FEBS Letters 586 (2012) 918–923

FEBS Letters

journal homepage: www.FEBSLetters.org

Rational design of DNA sequence-specific zinc fingers

Hidetoshi Kono [a,b,*], Miki Imanishi [b,c], Shigeru Negi [d], Kazuya Tatsutani [c], Yui Sakaeda [d], Ayaka Hashimoto [d], Chie Nakayama [d], Shiroh Futaki [c], Yukio Sugiura [d]

[a] Molecular Modeling and Simulation, Quantum Beam Science Directorate, Japan Atomic Energy Agency, Umemidai, Kizugawa, Kyoto 619-0215, Japan
[b] PRESTO, Japan Science and Technology Agency, Saitama 332-0012, Japan
[c] Institute for Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan
[d] Faculty of Pharmaceutical Sciences, Doshisha Women's University, Koudo, Kyotanabe-Shi, Kyoto 610-0395, Japan

ARTICLE INFO

Article history:
Received 7 November 2011
Revised 10 February 2012
Accepted 15 February 2012
Available online 23 February 2012

Edited by Robert B. Russell

Keywords:
DNA-binding protein
Protein design
Specificity
Affinity
Energy gap

ABSTRACT

We developed a rational scheme for designing DNA binding proteins. The scheme was applied for a zinc finger protein and the designed sequences were experimentally characterized with high DNA sequence specificity. Starting with the backbone of a known finger structure, we initially calculated amino acid sequences compatible with the expected structure and the secondary structures of the designed fingers were then experimentally confirmed. The DNA-binding function was added to the designed finger by reconsidering a section of the amino acid sequence and computationally selecting amino acids to have the lowest protein–DNA interaction energy for the target DNA sequences. Among the designed proteins, one had a gap between the lowest and second lowest protein–DNA interaction energies that was sufficient to give DNA sequence-specificity.

1. Introduction

Protein–DNA interactions are involved in a variety of fundamental processes, including transcription, recombination and DNA repair. DNA binding proteins, as exemplified by transcription factors, show high affinity and specificity for their target DNA sequences. To elucidate the mechanism underlying specific DNA recognition by DNA binding proteins, much effort has gone into determining the detailed structures of protein–DNA complexes. These structures suggest that there is not a perfect one-to-one correspondence between amino acids and bases in these interactions, though some preferences are indicated [1], in particular, zinc finger proteins show a strong preference [2]. To quantitatively evaluate the strength of protein–DNA interactions, we have developed statistical potential functions based on the spatial distributions of amino acids interacting with each base. Using these statistical potentials, we are able to discriminate target from non-target DNA sequences [3]. In addition, the potentials can also be used to find amino acids that will favorably interact with a given DNA sequence within the framework of a particular protein–DNA complex, and to design DNA binding proteins [3].

Recent successes in the computational design of various given backbone structures are remarkable [4–6]. The next goal is to design a protein with a desired function. One of the authors has developed a computer program called SCAD (statistical, computationally assisted design) for generating a sequence profile that fold into a desired backbone structure [7]. Thus far, SCAD has been successfully applied to design a four helix-bundle from a single chain [8], to make soluble a membrane protein [9] to design a thioredoxin mimic with redox activity [10] and to design an thermostable terpene syntase [11]. Here we report a rational scheme for designing DNA-binding proteins. The scheme was applied for designing of artificial zinc-finger proteins with specific DNA binding abilities and the designed amino acid sequences were experimentally validated by their secondary structures and their DNA binding abilities.

2. Materials and methods

2.1. Calculating a sequence profile for a zinc finger protein

We first used SCAD [7] to generate a sequence profile, a set of amino acid sequences that were compatible with the zinc finger structure (ββα), based on the backbone geometry of the second finger of Zif268 (PDB ID: 1AAY) [12]. Zif268 was used because its high-resolution X-ray structure is available, the structural basis

* Corresponding author at: Molecular Modeling and Simulation, Quantum Beam Science Directorate, Japan Atomic Energy Agency, Umemidai, Kizugawa, Kyoto 619-0215, Japan. Fax: +81 774 71 3460.
E-mail address: kono.hidetoshi@jaea.go.jp (H. Kono).

0014-5793/$36.00
doi:10.1016/j.febslet.2012.02.025

of its DNA binding mode is well-characterized [2,12–16] and its compact structure is amenable to computational examination. SCAD outputs probabilities for 20 amino acid residues at each position in the backbone, along with a backbone-dependent rotamer library [17] that is subject to certain constraints for fixing/limiting amino acids or their conformations based on prior information. From this output, we could easily model a tertiary structure by replacing the side chains of the second finger of Zif268 with the calculated ones. The amino acid sequences for the first and third fingers of Zif268 were retained. Because the spatial arrangement of the secondary structures generally depends on the loop structures connecting them, we imposed a constraint fixing the amino acid types within the loop structures, but allowed the conformation of the side chains to change. The amino acids for the loop and the second β-strand regions were assigned to be the ones most conserved in the zinc-finger fold according to the HSSP [18]. These sequences were determined to be EKPF (positions 1–4) for the arrangement of fingers 1 and 2, and GRAFS (positions 10–14) for the second β-strand and the loop to the α-helix within finger 2 (Appendix Fig. A1). We also fixed the Zn(II) ligands and their conformations to those of Zif268. Consequently, 15 of the 28 residues were allowed to freely mutate in the backbone geometry of the Zif268–DNA complex. The SCAD program was then used to calculate the amino acid sequences compatible with the backbone. The output of the program was a set of amino acid probabilities for each of the 15 positions. Amino acids with non-zero probability indicate that they are acceptable at the position. Among them, we picked up the most probable amino acid at each of 15 positions and obtained the representative amino acid sequence for the given geometry.

2.2. Designing the DNA-binding ability

Next, we considered to create the DNA-binding ability to amino acid sequence. In general, the amino acid preferences at positions exposed to the solvent tend to be much weaker than those in the protein interior [7]. This situation may allow us to mutate amino acids exposed to the solvent with little or no effect on the backbone structure. In fact, most of the residues that mediate specific binding of Zif268 to DNA are located at the protein surface, and the SCAD allowed these amino acids to vary considerably. For that reason, we carried out another calculation to introduce a specific DNA binding function into the T0 peptide.

Two key issues affecting protein–DNA interaction are affinity and specificity, which do not necessarily correlate with one another. Phage display methods have been used to select zinc finger proteins with different DNA-binding specificities. This approach enabled us to optimize the amino acid sequences to increase their affinity for target DNA, but it cannot completely guarantee sequence specificity because the optimized zinc finger has the potential to bind sequences other than the target. To achieve both high affinity and high specificity in DNA binding, we set up two criteria for selection of DNA-binding amino acids, as is exemplified by S_i in Fig. 1. First, the interaction energy should be low enough to ensure high affinity i.e., almost as low as with the wild-type protein. Second, the gap between the lowest interaction energy (target) and the second lowest (non-target) should be large enough to ensure high specificity.

2.3. CD spectrum measurements

CD experiments were performed on JASCO J-720 instrument with a temperature controller, using a capped 0.1 cm path length cuvette. The spectra were recorded from 195 to 260 nm by using continuous mode with 1 nm bandwidth, a 1 s response, and a scan speed of 50 nm min^{-1}. Each spectrum represents the average of twenty scans at 20 °C under a nitrogen atmosphere. Each peptide was dissolved in Tris–HCl buffer (10 mM, pH 7.5) containing NaCl (50 mM) and the concentration of peptide stock solutions were estimated by spectrophotometrically. All CD samples were prepared to 25 μM.

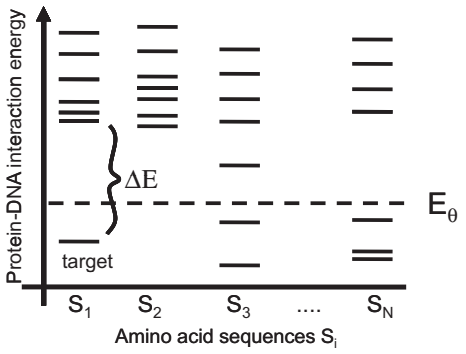

Fig. 1. Amino acid sequence selection criteria. For each amino acid sequence, the distribution of interaction energies for the DNA sequences is plotted. Ideal amino sequences should meet the following two criteria: (1) the protein–DNA interaction energy for one DNA sequence is sufficiently low (bellow E_θ) to achieve high binding affinity, and (2) there is an energy gap (ΔE) between the DNA sequences with the lowest and second lowest interaction energies that is sufficient to achieve high specificity. In the figure, S_1 is an ideal case.

2.4. Gel-shift assay

We measured the binding affinity for the target DNA sequences and sequence selectivity by gel-shift assay. The assay was carried out with the following conditions: 2.5 nM FITC-labeled DNAs, 10 mM Tris–HCl (pH 8.0), 50 mM NaCl, 1 mM DTT, 10 μM $ZnCl_2$, 20 ng/μL calf thymus DNA, 0.05% Nonidet P40, 5% glycerol, 40 ng/μL BSA, reaction time 30 min, electrophoresis in 8% polyacrylamide gel, 88 mM Tris–borate buffer. For rigorous examination of the binding selectivity, we carried out the gel-shift assays using 64 dsDNAs covering all possible triplet DNAs: 5′-TAG-TGGATCCGCGNNNGCGTGAATTCGA-3′ where NNN is the target triplet DNA.

3. Results and discussion

3.1. Design of amino acids folding into a zinc finger together with DNA-binding function

Using the SCAD program, we obtained a sequence profile for the zinc finger and the following sequence as a representative by selecting the most probable amino acid at each site; EKPFQCIIC-GRAFSKKWELKKHIFKHKG (we refer to this as T0, hereafter). For the 15 designed regions, T0 sequence has 33% (5/15) identity to the wild type (Fig. 2a). We emphasize that this sequence is just a representative of the sequences that are compatible with the given backbone geometry of zinc finger. Fig. 2b, for example, shows a part of the profile indicating that many amino acids are acceptable for some of the positions (positions 16, 17, 18 and 21) and different positions have a different amino acid preference. We first experimentally check the secondary structure of T0 and confirmed that the secondary structure of T0 be similar to that of wild-type peptide upon addition of Zn(II) by measuring CD spectra (Fig. 3). It is worth noting that in spite of its sequence identity of 33% for the designed section to the wild type T0 has a similar amount of secondary structures. In addition, two more amino acid sequences consistent with the calculated sequence profile were experimentally confirmed to have a similar amount of secondary structures

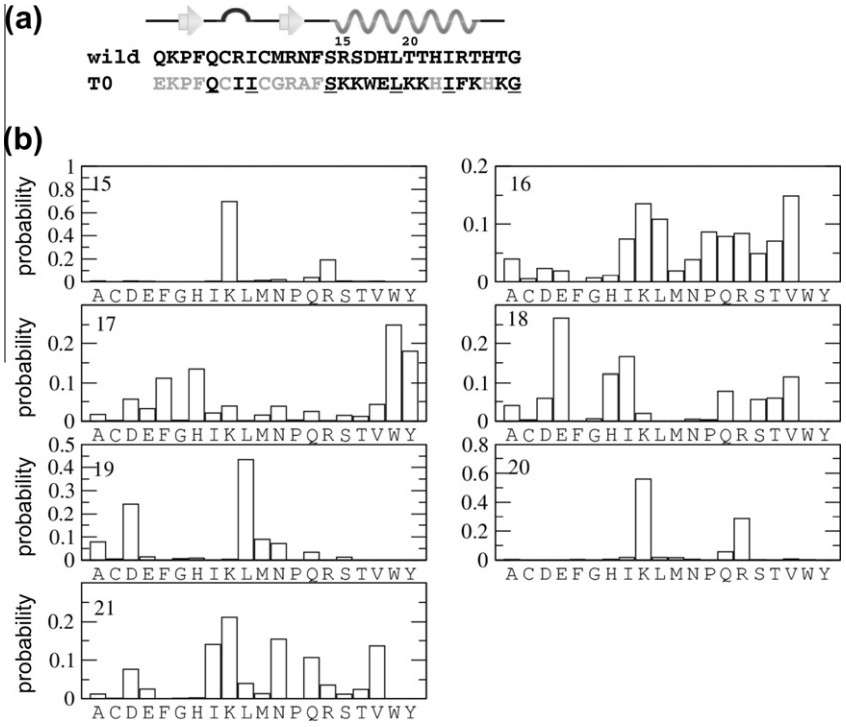

Fig. 2. (a) Amino acid sequence alignment of the wild type and T0. Top is the secondary structure of the wild type. Solid letter denote the sites where the amino acid was freely allowed to mutate (or design). The sites where the identical residue to that of the wild type came out computationally are marked by underline. (b) Amino acid probabilities for the section of DNA recognition sites calculated by the SCAD [7] program. Note that the ordinate is scaled differently for the clarity. The sites 16, 17, 18 and 21 do not show strong preferences for particular amino acids, indicating that these sites can structurally allow any type of amino acids and can be designed or mutated to bind to specific DNA sequences without corrupting the given backbone geometry.

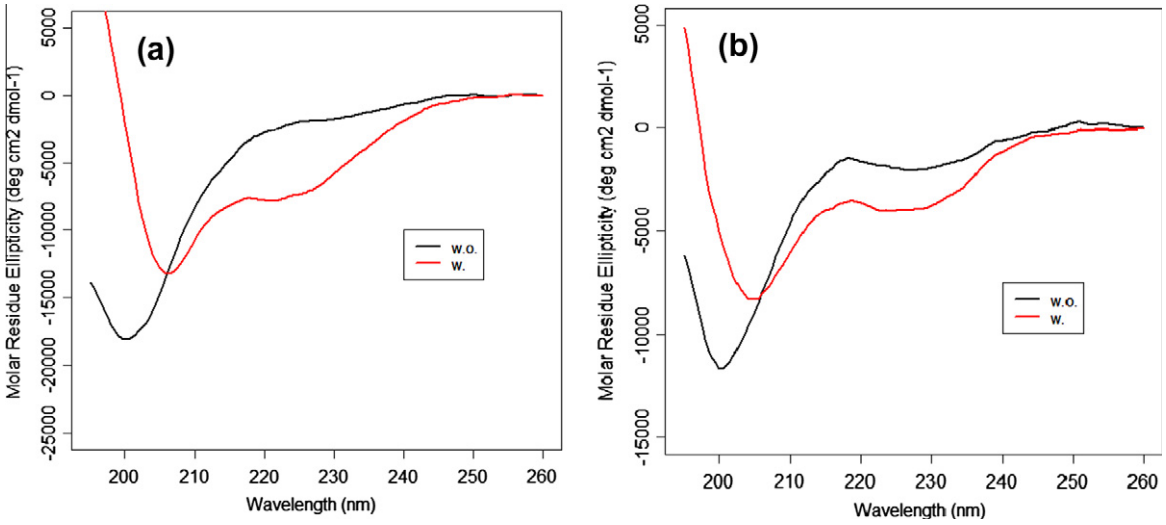

Fig. 3. CD spectra of T0 and T3 fingers (each 25 mM) in the absence (denoted by "w.o.") and presence of $ZnCl_2$ (50 mM) (denoted by "w.") in Tris–HCl (10 mM, pH 7.5) and NaCl (50 mM) at 20 °C.

to that of the wild type (Appendix Fig. A2). This demonstrates that the SCAD program can design amino acids consistent with the given backbone geometry.

Using the two criteria mentioned in the Section 2, we designed amino acid sequences with the ability to specifically bind DNA at residues 15 to 21 of T0 (shown by italic in Fig. 4a), which corresponds to the Zif268 sequence responsible for DNA binding [2,12–16]. Note that TGG (or GGG) is the preferred DNA triplet of the Zif268 Finger 2 [2]. Using the statistical potentials [3], we then computed protein–DNA interaction energies for all possible pairs of amino acid sequences for those seven positions in the second finger ($20^7 = 1.3 \times 10^9$) and all possible triplet DNA sequences ($64 = 4 \times 4 \times 4$), which served as targets [3]. The calculation took about one week with 4 CPUs operating at 2.4 GHz. The result showed that amino acid sequences with sufficiently low protein–DNA interaction energies exist only for TNG (TAG, TCG, TGG and TTG), AGG and CGG triplet DNAs. Among them, sequences only for TGG had a large energy gap between the lowest and the second lowest DNA targets. Then, we decided to synthesize peptides for TNG. Hereafter, we refer to the four peptides that rank the top in

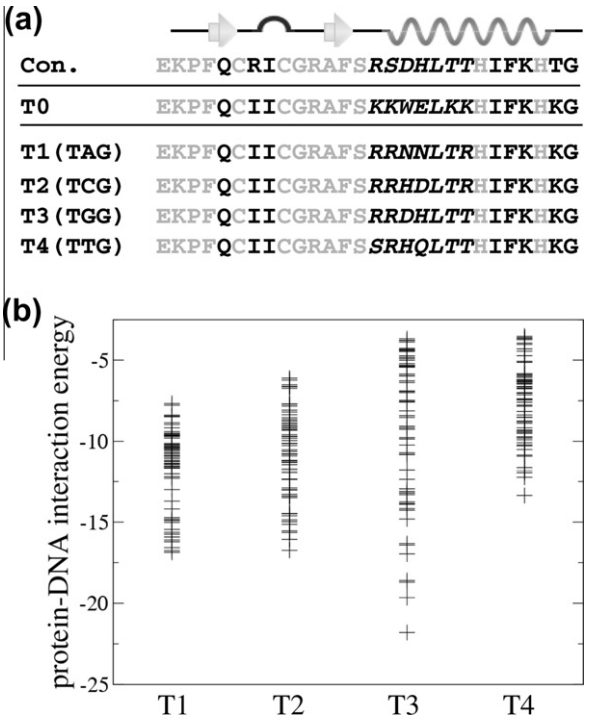

Fig. 4. (a) Five designed amino acid sequences. Out of the 28 positions, 15 positions depicted by solid letters were considered for design. Type of amino acid for positions shown by faint single letters was fixed to the evolutionary conserved one. The top sequence (Con.) is the most conserved amino acid sequence for the zinc finger motif with the secondary structures; the second top is most probable calculated amino acid sequence, given the backbone geometry of the second finger of Zif268. Based on this sequence, four fingers, T1 to T4, were designed. In the left column are the sequence names with the target DNA in parenthesis. Residues in italics were considered for designing DNA binding in the second calculation. (b) Distribution of protein–DNA interaction energies. For each of the designed proteins, interaction energies for all possible triplet DNA sequences ($64 = 4 \times 4 \times 4$) were calculated. The lowest bar denotes the energy for the target DNA. ZF(T3) has an ideal energy gap (see caption in Fig. 2).

terms of interaction energy and are predicted to bind TAG, TCG, TGG and TTG, as T1, T2, T3 and T4, respectively (Fig. 4a). This result seems quite reasonable, as the specificity of the binding of finger 2 to the first (T) base is dependent on the flanking Zif268 finger 3 [19,20]. Among the designed peptides, the amino acid sequence binding TGG (T3) was the only one that satisfied both of the aforementioned criteria (Fig. 4b), promising high specificity for TGG. Note that the wild-type sequence also meets the criteria and ranks the 27th among TGG targeted sequences in terms of interaction energy though the value itself was so close to that of T3, indicating that at least 27 peptides can specifically bind to TGG.

3.2. Experimental validation of the designed zinc finger proteins

To assess the Zn(II) binding and the secondary structures of the designed fingers, we used the Fmoc chemistry to synthesize T1 to T4 and measured their CD spectra in the absence and presence of Zn(II). The CD spectra of them all showed induction of a ββα structure similar to the wild type finger 2 upon addition of Zn(II). The stoichiometry of peptides to Zn(II) was 1:1, suggesting the designed peptides have a proper secondary structure and bind Zn(II) (Fig. 3 and Appendix Fig. A2).

We then used an E-coli system [21] to express the three-finger protein, ZF(T3) in which the amino acid sequences of fingers 1 and 3 were the same as those of Zif268, and finger 2 of Zif268 was replaced with T3, and measured the DNA binding affinities (for the details, see Appendix Plasmids Construction and Three-Finger Protein Expression). We initially used gel-shift assays to determine the binding abilities of all of the designed proteins for their target DNA and related DNAs (TNG). ZF(T3) (TGG binder) was confirmed to specifically bind to TGG with a high binding affinity (K_d = 33 nM; Table 1, and Appendix Fig. A3). Indeed, ZF(T3) and Zif268 had similar tendencies with respect to their binding affinities for TNG triplet DNAs (see the relative K_d in Table 1). ZF(T1), ZF(T2) and ZF(T4) were also created and examined their DNA binding abilities in the same way as ZF(T3). They, however, only weakly bound to their target DNAs and also bound to non-targets (data not shown), which agrees with our computational result indicating their unsatisfactory interaction energies and energy gaps (Fig. 4b). It might be argued that the affinity of ZF(T3) for its target is not due to finger 2, but to fingers 1 and 3; however, given that among the four designed ZFs, which all had the same flanking fingers, only ZF(T3) had a high affinity for its target, we conclude that it is T3 itself that has a high affinity for its target DNA.

We next rigorously investigated the DNA sequence specificity of ZF(T3) and Zif268 for all 64 ($4 \times 4 \times 4$) possible triplet DNAs (Fig. 5 and Appendix Fig. A4), as binding specificity has usually been characterized only for sequences similar to that of the target DNA. Fig. 5a and b show the ratio of the shifted DNA bound to the protein against the added DNA. Both ZF(T3) and Zif268 were found to bind not only to TGG, but also NPG (P = A or G), though they bound most strongly to TGG and TAG. ZF(T3) showed less binding to non-target DNA sequences, such as GGA, CCG, TAA, TGA and TGT, than did Zif268 (No. 09, 39, 49, 57 and 60 in Fig. 4a and b), indicating slightly better specificity. When we examined the order of the calculated protein–DNA interaction energies of ZF(T3) with the 64 triplets, we found it considerably consistent with that of the experimentally determined binding affinities. In particular, the triplet DNAs to which ZF(T3) most strongly bound (shown by red color in Fig. 5b) were all ranked within the top 15 of the 64 targets (Fig. 5c). This agreement confirms that our strategy for amino acid design is effective, and the binding affinity and specificity are predictable for sequences meeting the selection criteria.

Overall, the second fingers of ZF(T3) and Zif268 showed high affinities for KRG (where K is G or T, and R is G or A) triplets (Fig. 5a and b), which is the expected wild-type binding specificity. This indicates that our computational strategy resulted in a new zinc finger sequence with properties similar to the wild-type. When examining the affinities of finger 2 for each of the three sites, T was most favored as the first base of the triplet, followed by G, A and C. The relative affinities for the second and third bases of the triplet were G > A > T > C and G >> C > T > A, respectively. With the exception of CCC, CCT and GGC (for the explanation, see Appendix Fig. A5), these base preferences can explain nearly all of the affinities for the 64 triplets observed in this study and highlight the additivity of the base–amino acid interactions.

Finally, it should be noted that the neighboring zinc fingers may largely contribute to DNA binding specificity of the second finger. In addition, the stability and specificity of protein–DNA interactions not only depend on specific amino acid to nucleic acid contacts, but also to DNA deformability [22–24]. Next step should be the development of a method that considers such effects together in design.

Table 1
Dissociation constants $K_d \pm$ S.D. (nM) and relative values of K_d to the TGG binding (shown in parentheses) of Zif268 and ZF(T3) obtained from the gel-shift assays.

Target	Zif268	ZF3(T3)
TAG	5.1 ± 0.8 (2.0)	52 ± 5.6 (1.6)
TCG	49 ± 4.1 (19)	1350 ± 170 (41)
TGG	2.5 ± 1.0 (1.0)	33 ± 10 (1.0)
TTG	28 ± 5.0 (11)	370 ± 38 (11)

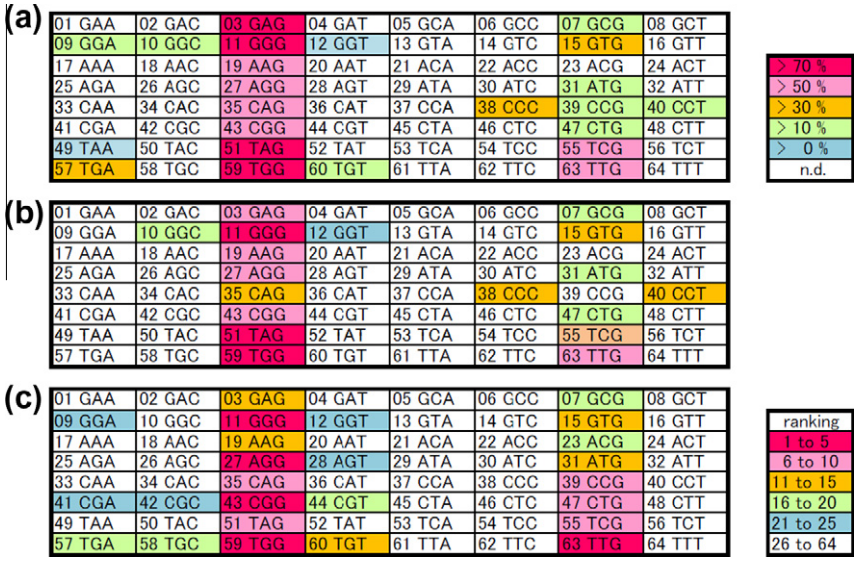

Fig. 5. Binding of Zif268 (a) and ZF(T3) (b) to all 64 possible triplet DNA sequences from 5′-FITC-CCAGACGGATCCTTGAAGCGNNN GCGTTTTCCGAATTCGATCG-3′, where N is A, T, G or C. Zif268 (25 nM) and ZF(T3) (100 nM) were used under the same experimental conditions described in the supporting information. The color code on the right shows the ratio of the shifted DNA bound to the protein to the added DNA; strong to weak binding corresponds to the colors red, pink, orange, light green and cyan. (c) Rankings of the calculated protein–DNA interaction energies of ZF(T3) for all 64 possible triplet DNAs are colored in descending order.

4. Conclusions

In summary, we provide a rational scheme for designing DNA binding proteins. With this scheme, we have computationally designed zinc finger proteins and succeeded in producing a TGG binder, ZF(T3). At the same time, our results suggest that a protein design for binding to DNA targets such as TAG, TCG and TTG is unlikely to be realized as far as we design based on the fixed backbone geometry of Zif268 though some small changes are implicitly allowed by using atom radii smaller than normal ones in the calculation. It is a great challenge in computational protein design to develop a method that can explicitly consider the changes in backbone geometry and we should go forward to such a direction. When designing DNA-binding proteins, one must consider both affinity and specificity. As illustrated in Fig. 4, this is accomplished by designing a protein with an appropriate gap between the lowest interaction energy (for the target) and the second lowest (non-target). The calculation in Fig. 4b shows that ZF(T3) has the lowest interaction energy with TGG and an ideal energy gap, which is in agreement with the experimental results for ZF(T3) showing high affinity and specificity. Artificial sequence-specific DNA binding proteins are useful for regulation of gene expression or recombination. We suggest that this rational design producing proteins with high affinity and specificity for their target DNA has great potential for future application.

Acknowledgements

We thank the advisory members in the PRESTO program for their critical and useful comments and encouragement throughout this work. This work was supported by the JST PRESTO program and JSPS KAKENHI (Grand-in-Aid for Exploratory Research).

Appendix A. Supplementary data

Supplementary data (details of the HSSP profile of the zinc finger fold, peptide synthesis, plasmid construction, protein expression, the data of CD measurement and gel-shift assays) associated with this article can be found, in the online version, at doi:10.1016/j.febslet.2012.02.025.

References

[1] Sarai, A. and Kono, H. (2005) Protein-DNA recognition patterns and predictions. Annu. Rev. Biophys. Biomol. Struct. 34, 379–398.
[2] Choo, Y. and Klug, A. (1997) Physical basis of a protein-DNA recognition code. Curr. Opin. Struct. Biol. 7, 117–125.
[3] Kono, H. and Sarai, A. (1999) Strucure-based prediction of DNA target sites by regulatory proteins. PROTEINS: Struct. Funct. Genet. 35, 114–131.
[4] Jiang, L. et al. (2008) De novo computational design of retro-aldol enzymes. Science 319, 1387–1391.
[5] Oelschlaeger, P. and Mayo, S.L. (2005) Hydroxyl groups in the (beta)beta sandwich of metallo-beta-lactamases favor enzyme activity: a computational protein design study. J. Mol. Biol. 350, 395–401.
[6] Rothlisberger, D. et al. (2008) Kemp elimination catalysts by computational enzyme design. Nature 453, 190–195.
[7] Kono, H. and Saven, J.G. (2001) Statistical theory for protein combinatorial libraries. Packing interactions, backbone flexibility, and the sequence variability of a main chain structure. J. Mol. Biol. 306, 607–627.
[8] Calhoun, J.R., Kono, H., Lahr, S., Wang, W., DeGrado, W.F. and Saven, J.G. (2003) Computational design and characterization of a monomeric helical dinuclear metalloprotein. J. Mol. Biol. 334, 1101–1115.
[9] Slovic, A.M., Kono, H., Lear, J.D., Saven, J.G. and DeGrado, W.F. (2004) Computational design of WSK, a water-soluble analogue of the potassium channel KcsA. Proc. Natl. Acad. Sci. USA 101, 1828–1833.
[10] Nanda, V., Rosenblatt, M.M., Osyczka, A., Kono, H., Getahun, Z., Dutton, P.L., Saven, J.G. and DeGrado, W.F. (2005) De novo design of a redox active minimal rubredoxin mimic. J. Am. Chem. Soc. 127, 5804–5805.
[11] Diaz, J.E. et al. (2011) Computational design and selections for an engineered, thermostable terpene synthase. Protein Sci. 20, 1597–1606.
[12] Pavletich, N.P. and Pabo, C.O. (1991) Zinc finger-DNA recognition: crystal structure of a Zif268–DNA complex at 2.1 A. Science 252, 809–817.
[13] Miller, J.C. and Pabo, C.O. (2001) Rearrangement of side-chains in a Zif268 mutant highlights the complexities of zinc finger-DNA recognition. J. Mol. Biol. 313, 309–315.
[14] Dreier, B., Beerli, R.R., Segal, D.J., Flippin, J.D. and Barbas 3rd, C.F. (2001) Development of zinc finger domains for recognition of the 5′-ANN-3′ family of DNA sequences and their use in the construction of artificial transcription factors. J. Biol. Chem. 276, 29466–29478.
[15] Dreier, B., Segal, D.J. and Barbas 3rd., C.F. (2000) Insights into the molecular recognition of the 5′-GNN-3′ family of DNA sequences by zinc finger domains. J. Mol. Biol. 303, 489–502.
[16] Segal, D.J., Dreier, B., Beerli, R.R. and Barbas 3rd., C.F. (1999) Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5′-GNN-3′ DNA target sequences. Proc. Natl. Acad. Sci. USA 96, 2758–2763.
[17] Dunbrack Jr., R.L. and Karplus, M. (1993) Backbone-dependent rotamer library for proteins. Application to side-chain prediction. J. Mol. Biol. 230, 543–574.
[18] Sander, C. and Schneider, R. (1993) The HSSP data base of protein structure-sequence alignments. Nucleic Acids Res. 21, 3105–3109.
[19] Isalan, M., Choo, Y. and Klug, A. (1997) Synergy between adjacent zinc fingers in sequence-specific DNA recognition. Proc. Natl. Acad. Sci. USA 94, 5617–5621.

[20] Isalan, M., Klug, A. and Choo, Y. (1998) Comprehensive DNA recognition through concerted interactions from adjacent zinc fingers. Biochemistry 37, 12026–12033.
[21] Morisaki, T., Imanishi, M., Futaki, S. and Sugiura, Y. (2008) Rapid transcriptional activity in vivo and slow DNA binding in vitro by an artificial multi-zinc finger protein. Biochemistry 47, 10171–10177.
[22] Olson, W.K., Gorin, A.A., Lu, X.J., Hock, L.M. and Zhurkin, V.B. (1998) DNA sequence-dependent deformability deduced from protein–DNA crystal complexes. Proc. Natl. Acad. Sci. USA 95, 11163–11168.
[23] Gromiha, M.M., Siebers, J.G., Selvaraj, S., Kono, H. and Sarai, A. (2004) Intermolecular and intramolecular readout mechanisms in protein-DNA recognition. J. Mol. Biol. 337, 285–294.
[24] Fujii, S., Kono, H., Takenaka, S., Go, N. and Sarai, A. (2007) Sequence-dependent DNA deformability studied using molecular dynamics simulations. Nucleic Acids Res. 35, 6063–6074.