Strong electrostatic loop-helix interactions in bundle motif protein structures

Kuo-Chen Chou* and Chong Zheng‡
*Computational Chemistry, Upjohn Research Laboratories, Kalamazoo, Michigan 49001; and
‡Department of Chemistry, Northern Illinois University, DeKalb, Illinois 60115

ABSTRACT Based on CHARMM potential (Brooks et al., 1983) an energetic analysis has been carried out for four typical 4-α-helix bundle proteins, i.e., methemerythrin, cytochrome *b*-562, cytochrome *c'*, and bovine somatotropin. The bovine somatotropin possesses long loops, but all the other three proteins have short loops. It was found that in all these four 4-α-helix bundle motif structures the interaction between loops and helices was much stronger than the interaction among the four helices themselves. Particularly for the electrostatic interaction energy, the loop-helix interaction is overwhelmingly stronger than the interhelix interaction although the latter involves the favorable helix dipole interaction due to the antiparallel arrangement of neighboring α-helices. The present study indicates that such a conclusion holds true regardless of what loops, long or short, are in the 4-α-helix bundle protein, and also regardless of which empirical potential, ECEPP or CHARMM, is used for calculations although in CHARMM the electrostatic energy is much more heavily emphasized than in ECEPP. Therefore, no appropriate conclusion can be drawn in arguing whether the dipole interaction among the four α-helices play a stabilizing role or destabilizing role for a 4-α-helix bundle protein without taking into consideration the effect of interaction between helices and loops. The calculated results reported here provide, from a different point of view, insights that might be useful for revealing the essence of the driving forces during the folding of proteins.

INTRODUCTION

As is well known, how to deal with loop structure is a very difficult problem in predicting the tertiary structure of proteins (Carlacci et al., 1991) and also a vitally important problem for the study of protein folding (Thornton et al., 1988; Chou and Carlacci, 1991; Urfer and Kirschner, 1991). Loops are also thought to assume important roles in molecular function and biological recognition (Leszczynski and Rose, 1986). Therefore, an investigation from the energetic point of view into the interaction between loops and the other part of a protein would provide important insights for the study of these areas.

Recent theoretical model studies indicated that the left-handed twisted feature as occurs in many 4-α-helix bundle proteins (Argos et al., 1977; Weber and Salemme, 1980; Richardson, 1980; Banner et al., 1987) could be explained in terms of nonbonded interactions between the constituent helices (Chou et al., 1988; Carlacci and Chou, 1990*a*). Meanwhile, it was also found that the interaction energy between helices and loops in stabilizing the bundle motif protein structures was stronger than that among the constituent helices themselves. Such a finding was further confirmed by a detailed energetic analysis of a 4-α-helix bundle in which each of its four helices was formed by poly$(Ala)_{12}$ and each of its three loops formed by poly$(Ala)_{10}$ (Carlacci and Chou, 1990*b*). It was found for such a theoretical 4-α-helix bundle structure that the electrostatic interaction between helices and loops was $\sim$5 times the size of that among the four helices themselves, and the corresponding nonbonded interaction for the former was almost two times the size of that for the latter (Carlacci and Chou, 1990*b*). Therefore, it seems that the interaction between loops and helices would play a dominant role in stabilizing the 4-α-helix bundle motif structures.

However, the above results were derived from a specific theoretical model, in which (*a*) a bundle was modeled with loops ten residues long, while keeping α-helices at a length of twelve, and (*b*) the poly-Ala sequence was used for all the α-helices. As is well known, real proteins which fold as regular 4-α-helix bundles usually consists of long helices (e.g., 25 residues in the repressor of primer named as the ROP protein (Banner et al., 1987)) and short loops (3–5 residues). Moreover, the use of poly-Ala sequence for the α-helices as adopted in the aforementioned theoretical model will greatly reduce the interaction energy among the four helices. Therefore, the question is naturally raised: does the above conclusion, i.e., that loop-helix interaction plays a dominant role in stabilizing the 4-α-helix bundle motif structure, still hold true if a more general model is adopted in calculation?

Interestingly, the strong interaction between loop and helix segments was found not only for the special theoretical model in which the constituent helices consist of only Ala residues, but also for bovine somatotropin, a real 4-α-helix bundle protein predicted recently by Carlacci et al. (1991) using the heuristic approach. In the case of bovine somatotropin, it was found that the interaction energy between loop and helix segments was -243.5 kcal/mol, whereas the total intersegment interaction energy among the four helices was only -121.1 kcal/mol (Carlacci et al., 1991), suggesting that the stabilization energy even for the real 4-α-helix bundle protein comes mainly from the interaction between the loops and the α-helices.

Nevertheless, the bovine somatotropin molecule is a quite exceptional 4-α-helix bundle, with extremely long (up to 40 residues) loops. For this unusual structure the importance of the loops is not surprising. Furthermore, the bovine somatotropin model based on which the in-

0006-3495/92/09/682/07 $2.00

teraction energies were calculated was not derived from a crystallographic structure, and hence, their reliability might be questioned. Particularly, the results reported in the bovine somatotropin paper by Carlacci et al. (1991) were calculated with ECEPP potential (Momany et al., 1975; Némethy et al., 1983). As is well known, in the ECEPP system only torsional degrees of freedom are allowed to vary, and the bond-stretching and angle-bending terms are omitted from the potential-energy function. Thus, a further question may be raised as asking: what will happen if the calculations are carried out in terms of other empirical potentials, such as those in CHARMM (Brooks et al., 1983) and AMBER (Weiner et al., 1984), where not only the bond-stretching and angle-bending terms are included in the potential-energy function, but the electrostatic energy is also much more heavily emphasized than in ECEPP?

The present study was initiated in an attempt to answer the above questions based on the CHARMM potential (Brooks et al., 1983). To realize this, calculations were carried out not only for the unusual, predicted bovine somatotropin molecule in which loops are much longer than usual, but also for three typical, crystallography-determined 4-α-helix bundle proteins. They are methemerythrin, cytochrome *b*-562 and cytochrome *c'*, in which most loops are short (3–5 residues).

METHODS

The computational procedure consisted of the following three parts, described below: (*a*) generation of the initial structures for each of the proteins investigated; (*b*) energy optimization of each protein molecule; (*c*) energy component analysis; and (*d*) geometric parameter analysis. Computations were carried out by using the residue geometry and the energy parameters of CHARMM (Brooks et al., 1983). The energy was calculated as the sum of electrostatic, nonbonded (including hydrogen-bond), bond, bond angle, and dihedral angle energies. Energy minimizations were based on the conjugated gradient method. The computations were carried out with a Silicon Graphics IRIS 4D/310GTX at Northern Illinois University and the CRAY Y-MP at the National Center for Supercomputing Applications at the University of Illinois (Urbana). The standard conventions for nomenclature of peptide conformations have been followed (IUPAC-IUB Commission on Biochemical Nomenclature, 1970).

Generation of the initial structures

The x-ray crystallographic heavy atom coordinates of methemerythrin determined by Stenkamp et al. (1983) at 2.0 Å resolution, those of cytochrome *b*-562 determined by Lederer et al. (1981) at 2.5 Å resolution, and those of cytochrome *c'* determined by Finzel et al. (1985) at 1.67 Å resolution were taken from Brookhaven Protein Data Bank (Bernstein et al., 1977). All these three proteins assume a typical 4-α-helix bundle motif. Dihedral angles corresponding to hydrogen atoms in the constituent residues were assigned according to their most common conformations as defined in CHARMM (Brook et al., 1983), whose arbitrariness was automatically taken care of by the energy minimization later on.

The atomic coordinates of bovine somatotropin were taken from the structure predicted by Carlacci et al. (1991) using a technique of combining heuristic approach and energy minimization. The bovine somatotropin structure is also a 4-α-helix bundle motif structure, but its three loops are much longer, with 40, 9, and 24 residues, respectively. Such a predicted structure was derived in the ECEPP system.

Energy minimization

The structures obtained above were used as starting points for energy minimizations. For each of these four structures, the energy minimization was carried out for 200 steps with the conjugated gradient method (Brooks et al., 1983). The initial step size used was 0.02 Å.

Energy component analysis

To compare and analyze the various interactions that contribute to overall stability of the 4-α-helix bundle proteins investigated, it is useful to separate the total conformational energy, E_{tot}, into components (Chou and Carlacci, 1991) as defined below:

$$\begin{cases} E^{\alpha}_{intra} = \text{sum of the energies of the four individual constituent } \alpha\text{-helices} \\ E^{\alpha}_{inter} = \text{total intersegment interaction energy among the four } \alpha\text{-helices} \\ E^{\alpha}_{tot} = E^{\alpha}_{intra} + E^{\alpha}_{inter} \\ E^{loop}_{intra} = \text{sum of the energies of the individual loop segments} \\ E^{loop}_{inter} = \text{total intersegment interaction energy among all the loop segments} \\ E^{loop}_{tot} = E^{loop}_{intra} + E^{loop}_{inter} \\ \epsilon = \text{the interaction energy between the loop segments and the } \alpha\text{-helices of the molecule} \\ E_{tot} = E^{\alpha}_{tot} + E^{loop}_{tot} + \epsilon. \end{cases} \quad (1)$$

In the above expressions, the NH_2- and COOH-terminal segments are treated as loops as well, and hence, their contributions to the interaction energy are also included. See Results for a discussion on the role of NH_2- and COOH-terminal segments.

Geometric parameter analysis

In the study of packing arrangement of secondary structures in proteins, the relative orientation of two secondary structures is usually expressed in terms of an orientation angle (Chothia et al., 1977, 1981; Chou et al., 1983*b*, 1984, 1990*b*), denoted by Ω. For two helices *i* and *j*, the orientation angle Ω_0 measures the tilting of the helix axes, with $\Omega_0 = 0°$ for parallel and $\Omega_0 = \pm 180°$ for antiparallel orientations, respectively. The orientation angle is positive if, starting from an initial parallel orientation of the helices ($\Omega_0 = 0°$), the helix far from the viewer is rotated clockwise relative to the one near the viewer; it is negative if the far helix is rotated in the counterclockwise sense. Expressed in terms of unit vectors $\mathbf{e}_i$ and $\mathbf{e}_j$, which each coincides with its corresponding helix axis (Chou et al., 1984) and points from the NH_2- towards the COOH-terminus,

$$\Omega_0 = \begin{cases} \cos^{-1}(\mathbf{e}_i \cdot \mathbf{e}_j), & \text{for clockwise rotation} \\ -\cos^{-1}(\mathbf{e}_i \cdot \mathbf{e}_j), & \text{for counterclockwise rotation.} \end{cases} \quad (2)$$

The sign and magnitude of Ω_0 are independent of which helix is chosen to be the near or far one. According to such a definition, for an antiparallel left-handed twisted 4-α-helix bundle structure, the orientation angle Ω_0 between any two of its adjacent helices should be within the range of $-155 \pm 15°$ (Weber and Salemme, 1980; Chou et al., 1988). If these angles, however, are within the range of $155 \pm 15°$, the corresponding structure would become an antiparallel right-handed twisted bundle.

Because the left twisting is a typical feature in the 4-α-helix bundle proteins (Weber and Salemme, 1980), it is useful to develop a quantitative definition of this, which can be accomplished as follows. First, define the central axis of the helix bundle in an analogous manner to that used by Chou et al. (1990*a*) in their study of the idealized β-barrels according to the following: (*a*) the center of the central axis is set at the

TABLE 1 **Sequence position of the four-α-helix bundle proteins investigated**

Protein	Helix-1	Helix-2	Helix-3	Helix-4	Loop-1	Loop-2	Loop-3	NH_2-terminal segment	COOH-terminal segment
Methemerythrin	19–37	41–64	70–85	91–108	38–40	65–69	86–90	1–18	109–113
Cytochrome *b*-562	3–19	23–42	56–80	84–104	20–22	43–55	81–83	1–2	105–106
Cytochrome *c'*	5–30	40–57	79–102	104–124	31–39	58–78	103–103	1–4	125–128
Bovine somatotropin	7–34	75–96	106–128	153–183	35–74	97–105	129–152	1–6	184–191

mass center of all the C^α atoms of the four α-helices that composed the bundle, and (b) the unit vector of the central axis is given by

$$\mathbf{e}_c = \frac{1}{4}\sum_{i=1}^{4} \tau_i \mathbf{e}_i \tag{3}$$

where $\mathbf{e}_i$ is the unit vector of the ith helix axis, and

$$\tau_i = \begin{cases} 1 & \text{for} \quad \mathbf{e}_1 \cdot \mathbf{e}_i \geq 0 \\ -1 & \text{for} \quad \mathbf{e}_1 \cdot \mathbf{e}_i < 0. \end{cases} \tag{4}$$

The above definition of τ_i ensures that the central axis is directed from the NH_2 to COOH terminus of the 1st helix (Chou et al., 1985). Thus, the twist of a 4-α-helix bundle can be defined as

$$\theta = -\frac{1}{4}\sum_{i=1}^{4} \Omega_i, \tag{5}$$

where Ω_i is the tilt angle of the ith helix with respect to the central axis and is given by

$$\Omega_i = \begin{cases} \Omega_0^i & \text{for} \quad -90° \leq \Omega_0^i \leq 90° \\ \Omega_0^i - 180° & \text{for} \quad 90° < \Omega_0^i \leq 180° \\ \Omega_0^i + 180° & \text{for} \quad -180° \leq \Omega_0^i < -90° \end{cases}, \tag{6}$$

where Ω_0^i is the orientation angle (cf. Eq. 2) between the ith helix axis and the central axis of the bundle. Note the difference between the tilt angle Ω_i defined in Eq. 6 and the orientation angle Ω_0 defined in Eq. 2: the latter ranges from −180 to 180°, but the former ranges from −90 to 90° so that a reversal of the axial direction of the ith helix is denoted by the same numerical value of Ω_i. In other words, when defining the orientation angle the directionality is assigned, but no such a directionality is assigned in defining the tilt angle. Obviously, in describing the twist of a 4-α-helix bundle, the directionalities of axes are not of interest. Therefore, in Eq. 5 we would rather adopt tilt angles than orientation angles. According to the definition as given by Eq. 5, left-handed twisted, nontwisted, and right-handed twisted 4-α-helix bundles are characterized by $\theta < 0$, $\theta = 0$, and $\theta > 0$, respectively. And the larger the value of $|\theta|$, the more the bundle is twisted.

RESULTS AND DISCUSSION

After energy minimization, the rms value of the protein backbone between the x-ray crystallographic structure and the energy-minimized structure for methemerythrin is 0.43 Å, that for cytochrome *b*-562 is 0.39 Å, and that for cytochrome *c'* is 0.44 Å. The small change indicates that the structure determined by x-ray technique is quite

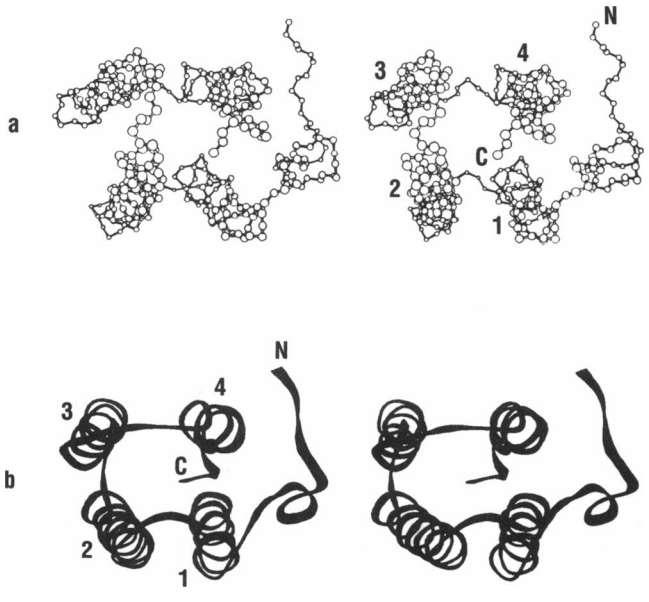

FIGURE 1 The stereo drawing of the energy-minimized methemerythrin: (a) stick-and-ball drawing, and (b) ribbon drawing. This is a antiparallel left-handed twisted four-helix bundle protein, whose four α-helices are marked by 1, 2, 3, and 4 along its sequence. N and C represent NH_2- and COOH-termini, respectively. See Table 1 for the sequence distribution of the constituent segments.

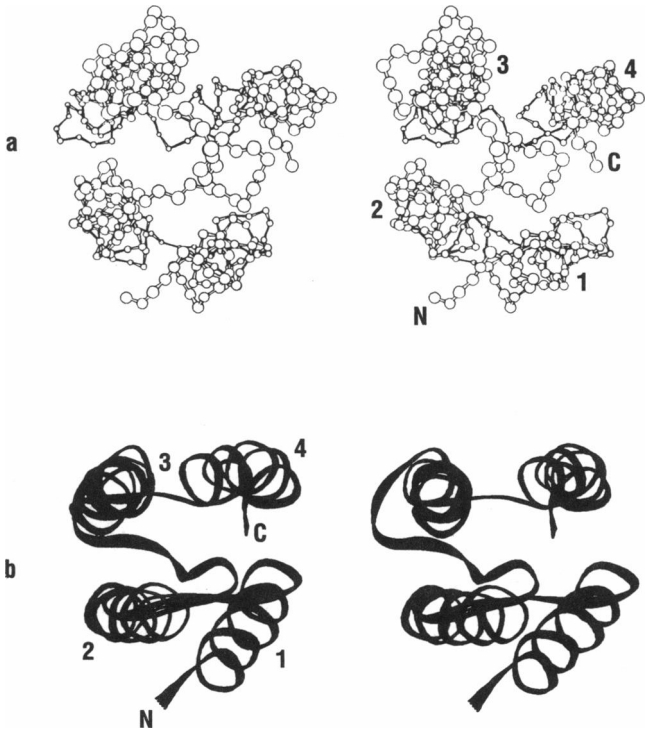

FIGURE 2 The stereo drawing of the energy-minimized cytochrome *b*-562: (a) stick-and-ball drawing, and (b) ribbon drawing. See the legend to Fig. 1 for further explanation.

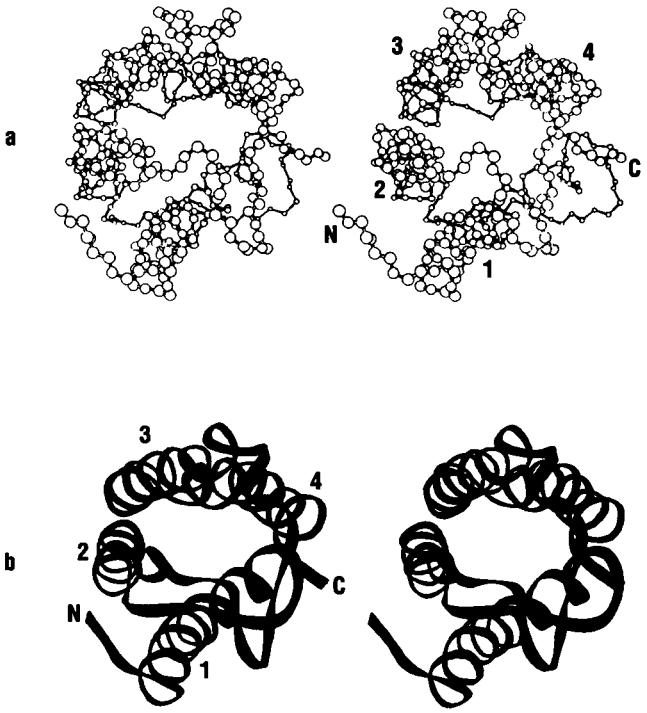

FIGURE 3 The stereo drawing of the energy-minimized cytochrome c': (a) stick-and-ball drawing, and (b) ribbon drawing. See the legend to Fig. 1 for further explanation.

close to that refined after energy minimization for all these three 4-α-helix bundle proteins although it was observed that during the course of energy minimization the energy dropped significantly (by more than 60%). The largest drop is in the electrostatic interaction energy in the loop regions. The energy drop in the helix regions is much smaller, because the helix structure is more rigid.

For bovine somatotropin, the heavy-atom rms value between the structure optimized by CHARMM and that by ECEPP was 1.11 Å.

If $(\phi, \psi) = (-68 \pm 30°, -38 \pm 30°)$ are defined as the dihedral angle region for helices (Chou et al., 1984), it was found after energy optimization that the sequence positions of the four main α-helices for methemerythrin are 19–37, 41–64, 70–85, and 91–108, those for cytochrome b-562 are 3–19, 23–42, 56–102, and 84–104, those for cytochrome c' are 5–30, 40–57, 79–102, and 104–124, and those for bovine somatotropin are 7–34, 75–96, 106–128, and 153–183. A complete information of the sequence position for each of these four 4-α-helix bundle proteins is given in Table 1, in which the sequences for loops 1–3, for the NH_2 attachment and COOH attachment, as well as the sequences of helices 1–4 are presented.

The corresponding stereo drawings for methemerythrin, cytochrome b-562, cytochrome c' and bovine somatotropin are shown in Figs. 1–4, respectively. In these figures, panel a is the backbone stick-and-ball drawing, and panel b the ribbon drawing.

Listed in Table 2 are the geometric parameters characterizing the four 4-α-helix bundle proteins. As shown there, the orientation angles for all the adjacent pairs of helices are within the range of $-155 \pm 15°$, meaning that all four proteins would assume a left-handed twisted bundle motif (Weber and Salemme, 1980; Chou et al., 1988). The left-handed twisted feature of the four bundle proteins is also clearly shown by the stereo drawings in Figs. 1–4. Furthermore, it is found in Table 2 that $\theta = -9°$, $-12°$, $-15°$, and $-20°$, for methemerythrin, cytochrome b-562, cytochrome c', and bovine somatotropin respectively, indicating that as far as the left-handed twisted extent is concerned we have the following order among the four bundle proteins: methemerythrin < cytochrome b-562 < cytochrome c' < bovine somatotropin.

The various energy terms as defined in Eq. 1 are given in Table 3. Those energy terms whose definitions are not explicitly given in Eq. 1 are defined in the footnotes of Table 3. As we can see there that, as far as the intrasegment energy is concerned, the intrahelix energy is much stronger than the intraloop energy in all the four 4-α-

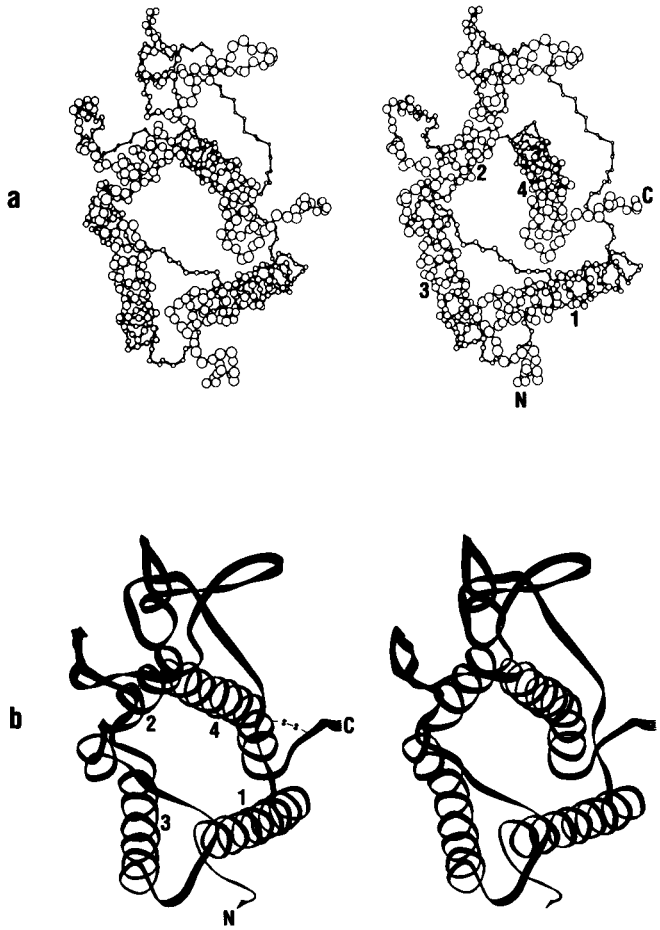

FIGURE 4 The stereo drawing of the energy-minimized bovine somatotropin: (a) stick-and-ball drawing, and (b) ribbon drawing. See the legend to Fig. 1 for further explanation.

TABLE 2 **Geometric parameters characterizing the four-helix bundle proteins investigated**

Relationship between helices[a]				Relationship between helix and the central axis[b]		
Adjacent pair	Diagonal pair	Ω_0 (°)[c]	D (Å)[d]	Helix[a]	Ω_i (°)[e]	R_i (Å)[f]
			Methemerythrin			
1–2		−158	8.9	1	13	6.1
2–3		−169	10.6	2	10	6.6
3–4		−168	11.7	3	10	8.2
4–1		−169	9.4	4	4	7.6
	1–3	23	15.0			
	2–4	14	14.2	θ^g	−9	
			Cytochrome *b*-562			
1–2		−162	8.7	1	18	7.2
2–3		−167	9.1	2	8	5.9
3–4		−170	7.7	3	9	6.8
4–1		−151	9.3	4	13	4.7
	1–3	26	14.0			
	2–4	20	10.6	θ^g	−12	
			Cytochrome *c'*			
1–2		−154	9.3	1	25	7.3
2–3		−167	7.0	2	7	5.8
3–4		−167	8.0	3	14	5.5
4–1		−144	8.8	4	14	5.5
	1–3	38	13.0			
	2–4	19	12.0	θ^g	−15	
			Bovine somatotropin			
1–3		−149	10.4	1	24	9.2
3–2		−159	12.2	2	23	8.6
2–4		−145	10.7	3	15	7.4
4–1		−152	11.2	4	19	6.9
	1–2	47	17.8			
	3–4	35	14.5	θ^g	−20	

[a] See Figs 1, 2, 3, and 4 for methemerythrin, cytochrome *b*-562, cytochrome *c'*, and bovine somatotropin, respectively. [b]The central axis of a four-helix bundle is defined by Eq. 3. [c]Ω_0 is the orientation angle between two helices as defined by Chou et al. (1983, 1984). An explanation of such a definition can also be found in Chou et al. (1988). [d]D is the distance of closest approach between two helix axes as defined by Chou et al. (1983). [e]Ω_i is the tilted angle of the ith helix to the central axis of the four-helix bundle. [f]R_i is the closest approach (Chou et al., 1983) of the ith helix axis to the central axis of the four-helix bundle. [g]The twisted angle of the four-helix bundle as defined in Eq. 5.

helix bundle proteins investigated. However, as far as intersegment energy is concerned, the loop-helix interaction energy is not only much stronger than the interloop interaction energy but also stronger than the interhelix interaction energy. This is clearly shown by the data listed in Table 3 that, compared with the interhelix interaction energies E^{α}_{inter} = −337.8, −399.0, −271.9, and −346.5 kcal/mol for methemerythrin, cytochrome *b*-562, cytochrome *c'*, and bovine somatotropin, we have the corresponding loop-helix interaction energies ϵ = −607.1, −626.4, −567.2, and −830.0 kcal/mol.

It has been pointed out that an antiparallel arrangement of neighboring α-helices is favored by electrostatic interactions between the helices (Sheridan et al., 1982; Chou et al., 1988) caused by the large dipole moment of the α-helices (Wada, 1976; Hol et al., 1978; Sheridan and Allen, 1980; Hol et al., 1981). However, as shown in Table 3, for methemerythrin, cytochrome *b*-562, cytochrome *c'*, and bovine somatotropin, we have ϵ_{ES} = −493.8, −562.6, −472.5, and −720.8 kcal/mol, respectively, whose magnitudes are much greater than the corresponding ϵ^{α}_{ES} = −255.2, −305.0, −168.4, and −289.7 kcal/mol. This indicates that even for the electrostatic interaction energy in the 4-α-helix bundle proteins, the interaction between loops and helices is stronger than that among the four α-helices themselves.

It should be pointed out that, as defined in Eq. 1, the data listed in Table 3 for loop-related energy also include the contributions from NH_2- and COOH-terminal segments. However, even if we removed the contributions of the two terminal segments from the loop-helix interaction energy and only considered the contributions of

TABLE 3 **Various energetic terms[a] of the four-α-helix bundle proteins investigated**

Proteins	Energy of α-helix set (kcal/mol)				
	Intrahelix	Interhelix			Total
	E^{α}_{intra}	ϵ^{α}_{ES} [b]	ϵ^{α}_{NB} [c]	E^{α}_{inter} [d]	E^{α}_{tot}
Methemerythrin	−2381.1	−255.2	−82.6	−337.8	−2718.9
Cytochrome *b*-562	−2771.1	−305.0	−94.0	−399.0	−3170.1
Cytochrome *c'*	−2447.1	−168.4	−103.5	−271.9	−2719.0
Bovine somatotropin	−3205.7	−289.7	−56.8	−346.5	−3552.2

Proteins	Energy of loop set (kcal/mol)				
	Intraloop	Interloop			Total
	E^{loop}_{intra}	ϵ^{loop}_{ES} [e]	ϵ^{loop}_{NB} [f]	E^{loop}_{inter} [g]	E^{loop}_{tot}
Methemerythrin	−778.5	−50.1	−13.8	−63.9	−842.4
Cytochrome *b*-562	−538.6	−33.4	−5.3	−38.7	−577.3
Cytochrome *c'*	−856.4	−69.6	−18.1	−87.7	−943.1
Bovine somatotropin	−2025.9	−59.9	−12.7	−72.6	−2098.5

Proteins	Loop-helix interaction energy (kcal/mol)			
	ϵ_{ES} [h]	ϵ_{NB} [i]	ϵ_{BD} [j]	ϵ [k]
Methemerythrin	−493.8	−114.7	1.4	−607.1
Cytochrome *b*-562	−562.6	−66.3	2.5	−626.4
Cytochrome *c'*	−472.5	−101.3	6.6	−567.2
Bovine somatotropin	−720.8	−116.4	7.2	−830.0

Proteins	Total energy (kcal/mol)
	$E_{tot} = E^{\alpha}_{tot} + E^{loop}_{tot} + \epsilon$
Methemerythrin	−4168.4
Cytochrome *b*-562	−4373.8
Cytochrome *c'*	−4229.3
Bovine somatotropin	−6480.7

[a] See Eq. 1 and the relevant footnote below for the definition of each of the energetic terms listed in this table. [b] Electrostatic interhelix energy. [c] Nonbonded interhelix energy. [d] $E^{\alpha}_{inter} = \epsilon^{\alpha}_{ES} + \epsilon^{\alpha}_{NB}$. [e] Electrostatic interloop energy. [f] Nonbonded interloop energy. [g] $E^{loop}_{inter} = \epsilon^{loop}_{ES} + \epsilon^{loop}_{NB}$. [h] Electrostatic loop-helix interaction energy. [i] Nonbonded loop-helix interaction energy. [j] Bonded loop-helix interaction energy. [k] $\epsilon = \epsilon_{ES} + \epsilon_{NB} + \epsilon_{BD}$.

the three loop segments, we found that it still had the values of $\epsilon = -378.4, -486.1, -437.2$, and -692.4 kcal/mol for methemerythrin, cytochrome *b*-562, cytochrome *c'*, and bovine somatotropin, respectively, indicating also stronger than the corresponding interhelix interaction energies.

Therefore, the loop-helix interaction plays a significant role in stabilizing the 4-α-helix bundle motif structure not only for those proteins containing long loops such as bovine somatotropin (Carlacci et al., 1991) but also for those proteins containing very short loops such as methemerythrin, cytochrome *b*-562, and cytochrome *c'*.

The current calculated results indicate that using CHARMM potential would not change the fact that the loop-helix interaction energy is much stronger than the interhelix interaction energy in the 4-α-helix bundle motif structures even though the electrostatic energy in CHARMM is much more heavily emphasized than in ECEPP.

WORK AHEAD

Since the solvent effect is not explicitly included in the calculation, the helix-loop interaction may not as strong as our estimate indicates. In all four proteins investigated, the percentage content of hydrophobic residues is higher in the helix region than in the loop and terminus region. For methemerythrin, there are 36.4% hydrophobic residues in the helix region and 30.6% in the rest of the protein. For cytochrome *b*-562, the numbers are 43.4% and 30.4%; for cytochrome *c'* 51.3% and 33.3%, and for bovine somatotropin 44.2% and 36.8%. Thus, the loop-helix interaction might be attenuated more than the helix–helix interaction when explicit solvent and ionic charge effects are included. Further studies, such as calculations in the presence of explicit solvent molecules, and efforts on improving the semi-empirical potentials (especially for more accurately reflecting ionic charge interactions and hydrophobic effects), are definitely needed to reveal the essence for understanding the forces that determine tertiary structure in proteins.

CONCLUSION

In a 4-α-helix bundle protein, as far as the intrasegment interaction is concerned, the intrahelix interaction energy is much stronger than the other intrasegment interaction energy; however, as far as intersegment interaction is concerned, the loop-helix interaction would play a dominant role in stabilizing the protein structure. Even for the part of electrostatic interaction energy, the interaction between loops and helices is stronger than that among the four α-helices themselves although the latter involves the interaction of the large dipole moments due to the antiparallel arrangement of neighboring α-helices (Wada, 1976; Hol et al., 1978; Sheridan and Allen, 1980; Hol et al., 1981). Such a fact holds true no matter what loops, long or short, the 4-α-helix bundle protein possesses, and also no matter which empirical potential, ECEPP or CHARMM, is used for calculations. Our calculated results reaffirm that no appropriate conclusion can be drawn in arguing whether the dipole interaction among the four α-helices play a stabilizing role or destabilizing role for a 4-α-helix bundle protein without taking into consideration the effect of interaction between helices and loops (Carlacci and Chou, 1990).

We wish to thank Dr. K. Kirschner for sending us the preprint of his group as cited in this paper. We also wish to thank Drs. C. D. Paliakasis and M. Kokkinidis whose questions raised in their letter stimulated the authors to initiate this work.

The research at Northern Illinois University has been supported in part by the National Science Foundation through the President Young Investigator program (CHE-91-57717), the National Center for Super-

computing Applications at University of Illinois at Urbana-Champaign, and Northern Illinois University Graduate School.

Received for publication 24 January 1992 and in final form 24 April 1992.

REFERENCES

Argos, P., M. G. Rossmann, and J. E. Johnson. 1977. A four-helical supper-secondary structure. *Biochem. Biophys. Res. Commun.* 75:83–86.

Banner, D. W., M. Kokkinidis, and D. Tsernoglou. 1987. Structure of the ColE1 Rop protein at 1.7 Å resolution. *J. Mol. Biol.* 196:657–675.

Bernstein, F. C., T. F. Koetzle, G. J. B. Williams, E. F. Meyer, Jr., M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi. 1977. The protein data bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* 112:535–542.

Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.* 4:187–217.

Carlacci, L., and K. C. Chou. 1990*a*. Energetic approach to the folding of four α-helices connected sequentially. *Protein Eng.* 3:509–514.

Carlacci, L., and K. C. Chou. 1990*b*. Electrostatic interactions between loops and α-helices in four-helix bundle proteins. *Protein Eng.* 4:225–227.

Carlacci, L., K. C. Chou, and G. M. Maggiora. 1991. A heuristic approach to predicting the tertiary structure of bovine somatotropin. *Biochemistry.* 30:4389–4398.

Chothia, C., M. Levitt, and D. Richardson. 1977. Structure of proteins: packing of α-helices and pleated sheets. *Proc. Natl. Acad. Sci. USA.* 74:4130–4134.

Chothia, C., M. Levitt, and D. Richardson. 1981. Helix to helix packing in proteins. *J. Mol. Biol.* 145:215–250.

Chou, K. C., and L. Carlacci. 1991. Energetic approach to the folding of α/β barrels. *Proteins: Struct. Funct. Genet.* 9:280–295.

Chou, K. C., L. Carlacci, and G. M. Maggiora. 1990*a*. Conformational and geometrical properties of idealized β-barrels in proteins. *J. Mol. Biol.* 213:315–326.

Chou, K. C., G. Némethy, and H. A. Scheraga. 1990*b*. Energetics of interactions of regular structural elements in proteins. *Accounts of Chemical Research.* 23:134–141.

Chou, K. C., G. Némethy, and H. A. Scheraga. 1983*a*. Role of interchain interactions in the stabilization of the right-handed twist of β-sheets. *J. Mol. Biol.* 168:389–407.

Chou, K. C., G. Némethy, and H. A. Scheraga. 1983*b*. Energetic approach to the packing of α-helices. 1. Equivalent helices. *J. Phys. Chem.* 87:2869–2881.

Chou, K. C., G. Némethy, and H. A. Scheraga. 1984. Energetic approach to the packing of α-helices. 2. General treatment of nonequivalent and nonregular helices. *J. Am. Chem. Soc.* 106:3161–3170.

Chou, K. C., G. Némethy, S. Rumsey, R. W. Tuttle, and H. A. Scheraga. 1985. Interactions between an α-helix and a β-sheet: energetics of α/β packing in proteins. *J. Mol. Biol.* 186:591–609.

Chou, K. C., G. M. Maggiora, G. Némethy, and H. A. Scheraga. 1988. Energetics of the structure of the four-α-helix bundle in proteins. *Proc. Natl. Acad. Sci. USA.* 85:4295–4299.

Finzel, B. C., P. C. Weber, K. D. Hardman, and F. R. Salamme. 1985. Structure of ferricytochrome c' from *Rhodospirillum Molischianum* at 1.67 Å Resolution. *J. Mol. Biol.* 186:627–643.

Lederer, F., A. Glatigny, P. H. Bethge, H. D. Bellamy, and F. S. Mathews. 1981. Improvement of the 2.5-Å resolution model of cytochrome b_{562} by redetermining the primary structure and using molecular graphics. *J. Mol. Biol.* 148:427–448.

Leszczynski, J. F., and G. D. Rose. 1986. Loops in globular proteins: a novel category of secondary structure. *Science (Wash. DC).* 234:849–855.

Hol, W. G. J., L. M. Halie, and C. Sander. 1981. Dipoles of the α-helix and β-sheet: their role in protein folding. *Nature (Lond.).* 294:532–536.

Hol, W. G. J., P. T. van Duijnen, and H. J. C. Berendsen. 1978. The α-helix dipole and the properties of proteins. *Nature (Lond.).* 273:443–446.

IUPAC-IUB Commission on Biochemical Nomenclature. 1970. Abbreviations and symbols for the description of the conformation of polypeptide chains. *Biochemistry.* 9:3471–3479.

Momany, F. A., R. F. McGuire, A. W. Burgess, and H. A. Scheraga. 1975. Energy parameters in polypeptides. VII. Geometric parameters, partial atomic charges, nonbonded interaction, hydrogen bond interactions, and intrinsic torsional potentials for the naturally occurring amino acids. *J. Phys. Chem.* 79:2361–2381.

Némethy, G., M. S. Pottle, and H. A. Scheraga. 1983. Energy parameters in polypeptides. 9. updating of geometrical parameters, nonbonded interaction, and hydrogen bond interactions for the naturally occurring amino acids. *J. Phys. Chem.* 87:1883–1887.

Richardson, J. S. 1981. The anatomy and taxonomy of protein structure. *Adv. Protein. Chem.* 34:167–339.

Sheridan, R. P., and L. C. Allen. 1980. The electrostatic potential of the α-helix. *Biophys. Chem.* 11:133–136.

Sheridan, R. P., R. M. Levy, and F. R. Salemme. 1982. α-Helix dipole model and electrostatic stabilization of 4-α-helical proteins. *Proc. Natl. Acad. Sci. USA.* 79:4545–4549.

Stenkamp, R. E., L. C. Sieker, and L. H. Jensen. 1983. Adjustment of restraints in the refinement of methemerythrin and azidomethemerythrin at 2.0 angstroms resolution. *Acta Crystallogr. Sect. B.* 39:697–703.

Thornton, J. M., B. L. Sibanda, M. S. Edwards, and D. J. Barlow. 1988. Analysis, design and modification of loop regions in proteins. *Bio Essays.* 8:63–69.

Urfer, R., and K. Kirschner. 1991. The importance of surface loops for stabilizing an eight-fold $\beta\alpha$ barrel protein. *Protein Sci.* 1:31–45.

Wada, A. 1976. The α-helix as an electric macro-dipole. *Adv. Biophys.* 9:1–63.

Weber, P. C., and F. R. Salemme. 1980. Structural and functional diversity in 4-α-helical proteins. *Nature (Lond.).* 287:82–84.

Weiner, S. J., P. A. Kollman, P. A. Case, U. Chandra Singh, C. Ghio, G. Alagona, S. Profeta, Jr., and P. Weiner. 1984. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* 106:765–784.