Biochemistry

Volume 14, Number 16 August 12, 1975

Conformational Energy Refinement of Horse-Heart Ferricytochrome *c*†

Paul K. Warme and Harold A. Scheraga*

ABSTRACT: The reported X-ray structure of horse-heart ferricytochrome *c* has been refined by conformational energy calculations, using a three-stage computational procedure. In stage I, the atomic positions are adjusted to conform to idealized bond lengths and bond angles characteristic of small amino acid derivatives, while yet remaining as close as possible to the X-ray coordinates. In stage II, atomic overlaps are eliminated by adjusting the backbone and side-chain dihedral angles to minimize the nonbonded energy, hydrogen-bonded energy, and rotational energy contributions. In the final stage of refinement, the electrostatic energy and a more accurate hydrogen-bonded energy treatment are considered, in addition to the energy contributions of stage II. A "fitting potential" of gradually decreasing strength is imposed in both stages II and III, in order to keep the computed structure as similar to the X-ray structure as is consistent with a low-energy conformation. The final computed structure of cytochrome *c* exhibits a very low conformational energy (−504 kcal/mol) and also closely resembles the X-ray structure (RMS deviation = 0.77 Å for all atoms). However, a special treatment was required in order to alter the location of the phenyl ring of phenylalanine-82. In contrast to the originally published X-ray structure, which shows the phenyl ring pointing away from the heme, the phenyl ring in the computed structure is tucked into the heme crevice, in a position similar to that observed in the reduced form of tuna cytochrome *c*, in the oxidized form of *Rhodospirillum rubrum* cytochrome c_2, and also in the recently determined structure of oxidized tuna cytochrome *c*.

The 2.8-Å resolution crystal structure of horse-heart ferricytochrome *c* has been determined by Dickerson et al. (1971). Unfortunately, efforts to improve the resolution beyond 2.8 Å have been hindered because of crystal disorder and defective isomorphism between the parent crystals and the heavy atom derivative crystals. Conformational energy calculations may provide the means for accurately establishing the positions of atoms which are not fully resolved by the X-ray diffraction studies, particularly in cases, such as this, where a higher resolution (<2 Å) structure cannot be determined. Detailed knowledge about the precise position and environment of each atom will undoubtedly be required in order to achieve a full understanding of the mechanism of electron transfers to and from cytochrome *c* and of its interactions with cytochrome oxidase and reductase.

This paper describes the use of conformational energy calculations to compute a low-energy structure for cytochrome *c*, which also meets the requirement that it remain as similar as possible to the X-ray structure. The calculations are based on methods which were developed for the refinement of lysozyme (Warme et al., 1972; Warme and Scheraga, 1973, 1974) and which were subsequently applied to actinomycin D (Ponnuswamy et al., 1973), α-lactalbumin (Warme et al., 1974), and rubredoxin (Rasse et al., 1974).

Methods

With the exceptions noted below, the methods used for refinement of cytochrome *c* closely parallel those applied to lysozyme. However, the presence of the heme group required the specialized treatment described below. The accuracy of the X-ray positions (R. E. Dickerson, personal communication) of the heme atoms in cytochrome *c* was assessed by comparison with the accurately known crystal structure of α-chlorohemin (Koenig, 1965). For this comparison, the propionic acid side chains of α-chlorohemin were first adjusted to approximate their conformation in cytochrome *c*, and then the two sets of heme atoms were superimposed by rigid-body rotations and translations until the best least-squares correspondence was achieved. The re-

† From the Department of Biochemistry, The Pennsylvania State University, University Park, Pennsylvania 16802, and the Department of Chemistry, Cornell University, Ithaca, New York 14850. *Received January 28, 1975.* This work was supported by research grants from the Research Corporation, from The Pennsylvania State University, and from the National Science Foundation (BMS74-00090) to P.K.W. and by grants to H.A.S. from the National Science Foundation (BMS71-00872 A04), from the National Institute of General Medical Sciences of the National Institutes of Health, U.S. Public Health Service (GM-14312), and from Walter and George Todd.

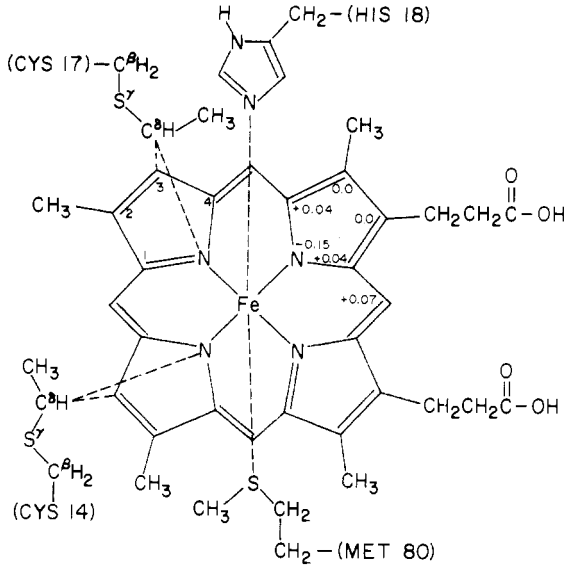

FIGURE 1: The heme group of cytochrome *c* is covalently attached to the protein by thioether linkages to Cys-14 and Cys-17, and the ligands to the heme iron are supplied by His-18 and Met-80. The distances used for fixing the positions of these atoms with respect to the heme (according to eq 1 and 2) are shown as dashed lines. The electrostatic charges used for the heme group are indicated on the upper right pyrrole ring (the same charges are used for the other three pyrrole rings; the methyl side chains are uncharged, and the charges for the propionic acid side chains are the same as glutamic acid).

sulting root mean square deviation for all atoms of the heme group (except for the vinyl side chains, which were treated as ethyl groups attached to cysteine residues 14 and 17, as described below) was only 0.27 Å; thus, the reported X-ray coordinates for the heme atoms of cytochrome *c* were judged to be accurate enough to permit their use without idealization of bond lengths and angles. Another reason for accepting the reported heme coordinates (without adjustment) is that the heme group is known to be somewhat flexible; thus, its special environment in cytochrome *c* might well lead to minor perturbations of its structure with respect to any other known model compounds.

Throughout these calculations, the heme group was held fixed in space while the protein backbone and side-chain atoms were allowed to move to positions of minimum potential energy. However, the movements of four atoms with respect to the heme group were constrained by a strong fitting potential in order to maintain their appropriate positions in relation to the heme group. The two ligands to the heme iron (imidazole N^{ϵ} of histidine-18, and S of methionine-80) were held close to their X-ray coordinates by a fitting potential of the form

$$E_{F1} = W\sum_{i}(D_i - D_{i0})^2 \qquad (1)$$

where D_i represents the computed distance between atom i and the heme iron atom, and D_{i0} represents the corresponding distance in the X-ray structure of cytochrome *c*. It should be noted that the weighting constant W used for the four atoms bound to the heme is considerably stronger (W = 1000 kcal/(mol Å^2)) than the normal fitting coefficient (W = 50 or 10 in stage II, W = 1 in stage III) applied to all other atoms of the protein. The other two atoms which were held in comparatively rigid positions relative to the heme group are the α-carbon atoms of its two vinyl side chains, which are covalently attached by thioether linkages to cysteine-14 and cysteine-17 in cytochrome *c*. Instead of constraining the bond angles and bond lengths associated with these thioether linkages, cysteine residues 14 and 17 were converted to *S*-ethylcysteine residues, thus incorporating the vinyl atoms of the heme in a manner which ensures retention of appropriate bond lengths and angles. The δ-carbon atoms of these *S*-ethylcysteine residues (Figure 1) were then constrained to remain at their X-ray positions by applying a fitting potential of the form

$$E_{F2} = W\sum_{i,j}[(D_i - D_{i0})^2 + (D_j - D_{j0})^2] \qquad (2)$$

with W = 1000 kcal/(mol Å^2) and where D_i and D_j represent the distances from the δ carbon of the *S*-ethylcysteine residue (corresponding to the CH-carbon atoms of its two vinyl side chains) to pyrrole atom 3 and to the pyrrole nitrogen atom, respectively, and D_{i0} and D_{j0} represent the corresponding distances in the X-ray structure. These constrained distances are illustrated as dashed lines in Figure 1. In other words, the δ carbons are forced to remain near the intersection of the spheres centered on these pyrrole atoms.

The bond lengths and bond angles of the amino acid residues and the energy parameters used in this work are based on crystal structures of amino acids and peptides (Momany et al., 1975); they are identical with those used in the refinement of lysozyme (Warme and Scheraga, 1973, 1974), except for the atoms of the heme group. For purposes of the nonbonded energy calculations, the pyrrole ring atoms were treated as aromatic carbons analogous to C^{γ} of phenylalanine, the methine bridge carbons were treated as aromatic carbons analogous to the phenyl ring atoms of phenylalanine, the pyrrole nitrogens were treated like the imidazole nitrogens of histidine, the methyl side chains were treated the same as alanine β-methyl groups, and the propionic acid side-chain atoms were treated the same as the analogous side-chain atoms of glutamic acid. The partial charges of the heme atoms were based on the extended Hückel molecular orbital calculations reported by Zerner and Gouterman (1966) for low-spin ferric porphine cyanide. Net charges of less than 0.01 electronics unit were condensed on adjacent atoms, and minor adjustments of charges were made in order to preserve electroneutrality of the heme group. The resulting partial charges used in this work for the ring atoms of the heme group are indicated in Figure 1. The methyl groups were given a net zero charge, and the propi-

Table I: RMS Deviations from X-Ray Coordinates at Each Stage of Refinement of Cytochrome *c*.

Stage	Iteration	RMS Deviations[a] (Å) Backbone Atoms	Side-Chain Atoms	All Atoms	W [kcal/(mol Å^2)]	ΣD_i^2 (Å^2)	E_{FP} (kcal/mol)
I		0.39	0.69	0.53		232	
II	1	0.57	0.64	0.60	50	297	14,850
II	2	0.57	0.65	0.60	10	300	3,000
II	3	0.53	0.63	0.57	10	266	2,660
III	1	0.63	0.86	0.73	1	432	432
III	2	0.68	0.90	0.77	1	475	475

[a] The RMS deviation is computed as the square root of the sum of squares of the deviations of the computed coordinates from the corresponding X-ray coordinates, divided by the number of atoms compared.

Table II: Energies of Cytochrome *c* Conformations at Each Stage of Refinement.

Stage	Iteration	Energy (kcal/mol)[a]					
		NB[b]	HB[b]	EL[b]	ROT[b]	F1 + F2[b]	Total
I		9.1×10^7	854	−800	865	157	9.1×10^7
II	1	3638	2	−815	943	88	3856
II	2	1191	−52	−819	877	15	1212
II	3	552	15	−839	792	15	535
III	1	−33	−53	−849	645	5	−285
III	2	−126	−60	−850	529	3	−504

[a] Per mole of protein, calculated with the energy functions used in stage III. [b] Energy abbreviations are: NB = nonbonded, HB = hydrogen-bonded, EL = electrostatic, ROT = rotational, F1 + F2 = fitting potentials (eq 1 and 2).

onic acid groups were given the same charges as the side-chain atoms of glutamic acid.

Unfortunately, theoretical calculations have not yet yielded precise information about the charge distribution for the total system of a heme group coordinated to methionine and histidine ligands, as found in cytochrome *c*. Thus, the approach taken in these calculations was to assign a net zero charge to the heme iron atom in order to maintain overall electroneutrality of the protein when considered as a whole. Although the ferric iron atom bears a net formal charge of +1, this charge is probably at least partially distributed over its six ligands (Zerner and Gouterman, 1966) and is probably further compensated by a bound anion somewhere in the vicinity (Dickerson, 1974). We have therefore relied on the fitting potentials (E_{F1} of eq 1) applied to the fifth and sixth position ligand atoms (N^{ϵ} of histidine-18 and S^{δ} of methionine-80) to maintain the integrity of the heme coordination sphere. It seems likely that the six ligands to the heme iron would quite effectively shield the remainder of the protein from any major perturbations due to the heme iron atoms.

Results

A. Stage I Refinement. The initial adjustment of coordinates to conform to standard bond lengths and bond angles produced a structure with an overall RMS deviation from the X-ray coordinates of only 0.53 Å (Table I). Thus, the initial fit was even better than that obtained previously for lysozyme (RMS = 0.66 Å, see Table III of Warme and Scheraga, 1974). However, as shown by the high value of E_{NB} in Table II, a number of major atomic overlaps resulted in a high nonbonded energy after stage I. The worst overlaps occurred between the side-chain atoms of phenylalanine-46 and the ring atoms of proline-30, which gave rise to essentially all (>99%) of the high nonbonded energy. These overlaps resulted from a poor fit (average deviation = 2.14 Å) of the side-chain atoms of phenylalanine-46 to the X-ray coordinates, even though the backbone atoms of residue 46 and the adjoining residues were very close. This problem apparently resulted from an unusually large N-C^{α}-C^{β} angle (128° instead of 109.5°) combined with a low C^{α}-C^{β}-C^{γ} angle (98° instead of 113°) in the X-ray structure for the side chain of phenylalanine-46, thus preventing a close fit to the side chain during stage I. As discussed below, a substantial movement of the backbone atoms of residues 46, 47, and 48 was required in order to eliminate atomic overlaps. In addition to these overlaps, 12 other overlaps contributed energies between 1000 and 6000 kcal, 23 overlaps contributed energies between 100 and 1000 kcal, and 102 overlaps contributed energies between 10 and 100 kcal/mol.

B. Stage II Refinement. In the first iteration of the stage II refinement (Warme and Scheraga, 1973), ten cycles (or more, as noted) of energy minimization were conducted on each of the following segments: 1–19, 18–37, 36–55 (15 cycles), 54–73 (15 cycles), 72–91, and 90–104 (15 cycles). In three of these sections, the energy was still comparatively high after ten cycles of minimization; hence, the minimization was continued for five additional cycles. The appreciable decrease in the nonbonded energy shown in Table II indicates that most of the major atomic overlaps have been relieved, at the expense of an increase in the overall RMS deviation from the X-ray coordinates (Table I).

In the second iteration of the stage II refinement, the energies of the following segments, which overlap the segments of iteration 1, were each minimized for ten cycles, with exceptions as noted: 1–14 (15 cycles), 13–32, 31–50 (15 cycles), 49–68, 67–86, and 85–104. As shown by Table II, another significant drop in energy resulted from this second iteration. However, the RMS deviation from the X-ray coordinates did not increase significantly (Table I), although the strength of the fitting coefficient W had been decreased from 50 kcal/(mol Å^2) in the first iteration to 10 in the second, thus allowing greater movement away from the X-ray coordinates, in some cases, to reduce the energy.

At this point, an attempt was made to continue on to stage III of refinement (see below), but the results were poor. Many atoms moved to positions quite distant (>1 Å) from their positions in the X-ray structure, apparently because the fitting coefficient W had been prematurely lowered to 1 (as is customary in stage III), while the interaction energies of these atoms were sufficiently high that the fitting constraint was overpowered, thus allowing these atoms to move too far from their X-ray positions. At this point, 34 atomic overlaps contributing nonbonded energies between 10 and 64 kcal/mol had not been eliminated, and these were quite widely distributed throughout the protein, involving 37 different residues (10, 20, 21, 22, 24, 30, 36, 38, 42, 43, 44, 46, 48, 58, 60, 61, 62, 63, 64, 65, 66, 68, 69, 71, 74, 80, 81, 82, 83, 84, 85, 91, 97, 98, 99, 100, and 102). Thus, a third iteration of stage II refinement was carried out before continuing with stage III.

In the final iteration of the stage II refinement, a number of liberties were exercised which were not taken in the more conventional first and second iterations. Segments 1–18, 31–37, 51–58, and 68–78, which did not contain significant overlaps, were effectively skipped over by simply constraining them to coincide with their positions at the end of the second iteration. Segment lengths were not restricted to 20 residues, and in some cases were as short as 7 residues. In other cases, it proved beneficial to refit a segment to the original X-ray coordinates by simultaneous adjustment of

both backbone and side-chain dihedral angles, before continuing the energy minimization. The energy minimization was not limited to 10 or 15 cycles as in previous iterations, but was continued until a suitably low-energy conformation was obtained. In several cases (described in the next paragraph) where high energy conformations coincided with regions of poor fit to the X-ray coordinates, it was necessary to rotate a peptide group (-CO-NH-) by 180° in order to achieve a low-energy conformation. Many of these alternative approaches were tried in various segments during this third iteration, but only those changes (outlined in the next paragraph) which improved the structure (by the criterion of lowering the energy while preserving a close correspondence to the X-ray coordinates) were incorporated into the final structure.

In order to relieve several overlaps within the tight bend involving residues 19–25, this segment was refitted to the X-ray coordinates by the stage I procedure, and then its energy was minimized for 40 cycles, resulting in an energy decrease of about 80 kcal and a 3% closer fit. The energy of segment 24–32 was minimized for 27 cycles, and then residues 31–37 were readjusted to conform to their positions at the end of iteration 2. Segment 38–50 contained several high-energy areas, involving overlaps between the side-chain atoms of arginine-38 and glutamine-42, overlapping backbone atoms of alanine-43 and proline-44, overlaps between proline-30 and the side-chain atoms of phenylalanine-46 and tyrosine-48, and side-chain atom overlaps between glutamine-42 and tyrosine-48. The energy of this segment was improved by 16 cycles of energy minimization subsequent to 180° rotation of the peptide groups joining residues 40–41 and 46–47, and by allowing the peptide group joining residues 42 and 43 to rotate more freely (by omitting the fitting constraint normally applied to the oxygen and nitrogen atoms of all of these peptide groups). These changes also improved the fit of this segment to the X-ray coordinates, since many of the atoms close to the sites of the overlaps mentioned above had moved quite far from their original positions in the course of relieving atomic overlaps during previous iterations of stage II refinement. The segment 51–58 was then readjusted to approximate its position at the end of iteration 2, and segment 57–68 was fitted to the X-ray coordinates prior to energy minimization for 15 cycles. Segment 68–78 was left in its position at the end of iteration 2, and then 15 cycles of energy minimization were carried out on each of the segments 79–92 and 91–104.

During the third iteration of the stage II refinement, the energy was significantly reduced, as shown in Table II. At this point, only eight atomic overlaps contributing nonbonded energies ranging between 10 and 21 kcal/mol remained in the computed structure. Furthermore, the overall RMS deviation from the X-ray coordinates was decreased by 5% during this third iteration (Table I), thus indicating that the changes which had been made led to a decrease in energy while maintaining a close correspondence to the X-ray structure of cytochrome *c*.

C. Stage III Refinement. The procedures already described for the case of lysozyme (Warme and Scheraga, 1974) were followed during the stage III refinement of cytochrome *c*, except that the computation time was reduced by performing only N cycles of energy minimization on each of the segments listed below, where N represents the number of variable dihedral angles in the segment. Although $2N$ cycles of minimization were carried out on each segment of lysozyme, experience showed that the energy decreases which occurred during the second N cycles were small compared to the energy decreases during the first N cycles. During the first iteration, the energies of segments 1–9, 8–17, 16–25, 24–33, 32–42, 41–50, 49–58, 57–66, 65–74, 73–82, 81–90, 89–98, and 97–104 were minimized. The results summarized in Table II indicate a substantial decrease in the energy of the computed structure during this iteration, most of this energy decrease being in the nonbonded energy term (E_{NB}). This reflects the fact that most of the minor atomic overlaps persisting through stage II have been eliminated at this point. However, the 28% increase in the RMS deviation from the X-ray coordinates (Table I) shows that this decrease in energy could be accomplished only by significant movements of some atoms away from their positions in the X-ray structure.

The energies of segments 1–5, 4–13, 12–21, 20–28, 27–36, 35–44, 43–53, 52–61, 60–69, 68–78, 77–85, 84–94, and 93–104 were minimized during iteration 2 of stage III refinement. After minimization of segment 43–53, it became apparent that rotation of the peptide group joining residue 46–47 (during iteration 3 of stage II) had resulted in an unfavorable rotational energy associated with the backbone dihedral angle ϕ of residue 47, in spite of the apparent improvement in the fit of these residues to the X-ray coordinates. Therefore, this peptide group was returned to the conformation existing at the end of iteration 2 of stage II before repeating the energy minimization on this segment.

A major alteration of the conformation of residues 81–85 was incorporated during this second iteration of stage III refinement, in response to recent reinterpretation of the X-ray structure in this region (Dickerson and Timkovich, 1975). In the oxidized form of horse-heart cytochrome *c*, phenylalanine-82 was originally thought to be swung out of the heme crevice, in contrast to the situation observed in cytochrome c_2 and tuna cytochrome *c* (both oxidized and reduced), in which this phenylalanine ring is located within the crevice between the side chains of lysine-13 and methionine-80. In light of these more recent (and more accurate) structure determinations, it now appears that phenylalanine-82 in horse cytochrome *c* should also be located in the heme crevice (Dickerson and Timkovich, 1975). In order to incorporate this conformation change in our computed structure, the fitting constraint on all atoms of isoleucine-81, phenylalanine-82, and alanine-83 was eliminated, thus allowing these atoms to move away from their X-ray positions. At the same time, a fitting potential was applied to the phenyl ring atoms of phenylalanine-82, in order to induce these atoms to move to a point centered between N^{ϵ} of lysine-13 and C^{γ} of methionine-80. Although all six ring atoms cannot occupy this same position, the optimum fit will be satisfied by any one of a large number of conformations in which these atoms are equidistant from the reference point. The resulting freedom for movement of the phenyl ring during the course of energy minimization is a desirable feature of this approach, since its exact position is uncertain. The conformation of segment 81–85 was therefore adjusted by a least-squares fitting procedure to optimize the fit of the phenyl ring atoms to this point within the heme crevice, while preserving the fit of residues 84 and 85 to the X-ray coordinates. The energy of the resulting conformation was then minimized by successive application of the stage II and stage III refinement procedures. These adjustments of segment 81–85 were carried out prior to the second iteration of the stage III refinement of residues 52–104

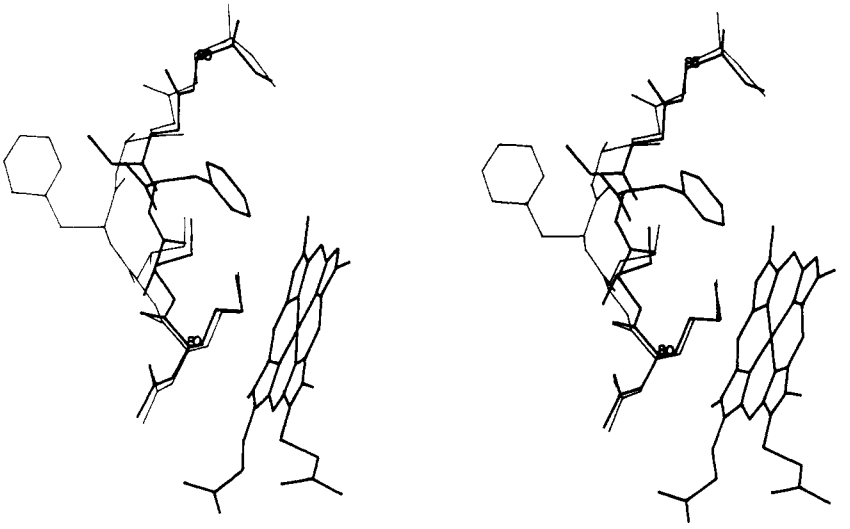

FIGURE 2: Stereo drawing of segment 80–85, comparing the computed structure (thick line) with the X-ray structure (thin line). Note the pronounced change in position of the phenyl ring of Phe-82. The sequence shown is 80-Met-Ile-Phe-Ala-Gly-Ile-85.

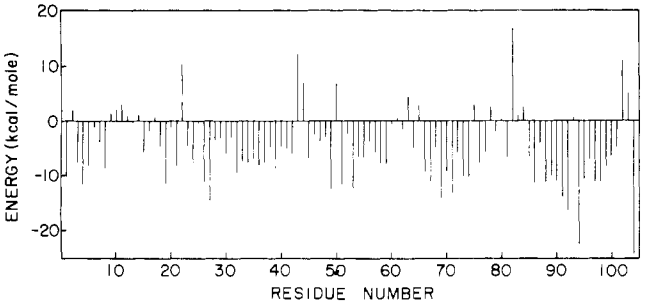

FIGURE 3: Plot of total energy ($E_{NB} + E_{HB} + E_{ROT} + E_{EL}$) for each residue of cytochrome *c* after stage III refinement, calculated as one-half of the sum of the energies of interaction with all other atoms of the entire protein which are within the cutoff distance (Warme and Scheraga, 1974).

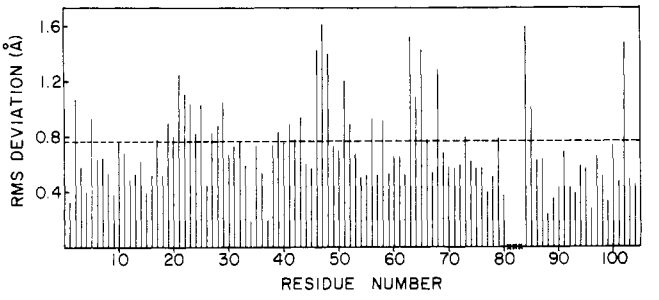

FIGURE 4: Plot of RMS deviation for each residue of cytochrome *c* after stage III refinement. The overall RMS deviation is shown as a dashed line.

so that other amino acid residues near segment 81–85 would have an opportunity to adapt to the new conformation in this region.

D. Characteristics of the Final Computed Structure. As shown in Table II, the energy again decreased in iteration 2 of stage III, but only by about one-fourth as much as in the first iteration, thus indicating that the energy is approaching its minimum value. There are no atomic overlaps (E_{NB} >10 kcal/mol), but 46 close contacts remain, each of which contributes a nonbonded energy within the range of 1–9 kcal/mol. However, 10 of these close contacts involve nitrogen and oxygen atoms participating in a hydrogen bond, and, in such cases, the favorable electrostatic and hydrogen bond energies more than compensate for the unfavorable nonbonded energy. Five additional close contacts involve interactions of the γ-carbon atoms of the branched aliphatic amino acids valine and isoleucine, with the hydrogen atom of the amino group on the same residue. Ten other close contacts involve backbone atoms of adjacent residues; however, some of these represent seven-member hydrogen-bonded ring conformations, which yield a net favorable energy when other electrostatic interactions are taken into account. The remaining 21 close contacts follow no established pattern, but it is worth noting that the highest energy for a single interaction (9.0 kcal/mol) involves a close approach of $C^{\delta 1}$ of phenylalanine-82 to the carbonyl oxygen of isoleucine-81. Two other close contacts (C^{γ} of phenylalanine-82 to O of isoleucine-81, and $C^{\epsilon 2}$ of phenylalanine-82 to H^{η} of lysine-13) also involve the phenyl ring of phenylalanine-82, thus suggesting that the position assigned to these atoms (for purposes of fitting the phenyl ring into the heme crevice) was not quite correct, although a movement of less than 0.3 Å could alleviate these close contacts. The X-ray positions of residues 80–85 are compared with their final computed positions in Figure 2.

The total energies ($E_{NB} + E_{HB} + E_{ROT} + E_{EL}$) for each amino acid residue of the final computed structure for cytochrome *c* are plotted in Figure 3. Most residues have negative energies, while only four exhibit total energies (for all interactions of every atom in the given residue) higher than 10 kcal/mol. The highest overall energy involves phenylalanine-82, primarily due to several close contacts discussed above. The fitting energies for the ligands histidine-18 and methionine-80 (E_{F1}) and for the thioether linkages of cysteine residues 14 and 17 (E_{F2}) are quite low in the final computed structure, as indicated by the low value of $E_{F1} + E_{F2}$ in Table II. The distances of these atoms from their respective reference atoms in the heme group are all within 0.07 Å of the corresponding distances in the X-ray structure.

The final RMS deviations summed over all atoms of each residue are presented in Figure 4. The deviations of residues 81–83 were not included in this figure, due to the fact that the X-ray coordinates of phenylalanine-82 are known to be incorrect, and the drastic change in its position necessitates considerable movements of the adjacent residues. The high RMS deviations of residues 84 and 85 indicate that these

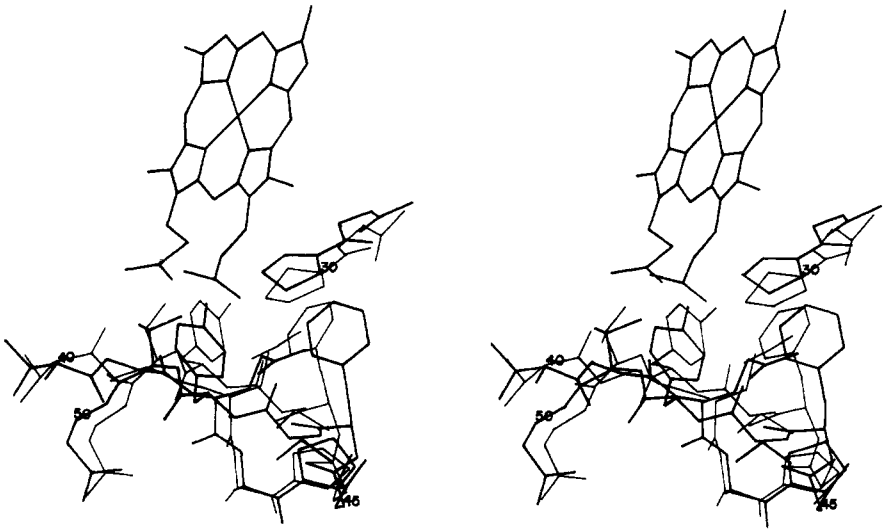

FIGURE 5: Stereo drawing of segment 40–50, comparing the computed structure (thick line) with the X-ray structure (thin line). Note the inversion of the peptide group joining residues 40 and 41, the significant shift of the Phe-46, Thr-47, and Tyr-48, and the close approach of Phe-46 to Pro-30 in the X-ray structure. The segments shown are 29-Gly-Pro-30 and 40-Thr-Gly-Gln-Ala-Pro-45-Gly-Phe-Thr-Tyr-Thr-Asp-50.

residues are also forced to reorient somewhat in order to compensate for the repositioning of phenylalanine-82. In general, the amino acid residues exhibiting the worst RMS deviations from their X-ray positions tended to fall in clusters which coincided with regions where difficulties had been experienced during the energy minimization. For example, two sharp bends occur in the backbone chain in the region of residues 19–24, and overlaps between backbone atoms in adjacent residues caused high energies in stage II, even though the fit to the X-ray coordinates in stage I was quite good (RMS deviations for residues 19–24 were 0.59, 0.59, 0.62, 0.17, 0.57, and 0.36, respectively). Two close contacts between backbone atoms (N_{21} to O_{19}, and N_{23} to O_{21}) persisted even in the final computed structure, and each of these contributed a nonbonded energy of 3.5 kcal/mol.

Another position where large RMS deviations in the final structure coincided with a region of high initial energy is segment 40–50. As described earlier, it proved necessary to rotate the peptide group joining residues 40 and 41 by approximately 180° and also to allow the peptide group joining residues 42 and 43 to rotate by approximately 120° in order to obtain a low energy structure in this region. Although the initially severe overlaps between phenylalanine-46 and proline-30 have been eliminated in the final computed structure, this was accomplished only at the expense of substantial movements of residues 46, 47, and 48 away from their X-ray positions, as indicated by their large RMS deviations in Figure 4. The final structure of segment 40–50 is compared to the X-ray structure in Figure 5, in order to illustrate the significant changes in this region.

The large RMS deviations of residues 63, 64, and 65 again reflect large movements induced by severe atomic overlaps which existed at the beginning of stage II, although the fit to these residues in stage I was quite reasonable. In this case, there is good reason to believe that the backbone dihedral angles of residues 61 and 62 are erroneous both in the X-ray structure and in the final computed structure (Table III), since prolines occur at each of these positions in *Crithidia* (Pettigrew, 1972) and *Candida krusei* (Narita and Titani, 1968; Lederer, 1972) cytochrome *c*, respectively. The dihedral angle ϕ of proline is rigidly fixed at about −70° and, thus, the X-ray ϕ values of −7 and 22° are quite unlikely if we assume that the overall conformations of these other species closely resemble that of cytochrome *c*. In contrast, the ϕ values of residues 3, 83, and 88 are consistent with the observation that proline occurs at each of these positions in one or more eucaryotic cytochromes *c*. In each of these cases, the ϕ values of the computed structure are within 30° of the expected ϕ value for proline; thus, relatively small conformational changes in adjacent residues could probably accommodate a proline substitution.

An attempt was made to reduce the overall RMS deviation by rigid-body rotations and translations of the computed structure with respect to the X-ray structure, but no improvement was achieved by such movements.

The final deviations of the computed dihedral angles for cytochrome *c* from the initial X-ray dihedral angles are given in Table III. The average deviation of the computed dihedral angles from their corresponding values in the X-ray structure was 25°, and the same average deviation applies for both backbone and side-chain dihedral angles when averaged separately. A complete listing of the Cartesian coordinates of the computed (stage III, iteration 2) structure of cytochrome *c* is included in the microfilm edition of this journal.

Discussion

A large body of experimental data derived from chemical and physical studies on cytochrome *c* has been reviewed by Margoliash and Schejter (1966), and more recent experimental results have been interpreted at length by the crystallographers (Dickerson et al., 1971; Salemme et al., 1973a,b; Takano et al., 1973). Amino acid sequences have been determined for the cytochromes *c* from at least 60 different eucaryotic organisms, as well as a growing number of procaryotic organisms (Margoliash, 1972; Dickerson and Timkovich, 1974), and sequence comparisons among these species have provided useful information concerning which amino acids may be essential for the biological functions of cytochrome *c*. Crystallographic studies have yielded complete three-dimensional structures for oxidized horse-heart cytochrome *c* at 2.8 Å (Dickerson et al., 1971), for reduced tuna-heart cytochrome *c* at 2.45 Å (Takano et al., 1973), for oxidized *Rhodospirillum rubrum* cytochrome c_2 at 2.45

Table III: Third-Stage Refined Dihedral Angles of Cytochrome *c* and Deviations from X-Ray Values.[a]

Residue	φ	ψ	χ_1	χ_2	χ_3	χ_4
1 GLY	-96(-4)	51(31)				
2 ASP	-72(-40)	-56(9)	-70(5)	75(-18)		
3 VAL	-40(23)	-58(-21)	-91(-136)			
4 GLU	-48(26)	-62(-27)	-78(5)	-67(26)	-73(-6)	
5 LYS	-53(5)	-74(-10)	-155(-6)	70(16)	-164(-57)	172(1)
6 GLY	-32(11)	-65(-1)				
7 LYS	-34(19)	-56(8)	-180(86)	-178(12)	77(-100)	178(3)
8 LYS	-61(-20)	-56(13)	-173(12)	-176(-1)	176(-11)	-180(6)
9 ILE	-57(3)	-16(-12)	-152(22)	172(-7)		
10 PHE	-80(24)	-78(-14)	-145(36)	53(-3)		
11 VAL	-37(12)	-48(-9)	-80(-104)			
12 GLN	-27(2)	-84(17)	-168(19)	-125(39)	127(155)	
13 LYS	-129(-22)	38(-1)	64(-5)	170(-3)	-175(7)	178(-5)
14 CYS	-143(6)	172(1)	-158(0)			
15 ALA	64(11)	13(-14)				
16 GLN	-94(14)	-45(17)	168(36)	-165(28)	-50(144)	
17 CYS	-104(-32)	14(10)	-45(-2)			
18 HIS	-141(-11)	172(-10)	-50(9)	-66(4)		
19 THR	-141(5)	85(-12)	-45(28)	154(****)		
20 VAL	-69(19)	81(65)	-71(-136)			
21 GLU	-171(-59)	99(-5)	161(81)	-159(-9)	-159(-66)	
22 LYS	-23(-10)	99(13)	31(41)	-139(25)	81(-98)	-167(-15)
23 GLY	116(11)	-95(-84)				
24 GLY	50(82)	-158(-56)				
25 LYS	-122(37)	160(15)	-43(-63)	118(27)	62(-117)	-179(-1)
26 HIS	-88(-50)	91(12)	-92(-31)	-14(-7)		
27 LYS	-84(-10)	-152(-26)	-64(-29)	-180(-8)	-168(7)	170(-6)
28 THR	-50(13)	-34(29)	-58(1)	165(****)		
29 GLY	-132(-14)	174(-9)				
30 PRO	-68(-15)	-170(4)				
31 ASN	-110(-14)	121(-18)	-62(-21)	47(47)		
32 LEU	-80(15)	73(6)	-79(6)	-22(-140)		
33 HIS	-109(1)	-61(-2)	-171(17)	-126(18)		
34 GLY	-95(-1)	42(19)				
35 LEU	-63(-25)	-57(32)	-146(-43)	-82(37)		
36 PHE	-79(-38)	124(4)	-70(-15)	-137(-19)		
37 GLY	113(9)	-54(-30)				
38 ARG	-63(41)	164(7)	-143(-5)	128(-68)	44(22)	74(-90)
39 LYS	-80(-19)	-83(-6)	-161(-35)	-172(13)	-71(99)	-179(-1)
40 THR	-148(-13)	170(-133)	24(36)	156(****)		
41 GLY	56(-170)	-155(-12)				
42 GLN	-116(0)	-109(121)	-107(-24)	-78(-15)	54(-3)	
43 ALA	167(-106)	-52(-8)				
44 PRO	-69(-8)	73(-10)				
45 GLY	56(-5)	-127(-46)				
46 PHE	45(49)	96(-11)	-179(-8)	-76(-17)		
47 THR	-44(31)	150(44)	-40(5)	150(****)		
48 TYR	-139(-37)	119(-15)	-119(-49)	99(17)	33(****)	
49 THR	-41(-8)	133(-10)	-37(-15)	165(****)		
50 ASP	-63(-25)	-16(36)	-66(23)	-107(-20)		
51 ALA	-54(-47)	-60(42)				
52 ASN	-88(-23)	-33(-5)	71(-2)	81(-180)		
53 LYS	-50(4)	-52(0)	-169(2)	-162(13)	173(-13)	-178(-3)
54 ASN	-52(-4)	-88(-53)	-168(-17)	-140(-2)		
55 LYS	-45(43)	149(-18)	74(31)	177(-6)	-62(-8)	-169(7)
56 GLY	-125(-25)	36(65)				
57 ILE	-42(-19)	143(8)	-130(-34)	77(38)		
58 THR	-59(-3)	139(-10)	48(-14)	161(****)		
59 TRP	-99(13)	-34(-16)	-102(10)	162(-27)		
60 LYS	-40(17)	100(-43)	45(41)	-174(1)	-166(5)	-172(12)
61 GLU	31(37)	-99(-23)	-29(-7)	175(-7)	91(-9)	
62 GLU	46(23)	-104(-4)	-156(-24)	-178(1)	111(35)	
63 THR	-99(-2)	66(55)	4(114)	163(****)		
64 LEU	-93(-57)	2(47)	-137(15)	-76(-22)		
65 MET	-109(-65)	5(-1)	55(69)	-146(-34)	48(-6)	
66 GLU	-92(26)	-38(37)	-76(14)	-139(44)	-141(-36)	
67 TYR	-61(-53)	-21(71)	-151(-30)	-89(12)	178(****)	
68 LEU	-90(-77)	-38(-2)	-55(28)	149(90)		
69 GLU	-63(33)	-54(-34)	-163(10)	176(-15)	22(6)	
70 ASN	-168(26)	87(13)	-173(-22)	19(28)		
71 PRO	-69(-9)	-24(4)				
72 LYS	-56(-26)	-39(12)	-151(-38)	177(-12)	-175(7)	-180(1)
73 LYS	-75(8)	-55(5)	171(71)	-170(-2)	76(-100)	179(-12)
74 TYR	-49(-10)	-41(1)	-133(-16)	72(-11)	179(****)	
75 ILE	-143(-8)	94(26)	38(6)	-97(1)		
76 PRO	-69(-16)	98(-14)				
77 GLY	139(7)	-22(-5)				
78 THR	-37(1)	103(14)	-95(4)	-55(****)		
79 LYS	-45(-14)	-34(2)	-81(-14)	163(-13)	-163(20)	166(-10)
80 MET	-48(12)	117(-135)	172(-4)	162(-23)	139(-5)	
81 ILE	-86(82)	-176(2)	67(8)	115(23)		
82 PHE	64(-174)	127(151)	107(-1)	67(38)		
83 ALA	-60(-20)	-92(-172)				
84 GLY	167(-57)	-156(42)				
85 ILE	-124(-41)	140(16)	28(31)	99(5)		
86 LYS	-96(-22)	-23(15)	179(50)	-169(-8)	167(-12)	-177(5)
87 LYS	-49(-45)	109(21)	-100(31)	-167(-3)	179(-5)	179(10)
88 LYS	-44(-1)	-59(7)	-162(-6)	74(-8)	-179(-2)	179(6)
89 THR	-46(-1)	-61(2)	26(12)	160(****)		
90 GLU	-47(4)	-49(-6)	-74(-43)	-169(13)	-47(25)	
91 ARG	-54(-3)	-59(-1)	-65(-105)	146(-33)	105(-75)	-176(1)
92 GLU	-47(-2)	-52(-7)	-72(-8)	180(-35)	-68(9)	
93 ASP	-44(14)	-58(4)	-56(-16)	138(-8)		
94 LEU	-49(-19)	-58(7)	-87(25)	164(-62)		
95 ILE	-46(11)	-63(-16)	-66(35)	126(-1)		
96 ALA	-40(17)	-70(-16)				
97 TYR	-37(5)	-60(4)	-169(-12)	75(-17)	-180(****)	
98 LEU	-45(-7)	-75(-11)	-99(-5)	66(-88)		
99 LYS	-36(18)	-55(-7)	-168(-5)	175(-7)	-175(11)	175(-14)
100 LYS	-58(19)	-85(-31)	-102(1)	58(22)	-154(24)	167(-13)
101 ALA	-71(4)	98(103)				
102 THR	155(-83)	-24(-13)	-138(123)	156(****)		
103 ASN	-34(7)	100(2)	-166(-28)	-143(25)		
104 GLU	44(-14)	18(-145)	-62(-21)	164(-21)	129(34)	

[a] Dihedral angles which were undefined by the X-ray coordinates are indicated by (****). The values in parentheses are the deviations from the dihedral angles computed from the X-ray coordinates.

Å (Salemme et al., 1973a) and for oxidized *Micrococcus denitrificans* cytochrome c_{550} at 4 Å (Timkovich and Dickerson, 1973). In spite of this wealth of information about cytochrome *c*, the mechanism of its participation in mitochondrial electron transfer reactions is still not very well understood. At least three different mechanisms for electron transfers to and from the heme iron of cytochrome *c* have been proposed (Dickerson et al., 1972; Takano et al., 1973; Salemme et al., 1973b), but unexplained discrepancies from experimental data cast doubt on each of these alternative mechanisms (Dickerson and Timkovich, 1975). Comparisons of the oxidized and reduced forms of cytochrome *c* have revealed substantial differences between their conformations (Takano et al., 1973), and this suggests rather strongly that conformational changes play a role in the electron transfer mechanism.

In light of these uncertainties regarding the mechanism of cytochrome *c*, precise information about the position and environment of each atom in cytochrome *c* is clearly needed. An accurate calculation of the conformational energy of a protein can be carried out only when the positions of all atoms are known with great accuracy, since movement of a single atom by only a few tenths of an ångstrom can make the difference between a high-energy structure and a low-energy one. Thus, comparison of the energies of two alternative conformations of a protein requires a very accurately known reference structure exhibiting a low conformational energy, such as the energy-refined structure presented in this paper. Comparisons of this type will be essential for understanding the complex conformational changes associated with changes of the oxidation state of cytochrome *c* and also for studying the effects of amino acid replacements on the conformation of cytochrome *c* from other species, as discussed in the accompanying paper (Warme, 1975).

The determination of X-ray coordinates of a protein from crystallographic data suffers from several sources of error, including perturbations of the X-ray structure by neighboring protein molecules in the crystal, by the high ionic strength of the crystallizing medium, and by the presence of the heavy atom in isomorphous derivatives. Crystal disorder often obscures the side chains of external amino acid residues, and other errors are also introduced during the stage at which a physical model of the protein is superimposed on the electron density map to measure the position of each atom. The atomic coordinates of cytochrome *c* used in this study were based on a 2.8-Å resolution map; thus, the provisional nature of these coordinates must be recognized. As this work was nearing completion, R. E. Dickerson (personal communication) informed us that more accurate coordinates will be obtained soon from the 2-Å structure of oxidized tuna cytochrome *c*. At 2.8-Å resolution, groups of atoms are resolved in the electron density map, but individual atoms are not. Thus, it is reasonable to expect that some of the reported atomic coordinates (R. E. Dickerson, personal communication) will differ substantially from their true positions. Since progress toward a higher resolution structure for horse-heart cytochrome *c* is hampered by problems of crystal disorder and lack of isomorphism between heavy atom derivatives (Dickerson et al., 1971), the need for alternative methods of refinement is apparent.

The computed structure of cytochrome *c* presented in this paper is not necessarily the one of lowest possible energy, but rather, it represents the structure corresponding to the local minimum of energy closest to the X-ray structure. In other words, the final dihedral angles of the computed structure are, with few exceptions (see Table III), within 60° of the original X-ray angles. The exceptions generally occur in residues where severe atomic overlaps required large changes in dihedral angles (e.g., residues 40–41, 42–43, and 80–84, as discussed in the Results section). Other exceptions reflect either difficulties in fitting the X-ray coordinates with a structure possessing idealized bond lengths and bond angles or else they reflect changes induced by unusually high or low energy interactions with other atoms in the vicinity. Because the computed structure was constrained to a local energy minimum, it is probably incorrect in those regions where the X-ray structure is grossly inaccurate. In cases where the X-ray coordinates are close to the most favorable local energy minimum (i.e., in the correct potential energy well), the computed structure should represent the true conformation of minimum energy accurately and, thus, should be more accurate than the X-ray structure.

The final computed structure is very similar to the original X-ray structure (with a few exceptions which are discussed in the Results section) and therefore, most of the structural interpretations previously offered by Dickerson et al. (1971) still pertain. For example, the RMS deviations of tryptophan-59, tyrosine-67, tyrosine-74, threonine-78, and methionine-80, all of which have been suggested to play a role in the electron transfer mechanism, are less than 0.60 Å, although the overall RMS deviation is 0.77 Å. In regions where the computed structure differs substantially from the X-ray one, the computed coordinates may be construed as predictions of the actual structure. Because of problems inherent in the X-ray diffraction studies on horse-heart cytochrome *c*, it may prove to be impossible to determine a more accurate structure experimentally. However, the validity of our predicted structure will be tested by comparing it to the 2-Å resolution structure of oxidized tuna cytochrome *c* when it becomes available. This type of analysis will provide information about the reliability of the energy refinement methods used in this work and will also serve as a guide for further efforts to refine the horse-heart structure. Of course, in comparing the structures of horse and tuna cytochrome *c* it is implicitly assumed that these proteins are very similar in structure. The strong similarities among the helix probability profiles for the cytochromes *c* from various species strengthen this assumption (Lewis and Scheraga, 1971). The accompanying paper further demonstrates that this is a reasonable assumption, based on conformational energy calculations concerning the influence of various naturally occurring amino acid substitutions on the reference structure of horse-heart cytochrome *c* (Warme, 1975).

Acknowledgments

We are indebted to Dr. R. E. Dickerson for supplying the preliminary unrefined X-ray coordinates of horse-heart ferricytochrome *c* (as of July 16, 1971) and for sending us a preprint of his review article for *Enzymes*.

Supplementary Material Available

A complete listing of the computed coordinates for cytochrome *c* after the third stage of refinement will appear following these pages in the microfilm edition of this volume of the journal. Photocopies of the supplementary material from this paper only or microfiche (105 × 148 mm, 24× reduction, negatives) containing all of the supplementary ma-

terial for the papers in this issue may be obtained from the Journals Department, American Chemical Society, 1155 16th St., N.W., Washington, D. C. 20036. Remit check or money order for $4.50 for photocopy or $2.50 for microfiche, referring to code number BIO-75-3509.

References

Dickerson, R. E. (1974), *Ann. N.Y. Acad. Sci. 227,* 599.

Dickerson, R. E., Takano, T., Eisenberg, D., Kallai, O. B., Samson, L., Cooper, A., and Margoliash, E. (1971), *J. Biol. Chem. 246,* 1511.

Dickerson, R. E., Takano, T., Kallai, O. B., and Samson, L. (1972), in Structure and Function of Oxidation Reduction Enzymes (Wenner-Gren Symposium, 1970), Akeson, A., and Ehrenberg, E., Eds., Oxford, Pergamon Press, p 69.

Dickerson, R. E., and Timkovich, R. (1975), *Enzymes, 3rd Ed.* (in press).

Koenig, D. F. (1965), *Acta Crystallogr. 18,* 663.

Lederer, F. (1972), *Eur. J. Biochem. 31,* 144.

Lewis, P. N., and Scheraga, H. A. (1971), *Arch. Biochem. Biophys. 144,* 576.

Margoliash, E. (1972), *Harvey Lect. 66,* 177.

Margoliash, E., and Schejter, A. (1966), *Adv. Protein Chem. 21,* 113.

Momany, F. A., McGuire, R. F., Burgess, A. W., and Scheraga, H. A. (1975), *J. Phys. Chem.,* submitted for publication.

Narita, K., and Titani, K. (1968), *J. Biochem. 63,* 226.

Pettigrew, G. W. (1972), *FEBS Lett. 22,* 64.

Ponnuswamy, P. K., McGuire, R. F., and Scheraga, H. A. (1973), *Int. J. Peptide Protein Res. 5,* 73.

Rasse, D., Warme, P. K., and Scheraga, H. A. (1974), *Proc. Natl. Acad. Sci. U.S.A. 71,* 3736.

Salemme, F. R., Freer, S. T., Xuong, N. H., Alden, R. A., and Kraut, J. (1973a), *J. Biol. Chem. 248,* 3910.

Salemme, F. R., Kraut, J., and Kamen, M. D. (1973b), *J. Biol. Chem. 248,* 7701.

Takano, T., Kallai, O. B., Swanson, R., and Dickerson, R. E. (1973), *J. Biol. Chem. 248,* 5234.

Timkovich, R., and Dickerson, R. E. (1973), *J. Mol. Biol. 79,* 39.

Warme, P. K. (1975), *Biochemistry,* following paper in this issue.

Warme, P. K., Go, N., and Scheraga, H. A. (1972), *J. Comput. Phys. 9,* 303.

Warme, P. K., Momany, F. A., Rumball, S. S., Tuttle, R. W., and Scheraga, H. A. (1974), *Biochemistry 13,* 768.

Warme, P. K., and Scheraga, H. A. (1973), *J. Comput. Phys. 12,* 49.

Warme, P. K., and Scheraga, H. A. (1974), *Biochemistry 13,* 757.

Zerner, M., and Gouterman, M. (1966), *Theor. Chim. Acta 4,* 44.