Experimental support for the evolution of symmetric protein architecture from a simple peptide motif

Jihun Lee and Michael Blaber[1]

Department of Biomedical Sciences, Florida State University, Tallahassee FL 32306-4300

Edited* by Brian W. Matthews, University of Oregon, Eugene, OR, and approved November 10, 2010 (received for review October 6, 2010)

The majority of protein architectures exhibit elements of structural symmetry, and "gene duplication and fusion" is the evolutionary mechanism generally hypothesized to be responsible for their emergence from simple peptide motifs. Despite the central importance of the gene duplication and fusion hypothesis, experimental support for a plausible evolutionary pathway for a specific protein architecture has yet to be effectively demonstrated. To address this question, a unique "top-down symmetric deconstruction" strategy was utilized to successfully identify a simple peptide motif capable of recapitulating, via gene duplication and fusion processes, a symmetric protein architecture (the threefold symmetric β-trefoil fold). The folding properties of intermediary forms in this deconstruction agree precisely with a previously proposed "conserved architecture" model for symmetric protein evolution. Furthermore, a route through foldable sequence-space between the simple peptide motif and extant protein fold is demonstrated. These results provide compelling experimental support for a plausible evolutionary pathway of symmetric protein architecture via gene duplication and fusion processes.

protein design | protein folding | protein symmetry

Symmetry is a central theme in protein structure, function, and evolution. Structural symmetry is observed in many different protein architectures, and gene duplication and fusion is the generally hypothesized mechanism for the emergence of symmetric architecture from simpler (i.e., archaic) peptide motifs (1–4). Such motifs, believed to represent the translational product of "genes of primordial life" (5), typically code for polypeptides of ~40–60 residues in length (5, 6) and may define elemental folding units (i.e., "foldons") (7). Two distinctly different evolutionary models for the emergence of symmetric protein architecture from a primordial peptide motif have been proposed (8–13). In the "emergent architecture" model, the primordial peptide autonomously folds to yield a simple architecture and the complex symmetric architecture emerges upon a final gene duplication and fusion event (Fig. 1*A*). In the conserved architecture model, the symmetric architecture (or integral units thereof) is present at each step of the evolutionary process (Fig. 1*B*).

Experimental studies to probe the evolutionary pathway responsible for symmetric protein architecture have focused principally upon the behavior of subdomain fragments of extant symmetric proteins and have asked whether such fragments fold independently or oligomerize. A fragmentation study of the two-fold symmetric aspartate racemase enzyme identified an independently folding monomeric subdomain, leading the authors to propose an emergent-architecture-type evolutionary model (10). Fragmentation studies of a fivefold symmetric "β-propeller" protein to identify an ancient folding motif showed evidence of multimeric assembly for ~100 amino acid fragments comprising two copies of the repeating motif, but not for ~50 amino acid single-motif fragments (11). Subsequent X-ray structure analysis showed that these duplicated fragments assembled to create two intact β-propeller pentamers, with each pentamer containing 2.5 dimeric domains (with one dimer spanning both pentamers) (14), supporting a conserved-architecture-type model. In the $(\beta\alpha)_8$-barrel architecture, studies of $(\beta\alpha)_4$ half-domain fragments showed that these subdomains assembled to form an intact $(\beta\alpha)_8$-barrel but could also exist as independently folded subdomains (15, 16). Computational and experimental studies designed to identify a corresponding $(\beta\alpha)_2$ "building block" identified a peptide capable of tetrameric oligomerization to create an intact $(\beta\alpha)_8$-barrel but could also exist as an independently folded subdomain (13), thus supporting both types of evolutionary models. However, such simple fragmentation studies of symmetric proteins have suspected limitations. Evolutionary divergence following a duplication/fusion event likely results in optimization of the unique interdo-

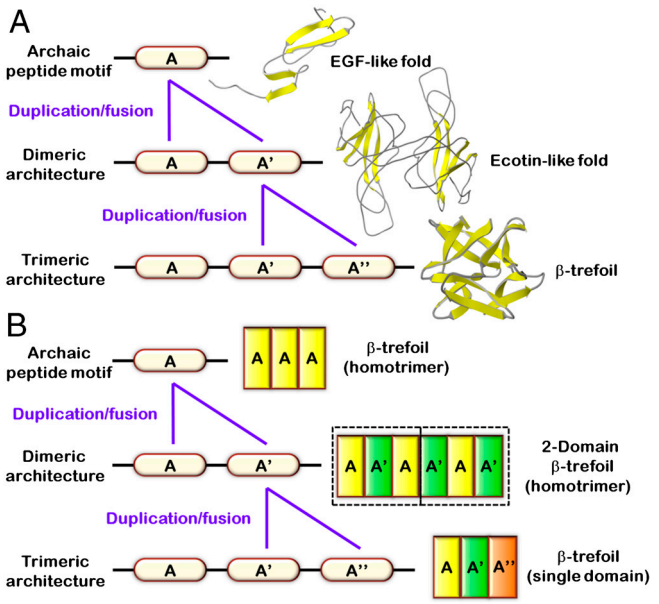

Fig. 1. Evolutionary models of symmetry protein architecture (e.g., the β-trefoil fold). (*A*) The emergent architecture model (8). The archaic peptide motif is autonomously folding, yielding simple single-domain architecture. Intermediate forms produced by gene duplication and fusion events have unique folds with increasing structural symmetry and complexity. The symmetric target architecture emerges upon the final gene duplication and fusion event. (*B*) The conserved architecture model (9). The archaic peptide is not an autonomously folding single-domain architecture; instead, it oligomerizes to yield the symmetric target architecture (e.g., β-trefoil fold). Intermediate forms produced by gene duplication and fusion events similarly oligomerize and reconstitute integral units of the symmetric target architecture. A final duplication and fusion event encodes the symmetric target architecture within a single polypeptide chain.

Author contributions: J.L. and M.B. designed research; J.L. performed research; J.L. and M.B. analyzed data; and J.L. and M.B. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Data deposition: The crystallography, atomic coordinates, and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 3O49, 3O4A,3O4B, 3O4C, 3O4D, 3OL0, 3OGF).

[1]To whom correspondence should be addressed. E-mail: michael.blaber@med.fsu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1015032108/-/DCSupplemental.

main interface, such that independent folding of fragments is unlikely; thus, solubility and folding properties of fragments of extant proteins are unlikely to recapitulate those of the archaic peptide motif (11).

In this report, we describe an experimental top-down symmetric deconstruction of symmetric protein architecture (the β-trefoil fold) with the goal of testing two competing hypotheses for its evolution from an archaic peptide motif. The starting point for this deconstruction is human fibroblast growth factor-1 (FGF-1), a 140 amino acid single-domain globular protein exhibiting the characteristic threefold symmetry of the β-trefoil architecture (where each ∼45 amino acid repeating subdomain is termed a trefoil-fold; ref. 17). As is commonly observed with symmetric proteins, the discernible tertiary structure symmetry of FGF-1 is substantially imperfect at both the primary and tertiary structure levels (Figs. 2*B* and 3*A*). The symmetric deconstruction involved sequential introduction of symmetric mutations (targeting core, reverse-turn, and β-strand secondary structure, respectively) until a purely threefold symmetric primary structure solution was achieved. This symmetric primary structure defines a repeating trefoil-fold polypeptide motif, whose folding and structural properties were studied along with dimer and monomer peptide fragments. The structure and folding properties for these polypeptides agree precisely with a previously hypothesized conserved architecture model of evolution; furthermore, each intermediate mutant in the symmetric deconstruction resides within foldable sequence space, providing thermodynamic support for the hypothesized evolutionary pathway.

Results

Top-Down Symmetric Deconstruction. A total of 18 mutant proteins were constructed and characterized in the process of successfully introducing a complete threefold symmetric primary structure constraint (involving 76 substitutions and 4 deleted positions total; Fig. 2 and Table S1). Detailed biophysical properties for each of these mutants, as well as details of unsuccessful mutations (principally resulting in poor folding cooperativity, solubility, or thermostability) will be reported elsewhere. The initial symmetric protein produced by the deconstruction is termed "Symfoil-1" (for symmetric β-trefoil protein 1). The Symfoil-1 protein has several notable properties in comparison to FGF-1. Symfoil-1 is more thermostable than FGF-1, and whereas FGF-1 irreversibly aggregates upon thermal denaturation, Symfoil-1 exhibits reversible two-state unfolding (Fig. S1). FGF-1 contains examples of all 20 common amino acids, whereas Symfoil-1 is devoid of Trp, Ala, Cys, and Met residues. Additionally, Symfoil-1 has no known FGF-1 functional properties (including FGF receptor-binding, heparin-binding, nuclear localization signal, mitogenic activity, etc.). Symfoil-1 is able to bind the C_3 symmetric molecule Tris on a threefold axis of structural symmetry (Fig. S2). The Symfoil-1 synthetic protein was subsequently optimized by mutation to yield a hyperthermophilic variant (Symfoil-4P mutant, Table S1). Monomeric and dimeric versions of the repeating peptide motif within Symfoil-4P were subsequently created by the introduction of stop codons at the appropriate locations, producing the Monofoil-4P (for monomer trefoil fold) and Difoil-4P (for dimer trefoil fold) polypeptides, respectively (Fig. 2*B*).

Properties of the Monofoil-4P, Difoil-4P, and Symfoil-4P Polypeptides. Both the Monofoil-4P and Difoil-4P polypeptides resolved as single peaks on calibrated size-exclusion chromatography, with apparent masses indicating a homotrimer (in both cases) and no detectable monomer (Fig. S3). Monofoil-4P and Difoil-4P proteins sedimented as homogeneous forms with $s_{20,w}$ values of 2.31 ± 0.036 *s* and 3.18 ± 0.021 *s*, respectively, and with corresponding molecular mass of 18.8 and 29.8 kDa, respectively. The molecular mass of the Monofoil-4P (6.6 kDa) and Difoil-4P (11.2 kDa) polypeptides therefore indicate homotrimer oligomerization for both polypeptides in solution. X-ray crystal structures were solved for Symfoil-1, 2, 4T, 4V, and 4P proteins as well as the Monofoil-4P and Difoil-4P polypeptides (Table S2). The peptide backbone of the Symfoil-1 protein exhibits a striking threefold rotational symmetry (Fig. 3*B*). The main-chain atoms of the three individual trefoil-fold subdomains of the Symfoil-1 protein overlay each other with an rms deviation of 0.2 Å, indicating essentially indistinguishable trefoil-fold subdomain structures. Furthermore, the side-chain rotamers for the set of core-packing residues (and essentially all surface residues with the exception of specific crystal contacts) are identical for a comparison of the three trefoil-fold subdomains (Fig. 3*C*); thus, the threefold symmetry of the Symfoil protein(s) appears exact. The Monofoil-4P structure exhibits a homotrimer assembly that recapitulates the intact β-trefoil architecture (Fig. 3*D*). An overlay of all main-chain atoms comprising the individual repeating trefoil-fold subdomains of the Monofoil-4P trimer with the Symfoil-4P structure yields an rms deviation of 0.5 Å, indicating essential structural identity. Similarly, analysis of side-chain rotamers indicates exact threefold symmetry in the Monofoil-4P structure. The Difoil-4P structure exhibits two intact β-trefoil folds related by a pseudo-twofold axis of symmetry and constructed from three polypeptide chains (Fig. 3*E*). Peptide chain B in this complex (colored green in Fig. 3*E*) adopts an entirely different conformation from the other two peptide chains. Despite this structural alteration, the main-chain atoms of the overall β-trefoil folds contained within the Difoil-4P homotrimer overlay those of the Symfoil-4P structure with an rmsd of ∼0.7 Å, indicating a highly conserved β-trefoil architecture (Fig. 3*F*). The two β-trefoil folds in Difoil-4P are constructed from three polypeptide chains via a "domain swapped" architecture (Fig. 3*G*). The first β-trefoil is comprised of chain "A" plus β-strands 1, 2, 3, and 8 of chain "B," whereas the second β-trefoil is comprised of chain "C" plus β-strands 4, 5, 6, and 7 of chain B. In this second β-trefoil, contiguous β-strands 4 and 5 of chain B take the place of canonical β-strands 12 and 9, respectively (Fig. 3*F*).

Discussion

Support for the Conserved Architecture Model of Symmetric Protein Evolution. Competing emergent architecture and conserved architecture models have been proposed for the evolutionary pathway of the threefold symmetric β-trefoil architecture. In the emergent architecture model of Mukhopadhyay (8), the ancient peptide

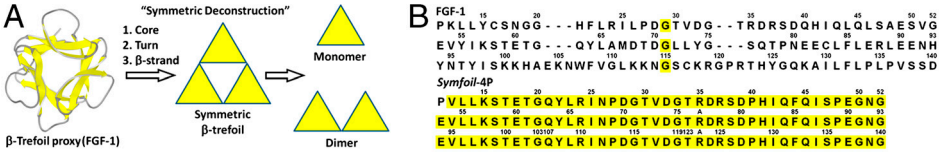

Fig. 2. Summary of the top-down symmetric deconstruction strategy and the result of its application to FGF-1. (*A*) The cumulative symmetric transforms applied to achieve a symmetric deconstruction of the β-trefoil architecture (beginning with FGF-1 and ending in the Monofoil and Difoil polypeptides). Details for the set of intermediary mutants comprising the deconstruction are provided in Table S1. (*B*) Primary structures of the FGF-1 starting protein and the resulting Symfoil-4P symmetric deconstruction. Positions of exact threefold primary structure symmetry are indicated in yellow. The Monofoil-4P and Difoil-4P polypeptides were generated by the introduction of stop codons at positions 53 and 94 of the Symfoil-4P protein, respectively.

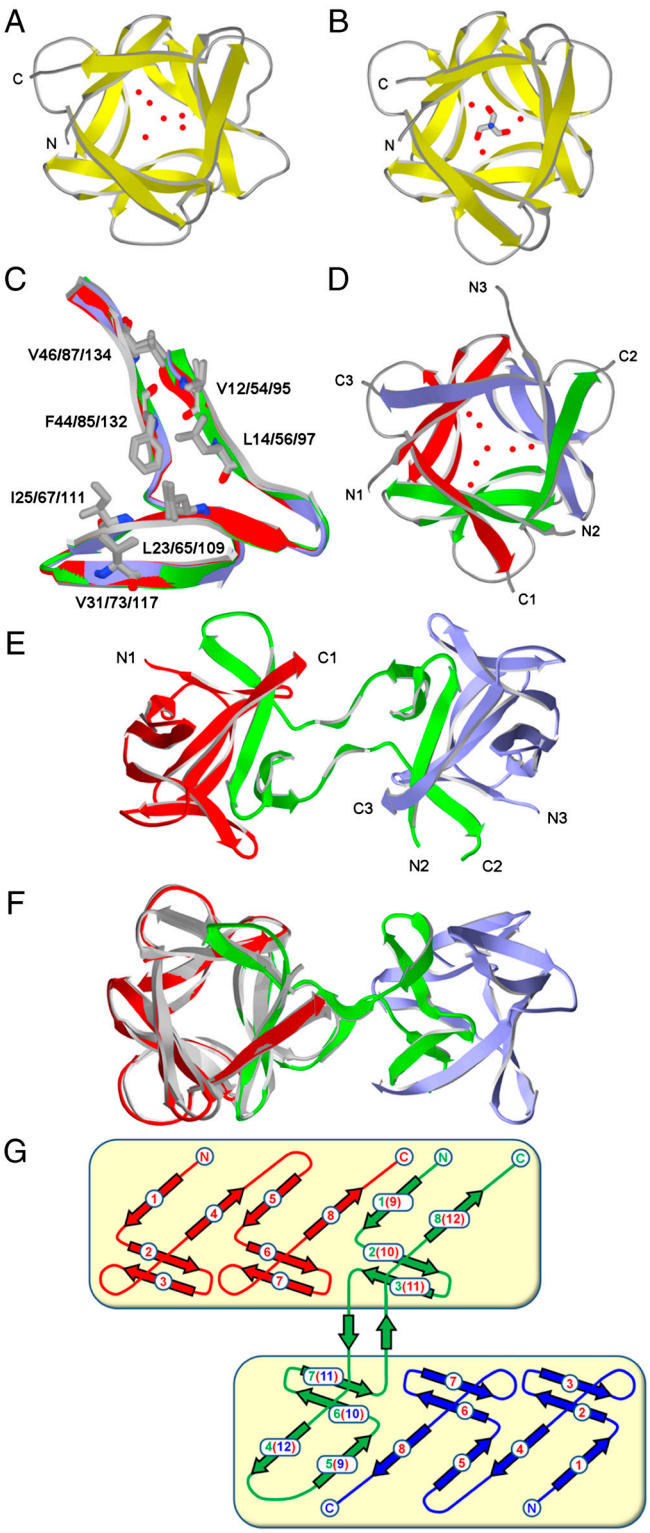

Fig. 3. X-ray structures of FGF-1, Symfoil-1, Monofoil-4P, and Difoil-4P proteins. (*A*) Ribbon representation of FGF-1 oriented down the threefold axis of symmetry and including select solvent structure. (*B*) Similar representation of the Symfoil-1 mutant (also showing the location of a bound Tris molecule). (*C*) Overlay of the repeating trefoil-fold subdomains of Symfoil-1. The main-chain atoms (ribbon representation) are colored red, green, and blue for subdomains 1, 2, and 3, respectively. The set of 21 hydrophobic core residues (Corey, Pauling, Koltun coloring) are in shown in wireframe representation. (*D*) The Monofoil-4P structure with select solvent; the individual Monofoil-4P peptides are colored red, green, and blue and their respective N and C termini are indicated. (*E*) The Difoil-4P structure (individual polypeptides colored as in *D*) and with respective N and C termini indicated. The view is down a twofold axis of symmetry relating the two intact β-trefoil folds present in the homotrimer Difoil-4P structure. (*F*) Fig. 3*E* rotated to view down the threefold axis of symmetry within the first β-trefoil domain and overlaid with the Symfoil-4P structure (gray). (*G*) Secondary structure schematic diagram of Difoil-4P (colored as in *E*). The boxed regions indicate the two β-trefoil domains.

motif was hypothesized to have been an autonomously folding single-domain, small globular protein, with a simple epidermal growth-factor-like fold. Dimerization driven by mutation(s) causing "domain-swapping" was postulated to produce a C_2 symmetric molecule with an overall Ecotin-like fold. The β-trefoil architecture was hypothesized to have emerged after a final duplication and fusion event. A distinctly different conserved architecture model for the evolution of β-trefoil architecture was proposed at essentially the same time by Ponting and Russell (9). In this model, the ancient peptide motif was hypothesized to have existed as a homotrimer with overall β-trefoil architecture. Duplication and fusion resulted in a dimer repeat of this polypeptide which also folded as a homotrimer, producing two complete β-trefoil folds (although structural details of exactly how this would be accomplished were not proposed). A subsequent duplication and fusion event produced a triplicate repeat of the trefoil-fold peptide, yielding a single polypeptide with β-trefoil architecture.

Symmetric deconstruction of the β-trefoil architecture was undertaken with no preconceived ideas regarding the sequence or properties of the ultimate polypeptide; however, the results show that the oligomerization and structural details of the Monofoil, Difoil, and Symfoil polypeptides agree precisely with the conserved architecture model of β-trefoil evolution (Figs. 1*B* and 3). The symmetric deconstruction of the FGF-1 protein was accomplished in a punctuated stepwise manner with each of the 18 intermediate mutant forms being foldable, thermostable, and soluble (Fig. 4 and Table S3); thus, the results also present a thermodynamically tenable pathway, through foldable sequence space, extending from FGF-1 to the Monofoil-4P peptide. Sequence analysis suggests that all extant β-trefoil proteins evolved from a common ancestor (9), because no homotrimer examples of the β-trefoil architecture (either Monofoil-type or Difoil-type) have been reported, the experimentally supported evolutionarily model describes a likely primordial event. Whereas the emergent architecture model implies "simple gene–simple protein architecture," the supported conserved architecture model enables complex protein architecture (e.g., a β-trefoil fold) via a simple gene (i.e., coding for a Monofoil-like 42-mer peptide). Thus, the results also support the hypothesis that complex protein architecture is possible in ancient life forms containing a substantially reduced genome complexity.

Implications for de Novo Protein Design. A long-standing goal in de novo protein design has been the exploitation of a "bottom-up" hierarchical design strategy utilizing peptide building blocks that can assemble (via oligomerization or concatenation) into target architecture (18–20) (in effect, paralleling aspects of the evolutionary pathway). The top-down symmetric deconstruction approach was remarkably successful in identifying a peptide building block solution (i.e., Monofoil-4P peptide) for the β-trefoil architecture. Engineered thermostable protein architecture can serve as a "function-competent" (21) scaffold for the incorporation of unique functionality (22, 23). Although the FGF-1 protein has low mesophilic stability (24, 25), inclusion of a stability screen in the deconstruction resulted in a hyperthermophilic β-trefoil (the Symfoil-4P protein), and may therefore be utilized as a departure point in bottom-up de novo protein design for the incorporation of unique functionality. The fundamental "symmetric deconstruction" concept is architecture independent and therefore applicable in identifying a peptide building block for other symmetric protein folds.

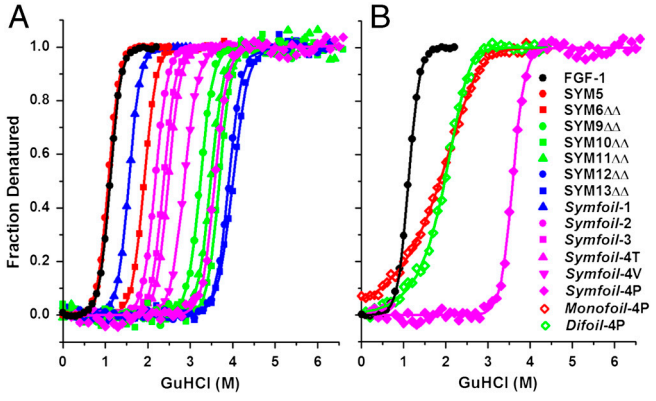

Fig. 4. Isothermal equilibrium denaturation profiles for FGF-1 and mutants comprising the top-down symmetric deconstruction. (*A*) Denaturation profiles for FGF-1 and all single-polypeptide mutant proteins (culminating in the Symfoil-4P mutant). The starting FGF-1 protein is the least stable in comparison to all mutant proteins. (*B*) Denaturation profiles comparing the homotrimer assemblies of Monofoil-4P and Difoil-4P peptides (10 μM each) with FGF-1 and Symfoil-4P proteins.

Materials and Methods

Top-Down Symmetric Deconstruction. Top-down symmetric deconstruction is the introduction (via mutagenesis) of a symmetric constraint upon the primary and tertiary structure of a naturally evolved symmetric protein fold with the goal of achieving a purely symmetric primary structure. Details of the method will be published elsewhere, but briefly, the deconstruction is accomplished via a series of cumulative symmetric mutagenic transforms. In the case of FGF-1, this deconstruction involved a cumulative symmetric constraint upon core, reverse-turn, and β-strand secondary structure elements, respectively (Fig. 2*A*). Symmetric deconstruction of FGF-1 requires substantial change to both the primary and tertiary structure of the protein (Fig. 2*B*). Mutations were combined with a screen for protein folding, stability, and solubility, and were retained if they were neutral or improved upon such properties. Maximum use was made of previously published data (26–34) as well as sequence analysis to identify useful symmetric mutations in FGF-1. A summary of the application of specific transforms in pursuing a symmetric deconstruction of FGF-1 follows.

Transform 1: Symmetric deconstruction of the hydrophobic core. Previously published work detailed the development of a series of symmetric core-packing mutations (SYM2-5, refs. 26 and 27; see Table S1 for mutant definitions) in FGF-1 that systematically introduced a symmetric primary structure constraint. Achieving a purely symmetric core while avoiding substantial destabilization required deletion of six amino acids within the third trefoil-fold subdomain, producing the SYM6ΔΔ mutant (28). Subsequent development of the SYM7ΔΔ mutant (29), involving symmetric mutations in additional buried hydrophobic positions of the protein represented essential completion of transform 1 and was the starting point for the application of transform 2.

Transform 2: Symmetric deconstruction of reverse turns. The SYM7ΔΔ mutant was initially modified by the inclusion of two previously described stabilizing point mutations (Lys12Val and Pro134Val) (30) to produce the SYM9ΔΔ mutant. These mutations are located in the adjacent first and last β-strands of the β-barrel of the architecture, respectively, and represent a "discontinuous" β-turn within the threefold architecture. A His93Gly mutation, located within β-turn 8 and known to stabilize FGF-1 (31, 35), was introduced into the SYM9ΔΔ mutant to produce the SYM10ΔΔ mutant. Prior studies of the role of the statistically preferred β-turn motif Asx-Pro-Asx-Gly at individual β-turns 2, 6, and 10 in FGF-1 indicated a consistently favorable effect upon thermostability with Asn-Xxx-Asp-Gly (where Xxx represents the retained FGF-1 residue at this position) (34). The canonical Asn-Pro-Asp-Gly sequence was therefore substituted into the SYM10ΔΔ mutant at each of these symmetry-related β-turns to produce the SYM11ΔΔ mutant. The symmetric deconstruction of β-turn regions was not taken to completion at this stage, and the SYM11ΔΔ mutant was considered a successful "proof-of-concept" and was utilized as the starting point for transform 3.

Transform 3: Symmetric deconstruction of β-strands. An Asn95Val mutation was introduced into SYM11ΔΔ resulting in a symmetric (Val) deconstruction of positions 12, 54, and 95 within β-strands 1, 5, and 9, respectively. Similarly, Leu46Val and Glu87Val mutations were introduced, resulting in a further symmetric constraint within symmetry-related β-strands 4, 8, and 12, respectively. The combination of these mutations resulted in the SYM12ΔΔ mutant. Ile56Leu and Tyr97Leu mutations were introduced into the SYM12ΔΔ mutant, resulting in a symmetric constraint upon symmetry-related β-strand

positions 14, 56, and 97 and producing the SYM13ΔΔ mutant. The symmetric deconstruction of β-strand secondary structure was not taken to completion at this stage, and the SYM13ΔΔ mutant was considered a successful proof-of-concept. Completion of transforms 2 and 3 was achieved simultaneously by a chimera construct strategy chosen to eliminate reactive free thiols at positions Cys16 and Cys83. The resulting chimera sequence, yielding a complete symmetric deconstruction, is termed Symfoil-1.

Transform 4: Fragmentation of the repeating motif. Optimization of Symfoil-1 stability was pursued prior to fragmentation of the repeating motif. Details of the optimization design will be published elsewhere, but involved development of the Symfoil-2, Symfoil-3, Symfoil-4T, 4V, and 4P mutants (Table S1). The Symfoil-4P mutant (the most stable Symfoil variant) was selected as the starting point for fragmentation studies of the repeating peptide motif. Monomer and dimer subdomains were constructed by mutagenesis to introduce stop codons at positions Glu53 or Glu94, respectively, of the Symfoil-4P construct (Fig. 2). The resulting monomeric and dimeric trefoil-fold peptides are referred to as Monofoil-4P and Difoil-4P, respectively (for monomeric trefoil fold and dimeric trefoil fold, respectively).

Mutagenesis, Protein Purification, and Characterization. Mutagenesis and protein purification and details of structural and biophysical characterization of mutant proteins followed previously published methods and are provided in *SI Materials and Methods*.

ACKNOWLEDGMENTS. This work was funded by Grant 0655133B from the American Heart Association and by additional research support provided by the Florida State University College of Medicine. Analytical ultracentrifugation calculations were performed on the UltraScan Laboratory Information Management System cluster at the Bioinformatics Core Facility at the University of Texas Health Science Center at San Antonio and the Lonestar cluster at the Texas Advanced Computing Center, supported by National Science Foundation Teragrid Grant MCB070038 (to B. Demeler). Use of the "mail-in crystallography" facility of Southeast Regional Collaborative Access Team for diffraction data collection is acknowledged. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract W-31-109-Eng-38. Use of the National Synchrotron Light Source, Brookhaven National Laboratory, is acknowledged, and financial support comes principally from the Office of Biological and Environmental Research and the Office of Basic Energy Sciences of the US Department of Energy, and the National Center for Research Resources of the National Institutes of Health.

1. Sepulveda P, Marciniszyn JJ, Liu D, Tang J (1975) Primary structure of porcine pepsin. III. Amino acid sequence of a cyanogen bromide fragment, CB2A, and the complete structure of porcine pepsin. *J Biol Chem* 250:5082–5088.
2. Tang J, James MN, Hsu IN, Jenkins JA, Blundell TL (1978) Structural evidence for gene duplication in the evolution of the acid proteases. *Nature* 271:618–621.
3. McLachlan AD (1979) Three-fold structural pattern in the soybean trypsin inhibitor (Kunitz). *J Mol Biol* 133:557–563.
4. Inana G, Piatigorsky J, Norman B, Slingsby C, Blundell T (1983) Gene and protein structure of a β-crystallin polypeptide in murine lens: Relationship of exons and structural motifs. *Nature* 302:310–315.
5. Tateno Y, et al. (1997) Evolutionary motif and its biological and structural significance. *J Mol Evol* 44:S38–S43.
6. Dorit RL, Schoenbach L, Gilbert W (1990) How big is the universe of exons? *Science* 250:1377–1382.
7. Panchenko AR, Luthey-Schulten Z, Cole R, Wolynes PG (1997) The foldon universe: A survey of structural similarity and self-recognition of independently folding units. *J Biol Chem* 272:95–105.
8. Mukhopadhyay D (2000) The molecular evolutionary history of a winged bean α-chymotrypsin inhibitor and modeling of its mutations through structural analysis. *J Mol Evol* 50:214–223.
9. Ponting CP, Russell RB (2000) Identification of distant homologues of fibroblast growth factors suggests a common ancestor for all beta-trefoil proteins. *J Mol Biol* 302:1041–1047.
10. Liu L, Iwata K, Yohda M, Miki K (2002) Structural insight into gene duplication, gene fusion and domain swapping in the evolution of PLP-independent amino acid racemases. *FEBS Lett* 528:114–118.
11. Yadid I, Tawfik DS (2007) Reconstruction of functional b-propeller lectins via homo-oligomeric assembly of shorter fragments. *J Mol Biol* 365:10–17.
12. Akanuma S, Matsuba T, Ueno E, Umeda N, Yamagishi A (2010) Mimicking the evolution of a thermally stable monomeric four-helix bundle by fusion of four identical single-helix peptides. *J Biochem* 147:371–379.
13. Richter M, et al. (2010) Computational and experimental evidence for the evolution of a $(\beta\alpha)_8$-barrel protein from an ancestral quarter-barrel stabilized by disulfide bonds. *J Mol Biol* 398:763–773.
14. Yadid I, Kirshenbaum N, Sharon M, Dym O, Tawfik DS (2010) Metamorphic proteins mediate evolutionary transitions of structure. *Proc Natl Acad Sci USA* 107:7287–7292.
15. Hocker B, Beismann-Driemeyer S, Hettwer S, Lustig A, Sterner R (2001) Dissection of a $(\beta\alpha)_8$-barrel enzyme into two folded halves. *Nat Struct Biol* 8:32–36.
16. Akanuma S, Yamagishi A (2008) Experimental evidence for the existence of a stable half-barrel subdomain in the $(\beta/\alpha)_8$-barrel fold. *J Mol Biol* 382:458–466.
17. Murzin AG, Lesk AM, Chothia C (1992) β-Trefoil fold. Patterns of structure and sequence in the kunitz inhibitors interleukins-1β and 1α and fibroblast growth factors. *J Mol Biol* 223:531–543.
18. DeGrado WF, Regan L, Ho SP (1987) The design of a four-helix bundle protein. *Cold Spring Harbor Symp Quant Biol* 52:521–526.
19. Regan L, DeGrado WF (1988) Characterization of a helical protein designed from first principles. *Science* 241:976–978.
20. Richardson J, Richardson DC (1989) The de novo design of protein structures. *Trends Biochem Sci* 14:304–309.
21. Soskine M, Tawfik DS (2010) Mutational effects and the evolution of new protein functions. *Nat Rev Genet* 11:572–582.
22. Gibney BR, Rabanal F, Dutton PL (1997) Synthesis of novel proteins. *Curr Opin Chem Biol* 1:537–542.
23. Schafmeister CE, Stroud RM (1998) Helical protein design. *Curr Opin Biotechnol* 9:350–353.
24. Blaber SI, Culajay JF, Khurana A, Blaber M (1999) Reversible thermal denaturation of human FGF-1 induced by low concentrations of guanidine hydrochloride. *Biophys J* 77:470–477.
25. Copeland RA, et al. (1991) The structure of human acidic fibroblast growth factor and its interaction with heparin. *Arch Biochem Biophys* 289:53–61.
26. Brych SR, Blaber SI, Logan TM, Blaber M (2001) Structure and stability effects of mutations designed to increase the primary sequence symmetry within the core region of a β-trefoil. *Protein Sci* 10:2587–2599.
27. Brych SR, Kim J, Logan TM, Blaber M (2003) Accommodation of a highly symmetric core within a symmetric protein superfold. *Protein Sci* 12:2704–2718.
28. Brych SR, et al. (2004) Symmetric primary and tertiary structure mutations within a symmetric superfold: A solution, not a constraint, to achieve a foldable polypeptide. *J Mol Biol* 344:769–780.
29. Dubey VK, Lee J, Blaber M (2005) Redesigning symmetry-related "mini-core" regions of FGF-1 to increase primary structure symmetry: Thermodynamic and functional consequences of structural symmetry. *Protein Sci* 14:2315–2323.
30. Dubey VK, Lee J, Somasundaram T, Blaber S, Blaber M (2007) Spackling the crack: Stabilizing human fibroblast growth factor-1 by targeting the N and C terminus beta-strand interactions. *J Mol Biol* 371:256–268.
31. Kim J, Brych SR, Lee J, Logan TM, Blaber M (2003) Identification of a key structural element for protein folding within β-hairpin turns. *J Mol Biol* 328:951–961.
32. Kim J, Lee J, Brych SR, Logan TM, Blaber M (2005) Sequence swapping does not result in conformation swapping for the b4/b5 and b8/b9 b-hairpin turns in human acidic fibroblast growth factor. *Protein Sci* 14:351–359.
33. Lee J, Blaber M (2009) Structural basis for conserved cysteine in the fibroblast growth factor family: Evidence for a vestigial half-cystine. *J Mol Biol* 393:128–139.
34. Lee J, Dubey VK, Longo LM, Blaber M (2008) A logical OR redundancy with the Asx-Pro-Asx-Gly type I β-turn motif. *J Mol Biol* 377:1251–1264.
35. Arakawa T, et al. (1993) Production and characterization of an analog of acidic fibroblast growth factor with enhanced stability and biological activity. *Protein Eng* 6:541–546.