Current Topics in Microbiology and Immunology

114

Springer-Verlag
Berlin Heidelberg New York Tokyo

With 39 Figures

ISBN-13:978-3-642-70229-7 e-ISBN-13:978-3-642-70227-3
DOI: 10.1007/978-3-642-70227-3

Library of Congress Catalog Card Number 15-12910

2123/3130-543210

Table of Contents

Indexed in Current Contents

List of Contributors

R.K. AKKINA, Jonsson Comprehensive Cancer Center and Department of Microbiology and Immunology, UCLA School of Medicine, Los Angeles, CA 90024, USA

D.H.L. BISHOP, Department of Microbiology, University of Alabama in Birmingham, Birmingham, AL 35294, USA

T.M. CHAMBERS, Jonsson Comprehensive Cancer Center and Department of Microbiology and Immunology, UCLA School of Medicine, Los Angeles, CA 90024, USA

P.M. COLMAN, CSIRO, Division of Protein Chemistry, 343 Royal Parade, Parkville 3052, Victoria, Australia

R.W. COMPANS, Department of Microbiology, University of Alabama in Birmingham, Birmingham, AL 35294, USA

D.P. NAYAK, Jonsson Comprehensive Cancer Center and Department of Microbiology and Immunology, UCLA School of Medicine, Los Angeles, CA 90024, USA

R. SCHNABEL, Max-Planck-Institut für Biochemie, D-8033 Martinsried

K.O. STETTER, Universität Regensburg Lehrstuhl für Mikrobiologie, D-8400 Regensburg

H.-P. VOSBERG, Max-Planck-Institut für Medizinische Forschung, Abteilung Molekulare Biologie, Jahnstraße 29, D-6900 Heidelberg

C.W. WARD, CSIRO, Division of Protein Chemistry, 343 Royal Parade, Parkville 3052, Victoria, Australia

W. ZILLIG, Max-Planck-Institut für Biochemie, D-8033 Martinsried

Archaebacteria and the Origin of the Eukaryotic Cytoplasm

W. ZILLIG[1], R. SCHNABEL[1], and K.O. STETTER[2]

1 Introduction

The recognition of a third urkingdom of life, the archaebacteria (WOESE and FOX 1977; WOESE et al. 1978), resulting in the division of the prokaryotic domain (FOX et al. 1980), arose from the perception of the large phylogenetic distance between methanogens and *Escherichia coli,* which placed these "bacteria" as far from each other as is, for example, the bacterium *E. coli* from the eukaryote yeast. Using the unique nature of archaebacterial lipids (KATES 1972; LANGWORTHY et al. 1974; DE ROSA et al. 1977; LANGWORTHY 1977; KATES 1978; TORNABENE 1979) as a guideline, but comparative cataloging of 16S rRNA (FOX et al. 1977) as the measure, it has been possible to expand the archaebacterial urkingdom to three orders of methanogens (BALCH et al. 1979), including the related extreme halophiles, the isolated genus *Thermoplasma* (DARLAND et al. 1970), and *Sulfolobus* (BROCK et al. 1972), the last two of which are often set apart from the larger part of the kingdom as the "thermoacidophiles." On the basis of the S_{AB} (identity coefficient $S_{AB}=2N_{AB}/(N_A+N_B)$ where N_A, N_B and N_{AB} are the total number of sequences and the identical sequences in two sequence catalogs to be compared) value of 0.17, they appear to be as isolated from each other as from the rest of the kingdom, however.

1 Max-Planck-Institut für Biochemie, D-8033 Martinsried
2 Universität Regensburg, Lehrstuhl für Mikrobiologie, D-8400 Regensburg

Current Topics in Microbiology and Immunology, Vol. 114

2 The Sulfur-Dependent Branch of Archaebacteria

The discovery of three families of extremely thermophilic, anaerobic, sulfur reducing archaebacteria, the *Thermoproteaceae* comprising the genera *Thermoproteus* (ZILLIG et al. 1981a) and *Thermofilum* (ZILLIG et al. 1983a), the *Desulfurococcaceae* (ZILLIG et al. 1982a) and the *Thermococcaceae* (ZILLIG et al. 1983b), combined in the novel order *Themoproteales* (ZILLIG et al. 1981a; 1981b) (Fig. 1), removed *Sulfolobus* from its isolated position. The Thermoproteales and Sulfolobales form a strong second branch of the kingdom, leaving *Thermoplasma* in an isolated or intermediate situation (ZILLIG et al. 1981; TU et al. 1982; ZILLIG et al. 1982b; WOESE et al. 1984) (Fig. 2). Since several isolates of the novel order grow well at neutral pH, we propose to term this group "sulfur-dependent" rather than "thermoacidophilic" archaebacteria (FISCHER et al. 1983), although *Sulfolobus* and the Thermoproteales utilize sulfur in opposite ways.

3 DNA-Dependent RNA Polymerases from Archaebacteria

The component patterns of DNA-dependent RNA polymerases of archaebacteria are clearly distinct from those of eubacteria (ZILLIG et al. 1982c; PRANGISHVILLI et al. 1982), but resemble those of eukaryotic nuclear RNA polymerases, especially RNA polymerase A(I) (Fig. 3). By challenging these patterns with antibodies raised against single components of the RNA polymerases from *Sulfolobus* and *Methanobacterium,* it has been possible to establish the homologies of the larger components, including component D and, within the sulfur-dependent branch, even component E (R. SCHNABEL et al. 1983). The two branches, one comprising methanogens and extreme halophiles, the other the *Thermoproteales* and *Sulfolobus,* are distinguished by different types of RNA polymerases, the AB′B″CD- and the BACD type, respectively. By this feature, *Thermoplasma* belongs to the latter rather than the first branch of the kingdom.

Fig. 1 a–q. Electron micrographs of Thermoproteales. **a–c** enrichment cultures: *a* one intact and one empty (filamentous) cell of *Thermofilum* with attached (spherical) *Desulfurococcus* cell; **b** two cells of *Thermofilum,* one with "preformed breaking zone" and one (thicker) of *Thermoproteus;* **c** *Thermofilum, Thermoproteus,* and *Desulfurococcus,* the last with attached slime filaments; **d–f** *Thermofilum pendens*: **d** end of cell with terminal pili; **e** three cells, one in "golf club" state; **f** thin section, exhibiting subunit cell wall; **g–i** *Thermococcus celer:* **g** flagellated cell; **h, i** thin section of constricted dividing cells; **j–m** *Thermoproteus tenax:* **j** one cell with pili, exhibiting regular structure of envelope; **k** thin section, showing subunit envelope; **l** golf club; **m** budding cell; **n–q** *Desulfurococcus*: **n** flagellated cells *of D. mobilis;* **o** *Desulfurococcus species* with slime filaments; **p** *D. mucosus* with solid slime layer; **q** thin section of *D. mobilis* showing how flagellae are attached. All specimens except the thin sections and **g,** which was negatively stained, were rotary-shadowed with Pt. Thin sections were doubly contrasted. *Bars* represent 1 μm each. All the electron micrographs were made by Davorin Janekovic

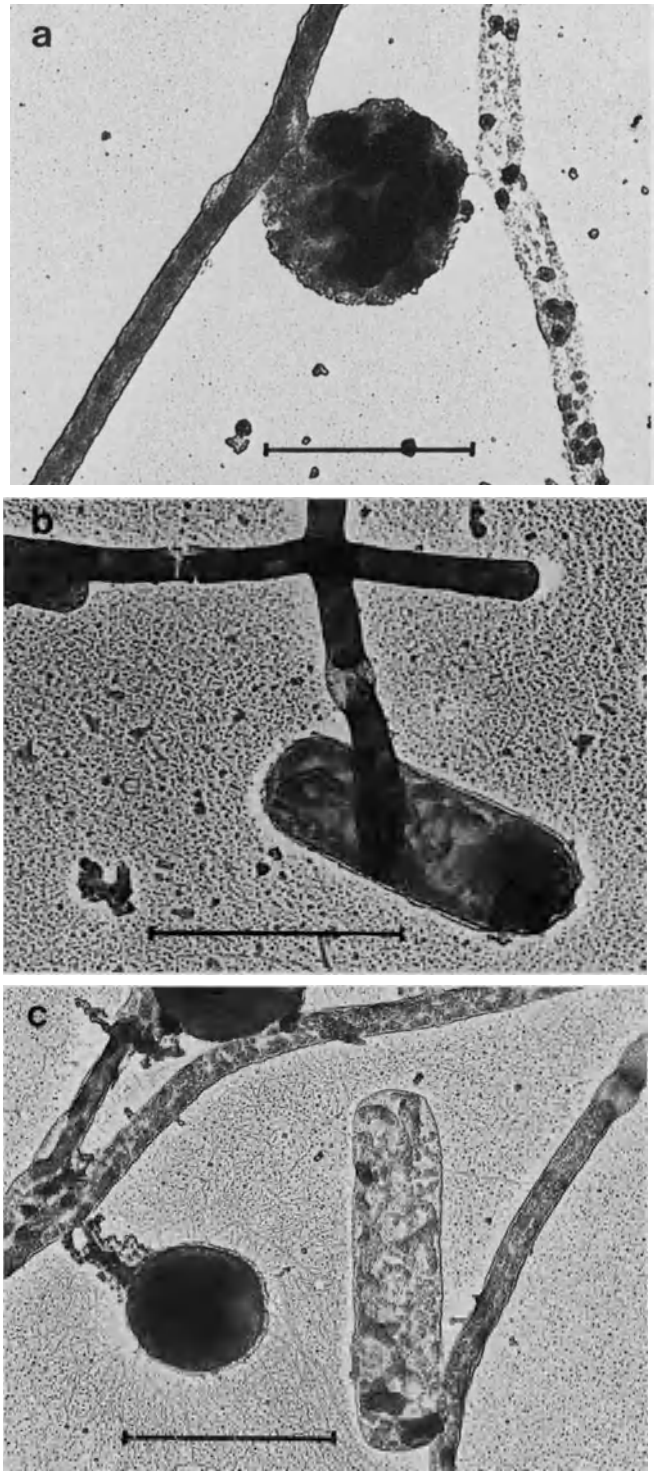
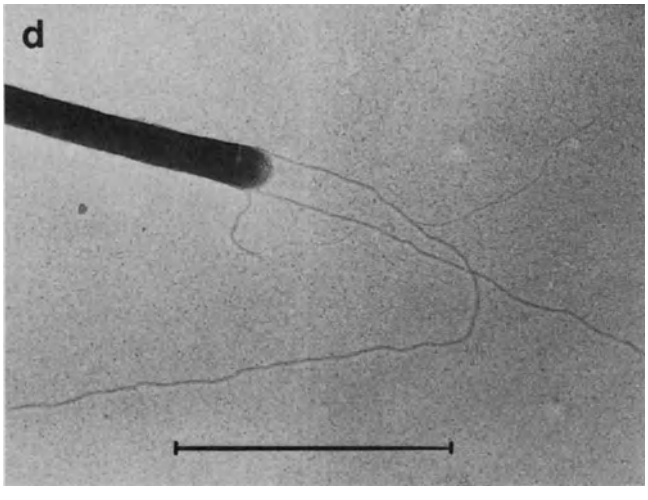
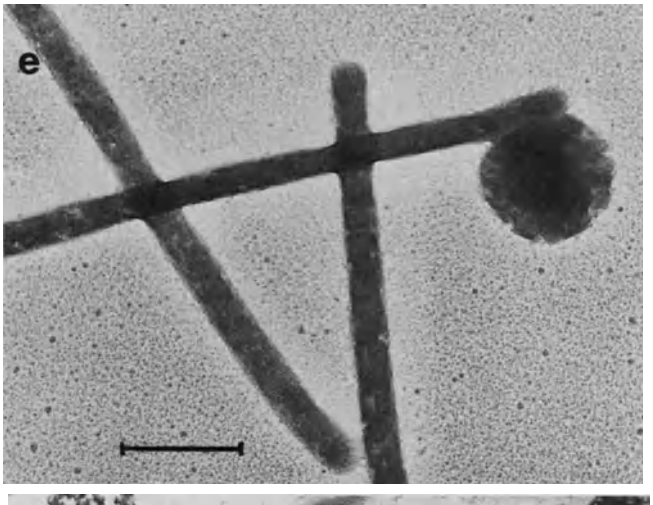


Fig. 1a–c





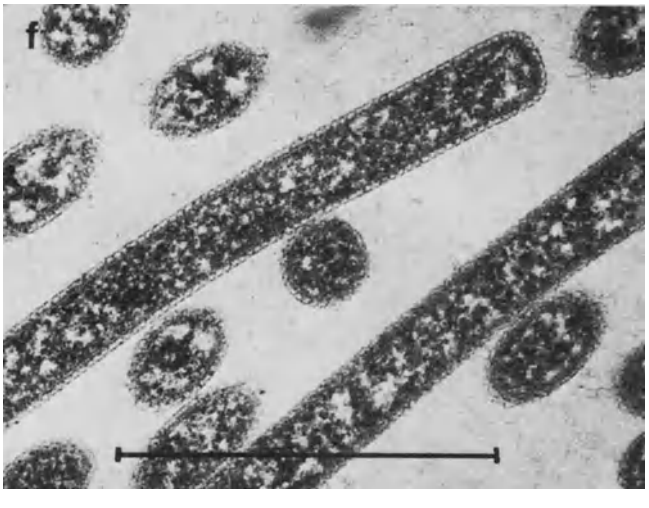


Fig. 1d–f

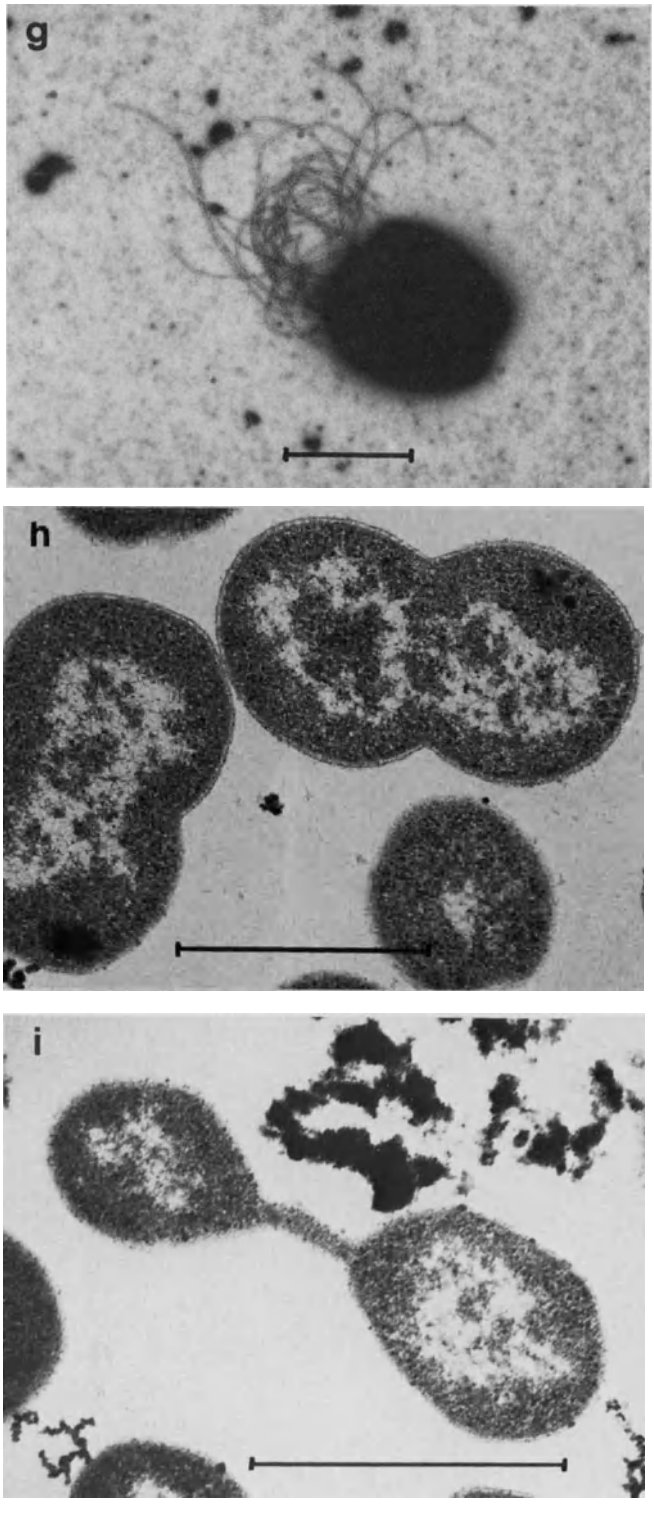


Fig. 1g–i

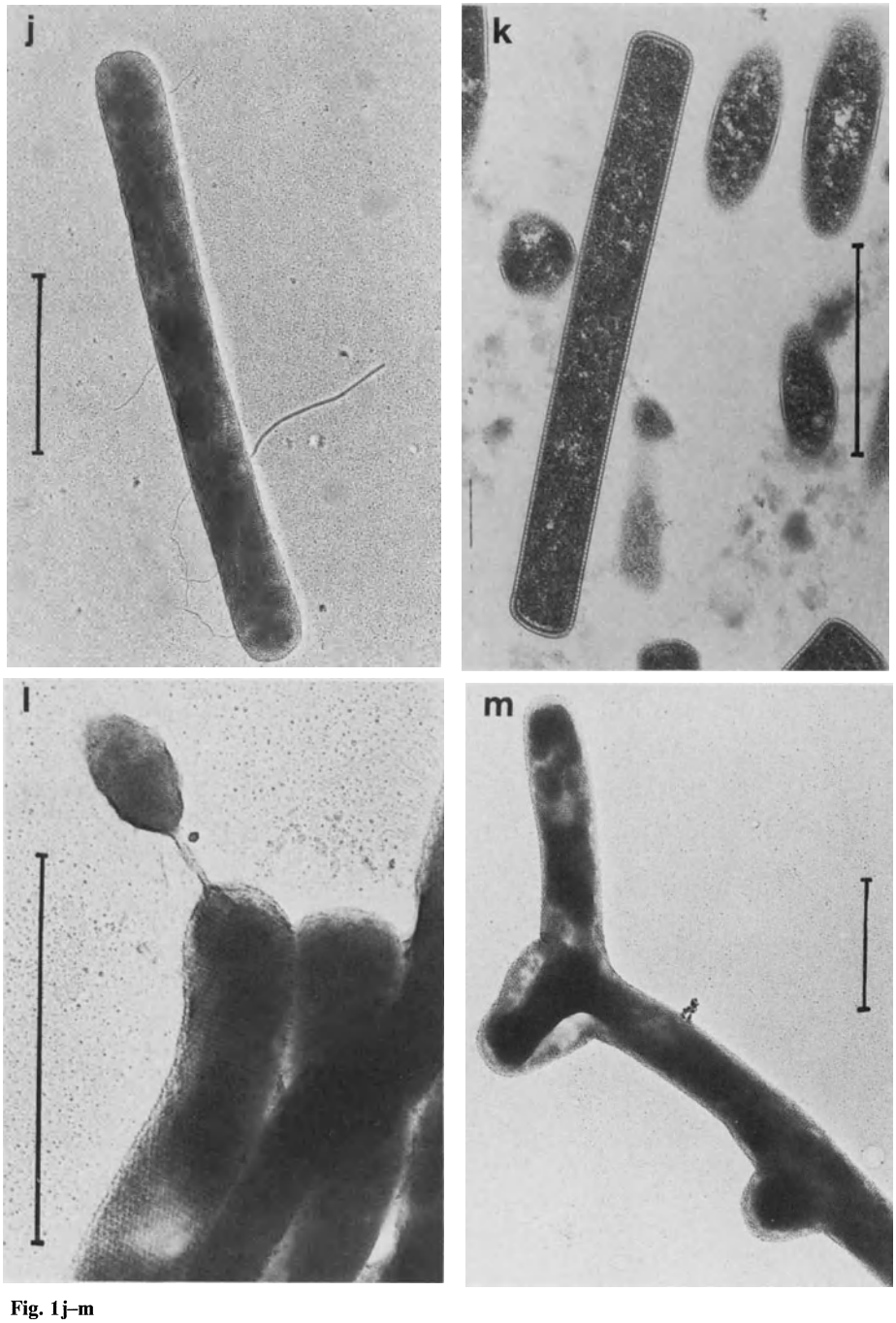


Fig. 1j–m

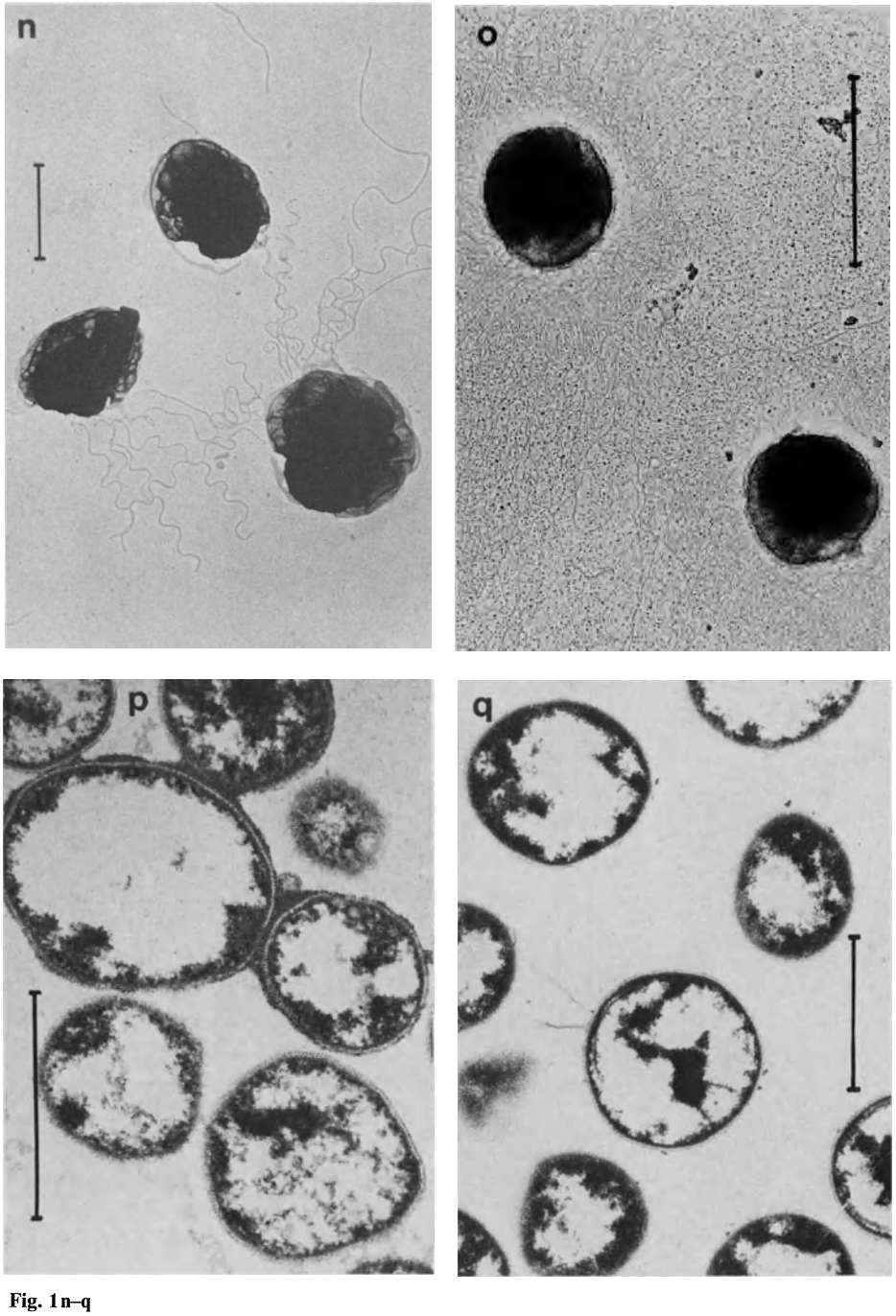


Fig. 1n–q

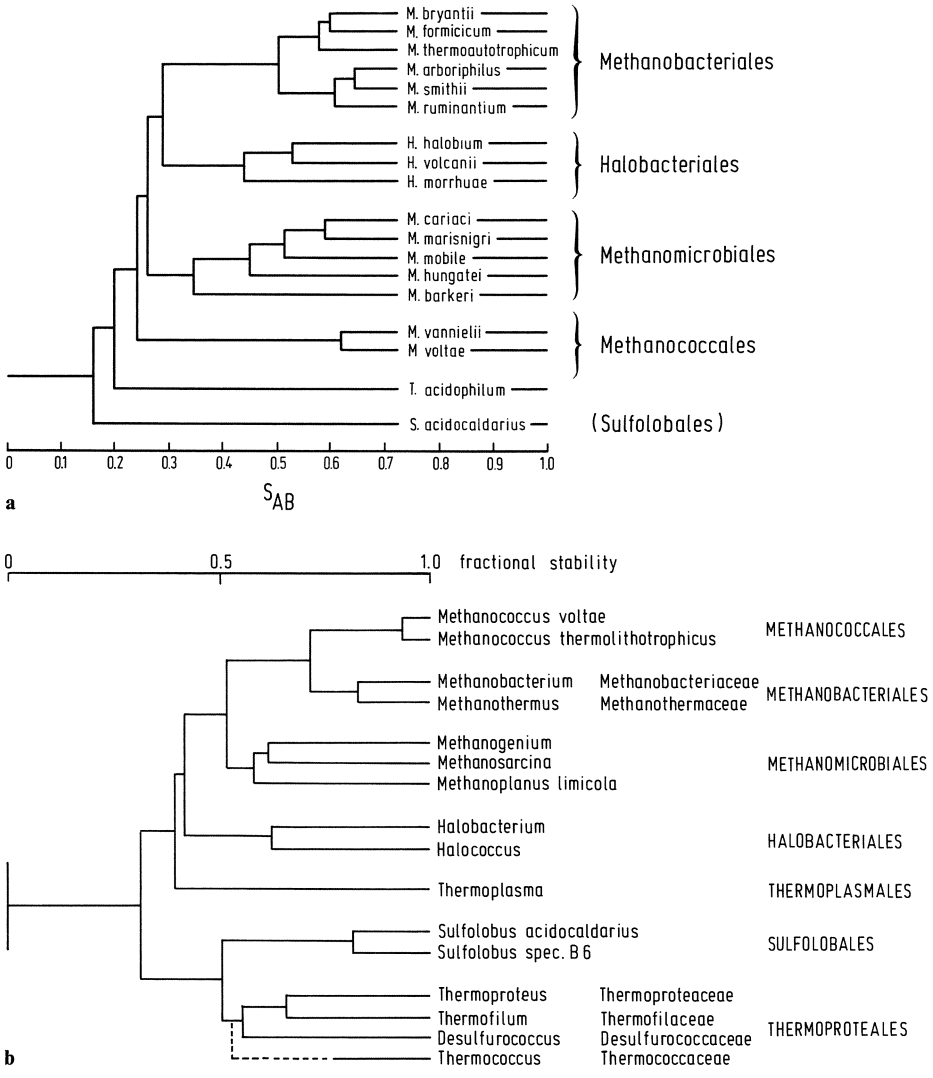


Fig. 2 a, b. Phylogenetic trees of the archaebacteria. **a** As derived by Fox et al. (1980) from S_{AB} values (16S rRNA sequence homology), not including the Thermoproteales and showing the large phylogenetic distance between *Sulfolobus* and the methanogens (including the halophiles); but also between *Sulfolobus* and *Thermoplasma*. **b** As derived by Tu et al. (1982) from 16S rRNA–DNA cross-hybridization data, including the *Thermoproteales*, showing the two major branches and the intermediate position of *Thermoplasma*. The association of *Thermoplasma* with the methanogens is not significant

By challenging component patterns or spots of DNA-dependent RNA polymerases from eubacteria and archaebacteria with antibodies raised against nuclear RNA polymerases from yeast and its components, it has been possible to substantiate the close similarity of eukaryotic and archaebacterial RNA polymerases (Huet et al. 1983) (Fig. 3). Accordingly, transcription by archaebacter-

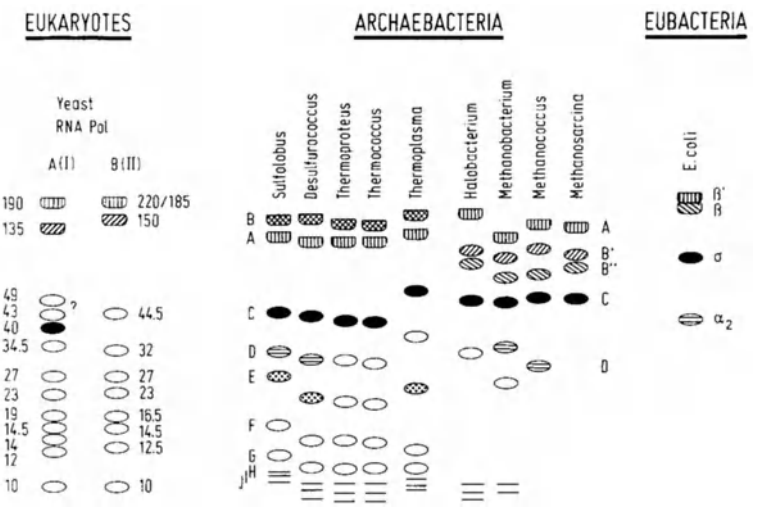


Fig. 3. Immunochemical homologies of components of DNA-dependent RNA polymerases of archaebacteria, eubacteria, and nuclei of eukaryotes (SCHNABEL et al. 1983; HUET et al. 1983), as indicated by equal designs in a schematic representation of SDS polyacrylamide gel electrophoresis patterns (for original patterns see ZILLIG et al. 1982)

ial RNA polymerases is not inhibited by rifampicin and streptolydigin. Like transcription by eukaryotic RNA polymerases A(I), it is stimulated significantly by the flavonolignane silybin (R. SCHNABEL et al. 1982).

The finding that archaebacterial and eukaryotic RNA polymerases are of the same type leads to the conclusion that both are derived from the same ancestor. The cross-reaction between homologous heavy components of archaebacterial and eukaryotic polymerases is stronger than that between homologous heavy components of the eukaryotic polymerases A(I) and B(II), indicating a larger phylogenetic distance between the latter two than between either of them and the archaebacterial polymerases. In the course of their separate evolution, the eukaryotic nuclear RNA polymerases have thus lost homologies with each other which have been conserved between one or the other and the archaebacterial enzymes.

The patterns of the components of DNA-dependent RNA polymerases from sulfur-dependent archaebacteria and *Thermoplasma* are strikingly similar to those from eukaryotic nuclear polymerases even in spacing. The main difference is the reversed order of the apparent molecular weights of the heavy components. The component patterns of polymerases from methanogens and extreme halophiles appear less related to those from eukaryotes, especially in the division of one of the heavy components, B, into two fragments, B and B′ and in a decreased complexity (loss of small components).

The distribution of antigenic determinants of the large B components of the sulfur-dependent archaebacteria between the B′ and B″ components of the methanogens and halophiles differs between the *Halobacteriales,* the *Methanomicrobiales* and the *Methanobacteriales* plus the *Methanococcales*. Some of the

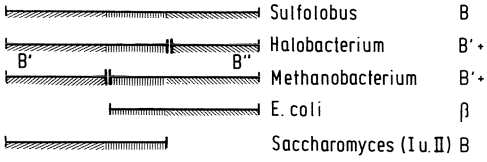


Fig. 4. Schematic representation of correspondence of *E. coli β* and *Saccharomyces* POL I B with B components of *Halobacterium, Methanobacterium* and *Sulfolobus*. Vertical lines connect corresponding structural elements i.e. antigenic sites. Horizontal lines do not give size of peptides but alignments of determinants

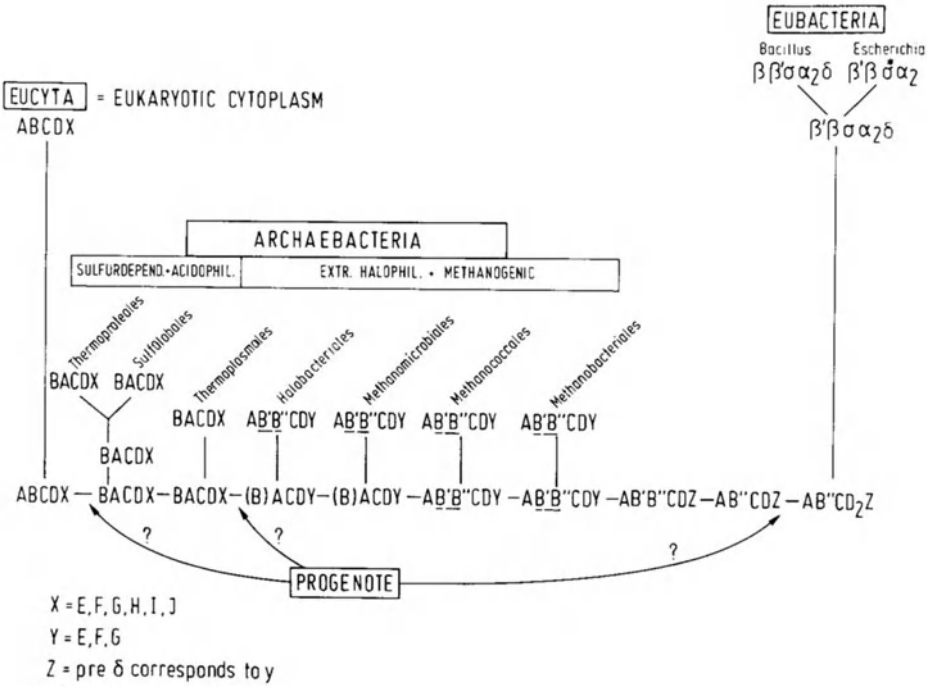


Fig. 5. Flow diagram of component patterns of different DNA dependent RNA polymerase types, corresponding to an unrooted tree. Most probable entering points for conversion to phylogenetic trees indicated by connecting lines from "progenote". The ABCDX (urkaryotic) type could also be placed between the branching points of *Thermoplasmales* and *Halobacteriales*. The branching points of *Methanococcales* and *Methanobacteriales* could be in reversed order or these orders could have divided *after* branching off the lineage leading to the eubacteria

determinants of the large B component of the sulfur-dependent branch which are also present in *Methanobacterium* and *Methanococcus* B″, are found in *E. coli β*, but not in the eukaryotic nuclear RNA polymerase I (and II) B subunit. Others are found in the latter, but not in *E. coli β*. A third group is conserved in both (ZILLIG et al. in press, a). These relations are schematically visualized in Fig. 4 and indicated by the marking of the B components in Fig. 3.

These data plus the homology between the archaebacterial C components and *E. coli σ* and the probable correspondence of *σ* with A40 of the yeast RNA polymerase I, and the possible homology of *E. coli α* and the archaebac-

terial D components are the basis of the flow diagram presented in Fig. 5, which suggests a complete sequence of single step transitions between the RNA polymerase types of the urkaryotes, the archaebacteria and the eubacteria (ZILLIG et al. in press).

Few variations of this type transition sequence appear as alternative possibilities: (1) Instead of the transition from the ABCDX to the BACDX type on the entire left of the diagram a transformation from BACDX to ABCDX (or vice versa) could be assumed between the branching points of the *Thermoplasmales* and the *Halobacteriales*. This position of the "urkaryote" would be in line with the conservation of B′ determinants and the loss of B″ determinants of *Halobacterium* in the B component of the yeast RNA polymerase I and II. (2) The *Methanococcales* and the *Methanobacteriales* would have branched off in reverse order or (3) they could have separated after branching off the lineage ending in the eubacteria.

The scheme corresponds to a complicated unrooted tree which derives the eubacteria from the *Methanococcales* or *Methanobacteriales* or their immediate ancestor (or the other way round), places the sulfur-dependent archaebacteria close to the urkaryote and *Thermoplasma* besides the *Halobacteriales*. It is the first consistent proposal of the relation of the two modern kingdoms to the two branches of the archaebacteria. As indicated in Fig. 5, this tree can be rooted in a number of ways, of which we prefer the rooting between *Thermoplasma* and the halophiles, since it best explains the finding that the archaebacteria share different determinants with the eukaryotes and the eubacteria, besides some common determinants. This would be expected if they would be witnesses of the ancestral layer (ZILLIG et al. in press) and is thus suggestive evidence for the archaic nature of the archaebacteria.

4 "Eukaryotic" Features of Archaebacteria

While several characteristics of archaebacteria, such as the occurrence of isopranyl ether instead of ester lipids and the sequences of the large ribosomal RNAs tend to set this urkingdom apart, a surprisingly high and apparently increasing number of features appear significantly "eukaryotic":

1. Sequences of ribosomal A proteins from archaebacteria resemble the sequence of yeast A protein rather than protein L_7, L_{12} of *E. coli* (MATHESON and YAGUCHI 1982).
2. Elongation factors EF II from 18 archaebacteria share a characteristic property of eukaryotic EF IIs in contrast to the homologous eubacterial EFGs: They are subject to ADP ribosylation by diphtheria toxin, revealing the conservation of the target sequence for this reaction in eukaryotes and archaebacteria (KESSEL and KLINK 1982).
3. The terminal base pair of the aminoacyl stem of initiator tRNA is AU, as in eukaryotes (KUCHINO et al. 1982). Methionyl-initiator tRNA is not formylated. The CCA termini of tRNA are not encoded (KAINE et al. 1983). Some tRNA genes contain introns (KAINE et al. 1983).

4. Archaebacteria possess aphidicolin sensitive replicating DNA polymerases, as eukaryotes do (FORTERRE et al. 1984; M. NAKAYAMA and M. KOHIYAMA, personal communication; R. SCHINZEL and K.J. BURGER, personal communication; H.P. ZABEL, J. WINTER, H. FISCHER, E. HOLLER, personal communication; D. PRANGISHVILLI and W. ZILLIG 1984).
5. Archaebacteria are normally insensitive to many antibiotics characteristically inhibiting eubacteria, e.g. streptomycin (SCHMID et al. 1982), rifamycin (ZILLIG et al. 1982c; PRANGISHVILLI et al. 1982), and vancomycin.

On the other hand, archaebacteria are clearly prokaryotes. (This term describes primitiveness of organisation rather than defining a phylogenetic entity.) However, some other properties of true bacteria, e.g. anti-Shine-Dalgarno sequences in the terminus of the 16S rRNA of methanogens (STEITZ 1978), and characteristic restriction enzymes (MCCONNELL et al. 1978; P. MCWILLIAM, unpublished work showing the occurrence of restriction enzymes in *Sulfolobus* and *Thermococcus* 1983), have also been observed.

5 Phylogenetic Depth of Archaebacteria

Several features of archaebacteria, however, are neither unique nor "eukaryotic" or "eubacterial," but vary in different phyla over the whole range separating the eubacteria and the eukaryotic cytoplasm.

The composition of the DNA-dependent RNA polymerases, which is a particularly significant feature of this type, has been discussed above. Another striking example is the secondary structure of 5S rRNA (FOX et al. 1982; HORI et al. 1982). Some representatives of the methanogenic halophilic branch possess 5S rRNAs with secondary structures closely resembling those characteristic of eubacteria. Others are intermediate. The structure of the 5S rRNA of *Thermoplasma* appears eukaryotic, and that of *Sulfolobus* represents an exaggerated eukaryotic type.

Similarly, though not as clearly, the sequence of the initiator tRNA of *Halococcus* resembles those of eubacteria more than those of eukaryotes, that of *Thermoplasma* is intermediary, and that of *Sulfolobus* is closer to those of eukaryotes (KUCHINO et al. 1982).

The shape of the small ribosomal subunit of archaebacteria is intermediate between those of eubacteria and the eukaryotic cytoplasm in exhibiting the "archaebacterial duckbill," which is absent in small subunits of eubacteria (LAKE et al. 1982), whereas small subunits of the eukaryotic cytoplasm show the duckbill plus an additional uniquely eukaryotic lobe. The shapes of the large subunits of archaebacteria range from eubacterial in the extreme halophiles to almost eukaryotic in *Thermoproteus,* thus spanning the gap between eubacteria and the eukaryotic cytoplasm (HENDERSON et al. 1984; CLARK et al. 1984). These observations are paralleled by the complexity of the protein composition of the ribosomal subunits, which is again higher in *Sulfolobus* than in *Halobacterium* (SCHMID et al. 1982). The extent of modification of the rRNAs in *Thermo-*

proteus and *Sulfolobus* is five times higher than in *E. coli* or as high as in eukaryotes (WOESE et al. 1984).

Archaebacteria have different types of cell walls e.g. glycoproteid or protein subunit envelopes, like *Sulfolobus* (WEISS 1974), the *Thermoproteales* (KÖNIG, personal communication) and *Halobacterium* (MESCHER and STROMINGER 1975; WIELAND et al. 1982), polysaccharide sacculi as *Halococcus* (SCHLEIFER et al. 1982), or even lack a cell wall, like Thermoplasma (DARLAND et al. 1970). They never have mureine sacculi. The gross resemblance of the pseudomureine wall of *Methanobacterium* (KÖNIG et al. 1982) and the mureine sacculus of eubacteria has been interpreted to have resulted from independent evolution rather than common origin (KANDLER 1982).

Archaebacteria of the methanogenic halophilic branch divide equally by septum formation, as the eubacteria do. In contrast, members of the sulfur-dependent branch divide by budding, e.g., *Thermoproteus* (ZILLIG et al. 1981a), *Desulfurococcus* (ZILLIG et al. 1982a); by constriction, e.g., *Thermococcus* (ZILLIG et al. 1983b); by branching, e.g., *Thermoproteus,* always without formation of septa; or by other yet unknown modes of "legitimate fragmentation," e.g., "golf club" formation in *Thermoproteus* and *Thermofilum* (ZILLIG et al. 1981a, 1983a) and "prepared breakage" in *Thermofilum.* Some of these division mechanisms are encountered among eukaryotes.

Several bacteriophages of *Halobacterium,* e.g., phage ϕ H (H. SCHNABEL et al. 1982a, b), resemble phages of eubacteria both in structural and functional aspects. A DNA-containing particle of *Sulfolobus spec.* B6, which crystallizes within the host cell in hexagonal dense array, resembles adenovirus in this respect (W. ZILLIG, P. MCWILLIAM, and S. YEATS, unpublished work on virus like particles in *Sulfolobus spec.* B6). Three DNA containing viruses multiplied upon sulfur starvation, in cultures of *Thermoproteus tenax,* strain Kra 1, show features of plant and animal viruses: rod shape, helical array of proteins, and an envelope (JANEKOVIC et al. 1983). A virus-like particle of *Sulfolobus solfataricus,* strain B12, exhibits a unique structure and has a lipid envelope (MARTIN et al. 1984).

Generally, in these features testifying to the phylogenetic depth and thus probably also to the age of this group, the *Thermoproteales* and *Sulfolobus* appear closer to the eukaryotic cytoplasm and the methanogens and halophiles less distant from the eubacteria.

6 Early Biotic Evolution

It thus appears that the ancestors of the eukaryotic cytoplasm will have resembled recent sulfur-dependent archaebacteria. The starting point of the independent evolution of the eubacteria would, on the other hand, have been close to that of the methanogens, particularly the *Methanococcales* and *Methanobacteriales.* However, even the methanogenic and halophilic archaebacteria appear to be more closely related to the eukaryotes than to the eubacteria in most features. The roots of the lineages of the eukaryotes and the eubacteria are probably

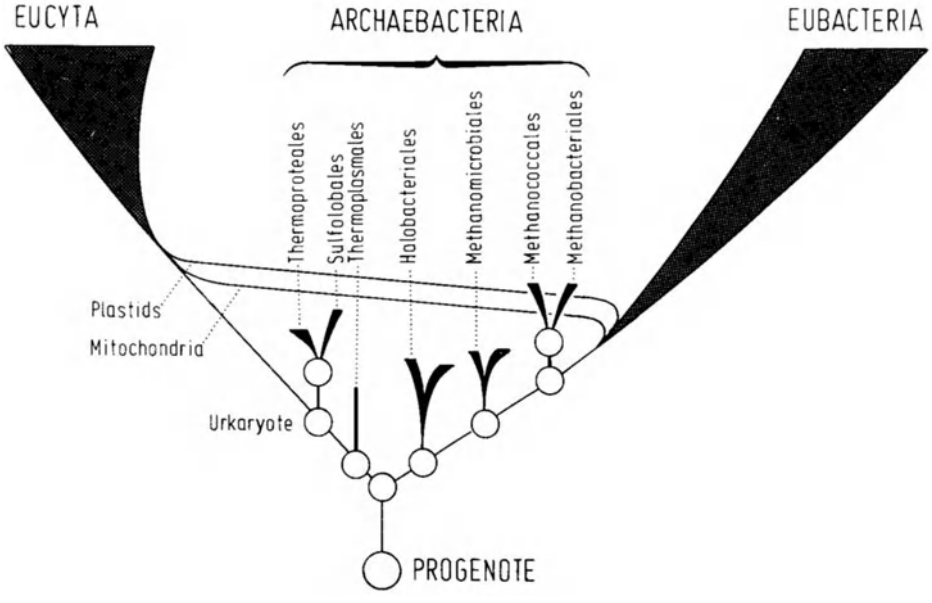


Fig. 6. Sketch of a probable evolutionary tree showing the relation of the three urkingdoms, especially the rise of the modern kingdoms from the primeval layer of which the present day archaebacteria are witnesses. The length of the branches crudely indicates evolutionary distance, the width the number of species within a group

not much farther apart than those of the two main branches of the recent archaebacteria of which the extreme halophiles plus methanogens appear as a polyphyletic group (Fig. 5). But the two branches of the archaebacteria show such a strong resemblance that they have been combined into one urkingdom. This apparent dilemma appears in a different light if one considers that the 16S rRNA of *Halobacterium volcanii* shows more sequence identity with those of *E. coli* and yeast than these two do with each other (GUPTA et al. 1983). This and the above-mentioned finding that the heavy components of different classes of eukaryotic nuclear RNA polymerases have less homology with each other than either of these with those from archaebacteria indicate that the archaebacteria are primitive in the sense that they have conserved primordial characteristics, possibly because the extreme conditions of their environment have severely restricted their evolution.

The roots of their "short" branches may be as far apart from each other as from the roots of the "long lineages" of the two other urkingdoms. It thus becomes somewhat arbitrary to combine them into one urkingdom, because their similarities appear to be due to their primordial character rather than to a common origin (Fig. 6). The apparently primitive mechanisms of cell division of the sulfur-dependent archaebacteria and their primitive anaerobic mode of energy conservation by sulfur respiration or H_2S formation, well suited to conditions on earth at the time when simple life forms were developing, and the genuine extreme thermostability of their "proteins", seem to support this line of argument.

Another cue comes from enzymology. Like eukaryotes and eubacteria, *Sulfolobus* and the *Thermoproteales* are able to synthesize glycogen. Whereas the glucosyl transferases of bacteria, plants, and animals utilize either ADP glucose or uridine diphosphate (UDP) glucose for this purpose, that of *Sulfolobus* uses both with almost equally low turnover numbers and high KMs, and is thus not only an inefficient but also a rather unspecific, possibly primitive enzyme (KÖNIG et al. 1982).

The limitations on functional perfection imposed by the structural requirements for protein stability at high temperatures or high salinity could well have retarded and even prevented evolution. Without a knowledge of the rules governing the stability of proteins under such conditions, one cannot be sure to what extent this occurred, however. Yet the molecular evidence indicates that the eubacteria and, to a lesser extent, the eukaryotes have evolved further than the archaebacteria. The recent eubacteria appear almost perfectly adapted to their biotopes. The eukaryotes are visibly in the process of rapid evolution. The archaebacteria seem to be almost as terminally adapted to their niches as the eubacteria, though to a much lower degree of perfection, probably due to the limitations discussed above.

The existence of fatty acid ester lipids of the same type in both the eubacteria and the eukaryotes, in contrast to the isopranyl ether lipids of the archaebacteria, could be taken as a serious argument against this hypothetical view of branching in early biotic evolution. Fatty acids do, however, exist in archaebacteria (see e.g., ZILLIG et al. 1981 a). One must therefore not necessarily assume that fatty acid ester lipids evolved twice independently: it can be argued that the isopranyl ether lipids represent the older type, which was abandoned in the lineages leading to the eubacteria and the eukaryotes.

7 Conclusions

Considering all these arguments, it appears that the archaebacteria reflect the level of biotic evolution which had been reached by the time of the separation of the three lineages. The early eubacteria are derived from ancestors which were considerably closer to the early methanogens than to the ancestors of the sulfur-dependent archaebacteria. The acquisition of their characteristic features should have occurred as the result of rapid evolution before this lineage split up into the multitude of forms known to date. The early methanogens, including the ancestors of the extreme halophiles, started their own line of descent some time later, followed by the ancestors of *Thermoplasma* and then by those of the sulfur utilizers from which the urkaryote (WOESE and FOX 1977) arose (Fig. 6).

With respect to a single feature, e.g. the shape of the ribosome (LAKE et al. 1982; HENDERSON et al. 1984; LAKE et al. 1984), other phylogenetic trees appear formally possible. But the comparative analysis of many independent sets of homologous structures, in a particularly lucid way those of RNA polymerase (ZILLIG et al., in press), and of 5S rRNA (HORI et al. 1982; FOX et al. 1982),

strongly supports the given tree, the exception being the lipids, as already discussed and some data on the interaction of ribosomes from the three kingdoms with different antibotics (BÖCK and KANDLER, in press). Both groups of the archaebacteria appear to be very close to the line connecting the bacteria and the eukaryotes, as expected for true intermediary forms. This underlines their primitive, i.e., witness characters.

This sketch does not account for possible horizontal gene transfer, which would counteract the formation of lineages. But evidence for this has only recently begun to accumulate. Even if occurring with significant frequency, it would probably rather retard than qualitatively direct evolution except in special instances.

Note Added in Proof

The discrepancy between the modes of autotrophic existence of the related orders *Thermoproteales* and *Sulfolobales* has been bridged by the discovery of *Sulfolobus* species which are able to conserve energy by sulfur reduction, like *Thermoproteus*, as well as by sulfur oxidation, as previously considered typical for *Sulfolobus* species (SEGERER et al. in press; ZILLIG et al. in press, b).

References

Balch WE, Fox CE, Magrum LJ, Woese CR, Wolfe RS (1979) Methanogens: reevaluation of a unique biological group. Microbiol Rev 43:260–296

Brock TD, Brock KM, Belley RT, Weiss RL (1972) Sulfolobus: a new genus of sulfur oxidizing bacteria living at low pH and high temperature. Arch Microbiol 84:54–68

Darland G, Brock TD, Samsonoff W, Conti SF (1970) A thermophilic acidophilic mycoplasm isolated from a CoA refuse pile. Science 170:1416–1418

De Rosa M, De Rosa S, Gambacorta A, Minale L, Bullock JD (1977) Chemical structure of the ether lipids of thermophilic acidophilic archaebacteria of the *Caldariella* group. Phytochemistry 19:249–254

Fischer F, Zillig W, Stetter KO, Schreiber G (1983) Chemolithoautotrophic metabolism of anaerobic extremely thermophilic archaebacteria. Nature 301:511–513

Fox GE, Pechmann KJ, Woese CR (1977) Comparative cataloging of 16S ribosomal RNA: molecular approach to procaryotic systematics. Int J System Bact 27:44–57

Fox GE, Stackebrandt E, Hespell RB, Gibson J, Maniloff J, Dyer TA, Wolfe RS, Balch WE, Tanner RS, Magrum LJ, Zablen LB, Blakemore R, Gupta R, Bonen L, Lewis BJ, Stahl DA, Luehrsen KR, Chen KN, Woese CR (1980) The phylogeny of prokaryotes. Science 209:457–463

Fox GE, Luehrsen KR, Woese CR (1982) Archaebacterial 5S ribosomal RNA. Zbl Bakt Hyg, I. Abt. Orig. C 3:330–354

Gupta R, Lanter JM, Woese CR (1983) Sequence of the 16S ribosomal RNA from Halobacterium volcanii, an archaebacterium. Science 221:656–659

Henderson E, Oakes M, Clark MW, Lake JA, Matheson AT, Zillig W (1984) A new ribosome structure. Science 225:510–512

Hori H, Itoh T, Osawa S (1982) The phylogenic structure of the metabacteria. Zbl Bakt Hyg, I. Abt. Orig. C 3:18–30

Huet J, Schnabel R, Sentenac A, Zillig W (1983) Archaebacteria and eukaryotes possess DNA-dependent RNA polymerases of a common type. EMBO J 2:1291–1294

Janekovic D, Wunderl S, Holz I, Zillig W, Gierl A, Neumann H (1983) TTV1, TTV2 and TTV3

a family of viruses of the extremely thermophilic, anaerobic, sulfur reducing archaebacterium Thermoproteus tenax. Mol Gen Genet

Kaine BP, Gupta R, Woese CR (1983) Putative introns in tRNA genes of prokaryotes. PNAS 80:3309–3312

Kandler O (1982) Cell wall structures and their phylogenetic implications. Zbl Bakt Hyg, I. Abt. Orig. C 3:149–160

Kates M (1972) Ether-linked lipids in extremely halophilic bacteria. In: Snyder F (ed) Ether lipids: Chemistry and biology. Academic, New York, pp 351–398

Kates M (1978) The phytanyl ether-linked polar lipids and isoprenoid neutral lipids of extremely halophilic bacteria. Progr Chem Fats Other Lipids 15:301–342

Kessel M, Klink F (1982) Identification and comparison of eighteen archaebacteria by means of the diphtheria toxin raction. Zbl Bakt Hyg, I. Abt. Orig. C 3:140–148

König H, Skorko R, Zillig W, Reiter WD (1982) Glycogen in thermoacidophilic archaebacteria of the genera *Sulfolobus, Thermoproteus, Desulfurococcus* and *Thermococcus*. Arch Microbiol 132:297–303

Kuchino Y, Hiara M, Yabusaki Y, Nishimura S (1982) Initiator tRNAs from archaebacteria show common unique sequence characteristics. Nature 298:684–685

Lake JA, Henderson E, Clark MW, Matheson AT (1982) Mapping evolution with ribosome structure: Intralineage constancy and interlineage variation. Proc Natl Acad Sci USA 79:5948–5952

Lake JA, Henderson E, Oakes M, Clark MW (1984) Eocytes: A new ribosome structure indicates a kingdom with a close relationship to eukaryotes. Proc Natl Acad Sci, USA 81:3786–3790

Langworthy TA (1977) Long-chain diglycerol tetraethers from *Thermoplasma acidophilum*. Biochim Biophys Acta 487:37–50

Langworthy TA, Mayberry WR, Smith PF (1974) Long-chain glycerol diether and polyol dialkyl glycerol triether lipids of *Sulfolobus acidocaldarius*. J Bact 119:106–116

Martin A, Yeats S, Janekovic D, Reiter WD, Aicher W, Zillig W (1984) SAV 1, a temperate u.v.-inducible DNA virus-like particle from the archaebacterium *Sulfolobus acidocaldarius* isolate B12. The EMBO J. 3:2165–2168

Matheson AT, Yaguchi M (1982) The evolution of the archaebacterial ribosome. Zbl Bakt Hyg, I Abt Orig C 3:192–199

McConnell DJ, Searcy DG, Sutcliffe JG (1978) A restriction enzyme Tha I from the thermophilic mycoplasma Thermoplasma acidophilum. Nucleic Acids Res 5:1729–1739

Mescher MF, Strominger JL (1975) Purification and characterization of a prokaryotic glycoprotein from the cell envelope of *Halobacterium salinarium*. J Biol Chem 251:2005–2014

Prangishvilli D, Zillig W, Gierl A, Biesert L, Holz I (1982) DNA-dependent RNA polymerases of thermoacidophilic archaebacteria. Eur J Biochem 122:471–477

Prangishvilli D, Zillig W (1984) DNA-dependent DNA-polymerases of thermoacidophilic archaebacterium *Sulfolobus acidocaldarius*. Poster presented at the FEMS Symposium "Evolution of Prokaryotes", held in Munich, September 16–18, 1984

Schleifer KH, Steber J, Mayer H (1982) Chemical composition and structure of the cell wall of *Halococcus morrhuae*. Zbl Bakt Hyg, I Abt Orig C3:171–178

Schmid G, Pecher T, Böck A (1982) Properties of the translational apparatus of archaebacteria. Zbl Bakt Hyg, I Abt Orig C3:209–217

Schnabel H, Zillig W, Pfäffle M, Schnabel R, Michel H, Delius H (1982a) *Halobacterium halobium* phage ϕ H. EMBO J 1:87–92

Schnabel H, Schramm E, Schnabel R, Zillig W (1982b) Structural variability in the genome of phage ϕ H of *Halobacterium halobium*. Mol Gen Genet 188:370–377

Schnabel R, Sonnenbichler J, Zillig W (1982) Stimulation by silybin, a eukaryotic feature of archaebacterial RNA polymerases (1982) FEBS L 150:400–402

Schnabel R, Thomm M, Gerardy-Schahn R, Zillig W, Stetter KO, Huet J (1983) Structural homology between different archaebacterial DNA-dependent RNA polymerases analyzed by immunological comparison of their components. EMBO J 2:751–755

Segerer A, Stetter KO, Klink F (1985) Two contrary modes of autotrophy in the same archaebacterium. Nature, in press

Steitz JA (1978) Methanogenic bacteria. Nature 273:10

Tornabene TG and Langworthy TA (1979) Diphytanyl and dibiphytanyl glycerol ether lipids of methanogenic archaebacteria. Science 203:51–53

Tu J, Prangishvilli P, Huber H, Wildgruber G, Zillig W, Stetter KO (1982) Taxonomic relations between archaebacteria including 6 novel genera examined by cross hybridization of DNAs and 16S rRNAs. J Mol Evol 18:109–114

Weiss RL (1974) Subunit cell wall of *Sulfolobus acidocaldarius*. J Bacteriol 118:275–284

Wieland F, Lechner J, Sumper MC (1982) The cell wall glycoprotein of *halobacteria*: structural, functional and biosynthetic aspects. Zbl Bakt Hyg, I Abt Orig C 3:161–170

Woese CR, Fox GE (1977) Phylogenetic structure of the prokaryotic domain: The primary kingdoms. Proc natl Acad Sci USA 74:5088–5090

Woese CR, Magrum LJ, Fox GE (1978) Archaebacteria. J Mol Evol 11:245–252

Woese CR, Gupta R, Hahn CM, Zillig W, Tu J (1984) The phylogenetic relationships of three sulfur dependent archaebacteria. System Appl Microbiol 5:97–105

Zillig W, Stetter KO, Schäfer W, Janekovic D, Wunderl S, Holz I, Palm P (1981a) Thermoproteales: a novel type of extremely thermoacidophilic anaerobic archaebacteria isolated from Icelandic solfataras. Zbl Bakt Hyg, I Abt Orig C 2:205–227

Zillig W, Tu J, Holz I (1981b) Thermoproteales – a third order of thermoacidophilic archaebacteria. Nature 293:85–86

Zillig W, Stetter KO, Prangishvilli D, Schäfer W, Wunderl S, Janekovic D, Holz I, Palm P (1982a) *Desulfurococcaceae,* the second family of the extremely thermophilic, anaerobic, sulfur-respiring *Thermoproteales*. Zbl Bakt Hyg, I Abt Orig C 3:304–317

Zillig W, Schnabel R, Tu J, Stetter KO (1982b) The phylogeny of archaebacteria, including novel anaerobic thermoacidophiles in the light of RNA polymerase structure. Naturwissenschaften 69:197–204

Zillig W, Stetter KO, Schnabel R, Madon J, Gierl A (1982c) Transcription in archaebacteria. Zbl Bakt Hyg, I Abt Orig C 3:218–227

Zillig W, Gierl A, Schreiber G, Wunderl S, Janekovic D, Stetter KO, Klenk HP (1983a) The archaebacterium *Thermofilum pendens* represents a novel genus of the thermophilic, anaerobic sulfur respiring *Thermoproteales*. System Appl Microbiol 4:79–87

Zillig W, Holz I, Janekovic D, Schäfer W, Reiter WD (1983b) The archaebacterium *Thermococcus celer* represents a novel genus within the thermophilic branch of the archaebacteria. System Appl Microbiol 4:88–94

Zillig W, Schnabel R, Stetter K, Thomm M, Gropp F, Reiter WD. The Evolution of the Transcription Apparatus. In "The Evolution of the Prokaryotes". K.H. Schleifer and E. Stackebrandt, eds., Academic Press, New York, in press

Zillig W, Yeats S, Holz, I, Böck A, Gropp F, Rettenberger M, Lutz S (1985) Plasmid-related anaerobic autotrophy of the novel archaebacterium *Sulfolobus ambivalens*. Nature, in press

DNA Topoisomerases: Enzymes That Control DNA Conformation

H.-P. VOSBERG

1 Introduction

It is common knowledge today that DNA molecules are not simple, rigid double helices, but molecules which have considerable conformational flexibility. This flexibility allows coiling of DNA into complex higher order structures. In most living cells or organelles, including many virus particles, these higher order structures occur on the basis of negative supercoiling or, as it is also called, negative superhelicity. Negative supercoiling is the conformational consequence of a winding deficit in DNA, i.e., the DNA under conditions of defined constraints forms in vivo or in vitro fewer helical rotations of one strand about the other than it assumes if left unconstrained.

Max-Planck-Institut für Medizinische Forschung, Abteilung Molekulare Biologie, Jahnstraße 29, D-6900 Heidelberg

Current Topics in Microbiology and Immunology, Vol. 114

Negative supercoiling in DNA was first identified by VINOGRAD and his co-workers as a property of closed circular DNA isolated from the animal virus polyoma (VINOGRAD et al. 1965). It was quickly found thereafter that closed circular DNAs from many other natural sources also have the conformation of a negative superhelix. Whereas many biochemical and biophysical properties of supercoiled DNA were successfully analyzed in a number of laboratories in the subsequent years, the biological significance of supercoiling was not immediately recognized.

This changed with two additional discoveries. The first, almost 15 years ago, was the detection of those enzymes which are now known as DNA topoisomerases. These enzymes catalyze conformational changes in DNA in vitro and in vivo. The second discovery was the independent finding that superhelicity is not a special feature of small DNA circles but is of far more general occurrence. This was initially shown for prokaryotic chromosomes (STONINGTON and PETTIJOHN 1971; WORCEL and BURGI 1972) and later also for the more complex eukaryotic chromosomes (COOK and BRAZELL 1975; BENYAJATI and WORCEL 1976; PINON and SALTS 1977).

DNA topoisomerases are functionally linked to DNA conformation by their ability to catalyze the concerted breaking and rejoining of DNA strands. This reaction is a prerequisite for the variation of topological and, hence, conformational properties of DNA. These changes affect the number of interstrand crossovers in a given duplex DNA, called the linking number (GLAUBIGER and HEARST 1967; CRICK 1976; BAUER 1978), and may also be reflected by the formation of intramolecular knots or intermolecular catenanes.

The first two topoisomerases were found in *Escherichia coli* (WANG 1971) and in embryonic mouse cells (CHAMPOUX and DULBECCO 1972). The prokaryotic prototype of this family of enzymes was designated ω-protein, whereas the eukaryotic enzyme, which differs in a variety of properties from the ω-protein, was by then called untwisting enzyme. Further topoisomerases with additional properties were found later. Topoisomerases have since been detected in a great number of pro- and eukaryotic organisms and also as components present in virus particles or encoded for by viral genomes. The original tentative designations, such as swivellase, nicking-closing enzyme, etc., have been given up in favor of a systematic nomenclature, which indicates that the action of these enzymes affects topological properties of DNA but normally not its primary structure (WANG and LIU 1979).

It is now, in principle, clear that topoisomerases serve an essential function in the macromolecular metabolism of DNA by providing a mechanism for the control of DNA conformation. This control has a variety of aspects. Many genetic processes require a defined DNA superstructure which has to be preserved. In addition, a number of DNA-dependent events are accompanied by transient conformational strain, which has to be released in order to restore the functional state of DNA. Furthermore, DNA has to be converted into condensed and complex superstructures during certain stages of the cell cycle and, finally, topologically linked DNA molecules have to be disentangled as, for instance, in the final stage of replication.

Conformation and shape of DNA are controlled not only by topoisomerases but also by other cellular components. Among them are enzymes as, for instance, DNA helicases and nonenzymatic proteins such as histones and DNA binding proteins. These proteins will not be considered in this article. For these proteins, reference is given to a number of relatively recent reviews some of which have also considered topoisomerases (CHAMPOUX 1978; GEIDER and HOFFMANN-BERLING 1981; DUGUET 1981; MCGHEE and FELSENFELD 1980).

Topoisomerases have been covered repeatedly by other review articles in the past, either in the context of structure and reactions of closed duplex DNA (BAUER 1978) or as topics in their own right, related to the whole area of topoisomerase research (WANG and LIU 1979; GELLERT 1981a; LIU 1983b) or to selected aspects of it (COZZARELLI 1980a, b; WANG 1981; GELLERT 1981b; WANG 1982b). In this study, a synopsis of the present knowledge of these enzymes is attempted, with special emphasis on the most recent developments. The field of topoisomerase-related research is rapidly expanding. One of the most significant achievements of the last 2 years is the progress in the understanding of the biological role of these enzymes: they solve problems of topological linkage wherever necessary, they are involved in compaction of DNA, and this, in turn, is tied into gene expression control mechanisms which operate in a *cis*-acting fashion by controlling DNA conformation.

Two figures depicting hypothetical models of topoisomerase mechanisms are included in this study. In addition, for the demonstration and discussion of models, in particular of DNA gyrase reactions, two articles by WANG (1982a, b) are recommended.

2 The Concept of Supercoiling

In the following brief exposition only general considerations are presented. For a rigorous treatment of the topic the reader is referred to CRICK (1976), BAUER (1978), and FULLER (1978).

The basic Watson-Crick structure of DNA in solution is constituted by the winding of one strand around the other with helical turns being repeated every 10.4 ± 0.2 base pairs (WANG 1979). In a circular, covalently closed duplex DNA molecule, the two strands are topologically linked. Covalent ring closure implies winding conservation, i.e., the number of strand rotations is invariant (and necessarily an integer), regardless of the conformation of the whole molecule. This number, which describes linking of complementary DNA strands quantitatively, is designated the topological winding number α (VINOGRAD et al. 1965) or, alternatively, the linking number (GLAUBIGER and HEARST 1967), denoted Lk or L (CRICK 1976). By convention, this number has a positive sign for right-handed duplex molecules.

If nicked circular DNA, which is under no torsional constraint, is closed by a DNA ligase, the resulting DNA molecules have topological winding numbers that roughly equal the number of duplex turns in the nicked DNA.

In this state, a closed DNA molecule is said to be relaxed. However, whereas nicked DNA is free to change the number of duplex turns in response to all conditions that affect the geometry of the DNA helix, such as counterion concentrations, pH, temperature, DNA binding agents etc. (WANG 1969a, b; BAUER 1978), the closed DNA does not change its linking number under the same conditions. Instead, it responds by altering its tertiary structure with the consequence that the DNA is supercoiled and no longer relaxed. A change in the linking number of such a supercoiled molecule would require a reversible chain scission in the DNA.

The linking number of relaxed DNA under physiological conditions, α_0, equals the number of base pairs per DNA molecule, N, divided by 10.4, the average number of base pairs per helical repeat (WANG 1979). Naturally occurring closed circular DNA molecules have linking numbers α that deviate from their respective α_0 values. The difference between these values, $\Delta\alpha = \alpha - \alpha_0$, was originally conceived as the number of superhelical turns in a given DNA molecule (VINOGRAD et al. 1968). However, since this concept did not strictly discriminate between topological and conformational properties of a DNA molecule, it has been rejected in favor of the concept of the linking difference (CRICK et al. 1979), $\Delta\alpha$, which is a quantity describing topological, not conformational, properties of a DNA molecule. Any linking difference in a DNA molecule affects its conformation, though only topological properties can be measured easily and accurately.

In order to compare topological properties in different DNA molecules, the expression of linking differences was normalized with respect to DNA length by referring to the specific linking difference $\Delta\alpha/\alpha_0$. This expression is also called the superhelical density of DNA. [The original definition of this quantity was $(\alpha-\beta)$ per ten base pairs, where β is the number of duplex turns (VINOGRAD and LEBOWITZ 1966). The concept of β was similar to that of α_0 by which it was later replaced.]

Values of $\Delta\alpha$ may be either positive or negative depending on whether a DNA is overwound or underwound with respect to its relaxed state under given conditions. If the topological winding number of a given molecule exceeds α_0, $\Delta\alpha$ is positive and the DNA is said to be positively supercoiled. Correspondingly, a negative value of $\Delta\alpha$ defines a negatively supercoiled DNA. All naturally occurring closed circular DNA molecules and also chromosomal DNA from pro- and eukaryotic organisms are negatively supercoiled. (Whether this conclusion is also true for archebacteria is as yet not known; see Sect. 6.1.3). Positively supercoiled DNA can be seen in vitro under certain conditions, e.g., as a consequence of dye binding. In vivo positive superhelicity may occur transiently in connection with DNA-dependent processes, such as DNA replication, but it is not a permanent property of DNA.

DNA samples with defined superhelical densities are not homogeneous with respect to the values of α_0 which the individual molecules have within these samples (KELLER 1975b; SHURE and VINOGRAD 1976; SHURE et al. 1977). Instead, these molecules show a distribution of topoisomers differing from each other by units of one. The values for α are, therefore, population averages.

Whereas α has to be an integer for a single DNA molecule, the average values for α need not be integers. For fully relaxed DNA, it has been shown that the distribution of topoisomers is Gaussian, centered essentially around $\Delta\alpha=0$ (PULLEYBLANK et al. 1975; DEPEW and WANG 1975).

Supercoiling has a variety of effects on the properties of DNA. All of these effects are intimately linked to the energetics of supercoiling (BAUER and VINOGRAD 1970; DAVIDSON 1972; HSIEH and WANG 1975; VOLOGODSKII et al. 1979). Since a winding deficit in DNA does not occur spontaneously, the formation of supercoils is associated with an increase in free energy. It has been experimentally determined that the difference in free energy between a particular topoisomer with a linking number α and a relaxed molecule with α_0 is proportional to $(\Delta\alpha)^2$ (PULLEYBLANK et al. 1975; DEPEW and WANG 1975). This relation is valid for superhelical densities up to -0.1. It has the corollary that all processes are favored which lead to unwinding of DNA and, hence, partially release the conformational strain associated with supercoiling. For example, the decrease in $\Delta\alpha$ by 1 in a DNA molecule with $\Delta\alpha/\alpha_0$ of -0.05 is accompanied by a change in free energy of approximately -6 kcal or -25 kJ.

Supercoiling has influences on physical, chemical, and biological properties of DNA. A supercoiled DNA molecule is much more compact than a relaxed molecule of the same size. This condition affects sedimentation coefficients in salt or sucrose solutions, mobility in gel electrophoresis, and other parameters that are affected by the degree of compactness of DNA molecules. Many of these parameters can be used to visualize changes in the topological winding number induced in vitro by topoisomerases. Chemical and biological effects of supercoiling pertain to the binding of ligands, the exposition of signal structures, such as cruciforms or left-handed turns of the helix, and the initiation of DNA-related catalytic events, e.g., transcription.

Basically, two different versions of supercoiling can be distinguished in living cells. The prototype of the first variety is the superhelical conformation of closed circular DNA. This conformation is under torsional strain. The second variety occurs in the bulk of the chromatin of eukaryotic cells, where DNA is wound around an octameric core of histones forming a left-handed toroidal superhelix (FINCH et al. 1977; MCGHEE and FELSENFELD 1980). This conformation is stabilized by the continuous presence of the histones. It is not under torsional strain and, therefore, does not require that both strands of the duplex are covalently closed. In closed circular DNA, the two forms of supercoiling can be interconverted by the addition or removal of histones (GERMOND et al. 1975). According to present knowledge, the DNA conformation in prokaryotes is predominantly of the first variety. It may be assumed that complexing of DNA with histonelike proteins in these organisms also leads to some supercoiling of the second variety.

The torsional stress of a closed circular supercoiled DNA can be relaxed (a) by local unwinding (denaturation) of the duplex; (b) by reducing the helix rotation angle, for instance, by the binding of intercalating dyes such as ethidium bromide; (c) by the introduction of at least one nick into one of the duplex strands; or (d) by the action of a topoisomerase. Unwinding of the duplex

and reduction of the helix rotation angle operate in the same direction without altering linking numbers. In contrast, nicking and the action of topoisomerases lead to a change in these numbers.

In addition, a different mode of at least partial reduction of superhelicity in circular DNA has been reported recently from studies on Z-DNA. Z-DNA is characterized by left-handed duplex turns with roughly 12 base pairs per helical repeat. This structure was identified in crystals of alternating deoxycytidilate-guanylate, d(CG), oligonucleotides (WANG et al. 1979; for an earlier analysis of Z-DNA properties see POHL and JOVIN 1972; for a recent review see RICH et al. 1984). If such sequences are present as artificial inserts in physiological DNA molecules, they undergo local transitions of the right-handed B form into the left-handed Z form in the presence of high levels of salt (KLYSIK et al. 1981) or under conditions of the conformational strain that is associated with superhelicity (PECK et al. 1982; NORDHEIM et al. 1982b). Natural DNA molecules also adopt left-handed conformations in the superhelical state (NORDHEIM et al. 1982b) or in the topological restrictions of form V-DNA (LANG et al. 1982; BRAHMS et al. 1982; POHL et al. 1982). [Form V-DNA results from the annealing of covalently closed complementary single strands (STETTLER et al. 1979); this DNA has a linking number of zero.] Z-DNA has been detected in interband regions of *Drosophila* polytene chromosomes (NORDHEIM et al. 1981) or in band regions of the polytene chromosomes of *Chironomus thummi thummi* (JOVIN et al. 1982). These observations suggest that left-handed turns in DNA may have a physiological significance, possibly as signal structures. Interesting results in support of this concept have recently been obtained showing that negatively supercoiled SV40 DNA contains Z-DNA segments within transcriptional enhancer sequences (NORDHEIM and RICH 1983).

In superhelical DNA, the local B-Z transition is accompanied by partial relaxation of the DNA circle without breaking and reclosing of phosphodiester bonds. The formation of left-handed turns is one of the structural alterations in superhelical DNA that are facilitated by the favorable change in free energy associated with the reduction in superhelicity. The introduction of one left-handed turn is equivalent to the removal of approximately two superhelical turns. The formation of left-handed turns in DNA can be detected by changes in sedimentation constants or gel electrophoretic mobilities (PECK et al. 1982). Other secondary structures which are stabilized by negative superhelicity due to their inherent tendency to relax torsional strain are hairpins or cruciforms in inverted repeat regions (LILLEY 1980; COUREY and WANG 1983; SINDEN and PETTIJOHN 1984) and those sequences which are characterized by increased sensitivity against single-strand specific nucleases such as S1 nuclease (WEINTRAUB 1983).

Topoisomerases both relax negative or positive supercoils, and they promote negative and, as is now known, also positive supercoiling. Not all members of the topoisomerase group are functionally identical. Some are only able to relax DNA, whereas others are also able to induce supercoil formation. More complicated topological reactions such as, for instance, formation of knots and catenation of separate DNA molecules are also catalyzed by these enzymes.

3 Classification and Terminology

It is important for an understanding of topoisomerase activity to appreciate that linking number changes may occur via two principally different routes: One includes changes in units of one or even multiples of one, the other one includes changes in units of two within a single reaction cycle. Changes in steps of one are the consequences of single-chain scissions and alterations of the number of srand crossovers, as will be discussed, either by rotation of the DNA helix about its long axis or some other kind of translocation of one strand relative to the other. Changes in steps of two require, as has been systematically shown by FULLER (1978), that a transient double-stranded break is produced in the DNA through which another segment of the same DNA molecule passes (passage may also involve two separate DNA molecules). Each single passage decreases or increases – depending on the direction of the event – the linking number by two.

On the basis of this distinction, two categories of topoisomerase are discriminated, type I and type II topoisomerases (LIU et al. 1980). The most conspicuous differences between these types are related to their mode of cleaving DNA and, in general, to their energy requirements. Type I enzymes act by introducing transient single-stranded cuts into duplex DNA molecules. These enzymes change linking numbers by units of one. Breaking and rejoining reactions by type II enzymes, in contrast, involve double-stranded interruptions of the DNA. In addition, they require ATP as a cofactor with the one known exception of topoisomerase II′ of *E. coli*. The linking number changes catalyzed by type II enzymes occur in units of two. The general properties of these enzymes are summarized in Table 1.

Both types of enzymes are present in pro- as well as in eukaryotic organisms. However, prokaryotic and eukaryotic type I enzymes differ from each other in a variety of respects, and the same has been found for the various type II enzymes. For the purpose of this review, the topoisomerases are, therefore, divided into four main classes, referring to the type of enzyme as well as to

Table 1. General properties of type I and type II topoisomerases

Type I topoisomerases	Type II topoisomerases
1. Transient single-strand breaks	Transient double-strand breaks
2. Change of linking in steps of one	Change of linking in steps of two
3. Knotting and catenation require nicked duplex molecules	Knotting and catenation occur with intact duplex molecules
4. ATP hydrolysis is not required	ATP hydrolysis is required[a]
5. Strand-passing mechanism obligatory for knotting and catenation. Relaxation may proceed by a stepwise swivel	Strand passing obligatory for all reactions

[a] The only known exception is topoisomerase II′ of *E. coli*

pro- or eukaryotic origin. Topoisomerases from mitochondria and chloroplasts as well as viral and recombinational topoisomerases are treated separately.

The first class of enzymes are the prokaryotic type I topoisomerases, which selectively remove negative superhelical turns from supercoiled DNA. The prototype of this class is the *E. coli* ω-protein first described by WANG (1971). The second class comprises the prokaryotic type II enzymes, also called DNA gyrases, which introduce, at the expense of ATP, negative superhelical turns into unconstrained closed circular DNA molecules (GELLERT et al. 1976a). Very recently, an exceptional gyrase was identified which generates positive supercoiling in DNA. This enzyme, which is called reverse gyrase, has been isolated from the archebacterium *Sulfolobus* (KIKUCHI and ASAI 1984). A third class consists of type I topoisomerases, which are associated with chromatin in eukaryotic cells. These enzymes act by relaxing both positively and negatively supercoiled DNA (CHAMPOUX and DULBECCO 1972). An additional class of eukaryotic topoisomerases has been described (BALDI et al. 1980; MILLER et al. 1981) which resemble the prokaryotic type II topoisomerases (gyrases) in their ATP requirement and also in their capacity to catalyze catenation and decatenation of closed DNA circles. However, these enzymes do not catalyze the principal reaction of gyrases, i.e., the introduction of negative superhelicity into DNA.

Various mitochondrial topoisomerases have been described, among them one from *Xenopus laevis* oocytes which clearly has the character of a eukaryotic type I enzyme (BRUN et al. 1981) and which is most probably identical to the nuclear type I topoisomerase of the same organism. The nature and origin of most other mitochondrial or chloroplast topoisomerase activities are less well defined.

Viral topoisomerases form a heterogeneous class of enzymes with regard to their functional and molecular details. At least one of them, the phage T4 topoisomerase, belongs to the prokaryotic type II class, others have type I properties.

The systematic designation of these enzymes as DNA topoisomerases (WANG and LIU 1979) has been generally agreed upon. A few of the older names, in particular DNA gyrase for the bacterial type II topoisomerase and ω-protein, or simply ω, for the topoisomerase I of *E. coli,* were sufficiently indicative to be retained. (The designation "gyrase" describes a swivel activity; this is in contrast to the now generally accepted view of the reaction mechanism of this enzyme.) The designations type I and type II topoisomerase refer to principal differences in reaction modes and requirements. If more than one category of topoisomerase is known to be present within a given biological source, this is indicated with the roman numerals I and II as suffixes. In general, topoisomerases I belong to the type I, and topoisomerases II belong to the type II category. Eukaryotic type I enzymes are occasionally still designated DNA nicking-closing enzymes.

4 Assay Procedures

All changes in the physical and chemical properties of a closed circular DNA molecule that occur as a consequence of an altered topological winding number can be used for measuring topoisomerase activity. Type I enzymes are normally monitored by utilizing properties which are specific for relaxed but not supercoiled DNA. Detection of type II enzymes is based either on the specific attributes of supercoiled DNA or on the dissolution of complex topological structures such as catenanes.

A simple qualitative test can demonstrate relaxation in the electron microscope, whereby superhelical DNA appears in a coiled form in contrast to relaxed DNA which is smoothly spread out (WANG 1971, 1974c). Identification of relaxed DNA can also be made by means of its hydrodynamic properties. The compact structure of supercoiled DNA involves a higher sedimentation coefficient than relaxed DNA in CsCl or sucrose solutions. CsCl zone sedimentation in the presence of ethidium bromide in a low concentration (3.5 μg/ml) was frequently used in the past to distinguish between relaxed and supercoiled DNA. Ethidium bromide intercalates into DNA and decreases the helix rotation angle upon intercalation by about 26° per dye molecule bound (WANG 1974a; KELLER 1975b; PULLEYBLANK and MORGAN 1975a). Titration of superhelical DNA with ethidium bromide causes first relaxation and then positive supercoiling (BAUER and VINOGRAD 1968). Thus, with a low dye concentration, superhelical DNA appears partially relaxed, and relaxed DNA becomes positively supercoiled with corresponding sedimentation coefficients for circular λ DNA (phage λ b2b5c) of 26 and 34 S, respectively (WANG 1971, 1974c).

CsCl buoyant density centrifugation in the presence of high ethidium bromide concentrations also provides a means of distinguishing different DNA superhelicities (GRAY et al. 1971; RADLOFF et al. 1967). With ethidium bromide in saturating concentrations, the distinction is based on the fact that DNA-dye complexes have lower buoyant densities than DNA alone. The amount of bound ethidium determines the extent of the shift. Since superhelical DNA binds more ethidium bromide under these conditions than relaxed DNA, the latter species bands at positions of higher density than the former (for details see BAUER 1978).

Another technique for monitoring topoisomerase action makes use of the fact that the fluorescence of ethidium is greatly enhanced upon binding to DNA (LEPECQ and PAOLETTI 1967). Any change, therefore, in the amount of ethidium bromide bound due to relaxation or supercoiling is associated with a change in the fluorescence signal. Relaxation decreases the signal, since relaxed DNA binds less dye than supercoiled DNA (MORGAN and PULLEYBLANK 1974; VOSBERG et al. 1975). Conversely, supercoiling leads to an enhanced signal (LIU and WANG 1978a). The principle of fluorescence enhancement by ethidium bromide has also been used for the development of a fluorescence spot test for topoisomerases (and other enzymes affecting DNA conformation), which allows the rapid handling of large numbers of samples (KOWALSKI 1980).

The most widely used assay for topoisomerases is based on the differential gel electrophoretic mobility of DNA topoisomers with molecular weights of

up to about 10^7. After electrophoresis, DNA in the gels is stained with ethidium bromide and visualized under UV light. The slow migration rate of relaxed DNA compared with that of superhelical DNA was first utilized by KELLER and WENDEL (1974) for the isolation of a topoisomerase from human cells. Gel electrophoresis is fast, inexpensive, and sensitive and provides the additional advantage that in agarose gels DNA topoisomers can be resolved that differ in their linking numbers by units of one (KELLER 1975b; SHURE and VINOGRAD 1976). This high resolving power has been of critical importance in assessing the different modes of action of type I and type II topoisomerases with respect to the basic change in DNA linking numbers induced by these enzymes. LEE et al. (1981) have shown that electrophoresis in two dimensions even allows positively and negatively supercoiled DNA to be distinguished. In the first dimension, the DNA migrates in the absence and in the second dimension in the presence of ethidium bromide or, alternatively, chloroquine (SHURE et al. 1977).

An alternative procedure for measuring in particular eukaryotic type I topoisomerases does not involve topological changes but the formation of intermediate covalent DNA-enzyme complexes, which can be trapped if single-stranded instead of double-stranded DNA is used as a substrate (PRELL and VOSBERG 1980). These complexes are bound to glass fiber filters through adsorption of their protein moieties (COOMBS and PEARSON 1978), and they can be recognized by radioactive label in the DNA component of these complexes. The most efficient complex formation has been obtained with circular single-stranded DNA such as that of phage fd. Since the enzyme does not turn over in this reaction, the number of enzyme molecules in the test has to exceed that of DNA molecules. The speed of this assay makes it specially valuable for the detection of topoisomerase activity during purification (SCHMITT et al. 1984).

A variation of this procedure has been developed which utilizes the precipitation of covalent DNA-protein complexes in the presence of SDS and KCl (TRASK et al. 1984). Uncomplexed nucleic acids do not precipitate. This assay could probably also be used for a rapid distinction between type I and type II topoisomerases, since it has been shown that topoisomerase II from *Drosophila melanogaster* is very inefficient in forming covalent complexes with DNA if EDTA is added prior to SDS (SANDER and HSIEH 1983).

The prokaryotic type II topoisomerases (DNA gyrase) are most conveniently measured by agarose gel electrophoretic analysis of relaxed and superhelical DNA (GELLERT et al. 1976a; SUGINO et al. 1977). These enzymes also catalyze catenation/decatenation and knotting/unknotting reactions with duplex DNA (see p. 49). Catenated DNA exists naturally in the form of huge networks of interlocked DNA molecules, called kinetoplast DNA, or kDNA, in the mitochondria of trypanosomes (BORST and HOEIJMAKERS 1979). Knotted DNA has been isolated from the truncated tailles capsids of the phages P4 (LIU et al. 1981a) and P2 (LIU et al. 1981b). These DNA species are suitable substrates for those type II topoisomerases that are unable to promote supercoiling, such as the topoisomerase of the phage T4 and the known eukaryotic type II topoisomerases, but which perform other type II specific reactions, such as decatenation and unknotting. In addition, these substrates are appropriate for monitor-

ing type II enzymes even in the presence of high concentrations of type I enzymes, since the latter are unable to catalyze these reactions (MILLER et al. 1981).

Quantitation of topoisomerase activity is tedious, but possible. Activity units are defined as those amounts of enzymes which relax (or supercoil) 50% of a given amount of DNA. Accurate evaluation may be achieved by densitometric scans of agarose gel photographs taken from the ethidium-DNA bands under UV light.

It should be added that the sensitivity of the electrophoretic procedure mentioned above allows a number of research applications in the analysis of proteins that alter the helix structure of DNA. The principle of this analysis involves ligating nicked circular DNA in the presence and in the absence of the protein of interest. After ligation, the closed circular product is analyzed on agarose gels for changes in the linking number of the test DNA (WANG et al. 1977; JAVAHERIAN et al. 1978; LIU and WANG 1978a). Instead of ligation, a topoisomerase reaction with relaxed closed circular DNA can also be applied in these analyses, which pertain to a broad spectrum of problems such as effects of DNA binding proteins (DUGUET et al. 1981), the unwinding activity of the SV40 large T antigen (GIACHERIO and HAGER 1980), the wrapping of DNA in nucleosomes (STEIN 1980), or the ATP-dependent unwinding of the double helix by the *E. coli rec*A protein (IWABUCHI et al. 1983).

It seems noteworthy that the problem of the helical repeat of DNA in solution could also be settled by applying the straightforward and unsophisticated technique of electrophoresis in agarose gels (WANG 1979).

5 Type I Topoisomerases

Enzymes of this category are characterized by a nicking-closing mechanism, which includes a transient single-strand nick. Pro- and eukaryotic type I topoisomerases differ in a variety of properties and are, therefore, treated separately except for the discussion of their reaction mechanism. Table 2 presents a comparison of the properties of pro- and eukaryotic type I topoisomerases

5.1 Prokaryotic Type I Topoisomerases

The first known topoisomerase, originally designated ω, was detected in *E. coli* (WANG 1971); it is now called *E. coli* DNA topoisomerase I. Since then, analogous enzymes have been isolated from the following bacterial species: *Bacillus megaterium* (BURRINGTON and MORGAN 1978), *Micrococcus luteus* (KUNG and WANG 1977; HECHT and THIELMANN 1977), *Agrobacterium tumefaciens* (LEBON et al. 1978), *Haemophilus gallinarum* (SHISHIDO and ANDO 1979), and *Salmonella typhimurium* (WANG and LIU 1979).

The most thoroughly studied enzyme of this class is the *E. coli* ω-protein. The following presentation is, therefore, restricted mainly to the properties of

Table 2. Reactions and properties of pro- and eukaryotic type I DNA topoisomerases

	Prokaryotic topo I[a]	Eukaryotic topo I
Relaxation of negative supercoils	+	+
Relaxation of positive supercoils	−	+
Knotting of single- or double-stranded DNA circles	+	n.d.[b]
Intertwining of complementary single-stranded circles	+	+
Catenation and decatenation	+	+
Interstrand transfer	−	+
Requirement for Mg^{2+}	+	−
Involvement of SH groups	−	+[c]
Site-specific binding	+	+
Enzyme linkage to 3′ ends of DNA	−	+
Enzyme linkage to 5′ ends of DNA	+	−
Phosphotyrosyl linkage between DNA and enzyme	+	+
Inhibition by neomycin[d]	+	−

[a] Most of the data for prokaryotic type I topoisomerases were obtained with *E. coli* ω-protein
[b] n.d., no data reported
[c] This result is not confirmed by TANG (1978)
[d] Unpublished results from this laboratory

this enzyme. The structural gene of this enzyme has been mapped recently. It is located at 28 min on the *E. coli* standard map (BACHMANN and LOW 1980) between the loci *trp* and *cys*B (STERNGLANZ et al. 1981; TRUCKSIS and DEPEW 1981). The gene is designated *top*A. Fine structure analysis of the cloned *E. coli* *top*A gene revealed that it is located immediately adjacent to *cys*B. Its distance from the *trp* locus is about 7 kbp (WANG and BECHERER 1983).

The ω-protein can easily be purified to homogeneity by standard techniques (DEPEW et al. 1978). From 1 kg packed cells 20 mg pure enzyme may be obtained. No direct evidence is available on the number of topoisomerase I molecules per cell. From the published purification data and from a comparison of specific activities in extracts and purified fractions, one might estimate about 500 enzyme molecules per cell. Immunological data, however, suggest that the relative content of ω-protein is lower by a factor of 10 (TRUCKSIS and DEPEW 1981). For preparative purposes overproducing strains carrying the cloned *top*A gene on a multicopy plasmid are available (TRUCKSIS and DEPEW 1981; WANG and BECHERER 1983).

The native enzyme is a single-subunit protein with a molecular weight of ca. 105000. It is insensitive to the sulfhydryl reagent N-ethylmaleimide. Mg^{2+} is required for activity. Recent studies in this laboratory indicate that ω belongs to the rare group of proteins that carry Mg^{2+} tightly bound. The strength of binding is illustrated by the fact that a fivefold stoichiometric excess of the chelating compound ethylenediamine tetraacetate (EDTA) over the number of enzyme molecules is required to destroy the relaxing activity of ω observed in the absence of added Mg^{2+}. Addition of 1 m*M* $MgCl_2$ stimulates the activity of ω to the level which is usually taken as the standard activity. Ca^{2+} or Co^{2+} may replace Mg^{2+} in this stimulation to some degree. High concentrations

of Mg^{2+} ($\geqq 20$ m*M*) are inhibitory (SCHÖLER 1982). Whether this result indicates a direct effect on the enzyme or an indirect effect due to a structural variation of the DNA is not known. A strong inhibitor of the relaxation reaction is single-stranded DNA (WANG 1971). Inhibition is also observed with two antibiotics, neomycin sulfate, which is otherwise known to inhibit protein biosynthesis, and actinomycin D (SCHÖLER 1982).

Four different topoisomerization reactions are catalyzed by *E. coli* topoisomerase I: (a) the relaxation of negatively supercoiled DNA (WANG 1971), (b) linking (or intertwining) of covalently closed single-stranded DNA rings containing complementary base sequences (KIRKEGAARD and WANG 1978), (c) the formation of topological knots in single-stranded DNA rings (LIU et al. 1976), and (d) the catenation and decatenation of double-stranded DNA circles (TSE and WANG 1980; LOW et al. 1984). The latter reaction occurs only if one of the reacting DNA molecules has a nick. Positively supercoiled DNA is essentially not relaxed by this enzyme. A fifth reaction may be mentioned, i.e., the conversion of paranemic into plectonemic joints (BIANCHI et al. 1983). Paranemic joints consist of homologous DNA strands which are paired without being topologically linked. By topoisomerization they are converted into truly interwound or plectonemic joints (WATSON and CRICK 1953). The reaction resembles the linking of covalently closed single-stranded rings.

A complete picture of the reaction mechanism of the ω-protein does not as yet exist. However, some detailed information is available on the following aspects of its activity: the requirements for binding of enzyme to DNA, the nature of the reaction intermediates, the status of the termini at the cleavage site, and the kinetics of the reaction.

Requirements for DNA Binding and Cleavage. The ω-protein is able to form stable complexes with single-stranded DNA (DEPEW et al. 1978). These complexes are stable even in molar concentrations of salt if Mg^{2+} is omitted from the solution. Subsequent addition of Mg^{2+} in low concentrations releases the enzyme from the DNA in an intact form. A similar complex was observed with double-stranded negatively supercoiled DNA. The strong dependence of complex formation on the degree of negative superhelicity suggests that local unwinding of the DNA is involved in binding of the enzyme to such DNA. A salt-stable complex is also found with nonsuperhelical DNA, but this complex is of a different nature, since it is not dissociated upon addition of Mg^{2+} (LIU and WANG 1979).

The formation of a covalent complex between protein and DNA, which occurs subsequent to binding, apparently does not depend on defined DNA nucleotide sequence. This has been shown for both the ω-protein from *E. coli* and the corresponding topoisomerase of *M. luteus* (TSE et al. 1980), using as substrate single-stranded DNA derived from pBR322 restriction fragments of defined length. Analysis of more than 40 cleavage points showed that some of them were markedly preferred to others. Comparison of sequences surrounding different cleavage points revealed no specific sequence on either the 5′ or the 3′ side of the break; the only fairly regular feature was that in 90% of the analyzed sequences there was a C residue in the fourth position on the

3′ side of the breakpoints. Some of the cleavage points were shared by the two enzymes. At present, it is not clear whether the nonrandom site distribution reflects specific recognition of single nucleotides or nucleotide sequences by these enzymes or whether recognition is linked to secondary or other structural features of the single-stranded DNA used in these experiments. In particular, it would be interesting to know whether the preferred sites on single-stranded DNA are identical with the sites on double-stranded DNA of the same sequence.

With oligonucleotides of defined length, it has been shown that oligo$(dA)_7$ and oligo$(dT)_8$ can be cleaved. Shorter oligonucleotides are not cleaved. The site of cleavage is in both cases four bases from the 3′ end. Oligo(dG) or oligo(dC) are very poorly or not cleaved at all (Tse-Dinh et al. 1983).

Nature of the Intermediates. If alkali is added to the salt-stable complexes, fragmentation of the DNA chain occurs. This result is obtained with single-stranded and superhelical DNA, but not with relaxed DNA (Depew et al. 1978; Liu and Wang 1979). Enzyme molecules remain attached to these fragments in a covalent mode. It is still open to question whether the breakage reaction occurs during complex formation but becomes apparent only after exposure of the complex to alkali, or whether DNA fragmentation is a consequence of alkali treatment. Two lines of evidence demonstrate quite clearly that fragments possess a free 3′-hydroxyl terminus and are linked to the enzyme by their 5′ ends. First, the fragments are susceptible to *E. coli* exonuclease I and the exonuclease activity of the T4 DNA polymerase, which both act in 3′ to 5′ direction. Secondly, single-stranded linear DNA was double labeled with ^3H at internal positions and with ^{32}P at the 5′ end. Upon cleavage, the enzyme was linked only to ^3H-labeled material indicating a bond to the 5′ side at the nick. This 5′ bonding was formed with single-stranded DNA as well as with superhelical DNA (Liu and Wang 1979).

The DNA 5′-phosphoryl group is linked to the O^4 position of a tyrosyl residue of the enzyme (Tse et al. 1980). Transient phosphodiester bond formation with a tyrosyl residue appears to be a general feature of topoisomerase reactions.

Kinetics. Very few and only qualitative data are available on the reaction kinetics. That the action of ω and other pro- and eukaryotic topoisomerases is catalytic, i.e., that the enzyme is able to act repeatedly, has been shown by various authors (for references see Wang and Liu 1979). If, with a freshly prepared enzyme, DNA is incubated at a molar ratio of enzyme – DNA less than one – then all DNA molecules will eventually be relaxed.

Reaction rates for the topoisomerization reactions have not yet been assessed. However, the relaxation of negatively supercoiled DNA shows a strong dependence on the degree of superhelicity of the substrate. Relaxation of weakly supercoiled DNA requires high enzyme concentrations and long incubation times to come to completion. It should be noted that this observation fits well with the data on alkali cleavage of DNA-enzyme complexes. Relaxed DNA forms salt-stable complexes with the enzyme, but it is, in contrast to superhelical DNA, not cleaved after addition of alkali; the complex merely dissociates.

The ω-protein of *E. coli* and the corresponding enzyme of *M. luteus* are both able to act in a distributive as well as in a processive fashion, depending on the degree of supercoiling in the DNA substrate and on the salt concentration of the medium. With highly supercoiled DNA and a K^+ concentration of less than 30 mM, the mode of action is predominantly processive. This can be concluded from a bimodal distribution of either superhelical or relaxed DNA molecules observed under conditions of substrate excess. With a 0.1 *M* K^+ concentration, an essentially monomodal distribution of partially relaxed DNA molecules can be identified during the course of the reaction under the same conditions. DNA with a low degree of supercoiling is always relaxed in a distributive fashion (KUNG and WANG 1977; WANG and LIU 1979).

The latter result and also the observation that partially relaxed DNA molecules appear under conditions of a slow overall reaction rate, e.g., at low temperature or in the absence of added Mg^{2+} (Schöler and Vosberg, unpublished observation), demonstrate quite clearly that the process of relaxation can be interrupted at any intermediate level of supercoiling. Although this result is not final proof, it suggests strongly that the ω-protein acts predominantly in a stepwise fashion and not by a single-hit mechanism.

This aspect and other mechanistic considerations of the reaction of the ω-protein will be discussed in Sect. 5.2 together with the mechanism of eukaryotic type I topoisomerases.

Properties of Other Bacterial Type I Topoisomerases. A new type I topoisomerase in *E. coli*, called topoisomerase III, was recently described in two independent reports (DEAN et al. 1982; SRIVENUGOPAL et al. 1984). This enzyme is a single polypeptide with a molecular weight of 74000. It changes linking numbers in steps of one. It differs from the ω-protein in its requirement for K^+ in addition to Mg^{2+} for optimal activity and in its affinity for novobiocin-Sepharose. Positive supercoils are not relaxed. ATP is neither cleaved, nor does it inhibit the enzyme. No inhibition is seen with oxolinic acid or novobiocin, the standard gyrase inhibitors. A potent inhibitor is single-stranded DNA. Topoisomerase III is a separate enzyme unrelated to topoisomerase I. This was shown with antitopoisomerase III antibodies which did not cross-react with topoisomerase I and by demonstrating topoisomerase III in extracts of *E. coli* cells in which the topoisomerase I gene (*top*A) had been deleted.

Another well-characterized prokaryotic topoisomerase I is that from *M. luteus* (KUNG and WANG 1977; HECHT and THIELMANN 1977). It has a molecular weight of 120000 and is very similar to the ω-protein in all its enzymatic characteristics. There is no immunological cross-reaction of the *M. luteus* enzyme with anti-ω-antibodies (KUNG and WANG 1977).

The other known bacterial type I topoisomerases from *B. megaterium*, *S. typhimurium*, *H. gallinarum*, and *A. tumefaciens* all have a molecular weight between 100000 and 120000 and are all rather similar in their properties to the *E. coli* ω-protein.

In *S. typhimurium*, a topoisomerase I has been genetically and biochemically defined as the product of the *sup*X gene (OVERBYE and MARGOLIN 1981; TRUCKSIS and DEPEW 1981; TRUCKSIS et al. 1981). *sup*X is located on the *Salmonella*

gene map between the *trp* and *cys*B loci, a position which was also found for the *top*A locus on the *E. coli* chromosome (STERNGLANZ et al. 1981). The phenotype of *sup*X (now called *top*A) reveals pleiotropic effects, which include suppression of a number of promoter mutations (OVERBYE and MARGOLIN 1981). The role of *top*A in the genetic control of DNA supercoiling in *Salmonella* has been thoroughly worked out recently (RICHARDSON et al. 1984).

A topoisomerase I from *A. tumefaciens* was the first topoisomerase to be subjected to a study involving monoclonal antibodies. Inhibition of catalytic activities by these antibodies always affected both the nicking and the closing reactions, a result which suggests that these functions reside in identical or closely associated sites within the enzyme molecule (LEBON et al. 1981).

5.2 Eukaryotic Type I Topoisomerases

Eukaryotic organisms or cells from which type I topoisomerases have been isolated include: lower eukaryotes like yeast (DURNFORD and CHAMPOUX 1978; BADARACCO et al. 1983; GOTO et al. 1984) or *Ustilago maydis*, another fungus (ROWE et al. 1981); the protozoon *Trypanosoma cruzi* (RIOU et al. 1983); tissue culture cells of mouse and human origin (KELLER 1975a; VOSBERG and VINOGRAD 1976; TANG 1978; LIU and MILLER 1981; DURBAN et al. 1983; ROSS et al. 1983; ISHII et al. 1983); rat and bovine liver (CHAMPOUX and MCCONAUGHY 1976; KOWALSKI 1980); calf thymus (PULLEYBLANK and MORGAN 1975b; PRELL and VOSBERG 1980; SCHMITT et al. 1984); avian erythrocytes (CAMERINI-OTERO and FELSENFELD 1977; BINA-STEIN et al. 1976; PULLEYBLANK and ELLISON 1982; TRICOLI and KOWALSKI 1983; TRASK and MULLER 1983); *D. melanogaster* eggs and embryos (BAASE and WANG 1974; HSIEH and BRUTLAG 1980; JAVAHERIAN et al. 1982); *X. laevis* (LASKEY et al. 1977; MATTOCIA et al. 1976; GANDINI-ATTARDI et al. 1981); salmon testis (ESKIN and MORGAN 1978); and sea urchins (POCCIA et al. 1978). In addition, two reports exist on similar enzymes in higher plants, wheat germ (DYNAN et al. 1981), and cauliflower (FUKATA and FUKASAWA 1982).

Purification and General Properties. Since topoisomerases are located in chromatin, purification usually starts with the isolation of nuclei from the tissues or cells of interest. The subsequent fractionation commonly involves treatment of nuclear extracts with high salt concentrations (e.g., 1 *M* KCl, or higher) in order to dissociate the enzyme from chromatin. DNA is removed by polyethylene glycol, and the enzyme is then purified by a sequence of chromatographic steps which are based on ion exchange (e.g., phosphocellulose, Bio-Rex 70), hydrophobic interaction (e.g., phenyl sepharose), affinity binding (e.g., sepharose-bound heparin, single- or double-stranded DNA coupled to cellulose), or separation according to size (e.g., gel filtration). Less frequently used methods include fractionation by polymin P and ammonium sulfate precipitation, or preparative electrofocusing. (For more details see, e.g., the following references: CHAMPOUX and MCCONAUGHY 1976; LIU and MILLER 1981; DYNAN et al. 1981; GANDINI-ATTARDI et al. 1981; LIU 1983a; SCHMITT et al. 1984.)

The type I topoisomerase is present in fairly high copy numbers per nucleus. A minimum of about 7×10^4 enzyme molecules has been estimated for the rat liver cell (CHAMPOUX and MCCONAUGHY 1976). Our own estimates for calf thymus nuclei based on purification from known numbers of nuclei range from 3×10^5 to 6×10^5 molecules per nucleus (which corresponds to 0.25%–0.5% of the total nonhistone protein in calf thymus nuclei) (PRELL 1980). An even higher value of up to about 1×10^6 copies per cell nucleus can be calculated from data obtained with HeLa cells (LIU 1980; LIU and MILLER 1981). This value would be equivalent to about one topoisomerase per 10–15 nucleosomes. Essentially similar numbers have recently been presented by others (TRICOLI and KOWALSKI 1983).

Different molecular weights have been reported for the eukaryotic type I topoisomerase. Some early reports indicated molecular weights of 66000–70000 for the native enzyme from rat, human KB, and mouse L cells (CHAMPOUX and MCCONAUGHY 1976; KELLER 1975a; VOSBERG and VINOGRAD 1976). Recent data for the enzyme of chicken erythrocytes (62000 daltons, see PULLEYBLANK and ELLISON 1982), *X. laevis* (67000 daltons, see GANDINI-ATTARDI et al. 1981), *T. cruzi* (65000 daltons, RIOU et al. 1983), and mouse leukemia cells (68000 daltons, ROSS et al. 1983) are essentially in agreement with the earlier results. However, evidence exists which indicates that the correct values for the native size of eukaryotic type I toposisomerases are about 100000 daltons, as demonstrated for the enzymes from HeLa cells (LIU and MILLER 1981), wheat germ (DYNAN et al. 1981), *Drosophila* embryos (HSIEH and BRUTLAG 1980), and calf thymus (SCHMITT et al. 1984). A still higher molecular weight of about 135000 has been reported recently for the *Drosophila* enzyme (JAVAHERIAN et al. 1982) and for a topoisomerase I of yeast (BADARACCO et al. 1983). The reason for the discrepancies between the early and the recent data is most probably proteolytic cleavage of the enzyme during purification and possibly also during storage of tissue materials. This proteolysis can be controlled to some degree by adding protease inhibitors, such as phenylmethylsulfonyl fluoride (PMSF) and Trasylol, and also by rapid preparation of the enzyme. LIU and MILLER (1981) showed that a catalytically active 67000-dalton component could be produced in vitro by limited proteolytic digestion of the native 100000-dalton enzyme from HeLa cells. In addition, further proteolysis of the 67000-dalton protein gave rise to subfragments which were also seen after digestion of the larger enzyme.

Purified calf thymus enzyme appears even under optimal conditions with fresh tissue as a set of five proteins with molecular weights of between 70000 and 100000, the 100000-dalton protein being one of the major components. After isolation of these proteins from sodium dodecyl sulfate (SDS) polyacrylamide gels and reactivation (according to HAGER and BRUGESS 1980), all five showed topoisomerase activity (SCHMITT et al. 1984).

A multiplicity of catalytically active forms of eukaryotic type I topoisomerases from a variety of sources has now been well documented (see TRASK and MULLER 1983; LIU 1983; TRICOLI and KOWALSKI 1983; MARTIN et al. 1983; ISHII et al. 1983).

Only one systematic comparison between the enzymatic properties of the 100000 dalton and the 67000-dalton enzyme has been reported; no differences

in the catalytic behavior of the two proteins could be detected (LIU and MILLER 1981). Thus, removal of about one-third of the native protein does not impair the enzymatic activity of the topoisomerase. This part of the protein molecule may conceivably serve additional functions which escape detection in the standard assays, such as interaction with other proteins in chromatin.

No gross differences have been found for the reaction requirements of the eukaryotic type I enzymes of different biological origin. Although systematic comparisons have not been made, it is likely that these enzymes have basically common properties and act according to a unique mechanism. They are, therefore, considered indiscriminately in the following section. The enzymes have optimal activity at concentrations of 150–200 m*M* Na^+ or other monovalent cations. Replacement of 200 m*M* Na^+ by 10 m*M* Mg^{2+} is accompanied by some loss of activity. Below a concentration of 100 m*M* Na^+, the enzyme acts in a processive fashion, and above 150 m*M* salt in a nonprocessive fashion (MCCONAUGHY et al. 1981). The enzyme of HeLa cells, *D. melanogaster* and yeast is stimulated significantly if 10 m*M* Mg^{2+} is added to the standard mixture containing 100 or 200 m*M* Na^+ (LIU and MILLER 1981; JAVAHERIAN et al. 1982; GOTO et al. 1984). A distinct pH optimum does not exist; activity is seen within a broad range between pH 6.5 and 8.

Inhibition by SH reagents like N-ethylmaleimide or p-hydroxy mercuribenzoate has been reported repeatedly (e.g., KELLER 1975a; VOSBERG et al. 1975; PULLEYBLANK and ELLISON 1982). Only one observation to the contrary has been reported (TANG 1978). Other inhibitors of these enzymes are the antibiotics actinomycin D and adriamycin (SCHÖLER and VOSBERG, unpublished results). The effects of these drugs may be explained by their known ability to complex with DNA. Thus, they presumably prevent access of the enzyme to its substrate. Furthermore, competitive inhibition is thought to occur with heparin (ISHII et al. 1982). Heparin as a polyanion is also known to bind to other DNA-dependent enzymes, e.g., RNA polymerase (STERNBACH et al. 1975). Use has already been made of binding to heparin for the purification of the enzyme (GANDINI-ATTARDI et al. 1981; SCHMITT et al. 1984). Inhibition by polydeoxyriboguanylic acid [poly(dG)] and polyriboguanylic acid [poly(rG)] (PRELL and VOSBERG 1980) may reflect a specific affinity of the enzyme for G-containing regions in DNA.

Eukaryotic topoisomerases carry out the following reactions. They relax both negatively and positively supercoiled DNA circles (CHAMPOUX and DULBECCO 1972; VOSBERG et al. 1975; TANG 1978). They also promote intertwining of complementary single-stranded DNA circles (CHAMPOUX 1977b) by forming relaxed duplex circles covalently closed in both strands. If denatured supercoiled PM2-DNA is used as a substrate, relaxed duplex circles as well as highly knotted molecules are seen among the products (LAU et al. 1981). These compact structures contain an unusual type of topological linkage in which a single-stranded segment in one region of the double-stranded molecule appears to pass between the strands of another region of the same molecule. The result is a structure with single- and double-stranded regions tied into a single DNA molecule. Catenane formation, a reaction specific for type II topoisomerases, has been reported for a eukaryotic type I topoisomerase of rat liver (BROWN and COZZAR-

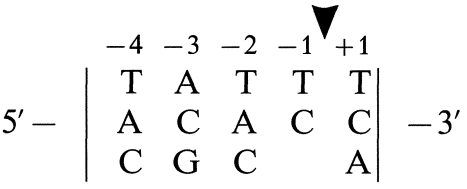


Fig. 1. A consensus sequence for the cleavage site of topo I from human and calf thymus cells. This sequence was deduced from cleavage data obtained with SV40 DNA. The *arrow* indicates the cleavage point. The probability for the occurrence of a particular nucleotide in a particular position decreases from top to bottem. Thus, the hierarchy in the position −4 is: T>A>C. (EDWARDS et al. 1982)

ELLI 1981) and yeast (BADARACCO et al. 1983). This reaction requires at least one nick in one of the DNA components to be catenated. Extensive sequence homology between the component circles is not involved. Spermidine, present at a concentration of 5 mM as a DNA condensing agent, is favorable or may even be essential for this reaction (KRASNOW and COZZARELLI 1982). Similar to the *E. coli* ω-protein, a eukaryotic topoisomerase I (of *U. maydis*) has been shown to catalyze the conversion of a paranemic to a plectonemic joint (KMIEC et al. 1983).

In addition to the classic reactions, a novel type of reaction of eukaryotic type I topoisomerases was recently observed in the course of which an enzyme-attached single-stranded fragment is transferred and covalently ligated to the 5′ end of a separate double-stranded DNA molecule (BEEN and CHAMPOUX 1981; HALLIGAN et al. 1982; TRASK and MULLER 1983). This intermolecular chain transfer will be discussed in detail in Sect. 5.2.3.

Sequence Recognition, Cleavage, and Reaction Intermediates. The eukaryotic topoisomerases recognize DNA sequences in a nonrandom fashion. This was first seen with single-stranded fd DNA, which is converted by calf thymus topoisomerase into a set of distinct fragments that predominantly carry T or A residues at their 5′ ends. The latter result was also found with double-stranded fd DNA (PRELL 1980). More direct evidence on specific recognition was recently provided by EDWARDS et al. (1982), who studied cleavage of SV40 DNA by topoisomerase I of HeLa cells and calf thymus. These authors found that both enzymes recognize the same sequences in single- and in double-stranded DNA. No unique sequence could be identified in the analysis of 68 cleavage sites of single-stranded DNA derived from SV40 DNA. A characteristic feature, however, was that certain nucleotides were frequently excluded from these sites within positions −4 and +1 about the point where DNA was nicked. The least frequent nucleotide found within this region was G. The most frequently cut dimer constellations were TT, TC, CT, or CC, with a striking preference for TT dimers. Cleavage between two purines was particularly infrequent within this collection of sites. A consensus sequence, which was derived from these data, is schematically depicted in Fig. 1. An essentially similar consensus sequence has been determined for the type I topoisomerases of both rat liver and wheat germ (BEEN et al. 1984a, b), however, with fewer ambiguities in the positions −4 to −1 and no nucleotide preference for the +1 position. These results show that the cleavage behavior of this enzyme from entirely unrelated organisms is very conservative.

A putative subset of topoisomerase I recognition and cleavage consensus sequences has been identified recently by Westergaard and his collaborators (Westergaard, pers. communication). These authors found in flanking regions of transcribed genes of different organisms a quadrodecameric consensus sequence

$$5'-\text{A C T T A G A}\begin{matrix}\text{G}\\\text{A}\end{matrix}\text{A A A}\begin{matrix}\text{A}\\\text{T}\end{matrix}\begin{matrix}\text{A}\\\text{T}\end{matrix}\begin{matrix}\text{A}\\\text{T}\end{matrix}-3',$$

with cleavage occuring between T and A in position 4 and 5 (results to be published by Westergaard and his collaborators). It should be noted that the 5′ end of this quadrodecameric motif agrees fairly well with the quadromeric consensus sequence for topoisomerase I cleavage as defined by EDWARDS et al. (1982) and BEEN et al. (1984) (for further details see p. 80).

In a study on base-specific inhibition of calf thymus topoisomerases, DARBY and VOSBERG (1985) recently observed that the relaxation of supercoiled DNA by topoisomerase I is effectively blocked if dGMP in the DNA is replaced by a phosphorothioate dGMP (dGMPαS). Other base-specific substitutions are much less effective. In considering earlier results about inhibition by poly(dG) or poly(rG) (PRELL and VOSBERG 1980) and the strikingly infrequent appearance of dGMP within the consensus sequence for cleavage, one might speculate that dGMP or G-containing sequences are involved in DNA recognition by topoisomerase remote from the cleavage site. However, a simple recognition pattern cannot be deduced from the published cleavage maps (EDWARDS et al. 1982; BEEN et al. 1984b).

Eukaryotic type I topoisomerases react covalently with the 3′ side of the single-stranded nick, which occurs in the course of the reaction (CHAMPOUX 1976, 1977a; PRELL and VOSBERG 1980; EDWARDS et al. 1982), in contrast to the *E. coli* ω-protein, which reacts with the 5′ side of the nick (DEPEW et al. 1978; PRELL and VOSBERG 1980).

The size of the binding site of eukaryotic topoisomerase I on duplex DNA has been determined using micrococcal nuclease digestion of unprotected DNA. Topoisomerase I blocks access to the helix over a distance of about 25 base pairs (TRASK and MULLER 1983). An oligonucleotide of about that size is also protected if complexes of single-stranded DNA and enzyme molecules are digested with pancreatic DNase (CHAMPOUX 1981; SCHMITT et al. 1984).

Mechanism of Topoisomerase I Action. In the following discussion, the reaction modes of pro- and eukaryotic type I topoisomerases will be considered together.

Relaxation of superhelical DNA by a eukaryotic topoisomerase I generates a Boltzmann distribution of topoisomers that is indistinguishable from a distribution generated by DNA ligase with DNA molecules containing single-stranded nicks (PULLEYBLANK et al. 1975; DEPEW and WANG 1975; DELEYS and JACKSON 1976). Topologically homogeneous DNA isolated from a single band within this distribution and rereacted with a topoisomerase shows odd and even linking numbers (PULLEYBLANK et al. 1975). This result shows that this eukaryotic enzyme relaxes DNA to completion and, more important, that it changes the linking number of closed DNA in units of one.

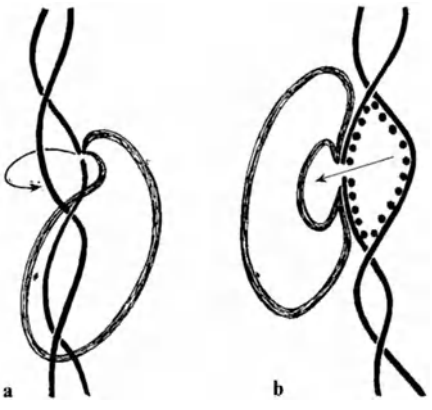


Fig. 2a, b. Possible mechanism of action of type I topoisomerases. Two different modes by which type I topoisomerases may relax supercoiled DNA are schematically drawn. **a** The swivel mode would begin with a single-stranded nick followed by rotation of the nicked strand about the long axis of the DNA helix. Two DNA recognition and binding sites are proposed for the enzyme, one reacting covalently with one end (3′ or 5′) of the nick and one reacting noncovalently with a region on the other side of the nick. There is no a priori need for a stepwise mechanism. However, if the interaction between the DNA and the noncovalent binding site of the enzyme is sufficiently strong, the nicking step would be followed by a closing step after limited rotation of the nicked strand for one full turn. No separation of the DNA strands would be required for this mode of action. **b** The strand-passing mode would also start with a single-stranded nick, which eventually would be broadened to a small gap. The enzyme would tightly span the gap by reacting covalently with one end and noncovalently with the other end. This bridging of the gap would prevent swiveling of the DNA. The intact strand opposite to the nick would slide through the gap into the hollow which is formed by the enzyme. Some local melting would conceivalbly be involved in this reaction. This mode would be strictly stepwise

Relaxation by the ω-protein does not normally go to completion. The lower the degree of superhecilicity, the slower is the rate of the reaction. However, this prokaryotic enzyme also changes the linking number of DNA in steps of one (BROWN and COZZARELLI 1979). This result is in agreement with notion that both groups of enzymes act via single-strand nicks. Corroborating evidence for single-strand nicks has been provided by LIU and WANG (1979) for ω and by CHAMPOUX (1976) for rat liver topoisomerase I. Type II topoisomerases, as will be discussed later (Sect. 6), act via double-strand breaks and change the linking number of DNA in steps of two.

In contrast to the latter enzymes, for which a general mode of action has been proposed that is characterized by strand passing and sign inversion (see below, Sect. 6.1.1), a similar unique model for type I topoisomerases does not exist. To explain the reactions of these enzymes, two different modes have been suggested in the past, the swivel mode and a strand-passing mechanism (see Fig. 2). Ingredients of the swivel are nicking of one DNA strand, rotation of the duplex about its long axis, and reclosure. In this model, the intact strand would be the stationary pivot for the rotational movement of the nicked strand. Local melting of the DNA duplex would not be required for this event. Theoreti-

cally, relaxation could proceed as a single-hit or also as a multi-hit event (VOSBERG and VINOGRAD 1975). A single-hit mechanism would involve complete removal of superhelical turns between nicking and reclosure of DNA. A multi-hit mechanism would proceed via repeated nicking-closing cycles with removal of only one or a few superhelical turns in a single cycle.

The strand-passing mechanism is more complicated than the swivel mode. The reaction begins with a single-stranded nick. The nick would be bridged by the enzyme in order to prevent swiveling of the DNA. The enzyme would then create a lock, or a hollow, and the intact strand would be translocated through the nick into this lock. This translocation would require broadening of the nick to a small gap, which in turn would involve some local melting in the region of the gap. Such a mechanism has been proposed by BROWN and COZZARELLI (1981). Relaxation according to this mode would necessarily be a multi-hit event. The driving force for either the swivel or for strand passing would be provided by the torsional strain inherent to superhelix conformation.

Strand passing in the course of a catenation or knotting reaction would involve a nick in a region opposite to a preexisting nick in a duplex circle, either directly opposite or in a staggered fashion some base pairs away. Such a reaction would create a transient double-strand interruption in the DNA bridged by the enzyme in order to prevent diffusion of the DNA ends. The enzyme together with the DNA ends, to which it would be attached, would then form a lock into which an intact duplex segment of the same or another DNA molecule could penetrate similar to what is envisaged for the relaxation reaction. Subsequent reclosure of the transient interruption would complete the topological linkage of DNA molecules. The reverse reaction of decatenation would proceed via the same route.

The ω-protein on one hand and the eukaryotic type I topoisomerases on the other have many properties in common, but they also differ in a variety of respects (see Table 2). Any concept of the mechanism has to account for the similarities as well as for the differences. Briefly, the two groups of enzymes share the following properties: (a) They both relax and catenate circular DNA. (The formation of knots, which may be considered as an intramolecular catenation event, has so far only been reported for the prokaryotic ω-protein.) (b) Both groups of enzymes act via transient single-strand nicks and intermediate covalent DNA enzyme complexes. (c) All type I topoisomerases may act in either a processive or a distributive fashion, depending on the composition of the reaction medium. (d) Single-stranded DNA is a potent inhibitor of relaxation reactions of all type I enzymes.

Differences exist in the following respects: (a) Prokaryotic topoisomerases attach to the 5′ end at the nick, whereas eukaryotic topoisomerases react with the 3′ end at the nick. (b) The ω-protein, in contrast to the eukaryotic type I topoisomerases, does not relax positive supercoils. (c) The ω-protein does not readily relax DNA to completion. In contrast, eukaryotic topoisomerases do so and, furthermore, accept fully relaxed DNA as a substrate for a nicking-closing reaction. (d) A novel type of topoisomerase reaction, the single-strand chain transfer has only been observed with eukaryotic type I topoisomerases but not with the ω-protein.

Although both groups of enzymes are inhibited by single-stranded DNA, the underlying mechanisms of this inhibition are apparently not identical. The reaction of eukaryotic type I topoisomerases with single-stranded DNA has been significantly clarified by the recently discovered chain transfer which is mediated by these enzymes. This reaction has been extensively reviewed by LIU (1983b).

It had been shown earlier that eukaryotic type I topoisomerases react with single-stranded fd or denatured SV40 DNA (PRELL and VOSBERG 1980; BEEN and CHAMPOUX 1980). These reactions lead to spontaneous fragmentation of the DNA. The appearance of fragments does not depend on protein denaturants or proteolytic treatment, as are required for demonstration of single-strand fragmentation by the ω-protein (DEPEW et al. 1978). Fragments generated by a eukaryotic topoisomerase I are covalently attached to enzyme protein at their 3′ ends, while the 5′ ends carry hydroxyl groups. The formation of DNA fragments may be considered as an abortive reaction cycle of "nicking" without "closing." The DNA-protein complex should inherently be able to complete the reaction cycle by reconstituting a phosphodiester bond with a free 5′-hydroxyl end either of the same fragment which carries the protein or of another fragment. Such a reaction would result in intramolecular cyclization or in the production of new linear DNA units. The original DNA sequences would be rearranged in both cases.

Circle formation and strand transfer reactions of this kind have been demonstrated (BEEN and CHAMPOUX 1981; HALLIGAN et al. 1982). In these reactions the enzyme-linked 3′ ends of the fragments function as donor sites and the 5′OH ends as acceptor sites. No specific bases are required at the 5′ ends. Single- and double-stranded 5′ ends function equally well as acceptors, regardless of whether the ends are recessed, protruding, or flush. Even internal nicks in duplex molecules are accepted.

Functionally, the intra- and intermolecular joining reaction may be considered equivalent to the second half-reaction of a normal "nicking-closing" cycle, which leads to the restitution of a broken phosphodiester bond in the DNA. Whether the mechanism of this joining reaction is exactly identical with that which normally occurs in double-stranded DNA is not known. An important implication of these results is that eukaryotic topoisomerases bind tightly only to one of the two ends of the nicks that they generate.

These reactions occur in competition to normal relaxation reactions. By fragmenting single-stranded DNA molecules, the enzyme is trapped in an abortive nicking-closing cycle. Chain transfer reactions are infrequent when DNA concentrations are low. The enzyme, therefore, does not turn over under these conditions. Competition experiments done in this laboratory with calf thymus topoisomerase (PRELL and VOSBERG, unpublished results) have shown that in the presence of increasing amounts of single-stranded DNA, the enzyme is progressively unavailable for the relaxation of superhelical DNA. Instead, it is quantitatively complexed with single-stranded DNA. These results do not imply that single-stranded regions are involved in relaxation by eukaryotic topoisomerases, since relaxed and positively supercoiled DNA molecules are readily accepted as substrates.

From the binding characteristics of ω to different DNA molecules, it was concluded that the formation of a catalytically functional complex between ω and DNA may require local single-stranded sections (LIU and WANG 1979). Such a requirement could explain why negatively supercoiled DNA is accepted as a substrate and why DNA, which does not readily expose local single-stranded regions such as relaxed or positively supercoiled DNA, is a poor substrate or no substrate at all. Topoisomerase I from *H. gallinarum* was recently demonstrated to cleave supercoiled DNA preferentialy at sites which are also recognized by the single-strand specific S1-nuclease (SHISHIDO et al. 1983). This result underlines the suggested correlation between activity and local single-stranded regions.

Inhibition of relaxation by ω in the presence of single-stranded DNA is probably also due to competition between different binding substrates. The details of the interaction between ω and single-stranded DNA are, however, not fully understood. Either the enzyme binds at a high on-rate and a very low off-rate without nicking, or the enzyme both binds and nicks but has an extremely low rate for reclosure and dissociation. In the latter case, one could conclude that ω interacts strongly with both ends of the nick, since neither spontaneous fragmentation of DNA nor interstrand chain transfer have been observed.

With regard to the presently known characteristics of relaxation reactions with either ω-protein or eukaryotic type I topoisomerases, one might tentatively conclude that a strand-passing mode is more compatible with the action of ω than with the action of the eukaryotic type I enzymes. The inability of the latter to span a transient nick in DNA tightly and also their apparent independence of local single-stranded regions make these enzymes intuitively more amenable for swiveling than for strand passing.

Strong evidence in favor of a strand-passing scheme for ω is provided by the catenation and knotting reactions of this enzyme. The only requirement for these reactions is the presence of nicked circular DNA; extensive sequence homology is not needed. Strand passing through intact duplex molecules does not occur. The molecular details of the passage are not known; however, the simplest explanation would be that the enzyme creates a transient interruption opposite the preexisting nick in one of the participating reaction components, although more complicated models are also conceivable. For further discussion on the mechanism of this process see TSE et al. (1980), BROWN and COZZARELLI (1981), and DEAN et al. (1982).

As was mentioned above, catenation with a eukaryotic topoisomerase I has been demonstrated twice (BROWN and COZZARELLI 1981; BADARACCO et al. 1983). However, since others failed to obtain catenated DNA with a similar enzyme (LIU and MILLER 1981), it seems possible that these topoisomerases realize the strand-passing mode under certain conditions, for instance in the presence of spermidine, but not as their standard pathway.

The conditions for a swivel as an alternative to strand passing have not yet been extensively investigated. A first, somewhat elaborate, approach presented by PULLEYBLANK and ELLISON (1982) favors a mechanism that involves stepwise rotation of the duplex, resulting in increments of one in the linking

number per single reaction cycle. This concept implies noncovalent binding of the enzyme to one region in the DNA with "nicking-closing" occurring at another region nearby. Attachment of the enzyme to the noncovalent binding site could provide a "hinge," limiting rotation of the duplex between "nicking" and "closing" (see Fig. 2). A principal advantage of this mechanism is that local melting of the duplex, which would impose an additional energy barrier to the reaction, is not required. Whether the reaction proceeds strictly in steps of one or whether linking number changes also occur in steps of more than one (VOSBERG and VINOGRAD 1975) depends on the strength of binding of the enzyme to the noncovalent recognition site. Theoretically at least, rotational slippage of the enzyme bound to this site is conceivable.

In summary, although a general mode of action of topoisomerases would be attractive, there is no a priori need for a unique model. Some reactions may occur by the strand-passing route, others may utilize a swivel. The currently available data, at least, are not compelling in favor of strand passing as the universal pathway of topoisomerases.

Modulation of Topoisomerase I Activities in Eukaryotes. It has been shown that calf thymus type I topoisomerase can be covalently modified in vitro by poly(ADP)ribosylation. The modifying enzyme, poly(ADP)ribosyl transferase, was also isolated from calf thymus. The modification leads to a loss in relaxing activity (FERRO et al. 1983, 1984; JONGSTRA-BILEN et al. 1983). Poly(ADP)ribosyl transferase is known to modify a number of intracellular proteins, among them some, such as histones, which are located in the nucleus. It is generally believed that these modifications are regulative measures of the cell (MANDEL et al. 1982).

Another covalent modification of topoisomerase I is phosphorylation. A phosphorylated topoisomerase I (molecular weight 110000) was detected in the nuclei of Novikoff ascites cells, HeLa and Namalwa cells, and fetal rat liver, but not in adult rat liver (DURBAN et al. 1981; DURBAN et al. 1983). A nuclear protein kinase which phosphorylates topoisomerase I predominantly at serine residues has also been identified (MILLS et al. 1982). Phosphorylation seems to be accompanied by an increase in topoisomerase activity (DURBAN et al. 1983). In conjunction with the above-mentioned poly(ADP)ribosylation, it appears that nuclei have a variety of means of regulating this enzyme and, thus, chromatin functions that are affected by topoisomerase I activity.

In addition to being covalently modified, topoisomerase I has been shown to be associated with nucleosomes and chromosomal proteins (WEISBROD 1982a; JAVAHERIAN and LIU 1983), in particular with histone H1 and high-mobility group (HMG) proteins. The protein HMG17 is reported to stimulate topoisomerase I in vitro by 10- to 20-fold (quoted by JAVAHERIAN and LIU 1983). Another, thus far unidentified acid soluble 30000-dalton protein with stimulating effect on topoisomerase I has been reported by ROSS et al. (1983). Although no detailed conclusions can be drawn from these observations, they point to intimate interactions between topoisomerase I and other chromosomal proteins in eukaryotic nuclei.

6 Type II Topoisomerases

Type II topoisomerases comprise a diverse group of enzymes which have in common the requirement for ATP, with one exception, and a reaction mechanism involving transient double-strand breaks in DNA. According to present knowledge, only the bacterial gyrases are able to promote negative supercoiling in DNA.

6.1 Prokaryotic Type II Topoisomerases

Within this category of enzymes, two varieties may be distinguished, the DNA gyrases, which have been identified in a number of different bacterial strains, and the topoisomerase II′, which has been found only in *E. coli* cells. The phage T4 coded topoisomerase II will be discussed under viral topoisomerases (Sect. 8).

6.1.1 DNA Gyrases

The principal reaction of these enzymes is the conversion of relaxed DNA into a negatively supercoiled form. Enzymes able to catalyze this type of reaction have thus far been detected in prokaryotes, but not in eukaryotes. The most extensively studied gyrases are those from *E. coli* (for reviews see COZZARELLI 1980a, 1980b; GELLERT 1981a and 1981b) and from *M. luteus* (LIU and WANG 1978a). Gyrases have also been obtained from *Bacillus subtilis* (SUGINO and BOTT 1980; ORR and STAUDENBAUER 1982) and from *Pseudomonas aeruginosa* (MILLER and SCURLOCK 1983).

Physical and Genetic Properties of Gyrase. The DNA gyrase of *E. coli* is composed of two different types of subunits which are called *gyr*A protein and *gyr*B protein. The component molecular weights of these two proteins under denaturing conditions are 105000 and 95000 for *gyr*A and *gyr*B subunits, respectively (SUGINO et al. 1977; MIZUUCHI et al. 1978; HIGGINS et al. 1978). Purification of the *E. coli* enzyme to near homogeneity is possible either by isolating the active complex, which contains the two proteins in stoichiometric amounts (MIZUUCHI et al. 1978), or by purifying the two components separately (SUGINO et al. 1977; HIGGINS et al. 1978). In this case, activity is reconstituted by mixing of the subunits. Purification yields are usually low but they can be significantly improved by chromatography on novobiocin-affinity columns (STAUDENBAUER and ORR 1981; OTTER and COZZARELLI 1983) or by using *E. coli* strains harboring plasmids with the separately cloned genes for the *gyr*A and *gyr*B subunits (OTTER and COZZARELLI 1983). These strains are overproducers for the respective gyrase subunits. From 2 liters of culture (or 12 g of cells), one can recover 25 mg of *gyr*A or 3 mg of *gyr*B protein (MIZUUCHI et al. 1984).

The *gyr*A subunit is the target protein for the antibiotics nalidixic acid and oxolinic acid (GELLERT et al. 1977; SUGINO et al. 1977), whereas *gyr*B pro-

tein is inhibited by novobiocin and coumermycin (GELLERT et al. 1976b). These drugs were originally identified as inhibitors of DNA replication (GOSS et al. 1965; GOSS and COOK 1975; STAUDENBAUER 1975, 1976a; RYAN 1976). In addition, *gyr*B is inhibited by the coumermycin-related compound chlorobiocin (FAIRWEATHER et al. 1980).[1] The two subunits are the products of the *gyr*A and *gyr*B genes, located at 48 and 82 min, respectively, on the *E. coli* chromosome (BACHMANN and LOW 1980). These map positions were originally assigned to the genetic loci mediating resistance to nalidixic acid and coumermycin (HANE and WOOD 1969; RYAN 1976).

The evidence for structural gene assignments of the two proteins is based on combined genetic and biochemical analysis. Subunit A from purified gyrase and radioactively labeled *gyr*A protein are identical in two-dimensional gel electrophoresis (PEEBLES et al. 1978). Furthermore, subunit A from a temperature-sensitive *gyr*A mutant is thermolabile after purification (KREUZER and COZZARELLI 1979).

Biochemical identity has also been demonstrated for the B subunit of gyrase and the gene product of the *gyr*B locus (HANSEN and VAN MEYENBURG 1979). In addition, extracts from a strain that had a conditionally lethal mutation that affected growth and was located in the *gyr*B gene were unable to replicate exogenous plasmid DNA in an extract system and could not maintain the superhelical conformation of this DNA. Both defects could be complemented by the addition of purified *gyr*B protein (ORR and STAUDENBAUER 1981). Finally, gyrase subunits from nalidixic acid or novobiocin-resistant mutants after reconstitution yield enzyme with the expected resistance in vitro (HIGGINS et al. 1978). The mutation for nalidixic acid resistance (*nal*^{r}) therefore affects the *gyr*A locus. One mutation for nalidixic acid resistance was found which is probably located in the *gyr*B gene (YAMAGISHI et al. 1981).

The number of copies of gyrase per *E. coli* cell is not exactly known. For subunit A, about 500 molecules have been estimated per cell, the number for subunit B may be lower by a factor of ten (HIGGINS et al. 1978; BROWN et al. 1979).

The DNA gyrase of *M. luteus* resembles the *E. coli* enzyme very closely (LIU and WANG 1978a). It is also composed of two subunits called A and B (or α and β) with molecular weights of 115000 and 95000, respectively. These two subunits are functionally equivalent to the *E. coli gyr*A and *gyr*B subunits. The subunit composition of the *M. luteus* gyrase is A_2B_2 upon binding to DNA (KLEVAN and WANG 1980). Data for *E. coli* gyrase show that this enzyme is also active as an A_2B_2 tetramer with a native molecular weight of about 400000 (SUGINO et al. 1980). Gyrase of the gram-positive *B. subtilis* resembles in its subunit composition the other known bacterial gyrases (ORR and STAUDENBAUER 1982). The *gyr*A gene of *B. subtilis* has been cloned in *E. coli*. Functional *gyr*A activity encoded by the cloned fragment complements *E. coli gyr*A mutations (LAMPE and BOTT 1984).

1 A comprehensive study of the effects of novobiocin, coumermycin, chlorobiocin, and 18 of their analogs on DNA gyrase of *E. coli* and bacterial growth has been published by HOOPER et al. (1982). In addition, the relationship between structure and inhibiting activity of oxolinic acid and derivatives thereof has recently been investigated (HOEGBERG et al. 1984)

Activities of Gyrase. DNA supercoiling, the principle reaction of this enzyme, requires ATP and Mg^{2+}. K^+ stimulates the reaction, Na^+ is ineffective. Spermidine enhances the reaction rate severalfold (GELLERT et al. 1976a). The ATP requirement for supercoiling can be partially substituted by dATP but not by other nucleotides.

A single gyrase molecule acts catalytically in a processive manner. One molecule produces a linking difference of about 100/min at 30° C (HIGGINS et al. 1978). Since supercoils are introduced in steps of two, this figure reflects a turnover number for the basic reaction cycle of about one per second. The degree of supercoiling which is achieved in vitro may exceed that found in vivo by a factor of about 1.5, as has been shown with ColE1 DNA (GELLERT et al. 1976a). The upper limit for the specific linking difference which is obtainable with gyrase in vitro is about -0.1. In vivo values for DNA supercoiling in *E. coli* are around -0.06. The difference between these values may be due to a limited action of gyrase or, more likely, to the counterproductive activity of relaxing topoisomerases in vivo. Negative turns can be introduced into both positively supercoiled and relaxed DNA (PEEBLES et al. 1978; BROWN et al. 1979).

The two groups of inhibiting antibiotics, nalidixic acid and oxolinic acid on the one hand and the coumarin derivatives novobiocin, coumermycin, and chlorobiocin on the other, have different mechanisms of action. The latter drugs primarily affect those steps that require ATP, whereas oxolinic acid and nalidixic acid distinctively block all activities which are involved in the breaking and rejoining of DNA. Oxolinic acid is required in fairly high concentrations (ca. 4×10^{-5} M) for 50% inhibition of supercoiling. Nalidixic acid is a rather inefficient inhibitor of gyrase activity in vitro (SUGINO et al. 1977; GELLERT et al. 1977). The enzyme is furthermore inhibited by N-ethylmaleimide, suggesting the participation of a sulfhydryl group in the catalytic cycle (GELLERT et al. 1976a). Inactivation of gyrase is also achieved by a chemical modification of tyrosyl residues with tetranitromethane (KLEVAN and TSE 1983). This modification can be prevented by preincubation of the enzyme with DNA. Presumably, the cleavage and rejoining site of subunit A is protected from modification if the enzyme is bound to DNA.

Among the many facets, which the supercoiling reaction of gyrase has, three are particularly intriguing: the breaking and rejoining of DNA, the disposition of energy needed for the generation of negatively superhelical turns, and the control over the direction of strand passage during the raction. In general, both subunits are required for all reactions of gyrase, and they are evidently in close contact during DNA supercoiling (BOGDANOVA et al. 1982). However, the specific roles of the two components in the reaction cycle can, to a certain extent, be discriminated. The reactions carried out by gyrase are summarized in Table 3. In the following sections, the separate subactivities of the enzyme will be treated first and the reaction mechanisms will be discussed subsequently.

Binding of Gyrase to DNA. The native enzyme forms a noncovalent and quite stable complex with DNA at nonrandom sites. Since subunit A is alone able to bind to DNA (STAUDENBAUER and ORR 1981; KLEVAN and TSE 1983; MOORE

Table 3. Activities of bacterial DNA gyrase

Type of reaction[a]	ATP requirement	Inhibition[b] by		Remarks
		Oxo or Nal	Novo or Cou	
1. Generation of negative supercoils	Yes	+	+	Reaction is processive
2. Binding to DNA	No	−	−	Specific sequences are recognized
3. Cleavage of DNA[c]	No[d]	−	−	Cleavage preferably involves a TG dinucleotide; staggered double-stranded breaks are produced
4. Catenation and unknotting	Yes	+	+	Sequence homology for catenation not required
5. Relaxation	No	+	−	Positive supercoils are relaxed in the presence of β,γ-imido ATP
6. ATP hydrolysis[e]	Yes	−	+	

[a] Gyrase activities generally require the presence of both subunits
[b] The gyrase inhibitors are: nalidixic acid, oxolinic acid, novobiocin, coumermycin
[c] Oxolinic acid and denaturing agents are required to detect cleavage
[d] ATP is not required, but in its presence other sequences may preferentially be cut than in its absence
[e] DNA is an effector of this reaction. Double-stranded DNA is about ten times more effective than single-stranded DNA. Subunit B catalyzes ATP hydrolysis also in the absence of subunit A (STAUDENBAUER and ORR 1981)

et al. 1983; LOTHER et al. 1984), this subunit may be considered as the DNA binding protein in the gyrase complex. Most of the gyrase molecules that are bound noncovalently are also able to cleave the DNA (MORRISON et al. 1980a). Redistribution of DNA-bound enzyme is highly unlikely since no exchange with competing free DNA has been observed (MORRISON et al. 1980a). The most plausible picture is, therefore, one of binding of enzyme at those sites where DNA is cleaved in the course of the subsequent reaction. The strength of binding depends on the tertiary structure of the DNA. Gyrase binds less tightly to negatively supercoiled DNA than to relaxed DNA (quoted in COZZARELLI 1980a). This result implies that the supercoiling reaction is limited by the superhelicity of its product.

Three different modes of binding of gyrase to DNA have been distinguished by means of retention of DNA-gyrase complexes on nitrocellulose filters (HIGGINS and COZZARELLI 1982). One complex consists of an A_2B_2 tetrameric enzyme bound noncovalently to specific sequences of DNA. At 23° C, a dissociation constant of about 10^{-10} M and a half-time of approximately 60 h have been estimated. A second complex is formed upon addition of oxolinic acid. Stability in high ionic strength solutions is characteristic for this type of complex. Both the noncovalent and salt-stable complexes require Mg^{2+} for optimal formation.

The second complex is converted into a third form on addition of protein denaturants. This denatured complex has been shown to contain A subunits covalently linked to DNA, which is interrupted at the attachment site.

Binding of gyrase of *M. luteus* and of *E. coli* to DNA protects sequences of approximately 140 base pairs in length from digestion by staphylococcal nuclease (LIU and WANG 1978b; GELLERT et al. 1980; FISHER et al. 1981; MORRISON and COZZARELLI 1981). If the DNA gyrase complex of the *M. luteus* enzyme is digested with pancreatic DNase I, a series of fragments is generated which differ in length by 10–11 bases (LIU and WANG 1978b; KIRKEGAARD and WANG 1981; FISHER et al. 1981). Pairs of DNase I-sensitive cleavage sites on complementary strands are typically staggered at two to four base pairs. This result is reminiscent of similar results with nucleosomes where DNA is known to be wrapped around a core of proteins (see MCGHEE and FELSENFELD 1980). It has further been shown that breakage and rejoining of the DNA by gyrase occurs within a central region of the staphylococcal nuclease-resistant segment of the DNA. ATP or the nonhydrolyzable analog β,γ-imido ATP do not alter this pattern, which implies that the DNA-enzyme topography as it is mapped by the nuclease is essentially not affected by ATP-induced conformational changes.

Additional evidence for the wrapping of DNA is obtained from the binding of *M. luteus* gyrase to nicked circular DNA and subsequent sealing of the DNA with DNA ligase (LIU and WANG 1978a). Upon removal of the gyrase the DNA was found to be wound, corresponding to a linking difference of $+1$ per enzyme tetramer. Thus, binding of DNA to the enzyme causes the formation of a positive supercoil. The simplest interpretation of these data is that the DNA wraps a full turn around the tetrameric gyrase complex with the appropriate handedness. [The handedness cannot be determined a priori; for a discussion see CRICK (1976) and LIU and WANG (1978a).]

Positively supercoiled plasmid DNA has recently also been obtained from living cells after treatment of *E. coli* bacteria with gyrase inhibitors (LOCKSHON and MORRIS 1983). Most probably, these positively supercoiled DNA molecules were formed in vivo essentially as they are made in vitro: gyrase is bound to DNA without being catalytically active. Concomitant relaxation of DNA by other topoisomerases leads to positively supercoiled DNA upon removal of gyrase.

Cleavage of DNA. Breaks in the DNA induced by *E. coli* gyrase become apparent by the addition of SDS to oxolinic-acid-treated DNA-enzyme complexes (GELLERT et al. 1977; MIZUUCHI et al. 1978; FISHER et al. 1981). Similar breakage by *M. luteus* gyrase has been demonstrated by treating such complexes with alkali instead of detergent (TSE et al. 1980). It is not known in detail how oxolinic acid interferes with the action of gyrase. Presumably, in the presence of this antibiotic, DNA is cleaved but not resealed by the *gyr*A subunit. Subsequent treatment with denaturing agents disrupts the protein complex and releases the fragments.

The breakpoints always have the same structure. The DNA is cut in a staggered fashion 4 base pairs apart (MORRISON and COZZARELLI 1979; GELLERT

et al. 1980; KIRKEGAARD and WANG 1981; FISHER et al. 1981). The 5′ ends protrude and are blocked by the covalent association of gyrase A subunit molecules. The 3′ ends have a free hydroxyl group and, thus, they are able to prime a DNA polymerase reaction (MORRISON and COZZARELLI 1979). The linkage between *gyr*A and DNA is a phosphotyrosine bond, which has also been identified for pro- and eukaryotic type I topoisomerases (TSE et al. 1980; CHAMPOUX 1981).

Cleavage by DNA gyrase is, like DNA binding, site specific. Strongly preferred cleavage sites are, however, rare. ColE1 DNA, for instance, (molecular weight 4.2×10^6) is normally cleaved at a single site (MORRISON and COZZARELLI 1979). Additional sites with lower and widely varying cleavage frequencies occur also so that DNA is potentially cleaved once in about 100 base pairs.

Nucleotide sequencing of several cleavage sites has not exhibited a clearly defined sequence pattern. In a series of cleavage sites, the dinucleotide TpG was commonly found in one of the strands with cleavage occurring between these two nucleotides (MORRISON and COZZARELLI 1979; MORRISON et al. 1980b). However, in other studies this result was not obtained (GELLERT et al. 1980; KIRKEGAARD and WANG 1981; FISHER et al. 1981). It is, therefore, not yet definite what the structural basis is for the site specificity of DNA gyrase.

Most of the data on cleavage by gyrases have been derived from analyses that involved the use of oxolinic acid. It was pointed out by COZZARELLI (1980a) that this drug does not create artificial cleavage sites but merely enhances the chance of their detection by preventing rejoining. Unperturbed reactions disclose the same sites although at a low frequency.

DNA gyrases from three different species, *E. coli*, *M. luteus*, and *B. subtilis*, produce very similar sets of fragments with a given DNA, suggesting similar recognition behavior of these enzymes (SUGINO and BOTT 1980).

Relaxation. In addition to its supercoiling reaction, gyrase catalyzes a number of other topological reactions with duplex DNA. It is able to relax negative supercoils if ATP is absent (SUGINO et al. 1977; GELLERT et al. 1977). Positive supercoils are not relaxed under these conditions. The rate of relaxation is lower than the rate of supercoiling by a factor of 20–40. This result indicates that relaxation is a minor reaction of the enzyme and probably without physiological significance. Relaxation is inhibited by oxolinic acid, but not by novobiocin or coumermycin. Thus, relaxation specifically requires subunit A with its ability to break and rejoin DNA, but not the energy-coupling reaction of subunit B. The increase in the linking number of the DNA during relaxation occurs in steps of two (BROWN and COZZARELLI 1979). This result demonstrates that relaxation and supercoiling proceed by the same mechanistic route, although in opposite directions.

In the presence of β,γ-imido ATP, gyrase is able to relax positively supercoiled DNA in a catalytic fashion at a rate which is similar to the relaxation of negatively supercoiled DNA by this enzyme (GELLERT et al. 1980).

Catenation, Decatenation, and Resolution of Topological Knots. Gyrase is able to catalyze the formation and separation of catenated DNA rings (KREUZER and COZZARELLI 1980; MIZUUCHI et al. 1980). The enzyme also resolves intra-

molecular topological knots in double-stranded DNA rings (LIU et al. 1980; MIZUUCHI et al. 1980). All of these reactions require ATP and are inhibited by novobiocin and oxolinic acid. Spermidine is needed for the formation of catenated rings. The function of this polycation is to maintain a high local concentration of DNA molecules (KRASNOW and COZZARELLI 1982). The rates of catenation reactions are comparable to the rate of supercoiling (COZZARELLI 1980b). DNA sequence homology is not required for catenation. Negatively supercoiled DNA is more readily catenated than nicked circular DNA.

These results have a bearing on the mechanism of gyrase action by providing clear evidence that gyrase acts by passing one DNA segment through a double-strand break of another segment of the same or of a separate molecule. Other explanations which could theoretically be based on the notion of a "progressive invasion" of one ring by another one via successive single-strand nicks are generally considered highly unlikely.

Role of ATP in Gyrase Action. The conditions of ATP hydrolysis will be discussed before an interpretation of the overall function of ATP is presented.

Duplex DNA is about ten times more effective in stimulating ATP hydrolysis to ADP and P_i than single-stranded DNA (SUGINO and COZZARELLI 1980). Synthetic double-stranded polydeoxynucleotide homopolymers also stimulate ATP cleavage. This result indicates that specific DNA sequences are not involved in this hydrolysis. Duplex DNA presumably stabilizes an enzyme conformation that promotes ATP hydrolysis. Since ATP hydrolysis is much easier to measure quantitatively than supercoiling, the only exact kinetic data on gyrase action pertain to this function of the enzyme.

The apparent K_m of gyrase for ATP was found to be 0.32 mM (SUGINO et al. 1978) or 0.5 mM (STAUDENBAUER and ORR 1981). Both values were determined in the presence of duplex DNA. In its absence, ATP is also hydrolyzed, but with a markedly higher K_m of about 1.7 mM (STAUDENBAUER and ORR 1981).

ATP hydrolysis is not necessarily coupled to supercoil formation. This conclusion is based on the observation that maximally supercoiled DNA ($\Delta\alpha/\alpha_0 = -0.10$) or DNA effectors that cannot be supercoiled trigger ATP cleavage. The number of superhelical turns introduced into an initially relaxed circular DNA was calculated to be approximately equivalent to the number of ATP molecules hydrolyzed in the reaction (SUGINO and COZZARELLI 1980).

As already mentioned, ATP interacts with the B subunit of gyrase. Early evidence for this conclusion was based on kinetic studies, which indicated that novobiocin and coumermycin both prevent ATP binding to the *gyr*B subunit in a competitive manner (SUGINO et al. 1978). It was further shown that the B subunit is covalently labeled with an ATP analog, the 2′,3′-dialdehyde ATP (oATP) (EASTERBROOK-SMITH et al. 1976). Novobiocin is a competitive inhibitor of this reaction (MIZUUCHI et al. 1978). The most direct evidence for the B subunit being the target of ATP is the finding that the *gyr*B subunit itself possesses novobiocin-sensitive ATPase activity (STAUDENBAUER and ORR 1981).

Gyrase binding to DNA causes breakage of the DNA upon the interference of oxolinic acid and sodium dodecylsulfate (or alkali). With a given DNA mole-

cule, the pattern of cleavage fragments can be changed by ATP or β,γ-imido ATP (SUGINO et al. 1978; FISHER et al. 1981), i.e., the presence or absence of ATP or its imido analog affects which sequences are preferentially cut – at least at some sites of gyrase action. This change by either ATP or its analog can be blocked by novobiocin, indicating that this drug interferes with ATP binding to gyrase rather than with ATP hydrolysis (SUGINO et al. 1978).

Inhibition constants (K_i) for the effect of novobiocin and coumermycin on ATP hydrolysis are quite low. Reported values are 1×10^{-8} M and 4×10^{-9} M for the two drugs, respectively (SUGINO et al. 1978). STAUDENBAUER and ORR (1981) found higher values of 3.5×10^{-7} M for novobiocin (and clorobiocin) and 8×10^{-8} M for coumermycin. Although the absolute values differ in these reports, the relative order of inhibitory strength is the same for these two drugs. Coumermycin is more effective than novobiocin.

ATP effects on gyrase activities can be summarized as follows: (a) ATP is required for supercoiling and all other topological changes of the DNA except for relaxation; (b) ATP is hydrolyzed to ADP and P_i by subunit B of the enzyme; novobiocin and coumermycin are competitive inhibitors which prevent binding of ATP to subunit B; (c) The nonhydrolyzable ATP analog β,γ-imido ATP permits a limited action of gyrase by promoting, on average, one cycle of supercoiling per enzyme molecule; (d) ATP is not required for the binding of gyrase to DNA and also not for cleaving and reclosing of DNA.

In combining the available data, the following picture about the role of ATP in the supercoiling process can be drawn. The association of gyrase with its primary binding site on DNA occurs in the absence of ATP. This step involves the wrapping of DNA around the enzyme. (The resulting DNA-gyrase complex will be called the "strand-passing gate" for the purpose of the following description.) According to recent results obtained in the electron microscope (MOORE et al. 1983), the strand-passing gate is presumably able to attract a distant DNA segment so that a ternary complex is formed which constitutes a DNA-DNA crossover at the site of the strand-passing gate. Formation of this complex does not require ATP, but in the absence of the nucleotide it is catalytically unproductive. The cleavage and reclosing event per se also does not depend on ATP, as is concluded from the fact that relaxation of negatively supercoiled DNA, in contrast to supercoiling, does occur in the absence of ATP.

In the presence of the nonhydrolyzable analog β,γ-imido ATP, one full round of supercoiling takes place, but then the reaction stops. According to MOORE et al. (1983), the DNA segment that is translocated in the presence of the analog is apparently not held in close contact with the strand-passing gate at the end of the reaction. Thus, dissociation of the translocated DNA segment from the gate after one catalytic cycle most likely does not require ATP cleavage.

From these results, it can be concluded that the critical event which depends on ATP binding is transport of the intact DNA segment through the strand-passing gate. ATP cleavage is conceivably needed for turning over of the enzyme so that it is able to attract the same or another distant DNA segment for a new round of supercoiling. ATP presumably exerts its effect by controlling the conformation of gyrase. At least two conformational transitions may be

postulated. The first one would be expected early in the reaction cycle when an additional DNA segment associated with the strand-passing gate is translocated only when ATP is also bound. The second transition would require ATP hydrolysis and dissociation of ADP in order to restore that conformation of the enzyme which is able to bind remote DNA sections so that a new productive DNA-DNA crossover can be formed.

The mechanism of relaxation of negatively supercoiled DNA in the absence of ATP is not entirely clear. The rate of relaxation is low; about 20 times more gyrase is required for relaxation than for supercoiling a given amount of DNA. Two facets of enzyme action may explain this result. First, gyrase binds more weakly to superhelical than to relaxed DNA. Second, for the initiation of supercoiling, gyrase actively brings distant DNA segments into proximity at the entry side of the strand-passing gate – presumably by specific DNA-binding affinities of the enzyme. For relaxation, formation of a DNA-DNA crossover is required at the exit side of the gate. This crossover is probably not stabilized by gyrase and relies, therefore, on random motion of different segments of a DNA molecule relative to each other. Thus, the occurrence of a spatial arrangement of DNA and gyrase appropriate for the relaxation reaction may be a rare event in which ATP has no role.

The distinctive behavior of the enzyme in supercoiling and relaxation reactions illustrates that gyrase has apparently tight control over the direction of strand passage. The supercoiling reaction is a one-way reaction, which is also true for the relaxation reaction. The preference for negatively supercoiled DNA in the absence of ATP can be switched in the presence of the uncleavable analog so that relaxation of positively supercoiled DNA is stimulated. It may be assumed that different conformations of these enzymes are associated with distinct binding preferences for DNA molecules containing supercoils of opposite handedness. This discriminative behavior is a unique property of the prokaryotic gyrase [and possibly also of the reverse gyrase which has been detected in the archebacterium *Sulfolobus* (Sect. 6.1.3)]. The topoisomerase II′ of *E. coli*, which is presumably a gyrase with a truncated fragment B (see the following section), relaxes both positively and negatively superhelical DNA molecules. It has lost the ability to discriminate between different conformations. The phage T4 type II topoisomerase and the eukaryotic type II topoisomerases, which both have no supercoiling activity, also do not distinguish between positive and negative supercoiling.

Models of the Mechanism. Although many details are known about the reaction requirements of gyrase and the nature of the products of its reaction, the molecular mechanics of the process by which supercoiling is achieved is not yet understood. Any explanation of this mechanism has to accommodate the duplex strand-passing character of the reaction which follows from the change in the winding number in units of two and, obviously, from the catenation reaction.

On the basis of known experimental facts, various concepts have been worked out to explain how gyrase might act (Brown and Cozzarelli 1979; Forterre 1980; Liu et al. 1980; Gellert et al. 1980; Wang et al. 1980; Morri-

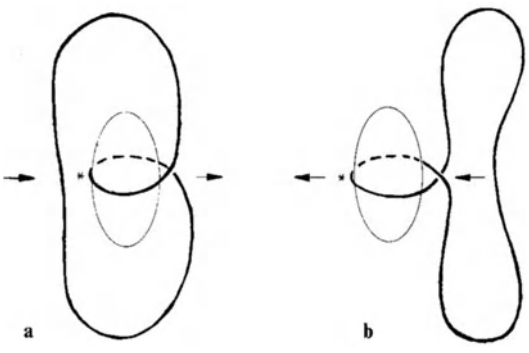


Fig. 3a, b. Model for the supercoiling activity of DNA gyrase. A basic feature of this model is formation of a DNA loop which results from DNA wrapping around the gyrase molecule. It is assumed that both the multimeric gyrase molecule and the DNA loop participate in the formation of a gate (called here the "strand-passing gate") which mediates the passage of an additional DNA segment through a transient break in the looped DNA. The position of the transient break is marked by an *asterisk*. The transport of the intact DNA segment may either occur **a** from outside of the loop to the inside or **b** in the opposite direction. The two possible directions of strand movement are indicated by the *arrows*. For the outside-inside transfer, DNA wrapping at the region of the strand-passing gate would be right-handed in order to achieve a decrease in the linking number by the strand-passing event. A transfer in the opposite direction would require left-handed wrapping for the same result. (For a treatment of the handedness problem see CRICK 1976)

SON et al. 1980b; WANG 1982a, b). Basically, two different models can be distinguished among the various attempts to predict the details of how DNA strands are translocated through molecular gates and how polarity or handedness of DNA binding to the enzyme is controlled. The first model is called the "sign inversion" model (BROWN and COZZARELLI 1979). This model involves juxtaposition of disparate DNA sections by the enzyme. The interaction between these DNA segments and the enzyme results in the formation of a crossover or a node, which has a positive superhelical sign. The enzyme would function as some kind of flexible allosteric three-dimensional matrix on the surface of which the whole process of sign inversion occurs. Binding of ATP is thought to provide the signal for transfer of an intact DNA segment through the DNA break at the site where primary binding of the enzyme occurs. One could envisage this passage as a translocation of the intact segment through the break. A consequence of this chain of breaking, passing, and rejoining would be a node of opposite handedness, hence the notion of sign inversion.

A model that is somewhat more predictive in its molecular details has been proposed by Wang and his colleagues (WANG et al. 1980; WANG 1982a, b). According to one version of this model (Fig. 3), DNA is wrapped around the enzyme in a right-handed or positive fashion. This wrapping essentially does not change during the reaction. Instead, the DNA segment to be translocated is passed through a transient break in the region of primary binding of gyrase into the interior of the complex designated above as the strand-passing gate,

which is constituted by the gyrase protein with DNA wrapped around it. The breakpoint in the DNA could be considered as the entry site of the strand-passing gate. The transfer of a distant DNA segment through this gate requires ATP, and the turnover of the enzyme after one cycle depends on ATP hydrolysis, as described above. Since the active gyrase is a tetrameric protein of A_2B_2 subunit composition, it could accommodate such a complicated process. The overall result of this reaction would be, as in the sign inversion model, a change in the linking number in steps of two. According to this version, the translocation of DNA would occur from the outside of the lock to the inside. As schematically depicted in Fig. 3, the opposite direction from inside to outside is also conceivable. In this case, however, wrapping of the DNA would be left-handed (CRICK 1976; LIU and WANG 1979).

Not much information is available as yet to choose between these proposals. Some results which bear on the gyrase mechanism were recently communicated by MOORE et al. (1983). These authors observed, as quoted above, in the electron microscope that in the absence of ATP gyrase is preferentially found at intramolecular DNA-DNA crossovers and that the number of enzyme molecules at these crossovers is significantly reduced if β,γ-imido ATP is present. As a simple interpretation of these findings, it was tentatively suggested that the crossovers in the absence of ATP represent an early state in the reaction cycle with the intact DNA segment placed in close proximity to the strand-passing gate inside the DNA loop that is wrapped around the enzyme. In the presence of β,γ-imido ATP, after one cycle of supercoiling the translocated segment would be outside the loop and no longer attached to the enzyme.

Although these results are no proof, they tend to favor a transport from inside to outside and at the same time left-handed wrapping of DNA around gyrase. (For a complete discussion of these results see MOORE et al. 1983.) Further direct probing of DNA-gyrase complexes will be required for a complete understanding of the mechanism of this enzyme.

6.1.2 DNA Topoisomerase II′

In addition to the normal A and B subunits of gyrase, purification of a protein from *E. coli* cells has been reported that complements with the A subunit to generate a complex, topoisomerase II′, which has a topoisomerase-related activity (BROWN et al. 1979; GELLERT et al. 1979). This protein, designated *v* in one of the reports, has a molecular weight of 50000. The number of *v* protein molecules per cell has been estimated to approximate that of the A subunit and to exceed that of the B subunit, roughly by a factor of ten.

Topoisomerase II′ resembles gyrase in a number of properties. It relaxes negative supercoils, but also positive supercoils without requirement for β,γ-imido ATP. Treatment with oxolinic acid causes double-strand breaks in DNA at gyrase-specific sites, and interaction of DNA involves formation of a local positive supercoil. Linking numbers are altered in steps of two (MORRISON et al. 1980b), demonstrating that the enzyme acts by a strand-passing mechanism

via transient double-strand breaks. This mechanism classifies the topoisomerase II′ as a type II topoisomerase.

Other properties are unlike those of gyrase. Most importantly, this enzyme does not induce negative supercoiling. ATP is not required for its limited activities, and ATP is not hydrolyzed. The gyrase inhibitor novobiocin does not affect this enzyme, in contrast to oxolinic acid, which is inhibitory.

When gyrase B subunit and *v* protein were subjected to limited proteolytic treatment, very similar or identical subfragments were obtained (GELLERT et al. 1979). This result suggests common primary structures of the two proteins, implying that *v* is a subfragment of B. The combined data on the catalytic activity of topoisomerase II′ and on the structural analysis suggest that the *v* protein has lost the ATP-binding domain responsible for energy transduction or enzyme turnover in the gyrase B subunit.

The nature of this protein is not yet definitely explained. Although antiproteolytic agents were included in purification buffers, the authors do not rigorously exclude proteolytic effects during isolation of this protein. The alternative possibility would be that the *v* protein really exists in living cells and comprises, in conjunction with the subunit A of gyrase, an additional topoisomerase in *E. coli* cells. If so, then the question arises how this B fragment is made. It could result from the B subunit of gyrase by proteolytic processing, but theoretically it could also be produced by differential transcription of the *gyr*B gene or differential translation of the *gyr*B mRNA. However, on the basis of the presently available data, all attempts to evaluate nature, mode of production, and possible biological role of this protein are preliminary.

6.1.3 Reverse Gyrase

A somewhat surprising discovery was recently made with the identification of a gyrase activity which introduces positively supercoiled turns into DNA (KIKUCHI and ASAI 1984). The authors called the enzyme which promotes this reaction reverse gyrase. It was found together with other topoisomerase activities in partially purified protein fractions of *Sulfolobus*, an acidothermophilic archebacterium. The enzyme is active at temperatures above 55° C and requires Mg^{2+} and ATP. Introduction of superhelical turns occurs in steps of two; hence, this enzyme is a type II topoisomerase. Other reactions catalyzed by this enzyme are relaxation of negative supercoils and also unknotting and catenation. The molecular weight of the native enzyme is about 500000.

At this early stage, one can only speculate about the reaction mechanism of this enzyme. A two-component system could be conceived, consisting of a putative binding protein able to promote a nucleosome-like structure with positive wrapping of the DNA around a protein core and a relaxing activity which removes compensating negative supercoiling in other regions of the DNA. If such a mechanism were operating, one could expect that free, unrestrained positive supercoiling would not be observed before the protein components are removed from the DNA. Since, however, a competing relaxing activity was able to release positive supercoiling in the presence of reverse gyrase, it

was concluded that such a two-component system is not very likely. Another possibility is the conversion of a normal gyrase into a reverse gyrase by association of an effector molecule, e.g., an additional protein subunit, which changes the direction of strand passage. It is, of course, also conceivable that reverse gyrase is a unique enzyme which developed independently of other topoisomerases.

Since positively supercoiled DNA is not observed in pro- or eukaryotic organisms, the question arises of what the function is of this enzyme in *Sulfolobus*. One simple, though attractive explanation would be that positive supercoiling protects the DNA of archebacteria from unscheduled denaturation, which could easily occur at the high temperatures to which these organisms are adapted. In this case, other archebacteria should also have this enzyme, which has as yet not been reported. The authors (KIKUCHI and ASAI 1984) suggested, therefore, a far more general significance of reverse gyrase. They proposed that this enzyme might be involved in the reversal of chromatin assembly by introducing a transient positive torsional constraint into internucleosomal sections of DNA, with the consequence that histones dissociate from nucleosomes, for instance in temporal and regional connection with gene expression events. Although this suggestion cannot be dismissed a priori, there is so far no evidence in support of a reverse gyrase in normal pro- or eukaryotic cells.

6.2 Eukaryotic Type II Topoisomerases

Several reports exist on ATP-dependent topoisomerases from a variety of eukaryotic sources. The first ATP-dependent in vitro catalyzed catenation of circular DNA was actually demonstrated with a cell extract from *X. laevis* oocytes (GANDINI-ATTARDI et al. 1976). Other early reports about similar activities in various higher organisms came from LIU et al. (1980). Enzymes of this type were subsequently obtained from *X. laevis* germinal vesicles (BALDI et al. 1980), *Drosophila* embryos (HSIEH and BRUTLAG 1980), HeLa cells (MILLER et al. 1981), yeast (GOTO and WANG 1982; GOTO et al. 1984), rat liver (DUGUET et al. 1983), and calf thymus (DARBY and VOSBERG 1985).

MILLER et al. (1981) reported the purification of about 4 μg enzyme from 20 g HeLa cells. This quantity amounted to less than 1% of the initial activity that was estimated to be present in the supernatant of a nuclear extract after removal of DNA. In considering the yield, the increase in purification by a factor of not more that 500, and the apparent homogeneity of the purified enzyme, one may estimate that it represents about 0.1% of the total nuclear protein. This value would amount to a copy number of roughly 100000 enzyme molecules or more per nucleus. This number is clearly below that estimated for topoisomerase I (up to 1×10^6 per nucleus), but not by more than a factor of ten. Thus, the number of topoisomerase II molecules per nucleus is, as that of topoisomerase I molecules, probably also rather high.

Purification of topoisomerase II in the presence of topoisomerase I was greatly facilitated by the availability of a DNA substrate that reacts only with topoisomerase II but not with topoisomerase I, i.e., mitochondrial DNA (kine-

toplast DNA) of *Crithidia fasciculata*, an insect trypanosomatid. This DNA exists in the form of networks of up to about 5000 interlocked covalently closed minicircles which are 2.5 kb long (see Sect. 7). Type II topoisomerases, in contrast to type I topoisomerases, decatenate these networks.

The biochemical and catalytic properties of these topoisomerases have been most thoroughly studied with the enzymes of HeLa cells (MILLER et al. 1981), yeast (GOTO and WANG 1982; GOTO et al. 1984), *D. melanogaster* (HSIEH 1983a; SANDER and HSIEH 1983; SHELTON et al. 1983; OSHEROFF et al. 1983), and *X. laevis* (BENEDETTI et al. 1983). The data obtained by the various authors converge to the picture that this topoisomerase exists in its native form as a homodimer of two subunits. The molecular weight of the subunit is between 166000 and 175000. The homodimer sediments with about 9.2 *S*. Similar to eukaryotic topoisomerase I, size variants of the subunits with smaller molecular weight have been reported (SHELTON et al. 1983; SANDER and HSIEH 1983) which participate in the topoisomerase cleavage reaction and form covalent complexes with the cleaved DNA (SANDER and HSIEH 1983).

Eukaryotic type II topoisomerases catalyze in an ATP-dependent fashion relaxation of negative and positive supercoils, knotting, unknotting, catenation, and decatenation of DNA circles. Supercoiling activity has not been detected. Linking number changes in DNA molecules occur in steps of two. A careful quantitative analysis of the relaxation reaction has been performed with the topoisomerase II from *D. melanogaster* (OSHEROFF et al. 1983). The relaxation is under conditions of maximal activity processive. Binding of ATP is required for the strand-passage event and hydrolysis of ATP for turning over of the enzyme. Removal of one superhelical turn is accompanied by hydrolysis of about four molecules of ATP. dATP but no other nucleotide replaces ATP to some extent. The corresponding apparent K_m values are 280 μM for ATP and 630 μM for dATP. Kinetic and inhibition studies indicated that the β and γ phosphate groups, the 2′-OH of the ribose, and the C_6-NH_2 of the adenine ring are required for the interaction of ATP with the enzyme. The antibiotics that inhibit prokaryotic gyrase very effectively, coumermycin A, novobiocin, oxolinic acid, and nalidixic acid, are also inhibitory for eukaryotic topoisomerase II, with coumermycin A being most and nalidixic acid being least effective. However, drug concentrations required to observe inhibition with the eukaryotic enzymes are much higher than those needed to inhibit gyrase. Relaxation is strongly inhibited by single-stranded DNA.

The processive mode of the reaction can be shifted to a distributive mode under three different conditions: high ionic strength ($\geqq$175 mM), high magnesium concentration ($\geqq$15 mM), and high pH ($\geqq$10) in glycine buffer. Although it has not been demonstrated, it seems reasonable to imply that under these conditions the ratio of the rate of relaxation to that of dissociation of the enzyme-DNA complex is shifted in favor of the latter (OSHEROFF et al. 1983).

Similar to gyrase, cleavage of DNA occurs in a staggered fashion with 5′-protruding ends, which are four nucleotides long. The 5′-phosphoryl end is covalently linked to the enzyme, the 3′ ends have free OH-groups. Cleavage is clearly nonrandom; it is, however, not sequence homology that determines cleavage specificity (SANDER and HSIEH 1983; LIU et al. 1983).

LIU and co-workers (1983) have found conditions under which the cleavage reaction (with the topoisomerase II of calf thymus) is apparently reversible. This reaction requires high enzyme concentration, addition of high concentrations of NaCl (up to 0.5 *M*), and stopping of the reaction by SDS. Under these conditions, both nicked form II and linear form III plasmid DNA which resulted from cleavage by the enzyme were converted back to superhelical form I DNA. No ATP is needed for the reaction. Inhibition of the catalytic activity of the enzyme by N-ethylmaleimide also does not impair this reversible cleavage. This result has a bearing on the details of the mechanism of the topoisomerase II action. It may suggest that the two broken DNA ends in the putative intermediate DNA topoisomerase complex are held tightly by the protein. However, it can at present not be excluded that cleavage occurs as a consequence of the application of SDS. If this were true, then the appearance of form I DNA at high salt concentrations would simply reflect ready dissociation of noncovalent enzyme-DNA complexes before SDS could be effective.

Some evidence has been reported indicating that eukaryotic topoisomerase II may, at least for some of its reactions or under certain conditions in vivo, cooperate with other proteins. According to an earlier report, the *Drosophila* enzyme requires an additional protein component for its catenation reaction. This protein is not identical to, but may be replaced by, histone H1 (HSIEH and BRUTLAG 1980). H1 is known to condense DNA and chromatin into compact structures (see, e.g., RENZ et al. 1977; KELLER et al. 1977; HSIANG and COLE 1977; MÜLLER et al. 1978). It also binds preferentially to superhelical DNA (VOGEL and SINGER 1976). Thus, the stimulation of catenation may be due to an increased local concentration of DNA strands, similar to effects of spermidine on catenation described by others (KRASNOW and COZZARELLI 1982). HSIEH (1983b) reported that the topoisomerase in high concentrations may itself act by condensing DNA. Catenation by the yeast topoisomerase II has also been shown to depend on a nuclear protein in addition to the topoisomerase (GOTO and WANG 1982).

Additional evidence for cooperation between topoisomerase II and other nuclear proteins is the DNA network formation promoted by the HMG17 protein, an effect similar to that observed with histone H1 (TSE et al. 1984). This result is, however, not specific for topoisomerase II; it is also obtained with pro- and eukaryotic type I topoisomerases.

Topoisomerase II of calf thymus can in vitro be modified by poly(ADP)ribosylation. The modification inhibits the activity (DARBY et al., to be published). This result is similar to what is known about topoisomerase I from the same source.

Much attention has been focussed recently on the effects which certain antitumor drugs exert on the activities of eukaryotic type II topoisomerases. Such studies have a bearing on the detailed understanding of topoisomerase reactions and functions as well as on the understanding of the action mechanism of these drugs in living organisms. It has been known for quite some time that a large number of antitumor drugs interact with DNA in vitro (see WARING 1981), many of them by an intercalative mode. On the basis of experimental evidence that had emerged by then, FILIPSKI (1983) hypothesized that competi-

tive inhibition of topoisomerases by DNA intercalators and similar compounds may explain a number of biological effects including inhibition of macromolecular synthesis, frameshift mutations, and various recombination events. The author implied that competition somehow affects the closing step of the topoisomerase reaction.

For some of the drugs in question, most noticeably for the weakly intercalating acridine derivative 4′-(9-acridinylamino)methansulfon-m-anisidide (m-AMSA) (NELSON et al. 1984; TEWEY et al. 1984) and ellipticine and derivatives thereof (DOUC-RASY et al. 1983; TEWEY et al. 1984), a biochemical analysis of their effects on mammalian topoisomerase II has been presented in the meantime.

The drug m-AMSA, but not its isomer o-AMSA, is a potent antitumor agent characterized by the ability to induce DNA strand breaks in mammalian cells with proteins tightly bound to the ends of the broken DNA (see, e.g., ROSS and BRADLEY 1981; ZWELLING et al. 1981). It had been suggested earlier that a DNA topoisomerase may be involved in the formation of these breaks (ROSS et al. 1979). The direct analysis with purified topoisomerase II has shown that m-AMSA and ellipticine very significantly stimulate the formation of DNA-topoisomerase complexes (NELSON et al. 1983; TEWEY et al. 1984). Single- and double-strand breaks are observed with topoisomerase II monomers covalently linked to the 5′ ends of the breaks. The isomer o-AMSA, which is not cytotoxic and which induces no significant DNA breakage in cultured cells, does not essentially affect topoisomerase II reactions in vitro. The mechanism of action of these drugs is possibly the stabilization of a "cleavable complex" formed between topoisomerase II and DNA. Visualization of the actual breakage requires, at least in vitro, treatment with a protein denaturant (TEWEY et al. 1984). These results indicate that topoisomerase II in mammalian cells is a prime target for these drugs, and they confirm similar, albeit indirect conclusions drawn from experiments in isolated cell nuclei and with isolated nucleoids which had been exposed to m-AMSA, ellipticine and, in addition, 5-iminodaunorubicin (POMMIER et al. 1984a, b). Further, novobiocin and coumermycin prevent the induction of DNA breaks if presented together with m-AMSA. In m-AMSA pretreated nuclei the two antibiotics did not interfere with rejoining of the induced breaks if added just prior to m-AMSA dilution to allow reclosure of DNA (MARSHALL et al. 1983).

Analogous inhibition of a catenating (topoisomerase II) activity by these drugs has been observed with a partially purified enzyme preparation of *T. cruzi* (DOUC-RASY et al. 1984). Further, the nonintercalating anticancer drugs VP16-213 and VM26 have very recently also been identified as potent topoisomerase II inhibitors in vitro (MINOCHA and LONG 1984).

Comparison with *E. coli* DNA Gyrase. Various aspects of eukaryotic topoisomerase II are reminiscent of *E. coli* DNA gyrase. Both require ATP for their activity, both are of composite, multimeric structure, and both share the essential features of their cleavage mechanisms, i.e., staggered breakage, formation of covalent links with the 5′ ends of the breaks, and site specificity of cleavage. For both categories of enzymes, inhibiting drugs are known which are involved in the

formation of DNA breaks – oxolinic acid for gyrase and m-AMSA and others for the eukaryotic topoisomerase II, respectively.

On the other hand, a number of clear differences exist between these enzymes. Most importantly, the eukaryotic enzymes do not promote negative supercoiling of DNA. In contrast to the *E. coli* gyrase, the eukaryotic type II enzymes relax both positively and negatively supercoiled DNA molecules. Strand passage always goes into the thermodynamically favored direction. In this respect, they resemble closely the type II topoisomerase of the phage T4 (see Sect. 8). It can only be speculated that these distinctions reflect fundamental differences in the quarternary structures of the respective DNA-enzyme complexes. Further differences exist with respect to the effect of single-stranded DNA on the action of these enzymes. Whereas this DNA is a potent inhibitor of the relaxation reaction of the eukaryotic topoisomerase II, it has no effect on the supercoiling reaction of gyrase. Distinct is also the mode of action of nucleoside triphosphates. The ATP binding sites of both enzymes are competitively inhibited by novobiocin and coumermycin; however, these competitions are much more efficient for gyrase than for eukaryotic topoisomerase II. In addition, gyrase has a more stringent requirement for ATP over dATP. Nucleotide binding to gyrase involved the α, β, and γ phosphates of ATP, whereas only the β and γ phosphates appeared to interact with the eukaryotic enzymes. This follows from the observation that ATP analogs, ADP, and AMP all inhibit reactions of DNA gyrase while only ATP analogs and ADP affect the eukaryotic topoisomerase II, as was exemplified with the topoisomerase II of *D. melanogaster* (Osheroff et al. 1983).

It is still a matter of debate whether negative supercoiling in eukaryotic chromosomes occurs as a consequence of the cooperative activity of known (and unknown) nuclear proteins, e.g., of histones and topoisomerase I and/or topoisomerase II, or whether it results from an as yet unidentified gyraselike enzyme (see, e.g., Mattern and Painter 1979; Edenberg 1980; Glikin et al. 1984; Ryoji and Worcel 1984). A unique eukaryotic gyrase cannot be exluded a priori. However, mechanisms for the generation of supercoiling in chromatin can be conceived which do not depend on such a unique enzyme. One might say that if eukaryotes do not have a gyrase, an assembly of cooperating nuclear proteins could afford the same reaction which would otherwise be catalyzed by a gyrase.

7 Topoisomerases in Mitochondria and Chloroplasts

That mitochondria may possess a DNA nicking-closing activity was inferred before topoimerases were known (see, e.g., Robberson et al. 1972). This conclusion was based on the observation that mitochondrial DNA in mouse L cells and in other organisms replicates essentially according to a modified Cairns model (Cairns 1963) via asymmetric extension of a displaced loop (D loop) with both strands remaining closed most of the time (Berk and Clayton 1974, 1976; for a review of that time see Kasamatsu and Vinograd 1974). Additional,

but still circumstantial evidence for topoisomerase activity in the mitochondria of mouse L cells was presented more recently by BOGENHAGEN and CLAYTON (1978) who found that newly synthesized and segregated daughter molecules have zero superhelix density. These molecules assume the supercoiled conformation only after a delay of about 1 h. From this result, the authors speculated upon the presence of gyraselike activity in the mitochondria.

Additional indirect evidence for a topoisomerase in mitochondria arises from the structure of the kinetoplast DNA (kDNA) in the mitochondria of trypanosomes. As mentioned above, this DNA consists of a network of many thousands of catenated minicircles (for a review see BORST and HOEIJMAKERS 1979). Formation and resolution of these networks in the course of kDNA replication and segregation conceivably involve a type II topoisomerase. It has been shown that kDNA is decatenated in vitro by type II topoisomerases of other sources, e.g., by *E. coli* DNA gyrase (MARINI et al. 1980). In the meantime, an endogenous type II topoisomerase has been identified in cell extracts of the trypanosomatid *C. fasciculata* (SHLOMAI and ZADOK 1983). This enzyme may not be uniquely mitochondrial; it decatenates, however, kDNA in vitro. It has properties typical for a eukaryotic type II topoisomerase. It requires ATP for its activity and is inhibited by novobiocin. A type I topoisomerase as that purified from *T. cruzi* (RIOU et al. 1983) is less likely to be involved in the catenation-decatenation of kDNA.

Topoisomerase activity has also been observed in rat liver mitochondria (FAIRFIELD et al. 1979). The relaxation reaction of this enzyme did not require Mg^{2+} and ATP; hence, it is presumably a type I topoisomerase. Since this enzyme was found to be inhibited by ethidium bromide and by the trypanocidal drug Berenil, in contrast to the nuclear topoisomerase of the same organism, the authors concluded that the mitochondrial topoisomerase was of distinct character.

DNA topoisomerase I from mitochondria of *X. laevis* oocytes (BRUN et al. 1981) has been purified and carefully studied. This enzyme comprises ca. 1% of the total topoisomerase activity in oocytes. Its molecular weight is 65000–70000. Hence, it may be a cleavage product of the nuclear type I topoisomerase. Its catalytic and physicochemical properties are apparently identical with those of the corresponding nuclear enzyme, including sensitivity to ethidium bromide and berenil. In *X. laevis* at least, the one identified topoisomerase, which is associated with mitochondria, does not, therefore, appear to be specifically mitochondrial. A similar conclusion has been reached for yeast mitochondria. Topoisomerase I and II, which both have been found associated with purified yeast mitochondria, are indistinguishable from the topoisomerases of yeast nuclei (GOTO et al. 1984). The yeast topoisomerase is not inhibited by berenil or ethidium bromide.

The existence of a DNA gyraselike activity was deduced from experiments with intact mitochondria (CASTORA and SIMPSON 1979) and partially purified protein fractions from rat liver mitochondria (CASTORA et al. 1983). The evidence for a mitochondrial gyrase was essentially based on the inhibition of DNA synthesis and on the failure of mitochondria to generate or maintain negative supercoiling of mitochondrial DNA in the presence of the *E. coli* gyrase inhibi-

tors nalidixic acid, oxolinic acid, novobiocin, and coumermycin. Biochemical data about the presumed type II topoisomerase from rat liver mitochondria are not yet available.

Recently a topoisomerase has been purified from spinach chloroplasts (SIEDLICKI et al. 1983). This enzyme seems to resemble the prokaryotic ω-protein in its catalytic and physical properties. The similarity was taken to suggest that this topoisomerase might be of prokaryotic origin and might have been acquired by the plant via endosymbiosis of ancient plastid precursors.

It is apparent from the different reports that nature and origin of mitochondrial and chloroplast topoisomerases are presently difficult to assess.

8 Viral Topoisomerases

Topoisomerases may be linked to viruses in two different ways. They are either encoded for by the viral genomes or they are of host origin and are found associated with viral components, such as minichromosomes or capsids. So far, two viruses have been reported to possess an endogenous topoisomerase; these are the bacteriophages T4 and the vaccinia virus. In the case of T4 the topoisomerase coding genes are known. With respect to vaccinia virus a topoisomerase which is different in many of its properties from the known eukaryotic topoisomerases has been detected as a virus-encapsidated enzyme. It is, strictly speaking, not known whether it is coded for by the virus. In addition, it has been suggested that phage Mu may code for a distinct topoisomerase. (The Int protein of the bacteriophage λ is treated separately in Sect. 9.) Association of host topoisomerases with viral components has been shown for SV40, polymoma, and Rous sarcoma viruses.

In those cases where a topoisomerase is genetically linked to a virus, a distinct role of the enzyme in viral functions is evident. For topoisomerases of host origin that are associated with virus components the same conclusion is not as obvious. Formation of such associates could simply reflect unspecific binding of host factors to viral chromosomes, and even inclusion into virus particles may be fortuitous.

The most extensively studied viral topoisomerase is that of the coliphage T4 (STETLER et al. 1979; LIU et al. 1979; LIU et al. 1980; KREUZER and JONGENEEL 1983; KREUZER and ALBERTS 1984). This enzyme is a multisubunit enzyme coded for by the T4 genes 39, 52, and 60 with molecular weights of 64000, 57000, and 16000, respectively (LIU et al. 1979). Mutations in these genes delay the initiation of viral DNA replication with the rate of fork movement essentially being unaffected (YEGIAN et al. 1971; MCCARTHY et al. 1976). These mutants do not grow in the presence of novobiocin or coumermycin. Since wild-type phage is resistant to these drugs, it is assumed that the gyrase of the host is able to replace the missing function (MCCARTHY 1979).

The enzyme catalyzes relaxation of positively and negatively supercoiled DNA, catenation/decatenation, and knotting/unknotting of circular DNA molecules. The latter two reactions and the observation of linking number changes

in steps of two show that the mechanism of this enzyme involves strand passing through double-strand breaks. Thus, this enzyme belongs to the category of type II topoisomerases. One to two ATP molecules are hydrolyzed in one reaction cycle during catalytic relaxation. A thiolated analog of ATP, adenosine-5′-0-(3-thiotriphosphate), or ATP-γ S, inhibits the reaction, presumably due to not being hydrolyzed (LIU et al. 1979). However, some relaxation, the extent of which could be related to the concentration of the analog, indicates that ATP hydrolysis is most probably needed for turning over of the enzyme rather than for promotion of the catalytic cycle itself. This effect of the thio analog may be similar to that of β-,γ-imido ATP on gyrase.

The small subunit (16000 daltons) supposedly functions by holding the two larger subunits in juxtaposition within the native protein complex (SEASHOLTZ and GREENBERG 1983). ROWE et al. (1984) demonstrated covalent binding of the 52000-dalton subunit to the 5′ end of broken DNA after stimulation of cleavage with the antitumor drug m-AMSA. This subunit is apparently comparable in its function with the subunit A of *E. coli* gyrase. Cleavage of native T4 DNA, which contains glucosylated hydroxymethylcytosine, is site specific (KREUZER and ALBERTS 1984). Cleavage of mutant, nonglucosylated DNA is less specific. In contrast, *E. coli* gyrase does not distinguish between normal and nonglucosylated T4 DNA; it recognizes both DNAs equally well. The T4 enzyme cleaves also single-strand DNA very efficiently (KREUZER 1984).

The T4 topoisomerase promotes initiation of the replication of the linear DNA. It has been speculated that this enzyme might induce some local supercoiling by forming a loop within the region of the replication origin (LIU et al. 1979, 1980). There is as yet no evidence for such a mechanism. It should be noted, however, that the generation of local supercoiling by a topoisomerase is principally conceivable, as was shown by MOORE et al. (1983) in their studies of gyrase binding to DNA.

With regard to phage Mu it has been suggested (GHELARDINI et al. 1982) that it requires a DNA nicking-closing activity for its integration into the chromosome of *E. coli*. Integration into host DNA is obligatory for the normal life cycle of Mu. Biochemical or genetic evidence for a Mu-coded topoisomerase does not as yet exist. However, if Mu lysogens are infected with a T4 phage that has an amber mutation in one of its topoisomerase genes, T4amG39, the genetic defect can be complemented by a Mu function in a nonsuppressing cell. The same authors (GHELARDINI et al. 1982) reported further that extracts of the Mu lysogenic strains, in contrast to extracts of nonlysogenic cells, contain a biochemically active T4 topoisomerase after infection with the above T4 topoisomerase amber mutant. The nature of the complementing agent is not known.

Another topoisomerase of presumed viral origin has been reported for vaccinia virus (BAUER et al. 1977); this enzyme could be isolated from the core particles of this virus. Each core contains about 250 copies of the enzyme monomer (molecular weight, 35000). This enzyme is a type I topoisomerase which relaxes negatively and positively supercoiled DNA. It has a number of unique properties. Although ATP is not required, it stimulates the activity about 2.5-fold. The enzyme has, however, no ATPase activity. ADP is inhibitory, as are the gyrase inhibitors novobiocin and coumermycin. Since these drugs

also inhibit transcription in vaccinia virus core particles, it has been suggested that this unusual topoisomerase has a function in this process (FOGLESONG and BAUER 1984). In addition to these results, a preliminary report has been published indicating the appearance of an ATP-dependent, vaccinia-virus-specific decatenating activity in HeLa cells which were infected with this virus (MILLER and ENGLUND 1981). No details are available about this activity.

Cells infected with herpes simplex virus (HSV) type I were found to contain a topoisomerase which was copurified with HSV-1-induced DNA polymerase (BISWAL et al. 1983). Whether this topoisomerase is of viral origin or not remains to be shown.

The attachment of a topoisomerase I to SV40 nucleoprotein was reported quite early (SEN and LEVINE 1974). Related observations have since been communicated by others for SV40 nucleoprotein complexes (KELLER et al. 1977; HAMELIN and YANIV 1979, 1980). Even the release of a topoisomerase from mature SV40 capsids has been reported (BINA et al. 1982). In addition, topoisomerase II has been identified as being complexed with SV40 minichromosomes (WALDECK et al. 1983). Similar associations of topoisomerase I and II with minichromosomes have also been found in studies of polyomavirus (GOURLIE et al. 1981; KRAUSS et al. 1984).

It is reasonable to assume that topoisomerases are involved in the propagation of these papovaviruses. Since replication follows the Cairns mode, topoisomerases could be required for strand elongation and for strand segregation at the end of one round of replication. SUNDIN and VARSHAVSKY (1981) were the first to conclude on the basis of their studies with SV40 replicative intermediates that a type II topoisomerase catalyzes segregation of DNA daughter molecules. For promotion of the condensed chromatin structure, a topoisomerase is probably also required.

Somewhat unexpected is the association of a topoisomerase with the RNA-containing Rous sarcoma virus (RSV) (WEIS and FARAS 1981). This topoisomerase was detected in a search for an activity which in disrupted virions of RSV may be able to facilitate – in concert with the polymerizing activity of reverse transcriptase – the formation of covalently closed circular DNA (CLAYMAN et al. 1979). The topoisomerase is most probably a type I enzyme with a molecular weight similar to that of a nuclear type I topoisomerase. Approximately five molecules have been estimated per RSV virion. Whether this topoisomerase is able to circularize DNA is open. It is unknown whether viral inclusion of this enzyme, if it has a function in the life cycle of RSV, reflects a "late" function in virus maturation or an "early" function required at the beginning of a new infectious cycle.

In addition, two bacteriophage coded proteins are known which resemble topoisomerases by their ability to break and rejoin phosphodiester bonds in DNA. However, in contrast to the standard reactions of topoisomerase, a chain transfer is associated with the principal reactions of these phage proteins. The normal nicking-closing cycles, which they also catalyze, probably have no physiological significance.

The two phage proteins are the ϕX174 gene A protein and the fd gene II protein. The gene A protein of the phage ϕX174 has a molecular weight

of 56000 (HENRY and KNIPPERS 1974). The coding sequence on the phage genome for gene A protein is known (SANGER et al. 1977, 1978). This protein functions as a strand-specific endonuclease by cutting the viral strand of a negatively supercoiled double-stranded φX174 DNA molecule at a unique position, defined as the origin of replication. The activities of gene A protein in φX174 DNA replication include cleavage at the origin, complexing with the 5′ end of the nicked strand to support strand separation in the rolling-circle-type replication, cleavage of the displaced strand at the origin after one round of synthesis, and ligation of the newly exposed 3′-OH end to the complexed 5′ end to generate circular viral molecules (EISENBERG and KORNBERG 1979). Only superhelical φX174 DNA is a substrate for gene A protein; relaxed or nicked duplex DNA is not cleaved (IKEDA et al. 1979). The finding that viral single strands of φX174 are also cleaved at the origin (LANGEVELD et al. 1979) is reasonably interpreted by assuming that this protein recognizes a defined local secondary structure which is present in single-stranded phage DNA and at least transiently in superhelical DNA, but not in relaxed DNA. A relaxation reaction has been observed with superhelical φX174 DNA after replacement of Mg^{2+} by Mn^{2+} (LANGEVELD et al. 1980). The action of gene A protein is not directly comparable with the action of standard topoisomerases. Breaking and rejoining are the end points of a complex chain of events which starts with superhelical double-stranded DNA as substrate and which terminates with the formation of a newly created circular single-stranded DNA molecule.

The gene A protein exists in a truncated form, called A* protein, which contains about 60% of the C-terminal sequence of the A protein. This A* protein has been shown to form a covalent complex through a tyrosyl-dAMP phosphodiester bond (SANHUEZA and EISENBERG 1984).

Gene II protein of phage fd has a similar function to φX174 gene A protein in the replication of this phage. This protein has a molecular weight of 46000 (MEYER and GEIDER 1979a), and its corresponding nucleotide sequence is known (BECK et al. 1978). As with φX174 gene A protein, it cleaves supercoiled fd DNA at a specific site (MEYER et al. 1979), but not relaxed fd DNA. Unlike φX174 gene A protein, which cleaves single-stranded φX174 DNA, gene II protein does not cleave single-stranded fd DNA. Thus, superhelicity is required for proper formation of the recognition site. As a side reaction, gene II protein is able to relax superhelical fd DNA in vitro, which is similar to the reaction of a topoisomerase (MEYER and GEIDER 1979b). Covalent DNA-protein complexes can be postulated, but they have thus far not been seen.

9 Recombinational Topoisomerases

Two topoisomerases have been described which have a defined function in site-specific recombination events, the λ Int protein (Int for integration) and the resolvase of the γ δ and Tn3 transposons.

The phage λ in its lysogenic pathway is able to integrate into the *E. coli* chromosome at a specific site. This integrative recombination is mediated by

the product of the λ Int gene and by a host-coded protein, integration host factor (IHF) (NASH and ROBERTSON 1981; for recent reviews see NASH 1981; WEISBERG and LANDY 1983). In cooperation with host protein factors and the λ Xis protein, Int is also able to promote excision of λ DNA from the host genome. Integration occurs by reciprocal recombination between two attachment sites, *att*P, the specific site on the phage chromosome, and *att*B, the corresponding site on the bacterial chromosome. The two sites share a core of a homologous sequence which is 15 bp long. The strand exchange reaction occurs within this region.

The Int protein has a molecular weight of 44000. The nucleotide sequence of the Int gene is known (HOESS et al. 1980). The reaction of Int involves the breakage of two double-stranded DNA molecules, those of the phage and the *E. coli* chromosome, at the *att* sites, and rejoining by a reciprocal crossover. Nash and his collaborators have presented a molecular interpretation of this event and its topological implications (NASH and POLLOCK 1983; POLLOCK and NASH 1983). Since no external energy is required for the rejoining step, a topoisomeraselike mechanism must be invoked which includes transient covalent DNA-protein complexes. Such complexes have been identified; they are formed between the Int protein and a 3′-phosphate at the breakpoint in the DNA. The 5′ end carries an OH-group (CRAIG and NASH 1983). This is the first prokaryotic topoisomerase which is linked to a 3′-phosphate. This type of linkage is otherwise typical for eukaryotic type I topoisomerases. Site-specific cleavage within the *att* site requires recognition of 9-bp-long sequences straddling the junctions of the core and the regions flanking the core on either side. The cleavage points are within the core exactly there, where the strand exchange occurs. They are staggered by 7 bp. Cleavage of duplex DNA by Int is always observed in one strand at a time (CRAIG and NASH 1983).

In addition to this principal function, Int is able to relax supercoiled DNA, albeit at a slow rate. This reaction appears to belong to the type I nicking-closing category, since the enzyme acts via transient single-strand nicks and linking numbers change in units of one (NASH et al. 1980). The specific *att* site is not needed for this reaction (KIKUCHI and NASH 1979). The relaxation is nevertheless probably a site-specific event, since it has been shown that Int binds to relatively short sequences which are similar to the recognition sequences within the *att* sites and common in non-*att* DNA (ROSS and LANDY 1983).

The second highly specialized recombinational topoisomerase is the resolvase of the γ δ and the Tn3 transposons (REED 1983; KRASNOW and COZZARELLI 1983). This protein is needed for a subfinal step in the transposition of the γ δ or Tn3. Transposition is essentially a two-step process with, first, formation of a fused cointegrate of two separate replicons and, subsequently, resolution of this composite structure. Resolution is catalyzed by resolvase. This protein is a strictly site-specific topoisomerase recognizing the *res*, or resolution sites in DNA. Recombination requires the presence of two *res* sites in directly repeated order within one DNA molecule. The distance between the *res* sites is of no critical importance (for an interpretation of these requirements see KRASNOW and COZZARELLI 1983). Resolvase is a small protein (21000 daltons) which acts as a type I topoisomerase. The resolvase reaction does not lead

to free rings, but to catenated dimers with each ring carrying one *res* site. A decatenation and frequently also an unknotting event (see KRASNOW and COZZARELLI 1983) presumably by gyrase will eventually complete the resolution of cointegrates.

10 Biological Functions of Topoisomerases

Until recently, the considerable amount of information about the in vitro reactions of topoisomerases was in striking contrast to the scanty knowledge about the functions they serve in living cells. The currently emerging concept for the in vivo role of these enzymes developed slowly, although not unexpected in many of its features. The functions of prokaryotic topoisomerases are more obvious than those of their eukaryotic counterparts. For both categories of enzymes, the picture is rather complex and many details are still missing. However, at least in principle, topoisomerases are no longer enzymes in search of a function. Critical progress was achieved in the studies of functions when pro- and even some eukaryotic mutants became available. Drug inhibition and, most importantly, the newly developed techniques of injecting DNA into single cells (frog oocytes) were also extremely useful in the analysis of functions.

A unique property of topoisomerases is their ability to act from the distance. Whereas some effects of these enzymes may be restricted to narrow regions within DNA, others can be transmitted to remote sections by using DNA as the transmitting medium. This condition implies at the same time pleiotropic effects on all events which are influenced by DNA tertiary structure.

Not all of the known in vitro reactions of the various topoisomerases may be of physiological significance. Whereas catenation and knotting of duplex DNA may be useful when tight packing of DNA is required, knotting of single-stranded DNA molecules, or production of highly contracted structures from denatured closed circles, may never occur in living cells. Relaxation of unrestrained positive supercoils can be catalyzed by many topoisomerases, but it is open to question whether such conformations occur in vivo, excepting transient intermediate DNA conformations and, possibly, DNA in archaebacteria. On the other hand, a reaction which may look like an in vitro artifact merely illustrating the catalytic capabilities of some of these enzymes, such as the interstrand transfer reaction of eukaryotic type I topoisomerases, might turn out to be instrumental in certain illegitimate recombination events in eukaryotic nuclei.

The following presentation first summarizes briefly the existing data on the intracellular conformation of DNA and then offers a detailed discussion on the functions of topoisomerases. Although pro- and eukaryotic topoisomerases have many features in common, they differ in many other respects. Therefore, pro- and eukaryotic topoisomerases will be treated separately. Within each class, type I and type II topoisomerases, which in vivo operate most likely in a concerted fashion, will be considered together. The known or presumed functions

of viral topoisomerases and of the recombinational topoisomerases Int and resolvase have already been discussed in Sects. 8 and 9.

10.1 Conformation of DNA In Vivo

The structural organization of DNA within living cells is characterized by negative supercoiling. The majority of the results reported in the past about the physical and chemical properties of DNA with this conformation were obtained in studies of circular closed duplex DNA. Such DNA can be isolated from a variety of pro- and eukaryotic sources. The extent of supercoiling in these DNA molecules ranges from about -0.03 to -0.09 (for a review see BAUER 1978). As already mentioned, chromosomal DNA of both pro- and eukaryotic origin is also negatively supercoiled. In *E. coli* cells, the chromosome is folded into a compact structure consisting of about 50 looped superhelical domains (STONINGTON and PETTIJOHN 1971; WORCEL and BURGI 1972), with each domain comprising a separate unit of roughly 100000 base pairs. Negative supercoiling within these domains corresponds to a specific linking difference (or superhelical density) of about 0.05 (SINDEN et al. 1980). A somewhat analogous situation was identified in the DNA of eukaryotic interphase genomes (COOK and BRAZELL 1975; BENYAJATI and WORCEL 1976; PINON and SALTS 1977) when investigating partially deproteinized chromosomes. These eukaryotic domains were found to comprise a length of about 85000 base pairs in the *Drosophila* genome (BENYAJATI and WORCEL 1976) and about 1×10^6 base pairs in the HeLa cell genome (COOK and BRAZELL 1975). The superhelical density in the eukaryotic DNA domains after partial deproteinization approximates that of the folded *E. coli* DNA.

Despite these similarities, principal differences exist in the mechanisms for the introduction of supercoils and also in the details of their molecular organization in eukaryotic and prokaryotic systems. In eukaryotes, negative supercoiling results from the left-handed wrapping of DNA with about $1^3/_4$ turns around a core of eight histone molecules (FINCH et al. 1977). This coiling constitutes the nucleosome as the basic structural unit of chromatin. Nucleosome formation has been demonstrated in vitro (GERMOND et al. 1975, 1979; LASKEY et al. 1977; RUIZ-CARILLO et al. 1979). Torsional tension, which occurs in a topologically fixed DNA molecule when part of it is wound around a protein core of histones or synthetic model polypeptides (DESANTIS et al. 1983), can be relieved by the action of a topoisomerase. Once the nucleosomes are formed, the DNA is supercoiled, but it is not torsionally stressed.

That this is also true for chromatin inside living cells was concluded from the rate of covalent photobinding of trimethylpsoralen to intracellular DNA. This rate is greater when DNA is under torsional stress than when it is relaxed. The photoreaction of DNA with this adduct within eukaryotic cells was found to be similar to that of relaxed DNA (SINDEN et al. 1980). The supercoil in nucleosomes is said to be restrained (PETTIJOHN and PFENNINGER 1980), and very little or no unrestrained supercoiling exists in internucleosomal regions. A similar result had already been obtained earlier in studies of SV40 nucleoprotein extracted from SV40-infected cells (YOUNG and CHAMPOUX 1978).

It should be noted that this picture is probably not complete. It is concluded from indirect evidence obtained in gene regulation and transcription studies that at least subregions of chromatin exist within cells in an unrestrained supercoiled state (HARLAND et al. 1983; LUCHNIK et al. 1982a, b; LARSEN and WEINTRAUB 1982; see also Sect. 11).

Histonelike proteins exist also in *E. coli* cells (ROUVIÈRE-YANIV and GROS 1975; BERTHOLD and GEIDER 1976; ROUVIÈRE-YANIV et al. 1979). They were previously called proteins HU, HD, and NS, and they are now classified as DNA binding protein II (for a review see GEIDER and HOFFMANN-BERLING 1981). It has been demonstrated that they are able to wind negative supercoils into DNA (ROUVIÈRE-YANIV et al. 1979), a result similar to that observed with histones. In addition, it had been shown earlier by electron microscopy that *E. coli* DNA, after careful disruption of cells on the spreading film, appears as a regularly shaped fiber structure, analogous to the nucleosome structure in chromatin (GRIFFITH 1976).

However, although the *E. coli* cell appears to have at least some potential to organize its DNA conformation in a restrained mode, evidence suggests that the majority of the bacterial DNA is not restrained. The trimethylpsoralen probe reacts with intracellular DNA as expected for purified superhelical DNA (SINDEN et al. 1980). It was also demonstrated that gyrase is responsible for the formation of unrestrained supercoils, since DNA was relaxed within the cells after treatment with coumermycin (DRLICA and SNYDER 1978; SINDEN et al. 1980). That some – possibly as much as 15% or even more – of the intracellular supercoiling in prokaryotes can be accounted for by mechanisms other than supercoiling by gyrase follows, e.g., from an early analysis of phage λ-DNA extracted from superinfected *E. coli* λ-lysogens. This DNA is circularized and supercoiled after infection, but it is not replicated. In the presence of coumermycin or oxolinic acid some, although limited, supercoiling is observed (GELLERT et al. 1976b; GELLERT et al. 1977). In summary, it can be concluded that the majority of the DNA in *E. coli* cells is under torsional tension and that gyrase is the factor which produces this tension.

10.2 Prokaryotic Topoisomerases

Only gyrase and the type I topoisomerase (ω-protein) are considered here. The functional relevance of the topoisomerases II′ and III (both found in *E. coli*) is not known. Gyrase is required for **DNA replication.** This follows, first, from the observation that several antibiotics which inhibit gyrase in vitro and in in vivo are also inhibitors of DNA replication. Second, temparature-sensitive mutations of the *gyr*A and *gyr*B genes in *E. coli* block DNA replication under nonpermissive conditions (ORR et al. 1979; KREUZER and COZZARELLI 1979). Novobiocin and coumermycin which both inhibit the gyrB subunit affect initiation and elongation of replicative synthesis in vivo (SMITH and DAVIES 1967; DRLICA and SNYDER 1978; RYAN 1976) and in permeabilized cells (STAUDENBAUER 1975; RYAN and WELLS 1976). Bacterial chromosomes and nonchromosomal circular and even some linear replicons require gyrase for their replication. Thus, initiation of ColE1 replication in cell-free extracts is blocked by coumer-

mycin and other antigyrase drugs (GELLERT et al. 1976b; STAUDENBAUER 1976b). In cell-free extracts, synthesis in the presence of novobiocin can be partially restored by the addition of gyrase from a novobiocin-resistant strain (GELLERT et al. 1976b).

A temperature-sensitive (ts) *gyr*B mutant has been isolated which at the nonpermissive temperature inhibits only initiation of replication, but not chain elongation (ORR et al. 1979; ORR and STAUDENBAUER 1981). This is in contrast to the action of coumermycin, which also arrests elongation (RYAN 1976; DRLICA and SNYDER 1978; ENGLE et al. 1982). Another ts *gyr*B mutant has been isolated that is deficient in chain elongation (FILUTOWICZ and JONCZYK 1983). The reason for the distinctive behavior of gyrase may be explained by a mode of action of gyrase which is conceivably not identical for initiation and elongation. With respect to mutation, one ts mutation may affect the role of gyrase preferentially in initiation, the other in elongation.

A thermosensitive mutation in the *gyr*A gene was found to lead to a rapid arrest of elongation (KREUZER and COZZARELLI 1979). This result suggests a specific requirement of subunit A for this stage in replication. Since the separate gyrase subunits are catalytically inactive in vitro – except for a residual ATPase activity of subunit B (STAUDENBAUER and ORR 1981) – it is plausible that the gyrase tetramer is required for both initiation and elongation in DNA replication. It has been suggested that gyrase exerts its function in elongation in close proximity to replication forks (DRLICA et al. 1980).

The precise role of gyrase in initiation of chromosomal replication is unknown. A direct participation of gyrase in the initiation process is suggested by the finding that this enzyme has a preferred binding site at the origin of replication (*ori*C) of *E. coli* (LOTHER et al. 1984). The binding is inhibited by oxolinic acid and the ATP analog β-γ-imido ATP. No binding to the *E. coli ori*C locus is seen with gyrase isolated from the gram-positive bacterium *B. subtilis*. This result is in apparent contrast to the observation of LAMPE and BOTT (1984) that a temperature sensitive *E. coli gyr*A mutation could be complemented by the corresponding *Bacillus subtilis* wild type gyrase gene introduced into the *E. coli* mutant as a plasmid cloned DNA (see p. 45). Since, however, the two systems tested are not directly comparable, these results are not necessarily contradictory. Conceivably the *E. coli* ts mutation does essentially not impair recognition of the *E. coli ori*C region, but predominantly other functions of gyrase which can be retrieved by the *Bacillus subtilis* gene.

An unusually high coumermycin sensitivity of the initiation step was found in a ts *dna*A mutant of *E. coli* (FILUTOWICZ 1980). The product of the *dna*A allele, which maps next to the *gyr*B gene, is involved in initiation of replication. Conceivably, gyrase-promoted supercoiling of the region where replication originates is required for binding of the *dna*A protein (KAGUNI and KORNBERG 1984a, b). The thermolabile *dna*A protein may function normally when gyrase activity is optimal, but a small impairment of this activity by low coumermycin concentrations, which are tolerated by the wild-type *dna*A protein, may not be acceptable for the mutant protein.

According to several reports, novobiocin and coumerycin are able to cure *E. coli* cells from plasmids (TAYLOR and LEVINE 1979; DANILEVSKAYA and GRA-

GEROV 1980; WOLFSON et al. 1982). Plasmid replication is apparently more sensitive to gyrase inhibition than replication of the host chromosome; however, selective degradation of the plasmid has also been suggested to explain this result (DANILEVSKAYA and GRAGEROV 1980).

Supercoiling and hence gyrase are also required for initiation of in vitro DNA replication of the phages φX174 (SUMIDA-YASUMOTO et al. 1976; MARIANS et al. 1977) and fd (GEIDER et al. 1982). As described earlier, replication of these DNA molecules requires specific initiation factors, the φX174-coded *cis*A protein and the fd-coded gene II proteins, respectively. Both proteins act as site-specific endonucleases and recognize superhelical substrate molecules only. In addition to this general and indirect effect of gyrase on phage DNA synthesis, a more specific and direct role of this enzyme in φX174 viral DNA synthesis has also been suggested (HAMATAKE et al. 1981). From inhibition of this process in vitro and in vivo by subunit A but not by subunit B antagonists, it was concluded that only subunit A of gyrase participates in this event. The action of gyrase as a swiveling device has been tentatively proposed to explain this observation.

A somewhat surprising effect of coumermycin in vivo is inhibition of replication of the linear T7 DNA (ITOH and TOMOZAWA 1977; DEWYNGAERT and HINKLE 1979). A topological problem should not be involved in the replication of this phage DNA. One could, however, tentatively argue that axial rotation of long linear DNA (T7 is about 30000 base pairs long) in the course of unwinding of parental strands during replication causes torsional stress and, hence, requires the action of gyrase.

The well-established inhibition of replication of *E. coli* DNA by oxolinic acid and nalidixic acid in vivo (COOK et al. 1966; DERMODY et al. 1974; CRUMPLIN and SMITH 1976) and in permeabilized cells (PEDRINI et al. 1972) is most probably explained not only by an interference with supercoiling but also by an interaction with gyrase, which leads to the formation of DNA-gyrase complexes and possibly other events that may ultimately lead to a loss of chromosomal integrity (for detailed review of the evidence see GELLERT 1981a). Briefly, it has been known for some time that nalidixic acid after prolonged application is able to trigger intracellular DNA fragmentation (RYAN 1976). It is not known whether gyrase plays a role in this process. Oxolinic acid is able, as described earlier, to induce DNA-gyrase aggregates, which upon incubation with SDS lead to fragmentation of DNA. This event occurs in vitro and also within intact cells (COOK et al. 1966; SNYDER and DRLICA 1979). Formation of these complexes at specific gyrase binding sites may be assumed. About 45 of these sites have been estimated in *E. coli*. Whether the approximate coincidence of this number with that of supercoiled loops in the *E. coli* chromosome is fortuitous or not cannot be decided at present. The recent observation that separation of loops with different degrees of superhelicity can be promoted by gyrase binding to DNA (MOORE et al. 1983; discussed in Sect. 6.1.1) lends support to the notion that gyrase binding is involved in the overall tertiary structure of the *E. coli* chromosome.

Gyrase is also required to complete DNA replication by segregating the daughter chromosomes through a decatenating event. It has been shown that

nucleoids from ts *gyr*B mutants of *E. coli* are obtained as doublets if isolated from cells grown under nonpermissive conditions. In vitro, these doublets could be resolved by the addition of gyrase (STECK and DRLICA 1984).

Other impairments due to *gyr*B mutations affect chromosomal organization by unscheduled relaxation of superhelical domains and cell division, which is either accompanied by the generation of anucleate cells with a high frequency (ORR et al. 1979; FAIRWEATHER et al. 1980) or even by a complete stop of cell division (MIRKIN and SHMERLING 1982).

A specific role of gyrase for initiation and presumably for elongation has also been shown for DNA replication in *B. subtilis* (OGASAWARA et al. 1979, 1981). These experiments were based on inhibition of synchronized DNA replication by novobiocin and nalidixic acid.

That topoisomerase I has apparently also a function in DNA replication of bacterial chromosomes was not generally anticipated. Indirect genetic evidence comes from a study showing that a ts *dna*A mutation can be suppressed by a deletion of the *top*A gene. This result implies that a change in the superhelical state of the chromosome compensates a deficiency of the *dna*A protein (LOUARN et al. 1984). From biochemical replication studies with purified proteins, it was concluded that topoisomerase I confers specificity in the replication of the *E. coli ori*C locus (KAGUNI and KORNBERG 1984a, b). This effect is explained by destabilization of replicative primers in regions other than *ori*C. Thus, the topoisomerase prevents random initiation, presumably by relaxing unspecific primer-template complexes, which causes dissociation of those RNA primers that are not tightly bound to DNA by the *ori*C-specific *dna*A protein.

Another area where topoisomerases, in particular gyrase, are apparently involved is **transcription**. It has been known for quite some time that the efficiency of transcription is linked to the conformation of DNA. Negatively supercoiled DNA was commonly found to enhance transcription compared with relaxed, nicked circular, or linear DNA (HAYASHI and HAYASHI 1971; BOTCHAN et al. 1973; WANG 1974b; BOTCHAN 1976; SEEBURG et al. 1977). The number of binding sites, the rate of binding of polymerase to DNA, as well as the stability of the complex between the two components were all found to be positively affected by negative superhelicity (RICHARDSON 1975a, b). The reading of the genetic message by RNA polymerase causes a small but defined unwinding of the DNA helix by approximately one turn (SAUCIER and WANG 1972; WANG et al. 1977). The local transition in the DNA helix induced by the enzyme is greatly favored by the superhelical conformation with its inherent tendency to release conformational stress by unwinding.

In agreement with this general notion are results which demonstrate a positive correlation between DNA supercoiling by gyrase and stimulation of transcription both in vitro (AKRIGG and COOK 1980) and in vivo (see, e.g., KUBO et al. 1979; KREUZER and COZZARELLI 1979; OOSTRA et al. 1980; KANO et al. 1981). However, the correlation is not a simple one. It was observed quite early that not all promoters respond in the same way to supercoiling. SMITH et al. (1978) found that transcription of the tryptophan (*trp*) operon integrated into the λ-phage genome is inhibited by nalidixic acid, novobiocin, and coumer-

mycin if it is under the control of the phage P_1 promoter but not if it is under the control of its own authentic promoter.

From another study, it was reported that the *E. coli* operons for maltose and lactose were more severely depressed by gyrase inhibitors than, for instance, the operons controlling threonine and tryptophan synthesis (SANZEY 1979). This result was taken to suggest that catabolite-sensitive promoters (maltose and lactose promoters) require supercoiling more than catabolite-insensitive promoters (threonine and tryptophan promotors).

The need for supercoiling can be reversed by promoter mutations. Thus, a mutation in the promoter of the *lac*-repressor gene (I^Q) and also a mutation in the promoter region of the *lac* operon (UV5) lead to increased expression upon inhibition of gyrase (SANZEY 1979). Further evidence for nonidentical effects of supercoiling on different promoters comes from expression studies of the plasmid pBR322 in *E. coli* minicells. Whereas novobiocin significantly depressed synthesis of the protein which mediates resistance against tetracycline, synthesis of the β-lactamase precursor protein was enhanced in the presence of the drug by a factor of two (GÓMEZ-EICHELMANN 1981).

A differential need for gyrase, or supercoiling, in transcription was also observed, although with conflicting results, for the expression of rRNA genes in *E. coli* cells. In some experiments, rRNA synthesis was found to depend on a high degree of DNA supercoiling (YANG et al. 1979; OOSTRA et al. 1980, 1981). In contrast, in another report, rRNA synthesis was suggested to be less dependent on DNA supercoiling than general RNA synthesis (WAHLE and MÜLLER 1980). No preferential stimulation of rRNA production was seen in vitro with purified gyrase (OOSTRA et al. 1981).

Complex effects of gyrase inhibition in *B. subtilis* on sporulation (VAZQUEZ-RAMOS and MANDELSTAM 1981) and on envelope protein synthesis and cell division in *E. coli* (HERRERO et al. 1982) have been interpreted as an involvement of gyrase in the transcription of genes specifically involved in these processes.

Most studies about the effects that supercoiling has on transcription focus on the role of gyrase. Since topoisomerase I also controls the conformation of DNA, it probably contributes to transcription activities in the cells, too. The properties of topoisomerase I-deficient (*top*A^-) mutants confirm this notion. In the original *top*A^- mutants, the most obvious single physiological effect was related to the expression of certain operons. Of four operons and genes tested, three had promoters sensitive to catabolite repression and one had a promoter unaffected by this regulation. In *top*A^- cells, the catabolite-sensitive promoters exhibited a higher level of expression than the insensitive promoter (for details see STERNGLANZ et al. 1981; DINARDO et al. 1982). The particular mutant strain, which was used, had a *top*A^- mutation with residual topoisomerase I activity and no compensating mutation in the gyrase genes. The degree of intracellular supercoiling in this mutant was slightly elevated over that of an isogenic *top*A^+ strain (DINARDO et al. 1982). The observed effect on transcription rates in the mutant is in agreement with the previously made suggestion according to which certain operons, among them catabolite-sensitive operons, are positively correlated with the degree of superhelicity within the cell (SANZEY

1979). It should be noted that WAHLE et al. (1984) reported data, which do not agree with this suggestion. These authors did not find a reduced activity of the *lac* promoter upon inhibition of a temperature-sensitive gyrase.

Analysis of the *top*A$^-$ mutants also provided an example for a negative correlation by the activation of a cryptic operon, the *bgl* operon, which controls the catabolism of β-glucosides. In normal *E. coli* K12 strains, this operon cannot be induced (PRASARD and SCHAEFLER 1974). A decrease in intracellular supercoiling in mutant strains which have a *top*A deletion and at the same time a less active gyrase due to a further mutation in the *gyr*B gene is accompanied by an activation of this normally silent operon (DINARDO et al. 1982).

A conclusion unifying the differential effects of DNA superhelicity on transcription cannot be drawn as yet. Promoter properties may respond to a loss or an increase in supercoiling. Local or general effects due to altered DNA association constants of represser molecules DNA are also conceivable. Recently, it was found that an impairment of transcription due to unfavorable supercoiling can be suppressed, for instance, by a compensatory mutation of the RNA polymerase subunit B (FILUTOWICZ and JONCZYK 1983).

Circularity of DNA is not always required for gyrase to function in transcription. Expression of late, but not of early, genes of the linear DNA of the phage T7 is in vivo efficiently suppressed after inhibition of gyrase (DE WYNGERT and HINKLE 1979). Stimulation of transcription with *E. coli* RNA polymerase by gyrase has also been observed in vitro with linear phage λDNA, but not with linear phage T7 DNA (SHMERLING and GRAGEROV 1982). Whether DNA-specific local supercoiling is involved, as it was suggested by the authors, is not known.

A particularly clear example for the role of supercoiling in transcription is the expression of the gyrase genes themselves. Increasing transcription rates have been observed in vivo for the two gyrase subunits after treatments that block gyrase activity or decrease the degree of intracellular supercoiling (MENZEL and GELLERT 1983). Thus the cell's instrument to control supercoiling is itself regulated by supercoiling (ADACHI et al. 1984).

Reduction of transcription by gyrase inhibitors is an effect on RNA chain initiation rather than on chain elongation. Evidence for this conclusion is based on the observation of sequential cessation of transcription in a polycistronic operon (*trp*) after addition of a gyrase inhibitor (oxolinic acid). Analysis of the polycistronic mRNA made after addition of the drug showed that promoter-proximal cistrons disappeared earlier than promoter-distant ones (KUBO et al. 1979). Further evidence for a role of gyrase in initiation was derived from in vitro transcription of ColE1 DNA. When RNA synthesis on supercoiled DNA templates is blocked by rifampicin, which specifically inhibits initiation, the growth of RNA chains, which were initiated prior to addition of rifampicin, is undisturbed in the presence of gyrase inhibitors (YANG et al. 1979).

The effect of oxolinic acid on inhibition of transcription may not only be due to a loss of supercoiling but also – as was suggested by KANO et al. (1981) – to trapping of the DNA gyrase complex in an intermediate form, a proposal similar to that made about the role of this drug in the inhibition of DNA replication.

DNA gyrase was discovered during the course of **recombination** studies as a component which activates DNA of the phage λ for site-specific integration into the *E. coli* chromosome (GELLERT et al. 1976a). The function of gyrase in this process is to supercoil the DNA molecule that carries the phage attachment site *att*P (see Sect. 9). The need for supercoiling can be bypassed by changing the reaction conditions (POLLOCK and ABREMSKI 1979); however, the efficiency of the process is always higher with supercoiled than with relaxed DNA. That superhelicity is a prerequisite also for general recombination in vivo may be concluded from inhibition of recombination in *E. coli* (HAYS and BOEHMER 1978) and *Streptococcus sanguis* cells (RAINA and RAVIN 1979) by nonlethal doses of coumermycin or oxolinic acid.

In addition, a type of *rec*A-independent, illegitimate recombination has been observed, which is apparently mediated by gyrase. This recombination promotes insertions or substitutions of plasmid DNA into λ-DNA in *E. coli* extracts (IKEDA et al. 1980, 1982). The recombinants were produced by a reciprocal strand exchange process with or without deletions in the component DNA molecules. This exchange reaction operates obviously with DNA molecules of heterologous origin. The analysis of a number of recombination sites of λ and pBR322 DNA revealed sequence homologies of not more than three base pairs (NAITO et al. 1984). It was found that oxolinic acid stimulates this process. With coumermycin the stimulation is abolished. Although a final conclusion on the mechanism of this reaction cannot be drawn as yet, it seems likely that intermediate DNA gyrase complexes are involved in the DNA strand exchange. The reaction occurs in the absence of oxolinic acid; however, stabilization of the intermediates by this drug leads to an increase in the frequency of this exchange event.

The ω-protein of *E. coli* can also – at least in vitro – be linked to recombination events. It was found by Radding and co-workers (CUNNINGHAM et al. 1981) that both, homologous pairing and topological linkage of separate DNA molecules is promoted in a combined action of *rec*A protein and topoisomerase I. The *rec*A protein promotes pairing of a single-stranded DNA with a complementary sequence in a duplex DNA to generate a joint in which the strands are base-paired, but not really interwound or linked. As mentioned earlier, this kind of pairing is called a paranemic joint (BIANCHI et al. 1983). The conversion of a paranemic to a plectonemic, or topologically linked, joint can be catalyzed by the ω-protein. This reaction could explain how a single-stranded DNA region is topologically linked to a duplex DNA region in the early steps of general recombination. The reaction occurs in a side-by-side fashion without need for a free end in any of the three strands, a result lending support to the notion that in general recombination duplex DNA molecules may pair before strand breakage occurs (WILSON 1979).

Topoisomerases also affect **transposition** frequencies. It was reported from a study with one of the first *top*A$^-$ mutants that the transposition frequency of the transposon Tn5 was decreased by a factor of 40 (STERNGLANZ et al. 1981). This result is in conflict with more recent data, which show that transposition of Tn5 is positively correlated with supercoiling of DNA and, hence, requires DNA gyrase. In a *topA*$^-$ strain (having a higher DNA superhelix density

than the corresponding wild-type strain), a slight increase in the transposition frequency of Tn5 has been observed (ISBERG and SYVANEN 1982). The discrepancy between these findings might at least be partially explained by the presence of compensatory mutations in the *topA*$^-$ strains of STERNGLANZ et al. (1981) (see unpublished data quoted by PRUSS et al. 1982).

That topoisomerase I controls **repair** functions to some extent can be concluded from the increased sensitivity of *top*A$^-$ cells to UV irradiation or treatment with methyl methanesulfonate (STERNGLANZ et al. 1981). A biochemical study with purified topoisomerase I from *M. luteus* has shown that this enzyme is inhibited in its relaxing activity by the presence of UV photoproducts (pyrimidine dimers) in supercoiled DNA molecules (PEDRINI and CIARROCCHI 1983). The mechanism of this inhibition is not clear. However, it suggests that repair deficiencies in *top*A$^-$ mutants are most probably not directly related just to a lack of the enzyme but to a more complicated failure, which may include deficient expression of genes involved in repair.

UV repair defects have also been observed in *S. typhimurium* strains lacking topoisomerase I. These mutants are called, as was mentioned earlier, *sup*X$^-$ or now *top*A$^-$. At least part of the increased UV sensitivity may be attributable to an increased level of *rec*BC DNase (exonuclease V), an enzyme degrading damaged DNA. Among the pleiotropic effects of the *sup*X$^-$ mutation was an almost total abolition of mutagenesis by UV or alkylating agents (OVERBYE et al. 1982).

A few reports indicate that gyrase also has a function in UV repair and mutagenesis in *E. coli* cells (HAYS and BOEHMER 1978; CRUMPLIN 1982). The evidence is based on the effects of mutations in *gyr*A or in *gyr*B. It should be noted that these reports are compatible with earlier results suggesting that gyrase is not involved in DNA repair synthesis in permeabilized cells (STAUDENBAUER 1975; RYAN and WELLS 1976). The repair polymerization process is most probably unaffected by gyrase inhibition, but viability of the cells requires a defined DNA conformation maintained by gyrase (CRUMPLIN 1981). It was recently observed that in *E. coli* cells a reduction in DNA superhelicity by *gyr*B mutations affects in vivo different repair capacities in nonidentical ways. Whereas excision repair processes are apparently enhanced compared with wild-type cells, postreplication repair operates with a reduced efficiency (VON WRIGHT and BRIDGES 1981).

If gyrase, which is an essential enzyme for the cell, is inhibited by oxolinic acid or coumermycin and DNA synthesis tops, the SOS system, the ultimate repair system of *E. coli* cells, is turned on. For the induction, the *rec*BC repair nuclease is needed. The induction with both inhibitors apparently responds to the DNA gyrase-inhibitor complexes rather than just to loss of gyrase activity (SMITH 1983).

A clue for the understanding of the role of topoisomerases in prokaryotic organisms resulted from the observation that in *E. coli* a simple deletion of the *top*A gene destroys viability of the cell. Carriers of *top*A deletions are viable only if a second mutation compensates for the loss of ω (DINARDO et al. 1982; PRUSS et al. 1982). Several of these compensatory mutations have been mapped at or near the gyrase genes *gyr*A and *gyr*B. Measurements of DNA supercoiling

by PRUSS et al. (1982) indicated a lower gyrase activity in some of these mutants. A *top*A deletion could not be transduced to strains with normal gyrase activity (DINARDO et al. 1982). It should be noted that incompatibility of an uncompensated *top*A gene deletion is not a general phenomenon. *S. typhimurium* cells tolerate the corresponding deletions without compensatory changes in their gyrase genes (RICHARDSON et al. 1984).

The data imply, first, at least in *E. coli*, that an excess of negative supercoiling due to a missing topoisomerase I is deleterious and, second, that the level of supercoiling is regulated within the cells. Excessive topoisomerase I activity is apparently less restrictive for the *E. coli* cells than a total and uncompensated lack of this enzyme. Overproduction of topoisomerase I in *E. coli*, which has been achieved by introducing extrachromosomal cloned *top*A gene copies into *E. coli* cells (WANG and BECHERER 1983), leads to a slower growth and smaller colonies on agar plates than obtained with normal cells. The intracellular level of topoisomerase I increased together with the copy number of the plasmids harboring the *top*A gene. This result suggests that the *top*A gene is not a target for strict regulation.

It is debatable whether gyrase and topoisomerase I are the only components active in regulating the degree of superhelicity. Other topoisomerases, topoisomerases III and II′, do exist, and they could also be involved in this regulation. DI NARDO et al. (1982) observed that reintroduction of an intact *top*A gene into a viable and, hence, gyrase-deficient *top*A deletion mutant did not further reduce the level of intracellular supercoiling. This result could indicate a regulatory scheme which is based on more than simple competition between gyrase and the ω-protein.

10.3 Eukaryotic Topoisomerases

The first identification of a type I topoisomerase in mouse cell extracts led immediately to the suggestion that this enzyme might function in vivo by providing a swivel during **replication** (CHAMPOUX and DULBECCO 1972). The need for a swiveling activity in DNA to facilitate unwinding of parental DNA strands was originally recognized in studies of the circular *E. coli* chromosome (CAIRNS 1963). This requirement was also deduced from studies of replicative SV40 DNA intermediates (SEBRING et al. 1971) and mitochondrial replication (BERK and CLAYTON 1974, 1976; KASAMATSU and VINOGRAD 1974). Replication of mitochondrial DNA has been briefly described (Sect. 7). Mitochondrial DNA does not form nucleosomes, the dissociation of which could produce some unrestrained negative supercoiling. Therefore, either a gyraselike enzyme preventing the formation of positive supercoils or a topoisomerase removing such supercoils would be required for the movement of the replication fork. Likewise, converging replication forks, which occur in bidirectionally replicating DNA molecules, would also require a swivel activity. Since replication involves either an excess of duplex winding ahead of the fork or preventive negative supercoiling before unwinding occurs, the ω-protein is no candidate for promotion of a swivel function in *E. coli* cells due to its inability to relax positive supercoils or to induce negative supercoiling. In contrast, eukaryotic type I topoisomerases,

which relax positive supercoils, may still be considered as potential replication factors in higher organisms. However, unequivocal evidence in favor of such a role does not exist.

In the absence of direct biochemical or genetic information, indirect attempts based on measuring topoisomerase activities during the cell cycle of sea urchin or mammalian cells have produced conflicting results. According to one report (ROSENBERG et al. 1976), synchronized human lymphocytes have an approximately 40-fold higher activity per cell in the S phase than in the G_0 or G_1 phase. A similar variation, although with lower amplitude, was found by POCCIA et al. (1978) in developing sea urchin embryos. In contrast, other data obtained with mouse cells suggested essentially no variation during the cell cycle (CHAMPOUX et al. 1978). In cells of regenerating rat liver, the topoisomerase I activity was also found to be invariant compared with nonregenerating liver cells. In this system, the topoisomerase II was found in higher activities per cell in regenerating liver (DUGUET et al. 1983). A general conclusion about the role of topoisomerase I in replication cannot be drawn from these data.

Association of topoisomerase with DNA-synthesizing nuclear membrane fragments has also been interpreted in terms of this enzyme being a replication factor (YOSHIDA et al. 1977). Since, however, the enzyme has in general a high binding affinity for DNA, the physiological significance of this association is not clear.

Two studies on topoisomerase-I-deficient mutants of *Saccharomyces cerevisiae* (yeast) (THRASH et al. 1984) and *Schizosaccharomyces pombe* (a fission yeast) (UEMURA and YANAGIDA 1984) both do not allow a clear interpretation about the role of topoisomerase I. In these studies, the yeast mutants had less than 1% of normal topoisomerase I activity, yet they grew normally. The only clearly recognized defect in yeast was loss of a double-stranded linear RNA, the killer RNA which has no essential function (for references see THRASH et al. 1984). These results do not indicate that topoisomerase I (*top*1) is dispensable, since residual activities could maintain essential functions and topoisomerase II (*top*2) may substitute for a deficient type I enzyme. Double mutants of the fission yeast which were temperature-sensitive in topoisomerase II showed a fast stop of growth under nonpermissive conditions, irrespective of the cell cycle stage. The authors (UEMURA and YANAGIDA 1984) inferred from a complex set of results that topoisomerase I might have a role in the control and maintenance of the chromatin structure throughout the cell cycle, whereas topoisomerase II is possibly involved in chromosome segregation at the end of replication prior to nuclear division.

It was reported from in vitro replication studies with adenovirus DNA that in addition to purified virus-coded proteins a nuclear protein of HeLa cells is required to initiate replication and to promote synthesis of adenovirus DNA in full length. The nuclear protein has topoisomerase I activity and can be replaced by a purified topoisomerase I from HeLa cells or calf thymus (NAGATA et al. 1983). It is, however, an unusually small protein ($\sim$30000 daltons). Thus, the nature os this enzyme is not quite clear. In any event, this study is the first to provide biochemical evidence for the suggested role of type I topoisomerases in replication.

A function of topoisomerase II in chromosome segregation was first considered by SUNDIN and VARSHAVSKY (1981) in their studies of SV40 replication. They suggested that the terminal replication step, which requires segregation of daughter molecules by passing one DNA helix through the other, is catalyzed by this enzyme. This conclusion is in good agreement with the results obtained recently with a temperature-sensitive topoisomerase II mutant of yeast (DINARDO et al. 1984). These authors showed that a recessive mutation of a single locus is responsible for temperature-sensitive growth and the enzymatic activity. Synchronized mutant cells can undergo one round of replication and then stop at the stage of nuclear division. The 2-μm plasmids from these cells were found in the form of dimeric catenated DNA circles. This result is highly suggestive of a requirement for topoisomerase II in the final stage of replication.

The essential character and the single copy nature of the topoisomerase II gene in yeast has very clearly been demonstrated in an elegant study by GOTO and WANG (1984). In this study, the *top*2 gene had been cloned and reintroduced into yeast cells in a truncated, nonfunctional version as a linear DNA molecule. The transformation resulted in a disruption of the endogeneous topoisomerase II gene due to recombination with the incoming DNA (ROTHSTEIN 1983). It was shown by tetrad analysis that integration of the nonfunctional exogenous *top*2 gene into the yeast genome occurred at the expense of viability of the affected progeny.

Additional, although rather indirect evidence for topoisomerase II being required for DNA replication comes from the biochemical analysis of a mutant mouse L cell line (ts A1S9) which is temperature-sensitive in replication (COLWILL and SHEININ 1983). The authors report that in this mutant probably not chain elongation is affected, but a mechanism which controls chromatin conformation and which is somehow coupled to replication. It was observed that nuclear novobiocin-binding proteins from normal cells contain a topoisomerase II activity, but the same protein fraction of mutant cells was devoid of such an activity when the cells had been kept at nonpermissive conditions prior to harvest. Based on electrophoretic analysis, it was deduced that a 30000-dalton protein is responsible for this loss of topoisomerase II activity. This protein was found in normal cells but not in the mutant cells under nonpermissive conditions. Whether this protein is a subunit (or a fragment) of the enzyme or an auxiliary protein is not known.

The only function which can be assigned at present to topoisomerase II with certainty is promotion of chromosome segregation. This, however, does not preclude a role of this enzyme in other stages of DNA replication also, e.g., in chain elongation. Topoisomerase I may substitute for a deficient topoisomerase II in chain elongation but not in termination of replication. In favor of a major and more general role of topoisomerase II in DNA replication is the observation, which was already mentioned, that in regenerating rat liver cells the type II activity per nucleus, but not the type I activity, is increased by a factor of at least ten, compared with unstimulated cells (DUGUET et al. 1983).

Not much is known about a function of eukaryotic topoisomerases in **transcription**. However, a few observations might be relevant. First, the nucleoli

of *X. laevis* which contain only ribosomal genes were found to be complexed with topoisomerase, presumably of type I, after isolation from oocytes (HIGASHINAKAGAWA et al. 1977). Since these nucleoli are involved in transcription only, but not in replication, this association could indicate that transient nicking and closing of DNA facilitates RNA chain growth, particularly in cases where transcription is initiated at a high rate in a tandem or head-to-head array (HAMKALO et al. 1973). Second, it was recently found that in vitro transcription of linear calf thymus DNA by eukaryotic RNA polymerase II is significantly stimulated by calf thymus topoisomerase I (BIALOJAN 1981). This stimulation, however, does not only depend on topoisomerase I, but also on other, thus far unidentified, chromatin components. Therefore, at present, these results cannot be interpreted in molecular terms.

Recently, GOCKE et al. (1983) reported on a site- and strand-specific nuclear activity which acts on rRNA genes of *Tetrahymena thermophila* in a manner reminiscent of topoisomerase I. Exposure of macronuclear chromatin of this organism to SDS causes cleavage at specific sites exclusively in the noncoding strand upstream from the transcription initiation point of the gene and close to known DNase-I-hypersensitive regions. The endogenous nuclease becomes covalently attached to the 3′ end of the nick, which is analogous to the result obtained with an abortive nicking reaction of topoisomerase I. Exogenous type topoisomerases of yeast, *Drosophila melanogaster*, chicken and calf thymus mimic this reaction with purified rDNA of *Tetrahymene thermophila*. Sequencing of ten of these recognition/cleavage sites has revealed a conserved quadrodecameric consensus sequence (for details see p. 38). Identical or closely related sequences of equal length were also found in flanking regions or rRNA genes of a number of other organisms including *Tetrahymena pyriformis*, *Dictyostelium discoideum*, *Saccharomyces cerevisiae* and *Drosophila melanogaster*. Moreover, these presumed topoisomerase I recognition sites were detected to frame a large number of genes transcribed by RNA polymerase I, II and III, respectively, in different eukaryotic species (Westergaard, personal communication). The occurrence of these sequences in flanking regions of transcribed genes and their close proximity to DNase hypersensitive sites were taken to consider the possibility that topoisomerase I might have a specific role in the transcription of at least certain genes (GOCKE et al. 1983; Westergaard, personal communication).

Finally, the investigation of *Drosophila* polytene chromosomes with purified antibodies against topoisomerase I revealed preferential association of this enzyme with transcriptionally active loci. This was shown by staining of puffs with immunofluorescent techniques (FLEISCHMANN et al. 1984). Topoisomerase I was also found in association with the nucleoli in these cells, a result reminiscent of the earlier observations made with *Xenopus* oocytes.

Two in vitro reactions of eukaryotic topoisomerases suggest that these enzymes might participate in **recombination** events. One reaction is based on the ability of these enzymes to intertwine complementary DNA strands which do not have free ends (CHAMPOUX 1977b), and the other reaction is interstrand transfer between covalent DNA enzyme complexes and DNA molecules carrying

free 5′-OH ends (BEEN and CHAMPOUX 1981; HALLIGAN et al. 1982; TRASK and MULLER 1983; reviewed by LIU 1983b).

The first reaction could trigger the initiation of strand exchange in general recombination by converting two homologous double strands into a four-strand synapsis (for details of a conceivable mechanism see CHAMPOUX 1977b). Conceptually, the reaction would also apply for a three-strand synapsis, as it could be conceived in connection with gene conversion events. A mechanism for local unwinding of the DNA sequences that participate in synapsis formation would be required in addition. Such a reaction would be similar in form to the in vitro cooperation of *rec*A and ω-protein of *E. coli* described above (CUNNINGHAM et al. 1981; BIANCHI et al. 1983).

In model studies with purified *rec*1 protein of the lower eukaryote *U. maydis* (a fungus) and topoisomerase I of the same organism, it has indeed been shown that a paranemic joint between two homologous DNA strands within a recombinational synapsis can be converted into a plectonemic joint by the topoisomerase I (KMIEC et al. 1983; KMIEC and HOLLOMAN 1984). The *rec*1 protein of *U. maydis* has a similar function in recombination as the *rec*A protein of *E. coli* (KMIEC and HOLLOMAN 1982).

Interstrand transfer reactions could theoretically promote illegitimate recombination, since no sequence homology is necessary. The only requirements for a reaction of this type are appropriate "donor" and "acceptor" DNA molecules (for details see Sect. 5.2). Since this reaction favors single-stranded DNA molecules, it is questionable whether it could occur in vivo.

Little is known about DNA **repair** functions of eukaryotic topoisomerases. It has been reported that novobiocin and nalidixic acid inhibit UV and other repair reactions in mammalian cells (COLLINS and JOHNSON 1979; MATTERN et al. 1982). These results presumably reflect the participation of a type II topoisomerase. That these enzymes are a target for a variety of antitumor drugs has been described in detail in Sect. 6.2. In addition, some complex observations have been reported, which may have a bearing on DNA-related functions of topoisomerases in higher organisms, including repair. They are quoted here, although no specific conclusions can be drawn as yet. It has been suggested that a recessive human genetic disease, Fanconi's anemia, is associated with an intracellular maldistribution of topoisomeraase I. One of the leading symptoms of this disease is chromosomal instability indicated by an increased frequency of DNA breaks, sister chromatid exchanges, and chromosomal translocations (SCHRÖDER 1982). In cells of the placenta of a newborn patient suffering from Fanconi's anemia, the topoisomerase I activity was detected to a considerable degree in the cytoplasm. In contrast, in normal placenta cells the enzyme was exclusively found in the nucleus (WUNDER et al. 1981). However, the same deficiency was not seen in fibroblast cells from patients with Fanconi's anemia (AUER et al. 1982).

A possibly related, though very preliminary observation has been reported according to which topoisomerase II is decreased in its normal activity in cells from patients affected by Bloom's syndrom, Fanconi's anemia, and other genetic disorders with an impaired DNA repair mechanism (BRELVI et al. 1983). Theo-

retically, it is conceivable that a lack of topoisomerase activities leads to chromosomal instability and other DNA-related deficiencies. It seems, however, premature to imply that these enzymes are involved in the pathogenesis of these disorders.

An important intranuclear process in which topoisomerases are most probably engaged is the **assembly of chromatin**. It was already mentioned that coiling of a topologically constrained DNA molecule around a histone core is accompanied by the occurrence of conformational strain, which arises in histone-free internucleosomal DNA regions during nucleosome formation. Thus, positive supercoiling is generated in compensation for negative left-handed supercoiling of DNA around the histone cores. The ability of eukaryotic type I topoisomerases to relax positive supercoils could, hence, facilitate nucleosome assembly. GERMOND et al. (1975, 1979) were the first to demonstrate in model studies with small circular DNAs that topoisomerase I can act in vitro in this way. This was later also shown by others (e.g., BALDI te al. 1978; NELSON et al. 1979, 1981). The arrangement of nucleosomes in these experiments resembled more or less closely that observed in native chromatin. It was suggested once that topoisomerase I itself might act as a factor promoting the assembly process directly (GERMOND et al. 1979), but this possibility was later excluded (NELSON et al. 1981; PULLEYBLANK and ELLISON 1982).

No comparable assembly experiments with histones have been documented with eukaryotic type II topoisomerases. WORCEL and his associates reported recently about a series of results (obtained with a different technique), which strongly suggest that topoisomerase II plays a critical role in nucleosome formation in vivo and also in vitro (RYOJI and WORCEL 1984; GLIKIN et al. 1984). These authors have followed the time course of chromatin assembly, DNA supercoiling, and transcription of a *Xenopus* 5 *S* RNA gene clone after injection into *X. laevis* oocytes. They found that the injected plasmid DNA is converted into two different types of chromatin. One is characterized by the absence of torsional strain, called "static" chromatin by the authors, and the other by the presence of such strain, called "dynamic" chromatin. The latter can be relaxed within the nuclei by injection of topoisomerase I or novobiocin. Relaxation prevents transcription of 5 *S* RNA, suggesting that only dynamic chromatin is transcriptionally active. In parallel in vitro experiments, the assembly has been shown to depend on ATP and Mg^{2+}. Moreover, it is inhibited by EDTA and novobiocin. These findings have led to the proposal that chromatin assembly is an active, ATP-driven process. Inhibition by novobiocin, which can be bypassed with an excess of ATP, shows that topoisomerase II is involved. The authors discussed the possibility that a gyraselike activity is responsible for the supercoiling of the dynamic chromatin. Although this suggestion may not be wrong, it is not compelling in the light of the experimental evidence. Neither the nature nor the number of factors is known which, in addition to topoisomerase II and histones, may take part in the formation of dynamic or static chromatin. It is, therefore, premature to speculate about the assembly mechanism. These experiments, however, set the frame for further analysis of chromatin assembly, DNA conformation, and gene expression in eukaryotes.

11 Supercoiling, DNA Structure, and Gene Expression

Two general categories can be discriminated among the numerous and divergent topoisomerase functions in pro- and eukaryotic organisms. The first one relates to those functions which require the action of a topoisomerase molecule at a particular region in the DNA. For instance, segregation of chromosomes as the terminal step in DNA replication requires breaking and rejoining of DNA in that region where the two daughter molecules are linked. All knotting and unknotting, catenation and decatenation events belong to this category of topoisomerase actions. This also applies to all recombination events, the conversion of paranemic into plectonemic joints, the illegitimate interstrand transfer (if it occurs in vivo), or other rearrangements. The existence of a gyrase binding site in the *ori*C region of the *E. coli* chromosome suggests a locally defined point of action for this enzyme in initiation of replication. If topoisomerase molecules were bound to travel together with the replication fork during elongation, this could also be considered as a confinement of the enzyme to a defined, although transient region. Whether the generation of nucleosomes in higher cells requires the presence of topoisomerase molecules in immediate proximity to nascent nucleosomes is not known, but would be conceivable.

The second category of topoisomerase functions includes a number of indirect effects. It has already been mentioned that topoisomerases can operate from the distance by using the DNA to transmit "allosteric" signals. The genetic and biochemical system which controls supercoiling in vivo has been extensively characterized for *E. coli* (see, e.g., GELLERT et al. 1982; MENZEL and GELLERT 1983) and to a similar degree also for *S. typhimurium* (RICHARDSON et al. 1984). Probably the most important control elements in this system are DNA gyrase (topoisomerase II) and ω-protein or its equivalents (topoisomerase I). It is likely that gyrase has the lead in this control circuit by actively introducing negative superhelical turns into DNA. The ω-protein modulates the extent of supercoiling by its relaxing potential. How the balance between the two counterproductive enzymes is controlled within the cells is not yet known. It is also not known whether additional factors, other topoisomerases, or DNA binding proteins are involved in the "homeostatic control of supercoiling" (MENZEL and GELLERT 1983). Strong indications that supercoiling is controlled were deduced from the properties of *top*A and *gyr* mutants of *E. coli*. Obviously, too much or not enough supercoiling are both disadvantageous for the cells. Two recent studies have convincingly shown that "fine tuning" in the control of supercoiling works very efficiently, either in response to a change in environmental conditions, e.g., temperature (GOLDSTEIN and DRLICA 1984), or as a consequence of topoisomerase mutations. These mutations change, one might say, the control values which define the degree of supercoiling (RICHARDSON et al. 1984).

The physiological functions of supercoiling are presumably compaction of DNA and modulation of gene expression in a *cis*-acting fashion. With respect to the latter, the expression of a normally silent *E. coli* operon (bgl) in *top*A mutants (DINARDO et al. 1982; WAHLE et al. 1984) and the autoregulation of the gyrase genes (MENZEL and GELLERT 1983) are particularly interesting exam-

ples of the regulatory potential of supercoiling. As mentioned, this control system works differentially; some operons are turned on by an increase in negative supercoiling and others by a decrease. Whether in prokaryotic cells different loci in separate domains of the chromosome can be controlled separately or whether the control acts uniformly on the chromosome as a whole is not known.

For eukaryotic organisms the picture is less clear. Organization of DNA in a compact form and regulation of gene activities are also essential objectives to be achieved. No gyraselike enzyme has been found so far. Topoisomerase-related genetic data are only available from yeast and, possibly, from mouse L cells. The involvement of topoisomerase II in the terminal steps of DNA replication and in chromatin assembly has been plausibly demonstrated. Most other functional assignments for topoisomerase I or II were derived from in vitro model studies or are at best circumstantial. It is reasonable to assume that eukaryotic topoisomerases, like their prokaryotic counterparts, play a role in all DNA-dependent events. Recent observations suggest that these enzymes are regulated in response to extracellular agonists. It was reported that topoisomerases can be induced by epidermal growth factor in human or mouse fibroblasts (MISKIMIS et al. 1983). In addition, epidermal growth factor receptor is able to nick supercoiled DNA in an ATP-stimulated manner (MROCZKOWSKI et al. 1984). No unambiguous conclusion can be drawn from these as from many other observations about topoisomerase-related effects in eukaryotes.

Much attention has been focussed recently on the role supercoiling may have for the control of gene expression in eukaryotes. Three lines of evidence support the underlying idea: (a) the above-mentioned influence of supercoiling on gene expression in prokaryotes, (b) the occurrence of DNase-I- and S1-nuclease-hypersensitive sites in the proximity of promoters or terminators of most active genes in eukaryotes (WEISBROD 1982b; WEINTRAUB 1983), and (c) the stabilization of altered DNA secondary structures by supercoiling. These structures include cruciforms in palindromic regions (LILLEY 1980; PANAYOTATOS and WELLS 1981; MIZUUCHI et al. 1982), left-handed DNA (Z-DNA) (JOVIN et al. 1982; RICH et al. 1984), and also the above-mentioned regions with increased sensitivity to DNase I, the single-strand-specific nuclease S1, and a single-strand-specific chemical reagent (WEINTRAUB 1983). These regions melt more easily than others. Stabilization of these secondary structures is due to the partial relief of torsional strain associated with the energetically unfavorable negative supercoiling.

DNA structures which exhibit hpyersensitivity against DNase I or S1 nuclease in pure supercoiled plasmids have been shown to act as *cis*-dominant elements after transfection into cultured mouse L cells. These hypersensitive sites are also seen in chromatin reconstituted in vitro with histones and supercoiled (but not linear) plasmid DNA (WEINTRAUB 1983). It has been suggested that these sites provide a signal in differentiated cells to guide the expression of certain genes. In cells not expressing these genes, the hypersensitive sites are presumably suppressed, possibly by DNA methylation or by higher order structures of the chromatin (WEINTRAUB 1983). Although cruciform structures have been demonstrated in vitro (MIZUUCHI et al. 1982; SINDEN and PETTIJOHN 1984), it is not clear whether they also exist in vivo and, if so, whether they have

a specific function (for a discussion see COUREY and WANG 1983). Z-DNA regions are found in natural chromatin (NORDHEIM et al. 1981; NORDHEIM and RICH 1983), and binding proteins which specifically recognize these regions have already been identified (NORDHEIM et al. 1982a). However, a defined biological role of Z-DNA has still to be assessed.

Since negative supercoiling appears to be the driving force for the formation of these altered secondary structures, the question arises as to what the superhelical status of chromatin in vivo is. The bulk of chromatin exists in a restrained form, as described in Sect. 10. It was, however, mentioned that at least subregions of the chromatin exist in a conformation which imposes torsional strain upon the DNA (LUCHNIK et al. 1982a, b; LARSEN and WEINTRAUB 1982; HARLAND et al. 1983). It seems plausible that the chromatin which has been called "dynamic chromatin" by RYOJI and WORCEL (1984) is identical with or closely related to this fraction of unrestrained chromatin. Experimental data suggest that transcription requires an unrestrained supercoiled DNA conformation in chromatin (see, e.g., RYOJI and WORCEL 1984).

It is not clear what the role of the two known eukaryotic topoisomerases in the formation of dynamic (or transcribed) or static (nontranscribed) chromatin is. The results of WORCEL and his co-workers (RYOJI and WORCEL 1984; GLIKIN et al. 1984) strongly invoke the requirement for topoisomerase II in the assembly of both types of chromatin. Whether these results reflect in reality, as the authors suggest, a gyraselike activity is uncertain. The participation of many other, hitherto unidentified factors in this process can be safely assumed.

In conclusion, prokaryotic and eukaryotic topoisomerases, which differ with respect to their structure and also in many details of their reaction mechanisms, have two important properties in common: first, they resolve DNA linkage problems wherever they occur in connection with DNA-dependent events, and second, they participate in the control of gene functions by controlling the conformation of DNA.

References

Adachi T, Mizuuchi K, Menzel R, Gellert M (1984) DNA sequence of the region upstream of the *E. coli gyr*B gene. Nucl Acids Res 12:6389–6395

Akrigg A, Cook PR (1980) DNA gyrase stimulates transcription. Nucl Acids Res 8:845–854

Auer B, Vosberg HP, Buhre U, Klocker H, Hirsch-Kauffmann M, Schweiger M (1982) Intracellular distribution of DNA topoisomerase I in fibroblasts from patients with Fanconi's anemia. Hum Genet 61:369–371

Baase WA, Wang JC (1974) An ω protein from *Drosophila melanogaster*. Biochemistry 13:4299–4303

Bachmann BJ, Low KB (1980) Linkage map of *Escherichia coli* K12, edition 6. Microbiol Rev 44:1–56

Badaracco G, Plevani P, Ruyechan WT, Chang LMS (1983) Purification and characterization of yeast topoisomerase I. J Biol Chem 258:2022–2026

Baldi MI, Mattocia E, Tocchini-Valentini GP (1978) DNA supercoiling by *Xenopus laevis* oocyte extracts: Requirement for a nuclear factor. Proc Natl Acad Sci USA 75:4873–4876

Baldi MI, Benedetti P, Mattocia E, Tocchini-Valentini GP (1980) In vitro catenation and decatenation of DNA and a novel eucaryotic ATP-dependent topoisomerase. Cell 20:461–467

Bauer WR (1978) Structure and reactions of closed duplex DNA. Ann Rev Biophys Bioeng 7:287–313

Bauer W, Vinograd J (1968) Interaction of closed circular DNA with intercalative dyes. I. The superhelix density of SV40 DNA in the presence and absence of dye. J Mol Biol 33:141–171

Bauer W, Vinograd J (1970) Interaction of closed circular DNA with intercalative dyes. II. The free energy of superhelix formation in SV40 DNA. J Mol Biol 47:419–435

Bauer WR, Ressner EC, Kates J, Patzke JV (1977) A DNA nicking-closing enzyme encapsidated in vaccinia virus: Partial purfication and properties. Proc Natl Acad Sci USA 74:1841–1845

Beck E, Sommer R, Auerswald EA, Kurz C, Zink B, Osterburg G, Schaller H (1978) Nucleotide sequence of bacteriophage fd DNA. Nucl Acids Res 5:4495–4503

Been MD, Champoux JJ (1980) Breakage of single-stranded DNA by rat liver nicking-closing enzyme with the formation of a DNA enzyme complex. Nucl Acids Res 8:6129–6143

Been MD, Champoux JJ (1981) DNA breakage and closure by rat liver type 1 topoisomerase: Separation of the half reactions by using a single-stranded DNA substrate. Proc Natl Acad Sci USA 78:2883–2887

Been MD, Burgess RR, Champoux JJ (1984a) DNA strand breakage by wheat germ type I topoisomerase. Biochim Biophys Acta 782:304–312

Been MD, Burgess RR, Champoux JJ (1984b) Nucleotide sequence preference at rat liver and wheat germ type I DNA topoisomerase breakage sites in duplex SV40 DNA. Nucl Acids Res 12:3097–3114

Benedetti P, Baldi MI, Mattoccia E, Tocchini-Valentini GP (1983) Purification and characterization of *Xenopus laevis* topoisomerase II. EMBO J 2:1303–1308

Benyajati C, Worcel A (1976) Isolation, characterisation and structure of the folded interphase genome of Drosophila melanogaster. Cell 9:393–407

Berk AJ, Clayton DA (1974) Mechanism of mitochondrial DNA replication in Mouse L cells: Asynchronous replication of strands, segregation of circular daughter molecules, aspects of topology and turnover of an initiation sequence. J Mol Biol 86:801–824

Berk AJ, Clayton CA (1976) Mechanism of mitochondrial DNA replication in mouse L cells: Topology of circular daughter molecules and dynamics of catenated oligomer formation. J Mol Biol 100:85–102

Berthold V, Geider K (1976) Interaction of DNA with DNA-binding proteins. The characterisation of protein HD from *Escherichia coli* and its nucleic acid complexes. Eur J Biochem 71:443–449

Bialojan C (1981) Einfluß der DNA-Topoisomerase I auf die *in vitro* Transcription mit RNA-Polymerase B aus *Drosophila melanogaster*. Diplomarbeit (Diploma thesis), University of Heidelberg, Faculty of Biology

Bianchi M, DasGupta C, Radding CM (1983) Synapsis and the formation of paranemic joints by *E. coli rec*A protein. Cell 34:931–939

Bina-Stein M, Vogel T, Singer DS, Singer MF (1976) H5 histone and DNA-relaxing enzyme of chicken erythrocytes. Interaction with superhelical DNA. J Biol Chem 251:7363–7366

Bina M, Beecher S, Blasquez V (1982) Stability and components of mature simian virus 40. Biochemistry 21:3057–3063

Biswal N, Feldan P, Levy CC (1983) A DNA topoisomerase activity copurifies with the DNA polymerase induced by herpes simplex virus. Biochim Biophys Acta 740:379–389

Bogdanova ES, Mirkin SM, Shmerling ZG (1982) Changed properties of the A subunit in DNA gyrase with a B subunit mutation. Mol Gen Genet 186:572–574

Bogenhagen PO, Clayton CA (1978) Mechanism of mitochondrial DNA replication in mouse L-cells: Introduction of superhelical turns into newly replicated molecules. J Mol Biol 119:69–81

Borst P, Hoeijmakers JHJ (1979) Kinetoplast DNA. Plasmid 2:20–40

Botchan P (1976) An electron microscopic comparison of transcription on linear and superhelical DNA. J Mol Biol 105:161–176

Botchan P, Wang JC, Echols H (1973) Effect of circularity and superhelicity on transcription from bacteriophage λ DNA. Proc Natl Acad Sci USA 70:3077–3081

Brahms S, Vergne J, Brahms JG, Di Capua E, Bucher P, Koller T (1982) Natural DNA sequences can form left-handed helices in low salt solution under conditions of topological constraint. J Mol Biol 162:473–493

Brelvi Z, Fernandez C, Studzinski GP (1983) Comparison of nucleolar DNA topoisomerase II-like activity in normal human cultured fibroblasts and genetic variants predisposing to neoplasia. Fed Proc 42:1295 (Abstract)

Brown PO, Cozzarelli NR (1979) A sign inversion mechanism for enzymatic supercoiling of DNA. Science 206:1081–1083
Brown PO, Cozzarelli NR (1981) Catenation and knotting of duplex DNA by type 1 topoisomerases: A mechanistic parallel with type 2 topoisomerases. Proc Natl Acad Sci USA 78:843–847
Brown PO, Peebles CL, Cozzarelli NR (1979) A topoisomerase from *Escherichia coli* related to DNA gyrase. Proc Natl Acad Aci USA 76:6110–6114
Brun G, Vannier P, Scovassi I, Callen JC (1981) DNA topoisomerase I from mitochondria of *Xenopus laevis* oocytes. Eur J Biochem 118:407–415
Burrington MG, Morgan AR (1978) The purification and characterization of a DNA nicking-closing enzyme from Bacillus megaterium. Can J Biochem 56:123–128
Cairns J (1963) The bacterial chromosome and its manner of replication as seen by autoradiography. J Mol Biol 6:208–213
Camerini-Otero RD, Felsenfeld G (1977) Supercoiling energy and nucleosome formation: The role of the arginine-rich histone kernel. Nucl Acids Res 5:1159–1181
Castora FJ, Simpson MV (1979) Search for a DNA gyrase in mammalian mitochondria. J Biol Chem 254:11193–11195
Castora FJ, Vissering FF, Simpson MV (1983) The effect of bacterial gyrase inhibitors on DNA synthesis in mammalian mitochondria. Biochim Biophys Acta 740:417–427
Champoux JJ (1976) Evidence for an intermediate with a single-strand break in the reaction catalysed by the DNA untwisting enzyme. Proc Natl Acad Sci USA 73:3488–3491
Champoux JJ (1977a) Strand breakage by the DNA untwisting enzyme results in covalent attachment of the enzyme to DNA. Proc Natl Acad Sci USA 74:3800–3804
Champoux JJ (1977b) Renaturation of complementary single-stranded DNA circles: Complete rewinding facilitated by the DNA untwisting enzyme. Proc Natl Acad Sci USA 74:5328–5332
Champoux JJ (1978) Proteins that effect DNA conformation. Ann Rev Biochem 47:449–479
Champoux JJ (1981) DNA is linked to the rat liver DNA nicking-closing enzyme by a phosphodiester bond to tyrosine. J Biol Chem 256:4805–4809
Champoux JJ, Dulbecco R (1972) An activity from mammalian cells that untwists superhelical DNA – A possible swivel for DNA replication. Proc Natl Acad Sci USA 69:143–146
Champoux JJ, McConaughy BL (1976) Purification and characterization of the DNA untwisting enzyme from rat liver. Biochemistry 15:4638–4642
Champoux JJ, Young LS, Been MD (1978) Studies on the regulation and specificity of the DNA-untwisting enzyme. Cold Spring Harbor Symp Quant Biol 43:53–58
Clayman CH, Mosharaffa ET, Anderson DI, Faras AJ (1979) Circular forms of DNA synthesized by Rous sarcoma virus in vitro. Science 206:582–584
Collins A, Johnson R (1979) Novobiocin, an inhibitor of the repair of UV-induced but not X-ray-induced damage in mammalian cells. Nucl Acids Res 7:1311–1320
Colwill RW, Sheinin R (1983) *ts* A1S9 locus in mouse L cells may encode a novobiocin binding protein that is required for DNA topoisomerase II activity. Proc Natl Acad Sci USA 80:4644–4648
Cook TM, Brown KG, Boyle JV, Goss WA (1966) Bactericidal action of nalidixic acid on *Bacillus subtilis*. J Bact 92:1510–1514
Cook PR, Brazell IA (1975) Supercoils in human DNA. J Cell Sci 19:261–279
Coombs DH, Pearson GD (1978) Filter-binding assay for covalent DNA-protein complexes: Adenovirus DNA-terminal protein complexes. Proc Natl Acad Sci USA 75:5291–5295
Courey AJ, Wang JC (1983) Cruciform formation in a negatively supercoiled DNA may be kinetically forbidden under physiological conditions. Cell 33:817–829
Cozzarelli NR (1980a) DNA gyrase and the supercoiling of DNA. Science 207:953–960
Cozzarelli NR (1980b) DNA topoisomerases. Cell 22:327–328
Craig NL, Nash HA (1983) The mechanism of phage λ site-specific recombination: Site-specific breakage of DNA by Int topoisomerase. Cell 35:795–803
Crick FHC (1976) Linking numbers and nucleosomes. Proc Natl Acad Sci USA 73:2639–2643
Crick FHC, Wang JC, Bauer WR (1979) Is DNA really a double helix? J Mol Biol 129:449–461
Crumplin GC, Smith JT (1976) Nalidixic acid and bacterial chromosome replication. Nature 260:643–645
Crumplin GC (1981) The involvement of DNA topoisomerases in DNA repair and mutagenesis. Carcinogenesis 2:157–160

Cunningham RP, Wu AM, Shibata T, DasGupta C, Radding CM (1981) Homologous pairing and topological linkage of DNA molecules by combined action of E. coli recA protein and topoisomerase I. Cell 24:213–223

Danilevskaya ON, Gragerov AJ (1980) Curing of *Escherichia coli* K12 plasmids by coumermycin. Molec Gen Genet 178:233–236

Darby MK, Vosberg H-P (1985) Relaxation of supercoiled phosphorothioate DNA by topoisomerase is inhibited in a base-specific manner. J Biol Chem (in press)

Davidson N (1972) Effect of DNA length on the free energy of binding of an unwinding agent to a supercoiled DNA. J Mol Biol 66:307–309

Dean F, Krasnow MA, Otter R, Matzuk MM, Spengler SJ, Cozzarelli NR (1982) *Escherichia coli* type-I topoisomerases: Identification, mechanism, and role in recombination. Cold Spring Harbor Symp Quant Biol 47:769–777

DeLeys RJ, Jackson DA (1976) Electrophoretic analysis of covalently closed SV40 DNA: Boltzmann distributions of DNA species. Nucl Acids Res 3:641–652

Depew RE, Wang JC (1975) Conformational fluctuations of the DNA helix. Proc Natl Acad Sci USA 72:4275–4279

Depew RE, Liu LF, Wang JC (1978) Interaction between DNA and *Escherichia coli* protein ω. Formation of a complex between single-stranded DNA and protein ω. J Biol Chem 253:511–518

Dermody JJ, Bourguignon GJ, Foglesong PD, Sternglanz R (1974) Nalidixic acid-sensitive and resistant modes of DNA replication in *Escherichia coli*. Biochem Biophys Res Comm 61:1340–1347

DeSantis P, Falcioni M, Morosetti S, Savino M (1983) DNA supercoiling induced by a synthetic polypeptide. Biopolymers 22:2517–2521

De Wyngaert MA, Hinkle DC (1979) Involvement of DNA gyrase in replication and transcription of bacteriophage T7 DNA. J Virol 29:529–535

DiNardo S, Voelkel KA, Sternglanz R, Reynolds AE, Wright A (1982) Escherichia coli DNA topoisomerase I mutants have compensatory mutations in DNA gyrase genes. Cell 31:43–51

DiNardo S, Voelkel K, Sternglanz R (1984) DNA topoisomerase II mutant of *Saccharomyces cerevisiae:* Topoisomerase II is required for segregation of daughter molecules at the termination of DNA replication. Proc Natl Acad Sci USA 81:2616–2620

Douc-Rasy S, Multon E, Kayser A, Riou G (1983) Inhibition par les derives de l'ellipticine des reactions catalysee par les topoisomerases: Inhibition preferentielle d'une topoisomerase II. C R Acad Sc Paris 296:899–904

Douc-Rasy S, Kayser A, Riou G (1984) Inhibition of the reactions catalysed by a type I topoisomerase and a catenating enzyme of *Trypnosoma cruzi* by DNA-intercalating drugs. Preferential inhibition of the catenating reaction. EMBO J 3:11–16

Drlica K, Snyder M (1978) Superhelical *Escherichia coli* DNA: Relaxation by coumermycin. J Mol Biol 120:145–154

Drlica K, Engle EC, Manes SH (1980) DNA gyrase on the bacterial chromosome: Possibility of two levels of action. Proc Natl Acad Sci USA 77:6879–6883

Duguet M (1981) Proteins that shape DNA. Biochimie 63:649–669

Duguet M, Bonne C, de Recondo AM (1981) Single-strand deoxyribonucleic acid binding protein from rat liver changes the helical structure of deoxyribonucleic acid. Biochemistry 20:3598–3603

Duguet M, Lavenot C, Harper F, Mirambeau G, De Recondo AM (1983) DNA topoisomerases from rat liver: physiological variations. Nucl Acids Res 11:1059–1075

Durban E, Roll D, Beckner G, Busch H (1981) Purification and characterization of a nuclear DNA-binding phosphoprotein in fetal and tumor tissues. Cancer Res 41:537–545

Durban E, Mills JS, Roll D, Busch H (1983) Phosphorylation of purified Novikoff Hepatoma topoisomerase I. Biochem Biophys Res Comm 111:897–905

Durnford JM, Champoux JJ (1978) The DNA untwisting enzyme from *Saccharomyces cerevisiae,* partial purification and characterization. J Biol Chem 253:1086–1089

Dynan WS, Jendrisak JJ, Hager DA, Burgess RR (1981) Purification and characterization of wheat germ DNA topoisomerase I (nicking-closing enzyme). J Biol Chem 256:5860–5865

Easterbrook-Smith SB, Wallace JC, Keech BD (1976) Pyruvate carboxylase: Affinity labelling of the magnesium adenosine triphosphate binding site. Eur J Biochem 62:125–130

Edenberg HJ (1980) Novobiocin inhibition of simian virus 40 DNA replication. Nature 286:529–531

Edwards KA, Halligan BD, Davis JL, Nivera NL, Liu LF (1982) Recognition sites of eukaryotic DNA topoisomerase I: DNA nucleotide sequencing analysis of topo I cleavage sites on SV40 DNA. Nucl Acids Res 10:2565–2576

Eisenberg S, Kornberg A (1979) Purification and characterization of φX174 gene A protein. J Biol Chem 254:5328–5332

Engle EC, Manes SH, Drlica K (1982) Differential effects of antibiotics inhibiting gyrase. J Bact 149:92–98

Eskin B, Morgan AR (1978) DNA nicking-closing activity from salmon testis. Can J Biochem 56:89–91

Fairfield FR, Bauer WR, Simpson MV (1979) Mitochondria contain a distinct DNA topoisomerase. J Biol Chem 254:9352–9354

Fairweather NF, Orr E, Holland IB (1980) Inhibition of deoxyribonucleic acid gyrase: Effects on nucleic acid synthesis and cell division in *Escherichia coli* K12. J Bact 142:153–161

Ferro AM, Higgins NP, Olivera BM (1983) Poly(ADP-ribosylation) of a DNA topoisomerase. J Biol Chem 258:6000–6003

Ferro AM, Olivera BM (1984) Poly(ADP-ribosylation) of DNA topoisomerase I from calf thymus. J Biol Chem 259:547–554

Filipski J (1983) Competitive inhibition of nicking-closing enzymes may explain some biological effects of DNA intercalators. FEBS Letters 159:6–12

Filutowicz M (1980) Requirement of DNA gyrase for the initiation of chromosome replication in *Escherichia coli*. Mol Gen Genet 177:301–309

Filutowicz M, Jonczyk P (1983) The *gyrB* gene product functions in both initiation and chain polymerization of *Escherichia coli* chromosome replication: Suppression of the initiation deficiency by a class of *rpoB* mutations. Mol Gen Genet 191:282–287

Finch JT, Lutter LC, Rhodes D, Brown RS, Rushton B, Levitt M, Klug A (1977) Structure of nucleosome core particles of chromatin. Nature 269:29–36

Fisher LM, Mizuuchi K, O'Dea MH, Ohmori H, Gellert M (1981) Site-specific interaction of DNA gyrase with DNA. Proc Natl Acad Sci USA 78:4165–4169

Fleischmann G, Pflugfelder G, Steiner EK, Javaherian K, Howard GC, Wang JC, Elgin SCR (1984) *Drosophila DNA topoisomerase I is* associated with transcriptionally active regions of the genome. Proc Natl Acad Sci USA 81:6958–6962

Foglesong PD, Bauer WR (1984) Effects of ATP and inhibitory factors on the activity of vaccinia virus type I topoisomerase. J Virol 49:1–8

Forterre P (1980) Model for the supercoiling reaction catalysed by DNA gyrase. J Theor Biol 82:255–269

Fukata H, Fukasawa H (1982) Isolation and partial characterisation of two distinct DNA topoisomerases from cauliflower inflorescence. J Biochem 91:1337–1342

Fuller FB (1978) Decomposition of the linking number of a closed ribbon: A problem from molecular biology. Proc Natl Acad Sci USA 75:3557–3561

Gandini-Attardi D, Martini G, Mattocia E, Tocchini-Valentini GP (1976) Effect of *Xenopus laevis* oocyte extract on supercoiled simian virus 40 DNA: Formation of complex DNA. Proc Natl Acad Sci USA 73:554–558

Gandini-Attardi D, De Paolis A, Tocchini-Valentini GP (1981) Purification and characterisation of *Xenopus laevis* type I topoisomerase. J Biol Chem 256:3654–3661

Geider K, Hoffmann-Berling H (1981) Proteins controlling the helical structure of DNA. Ann Rev Biochem 50:233–260

Geider K, Bäumel I, Meyer TF (1982) Intermediate stages in enzymatic replication of bacteriophage fd duplex DNA. J Biol Chem 257:6488–6493

Gellert M (1981a) DNA topoisomerases. Ann Rev Biochem 50:879–910

Gellert M (1981b) DNA gyrase and other type II topoisomerases. In: "The enzymes", Vol 14 (Boyer P, ed.), pp 345–366, Academic Press, New York

Gellert M, Mizuuchi K, O'Dea MH, Nash HA (1976a) DNA gyrase: An enzyme that introduces superhelical turns into DNA. Proc Natl Acad Sci USA 73:3872–3876

Gellert M, O'Dea MH, Itoh T, Tomizawa J (1976b) Novobiocin and coumermycin inhibit DNA supercoiling catalysed by DNA gyrase. Proc Natl Acad Sci USA 73:4474–4478

Gellert M, Mizuuchi K, O'Dea MH, Itoh T, Tomizawa J (1977) Nalidixic acid resistance: A second genetic character involved in DNA gyrase activity. Proc Natl Acad Sci USA 74:4772–4776

Gellert M, Fisher LM, O'Dea MH (1979) DNA gyrase: Purification and catalytic properties of a fragment of gyrase B protein. Proc Natl Acad Sci USA 76:6289–6293

Gellert M, Fisher LM, Ohmori H, O'Dea MH, Mizuuchi K (1980) DNA gyrase: Site-specific interactions and transient double-strand breakage of DNA. Cold Spring Harbor Symp Quant Biol 45:391–399

Gellert M, Menzel R, Mizuuchi K, O'Dea MH, Friedman DI (1982) Regulation of DNA supercoiling in *Escherichia coli*. Cold Spring Harbor Symp Quant Biol 47:763–767

Germond JE, Hirt B, Oudet P, Gross-Belard M, Chambon P (1975) Folding of the DNA double helix in chromatin-like structures from simian virus 40. Proc Natl Acad Sci USA 72:1843–1847

Germond JE, Rouvière-Yaniv J, Yaniv M, Brutlag D (1979) Nicking-closing enzyme assembles nucleosome-like structures *in vitro*. Proc Natl Acad Sci USA 76:3779–3783

Ghelardini P, Pedrini AM, Paolozzi L (1982) The topoisomerase activity of T4*am*G39 mutant is restored in Mu lysogens. FEBS Letters 137:49–52

Giacherio D, Hager LP (1980) A specific unwinding activity associated with SV40 large T antigen. J Biol Chem 255:8963–8966

Glaubiger D, Hearst JE (1967) Effect of superhelical structure on the secondary structure of DNA rings. Biopolymers 5:691–696

Glikin GC, Ruberti I, Worcel A (1984) Chromatin assembly in *Xenopus* oocytes: In vitro studies. Cell 37:33–41

Gocke E, Bonven BJ, Westergaard O (1983) A site and strand specific nuclease activity with analogies to topoisomerase I frames the rRNA gene of *Tetrahymena*. Nucl Acids Res 11:7661–7678

Goldstein E, Drlica K (1984) Regulation of bacterial DNA supercoiling: Plasmid linking numbers vary with growth temperature. Proc Natl Acad Sci USA 81:4046–4050

Gómez-Eichelmann MC (1981) Effect of nalidixic acid and novobiocin on pBR322 genetic expression in *Escherichia coli* minicells. J Bact 148:745–752

Goss WA, Cook TM (1975) Nalidixic acid – mode of action. In: "Antibiotics" (Corcoran JW, Hahn FE, eds) Vol III. Springer-Verlag, New York, pp 174–196

Goss WA, Deitz WH, Cook TM (1965) Mechanism of action of nalidixic acid on *Escherichia coli*. II. Inhibition of deoxyribonucleic acid synthesis. J Bact 89:1068–1074

Goto T, Wang JC (1982) Yeast DNA topoisomerase II. An ATP-dependent type II topoisomerase that catalyses the catenation, decatenation, unknotting, and relaxation of double-stranded DNA rings. J Biol Chem 257:5866–5872

Goto T, Wang JC (1984) Yeast DNA topoisomerase II is encoded by a single-copy, essential gene. Cell 36:1073–1080

Goto T, Laipis P, Wang JC (1984) The purification and characterization of DNA topoisomerase I and II of the yeast *Saccharomyces cerevisiae*. J Biol Chem 259:10422–10429

Gourlie BB, Pigiet V, Breaux CB, Krauss MR, King CR, Benbow RM (1981) Polyoma virus minichromosomes: Associated enzyme activities. J Virol 38:826–832

Gray HB, Upholt WB, Vinograd J (1971) A Buoyant method for the determination of the superhelix density of closed circular DNA. J Mol Biol 62:1–19

Griffith JD (1976) Visualization of procaryotic DNA in a regularly condensed chromatin-like fiber. Proc Natl Acad Sci USA 73:563–567

Hager DA, Burgess RR (1980) Elution of proteins from sodium dodecyl sulfate-polyacrylamide gels, removal of sodium dodecyl sulfate, and renaturation of enzymatic activity: Results with sigma subunit of Escherichia coli RNA polymerase, wheat germ DNA topoisomerase, and other enzymes. Anal Biochem 109:76–86

Halligan BD, Davis JL, Edwards KA, Liu LF (1982) Intra- and intermolecular strand transfer by HeLa DNA topoisomerase I. J Biol Chem 257:3995–4000

Hamatake RK, Mukai R, Hayashi M (1981) Role of DNA gyrase subunits in synthesis of bacteriophage φX174 viral DNA. Proc Natl Acad Sci USA 78:1532–1536

Hamelin C, Yaniv M (1979) Nicking-closing enzyme is associated with SV40 DNA in vivo as a sodium dodecyl sulfate resistent complex. Nucl Acids Res 7:679–687

Hamelin C, Yaniv M (1980) Coprecipitation of topoisomerase activity with simian virus 40 nucleoprotein complexes by divalent cations. Biochimie 62:261–265

Hamkalo BA, Miller OL, Bakken AH (1973) Ultrastructure of active eukaryotic genomes. Cold Spring Harbor Symp Quant Biol 38:915–919

Hane MW, Wood TH (1969) *Escherichia coli* K-12 mutants resistant to nalidixic acid: Genetic mapping and dominance studies. J Bact 99:238–241

Hansen FG, von Meyenburg K (1979) Characterisation of the *dna*A, *gyr*B and other genes in the *dna*A region of the *Escherichia coli* chromosome on specialised transducing phages λ *tna*. Mol Gen Genet 175:135–144

Harland RM, Weintraub H, McKnight SL (1983) Transcription of DNA injected into *Xenopus* oocytes is influenced by template topology. Nature 302:38–43

Hayashi Y, Hayashi M (1971) Template activities of the φX174 replicative allomorphic deoxyribonucleotiacid. Biochemistry 10:4212–4218

Hays JB, Boehmer S (1978) Antagonists of DNA gyrase inhibit repair and recombinantion of UV irradiated phage λ. Proc Natl Acad Sci USA 75:4125–4129

Hecht R, Thielmann HW (1977) DNA-relaxing enzyme from *Micrococcus luteus*. Nucl Acids Res 4:4235–4247

Henry TJ, Knippers R (1974) Isolation and function of the gene A initiator of bacteriophage φX174, a highly specific DNA endonuclease. Proc Natl Acad Sci USA 71:1549–1553

Herrero E, Fairweather NF, Holland IB (1982) Envelope protein synthesis and inhibition of cell division in *Escherichia coli* during inactivation of the B subunit of DNA gyrase. J Gener Microbiol 128:361–369

Higashinakagawa T, Wahn H, Reeder RH (1977) Isolation of ribosomal gene chromatin. Develop Biol 55:375–386

Higgins NP, Cozzarelli NR (1982) The binding of gyrase to DNA: Analysis by retention by nitrocellulose filters. Nucl Acids Res 10:6833–6847

Higgins NP, Peebles CL, Sugino A, Cozzarelli NR (1978) Purification of subunits of *Escherichia coli* DNA gyrase and reconstitution of enzymatic activity. Proc Natl Acad Sci USA 75:1773–1777

Hoegberg T, Khanna I, Drake SD, Mitscher LA, Shen LL (1984) Structure-activity relationships among DNA gyrase inhibitors. Synthesis and biological evaluation of 1,2-dihydro-4,4-dimethyl-1-oxo-2-naphthalenecarboxylic acids as 1-carba bioisosteres of oxolinic acid. J Med Chem 27:306–310

Hoess RH, Foeller C, Bidwell K, Landy A (1980) Site specific recombination functions of bacteriophage λ: DNA sequence of regulatory regions and overlapping structural genes for Int and Xis. Proc Natl Acad Sci USA 77:2482–2486

Hooper DC, Wolfson JS, McHugh GL, Winters MB, Swartz MN (1982) Effects of novobiocin, coumermycin A1, clorobiocin, and their analogs on *Escherichia coli* DNA gyrase and bacterial growth. Antimicrob Agents Chemother 22:662–671

Hsiang MW, Cole RD (1977) Structure of histone H1-DNA complex: Effect of histone H1 on DNA condensation. Proc Natl Acad Sci USA 74:4852–4856

Hsieh T (1983a) Purification and properties of type II DNA topoisomerase from embryos of *Drosophila melanogaster*. In: "Methods in Enzymology" (Wu R, Grossman L, Moldave K, eds) Vol 100, pp 161–170

Hsieh T (1983b) Knotting of the circular duplex DNA by type II DNA topoisomerase from *Drosophila melanogaster*. J Biol Chem 256:8413–8420

Hsieh T, Wang JC (1975) Thermodynamic properties of superhelical DNAs. Biochemistry 14:527–535

Hsieh T, Brutlag D (1980) ATP-dependent DNA topoisomerase from D. melanogaster reversibly catenates duplex DNA rings. Cell 21:115–125

Ikeda J, Yudelevich A, Shimamoto N, Hurwitz J (1979) Role of polymeric forms of the bacteriophage φX174 coded gene A protein in φXRFI DNA cleavage. J Biol Chem 254:9416–9428

Ikeda H, Moriya K, Matsumoto T (1980) In vitro study of illegitimate recombination: Involvement of DNA gyrase. Cold Spring Harbor Symp Quant Biol 45:399–402

Ikeda H, Aoki K, Naito A (1982) Illegitimate recombination mediated in vitro by DNA gyrase of *Escherichia coli:* Structure of recombinant DNA molecules. Proc Natl Acad Sci USA 79:3724–3728

Isberg RR, Syvanen M (1982) DNA gyrase is a host factor required for transposition of Tn5. Cell 30:9–18

Ishii K, Katase A, Andoh T, Seno N (1982) Inhibition of topoisomerase I by heparin. Biochem Biophys Res Comm 104:541–547

Ishii K, Hasegawa T, Fujisawa K, Andoh T (1983) Rapid purification and characterisation of DNA

topoisomerase I from cultured mouse mammary carcinoma FM3A cells. J Biol Chem 258:12728–12732
Itoh T, Tomizawa JI (1977) Involvement of DNA gyrase in bacteriophage T7 DNA replication. Nature 270:78–80
Iwabuchi M, Shibata T, Ohtani T, Natori M, Ando T (1983) ATP-dependent unwinding of the double helix and extensive supercoiling by *Escherichia coli* recA protein in the presence of topoisomerase. J Biol Chem 258:12394–12404
Javaherian K, Liu LF (1983) Association of eukaryotic DNA topoisomerase I with nucleosomes and chromosomal proteins. Nucl Acids Res 11:461–471
Javaherian K, Liu LF, Wang JC (1978) Nonhistone proteins HMG_1 and HMG_2 change the DNA helical structure. Science 199:1345–1346
Javaherian K, Tse YC, Vega J (1982) Drosophila topoisomerase I: Isolation, purification and characterization. Nucl Acids Res 10:6945–6955
Jongstra-Bilen J, Ittel M-E, Niedergang C, Vosberg H-P, Mandel P (1983) DNA topoisomerase I from calf thymus is inhibited *in vitro* by poly(ADP-ribosylation). Eur J Biochem 136:391–396
Jovin TM, van de Sande JH, Zarling DA, Arndt-Jovin DJ, Eckstein F, Füldner HH, Greider C, Grieger I, Hamori E, Kalisch B, McIntosh LP, Robert-Nicoud M (1982) Generation of left-handed Z-DNA in solution and visualization in polytene chromosomes by immunofluorescence. Cold Spring Harbor Symp Quant Biol 47:143–154
Kaguni JM, Kornberg A (1984a) Topoisomerase I confers specificity in enzymatic replication of the *Escherichia coli* chromosomal origin. J Biol Chem 259:8578–8583
Kaguni JM, Kornberg A (1984b) Replication initiated at the origin (*ori*C) of the *E. coli* chromosome reconstituted with purified enzymes. Cell 38:183–190
Kano Y, Miyashita T, Nakamura H, Kuroki K, Nagata A, Imamoto F (1981) In vivo correlation between DNA supercoiling and transcription. Gene 13:173–184
Kasamatsu H, Vinograd J (1974) Replication of circular DNA in eukaryotic cells. Ann Rev Biochem 43:695–719
Keller W (1975a) Characterization of purified DNA-relaxing enzyme from human tissue culture cells. Proc Natl Acad Sci USA 72:2550–2554
Keller W (1975b) Determination of the number of superhelical turns in simian virus 40 DNA by gel electrophoresis. Proc Natl Acad Sci USA 72:4876–4880
Keller W, Wendel I (1974) Stepwise relaxation of supercoiled SV40 DNA. Cold Spring Harbor Symp Quant Biol 39:199–208
Keller W, Müller U, Eicken I, Wendel I, Zentgraf H (1977) Biochemical and ultrastructural analysis of SV40 chromatin. Cold Spring Harbor Symp Quant Biol 42:227–243
Kikuchi Y, Nash HA (1979) Nicking-closing activity associated with bacteriophage λ *int* gene product. Proc Natl Acad Sci USA 76:3760–3764
Kikuchi A, Asai K (1984) Reverse gyrase – a topoisomerase which introduces positive superhelical turns into DNA. Nature 309:677–681
Kirkegaard K, Wang JC (1978) *Escherichia coli* DNA topoisomerase I catalysed linking of single-stranded rings of complementary base sequences. Nucl Acids Res 5:3811–3820
Kirkegaard K, Wang JC (1981) Mapping the topography of DNA wrapped around gyrase by nucleolytic and chemical probing of complexes of unique DNA sequences. Cell 23:721–729
Klevan L, Wang JC (1980) Deoxyribonucleic acid gyrase-deoxyribonucleic acid complex containing 140 base pairs of deoxyribonucleic acid and an $\alpha_2\beta_2$ protein core. Biochemistry 19:5229–5234
Klevan L, Tse Y (1983) Chemical modification of essential tyrosine residues in DNA topoisomerases. Biochim Biophys Acta 745:175–180
Klysik J, Stirdivant SM, Larson JE, Hart PA, Wells RD (1981) Lefthanded DNA in restriction fragments and a recombinant plasmid. Nature 290:672–677
Kmiec EB, Holloman WK (1982) Homologous pairing of DNA molecules promoted by a protein from *Ustilago*. Cell 29:367–374
Kmiec EB, Holloman WK (1984) Synapsis promoted by *Ustilago* rec1 protein. Cell 36:593–598
Kmiec EB, Kroeger PE, Brougham MJ, Holloman WK (1983) Topological linkage of circular DNA molecules promoted by *Ustilago* rec1 protein and topoisomerase. Cell 34:919–929
Kowalski D (1980) Fluorescence spot test for DNA endonuclease, ligase and topoisomerase activities. Anal Biochem 107:311–313
Krasnow MA, Cozzarelli NR (1982) Catenation of DNA rings by topoisomerases, mechanism of control by spermidine. J Biol Chem 257:2687–2693

Krasnow MA, Cozzarelli NR (1983) Site-specific relaxation and recombination by the Tn3 resolvase: Recognition of the DNA path between oriented *res* sites. Cell 32:1313–1324

Krauss MR, Gourlie BB, Bayne ML, Benbow RM (1984) Polyomavirus minichromosomes: Associated DNA topoisomerase II and DNA ligase activities. J Virol 49:333–342

Kreuzer KN (1984) Recognition of single-stranded DNA by the bacteriophage T4-induced type II topoisomerase. J Biol Chem 259:5347–5354

Kreuzer KN, Cozzarelli NR (1979) *Escherichia coli* mutants thermosensitive for deoxyribonucleic acid gyrase subunit A: Effects on deoxyribonucleic acid replication, transcription and bacteriophage growth. J Bact 140:424–435

Kreuzer KN, Cozzarelli NR (1980) Formation and resolution of DNA catenanes by DNA gyrase. Cell 20:245–254

Kreuzer KN, Jongeneel CV (1983) *Escherichia coli* phage T4 topoisomerase. In: "Methods in Enzymology" (Wu R, Grossman L, Moldave K, eds) Vol 100, pp 144–160

Kreuzer KN, Alberts BM (1984) Site-specific recognition of bacteriophage T4 DNA by T4 type II DNA topoisomerase and *Escherichia coli* gyrase. J Biol Chem 259:5339–5346

Kubo M, Kano Y, Nakamura H, Nagata A, Imamoto F (1979) *In vivo* enhancement of general and specific transcription in *Escherichia coli* by DNA gyrase activity. Gene 7:153–171

Kung VT, Wang JC (1977) Purification and characterization of an ω protein from *Micrococcus luteus*. J Biol Chem 252:5398–5402

Lampe MF, Bott KF (1984) Cloning of *gyr*A gene of *Bacillus subtilis*. Nucl Acids Res 12:6307–6323

Lang MC, Malfoy B, Freund AM, Daune M, Leng M (1982) Visualisation of Z sequences in form V of pBR322 by immuno-electron microscopy. The EMBO J 1:1149–1153

Langeveld SA, van Mansfeld ADM, de Winter JM, Weisbeek PJ (1979) Cleavage of single-stranded DNA by the A and A* proteins of bacteriophage φX174. Nucl Acids Res 7:2177–2188

Langeveld SA, van Arkel GA, Weisbeek PJ (1980) Improved method for the isolation of the A and A* proteins of bacteriophage φX174. FEBS Letters 114:269–272

Larsen A, Weintraub H (1982) An altered DNA conformation detected by S1 nuclease occurs at specific regions in active chick globin chromatin. Cell 29:609–622

Laskey RA, Mills AD, Morris NR (1977) Assembly of SV40 chromatin in a cell-free system from Xenopus eggs. Cell 10:237–243

Lau PP, Gray HB, Wei CF, Legerski RJ, Robberson DL (1981) Type I topoisomerases from mammalian cell nuclei interlock strands and promote renaturation of denatured close circular PM2 DNA. Biochim Biophys Acta 655:199–209

LeBon JM, Kado CI, Rosenthal LJ, Chirikjian JG (1978) DNA modifying enzymes of *Agrobacterium tumefaciens:* Effect of DNA topoisomerase, restriction endonuclease, and unique DNA endonuclease on plasmid and plant DNA. Proc Natl Acad Sci USA 75:4097–4101

LeBon JM, Agarwal S, Chirikjian JG (1981) DNA topoisomerase from *Agrobacterium tumefaciens:* Purification and catalytic properties. Nucl Acids Res 9:909–920

Lee C-H, Mizusawa H, Kakefuda T (1981) Unwinding of double-stranded DNA helix by dehydration. Proc Natl Acad Sci USA 78:2838–2842

LePecq JB, Paoletti C (1967) A fluorescent complex between ethidium bromide and nucleic acids. Physical-chemical characterization. J Mol Biol 27:87–106

Lilley DM (1980) The inverted repeat as a recognizable structural feature in supercoiled DNA molecules. Proc Natl Acad Sci USA 77:6468–6472

Liu LF (1980) DNA strand passing and the function of type II DNA topoisomerases. In: "Mechanistic Studies of DNA Replication and Genetic Recombination" (BM Alberts and CF Fox, eds), ICN-UCLA Symp Mol Cellular Biol, Vol 19. Academic Press, New York, pp 817–831

Liu LF (1983a) HeLa topoisomerase I. In "Methods in Enzymology" (Wu R, Grossman L, Moldave K, eds) Vol 100, pp 133–137

Liu LF (1983b) DNA topoisomerases – Enzymes that catalyze the breaking and rejoining of DNA. Crit Rev Biochem 15:1–24

Liu LF, Wang JC (1978a) *Micrococcus luteus* gyrase: Active components and a model for its supercoiling of DNA. Proc Natl Acad Sci USA 75:2098–2102

Liu LF, Wang JC (1978b) DNA-DNA gyrase complex: The wrapping of the DNA duplex outside the enzyme. Cell 15:979–984

Liu LF, Wang JC (1979) Interaction between DNA and *Escherichia coli* DNA topoisomerase I. Formation of complexes between the protein and superhelical and nonsuperhelical duplex DNAs. J Biol Chem 254:11082–11088

Liu LF, Miller KG (1981) Eukaryotic DNA topoisomerases: Two forms of type I DNA topoisomerases from HeLa cell nuclei. Proc Natl Acad Sci USA 78:3487–3491

Liu LF, Depew RE, Wang JC (1976) Knotted single-stranded DNA rings: A novel topological isomer of circular single-stranded DNA formed by treatment with *Escherichia coli* ω protein. J Mol Biol 106:439–452

Liu LF, Liu CC, Alberts BM (1979) T4 DNA topoisomerase: A new ATP-dependent enzyme essential for initiation of T4 bacteriophage DNA replication. Nature 281:456–461

Liu LF, Liu CC, Alberts BM (1980) Type II DNA topoisomerases: Enzymes that can unknot a topologically knotted DNA molecule via a reversible double-stranded break. Cell 19:697–707

Liu LF, Davis JL, Calendar R (1981a) Novel topologically knotted DNA from bacteriophage P4 capsids: Studies with DNA topoisomerases. Nucl Acids Res 9:3979–3989

Liu LF, Perkocha L, Calendar R, Wang JC (1981b) Knotted DNA from bacteriophage capsids. Proc Natl Acad Sci USA 78:5498–5502

Liu LF, Rowe TC, Yang L, Tewey KM, Chen GL (1983) Cleavage of DNA by mammalian topoisomerase II. J Biol Chem 258:15365–15370

Lockshon D, Morris DR (1983) Positively supercoiled plasmid DNA is produced by treatment of *Escherichia coli* with DNA gyrase inhibitors. Nucl Acids Res 11:2999–3017

Lother H, Lurz R, Orr E (1984) DNA binding and antigenic specifications of DNA gyrase. Nucl Acids Res 12:901–914

Louarn J, Bouche J-P, Patte J, Louarn J-M (1984) Genetic inactivation of topoisomerase I suppresses a defect in initiation of chromosome replication in *Escherichia coli*. Mol Gen Genet 195:170–174

Low RL, Kaguni JM, Kornberg A (1984) Potent catenation of supercoiled and gapped DNA circles by topoisomerase I in the presence of a hydrophilic polymer. J Biol Chem 259:4576–4581

Luchnik AN, Bakayev VV, Glaser VM (1982a) DNA supercoiling: Changes during cellular differentiation and activation of chromatin transcription. Cold Spring Harbor Symp Quant Biol 47:793–801

Luchnik AN, Bakayev VV, Zbarsky IB, Georgiev GP (1982b) Elastic torsional strain in DNA within a fraction of SV40 minichromosomes: Relation to transcriptionally active chromatin. EMBO J 1:1353–1358

Mandel P, Okazaki H, Niedergang C (1982) Poly(adenosine diphosphate ribose). Progr Nucl Acid Res Mol Biol 27:1–51

Marians KJ, Ikeda JE, Schlagman S, Hurwitz J (1977) Role of DNA gyrase in φX replicative-form replication in vitro. Proc Natl Acad Sci USA 74:1965–1968

Marini JC, Miller KG, Englund PT (1980) Decatenation of kinetoplast DNA by topoisomerases. J Biol Chem 255:4976–4979

Marshall B, Darkin S, Ralph RK (1983) Evidence that mAMSA induces topoisomerase action. FEBS Letters 161:75–78

Martin SR, McCoubrey WK Jr, Mc, McConaughy BL, Young LS, Been MD, Brewer BJ, Champoux JJ (1983) In "Methods in Enzymology" (Wu R, Grossman L, Moldave K, eds) Vol 100, pp 137–144

Mattern MR, Painter RB (1979) Dependence of mammalian DNA replication on DNA supercoiling. II. Effects of novobiocin on DNA synthesis in chinese hamster ovary cells. Biochim Biophys Acta 563:306–312

Mattern MR, Paone RF, Day III RS (1982) Eukaryotic DNA repair is blocked at different steps by inhibitors of DNA topoisomerases and of DNA polymerases α and β. Biochim Biophys Acta 697:6–13

Mattocia E, Gandini Attardi D, Tocchini-Valentini GP (1976) DNA-relaxing activity and enclonuclease activity in *Xenopus laevis* oocytes. Proc Natl Acad Sci USA 73:4551–4554

McCarthy D, Minner C, Bernstein H, Bernstein C (1976) DNA elongation rates and growing point distribution of wild-type phage T4 and a DNA-delay amber mutant. J Mol Biol 106:963–981

McCarthy D (1979) Gyrase-dependent initiation of bacteriophage T4 DNA replication: Interactions of *Escherichia coli* gyrase with novobiocin, coumermycin and phage DNA-delay gene products. J Mol Biol 127:265–283

McConaughy BL, Young LS, Champoux JJ (1981) The effect of salt on the binding of the eucaryotic DNA nicking-closing enzyme to DNA and chromatin. Biochim Biophys Acta 655:1–8

McGhee J, Felsenfeld G (1980) Nucleosome structure. Ann Rev Biochem 49:1115–1156

Menzel R, Gellert M (1983) Regulation of the genes for *E.* coli DNA gyrase: Homeostatic control of DNA supercoiling. Cell 34:105–113

Meyer TF, Geider K (1979a) Bacteriophage fd gene II protein. I. Purification, involvement in RF replication, and the expression of gene II. J Biol Chem 254:12636–12641
Meyer TF, Geider K (1979b) Bacteriophage fd gene II protein. II. Specific cleavage and relaxation of supercoiled RF from filamentous phages. J Biol Chem 254:12642–12646
Meyer TF, Geider K, Kurz C, Schaller H (1979) Cleavage site of bacteriophage fd gene II-protein in the origin of viral strand replication. Nature 278:365–367
Miller RV, Scurlock TR (1983) DNA gyrase (topoisomerase II) from *Pseudomonas aeruginosa*. Biochem Biophys Res Comm 110:694–700
Miller KG, Englund PT (1981) A topoisomerase induced by vaccinia virus. Fed Proc 40:1876 (Abstr)
Miller KG, Liu LF, Englund PT (1981) A homogeneous type II topoisomerase from HeLa cell nuclei. J Biol Chem 256:9334–9339
Mills JS, Busch H, Durban E (1982) Purification of a protein kinase from human Namalwa cells that phosphorylates topoisomerase I. Biochem Biophys Res Comm 109:1222–1227
Minocha A, Long BH (1984) Inhibition of the DNA catenation activity of type II topoisomerase by VP16-213 and VM26. Biochem Biophys Res Comm 122:165–170
Mirkin M, Shmerling ZG (1982) DNA replication and transcription in a temperature-sensitive mutant of *E.* coli with a defective DNA gyrase subunit b. Mol Gen Genet 188:91–95
Miskimins R, Miskimins WK, Bernstein H, Shimizu N (1983) Epidermal growth factor-induced topoisomerase(s). Intracellular translocation and relation to DNA synthesis. Exp Cell Res 146:53–62
Mizuuchi K, Mizuuchi M, O'Dea MH, Gellert M (1984) Cloning and simplified purification of *Escherichia coli* DNA gyrase A and B proteins. J Biol Chem 259:9199–9201
Mizuuchi K, O'Dea MH, Gellert M (1978) DNA gyrase: Subunit structure and ATPase activity of the purified enzyme. Proc Natl Acad Sci USA 75:5960–5963
Mizuuchi K, Fisher LM, O'Dea ML, Gellert M (1980) DNA gyrase action involves the introduction of transient double-strand breaks into DNA. Proc Natl Acad Sci USA 77:1847–1851
Mizuuchi K, Mizuuchi M, Gellert M (1982) Cruciform structures in palindromic DNA are favored by DNA supercoiling. J Mol Biol 156:229–243
Moore CL, Klevan L, Wang JC, Griffith JD (1983) Gyrase-DNA complexes visualized as looped structures by electron microscopy. J Biol Chem 258:4612–4617
Morgan AR, Pulleyblank DE (1974) Native and denatured DNA, cross-linked and palindromic DNA and circular covalently closed DNA analysed by a sensitive fluorometric procedure. Biochem Biophys Res Comm 61:396–403
Morrison A, Cozzarelli NR (1979) Site-specific cleavage of DNA by *E. coli* DNA gyrase. Cell 17:175–184
Morrison A, Cozzarelli NR (1981) Contacts between DNA gyrase and its binding site on DNA: Features of symmetry and asymmetry revealed by protection from nucleases. Proc Natl Acad Sci USA 78:1416–1420
Morrison A, Higgins NP, Cozzarelli NR (1980a) Interaction between DNA gyrase and its cleavage site on DNA. J Biol Chem 255:2211–2219
Morrison A, Brown PO, Kreuzer KN, Otter R, Gerrard SP, Cozzarelli NR (1980b) Mechanisms of DNA topoisomerases. In: "Mechanistic Studies of DNA Replication and Genetic Recombination" (BM Alberts and CF Fox, eds) ICN-UCLA Symp Mol Cellular Biol, Vol 19. Academic Press, New York, pp 785–807
Mroczkowski B, Mosig G, Cohen S (1984) ATP-stimulated interaction between epidermal growth factor receptor and supercoiled DNA. Nature 309:270–273
Müller U, Zentgraf H, Eicken I, Keller W (1978) Higher oder structure of simian virus 40 chromatin. Science 201:406–415
Nagata K, Guggenheimer RA, Hurwitz J (1983) Adenovirus DNA replication in vitro: Synthesis of full-length DNA with purified proteins. Proc Natl Acad Sci USA 80:4266–4270
Naito A, Naito S, Ikeda H (1984) Homology is not required for recombination mediated by DNA gyrase of *Escherichia coli*. Mol Gen Genet 193:238–243
Nakayama K, Sugino A (1980) Novobiocin and nalidixic acid target proteins in yeast. Biochem Biophys Res Comm 96:306–312
Nash HA (1981) Site-specific recombination protein of phage lambda. In: "The Enzymes", Vol 14 (Boyer P, ed). Academic Press, New York, pp 471–480
Nash HA, Robertson CA (1981) Purification and properties of the *Escherichia coli* protein factor required for λ integrative recombination. J Biol Chem 256:9246–9253

Nash HA, Pollock TJ (1983) Site-specific recombination of bacteriophage lambda. The change in topological linking number associated with exchange of DNA strands. J Mol Biol 170:19–38

Nash HA, Mizuuchi K, Enquist LW, Weisberg RA (1980) Strand exchange in λ integrative recombination: Genetics, biochemistry, and models. Gold Spring Harbor Symp Quant Biol 45:417–428

Nelson T, Hsieh TS, Brutlag D (1979) Extracts of *Drosophila* embryos mediate chromatin assembly in vitro. Proc Natl Acad Sci USA 76:5510–5514

Nelson T, Wiegand R, Brutlag D (1981) Ribonucleic acid and other polyanions facilitate chromatin assembly. Biochemistry 20:2594–2601

Nelson EM, Tewey KM, Liu LF (1984) Mechanism of antitumor drug action: Poisening of mammalian DNA topoisomerase II on DNA by 4′-(9-acridinylamino)-methanesulfon-m-anisidide. Proc Natl Acad Sci USA 81:1361–1365

Nordheim A, Rich A (1983) Negatively supercoiled simian virus 40 DNA contains Z-DNA segments within transcriptional enhancer sequences. Nature 303:674–679

Nordheim A, Pardue ML, Lafer EM, Möller A, Stollar BD, Rich A (1981) Antibodies to left-handed Z-DNA bind to interband regions of *Drosophila* polytene chromosomes. Nature 294:417–422

Nordheim A, Tesser P, Azorin F, Kwon YH, Möller A, Rich A (1982a) Isolation of *Drosophila* proteins that bind selectively to left-handed Z-DNA. Proc Natl Acad Sci USA 79:7729–7733

Nordheim A, Lafer EM, Peck LJ, Wang JC, Stollar BD, Rich A (1982b) Negatively supercoiled plasmids contain left-handed Z-DNA segments as detected by specific antibody binding. Cell 31:309–318

Ogasawara N, Seiki M, Yoshikawa H (1979) Effect of novobiocin on initiation of DNA replication in *Bacillus subtilis*. Nature 281:702–704

Ogasawara N, Seiki M, Yoshikawa H (1981) Initiation of DNA replication in *Bacillus subtilis*. V. Role of DNA gyrase and superhelical structure in initiation. Mol Gen Genet 181:332–337

Oostra BA, Ab G, Gruber M (1980) Involvement of DNA gyrase in the transcription of ribosomal RNA. Nucl Acids Res 8:4235–4246

Oostra BA, van Vliet AJ, Ab G, Gruber M (1981) Enhancement of ribosomal ribonucleic acid synthesis by deoxyribonucleic acid gyrase activity in *Escherichia coli*. J Bacteriol 148:782–787

Orr E, Fairweather NF, Holland B, Pritchard RH (1979) Isolation and characterisation of a strain carrying a conditional lethal mutation in the *cou* gene of *Escherichia coli* K12. Mol Gen Genet 177:103–112

Orr E, Staudenbauer WL (1981) An *Escherichia coli* mutant thermosensitive in the B subunit of DNA gyrase: Effect on the structure and replication of the colicin E1 plasmid in vitro. Mol Gen Genet 181:52–56

Orr E, Staudenbauer WL (1982) *Bacillus subtilis* DNA gyrase: Purification of subunits and reconstitution of supercoiling activity. J Bacteriol 151:524–527

Osheroff N, Shelton ER, Brutlag DL (1983) DNA topoisomerase II from Drosophila melanogaster. J Biol Chem 258:9536–9543

Otter R, Cozzarelli NR (1983) *Escherichia coli* DNA gyrase. In: "Methods in Enzymology" (Wu R, Grossman L, Moldave K, eds) Vol 100, pp 171–180

Overbye KM, Margolin P (1981) Role of the *sup*X gene in ultraviolet light-induced mutagenesis in *Salmonella typhimurium*. J Bact 146:170–178

Overbye KM, Basu SK, Margolin P (1982) Loss of DNA topoisomerase I activity alters many cellular functions in Salmonella typhimurium. Cold Spring Harbor Symp Quant Biol 47:785– 791

Panayotatos N, Wells RD (1981) Cruciform structures in supercoiled DNA. Nature 289:466–470

Peck LJ, Nordheim A, Rich A, Wang JC (1982) Flipping of cloned d(pCpG) $d(pCpG)_n \cdot d(pCpG)_n$ DNA sequences from right- to left-handed helical structure by salt, Co(III), or negative supercoiling. Proc Natl Acad Sci USA 79:4560–4564

Pedrini AM, Ciarrocchi G (1983) Inhibition of *Micrococcus luteus* DNA topoisomerase I by UV photoproducts. Proc Natl Acad Sci USA 80:1787–1791

Pedrini AM, Geroldi D, Siccardi A, Falaschi A (1972) Studies on the mode of action of nalidixic acid. Eur J Biochem 25:359–365

Peebles CL, Higgins NP, Kreuzer KN, Morrison A, Brown PO, Sugino A, Cozzarelli NR (1978) Structure and activities of *Escherichia coli* DNA gyrase. Cold Spring Harbor Symp Quant Biol 43:41–52

Pettijohn DE, Pfenninger O (1980) Supercoils in prokaryotic DNA restrained *in vivo*. Proc Natl Acad Sci USA 77:1331–1335

Piñon R, Salts Y (1977) Isolation of folded chromosomes from the yeast *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 74:2850–2854

Poccia DL, LeVine D, Wang JC (1978) Activity of a DNA topoisomerase (nicking-closing enzyme) during sea urchin development and the cell cycle. Develop Biol 64:273–283

Pohl FM, Jovin TM (1972) Salt-induced cooperative conformational change of a synthetic DNA: Equilibrium and kinetic studies with poly(dG-dC). J Mol Biol 67:375–396

Pohl FM, Thomae R, Di Capua E (1982) Antibodies to Z-DNA interact with form V DNA. Nature 300:545–546

Pollock TJ, Abremski K (1979) DNA without supertwists can be an *in vitro* substrate for site-specific recombination of bacteriophage λ. J Mol Biol 131:651–654

Pollock TJ, Nash HA (1983) Knotting of DNA caused by a genetic rearrangement. Evidence for a nucleosome-like structure in site-specific recombination of bacteriophage lambda. J Mol Biol 170:1–18

Pommier Y, Mattern MR, Schwartz RE, Zwelling LA (1984a) Absence of swivelling at sites of intercalator-induced protein-associated deoxyribonucleic acid strand breaks in mammalian cell nucleoids. Biochemistry 23:2922–2927

Pommier Y, Schwartz RE, Kohn KW, Zwelling LA (1984b) Formation and rejoining of deoxyribonucleic acid double-strand breaks induced in isolated cell nuclei by antineoplastic intercalating agents. Biochemistry 23:3194–3201

Prasard I, Schaefer S (1974) Regulation of the β-glucoside system in *Escherichia coli* K-12. J Bact 120:638–650

Prell B (1980) Untersuchungen zum Reaktionsmechanismus der DNA-Topoisomerase I aus Kalbsthymus. PhD-Thesis, University of Heidelberg

Prell B, Vosberg HP (1980) Analysis of covalent complexes formed between calf thymus DNA topoisomerase and single-stranded DNA. Eur J Biochem 108:389–398

Pruss GJ, Manes SH, Drlica K (1982) *Escherichia coli* DNA topoisomerase I mutants: Increased supercoiling is corrected by mutations near gyrase genes. Cell 31:35–42

Pulleyblank DE, Morgan AR (1975a) The sense of naturally occurring superhelices and the unwinding angle of intercalated ethidium. J Mol Biol 91:1–13

Pulleyblank DE, Morgan AR (1975b) Partial purification of "ω" protein from calf thymus. Biochemistry 14:5206–5209

Pulleyblank DE, Ellison MJ (1982) Purification and properties of type I topoisomerase from chicken erythrocytes: Mechanism of eukaryotic topoisomerase action. Biochemistry 21:1155–1161

Pulleyblank DE, Shure M, Tang D, Vinograd J, Vosberg HP (1975) Action of the nicking-closing enzyme on supercoiled and non-supercoiled closed circular DNA: Formation of a Boltzmann distribution of topological isomers. Proc Natl Acad Sci USA 72:4280–4284

Radloff R, Bauer W, Vinograd J (1967) A dye-buoyant-density method for the detection and isolation of closed circular duplex DNA: The closed circular DNA in HeLa cells. Proc Natl Acad Sci USA 57:1514–1521

Raina JL, Ravin AW (1979) Superhelical DNA in *Streptococcus sanguis:* Role in recombination in vivo. Mol Gen Genet 176:171–181

Reed RR (1983) The resolvase protein of the transposon $\gamma\delta$. In: "Methods in Enzymology" (Wu R, Grossman L, Moldave K, eds) Vol 100, pp 191–196

Renz M, Nehls P, Hozier J (1977) Involvement of histone H1 in the organization of the chromosome fiber. Proc Natl Acad Sci USA 74:1879–1883

Rich A, Nordheim A, Wang AHJ (1984) The chemistry and biology of left-handed Z-DNA. Ann Rev Biochem 53:791–846

Richardson JP (1975a) Initiation of transcription by *Escherichia coli* RNA polymerase from supercoiled and non-supercoiled bacteriophage PM2-DNA. J Mol Biol 91:477–487

Richardson JP (1975b) Attachment of nascent RNA molecules to superhelical DNA. J Mol Biol 98:565–579

Richardson SMH, Higgins CF, Lilley DMJ (1984) The genetic control of DNA supercoiling in *Salmonella typhimurium*. EMBO J 3:1745–1752

Riou GF, Gabillot M, Douc-Rasy S, Kayser A, Barrois M (1983) A type I topoisomerase from *Trypanosoma cruzi*. Eur J Biochem 134:479–484

Robberson DL, Kasamatsu H, Vinograd J (1972) Replication of mitochondrial DNA. Circular replicative intermediates in mouse L cells. Proc Natl Acad Sci USA 69:737–741

Rosenberg BH, Ungers G, Deutsch JF (1976) Variation in DNA swivel enzyme activity during the mammalian cell cycle. Nucl Acids Res 3:3305–3311

Ross W, Landy A (1983) Patterns of λ Int recognition in the regions of strand exchange. Cell 33:261–272

Ross WE, Bradley MO (1981) DNA double-strand breaks in mammalian cells after exposure to intercalating agents. Biochim Biophys Acta 654:129–134

Ross WE, Glaubiger D, Kohn KW (1979) Qualitative and quantitative aspects of intercalator-induced DNA strand breaks. Biochim Biophys Acta 562:41–50

Ross CF, Brougham MJ, Holloman WK, Ross WE (1983) Properties of a purified nuclear topoisomerase from L1210 cells. Biochim Biophys Acta 741:230–236

Rothstein RJ (1983) One-step gene disruption in yeast. In: "Methods in Enzymology" (Wu R, Grossman L, Moldave K, eds) Vol 101, pp 202–211

Rouvière-Yaniv J, Gros F (1975) Characterization of a novel, low-molecular-weight DNA-binding protein from *Escherichia coli*. Proc Natl Acad Sci USA 72:3428–3432

Rouvière-Yaniv J, Yaniv M, Germond JE (1979) *E. coli* DNA binding protein HU forms nucleosome-like structure with circular double-stranded DNA. Cell 17:265–274

Rowe TC, Tewey KM, Liu LF (1984) Identification of the breakage-reunion subunit of T4 DNA topoisomerase. J Biol Chem 259:9177–9181

Rowe TC, Rusche JR, Brougham MJ, Holloman WK (1981) Purification and properties of a topoisomerase from *Ustilago maydis*. J Biol Chem 256:10354–10361

Ruiz-Carillo A, Jorcano JL, Eder G, Lurz R (1979) *In vitro* core particle and nucleosome assembly at physiological ionic strength. Proc Natl Acad Sci USA 76:3284–3288

Ryan MJ (1976) Coumermycin A1: A preferential inhibitor of replicative DNA synthesis in *Escherichia coli*. I. In vivo characterization. Biochemistry 15:3769–3777

Ryan MJ, Wells RD (1976) Coumermycin A1: A preferential inhibitor of replicative DNA synthesis in *Escherichia coli*. II. In vitro characterization. Biochemistry 15:3778–3782

Ryoji M, Worcel A (1984) Chromatin assembly in *Xenopus* oocytes: In vivo studies. Cell 37:21–32

Sander M, Hsieh T (1983) Double strand DNA cleavage by type II DNA topoisomerase from *Drosophila melanogaster*. J Biol Chem 258:8421–8428

Sanger F, Air GM, Barrell BG, Brown NL, Coulson AR, Fiddes JC, Hutchison III CA, Slocombe PM, Smith M (1977) Nucleotide sequence of bacteriophage φX174 DNA. Nature 265:687–695

Sanger F, Coulson AR, Friedmann T, Air GM, Barrell BG, Brown NL, Fiddes JC, Hutchison III CA (1978) The nucleotide sequence of bacteriophage φX174. J Mol Biol 125:225–246

Sanhueza S, Eisenberg S (1984) Cleavage of single-stranded DNA by the φX174 A* protein: The A*-single-stranded DNA covalent linkage. Proc Natl Acad Sci USA 81:4285–4289

Sanzay B (1979) Modulation of gene expression by drugs affecting DNA gyrase. J Bact 138:40–47

Saucier JM, Wang JC (1972) Angular alteration of the DNA helix by *E. coli* RNA polymerase. Nature New Biol 239:167–170

Schmitt B, Buhre U, Vosberg H-P (1984) Characterisation of size variants of type I topoisomerase isolated from calf thymus. Eur J Biochem 144:127–134

Schöler HR (1982) Untersuchungen zum Wirkungsmechanismus der *Eco* DNA Topoisomerase I (ω). Diplomarbeit (Diploma thesis), University of Heidelberg, Faculty of Biology

Schroeder TM (1982) Genetically determined chromosome instability syndromes. Cytogenet Cell Genet 33:119–132

Seasholtz AF, Greenberg GR (1983) Identification of bacteriophage T4 gene 60 product and a role for this protein in DNA topoisomerase. J Biol Chem 258:1221–1226

Sebring ED, Kelly TJ Jr, Thoren MM, Salzman NP (1971) Structure of simian virus 40 deoxyribonucleotide acid molecules. J Virol 8:478–490

Seeburg PH, Nüsslein C, Schaller H (1977) Interaction of RNA polymerase with promoters from bacteriophage fd. Eur J Biochem 74:107–113

Sen A, Levine AJ (1974) SV40 nucleoprotein complex activity unwinds superhelical turns in SV40 DNA. Nature 249:343–344

Shelton ER, Osheroff N, Brutlag DL (1983) DNA topoisomerase II from *Drosophila melanogaster*. Purification and physical characterization. J Biol Chem 258:9530–9535

Shishido K, Ando T (1979) Purification and characterisation of DNA-relaxing enzyme from *Haemophilus gallinarum*. Biochim Biophys Acta 563:261–265

Shishido K, Noguchi N, Ando T (1983) Correlation of enzyme-induced cleavage sites on negatively

superhelical DNA between prokaryotic topoisomerase I and S1 nuclease. Biochim Biophys Acta 740:108–117

Shlomai J, Zadok A (1983) Reversible decatenation of kinetoplast DNA by a DNA topoisomerase from trypanosomatids. Nucl Acids Res 11:4019–4034

Shmerling ZG, Gragerov AI (1982) The influence of DNA gyrase on the transcription of linear DNA *in vitro*. FEBS Lett 140:260–262

Shure M, Vinograd J (1976) The number of superhelical turns in native virion SV40 DNA and minocol DNA determined by the band counting method. Cell 8:215–226

Shure M, Pulleyblank DE, Vinograd J (1977) The problems of eukaryotic and prokaryotic DNA packaging and in vivo conformation posed by superhelix density heterogeneity. Nucl Acids Res 4:1183–1205

Siedlecki J, Zimmermann W, Weissbach A (1983) Characterization of a prokaryotic topoisomerase I activity in chloroplast extracts from spinach. Nucl Acids Res 11:1523–1536

Sinden RR, Pettijohn DE (1984) Cruiform transitions in DNA. J Biol Chem 259:6593–6600

Sinden RR, Carlson JO, Pettijohn DE (1980) Torsional tension in the DNA double helix measured with trimethylpsoralen in living *E. coli* cells: Analogous measurements in insect and human cells. Cell 21:773–783

Smith CL (1983) *refF*-dependent induction of *recA* synthesis by coumermycin, a specific inhibitor of the B subunit of DNA gyrase. Proc Natl Acad Sci USA 80:2510–2513

Smith DH, Davies BD (1967) Mode of action of novobiocin in *Escherichia coli*. J Bact 93:71–79

Smith CL, Kubo M, Imamoto F (1978) Promoter specific inhibition of transcription by antibiotics which act on DNA gyrase. Nature 275:420–423

Snyder M, Drlica K (1979) DNA gyrase on the bacterial chromosome: DNA cleavage induced by oxolinic acid. J Mol Biol 131:287–302

Srivenugopal KS, Lockshon D, Morris DR (1984) Escherichia coli DNA topoisomerase III: Purification and characterization of a new type I enzyme. Biochemistry 23:1899–1906

Staudenbauer WL (1975) Novobiocin – a specific inhibitor of semiconservative DNA replication in permeabilized *Escherichia coli* cells. J Mol Biol 96:201–205

Staudenbauer WL (1976a) Replication of *Escherichia coli* DNA *in vitro:* Inhibition by oxolinic acid. Eur J Biochem 62:491–497

Staudenbauer WL (1976b) Replication of small plasmids in extracts of *Escherichia coli*. Mol Gen Genet 145:273–280

Staudenbauer WL, Orr E (1981) DNA gyrase: Affinity chromatography on novobiocin-sepharose and catalytic properties. Nucl Acids Res 9:3589–3602

Steck TR, Drlica K (1984) Bacterial chromosome segregation: Evidence for DNA gyrase involvement in decatenation. Cell 36:1081–1088

Stein A (1980) DNA wrapping in nucleosomes. The linking number problem re-examined. Nucl Acids Res 8:4803–4820

Sternbach H, Engelhardt R, Lezius AG (1975) Rapid isolation of highly active RNA polymerase from *Escherichia coli* and its subunits by matrix-bound heparin. Eur J Biochem 60:51–55

Sternglanz R, DiNardo S, Voelkel KA, Nishimura Y, Hirota Y, Becherer K, Zumstein L, Wang JC (1981) Mutations in the gene coding for *Escherichia coli* DNA topoisomerase I affect transcription and transposition. Proc Natl Acad Sci USA 78:2747–2751

Stetler GL, King GJ, Huang WM (1979) T4 DNA-delay proteins, required for specific DNA replication, form a complex that has ATP-dependent DNA topoisomerase activity. Proc Natl Acad Sci USA 76:3737–3741

Stettler UH, Weber H, Koller T, Weissmann C (1979) Preparation and characterization of form V DNA, the duplex DNA resulting from association of complementary, circular single-stranded DNA. J Mol Biol 131:21–40

Stonington GO, Pettijohn DE (1971) The folded genome of *Escherichia coli* isolated in a protein-DNA-RNA complex. Proc Natl Acad Sci USA 68:6–9

Sugino A, Cozzarelli NR (1980) The intrinsic ATPase of DNA gyrase. J Biol Chem 255:6299–6306

Sugino A, Bott K (1980) *Bacillus subtilis* deoxyribonucleic acid gyrase. J Bact 141:1331–1339

Sugino A, Peebles CL, Kreuzer KN, Cozzarelli NR (1977) Mechanism of action of nalidixic acid: Purification of *Escherichia coli nal*A gene product and its relationship to DNA gyrase and a novel nicking-closing enzyme. Proc Natl Acad Sci USA 74:4767–4771

Sugino A, Higgins NP, Brown PO, Peebles CL, Cozzarelli NR (1978) Energy coupling in DNA gyrase and the mechanism of action of novobiocin. Proc Natl Acad Sci USA 75:4838–4842

Sugino A, Higgins NP, Cozzarelli NR (1980) DNA gyrase subunit stoichiometry and the covalent attachment of subunit A to DNA during DNA cleavage. Nucl Acids Res 8:3865–3875

Sumida-Yasumoto C, Yudelevich A, Hurwitz G (1976) DNA synthesis *in vitro* dependent on φX174 replicative form I DNA. Proc Natl Acad Sci USA 73:1887–1891

Sundin O, Varshavsky A (1981) Arrest of segregation leads to accumulation of highly intertwined catenated dimers: Dissection of the final stages of SV40 DNA replication. Cell 25:659–669

Tang D (1978) Purification of a DNA nicking-closing enzyme from mouse L-cells. Nucl Acids Res 5:2861–2875

Taylor DE, Levine JG (1979) Characterization of a plasmid mutation affecting maintenance, transfer and elimination by novobiocin. Evidence that bacterial DNA gyrase is required for replication of H group plasmids. Mol Gen Genet 174:127–133

Tewey KM, Chen GL, Nelson EM, Liu LF (1984) Intercalative antitumor drugs interfere with the breakage-reunion reaction of mammalian DNA topoisomerase II. J Biol Chem 259:9182–9187

Thrash C, Voelkel K, DiNardo S, Sternglanz R (1984) Identification of *Saccharomyces cerevisiae* mutants deficient in DNA topoisomerase I activity. J Biol Chem 259:1375–1377

Trask DK, Muller MT (1983) Biochemical characterization of topoisomerase I purified from avian erythrocytes. Nucl Acids Res 11:2779–2800

Trask DK, DiDonato JA, Muller MT (1984) Rapid detection and isolation of covalent DNA/protein complexes: Application to topoisomerase I and II. EMBO J 3:671–676

Tricoli JV, Kowalski D (1983) Topoisomerase I from chicken erythrocytes: Purification, characterization, and detection by a deoxyribonucleic acid binding assay. Biochemistry 22:2025–2031

Trucksis M, Depew RE (1981) Identification and localization of a gene that specifies production of *Escherichia coli* DNA topoisomerase I. Proc Natl Acad Sci USA 78:2164–2168

Trucksis M, Golub EI, Zabel DJ, Depew RE (1981) *Escherichia coli* and *Salmonella typhimurium supX* genes specify deoxuribonucleic acid topoisomerase I. J Bact 147:679–681

Tse YC, Wang JC (1980) *E. coli* and *M. luteus* DNA topoisomerase I can catalyze catenation or decatenation of double-stranded DNA rings. Cell 22:269–276

Tse YC, Kirkegaard K, Wang JC (1980) Covalent bonds between protein and DNA: Formation of phosphotyrosine linkage between certain DNA topoisomerases and DNA. J Biol Chem 255:5560–5565

Tse YC, Javaherian KJ, Wang JC (1984) HMG17 protein facilitates the DNA catenation reaction catalysed by DNA topoisomerases. Arch Biochem Biophys 231:169–174

Tse-Dinh Y-C, McCarron BGH, Arentzen R, Chowdry V (1983) Mechanistic study of *E. coli* DNA topoisomerase I: Cleavage of oligonucleotides. Nucl Acids Res 11:8691–8701

Uemura T, Yanagida M (1984) Isolation of type I and II DNA topoisomerase mutants from fission yeast: Single and double mutants show different phenotypes in cell growth and chromatin organisation. EMBO J 3:1737–1744

Vazquez-Ramos JM, Mandelstam J (1981) Inhibition of sporulation by DNA gyrase inhibitors. J Gen Microbiol 127:11–17

Vinograd J, Lebowitz J (1966) Physical and topological properties of circular DNA. J Gen Physiol 49 (Suppl):103–125

Vinograd J, Lebowitz J, Radloff R, Watson R, Laipis P (1965) The twisted circular form of polyoma viral DNA. Proc Natl Acad Sci USA 53:1104–1111

Vinograd J, Lebowitz J, Watson R (1968) Early and late helix-coil transitions in closed circular DNA. The number of superhelical turns in polyoma DNA. J Mol Biol 33:173–197

Vogel T, Singer M (1976) The effect of superhelicity on the interaction of histone f1 with closed circular DNA. J Biol Chem 251:2334–2338

Vologodskii AV, Lukashin AV, Anshelevich VV, Frank-Kamenetskii MD (1979) Fluctuations in superhelical DNA. Nucl Acids Res 6:967–992

Vosberg HP, Vinograd J (1975) Isolation and properties of a nicking-closing protein from mammalian nuclei. In: "DNA Synthesis and its Regulation" (Goulian M, Hanawalt P, eds), ICN-UCLA Symp Mol Cellular Biol, Vol 3, Benjamin, Menlo Park, pp 94–120

Vosberg HP, Vinograd J (1976) Purification and demonstration of the enzymatic character of the nicking-closing protein from mouse L cells. Biochem Biophys Res Comm 68:456–464

Vosberg HP, Grossman LI, Vinograd J (1975) Isolation and partial characterisation of the relaxation protein from nuclei of cultured mouse and human cells. Eur J Biochem 55:79–93
Wahle E, Mueller K (1980) Involvement of DNA gyrase in rRNA synthesis in vivo. Mol Gen Genet 179:661–667
Wahle E, Mueller K, Orr E (1984) Gene expression in a temperature-sensitive gyrB mutant of Escherichia coli. EMBO J 3:315–320
Waldeck W, Theobald M, Zentgraf H (1983) Catenation of DNA by eucaryotic topoisomerase II associated with simian virus 40 minichromosomes. EMBO J 2:1255–1261
Wang JC (1969a) Variation of the average rotation angle of the DNA helix and the superhelical turns of covalently closed cyclic DNA. J Mol Biol 43:25–39
Wang JC (1969b) Degree of superhelicity of covalently closed cyclic DNAs from *Escherichia coli*. J Mol Biol 43:263–272
Wang JC (1971) Interaction between DNA and an *Escherichia coli* protein ω. J Mol Biol 55:523–533
Wang JC (1974a) The degree of unwinding of the DNA helix by ethidium. I. Titration of twisted PM2 DNA molecules in alkaline cesium chloride density gradients. J Mol Biol 89:783–801
Wang JC (1974b) Interactions between twisted DNAs and enzymes: The effects of superhelical turns. J Mol Biol 87:797–816
Wang JC (1974c) Protein ω from *Escherichia coli*. In: "Methods in Enzymology" (Grossman L, Moldave K, eds) Vol 29, pp 197–203
Wang JC (1979) Helical repeat of DNA in solution. Proc Natl Acad Sci USA 76:200–203
Wang JC (1981) Type I DNA topoisomerases. In: "The Enzymes" (Boyer P, ed), Vol 14. Academic Press, New York, pp 331–344
Wang JC (1982a) DNA topoisomerases. Sci Am 247:84–95
Wang JC (1982b) DNA topoisomerases. In: "Nucleases" (Linn SM, Roberts RJ, eds). Cold Spring Harbor Lab., New York, pp 41–57
Wang JC, Becherer K (1983) Cloning of the gene *top*A encoding for DNA topoisomerase I and the physical mapping of the *cys*B-*top*A-*trp* region of *Escherichia coli*. Nucl Acids Res 11:1773–1790
Wang JC, Liu LF (1979) DNA topoisomerases: Enzymes which catalyze the concerted breaking and rejoining of DNA backbone bonds. In: "Molecular Genetics" Part III (Taylor JH, ed). Academic Press, New York, pp 65–88
Wang JC, Jacobsen JH, Saucier J-M (1977) Physicochemical studies on interactions between DNA and RNA polymerase. Unwinding of the DNA helix by Escherichia coli RNA polymerase. Nucl Acids Res 4:1225–1241
Wang AHJ, Quigley GJ, Kolpak FJ, Crawford JL, van Boom JH, van der Marcel G, Rich A (1979) Molecular structure of a left-handed double helical DNA fragment at atomic resolution. Nature 282:680–686
Wang JC, Gumport RI, Javaherian K, Kirkegaard K, Klevan L, Kotewicz KL, Tse YC (1980) DNA topoisomerases. In: "Mechanistic Studies of DNA Replication and Genetic Recombination" (Alberts BM, Fox CF, eds), ICN-UCLA Symp Mol Cellular Biol, Vol 19. Academic Press, New York, pp 769–784
Waring MJ (1981) DNA modification and cancer. Ann Rev Biochem 50:159–192
Watson JD, Crick FHC (1953) The structure of DNA. Cold Spring Harbor Symp Quant Biol 18:123–131
Weintraub H (1983) A dominant role for DNA secondary structure in forming hypersensitive structures in chromatin. Cell 32:1191–1203
Weis JH, Faras AJ (1981) DNA topoisomerase activity associated with Rous sarcoma virus. Virology 114:563–566
Weisberg R, Landy A (1983) Site-specific recombination in phage lambda. In: "Lambda II" (Hendrix RW, Roberts JW, Stahl FW, Weisberg RA, eds). Cold Spring Harbor, N.Y., pp 211–250
Weisbrod ST (1982a) Properties of active nucleosomes as revealed by HMG14 and 17 chromatography. Nucl Acids Res 10:2017–2042
Weisbrod ST (1982b) Active chromatin. Nature 297:289–295
Wilson JH (1979) Nick-free formation of reciprocal heteroduplexes: A simple solution to the topological problem. Proc Natl Acad Sci USA 76:3641–3645
Wolfson JS, Hooper DC, Swartz MN, McHugh GL (1982) Antagonism of the B subunit of DNA gyrase eliminates plasmids pBR322 and pMG110 from *Escherichia coli*. J Bacteriol 152:338–344

Worcel A, Burgi E (1972) On the structure of the folded chromosome of *Escherichia coli*. J Mol Biol 71:127–147

Wright A von, Bridges BA (1981) Effect of *gyr*B-mediated changes in chromosome structure on killing of *Escherichia coli* by ultraviolet light: Experiments with strains differing in deoxyribonucleic acid repair capacity. J Bact 146:18–23

Wunder E, Burghardt U, Lang B, Hamilton L (1981) Fanconi's anemia: Anomaly of enzyme passage through the nuclear membrane? Anomalous intracellular districution of topoisomerase activity in placental extracts in a case of Fanconi's anemia. Hum Genet 58:149–155

Yamagishi JI, Furutani Y, Inoue S, Ohue T, Nakamura S, Shimizu M (1981) New nalidixic acid resistance mutations related to deoxyribonucleic acid gyrase activity. J Bact 148:450–458

Yang HL, Heller K, Gellert M, Zubay G (1979) Differential sensitivity of gene expression *in vitro* to inhibitors of DNA gyrase. Proc Natl Acad Sci USA 76:3304–3308

Yegian CD, Mueller M, Selzer G, Russo V, Stahl FW (1971) Properties of the DNA-delay mutants of bacteriophage T4. Virology 46:900–919

Yoshida S, Ungers G, Rosenberg BH (1977) DNA swivel enzyme activity in a nuclear fraction. Nucl Acids Res 4:223–228

Young LS, Champoux JJ (1978) Interaction of the DNA untwisting enzyme with the SV40 nucleoprotein complex. Nucl Acids Res 5:623–635

Zwelling LA, Michaels S, Erickson LC, Ungerleider RS, Nichols M, Kohn KW (1981) Protein-associated deoxyribonucleic acid strand breaks in L1210 cells treated with the deoxyribonucleic acid intercalating agents 4′-(9-acridinylamino)methanesulfon-m-anisidide and adriamycin. Biochemistry 20:6553–6563

Defective-Interfering (DI) RNAs of Influenza Viruses: Origin, Structure, Expression, and Interference

D.P. NAYAK, T.M. CHAMBERS, and R.K. AKKINA

List of Abbreviations

cDNA	complementary DNA	M	membrane protein
CEF	chicken embryo fibroblast	MDBK	Madin-Darby bovine kidney
cRNA	complementary RNA	MDCK	Madin-Darby canine kidney
DI	defective-interfering	MOI	multiplicity of infection
DIU	defective-interfering units	mRNA	messenger RNA
DNI	defective-noninterfering	N	nucleoprotein (VSV)
EID_{50}	egg infectivity titer (50%)	NA	neuraminidase
HA	hemagglutinin	NDI	nondefective-interfering
HAU	hemagglutinating units	NP	nucleoprotein (influenza)

Jonsson Comprehensive Cancer Center and Department of Microbiology and Immunology, UCLA School of Medicine, Los Angeles, CA 90024, USA

Current Topics in Microbiology and Immunology, Vol. 114

NS	nonstructural protein
PAGE	polyacrylamide gel electrophoresis
PFU	plaque-forming units
poly(A)	3′ polyadenosine
RNP	ribonucleoprotein
tRNA	transfer RNA
vRNA	viral RNA
VSV	vesicular stomatitis virus
WSN	Wilson-Smith neurotropic virus

1 Introduction

Influenza virus was the first virus for which defective-interfering particles were reported. VON MAGNUS (1947) observed that when he passaged influenza viruses serially undiluted in embryonated chicken eggs, both the amount of total virus particles as assayed by hemagglutinating units (HAU) and the amount of infectious virus particles as assayed by egg infectivity titer (EID_{50}) decreased. However, he observed that the proportion of infectious particles as assayed by EID_{50} decreased precipitously when compared to the total virus particle (HAU) production. During the undiluted passages many particles were produced which were noninfectious. These noninfectious influenza virus particles produced at high multiplicity were called "incomplete" particles. VON MAGNUS (1951a–c) also described the phenomenon of interference when he reported that these "incomplete" particles interfered with the multiplication of complete (infectious, nondefective, or standard) virus particles. This phenomenon of multiplicity-dependent production of noninfectious virus particles is also called the "von Magnus phenomenon." Since this observation was made with influenza viruses, similar multiplicity-dependent formation of noninfectious particles has been observed with almost all animal viruses carefully studied to date; indeed, formation of such particles has also been reported for plant, yeast, and bacterial viruses (KANE et al. 1979; MILLS et al. 1967) and, thus, probably represents a general phenomenon for all viruses. The term "defective-interfering" (DI) particles was introduced by HUANG and BALTIMORE (1970) to describe the definitive characteristics of these noninfectious particles and to separate them from other noninfectious particles which may be noninterfering.

All DI particles possess the following properties. (a) They are noninfectious and, therefore, need the helper function of standard virus particles for replication. (b) They are not only defective but also interfering, i.e., when the same cell is infected with both standard and DI particles, DI particles replicate at the expense of the standard particles. (c) DI particles are defective because they are deletion mutants, i.e., they do not possess the complete viral genome but a shortened genome (compared to that of the standard virus). However, although almost all DI particles appear to contain a shortened genome, it is possible to envision a DI particle containing a genome similar in size to the standard genome, or even one larger than the standard genome. Alterations, such as base substitution or insertion rather than deletion of the standard genome, could make the viral genome defective as well as interfering.

This review is a summary of current knowledge of influenza DI particles with special emphasis on the origin, structure, and expression of the DI genome and the possible role of influenza DI particles in virus evolution and disease.

Some aspects of these and other studies have been presented in two recent reviews (NAYAK 1980; NAYAK and SIVASUBRAMANIAN 1983) and will not be discussed in depth here. In addition, readers may consult a recent comprehensive publication on the biology of influenza viruses which encompasses among other things, structure of the viral genome, transcription, replication, genetics, and epidemiology (see PALESE and KINGSBURY 1983).

2 Definition

Originally, VON MAGNUS (1951 b) used the term "incomplete" virus to describe the noninfectious particles produced during undiluted passages of A/PR/8/34 virus in embryonated chicken eggs. HUANG and BALTIMORE (1970) introduced the term "defective-interfering" (DI) virus particles to describe the characteristics of these noninfectious particles produced by high-multiplicity passages. The term "DI" particles has been used extensively in describing these noninfectious particles of other viruses. However, with influenza viruses, use of the term "DI" is not as common as with other viruses for a number of reasons. (1) Until 1978 some confusion existed about the genetic nature of the "incomplete" influenza particles. Since these particles were noninfectious, since there was no clear evidence for the intragenic deletion and, since the interference by these influenza particles was not easily assayed, they were called "incomplete" particles rather than DI particles. (2) In the influenza virus population, at least three types of particles exist: (a) complete infectious or standard particles, (b) DI particles (see NAYAK 1980), and (c) defective-noninterfering (DNI) particles (HIRST and PONS 1973; CARTER and MAHY 1982a). These different particles cannot be physically separated from each other and remain as a mixture, although the proportion of each particle may vary in different virus preparations.

In 1978 it was demonstrated that there is a class of influenza particles satisfying all five functional and genetic criteria for the definition of DI particles (NAYAK et al. 1978). Like other DI particles, they (a) are produced and amplified by high-multiplicity passages, (b) are defective and need the helper function of standard virus particles for replication, (c) can cause interference with the homologous standard virus particles, and (d) possess deleted RNA (DI RNA) segments which (e) are responsible for interference. For the sake of generality, therefore, we prefer the term "DI" to describe this class of influenza defective particles. The phenomenon of interference and the presence of deleted RNA are unique to DI particles and not to other classes of "defective" noninterfering (DNI) particles.

Two terms, "DI RNAs" and "subgenomic RNAs," have been used to describe deleted RNAs. As discussed in Sect. 9, the term "DI RNA" should be restricted to describing one or more predominant RNA species in a DI preparation which is known to possess interfering ability. The other deleted RNA segments for which interference has not been demonstrated should be called subgenomic RNAs, i.e., all subgenomic RNAs are not DI RNAs, while all DI RNAs are subgenomic RNAs. Probably only a few of the subgenomic

RNAs become predominant species because of some as yet undefined intrinsic properties which give these RNA species an advantage in replication at the expense of other genomic or subgenomic RNA segments (see Sect. 9).

Some influenza virus preparations may contain a substantial amount of DNI particles (CARTER and MAHY 1982a). Although the genetic nature of these particles has not been clearly defined, they are expected to lack one or more standard segments and may possess no DI RNA segment. The factors that contribute to the generation of DNI particles are not fully understood, but they may be present in both DI and standard virus preparations. It is, however, possible to obtain virus preparations in which over 99% of standard particles are replaced by DI particles (JANDA et al. 1979).

Another type of particle, called "nondefective-interfering" (NDI), may be produced. NDI particles possess a complete genomic complement and one or more DI RNA segments. Such NDI particles do not produce plaques, but express all viral antigens, and perhaps even produce virus particles. Whether NDI particles are actually produced or not depends on two factors. (a) Such particles should be able to package more than eight RNA molecules including more than one molecule of the same gene, e.g., standard PB1 and deleted PB1 gene, in the same particle. The assembly and packaging of standard RNA segments is yet to be understood. (b) Enough standard virus RNA segments and DI RNA segments must be produced so that a fraction of particles possesses both a set of complete standard RNA segments and one or more DI RNAs. However, since the molar ratios of standard RNA segments – particularly those of the largest RNA segments, e.g., polymerase genes – are reduced in DI preparations, it is expected that only a few, if any, particles possess a complete set of standard RNA segments as well as one or more DI RNAs.

In the older literature the term "incomplete" particles was extensively used. Since these "incomplete" particles were produced by the same procedure that was used by VON MAGNUS and since VON MAGNUS (1951a–c) demonstrated in his early studies that these "incomplete" influenza particles were interfering, we assume that all "incomplete" particles generated by the von Magnus procedure are procedure are predominantly DI particles.

In summary, a variety of terms, such as "incomplete," "defective," "noninfectious," "von Magnus," "defective-interfering," have been used to describe DI particles. Similarly, "standard," "infectious," "complete," "nondefective," have been used to describe infectious virus particles. In this review we shall use either "incomplete" or "defective-interfering" particles and "standard" particles to describe these two classes of virus particles.

3 Generation and Amplification of Influenza DI Particles

In the initial studies, DI particles were produced by repeated passages of the A/PR/8/34 strain of influenza virus in the chorioallantoic sac of chicken embryos (VON MAGNUS 1951a–c). FAZEKAS DE ST. GROTH and GRAHAM (1954a, b) and, later, VON MAGNUS (1965) showed that "incomplete" particles can be produced

with B/Lee virus by undiluted passage but required more serial passages (usually six or more) in embryonated chicken eggs. WERNER (1956) demonstrated that "incomplete" influenza particles are produced in the amniotic sac of chicken embryos and GINSBERG (1954) reported the generation of "incomplete" particles by infecting mice with large doses of viruses.

"Incomplete" influenza particles were successfully produced using undiluted inocula but not using highly diluted inocula. Although CAIRNS and EDNEY (1952) claimed that multiple infection was not necessary and that "incomplete" particles were produced even when only 1% of cells were infected, it became well established that multiple infection was crucial for the production of "incomplete" particles. Recent observations confirm that the amplification of DI particles depends on the coinfection of the same cell by DI particles as well as by standard virus (NAYAK et al. 1978). Generation of DI particles can take place even in a single cycle of infection by standard virus alone but, unless subsequently amplified by multiple infection with standard viruses, these DI particles will not replicate and therefore will not be detected (JANDA et al. 1979).

Two factors in particular appear to be crucial in the generation and amplification of influenza DI particles. The first is a host system capable of allowing growth of the standard virus to a high titer. Therefore, host cells such as HeLa cells or L cells (which cause abortive replication of influenza viruses) are not suitable for the production of influenza DI particles. With influenza viruses three host systems have often been used to study DI particles. (a) As studied initially by VON MAGNUS (1951a) and repeated later by others, DI particles have been produced by undiluted passages in embryonated chicken eggs (HENLE 1953; HOYLE 1968). However, after the third or fourth serial passage of undiluted viruses, the total virus production is greatly reduced. (b) For fowl plague virus, which grows to a high titer in chicken embryos as well as in primary chicken embryo fibroblast (CEF) cells, DI particles can be generated by passaging the virus at high multiplicity in either embryonated chicken eggs or in primary CEF cultures (ROTT and SCHÄFER 1960; CARTER and MAHY 1982a). (c) With Wilson-Smith neurotropic (WSN) virus, either Madin-Darby bovine kidney (MDBK) cells or Madin-Darby canine kidney (MDCK) cells can be used in the generation and amplification of DI particles (see NAYAK 1980). MDCK cells also appear to be the host of choice for producing DI particles of influenza B virus (DE and NAYAK 1980).

The second factor, as discussed above, is the multiplicity of infection – which should be high. Generation and amplification of DI particles are two independent phenomena. Generation of DI RNAs depends on the aberrant replication of viral RNA and, therefore, does not depend on the multiplicity of infection. These DI RNAs are then incorporated into DI particles. Subsequent amplification of DI particles during the next round of infection depends on the availability of the helper function of the standard viruses. The high multiplicity of infection is required for amplification to insure that the same cell will be infected by one or more DI particles along with one or more standard particles. The ratio of DI particles/standard particles in the virus yield from co-infected cells depends on the interfering ability of the specific DI RNA with the replication of standard viral RNAs. The factors that affect such an outcome

are not properly understood. The number of DI particles obtained from clones after repeated passages appears to vary from 2×10^6 to 5×10^6 defective interfering units (DIU) per millilitre (JANDA et al. 1979; CARTER and MAHY 1982a). However, the absolute yield of DI particles is not affected by the contaminating standard particles. The amplification and production of DI particles may depend on the replicating ability of a DI RNA and may be different from its interfering ability which, instead, will determine the number of contaminating standard particles. In fact, usually a higher yield of DIU/ml is observed when the preparation also contains a relatively large number of standard particles.

When passed at a high multiplicity, 99% of the standard particles are replaced by DI particles by passage 3 to passage 5. Further amplification of DI particles is done by coinfecting cells with 2–3 DIU/cell and 3–5 plaque-forming units (PFU) per cell. Such a procedure yields a large amount of DI particles with uniform characteristics. By screening DI particles produced from a number of individual plaques, one may select a DI preparation with the desired characteristics and amplify such a DI particle by coinfection with standard particles for further studies. A number of other factors, including the presence of heat-inactivated virus (PAUCKER and HENLE 1955) and premature release of virus particles (VON MAGNUS 1952), have been postulated as being involved in the production of "incomplete" particles. However, critical experiments with appropriate controls implying definitive involvement of these factors in the generation and amplification of DI particles are lacking.

4 Physical and Biological Characteristics

Soon after the discovery of influenza DI particles by VON MAGNUS (1947), attempts were made to isolate and separate them from standard particles and to compare their properties with those of standard particles. In addition to the two important biological properties of noninfectivity and interfering ability, influenza DI particles have been compared to standard particles with respect to, among other things, morphology, size, density, lipids, structural proteins, antigenicity, and nucleic acid content. Reports on the separation of influenza DI particles from standard particles are conflicting. Von Magnus used velocity centrifugation to separate influenza DI particles and identified a 380S component for incomplete particles as compared to a 660S–750S for standard particles (GARD and VON MAGNUS 1947; GARD et al. 1952). Estimated S-values of incomplete particles have ranged from 300S to 600S (HANIG and BERNKOPF 1950; VON MAGNUS 1954; ROTT and SCHÄFER 1960). YOSHISHITA et al. (1959) claimed clear separation of influenza DI particles from standard viruses. However, the experiments are often not reproducible and clear-cut separation of DI and standard influenza particles has not been obtained in most laboratories. We and others (see NAYAK 1980) have found more than one peak of influenza virus particles in sucrose gradients. However, these peaks are not consistently produced and, moreover, they often do not show the reduced PFU/HA ratio which is indicative of DI particles.

A number of factors contribute to the difficulty of separating influenza DI particles from standard particles. Because standard influenza particles vary in size and shape, DI particles although reportedly more heterogeneous than standard viruses, cannot be physically separated from standard particles. Since nucleic acid constitutes only a small fraction (about 0.8%) of the total mass of the standard particle (ADA and PERRY 1955), a slight reduction of RNA content in DI particles (compared with that in standard particles) does not change the density enough for them to be separated from heterogeneous standard particles. Variation in RNA content among different DI particles makes separation even more difficult. Furthermore, any comparative analyses of the properties of the DI particles versus standard particles are complicated by the presence of DNI particles (CARTER and MAHY 1982a).

Although "incomplete" particles could not be separated from the standard particles, gradient centrifugation suggested that "incomplete" particles are heterogeneous and less dense. UHLER and GARD (1954) reported that "incomplete" particles contain more lipid than is present in standard particles and suggested that increased amounts of lipid may contribute to the change in sedimentation property of "incomplete" particles observed in sucrose density gradient analyses. Increased heterogeneity has also been observed in the morphology of "incomplete" particles (SHIOTA 1956). "Incomplete" particles tend to be more pleomorphic, and often lack electron-dense internal bodies (BIRCH-ANDERSEN and PAUCKER 1959; MOORE et al. 1962; MORGAN et al. 1962; PAUCKER et al. 1959; PETERS 1959; WERNER and SCHLESINGER 1954; WATERSON et al. 1961). "Incomplete" particles have been reported to contain varying amounts of the envelope antigens, HA and neuraminidase (NA), and membrane protein (M) (SETO 1964; LENARD and COMPANS 1975; NAYAK et al. 1978). In addition, reduced polymerase activity and nucleoprotein (NP) content produced have been reported in DI particles (ADA and PERRY 1955; SCHOLTISSEK et al. 1966; LENARD and COMPANS 1975; NAYAK et al. 1978). In summary, influenza DI or "incomplete" particles essentially possess the same antigenic determinants and protein components as standard viruses, although the amounts of individual components vary quantitatively in different DI preparations. The observed variation in the protein composition of DI particles possibly may represent heterogeneity among influenza DI particles as well as varying contamination with standard particles.

VON MAGNUS (1954) observed two biological properties that made these influenza particles different from standard particles – defectiveness (i.e., the lack of infectivity) and interfering ability. In early experiments, particles with reduced infectivity as measured by PFU/HA or EID_{50}/HA were equated with DI particles, i.e., the measure of defectiveness was used as a measure of interfering ability. Indeed, this would be true for all nonsegmented RNA viruses in which a noninfectious particle is an interfering particle, although some defective particles may be more interfering than others (RAO and HUANG 1982). However, with influenza virus all defective particles may not be interfering particles (NAYAK 1980). Since the genome of influenza virus is segmented, one can obtain virus particles in which one or more standard RNA segments are missing. Such particles will be defective but noninterfering. Indeed, these DNI particles are

produced in significant numbers during both low-multiplicity (HIRST and PONS 1973) and high-multiplicity passages (CARTER and MAHY 1982a). Furthermore, since interfering particles could be separated neither from DNI nor from standard influenza particles, physical or chemical quantitation could not be used to determine the number of interfering particles in a given preparation. In addition, PFU/HA or EID_{50}/HA ratios are not sensitive enough to detect a relatively small number of DI particles present in a standard virus preparation. Therefore, a direct assay for quantitating influenza DI particles has been developed, using infectious center reduction (JANDA et al. 1979). In this procedure, a cell coinfected with both standard and one or more DI particles does not produce a visible plaque. Since only the fraction of cells infected with standard particles but not coinfected with DI particles produces plaques, the multiplicity of DI particles can be precisely calculated using the Poisson distribution formula, $P(0)=e^{-m}$ or $m=-ln\ P(0)$, where m is the multiplicity of DI particles and $P(0)$ is the fraction of cells not receiving DI particles as determined from the number of plaques in an infectious center assay. Using this assay, one can calculate the concentration of biologically active DI particles in a given preparation as DIU/ml (JANDA et al. 1979; CARTER and MAHY 1982b). It was further observed that a single DI particle can prevent visible plaque formation by standard particles. In addition, one could assay the contaminating standard particles in the same preparation by analyzing the PFU/ml at a very low multiplicity. Since the interfering activity is highly resistant to ultraviolet light (NAYAK et al. 1978), one can essentially eliminate the contaminating standard virus by ultraviolet irradiation and determine the DIU/ml in the same virus preparation. Thus, the same virus preparation can be assayed for PFU/ml, DIU/ml, and HAU/ml.

5 Genome of Influenza DI Particles

Since the genome of influenza virus is RNA and since "incomplete" particles lack infectivity, it was appropriate to investigate the RNA content of "incomplete" particles to account for the loss or reduction of infectivity. Early analyses indicated that noninfectious or "incomplete" particles contain little or no nucleic acid (see HOYLE 1968). Such results have not been reproduced (NAYAK and BALUDA 1967) and may reflect the lack of purity of virus material and a lack of sensitivity of early assays. ADA and PERRY (1955, 1956) were the first to critically analyze the RNA content of both standard and "incomplete" virus particles. They reported that standard viruses contain approximately 0.8% nucleic acid and that "incomplete" particles contain less nucleic acid than is present in standard viruses. The reduction in infectivity plotted on a logarithmic scale correlated with the reduction of nucleic acid on a linear scale. Even these early studies indicated that reduction in infectivity was not directly proportional to the reduction of nucleic acid content. The amount of S, or ribonucleoprotein (RNP), antigen in "incomplete" particles compared with that in standard virus was reduced (LIEF et al. 1956; ROTT and SCHÄFER 1961; ROTT and SCHOLTISSEK 1963). Essentially similar conclusions were drawn about the nucleoprotein (NP)

content of DI particles by LENARD and COMPANS (1975) from polyacrylamide gel electrophoresis (PAGE) analyses of DI and standard viral proteins. It was also found that one could not obtain either the multiplicity-dependent reactivation or marker rescue using "incomplete" particles at high multiplicity (BARRY 1961; ROTT and SCHOLTISSEK 1963). Later, both sucrose gradient analyses and PAGE analyses implied that the influenza viral RNA genome was segmented. Comparative RNA analyses showed that an overall reduction of high-molecular-weight RNAs and an increase of low-molecular-weight RNAs occurs in "incomplete" particles when compared to that in standard particles (DUESBERG 1968; CHOPPIN and PONS 1970; NAYAK 1969, 1972). It was further implied (CHOPPIN and PONS 1970) that one specific segment of RNA was missing in incomplete particles, which would account for the loss of infectivity, lack of multiplicity-dependent reactivation, or lack of marker rescue among DI particles (BARRY 1961). However, using a more careful analysis of RNA by PAGE, BEAN and SIMPSON (1976) could not find a direct correlation between the loss of a single RNA segment and the reduction of infectivity. They concluded that factors other than the loss of a specific RNA segment were involved in the drastic reduction of infectivity observed in DI particles. However, none of these observations could account for the interfering ability of influenza DI particles.

Once the eight segments of the influenza viral genome and the proteins coded by each segment of the genome were identified (see PALESE 1977), the genomic content of the influenza DI particles was reinvestigated. Since influenza DI particles could not be separated from standard particles, it was important to obtain DI preparations in which contaminating standard particles would constitute a very small fraction and to obtain enough such DI particles for repeated analyses. These conditions were satisfied when it was found that influenza DI particles could be produced in large amounts in MDBK cells, known for favoring the growth of standard particles. Furthermore, it was found that as with DI particles of other viruses, influenza DI particles could be amplified by coinfecting cells with DI particles and standard helper viruses (NAYAK et al. 1978). When the RNA genomes of such DI preparations were examined by gel analyses, it was found that the DI genome contained – in addition to the eight standard RNA segments – one or more novel RNA segments (Fig. 1) not present in plaque-purified standard virus preparations (NAYAK et al. 1978). These small RNA segments became increasingly pronounced at each serially undiluted passage. Similar small RNA segments have been found in other influenza DI preparations as well as in many standard virus preparations (CRUMPTON et al. 1978; JANDA et al. 1979; DE and NAYAK 1980; NAKAJIMA et al. 1979). Over the past 5 years a great deal of effort has been made to understand the nature of the DI RNA segments, their relationship to the progenitor RNA, and their role in interference.

Hybridization studies soon established that these RNAs are virus specific, of the same polarity (i.e., negative strand) as the viral RNA, and predominantly of polymerase gene origin (DAVIS and NAYAK 1979; NAKAJIMA et al. 1979). Each DI preparation obtained by repeated passage of individual plaques contains a unique set of DI RNAs. A single DI preparation may contain multiple DI RNA segments, though not in equimolar ratios. Once produced, these DI

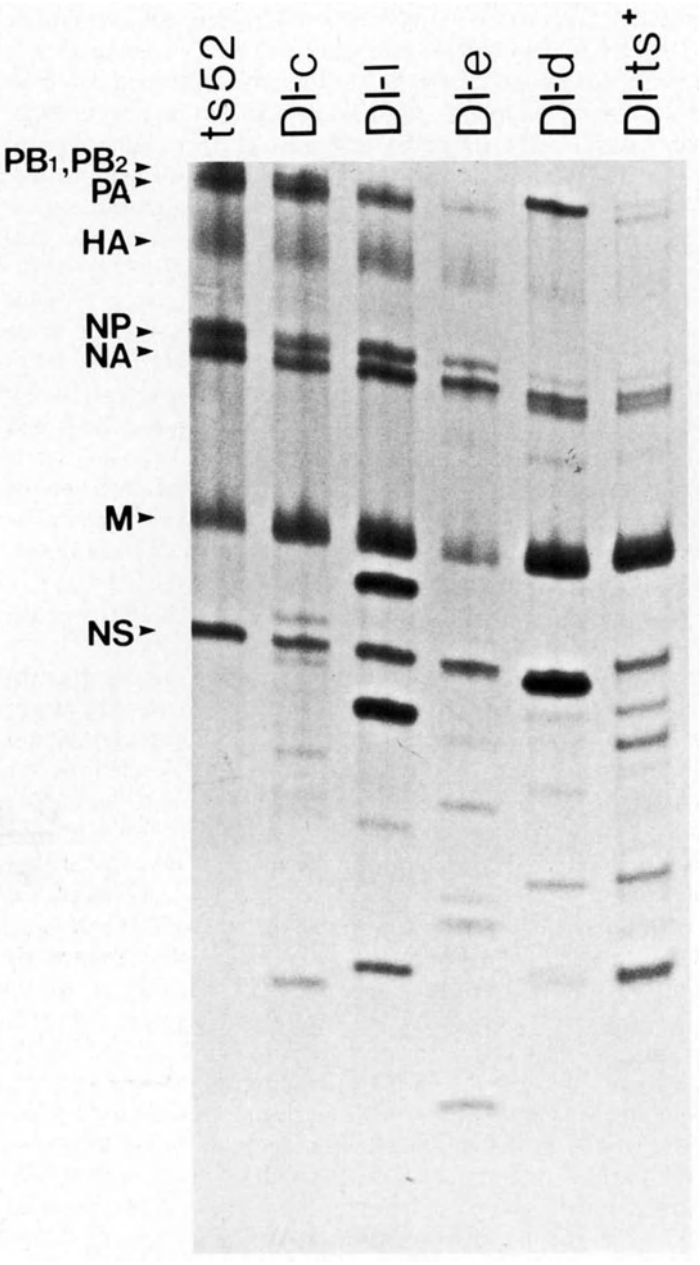


Fig. 1. Analysis of ^{32}P-labeled RNAs of influenza DI preparations. MDBK cells were infected with WSN ts-52 standard virus or coinfected with standard virus and different DI preparations – DI-c, DI-l, DI-e, DI-d, or DI-ts$^+$ – obtained by serial passage of individual plaques without dilution. Standard viral RNAs are indicated at *the left*. Electrophoresis was in a 2.2% polyacrylamide/0.6% agarose/6*M* urea gel. (From JANDA et al. 1979)

RNAs replicate, amplify, and become the predominant species in a given DI preparation. Upon long-term passage, however, new DI RNA segments are generated and become the predominant DI species replacing earlier DI RNAs (DE and NAYAK 1980). In some cases, DI RNA species found in later passages are of the same progenitor gene as the earlier DI RNA but smaller than the DI RNA present in earlier passages.

When it was established that these small RNA segments are DI-specific and can replicate in coinfected cells, it became important to determine the role of these DI RNAs in interference with standard virus replication. With DI particles of other viruses, such as vesicular stomatitis virus (VSV), definitive studies have been possible because pure DI particles and standard virus particles can be separated from each other and because DI particles contain only the shortened nucleic acid. With influenza viruses, it is not possible to obtain direct experimental evidence of interference because influenza DI particles, as discussed earlier, cannot be completely separated from standard particles and influenza DI particles contain standard RNA segments in addition to DI RNA segments. Initially, several lines of evidence suggested that these DI RNA segments, rather than standard RNA segments, were responsible for the interference with standard virus replication. It was shown that DI RNA segments, although present in many standard virus preparations, can be excluded by repeated cloning and, therefore, do not play a role in standard viral replication. DI RNAs, on the other hand, become amplified with high-multiplicity passages and in a given DI preparation the amount of DI RNA fairly correlates with its interfering ability (DIU/ml) (NAYAK et al. 1978; JANDA et al. 1979). Ultraviolet light sensitivity data suggested that the target size of the interfering molecule is small and approximates the size of the predominant DI RNA species (NAYAK et al. 1978). Finally, it was shown that RNP complexes containing DI RNA species possess interfering ability, while other RNP complexes from either the standard or DI particles – which do not contain DI RNA – do not interfere (JANDA and NAYAK 1979). Furthermore, this interfering ability was ribonuclease-sensitive. These data provide direct evidence that the DI RNA in the RNP complex is responsible for DI-mediated interference.

Analyses of RNA from different DI preparations show the following characteristics:

1. Each contains one or more DI RNA segments (see NAYAK 1980; NAYAK and SIVASUBRAMANIAN 1983). These DI RNA segments are not in equimolar concentration, suggesting heterogeneity among DI particles in the same preparation with respect to DI RNA content. Since we do not know the genomic content of individual DI particles, it is not possible to assess the number of DI RNA segments per DI particle. We assume that a DI particle must contain at least one DI RNA segment to provide the interfering ability.
2. The standard genomic segments of the DI RNA population also vary among different DI RNA preparations. Large molecular weight RNA segments, particularly polymerase genes, are reduced in most DI preparations (BEAN and SIMPSON 1976; NAYAK et al. 1978; CRUMPTON et al. 1978, 1979, 1981). In some DI preparations individual specific gene segments appear to be

more drastically reduced than other polymerase genes (CHOPPIN and PONS 1970; UEDA et al. 1980; JANDA et al. 1979). Molar ratio estimates of different standard RNA segments indicate that particles in DI preparations are likely to lack one or more standard RNA segments, i.e., they may not possess the full complement of standard RNA segments present in standard infectious virus. In addition, these DI particles contain one or more DI RNA segments.

The lack of a full complement of standard RNA segments, which may vary among DI preparations, would make these particles noninfectious. In the majority of DI preparations, the loss of a random, rather than a specific, RNA segment would account for the drastic loss of infectivity often observed among DI particles (NAYAK et al. 1978). At a high multiplicity of DI particles, although cells may get a full complement of standard genomic segments, the presence of one or more DI RNAs would interfere with multiplicity-dependent reactivation as well as marker rescue, which would otherwise be expected if DI particles were formed due to a random loss of one or more RNA segments. However, one would expect a partial expression of viral genes at a high multiplicity of DI particles. KAVERIN et al. (1980) reported such results at a higher multiplicity of DI particles in support of this hypothesis. However, the presence of a significant number of DNI particles would complicate the interpretation of these data (CARTER and MAHY 1982a).

In summary, the genomic content of influenza DI particles varies among different preparations; they appear to contain less than a full complement of standard RNA segments, but contain one or more DI RNA segments. However, it should be noted that infectious center reduction data show that even a particle which contains a full complement of standard RNA segments plus one or more DI RNA segments would not produce a plaque and, therefore, would not behave as a standard infectious virus.

6 Structure of DI RNAs

When it became obvious that the small RNA segments found in influenza DI preparations were not the result of degradation of standard segments but represented replicative molecules responsible for interference, it became important to determine the origin and structure of these DI RNAs and their relationship to the standard progenitor RNA segment. Initial studies involving hybridization, oligonucleotide mapping, and partial direct RNA sequencing of the 5′ and 3′ termini showed that all DI RNAs possess both genomic termini and thus arise by internal deletion (NAYAK 1980). Furthermore, it was evident that the same progenitor gene can give rise to multiple DI RNAs of varying length and that the smaller DI RNAs are not always the subset of the larger ones (DAVIS and NAYAK 1979; DAVIS et al. 1980). However, precise data showing the exact site of deletion could not be obtained using the above techniques. With the advent of recent developments in cDNA cloning and direct DNA sequencing, it has become possible to determine the nucleotide sequence of

cDNA clones of influenza virus standard gene segments as well as of DI RNAs arising from these genes.

A number of different DI RNAs of a variety of negative and positive strand RNA viruses have recently been studied in considerable detail using direct RNA and DNA sequencing. Among the best studied are the DI RNAs of VSV (YANG and LAZZARINI 1983), Sindbis virus (MONROE and SCHLESINGER 1983; MONROE et al. 1982), Semliki forest virus (PETTERSON 1981; LEHTOVAARA et al. 1981), and influenza virus (NAYAK and SIVASUBRAMANIAN 1983; JENNINGS et al. 1983). Based on the sequence data derived from these viruses, all DI RNAs appear to contain only a portion of the standard genome and these DI RNAs can be classified into two major groups (LAZZARINI et al. 1981):

1. 5′ DI RNAs: DI RNAs belonging to this class have preserved the 5′ end of the parental strand, followed by a significant loss of the 3′ portion including the entire 3′ progenitor terminus. The majority of DI RNAs of the nonsegmented negative strand viruses (VSV and Sendai viruses) belong to this class. The 3′ terminus of such a DI RNA will be generated by copying its own 5′ terminus. Thus, a stretch of nucleotides at the 3′ end will be complementary to the 5′ end of the genome. Because of their terminal complementarity, these DI RNAs tend to circularize after deproteinization and form stem-like structures called panhandles. It is proposed that these DI RNAs arise by a copy-back mechanism (KOLAKOFSKY 1976; HUANG 1977). When copying the positive strand of the genome, the viral polymerase becomes detached from the template at the deletion point and copies a portion of the 5′ end of the daughter strand, thereby giving rise to complementary termini and a panhandle structure (PERRAULT 1981; LAZZARINI et al. 1981). Since this class of DI RNAs lack the 3′ terminus, including the leader sequence which is recognized by the polymerase for transcription, messenger RNAs (mRNAs) are not produced. These are called nontranscribing DI RNAs. Another variation of the 5′ DI RNAs are the so-called snap-back or hairpin DI RNAs. The 5′ half of a DI RNA of this type is identical to the progenitor genome but the 3′ half is the exact complement of its 5′ half. Such a DI RNA, therefore, can self-hybridize almost 100%, giving rise to hairpin-like structures when the DI RNA is deproteinized. VSV DI-011 is an example of this type (see LAZZARINI et al. 1981; PERRAULT 1981).

2. 5′-3′ DI RNAs: All DI RNAs belonging to this class have one common characteristic, i.e., they retain both the 3′ and 5′ termini of the progenitor gene (or genome). Influenza DI RNAs, some of the DI RNAs of VSV (YANG and LAZZARINI 1983), and Sendai (AMESS et al. 1982) viruses, as well as the DI RNAs of positive strand viruses (poliovirus, Semliki forest virus) belong to this class. All influenza DI RNAs studied to date are 5′-3′ DI RNAs (JENNINGS et al. 1983; NAYAK and SIVASUBRAMANIAN 1983). This class can be further grouped into subclasses:

   a) 5′-3′ single deletion: DI RNAs that arise by single internal deletion of the progenitor genome; the size of the internal deletion may vary and determines the size of the DI RNA.

b) 5′-3′ multiple deletion: DI RNAs that retain both genomic termini but contain more than one internal deletion.
c) 5′-3′ complex: DI RNAs that contain one or more deletions in addition to more extensive changes such as, insertion, transposition, base changes, or even new sequences not present in the progenitor gene.
d) 5′-3′ mosaic DI RNA: DI RNAs that involve deletion as well as a true intersegmental recombination between two or more RNA segments (or genomes). Recent evidence indicates that intersegmental recombination between two different RNA molecules is possible and has been reported with picornaviruses (KING et al. 1982). Based on the DNA sequence data, one influenza DI RNA of mosaic type has been reported (MOSS and BROWNLEE 1981). However, this has not yet been confirmed by direct RNA sequencing and, therefore, it is possible that this mosaic structure is a cloning artifact (BUONAGURIO et al. 1984).

In addition to these types of DI RNAs, another type of unusual DI RNA of Sindbis virus containing a host transfer RNA (tRNA) sequence at the 5′ end has been reported (MONROE and SCHLESINGER 1983), suggesting that neither the 3′ nor the 5′ genomic terminus is absolutely essential for replication.

7 Complete Nucleotide Sequence of DI RNAs

Complete nucleotide sequences of a few DI RNAs of PB1 and PB2 origin are shown in Figs. 2 and 3, respectively. In addition, 35 other subgenomic RNAs of A/PR/8/34 virus have been recently sequenced using M13 cDNA cloning (JENNINGS et al. 1983). Taken together, complete sequences of 41 subgenomic RNA segments have been determined and this represents the largest number of DI RNAs with known sequence in any virus group to date. These sequence analyses have revealed the following characteristics:

1. Distribution of origin: Only six of 41 subgenomic RNA segments arose from genes other than polymerase genes. Among these six RNAs, three arose from HA and one each from NA, NP, and nonstructural protein (NS) genes. Of the 35 which are of polymerase gene origin, an overwhelmingly large number (26) are of PB2 origin, while three and five came from PB1 and PA, respectively. Previous studies of 16 influenza DI RNA segments (Nayak 1980; Nayak unpublished data) indicated that all of the predominant DI RNA segments present in six DI preparations are of polymerase gene origin. Recent sequence studies also demonstrate the paucity of subgenomic RNAs arising from nonpolymerase RNA segments (JENNINGS et al. 1983). However, there are as yet no clues that would explain this asymmetric distribution in the origin of DI RNA segments.
2. Size: Although large DI RNAs have been noted (NAYAK 1980), these have not yet been sequenced. The ones that have been completely sequenced fall within a range of 178–859 nucleotides, with the majority (84%) being 300–500

Fig. 2. Complete nucleotide sequence of L2b and L3 DI RNAs of PB1 origin. *Broken lines* represent regions that are absent in L2b and L3 DI RNAs; *N-N-N-N*, a sequence of 1658 nucleotides in the PB1 gene. (From NAYAK et al. 1982b)

nucleotides. This distribution in size may have been biased because M13 cloning favors insertion of cDNA in this size range.

3. Structure: Nearly 90% (36/41) of subgenomic RNAs arose by single internal deletion; three possessed two internal deletions and only one each was of the complex or mosaic type. The extent of deletion in the progenitor gene


5′
P3 cRNA
L2a-7
L2a-17
C
A & B
3′

ranges from 1977 (in PB1) to 712 (in NS) nucleotides. The second deletion, when found, is usually smaller (approximately 50 nucleotides). Furthermore, DI RNAs arising from each gene also vary in size. All DI RNAs appear to possess the 5′ and 3′ terminal sequences of the progenitor gene. JENNINGS et al. (1983) reported a relatively symmetrical contribution from the 5′ and 3′ ends of the progenitor gene. The range of contribution from the 5′ end (positive sense) varies from 83 to 445 nucleotides and that from the 3′ end from 95 to 413 nucleotides.

4. Except for the internal deletion, sequence alteration among the influenza DI RNAs, unlike the DI RNAs of other viruses, is relatively uncommon. A few base mismatches have been observed and only rarely have any drastic changes been found, such as insertion or transpositions in sequence. Most likely the base mismatches, as well as other alterations of nucleotide sequence in the DI genome, occur during the multiple passages of the DI particles rather than during the generation of DI RNAs from the progenitor gene. Occurrence of only a few base substitutions observed among the 41 subgenomic influenza RNA segments (JENNINGS et al. 1983) probably reflects the recent origin of these DI RNA segments from the progenitor gene.
5. After the deletion point, DI RNAs do not show any preference for maintaining one reading frame over others. About one-third of DI RNAs sequenced maintain the same reading frame as the progenitor gene after the deletion point and therefore these DI polypeptides, if produced, are expected to retain the NH_2 and COOH termini of the parent polypeptide; whereas, the other two thirds will possess the same NH_2 terminus but a different COOH terminus. DI-specific polypeptides have been detected with some DI preparations (see Sect. 10). However, we do not know if these DI-specific polypeptides were transcribed from DI RNAs that maintained the same reading frame as the progenitor gene. Furthermore, the role, if any, of these defective proteins remains undetermined (see Sect. 11).
6. The sequences at the junction points do not resemble the consensus splicing sequences of eukaryotic RNAs. We cannot detect any unique sequence for the detachment or reattachment of the polymerase on the template of plus and minus progenitor RNA strands (Fig. 4).

In general, sequence analyses for relatively large numbers of influenza DI RNAs indicate that the majority of DI RNAs arise by internal deletion and from polymerase genes. However, they provided no clue as to why the majority of DI RNAs arise from polymerase genes. It has been suggested that because of their large size polymerase genes are more prone to errors in replication.

Fig. 3. Complete nucleotide sequence of DI RNAs of PB2 origin. *Squares* indicate mismatches with the progenitor gene; *solid lines* show the common nucleotide sequence (GGA) at the deletion point; *broken lines* show regions that are absent in the DI RNAs; *N-N-N-N*, a sequence of 1626 nucleotides in the PB2 gene. PB2, L2a-7, and L2a-17 are of WSN origin (KAPTEIN and NAYAK 1982; SIVASUBRAMANIAN and NAYAK 1983); whereas, C and A, B were obtained from A/PR/8 virus (WINTER et al. 1981). (From NAYAK and SIVASUBRAMANIAN 1983)

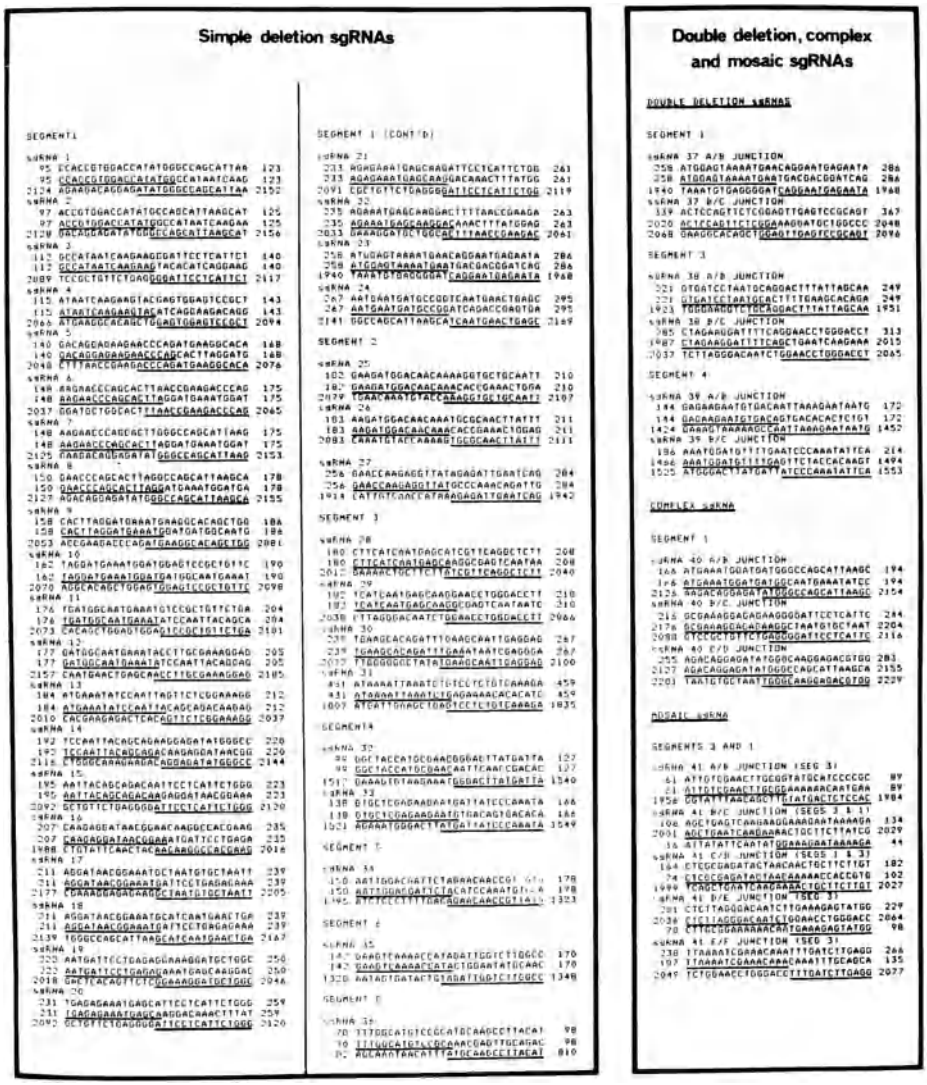


Fig. 4. Nucleotide sequences of subgenomic RNAs at the junction points. The top line of each set is the subgenomic RNA sequence at the junction point, aligned with the corresponding parental sequence from the 5′ (*middle line*) and 3′ (*bottom line*) ends of the progenitor segments (mRNA sense throughout). The *underlined regions* of parental sequence are identical with the subgenomic RNA sequence. (From JENNINGS et al. 1983)

However, this cannot be the only reason because, although polymerase genes are larger, their relative molarity (even in standard virus preparations) often is lower than that of smaller genes. Furthermore, intracellular replication of polymerase RNA segments is much lower than that of M or NS segments (SMITH and HAY 1982). Since the absolute number of nucleotides synthesized in an infected cell is the size (nucleotide number) of the RNA times the number

of molecules synthesized, it appears that the total number of nucleotides of a polymerase gene synthesized may not be very different from that of the M or NS gene in infected cells. Therefore, if the generation and the selection of all subgenomic RNAs are based on random errors of the polymerase complex, one would expect a more even distribution of DI RNAs in all standard RNA segments than is in fact found. Alternatively, DI RNAs may arise from other segments with equal frequency, but since they are not able to replicate as efficiently as polymerase DI RNAs they are eventually eliminated. Sequence analyses did not reveal any unique sequence(s) that may favor the generation of influenza DI RNAs (see Sect. 8).

8 Generation of Influenza DI RNAs

Defective interfering RNAs are consistently observed in influenza virus passaged at high multiplicity. They are also present even in clonal stocks of the virus, although at much lower levels. These DI RNAs become amplified during subsequent passages at high multiplicity. How and why these smaller RNAs, with no apparently useful function in the biology of virus replication, arise with such high frequency is not well understood. Although influenza virus was the first virus for which defective particles were described, only recently have influenza DI particles been sufficiently characterized for any meaningful interpretations to be deduced about as their biology, structure, or function. One has to consider the following facts and observations before formulating a general model for their origin. Even though a considerable number of influenza DI RNAs have been sequenced to date, they may not represent examples of all the possible varieties of DI RNAs that may exist or arise in nature. The sequences of DI RNAs indicate a wide variation in structure, e.g., junction point or size. Furthermore, DI RNAs studied so far may not be the initial products of generation but, rather, represent the predominant species which have survived selection pressure. These DI RNAs may not be the end-product of evolution either, because it has been shown that any given DI RNA may further evolve or even disappear during subsequent growth-cycles of the virus (De and Nayak 1980; Nayak unpublished data). In fact, it is not unreasonable to assume that the majority of DI RNAs studied to date fall into the intermediate or evolving species (see Sect. 9).

Factors affecting the generation of influenza DI RNAs are poorly understood, though it would be possible to define many of them if one could generate DI RNAs in a more controllable in vitro replication system. Such a system is not yet available for influenza virus. The detailed steps involved in the replication of the standard virus are not clear at present, which makes it difficult to draw a model depicting the specific steps involved in generation of DI RNAs during replication of standard RNA segments. Sequence studies rule out the normal eukaryotic splicing mechanism or the post-transcriptional processing of the nascent RNA in the generation of DI RNAs (Nayak et al. 1982b); therefore, it was suggested that an aberrant replicative event(s), in which the

polymerase by some mechanism skips a portion of the template, is involved. Furthermore, it appears that such a process is not restricted to generating DI RNAs, but may also produce RNA segments which exhibit altered phenotypic characteristics. BUONAGURIO et al. (1984) reported that a 36-nucleotide internal deletion of the NS gene of influenza A/Alaska/6/77 produces a host-range variant virus; it is not known whether the deleted NS gene possesses interfering properties.

Two general mechanisms for causing internal deletion(s) can be suggested. The first is the jumping polymerase model which involves detachment and reattachment of the polymerase complex to the template (Fig. 5A). In this scheme, during replication of either the plus or the minus RNA strand, the polymerase complex detaches from the template and reattaches to it downstream, giving rise to shortened nascent RNA strands possessing internal deletions. Such a process requires that when the polymerase is detached from the template the nascent RNA strand remains attached to the polymerase. An essentially similar mechanism of detachment has been proposed to explain the copy-back synthesis of Sendai and VSV DI RNAs (KOLAKOFSKY 1976; HUANG 1977).

If we accept this hypothesis for the generation of influenza DI RNAs, the question arises: Is there any sequence or secondary structure which favors detachment or reattachment of the polymerase? Sequence studies show that the junction and the flanking sequences vary widely (Fig. 4) and do not reveal the presence of any unique sequence or obvious secondary RNA structure at either the junction or the flanking regions. It has been proposed that uracil-rich regions are possible sites where a polymerase may pause, detach, and reattach at a different site for reinitiation (FIELDS and WINTER 1982). However, uracil-rich regions are not present in all DI RNAs (in either plus or minus sense). The occurrence of uracil-rich regions in some DI RNAs may be a coincidence, since influenza RNAs are generally rich in uracil. If uracil-rich regions are more prone to cause polymerase detachment, then one would expect the same frequency of generation of DI RNAs from other segments of influenza RNA (in addition to polymerase genes), but this appears not to be the case. Polymerase is believed to pause and stutter at uracil-rich sequences to add adenine residues to the daughter RNA strand (ROBERTSON et al. 1981). No additional adenine residues are found at the deletion points in DI RNAs (NAYAK and SIVASUBRAMANIAN 1983). JENNINGS et al. (1983) observed a high frequency of GAA and CAA sequences at the junction points of DI RNAs; such sequences frequently occur in other regions of the viral RNAs as well. A limited homology at the junction of only a few DI RNAs ruled out hydrogen bonding between the nascent strand and the template RNA as the general mechanism for finding the attachment site after detachment. Therefore, detachment and reattachment of the polymerase (from the template RNA involved in the generation of influenza DI RNAs) appear to occur at random or recognize some unique feature of the RNP complex which has yet to be identified.

The second scheme in generating deletion can be described as the rolling polymerase model, which involves looping out of the template (Fig. 5B). In this model, the polymerase complex does not completely detach from the template; instead, the polymerase with the attached nascent daughter strand rolls

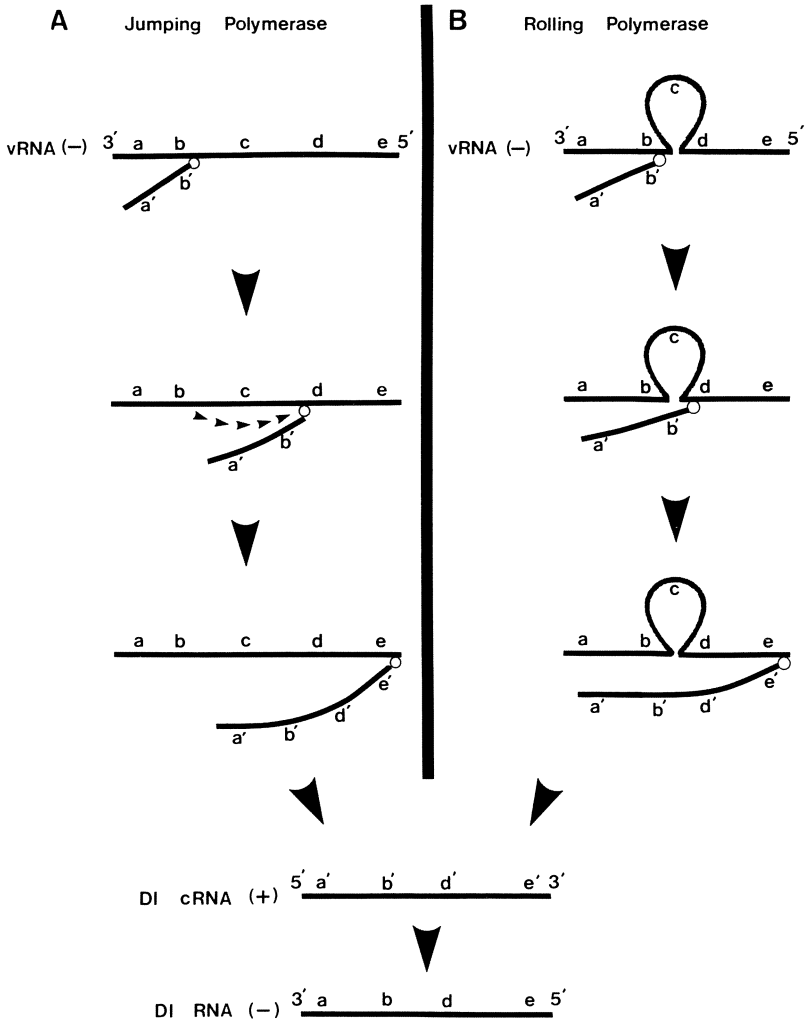


Fig. 5A, B. Models proposed for the generation of influenza DI RNAs from standard viral RNAs. Shown are negative-strand templates with sequence regions *a*, *b*, etc., and nascent positive-strand DI RNAs with complementary sequence regions a′, b′, etc. The *open circles* represent the functioning polymerase complex. Generation of DI RNAs could equally occur during replication of the positive-strand template

over to a new site of the template brought into juxtaposition (NAYAK and SIVASUBRAMANIAN 1983). The fact that the majority, if not all, of influenza DI RNAs are monogenic and not polygenic in origin also argues against complete detachment of the polymerase complex containing the nascent strand from the template and reattachment to a new site, and favors a roll-over/looping-out model. How the two sites of the RNA template are brought close to each other for the roll-over of the polymerase complex is not clear. One possibility is that such a juxtaposition of the RNA template is caused by the formation

of transient secondary structures of RNA. Influenza RNP complex is not a rigid structure but is flexible and, therefore, transient secondary structure(s) in the RNA template may occur during RNA replication because of the dynamic nature of the influenza RNA/protein interactions. A variation of the looping-out model has been suggested by JENNINGS et al. (1983) who propose that nucleoprotein structures of influenza RNP bring the 5′ and 3′ genomic termini close to each other making it possible for the polymerase complex to roll over from one region of the template to another.

9 Evolution of DI RNAs

Since none of the models presented above discriminate one standard RNA template from another, they can not explain why DI RNAs are predominantly formed from polymerase genes. Although a few subgenomic RNAs from genes other than polymerase (JENNINGS et al. 1983) have been found, none has yet been found to form a major DI RNA species or to possess either the replicative advantage or the interfering ability expected for a DI RNA (NAYAK 1980). In most DI preparations relatively few subgenomic RNA segments become the predominant RNA species. When more than one DI RNA species is present, they are not in equimolar concentration. Therefore, only some of the subgenomic RNA species possess a replicative advantage over other standard as well as other subgenomic RNA species. From these observations, three intriguing questions arise: (a) What features provide a subgenomic RNA with a replicative advantage? (b) What characteristics provide better interfering ability? (c) How are these subgenomic RNAs, with unique features of replicative advantage as well as interfering ability, selected?

Since only a few subgenomic RNA species become predominant in a DI preparation, DI RNAs are likely to represent a unique subgroup of the large number of subgenomic RNA segments that may be randomly generated. Sequence data have not revealed any unique characteristics of this subgroup of subgenomic RNAs. The smaller size of the DI RNAs may provide some replicative advantage over the larger RNA species. However, this cannot be the sole characteristic of a predominant DI RNA species for two reasons. (a) In DI preparations smaller RNA species are not always predominantly in molar excess over the larger DI RNAs. (b) All else being equal, smaller subgenomic RNAs are likely to be generated more frequently from small standard RNA genes, such as M or NS than from the larger polymerase genes. One would expect therefore to find more often a small predominant DI RNA species of M and NS gene origin than of polymerase gene origin; however, this is not the case. Sixteen predominant DI RNA species that have been examined are all of polymerase origin (NAYAK 1980), suggesting that subgenomic RNA species of polymerase origin possess some replicative advantage over other subgenomic RNA species. A possible explanation is that secondary structural features of some DI RNA species provide them with a specific replicative advantage so that they evolve into major DI RNA species. However, analysis and identification

of such secondary structural features of these single-stranded RNAs is still a major and as yet unsolved problem.

Although the specific attributes of subgenomic RNAs which are likely to become major DI RNA species are not yet known, two processes may be involved in such selection. (a) All DI RNAs, large or small, are generated directly from the progenitor gene, and the ones with the replicative and interfering advantages will become the predominant DI RNA species. (b) Some of the DI RNAs, if not all of the major DI RNA species, do not originate directly from the progenitor genes but from a precursor DI RNA. Some of the DI RNAs (the first generation product) must arise directly from the progenitor gene. The largest DI RNAs are good candidates for this pathway. Unfortunately, complete sequences of the largest DI RNAs, particularly those generated in the first or second high-multiplicity passage, have not been determined. However, since even the largest DI RNAs studied to date do not possess all the sequences that some of the smaller DI RNAs do (DAVIS and NAYAK 1979), some of the small DI RNAs may have originated directly from the standard progenitor gene.

There is a good probability that some influenza DI RNAs are generated from a precursor DI RNA. Since two DI RNAs, L2a-7 and L2a-17 (Fig. 3), contained an identical deletion (SIVASUBRAMANIAN and NAYAK 1983) it is unlikely that they arose independently from the progenitor PB2 gene. Rather, they may have come from a single precursor DI RNA which contained the common deletion and subsequently evolved into two DI RNA species. If this were the case, it would indicate that influenza DI RNAs undergo further evolution after they are formed. In fact, generation of progressively smaller DI RNAs after many passages suggests generation of DI RNAs from precursor DI RNAs and selection of a subclass of DI RNA species with a better replicative advantage and interfering ability. Clearly, some DI RNA species possess a better replicative advantage than others do (NAYAK and CHAMBERS unpublished data).

In summary, although subgenomic RNA species can be formed from all standard RNA segments, only a few, predominantly those of polymerase origin, have the ability to become predominant DI RNA species. These possess a replicative advantage and interfering ability and probably represent a subclass among the subgenomic RNAs. Sequence analyses have not yet revealed the special attributes of this unique subclass of subgenomic RNA species. Both pathways, i.e. independent generation of DI RNA species directly from the standard progenitor gene as well as progressive evolution of subgenomic RNAs in the process of becoming a major DI RNA species, may be operating in influenza virus replication and the selection of influenza DI particles.

10 Replication, Transcription, and Translation of Influenza DI RNAs

Like all other DI RNAs, influenza DI RNAs possess two unique biological characteristics: replicative advantage and interfering ability. Both of these properties must be attributed to either the unique feature(s) of DI RNAs themselves

or their transcriptional or translational products, or both. We have attempted (Sects. 8 and 9) to elucidate some of the unique structural features of DI RNAs. In this section, we shall try to assess whether influenza DI RNAs are capable of producing functional transcripts and whether these DI transcripts are translated into DI-specific polypeptides. A great deal of information on the transcription, translation, and replication of RNAs of the standard influenza virus has recently become available and was reviewed in detail by KRUG (1983) and MCCAULEY and MAHY (1983).

Early studies of influenza DI particles demonstrated that they contain decreased amounts of nucleic acid (ADA and PERRY 1956) and a reduced level of RNA polymerase activity (CHOW and SIMPSON 1971), suggesting that interference may be the product of these defects. BEAN and SIMPSON (1976) examined the transcription of purified high-multiplicity passaged virus in vitro and observed that, while these particles contain reduced levels of polymerase activity, they transcribe all eight segments of their genome.

Furthermore, sequence analyses (JENNINGS et al. 1983; NAYAK et al. 1982a, b; NAYAK and SIVASUBRAMANIAN 1983) revealed that most DI RNAs retain both the 3′ and 5′ ends of their parent gene segment intact, including usually 200–400 nucleotides from each end (see Sect. 7). Presumbably, these large homologous regions include the recognition sites for polymerase binding, transcription initiation and termination, and 3′ polyadenosine (poly(A)) addition. On this basis it was hypothesized (NAYAK et al. 1982b) that, unlike the 5′ DI RNAs of VSV which are not transcribed into mRNAs, influenza DI RNAs ought to be capable of undergoing transcription into mRNAs which, in turn, might be translated into DI-specific proteins. CHANDA et al. (1983) demonstrated that DI-particle preparations are capable of transcribing poly(A)$^+$ cRNAs corresponding to each DI RNA segment, as well as each standard segment, in vitro. The DI cRNAs transcribed in vitro are similar to the transcripts of standard virus RNAs in that (a) they require ApG or capped mRNA as a primer for transcription initiation; (b) they contain poly(A) of heterogeneous length at the 3′ end; and, (c) as expected, they are slightly smaller than the full-length DI RNAs when the poly(A) is removed. Hybridization analysis indicated that the DI cRNAs are otherwise faithful complements of the DI RNAs. While there are great differences among DI preparations, both the rate and extent of DI viral polymerase activity in general are reduced compared with those of standard virus preparations.

Transcripts of the DI-specific RNAs are also found in DI-infected cells. The analysis of influenza transcription in CEF cells infected at high multiplicity with fowl plague virus by HAY et al. (1977) demonstrated the presence of poly(A)$^+$ cRNAs for gene segments 9 and 10, which are probably DI RNAs. These cRNAs are maximally synthesized during the late period of secondary transcription. In particular, the transcription and replication of segment 9 increases steadily throughout infection and persists after the synthesis of standard RNA segments has declined. The time course of transcription and replication of segment 10 RNA, however, resembles that of the "late" viral gene segments, HA, NA, and M. Similarly, PONS (1980) observed the presence of extra RNAs in both the polysome-associated and nonpolysome-associated fractions of infected CEF cells following high-multiplicity passages of WSN virus.

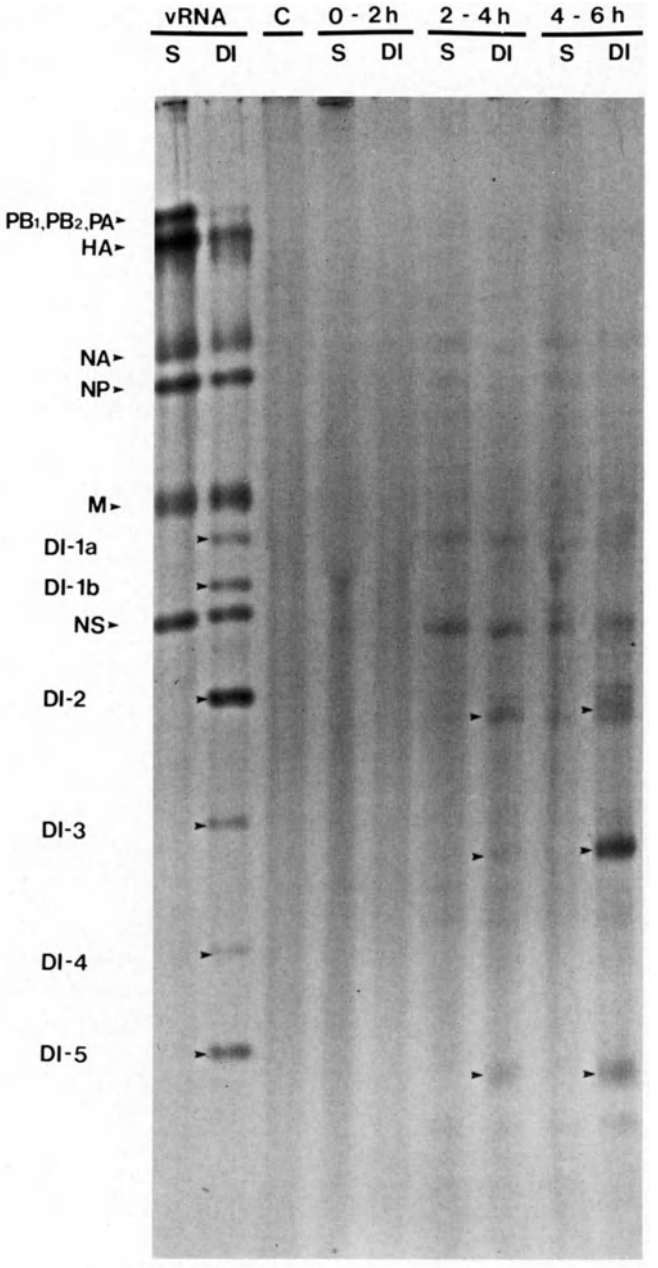


Fig. 6. Poly(A)$^+$ cytoplasmic cRNA from cells infected with standard and DI particles. S = RNA from infection of MDBK cells with WSN standard virus. DI = RNA from cells coinfected with standard virus and DI-ts$^+$ (Tobita). C = RNA from mock-infected cells. vRNA = marker RNA extracted from standard and DI particles. The indicated times are the ^{32}P labeling periods post-infection. At the end of the labeling period the cells were lysed and cytoplasmic RNA extracted. The poly(A)$^+$ and poly(A)$^-$ RNA fractions were separated by oligo-dT cellulose chromatography. Subsequently, poly(A) was removed from the poly(A)$^+$ fraction using poly(dT) and ribonuclease H (ETKIND et al. 1977). Electrophoresis was on a 3% polyacrylamide/6*M* urea gel. *Arrows* indicate positions of DI-specific poly(A)$^+$ cRNAs corresponding to DI RNAs. (From AKKINA et al. 1984a)

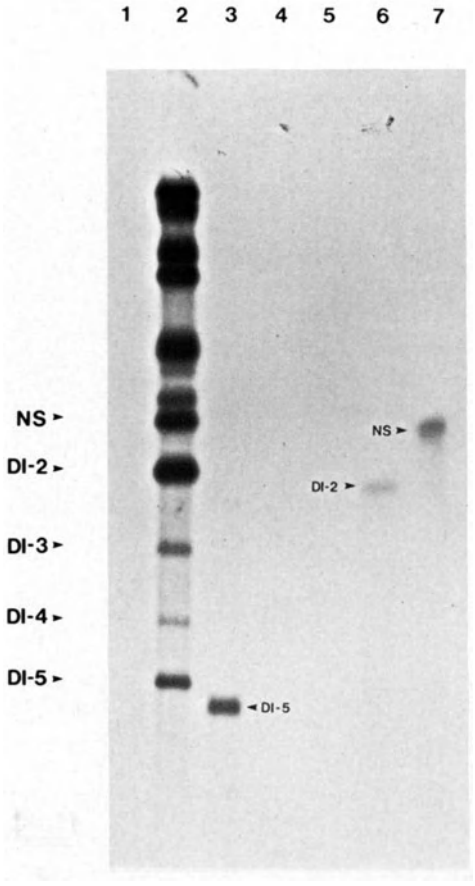


Fig. 7. Ribonuclease resistance of individual DI-specific poly(A)$^+$ cytoplasmic cRNAs after hybridization to the corresponding DI RNA segments. Individual ^{32}P-labeled cRNA bands and their corresponding unlabeled vRNA bands were eluted from gels (see Fig. 6) and specifically hybridized to each other. The hybrids were treatd with ribonucleases A and T1, then phenol/chloroform extracted, denatured, and electrophoresed in a 3% polyacrylamide/6*M* urea gel. Lane 2 is marker vRNA extracted from DI-ts$^+$(Tobita) particles. Lanes 3–7 represent specific hybrids of cRNA and vRNA of the DI-5, DI-4, DI-3, DI-2, and NS segments, respectively, with *arrows* indicating the positions of visible protected cRNAs. Lane 1 contains DI-2 cRNA equivalent to that in lane 6 but hybridized with itself, then treated as above. (From AKKINA et al. 1984a)

Our latoratory (AKKINA et al. 1984a; CHAMBERS et al. 1984) has recently analyzed mRNA synthesis in MDBK cells coinfected with WSN standard virus and DI preparations. As was observed in vitro, poly(A)$^+$ cRNAs corresponding in size to each DI RNA segment as well as standard segment mRNAs are present in coinfected cells (Fig. 6). Although it has not been shown, these DI-specific poly(A)$^+$ cRNAs are likely to also possess host-derived 5′ cap structures. Hybrids of specific DI transcripts and their corresponding minus-strand DI RNAs are resistant to ribonuclease digestion, indicating that the DI-specific transcripts were true complements of the DI RNAs and not merely incomplete transcripts of standard RNA segments (Fig. 7). Molar ratio analyses (Table 1) indicate that most DI RNAs are transcribed in molar excess over standard gene segments. The DI-specific RNAs in the poly(A)$^-$ cytoplasmic RNA fraction of these cells are also present in excess over the standard RNAs (Fig. 8). PENN and MAHY (1984) report the synthesis in vivo of mRNAs corresponding to subgenomic RNAs of PA and PB2 origin.

Because the DI RNAs are transcribed both in vitro and in vivo into poly(A)$^+$ cRNAs, it is likely that some of these may be translated into DI-specific polypep-

Table 1. Molar ratios of in vivo poly(A)$^+$ cRNAs of specific viral and DI RNA segments[a] (From CHAMBERS et al. 1984)

RNA	Viral segments					
	PB1+PB2+PA	HA	NP	NA	M	NS
DI-ts$^+$(Tobita) viral RNA	0.09	0.39	0.43	0.55	1.0	0.95
poly(A)$^+$ cRNA[b]	0.07	0.43	0.34	0.34	1.0	1.62
Standard viral RNA	0.29	0.62	0.58	0.71	1.0	1.10
poly(A)$^+$ cRNA[b]	0.09	1.07	0.63	0.56	1.0	1.54
RNA	DI segments					
	1a+1b	2	3	4	5	
DI-ts$^+$(Tobita) viral RNA	0.73	1.68	0.65	0.59	1.04	
poly(A)$^+$ cRNA[b]	1.46	6.19	1.81	2.80	4.17	

[a] Molar ratios normalized to M segment: $\frac{\text{(segment cpm/estimated nucleotide length)}}{\text{(M segment cpm/nucleotide length of M)}}$

[b] Cells labeled with ^{32}P, 3–6 h postinfection

tides. We and others (AKKINA et al. 1984a; CHAMBERS et al. 1984; PENN and MAHY 1984) detected several additional small polypeptides in cells infected with DI particles. Two DI influenza preparations of WSN origin were found to produce DI-specific small polypeptides (Fig. 9). Both of these DI preparations possess interfering properties as determined by infectious center reduction assays. The DI RNAs of these viruses are transcribed into poly(A)$^+$ cRNAs both in vitro and in vivo. One DI preparation contains a prominent DI RNA of about 700 nucleotides derived from the PB2 segment. MDBK cells coinfected with this DI preparation contained an additional polypeptide of about 22000 daltons (Fig. 9, lane 3). The second DI preparation contains minor and major DI RNAs of about 600 and 500 nucleotides derived from PB1 and PB2, respectively. Cells coinfected with this DI preparation contain an additional polypeptide of about 8000 mol. wt. (Fig. 9, lane 4). Both polypeptides are produced by in vitro translation using poly(A)$^+$ cytoplasmic RNA from the corresponding DI-coinfected cells, but not from standard virus-infected cells. Sucrose gradient analysis has shown that these DI-specific polypeptides are translated from small mRNAs which correspond to the size of DI RNAs, and therefore are not products of degradation of standard polypeptides or of premature translation termination of standard mRNAs. Hybrid-selected translation has shown that both of these DI-specific polypeptides are of PB2 origin, as expected. We have not, however, detected additional polypeptides in several other DI preparations. PENN and MAHY (1984) also report the presence of subgenomic mRNAs and small polypeptides in cells infected with a virus mutant containing subgenomic RNAs. These small polypeptides are related to the PB2 or PA proteins, as determined by tryptic peptide mapping, and are encoded by small mRNAs corresponding to the size of the subgenomic RNAs.

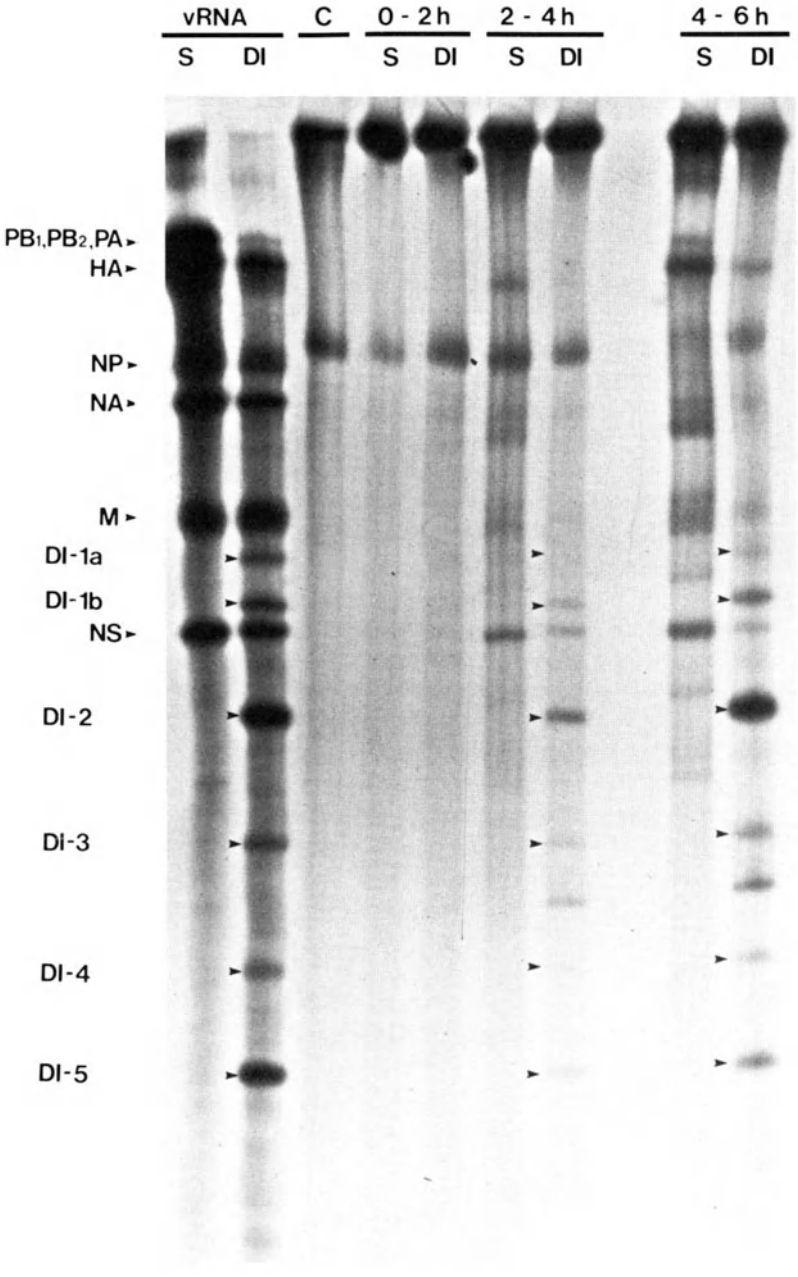


Fig. 8. Poly(A)$^-$ cytoplasmic RNA from cells coinfected with standard and DI particles. Figure labels and methods are as described in legend to Fig. 6. *Arrows* indicate the positions of DI-specific poly(A)$^-$ RNAs corresponding to DI RNAs. (From AKKINA et al. 1984a)

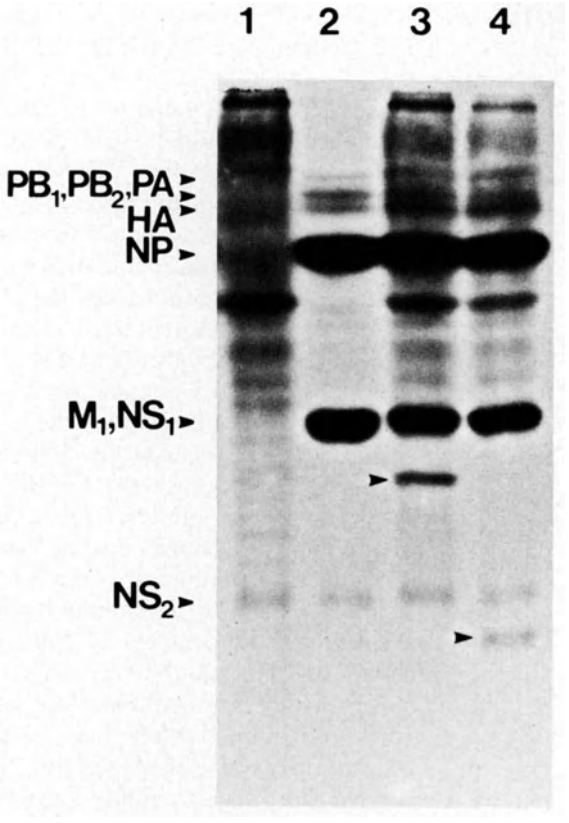


Fig. 9. DI-specific polypeptides. MDBK cells were mock-infected (lane 1), infected with WSN standard virus (lane 2), or coinfected with standard virus and DI-3 (lane 3) or DI-7 (lane 4). Polypeptides were labeled with ^{35}S-methionine, 4–5 h post-infection. Labeled proteins were analyzed by electrophoresis on a 13% polyacrylamide/4*M* urea gel. *Arrows* indicate the positions of DI-specific polypeptides

11 Mechanisms of Interference

11.1 Models Proposed for Other Virus Systems

HUANG (1977) summarized the important features of DI-mediated interference in animal viruses as follows: (a) Interference occurs intracellularly and not at the cell surface. (b) Interference is strongest against the standard virus from which the DI particles were derived, less strong against related viruses, and nonexistent against unrelated viruses. (c) Interference by DI particles is not mediated by interferon. These features also hold for interference mediated by the DI particles of influenza. However, a detailed mechanism of DI-mediated interference has yet to be elucidated for any virus group. Furthermore, the interfering mechanism may be different in different virus groups and even among different DI particles in the same group (PERRAULT 1981).

COLE (1975) suggested that DI-mediated interference in picornaviruses is caused by competition because the total amount of viral RNA, both standard and DI, synthesized in a DI-coinfected cell is the same as that in a cell infected with standard virus only. Consequently, less standard RNA is transcribed in

DI-coinfected cells and therefore less capsid protein is made. Furthermore, standard and DI RNAs compete for the limited amount of capsid protein during assembly. Alternatively, LUNDQUIST et al. (1979) argued that the enrichment of DI polioviruses is directly related to the size of the deleted genome segment. DI RNAs replicate faster than standard RNAs since the shorter DI genomes are transcribed and translated in a shorter time. In this model, because of initial compartmentalization of the input viral genomes, DI RNAs that retain the capacity to transcribe and translate a functional viral protease and polymerase are selected, while the capsid region of the genome may be deleted.

Several models have been forwarded to explain DI-mediated interference in VSV. Here it is known that (a) interference does not occur solely by virtue of the small size of the DI genome, (b) VSV DI RNAs do not inhibit primary transcription by the standard virus, and (c) most VSV DI RNAs are nontranscribing. The majority of VSV DI RNAs are the 5′ type and possess a 3′ terminus which is the inverted complement of the 5′ standard RNA terminus. PERRAULT et al. (1978) speculated that since the DI RNA possesses 3′ termini of only the plus-strand type on both plus and minus RNA strands, it must recognize only one kind of polymerase complex, while the standard RNA would recognize two kinds of polymerase complexes – one for transcription and the other for replication. If the synthesis by the minus-strand replicase is rate-limiting, the DI RNA gains a kinetic advantage, i.e., it can undergo rapid plus- and minus-strand synthesis while the standard genome is occupied with both template plus strand as well as mRNA synthesis. Alternatively, rates of nucleotide addition on the plus and minus strands may be the same, while the plus-strand type of 3′ terminus may have a higher polymerase binding affinity than the minus-strand type, thus, also providing the DI RNA with a kinetic advantage in replication (HUANG et al. 1978; HUANG 1982).

BLUMBERG and KOLAKOFSKY (1983) recently published a mathematical analysis of interference in VSV which supports a model for modulation between transcription and replication of the viral genome in response to the intracellular concentration of VSV nucleoprotein (N) (KINGSBURY 1974; LEPPERT et al. 1979). In this model, at low N concentration, the viral polymerase terminates at the end of the 46-mer 3′ leader sequence releasing the leader RNA, and then reinitiates to transcribe the VSV mRNAs in succession. At high concentration, N binds to the leader sequence, allowing the polymerase to read through the termination signal and complete the synthesis of full-length viral plus-strand antigenome. Read-through of the antigenomic leader sequence to synthesize progeny minus-strand genome proceeds via the same mechanism. The central assumption of this model is that the minus-strand genomic terminator signal requires a higher intracellular N concentration to be read through than does its plus-strand counterpart. Thus, the 5′ DI RNAs of VSV have a replicative advantage over standard RNAs because DI RNAs possess the plus-strand antigenomic type of terminator rather than the genomic terminator in both plus and minus strands. DI RNAs are replicated and encapsidated at lower N concentrations adequate for only transcription of the standard viral genome, preventing the N concentration from rising to a higher level sufficient for replication of standard antigenomes. The mathematical analysis of this model fits the reciprocal

exponential relationship between the multiplicity of infecting DI particles and the level of release of infectious progeny virus that was observed experimentally by BELLETT and COOPER (1959). RAO and HUANG (1982) described interference among different VSV DI RNAs. The hierarchy of interference was independent of input multiplicity and correlated with the length of DI RNA 3′ sequences complementary to the standard 5′ end – the greater the extent of complementarity in a 12-nucleotide region adjacent to the leader sequence, the greater the competitive advantage of that DI RNA. This can also be understood in the context of regulation by viral protein binding to this region.

However, PERRAULT et al. (1983) recently described VSV mutants in which transcription termination at leader sites is specifically suppressed. Reconstitution experiments implicate the N of the template nucleocapsid, rather than newly synthesized N, as being responsible for the read-through property. Therefore, they proposed an alternative model in which modulation from transcription to replication is due to a modification of the N of the template nucleocapsid and not to homeostatic regulation of N synthesis. However, the DI RNAs may still have the other kinetic advantages discussed previously.

A few VSV DI RNAs, such as DI-LT, are the 5′-3′ type and are generated by internal deletion rather than by a copy-back mechanism. DI-LT is self-transcribing, interferes with standard virus primary transcription, and also interferes heterotypically. Analysis of temperature-sensitive mutants of these particles implied that in cells coinfected with DI-LT and standard virus, replication of DI-LT can take place in the absence of DI-LT primary transcription or of DI-LT-mediated interference. However, self-transcription of the DI-LT genome is necessary for interference with standard virus primary transcription (BAY and REICHMANN 1982). This interference is dependent on the relative concentration of input DI and standard viruses, and was interpreted to be the result of competition during primary transcription for a limited number of polymerase molecules. It was further hypothesized that the advantage of the DI-LT RNA lies in a decreased rate of polymerase dissociation from the template during the course of elongation, although the polymerase affinity at the binding site of the DI RNA is the same as in the standard RNA.

11.2 The Mechanism of DI-Mediated Interference in Influenza Virus

The mechanism of interference by influenza DI particles remains undetermined. Experiments by NAYAK et al. (1978) in which MDBK cells were superinfected with DI particles at different times after infection with standard virus indicate that the occurrence of interference may not be restricted to the earliest stages of infection. DI particle superinfection as late as 3 h after standard virus infection caused 80% inhibition in standard PFU released; even as late as 6 h after standard virus infection, DI particle superinfection inhibited release of infectious virus by 50%.

As discussed in Sect. 5, we have shown that the influenza DI RNAs are specifically responsible for interference (JANDA and NAYAK 1979). Interference in influenza is therefore most likely to be associated with the transcription,

translation, or replication of DI RNAs, as well as with the assembly of vRNAs into progeny virions – a process that must be substantially different from that of the nonsegmented viruses. The mechanism of assembly and packaging of eight different standard viral RNAs into an infectious viral particle remains unknown, although our recent experiments (AKKINA et al. 1984b) suggest selective rather than random packaging of RNP complexes during virion assembly (see sect. 11.2.5).

11.2.1 Effect of DI Particles on Transcription and Replication of Standard RNAs

It is clear that influenza DI particles cause some alteration in the mRNA transcription of the standard gene segments. Virus particles purified after repeated high-multiplicity passage exhibit a two-to-sixfold decrease in RNA transcriptase activity (BEAN and SIMPSON 1976; CARTER and MAHY 1982a; CHANDA et al. 1983; PONS 1980), although virus from the very first undiluted passage sometimes shows increased activity (CARTER and MAHY 1982a; PONS 1980). This increase is believed to reflect early amplification of DNI particles and their subsequent replacement by DI particles (CARTER and MAHY 1982a). Repeated high-multiplicity passages result in reduced synthesis of total viral RNA and polysome-associated RNA in infected cells (PONS and HIRST 1969; PONS 1980).

Most striking is the decreased synthesis of polymerase gene transcripts in cells coinfected with standard and DI particles (CHAMBERS et al. 1984) and in cells infected with high-multiplicity-passaged virus (PONS and HIRST 1969; PONS 1980). Molar ratio analyses indicate that in DI-coinfected cells the DI RNAs are preferentially transcribed, accounting for about 80% of total viral transcript molecules in the cytoplasms of these cells (Table 1) (CHAMBERS et al. 1984). However, in vitro transcription of purified DI particles showed no evidence for preferential primary transcription of DI RNAs; the relative levels of the standard and DI mRNAs produced correspond closely to those of the viral RNAs in the DI particles (CHANDA et al. 1983). Therefore, reduction of standard RNA transcripts in vivo may reflect the effect of DI particles on the secondary transcription, which, in turn, may be due to interference in replication rather than transcription of standard RNAs.

Replication of the standard viral segments is also altered after coinfection with DI particles. Again, the DI RNAs appear to be replicated in preference to standard gene segments (AKKINA et al. 1984a; CHAMBERS et al. 1984) and replication of the polymerase genes is reduced (CARTER and MAHY 1982c; PONS 1980). Analysis of intracellular poly(A)$^-$ cRNAs synthesized in successive undiluted passages indicates that the inhibition of polymerase gene replication is an early event occurring in the very first high-multiplicity passage (CARTER and MAHY 1982c). The standard polymerase gene content of the progeny virus from undiluted passage is also reduced (CARTER and MAHY 1982c; CHOPPIN and PONS 1970; CRUMPTON et al. 1978; DUESBERG 1968; JANDA et al. 1979; PONS 1980). With different DI preparations, a preferential inhibition of PB1 (UEDA et al. 1980), PB2 (NAKAJIMA et al. 1979; PONS 1980), or PA (JANDA et al. 1979) has been observed, while in others all three polymerase segments

are reduced in almost equal proportions (CRUMPTON et al. 1978, 1981; BEAN and SIMPSON 1976; NAYAK et al. 1978). In addition, some DI preparations show decreased levels of the HA, NP, and NA genes as well as the polymerase genes (CRUMPTON et al. 1981; NAYAK et al. 1978). In summary, although a reduction in the synthesis of standard vRNA segments (particularly polymerase genes) is observed in DI-coinfected cells, the mechanism of this reduction is not understood. Furthermore, as yet there has been no detailed analysis of the effect of coinfecting DI particles on the rates of transcription and replication of individual standard viral segments.

11.2.2 Effect of DI Particles on Translation of Standard Polypeptides

DI particles may also affect the levels of synthesis of the intracellular virus-specific proteins although results vary with different DI preparations. For example, LENARD and COMPANS (1975) and PONS and HIRST (1969) have earlier reported slight or occasional reductions in the total level of polymerase proteins as well as other proteins in DI particles compared to that in standard virus particles. However, these workers did not determine levels of individual polymerase proteins. Since polymerase proteins play a pivotal role in replication and consequently, in DI-mediated interference, we have studied the synthesis of specific polymerase proteins in cells coinfected with standard and DI particles (AKKINA et al. 1984b). In cells infected with standard virus particles the synthesis of the three polymerase proteins was nearly equimolar but in cells coinfected with standard and DI particles, the synthesis of specific polymerase proteins (PA, PB1 and PB2) relative to each other varied with different DI preparations (Fig. 10). With DI_3, DI_7 and DI-Tobita, both PA and PB2 were reduced whereas PB1 appeared to be synthesized in normal amounts. Similar patterns of reduction were observed in the in vitro translation products of cytoplasmic mRNAs from DI-coinfected cells, suggesting that the reduction of specific polymerase proteins in cells coinfected with standard and DI particles was due to decreased levels of specific polymerase mRNAs rather than the inhibition of translation of specific polymerase mRNAs. Furthermore, the decrease in specific polymerase mRNAs occurred also at the level of primary transcription which in turn reflected the ratio of the specific polymerase genes in a given DI preparation.

We also showed that although there was a great variation among the intracellular concentrations of the three polymerase proteins, they were in nearly equimolar ratio in the DI particles released from coinfected cells. This would suggest that all three polymerase proteins are present in viral RNP in the form of a complex with a ratio of 1:1:1, and that before RNPs are used for virus assembly they are likely to be screened for the completeness of the polymerase complex (AKKINA et al. 1984b).

11.2.3 Role of DI-Specific Polypeptides in Interference

Although DI-specific polypeptides have been detected in some DI preparations, their role, if any, in interference is not clear. As discussed above, the level

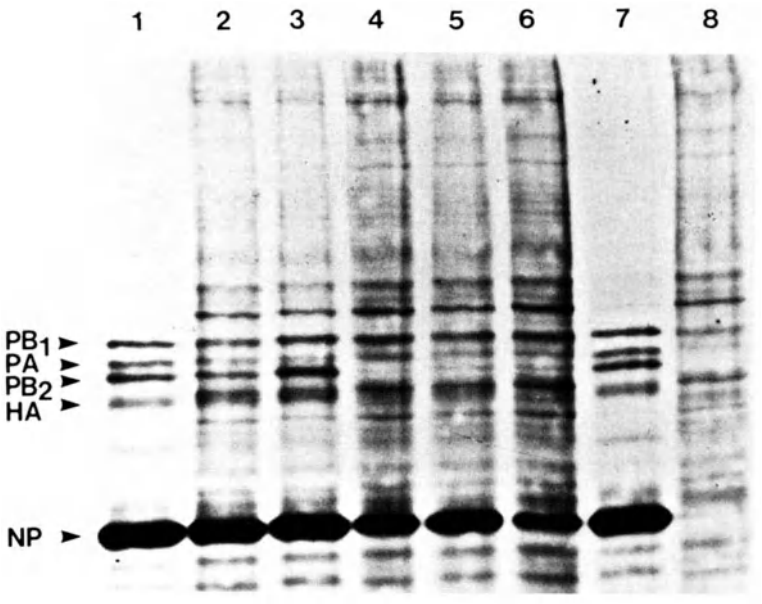


Fig. 10. Analysis of polymerase protein synthesis in DI-infected cells. MDBK cells were infected with WSN standard virus (lanes 1 and 7), or coinfected with standard virus and DI-4 (lane 2), DI-6 (lane 3), DI-7 (lane 4), DI-ts$^+$(Tobita) (lane 5), or DI-3 (lane 6); or mock-infected (lane 8). Intracellular polypeptides were labeled with ^{35}S-methionine, 2.5–3.5 h postinfection. Labeled proteins were analyzed by electrophoresis in an 8% polyacrylamide/4*M* urea gel. Only the top portion of the gel is shown. PA is relatively decreased in lanes 3 and 5, and PB2 in lanes 4, 5, and 6. A host cell protein (lane 8) closely comigrates with PB1 in the DI preparations

of different standard proteins in DI-coinfected cells is also altered and reduction in the level of standard proteins reflects the reduced level of mRNAs rather than a direct interference at the level of translation.

If the temporal regulation of influenza RNA synthesis is mediated by intracellular concentrations of viral proteins, as suggested (McCauley and Mahy 1983; Smith and Hay 1982), then conceivably this mechanism could be usurped by DI-specific polypeptides to give preference to DI RNA synthesis at the expense of polymerase gene synthesis. The polypeptides encoded by DI RNAs of polymerase gene origin would possess amino acid sequences like those of polymerase proteins at the amino-terminal region up to the point of deletion, and might similarly bind to and thus block replication initiation sites. However, not every DI preparation produces detectable DI polypeptides (Akkina et al. 1984a; Chambers et al. 1984; Penn and Mahy 1984), nor is there evidence that polypeptide-producing DI particles have an advantage in interference over other DI particles. All DI preparations, whether or not expressing visible DI-specific polypeptides, possessed interfering activity. In experiments where cells were doubly infected with equal DIU of two DI preparations, one of which produces DI-specific polypeptides, analysis of progeny particle-RNA showed that replication of the DI particle producing a prominent DI polypeptide is not favored over the other DI particle (Akkina et al. 1984a). In addition, there

Table 2. Dose-dependent reversal of DI-mediated interference[a] (From AKKINA et al. 1984b)

Infecting virus		Progeny virus		
DIU/cell	PFU/cell[b]	PFU/ml	HAU/ml	PFU/HAU ratio
0	0.5	3.6×10^8	4096	90000
4	0.01[c]	1.1×10^4	1536	7
4	0.5	2.1×10^4	1536	14
4	1.0	4.3×10^4	2048	21
4	2.5	1.3×10^5	2048	63
4	6.25	3.6×10^5	3072	120

[a] MDBK cells were infected with standard virus and DI-Tobita. Progeny virus was harvested for the above assays at 24 h pi
[b] Helper standard virus supplied
[c] Contamination of DI preparation with residual standard virus

is no evidence for selection of DI RNAs which maintain the translation reading frames in-phase through the deletion point (JENNINGS et al. 1983). Lastly, although DI_3 and DI_7 polypeptides are produced relatively in abundance in DI-coinfected cells (AKKINA et al. 1984a) we have been unable to detect their presence in DI particles released from DI-coinfected cells. It is, however, possible that our analytical methods may not be sensitive enough to detect the presence of a few DI-specific polypeptides in DI particles. In summary, most DI polypeptides are expected to be short, compared to full-length polymerase proteins, and to possess at the amino-terminal region the same amino acid sequences as does the standard polypeptide. However, the role, if any, of DI-specific polypeptides in interference remains undetermined.

11.2.4 Partial Reversal of DI-Mediated Interference with Increased Concentration of Standard Virus Particles

Our laboratory has performed experiments to determine if increased multiplicity of standard virus would overcome DI-mediated interference (AKKINA et al. 1984b). The results (Table 2) show that the interference with the production of standard virus particles by DI-Tobita was partially reversed with increasing amounts of standard virus. The reversal of DI-mediated interference was clearly dose-dependent. When standard virus was increased twelvefold, there was a seventeenfold increase in PFU/HAU ratio. Similar results were also obtained with DI_7 (data not shown). Also, with increasing standard virus concentration, the synthesis of polymerase proteins became more equimolar as in standard virus-infected cells. These data suggest that interference by DI particles is the outcome of competition for transcription and replication between the DI and standard viral genomes in the coinfected cell. However, it is likely that interference can never be totally reversed in such experiments, because very high standard virus MOI may result in the generation of new DI RNAs.

11.2.5 Effect of DI RNA on Virus Assembly

The mechanism of assembly and packaging of eight different influenza gene segments into an infectious viral particle remains unknown. The genomic content of an individual influenza virion, standard or DI, is also unknown, save as the mean values of a heterogeneous population. Although all eight gene segments are required for infectivity, it is not known if the eight standard RNA segments are incorporated into virions by a mechanism involving selective or random packaging. SMITH and HAY (1982) have raised the possibility of a selective packaging mechanism, to explain the differences they observed between the relative proportions of intracellular vRNAs and progeny particle vRNAs. Furthermore, they have speculated that packaging of "incomplete" RNAs may be related to reduced packaging of polymerase vRNAs. As discussed above (Sect. 11.2.1), this reduction in polymerase gene content is specifically characteristic of different influenza DI preparations. However, any competitive advantage of a given DI RNA in transcription and replication would cause a general reduction in the synthesis of standard RNA segments but would not explain the reduction of a specific polymerase gene segment in a given DI preparation. Because the DI RNAs are mutants of polymerase RNAs, it can be speculated that during the virus assembly process DI RNAs are somehow mistaken for polymerase RNAs and packaged in their place. Interference would result if the DI RNA were to specifically displace its progenitor gene from the progeny particle during assembly. Evidence indicates that this may be so; we and others have observed that with several DI preparations, the presence of major DI RNAs of, for example, PB2 origin was correlated with specifically decreased levels of PB2 vRNA in that preparation (AKKINA et al. 1984b; NAKAJIMA et al. 1979). However, whether this correlation would be a general rule in explaining the reduction of polymerase genes in influenza DI particles requires more careful analysis. Any advantage of the DI RNA operating during virus assembly may be critically important for the perpetuation of DI particles, because the competitive advantage during transcription and replication might not be very effective in producing interference. Furthermore, such a mechanism would favor selective rather than random packaging in the assembly of a standard virion containing eight different RNA segments. Thus a detailed inquiry into the events leading to the genesis of DI particles may elucidate the process involved in assembly and budding of influenza virus particles.

11.2.6 Possible Mechanisms of Interference

Like other DI viruses, the mechanism of interference with standard influenza particles by influenza DI particles is complex and involves multiple steps in the virus life cycle. Any proposed mechanism of interference must first account for the generation of DI RNA segments which appear to be produced due to an error during replication (see Sect. 8). Secondly, such a mechanism must be able to account for preferential replication of DI RNA segments at the expense of standard viral RNA segments. However it is unlikely that the replica-

tion of a specific standard RNA segment (e.g. PB1, PB2 or PA) will be interfered with by a given DI RNA segment because the same polymerase complex is involved in the replication process of all eight standard as well as DI RNA segments. Finally, the proposed mechanism must take into account the process of packaging eight RNA segments into a virion. This step of packaging multiple RNA segments into virions must clearly be different from that in the nonsegmented RNA viruses, which package only one RNA segment – either the DI or the standard viral RNA – into a virus particle.

The Kingsbury-Kolakofsky model for interference in VSV (KINGSBURY 1974; LEPPERT et al. 1979) (see Sect. 11.1) does not explain the mode of interference in influenza. First, no copy-back influenza DI RNAs have been discovered. The DI RNAs that have been sequenced retain both the 5′ and 3′ terminal regions of their standard progenitor genes and, therefore, have no known regulatory advantage. Second, because of the segmented nature of the influenza genome and the structural difference between its replicative templates and mRNAs, modulation between transcriptive and replicative polymerase activity must occur at initiation rather than during elongation. That this modulation is a function of levels of particular viral proteins (MCCAULEY and MAHY 1983; SMITH and HAY 1982) is an attractive hypothesis, but as yet there is no evidence for it. Third, the relation between the multiplicity of input DI particles and the yield of progeny standard virions is a reciprocal exponential relationship in VSV (BELLETT and COOPER 1959), whereas with influenza it is an inverse proportionality (NAYAK et al. 1978). This would support a model for interference based on competition among templates for initiation. Fourth, interference in VSV is overcome by delaying the superinfection with DI particles until after the onset of secondary transcription (i.e., at 2.5 h), after which both DI particles and the full yield of standard progeny virions are released (HUANG and WAGNER 1966). In influenza, as mentioned above, there is 50% inhibition of standard virus release even when DI superinfection is delayed by 6 h (NAYAK et al. 1978). This would also support a model involving competition for initiation.

Two possibilities exist to explain the competitive advantage of the DI RNA:

1. There may be as yet unrecognized internal regulatory features in the standard gene segments that are either lacking altogether in the DI RNA or are altered by the deletion. By this hypothesis, the size and location of the deletion is as critical as the total size of the RNA; DI RNAs would be selected on the basis of structure. We have observed (JANDA et al. 1979) that the smallest DI RNAs are not always the most strongly replicated. Conceivably, the binding affinity of the viral polymerase may be partly dependent on the adjacent secondary structure of the RNA as well as its 3′-terminal sequence. Such an internal polymerase recognition structure has been found in $Q\beta$ bacteriophage (BLUMENTHAL and CARMICHAEL 1979). Similar polymerase recognition structures may be present among different influenza viral RNA segments and may be involved in the temporal regulation of viral RNA synthesis. The polymerase RNAs are transcribed and replicated at the lowest level of all the viral segments; it is thus possible they possess some internal attenuating feature which is missing in the DI RNAs. However,

there is as yet, no evidence for such an internal attenuation site in influenza RNA.

2. DI RNAs may be identical to their parent genes from the standpoint of regulation of RNA synthesis and the only competitive advantage of the DI RNAs would be in their small size. A similar model has been proposed to account for the advantage of the subgenomic RNAs of $Q\beta$ bacteriophage (MILLS et al. 1967) and has also been proposed for picornavirus DI particles (LUNDQUIST et al. 1979) (see Sect. 11.1). If so, although the DI RNAs would be synthesized in molar excess, the replication of standard genes such as polymerase genes may be reduced but never entirely eliminated. Therefore, a competitive edge for replication and transcription primarily based on smaller size of DI RNA may be the major factor in influenza DI-mediated interference. Partial reversal of DI-mediated interference by increased multiplicity of standard particles would also support this hypothesis. Such a mechanism would predict that the molar concentration of DI RNAs will be strictly size-dependent, which, however, is not always the case (JANDA et al. 1979). These exceptions could be explained by supposing that in such cases the larger DI RNAs were generated earlier and thus had the advantage of a higher input ratio. If passage is continued for a large number of generations, smaller DI RNAs will eventually overtake the larger DI RNAs (DE and NAYAK 1980) (see Sect. 9). When cells are infected with more than one DI particle, the input multiplicity along with the size will determine which DI RNA will be predominant in the yield.

In addition to any competitive advantage in amplification due to regulatory features or to smaller size, DI RNA may be further selected during virus assembly if a particle containing a DI RNA were to exclude its progenitor gene from the same particle (see Sect. 11.2.5). Such a mechanism of exclusion of progenitor RNA due to the presence of DI RNA may explain why, in some DI preparations, specific polymerase genes are greatly reduced or lacking (NAKAJIMA et al. 1979; JANDA et al. 1979; UEDA et al. 1980; PONS 1980). The study of DI-mediated interference may therefore aid in the understanding of the important influenza viral assembly process.

12 Role of DI Particles in Viral Pathogenesis and Virus Evolution

12.1 Viral Pathogenesis

Since DI particles were discovered over 35 years ago with influenza viruses, virus particles possessing similar properties have been reported with almost all DNA or RNA viruses irrespective of whether they contain single-stranded or double-stranded, positive sense or negative sense, unsegmented or segmented genomes (HUANG 1975; HUANG and BALTIMORE 1977; PERRAULT 1981; HOLLAND et al. 1980). Similarly, they are found in viruses with different biological properties: either lytic or nonlytic, oncogenic or nononcogenic. The majority

of the highly transforming oncogenic RNA viruses are defective, i.e., they need the helper function of a nontransforming replication-competent virus. However, unlike DI particles these defective oncogenic viruses do not interfere with the growth of replication-competent viruses and are usually outnumbered by them. DI particles also have been observed among plant, yeast (KANE et al. 1979) and bacterial viruses (MILLS et al. 1967). Therefore, generation of DI particles is not restricted to influenza or to a particular group of viruses; rather, it is a universal phenomenon among all viruses. Neither is the generation of DI particles restricted to a group of hosts. Although host cells play an important role in the generation and amplification of DI particles, DI particles are produced in most permissive host cells. With influenza viruses DI particles have been produced in embryonated chicken eggs and a variety of cell cultures (MDBK, MDCK, CEF). Influenza DI particles have been generated in animals. GINSBERG (1954) reported the production of "incomplete" virus in mice as a result of intranasal inoculation of large doses of virus. Generation of DI particles both in cell cultures and in experimental animals has been reported for VSV, Sendai, reoviruses, rabies, lymphocytic choriomeningitis, and many other viruses as well (HUANG 1975; HUANG and BALTIMORE 1977; HOLLAND et al. 1980). In short, any host system that will support the growth of a standard virus will also produce and amplify DI particles. In fact, our studies show that the growth of DI and standard particles is not always mutually exclusive (DE and NAYAK 1980). The standard virus has to replicate to a certain level to provide sufficient factors necessary for the optimum amplification of DI particles. The virus preparations that contain a relatively large amount of DI particles often also contain a relatively large amount of standard particles.

Since in the natural mode of infection in influenza the multiplicity of infection (MOI) is very low, how are DI particles generated and amplified? This is a critical question. Our data in cell culture show that DI particles can be generated and amplified at very low MOI and even in individual viral plaques. In naturally occurring influenza infections, although the initial infection occurs at a low MOI, subsequent infection of adjacent cells will occur at a high MOI. Therefore, any DI particle generated during the early phase of infection is likely to be amplified in the neighboring cells coinfected with both standard and DI particles, since many cycles of infection are required before the virus can produce pathognomonic syndromes and lesions. DI particles may exert their influence in modifying the outcome of the disease by interfering with the growth of the standard virus during the multicycle replication. What is the likelihood of generating DI particles during an infection? We have shown that it is most difficult to obtain a DI particle-free influenza virus stock and to maintain such a stock. Almost every stock of influenza A and B viruses we have examined contains visible DI RNA bands in gels and therefore, must contain a relatively large amount of DI particles (JANDA et al. 1979; DE and NAYAK 1980). Our data show that DI particles are being continuously generated. The generation of influenza DI particles is therefore a common and natural phenomenon almost always occurring during the replication of standard viruses. The structure of influenza RNA or the nature of the polymerase may be responsible for the frequent generation of DI RNAs. However, the amplification of newly generated

DI particles will depend on other factors including subsequent infection at high multiplicity and the presence of preexisting DI particles. Therefore, although no definitive studies have been done to demonstrate and quantify the number of DI particles produced in influenza lesions during natural infection, DI particles are likely to be produced and amplified in individual animals or humans during the course of the disease even though the infection is transmitted at a very low multiplicity from person to person.

If we accept the premise that DI particles are produced and amplified during the course of influenza infection we may ask the next question: What is their role in the outcome of the disease and the biology of the viruses? Here again, we can draw some conclusions on the basis of experimental data obtained both in cell culture and in experimental animals. In cell culture, DI particles ameliorate the effect of lytic virus by reducing cytopathic effect. Cells coinfected with DI particles can survive the cytolytic effect of standard viruses. Coinfection by DI particles also reduces the drastic inhibitory effect of standard influenza virus on cellular protein synthesis (Figs. 9 and 10). Such a coinfection of DI and standard particles will have two important sequelae. (a) Cells will remain healthier for a longer period of time, and (b) fewer standard virus particles will be produced. In cell culture, the outcome of such a process will help in establishing a persistent viral infection (DE and NAYAK 1980), while in humans or animals, the effect may be more complicated. Because of the healthier cells and fewer standard viruses, DI particles may limit the disease process and favor recovery by stimulating host defenses. Alternatively, such virus-infected cells may survive and evade the host immune mechanism, producing a chronic persistent viral disease (HOLLAND et al. 1980).

Soon after the discovery of incomplete particles by VON MAGNUS (1947), it was reported that undiluted A/PR/8/34 virus enriched in "incomplete" particles resulted in a reduction in mortality in mice when administered intranasally, as well as in a decrease in virus titer in lungs compared with the effect of standard virus produced at a low multiplicity (VON MAGNUS 1951b; GINSBERG 1954; HORSFALL 1954, 1955). DOYLE and HOLLAND (1973), in their studies with influenza DI particles, also reported some reduction in viral pathogenicity with increased survival time and reduced virus titer in the lungs of mice coinfected with influenza DI particles. However, mice inoculated intranasally with the DI preparation invariably died. It is quite possible that in these studies coinfection with DI particles was insufficient to protect against the effect of a relatively large amount of standard virus. Factors such as strain and age of mice as well as the ratio of DI particles to standard particles were studied by RABINOWITZ and HUPRIKAR (1979) in an attempt to further define the role of DI particles in viral disease. They showed that resistance to DI-enriched influenza virus was age-dependent and varied with the mouse strain. Seven week old Swiss and four week old C57B16Cr mice survived whereas three week old Swiss mice succumbed to intranasal inoculation with virus. They found that the DI particle/standard particle ratio was critical for protection, as was reported by DOYLE and HOLLAND (1973) for VSV. With increased enrichment in DI particles, influenza virus grew to a lower titer and lungs contained reduced lesions and reduced amounts of viral antigen, as demonstrated using immunofluorescence. Further-

more, the authors reported that mice inoculated with virus enriched in DI particles developed humoral immunity earlier and to a higher antibody titer than those inoculated with standard virus. Therefore, they concluded that humoral immunity rather than homologous autointerference with the standard virus was responsible for the increased protection produced by DI particles with influenza virus as with many other viruses.

Influenza DI particles appear to play an important role in the establishment of persistently infected cultures (DE and NAYAK 1980; FRIELLE et al. 1984). These persistently infected cells are resistant to superinfection by standard homologous virus, but not to the heterologous virus, and persistently infected cells show little cytopathic effect after superinfection with standard virus. Most of these cells possess viral antigens on the cell surface. However, unlike persistent infection established with VSV, DI particles could not be demonstrated after the persistent influenza infection was established in cell cultures. Although DI particles help in the establishment of persistent infection in cell culture, persistent infection in either humans or animals with influenza viruses appears to be rare. Therefore, the contribution, if any, of influenza DI particles in a natural infection would be to reduce the disease syndrome and aid recovery rather than to prolong the disease process.

DI particles can modulate host response in a number of ways. They can reduce growth of standard virus, help in mounting a specific immune response (GAMBOA et al. 1975), and modulate expression of virus-specific antigen on cell surfaces (WELSH et al. 1977). "Incomplete" influenza particles were found to be less toxic via either the intravenous or intracerebral route than the standard influenza virus (MCKEE 1951; MANIRE 1957) DI influenza particles would therefore make better vaccines. Indeed, it can be postulated that an ideal vaccine would consist of live viruses which, upon intranasal inoculation, would produce only "incomplete" particles; such a vaccine would not cause disease but would stimulate both local and systemic humoral antibodies as well as a cellular immune response.

Finally, although it remains unclear whether influenza DI particles play an important role in modulating viral pathogenesis in nature, recently it has been reported that DI RNA segments were found to be associated with relatively mildly pathogenic chicken virus isolates but were absent from highly pathogenic virus isolates (WEBSTER et al. 1984) which caused a loss of over $60 million in the poultry industry in 1983. Although the reason for the conversion of this relatively mildly virulent virus into a highly virulent virus has yet to be determined, the role of DI virus in modulating viral pathogenesis should be considered.

12.2 Virus Evolution

If we assume, as discussed above, that DI particles are also produced and amplified in natural infections during the replicative cycles of viruses, the next question arises: What is the role of these particles in the natural evolution of viruses? We assume that generation and amplification of DI particles would

not be built into the replication process of viruses unless an advantage were thereby provided in their evolution. Of course, reducing the pathogenesis of standard viruses would be beneficial to the host in recovery and in elimination of the virus, but such an effect does not offer an advantage to the survival and propagation of the virus itself. On the other hand, the establishment of long-term persistent infections in humans or animals may help the survival of a virus and enable it to find another susceptible host for replication. However, influenza virus is not known to cause long-term persistent infections in human beings or in animals, except for ducks and water fowl (HINSHAW et al. 1980). Influenza infection in most animals and humans is of relatively short duration and the virus is eliminated from the host with recovery from the disease. Thus, one might ask: How would these DI particles offer an advantage in the survival of influenza virus in these host-virus systems?

Influenza virus is perpetuated in nature by evading the host immunity. It uses two mechanisms, antigenic shift and antigenic drift, to change its outer envelope proteins (HA and NA) and, therefore, to escape the effect of host immunity. Antigenic shift, a much more drastic change, requires acquisition of genes for the new envelope antigen(s) and is believed to originate from gene exchange with a nonhuman influenza strain or from the reemergence of an older strain. DI particles are unlikely to play any major role in antigenic shift. Antigenic drift, on the other hand, is a gradual change involving mutation in the nucleic acid sequence and selection of appropriate mutants. A number of factors, particularly host immunity, are involved in the selection of mutants responsible for antigenic drift; however, the role of other factors cannot be eliminated. KILBOURNE et al. (1983) postulated from their studies on genetic dimorphism that other factors, such as variation at the receptor site, may be involved in the selection of antigenic variants. It is possible that DI particles which interfere with the growth of homologous viruses play a role in the selection of variants. DI particles may be involved in causing or selecting mutants. Although there is no published data showing that DI particles can act as mutagens and enhance mutation rates, HOLLAND et al. (1979, 1980, 1982) and ROWLANDS et al. (1980) showed that in cells persistently infected with VSV progeny viruses undergo mutation with time. This change is reflected in the nucleic acid sequence, amino acid sequence of the proteins, and the biological behavior of the viruses. These authors also hypothesize that DI particles are responsible for the observed alteration in the viral genome in persistently infected cells since alteration could not be demonstrated in acute infection with standard VSV. WEISS and SCHLESINGER (1981) have demonstrated the emergence of mutant viruses resistant to DI particles of the parent viruses. FRIELLE et al. (1984) have recently reported a number of phenotypic changes in the virus isolated from persistently WSN virus-infected cells. These included decreased plaque size, decreased hemagglutinating ability, appearance of temperature-sensitive mutants, etc. Although DI particles were not detected during the maintenance of the persistent influenza infection, they are crucially important at the initial establishment of persistent infection (DE and NAYAK 1980; FRIELLE et al. 1984).

Since DI particles suppress the replication of the homologous viruses, it is probable that they would produce a selective pressure for the emergence

of variant viruses which would be more resistant to DI particles (HOLLAND et al. 1982). As discussed earlier, the generation and perpetuation of DI particles is so common in influenza infection that influenza virus replication is probably constantly subjected to the selective pressure exerted by the homologous DI particles. Therefore, DI particles, along with immunological and other host and environmental factors may play an important role in the evolution of influenza viruses in nature and may aid in producing variant viruses. ROTT et al. (1983) have recently reported an interesting observation which may be pertinent to the role of DI particles in virulence. They found that nonpathogenic recombinant influenza viruses, when passaged at high multiplicity at 41° C, produced pathogenic progeny viruses. They further showed that the gene constellations of the pathogenic viruses were the same as those of the parent recombinant viruses, indicating that mutations in the parent gene(s) were responsible for the regaining of virulence. Although the mechanism of reactivation remains unknown, it is possible that multicycle high-multiplicity infections, which yield DI particles, may have aided in the selection process. Recent findings that mildly pathogenic chicken virus isolates contained DI particles but highly virulent viruses isolated later did not (Webster et al. 1984) suggest that the loss of DI particles may also be involved in the emergence of a highly virulent virus.

Since the influenza virus genome is segmented, it should be possible to discover (a) what effect a specific DI preparation has on the changes observed in different genes and (b) which gene or gene constellation of a virus is responsible for the susceptibility or resistance of a standard virus to a given DI preparation. Such experiments would determine whether a specific DI particle produces a uniform selective pressure against all genes or whether some specific genes and gene products are more vulnerable to the interfering effect of a specific DI preparation.

In summary, DI particles interfering with the growth of standard virus may exert selective pressure against the homologous virus. Influenza DI particles, which are ubiquitous, may, along with other factors, aid in the generation of variant viruses.

13 Conclusion

Research in the past 5 years has illuminated many aspects of the biology of influenza DI particles, the highlights of which are as follows:

1. Influenza DI particles contain reduced amounts of standard viral gene segments, usually with pronounced reduction in the level of one or more polymerase genes; this is the reason for their defective nature.
2. DI particles also contain novel small RNA segments, called DI RNAs, which are not required for the replication of standard viruses and which are derived from standard gene segments, predominantly from one of the polymerase genes, by an internal deletion mechanism; DI RNAs can undergo further evolution.

3. DI RNAs are specifically responsible for interference. Unlike most DI RNAs of other negative strand RNA viruses, influenza DI RNAs are of the 5′-3′ type and are capable of producing poly(A)$^+$ mRNAs. Some DI preparations produce DI-specific polypeptides (the translation products of DI transcripts) in infected cells.
4. Recent data suggest a simple competition model based on the smaller size of DI RNA for the amplification of DI RNA over standard viral RNA segments. Furthermore, packaging of DI RNP may account for the reduction in the packaging of its progenitor gene in DI virus particles and suggest the possibility of a selective rather than random packaging of eight standard segmented RNA segments during virus assembly. DI-mediated interference with influenza viruses therefore probably affects processes involved in replication and transcription of viral RNA as well as assembly of RNP complexes into viruses.
5. DI particles can be generated in any host system that allows growth of the standard virus; it is therefore probable that they are generated in natural infections as well as in laboratory experiments.
6. DI particles may serve to attenuate the pathogenic effects of the standard virus infection. They may thus prove useful as the basis of an effective influenza vaccine. By playing a role in the emergence of variant viruses, their occurrence may also be advantageous to the evolution of the standard virus.

Many questions remain unanswered. Nevertheless, since DI particles have been found in every major virus-host system that has been studied, it appears that they are not an extraneous phenomenon but the outcome of fundamental processes involved in virus replication. The ongoing study of DI particles will help shed light on these processes.

Acknowledgements. Work from our laboratory described in this review was supported by research grants from the National Institute of Allergy and Infectious Diseases (R01 AI 12749, R01 AI 16348) and the National Science Foundation (PCM 81-292). Thomas M. Chambers was a recipient of a U.S. Public Health Service National Research Service Award (1 F32 AI 06945-01). We are grateful to N. Sivasubramanian for helpful discussions.

References

Ada GL, Perry BT (1955) Infectivity and nucleic acid content of influenza virus. Nature (London) 175:209–210

Ada GL, Perry BT (1956) Influenza virus nucleic acid: relationship between biological characteristics of the virus particle and properties of the nucleic acid. J Gen Microbiol 14:623–633

Akkina RK, Chambers TM, Nayak DP (1984a) Expression of defective-interfering influenza virus-specific transcripts and polypeptides in infected cells. J Virology 51:395–403

Akkina RK, Chambers TM, Nayak DP (1984b) Mechanism of interference by defective-interfering particles of influenza virus: Differential reduction of intracellular synthesis of specific polymerase proteins. Virus Res 1:687–702

Amess LS, Pridgen CL, Kingsbury DW (1982) Sendai virus DI RNA species with conserved virus genome termini and extensive internal deletions. Virology 118:17–27

Barry RD (1961) The multiplication of influenza virus. 2. Multiplicity reactivation of ultraviolet irradiated virus. Virology 14:398–405

Bay PHS, Reichmann ME (1982) In vitro and in vivo inhibition of primary transcription of vesicular stomatitis virus by a defective interfering particle. J Virol 41:172–182
Bean WJ, Simpson RW (1976) Transcriptase activity and genome composition of defective influenza virus. J Virol 18:365–369
Bellett AJD, Cooper PD (1959) Some properties of the transmissible interfering component of vesicular stomatitis virus preparations. J Gen Microbiol 21:498–509
Birch-Andersen A, Paucker K (1959) Studies on the structure of influenza virus. 2. Ultrathin sections of infectious and non-infectious particles. Virology 8:21–40
Blumberg BM, Kolakofsky D (1983) An analytical review of defective infections of vesicular stomatitis virus. J Gen Virol 64:1839–1847
Blumenthal T, Carmichael GG (1979) RNA replication: function and structure of Qβ replicase. Ann Rev Biochem 48:525–548
Buonagurio DA, Krystal M, Palese P, DeBorde DC, Maassab HF (1984) Analysis of an influenza A virus mutant with a deletion in the NS segment. J Virol 49:418–425
Cairns HJF, Edney M (1952) Quantitative aspects of influenza virus multiplication. I. Production of incomplete virus. J Immunol 69:155–160
Carter MJ, Mahy BWJ (1982a) Incomplete avian influenza virus contains a defective noninterfering component. Arch Virol 71:12–25
Carter MJ, Mahy BWJ (1982b) Incomplete avian influenza A virus displays anomalous interference. Arch Virol 74:71–76
Carter MJ, Mahy BWJ (1982c) Synthesis of RNA segments 1–3 during generation of incomplete influenza A (fowl plague) virus. Arch Virol 73:109–119
Chambers TM, Akkina RK, Nayak DP (1984) In vivo transcription and translation of defective interfering particle – specific RNAs of influenza virus. In: Compans RW, Bishop DHL (eds) Segmented Negative Strand Viruses, Academic Orlando, pp 85–91
Chanda PK, Chambers TM, Nayak DP (1983) In vitro transcription of defective interfering particles of influenza virus produces poly(A) containing complementary RNAs. J Virol 45:55–61
Choppin PW, Pons MW (1970) The RNAs of infective and incomplete influenza virions grown in MDBK and HeLa cells. Virology 42:603–610
Chow N, Simpson RW (1971) RNA-dependent RNA polymerase activity associated with virions and subviral particles of myxoviruses. Proc Natl Acad Sci USA 68:752–756
Cole CN (1975) Defective interfering (DI) particles of poliovirus. Prog Med Virol 20:180–207
Crumpton WM, Dimmock NJ, Minor PD, Avery RJ (1978) The RNAs of defective interfering influenza virus. Virology 90:370–373
Crumpton WM, Clewley JP, Dimmock NJ, Avery RJ (1979) Origin of subgenomic RNAs in defective interfering influenza virus. FEMS Microbiol Lett 6:431–434
Crumpton WM, Avery RJ, Dimmock NJ (1981) Influence of the host cell on the genomic and subgenomic RNA content of defective interfering influenza virus. J Gen Virol 53:173–177
Davis AR, Nayak DP (1979) Sequence relationships among defective interfering influenza viral RNAs. Proc Natl Acad Sci USA 76:3092–3096
Davis AR, Hiti AL, Nayak DP (1980) Influenza defective interfering viral RNA is formed by internal deletion of genomic RNA. Proc Natl Acad Sci USA 77:215–219
De BK, Nayak DP (1980) Defective interfering influenza viruses and host cells: Establishment and maintenance of persistent influenza virus infection in MDBK and HeLa cells. J Virol 36:847–859
Doyle M, Holland JJ (1973) Prophylaxis and immunization in mice by use of virus-free defective T particles to protect against intracerebral infection by vesicular stomatitis virus. Proc Natl Acad Sci USA 70:2105–2108
Duesberg PH (1968) The RNAs of influenza virus. Proc Natl Acad Sci USA 59:930–937
Etkind PR, Buchhagen DL, Herz C, Broni BB, Krug RM (1977) The segments of influenza viral mRNA. J Virol 22:346–352
Fazekas De St Groth S, Graham DM (1954a) The production of incomplete virus particles among influenza strains. Experiments in eggs. Brit J Exp Path 35:60–74
Fazekas De St Groth S, Graham DM (1954b) Artificial production of incomplete influenza virus. Nature (London) 173:637–638
Fields S, Winter G (1982) Nucleotide sequences of influenza virus segments 1 and 3 reveal mosaic structure of a small viral RNA segment. Cell 28:303–313
Frielle DW, Huang DD, Youngner JS (1984) Persistent infection with influenza A virus: Evolution of virus mutants. Virology 138:103–107

Gamboa ET, Harter DH, Daffy PE, Hsu KC (1975) Murine influenza virus encephalomyelitis. III. Effect of defective interfering particles. Acta Neuropathol 34:157–169

Gard S, von Magnus P (1947) Studies on interference in experimental influenza. 2. Purification and centrifugation experiments. Ark Kemi Mineral Geol 24b, No. 8

Gard S, von Magnus P, Svedmyr A, Birch-Andersen A (1952) Studies on the sedimentation of influenza virus. Arch Ges Virusforsch 4:591–611

Ginsberg HS (1954) Formation of non-infectious influenza virus in mouse lungs: Its dependence upon extensive pulmonary consolidation initiated by the viral inoculum. J Exp Med 100:581–603

Hanig M, Bernkopf H (1950) The sedimentable components of influenza virus propagated in de-embryonated eggs. J Immunol 65:585–590

Hay AJ, Lomniczi B, Bellamy AR, Skehel JJ (1977) Transcription of the influenza virus genome. Virology 83:337–355

Henle W (1953) Multiplication of influenza virus in the entodermal cells of the allantois of chick embryo. Adv Virus Res 1:141–227

Hinshaw VS, Bean WJ, Webster RG, Sriram G (1980) Genetic reassortment of influenza A viruses in the intestinal tract of ducks. Virology 102:412–419

Hirst GK, Pons MW (1973) Mechanism of influenza recombination. II. Virus aggregation and its effect on plaque formation by so-called noninfective virus. Virology 56:620–631

Holland J, Grabau EA, Jones CL, Semler BL (1979) Evolution of multiple genome mutations during long-term persistent infection by vesicular stomatitis virus. Cell 16:495–504

Holland J, Kennedy SIT, Semler BL, Jones CL, Roux L, Grabau EA (1980) Defective interfering RNA viruses and host cell response. In: Fraenkel-Conrat H, Wagner RR (eds) Comprehensive Virology, Vol. 16. Plenum New York, pp 137–192

Holland J, Spindler K, Horodyski F, Grabau E, Nichol S, VandePol S (1982) Rapid evolution of RNA genomes. Science 215:1577–1585

Horsfall FL (1954) On the reproduction of influenza virus. Quantitative studies with procedures which enumerate infective and haemagglutinating virus particles. J Exp Med 100:135–161

Horsfall FL (1955) Reproduction of influenza viruses. Quantitative investigations with particle enumeration procedures on the dynamics of influenza A and B virus reproduction. J Exp Med 102:441–473

Hoyle L (1968) The influenza viruses. Virology Monographs Vol 4, Springer-Verlag, New York

Huang AS (1975) Defective interfering viruses. Ann Rev Microbiol 27:101–117

Huang AS (1977) Viral pathogenesis and molecular biology. Bacteriol Rev 41:811–821

Huang AS (1982) Significance of sequence rearrangements in rhabdovirus. ASM News 48:148–151

Huang AS, Wagner RR (1966) Defective T particles of vesicular stomatitis virus. II. Biological role in homologous interference. Virology 30:173–181

Huang AS, Baltimore D (1970) Defective viral particles and viral disease processes. Nature (London) 226:325–327

Huang AS, Baltimore D (1977) Defective interfering animal viruses. In: Fraenkel-Conrat H, Wagner RR (eds). Comprehensive Virology, Vol 10. Plenum, New York, pp 73–116

Huang AS, Little SP, Oldstone MBA, Rao D (1978) Defective interfering particles: their effect on gene expression and replication of vesicular stomatitis virus. In: Stevens JG, Todaro GH, Fox CF (eds) Persistent viruses. ICN-UCLA Symposium on Molecular and Cellular Biology, Vol VI. Academic New York, pp 399–408

Janda JM, Nayak DP (1979) Defective influenza viral ribonucleoproteins cause interference. J Virol 32:697–702

Janda JM, Davis AR, Nayak DP, De BK (1979) Diversity and generation of defective interfering influenza virus particles. Virology 95:48–58

Jennings PA, Finch JT, Winter G, Robertson JS (1983) Does the higher order structure of the influenza virus ribonucleoprotein guide sequence rearrangements in influenza viral RNA? Cell 34:619–627

Kane WP, Pietras DF, Bruenn JA (1979) Evolution of defective-interfering double stranded RNAs of yeast killer virus. J Virol 32:692–696

Kaptein J, Nayak DP (1982) Complete nucleotide sequence of the polymerase 3 (P3) gene of human influenza virus A/WSN/33. J Virol 42:55–63

Kaverin N, Kolomietz L, Rudneva I (1980) Incomplete influenza virus. Partial functional complementation as revealed by hemadsorbing cell count test. J Virol 34:506–511

Kilbourne ED, Gerhard W, Whitaker CW (1983) Monoclonal antibodies to the hemagglutinin Sa antigenic site of A/PR/8/34 influenza virus distinguish biologic mutants of swine influenza virus. Proc Natl Acad Sci USA 80:6399–6402

King AMQ, McCahon D, Slade WR, Newman JWI (1982) Recombination in RNA. Cell 29:921–928

Kingsbury DW (1974) The molecular biology of paramyxoviruses. Med Microbiol Immunol 160:73–83

Kolakofsky D (1976) Isolation and characterization of Sendai virus DI RNAs. Cell 8:547–555

Krug RM (1983) Transcription and replication of influenza viruses. In: Palese P, Kingsbury DW (eds) Genetics of Influenza Viruses,. Springer, Vienna, pp 70–98

Lazzarini RA, Keene JD, Schubert M (1981) The origin of defective interfering particles of the negative strand RNA viruses. Cell 26:145–154

Lehtovaara P, Soderlund H, Kernen S, Petterson RF, Kaarianen L (1981) 18 S defective interfering RNA of Semliki forest virus contains a triplicate linear repeat. Proc Natl Sci USA 78:5353–5357

Lenard J, Compans RW (1975) Polypeptide composition of incomplete influenza virus grown in MDBK cells. Virology 65:418–426

Leppert M, Rittenhouse L, Perrault J, Summers D, Kolakofsky D (1979) Plus and minus strand leader RNAs in negative strand virus-infected cells. Cell 18:735–747

Lief FS, Fabiyi A, Henle W (1956) The decreased incorporation of S antigen into elementary bodies of increasing incompleteness. Virology 2:782–797

Lundquist RE, Sullivan M, Maizel JV Jr (1979) Characterization of a new isolate of poliovirus defective interfering particles. Cell 18:759–769

von Magnus P (1947) Studies on interference in experimental influenza. I. Biological observations. Ark Kemi Mineral Geol 24b(7):1

von Magnus P (1951a) Propagation of the PR8 strain of influenza virus in chick embryos. I. The influence of various experimental conditions on virus multiplication. Acta Path Microbiol Scand 28:250–277

von Magnus P (1951b) Propagation of the PR8 strain of influenza virus in chick embryos. II. The formation of "incomplete" virus following the inoculation of large doses of seed virus. Acta Path Microbiol Scand 28:278–293

von Magnus P (1951c) Propagation of the PR8 strain of influenza virus in chick embryos. III. Properties of the incomplete virus produced in serial passages of undiluted virus. Acta Path Microbiol Scand 29:157–181

von Magnus P (1952) Propagation of the PR8 strain of influenza virus in chick embryos. IV. Studies on the factors involved in the formation of incomplete virus upon serial passage of undiluted virus. Acta Path Microbiol Scand 30:311–335

von Magnus P (1954) Incomplete forms of influenza virus. Adv Virus Res 2:59–78

von Magnus P (1965) The in ovo production of incomplete virus by B/Lee and A/PR8 influenza viruses. Arch Ges Virusforsch 17:414–423

Manire GP (1957) Studies on the toxicity for mice of incomplete influenza virus. Acta Path Microbiol Scand 40:501–510

McCauley JW, Mahy BWJ (1983) Structure and function of the influenza virus genome. Biochem J 211:281–294

McKee AP (1951) Non-toxic influenza virus. J Immunol 66:151–167

Mills DR, Peterson RI, Spiegelman S (1967) An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. Proc Natl Acad Sci USA 58:217–224

Monroe SS, Ou G-S, Rice CM, Schlesinger S, Strauss E, Strauss JH (1982) Sequence analysis of cDNAs derived from the RNA of Sindbis virions and of defective interfering particles. J Virol 41:153–162

Monroe SS, Schlesinger S (1983) RNAs of two independently isolated defective interfering particles of Sindbis virus contains a cellular tRNA sequence at their 5′ ends. Proc Natl Acad Sci USA 80:3279–3283

Moore DH, Davies MC, Levine S, Englert ME (1962) Correlation of structure with infectivity of influenza virus. Virology 17:470–479

Morgan C, Hsu KC, Rose HM (1962) Structure and development of viruses as observed in the electron microscope. VII. Incomplete influenza virus. J Exp Med 116:553–564

Moss BA, Brownlee GG (1981) Sequence of DNA complementary to a small RNA segment of influenza virus A/NT/60/68. Nucleic Acids Res 9:1941–1947

Nakajima K, Ueda M, Sugiura A (1979) Origin of small RNA in von Magnus particles of influenza virus. J Virol 29:1142–1148

Nayak DP (1969) Influenza viruses: Structure, replication and defectiveness. Fed Proc 28:1858–1865

Nayak DP (1972) Defective virus RNA synthesis and production of incomplete influenza virus in chick embryo cells. J Gen Virol 14:63–67

Nayak DP (1980) Defective interfering influenza viruses. Ann Rev Microbiol 34:619–644

Nayak DP, Baluda MA (1967) Isolation and partial characterization of nucleic acid of influenza virus. J Virol 1:1217–1223

Nayak DP, Sivasubramanian N (1983) The structure of the influenza defective interfering (DI) RNAs and their progenitor genes. In: Palese P, Kingsbury DW (eds) Genetics of Influenza Viruses. Springer-Verlag Vienna, pp 255–279

Nayak DP, Tobita K, Janda JM, Davis AR, De BK (1978) Homologous interference mediated by defective interfering influenza virus derived from a temperature-sensitive mutant of influenza virus. J Virol 28:375–386

Nayak DP, Davis AR, Cortini R (1982a) Defective interfering influenza viruses: complete sequence analysis of a DI RNA. In: Nayak DP (ed) Genetic variation Among Influenza Viruses, Academic New York, pp 77–92

Nayak DP, Sivasubramanian N, Davis AR, Cortini R, Sung J (1982b) Complete sequence analyses show that two defective interfering influenza viral RNAs contain a single internal deletion of a polymerase gene. Proc Natl Acad Sci USA 79:2216–2220

Palese P (1977) The genes of influenza virus. Cell 10:1–10

Palese P, Kingsbury DW, eds (1983) Genetics of Influenza Viruses. Springer Vienna

Paucker K, Henle W (1955) Studies on host-virus interaction in the chick embryo influenza virus system. XII. Further analysis of yields derived from heat-inactivated standard seeds. J Exp Med 101:493–506

Paucker K, Birch-Andersen A, von Magnus P (1959) Studies on the structure of influenza virus. I. Components of infectious and incomplete particles. Virology 8:1–20

Penn CR, Mahy BWJ (1984) Expression of influenza virus subgenomic virion RNAs in infected cells. In: Compans RW, Bishop DHL (eds) Segmented Negative Strand Viruses, Academic Orlando, pp 173–178

Perrault J (1981) Origin and replication of defective interfering particles. Current Topics Microbiol Immunol 93:151–207

Perrault J, Semler BW, Leavitt RW, Holland JJ (1978) Inverted complementary terminal sequences in defective interfering particle RNAs of vesicular stomatitis virus and their possible role in autointerference. In: Mahy BWJ, Barry RD (eds) Negative Strand Viruses and the Host Cell. Academic New York, pp 527–538

Perrault J, Clinton GM, McClure MA (1983) RNP template of vesicular stomatitis virus regulates transcription and replication functions. Cell 35:175–185

Peters D (1959) Morphology of viruses pathogenic for man and animals (German). Zbl. Bakt I Abt Orig 176:259–294

Pettersson RF (1981) 5′-terminal nucleotide sequence of Semliki forest virus 18S defective interfering RNA is heterogeneous and different from the genomic 42S RNA. Proc Natl Acad Sci USA 78:115–119

Pons M, Hirst GK (1969) The single and double-stranded RNAs and the proteins of incomplete influenza virus. Virology 38:68–72

Pons MW (1980) The genome of incomplete influenza virus. Virology 100:43–52

Rabinowitz SG, Huprikar J (1979) Influence of defective-interfering particles of the PR8 strain of influenza A virus on the pathogenesis of pulmonary infection in mice. J Infect Dis 140:305–315

Rao DD, Huang AS (1982) Interference among defective interfering particles of vesicular stomatitis virus. J Virol 41:210–221

Robertson JS, Schubert M, Lazzarini RA (1981) Polyadenylation sites for influenza virus mRNA. J Virol 38:157–163

Rott R, Schäfer W (1960) Untersuchungen über die hämagglutinierenden-nichtinfektiösen Teilchen der Influenza-Viren. I. Die Erzeugung von „inkompletten Formen" beim Virus der klassischen Geflügelpest (v. Magnus-Phänomen). Z Naturforsch 150:691–693

Rott R, Schäfer W (1961) Untersuchungen über die hämagglutinierenden nichtinfektiosen Teilchen der Influenza-Viren (German). Z Naturforsch 16b:310–321

Rott R, Scholtissek C (1963) Investigations about the formation of incomplete forms of fowl plague virus. J Gen Microbiol 33:303–312

Rott R, Orlich M, Scholtissek C (1983) Pathogenicity reactivation of nonpathogenic influenza virus recombinants under von Magnus conditions. Virology 126:459–465

Rowlands D, Grabau E, Spindler K, Jones C, Semler B, Holland J (1980) Virus protein changes and RNA termini alterations evolving during persistent infection. Cell 19:871–880

Schafer W (1955) Sero-immunologic studies on incomplete forms of the virus of classical fowl plague (German). Arch Exp Vet Med 9:218–230

Scholtissek C, Drzeniek R, Rott R (1966) Stepwise inactivation of an influenza virus by serial undiluted passages. Virology 30:313–318

Seto JT (1964) Sialidase (neuraminidase) activity of standard and incomplete virus. Biochem Biophys Acta 90:420–422

Shiota O (1956) Studies on the incomplete form of influenza virus. II. The morphological characteristics of so-called incomplete particles of the virus (Japanese). Virus 6:193–206

Sivasubramanian N, Nayak DP (1983) Defective interfering influenza RNAs of polymerase 3 gene contain single as well as multiple internal deletions. Virology 124:232–237

Smith GL, Hay AJ (1982) Replication of the influenza virus genome. Virology 118:96–108

Ueda M, Nakajima K, Suguira A (1980) Extra RNAs of von Magnus particles of influenza virus cause reduction of particular polymerase genes. J Virol 34:1–8

Uhler M, Gard S (1954) Lipid content of standard and incomplete influenza A virus. Nature (London) 173:1041–1042

Waterson AP, Rott R, Schäfer W (1961) The structure of fowl plague virus and virus N. Z Naturforsch 16b:154–156

Webster R, Kawaoka Y, Bean WJ, Naeve CW, Wood JM (1984) Characterization of virulent and avirulent A/chick/Pennsylvania/83 (H_5N_2). An International Influenza Virus Hemagglutinin Meeting, Green Pia Miki, Japan

Weiss B, Schlesinger S (1981) Defective interfering particles of Sindbis virus do not interfere with homologous virus obtained from persistently infected BHK cells but do interfere with Semliki Forest virus. J Virol 37:840–844

Welsh RM, Lampert PW, Oldstone MBA (1977) Prevention of virus-induced cerebellar disease by defective interfering lymphocytic choriomeningitis virus. J Infect Dis 136:391–399

Werner GH (1956) Quantitive studies on influenza virus infection of the chick embryo by the amniotic route. J Bact 71:505–515

Werner GH, Schlesinger RW (1954) Morphological and quantitative comparison between infectious and non-infectious forms of the influenza virus. J Exp Med 100:203–215

Winter G, Fields S, Ratti G (1981) The structure of the two subgenomic RNAs from human influenza virus A/PR/8/34. Nucleic Acids Res 9:6907–6915

Yang F, Lazzarini RA (1983) Analysis of the recombination event generating a vesicular stomatitis virus deletion defective interfering particle. J Virol 45:766–772

Yoshishita T, Kawai K, Fukai K, Ito R (1959) Analysis of infective particles in incomplete virus preparations of the von Magnus type. Biken J 2:25

Biochemistry of Arenaviruses

R.W. COMPANS and D.H.L. BISHOP

1 Introduction

Although the isolation of lymphocytic choriomeningitis virus (LCMV), the prototype member of the arenavirus group, was initially reported by ARMSTRONG and LILLIE in 1934, the virus remained unclassified for many years because of the lack of information on its biochemical properties and morphology. DALTON and co-workers (1968) first described the ultrastructure of the LCM virion; virus particles were found to be enveloped and varied greatly in size, ranging from 50 to over 200 nm in diameter. The most distinctive ultrastructural feature was the presence of one to eight or more electron-dense granules, which have subsequently been demonstrated to be host cell-derived ribosomes (see below). Following this report, MURPHY and co-workers (1969) demonstrated striking

Department of Microbiology, University of Alabama in Birmingham, Birmingham, AL 35294, USA

Current Topics in Microbiology and Immunology, Vol. 114

Table 1. The arenaviridae

Virus	Distribution	Principal vertebrate host	References
Old world species:			
LCM	Worldwide	*Mus musculus*	ARMSTRONG and LILLIE (1934)
Lassa	West Africa	*Mastomys natalensis*	BUCKLEY and CASALS (1970) FRAME et al. (1970)
Mobala	Central Africa Republic	*Praomys jacksonii*	GONSALEZ et al. (1983)
Mopeia	Mozambique	*Mastomys natalensis*	WULFF et al. (1977)
New world species: Tacaribe complex			
Amapari	Brazil	*Oryzomys goeldi* *Neacomys guianae*	PINHEIRO et al. (1966)
Flexal	Brazil	*Oryzomys* species	PINHEIRO et al. (1977)
Junin	Argentina	*Calomys laucha* *Calomys musculinus* *Akodon azarae*	PARODI et al. (1958)
Latino	Bolivia	*Calomys callosus*	JOHNSON et al. (1973)
Machupo	Bolovia	*Calomys callosus*	JOHNSON et al. (1966)
Parana	Paraguay	*Oryzomys buccinatus*	WEBB et al. (1970)
Pichinde	Colombia	*Oryzomys albigularis*	TRAPIDO and SANMARTIN (1971)
Tacaribe	Trinidad	*Antibesis literates* *Antibesis jamaicensis*	DOWNS et al. (1963)
Tamiami	Florida	*Sigmodon hispideis*	CALISHER et al. (1970)

morphological similarities between LCMV and Machupo and Tacaribe viruses, two members of a group of antigenically related viruses termed "the Tacaribe complex." These investigators proposed that the viruses of the Tacaribe complex and LCMV be placed in a new taxonomic group based on their morphological similarity. A serological relationship of LCMV to the members of the Tacaribe complex was also demonstrated using immunofluorescence (ROWE et al. 1970b). The name "arenovirus" (from *arenosus,* L., "sandy") was initially proposed to reflect the characteristic granules seen in the virions in ultrathin sections (ROWE et al. 1970a). It was noted that the virions were morphologically distinct from other RNA-containing enveloped viruses because of the lack of a well-defined helical nucleocapsid and the presence of the electron-dense internal granules. The designation "arenavirus" was subsequently accepted for the taxonomic group (PFAU et al. 1974) and the viruses established as the family Arenaviridae (MATTHEWS 1982). The members of the group which have now been identified are listed in Table 1.

Extensive information is available on the biological properties of arenaviruses and the human diseases caused by these agents. This has been reviewed elsewhere (LEHMANN-GRUBE 1971; CASALS 1975; RAWLS and LEUNG 1979; PEDERSEN 1979; HOWARD and SIMPSON 1980; BUCHMEIER et al. 1980; LEHMANN-GRUBE et al. 1983) and will not be described here. Important features of arenavirus biology include the following: (a) Several members of the group cause severe and sometimes fatal diseases in man – including lymphocytic choriomeningitis, Argentine and Bolivian hemorrhagic fevers, and Lassa fever. (b) Trans-

mission of arenaviruses to man occurs through exposure to virus carried by a rodent, the host in which the viruses characteristically produce persistent infections. Persistently infected rodents continue to excrete virus in urine and saliva throughout life. Persistent infection by arenaviruses in cell culture is also readily established. (c) There are significant immunopathological events in arenavirus infections. In acute LCM virus infections of mice, the disease symptoms are a consequence of the immune response, rather than a direct effect of the virus on host tissues. Certain mouse strains persistently infected with LCM virus develop a late-onset disease involving the deposition of antigen-antibody complexes, leading to chronic glomerulonephritis (OLDSTONE and DIXON 1967, 1969; BUCHMEIER and OLDSTONE 1978b).

2 Virion Structure

2.1 Morphology

In negatively stained preparations, the surfaces of arenavirus particles were found to be covered with distinct spikes, which appear to vary in spacing on the envelope (MURPHY et al. 1970). Some of the surface projections appeared triangular when observed end-on; however, this was not uniformly observed. At high magnification, the surface spikes on Pichinde virions appeared to be club shaped (VEZZA et al. 1977). Virus particles are generally spherical, although some pleomorphism is evident, particularly in the larger particles in virus preparations. No filamentous virions of the type seen with orthomyxoviruses and paramyxoviruses have been observed. Internal structural details have been difficult to resolve; after disruption of the viral envelope, strand-like nucleoprotein complexes are observed (see below).

When examined in thin sections, virus particles were found to range from 60 to 280 nm in diameter and to consist of an electron-dense unit membrane upon which the surface projections can be resolved. In the interior, one or more electron-dense granules 20–25 nm in diameter were observed in most particles, although occasional particles appear to lack such structures.

Examples of virus particles visualized by negative staining or thin-section electron microscopy are shown in Fig. 1. A schematic of an arenavirus particle is provided in Fig. 2.

2.2 Composition

2.2.1 Particle Mass and Composition

No direct measurements have been made of virus particle mass or chemical composition; however, a rough estimate may be made based on the finding of an RNP:protein ratio of 1:31.6 for a Pichinde virus preparation containing no detectable ribosomes (VEZZA et al. 1977) and an assumption of a protein

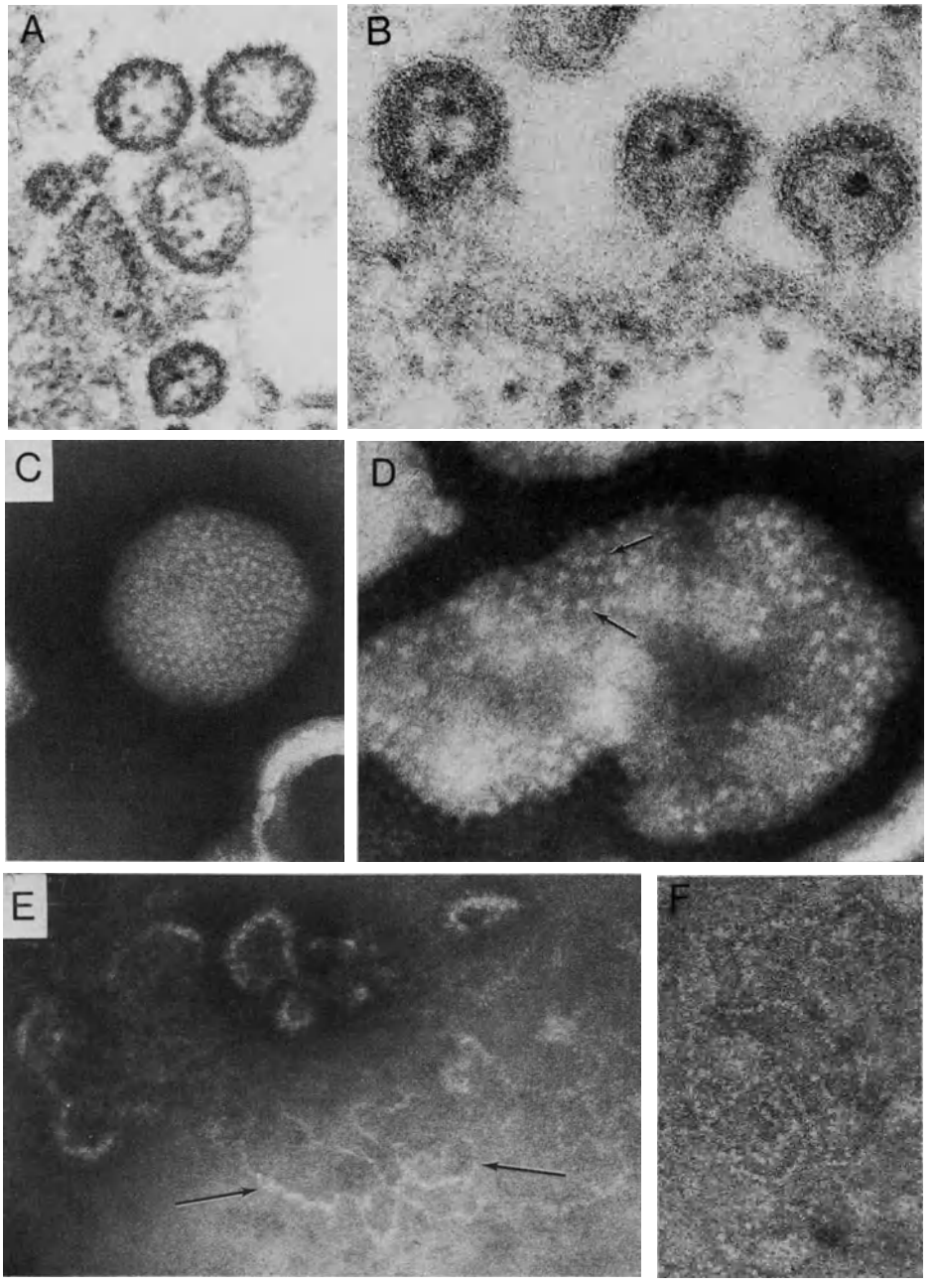


Fig. 1 A–F. Electron micrographs of arenavirus particles and subviral components. **A** Thin section of Pichinde virus particles released from infected BHK21 cells. Many ribosomes are evident in the interior. **B** Tacaribe virions, containing one or two ribosomes, in the process of budding from an infected Vero cell. A diffuse layer of electron-dense material, which may represent the viral nucleoprotein, is localized beneath the envelope. **C** A Tacaribe virion covered with surface spikes. **D** Higher magnification of a pleomorphic Tacaribe virion; some of the surface spikes (*arrows*) appear triangular in shape. **E** Spontaneously disrupted Tacaribe virions, revealing fragments of the viral envelope and strand-like components believed to be the internal ribonucleoproteins (*arrows*). **F** A preparation of Tamiami ribonucleoproteins purified in a CsCl equilibrium gradient, showing beaded strand-like components. **C–F** Negative staining with potassium phosphotungstate. (**C, E,** and **F** are from GARD et al. 1977.) **A,** ×100000; **B,** ×180000; **C,** ×180000; **D,** ×340000; **E,** ×180000; **F,** ×220000

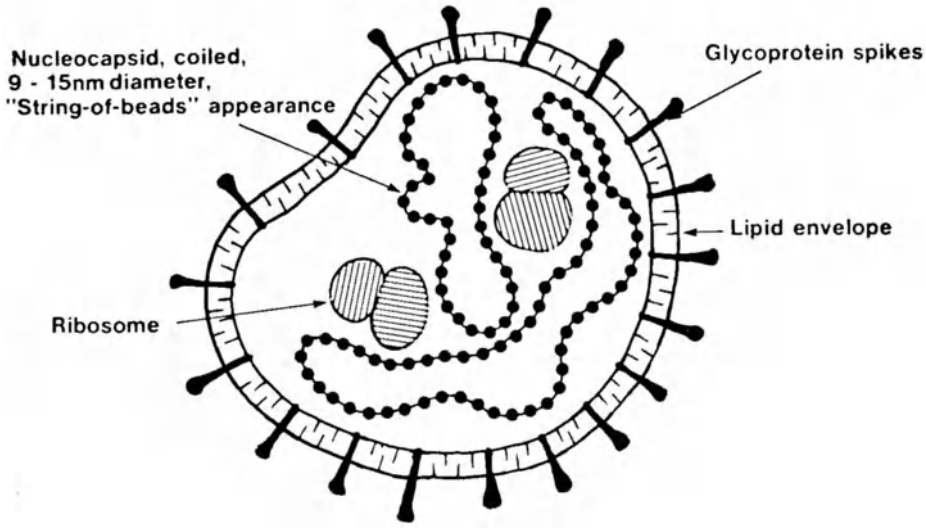


Fig. 2. Schematic of an arenavirus particle

content of about 70%, which corresponds to that observed for several other lipid-containing RNA viruses (COMPANS and KLENK 1979). These values would indicate an RNA content of 2.2% in arenavirus particles. Using the value of 3.6×10^6 for the total mol. wt. of the RNA in one complete viral genome (see below), a minimum particle weight of 165×10^6 is obtained for a virion containing one genome and no ribosomes. Obviously, particle weights would be significantly larger for the large, pleomorphic particles found in virus preparations.

2.2.2 RNA

Arenaviruses are unique in that host ribosomes and small RNA species (4–5.5 *S*) are frequently incorporated in mature virus particles, in addition to the viral genomic RNA species. The presence of 28 *S* and 18 *S* ribosomal RNA species in viral extracts was initially demonstrated by PEDERSEN (1971) in studies with LCM virus, and subsequently confirmed and extended to Junin, Parana, Pichinde, Tacaribe, and Tamiami arenaviruses (PEDERSEN 1973; CARTER et al. 1973a; ANON et al. 1976; VEZZA et al. 1977, 1978; DUTKO et al. 1976, 1981; GIMENEZ and COMPANS 1980; RAMSINGH et al. 1980; CHINAULT et al. 1981; BUCHMEIER et al. 1980). Using actinomycin D at doses that inhibited the synthesis of host ribosomal RNA, it was shown that the incorporation of radioactive precursors into viral 28 *S* and 18 *S* RNA species was selectively inhibited (PEDERSEN 1971; CARTER et al. 1973b). It has been reported that some Pichinde virus stocks minimally incorporate preformed or newly synthesized ribosomes, although the results depend on the cell substrates employed (Fig. 3; VEZZA et al. 1978).

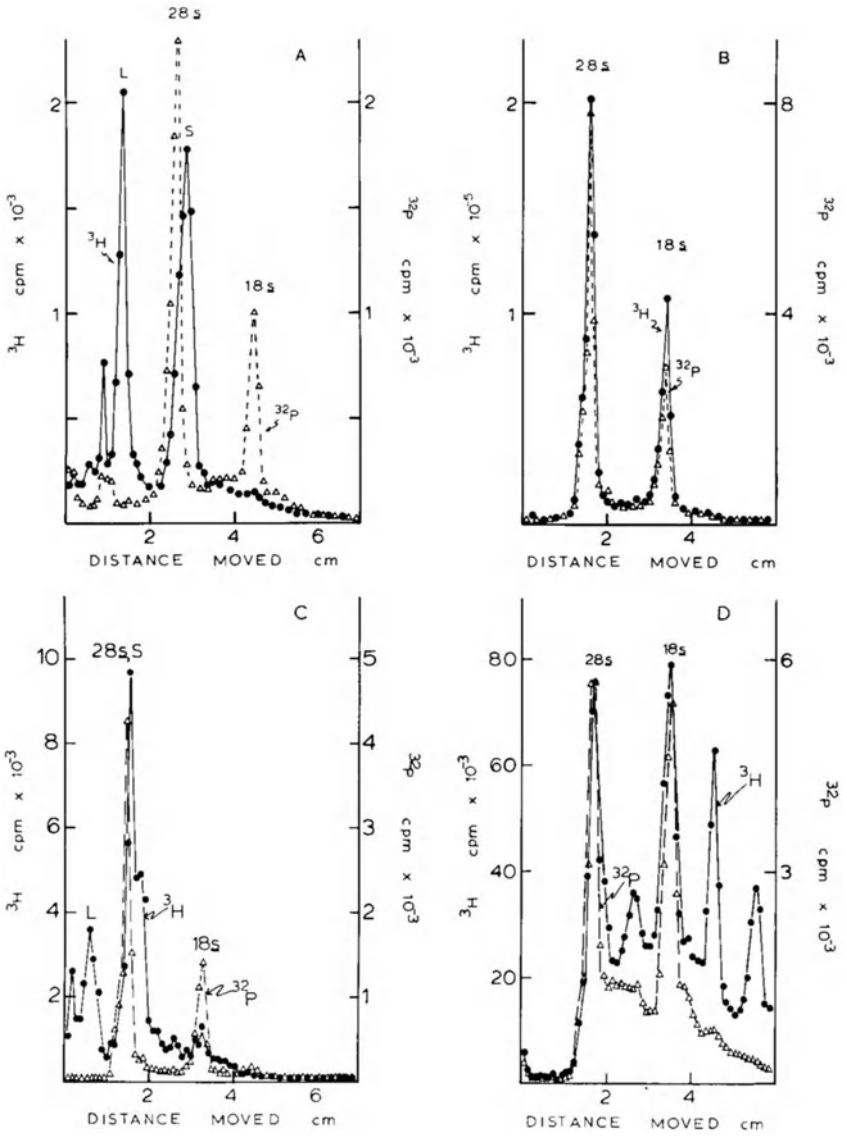


Fig. 3A–D. RNA species of Pichinde virus preparations. In **A** a preparation of (^{3}H)uridine-labeled Pichinde viral RNA was melted and then mixed with ^{32}P-labeled uninfected BHK21 cell ribosomal 28*S* and 18*S* RNA, and resolved by polyacrylamide gel electrophoresis (VEZZA et al. 1977). In **B** (^{3}H)uridine infected cell and ^{32}P-labeled uninfected BHK21 cell ribosomal RNA was coelectrophoresed. **C** and **D** Profiles of (^{3}H)uridine-labeled Tacaribe viral and ^{32}P-BHK21 cellular RNA for early (**C,** 2 days) and late (**D,** 10 days) virus harvests

The genomic information of arenaviruses is contained in two RNA species designated large (L) and small (S). Estimation of the sizes of these RNA species vary depending on the method of analysis employed and on the viral serotype. Sizes on the order of $2.5–3.6 \times 10^{6}$ daltons (L) and $1.1–1.6 \times 10^{6}$ daltons (S)

have been reported (ANON et al. 1976; CARTER et al. 1973a; DUTKO et al. 1976, 1981; KILEY et al. 1981; PEDERSEN 1970, 1971, 1973, 1979; RAMSINGH et al. 1980; VEZZA et al. 1978). DNA cloning analyses of the Pichinde S RNA species have established that it has a size of 1.1×10^6 daltons (AUPERIN et al. 1984c).

Additional RNA species have been demonstrated in virus preparations, including a 15*S* species (FARBER and RAWLS 1975; DUTKO et al. 1976; DIMOCK et al. 1982). Virus recovered from persistently infected cells, or late in an infection course, also yield additional RNA species (Fig. 3). The relationship of such RNA species to the viral L and S RNA species, or to defective virus particles, has not yet been elucidated.

Analyses of the end sequences of LCM, Pichinde, and Tacaribe viruses have revealed consensus conserved 3′ end sequences for the L and S RNA species of the three viruses (AUPERIN et al. 1981, 1982, 1984a, b). The 3′ends are not polyadenylated. The viral L and S RNA species have neither a 5′cap structure, nor methylated bases; they do not program the synthesis of protein in vitro (CARTER et al. 1973a; LEUNG et al. 1977; LEUNG 1978). Evidence has been obtained that the 3′ half of the viral S RNA species is complementary in composition to one of the viral induced, subgenomic, nonpolyadenylated, intracellular messenger RNA species, the N mRNA (LEUNG et al. 1977; D. AUPERIN, V. ROMANOWSKI, M. GALINSKI, and D.H.L. BISHOP, to be published), but in its 5′ half it is identical to a viral induced, subgenomic GPC mRNA species. Thus, although partially negative stranded, the term "ambisense RNA" has been coined to describe the plus and minus sense coding strategy of the S RNA (D. AUPERIN, V. ROMANOWSKI, M. GALINSKI, and D.H.L. BISHOP, to be published).

Sequence analyses have shown that the L RNA of arenaviruses codes for a gene product in the 3′ viral complementary sequence of the RNA. Whether nonstructural proteins are also coded by the L or S RNA species has not been determined.

The fact that the L and S RNA species encode distinct genetic information has also been demonstrated by oligonucleotide fingerprinting, hybridization, and genetic analyses (BISHOP 1980; BISHOP et al. 1980; AUPERIN et al. 1981, 1982, 1984a, b; VEZZA et al. 1978; LEUNG et al. 1981; VEZZA and BISHOP 1977; VEZZA et al. 1980; KIRK et al. 1980; GIMENEZ and COMPANS 1981; ROMANOWSKI and BISHOP 1983).

The molar ratios of the RNA species recovered from virus preparations are often far from equal. This is true for L compared both with S and with the ribosomal 28*S* and 18*S* species. The genetic implication of such results is either that the population of viruses includes particles lacking certain RNA species (or ribosomes), or that some viruses are multiploid. Recent studies have established that genetically diploid arenaviruses can be produced in virus infections (ROMANOWSKI and BISHOP 1983). Using S RNA temperature-sensitive (*ts*) mutants of two different strains of LCM virus (distinguishable by their oligonucleotide fingerprints), plaque-cloned, wild-type progeny were obtained possessing both types of S RNA species (in addition to L, 28*S*, and 18*S* RNA). On clonal analyses, such diploid viruses segregated *ts* progeny, but could be passaged at nonpermissive temperatures without loss of the wild-type phenotype. It is

Table 2. Structural polypeptides of arenaviruses[a]

	Virus						
	LCMV	Pichinde	Tacaribe	Tamiami	Junin	Machupo	Lassa
Nucleocapsid (N) Protein	63000	64000–72000	68000	66000	60000–64000	68000	72000
Glycoproteins	54000 35000	64000–72000 34000–38000	38000–42000	44000	35000–39000	50000 41000	52000 39000
Other polypeptides		200000	79000	77000			

[a] The values listed are estimated molecular weights for the major polypeptides that have been reported; a number of other minor polypeptides have also been identified in some virus preparations, but it is uncertain that they are virus-coded. The data were obtained from the following sources: LCMV, BUCHMEIER et al. 1978; Pichinde, RAMOS et al. 1972; VEZZA et al. 1977; HARNISH et al. 1981; Tacaribe and Tamiami, GARD et al. 1977; Junin, MARTINEZ SEGOVIA and DE MITRI 1977; GRAU et al. 1981; Machupo, GANGEMI et al. 1978; Lassa, KILEY et al. 1981)

possible that the presence of multiple copies of genetic information may be related to the process of viral morphogenesis and the synthesis of pleomorphic virus particles.

2.2.3 Proteins

The polypeptide composition of a number of arenaviruses has been investigated, and the major species identified are summarized in Table 2. The most abundant polypeptide, comprising an estimated 56%–66% of the total proteins has been designated N or NP (Fig. 4, $60–72 \times 10^3$ daltons) and is the major structural protein of the nucleocapsid. Approximately 1500 molecules of N have been estimated to be present in Pichinde virions (VEZZA et al. 1977). The N polypeptide is firmly associated with the viral RNA and can be purified in association with RNA by separation of nucleocapsids from detergent–disrupted virus, using equilibrium centrifugation in CsCl, or phase separation with polyethylene glycol-dextran (VEZZA et al. 1977; GARD et al. 1977).

For LCM and Pichinde viruses, two glycosylated proteins designated GP-1 and GP-2 are present in virions (Fig. 4, $35–72 \times 10^3$ daltons); they have distinct peptide maps, indicating that they have different amino acid sequences (BUCHMEIER and OLDSTONE 1979; HARNISH et al. 1981 a, b), and can be removed from virions by proteolytic digestion yielding spikeless particles (Fig. 4; VEZZA et al. 1977). No evidence for a disulfide linkage between the two polypeptides has been obtained. For Tacaribe, Tamiami, and Junin virions, only a single major glycoprotein size class has been identified, although the Junin virus was found to contain a minor glycoprotein with a mol. wt. of 44×10^3 (GARD et al. 1977; GRAU et al. 1981). Two glycosylated polypeptides were resolved in Machupo and Lassa virions (KILEY et al. 1981). In the case of Pichinde virus, the two glycoproteins were estimated to be present in virions in approximately equal

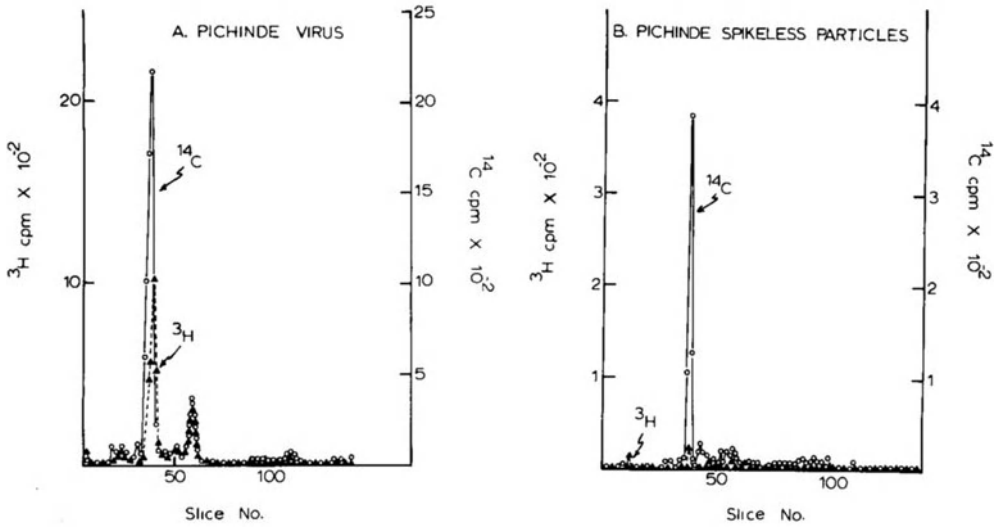


Fig. 4A, B. Characterization of Pichinde virus polypeptides. In **A** a preparation of Pichinde virus labeled by (^{3}H)glucosamine and ^{14}C-amino acids was dissociated by SDS, and the distribution of radioactivity was determined after resolution by polyacrylamide gel electrophoresis. A sample of the (^{3}H)glucosamine and ^{14}C-amino acid-labeled Pichinde was digested with Pronase and bromelain (VEZZA et al. 1977) and repurified by centrifugation in a 20%–70% sucrose gradient. After SDS lysis of rebanded protease-treated material, the distribution of radioactivity after polyacrylamide gel electrophoresis was determined in panel **B**

molar amounts, with 390 molecules of GP-1 and 440 molecules of GP-2 in a typical virion (VEZZA et al. 1977). Some information on the functions of the LCM viral glycoproteins has been obtained using monoclonal antibodies (BUCHMEIER et al. 1981). Antibodies to GP-1 had virus-neutralizing activity comparable to that of guinea pig antibody to LCMV, while a single antibody to GP-2 neutralized poorly. No neutralization was observed with antibody to the NP protein.

The glycoprotein of Tacaribe virions was separated into multiple components by two-dimensional electrophoresis (BOERSMA et al. 1982). The most acidic glycoprotein species were found to differ from less acidic species primarily in their sialic acid content; after neuraminidase treatment, similar isoelectric points were observed, and no significant differences were detected in peptide profiles obtained by limited proteolysis. Thus, only a single glycosylated polypeptide appears to be present in this virus.

Other polypeptides reported in arenaviruses include a large polypeptide of about 200×10^3 daltons, designated L protein, in Pichinde virus (HARNISH et al. 1981). This polypeptide was found in immune precipitates of infected (but not of mock-infected) cells, using antisera from infected hamsters, and was found to be distinct from other virion proteins using peptide mapping. A polypeptide of about 72×10^3 daltons, designated P, has been observed in Tacaribe and Tamiami virions in association with the ribonucleoprotein complex (GARD et al. 1977). It is probably related in sequence to the N protein, as similar peptide patterns were observed following limited proteolytic digestion (COMPANS et al.

1981). Two minor polypeptides of 79×10^3 and 105×10^3 daltons designated p 79 and p 105) have also been found in Tacaribe virions and virus-infected cells (GIMENEZ et al. 1983; see Sect. 3.4). Variable numbers of low mol.wt. polypeptides have been observed in some arenavirus preparations, such as one of 12×10^3 daltons reported in Pichinde virus (RAMOS et al. 1972; VEZZA et al. 1977). Further data, including those from sequence analysis and in vitro translation studies, are needed to define the proteins for each of the arenaviruses.

Antigenic cross-reactivity between specific polypeptides of different arenaviruses has been demonstrated with the use of monoclonal antibodies (BUCHMEIER and OLDSTONE 1978a; BUCHMEIER et al. 1981). One monoclonal antibody to the N polypeptide of Pichinde virus reacted with various members of the Tacaribe complex as well as with LCMV and Lassa viruses, two old-world arenaviruses. Other monoclonals showed cross-reactivity with one or more of the Tacaribe complex viruses. Cross-reactivity was also detected between LCMV and Mopeia virus using a monoclonal antibody to the G2 glycoprotein. Five monoclonal antibodies to the NP polypeptide of LCMV were reactive with both Lassa and Mopeia viruses.

2.2.4 Lipids

No direct analysis of the lipids of arenavirus particles has been reported. The presence of a lipid-containing membrane can be inferred from the process of virus assembly involving budding at the cell surface (see below) and from the sensitivity of virus infectivity to detergents and lipid solvents (PFAU et al. 1974).

2.2.5 Carbohydrates

The presence of carbohydrate components covalently linked to arenavirus proteins has been demonstrated by metabolic labeling with radiolabeled sugar precursors. The oligosaccharide chains linked to Tacaribe viral glycoproteins have been further characterized by gel filtration following extensive digestion of the polypeptide chains with pronase (BOERSMA et al. 1982). The elution profile of ^{3}H-glucosamine-labeled glycopeptides from the Tacaribe G protein revealed a heterogeneous major peak, with an approximate size of 3.3×10^3 daltons. After treatment with neuraminidase, the major peak was shifted to a position corresponding to a mol.wt. of about 2.5×10^3. Analysis of glycopeptides labeled with ^{3}H-mannose revealed two major glycopeptide size classes. The larger species was approximately 3.3×10^3 daltons, corresponding to the glycopeptides observed using glucosamine label, whereas the smaller species ($1.6–1.9 \times 10^3$ daltons) was resolved as a distinct peak only when mannose label was used, indicating that these are mannose-rich species. The low-molecular-weight, ^{3}H-mannose-labeled glycopeptides were found to be sensitive to endoglycosidase H, whereas the larger size class of glycopeptides were resistant to this treatment. Under conditions in which significant amounts of sulfate were found to be incorporated into the glycoproteins of influenza virus or VSV, no sulfate was detected in the glycoproteins of Pichinde virus (VEZZA et al. 1977).

2.2.6 Virus-associated Ribosomes

It has been established that the viral ribosomes are comparable in every respect to their cellular counterparts. Ribosomal 80*S* monosomes and 60*S* and 40*S* subunits have been isolated from disrupted virions and shown to have properties similar to those of cell ribosomes. They are sensitive to EDTA and contain 28*S* and 18*S* ribosomal RNA species as judged using sedimentation analyses, oligonucleotide fingerprinting, and by the presence of methylated bases (FARBER and RAWLS 1971; PEDERSEN and KONIGSHOFER 1976; VEZZA et al. 1978). The ribosomes have been shown to be capable of directing the synthesis of protein in vitro, provided they are supplied with mRNA and the appropriate factors to promote translation (CHINAULT et al. 1981).

The reason for the characteristic presence of ribosomes in virus particles is not known. No evidence has been presented that they are required for either viral morphogenesis or the establishment of a productive infection. It is generally held that ribosomes are present in arenaviruses fortuitously, presumably being incorporated at the site of virus budding simply because of the lack of a specific mechanism to exclude them. Presumably, the synthesis of pleomorphic virus particles is conducive to ribosome incorporation. The finding of infectious virus particles lacking detectable ribosomes (VEZZA et al. 1978) and results of studies employing cells with *ts* ribosomes (LEUNG and RAWLS 1977) argue against a necessary role for them. LEUNG and RAWLS (1977) using an embryonic hamster lung cell line with a 60*S* ribosomal *ts* lesion, showed that virus propagated twice in these cells and then in wild-type cells grew essentially equally well at either high or low temperatures.

2.2.7 Viral Enzymes

RNA polymerase activities associated with arenavirus particles are discussed in Sect. 3.3. In addition to these enzymes, a virus-associated cyclic AMP-independent protein kinase activity has been identified in purified preparations of LCMV (HOWARD and BUCHMEIER 1983). The enzyme activity was substantially enhanced by detergent treatment of the virus and was associated with nucleocapsids purified in a Renografin gradient. The major phosphorylated species was found to be the N polypeptide, in which both serine and threonine residues were found to be phosphorylated. Neither the protein responsible for this activity nor its origin have been determined.

2.3 Arrangement of Components

2.3.1 Envelope Structure

The glycoproteins of arenaviruses have been shown to be located on the external surface of the virion, forming the spikelike projections seen with electron microscopy. Treatment of Pichinde, Tacaribe, or Tamiami virions with proteolytic

enzymes yielded spikeless particles which lacked the glycoproteins and contained the internal, nonglycosylated polypeptides (VEZZA et al. 1977; GARD et al. 1977). By analogy with other enveloped viruses, it is likely that the glycoproteins extend through the lipid bilayer and possess hydrophobic segments embedded in the viral membrane; however, no evidence for such structural features has yet been obtained. It is also not known how many polypeptide chains comprise each of the morphological spikes on arenavirus particles. However, the finding that some monoclonal antibodies precipitate both GP-1 and GP-2 of LCMV (BUCHMEIER et al. 1981) suggests that these polypeptides may be part of the same spike structure. The fact that monoclonal antibodies against GP-1 exhibit the strongest neutralizing activity (BUCHMEIER et al. 1981) might indicate that this polypeptide comprises the distal portion of the spike structure.

The lipids of arenavirus particles appear to be organized in a bilayer structure. Analysis of Tacaribe virions by electron spin resonance measurements using spin-labeled fatty acid probes revealed a "fluidity gradient" in the lipid layer, as expected for a bilayer structure (F. LANDSBERGER, G. GARD, and R.W. COMPANS, unpublished).

2.3.2 Nucleocapsid Structure

Viral ribonucleoproteins released from spontaneously disrupted Tacaribe or Tamiami virus particles were found to appear as convoluted strands 3–4 nm in diameter, without an obvious helical symmetry (GARD et al. 1977). Similar strands of somewhat larger diameter were found released from disrupted Pichinde virus (VEZZA et al. 1977). The Tamiami ribonucleoproteins, after purification in a CsCl equilization gradient, had a beaded appearance, suggesting that the strands contained a series of globular subunits spaced with a periodicity of about 4.5 nm. The major polypeptide in the ribonucleoprotein complexes was the N protein. The viral RNA in the nucleoprotein complexes was sensitive to digestion with ribonuclease, indicating a loosely packed arrangement of proteins. No evidence has been obtained for direct association of viral ribosomes with the viral nucleoproteins.

Positive staining of Tacaribe nucleoproteins with uranyl acetate revealed predominantly closed circular structures (PALMER et al. 1977). Length distribution analysis revealed two size classes with lengths of 640 nm and 1300 nm, presumably corresponding to the two distinct genome RNA species. Two viral ribonucleoprotein classes with distinct sedimentation properties were also resolved by analysis of detergent-disrupted LCM virus in sucrose density gradients (PEDERSEN and KONIGSHOFER 1976).

Examination of Pichinde virus particles disrupted by osmotic shock revealed strands about 12 nm in diameter apparently consisting of helical coils of the 4- to 5-nm nucleocapsid strands (YOUNG and HOWARD 1983). Isolated nucleocapsids in high-ionic-strength buffers were found to be further condensed into beadlike structures about 15 nm in diameter, that appeared to unfold as the ionic strength was reduced. It is not known whether these structures correspond to the arrangement of the nucleoprotein within the intact virion.

3 Replication

3.1 General Features of the Replication Cycle

The kinetics of replication in cell culture have been determined for several arenaviruses; in general, their growth cycle is longer than that of most other RNA viruses. For BHK21 cells infected under single-cycle conditions with Pichinde virus (Fig. 5; VEZZA et al. 1977) or Tacaribe virus (SALEH et al. 1979), the latent period was about 12 h, and maximum virus yields were detected 36–48 h post-infection (p.i.). The replication cycle of LCMV in BHK21 cells was significantly more rapid, with a latent period of 6 h, and a progressive rise in virus titer until 24 h p.i. (BUCHMEIER et al. 1978). The titers of released virus and cell-associated virus were observed to be approximately equivalent throughout the growth cycle of Tacaribe virus (SALEH et al. 1979). Infectious virus release continues for several days, and cytopathic effects develop very slowly. These observations are consistant with findings that host cell biosynthesis is virtually unaffected by arenavirus infection, and that synthesis of virus-specific components involves only a minor fraction of the total macromolecular synthesis in infected cells (BUCHMEIER et al. 1978; SALEH et al. 1979; GIMENEZ et al. 1983).

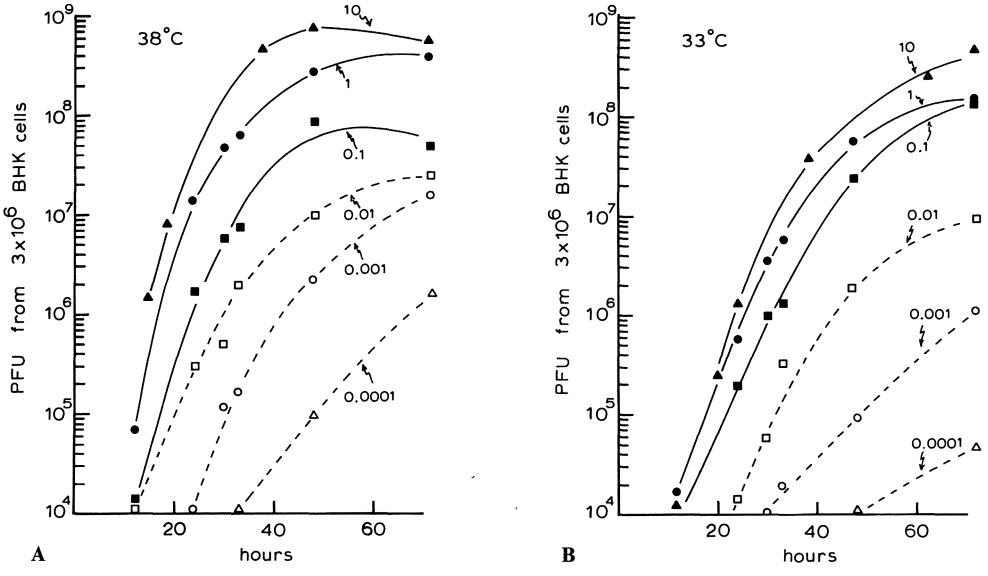


Fig. 5A, B. Growth of Pichinde virus as a function of the input multiplicity of infection and the temperature of incubation. Monolayers of BHK21 cells were infected at input multiplicities of infection varying from 10 to 0.0001 PFU/cell, and, after removal of the virus inoculum, the cells were incubated at 33° or 38° C; 2 h post-infection, the supernatant fluids were discarded to remove desorbed inoculum virus and the incubations were continued. At specified times, the PFU in the supernatant fluids were determined by using monolayers of Vero cells for plaque assays

3.2 Adsorption and Penetration

No information has been obtained on the nature of the cellular receptors for arenaviruses. The kinetics of adsorption and penetration have been examined for several arenaviruses (reviewed by RAWLS and LEUNG 1979). Maximum adsorption was observed within 2 h, and uncoating of LCM virus also occurred within this interval. Since it is likely that the viral ribonucleoproteins possess transcriptase activity, the uncoating process presumably involves release of these components into the cytoplasm.

3.3 RNA Transcription and Replication and RNA Polymerases

The presence of an RNA-directed RNA polymerase in preparations of Pichinde virus was initially reported by CARTER and associates (1974). These observations, together with evidence that viral complementary RNA extracted from infected cells (unlike virion RNA) could be translated in vitro (LEUNG et al. 1977), allowed arenaviruses to be classified as negative-strand viruses. RNA sequence studies (AUPERIN et al. 1981, 1982) and cloned DNA analyses (AUPERIN et al. 1984a, b) have confirmed that a subgenomic viral N mRNA is complementary in sequence to the 3′ half of the S RNA. By contrast the S-coded, subgenomic glycoprotein precursor (GPC) mRNA is identical in sequence to the 5′ half of the RNA (AUPERIN et al. 1984c).

It has not yet been established that the arenavirus particle contains a virus-coded transcriptase. In part this is because the viral transcriptase activity has a very low specific activity, and no *ts* mutants of the viral transcriptase have been identified. Furthermore, additional RNA polymerase activities, including homopolymeric polymerases characteristic of cellular enzymes, have been recov-

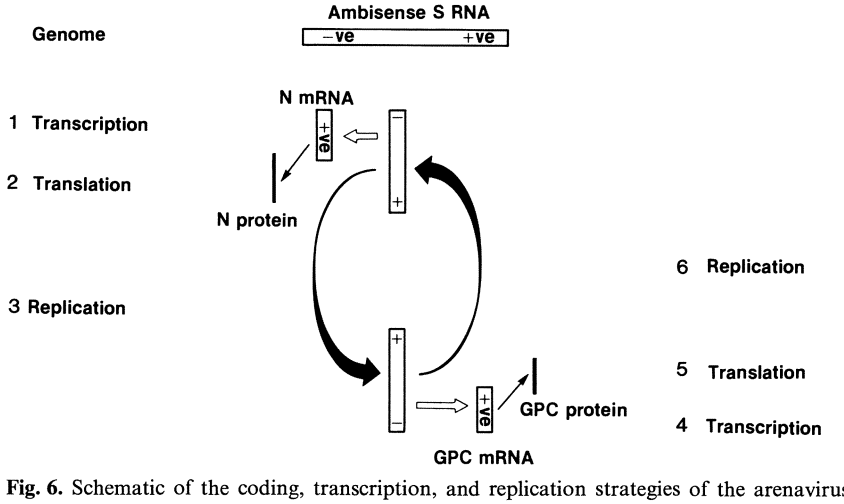


Fig. 6. Schematic of the coding, transcription, and replication strategies of the arenavirus S RNA species

ered from Pichinde viral preparations (LEUNG et al. 1979). A candidate virion protein for the viral transcriptase is the L protein in the case of Pichinde virus. Like transcriptases of other negative-strand viruses, it is present only in small quantities in the virus particle. The mechanism of synthesis of the virus-induced mRNA species has not yet been elucidated.

If arenaviruses are like other negative-strand viruses, the transcriptase is probably activated shortly after virus uncoating in the infected cell and yields multiple copies of the N mRNA and the L-coded mRNA species. Since the S RNA has an ambisense coding strategy, GPC mRNA cannot be made until after RNA replication has commenced (Fig. 6). The lower abundance of GPC mRNA, in comparison with N mRNA, in lytic infections M. GALINSKI and D.H.L. BISHOP, unpublished data), agrees with the conclusion that the two gene products are regulated independently. Independent regulation of GPC mRNA may be important in the establishment of viral persistence in vivo and in vitro. Whether or not the transcriptase also contributes to the RNA replication process is not known. No information is available concerning the process of viral RNA replication in infected cells.

3.4 Translation and Protein Processing

Since arenavirus-specific polypeptide synthesis comprises only a small fraction of the total cellular protein synthesis, it has been difficult to identify viral proteins by direct analysis of infected cells. In BHK21 cells infected with Tacaribe virus, the two major structural proteins N and G were observed above the host cell background by pulse labeling with amino acid or sugar precursors (SALEH et al. 1979). The nucleoprotein was first detected around 24 h p.i., and its rate of synthesis increased until 48 h p.i., followed by a slight decrease by 72 h p.i. The glycoprotein was not detected until 48 h and its rate of synthesis increased up to 60 h p.i. The nucleoprotein subunit of LCM virus was detected in infected BHK21 cells as early as 6 h p.i. and was produced in relatively constant amounts throughout the growth cycle (BUCHMEIER et al. 1978). In amino acid-labeled polypeptide patterns of Tacaribe or LCM virus-infected cells, the nucleoprotein was present in much larger molar amounts than was the glycoprotein, which was only resolved clearly with the use of sugar precursors. This result, as well as the difference in kinetics of appearance of the polypeptide species, agrees with the postulate that the nucleoprotein and glycoproteins are translated from distinct mRNA species.

Using procedures designed to suppress the background of host cell polypeptide synthesis, two additional polypeptide species, designated p79 (mol. wt. 79×10^3) and p105 (mol. wt. 105×10^3), were detected in BHK21 or Vero cells infected with Tacaribe virus analyzed by one- or two-dimensional polyacrylamide gel electrophoresis (GIMENEZ et al. 1983). Two-dimensional gel analysis indicated that p79 and p105 were acidic polypeptides that did not comigrate with any polypeptides detected in uninfected cells. The p79 polypeptide was further characterized using peptide mapping and appeared to be unrelated to the N protein in its primary sequence.

Immune precipitation has been used to identify virus-specific polypeptides in cells infected with LCM, Pichinde, and Tacaribe viruses (BUCHMEIER and OLDSTONE 1979; SALEH et al. 1979; HARNISH et al. 1981a, b). A polypeptide of approximately 200×10^3 daltons, designated the L protein, was detected in immune precipitates of Pichinde virus-infected cells (HARNISH et al. 1981a, b). The L polypeptide was stable in pulse-chase experiments and was distinct from the other virion polypeptides by peptide mapping. The two polypeptides designated p79 and p105 were also found in immune precipitates of Tacaribe virus-infected cells, using virus-specific rabbit or mouse antisera (GIMENEZ et al. 1983). A glycoprotein precursor was detected for both LCM virus (BUCHMEIER and OLDSTONE 1979) and Pichinde virus (HARNISH et al. 1981a, b). In both cases, the glycoprotein precursors were found to be related to the virion glycoproteins, G1 and G2, using peptide mapping and pulse-chase analysis. A similar precursor of mol. wt. 70×10^3 has been found in Tacaribe virus-infected cell lysates or immune precipitates using anti-Tacaribe antiserum; however, only a single cleavage product mol. wt. 42×10^3 has been identified (SALEH et al. 1979). The fate of the other fragment of about 30×10^3 daltons has not been determined.

Apart from the cleavage of the glycoprotein precursor, no evidence for an essential proteolytic processing step has been obtained for arenavirus polypeptides. Several cleavage products of the N protein have been detected in immune precipitation analyses with Pichinde or Tacaribe virus (HARNISH et al. 1981a, b; GIMENEZ et al. 1983) but were not found in Tacaribe-infected cell lysates analyzed directly, suggesting that breakdown occurred following cell lysis and during manipulations involved in immune precipitation.

The modifications of arenavirus polypeptides that have been detected include phosphorylation and glycosylation. The polypeptides of Tacaribe virions were analyzed using virions grown in the presence of $^{32}PO_4$, and a fraction of the N protein was found to be phosphorylated (GIMENEZ et al. 1983). No phosphorylation of other viral polypeptides was detected.

The carbohydrate components of virion glycoproteins have been described (see Sect. 2.2.5). The glycoprotein precursors of LCM and Tacaribe virus were found to be rich in mannose and glucosamine and to contain little galactose or fucose (BUCHMEIER and OLDSTONE 1979; BOERSMA et al. 1982). It is likely that processing of the viral oligosaccharides occurs during the migration of the glycoproteins from their site of synthesis to the cell surface; as with other viral glycoproteins, this probably involves trimming of mannose residues and addition of galactose, fucose, and sialic acid to form complex carbohydrate chains (HUNT et al. 1978; TABAS et al. 1978; NAKAMURA and COMPANS 1978).

3.5 Assembly and Release

The only stage in virus assembly that has been visualized using electron microscopy is the final process of budding at the plasma membrane (MURPHY et al. 1970). Presumably, viral nucleocapsids are assembled in the cytoplasm, since the nucleoprotein antigen can be identified throughout the cytoplasm using immunofluorescence. By analogy with other enveloped viruses, it is likely that

the viral glycoproteins are synthesized on membrane-bound polyribosomes in the rough endoplasmic reticulum and are transported through the Golgi complex, where simple, high mannose oligosaccharides are processed into complex carbohydrate chains. The final movement of glycoproteins to the plasma membrane is presumably mediated by a vesicular transport process, as has been observed with other enveloped viruses (ROTHMAN and FINE 1980).

Since the major virus-specified polypeptides that have been identified are the glycoproteins and the nucleoprotein, assembly of virions may occur by direct recognition between the nucleoprotein and an exposed, cytoplasmic tail on the glycoproteins. The possibility that an additional minor polypeptide participates in assembly (such as the low-mol.-wt. species that have occasionally been identified) cannot be excluded, however. If assembly occurs by interactions between the nucleoprotein and glycoproteins alone, the process might be expected to be less precise than that of other enveloped viruses which have a membrane (M) protein that plays a major role in aligning viral components in a domain at the plasma membrane. Lack of precision in virus assembly is consistent with the observed pleomorphism of the virions and the incorporation of host components, such as ribosomes.

Membrane changes seen at the sites of virus budding include an increase in density of both membrane lamellae in discrete areas large enough to form a viral envelope (MURPHY et al. 1970). Surface projections were observed on the exterior of emerging virus particles. In some cells observed at late stages of infection, extensive regions of the plasma membrane were involved in virus assembly.

In addition to the assembly of virions, electron-microscopic studies of arenavirus-infected Vero cells revealed distinctive intracytoplasmic inclusions consisting of aggregations of electron-dense granules with the appearance of ribosomes (MURPHY et al. 1970). The inclusions were variable in their size and shape and appeared to become progressively denser during the course of infection. The functional significance of these inclusions has not been determined.

4 Genetics

The segmented structure of the viral genetic information allows the possibility for recombinant viruses to be produced by RNA segment reassortment involving two genetically compatible viruses. High-frequency recombination was initially demonstrated using *ts* mutants of a high-temperature-adapted strain of Pichinde virus (VEZZA and BISHOP 1977). The mutants were categorized into two groups on the basis of their recombination abilities. Dual virus infections including mutants of only one group do not yield recombinant viruses. Infections involving two mutants representing both groups do yield recombinants. By performing dual virus infections with *ts* mutants of Pichinde and an alternate Pichinde strain, Pichinde Munchique, the RNA segment assignments of the mutants were determined and intertypic recombinant viruses were produced (VEZZA et al. 1980). This was accomplished by virtue of the fact that using oligonucleotide

fingerprinting the L and S RNA species of Munchique could be easily distinguished from those of prototype Pichinde virus, allowing the origins of the recombinant viral RNA segments to be determined (VEZZA et al. 1980). Tryptic peptide analyses of Pichinde and Munchique viral proteins have shown that the L, N, and GPC proteins of the respective viruses can also be distinguished. Thus, analyses of a Pichinde L-Munchique S recombinant virus demonstrated that the L RNA coded for the viral L protein and the S RNA for both N and GPC (BISHOP 1980; BISHOP et al. 1980; VEZZA et al. 1980; HARNISH et al. 1983).

Intertypic recombinant (reassortant) viruses have also been obtained from dual wild-type (KIRK et al. 1980) and dual *ts mutant LCM virus coinfections* (ROMANOWSKI and BISHOP 1983; AUPERIN et al. 1984b). Analyses of the virulence phenotypes of the intertypic reassortant arenaviruses, in comparison with those of the parental viruses, have established that the S RNA gene products determine virulence (KIRK et al. 1980; VEZZA et al. 1980).

The LCM *ts* mutants have been categorized into two recombination groups (ROMANOWSKI and BISHOP 1983; AUPERIN et al. 1984b). However, it has been shown that some of the LCM S RNA *ts* mutants can complement other LCM S RNA *ts* mutants. This result is understood in terms of their representing different S-coded gene products that are expressed independently. Dual-virus infections involving the complementing *ts* mutants also yield phenotypically wild-type, diploid viruses that are genetically unstable and on passage shed large numbers of *ts* progeny (ROMANOWSKI and BISHOP 1983). Diploidy was also demonstrated, using oligonucleotide fingerprint analyses of virus passaged at nonpermissive temperatures, by the presence in the recombinants of both parental mutant virus S RNA species (i.e., WE S RNA and ARM S RNA). The biological significance of diploid (multiploid) virus formation for arenaviruses is unknown; it may correlate with the fact that in virus extracts usually the S RNA is in molar excess over the L RNA. The inclusion of multiple copies of a viral RNA species in a virus particle may confer genetic stability and be understood in terms of the pleomorphic character of arenaviruses.

5 Defective Viruses and Persistent Infections

Persistent infections of cell cultures can be readily established with arenaviruses by infection under single-cycle conditions, followed by passage of the cells when the monolayers begin to show cytopathic effects. The resulting cells express viral antigens, exhibit normal morphology after several passages, and produce defective-interfering (DI) virus particles (WELSH et al. 1975; WELSH and BUCHMEIER 1979; GIMENEZ and COMPANS 1980). The most extensive biological studies on persistent infection in cell culture and in vivo have been carried out with LCM virus and have been reviewed elsewhere (BUCHMEIER et al. 1980; LEHMANN-GRUBE et al. 1983).

In general, arenavirus DI particles contain the normal structural polypeptides of the homologous parent, but altered RNA species. However, the exact

alteration found in the RNA profile has varied considerably in published reports. In sucrose gradient analyses, loss of one or more RNA segments and an increased amount of smaller RNA species have been observed (DUTKO et al. 1976; PEDERSEN 1979). Analysis of the RNA species of DI particles released from BHK21 cells persistently infected with Tacaribe virus revealed no detectable L or S RNA molecules; however, five species of RNA were observed that were smaller than any of the RNA species of standard virions, varying in size from 1.1×10^6 to 0.2×10^6 daltons (GIMENEZ and COMPANS 1980). Recent studies on the RNA species of DI influenza virus have also shown the presence of multiple new RNA species which are smaller than any of the normal genome segments (NAYAK et al. 1978; CRUMPTON et al. 1978). These RNAs contain the normal terminal sequences, but have internal deletions (NAYAK et al. 1982). Similarly, the smaller RNA species in Tacaribe DI particles are likely to result from deletions of parts of the L or S RNA segments.

Examination of Tacaribe DI particles, using negative staining, indicated a morphology similar to that of standard virions, with characteristic spikes on the surfaces (GIMENEZ and COMPANS 1980). However, the mean particle diameter was significantly smaller, being 55 nm as compared with 95 nm for the standard virus. A smaller size was also found for DI particles of LCMV (LEHMANN-GRUBE et al. 1983). The buoyant density of DI particles of LCM virus was found to be slightly lower than that of the corresponding standard virions (WELSH and BUCHMEIER 1979). The N polypeptide of DI particles obtained from persistently infected BHK21 cells showed a slight reduction in electrophoretic mobility compared with the N polypeptide of wild-type virions; however, no differences were detected in peptide profiles obtained by limited digestion of the two polypeptides with *Staphylococcus aureus* V8 protease (GIMENEZ and COMPANS 1980).

DI particles obtained from persistently infected cells block the replication of homologous virus; no interference is observed against heterologous viruses such as VSV (WELSH and BUCHMEIER 1979; GIMENEZ and COMPANS 1980). The persistently infected cells exhibit similar interference characteristics (LEHMANN-GRUBE 1967). Immune serum or UV-irradiation is able to prevent the interference properties of DI virus, and no interferon-like activity has been detected in DI virus preparations. Pretreatment of BHK21 cells with DI-LCMV was found to prevent the synthesis of viral proteins upon subsequent infection with standard virus (WELSH and BUCHMEIER 1979).

In several systems, production of viral antigens and DI particles has been observed in persistently infected cells which do not score as infectious centers and from which no infectious virus production has been detected (WELSH and BUCHMEIER 1979; GIMENEZ and COMPANS 1980). It is of interest to consider the mechanism of persistence in this situation. One possibility is that infection is maintained by an infectious virus variant that is difficult to detect using standard assay techniques (HOTCHIN 1974). Alternatively, replication of subgenomic RNA species may occur by a mechanism not involving fully infectious virus, although there is no precedent for such a mechanism of persistence in other virus systems. In addition, the fact that arenavirus DI particles are not enriched following coinfection of cells with standard virus and DI particles,

or by serial undiluted virus passage (LEHMANN-GRUBE et al. 1983), as has been observed for DI particles produced by other virus families (HUANG and BALTIMORE 1977), suggests that an unusual mechanism may be involved in DI particle production by arenaviruses.

Acknowledgements. These studies were supported in part by NIH grants AI-14183 and CA-18611.

References

Anon MC, Grau O, Martinez-Segovia Z, Franze-Fernandez MT (1976) RNA composition of Junin virus. J Virol 18:833–838

Armstrong C, Lillie RD (1934) Experimental lymphocyte choriomeningitis of monkeys and mice produced by a virus encountered in studies of the 1933 St. Louis encephalitis epidemic. Public Health Reports 49:1019–1027

Auperin D, Dimock K, Cash P, Rawls WE, Leung W-C, Bishop DHL (1981) Analyses of the genomes of prototype Pichinde and a virulent derivative of Pichinde Munchique arenaviruses: evidence for sequence conservation at the 3′ termini of their viral RNA species. Virology 116:363–367

Auperin D, Compans RW, Bishop DHL (1982) Nucleotide sequence conservation at the 3′ termini of the virion RNA species of New World and Old World arenaviruses. Virology 121:200–203

Auperin DD, Galinski M, Bishop DHL (1984a) The sequences of the N protein gene and intergenic region of the S RNA of Pichinde arenaviruses. Virology 134:208–219

Auperin DD, Romanowski V, Bishop DHL (1984b) Genetic analyses of Pichinde and LCM arenaviruses; evidence for a unique organization for Pichinde S RNA. In: Compans RW, Bishop DHL (eds) Segmented negative strand viruses: Arenaviruses, bunyaviruses and orthomyxoviruses. Academic, New York, pp 51–57

Auperin DD, Romanowski V, Galinski M, Bishop DHL (1984c) Sequencing studies of Pichinde, Arenavirus S RNA indicate a novel coding strategy, ambisense viral S RNA. J Virol 52:397–904

Bishop DHL (1980) Propiedades moleculares y genéticas de los arenavirus. Medicina 40:275–288

Bishop DHL, Beaty BJ, Shope RE (1980) Recombination and gene coding assignments of bunyaviruses and arenaviruses. Ann N Y Acad Sci 354:84–106

Boersma DP, Saleh F, Nakamura K, Compans RW (1982) Structure and glycosylation of Tacaribe viral glycoproteins. Virology 123:452–456

Buchmeier MJ, Oldstone MBA (1978a) Identity of the viral protein responsible for serological cross-reactivity among the Tacaribe complex arenaviruses. In: Mahy BWJ, Barry RD (eds) Negative strand viruses and the host cell. Academic, New York, pp 91–97

Buchmeier MJ, Oldstone MBA (1978b) Virus-induced immune complex disease: Identification of specific viral antigens and antibodies deposited in complexes during chronic lymphocytic choriomeningitis virus infection. J Immunol 120:1297–1304

Buchmeier MJ, Oldstone MBA (1979) Protein structure of lymphocytic choriomeningitis virus: evidence for a cell-associated precursor of the virion glycopeptides. Virology 99:111–120

Buchmeier MJ, Elder JH, Oldstone MBA (1978) Protein structure of lymphocytic choriomeningitis virus: identification of the virus structural and cell associated polypeptides. Virology 89:133–145

Buchmeier MJ, Welsh R, Dutko F, Oldstone MBA (1980) The virology and immunobiology of lymphocytic choriomeningitis virus infection. Adv Immunol 30:275–331

Buchmeier MJ, Lewicki HA, Tomori O, Oldstone MBA (1981) Monoclonal antibodies to lymphocytic choriomeningitis and Pichinde viruses: generation, characterization, and cross-reactivity with other arenaviruses. Virology 113:73–85

Buckley SM, Casals J (1970) Lassa fever, a new virus disease of man from West Africa. III. Isolation and characterization of the virus. Am J Trop Med Hyg 19:680–691

Calisher CH, Tzianabos T, Lord RD, Coleman PH (1970) Tamiami virus, a new member of the Tacaribe group. Am J Trop Med Hyg 19:520–526

Carter MF, Biswal N, Rawls WE (1973a) Characterization of nucleic acid of Pichinde virus. J Virol 11:61–68

Carter MF, Murphy FA, Brunschwig JP, Noonan C, Rawls WE (1973b) Effects of actinomycin D and ultraviolet and ionizing radiation on Pichinde virus. J Virol 12:33–38
Carter MF, Biswal N, Rawls WE (1974) Polymerase activity of Pichinde virus. J Virol 13:577–583
Casals J (1975) Arenaviruses. Yale J Biol Med 48:115–140
Chinault DN, Thompson HA, Gangemi JD (1981) Polypeptide synthesis catalysed by components of Pichinde virus disrupted by detergent. J Gen Virol 55:213–217
Compans RW, Klenk HD (1979) Viral membranes. In: Comprehensive virology (Fraenkel-Conrat H. Wagner RR eds), Plenum Press, Vol 13, pp 293–407
Compans RW, Boersma DP, Cash P, Clerx JPM, Gimenez HB, Kirk WE, Peters CJ, Vezza AC, Bishop DHL (1981) Molecular and genetic studies of Tacaribe, Pichinde, and lymphocytic choriomeningitis viruses. In: Bishop DHL, Compans RW (eds) Replication of negative strand viruses. Elsevier, New York, pp 31–42
Crumpton WM, Dimmock NJ, Minor PD, Avery RJ (1978) The RNAs of defective-interfering influenza virus. Virology 90:370–373
Dalton AJ, Rowe WP, Smith GH, Wilsnack RE, Pugh WE (1968) Morphological and cytochemical studies on lymphocytic choriomeningitis virus. J Virol 2:1465–1478
Dimock K, Harnish D, Sisson G, Leung W-C, Rawls WE (1982) Synthesis of viral specific polypeptides and genomic RNA during the replicative cycle of Pichinde virus. J Virol 43:273–283
Downs WG, Andersen CR, Spence TH, Aitken G, Greenhall AH (1963) Tacaribe virus, a new agent isolated from antibesis bats and mosquitoes in Trinidad, West Indies. Am J Trop Med Hyg 12:640–646
Dutko FJ, Pfau CJ (1978) Arenavirus defective-interfering particles mask the cell-killing potential of standard virus. J Gen Virol 38:195–208
Dutko FJ, Wright EA, Pfau CJ (1976) The RNAs of defective-interfering Pichinde virus. J Gen Virol 31:417–427
Dutko FJ, Kennedy SIT, Oldstone MBA (1981) Genome structure of lymphocytic choriomeningitis virus: cohesive complementary termini? In: Bishop DHL, Compans RW (eds) Replication of negative strand viruses. Elsevier North-Holland, New York, pp 43–47
Farber FE, Rawls WE (1975) Isolation of ribosome-like structure from Pichinde virus. J Gen Virol 26:21–31
Frame JE, Baldwin JM, Gocke DJ, Troup JM (1970) Lassa fever, a new virus disease of man from West Africa. 1. Clinical description and pathological findings. Am J Trop Med Hyg 19:670–676
Gangemi JD, Rosato RR, Connell EV, Johnson EM, Eddy GA (1978) Structural polypeptides of Machupo virus. J Gen Virol 41:183–188
Gard GP, Vezza AC, Bishop DHL, Compans RW (1977) Structural proteins of Tacaribe and Tamiami virions. Virology 83:84–95
Gimenez HB, Compans RW (1980) Defective-interfering Tacaribe virus and persistently infected cells. Virology 107:229–239
Gimenez HB, Compans RW (1981) Oligonucleotide fingerprint analysis of Tacaribe virion RNA. J Gen Virol 55:219–222
Gimenez HB, Boersma DP, Compans RW (1983) Analysis of polypeptides in Tacaribe virus-infected cells. Virology 128:469–473
Gonsalez JP, McCormick JB, Saluzzo JF, Herve JP, Georges AJ, Johnson KM (1983) An arenavirus isolated from wild-caught rodents (*Praomys* species) in the Central African Republic. Interviriology 19:105–112
Grau O, Franze-Fernandez MT, Romanowski V, Rustici SM, Rosas MF (1981) Junin virus structure. In: Bishop DHL, Compans RW (eds) The replication of negative strand viruses. Elsevier, New York, pp 11–14
Harnish DG, Dimock K, Leung W-C, Rawls WE (1981a) Immunoprecipitable polypeptides in Pichinde virus-infected BHK-21 cells. In: Bishop DHL, Compans RW (eds) The replication of negative strand viruses. Elsevier, New York, pp 23–29
Harnish DG, Leung W-C, Rawls WE (1981b) Characterization of polypeptides immunoprecipitable from Pichinde virus-infected BHK21 cells. J Virol 38:840–848
Harnish DG, Dimock K, Bishop DHL, Rawls WE (1983) Gene mapping in Pichinde virus: assignment of viral polypeptides to genomic L and S RNAs. J Virol 46:638–641
Hotchin J (1974) The role of transient infection in arenavirus persistence. Progr Med Virol 18:81–93
Howard CR, Buchmeier MJ (1983) A protein kinase activity in lymphocytic choriomeningitis virus

and identification of the phosphorylated product using monoclonal antibody. Virology 126:538–547
Howard CR, Simpson DI (1980) The biology of arenaviruses. J Gen Virol 51:1–14
Huang AS, Baltimore D (1977) Defective-interfering animal viruses. In: Fraenkel-Conrat H, Wagner RR (eds) Comprehensive virology, vol 10. Plenum, New York, pp 73–116
Huang AS, Baltimore D (1977) Defective interfering animal viruses. In: Comprehensive Virology (Fraenkel-Conrat H, Wagner RR eds) Plenum Press, Vol 10, pp 73–116
Hunt LA, Etchinson JR, Summers DF (1978) Oligosaccharide chains are trimmed during synthesis of the envelope glycoprotein of vesicular stomatitis virus. Proc Natl Acad Sci USA 75:754–758
Johnson KM, Kuns ML, Mackenzie RB, Webb PA, Yunker CE (1966) Isolation of Machupo virus from wild rodent *Calomys callosus*. Am J Trop Med Hyg 15:103–106
Johnson KM, Webb PA, Justines G (1973) Biology of Tacaribe complex viruses. In: Lehmann-Grube F (eds) Lymphocytic choriomeningitis and other arenaviruses. Springer, Berlin Heidelberg New York, pp 241–258
Kiley MP, Tomori O, Regnery RL, Johnson KM (1981) Characterization of the arenaviruses Lassa and Mozambique. In: Bishop DHL, Compans RW (eds) The replication of negative strand viruses. Elsevier, New York, pp 1–9
Kirk WE, Cash P, Peters CJ, Bishop DHL (1980) Formation and characterization of an intertypic lymphocytic choriomeningitis recombinant virus. J Gen Virol 51:213–218
Lehmann-Grube F (1967) A carrier state of lymphocytic choriomeningitis virus in L cell cultures. Nature 213:770–773
Lehmann-Grube F (1971) Lymphocytic choriomeningitis virus. Virol Monogr 10:1–173
Lehmann-Grube F, Peralta LM, Bruns M, Lohler J (1983) Persistent infection of mice with the lymphocytic choriomeningitis virus. In: Fraenkel-Conrat H, Wagner RR (eds) Comprehensive virology, vol 18. Plenum, New York, pp 43–103
Leung W-C (1978) Considerations of the replication of Pichinde virus – an arenavirus. In: Mahy BWJ, Barry RD (eds) Negative strand viruses and the host cell. Academic Press, New York, pp 415–426
Leung W-C, Rawls WE (1977) Virion-associated ribosomes are not required for the replication of Pichinde virus. Virology 81:174–176
Leung W-C, Ghosh HP, Rawls WE (1977) Strandedness of Pichinde virus RNA. J Virol 22:235–237
Leung W-C, Leung MFKL, Rawls WE (1979) Distinctive RNA transcriptase, polyadenylic acid polymerase, and polyuridylic acid polymerase activities associated with Pichinde virus. J Virol 30:98–107
Leung W-C, Harnish D, Ramsingh A, Dimock K, Rawls WE (1981) Gene mapping in Pichinde virus. In: Bishop DHL, Compans RW (eds) The replication of negative strand viruses. Elsevier, New York, pp 51–57
Martinez-Segovia ZM, De Mitri MI (1977) Junin virus structural proteins. J Virol 21:579–583
Matthews REF (1982) Classification and nomenclature of viruses. Intervirol 17:1–200
Murphy FA, Webb PA, Johnson KM, Whitfield SG (1969) Morphological comparison of Machupo with lymphocytic choriomeningitis virus: basis for a new taxonomic group. J Virol 4:535–541
Murphy FA, Webb PA, Johnson KM, Whitfield SG, Chappell WA (1970) Arenaviruses in Vero cells: ultrastructural studies. J Virol 6:507–518
Nakamura K, Compans RW (1978) Glycopeptide components of influenza viral glycoproteins. Virology 86:432–442
Nayak DP, Tobita K, Janda JM, Davis AR, De BK (1978) Homologous interference mediated by defective-interfering influenza virus derived from a temperature-sensitive mutant of influenza virus. J Virol 28:374–386
Nayak DP, Sivasubramanian N, Davis AR, Cortini R, Sung J (1982) Complete sequence analyses show that two defective-interfering influenza viral RNAs contain a single internal deletion of a polymerase gene. Proc Natl Acad Sci USA 79:2216–2220
Oldstone MBA, Dixon FJ (1969) Pathogenesis of chronic disease associated with persistent lymphocytic choriomeningitis viral infection. J Exp Med 129:483–505
Palmer EL, Obijeski JF, Webb PA, Johnson KM (1977) The circular, segmented nucleocapsid of an arenavirus-Tacaribe virus. J Gen Virol 36:541–545
Parodi AJ, Greenway DJ, Rugiero HR, Rivers S, Figerio M, de la Barrera JM, Mattler N, Garzon F, Boxaca M, de Guerrero L, Nota N (1958) Sobre la etiología del brote epidermico de Junin. El Dia Medico 30:2300–2302

Pedersen IR (1970) Density gradient centrifugation studies on lymphocyte choriomeningitis virus and on viral ribonucleic acid. J Virol 6:414–420
Pedersen IR (1971) Lymphocytic choriomeningitis virus RNAs. Nature 234:112–114
Pedersen IR (1973) LCM virus: its purification and its chemical and physical properties. In: Lehmann-Grube F (ed) Lymphocytic choriomeningitis and other arenaviruses. Springer, Berlin Heidelberg New York, pp 13–24
Pedersen IR (1979) Structural components and replication of arenaviruses. Adv Virus Res 24:277–330
Pedersen IR, Konigshofer EP (1976) Characterization of ribonucleoproteins and ribosomes isolated from lymphocytic choriomeningitis virus. J Virol 20:14–21
Pfau CJ, Bergold GH, Casals J, Johnson KM, Murphy FA, Pedersen IR, Rawls WE, Rowe WP, Webb PA, Weissenbacher MC (1974) Arenaviruses. Intervirology 4:207–213
Pinheiro FP, Shope RE, de Androde AHP, Bensabeth G, Cacios GV, Casals J (1966) Amapari, a new virus of the Tacaribe group from rodents and mites of Amapa territory, Brazil. Proc Soc Exp Biol Med 122:531–535
Pinheiro FP, Woodall JP, Travassos da Rosa APA, Travassos da Rosa JF (1977) Studies of arenaviruses in Brazil. Medicina (Buenos Aires) 37:175–181
Ramos BA, Courtney RJ, Rawls WE (1972) Structural proteins of Pichinde virus. J Virol 10:661–667
Ramsingh AI, Dimock K, Rawls WE, Leung W-C (1980) Size estimation of Pichinde virus RNA by gel electrophoresis under denaturing conditions. Intervirology 14:31–36
Rawls WE, Leung W-C (1979) Arenaviruses. In: Fraenkel-Conrat H, Wagner RR (eds) Comprehensive virology, vol 14. Plenum, New York, pp 157–192
Romanowski V, Bishop DHL (1983) The formation of arenaviruses that are genetically diploid. Virology 126:87–95
Rothman JE, Fine RE (1980) Coated vesicles transport newly synthesized membrane glycoproteins from endoplasmic reticulum to the plasma membrane in two successive stages. Proc Natl Acad Sci, USA 77:780–784
Rowe WP, Murphy FA, Bergold GH, Casals J, Hotchin J, Johnson KM, Lehmann-Grube F, Mims CA, Traub E, Webb PA (1970a) Arenoviruses: proposed name for a newly defined virus group. J Virol 5:651–652
Rowe WP, Pugh WF, Webb PA, Peters CJ (1970b) Serological relationships of the Tacaribe complex of viruses to lymphocytic choriomeningitis virus. J Virol 5:289–292
Saleh F, Gard GP, Compans RW (1979) Synthesis of Tacaribe viral proteins. Virology 93:369–376
Tabas I, Schlesinger S, Kornfeld S (1978) Processing of high-mannose oligosaccharides to form complex-type oligosaccharides on the newly synthesized polypeptide of the vesicular stomatitis virus G protein and the IgG heavy chain. J Biochem 253:716–722
Trapido H, Sanmartin C (1971) Pichinde virus: a new virus of the Tacaribe group from Colombia. Am J Trop Med Hyg 20:631–641
Vezza AC, Bishop DHL (1977) Recombination between temperature-sensitive mutants of the arenavirus Pichinde. J Virol 24:712–715
Vezza AC, Gard GP, Compans RW, Bishop DHL (1977) Structural components of the arenavirus Pichinde. J Virol 23:776–786
Vezza AC, Clewley JP, Gard GP, Abraham NZ, Compans RW, Bishop DHL (1978) Virion RNA species of the arenaviruses Pichinde, Tacaribe, and Tamiami. J Virol 26:485–497
Vezza AC, Cash P, Jahrling P, Eddy G, Bishop DHL (1980) Arenavirus recombination: the formation of recombinants between prototype Pichinde and Pichinde Munchique viruses and evidence that arenavirus S RNA codes for N polypeptide. Virology 106:250–260
Webb PA, Johnson KM, Hibbs JB, Kuns ML (1970) Parana, a new Tacaribe complex virus from Paraguay. Arch Gesamte Virusforsch 32:379–388
Welsh RM Jr, Buchmeier MJ (1979) Protein analysis of defective interfering lymphocytic choriomeningitis virus and persistently infected cells. Virology 96:503–515
Welsh RM, Burner PA, Holland JJ, Oldstone MBA, Thompson HA, Villarreal LP (1975) A comparison of biochemical and biological properties of standard and defective lymphocytic choriomeningitis viruses. Bull WHO 52:403–408
Wulff H, McIntosh BM, Hammer DB, Johnson KM (1977) Isolation of an arenavirus closely related to Lassa virus from *Mastomys natalensis* in Southeast Africa. Bull WHO 55:441–444
Young PR, Howard CR (1983) Fine structure analysis of Pichinde virus nucleocapsids. J Gen Virol 64:833–842

Structure and Diversity of Influenza Virus Neuraminidase

P.M. COLMAN and C.W. WARD

CSIRO, Division of Protein Chemistry, 343 Royal Parade, Parkville 3052, Victoria, Australia

Current Topics in Microbiology and Immunology, Vol. 114

1 Introduction

Despite all that has been learned about influenza virus in the years since its isolation (SMITH et al. 1933), no method of cure or control has yet been found. Vaccination, so successful against polio and smallpox, is frustrated by antigenic variation of the two surface glycoproteins of the virus, and prospects for chemotherapy are uncertain.

Enzyme activity on the viral surface was discovered by HIRST (1942), who observed that cells once agglutinated by influenza could not again be agglutinated by either the eluted virus or fresh vürus preparations. The alteration in the cells is now known to be the result of the destruction of receptors for viral hemagglutinin by viral neuraminidase. A similar activity was subsequently found in culture filtrates of the bacterium *Vibrio cholerae* by BURNET and STONE (1947).

In this article we review recent studies on the primary and three-dimensional structure of influenza neuraminidase and discuss the bearing of these results on the current understanding of antigenic variation in influenza. Work on hemagglutinin has already been reviewed (WARD 1981) in this series. Recent reviews of neuraminidases include COLMAN (1984), ROSENBERG and SCHENGRUND (1976), and BUCHER and PALESE (1975).

1.1 Role of Neuraminidase in Influenza Life Cycle

Influenza is an enveloped virus. Like other orthomyxoviruses, it has two distinct, integral membrane glycoproteins, a hemagglutinin and a neuraminidase (LAVER and KILBOURNE 1966; LAVER and VALENTINE 1969). The hemagglutinin is specific for *N*-acetylneuraminic acid-containing receptors on the surface of susceptible cells (see, e.g., ROGERS and PAULSON 1983). Attachment to these cells via these receptors provides the initial contact between virus and host cell and is followed by the fusion of the viral envelope with cellullar membranes (KLENK and ROTT 1980; MATLIN et al. 1981). The fusion function is also a property of the hemagglutinin polypeptide (KLENK et al. 1975; LAZAROWITZ and CHOPPIN 1975).

Viral neuraminidase is an exoglycosidase, specifically removing terminal sialic acid (*N*-acetylneuraminic acid) from an α-ketosidic linkage to an adjacent sugar residue (KLENK et al. 1955; GOTTSCHALK 1957). The biological role of this enzyme activity has been the focus of some dispute. From this debate the view is emerging that the general function of viral neuraminidase is to mobilize the virus both to and from the site of infection. Mucosal secretions are rich in sialic acid-containing macromolecules (GOTTSCHALK 1958, 1972). Without neuraminidase activity, entrapment of the virus in these secretions of the upper respiratory tract is likely (BURNET et al. 1947; BURNET 1948), though not proven. Following a single cycle of replication in the host cell, progeny virions bud out from the cell plasma membrane, at which point two potentially immobilizing influences come into play. One is the sialic acid of the host cell membrane, and the other is terminal sialic acid on the carbohydrate moieties of the newly synthesized hemagglutinin and neuraminidase glycoproteins. Inactivation or inhibition of neuraminidase during budding has been observed to result in the aggregation of virions on the cell surface (PALESE et al. 1974; PALESE and COMPANS 1976; GRIFFIN and COMPANS 1979). A recent study of factors influencing influenza virus release from infected cells has shown that even when the viral envelope contains no sialic acid, as is the case when virus is grown under conditions of hexose starvation, virions aggregate at the cell surface (GRIFFIN et al. 1983). Under these conditions, virions carry a functional hemagglutinin but an inactive neuraminidase. The possible influence of the carbohydrate moiety of neuraminidase on the activity of the latter will be discussed in Sect. 4.4.

It has been reported (HUANG et al. 1980) that neuraminidase plays an essential role in membrane fusion. Liposomes bearing viral glycoproteins were observed to fuse with chick embryo cells only when neuraminidase was present, either bound to the liposome or in soluble form. However, other studies do not support the requirement for neuraminidase in fusion. WHITE et al. (1982a), using hemagglutinin expressed at the surface of simian CV-1 cells (infected with a recombinant SV40 virus containing the hemagglutinin gene of influenza A/Japan/305/57), showed that hemagglutinin is necessary and sufficient for pH-dependent fusion of these cells with each other. Not even sialic acid receptors on the target cell surface are required for fusion to occur (WHITE and HELENIUS 1980; WHITE et al. 1982b).

1.2 Biological Consequences of Neuraminidase Activity During Influenza Infection

Sialic acids are now recognized to play a key role in the regulation of many biological phenomena (ROSENBERG and SCHENGRUND 1976; SCHAUER 1982). They exert this control through their charge, by masking other structures on cell surfaces, or by active involvement as receptors (SCHAUER 1982). Neuraminidases in mammals are usually membrane-bound (see SCHAUER 1982; COLMAN 1984). The introduction of a mobile, exogenous neuraminidase during influenza infection will result in desialylation of cells contacted by the virus, with a concomitant disturbance of the regulatory processes mediated by sialic acid on those cells.

An exhaustive summary here is not appropriate; we choose only to touch on some topics which could be of relevance to influenza and its clearance.

1.2.1 The Alternate Complement Pathway

One function of viral neuraminidase is to remove sialic acid from the surface glycoproteins of the virus. A strong negative correlation has now been established by several authors between the sialic acid content of a viral envelope and the capacity of that virus to activate the alternate complement pathway. Membrane-bound sialic acid interferes with the C3b feedback cycle by promoting attachment of the regulatory protein β1H to C3b (KAZATCHKINE et al. 1979). Newcastle disease virus, devoid of surface sialic acid because of its neuraminidase activity, is inactivated by the alternate complement pathway (WEDGWOOD et al. 1956). Similarly, influenza and the parainfluenza virus simian virus 5 (SV5), together with vesicular stomatitis virus (VSV) grown in the presence of SV5, all of which lack surface sialic acids, were the most effective activators of the alternate complement pathway studied by MCSHARRY et al. (1981). On the other hand, VSV grown in MDBK cells and Sindbis virus grown in BHK21-F cells, containing 5 and 13 μg of sialic acid per mg of protein respectively were poor activators of the alternate complement pathway. Sindbis virus was the least effective activator, while intermediate levels of activation were observed for VSV (MCSHARRY et al. 1981). Similarly, Sindbis virus grown in different cell types or enzymatically modified by neuraminidase treatment shows a quantitatively inverse relationship between envelope sialic acid content and ability to activate the alternate complement pathway (HIRSCH et al. 1980, 1981). This phenomenon is not restricted to enveloped viruses. Neuraminidase-treated guinea pig erythrocytes also activate this complement pathway (LAMBRÉ et al. 1982). The polyanion heparin modulates the alternate pathway in a way similar to cell surface sialic acid (WEILER et al. 1978; MAILLET and KAZATCHKINE 1983), indicating that the control is likely to be based on charge effects rather than stereochemical specificity of sialic acid.

1.2.2 Unmasking of Structures by Neuraminidase

Sialic acid on serum glycoproteins masks D-galactosyl (D-gal) residues, and desialylation of these proteins triggers their removal from circulation as a result of their binding to specific D-gal receptors on the surface of hepatocytes (ASHWELL and MORELL 1974). Carbohydrates having a common specificity with ABH-blood groups are unmasked on the surface of erythrocytes by treatment with neuraminidase (SABER et al. 1965). Similarly, the lifetime of erythrocytes is dramatically reduced after neuraminidase treatment (JANCIK et al. 1975), as a result of phagocytosis by liver Kupffer cells which express lectins with D-gal specificity on their surface (SCHAUER 1982). Guinea pig erythrocytes exposed to influenza virus or *Vibrio cholerae* neuraminidase activate the classical complement pathway in autologous serum (LAMBRÉ et al. 1983). Antigenic determinants on the red cells are exposed by the removal of sialic acid and immunoglobulin M (IgM) autoantibodies fix to those determinants. These observations led to the suggestion that membrane desialylation during influenza virus infection could lead to autoimmunity and might assist recovery by the elimination of virus-modified cells.

Desialylation of lymphocytes and antibodies may prejudice the performance of the immune system. Lymphocytes treated with neuraminidase bind to hepatocytes and exhibit increased cytotoxicity (KOLB-BACHOFEN and KOLB 1979). WOODRUFF and GESNER (1969) observed that lymphocytes stripped of sialic acid are reversibly trapped in the liver. Based on turnover rates of sialic acid in cell membranes (KREISEL et al. 1980), SCHAUER (1982) has postulated that these lymphocytes are only released after resialylation of their membrane glycoconjugates. The sialic acid residues of immunoglobulins are also of some importance. In one instance, their removal led immunoglobulin G (IgG) to become immunogenic and arthritogenic (GALLOWAY et al. 1983) and resulted in joint lesions resembling those found in rheumatoid arthritis.

Given the pace of recent discoveries in this area, it is likely that the role of terminal sialic acid residues on soluble or cell-bound proteins is not yet fully understood and that the list of potential biological disturbances caused by viral neuraminidase during influenza infection is incomplete.

2 Primary Structure

Since previous reviews (LAVER 1973; WHITE 1974; BUCHER and PALESE 1975; KENDAL and KILEY 1975) have described much of the background information on influenza neuraminidase, these aspects will be summarized only briefly here. This will be followed by detailed descriptions of the protein and gene structures for two of the nine type A influenza neuraminidase subtypes (N1 and N2) and one type-B influenza virus neuraminidase. The amino acid and nucleotide sequence changes associated with antigenic shift and drift will be discussed in the section on antigenic structure (Sect. 5).

2.1 Background Information

Neuraminidase accounts for approximately 10% of the visible spikes projecting out from the viral surface (see WHITE 1974). Estimates suggest that there are from 50 to 100 neuraminidase spikes per virion (BUCHER and PALESE 1975), and antibody binding studies suggest these may not be distributed evenly over the surface of the virus (COMPANS et al. 1969; KENDAL and MADELEY 1970). Neuraminidase spikes can be isolated from influenza virus by treatment with detergents (LAVER 1963; WEBSTER and DARLINGTON 1969; GREGORIADES 1972; STANLEY et al. 1973), lipid solvents (HOYLE 1952; SETO et al. 1966), or proteolytic enzymes such as trypsin (NOLL et al. 1962; HASLAM et al. 1970), subtilisin (KENDAL and ECKERT 1972), and the proteolytic complex pronase (SETO et al. 1966; ROTT et al. 1972).

When viewed under the electron microscope, detergent-released neuraminidase spikes exhibit a mushroom-shaped appearance (LAVER and VALENTINE 1969; WRIGLEY et al. 1973; GRIFFITH 1975). They have a box-like head $80 \times 80 \times 40$ Å and a narrow, centrally attached fiber or stalk 15 Å wide and 100 Å in length. At the base of the stalk is a very hydrophobic knob that causes aggregation of isolated neuraminidase peplomers in the absence of detergent and is believed to anchor neuraminidase in the viral membrane (LAVER and VALENTINE 1969). Neuraminidase spikes released by proteolytic digestion can be seen to have lost the stalk region of the peplomer (WRIGLEY et al. 1973) and no longer aggregate. The resultant "heads" are antigenically and enzymatically intact (DRZENIEK et al. 1968; WRIGLEY et al. 1977; LAVER 1978).

Molecular weight estimates of detergent-released neuraminidase preparations before (BUCHER and KILBOURNE 1972) and after (HASLAM et al. 1970; WEBSTER 1970; WRIGLEY et al. 1973) denaturation and reduction indicate that the neuraminidase spike has a total molecular weight of 240000 and is a tetramer of four 60000-dalton polypeptides. The protease-released heads have a lower molecular weight of 200000 (NOLL et al. 1962; WRIGLEY et al. 1973; BLOK et al. 1982) and are made up of four 50000-dalton polypeptide subunits (KENDAL and ECKERT 1972; WRIGLEY et al. 1973; LAVER 1978; BLOK et al. 1982), indicating that the membrane-embedded stalk region comprises approximately 17% (40000 daltons) of the neuraminidase tetramer. At one stage it was believed that neuraminidase might be composed of two different kinds of polypeptide chains, since two bands had been observed on acrylamide gel analysis of detergent-purified neuraminidase from the recombinant virus X7.F1 (WEBSTER 1970; SKEHEL and SCHILD 1971; LAVER and BAKER 1972; BUCHER and KILBOURNE 1972). However, peptide maps (KENDAL and KILEY 1975; LAVER 1978), the more recent amino acid sequence (BLOK et al. 1982; WARD et al. 1982; LAVER et al. 1982) and gene sequence data (FIELDS et al. 1981; HITI and NAYAK 1982; ELLEMAN et al. 1982; MARKOFF and LAI 1982; BENTLEY and BROWNLEE 1982; VAN ROMPUY et al. 1982; SHAW et al. 1982), and the recently completed three-dimensional crystal structure (VARGHESE et al. 1983) show that influenza virus neuraminidase is a tetramer made up of four identical polypeptide chains.

The neuraminidase monomer is coded for by a single segment of RNA. In most strains, the neuraminidase gene is segment 6 (numbering from lowest

mobility to highest), although in some gel systems the position of the neuraminidase gene and the nucleoprotein gene (fifth gene segment) are reversed (see PALESE 1977 and SCHOLTISSEK 1978a for reviews). Neuraminidase is synthesized as a single polypeptide chain, and the steps involved in its production on the rough endoplasmic reticulum and transfer to the plasma membrane are believed to be similar to those observed for hemagglutinin (COMPANS and KLENK 1979; KLENK and ROTT 1980). Cotranslational and post-translational processing involve glycosylation by host cell glycosyltransferases. Although the primary and tertiary structure of neuraminidase dictates the availability of potential glycosylation sites (CLAMP 1975; WAGH and BAHL 1981), the size and composition of the attached carbohydrate units are influenced by the host cell (SCHWARZ et al. 1977; NAKAMURA and COMPANS 1978). Chemical stability studies (ALLEN et al. 1977; Keil et al. 1979) and direct amino acid sequence analysis (WARD et al. 1982) show that neuraminidase carbohydrate side chains, like those on hemagglutinin (see WARD 1981), are attached in *N*-glycosidic linkage to asparagine residues. Qualitative experiments with radiolabeled sugars (SCHWARZ et al. 1977; NAKAMURA and COMPANS 1978) and quantitative analyses (WARD et al. 1983b) show the presence of both *N*-acetyllactosamine-type and oligomannoside-type carbohydrate side chains. Compositional analysis of detergent-released whole neuraminidase and trypsin-released neuraminidase heads from the B/Lee/40 strain of influenza virus shows that more than 50% of the carbohydrate is associated with the stalk region of the protein (ALLEN et al. 1977).

In contrast to hemagglutinin, neuraminidase protein is not subject to proteolytic processing during biosynthesis. There is neither proteolytic removal of the N-terminal methionine residue or of an N-terminal hydrophobic signal peptide (BLOK et al. 1982), nor is there any proteolytic processing internally or at the C-terminal end (ELLEMAN et al. 1982; WARD et al. 1982). This is in marked contrast to influenza virus hemagglutinin, which loses a 15–18-residue signal peptide from its N-terminus (MCCAULEY et al. 1979; AIR 1981) and is cleaved internally to yield the two polypeptide chains HA1 and HA2 (LAZAROWITZ et al. 1971; KLENK et al. 1972; SKEHEL 1972; SKEHEL and WATERFIELD 1975; PORTER et al. 1979).

2.2 Type A N2 Neuraminidase

N2 neuraminidase will be described first and used as the prototype structure since it has been characterized by both gene sequence (MARKOFF and LAI 1982; ELLEMAN et al. 1982; BENTLEY and BROWNLEE 1982; VAN ROMPUY et al. 1982) and protein sequence (BLOK et al. 1982; WARD et al. 1982) analysis; its disulfide-bond arrangements (WARD et al. 1983a) and oligosaccharide distribution (WARD et al. 1983b) have been established chemically, and its three-dimensional structure recently resolved (VARGHESE et al. 1983). The N2 neuraminidase structure discussed in this section will be that of A/RI/5^{-}/57, one of several strains isolated at the beginning of the Asian influenza pandemic of 1957 (see WEBSTER et al. 1982). It represents the initial structure from which antigenic drift in N2 neuraminidase proceeded. This section will also describe the protein chemical studies

that established the nature of the primary gene product, its orientation in the viral membrane, and the extent of cotranslational and posttranslational processing. The structures of neuraminidases of other N2 strains will be discussed in the section on antigenic drift (Sect. 5.2.3).

2.2.1 Nucleotide Sequence of the Neuraminidase Gene of A/RI/5⁻/57

The structure of the A/RI/5⁻/57 neuraminidase gene was determined by cloning and sequencing double-stranded cDNA copies of the vRNA (ELLEMAN et al. 1982). Single-stranded cDNA was produced by reverse transcription of unfractionated vRNA, using the commercially available dodecanucleotide primer d(pAGCAAAAGCAGG) that is complementary to the 3′ end of all eight influenza virus gene segments (ROBERTSON 1979). These cDNA transcripts were made double stranded by 3′ end self-priming, the hairpin loops removed by S1 nuclease digestion, the double-stranded cDNA segments poly(dC)-tailed, inserted into the *Pst*I site of poly(dG)-tailed, linearized plasmid pBR322, and cloned into *Escherichia coli* strain RR1.

No full-length cloned copies were obtained, the longest cloned fragment being 84% of the gene. The sequence was determined for this fragment, which corresponds to the first 1235 nucleotides of the A/RI/5⁻/57 neuraminidase gene. The structure was completed by using specific restriction fragments (nucleotides 1055-1124 and 1125-1235) as primers for sequencing reactions on unfractionated vRNA and by back-sequencing on restriction fragment-primed, single-stranded cDNA template, using the commercially available dodecanucleotide d(pAGTAGAAACAAG) as primer (ELLEMAN et al. 1982). This oligonucleotide is equivalent to the constant 5′ end of all influenza vRNA segments (ROBERTSON 1979) and should only prime on cDNA copies that extend to the end of the gene. A second, shorter clone was also sequenced for 894 nucleotides, and T/C discrepancies were found between the two clones at nucleotides 253, 478, and 669. Dideoxy sequencing of vRNA using external restriction fragments as primers showed that these discrepancies all resulted from incorrect pairing of T with viral G (ELLEMAN et al. 1982). Such differences are not unexpected since the estimated error frequency for G/T mismatching with reverse transcriptase is approximately 1 in 1000 (BATTULA and LOEB 1974; GOPINATHAN et al. 1979; FIELDS and WINTER 1981).

The structure of the A/RI/5⁻/57 neuraminidase gene is shown in Table 1. Sequence is shown for the positive-strand cRNA, which has the same sense as mRNA. The sequence contains a total of 1467 nucleotides. Nineteen bases precede the initiating codon AUG; 1407 bases code without interruption for the N2 neuraminidase protein; there is a single stop codon UAA followed by 38 bases to the end of the gene. There is only one long reading frame in the cRNA sequence, and it codes for a primary gene product of 469 amino acids. The five alternative reading frames are blocked by numerous termination codons.

Previous sequence studies on the neuraminidase gene of A/RI/5⁻/57 have been limited to a 151-nucleotide partial sequence from the 3′ end of vRNA

(BLOK and AIR 1980). This partial sequence agrees with that shown in Table 1 except for nucleotides 103, 106, and 108. The cDNA clone shows ATC (Ile) CTG (Leu) GCA (Ala) in codons 28, 29, and 30 respectively, while the 3′ end partial sequence is ATG (Met) CTA (Leu) GTA (Val). Direct amino acid sequence analysis (BLOK et al. 1982) of detergent-released, intact neuraminidase from the closely related strain A/RI/5^+/57 (X7.F1) shows that the amino acid residues at positions 28–30 were Ile Leu Ala, in agreement with the sequence shown in Table 1.

2.2.2 Amino Acid Sequence of A/RI/5^-/57 Neuraminidase

The amino acid sequence deduced from the nucleic acid sequence of the A/RI/5^-/57 neuraminidase gene is also shown in Table 1. Initiation codons occur at nucleotides 20–22 and 89–91; both are in phase with the long, open reading frame. Automated amino acid sequence analysis of intact detergent-released neuraminidase from A/RI/5^+/57 showed that protein synthesis did start at the first AUG codon (BLOK et al. 1982). The sequence of the first 31 amino acid residues was identical to that predicted from the nucleic acid sequence (Table 1) except for the known Ile/Thr substitution at residue 7 (BLOK and AIR 1980).

This amino acid sequence data also showed that there was no proteolytic processing at the N-terminal end of the molecule. The initiating methionine residue was still present on this viral glycoprotein, and no hydrophobic signal peptide was removed during biosynthesis. This is in marked contrast to the proteolytic processing which accompanies hemagglutinin biosynthesis (AIR 1979; MCCAULEY et al. 1979). There was also no proteolytic processing at the C-terminal end of the molecule since detergent-released A/RI/5^+/57 neuraminidase had the predicted C-terminal cyanogen bromide dipeptide Pro Ile (BLOK et al. 1982).

The neuraminidase of A/RI/5^-/57 contains 469 amino acid residues for an apoprotein molecular weight of 52006 (Table 2). It contains eight methionine residues at positions 1, 24, 51, 160, 241, 307, 362, and 467; 22 half-cystine residues at positions 21, 42, 53, 78, 92, 124, 129, 175, 183, 193, 230, 232, 237, 278, 280, 289, 291, 318, 337, 417, 421, and 447; and eight potential glycosylation sites at asparagine residues 61, 69, 70, 86, 146, 200, 234, and 402.

2.2.3 Membrane Orientation and Pronase Cleavage Site

Neuraminidase, like hemagglutinin, is an integral membrane protein anchored in the viral membrane. Hemagglutinin is embedded in the viral membrane by a hydrophobic 26-residue section (amino acids 185–210) near the C-terminal end of HA2 (SKEHEL and WATERFIELD 1975; DOPHEIDE and WARD 1981). Similarly, neuraminidase is embedded in the viral membrane by a hydrophobic peptide region at the base of the stalk (LAVER and VALENTINE 1969), this stalk region being removed during proteolytic release of neuraminidase heads from virus particles (DRZENIEK et al. 1968; WRIGLEY et al. 1977). Since these protease-

Table 1. Gene sequence and the translated protein sequence of A/RI/5^-/57 neuraminidase (Elleman et al. 1982) and protein sequence of A/Tokyo/3/67 neuraminidase (Ward et al. 1982)[a]

										cRNA 5'-AGCAAAAGCAGGAGUGAAA	AUG	AAU	CCA	AAU	CAA	AAG	ACA	AUA	ACA	AUU	GGC	UCU	GUC	UCU	CUC
RI/5^-/57											Met	Asn	Pro	Asn	Gln	Lys	Thr	Ile	Thr	Ile	Gly	Ser	Val	Ser	Leu
											1				5					10					15
												100													
	ACC	AUU	GCA	ACA	GUA	UGC	UUC	CUC	AUG	CAG	AUU	GCC	AUC	CUG	GCA	ACU	ACU	GUG	ACA	UUG	CAU	UUU	AAA	CAA	CAU
RI/5^-/57	Thr	Ile	Ala	Thr	Val	Cys	Phe	Leu	Met	Gln	Ile	Ala	Ile	Leu	Ala	Thr	Thr	Val	Thr	Leu	His	Phe	Lys	Gln	His
					20					25					30					35					40
																					200				
	GAC	UGC	GAC	UCC	CCC	GCG	AGC	AAC	CAA	GUA	AUG	CCA	UGU	GAA	CCA	AUA	AUA	AUA	GAA	AGG	AAC	AUA	ACA	GAC	AUA
RI/5^-/57	Glu	Cys	Asp	Ser	Pro	Ala	Ser	Asn	Gln	Val	Met	Pro	Cys	Glu	Pro	Ile	Ile	Ile	Glu	Arg	Asn	Ile	Thr	Glu	Ile
					45					50					55					60					65
	GUG	UAU	UUG	AAU	AAC	ACC	ACC	AUA	GAG	AAA	GAG	AUU	UGC	CCC	GAA	GUA	GUG	GAA	UAC	AGA	AAU	UGG	UCA	AAG	CCG
RI/5^-/57	Val	Tyr	Leu	Asn	Asn	Thr	Thr	Ile	Glu	Lys	Glu	Ile	Cys	Pro	Glu	Val	Val	Glu	Tyr	Arg	Asn	Trp	Ser	Lys	Pro
Tokyo/67									Glu	Lys	Glu	Ile	Cys	Pro	**Lys**	Val	Val	Glu	Tyr	Arg	Asn	Trp	Ser	Lys	Pro
					70					75					80					85					90
				300																					
	CAA	UGU	CAA	AUU	ACA	GGA	UUU	GCA	CCU	UUU	UCU	AAG	GAC	AAU	UCA	AUC	CGG	CUU	UCU	GCU	GGU	GGG	GAC	AUU	UGG
RI/5^-/57	Gln	Cys	Gln	Ile	Thr	Gly	Phe	Ala	Pro	Phe	Ser	Lys	Asp	Asn	Ser	Ile	Arg	Leu	Ser	Ala	Gly	Gly	Asp	Ile	Trp
Tokyo/67	Gln	Cys	Gln	Ile	Thr	Gly	Phe	Ala	Pro	Phe	Ser	Lys	Asp	Asn	Ser	Ile	Arg	Leu	Ser	Ala	Gly	Gly	Asp	Ile	Trp
					95					100					105					110					115
												400													
	GUG	ACC	AGA	GAA	CCU	UAU	GUG	UCA	UGC	GAC	CCU	GGC	AAG	UGU	UAU	CAA	UUU	GCA	CUC	GGG	CAG	GGG	ACC	ACA	CUA
RI/5^-/57	Val	Thr	Arg	Glu	Pro	Tyr	Val	Ser	Cys	Asp	Pro	Gly	Lys	Cys	Tyr	Gln	Phe	Ala	Leu	Gly	Gln	Gly	Thr	Thr	Leu
Tokyo/67	Val	Thr	Arg	Glu	Pro	Tyr	Val	Ser	Cys	Asp	Pro	**Val**	Lys	Cys	Tyr	Gln	Phe	Ala	Leu	Gly	Gln	Gly	Thr	Thr	Leu
					120					125					130					135					140
																					500				
	GAC	AAC	AAA	CAU	UCA	AAU	GGC	ACA	AUA	CAU	GAU	AGA	AUC	CCU	CAC	CGA	ACC	CUA	UUA	AUG	AAU	GAG	UUG	GGU	GUU
RI/5^-/57	Asp	Asn	Lys	His	Ser	Asn	Gly	Thr	Ile	His	Asp	Arg	Ile	Pro	His	Arg	Thr	Leu	Leu	Met	Asn	Glu	Leu	Gly	Val
Tokyo/67	Asp	Asn	Lys	His	Ser	Asn	**Asp**	Thr	**Val**	His	Asp	Arg	Ile	Pro	His	Arg	Thr	Leu	Leu	Met	Asn	Glu	Leu	Gly	Val
					145					150					155					160					165
	CCA	UUU	CAU	UUA	GGA	ACC	AAA	CAA	GUG	UGU	GUA	GCA	UGG	UCC	AGC	UCA	AGU	UGU	CAC	GAU	GGA	AAA	GCA	UGG	UUG
RI/5^-/57	Pro	Phe	His	Leu	Gly	Thr	Lys	Gln	Val	Cys	Val	Ala	Trp	Ser	Ser	Ser	Ser	Cys	His	Asp	Gly	Lys	Ala	Trp	Leu
Tokyo/67	Pro	Phe	His	Leu	Gly	Thr	**Arg**	Gln	Val	Cys	**Ile**	Ala	Trp	Ser	Ser	Ser	Ser	Cys	His	Asp	Gly	Lys	Ala	Trp	Leu
					170					175					180					185					190
				600																					
	CAU	GUU	UGU	GUC	ACU	GGG	GAU	GAU	AGA	AAU	GCG	ACU	GCU	AGC	UUC	AUU	UAU	GAC	GGG	AGG	CUU	GUG	GAC	AGU	AUU
RI/5^-/57	His	Val	Cys	Val	Thr	Gly	Asp	Asp	Arg	Asn	Ala	Thr	Ala	Ser	Phe	Ile	Tyr	Asp	Gly	Arg	Leu	Val	Asp	Ser	Ile
Tokyo/67	His	Val	Cys	**Ile**	Thr	Gly	Asp	Asp	**Asp**	Asn	Ala	Thr	Ala	Ser	Phe	Ile	Tyr	Asp	Gly	Arg	Leu	Val	Asp	Ser	Ile
					195					200					205					210					215
												700													
	GCU	UCA	UGG	UCU	CAA	AAU	AUC	CUC	AGG	ACC	CAG	GAG	UCG	GAA	UGC	GUU	UGU	AUC	AAU	GGG	ACU	UGC	ACA	GUA	GUA
RI/5^-/57	Gly	Ser	Trp	Ser	Gln	Asn	Ile	Leu	Arg	Thr	Gln	Glu	Ser	Glu	Cys	Val	Cys	Ile	Asn	Gly	Thr	Cys	Thr	Val	Val
Tokyo/67	Gly	Ser	Trp	Ser	Gln	Asn	Ile	Leu	Arg	Thr	Gln	Glu	Ser	Glu	Cys	Val	Cys	Ile	Asn	Gly	Thr	Cys	Thr	Val	Val
					220					225					230					235					240
																					800				
	AUG	ACU	GAU	GGA	AGU	GCA	UCA	GGA	AGA	GCC	GAU	ACU	AGA	AUA	CUA	UUC	AUU	AAA	GAG	GGG	AAA	AUU	GUC	CAU	AUU
RI/5^-/57	Met	Thr	Asp	Gly	Ser	Ala	Ser	Gly	Arg	Ala	Asp	Thr	Arg	Ile	Leu	Phe	Ile	Lys	Glu	Gly	Lys	Ile	Val	His	Ile
Tokyo/67	Met	Thr	Asp	Gly	Ser	Ala	Ser	Gly	Arg	Ala	Asp	Thr	Arg	Ile	Leu	Phe	Ile	**Glu**	Glu	Gly	Lys	Ile	Val	His	Ile
					245					250					255					260					265
	AGC	CCA	UUG	UCA	GGA	AGU	GCU	CAG	CAU	AUA	GAG	GAG	UGU	UCC	UGU	UAC	CCU	CGA	UAU	CCU	GAC	GUA	AGA	UGU	AUC
RI/5^-/57	Ser	Pro	Leu	Ser	Gly	Ser	Ala	Gln	His	Ile	Glu	Glu	Cys	Ser	Cys	Tyr	Pro	Arg	Tyr	Pro	Asp	Val	Arg	Cys	Ile
Tokyo/67	Ser	Pro	Leu	**Ala**	Gly	Ser	Ala	Gln	His	**Val**	Glu	Glu	Cys	Ser	Cys	Tyr	Pro	Arg	Tyr	Pro	**Gly**	Val	Arg	Cys	Ile
					270					275					280					285					290
				900																					
	UGC	AGA	GAC	AAC	UGG	AAA	GGC	UCU	AAU	AGG	CCC	GUU	AUA	GAC	AUA	AAU	AUG	GAA	GAU	UAU	AGC	AUU	GAU	UCC	AGU
RI/5^-/57	Cys	Arg	Asp	Asn	Trp	Lys	Gly	Ser	Asn	Arg	Pro	Val	Ile	Asp	Ile	Asn	Met	Glu	Asp	Tyr	Ser	Ile	Asp	Ser	Ser
Tokyo/67	Cys	Arg	Asp	Asn	Trp	Lys	Gly	Ser	Asn	Arg	Pro	Val	**Val**	Asp	Ile	Asn	Met	Glu	Asp	Tyr	Ser	Ile	Asp	Ser	Ser
					295					300					305					310					315
												1000													
	UAU	GUG	UGC	UCA	GGC	CUU	GUU	GGC	GAC	ACA	CCC	AGG	AAC	GAC	GAC	AGC	UCU	AGC	AAU	AGC	AAU	UGC	AGG	GAU	CCU
RI/5^-/57	Tyr	Val	Cys	Ser	Gly	Leu	Val	Gly	Asp	Thr	Pro	Arg	Asn	Asp	Asp	Ser	Ser	Ser	Asn	Ser	Asn	Cys	Arg	Asp	Pro
Tokyo/67	Tyr	Val	Cys	Ser	Gly	Leu	Val	Gly	Asp	Thr	Pro	Arg	Asn	Asp	Asp	**Arg**	Ser	Ser	Asn	Ser	Asn	Cys	Arg	**Asn**	Pro
					320					325					330					335					340
																					1100				
	AAC	AAU	GAA	AGA	GGG	AAU	CCA	GGA	GUC	AAA	GGC	UGG	GCC	UUU	GAC	AAU	GGA	GAU	GAU	GUA	UGG	AUG	GGA	AGA	ACA
RI/5^-/57	Asn	Asn	Glu	Arg	Gly	Asn	Pro	Gly	Val	Lys	Gly	Trp	Ala	Phe	Asp	Asn	Gly	Asp	Asp	Val	Trp	Met	Gly	Arg	Thr
Tokyo/67	Asn	Asn	Glu	Arg	Gly	**Thr**	**Gln**	Gly	Val	Lys	Gly	Trp	Ala	Phe	Asp	Asn	Gly	**Asn**	Asp	**Leu**	Trp	Met	Gly	Arg	Thr
					345					350					355					360					365
	AUC	AAC	AAA	GAA	UCA	CGC	UCA	GGU	UAU	GAA	ACU	UUC	AAA	GUC	AUU	GGU	GGU	UGG	UCC	ACA	CCU	AAU	UCC	AAA	UCG
RI/5^-/57	Ile	Asn	Lys	Glu	Ser	Arg	Ser	Gly	Tyr	Glu	Thr	Phe	Lys	Val	Ile	Gly	Gly	Trp	Ser	Thr	Pro	Asn	Ser	Lys	Ser
Tokyo/67	Ile	**Ser**	Lys	**Asp**	**Leu**	Arg	Ser	Gly	Tyr	Glu	Thr	Phe	Lys	Val	Ile	Gly	Gly	Trp	Ser	Thr	Pro	Asn	Ser	Lys	Ser
					370					375					380					385					390

[a] The amino acid substitutions that have accumulated since 1957 are shaded

Table 1 (continued)

```
                                                1200
                CAG GUC AAU AGA CAG GUC AUA GUU GAC AAC AAU AAU UGG UCU GGU UAC UCU GGU AUU UUC UCU GUU GAG CGC AAA
RI/5⁻/57        Gln Val Asn Arg Gln Val Ile Val Asp Asn Asn Asn Trp Ser Gly Tyr Ser Gly Ile Phe Ser Val Glu Gly Lys
Tokyo/67        Gln Ile Asn Arg Gln Val Ile Val Asp Ser Asp Asn Arg Ser Gly Tyr Ser Gly Ile Phe Ser Val Glu Gly Lys
                            395                 400                 405                 410                 415

                                                                1300
                AGC UGC AUC AAU AGG UGC UUU UAU GUG GAG UUG AUA AGG GGA ACG CCA CAG GAG ACU AGA GUA UGG UGG ACC UCA
RI/5⁻/57        Ser Cys Ile Asn Arg Cys Phe Tyr Val Glu Leu Ile Arg Gly Arg Pro Gln Glu Thr Arg Val Trp Trp Thr Ser
Tokyo/67        Ser Cys Ile Asn Arg Cys Phe Tyr Val Glu Leu Ile Arg Gly Arg Lys Gln Glu Thr Arg Val Trp Trp Thr Ser
                            420                 425                 430                 435                 440

                                                                                                1400
                AAC AGU AUU GUU GUG UUU UGU GGC ACU UCA GGU ACU UAU GGA ACA GGC UCA UGG CCU GAU GGG GCG AAC AUC AAU
RI/5⁻/57        Asn Ser Ile Val Val Phe Cys Gly Thr Ser Gly Thr Tyr Gly Thr Gly Ser Trp Pro Asp Gly Ala Asn Ile Asn
Tokyo/67        Asn Ser Ile Val Val Phe Cys Gly Thr Ser Gly Thr Tyr Gly Thr Gly Ser Trp Pro Asp Gly Ala Asn Ile Asn
                            445                 450                 455                 460                 465

                UUC AUG CCU AUA UAA CGUUUCGCAAUUUUAGAAAAAAACUCCUUUGUUUCUACU-3'
RI/5⁻/57        Phe Met Pro Ile
Tokyo/67        Phe Met Pro Ile
                        469
```

released heads retain the enzymatic and antigenic properties of neuraminidase and are the portion of the peplomer used for X-ray crystallographic structure analyses, it is important to establish which portion of the neuraminidase sequence corresponds to that of the free heads and which corresponds to that associated with the stalk and membrane associated region.

An examination of the data in Table 1 shows that there are two long hydrophobic stretches in neuraminidase, a 29-residue section near the N-terminus from amino acid 7–35 and a 24-residue section near the C-terminus from residues 436–459. Both are similar in length to the membrane-embedded region of hemagglutinin (DOPHEIDE and WARD 1981).

The membrane orientation of neuraminidase was established by direct amino acid sequence analysis of pronase-released heads derived from the late Asian influenza variant A/Tokyo/3/67 (WARD et al. 1982). The sequence data (Table 1) show that pronase cuts N2 neuraminidase in at least two positions to give a ragged N-terminal end starting at residue 74 or 77. The sequence then extends to the C-terminal Ile residue at position 469, indicating that pronase neither produces additional internal cleavages nor modifies the C-terminal end of this neuraminidase.

Table 2. Amino acid composition of A/RI/5⁻/57 neuraminidase (ELLEMAN et al. 1982)

Amino acid	Residues/mole A/RI/5⁻/57	Amino acid	Residues/mole A/RI/5⁻/57	Amino acid	Residues/mole A/RI/5⁻/57
Asn	34	Gly	39	Tyr	13
Asp	27	Ala	16	Phe	14
Thr	33	Cys	22	Trp	13
Ser	47	Val	34	His	10
Gln	15	Met	8	Lys	17
Glu	22	Ile	39	Arg	25
Pro	22	Leu	19		

Total residues 469
Apoprotein molecular weight 52006 daltons

Thus, the membrane-embedded stalk region of influenza neuraminidase comes from the N-terminal end of the molecule. Residues 7–35 are the favored hydrophobic transmembrane region, residues 36–73 comprise the neuraminidase stalk, and residues 74–469 make up the heads. Of the 22 half-cystine residues in N2 neuraminidase, one (residue 21) occurs in the proposed transmembrane sequence, and three (residues 42, 53, and 78) occur in the stalk region.

FIELDS et al. (1981) were the first to suggest that influenza neuraminidase may be embedded in the viral membrane at its N-terminal rather than C-terminal end. Cytochrome $P\text{-}450_{LM2}$ is also membrane-bound via an unprocessed N-terminal hydrophobic sequence (HAUGEN et al. 1970; CLEMENT 1983).

2.2.4 Oligosaccharide Distribution and Properties

The nucleic acid sequence of A/RI/5$^-$/57 NA showed there are eight potential glycosylation sites in this protein (Table 1). Three occur in the stalk region of the molecule at asparagine residues 61, 69, and 70, and five occur in the head at asparagine residues 86, 146, 200, 234, and 402. Amino acid sequence studies on Tokyo/67 heads have shown that only four of these five potential glycosylation sites contain carbohydrate (WARD et al. 1982, 1983b). The asparagine residue at position 402 is not glycosylated.

The monosaccharide compositions of the four oligosaccharide side chains on pronase-released heads of A/Tokyo/3/67 neuraminidase are shown in Table 3. Although stereochemical models show that any amino acid could occur in the X position of the glycosylation sequence Asn-X-$\begin{matrix}\text{Ser}\\\text{Thr}\end{matrix}$ (MARSHALL 1972), tryptophan, found here at residue 87, and aspartic acid, found at 147, occur rarely in such sequences (WAGH and BAHL 1981; STRUCK and LENNARZ 1980). As shown in Table 3, the oligosaccharide units at Asn_{86} and Asn_{200} contain approximately two residues of *N*-acetylglucosamine and five residues of mannose and are essentially of the oligomannoside type (simple or type II). The other two carbohydrate units at Asn_{146} and Asn_{234} are of the *N*-acetyllactosamine type (complex, type I) and contain the four sugars *N*-acetylglucosamine, mannose, galactose, and fucose previously reported for the complex carbohydrate units on the H1 (BASAK et al. 1981), H2 (BROWN et al. 1982), and H3 (WARD et al. 1980; WARD and DOPHEIDE 1981a, 1981b) hemagglutinins. How-

Table 3. Carbohydrate composition of the four oligosaccharide units on A/Tokyo/3/67 neuraminidase heads (WARD et al. 1983b)

Peptide	Asparagine residue	Sugar composition (residues/mole)					Carbohydrate type	Host antigen cross reaction
		GlcNAc	GalNac	Man	Gal	Fuc		
CN1.T3	86	1.9	–	5.0	–	–	Simple	No
CN1.T8	146	3.1	1.4	1.6	0.3	1.1	Complex	Yes
CN2.C4	200	1.6	–	4.6	–	–	Simple	No
CN2.T5.Th3	234	2.4	–	3.9	0.7	0.3	Complex	No

ever, the complex unit at Asn_{146} differs from that at Asn_{234} and from those on the different hemagglutinin polypeptides in that it contains a significant amount of *N*-acetylgalactosamine.

N-acetylgalactosamine has not been found in other influenza virus glycoproteins. Carbohydrate analyses of whole virus (ADA and GOTTSCHALK 1956), whole hemagglutinin (WARD and DOPHEIDE 1976; COLLINS and KNIGHT 1978; WATERFIELD et al. 1979), the glycopeptides derived by enzymatic digestion of hemagglutinin molecules (WARD et al. 1980; WARD and DOPHEIDE 1981a, b; BASAK et al. 1981), and whole B/Lee/40 neuraminidase (ALLEN et al. 1977) did not reveal the presence of *N*-acetylgalactosamine. The presence of *N*-acetyl-galactosamine in *N*-glycosidic carbohydrate units is rare but has been reported (see WAGH and BAHL 1981).

Antigenic analyses of each of the four isolated glycopeptides in Tokyo/67 neuraminidase show that only the carbohydrate unit at Asn_{146} is antigenically related to chick embryo "host antigen" (WARD et al. 1982, 1983b). Nothing is known about the number or nature of the carbohydrate units attached to the stalk but, by analogy to H2 hemagglutinin (WATERFIELD et al. 1980; BROWN et al. 1982), one might expect only one of the asparagine residues in the doublet at positions 69 and 70 to be glycosylated.

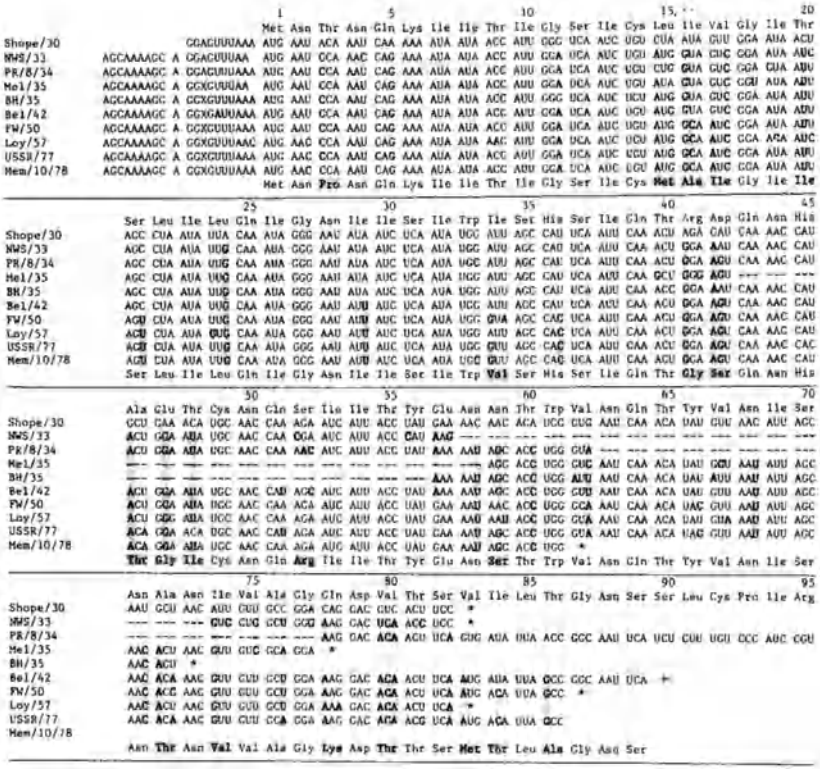
2.3 Type A N1 Neuraminidase

The primary structure of N1 neuraminidase has been examined by partial nucleotide sequences from the 3′ end of the vRNA (BLOK and AIR 1980, 1982a), complete nucleotide sequences of A/PR/8/34 (FIELDS et al. 1981) and NWS/33 (HITI and NAYAK 1982), and comparative peptide map amino acid analyses for A/Bel/42 (BLOK et al. 1982; BLOK and AIR 1982a).

2.3.1 Stalk Region Deletions in Some N1 Neuraminidases

Partial sequences (206–287 nucleotides) obtained from the 3′ end of the neuraminidase gene in 12 N1 strains isolated in the periods 1930–1957 and 1976–1978 revealed that four strains, NWS/33, PR/8/34, BH/35, and Mel/35, had deletions of 48, 45, 33, and 48 nucleotides respectively (BLOK and AIR 1982a). As shown in Table 4, the positions of these deletions were different in each of the four strains, but all occurred in that region of the sequence (amino acids 43 to 77) which corresponds to the stalk of the neuraminidase. These deletions of 11–16 amino acids would thus not be expected to affect the enzymatic or antigenic characteristics of this protein significantly (BLOK and AIR 1980, 1982c). Sequence studies of subgenomic influenza RNA molecules suggest that they may be generated by a jumping of the viral polymerase between adjacent segments of the RNA template in the three-dimensional ribonucleo-protein structure (JENNINGS et al. 1983), and the block deletions in the neuraminidase gene are presumably generated by a similar mechanism (JENNINGS et al. 1983; FIELDS and WINTER 1982).

Table 4. Partial gene sequences from the 3′ end of N1 neuraminidase showing deletions in regions corresponding to the stalk of the protein (BLOK and AIR 1982c)[a]



[a] Nucleotide and amino acid sequence differences are shaded

2.3.2 *Nucleotide Sequence of the Neuraminidase Gene of A/PR/8/34 (N1)*

For the structural determination of the neuraminidase gene of A/PR/8/34, all eight segments of unfractionated vRNA were polyadenylated and converted to double-stranded cDNA by oligo(dT)$_{12-18}$ priming and either 3′ end loopback or primed, second-strand synthesis with the 13 nucleotide primer d(pAGTAGAAACAAGG) which is complementary to the 3′ end of full-length, single-strand cDNA (FIELDS et al. 1981). The nucleotide sequence was obtained by: (a) shotgun-cloning restriction fragments derived from the whole genome; (b) cloning restriction fragments derived from the cDNA copies of the isolated neuraminidase gene segment; and (c) using restriction fragments, derived from the replicative form of selected M13 clones, as primers for dideoxy sequencing on a total vRNA template (FIELDS et al. 1981).

The structure of the A/PR/8/34 neuraminidase gene is shown in Table 5. It contains 1413 nucleotides. Twenty bases precede the first AUG codon, 1362 bases code without interruption for the N1 neuraminidase protein; there is

Table 5. Gene sequence and translated protein sequence for the N1 neuraminidase PR/8/34 (FIELDS et al. 1981)

```
                5'-AGCGAAAGCAGGGGUUU AAA AUG AAU CCA AAU CAG AAA AUA AUA ACC AUU GGA UCA AUC UGU CUG
                                       Met Asn Pro Asn Gln Lys Ile Ile Thr Ile Gly Ser Ile Cys Leu
                                         1                   5                  10                  15
                                                        100
GUA GUC GGA CUA AUU AGC CUA AUA UUG CAA AUA GGG AAU AUA AUC UCA AUA UGG AUU AGC CAU UCA AUU CAA ACU
Val Val Gly Leu Ile Ser Leu Ile Leu Gln Ile Gly Asn Ile Ile Ser Ile Trp Ile Ser His Ser Ile Gln Thr
                  20                  25                  30                  35                  40
                                                                          200
GGA AGU CAA AAC CAU ACU GGA AUA UGC AAC CAA AAC AUC AUU ACC UAU AAA AAU AGC ACC UGG GUA AAG GAC ACA
Gly Ser Gln Asn His Thr Gly Ile Cys Asn Gln Asn Ile Ile Thr Tyr Lys Asn Ser Thr Trp Val Lys Asp Thr
                  45                  50                  55                  60                  65

ACU UCA GUG AUA UUA ACC GGC AAU UCA UCU CUU UGU CCC AUC CGU GGG UGG GCU AUA UAC AGC AAA GAC AAU AGC
Thr Ser Val Ile Leu Thr Gly Asn Ser Ser Leu Cys Pro Ile Arg Gly Trp Ala Ile Tyr Ser Lys Asp Asn Ser
                  70                  75                  80                  85                  90
          300
AUA AGA AUU GCU UCC AAA GGA GAC CUU UUU GUC AUA AGA GAG CCC UUU AUU UCA UGU UCU CAC UUG GAA UGC AGG
Ile Arg Ile Gly Ser Lys Gly Asp Val Phe Val Ile Arg Glu Pro Phe Ile Ser Cys Ser His Leu Glu Cys Arg
                  95                 100                 105                 110                 115
                                                        400
ACC UUU UUU CUG ACC CAA GGU GCC UUA CUG AAU GAC AGG CAU UCA AAU GGG ACU GUU AAG GAC AGA AGC CCU UAU
Thr Phe Phe Leu Thr Gln Gly Ala Leu Leu Asn Asp Arg His Ser Asn Gly Thr Val Lys Asp Arg Ser Pro Tyr
                 120                 125                 130                 135                 140
                                                                          500
AGG GCC UUA AUG AGC UGC CCU GUC GGU GAA GCU CCG UCC CCG UAC AAU UCA AGA UUU GAA UCG GUU GCU UGG UCA
Arg Ala Leu Met Ser Cys Pro Val Gly Glu Ala Pro Ser Pro Tyr Asn Ser Arg Phe Glu Ser Val Ala Trp Ser
                 145                 150                 155                 160                 165

GCA AGU GCA UGU CAU GAU GGC AUG GGC UGG CUA ACA AUC GGA AUU UCA GGU CCA GAU AAU GGA GCA GUG GCU GUA
Ala Ser Ala Cys His Asp Gly Met Gly Trp Leu Thr Ile Gly Ile Ser Gly Pro Asp Asn Gly Ala Val Ala Val
                 170                 175                 180                 185                 190
          600
UUA AAA UAC AAC GGC AUA AUA ACU GAA ACC AUA AAA AGU UGG AGG AAG AAA AUA UUG AGG ACA CAA GAG UCU GAA
Leu Lys Tyr Asn Gly Ile Ile Thr Glu Thr Ile Lys Ser Trp Arg Lys Lys Ile Leu Arg Thr Gln Glu Ser Glu
                 195                 200                 205                 210                 215
                                                        700
UGU GCC UGU GUA AAU GGU UCA UGU UUU ACU AUA AUG ACU GAU GGC CCG AGU GAU GGG CUG GCC UCG UAC AAA AUU
Cys Ala Cys Val Asn Gly Ser Cys Phe Thr Ile Met Thr Asp Gly Pro Ser Asp Gly Leu Ala Ser Tyr Lys Ile
                 220                 225                 230                 235                 240
                                                                          800
UUC AAG AUC GAA AAG GGG AAG CUU ACU AAA UCA AUA GAG UUG AAU GCA CCU AAU UCU CAC UAU GAG GAA UGU UCC
Phe Lys Ile Glu Lys Gly Lys Val Thr Lys Ser Ile Glu Leu Asn Ala Pro Asn Ser His Tyr Glu Glu Cys Ser
                 245                 250                 255                 260                 265

UGU UAC CCU GAU ACC GGC AAA GUG AUG UGU GUG UGC AGA GAC AAU UGG CAU GGU UCG AAC CGG CCA UGG GUG UCU
Cys Tyr Pro Asp Thr Gly Lys Val Met Cys Val Cys Arg Asp Asn Trp His Gly Ser Asn Arg Pro Trp Val Ser
                 270                 275                 280                 285                 290
          900
UUC GAU CAA AAC CUG GAU UAU CAA AUA GGA UAC AUC UGC AGU GGG GUU UUC GGU GAC AAC CCG CGU CCC AAA GAU
Phe Asp Gln Asn Leu Asp Tyr Gln Ile Gly Tyr Ile Cys Ser Gly Val Phe Gly Asp Asn Pro Arg Pro Lys Asp
                 295                 300                 305                 310                 315
                                                       1000
GGA ACA GGC AGC UGU GGU CCA GUG UAU GUU GAU GGA GCA AAC GGA GUA AAG GGA UUU UCA UAU AGG UAU GGU AAU
Gly Thr Gly Ser Cys Gly Pro Val Tyr Val Asp Gly Ala Asn Gly Val Lys Gly Phe Ser Tyr Arg Tyr Gly Asn
                 320                 325                 330                 335                 340
                                                                         1100
GGU GUU UGG AUA GGA AGG ACC AAA AGU CAC AGU UCC AGA CAU GGG UUU GAG AUG AUU UGG GAU CCU AAU GGA UGG
Gly Val Trp Ile Gly Arg Thr Lys Ser His Ser Ser Arg His Gly Phe Glu Met Ile Trp Asp Pro Asn Gly Trp
                 345                 350                 355                 360                 365

ACA GAG ACU GAU AGU AAG UUC UCU GUG AGG CAA GAU GUU GUG GCA AUG ACU GAU UGG UCA GGG UAU AGC GGG AGU
Thr Glu Thr Asp Ser Lys Phe Ser Val Arg Gln Asp Val Val Ala Met Thr Asp Thr Ser Gly Tyr Ser Gly Ser
                 370                 375                 380                 385                 390
          1200
UUC GUU CAA CAU CCU GAG CUA ACA GGG CUA GAC UGU AUA AGG CCG UGC UUC UGG GUU GAA UUA AUC AGG GGA CGA
Phe Val Gln His Pro Glu Leu Thr Gly Leu Asp Cys Ile Arg Pro Cys Phe Trp Val Glu Leu Ile Arg Gly Arg
                 395                 400                 405                 410                 415
                                                       1300
CCU AAA GAA AAA ACA AUC UGG ACU AGU GCG AGC AGC AUU UCU UUU UGU GGC GUG AAU AGU GAU ACU GUA GAU UGG
Pro Lys Glu Lys Thr Ile Trp Thr Ser Ala Ser Ser Ile Ser Phe Cys Gly Val Asn Ser Asp Thr Val Asp Trp
                 420                 425                 430                 435                 440
                                                                         1400
UCU UGG CCA GAC GGU GCU GAG UUG CCA UUC ACC AUU GAC AAG UAG UCUGUUCAAAAAACUCCUUGUUUCUACU-3'
Ser Trp Pro Asp Gly Ala Glu Leu Pro Phe Thr Ile Asp Lys
                 445                 450
```

then a single stop codon UAG followed by 28 nucleotides to the end of the gene. There is only one long, open reading frame in the cRNA, and it codes for a protein of 454 amino acids. The alternative reading frames are blocked by numerous termination codons.

Previous sequencing of the A/PR/8/34 neuraminidase gene was limited to the first 275 nucleotides at the 5′ end of the cRNA (BLOK and AIR 1980, 1982a) and the 22 nucleotides corresponding to the 3′ end of the cRNA (DESSELBERGER et al. 1980). These sequences differed from that shown in Table 6 at nucleotides 4 (A for G) and 13 (A for G) in the 5′ noncoding region of the cRNA as well as at codon 173 (CAT for CAA), which produces an amino acid change of His for Gln (BLOK and AIR 1980, 1982a). As pointed out in Sect. 2.3.1, the A/PR/8/34 neuraminidase gene contains a 45-nucleotide deletion between codons 62 and 63 when compared with other N1 strains (BLOK and AIR 1982a).

2.3.3 Amino Acid Sequence and Oligosaccharide Distribution

The predicted amino acid sequence for A/PR/8/34 neuraminidase is shown in Table 5. It contains 454 amino acids for an apoprotein molecular weight of 50087 (Table 6). As pointed out by FIELDS et al. (1981), there is only one long hydrophobic sequence from residues 8 to 35, and this is involved in membrane attachment. The second long hydrophobic sequence equivalent to that near the C-terminal end of N2 neuraminidase (Sect. 2.2.3) extends for only 16 residues in PR/8/34, being interrupted by acidic residues at positions 436, 439, 444, and 447.

There are seven methionine residues at positions 1, 144, 173, 227, 274, 358, and 381; 19 half-cystine residues; and five potential glycosylation sites. Of the 19 half-cystines, one is found in the transmembrane region (residue 14), one is found in the stalk region (residue 49), and 17 are found in the head (residues 77, 109, 114, 146, 169, 216, 218, 223, 264, 266, 275, 277, 303, 320, 402, 406, and 421). Two of the glycosylation sites occur in the stalk region of A/PR/8/35 (residues 44 and 58), while three occur in the heads (residues 73, 131, and 220).

Complex and simple oligosaccharide units are attached to the N1 PR/8/34 neuraminidase molecule (SCHWARZ and KLENK 1981); all attachment sites probably carry carbohydrate. They occur in positions similar to those in N2 neur-

Table 6. Amino acid composition of A/PR/8/34 (N1) neuraminidase (FIELDS et al. 1981)

Amino acid	Residues/mole A/PR/8/34	Amino acid	Residues/mole A/PR/8/34	Amino acid	Residues/mole A/PR/8/34
Asn	25	Gly	44	Tyr	14
Asp	25	Ala	16	Phe	16
Thr	29	Cys	19	Trp	16
Ser	51	Val	28	His	10
Gln	11	Met	7	Lys	23
Glu	17	Ile	41	Arg	20
Pro	21	Leu	21		

Total residues 454
Apoprotein molecular weight 50087 daltons

aminidase, they contain tripeptide sequences that can be glycosylated, and they all occur in positions that would be expected to be on the surface of the molecule (see Sect. 4).

2.4 Type B Neuraminidase

The primary structure of B/Lee/40 neuraminidase is derived entirely from the gene sequence of SHAW et al. (1982).

2.4.1 Nucleotide Sequence of the Neuraminidase Gene of B/Lee/40

The approach to the preparation and cloning of the B/Lee/40 neuraminidase gene was based on that described by LAI et al. (1980), where double-stranded cDNA was produced by hybridization of full-length cDNA copies of the minus-sense vRNA and the positive-sense mRNA. The single-stranded cDNA copies were prepared by oligo$(dT)_{12-18}$-primed reverse transcription of influenza B virus-specific mRNA or polyA-tailed, unfractionated vRNA (BREIDIS and LAMB 1982). Both (+) and (−) single-stranded cDNA copies were poly(dC)-tailed in separate reactions, size-selected on 4% acrylamide gels, and hybridized before insertion into oligo(dG)-tailed, *Pst*I-cleaved pBR322 for cloning into *E. coli* strain HB101.

The structure of the B/Lee/40 neuraminidase gene is shown in Table 7. It contains 1557 nucleotides and is 90 nucleotides longer than the type A N2 neuraminidase gene. A long open reading frame extends from nucleotide 54 to 1451, and there are two adjacent termination codons UAA and UAG at bases 1452–1457, followed by another 100 bases to the end of the gene. The major open reading frame codes for a primary gene product of 466 amino acids, which is three fewer than the type A N2 protein.

It was also observed (SHAW et al. 1982) that there is a second shorter reading frame extending from nucleotide 47 to 346 which overlaps the main reading frame for most of its length and could code for a protein of 100 amino acid residues (molecular weight 11240). Evidence for this minor gene product in infected cells has recently been found (LAMB et al. 1983), but its function remains unknown. In this respect the B/Lee/40 neuraminidase gene (segment 6) resembles the gene segments 7 (ALLEN et al. 1980; WINTER and FIELDS 1980; LAMB and LAI 1981; LAMB and CHOPPIN 1981; BREIDIS et al. 1982) and 8 (LAMB and CHOPPIN 1979; INGLIS et al. 1979; LAMB and LAI 1980; BREIDIS and LAMB 1982) of both Type A and Type B influenza viruses in coding for more than one protein on overlapping reading frames.

2.4.2 Amino Acid Sequence and Oligosaccharide Distribution

The predicted amino acid sequence of B/Lee/40 neuraminidase is shown in Table 7. It consists of 466 amino acids for an apoprotein molecular weight of 51441. The amino acid composition of B/Lee/40 neuraminidase, based on

Table 7. Gene sequence and translated protein sequence for B/Lee/40 neuraminidase (SHAW et al. 1982)

```
5'-AGCAGAAGCAGAGCAUAUUCUUAGAACUGAAGUGAACAGGCCAAAAAUGAACA AUG CUA CCU UCA ACU GUA CAA ACA UUA ACC
                                                          Met Leu Pro Ser Thr Val Gln Thr Leu Thr
                                                                                    5                  10

                                100
CUA UUA CUC ACA UCA GGG GGA GUA UUA UUA UCA CUA UAU GUG UCA GCC UCA UUG UCA UAC UUA UUG UAU UCG GAU
Leu Leu Leu Thr Ser Gly Gly Val Leu Leu Ser Leu Tyr Val Ser Ala Ser Leu Ser Tyr Leu Leu Tyr Ser Asp
                  15                  20                  25                  30                  35

                                                      200
GUA UUG CUA AAA UUU UCA UCA ACA AAA ACA ACU GCA CCA ACA AUG UCA UUA GAG UGC ACA AAC GCA UCA AAU GCC
Val Leu Leu Lys Phe Ser Ser Thr Lys Thr Thr Ala Pro Thr Met Ser Leu Glu Cys Thr Asn Ala Ser Asn Ala
                  40                  45                  50                  55                  60

                                                                                          300
CAG ACU GUG AAC CAU UCU GCA ACA AAA GAG AUG ACA UUU CCA CCC CCA GAG CCG GAG UGG ACA UAC CCU CGU UUA
Gln Thr Val Asn His Ser Ala Thr Lys Glu Met Thr Phe Pro Pro Pro Glu Pro Glu Trp Thr Tyr Pro Arg Leu
                  65                  70                  75                  80                  85

UCU UGC CAG GGC UCA ACC UUU CAG AAG GCA CUC CUA AUU AGC CCU CAU AGG UUC GGA GAG AUC AAA GGA AAC UCA
Ser Cys Gln Gly Ser Thr Phe Gln Lys Ala Leu Leu Ile Ser Pro His Arg Phe Gly Glu Ile Lys Gly Asn Ser
                  90                  95                 100                 105                 110

                          400
GCU CCC UUG AUA AUA AGA GAA CCU UUU GUU GCU UGU GGA CCA AAA GAA UGC AGA CAC UUU GCU CUG ACC CAU UAU
Ala Pro Leu Ile Ile Arg Glu Pro Phe Val Ala Cys Gly Pro Lys Glu Cys Arg His Phe Ala Leu Thr His Tyr
                 115                 120                 125                 130                 135

                                                    500
GCA GCU CAG CCG GGG GGA UAC UAC AAU GGA ACA AGA AAG GAC AGA AAC AAG CUG AGG CAU CUA GUA UCA GUC AAA
Ala Ala Gln Pro Gly Gly Tyr Tyr Asn Gly Thr Arg Lys Asp Arg Asn Lys Leu Arg His Leu Val Ser Val Lys
                 140                 145                 150                 155                 160

                                                                                    600
UUG GGA AAA AUC CCA ACU GUG GAA AAC UCC AUU UUC CAC AUG GCA GCU UGG AGC GGA UCC GCA UGC CAU GAU GGU
Leu Gly Lys Ile Pro Thr Val Glu Asn Ser Ile Phe His Met Ala Ala Trp Ser Gly Ser Ala Cys His Asp Gly
                 165                 170                 175                 180                 185

AGA GAA UGG ACA UAU AUC GGA GUU GAU GGU CCU GAC AAU GAU GCA UUG GUC AAA AUA AAA UAU GGA GAA GCA UAU
Arg Glu Trp Thr Tyr Ile Gly Val Asp Gly Pro Asp Asn Asp Ala Leu Val Lys Ile Lys Tyr Gly Glu Ala Tyr
                 190                 195                 200                 205                 210

                          700
ACU GAC ACA UAU CAU UCC UAU GCA CAC AAC AUC CUA AGA ACA CAA GAA AGU GCC UGC AAU UGC AUC GGG GGA GAU
Thr Asp Thr Tyr His Ser Tyr Ala His Asn Ile Leu Arg Thr Gln Glu Ser Ala Cys Asn Cys Ile Gly Gly Asp
                 215                 220                 225                 230                 235

                                                    800
UGU UAU CUU AUG AUA ACA GAC GGC UCA GCU UCA GGA AUU AGU AAA UGC AGA UUU CUU AAA AUU AGA GAG GGU CGA
Cys Tyr Leu Met Ile Thr Asp Gly Ser Ala Ser Gly Ile Ser Lys Cys Arg Phe Leu Lys Ile Arg Glu Gly Arg
                 240                 245                 250                 255                 260

                                                                                    900
AUA AUA AAA GAA AUA CUU CCA ACA GGA AGA GUG GAG CAC ACU GAA GAG UGC ACA UGC GGG UUC GCC AGC AAU AAA
Ile Ile Lys Glu Ile Leu Pro Thr Gly Arg Val Glu His Thr Glu Glu Cys Thr Cys Gly Phe Ala Ser Asn Lys
                 265                 270                 275                 280                 285

ACC AUA GAA UGU GCC UGU AGA GAC AAC AGU UAC ACA GCA AAA AGA CCC UUU GUC AAA UUA AAU GUG GAA ACU GAU
Thr Ile Glu Cys Ala Cys Arg Asp Asn Ser Tyr Thr Ala Lys Arg Pro Phe Val Lys Leu Asn Val Glu Thr Asp
                 290                 295                 300                 305                 310

                         1000
ACA GCU GAA AUA AGA UUG AUG UGC ACA AAG ACU UAU CUA GAC ACU CCC AGA CCG GAU GAU GGA AGC AUA GCA GGG
Thr Ala Glu Ile Arg Leu Met Cys Thr Lys Thr Tyr Leu Asp Thr Pro Arg Pro Asp Asp Gly Ser Ile Ala Gly
                 315                 320                 325                 330                 335

                                                   1100
CCU UGC GAA UCU AAU GGA GAC AAG UGG CUU GGA GGC AUC AAA GGA GGA UUC GUC CAU CAA AGA AUG GCA UCU AAG
Pro Cys Glu Ser Asn Gly Asp Lys Trp Leu Gly Gly Ile Lys Gly Gly Phe Val His Gln Arg Met Ala Ser Lys
                 340                 345                 350                 355                 360

                                                                                    1200
AUU GGA AGA UGG UAC UCC CGA ACG AUG UCU AAA ACU AAC AGA AUG GGG AUG GAA CUG UAU GUA AAG UAU GAU GGU
Ile Gly Arg Trp Tyr Ser Arg Thr Met Ser Lys Thr Asn Arg Met Gly Met Glu Leu Tyr Val Lys Tyr Asp Gly
                 365                 370                 375                 380                 385

GAC CCA UGG ACU GAC AGU GAU GCU CUU ACU CUU AGU GGA GUA AUG GUU UCC AUA GAA GAA CCU GGU UGG UAU UCU
Asp Pro Trp Thr Asp Ser Asp Ala Leu Thr Leu Ser Gly Val Met Val Ser Ile Glu Glu Pro Gly Trp Tyr Ser
                 390                 395                 400                 405                 410

                         1300
UUU GGC UUC GAA AUA AAG GAC AAG AAA UGU GAU GUC CCU UGU AUU GGG AUA GAG AUG GUA CAC GAU GGU GGA AAA
Phe Gly Phe Glu Ile Lys Asp Lys Lys Cys Asp Val Pro Cys Ile Gly Ile Glu Met Val His Asp Gly Gly Lys
                 415                 420                 425                 430                 435

                                                   1400
GAU ACU UGG CAU UCA GCU GCA ACA GCC AUU UAC UGU UUG AUG GGC UCA GGA CAA UUG CUA UGG GAC ACU GUC ACA
Asp Thr Trp His Ser Ala Ala Thr Ala Ile Tyr Cys Leu Met Gly Ser Gly Gln Leu Leu Trp Asp Thr Val Thr
                 440                 445                 450                 455                 460

                                                                                    1500
GGC GUU GAU AUG GCU UUA UAA UAG AGGAAUGGUUGGAUCUGUUCUAAACCCUUUGUUCCUAUUUUAUUUGAACAGUUGUUCUUACUAGAUU
Gly Val Asp Met Ala Leu
                 465

                                  1557
UAAUUGUUUCUGAAAAAUGCUCUUGUUACUACU-3'
```

Table 8. Amino acid and carbohydrate composition of B/Lee/40 neuraminidase. Composition from sequence data (SHAW et al. 1982); analysis 1 (LAVER and BAKER 1972); analysis 2 (ALLEN et al. 1977)

Residue	Intact neuraminidase			Heads		Stalk
	sequence	analysis 1	analysis 2	sequence	analysis 2	sequence
Amino acid composition						
Aspartic acid	25	46	45	24	37	1
Asparagine	15			12		3
Threonine	40	37	36	29	28	11
Serine	38	35	42	26	31	12
Glutamic acid	25	36	44	24	36	1
Glutamine	8			6		2
Proline	22	21	24	20	21	2
Glycine	41	44	48	39	43	2
Alanine	30	30	31	25	26	5
$^1/_2$ cystine	18	12	23	17	15	1
Valine	22	23	23	17	17	5
Methionine	14	11	11	12	10	2
Isoleucine	26	26	21	26	25	0
Leucine	38	40	27	24	24	14
Tyrosine	20	18	13	17	15	3
Phenylalanine	13	13	13	12	11	1
Tryptophan	9	ND	7	9	5	0
Histidine	13	12	17	12	13	1
Lysine	28	31	23	26	22	3
Arginine	21	22	18	21	18	0
Total residues	466	457+Trp	466	397	397	69
Apoprotein molecular weight	51441			44174		7285
Carbohydrate composition						
N-Acetylglucosamine			15		8	7
Mannose			14		12	2
Galactose			9		3	6
Fucose			4		1	3
Total residues			42		24	18
Carbohydrate molecular weight			7592		4273	3319
Glycoprotein molecular weight			59033		48447	10604

The data for analysis 1 has been expressed as residues per mole, not residues % as in the original manuscript (LAVER and BAKER 1972). The data for analysis 2 has been recalculated for 466 amino acid residues (397 in heads, 69 in stalk) not 549 amino acids (401 in heads, 148 in stalk), as in the original reference (ALLEN et al. 1977)

the nucleotide sequence data, is compared in Table 8 with the experimentally determined values for the detergent-released enzyme, the trypsin-released heads, and the residual membrane-attached stalk region (ALLEN et al. 1977). Although the position of trypsin cleavage during solubilization of the B/Lee/40 heads has not been established by direct amino acid sequence analysis, the composition data in Table 8 and the known point of pronase cleavage in N2 neuraminidase (WARD et al. 1982) suggest that the site of trypsin action is Lys_{69} (see Table 7). The high proline content of the B/Lee/40 heads suggests cleavage does not occur at Arg_{84}. The high carbohydrate content of the removed stalk region (Tables 7 and 8) and the fact that trypsin-released heads of B/Lee/40 are monomers (LAZDINS et al. 1972; ALLEN et al. 1977), not disulphide-linked dimers, indicate that cleavage cannot occur at Lys_{39} or Lys_{44}.

Of the 18 half-cystine residues in B/Lee/40 neuraminidase, none occurs in the hydrophobic transmembrane region, one (residue 54) occurs in the stalk, and the remaining 17 (residues 87, 122, 127, 182, 229, 231, 236, 251, 277, 279, 289, 291, 318, 337, 420, 424, and 447) occur in the heads.

There are four potential glycosylation sites in B/Lee/40 neuraminidase, two in the stalk region (asparagine residues 56 and 64), and two in the heads (residues 144 and 284). From the composition data (ALLEN et al. 1977) shown in Table 8 for intact B/Lee/40 neuraminidase and the trypsin-released heads, it would appear that all four potential sites are glycosylated. The composition data suggest the two oligosaccharide units in the stalk would be of the *N*-acetyl-lactosamine (complex, Type I) type while the two sites in the heads carry one *N*-acetyllactosamine type and one oligosaccharide (simple, Type II) type sugar moiety.

3 Secondary Structure

3.1 The Disulfide Bonds

The primary structure data discussed in the previous section have shown that the type A N1, type A N2, and type B neuraminidase monomers contain 19, 22, and 18 half-cystine residues respectively. As shown in Table 9, the number and location of the half-cystine residues in the hydrophobic transmembrane portion and in the stalk are variable, while most of the half-cystine residues in the heads occur in relatively equivalent positions. The half-cystine residue numbers are not exactly equivalent because the N1 strain used for comparison, A/PR/8/34, has a 15-residue deletion in the stalk between residues 62 and 63

Table 9. Comparative locations of the half-cystine residues in the neuraminidase of A/PR/8/34 (N1) (FIELDS et al. 1981), A/RI/5$^-$/57 (N2) (ELLEMAN et al. 1982), and B/Lee/40 (SHAW et al. 1982)

Type A N1	14,				49,						77,		109,		114,		146,				169,	
(Corrected + 15)											(92)		(124)		(129)		(161)				(184)	
Type A N2			21,		42,				53,	78,	92,		124,		129,				175,		183,	193,
Type B									54,		87,		122,		127,						182,	

when compared with other N1 strains (BLOK and AIR 1982a). In addition, several single or double insertions or deletions have been made to maximize the homology between the N1, N2, and type B neuraminidases (see Sect. 5.4).

3.1.1 Disulfide Bonds in N2 Neuraminidase

The arrangement of these half-cystine residues in disulfide bridges has been examined only in the heads of the N2 enzyme. All 19 half-cystine residues are involved in disulfide bonds, since no free sulfydryl groups could be detected (KENDAL and ECKERT 1972) in the subtilisin-released heads of A/R1/5$^+$/57 neuraminidase (X7F1). Recent chemical (WARD et al. 1983a) and crystallographic (VARGHESE et al. 1983) analyses of the pronase-released neuraminidase of A/Tokyo/3/67 have shown that 18 of these Cys residues are involved in intrachain linkages within the heads and that one is involved in an interchain bridge that links pairs of monomers at the distal end of the stalk region in the neuraminidase tetramer.

In chemical studies (WARD et al. 1983a), cyanogen bromide digestion of Tokyo/67 heads showed that CN1 (residues 78–160, containing Cys residues 78, 92, 124, and 129) was linked by at least one disulfide bridge to CN4 (residues 363–467, containing Cys residues 417, 421, and 447) and that neither CN2 (residues 161–306, containing ten Cys residues at 175, 183, 193, 230, 232, 237, 278, 280, 289, and 291) nor CN3 (which contains Cys residues 318 and 337) was disulfide-linked to any other fragment. Enzymatic digestion and diagonal peptide mapping of the cyanogen bromide fragments confirmed the Cys_{318}-Cys_{337} bond in CN3; established three of the five disulfide bridges in CN2, Cys_{183}-Cys_{230}, Cys_{278}-Cys_{291}, and Cys_{280}-Cys_{289}; established the Cys_{92}-Cys_{417} bond, the only link between the N-terminal cyanogen bromide peptide CN1 and the penultimate C-terminal peptide CN4; and provided evidence consistent with but not proving the disulfide bridges Cys_{124}-Cys_{129} in CN1 and Cys_{175}-Cys_{193} in CN2 (WARD et al. 1983a). The three-dimensional crystal structure confirmed these assignments and provided the evidence for the remaining intrachain linkages Cys_{232}-Cys_{237} in CN2 and Cys_{421}-Cys_{447} in CN4 (VARGHESE et al. 1983).

The 19th Cys residue at position 78 in the N2 enzyme was not found in any of the intrachain disulfide bonds but appears to be involved in an interchain disulfide bond between adjacent monomeric units in the neuraminidase tetramer. This was first suspected when different preparations of pronase-released heads were found to contain variable proportions of monomer to dimer when exam-

216,	218,	223,		264,	266,	275,	277,	303,	320,	402,	406,	431
(231)	(233)	(238)		(279)	(281)	(290)	(292)	(318)	(335)	(417)	(421)	(446)
230,	232,	237,		278,	280,	289,	291,	318,	337,	417,	421,	447
229,	231,	236,	251,	277,	279,	289,	291,	318,	337,	420,	424,	447

ined on sodium dodecyl sulfate (SDS) gels without reduction (BLOCK et al. 1982). Since pronase digestion was known to produce ragged-end cleavage in front of residues 74 and 76 (WARD et al. 1982), it seemed possible that in some preparations pronase digestion could include some cleavage after Cys_{78} resulting in the absence of disulfide-linked dimers in these neuraminidase head preparations. In the chemical study of the disulfide bonds of A/Tokyo/67 (WARD et al. 1983a), no partner was found for the peptide that contained Cys_{78} in the enzymatic digest of CN1-CN4, and the elution position of the CN1-CN4 complex on Sephacryl S300 suggested it was present as the dimer. In addition, the comparative sequences (see ELLEMAN et al. 1982) and three-dimensional structure (VARGHESE et al. 1983) suggest that Cys_{78} occurs at the top of the stalk region of N2 neuraminidase where it can participate in disulfide linkage with the homologous residue on another monomer. The arrangement of these intra- and interchain disulfide bridges in the three-dimensional structure is discussed in Sect. 4.3.

The disulfide bond arrangements of the remaining Cys residues in the membrane-embedded region (Cys_{21}) and the stalk (Cys_{42} and Cys_{53}) of A/Tokyo/67 neuraminidase were not examined.

3.1.2 Disulfide Bonds in Other Neuraminidases

The disulfide bond arrangements for N1, type B, and other neuraminidases have not been investigated chemically. However, the 16 Cys residues found in homologous positions in the N1 and B neuraminidase heads would be expected to be involved in the same eight disulfide bonds as their counterparts in the N2 structure (Table 10). As shown in Tables 9 and 10, one of the N2 disulfides, Cys_{175} and Cys_{193}, has no counterpart in either N1 or type B neur-

Table 10. Disulfide bonds in N2, N1, and type B neuraminidase heads

A/Tokyo/3/67 (N2)	A/PR/8/34 (N1)	B/Lee/40 (type B)
92–417	77–402	87–420
124–129	109–114	122–127
175–193	absent	absent
183–230	169–216	182–229
232–237	218–223	231–236
278–291	264–277	277–291
280–289	266–275	279–289
318–337	303–320	318–337
421–447	406–431	424–447

The N2 bonds were established experimentally (WARD et al. 1983a; VARGHESE et al. 1983). The N1 and type B bonds are based on homology (see Sect. 5.5.5)

aminidase. The status of the odd Cys residues at position 146 in A/PR/8/34 and 251 in B/Lee/40 (Table 9) is not known. The three-dimensional structure (see Sect. 4) suggests that it is possible for Cys_{146} in the A/PR/8/34 structure to be involved in dimer cross-linking, but not for Cys_{251} in the B/Lee/40 structure. The latter Cys residue should exist as free cysteine, the presence of which has not been investigated in B/Lee/40 neuraminidase.

In the stalk region, N1 and type B neuraminidases lack the Cys residue equivalent to Cys_{78} in the N2 sequence but contain single Cys residues at positions 49 and 54 respectively. In the case of B/Lee/40, a comparison of the subunit structure of the detergent-released and trypsin-released neuraminidases showed that Cys_{54} was involved in the formation of interchain cross-linked dimers (LAZDINS et al. 1972; ALLEN et al. 1977).

Other type A neuraminidases show variable numbers of Cys residues in this stalk region, and these could contribute to multiple disulfide bridges between the extended polypeptide chains of the four identical monomers that make up the neuraminidase tetramer (BLOK and AIR 1982b). It is not known whether the tetramers of those subtype neuraminidases with multiple half-cystine residues of the stalk are assembled from dimers linked by several disulfides or whether disulfide bridges in the stalk region are formed at random across all four monomers.

3.2 Empirical Predictions of Secondary Structure

While experimental evidence has shown that all the information required for a protein to acquire its three-dimensional structure is stored in its amino acid sequence (ANFINSEN 1973), attempts to predict tertiary structure by various algorithms have in general been unsuccessful. However, such predictive methods do have some value in indicating local regions of secondary structure.

Two schemes for predicting secondary structure from amino acid sequence data have been applied to A/RI/5^-/57 N2 neuraminidase, and the conformational assignments are shown in Table 11. The Chou and Fasman procedure (AZAD et al. 1983) showed that the neuraminidase heads were very rich in potential β-structure. There were 19 β-strands varying in length from five to nine residues, 46 β-turns, and only three short regions predicted to be α-helical. Large loops of random coil/β-turn structure were predicted to occur between residues 134 and 155 of CN1, residues 278–302 of CN2, and residues 307–360, which accounts for most of the CN3 sequence (see Sect. 2.2.2).

The Robson prediction algorithm (GARNIER et al. 1978) indicates a similar pattern of 19 β-strands, 33 β-bends, and two short sections of α-helix (Table 11). The two prediction algorithms were not identical in their structural assignments; most of the β-turns, but only 12 of the 19 β-strands and one of the helical sections were predicted by both procedures.

When compared with the three-dimensional structure determined by X-ray crystallographic analysis (see Sect. 4), the Chou and Fasman procedure identified 14 of the 24 β-strands found in the head region of the neuraminidase monomer (Table 12), while the Robson procedure predicted 18 of the 24 β-

Table 11. Predicted secondary structure for A/RI/5^{-}/57 neuraminidase. The data using the Chou and Fasman method is from AZAD et al. (1983)

Protein region	Chou and Fasman method			Robson method		
	α-helix	β-sheet	β-turn	α-helix	β-sheet	β-turn
Transmembrane region		5–10	2–5		6–11	
	13–18				15–19	
	20–31	13–35		20–38		
Stalk	33–41	47–53	40–44		50–55	39–43, 45–49
		56–60		56–60		
		61–73			63–67	
	73–78		78–81	77–83		
Heads			84–87, 89–91			84–87
		91–98	101–104		92–97	100–104, 109–112
		114–118			114–118	120–123
		129–134	124–128		129–134	124–128
			135–138			
			142–148			145–148
		155–161			157–160	
			170–173	165–168		169–172
		173–179	179–182, 183–187		173–177	178–182, 183–187
		188–195	197–201		190–194	195–198
		202–207	207–210		203–207	
		218–226	227–230		210–214	224–227, 228–231
		230–234	234–237			232–236
		236–242	242–245		237–242	
			245–248			
		252–257				
	258–265		266–271	256–261	262–266	
	271–278		278–282, 283–287			277–281, 282–286
					287–291	
			291–295, 295–298			292–295, 296–298
			297–300			
		302–307	309–312, 313–316		302–308	309–312, 312–315
			318–321		316–322	
			325–329, 330–333			327–330, 331–334
			334–337, 337–340			335–338
			341–345			340–345
			354–359		350–354	355–358
		360–366	367–370, 371–374			371–374
		376–380	380–383		376–379	380–383
			385–388, 388–391			
		391–398	399–403, 403–406		392–399	400–403
			414–416		409–412	413–416
		417–424				417–420
	425–430		431–434		421–426	
		436–440	439–442			440–443
		443–447	448–451		444–447	
			451–457, 458–461			451–456
		462–467			466–469	

Table 12. Predicted and observed β-structure in N2 neuraminidase. The Chou and Fasman data is from AZAD et al. (1983) and the Robson data from COLMAN (1984)

Observed β-structure in N2 neuraminidase heads		Predicted	
		CHOU and FASMAN	ROBSON
β_1S_1	120–123	114–118	114–118
β_1S_2	130–134	129–134	129–134
β_1S_3	157–162	155–161	157–160
β_1S_4	170–174	173–179	173–177
β_2S_1	178–185		
β_2S_2	188–194	188–195	190–194
β_2S_3	200–207	202–207	203–207
β_2S_4	210–217		210–214
β_3S_1	228–233	218–226, 230–234	
β_3S_2	237–242	236–242	237–242
β_3S_3	251–259	252–257	
β_3S_4	262–268		262–266
β_4S_1	278–283		
β_4S_2	286–291		287–291
β_4S_3	296–304	302–307	302–308
β_4S_4	307–314		316–322
β_5S_1	351–356		350–354
β_5S_2	359–366	360–366	
β_5S_3	372–382	376–380	376–379
β_5S_4	387–398	391–398	392–399
β_6S_1	404–414		409–412
β_6S_2	417–428	417–424	421–426
β_6S_3	438–447	436–440, 443–447	444–447
β_6S_4	95–106	91–98	92–97

strands. With regard to the predictions of α-helix, only one of the four predictions appears to be correct. The three-dimensional structure (VARGHESE et al. 1983) shows that residues 429–437, which form the loop between the second and third strands of the sixth sheet, approximate α-helical configuration.

There is no three-dimensional X-ray data for the stalk region, although electron micrographs show that it is a long, slender, filamentous structure 12–15 Å in diameter and approximately 100 Å in length (LAVER and VALENTINE 1969; WRIGLEY et al. 1973). The Chou and Fasman structural predictions suggest that the residues constituting the stalk region include a 27-residue region of β-structure (residues 47–73). This β-strand region in each monomer could be arranged in the neuraminidase tetramer either as a four-chain β-barrel (HARRISON et al. 1978; WEBER et al. 1981) or as two double-stranded parallel β-sheets wound round each other (CHOTHIA et al. 1977). The latter arrangement would be more likely if the neuraminidase tetramer was assembled from two disulfide-

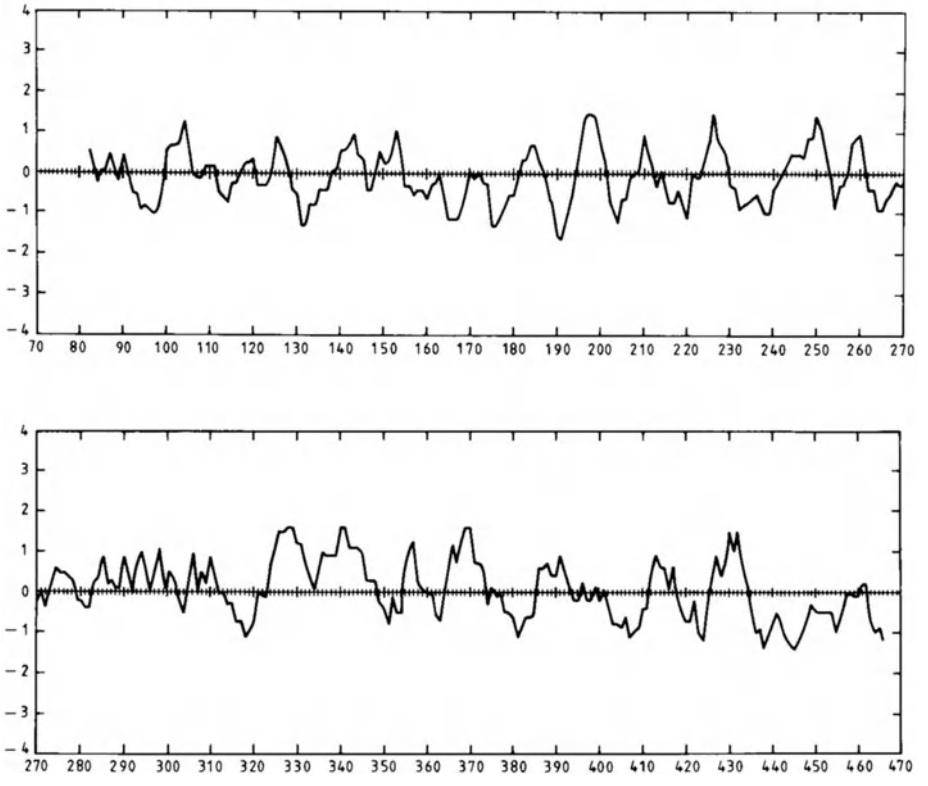


Fig. 1. Hydrophilicity profile (*ordinate*) of the N2 neuraminidase head region amino acid sequence (COLMAN, 1984)

linked dimers (ALLEN et al. 1977). Such extended four-chain β-structures would have an expected diameter of 10–12 Å (LEVITT and CHOTHIA 1976) and a length of 95 Å consistent with the electron micrographs (LAVER and VALENTINE 1969; WRIGLEY et al. 1973). The structure of the four polypeptide chains in the stalk region could be stabilized by the presence of disulfide bonds between the homologous half-cystine residues on adjacent monomers.

The hydrophobic sequence (residues 7 to 35) that constitutes the membrane-embedded region of the neuraminidase molecule probably occurs in α-helical conformation. The length of this region (27 residues) is similar to that found in influenza hemagglutinin (see WARD 1981) and other membrane-embedded proteins (see FIELDS et al. 1981 for references) and is sufficient to allow the polypeptide chain to span the membrane once. The Robson algorithm predicts an α-helical conformation for residues 20–38, while the Chou and Fasman procedure supports either α-helix or β-sheet structure for the region (Table 11). It is interesting to note that the three-dimensional structure of influenza virus hemagglutinin shows α-helical conformation for the polypeptide chain just preceding the membrane-embedded region (WILSON et al. 1981), but no data are

available to indicate whether that helix continues into the hydrophobic C-terminal region of the protein.

3.3 Hydrophilicity Profile

A plot of the hydrophilicity index of the N2 head sequence is shown in Fig. 1 (HOPP and WOODS 1981). Peaks in this profile are expected to correlate with surface residues in the three-dimensional structure of the protein, and these in turn may also correlate with antigenic determinants in some cases. The maximum values of this function are at residues 198, 225, 250, 328, 341, 370, and 430, in good agreement with surface residues (Sect. 4) and fair agreement with antigenic regions (Sect. 5).

4 Tertiary Structure

There are four functionally distinct structural regions in influenza neuraminidase. The six N-terminal amino acids, which are strictly conserved in all influenza A viruses, are likely to reside on the cytoplasmic side of the viral membrane (BLOK and AIR 1982b). These are followed by 29 hydrophobic residues (30 and 31 amino acids in neuraminidase subtypes N8 and N5 respectively, which are believed to constitute the transmembrane segment of the protein. A linker region begins at residue 36 and extends to residue 77 in N2 strains. The C-terminus of this peptide is defined by the site at which pronase cleavage occurs (WARD et al. 1982), liberating soluble heads of neuraminidase which are enzymatically and antigenically indistinguishable from the intact integral membrane protein (LAVER 1978). The only crystal forms of neuraminidase that have been grown so far are from pronase-liberated heads. Two problems confront attempts to crystallize intact molecules. One is the insoluble nature of the whole protein and the requirement that detergent be contained in the crystallization medium (MICHEL 1983). The other is the nature of the linker or stalk peptide, whose length and sequence show considerable variation between subtypes of neuraminidase, and whose structure is likely to be flexible.

Two N2 neuraminidases have been crystallized. A/Tokyo/3/67 neuraminidase crystallizes in low ionic strength buffers (LAVER 1978; WRIGHT and LAVER 1978) in space group I422 with $a=139.6$ Å, $c=191.0$ Å (COLMAN and LAVER 1981). The tetrameric protein in those crystals is centered around the crystallographic fourfold axis with one protein subunit in the crystallographic asymmetric unit. A second N2 enzyme, RI/5^+/57 neuraminidase, (VARGHESE et al. 1983) crystallizes from 0.75 *M* sodium citrate at pH 7.0 as tetragonal bipyramids in space group $P4_321$ with $a=124.1$ Å, $c=181.2$ Å. In these crystals the tetrameric enzyme is centered around the crystallographic twofold axis, and there are two copies of the subunit in the crystallographic asymmetric unit. Both crystal forms are loosely packed, containing by volume only 26% and 34% protein respectively. The three-dimensional structures of the two N2 enzymes have been deter-

mined by multiple isomorphous replacement and non-crystallographic symmetry averaging (VARGHESE et al. 1983). The current image of the protein molecule is an average of the two structures. Independent crystallographic refinement of both structures will show the structural changes associated with antigenic and other variation between 1957 and 1967.

A third crystalline influenza neuraminidase has recently been reported (LAVER et al., 1984). This neuraminidase is from the N9 subtype and crystallized in 1.9 *M* potassium phosphate at pH 6.8. The crystals are rhombic dodecahedra belonging to space group I432. The molecular fourfold axis is the crystallographic tetrad. The cell edge is 181.4 Å, and there is one subunit per asymmetric unit. These are the most densely packed neuraminidase crystals observed yet, containing by volume 48% protein, and measurable X-ray diffraction data extend to beyond 2 Å resolution. That structure analysis is in progress and is of special interest because the neuraminidase shows hemagglutinating activity (LAVER et al. 1984).

Microcrystals from neuraminidase heads of other N2 strains have also been grown, and in the case of Ned/84/68, characterized by electron microscopy (COLMAN et al. 1980). That study indicates that these crystals are probably isomorphous with Tokyo/3/67 crystals. The unit cell is body-centered, and $a = b \sim 145$ Å. The *c*-axis spacing has not been determined. Crystals of this strain suitable for high-resolution X-ray diffraction analysis have not yet been grown.

4.1 Molecular Shape

The neuraminidase tetramer is a mushroom-shaped object comprising a box-shaped head attached to a filamentous stalk (LAVER and VALENTINE 1969; WRIGLEY et al. 1973). The dimensions of the head are $100 \times 100 \times 60$ Å (VARGHESE et al. 1983). The length of the stalk varies. In the N2 subtype, the 40 amino acids in the stalk measure 140 Å in fully extended conformation. Electron micrographs (LAVER and VALENTINE 1969; WRIGLEY 1979) suggest lengths of the order of 100 Å for this peptide. Thus, the distal end of neuraminidase may be as far as 160 Å from the viral membrane, a distance on the order of that estimated for hemagglutinin, the external length of which is 135 Å (WILSON et al. 1981).

4.2 Polypeptide Folding

The backbone chain tracing of neuraminidase has revealed a completely new topological class of structure, a β-sheet propeller (VARGHESE et al. 1983). A stereo view of the C^{α} positions in the monomer is shown in Fig. 2, and diagrams of the chain-folding are given in Figs. 2–4. The polypeptide backbone passes sequentially through six four-stranded, antiparallel β-sheets in counter clockwise order as the subunit is viewed into its upper face (i.e., that outermost from the viral membrane). Each β-sheet has the topology of a "W", or +1, +1, +1 in RICHARDSON's (1977) notation. The first strand of each sheet to be built

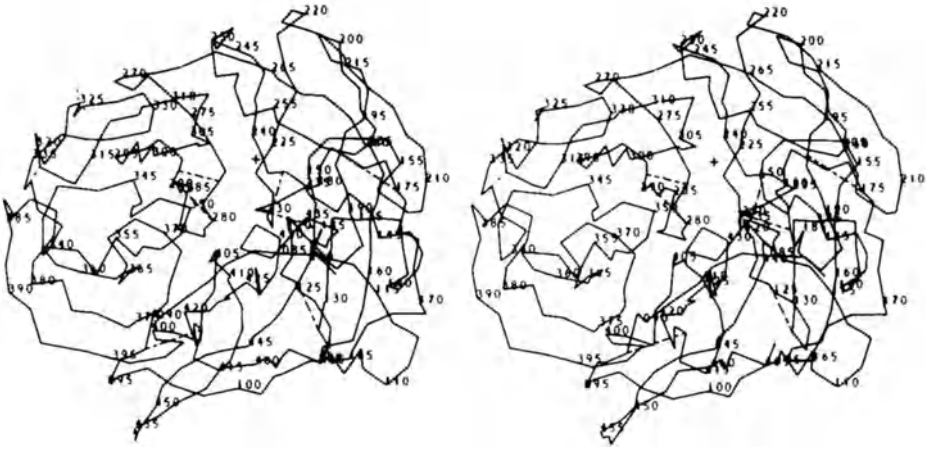


Fig. 2. Stereo view of α-carbon skeleton of neuraminidase subunit. Disulfide bonds are *dashed*. Active site is marked with a *cross*. View is from above with the viral membrane to the rear (VARGHESE et al. 1983). Reprinted by permission from *Nature*. Vol. 303, pp 35–40, 1983

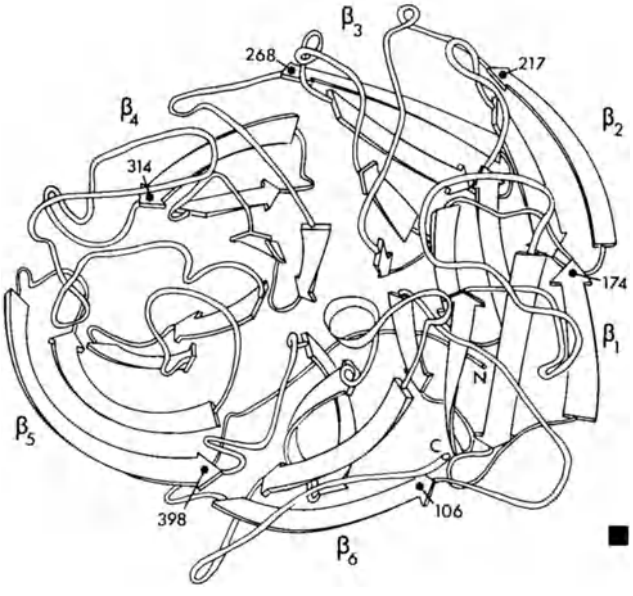


Fig. 3. Diagram of polypeptide folding in neuraminidase subunit. View as in Fig. 2, namely from above looking down the fourfold symmetry axis shown in the bottom right-hand corner. The six β-sheets of the propellor fold are labeled, as are the N- and C-terminii of the head (VARGHESE et al. 1983). Reprinted by permission from *Nature*. Vol. 303, pp 35–40, 1983

is that nearest the subunit center, and the last is at the subunit surface. All sheets show the typical, right-handed twist (CHOTHIA 1973), with the interior strands nearly parallel to the molecular four-fold symmetry axis and the exterior strands inclined to be more nearly in the plane of the tetrameric assembly.

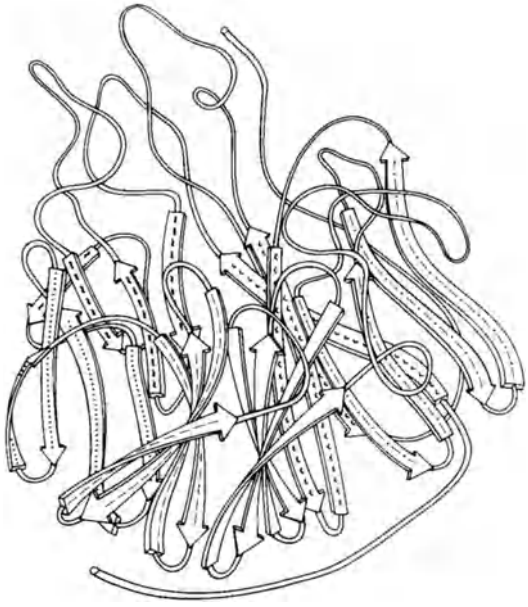


Fig. 4. Diagram of polypeptide folding in neuraminidase subunit viewed from the side at right angles to the molecular symmetry axis, here located at the left rear of the view. The six sheets are embellished with different markings

In each sheet, two of the loops connecting strands are on the upper surface (L_{01} connecting adjacent sheets and L_{23} connecting strands 2 and 3 within sheets), and two are on the bottom surface (L_{12} and L_{34}). Notating the *j*th strand of the *i*th sheet as $\beta_i S_j$, the folding pattern can be described as follows (see also Table 13):

N-terminal arm on bottom of subunit

$\beta_6 S_4$
$\beta_1 L_{01}, \beta_1 S_1, \beta_1 L_{12}, \beta_1 S_2, \beta_1 L_{23}, \beta_1 S_3, \beta_1 L_{34}, \beta_1 S_4$
$\beta_2 L_{01}, \ldots$
$\ldots \beta_6 S_2, \beta_6 L_{23}, \beta_6 S_3,$

C-terminal arm at subunit interface (VARGHESE et al. 1983).

Table 13. Assignment of sheet and loop structure in N2 influenza virus neuraminidase

	i=1	2	3	4	5	6
$\beta_i L_{01}$	107–119	175–177	218–227	269–277	315–350	399–403
$\beta_i S_1$	120–123	178–185	228–233	278–283	351–356	404–414
$\beta_i L_{12}$	124–129	186–187	234–236	284–285	357–358	415–416
$\beta_i S_2$	130–134	188–194	237–242	286–291	359–366	417–428
$\beta_i L_{23}$	135–156	195–199	243–250	292–295	367–371	429–437
$\beta_i S_3$	157–162	200–207	251–259	296–304	372–382	438–447
$\beta_i L_{34}$	163–169	208–209	260–261	305–306	383–386	448–450
$\beta_i S_4$	170–174	210–217	262–268	307–314	387–398	95–106

The N-terminal strand (84–94) and the C-terminal strand 451–469 are appendages at the bottom of the subunit and at the subunit interface respectively. Not all of the residues assigned to the strands of the sheet are in regular β-conformation

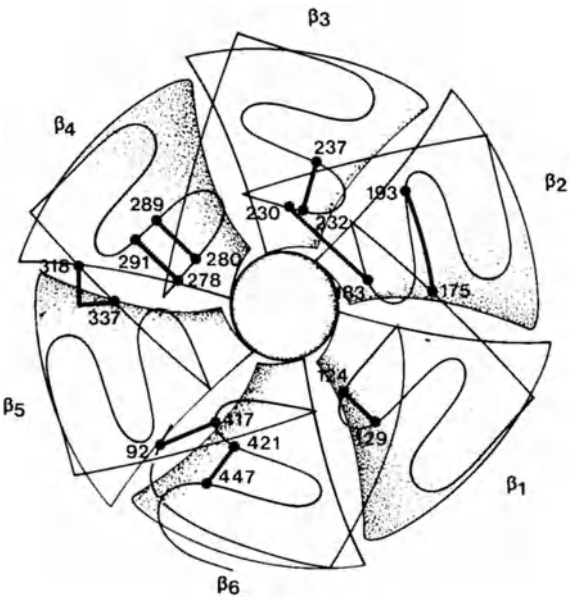


Fig. 5. Disulfide bonds in N2 neuraminidase overlaid on the sheets of the propellor fold

These loop structures connecting the strands are variable in length and carry amino acids implicated in enzyme activity and antigenic variation (COLMAN et al. 1983). The conservation of small amino acids at 164 (gly, β_1L_{34}), 177 (ala or gly, β_2L_{01}), 186 (gly, β_2L_{12}), 196 (gly, β_2L_{23}), 235 (gly, β_3L_{12}), 244 (gly, β_3L_{23}), 261 (gly, β_3L_{34}), 270 (gly or ala, β_4L_{01}), and 348 (gly, β_5L_{01}) in some loops points to their importance in mediating reversal of chain direction. Other loops are bounded by disulfide bonds, in particular β_1L_{12}, β_3L_{12}, β_4L_{12}, and β_6L_{23} (Fig. 5).

A preliminary table of secondary structure assignments is given in Table 13 (VARGHESE et al. 1983) and is the basis for Figs. 2 and 3. Refinement of the structure will be necessary before interstrand hydrogen-bonding assignments can be made. The structure in β_4S_1 and β_4S_2 is especially interesting because it includes the disulfide linkages 278–291 and 280–289. Since disulfides may not join neighboring strands in a regular β-sheet (RICHARDSON 1981), the backbone conformation in this region of neuraminidase is presumably relaxed from regular β-strand to allow these two covalent linkages.

Between all pairs of sheets there are cores of hydrophobic amino acids, which are listed in Table 14. Many of these residues are either conserved across influenza subtypes or conservatively exchanged. Some of these cores appear to be more extensive than others. At the center of the subunit, the S_1 strands on all sheets run parallel to each other. Viewed from above the subunit, the six strands project onto the corners of a distorted hexagon such that β_2S_1 and β_5S_1 are further apart than any other pairs. The possibility of some parallel-sheet structure linking β_1 to β_3 and β_4 to β_6 cannot yet be excluded, but the chains appear too far apart for this type of interaction.

Two classes of sheet-sheet packing, aligned and orthogonal, have been identified in globular proteins (CHOTHIA et al. 1977). Neuraminidase is clearly in

Table 14. Amino acids in the hydrophobic cores between adjacent sheets in the N2 neuraminidase propeller

Core	Contributing amino acids
β_1–β_2	Phe 132, Leu 134, Leu 158 Trp 189
β_2–β_3	Val 192, Phe 205, Tyr 207 Met 241, Leu 255, Ile 257, Ile 262
β_3–β_4	Val 240, Ile 254 Cys 278, Cys 280, Cys 289, Cys 291
β_4–β_5	Ile 290 Trp 383
β_5–β_6	Trp 352, Phe 354, Trp 361, Tyr 374 Phe 422, Val 424, Phe 446, Phe 97
β_6–β_1	Phe 410, Tyr 423, Phe 100

the former category. Furthermore, the geometrical arrangement of the sheets in the propeller requires that the intersheet angle be negative (CHOTHIA and JANIN 1981; COHEN et al. 1981).

Two different cleavage sites for pronase in Tokyo neuraminidase have been positively identified at residues 74 and 77 (WARD et al. 1982). A third site on the C-terminal side of Cys 78 can be inferred from the observation that the subunits in the neuraminidase heads are sometimes disulfide-bonded in pairs and sometimes not. Cys 78 presumably mediates this covalent dimerization. Electron density for Tyr 84 is clearly visible in the present map 19 Å from the four-fold axis on the underside of the subunit. Beyond this point, the density falls to background levels. A tentative assignment of the chain through weak density from Tyr 84 back to Pro 79 in strong density has been made, but this chain segment is not included in any of the present figures. Such a structure would give rise to a loop in intact neuraminidase, with Cys 78 at the top of the loop and at the top of a column of electron density seen on and around the fourfold axis in a 5 Å resolution map of Tokyo neuraminidase (COLMAN and LAVER 1981). This feature is not prominent in the high-resolution image. A number of factors may be responsible, including heterogeneity of the pronase cleavage, sequence variability between RI/5$^+$ and Tokyo neuraminidase, and failure of this segment to observe strictly the fourfold symmetry axis.

4.3 Disulfide Bonds

The complete list of intrasubunit disulfide bonds consists of Cys 92-Cys 417, Cys 124–129, Cys 175-Cys 193, Cys 183-Cys 230, Cys 232-Cys 237, Cys 278-Cys 291, Cys 280-Cys 289, Cys 318-Cys 337, and Cys 421-Cys 447 (Fig. 5). Chemical studies have confirmed six of these pairs (WARD et al. 1983a). This pattern of disulfide bonding reflects the sequential arrangement of β-sheets from

N- to C-termini. The only bridge between linearly distant residues originates near the two termini of the heads (residues 92 and 417) and connects the N-terminal strand to the bottom of sheet 6 at $\beta_6 L_{12}$. It is particularly interesting that the disulfide-rich regions 230–237 and 278–291 are adjacent to residues implicated in catalytic activity (COLMAN et al. 1983). These regions contain disulfide bonds between $\beta_3 S_1$ and $\beta_3 S_2$ and between $\beta_4 S_1$ and $\beta_4 S_2$ respectively, and the loops $\beta_3 L_{01}$ and $\beta_4 L_{01}$ contain conserved charged residues.

N1 type neuraminidase contains an additional cysteine at residue 146 (FIELDS et al. 1981; HITI and NAYAK 1982), which corresponds to residue 161 in the N2 type sequence. This residue is possibly disulfide-bonded to its counterpart on a neighboring monomer. The bond between residues 175 and 193 is not found in N1 (FIELDS et al. 1981; HITI and NAYAK 1982) or influenza B (SHAW et al. 1982) neuraminidase. It joins the end of β-sheet 1 ($\beta_1 S_4$) with the center of β-sheet 2 ($\beta_2 S_2$) in the N2 structure. An additional cysteine at a position analogous to residue 251 in the N2 sequence has been found in influenza B neuraminidase (SHAW et al. 1982). This residue is distant from the four-fold axis and would have to exist as free cysteine, if present.

4.4 Surface Distribution of Carbohydrate

Oligosaccharides are attached at asparagine residues 86, 146, 200, and 234 (WARD et al. 1982). As shown in Fig. 6, a fifth potential glycosylation site, Asn 402, carries no carbohydrate in Tokyo neuraminidase (WARD et al. 1983b). Residues 86 and 234 are on the bottom face of the monomer, approximately 11 Å apart. Little if any electron density is visible for the sugars in either case. Residue 146 is on the top surface of the molecule in $\beta_1 L_{23}$, and the oligosaccharide attached there is of the complex type and is serologically cross-reactive with chick cell host antigen (WARD et al. 1983b). This glycosylation site is conserved in all known sequences except that of the neurovirulent virus WSN/33 (HITI and NAYAK 1982). It is also unique in containing a significant amount of *N*-acetylgalactosamine (WARD et al. 1983b). The only carbohydrate plainly visible in the present electron density map is that attached at residue 200 on $\beta_2 S_3$. This site is near the surface of the tetramer on a subunit interface, and the sugar residues appear to interact primarily with surface residues on the neighboring monomer on $\beta_5 S_4$.

We note that the other two glycoproteins whose three-dimensional structures are known also have carbohydrate attached at a subunit interface. In immunoglobulin γ_1 heavy chains, carbohydrate is attached to the C_H2 domain in such a way that the association of the two C_H2 domains in the dimeric molecule is different from that found between other pairs of constant domains (DEISENHOFER et al. 1976). The hemagglutinin of influenza virus, like the neuraminidase reported here, has a number of glycosylation sites, and one of them occurs at a subunit interface (WILSON et al. 1981). Approximately half of the carbohydrate associated with neuraminidase is located in the N-terminal membrane-bound stalk (ALLEN et al. 1977; WARD et al. 1983b), where it might, at least in part, be involved in interchain contacts.

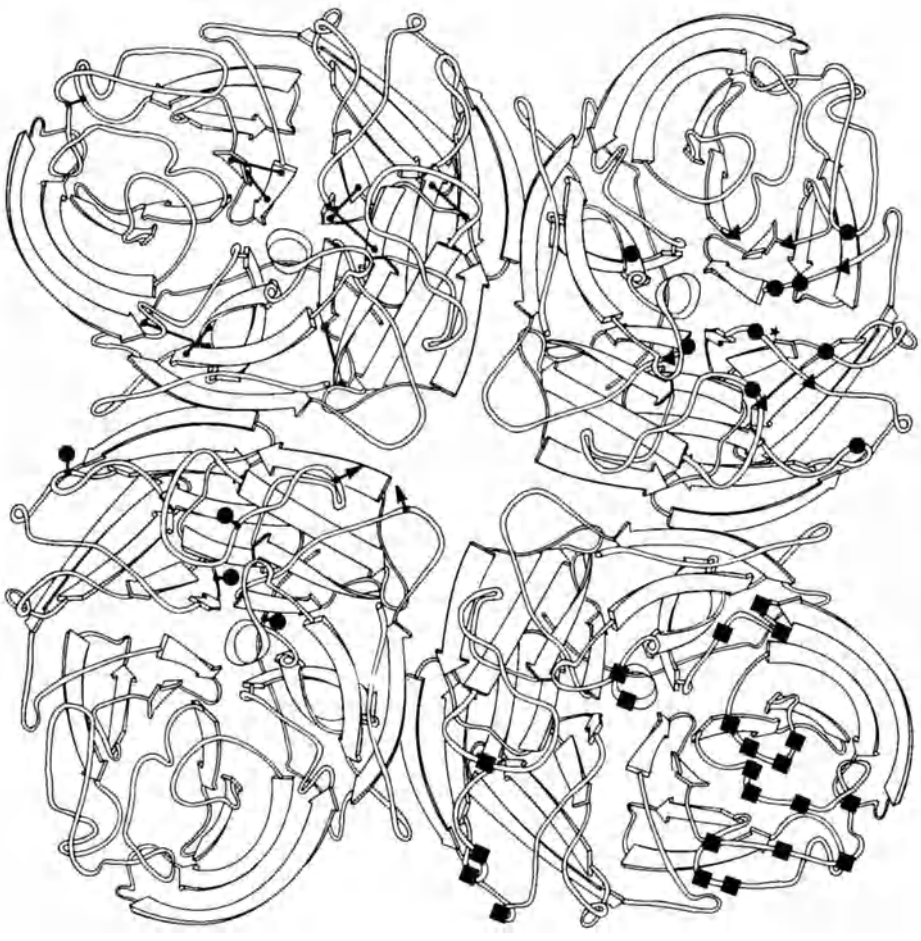


Fig. 6. The neuraminidase tetramer showing, anticlockwise from top left, disulfide bonds, carbohydrate attachment sites and putative calcium ligands, selected regions of upper surface variation, and conserved active site residues surrounding the sialic acid binding site (*)

It is not yet possible to describe in detail the polypeptide backbone conformation around the glycosylation sites of neuraminidase. Asparagine residues 86, 146, and 200 are in chain segments with extended configurations, while residue 234 is in a bend (Fig. 6).

The glycosylation triplet from Asn 402 contains no oligosaccharide (WARD et al. 1982). The chain configuration here is a turn which exposes the middle residue of the triplet (Trp 403) in a way that would hinder simultaneous recognition of Asn 402 and Ser 404 by transferases. Apparently, this structure is formed before glycosylation can occur. The transfer of sugar to protein is cotranslational (ROTHMAN and LODISH 1977; BERGMAN and KUEHL 1978), although in some cases transfer may occur to the native folded polypeptide chain. In vitro glycosylation of nonglycosylated sequences is possible after denaturation (STRUCK and

LENNARZ 1980). Clearly, the three-dimensional protein structure may influence the accessibility of a glycosylation triplet, even if the nascent polypeptide chain is the recipient of the oligosaccharide. Only 49 out of 159 tripeptide glycosylation sequences analyzed in eukaryotes were actually glycosylated (STRUCK and LENNARZ 1980). In this respect, glycosylation sequences near the protein C-terminus may be more at risk than early sequences. The distance lag between synthesis and sugar transfer is believed to be 45–80 residues (ROTHMAN and LODISH 1977; HUBBARD and IVATT 1981).

The structure of the sugar at Asn 146 is especially interesting for two reasons. First, this location is the most common glycosylation site in known neuraminidase amino acid sequences, and secondly, carbohydrate is implicated, directly or otherwise, in enzyme activity (GRIFFIN et al. 1983). A thirtyfold decrease in the activity of both neuraminidase and hemagglutinin has been reported for influenza virions with nonglycosylated surface glycoproteins (BASAK and COMPANS 1983).

Influenza B strain neuraminidase has no glycosylation sequences on the N-terminal arm on the bottom surface of the subunit. However, a compensating sequence is found nearby on the bottom surface at residue 285 (N2 numbering). Thus, N1, N2, and B neuraminidase structures all have two potential sites for carbohydrate attachment on the bottom face of the subunit.

4.5 The Enzyme Active Center

The active site of neuraminidase has been located by soaking sialic acid, a known inhibitor, into the crystals and using difference Fourier methods to determine the binding site (COLMAN et al. 1983). As seen looking into the top face of the subunit, the binding site is almost directly above the first strands of sheets 3 and 4 (β_3S_1, β_4S_1) and is located in a large pocket or depression in the surface of the molecule (Fig. 7). Like other enzymes catalyzing the removal of end groups rather than internal chain cleavage, the morphology of the site is more like a pocket than a cleft. The walls of the pocket are lined by residues from the lopps β_1L_{01}, β_1L_{23}, β_2L_{23}, β_3L_{01}, β_3L_{23}, β_4L_{01}, β_4L_{23}, and β_5L_{01}. On these loops there are an extraordinarily large number of conserved, charged amino acids whose orientations are seen to be toward the sialic acid binding site (Fig. 7).

The binding of heavy metals to some of these residues during the multiple isomorphous replacement phasing procedure gives some indication of their reactivity. Diaminodinitroplatinum binds to His 150, and mercury phenyl-glyoxal binds to Arg 224, supporting the observed inhibition of neuraminidase by phenylglyoxal derivatives (EDMOND et al. 1966). Chemical modification of tryptophan also abolishes activity (BACHMEYER 1972). The role of some sulfhydryl reagents as inhibitors is presumably an indirect one and may result from opening up the disulfide linkages in the stepladder between β_4S_1 and β_4S_2. There is so far no evidence for the presence of a calcium ion in the active site, although this cannot yet be excluded. If indeed there is none, then the inhibitory role of ethylenediaminetetraacetate (EDTA) on enzyme activity must also be indirect,

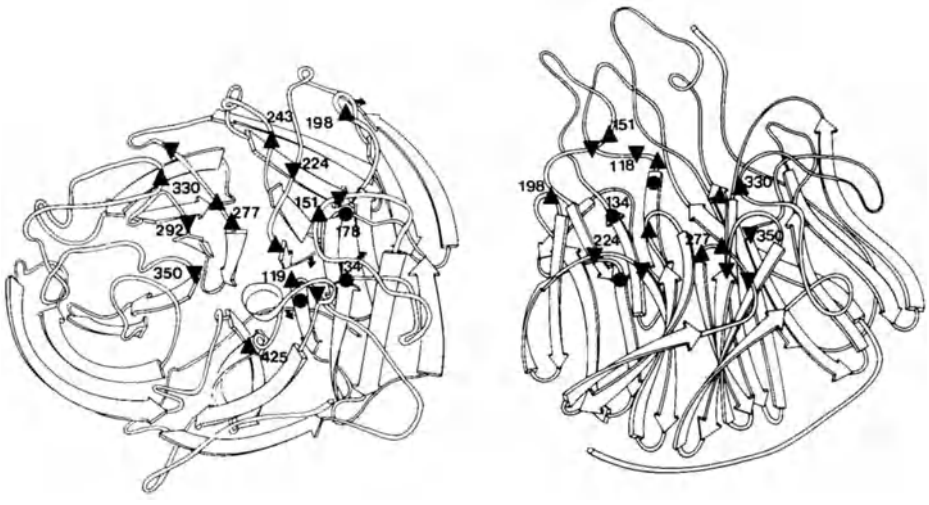


Fig. 7. Conserved amino acid residues surrounding the sialic acid binding site in N2 neuraminidase. ▲ glu 119, asp 151, asp 198, glu 227, asp 243, glu 276, glu 277, asp 330, glu 425. ▼ arg 118, arg 152, arg 224, his 274, arg 292, lys 350. ● tyr 121, leu 134, trp 178

because the putative calcium site on the fourfold symmetry axis is approximately 30 Å distant from the sialic acid binding site.

Sigmoid kinetics have been reported for influenza (MOUNTFORD et al. 1982) and human brain (TETTAMANTI et al. 1978) neuraminidase. The four active centers of the influenza enzyme are ~40 Å apart and thus, like other proteins showing interdependent ligand binding to identical subunits, the effect would have to be mediated in some way through the subunit interface. No gross reorganization around this interface occurs when sialic acid binds, although the binding affinity in that case is weak (~1 m*M*) and possibly insufficient to trigger a quaternary structure change.

There are some superficial similarities between the active site of neuraminidase and the receptor binding pocket of hemagglutinin (WILSON et al. 1981), both of which contain Tyr 98, His 193, Glu 190, Trp 153, and Leu 194. High-resolution studies of sialic acid binding to both proteins are required to determine whether or not the binding modes share common features. It has been reported (ROGERS et al. 1983) that the single amino acid change of Gln to Leu at residue number 226 in hemagglutinin changes the specificity from NeuAcα2→3 Gal- to NeuAcα2→6 Gal-linkages. Influenza neuraminidases cleave preferentially, but not exclusively, the α2→3 linkage (CORFIELD et al. 1982, 1983; PAULSON et al. 1982).

4.6 Subunit Interfaces

The arrangement of subunits in the tetramer is shown diagrammatically in Fig. 6. The single subunit interface consists primarily of residues from the out-

side strands (S4) of sheets β_6, β_1, and β_2; the C-terminal strand; and the loops L_{01}, L_{23}, and L_{34} on β_1 (VARGHESE et al. 1983). Hydrophobic and hydrophilic interactions are found across the interface. Two possibilities are envisaged at present for continuing the sheet secondary structure across the interface. Both include $\beta_1 S_4$, in one case in conjunction with $\beta_6 S_4$ and in the other with $\beta_1 L_{34}$, a part of which may be antiparallel to $\beta_1 S_4$ on a neighboring subunit. Around the fourfold symmetry axis, His 168 makes contact with its neighbors in a way which is unique to N2 structures. Nearer the upper surface of the head, eight acidic groups (Asp 113, Asp 141 on all four subunits; Fig. 6) cluster around a site which can be labeled by Sm^{3+}. Calcium ions are presumed to be bound here in the native enzyme, by analogy with other systems (COLMAN et al. 1972) and in keeping with observations of calcium requirements for neuraminidase stability (BAKER and GANDHI 1976; CARROLL and PAULSON 1982). The carbohydrate at Asn 200 interacts with surface residues of the neighboring subunit, thereby contributing to the stabilization of the tetramer (VARGHESE et al. 1983).

5 Antigenic Structure

Antigenic variation in influenza virus is confined to the two surface glycoproteins hemagglutinin and neuraminidase. It involves major changes (antigenic shift) to new subtypes at irregular intervals, where one or both coat proteins are altered and minor changes (antigenic drift) within subtypes, where both coat proteins change gradually. Since 1933, major antigenic shifts have occurred in 1957 (Asian flu), 1968 (Hong Kong flu), and 1977 (Russian flu).

The antigenic shift in 1957 saw a change in both hemagglutinin (H1 to H2) and neuraminidase (N1 to N2) as a result of genetic reassortment (SCHOLTISSEK et al. 1978a), and these coat proteins accumulated additional changes by antigenic drift in the period 1957–1968. In 1968, the Hong Kong influenza subtype appeared in man following genetic reassortment of just the hemagglutinin gene (LAVER and WEBSTER 1973; SCHOLTISSEK et al. 1978a; WARD and DOPHEIDE 1981b). The Hong Kong subtype retained the N2 neuraminidase, and this has continued to undergo further antigenic drift since 1968. In 1977, the Russian influenza subtype appeared, which very closely resembled the genetic constitution of the H1N1 viruses prevalent in man in the early 1950s (SCHOLTISSEK et al. 1978b; NAKAJIMA et al. 1978).

The structural changes associated with antigenic shift and drift in hemagglutinin have been established by protein sequence, nucleotide sequence, and X-ray crystallographic analyses of several natural and artificially selected influenza variants (see WARD 1981 for review; GREEN et al. 1982; CATON et al. 1982; ROGERS et al. 1983). In a similar way, the structural data described in this review have been determined to establish the chemical basis of antigenic variation in the neuraminidase of influenza virus.

Three approaches can be adopted to characterize the number and nature of antigenic determinants in neuraminidase. The first is the examination of the antigenic and immunogenic properties of neuraminidase fragments produced

by peptide synthesis or chemical and enzymatic cleavage. The second is the examination of antigenic variants with monoclonal and polyclonal antibodies. A third approach is to examine the antigenic and immunogenic properties of chemically modified, artificially selected, or naturally occurring variants and to correlate these changes in antigenicity with specific changes in their primary and tertiary structure.

5.1 Studies with Neuraminidase Fragments

Early studies of detergent-released and protease-solubilized neuraminidase demonstrated that the enzyme activity and major antigenic determinants were associated with the globular heads (SETO and ROTT 1966; LAVER 1978). These heads consist of the 400 C-terminal residues of the molecule (WARD et al. 1982).

These protease-released heads can be further broken down by cyanogen bromide cleavage to yield disulfide-bonded fragments that can be tested for antigenicity. In studies on the disulfide bond arrangements of A/Tokyo/3/67 N2 neuraminidase heads, WARD et al. (1983a) obtained three large fragments which were readily separated by gel filtration. The largest fragment consisted of CN1 (residues 74–162) linked to CN4 (residues 363–467), the second peak contained CN2 (residues 163–308) with five intrafragment disulfide bonds, and the third peak contained CN3 (residues 309–362) with one intrafragment disulfide bond.

This simple separation provided an easy means of generating large disulfide-bonded fragments which, by analogy with the studies on hemagglutinin (JACKSON et al. 1979), would be expected to retain some of their antigenic characteristics. Studies are in progress (D.C. JACKSON et al., unpublished) to establish the antigenic and immunogenic properties of these N2 neuraminidase fragments.

An examination of the distribution of methionine residues in N1 and type B neuraminidase (Sect. 5.2) and the predicted locations of the disulfide bonds in these molecules shows that cyanogen bromide cleavage would yield three fragments from N1 neuraminidase (two of which would be difficult to separate) and a complex mixture of four major fragments and four smaller fragments from the B/Lee/40 protein.

5.2 Amino Acid Sequence Changes and Antigenic Drift

5.2.1 Variants Selected with Monoclonal Antibodies

Monoclonal antibodies have been used in two major studies aimed at delineating antigenic maps of the N2 neuraminidases of A/Tokyo/3/67 (WEBSTER et al. 1982; LAVER et al. 1982; JACKSON and WEBSTER 1982) and A/RI/5^+/57 (WEBSTER et al. 1984).

In the first study, monoclonal antibodies were raised against five strains of N2 neuraminidase, A/Jap/57, A/Tokyo/3/67, A/Port Chalmers/1/73, A/Victoria/3/75, and A/Texas/1/77, and used to examine the extent of antigenic drift

Table 15. Inhibition of field strain neuraminidases with monoclonal antibodies (WEBSTER et al. 1982)

Monoclonal antibody to neuraminidase (N2) of		Jap/57	Sing/57	Okuda/57	RI/5$^+$/57	Eng/61	Nth. Carolina/63	Ned/63	Taiwan/64	Tokyo/67	Hong Kong/68	Eng/72	Udorn/72	Port Ch/73	Vic/75	Tex/77	Bang/79
Jap/305/57	78/4	+	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−
	102/2	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	113/2	+	−	−	+	+	−	+	+	+	+	−	−	−	−	−	−
	117/2	+	−	+	−	−	−	−	−	−	−	−	−	−	−	−	−
	136/5	+	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−
	152/6	+	+	+	+	−	−	−	−	−	−	−	−	−	−	−	−
Tokyo/3/67	S10/1	−	−	−	−	+	+	+	+	+	+	+	−	−	−	−	−
	S25/3	+	−	−	−	+	−	+	+	+	+	+	−	−	−	−	−
	25/4	+	+	−	−	+	+	+	+	+	+	+	−	−	−	−	−
	S32/3	−	−	−	−	+	+	+	+	+	+	+	−	−	−	−	−
	16/8	−	−	−	−	+	+	+	+	+	+	+	−	−	−	−	−
	23/9	−	−	−	−	+	+	+	+	+	+	−	−	−	−	−	−
Port Chalmers/1/73	27/5	−	−	−	−	−	−	−	−	−	−	+	+	+	+	−	−
Vic/3/75	12/2	−	−	−	−	−	−	−	−	−	−	+	+	+	+	−	−
	21/3	−	−	−	−	−	−	−	−	−	−	+	+	+	+	−	−
	22/3	−	−	−	−	−	−	−	−	−	−	+	+	+	+	−	−
Texas/1/77	18/1	−	−	−	−	−	−	−	−	+	+	+	+	+	+	+	+
	19/1	−	−	−	−	−	−	−	−	−	−	−	+	+	−	+	+
	67/1	+	−	−	−	−	−	+	+	+	+	+	+	+	+	+	+
	69/1	−	−	−	−	−	−	−	−	−	−	−	+	+	+	+	+
	78/1	−	−	−	−	−	−	−	−	−	−	−	−	+	+	+	+
	88/1	−	−	−	−	−	−	−	−	−	−	+	−	+	+	+	+
	123/1	−	−	−	−	−	−	−	−	−	−	−	+	+	+	+	+

+ greater than 90% inhibition; − no inhibition

in N2 neuraminidase from 1957 to 1979 (WEBSTER et al. 1982). As shown in Table 15, there was a progressive change in the antigenic profiles of these N2 field strains. The antigenic regions recognized by the monoclonal antibodies raised against A/Jap/57 had all changed by 1963, with the exception of that recognized by 113/2, which remained until 1968. The antibodies raised against A/Tokyo/67 did not in general recognize field strains isolated before 1961 or after 1972. The monoclonal antibodies produced against Port Chalmers/1/73 and Vic/3/75 behaved identically and bound to a restricted set of field strains isolated between 1972 and 1975. The antibodies raised against A/Texas/1/77 generally recognized strains isolated between 1972 and 1979, although two of these, 18/1 and 67/1, recognized determinants that were present in earlier N2 field strains.

The data from this study (WEBSTER et al. 1982) also indicated that the neuraminidase molecules on different influenza viruses isolated in 1957, at the beginning of the Asian flu pandemic, were antigenically heterogeneous and showed

different reactivity patterns with the panel of 23 monoclonal antibodies (Table 15).

While the data in Table 15 show evidence of significant antigenic drift in human influenza field strains, only limited antigenic variation was detected in the N2 neuraminidase of avian strains over a similar time period, all of the avian N2 neuraminidases being closely related to A/Japan/305/57 (WEBSTER et al. 1982).

In the second part of this analysis, WEBSTER et al. (1982) showed for the first time that these antineuraminidase monoclonal antibodies could be used to select antigenic variants. The antibodies used had been raised against A/Jap/305/57, A/Tokyo/3/67, or A/Texas/1/77 (see Table 16) and were all capable of inhibiting the neuraminidase of A/Tokyo/3/67 (Table 15). WEBSTER et al. (1982) showed that the frequency of variant selection in neuraminidase did not differ from that observed in hemagglutinin (LAVER et al. 1979), suggesting that antigenic variants in neuraminidase were present in all influenza virus preparations.

When these antigenic variants were tested for their ability to react with the panel of monoclonal antibodies, the latter could be sorted into three groups based on neuraminidase inhibition assays. As shown in Table 16, the variants no longer reacted with the antibody used for their selection and in general did not react with related antibodies from the same group. The exceptions were the group 1 variants V1, selected with S10/1, and V1 and V3, selected with 23/9, which did react strongly with some of the other group 1 monoclones. Most variants reacted to the same titer as the parent virus when tested with monoclonal antibodies from the other two groups. The only exceptions to this were the group 1 variants V2 and V3, selected with monoclone 23/9, which no longer reacted with the anti-Texas antibody 18/1 or in the case of V2 with the anti-Jap/57 antibody 113/2.

When the antigenic variants were screened by Elisa assay with the same panel of monoclonal antibodies shown in Table 16, there were some differences from the reactivity patterns obtained with the neuraminidase inhibition assay (WEBSTER et al. 1982). Some variants bound antibody, as judged by Elisa assay, but their enzyme activity was not inhibited. Using Elisa assays, the monoclonal antibodies S10/1, S25/3, S25/4, and S32/3 were assigned into one subset group Ia, while antibodies 16/8 and 23/9 were assigned to group Ib (WEBSTER et al. 1982).

Competitive radioimmunoassays of the anti-Jap/57 antibody 113/2 and the group I monoclones S10/1, S25/3, 25/4, S32/3, 16/8, and 23/9 confirmed the existence of at least three overlapping antigenic regions on A/Tokyo/3/67 neuraminidase (JACKSON and WEBSTER 1982). One site is that recognized by the four antibodies S25/3, 25/4, S32/3, and 23/9. This region is close to the enzyme active site, as all four antibodies inhibit neuraminidase activity against fetuin (molecular weight 50000) and the small substrate *N*-acetyl-neuraminyllactose (molecular weight 600). A second site is that recognized by the anti-Jap/57 monoclone 113/2. It is adjacent to the region recognized by antibodies S25/3, 25/4, S32/3, and 23/9, as judged by their ability to compete partially with monoclone 113/2. The group 1 monoclones S10/1 and 16/8 did not interfere with

Table 16. Neuraminidase inhibition of Tokyo/67 variants selected with monoclonal antibodies (WEBSTER et al. 1982)

Monoclonal antibody		Parent virus	Group II	Group I										Group III	
			Jap 113/2	S10/1	S25/3	25/4	S32/3		16/8		23/9			Tx18/1	Tx67/1
		(Tok/67)	V1	V1	V1	V1	V1	V2	V1	V2	V1	V2	V3	V1	V1
Anti-Jap/57	113/2	+	−	+	+	+	+	+	+	+	+	−	+	+	+
Anti-Tokyo/67	S10/1	+	+	−	−	−	−	−	−	−	+	−	−	+	+
	S25/3	+	+	−	−	−	−	−	−	−	+	−	+	+	+
	25/4	+	+	+	−	−	−	−	−	−	−	−	−	+	+
	S32/3	+	+	+	−	−	−	−	−	−	−	−	−	+	+
	16/8	+	+	+	−	−	−	−	−	−	+	−	−	+	+
	23/9	+	+	+	−	−	−	−	−	−	−	−	−	+	+
Anti-Texas/77	Tx18/1	+	+	+	+	+	+	+	+	+	+	−	−	−	−
	Tx67/1	+	+	+	+	+	+	+	+	+	+	+	+	−	−

+ neuraminidase inhibited; − neuraminidase not inhibited

the binding of 113/2 and thus recognize a third antigenic region remote from the binding site for the group II monoclone 113/2, but overlapping with the regions recognized by the other group I antibodies. The two sites recognized by 113/2 and the group I pair S10/1, 16/8 are further away from the enzyme active center than the region recognized by the other Group I antibodies, as 113/2, S10/1, and 16/8 did not inhibit the hydrolysis of the small synthetic substrate neuraminyllactose but did inhibit hydrolysis of the fetuin (JACKSON and WEBSTER 1982). No competitive studies or differential substrate utilization analyses were carried out with the group III anti-Texas monoclones.

In the final stage of this antigenic analysis of A/Tokyo/3/67 neuraminidase, the antigenic variants shown in Table 16 were analyzed for amino acid sequence changes by comparative peptide mapping and amino acid analysis (LAVER et al. 1982). The results are shown in Table 17. Four of the variants selected with Group I monoclones showed the same single amino acid sequence change of Arg to Ile at position 344. These variants all behaved identically in the original screening with monoclonal antibodies (Table 16). The group I variant selected with S10/1 antibody did not exhibit this change but had a different amino acid substitution Lys to Glu at a different locus, residue 368. This residue is thus on the periphery of the antigenic region recognized by the group I monoclones, as the Lys to Glu change does not affect the binding of monoclones 25/4, S32/3, 16/8, and 23/9 (Table 16).

The variants selected with group II and group III monoclonal antibodies (see Table 16) have also been examined for sequence changes, but so far only that present in the 18/1 variant V1 has been identified. In this variant, the Asn residue at position 221 changes to His. This locus is far removed in the linear sequence from the two positions that change in the group I variants. The three-dimensional structure (Sect. 4) shows that residue 221 is located on one side of the top surface of the neuraminidase monomer in loop L_{01} of sheet 3, while residues 344 and 368 are on the other side of the top surface of the monomer on loops L_{01} and L_{23} of sheet 5.

Recently the neuraminidase gene sequence of A/Tokyo/3/67 has been determined, and further monoclonal variants of that protein have been characterized (LENTZ et al. 1984). Single amino acid sequence changes in those variants are found at Arg 344, changing to Lys, Gly, Thr, and Ser in four different variants, and at Arg 253, changing to Ser. All of the Tokyo variants selected by monoclonal antibody to Tokyo neuraminidase show changes at either 344 or 368. Monoclonal antibody to Texas/77 selects the 221 variant and to Jap/57, the 253 variant.

The monoclonal antibodies used to map the antigenic regions of A/Tokyo/67 neuraminidase were all selected on the basis of neuraminidase inhibition, and the sites they recognize all appear to be on the top surface of the molecule. A more extensive panel of 40 monoclonal antibodies has been produced recently against the early N2 strain A/RI/5^+/57 (WEBSTER et al. 1984) to map additional sites on neuraminidase.

As shown in Table 18, competitive radioimmunoassays indicate four antigenic regions on A/RI/5^+/57 neuraminidase. The group 1 monoclonal antibodies do not interfere with the binding of group 4 antibodies; the reverse

Table 17. Amino acid sequence changes in Tokyo/67 antigenic variants selected with monoclonal antibodies (LAVER et al. 1982)

Group I *Anti-Tokyo/67 Variants*

Tokyo/67	Asn	Pro	Asn	Asn	Glu	Arg	Gly	Thr	Gln	Gly	Val	Lys	Gly	Trp	Ala	Phe	Asp	Asn	Gly	Asn	Asp	Leu	Trp	Met	Gly	Arg	Thr	Ile	Ser	Lys	Asp	Leu	Arg
S10/1 V1	Asn	Pro	Asn	Asn	Glu	Arg	Gly	Thr	Gln	Gly	Val	Lys	Gly	Trp	Ala	Phe	Asp	Asn	Gly	Asn	Asp	Leu	Trp	Met	Gly	Arg	Thr	Ile	Ser	**Glu**	Asp	Leu	Arg
S25/3 V1	Asn	Pro	Asn	Asn	Glu	**Ile**	Gly	Thr	Gln	Gly	Val	Lys	Gly	Trp	Ala	Phe	Asp	Asn	Gly	Asn	Asp	Leu	Trp	Met	Gly	Arg	Thr	Ile	Ser	Lys	Asp	Leu	Arg
S25/4 V1	Asn	Pro	Asn	Asn	Glu	**Ile**	Gly	Thr	Gln	Gly	Val	Lys	Gly	Trp	Ala	Phe	Asp	Asn	Gly	Asn	Asp	Leu	Trp	Met	Gly	Arg	Thr	Ile	Ser	Lys	Asp	Leu	Arg
S32/3 V1	Asn	Pro	Asn	Asn	Glu	**Ile**	Gly	Thr	Gln	Gly	Val	Lys	Gly	Trp	Ala	Phe	Asp	Asn	Gly	Asn	Asp	Leu	Trp	Met	Gly	Arg	Thr	Ile	Ser	Lys	Asp	Leu	Arg
V2	Asn	Pro	Asn	Asn	Glu	**Ile**	Gly	Thr	Gln	Gly	Val	Lys	Gly	Trp	Ala	Phe	Asp	Asn	Gly	Asn	Asp	Leu	Trp	Met	Gly	Arg	Thr	Ile	Ser	Lys	Asp	Leu	Arg
		340					345					350					355					360					365					370	

Group III Anti-Texas/77 Variants

Tokyo/67	Leu	Val	Asp	Ser	Ile	Gly	Ser	Trp	Ser	Gln	Asn	Ile	Leu	Arg
Tx18/1 V1	Leu	Val	Asp	Ser	Ile	Gly	Ser	Trp	Ser	Gln	**His**	Ile	Leu	Arg
Tx67/1 V1	Leu	Val	Asp	Ser	Ile	Gly	Ser	Trp	Ser	Gln	Asn	Ile	Leu	Arg
	211				215					220				224

Sequence change not found.

Table 18. Topographical mapping of N2 (A/RI/5$^+$/57) neuraminidase using monoclonal antibodies in competitive radioimmunoassays (WEBSTER et al. 1984)

Radio-labeled IgG	Competition with the following monoclonal antibodies 1						2																											3		4				
	254/1	239/3	552/5	618/1	562/1	129/2	11/1	193/2	474/1	490/3	415/1	220/2	103/1	561/3	443/2	307/2	93/1	73/1	132/6	81/4	438/1	514/1	101/5	664/1	112/2	342/4	513/1	145/1	509/1	266/4	480/2	560/1	233/3	72/1	100/1	190/1	441/2	157/1	740/1	658/1
254/1	+	+	+	+	+	+								−						−																				
239/3	+	+	+	+	+	+								−															−											
509/1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		+	+	+	+	+	+	+	+	+	+	+	+	+	−	+	+					
145/1						−	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	−	−	+	+					
342/4							+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	−	−	−	−	−	−	−	−	−					
474/1							+	+	+	+	+	+	+	+	+	+	+	+	+		+	+	+	−	−	−				−	−		+	−	−					
561/3	−	−				−	+	+	+	+	+	+	+	+	−	−	−	−			−	−	−	−	−								−	+	+					−
220/2							+	+	+	+	+	+	+	+	+	+	−	−			−	−	−	−	−	−				−			−	−	−					
560/1							+	−	−	−	−	−	−	−	−	−	+	+	+	+	−	−	−	−	−	−		−		−	−		−	−	−					
72/1																																		+	+	−	−	−	−	−
100/1																																		+	+	−	−	−	−	−
157/1																																		+	+	+	+	+	+	+
658/1																																		−	−	−	−	−	+	+

+ the binding of radiolabeled IgG was completely inhibited; − the binding of radiolabled IgG was significantly inhibited, but 100% competition was not obtained with the concentrations of unlabeled antibody used as competitor; no entry indicates no competition

and vice versa, suggesting these regions were topographically well separated. The group 3 antibodies recognize sites close to the group 4 antibodies, but again distant from group 1. This is apparent from the strong competition between group 3 and 4 monoclones, contrasted with the lack of competition between the antibodies in groups 1 and 3. The antigenic regions recognized by the large panel of group 2 antibodies overlap extensively with the group 3 regions, partially with the group 1 regions, and in only one instance is there some competition with a group 4 antibody (Table 18).

With regard to their effect on enzyme activity, the antibodies in groups 2 and 3 show strong inhibition of neuraminidase, with fetuin acting as substrate, while antibodies in groups 1 and 4 inhibit weakly or not at all (Table 19). A few group 2 monoclones also inhibit neuraminidase activity on the small substrate *N*-acetyl-neuraminyllactose.

With regard to their effect on virus replication, all of the monoclonal antibodies inhibit virus release from MDCK cells when incorporated in an agar overlay, but only those in groups 2 and 3 can effectively neutralize infectivity and permit the selection of antigenic variants (WEBSTER et al. 1984).

As shown in Table 18, the majority of monoclonal antibodies to A/RI/5$^+$/57 neuraminidase are assigned to reactivity group 2. WEBSTER et al. (1984) suggested that this is because the region recognized by these monoclones is either the most immunogenic region of the molecule or a topographically large area. In order to further delineate the group 2 antigenic region, WEBSTER et al. (1984) selected antigenic variants with 14 different group 2 antibodies and examined their reactivities with a panel of 27 group 2 and two group 3 antibodies. As shown in Table 20, the antibodies in group 2 can be subdivided into four subgroups 2a, 2b, 2c, and 2d.

From the reactivity profiles and biological properties of anti-A/RI/5$^+$/57 monoclonal antibodies, WEBSTER et al. (1984) proposed that antigenic regions 1 and 4 are spatially separated and located on the bottom surface of the molecule, while region 3 is on the top surface but at some distance from the catalytic center. They suggested that region 2, which is equivalent to regions 1, 2, and 3 described in the antigenic analysis of A/Tokyo/3/67 neuraminidase (WEBSTER et al. 1982; LAVER et al. 1982; JACKSON and WEBSTER 1982; LENTZ et al. 1984), encompasses most of the top surface of the molecule. Subregion 2d is thought to be closest to the catalytic center with subregions 2a and 2b adjacent to it.

When tested against four N2 field strains (Table 21), it could be seen that region 1 did not undergo significant antigenic drift until after 1968, while regions 2, 3, and 4 had all changed by 1967 (WEBSTER et al. 1984).

Monoclonal variants of A/RI/5$^+$/57 neuraminidase are now being characterized, and changes have been determined at Asn 334 (to Ser), His 150 (to Asn), and Trp 403 (to Arg) (G.M. Air, personal communication).

5.2.2 Sequence Changes in N1 Field Strains

Sequence studies on antigenic drift in the N1 subtype have been restricted to partial sequences (206–287 nucleotides long) from the 3′ end of the vRNA

Table 19. Effect of monoclonal antibodies on neuraminidase activity with fetuin and *N*-acetyl-neuraminyllactose (WEBSTER et al. 1984)

Antibody group	Monoclonal antibody	Neuraminidase inhibition with following substrates	
		Fetuin	NAL
1	254/1	+	−
	239/3	+/−	−
	552/6	tr	−
	618/1	+/−	−
	562/1	+/−	−
	129/2	+	−
2A	81/4	++	−
	93/1	++	+/−
2B	438/1	++	−
	415/1	++	−
	443/2	++	+/−
	193/2	++	−
	474/1	++	−
	490/3	++	−
	220/2	++	+/−
	342/4	++	+/−
	101/5	++	−
	664/1	++	−
	513/1	++	−
	514/1	++	−
	561/3	++	−
	11/1	++	+/−
	112/2	++	+/−
	103/1	++	−
	233/3	++	−
	132/6	++	−
2C	73/1	++	−
	560/1	++	−
2D	145/1	++	+/−
	509/1	++	+
	266/4	++	+
	480/2	++	+/−
	307/2	++	+/−
3	72/1	++	−
	100/1	++	−
4	190/1	tr	−
	441/2	−	−
	157/1	tr	−
	740/1	+/−	−
	658/1	+/−	−

++ complete inhibition (100%); + more than 50% inhibition (50%–80%); +/− less than 50% inhibition (30%–40%); tr less than 20% inhibition (<20%); − no inhibition

Table 20. Operational antigenic map of groups 2 and 3. (WEBSTER et al. 1984)

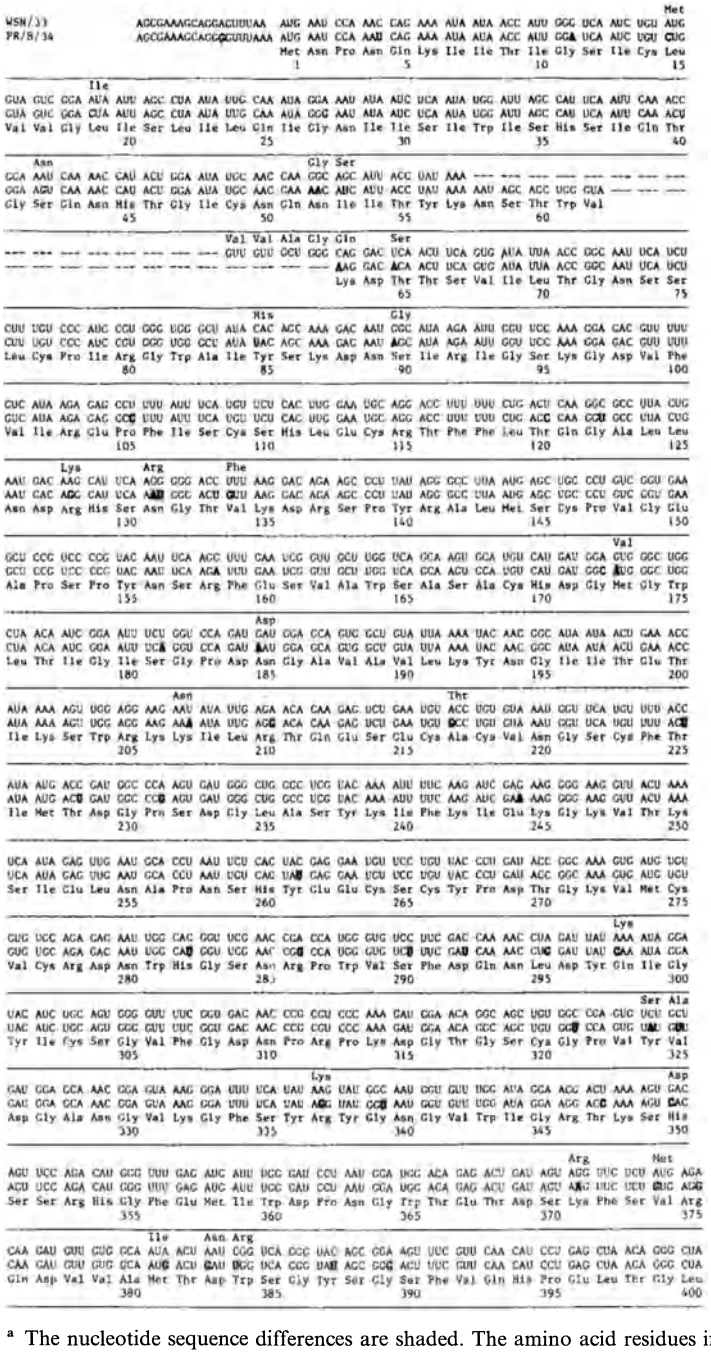
Monoclonal variants	Reactivity patterns of the monoclonal antibodies to N2 in ELISA and neuraminidase inhibition																												
	2A		2B																		2C		2D					3	
	81/4	93/1	438/1	415/1	443/2	193/2	474/1	490/3	220/3	342/4	101/5	664/1	513/1	514/1	561/3	11/1	112/2	103/1	233/3	132/6	73/1	560/1	145/1	509/1	266/4	480/2	307/2	72/1	100/1
V81/4	+	−																											
V514/1			+	+	+	+	+	+	+	+			+	+															
V438/1			+	−	+	−	+		+	+				−															
VM664/1			+	+	+	+	+	+	+	−		+			+				+										
VM415/1	−		−	−		−	+	+	+	−	+	−		+	−				−										
V561/3			+	+	+	+	+	+	−		+	+			+				−										
V112/2			+	+	+	+	+	+	+	+	+	+	+	+		+	+	+		−									
VM193/2			+	+	+	+	+	+	−		−		+			+													
VM220/2			+	+	+	+	+	−	+	+			+	+															
VM474/1			+	+	+	+	+	+	+	+	+	+	+	+	−	+	+	+											
VM490/3			+	+	+	+	+	+	+	+		+	+	+	−	+	+	+											
V509/1																			−					+	+			−	−
V145/1																							+	+	+	+	+	−	−
V480/2																							+	+	+	+			

V, variants selected in eggs; VM variants selected in MDCK cells; + binding of the monoclonal antibodies to variant virus in ELISA was decreased to below control binding and neuraminidase inhibition activity was completely inhibited using dilutions of antibody that were highly reactive with the parent X-7F1 virus; − neuraminidase inhibition activity abolished, binding to variant in ELISA was no different or only slightly decreased compared with binding to X-7F1; no entry indicates that the virus is indistinguishable from the parental X-7F1

Table 21. Reactivity of monoclonal antibodies with different influenza virus strains (WEBSTER et al. 1984)

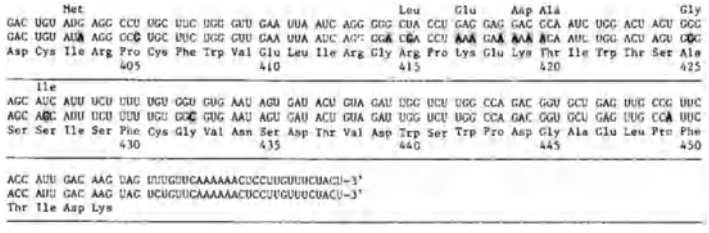
Antibody group	Monoclonal antibody	Binding of monoclonal antibodies in ELISA to the following viruses			
		X-7F1/57	NWS-Tokyo/67	Aichi/68	Udorn/72
1	254/1	+	+	+	−
	239/3	+	+	+	+
	552/5	+	+	+	−
	618/1	+	+	+	−
	562/1	+	+	+	−
	129/2	+	+	+	−
2	27 antibodies	+	−	−	−
3	2 antibodies	+	−	−	−
4	5 antibodies	+	−	−	−

Table 22. Comparison of nucleotide and amino acid sequences of A/WSN/33 (HITI and NAYAK 1982) and A/PR/8/34 (FIELDS et al. 1981)[a]



[a] The nucleotide sequence differences are shaded. The amino acid residues in A/WSN/33 that differ from A/PR/8/34 are indicated above the gene sequence

Table 22 (continued)



neuraminidase gene segment of some 12 N1 strains (BLOK and AIR 1980, 1982a), complete sequences for the neuraminidase genes of A/PR/8/34 (FIELDS et al. 1981) and A/WSN/33 (HITI and NAYAK 1982), and comparative peptide map analyses for A/Bel/42 (BLOK et al. 1982; BLOK and AIR 1982a). The sequence changes in the N-terminal transmembrane and stalk regions have already been discussed in Sect. 2.3.1.

Only one unequivocal change is indicated by the peptide map comparison of A/PR/8/34 and A/Bel/42 (BLOK and AIR 1982a), the change from Ser 90 in the 1934 strain to Gly 90 in 1942. Substantial amino acid sequence changes were observed in the tryptic peptides 376–413 and 420–454.

The complete neuraminidase sequences for A/WSN/33 and A/PR/8/34 are shown in Table 22. The WSN virus is not a naturally occurring field strain. Rather it has been selected by multiple passages of the prototype human influenza, WS/33, in mouse brain. Similar attempts to select a neurovirulent derivative of PR/8/34 were not successful (FRANCIS and MOORE 1940). A/WSN/33 is the only human influenza virus capable of multicycle replication in MDBK cells (CHOPPIN 1969), and this uniqueness is a property of viral neuraminidase, not hemagglutinin (SCHULMAN and PALESE 1977; SUGIURA and UEDA 1980; NAKAJIMA and SUGIURA 1980). Thus sequence changes shown in Table 22 reflect not only antigenic drift between 1933 and 1934 but also the biological distinctiveness of WSN/33. The alignment at the C-terminal end of the stalk region, Lys 63 in PR/8/34, follows the block deletion study of BLOK and AIR (1982c). Eleven amino acid changes occur in the stalk and membrane regions of the sequence, and 27 occur in the heads. In particular, the glycosylated residue Asn 131 changes to arginine in the WSN sequence. This is the only neuraminidase sequence in which this glycosylation triplet is not preserved. After allowing for the 15-residue stalk deletion, it can be seen that this is the analogue of Asn 146 in N2 strains. With the exception of segments 324–325, 371–374, 381–384, 415–420, and 425–427, the changes are not clustered but distributed throughout the sequence. Their location in the three-dimensional structure is discussed in Sect. 5.5.2.

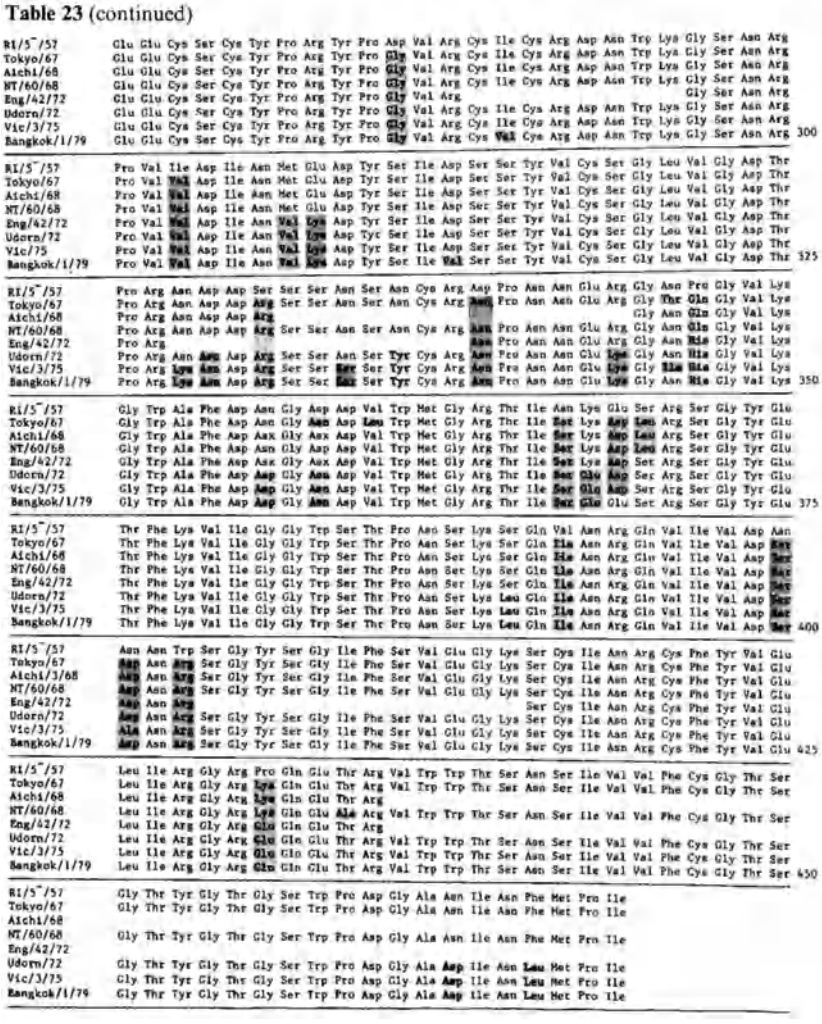
5.2.3 Sequence Changes in N2 Field Strains

The chemical changes associated with antigenic drift have been investigated most extensively in the N2 subtype. The amino acid sequence data (Table 23)

Table 23. Comparison of protein sequences in eight N2 neuraminidases: A/RI/5⁻/57 (ELLEMAN et al. 1982); A/Tokyo/3/67 (WARD et al. 1982, LENTZ et al. 1984); A/Aichi/2/68 (LAVER et al. 1982); A/NT/60/68 (BENTLEY and BROWNLEE 1982); A/Eng/42/72 (LAVER et al. 1982); A/Udorn/307/72 (MARKOFF and LAI 1982); A/Vic/3/75 (VAN ROMPUY et al. 1982); and A/Bangkok/1/79 (MARTINEZ et al. 1983)[a]

```
RI/5¯/57        Met Asn Pro Asn Gln Lys Thr Ile Thr Ile Gly Ser Val Ser Leu Thr Ile Ala Thr Val Cys Phe Leu Met Gln
Tokyo/3/67      Met Asn Pro Asn Gln Lys Ile Ile Thr Ile Gly Ser Val Ser Leu Thr Ile Ala Thr Val Cys Phe Leu Met Gln
Aichi/2/68
NT/60/68        Met Asn Pro Asn Gln Lys Ile Ile Thr Ile Gly Ser Val Ser Leu Thr Ile Ala Thr Val Cys Phe Leu Met Gln
Eng/42/72
Udorn/307/72    Met Asn Pro Asn Gln Lys Ile Ile Thr Ile Gly Ser Val Ser Leu Thr Ile Ala Thr Ile Cys Phe Leu Met Gln
Vic/3/75        Met Asn Pro Asn Gln Lys Ile Ile Thr Ile Gly Ser Val Ser Leu Thr Ile Ala Thr Ile Cys Phe Leu Met Gln
Bangkok/1/79    Met Asn Pro Asn Gln Lys Ile Ile Thr Ile Gly Ser Val Ser Leu Thr Ile Ala Thr Ile Cys Phe Leu Met Gln  25

RI/5¯/57        Ile Ala Ile Leu Ala Thr Thr Val Thr Leu His Phe Lys Gln His Glu Cys Asp Ser Pro Ala Ser Asn Gln Val
Tokyo/3/67      Ile Ala Ile Leu Val Thr Thr Val Thr Leu His Phe Lys Gln His Glu Cys Asp Ser Pro Ala Ser Asn Gln Val
Aichi/2/68
NT/60/68        Ile Ala Ile Leu Val Thr Thr Val Thr Leu His Phe Lys Gln Tyr Glu Cys Asp Ser Pro Ala Ser Asn Gln Val
Eng/42/72
Udorn/307/72    Ile Ala Ile Gln Val Thr Thr Val Thr Leu His Phe Lys Gln Tyr Glu Cys Asp Ser Pro Ala Asn Asn Gln Val
Vic/3/75        Ile Ala Ile Leu Val Thr Thr Val Thr Leu His Phe Lys Gln Tyr Glu Cys Asp Ser Pro Ala Asn Asn Gln Val
Bangkok/1/79    Ile Ala Ile Leu Val Thr Thr Val Thr Leu His Phe Lys Gln Tyr Glu Cys Ser Ser Pro Pro Asn Asn Gln Val  50

RI/5¯/57        Met Pro Cys Glu Pro Ile Ile Ile Glu Arg Asn Ile Thr Glu Ile Val Tyr Leu Asn Asn Thr Thr Ile Glu Lys
Tokyo/67        Met Pro Cys Glu Pro Ile Ile Ile Glu Arg Asn Ile Thr Glu Ile Val Tyr Leu Asn Asn Thr Thr Ile Glu Lys
Aichi/2/68                                                                                          Glu Lys
NT/60/68        Met Pro Cys Glu Pro Ile Ile Ile Glu Arg Asn Ile Thr Glu Ile Val Tyr Leu Asn Asn Thr Thr Ile Glu Lys
Eng/42/72
Udorn/72        Met Pro Cys Glu Pro Ile Ile Ile Glu Arg Asn Ile Thr Glu Ile Val Tyr Leu Thr Asn Thr Thr Ile Glu Lys
Vic/3/75        Met Pro Cys Glu Pro Ile Ile Ser Ile Glu Arg Asn Ile Thr Glu Ile Val Tyr Leu Thr Asn Thr Thr Ile Glu Lys
Bangkok/1/79    Met Pro Cys Glu Pro Ile Ile Ile Glu Arg Asn Ile Thr Glu Ile Val Tyr Leu Thr Asn Thr Thr Ile Glu Lys  75

RI/5¯/57        Glu Ile Cys Pro Glu Val Val Glu Tyr Arg Asn Trp Ser Lys Pro Gln Cys Gln Ile Thr Gly Phe Ala Pro Phe
Tokyo/67        Glu Ile Cys Pro Lys Val Val Glu Tyr Arg Asn Trp Ser Lys Pro Glu Cys Gln Ile Thr Gly Phe Ala Pro Phe
Aichi/68        Glu Ile Cys Pro Lys Val Val Glu Tyr Arg Asn Trp Ser Lys Pro Gln Cys Gln Ile Thr Gly Phe Ala Pro Phe
NT/60/68        Glu Ile Cys Pro Lys Val Val Glu Tyr Arg Asn Trp Ser Lys Pro Gln Cys Gln Ile Thr Gly Phe Ala Pro Phe
Eng/42/72                       Leu Val Glu Tyr Arg Lys Trp Ser Lys Pro Gln Cys Lys Ile Thr Gly Phe Ala Pro Phe
Udorn/72        Gly Ile Cys Pro Lys Leu Val Glu Tyr Arg Asn Trp Ser Lys Pro Gln Cys Lys Ile Thr Gly Phe Ala Pro Phe
Vic/75          Gly Ile Cys Pro Lys Leu Val Glu Tyr Arg Asn Trp Ser Lys Pro Gln Cys Lys Ile Thr Gly Phe Ala Pro Phe
Bangkok/1/79    Glu Ile Cys Pro Lys Leu Val Glu Tyr Arg Asn Trp Ser Lys Pro Gln Cys Lys Ile Thr Gly Phe Ala Pro Phe 100

RI/5¯/57        Ser Lys Asp Asn Ser Ile Arg Leu Ser Ala Gly Gly Asp Ile Trp Val Thr Arg Glu Pro Tyr Val Ser Cys Asp
Tokyo/67        Ser Lys Asp Asn Ser Ile Arg Leu Ser Ala Gly Gly Asp Ile Trp Val Thr Arg Glu Pro Tyr Val Ser Cys Asp
Aichi/68        Ser Lys Asp Asn Ser Ile Arg Leu Ser Ala Gly Gly Asp Ile Trp Val Thr Arg
NT/60/68        Ser Lys Asp Asn Ser Ile Arg Leu Ser Ala Gly Gly Asp Ile Trp Val Thr Arg Glu Pro Tyr Val Ser Cys Asp
Eng/42/72       Ser Lys Asp Asn Ser Ile Arg Leu Ser Ala Gly Gly Asp Ile Trp Val Thr Arg Glu Pro Tyr Val Ser Cys Asp
Udorn/72        Ser Lys Asp Asn Ser Ile Arg Leu Ser Ala Gly Gly Asp Ile Trp Val Thr Arg Glu Pro Tyr Val Ser Cys Asp
Vic/3/75        Ser Lys Asp Asn Ser Ile Arg Leu Ser Ala Gly Gly Asp Ile Trp Val Thr Arg Glu Pro Tyr Val Ser Cys Asp
Bangkok/1/79    Ser Lys Asp Asn Ser Ile Arg Leu Ser Ala Gly Gly Gly Ile Trp Val Thr Arg Glu Pro Tyr Val Ser Cys Asp 125

RI/5¯/57        Pro Gly Lys Cys Tyr Gln Phe Ala Leu Gly Gln Gly Thr Thr Leu Asp Asn Lys His Ser Asn Gly Thr Ile His
Tokyo/67        Pro Val Lys Cys Tyr Gln Phe Ala Leu Gly Gln Gly Thr Thr Leu Asp Asn Lys His Ser Asn Asp Thr Val His
Aichi/68                Cys Tyr Gln Phe Ala Leu Gly Gln Gly Thr Thr Leu Asp Asn Lys His Ser Asn Asp Thr Ile His
NT/60/68        His Gly Lys Cys Tyr Gln Phe Ala Leu Gly Gln Gly Thr Thr Leu Asp Asn Lys His Ser Asn Asp Thr Ile His
Eng/42/72       Pro Gly Lys
Udorn/72        Pro Gly Lys Cys Tyr Gln Phe Ala Leu Gly Gln Gly Thr Thr Leu Asp Asn Lys His Ser Asn Asp Thr Ile His
Vic/3/75        Pro Arg Lys Cys Tyr Gln Phe Ala Leu Gly Gln Gly Thr Thr Leu Glu Asn Lys His Ser Asn Asp Thr Ile His
Bangkok/1/79    Pro Gly Lys Cys Tyr Gln Phe Ala Leu Gly Gln Gly Thr Thr Leu Asp Asn Lys His Ser Asn Asp Thr Ile His 150

RI/5¯/57        Asp Arg Ile Pro His Arg Thr Leu Leu Met Asn Glu Leu Gly Val Pro Phe His Leu Gly Thr Lys Gln Val Cys
Tokyo/57        Asp Arg Ile Pro His Arg Thr Leu Leu Met Asn Glu Leu Gly Val Pro Phe His Leu Gly Thr Arg Gln Val Cys
Aichi/2/68      Asp Arg Ile Pro His Arg Thr Leu Leu Met Asn Glu Leu Gly Val Pro Phe His Leu Gly Thr Arg
NT/60/68        Asp Arg Ile Pro His Arg Thr Leu Leu Met Asn Glu Leu Gly Val Pro Phe His Leu Gly Thr Arg Gln Val Cys
Eng/42/72               Ile Pro His Arg Thr Leu Leu Met Asn Glu Leu Gly Val Pro Phe His Leu Gly Thr Arg
Udorn/72        Asp Arg Thr Pro His Arg Thr Leu Leu Met Asn Glu Leu Gly Val Pro Phe His Leu Gly Thr Arg Gln Val Cys
Vic/3/75        Asp Arg Thr Pro His Arg Thr Leu Leu Met Asn Glu Leu Gly Val Pro Phe His Leu Gly Thr Arg Gln Val Cys
Bangkok/1/79    Asp Arg Thr Pro Tyr Arg Thr Leu Leu Met Asn Glu Leu Gly Val Pro Phe His Leu Gly Thr Arg Gln Val Cys 175

RI/5¯/57        Val Ala Trp Ser Ser Ser Ser Ser Cys His Asp Gly Lys Ala Trp Leu His Val Cys Val Thr Gly Asp Asp Arg Asn
Tokyo/67        Ile Ala Trp Ser Ser Ser Ser Cys His Asp Gly Lys Ala Trp Leu His Val Cys Ile Thr Gly Asp Asp Lys Asn
Aichi/3/68
NT/60/68        Ile Ala Trp Ser Ser Ser Ser Cys His Asp Gly Lys Ala Trp Leu His Val Cys Ile Thr Gly Asp Asp Lys Asn
Eng/42/72
Udorn/72        Ile Gly Trp Ser Ser Ser Ser Cys His Asp Gly Lys Ala Trp Leu His Val Cys Val Thr Gly Tyr Asp Lys Asn
Vic/3/75        Ile Ala Trp Ser Ser Ser Ser Cys His Asp Gly Lys Ala Trp Leu His Val Cys Val Thr Gly Tyr Asp Lys Asn
Bangkok/1/79    Ile Ala Trp Ser Ser Ser Ser Cys His Asp Gly Lys Ala Trp Leu His Val Cys Val Thr Gly Tyr Asp Lys Asn 200

RI/5¯/57        Ala Thr Ala Ser Phe Ile Tyr Asp Gly Arg Leu Val Asp Ser Ile Gly Ser Trp Ser Gln Asn Ile Leu Arg Thr
Tokyo/67        Ala Thr Ala Ser Phe Ile Tyr Asp Gly Arg Leu Val Asp Ser Ile Gly Ser Trp Ser Gln Asn Ile Leu Arg Thr
Aichi/68                                        Leu Val Asp Ser Ile Gly Ser Trp Ser Gln Asn Ile Leu Arg Thr
NT/60/68        Ala Thr Ala Ser Phe Ile Tyr Asp Gly Arg Leu Val Asp Ser Ile Gly Ser Trp Ser Gln Asn Ile Leu Arg Thr
Eng/42/72                                       Leu Val Asp Ser Ile Gly Ser Trp Ser Gln Asn Ile Leu Arg Thr
Udorn/72        Ala Thr Ala Ser Phe Ile Tyr Asp Gly Arg Leu Val Asp Ser Ile Gly Ser Trp Ser Gln Asn Ile Leu Arg Thr
Vic/3/75        Ala Thr Ala Ser Phe Ile Tyr Asp Gly Arg Leu Val Asp Ser Ile Gly Ser Trp Ser Gln Asn Ile Leu Arg Thr
Bangkok/1/79    Ala Thr Ala Ser Phe Ile Tyr Asp Gly Arg Leu Val Asp Ser Ile Gly Ser Trp Ser Lys Asn Ile Leu Arg Thr 225

RI/5¯/57        Gln Glu Ser Glu Cys Val Cys Ile Asn Gly Thr Cys Thr Val Val Met Thr Asp Gly Ser Ala Ser Gly Arg Ala
Tokyo/67        Gln Glu Ser Glu Cys Val Cys Ile Asn Gly Thr Cys Thr Val Val Met Thr Asp Gly Ser Ala Ser Gly Arg Ala
Aichi/68        Gln Glu Ser Glu Cys Val Cys Ile Asn Gly Thr Cys Thr Val Val Met Thr Asp Gly Ser Ala Ser Gly Arg Ala
NT/60/68        Gln Glu Ser Glu Cys Val Cys Ile Asn Gly Thr Cys Thr Val Val Met Thr Asp Gly Ser Ala Ser Gly Arg Ala
Eng/42/72       Gln Glu Ser Glu Cys Val Cys Ile Asn Gly Thr Cys Thr Val Val Met Thr Asp Gly Ser Ala Ser Gly Arg Ala
Udorn/72        Gln Glu Ser Glu Cys Val Cys Ile Asn Gly Thr Cys Thr Val Val Met Thr Asp Gly Ser Ala Ser Gly Arg Ala
Vic/3/75        Gln Glu Ser Glu Cys Val Cys Ile Asn Gly Thr Cys Thr Val Val Met Thr Asp Gly Ser Ala Ser Gly Arg Ala
Bangkok/1/79    Gln Glu Ser Glu Cys Val Cys Ile Asn Gly Thr Cys Thr Val Val Met Thr Asp Gly Ser Ala Ser Glu Arg Ala 250

RI/5¯/57        Asp Thr Arg Ile Leu Phe Ile Lys Glu Gly Lys Ile Val His Ile Ser Pro Leu Ser Gly Ser Ala Gln His Ile
Tokyo/67        Asp Thr Arg Ile Leu Phe Ile Glu Glu Gly Lys Ile Val His Ile Ser Pro Leu Ser Gly Ser Ala Gln His Val
Aichi/68        Asp Thr Arg Ile Leu Phe Ile Glu Glu Gly Lys Ile Val His Ile Ser Pro Leu Ser Gly Ser Ala Gln His Val
NT/60/68        Asp Thr Arg Ile Leu Phe Ile Glu Glu Gly Lys Ile Val His Ile Ser Pro Leu Ala Gly Ser Ala Gln His Val
Eng/42/72       Asp Thr Arg Ile Leu Phe Ile Glu Glu Gly Lys Ile Val His Ile Ser Pro Leu Ser Gly Ser Ala Gln His Val
Udorn/72        Asp Thr Lys Ile Leu Phe Ile Glu Glu Gly Lys Ile Val His Ile Ser Pro Leu Ser Gly Ser Ala Gln His Val
Vic/3/75        Asp Thr Lys Ile Leu Phe Ile Glu Glu Gly Lys Ile Val His Ile Ser Pro Leu Ser Gly Ser Ala Gln His Val
Bangkok/1/79    Asp Thr Lys Ile Leu Phe Ile Glu Glu Gly Lys Ile Val His Ile Ser Pro Leu Ser Gly Ser Ala Gln His Val 275
```

Table 23 (continued)



[a] The accumulated amino acid sequence changes are shaded

shows that antigenic drift in N2 neuraminidase involves an accumulation of point mutations at a rate similar to that found in Hong Kong hemagglutinin. By 1967, 27 amino acid substitutions had accumulated in the heads of neuraminidase. This represents a mutation rate during the period of the Asian influenza epidemic of 0.68 residues/year/100 amino acid residues. This is comparable to the rate of 0.66 found for influenza hemagglutinin in the period 1968 to 1979.

The Hong Kong influenza subtype, which appeared in man in 1968, retained the Asian N2 neuraminidase, and this has continued to undergo antigenic drift. Comparison of the A/Tokyo/67 and NT/60/68 sequences shows that the Hong Kong strain retained 22 of the sequence changes that had accumulated during the Asian flu period. There were only eight differences between these two strains

(residues 42, 127, 149, 199, 269, 346, 358, and 434), of which three were substitutions found only in Tokyo/67, and two found only in NT/60/68. The difference at position 199 is now also doubtful. The amino acid sequence data for Tokyo/67 heads (WARD et al. 1982) indicated Asp at this position, while the nucleotide sequence using vRNA template from the same virus preparation indicated Lys (LENTZ et al. 1984) such as was found in all subsequent N2 strains. As shown in Table 23, the two 1968 isolates were very similar except for the substitutions found only in NT/60/68 (positions 269 and 434).

From 1968 to 1972, new substitutions appeared at positions 81, 86, 93, 307, 308, 347, and 431 (Eng/42/72), with additional changes apparent at 20, 29, 47, 69, 153, 253, 344, 368, and 390 in Udorn/72. A further seven changes had accumulated at positions 57, 137, 141, 328, 334, 346, and 401 by 1975. A matrix of amino acid sequence relationships is shown in Table 24. The location of these accumulated substitutions on the three-dimensional structure of N2 neuraminidase is discussed in Sect. 5.5.2. Only the 1979 field strain showed substitutions around residue 221 which changed in the group III N2 monoclonal variants (Table 18). Most field strains, however, show substitutions at and near residues 344 and 368 which changed in the variants selected with group I monoclonal antibodies.

Table 24. Matrix of amino acid and nucleotide sequence homologies between N2 neuraminidases. Entries show the number of different amino acids (lower left triangle) or nucleotides (upper right triangle). Strains used are RI/5$^-$/57, Tokyo/3/67, NT/60/68, Udorn/72, Vic/3/75 and Bangkok/1/79

	1957	1967	1968	1972	1975	1979
1957		60	61	91	104	109
1967	28		24	62	74	83
1968	27	9		60	75	79
1972	44	30	29		21	31
1975	48	33	34	9		42
1979	50	38	37	14	15	

The complete gene sequences of six N2 field strains, RI/5$^-$/57, Tokyo/3/67, NT/60/68, Udorn/72, Vic/3/75, and Bangkok/1/79 are shown in Table 25. The non-coding regions are conserved through the N2 subtype with the exception of nucleotide 19, changing from A to G between 1968 and 1972, and the excision of one nucleotide in the 5′ region of the A/Udorn/72 and A/Bangkok/79 sequences, resulting in the sequence -AAAAAA- compared to -AAAAAAA- in other strains. Nucleotide changes occur at 121 positions in the sequence. Changes at 85 of the 121 positions have become fixed in later strains, and 18 revert

Table 25. Comparison of gene sequences in six N2 neuraminidases: A/RI/5⁻/57 (ELLEMAN et al. 1982); A/Tokyo/3/67 (WARD et al. 1982, LENTZ et al., to be published); A/NT/60/68 (BENTLEY and BROWNLEE 1982); A/Udorn/307/72 (MARKOFF and LAI 1982); A/Vic/3/75 (VAN ROMPUY et al. 1982); and A/Bangkok/1/79 (MARTINEZ et al. 1983). The nucleotide and amino acid sequence differences are shaded

	5	10	15	20
	Met Asn Pro Asn Gln Lys Thr Ile Thr Ile Gly Ser Val Ser Leu Thr Ile Ala Thr Val			
RI/5⁻/57	3'-AGCAAAAGCAGGAG/GAAA AUG AAU CCA AAU CAA AAG ACA AUA ACA AUU GGC UCU GUC UCU CUC ACC AUU GCA ACA GUA			
Tokyo/3/67	AGCAAAAGCAGGAG/GAAA AUG AAU CCA AAU CAA AAG AUA AUA ACA AUU GGC UCU GUC UCU CUC ACC AUU GCA ACA GUA			
NT/60/68	AGCAAAAGCAGGAG/GAAA AUG AAU CCA AAU CAA AAG AUA AUA ACA AUU GCC UCU GUC UCU CUC ACC AUU GCA ACA GUA			
Udorn/72	AGCAAAAGCAGGAG/GAAG AUG AAU CCA AAU CAA AAG AUA AUA ACA AUU GGC UCU GUC UCU CUC ACC AUU GCA ACA AUA			
Vic/3/75	AGCAAAAGCAGGAG/GAAG AUG AAU CCA AAU CAA AAG AUA AUA ACA AUU GGC UCU GUC UCU CUC ACC AUU GCA ACA AUA			
Bangkok/1/79	AGCAAAAGCAGGAG/GAAG AUG AAU CCA AAU CAA AAG AUA AUA ACA AUU GGC UCU GUC UCU CUC ACC AUU GCA ACA AUA			
	Met Asn Pro Asn Gln Lys Ile Ile Thr Ile Gly Ser Val Ser Leu Thr Ile Ala Thr Ile			

	25	30	35	40	45
	Cys Phe Leu Met Gln Ile Ala Ile Leu Ala Thr Thr Val Thr Leu His Phe Lys Gln His Glu Cys Asp Ser Pro				
RI/5⁻/57	UGC UUC CUC AUG CAG AUU GCC AUC CUG GCA ACU ACU GUG ACA UUG CAU UUU AAA CAA CAU GAG UGC GAC UCC CCC				
Tokyo/3/67	UGC UUU CUC AUG CAG AUU GCC AUC UUG GCA ACU ACU GUA ACA UUC CAC UUU AAG CAA CAU GAG UGC GAC UCC CCC				
NT/60/68	UGC UUC CUC AUG CAG AUU GCC AUC CUU GCA ACG ACU GUA ACA UUC CAU UUU AAG CAA UAU GAG UGC GAC UCC CCU				
Udorn/72	UGC UUC CUC AUG CAG AUU GCC AUC CAG GCA ACU ACU GUA ACA UUG CAU UUC AAG CAA UAU GAG UGC GAC UCC CCC				
Vic/3/75	UGC UUC CUC AUG CAG AUU GCC AUC CUG GCA ACU ACU GUA ACA UUG CAU UUC AAG CAA UAU GAG UGC GAC UCC CCC				
Bangkok/1/79	UGC UUC CUU AUG CAA AUU GCC AUC CUG GUA ACU ACU GUA ACA UUG CAU UUC AAG CAA UAU GAG UGC AGC UCC CCC				
	Cys Phe Leu Met Gln Ile Ala Ile Leu Val Thr Thr Val Thr Leu His Phe Lys Gln Tyr Glu Cys Ser Ser Pro				

	50	55	60	65	70
	Ala Ser Asn Gln Val Met Pro Cys Glu Pro Ile Ile Ile Glu Arg Asn Ile Thr Glu Ile Val Tyr Leu Asn Asn				
RI/5⁻/57	GCG AGC AAC CAA GUA AUG CCA UGU GAA CCA AUA AUA AUA GAA AGG AAC AUA ACA GAG AUA GUG UAU UUG AAU AAC				
Tokyo/3/67	GCG AGC AAC CAA GUA AUG CCC UGU GAA CCA AUA AUA AUA GAA AGG AAC AUA ACA GAG AUA GUG UAU UUG AAU AAC				
NT/60/68	GCG AGC AAC CAA GUA AUG CCC UGU GAA CCA AUA AUA AUA GAA AGG AAC AUA ACA GAC AUA GUG UAU UUG AAU AAC				
Udorn/72	GCC AGC AAC CAA GUA AUG CCG UGU GAA CCA AUA AUA AUA GAA AGC AAC AUA ACA GAC AUA GUG UAU UUG AGU AAC				
Vic/3/75	GCG AAC AAC CAA GUA AUG CCC UGU GAA CCA AUA UCA AUA GAA AGC AAC AUA ACA GAC AUA GUC UAU UUG ACU AAC				
Bangkok/1/79	GCA AAC AAC CAA GUA AUG CCC UGU GAA CCA AUA AUA AUA GAA AGA AAC AUA ACA GAG AUA GUC UAU UUG AGU AAC				
	Pro Asn Asn Gln Val Met Pro Cys Glu Pro Ile Ile Ile Glu Arg Asn Ile Thr Glu Ile Val Tyr Leu Thr Asn				

	75	80	85	90	95
	Thr Thr Ile Glu Lys Glu Ile Cys Pro Glu Val Val Glu Tyr Arg Asn Trp Ser Lys Pro Gln Cys Gln Ile Thr				
RI/5⁻/57	ACC ACC AUA GAG AAA GAG AUU UGC CCC GAA GUA GUG GAA UAC AGA AAU UGG UCA AAG CCG CAA UGU CAA AUU ACA				
Tokyo/3/67	ACC ACC AUA GAG AAA GAG AUA UGC CCC AAA GUA GUG GAA UAC AGA AAU UGG UCA AAG CCG CAA UGU CAA AUU ACA				
NT/60/68	ACC ACC AUA GAG AAA GAG AUA UGC CCC AAA GUA GUG GAA UAC AGA AAU UGG UCA AAG CCG CAA UGU CAA AUU ACA				
Udorn/72	ACC ACC AUA GAG AAA GAG AUA UGC CCC AAA UUA GUG GAA UAC ACG AAU UGG UCA AAG CCG CAA UGU GAA AUU ACA				
Vic/3/75	ACC ACC AUA GAA AAA GAG AUA UGC CCC AAA UUA GUG GAA UAC ACG AAU UGG UCA AAC CCA CAA UCU GAA AUU ACA				
Bangkok/1/79	ACC ACC AUA GAC AAA GAG AUA UGC CCC AAA UUA GUG GAA UAC AGA AAC UGG UCA AAG CCG CAA UGU GAA AUU ACA				
	Thr Thr Ile Glu Lys Glu Ile Cys Pro Lys Leu Val Glu Tyr Arg Asn Trp Ser Lys Pro Gln Cys Lys Ile Thr				

	100	105	110	115	120
	Gly Phe Ala Pro Phe Ser Lys Asp Asn Ser Ile Arg Leu Ser Ala Gly Gly Asp Ile Trp Val Thr Arg Glu Pro				
RI/5⁻/57	GGA UUU GCA CCU UUU UCU AAG GAC AAU UCA AUC CGG CUU UCU GCU GGU GGG GAC AUU UGG GUG ACG AGA GAA CCU				
Tokyo/3/67	GGA UUU GCA CCU UUU UCU AAG GAC AAU UCA AUC CGG CUU UCU GCU GGU GGG GAC AUU UGG GUG ACG AGA GAA CCU				
NT/60/68	GGA UUU GCA CCU UUU UCU AAG GAC AAU UCA AUC CGG CUU UCU GCU GGU GGG GAC AUU UGG GUG ACG AGA GAA CCU				
Udorn/72	GGA UUU GCA CCU UUU UCU AAG GAC AAU UCA ACU CGG CUU UCU GCU GGU GGG GAC AUU UGG GUG ACG AGA GAA CCU				
Vic/3/75	GGA UUU GCA CCU UUC UCU AAG GAC AAU UCA AUU CGG CUU UCU GCU GGU GGG GAC AUU UGG GUG ACG AGA GAA CCU				
Bangkok/1/79	GGA UUU GCA CCU UUU UCU AAG GAC AAU UCA AUU CGG CUU UCU GCU GGU GGG GGC AUU UGG GUG ACG AGA GAA CCU				
	Gly Phe Ala Pro Phe Ser Lys Asp Asn Ser Ile Arg Leu Ser Ala Gly Gly Gly Ile Trp Val Thr Arg Glu Pro				

	125	130	135	140	145
	Tyr Val Ser Cys Asp Pro Gly Lys Cys Tyr Gln Phe Ala Leu Gly Gln Gly Thr Thr Leu Asp Asn Lys His Ser				
RI/5⁻/57	UAU GUG UCA UGC GAC CCU GGG AAG UGU UAU CAA UUU GCA CUC GGG CAG GGG ACC ACA CUA GAC AAC AAA CAU UCA				
Tokyo/3/67	UAU GUG UCA UGC GAU CCU GGC AAG UGU UAU CAA UUU GCA CUC GGG CAG GGG ACC ACA CUA GAC AAC AAA CAU UCA				
NT/60/68	UAU GUG UCA UGC GAU CAU GGC AAG UGU UAU CAA UUU GCA CUC GGG CAG GGG ACC ACA CUA GAC AAC AAA CAU UCA				
Udorn/72	UAU GUG UCA UGC GAU CCU GGC AAG UGU UAU CAA UUU GCA CUC GGG CAG GGG ACC ACA CUA GAC AAC AAA CAU UCA				
Vic/3/75	UAU GUG UCA UGC GAU CCU GGC AAA UGU UAU CAA UUU GCA CUC GGG CAG GGG ACC ACA CUA GAA AAC AAA CAU UCA				
Bangkok/1/79	UAU GUG UCA UGC GAU CCU GGC AAG UGU UAU CAA UUU GCA CUC GGG CAG GGG ACC ACA CUA GAC AAG AAG CAU UCA				
	Tyr Val Ser Cys Asp Pro Gly Lys Cys Tyr Gln Phe Ala Leu Gly Gln Gly Thr Thr Leu Asp Asn Lys His Ser				

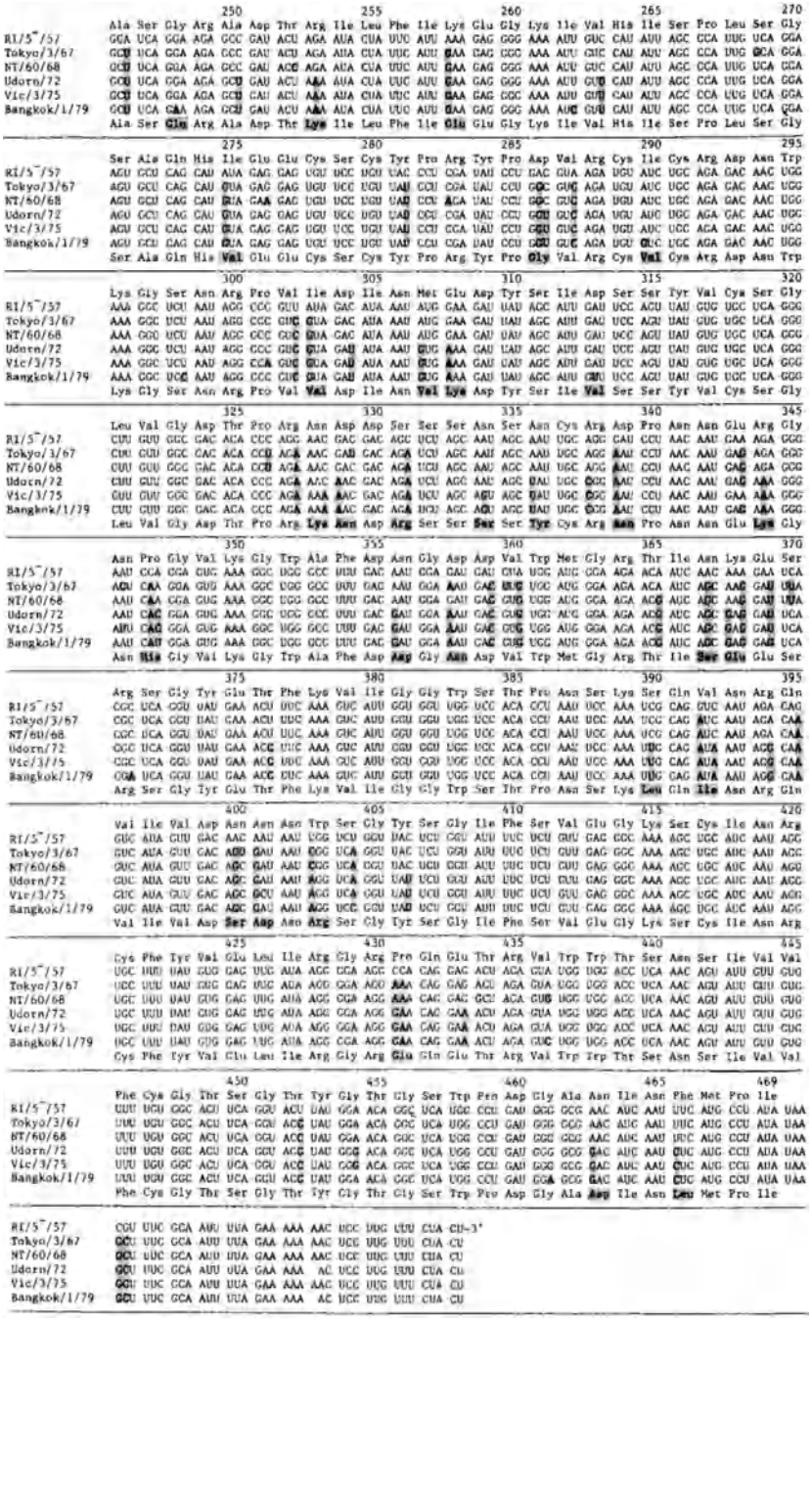
	150	155	160	165	170
	Asn Gly Thr Ile His Asp Arg Ile Pro His Arg Thr Leu Leu Met Asn Glu Leu Gly Val Pro Phe His Leu Gly				
RI/5⁻/57	AAU GGC ACA AUA CAU GAU AGA AUC CCU CAC CGA ACC CUA UUA AUG AAU GAG UUG GGU GUU CCA UUU CAU UUA GGA				
Tokyo/3/67	AAU GAC ACA GUA CAU GAU AGA AUC CCU CAU CGA ACC CUA UUA AUG AAU GAG UUG GGU GUU CCA UUU CAC UUA GGA				
NT/60/68	AAU GAC ACA AUA CAU GAU AGA AUC CCU CAU CGA ACC CUA UUA AUG AAU GAG UUG GGU GUU CCA UUU CAU UUA GGA				
Udorn/72	AAU GAC ACA AUA CAU GAU AGA AGC CCU CAU CGA ACC CUA UUG AUG AAU GAG UUG GGU GUU CCA UUU CAU UUG GGA				
Vic/3/75	AAU GAC ACA AUA CAU GAU AGA AGC CCU CAU CGA ACC CUA UUG AUG AAU GAG UUG GGU GUU CCA UUU CAU UUG GGA				
Bangkok/1/79	AAU GAC ACA AUA CAU GAU AGG AGC CCU UAU CGA ACC CUA UUG AUG AAU GAG UUG GGU GUU CCA UUU CAU UUG GGA				
	Asn Asp Thr Ile His Asp Arg Thr Pro Tyr Arg Thr Leu Leu Met Asn Glu Leu Gly Val Pro Phe His Leu Gly				

	175	180	185	190	195
	Thr Lys Gln Val Cys Val Ala Trp Ser Ser Ser Ser Cys His Asp Gly Lys Ala Trp Leu His Val Cys Val Thr				
RI/5⁻/57	ACC AAA CAA GUG UGU GUA GCA UGG UCC AGC UCA AGU UGU CAC GAU GGA AAA GCA UGG UUG CAU GUU UGU GUC ACU				
Tokyo/3/67	ACC AGG CAA GUG UGU AUA GCA UGG UCC AGC UCA AGU UGU CAC GAU GGA AAA GCA UGG CUG CAU GUU UGU AUC ACU				
NT/60/68	ACC AGG CAA GUG UGU AUA GCA UGG UCC AGC UCA AGU UGU CAC GAU GGA AAA GCA UGG CUG CAU GUU UGU AUC ACU				
Udorn/72	ACC AGG CAA GUG UGU AUA GGA UGG UCC AGC UCA AGU UGU CAC GAU GGA AAA GCA UGG CUG CAU CUU UGU GUC ACU				
Vic/3/75	ACC AGC CAA GUG UGU AUA GCA UGG UCC AGC UCA AGU UGU CAC GAU GGA AAA GCA UCG CUG CAU CUU UGU GUC ACU				
Bangkok/1/79	ACC AGC CAA GUG UGU AUA GCA UGG UCC AGC UCA AGU UGU CAC GAU GGA AAA GCA UGG CUG CAU CUU UGU GUC ACU				
	Thr Arg Gln Val Cys Ile Ala Trp Ser Ser Ser Cys His Asp Gly Lys Ala Trp Leu His Val Cys Val Thr				

	200	205	210	215	220
	Gly Asp Asp Arg Asn Ala Thr Ala Ser Phe Ile Tyr Asp Gly Arg Leu Val Asp Ser Ile Gly Ser Trp Ser Gln				
RI/5⁻/57	GGG GAU GAU AGA AAU GCA ACU GCU AGC UUC AUU UAU GAC GGG AGG CUU GUC GAC AGU AUU GGU UCA UGG UCU CAA				
Tokyo/3/67	GGG GAU GAU AGA AAU GCA ACU GCU AGC UUC AUU UAU GAC GGG AGG CUU GUC GAC AGU AUU GGU UCA UGG UCU CAA				
NT/60/68	GGG GAU GAC AGA AAU GCA ACU GCU AGC UUC AUU UAU GAC GGG AGG CUU GUC GAC AGU AUU GGU UCA UGG UCU CAA				
Udorn/72	GGG GAU GAU AGA AAU GCA ACU GCU AGC UUC AUU UAC GAU GGG AGC CUU GUA GAC AGU AUU GGU UCA UGG UCU CAA				
Vic/3/75	GGG GAU GAU AGA AAU GCA ACU GCU AGC UUC AUU UAC GAU GGG AGG CUU GUA GAC AGC AUU GGU UCA UGG UCU CAA				
Bangkok/1/79	GGG GAU GAU AGA AAU GCA ACU GCU AGC UUC AUU UAC GAU GGG AGG CUU GUA GAC AGU AUU GGU UCA UGG UCC AAA				
	Gly Asp Asp Arg Asn Ala Thr Ala Ser Phe Ile Tyr Asp Gly Arg Leu Val Asp Ser Ile Gly Ser Trp Ser Lys				

	225	230	235	240	245
	Asn Ile Leu Arg Thr Gln Glu Ser Glu Cys Val Cys Ile Asn Gly Thr Cys Thr Val Val Met Thr Asp Gly Ser				
RI/5⁻/57	AAU AUC CUC AGG ACC CAG GAG UCG GAA UGC GUU UGU AUC AAU GGG ACU UGC ACA GUA GUA AUG ACU GAU GGA AGU				
Tokyo/3/67	AAU AUC CUC AGA ACC CAG GAG UCG GAA UGC GUU UGU AUC AAU GGG ACU UGC ACA GUA GUA AUG ACU GAU GGA AGU				
NT/60/68	AAU AUC CUC AGA ACC CAG GAG UCG GAA UGC GUU UGU AUC AAU GGG ACU UGC ACA GUA GUA AUG ACU GAU GGA AGU				
Udorn/72	AAU AUC CUC AGG ACC CAA GAG UCG GAA UGC GUU UGU AUC AAU GGG ACU UGU ACA GUA GUA AUG ACU GAU GGA AGU				
Vic/3/75	AAU AUC CUC AGG ACC CAG GAC UCG GAA UGU GUU UGU AUC AAC GGG ACU UGU ACA GUA GUA AUG ACU GAU GGA AGU				
Bangkok/1/79	AAU AUC CUC AGG ACC CAG GAC UCG GAA UGC GUU UGU AUC AAU GGC ACU UGU ACA GUA GUA AUG ACU GAU GGA AGU				
	Asn Ile Leu Arg Thr Gln Glu Ser Glu Cys Val Cys Ile Asn Gly Thr Cys Thr Val Val Met Thr Asp Gly Ser				

Table 25 (continued)



to the earlier strain sequence, indicating that strains isolated in different parts of the world do not necessarily represent a direct geneological lineage. Two positions show progressive change, in codons 403 and 431. A matrix of nucleotide sequence relationships is shown in Table 24. The nucleotide sequence of A/RI/5$^+$/57, from nucleotide 250 at the 5′ end of the cDNA to the end, has now been determined (G.M. Air, personal communication) and, contrary to the interpretation of the peptide map data (LAVER et al. 1982), the sequence in the head region is identical to that of RI/5$^-$/57 neuraminidase, except at nucleotide 962, where G substitutes for A and Ser 315 becomes Gly.

5.3 Sequence Changes and Antigenic Shift

Base sequence homology studies (SCHOLTISSEK et al. 1978a) and comparative structural analyses show that antigenic shifts to new pandemic strains of influenza virus result from genetic reassortment, not point mutation as proposed by FAZEKAS DE ST. GROTH (1978). In general the nonstructural proteins (LAMB and LAI 1980; WINTER et al. 1981), matrix protein (MCCAULEY et al. 1982), and the three P proteins (FIELDS and WINTER 1982; WINTER and FIELDS 1982; KAPTEIN and NAYAK 1982; BISHOP et al. 1982) involved in RNA transcription (ULMANEN et al. 1981) are highly conserved among type A influenza strains. In contrast, the hemagglutinins (see WARD 1981 for review) and neuraminidases of different subtype viruses vary by 30%–60%.

Complete sequence comparisons can only be made for the N1 and N2 subtypes (Table 26). To maximize the structural relationships between the two subtypes, gaps have been inserted in the aligned sequences. Using the N2 sequence as a reference, the N1 sequence contains deletions at positions 63 to 77, 158, 314, 315, 334, 336, 338, and 436; single insertions between residues 162/163, 164/165, and 308/309; and double insertions between residues 410/411 and 413/414.

As shown in Table 26, the nontranslated region at the 5′ end of the cDNA contains 20 nucleotides in A/PR/8/34 (N1) and 19 in A/RI/5$^-$/57 (N2), the coding regions contain 1362 and 1407 bases respectively, and the nucleotides at the 3′ end of the cDNA number 31 in A/PR/8/34 and 41 in A/RI/5$^-$/57. There is a 45-nucleotide deletion in A/PR/8/34 which, by comparison with the partial sequences of other N1 neuraminidase genes (BLOK and AIR 1982a), involves nucleotides 207–251. Other N1 strains isolated between 1933 and 1935 show similar deletions which differ in size and position (see Sect. 2.3).

The nucleotide sequences of A/PR/8/34 and A/RI/5$^-$/57 differ by approximately 50% with the expected (ROBERTSON 1979) high conservation of sequence at the 5′ noncoding region of the cDNA and at the 3′ noncoding end (21 identical residues). Other regions of sequence that are highly conserved between A/PR/8/34 (N1) and A/RI/5$^-$/57 (N2) are nucleotides 20–54, 449–459, 738–750, 845–865, 890–904, and to a lesser extent, 689–708 (ELLEMAN et al. 1982).

The overall amino acid sequence homology between the N1 and N2 proteins is 40%. As shown in Table 26, the sequence homology is not distributed evenly over the length of the molecule. As previously reported (BLOK and AIR 1982b),

Table 26. Comparison of gene and protein sequences of an N2 (ELLEMAN et al. 1982) and an N1 (FIELDS et al. 1981) neuraminidase. The nucleotide and amino acid sequence differences are shaded

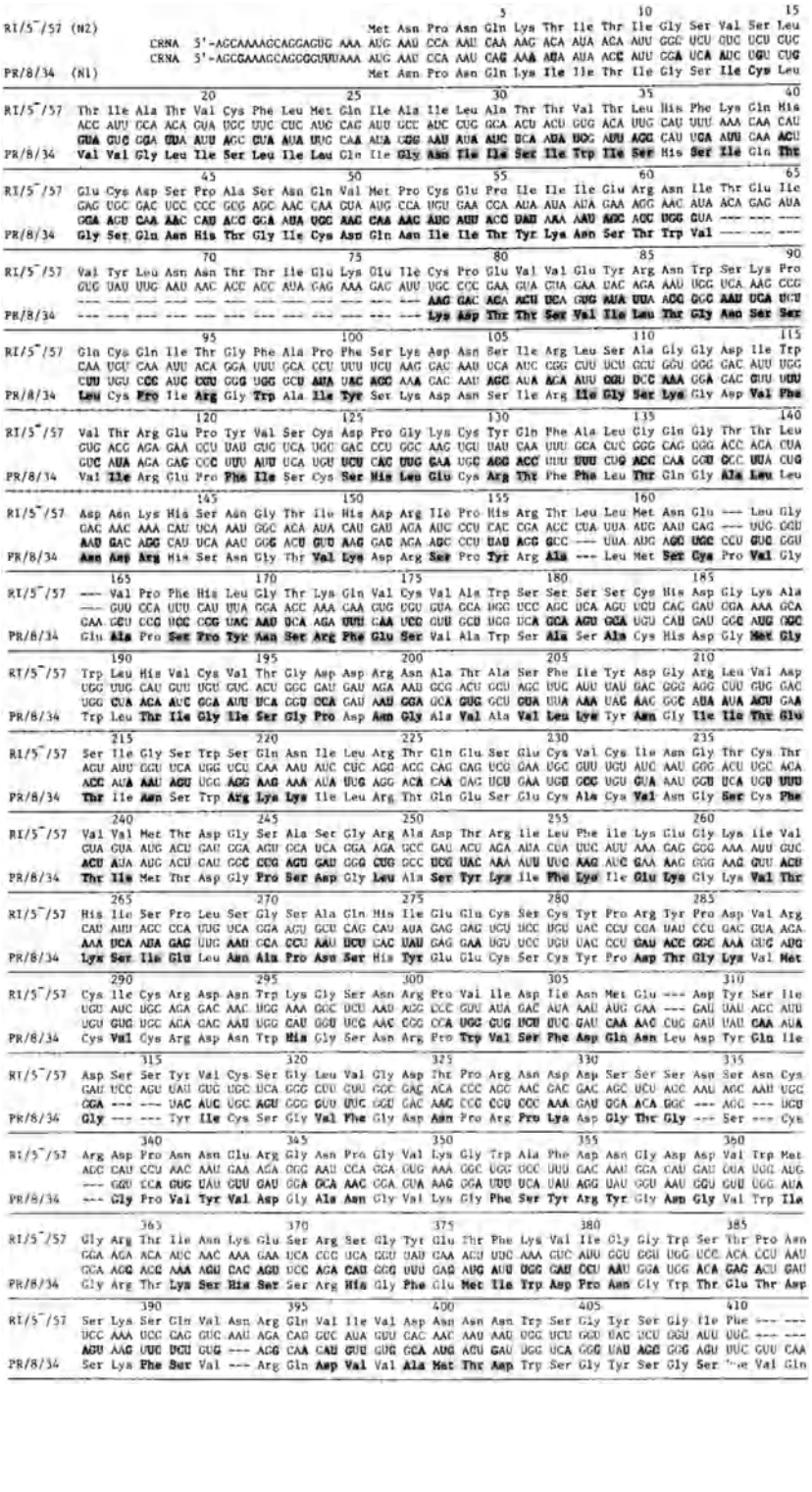


Table 26 (continued)

```
                                  415                 420                 425                 430
RI/5⁻/57  Ser Val Glu --- --- Gly Lys Ser Cys Ile Asn Arg Cys Phe Tyr Val Glu Leu Ile Arg Gly Arg Pro Gln Glu
          UCU GUU GAG --- --- GGC AAA AGC UGC AUC AAU AGG UGC UUU UAU GUG GAG UUG AUA AGC GGA AGG CCA CAG GAG
          GAU GCU GAG CUA ACA GCG GUA GAC UCU AUA AGG CCG UGC UUC UGG GUU GAA UUA AUC AGC GGA DGA CCU AAA GAA
PR/8/34   His Pro Glu Leu Thr Gly Leu Asp Cys Ile Arg Pro Cys Phe Trp Val Glu Leu Ile Arg Gly Arg Pro Lys Glu

          435                 440                 445                 450                 455
RI/5⁻/57  Thr Arg Val Trp Trp Thr Ser Asn Ser Ile Val Val Phe Cys Gly Thr Ser Gly Thr Tyr Gly Thr Gly Ser Trp
          ACU AGA GUA UGG UGG ACC UCA AAC AGU AUU GUU GUG UUU UGU GGC ACU UCA GGU ACU UAU GGA ACA GGC UCA UGG
          AAA ACA AUC --- UGG ACU AGU GCG ACG AGE AUU UCU UUU UGU GGC GUG AAU AGU GAU ACU GUA GAU UGG UCU UGG
PR/8/34   Lys Thr Ile --- Trp Thr Ser Ala Ser Ser Ile Ser Phe Cys Gly Val Asn Ser Asp Thr Val Asp Trp Ser Trp

          460                 465
RI/5⁻/57  Pro Asp Gly Ala Asn Ile Asn Phe Met Pro Ile
          CCU GAU GGG GCG AAC AUC AAU UUC AUG CCU AUA UAA GUU UCG CAA UUU UAG AAA AAA ACUCGUUGUUUCUACU-3'
          CCA GAC GGU GCU GAG UUG CCA UUC ACE AUU GAC AAG UAG UCU GUU C-- --- AA AAA ACUCCUUGUUUCUACU
PR/8/34   Pro Asp Gly Ala Glu Leu Pro Phe Thr Ile Asp Lys
```

the first six amino acids are conserved in all subtypes, and the next six residues conserved in most subtypes. From then on there are major differences in sequence. There is little homology in the hydrophobic membrane attachment sequence and no homology in the region (residues 37–73) which comprises the narrow-stalk region of the protein. In this stalk region neither the half-cystine residues nor the potential glycosylation sites are equal in number, nor do they occur in equivalent positions. The length of the stalk also varies, as some N1 strains contain deletions of up to 16 amino acid residues in this region (BLOK and AIR 1982a).

In contrast, the amino acid sequences of the globular head regions of the N1 and N2 neuraminidases are similar (Table 26), the homology from residues 92 to the C-terminus being 46%. The half-cystine residues are highly conserved in the heads, with 16 of the 17 N1 Cys residues present in identical positions in the N2 sequence. There is no Cys residue at position 162 in N2, and there are three Cys residues at positions 78, 175, and 193 in N2 that are absent from the reported N1 structure.

The oligosaccharide distribution over the N1 and N2 heads is also similar (Table 26). The glycosylation sites at residues 146 and 234 in the N2 protein are present at the same position in the N1 sequence, while that at residue 86 in N2 is within two residues of a potential site in N1. The fourth glycosylated asparagine at position 200 is not present in the N1 enzyme. In contrast to N2 neuraminidase (WARD et al. 1983b), nothing is known about the chemical composition or antigenic characteristics of the carbohydrate side chains of N1 neuraminidase.

The new classification of influenza viruses recognizes 13 subtypes of hemagglutinin and 9 subtypes of neuraminidase. BLOK and AIR (1980, 1982b) have studied the nucleic acid sequences (257 to 341 nucleotides) at the 3′ end of the neuraminidase genes in eight of these nine neuraminidase subtypes. The sequences represent part of the noncoding sequence, the sequence coding for the hydrophobic transmembrane portion of the molecule, and variable regions of the stalk. The predicted amino acid sequences for the first 94 residues are shown in Table 27. The most striking feature is that the first six amino acids of all neuraminidase subtypes are identical, while the next six amino acids are the same in five of the eight subtypes. The N3 subtype differs at position 12, the N6 at positions 9, 11, and 12, while the N7 differs in all six amino acids

Table 27. N-terminal sequences of eight different neuraminidase subtypes. Data based on cDNA sequences of BLOK and AIR (1982c). Regions of sequence homology across at least four subtypes are shaded

N1	Met	Asn	Pro	Asn	Gln	Lys	Ile	Ile	Thr	Ile	Gly	Ser	Ile	Cys	Met	Val	Val	Gly	Ile	Ile	Ser	Leu	Ile	Leu	Gln
N4	Met	Asn	Pro	Asn	Gln	Lys	Ile	Ile	Thr	Ile	Gly	Ser	Ala	Ser	Ile	Val	Leu	Thr	Thr	Ile	Gly	Leu	Leu	Leu	Pro
N5	Met	Asn	Pro	Asn	Gln	Lys	Ile	Ile	Thr	Ile	Gly	Ser	Ala	Ser	Leu	Gly	Leu	Val	Ile	Phe	Asn	Ile	Leu	Leu	His
N8	Met	Asn	Pro	Asn	Gln	Lys	Ile	Ile	Thr	Ile	Gly	Ser	Val	Ser	Leu	Gly	Leu	Val	Cys	Leu	Asp	Ile	Leu	Leu	His
N2	Met	Asn	Pro	Asn	Glu	Lys	Ile	Ile	Thr	Ile	Gly	Ser	Val	Ser	Leu	Thr	Ile	Ala	Thr	Ile	Cys	Phe	Leu	Met	Gln
N3	Met	Asn	Pro	Asn	Gln	Lys	Ile	Ile	Thr	Ile	Gly	Val	Val	Asn	Thr	Thr	Leu	Ser	Thr	Ile	Ala	Leu	Leu	Ile	Gly
N6	Met	Asn	Pro	Asn	Gln	Lys	Ile	Ile	Cys	Ile	Ser	Ala	Thr	Gly	Met	Thr	Leu	Ser	Val	Val	Ser	Gln	Leu	Ile	Gly
N7	Met	Asn	Pro	Asn	Gln	Lys	Leu	Phe	Ala	Leu	Ser	Gly	Val	Ala	Ile	Ala	Leu	Ser	Val	Met	Asn	Leu	Leu	Ile	Gly
					5					10					15					20				25

N1	Ile	Gly	Asn	Ile	Ile	Ser	Ile	Trp	Ile	Ser	His	Ser	Ile	Gln	Thr	Gly	Ser	Gln	Asn*	His	Thr	Gly	Ile	Cys	Asn*
N4	Ile	Thr	Ser	Leu	Cys	Ser	Ile	Trp	Phe	Ser	His	Tyr	Asn	Gln	Gly	Thr	Gln	Pro	His	Glu	Gln	Ala	Cys	Ser	Thr
N5	Gly	Ala	Ser	Ile	Thr	Trp	Gly	Thr	Ile	Ser	Val	Thr	Lys	Asp	Asn	Lys	Val	His	Ile	Cys	Asn*	Thr	Thr	-	-
N8	Ile	Ile	Ser	Ile	Thr	Ile	Thr	Val	Leu	Gly	Leu	His	Lys	Asn	Gly	Lys	Gln	Arg	Arg	Cys	Asn*	Glu	Thr	Val	Ile
N2	Ile	Ala	Met	Leu	Val	Thr	Thr	Val	Thr	Leu	His	Phe	Lys	Gln	Tyr	Glu	Cys	Asp	Ser	Pro	Gly	Asn	Asn	Gln	Val
N3	Val	Gly	Asn	Leu	Val	Phe	Asn	Thr	Val	Ile	His	Glu	Lys	Ile	Gly	Asn	His	Gln	Thr	Val	Ile	His	Pro	Thr	Ile
N6	Leu	Ala	Asn	Leu	Gly	Leu	Asn	Ile	Gly	Leu	His	Phe	Lys	Val	Gly	Glu	Thr	Pro	Glu	Ile	Gly	Thr	Pro	Ser	Val
N7	Ile	Ser	Asn	Val	Gly	Leu	Asn	Val	Ser	Leu	His	Leu	Lys	Glu	Lys	Gly	Thr	Lys	Gln	Glu	Glu	Asn*	Leu	Thr	Cys
					30					35					40					45				50

N1	His	Ser	Ile	Ile	Thr	Tyr	Lys	Asn*	Ser	Thr	Trp	Val	Asn*	Gln	Thr	Tyr	Val	Asn*	Ile	Ser	Asn	Thr	Asn	Val	Val
N4	Thr	Gln	Arg	Ile	Thr	Ile	Asn*	Glu	Thr	Phe	Val	Asn*	Val	Thr	Asn	Val	Gln	Asn	Asn*	Tyr	Thr	Thr	Ile	Ile	Asp
N5	Glu	Ala	Tyr	Asn*	Glu	Thr	Ala	Arg	Ala	Gln	Lys	Val	Val	Ile	Pro	Val	Asn*	Asn	Thr	Ile	His	Ser	Asn	His	Glu
N8	Arg	Glu	Asp	Asn*	Glu	Thr	Val	Arg	Ile	Glu	Lys	Val	Thr	Gln	Trp	His	Asn	Thr	Asn	Val	Ile	Glu	Tyr	Ile	Glu
N2	Thr	Leu	Cys	Glu	Pro	Ile	Ile	Ile	Glu	Arg	Asn*	Ile	Thr	Glu	Ile	Val	His	Leu	Thr	Asn*	Thr	Thr	Ile	Glu	Lys
N3	Thr	Thr	Pro	Ala	Val	Pro	Asn*	Cys	Ser	Asp	Thr	Ile	Ile	Thr	Tyr	Asn	Asn	Thr	Val	Ile	Asn	Asn*	Ile	Thr	Thr
N6	Asn*	Glu	Thr	Asn*	Ser	Thr	Thr	Thr	Ile	Ile	Asn	Tyr	Asn	Thr	Gln	Asn	Asn*	Phe	Thr	Asn*	Val	Thr	Asn	Ile	Val
N7	Thr	Thr	Ile	Thr	Gln	Asn*	Asn*	Thr	Thr	Val	Val	Glu	Asn	Thr	Tyr	Val	Asn*	Asn*	Thr	Thr	Ile	Ile	Thr	Lys	Glu
					55					60					65					70				75

N1	Ala	Gly	-	-	-	-	-	-	-	-	-	-	-	-	-	Lys	Asp	Thr							
N4	Pro	Gln	Pro	Pro																					
N5	Pro	Glu	-	-	-	-	-	-	Phe	Leu	Asn*	Asn	Thr	Glu	Pro	Leu	Cys	Asp	Val						
N8	Lys	Leu	Glu	Gly	Asp	His	-	-	Phe	Met	Asn*	Asn	Thr	Glu	Pro	Leu	Cys	His	Ala						
N2	Glu	Ile	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Cys	Pro	Lys						
N3	Thr	Ile	Ile	Thr	Glu	Ala	Glu	Arg	Leu	Phe	Lys	Pro	Pro	Leu	Pro	Leu	Cys	Pro	Ser						
N6	Leu	Ile	Lys	Glu	Glu	Asp	Glu	Met	Phe	Thr	Asn*	Leu	Ser	Lys	Pro	Leu	Cys	Glu	Val						
N7	Pro	Asp	Leu	Lys	Ala	Pro	Ser	Tyr	Leu	Leu	Leu	Asn*	His	Ser	-	Leu	Cys	Ser	Val						
					80					85					90					95				100

(positions 7–12). From here on, the differences between subtypes increase, and BLOK and AIR (1982b) concluded that the eight subtypes were totally different in this region of the molecule, although some homology can be seen in Table 27 extending through the membrane-spanning peptide to about residue 43. Beyond this point, alignment of the sequences is no longer meaningful, with the possible exception of N5 and N8, which do show some homology. Sequence homology appears again at residue 84 (N3 and N6 numbering), near the beginning of the head sequence.

In their extended sequence analysis, BLOK and AIR (1982b) observed that all eight subtypes had at least one Cys residue in the first 75 amino acids

and several subtypes contained more than one cysteine. However, there was no apparent homology in the sequences surrounding these Cys residues, with the exception of the Cys residues near the end of the stalk region, where a consensus sequence Pro Leu Cys X X X Gly Phe/Trp X Ile was reported (BLOK and AIR 1982b).

In their alignment, BLOK and AIR (1982b) included the N1 PR/8/34 sequence around Cys 77. This residue is more likely to be homologous to Cys 92 in N2 sequences (see Tables 4 and 26) and is part of the globular head structure, being disulfide-bonded to Cys 417.

5.4 Sequence Differences Between Type A and Type B Neuraminidase

The nucleotide sequence in the neuraminidase gene of B/Lee/40 shows several features distinguishing it from the A strain structures (SHAW et al. 1982). The first AUG codon, nucleotides 47–49, has an open reading frame of only 300 nucleotides (see Table 7). The second AUG codon, 54–56, is followed by an open reading frame coding for a protein, neuraminidase, of 466 amino acids. The use of the second AUG codon as the site for initiation of translation differs from other influenza genes so far sequenced and is uncommon, though not unprecedented, in eukaryotic mRNAs (KOZAK 1981).

Perhaps the most striking feature of the B gene structure is the departure from the strict homology seen in A sequences over the first six amino acids (Table 28). It may be significant that reassortments of A and B genes do not occur and that these six amino acids in the neuraminidase structure are believed to be on the cytoplasmic side of the membrane.

The nontranslated region at the 5′ end of the cDNA contains 46 nucleotides compared with 20 and 19 in the N1 and N2 strains, and at the 3′ end 100 nucleotides, compared with 31 and 41 in N1 and N2 (Tables 7 and 26). In A and B strains, polyA sites are located 15 and 16 nucleotides respectively from the 3′ end.

SHAW et al. (1982) have compared the complete nucleotide and protein sequence of B/Lee/40 neuraminidase with both N1, A/PR/8/34 and WSN/33, and N2, Udorn/72/sequences. The most extensive homology they detect is in the central 744 nucleotides of B/Lee/40, encoding amino acid residues 116 to 363. In this region, 87 of the 247 amino acids are unchanged between B and N2, including 12 half-cysteine residues and the glycosylation site at Asn/144 (146 in N2 numbering). Within the head regions alone, the pairwise homology of B/N1 and B/N2 sequences is 29% and 26% respectively, compared with 43% for N1/N2 (COLMAN 1984). The lower level of homology between the A and B strain sequences is largely due to differences in the 100 C-terminal residues, where N2 and B have only eight amino acids in common, and N1 and B only 13 (COLMAN, 1984). Three of the conserved amino acids in this region are cysteine residues. Of residues believed to be important in enzyme activity, only Glu 425 (N2) is found in this region, and it is conserved across

Table 28. Amino acid sequences of N2 (ELLEMAN et al. 1982), N1 (FIELDS et al. 1981), and B (SHAW et al. 1982) neuraminidase. Regions of sequence homology are shaded

```
                          20                  40                  60
N2 (A/RI/5~/57)   NNPNQKIITIGSVSLTIATV CPLMQIAILATTVTLHFKQH ECDSPASNQVMPCEPIIIER
N1 (A/PR/8/34)    MNPNQKIITIGSICLVVGLI SLILQIGNIISIWISHSIQT GSQNHTGICNQNIITYKNST
B  (B/Lee/40)     MLPSTVQ-TLTLLLTSGGVL LSLYVSASLSYLLYSDVLLK FSSTKTTAPTMSLECTNASN

                          60                  100                 120
N2 (A/RI/5~/57)   NITEIVYLNNTTIEKEICPE VVEYRNWSKPQCQ---ITGFAPF SKDNSIRLSAGGDIWVTREP
N1 (A/PR/8/34)    WV----------------KDT TSVILTGNSSLCP---IRGWAIY SKDNSIRIGSKGDVFVIREP
B  (B/Lee/40)     AQTVNHSATKEMTF----PP PEPEWTYFRLSCQGSIFQKALLI SPHRFGEIKGNSAPLIIREP

                          140                 160                 180
N2 (A/RI/5~/57)   YVSCDPGKCYQFALGQGTTL DNKHSNGTIHDRIPHRTLLM NELG-VPFHLGTKEVCVAWSS
N1 (A/PR/8/34)    FISCSHLECRTFFLTQGALL NDRHSNGTVKDRSPYRALMS CPVGEAPSPYNSRFESVAWSA
B  (B/Lee/40)     FVACGPKECRHFALTHYAAQ PGGYYNGTRKDRNKLRHLVS VKLGKIPTVENSIFHMAAWSG

                          200                 220                 240
N2 (A/RI/5~/57)   SSCHDGKAWLHVCVTGDDRN ATASFLYDGRLVDSIGSWSQ NILRTQESECVCINGTCTVV
N1 (A/PR/8/34)    SACHDGMGWLTIGISGPDNG AVAVLKYNGIITETIKSWRK KILRTQESECACVNGSCFTI
B  (B/Lee/40)     SACHDGKEWTYIGVDGPDND ALVKIKYGEAYTDTYHSYAH NILRTQESACNCIGGDCYLM

                          260                 280                 300
N2 (A/RI/5~/57)   MTDGSASGRADTRILFIKEG KIVHISPLSGSAQHIEECSC YPRYPDV-RCICRDNWKGSNR
N1 (A/PR/8/34)    MTDGPSDGLASYKIFKIEKG KVTKSIELNAPNSHYEECSC YPDTGKVM-CVCRDNWHGSNR
B  (B/Lee/40)     ITDGSASGISKCRFLKIREG RIIKEILPTGRVEHTEECTC GFASNKTIECACRDNSYTAKR

                          320                 340                 360
N2 (A/RI/5~/57)   PVIDINME-DYSIDSSYVCSG LVGDTPRNDDSSSNSNCRDP NNERGNPGVKGWAFDNGDDV
N1 (A/PR/8/34)    PWVSFDQNLDYQIG--YICSG VFGDNPRPKDGT--GSCG-P VYVDGANGVKGFSYRYGNGV
B  (B/Lee/40)     PFVKLNVETDTAEI-RLMCTK TYLDTPRPDDGSIAGPCES- NGDKWLGGIKGGFVHQRMAS

                          380                 400                 420
N2 (A/RI/5~/57)   WMGRTINKESRSGYETFKVI GGWSTPNSKSQVNRQVIVDN SNWSGYSGIF---SVEGK--SCINR
N1 (A/PR/8/34)    WIGRTKSHSSRHGFEMIWDP NGWTETDSKFSV-RQDVVAM IDWSGYSGSFVQHPELTG-LDCIRP
B  (B/Lee/40)     KIGRWYSRTMSKTNRMGMEL YVKYDGDPWTDS-DALTLSG VMVSIEEPGWYSFGFEIKDKECDVP

                          440                 460
N2 (A/RI5~/57)    CFYVELIRGRPQETRVWWTS NSIVVFCGTSGTYGTGSWPD GANINFMPI-
N1 (A/PR/8/34)    CPWVELIRGRPKETTI-WTS ASSISFCGVNSDTVDWSWPD GAELPFTIDK
B  (B/Lee/40)     CIGIEMVHDGGKDT---WHS AATAIYCLMGSGQLLWDTVT GVDMAL----
```

A and B sequences. Table 28 aligns the N1, N2, and B strain neuraminidase sequences. Amino acid insertions and deletions are discussed in Sect. 5.5.4.

5.5 The Three-Dimensional Antigenic Structure

COLMAN et al. (1983) have analyzed the three-dimensional structure of N2 neuraminidase in the light of sequence changes observed in field strains between 1957 and 1975 and in monoclonal variants. In a similar study of the hemagglutinin molecule, WILEY et al. (1981) used the data for strains with the greatest epidemic impact to underscore the observation that between these epidemics amino acid sequence changes occur in each of four major antigenic regions. The neuraminidase analysis, originally covering 18 years of antigenic drift (COLMAN et al. 1983), is less straightforward for two reasons. Firstly, no clustering of amino acid sequence changes into distinct regions is observed. Rather, the sequence variation is seen to be primarily on surface loops of the protein structure, encircling the catalytic site. In the original hemagglutinin analysis (WILEY et al. 1981), sequence data was available for the period 1968–1975. As more amino acid sequence data become available for hemagglutinin, the original antigenic sites are being seen as less distinct than first thought. Apparently, most of the solvent-accessible surface on the top of the globular domain of the hemagglutinin structure has undergone amino acid substitutions, either in naturally occurring field strains or in monoclonal variants (DANIELS et al. 1983). Sec-

ondly, the role of antibodies to neuraminidase in protecting against infection is less clear-cut than it is with antihemagglutinin antibodies which are known to neutralize infectivity. Antineuraminidase antibodies significantly modify the disease in favor of the host. It has been observed (SCHULMAN 1975) that reduced titers of virus and numbers of lung lesions result from protection with antibodies to neuraminidase. These data are consistent with the notion (Sect. 1) that neuraminidase is important for dispersing progeny virus particles from the initial site of infection, and they suggest an important epidemiological role for antineuraminidase antibodies. The selection of monoclonal variants of neuraminidase by monoclonal antibodies to neuraminidase (Sect. 5.2.1) is further evidence of the role such antibodies play in effectively neutralizing a particular virus strain.

5.5.1 Spatial Extent of an Antigenic Epitope

No data addressing this question directly are available. The three-dimensional structures of a number of Fab fragments from antibodies are known (POLJAK et al. 1973; SEGAL et al. 1974; MATSUSHIMA et al. 1978; MARQUART et al. 1980). The specificities of these antibodies, being in all cases myeloma proteins, are not known, although in some cases small haptens have been observed to bind them with modest affinity ($K_a \sim 10^{-5}$ *M*). Recently, some success has been reported in crystallizing Fab fragments from monoclonal antibodies of known protein specificity (COLMAN et al. 1981). The three-dimensional structure of the antineuraminidase Fab fragment (COLMAN et al. 1981) has now been determined at 3.5 Å resolution (P.M. COLMAN, unpublished). The amino acid sequence for that Fab fragment is currently being determined (G.M. AIR, personal communication). Crystals of antilysozyme Fab complexed to lysozyme have also been reported (MARIUZZA et al. 1983).

In all of the X-ray structures of Fab fragments that have been published, the six hypervariable loops on the variable domains of heavy and light chains are observed to cluster around the distal end of the Fab arm on the antibody molecule. In most cases these loops surround a cavity in which haptens such as phosphocholine (SEGAL et al. 1974) or vitamin K_1OH (AMZEL et al. 1974) have been observed to bind. In one case (MARQUART and DEISENHOFER 1982), there is no cavity for hapten binding. The KOL myeloma protein (COLMAN et al. 1976; SCHMIDT et al. 1983) has a large insertion in the third complementarity-determining region of the heavy chain, and this extra peptide is folded back into the surface of the protein to form a blunt-ended Fab structure (MARQUART and DEISENHOFER 1982). Despite differences in the size and shape of the antigen-combining region of antibodies, all structures show a common feature. The complementarity-determining regions form a surface which is approximately 30 Å in diameter (COLMAN et al. 1983), suggesting that the area of the interface between antigen and antibody might be of the order of $\pi \times 15^2$ Å^2, or 700 Å^2.

Independent data on protein-protein interactions suggest that this figure is not unreasonable, at least for protein antigens. CHOTHIA and JANIN (1975) conclude that interfaces between stable protein-protein complexes need to bury

approximately 600 Å^2 of surface area from each contributor, giving rise to a total of about 1200 Å^2 of surface removed from solvent, or a contribution of $\sim$30 kcal mol^{-1} of hydrophobic energy to the total free energy of interaction. Typical figures for some protein-protein interfaces are 1130 Å^2 for the insulin dimer, 1390 Å^2 for the pancreatic trypsin inhibitor-trypsin complexes, and 1720 Å^2 for α-β subunits of hemoglobin (CHOTHIA and JANIN 1975). For the subunits, or domains, of antibodies, figures are 1760 Å^2 for V_L-V_H domain pairing, 1923 Å^2 for C_L-C_H1 pairing, and 2180 Å^2 for C_H3-C_H3 dimerization (MARQUART and DEISENHOFER 1982). These data for the interfaces of domains that measure $\sim 40 \times 20 \times 20$ Å also support the notion that, as a first approximation, the buried area can be calculated from the dimensions of the contact surface of each domain, in this case, $\sim 40 \times 20$ Å. In contrast, the contact between C_H2 and C_H3 in the Fc region of antibodies, known to be a "soft" contact, the total surface area excluded from both domains in the interface is only 778 Å^2 (DEISENHOFER 1981).

5.5.2 Location of Regions of Field Strain Variation on the Three-Dimensional Structure

Figure 8 shows diagrams of the three-dimensional structure indicating all surface amino acids which changed in the period 1957–1979 (data are from Table 1 of COLMAN et al. 1983), including residues 220, 248, and 313, which changed between 1975 and 1979 (MARTINEZ et al. 1983). Apart from some clustering on the underneath surface of the neuraminidase head (residues 126–127, 258, 308, 356–358) and some scattered changes on the side surface (residues 253, 269, 390), most of the field strain sequence variation occurs on the upper surface, where it encircles the enzyme active site (compare Fig. 8 with Fig. 7). These

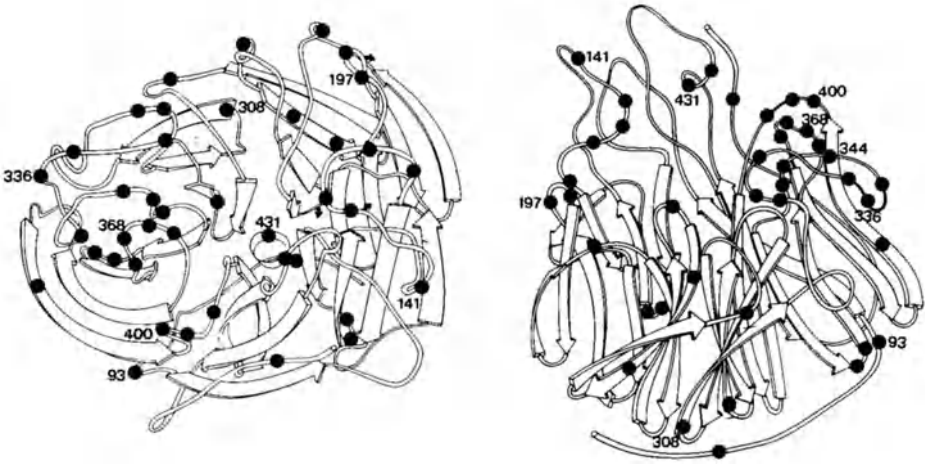


Fig. 8. Location of all surface amino acid changes documented between 1957 and 1979 N2 neuraminidase sequences

Table 29. Shortest distances (Å) between C^{α} atoms of the seven variable segments in the three-dimensional structure

Family	I	II	III	IV	V	VI	VII
Chain segments	328–336	339–347	367–370	400–403	431–434	197–199	153
328–336	–	6	11	20	19	>	>
339–347	6	–	6	15	16	>	>
367–370	11	6	–	9	14	>	>
400–403	20	15	9	–	12	(16)	(17)
431–434	19	16	14	12	–	>	>
197–199	>	>	>	(16)	>	–	8
153	>	>	>	(17)	>	8	–
NANA	10	10	10	13	11	10	13

Distances greater than 20 Å are shown >. Numbers in parentheses indicate distances measured across the subunit interface. Also shown are the closest distances from the observed binding site of sialic acid (NANA) on the enzyme. Roman numerals refer to the family of the antigenic determinant defined as the association of a particular chain segment in the binding pocket of an antibody

changes occur on the loops L_{01} and L_{34} connecting the various strands of β-sheets in the propeller framework of the structure.

As a first approximation, COLMAN et al. (1983) have described the upper surface sequence variation in N2 neuraminidase as seven overlapping antigenic determinants. Seven chain segments of field strain variation (328–336, 339–347, 367–370, 400–403, 431–434, 197–199, and 153) are the basis for the seven determinants. Table 29 shows the closest distances between these chain segments together with their closest approach to the sialic acid binding site in the enzyme active center (COLMAN et al. 1983). Each chain segment is envisaged as having the capacity to bind to the central cavity or region of the complementarity-determining surface of the antibody. Neighboring chain segments within a radius of 15 Å are expected to be capable of forming part of that particular antigenic determinant. They would interact with the periphery of the complementarity-determining surface of the antibody. Broadly speaking, the families separate into two spatially distinct regions to the left- and right-hand side of the active site as seen in Fig. 8, although they are tenuously connected across the subunit interface as indicated in Table 29.

All of these chain segments are, in principle, sufficiently close to the active site to block enzyme activity on neuraminyllactose (Table 29). In practice, certain binding modes of antibodies might not be inhibitory.

Following WILEY et al. (1981), Fig. 9 shows the location of amino acid sequence changes between 1968 and 1972 and between 1972 and 1975. The strains used are NT/60/68, Udorn/72, and Vic/3/75, and only those changes which are presently adjudged to be on or in the surface of the three-dimensional structure are shown (see Table 1 in COLMAN et al. 1983). Against the backdrop of all the changes occurring between 1957 and 1975 (Fig. 8), the changes taking

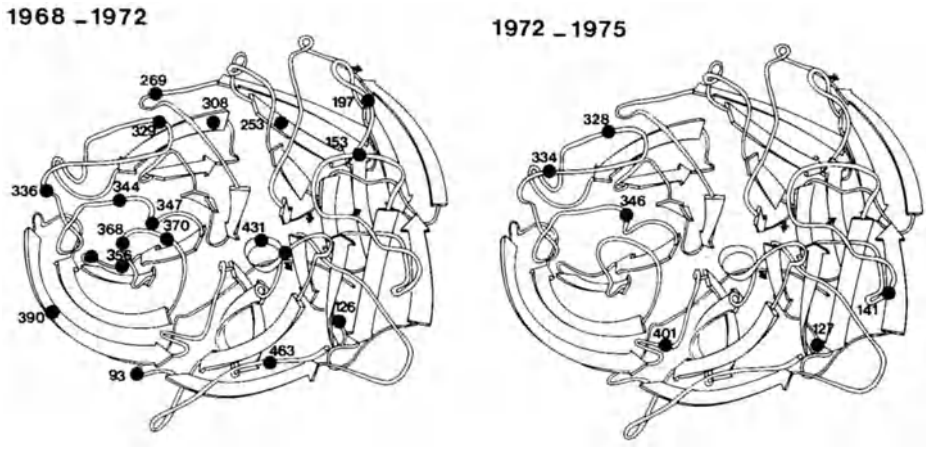


Fig. 9. Location of surface amino acid changes between 1968 and 1972 and between 1972 and 1975

place between 1968 and 1972 are seen to involve all of the regions of change on the upper surface of the molecule, while those occurring between 1972 and 1975 are both fewer in number and exclude chain segments around residues 153 and 198. Those chain segments, however, are only some 16 Å distant from the segment 400–403 on a neighboring subunit and could form part of a single antigenic determinant for some antibodies sensitive to changes in that region (COLMAN et al. 1983). Thus, the analysis of these three epidemic periods shows that none of the variable regions around the active site pocket escapes without some alteration.

The significance of variation on the outer and under surfaces of the molecule, where binding of antibody would not be expected to interfere with the enzyme activity, is not clear. Presumably there is no single mechanism by which antibodies neutralize infection. Direct inhibition of some viral function, such as receptor binding or neuraminidase activity, is one mechanism. Antibody-dependent or -independent complement-mediated lysis is another (HIRSCH 1982).

The recently determined sequence of the neuraminidase gene of A/Bangkok/1/79 (MARTINEZ et al. 1983) has added to the previous picture of variation by indicating a spatial clustering of variable residues near amino acid 269 at positions 248 and 313. Antibodies attached to this region could marginally affect enzyme activity. Antibodies binding to the bottom-surface variable loops will not affect enzyme activity, at least not directly.

Amino acid sequence differences between the 1933 and 1934 N1 neuraminidases are shown mapped onto the N2 structure in Fig. 10 and, as remarked earlier (Sect. 5.2.2), reflect biological and antigenic differences between these two proteins. Variation at or around 149, 199, 221, 343, 368, 400, and 432 (N2 numbering) is consistent with the assignment of these chain segments to antigenic sites on the basis of N2 sequences and structure. The grounds for using the N2 structure as a framework for mapping N1 sequence variation are discussed in Sect. 5.5.5. The loss of the glycosylation sequence at residue

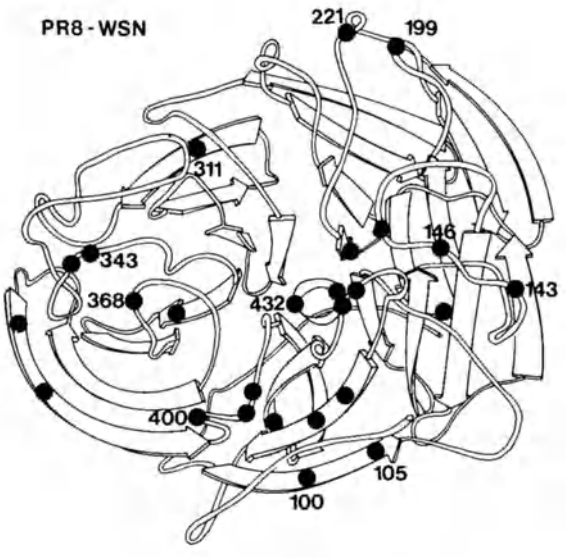


Fig. 10. Location of amino acid sequence differences between the two N1 neuraminidase strains PR/8/34 and WSN/33 mapped onto a diagram of the N2 neuraminidase structure

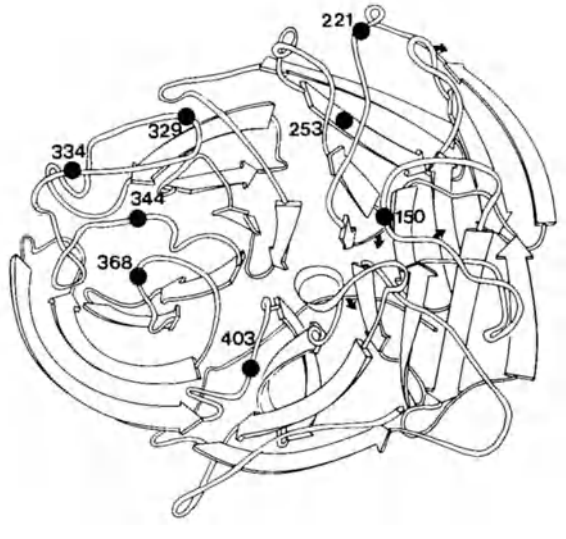


Fig. 11. Sites of variation characterized in monoclonal variants of N2 neuraminidase

146 (N2 numbering) might expose new antigenic regions on the protein surface or modify the accessibility of nearby amino acids to antibodies. Carbohydrate is believed to play a role in masking antigenic determinants in hemagglutinin (WILEY et al. 1981; DANIELS et al. 1983; CATON et al. 1982).

5.5.3 Location of Monoclonal Variant Substitutions on the Three-Dimensional Structure

The positions of amino acid substitutions in all the monoclonal variants of neuraminidase are indicated in Fig. 11. In all cases, they map to regions of

the structure showing amino acid sequence variation in field strains. Some of the monoclonal antibody data shown in Table 15 can now be rationalized. For example, the ability of the anti-Tokyo/67 neuraminidase monoclone S10/1 to distinguish between A/England/42/72 and A/Udorn/72 neuraminidase is correlated with the change of Lys 368 to Glu 368 and the selection of a monoclonal variant (POT, see LAVER et al. 1982) by that antibody showing the same amino acid sequence change. Similarly, the anti-Tokyo/67 monoclones 25/4, S25/3, and S32/3 distinguish those two 1972 field strains, and they all select variants at residue 344 of Arg to Ile. The two field strains differ at that amino acid by Arg to Lys, but both differ from the A/Tokyo/3/67 sequence at residues 346 and 347 by Thr-Gln to Asn-His. Either the latter two changes are not sensed by the binding of any of those antibodies to a site on the neuraminidase which includes residue 344, or else that binding can accommodate these changes. The substitution at 368 of Lys for Glu between England/72 and Udorn/72 may be the determining factor in differentiating these two strains to those monoclonal antibodies selecting variants at residue 344. Positions 344 and 368 are ~ 6 Å apart in the three-dimensional structure.

The anti-Texas/1/77 monoclonal antibody 18/1 recognizes a determinant that is unaltered between 1967 and 1979 (WEBSTER et al. 1982). There is some question as to the location of the field strain variation that gives rise to this reactivity pattern (Table 15). That antibody selects a variant at residue 221 of Asn to His in a chain segment that shows no field strain variation between 1957 and 1975. The Bangkok/1/79 sequence has a substitution at residue 220 of Gln to Lys not sensed by that antibody. Nearby in the three-dimensional structure, the variable segment 197–199 undergoes substitution at position 197 (Asp to Tyr) between 1968 and 1972. This segment is therefore not an important part of the determinant recognized by 18/1. Residue 199 changes from Arg to Lys between 1957 and 1967 and is thereafter unchanged. With present data, the only explanation for the reactivity pattern of the 18/1 monoclone is that it is sensitive to this Arg to Lys substitution, a change normally considered conservative but which has nevertheless been observed in a monoclonal variant of Tokyo/67 neuraminidase.

The panel of monoclonal antibodies to A/RI/5$^-$/57 (WEBSTER et al. 1984) imply four separate antigenic areas. Group 1 antibodies recognize a region of neuraminidase that is unaltered between 1957 and 1968 (Table 21). The site is believed to be on the bottom surface of the subunit (WEBSTER et al. 1984) and residues 307 and 308 fit these requirements. Like group 1 antibodies, group 4 antibodies do not inhibit neuraminidase activity on fetuin, and they are also believed to recognize a determinant on the bottom surface of the subunit. In this case the determinant is altered between 1957 and 1967 strains. The pattern of change at residue 358 satisfies this condition, as does the observation that group 1 and 4 antibodies do not complete with each other, being approximately 22 Å distant from the 307–308 peptide. Group 3 antibodies inhibit enzyme activity on fetuin but not residue 358 neuraminyllactose and are thus thought to be near, but not too near, the catalytic site. Region 2 and its various subregions are also on the upper surface. The three monoclonal variants so far identified (G.M. AIR, personal communication) map to regions

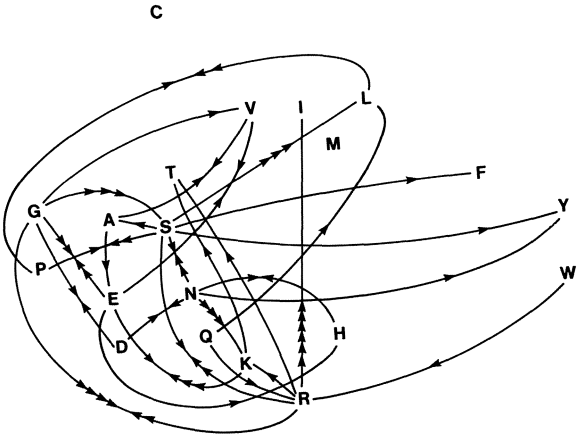


Fig. 12. Substitutions in monoclonal variants of influenza virus hemagglutinin and neuraminidase. Each *arrowhead* indicates one observation of a particular substitution. The distribution of amino acids in the map follows FRENCH and ROBSON (1983)

2a (403, Trp to Arg), 2b (334 Asn to Ser), and 2d (150 His to Asn), being selected with monoclonal antibodies 81/4, 438/1, and 145/1 respectively (Table 19) (G.M. AIR, personal communication). The association of monoclonal antibody 145/1 with residue 150 and its inhibition of enzyme activity on neuraminyllactose is in excellent agreement with the observation that that amino acid residue is immediately adjacent to the catalytic site (COLMAN et al. 1983).

The question of how the single amino acid sequence changes in monoclonal variants suffice to abolish binding by the antibody used for their selection is a complex one. While hydrophobicity contributes primarily to the affinity of a protein-protein complex, specificity requires complementarity (CHOTHIA and JANIN 1975), and the initial docking process will be determined by longer-range forces, notably polar interactions. Flexibility is likely to be an important aspect of the exploration of docking modes (BLUNDELL 1981). In the absence of any hard data on the nature of an antibody-antigen interface, it is useful to compile a list of all the single amino acid substitutions associated with the abolition of such an interaction. Data in Fig. 12 show all of the hemagglutinin monoclonal variants of Aichi/2/68 and PR/8/34 (DANIELS et al. 1983) and neuraminidase monoclonal variants of Tokyo/3/67 (LENTZ et al. 1984) and RI/5$^+$/57 (G.M. AIR, personal communication). The changes are indicated on a map of amino acid exchangeability calculated by FRENCH and ROBSON (1983). Amino acids close together on this map exchange with each other more frequently than those farther apart in homologous protein structures. A feature of the map is the observation that amino acid residues tend to group together into particular secondary structure forming groups (FRENCH and ROBSON 1983). The 63 substitutions shown in Fig. 12 include charge changes of polar to apolar and reverse, polar to polar, and apolar to apolar. Some amino acids occur infrequently on the surface of proteins, and these residues will not be the source of many substitutions. The greater degree of substitution within the polar group of amino acids reflects this bias. Many of the substitutions would be categorized as conservative in protein structure studies, none more so than the Arg to Lys variant

of Tokyo/3/67 neuraminidase (LENTZ et al. 1984). Some of these changes will clearly affect docking of antigen and antibody, and others will prejudice their affinity. It seems likely that in some cases this affinity will be lowered rather than abolished, although there are presently no data to support this. In summary, alteration of either short-range or long-range interactions can prejudice binding of an antibody to an antigen.

5.5.4 Juxtaposition of Catalytic and Antigenic Sites

The enzyme active center appears to enjoy a stereochemical privelege which effectively makes it invisible to the host immune system. That privelege might arise in one of three ways (COLMAN et al. 1983). Firstly, the active site might be inaccessible to antibodies. Indeed, the concave nature of most enzyme active sites makes them unlikely targets of antibodies with concave antigen-binding sites. Secondly, the active site might form part of antigenic determinants, but only in such a way that nonessential, variable amino acids also form part of the antigen-antibody complex. In this way, variation of the nonessential amino acids in the determinant would mask future strains from antibodies raised against a current strain. Thirdly, the active site might not be immunogenic. This situation would be most desirable if host neuraminidases had an active site similar to that of the viral enzyme.

Figure 13 shows stereo views of the active site and the surrounding region of field strain variation. On either side, two of the active-site residues, Asp 198 and Asp 330, are adjacent to amino acids which vary in field strain sequences. In both cases the active-site residues are oriented inward, towards the site of sialic acid binding, and the variable residues in the surface are oriented outward.

A careful study of the possible interaction of the active site with an antibody is not yet feasible. Certainly, some of the conserved catalytic residues could make contact with antibodies, but only under circumstances where functionally nonessential amino acids were also part of the same antigenic determinant.

5.5.5 The Three-Dimensional Structures of Type N1 and B Neuraminidase

Within the globular head sequences (residues 74–469 of N2 strains), there are many features pointing to similarity of structure for the three neuraminidase subtypes. Firstly, the overall level of homology, as shown in Table 28 and discussed in Sect. 5.4, is sufficient to direct folding into similar topologies. All of the half-cysteine residues of the head region are in alignment. N2 sequences have one additional pair of cysteine residues bridging residues 175 and 193. An additional cysteine in the B sequence at residue 251 (N2 numbering) is remote from the fourfold axis and cannot be involved in disulfide formation. Apart from conserved, catalytically important residues, clusters of amino acids conserved in all three sequences are found within the segments 146–148, 177–186, 222–237, 242–248, 276–280, and 289–294. The last four segments are presumably critical features of β-sheets 3 and 4. The upper-surface loops of these sheets are rich in residues implicated in catalytic activity (Sect. 4.5).

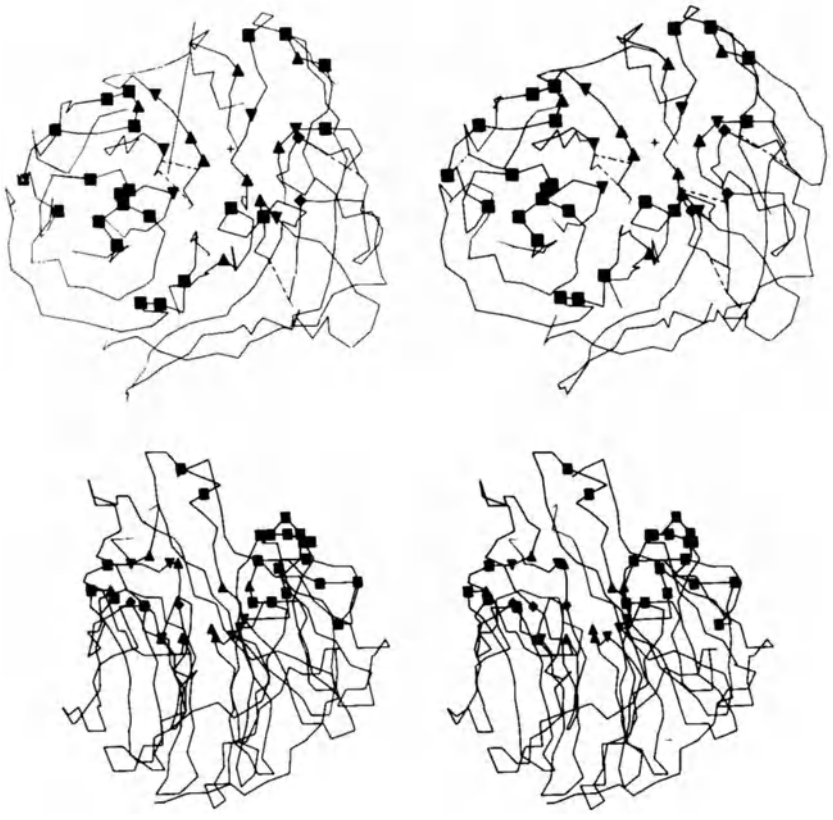


Fig. 13a, b. Stereo views of active site residues (flagged as in Fig. 7) and sites of upper surface variation (*squares*) showing juxtaposition of active site and antigenic regions. **a** View down the molecular fourfold axis; **b** view at right angles to the symmetry axis (COLMAN et al. 1983). Reprinted by permission from *Nature*. Vol. 303, pp 41–44, 1983

The insertions and deletions found in N1 and B structures relative to N2 are most commonly located in surface loops and are shown in Fig. 14.

Three residues are inserted into the B sequence at position 93. The homology with N1 and N2 in the following segment is low, but insertion here rather than elsewhere in the next heptapeptide allows the additional structure to extend the N-terminal arm prior to its entry into $\beta_6 S_4$. N1 and B both have an additional charged residue in $\beta_1 L_{34}$. Its role is unclear. $\beta_4 S_2$ (287) has an insertion and a deletion in N1 and an insertion in B. In N2, Arg 288 may be in a salt link with Asp 304. In the B structure, the same residues may be linked, Glu 288-Lys 304. The salt bridge is not present in N1, where 288 is deleted and residue 304 is Ser. On the opposite side of β_4, N1 and B build a more extensive interface with β_5 than does N2. The inserted residue 287a in these structures might contact Phe 238 or Tyr 238 on β_3, at which point N2 has Thr. N1 and B structures insert one residue at position 308 at the start of $\beta_4 S_4$. At the end of $\beta_4 S_4$, 314 is deleted in B, while 314 and 315 are deleted in N1. The disulfide bridge

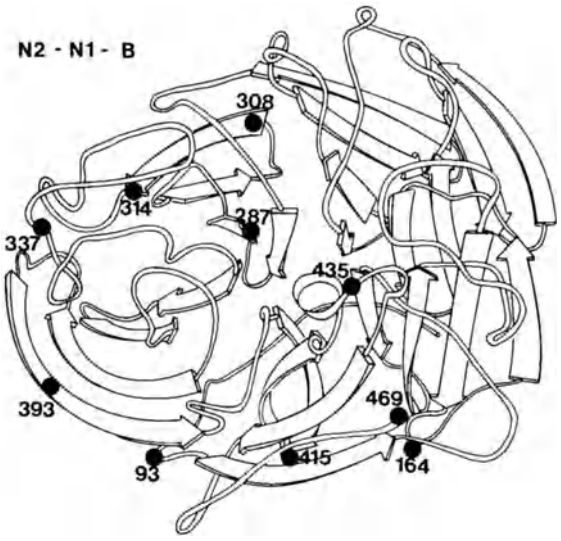


Fig. 14. Sites of insertion or deletion of N1 and B subtype neuraminidases mapped onto a diagram of the N2 structure (COLMAN 1984)

between Cys 318 and Cys 337 in $\beta_5 L_{01}$ may assist in the stabilization of a structure whose sequence varies considerably within subtypes. Deletions in N1 and B are found either side of Cys 337. $\beta_5 S_4$ demonstrates a β-bulge (RICHARDSON et al. 1978) at residue Asn 392. N1 and B have a deletion at this point. N1 and B insert four and five residues respectively in $\beta_6 L_{12}$. One and three deletions respectively are found in N1 and B at residue 435. This loop, $\beta_6 L_{23}$, represents the most distal extent of the structure from the viral membrane. Slight variations in length are found at the C-terminus.

In summary, the structures of N1 and B neuraminidase belong to the family of β-sheet propellers. Furthermore, they also display some of the more subtle features of this structure, e.g., disulfide bonding, sheet-sheet stacking, lengths of loops, and residue orientation within the catalytic center.

6 Conclusion

The studies reviewed here on influenza neuraminidase complement others on hemagglutinin (see WARD 1981 for review). In both cases, gene sequencing, protein sequencing, and three-dimensional structure analyses have provided a foundation for understanding not only the mechanism of antigenic variation of influenza, but also the biological functions of these two membrane glycoproteins.

Patterns of amino acid sequence variation in the neuraminidase of different field strains and monoclonal variants have, when mapped onto the three-dimensional structure, indicated regions of the protein surface involved in eliciting an antibody response from the host. Further studies of this type are required to consolidate the picture. The rate of change in the amino acid sequence in

these regions of neuraminidase is similar to that found in hemagglutinin, suggesting that similar selection pressures are exerted on both proteins. The mechanism of antigenic drift in the two proteins is also similar, being largely a result of single base changes between parent and derivative strains, with subsequent changes to later strains most commonly, but not exclusively, accumulating at other loci.

Among subtypes of neuraminidase, the level of sequence homology is similar to that existing among subtypes of hemagglutinin. This level of homology, on the order of 50%, is high enough to enable confident prediction that the three-dimensional structure of the surface glycoproteins of other subtypes is very similar to those structures determined so far. For neuraminidase, that prediction will be tested by the study in progress of crystals of the N9 subtype.

The catalytic site of neuraminidase has been directly identified, and further studies should permit a description of the mechanism of the enzyme. The location and shape of the active site suggest that host antibodies will not be able to see it to the exclusion of other surface structures. Viable, enzymatically active mutants therefore appear to be comfortably within the repertoire of variation accessible to and required by the virus to escape neutralization by antibodies to earlier strains.

Acknowledgements. We thank G.M. AIR, L.E. BROWN, W.G. LAVER, and R.G. WEBSTER for data in advance of publication, and B.M. WOOD, R.E. THOMAS, S.M. RICHARDSON, and H.C.H. NGUYEN for assistance in preparing the manuscript.

References

Ada GL, Gottschalk A (1956) The component sugars of the influenza virus particle. Biochem J 62:686–689

Air GM (1979) Nucleotide sequence for the signal peptide and N-terminus of the hemagglutinin from an Asian (H2N2) strain of influenza virus. Virology 97:468–472

Air GM (1981) Sequence relationships among the hemagglutinin genes of 12 subtypes of influenza A virus. Proc Natl Acad Sci USA 78:7639–7643

Allen AK, Skehel JJ, Yuferof VJ (1977) The amino acid and carbohydrate composition of the neuraminidase of B/Lee/40 influenza virus. J Gen Virol 37:625–628

Allen H, McCauley J, Waterfield MD, Gething MJ (1980) Influenza virus RNA segment 7 has the coding capacity for two polypeptides. Virology 107:548–551

Amzel LM, Polzak RJ, Saul F, Vargha JM, Richards FF (1974) The three-dimensional structure of a combining region – ligand complex of immunoglobulin NEW at 3.5 Å resolution. Proc Natl Acad Sci USA 71:1427–1430

Anfinsen CB (1973) Principles that govern the folding of protein chains. Science 181:223–230

Ashwell G, Morrell AG (1974) The role of surface carbohydrates in the hepatic recognition and transport circulating glycoproteins, Adv Enzymol 41:99–128

Azad AA, Elleman TC, Laver WG, Ward CW (1983) Sequence changes associated with antigenic shift and drift in influenza virus neuraminidase. In: Laver WG (ed) The origins of pandemic influenza viruses. Elsevier, New York, pp 59–76

Bachmeyer H (1972) Effect of tryptophan modification on the activity of bacterial and viral neuraminidase. FEBS Lett 23:217–219

Baker NJ, Gandhi SS (1976) Effect of Ca^{++} on the stability of influenza virus neuraminidase. Arch Virol 52:7–18

Basak S, Compans RW (1983) Studies on the role of glycosylation in the function and antigenic properties of influenza virus glycoproteins. Virology 128:77–91

Basak S, Pritchard DG, Bhown AS, Compans RW (1981) Glycosylation sites of influenza viral glycoproteins – characterization of tryptic glycopeptides from the A/USSR (H1N1) hemagglutinin glycoprotein. J Virol 37:549–558

Battula N, Loeb L (1974) The infidelity of avian myeloblastosis virus deoxyribonucleic acid polymerase in polynucleotide replication. J Biol Chem 249:4086–4093

Bentley DR, Brownlee GG (1982) Sequence of the N2 neuraminidase from influenza virus A/NT/60/68. Nucleic Acids Res 10:5033–5042

Bergman LW, Kuehl WM (1978) Temporal relationship of translation and glycosylation of immunoglobulin heavy and light chains. Biochemistry 17:5174–5180

Bishop DHL, Huddleston JA, Brownlee GG (1982) The complete sequence of RNA segment 2 of influenza A/NT/60/68 and its encoded P1 protein. Nucleic Acids Res 10:1335–1343

Blok J, Air GM (1980) Comparative nucleotide sequences at the 3′ end of the neuraminidase gene from eleven influenza type A viruses. Virology 107:50–60

Blok J, Air GM (1982a) Block deletions in the neuraminidase genes from some influenza A viruses of the N1 subtype. Virology 118:229–234

Blok J, Air GM (1982b) Variation in the membrane – insertion and "stalk" sequences in eight subtypes of influenza type A virus neuraminidase. Biochemistry 21:4001–4007

Blok J, Air GM (1982c) Sequence variation at the 3′ end of the neuraminidase gene from 39 influenza type A viruses. Virology 121:211–229

Blok J, Air GM, Laver WG, Ward CW, Lilley GG, Woods EF, Roxburgh CM, Inglis AS (1982) Studies on the size, chemical composition and partial sequence of the neuraminidase (NA) from type A influenza virus show that the N-terminal region of the NA is not processed and serves to anchor the NA in the viral membrane. Virology 119:109–121

Blundell TL (1981) Protein-protein recognition and assembly in structural aspects of recognition and assembly in biological macromolecules. In: Balaban M (ed) International Science Services, Rehovot, pp 281–286

Breidis DJ, Lamb RA (1982) Influenza B virus genome: sequences and structural organization of RNA segment 8 and the mRNAs coding for the NS_1 and NS_2 proteins. J Virol 42:186–193

Breidis DJ, Lamb RA, Choppin PW (1982) Sequence of RNA segment 7 of the influenza B virus genome; partial amino acid homology between the membrane proteins (M_1) of the influenza A and B viruses and conservation of a second open reading frame. Virology 116:581–588

Brown LE, Ward CW, Jackson DC (1982) Antigenic determinants of influenza virus hemagglutinin. IX. The carbohydrate side chains from an Asian strain. Mol Immunol 19:329–338

Bucher DJ, Kilbourne ED (1972) A_2 (N2) neuraminidase of the X-7 influenza virus recombinant: determination of molecular size and subunit composition of the active unit. J Virol 10:60–66

Bucher DJ, Palese P (1975) The biologically active proteins of influenza virus neuraminidase. In: Kilbourne ED (ed) Influenza virus and influenza. Academic, New York, pp 83–123

Burnet FM (1948) Mucins and mucoids in relation to influenza virus action. IV. Inhibition by purified mucoid of infection and haemagglutinin with the virus strain WSE. Aust J Exp Biol Med Sci 26:381–387

Burnet FM, Stone JD (1947) The receptor-destroying enzyme of V. cholerae. Aust J Exp Biol Med Sci 25:227–233

Burnet FM, McCrea JF, Anderson SG (1947) Mucin as substrate of enzyme action by viruses of the mumps influenza group. Nature 160:404–405

Carroll SM, Paulson JC (1982) Complete metal ion requirement of influenza virus N_1 neuraminidases. Arch Virol 71:273–277

Caton AJ, Brownlee GG, Yewdell JW, Gerhard W (1982) The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). Cell 31:417–427

Choppin PW (1969) Replication on influenza virus in a continuous cell line: high yield of infective virus from cells infected at high multiplicity. Virology 38:130–134

Chothia C (1973) Conformation of twisted β-pleated sheets in proteins. J Mol Biol 75:295–302

Chothia C, Janin J (1975) Principles of protein-protein recognition. Nature 256:705–708

Chothia C, Janin J (1981) Relative orientation of close-packed β-pleated sheets in proteins (twisted β-sheets/protein secondary and tertiary structure). Proc Natl Acad Sci USA 78:4146–4150

Chothia C, Levitt M, Richardson D (1977) Structure of proteins: packing of α-helices and pleated sheets. Proc Natl Acad Sci USA 74:4130–4134

Clamp JR (1975) Structure and function of glycoproteins. In: Puttnam FW (ed) The plasma proteins. 2nd ed, vol 2. Academic New York, pp 163–211

Clement J-M (1983) On the insertion of proteins into membranes. Biochemistry 65:325–338

Cohen FE, Sternberg MJE, Taylor WR (1981) Analysis of the tertiary structure of protein β-sheet sandwiches. J Mol Biol 148:253–272

Collins JK, Knight CA (1978) Purification of the influenza hemagglutinin glycoprotein and characterization of its carbohydrate components. J Virol 26:457–467

Colman PM (1984) The structure and function of neuraminidase. Peptide Protein Rev 4:215–255

Colman PM, Laver WG (1981) The structure of influenza virus neuraminidase heads at 5 Å resolution. In: Balaban M (ed) Structural aspects of recognition and assembly in biological macromolecules. I.S.S., Rehovot, pp 869–872

Colman PM, Jansonius JN, Matthews BW (1972) The structure of thermolysin: an electron density map at 2.3 Å resolution. J Mol Biol 70:701–724

Colman PM, Deisenhofer J, Huber R, Palm W (1976) The structure of the human antibody molecule KOL: An electron density map at 5 Å resolution. J Mol Biol 100:257–282

Colman PM, Tulloch PA, Laver WG (1980) Preliminary structural studies on two influenza virus neuraminidases. In: Laver WG, Air GM (eds) Structure and variation in influenza virus. Elsevier, New York, pp 351–356

Colman PM, Gough KH, Lilley GG, Blagrove RJ, Webster RG, Laver WG (1981) Crystalline monoclonal Fab fragment with specificity towards an influenza virus neuraminidase. J Mol Biol 152:609–614

Colman PM, Varghese JN, Laver WG (1983) Structure of the catalytic and antigenic sites in influenza virus neuraminidase. Nature 303:41–44

Compans RW, Klenk H-D (1979) Viral membranes. Compr Virol 13:293–377

Compans RW, Dimmock NJ, Meier-Ewert H (1969) Effect of antibody to neuraminidase on the maturation and hemagglutinating activity of influenza A_2 virus. J Virol 4:528–534

Corfield AP, Wember M, Schauer R, Rott R (1982) The specificity of viral sialidases. The use of oligosaccharide substrates to probe enzyme characteristics and strain specific differences. Eur J Biochem 124:521–525

Corfield AP, Higa H, Paulson JC, Schauer R (1983) The specificity of viral and bacterial sialidases for α(2–3) and α(2–6) linked sialic acids in glycoproteins. Biochim Biophys Acta 744:121–126

Daniels RS, Douglas AR, Gonsalues-Scarano F, Palu G, Skehel JJ, Brown E, Knossow M, Wilson IA, Wiley DC (1983) Antigenic structure of influenza virus hemagglutinin. In: Laver WG (ed) Origin of pandemic influenza viruses. Elsevier, New York, pp 9–18

Deisenhofer J (1981) Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from Staphylococcus aureus at 2.9 and 2.8 Å resolution. Biochemistry 20:2361–2370

Deisenhofer J, Colman PM, Epp O, Huber R (1976) Crystallographic studies of a human Fc fragment. II. A complete model based on a Fourier map at 3.5 Å resolution. Hoppe-Seylers Z Physiol Chem 357:1421–1434

Desselberger U, Racaniello VR, Zazra JJ, Palese P (1980) The 3′ and 5′-terminal sequences of influenza A, B and C virus RNA segments are highly conserved and show partial inverted complementarity. Gene 8:315–328

Dopheide TAA, Ward CW (1981) The location of the bromelain-cleavage site in a Hong Kong influenza virus hemagglutinin. J Gen Virol 52:367–370

Drzenick R, Frank H, Rott R (1968) Electron microscopy of purified influenza virus neuraminidase. Virology 36:703–707

Edmond JD, Johnston RG, Kidd D, Rylance HJ, Sommerville RG (1966) Inhibition of neuraminidase and antiviral action. Br J Pharmacol Chemother 27:415–426

Elleman TC, Azad AA, Ward CW (1982) Neuraminidase gene from the early Asian strain of human influenza virus A/RI/5^-/57 (H2N2). Nucleic Acids Res 10:7005–7015

Fazekas de St. Groth S (1978) Antigenic, adaptive and adsorptive variants of the influenza A hemagglutinin. In: Laver WG, Bachmayer H, Weil R (eds) The influenza virus hemagglutinin. Springer, Vienna New York, pp 25–48

Fields S, Winter G (1981) Nucleotide sequence heterogeneity and sequence rearrangement in influenza virus cDNA. Gene 15:207–214

Fields S, Winter G (1982) Nucleotide sequences of influenza virus segments 1 and 3 reveal mosaic structure of a small viral RNA segment. Cell 28:303–313

Fields S, Winter G, Brownlee GG (1981) Structure of the neuraminidase in human influenza virus A/PR/8/34. Nature 290:213–217

Francis T, Moore AE (1940) A study of the neurotropic tendency in strains of the virus of epidemic influenza. J Exp Med 72:717–728

French S, Robson B (1983) What is a conservation substitution? J Mol Evol 19:171–175

Galloway G, Leung AYT, Hunneyball IM, Stanworth DR (1983) The successful use of asialylated IgG as an immunogen and arthritogen in the rabbit. Immunology 49:511–518

Garnier J, Osguthorpe DJ, Robson B (1978) Analysis of the accuracy and implication of simple methods for predicting the secondary structure of globular proteins. J Mol Biol 120:97–120

Gopinathan KP, Weymouth LA, Kunkel TA, Loeb LA (1979) Mutagenesis in vitro by DNA polymerase from an RNA tumour virus. Nature 278:857–859

Gottschalk A (1957) Neuraminidase: the specific enzyme of influenza virus and vibrio cholerae. Biochim Biophys Acta 23:645–646

Gottschalk A (1958) Neuraminidase: its substrate and mode of action. Adv Enzymol 20:135–145

Gottschalk A (1972) Historical introduction. In: Gottschalk A (ed) Glycoproteins. Their composition, structure and function. Elsevier, Amsterdam, pp 2–23

Green N, Alexander H, Olson A, Alexander S, Shinnick TM, Sutcliffe JG, Lerner RA (1982) Immunogenic structure of the influenza virus hemagglutinin. Cell 28:477–487

Gregoriades A (1972) Isolation of neuraminidase from the WSN strain of influenza virus. Virology 49:333–336

Griffin JA, Compans RW (1979) Effect of cytochalasin B on the maturation of enveloped viruses. J Exp Med 150:379–391

Griffin JA, Basak S, Compans RW (1983) Effects of hexose starvation and the role of sialic acid in influenza virus release. Virology 125:324–334

Griffith IP (1975) The fine structure of influenza virus. In: Mahy BWJ, Barry RD (eds) Negative strand viruses, vol 1. Academic, London, pp 121–132

Harrison SC, Olson AJ, Schutt CE, Winkler FK, Bricogne G (1978) Tomato bushy stunt virus at 2.9 Å resolution. Nature 276:368–373

Haslam EA, Hampson AW, Radiskevics I, White DO (1970) The polypeptides of influenza virus. III. Identification of the hemagglutinin, neuraminidase and nucleocapsid proteins. Virology 42:566–575

Haugen DA, Armes LG, Yasunobu KT, Coon MJ (1977) Amino-terminal sequence of phenobarbital-inducible cytochrome P-450 from rabbit liver microsomes: similarity to hydrophobic amino-terminal segments of pre-proteins. Biochem Biophys Res Comm 77:967–973

Hirsch RL (1982) The complement system: its importance in the host response to viral infection. Microbiol Rev 46:71–85

Hirsch RL, Winkelstein JA, Griffin DE (1980) The role of complement in viral infections. III. Activation of the classical and alternate pathways by Sindbis virus. J Immunol 124:2507–2510

Hirsch RL, Griffin DE, Winkelstein JA (1981) Host modification of Sindbis virus sialic acid content influences alternative complement pathway activation and virus clearance. J Immunol 127:1740–1743

Hirst GK (1942) Adsorption of influenza hemagglutinins and virus by red blood cells. J Exp Med 76:195–209

Hiti AL, Nayak DP (1982) Complete nucleotide sequence of the neuraminidase gene of human influenza virus A/WSN/33. J Virol 41:730–734

Hopp TP, Woods KR (1981) Prediction of protein antigenic determinants from amino acid sequences. Proc Natl Acad Sci USA 78:3824–3828

Hoyle L (1952) Structure of the influenza virus. The relation between biological activity and chemical structure of virus fractions. J Hyg (Camb) 50:229–245

Huang RTC, Rott R, Wahn K, Klenk H-D, Kohama T (1980) The function of the neuraminidase in membrane fusion induced by myxoviruses. Virology 107:313–319

Hubbard SC, Ivatt RJ (1981) Synthesis and processing of asparagine-linked oligosaccharides[1,2]. Annu Rev Biochem 50:555–583

Inglis SC, Barrett T, Brown CM, Almond JW (1979) The smallest genome RNA segment of influenza virus contains two genes that may overlap. Proc Natl Acad Sci USA 76:3790–3794

Jackson DC, Webster RG (1982) A topographic map of the enzyme active center and antigenic sites on the neuraminidase of influenza virus A/Tokyo/3/67 (H2N2). Virology 123:69–77

Jackson DC, Dopheide TAA, Russell RJ, White DO, Ward CW (1979) Antigenic determinants of influenza virus hemagglutinin. II. Antigenic reactivity of the isolated N-terminal cyanogen bromide peptide of A/Memphis/72 hemagglutinin heavy chain. Virology 93:458–465

Jancik JM, Schauer R, Streicher H-J (1975) Influence of membrane-bound N-acetylneuraminic acid on the survival of erythrocytes in man. Z Physiol Chem 356:1329–1331

Jennings PA, Finch JT, Winter G, Robertson JS (1983) Does the higher order structure of the influenza virus ribonucleoprotein guide sequence rearrangements in influenza viral RNA? Cell 34:619–627

Kaptein JS, Nayak DP (1982) Complete nucleotide sequence of the polymerase 3 gene of human influenza virus A/WSN/33. J Virol 42:55–63

Kazatchkine MD, Fearon DT, Austen KF (1979) Human alternative complement pathway: Membrane associated sialic acid regulates the competition between β and β1H for cell bound C3b. J Immunol 122:75–81

Keil W, Klenk H-D, Schwarz RT (1979) Carbohydrates of influenza virus. III. Nature of oligosaccharide-protein linkage in viral glycoprotein. J Virol 31:253–256

Kendal AP, Eckert EA (1972) The preparation and properties of ^{14}C-carboxamido-methylated subunits from A_2/1957 influenza neuraminidase. Biochim Biophys Acta 258:484–495

Kendal AP, Kiley MP (1975) Structural comparisons of influenza A neuraminidase. In: Mahy BWJ, Barry RD (eds) Negative strand viruses, vol 1. Academic Press, London. pp 145–159

Kendal AP, Madeley CR (1970) Flocculation of influenza virus by specific anti-neuraminidase antibody. Arch Gesamte Virusforsch 31:219–229

Klenk E, Faillard H, Lempfrid H (195) Über die enzymatische Wirkung von Influenzavirus. Z Physiol Chem 301:235–246

Klenk H-D, Rott R (1980) Cotranslational and posttranslational processing of viral glycoproteins. Curr Top Microbiol Immunol 90:19–48

Klenk H-D, Scholtissek C, Rott R (1972) Inhibition of glycoprotein biosynthesis of influenza virus by D-glucosamine and 2-deoxy-D-glucose. Virology 49:723–734

Klenk H-D, Rott, R, Orlich M, Blodorn J (1975) Activation of influenza A viruses by trypsin treatment. Virology 68:426–439

Kolb-Bachofen V, Kolb H (1979) Autoimmune reactions against liver cells by syngeneic neuraminidase-treated lymphocytes. J Immunol 123:2830–2834

Kozak M (1981) Mechanism of mRNA recognition by eubaryotic ribosomes during initiation of protein synthesis. Curr Top Microbiol Immunol 93:81–123

Kreisel W, Volk BA, Buechsel R, Reutter W (1980) Different half-lives of the carbohydrate and protein moieties of a 110000-dalton glycoprotein isolated from the plasma membranes of rat liver. Proc Natl Acad Sci USA 77:1828–1831

Lai C-J, Markoff LJ, Sveda M, Dhar R, Chanock RM (1980) DNA sequences derived from genomic and mRNA species that code for the haemagglutinin and the neuraminidase of influenza A virus. In: Laver G, Air GM (eds) Structure and variation in influenza virus. Elsevier, New York, pp 115–124

Lamb RA, Choppin PW (1979) Segment 8 of the influenza virus genome is unique in coding for 2 polypeptides. Proc Natl Acad Sci USA 76:4908–4912

Lamb RA, Choppin PW (1981) Identification of a second protein (M_2) encoded by RNA segment 7 of influenza virus. Virology 112:729–737

Lamb RA, Lai CJ (1980) Sequence of interrupted and uninterrupted mRNAs and cloned DNA coding for the two overlapping nonstructural proteins of influenza virus. Cell 21:475–485

Lamb RA, Lai CJ (1981) Conservation of the influenza virus membrane protein (M_1) amino acid sequence and an open reading frame of RNA segment 7 encoding a second protein (M_2) in H1H1 and H3N2 strains. Virology 112:746–751

Lamb RA, Shaw MW, Breidis DJ, Choppin PW (1983) The nucleotide sequence of the neuraminidase gene of influenza B virus reveals two overlapping reading frames. In: Laver WG (ed) The origin of pandemic influenza viruses. Elsevier, New York, pp 77–86

Lambré CR, Kazatchkine MD, Maillet F, Thibon M (1982) Guinea pig erythrocytes after their

contact with influenza virus acquire the ability to activate the alternative complement pathway through virus induced desialation of the cells. J Immunol 128:629–634
Lambré CR, Thibon M, Le Maho S, Di Bella G (1983) Auto antibody dependent activation of the autologous classical complement pathway in guinea pig red cells treated with influenza virus or neuraminidase: in vitro and in vivo study. Virology 49:311–319
Laver WG (1963) The structure of influenza viruses. 3. Disruption of the virus particle and separation of neuraminidase activity. Virology 20:251–262
Laver WG (1973) The polypeptides of influenza viruses. Adv Virus Res 18:57–103
Laver WG (1978) Crystallization and peptide maps of neuraminidase "heads" from H2N2 and H3N2 influenza virus strains. Virology 86:78–87
Laver WG, Baker N (1972) Amino acid composition of polypeptides from influenza virus particles. J Gen Virol 17:61–68
Laver WG, Kilbourne ED (1966) Identification in a recombinant influenza virus of structural proteins derived from both parents. Virology 30:493–501
Laver WG, Valentine RC (1969) Morphology of the isolated hemagglutinin and neuraminidase subunits of influenza virus. Virology 38:105–119
Laver WG, Webster RG (1973) Studies on the origin of pandemic influenza. III. Evidence implicating duck and equine influenza viruses as possible progenitors of the Hong Kong strain of human influenza. Virology 51:383–391
Laver WG, Air GM, Webster RG, Gerhard W, Ward CW, Dopheide TAA (1979) Antigenic drift in type A influenza virus: sequence differences in the hemagglutinin of Hong Kong (H3N2) variants selected with monoclonal hybridoma antibodies. Virology 98:226–237
Laver WG, Air GM, Webster RG, Markoff LJ (1982) Amino acid sequence changes in antigenic variants of type A influenza virus N2 neuraminidase. Virology 122:450–460
Laver WG, Colman PM, Webster RG, Hinshaw VS, Air GM (1984) Influenza virus neuraminidase with haemagglutinin activity. Virology
Lazarowitz SG, Choppin PW (1975) Enhancement of infectivity of influenza A and B viruses by proteolytic cleavage of the hemagglutinin polypeptide. Virology 68:440–454
Lazarowitz SG, Compans RW, Choppin PW (1971) Influenza virus structural and nonstructural proteins in infected cells and their plasma membranes. Virology 46:830–843
Lazdins I, Haslam EA, White DO (1972) The polypeptides of influenza virus. VI. Composition of the neuraminidase. Virology 49:758–765
Lentz MR, Air GM, Laver WG, Webster RG (1984) Sequence of the neuraminidase gene from influenza virus A/Tokyo/3/67 and previously uncharacterised monoclonal variants. Virology 135:257–265
Levitt M, Chothia C (1976) Structural patterns in globular proteins. Nature 261:552–558
Maillet F, Kazatchkine MD (1983) Modulation of the formation of the human amplification C3 convertase of complement by polycations. Immunology 50:27–33
Mariuzza RA, Jankovic DL, Boulot G, Amit AG, Saludjian P, Le Guern A, Mazie JC, Poljak RJ (1983) Preliminary crystallographic study of the complex between the Fab fragment of a monoclonal antilysozyme antibody and its antigen. J Mol Biol 170:1055–1088
Markoff L, Lai CJ (1982) Sequence of the influenza A/Udorn/72 (H3N2) virus neuraminidase gene as determined from cloned full-length DNA. Virology 119:288–297
Marquart M, Deisenhofer J (1982) The three-dimensional structure of antibodies. Immunol Today 3:160–166
Marquart M, Deisenhofer J, Huber R, Palm W (1980) Crystallographic refinement and atomic models of the intact immunoglobulin molecule Kol and its antigen binding fragment at 3.0 and 2.9 Å resolution. J Mol Biol 141:369–392
Marshall RD (1972) Glycoproteins. Annu Rev Biochem 41:673–702
Martinez C, Del Rio L, Portelo A, Domingo E, Ortin J (1983) Evolution of the influenza virus neuraminidase gene during drift of the N2 subtype. Virology 130:539–545
Matlin K, Reggio H, Helenius A, Simons KJ (1981) Infectious entry pathway of influenza virus in a canine kidney cell line. J Cell Biol 91:601–631
Matsushima M, Marquart M, Jones TA, Colman PM, Bartels K, Huber R (1978) Crystal structure of the human Fab fragment Kol and its comparison with the intact Kol molecule. J Mol Biol 121:441–459
McCauley JW, Bye J, Elder K, Gething MJ, Skehel JJ, Smith A, Waterfield MD (1979) Influenza virus hemagglutinin signal sequence. FEBS Lett 108:422–428

McCauley JW, Mahy BWJ, Inglis SC (1982) Nucleotide sequence of fowl plague virus RNA segment 7. J Gen Virol 58:211–215

McSharry JJ, Pickering RJ, Caliguiri LA (1981) Activation of the alternate complement pathway by enveloped viruses containing limited amounts of sialic acid. Virology 114:507–515

Michel H (1983) Crystallisation of membrane proteins. Trends in Biomedical Sciences 8:56–59

Mountford CE, Grossman G, Holmes KT, O'Sullivan WJ, Hampson AW, Raison RL, Webster RG (1982) Effect of monoclonal antineuraminidase antibodies on the kinetic behaviour of influenza virus neuraminidase. Mol Immunol 19:811–816

Nakajima K, Desselberger U, Palese P (1978) Recent human influenza A (H1N1) viruses are closely related genetically to strains isolated in 1950. Nature 274:334–339

Nakajima S, Sugiura A (1980) Neuroviralence of influenza in mice. II. Mechanism of virulence as studied in a neuroblastoma cell line. Virology 101:450–457

Nakamura K, Compans RW (1978) Glycopeptide components of influenza viral glycoproteins. Virology 86:432–442

Noll H, Aoyagi T, Orlando J (1962) The structural relationship of sialidase to the influenza virus surface. Virology 18:154–157

Palese P (1977) The genes of influenza virus. Cell 10:1–10

Palese P, Compans RW (1976) Inhibition of influenza virus replication in tissue culture by 2-deoxy-2,3-dehydro-N-trifluoro-acetyl-neuraminic acid (FANA): mechanism of action. J Gen Virol 33:159–163

Palese P, Tobita K, Ueda M, Compans RW (1974) Characterization of temperature sensitive influenza virus mutants defective in neuraminidase. Virology 61:397–410

Paulson JC, Weinstein J, Dorland L, Van Halbeek H, Vliegenthart JFG (1982) Newcastle disease virus contains a linkage-specific glycoprotein sialidase. J Biol Chem 257:12734–12738

Poljak RJ, Amzel LM, Avey HP, Chen BL, Phizackerly RP, Saul F (1973) Three dimensional structure of the Fab[1] fragment of a human immunoglobulin at 2.8 Å resolution. Proc Natl Acad Sci USA 70:3305–3310

Porter AG, Barber C, Carey NH, Hallewell RA, Threlfall G, Emtage JS (1979) Complete nucleotide sequence of an influenza virus hemagglutinin gene from cloned DNA. Nature 282:471–477

Richardson JS (1977) β-sheet topology and the relatedness of proteins. Nature 268:495–500

Richardson JS (1981) The anatomy and taxonomy of protein structure. Adv Prot Chem 34:167–339

Richardson JS, Getzoff ED, Richardson DC (1978) The β-bulge: A common small unit of nonrepetitive protein structure. Proc Natl Acad Sci USA 75:2574–2578

Robertson JS (1979) 5′ and 3′ terminal nucleotide sequences of the RNA genome segments of influenza virus. Nucleic Acids Res 6:3745–3757

Rogers GN, Paulson JC (1983) Receptor determinants of human and animal influenza virus isolates: difference in receptor specificity of the H3 haemagglutinin based on species of origin. Virology 127:361–373

Rogers GN, Paulson JC, Daniels RS, Skehel JJ, Wilson IA, Wiley DC (1983) Single amino acid substitutions in influenza hemagglutinin change receptor binding specificity. Nature 304:76–78

Rosenberg A, Schengrund C-L (1976) Sialidases. In: Rosenburg A, Schengrund C-L (eds) The biological roles of sialic acid Plenum, New York, pp 295–359

Rothman JE, Lodish HF (1977) Synchronised transmembrane insertion and glycosylation of a nascent membrane protein. Nature 269:775–780

Rott R, Becht H, Klenk H-D, Scholtissek C (1972) Interactions of concanavalin A with the membrane of influenza virus infected cells and with envelope components of the virus particle. Z Naturforsch 27b:227–233

Saber MS, Drzenick R, Krüpe M (1965) Freilegung von ABH-Blutgruppen determinierenden Kohlenhydraten an Erythrozyten durch Neuraminidase-Einwirkung. Z Naturforsch 206:965–973

Schauer R (1982) Sialic acids. Adv Carbohydr Chem Biochem 40:131–234

Schmidt WE, Jung H-D, Palm W, Hilschman N (1983) The primary structure of the crystallisable monoclonal immunoglobulin IgGl, KOL. Hoppe Seylers Z Physiol Chemie 364:713–747

Scholtissek C (1978) The genome of the influenza virus. Curr Top Microbiol Immunol 80:139–169

Scholtissek C, Rohde W, von Hoyningen V, Rott R (1978a) On the origin of the human influenza virus subtypes H2N2 and H3N2. Virology 87:13–20

Scholtissek C, von Hoyningen V, Rott R (1978b) Genetic relatedness between the new 1977 epidemic strains (H1N1) of influenza and human influenza and human influenza strains isolated between 1947 and 1957 (H1N1). Virology 89:613–617

Schulman JL (1975) Immunology of influenza. In: Kilbourne ED (ed) The influenza viruses and influenza. Academic, New York, pp 373–393

Schulman JL, Palese P (1977) Virulence factors of influenza A viruses: WSN virus neuraminidase required for plaque production in MDBK cells. J Virol 24:170–176

Schwarz RT, Klenk H-D (1981) Carbohydrates of influenza virus. IV. Strain dependent variations. Virology 113:584–593

Schwarz RT, Schmidt MFG, Anwer U, Klenk H-D (1977) Carbohydrates of influenza virus. I. Glycopeptides derived from viral glycoproteins after labelling with radioactive sugars. J Virol 23:217–226

Segal DM, Paplan EA, Cohen GH, Rudikoff S, Potter M, Davies DR (1974) The three-dimensional structure of a phosphorylcholine-binding mouse immunoglobulin Fab and the nature of the antigen binding site. Proc Natl Acad Sci USA 71:4298–4302

Seto JT, Rott R (1966) Functional significance of sialidase during influenza virus multiplication. Virology 30:731–737

Seto JT, Drzeniek R, Rott R (1966) Isolation of a low molecular weight neuraminidase from influenza virus. Biochim Biophys Acta 113:402–404

Shaw MW, Lamb RA, Erickson BW, Breidis DJ, Choppin PW (1982) Complete nucleotide sequence of the neuraminidase gene of influenza B virus. Proc Natl Acad Sci USA 79:6817–6821

Skehel JJ (1972) Polypeptide synthesis in influenza virus-infected cells. Virology 49:23–36

Skehel JJ, Schild GC (1971) The polypeptide composition of influenza A viruses. Virology 44:396–408

Skehel JJ, Waterfield MD (1975) Studies on the primary structure of the influenza virus hemagglutinin. Proc Natl Acad Sci USA 72:93–97

Smith W, Andrewes CH, Laidlaw PP (1933) A virus obtained from influenza patients. Lancet 2:66–68

Stanley PM, Crook NE, Streader LG, Davidson BE (1973) The polypeptides of influenza virus. VIII. Large-scale purification of hemagglutinin. Virology 56:640–645

Struck DK, Lennarz WJ (1980) The function of saccharide lipids in synthesis of glycoproteins. In: Lennarz W (ed) The biochemistry of glycoproteins and proteoglycans. Plenum, New York, pp 35–83

Sugiuna A, Ueda M (1980) Neurovirulence of influenza in mice. I. Neurovirulence of recombinants between virulent and avirulent viral strains. Virology 101:440–449

Tettamanti G, Cestaro B, Venerendo B, Preti A (1978) Neuraminidase gangliosides interactions. In: Gatt S, Freysz L, Mandel P (eds) Enzymes of lipid metabolism. Plenum, New York, pp 417–437

Ulmanen I, Broni BA, Krug RM (1981) Role of two of the influenza virus core P proteins in recognizing cap 1 structure (m^7 Gppp Nm) on RNAs. Proc Natl Acad Sci USA 78:7355–7359

Van Rompuy, Min-Ion W, Huylebroeck D, Fiers W (1982) Complete nucleotide sequence of a human influenza neuraminidase gene of subtype N2 (A/Vic/3/75). J Mol Biol 161:1–11

Varghese JN, Laver WG, Colman PM (1983) Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å resolution. Nature 303:35–40

Wagh PV, Bahl OP (1981) Sugar residues on proteins. Crit Rev Biochem 10:307–377

Ward CW (1981) Structure of the influenza virus hemagglutinin. Curr Top Microbiol Immunol 94/95:1–74

Ward CW, Dopheide TAA (1976) Size and chemical composition of influenza virus hemagglutinin chains. FEBS Lett 65:365–368

Ward CW, Dopheide TAA (1981a) Amino acid sequence and oligosaccharide distribution of the hemagglutinin from an early Hong Kong variant A/Aichi/2/68(x-31). Biochem J 193:953–962

Ward CW, Dopheide TAA (1981b) Evolution of the Hong Kong influenza A sub-type. Structural relationships between the hemagglutinins from A/duck/Ukraine/1/63 (Hav7) and the Hong Kong (H3) hemagglutinins. Biochem J 195:337–340

Ward CW, Gleeson PA, Dopheide TAA (1980) Carbohydrate composition of the oligosaccharide units on the hemagglutinin from the Hong Kong influenza variant A/Memphis/102/72. Biochem J 189:649–652

Ward CW, Elleman TC, Azad AA (1982) Amino acid sequence of the pronase-released heads of neuraminidase subtype N2 from the Asian strain A/Tokyo/3/67 of influenza virus. Biochem J 207:91–95

Ward CW, Colman PM, Laver WG (1983a) The disulphide bonds of an Asian influenza virus neuraminidase. FEBS Lett 153:29–33

Ward CW, Murray JM, Roxburgh CM, Jackson DC (1983b) Chemical and antigenic characterization of the carbohydrate side chains of an Asian (N2) influenza virus neuraminidase. Virology 126:370–375

Waterfield MD, Espelie K, Elder K, Skehel JJ (1979) Structure of the hemagglutinin of influenza virus. Br Med Bull 35:57–63

Waterfield MD, Gething MJ, Scrace G, Skehel JJ (1980) The carbohydrate side chains and disulphide bonds of the hemagglutinin of the influenza virus A/Japan/305/57(H2N1). In: Laver G, Air G (eds) Structure and variation in influenza virus. ElsevierNorth Holland, New York, pp 11–20

Weber E, Papamokos E, Bode W, Huber R, Kato I, Laskowski M (1981) Crystallization, crystal structure analysis and molecular model of the third domain of Japanese quail ovomucoid, a kazal type inhibitor. J Mol Biol 149:109–123

Webster RG (1970) Estimation of the molecular weights of the polypeptide chains from the isolated hemagglutinin and neuraminidase subunits of influenza viruses. Virology 40:543–554

Webster RG, Darlington RW (1969) Disruption of myxoviruses with Tween 20 and isolation of biologically active hemagglutinin and neuraminidase subunits. J Virol 4:182–187

Webster RG, Hinshaw VS, Laver WG (1982) Selection and analysis of antigenic variants of the neuraminidase of N2 influenza viruses with monoclonal antibodies. Virology 117:93–104

Webster RG, Brown LE, Laver WG (1984) Antigenic and biological characterisation of influenza virus neuraminidase (N2) with monoclonal antibodies. Virology 135:30–42

Wedgwood RJ, Ginsberg HS, Pillemer L (1956) The properdin system and immunity. VI. The inactivation of Newcastle disease virus by the properdin system. J Exp Med 104:707–725

Weiler JM, Yurt RM, Fearon DT, Austen KF (1978) Modulation of the formation of the amplification convertase of complement, C3b, Bb, by native and commercial heparin. J Exp Med 147:409–421

White DO (1974) Influenza viral proteins: identification and synthesis. Curr Top Microbiol Immunol 63:1–48

White J, Helenius A (1980) pH-dependent fusion between the Semliki forest virus membrane and liposomes. Proc Natl Acad Sci USA 77:3273–3277

White J, Helenius A, Gething M-J (1982a) Haemagglutinin of influenza virus expressed from a cloned gene promotes membrane fusion. Nature 300:658–659

White J, Kartenbeck J, Helenius A (1982b) Membrane fusion activity of influenza virus. EMBO J 1:217–222

Wiley DC, Wilso JA, Skehel JJ (1981) Structural identification of the antibody binding sites of the Hong Kong influenza haemagglutinin and their involvement in antigenic variation. Nature 298:373–378

Wilson IA, Skehel JJ, Wiley DC (1981) Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. Nature 289:366–373

Winter G, Fields S (1980) Cloning of influenza cDNA into M13: the sequence of the RNA segment encoding the A/PR/8/34 matrix protein. Nucleic Acids Res 8:1965–1974

Winter G, Fields S (1982) Nucleotide sequence of human influenza A/PR/8/34 segment 2. Nucleic Acids Res 10:2135–2143

Winter G, Fields S, Gait MJ (1981) The use of synthetic oligodeoxynucleotide primers in cloning and sequencing segment 8 of influenza virus (A/PR/8/34). Nucleic Acids Res 9:237–245

Woodruff JJ, Gesner BM (1969) The effect of neuraminidase on the fate of transfused lymphocytes. J Exp Med 129:551–567

Wright CE, Laver WG (1978) Preliminary crystallographic data for influenza virus neuraminidase "heads". J Mol Biol 120:133–136

Wrigley NG (1979) Electron microscopy of influenza virus. Br Med Bull 35:35–38

Wrigley NG, Laver WG, Downie JC (1977) Binding of antibodies to isolated hemagglutinin and neuraminidase molecules of influenza virus observed in the electron microscope. J Mol Biol 109:405–421

Wrigley NG, Skehel JJ, Charlwood PA, Brand CM (1973) The size and shape of influenza virus neuraminidase. Virology 51:525–529