A mini-protein designed by removing a module from barnase: molecular modeling and NMR measurements of the conformation

Ken-ichi Takahashi1,2, Tosiyuki Noguti1, Hironobu Hojo1,3,4, Kiyoshi Yamauchi5, Masayoshi Kinoshita3, Saburo Aimoto5, Tadayasu Ohkubo6,7 and Mitiko Gō1,8

1Division of Biological Science, Graduate School of Science, Nagoya University, Furo-cho, Chikusa, Nagoya 464-8602, 2Department of Biapplied Chemistry, Faculty of Engineering, Osaka City University, 3-3-138 Sugimoto, Sumiyoshi, Osaka 558-8585, 3Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita 565-0871 and 4Japan Advanced Institute of Science and Technology, Hokuriku, 15 Asahidai, Tatsunokuchi, Ishikawa 923-1292, Japan

5Division of Biological Science, Graduate School of Science, Nagoya University, Furo-cho, Chikusa, Nagoya 464-8602, 6Department of Biapplied Chemistry, Faculty of Engineering, Osaka City University, 3-3-138 Sugimoto, Sumiyoshi, Osaka 558-8585, 7Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita 565-0871 and 8Japan Advanced Institute of Science and Technology, Hokuriku, 15 Asahidai, Tatsunokuchi, Ishikawa 923-1292, Japan

8To whom correspondence should be addressed

© Oxford University Press 1999

673

Protein Engineering vol.12 no.8 pp.673–679, 1999

Introduction

Protein designs are often a modification of local protein structures produced by replacing one or more amino acid residues (Shao and Arnold, 1996). De novo design, in which an amino acid sequence is designed with the expectation that the polypeptide chain folds into a template structure, has also been challenged. Some β-hairpins (Ramirez-Alvarado et al., 1996), bundle structures of α-helices (Betz et al., 1995) and a βα motif (Dahiyat and Mayo, 1997) were indeed designed de novo. Polypeptides designed to fold into mixed α/β- or all β-structures had the expected secondary structural features but their tertiary structures were molten and globule-like rather than rigid (Betz et al., 1993).

The ancestors of present-day proteins were produced during the course of pre- and early biological evolution. Studies of putative mechanisms of protein evolution should provide insights into protein design. The globular domain of a protein consists of structural units, termed modules, which are defined as compact and contiguous segments of 10–30 amino acid residues (Gō, 1981, 1983). There are many cases where the boundaries of a module correlate with intron positions within the gene; for example, hemoglobin (Gō, 1981; Jensen et al., 1981), lysozyme (Gō, 1983; Isaacs et al., 1985), triose phosphate isomerase (Gilbert et al., 1986; Gō and Nosaka, 1987; Gilbert and Glynias, 1993; Tittiger et al., 1993) and so forth (Gō, 1985). The observed correlation suggests that each module of a given ancestral protein was encoded by an exon, and the proteins evolved by assembling modules through exon shuffling mediated by introns (Gō and Nosaka, 1987).

Barnase, a bacterial RNase from Bacillus amyloliquefaciens, is a single domain protein consisting of 110 amino acid residues (Nishimura and Nomura, 1959; Hartley and Barker, 1972). Its conformation has been elucidated by X-ray crystallography (Mauguen et al., 1982) and by NMR spectroscopy (Bycroft et al., 1991) and shows three α-helices, a two-stranded parallel β-sheet and a five-stranded anti-parallel β-sheet. The boundaries of the modules were assigned to residues 24, 52, 73, 88 and 98 by Noguti et al. (1993). The six modules are M1, 1–24; M2, 25–52; M3, 53–73; M4, 74–88; M5, 89–98 and M6, 99–110 (Figure 1a). Modules M1 and M2 have one and two α-helices, respectively. Module M2 forms a parallel β-sheet near the N- and C-terminal ends.

When we examined the atomic interactions of barnase, we found that the hydrogen bonds were mainly localized within each module (Noguti et al., 1993). Solution structures of chemically synthesized modules M2 and M3 were studied using two-dimensional NMR. The formation of secondary structures in the modules was shown to be at similar positions as those in barnase (Ikura et al., 1993). Module M1 was found to assemble together in aqueous solution and to form a stable, filamentous structure. Scanning tunneling microscopy showed that each M1 module in the assembly took on a compact conformation similar to that seen in barnase (Yoshida et al., 1993). These observations indicate that the compact conformation of a module is specified predominantly by interactions within the module itself and the conformation is further stabilized by interactions with associated modules. Thus modules may have served as the building blocks of globular domains during protein evolution.

If this view of protein evolution is valid, one could acquire an artificial protein by removing one or more modules from a globular domain of a present-day protein; the polypeptide chain consisting of the remaining modules will be able to fold into a stable conformation similar to that of the intact protein, except for the absent parts. Design of such a mini-protein is...
Among the six modules of barnase, M2 forms the least number of hydrogen bonds and hydrophobic contacts with the remaining protein (Noguti et al., 1993; Noguti and Go, 1995). The influence of the excision of one module on conformational stability of the remaining portion of barnase was expected to be the least when M2 is removed. We designed a mini-protein by removing the M2 module from barnase. This mini-barnase was chemically synthesized and its structure was examined by CD and NMR. NMR revealed that the mini-barnase folded into a stable conformation with a similar hydrophobic core structure to that of barnase.

Materials and methods

Design of mini-barnase

When we removed module M2 (segment 25–52) from barnase two chains, M1 and M3–M6, remained. The carbonyl carbon atom at the C-terminus of M1 in barnase is 6.5 Å away from the amide nitrogen atom at the N-terminus of M3. If the two atoms are connected chemically, much stress will be placed on the conformation of the M1 and M3–M6 chains and the conformation will be unstable.

Because a module has a compact conformation, the ends are relatively close in proximity. When we calculated the distance between a carbonyl carbon atom in the vicinity of the boundary of the module M1 and M2, and an amide nitrogen atom in the vicinity of the boundary of M2 and M3, the distance between the carbonyl carbon atom of Ile25 and the amide nitrogen atom of Gly52 was 3.9 Å, the shortest. To acquire a smooth join between the C-terminal end of M1 and the N-terminal end of M3, terminal residues 25 and 52 were included in the design of the mini-barnase.

The conformation of mini-barnase was modeled in the following procedures using the program BIOGRAF (Version 3.21, Molecular Simulations, Inc., Burlington, MA, 1992). First, segment 26–51 was removed from the crystal structure of barnase [Baudet and Janin, 1991; Brookhaven Protein Data Bank (Bernstein et al., 1977), entry 1RNB]. Secondly, residues 25 and 52 were connected. Then energy minimization was performed, relaxing residues 23–25 and 52–54. Irregular geometry of chemical bonds and collision of atoms were never observed in the modeled conformation. Thirdly, conformation of the first residue and that of the second side chain, which were lacking in the crystal structure of barnase, were modeled using atomic coordinates of the crystal structure of the complex of barnase and barstar (Guillet et al., 1993; Brookhaven Protein Data Bank, entry 1BGS) after superimposing the α carbon, carbonyl carbon and amide nitrogen atoms of Glu2 in the two crystal structures of barnase (1RNB and 1BGS). To increase thermodynamic conformational stability and solubility of minibarnase, Ile25 and Phe82 were replaced by Thr and Ala, respectively (see Results). Hereafter we refer to mini-barnase prior to the replacement of these residues as ‘mini-barnase (unmodified)’ and that after the replacements as ‘mini-barnase (modified)’, in order to distinguish between the two versions.

Four layers of water molecules were put on the modeled mini-barnase (unmodified) and energy related to water and protein system was minimized until the root mean square (r.m.s.) of atomic forces was less than 0.01 kcal mol⁻¹ Å⁻¹.

Molecular dynamics simulation was performed on this system for 115 ps. The temperature was maintained at 298 K by the heat-bath-coupling method (τ = 0.2 ps) (Berendsen et al., 1984); the time step was 1 fs. In minimization and molecular dynamics simulation a linear distance dependent dielectric function, ε = r, was applied and nonbonded interactions were truncated at 9 Å distance. Hydrogen atoms that can form hydrogen bonds were treated explicitly but other hydrogen atoms were united with heavy atoms. Side chains of Lys, Arg, Asp and Glu, and both ends of the main chain were ionized.

During the simulation some water molecules vaporized out of the water shell. After execution of the first 25 ps simulation, we filled defects of the four-layered water shell with water molecules. The water shell was equilibrated with 0.8 ps molecular dynamics simulation after 200 cycles of conjugate gradient minimization, then the simulation was continued. Supplement of water molecules was repeated in this manner every 10 ps after the first 25 ps. We used the program and energy parameters of BIOGRAF. Molecular dynamic simulation of mini-barnase (modified) was carried out in the same manner as that of mini-barnase (unmodified), except for a simulation time length of 75 ps.

Oobatake and Ooi (1993) developed a method for predicting thermodynamics of protein unfolding, whereby Gibbs free energy change in protein unfolding, ΔG°U, was calculated as a linear function of changes in accessible surface areas (ASAs) from those of the folded conformation to those of the extended conformation. Using this method, we estimated the contribution of residues to ΔG°U for mini-barnase. Contribution of residue i to ΔG°U was calculated using the following equation

\[ \Delta G°U_i = \Sigma_i (g_{jb} + g_{jc}) (ASA^U_i - ASA^F_i), \]

where ΔG°U is the contribution of residue i, g JB and g JC are the hydration and chain free energy per unit of ASA of atom j, respectively, and ASA^U and ASA^F are the values of ASA of atom j in unfolded and folded states, respectively. The summation was done over all heavy atoms of residue i. No hydrogen atom was treated explicitly in the calculation of atomic ASAs. ΔG°U is the sum of ΔG°U_i over all residue i's. Oobatake and Ooi (1993) evaluated hydration and chain free energy per unit of ASA for seven atomic groups, aliphatic carbon, aromatic

![Fig. 1. Modules of barnase and the designed mini-barnase. Six modules, M1–M6, are shown in color: M1, sky blue; M2, red; M3, magenta; M4, green; M5, blue and M6, yellow. (a) The modules of barnase in tube (left) and space-filling models (right) are shown in the direction of its catalytic sites. (b) Conformation of the modeled mini-barnase by removing module M2, sampled at 115 ps in molecular dynamic simulation. N- and C-terminal residues of M2 left in mini-barnase are shown in red.](image-url)
Design of a mini-protein

carbon, hydroxyl oxygen, amide nitrogen, carbonyl carbon, carbonyl oxygen and sulfur, by least-squares fitting to experimental free energy changes of solution of small monofunctional molecules and those of unfolding of proteins. They assumed a linear relation between changes in free energies and those of atomic ASAs. The same set of values at 298 K was used here. Atomic ASAs in the unfolded state were represented by those of the extended polypeptide chain in the studies of Oobatake and Ooi (1993). We calculated atomic ASAs of the unfolded state of mini-barnase using their method. The van der Waals radii of heavy atoms and water were the same as those in their studies. They used X-ray structures of proteins for calculating atomic ASAs in the folded states of proteins. Since we wanted to design a protein whose X-ray or NMR structure did not exist, we used conformations in the records of the molecular dynamics simulations of the molecularly-modeled mini-barnase. ASA values of atom j in conformations sampled every 0.1 ps in a period of 5 ps were averaged to represent \( \text{ASA}_j \) in the period of the simulation. The solubility of a protein is determined by the Gibbs free energy change of aggregation, \( \Delta G^A \). We predicted the contribution of residue i to \( \Delta G^A \), using the following equation

\[
\Delta G_i^A = \sum_j \left( g_{ij} + g_{ji} \right) \left( \text{ASA}_j^0 - \text{ASA}_j^A \right),
\]

where \( \Delta G_i^A \) is the contribution of residue i, and \( \text{ASA}_j^A \) is the value of ASA of atom j in the aggregated state. In this state, molecules assemble and their access to water molecules is interrupted. Therefore, we assumed that atomic ASAs were essentially zero in the aggregated states, i.e. \( \text{ASA}_j^A = 0 \). \( \Delta G^A \) is the sum of \( \Delta G_i^A \)'s over all residue i’s.

Atomic ASA was calculated by the method of Shrake and Rupley (1973). The ASA value obtained contains an error caused by incompleteness of uniform distribution of 92 points over the surface of a sphere. The error is not negligible in the calculation of free energies (Oobatake and Ooi, 1993). We improved the accuracy of atomic ASA calculation using the following procedure. We generated a hundred different sets of atomic coordinates by rotating a molecule randomly in the same conformation, and calculated atomic ASAs for each set of the coordinates by the method of Shrake and Rupley (1973). Values of the respective atomic ASAs that were sufficiently precise for the calculation of the free energies were obtained by averaging the 100 sets of calculated ASAs.

Accessibility of a residue was defined as the ratio between the sum of the ASAs of atoms of the residue in the folded conformation and that in the extended form (Gö and Miyazawa, 1980), a value which is indicative of whether the residue is located on the surface or in the interior of the molecule.

Chemical synthesis, and NMR and CD measurements of mini-barnase

Boc-amino acid derivatives, reagents for solid-phase peptide synthesis and \(^{15}\)N-labeled amino acids were purchased from Peptide Institute, Inc. (Osaka, Japan), Watanabe Chemical Ind., Ltd (Hiroshima, Japan) and Nihon Shoji Co., Ltd (Osaka, Japan), respectively. Chemical synthesis of the designed mini-barnase (modified) was carried out as previously described (Hojo and Aimoto, 1993). Briefly, partially protected peptide thioesters corresponding to mini-barnase (1–25 + 52), (53–81) and partially protected peptide (82–110) were prepared via a solid-phase method, using Boc-amino acids. For introduction of the residues Ala11, Tyr13, Leu14, Leu89, Tyr90, Ser91, Ser92, Asp93, Leu95 and Tyr97, \(^{15}\)N-labeled amino acid derivatives were used. The peptide segments were successively condensed by activation of thioester groups with silver ions to give a protected form of mini-barnase. After deprotection, mini-barnase was highly purified by HPLC. The purity was confirmed by analytical HPLC, amino acid analysis and matrix-assisted laser desorption ionization (MALDI) mass spectrometry.

Lyophilized mini-barnase was dissolved in 90% H₂O/10% D₂O. Acids remaining after the purification were excluded by ultrafiltration until the pH of solution was 5.1; simultaneously, mini-barnase was concentrated to 0.1 mM. Two-dimensional \(^1\)H–\(^{15}\)N HSQC spectrum (Boyd et al., 1992) was measured at 25°C on a Varian Unity-plus 750 spectrometer. Resonance of the solvent water was suppressed by presaturation during the relaxation delay. The spectrum consisted of 128 complex t1 points with each free induction decay having 2048 complex t2 points. 160 transients were acquired per t1 value to gain a signal to noise ratio. The spectral width for \(^1\)H observation was 12 000 Hz and that for \(^{15}\)N was 3000 Hz. The spectrum was zero-filled to 2048 × 512 real points and was processed with shifted sinebell window function, in both dimensions. Linear prediction was applied to raise the spectral resolution of the \(^{15}\)N dimension. \(^1\)H and \(^{15}\)N chemical shifts were referenced to external TSP (0 p.p.m.) and NH₄Cl (24.93 p.p.m.), respectively.

CD spectra were measured using a Jasco J-600 spectropolarimeter at room temperature in a quartz cell with a light path length of 1.0 cm. The concentration of mini-barnase was 0.6 \(\mu\)M, determined by measurement of UV absorbance at 280 nm. The extinction coefficient of the mini-barnase was estimated to be 21 750 M⁻¹ cm⁻¹ on the basis of the number of tryptophan and tyrosine residues in the peptide (Kuramitsu et al., 1990). The pH was adjusted with NaOH or HCl.

Results

Molecular modeling of mini-barnase

To acquire a mini-barnase we chose to remove module M2. By removing the segment 26–51 and connecting residues 25 and 52 of the remaining protein, we prepared a molecular model of mini-barnase lacking module M2 (Figure 1b). In this report, the same numbering of residues and modules have been used for mini-barnase as those for barnase, to facilitate correspondence. We considered the following points: (i) a conformationally smooth connection of polypeptide chains of both ends of the module M2; (ii) mechanical stability of the designed conformation; and (iii) thermodynamic stability of the conformation and solubility of the acquired model.

Mechanical stability of modeled mini-barnase

The conformation of a native protein is mechanically stable as the balance of all forces on its atoms is maintained at small amplitude thermal perturbations. The absence of 26 amino acid residues might perhaps disturb the balance of atomic forces and the mini-barnase would be unable to maintain conformation. We examined the mechanical stability of the modeled mini-barnase conformation by studying the molecular dynamics. The conformation of mini-barnase was retained during 115 ps in a water environment, except for the loop of module M4 (Figure 1). \(C_\alpha\) atoms in the conformation on the trajectory of the simulation at 115 ps were superimposed on the corresponding atoms in the 3D structure of barnase, except for those in the loop formed by residues 79–83. The root mean square deviations (r.m.s.d.) of the superimposed \(C_\alpha\) atoms was
1.3 Å. Thus, conformation of the modeled mini-barnase was essentially mechanically stable. For Cα atoms in the loop, the r.m.s.d. value was 4.7 Å. Movement of the loop narrowed the cleft between modules M1 and M4 (Figure 1).

**Increase of thermodynamic conformational stability and solubility of mini-barnase by amino acid replacement**

An appropriate method for the prediction of thermodynamic stability of a protein conformation has yet to be established. Whether or not the designed mini-barnase takes on the expected conformation has to be examined experimentally. By doing so, one can test the prediction and improve methods used for prediction. In the process of designing a mini-barnase, we directed attention to the amino acid residues that might contribute to the decrease in thermodynamic conformational stability or solubility of the molecule, and to substitute suitable residues in their place. High solubility in aqueous solution is a sine qua non for NMR measurements.

After excision of module M2, the interface between M2 and the other modules in barnase becomes part of the surface of the mini-barnase. Twelve residues of the mini-barnase gained solvent accessibility by more than 0.1 from those of barnase. Of the 12, four were hydrophobic: Val3, Ile25, Tyr78 and Phe82 (Figure 2). Exposure of a hydrophobic residue on the protein surface usually decreases thermodynamic conformational stability and/or solubility of the molecule, and to substitute suitable residues in their place. High solubility in aqueous solution is a sine qua non for NMR measurements.

After excision of module M2, the interface between M2 and the other modules in barnase becomes part of the surface of the mini-barnase. Twelve residues of the mini-barnase gained solvent accessibility by more than 0.1 from those of barnase. Of the 12, four were hydrophobic: Val3, Ile25, Tyr78 and Phe82 (Figure 2). Exposure of a hydrophobic residue on the protein surface usually decreases thermodynamic conformational stability and/or solubility of the molecule, and to substitute suitable residues in their place. High solubility in aqueous solution is a sine qua non for NMR measurements.

1.3 Å. Thus, conformation of the modeled mini-barnase was essentially mechanically stable. For Cα atoms in the loop, the r.m.s.d. value was 4.7 Å. Movement of the loop narrowed the cleft between modules M1 and M4 (Figure 1).

**Increase of thermodynamic conformational stability and solubility of mini-barnase by amino acid replacement**

An appropriate method for the prediction of thermodynamic stability of a protein conformation has yet to be established. Whether or not the designed mini-barnase takes on the expected conformation has to be examined experimentally. By doing so, one can test the prediction and improve methods used for prediction. In the process of designing a mini-barnase, we directed attention to the amino acid residues that might contribute to the decrease in thermodynamic conformational stability or solubility of the molecule, and to substitute suitable residues in their place. High solubility in aqueous solution is a sine qua non for NMR measurements.

After excision of module M2, the interface between M2 and the other modules in barnase becomes part of the surface of the mini-barnase. Twelve residues of the mini-barnase gained solvent accessibility by more than 0.1 from those of barnase. Of the 12, four were hydrophobic: Val3, Ile25, Tyr78 and Phe82 (Figure 2). Exposure of a hydrophobic residue on the protein surface usually decreases thermodynamic conformational stability and/or solubility of the molecule, and to substitute suitable residues in their place. High solubility in aqueous solution is a sine qua non for NMR measurements. After excision of module M2, the interface between M2 and the other modules in barnase becomes part of the surface of the mini-barnase. Twelve residues of the mini-barnase gained solvent accessibility by more than 0.1 from those of barnase. Of the 12, four were hydrophobic: Val3, Ile25, Tyr78 and Phe82 (Figure 2). Exposure of a hydrophobic residue on the protein surface usually decreases thermodynamic conformational stability and/or solubility of the molecule, and to substitute suitable residues in their place. High solubility in aqueous solution is a sine qua non for NMR measurements.

After excision of module M2, the interface between M2 and the other modules in barnase becomes part of the surface of the mini-barnase. Twelve residues of the mini-barnase gained solvent accessibility by more than 0.1 from those of barnase. Of the 12, four were hydrophobic: Val3, Ile25, Tyr78 and Phe82 (Figure 2). Exposure of a hydrophobic residue on the protein surface usually decreases thermodynamic conformational stability and/or solubility of the molecule, and to substitute suitable residues in their place. High solubility in aqueous solution is a sine qua non for NMR measurements. After excision of module M2, the interface between M2 and the other modules in barnase becomes part of the surface of the mini-barnase. Twelve residues of the mini-barnase gained solvent accessibility by more than 0.1 from those of barnase. Of the 12, four were hydrophobic: Val3, Ile25, Tyr78 and Phe82 (Figure 2). Exposure of a hydrophobic residue on the protein surface usually decreases thermodynamic conformational stability and/or solubility of the molecule, and to substitute suitable residues in their place. High solubility in aqueous solution is a sine qua non for NMR measurements. After excision of module M2, the interface between M2 and the other modules in barnase becomes part of the surface of the mini-barnase. Twelve residues of the mini-barnase gained solvent accessibility by more than 0.1 from those of barnase. Of the 12, four were hydrophobic: Val3, Ile25, Tyr78 and Phe82 (Figure 2). Exposure of a hydrophobic residue on the protein surface usually decreases thermodynamic conformational stability and/or solubility of the molecule, and to substitute suitable residues in their place. High solubility in aqueous solution is a sine qua non for NMR measurements. After excision of module M2, the interface between M2 and the other modules in barnase becomes part of the surface of the mini-barnase. Twelve residues of the mini-barnase gained solvent accessibility by more than 0.1 from those of barnase. Of the 12, four were hydrophobic: Val3, Ile25, Tyr78 and Phe82 (Figure 2). Exposure of a hydrophobic residue on the protein surface usually decreases thermodynamic conformational stability and/or solubility of the molecule, and to substitute suitable residues in their place. High solubility in aqueous solution is a sine qua non for NMR measurements. After excision of module M2, the interface between M2 and the other modules in barnase becomes part of the surface of the mini-barnase. Twelve residues of the mini-barnase gained solvent accessibility by more than 0.1 from those of barnase. Of the 12, four were hydrophobic: Val3, Ile25, Tyr78 and Phe82 (Figure 2). Exposure of a hydrophobic residue on the protein surface usually decreases thermodynamic conformational stability and/or solubility of the molecule, and to substitute suitable residues in their place. High solubility in aqueous solution is a sine qua non for NMR measurements. After excision of module M2, the interface between M2 and the other modules in barnase becomes part of the surface of the mini-barnase. Twelve residues of the mini-barnase gained solvent accessibility by more than 0.1 from those of barnase. Of the 12, four were hydrophobic: Val3, Ile25, Tyr78 and Phe82 (Figure 2). Exposure of a hydrophobic residue on the protein surface usually decreases thermodynamic conformational stability and/or solubility of the molecule, and to substitute suitable residues in their place. High solubility in aqueous solution is a sine qua non for NMR measurements. After excision of module M2, the interface between M2 and the other modules in barnase becomes part of the surface of the mini-barnase. Twelve residues of the mini-barnase gained solvent accessibility by more than 0.1 from those of barnase. Of the 12, four were hydrophobic: Val3, Ile25, Tyr78 and Phe82 (Figure 2). Exposure of a hydrophobic residue on the protein surface usually decreases thermodynamic conformational stability and/or solubility of the molecule, and to substitute suitable residues in their place. High solubility in aqueous solution is a sine qua non for NMR measurements. After excision of module M2, the interface between M2 and the other modules in barnase becomes part of the surface of the mini-barnase. Twelve residues of the mini-barnase gained solvent accessibility by more than 0.1 from those of barnase. Of the 12, four were hydrophobic: Val3, Ile25, Tyr78 and Phe82 (Figure 2). Exposure of a hydrophobic residue on the protein surface usually decreases thermodynamic conformational stability and/or solubility of the molecule, and to substitute suitable residues in their place. High solubility in aqueous solution is a sine qua non for NMR measurements. After excision of module M2, the interface between M2 and the other modules in barnase becomes part of the surface of the mini-barnase. Twelve residues of the mini-barnase gained solvent accessibility by more than 0.1 from those of barnase. Of the 12, four were hydrophobic: Val3, Ile25, Tyr78 and Phe82 (Figure 2). Exposure of a hydrophobic residue on the protein surface usually decreases thermodynamic conformational stability and/or solubility of the molecule, and to substitute suitable residues in their place. High solubility in aqueous solution is a sine qua non for NMR measurements.
Design of a mini-protein

Fig. 3. CD spectra of mini-barnase measured at pH 2.5 ( ), 3.0 ( ), 3.5 ( ), 4.5 ( ), 5.3 ( ) and 6.4 ( ).

Fig. 4. 1H-15N HSQC spectrum of mini-barnase labeled with 15N site-specifically. Cross-peaks of amide protons and 15Ns of 10 residues, Ala11, Tyr13, Leu14, Leu89, Tyr90, Ser91, Ser92, Asp93, Leu95 and Tyr97, are shown but not assigned.

Table II. Predicted contributions of two hydrophobic residues to Gibbs free energy changes of unfolding and aggregation of mini-barnase after amino acid replacement

<table>
<thead>
<tr>
<th>Residue</th>
<th>Contribution to ΔG_U (kcal mol⁻¹)</th>
<th>Contribution to ΔG_A (kcal mol⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>40–45 ps</td>
<td>70–75 ps</td>
</tr>
<tr>
<td>Thr25</td>
<td>0.12</td>
<td>0.11</td>
</tr>
<tr>
<td>Ala82</td>
<td>0.05</td>
<td>0.06</td>
</tr>
</tbody>
</table>

*a* See footnote of Table I.

*b* Values in the columns of 40–45 ps were calculated by equations (1) and (2) (see Materials and methods), where atomic ASAs in the folded state are the averages of respective atomic ASAs of conformations in the period 40–45 ps in molecular dynamics of mini-barnase after amino acid replacement. Values in the 70–75 ps columns were calculated in the same manner, using the conformations in the period 70–75 ps.

Tyr97 (Figure 5). Dispersion of the cross-peaks of the labeled residues in mini-barnase was 10.2–7.9 p.p.m. in the 1H axis and 129–114 p.p.m. in the 15N axis. It was much larger than that of the cross-peaks of Ala, Tyr, Leu, Ser and Asp in a random coil state: 8.4–8.0 p.p.m. in the 1H axis and 123–116 p.p.m. in the 15N axis (Bundi and Wüthrich, 1979; Wishart et al., 1991). Thus, mini-barnase was not in a random coil or molten globule state, rather a stable and specific conformation was formed.

The 1H-15N HSQC spectrum of uniformly 15N-labeled barnase was measured at pH 6.6, 30°C and the complete assignment of the cross-peaks has been reported (Jones et al., 1993). We picked up 10 cross-peaks from the corresponding residues to the 15N-labeled residues in mini-barnase, and compared them with the unassigned cross-peaks in the 1H-15N HSQC spectrum of mini-barnase. Because of the different solvent conditions of pH and temperature, the two measurements have systematic differences in values of chemical shifts. The 1H-15N HSQC spectrum of mini-barnase could be superimposed on that of the picked-up cross-peaks of barnase by shifting 0.1 and 4.6 p.p.m. along the 1H and 15N axes, respectively (Figure 6). Patterns of cross-peak locations in the two spectra were so alike that one-to-one correspondence was
observed between the cross-peaks of mini-barnase and those of barnase. Using this correspondence we could assign all 10 cross-peaks to the $^{15}$N-labeled residues of the mini-barnase (Table III). Chemical shift differences of the corresponding cross-peaks were slight as shown in Table III.

If the three-dimensional (3D) structure of mini-barnase is similar to that of the corresponding portion of barnase, local magnetic fields of the residues of mini-barnase will be similar to those of the corresponding residues in barnase, except for residues near the module M2, then chemical shifts of respective residues will also be similar for the two proteins. The good correspondence of the cross-peaks of HSQC spectra between mini-barnase and barnase is consistent with the conformational similarity in the simulated model of mini-barnase and the intact barnase. Three of the 10 labeled residues, Ala11, Tyr13 and Leu14, and the other seven residues, Leu89, Tyr90, Ser91, Ser92, Asp93, Leu95 and Tyr97, are located on the buried face of the N-terminal $\alpha$-helix in module M1 and in the $\beta$-hairpin in M5, respectively, in the 3D structure of barnase and the molecular model of mini-barnase (Figure 5). The $\beta$-hairpin is the central part of the five-stranded anti-parallel $\beta$-sheet, which forms a hydrophobic core with the N-terminal $\alpha$-helix. Our NMR measurements indicate that a stable conformation of mini-barnase was formed, at least in regions around the hydrophobic core sandwiched between the $\alpha$-helix and the $\beta$-sheet.

**Discussion**

By removing a module from barnase, we designed a mini-barnase of 84 residues which was then chemically synthesized by site-specifically labeling with $^{15}$N. NMR spectrum indicated that the mini-barnase folded into a stable specific conformation that has a similar hydrophobic core structure to that of barnase.

Module M2 had the most independent conformation among the six modules of barnase (Noguti et al., 1993; Noguti and Gō, 1995). Twenty-six residues of M2 were removed from barnase; deletion of such a fragment from a globular protein would appear difficult in molecular modeling. We attained success in modeling the conformation of mini-barnase which proved to be mechanically stable, presumably for the following reasons. First, N- and C-terminal residues of a module are often located close in space because of a compact and small conformation of a module. This characteristic of modules facilitates the connection of two chains that remain after removal of a module by shifting the excision points one or two residues from module boundaries. Secondly, most of the module conformations in barnase are mechanically stable (Takahashi et al., 1997). Removal of a module from a globular protein seems to maintain the conformation of the remaining parts and mechanical stability is retained.

De Sanctis et al. (1988) made a mini-myoglobin of 108 amino acid residues, a peptide excised from myoglobin and which closely corresponds to the peptide encoded by the central exon but contains an additional 30 residues at the C-terminal end. The peptide binds heme stoichiometrically, forms a helical structure and exhibits the reversible oxygen-binding function of myoglobin itself. Pessi et al. (1993) designed a mini-protein of 61 amino acid residues, referred to as a minibody, using the central portion of the heavy chain variable domain of an immunoglobulin as a template, which constitutes a compact sub-domain consisting of a $\beta$-sandwich of two triple-stranded $\beta$-sheets and contains the hypervariable loops H1 and H2. Hydrophobic residues buried in the parent molecule, but exposed in the minibody, were suitably replaced and a metal-binding site was engineered by introducing one histidine in H1 and two in H2. The minibody exhibited all-$\beta$ structural features following CD spectra, cooperative transition in urea-induced unfolding and Zn-binding activity. The design of mini-proteins will become an important technology in protein engineering, because the small size of designed proteins facilitates theoretical and experimental treatments such as molecular dynamics simulation and structure determination by NMR.

Exchange of modules or exons between homologous proteins has been previously reported: Kumagai et al. (1992) succeeded in functional conversion of $\alpha$-lactalbumin and the c-type lysozyme by exon exchange. These two proteins are functionally different but structurally homologous and their genes have the same exon–intron organization. Exchange of the second exon of $\alpha$-lactalbumin with that of lysozyme produced a chimeric $\alpha$-lactalbumin with lysozyme activity, whereas the replacement of 10 amino acids, which are directly involved in substrate-binding, did not produce $\alpha$-lactalbumin with lysozyme activity. Yano et al. (1996) successfully converted the coenzyme specificity of isocitrate dehydrogenase (ICDH) from NADP-dependent to NAD-dependent by replacing a module involved in coenzyme binding with the corresponding module of NAD-dependent enzyme 3-isopropylmalate dehydrogenase.
Design of a mini-protein

(IPMDH). Replacement of a few residues interacting with the 2’-phosphate group of NADP did not result in conversion of coenzyme specificity and seemed to destabilize the structure of ICDH. The module-replaced ICDH was as stable as the wild-type ICDH. Wakasugi et al. (1994) in making a chimeric globin replaced the fourth module of the hemoglobin β-subunit by the fourth module of the hemoglobin α-subunit. The chimeric globin associated with the hemoglobin β-subunit through interactions between the fourth module of the chimera and that of the β-subunit. These studies demonstrate designs of functionally chimeric proteins by exchanging modules. Indeed modules are functional units as well as structural units of proteins.

Fersht’s group has studied and obtained valuable information on the folding processes and conformational stability of barnase, mainly using systematic amino acid replacements (Matouschek et al., 1989). They also studied truncated barnase in which the N-terminal 22 or 36 residues were excised (Sancho and Fersht, 1992; Kippen et al., 1994). The truncated barnases did not appear to form stable conformations in the isolated state. Our approach to designing a mini-protein was to remove a single module of the protein. To confirm the success of the design of mini-barnase, we are in the process of determining its 3D structure.

Acknowledgements

We would like to thank Drs Y.Mukohata, Y.Sagiya and A.Mizoguchi for use of the CD spectrometer and centrifuge at their laboratories, K.Tanaka for help in calculation of mini-barnase design, and M.Ohara for helpful comments on the manuscript. This work was supported by Grants-in-Aid for Scientific Research on Priority Areas and for Scientific Research (A) from the Ministry of Education Science, Sports and Culture of Japan.

References