Complementary Nucleobase Interaction Enhances Peptide – Peptide Recognition and Self-Replicating Catalysis

Sachiko Matsumura, Tsuyoshi Takahashi, Akihiko Ueno,† and Hisakazu Mihara*[a]

This paper is dedicated to the memory of Prof. Akihiko Ueno, who died on March 23, 2003.

Abstract: The availability of the complementary interaction of nucleobases for influencing the formation of peptide architectures was explored. Nucleobases were incorporated as additional recognition elements in coiled-coil peptides by employing nucleobase amino acids (NBAs), which are artificial L-α-amino γ-nucleobase-butyric acids. The effect of the base-pair interaction on intermolecular recognition between peptides was evaluated through a self-replication reaction. The self-replication reactions of the peptides with complementary base pairs such as thymine – adenine or guanine – cytosine at the $g-g'$ heptad positions were accelerated in comparison with those of the peptides with mismatched base pairs or without nucleobases. However, thymine – adenine pairs at the $e-e'$ positions did not enhance the self-replication. In the presence of a denaturant, the enhancement effects of complementary base pairs on the reaction disappeared. Thermal denaturation studies showed that the thymine – adenine pairs contributed to stabilization of the coiled-coil structure and that the pairs at the $g-g'$ positions were more effective than those at the $e-e'$ positions. The peptide – peptide interaction was reinforced by complementary nucleobase interactions appropriately arranged in the peptide structure; these led to acceleration of the self-replication reactions.

Keywords: coiled-coil structures · complementary interactions · nucleobases · peptides · self-replication

Introduction

Complementary interactions are required to recognize and assemble molecules for creation of a specific structure or function and are essential for various chemical and biological processes in life systems. Natural proteins utilize various interactions including hydrophobic, electrostatic, and hydrogen-bonding interactions for complementary recognition. Interactions in proteins occur through a variety of amino acid functions and their arrangements in the highly ordered structure.

Advances in de novo peptide and protein design have made possible the construction of various structures, including native-like structures of proteins and those with nonnatural components.[1–10] De novo designed peptides have been demonstrated to provide a simple model for investigating sequence – structure and/or structure – function relationships and to create a sophisticated artificial system that resembles the natural system. However, at present, peptide de novo design cannot completely reach the elegance of nature. In many cases of de novo peptide architecture, the development of complementary interactions has not been successfully carried out for the creation of structural and functional specificity.

On the other hand, DNA and RNA can form complementary base pairs through hydrogen bonds and can achieve the specific recognition necessary for the ensuing genetic information transfer. The simple and elegant strategy of base pairing has been widely applied to molecular design, not only in biology but also in engineering, for example, in DNA-directed nanoscaled assemblies.[11] In the field of chemical biology, various peptide – nucleobase conjugates have been developed.[12–16] The most famous one is a peptide nucleic acid (PNA), which has a high ability to hybridize strongly and specifically with DNA, RNA, and PNA.[12] Although it contains an amide bond in the skeleton, PNA is not regarded as a peptide that constructs a protein-like tertiary structure. Unlike simple analogues of DNA, peptide – nucleobase conjugates are regarded as material with the structural characteristics of peptides and the complementary recognition of nucleic acids. By using secondary-structured peptides in which nucleobases were arranged, DNA and RNA recognition were demonstrated.[14, 15] However, most peptide – nucleobase conjugates, including these examples, are directed at targeting DNA and RNA.

[a] Prof. H. Mihara, Dr. S. Matsumura, Dr. T. Takahashi, Prof. A. Ueno†
Department of Bioengineering
Graduate School of Bioscience and Biotechnology
Tokyo Institute of Technology
4259 Nagatsuta-cho, Midori-ku, Yokohama 226-8501 (Japan)
Fax: (+81) 45-924-5833

 DOI: 10.1002/chem.200305121

Employing the base-pair interaction in de novo peptide architecture is promising for the creation of novel structures and functions equipped with manipulatable and/or multifunctional abilities. In this study, we have attempted to employ a nucleobase interaction as a functional factor for a peptide–peptide interaction. In order to evaluate the utility of the complementary interaction of nucleobases between peptides, we selected a self-replication system (Figure 1), because the

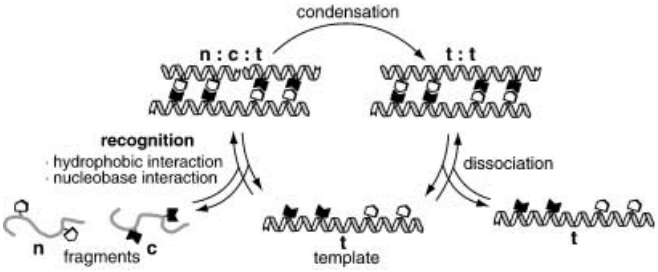

Figure 1. Schematic illustration of the self-replication system for the nucleobase-combined peptides. The template (**t**) is assumed to assemble with N- and C-terminal fragments (**n** and **c**) in the antiparallel orientation.

effects of the nucleobase interaction can be amplified through the replication reaction. Self-replication systems are chemical systems capable of templating and catalyzing their own synthesis, and complementary recognition is necessary for the catalytic reactions.[17] Throughout these systems, information is amplified by product formation. Many self-replicating molecular systems have been studied from the viewpoint of the nature of living systems,[18–20] and self-replicating peptides have also been developed.[21, 22] These peptide systems are based on coiled-coil structures, and hydrophobic interactions and adjacent charged residues mainly act for the molecular recognition. On the basis of the coiled-coil structure, additional incorporation of nucleobases is expected to encourage the structural complementarity. If a self-replicating peptide can obtain complementary recognition ability like that of DNA and RNA, the peptide, which is equipped with additional information, will greatly expand the possibilities of molecular evolution.

To combine nucleobases with a peptide structure, we used artificial L-α-amino γ-nucleobase-butyric acids (nucleobase amino acids, NBAs; Figure 2b). NBAs would assuredly not interrupt the secondary structure of the peptide and can be incorporated at any position in a peptide sequence by conventional solid-phase peptide synthesis.[15, 16] We demonstrated that peptides containing NBAs can take an α-helix form or a zinc-finger structure and that they can interact with the target RNA specifically.[15] In another case with a short two-α-helix peptide containing NBAs, the structure was stabilized with the base-pair interaction.[16] Here, we incorporate NBAs as additional recognition units into a coiled-coil structure, and evaluate the base-pair interaction for intermolecular peptide–peptide recognition by using a peptide self-replication system. Several peptide systems were constructed that varied in the NBA species or positions. We describe how complementary base pairs incorporated as additional recognition elements perform peptide–peptide recognition and lead to acceleration of the replicating catalysis.

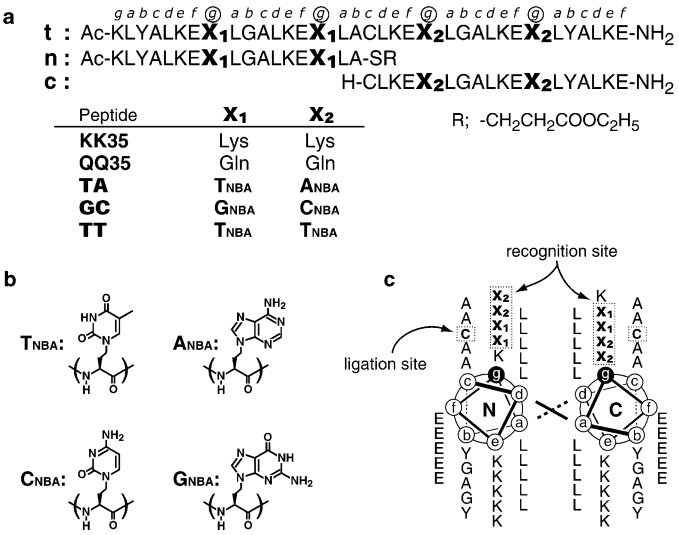

Figure 2. Designed structures of 35-residue peptide systems. a) Amino acid sequences of the template and fragment peptides. KK35 is the same peptide as K1K2 reported by Yao and co-workers.[22a] b) Chemical structures of the thymine (T_{NBA}), adenine (A_{NBA}), cytosine (C_{NBA}), and guanine (G_{NBA}) residues within the peptide chain. c) A helix wheel drawing of the template (product) peptide in the dimeric antiparallel coiled-coil form.

Results

Design of 35-residue peptide replication systems: In order to examine the nucleobase interaction in peptide–peptide recognition, we used the self-replication peptides reported by Yao and co-workers as a first model system; this system was composed of a 35-residue template peptide (K1K2) and the fragment peptides (K1 and K2).[22a] These peptides were designed to form a coiled-coil structure. The amino acid sequences of coiled-coil structures are characterized by a seven-residue periodicity (heptad repeat) denoted as $(abcdefg)_n$. The *a* and *d* positions are usually occupied by hydrophobic residues and form a hydrophobic interface required for helix–helix interaction. The *e* and g positions flanking the hydrophobic face are often charged residues that participate in interhelical interaction. K1K2 was composed of five heptad repeats with Leu residues at the *a* and *d* positions and Lys residues at the *e* and *g* positions. The N-terminal fragment K1 was activated at its C terminus as a thiol ester, and the C-terminal fragment K2 had a Cys residue at its N terminus. The coupling reaction of these fragments produced the template K1K2 by the native chemical ligation method.[23] The Lys residues at the *e* and g positions in K1K2 were reported to adjust the coiled-coil formation of template and fragment peptides.[22a] Since these peptide sequences were simple and would be suitable to evaluate an advantage of NBAs, we employed this peptide system as a reference. The K1K2 template is denoted as KK35-t in this study, and the N- and C-terminal fragments K1 and K2 are denoted as KK35-n and KK35-c, respectively (Figure 2).

For the evaluation of the advantages of NBA units on peptide–peptide recognition, NBAs were incorporated at the *g* and *g*′ positions of this system. Since the Cα–Cβ vectors of

the pair of amino acids at the $g-g'$ positions are directed toward each other in an antiparallel coiled-coil dimer, molecular modeling suggests that hydrogen bonds can be formed between nucleobases. The nucleobases at these positions are expected to work as interhelical recognition elements through specific base-pair interaction. Indeed, incorporation of complementary base pairs at the $g-g'$ positions in antiparallel two-α-helix peptides contributed to α-helix formation and stabilization.[16] Hence, three types of NBA-incorporated peptide system were designed (Figure 2). Two systems contain complementary base pairs of thymine – adenine (the TA system) or guanine – cytosine (the GC system), and another has mismatched base pairs of thymine – thymine (the TT system). Considering the effects of charged residues at the g and g' positions, we also designed the QQ35 system, in which the NBAs were replaced with Gln residues. In the 35-residue systems, since each template contains four bases and each of the two fragments has two bases, two by two base-pair formation is expected between one template and each of the two fragments.

Design of 28-residue peptide replication systems: To clarify the positional preference of NBAs for their interaction between peptides, 28-residue peptide systems with NBA pairs at the $g-g'$ positions or the $e-e'$ positions were designed (Figure 3). As a basic template peptide, 28-residue peptide, QQ28, was designed. QQ28 was composed of four heptad repeats with Gln residues at both the e and g positions to allow easy comparison upon substitution with NBA. The length was shortened relative to the above systems; the aim of this was to avoid stronger peptide association. The N- and C-terminal fragments are denoted as QQ-n and QQ-c, respectively. Two systems of peptides incorporating NBAs were designed. The eTeA template is designed to form two thymine – adenine pairs at the $e-e'$ positions in an antiparallel fashion, and the N- and C-terminal fragments are denoted as eT-n and eA-c, respectively. Similarly, the gTgA template is designed to have two thymine – adenine pairs at the $g-g'$ positions, and the N- and C-terminal fragments are denoted as gT-n and gA-c, respectively. Since each template peptide contains two bases and each of the two fragments has one base, one by one base-pair formation is expected between one template and each of the two fragments.

Furthermore, in order to examine whether base pairs can work in either parallel or antiparallel orientation, the gAgT system was also designed. The difference between gAgT and gTgA is the alignment of nucleobases, that is, the order of NBAs from N to C terminus is adenine and thymine in gAgT and thymine and adenine in gTgA. gAgT was designed to associate with eTeA into a parallel coiled-coil structure with two thymine – adenine pairs, that is, gAgT and eTeA would make a heteroassembly. Of course, gAgT can also associate with itself in an antiparallel coiled-coil structure (Figure 3b). The gAgT system includes the corresponding two fragment peptides, gA-n as the N-terminal fragment and gT-c as the C-terminal one.

Self-replication of 35-residue peptide systems: The self-replication reactions of 35-residue systems were carried out to estimate the influence of pairs of nucleobases at the $g-g'$ positions on the peptide – peptide interaction. In all the designed peptide systems, the ligation reaction of the corresponding two fragments (100 μM each) produced product identical to the template peptide. The rate was dependent on the added initial template concentration (0 – 20 μM; Figure 4a); this indicates the self-replicating reaction.[21, 22] The product formation ability of each peptide system was compared (Figure 4b). In the presence of 10 μM concentrations of the initial template, after 9 h, the product formation of the KK35 system (16 μM) was almost equal to that of the QQ35 system, a fact indicating that the effect of charged residues at the g and g' positions was not significant in the conditions used in our study. Compared to the peptides without NBAs, a significant reaction enhancement was observed in the TA and GC systems with their complementary base pairs. The TA system showed a 3.5-fold increase in production over the KK35 system, and, furthermore, the increase for the GC system relative to KK35 was

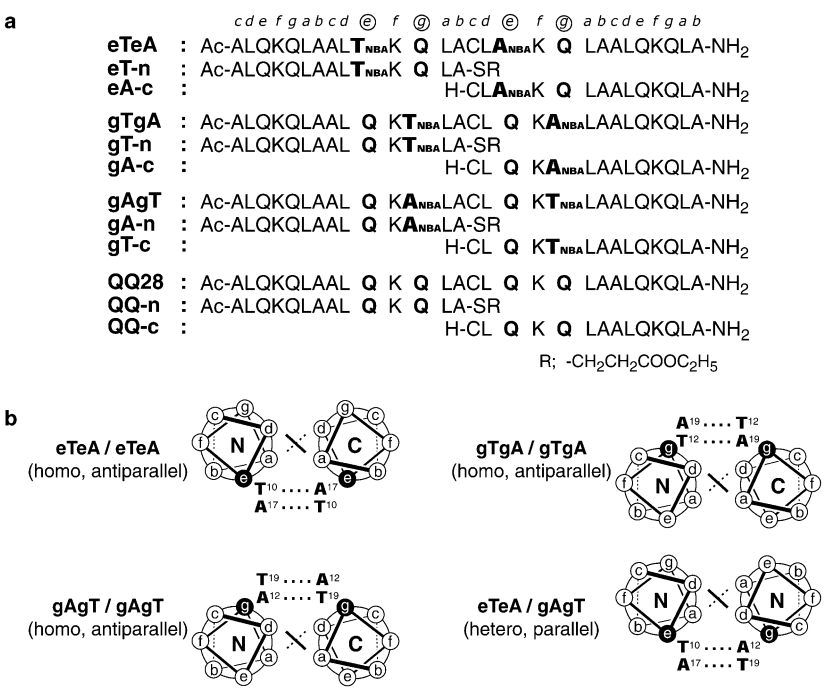

Figure 3. Designed structures of 28-residue peptide systems. a) Amino acid sequences of the template and fragment peptides. b) Helix wheel drawings of the template (product) peptide in the dimeric coiled-coil form. Prospective complementary nucleobase pairs are shown in three species of homodimers in an antiparallel orientation and one heterodimer in a parallel orientation.

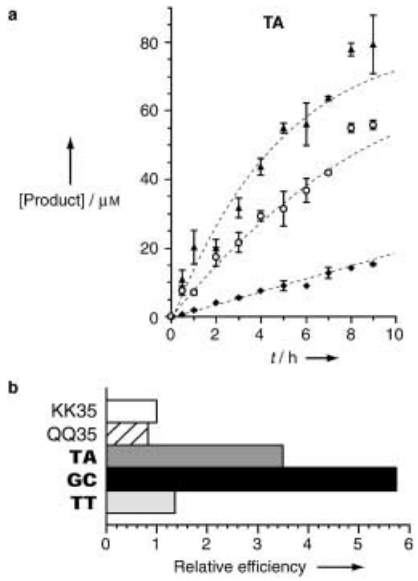

Figure 4. Self-replication reaction of 35-residue peptide systems. a) Product formation of the TA system is dependent on the initial template concentrations. The reactions were performed in buffer with 100 μM concentrations of each of the N- and C-terminal fragments and with the template peptide at a concentration of 0 μM (◆), 10 μM (○), or 20 μM (▲) (see Experimental Section for details). b) Relative efficiency of the reactions with 10 μM template peptide in different peptide systems after 9 h, as compared with the KK35 system.

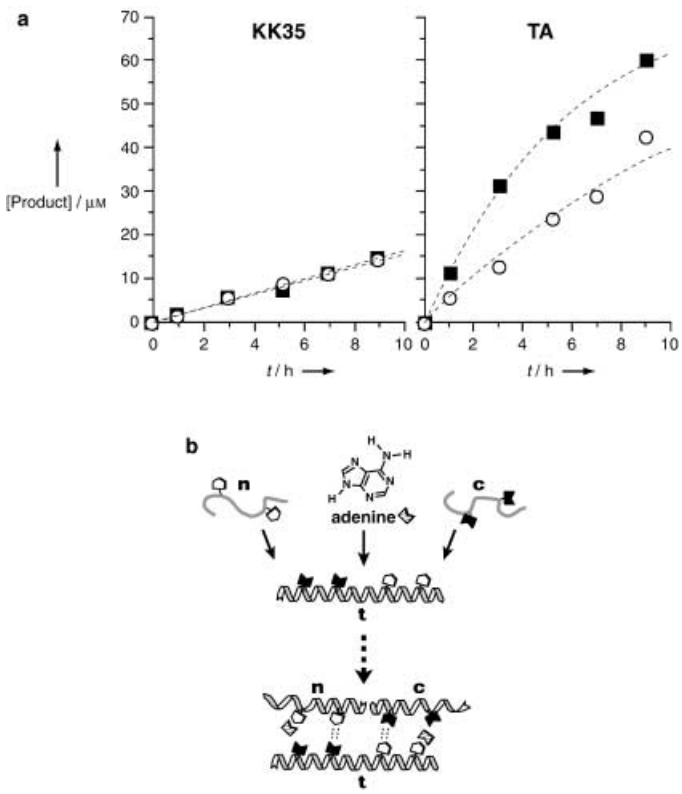

Figure 5. Inhibition of the self-replication reaction by addition of adenine. a) Product formation for reaction mixtures initially containing 10 μM template peptide in the absence (■) and presence (○) of 0.5 mM adenine. b) Illustration of inhibition of coiled-coil formation by added adenine.

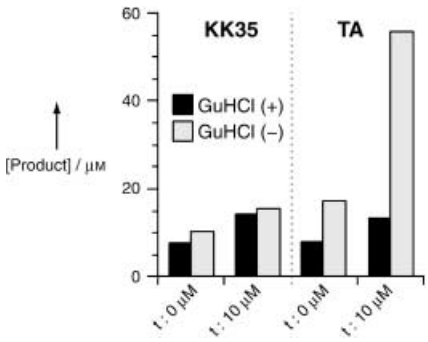

Figure 6. Decrease in product formation induced by denaturation of the coiled-coil structure by addition of GuHCl. Product formation was measured after 9 h in the absence (gray bar) and presence of 4.4 M GuHCl (black bar) and with and without 10 μM template peptide (t).

5.8 times. These results imply that the thymine – adenine and guanine – cytosine interactions contribute to the assembly of the template and fragments into the coiled-coil formation, thereby enhancing the replication reaction. The base-pair interactions seemed to contain hydrogen-bonding interactions and hydrophobic interactions. The TT system (with mismatched base pairs) showed a production increase of only 1.4 times over the KK35 system. This indicated that nonspecific base-pair interactions, such as hydrophobic interactions, did not have large influences on the enhancement of the TA system.

In the presence of exogenous free adenine, the self-replicating reaction in the TA system was inhibited to some extent. Addition of excess adenine (0.5 mM) into the reaction solution with 10 μM concentrations of template reduced product formation by at least 30% relative to the reaction without adenine (Figure 5). This result suggests that the interaction between thymine and adenine in the coiled-coil structure is disturbed by exogenous adenine; the monomeric base did not influence the reaction of KK35. Thus, the base-pair interaction was suggested for recognition between the peptides.

The importance of the coiled-coil structure with the nucleobase interaction to the reaction was ascertained by ligation reactions in the presence of a denaturant (Figure 6). In the TA system with 10 μM template, the product formation after 9 h (56 μM) was reduced to 13 μM in the reaction in the presence of 4.4 M guanidine hydrochloride (GuHCl). In the absence of the template, added GuHCl also reduced the product formation, from 17 to 8 μM. In the presence of 4.4 M GuHCl, the product formation of the TA system was almost equivalent to that of the KK35 system, in spite of the difference of peptides with or without NBAs. These results confirmed that template-assisted preorganization of reactant peptides derived from the coiled-coil structure was required for efficient catalysis.

Self-replication of 28-residue peptide systems: The aim of these 28-residue systems was to investigate the position of base pairs appropriate for effective interaction between peptide strands (Figure 3). At first, we compared the effects of pairs of NBAs at the $g-g'$ or $e-e'$ positions. The gTgA system was designed to have one thymine – adenine pair at the $g-g'$ positions between the template and each of the fragments, thus two pairs would be formed between the template

and the product. Similarly, the eTeA system may form thymine – adenine pair(s) at the $e-e'$ positions. The gTgA, eTeA, and QQ28 (without NBAs) systems showed a self-replication profile, since template production was accelerated when larger amounts of the template peptide were initially added to the reaction mixture (Figure 7). In the gTgA system the effect of NBAs on product formation was clearly observed, even though the peptides had only two thymine – adenine pairs. Figure 7b shows the relative product formation of the systems after 8 h in the presence of 5 μM initial template. The gTgA system generated twice as much product as the QQ28 system did. In contrast, the product formation in the eTeA system was almost the same as that in the QQ28 system, although the eTeA system was designed to have thymine – adenine pairs at the $e-e'$ positions. As was consistent with the results of 35-residue systems, pairs of NBAs at the $g-g'$ positions could contribute to the reaction, probably by effective assembly of peptides due to the additional interaction between thymine and adenine. In the $e-e'$ positions the pairs of NBAs did not exhibit an effect on the replication reaction under the conditions used in this study.

Reactions in the presence of GuHCl confirmed the requirement of the coiled-coil structure (data not shown). In all systems where 6 M GuHCl was added, product formation was reduced by approximately 80%, in both the absence and presence of an initial template and regardless of the absence or presence and positions of NBAs. These results suggested that the coiled-coil structure was responsible for the catalytic reaction; they further suggested that the nucleobase interaction was associated with the positioning of nucleobases in the restricted structure and that the interaction would not occur in random (denatured) peptide chains.

The specificity of the template in the reaction was demonstrated. The production of gTgA by condensation of the fragments, gT-n and gA-c (75 μM each), was examined in the presence of 20 μM concentrations of inappropriate template, QQ28 or eTeA, compared with the same amount of appropriate gTgA template (Figure 8). The appropriate template gTgA clearly increased product formation more than eTeA or QQ28, although the reaction with these inappropriate templates showed a little increase over the case in the absence of template. Since the peptide self-

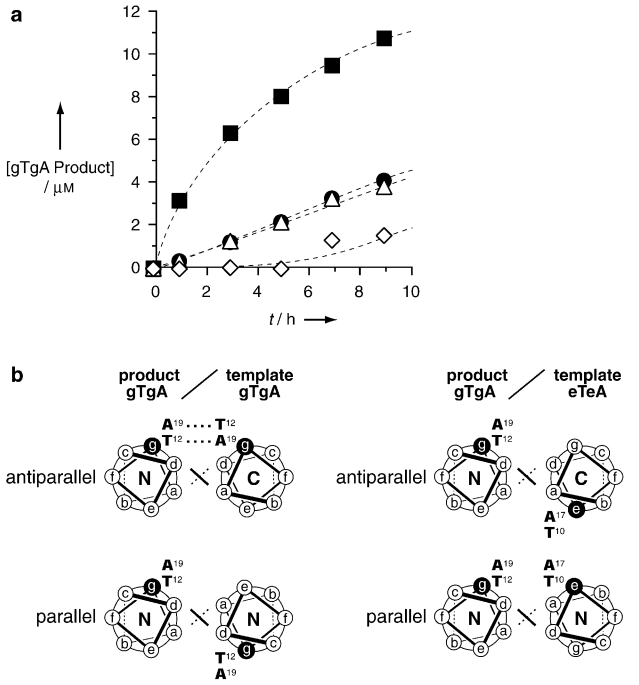

Figure 8. Formation of the gTgA product depending on the template species. a) The production of gTgA by ligation between gT-n and gA-c in the presence of template gTgA (■), eTeA (●), or QQ28 (△), or in the absence of template (◇). The reactions were performed in buffer with 75 μM concentraions of each of the N- and C-terminal fragments and with the template peptide at a concentration of 20 μM (see Experimental Section for details). b) Helix wheel drawings of potential coiled-coil dimers composed of the product gTgA and the template gTgA (left side) or the product gTgA and the template eTeA (right side) in the ligation reaction of gT-n and gA-c.

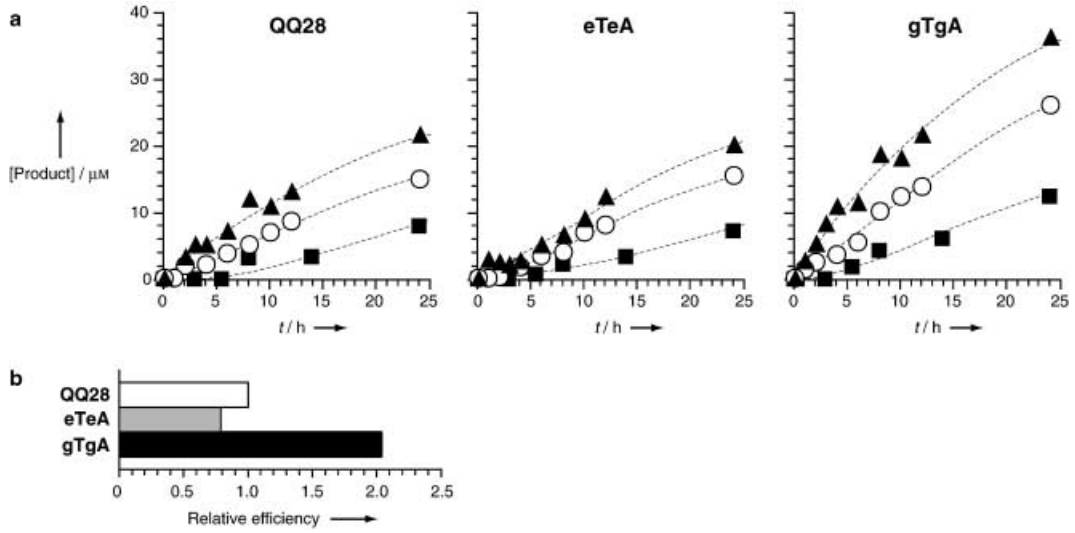

Figure 7. Self-replication reaction of the 28-residue peptide systems QQ28, eTeA, and gTgA. a) The reactions were performed in buffer with 100 μM concentrations of each of the N- and C-terminal fragments and with the template peptide at a concentration of 0 μM (■), 5 μM (○), and 10 μM (▲), (see Experimental Section for details). b) Relative efficiency of the reactions with 5 μM template peptide in different peptide systems after 8 h, as compared with the QQ28 system.

replication reaction was, in itself, based on recognition by the hydrophobic interface derived from the coiled-coil structure, it was reasonable that QQ28 had some template activity. It was noteworthy that the eTeA showed the same effect on the reaction as QQ28 did, in spite of the presence of NBAs. These results indicated that the NBAs at the *e* positions did not contribute to the reaction of fragment peptides with NBAs at the *g* positions. For the ligation of fragments with NBAs at the *g* position, the template eTeA is required to orient itself parallel to the fragments to afford $e-g$ and $e'-g'$ interactions, and in this case mismatched base pairs, adenine–adenine and thymine–thymine, are predicted (Figure 8b). In an antiparallel orientation, the nucleobase interaction is not expected between this template and the fragments. Considering that the ability of eTeA and QQ28 to influence the reaction was almost same, the antiparallel orientation was dominant, and the mismatched base pairs in a parallel orientation did not affect the reaction.

To examine the above supposition and determine whether the nucleobase interactions occur in a parallel or antiparallel orientation, we designed the gAgT system (Figure 3). As with the above systems, the gAgT system is expected to self-associate in an antiparallel coiled-coil form with thymine–adenine pairs. In fact, the gAgT system showed the self-replication and the reaction was superior to that of eTeA. Additionally, the new gAgT system may associate with eTeA through parallel coiled-coil formation with thymine–adenine pairs. In the production of gAgT by ligation of gA-n and gT-c, the effect of each template was estimated by using the gAgT template for the formation of base pairs in an antiparallel coiled-coil structure, the eTeA template for a parallel coiled-coil structure, and the QQ28 template for the reaction in the absence of nucleobases (Figure 9). The presence of any template increased product formation because of the hydrophobic interactions in the coiled-coil structure as described above. Among the three template species, gAgT acted as the most effective template, and the effect of eTeA was nearly identical to that of QQ28. These results again suggested that the nucleobase interaction in an antiparallel orientation was preferable.

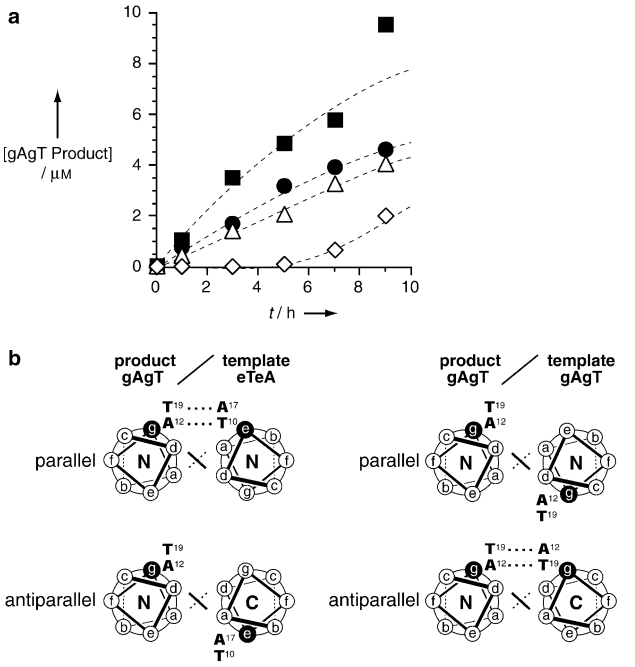

Figure 9. Formation of gAgT product depending on the template species and the helix strand orientation. a) Production of gAgT by ligation between gA-n and gT-c in the presence of the template gAgT (■), eTeA (●), or QQ28 (△), or in the absence of template (◇). The reaction conditions were the same as those in Figure 8. b) Helix wheel drawings of potential coiled-coil dimers between the product gAgT and the template eTeA (left side) and the product gAgT and the template gAgT (right side) in the ligation reaction of gA-n and gT-c.

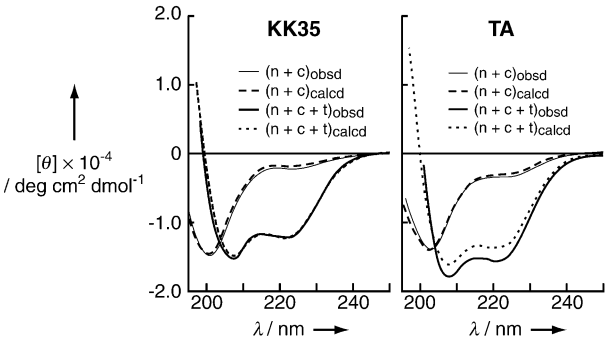

Figure 10. CD spectra of the mixture of fragment and template peptides in the 35-residue peptide systems. Thin lines represent observed spectra of a mixture of N- and C-terminal fragments (50 μM each), and dashed lines are the calculated additional spectra of the N and C fragments. Bold lines represent observed spectra of a mixture of N- and C-terminal fragments with the template (50 μM each), and dotted lines are the calculated additional spectra of N and C fragments with the template. For further details, see the Experimental Section.

Circular dichroism study: In order to evaluate the preorganization of the template and fragments (coiled-coil formation), circular dichroism (CD) spectra were measured (Figure 10). Template peptides took a highly helical structure, as evaluated from the bimodal spectra with two negative peaks at 208 and 222 nm. The CD spectra of individual fragments and equimolar mixtures (50 μM each) showed that fragment peptides were predominantly in random-coil formation, with no preorganization indicated between the two fragments. In the TA system, when the two fragments and the template were mixed, the coiled-coil formation was induced. The observed CD signal at 222 nm was $-15300\ \mathrm{deg\,cm^2\,dmol^{-1}}$, whereas the calculated signal of the additional spectrum of each peptide was $-13600\ \mathrm{deg\,cm^2\,dmol^{-1}}$. The signal increase ($\Delta[\theta]_{222} = 1700\ \mathrm{deg\,cm^2\,dmol^{-1}}$) suggests that the TA template induced the α-helical formation of fragments. Therefore, we considered that the thymine–adenine pairs offered advantages for peptide recognition and association followed by coiled-coil formation of the template and both fragments. In contrast, in the KK35 and QQ35 systems, the helix induction in the mixture of fragments and the template was not observed.

The nucleobase interaction also affected peptide stability. Thermal denaturation studies were carried out with template peptides in which the thiol side chain of the Cys residue was alkylated to prevent interhelical disulfide formation during the measurement. We considered that the alkylation did not affect the peptide structure, since the spectra of the alkylated

peptides were same as those of the original peptides measured immediately after the solution preparation. The thermal denaturation curves of heating and cooling were superimposable, which indeicates that reversible denaturation occurred. The melting curves were fitted by using an equation assuming a two-state transition (Figure 11a) to afford the melting

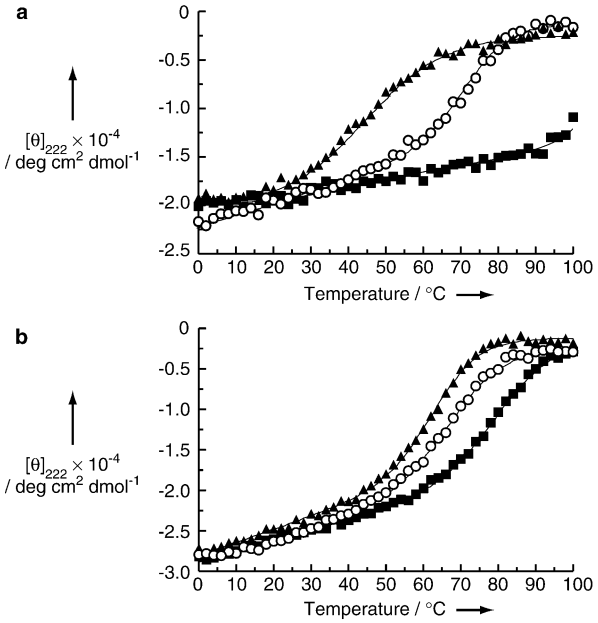

Figure 11. Thermal denaturation profiles of the template peptides estimated by thermal dependence of the ellipticities at 222 nm for the peptides. a) 35-residue peptide systems: KK35-t (▲), QQ35-t (○), and TA-t (■) in 100 mM sodium phosphate buffer (pH 7.5). [Peptide] = 9 μM. b) 28-residue peptide systems: QQ28 (▲), eTeA (○), and gTgA (■) in 100 mM sodium phosphate buffer (pH 7.5) containing 2.5 M GuHCl. [Peptide] = 10 μM. Fitting curves were obtained assuming the two-state transition.

temperatures, T_m, for the 35-residue template peptides.[24, 25] The T_m value of the template TA-t (with four nucleobases) was over 100 °C, much higher than those of KK35-t and QQ35-t (44 and 73 °C, respectively). These results suggested that the nucleobase interaction increased the structural stability. Moreover, in the 28-residue templates (with two nucleobases), a significant difference with respect to the position of the NBAs was observed (Figure 11b). The T_m value of the gTgA template was 80 °C, which was higher than the values of 68 °C for the eTeA template and 63 °C for the QQ28 template (T_m values in the presence of 2.5 M GuHCl). The thymine – adenine pairs at the $g-g'$ positions (in gTgA) raised the T_m value by 17 °C compared with the T_m value of QQ28, whereas the thymine – adenine pairs at the $e-e'$ positions (in eTeA) showed a small increase of 5 °C. The difference between gTgA and eTeA implied the advantage of the $g-g'$ positions for the nucleobase interaction, and this implication was compatible with the results of the self-replication reactions.

Discussion

By using the self-replication system, we successfully demonstrated that nucleobase interactions in the peptide structures effectively contributed to peptide – peptide recognition. The self-replication reaction was based on recognition through the coiled-coil formation of template and fragment peptides. Additional complementary nucleobase interaction reinforced peptide – peptide recognition and accelerated the reaction. In the presence of the denaturant GuHCl, the replication reactions were largely suppressed, a fact indicating that the formation of the coiled-coil structure was responsible for the recognition and that the nucleobases could work in the folded structure. We found that the nucleobases are required to be arranged in the peptide structure; however, there was no direct evidence for what type of interaction (such as hydrogen-bonding or hydrophobic (base-stacking) interactions) dominantly worked between the nucleobases. It was shown that the TT system (having mismatched base pairs) was not effective in the self-replication reaction. Our previous study had demonstrated that at the $g-g'$ positions one or two thymine – adenine pair(s) was more effective than two thymine – thymine or adenine – adenine pairs in formation of a coiled-coil structure.[16] We also showed that interactions of pairs of NBAs at the $g-g'$ positions were preferable to those at the $e-e'$ or $e-g'$ positions. Considering these results, it can be concluded that some specific interactions other than hydrophobic ones were at work.

The selectivity of the nucleobase position was very suggestive. The fact that the $g-g'$ positions were superior to the $e-e'$ positions for the peptide – peptide interaction might be explained by the difference in the direction of the amino acid side chains. In an antiparallel coiled-coil structure, the directions at $g-g'$ positions are pointed toward each other whereas those at $e-e'$ positions are pointed away from each other. Therefore, the nucleobases preferably interacted between the g and g' positions in the antiparallel coiled-coil form. This suggestion about the strand orientation was further supported by the demonstration with the gAgT system. The fragments gA-n and gT-c were designed to associate with the gAgT template and also with the eTeA template (Figure 9). By use of the gAgT template, two thymine – adenine pairs can be performed between the $g-g'$ positions in an antiparallel coiled-coil structure. On the other hand, with the template eTeA, two thymine – adenine pairs may be formed at the $e-g'$ positions in a parallel coiled-coil formation. The results showed that only the gAgT template could give the acceleration effect derived from nucleobase interaction, since the effect of the eTeA template was almost the same as that of the QQ28 template which does not have nucleobases. These results allowed us to conclude that the nucleobase interaction could work between the $g-g'$ positions in an antiparallel coiled-coil structure. One of the reasons why an antiparallel coiled-coil formation was preferred may be the inherent properties of the coiled-coil peptide. Without NBAs, coiled-coil peptides used in this study are predicted to be more stable in an antiparallel orientation, since coiled-coil peptides in which Leu residues were placed at both the a and d positions were reported to be more stable in antiparallel orientation than in parallel orientation,[26] due to the dipole moment from the C terminus to the N terminus in the α-helical peptide.[27] Additionally, the advantage of the $g-g'$ position compared with the $e-e'$ positions was shown even in reactions with a

mixture of four kinds of fragments, such as gA-n, gT-c, eT-n, and eA-c (data not shown). Although the four fragments generated four species of products (gAgT, eTeA, gAeA, and eTgT), the production of gAgT was most active. Furthermore, by addition of gAgT as a template in this mixture, production of gAgT was increased more than by addition of eTeA. From consideration of the apparent preference for the $g-g'$ positions as shown, the strand orientation is assumed to be dominantly antiparallel, although the aggregation state might not be limited to only a dimer. Through this orientation the nucelobases are arranged by the peptide structure to be able to interact specifically with each other.

The replication systems previously reported were kinetically analyzed by using the program SimFit.[20–22] Following the reactions, we had attempted the analysis, and a fitting curve for each reaction was obtained (Figure 4a). The apparent rate constants for the autocatalytic reactions of the KK-35 and QQ-35 systems were 12 and 6 $M^{-3/2}s^{-1}$, and the rate constants for the uncatalyzed reactions were 0.01 and 0.02 $M^{-1}s^{-1}$, respectively. These values are comparable to those of reported reactions.[21, 22] However, in the case of NBA-containing systems such as the TA system, the rate constants in the system varied at each template concentration. In the TA system, the rate constant for the autocatalytic reaction ranged from 11–75 $M^{-3/2}s^{-1}$ for the initial template concentrations of 0–20 μM. These results implied that the association state of the template (product) was not simple, depending on the peptide concentration, and further implied the probability of product inhibition. Product inhibition is a serious problem related to the catalytic efficiency and causes parabolic growth curves.[22a, b, 28] Product inhibition is related to the stability of the product–template complex and is probably dependent on the peptide concentration. In some cases, the deviation between fitting curves and actual product formation plots became larger at higher template concentrations.[22b, 28] To reduce product inhibition, destabilization of the coiled-coil structure was reported to be effective.[22c] In this study, improved thermal stability of template peptides by nucleobase interactions was indicated. Therefore, the interactions might induce additional association and product inhibition, although the K1K2 peptide (corresponding to KK35-t) was reported to be in both the monomeric and dimeric forms.[22a] Peptide association might not be simple, thereby preventing the satisfactory kinetic analysis of the TA system. Although the aggregation states were still unclear, as described above, the preference in the nucleobase positions for the interaction and reaction strongly implies that the helical strands in an active species were at least not in random orientation but regulated in antiparallel orientation.

We demonstrated that the thermal stability of coiled-coil peptides was improved by incorporated nucleobase interactions. Previously, we had shown that complementary base pairs stabilized the coiled-coil peptide in which two helical strands were linked through a disulfide bond.[16] In this study, we found that nucleobase interactions can work between peptides not linked with each other. The positional preference and base selectivity of incorporated nucleobases, in a sense, implied regulated association. These demonstrations indicated the availability of nucleobases for the enhancement of structural complementarity for designing functional peptide assembly.

Conclusion

Peptides with complementary base pairs facilitated the self-replicating reaction. These results successfully demonstrated an advantage of complementary nucleobase interaction for peptide–peptide interactions. Peptide assembly was largely attributed to the nucleobase interactions appropriately arranged in the coiled-coil structure. Thus, the structured peptide and nucleobases cooperatively functioned to reinforce the structural complementarity. Complementarity is indispensable for the regulation of molecular assembly. The combination of nucleobase and peptide has the potential to create a specific and stable structure, more complicated than a coiled-coil formation, in a designed peptide assembly; this could lead to the design of unique functions. Moreover, this demonstration may provide inspiration for the design of peptides together with nucleic acid functions. Nucleic acids fill the role of information storage and transfer as genetic materials. In chemical evolution on the early earth, molecules to amplify and transfer information might have existed. Peptides equipped with self-replicating ability and nucleobase information may be candidates for the prebiotic molecule.[29, 30]

Experimental Section

Peptide synthesis: 9-Fluorenylmethoxycarbonyl (Fmoc)-protected NBA monomers were synthesized according to the reported methods.[15] Peptides, including NBA-combined peptides, were synthesized by the solid-phase method with Fmoc strategy[31] on an Advanced Chemtech BenchMark 348 multiple peptide synthesizer by using the following Fmoc-protected amino acid derivatives: Fmoc-Ala-OH, Fmoc-Cys(Trt)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Glu(O*t*Bu)-OH, Fmoc-Gly-OH, Fmoc-Lys(Boc)-OH, Fmoc-Leu-OH, Fmoc-Tyr(*t*Bu)-OH, Fmoc-A_{NBA}(Z)-OH, Fmoc-C_{NBA}(Z)-OH, Fmoc-G_{NBA}-OH, Fmoc-T_{NBA}-OH (Trt = triphenylmethyl, *t*Bu = *tert*-butyl, Boc = *tert*-butyloxycarbonyl, Z = benzyloxycarbonyl).

The C-terminal fragment peptides were synthesized on the 4-(2′,4′-dimethoxyphenyl-aminomethyl)phenoxy resin (Rink amide resin),[32] by using Fmoc-protected amino acid derivatives (6 equiv), *N*,*N*-diisopropylethylamine (DIEA, 12 equiv), 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluroniumhexafluorophosphate (HBTU, 6 equiv), and 1-hydroxybenzotriazole hydrate (HOBt·H_2O, 6 equiv) in *N*-methylpyrrolidone (NMP) for coupling, and 25% piperidine/NMP for Fmoc removal. To cleave the peptide from the resin and to remove the side-chain protecting groups, the peptide–resin was treated with trifluoroacetic acid (TFA) in the presence of *m*-cresol, thioanisole, and ethanedithiol. In the case of peptide fragments containing A_{NBA}, for the complete removal of the Z group on the adenine moiety, the peptide was further treated with 1 M trimethylsilyl trifluoromethanesulphonate (TMSOTf) in TFA in the presence of *m*-cresol, thioanisole, and ethanedithiol at 0 °C.[33]

The N-terminal peptides were synthesized on a 2-chlorotrityl chloride resin (Cl-Trt resin).[34] The elongated peptide was detached from the resin by treatment with acetic acid/2,2,2-trifluoroethanol/dichloromethane (1:1:3) for 2 h. The resulting peptide was treated with 3-mercaptopropionic acid ethyl ester (20 equiv), 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC·HCl, 10 equiv), and HOBt·H_2O (10 equiv) in *N*,*N*-dimethylformamide at 0 °C overnight.[35] The removal of protecting groups

from amino acid side chains was carried out by TFA or TMSOTf treatment as described above.

The template peptides were synthesized by native chemical ligation[22] between the N-terminal thiol ester fragment and the C-terminal fragment. Chemical ligation of the fragment peptides was carried out in 100 mM sodium phosphate buffer (pH 7.5) containing thiophenol (4% v/v) and 6 M GuHCl under a nitrogen atmosphere.

All peptides were purified by semipreparative reversed-phase HPLC and identified in satisfactory results by MALDI-TOF MS and amino acid analyses. The amino acid analyses were also utilized to determine the peptide concentration of the stock solutions.

CD measurements: CD measurements were performed on a Jasco J-720WI spectropolarimeter equipped with a thermoregulator and by using a quartz cell with 1.0 or 2.0 mm pathlength. The stock solution of each peptide was diluted with 100 mM sodium phosphate buffer (pH 7.5) containing 0.5 mM dithiothreitol (DTT) and 0.5 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP). For the thermal denaturation studies, to prevent the formation of a disulfide bond during the measurement, template peptides with the Cys residue alkylated by treatment with bromoacetic acid methyl ester were prepared.

Self-replication reactions: All fragment and template peptides were dissolved in 0.1% aqueous TFA. The appropriate peptide solutions were mixed, then the reactions were initiated by adding to the peptide mixture 250 mM 3-morpholinopropanesulfonic acid (MOPS) buffer (pH 7.5) containing 1% (v/v) 3-mercaptopropionic acid ethyl ester and 0.4 mM 4-dimethylaminobenzoic acid as an internal standard; the buffer had been bubbled with nitrogen prior to addition. Average final concentrations: [MOPS] = 140 mM, [4-dimethylaminobenzoic acid] = 0.2 mM, 0.7% (v/v) 3-mercaptopropionic acid ethyl ester. The reaction temperature was maintained at 4 or 22 °C. Aliquots (15 μL) were removed from the reaction solution at each time point, immediately quenched with 5% aqueous TFA (20 μL), and stored at −20 °C prior to HPLC analysis. Product formation was analyzed quantitatively by reversed-phase HPLC against a standard curve of the product. HPLC peaks were detected at 220 and 260 nm and identified by comparing the retention times with those of the standard samples. This was followed by MALDI-TOF MS.

Acknowledgements

We are grateful to Prof. G. von Kiedrowski for the generous gift of the software program SimFit.

[1] a) R. B. Hill, D. P. Raleigh, A. Lombardi, W. F. DeGrado, *Acc. Chem. Res.* **2000**, *33*, 745 – 754; b) C. M. Summa, A. Lombardi, M. Lewis, W. F. DeGrado, *Curr. Opin. Struct. Biol.* **1999**, *9*, 500 – 508; c) W. F. DeGrado, C. M. Summa, V. Pavone, F. Nastri, A. Lombardi, *Annu. Rev. Biochem.* **1999**, *68*, 779 – 819.

[2] a) L. Baltzer, H. Nilsson, J. Nilsson, *Chem. Rev.* **2001**, *101*, 3153 – 3163; b) L. Baltzer, J. Nilsson, *Curr. Opin. Biotechnol.* **2001**, *12*, 355 – 360; c) L. Baltzer, *Curr. Opin. Struct. Biol.* **1998**, *8*, 466 – 470.

[3] a) M. Mutter, *Chimia*, **2000**, *54*, 552 – 557; b) G. Tuchscherer, L. Scheibler, P. Dumy, M. Mutter, *Biopolymers* **1998**, *47*, 63 – 73.

[4] K. H. Mayo, *Trends Biotechnol.* **2000**, *18*, 212 – 217.

[5] B. R. Gibney, F. Rabanal, P. L. Dutton, *Curr. Opin. Chem. Biol.* **1997**, *1*, 537 – 542.

[6] G. Xing, V. J. DeRose, *Curr. Opin. Chem. Biol.* **2001**, *5*, 196 – 200.

[7] W. D. Kohn, R. S. Hodges, *Trends Biotechnol.* **1998**, *16*, 379 – 389.

[8] J. R. Beasley, M. H. Hecht, *J. Biol. Chem.* **1997**, *272*, 2031 – 2034.

[9] F. Offredi, F. Dubail, P. Kischel, K. Sarinski, A. S. Stern, C. Van de Weerdt, J. C. Hoch, C. Prosperi, J. M. Francois, S. L. Mayo, J. A. Martial, *J. Mol. Biol.* **2003**, *325*, 163 – 174.

[10] S. Matile, *Chem. Soc. Rev.* **2001**, *30*, 158 – 167.

[11] C. M. Niemeyer, *Angew. Chem.* **2001**, *113*, 4254 – 4278; *Angew. Chem. Int. Ed.* **2001**, *40*, 4128 – 4158.

[12] a) P. E. Nielsen, M. Egholm, R. H. Berg, O. Buchardt, *Science* **1991**, *254*, 1497 – 1500; b) M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. Driver, R. H. Berg, S. K. Kim, B. Norden, P. E. Nielsen, *Nature* **1993**, *365*, 566 – 568; c) P. Wittung, P. E. Nielsen, O. Buchardt, M. Egholm, B. Norden, *Nature* **1994**, *368*, 561 – 563; d) P. E. Nielsen, *Curr. Opin. Biotechnol.* **2001**, *12*, 16 – 20; e) P. E. Nielsen, *Acc. Chem. Res.* **1999**, *32*, 624 – 630.

[13] a) U. Diederichsen, *Angew. Chem.* **1996**, *108*, 458 – 461; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 445 – 448; b) U. Diederichsen, H. W. Schmitt, *Angew. Chem.* **1998**, *110*, 312 – 315; *Angew. Chem. Int. Ed.* **1998**, *37*, 302 – 305; c) M. F. H. Hoffmann, A. M. Brückner, T. Hupp, B. Engels, U. Diederichsen, *Helv. Chim. Acta* **2000**, *83*, 2580 – 2593.

[14] a) P. Garner, Y. Huang, S. Dey, *ChemBioChem* **2001**, 224 – 226; b) P. Garner, S. Dey, Y. Huang, *J. Am. Chem. Soc.* **2000**, *122*, 2405 – 2406.

[15] a) T. Takahashi, A. Ueno, H. Mihara, *ChemBioChem* **2002**, 543 – 549; b) T. Takahashi, K. Hamasaki, A. Ueno, H. Mihara, *Bioorg. Med. Chem.* **2001**, *9*, 991 – 1000; c) T. Takahashi, K. Hamasaki, I. Kumagai, A. Ueno, H. Mihara, *Chem. Commun.* **2000**, 349 – 350.

[16] S. Matsumura, A. Ueno, H. Mihara, *Chem. Commun.* **2000**, 1615 – 1616.

[17] A. Robertson, A. J. Sinclair, D. Philp, *Chem. Soc. Rev.* **2000**, *29*, 141 – 152.

[18] D. H. Lee, K. Severin, M. R. Ghadiri, *Curr. Opin. Chem. Biol.* **1997**, *1*, 491 – 496.

[19] L. E. Orgel, *Nature* **1992**, *358*, 203 – 209.

[20] a) G. von Kiedrowski, *Angew. Chem.* **1986**, *98*, 932 – 934; *Angew. Chem. Int. Ed. Engl.* **1986**, *25*, 932 – 935; b) D. Sievers, G. von Kiedrowski, *Nature* **1994**, *369*, 221 – 224; c) D. Sievers, G. von Kiedrowski, *Chem. Eur. J.* **1998**, *4*, 629 – 641; d) H. Schöneborn, J. Bülle, G. von Kiedrowski, *ChemBioChem* **2001**, 922 – 927.

[21] a) D. H. Lee, J. R. Granja, J. A. Martinez, K. Severin, M. R. Ghadiri, *Nature* **1996**, *382*, 525 – 528; b) K. Severin, D. H. Lee, J. A. Martinez, M. R. Ghadiri, *Chem. Eur. J.* **1997**, *3*, 1017 – 1024; c) K. Severin, D. H. Lee, J. A. Martinez, M. Vieth, M. R. Ghadiri, *Angew. Chem.* **1998**, *110*, 133 – 135; *Angew. Chem. Int. Ed.* **1998**, *37*, 126 – 128; d) A. Saghatelian, Y. Yokobayashi, K. Soltani, M. R. Ghadiri, *Nature* **2001**, *409*, 797 – 801.

[22] a) S. Yao, I. Ghosh, R. Zutshi, J. Chmielewski, *Angew. Chem.* **1998**, *110*, 489 – 492; *Angew. Chem. Int. Ed.* **1998**, *37*, 478 – 481; b) S. Yao, I. Ghosh, R. Zutshi, J. Chmielewski, *J. Am. Chem. Soc.* **1997**, *119*, 10559 – 10560; c) R. Issac, J. Chmielewski, *J. Am. Chem. Soc.* **2002**, *124*, 6808 – 6809.

[23] a) P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science* **1994**, *266*, 776 – 779; b) G. G. Kochendoerfer, S. B. Kent, *Curr. Opin. Chem. Biol.* **1999**, *3*, 665 – 671; c) P. E. Dawson, M. J. Churchill, M. R. Ghadiri, S. B. H. Kent, *J. Am. Chem. Soc.* **1997**, *119*, 4325 – 4329.

[24] N. E. Zhou, C. M. Kay, B. D. Sykes, R. S. Hodges, *Biochemistry* **1993**, *32*, 6190 – 6197.

[25] W. J. Becktel, J. A. Schellman, *Biopolymers* **1987**, *26*, 1859 – 1877.

[26] O. D. Monera, C. M. Kay, R. S. Hodges, *Biochemistry* **1994**, *33*, 3862 – 3871.

[27] a) K. R. Shoemaker, P. S. Kim, E. J. York, J. M. Stewart, R. L. Baldwin, *Nature* **1987**, *326*, 563 – 567; b) K. R. Shoemaker, P. S. Kim, D. N. Brems, S. Marqusee, E. J. York, I. M. Chaiken, J. M. Stewart, R. L. Baldwin, *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 2349 – 2353.

[28] a) S. Yao, I. Ghosh, R. Zutshi, J. Chmielewski, *Nature* **1998**, *396*, 447 – 450, b) S. Yao, J. Chmielewski, *Biopolymers* **1999**, *51*, 370 – 375.

[29] a) K. E. Nelson, M. Levy, S. L. Miller, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 3868 – 3871; b) S. L. Miller, *Nature Struct. Biol.* **1997**, 167 – 169.

[30] B. M. Rode, *Peptides* **1999**, 20, 773 – 786.

[31] W. C. Chan, P. D. White in *Fmoc Solid Phase Peptide Synthesis: A Practical Approach* (Eds.: W. C. Chan, P. D. White), Oxford University Press, New York, **2000**, pp. 41 – 76.

[32] H. Rink, *Tetrahedron Lett.* **1987**, *28*, 3787 – 3790.

[33] N. Fujii, A. Otaka, O. Ikemura, K. Akagi, S. Funakoshi, Y. Hayashi, Y. Kuroda, H. Yajima, *J. Chem. Soc. Chem. Commun.* **1987**, 274 – 275.

[34] a) K.Barlos, D. Gatos, S. Kapolos, G. Paraphotiu, W. Schäfer, Y. Wenqing, *Tetrahedron Lett.* **1989**, *30*, 3947 – 3950; b) K. Barlos, O. Chatzi, D. Gatos, G. Stavropoulos, *Int. J. Peptide Protein Res.* **1991**, *37*, 513 – 520.

[35] S. Futaki, K. Sogawa, J. Maruyama, T. Asahara, M. Niwa, H. Hojo, *Tetrahedron Lett.* **1997**, *38*, 6237 – 6240.

Received: May 9, 2003 [F5121]

