Supporting Information

An interface-driven design strategy yields a novel corrugated protein architecture

Mohammad ElGamacy, Murray Coles, Patrick Ernst[†], Hongbo Zhu, Marcus D. Hartmann, Andreas Plückthun[†] and Andrei N. Lupas[*]

Dept. of Protein Evolution, Max-Planck-Institute for Developmental Biology, 72076 Tübingen, Germany

[†]Dept. of Biochemistry, University of Zurich, 8057 Zurich, Switzerland

.

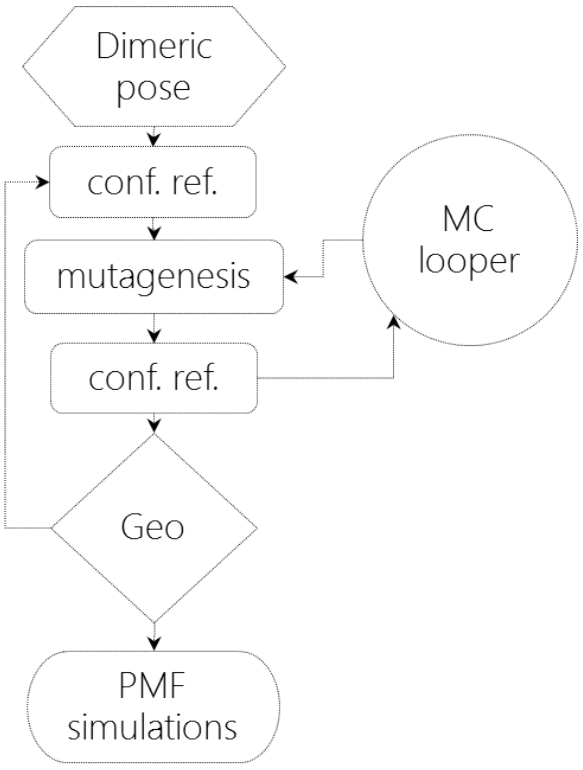

Figure S1. Interface design scheme. The conformational refinement steps consist of a FastRelax mover (repacking only in 2 rounds), a Backrub mover, and a DockingProtocol mover (repacking only with local refinement using soft_rep scoring, with a maximum rigid body perturbation of 3 Å and 2° per round). The mutagenesis step comprised three consecutive RepackMinimize movers (initially with soft_rep scoring and no backbone refinement, and later using talaris13 scoring and performing backbone refinement) and a BackrubDD mover. The GenericMonteCarlo mover was run for 5 loops and all of the output decoys were filtered geometrically and the top few were reused as interface design input. After 4 rounds for BRIC1, 5 rounds for BRIC2 and 7 rounds for BRIC3, the decoys generated were evaluated by the PMF simulations, where the ones exhibiting the highest dissociation work were chosen for loop design.

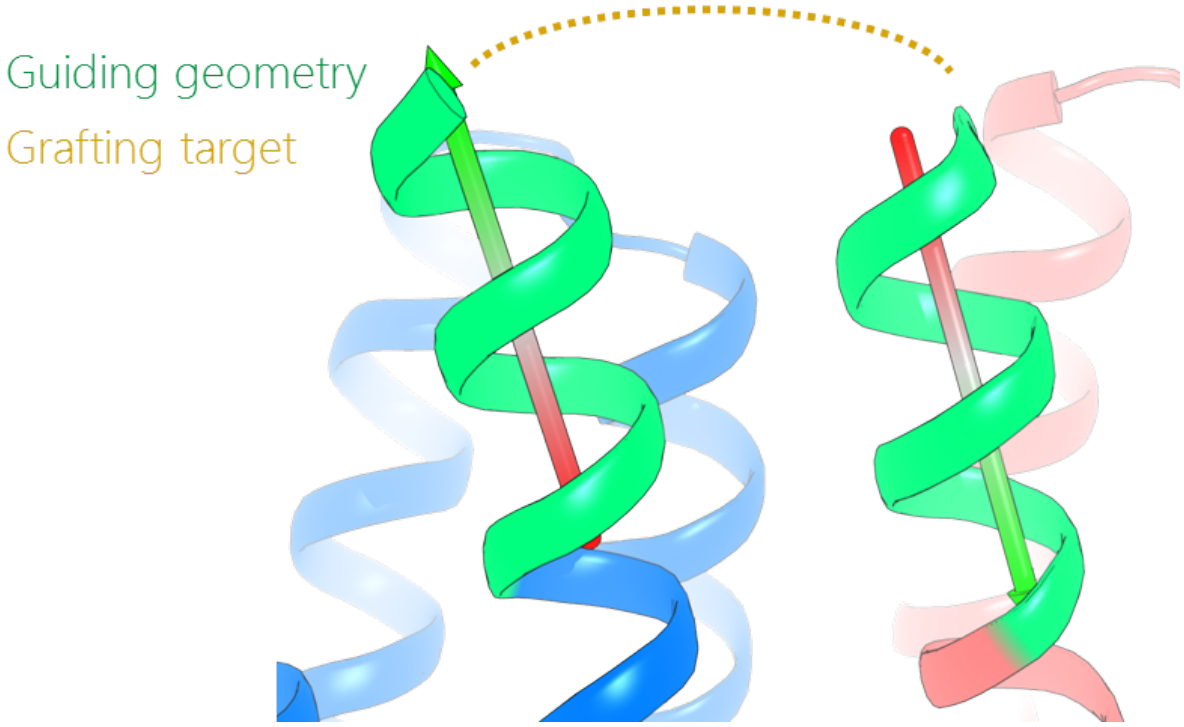

Figure S2. Loop grafting step. Initial loop compositions were fetched from natural structures using an alignment-free geometric matching method that compares the landing site vectors and their associated backbone dihedrals profile (The green stretches of the design interface represent the landing sites).

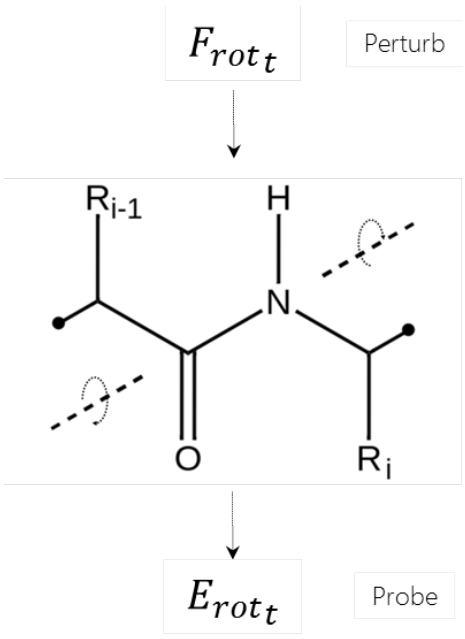

Figure S3. The rotational perturbation axis was taken to be the peptide bond axis, which would force dihedral sampling by diagonal traversal of the shifted-Ramachandran space (i.e. the $(\boldsymbol{\psi_{i-1}}, \boldsymbol{\phi_i})$-space), which was performed while monitoring the instantaneous kinetic energy response**.**

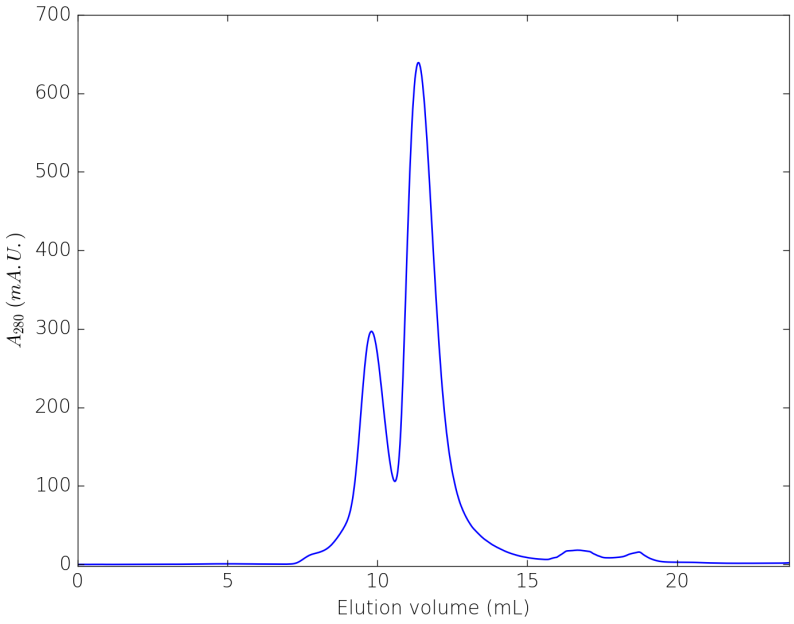

Figure S4. Analytical gel filtration of BRIC1 showing a major monomeric species and a minor dimeric species, using GE Superdex™ 75 10/300.

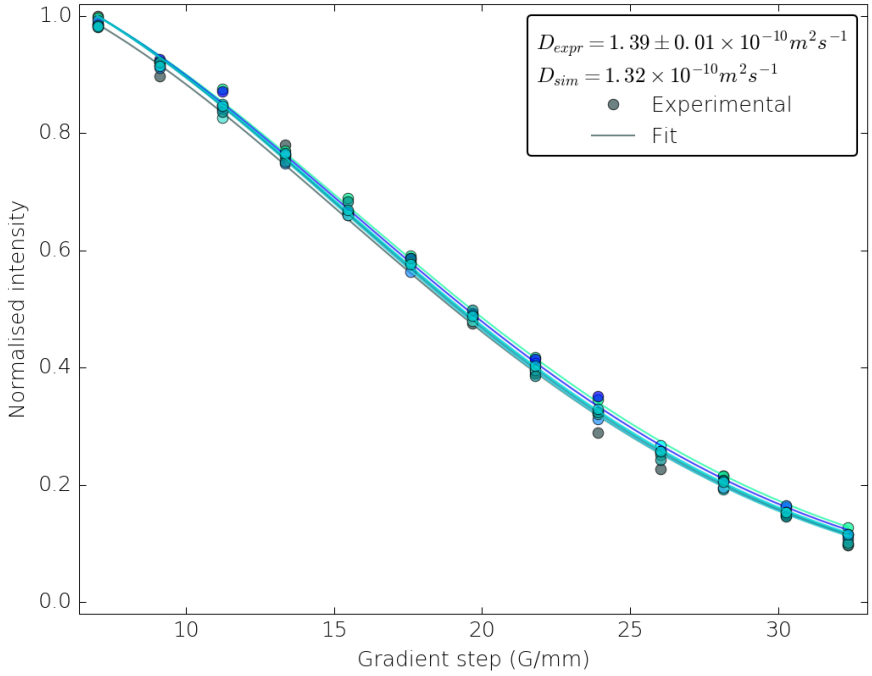

Figure S5. Designed coordinates best explain the diffusion-ordered spectroscopy experiments on BRIC1. Different colours designate data collected for eight different aliphatic proton peaks, where the legend shows the average and standard deviation values of the diffusion coefficient. The predicted translational diffusion coefficient value using the designed coordinates at the same temperature was 1.32×10^{-10}, while that of the swapped dimeric and the swapped monomeric forms were 6.81×10^{-11} and 1.18×10^{-10}, respectively.

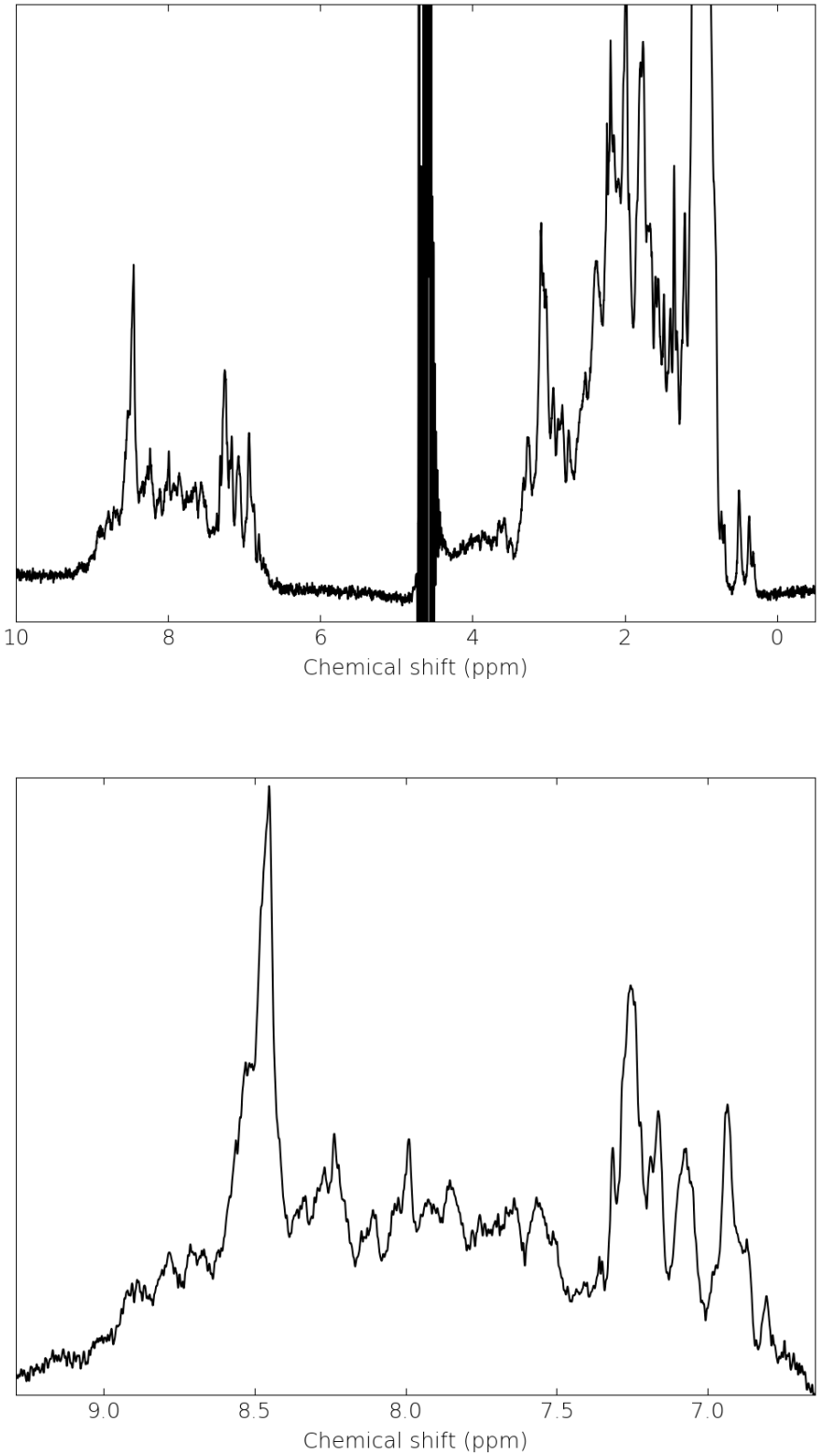

Figure S6. 1D ^{1}H NMR spectrum of BRIC2. The panes show dispersions along the full and the amide spectral ranges, respectively.

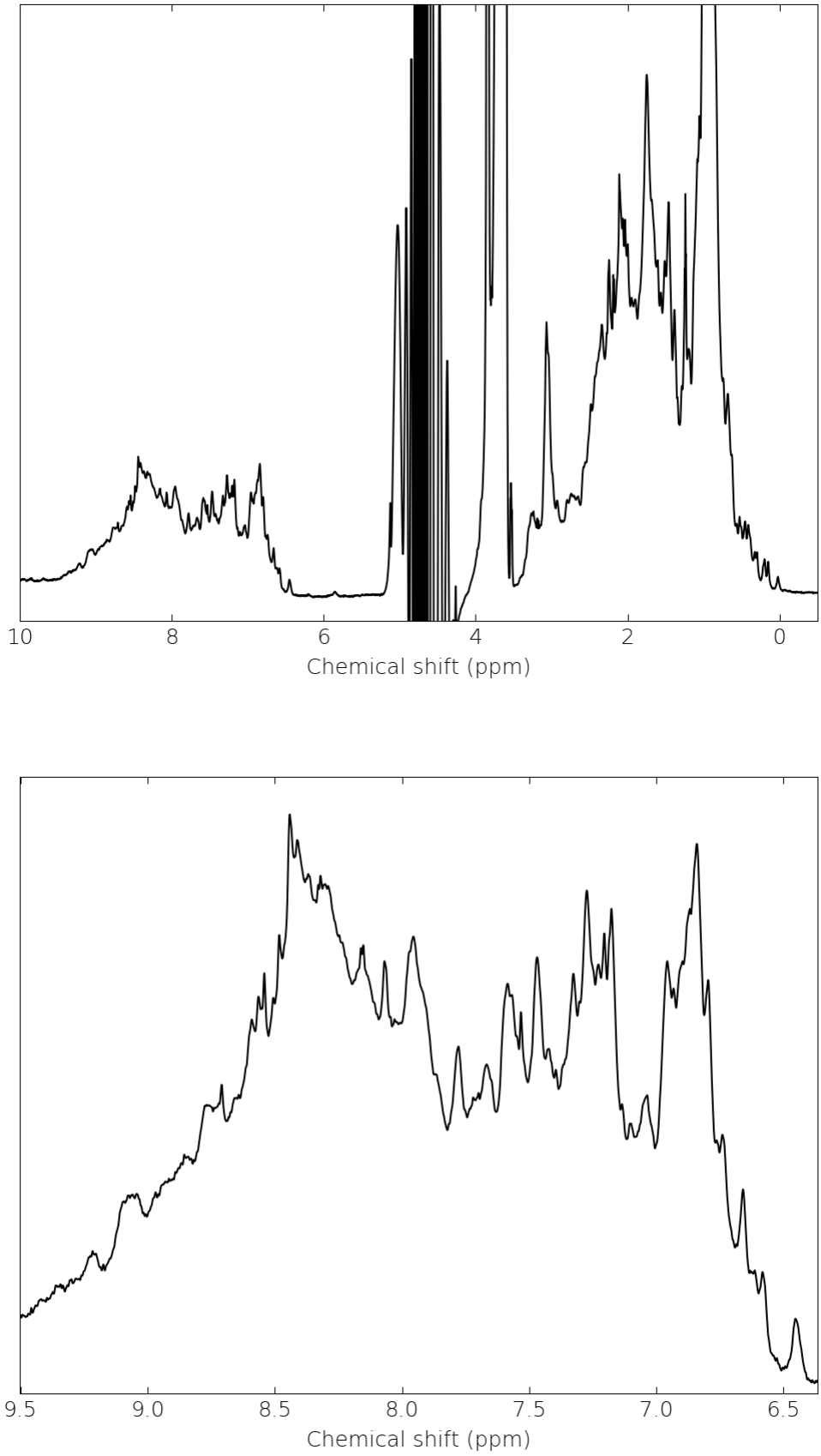

Figure S7. 1D ^{1}H NMR spectrum of BRIC3. The panes show dispersions along the full and the amide spectral ranges, respectively.

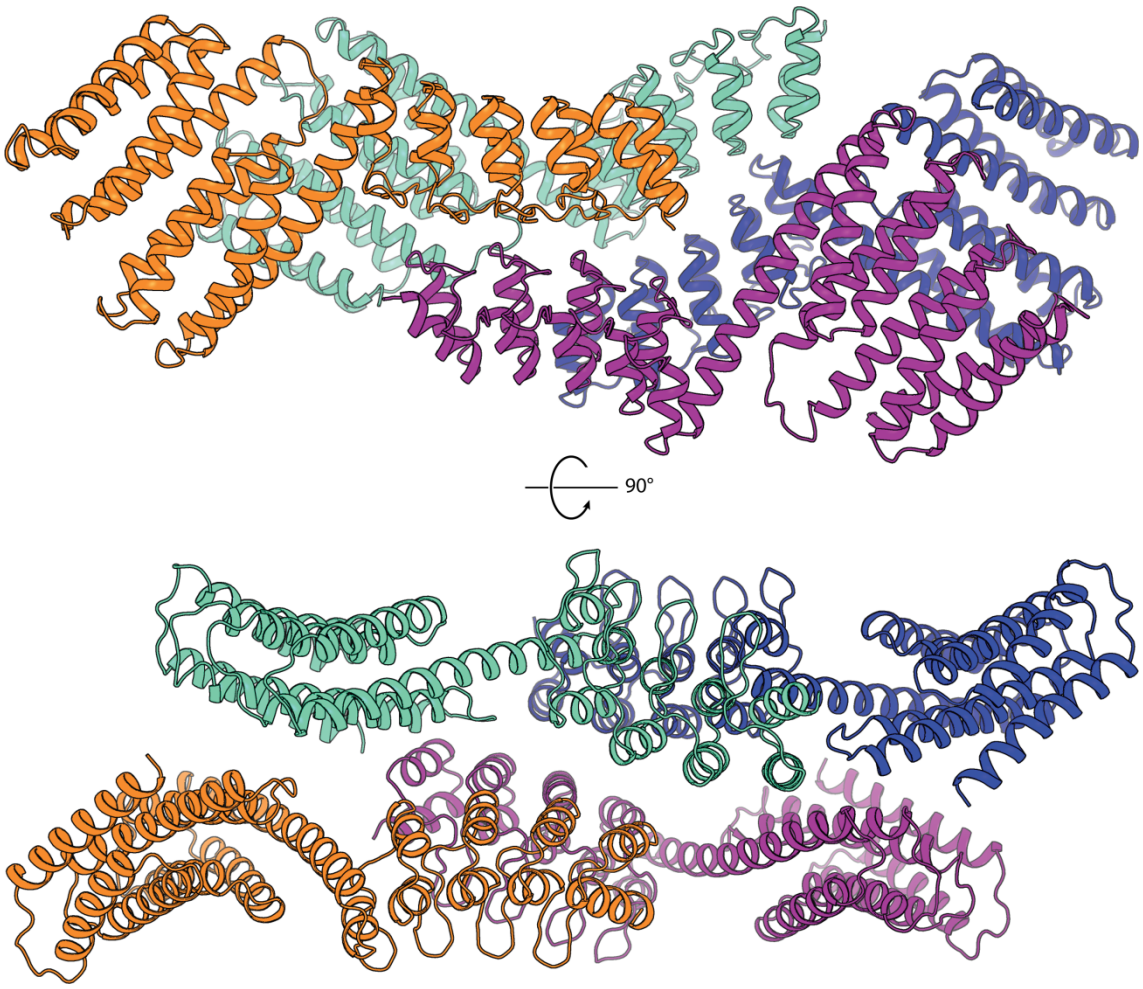

Figure S8. The asymmetric unit of the D12-BRIC2 fusion crystal structure. Different colours show the four chains in the asymmetric unit from two different views.

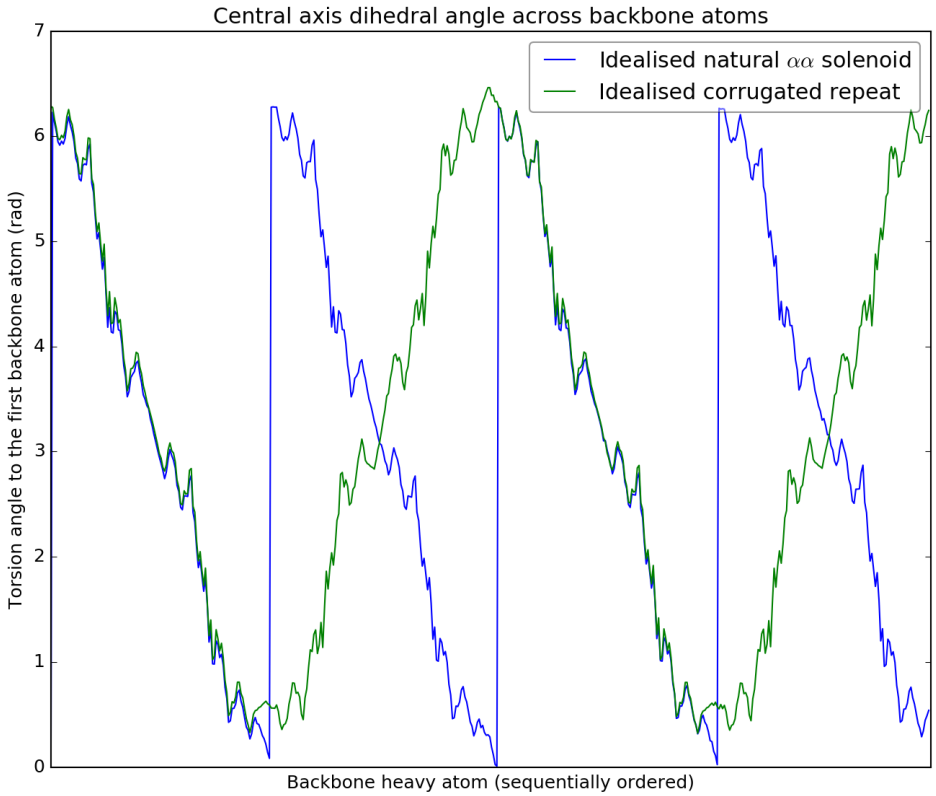

Figure S9. Superhelical axis torsion waveforms of α/α-solenoid *vs.* corrugated fold. The α/α-solenoid shows a clear saw-tooth pattern as compared to the triangle-wave pattern emergent from the more complex corrugated topology. The corrugated topology waveform exhibits double the phase-cycle of the equivalent solenoid illustrating the complexity increment. Both waveforms are distorted by high-frequency interference from the local α-helical pattern.

Table S1. Sequences of the final design constructs aligned structurally against their respective starting templates, with the highlight and underline designating the designed loop and C-terminal cap, respectively.

```
BRIC1           FYQTFFDEADELLADMEQHLLDLVPESPDAEQLNAIFRAAHSIKGGAGTFGFTMLQYAVE
1i5n_tmplt      FYQTFFDEADELLADMEQHLLDLVPESPDAEQLNAIFRAAHSIKGGAGTFGFTILQETTH
                ********************************************************:** :..

BRIC1           LMENMLDFARRGEMQLNTDIINLFLELKDLMQRMLDYYKKPQPCFYQAFFDMADVMLKVM
1i5n_tmplt      LMENLLDEARRGEMQLNTDIINLFLETKDIMQEQLDAYK-----FYQTFFDEADELLADM
                ****:** ****************** **:**. ** **     ***:*** ** :*  *

BRIC1           EQLLKLLVPESPDAAMLNAIFRAAHFIKGAAGTFGFTILQETTHLMENLLDEARRGEMQL
1i5n_tmplt      EQHLLDLVPESPDAEQLNAIFRAAHSIKGGAGTFGFTILQETTHLMENLLDEARRGEMQL
                ** *  ********  ********* ***.****************************

BRIC1           NTDIINLFLETKDIMQEQLDAYKNSEEPDAASFEYICNALRQLALEA
1i5n_tmplt      NTDIINLFLETKDIMQEQLDAYKNSEEPDAASFEYICNALRQLALEA
                *********************************************

BRIC2           YIKKVVDELKELIQNVNDDIKEVEKNPEDMEYWNKIYRLVHTMKEITETMGFSPVALVLE
2lch_tmplt      YIKKVTDELKELIQNVNDDIKEVEKNPEDMEYWNKIYRLVHTMKEITETMGFSSVAKVLH
                *****.*********************************************** ** **.

BRIC2           AIMMLVKLMLNSEIKITSDLIDAVKKMLDMVTRLLDLMVDPNLNEEQYIKMVVDALKILI
2lch_tmplt      TIMNLVDKMLNSEIKITSDLIDKVKKLDMVTRELDKKVS-------YIKKVTDELKELI
                :** **. ************** *** ****** **  *        *** *.* ** **

BRIC2           EAVNVLIKMVEKNPEDMEFWNLIYRLVHVMKEVTETMGFSSVAKVLHTIMNLVDKMLNSE
2lch_tmplt      QNVNDDIKEVEKNPEDMEYWNKIYRLVHTMKEITETMGFSSVAKVLHTIMNLVDKMLNSE
                : **  ** **********:** ******.***:**************************

BRIC2           IKITSDLIDKVKKKLDMVTRELDKMVS
2lch_tmplt      IKITSDLIDKVKKKLDMVTRELDKKVS
                ************************ **

BRIC3           DKVYENVTGLVKAVIEMSSKIQPAPPEEYVPMVKEVGLALRTLLATVDETIPLLPASTHR
3b71_tmplt      DKVYENVTGLVKAVIEMSSKIQPAPPEEYVPMVKEVGLALRTLLATVDETIPLLPASTHR
                ************************************************************

BRIC3           AIELMQELLNIALQLLEIAMKLAQQYVMTSAQQEHKKMMLMAAQVLAEIAKFLLDCITSP
3b71_tmplt      EIEMAQKLLNSDLGELINKMKLAQQYVMTSLQQEYKKQMLTAAHALAVDAKNLLD-----
                 **: *:***  *  *   *********** ***:** ** **:.**  ** ***

BRIC3           CVVYAAVQILVKFVEFMSKFIQPAPPELYVAMVKAVGKALRVLLAIVDMTIPLLPASTHR
3b71_tmplt      -KVYENVTGLVKAVIEMSSKIQPAPPEEYVPMVKEVGLALRTLLATVDETIPLLPASTHR
                  **  *  *** *  **. ******* ** *** ** ***.*** ** ***********

BRIC3           EIEMAQKLLNSDLGELINKMKLAQQYVMTSLQQEYKKQMLTAAHALAVDAKNLLDVIDQ
3b71_tmplt      EIEMAQKLLNSDLGELINKMKLAQQYVMTSLQQEYKKQMLTAAHALAVDAKNLLDVIDQ
                ************************************************************
```

Table S2. Crystallographic structure statistics for BRIC1

Data collection	
Space group	C2
Cell dimensions	
a, b, c (Å)	113.66, 41.95, 58.26
α, β, γ (°)	90.0, 90.46, 90.0
Resolution (Å)	40.84 – 2.50 (2.65 – 2.50)[a]
R_{merge}	0.067 (0.787)
$\langle I \rangle / \langle \sigma(I) \rangle$	12.1 (1.67)
Completeness (%)	99.7 (98.8)
Redundancy	6.55 (6.16)
Refinement	
Resolution (Å)	40.8 – 2.50 (2.79 – 2.50)
No. reflections	9739 (2721)
R_{work} / R_{free} (%)	22.9 / 27.9 (24.9 / 30.2)
No. atoms	1791
Protein	1791
<B-factors> (Å^2)	116.0
Protein	116.0
RMS deviations	
Bond lengths (Å)	0.010
Bond angles (°)	1.03

[a]Values in parentheses are for highest-resolution shell.

Table S3. Solution structure statistics for BRIC1

Restraint Violations[1]	
Inter-Helical Distance restraints (Å)	
All (39)	0.018 ± 0.003
N-terminal interface (11)	
C-terminal interface (15)	
Designed interface (13)	
Dihedral restraints (°)	
All (314)	0.082 ± 0.014
Covalent Geometry	
Bonds (Å × 10^{-3})	1.91 ± 0.04
Angles (°)	0.54 ± 0.01
Impropers (°)	0.88 ± 0.01
Structure Quality Indicators[2]	
Ramachandran Map (%)	98.2 / 1.6 / 0.2
Backbone Heavy Atom R.M.S.D (Å)[3]	
E vs <E>	0.84 ± 0.06
E vs design	1.78 ± 0.11
<E> vs design	1.57

[1] Violations are expressed as RMSD ± SD. The number of each restraint type is shown in brackets.

[2] Defined as the percentage of residues in the favored/allowed/outlier regions of the Ramachandran map as determined by MOLPROBITY (http://molprobity.biochem.duke.edu).

[3] Structures are labeled as follows: E, the final ensemble of 26 structures; <E>, the mean structure calculated by averaging the coordinates of E structures after fitting over ordered residues. RMSD values are based on superimpositions over ordered residues (defined as F1-L225)

Table S4. Crystallographic structure statistics for D12-BRIC2*

Data collection	
Space group	P1
Cell dimensions	
a, b, c (Å)	52.99, 63, 133.54
α, β, γ (°)	97.991, 91.923, 110.815
Resolution (Å)	49.33 – 3 (3.08 - 3.0)
R_{merge}	0.177 (1.308)
<I>/<σ(I)>	7.6 (2.15)
Completeness (%)	99.0 (99.8)
Redundancy	5.6 (5.8)
Refinement	
Resolution (Å)	44.9 – 3 (3.0969 - 3.0)
No. reflections	31544 (2901)
R_{work} / R_{free} (%)	25.09 / 29.19 (35.56 / 35.97)
No. atoms	11544
Protein	11420
Water	51
Ligands	73
<B-factors> (Å2)	97.91
Protein	98.06
RMS deviations	
Bond lengths (Å)	0.003
Bond angles (°)	0.66
Ramachandran plot	
Favoured	95.37
Allowed	4.22
Outliers	0.41

[a]Values in parentheses are for highest-resolution shell.

*Diffraction data from two crystals from the same drop were merged to improve completeness.