Proc. Nat. Acad. Sci. USA
Vol. 72, No. 4, pp. 1221–1225, April 1975

Assessment of Some Problems Associated with Prediction of the Three-Dimensional Structure of a Protein from Its Amino-Acid Sequence

(conformational states/three-dimensional prediction algorithm/energy minimization)

ANTONY W. BURGESS AND HAROLD A. SCHERAGA*

Department of Chemistry, Cornell University, Ithaca, New York 14853; and Department of Biophysics, Weizmann Institute of Science, Rehovoth, Israel

Contributed by Harold A. Scheraga, January 8, 1975

ABSTRACT **It is shown that most present empirical prediction algorithms provide information about the conformational states of individual residues, but give little information about the three-dimensional structure of a protein. It is necessary to predict the conformational state of every residue before the resulting structure can serve as a starting conformation to compute the native structure. It is also shown that even a perfect five-state algorithm (which does not include long-range interactions from disulfide loop closing or solvation) will not lead to a globular structure resembling the native one. However, starting from the results of a *perfect* prediction algorithm, it appears that conformational energy minimization (with long-range interactions included) can lead to a structure having the general features of the native protein.**

In this report, an attempt is made to delineate the *problems* involved in the prediction of the three-dimensional structure of a protein from its amino-acid sequence. Our basic approach is to try to circumvent the multiple-minimum problem (1) by using empirical prediction algorithms to obtain a proper starting conformation (topographical structure) from which energy minimization should lead to the native structure (1). We present here an assessment of several empirical algorithms for the prediction of the conformations of individual amino-acid residues (2–5), and show that, even if they were perfect, they would not lead to the native structure without the introduction of additional information. We have also tested the utility of an ideal (perfect) prediction algorithm, by demonstrating, that it can provide a useful starting conformation, from which it might be possible to attain the native structure by using conformational energy calculations. As an example, these procedures are applied to bovine pancreatic trypsin inhibitor (BPTI) (6).

METHODS

Topographical Structure† *for BPTI.* From the x-ray coordinates for the atomic positions determined at 1.5 Å for BPTI (R. Huber, private communication), the corresponding ϕ, ψ, and χ dihedral angles were calculated for all 58 residues. These values of ϕ and ψ were used to assign the conformational state of each residue (2), i.e., each residue was assigned to one of five conformational states: α_R, α_L, ζ_R, ζ_L, ϵ. The distribution of conformational states for each amino-acid residue in eight proteins was then used to assign average values of ϕ, ψ for each residue (Table III of ref. 2). For example, a phenylalanine residue in an extended (ϵ) conformational state is assigned ϕ and ψ values of −98° and 133°, respectively. But a threonine residue in the extended (ϵ) state would be assigned ϕ, ψ values of −110°, 152°, respectively. There is no method for predicting the conformations of amino-acid side chains at the present time; therefore, each χ value in the topographical structure was assigned the experimental value from the x-ray structure, but was rounded off to the nearest 5°. All peptide bonds were held fixed in the planar *trans* conformation. Bond angles and bond lengths for all amino-acid residues were taken from a recent compilation‡. The average values of ϕ, ψ and the molecular geometry for each residue were used to generate the positions of all the atoms in the BPTI polypeptide chain (i.e., its topographical structure). A FORTRAN IV program based on an empirical conformational energy program for polypeptides (ECEPP), but modified to omit the nonpolar hydrogen atoms§, was used to calculate these atomic positions (see forthcoming paper‡ for procedure to obtain this program).

Three-Dimensional Folding Algorithms. Three algorithms based on the optimization of *dihedral angles* in the topographical structure of BPTI were tested in an attempt to fold the topographical structure into a three-dimensional conformation which resembles the native one. (*i*) In the simplest algorithm the values of ϕ and ψ were adjusted to minimize the function F:

$$F = \sum_{i=1}^{i=3} [(r_{C^{\alpha}_i} - \langle r_{C^{\alpha}_0}\rangle)^2 + (r_{C^{\beta}_i} - \langle r_{C^{\beta}_0}\rangle)^2] \qquad [1]$$

where $r_{C^{\alpha}_i}$ is the distance between the two C^{α} atoms of the half-cystine residues forming the *i*th disulfide bond, $\langle r_{C^{\alpha}_0}\rangle$ = 6 Å, the average $r_{C^{\alpha}}$ distance in BPTI. $r_{C^{\beta}_i}$ is the corresponding distance between C^{β} atoms of the residues involved in the *i*th disulfide bond, and $\langle r_{C^{\beta}_0}\rangle$ = 4 Å. The minimum value of F with respect to ϕ, ψ of residues 5–55 was found using the method of Powell (7). (*ii*) Algorithm (*i*) was modified by introducing two constraints such that:

$$G = F + A \sum_{i=1}^{i=57} \sum_{l=1}^{l=M_i} \sum_{j=i+1}^{j=58} \sum_{k=1}^{k=M_j} r_{il,jk}{}^{-12} + B \sum_{i=1}^{i=N} (\theta_i - \theta_{i,0})^2 \qquad [2]$$

Abbreviation: BPTI, bovine pancreatic trypsin inhibitor.

[*] To whom requests for reprints should be addressed, at Cornell University.

[†] The meaning of "topographical structure" is explained in ref. 2.

[‡] F. A. Momany, R. F. McGuire, A. W. Burgess, and H. A. Scheraga, manuscript submitted.

[§] A. W. Burgess and H. A. Scheraga, manuscript to be submitted.

TABLE 1. *Comparison of predicted and experimental conformational states[a] for BPTI*

Residue no.	Type[b]	Prediction				Exp.	Residue no.	Type[b]	Prediction				Exp.	Residue no.	Type[b]	Prediction				Exp.
		A[c]	*B*[d]	*C*[e]	*D*[f]	*E*[g]			*A*[c]	*B*[d]	*C*[e]	*D*[f]	*E*[g]			*A*[c]	*B*[d]	*C*[e]	*D*[f]	*E*[g]
1	R	*c*	α_R	*c*	*c*	ϵ	21	Y	ϵ	ϵ	ϵ	ϵ	ϵ	41	K	*b*	α_R	*c*	*c*	ϵ
2	P	*c*	ϵ	α_R	*c*	ϵ	22	F	ϵ	α_R	ϵ	ϵ	ϵ	42	R	*b*	α_R	*b*	*c*	α_R
3	D	*c*	α_R	α_R	*c*	α_R	23	Y	ϵ	ϵ	ϵ	ϵ	ϵ	43	N	*c*	ζ_R	*b*	*c*	ζ_R
4	F	*c*	α_R	α_R	ϵ	ζ_R	24	N	ϵ	α_R	*c*	*c*	ϵ	44	N	*b*	ζ_R	*c*	*c*	ϵ
5	C	*c*	ϵ	α_R	ϵ	α_R	25	A	*c*	α_R	*c*	*c*	α_R	45	F	*b*	α_R	α_R	*c*	ϵ
6	L	*c*	α_R	α_R	ϵ	ζ_R	26	K	*c*	α_R	*c*	*c*	α_R	46	K	*c*	α_R	α_R	*c*	α_R
7	E	*c*	α_R	α_R	*c*	ϵ	27	A	*c*	α_R	ϵ	*c*	α_R	47	S	*c*	ϵ	α_R	*c*	ϵ
8	P	*b*	ϵ	*c*	*c*	ϵ	28	G	ϵ	ζ_L	ϵ	*c*	α_L	48	A	*c*	α_R	α_R	*c*	α_R
9	P	*b*	ϵ	*b*	*c*	ϵ	29	L	ϵ	α_R	ϵ	*c*	ϵ	49	E	*c*	α_R	α_R	α_R	α_R
10	Y	*b*	ϵ	*b*	*c*	ϵ	30	C	ϵ	ϵ	ϵ	*c*	ϵ	50	D	*c*	α_R	α_R	α_R	α_R
11	T	*b*	ϵ	*b*	*c*	α_R	31	Q	ϵ	ϵ	ϵ	ϵ	ϵ	51	C	ϵ	ϵ	α_R	α_R	α_R
12	G	ϵ	ζ_L	*b*	*c*	ϵ	32	T	ϵ	ϵ	ϵ	ϵ	ϵ	52	M	ϵ	α_R	α_R	α_R	α_R
13	P	ϵ	ϵ	*b*	*c*	ζ_R	33	F	ϵ	α_R	ϵ	ϵ	ϵ	53	R	ϵ	α_R	α_R	α_R	α_R
14	C	ϵ	ϵ	*b*	ϵ	ϵ	34	V	ϵ	ϵ	ϵ	ϵ	ϵ	54	T	ϵ	ϵ	α_R	α_R	α_R
15	K	ϵ	α_R	*c*	ϵ	ζ_R	35	Y	ϵ	ϵ	ϵ	ϵ	ϵ	55	C	*c*	ϵ	*c*	α_R	ζ_R
16	A	*c*	α_R	ϵ	ϵ	ϵ	36	G	*b*	ζ_L	*b*	*c*	α_R	56	G	*c*	ζ_L	*c*	*c*	α_R
17	R	ϵ	α_R	ϵ	ϵ	ζ_R	37	G	*b*	ζ_L	*b*	*c*	ζ_R	57	G	*c*	ζ_L	*c*	*c*	ϵ
18	I	ϵ	α_R	ϵ	ϵ	ϵ	38	C	*c*	ϵ	*c*	*c*	ϵ	58	A	*c*	α_R	*c*	*c*	ϵ
19	I	ϵ	α_R	ϵ	ϵ	ϵ	39	R	*c*	α_R	*c*	*c*	α_L							
20	R	ϵ	α_R	ϵ	ϵ	ϵ	40	A	*c*	α_R	*c*	*c*	ϵ							

c = no conformational state assigned; *b* = residue participates in a bend, but no conformational state assigned.

[a] α_R, α_L, ζ_R, ζ_L, and ϵ conformational states are defined in ref. 2; [b] the IUPAC–IUB one-letter abbreviations (9); [c] ref. 2; [d] ref. 5; [e] ref. 3; [f] ref. 4; [g] computed from data of R. Huber (private communication).

where M_i is the number of atoms in residue i, $r_{il,jk}$ is the distance between atoms l and k in residues i and j, respectively, θ_i is the value of each variable dihedral angle ϕ or ψ, and $\theta_{i,0}$ is the value of that dihedral angle in the topographical structure. A minimum in G was found with respect to ϕ, ψ of residues 1–58 using the method of Powell (7), and the results were insensitive to the values of A and B in the range of 1 to 100. (*iii*) The final algorithm considered here used empirical conformational energy calculations, developed for polypeptides‡ and modified to approximate the nonpolar carbon atoms and their attached hydrogen atoms as a single group.§ The conformational energy of the topographical BPTI structure was minimized with respect to ϕ, ψ, ω, and χ's of each residue. The energy function included contributions from nonbonding, electrostatic, and hydrogen-bonding interactions, and intrinsic torsional potentials for rotation around each bond. The united atom approximation used for these calculations simulates the presence of the nonpolar hydrogen atoms by using group parameters for the methyl and methylene groups, and Fourier series to represent the interactions across each bond§. A specific loop closing potential is included‡ to direct the formation of the correct disulfide bonds, and to optimize the interatomic interactions to give a stable conformation. The dihedral angles were optimized in two different ways: (*A*) Using *non-overlapping* nine-residue segments of the polypeptide, all of the ϕ, ψ, ω, χ^1, and χ^2 values in the segment were allowed to vary during the minimization of the conformational energy of the *whole* topographical structure, i.e., the conformational energy was evaluated within the segment, between the segment and the rest of the polypeptide, and between the two parts of the polypeptide chain on either side of the segment. After a given segment had reached a local minimum on the conformational energy surface, the next frame of nine residues was treated in the same way. In all cases, the subsequent minimization proceeded by using the optimized dihedral angles of previous segments. Segments were chosen so that all dihedral angles of the whole molecule were optimized. This procedure was then repeated on the optimized BPTI polypeptide. (*B*) Using *overlapping* segments of nine residues, the conformational energy of only the segment was minimized (1). The values of ϕ, ψ, ω, and all χ's of only the central residue were optimized (8) for each segment, with the conformations of the other eight residues maintained fixed in their topographical conformations. When the local minimum for a particular ninemer was reached, the next segment of nine residues in the polypeptide chain [shifted toward the C-terminus by one residue (1)] was considered. Again the dihedral angles of the central residue were optimized, but the preceding residues were fixed in the conformations found by energy minimization, and the succeeding residues were kept in their topographical conformations (1).

RESULTS AND DISCUSSION

Empirical Prediction Algorithms. The results of algorithms that attempt to predict extended structures, α-helices, and bends (2–4), and one that was used to predict the conformational state of every residue in BPTI (5) are shown in Table 1. Although algorithms *A* and *C* predict the positions of bends in the polypeptide chain, these bends can be formed by many different combinations of conformational states, and a clear choice for the conformational states is not available from the predictions. In algorithms *A*, *C*, and *D*, approximately 50% of the residues are not assigned a particular conformational state, so that another technique would be needed before a topographical structure could even be generated. Algorithm *B* provides a strong correlation between the predicted and observed conformational states, but assigns the conformational states of almost half of the residues incorrectly. Thus, the prediction algorithms currently available make a considerable number of errors, and many of the residues are not even assigned a conformational state. The former is a more serious limitation than the latter, because unassigned conformational states can be treated by alternative procedures. Even if the results reported for these prediction algorithms (2–5) can be repeated (which is not easy in the case of algorithms *C* and *D*), the amount of useful information for predicting the three-dimensional structure of a protein from its sequence is small.

It is interesting to consider what the BPTI molecule would look like if an empirical prediction algorithm (which assigns one of five possible conformational states to each residue) yielded *perfect* results. Such a topography would be represented by the data in column E of Table 1. When the mean values of ϕ and ψ for each residue in those conformational states (Table III of ref. 2) are used, together with the standard geometry for each residue, the C^{α} atoms of the topographical structure for BPTI can be generated (Fig. 1A). This should be compared to the experimental structure (Fig. 1C). Some of the C^{α} atoms of the topographical structure for BPTI are separated by more than 50 Å (compare the experimental structure where the maximum separation between two C^{α}'s is approximately 15 Å). Before *any* use can be made of the predicted structure, at the very least the disulfide bond lengths and bond angles must be adjusted to reasonable values[‡]; however, it is shown in the next section that a more detailed algorithm is needed before a three-dimensional structure that is likely to resemble the native conformation can be reached. It should be noted that mistakes in a single ϕ or ψ dihedral angle can lead to quite different topographical conformations.

It is in the context of the above remarks that the current status of prediction algorithms must be considered. At best, there is a reasonable correlation between the predicted and experimental positions of α-helices and extended structures. Although some prediction algorithms claim to be easy to apply (3), despite obvious ambiguities in the rules for prediction of conformations, the amount of reliable information available from these algorithms (2–4) is limited. Certainly, there is no reason to expect that these algorithms can yield reliable information relevant to the three-dimensional *globular* structure of a protein. Speculative reports, which claim that prediction algorithms can lead to information about "protein binding sites to membranes, nucleic acids, and so on" (10), or even that present prediction algorithms "will be of assistance to all those interested in studying the correlation between protein conformation and biological activity" (3), serve only to misrepresent the power of these algorithms. The biological properties of proteins depend on their unique arrangements of atoms at a molecular level, and even small (about 2 Å) disturbances in the relationships between parts of the molecule can destroy the activity of a protein. The gross features sought by many prediction algorithms (e.g., refs. 2–4) can only hope to make the analysis of the conformations available to a protein a little easier. However, to be useful for this purpose, their power and accuracy must be improved considerably, e.g., by a statistical mechanical treatment of a multi-state model which incorporates the longer-range interactions required to produce globularity.

TABLE 2. *Conformational energy minimization of central residue in ninemer segments from BPTI*

Sequence numbers of segment	Central residue		Conformation (degrees)			
	Number	Type	Exp.		Lowest energy	
			ϕ	ψ	ϕ	ψ
4–12	8	Pro	−68	157	−68	140
5–13	9	Pro	−63	145	−63	−26
8–16	12	Gly	94	−180	97	−140
9–17	13	Pro	−89	−9	−89	82
19–27	23	Tyr	−81	129	−79	126
20–28	24	Asn	−108	103	−132	116
22–30	26	Lys	−66	−34	−112	−39
23–31	27	Ala	−95	−21	−92	−30
24–32	28	Gly	83	14	86	11
30–38	34	Val	−95	119	−83	108
32–40	36	Gly	−68	−14	−35	97
33–41	37	Gly	106	−3	122	−47
41–49	45	Phe	−126	159	−137	−150

Medium-Range Interactions. Previously, conformational energy calculations on lysozyme (8) indicated that the conformational state of each residue in a protein chain is determined not only by interatomic interactions within the immediate vicinity of the residue, but also by interactions with up to four residues on either side. Similar calculations as those in ref. 8 were performed here for nine-residue segments of the BPTI molecule except that: (*i*) a united atom approximation was used to calculate the conformational energy of each peptide[§], (*ii*) up to 25 different starting conformations per central residue were subjected to energy minimization; the starting conformations included all of the likely ϕ, ψ energy minima (11) and the observed conformation of the residue, and several combinations of low energy side-chain conformations (11) at each value of ϕ and ψ, (*iii*) the x-ray bond lengths and bond angles of BPTI were used for the conformational energy calculations. In all cases, the conformational state corresponding to the experimental structure had either the lowest or the next to lowest conformational energy. In Table 2 the ϕ, ψ values for the lowest energy conformations are compared to their corresponding experimental values. The correlation between the conformation of the calculated local energy minimum for these thirteen residues (Table 2) and the observed conformation in BPTI is quite good, although there is a marked discrepancy between the calculated and observed conformations for two of the three proline residues. However, the agreement between nine of the calculated and observed ϕ, ψ pairs is good evidence that the local and medium-range interactions dominate the formation of the conformational state for a given residue. However, such a correlation was not found for the calculated and observed conformations of the side chains of these residues. It appears that short- and medium-range interactions may determine the conformational state of the backbone dihedral angles, but that the dihedral angles for the side chains must be influenced by long-range forces (i.e., interactions with residues separated by more than four in the linear sequence). Since different side-chain conformations usually differ very little in energy, long-range packing arrangements are likely to be the determining factors for the native conformations of the side chains; hence, it is probable that an empirical correlation between a particular ϕ, ψ pair for a given residue and χ values corresponding to this state will be difficult to achieve.

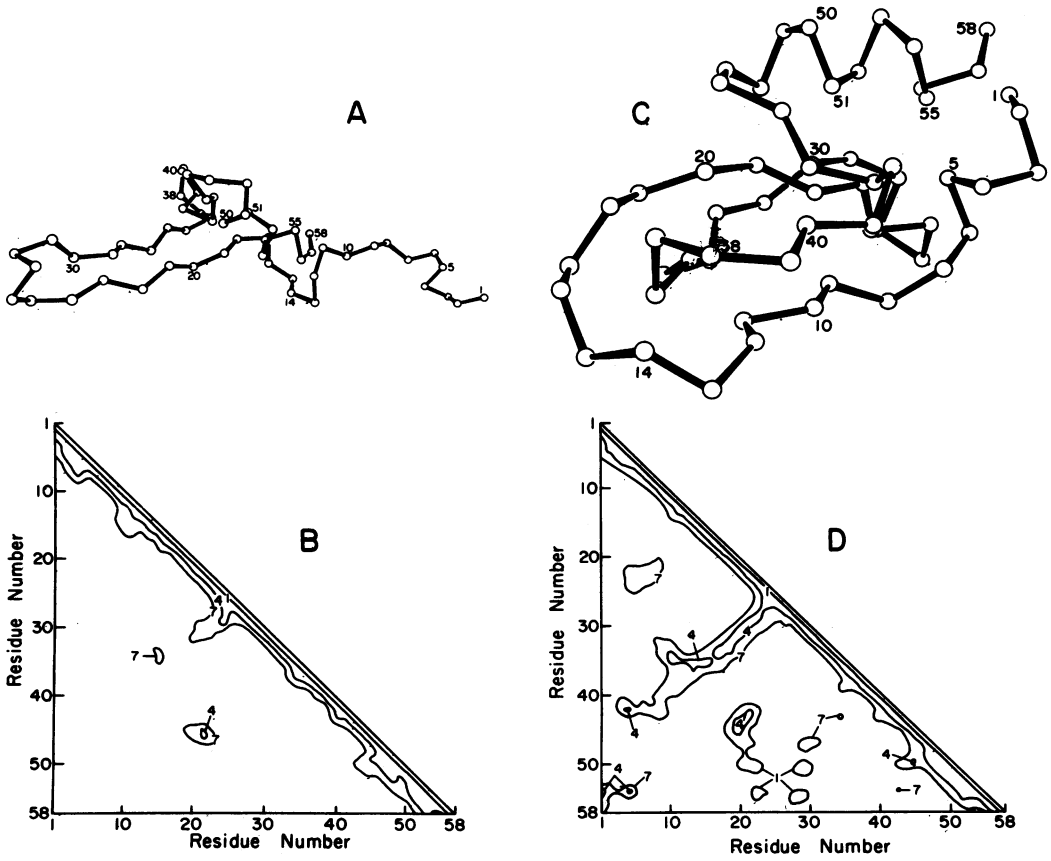

FIG. 1. (A) Topographical structure (C_α atoms) of BPTI. (B) Distance contours (12), in Å, of structure in (A). (C) X-ray structure (C^α atoms) of BPTI (R. Huber, private communication). (D) Distance contours (12), in Å, of structure in (C).

Optimization of Dihedral Angles. All of the "folding" studies started from the topographical BPTI structure shown in Fig. 1A. A contour plot of the $C^\alpha{}_i \cdots C^\alpha{}_j$ distances (12) for this conformation is shown in Fig. 1B. C^α diagrams were found to be particularly useful for monitoring the overall conformation of the protein during these conformational energy calculations. The *observed* backbone conformation of BPTI is represented by its C^α atoms in Fig. 1C, and the $C^\alpha{}_i \cdots C^\alpha{}_j$ distance plot corresponding to this structure in Fig. 1D. Although the *local* conformations of these chains are similar (i.e., the contours close to the diagonal in Fig. 1B and D are similar), the long-range correlation between residues (represented by the contours off the diagonal) is nonexistent. In Fig. 1D, the antiparallel β structure is represented by the contour lines running perpendicular to the diagonal; the short stretch of α-helix towards the C-terminus is indicated by the flaring of the contours away from, but running parallel to, the diagonal. The close contacts between those parts of the chain linked by disulfide bonds are indicated by the 4 Å distance contours in the regions near residues 5 and 55, 14 and 38, 30 and 51, and the proximity of the N- and C-termini by the 4 Å contour in the lower left-hand corner.

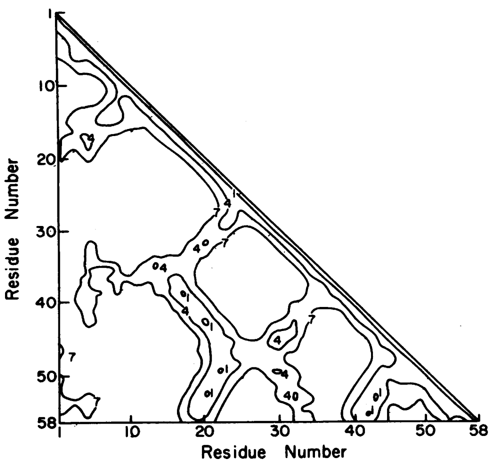

FIG. 2. Distance contours (12), in Å, of C^α atoms, obtained from procedure (*i*).

(*i*) When the disulfide bonds of BPTI are closed using Eq. 1, without consideration of atomic overlaps, the disulfide

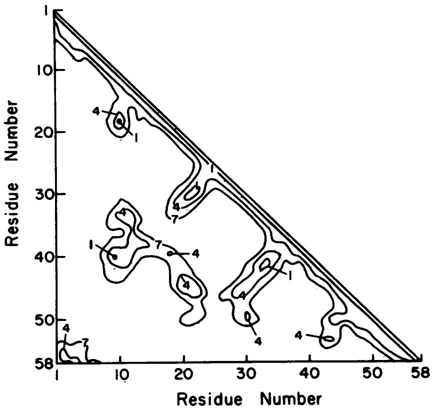

FIG. 3. Distance contours (12), in Å, of C^α atoms, obtained from procedure (*ii*).

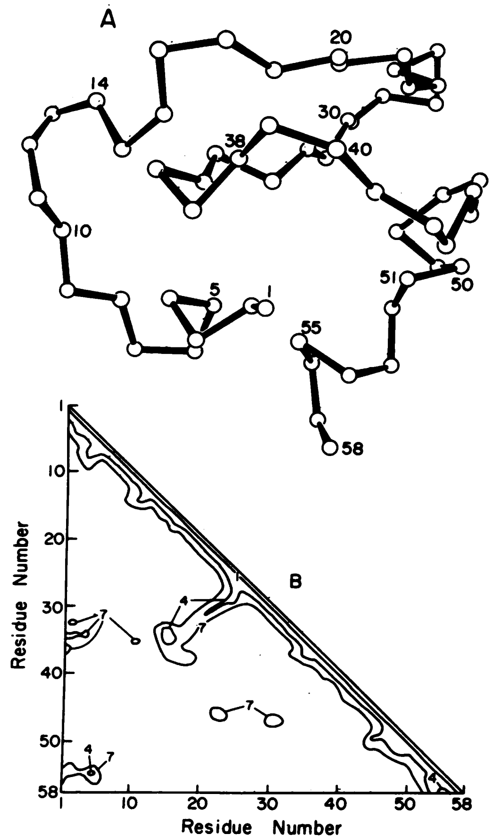

FIG. 4. (A) Structure of BPTI (C^{α} atoms) computed by procedure *iii*A. (B) Distance contours (12), in Å, of structure in (A).

bonds can be formed easily. The resulting structure, in which all of the disulfide bonds are formed, gives the distance plot shown in Fig. 2. Although the C^{α}'s for the half-cystines are separated by the expected distances, other parts of the chain obviously intersect (note the 1 Å contour lines off the diagonal). The conformational energy of this structure is extremely high, and there is no possibility of relieving all of the bad atomic overlaps by energy minimization. (*ii*) If some attempt is made to avoid atomic overlaps at the same time as the disulfide bonds are closed (using Eq. 2), a structure with all disulfide bonds formed, but with no intersecting chains, can be generated (Fig. 3). However, in the absence of hydrogen bonding and electrostatic forces, the polypeptide fails to form the correct juxtaposition of chains so that the antiparallel β structure present in the observed conformation does not form, and different orientations of other parts of the chain occur. (*iii*A) The preliminary results using the united atom conformational energy algorithm were encouraging. The calculations were not taken to completion because of the expense involved in computer time; however, the structure resulting from two applications of procedure *iii*A is shown, together with the corresponding $C^{\alpha}_i \cdots C^{\alpha}_j$ distance map, in Fig. 4 (compare to Fig. 1D). The correct antiparallel β structure has started to form, the regions involved in the disulfide crosslinks are close together, the conformation of the α-helix at the C-terminus has been preserved, the N- and C-termini are near each other, and no high-energy contacts occur between distant parts of the chain. In procedure *iii*B, each residue reaches a local minimum (in the vicinity of the starting conformation) very quickly, but the neglect of juxtaposition of residues outside the ninemer segment introduces severe atomic overlaps which cannot be removed by energy minimization of the whole molecule. Thus, it appears that, in the computations, the whole chain must be present (as in procedure *iii*A), and then some variant of procedure *iii*B might work. At present, procedure *iii*A (with an increase in the rate of convergence, or preceded by a variant of procedure *ii*) offers the hope of folding a protein from its topographical structure. For proteins without disulfide bonds (and, hence, without the help of a disulfide closing algorithm), the long-range interactions required to achieve globularity can be introduced into the energy function by solvation parameters (13) which tend to force nonpolar residues to the interior and polar ones to the exterior.

Apart from the accuracy of the potential functions used for these calculations, the main difficulties are still likely to arise from the location of a set of suitable topographical structures from which to start the calculations. It is important to develop powerful and reliable prediction algorithms for the conformational state of each residue in a polypeptide chain. Although significant progress has been made in this field, it is not likely to come from oversimplified sets of arbitrary rules, but rather from a basic understanding of the information that can be obtained from our present knowledge of protein structures, combined with a knowledge of the interactions that determine protein structures.

This work was supported by grants from the National Institutes of Health and from the National Science Foundation. A.W.B. gratefully acknowledges an Australian-American Travel Grant.

1. Scheraga, H. A. (1974) in *Current Topics in Biochemistry 1973* eds. Anfinsen, C. B. & Schechter, A. N. (Academic Press, New York), pp. 1–42.
2. Burgess, A. W., Ponnuswamy, P. K. & Scheraga, H. A. (1974) *Isr. J. Chem.* **12**, 239–286.
3. Chou, P. Y. & Fasman, G. D. (1974) *Biochemistry* **13**, 222-245.
4. Lim, V. I. (1974) *J. Mol. Biol.* **88**, 873–894.
5. Robson, B. & Pain, R. H. (1974) *Biochem. J.* **141**, 869–882.
6. Huber, R., Kukla, D., Rüllmann, A. & Steigemann, W. (1972) *Cold Spring Harbor Symp. Quant. Biol.* **36**, 141–150.
7. Powell, M. J. D. (1964) *Comput. J.* **7**, 155–162.
8. Ponnuswamy, P. K., Warme, P. K. & Scheraga, H. A. (1973) *Proc. Nat. Acad. Sci. USA* **70**, 830–833.
9. IUPAC-IUB Commission on Biochemical Nomenclature (1968) *Arch. Biochem. Biophys.* **125**, i-v.
10. Fasman, G. D. & Chou, P. Y. (1974) in *Peptides, Polypeptides and Proteins*, eds. Blout, E. R., Bovey, F. A., Goodman, M. & Lotan, N. (John Wiley, New York), pp. 114–125.
11. Lewis, P. N., Momany, F. A. & Scheraga, H. A. (1973) *Isr. J. Chem.* **11**, 121–152.
12. Ooi, T. & Nishikawa, K. (1973) in *Conformation of Biological Molecules and Polymers* (Jerusalem Symp. Quantum Chem. and Biochem.), eds. Bergmann, E. D. & Pullman, B. (Academic Press, New York), Vol. **5**, pp. 173–187.
13. Gibson, K. D. & Scheraga, H. A. (1967) *Proc. Nat. Acad. Sci. USA* **58**, 420–427.