Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Molecular Catalysis B: Enzymatic 28 (2004) 145–153
JOURNAL OF MOLECULAR CATALYSIS B: ENZYMATIC
www.elsevier.com/locate/molcatb

Review

MolCraft: a hierarchical approach to the synthesis of artificial proteins

Kiyotaka Shiba*

Department of Protein Engineering, Cancer Institute, Japanese Foundation for Cancer Research, and CREST, JST, 1-37-1 Kami-Ikebukuro, Toshima-ku, Tokyo 170-8455, Japan

Received 9 October 2003; received in revised form 5 December 2003; accepted 20 December 2003

Abstract

The modular structures of extant proteins and genes suggest that modern genes developed hierarchically from combinatorial assemblages of smaller primordial genetic units (microgenes). The **MolCraft** system described in this review is the new type of in vitro protein evolution system whose underlying concept is the hierarchical evolution of genes. In **MolCraft**, a microgene is initially evolved in silico and then tandemly polymerized with insertion or deletion mutations at the junctions between microgene units. Because of the junctional perturbations, proteins translated from a single microgene polymer are molecularly diverse, originating from the combinatorics of three reading frames, and are thus combinatorial polymers of three peptides. Notably, repetitiousness retained in the overall structure of proteins contributes to the formation of ordered structures, and enhances the chances of reconstituting biological activity rationally encrypted in the microgene unit. Applications of this new technology are discussed.

Keywords: Protein engineering; In vitro evolution; Functional motifs; Evolutionary molecular engineering; Combinatorial bioengineering

1. Introduction

The beginning of the 21st century is marked by the massive growth of information on the structures and functions of proteins coming from research into genomics and proteomics [1–3]. And as our knowledge about the relationship between protein structure and function grows, so grows the expectation of the rational synthesis of novel proteins. This ambition for rational creation of novel proteins is, however, a revisit to the "protein engineering" that rose to prominence at beginning of the 1980s, when analysis of protein crystal structures began in earnest, and recombinant technologies reached full maturity, and computational capacity reached new heights [4,5]. But although the fruits of protein engineering are certainly not negligible [6–12], one would have to say they have fallen far short of expectations. This is simply because deciphering the relationship between protein structure and function is an intricate problem and difficult to attack. On the other hand, the irrational approach known as "combinatorial bioengineering" has made remarkable progress over the past 10 years [13–19]. In one of the formulas for such combinatorial bioengineering, a large pool of random sequences is prepared by combinatorial polymerization of nucleotide units, after which functional clones are selected from a library. This methodology has shown its greatest utility in the de novo synthesis of ribozymes [20–27], RNA aptamer [28,29], DNA aptamer [30] and peptide aptamer [31–35], and has recently been applied to the de novo protein evolution [36].

Still, selection from random sequences has an inherent drawback that renders it incapable of taking advantage of the output of ongoing genome research: in this system molecules cannot be created in a rational manner. In the emerging genome-era, the emphasis is on "rationality" in protein creation, and the novel system for protein creation is expected to rationally utilize the accumulated knowledge obtained from genome research. It is within that context that I describe **MolCraft**, our approach to protein creation, in which a short rationally designed DNA sequence (microgene) is used instead of nucleotides as the building block for constructing combinatorial pools. This novel type of combinatorial bioengineering system is expected to make a significant contribution in creating new catalyst, bio-nano-materials and pharmaceuticals.

* Tel.: +81-3-5394-3903; fax: +81-3-5394-3903.
E-mail address: kshiba@jfcr.or.jp (K. Shiba).

1381-1177/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.molcatb.2003.12.027

2. A hierarchical approach to in vitro protein evolution

De novo evolution of functional proteins from a pool of random sequences is theoretically underpinned by the concept that genes originated from random sequences of nucleotides [37,38]. Although small, primordial genes might have arisen from random sequences, it is now widely accepted that larger modern genes evolved from assemblages of preexisting smaller primordial genetic units [39–44]. The "exon theory of genes" proposed by Gilbert in 1987 [45] is fully compatible with this notion. In the exon theory, it is proposed that polymerization of exons via their flanking introns ("exon shuffling") gave rise to the first set of genes. Although the engagement of exon shuffling in the early stages of gene evolution is a matter of debate [46,47], it is now evident from analytical studies of existing genes that exon shuffling did indeed play a pivotal role in evolving new genes in metazoan phyla [48,49]. Thus, a plausible scenario for protein evolution is a hierarchal one in which primordial microgenes endowed with rudimentary activities initially emerged from random nucleotide sequences or from repeats of short oligonucleotides, after which these microgenes served as building blocks for the larger, more exquisite genes that evolved from their combinatorial assemblages.

The advantages of a hierarchal approach are also apparent from a theoretical viewpoint. Bogarad and Deem compared the efficiencies of generating new protein folds using several evolution models in silico and observed that combinatorial linkage between preexisting gene blocks via nonhomologous DNA recombination, rearrangement, or desertion provided the greatest likelihood of generating new protein folds [50]. Fig. 1 depicts an idealized directed protein evolution system whose underlying concept is based on exon shuffling, or a hierarchical approach. In this system, small gene blocks (microgenes), rather than nucleotides, are used as building units, and molecularly diverse gene libraries are prepared by combinatorial polymerization of these microgenes.

Efforts are ongoing in many laboratories to establish a versatile system based on hierarchical evolution [51–63]. "DNA shuffling", in which genetic segments are shuffled among a family of related genes, enabling a sparser sampling of protein sequence space, is one such hierarchical evolution system [64–66]. Because this method is dependent on homologous recombination of DNAs, however, assemblage between gene blocks that do not share any sequence similarity is difficult to achieve. Therefore, several artifices have been proposed to address this problem: (i) shuffling RNA cassettes of exons using the *trans*-splicing activity of engineered group II introns [51]; (ii) assembling exons using the lox-Cre recombination system [52]; (iii) assembling exons by PCR from gene blocks pre-made so that they contain regions that partially [53,58] or completely [56,57] overlap regions in other blocks; (iv) using the "SCRATCHY" method [59], which enables creation of multiple-crossover DNA libraries in which nonhomologous genes are recombined using incremental truncation for creating hybrid enzymes ("ITCHY") [55] or sequence homology-independent protein recombination ("SHIPREC") [60], and the resultant fusions are subjected to DNA shuffling [64–66]; (v) assembling microgene blocks using thermostable DNA ligase and one or two guide oligonucleotides [61]; and (vi) using "Y-ligation", which uses T4 RNA ligase to ligate single stranded DNA blocks that form one stem and two branches (Y-shape) [54,62].

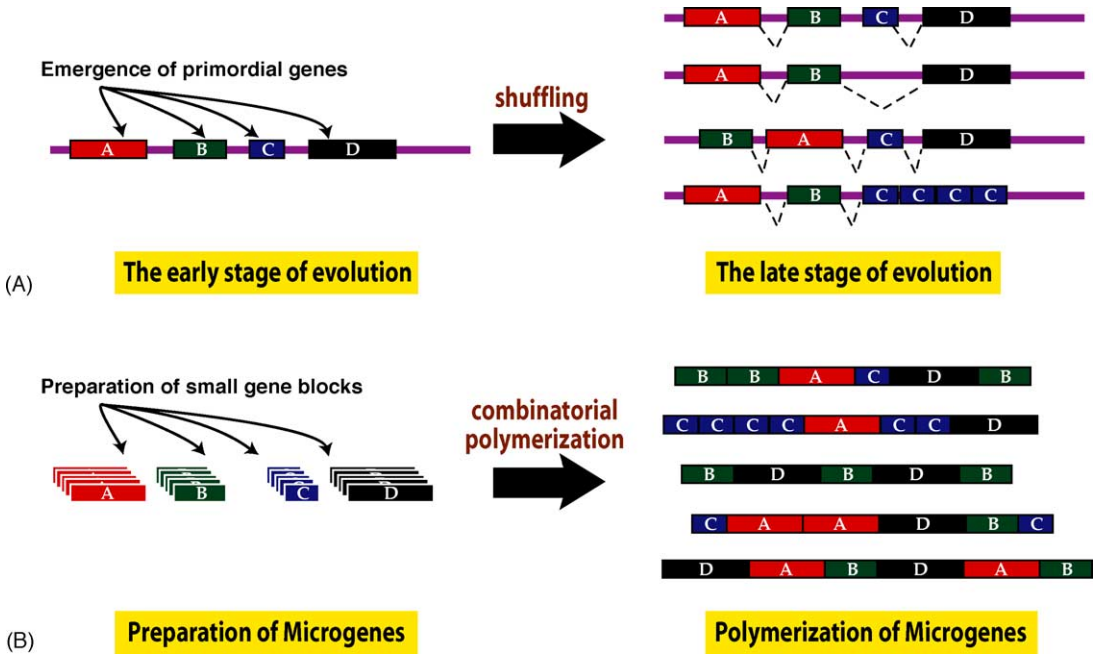

Fig. 1. The hierarchical evolution of proteins (A) and an in vitro protein evolution system that mimics hierarchical evolution (B).

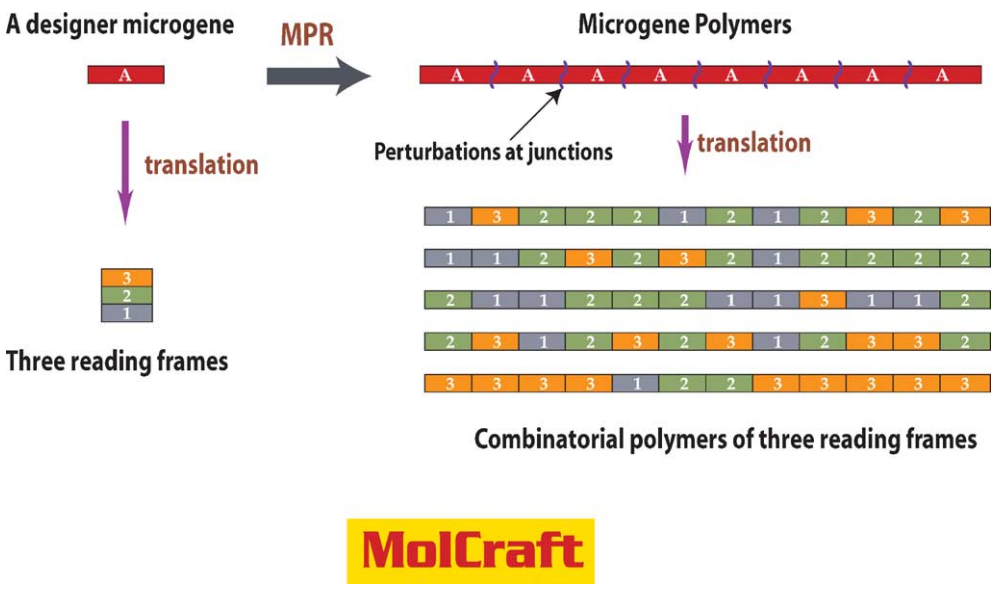

Fig. 2. Outline of protein evolution in **MolCraft**.

MolCraft is a simplified hierarchical protein evolution system in which a single gene block is used to make larger genes (Fig. 2). By limiting the number of gene blocks to one, we have provided this system with unparalleled advantages.

3. The contribution of repetitiousness in MolCraft

As mentioned above, one view of the origin of proteins is that they arose from a pool of random amino acid sequences [37,38]. According to this "emergence from random" hypothesis, small primordial genes would have evolved from random nucleotide sequences. On the other hand, Ohno recognized the repetitious nature of coding sequences and concluded that random sequences would be an unusual state for RNA or DNA. He therefore proposed his alternative "emergence from repeats" theory, in which he asserted that "periodic structures" formed during the course of semi-conservative replication are more reflective of the inherent nature of nucleic acids, and thus primordial genes more likely arose from repeats of base oligomers [67–70]. Two advantages of the emergence from repeats theory are that (i) genes would be relatively tolerant to frame shift mutations if they emerged from repeating sequences devoid of termination codons; and (ii) the proteins translated from repeated DNA sequences would likely have a higher propensity to form secondary structures [68]. The latter possibility is supported by an earlier experiment in which de novo-designed repetitive polypeptides, including $((GlyAla)_3GlyGlu)_n$, were shown to form stable secondary structures [71,72].

Recent innovations in genomics, proteomics and structural genomics have increased our appreciation of the scope of periodicity. Genome structures have been shown to contain a multitude of periodicities, ranging from those on a scale of one to several nucleotides [73] to those on a sub-chromosomal scale [74,75]. One surprising discovery to come out of the sequencing of the human genome was that various repeated sequences account for more than 50% of the genome [76]. Proteins also exhibit levels of repeats in their primary and tertiary structures. For instance, piscine antifreeze proteins consist of near exact copies of oligopeptide sequences [69], and the tertiary structure of porcine ribonuclease inhibitor shows near-perfect periodicity, apparently evolving from the reiteration of a DNA sequence 87 bp in length or less [77]. From these observations and others, we can safely conclude that "periodicity", or "reiteration", plays a critical role in the origin and evolution of proteins.

In my laboratory, we have been focusing on the periodicity observed in proteins and genes with the aim of evaluating its role in the birth of proteins. My coworkers and I have used our **MolCraft** system to create repetitious DNAs through polymerization of arbitrarily chosen, short DNA sequences (microgenes), and characterized their translation products [78,79]. Interestingly, although the proteins created with **MolCraft** are combinatorial polymers of three peptide sequences, they are inherently periodic and frequently have ordered structures, which is in sharp contrast to proteins with random sequences, which lack secondary structures.

4. Designer microgenes for rational protein construction

During the first stage of hierarchical evolution of proteins (Fig. 1A), the small primordial genes must have evolved rudimentary biological functionality [80]. Similarly, in protein evolution systems that mimic hierarchical evolution, it is desirable that microgene blocks are prepared so that their translation products will be related to biological activities or ordered structures. Exon units or module units found in nat-

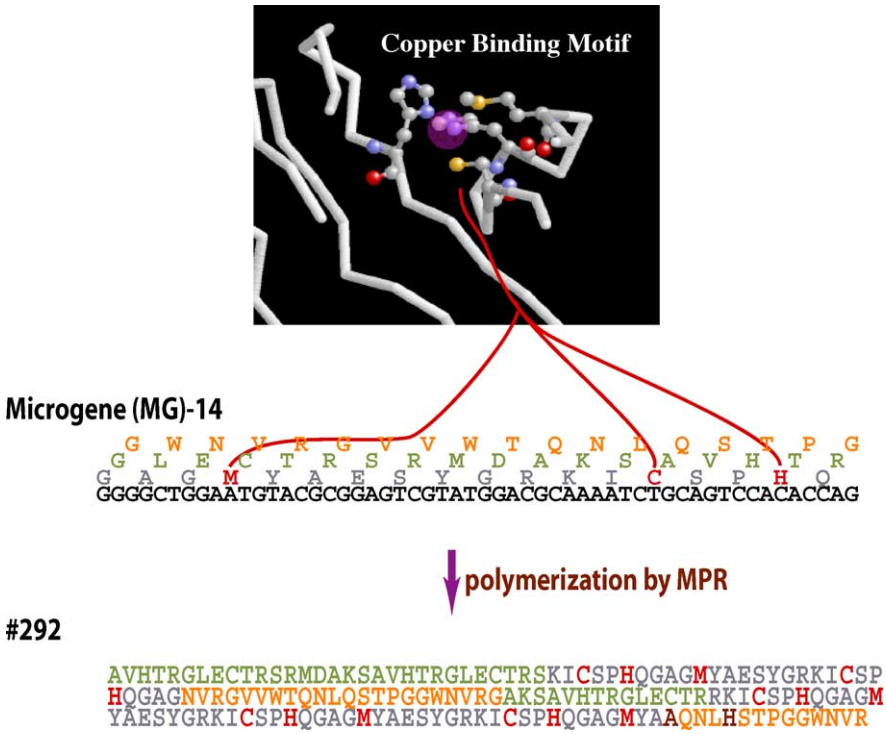

Fig. 3. Microgene MG-14, which is encrypted with a copper binding motif, and an artificial protein (#292) that was created by its polymerization.

ural proteins can serve as such building blocks [40,81–85]. Alternatively, we can operationally define the small genetic units through fragment complementation [61,86,87]. In **MolCraft**, moreover, a microgene is rationally designed with the help of computation so that the translation products from its three reading frames are related to biological functions and/or structures.

The most straight forward way to embed a specific function within a microgene is to apply "peptide motifs" [88,89] that have been previously determined to be associated with functions in natural proteins. For an instance, "H. . . [46aa] . . .CSPHQGAGM" has been identified as a copper binding motif in plastocyanin [90]. We therefore designed a microgene, MG-14 [78], that encodes the peptide GAGMYAESYGRKICSPHQ, which with head-to-tail concatenation, recreates part of the above copper-binding motif (Fig. 3). When artificial proteins were made by polymerizing MG-14, we found that some had an absorption maximum at 630 nm in the presence of copper ions, indicating association between the proteins and Cu^{2+} [79]. In addition to the embedding of peptide motifs, by designing microgenes so that its coded peptides have propensity to form secondary structures, we have also shown that structures can be encrypted in a microgene and reconstituted in **MolCraft** proteins [78,79].

Thus, with the hierarchal approach of the **MolCraft** system, a microgene is first rationally designed so that the encoded peptides are related to specific function(s) and/or structure(s), after which the designer microgene is polymerized to yield a library of larger genes that are combinatorial polymers of three reading frames (Fig. 2).

5. Exploitation of information in hidden reading frames

In **MolCraft**, a designer microgene is polymerized using a method called microgene polymerization reaction (MPR; Fig. 4) [91]. First, MPR primer pairs are designed from a custom-made microgene sequence. These primers contain complementary bases in their 3′ region and a mismatched base at their 3′-OH end, which enables efficient polymerization of the microgenes [91]. Thermal cycle reactions with MPR primers, the four dNTPs and thermostable exo^+ DNA polymerase yields head-to-tail polymers of the microgene in sizes that can be controlled by varying the reaction conditions. One intriguing feature of MPR is that the reaction randomly inserts or deletes nucleotides at end-joining junctions, thereby changing the reading frame of the polymers [91]. In other words, with **MolCraft**, all three reading frames coded by a single microgene are used to construct proteins.

To explore all the information possessed by a single microgene, we developed the microgene design program "CyberGene" [92]. Using CyberGene, we can design "multi-functional microgenes" as illustrated in Fig. 5. In this example, the peptide sequence Ile–Arg–Ile–Gln–Arg–Gly–Pro–Gly–Arg–Thr–Phe–Val–Thr serves as a primary motif to be embedded in a microgene. The first isoleucine can be

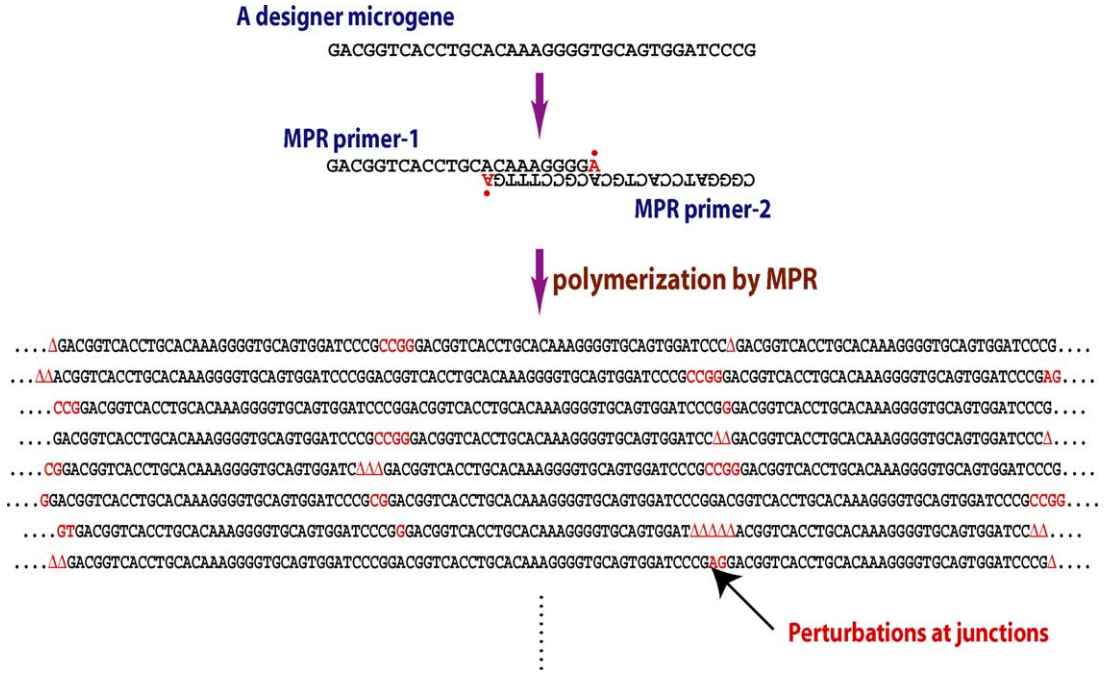

Fig. 4. Polymerization of a microgene using MPR.

coded by three triplet sequences, any of which can be used for a microgene sequence. Similarly, the second arginine can be coded by any of six triplet codons. In this manner, the total number of DNA sequences that can code the given 13-mer peptide is 5×10^8 ($3 \times 6 \times 3 \times 2 \times 6 \times 4 \times 6 \times 4 \times 2 \times 4 \times 4$). Among this pool of sequences, approximately 1.1×10^7 microgenes that did not have termination codons in each of its three reading frames were selected in the processor. Next, all the amino acid sequences that were encoded by the other two reading frames of the DNA sequence ($2 \times 1.1 \times 10^7 = 2.2 \times 10^7$) were constructed in the processor. After then excluding all duplicated sequences, approximately 1.5×10^7 variant peptide pools, each having a different sequence, were selected and assessed for their physicochemical properties. For instance, if we wanted to embed a propensity for α-helix formation, the pools were assessed for their potential to form α-helical structure using a secondary structure prediction program. Then a peptide,

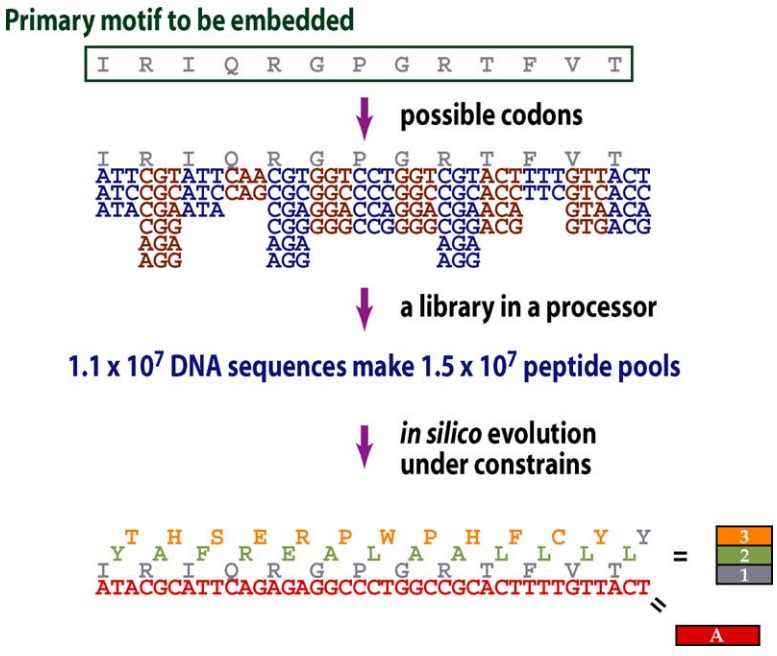

Fig. 5. Design of a multifunctional microgene using CyberGene.

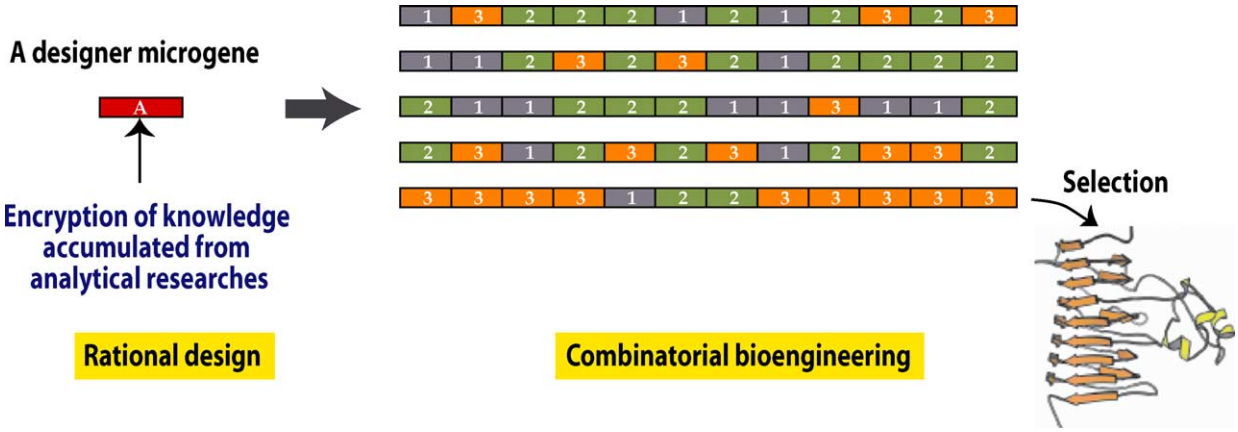

Fig. 6. Rational and irrational aspects of **MolCraft**.

Tyr–Ala–Phe–Arg–Glu–Ala–Leu–Ala–Ala–Leu–Leu–Leu–Leu, was selected from among the peptides that were predicted to have a high potential for forming α-helical secondary structure. This peptide was encoded by the second reading frame of the microgene, ATACGCATTCAGAGAGGCCCTGGCCGCACTTTTGTTACT, which was selected as the microgene to encode the aforementioned primary motif in its first reading frame.

In the above example, we focused only on the secondary structures of the pool of peptides coded by the hidden frames. However, we are able constrain virtually any physicochemical property (e.g., hydrophobicity, amino acid composition and isoelectric point, among many others) on the pool of peptide sequences and evolve the desired microgenes in silico.

6. Rationalism versus irrationalism

The design of a custom-made microgene represents the rational aspect of **MolCraft** and run in the blood protein engineering. We can extract sequences of functional motifs from natural proteins and embed them within a purposefully designed microgene, while at the same time encrypting biophysical characteristics into the microgene so that the proteins produced from microgene polymers have directed propensities. Despite this rationality, **MolCraft** also exhibits an irrational aspect—i.e., we cannot empirically predict the combinations of the three reading frames that will give the desired functions and structures. For instance, artificial proteins #320 and #334 were, respectively, created from combinations of $(5.5 \times \text{frame } 1) - (1 \times \text{frame } 3)$ and $(9.5 \times \text{frame } 1)$ of the designer microgene MG-15 [79]. Both proteins were α-helix-rich at neutral pH, but #320 lost its secondary structure at acidic pH, while #344 retained its secondary structure, even at pH 2.0. Furthermore, #344 precipitated at pH 8.0, whereas #320 remained soluble at that pH. Thus, variation in reading frame and/or the number of repeats can modulate the biochemical characteristics of microgene polymer proteins, in this case pH dependency. These characteristics cannot be rationally predicted from our knowledge of structure-function relationships. This lack of predictability can be overcome, however, by using a "selection" step, in which clones having desired functions and/or structures are chosen from the pools of combinatorial libraries of polymers of the three reading frames (Fig. 6). Thus, **MolCraft** simultaneously exhibits both the rationality of protein engineering and the irrationality of combinatorial bioengineering.

7. Applications in bio-nanotechnology and contributions in basic science

Motifs embedded within a microgene are not necessarily ones found in natural proteins. By virtue of the progress made in combinatorial engineering, peptide aptamers that specifically recognize target molecules are now routinely created using peptide-phage displaying systems. These systems were originally developed for isolating peptides that specifically recognize biomacromolecules (e.g., proteinous receptors) [31], but are now also being used to acquire peptide motifs that bind to inorganic materials, such as semiconductors [93,94], silver [95], carbon nanotube [96], carbon nanohorns [97] and titanium [98]. These peptide aptamers for inorganic materials expand of the applicable field for **MolCraft** from molecular biology to bio-nanotechnology [99]. In my laboratory, a new project is ongoing in which the functionalization of titanium (Ti) is the reason for using the **MolCraft** system [98]. Although titanous materials are now routinely used for dental implants and artificial joints, problems with their biocompatibility persist [100,101]. Our aim is to improve the biocompatibility of Ti surfaces using artificial proteins (Fig. 7). We plan to encrypt motifs that (i) bind to the surface of Ti; (ii) stimulate osteogenesis; and (iii) regulate biomineralization of calcium phosphate. Short peptide motifs that stimulate osteogenesis are being explored in natural proteins, such as bone morphogenetic

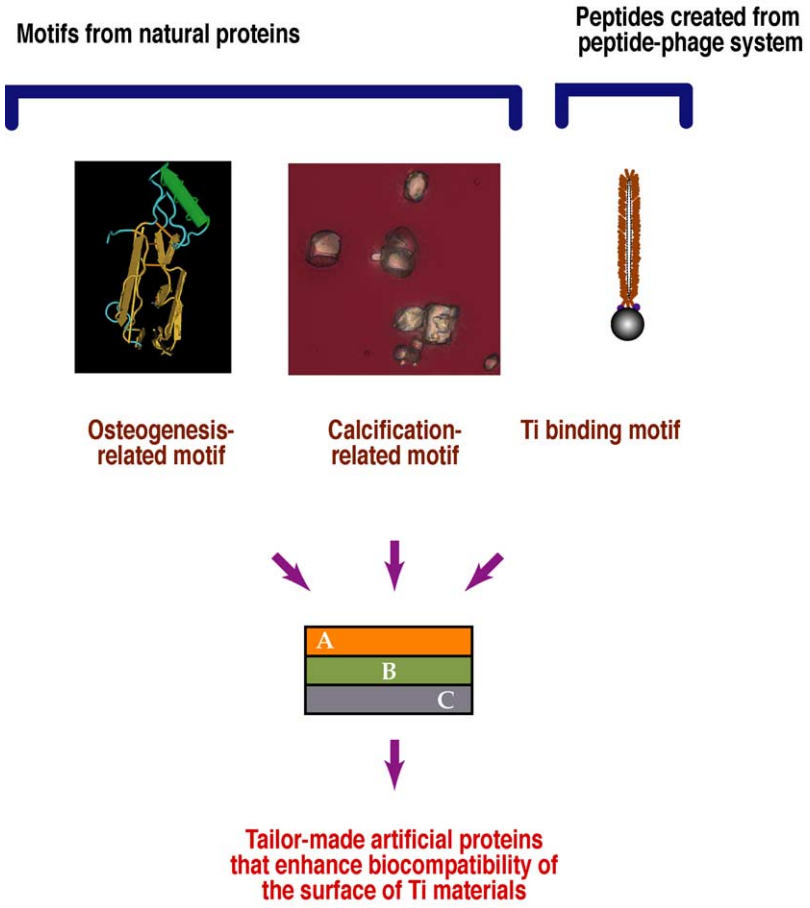

Fig. 7. Towards the creation of artificial proteins that enhance the biocompatibility of the surface of titanous materials.

proteins (BMPs) [102], and an artificial peptide that binds to the surface of Ti was recently isolated in our laboratory using a peptide-phage system [98]. We have also shown that a protein created from **MolCraft** modulates the growth of some crystals [103], suggesting the possibility that artificial proteins could be used to regulate calcification. By integrating these short sequences, multi-functional microgenes will be designed to create artificial proteins.

Creation of artificial proteins is a synthetic approach to understanding the dynamic behavior of biological systems and contributes to our understanding of the principles by which proteins and genes arose [104]. Moreover, this synthetic strategy should increasingly strengthen our understanding of such convoluted biological systems as genetic networks and signal transduction pathways [105]. By embedding motifs found in proteins involved in signal transduction, we are creating artificial proteins that can rewire signaling networks [106]. Clearly, the creation of artificial signaling proteins represents a new approach to the development of tailor-made medications. Thus, combinatorial engineering of artificial proteins shows potentially unparalleled promise for both applied and basic sciences.

Acknowledgements

My thanks to Drs. H. Saito, K. Kashiwagi, K. Sano, M. Nakamura, Y. Takaoka, D. Kase, T. Okamura, T. Murakami, T. Shirai and T. Noda, and to Ms. T. Minamisawa, T. Honma and Y. Takahashi for their contributions on the establishment of **MolCraft**.

References

[1] M. Tyers, M. Mann, Nature 422 (2003) 193.
[2] F.S. Collins, M. Morgan, A. Patrinos, Science 300 (2003) 286.
[3] F.S. Collins, E.D. Green, A.E. Guttmacher, M.S. Guyer, Nature 422 (2003) 835.
[4] K.M. Ulmer, Science 219 (1983) 666.
[5] C. Pabo, Nature 301 (1983) 200.
[6] B.I. Dahiyat, S.L. Mayo, Science 278 (1997) 82.
[7] L. Regan, W.F. DeGrado, Science 241 (1988) 976.
[8] M.H. Hecht, J.S. Richardson, D.C. Richardson, R.C. Ogden, Science 249 (1990) 884.
[9] M. Ramirez-Alvarado, F.J. Blanco, L. Serrano, Nat. Struct. Biol. 3 (1996) 604.
[10] T. Kortemme, M. RamirezAlvarado, L. Serrano, Science 281 (1998) 253.

[11] P.B. Harbury, J.J. Plecs, B. Tidor, T. Alber, P.S. Kim, Science 282 (1998) 1462.
[12] C.M. Kraemer-Pecore, A.M. Wollacott, J.R. Desjarlais, Curr. Opin. Chem. Biol. 5 (2001) 690.
[13] G.F. Joyce, Gene 82 (1989) 83.
[14] J.W. Szostak, Trends Biochem. Sci. 17 (1992) 89.
[15] J.D. Marks, H.R. Hoogenboom, A.D. Griffiths, G. Winter, J. Biol. Chem. 267 (1992) 16007.
[16] S.A. Kauffman, J. Theor. Biol. 157 (1992) 1.
[17] M.A. Gallop, R.W. Barrett, W.J. Dower, S.P.A. Fodor, E.M. Gordon, J. Med. Chem. 37 (1994) 1233.
[18] D.J. Kenan, D.E. Tsai, J.D. Keene, Trends Biochem. Sci. 19 (1994) 57.
[19] M.B. Zwick, J.Q. Shen, J. Scott, Curr. Opin. Biotechnol. 9 (1998) 427.
[20] J.R. Lorsch, J.W. Szostak, Biochemistry 33 (1994) 973.
[21] J.R. Prudent, T. Uno, P.G. Schultz, Science 264 (1994) 1924.
[22] E.H. Ekland, J.W. Szostak, D.P. Bartel, Science 269 (1995) 364.
[23] M. Illangasekare, G. Sanchez, T. Nickles, M. Yarus, Science 267 (1995) 643.
[24] X. Dai, A. De Mesmaeker, G.F. Joyce, Science 267 (1995) 237.
[25] C. Wilson, J.W. Szostak, Nature 374 (1995) 777.
[26] N. Lee, Y. Bessho, K. Wei, J.W. Szostak, H. Suga, Nat. Struct. Biol. 7 (2000) 28.
[27] H. Saito, D. Kourouklis, H. Suga, EMBO J. 20 (2001) 1797.
[28] A.D. Ellington, J.W. Szostak, Nature 346 (1990) 818.
[29] M. Sassanfar, J.W. Szostak, Nature 364 (1993) 550.
[30] L.C. Bock, L.C. Griffin, J.A. Latham, E.H. Vermaas, J.J. Toole, Nature 355 (1992) 564.
[31] J.K. Scott, G.P. Smith, Science 249 (1990) 386.
[32] K. Harada, S.S. Martin, A.D. Frankel, Nature 380 (1996) 175.
[33] R. Pasqualini, E. Ruoslahti, Nature 380 (1996) 364.
[34] P. Colas, B. Cohen, T. Jessen, I. Grishina, J. Mccoy, R. Brent, Nature 380 (1996) 548.
[35] M. Frugier, P. Schimmel, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 11291.
[36] A.D. Keefe, J.W. Szostak, Nature 410 (2001) 715.
[37] O.B. Ptitsyn, M.V. Volkenstein, J. Biomol. Struct. Dyn. 4 (1986) 137.
[38] S.H. White, R.E. Jacobs, J. Mol. Evol. 36 (1993) 79.
[39] M.G. Rossmann, D. Moras, K.W. Olsen, Nature 250 (1974) 194.
[40] M. Go, Nature 291 (1981) 90.
[41] M. Jasin, L. Regan, P. Schimmel, Nature 306 (1983) 441.
[42] L. Patthy, Curr. Opin. Struct. Biol. 1 (1991) 351.
[43] M. Baron, D.G. Norman, I.D. Campbell, Trends Biochem. Sci. 16 (1991) 13.
[44] R.F. Doolittle, Annu. Rev. Biochem. 64 (1995) 287.
[45] W. Gilbert, Cold Spring Harbor Symp. Quant. Biol. 52 (1987) 901.
[46] J.M. Logsdon, M.G. Tyshenko, C. Dixon, J.D. Jafari, V.K. Walker, J.D. Palmer, Proc. Natl. Acad. Sci. U.S.A. 92 (1995) 8507.
[47] S.J. de Souza, M. Long, R.J. Klein, S. Roy, S. Lin, W. Gilbert, Proc. Natl. Acad. Sci. U.S.A. 95 (1998) 5094.
[48] L. Patthy, Gene 238 (1999) 103.
[49] M. Long, Curr. Opin. Genet. Dev. 11 (2001) 673.
[50] L.D. Bogarad, M.W. Deem, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 2591.
[51] S. Mikheeva, K.A. Jarrell, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 7486.
[52] I. Fisch, R.E. Kontermann, R. Finnern, O. Hartley, A.S. Solergonzalez, A.D. Griffiths, G. Winter, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 7761.
[53] A. Crameri, S. Cwirla, W.P.C. Stemmer, Nat. Med. 2 (1996) 100.
[54] K. Nishigaki, K. Taguchi, Y. Kinoshita, T. Aita, Y. Husimi, Mol. Divers. 4 (1998) 187.
[55] M. Ostermeier, J.H. Shim, S.J. Benkovic, Nat. Biotech. 17 (1999) 1205.
[56] T. Tsuji, K. Kobayashi, H. Yanagawa, FEBS Lett. 453 (1999) 145.
[57] T. Tsuji, M. Onimaru, H. Yanagawa, Nucleic Acids Res. 29 (2001) E97.
[58] J.A. Kolkman, W.P. Stemmer, Nat. Biotech. 19 (2001) 423.
[59] S. Lutz, M. Ostermeier, G.L. Moore, C.D. Maranas, S.J. Benkovic, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 11248.
[60] V. Sieber, C.A. Martinez, F.H. Arnold, Nat. Biotech. 19 (2001) 456.
[61] K. Shiba, T. Hatada, Y. Takahashi, T. Noda, J. Biochem. (Tokyo) 132 (2002) 689.
[62] K. Kitamura, Y. Kinoshita, S. Narasaki, N. Nemoto, Y. Husimi, K. Nishigaki, Protein Eng. 15 (2002) 843.
[63] K. Kashiwagi, K. Shiba, J. Mol. Catal. B (2003) xxx.
[64] W.P.C. Stemmer, Nature 370 (1994) 389.
[65] W.P.C. Stemmer, BioTechnology 13 (1995) 549.
[66] A. Crameri, S.A. Raillard, E. Bermudez, W.P.C. Stemmer, Nature 391 (1998) 288.
[67] S. Ohno, Proc. Natl. Acad. Sci. U.S.A. 78 (1981) 7657.
[68] S. Ohno, J.T. Epplen, Proc. Natl. Acad. Sci. U.S.A. 80 (1983) 3391.
[69] S. Ohno, J. Mol. Evol. 20 (1984) 313.
[70] S. Ohno, J. Mol. Evol. 25 (1987) 325.
[71] M.T. Krejchi, E.D.T. Atkins, A.J. Waddon, M.J. Fournier, T.L. Mason, D.A. Tirrell, Science 265 (1994) 1427.
[72] K.P. McGrath, M.J. Fournier, T.L. Mason, D.A. Tirrell, J. Am. Chem. Soc. 114 (1992) 727.
[73] G.R. Sutherland, R.I. Richards, Proc. Natl. Acad. Sci. U.S.A. 92 (1995) 3636.
[74] K.H. Wolfe, D.C. Shields, Nature 387 (1997) 708.
[75] J.A. Bailey, Z. Gu, R.A. Clark, K. Reinert, R.V. Samonte, S. Schwartz, M.D. Adams, E.W. Myers, P.W. Li, E.E. Eichler, Science 297 (2002) 1003.
[76] E.S. Lander, L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, et al., Nature 409 (2001) 860.
[77] B. Kobe, J. Deisenhofer, Nature 366 (1993) 751.
[78] K. Shiba, Y. Takahashi, T. Noda, J. Mol. Biol. 320 (2002) 833.
[79] K. Shiba, T. Shirai, T. Honma, T. Noda, Protein Eng. 16 (2003) 57.
[80] D.L. Rohlfing, S.W. Fox, Adv. Catal. 20 (1969) 373.
[81] I. Kumagai, S. Takeda, K.-i. Miura, Proc. Natl. Acad. Sci. U.S.A. 89 (1992) 5887.
[82] K. Wakasugi, K. Ishimori, K. Imai, Y. Wada, I. Morishima, J. Biol. Chem. 269 (1994) 18750.
[83] K. Shiba, in: M. Go, P. Schimmel (Eds.), Tracing Biological Evolution in Protein and Gene Structures, Elsevier, Amsterdam, 1995, p. 11.
[84] K. Inaba, K. Wakasugi, K. Ishimori, T. Konno, M. Kataoka, I. Morishima, J. Biol. Chem. 272 (1997) 30054.
[85] T. Tsuji, K. Yoshida, A. Satoh, T. Kohno, K. Kobayashi, H. Yanagawa, J. Mol. Biol. 286 (1999) 1581.
[86] K. Shiba, P. Schimmel, Proc. Natl. Acad. Sci. U.S.A. 89 (1992) 1880.
[87] K. Shiba, P. Schimmel, J. Biol. Chem. 267 (1992) 22703.
[88] J. Schultz, F. Milpetz, P. Bork, C.P. Ponting, Proc. Natl. Acad. Sci. U.S.A. 95 (1998) 5857.
[89] L. Falquet, M. Pagni, P. Bucher, N. Hulo, C.J. Sigrist, K. Hofmann, A. Bairoch, Nucleic Acids Res. 30 (2002) 235.
[90] E.T. Adman, Curr. Opin. Struct. Biol. 1 (1991) 895.
[91] K. Shiba, T. Takahashi, T. Noda, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 3805.
[92] K. Shiba, in preparation.
[93] S.R. Whaley, D.S. English, E.L. Hu, P.F. Barbara, A.M. Belcher, Nature 405 (2000) 665.
[94] S.W. Lee, C. Mao, C.E. Flynn, A.M. Belcher, Science 296 (2002) 892.
[95] R.R. Naik, S.J. Stringer, G. Agarwal, S.E. Jones, M.O. Stone, Nat. Mater. 1 (2002) 169.
[96] S. Wang, E.S. Humphreys, S.Y. Chung, D.F. Delduco, S.R. Lustig, H. Wang, K.N. Parker, N.W. Rizzo, S. Subramoney, Y.M. Chiang, A. Jagota, Nat. Mater. 2 (2003) 196.

[97] J. Zhu, D. Kase, K. Shiba, D. Kasuya, M. Yudasaka, S. Iijima, Nano Lett. 3 (2003) 1033.
[98] K. Sano, K. Shiba, J. Am. Chem. Soc. 125 (2003) 14234.
[99] M. Sarikaya, C. Tamerler, A.K. Jen, K. Schulten, F. Baneyx, Nat. Mater. 2 (2003) 577.
[100] N. Sykaras, A.M. Iacopino, V.A. Marker, R.G. Triplett, R.D. Woody, Int. J. Oral Maxillofac. Implants 15 (2000) 675.
[101] D.A. Puleo, A. Nanci, Biomaterials 20 (1999) 2311.
[102] Y. Suzuki, M. Tanihara, K. Suzuki, A. Saitou, W. Sufan, Y. Nishimura, J. Biomed. Mater. Res. 50 (2000) 405.
[103] K. Shiba, T. Honma, T. Minamisawa, K. Nishiguchi, T. Noda, EMBO Rep. 4 (2003) 148.
[104] K. Shiba, J. Biochem. Mol. Biol. 31 (1998) 209.
[105] S.A. Benner, Nature 421 (2003) 118.
[106] H. Saito, T. Honma, T. Minamisawa, K. Yamazaki, T. Noda, T. Yamori, K. Shiba, Chem. Biol. (2004) in press.