Chemistry & Biology, Vol. 11, 765–773, June, 2004, ©2004 Elsevier Ltd. All rights reserved. DOI 10.1016/j.chembiol.2004.03.032

Synthesis of Functional Proteins by Mixing Peptide Motifs

Hirohide Saito,[1] Takako Honma,[2] Tamiko Minamisawa,[1] Kanami Yamazaki,[3] Tetsuo Noda,[2,4] Takao Yamori,[3] and Kiyotaka Shiba[1,5,*]

[1]Department of Protein Engineering
[2]Department of Cell Biology
[3]Division of Molecular Pharmacology and Cancer Chemotherapy Center
Cancer Institute
Japanese Foundation for Cancer Research
Toshima, Tokyo 170-8455
[4]Department of Molecular Genetics
University of Tohoku School of Medicine
Aoba Sendai 980-8575
Japan
[5]CREST, JST

Summary

Here, we describe a synthetic approach for generating artificial proteins by the assemblage of naturally occurring peptide motifs. Two motifs respectively related to apoptosis induction and protein transduction were encrypted into different reading frames of an artificial gene (microgene), which was then polymerized; random frame shifts at the junctions between the microgene units yielded combinatorial polymers of three reading frames. Among the proteins created, #284 was found to penetrate through cell membranes and exert a strong apoptotic effect on several cancer cell lines. Because a simple linkage of these motifs was not sufficient to construct a bifunctional peptide, and the successful reconstitution was dependent on how they were joined together, the combinatorial strategy is important for reconstituting functions from mixtures of motifs. This microgene-based approach represents a novel system for creating proteins with desired functions.

Introduction

Recently, with the advent of high-throughput proteomics and genomics, the identification of "peptide motifs" associated with biological functions, protein structures, or evolutionary history has been drastically accelerated [1, 2]. Examples include the Bcl-2 homology 1-4 peptide motifs (BH1-BH4 motifs), which have been found from proteins involved in the signal transduction of programmed cell death (apoptosis) pathways [3]. Complex interplays of these motifs underpin the highly regulated cell proliferation, and aberration in the network often results in cell malignancy.

Interestingly, several peptide motifs recapture biological functions even when they are isolated from their parental bodies. For example, a 16-amino acid peptide comprising the BH3 motif of the proapoptotic Bax protein (192 amino acids) [4] is sufficient to induce apoptosis when microinjected into Rat-1 fibroblasts [5]. Moreover, the synthesis of multifunctional polypeptides was successfully achieved from the conjugation of peptide motifs in some cases [6–9]. However, the simple linkage of multiple motifs does not always result in reconstitution of expected functions. A conjugant of the 8-amino acid polyarginine, which has a protein translocation ability, and the 25-amino acid BH3 domain of the proapoptotic Bad did not induce apoptosis in Jurkat leukemic cells; on the other hand, a conjugant of polyarginine and the 20-amino acid BH3 domain of proapoptotic Bid did [10].

The capriciousness of the reconstitution from peptide motifs becomes more problematic as the length of the peptide motifs become shorter. Frugier et al. [9] recently constructed artificial tRNA recognition domains comprised of polymers of a 10-amino acid MFβ2 motif—a selected peptide as an aptamer to $tRNA^{Ala}$ from a phage display library [11]—and a 5-amino acid RGG-box motif, previously identified as a general RNA binding enhancer [12]. They found that the order and number of the motifs significantly influenced the effective reconstitutions of the active proteins. Motifs assembled in the order R (RGG-box)-M (MFβ2) acquired the $tRNA^{Ala}$ recognition ability, whereas assembly in the order of M-R did not. The most active construct was created by fusion of a 5-mer polymer, R-R-M-R-R, though R-M, a 2-mer polymer, exerted a stronger effect than the 3-mer polymers R-M-R and R-R-M, highlighting the unpredictability of the functional reconstitution from the motif assembly.

Under these circumstances, it is desired to establish a novel methodology to efficiently reconstitute multiple functions from the assemblages of short peptide motifs. We have previously developed a protein evolution system in which three different peptides coded by a single microgene were combinatorially polymerized to give a large protein library [13, 14]. This strategy exhibits several advantages for constructing artificial proteins: (1) one simple reaction allows the construction of a combinatorial library composed of three reading frames coded by a single microgene; (2) large open reading frames are readily prepared from the polymerization of a microgene that is devoid of a termination codon in all of its reading frames; (3) the repetitiousness retained in the constructed proteins has been shown to contribute to the emergence of ordered proteins [14]; and (4) desired functions and structures can be rationally encrypted into a microgene, giving a tailor-made combinatorial library. In the present study, we demonstrate the synthesis of artificial proteins, which penetrate through membranes and induce apoptosis in several cancer cell lines, by the combinatorial assemblage of naturally occurring peptide motifs.

Results

Reconstruction of Artificial Proteins by a Combinatorial Approach

We began by constructing artificial proteins using two related short peptide sequences, the protein transduc-

*Correspondence: kshiba@jfcr.or.jp

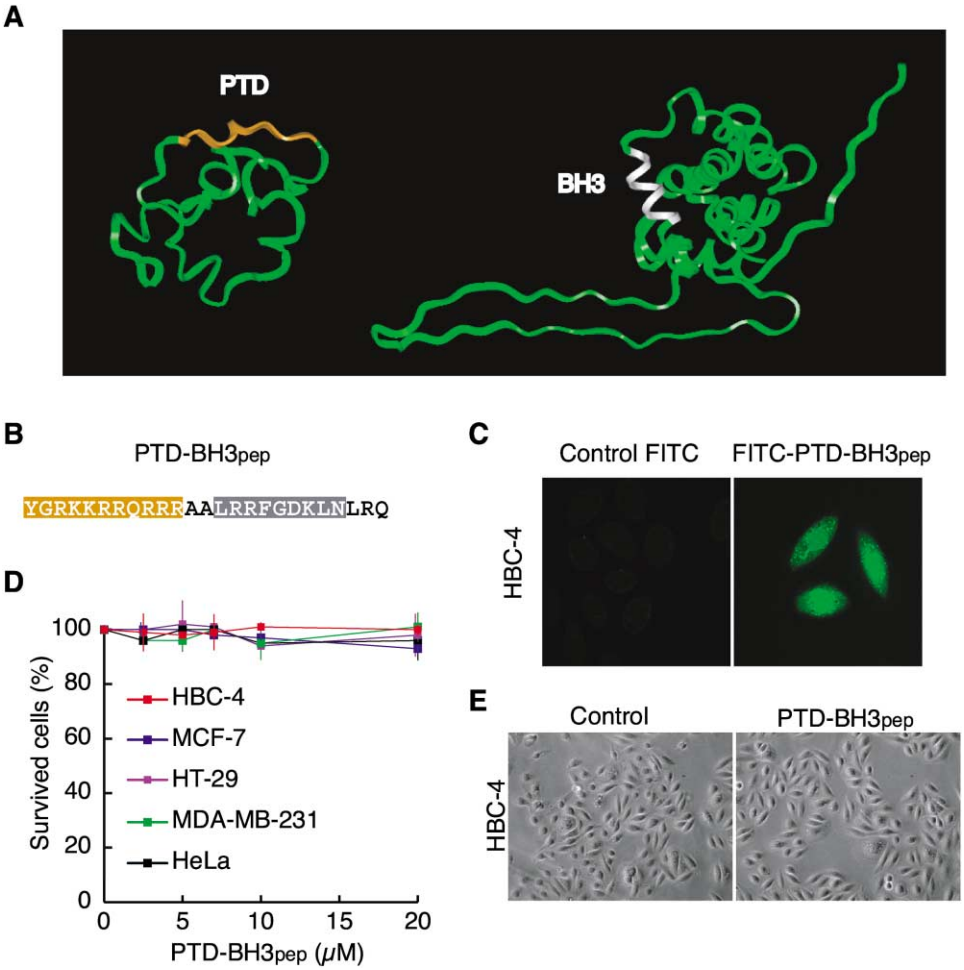

Figure 1. Functional Analysis of a Synthetic Peptide Composed of Two Motifs

(A) Two naturally occurring peptide motifs, PTD^{Tat} (^{47}YGRKKRRQRRR^{57}) and $BH3^{Noxa}$ (^{29}LRRFGDKLN^{37}), were selected in this study. On the left is the structure of the HIV tat protein [34], in which the PTD motif is colored orange. Because the structure of Noxa has not yet been determined, its sibling, Bid [35], is shown on the right. The BH3 motif (^{90}LAQIGDEMD^{98}) of this protein is colored gray. (B) The sequence of PTD-$BH3_{pep}$. PTD^{Tat} and $BH3^{Noxa}$ were connected via two alanines, and an extra tripeptide (LRQ) was added at its C terminus.

(C) Translocation of PTD-$BH3_{pep}$ across the cell membrane. Thirty minutes after the addition of the FITC-labeled PTD-$BH3_{pep}$ (15 μM) to the culture medium, cells were fixed in 4% paraformaldehyde and examined using a confocal laser microscope to localize the peptide (right). As a control, free FITC was added to the medium (left).

(D) Effects of PTD-$BH3_{pep}$ on cell viability. Five cancer cell lines (HBC-4, MCF-7, HT-29, MDA-MB-231, and HeLa) were incubated for 3 hr with the indicated concentrations of the indicated peptides, after which the metabolic activity was assessed using WST-1 assays. Data for HBC-4 were obtained from three independent experiments, and the bar shows standard deviation. For other cells, experiments were repeated twice and data shown are average values of two independent experiments. Bars represent observed values.

(E) Morphology of PTD-$BH3_{pep}$-treated HBC-4 cells. There were no obvious morphological differences between the control cells (left) and those incubated for 20 hr with 50 μM PTD-$BH3_{pep}$ (right).

tion domain of HIV Tat (PTD^{Tat}) and the BH3 motif from human Noxa ($BH3^{Noxa}$) (Figure 1A). The 11-amino acid PTD^{Tat} peptide has been used for the translocation of peptides, proteins, nucleic acids, and other molecules across the plasma membrane [15]. The 12-amino acid $BH3^{Noxa}$ motif, which is comprised of a 9-amino acid BH3 core sequence and an extra tripeptide at its C terminus, is derived from a human proapoptotic BH3-only protein, Noxa (54-amino acids), and has been shown to mediate p53-dependent apoptosis [16]. Previous work demonstrated that 20-amino acid mouse Noxa BH3 domain, which contains the 9-amino acid BH3 core sequence, failed to recapitulate proapoptotic functions such as the release of cytochrome c from mitochondria or the interaction with antiapoptotic Bcl-2 protein [10].

We initially investigated the function of a 25-amino acid synthetic peptide in which PTD^{Tat} and $BH3^{Noxa}$ were connected via two alanines (PTD-$BH3_{pep}$, Figure 1B). Although translocation of PTD-$BH3_{pep}$ into cells was confirmed (Figure 1C), it had no inhibitory effect on the growth of several tumor cell lines, nor did it affect the cell morphology (Figures 1D and 1E). Thus, simple conjugation of PTD^{Tat} and $BH3^{Noxa}$ was not sufficient to create a bifunctional peptide.

We next aimed to reconstitute apoptosis-inducing activity by assembling polymers of these motifs. We re-

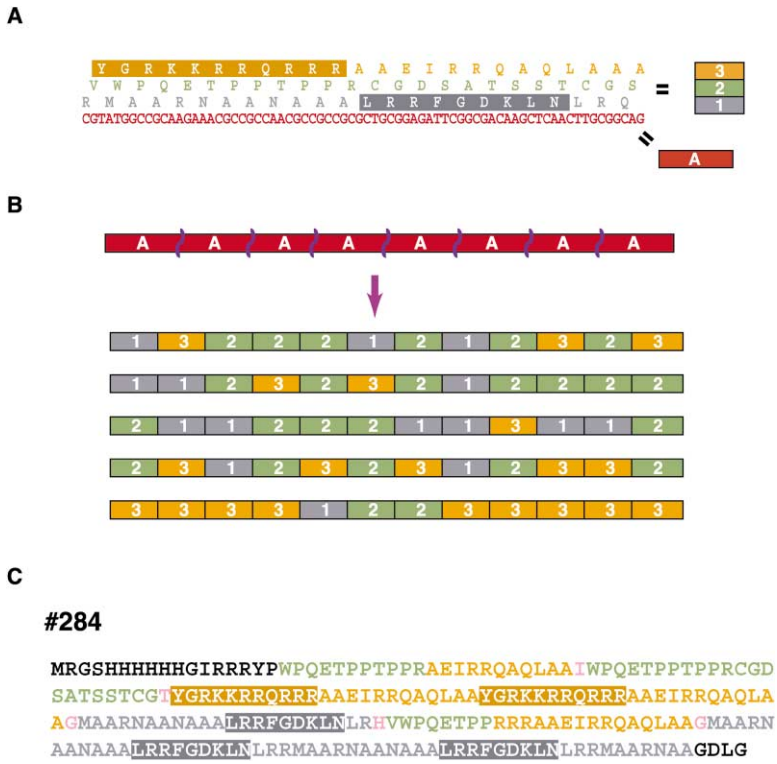

Figure 2. Synthesis of Artificial Proteins by Combinatorial Assemblage of Peptide Motifs

(A) The 72 bp microgene (MG-27) was designed using the CyberGene program (data not shown) so that: (1) the third reading frame of the 5′ half would code for the PTD^{Tat} motif (shown by the orange box); (2) the first reading frame of the 3′ half would code for the BH3^{Noxa} motif (shown by the gray box); and (3) both of these reading frames would code for peptides with a propensity to form α helix. None of the three reading frames contain a termination codon.

(B) Polymers of MG-27 were prepared using the MPR method [13], with which insertions or deletions of nucleotides randomly occur at junctions between the microgene units (shown by wavy lines), endowing the translated products with combinatorial molecular diversity.

(C) Primary structure of artificial protein 284, which is composed of two PTD^{Tat} (orange box) and three BH3^{Noxa} (gray box) motifs. The residues shown in pink are the result of the mutations leading to frame shifts. Residues in black are derived from a vector. The length, molecular weight, and pI of 284 are shown in Table 1.

cently developed a protein evolution system in which three different peptides coded by a single microgene were combinatorially polymerized to give a large repetitious protein library [13, 14]. Using this system, we were able to prepare combinatorial polymers of PTD^{Tat} and BH3^{Noxa} by encoding them in different frames of a single 72 bp microgene (MG-27) (Figure 2A). In addition, we selected the codons for these motifs so that the different reading frames coded for peptides containing an α helix, which we expected would help the structural formation of artificial proteins. Starting with MG-27 as a building block, long DNA polymers were generated using the "microgene polymerization reaction" (MPR) [13]. Two MPR primers, which contain eight complementary bases in their 3′ region and a mismatched base at their 3′-OH end, were designed so that primer dimers recreated the microgene sequence. Cycle of denaturation and elongation reactions with a pair of primers, four dNTPs, and 3′-5′ exo$^+$ thermostable DNA polymerase gave head-to-tail polymers of the microgene. Because MPR randomly inserts or deletes nucleotides at junctions between microgene units, the translational frame of the microgene is randomly shifted at junctions. This property of MPR enabled us to construct a combinatorial library composed of PTD^{Tat} and BH3^{Noxa} motifs (Figure 2B). For example, polymers consisting of 10-mer of the microgene (720 bp) could have $3^{10} = \sim 0.6 \times 10^5$ molecular diversity. The diversity of the population would be much higher when deletion and insertion at junction points are taken into account. In this study, we have randomly chosen 20 clones and expressed their proteins in *E. coli* after recloning them into an expression vector. Nine out of 20 clones were stably expressed within cells and were further characterized, which ranged from 73 to 214 amino acids and contained various combinations of the PTD^{Tat} and BH3^{Noxa} motifs (see Figure S1A in the Supplemental Data available with this article online).

An Artificial Protein with Proapoptotic Function

Initial characterization using three human hepatoma cell lines indicated that protein #284, which was composed of 214 amino acids and contained two PTD^{Tat} motifs followed by three BH3^{Noxa} motifs (Figure 2C and Table 1), exerted a strong inhibitory effect on cell growth (Figure S2A). Subsequent analysis of 39 other cancer cell lines showed that 284 inhibits the growth of several cell lines (Figure S2B). Analysis of the circular dichroism (CD) spectra showed 284 to contain a secondary structure that was not observed in PTD-BH3$_{pep}$ (Figure 3A). Further examination of the inhibitory effect of 284 confirmed that, whereas PTD-BH3$_{pep}$ did not affect cell growth even after prolonged incubation (Figure 3B), 284 inhibited the growth of several cell lines (HBC-4, MCF-7, MDA-MB-231, and HeLa) in a concentration-dependent manner (Figure 3C). Moreover, the morphological changes induced by 284, including membrane blebbing and condensation of the nucleus, were indicative of apoptosis (Figure 3D), which was also confirmed by TUNEL staining (Figure 3E), caspase activation (Figure 3F), and chromosomal DNA fragmentation (Figure S2C). These results have clearly indicated that 284 induced apoptosis in several types of cancer cell lines.

Characterization of Derivatives of 284

To determine the extent to which the embedded PTD^{Tat} and BH3^{Noxa} motifs were responsible for the induction of apoptosis, we constructed three derivatives by deletion of amino acids from the N or C terminus of 284

Table 1. Properties of Artificial Proteins Used in This Study

Name	Total Residues	Mw[a]	pI[b]	Amino Acid Composition	
				Positive Charged[c]	Negative Charged[d]
284	214	24,205	12.16	0.24	0.06
172	192	21,690	11.91	0.22	0.07
284ΔBH3	107	12,614	11.93	0.27	0.06
284ΔPTD	141	15,518	12.35	0.21	0.05
214	96	10,684	11.85	0.23	0.06
216	73	8,612	12.16	0.30	0.05

[a] Calculated from nucleotide sequences. For 284, 214, and 216, the molecular weights were determined to be 24,074, 10,627, and 8,574, respectively, by Mass spectrometry (Ciphergen, USA).
[b] Calculated from http://tw.expasy.org/tools/protparam.html.
[c] Lys and Arg.
[d] Asp and Glu.

(Figure 4A). Analysis of the CD spectra showed that 284 and its derivatives have similar adsorptions at 222 nm (data not shown), and the content rate of positively charged residues was also similar among these proteins (Table 1). However, these proteins exhibit different phenotypes in cancer cells. Cell proliferation assays showed that 172, in which the last two $BH3^{Noxa}$ motifs were deleted, leaving two PTD^{Tat} and one $BH3^{Noxa}$ motifs, still inhibited cell growth in a similar manner to 284 (Figure 4B). In contrast, when all of either motif was deleted

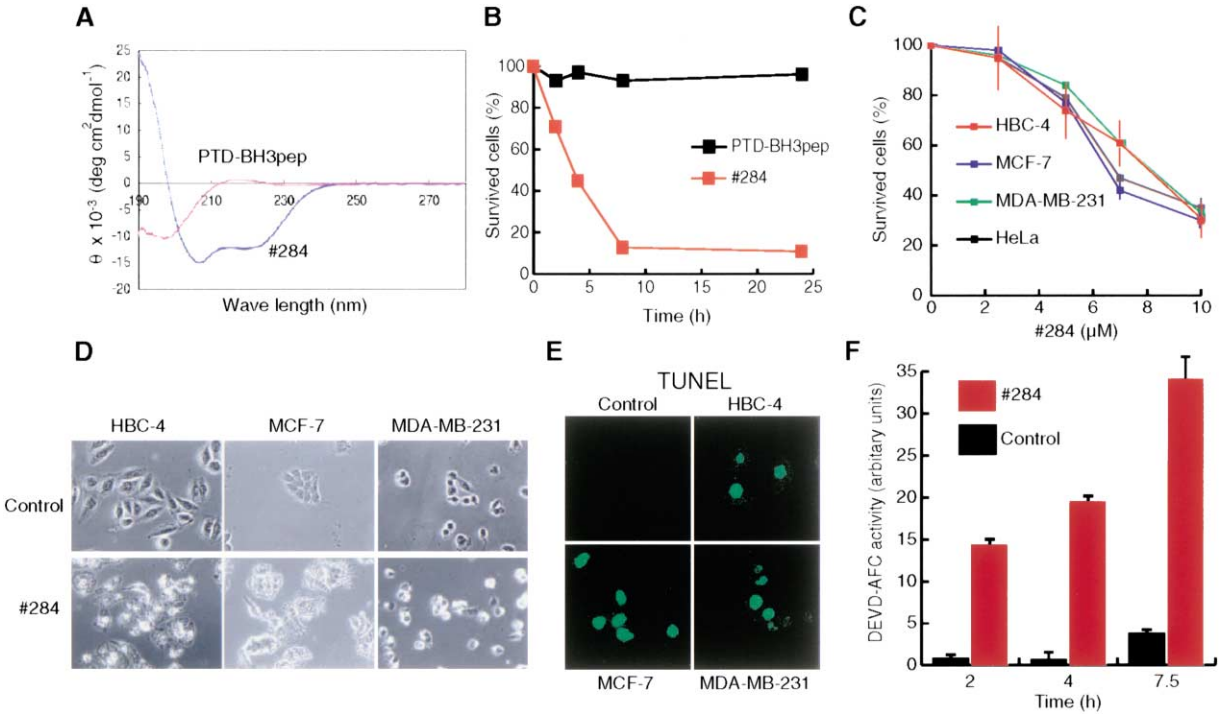

Figure 3. Characterization of 284

(A) Far UV CD spectra of PTD-$BH3_{pep}$ (red) and 284 (blue).
(B) Effect of PTD-$BH3_{pep}$ and 284 on HBC-4 cell viability. PTD-$BH3_{pep}$ (10 μM) or 284 (10 μM) was added to the cell medium at time = 0. After 2, 4, 8, and 24 hr, the number of viable cells was determined by trypan blue exclusion.
(C) Concentration-dependent effect of 284. Four cell lines (HBC-4, MCF-7, MDA-MB-231, and HeLa) were incubated for 3 hr in the indicated concentration of 284, after which the metabolic activities were assessed using WST-1 assays. Data for HBC-4 were obtained from three independent experiments, and the bar shows standard deviation. For other cells, experiments were repeated twice and bars represent observed values.
(D) Effect of 284 on cell morphology. After incubating for 1 hr in the presence or absence of 284 (10 μM), the cells (HBC-4, MCF-7, and MDA-MB-231) were fixed in 4% paraformaldehyde in PBS and observed under a light microscope.
(E) TUNEL staining. After incubating 1.5 hr with 284 (10 μM), the cells (HBC-4, MCF-7, and MDA-MB-231) were fixed in 4% paraformaldehyde, and the apoptotic cells were detected by TUNEL staining.
(F) Measurement of caspase activity. HBC-4 cells were incubated in the presence or absence of 284 (10 μM) for the indicated times. At time = 0, 2, 4, and 7.5 hr, the caspase-3-like activity was measured using a substrate analog, DEVD-AFC. Bars represent observed values of two independent experiments.

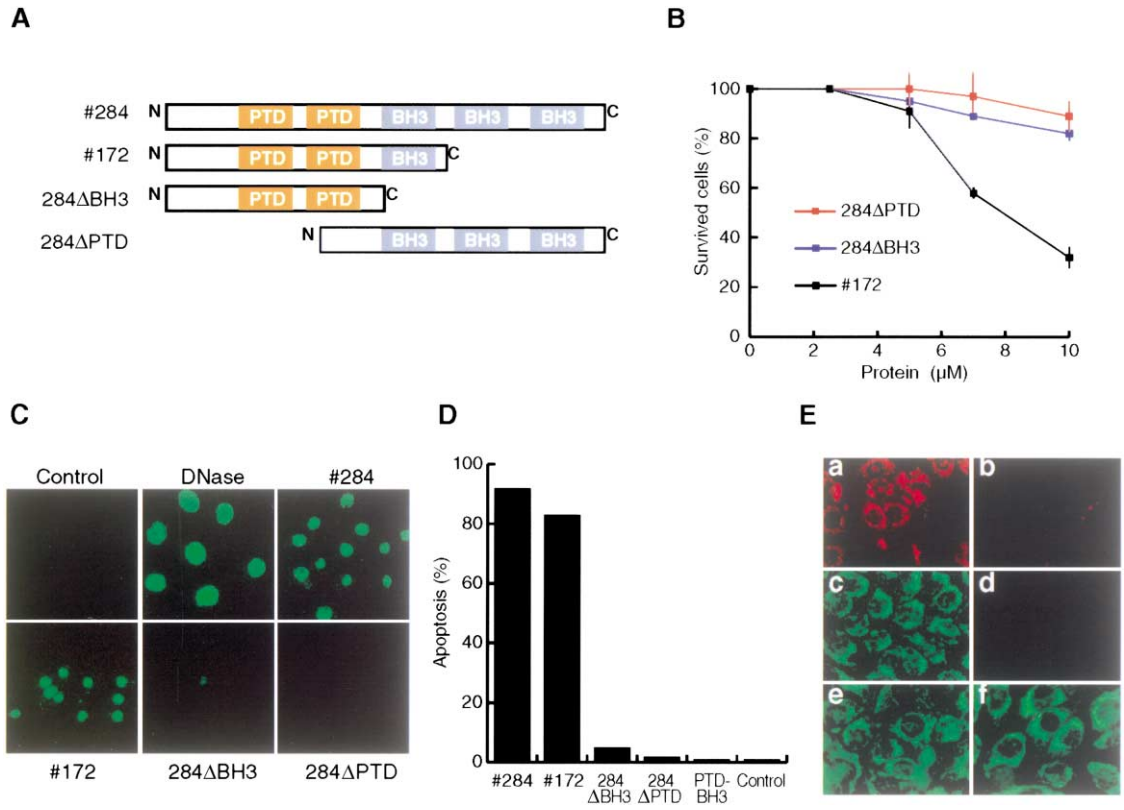

Figure 4. Mutational Assays

(A) Schematic drawing of mutant derivatives of 284.

(B) Effects of mutant proteins on cell viability. HBC-4 cells were incubated for 3 hr with the indicated concentration of the 284 derivatives, after which the metabolic activity was determined using WST-1 assays. Data shown are average values of two independent experiments. Bars represent observed values.

(C) Induction of apoptosis by 284 derivatives. HBC-4 cells were incubated for 20 hr with the indicated protein (10 μM), after which the apoptotic cells were visualized by TUNEL staining. As a positive control, the cells were treated with 64 mg/ml of DNaseI for 10 min at room temperature.

(D) The ratios of apoptotic cells in (C) were calculated as the number of TUNEL-positive cells per 300 cells counted. 50 μM PTD-$BH3_{pep}$ was used for PTD-BH3 treatment.

(E) Mitochondria membrane potential assessed by JC-1 and rhodamine 123 staining. Panels a and c, control HBC-4 cells stained with JC-1 or rhodamine 123, respectively; panels b and d, 284-treated cells (10 μM for 3 hr) stained with JC-1 or rhodamine 123, respectively; panel e, 284ΔBH3-treated cells (10 μM for 3 hr) stained by rhodamine 123; panel f, PTD-$BH3_{pep}$-treated cells (100 μM for 3 hr) stained with rhodamine 123.

from the protein (284ΔBH3 or 284ΔPTD, see Figure 4A), only minor impairment of cell growth was observed (Figure 4B). Consistent with these findings, TUNEL staining showed that both 284 and 172 induced apoptosis, whereas 284ΔBH3 and 284ΔPTD significantly reduced the proapoptotic function (Figures 4C and 4D). We also confirmed that 284ΔBH3 effectively entered the cells (Figure S2D), although it did not induce apoptosis. These results indicate that the $BH3^{Noxa}$ motif in 284 is indeed responsible for the induction of apoptosis.

BH3-only proteins have been shown to induce apoptosis by collapsing the mitochondrial membrane potential ($\Delta\Psi_m$) [17, 18], which was followed by caspases activation. Likewise, in the present study, measurements of $\Delta\Psi_m$ using the fluorescent probes JC-1 and rhodamine 123 showed that 284 drastically reduced $\Delta\Psi_m$, whereas 284ΔBH3 or PTD-$BH3_{pep}$ did not (Figures 4Ea–4Ef). Taken together, these results indicate that the $BH3^{Noxa}$ motif located in 284 is critical for causing the mitochondrial dysfunction caused by the potential loss of the membrane.

Functional Differences between Combinatorial Polymers

Although the conditions used for the initial characterization apparently dictated that 284 would be the only clone able to induce apoptosis in human hepatoma cell lines (Figure S2A), we found that other MG-27 polymers also exerted inhibitory effects on growth under different conditions. To further clarify the combinatorial nature of the motif-mixing experiment, we selected two polymers, 214 and 216, that share the same number of motifs (one PTD^{Tat} and one $BH3^{Noxa}$) and have similar pIs and molecular weights (Figure 5A and Table 1), but have different effects on cell growth and morphology, i.e., 216 exerted a much stronger inhibitory effect on growth of HBC-4 than 214 (Figures 5B and 5C). The TUNEL staining has indicated that 216, which has a stronger inhibitory effect on cell growth than 214, could induce apoptosis, whereas 214 could not (Figure 5D), confirming the functional differentiation in these two artificial proteins.

We initially considered the possibility that 214 might have entered the cell, but for some reason was unable

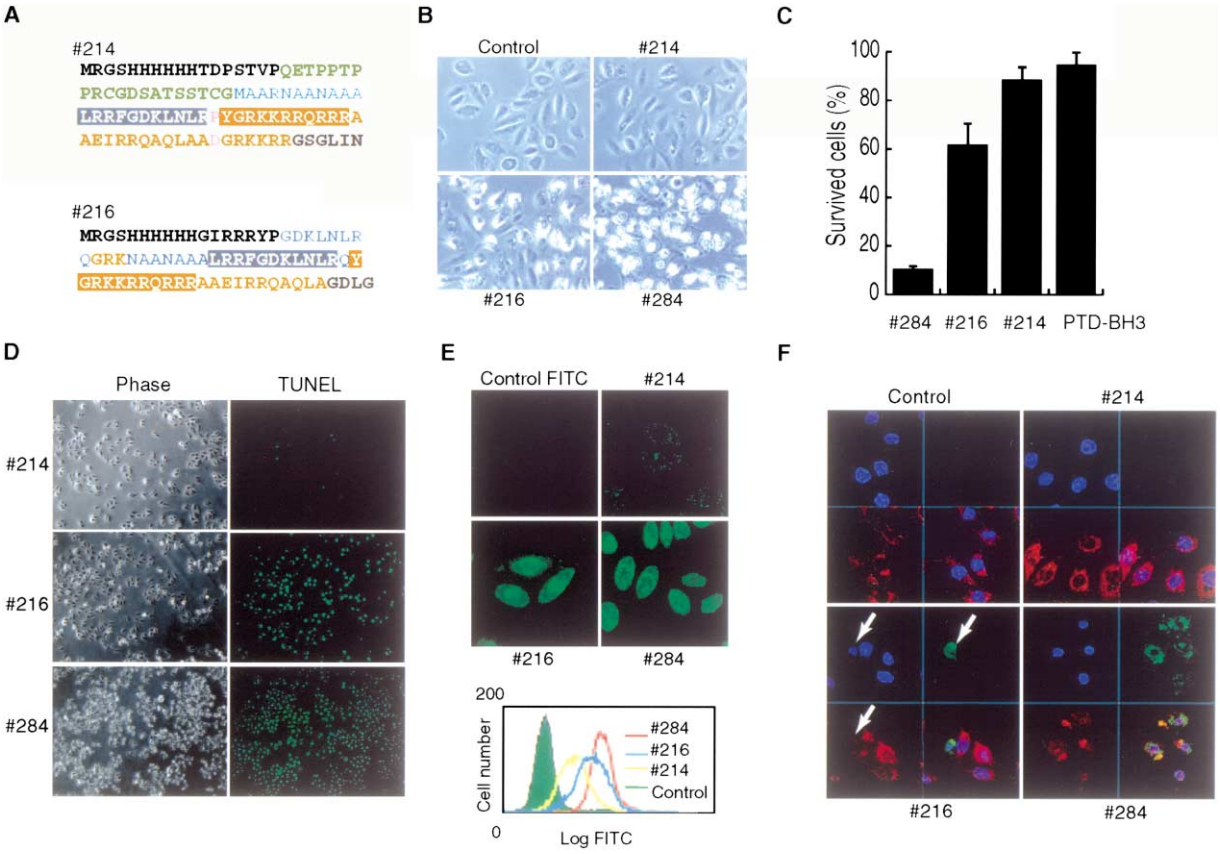

Figure 5. Characterization of 214 and 216

(A) Primary sequences of 214 and 216. See the legend to Figure 1 for coloring scheme.

(B) Morphological changes. HBC-4 cells were incubated for 1 hr with each protein (10 μM), fixed in 4% paraformaldehyde, and observed under a light microscope.

(C) Effect on cell viability. After incubating HBC-4 cells for 3 hr with 20 μM 284, 214, or 216, the metabolic activity was assessed using WST-1 assays. Data shown are average values of two independent experiments. Bars represent observed values.

(D) Induction of apoptosis. HBC-4 cells were treated with each protein (10 μM) for 1.5 hr and observed under a light microscope (left panels). TUNEL analysis was also carried out (right panels).

(E) Incorporation of FITC-labeled proteins. HBC-4 cells were incubated for 30 min with the indicated FITC-labeled protein, fixed in 4% paraformaldehyde, and analyzed by either confocal microscopy (top) or FACS (bottom).

(F) Effects of artificial proteins on the morphology of mitochondria and the nucleus. Cells of HBC-4 cells were incubated with the indicated protein (20 μM for 3 hr), fixed in 2% paraformaldehyde, and examined using a confocal laser microscope. Nuclei were stained with DAPI (blue, top left), proteins were immunodetected using penta-His monoclonal antibody conjugated with Alexa Fluor 488 (green, top right), and mitochondria were stained with CMTMRos (orange, bottom left). The three images were merged in the bottom right panel. Note that 214 does not penetrate through the cell membranes and maintains the mitochondrial membrane potential. In contrast, 284 and 216 (shown by white arrows) effectively penetrate through the membranes and induce the potential loss of the mitochondrial membrane and condensation of the nucleus.

to induce apoptosis, as was the case with PTD-$BH3_{pep}$ (Figure 1). We soon found, however, that 214's inability to induce apoptosis was actually attributable to an inability to translocate into cells. Whereas FITC-labeled 216 entered cells in a manner similar to 284, 214 showed much less ability to do so (Figure 5E). Multiple staining with DAPI (for nucleus), Mitotracker (for active mitochondria), and anti-penta His antibody conjugated with Alexa Fluor 488 (for artificial proteins) confirmed that 216 entered the cells, leading to a decrease in $\Delta\Psi_m$ and condensation of the nucleus, but this did not occur with 214 (Figure 5F). It is now known that the structural [19] or electrical properties [20] of the PTD motif, along with its flanking sequences, are key determinants of the motif's transductional activity. This may explain the difference in the protein transductional abilities observed in 214 and 216. Although further biochemical and structural experiments are required to elucidate the functional differences between the two proteins, it is clear that the combinatorial polymerization approach facilitates the isolation of proteins with the desired phenotypes from multiple peptide motifs.

Discussion

Combinatorial Approach for Generating Functional Proteins

In the present study, we created artificial proteins that can induce apoptosis in cancerous cell lines by mixing two naturally occurring peptide motifs, PTD^{Tat} and

$BH3^{Noxa}$, which are respectively related to protein transduction [15] and apoptosis induction [16]. The simple conjugant of these two motifs has been shown to penetrate into cells but fails to induce apoptosis (Figure 1), indicating that the activity of the $BH3^{Noxa}$ motif is somehow latent and proper configurations of the proteins are required to actualize the $BH3^{Noxa}$ activity. Several reports have already indicated that the isolated BH3 peptide had apoptotic activity by itself [5] or by conjugation with a protein transduction peptide [6, 7, 8]. Differences between these previous experiments and ours are as follows: (1) in previous reports, a longer peptide sequence (16- to 24-amino acids) was used as a BH3 motif, whereas we used a shorter 12-amino acid $BH3^{Noxa}$ motif that is comprised of the 9-amino acid BH3 core sequence and an extra tripeptide; (2) previous reports have extracted the BH3 motif from Bak [6], Bad [8], Bid [10], or Bax to make fusions with PTD, whereas we used the BH3 motif from Noxa, whose motif sequence somehow deviated from the other members [16]. Furthermore, a 20-amino acid peptide containing the Noxa BH3 domain has been shown to neither release cytochrome c from mitochondria nor interact with Bcl-2 [10]. We believe that these differences would underlie the latency of $BH3^{Noxa}$ that we used, and for this very reason we were able to investigate the combinatorial effects on the elicitation of the $BH3^{Noxa}$ activity in artificial proteins.

Among the artificial proteins, 284 has shown the strongest apoptotic activity accompanied by caspase activation (Figure 3) and the potential loss of the mitochondrial membrane (Figure 4E). These observations indicated that 284 activated the mitochondria pathway of apoptosis, where the BH3 motif-containing proteins play a pivotal role. Although 284 has three $BH3^{Noxa}$ motifs in its sequence, this multiplicity could not explain the elicitation of the latent activity of the $BH3^{Noxa}$, because one of the deletion derivatives of 284, 172, retained the apoptotic activity irrespective of its loss of the last two $BH3^{Noxa}$ motifs (Figure 4). Furthermore, a simpler polymer, 216, which has one PTD^{Tat} motif and one $BH3^{Noxa}$ motif, also exhibited a weak but significant proapoptotic function (Figure 5D). Interestingly, in simpler polymers, the combinatorics of three reading frames was profound, i.e., the 214 protein, which also has the same number of motifs as does 216, has been shown to be defective in inducing apoptosis (Figure 5D). In these proteins, sequences derived from the different reading frames were connected with the two motif sequences. In 214, a second reading frame (green, Figure 5A) precedes the PTD^{Tat}-$BH3^{Noxa}$ motifs, whereas 216 has a first reading frame (gray) and a part of the third reading frame (GRK, orange) in its N-terminal region. Moreover, in 214, a single mutation occurs between the two motifs (Figure 5A, P, pink). Although further investigations will be required to elucidate the phenotypic difference between 214 and 216, especially in terms of the relationship to their structures, these results showed that the presence of two motifs per se is not sufficient to reconstitute a bifunctional molecule, but the arrangements of the motifs in proper configurations would be critical for obtaining functional proteins. Thus, the combinatorial approach from a library composed of short sequence motifs is a promising way to create functional proteins.

A Hierarchal Approach for Protein Evolution

The synthesis of artificial proteins that regulate signal transduction pathways is a "bottom-up" approach for understanding dynamic behaviors of the signal network and its evolution [21]. It has already been shown that such a synthetic approach for nucleic acids has given us profound insights into the evolvability of RNA and DNA [22]. In these nucleic acid experiments, a pool of random sequences was first prepared by the combinatorial assemblages of four nucleotides from which novel molecules having significant catalytic activities such as tRNA aminoacylation activity [23] or RNA polymerization [24] have been selected. In contrast to the remarkable achievements of this "emergence from random sequence" approach in RNA evolution, the creation of novel proteins from random sequences has been limited to the one that has ATP binding activity [25], and we do not know whether a similar "random sequence" approach would be applicable for the in vitro synthesis of larger functional proteins.

The modular structures of extant proteins suggest that they evolved from duplications or recombination of genetic units corresponding to "exon" [26], "module" [27], or "domain." It is thus plausible that the evolution of proteins has proceeded in a hierarchical manner, i.e., first, a smaller primordial microgene, whose translational products (polypeptides) had only weak biological activities, arose from a random sequence, and then combinatorial assemblages of microgenes gave birth to the larger modern genes having more sophisticated activities. The relevance of the hypothesis of "a hierarchical evolution of proteins" has already been proven by theoretical experiments in which such an approach has been shown to allow a much more efficient search of the protein sequence space and to provide a greater likelihood of generating new protein folds [28]. Efforts are ongoing in many laboratories to establish a versatile in vitro protein synthesis system whose underling concept is a hierarchical evolution using a constrained DNA library rather than random sequences [14, 29–30]. Although "DNA shuffling," invented by Stemmer, is one such hierarchical evolution system and has been successfully used for the directed evolution of existing proteins [31], the assemblage between gene blocks that do not share any sequence similarity is difficult to achieve by this methodology. Methods to create artificial protein that rival the random sequence approach include the "constrained libraries" approach, in which a pool of biased sequences are used for starting libraries [32–33].

In this study, we started from smaller primordial units of sorts, peptide motifs, and created larger functional proteins by their combinatorial polymerization. For the polymerization, we first embedded two motifs in different reading frames of a single microgene, and then prepared tandem polymers of the microgene with genetic noise at the junctions. Translational products of the resultant microgene polymers were combinatorial polymers of three reading frames, because the noise at the junctions randomly changes the reading frames (Figure 2). This strategy unambiguously belongs to the hierarchal approach for an in vitro protein evolution system and is relevant to the constrained libraries approach. Although the number of motifs that can be embedded

into a single microgene would be limited, it has the advantage that the periodic property of created proteins contributes to the emergence of ordered proteins [14].

In conclusion, we developed a new strategy for synthesizing functional proteins by the combinatorial assemblage of short peptide motifs. Since we can incorporate desired peptide motifs in a single microgene, this approach, by combining with a high-throughput screening, would represent a novel system for creating multifunctional proteins. Accumulating information on natural peptide motifs in existing proteins as well as artificial peptides selected from peptide libraries should provide us with the flexibility to create proteins having a wide variety of functions.

Significance

In this study, we created artificial proteins that exert a proapoptotic effect when added to culture medium by combinatorially polymerizing two short peptide motifs respectively related to induction of apoptosis ($BH3^{Noxa}$) and protein transduction (PTD^{Tat}). Because simple linkage of these motifs was not sufficient to create a bifunctional peptide, and the successful reconstitution was dependent on how these motifs were joined together, the combinatorial polymerization strategy was shown to be important for reconstitution of function from mixtures of multiple short sequence motifs. Given that the effort to identify peptide motifs associated with protein functions or structures has been markedly accelerated by post-genomic research, the synthetic approach adopted in this study is undoubtedly relevant not only to basic science but also to applied science.

Experimental Procedures

Cell Culture

All cells except the HeLa cells were cultured in RPMI 1640 (Nissui) supplemented with 10% fetal bovine serum (Morigate) and antibiotic/antimycotic solution (SIGMA, A5955) at 37°C in humidified air containing 5% CO_2. Hela cells were cultured in Dulbecco's modified Eagle's medium (SIGMA) supplemented with 10% fetal bovine serum (JRH Biosciences). All assays were carried out in the presence of serum.

Construction and Purification of Artificial Proteins

A 72 bp microgene, MG-27 (Figure 2A), was designed as described in the text. Based on its sequence, two MPR primers, 5′-CGTATGG CCGCAAGAAACGCCGCCAACG CCGCCGCGCTGCA-3′ and 5′-CTGCCGCAAGTTGAGCTTGTCGCCGAATCTCC GCAGCGCGA-3′, were synthesized to have double mismatch pairs at their 3′-OH, after which the polymerization of MG-27 was carried out as previously described [13]. The resultant microgene polymers ligated into the SmaI site of pTZ19R.

For expression, BamHI-Asp718 fragments from pTZ19R, which contained the cloned microgene polymers, were subcloned into the BglII-Asp718 sites of vectors pKS600-605, each of which can translate one of six coding frames of the microgene polymers as an N-terminal His-tagged fusion protein. Proteins were expressed in the XL1Blue strain (Strategene, La Jolla) and were purified using TALON resin (Clontech, Palo Alto) as described previously [14]. The purified proteins run as a single band in SDS-PAGE and Coomassie Brilliant Blue staining (Figure S1B). Purified proteins were dialyzed against TNE buffer (50 mM Tris-acetate, 100 mM NaCl, 1 mM EDTA [pH 4.0]) and then kept frozen until their identities were confirmed by mass spectrometry (CIPHERGEN, USA). For the circular dichroism spectroscopy, proteins were dialyzed against 10 mM phosphate buffer (pH 7.0) at 4°C, after which 10 μM of each protein was used for analysis. The CD spectra were recorded at 5°C using a Jasco J-725.

Cell Proliferation Assay

For the trypan blue exclusion assay, HBC-4 cells (2.5×10^4 cells/well) were plated in 96-well plates. After incubating for 24 hr at room temperature, the cells were treated with 10 μM 284 or PTD-$BH3_{pep}$ and then stained with 0.2% trypan blue (SIGMA) in PBS. The number of blue cells and the total number of cells were counted under a microscope. For the WST-1 assay, cells (1×10^4 cells/well) were plated in 96-well plates. After incubating for 24 hr at room temperature, cells were treated with the indicated proteins, and the mitochondrial dehydrogenase activity, which served as an index of the cell metabolic activity, was evaluated using the tetrazolium salt WST-1 according to the manufacturer's (Roche) instructions. For the TUNEL staining analysis, cells were treated with 284 or its variants for the indicated times and then fixed in 4% paraformaldehyde. An in situ cell death detection kit (Roche) was then used according to the manufacturer's instructions. For the measurement of caspase-3-related activity, HBC-4 cells (4×10^5 cells/well) were preincubated for 24 hr, and the caspase activity was measured using a CPP32/Caspase3 Fluorometric Proteinase Assay Kit (MBL), in which a fluorogenic peptide, DEVD-7-amino-4-trifluoromethyl coumarin (DEVD-AFC), served as a substrate.

Assessment of Active Mitochondria

Changes in ψ_m were assessed by staining cells with the indicator dye, 5,5′,6,6′-tetrachloro-1,1′,3,3′,-tetraethylbenzimidazolocarbocyanine iodide (JC-1), or Rhodamine 123. Cells were incubated for 15 min with JC-1 (1μM) or Rhodamine 123 (10 μM) added directly to the culture medium, after which the medium was replaced with PBS, and the cells were observed under a fluorescence microscope. To analyze the state of the mitochondria after fixation, cells were incubated for 45 min with 100 nM MitoTracker Orange CMTMRos (Molecular Probes) added to the culture medium.

Supplemental Data

Supplemental Data including two figures and experimental procedures are available at http://www.chembiol.com/cgi/content/full/11/6/765/DC1.

Acknowledgments

We thank Drs. H. Suga and P. Schimmel for their critical proofreading. We also thank Drs. Y. Sirakata, R. Yao, K. Kumada, T. Mashima, T. Kawaguchi, and S. Futaki for their helpful comments on the experiments. This work was partly supported by a HFSP grant to K.S. H.S. acknowledges the JSPS Research Fellowships for Young Scientists for their generous support.

Received: January 20, 2004
Revised: March 4, 2004
Accepted: March 11, 2004
Published: June 25, 2004

References

1. Falquet, L., Pagni, M., Bucher, P., Hulo, N., Sigrist, C.J., Hofmann, K., and Bairoch, A. (2002). The PROSITE database, its status in 2002. Nucleic Acids Res. *30*, 235–238.
2. Schultz, J., Milpetz, F., Bork, P., and Ponting, C.P. (1998). SMART, a simple modular architecture research tool: identification of signaling domains. Proc. Natl. Acad. Sci. USA *95*, 5857–5864.
3. Opferman, J.T., and Korsmeyer, S.J. (2003). Apoptosis in the development and maintenance of the immune system. Nat. Immunol. *4*, 410–415.
4. Oltvai, Z.N., Milliman, C.L., and Korsmeyer, S.J. (1993). Bcl-2 heterodimerizes in vivo with a conserved homolog, Bax, that accelerates programmed cell death. Cell *74*, 609–619.
5. Moreau, C., Cartron, P.F., Hunt, A., Meflah, K., Green, D.R.,

Evan, G., Vallette, F.M., and Juin, P. (2003). Minimal BH3 peptides promote cell death by antagonizing anti-apoptotic proteins. J. Biol. Chem. *278*, 19426–19435.

6. Holinger, E.P., Chittenden, T., and Lutz, R.J. (1999). Bak BH3 peptides antagonize Bcl-xL function and induce apoptosis through cytochrome c–independent activation of caspases. J. Biol. Chem. *274*, 13298–13304.
7. Arnt, C.R., Chiorean, M.V., Heldebrant, M.P., Gores, G.J., and Kaufmann, S.H. (2002). Synthetic Smac/DIABLO peptides enhance the effects of chemotherapeutic agents by binding XIAP and cIAP1 in situ. J. Biol. Chem. *277*, 44236–44243.
8. Schimmer, A.D., Hedley, D.W., Chow, S., Pham, N.A., Chakrabartty, A., Bouchard, D., Mak, T.W., Trus, M.R., and Minden, M.D. (2001). The BH3 domain of BAD fused to the Antennapedia peptide induces apoptosis via its alpha helical structure and independent of Bcl-2. Cell Death Differ. *8*, 725–733.
9. Frugier, M., Giege, R., and Schimmel, P. (2003). RNA recognition by designed peptide fusion creates "artificial" tRNA synthetase. Proc. Natl. Acad. Sci. USA *100*, 7471–7475.
10. Letai, A., Bassik, M.C., Walensky, L.D., Sorcinelli, M.D., Weiler, S., and Korsmeyer, S.J. (2002). Distinct BH3 domains either sensitize or activate mitochondrial apoptosis, serving as prototype cancer therapeutics. Cancer Cell *2*, 183–192.
11. Frugier, M., and Schimmel, P. (1997). Subtle atomic group discrimination in the RNA minor groove. Proc. Natl. Acad. Sci. USA *94*, 11291–11294.
12. Burd, C.G., and Dreyfuss, G. (1994). Conserved structures and diversity of functions of RNA-binding proteins. Science *265*, 615–621.
13. Shiba, K., Takahashi, T., and Noda, T. (1997). Creation of libraries with long open reading frames by polymerization of a microgene. Proc. Natl. Acad. Sci. USA *94*, 3805–3810.
14. Shiba, K., Takahashi, Y., and Noda, T. (2002). On the role of periodism in the origin of proteins. J. Mol. Biol. *320*, 833–840.
15. Vocero-Akbani, A., Chellaiah, M.A., Hruska, K.A., and Dowdy, S.F. (2001). Protein transduction: delivery of Tat-GTPase fusion proteins into mammalian cells. Methods Enzymol. *332*, 36–49.
16. Oda, E., Ohki, R., Murasawa, H., Nemoto, J., Shibue, T., Yamashita, T., Tokino, T., Taniguchi, T., and Tanaka, N. (2000). Noxa, a BH3-only member of the Bcl-2 family and candidate mediator of p53-induced apoptosis. Science *288*, 1053–1058.
17. Jurgensmeier, J.M., Xie, Z., Deveraux, Q., Ellerby, L., Bredesen, D., and Reed, J.C. (1998). Bax directly induces release of cytochrome c from isolated mitochondria. Proc. Natl. Acad. Sci. USA *95*, 4997–5002.
18. Narita, M., Shimizu, S., Ito, T., Chittenden, T., Lutz, R.J., Matsuda, H., and Tsujimoto, Y. (1998). Bax interacts with the permeability transition pore to induce permeability transition and cytochrome c release in isolated mitochondria. Proc. Natl. Acad. Sci. USA *95*, 14681–14686.
19. Ho, A., Schwarze, S.R., Mermelstein, S.J., Waksman, G., and Dowdy, S.F. (2001). Synthetic protein transduction domains: enhanced transduction potential in vitro and in vivo. Cancer Res. *61*, 474–477.
20. Futaki, S., Suzuki, T., Ohashi, W., Yagami, T., Tanaka, S., Ueda, K., and Sugiura, Y. (2001). Arginine-rich peptides. An abundant source of membrane-permeable peptides having potential as carriers for intracellular protein delivery. J. Biol. Chem. *276*, 5836–5840.
21. Benner, S.A. (2003). Synthetic biology: Act natural. Nature *421*, 118.
22. Wilson, D.S., and Szostak, J.W. (1999). In vitro selection of functional nucleic acids. Annu. Rev. Biochem. *68*, 611–647.
23. Saito, H., Kourouklis, D., and Suga, H. (2001). An in vitro evolved precursor tRNA with aminoacylation activity. EMBO J. *20*, 1797–1806.
24. Johnston, W.K., Unrau, P.J., Lawrence, M.S., Glasner, M.E., and Bartel, D.P. (2001). RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. Science *292*, 1319–1325.
25. Keefe, A.D., and Szostak, J.W. (2001). Functional proteins from a random-sequence library. Nature *410*, 715–718.
26. Gilbert, W. (1978). Why genes in pieces? Nature *271*, 501.
27. Go, M. (1983). Modular structural units, exons, and function in chicken lysozyme. Proc. Natl. Acad. Sci. USA *80*, 1964–1968.
28. Bogarad, L.D., and Deem, M.W. (1999). A hierarchical approach to protein molecular evolution. Proc. Natl. Acad. Sci. USA *96*, 2591–2595.
29. Fisch, I., Kontermann, R.E., Finnern, R., Hartley, O., Solergonzalez, A.S., Griffiths, A.D., and Winter, G. (1996). A strategy of exon shuffling for making large peptide repertoires displayed on filamentous bacteriophage. Proc. Natl. Acad. Sci. USA *93*, 7761–7766.
30. Sieber, V., Martinez, C.A., and Arnold, F.H. (2001). Libraries of hybrid proteins from distantly related sequences. Nat. Biotechnol. *19*, 456–460.
31. Stemmer, W.P.C. (1994). Rapid evolution of a protein *in vitro* by DNA shuffling. Nature *370*, 389–391.
32. Wei, Y., Kim, S., Fela, D., Baum, J., and Hecht, M.H. (2003). Solution structure of a *de novo* protein from a designed combinatorial library. Proc. Natl. Acad. Sci. USA *100*, 13270–13273.
33. Wei, Y., and Hecht, M.H. (2004). Enzyme-like proteins from unselected library of designed amino acid sequences. Protein Eng. *17*, 67–75.
34. Bayer, P., Kraft, M., Ejchart, A., Westendorp, M., Frank, R., and Rosch, P. (1995). Structural studies of HIV-1 Tat protein. J. Mol. Biol. *247*, 529–535.
35. McDonnell, J.M., Fushman, D., Milliman, C.L., Korsmeyer, S.J., and Cowburn, D. (1999). Solution structure of the proapoptotic molecule BID: a structural basis for apoptotic agonists and antagonists. Cell *96*, 625–634.