Permutation of modules or secondary structure units creates proteins with basal enzymatic properties

Toru Tsuji[a,b], Kensei Kobayashi[a], Hiroshi Yanagawa[b,*]

[a]*Department of Chemistry and Biotechnology, Yokohama National University, Tokiwadai, Hodogaya-ku, Yokohama 240, Japan*
[b]*Mitsubishi Kasei Institute of Life Sciences, Minamiooya, Machida, Tokyo 194, Japan*

Abstract The RNase activity of barnase mutants obtained by the permutation of modules or secondary structure units was investigated. Four of the 45 mutants had weak but distinct RNase activity, and they had unique optimum pHs and temperatures like natural enzymes. One of the active mutants had an ordered conformation, but the others did not. An active mutant having disordered conformation formed an ordered conformation in the presence of GMP, which is an inhibitor of this mutant. These results indicate that the amino acid sequences derived from barnase have sufficient plasticity to be rearranged into different proteins with basal enzymatic properties.
© 1999 Federation of European Biochemical Societies.

Key words: Global sequence space; Directed evolution; Combinatorial method; Induced folding; Barnase; Protein evolution

1. Introduction

Functional nucleic acids consisting of completely artificial sequences have been obtained by directed in vitro evolution from random sequence libraries [1,2]. However, the difficulty of genotype (nucleic acid) assignment to phenotype (protein) has limited the size of a protein library considerably, therefore, it is difficult to use random sequence libraries [3,4] for directed evolution of proteins. The present directed evolution of proteins can only improve or modify ready-made proteins [5,6].

Recently we have succeeded in constructing a molecule assigning a genotype to a phenotype, 'in vitro virus', in which in vitro synthesized polypeptides can be directly attached to its encoded message through a puromycin derivative that is synthetically coupled to the 3′ end of the mRNA [7]. The in vitro virus can expand the protein library size to $\sim 10^{12}$ [7,8]. Therefore, it is timely to establish the strategy for the searching of global sequence space. The library constructed with the

combinatorial method using peptide fragments derived from natural protein sequences should be useful for the searching of semi-global sequence space.

For the initial attempt, barnase (a bacterial ribonuclease consisting of 110 amino acids [9]) was chosen as a model protein, because its functional [10,11] and folding mechanisms [12] have been studied extensively, and it has been the subject of protein anatomy studies [13,14]. Barnase can be divided into six modules (M1–M6), which are contiguous peptide chains forming compact structures [15], or six secondary structure units (S1–S6), which are also contiguous peptide chains, but forming tertiary extended regions [16].

To examine whether amino acid sequences corresponding to these modules or secondary structure units have sufficient plasticity to be rearranged into novel enzymes, the RNase activity of 23 module mutants and 22 secondary structure unit mutants was investigated. These mutants were obtained by the permutation of the internal four of six modules or secondary structure units (Fig. 1).

2. Materials and methods

2.1. Construction of barnase mutants
The procedures for the construction of mutant genes, protein expression and protein purification were described in our previous paper [16].

2.2. Activity staining
Activity staining [13] was performed on SDS-PAGE with a separate gel containing poly A or poly G. To remove SDS, after SDS-PAGE the gel was washed for 2 h in 5 mM bis-Tris-HCl buffer (pH 6.0). The gel was further left to allow cleavage of the substrate for 12 h in the same buffer, and then stained with buffer containing 0.2% toluidine blue O (Merck). The procedures described above were performed at 4°C and 25°C.

2.3. Hydrolysis of polymers of 1-N^6-ethenoadenosine 5′-monophosphate (poly eAp)
The hydrolysis was started by adding enzyme solution (5–100 µl of 5 µM mutant or 5 µl of 0.18 µM wild-type barnase) to buffer solution containing 0.2 mM poly eAp in a thermostatted cell holder. Each solution was equilibrated at each temperature for 10 min before starting the reaction. The increment of fluorescence was monitored as a function of time with a spectrofluorophotometer (Shimadzu RF-502). The initial velocity was calculated based on the initial linear portion of the increment of fluorescence. When the optimum pH of enzymes was examined, the increments of the fluorescence at different pHs were calibrated based on the fluorescence intensity of monomers of eAp (Sigma). Buffers were all at same ionic strength ($I = 0.05$); sodium acetate/acetic acid was used at pH 5.0, bis-Tris-HCl at pH 5.5–7.0, and Tris-HCl at pH 7.5–9.0. Excitation was at 320 nm with emission measured at 410 nm. Poly eAp was synthesized by the reaction of poly A with chloroacetaldehyde in phosphate buffer at pH 6.8 and 37°C for 48 h.

2.4. CD measurements
CD measurements were made in a Jasco J-725 spectropolarimeter

fitted with a thermostatted cell holder. Ellipticities are expressed in units of deg cm^2 dmol^{-1} using the mean residue concentration. The baselines due to the buffer (5 mM bis-Tris-HCl, pH 6.0) or buffer containing nucleotide were subtracted from the spectra. Thermal unfolding experiments were performed at a heating rate of 50°C/h using a response time of 4 s. The effect of 2′ (3′)-GMP (Yamasa) on the CD spectra of the mutants was examined at a protein concentration of 3 μM and a 2′ (3′)-GMP concentration of 15 μM, at 5°C.

3. Results and discussions

3.1. Activity staining

To examine whether module mutants and secondary structure unit mutants have RNase activity or not, and to exclude the possibility of contamination with RNases, the activity was confirmed by activity staining [13] using poly A and poly G embedded in SDS-polyacrylamide gel, because wild-type barnase is known to hydrolyze poly A or poly G efficiently [17]. Some of the results are shown in Fig. 2: poly A was used as a substrate, and five mutants (100 pmol) were loaded on the gel. A module mutant, M2354, and three secondary structure unit mutants, S2435, S2543, and S4523, gave distinct bands at 4°C (Fig. 2A). These bands are regions where poly A was cleaved and eliminated from the stained gel. They coincided well with the corresponding protein staining bands (Fig. 2C), indicating that the RNase activity of the four mutants is not due to contamination with other RNases, but the mutants possess real hydrolytic activities toward poly A. M5342 (Fig. 2A,B) and 40 other mutants gave no band on an activity staining gel under the same conditions. Similar results were obtained when poly G was used as a substrate.

Only M2354 gave a clear band when the reaction was performed at 25°C (Fig. 2B). This result indicates that S2435, S2543, and S4523 are more active at lower temperatures. This result is in contrast to the case of three peptides corresponding to M2, M3, and M6 of wild-type barnase which hydrolyzed RNA at higher temperatures (>55°C) as simple acid-base catalysts [14], suggesting that the activities of the mutants depend on their conformational stability, as in the case of natural enzymes. As little as 0.01 pmol of wild-type barnase gave a clear band at 25°C.

3.2. Characterization of hydrolytic activity of module mutants and secondary structure unit mutants

Because hydrolysis of poly eAp can be measured in terms of the increment of fluorescence intensity, poly eAp is convenient for the measurement of the weak hydrolytic activities of the mutants. Furthermore, wild-type barnase can hydrolyze poly eAp [18]. Therefore we used poly eAp as a substrate for further characterization of the hydrolytic activities of the four mutants.

The activities of the four mutants were measured in the pH range of 5–9 at 5°C. All of them were found to hydrolyze poly eAp, and each mutant showed a distinct bell-shaped curve and optimum pH (Table 1). This suggests that these mutants have different active residues or active residues placed in different environments, although the active sites of the mutants have not been determined. The optimum pH of wild-type barnase was 7.5, which coincided with that reported by Nishimura et al. [19] for RNA as a substrate.

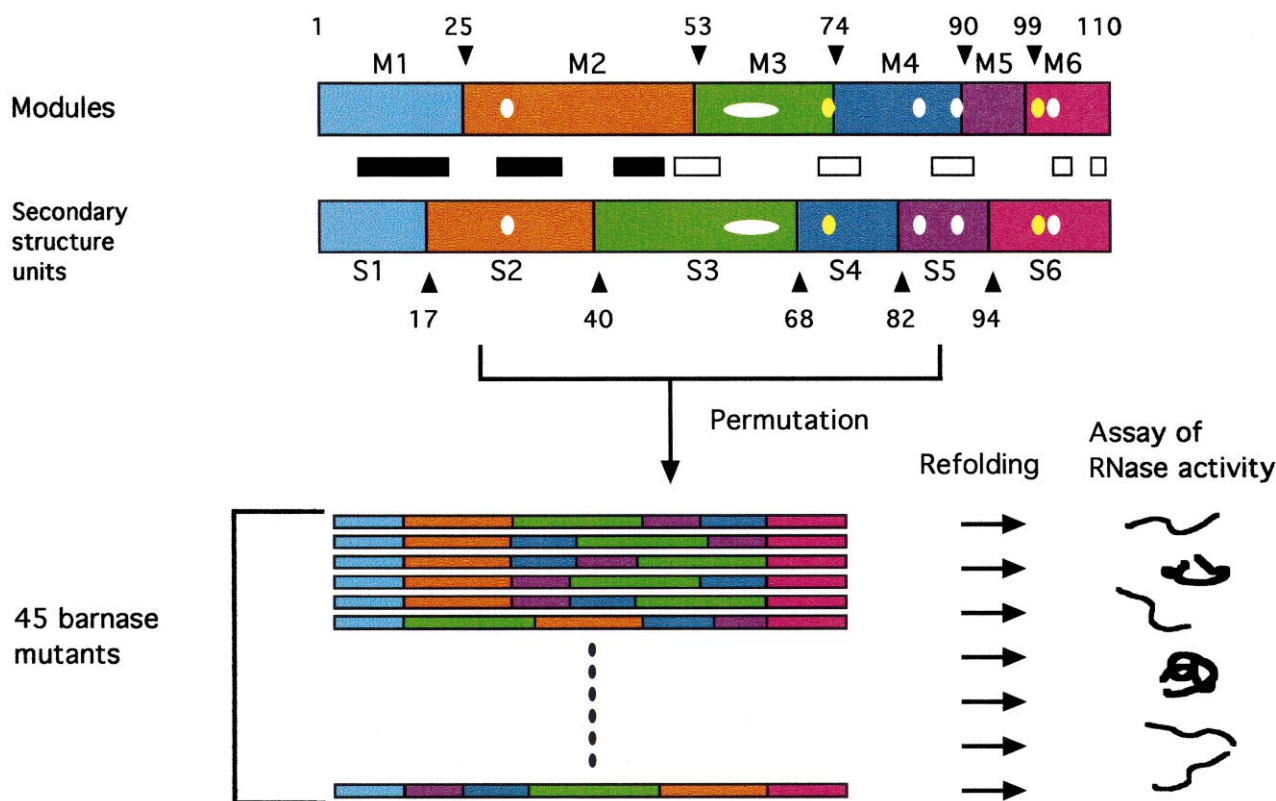Fig. 3 shows the effect of temperature on the hydrolytic

Fig. 1. Modules and secondary structure units of wild-type barnase and their permutations. The general acid and base residues of the enzyme, His-102 and Glu-73, are shown by yellow circles [8]. Amino acids organizing the active site of the enzymes, Lys-27, Arg-83, 87, Tyr-103, and Phe-56 to Leu-63 in a loop, are shown by white circles [8,9,21]. Black and white boxes indicate α-helices and β-sheets, respectively [21].
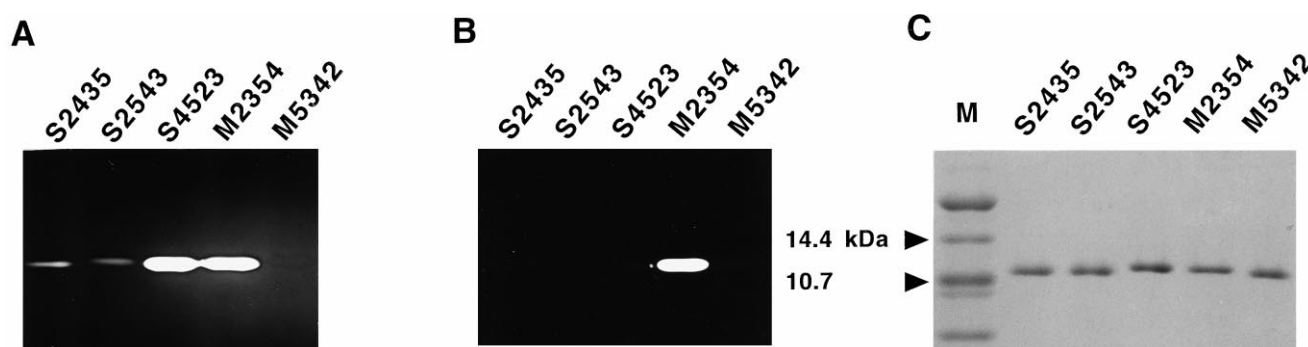
Fig. 2. SDS-PAGE of barnase mutants. Activity staining performed at 4°C (A) and 25°C (B), and Coomassie brilliant blue staining (C). Lane M, molecular weight markers. Bands of the mutants were positioned between 10.7 and 14.4 kDa, which is consistent with the molecular weight of wild-type barnase (12 382 Da) [7].

activities of the four mutants. The results indicate that each mutant has a definite optimum temperature, and the activity of the mutants dropped sharply above the optimum temperature. These cooperative denaturation profiles are similar to those of general enzymes, so the profiles support the idea that these mutants are real enzymes, not simple chemical catalysts. The activity of wild-type barnase dropped sharply at 47°C and above (data not shown). This temperature at which wild-type barnase became unstable coincided well with previous findings on the heat stability of activity [19] and conformation [12,20] of the enzyme. Thus the low melting temperatures of the mutants are not experimental errors, but appear to reflect unique properties of the mutants. Furthermore, the denaturation profiles can account for the results of activity staining experiments in which only M2354 hydrolyzed poly A or poly G at 25°C (Fig. 2). The conformation around the putative active site of M2354 seems to be more stable than that of the other mutants.

The hydrolytic activity of the mutants toward poly eAp at optimum temperature was compared to that of wild-type barnase at 25°C (Table 1). Although S4523 has an amino acid sequence with the lowest homology to that of wild-type barnase, the activity of this mutant is the highest of all the mutants, indicating that it is very difficult to predict the relative activity of mutants constructed by drastic mutation including some active residues. The activities of the mutants were 1/500 to 1/3000 that of wild-type barnase under the conditions used. These values are consistent with the value required for the basal enzymatic activity estimated by Axe et al. [21]. The relative activity of M2354 was lower than that expected from the activity staining bands (Fig. 2). This may be due to unknown factors affecting the refolding process of the mutant in the gel.

The activities of S2543 and S4523 were inhibited by addition of 2′ (3′)-GMP, like that of wild-type barnase (data not shown [22]), but those of S2435 and M2354 were not inhibited

by addition of 2′ (3′)-GMP under any of the conditions examined. Other nucleotides, 2′ (3′)-AMP, 2′ (3′)-UMP, and 2′ (3′)-CMP, did not inhibit the activities of the four mutants or wild-type barnase (data not shown). This indicates that S2543 and S4523 can recognize such small substances, like wild-type barnase. However, catalytic activities of the mutants were not detectable with dinucleotides, and K_m and k_{cat} values could not be obtained. This is in line with our expectation, because the activity of E73A, a mutant of barnase which has 0.2% of the RNA hydrolytic activity of the wild-type enzyme, was not detectable with dinucleotides [10].

3.3. Structural properties of active barnase mutants

Fig. 4A,B shows far- and near-UV CD spectra of S2543 at different temperatures. Broad negative bands over 208–230 nm were observed in the far-UV CD spectrum measured at 5°C, but not 25°C. Similarly, distinct positive bands were observed in the near-UV CD spectrum measured at 5°C, but not at 25°C. These spectra indicate that secondary and tertiary structures of S2543 formed at lower temperatures unfolded at 25°C. The thermal unfolding profiles of secondary and tertiary structures of S2543 monitored in terms of CD intensity at 228 nm and 280 nm indicate that the structures unfolded rapidly above 10°C and completely at 25°C (Fig. 4C,D). The tendency of these thermal unfolding profiles of S2543 was similar to that of temperature dependence of its hydrolytic activity (Figs. 2B and 3).

The previous study showed that S2435, S4523, and M2354 had no ordered backbone conformation or tertiary structures even at 5°C [16]. However, the thermal denaturation profiles shown in Fig. 3 suggest that these mutants have catalytic activities which depend on the stability of the conformation around the putative active sites. Therefore, these mutants may have active sites with local structures which are not substantial enough to be detected by macroscopic analysis, such as CD measurement, or the proper conformations around active

Table 1
Summary of hydrolytic activity of barnase mutants toward poly eAp

Mutant	Sequence homology (%)	Optimum pH	Optimum temperature (°C)	Relative activity[a]
S2435	65	6.5	7	4×10^{-4}
S3543	57	6.0	10	3×10^{-4}
S4523	35	7.5	8	2×10^{-3}
M2354	77	7.0	27	5×10^{-4}
Wild-type	100	7.5	47	1

[a]Activity of barnase mutants at their optimum temperature was compared to that of wild-type barnase at 25°C.
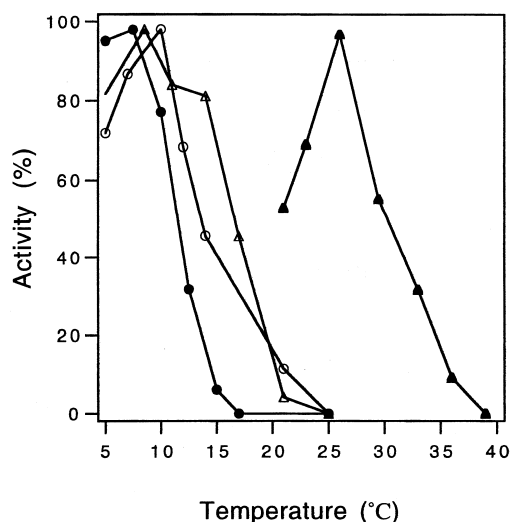
Fig. 3. Effect of temperature on hydrolytic activity of barnase mutants toward polymers of poly eAp. The activity of each mutant at its optimum temperature is defined as 100% activity. ● S2435, ○ S2543, △ S4523, ▲ M2354.

sites may be generated only when the substrate binds to them, namely an induced folding mechanism [23–25].

In the latter case, because S4523 can interact with 2′ (3′)-GMP as mentioned above, it is possible that binding of 2′ (3′)-GMP to the mutant induces conformational changes of the mutant. Fig. 5A shows the far-UV CD spectra of S4523 in

the absence (dashed line) and presence (solid line) of 2′ (3′)-GMP at 5°C. The spectrum in the absence of 2′ (3′)-GMP shows a minimum in the vicinity of 200 nm, indicating that S4523 is unfolded even at low temperature. However, in the presence of 2′ (3′)-GMP, the CD intensity at 200 nm drastically decreased, and the minimum in the vicinity of 230 nm became clearer, indicating that conformational change was induced by 2′ (3′)-GMP. The minimum at 230 nm is seen in the far-UV CD spectra of wild-type barnase and S2543 [16], and it is known that Trp-94 contributes to it in the case of wild-type barnase [26]. Therefore, it is plausible that the conformation around Trp-94 of S4523 is structured by the binding of 2′ (3′)-GMP to the mutant. The presence of a minimum at 230 nm in the far-UV CD spectrum does not necessarily mean the formation of an active site, because M3245 having such a minimum does not have RNase activity [16]. A similar effect of addition of 2′ (3′)-GMP was also observed in the near-UV CD spectrum of S4523 (Fig. 5B). Thus, 2′ (3′)-GMP probably induced ordered backbone and tertiary structures of S4523. The presence of 2′ (3′)-GMP did not change the far- and near-UV CD spectra of other mutants such as M5342, confirming that the results described above are not experimental errors (Fig. 5C,D).

Although the binding of 3′-GMP to wild-type barnase induces a change of local conformation around the active site, the overall conformation is not affected by the binding [27]. This is probably due to the stable and rigid structure of the enzyme. On the other hand, because S4523 does not have such a stable conformation, the overall structure of the mutant
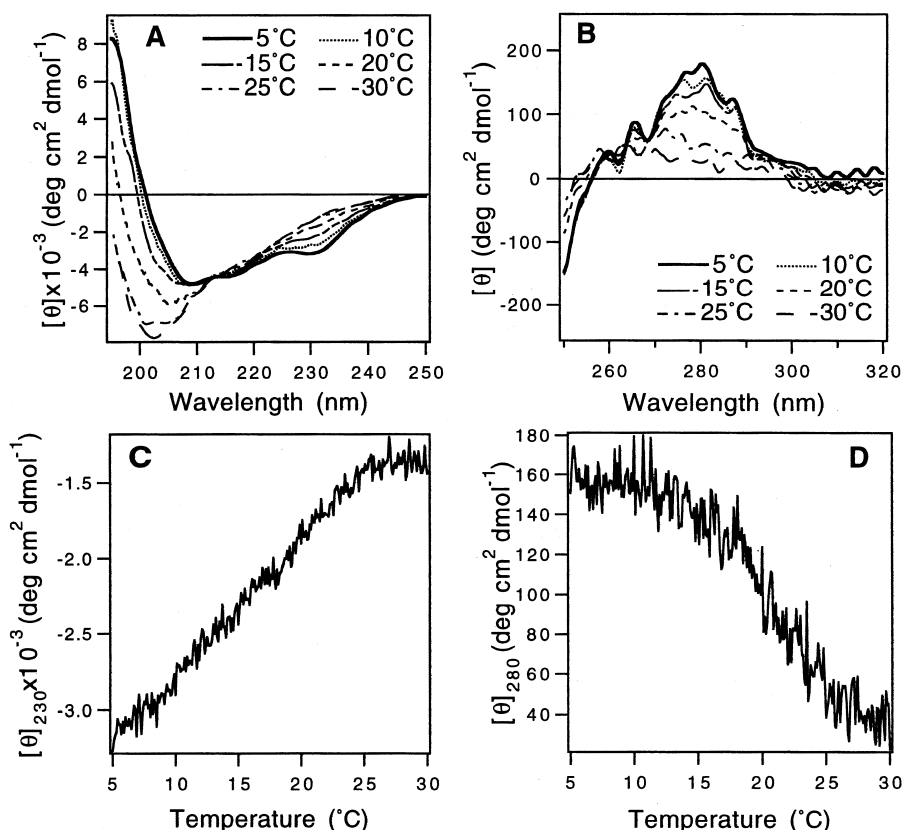

Fig. 4. Thermal unfolding of secondary and tertiary structures of S2543 monitored by far- and near-UV CD spectroscopy. Far-UV CD spectra (A) and near-UV CD spectra (B) of S2543 at different temperatures. The thermal unfolding profiles of secondary (C) and tertiary (D) structures of S2543 were monitored by measuring the CD intensity at 230 nm and 280 nm.
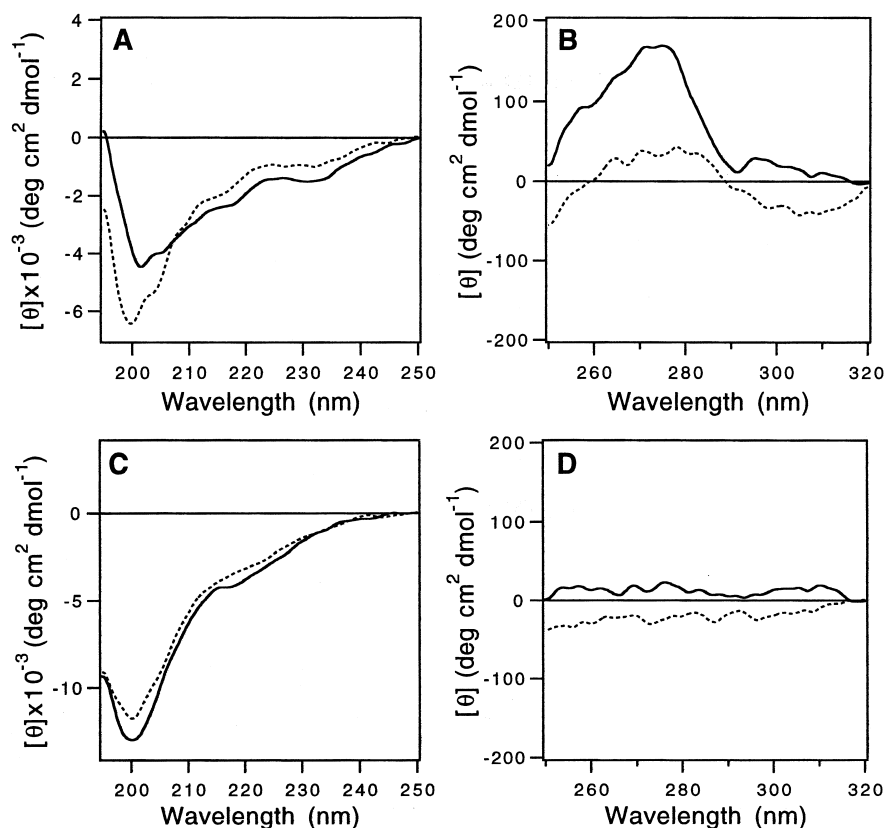
Fig. 5. Effect of 2′ (3′)-GMP on the CD spectra of S4523 and M5342. Far-UV CD (A) and near-UV CD (B) spectra of S4523 (3 μM) in the presence (solid line) and absence (dashed line) of 2′ (3′)-GMP (15 μM). Far-UV CD (C) and near-UV CD (D) spectra of M5342 (3 μM) in the presence (solid line) and absence (dashed line) of 2′ (3′)-GMP (15 μM). All CD spectra were obtained at 5°C.

would be easily affected by the conformational change of the active site. The near-UV CD spectrum of S2543, whose activity is inhibited by 2′ (3′)-GMP, changed in the presence of 2′ (3′)-GMP, but the far-UV CD spectrum did not (data not shown), suggesting that the backbone conformation is more stable than the tertiary structure in the mutant.

As chemical reactions proceed efficiently at higher temperatures, enzymes having low melting temperatures must have immature catalytic activities. Probably amino acids not involved directly in catalysis have been selected as scaffolds to generate active sites having appropriate conformation and thermal stability in the course of globular protein evolution [28]. The barnase mutants which have active sites but not such scaffolds can be considered globular proteins in 'evolutionarily intermediate states'.

Although S2543 has the most ordered conformation of all the mutants obtained [16], its activity is the lowest among the active mutants. Active mutants which have disordered conformations have the possibility to evolve, if they acquire appropriate scaffolds in a further evolutionary process, namely directed evolution. On the other hand, S2543 may not readily evolve, because this mutant already has a rigid scaffold to some extent. The three active mutants having a disordered conformation may therefore be located at the foot of a high mountain, while the weak active mutant having an ordered conformation may be located near the top of a low mountain in the fitness landscape. This is similar to the situation suggested by theoretical studies, in which kinetic intermediates trapped at local minima have to break non-corrective bonds

to attain a global minimum on the energy landscape in protein folding reactions [29].

This study indicates that the amino acid sequence of wild-type barnase has sufficient plasticity to generate new RNases upon permutation of its modules or secondary structure units. Therefore, it is possible that many kinds of immature proteins with different functions can be created by the combination of peptide fragments derived from different natural proteins, namely from a semi-global sequence space. These proteins might evolve to mature proteins by further directed evolution.

References

[1] Ellington, A.D. and Szostak, J.W. (1990) Nature 346, 818–822.
[2] Ellington, A.D. and Szostak, J.W. (1992) Nature 355, 850–852.
[3] Yamauchi, A., Yomo, T., Tanaka, F., Prijambada, I.D., Ohhashi, S., Yamamoto, K., Shima, Y., Ogasahara, K., Yutani, K., Kataoka, M. and Urabe, I. (1998) FEBS Lett. 421, 147–151.
[4] Prijambada, I.D., Yomo, T., Tanaka, F., Kawama, T., Yamamoto, K., Hasegawa, A., Shima, Y., Negoro, S. and Urabe, I. (1996) FEBS Lett. 382, 21–25.
[5] Crameri, A., Raillard, S.-A., Bermudez, E. and Stemmer, W.P.C. (1998) Nature 391, 288–291.
[6] Zhao, H., Giver, L., Shao, Z., Affholter, J.A. and Arnold, F.H. (1998) Nature Biotechnol. 16, 258–261.

 [7] Nemoto, N., Miyamoto-Sato, E., Hushimi, Y. and Yanagawa, H. (1997) FEBS Lett. 414, 405–408.

 [8] Robert, R.W. and Szostak, J.W. (1997) Proc. Natl. Acad. Sci. USA 94, 12297–12302.

 [9] Hartley, R.W. (1988) J. Mol. Biol. 202, 913–915.

[10] Mossakowska, D.E., Nyberg, K. and Fersht, A.R. (1989) Biochemistry 28, 3843–3850.

[11] Meiering, E.M., Serrano, L. and Fersht, A.R. (1992) J. Mol. Biol. 225, 585–589.

[12] Neira, J.L. and Fersht, A.R. (1999) J. Mol. Biol. 285, 1309–1333.

[13] Yanagawa, H., Yoshida, K., Torigoe, C., Park, J.S., Sato, K., Shirai, T. and Go, M. (1993) J. Biol. Chem. 268, 5861–5865.

[14] Yoshida, K., Shibata, T., Masai, J., Sato, K., Noguti, T., Go, M. and Yanagawa, H. (1993) Biochemistry 32, 2162–2166.

[15] Noguti, T., Sakakibara, H. and Go, M. (1993) Proteins 16, 357–363.

[16] Tsuji, T., Yoshida, K., Satoh, A., Kohno, T., Kobayashi, K. and Yanagawa, H. (1999) J. Mol. Biol. 286, 1581–1596.

[17] Fitzgerald, P.C. and Hartley, R.W. (1993) Anal. Biochem. 214, 544–547.

[18] Whitfeld, P.R. and Witzel, H. (1963) Biochim. Biophys. Acta 72, 362–371.

[19] Nishimura, S. and Nomura, M. (1959) J. Biochem. 46, 161–167.

[20] Hartley, R.W. (1968) Biochemistry 7, 2401–2408.

[21] Axe, D.D., Foster, N.W. and Fersht, A.R. (1998) Biochemistry 37, 7157–7166.

[22] Guillet, V., Lapthorn, A. and Mauguen, Y. (1993) FEBS Lett. 330, 137–140.

[23] He, Z., Dunker, A.K., Wesson, C.R. and Trumble, W.R. (1993) J. Biol. Chem. 268, 24635–24641.

[24] Daughdrill, G.W., Chadsey, M.S., Karlinsey, J.E., Hughes, K.T. and Dahlquist, F.W. (1997) Nature Struct. Biol. 4, 285–291.

[25] Tarkka, T., Oikarinen, J. and Grundstrom, T. (1997) FEBS Lett. 406, 56–60.

[26] Vuilleumier, S., Sancho, J., Loewenthal, R. and Fersht, A.R. (1993) Biochemistry 32, 10303–10313.

[27] Meiering, E.M., Bycroft, M., Lubienski, M.J. and Fersht, A.R. (1993) Biochemistry 32, 10975–10987.

[28] Saito, S., Sasai, M. and Yomo, T. (1997) Proc. Natl. Acad. Sci. USA 94, 11324–11328.

[29] Dill, K.A. and Chan, H.S. (1997) Nature Struct. Biol. 4, 10–19.