Supporting Information

Lee and Blaber 10.1073/pnas.1015032108

SI Materials and Methods

Mutagenesis and Protein Purification. Construction of all the "SYMΔΔ" mutants utilized a synthetic gene for the 140 amino acid form of human fibroblast growth factor-1 (FGF-1) (1–4) containing an additional amino-terminal six His tag and following previously described procedures (5). Construction of the Symfoil-1 (for symmetric β-trefoil protein 1) mutant involved complete gene synthesis utilizing unique codons at symmetry-related positions. Expression and purification of recombinant proteins followed previously published procedures (5). Purified protein was exchanged into 50 mM sodium phosphate, 0.1 M NaCl, 10 mM $(NH_4)_2SO_4$, 2 mM DTT, pH 7.5 ("crystallization buffer") for crystallization studies or 20 mM N-(2-acetamido)iminodiacetic acid (ADA), 0.1 M NaCl, 2 mM DTT, pH6.6 ("ADA buffer") for biophysical studies (DTT was omitted in all Symfoil mutants). An extinction coefficient of E_{280} nm (0.1%,1 cm) = 1.26 (6, 7) was used for FGF-1 and the extinction coefficient for all mutant forms was determined by the method of Gill and von Hippel (8).

Isothermal Equilibrium Denaturation. Isothermal equilibrium denaturation by guanidine HCl (GuHCl) was performed using either fluorescence or CD as previously described (9, 10). The effect of mutation upon protein stability (ΔΔG) was calculated by taking the difference between the midpoint of denaturation (C_m value) for reference and mutant proteins and multiplying by the average m value as described by Pace and Scholtz (11), where a negative value indicates the mutation is stabilizing. In the case of Monofoil-4P and Difoil-4P proteins, the data were analyzed using a trimer-to-monomer isothermal equilibrium denaturation model (12, 13).

Differential scanning calorimetry (DSC). All DSC data were collected on a VP-DSC microcalorimeter (GE Healthcare) as previously described (9). Molar heat capacity data were analyzed using a two-state model as implemented in the DSCfit software package (14). The Monofoil-4P and Difoil-4P mutants were analyzed using a trimer-to-monomer thermal denaturation model (12, 13) implemented using the DataFit nonlinear least-squares-fit software package (Oakdale Engineering).

X-Ray Crystallization and Structure Determination. Purified mutant protein in crystallization buffer was concentrated to 9–15 mg/mL and crystals were grown using either the hanging-drop or sitting-drop vapor diffusion method at room temperature. Crystals of Symfoil-1, Symfoil-2, Symfoil-4T, Symfoil-4V, and Symfoil-4P proteins grew in 1–3 wk from vapor diffusion against 1.8–2.3 M ammonium sulfate, 0.1–0.2 M lithium sulfate, and 0.1 M Tris pH 7.0. To minimize potential flexibility of the N-terminal region of Monofoil-4P and Difoil-4P polypeptides, residues from Phe1 to Lys10 were deleted (producing the Monofoil-4PΔ1-10 and Difoil-4PΔ1-10 mutants, respectively). Crystals of Monofoil-4PΔ1-10 and Difoil-4PΔ1-10 grew in 2 M ammonium sulfate, 0.1 M Na citrate pH 5.5. Crystals were mounted in a stream of gaseous nitrogen at 100 K and diffraction data were collected at either the Southeast Regional Collaborative Access Team 22-BM beam line ($\lambda = 1.00$ Å) at the Advanced Photon Source, Argonne National Laboratory, using a MarCCD 300 detector (Mar USA) or at the X25 beam line of the National Synchrotron Light Source at Brookhaven National Laboratory, using an ADSC Q315 CCD detector. Each dataset was collected from a single crystal except Monofoil-4PΔ1-10 and Difoil-4PΔ1-10. Datasets from two crystals of Monofoil-4PΔ1-10 were combined for better completion and redundancy. Difoil-4PΔ1-10 crystals exhibited radiation sensitivity and diffraction data from three Difoil-4PΔ1-10 crystals were combined to yield acceptable completeness.

Diffraction data were indexed, integrated, and scaled using the HKL2000 software package (15, 16). Molecular replacement and model building utilized the PHENIX software package (17), with 5% of the data in the reflection files set aside for R_{free} calculations (18). Model building and visualization utilized the COOT molecular graphics software package (19). His-tagged FGF-1 (Protein Data Bank ID code 1JQZ) was used as the search model in molecular replacement for all Symfoil mutant proteins. The resulting Symfoil-4P X-ray structure was used as a search model in molecular replacement with the Monofoil-4PΔ1-10 X-ray data. The correctly positioned Symfoil-4P structure was subsequently divided into three chains (A through C) representing the three individual Monofoil-4PΔ1-10 polypeptides. The Symfoil-4P X-ray structure was also used as the search model in molecular replacement with the Difoil-4PΔ1-10 X-ray data, yielding an acceptable solution with two independent copies of the β-trefoil search model in the asymmetric unit. Refinement of the Difoil-4PΔ1-10 structure initially utilized a twofold noncrystallographic symmetry (NCS) constraint (relating the two independent β-trefoil solutions). Chain definitions for the three independent polypeptide chains comprising the two intact β-trefoil folds were assigned based upon contiguous density in the $2F_o$-F_c omit map. Refinement of the Difoil-4PΔ1-10 structure subsequently utilized a threefold NCS constraint for residue segments 11–38 of chains A, B, and C, and residue segments 52–79 of chains A, B, and C. Final refinement of the Difoil-4PΔ1-10 model did not utilize any NCS constraint. All X-ray structures, with the exception of Difoil-4P, yielded refined coordinates with >90% of residues in the most favored region of the Ramachandran plot and no residues in disallowed regions. The Difoil-4P structure yielded refined coordinates with 79% of residues in the most favored region and no residues in disallowed regions.

Calibrated Size-Exclusion Chromatography. Calibrated size-exclusion chromatography was performed on a Hi-Load Superdex 75 26/60 column (318 mL column volume; GE Healthcare) on an AKTA FPLC system (GE Healthcare) in crystallization buffer at a flow rate of 2.5 mL/min. The column was calibrated using mass standards of bovine serum albumin (66.0 kDa), carbonic anhydrase (33.0 kDa), and cytochrome C (12.5 kDa) and a standard curve was fit to elution volume versus log(mw). Two-milliliter samples of 50 μM FGF-1 and mutant proteins were resolved and the apparent molecular masses were determined by comparison to the mass standard curve.

Analytical Ultracentrifugation. Sedimentation velocity analytical ultracentrifugation experiments were performed in a Beckman XL-I centrifuge (Beckman Coulter, Inc.) using absorbance optics and measuring intensity scans at 280 nm. The experiments were performed at 20 °C in two-channel Epon centerpieces with an AN60 Ti rotor at 58,000 rpm and using 142 μM of Monofoil-4P, and 65 μM of Difoil-4P, equilibrated in ADA buffer. Data were analyzed using the UltraScan II version 9.9 software suite (20, 21). All computations were performed on the TIGRE cluster at the University of Texas Health Science Center at San Antonio and the Texas Advanced Computing Center at the University of Texas in Austin. All data were first analyzed by two-dimensional spectrum analysis (22) with simultaneous removal of time-invariant noise, and then by genetic algorithm refinement (23),

followed by Monte Carlo analysis (24). The partial specific volumes at 20 °C of the Monofoil-4P (0.716 cm^3/g) and Difoil-4P (0.715 cm^3/g) proteins were estimated from peptide sequence as described by Durchschlag (25).

1. Gimenez-Gallego G, Conn G, Hatcher VB, Thomas KA (1986) The complete amino acid sequence of human brain-derived acidic fibroblast growth factor. *Biochem Biophys Res Commun* 128:611–617.
2. Linemeyer DL, et al. (1990) Disulfide bonds are neither required, present, nor compatible with full activity of human recombinant acidic fibroblast growth factor. *Growth Factors* 3:287–298.
3. Ortega S, et al. (1991) Conversion of cysteine to serine residues alters the activity, stability, and heparin dependence of acidic fibroblast growth factor. *J Biol Chem.* 266:5842–5846.
4. Blaber M, DiSalvo J, Thomas KA (1996) X-ray crystal structure of human acidic fibroblast growth factor. *Biochemistry* 35:2086–2094.
5. 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.
6. Zazo M, et al. (1992) High-level synthesis in *Escherichia coli* of a shortened and full-length human acidic fibroblast growth factor and purification in a form stable in aqueous solutions. *Gene* 113:231–238.
7. Tsai PK, et al. (1993) Formulation design of acidic fibroblast growth factor. *Pharm Res* 10:649–659.
8. Gill SC, von Hippel PH (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem* 182:319–326.
9. 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.
10. 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.
11. Pace CN, Scholtz JM (1997) Measuring the conformational stability of a protein. *Protein Structure: A Practical Approach*, ed Creighton TE (Oxford Univ Press, Oxford), pp 299–321.
12. Backmann J, Schafer G, Wyns L, Bonisch H (1998) Thermodynamics and kinetics of unfolding of the thermostable trimeric adenylate kinase from the archaeon *Sulfolobus acidocaldarius*. *J Mol Biol* 284:817–833.
13. Jelesarov I, Lu M (2001) Thermodynamics of trimer-of-hairpins formation by the SIV gp41 envelope protein. *J Mol Biol* 307:637–656.
14. Grek SB, Davis JK, Blaber M (2001) An efficient, flexible-model program for the analyis of differential scanning calorimetry protein denaturation data. *Protein Pept Lett* 8:429–436.
15. Otwinowski Z (1993) Oscillation data reduction program. *Proceedings of the CCP4 Study Weekend: "Data Collection and Processing"*, eds Sawyer L, Isaacs N, Bailey S (Science and Engineering Research Council, Daresbury Laboratory, England), pp 56–62.
16. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Meth Enzymol* 276:307–326.
17. Zwart PH, et al. (2008) Automated structure solution with the PHENIX suite. *Methods Mol Biol* 426:419–435.
18. Brunger AT (1992) Free *R* value: A novel statistical quantity for assessing the accuracy of crystal structures. *Nature* 355:472–475.
19. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr, Sect D: Biol Crystallogr* 60:2126–2132.
20. Demeler B (2005) UltraScan—a comprehensive data analysis software package for analytical ultracentrifugation experiments. *Modern Analytical Ultracentrifugation: Techniques and Methods*, eds Scott DJ, Harding SE, Rowe AJ (Royal Society of Chemistry, Cambridge, UK), pp 210–229.
21. Schuck P, Demeler B (1999) Direct sedimentation analysis of interference optical data in analytical ultracentrifugation. *Biophys J* 76:2288–2296.
22. Brookes E, Cao W, Demeler B (2010) A two-dimensional spectrum analysis for sedimentation velocity experiments of mixtures with heterogeneity in molecular weight and shape. *Eur Biophys J* 39:405–414.
23. Brookes EH, Demeler B (2007) Parsimonious regularization using genetic algorithms applied to the analysis of analytical ultracentrifugation experiments. *Proceedings of the Ninth Annual Conference on Genetic and Evolutionary Computation* (Association for Computing Machinery, London), pp 361–368.
24. Demeler B, Brookes E (2008) Monte Carlo analysis of sedimentation experiments. *Colloid Polym Sci* 286:129–137.
25. Durchschlag H (1986) *Thermodynamic Data for Biochemistry and Biotechnology*, ed Hinz H-J (Springer, Berlin), pp 45–128.

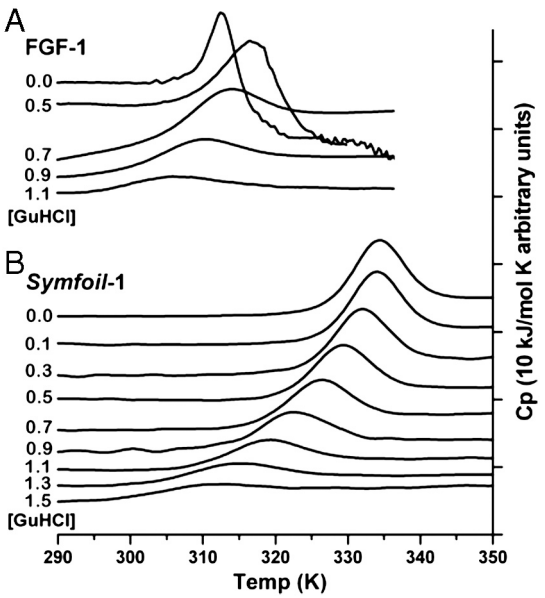

Fig. S1. Differential scanning calorimetry endotherms of FGF-1 and Symfoil-1 proteins in the presence of varying concentrations of GuHCl. (*A*) DSC endotherms of FGF-1 in 20 mM N-(2-acetamido)iminodiacetic acid, 0.1 M NaCl, pH6.6, and with the indicated concentrations of guanidine hydrochloride. FGF-1 undergoes irreversible thermal denaturation in the absence of ~0.6 M GuHCl. (*B*) DSC endotherms for the Symfoil-1 mutant in the same buffer conditions as in *A*. Unlike FGF-1, the Symfoil-1 mutant exhibits reversible, two-state thermal denaturation under all buffer conditions.

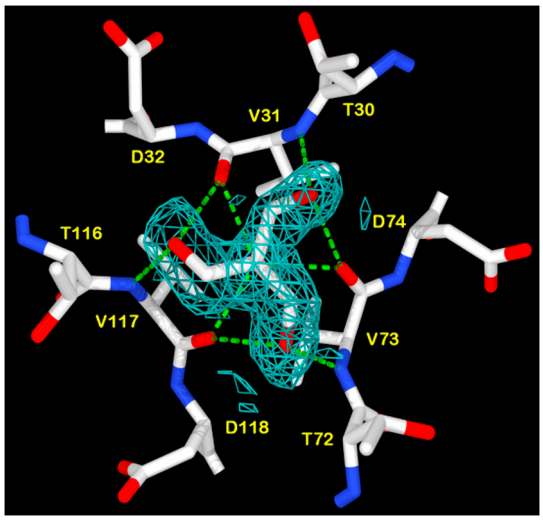

Fig. S2. Symfoil-1 mutant X-ray data omit-map contoured at 1.0σ and showing nonprotein density in the region of the threefold axis of structural symmetry at the base of the central β-barrel. A Tris molecule modeled into this density is shown, along with the H-bond interactions between the Tris polar groups and Symfoil-1 protein.

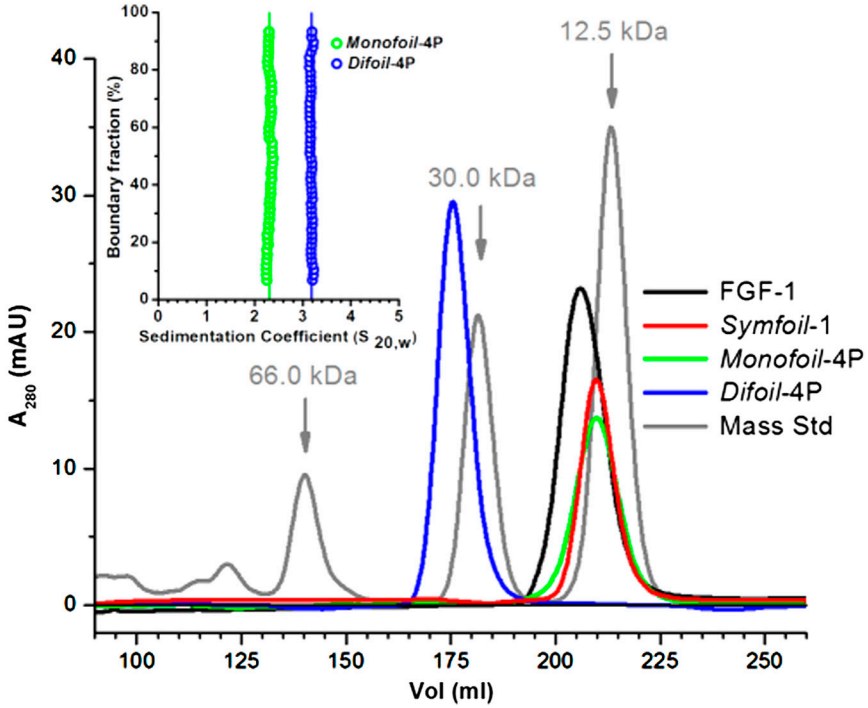

Fig. S3. Calibrated size-exclusion chromatography of FGF-1, Symfoil-1, Monofoil-4P, and Difoil-4P proteins. The Monofoil-4P polypeptide retention time is equivalent to that of the Symfoil-1 protein, indicating homotrimer assembly in solution. The Difoil-4P polypeptide retention time is approximately 36 kDa, also indicating homotrimer assembly in solution. The inset diagram shows the analytical ultracentrifuge sedimentation coefficients determined for the Monofoil-4P and Difoil-4P proteins and indicates homogeneous trimer assemblies for both polypeptides, with no evidence for monomeric or other multimeric assemblies.

Table S1. Nomenclature for the β-trefoil mutant proteins comprising the top-down symmetric deconstruction of the FGF-1 protein and leading to the Monofoil-4P peptide

Mutant	Composition
SYM2 (1, 2)	FGF-1/Leu73Val/Val109Leu
SYM3 (1, 2)	SYM2/Leu44Phe
SYM4 (2)	SYM3/Cys117Val
SYM5 (2)	SYM4/Leu111Ile
SYM6 (2)	SYM5/Met67Ile
SYM6ΔΔ (3)	SYM6/Ala103Gly/Δ104-106/Arg119Gly/Δ120-122
SYM7ΔΔ (4)	SYM6ΔΔ/Phe22Tyr/Phe108Tyr
SYM9ΔΔ	SYM7ΔΔ/Lys12Val/Pro134Val
SYM10ΔΔ	SYM9ΔΔ/His93Gly
SYM11ΔΔ	SYM10ΔΔ/Leu26Asn/Asp68Asn/Thr69Pro/Lys112Asn/Lys113Pro/Asn114Asp
SYM12ΔΔ	SYM11ΔΔ/Asn95Val/Leu46Val/Glu87Val
SYM13ΔΔ	SYM12ΔΔ/Ile56Leu/Tyr97Leu
Symfoil-1	SYM13ΔΔ chimera:{53}{12–14}{57–65}{24–47}{136}{90}{GGG}
Symfoil-2	Symfoil-1/Val46Ile/Val87Ile/Val134Ile
Symfoil-3	Symfoil-2/Gly51Asn/Gly92Asn/Gly139Asn
Symfoil-4T	Symfoil-3/Gln40Thr/Gln81Thr/Gln128Thr
Symfoil-4V	Symfoil-3/Gln40Val/Gln81Val/Gln128Val
Symfoil-4P	Symfoil-3/Gln40Pro/Gln81Pro/Gln128Pro
Difoil-4P	Symfoil-4P/Glu94stop
Monofoil-4P	Symfoil-4P/Glu53stop

1 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.

2 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.

3 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.

4 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.

Table S2. Crystallographic data collection and refinement statistics

	Symfoil-1	Symfoil-2	Symfoil-4T	Symfoil-4V	Symfoil-4P	Monofoil-4P*	Difoil-4P*
Data collection							
Space group	I222	$P2_1$	I222	I222	I222	$P2_12_12_1$	$I2_12_12_1$
Cell dimensions							
a, b, c, Å	50.4, 53.4, 85.2	50.7, 53.6, 85.3	50.6, 53.5, 85.0	50.8, 53.7, 85.6	50.4, 53.2, 84.8	49.5, 53.5, 65.9	81.2, 85.4, 86.1
α, β, γ, °	90.0, 90.0, 90.0	90.0, 90.1, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution, Å	50.00–1.45 (1.48–1.45)	50.00–1.45 (1.48–1.45)	50.00–1.80 (1.83–1.80)	50.00–1.75 (1.78–1.75)	50.00–1.65 (1.68–1.65)	50.00–1.48 (1.51–1.48)	50.00–2.85 (2.90–2.85)
R_{merge}	5.1 (37.8)	8.3 (39.8)	9.3 (37.9)	10.2 (34.5)	5.0 (33.3)	6.9 (35.8)	11.9 (36.2)
$I/\sigma I$	53.6 (3.6)	34.1 (2.9)	67.1 (10.1)	50.2 (8.1)	59.7 (4.5)	82.4 (8.4)	33.0 (4.3)
Completeness, %	98.6 (85.8)	97.8 (95.1)	99.0 (98.2)	98.3 (100)	96.0 (73.8)	99.0 (98.8)	94.5 (61.8)
Redundancy	7.0 (5.8)	5.1 (2.5)	9.7 (9.8)	5.4 (5.0)	9.3 (7.0)	19.4 (16.6)	6.1(3.2)
Refinement							
Resolution, Å	45.21–1.45	45.41–1.45	36.75–1.80	45.51–1.75	42.39–1.65	41.52–1.48	43.04–2.86
No. reflections	20,483	79,285	10,908	11,947	13,546	29,337	6,757
R_{work}/R_{free}	19.6/22.7	18.6/21.5	17.7/21.8	19.3/22.4	19.0/22.6	18.1/20.4	22.7/31.1
No. atoms							
Protein	998	4,209	987	992	996	1,062	1,912
Ligand/ion	13	92	19	20	19	25	15
Water	154	624	86	78	93	203	0
B factor							
Protein	23.4	19.2	26.3	29.3	26.8	18.1	60.6
Ligand/ion	27.5	31.6	39.7	40.9	38.7	52.5	56.5
Water	34.8	33.9	36.4	38.5	37.0	37.5	—
rms deviations							
Bond length, Å	0.007	0.006	0.007	0.006	0.007	0.006	0.002
Bond angle, °	1.14	1.09	1.10	1.05	1.12	1.07	0.549
PDB ID code	3O49	3O4A	3O4B	3O4C	3O4D	3OL0	3OGF

Each dataset was collected from a single crystal except Monofoil-4P and Difoil-4P. Two crystals were used for Monofoil-4P and three crystals were used for Difoil-4P dataset. Values in parentheses are for the highest-resolution shell. PDB, Protein Data Bank.
*Δ1-10 mutant form.

Table S3. Thermodynamic parameters for FGF-1, SYM, and Symfoil mutant proteins determined from isothermal equilibrium denaturation by GuHCl in ADA buffer

Protein	ΔG, kJ/mol	*m* value, kJ/mol M	C_m, M	ΔΔG, kJ/mol
FGF-1	21.1 ± 0.6	18.9 ± 0.6	1.11 ± 0.01	—
FGF-1* (1)	26.6 ± 0.9	20.3 ± 0.7	1.29 ± 0.01	—
Transform 1				
SYM5 (2)	20.8 ± 0.5	19.4 ± 0.1	1.07 ± 0.02	0.8
SYM6ΔΔ (3)	33.9 ± 0.6	17.7 ± 0.4	1.91 ± 0.02	−14.6
SYM7ΔΔ* (1)	41.2 ± 0.6	19.2 ± 0.8	2.14 ± 0.01	−16.8[†]
Transform 2				
SYM9ΔΔ	54.0 ± 1.4	16.7 ± 0.4	3.24 ± 0.01	−37.9
SYM10ΔΔ	66.0 ± 3.1	17.9 ± 0.9	3.68 ± 0.01	−47.3
SYM11ΔΔ	56.4 ± 3.2	16.2 ± 0.8	3.48 ± 0.02	−41.6
Transform 3				
SYM12ΔΔ	64.6 ± 0.7	16.5 ± 0.2	3.91 ± 0.01	−49.6
SYM13ΔΔ	55.6 ± 2.5	13.9 ± 0.5	4.00 ± 0.03	−47.4
Symfoil-1	28.8 ± 0.3	18.5 ± 0.2	1.55 ± 0.01	−8.2
		Stability optimization		
Symfoil-2	37.9 ± 0.2	17.4 ± 0.1	2.18 ± 0.01	−19.4
Symfoil-3	41.0 ± 0.1	17.3 ± 0.1	2.37 ± 0.01	−22.8
Symfoil-4T	43.2 ± 0.3	17.4 ± 0.1	2.48 ± 0.01	−24.9
Symfoil-4V	44.2 ± 0.1	15.4 ± 0.1	2.86 ± 0.01	−30.0
Symfoil-4P	55.1 ± 0.7	15.3 ± 0.3	3.60 ± 0.01	−42.6
Transform 4	**ΔG_0^0, kJ/mol**	***m* value, kJ/mol M**	**C_m, M**	
Monofoil-4P				
2 μM	71.8 ± 0.8	7.8 ± 0.4	1.64 ± 0.01	
4 μM	70.6 ± 0.2	7.9 ± 0.1	1.71 ± 0.04	
10 μM	67.4 ± 1.0	7.4 ± 0.5	1.64 ± 0.03	
Difoil-4P				
2 μM	83.8 ± 0.2	13.5 ± 0.1	1.72 ± 0.02	
4 μM	82.2 ± 0.8	13.8 ± 0.3	1.82 ± 0.02	
10 μM	82.1 ± 1.3	15.1 ± 0.4	1.96 ± 0.03	

1 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.

2 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.

3 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.

*Determined in crystallization buffer.

[†]ΔΔG in comparison to FGF-1 in crystallization buffer.