Biophysical Chemistry, 44 (1992) 113–127
Elsevier Science Publishers B.V., Amsterdam

BIOCHE 01680

Mechanism of protein folding. IV. Forming and breaking of disulfide bonds in bovine pancreatic tripsin inhibitor

Yukio Kobayashi [1], Hiroyuki Sasabe [1], Takao Akutsu [2] and Nobuhiko Saitô [2]

[1] *Bioelectronic Materials, Frontier Research Program, RIKEN Institute, Hirosawa 2-1, Wako, Saitama 351-01 (Japan)*
[2] *Department of Applied Physics, Waseda University, Okubo 3-4-1, Shinjuku, Tokyo 169 (Japan)*

(Recieved 28 February 1992; accepted 6 April 1992)

Abstract

The folding mechanism of bovine pancreatic tripsin inhibitor (BPTI) is explained theoretically on the basis of the island model, where the driving force of folding is hydrophobic interaction. For this purpose, we take a look at the formation and breaking of disulfide bonds during the folding process of BPTI. The intermediate conformations and the native one are successfully obtained, which satisfy the so-called "lampshade" geometrical criterion for the formation of the disulfide bonds. The folding pathway is consistent with the renaturation experiment by Creighton. In addition, an elaborate treatment of side chains of amino acid residues by the software programme CHARMM confirms quantitatively the formation of disulfide bridges.

Keywords: Bovine pancreatic tripsin inhibitor; Creighton's renaturation experiment; Disulfide bond; Intermediate state; Island model

1. Introduction

In a series of papers [1–6], we elucidated the mechanism of protein folding for some proteins such as myoglobin, lysozyme and phospholipase on the basis of the "island model". The principle of protein folding discovered and developed by us can be summarized as follows: (1) The formation of α-helices and β-structures (or β-sheets) in the first stage. (2) The pairs of hydrophobic residues located closely on the amino acid sequence are bound by the long-range hydrophobic interaction, thus yielding local structures. The simulation of this process is carried out in a sequential way from nearest pairs to more remote ones, by considering, besides the hydrophobic interactions, Lennard–Jones potentials between nearby residues to prevent the collapse of molecules. (3) The local structures obtained in this way can bring other pairs of hydrophobic residues or the pairs of cysteines close to each other which, when bound, can modify the local structures. (4) The various interactions ignored in the above calculations are taken into account to obtain the final tertiary structure. The whole process described above is called the "island model". The further details will be described later. The criterion for the formation of a disulfide bond was also proposed [4]. In the present paper, we attempt to explain the mechanism of the formation and breaking of the disulfide bonds during the folding process of bovine pancreatic tripsin inhibitor (BPTI).

Correspondence to: Dr. Y. Kobayashi, Bioelectronic Materials, Frontier Research Program, RIKEN Institute, Hirosawa 2-1, Wako, Saitama 351-01 (Japan).

BPTI is a small protein consisting of 58 amino acid residues, and for that reason it has frequently been selected in various computational works [7–10] on protein conformation. The folding mechanism of BPTI, however, has not yet been fully explained owing to its remarkable

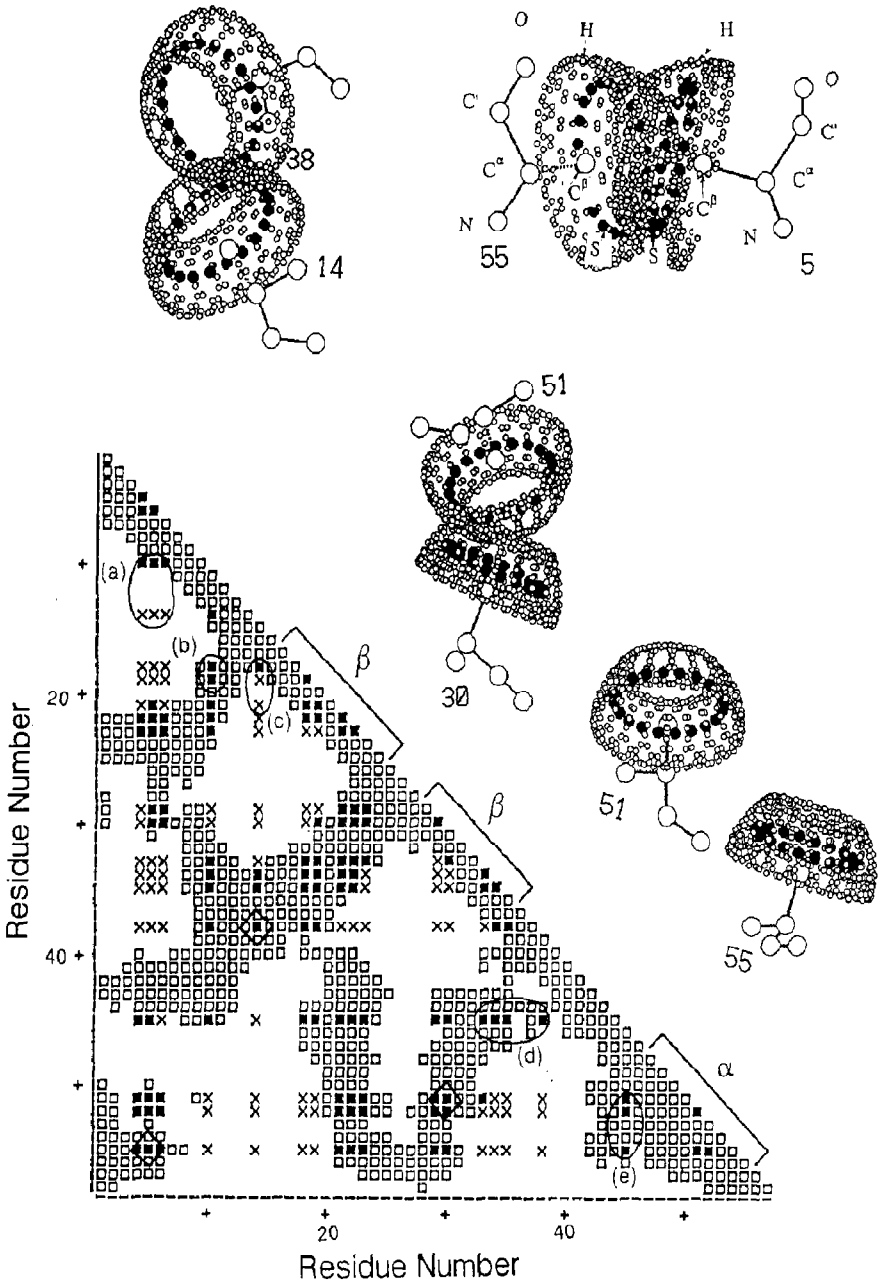

Fig. 1. The distance map for the native structure of BPTI. Open squares (□) indicate the amino acid pairs with a distance between their C^{α}'s < 13 Å, and solid squares (■) and (×) indicate the hydrophobic pairs with a distance < 13 Å and > 13 Å, respectively. Disulfide bonds are indicated by (◇). Important hydrophobic pairs (a), (b), (c), (d) and (e) are encircled. Lampshades for cysteine pairs are also shown. Cysteine pairs (30–51), (14–38) and (30–51) form disulfide bonds, but for example (51–55) does not form a disulfide bond.

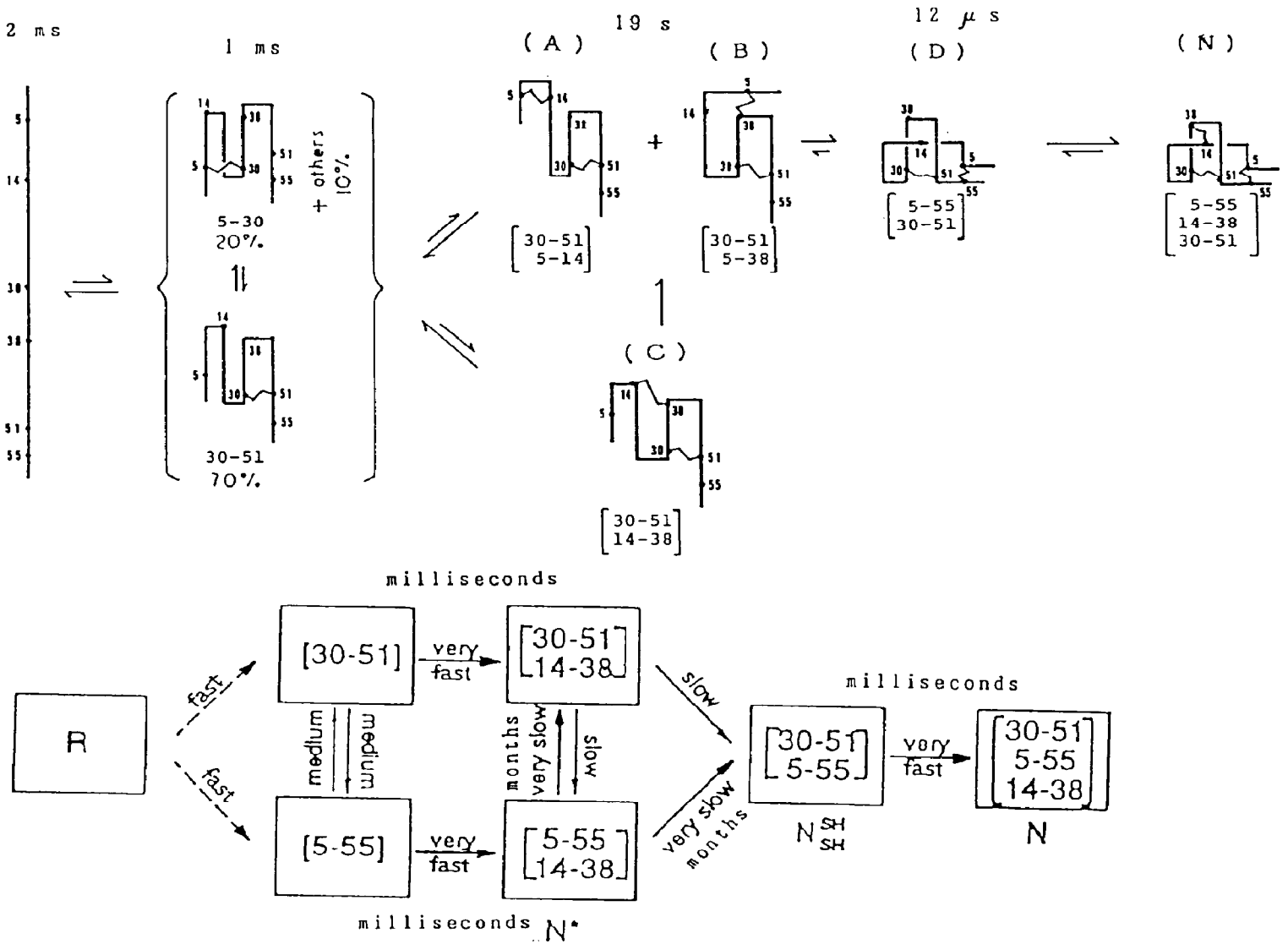

Fig. 2. Schematic diagram of the folding pathway of BPTI that accompanies disulfide bond formation and breakage. (a) The folding pathway proposed by Creighton (adapted from Fig. 7 of Ref. [13]). Intermediate and final (native) states are denoted as (A), (B), (C), (D) and (N), respectively. The half lives of the respective conformations are shown. (b) The folding pathway proposed by Weissman and Kim (adapted from Fig. 10 of Ref. [30]).

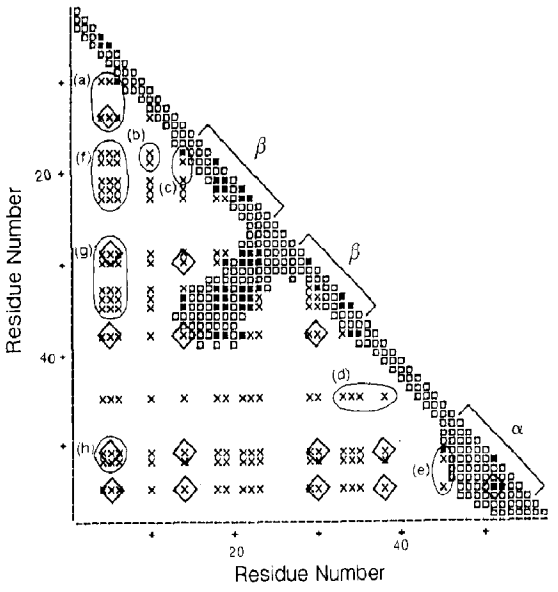

Fig. 3. Initial extended structure of BPTI with only α-helix and β-structure at their native locations. Possible pairs of cysteine residues are indicated by (◇) (not necessarily to be bonded). Important hydrophobic pairs (a), (b), (c), (d), (e), (f), (g) and (h) are encircled.

character: BPTI seems to be an exception within the island model. We take Trp, Ile, Leu, Val, Phe, Met and Cys as hydrophobic residues. Inspecting the native conformations of many proteins, we see that the pairs of hydrophobic residues which connect the neighbouring secondary structures are those at the shortest distance on the chain. For example see Ref. [2] for myoglobin and Ref. [3] for lysozyme and phospholipase. This fact enabled us to establish the folding principles mentioned above. For BPTI, however, this is not the case. Notice the hydrophobic pairs (a), (b), (c), (d) and (e) encircled in Figs. 1 and 3. We see that they contain rather favourable hydrophobic pairs (a) that nevertheless remain unbound. Therefore, when hydrophobic interactions are introduced successively in accordance with the general principles of protein folding, conformations different from the native one will be obtained. This suggests the existence of intermediates which do not passess the native disulfide bridge. A preliminary study was reported by Saitô [1].

The renaturation experiments of BPTI by Creighton [11–18] are thus expected to be explained in reasonable way by the island model. BPTI has three disulfide bonds in positions (30–51), (14–38) and (5–55), respectively. According to Creighton's experimental results, the native disulfide bonds are not formed directly. When BPTI refolds from the fully reduced conformation, it proceeds to its native conformation via one- and two-disulfide intermediates (Fig. 2(a)). The one-disulfide intermediates are the first products. About 70% of these have the disulfide bond (30–51). The two-disulfide intermediates ((A), (B), (C) and (D) in Fig. 2(a)) contain disulfide bonds, (5–14), (5–38), (14–38) or (5–55) in addition to (30–51). The (30–51) disulfide bond is present in the refolded and the native conformations, while (5–14) and (5–38) are non-native. Only the intermediate (D) with the disulfide bond (5–55) can proceed to the native conformation. The other three intermediates (A), (B) and (C) involve disulfide interchange before completing refolding, although the intermediate (C) has two native disulfide bonds (14–38) and (30–51). The experimental results indicate that there exists a definite folding pathway of BPTI.

Is the island model valid for BPTI, which has several intermediates? The aim of the present work is to give an answer to this query. We focus our attention on the disulfide bonds of the intermediate conformations and the native one. In Section 2, we will describe how to simulate the folding pathway on the basis of the island model. In Section 3, several intermediates and the final refolded conformations will be obtained with the aid of this model. At the early stage, the side chains are treated as rigid spheres, but we pay special attention to the orientations of the "lampshades" of cysteine residues, which come close to each other, to see if they can form disulfide bonds [4]. At the next stage, CHARMM [1] energy will be minimized, after the coordinates of the atoms in the side chains being generated, for carrying out the process (4) mentioned at the

[1] CHARMm is a general and flexible software application for modelling the structure and behaviour of molecular systems.

beginning of this section. The obtained conformations will be quantitatively examined on the basis of the values of geometrical parameters of disulfide bond. Section 4 is devoted to discussion, and our results will be compared with other works.

2. Method of folding

BPTI is an α- and β-type protein whose two β-strands form an anti-parallel β-sheet. The first step of the folding process is the formation of the secondary structures. In accordance with the island model, we examine the mechanism of packing of these secondary structures into the tertiary structure. To do this, we start from the conformation with native secondary structures and the extended conformation in the other region, and then search for the conformation of the minimum energy by changing the dihedral angles of the amino acid residues in the extended regions.

The driving force of packing is hydrophobic interaction, which has two aspects: the long-range nature and the specificity of pairing after formation of the secondary structures. We treat hydrophobic interaction as a simple attractive potential between hydrophobic side chains, since it is extremely difficult to obtain the free energy of the hydrophobic interaction considering the molecular arrangement of intervening water. Our problem is to search for the minimum of the conformational energy which is calculated as the sum of hydrophobic interactions, Lennard–Jones 6–12 potentials, and disulfide binding. The Lennard–Jones 6–12 potentials between non-bonded atoms (except H), of course, are taken into account in order to avoid the collapse of the molecule. The disulfide bonds are introduced in the packing process when two cysteine residues come close together and the orientations of their side chains are appropriate for binding. At the initial stage, other interactions, such as hydrogen bonding at regions other than the secondary structures, the electrostatic interactions, and torsional energies of dihedral angles are not considered. These interactions will be taken into account at the final stage.

The energies under consideration are calculated as follows. In order to reduce the number of the variables, each side chain is replaced by a rigid sphere of Van der Waals radius whose center is located at a distance of r_G from the C^α in the direction of C^β, r_G being the distance between C^α and the center of mass of the side chain [2]. The generation of real side chains will be done later as mentioned earlier. The hydrophobic interaction is of the long-range type by virtue of the presence of ice-like water layers around the hydrophobic residues and thus assumed to be:

$$u(r) = \begin{cases} u_0 & r \le r_0 \\ u_0 \exp[-0.1(r - r_0)] & r_0 \le r \le r_0 + 12 \\ 0 & r_0 + 12 \le r, \end{cases} \quad (1)$$

where r is the distance between the centers of the spheres representing the side chains and r_0 is the sum of the Van der Waals radii of the side chains [2]. We take $u_0 = -3.0$ kcal/mol for any pair of hydrophobic residues. A decay length of 10 Å in formula (1) is adopted in accordance with the experimental observations [19], which showed that the interaction between two crossed rods composed of hydrophobic molecules is an exponentially decaying function with a decay length of 10 Å. This also implies that the elementary interaction between hydrophobic molecules will be similar. The functional form of the disulfide bonding is assumed to be:

$$u(r) = \begin{cases} 20 & r < 3.2 \\ -10\left[1 - (r - 4.2)^2\right] & 3.2 \le r \le 5.0 \\ -3.6\left[1 - (r - 5.0)^2/25\right] & 5.0 < r < 10 \\ 0 & 10.0 \le r. \end{cases} \quad (2)$$

This is a potential composed of a short-range part for the S–S bond of a disulfide and another part with a long-range tail [3]. This would seem rather

[2] The units of length and energy are taken here as Å and kcal/mol, respectively.

artificial, but the fine details of the interaction are irrelevant to our approximate calculation of folding, except for the long range character and potential well depth.

Before carrying out energy calculation and search for its minimum, we need to say a few words on the initial structure. First, the initial structure consists of one α-helix and an antiparallel β-sheet, i.e. the native structure. Second, in the initial structure, we set the dihedral angles to 180° in regions these secondary structures. This is because we should start from the initial structure without steric effects in the numerical calculation. If steric hindrance is present and noticeable in the initial structure, the energy gain due to the Lennard–Jones potentials might mask that due to the hydrophobic interactions.

Energy minimization is performed as follows. The variables considered are only the dihedral angles, ϕ's and ψ's, except ϕ's of prolines. Bond lengths and bond angles are assumed to be invariant during the folding process. The various interactions are introduced successively, starting from short-distance pairs to long-distance ones in accordance with the island model for calculating energies of local structures. Next, a search for the set of dihedral angles of relevant amino acid residues with minimum energy is performed by the procedure of Bremermann [20] each time additional pairs are considered. As folding proceeds, new hydrophobic residues come close together which can undergo hydrophobic interactions resulting in a succession of many pairs of atoms making contact. So, in addition, the interactions are introduced to these new pairs and energy minimization is performed.

In the following we exclusively calculate the part of interaction energy in the change of free energy during refolding, neglecting the entropy part. While refolding proceeds, the chain entropy decreases. In the island model the folding proceeds by the interaction between short-distance pairs and thus the accompanying entropy decrease is not large, keeping the free energy as always dereasing. This also ensures multiple minima are avoided in search for the minimum free energy [3].

3. Results of folding

3.1 Folding pathway to intermediate and native conformations

In this section, we search for the folding pathway via the intermediate conformations to reach the native one. For this purpose, we consider the disulfide bonds as specific folding indicators. We consider only four species, (30–51, 5–14), (30–51, 5–38), (30–51, 14–38) and (30–51, 5–55) as intermediates (denoted as (A), (B), (C) and (D), respectively), and disregard one-disulfide intermediates ((30–51) and (5–30)), which can be formed by random disulfide bonding and can also be separated from each other experimentally. This is because disulfide bond formation and breaking of these one-disulfide intermediates occur within milliseconds, therefore they are extremely unstable in comparison with two-disulfide intermediates except for (D) [12]. We take account of (D), since the extremely short half-life of (D) indicates that the conformation of (D) is almost the same as that of the native BPTI (N).

Before entering into the folding simulation, we recapitulate the geometrical nature of a disulfide bond as discussed by Watanabe et al. [4]. In considering the conformation of C^{α}–C^{β}–S–H, it is assumed that the positons of C^{α} and C^{β} are fixed. Then the S-atom lies on a cone with the axis C^{α}–C^{β}, and the H-atom lies on another cone, where the length of S–H is taken as 1 Å, half the S–S bond length. Thus the locus of the H-atom looks like a lampshade (see Fig. 1). In the real disulfide bond, two lampshades of two cysteine residues must be in face to face contact. This additional geometrical criterion is useful for selecting two cysteines for disulfide bonding. Inspecting the 15 pairs of cysteine residues found in the native structure (Fig. 3), we see that in the proper pairs, (30–51), (14–38), and (5–55), two cysteine residues are close together while the two lampshades face each other. Thus Cys5 and Cys55 are bonded (Fig. 1), but the pair (51–55) cannot make a disulfide bond, because although Cys51 and Cys55 are close together, the two lampshades do not face each other (Fig. 1). In this way, when

two cysteine residues are in close contact, the mutual orientation of their lampshades is a key factor for selecting the proper pairs of cysteine residues for disulfide bonding. Thus, we must notice the geometrical relation for the pairs of cysteine residues, if we want to attain to an intermediate or the final tertiary structure by calculation.

The details of the search for the conformations of minimum energy are as follows. First we introduce the hydrophobic interactions between nearest hydrophobic pairs (encircled (a), (b), (c), (d)

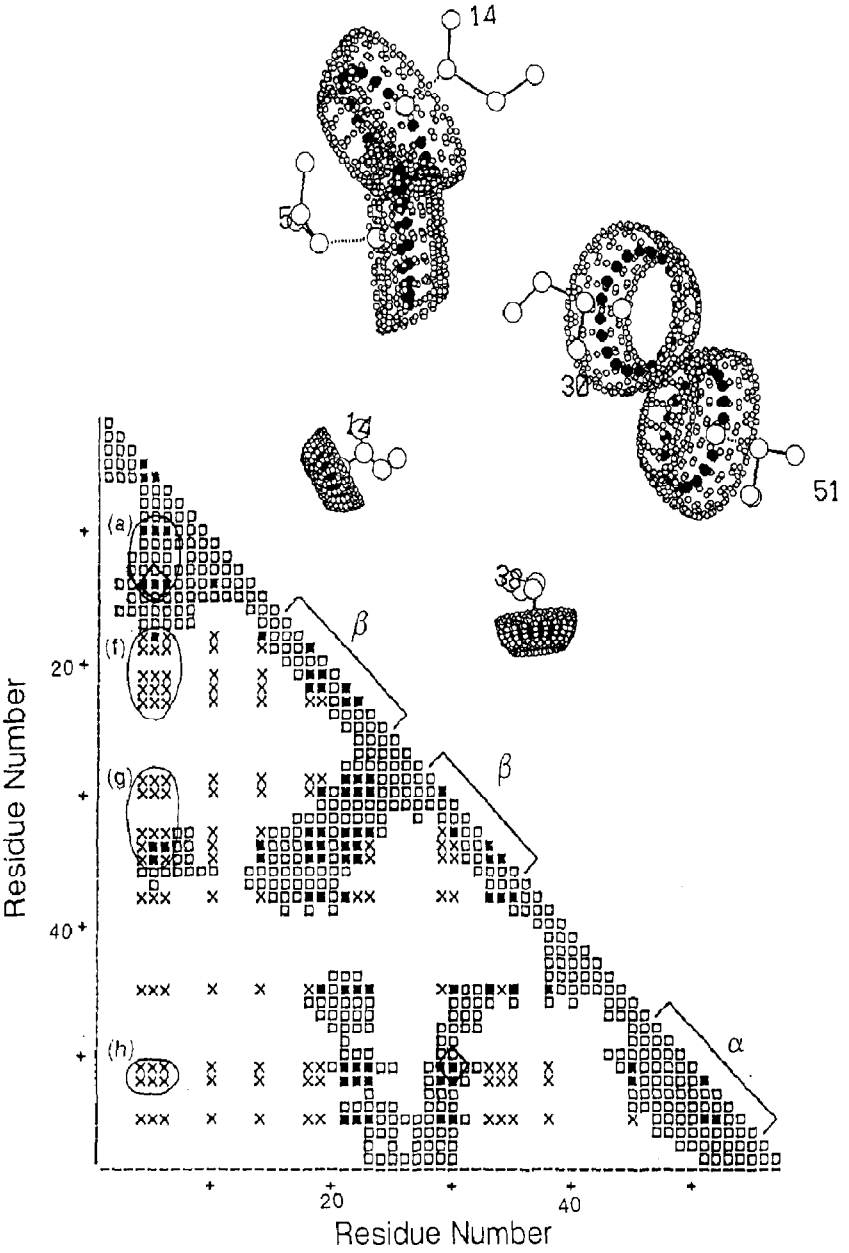

Fig. 4. Intermediate (A) of BPTI. Lampshades for (5–14), (30–51) and (14–38) are also shown. Cys14 and Cys38 do not form a disulfide bond.

and (e) in Fig. 3). The hydrophobic interactions in groups (d) and (e) easily yield the right-hand side of the structure shown in Fig. 4 which forms the disulfide bond (30–51), since Cys30 and Cys51 are close and have facing lampshades. The hydrophobic interactions in the groups (a), (b) and (c) give rise to the structure shown at the left-hand side in Fig. 4, which shown the (5–14) cysteine (facing lampshades) and the (14–38) pair (non-facing lampshades). Thus this structure, when

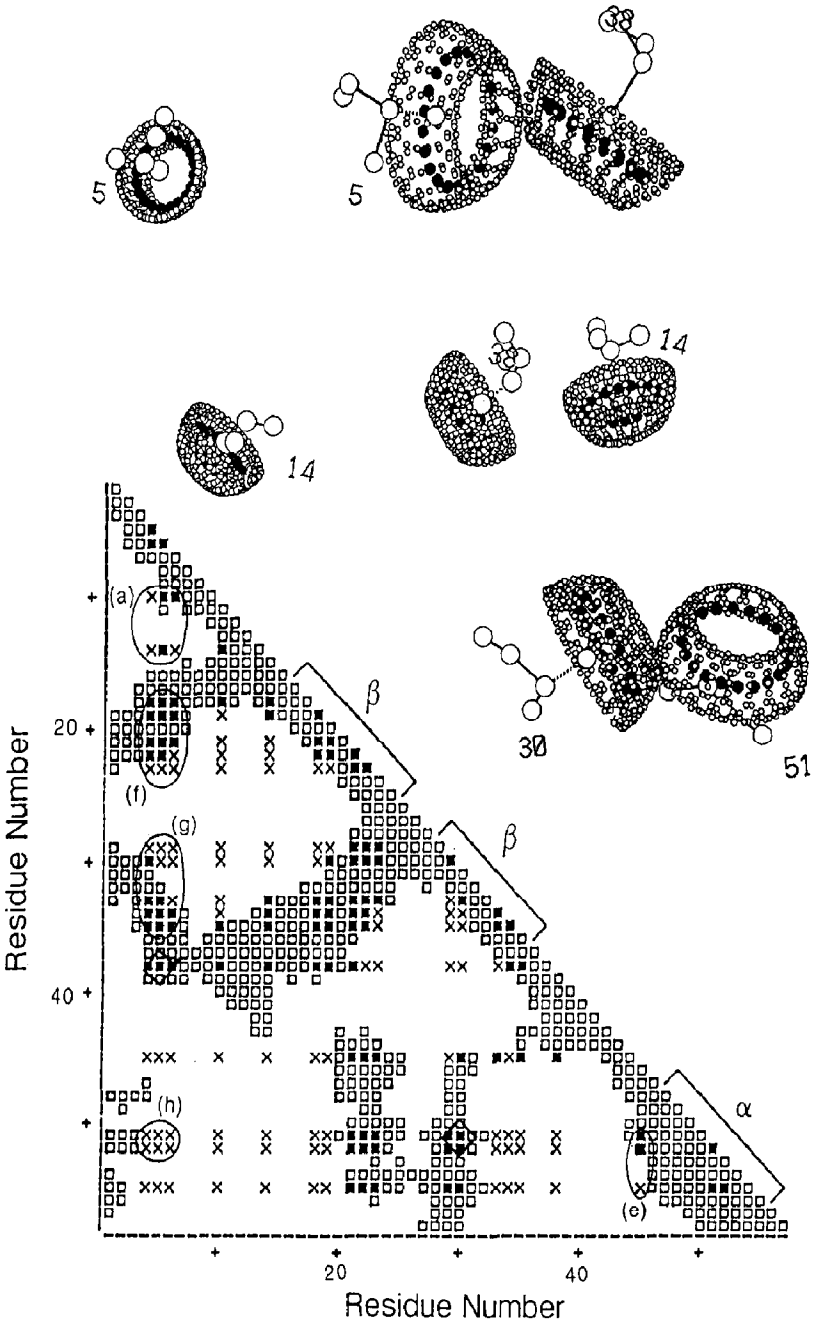

Fig. 5. Intermediate (B) of BPTI. Lampshades for (5–14), (14–38), (5–38) and (30–51) are also shown. Cys14 does not form a disulfide bond with Cys5 and Cys38.

disulfide bonding is introduced between Cys5 and Cys14, yields intermediate (A) defined in Fig. 2(a). In intermediate (A) the hydrophobic pairs in groups (b) and (c) are not fully combined by virtue of the steric hindrance and a main role is played by the pairs in group (a). However, according to the principle (3) mentioned in Section 1, we should consider weakly hydrophobic Cys14 and Cys38 as well, because they are close to each other in real space due to the assumption of the forming of the antiparallel β-sheet (Fig. 3). Furthermore there exists the possibility to make a disulfide bond (14–38). The two disulfide bonds (5–14) and (14–38) have the 14th residue in common. If either one of them is bonded, the other becomes unbonded. It may happen that neither (5–14) nor (14–38) is bonded but (5–38) can be bonded. With these possibilities in mind, we accomplished the energy minimization on the basis of the sets of dihedral angles determined from different random numbers and obtained the conformations in which Cys5 and Cys14, Cys14 and Cys38 or Cys5 and Cys38 are close to each other and the lampshades of the respective pairs face each other (Figs. 4, 6 and 5, respectively). Consequently, in the respective conformations, we introduced the disulfide bonding to the pair (5–14), (14–38) or (5–38) and the hydrophobic interactions to the relevant hydrophobic pairs, for minimizing the total energy. In this way, the intermediates, (A), (C) and (B) were obtained. Note that as one can see in Fig. 4, for intermediate (A) some of the residues in group (f) are close even though interactions in (f) are not considered. This implies that the interactions between the residues in group (f) have to be considered and when they

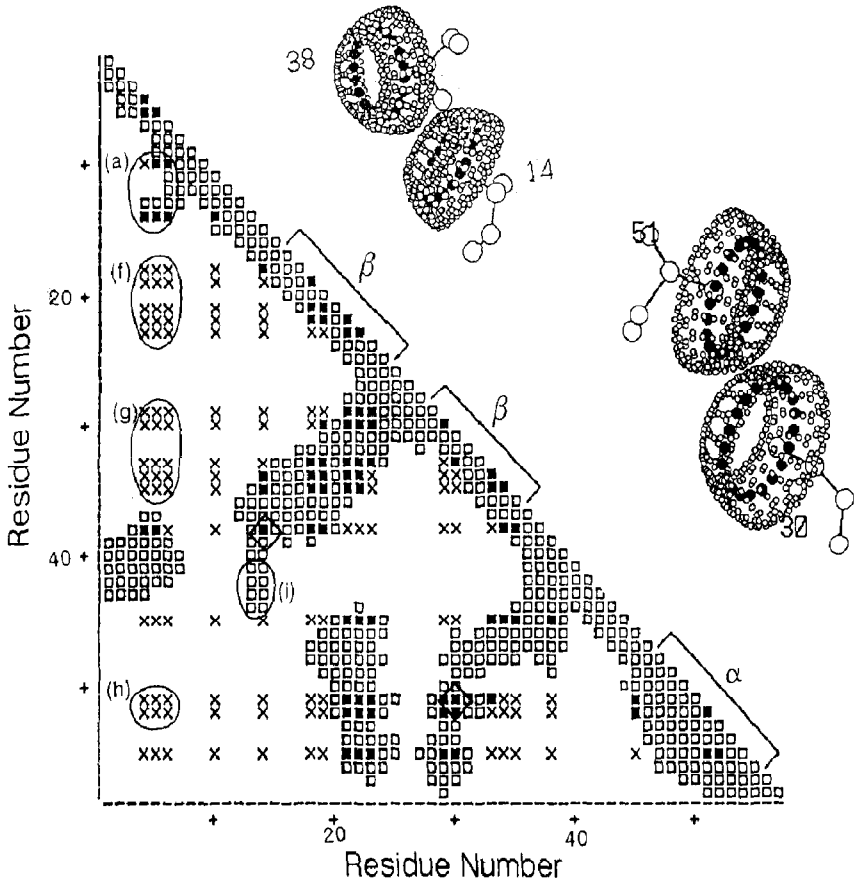

Fig. 6. Intermediate (C) of BPTI. Lampshades for (14–38) and (30–51) are also shown.

become dominant they yield conformation (B), as mentioned above in which Cys5 and Cys38 are bonded (Fig. 5). However, in this conformation (B), Cys14 forms the disulfide bond neither with Cys5 nor Cys38 (Fig. 5), although the pair of Cys5 and Cys14 as well as that of Cys14 and Cys38 lie in close range (Fig. 5), because the lampshades for both pairs are not appropriate for disulfide bonding to occur.

Above computer calculations using different random numbers indicate that we can obtain three intermediates (A), (B) and (C) depending on which of the hydrophobic groups (a), (f) and (Cys14–Cys38) plays a role in packing, respectively. We cannot select one from the others. This seems contradictory to the idea of the island model for a uniquely determined pathway, but in the case of BPTI, hydrophobic pairs located at short distance on the chain and hydrophobic pairs close to each other in real space exist simultaneously. Coexistence of these pairs gives rise to the intermediates (A), (B) and (C). The distance map of three types of the intermediates are shown in Figs. 4, 5 and 6.

Next, we proceed to search for the pathway from the intermediate (A) to the native conformation. From Fig. 4, at the right-hand side in the map, the hydrophobic pairs are bound except the pairs which are not bound due to steric effects, and a stable local structure is formed. On the other hand, at the left-hand side, a lot of remote pairs along the chain are not bound. When the

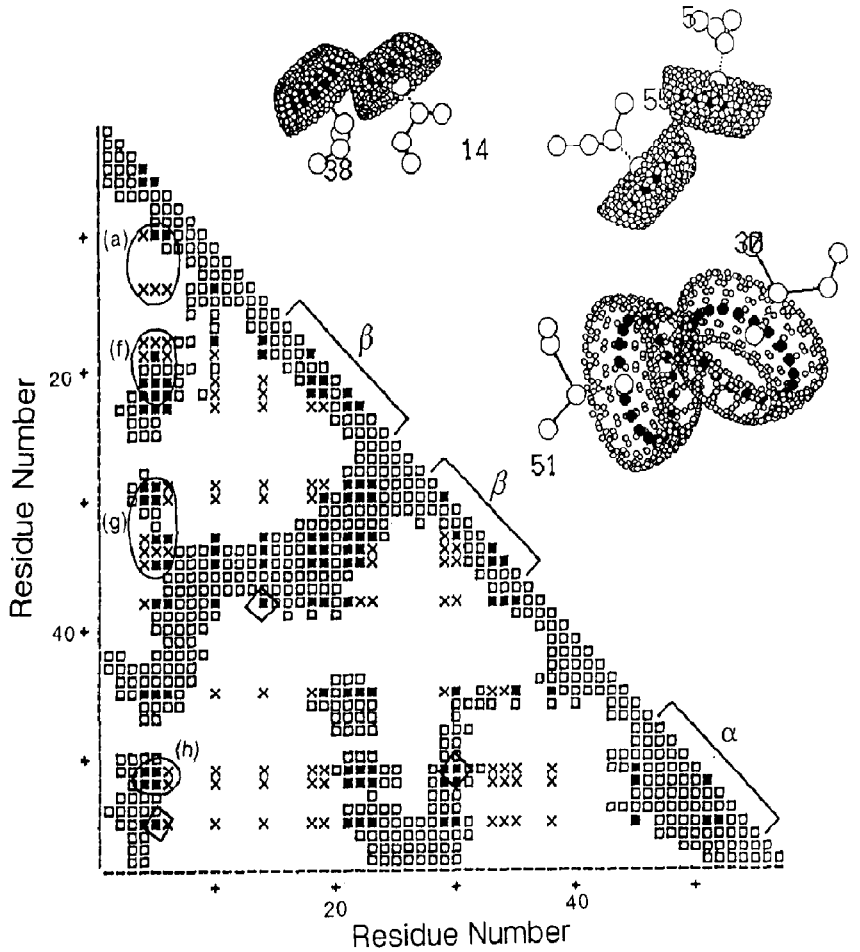

Fig. 7. Refolded structure of BPTI obtained from the intermediate (A). This corresponds to the native structure (N) (30–51, 14–38, 5–55). Disulfide bonds are indicated by (○). Lampshades for (14–38), (30–51) and (5–55) are also shown.

disulfide bond between Cys5 and Cys14 is broken at an instant due to the flexibility of the cysteine residues, the local structure at the upper part of the left-hand side may be loosen, and folding may proceed for the remote pairs (f), (g) and (h) in Fig. 3 to be bound. The hydrophobic interactions of the remote pairs can make up for the energy increment due to breaking of the disulfide bond. As a result, the energy of conformation lower by about −94 kcal/mol (*not* the free energy, see above) was obtained. In this conformation, the local structure is formed at the lower part of the left-hand side in Fig. 7 and Cys5 and Cys55 are close together. The lampshades of these cysteine residues face each other (Fig. 7). From Figs. 1 and 7, this is similar to the native structure. Starting from the intermediate (B), when the hydrophobic pairs (h) in Fig. 5 are bound and the disulfide bond (5–38) in broken, we reach the conformation (D) in Fig. 2(a) as shown in Fig. 8. This conformation can yield the disulfide bonds (5–55) in addition to (30–51). The lampshades for (5–55) is shown in Fig. 8. The folding simulation starting from the intermediate (C) could not yield a conformation similar to the native one by the same method as above. In this sense, our result is consistent with that from experiments [11–18], which is summarized in Fig. 2(a).

We point out three important features. First, Cys30 and Cys51 are always bonded. Second,

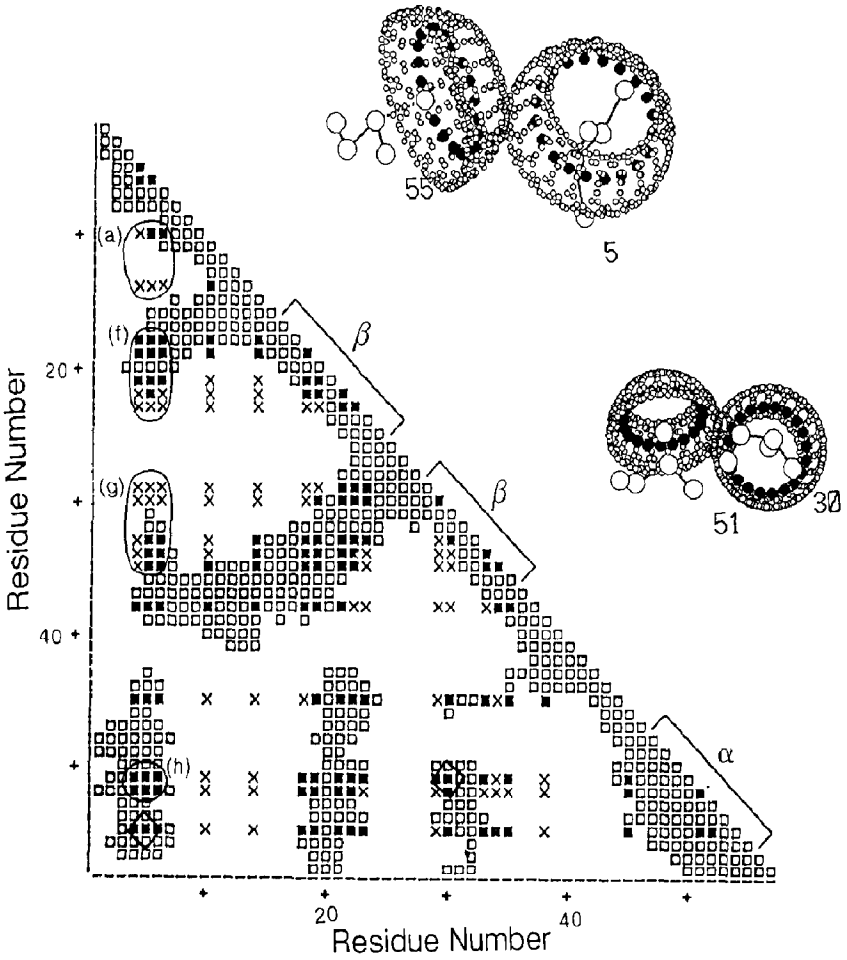

Fig. 8. Refolded structure (D) of BPTI obtained via the intermediate (B). Lampshades for (30–51) and (5–55) are also shown.

Cys5 changes its partner in the disulfide bond in the folding process from 14 to 55 ((A) → (N)) or from 38 to 55 ((B) → (D)). Third, in every conformation ((A), (B), (C), (D) and (N)), the distance between C^{α}'s of Cys51 and Cys55 is only 5.9 Å, but the geometrical relation is not appropriate for the formation of the disulfide bond. These may be explained as follows. From the distance map of the initial conformation, important hydrophobic pairs are distributed into almost three parts at the left-hand side in the map (encircled parts (a), (f)–(g) and (h) in Fig. 3) and (Cys14–Cys38). The local structures may be formed corresponding to this distribution, since the remote hydrophobic residues on the chain come close together in real space. The respective parts have, in their neighbours, the disulfide bond (5–14), (5–38), (5–55) and (14–38), which play an important role in fastening tightly the local structure. In the intermediate (B), when Cys5 and Cys38 are bonded, most of the hydrophobic pairs at the upper and lower parts ((a) and (h)) of the left-hand side in the distance map (Fig. 5) are not bound. They correspond to the nearer pairs and the remote ones along the chain, respectively. In the same way, in the intermediate (A) (Fig. 4), the hydrophobic group (a) is bound, while (f)–(g) and (h) are almost unbound. In the intermediate (D) (Fig. 8) and the final state (N) (Fig. 7), the hydrophobic group (f)–(g) and (h) are bound, while (a) is almost unbound. On the other hand, at the right-hand side in either map of Figs. 4, 5, 6, 7 or 8, an extremely stable local structure is formed by virtue of the hydrophobic interactions in each of the groups (d) and (e) in Fig. 3. This results in the formation of the stable disulfide bond between Cys30 on a pleated sheet and Cys51 on an α-helix. This mechanism given above was verified experimentally by Oas and Kim [21]. Therefore Cys51 cannot make a bond to Cys55.

The reason why the native conformation may not be attained from the intermediate (C) is explained as follows. The distance map of the conformation (C) (Fig. 6) shows that the pairs denoted as (i) prevent the formation of the disulfide bond (5–55). Thus the native structure cannot be attained. After a long time, when the disulfide bond (14–38) is loosen the interactions in the group (i) may happen to be released and yield the disulfide bond (5–55) instead (see Section 4).

3.2 Disulfide bridge formation in refined conformation

In Section 3.1, the intermediate conformations and the native one were successfully obtained by the simple method based on the island model. The disulfide bonds were considered on the basis of the lampshades. In this section, we reconsider the disulfide bridge formation quantitatively, referring to quantum chemical calculation by Aida and Nagata [22].

For this purpose, the calculated conformation (N) should be refined, after generating the atoms in the side chains by CHARMM [23], because the side chains were replaced by appropriate spheres ill Section 3.1. The Cartesian coordinates for all unspecified atoms are calculated from the data in the CHARMM internal coordinate table, since the coordinates of atoms (except C^{β}) in the side chains are not determined in Section 3.1. In addition, to treat disulfide bridges, CHARMM deletes an H-atom in the SH group and forms the disulfide bridges. Then, CHARMM empirical energy of this conformation was minimized to achieve a better set of the atom coordinates. CHARMM empirical energy function is a summation of many individual energy terms. Internal energy terms include bond energy, bond angle energy, and dihedral energy. External non-bonding energy terms include electrostatic and Van der Waals interactions. The method of minimization is as follows. One hundred steps of steepest descent minimization [24] is performed, followed by the adopted basis Newton–Raphson method of 600 steps. The value of CHARMM energy decreased from 3.50×10^8 (kcal/mol) to 2.18×10^3 (kcal/mol).

For the pairs of Cys5 and Cys55, Cys14 and Cys38, and Cys30 and Cys51, the dihedral angles $C_i^{\beta}S_iS_jC_j^{\beta}$ ($\chi^{S_iS_j}$) were constrained during energy evaluation steps. These angles were set to $-90°$, $90°$, and $-90°$, respectively. According to the calculation of Aida and Nagata (1984), the S–S bond length increases as the S–S dihedral angle

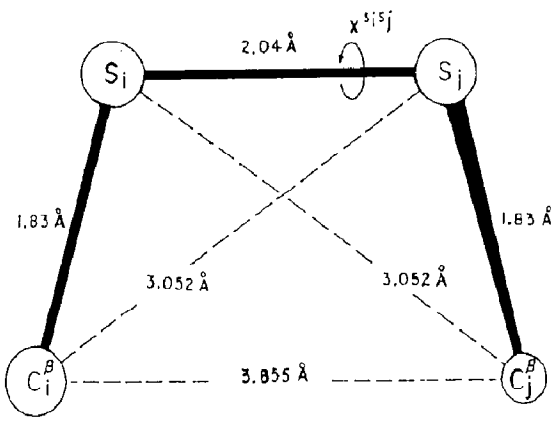

Fig. 9. Relative positions of atoms involved in formation of disulfide bond (adapted from Fig. 16 of Ref. [25]). The values indicate the experimental values for a conformation in which $\chi^{S_iS_j} = \pm 90°$.

varies from near 90° to 0°. We calculated the distances between atom pairs, S_iS_j, $C_i^\beta S_j$, $S_iC_j^\beta$, and $C_i^\beta C_j^\beta$, and the bond angles, $\tau(C_i^\beta S_i S_j)$ and $\tau(S_iS_jC_j^\beta)$ (Fig. 9), for the obtained conformation. The results are tabulated in Table 1. The calculated values are close to the corresponding standard ones [25,26] for the pairs of Cys30 and Cys51, Cys14 and Cys38, and Cys5 and Cys55. In this way, the refinement of the conformation which was obtained by the island model yields certainly the disulfide bridge formation. This indicates that the island model and the lampshade criterion are valid at the initial stage of disulfide bridge formation.

4. Discussion

We have successfully explained the disulfide patterns of BPTI in the present work. Computer simulations of folding of BPTI were also performed by Levitt and Warshel [7], and Tanaka and Scheraga [9]. However, forming and breaking of disulfide bonds were not considered in either work. The method adopted by Levitt and Warshel [9] is as follows. The virtual C^α bond angle is taken as a function of the C^α torsion angle, which is an independent variable. The simulation of the folding process was started with a chain in an extended quasi-random conformation except for the native α-helices which were considered to be performed. The chain was allowed to run down the free energy gradient. The chain was stirred up only if it had settled in an energy minimum, in order to release it from the trap if the minimum was only a local one. As a result, the chain of BPTI folded with about 50% probability to a conformation with a root mean square C^α distance of about 6 Å to the true native structure. Even in the best result, however, β-sheet topology was wrong. In addition, these simulations did not account for the specific disulfide patterns assumed by BPTI during the folding process. On the other hand, Tanaka and Scheraga [9] proposed a hypothesis that the native structure is formed by a three-step mechanism. First, formation of ordered backbone structures by short-range interactions. Second, formation of small contact regions by medium-range interactions. Third, association of the small contact regions into the native structure by long-range interactions. They did not consider disulfide bonds, although they obtained the conformation similar to the native one. However, they did not reveal the mechanism of the formation of various intermediates from either result.

Table 1

Values of geometrical parameters for the disulfide bonds (lengths in angstroms and angles in degrees)

	Pairs	S_iS_j	$C_i^\beta S_j$	$S_iC_j^\beta$	$C_i^\beta C_j^\beta$	$\tau(C_i^\beta S_iS_j)$	$\tau(S_iS_jC_j^\beta)$	$\chi^{S_iS_j}$
Native	30–51	2.02	3.02	3.03	3.79	100.38	105.68	−87.87
	14–38	2.09	3.14	3.09	4.02	107.50	105.85	95.99
	5–55	2.05	3.02	3.00	3.65	96.33	101.94	−81.44
Calculated	30–51	2.04	3.33	3.30	4.39	113.69	115.95	−90.90
	14–38	2.04	3.36	3.29	4.42	116.29	115.06	91.68
	5–55	2.04	3.35	3.29	4.39	113.79	115.22	−90.22
Standard		2.04	3.05	3.05	3.86	104.00	104.00	±90.00

Next, we compare our results with the experiments in order to discuss the role of Cys14 and Cys38 in the folding process. According to the experiment [11–18], the intermediate (C) does not directly fold into the native conformation. On the other hand, mutants of BPTI in which Cys14 and Cys38 were replaced by Ala or Thr were made by Marks et al. [27]. It was shown that BPTI lacking Cys14 and Cys38 can refold properly. This indicates that Cys14 and Cys38 do not contribute to folding of backbone, since these cysteine residues form the disulfide bond at the final stage of folding. In addition, according to Kress and Laskowski [28], partial reduction of BPTI with sodium borohydride results in the selective cleavage of the disulfide bridge (14–38). Huber et al. [29] explained this as follows. The disulfide bridges, (5–55) and (30–51), are buried in the interior of the molecule, but (14–38) connects the loops forming the top of the molecule. From these points, (14–38) can be easily reduced. Our result indicates that the intermediate (C) did not fold into the native conformation and the disulfide bond (14–38) is bonded at the last stage, and is easily attacked by borohydride.

In concluding this article, we would like to assert that the folding pathway of BPTI is explained reasonably by extending the island model. This indicates the following. The pathway from the intermediates to the native conformation is determined by the distribution of hydrophobic pairs, although this pathway is sometimes bifurcated and the reformations of hydrophobic bindings are necessary in the case of BPTI by virtue of the presence of the local conformations (or intermediate structures) which are not uniquely selected while folding the chain. In addition, our geometrical criterion for the formation of disulfide bonds is valid.

Recent experiments by Weissman and Kim [30] showed that the non-native disulfide bonds cannot be found in the refolding processes, contrary to Creighton's experiments, and they proposed a different pathway (Fig. 2(b)). The intermediates (A) and (B) were observed only for a few seconds in Creighton's experiments. Weissman and Kim, however, gave the experimental results after few minutes, during which the intermediates with non-native disulfide bonds are supposed to have decayed. In their figures (see, for example, Fig. 5A in Ref. [30], the traces of the intermediates with non-native disulfide bonds are observed. They also indicated that the native structure can be obtained from the intermediate (C) after a long time (2 months). This is expected as described above (Section 3.1).

The information on the principles of protein architecture will enable us to refine the design of proteins, with the aim of creating proteins with specific properties such as conformational stability. A study along this line is in progress.

Acknowledgements

One of the authors (Y.K.) would like to thank Ms. M. Tanaka and Mr. H. Imamura of C. Itoh Techno-Science Co., Ltd. for their kindness to teach how to use CHARMm.

References

1 N. Saitô, Mem. Sch. Sci. Eng. Waseda Univ. 46 (1982) 295.
2 N. Saitô, T. Shigaki, Y. Kobayashi and M. Yamamoto, Proteins Struct. Funct. Genet. 3 (1988) 199.
3 T. Yoshimura, H. Noguchi, T. Inoue and N. Saitô, Biophys. Chem. 40 (1991) 277.
4 K. Watanabe, A. Nakamura, Y. Fukuda and N. Saitô, Biophys. Chem. 40 (1991) 293.
5 N. Saitô, Adv. Biophys. 25 (1989) 95.
6 N. Saitô, K. Yura and Y. Fukuda, Protein structural analysis, folding and design, ed. M. Hatano (Japan, Scientific Societies Press, Tokyo and Elsevier, Amsterdam, 1990) p. 19.
7 M. Levitt and A. Warshel, Nature 253 (1975) 694.
8 M. Levitt, J. Mol. Biol. 104 (1976) 59.
9 S. Tanaka and H.A. Scheraga, Proc. Natl. Acad. Sci. USA 72 (1975) 3802.
10 H.A. Scheraga, Biopolymers 22 (1983) 1.
11 T.E. Creighton, Proteins, structure and molecular principles (W.H. Freeman, New York, NY, 1983) p. 305.
12 T.E. Creighton, J. Mol. Biol. 95 (1975) 167.
13 T.E. Creighton, J. Mol. Biol. 113 (1977) 275.
14 T.E. Creighton, Prog. Biophys..Mol. Biol. 33 (1978) 231.
15 T.E. Creighton and D.P. Goldenberg, J. Mol. Biol. 179 (1984) 497.
16 D.J. States, T.E. Creighton, C.M. Dobson and M. Karplus, J. Mol. Biol. 195 (1987) 731.
17 T.E. Creighton, Biochem. J. 270 (1990) 1.

18 N.J. Darby, C.P.M. Van Mierlo, T.E. Creighton, FEBS Lett. 279 (1991) 61.
19 J. Israelachvili and R. Pashly, Nature 300 (1982) 342.
20 H.A. Bremermann, Math. Biosci. 9 (1970) 1.
21 T.G. Oas and P.S. Kim, Nature 336 (1988) 42.
22 M. Aida and C. Nagata, Theor. Chim. Acta. 70 (1986) 73.
23 B.R. Broks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan and M. Karplus, J. Comp. Chem. 4 (1983) 187.
24 M. Levitt and S. Lifson, J. Mol. Biol. 46 (1969) 269.
25 F.A. Momany, R.F. McGuire, A.W. Burgess and H.A. Scheraga, J. Phys. Chem. 79 (1975) 2361.
26 D.D. Jones, I. Bernel, M.N. Fey and T.F. Koetzie, Acta Crystallogr. B30 (1974) 1220.
27 C.B. Marks, H. Naderi, P.A. Kosen, I.D. Kuntz and S. Anderson, Science 235 (1987) 1370.
28 L.F. Kess and M.Sr. Laskowski, J. Biol. Chem. 242 (1967) 4925.
29 R. Huber, D. Kukla, A. Rühlmann, O. Epp and H. Formanek, Naturwissenschaften 57 (1970) 389.
30 J.S. Weissman and P.S. Kim, Science 253 (1991) 1386.

Note added after submission of the paper

In his recent papers (T.E. Ceighton, BioEssays 14, March 1992 and Science, in press) , Creighton discussed the disulfide folding pathway in BPTI proposed by Weissman and Kim, and showed that (1) the kinetic folding pathway is not elucidated by the relative levels of accumulation of intermediates, because, for example, the intermediate with less stability and easy to convert to another intermediate cannot accumulate, and (2) the crucial experimental evidences cannot explain their pathway, but confirm the originally proposed pathway of Creighton. The fact pointed out by us in Section 4 for discussion is consistent with the view (1).

The authors are grateful to Dr. Creighton for kindly sending them the above papers before publication.