USE OF HELICAL WHEELS TO REPRESENT THE STRUCTURES OF PROTEINS AND TO IDENTIFY SEGMENTS WITH HELICAL POTENTIAL

MARIANNE SCHIFFER *and* ALLEN B. EDMUNDSON

From the Division of Biological and Medical Research, Argonne National Laboratory, Argonne, Illinois

ABSTRACT The three-dimensional structures of α-helices can be represented by two-dimensional projections which we call helical wheels. Initially, the wheels were employed as graphical restatements of the known structures determined by Kendrew, Perutz, Watson, and their colleagues at the University of Cambridge and by Phillips and his coworkers at The Royal Institution. The characteristics of the helices, discussed by Perutz et al. (1965), and Blake et al. (1965), can be readily visualized by examination of these wheels. For example, the projections for most helical segments of myoglobin, hemoglobin, and lysozyme have distinctive hydrophobic arcs. Moreover, the hydrophobic residues tend to be clustered in the $n \pm 3, n, n \pm 4$ positions of adjacent helical turns. Such hydrophobic arcs are not observed when the sequences of nonhelical segments are plotted on the wheels. Since the features of these projections are also distinctive, however, the wheels can be used to divide sequences into segments with either helical or nonhelical *potential*. The sequences of insulin, cytochrome *c*, ribonuclease A, chymotrypsinogen A, tobacco mosaic virus protein, and human growth hormone were chosen for application of the wheels for this purpose.

HELICAL WHEELS AND HYDROPHOBIC ARCS

While attempting to correlate the amino acid sequences of sperm whale myoglobin (Edmundson, 1965) and egg white lysozyme (Canfield, 1963; Jollès et al., 1963, 1964) with their three-dimensional structures (Kendrew et al., 1960, 1961; Kendrew, 1962, 1963; Blake et al., 1965), we found that it was undesirable to represent the helical segments by the usual linear array of abbreviations. Possible side-chain interactions and the general characteristics of helices can be more readily visualized if the sequence is plotted on two-dimensional figures which we call "helical wheels." The wheels are projections of the amino acid side-chains onto a plane perpendicular to the axis of the helix. The perimeter of each wheel corresponds to the backbone of the polypeptide chain and the external spokes to the side-chains. For an α-helix with 3.6 residues per turn, adjacent side-chains in the sequence are separated by 100° of arc on the wheel.[1]

[1] As J. L. Oncley pointed out, Haggis (1964) employed a similar type of projection to represent a segment of insulin. His projection differed from our wheels in the choice of a period of 11 rather than the usual 18 residues in an α-helix.

The eight helical wheels for myoglobin (Kendrew et al., 1961) are presented in Fig. 1. For comparison, the nonhelical segments containing more than two residues were treated *as if* they were helical, and the resulting wheels are shown in Fig. 2. The N-terminal two residues, Val·Leu, and the Ala residue located in the AB corner are omitted. The residues in the wheels are numbered by the crystallographic system of Watson and Kendrew (1961; also see Edmundson, 1965; Perutz et al., 1965). Helices are designated by a single letter, while the nonhelical connecting segments are denoted by two letters corresponding to those of the helices joined. The inclusive numbers of the residues in the linear sequence (Edmundson, 1965) are listed above the wheels.

The comparison of the helical with nonhelical wheels provides an illustration of what Perutz, Kendrew, and Watson (1965) emphasize in their analyses of the myoglobins and hemoglobins, namely, that the sequences assigned to the two types of structures are substantially different. In the long helices, A, B, E, G, and H, consisting of four to seven turns, the percentage of hydrophobic side-chains, i.e. those of Leu, Ilu, Val, Met, Phe, Tyr, Try, and Ala side-chains (also Cys in proteins with disulfide bonds), is significantly higher than in the nonhelical regions or in intact myoglobin. Conversely, the nonhelical regions have relatively more polar side-chains, both charged and uncharged, i.e. Lys, Arg, His, Asp, Glu, Asn, Gln, Ser, and Thr. The actual percentages are given in Table I.

Unlike myoglobin, the polypeptide chain of lysozyme is cross-linked by disulfide bonds, and the helical content is also lower (Blake et al., 1965). The percentages of hydrophobic residues in the helical and nonhelical segments, as determined by Blake et al. (1965), are listed in Table I. These results indicate that the differences in the composition of the two types of segments in lysozyme are even more pronounced than those for myoglobin. No attempt has as yet been made to use the wheels to compare the α-helices with the 3_{10} helices found in lysozyme (Phillips, 1966). The external spokes on wheels for the latter are spaced 120° apart.

In each α-helix, except C and D in myoglobin, there is a cluster of at least three hydrophobic residues on the same side of the wheel. We call these clusters stabilization arcs, since the presence of hydrophobic residues is believed to stabilize a helix (Scheraga, 1961; Némethy and Scheraga, 1962). Kendrew et al. (1961, 1962) found that hydrophobic residues tend to be directed away from the solvent and toward the "inside" of the molecule, while almost all polar residues are located on the "surface." One of the few exceptions is the heme-linked histidine residue F8 (Kendrew et al., 1961), which is the only polar residue along what is otherwise a completely hydrophobic arc.

When the wheels are oriented in positions required by the three-dimensional models of Watson and Kendrew (personal communication), it is evident that the hydrophobic arcs collectively form a stabilizing internal "core," in which the possibilities for *inter*helical interactions are maximized (see Perutz et al., 1965). Within each helical segment, the side-chains of residues numbered $n \pm 3$, n, $n \pm 4$ are in the

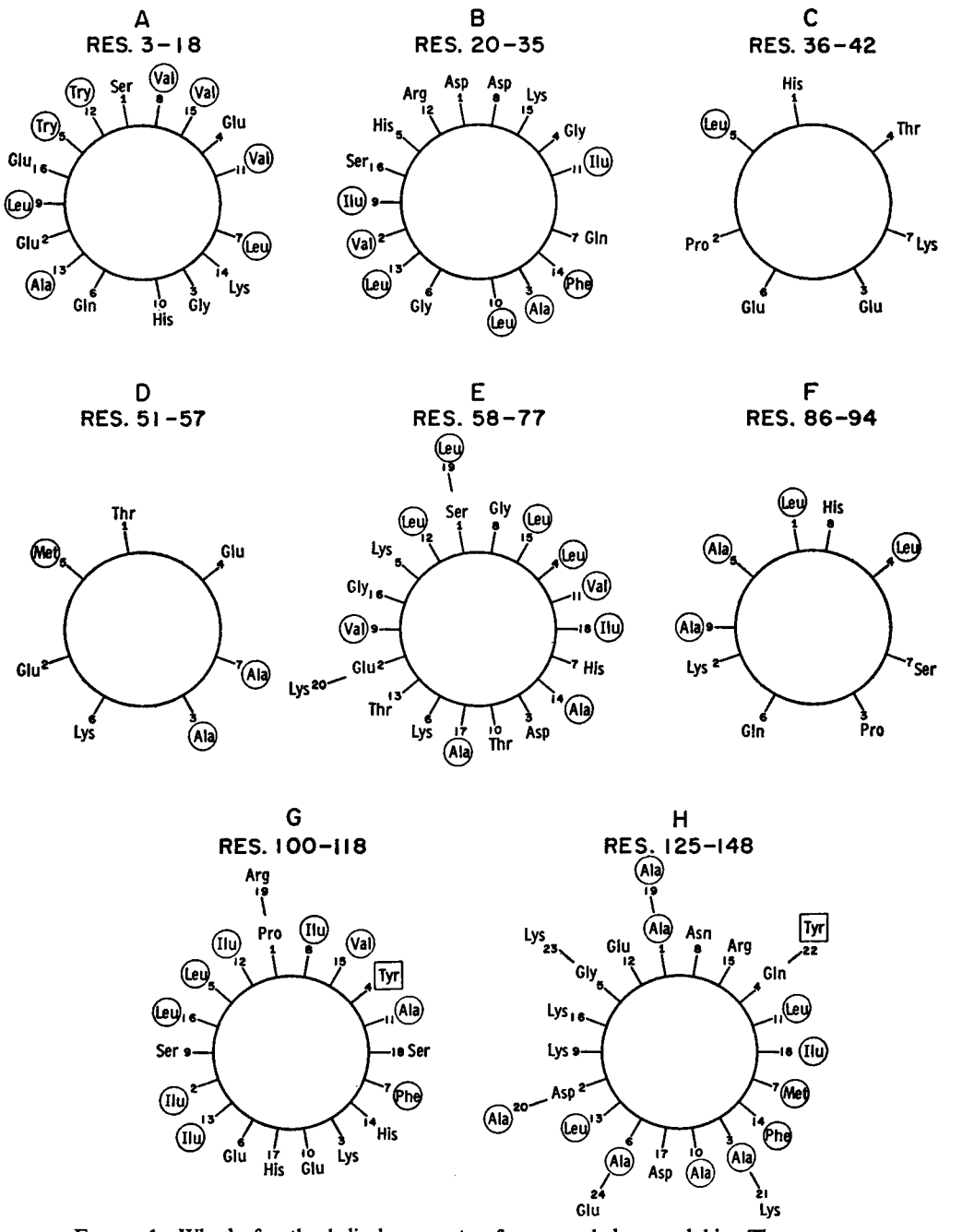

FIGURE 1 Wheels for the helical segments of sperm whale myoglobin. The sequences (Edmundson, 1965) found in helices A through H (Watson and Kendrew, 1961) are represented by wheels which are projections of the amino acid side-chains onto planes perpendicular to the long axes of the helices. The perimeter of each wheel corresponds to the backbone of the polypeptide chain, and the external spokes, to the side-chains. The latter are denoted by three-letter abbreviations, and are numbered by the crystallographic system of Watson and Kendrew (1961; Edmundson, 1965). Since an α-helix contains 3.6 residues per turn, side-chains adjacent in the linear sequence are separated by 100° of arc on the wheel. Residues with hydrophobic side-chains are circled to assist in visualization of the hydrophobic arcs in helics or their absence in nonhelical segments (see text). The abbreviation for tyrosine is enclosed in a rectangle because in exceptional cases its side-chain may not be predominantly hydrophobic in character.

most favorable positions to interact (Némethy and Scheraga, 1962). Némethy and Scheraga (1962) have pointed out that *intra*helical interactions can involve only two residues. Without specific knowledge of the crystal structure, however, it is desirable to consider all three residues in a cluster as possible participants in such interactions. In intact proteins like myoglobin the *inter*helical hydrophobic interactions appear

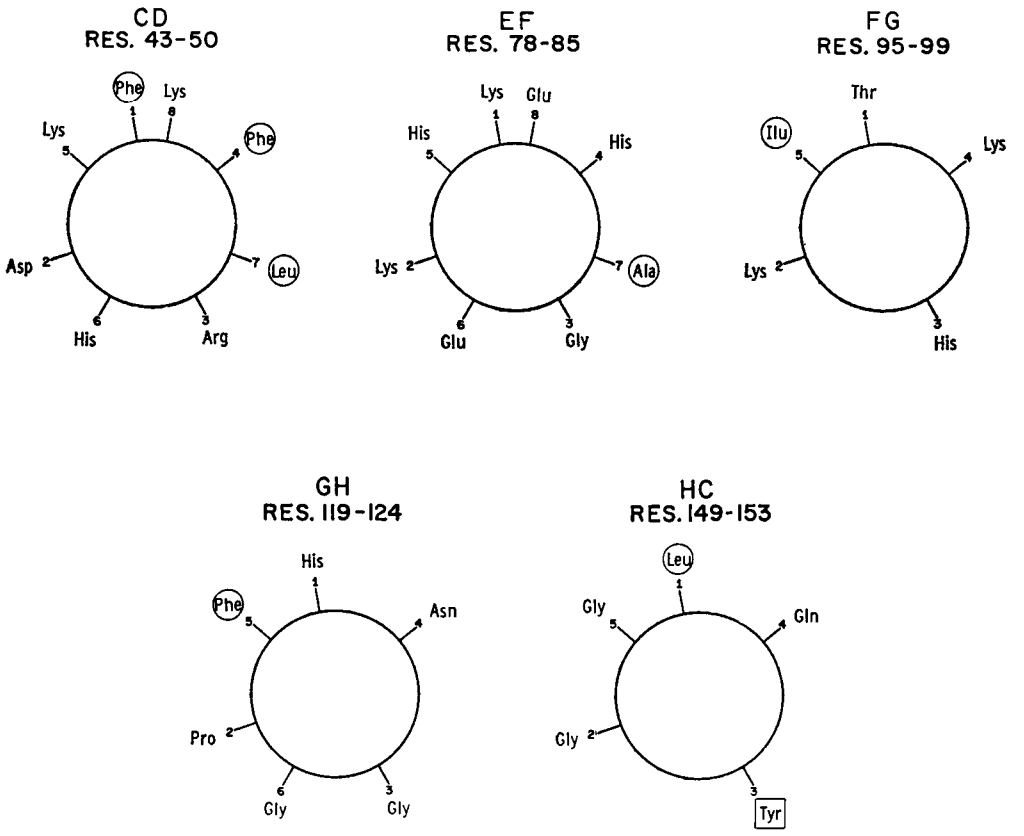

FIGURE 2 Wheels for the nonhelical segments of sperm whale myoglobin. The sequences of nonhelical segments containing more than two residues were treated *as if* they were helical and were plotted on the wheels by the procedure outlined in the caption for Fig. 1. These connecting segments are designated by two letters corresponding to those of the helices joined. The C-terminal segment following helix H is denoted by the abbreviation HC. The characteristics of these wheels are discussed in the text.

to be the major factors required for stabilization, but in an isolated helix the *intra*helical interactions must also be important for thermodynamic stability (see Guzzo, 1965; Némethy and Scheraga, 1962). In the prediction of helical and nonhelical structures with the wheels, therefore, we try to identify segments likely to have thermodynamic stability by themselves and at the same time to have capability for *inter*helical interactions in the protein. This duality is served when the hydrophobic

residues are located in the n, $n \pm 3$, $n \pm 4$ positions, and we consequently consider such segments to have the greatest potential for helicity. A case in point is a segment from ribonuclease. We predicted that residues 43–58 might be in helical configuration in the protein, and Prothero (1966) came to a similar conclusion by a statistical study. Shearer et al. (1966) isolated a peptide from this region of ribonuclease and demonstrated by optical rotatory dispersion methods that a small amount of helical structure was present in solution.

Segments with distinct polar and hydrophobic arcs, but without the n, $n \pm 3$, $n \pm 4$ distribution, may be helical in the protein and participate principally in *inter*-helical interactions. They are less likely to form a stable helix in solution when isolated. In nonhelical regions, the opportunities for both types of hydrophobic inter-

TABLE I

DISTRIBUTION OF POLAR AND HYDROPHOBIC RESIDUES IN HELICAL AND NONHELICAL SEGMENTS OF MYOGLOBIN AND LYSOZYME

The results for myoglobin were compiled from the data of Kendrew and his colleagues (1961, 1962, 1963; see Edmundson, (1965), and those for lysozyme were obtained from Fig. 6 of the article written by Blake et al. (1965).

Segments	No. of residues	Polar residues		Hydrophobic residues (includingAla)		Alanine residues	
		No.	% of total	No.	% of total	No.	% of total
Myoglobin							
Helical	118	55	47	54	46	15	13
Nonhelical	35	18	51	11	31	2	6
Total	153	73	48	65	42	17	11
Lysozyme							
Helical	56	22	39	32	57	10	18
Nonhelical	73	39	53	22	30	2	3
Total	129	61	47	54	42	12	9

actions are generally less numerous, as evidenced by the absence of well-defined hydrophobic arcs when these sequences are plotted on the wheels (see Fig. 2).

Application of these criteria should result in the selection of many segments with helical potential, but there are also other factors leading to the formation of helices. For example, the tendency of polymers of alanine to form α-helices in aqueous solutions (Gratzer and Doty, 1963) is strongly reflected in the helical segments of myoglobin and lysozyme. The percentages assigned to helical and to nonhelical segments are given in Table I. The results for myoglobin indicate that alanine represents 13% of the residues in the helical regions and only 6% of those in nonhelical segments. For lysozyme the corresponding figures are 18 and 3%. Within the helices, particularly the H helix of myoglobin, the alanine residues frequently are constituents of the hydrophobic arcs. We have included alanine in these arcs for other proteins, while realizing that the methyl side chain is not very hydrophobic.

Moreover, the presence of other specific residues like proline, β-forked residues, serine, threonine, and to a lesser extent aspartic and glutamic acid and their amides, also affects the type of structure possible in a given region. These effects have been previously discussed in detail (Kendrew, 1962; Perutz et al., 1965; Watson and Kendrew, 1961; Blout et al., 1960; Harrison and Blout, 1965; Davies, 1964; Guzzo, 1965; Havsteen, 1966; Prothero, 1966).

EXAMINATION OF SEQUENCES OF INSULIN, CYTOCHROME c, RIBONUCLEASE A, TMV PROTEIN, HUMAN GROWTH HORMONE, AND CHYMOTRYPSINOGEN A

The amino acid sequences of proteins are genetically controlled, while the three-dimensional structures appear to be determined exclusively by the sequences (Anfinsen, 1962; Harrison and Blout, 1965). By extrapolation from known structures like myoglobin, hemoglobin, and lysozyme, therefore, it may be possible to deduce some of the three-dimensional features of proteins for which only the sequences are known. With the use of helical wheels and the general concepts elucidated by Kendrew, Perutz, Watson, Phillips, and their colleagues, we examined the sequences of insulin (Sanger, 1960; Sanger and Tuppy, 1951; Sanger and Thompson, 1953; Ryle et al., 1955; Smith, 1962), cytochrome *c* (Margoliash, 1963; Margoliash and Smith, 1965; Chan and Margoliash, 1966), ribonuclease A (Hirs et al., 1960; Spackman et al., 1960; Smyth et al., 1963; Anfinsen, 1962), chymotrypsinogen A (Hartley et al., 1964, 1965), tobacco mosaic virus protein (Tsugita et al., 1960; Anderer et al., 1960; Funatsu et al., 1964), and human growth hormone (Li et al., 1966). We attempted to distinguish between segments with characteristics similar to those attributed to the α-helices of myoglobin and lysozyme, and those with the properties of the nonhelical segments. The following criteria are considered in the selection process:

1. Initially, the sequences are scanned to find all hydrophobic residues in triple n, $n \pm 3$, $n \pm 4$ type distributions. Each set is plotted on a mimeographed copy of a standard wheel which has 24 numbered spokes, each set 100° from its predecessor. After residue n is first placed on one of the spokes in the midrange of 8–16, the plot is extended to include the series of amino acids on both the amino and the carboxyl sides of residue n. The process is continued until the n, $n \pm 3$, $n \pm 4$ type of hydrophobic arc is interrupted by a polar residue, or until a helix disruptor like proline appears. However, proline residues may be located on the first turn of a helix, as in the C, F, and G helices of myoglobin (Kendrew et al., 1961). At the completion of this procedure, the selected sequence is replotted, with the first residue shifted to spoke 1. This group represents our prediction of the type of segment most likely to be helical.

2. The remainder of the sequence is next placed on the wheels in groups of eight residues, corresponding to approximately two turns of α-helix. If polar groups are present around the entire circumference, or if the hydrophobic residues are not clustered to form a distinct arc, the segment is assigned to the nonhelical type.

3. The most difficult assignments are associated with segments in which three of eight residues are hydrophobic and clustered, but are not distributed in n, $n \pm 3$, $n \pm 4$ positions. An example of a helical segment with these properties is the region containing residues 108–115 in lysozyme (with Try, Try, Cys in the 1, 4, 8 positions; see Blake et al., 1965). Segment CD of myoglobin is compact but nonhelical, with Phe, Phe, Leu in a 1, 4, 7 distribution (see Fig. 2). The latter arc is interrupted by a lysine residue, and in contrast to the segment in lysozyme, the remaining residues on the wheel are all polar. The third type of segment, exemplified by residues 55–62 in lysozyme, has a 1, 4, 8 distribution of Ilu, Ilu, Try, but is neither compact nor helical. Because of these inconsistencies, we have not attempted to distinguish among the three.

4. It was assumed that helical structures would not be seriously affected by steric interference between β-forked residues.

5. Since serine, threonine, and to a lesser extent, aspartic and glutamic acid and their amides may exert a terminating effect on helical segments (Kendrew, 1962; Davies, 1964; Perutz et al., 1965; Guzzo, 1966), we try whenever possible to place serine and threonine residues, or 1, 4 combinations of these with the other residues mentioned, on the first or last turns. In addition to the terminating effects, the observations may also indicate a stabilization of the first and last turns by the formation of additional hydrogen bonds involving the β-hydroxyl group, at least in the crystals. In the cases of the C and D helices of myoglobin, it is difficult to cite any other features with which the existence of the helical structures can be explained.

6. In the myoglobins and hemoglobins, tyrosine is included in the hydrophobic category (Perutz et al., 1965), but it may not always be an important member of hydrophobic clusters in other proteins because of its polar phenolic group.

7. If the regions with the characteristics cited above are in fact helical, the hydrophobic arcs should not be interrupted by the imposition of polar residues in the corresponding protein from a different species. Substitution of one hydrophobic residue by another, or of one polar residue by another, is not expected to produce any major changes in the helical structure.

Insulin

One of the pitfalls of the present approach is illustrated by the wheels for the A chain of insulin (see Fig. 3). The sequence of the first 11 residues meets the criteria for helical structure, but the disulfide bond between residues 6 and 11 probably causes the intervening segment to mushroom out into a nonhelical configuration. Many of the genetically controlled variations are manifest in this loop (Smith, 1962; Young, 1962). The first six residues by themselves also fulfill the requirements we set for helicity, but would occupy only 1½ turns of such a helix. The wheel for residues 12–20, presented in Fig. 3, has both a polar and a hydrophobic arc, but not an $n \pm 3$, n, $n \pm 4$ type distribution of hydrophobic residues, i.e., unless tyrosine in position 8 acts as a hydrophobic residue. The possibility of a 1, 4 hydrogen bond

between serine and glutamine also exists. Although not completely satisfying the most rigorous conditions, this segment may be helical. In the B chain, the middle one-third (residues 9–19; see Fig. 3) has a hydrophobic arc, with two overlapping

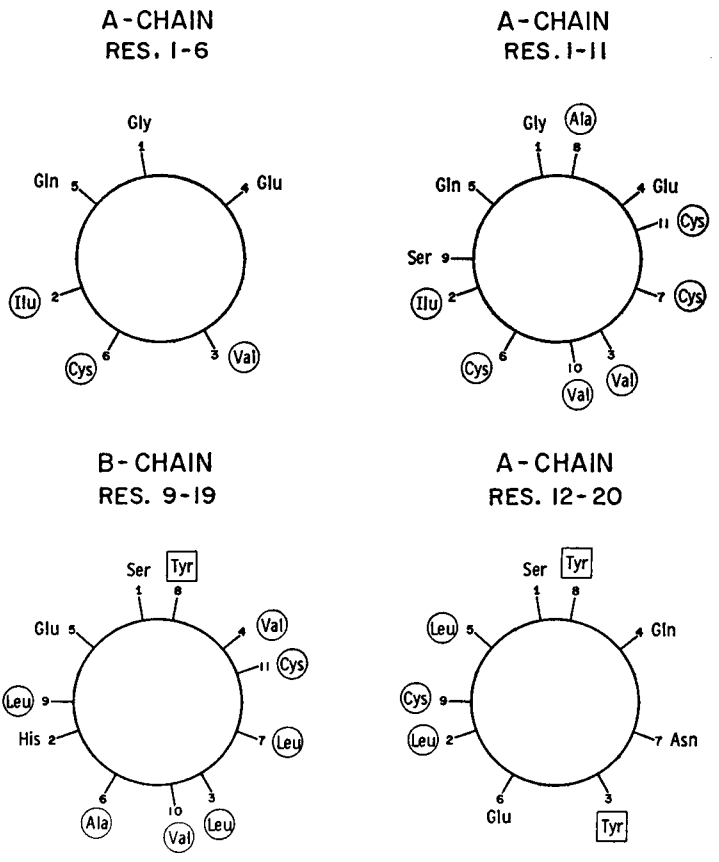

FIGURE 3 Wheels for bovine insulin. The wheels for the first 11 residues of the A chain (Sanger and Thompson, 1953; Ryle et al., 1955) illustrate one of the problems associated with the prediction of possible helical structures. Segments containing either residues 1–6 or 1–11 have hydrophobic arcs and the favorable *n*, 3, 4 type distribution of hydrophobic residues. However, the disulfide bond between residues 6 and 11 probably causes the intervening segment to mushroom out into a nonhelical configuration. The two remaining wheels for segments with helical potential are similar to those for myoglobin and lysozyme and are discussed in the text. Wheels for segments with the characteristics of nonhelical structures are not shown.

triplet groups (3, 7, 4 and 10, 7, 11) in positions for possible stabilizing interactions. If this segment is included with that of residues 12–20 of the A chain, the calculated helical content of insulin is 39%, in approximate agreement with the published values of 38–59% (Urnes and Doty, 1961).

Cytochrome c

In equine cytochrome *c*, five segments containing 45 residues have properties similar to those of known helices, and the remaining segments appear to be nonhelical in character. Only the wheels for the proposed helices are illustrated in Fig. 4. The

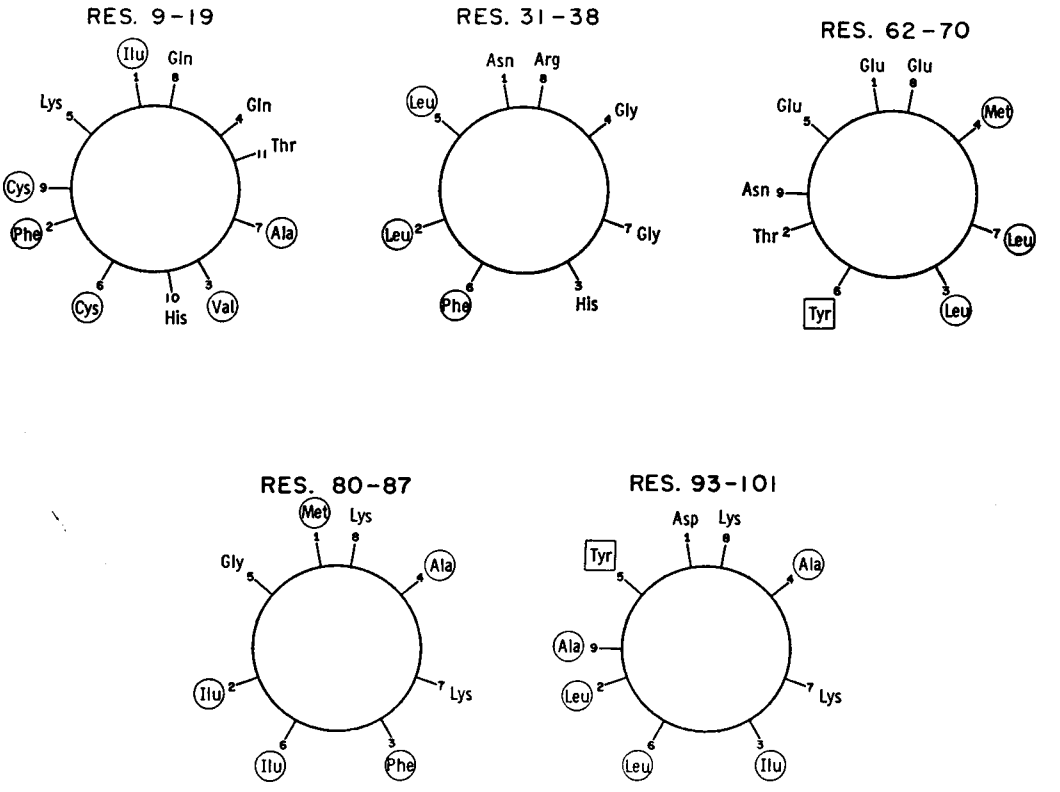

FIGURE 4 Wheels for horse heart cytochrome *c* (Margoliash and Smith, 1965). Only the wheels for segments with helical potential are presented. The sequence plotted on the wheel in the upper left-hand corner contains the heme-linked histidine residue (see position 10), the only polar residue in an otherwise hydrophobic arc (compare with residue 8 in the F helix of myoglobin). The remaining wheels, although representing only two turns of helix in each case, meet all the criteria we set for helicity.

calculated helical content of 43% is somewhat higher than the values (27–34%) obtained by optical rotatory dispersion methods (Urry and Doty, 1965).

As in myoglobin, the histidine residue (No. 18; see position 10 on the wheel), probably linked to the heme group, is the only polar residue in the hydrophobic arc. After building models of the analogous hemopeptide, Ehrenberg and Theorell in 1955 concluded that it is sterically possible for the heme group to be linked to His-18 and the two Cys residues (positions 6 and 9 on the wheel), if they are members of

an α-helix. In the regions with helical potential, we did not find any insertions of polar residues into the hydrophobic arcs of the vertebrate proteins examined by Margoliash and his coworkers (1963, 1964; Chan and Margoliash, 1966). In yeast cytochrome *c* (Narita et al., 1963), however, the hydrophobic arc of the heme-linked segment is weakened, but not completely destroyed by the substitution of a lysine for a valine residue (see position 3 on wheel).

Ribonuclease A

The analysis of ribonuclease indicates less helical potential than that found in either insulin or cytochrome *c*. Only three wheels of eight or more residues have hydrophobic arcs in which the $n \pm 3$, n, $n \pm 4$ distribution occurs (see Fig. 5). The segment, including residues 43–58, can be divided into two halves, with the serine residue in position 8 considered as the first residue of the second half. Most residues known to be essential for the activity of the enzyme, including lysine-*41* (Hirs, 1962), histidine-*119*, and histidine-*12* (Crestfield et al., 1963), are located in segments which are probably not helical. The segments most likely to be helical contain about 33 residues, or 27% of the molecule. The reported values range from 17 to 40% (Yang and Doty, 1957).

To test the restrictions imposed by the combination of the presumed helices and the disulfide bonds, we constructed a model made deliberately flexible by stringing pieces of rubber tubing on a section of plastic electrical tubing (diameter 2 mm). The length was selected to allow approximately 3.5 A per residue in a nonhelical segment, on a scale of 1 cm = 3.5 A. On the same scale, the peptide grouping for each residue was represented by tubing equivalent to 4 A in diameter, but slightly less then 3.5 A in thickness for increased flexibility. Side chains constructed from colored beads were added as needed. Helices were represented by cylinders of white tubing equivalent to 10 A in diameter (to include side chains), and their constituent residues were placed 1.5 A apart. Electrical tubing with rubber spacers was used as disulfide bonds (4 A).

To have the model conform to the axial ratios of 5:3:2 (Haggis, 1964, p. 64), we assumed that the molecule was compact and that it consisted of two layers of polypeptide chain segments. We found that *inter*helical interactions, involving the segments listed in Fig. 5, could be maximized without difficulty. For example, segment 1 (res. 24–31) in one layer could be brought close to segments 2 and 3 (res. 43–58) in the second layer, and segment 4 (res. 104–112) is next to segments 2 and 3 in the second layer. With these restrictions, in combination with those imposed by the disulfide bonds, only the two ends remained free. We aligned these in such a way that the His-119 of the carboxyl end was over the His-12 of the amino end (Crestfield et al., 1963), and moved both ends until Lys-7 was sufficiently close to Lys-41 (Hirs, 1962) to permit bridging by a bi-functional reagent (Marfey et al., 1965). In this model, there is an opening for entry of substrates, with access to the residues believed to be involved in the active site.

Such models, while highly speculative and doubtlessly incorrect in detail, if not in toto, can be constructed rapidly to check the possibility that either helices or other types of structures exist. If the helices can be predicted accurately, the possible variations in the model are considerably decreased, especially in proteins cross-linked by disulfide bonds.

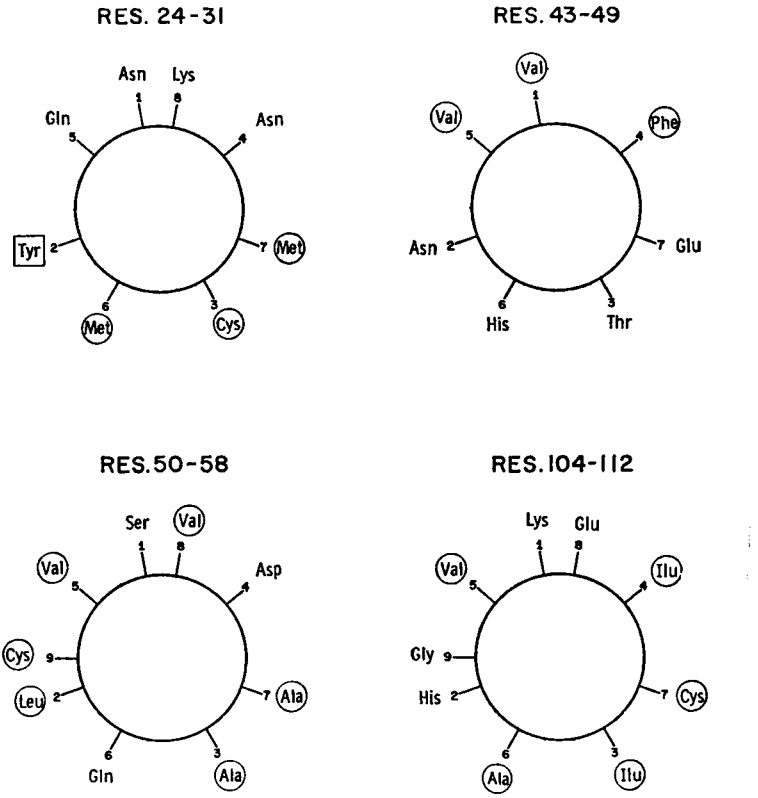

FIGURE 5 Wheels for bovine ribonuclease A (Hirs et al., 1960; Spackman et al., 1960; Smyth et al., 1963; Anfinsen, 1962). The amino acid sequence of ribonuclease includes relatively few segments with properties we have ascribed to α-helices. Only the wheels for these segments are shown, but it should be emphasized that the wheels for the other segments also have distinctive features similar to those for the nonhelical regions of myoglobin and lysozyme. The sequence of residues 43–58 is divided into two segments for reasons given in the text.

TMV Protein, Human Growth Hormone, and Chymotrypsinogen A

The sequences of these three proteins have been examined by the methods used in the preceding cases, and the results are summarized along with the others in Table II. The wheels are not shown, but are analogous to those presented in Figs. 1–5. We shall discuss the results briefly while comparing our procedure with Prothero's statistical method (1966), which appeared in this journal after initial submission of our article.

TABLE II

SEGMENTS WITH HELICAL POTENTIAL; COMPARISON OF OUR RESULTS WITH THOSE OF PROTHERO

The results obtained with the aid of the helical wheels are summarized in the order in which the proteins are discussed in the text. Only segments with helical potential are listed, and they are denoted by the numbers of their first and last constituent amino acid residues. The results obtained by crystallographic methods are given for myoglobin (Kendrew et al., 1961) and lysozyme (Blake et al., 1965). Prothero's published values (1966) are presented below ours. We also applied Prothero's procedure to proteins not considered in his article. The values for these are listed after the term Prothero's method. Our wheels for TMV protein are drawn from a later and slightly different version of the sequence (Funatsu et al., 1964), and Prothero's results are therefore altered for direct comparison. The other published values were also checked, and our figures, which differ in some places from those of Prothero, appear in the table. In the case of the A chain of insulin, the loop between residues 6 and 11 is not included in either type of analysis (see text).

Myoglobin	Kendrew et al.	3–18	20–35	36–42	51–57	58–77	86–94	100–118	124–148
	Our predictions	3–17	27–35	—	—	65–77	86–92	98–116	126–140
									141–146
	Prothero	3–23	—	36–44	47–61	63–78	81–86	103–116	124–139
							88–95		142–148
Lysozyme	Blake et al.	4–15	24–36	80–85	88–95	108–115	119–125		
					99–100				
	Our predictions	1–12	24–34	77–85	92–101	104–111	118–128		
	Prothero	5–13	28–35	80–95		106–111	—		
Insulin									
A chain	Our predictions	1–6							
		12–20							
	Prothero's method	12–18							
B chain	Our predictions	9–19							
	Prothero's method	9–21							

Equine cytochrome *c*	Our predictions	9–19	31–38	62–70	80–87	93–101					
	Prothero's method	15–21	—	58–70	—	88–101					
Ribonuclease A	Our predictions	—	24–31	43–49 50–58	104–112						
	Prothero	2–11	—	43–61	105–113						
TMV protein	Our predictions	10–18	19–33	—	45–53	—	77–89	117–125 126–135			
	Prothero's method	—	20–32	39–46	—	72–77	78–89	104–136			
Human growth hormone	Our predictions	5–20 21–26	29–41	68–74	87–94	95–105	106–126	—	157–165	170–178	
	Prothero's method	12–27	28–51	—	83–93	95–102	107–126	150–156	—	168–183	
Chymotrypsinogen	Our predictions	6–20	50–60 67–74	83–89	113–122	—	—	176–185	186–194	204–213	225–241
	Prothero	9–13 15–23	43–73	82–89	102–123	131–137	160–166	179–185	—	204–215	227–240

In extending Guzzo's method (1965), Prothero (1966) proposed that (*a*) any region of five residues will be helical if at least three include Ala, Val, Leu, or Glu; and (*b*) any segment of seven residues will be helical if at least three are Ala, Val, Leu, or Glu and if one more is Ilu, Thr, or Gln. These criteria are similar to ours, since the Ala, Val, Leu, and Ilu residues are also important members of our helical wheels. The two methods differ, however, in that the statistical procedure does not directly take into account either the relative positions of these residues within the segment or the physical significance resulting from such distributions. Moreover, we assume that other hydrophobic residues, such as Try, Phe, Met, and Cys, are also important in the stabilization of helices. The myoglobins and hemoglobins have no disulfide bonds and the combined number of Try, Phe, and Met residues is substantially lower than that of the more prevalent amino acids cited by Prothero. In these examples the statistical method, unlike ours, is not markedly affected by the presence of the additional hydrophobic residues, but this will not hold for all proteins.

A more pronounced difference in the two procedures is found in the emphasis placed on Glu, Thr, and Gln residues. These are important to us only insofar as they contribute to the polar arcs, interrupt the hydrophobic arcs, or participate in the first and last turns of helices. For proteins in which there are high concentrations of these amino acids, as in TMV protein, the statistical procedure consequently leads to different results from ours (see Table II). A prime reason for inclusion of these residues in the statistical method was their presence in helices C and D of myoglobin (see Fig. 2), and we can only wait to see if these helices represent general or exceptional cases.

In the remaining segments of myoglobin, as well as those of lysozyme and the B chain of insulin (see Table II), similar but not identical results can be obtained with the two methods. The sequences chosen usually are derived from the same region, but differ in length. In ribonuclease, growth hormone, and chymotrypsinogen, there are examples of both close agreement and wide disagreement. Nevertheless, it is probable that we are rapidly approaching the time when the application of thermodynamic principles and model building can be successfully combined with the statistical data derived from known structures.

Note Added in Proof The values given for ribonuclease in Table II and the wheels in Fig. 5 should have included the segment containing residues 3–11.

This work was supported under the auspices of the U. S. Atomic Energy Commission.

Received for publication 14 September 1966.

BIBLIOGRAPHY

ANDERER, F. A., H. UHLIG, E. WEBER, and G. SCHRAMM. 1960. *Nature.* **186:** 922.
ANFINSEN, C. B. 1962. *Brookhaven Symp. Biol.* **15:** 184.

BLAKE, C. C. F., D. F. KOENIG, G. A. MAIR, A. C. T. NORTH, D. C. PHILLIPS, and V. R. SARMA. 1965. *Nature.* **206:** 757.
BLOUT, E. R., C. DE LOZÉ, S. M. BLOOM, and G. D. FASMANN. 1960. *J. Am. Chem. Soc.* **82:** 3787.
CANFIELD, R. E. 1963. *J. Biol. Chem.* **238:** 2698.
CHAN, S. K., and E. MARGOLIASH. 1966. *J. Biol. Chem.* **241:** 335, 507.
CRESTFIELD, A. M., W. H. STEIN, and S. MOORE. 1963. *J. Biol. Chem.* **238:** 2413, 2421.
DAVIES, D. R. 1964. *J. Mol. Biol.* **9:** 605.
EDMUNDSON, A. B. 1965. *Nature.* **205:** 883.
EHRENBERG, A., and H. THEORELL. 1955. *Acta Chem. Scand.* **9:** 1193.
FUNATSU, G., A. TSUGITA, and H. FRAENKEL-CONRAT. 1964. *Arch. Biochem. Biophys.* **105:** 25.
GRATZER, W. B., and P. DOTY. 1963. *J. Am. Chem. Soc.* **85:** 1193.
GUZZO, A. V. 1965. *Biophys. J.* **5:** 809.
HAGGIS, G. H. 1964. Introduction to Molecular Biology. Longmans, Green and Company, London. 73.
HARRISON, S. C., and E. R. BLOUT. 1965. *J. Biol. Chem.* **240:** 299.
HARTLEY, B. S. 1964. *Nature.* **201:** 1285.
HARTLEY, B. S., J. R. BROWN, D. L. KAUFFMAN, and L. B. SMILLIE. 1965. *Nature.* **207:** 1157.
HAVSTEEN, B. H. 1966. *J. Theoret. Biol.* **10:** 1.
HIRS, C. H. W. 1962. *Brookhaven Symp. Biol.* **15:** 154.
HIRS, C. H. W., S. MOORE, and W. H. STEIN. 1960. *J. Biol. Chem.* **235:** 633.
JOLLÈS, J., J. JAUREGUI-ADELL, and P. JOLLÈS. 1963. *Biochim. Biophys. Acta* **78:** 668.
JOLLÈS, P., J. JAUREGUI-ADELL, and J. JOLLÈS. 1964. *Compt. Rend. Acad. Sci.* **258:** 3926.
KENDREW, J. C. 1962. *Brookhaven Symp. Biol.* **15:** 216.
KENDREW, J. C. 1963. *Science.* **139:** 1259.
KENDREW, J. C., R. E. DICKERSON, B. E. STRANDBERG, R. G. HART, D. R. DAVIES, D. C. PHILLIPS, and V. C. SHORE. 1960. *Nature.* **185:** 422.
KENDREW, J. C., H. C. WATSON, B. E. STRANDBERG, R. E. DICKERSON, D. C. PHILLIPS, and V. C. SHORE. 1961. *Nature.* **190:** 666.
LI, C. H., W. K. LIU, and J. S. DIXON. 1966. *J. Am. Chem. Soc.* **88:** 2050.
MARFEY, P. S., M. UZIEL, and J. LITTLE. 1965. *J. Biol. Chem.* **240:** 3270.
MARGOLIASH, E. 1963. *Proc. Natl. Acad. Sci. U. S.* **50:** 672.
MARGOLIASH, E., and E. L. SMITH. 1965. Rutgers University Symposium on Evolving Genes and Proteins. V. Bryson and H. J. Vogel, editors. Academic Press Inc., N. Y. 221.
NARITA, K., K. TITANI, Y. YAOI, and H. MURAKAMI. 1963. *Biochim. Biophys. Acta.* **77:** 688.
NÉMETHY, G., and H. A. SCHERAGA. 1962. *J. Phys. Chem.* **66:** 1773.
PERUTZ, M. F., J. C. KENDREW, and H. C. WATSON. 1965. *J. Mol. Biol.* **13:** 669.
PHILLIPS, D. C. 1966. *Sci. Am.* **215**(5): 78.
PROTHERO, J. W. 1966. *Biophys. J.* **6:** 367.
RYLE, A. P., F. SANGER, L. F. SMITH, and R. KITAI. 1955. *Biochem. J.* **60:** 541.
SANGER, F. 1960. *Brit. Med. Bull.* **16:** 183.
SANGER, F., and E. O. P. THOMPSON. 1953. *Biochem. J.* **53:** 366.
SANGER, F., and H. TUPPY. 1951. *Biochem. J.* **49:** 481.
SCHERAGA, H. A. 1961. *J. Phys. Chem.* **65:** 1071.
SHEARER, W. T., R. K. BROWN, G. F. BRYCE, and F. R. N. GURD. 1966. *J. Biol. Chem.* **241:** 2665.
SMITH, L. F. (reference to Young, F. G. (1962)).
SMYTH, D. G., W. H. STEIN, and S. MOORE. 1963. *J. Biol. Chem.* **238:** 227.
SPACKMAN, D. H., W. H. STEIN, and S. MOORE. 1960. *J. Biol. Chem.* **235:** 648.
TSUGITA, A., D. T. GISH, J. YOUNG, H. FRAENKEL-CONRAT, C. A. KNIGHT, and W. M. STANLEY. 1960. *Proc. Natl. Acad. Sci. U. S.* **46:** 1463.
URNES, P., and P. DOTY. 1961. *Advan. Protein Chem.* **16:** 401.
URRY, D. W., and P. DOTY. 1965. *J. Am. Chem. Soc.* **87:** 2756.
WATSON, H. C., and J. C. KENDREW. 1961. *Nature.* **190:** 670.
YANG, J. T., and P. DOTY. 1957. *J. Am. Chem. Soc.* **79:** 761.
YOUNG, F. G. 1962. *Proc. Roy. Soc.* (*London*), *Ser. B.* **157:** 1.