Zeitschrift: Archives des sciences et compte rendu des séances de la Société
Herausgeber: Société de Physique et d'Histoire Naturelle de Genève
Band: 42 (1989)
Heft: 1: Archives des Sciences

Artikel: Polarity : from dipoles to biopolarization
Autor: Turian, Gilbert
Kapitel: III: Macromolecular polarities
DOI: https://doi.org/10.5169/seals-740080

Nutzungsbedingungen
Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation
L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use
The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 22.11.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

III. MACROMOLECULAR POLARITIES

A. FREE MACROMOLECULES

1. NUCLEIC ACIDS

a) *Deoxyribonucleic acid (DNA)*

a[1] *Structure*

According to the celebrated model of Watson-Crick (1953), both single strands of the antiparallel DNA helix have inverse polarity (5'→3' and 3'→5', Fig. 5).

The two major A and B conformational DNAs are known to depend on the salt content of the fiber while the C conformation is a low salt form of sodium DNA. The recently studied H-DNA is a structure containing both triple-stranded and single-

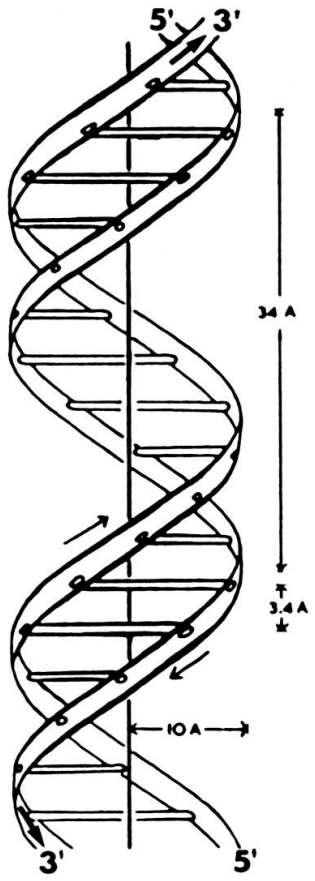

FIG. 5.

Inverse polarity of the two strands (reading direction, 5'-3' arrows) of the double DNA helix. The hydrogen-bonded purines and pyrimidines of the ladder's bars are also poles of a complementarity. After Watson and Crick, 1953.

stranded regions which have been modelled on the basis of reactivity and topological arguments (Htun and Dahlberg, 1988). A prediction of this model is that DNA should introduce a severe kink in DNA molecules which could serve to bring DNA binding proteins into proximity with each other. A conversion from H to B form was found to occur during electrophoresis at progressively higher pHs.

There are three major families of A-, B- and Z-DNA duplexes (Dickerson *et al.*, 1982), which have a common feature, the antiparallel disposition of their constituent strands (Saenger, 1984). Normal Watson-Crick base-pairing is incompatible with parallel stranded duplexes. Recently, however, alternative conformations of DNA molecules have been produced in which hemiprotonation and protonation stabilizing bindings have rended feasible the existence of parallel stranded double helices (Van de Sande *et al.*, 1988). Two of the four hairpin deoxyribonucleotide molecules synthesized "contain either a 3' -P-3' or 5' -P-5' internucleotide linkage in the loop, so that the strands in the stem have the same, that is, parallel polarity". This reversal of polarity necessitated some steps of chemical synthesis in the 5' to 3' direction, that is, opposite to the conventional 3'-5' polarity (Van de Sande *et al.*, 1988).

The classic B-DNA double helix (Watson and Crick, 1953) in its lattice configuration (stable at high relative humidity) consists of two *right*-handed helical polynucleotide chains of opposite polarity — the internucleotide linkage in one strand is 3'→5' and in the other 5'→3'. The pairing rules require that the bases in the two chains are complementary and the nucleoside pairs in the interior of the helix have the planes of their aromatic rings at right angles to the helix axis, and their apolar bulk stacked on top of one another (Drew *et al.*, 1988). The highly negatively charged phosphodiester backbone conversely faces outward, and its strongly polar groups can thus interact with the components of the aqueous environment (Mahler and Cordes, 1969). Oppositely, the helical Z-DNA, discovered by Rich and his colleagues (Wang *et al.*, 1981), characterized by a back-bone which zigzags down the molecule, has a *left*-handed conformation. Chiral metal complexes have provided very sensitive probes for both these left- and right-handed helical structures (Barton, 1986). The DNA double helix is not a regular, featureless barberpole molecule (Privé *et al.*, 1987). Different base sequences "have their own special signature, in the way that they influence groove width, helical twist, bending, and mechanical rigidity or resistance to bending. These special features probably help other molecules such as repressors to read and recognize one base sequence in preference to another".

Polarized unique sequences between inverted repeats have been determined by restriction fragment analysis of chloroplast DNA molecules. A large inverted repeat segment, characteristic of the circular chloroplast genome, has been considered as responsible for a number of its physical properties, among which formation of head-to-head dimers and copy-correction between the inverted repeats. However, "one property which an inverted repeat might be expected to confer — reversal of polarity

of the single copy sequences located between the repeats — has not yet been demonstrated for the chloroplast genome" (Palmer, 1983).

For its central functions of storage and retrieval, DNA can be quite adequately approximated as a one-dimensional line of nucleotides. However, this approximation does not suffice to explain transcriptional regulation. In some regulatory systems, the DNA loops out of one dimension to bring two specific well-separated DNA sites into close proximity (Ptashne, 1986).

Palindromic sequences, long or short, have been considered as being looped out of the linear double-helical structure of DNA. The hairpin(s) or stem- and loop structures could serve as special recognition sites on DNA. In testing ideas about the looping-out process, a fully palindromic circular DNA with head to-head dimer was constructed; linearization (endonuclease) — annealation — ligation then gave a polarity mixture of head-to-tail, head-to-head, and tail-to-tail oligomers (Gellert *et al.*, 1978). DNA gyrase catalyzes a supercoiling reaction which has an intrinsic directionality. In a mechanical model, it left the DNA with a net deficit of helical turns or negative supercoiling (Gellert *et al.*, 1978). In an alternate version of the model, the DNA is translocated without rotation over the enzyme bound to a second DNA segment. The enzyme-DNA complex is rotated by the translocation and "if the polarity of translocation is such that the nicking-closing function sits in the over-wound loop and relaxes those turns selectively, a negatively supercoiled molecule will once again result" (Gellert *et al.*, 1978). Another model was proposed (Liu and Wang, 1978) in which the polarity of translocation is determined by a superhelical wrapping of DNA on the enzyme. However, Gellert *et al.* (1978) preferred to "consider the polarity as arising from the arrangement of the enzymatic sites controlling the rolling motion".

The formation of positively and negatively supercoiled DNA loops by tracking is an example of distant action along a DNA. When both positively and negatively supercoiled loops are generated at the same time by a tracking process, DNA topoisomerase I and DNA gyrase may act differentially in the loop (Liu and Wang, 1987; Wang and Giaever, 1988). Moreover "because of the dyad symmetry of the DNA double helix, the directionality of tracking is usually determined by the asymmetry of the polymerase protein" (Wang and Giaever, 1988).

There is a possibility that solitons (solitary waves) play a role in the properties of DNA and provide a mechanism for transferring energy in that macromolecule (Maddox, 1987). As localized non linear waves propagating and interacting-like particules, the solitons simultaneously create both the environment necessary to support excitation energy or charge, and the conditions for its translation. This proposal is reminiscent of the solvated electron concept in which the electron couples with adjacent water molecules to form a comparatively stable "polaron". The effectively one-dimensional nature of DNA permits a translational wave to propagate, so transferring the energy either as an exciton (coupled electron-hole pair)

or as a charge (polaron). Energy localization, and hence damage, "may result from the collapse of the soliton at some discontinuity in the structure of the DNA molecule, such as a B/Z transition, an intercalated or otherwise bound molecule (particularly with energy or charge accepting character) or as the result of the coincidence of two passing solitons" (Baverstock and Cundall, 1988).

a^2 *Replication*

The macromolecule of DNA can be considered as a duplex of two inversely polarized strands distinctive by their 5' or 3' strands. DNA synthesis is highly asymmetric. This asymmetry is inherent in the opposite-running polarity of the complementary DNA strands: the two complementary DNA strands run in opposite directions (3' to 5' on one strand, 5' to 3' on the other), and the two branches of the DNA replication fork are not identical; one strand, called the lagging strand, makes short fragments of new DNA that subsequently are joined into a continuous strand, whereas the other strand is elongated from the outset as a continuous piece. Replication proceeds polarly along both template strands in the 3'- to -5' direction and, therefore, the process is discontinuous for one of the strands.

The circular bacterial chromosome is replicated semiconservatively and *bidirectionally* starting from a single origin point. DNA polymerase I of *E. coli* makes a strand of DNA complementary to the template strand and the polarity of the new strand (chain growth in the 5'- to -3' direction) is opposite to that of the template strand. The leading strand of DNA is replicated continuously in that 5'- to -3' direction whereas the lagging strand is replicated in short pieces called Okazaki fragments made in the direction opposite to that of the movement of the replication fork (see Lehninger, 1982).

Polarity of DNA replication has recently been shown not to change as a function of transcriptional activity (James and Leffak, 1986). This finding is consistent with earlier suggestions that DNA replication in the transcriptional direction may be a necessary but not sufficient condition for gene expression. Replication proceeds from upstream (5') origins through the alpha-globin locus irrespective of transcriptional activity in erythrocytes (James and Leffak, 1986). Active genes are replicated in the transcriptional direction and cells may switch the replication polarity of genes in concert with gene activation or repression. The replication in the early synthetic (S) phase of genes whose transcription is activated by *trans*-acting factors and/or chromosome translocation suggests that protein binding and chromosome position may also influence replication polarity (Calza *et al.*, 1984).

a^3 *Transcription*

Transcription begins when an enzyme called RNA polymerase binds to a specific base sequence on DNA, the promotor. The RNA molecule always polarly grows in

the 5'- to -3' direction. The polymerase continues along the DNA chain until its passes another special base sequence, the termination signal (Darnell *et al.*, 1986).

In prokaryotes, the degree of supercoiling of DNA can profoundly influence the use of specific promoters. DNA topology and supercoiling can be important parameters when considering DNA wrapping around histones into a left-handed toroidal coil as topologically equivalent to DNA writhe having a negative sense (Drlica and Rouvière-Yaniv, 1987). Supercoiling or topoisomerase activity or both can affect numerous chromosomal processes, such as replication, transcription, and recombination (Bjornsti *et al.*, 1986). Thus, DNA in bacterial cells would be under negative superhelical tension arisen from the action of a topoisomerase (Drlica and Rouvière-Yaniv, 1987). In eukaryotes, a variety of indirect observations suggest that DNA topology has a similar importance in proper gene expression. Much attention has therefore been focused on the cellular proteins that control DNA supercoiling, among which are the enzymes topoisomerase I and II.

The higher order intertwining of DNA generally takes the form of supercoiling of the helix axes. Catenation-induced supercoiling occurs by the wrapping of one DNA duplex around another in process involving chirality: left-handed and right-handed toroidal catenanes were observed (Wasserman *et al.*, 1988). While there is left-hand coiling of DNA around histones in the nucleosomes (Germond *et al.*, 1975), intertwined catenanes of DNA substrates containing bacterial and λ-phage attachment sites are exclusively right-handed (Spengler *et al.*, 1985).

Peptides with an excess of negatively charged side chains forming "acid blobs" can stimulate the formation and/or action of transcriptional preinitiation complexes. There is wide variety of such transcriptional-regulatory proteins that bind to specific DNA targets called "enhancers" or "upstream activating sequences". According to Sigler (1988) "The position and polarity of these *cis* regulatory sequences are not stringently defined with regard to the promotor that they activate".

The traditional view of transcription of bacterial operons is that they are synthesized into long polycistronic mRNAs — which represent the sequences of all the genes of the operon. The messengers are sequentially translated into the various proteins for which they code. Termination of DNA transcription has been considered as a plausible mechanism for controlling gene expression during phage infection. This termination is catalysed by one of the innumerable protein factors of macromolecular metabolism in bacteria, the *rho* factor known to act as a template of phage DNA. The immediate significance of *rho*-dependent termination within an operon is that this mechanism may thus explain the natural polarity of some operons, in which later genes direct synthesis of less protein than earlier genes.

Polarity of transcription-translation is revealed when a mutation which prematuraly interrupts translation of any gene of a typical bacterial operon is apt to depress the translation of all genes which come later in order of transcription. The strength of this polar effect depends, in general, on the distance between the

mutant chain-terminating codon, and the normal terminator at the end of the gene — the less this distance the weaker the polar effect tends to be. Both these phenomena — coordinate regulation and polarity — have often been interpreted as meaning that the operon is a unit not only of transcription but of translation as well.

Polarity in prokaryotic organisms reflects their obligate coupling of transcription to translation. By contrast, in eukaryotes, translation occurs in the cytoplasm on monocistronic units of mRNA and reduction of distal expression would be differently operated (Bigelis *et al.*, 1977; Steege and Söll, 1979).

a[4] *Mutations*

Polar mutants are due to base substitution in the DNA, but instead of coding for a different amino acid, the altered triplet now codes for a chain termination codon, i.e. a triplet preventing further translation in a strand of messenger RNA.

Mutations which prematurely interrupt translation of any gene of a typical bacterial operon are apt to depress the translation of all genes which come later in order of transcription. The strength of this polar effect depends, in general, on the distance between the mutant chain-terminating codon, and the normal terminator at the end of the gene — the less this distance the weaker the polar effect tends to be.

The presence of nonsense mutations and the insertion of foreign DNA into the operon are also a cause of polarity. Not all nonsense mutations are polar and, in general, only those at the 5' ends of a given gene are extreme polars, with a gradient of polarity extending toward the 3' end (Watson, 1970). As for insertion mutants, they differ from nonsense mutants in that they are always extremely polar and their polarity does not show the dependence on position within the gene characteristic of nonsense mutations. The polarity of nonsense mutations might be explained by a similar mechanism to that of polarity resulting from *rho*-dependent termination because nonsense codons constitute part of the DNA sequence recognized by *rho* factor. However, polar nonsense mutants in the galactose operon proved to have no effect on transcription in either the absence or presence of *rho*. The polarity of nonsense mutants would thus depend on the failure of translation as such and not on the sequence of the nonsense codons themselves.

Certain mutations in the structural gene of an operon have a dual (pleiotropic) effect. As explained by Mahler and Cordes (1969) "they not only control the structure of the resultant polypeptide but they also effect a decrease in the amount of all enzymes controlled by genes located beyond the gene primarily affected and distal from the operon (-4-3*-2-1-*0*). Moreover, mutation (*) in 3 will affect the amount of enzyme produced by 4,5, etc. Such polarity mutations were first observed in the *His* operon. Only the genes away from the operator were affected. Polarity mutants have been observed in other systems among which the best studied *Lac* operon (Mahler and Cordes, 1969): mutants in the *z* gene produce less of the enzymes con-

trolled by *y* and *x*, and mutants in *y*, the permease gene, produce decreased levels of transacetylase *(x)* but have no effect on galactosidase *(z)*.

Polarity is also revealed by observation of polar effects of nonsense mutations on distal gene expression. For example, when an operon such as *trp* (tryptophan) is transcribed from its normal promotor, amber mutations in the *tryp E* gene can sharply reduce expression of downstream genes. The reduced yields of the distal polypeptides of an operon could be a direct result of altered translational controls; alternatively, they could simply be a consequence of the fact that distal mRNA templates are not present for translation. Evidence for this last proposal has been obtained with the *E. coli* tryptophan operon in which the limitation in gene expression distal to a polar expression derived from the markedly reduced levels of the relevant mRNA sequences (Morse and Yanofsky, 1969; Steege and Söll, 1979). Inversely, polarity supressors selected to relieve mutational polarity in both the *tryp* and *lac* operons simultaneously increased wild type *tryp* mRNA synthesis (Korn and Yanofsky, 1976; Steege and Söll, 1979).

The recently studied crystal structure of the *trp* repressor/operator complex has shown an extensive contact surface including solvent-mediated hydrogen bonds to the phosphate groups of the DNA; no direct hydrogen-bonded or non-polar contact with the bases were detected. This absence of direct hydrogen-bonds implies that "in *trp* system, specificity is not due to direct hydrogen-bonded contact of the major groove's polar groups by the hydrogen bonding groups of the protein" (von Hippel *et al.*, 1986; Otwinowski *et al.*, 1988). A certain degree of helical twist to accomodate both polar and non-polar stabilizing contacts is needed. Three polar contacts mediated by well ordered water molecules bridge the functional groups of mutationally sensitive base pairs to the protein and thereby contribute to "direct readout" of the operator sequence (Otwinowski *et al.*, 1988).

Transposition mutagenesis by so-called transposons has been used successfully for the genetic analysis of cloned genes. As with transposons, insertional mutagenesis with a selectable DNA fragment or interposon has strong polar effects in a wide range of bacterial species. The strong polarity of the interposon mutation manifested in Gram-negative bacteria was probably due to their flanking *rho*-independent transcriptional terminator sequence which comes from a bacteriophage T40 gene (Prentki and Krisch, 1984).

Forbidden inversions have recently been described in the polar rearrangement of the chromosome of enteric bacteria: "sequences at many pairs of sites sequences (permissive) do recombine to generate the expected inversion, while the same sequences placed at other pairs of sites (non permissive) do not form an inversion" (Segall *et al.*, 1988).

a[5] *Gene conversions*

There are two major mechanisms for the control of gene expression: DNA inversion and DNA transposition. In the first or flip-flop control principles, gene expression is either controlled by an invertible DNA segment or activated by the transposition of a copy of the silent gene to expression site, where it can be transcripted (Borst and Greaves, 1987). The duplicative transposition occurred by a gene conversion, in which a gene in the expression site was displaced by the incoming gene copy and destroyed ("cassette model", see Hicks *et al.,* 1977). Duplicative gene conversion can result in the switching of yeast mating-type; in this event, there is replacement of the α sequence of MATα with *a* sequence from HMR*a*.

In certain eukaryotic genes, called polarons, there appears to be a gradient of conversion frequency from one end of the gene to the other as discovered in the fungus *Ascobolus* (Lissouba and Rizet, 1960, in Fincham *et al.,* 1979). The more frequently observed polaron effect is a strong tendency, in inter-mutant crosses, for recombinants to arise by conversion of the more distal (or in other polarons the more proximal) site rather than the other one. According to Fincham *et al.* (1979), "the polaron effect can be accommodated within a hybrid DNA hypothesis by making the additional assumption that hybrid DNA is not formed in a completely random way but tends to be initiated at certain fixed points, if at different points, within rather well-defined short regions".

Among tetrads in which wild-type segregants have arisen by nonreciprocal interallelic recombination it is often observed that conversion occurs at different rates at the two heterozygous allelic sites. More surprisingly, observed conversion rates frequently depend less upon the specific site undergoing conversion than upon the particular second heterozygous site with which it is associated (Emerson, 1966). The most striking examples of this sort have been found at loci affecting ascospore color in *A. immersus,* in which there exists a strongly polarized control of gene conversion (Rizet's group, 1960, 1961, in Fincham *et al.,* 1979). Complete polarization may exist in a large section of a polaron gene. For a given locus, six allelic sites constituting such a polaron were mapped in linear order by the frequencies of recombination occurring between each pair of sites.

In *Neurospora crassa* there are also indications of polarity in interallelic recombination between members of two groups of sites within a locus (Case and Giles, 1958, 1964; Stadler and Towe, 1963), but to a much lesser degree than in *Ascobolus.*

b) *Ribonucleic acid (RNA)*

RNA is involved in many biological functions, ranging from information storage and transfer to the catalysis of reactions implicating both nucleic acids and proteins.

Transcription normally occurs from DNA in which it begins where an enzyme, RNA polymerase, binds to a specific base sequence (the promoter) on one strand

of the DNA double helix. The RNA molecule always grows in the 5'- to 3' direction. The sequence TATA is involved in positioning the RNA polymerase correctly and transcription of DNA begins between 20-30 nucleotides since past that site. A chemically protective "cap" consisting of a methylated guanosine is linked to the first nucleotide by a triphosphate bridge at the beginning (5'-) of the chain in all eukaryotic messenger RNA (mRNA).

In the problem of recognition between mRNA and amino acyl- *t*RNA, one of the questions arisen is "what is the actual sequence of bases for various codons?". Its answer is that polarity is involved in the highly specific selection of the sequence isomers and their recognition examplified as follows (Mahler and Cordes, 1969): "G_pU_pU induces Val-*t*RNA binding while U_pU_pG does not, and in fact produces binding of Leu-*t*RNA".

Processing of nuclear messenger RNA precursors (pre-mRNA) involves two important steps: the removal of introns — intervening sequences separating the protein coding portions of the gene — by splicing — mechanism selecting which combination of exons contributes to the coding — and the generation of the mature 3' ends by cleavage / polyadenylation. Splicing requires the ordered assembly of large multicomponent complexes, the spliceosomes (Darnell, 1985).

Prebiotic molecules of short RNAs were first carriers of genetic information. This information was encoded in RNA and transferred by synthesis of a complementary replica strand, with the original strand as a template along which complementary nucleotides are assembled according to the base-pairing rules. As proposed by Eigen *et al.* (1981), in prebiotic translation the first code had to establish the direction of readoff and the punctuation of the message by defining a "reading frame". According to Crick *et al.* (1976), directionality and framing were initially fixed by translating only triplets that had a special sequence. The first, primordial protein was made of four amino acids that had a negative electric charge and "which would not readily associate with a negatively charged RNA species unless specific forces stabilized a particular interaction" (Eigen *et al.,* 1981).

In similar fashion to the pairing of the two strands of the DNA double helix two stretches of RNA can form a helical region if they contain complementary nucleotides. By pairing with one another, conserved sequences might determine that the intron folds up into a particular shape. Cech and Bass (1986) have shown, in the protozoan *Tetrahymena,* that this shape helps the RNA to remove the intron portion of its own length which could function as "ribozyme".

Double-strand RNA which is the only genetic information of many groups of viruses (retroviruses, etc.) can be considered as homologous to DNA by reference to the inverse polarity of its two strands. In double stranded RNA viruses, it is also the minus strand which acts as a template for the synthesis of a plus strand acting as mRNA. Retroviruses contain two copies of the plus stranded viral RNA genome. A single RNA molecule may be sufficient for the generation of a complete viral DNA

copy during reverse transcription by strand transfer from the 5' to the 3' end of the same template. There are, however, alternatives to this mechanism (Panganiban and Fiore, 1988).

In the process of translation of mRNAs, proteins are synthesized from their N-terminus to their C-terminus by translating their mRNA polarly from the 5' end of the RNA toward the 3' end. Thus, the 5' half of an mRNA molecule codes for the N-terminal half of a protein and the 3' half codes for the C-terminal half.

Initiation of protein synthesis in eukaryotes requires over a dozen proteins. One of its first steps is the recognition of the cap structure at the 5' end of mRNA by the preinitiation complex. This complex then moves on the mRNA in the 3' direction until an AUG codon is encountered. There, initiation factors are released, elongation factors and the large ribosomal subunit join and protein synthesis begins.

RNA that contains the complementary sequence to a given RNA is termed antisense RNA. The use of this "inverse polarity RNA" permitted to produce "mutants without mutations" (North, 1985). In such mimic mutations the expression of the gene is blocked by preventing the translation of its sense transcripts and phenocopies of mutants are produced (Herskowitz, 1987).

2. Proteins

There is a strong tendency for the polar side chains of proteins to seek a polar environment such as water and for the nonpolar ones to be segregated in nonpolar areas. Except in cell membranes, a protein chain will therefore tend to fold so that polar side chains are on the exposed surface and nonpolar ones are inside.

As first shown by Perutz, electrostatic interactions are of considerable importance in protein structure and function, and in a variety of cellular and biochemical processes (Perutz, 1978; Matthew, 1985). The electrostatic attraction between a polar side chain and water is a form of hydrogen bonding, in which a hydrogen atom acts as a bridge between charged oxygen or nitrogen atoms. In the secondary structure or alpha helix of Linus Pauling, hydrogen bonding between peptide units reduces the energy, but if the helix were unraveled, the same sites would form hydrogen bonds with water. Local interactions between nearby amino acids give rise not only to alpha helices, but also to beta sheets (Pauling, 1960) or other forms of secondary structure. These subassemblies, acting as more or less coherent units, organize themselves into domains of the tertiary structure (Doolittle, 1985). Protein structure is stabilized by interaction of α-helix dipole with a charged side chain. This dipole moment results from the alignment of dipoles of peptide bonds (Fig. 6) which can perturb the pKas of ionizing groups (Săli *et al.*, 1988). Folding of small proteins occurs very rapidly even if it is slowed down by the need for proline isomerization or disulphide crosslinkage. Recently, many proteins have shown to contain Zn^{2+} finger domains

where a short polypeptide could be held in a folded structure by the chelation of Zn^{2+} by two cysteines and two histidines appropriately spaced (Sharp and Eisenberg, 1987).

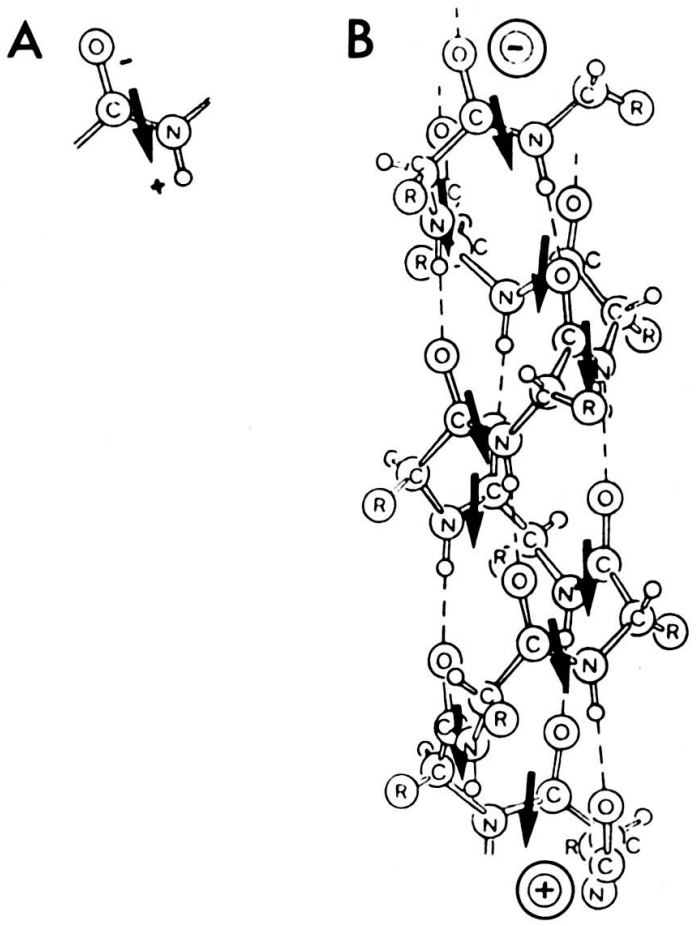

FIG. 6.

(A) Single peptide-bond dipole. (B) Peptide-bond dipoles aligned nearly head to tail within an α-helical protein. Adapted from Creighton, *Nature* 326: 547 (1987).

Recent evidence has shown that the stabilities of helical synthetic polypeptides are influenced by changes in charge at the ends of the helix. Thermostability of a protein such as T4 lysozyme can be increased by introducing charged residues at sites designed to interact with α-helix dipoles. Two mutant lysozymes (Ser Asp; Asn Asp) were designed to increase thermostability by enhanced helix-dipole interactions. An interpretation for each replacement could be that the mutant structure is stabilized by an electrostatic interaction between the negatively-charged side chain and the positive charge attributed to the α-helix dipole (Nicholson *et al.*, 1988).

The peptide-peptide hydrogen bond is the structural element underlying the closely packed helices and beta sheet structure composing the interiors of most

proteins. Net electric dipole moments arise from the uneven distribution of electrons in the carbon-oxygen and nitrogen-hydrogen bonds of each peptide (Matthew, 1985). In the constrained geometry of the protein hydrogen bond network, the peptide dipoles often aligned coherently and parallel to a helix axis may give rise to a macrodipole (Wada, 1976). However, according to Matthew (1985) "it is likely incorrect to uniformly assign the same set of partial charges to peptide dipoles throughout the structures", a problem which should be solved by "an all-atom computation including discrete solvent atoms and individual atomic polarisabilities".

According to Quiocho *et al.* (1987), highly polarized peptide bonds are the net effect of three features involving hydrogen-bonds and the association of peptide units with charged groups which are in turn coupled to hydrogen bond arrays. These authors proposed a general mechanism in which the isolated charges on the various buried, desolvated ionic groups are stabilized by these polarized peptide units.

L-Arabinose-, sulphate- and leucine/isoleucine/valine binding proteins are members of a large group of periplasmic proteins (collectively called "binding proteins") which serve as initial high-affinity receptors of bacterial active transport systems for a variety of nutrients and of sugar chemotaxis (see V). The presence of several helices in the binding-protein structures, together with the concept that the alignment of peptide dipoles parallel to the helix axis may give rise to a macrodipole with partial charges, suggests that helix macrodipoles could provide some means of stabilizing the charges on various ionic groups. Helix macrodipoles could assist in the binding of sulphate, although they have no effect on Arg 151 and the bound leucine. This mechanism could also be applied to the activation of catalytic residues (Quiocho *et al.*, 1987).

At their high level of organization, some proteins can have nonpeptide components such as metal ions required for the activity of certain enzymes or the porphyrin ring found in chlorophyll (Mg) and hemoglobin (Fe). Iron is itself the active center of the group interacting with oxygen. The oxygen-carrying plate-like haem of the globular protein myoglobin has been shown by classical crystallographic studies (Kendrew *et al.*, 1960) to be embedded in a pocket that is lined largely with nonpolar amino acids. However, on the basis of the often minor changes in acid dissociation constants of proteic groups in water and inside certain proteins, Warshel and Russell (1984) have recently deduced that many protein cavities have a high effective polarity despite the lining of nonpolar amino-acid residues. Warshel (1987), in reexamining the theories concerning the interactions of charges located within proteins, points out that the similarity of the dissociation constants of proton-bearing groups within the protein and in water force the conclusion that the protein is polar in their immediate vicinity, and that previous theoretical treatments had neglected the role of peptide bonds in the stabilisation of charges.

The translation of the three dimensional-structure of proteins to electrostatic free energies is far from simple. According to Warshel (1987) "By genetic engineering

techniques, it is now possible to examine theoretical predictions directly by experimentally creating or removing charged residues, and even changing the hydrogen-bonding pattern around such residues''. Modelling of electrostatic effects in proteins — by algorithm — requires consideration of the atomic polarizabilities of the heterogeneous protein and the solvent, including both water and counterions (Warshel and Russell, 1984). The model was founded in classical electrostatic theory, which treats materials as homogeneous dielectric media that may be polarized by electrical charges. Proteins are known to have low dielectrical constants (about 2-3), because reorientation of their dipolar groups is severely restricted. Because water has a very high dielectric constant (± 80), the protein-water interface constitutes a boundary between two dielectric media. The Poisson-Boltzmann equation describes systems composed of charges and dielectric media. To solve it, numerical methods on high-speed computers are implemented (Warshel, 1987). They have already been applied to proteins in studies of substrate diffusion in the electric field of Cu, Zn-superoxide dismutase, the interactions of Klenow fragment of DNA polymerase with B-DNA and the effect of a specific charge on the redox properties of cytochrome C_{551}. As recently put by Warshel (1987) ''It seems clear that only models that take into account both the polarity of the water and of the protein could reproduce the observed energetics''.

Preceding authors (in Warshel, 1987) have attributed electrostatic effects to the macrodipoles of the protein helices (which are in fact almost completely shielded by large dielectric effects) rather than to local polarity. However, isolated ionized groups have been found in protein interiors, for example the aspartate 102 in chymotrypsin (Warshel and Russell, 1984). The recent work by Quiocho *et al.* (1987) in which were identified highly refined cases of uncompensated charges stabilized by the surrounding protein dipoles (positive groups by the C = O dipoles and negative groups by the N-H dipoles) has confirmed this study and thus the importance of the local dipolar components (Warshel, 1987).

The true polarity of the interior of proteins has further been confirmed by the use of bound fluorophores, the spectroscopic properties of which reflect the polarity of their environment. Polar chromophores and fluorophores are molecules that exhibit large differences in the electronic distribution of the ground and fluorescent states. The local polarity of a biological system harbouring such molecules can be estimated by comparison of the spectral properties in that environment with those of the probe in solvents of known polar characteristics. The ideal polar fluorescent probe should have no sources of charge besides two readily identifiable monopoles, and the dipole moment that they determine should change strength but not direction in the excited state. Most important, the probe should be soluble in a range of solvents that cover the polarity scale, from apolar solvents like cyclohexane to the most polar ones like alcohol or water. Optical spectroscopy studies with fluorescent probes such as amino- naphtalene sulfonate (MacGregor and Weber, 1986) have confirmed that

the myoglobin haem cleft has a polar interior and that polarity can be accounted for by peptide amine dipoles (Cowley, 1986).

Chromoprotein molecules such as those of hemoglobin can be associated into fibers. Mutational abnormalities in these fibers can result in abnormal shape of the cell as that of erythrocytes in the sickle (S) cell anemia (Pauling *et al.*, 1949). The only molecular difference between normal and abnormal (S) hemoglobins is the replacement in position 6 of the two β chains of a polar amino acid, glutamic acid, by a non polar one, valine (Ingram, 1956). In sickle cells, the 7 pairs of strands of hemoglobin (HbS) fibers are polarly oriented. Improved, three-dimensional reconstructions from electron micrographs recently showed 3 pairs in one orientation and 4 in opposite orientation (Rodgers *et al.*, 1987). The alignment and the association of these long double-stranded fibers of opposite polarities could thus concour to the distorted shape of the mutated erythrocytes S cells.

Membrane proteins have a number of charged groups on their surface and polarized bonds between their atoms which give rise to bond dipoles. An electric field can therefore exert conformational forces on their molecular structure. Protein's net electric charge might also control polarized transport of proteins as suggested by Woodruff and Telfer (1980) who microinjected fluorescently labelled proteins with different pKs into either the nurse cell or the oocyte of *Drosophila* and followed their movements across the cytoplasmic bridge. Injected electronegative protein, serum globulin, moved only from nurse cell to oocyte, whereas electropositive lysosome only moved from oocyte to nurse cell. These observations suggest that the protein's charge is the main factor determining the direction of intercellular transport (Nuccitelli, 1983). However, by microinjecting fluorescently labelled acidic and basic proteins into the nurse cells or oocyte of vitellogenic *Drosophila* follicles, Bohrmann and Gutzeit (1987) failed to obtain evidence for charge-dependent migration of these molecules.

Electric current can also influence protein biosynthesis. Using collagen matrices of mammalian fibroblasts, McLeod *et al.* (1987) found a reduced rate of incorporation of proline into extra- and intracellular protein depending of the frequency and amplitude of the applied electric field. In tissues containing cells aligned either parallel or perpendicular to the electric field, this response was dependent on the orientation of the cells relative to the electric direction. It was thus demonstrated that "currents of physiological strength can stimulate alterations in protein biosynthesis and thereby may influence tissue growth, remodelling, and repair".

3. Lipids

Neutral fats or triacylglycerols, the major type of lipid stored in cells, are entirely unpolar.

Glycerophosphatides or phospholipids are mainly represented by derivatives of phosphatidic acid in which a phosphate group replaces one of the fatty acids.

Consequently, they are amphipathic, that is one end of the molecule is strongly hydrophobic (i.e. the end containing the hydrocarbon chains) while the other end is hydrophilic due to the charged nature of the dissociated phosphate group and other substituents.

Amphiphiles are surface-active substances or "surfactants" which tend to accumulate at the water surface where they lead to a strong decrease of the surface tension of the solution. They are organic, lipid-type molecules such as soaps and detergents that consist of a hydrophobic, head group, and an oleophilic, hydrocarbon chain; at ambiant temperature, the hydrophilic head groups form hydrogen bonds with water and they can mix water and oil into stable, homogeneous solutions called "microemulsions". There are two classes of amphiphiles: ionics such as the single-tailed sodium dodecyl sulfate (SDS) or the more oleophilic double-tailed compounds; nonionic (uncharged) amphiphiles are not equally well soluble in water and oil (Kahlweit, 1988). At low temperatures, such amphiphiles are more soluble in water than in oil while the reverse is true at elevated temperatures. Nonionic amphiphilic molecules can thus be viewed as "chemical dipoles" rather than electric dipoles. The higher their "dipole moment", that is, the stronger their amphiphilicity, the stronger the tendency of their micelles to form more complex structures such as hexagonal and then lamellar liquid crystals (Kahlweit, 1988).

Of special significance are special membrane structures formed by phospholipids in aqueous solution such as spherical micelles, with hydrophobic interior of fatty acyl side chains, sheets of phospholipids in a bilayer and spherical liposomes comprising one phospholipid bilayer.

Compounds such as steroids and fats, i.e. compounds that have a preponderance of hydrocarbon groupings, which makes them only slightly soluble in water, also have the ability to spread on surfaces and to have a particular orientation so that their polar groups will, in general, be in contact with water or with other polar groups, and those parts that have hydrocarbon groups will be out of contact with the water. Furthermore, lipoprotein films which combine both these two types of compounds, can even be formed artificially, for instance cholesterol and the protein gliadin have been used to produce such a film.

In membranes of fungal hyphae, additionally to phospholipids, there are sphingolipids which are also polar molecules containing hydrophilic "head" and a long hydrophobic "tail" (Weete, 1974). Phosphatidylcholine and phosphatidylethanolamine are the most common phospholipids with phosphatidylserine and phosphatidylinositol present in smaller amounts.

Glycophospholipids are also amphipathic molecules which can anchor proteins to the plasma membrane at the cell surface. A molecule with a bulky polar head group such as glycosylphosphatidylinositol might serve as surface-oriented lipid protein anchor (Low and Saltiel, 1988).

4. Polysaccharides

Homopolysaccharides contain only one type of sugar residue. Oligomers or polymers of uronic acids known as uronides are constituents of pectic substances. Owing to the presence of carboxyl groups, they are polar compounds capable of forming salts (insoluble Ca-pectates). Homopolysaccharides can also become polar by phosphorylation: phosphomannans secreted by yeasts of the genus *Hansenula* have phosphodiester linkages between their mannose residues (Rose, 1968). In cellulose, the successive links of β-glucose are rotated through 180° with respect to each other, but in starch the α-glucose residues can react with each other without rotation. The cellulose chains have a twofold screw axis as an element of symmetry; consequently, they are more stable than the starch molecules which are deprived of axial symmetry (Frey-Wyssling and Mühlethaler, 1965). According to structure analysis, the cellulose chains in the crystal lattice run anti-parallel with respect to each other, so that at the end of a polar fibril the glucose residues must be added in two different ways to form 1,4-bonds with half of the chains, and 4,1-bonds with the remaining chains (Frey-Wyssling and Mühlethaler, 1965).

Heteropolysaccharides contain more than one type of sugar monomers. The capsular heteropolysaccharide responsible for the serological specificity of type III of pneumococci is negatively charged by the carboxyl group of its glucopyranosyl-glucosyl unit. Such polar property of the heteropolysaccharide might intervene in the ionic bonding suggested to be largely responsible for "trapping" the capsular layer around the cell wall (Rose, 1968).

Acid mucopolysaccharides such as hyaluronic acid (alternate group units of glucuronic acid and acetyl-glucosamine), chondroitin, and keratan sulfates have also polar characteristics which might be related to their role in the cells.

5. Enzymes

Enzyme processes can be described by energy surface. The reaction's activation energy is a difference in height of the points on the energy surface representing the starting materials and the transition state. Enzymes can lower the reaction's energy barrier by binding most strongly not to the reactants but the transition state which is thereby stabilized (Pauling, 1960). Hydrolysis of an ester passes through an unstable state whose distinctive shape and charge can be mimicked by a stable molecule. The transition state is tetrahedral and its polarized partial negative charge is concentrated at one apex (Fig. 7). It was recently found that crossed asymmetric insertion — by fusion of encoding DNA — of polar domains of the two aspartate and ornithine transcarbamoylases (AT- and OTCases) from *E. coli* can switch substrate specificity (Houghton *et al.*, 1989).

From his study of the kinetics and specific inhibition of acetylcholine esterase, Wilson (1954) has proposed a bipolar interpretation of the interaction between active

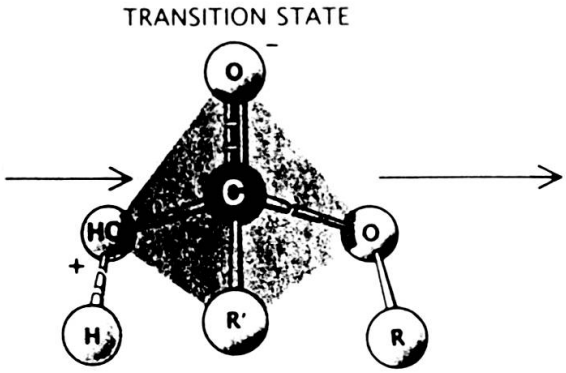

FIG. 7.

Polarized tetrahedral transition state in the hydrolytic process of an ester. (R and R' = chemical groups not taking part in the reaction). From R. A. Lerner and A. Tramontano, Scientific American, vol. 258 (3), p. 46 (1988), with authorization.

groups of the enzyme and its substrate occurring at two sites: the "esterasic site" in which a basic group G donates an electron pair becomes (+) and binds the acyl-group (RCO-) while liberating choline (R'OH); the "anionic group", responsible for the specificity of the enzyme in which a negatively charged group on the enzyme is thought to combine with the positively charged trimethylammonium group of the substrate. This hypothesis of the intermediate formation of acyl-enzymes has received support from several other enzymes among which chymotrypsin. Kinetic studies of several of these enzymes that act at ester and amide bonds have suggested that the imidazolyl group of a histidine residue may serve as the basic electron donor at the positive "esterasic site". In some enzymatic reactions, both general acid and general base are involved as catalysts in concerted proton tranfers. In such polarized reactions, the general acid donates a proton to the transition-state species and the general base accepts one (Lehninger, 1975; Koshland, 1987).

The decarboxylation of β-keto acids such as dimethyl-oxaloacetic acid is catalyzed by metal ions such as Cu^{2+}; this has been attributed to the formation of a chelated intermediate and to the attraction by the (+) charged copper ion of electron away from the β-carboxyl group, with the liberation of CO_2 (Steinberger and Westheimer, 1951). This mechanism may simulate the enzymatic decarboxylation of β-keto acids, which is known to be dependent on the presence of metal ions.

Most of the enzymes catalyzing biosynthetic reactions of phospholipids are amphipathic molecules (Darnell *et al.*, 1986); thus, the introduction of a double bond in stearoyl-CoA to form the CoA derivative of oleic acid is catalyzed by stearoyl-CoA desaturase which has the hydrophilic part of its active site facing the cytosol and the hydrophobic part associated with the endoplasmic reticulum.

Heme-containing enzymes are versatile biocatalysts. They have the ability to incorporate oxygen atoms from O_2 into organic substrates (the oxygenase activity) and to use H_2O_2 and other peroxides to oxidize substrates (the peroxidase activity).

Oxygenases and peroxidases use O_2 and H_2O_2 respectively, through variations in the nature of the proximal axial ligand, in the polarity of the proximal and distal heme environments, and in the degree of accessibility of the heme iron center and the heme edge (Dawson, 1988). The mono-oxygenase cytochrome P-450 has a thiolate axial ligand and a very nonpolar environment for the heme and for the organic-O_2 substrates. The horseradish peroxidase has an imidazole axial ligand and a charged and polar heme environment to facilitate cleavage of the peroxide O-O bond. As for the chlorperoxidase, which can carry out one-electron (peroxidase) as well as two-electron (peroxygenase) reactions, it has a mechanistic activity falling between that of P-450 and that of horseradish peroxidase. The electrostatic potential in and around the active site channel of superoxide dismutase (SOD) has been calculated and its deep channel found to be positively charged, with residues critically placed to guide an incoming negatively charged ion towards the copper ion in the active site (Getzoff *et al.*, 1983). The metalloenzyme SOD which catalyzes the dismutation of O_2- thus provides a particularly good example of an asymmetric distribution of charged groups on the surface of a protein as also occurring in the dipolar cytochrome c (Margoliash *et al.*, 1982) and ribonuclease T1 (Heinemann and Saenger, 1982).

The internal motions that underly the workings of enzymes are best explored in computer simulations (Karplus and McCammon, 1986). These simulations applied to ribonuclease have shown that a number of amino acids in these enzymes make a network of hydrogen bonds to the substrate. In its absence, "water molecules occupy the positions of the polar substrate atoms in the binding sites and produce a network of hydrogen bonds that mimic the substrate interactions and stabilize the positions of binding-site amino acids".

6. ANTIGENS-ANTIBODIES

In the immune reaction, antigen-antibody reactions involve a specific molecular union between the two reagents. Antigens are normally macromolecules such as proteins. Antibodies are specific not only for a given antigen molecule but for certain regions of that antigen, the antigenic determinants or haptens. Hydrophobic haptenic groups tend to form more stable complexes with antibody than do water-soluble groups. Not only hydrophobicity but also polarity have been suggested to play a role in the interaction between antibody and the antigenic determinant (Levy *et al.*, 1973).

Every individual antibody molecule has one type of light (L) chain and one type of heavy (H) chain. The chains are held together by disulfide (-S-S-) bonds to form a monomer; two monomers are linked by -S-S- bonds to form the basic dimeric structure of the antibody. Within each chain, units made up of small peptides of about 110 amino acids fold up to form compact domains. The N-terminal domain is called the variable region and the C-terminal domain is called the constant region.

The variable domains of L and H chains are bound to one another and interact closely to form a single compact unit which is the antibody-binding site to antigen. The three amino acid segments of the variable region contribute to that antigen binding; they are loops that extend from the so-called immunoglobulin fold (Darnell *et al.*, 1986). The binding site being complementary in structure to the antigen, its three segments are therefore called complementarity-determining regions (CDRs). In the CDR regions, amino acids contact the hapten phosphorylcholine and form a cleft or pocket into which the phosphorylcholine hapten fits. A combination of electrostatic, hydrogen-bonding and van der Waals forces hold the hapten in the cleft (Capra and Edmundson, 1977).

The antigen-combining site of antibody molecules consists of six separate loops supported by a conserved β-sheet framework. Modelling of these polypeptide loops with special regard to antibody-binding site modelling are promising X-rays crystallographic technics to allow selection of better-quality antigenic conformations (Bruccoleri *et al.*, 1988). In such model constructions there should be used distance-dependent dielectric constant in the evaluation of electrostatic interactions and given consideration to polarity relationships between backbones of the framework residues in the antigen combining sites.

7. Synthetic polymers

Polymer scientists have designed molecular chains to increase their mutual attractions and thermoplastic's toughness. In pure polystyrene, which consists of a backbone of covalently bonded carbon atoms surrounded by hydrogen atoms and bearing a six-carbon benzene ring on every second backbone atom, the molecule is nonpolar: it lacks localized concentrations of positive and negative charges, which could give rise to electrostatic forces among the chains.

If more polar monomers-monomers with a less symmetric distribution of charges are placed along the polymer chain, electrostatic attraction will contribute to the intermolecular forces. For example, the nonpolar styrene monomers can be interspersed with more polar acrylonitriles monomers, a step that increases the cohesion of the molecules and hence the solvent resistance of the bulk material. The copolymerization of styrene and acrylonitrile is now routine, and many other commercial polymers similarly incorporate two or more different monomer building blocks.

Formaldehyde (CH_2O) polymers can be synthesized in the deep cold at liquid helium temperatures (4.2°K) by a quantum-mechanical effect known as tunnelling which allows some forbidden reactions to take place near absolute zero. Tunnelling results of the particle-wave duality of all matter and radiation and could thus also dominate the behaviour of "particles" such as atoms and molecules (Goldanskii, 1986). The quantum-tunnelling of particles would provide the additional activation

energy to the reacting molecules, an effect also shown to play an important role in redox reactions at low temperature.

A salient feature of polymers is their alternating single and double bonds, which create an electronic configuration that is susceptible to doping. It has recently been found that doping of ordinary polymers by either chemical or electrochemical processes can change their electronic profile of insulators and semi-conductors (Kaner and MacDiarmid, 1988). Electrons are removed from the valence band (positive doping by an iodine solution, for example), or added to the conduction band (negative with a solution of sodium metal in mercury, for example). The charge conferred on a polymer backbone by doping causes a slight but important change in the position of its atoms which prompts the formation of one of three types of charge "islands" called solitons, polarons and bipolarons. Electrons can freely flow between these adjoining islands and so, solitons, polarons and bipolarons are responsible for making polymers conduct. This is well examplified by plastics such as the positively-doped polyacetylene chain which has a delocalized positive charge that enables it to conduct like a metal (Kaner and MacDiarmid, 1988). These recent findings open the interesting possibility of artificial nerves made of polarly conducting polymers.

B. Aggregates

1. Crystals-quasicrystals

A conventional crystal is a particularly well-ordered arrangement of atoms or molecules. It is a lattice-work, in which identical "unit cells" fit together regularly and periodically to fill space. Inherent in every crystal structure are certain symmetries: three- four- sixfold rotational symmetry for equilateral triangle, square, hexagon respectively. The fivefold symmetry is displayed only by quasicrystals which embody a novel kind of order between crystalline and amorphous (Nelson, 1986).

In ionic lattices of crystals, oppositely charged ions exert a polarizing effect upon each other which can lead to deformations of the electron orbits (Fajans, 1925); then the ions can no longer be regarded as spherical and ionic lattices often possess a low symmetry. The ionic lattice is broken down and the solid crystal dissolved when ionic bonds are broken by polar liquids (water, ammonia) and by hydrogen or hydroxyl ions.

Antifreeze polypeptides are aligned at the surface of growing ice crystals. Dipole interactions between the α-helical polypeptide molecules and water molecules in the crystal lattice were found to be maximized for antiparallel alignment (Yang *et al.*, 1988). Adsorption of polar additives will be unfavoured on the basal and prism faces which are the ones with the lower energy. According to the mechanism proposed for the antifreeze effect, the relative orientations of the ice dipoles and the prism faces direct the binding of the protein. The ordering of the ice dipoles by the helices on the surfaces would make these sites less favourable for further growth of the ice crystals.

2. VIRAL ASSEMBLIES

Viruses are built from many protein subunits which are shape determining. Nucleic acids are dispensable in this process as demonstrated by assembly of nucleic acid free viral shells (capsids). Therefore, the ability to determine the shape of the assembly product or aggregate can only be explained by the binding properties and geometry of these protein subunits (Kellenberger, 1984).

a) *Spatial symmetry*

As assemblies of fully shape determining subunits viruses can be considered as the first level of complexity in biological form. One of the simplest is the Tobacco Mosaic Virus (TMV) with its filamentous capsid built as a *cylindrical* assembly of identical protein subunits.

Spherical virions can also be built from at least 60 identical subunits when assuming icosahedral symmetry and the principle of quasi-equivalence for the positions of each protein subunit with respect to each other of the same virus (Caspar and Klug, 1962). Building of larger, closed particles related to icosahedral symmetry rather obey to a quasi-relevance principle, the shape of the proteins being slightly different according to their positioning on the faces, edges or vertices of the icosahedron (Caspar, 1980). The symmetry is therefore only approximately fulfilled. According to Kellenberger (1984), "close contact between subunits is essential, ascertaining the mechanically required, relatively strong interactions. The interactions are the sum of so-called weak (non covalent) interactions that occur in the area of contact".

A very significant feature of organized structures built by use of the specificity of the relatively weak non covalent interactions that are possible between macromolecules (proteins, etc.) is that their design and stability can be determined completely by the bonding properties of their constituent units. Thus, once the component parts are made they "assemble themselves" (Caspar and Klug, 1962) without a template or other specific external control, although, in the case of certain enzymes and in the sheath of the T4 bacteriophage tails, a controlling function may be exerted by the binding of some other small molecule(s); this changes the configuration of the protein molecules so that they have the appropriate bonding properties.

As explained by Klug (1969), the possible types of spatial symmetry are restricted by the geometry of space. Since biological structures are built of molecules which are different from their mirror image, mirror or inversion symmetry is not possible at the molecular level of organization. The only kinds of spatial symmetry operations possible are rotations and translations.

b) *Polar elongative assembly*

The length of a helical structure such as TMV is not determined by symmetry since the structure can repeat indefinitely along a line (the helix axis). A helical arrangement of subunits is therefore theoretically infinite in extent, but in the case of TMV the length of the particle is determined by the length of the RNA.

Electric bipolarization intervenes in the interaction of RNA and protein in the best known of the self-assembly systems, that of the single-stranded TMV: the negative phosphates in the RNA chain are held in place by salt bridges to the positive side chains of two arginines (Arg-90 and Arg-92) of the protein (Klug *et al.*, 1979).

Phage tails are also suitable model-systems for exploring length determination (King *et al.*, 1978; Hohn and Katsura, 1977). In their monopolar assembly lengthening, the central tube or core plays an important role. It has been suggested that "a single gene product is producing a fibrous protein which determines the length by acting as a ruler. The protein would be attached to the base plate and guides the assembly of the subunits around it. The end of the fiber would be the signal for termination of the polymerization" (Kellenberger, 1984).

P66 is a most interesting morphopoietic factor: indeed, if its encoding gene 66 only is mutated, but all other genes active, then appears a highly increased amount of the shortheaded, isometric variant of T4 (references in Kellenberger, 1966). Therefore, this gene product has an "elongating" morphopoietic function.

c) *Polar viral morphopoiesis*

The active process of viral form determination is very precisely genetically predetermined. Contrarily to standard morphogenesis, it is not under time-control of transcription or translation but the consequence of protein-protein, possibly protein-nucleic acid interactions. Therefore, proteins of a precursor particle must first reach a certain conformation to provide binding sites for an additional protein.

The form of bacteriophage pre- or prohead (DNA-filled when mature) has been shown to be determined in a large part by the scaffolding action of their proteineous cores (Kellenberger, 1980); such "scaffolding", "assembly" or "morphopoietic" cores are transient structures which carry form-determining information. The core of the bacteriophage T_4 is especially rich in such information and thus, by the action of drugs or by mutations could be diverted from its normally slightly elongating, so-called prolate icosahedral form to many form variants: short or isometric heads, head length variants and tubular forms, the so-called polyheads produced in the absence of core proteins. In extreme cases, giant tubular phages with tails at both extremities are even produced; this should imply a surprising and still unexplained inversion of assembly polarity or "neobipolarity".

Interactions of electric charges can control the length of phage particles as recently shown by Hunter *et al.* (1987) in the filamentous bacteriophage fd. This

phage comprises a circular, single-stranded DNA molecule enclosed in a cylindrical protein sheath to form a flexible particle. Its coat protein subunit is 50 amino-acid residues in length and, in the virus particle, adopts a largely α-helical conformation, with the long axis of the helix aligned close to the long axis of the filament. This protein is arranged with its negatively charged N-terminal region on the outside of the filament and its positively charged C-terminal region on the inside abutting the DNA. The positive charge on one of the four lysine side chains in the latter region has a direct effect on DNA packaging, because when this charge is absent, elongated particles are produced with lengths that can be correlated with the residual positive charge in the C-terminal region of the coat protein subunit. There are four lysine residues in the C-terminal region of the major coat protein and, according to the same authors, "it is reasonable to suppose that some or all of these have an important function in neutralizing the negatively charged phosphodiester links in the encapsidated DNA".