The role of the α-helix dipole in protein function and structure

Hol, Wim G. J.

Published in:
Progress in Biophysics & Molecular Biology

DOI:
10.1016/0079-6107(85)90001-X

Publication date:
1985

Prog. Biophys. molec. Biol., Vol. 45, pp. 149–195, 1985.
Printed in Great Britain.

0079-6107/85 $5.00 + .50

THE ROLE OF THE α-HELIX DIPOLE IN PROTEIN FUNCTION AND STRUCTURE

WIM G. J. HOL

Laboratory of Chemical Physics, University of Groningen, Nijenborgh 16, 9747 *AG Groningen, The Netherlands*

CONTENTS

I. INTRODUCTION

Globular proteins are the sophisticated, molecular "chips" of living organisms performing an incredible range of different functions in each cell. Some guide the flow of single electrons, others interact with large macromolecules, whereas fascinating multiprotein complexes are the essential features of such processes as photosynthesis, oxidative phosphorylation, protein synthesis and DNA replication. Therefore, understanding how globular proteins work and

why they fold into compact globular units is one of the central issues of molecular, or perhaps better, atomic biology.

A frequently encountered building block of proteins is the α-helix. This regular arrangement of the polypeptide chain carries a large macrodipole which has been known for a long time but which has often been ignored when discussing the importance of α-helices for biological systems. As the physical aspects of the α-helix dipole have been reviewed by Wada (1976), this review will focus on the effects of the α-helix dipole on the function and on the structure of proteins. It will be shown that the α-helix dipole plays an important role in modulating the properties of several enzymes and in defining the mode of coenzyme binding by numerous proteins. Further, the intriguing possibility that the α-helix dipole, together with the dipole of parallel β-strands, may explain the most frequently observed folding patterns in proteins will be considered in detail. This also leads to a discussion of the electrostatic interactions in proteins, still poorly understood on a quantitative level, but nevertheless crucial for protein function and folding. In this context, a brief overview of crystal structures of uncharged helical oligopeptides is given, in order to see if the helix dipole affects the helix packing. Finally, two quite different processes: (i) charge separation in photosynthesis and (ii) proton and ion transport across membranes, will be briefly considered, since hypotheses based on the α-helix dipole have been put forward as an explanation of these phenomena.

II. THE α-HELIX DIPOLE: ITS ORIGIN, SIZE AND DIRECTION

The α-helix dipole originates from the dipole moment of the individual peptide unit. This dipole moment is considerable due to the partial double bond character of the N–C bond (Schulz and Schirmer, 1979; Pauling, 1960). A commonly accepted value for this dipole moment is 3.5 Debye which is equivalent to 0.7 e Å or 1.2×10^{-29} C m (Hol *et al.*, 1978). A charge distribution which represents this dipole is given in Fig. 1, with its direction parallel to

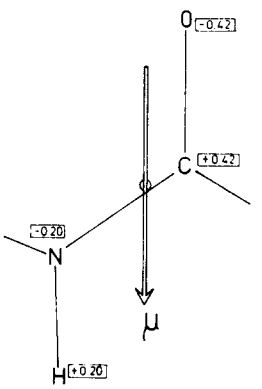


FIG. 1. Charge distribution of the peptide unit according to Hol *et al.* (1978). The dipole moment with the partial charges indicated amounts to 3.46 D = 0.72 eÅ = 1.155×10^{-29} Cm. Numbers in boxes give the approximate fractional charges (in units of the elementary charge). (Fig. from Hol *et al.*, 1978.)

the C=O and N–H bonds. Other peptide charge distributions are used, but they give only marginally different values for the helix dipole moment and will be discussed in a later section, in connection with β-strands.

In an α-helix the peptide units are aligned in such a manner that ~97% of the peptide dipole moments point in the direction of the helix axis (Wada, 1976 and Fig. 2). Wada (1976) has shown that the percentage of the peptide dipole in the direction of the helix axis is quite insensitive to the φ,ψ angles. This means that, in a first approximation, the dipole of a helix of n residues is $n \times 3.5$ D. The N-terminus of a helix is the positive end of the dipole, the C-terminus the negative end.

It has been demonstrated that the helix dipole can be quite well approximated by placing

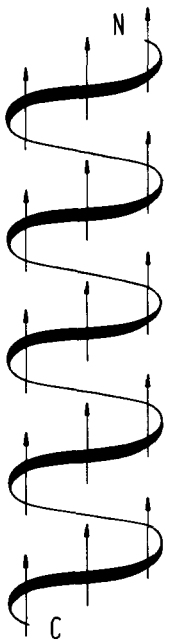


FIG. 2. Idealized and schematic view of the alignment of the peptide dipole moments parallel to the α-helix axis. A more detailed picture is given in Fig. 4 of Wada's review (1976). (Fig. from Hol and Wierenga, 1981.)

half a positive unit charge near the N-terminus and half a negative unit charge near the C-terminus of the helix (Hol *et al.*, 1978; Sheridan and Allen, 1980; Sheridan *et al.*, 1982; Ovchinnikov and Ukrainskii, 1978). This can easily be understood in the following way. Each individual peptide has a dipole moment of ~ 3.5 D $= 0.72$ e Å $\cong 0.5$ e $\times 1.5$ Å. As the axial shift per residue in an α-helix is also 1.5 Å (see e.g. Schulz and Schirmer, 1979, p. 67), it is clear that all dipole charges cancel except for the N-terminal and C-terminal ones.

The dipole moment of a molecule is not a fixed value but depends on the molecular conformation and on its local environment. In particular hydrogen bonds can give rise to significant increases in dipole moments. These effects are large as is shown by experimental determination and quantum mechanical calculations of the water dimer (Yamabe and Morokuma, 1975; Dyke and Muenter, 1974). The dipole moment of the individual water molecule in the dimer appears to be increased by 20% compared with that of the isolated water molecule. In ice, an increased dipole moment of 50% for the water molecule has been reported by several investigators (Coulson and Eisenberg, 1966a,b; Adams, 1981; Barnes *et al.*, 1980). Cooperative effects in simulated water gives similar results (Barnes *et al.*, 1979; Goodfellow, 1982). It is therefore quite likely that the dipole moment of the peptide unit in an α-helix is also considerably larger than that of an isolated peptide unit. Wada (1976) has provided evidence that the individual dipole moment of a peptide unit in an α-helix may be increased up to ~ 5 D, i.e. by roughly 50%. This is supported by dielectric measurements of Applequist and Mahr (1966). Large-scale *ab initio* quantum mechanical calculations (Van Duijnen and Thole, 1982) have shown that an increase of 25% occurs. These calculations concerned a short, nine-residue, helix, and were carried out with a minimal basis set, but the increase nevertheless corresponds well with the estimate of other investigators. It can therefore be concluded that the dipole moment of an α-helix is 25–50% larger than based upon a simple addition of static peptide dipoles.

To summarize, the electrostatic effect of the α-helix dipole is roughly equivalent to the effect of two partial charges of opposite sign, placed at the ends of the helix. The size of these charges is $+0.5$ to $+0.75$ unit charge at the N-terminus, and -0.5 to -0.75 unit charge at the C-terminus. In the sections which follow, we shall look into a number of quite different biochemical phenomena in which the electrostatic field of the α-helix dipole plays a crucial role, or is proposed to do so.

III. THE α-HELIX DIPOLE AND ENZYME ACTION

1. *Introduction*

X-ray diffraction has revealed the three-dimensional structure of a large number of enzymes. Nevertheless, even for the best studied of these, the catalytic mechanism is often still subject to considerable debate. We will see, however, that in several instances a helix dipole is very likely to be involved in the enzymatic process. It is obvious that within the scope of the present review it is impossible to discuss the various mechanisms in detail, let alone to provide a quantitative assessment of the effect of the helix field on reaction rates. We will therefore limit ourselves to describe briefly the mode of action of those enzymes where the active site is located close to a helix terminus.

2. *Active Site Helices in Enzymes*

Ten "active site helices" occurring in nine completely different enzymes have been discovered by X-ray analysis so far (Table 1). From a recent review (Richardson, 1981) it can be estimated that for about 36 evolutionary unrelated enzymes the three-dimensional structure is known. Hence, one in four enzymes of known conformation has an α-helix dipole affecting the electrical field in its active site. Three-dimensional views of all ten "active site helices" are given in Fig. 3. The best studied example, the sulphydryl protease papain and its relative, actinidin, will be discussed in Section III.2.(j). First we will look at the other nine enzymes—thioredoxin being an "honorary" enzyme which seems to be justified in the present context.

(a) *Rhodanese*

Rhodanese is a ubiquitous enzyme occurring in many different organisms (Westley, 1973). In mammals the highest concentration is found in liver and kidney where it is located in the mitochondria. The enzyme is probably involved in cyanide-detoxication as it catalyzes the following reaction:

$$CN^- + S_2O_3^{2-} \rightleftharpoons SCN^- + SO_3^{2-} \quad (1)$$

Several other functions have been suggested, however, one of which is that rhodanese functions as a "sulphur-insertase" for the iron–sulphur complexes observed in many electron-transport proteins (Westley, 1973; Bonomi *et al.*, 1977a,b; Pagani and Galante, 1983).

The essential group in the catalytic mechanism of this enzyme is the sulphydryl of Cys-247 which transfers the outer sulphur atom from thiosulphate to cyanide in a two-step process (Westley, 1973; Ploegman *et al.*, 1978, 1979). The pK of this sulphydryl group is abnormally low. Westley and coworkers have reported a value of 6.5 (Schlesinger and Westley, 1974). In the three-dimensional structure of rhodanese, the charged group nearest to Cys-247 is Asp-180 (Ploegman *et al.*, 1979). This is hard to reconcile with the low pK of the catalytic group. It appears, however, that Cys-247 is located at the N-terminus of a long, central α-helix (αD′) [Fig. 3(a)]. Moreover, a second α-helix (αE′), which is extended by the peptide units of residues 247–249, also points its N-terminus to the essential sulphydryl group [Fig. 3(a)]. It therefore appears likely that the electrical field due to the central and to the elongated helix, is responsible for the low pK of the essential sulphydryl group of rhodanese.

(b) *Glutathione peroxidase*

The seleno-enzyme glutathione peroxidase (GSHP) catalyzes the reduction of a variety of hydroperoxides, including lipid hydroperoxides, and is believed to protect cells against the dangerous effects of the oxygen molecule. During the reaction with peroxides glutathione (GSH) is oxidized:

$$ROOH + 2GSH \longrightarrow ROH + H_2O + GSSG \quad (2)$$

GSHP contains one selenocysteine residue per subunit: SeCys-35 (Forstom *et al.*, 1978; Epp *et al.*, 1983). Although the amino acid sequence of this enzyme is largely unknown, and,

TABLE 1. ACTIVE SITE HELICES

Enzyme	Active Site Residue	Helix	Reference
Rhodanese	Cys-247	251–264	Ploegman *et al.*, 1978, 1979
Glutathione peroxidase	Seleno Cys-35	36–50	Epp *et al.*, 1983
Subtilisin	Ser-221	220–238	Wright *et al.*, 1969; Drenth *et al.*, 1971a; Kraut, 1977; Robertus *et al.*, 1972; Matthews *et al.*, 1975
Triose phosphate isomerase	His-95	95–101	Banner *et al.*, 1975; Phillips *et al.*, 1977; Alber *et al.*, 1981
Glyceraldehyde-3-phosphate dehydrogenase	Cys-149	149–162	Moras *et al.*, 1975; Biesecker *et al.*, 1977
p-Hydroxybenzoate hydroxylase	isoalloxazine of FAD	298–318	Wierenga *et al.*, 1979; Weijer *et al.*, 1983
Glutathione reductase	isoalloxazine of FAD	338–354	Pai and Schulz, 1983; Schulz *et al.*, 1978
Glutathione reductase	Cys-58–Cys-63	63–80	Pai and Schulz, 1983; Schulz *et al.*, 1978
Thioredoxin	Cys-14–Cys-17	14–28	Holmgren *et al.*, 1975; Söderberg *et al.*, 1978
Papain	Cys-25	24–42	Drenth *et al.*, 1971b, 1975, 1976.

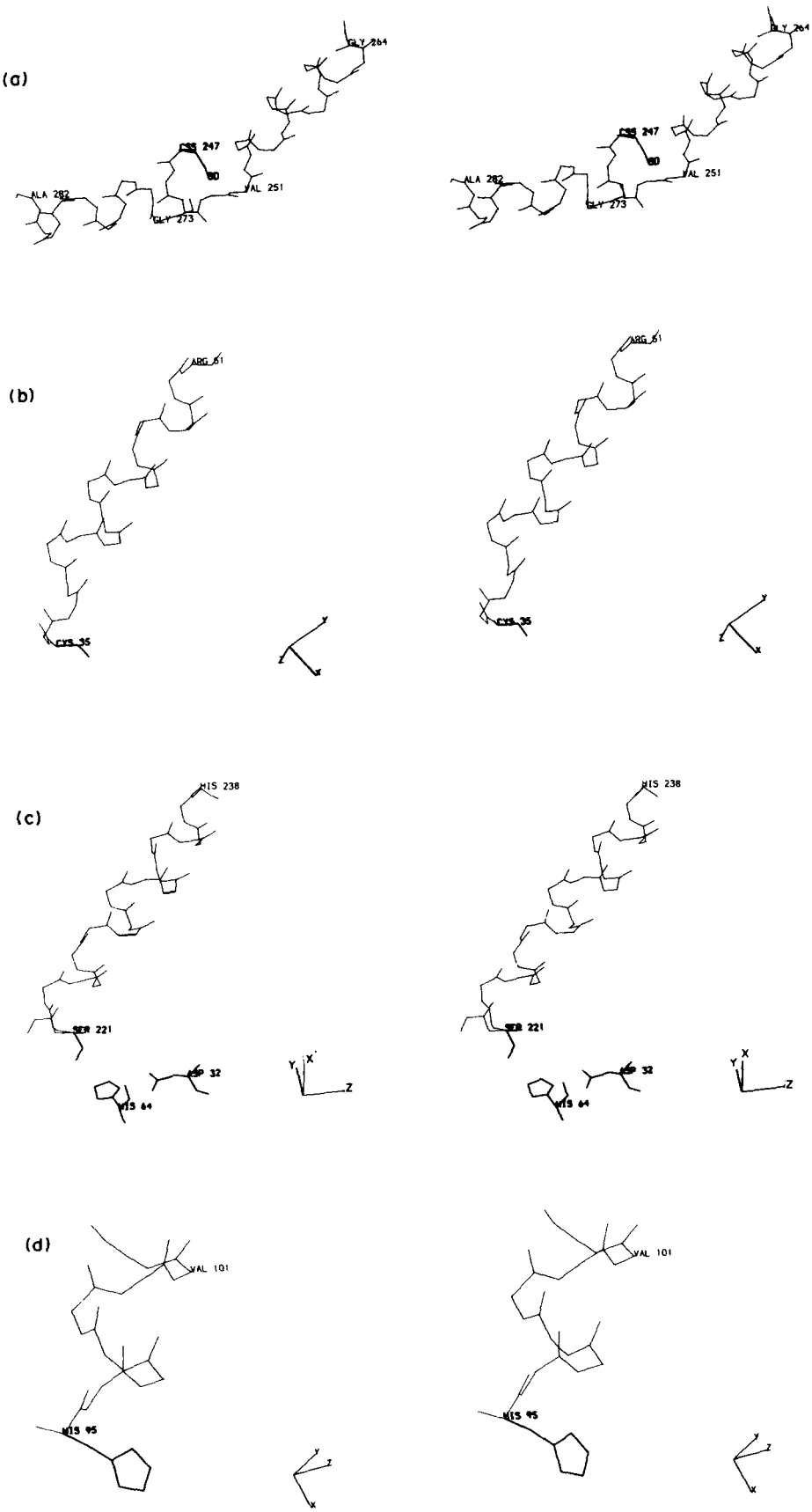

(a)
GLY 264
CYS 247
SG
VAL 251
ALA 282
GLY 273
(b)
ARG 61
CYS 35
(c)
HIS 238
SER 221
ASP 32
HIS 64
(d)
VAL 101
HIS 95

FIG. 3.

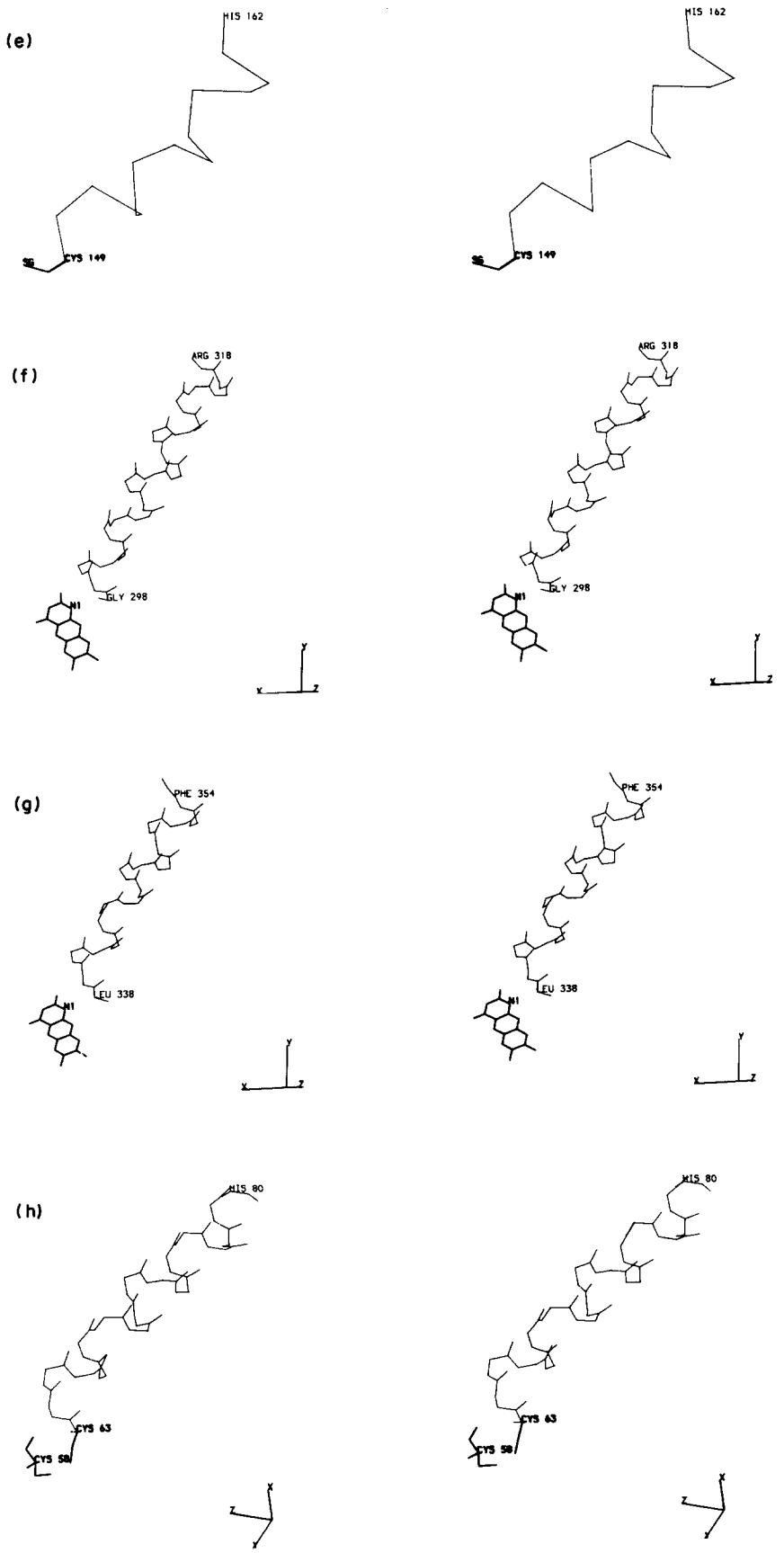

(e)
HIS 162
CYS 149
HIS 162
CYS 149
(f)
ARG 318
GLY 298
N1
ARG 318
GLY 298
N1
(g)
PHE 354
LEU 338
N1
PHE 354
LEU 338
N1
(h)
HIS 80
CYS 63
CYS 58
HIS 80
CYS 63
CYS 58

FIG. 3.

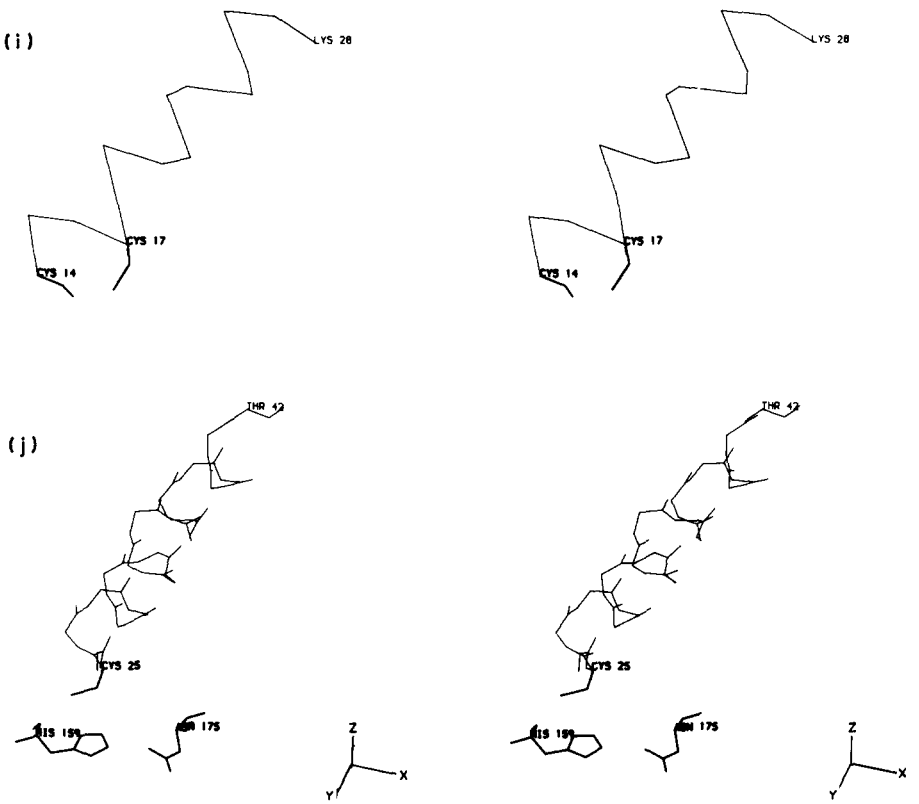


FIG. 3. Active site helices in proteins. The coordinates used came from the Protein Data Bank (Bernstein *et al.*, 1977) except when indicated otherwise. The stereo-figures in this Fig. and in Fig. 6, were made by using the interactive graphics program GUIDE (Brandenburg *et al.*, 1981).

(a) Bovine liver rhodanese (Ploegman *et al.*, 1978, 1979). The thiol group of the essential catalytic residue, Cys-247, lies close to the N-terminus of helix 251–264 (αD'). It is also on the axis of helix 273–283 (αE') which is "extended" by the peptide units formed by residues 247–249.

(b) Glutathione peroxidase from bovine erythrocytes (Epp *et al.*, 1983). Selenocysteine 35 contains a selenolate anion during the catalytic cycle [3] and it is situated near the N-terminus of helix 37–51 (coordinates were kindly provided by Drs. Epp, Ladenstein and Huber).

(c) Subtilisin (Wright *et al.*, 1969; Drenth *et al.*, 1971a). The catalytic triad Ser-221, His-64 and Asp-32 near the N-terminus of helix 220–238.

(d) Chicken triose-phosphate isomerase (Banner *et al.*, 1975; Phillips *et al.*, 1977; Alber *et al.*, 1981): active site helix 95–101 and catalytic residue His-95.

(e) Lobster glyceraldehyde-3-phosphate dehydrogenase (Moras *et al.*, 1975); the essential S^{γ} of Cys-149 lies close to and virtually on the axis of its active site helix.

(f) *p*-Hydroxybenzoate hydroxylase from *Pseudomonas fluorescens* (Wierenga *et al.*, 1979; Weijer *et al.*, 1983). In the reduced state, the N1-atom of the isoalloxazine ring of the bound FAD is negatively charged (Entsch *et al.*, 1976) and is close to the N-terminus of helix 298–318.

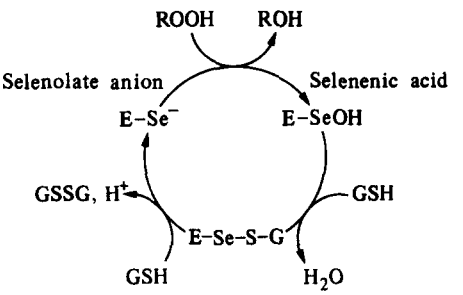
(g) Human erythrocyte glutathione reductase (Pai and Schulz, 1983; Schulz *et al.*, 1978). Just like in *p*-hydroxybenzoate hydroxylase (Fig. 3f) the N-1 atom of the isoalloxazine ring is close to an active site helix.

(h) Human erythrocyte glutathione reductase (Pai and Schulz, 1983; Schulz *et al.*, 1978): the reactive disulphide bridge is near the N-terminus of helix 63–80.

(i) Thioredoxin from bacteriophage T_4 (Söderberg *et al.*, 1978): just like in glutathione reductase this reactive disulphide bridge is very close to an active site helix (coordinates were kindly provided by Drs. Söderberg, Eklund and Brändén).

(j) The essential thiol group of papain (Drenth *et al.*, 1971b, 1975, 1976), close to the N-terminus of helix 25–42. The catalytic triad in this protease consists of Cys-25, His-159 and Asn-175.

hence, several details of the catalytic steps remain uncertain, the hypothetical mechanism of Epp *et al.* (1983) shows that the selenolate anion is crucial for the functioning of this enzyme:

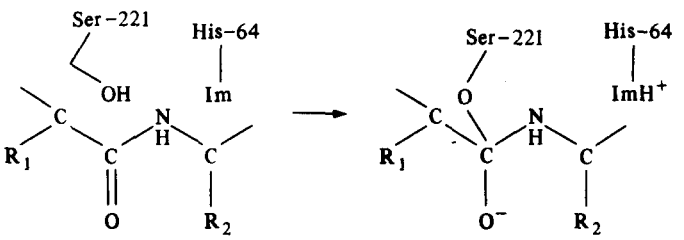


(3)

It appears that this essential anion is located near the N-terminus of the long helix α_1 [Epp *et al.* and Fig. 3(b)], and it is well possible that the electrical field of this helix will stabilize the active site selenolate and enhance the nucleophilic reactivity of this unusual catalytic group.

(c) *Subtilisin*

Subtilisin is a well characterized protease which contains three essential residues in its active centre: Ser-221, His-64 and Asp-32 (Wright *et al.*, 1969; Drenth *et al.*, 1971a). The hydroxyl group of Ser-221 is the nucleophilic agent in the first step of the catalytic process:



(4)

whereby the proton of the hydroxyl group is transferred to the imidazole ring of His-64 (Kraut, 1977). Ser-221 resides at the N-terminus of a long α-helix [Fig. 3(c)] which exerts an electrical effect such that transfer of the proton of its hydroxyl group in the direction of His-64 is promoted (Hol *et al.*, 1978). In addition, the negatively charged carbonyl oxygen of the tetrahedral intermediate shown in reaction (4) is hydrogen bonded to the NH group of Ser-221 (Robertus *et al.*, 1972; Matthews *et al.*, 1975). This arrangement in which the field of the α-helix can optimally interact with the charge of the intermediate. It appears therefore that α-helix 220–238 in subtilisin contributes in two ways to the efficiency of the catalytic action of this enzyme.

One problem remains, however. The trypsin-related serine proteases are also efficient catalysts and have spatially a virtually identical arrangement of a serine, histidine, aspartic acid triad in their catalytic centre: the classic example of convergent evolution at the atomic level (Kraut, 1977; Robertus *et al.*, 1972). In these proteases no helix occurs near the active site and one may therefore wonder if the active site helix of subtilisin is of any significance for the catalytic process. This question is unresolved as yet, but the following points can be made:

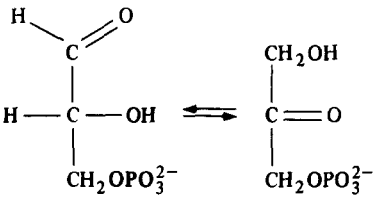
(i) in subtilisin, histidine-64 is also located at the N-terminus of a helix (Wright *et al.*, 1969; Drenth *et al.*, 1971a), and the field of the "Ser-221-helix" may be required to overcome the unfavourable effect of the "His-64-helix";

(ii) in the trypsin-related proteases the electrical field on the line O^{γ} (Ser-221)$\cdots N^{\epsilon 2}$ (His-64) due to backbone and uncharged side chain atoms is quite similar to the field of these atoms in subtilisin (Johannin, 1979). This points to a conservation of the direction of this field in all serine proteases. Great uncertainty exists, however, as to the effect of charged residues on the electrical field along this line in the active site. Johannin (1979) finds very large variations of this field for closely related serine proteases, but has not taken any dielectric screening effects for the fully exposed charges into account. Dielectric screening of fully exposed residues reduces electric

effects of these residues inside the protein by one to two orders of magnitude (Rees, 1980). Therefore, the two groups of serine proteases (subtilisin-related and trypsin-related) might not only have converged to the same spatial configuration of the three catalytic residues, but also to the same electrical field along the line which a crucial proton has to traverse in the first step of the catalytic process.

(d) *Triose phosphate isomerase*

Triose phosphate isomerase (TIM) is a well-known glycolytic enzyme which catalyzes, very efficiently, the interconversion of D-glyceraldehyde-3-phosphate and dihydroxyacetone phosphate:



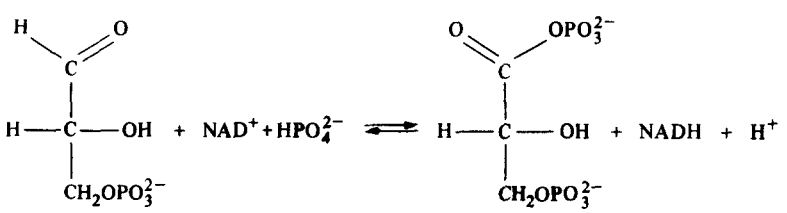
(5)

From the three-dimensional structure of this enzyme (Banner *et al.*, 1975; Phillips *et al.*, 1977; Alber *et al.*, 1981), it appeared that the imidazole ring of His-95 is equidistant from the carbonyl oxygen and the hydroxyl oxygen of the substrate glyceraldehyde-3-phosphate, so that it can shuttle a proton between these two atoms. As shown in Fig. 3(d), this histidine side chain lies at the N-terminus of a short α-helix. When discussing the mechanism of action of this enzyme, Fersht (1977, pp. 338–339) noted that the precise role of His-95 is not yet clear as its ionization is not seen in the pH-dependence of k_{cat}. It may well be that the location of this imidazole ring squarely at the N-terminus of an α-helix may have lowered the pK of this residue considerably. Hence, His-95 does not affect the bell-shaped curve of k_{cat}/K_M *versus* pH governed by pK_a values of 6 and 9; the lower pK_a value being the pK_a of the phosphate group of the substrate (Fersht, 1977; Hartmann *et al.*, 1975) and not of this special imidazole ring.

It may be mentioned here that the remarkable, symmetrical, eight-stranded α/β-barrel of TIM also occurs in four other enzymes (Muirhead, 1983): pyruvate kinase (Stuart *et al.*, 1979), 2-keto-3-deoxy-6-phosphogluconate aldolase (Mavridis *et al.*, 1982), taka-amylase A (Matssuura *et al.*, 1980) and glycolate oxidase (Lindquist and Branden, 1980). Although these structurally related enzymes have quite different catalytic functions, the active site corresponds in all cases roughly in position with the catalytic centre of TIM (Muirhead, 1983; Lebodia *et al.*, 1982). Therefore, it is to be expected that higher resolution studies will reveal that an α-helix dipole plays a role in the catalytic mechanism of one or more of the other members of this family of enzymes.

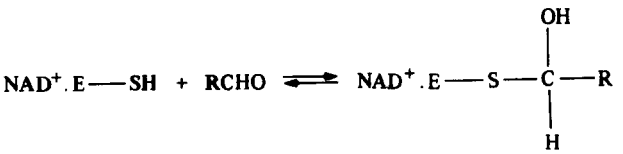
(e) *Glyceraldehyde phosphate dehydrogenase*

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) catalyzes the reversible oxidative phosphorylation of glyceraldehyde-3-phosphate to 1,3-diphosphoglycerate, using NAD^+ as coenzyme:



(6)

The reaction pathway consists of series of reactions (see e.g. Fersht, 1977, p. 298) which need not be discussed here in detail. The enzyme has a reactive cysteine residue which is crucial for the catalytic process, as is clear from the first reactions step:



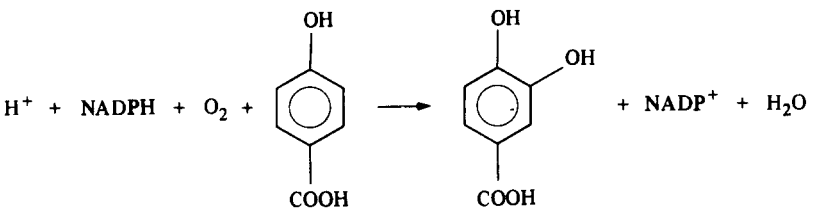
(7)

As shown in Fig. 3(e), this essential Cys-149 is located at the N-terminus of an α-helix (Moras *et al.*, 1975; Biesecker *et al.*, 1977) which stabilizes the thiolate anion relative to the neutral form of this side chain and consequently enhances its nucleophilicity.

This situation resembles the arrangement of the essential Cys-25 in papain which is also located at the N-terminus of an α-helix [Section III.2.(j)] and also forms a covalent bond with a carbonyl carbon atom during an initial catalytic step. As the three-dimensional structures of the sulphydryl protease papain and the glycolytic enzyme GAPDH are entirely different, it appears that here another example of convergent evolution in enzyme catalysis is observed. During the evolutionary process, it has been discovered at least two times, that a sulphydryl group situated precisely at the N-terminus of an α-helix has special catalytic properties.

(f) p-*Hydroxybenzoate hydroxylase*

p-Hydroxybenzoate hydroxylase (PHBH) is the best studied enzyme of a group of NAD(P)H-dependent flavin-hydroxylases (Entsch *et al.*, 1976; Wierenga *et al.*, 1979; Weijer *et al.*, 1983) which catalyzes the following reaction:



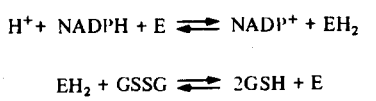
(8)

The holoenzyme contains the coenzyme FAD. The overall reaction consists of two essential steps. First, the reduction of FAD by NADPH, and, second, the hydroxylation of the substrate. The spectral properties of the reduced FAD in the holoenzyme–substrate complex favoured the interpretation that the reduced flavin is bound in the anionic form with a negative charge at N1 of the isoalloxazine ring system (Entsch *et al.*, 1976). This corresponds well with the three-dimensional structure of the holoenzyme since the O2-atom of the flavin ring makes a hydrogen bond with the N-terminus of helix H5 [Weijer *et al.*, 1983 and Fig. 3(f)]. Thus, the dipole of this α-helix interacts favourably with a negative charge on the region N1–O2 of the flavin ring.

PHBH is presently the only flavin-monooxygenase with known three-dimensional structure. It remains therefore to be seen if other members of this group of NADP-dependent flavin-hydroxylases, such as salicylate hydroxylase and melilotate hydroxylase (Flashner and Massey, 1974), contain a similar active site helix. The same question remains to be answered for the important pteridine-dependent hydroxylases phenylalanine hydroxylase and tyrosine hydroxylase (Massey and Hemmerich, 1975). For another enzyme, glutathione reductase which shares catalytic features with PHBH, the three-dimensional structure is known at high resolution and there, indeed, a similar active site helix is observed, to be discussed in the next section.

(g) *Glutathione reductase: the "reduced flavin helix"*

Glutathione reductase (GTHR) is important for a broad range of cellular activities, but its prime function is to keep the concentration of glutathione (GSH) high and that of its oxidized form (GSSG) low (Pai and Schulz, 1983). This ubiquitous FAD-containing enzyme catalyzes the reactions:



(9)

The three-dimensional structure of the enzyme from human erythrocytes is known in detail (Pai and Schulz, 1983; Schulz *et al.*, 1978). The N-terminus of α-helix 338–354 is near to the N1 and O2 atoms of the flavin ring [Pai and Schulz, 1983 and Fig. 3(g)] and this helix dipole interacts favourably with the negative charge which occurs in this region in anionic flavin states (Müller *et al.*, 1970). In this respect GTHR is very similar to PHBH. A comparison of the three-dimensional structures of the FAD-binding domains of these two enzymes reveals a close structural relationship for 110 residues with very little sequence homology (Wierenga *et al.*, 1983). The authors conclude, nevertheless, that these structural and functional similarities are the result of divergent evolution.

Recently, the amino acid sequences of the flavoproteins mercuric reductase and lipoamide dehydrogenase have been reported (Fox and Walsh, 1983; Brown *et al.*, 1983; Stephens *et al.*, 1983). The primary structures show extensive homology with glutathione reductase and with each other. It appears therefore very likely that in four flavin enzymes—PHBH, GTHR, mercuric reductase and lipoamide dehydrogenase—a negative charge which develops after reduction of the isolloxazine ring is stabilized by an α-helix dipole.

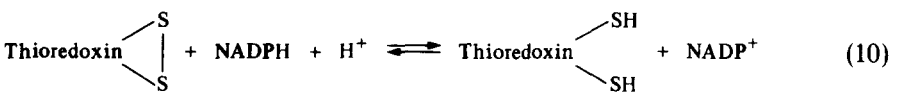
(h) *Glutathione reductase: the "reactive disulphide helix"*

In glutathione reductase a second active site helix occurs in addition to the helix described in the previous section. This is helix 63–80 containing the reactive disulphide Cys-58–Cys-63 near its N-terminus (Pai and Schulz, 1983). During the complex catalytic mechanism (Pai and Schulz, 1983), Cys-63 forms a thiolate anion which is stabilized by the helix dipole. A stereopicture of this special disulphide bridge and its helix in GTHR is given in Fig. 3(h).

In view of the strong sequence homology of GTHR with mercuric reductase and lipoamide dehydrogenase (Fox and Walsh, 1983; Brown *et al.*, 1983; Stephens *et al.*, 1983), almost certainly a similar "disulphide helix" occurs in the latter two enzymes as well.

(i) *Thioredoxin*

Thioredoxin is a small 12 K dalton protein which shuttles reducing power from NADPH to ribonucleotide reductase—a crucial enzyme in the synthesis of deoxyribonucleotides. One of the reactions in which thioredoxin is a substrate, is catalyzed by thioredoxin reductase (see e.g. Stryer, 1981, p. 525):



(10)

The reactive disulphide bridge of thioredoxin is, just like in GTHR, located at the N-terminus of an α-helix [Holmgren *et al.*, 1975; Söderberg *et al.*, 1978, Fig. 3(i)]. This may be another example of convergent evolution as the three-dimensional structures as well as the biochemical functions of GTHR and thioredoxin are very different indeed. Interestingly, it appears that in both proteins the "first" cysteine residue, i.e. Cys-58 in GTHR and Cys-35 in thioredoxin, is the most reactive of the two (Pai and Schulz, 1983; Kallis and Holmgren, 1980).

(j) *Papain and related sulphydryl proteases*

The effect of the electrical field of the helix dipole on the enzymatic action of papain and the related sulphydryl protease actinidin has been investigated in considerable detail by Van Duijnen and coworkers (Broer *et al.*, 1976; Van Duijnen *et al.*, 1979, 1980; Thole and Van Duijnen, 1983). They used large scale, *ab initio*, quantum mechanical calculations. The essential active site residues in papain are Cys-25, His-159 and Asn-175 (Drenth *et al.*, 1971b). The disposition of this catalytic triad with respect to the active site helix 24–42 is shown in Fig. 3(j). The mechanism of papain has been described by Drenth and coworkers (Drenth *et al.*, 1975) and is schematically presented in Fig. 4. The point to focus on in the present context is the existence of the ion pair $S^- \cdots Im^+ H$ at pH values as low as 6 (Drenth *et al.*, 1975). Van Duijnen and coworkers first found that the effect of the hydrogen bond between imidazole-159 and Asn-175 upon ion-pair formation was negligible (Broer *et al.*, 1976).

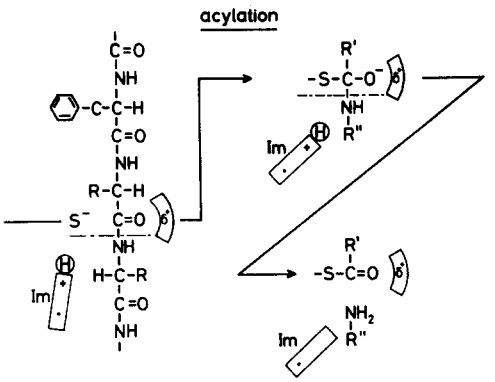


FIG. 4. the acylation step of the sulphydryl protease papain. The nucleophile, S^- of Cys-25, attacks the carbonyl carbon of the peptide bond to be split. The tetrahedral intermediate formed is stabilized by two hydrogen bonds ($\delta+$) with the negatively charged oxygen atom. Next, the hydrogen is transferred from the active site imidazolium ion (residue 159) to the NH group of the substrate and the peptide bond is split. In the deacylation, a water molecule replaces the product NH_2-R and the reaction steps occur in reverse order. (Drawing kindly provided by Dr. J. Drenth.)

Consequently, a different explanation for the existence of the ion-pair had to be discovered. It appeared that upon including the dipole field of the α-helix 24–42 [Fig. 5(a)] into the quantum mechanical calculations, the energetics changed considerably and an ion-pair (Van Duijnen *et al.*, 1979, 1980) is at least as favourable as the uncharged situation [Fig. 5(b)].

A second, even more sophisticated, investigation pertaining to the same question is a study

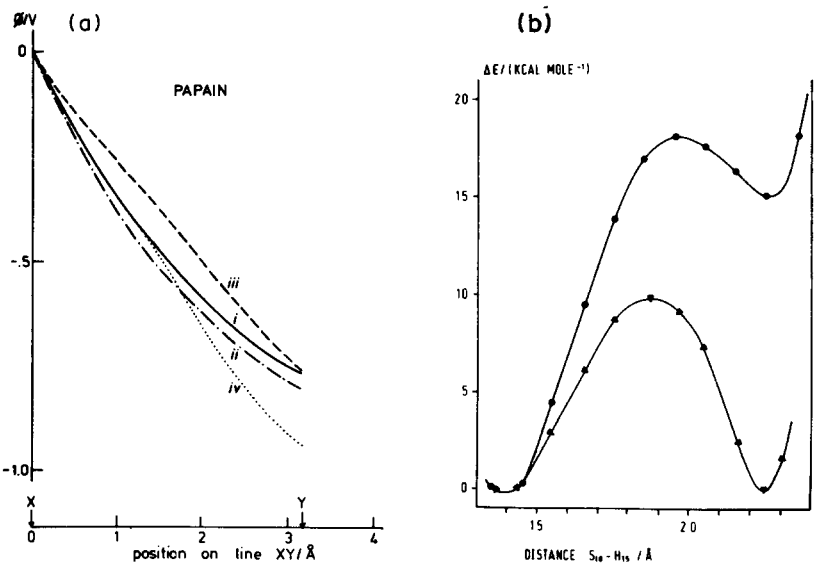


FIG. 5. The effect of the active site helix of papain.

(a) Electrostatic potential along the line S^{γ} of Cys-25 (indicated by X) to $N^{\delta 1}$ of His-159 (indicated by Y) in the active centre of papain. The different potential curves are due to: (i) backbone of the active site helix (24–43); (ii) all atoms of the active site helix, except for the side chain of Cys-25; (iii) backbone of all papain residues; (iv) all papain atoms, except for the side chains of Cys-25, His-159 and Asn-175. Partial atomic charges for calculation with complete residues were taken from Poland and Scheraga (1967). In order to simulate the effect of solvation, the charges of the exposed Asp, Glu, Arg and Lys side chains were reduced to half their normal values. For calculations using the backbone only, the charges of Fig. 1 were used. The dielectric constant was taken as $\varepsilon = 1 + (R - R_A)/R$, R being the distance between a point on the line XY and a contributing atom; $R_A = 1.5$ Å; for $R < R_A$: $\varepsilon = 1$. The potential at X was taken as (arbitrary) zero point. (Fig. from Hol *et al.*, 1978.)

(b) The total energy of a model system of the three essential residues in the active site of papain on the line $S^{\gamma} \ldots N^{\delta 1}$, i.e. the line XY in Fig. 5(a). The energy of the first point, at a distance of 1.35 Å from the S^{γ} atom, is taken as arbitrary zero point. The active site model consists of methane thiol, imidazole and formaldehyde. Upper curve: no helix field. Lower curve: helix field included. (Fig. from Van Duijnen *et al.*, 1979.)

of Thole and Van Duijnen (1983) on actinidin. This is another SH-protease with essentially an identical structure as papain (Baker, 1980). Here, the authors incorporate, in addition to the helix dipole field, also the atomic polarizabilities of the protein atoms surrounding the active site in their quantum mechanical calculations. The effect of the helix dipole is then less than in their investigation on papain. However, the two studies cannot be directly compared as the sulphur–nitrogen distance is smaller in the actinidin study which means a smaller dipole of the ion pair and thus a smaller energy gain in the field of the α-helix. Nevertheless, a distinct effect of the α-helix clearly remains in the active site of actinidin.

A second effect of the helix dipole upon the catalysis by papain may be stabilization of the negative charge on the carbonyl oxygen in the tetrahedral intermediate (Fig. 4). One of the NH peptide hydrogen bond donors to this charged oxygen is provided by Cys-25: at the very end of the active site helix 24–42 (Drenth *et al.*, 1976). A similar favourable interaction between an α-helix dipole and a charged atom of an intermediate in the catalytic process is observed in subtilisin [Section III.2.(c)], although otherwise these two proteases have completely different structures.

3. *Conclusion*

Although many mechanistic and theoretical points are not definitely settled, it can be safely concluded from the examples given above, that numerous active sites, in a wide variety of enzymes, are more effective catalytic centres due to the electrical field of the helix dipole.

IV. ANION-BINDING BY HELIX DIPOLES

1. *Introduction*

Many proteins interact with charged molecules; in particular phosphate groups are frequently bound. This is the consequence of the vast array of phosphate-containing molecules belonging to the most essential components of living cells: DNA, RNA, ATP and several other nucleotides, phosphorylated substrates in glycolysis and membrane transport, thiamine pyrophosphate, pyridoxal phosphate, phospholipids, etc. For about twenty proteins which bind *low* molecular weight phosphate-containing ligands, the three-dimensional structure has been determined by X-ray diffraction at reasonably high resolution with the positions of bound phosphate groups elucidated (Tables 2 and 3). For these low molecular weight ligands the α-helix dipole is very often, i.e. in about 60% of the cases, employed for binding phosphate moieties (Tables 2 and 3). For eleven proteins, binding *high* molecular weight phosphate-containing ligands, the three-dimensional structure is also known (Table 4). It appears that for these cases the α-helix is *never* involved in binding the phosphates of DNA or RNA. The reason for this distinct difference in the use of the α-helix dipole for binding high and low molecular weight phosphate-containing molecules does not seem to be obvious and it is probably best to refrain here from speculations.

2. *Phosphate-Binding Helices in Proteins*

About 20 "phosphate-binding helices" are known at present (Table 2). The phosphate-containing molecules bound vary considerably in size and shape as do the functions and structures of the proteins employing the helix dipole in binding these charged molecules. Among the proteins we find small electron-transport proteins as flavodoxins, a group of dehydrogenases, a ribosomal elongation factor, and phosphorylase, the largest polypeptide chain with known tertiary structure to date. For those instances where coordinates were available, stereo-pictures of helices binding phosphate groups are presented in Fig. 6.

In the majority of cases the phosphate group is bound between the N-terminus of an α-helix and positively charged side chains (Table 2). A schematic representation of this binding mode is given in Fig. 7. Apparently, this is a quite favourable arrangement for the phosphate group. In this manner, the protein is probably "solvating" this charged group better than solvent molecules and counter ions are able to do. The interactions of the half positive charge of the α-helix dipole and the full positive charge of the side chains with the

TABLE 2. PROTEINS BINDING PHOSPHATE MOIETIES OF LOW MOLECULAR WEIGHT LIGANDS BY USING AN α-HELIX DIPOLE

Protein	Phosphate-Containing Ligand Bound	Phosphate Binding Helix	Positively Charged Residue(s) also Involved in Phosphate Binding?	Reference
Alcohol dehydrogenase	NAD	201–215	yes	Rossmann *et al.*, 1975; Eklund *et al.*, 1981
Lactate dehydrogenase	NAD	30–42	yes	Rossmann *et al.*, 1975; Grau *et al.*, 1981
Malate dehydrogenase	NAD	15–26	?	Rossmann *et al.*, 1975; Hill *et al.*, 1972
Glyceraldehyde-3-phosphate dehydrogenase	NAD	9–22	no	Moras *et al.*, 1975; Biesecker *et al.*, 1977; Rossmann *et al.*, 1975; Wierenga *et al.*, 1984
Flavodoxin	FMN	10–26	no	Watenpaugh *et al.*, 1973; Smith *et al.*, 1977, 1983; Ludwig *et al.*, 1982
p-Hydroxybenzoate hydroxylase	FAD	11–25	yes	Wierenga *et al.*, 1979, 1984; Weijer *et al.*, 1983
Glutathione reductase	FAD	29–42	no	Wierenga *et al.*, 1983, 1984; Schulz *et al.*, 1982
Glutathione reductase	NADP	196–210	no	Pai and Schulz, 1983; Wierenga *et al.*, 1983, 1984
Dihydrofolate reductase	NADP	42–48 and 99–106	yes	Wierenga *et al.*, 1984; Filman *et al.*, 1982; Volz *et al.*, 1972
Adenylate kinase	ATP	23–30	yes	Pai *et al.*, 1977
Elongation factor Tu	GTP	227–242	no?	Duisterwinkel *et al.*, 1984
Aspartate amino transferase	pyridoxal phosphate	107–124	yes	Ford *et al.*, 1980
Phosphorylase	pyridoxal phosphate	675–682	yes	Johnson *et al.*, 1980
Phosphorylase	glucose-1-phosphate	134–140	yes	Johnson *et al.*, 1980
Triose phosphate isomerase	dihydroxy-acetone-P	232–237	no	Alber *et al.*, 1981
Phospho glycerate kinase	ATP	372–380	? (Mg^{2+})	Watson *et al.*, 1982
Phospho glycerate kinase	3-phospho glycerate	393–401	yes	Watson *et al.*, 1982
Phospho fructo kinase	ADP (Site B)	"helix 5"	? (Mg^{2+})	Evans and Hudson, 1979
Glyceraldehyde-3-phosphate dehydrogenase	glyceraldehyde-3-phosphate	149–162*	yes	Dalziel *et al.*, 1981
Phosphorylase	NAD (I-site)	614–631†	no	Stura *et al.*, 1983
Aspartate carbamoyltransferase pyrophosphate, (catalytic chain)	ATP, CTP, GTP	52–65‡	yes	Honzatko and Lipscomb, 1982; Honzatko *et al.*, 1982

*The phosphate moiety of the substrate is bound by the side chains of Ser 148 and Thr 150 near to the N-terminus of helix 149–162.
†The pyrophosphate is situated at 5 Å from the N-terminus of the helix.
‡The pyrophosphate moiety may be positioned quite far off the helix axis (K. Volz, personal communication).

TABLE 3. PROTEINS BINDING PHOSPHATE MOIETIES OF LOW MOLECULAR WEIGHT LIGANDS WITHOUT THE USE OF AN α-HELIX DIPOLE

Protein	Phosphate Containing Ligand Bound	Reference
Hexokinase	8-Br ADP	Schoham and Staitz, 1980
Glyceraldehyde-3-phosphate dehydrogenase	Glyceraldehyde-3-phosphate	Dalziel *et al.*, 1981
Phospho fructo kinase	Fructose-6-phosphate	Evans and Hudson, 1979
Phospho fructo kinase	ADP (Site C)	Evans and Hudson, 1979
Phosphorylase	NADP	Jenkins *et al.*, 1981
Tyrosyl-tRNA synthetase	Tyrosyladenylate*	Bhat *et al.*, 1982
Catabolic gene activator protein	cAMP	McKay *et al.*, 1982
Dihydrofolate reductase	NADP (2′ phosphate)	Filman *et al.*, 1982; Volz *et al.*, 1972
Glutathione reductase	NADP (2′ phosphate)	Pai and Schulz, 1983; Wierenga *et al.*, 1983
Phosphorylase	AMP (I-site)	Stura *et al.*, 1983
Phosphorylase	AMP, NAD (N-site)	Stura *et al.*, 1983
Aspartate carbamoyl transferase (regulatory chain)	CTP	Honzatko and Lipscomb, 1982; Honzatko *et al.*, 1982
Citrate synthase	Coenzyme A	Remington *et al.*, 1982

*In this case the phosphate moiety is about 6 Å removed from a helix N-terminus.

TABLE 4. PROTEINS BINDING PHOSPHATE-CONTAINING MOLECULES OF HIGH MOLECULAR WEIGHT

Protein	Molecule to be Bound	Footnote	Reference
Pancreatic ribonuclease	RNA	*	Wodak *et al.*, 1977
Microbial ribonuclease	RNA	*	Mauguen *et al.*, 1982; Heinemann and Saenger, 1982; Nakamura *et al.*, 1982
Staphylococcal nuclease	RNA and DNA	*	Coppom *et al.*, 1979
DNA unwinding protein	DNA	†	McPherson *et al.*, 1979
Cro-repressor	DNA	‡	Anderson *et al.*, 1981; Ohlendorf *et al.*, 1982; Steitz *et al.*, 1982
λ-repressor	DNA	‡	Pabo and Lewis, 1982; Sauer *et al.*, 1982
Catabolic gene activator protein	DNA	‡	Steitz *et al.*, 1982; McKay and Steitz, 1981; Salemme, 1982
Southern bean mosaic virus	RNA	§	Abad-Zapatero *et al.*, 1980
Tomato bushy stunt virus	RNA	§	Harrison *et al.*, 1978
Satellite tobacco necrosis virus	RNA	§	Liljas *et al.*, 1982
Tobacco mosaic virus	RNA	‖	Stubbs *et al.*, 1977; Bloomer *et al.*, 1978

The column "footnote" indicates the evidence for the lack of involvement of a helix dipole in the interaction with nucleic acids:

*the structure of a complex with a low molecular weight nucleic acid analogue has been determined;

†the protein does not contain any α-helix;

‡detailed models of protein-DNA interactions have been described; for the λ-repressor a helix N-terminus might be involved in DNA-binding, however (Pabo and Lewis, 1982).

§flexible, or disordered, segments of the polypeptide chain interact very likely with the RNA phosphates via positively charged side chains;

‖in the structure determination the protein RNA interactions could be determined.

phosphate group are difficult to quantify. The helix terminus is deeper in the active site cleft and moreover, in the complex, more buried than the charged side chain. It may therefore well be that the helix-phosphate interaction is more favourable than the charge-phosphate interaction.

In a number of instances, the phosphate moiety is bound by an α-helix dipole only (Table 2). A particularly interesting case is flavodoxin where an α-helix plus a number of peptide NH groups interact with the phosphate group. The nearest charged residues are all carboxyl groups which are located predominantly at the FMN-binding "face" of the molecule. Again,

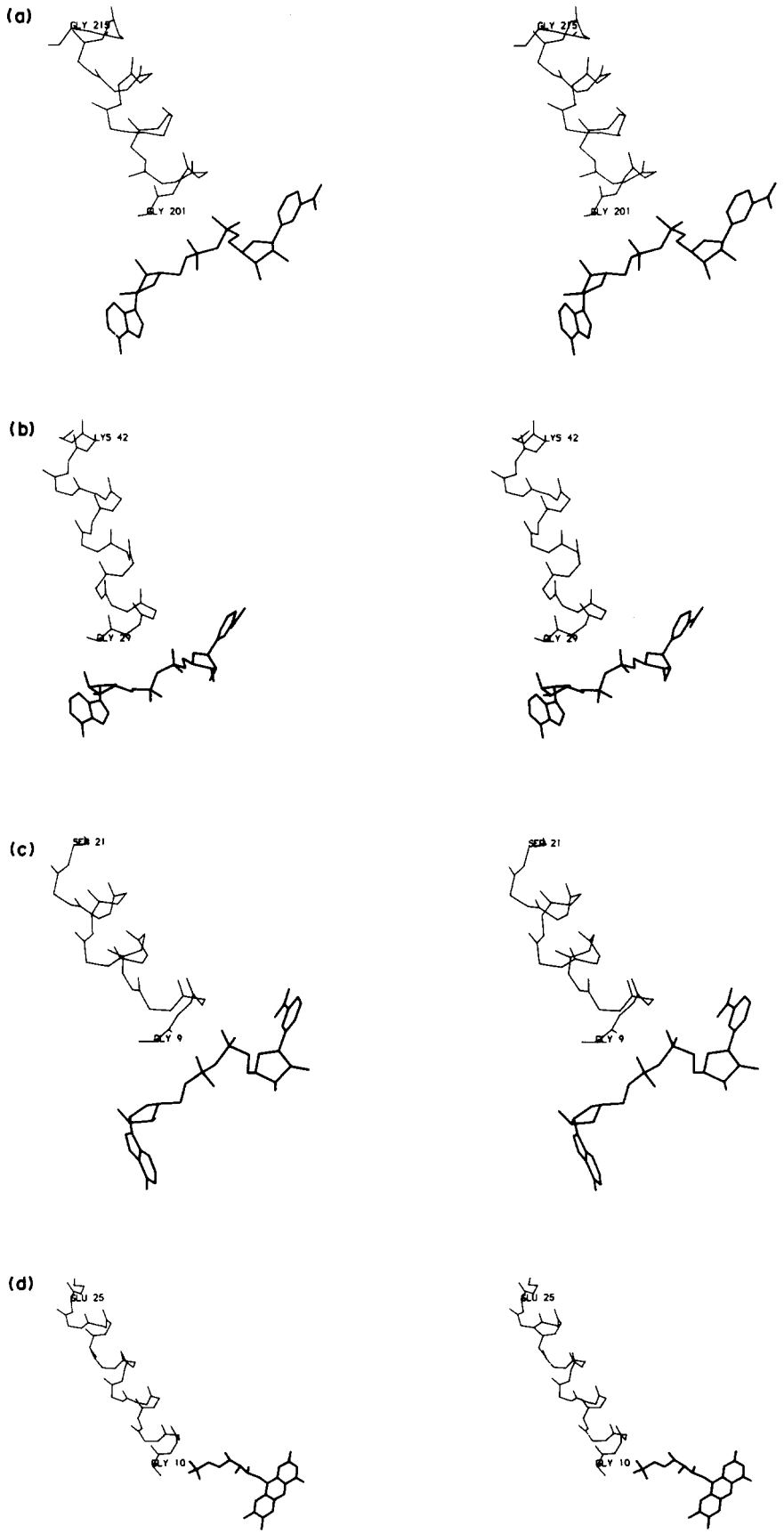

(a)
GLY 215
GLY 201
GLY 215
GLY 201
(b)
LYS 42
LYS 42
(c)
SER 21
SER 21
(d)
GLU 25
GLU 25

FIG. 6.

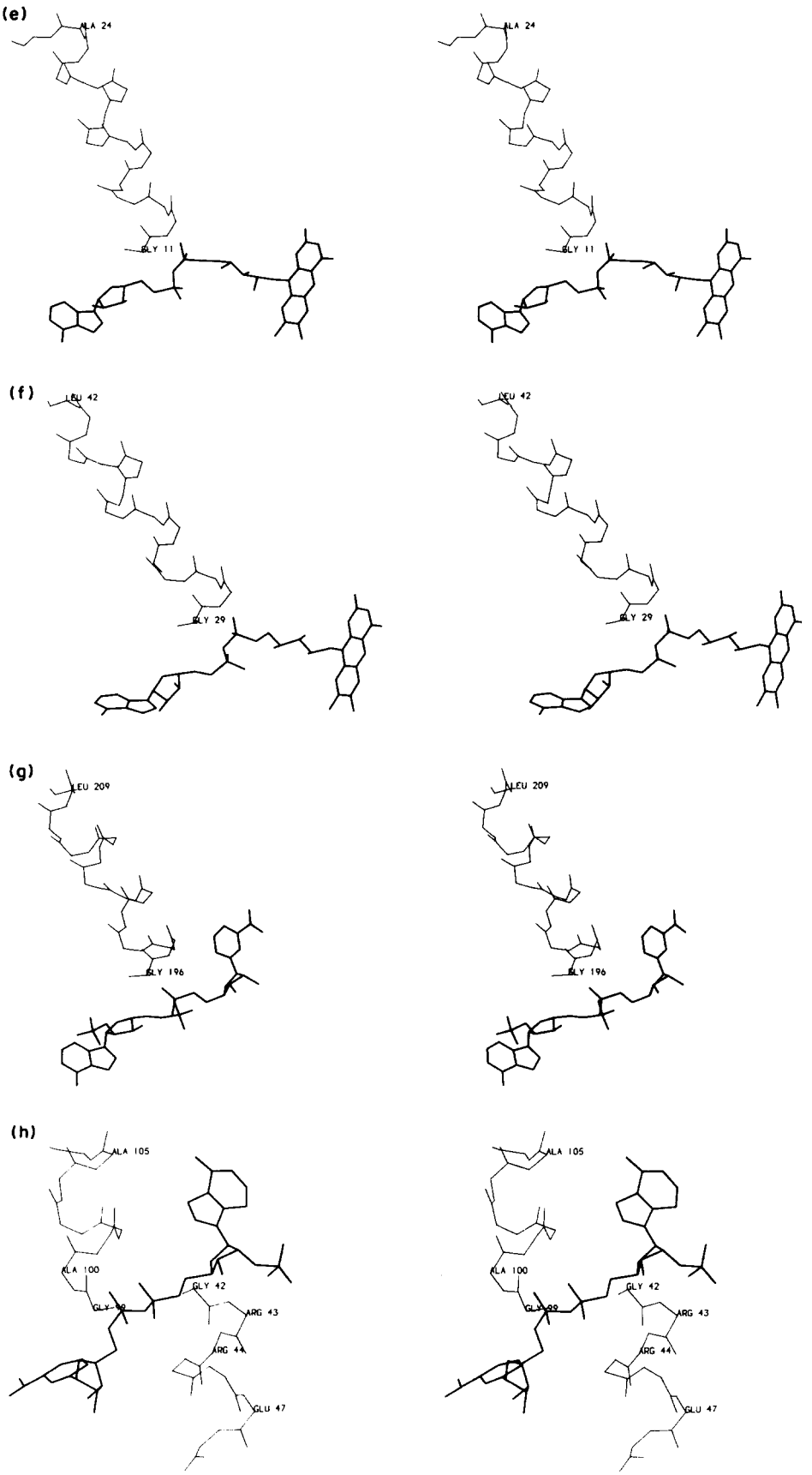

(e)
ALA 24
GLY 11
ALA 24
GLY 11
(f)
LEU 42
GLY 29
LEU 42
GLY 29
(g)
LEU 209
GLY 196
LEU 209
GLY 196
(h)
ALA 105
ALA 100
GLY 42
ARG 43
ARG 44
GLU 47
ALA 105
ALA 100
GLY 42
ARG 43
ARG 44
GLU 47

FIG. 6.

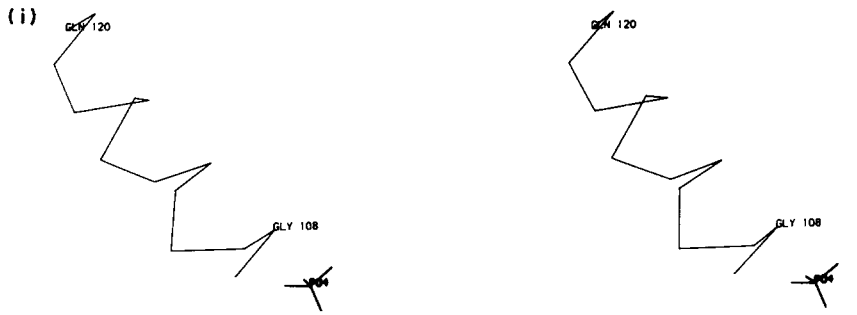


FIG. 6. Gallery of "phosphate-binding" helices. The coordinates are taken from the Protein Data Bank (Bernstein *et al.*, 1977), unless otherwise indicated.

(a) The pyrophosphate moiety of NAD bound by helix 201–215 of horse liver alcohol dehydrogenase (Eklund *et al.*, 1981). Coordinates were kindly provided by Drs. H. Eklund and C.-I. Brändén.

(b) The pyrophosphate moiety of NAD bound by helix 29–42 of dogfish lactate dehydrogenase (Grau *et al.*, 1981).

(c) The pyrophosphate moiety of NAD bound by helix 9–21 in lobster glyceraldehyde-3-phosphate dehydrogenase (Moras *et al.*, 1975).

(d) The phosphate group of FMN interacting with helix 10–25 of *Clostridium* flavodoxin (Smith *et al.*, 1977).

(e) FAD with its pyrophosphate close to helix 11–24 of p-hydroxybenzoate hydroxylase from *Pseudomonas fluorescens* (Wierenga *et al.*, 1979; Weijer *et al.*, 1983). (Coordinates provided by Drs. R. K. Wierenga and J. Drenth.)

(f) FAD with its pyrophosphate close to helix 29–42 of human erythrocyte glutathione reductase (Schulz *et al.*, 1982).

(g) The pyrophosphate group of NADP near helix 196–209 of human erythrocyte glutathione reductase (Pai and Schulz, 1983; Wierenga *et al.*, 1983). The coordinates of NADP were kindly provided by Dr. G. E. Schulz.

(h) Helices 99–105 and 42–47 of dihydrofolate reductase from *Lactobacillus casei* interacting with the pyrophosphate moiety of NADP (Filman *et al.*, 1982).

(i) The phosphate group of pyridoxal phosphate in close proximity of helix 107–120 of aspartate amino transferase from chicken mitochondria (Ford *et al.*, 1980). Coordinates kindly provided by Dr. J. N. Jansonius.

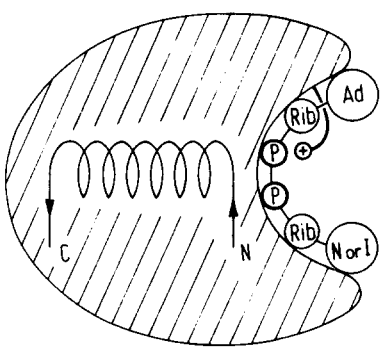


FIG. 7. Schematic drawing of dinucleotide binding by numerous proteins. The pyrophosphate moiety is close to the N-terminus of an α-helix while the negative charges are also compensated by one, or more, positively charged side chains. As shown in Table 2, several instances are known when such a positively charged residue is absent. (Fig. from Hol and Wierenga, 1984.)

the electrostatic interactions are hard to quantify, but it seems justified to conclude that the α-helix dipole is absolutely essential for binding phosphate moieties at this position to the flavodoxins.

In glutathione reductase, the charged residue nearest to the pyrophosphate moiety of FAD is an aspartic acid. In this example no positive charges occur within 4 Å of the pyrophosphate atoms and the helix dipole is presumably the only compensating factor (although Schulz *et al.*, 1982, have discussed the possibility of a metal ion near the phosphates). It appears therefore that great variability exists as to the presence of positively charged residues near the phosphate moieties bound by helix dipoles.

The position of the phosphate group with respect to the α-helix dipole is remarkably

constant (Fig. 6 and Wierenga *et al.*, 1984). This can also be nicely illustrated by considering different flavodoxin structures (Ludwig *et al.*, 1982). In *A. Nidulans* flavodoxin (Smith *et al.*, 1983) the isoalloxazine ring is quite differently oriented than in *Clostridium* and *Vulgaris* flavodoxins (Watenpaugh *et al.*, 1973; Smith *et al.*, 1977), but the positioning of the phosphate moiety with respect to the α-helix is virtually the same.

The phosphate-binding helices form often part of a βαβ-unit in proteins (Wierenga *et al.*, 1983, 1984; Rossmann *et al.*, 1975; Brändén, 1980; Wierenga and Hol, 1983). One might expect that this common feature is simply a reflection of significant similarities in the amino acid sequence in these phosphate binding βαβ-units. Table 5 contains an alignment of amino acid sequences of a number of dinucleotide binding helices and adjacent β-strands. Although a "fingerprint" appears to be present (Table 5; Wierenga and Hol, 1983), it is surprising that the amino acid sequences of these closely related folding units are so very different (Table 6). It does not seem obvious that a specific amino acid sequence is responsible for phosphate binding. The second invariant glycine, however, is particularly well conserved. This residue appears to provide space for the approaching phosphate moiety and may be an essential feature of phosphate binding helices.

The nucleotide-binding βαβ-units are usually part of larger parallel β-sheet. Brändén (1980) has suggested that the "switch points" in these parallel β-structures form crevices

TABLE 5. SEQUENCE ALIGNMENT OF DINUCLEOTIDE-BINDING βαβ-UNITS IN PROTEINS (FROM WIERENGA AND HOL, 1983)

	β	β	β	β	β	β		α	α	α	α	α	α	α	α	α	α	α	α	α	α					β	β	β	β	β	β	
	194										204										214										223	
ADH (NAD)	T	C	A	V	F	G	L	G	G	V	G	L	S	V	I	M	G	C	K	A	A	G	A	A	—	R	I	I	G	V	D	I
	△	□		□		●		●			●				□			□									□		□		⊖	
	22										32										42										52	
LDH (NAD)	K	I	T	V	V	G	V	G	A	V	G	M	A	C	A	I	S	I	L	M	K	D	L	A	D	E	V	A	L	V	D	V
	△	□		□		●		●			●				□			□									□		□		⊖	
	2										12										22										31	
GAPDH (NAD)	K	I	G	I	D	G	F	G	R	I	G	R	L	V	L	R	A	A	L	S	C	G	A	Q	—	V	V	A	V	N	D	P
	△	□		□		●		●			●				□			□									□		□		⊖	
	4										14										24										32	
PHBH (FAD)	Q	V	A	I	I	G	A	G	P	S	G	L	L	L	G	Q	L	L	H	K	A	G	I	—	—	D	*N*	V	I	L	E	R
	△	□		□		●		●			●				□			□									□		□		⊖	
	22										32										42										50	
GTHR (FAD)	D	Y	L	V	I	G	G	G	S	G	G	L	A	S	A	R	R	A	A	E	L	G	A	—	—	R	A	A	V	V	E	S
	△	□		□		●		●			●				□			□									□		□		⊖	
	189										199										209										217	
GTHR (NADP)	R	*S*	V	I	V	G	A	G	Y	I	*A*	V	E	M	A	G	I	L	S	A	L	G	S	—	—	K	*T*	S	L	M	*I*	R
	△	□		□		●		●			●				□			□									□		□		⊖	

Although these folding units have very similar structure, their sequences differ considerably (Table 6). Nevertheless, a common "fingerprint" can be distinguished (Wierenga and Hol, 1983) and is indicated by the following symbols: ● invariant glycines; ⊖ negatively charged side chain; □ usually a small and hydrophobic residue; △ hydrophylic side chain. The secondary structure is indicated above the aligned sequences.

TABLE 6. AMINO ACID SEQUENCE DISSIMILARITY AND STRUCTURAL SIMILARITY OF DINUCLEOTIDE BINDING βαβ-UNITS (FROM WIERENGA *et al.*, 1984)

	NAD ADH	NAD LDH	NAD GAPDH	FAD PHBH	FAD GTHR	NADP GTHR
ADH	*	7	5	6	7	3
LDH	1.1	*	9	3	8	5
GAPDH	1.2	1.2	*	5	7	4
PHBH	1.2	1.4	1.4	*	6	5
GTHR (FAD)	1.0	1.4	1.5	0.9	*	4
GTHR (NADP)	1.1	1.3	1.5	0.8	0.9	*

The upper half of the matrix contains the number of identical residues at equivalent positions for each pair of βαβ-units. The number of equivalent positions was in each case 27.

The lower half of the matrix gives the r.m.s. deviations of the 27 C^{α}-atoms for all βαβ-units, after optimal superposition. The most remarkable pairs are ADH/GTHR (NADP) and LDH/PHBH with both only 11% sequence identity while the C^{α}-atoms fall nevertheless within 1.1 Å resp. 1.4 Å of each other.

which are preferred binding sites of adenine and nicotinamide rings of nucleotides and dinucleotides. Wierenga *et al.* (1984) found that this certainly holds for the adenine ring which is bound in analogous positions in a variety of proteins. It may therefore be concluded that:

(i) the $\beta\alpha\beta$-unit appears to be a particularly stable structure;
(ii) the parallel β-sheet provides switch points which are preferred binding sites for ring systems ("Brändén crevices");
(iii) the α-helix dipole provides an excellent interaction with the phosphate group; and
(iv) the phosphate moiety can approach these helices closely because of a space-providing N-terminal glycine residue.

These four factors together may explain why the (di)nucleotide binding $\beta\alpha\beta$-unit is so frequently observed.

Just to be complete, it may be mentioned that in myoglobin crystals two helix dipoles from different molecules bind a sulphate ion (Phillips, 1980). Finally, it has been reported that in D-Ala–D-Ala-carboxypeptidase transpeptidase (Kelly *et al.*, 1982), the antibiotic penicillin is bound near the N-termini of helices. It has been suggested that the carboxyl group of penicillin may interact with an α-helix N-terminus. Higher resolution studies are awaited with great interest to see in detail the mode of binding of this important drug to its target enzyme.

3. *Conclusion*

From the overview above it is clear that a wide range of proteins utilize an α-helix in binding a variety of phosphate-containing molecules. It is also clear that a number of proteins employ more than one helix for this purpose. For instance, in dihydrofolate reductase two helices point their N-terminus to the same pyrophosphate group of a bound NADP molecule. In glutathione reductase two helices are involved in binding two different dinucleotides.

Very recently, it has been discovered that in a so-called "sulphate binding protein" the sulphate ion is bound internally, with not less than three helix dipoles interacting favourably with the ion (the nearest charged residue is an aspartic acid side chain!; J. Pflugrat, personal communication). This appears to be the most illustrious example of helix dipole–charge interactions in proteins to date.

One may raise the question if the binding of sulphate or phosphate ions by α-helices is not simply the result of a unique arrangement in which several oxygen atoms of the anion are very favourably positioned to interact with the NH groups of the helix terminus. If this were the case then one would expect a rather large number of short hydrogen bonds. This point has been looked into recently by Wierenga *et al.* (1984) when they studied details of eight α-helices, of six different proteins with known three-dimensional structure, interacting with dinucleotides. It was found that: (i) for *six* of the eight helices only *one* hydrogen bond was formed with the pyrophosphate moiety, while in *four* of these cases this hydrogen bond was surprisingly long, i.e. 3.3–3.4 Å; (ii) helix 42–48 in dihydrofolate reductase formed two hydrogen bonds, but of length 3.2 and 3.4 Å; and, (iii) only helix 99–106 in dihydrofolate formed several hydrogen bonds with the pyrophosphate group of NADP (it may be pointed out, however, that this was the shortest of the eight helices where the dipole–anion interactions would be the weakest). For these eight examples there appeared to be no characteristic set of a large number of strong hydrogen bonds between helix and anion, supporting the presence of a general electrostatic interaction between the dipole of the helix and the monopole of the anion.

Several enzymes also use an "active site helix" in addition to a "phosphate binding helix". Examples are PHBH and GAPDH. A fascinating example of the use of helix dipoles by enzymes is glutathione reductase which harbours, in addition to the two phosphate-binding helices mentioned in the previous paragraphs [Figs 3(g), (h)], also two active site helices [Figs 6(f), (g)]. With this example we will leave the effect of the helix dipole on the functional properties of proteins—although this subject will return in the last section of this review in connection with membrane processes—and turn now to the folding of globular proteins.

V. THE PREFERRED POSITION OF CHARGED AMINO ACIDS WITH RESPECT TO α-HELICES

A distinct asymmetry in the distribution of charged residues α-helices in globular proteins has been discovered (Crawford *et al.*, 1973; Chou and Fasman, 1974; and Table 5). Blagdon and Goodman (1975) pointed out that this was due to an interaction of the α-helix dipole with the charged residue, leading to a polar mechanism for helix initiation, which can proceed from either a C-terminus or an N-terminus. Whether this mechanism involves a specific interaction of the charged side chain group via hydrogen bonds with the backbone of a helix terminus during the folding process, or is a more general electrostatic interaction without a well-defined structural intermediate, remains to be established. In fact both effects may act simultaneously. First a kinetic effect upon the folding process may occur, followed by an electrostatic stabilization of the α-helix.

In a study on charge–charge and charge–helix dipole interactions in a large number of proteins, Wada and Nakamura (1981) also observed a favourable interaction between charged residues and helix dipoles. This implies that negatively charged residues have a tendency to cluster near N-termini of helices and positively charged residues near C-termini.

TABLE 7. FREQUENCY OF HELICAL BOUNDARY RESIDUES IN 15 PROTEINS*

	f_{hN}†		f_{hC}‡
Pro	0.212	His(+)	0.216
Asp(−)	0.207	Lys(+)	0.160
Glu(−)	0.195	Gln	0.158
Ala	0.140	Arg(+)	0.154
Trp	0.136	Cys	0.148
Thr	0.122	Met	0.143
Gln	0.116	Glu(−)	0.124
Phe	0.098	Ala	0.118
Asn	0.090	Val	0.116
Ser	0.079	Phe	0.110
Cys	0.074	Leu	0.102
Met	0.071	Asn	0.090
Tyr	0.070	Ser	0.084
Ile	0.066	Ile	0.075
Val	0.061	Asp(−)	0.054
Gly	0.060	Tyr	0.050
Lys(+)	0.057	Thr	0.045
Leu	0.056	Trp	0.045
His(+)	0.054	Gly	0.039
Arg(+)	0.038	Pro	0.000

*Taken from Table IV in Chou and Fasman (1974).

†Frequency of residues in the N-terminal helix region (comprising 3 residues).

‡Frequency of residues in the C-terminal helix region (comprising 3 residues).

The same conclusion is obtained by a study of Van der Plaats and Hol (unpublished) who looked into charge–helix dipole interactions in eleven proteins (Table 8). The energies given should only be considered as indications for "favourable" or "unfavourable" as the important screening effects of the solvent have not been taken into account. This table shows that also for charged residues which do not belong to an α-helix, a favourable interaction with α-helix dipoles exists, in agreement with Wada and Nakamura's very general analysis (Wada and Nakamura, 1981).

It is worthwhile to point out here that the distribution of charged residues along the α-helix (Table 7), shows the effect of the positive as well as the negative end of the dipole. This is different from the examples given in the previous section, where virtually only phosphate ions near N-termini of helices were discussed. This bias in the previous section is undoubtedly due to the large number of phosphate binding proteins known which, in its turn, is a result of the great importance of phosphate-containing molecules for the living cell.

TABLE 8. ELECTROSTATIC INTERACTIONS BETWEEN CHARGED SIDE CHAINS AND HELIX DIPOLES IN ELEVEN GLOBULAR PROTEINS (VAN DER PLAATS AND HOL, UNPUBLISHED RESULTS)

Protein	Number of Helices	$U_{z\text{Glu}}+U_{z\text{Asp}}$ $+U_{z\text{Lys}}+U_{z\text{Arg}}$	$U_{z\text{His}}$	$U_{z\text{Cys}}$	U_{zq}*
Alcohol dehydrogenase	12	−83	+51	−63	−95
Adenylate kinase	9	−103	+11	—	−92
Carboxypepidase A	9	−17	+105	—	+88
Flavodoxin	5	−174	—	—	−174
Glyceraldehyde-3-phosphate dehydrogenase	9	−75	+31	−29	−73
Lactate dehydrogenase	9	−129	—	—	−129
Lysozyme (HEW)	6	+70	—	—	+70
Lysozyme (T_4 phage)	9	−50	−4	—	−54
Papain	6	−101	+17	−11	−95
Subtilisin	8	−146	−4	—	−150
Rhodanese	10	−137	—	—	−137

Energies are given in kcal mol^{-1}. The dielectric constant was taken as unity, i.e. no dielectric screening was taken into account. For the helix backbone, the HN atom was generated according to standard geometry and the partial charges of Fig. 1 were used. For Asp and Glu the charge was positioned between the two carboxyl oxygen atoms; for Arg between the two N^{η}-atoms; for His between the two ring nitrogens; for Cys on S^{γ}; for Lys on N^{ζ}.

If the distance between the position of a charge and a backbone atom of a helix was smaller than 3.6 Å then the interaction between this helix residue and the charge was excluded from the calculation. The interactions between a charged residue occurring in a helix and its "own" helix are also excluded.

*Sum of the three previous columns.

VI. THE α-HELIX AND β-STRAND DIPOLES AND PREFERRED FOLDING PATTERNS IN GLOBULAR PROTEINS

1. *Protein Folding*

Globular proteins perform an immense range of different functions in biological systems. These functions can, however, only be carried out correctly if the protein molecule has assumed a well-defined globular state—only then the various essential side chains are in the correct spatial position with respect to each other. Consequently, understanding why protein molecules adopt their "native" conformations is a central issue in understanding living organisms.

This important question is often called the problem of "protein folding". This term does, however, not have a very precise meaning. At least four groups of questions covered by this term can be distinguished. One is the pathway by which the extended polypeptide chain is folded into a globular unit. Also, the time necessary for this process is then considered. A second protein folding question is: why do proteins assume well-defined globular structures at all? What are the factors which overcome the enormous entropy loss when the coiling polypeptide chain is assuming one rather well defined conformation? A third aspect of protein folding concerns general folding patterns, which are sequence-independent. One may wonder why are helices right-handed? Why are β-sheets left-handed? Are there preferred folding units of a polypeptide chain and, if so, why? A fourth and highly important question is: what is the link between the amino acid sequence and the observed three-dimensional structure?

The first aspect of protein folding will not be discussed in this review, although one should keep in mind that possibly certain "preferred folding patterns" may also be determined by kinetic factors concerning the folding process. In fact, we do not know whether globular proteins assume a global minimum in free energy, or are kinetically trapped in one out of a multitude of local free energy minima.

The overall aspect of protein folding is also poorly understood. One may quote Tanford (1970, p. 48) in this respect: "Were it necessary to make a prediction in the absence of experimental knowledge, one would probably conclude that the native state should not exist". The problem is, as Finney *et al.* (1980) have also pointed out, that we have to determine the small difference between two large numbers—and we do not know these large numbers with great accuracy. The small number is the difference between the free energies of the folded

protein and of the coiling polypeptide chain. This difference is in the order of 5–15 kcal mol^{-1} (Privalov, 1979). Entropy terms of the main chain and of the water molecules surrounding hydrophobic residues appear to be the crucial factors—with the latter term compensating the first when a globular unit is formed. Kauzmann (1959) has suggested, before any three-dimensional structure of a protein was determined, that the hydrophobic interaction, i.e. the entropy gain of water molecules, is the most important factor in lowering the free energy of a native conformation in aqueous solution. This appears to be in agreement with later observations that proteins have one or more hydrophobic cores in their interior (see e.g. Meirovitch and Scheraga, 1980, and references therein). However, *quantitatively* our understanding of the complex factors driving the polypeptide chain to a folded state is rather limited (Tanford, 1970; Finney *et al.*, 1980; Richards and Richmond, 1978).

The third question is probably the most easy to approach with current computational techniques. In considering a number of different possible conformations, one disregards entropic factors, and calculates energies. Although the absolute values of these energies are quite uncertain, it appears that energy *differences* can be calculated, the *sign* of which is rather independent of the computational procedure (Bernstein and Hagler, 1978; Hagler and Bernstein, 1978; Bar and Bernstein, 1982). This approach has revealed the reasons why α-helices are right-handed (see e.g. Ramachandran *et al.*, 1966), while understanding the left-handedness of β-sheet structures is an area of much current debate (Chothia, 1973; Nishikawa and Scheraga, 1976; Salemme, 1981; Chou *et al.*, 1982, 1983). In this third category also fall the explanations based on general steric effects which have been given for preferred modes of helix–helix, helix–sheet and sheet–sheet packing observed in proteins (Chothia *et al.*, 1977; Janin and Chothia, 1980; Chothia and Janin, 1981).

The fourth question is obviously of very great significance, but the relationship between the amino acid sequence and the three-dimensional structure of proteins seems to be elusive. One is struck by the ever-increasing number of examples where amino acid sequences are vastly different and three-dimensional structures are remarkably similar. One set of examples can be found in Tables 5 and 6. On top of this, at least one example is known where a stretch of eleven residues assumes completely different conformations in two proteins, although only a single residue is different out of eleven (Dijkstra *et al.*, 1983a, b). It may therefore take a long time before real progress is made in this area. We will not discuss it further in this review, in spite of its great challenge and importance.

2. *A Possible Explanation for Preferred Folding Patterns in Proteins*

Here, we will look into a number of questions which fall under category three discussed above. In their elegant analysis of protein structures, Levitt and Chothia (1976) have shown that the large majority of protein domains fall into three categories: (i) all-helical proteins; (ii) all-β proteins where the β-strands are virtually all anti-parallel (Richardson, 1977; Hol *et al.*, 1981); and (iii) α/β proteins which contain a central parallel twisted β-sheet, flanked by α-helices on both sides. These authors also report the occurrence of three preferred folding units: the αα, the ββ and the βαβ-unit (Fig. 8).

Hol *et al.* (1981) have made a proposal aimed at explaining these preferred folding patterns. Making the assumption that hydrophobic interactions are insensitive to the direction of β-strands and α-helices in folded proteins, and realizing that great variations in amino acid sequences are allowed for proteins assuming virtually identical tertiary structures, the authors focused on the interactions between peptide dipoles of secondary structure elements in globular proteins.

The α-helix contains a considerable dipole moment as discussed in previous sections. As illustrated in Fig. 9, the peptide units in a parallel β-structure are arranged in such a manner that a resultant dipole moment arises if the dipole moment of each peptide unit is parallel to the C=O and N–H bonds. Then, approximately 30% of the peptide dipole moment points along the axis of the β-strand with the N-terminus being the positive end of the β-dipole (Hol *et al.*, 1981). Figure 9 also shows that the β-strands in an anti-parallel arrangement have a negligible dipole moment.

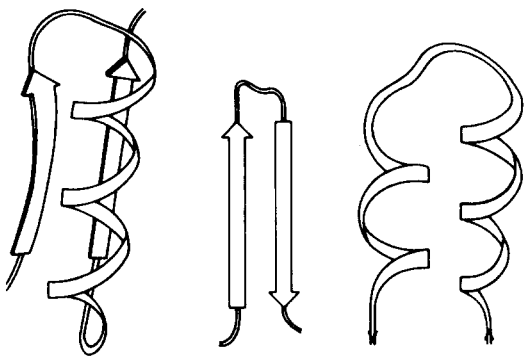


FIG. 8. Three folding units frequently observed in globular proteins: the αα, ββ and βαβ units (from right to left). (Fig. from Levitt and Chothia, 1976.)

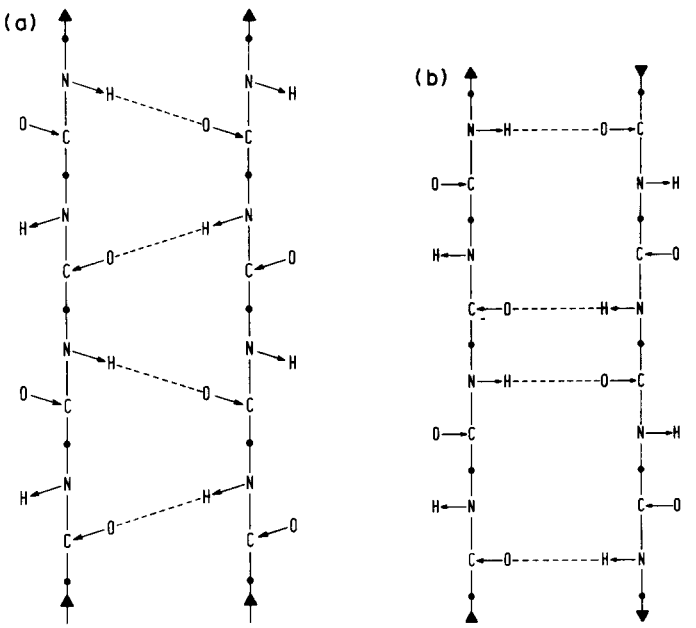


FIG. 9. Schematic view of the direction of N→H and O→C dipoles in β-sheets. Dotted lines are hydrogen bonds.
(a) Parallel β-sheet, with the N→H and O→C directions making an angle of about 20° with the normal to the strand directions.
(b) Anti-parallel β-sheet, with the N–H and C–O bonds perpendicular to the strand directions. (Fig. from Hol *et al.*, 1981.)

With the α-helix and parallel β-strand dipoles, the explanation for the observed preferred folding patterns in proteins which has been put forward (Hol *et al.*, 1981) is as follows:

(i) In all-helical proteins, helices tend to be anti-parallel because of a favourable interaction between α-helix dipoles. This also explains the frequent occurrence of αα-units;

(ii) In α/β-proteins, a favourable interaction occurs between the β-dipoles in the centre of the molecule and the α-helix dipoles arranged in an anti-parallel way around the parallel β-strands. Such an interaction would also explain the stability of the βαβ-unit;

(iii) In all-β proteins the favourable α-β dipole interactions are absent and β-strands tend to be anti-parallel as an unfavourable interaction between parallel β-dipoles would otherwise occur.

This is a simple, "unifying", hypothesis, but can it be proved? Let us proceed in three steps. First, calculate electrostatic energies between secondary structure elements, in vacuum, with

the peptide charge distribution given in Fig. 1. Second, discuss the effect of different peptide charge distributions on the electrostatic interactions. Third, try to estimate dielectric screening effects, in particular that of the water molecules surrounding the protein.

(a) *Calculations of electrostatic interactions, in vacuum, between secondary structure elements*

The interactions between two α-helix dipoles in vacuum have been calculated by Hol *et al.* (1981). The most important parameters appear to be: (i) the helix length; (ii) the distance d between the helices; and (iii) the angle $\Omega_{\alpha\alpha}$ (Fig. 10). In Fig. 11 some results of the model calculations are indicated. It may be noted here that distances between axes of helices in contact with each other in actual proteins vary from 7 to 11 Å (Chothia *et al.*, 1981). Figure 11(c) shows that $\Delta U_{\alpha\alpha}$, i.e. the difference in electrostatic interaction energy between,

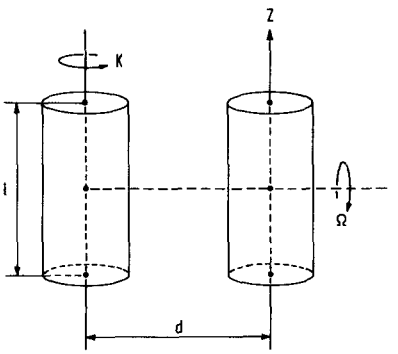


FIG. 10. Definition of parameters describing the mutual orientation and position of two helices. A helix can be rotated about its own axis by an angle K (with negligible variation in electrostatic energy) and about the inter-helical axis by an angle $\Omega_{\alpha\alpha}$. l is the length of both helices: $l = N \times 1.5$ Å where N is the number of residues in each helix. d is the distance between the helix axes. (Fig. from Hol *et al.*, 1981.)

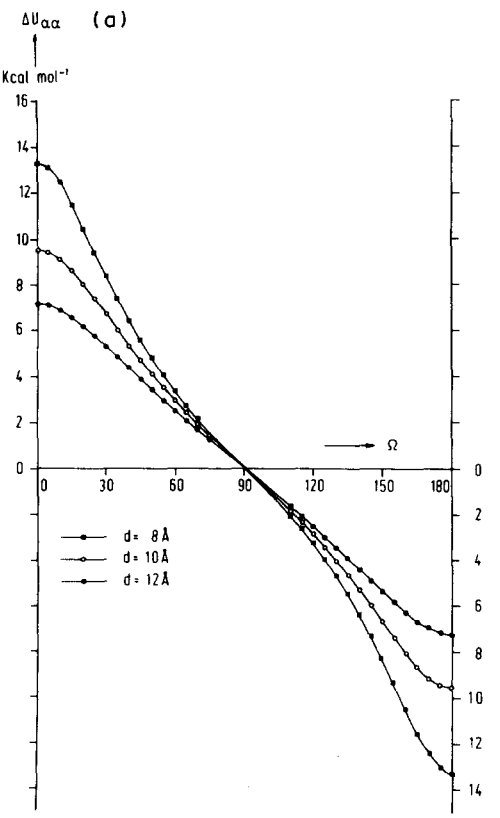


FIG. 11.

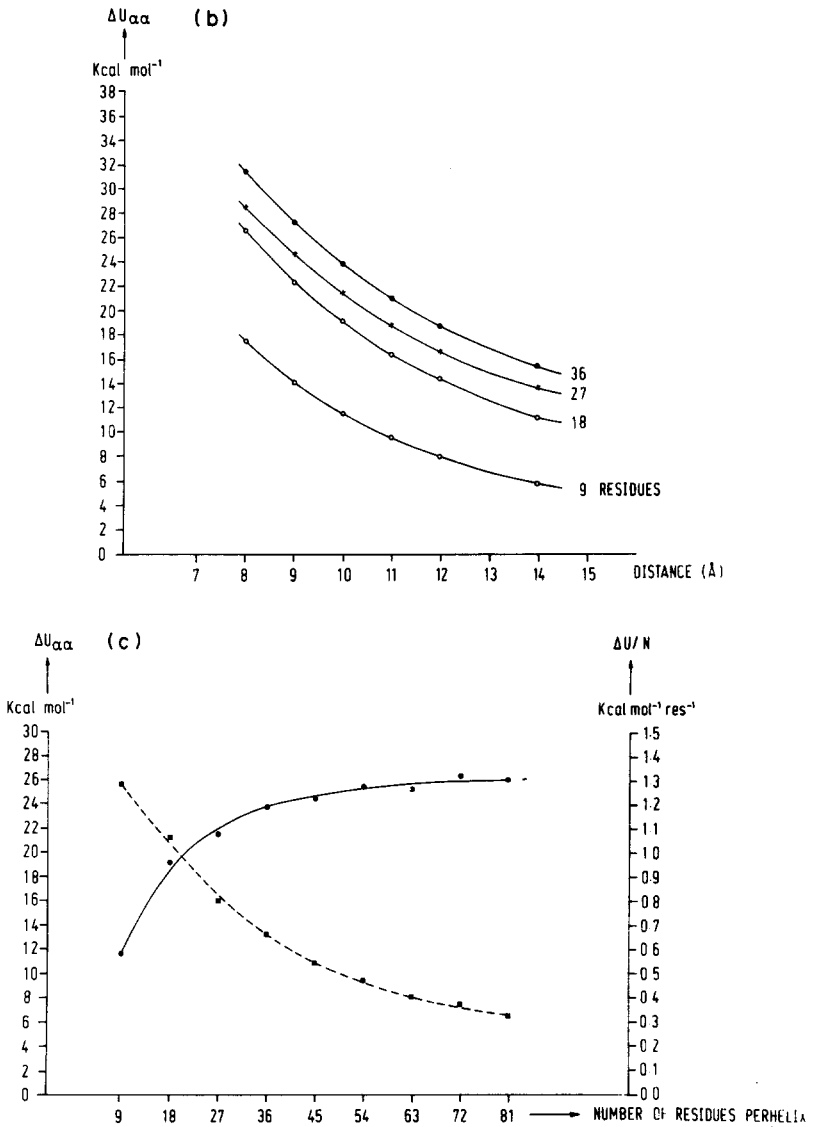


FIG. 11. The electrostatic interaction energy of two helices as function of rotational and translational parameters (defined in Fig. 10). No dielectric screening taken into account. Only partial charges of the backbone, as shown in Fig. 1, are included. The torsion angles (φ,ψ) of the helical residues are $(-57°, -48°)$. As Wada (Wada, 1976, Fig. 3) has shown, little variation in U_{zz} is to be expected if the (φ,ψ) angles are different, as long as they are in the helical region of the Ramachandran plot.

(a) The variation of U_{zz} as function of angle Ω and of interaxial distance d. The length of both helices was five turns, or 27 Å, i.e. 18 residues.

(b) The variation of ΔU_{zz}, i.e. the electrostatic energy difference between two parallel and two anti-parallel helices, as function of interaxial distance, d, for different helix lengths.

(c) The dependence of $\Delta U_{\alpha\alpha}$ (solid line) on the number of residues per helix, which levels off rapidly beyond 27 residues per helix. The interaxial distance is 10 Å. The figure illustrates that the electrostatic energy difference becomes less important on a per residue basis for longer helices (broken line). (Fig. taken from Hol *et al.*, 1981.)

on the one hand, two parallel and, on the other hand, two anti-parallel helices, levels off quite rapidly when the helix length is increased.

In systems containing more than one helix pair, helix–dipole interactions give rise to considerable energy differences between parallel and anti-parallel arrangements of helices, the latter being clearly preferred as shown in Fig. 12. It should be emphasized here that these calculations are all performed without taking dielectric screening effects into account. Nevertheless, even if the effective dielectric constant would be in the order of 10 then, among

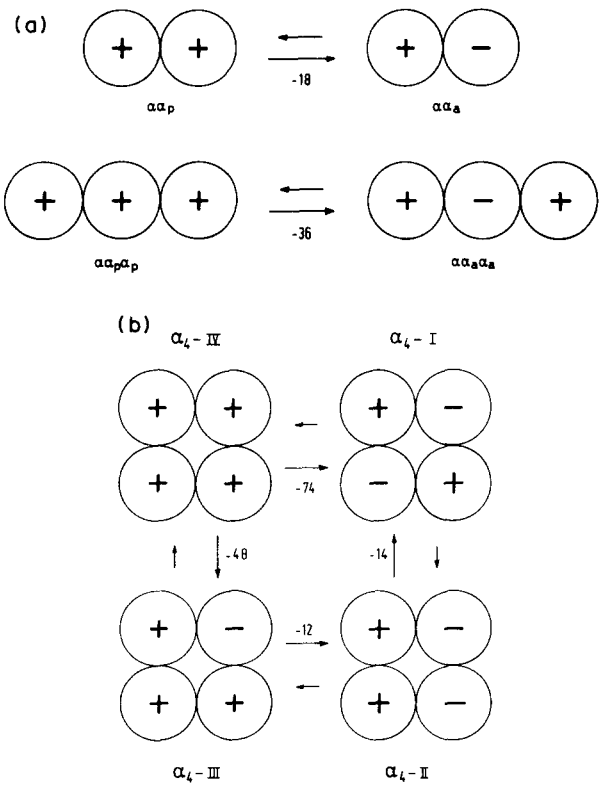


FIG. 12. The electrostatic energy difference between parallel and anti-parallel helical arrangements, in kcal mol^{-1}. Calculated from Fig. 11, with $d = 10$ Å for nearest neighbours, and $l = 27$ Å, or 18 residues. The + sign indicates the N-terminus of a helix.
(a) The difference between two and three parallel and anti-parallel helices with their axes lying in one plane.
(b) The energy differences for the possible arrangements with four helices.

others, anti-parallel 4-helical bundles would still be greatly preferred over all-parallel arrangements (Fig. 12).

The same conclusion was reached by Sheridan *et al.* (1982). Weill and André (1978) have also looked into the importance of the helix dipole for "head-to-tail" *versus* "side-by-side" anti-parallel aggregations of helices. Silverman and Scheraga (1972), in a study of interhelical interactions of poly-alanine in water, conclude that two anti-parallel helices are stabilized by electrostatic interactions between the backbones to the extent of −0.35 kcal/mole residue. The stabilization due to hydrophobic bonding was estimated as −0.25 kcal/mole residue, i.e. smaller than the favourable interactions between the helix dipoles. These authors suggest that these interactions play a role in the formation of incipient globularity in some proteins, whose folding may be initiated around such an anti-parallel array of α-helices. It should be mentioned, however, that in actual proteins side chains are on the average larger than that of alanine. Therefore hydrophobic interactions would in all likelihood be more important than the electrostatic interactions between the backbone peptide units.

Calculations on the electrostatic interactions between peptide units in model $\beta\alpha\beta$-units have also been carried out (W. G. J. Hol, to be published). The geometry used is shown in Fig. 13. The angle $\Omega_{\alpha\beta}$ is a crucial parameter: for $\Omega_{\alpha\beta} = 0°$, the helix runs parallel to the two β-strands; if $\Omega_{\alpha\beta} = 180°$ the helix runs anti-parallel and a model of the $\beta\alpha\beta$-units observed in proteins is obtained. The quantity calculated, $U_{\beta\alpha\beta}$, is the interaction energy between the peptide atoms in the α-helix, on the one hand, and the peptide atoms in the two β-strands, on the other hand. Table 9 gives, for a number of (φ,ψ)-combinations for the β-strands, the electrostatic energy difference $\Delta U_{\beta\alpha\beta} = U_{\beta\alpha\beta}\,(\Omega_{\alpha\beta} = 0°) - U_{\beta\alpha\beta}\,(\Omega_{\alpha\beta} = 180°)$. A positive value for $\Delta U_{\beta\alpha\beta}$ means that the $\beta\alpha\beta$-unit with the helix anti-parallel to the β-strands is more favourable than an arrangement with the helix parallel to the strands. This is observed (Table

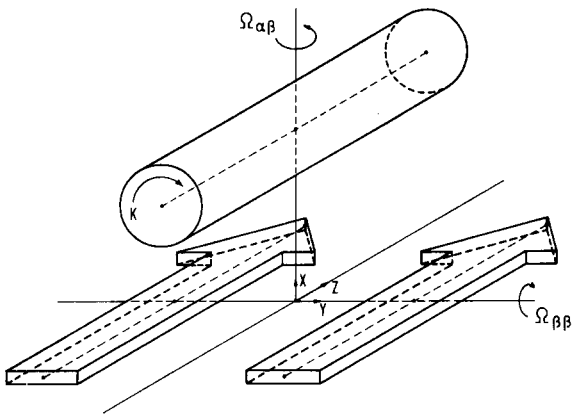


FIG. 13. Definition of rotational and translational parameters of a model βαβ-unit. $\Omega_{\alpha\beta}$ is the angle between the helix axis and the direction of the β-strands. For $\Omega_{\alpha\beta}=0°$, the helix runs parallel to the strands. If $\Omega_{\alpha\beta}=180°$, the helix runs anti-parallel and a model of βαβ-units frequently occurring in proteins (Fig. 8) is obtained. The variation of K, rotating the helix about its own axis, hardly affects $U_{\beta\alpha\beta}$. Also small variations in $\Omega_{\beta\beta}$ (carried out by rotating one β-strand $+\frac{1}{2}\Omega_{\beta\beta}$ about the Y-axis, and the second β-strand $-\frac{1}{2}\Omega_{\beta\beta}$ about this axis) changes $U_{\beta\alpha\beta}$ only marginally. In the calculations performed for Tables 9 and 11, the helix centre was kept at a constant distance of 8 Å along X above the origin of the axial system, while the centres of the β-strands were separated 4.7 Å (Pauling and Corey, 1953) from each other along Y.

TABLE 9. ELECTROSTATIC INTERACTION ENERGIES OF MODEL βαβ-UNITS

φ_B, ψ_B of the Two β-Strands (°)	$\Delta U_{\beta\alpha\beta}$ (kcal mol^{-1})	The φ_B, ψ_B Combination is Typical for
−119, +113	+2.6	a parallel β-strand
−129, +124	+1.0	intermediate
−139, +146	+0.4	an anti-parallel β-strand

A positive value of $\Delta U_{\beta\alpha\beta}$, defined in the text, implies that a βαβ-unit with the helix anti-parallel to the β-strands is electrostatically more stable than the situation with the helix parallel to the strands (Fig. 13). Further details of the calculations (W. G. J. Hol, to be published): (i) the length of the β-strands was 6 residues, that of the helix 14 residues. These are the average values found in α/β proteins (Janin and Chothia, 1980); (ii) the partial charges of the main chain peptide units were taken as in Fig. 1; (iii) no dielectric screening was taken into account; (iv) the (φ, ψ) angles for the model α-helix were (−57°, −48°); (v) the centre of the helix was 8 Å above the plane of the β-strands; (vi) the β-strands were 4.7 Å apart (Pauling & Corey, 1953); (vii) the interactions between the peptide units of the β-strands were ignored as these do not depend on $\Omega_{\alpha\beta}$; and (vii) $U_{\beta\alpha\beta}$ appeared to be hardly affected by varying K, i.e. rotating the helix about its own axis (Fig. 13).

9) if the φ,ψ-angles of the β-strands correspond with the values for parallel-strands. Thus the actual βαβ-unit as frequently encountered in proteins is electrostatically preferred by the backbone peptide units.

A characteristic feature of actual α/β proteins is a considerable twist of the parallel β-sheet (Janin and Chothia, 1980). The average angle between adjacent β-strand ($\Omega_{\beta\beta}$ in Fig. 13) is not 0°, as assumed in the model calculations, but about 20°. However, the calculations of $U_{\beta\alpha\beta}$ are hardly affected by this difference (W. G. J. Hol, not shown). Concomitant with the twist of the β-sheet, the helices in α/β proteins are also rotated considerably with respect to each other. In these proteins the torsion angle between two adjacent "parallel" helices, $\Omega_{\alpha\alpha}$, is typically of the order of 40° (Janin and Chothia, 1980). Inspection of Fig. 11(a) shows that this angle between helix axes leads to a considerable reduction of the unfavourable $U_{\alpha\alpha}$ which would have occurred if the helices were perfectly parallel. Therefore, in α/β proteins two electrostatic factors may be operative: (i) a favourable interaction between main chain dipoles in the βαβ-units; and, (ii) an avoidance of the unfavourable interaction of truly parallel helices by a considerable rotation of helices with respect to each other.

Although Table 9 shows considerably smaller numbers for $\Delta U_{\beta\alpha\beta}$ than Fig. 11 for $\Delta U_{\alpha\alpha}$, it should be mentioned that the parallel β-sheets are buried in the interior of α/β-proteins (Richardson, 1977). The dielectric screening effects of the solvent for the α–β interactions are

therefore less than for the interaction between exposed helices. Consequently, it is to be expected that the effective dielectric constant for $\Delta U_{\beta x \beta}$ will be smaller than for ΔU_{xx}.

Calculation of the interaction energies between peptide units in adjacent β-strands is difficult because the partially charged atoms involved come near to each other and small differences in geometry greatly affect $U_{\beta\beta}$. It has been calculated (Hol *et al.*, 1981) by a line dipole approximation of the parallel β-strand, that the anti-parallel arrangement of two β-strands is preferred by 0.4 kcal mol^{-1} with respect to a parallel arrangement. Extending this simple approximation to sheets with a larger number of strands, and ignoring interactions between non-nearest neighbours, energy differences as given in Fig. 14 are obtained. Although these energies are again smaller than for the $\beta\alpha\beta$-units, it may be noted that the average distance between the peptide atoms involved is considerably smaller than in $\beta\alpha\beta$ and $\alpha\alpha$ units and that consequently the effective dielectric constant is also smaller.

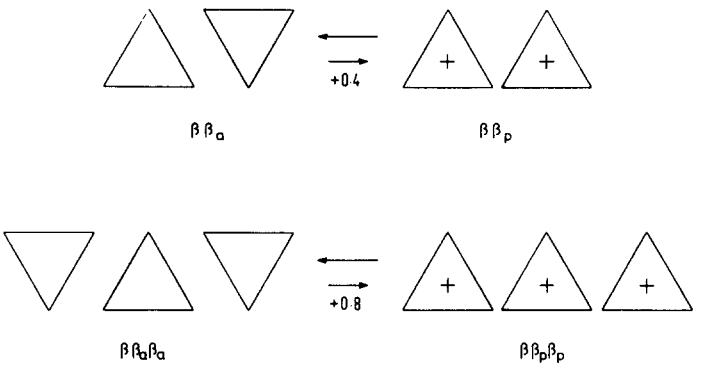


Fig. 14. Estimates of electrostatic energy differences, in kcal mol^{-1}, between two- and three-stranded parallel and anti-parallel arrangements. No dielectric screening effect was taken into account. The β-strands have a length of four residues, or about 14 Å, and are separated by 4.7 Å (Pauling and Corey, 1953). An anti-parallel strand was given a zero dipole moment, while the dipole moment of a parallel strand was approximated by placing $+1/15$ elementary charge at its N-terminus and $-1/15$ elementary charge at the C-terminus. These values are obtained by assuming (i) the charge distribution of Fig. 1 for the peptide atoms, and (ii) an angle of $\sim 20°$ between the C→O and H→N bonds and the normal to the strand direction [Fig. 9(a)]. This means that about 1/3 of the dipole moment runs parallel to the parallel β-strand. Hence, the component along the strand direction is: $\mu_{ax} = 1/3\mu_{peptide} = 1/3 \times 3.5$ D $\cong 1.15$ D $\cong 0.23$ e.Å. This may also be written as $\mu_{ax} = 1/15e \times 3.5$ Å. As the repeat distance along a β-strand is about 3.5 Å (Pauling and Corey, 1953), all partial charges cancel to a first approximation and only $+1/15$ elementary charge at the N-terminus and $-1/15$ elementary charge at the C-terminus remains. It should be pointed out, however, that uncertainty exists as to the charge distribution of peptide units (see Table 11 and text).

So much for model calculations, how about real proteins? In Table 10 the results of Hol *et al.* (1981) are given for a variety of all-helical and α/β proteins. It appears that in all-helical proteins the U_{xx} values are invariably negative. In the α/β proteins the U_{xx} values are of the same order as the $U_{\alpha\beta}$ values. As mentioned above, the interaction between the β-strands and the helices is probably less affected by the screening effects of solvent than the interactions between helices, which makes it reasonable to assume that the $U_{\alpha\alpha}$ term is overcompensated by the $U_{\alpha\beta}$ term when more sophisticated calculations, taking dielectric effects into account, are carried out.

(b) *The effect of utilizing different peptide charge distributions*

In the calculations above, the charge distribution of the peptide unit as shown in Fig. 1 is used. It appears from model calculations that U_{xx} is rather insensitive to variations in this charge distribution. $U_{\beta x \beta}$, however, appears to be very sensitive to the charge distribution chosen. This becomes clear from an inspection of Fig. 9: a small deviation in the *direction* of the dipole moment has a dramatic effect on the component of the dipole moment along the axis of the β-strand and thus on the value of $U_{\beta x \beta}$. The sensitivity of $\Delta U_{\beta x \beta}$ for the peptide dipole direction is quantified in Table 11 where several peptide charge distributions are compared. The first is a distribution used by Ferro and coworkers (1980) and has proved to

TABLE 10. ELECTROSTATIC INTERACTION ENERGIES BETWEEN α-HELICES AND β-STRANDS IN ACTUAL PROTEINS

	Type of Protein	$U_{\alpha\alpha}$ (kcal mol^{-1})	$U_{\alpha\beta}$ (kcal mol^{-1})
Myoglobin	α_a	−22.7	—
TMV	α_a	−14.3	—
Parvalbumin	α_a	−11.8	—
Cytochrome b_5	α_a	−4.7	—
Phospholipase A_2	$\alpha_a\beta_a$	−8.3	−0.8
Cytochrome *c*	$\alpha_a\beta_a$	−3.4	−1.4
Hen egg-white lysozyme	$\alpha_a\beta_a$	−12.8	−1.6
Phage T_4 lysozyme	$\alpha_a\beta_a$	−23.5	−0.7
Thermolysin	$\alpha_a\beta$	−3.6	−5.4
Lactate dehydrogenase	$\alpha_p\beta_p$	+14.0	−14.0
Alcohol dehydrogenase	$\alpha_p\beta_p$	+14.2	−15.3
Glyceraldehyde phosphate dehydrogenase	$\alpha_p\beta_p$	+5.8	−5.7
Adenylate kinase	$\alpha_p\beta_p$	+4.9	−10.5
Rhodanese	$\alpha_p\beta_p$	+13.5	−10.3
Subtilisin	$\alpha_p\beta_p$	+5.2	−15.0
Dihydrofolate reductase	$\alpha_p\beta_p$	+3.2	−3.8
Flavodoxin	$\alpha_p\beta_p$	+13.3	−4.9
Triose phosphate isomerase	$\alpha_p\beta_p$	+6.6	−7.0
Carboxypeptidase	$\alpha_p\beta_p$	+11.4	−11.0
Papain/actinidin		+13.3	−6.6

Taken from Hol *et al.* (1981). Only main chain peptide atoms were considered with a charge distribution as given in Fig. 1. No dielectric screening was taken into account. For further details, see Hol *et al.* (1981).
α_a: all-helical proteins in which the helices run mainly anti-parallel;
$\alpha_a\beta_a$: proteins with anti-parallel α-helices and anti-parallel β-strands;
$\alpha_p\beta_p$: proteins with a central mainly parallel β-sheet surrounded by α-helices which run "parallel" to each other, but with considerable inter-helical angles (Janin and Chothia, 1980).

TABLE 11. ELECTROSTATIC INTERACTION ENERGY OF THE MAIN CHAIN PEPTIDE UNITS IN A MODEL βαβ-UNIT AS FUNCTION OF THE PEPTIDE CHARGE DISTRIBUTION (W. G. J. HOL, TO BE PUBLISHED)

Charge Distribution Group	Partial Charges (in Electrons) HN	N	HA	CA	C	O	$\Delta U_{\beta\alpha\beta}$ (kcal mol^{-1})	Footnote
I	+0.204	−0.204	+0.046	+0.056	+0.318	+0.420	+7.2	*
II	+0.280	−0.280	0.000	0.000	+0.380	−0.380	+2.8	†
II	+0.200	−0.200	0.000	0.000	+0.420	−0.420	+2.6	‡
III	+0.250	−0.350	0.000	+0.100	+0.450	−0.450	+0.0	§
IV	+0.176	−0.356	0.000	+0.114	+0.450	−0.384	−3.3	‖
IV	+0.350	−0.650	−0.040	+0.360	+0.680	−0.700	−4.5	¶

The calculations are carried out on a model βαβ-unit as depicted in Fig. 14 and described in the legend to Table 9. The (φ, ψ)-values of the β-strands were (−119°, +113°).
Footnotes:
*Charges used by the energy-minimization program REFINE2 of Ferro *et al.* (1980).
†Charges used in the molecular dynamics simulation by Van Gunsteren (1984).
‡Charges depicted in Fig. 1 and used by Hol *et al.* (1978, 1981).
§Charges used in the molecular dynamics simulations of the Harvard group (Brooks *et al.*, 1983).
‖Charges from Momany *et al.* (1975).
¶Charges obtained by *ab initio* quantum mechanical calculations on β-strands (Van Duijnen, De Jager and Thole, in preparation).

be quite successful in the energy-refinement of rubredoxin: the ω-values after energy minimization appeared to correlate well with the ω-values after further crystallographic refinement. In the second group of charge distributions used in Table 11, the dipole moment runs parallel to the N–H and C=O bonds. This direction for the peptide dipole moment is used by a large number of investigators (Hol *et al.*, 1978; Van Gunsteren, 1984; Prasad and Sasisekharan, 1979; Cantor and Schimmel, 1980; Lifson *et al.*, 1979; Warshel and Levitt, 1976; Jernigan *et al.*, 1980; Brant *et al.*, 1967) and is in agreement with the concept of bond dipole moment. These two groups of charge distribution give positive values for $\Delta U_{\beta\alpha\beta}$ (Table 11), i.e. the βαβ-unit as found in actual proteins is electrostatically preferred. The third group in Table 11 does not show such a preference while for the charge distributions in Group IV an

arrangement with the helix parallel to the β-strands would be electrostatically preferred. It is clear that $\Delta U_{\beta\alpha\beta}$ is very sensitive to the charge distribution of the peptide unit used and also that no consensus exists among workers in the field as to what the correct peptide charge distribution actually is. This is also concluded by Hagler and Lapiccirella (1976) who have looked carefully into electron distributions of peptides and related compounds. These authors caution for the use of partial charges obtained by population analysis of quantum mechanical calculations in conformational analyses. It is clear that further computational and, if possible, real experiments on the backbone dipole moment in β-strands and β-sheets are very important in order to provide a solid base for the calculation of electrostatic interactions in proteins.

(c) *Dielectric screening by the solvent*

In order to calculate electrostatic interactions on globular proteins in solution, the effect of the surrounding water molecules has to be taken into account. Bulk water has special dielectric properties which are due to: (i) the large dipole moment of about 1.85 D of the water molecule (Eisenberg and Kauzmann, 1969) and (ii) the rotational mobility of these molecules. The importance of the dipole moment of a molecule for the dielectric constant of a liquid is shown by the values of 80 for water and 109 for liquid formamide in contrast with values of e.g. 1.9, 2.0 and 2.3 for the non-polar liquids *n*-hexane, cyclo-hexane and benzene, respectively (Handbook of Chemistry and Physics, 1968–1969). The necessity of rotational freedom of the molecules is amply demonstrated by: (i) the difference in dielectric constant of 109 for liquid formamide *versus* 4.0 for solid acetamide (Handbook of Chemistry and Physics, 1968–1969), (ii) the contrast between the value of 80 for water and 4 for ice II and (iii) very clearly by the dramatic difference between, on the one hand, the large dielectric constants, with values ranging from 100–200, of ices I, III, V, VI and VII where the water molecules constantly change their orientations and, on the other hand, the values of ~ 4 for ices II and VIII with immobile waters (see e.g. Eisenberg and Kauzmann, 1969, pp. 106–107).

Quantitative calculations of electrostatic energies between charges and dipoles of dissolved protein molecules are difficult because the dynamic aspects of water and protein have to be taken into account. The following regions of the system under consideration can be distinguished:

(i) the "bulk" protein with a densely packed array of main chain and side chain atoms (Richards, 1977), a low mobility and thus a low dielectric screening effect;
(ii) the outer surface of the protein with often great mobility for side chains extending into the solvent;
(iii) the first hydration shell of the protein with quite unknown dielectric properties as e.g. Warshel (Warshel, 1979) has stressed. This third region may also include water molecules in clefts and crevices of protein molecules;
(iv) the bulk water molecules which can freely rotate and orient themselves statistically in preferred orientations in response to the electrical field generated by the monopolies and dipoles of the protein molecule.

The concept of a "dielectric constant" breaks down in such a system. Also a distance-dependent dielectric constant is incorrect as it is obvious that the electrostatic interaction between two charges separated by a distance R in the interior of the protein is quite different from the electrostatic interaction between two charges separated by the same distance R at the surface of the protein. It seems best to speak of a general dielectric screening effect which is different, in principle, for every different process one is considering in such a system. Following Rees (1980), the screening effect can be quantified by introducing an "effective dielectric constant", ε_{eff}, which is simply the ratio between the electrostatic interaction energy of a process in vacuum and the same process in the dielectric system under consideration. However, this "effective dielectric constant" is no constant at all (Rees, 1980) and in order to avoid confusion the term "dielectric screening factor", S, is suggested. Completely analogous to the definition of ε_{eff} by Rees (1980), S is given by:

$$S = \frac{U^{\text{electrostatic}}_{\text{vacuum}}}{U^{\text{electrostatic}}_{\text{dielectric}}}, \tag{11}$$

where

$$U_{\text{vacuum}}^{\text{electrostatic}} = \frac{332\, q_1 \cdot q_2}{r_{12}} \text{ kcal mol}^{-1} \tag{12}$$

for two charges q_1 and q_2 (expressed as "fractional charges", i.e. one electron has charge -1) at a mutual distance of r_{12} Å. The variability of S can be demonstrated by the approximation of Northrup *et al.* (1981) used for the electrostatic interactions in molecular dynamics calculations of proteins. Their reduction of the partial charges of the atoms as function of the distance to the protein centre has the following effect. For two charges near the centre $S \cong 1.2$. For two charges, equally far apart, near the surface of the molecule $S \cong 11$. The electrostatic interaction energy does not only depend on the distance between the charges, but also on the distances of the charges from the surrounding solvent. What we like to obtain is an estimate of S for the interaction between helix dipoles in a protein dissolved in water.

There are two, fundamentally different, approaches for calculating the electrostatic energies required for our purposes. The first approach consists of methods which treats the entire system microscopically, i.e. the water molecules as well as the protein atoms are explicitly taken into account. In this manner the different rotational mobilities of water molecules near to and far from the protein surface are incorporated into the calculations. The methods available for this approach are computationally quite complex. They include molecular dynamics and Monte Carlo calculations on a protein molecule surrounded by a large number of water molecules. At present, only one molecular dynamics calculation has been performed on a protein molecule surrounded by water (Van Gunsteren *et al.*, 1983) and even in that case a relatively small number of water molecules was taken into account as a hydrated protein crystal was simulated. Another microscopic method is the "surface constrained soft sphere dipole" water model developed by Warshel (1979).

The second approach for calculating electrostatic energies is the use of a continuum for the solvent molecules surrounding the protein which itself is treated microscopically. Analytical solutions for this electrostatic problem are only available for systems consisting of spherical or ellipsoidal cavities (Tanford and Kirkwood, 1957; Ehrenson, 1976; Bötcher, 1973) which do not allow the convoluted protein surface to be taken into account, or only in an approximate manner (Shire *et al.*, 1974). These methods do also not seem suitable for coping with the special properties of the first shell of hydration. Recently, Warwicker and Watson (1982) have introduced a method employing a grid for the solvent region which allowed calculations on systems of any shape. This method is also computationally expensive but may be extended to include regions with "bound" water with lower rotational mobility near the surface and in crevices of protein molecules. For present purposes, i.e. obtaining a general impression of the screening effect of the solvent, Friedman's "image charge method" (1975) is probably sufficiently accurate and will be used in the discussion below.

In Friedman's approximation we consider a sphere of low dielectric constant embedded in a medium with high dielectric constant (Friedman, 1975). Consider a "source charge", q_s, at a distance r from the centre of the sphere with a radius a. The effect of the orientation of the water molecules, or the "reaction field", can then be approximated by an "image charge", q^{im}, which is situated outside the sphere on a line from the centre of the sphere through q_s, as indicated in Fig. 15. The distance of q^{im} to the centre of the sphere is given by:

$$r_{\text{im}} = \frac{a}{r} \times a \tag{13}$$

and the magnitude of the image charge is:

$$q^{\text{im}} = -\frac{(\varepsilon_{\text{outer}} - \varepsilon_{\text{inner}})}{(\varepsilon_{\text{outer}} + \varepsilon_{\text{inner}})} \times \frac{a}{r} \times q_s. \tag{14}$$

This image charge approximation does not hold for charges close to the surface of the sphere (Friedman, 1975).

The electrostatic energy of two charges within the sphere is equal to:

$$U_{q_1 q_2} = q_1 \cdot \phi_2 + q_1 \cdot \phi_2^{\text{im}} + \tfrac{1}{2} q_1 \cdot \phi_1^{\text{im}} + \tfrac{1}{2} q_2 \cdot \phi_2^{\text{im}} \tag{15}$$

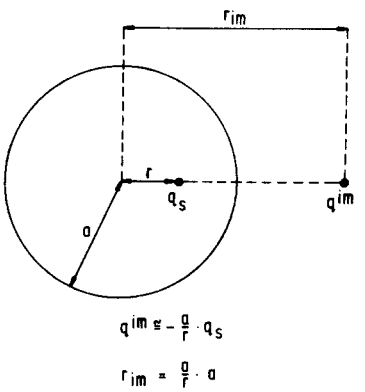


FIG. 15. Friedman's image charge approximation to the reaction field (Friedman, 1975). The reaction field of a medium with high dielectric constant surrounding a sphere with low dielectric constant containing a "source charge", q_s, can be approximated by the effect of an image charge, q^{im}, at the position shown in the figure and discussed in the text. The formula for q^{im} in the figure is a good approximation for eqn (14) in the text when $\varepsilon_{outer} \gg \varepsilon_{inner}$.

where ϕ_2 is the potential due to the source charge q_2, while ϕ_1^{im} and ϕ_2^{im} are the potentials due to the image charges q_1^{im} and q_2^{im}, respectively. The first two terms of eqn (15) give the energy required to bring q_1 from infinity to its position within the sphere in the presence of q_2. The second and third terms of eqn (15) are the energies of q_1 resp. q_2 in their own reaction field, or "solvation energies". The factor $\frac{1}{2}$ of these two terms is due to the fact that it costs energy to create the reaction field or, in other words, to orient the water molecules (see e.g. pp. 141–144 of Bötcher, 1973).

Turning now to the electrostatic energy of the main chain atoms of two helices, $U_{\alpha\alpha}$, then six terms have to be considered:

$$\left.\begin{aligned} U_{\alpha\alpha} = {} & \text{energy of helix 1 in } \phi \text{ of helix 2} \\ & + \text{energy of helix 1 in } \phi^{im} \text{ of helix 2} \end{aligned}\right\} U_{\alpha\alpha}^{int}$$
$$\left.\begin{aligned} & + \text{self energy of helix 1} \\ & + \text{self energy of helix 2} \end{aligned}\right\} U_{\alpha\alpha}^{self}$$
$$\left.\begin{aligned} & + \text{solvation energy of helix 1} \\ & + \text{solvation energy of helix 2} \end{aligned}\right\} U_{\alpha\alpha}^{solv} \qquad (16)$$

As we are interested in preferred folding patterns of α-helices, the value of $\Delta U_{\alpha\alpha}$ as function of $\Omega_{\alpha\alpha}$ is the significant quantity. For helices of constant geometry, the value of $U_{\alpha\alpha}^{self}$ in eqn (16), i.e. the energy required to assemble each α-helix in vacuum, is constant and can be ignored for present purposes. For the arrangement of two helices as shown in Fig. 10, the solvation energies of the helices can be considered constant when $\Omega_{\alpha\alpha}$ is varied, provided that the centroid of the two helices coincides with the centre of the sphere. Then $\Delta U_{\alpha\alpha} \cong \Delta U_{\alpha\alpha}^{int}$, and $U_{\alpha\alpha}^{int}$ can be calculated as:

$$U_{\alpha\alpha}^{int} = \sum_{i=1}^{N} \sum_{j=1}^{N} q_{1,i} \cdot \phi_{2,j} + \sum_{i=1}^{N} \sum_{j=1}^{N} q_{1,i} \phi_{2,j}^{im} \qquad (17)$$

where i denotes a summation of the atoms of helix 1, and j of helix 2.

Figure 16 shows the variation of $U_{\alpha\alpha}^{int}$ as a function of $\Omega_{\alpha\alpha}$ for a rather small radius of the sphere, together with the value of $U_{\alpha\alpha}^{int}$ in vacuum. It appears that the interaction energy is considerably reduced by the surrounding solvent [similar results have been obtained, employing the method of Warwicker and Watson, 1982, by N. Rogers and M. Sternberg (personal communication)]. Nevertheless, even for this small radius, which is about 3 Å more than the outermost atoms of the helix backbone, the reduction of $U_{\alpha\alpha}^{int}$ is only a factor of 2–3, depending on the angle $\Omega_{\alpha\alpha}$. This means that the actual energies differences between helical arrangements as shown in Fig. 12 would still be highly significant. It may be pointed out that further decrease of the sphere radius leads to a rapid further reduction of $U_{\alpha\alpha}^{int}$ (W. G. J. Hol, unpublished; Rogers and Sternberg, personal communication) by another factor of 2–3. It is

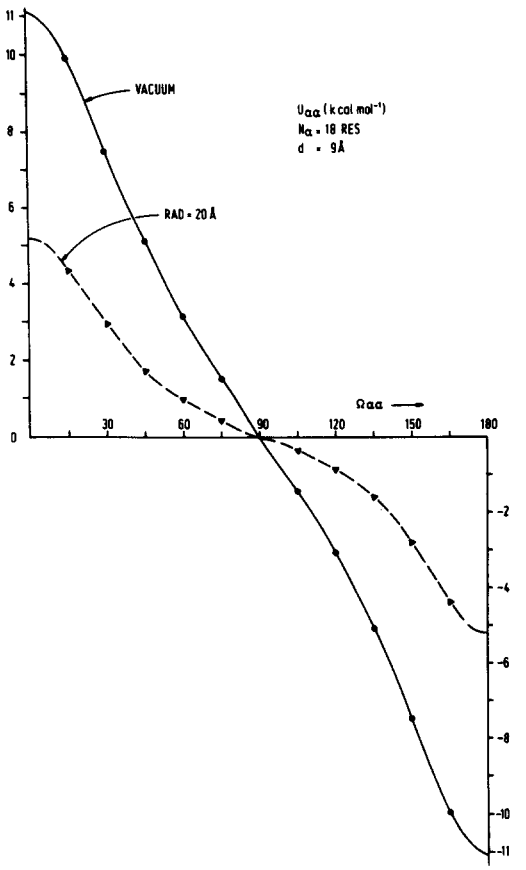


FIG. 16. Effect of the solvent reaction field on the interaction energy $U_{\alpha\alpha}$ of two helix backbone dipoles. Solid curve: results in vacuum; dashed curve: interaction energy incorporating the reaction field. The "solvation" and "self" energies of the helices [eqn (16)] have been ignored (see also text). The angle $\Omega_{\alpha\alpha}$ is defined in Fig. 10, as is the interhelical distance d. The calculations for the solid curve were carried out as explained in the legend of Fig. 11(a). For the calculation of the effect of the reaction field, Friedman's image charge method was used as explained in Fig. 15 and the text. The sphere had a radius of 20 Å and its origin was placed at the centre of gravity of *both* helices. The distance of the outer atom of the helix to the sphere centre was 16.4 Å, i.e. the reaction field sphere was 3.6 Å removed from this atom in order to simulate one hydration layer of solvent molecules surrounding both helices. The dielectric constant of the sphere was taken as unity; the magnitude of the image charge was calculated by the approximate equation shown in Fig. 15.

questionable, however if modelling the electrostatic properties of the first hydration layer of a protein by those of bulk water is physically meaningful.

Analogous calculations have been carried out for the $\beta\alpha\beta$-unit (W. G. J. Hol, unpublished). The reaction field appears to reduce the $\Delta U_{\beta\alpha\beta}$ approximately by a factor of 1.5. This is a smaller reduction, as expected, than for the interaction energy between two helices.

(d) *Dielectric effects of the protein atoms*

In the absence of molecular rotational freedom, the dielectric response of a medium is determined by the polarizabilities of the atoms. In protein molecules, the interior is densely packed (Richards, 1979) with seriously restricted motions as shown by the low temperature factors for internal residues obtained by crystallographic refinement procedures. Consequently, the dielectric constant of a medium with properties equivalent to that of a protein interior must be quite low. Pethig (1979) concludes that atomic polarizabilities of protein atoms would give rise to a dielectric constant of ~ 2. Electrostatic interactions between monopoles and dipoles in proteins have been calculated with dielectric "constants" ranging from 1 to 4 (see Hol *et al.*, 1978; Sheridan *et al.*, 1982; Ferro *et al.*, 1980; Hermans *et al.*, 1984; Jernigan *et al.*, 1980; Brant *et al.*, 1967; Brooks *et al.*, 1983; Van

Gunsteren *et al.*, 1983; Warwicker and Watson, 1982; Ralston and De Coen, 1974; Robson and Osguthorpe, 1979). Clearly, considerable uncertainty exists as to the value which should be taken for the dielectric screening factor due to the protein atomic polarizabilities. Making the assumption that the interior of a protein resembles, with respect to its *dielectric* properties, an average of solid acetamide with a dielectric constant of ~ 4 (Handbook of Chemistry and Physics, 1968–1969), and liquid benzene or hexane with dielectric constants of ~ 2 (Handbook of Chemistry and Physics, 1968–1969), then one arrives at an estimate of ~ 3 for the dielectric screening factor of the protein interior.

We may recall here that cooperative polarization effects of the peptide units in an α-helix (see Section II) cause an increase of the helix dipole moment by 25–50%. A similar increase of the dipole moment is probably operative in β-strands. It may therefore be concluded that these effects, which increase $\alpha\alpha$ and $\alpha\beta$ interactions by a factor of about 2, largely cancel the attenuation of the electrostatic interactions by the atomic polarizabilities of the rest of the protein atoms.

(e) *The importance of the protein hydration layer*

A particularly difficult subject is the dielectric screening effect due to the first hydration layer surrounding a protein molecule. A well-written and extensive review on the location and dynamics of water in protein systems has appeared recently (Edsall and McKenzie, 1983). It is clear that high resolution X-ray and neutron diffraction studies of proteins have detected numerous positions near protein surfaces where water molecules bind preferentially (Edsall and McKenzie, 1983). Uncertainty exists, however, as to the translational and rotational mobilities of these water molecules. This may be illustrated by the apparent contradiction between the conclusions reached by different authors. On the one hand, Rossky and Karplus (1979) deduce, from a molecular-dynamics study of a dipeptide analogue in water, that the polar (C=O, N–H) groups of the dipeptide have little influence on the mobility of the solvent but that near the non-polar methyl groups substantial hindrance of the translational and rotational motion of the water molecules occurs. On the other hand, Edsall and McKenzie (1983) conclude at the end of their review that water molecules adjacent to non-polar surface atoms remain highly mobile, but those hydrogen-bonded to polar groups are more retarded in mobility. Further studies on the mobility and dielectric properties of this layer of water molecules are crucial for a quantitative understanding of electrostatic interactions between α-helices in proteins, because of the observation (W. G. J. Hol, unpublished; Rogers and Sternberg, personal communication) that the dielectric screening factor for U_{zz} is very sensitive to the dielectric constant assumed for the first hydration shell. Clearly, more investigations into the mobilities of protein bound water molecules are of great importance for assessing electrostatic interactions in proteins. High resolution neutron diffraction studies might provide new insights as this technique is, in principle, capable of revealing the positions of hydrogen atoms of water molecules.

3. *Conclusion*

From the discussions above it is evident that many fundamental problems have still to be solved before the role of the electrostatic interactions between the peptide units of α-helices and parallel β-strands upon protein folding will be firmly established. It seems very likely however that proteins consisting of anti-parallel helical bundles are significantly stabilized by the favourable interaction between the α-helix dipoles [Fig. 13(a) (Hol *et al.*, 1981; Sheridan *et al.*, 1982)]. It is also probable that the α-helix dipole is the explanation for the observation by J. S. Richardson (1981, p. 190) that, in known protein structures, neighbouring helices with a small angle between their axes have a strong tendency to be anti-parallel, while "parallel" helices have much larger inter-axial angles.

As to the possible electrostatic stabilization of $\beta\alpha\beta$-units, much depends on the direction of the peptide dipole moment in parallel β-sheets. Accurate calculations of the direction and size of the dipole moments in such large systems remain a tremendous challenge for quantum chemists. The ultimate answer to the question of electrostatic interactions of solvated

proteins may come from molecular dynamics simulations, for long time periods and including atomic polarizabilities for the protein as well as for the water molecules.

VII. THE MOLECULAR PACKING IN CRYSTALS OF HELICAL OLIGOPEPTIDES

As the simulations alluded to above will take some time to be realized, it is worthwhile to search for simpler systems in which the electrostatic effect of the α-helix dipole may become evident. Crystals of uncharged α-helical molecules allow, in principle, an assessment of the effect of the α-helix dipole on helix packing. No effect of the surrounding solvent needs to be taken into account in such crystals. In addition, cooperative polarization within the helical peptide units and polarization effects of the remaining atoms cancel each other to a large extent [see Section VI.2.(d)] so that calculations with a dielectric constant of one, give a fair impression of the electrostatic energies involved.

In crystals, it seems reasonable to assume that the Van der Waals interactions will vary little in the various possible arrangements. This is borne out by the small differences in "packing energy" of polymorphic crystals observed by Bernstein and coworkers (Bernstein and Hagler, 1978; Hagler and Bernstein, 1978; Barr and Bernstein, 1982). Obviously, hydrogen bonds between helices can play an important role but the number of inter-helical hydrogen bonds which can be formed is, on the average, no source of bias for parallel, anti-parallel or any other mode of helix packing. It does, however, mean that firm conclusions can only be drawn when a large number of such crystals is considered. We shall see that this is, unfortunately, not yet possible.

The number of known crystal structures of helical oligopeptides, which are uncharged and contain no bulk solvent regions, is rather small. Although the helices, e.g. in the avian pancreatic polypeptide dimer (Blundel *et al.*, 1981) and in the mellitin tetramer (Terwilliger and Eisenberg, 1982) run clearly anti-parallel to each other, they fall outside the scope of this section because both are charged and contain aqueous solvent regions. Also the structure of alamethicin (Fox and Richards, 1982), which contains several charged residues, does not serve our purposes.

Fortunately, the structures of a number of α-aminoisobutyric acid (Aib) containing synthetic oligopeptides have become available recently and appear to be excellent objects for evaluating the effect of helix dipole interactions on helix packing. The three helical crystal structures of particular interest are:

(1) Boc-Aib-Pro-Val-Aib-Val-OMe (Francis *et al.*, 1982). The individual molecules fold in a partly 3_{10}, partly α-helical conformation and stack in a head-to-tail fashion along a direction parallel to the helix axis. The stacking of the helical columns is complex and has not yet been studied in detail. They are also very short helices—barely a single turn.

(2) Boc-Ala-Aib-Ala-Aib-OMe (Francis *et al.*, 1983) forms also columns of α-helices in a head-to-tail fashion. Although no calculations on dipole–dipole interactions have yet been performed, these columns seem to be arranged in an optimally anti-parallel fashion as shown in Fig. 17. Each column is in contact with six surrounding columns. Of the six, two have the same direction as that of the central one, whereas the remaining four propagate in the opposite direction.

(3) The undecapeptide Boc-Ala-Aib-Ala-Aib-Ala-Glu(OBz)-Ala-Aib-Ala-Aib-Ala-OMe also forms columns of α-helices (Butters *et al.*, 1981). The helix-dipole interactions in this crystal form have been analyzed (Hol and De Maeyer, 1984) and the results are presented in Table 12. It appears, not surprisingly, that the major interaction exists between two adjacent helices in a column. The remainder of the helical interactions is also favourable, and fluctuates considerably until about a 1000 "neighbouring" helices are included (Table 12). This table also shows that electrostatic effects in crystals of these uncharged molecules appear to be surprisingly large: the interaction energy of one helix dipole with its surrounding dipoles amounts to about 20 kcal mol^{-1}.

Because of the limited number of examples available, it is still too early to conclude that the α-helix dipole is in general an important factor in governing the packing of helical

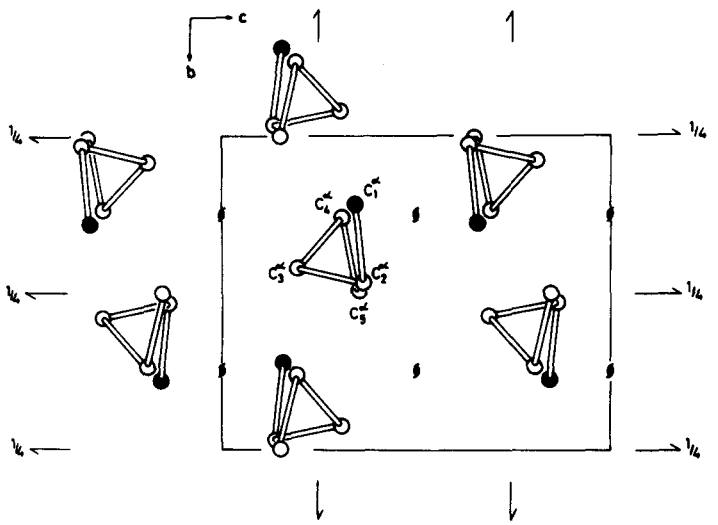


FIG. 17. Packing of helices in crystals of the pentapeptide Boc-Ala-Aib-Ala-Aib-Aib-OMe (Francis *et al.*, 1983). (Aib = α-amino isobutyric acid). Each helix shown is representing a column of helices, aligned in head-to-tail fashion, the axis of which is perpendicular to the plane of the paper. Each column is surrounded by six other helix columns, four of which run anti-parallel and two parallel to the central column. The N-terminal C^α-atoms are shown in black. (Adapted from Fig. 3 in Francis *et al.*, 1983.)

TABLE 12. ELECTROSTATIC ENERGY (BACKBONE ONLY) OF A "CENTRAL" HELIX IN THE CRYSTALS OF AN UNDECAPEPTIDE, WITH THE SURROUNDING HELICES—AS FUNCTION OF CUTOFF RADIUS (HOL AND DE MAEYER, 1984). A "SURROUNDING" HELIX IS INCORPORATED IF THE CENTRE OF GRAVITY OF THE HELIX LIES WITHIN THE CUTOFF RADIUS FROM THE CENTRE OF GRAVITY OF THE CENTRAL HELIX. CHARGE DISTRIBUTION AS IN FIG 1. NO DIELECTRIC SCREENING EFFECTS TAKEN INTO ACCOUNT

Radius	Number of Helices	Electrostatic Interaction Energy (kcal mol^{-1})
25	34	−23.1
30	64	−21.8
35	94	−19.5
40	148	−17.0
45	218	−21.1
50	292	−21.6
55	390	−19.1
60	492	−18.8
65	649	−19.7
70	800	−20.2
75	988	−20.1

oligopeptides in the crystalline state. The first investigations of this aspect of helix–dipole interactions, however, seem to be in agreement with such a proposition.

VIII. THE POSSIBLE ROLE OF THE α-HELIX DIPOLE IN PHOTOSYNTHESIS

In the last two sections of this review, we will look briefly into two different phenomena in biochemistry for which models have been proposed involving the electrostatic effect of the α-helix dipole. In photosynthesis the crucial initial events are the following (see e.g. Bearden and Malkin, 1975; Gerischer and Katz, 1979): via a light-harvesting system, a quantum of light reaches a "special pair" of chlorophyll molecules in a so-called "reaction centre" and causes an exceedingly rapid separation of charges across a membrane. The energy thus stored is, subsequently, *via* a complex series of reactions, used for the generation of ATP. This charge separation in one of the most important biological systems has received considerable

attention from experimentalists (see e.g. Gerischer and Katz, 1979) and theoreticians (see e.g. Warshel and Schlosser, 1981). Here we will briefly describe a model of the photosynthetic reaction centre which has been proposed by Rypma, Van Duijnen, Hol and Zandstra (to be submitted).

This model is based upon the following observations:

(i) photosynthetic membranes can be oriented in a magnetic field (Geacintov *et al.*, 1972a,b; Breton, 1974; Clement-Metral, 1975). Rypma *et al.* propose that this alignment is due to an interaction of the magnetic field with α-helices. This interaction has nothing to do with the electrical dipole of an α-helix but arises from the diamagnetic anisotropy of the peptide units (Worcester, 1978; Pauling, 1979). The result is that α-helices align themselves with their axes parallel to the magnetic field (Worcester, 1978)—independent of the direction of the electrical dipole. As photosynthetic membranes are aligned with their surface perpendicular to the magnetic field (Geacintov *et al.*, 1972a,b; Breton, 1974; Clement-Metral, 1975), this suggests that α-helices run also perpendicular to the plane of the membrane;

(ii) the spectrum of the special pair shows a red shift compared to spectra of chlorophyll molecules free in solution (Sauer, 1978). It is proposed by Rypma *et al.* that this spectral shift is due to an electrical field which arises from α-helices surrounding the special pair in the photosynthetic reaction centre as shown in Fig. 18.

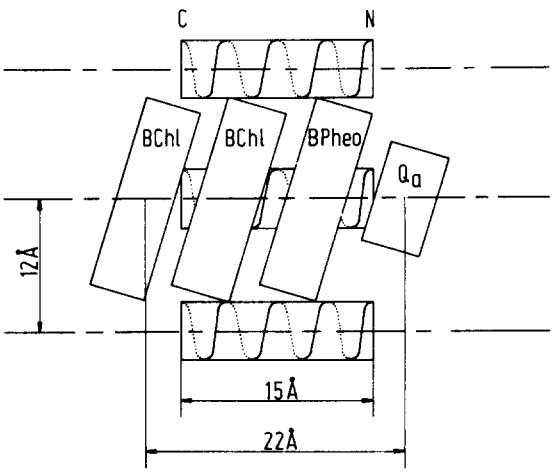


FIG. 18. Schematic drawing of the model of a photosynthetic reaction centre proposed by Rypma *et al.* (in preparation). The electron transferring units depicted consist of the bacteriochlorophyll dimer ($BChl_c$), a bacteriopheophytin (BPheo) and a quinone molecule (Q_a). The units are surrounded by parallel helices, in this case four of which the one closest to the viewer is not shown, which are part of the reaction centre protein. In this case the N-termini of the helices are at the right side of the figure and the length of the helices is 15 Å.

In this model, several, e.g. four, parallel helices surround two chlorophyll molecules. When a light quantum reaches the special pair, an electron is excited and the permanent field of the helices guides the electron towards other receptor molecules. The electrical field along a number of lines parallel to the helix axis is shown in Fig. 19. Even with a dielectric screening factor of four, a considerable field exists when several such helices are considered as in the model of Fig. 18. Such a low screening factor seems reasonable for processes occurring in photosynthetic membrane with a thickness of approximately 70 Å (Miller, 1982). It should be mentioned that the parallel arrangement of helices proposed may well be a subset of a cluster of largely anti-parallel helices which go "up and down" in the membrane as is also observed in the bacteriorhodopsin membrane (Engelman *et al.*, 1980).

Obviously, this model is a rather speculative one. Nevertheless, it will be highly interesting to see whether it approaches the actual situation in photosynthetic reaction centres. This will be known in due course and in magnificent detail from crystallographic studies which are currently under way (Michel, 1982).

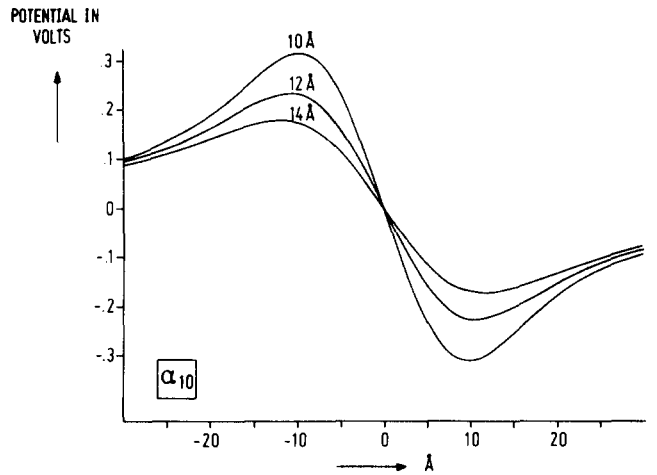


FIG. 19. Electrostatic potentials due to the backbone peptide units of a ten-residue α-helix along lines which run parallel to the helix axis at distances of 10, 12 and 14 Å from this axis. The partial atomic charges used are those of Fig. 1, and no dielectric screening was taken into account. The N-terminus of the helix is at the left-hand side of the picture. The origin of the horizontal axis is the projection of the centre of gravity of the helix onto the line along which the potential is calculated. (Taken from Rypma *et al.*, in preparation.)

IX. PROTON AND ION CONDUCTION ACROSS LIPID BILAYERS

1. *Proton and Ion Conduction Along the Helix Axis*

Proton and ion transport through biological membranes are essential features of several key processes of living organisms, such as oxidative phosphorylation and nerve conduction. Numerous different mechanisms have been proposed for the transport processes involved and it would not be surprising if nature would indeed employ a variety of different principles to control membrane transport. We will limit ourselves here to two proposed mechanisms in which the electrical properties of the α-helix dipole play an important role.

Van Duijnen and Thole (1981) have investigated, by means of *ab initio* quantum mechanical calculations, the energetics of a proton travelling along the axis of an α-helix. For seven positions of the proton, the system, consisting of a helical octa-alanine molecule plus the proton, was allowed to relax electronically. As shown in Fig. 20, an energy minimum is observed for the proton near the C-terminus of the helix, whereas the barrier for obtaining that position from the N-terminus is $\sim$20 kcal mol^{-1}. No relaxation of the geometry has

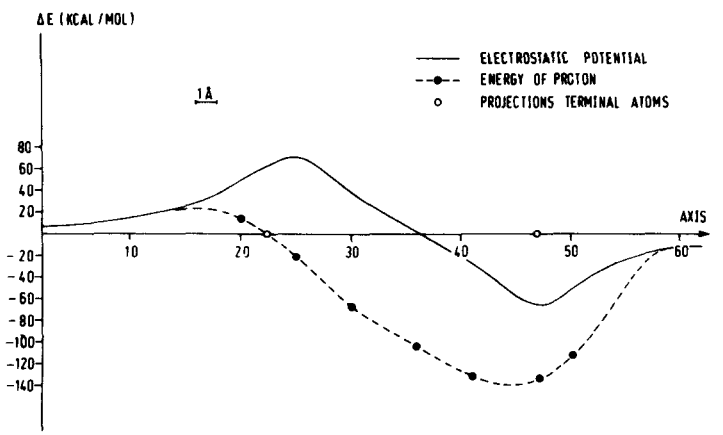


FIG. 20. Electrostatic potentials on the axis of an octa-alanine α-helix. The solid line—representing the potential of the undisturbed charge distribution in the helix—was obtained by calculating this potential at 65 points, spaced 1 bohr ($\approx$0.53 Å) apart. The dashed line gives the relative energy of a proton as obtained from separate molecular orbital calculation for the seven positions of the proton indicated with black dots. (Fig. from Van Duijnen and Thole, 1981.)

been carried out, so one may expect that the actual energy barrier will be substantially lower. The authors suggest that the internal electric field of the α-helix could be an essential part of proton- and even ion-transducing systems. The same authors mention, furthermore, the possibility that the field of the α-helix may also play a role in ion transport alongside α-helices.

Very similar suggestions have been made by Ovchinnikov and Ukrainskii (1978, 1979) on the basis of straightforward electrostatic approximations taking the solvent reaction field into account by the method of image charges. These authors propose that electrons and protons may travel along the axis of α-helices, and that such processes occur in the mitochondrial proton pump. For ions as K^+ and Na^+ steric hindrance will occur when these move inside the helix but Ovchinnikov and Ukrainskii discuss the possibility of a cavity surrounded by α-helices in a parallel orientation as a membrane ion channel. This interesting possibility is closely related to the model, discussed in the previous section, for charge separation in photosynthesis. An important point still to be solved appears to be the mechanism by which the protons and ions overcome the barriers when entering and leaving the helix field.

2. *Voltage-Dependent Pore-Formation: a Flip-Flopping Helix Model*

An intriguing membrane process is the voltage-dependent pore-formation in lipid bilayers by helical oligopeptides such as alamethicin and mellitin. Conductance measurements reveal that ions can only cross the membrane when a voltage is applied, implying that the external electrical field somehow affects the arrangement of the helices in the membrane (Mueller and Rudin, 1968; Gordon and Haydon, 1972; Eisenberg *et al.*, 1973; Hanke and Boheim, 1980; Hanke *et al.*, 1983). Several explanations at the molecular level have recently been put forward to explain this phenomenon (Fox and Richards, 1982; Boheim *et al.*, 1983). Only the proposal of Boheim *et al.* (1983) will be considered here further as the α-helix dipole is a central part of it. This mechanism, as shown in Fig. 21, involves a flip-flop of the helical molecules in the membrane as a response to the applied field. One would like to have an insight into the energetics of the proposed process. A first glimpse of this may be obtained by the following calculation. The energy U of a dipole $\boldsymbol{\mu}$ in an electrical field $\mathbf{E}$ is: $U = -\boldsymbol{\mu} \cdot \mathbf{E}$. The energy difference between the most favourable and most unfavourable position of an α-helix dipole in a homogeneous field is therefore: $\Delta U = 2\mu \cdot E$. For the nonadecapeptide alamethicin in helical conformation a value of $\mu \approx 70$ D is expected on the basis of the approximate alignment of peptide dipole moments (Section I). Moreover, similar values have been measured for the alamethicin dipole in octanol/dioxane (Yantorno *et al.*, 1982) and in ethanol/dioxane (Schwarz and Sarko, 1982) mixtures. E is in the order of 120 mV over a distance of approximately 30 Å (Mueller and Rudin, 1968; Gordon and Hayden, 1972; Eisenberg *et al.*, 1973; Hanke and Boheim, 1980; Hanke *et al.*, 1983), i.e. $E = 120/30$ mV.Å^{-1} $= 4$ mVÅ^{-1}. Then, $\Delta U = 2 \times 14$ e.Å $\times 4$ mV.Å^{-1} $\simeq 110$ meV $\simeq 2.5$ kcal mol^{-1} $\simeq 4$ kT. It appears, therefore, quite possible that flip-flop of helices may occur in this voltage-dependent conductance process.

It may be pointed out that in the membrane pores of Boheim's model [Fig. 21(e)] the field originating from the parallel helices counteracts the external electrical field. As it is likely that the ion channels are filled with solvent molecules, the field due to the helix dipoles will in such cases hardly affect the ion transport process due to the efficient dielectric screening effect of water. In other cases, however, where the parallel helices are embedded in hydrophobic environments, the intrinsic electrical field of the helix dipoles may be of crucial importance, such as in the proposed model of the photosynthetic reaction centre (Figs. 18,19). The degree of dielectric screening by the solvent is as crucial a factor in these membrane processes as it appeared to be in protein folding [Section VI.2.(d)].

X. CONCLUSIONS

Perutz (1978) has shown that electrostatic interactions between charged amino acids are important for the functioning of many protein molecules. From the present review, it appears that the α-helix dipole is another important source of electrical fields generated by proteins.

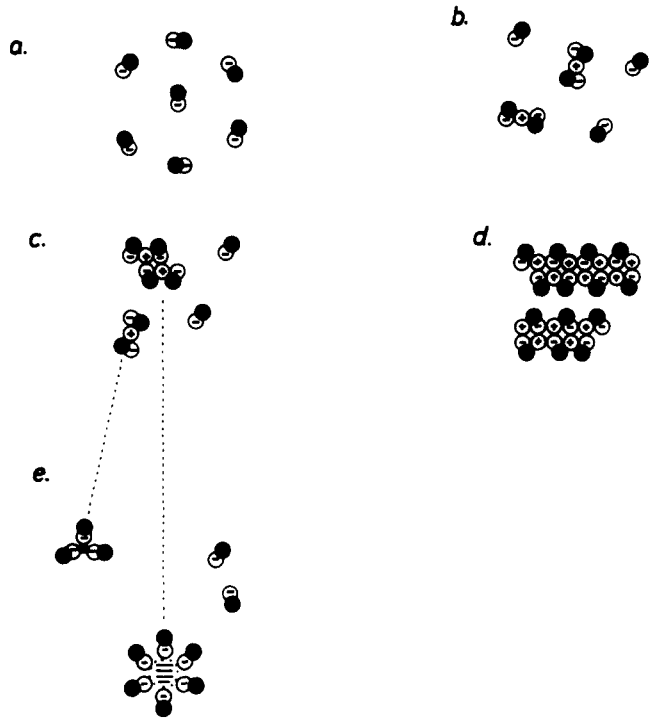


FIG. 21. Alamethicin pore formation: voltage-dependent flip-flop of α-helix dipoles in membranes as proposed by Boheim *et al.* (1983). The membrane (in the plane of the paper) is separating the system in two compartments. Let alamethicin be added to only one aqueous compartment at the *cis*-side of the membrane.

(a) The oligopeptide diffuses into the membrane and is incorporated with its α-helical N-terminal part. The N-terminus points to the *trans*-side and the C-terminus to the *cis*-side, indicated by ⊖ (*n*-dipole) as this is the negative end of the helix dipole. The C-terminal non-helical residues are indicated by a black dot. Due to the unfavourable (electrostatic) interaction between parallel helix dipoles, the alamethicin molecules tend to be present as monomers.

(b) If an *n*-dipole flips, it becomes a *p*-dipole ⊕ with the positive end pointing to the *cis*-side. The *p*-dipole and *n*-dipole form trimers because of favourable electrostatic interactions.

(c) After sufficient time also larger aggregates form and a temperature-dependent distribution of aggregate-sizes will be obtained.

(d) In case of a frozen lipid matrix extended structures may build up in form of regular linear arrays of *p*- and *n*-dipoles. Such a linear array may be the result of the asymmetry due to the C-terminal non-helical part of the molecules.

(e) When a potential difference is applied across the membrane, with the more positive potential chosen on the *cis*-side, *p*-dipoles may flip to become *n*-dipoles. Within an aggregate, neighbouring *n*-dipoles repel each other thus creating a hole. This hole may be filled by lipid or by solvent molecules. The number of helices surrounding the hole may vary and a pore of variable diameter is obtained. (Fig. from Boheim *et al.*, 1983.)

This is shown by the following:

(i) Ten enzymes are known where the active site is close to a helix terminus. This means that about one out of every four enzymes with known three-dimensional structure is employing the helix field for enhancing reaction rates;

(ii) About twenty phosphate binding helices have been discovered among the proteins with known tertiary structure. Phosphate moieties of low molecular weight ligands, when bound by proteins, are in ~60% of the known examples interacting with the positive end of a helix dipole;

(iii) On a statistical basis, the α-helix dipole in globular proteins appears to interact favourably with positively and negatively charged amino acid side chains;

(iv) In globular proteins, the α-helix dipole is the probable cause of the preference for anti-parallel alignments of neighbouring helices with small inter-axial angles.

In addition, the α-helix dipole is currently a source of inspiration for proposing models for several intriguing biochemical processes, such as photosynthesis and membrane transport of

ions. Numerous aspects of electrostatic interactions in proteins surrounded by the solvent water are still poorly understood at the present time, however. Therefore, the value of the hypothesis that the preferred $\alpha\alpha$, $\beta\alpha\beta$ and $\beta\beta$ folding patterns in proteins are due to the electrostatic interactions between peptide dipoles in α-helices and β-strands cannot be assessed at this moment. In simpler systems, such as water-free crystals of helical oligopeptides, the helix dipole probably plays an important role in the packing of the helices. Unfortunately, the number of crystal structures known for this type of molecules is too small to draw firm conclusions yet.

In conclusion, it can be said that the α-helix, proposed by Pauling *et al.* (1951) many years before any protein structure had been elucidated, has not only been actually observed in numerous protein molecules but contains also a considerable dipole moment which is employed in a wide variety of biochemical processes.

ACKNOWLEDGEMENTS

I am grateful to Dr. J. Drenth for his continuous interest for my work on the α-helix dipole. I am also indebted to Drs. H. J. C. Berendsen, P.Th. van Duijnen, C. Sander, R. K. Wierenga, Y. Rypma and G. Boheim for many stimulating discussions. I like to thank Louis Halie and Gerrit van der Plaats for carrying out calculations on actual proteins, Roelof Kloosterman and Myra Swarte for preparation of the drawings, Roeline Hogenkamp for excellent assistance in preparing the manuscript and Dr. J. N. Jansonius (Biozentrum, Basel) for suggestions for improvement of this review. The generous support of the Dutch Organization for Chemical Research (SON) has been crucial for creating a stimulating scientific climate in our research group. Finally, I like to mention with gratitude that it was Dr. J. Kraut, of the University of California in San Diego, who made the initial suggestion to look into the possible importance of the α-helix dipole for the functioning of proteins.

REFERENCES

ABAD-ZAPATERO, C., ABEL-MEGUID, S. S., JOHNSON, J. E., LESLIE, A. G. W., RAYMENT, I., ROSSMANN, M. G., SUCK, D. and TSUKIHARA, T. (1980) *Nature, Lond.* **286,** 33–39.

ADAMS, D. J. (1981) *Nature, Lond.* **293,** 447–448.

ALBER, T., BANNER, D. W., BLOOMER, A. C., PETSKO, G. A., PHILLIPS, D. C., RIVERS, P. S. and WILSON, I. A. (1981) *Phil. Trans. R. Soc. Lond.* **B293,** 159–171.

ANDERSON, W. F., OHLENDORF, D. H., TAKEDO, Y. and MATTHEWS, B. W. (1981) *Nature, Lond.* **290,** 754–758.

APPLEQUIST, J. and MAHR, T. G. (1966) *J. Am. chem. Soc.* **88,** 5419–5429.

BAKER, E. N. (1980) *J. molec. Biol.* **141,** 441–484.

BANNER, D. W., BLOOMER, A. C., PETSKO, G. A., PHILLIPS, D. C., POGSON, C. I. and WILSON, I. A. (1975) *Nature, Lond.* **255,** 609–614.

BAR, I. and BERNSTEIN, J. (1982) *J. phys. Chem.* **86,** 3223–3231.

BARNES, P., BLISS, D. V., FINNEY, J. L. and QUINN, J. E. (1980) *Faraday Disc. Roy. Soc. Chem.* **69,** 210–220.

BARNES, P., FINNEY, L. J., NICHOLAS, J. D. and QUINN, J. E. (1979) *Nature,* **282,** 459–464.

BEARDEN, A. and MALKIN, R. (1975) *Q. Rev. Biophys.* **7,** 131–177.

BERNSTEIN, J. and HAGLER, A. T. (1978) *J. Am. chem. Soc.* **100,** 673–681.

BERNSTEIN, F. C., KOETZLE, T. F., WILLIAMS, G. J. B., MEYER, E. F., BRICE, M. D., RODGERS, J. R., KENNARD, O., SHIMANOUCKI, T. and TASUMI, M. (1977) *J. molec. Biol.* **112,** 535–542.

BHAT, T. N., BLOW, D. M., BRICK, P. and NYBORG, J. (1982) *J. molec. Biol.* **158,** 699–709.

BIESECKER, G., IEUAN HARRIS, J., THIERRY, J. C., WALKER, J. E. and WONACOTT, A. J. (1977) *Nature* **266,** 328–333.

BLAGDON, D. E. and GOODMAN, M. (1975) *Biopolymers* **14,** 241–245.

BLOOMER, A. C., CHAMPNESS, J. N., BRICOGNE, G., STADEN, R. and KLUG, A. (1978) *Nature, Lond.* **276,** 362–373.

BLUNDELL, T. L., PITTS, J. E., TICKLE, I. J., WOOD, S. P. and WU, C.-W. (1981) *Proc. natn. Acad. Sci. U.S.A.* **78,** 4175–4179.

BOHEIM, G., HANKE, W. and JUNG, G. (1983) *Biophys. struct. Mech.* **9,** 181–191.

BONOMI, F., PAGANI, S. and CERLETTI, P. (1977a) *FEBS Lett.* **84,** 149–152.

BONOMI, F., PAGANI, S., CERLETTI, P. and CANELLA, C. (1977b) *Eur. J. Biochem.* **72,** 17–24.

BÖTCHER, C. J. F. (1973) *Theory of Electric Polarization,* Vol. I, 2nd Edn. Elsevier, Amsterdam.

BRÄNDÉN, C.-I. (1980) *Q. Rev. Biophys.* **13,** 317–338.

BRANDENBURG, N. P., DEMPSEY, S., DIJKSTRA, B. W., LIJK, L. J. and HOL, W. G. J. (1981) *J. appl. Crystallogr.* **14,** 274–279.

BRANT, D. A., MILLER, W. G. and FLORY, P. J. (1967) *J. molec. Biol.* **23,** 47–65.

BRETON, J. (1974) *Biochem. Biophys. Res. Commun.* **59,** 1011–1017.

BROER, R., VAN DUIJNEN, P. TH. and NIEUWPOORT, W. C. (1976) *Chem. Phys. Lett.* **42,** 525–529.

BROOKS, B. R., BRUCCOLERI, R. E., OLAFSON, B. D., STATES, D. J., SWAMINATHAN, S. and KARPLUS, M. (1983) *J. comp. Chem.* **4,** 187–217.

Brown, N. L., Ford, S. J., Pridmore, R. D. and Fritzinger, D. C. (1983) *Biochemistry* **22,** 4089–4095.
Butters, T., Hütter, P., Jung, G., Pauls, N., Schmitt, H., Sheldrick, G. M. and Winter, W. (1981) *Angew. Chem.* (Eng. Edn.) **20,** 889–890.
Cantor, C. R. and Schimmel, P. R. (1980) *Biophysical Chemistry*, Part I, p. 263. W. H. Freeman, San Francisco.
Chothia, C., Levitt, M. and Richardson, D. (1977) *Proc. natn. Acad. Sci. U.S.A.* **74,** 4130–4134.
Chotia, C. and Janin, J. (1981) *Proc. natn. Acad. Sci. U.S.A.* **78,** 4146–4150.
Chothia, C., Levitt, M. and Richardson, D. (1981) *J. molec. Biol.* **145,** 215–250.
Chothia, C. (1973) *J. molec. Biol.* **75,** 295–302.
Chou, K.-C., Pottle, M., Némethy, G., Ueda, Y. and Scheraga, H. A. (1982) *J. molec. Biol.* **162,** 89–112.
Chou, K.-C., Némethy, G. and Scheraga, H. A. (1985) *J. molec. Biol.* **168,** 389–407.
Chou, P. Y. and Fasman, G. D. (1974) *Biochemistry* **13,** 211–222.
Clement-Metral, J. D. (1975) *FEBS Lett.* **50,** 257–260.
Cotton, F. A., Hazen, E. E. and Legg, M. J. (1979) *Proc. natn. Acad. Sci. U.S.A.* **76,** 2551–2555.
Coulson, C. A. and Eisenberg, D. (1966a) *Proc. R. Soc.* **A291,** 445–453.
Coulson, C. A. and Eisenberg, D. (1966b) *Proc. R. Soc.* **A291,** 454–459.
Crawford, J. L., Lipscomb, W. N. and Schellman, C. G. (1973) *Proc. natn. Acad. Sci. U.S.A.* **70,** 538–542.
Dalziel, K., McFerran, N. V. and Wonacott, J. (1981) *Phil. Trans. R. Soc. Lond.* **B293,** 105–118.
Dijkstra, B. W., Renetseder, R., Kalk, K. H., Hol, W. G. J. and Drenth, J. (1983a) *J. molec. Biol.* **168,** 163–179.
Dijkstra, B. W., Weijer, W. J. and Wierenga, R. K. (1983b) *FEBS Lett.* **164,** 25–27.
Drenth, J., Hol, W. G. J., Jansonius, J. N. and Koekoek, R. (1971a) *eur. J. Biochem.* **26,** 177–181.
Drenth, J., Jansonius, J. N., Koekoek, R. and Wolters, B. G. (1971b) *Adv. Prot. Chem.* **25,** 79–115.
Drenth, J., Swen, H. M., Hoogenstraaten, W. and Sluyterman, L. E. A. E. (1975) *Proc. Koninkl. Akad. Wetensch.* **C78,** 104–110.
Drenth, J., Kalk, K. H. and Swen, H. M. (1976) *Biochemistry* **15,** 3731–3738.
Duisterwinkel, F. J., Kraal, B., De Graaf, M. T., Talens, A., Bosch, L., Swart, G. W. M., Parmeggiani, A., La Cour, T. F. M., Nyborg, J. and Clark, B. F. C. (1984) *EMBO J.* **3,** 113–120.
Dyke, T. R. and Muenter, J. S. (1974) *J. chem. Phys.* **60,** 2929–2930.
Edsall, J. T. and McKenzie, H. A. (1983) *Adv. Biophys.* **16,** 53–183.
Ehrenson, S. (1976) *J. Am. chem. Soc.* **98,** 7510–7514.
Eisenberg, M., Hall, J. E. and Mead, C. A. (1973) *J. Membr. Biol.* **7,** 325–344.
Eisenberg, D. and Kauzmann, W. (1969) *The Structure and Properties of Water*, Clarendon Press, Oxford.
Eklund, H., Samama, J.-P., Wallén, L., Brändén, C.-I., Åkeson, Å. and Jones, A. T. (1981) *J. molec. Biol.* **146,** 561–587.
Engelman, D. M., Henderson, R., McLachlan, A. D. and Wallace, B. A. (1980) *Proc. natn. Acad. Sci. U.S.A.* **77,** 2023–2027.
Entsch, B., Ballou, D. P. and Massey, V. (1976) *J. biol. Chem.* **251,** 2550–2563.
Epp, O., Ladenstein, R. and Wendel, A. (1983) *Eur. J. Biochem.* **133,** 51–69.
Evans, P. R. and Hudson, P. J. (1979) *Nature* **279,** 500–504.
Ferro, D. R., McQueen, J. E., McCown, J. T. and Hermans, J. (1980) *J. molec. Biol.* **136,** 1–18.
Fersht, A. (1977) *Enzyme Structure and Mechanism*, W. H. Freeman and Co., Reading.
Filman, D. J., Bolin, J. T., Matthews, D. A. and Kraut, J. (1982) *J. Biol. Chem.* **257,** 13663–13672.
Finney, J. L., Gellatly, B. J., Golton, I. C. and Goodfellow, J. (1980) *Biophys. J.* **32,** 17–33.
Flashner, M. S. and Massey, V. (1974) In *Molecular Mechanisms of Oxygen Activation*, (ed. O. Hayaishi), pp. 245–284, Academic Press, New York.
Ford, G. C., Eichele, G. and Jansonius, J. N. (1980) *Proc. natn. Acad. Sci. U.S.A.* **77,** 2559–2563.
Forstom, J. W., Zakowski, J. J. and Tappel, A. L. (1978) *Biochemistry* **17,** 2639–2644.
Fox, B. S. and Walsh, C. T. (1983) *Biochemistry* **22,** 4082–4088.
Fox, R. O. and Richards, F. M. (1982) *Nature, Lond.* **300,** 325–330.
Francis, A. K., Iqbal, M., Balaram, P. and Vijayan, M. (1982) *J. Chem. Soc. Perkin Trans.* **2,** 1235–1239.
Francis, A. K., Iqbal, M., Balaram, P. and Vijayan, M. (1983) *Biopolymers* **22,** 1499–1505.
Friedman, H. L. (1975) *Molec. Phys.* **29,** 1533–1543.
Geacintov, N. E., Van Nostrand, F., Pope, M. and Tinkel, J. B. (1972a) *Biochim. biophys. Acta* **226,** 486–491.
Geacintov, N. E., Van Nostrand, F., Becker, J. F. and Tinkel, J. B. (1972b) *Biochim. biophys. Acta* **267,** 65–79.
Gerischer, H. and Katz, J. J. (eds.) (1979) *Light-Induced Charge Separation in Biology and Chemistry*, Verlag Chemie, Berlin.
Goodfellow, J. M. (1982) *Proc. natn. Acad. Sci. U.S.A.* **79,** 4977–4979.
Gordon, L. G. M. and Haydon, D. A. (1972) *Biochim. biophys. Acta* **255,** 1014–1018.
Grau, U. M., Trommer, W. E. and Rossmann, M. G. (1981) *J. molec. Biol.* **151,** 289–307.
Hagler, A. T. and Bernstein, J. (1978) *J. Am. chem. Soc.* **100,** 6349–6354.
Hagler, A. T. and Lapiccirella, A. (1976) *Biopolymers* **15,** 1167–1200.
Handbook of Chemistry and Physics, 49th Edn. (1968–1969) pp. E 58–61. CRC Press, Cleveland Ohio.
Hanke, W. and Boheim, G. (1980) *Biochim. biophys. Acta* **596,** 456–462.
Hanke, W., Methfessel, C., Wilmsen, H.-U., Katz, E., Jung, G. and Boheim, G. (1983) *Biochim. biophys. Acta* **727,** 108–114.
Harrison, S. C., Olson, A. J., Schutt, C. E. and Winkler, F. K. (1978) *Nature, Lond.* **276,** 368–373.
Hartmann, F. C., La Muraglia, G. M., Tomozawa, Y. and Wolfenden, R. (1975) *Biochemistry* **14,** 5274–5279.
Heinemann, U. and Saenger, W. (1982) *Nature, Lond.* **299,** 27–31.
Hermans, J., Berendsen, H. J. C., Van Gunsteren, W. F. and Postma, J. P. M. (1984) *Biopolymers*, in press.
Hill, E., Tsernoglou, D., Webb, L. and Banaszak, L. J. (1972) *J. molec. Biol.* **72,** 577–591.
Hol, W. G. J., Halie, L. M. and Sander, C. (1981) *Nature* **294,** 532–536.
Hol, W. G. J. and De Maeyer, M. C. H. (1984) *Biopolymers* **23,** 809–817.
Hol, W. G. J., Van Duijnen, P. Th. and Berendsen, H. J. C. (1978) *Nature* **273,** 443–446.
Hol, W. G. J. and Wierenga, R. K. (1984) In *Drug action and X-ray Crystallography: Current Perspectives* (eds. A. S. Horn and C. J. de Ranter) pp. 151–168. Clarendon Press, Oxford.

HOLMGREN, A., SÖDERBERG, B.-O., EKLUND, H. and BRÄNDÉN, C.-I. (1975) *Proc. natn. Acad. Sci U.S.A.* **72,** 2305–2309.
HONZATKO, R. B. and LIPSCOMB, W. N. (1982) *J. molec. Biol.* **160,** 265–286.
HONZATKO, R. B., CRAWFORD, J. L., MONACO, H. L., LADNER, J. L., EWARDS, B. F. P., EVANS, D. R., WARREN, S. G., WILEY, D. C., LADNER, R. C. and LIPSCOMB, W. N. (1982) *J. molec. Biol.* **160,** 219–263.
JANIN, J. and CHOTHIA, C. (1980) *J. molec. Biol.* **143,** 95–128.
JENKINS, J. A., JOHNSON, L. N., STUART, D. I., STURA, E. A., WILSON, K. S. and ZANOTTI, G. (1981) *Phil. Trans. R. Soc. Lond.* **B293,** 23–41.
JERNIGAN, R. L., MIYAZAWA, S. and SZU, S. C. (1980) *Macromolecules* **13,** 518–525.
JOHANNIN, G. (1979) In *Simulation of Enzyme Catalysis* (eds. P. TH. VAN DUIJNEN and W. G. J. HOL) pp. 124–140. CECAM, Orsay, France.
JOHNSON, L. N., JENKINS, J. A., WILSON, K. S., STURA, F. A. and ZANOTTI, G. (1980) *J. molec. Biol.* **140,** 565–580.
KALLIS, G.-B. and HOLMGREN, A. (1980) *J. biol. Chem.* **255,** 10261–10265.
KAUZMANN, W. (1959) *Adv. Prot. Chem.* **14,** 1–63.
KELLY, J. A., MOEWS, P. C., KNOX, J. R., FRÈRE, J.-M. and GHUYSEN, J.-M. (1982) *Science* **218,** 479–481.
KRAUT, J. (1977) *An. Rev. Biochem.* **46,** 331–358.
LEBODIA, L., HATADA, M. H., TULINSKY, A. and MAVRIDIS, I. M. (1982) *J. molec. Biol.* **162,** 445–458.
LEVITT, M. and CHOTHIA, C. (1976) *Nature* **261,** 552–558.
LIFSON, S., HAGLER, A. T. and DAUBER, P. (1979) *J. Am. chem. Soc.* **101,** 5111–5121.
LILJAS, L., UNGE, T., JONES, T. A., FRIDBORG, K., LÖVGREN, S., SKOGLUND, U. and STRANDBERG, B. (1982) *J. molec. Biol.* **159,** 93–108.
LINDQUIST, Y. and BRÄNDÉN, C.-I. (1980) *J. molec. Biol.* **143,** 201–211.
LUDWIG, M. L., PATTRIDGE, K. A., SMITH, W. W., JENSEN, L. H. and WATENPAUGH, K. D. (1982) In *Flavins and Flavoproteins* (eds. V. MASSEY and C. H. WILLIAMS, JR.) pp. 19–27. Elsevier North Holland.
MASSEY, V. and HEMMERICH, P. (1975) *The Enzymes* **12,** 191–252.
MATSSUURA, Y., KUSUNOKI, M., HARADA, W., TANAKA, N., IGA, Y., YASUOKA, N., TODA, H., NARITA, K. and KAKUDO, M. (1980) *J. Biochem.* **87,** 1555–1558.
MATTHEWS, D. A., ALDEN, R. A., BIRKTOFT, J. J., FREER, S. T. and KRAUT, J. (1975) *J. biol. Chem.* **18,** 7120–7126.
MAUGUEN, Y., HARTLEY, R. W., DODSON, E. J., DODSON, G. G., BRICOGNE, G., CHOTHIA, C. and JACK, A. (1982) *Nature, Lond.* **297,** 162–164.
MAVRIDIS, I. M., HATADA, M. H., TULINSKY, A. and LEBIODA, L. (1982) *J. molec. Biol.* **162,** 419–444.
MCKAY, D. B., WEBER, I. T. and STEITZ, T. A. (1982) *J. biol. Chem.* **257,** 9518–9524.
MCKAY, D. B. and STEITZ, T. A. (1981) *Nature, Lond.* **290,** 744–749.
MCPHERSON, A., JURNAK, F. A., WANG, A. H. J., MOLINEUX, I. and RICH, A. (1979) *J. molec. Biol.* **134,** 379–400.
MEIROVITCH, H. and SCHERAGA, H. A. (1980) *Macromolecules* **13,** 1406–1414.
MICHEL, H. (1982) *J. molec. Biol.* **158,** 567–572.
MILLER, K. R. (1982) *Nature, Lond.* **300,** 53–55.
MOMANY, F. A., MCGUIRE, R. F., BURGESS, A. W. and SCHERAGA, H. A. (1975) *J. phys. Chem.* **22,** 2361–2381.
MORAS, D., OLSEN, K. W., SABESAN, M. N., BUEHNER, M. N., FORD, G. C. and ROSSMANN, M. G. (1975) *J. Biol. Chem.* **250,** 9137–9162.
MUELLER, P. and RUDIN, D. O. (1968) *Nature, Lond.* **217,** 713–719.
MUIRHEAD, H. (1983) *Trends biochem. Sci.* **8,** 326–330.
MÜLLER, F., HEMMERICH, P., EHRENBERG, A., PALMER, G. and MASSEY, V. (1970) *Eur. J. Biochem.* **14,** 185–196.
NAKAMURA, K. T., IWAHASHI, K., YAMAMOTO, Y., IITAKA, Y., YOSHIDA, N. and MITSUI, Y. (1982) *Nature, Lond.* **299,** 564–566.
NISHIKAWA, K. and SCHERAGA, H. A. (1976) *Macromolecules* **9,** 395–407.
NORTHRUP, S. H., PEAR, M. R., MORGAN, J. D., MCCAMMON, J. A. and KARPLUS, M. (1981) *J. molec. Biol.* **153,** 1087–1109.
OHLENDORF, D. H., ANDERSON, W. F., FISHER, R. G., TAKEDA, Y. and MATTHEWS, B. W. (1982) *Nature, Lond.* **298,** 718–723.
OVCHINNIKOV, A. A. and UKRAINSKII, I. I. (1978) Acad. Sciences Ukrainian SSR, Institute of Theoretical Physics, Preprint ITP-78-150E.
OVCHINNIKOV, A. A. and UKRAINSKII, I. I. (1979) *Doklady Acad. Sci. USSR* **244,** 751–754.
PABO, C. O. and LEWIS, M. (1982) *Nature, Lond.* **298,** 443–447.
PAGANI, S. and GALANTE, Y. M. (1983) *Biochim. biophys. Acta* **742,** 278–284.
PAI, E. F. and SCHULZ, G. E. (1983) *J. biol. Chem.* **258,** 1752–1757.
PAI, E. F., SACHSENHEIMER, W., SCHIRMER, R. H. and SCHULZ, G. E. (1977) *J. molec. Biol.* **114,** 37–45.
PAULING, L. (1960) *The Nature of the Chemical Bond,* p. 281. Cornell University Press, Ithaca, New York.
PAULING, L. (1979) *Proc. natn. Acad. Sci. U.S.A.* **76,** 2293–2294.
PAULING, L. and COREY, R. B. (1953) *Proc. natn. Acad. Sci. U.S.A.* **39,** 253–256.
PAULING, L., COREY, R. B. and BRANSON, H. R. (1951) *Proc. natn. Acad. Sci. U.S.A.* **37,** 205–211.
PERUTZ, M. F. (1978) *Science* **201,** 1187–1191.
PETHIG, R. (1979) *Dielectric and Electronic Properties of Biological Materials,* p. 65. John Wiley, New York.
PHILLIPS, D. C., RIVERS, P. S., STERNBERG, M. J. E., THORNTON, J. M. and WILSON, I. A. (1977) *Biochem. Soc. Trans.* **5,** 642–647.
PHILLIPS, S. E. V. (1980) *J. molec. Biol.* **142,** 531–554.
PLOEGMAN, J. H., DRENT, G., KALK, K. H., HOL, W. G. J., HEINRIKSON, R. L., KEIM, P., WENG, L. and RUSSELL, J. (1978) *Nature, Lond.* **273,** 124–129.
PLOEGMAN, J. H., DRENT, G., KALK, K. H. and HOL, W. G. J. (1979) *J. molec. Biol.* **127,** 149–162.
POLAND, D. and SCHERAGA, H. A. (1967) *Biochemistry* **6,** 3791–3800.
PRASAD, B. V. V. and SASISEKHARAN, V. (1979) *J. Am. chem. Soc.* **12,** 1107–1110.
PRIVALOV, P. L. (1979) *Adv. Prot. Chem.* **33,** 167–241.
RALSTON, E. and DE COEN, J.-L. (1974) *J. molec. Biol.* **83,** 393–420.

Ramachandran, G. N., Venkatalacham, C. M. and Krimm, S. (1966) *Biophys. J.* **6,** 849–872.
Rees, D. C. (1980) *J. molec. Biol.* **141,** 323–326.
Remington, S., Wiegand, G. and Huber, R. (1982) *J. molec. Biol.* **158,** 111–152.
Richards, F. M. (1977) *Ann. Rev. Biophys. Bioeng.* **6,** 151–176.
Richards, F. M. and Richmond, T. (1978) In *Molecular Interactions and Activity in Proteins* (eds. R. Porter and D. W. Fitzsimons), p. 25, Ciba Foundation Symposium 60, Exerpta Medica, Amsterdam.
Richardson, J. S. (1977) *Nature* **268,** 495–500.
Richardson, J. S. (1981) *Adv. Prot. Chem.* **34,** 167–339.
Robertus, J. D., Kraut, J., Alden, R. A. and Birktoft, J. J. (1972) *Biochemistry* **11,** 4293–4303.
Robson, B. and Osguthorpe, D. J. (1979) *J. molec. Biol.* **132,** 19–51.
Rossky, P. J. and Karplus, M. (1979) *J. Am. chem. Soc.* **8,** 1913–1937.
Rossmann, M. G., Liljas, A., Brändén, C.-I. and Banaszak, L. J. (1975) *The Enzymes* **11,** 61–102.
Salemme, F. R. (1982) *Proc. natn. Acad. Sci. U.S.A.* **79,** 5263–5267.
Salemme, F. R. (1981) *J. molec. Biol.* **146,** 143–156.
Sauer, K. *Acc. Chem. Res.* **11,** 257–264.
Sauer, R. T., Yocum, R. R., Doolittle, R. F., Lewis, M. and Pabo, C. O. (1982) *Nature, Lond.* **298,** 447–451.
Schlesinger, P. and Westley, J. (1974) *J. biol. Chem.* **10,** 780–788.
Schulz, G. E., Schirmer, R. H., Sachsenheimer, W. and Pai, E. F. (1978) *Nature, Lond.* **273,** 120–124.
Schulz, G. E. and Schirmer, R. H. (1979) *Principles of Protein Structure*, Springer-Verlag, New York.
Schulz, G. E., Schirmer, R. H. and Pai, E. F. (1982) *J. molec. Biol.* **160,** 287–308.
Schwarz, G. and Sarko, P. (1982) *Biophys. J.* **39,** 211–219.
Sheridan, R. P. and Allen, L. C. (1980) *Biophys. Chem.* **11,** 133–136.
Sheridan, R. P., Levy, R. M. and Salemme, F. R. (1982) *Proc. natn. Acad. Sci. U.S.A.* **79,** 4545–4549.
Shire, S. J., Hanania, G. I. H. and Gurd, F. R. N. (1974) *Biochemistry* **13,** 2967–2974.
Shoham, M. and Steitz, T. A. (1980) *J. molec. Biol.* **140,** 1–14.
Silverman, D. N. and Scheraga, H. A. (1972) *Arch. biochem. Biophys.* **153,** 449–456.
Smith, W. W., Burnett, R. M., Darling, G. D. and Ludwig, M. L. (1977) *J. molec. Biol.* **117,** 195–225.
Smith, W. W., Pattridge, K. A., Ludwig, M. L., Petsko, G. A., Tsernoglou, D., Tanaka, M. and Yasunobu, T. (1983) *J. molec. Biol.* **165,** 737–755.
Söderberg, B.-O., Sjöberg, B.-M., Sonnerstram, U. and Brändén, C.-I. (1978) *Proc. natn. Acad. Sci. U.S.A.* **75,** 5827–5830.
Steitz, T. A., Ohlendorf, D. H., McKay, D. B., Anderson, W. F. and Matthews, B. W. (1982) *Proc. natn. Acad. Sci. U.S.A.* **79,** 3079–3100.
Stephens, P. E., Lewis, H. M., Darlison, M. G. and Guest, J. R. (1983) *Eur. J. Biochem.* **135,** 519–527.
Stryer, L. (1981) *Biochemistry*, 2nd Edn., Freeman, San Francisco.
Stuart, D. I., Levine, M., Muirhead, H. and Stammers, D. K. (1979) *J. molec. Biol.* **134,** 109–142.
Stubbs, G., Warren, S. and Holmes, K. (1977) *Nature, Lond.* **267,** 216–221.
Stura, E. A., Zanotti, G., Babu, Y. S., Sansom, M. S. P., Stuart, D. I., Wilson, K. S. and Johnson, J. N. (1983) *J. molec. Biol.* **170,** 529–565.
Tanford, C. (1970) *Adv. Prot. Chem.* **24,** 1–95.
Tanford, C. and Kirkwood, G. (1957) *J. Am. chem. Soc.* **79,** 5333–5339.
Terwilliger, T. C. and Eisenberg, D. (1982) *J. biol. Chem.* **257,** 6016–6022.
Thole, B. T. and Van Duijnen, P. Th. (1983) *Biophys. Chem.* **18,** 53–59.
Van Duijnen, P. Th. and Thole, B. T. (1982) *Biopolymers* **21,** 1749–1754.
Van Duijnen, P. Th. and Thole, B. T. (1981) *Chem. Phys. Lett.* **83,** 129–133.
Van Duijnen, P. Th., Thole, B. T., Broer, R. and Nieuwpoort, W. C. (1980) *Int. J. Quant. Chem.* **17,** 651–671.
Van Duijnen, P. Th., Thole, B. T. and Hol, W. G. J. (1979) *Biophys. Chem.* **9,** 273–280.
Van Gunsteren, W. F., Berendsen, H. J. C., Hermans, J., Hol, W. G. J. and Postma, J. P. M. (1983) *Proc. natn. Acad. Sci. U.S.A.* **80,** 4315–4319.
Volz, K. W., Matthews, D. A., Alden, R. A., Freer, S. T., Hansch, C., Kaufman, B. T. and Kraut, J. (1972) *J. biol. Chem.* **257,** 2528–2536.
Wada, A. (1976) *Adv. Biophys.* **9,** 1–63.
Wada, A. and Nakamura, H. (1981) *Nature* **293,** 757–758.
Warshel, A. and Schlosser, D. W. (1981) *Proc. natn. Acad. Sci. U.S.A.* **78,** 5564–5568.
Warshel, A. and Levitt, M. (1976) *J. molec. Bol.* **103,** 227–249.
Warshel, A. (1979) *J. phys. Chem.* **83,** 1640–1652.
Warwicker, J. and Watson, H. C. (1982) *J. molec. Biol.* **157,** 671–679.
Watenpaugh, K. D., Sieker, L. C. and Jensen, L. H. (1973) *Proc. natn. Acad. Sci. U.S.A.* **70,** 3857–3860.
Watson, H. C., Walker, N. P. C., Shaw, P. J., Bryant, T. N., Wendell, P. L., Fothergill, L. A., Perkins, R. E., Conroy, S. C., Dobson, M. J., Tuite, M. F., Kingsman, A. J. and Kingsman, S. M. (1982) *EMBO J.* **1,** 1635–1640.
Westley, J. (1973) *Adv. Enzymol.* **39,** 327–368.
Weijer, W. J., Hoefsteenge, J., Beintema, J. J., Wierenga, R. K. and Drenth, J. (1983) *Eur. J. Biochem.* **133,** 109–118.
Weill, G. and André, J. J. (1978) *Biopolymers* **17,** 811–814.
Whalley, E. (1978) *Chem. Phys. Lett.* **53,** 449–451.
Wierenga, R. K., De Jong, R. J., Kalk, K. H., Hol, W. G. J. and Drenth, J. (1979) *J. molec. Biol.* **131,** 55–73.
Wierenga, R. K. and Hol, W. G. J. (1983) *Nature, Lond.* **302,** 842–844.
Wierenga, R. K., Drenth, J. and Schulz, G. E. (1983) *J. molec. Biol.* **167,** 725–739.
Wierenga, R. K., De Maeyer, M. C. H. and Hol, W. G. J. (1984) *Biochemistry*, in press.
Wodak, S. Y., Liu, M. Y. and Wyckoff, H. W. (1977) *J. molec. Biol.* **116,** 855–875.
Worcester, D. L. (1978) *Proc. natn. Acad. Sci. U.S.A.* **75,** 5475–5477.

WRIGHT, C. S., ALDEN, R. A. and KRAUT, J. (1969) *Nature* **221,** 235–242.
YAMABE, S. and MOROKUMA, K. (1975) *J. Am. chem. Soc.* **97,** 4458–4465.
YANTORNO, R., TAKASHIMA, S. and MUELLER, P. (1982) *Biophys. J.* **38,** 105–110.