View Online / Journal Homepage
Dynamic Article Links

Chemical Science

Cite this: DOI: 10.1039/c2sc21117c

www.rsc.org/chemicalscience

EDGE ARTICLE

Remodeling a β-peptide bundle†

Matthew A. Molski,[a] Jessica L. Goodman,[b] Fang-Chieh Chou,[c] David Baker,*[d] Rhiju Das*[c] and Alanna Schepartz*[ae]

Received 7th April 2012, Accepted 10th September 2012
DOI: 10.1039/c2sc21117c

Published on 21 September 2012 on http://pubs.rsc.org | doi:10.1039/C2SC21117C
Downloaded by Stanford University on 24 September 2012

Natural biopolymers fold with fidelity, burying diverse side chains into well-packed cores and protecting their backbones from solvent. Certain β-peptide oligomers assemble into bundles of defined octameric stoichiometry that resemble natural proteins in many respects. These β-peptide bundles are thermostable, fold cooperatively, exchange interior amide N–H protons slowly, exclude hydrophobic dyes, and can be characterized at high resolution using X-ray crystallography – just like many proteins found in nature. But unlike natural proteins, all octameric β-peptide bundles contain a sequence-uniform hydrophobic core composed of 32 leucine side chains. Here we apply rational design principles, including the Rosetta computational design methodology, to introduce sequence diversity into the bundle core while retaining the characteristic β-peptide bundle fold. Using circular dichroism spectroscopy and analytical ultracentrifugation, we confirmed the prediction that an octameric bundle still assembles upon a major remodelling of its core: the mutation of sixteen core β-homo-leucine side chains into sixteen β-homo-phenylalanine side chains. Nevertheless, the bundle containing a partially β-homo-phenylalanine core poorly protects interior amide protons from exchange, suggesting molten-globule-like properties. We further improve stability by the incorporation of eight β-homo-pentafluorophenyalanine side chains, giving an assembly with amide protection factors comparable to prior well-structured bundles. By demonstrating that their cores tolerate significant sequence variation, the β-peptide bundles reported here represent a starting point for the "bottom-up" construction of β-peptide assemblies possessing both structure and sophisticated function.

Introduction

Natural biopolymers fold with fidelity, can exist as oligomers or discrete complexes, and possess kinetic and thermodynamic signatures that distinguish them from most non-biological polymers and smaller molecules. In 2007, we reported that certain oligomers of β^3-amino acids (β-peptides) fold into bundles of defined stoichiometry that resemble natural proteins in many respects.[1–9] The high-resolution structures of four β-peptide bundles[3,4,8,9] reveal a shared octameric fold composed of parallel and anti-parallel 3_{14}-helices, a salt-bridge-rich exterior, and a close-packed hydrophobic core. These β-peptide bundles are thermostable, undergo cooperative folding transitions, exchange interior amide N–H protons slowly, and exclude hydrophobic dyes, but contain a sequence-uniform core of 32 leucine side chains. Eliminating this side chain uniformity is a critical step toward the "bottom-up" construction of heterogeneous β-peptide assemblies possessing defined sizes, reproducible structures, and sophisticated function.[10–12] Computational methods have also recently been used to predict β-peptide sequences that assemble into stable quaternary assemblies.[13] Although these oligomers are not yet characterized structurally at high resolution, their sequences imply that they too possess uniform hydrophobic cores.[13–28] Structurally characterized bundles composed of both α- and β-amino acids have also been reported.[17,18,25]

In this work, we applied the Rosetta software package[29,30] to predict β-peptide sequences that could effectively recapitulate the structurally characterized β-peptide bundle core using a mixture of leucine and non-leucine side chains. One such sequence (Acid-1Y^{FF}), containing an equal number of core β-homo-phenylalanine and β-homo-leucine residues, assembles into a 3_{14}-helical,

[a]*Department of Chemistry, Yale University, New Haven, CT 06511, USA. E-mail: matthew.molski@yale.edu; Fax: +1 203-432-3486; Tel: +1 203-432-8276*
[b]*Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, MA 02142, USA. E-mail: jgoodman@wi.mit.edu; Fax: +1 617-258-7226; Tel: +1 617-258-5184*
[c]*Department of Biochemistry, Stanford University, Stanford, CA 94305, USA. E-mail: fcchou@stanford.edu; rhiju@stanford.edu; Fax: +1 650-723-5976; Tel: +1 650-723-7310*
[d]*Department of Biochemistry, University of Washington, Seattle, WA 98195, USA. E-mail: dabaker@u.washington.edu; Fax: +1 206-685-1792; Tel: +1 206-543-1295*
[e]*Departments of Chemistry and of Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT 06511, USA. E-mail: alanna.schepartz@yale.edu; Fax: +1 203-432-3486; Tel: +1 203-432-5094*

† Electronic supplementary information (ESI) available: β-Peptide design and synthesis and characterization using CD, SE-AU, and NMR. See DOI: 10.1039/c2sc21117c

Published on 21 September 2012 on http://pubs.rsc.org | doi:10.1039/C2SC21117C
Downloaded by Stanford University on 24 September 2012

relatively thermostable, octameric bundle. Despite this stability, Acid-1Y^{FF} displayed two properties associated with a molten globule state: rapid amide NH exchange and conformational heterogeneity as judged by NMR, indicating the potential for improved side chain organization within the core. It is well known that mixtures of phenylalanine and penta-fluorophenylalanine can improve protein stability when introduced into an otherwise all-hydrocarbon protein core.[31–40] To test whether the thermodynamic stability and structural uniqueness of Acid-1Y^{FF} could be improved by fluorocarbon substitution, we synthesized an analogue (Acid-1Y^{FF*}) of Acid-1Y^{FF} containing β-pentafluoro-homo-phenylalanine (F_5βPhe) at position 8. Acid-1Y^{FF*} displays improved folding properties, resulting in a more conformationally distinct and stable core as judged by NMR. Acid-1Y^{FF} and Acid-1Y^{FF*} are the first β-peptide bundles containing mixed sequence hydrophobic cores, suggesting further use of Rosetta and rational design principles as tools to remodel novel β-peptide bundles.

Results and discussion

The Rosetta software package[29,30] has been applied successfully to improve protein thermal stability,[41] design novel protein folds[42] and enzymes[43] and predict protein structure. More recently, the Rosetta approach has been extended to model and design RNA, another natural biopolymer.[44–46] We began this work by using Rosetta to evaluate the core residue preference of the β-peptide bundle Acid-1Y, which had been characterized previously using X-ray crystallography. Acid-1Y assembles into a D_2 symmetric octamer with two β-peptides in the asymmetric unit, each with three symmetry mates (Fig. 1A).[4] We stripped out the core side chains and performed a full side chain conformer search to evaluate potential variants containing any of the twenty canonical amino acid side chains at any of eight core positions (residues 2, 5, 8 and 11 on the structurally two non-equivalent beta-peptides; the other 24 of the 32 core positions were constrained by symmetry).[47] Rosetta predicted a uniform, all-leucine core as the most stable bundle among these 20 (ref. 8) variants, recovering the known sequence of Acid-1Y. Moreover, the Rosetta-modeled side-chain conformers superimposed with those seen in the Acid-1Y crystal structure in atomic detail (Table S1†), supporting use of Rosetta for further modeling.

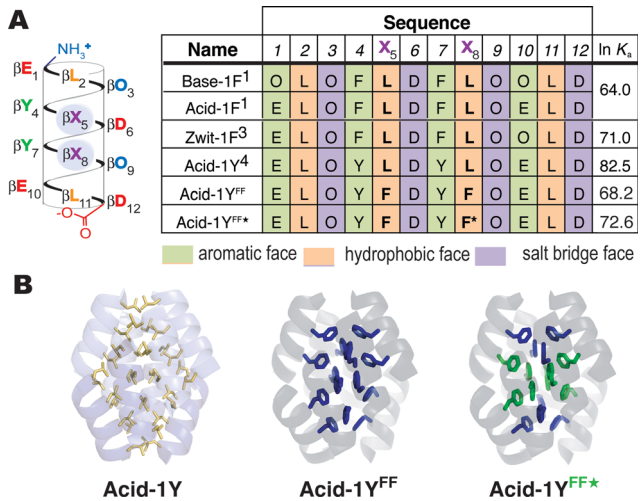

Fig. 1 β-peptide bundles studied in this work. (A) Helical net diagram and sequences of Base-1F, Acid-1F, Zwit-1F, Acid-1Y, Acid-1Y^{FF}, and Acid-1Y^{FF*}. F represents β-homo-phenylalanine; F* represents β-homo-pentafluorophenylalanine. Colors distinguish side chains on the aromatic (β-hY-containing), hydrophobic (β-hL containing) and salt bridge (β-hO- and β-hD-containing) faces. Also shown are ln K_a values characterizing each octameric assembly, as determined previously or in this work by SE-AU. (B) Ribbon representation of the crystal structure of Acid-1Y highlighting the packing of the leucine side chains, along with computationally predicted structure of Acid-1Y^{FF} and a color-coded guide to the locations of F_5β-hPhe (green) in Acid-1Y^{FF*}.

Next we searched for alternative side chains that could be accommodated in place of some or all of the β-homo-leucines within the bundle interior. While the fixed-backbone conformational search described above gave the all-leucine core as the optimal solution, a small void observed in the bundle center suggested that larger side chains might be tolerated at positions 5, 8 and/or 11 with minor backbone adjustments. We modeled the effects of introducing between four and eight alternative, proteinogenic side chains at these positions within Acid-1Y, and optimized the backbone torsion angles with a constraint potential tethering the angles to within ~10° of their crystallographic values. Rosetta calculations predicted variant bundles would generally be poorer in energy than the starting bundle. However, a few were predicted to possess better hydrophobic packing of non-polar side-chains, as assessed by the sum of van der Waals and solvation energies. In particular, a variant containing β-homo-phenylalanine residues at positions 5 and 8 (Acid-1Y^{FF}, Fig. 1A) gave more favorable hydrophobic packing energy by ~1 kT/monomer. Acid-1Y^{FF} contains 16 β-homo-leucine to β-homo-phenylalanine substitutions, placing an additional 48 carbon atoms into the bundle core. Nevertheless, the backbone atoms of the Acid-1Y^{FF*} bundle were shifted by less than 0.8 Å RMSD from those in the starting bundle (Fig. 1B), suggesting that the bulky aromatic side chains in the interior could be accommodated without disrupting the octamer.

The β-peptide monomer Acid-1Y^{FF} was prepared using solid phase, microwave-assisted methods, purified to homogeneity by HPLC, and characterized initially using wavelength-dependent circular dichroism (CD) spectroscopy. As predicted, Acid-1Y^{FF} underwent a concentration-dependent increase in 3_{14}-helical structure (as judged by the molar residue ellipticity at 209 nm, MRE_{209})[48] between 12 and 200 μM (Fig. 2a), consistent with an equilibrium between a partially structured monomer and a folded oligomer. A plot of MRE_{209} *vs.* [Acid-1Y^{FF}] was first fit to a monomer–octamer equilibrium with ln K_a = 66.9 ± 0.5, suggesting that oligomerization of Acid-1Y^{FF} was less favorable than that of the Acid-1Y bundle (ln K_a = 82.5 ± 1.8).[4] The fit, however, was imperfect [$P = 2 \times 10^{-10}$; see Table S3 and Fig. S1†], with a closer agreement at higher concentrations ([Acid-1Y^{FF}] > 50 μM) than at lower concentrations ([Acid-1Y^{FF}] < 50 μM). A plot of MRE_{209} *vs.* [Acid-1Y^{FF}] was subsequently fit to alternative, three-state models containing either a dimeric (1–2–8) or tetrameric (1–4–8) intermediate (Fig. 2a). The resulting association constants for the 1–2–8 model were ln K_{a1} = 18.4 ± 4.1 and ln K_{a2} = 79.2 ± 9.8, and those for the 1–4–8 model were ln K_{a1} = 36.8 ± 3.2 and ln K_{a2} = 73.7 ± 11.2. Both three-state models fit the CD data substantially better than the original

This journal is © The Royal Society of Chemistry 2012

Published on 21 September 2012 on http://pubs.rsc.org | doi:10.1039/C2SC21117C
Downloaded by Stanford University on 24 September 2012

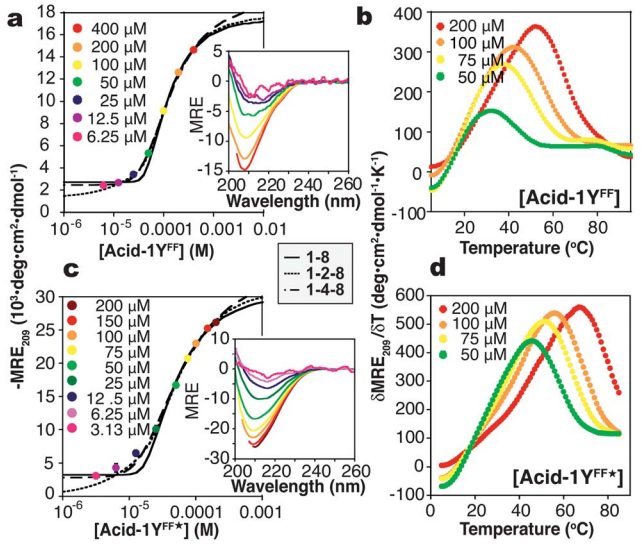

Fig. 2 Self-association of designed β-peptide bundles. Circular dichroism spectra of Acid-1Y^{FF} (a and b) and Acid-1Y^{FF*} (c and d) as a function of concentration (a and c) and temperature (b and d). Plots of MRE_{209} as a function of [β-peptide] were fit to a monomer–dimer–octamer (1–2–8) equilibrium (dotted line), a monomer–tetramer–octamer (1–4–8) equilibrium (dashed line) or monomer–octamer equilibrium (solid line). Inset: wavelength-dependent CD spectra of Acid-1Y^{FF} and Acid-1Y^{FF*} (MRE in units of 10^3 deg cm^2 dmol^{-1}). The T_M is defined as the maximum of a plot of $\delta MRE_{209} \cdot \delta T^{-1}$ *versus* temperature.

two-state monomer–octamer model (1–8), albeit with greater uncertainties in equilibrium constants due to the added fit parameters. Both fits gave predictions within the error of the data (Table S3 and Fig. S1†). Temperature-dependent CD studies also supported formation of a relatively stable Acid-1Y^{FF} bundle; the T_M of a 200 μM solution of Acid-1Y^{FF} was 52 °C (Fig. 2b); this value is lower than that of Acid-1Y, whose T_M was 82 °C at 150 μM.[4] Taken together, these CD data established the relative stability of the Acid-1Y^{FF} bundle, but could not precisely define the stoichiometry of the putative intermediate or the difference in stability between Acid-1Y and Acid-1Y^{FF}.

We therefore turned to sedimentation equilibrium analytical ultracentrifugation to more precisely characterize the stoichiometry and relative stability of the Acid-1Y^{FF} bundle. Sedimentation of Acid-1Y^{FF} at concentrations of 20, 80 and 200 μM was monitored at four speeds (36 000, 42 000, 50 000 and 60 000 RPM). The AU data was fit to both two-state monomer-*n*-mer equilibrium models as well as three-state models proceeding through a dimer or tetramer intermediate (1–2–8 or 1–4–8, respectively). Poor fits with high RMSD values and larger and more systematic residuals were observed when *n* in the two-state model was set to any value other than 8 between 2 and 10 (see Fig. S2†). Both the 1–2–8 and the 1–4–8 model fit the AU data better than the two state 1–8 model ($P < 10^{-6}$; *F*-test, see ESI†), in agreement with the CD analysis, and the 1–4–8 model produced the best fit. The ln K_a values calculated from the monomer–tetramer–octamer fit were ln $K_{a1} = 27.6 \pm 0.2$ and ln $K_{a2} = 68.2 \pm 0.3$, and the position of the monomer–octamer equilibrium (ln K_{a2}, 68.2 ± 0.3) agrees within error with the ln K_{a2} value determined by CD (73.7 ± 11.2). Taken together, the CD and AU data suggest that Acid-1Y^{FF} assembly is a three state process involving a tetrameric intermediate towards a final octameric assembly, and that the oligomerization of Acid-1Y^{FF} is less favorable than that of Acid-1Y (ln $K_a = 82.5 \pm 1.8$).[4]

Many *de novo* designed proteins exist as molten globules, and thus we hypothesized that the lower thermodynamic stability of the Acid-1Y^{FF} bundle compared to Acid-1Y might signal the presence of an undefined or heterogeneous hydrophobic core.[49] Molten globules often bind and increase the fluorescence of dyes such as 1-anilino-8-naphthalenesulfonate (ANS) by factors $\Delta F > 100$.[49] By contrast, well-folded or unfolded proteins do not provide favorable ANS binding sites, and elicit little or no change in ANS fluorescence ($\Delta F < 10$). All octameric β-peptide bundles reported thus far, including the Acid-1Y bundle, behave like well-folded proteins, causing minimal (<2-fold) changes in ANS fluorescence. Like previously characterized β-peptide bundles, even at concentrations as high as 200 μM (70% octamer), Acid-1Y^{FF} had little or no effect on the fluorescence of ANS ($\Delta F = 2$) (Fig. S4†). This analysis suggests that the Acid-1Y^{FF} β-peptide bundle provides only limited access of solvent to its hydrophobic core.

Greater insight into the differences between the Acid-1Y^{FF} bundle relative to Acid-1Y was revealed by NMR experiments that monitored the rate of amide NH hydrogen/deuterium exchange. Unlike all previously characterized β-peptide bundles,[4–6,9] the Acid-1Y^{FF} spectrum revealed no slowly exchanging amide NH protons; no amide NH resonances were visible as soon as 15 min after addition of D_2O at 25 °C (Fig. 3a). In addition, relative to the NMR spectrum of Acid-1Y,[4] the NMR spectrum of Acid-1Y^{FF} showed significant line broadening in the adjacent aromatic region. These data suggest that the aromatic sub-core of the Acid-1Y^{FF} bundle, although octameric and relatively thermostable, possesses significant conformational heterogeneity on the NMR time scale.

Fluorocarbon side chains are more hydrophobic than their hydrocarbon counterparts[50] and mixtures of phenylalanine and penta-fluorophenylalanine can improve protein stability when introduced into an otherwise all-hydrocarbon protein core.[31–40] In certain cases, favorable face-to-face interactions between phenylalanine and pentafluorophenylalanine side chains account for improved stability,[51–54] while in others steric and/or hydrophobic

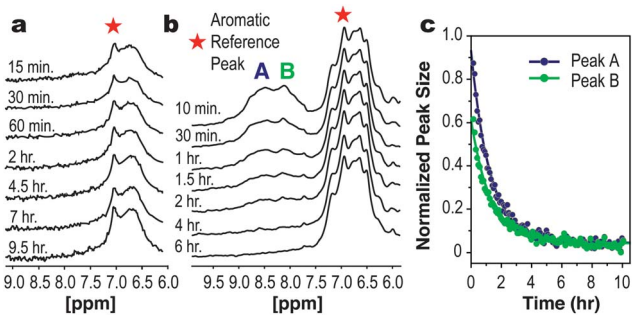

Fig. 3 Selective substitution of β-homo-pentafluorophenylalanine improves β-peptide bundle stability as judged by NMR. (a) The amide NH regions of the Acid-1Y^{FF} and (b) Acid-1Y^{FF*} β-peptide bundles after dissolution in D_2O. In the case of Acid-1Y^{FF}, no amide NH resonances are visible within 15 min after addition of D_2O. In the case of Acid-1Y^{FF*}, the amide signals persist for more than 1 h. (c) Integration of the resonances indicated in (B) normalized to the aromatic reference peak and fit to exponential decays to determine average exchange rate constants.

Published on 21 September 2012 on http://pubs.rsc.org | doi:10.1039/C2SC21117C
Downloaded by Stanford University on 24 September 2012

effects have been invoked.[33,55] To evaluate whether the thermodynamic stability and structural uniqueness of Acid-1Y^{FF} could be improved by fluorocarbon substitution, we synthesized an analogue of Acid-1Y^{FF} containing β-pentafluoro-homo-phenylalanine (F_5βPhe) at position 8 (Acid-1Y^{FF*}). Previous work has shown that β-peptide bundles containing β-homo-hexafluoroleucine at position 8 contain a discrete fluorous sub-domain and retain the characteristic β-peptide bundle fold.[8]

The Acid-1Y^{FF*} sequence was synthesized and its assembly characterized by CD, SE-AU, ANS binding, and NMR to evaluate the effects of β-homo-pentafluorophenylalanine substitution on β-peptide bundle structure and stability. As expected, the CD spectrum of Acid-1Y^{FF*} was characterized by concentration-dependent changes in MRE_{212}, and the data was first fit to a monomer–octamer equilibrium with a ln K_a of 73.0 ± 0.5. Again the fits deviated from the experimental data in the lower concentration range ([Acid-1Y^{FF*}] < 25 μM (Fig. S1†) and could be improved by including a dimer or tetramer intermediate along the folding pathway. Fitting to a monomer–dimer–octamer equilibrium (1–2–8) resulted in ln K_a values of 18.4 ± 4.0 (ln K_{a1}) and 82.9 ± 2.3 (ln K_{a2}) (Fig. 2c), and fitting to a monomer–tetramer–octamer equilibrium (1–4–8) yielded ln K_a values of 41.4 ± 4.6 (ln K_{a1}) and 84.7 ± 3.0 (ln K_{a2}). Both fits agreed with the CD data within error (Fig. S1 and Table S3†). Although measured with large errors, the ln K_{a2} values suggested that the Acid-1Y^{FF*} octamer stability was greater than the initial Acid-1Y^{FF} design and now comparable to the starting Acid-1Y bundle with the all-leucine core (ln K_a of 82.5 ± 1.8). The temperature-dependence of the Acid-1Y^{FF*} CD signal was consistent with improved stability; the T_M of Acid-1Y^{FF*} at 200 μM was 67 °C (Fig. 2d), a significantly higher value than the T_M of 52 °C for Acid-1Y^{FF} at the same concentration.

We turned again to SE-AU to more precisely and accurately determine the relative association constants of the Acid-1Y^{FF} and Acid-1Y^{FF*} bundles. The equilibrium sedimentation of Acid-1Y^{FF*}, performed at 20, 80 and 200 μM, was fit to two-state and three-state equilibrium models as described previously. Once again, the data fit best to a three-state model with a tetramer intermediate (1–4–8 model), although the 1–2–8 model was also an improvement over a two-state monomer–octamer equilibrium (P-value < 10^{-6} for both three state models according to the F-test; see ESI†). The ln K_a values calculated from the 1–2–8 fit were ln K_{a1} = 11.6 ± 0.4 and ln K_{a2} = 79.7 ± 1.3, and those from the 1–4–8 fit were ln K_{a1} = 29.7 ± 0.2 and ln K_{a2} = 72.6 ± 0.2 (Fig. S3†). Comparison to the AU analysis of the initial design indicates that, irrespective of model, the Acid-1Y^{FF*} bundle is significantly more stable than the Acid-1Y^{FF} bundle.

The effects of perfluoro substitution on the aromatic β-peptide bundle core were studied further by monitoring the effect of Acid-1Y^{FF*} on the intrinsic fluorescence of ANS (see ESI†). In previous work, designed α-amino acid proteins containing fluorous cores induced larger increases in ANS fluorescence than analogous proteins containing hydrocarbon cores; for example, the perfluorinated helical bundle α4-F2 increased ANS fluorescence by 5-fold.[56] This value is greater than that observed in the presence of α4 (ΔF = 3), but less than that seen in the presence of classic molten globule states such as lactalbumin ($\Delta F \geq 135$).[57,58] Whereas the Acid-1Y^{FF} bundle, like the Acid-1Y bundle, led to little or no increase in ANS fluorescence intensity (ΔF = 2-fold at [Acid-1Y^{FF}] = 200 μM), at an equivalent concentration the Acid-1Y^{FF*} bundle caused ANS fluorescence to increase >19-fold (Fig. S4†). Interestingly, the ΔF of ANS observed in the presence of Acid-1Y^{FF*} is comparable to that observed in the presence of the Zwit-8L*, in which eight leucine side chains were substituted by hexafluoroleucine (ΔF = 15), and whose structure is known at atomic detail.[8] Thus, it may be that the increased fluorescence of ANS in the presence of the Acid-1Y^{FF*} and Zwit-8L* results from an intrinsic affinity of ANS for fluorous domains, and not necessarily from a molten core.

Finally, H/D exchange NMR experiments gave the most incisive assessment of whether introduction of β-homo-pentafluorophenylalanine into the Acid-1Y^{FF} β-peptide bundle core would impart greater stability and structural uniqueness. Whereas the amide NH resonances in the Acid-1Y^{FF} bundle exchanged completely within 15 min, those of Acid-1Y^{FF*} were better resolved and exchanged over the course of hours (Fig. 3b and c). The average rate of H/D exchange for two regions of the Acid-1Y^{FF*} spectrum each fit a first-order decay function with a rate constant (k_{ex}) of 2.2 × 10^{-4} s^{-1}. By comparison, the exchange rate constant for the random coil model poly-β-alanine (β^3-homo-glycine) (k_{rc}) is 4.4 s^{-1}. The ratio of these values (k_{rc}/k_{ex}), often defined as a protection factor P, is 2.0 × 10^4, a value comparable to that of previously reported β-peptide bundles (P between 9 × 10^3 and 6 × 10^4),[5,6] including Acid-1Y (P = 6 × 10^4).[4] This equivalence of the P values calculated for Acid-1Y^{FF*} underscores the effect that the F_5βPhe substitution has on the stability of the diversified hydrophobic core. The octamer bundle containing a mixed core of β-homo-leucine, β-homo-phenylalanine, and β-homo-pentafluorophenylalanine is characterized by monomer–octamer equilibrium constants and NMR amide hydrogen exchange rates that are comparable to previous bundles containing homogeneous cores.

Conclusions

The design of interfaces within a β-peptide bundle – a cooperatively folded structure that lacks even a single natural α-amino acid – is perhaps the most stringent test of our understanding of the principles that guide interactions between proteins. In this report, we apply the Rosetta computational algorithm and rational design to introduce sequence diversity into an octameric β-peptide bundle core that contains a uniform array of 32 leucine side chains. Using circular dichroism spectroscopy and analytical ultracentrifugation, we confirm that the Rosetta-remodeled bundle remains octameric upon changing sixteen core β-homo-leucine side chains into sixteen β-homo-phenylalanine side chains. We improve this design, bringing its stability to the level of the starting β-homo-leucine core, by introducing penta-fluorophenylalanine at selective positions within the bundle core. The repacking of *de novo* designed helical bundle proteins has been reported previously,[59] although never in the context of β-peptides. This work represents the first example in which Rosetta is applied successfully to design a wholly non-biological polymer.[60,61] By demonstrating that their hydrophobic cores will tolerate significant sequence variation, the β-peptide bundles reported here represent a starting point for the "bottom-up" construction of β-peptide assemblies possessing defined sizes, reproducible structures, and sophisticated function.

This journal is © The Royal Society of Chemistry 2012

Published on 21 September 2012 on http://pubs.rsc.org | doi:10.1039/C2SC21117C
Downloaded by Stanford University on 24 September 2012

Notes and references

1 J. X. Qiu, E. J. Petersson, E. E. Matthews and A. Schepartz, *J. Am. Chem. Soc.*, 2006, **128**, 11338.
2 A. D. Bautista, C. J. Craig, E. A. Harker and A. Schepartz, *Curr. Opin. Chem. Biol.*, 2007, **11**, 685.
3 D. S. Daniels, E. J. Petersson, J. X. Qiu and A. Schepartz, *J. Am. Chem. Soc.*, 2007, **129**, 1532.
4 J. L. Goodman, E. J. Petersson, D. S. Daniels, J. X. Qiu and A. Schepartz, *J. Am. Chem. Soc.*, 2007, **129**, 14746.
5 E. J. Petersson, C. J. Craig, D. S. Daniels, J. X. Qiu and A. Schepartz, *J. Am. Chem. Soc.*, 2007, **129**, 5344.
6 J. L. Goodman, M. A. Molski, J. Qiu and A. Schepartz, *ChemBioChem*, 2008, **9**, 1576.
7 E. J. Petersson and A. Schepartz, *J. Am. Chem. Soc.*, 2008, **130**, 821.
8 M. A. Molski, J. L. Goodman, C. J. Craig, H. Meng, K. Kumar and A. Schepartz, *J. Am. Chem. Soc.*, 2010, **132**, 3658.
9 C. J. Craig, J. L. Goodman and A. Schepartz, *ChemBioChem*, 2011, **12**, 1035.
10 V. Nanda and R. L. Koder, *Nat. Chem.*, 2009, **2**, 15.
11 J. W. Bryson, S. F. Betz, H. S. Lu, D. J. Suich, H. X. Zhou, K. T. O'Neil and W. F. DeGrado, *Science*, 1995, **270**, 935.
12 P. B. Harbury, J. J. Plecs, B. Tidor, T. Alber and P. S. Kim, *Science*, 1998, **282**, 1462.
13 I. V. Korendovych, Y. H. Kim, A. H. Ryan, J. D. Lear, W. F. Degrado and S. J. Shandler, *Org. Lett.*, 2010, **12**, 5142.
14 S. Azeroual, J. Surprenant, T. D. Lazzara, M. Kocun, Y. Tao, L. A. Cuccia and J. M. Lehn, *Chem. Commun.*, 2012, **48**, 2292.
15 L. Fischer and G. Guichard, *Org. Biomol. Chem.*, 2010, **8**, 3101.
16 J. M. Fletcher, A. L. Boyle, M. Bruning, G. J. Bartlett, T. L. Vincent, N. R. Zaccai, C. T. Armstrong, E. H. C. Bromley, P. J. Booth, R. L. Brady, A. R. Thomson and D. N. Woolfson, *ACS Synth. Biol.*, 2012, **1**, 240.
17 M. W. Giuliano, W. S. Horne and S. H. Gellman, *J. Am. Chem. Soc.*, 2009, **131**, 9860.
18 W. S. Horne, J. L. Price, J. L. Keck and S. H. Gellman, *J. Am. Chem. Soc.*, 2007, **129**, 4178.
19 R. Kudirka, H. Tran, B. Sanii, K. T. Nam, P. H. Choi, N. Venkateswaran, R. Chen, S. Whitelam and R. N. Zuckermann, *Biopolymers*, 2011, **96**, 586.
20 S. Kwon, A. Jeon, S. H. Yoo, I. S. Chung and H. S. Lee, *Angew. Chem., Int. Ed.*, 2010, **49**, 8232.
21 S. Kwon, K. Kang, A. Jeon, J. H. Park, I. S. Choi and H. S. Lee, *Tetrahedron*, 2012, **68**, 4368.
22 S. Kwon, H. S. Shin, J. Gong, J. H. Eom, A. Jeon, S. H. Yoo, I. S. Chung, S. J. Cho and H. S. Lee, *J. Am. Chem. Soc.*, 2011, **133**, 17618.
23 I. M. Mandity, L. Fulop, E. Vass, G. K. Toth, T. A. Martinek and F. Fulop, *Org. Lett.*, 2010, **12**, 5584.
24 W. C. Pomerantz, V. M. Yuwono, R. Drake, J. D. Hartgerink, N. L. Abbott and S. H. Gellman, *J. Am. Chem. Soc.*, 2011, **133**, 13604.
25 J. L. Price, W. S. Horne and S. H. Gellman, *J. Am. Chem. Soc.*, 2007, **129**, 6376.
26 S. Segman-Magidovich, M. R. Lee, V. Vaiser, B. Struth, S. H. Gellman and H. Rapaport, *Chem.–Eur. J.*, 2011, **17**, 14857.
27 E. Torres, E. Gorrea, K. K. Burusco, E. Da Silva, P. Nolis, F. Rua, S. Boussert, I. Diez-Perez, S. Dannenberg, S. Izquierdo, E. Giralt, C. Jaime, V. Branchadell and R. M. Ortuno, *Org. Biomol. Chem.*, 2010, **8**, 564.
28 E. Torres, J. Puigmarti-Luis, A. P. del Pino, R. M. Ortuno and D. B. Amabilino, *Org. Biomol. Chem.*, 2010, **8**, 1661.
29 R. Das and D. Baker, *Annu. Rev. Biochem.*, 2008, **77**, 363.
30 A. Leaver-Fay, M. Tyka, S. M. Lewis, O. F. Lange, J. Thompson, R. Jacak, K. W. Kaufmann, P. D. Renfrew, C. A. Smith, W. Sheffler, I. W. Davis, S. Cooper, A. Treuille, D. J. Mandell, F. Richter, Y.-E. Ban, S. J. Fleishman, J. E. Corn, D. E. Kim, S. Lyskov, M. Berrondo, S. Mentzer, Z. Popovic, J. J. Havranek, J. Karanicolas, R. Das, J. Meiler, T. Kortemme, J. J. Gray, B. Kuhlman, D. Baker and P. Bradley, *Methods Enzymol.*, 2011, **487**, 545.
31 B. Bilgicer and K. Kumar, *Tetrahedron*, 2002, **58**, 4105.
32 B. Bilgicer, X. Xing and K. Kumar, *J. Am. Chem. Soc.*, 2001, **123**, 11815.
33 B. C. Buer, R. de la Salud-Bea, H. M. A. Hashimi and E. N. G. Marsh, *Biochemistry*, 2009, **48**, 10810.
34 B. C. Buer, J. L. Meagher, J. A. Stuckey and E. N. G. Marsh, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 4810.
35 H. Y. Lee, K. H. Lee, H. M. Al-Hashimi and E. N. G. Marsh, *J. Am. Chem. Soc.*, 2006, **128**, 337.
36 K. H. Lee, H. Y. Lee, M. M. Slutsky, J. T. Anderson and E. N. G. Marsh, *Biochemistry*, 2004, **43**, 16277.
37 A. J. Link, M. L. Mock and D. A. Tirrell, *Curr. Opin. Biotechnol.*, 2003, **14**, 603.
38 E. Neil and G. Marsh, *Chem. Biol.*, 2000, **7**, R153.
39 Y. Tang, G. Ghirlanda, W. A. Petka, T. Nakajima, W. F. DeGrado and D. A. Tirrell, *Angew. Chem., Int. Ed.*, 2001, **40**, 1494.
40 N. C. Yoder, D. Yuksel, L. Dafik and K. Kumar, *Curr. Opin. Chem. Biol.*, 2006, **10**, 576.
41 G. Dantas, B. Kuhlman, D. Callender, M. Wong and D. Baker, *J. Mol. Biol.*, 2003, **332**, 449.
42 B. Kuhlman, G. Dantas, G. C. Ireton, G. Varani, B. L. Stoddard and D. Baker, *Science*, 2003, **302**, 1364.
43 J. B. Siegel, A. Zanghellini, H. M. Lovick, G. Kiss, A. R. Lambert, J. L. St.Clair, J. L. Gallaher, D. Hilvert, M. H. Gelb, B. L. Stoddard, K. N. Houk, F. E. Michael and D. Baker, *Science*, 2010, **329**, 309.
44 P. Barth, J. Schonbrun and D. Baker, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 15682.
45 R. Das and D. Baker, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 14664.
46 R. Das, J. Karanicolas and D. Baker, *Nat. Methods*, 2010, **7**, 291.
47 I. André, P. Bradley, C. Wang and D. Baker, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 17656.
48 R. P. Cheng, S. H. Gellman and W. F. DeGrado, *Chem. Rev.*, 2001, **101**, 3219.
49 S. F. Betz, D. P. Raleigh and W. F. Degrado, *Curr. Opin. Struct. Biol.*, 1993, **3**, 601.
50 V. H. Dalvi and P. J. Rossky, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 13603.
51 G. Cornilescu, E. B. Hadley, M. G. Woll, J. L. Markley, S. H. Gellman and C. C. Cornilescu, *Protein Sci.*, 2006, **16**, 14.
52 H. Zheng, K. Comeforo and J. M. Gao, *J. Am. Chem. Soc.*, 2009, **131**, 18.
53 H. Zheng and J. M. Gao, *Angew. Chem., Int. Ed.*, 2010, **49**, 8635.
54 B. C. Gorske and H. E. Blackwell, *J. Am. Chem. Soc.*, 2006, **128**, 14378.
55 M. G. Woll, E. B. Hadley, S. Mecozzi and S. H. Gellman, *J. Am. Chem. Soc.*, 2006, **128**, 15932.
56 R. Li, R. Gorelik, V. Nanda, P. B. Law, J. D. Lear, W. F. DeGrado and J. S. Bennett, *J. Biol. Chem.*, 2004, **279**, 26666.
57 K. J. Lumb and P. S. Kim, *Biochemistry*, 1995, **34**, 8642.
58 K. H. Lee, H. Y. Lee, M. M. Slutsky, J. T. Anderson and E. N. Marsh, *Biochemistry*, 2004, **43**, 16277.
59 J. W. Bryson, J. R. Desjarlais, T. M. Handel and W. F. DeGrado, *Protein Sci.*, 1998, **7**, 1404.
60 S. A. Sievers, J. Karanicolas, H. W. Chang, A. Zhao, L. Jiang, O. Zirafi, J. T. Stevens, J. Münch, D. Baker and D. Eisenberg, *Nature*, 2011, **475**, 96.
61 P. D. Renfrew, E. J. Choi, R. Bonneau and B. Kuhlman, *PLoS One*, 2012, **7**, e32637.

Electronic Supplementary Material (ESI) for Chemical Science
This journal is © The Royal Society of Chemistry 2012

Supplementary Information for:

Remodeling a β-peptide bundle

Matthew A. Molski,*[a] Jessica L. Goodman,[b] Fang-Chieh Chou,[c] David Baker,[d] Rhiju Das,*[c] and Alanna Schepartz*[a,e]

[a]Department of Chemistry, Yale University, New Haven, CT 06520-8107, [b]Whitehead Institute for Biomedical Research, Cambridge, MA 02142, [c]Department of Biochemistry, Stanford University, Stanford, CA 94305, [d]Department of Biochemistry, University of Washington, Seattle, WA 98195, [e]Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, CT 06520-8107.

Methods

Tables

Figures

This journal is © The Royal Society of Chemistry 2012

General Methods. Fmoc-protected amino acids, *O*-Benzotriazole-*N,N,N'N'*-tetramethyl-uronium-hexafluoro-phosphate (HBTU), and Wang resin were purchased from Novabiochem (San Diego, CA). (7-Azabenzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (PyAOP) was purchased from Oakwood Products, Inc. (West Columbia, SC). 1-Hydroxy-7-azabenzotriazole (HOAt) was purchased from Chempep, Inc. (Miami, FL). Fmoc-L-pentafluorophenylalanine was purchased from Anaspec (Fremont, CA). Dimethylformamide (DMF), *N*-methyl-2-pyrrolidone (NMP), *N*-methyl morpholine (NMM), trifluoroacetic acid (TFA) and piperdine (Pip) were purchased from American Bioanalytical (Natick, MA). All other reagents were purchased from Sigma-Aldrich (St. Louis, MO). Mass spectra were acquired with an Applied Biosystems Voyager-DE Pro MALDI-TOF mass spectrometer (Foster City, CA). Reverse-phase HPLC was performed using a Varian Prostar HPLC and Vydac analytical (C8, 300 Å, 5 μm, 4.6 mm X 150 mm), semi-preparative (C8, 300 Å, 10 μm, 10 mm X 250 mm), or preparative (C8, 300 Å, 5 μm, 25 mm X 250 mm) columns, using water/acetonitrile gradients containing 0.1% TFA. Circular dichroism (CD) spectra were acquired with a Jasco J-810 Spectropolarimeter (Jasco, Tokyo, Japan) equipped with a Peltier temperature control module. Analytical ultracentrifugation (AU) was performed using a Beckman XL-I instrument. ^{1}H NMR experiments were performed on Bruker 500 MHz and Varian Inova 800 MHz instruments. Sedimentation Equilibrium (SE)-AU, CD, fluorescence, and NMR experiments were performed in phosphate buffer (10 mM NaH_2PO_4, 200 mM NaCl, pH adjusted to 7.1 with NaOH).

β-Peptide Bundle Design using Rosetta. The Rosetta macromolecule modeling software[1-4] was used to identify variants of the β^3-peptide Acid-1Y bundles that would potentially assemble into well-folded, octameric β-peptide bundles. The only Rosetta extension required for modeling β^3-peptides was the specification of initial bond lengths and angles, which were derived from the Acid-1Y crystallographic model (for the backbone) and from natural proteins (for the side chains). All eight non-equivalent homoleucine core residues in the crystallographic model (positions 2, 5, 8 and 11 in two non-equivalent monomers) were remodeled. These side chains were allowed to sample β-peptide

This journal is © The Royal Society of Chemistry 2012

variants of the twenty proteinaceous amino acids with Rosetta's fixed backbone design algorithm *fixbb*.[5] Symmetry was enforced across all 32 side chains using Rosetta's *FixbbLinkingRotamerSimAnnealer*.[6] All calculations were carried out in the recently developed Rosetta 3 framework.[7] Example command-lines are given below; and all calculations are being made available in the β-peptide_modeling application contained in Rosetta 3.4.2 (freely available to academics at http://www.rosettacommons.org).

Table S1. Redesigned sequences from Rosetta fixed-backbone designs.

Score terms[a]	**Symmetry Enforced?**	**Sequence**
Old score	Yes	(LLLL LLLL) × 4
Current	Yes	(LLLL LLLL) × 4
MM	Yes	(LLLL LLLL) × 4
Old score + softrep	Yes	(MILI FLLL) × 4
Current + softrep	Yes	(MIAI FLFL) × 4
MM + softrep	Yes	(MILI FLLL) × 4
Old score	No	LIFI LLLL LLLL LLLL RLLI FITI LILI LMML
Current	No	LLLI LLLL LLLI LLLL RILI FLLL LILI LLLL
MM	No	LLLI LLLL LLLL LLLL LLLI LLML LILL LLAL
Old score + softrep	No	MILI FLVI LPFL LFLL MIFI FLLL MILL FLVI
Current + softrep	No	MILI FMLL MPAI FFFI MPLI FMML MILI FLVI
MM + softrep	No	MILI FLLL LLAI LLFL MILI FMML MILL FLVM

[a] 'Old score' refers to Rosetta all-atom energy function available in pre-release versions of Rosetta 3.0; 'Current' refers to Rosetta 3.4 all-atom energy function. 'softrep' refers to smoothed potential with lower penalties for clashes, used in early stages of modeling or design[8]. 'MM' refers to calculations with additional terms *mm_twist, mm_lj_intra_rep*, and *mm_lj_intra_atr* introduced in reference.[9]

This journal is © The Royal Society of Chemistry 2012

The Rosetta all-atom energy function for protein design was used for the optimization; it accounts for van der Waals interactions between side chains (attractive and repulsive forces), the hydrophobic effect and polar burial penalties (solvation) and hydrogen bonding (here, primarily fixed backbone-backbone interactions). To avoid incorrect biases from natural protein crystals, the backbone-independent version of Dunbrack's rotamer library[7] was used for the side chain search and a knowledge-based term favoring frequently observed rotamers in natural proteins was omitted. These approximations will necessarily lead to inaccuracies in energy estimates. Nevertheless, by modeling the non-polar interactions of the molecule's side chains at an all-atom level, the calculations were still useful in suggesting novel variants of well-packed cores. For example, as a basic consistency check, the original sequence (native core: L_2, L_5, L_8, L_{11} in both non-equivalent monomers) was returned as the optimal core sequence in a symmetric complete re-design of the Acid-1Y core with the standard Rosetta protein energy function.[10]

The Rosetta energy function has been improved[7] since our original calculations (which made use of an early version of Rosetta 3.0, internal revision number 17858). These improvements were also systematically explored in this symmetric sequence re-design test. First, the van der Waals potential has been smoothed to have continuous derivatives. Second, additional terms *mm_twist, mm_lj_intra_rep*, and *mm_lj_intra_atr* (called herein the MM potential) have been introduced and calibrated to better model non-canonical amino acids.[9] Redesigning the eight Acid-1Y core positions returned the all-homoleucine core. In each of these calculations, symmetry constraints required that the substitution of side-chain identity or rotameric conformation at one of the eight core positions be copied to three other positions related by symmetry. Lifting this symmetry constraint led to sequences that were not all-leucine cores (see Table S1). These sequences have not yet been explored due to the complications of preparing and mixing multiple different β-peptide components in exact stoichiometries.

This journal is © The Royal Society of Chemistry 2012

For calculations exploring core substitutions with backbone flexibility, side chains were repacked with the *fixbb* algorithm and the *softrep* energy function,[8] followed by continuous minimization of all backbone and side chain torsions and rigid-body degrees of freedom positioning the eight monomers. For the minimization, the standard Rosetta energy function rather than the *softrep* variant was employed to better discriminate precise hydrophobic packing; weak harmonic constraint potentials with spring constants of $K = 0.08$ kT/rad^2 were applied to all backbone torsional degrees of freedom (four per β-peptide residue) to prevent large conformational excursions (given as *dihedral* in Table S2). The total Rosetta energy of the homophenylalanine-containing variant Acid-1Y^{FF} was not as low as the starting conformations, mainly due to increases in *dihedral*; however the lower energies for summed Lennard-Jones terms *fa_atr* and *fa_rep* and solvation *fa_sol* suggested that the packing of non-polar residues was better than the starting homoleucine-containing bundle Acid-1Y. As above, these calculations have been repeated in the more recent Rosetta 4.3 framework, including use of terms *mm_twist, mm_lj_intra_rep*, and *mm_lj_intra_atr* to replace the *dihedral* constraints. In these repeated calculations, the summed non-polar packing terms were comparable in Acid-1Y and Acid-1Y^{FF} (within 2 Rosetta units; approximately 1 k_BT).

Table S2. Rosetta terms for non-polar atom-atom packing in the core sequence redesign of the Acid-1Y octameric β-peptide bundle. The van der Waals forces are accounted for in the attractive (atr) and repulsive (rep) terms. The hydrophobic effect is represented by the solvation term (sol). Energies are given in Rosetta units, which approximately corresponds to 1 k_BT.[11]

Variant	Core Sequence[a]	*fa_atr*	*fa_rep*	*fa_sol*	nonpolar sum[b]	*hbond*	*dihedral / MM*[c]	sum
Acid-1Y (old score)	LLLL LLLL	-409	29	181	-200	-89	0	-298
Acid-1Y^{FF} (old score)	LFFL LFFL	-436	25	187	-224	-102	134	-192
Acid-1Y (current)	LLLL LLLL	–440	53	185	-202	-90	3	-288
Acid-1Y^{FF} (current)	LFFL LFFL	-442	58	186	-198	-89	109	-178

This journal is © The Royal Society of Chemistry 2012

Acid-1Y (MM)	LLLL LLLL	-499	65	197	-237	-86	106	-218
Acid-1Y^{FF} (MM)	LFFL LFFL	-528	87	206	-235	-84	193	-126

[a] The sequences are positions 2,5,8, and 11 of the non-equivalent "outer" and "inner" monomers, respectively.
[b] Sum of *fa_atr*, *fa_rep*, and *fa_sol*.
[c] For the first four entries, the dihedral constraint score us shown. For the last two entires, we show the sum of the molecular mechanics scores (*mm_twist, mm_lj_intra_rep*, and *mm_lj_intra_atr*) since dihedral constaints are not used in those calculations.

After our work, a study from Shandler and colleagues[12] proposed a molecular-mechanics-based potential for β-peptide side chains. Interestingly, our re-designed core β^3-phenylalanine rotamer conformations turn out to be similar to conformations proposed in that study to be energetic minima in the 3_{14} helix context. These rotamers are also similar to those observed in the previous Zwit-1F crystal structure, but at non-core positions. The four (χ_1, χ_2) values in our new designed core phenylalanine side chains are (–53°, 112°), (–57°, 138°), (–59°, 139°), and (–76°, 176°), with other positions related by symmetry. The first 3 conformers fall within the dominant energy well in the molecular mechanics-based prediction (–80° to –40°, 90° to 135°), and the fourth conformer is close by.

Rosetta command lines. The Rosetta command lines used in this work are listed below for reference.

1. β^3-peptide peptide redesign

```
beta_peptide_modeling.<exe> -database <database> -force_field
beta_peptide -native acdy_LLLL_LLLL.pdb -algorithm redesign -ex1 -ex2 -
packing::pack_missing_sidechains false -packing::extrachi_cutoff 0 -
repack_res 2 5 8 11 14 17 20 23 -n_repeat 4 -repeat_size 24
```

2. Repacking and minimization for mutated sequence

   Repack:

```
beta_peptide_modeling.<exe> -database <database> -force_field
beta_peptide_soft_rep_design -native acdy_LFFL_LFFL.pdb -algorithm
```

This journal is © The Royal Society of Chemistry 2012

```
repack -ex1 -ex2 -packing::pack_missing_sidechains false -
packing::extrachi_cutoff 0 -repack_res 2 5 8 11 14 17 20 23 -n_repeat 4
-repeat_size 24
```

Minimize:

```
beta_peptide_modeling.<exe> -database <database> -force_field
beta_peptide -native acdy_LFFL_LFFL_repack.pdb -algorithm minimize -ex1
-ex2 -packing::pack_missing_sidechains false -packing::extrachi_cutoff 0
```

The two commands above are run sequentially to obtain the final model for the redesigned sequence.

All the commands listed above used the standard Rosetta protein scoring function with symmetry enforced. Extra flags/modifications that are used for old scoring (no smoothing), non-symmetric redesign, and MM potential are given below.

`-score::no_smooth_etables true` Use the old Rosetta attractive/repulsive score without smoothing.

`-no_symmetry true` Disable the symmetry-based redesigning.

`-force_field beta_peptide_mm / -force_field beta_peptide_soft_rep_mm` Use the molecular mechanics-based Rosetta scoring files. Replace the corresponding "`-force_field`" flags with the MM ones in the above command lines. In addition, the extra flags "`-apply_dihedral_cst false`" should be used to turn off dihedral constraint when the MM potential is used.

β-Peptide synthesis and purification. The Arndt-Eistert homologation method[13] was used to synthesize *N*-Fmoc-(*S*)-β^3-amino acids from the corresponding α-amino acid precursor as described previously.[14] β-peptides were synthesized in a CEM MARS microwave reactor on pre-loaded Fmoc-β^3-Asp(*t*Bu)-OH Wang resin[15] (25 μmol scale) in a glass reaction vessel (Ace Glass, Vineland, NJ) that was pretreated with SigmaCote (Sigma Aldrich). The resin (46.6 mg of 0.537 mmol/g) was allowed to swell for 1 h in 8 mL DMF on a rotary shaker. After swelling, the resin was washed 6 times

This journal is © The Royal Society of Chemistry 2012

with 4 mL DMF, including 3 times of swirling the reaction vessel (RV) to break up any resin clumps. Subsequently, the Fmoc protecting group was removed using 4 mL of 20% piperdine/DMF and microwave irradiation (microwave deprotection program: 50% power at 400 W maximum; ramp to 70 °C for 2 min.; hold at 70 °C for 4 min.; cool for 5 min) with magnetic stirring. Following deprotection, the resin was washed using the same wash procedure as above (6 x 4 mL DMF, including 3 x with vigorous shaking of the RV). Coupling solution was prepared in a separate vial. This cocktail consisted of the appropriate Fmoc-β^3-amino acid (75 µmol), PyAOP (75 µmol), and HOAT (75 µmol) resuspended in 2 mL DMF. To this was added 35 µL diisopropylethylamine (DIPEA), and the 2 mL solution was added to the resin in the RV. The reaction mixture was then exposed to microwave irradiation with magnetic stirring (coupling program: 50% power at 400 W maximum; ramp to 60 °C for 2 min.; hold at 60 °C for 6 min.; cool for 5 min). After coupling, the resin in the RV was extensively washed with DMF. To prepare for another coupling, the resin was deprotected as previously described. The cycle of deprotection-wash-coupling-wash was performed until the peptide was synthesized in full. The solution used for deprotection was always prepared immediately before use.

At the end of the synthesis, the last Fmoc protecting group was removed using the deprotection protocol. The resin was washed alternately with DMF and methanol for a total of 16 washes, followed by an additional 8 consecutive methanol washes, and dried at least 30 min under N_2. The dry resin was treated with three 4 mL portions of cleavage solution (TFA with 1% v/v each of H_2O, phenol, and triisopropyl silane) for 1 hr, 1 hr, and 30 min. The cleavage solution was collected in a 50 mL round bottom flask and was rotovaped to remove the TFA. The remaining 1 mL of solution was resuspended in 50:50 CH_3CN/H_2O for HPLC purification.

The success of each synthesis was assessed by both HPLC and MALDI-TOF analysis of the crude reaction mixture. β-peptides were then purified to homogeneity by reverse-phase HPLC. The identities and purities of the purified β-peptides were assessed by analytical HPLC and mass spectrometry. MALDI mass spectra were obtained using peptide samples in α-cyano-4-hydroxycinnaminic acid (CHCA) matrix. The masses found were: Acid-1Y^{FF} (m/z observed,

This journal is © The Royal Society of Chemistry 2012

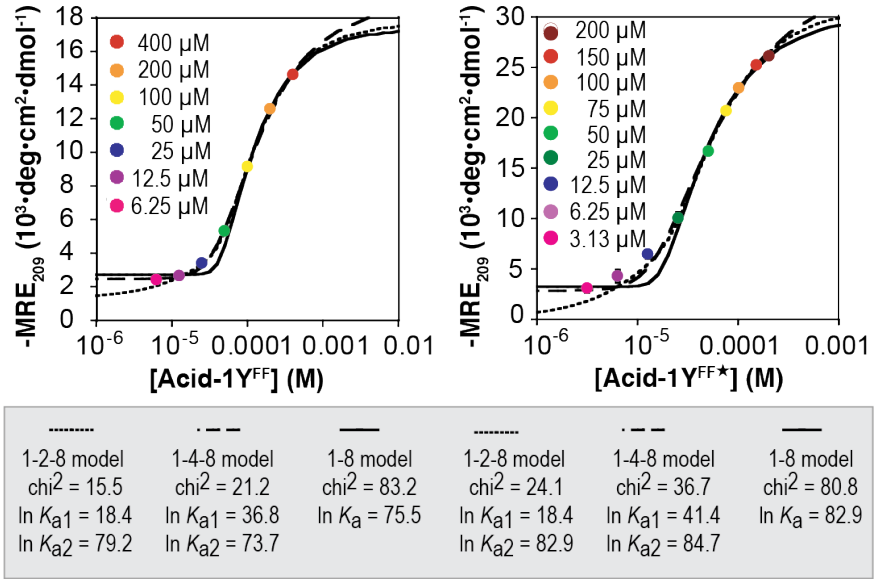

Fig. S1. Self-association of the Acid-1Y^{FF} and Acid-1Y^{FF*} bundles showing data fits. Circular dichroism spectra of Acid-1Y^{FF} and Acid-1Y^{FF*} as a function of concentration. Plots of MRE_{209} as a function of [β-peptide] were fit to monomer-dimer-octamer (1-2-8), monomer-tetramer-octamer (1-4-8) or monomer-octamer (1-8) equilibria; error analysis and calculated values are shown.

calculated): 1748, 1748; and Acid-1Y^{FF}* (m/z observed, calculated): 1840, 1839. Following purification, β-peptides were lyophilized and reconstituted in buffer (described above) for characterization. Concentrations were determined by UV absorbance at 280 nm, where 1490 M^{-1} cm^{-1} was used as the molar extinction coefficient for tyrosine. (http://encorbio.com/protocols/Prot-MW-Abs.htm)

Circular Dichroism (CD). The wavelength dependent CD spectra of Acid-1Y^{FF} and Acid-1Y^{FF}* at concentrations between 3.13 and 400 μM were acquired at 25 °C using a Jasco J-810 Spectropolarimeter (Jasco, Tokyo, Japan). Relevant machine settings include: continuous scan mode, 1.0 nm data pitch, 50 nm/min scanning speed, 4 sec response, 0.5 nm band width and 3 accumulations. The concentration dependence of the molar residue ellipticity (MRE) minimum (at 209 nm for Acid-1Y^{FF} and 212 nm for Acid-1Y^{FF}*) was determined by fitting of the total peptide monomer concentration ($[Peptide]_{Total}$) as a function of the experimentally determined MRE using the following equation in Kaleidagraph, Version 3.6 (Synergy Software; Reading, PA).[16] MRE_{Mon} represents the MRE of the β-peptide monomer; MRE_{nmer} represents the MRE of the β-peptide oligomer, where n = the oligomerization state.

This journal is © The Royal Society of Chemistry 2012

$$[Peptide]_{Total} = \left\{ \frac{(MRE_{Exp} - MRE_{Mon})(1/K_a)}{n(MRE_{Nmer} - MRE_{Mon})\left[1 - \left(\frac{MRE_{Exp} - MRE_{Mon}}{MRE_{Nmer} - MRE_{Mon}}\right)\right]^n} \right\}^{1/(n-1)}$$

Using this equation and the fixed parameters MRE_{Mon} = 2991 ± 300 deg•cm^2•dmol^{-1} and MRE_{Nmer} = 17,052 ± 150 deg•cm^2•dmol^{-1}, we calculate a ln K_a = 75.5 ± 0.5 for the Acid-1Y^{FF} bundle. For the Acid-1Y^{FF*} bundle, the ln K_a calculated was 73.0 ± 0.5 (MRE_{Mon} = 3000 deg•cm^2•dmol^{-1}; MRE_{Nmer} = 30,014 ± 329 deg•cm^2•dmol^{-1}).

Fits of peptide circular dichroism data were also carried out to a more general equilibrium involving multiple species e.g., a monomer, tetramer, and an octamer. If the fraction of peptide within each n-mer species is f_n, these values satisfy the equations:

$$f_n = n f_1 \left(p f_1 / \kappa_n \right)^{n-1} \qquad (1)$$

where p is the peptide concentration and the constants κ_n are related to standard association constants by:

$$\ln K_a = -n \ln(\kappa_n / 1\ \mathrm{M}) \qquad (2)$$

The fractions in each species sum to 1:

$$f_1 = 1 - \sum_{n>1} f_n \qquad (3)$$

For given κ_n, the equations were solved numerically in MATLAB by iteratively computing updates to the f_n (for $n > 1$) with eq. 1 and updating f_1 with eq. 3. To optimize over κ_n, grid searches were carried out over values from 10^{-6} M to 10 M (in $10^{0.25}$ M increments). Errors were estimated by bootstrapping the data, i.e., by repeating the entire fitting analysis with replicates generated by random selection of available data with replacement.[17] Association constants calculated were as follows: Acid-1Y^{FF} 1mer-2mer-8mer model, ln K_1 = 18.4 ± 4.1 and ln K_2 = 79.2 ± 9.8; Acid-1Y^{FF} 1mer-4mer-8mer model, ln K_1 = 36.8 ± 3.2 and ln K_2 = 73.7 ± 11.2; Acid-1Y^{FF*} 1mer-2mer-8mer model, ln K_1 = 18.4 ± 4.0 and ln K_2 = 82.9 ± 2.3; Acid-1Y^{FF*} 1mer-4mer-8mer model, ln K_1 = 41.4 ± 4.6 and ln K_2 = 84.7 ± 3.0.

This journal is © The Royal Society of Chemistry 2012

Table S3. Fits of isothermal circular dichroism data to monomer-octamer (1-8), monomer-dimer-octamer (1-2-8) and monomer-tetramer-octamer (1-4-8) models. *P* values below 1 x 10^{-3} for the simplest monomer-octamer (1-8) model strongly disfavor those fits, while *P* values above 0.1 for monomer-dimer-octamer (1-2-8) and monomer-tetramer-octamer (1-4-8) models show that either model is consistent with the data. **#** = number of parameters

Construct	Fit	ln K_1	ln K_2	c^2	*N*	#	*P*
Acid-1Y^{FF}	1-8	–	75.5	83.2	21	3	2.3 x 10^{-10}
Acid-1Y^{FF}	1-2-8	18.4	79.2	15.5	21	5	0.48
Acid-1Y^{FF}	1-4-8	36.8	73.7	21.2	21	5	0.17
Acid-1Y^{FF}*	1-8	–	82.9	80.8	36	3	6.9 x 10^{-6}
Acid-1Y^{FF}*	1-2-8	18.4	82.9	24.1	36	5	0.81

Table S4. Fits of sedimentation equilibrium analytical ultracentrifugation data for Acid-iY^{FF} and Acid-1Y^{FF*} to monomer-*n*-mer models

Acid-1Y^{FF}			Acid-1Y^{FF*}		
n	ln *K*	RMSD	n	ln *K*	RMSD
2	15.496	0.04950	2	15.872	0.04926
3	25.883	0.03906	3	27.349	0.03880
4	51.271	0.02872	4	40.714	0.02894
5	110.564	0.01973	5	53.217	0.02039
6	54.738	0.01294	6	54.896	0.01320

This journal is © The Royal Society of Chemistry 2012

7	58.615	0.00821	7	62.379	0.00766
8	64.763	0.00778	8	70.662	0.000702
9	69.020	0.01026	9	79.094	0.01122
10	74.962	0.01246	10	84.711	0.01650
Float n	62.091	0.00731	Float n	67.223	0.00646
	n = 7.57			n = 7.59	

Temperature-dependent CD spectra were obtained at 209 nm and 212 nm (for Acid-1Y^{FF} and Acid-1Y^{FF*}, respectively) between 5 and 95 ^{o}C, using the Peltier temperature control module provided with the instrument. Relevant settings include: 1 ^{o}C data pitch, 5 sec delay time, 1 ^{o}C/min temperature slope, 4 sec response time, and 1 nm band width. In the case of both Acid-1Y^{FF} and Acid-1Y^{FF*}, the first derivative of the MRE was calculated at each of four β-peptide (monomer) concentrations. The T_m values reported represent the maximum of the $\delta MRE_{min}/\delta T$ plots. Fig. 2b,d in the main text show plots of $\delta MRE_{min}/\delta T$ as a function of temperature for Acid-1Y^{FF} and Acid-1Y^{FF*}.

This journal is © The Royal Society of Chemistry 2012

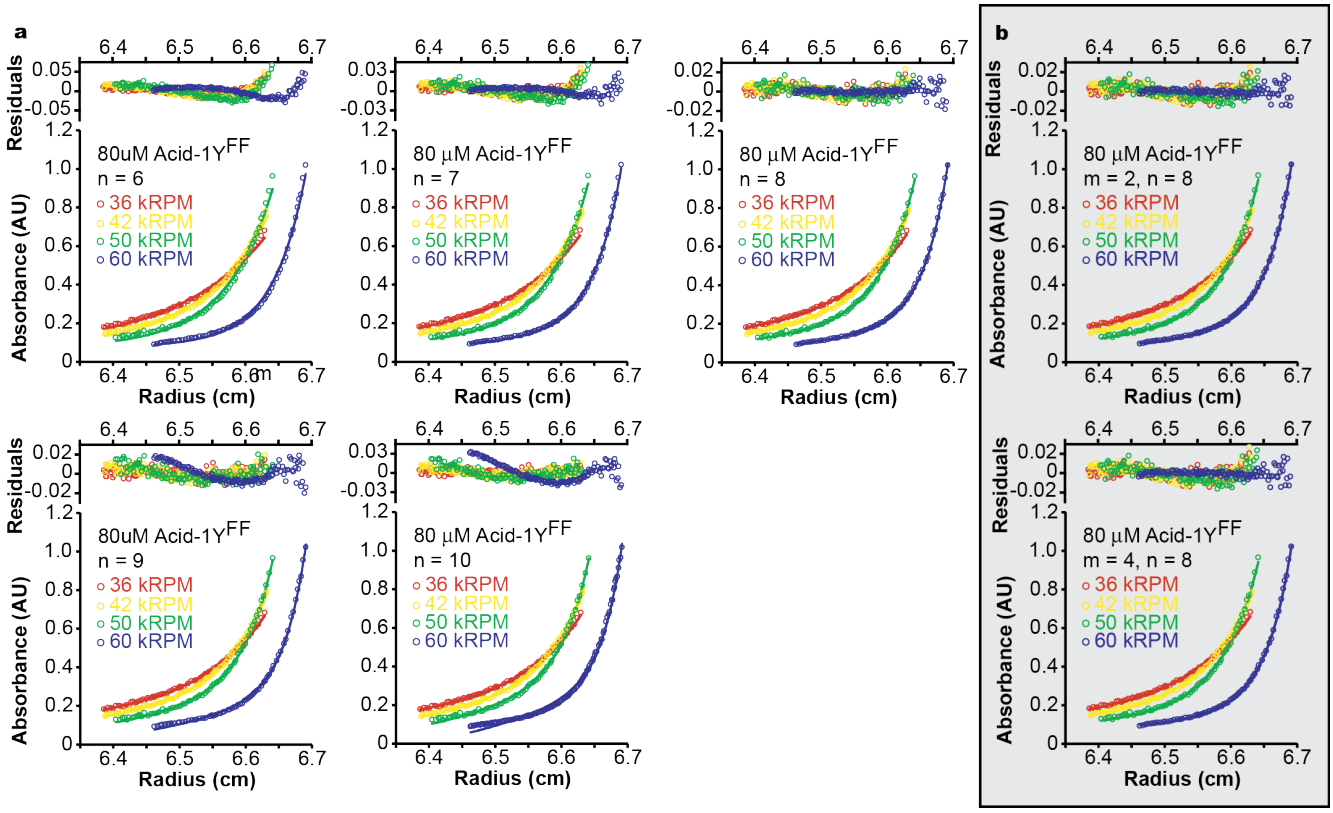

Fig. S2. Sedimentation equilibrium analytical ultracentrifugation analysis of bundle stoichiometry and association constant: Acid-1Y^{FF}. Plots illustrating monomer-*n*-mer equilibrium fits to experimental data at 80 μM concentrations (monomer) of Acid-1Y^{FF}. (**A**) Representative data fits and residuals of Acid-1Y^{FF} monomer-*n*-mer and (**B**) monomer-*m*-mer-*n*-mer equilibrium models. See also Table S4.

Sedimentation Equilibrium Analytical Ultracentrifugation. HPLC-purified, lyophilized samples of Acid-1Y^{FF} and Acid-1Y^{FF}⋆ were resuspended in phosphate buffer (10 mM NaH_2PO_4, 200 mM NaCl, pH adjusted to 7.1 with NaOH) at concentrations of 20, 80 and 200 μM, and centrifuged at 25 °C to equilibrium at four speeds (36,000, 42,000, 50,000 and 60,000 rpm). Centrifugation was performed using an AN 60-Ti 4-hole rotor equipped with six-channel, carbon-epoxy composite centerpieces (Beckman). Absorbance was monitored at both 230 and 280 nm. Data were collected with a step size of 0.001 cm with scans occurring at 1 h intervals. Samples were determined to have reached equilibrium when no significant changes in radial concentration were observed in three successive scans, as determined using the Match sub-routine within the Heteroanalysis software suite (available

from the National Analytical Ultracentrifugation Facility website,

http://vm.uconn.edu~wwwbiotc/uaf.html).

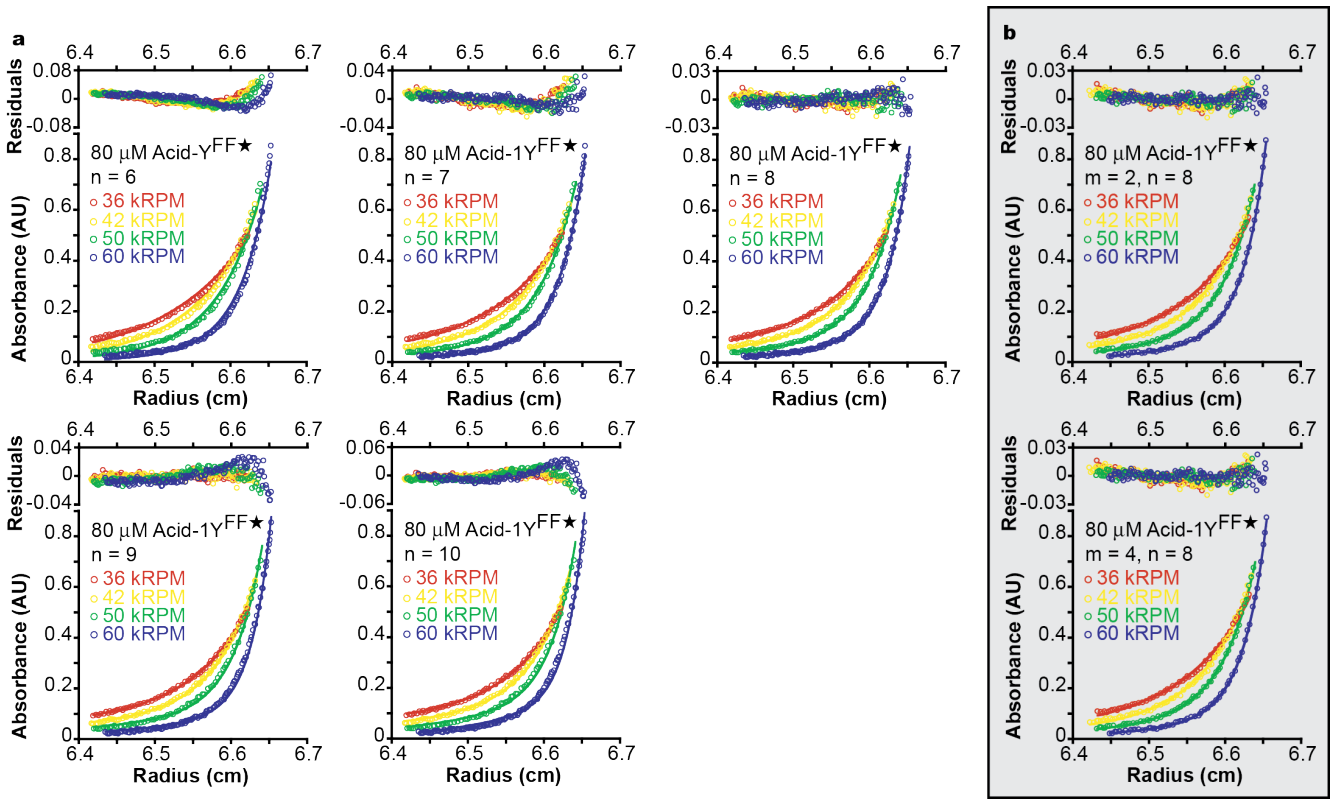

Fig. S3. Sedimentation equilibrium analytical ultracentrifugation analysis of bundle stoichiometry and association constant: Acid-1Y^{FF*}. Plots illustrating monomer-*n*-mer equilibrium fits to experimental data at 80 μM concentrations (monomer) of Acid-1Y^{FF*}. **(A)** Representative data fits and residuals of Acid-1Y^{FF*} monomer-*n*-mer and (**B**) monomer-*m*-mer-*n*-mer equilibrium models. See also Table S4.

The partial specific volume of each β-peptide was calculated on the basis of functional group composition according to Durchschlag and Zipper.[18] The data were initially fit to a monomer-*n*-mer equilibrium model using Heteroanalysis software, assuming fast monomer-*n*-mer equilibration compared to diffusion within the ultracentrifuge profile. Fixed parameters for Acid-1Y^{FF}: monomer MW = 1748 Da; v-bar = 0.83070 cm^3/g; d = 1.00674 g/mL; ε_{280} = 2906 (accounting for the 1.2 cm path length) M^{-1}•cm^{-1}; and n = 8. Fitted parameters for Acid-1Y^{FF}: ln K_a = 64.8 ± 0.1; root mean square deviation (RMSD) = 0.00778; and baseline deviation < 0.02. Fixed parameters for Acid-1Y^{FF*}: monomer MW = 1839 Da; v-bar = 0.77591 cm^3/g; d = 1.00674 g/mL; ε_{280} = 2906 (accounting for the

This journal is © The Royal Society of Chemistry 2012

1.2 cm path length) $M^{-1}\bullet cm^{-1}$; and $n = 8$. Fitted parameters for Acid-1Y^{FF}★: ln $K_a = 70.7 \pm 0.1$; RMSD = 0.00702; and baseline deviation < 0.03. For both Acid-1Y^{FF} and Acid-1Y^{FF}★, fitting to all other values of *n* between 2 and 10 resulted in larger RMSD values, and greater and more systematic residuals (Fig. S2a and S3a). When *n* was allowed to float, it fit best to 7.57 for Acid-1Y^{FF} and 7.59 for Acid-1Y^{FF}★.

The data were also fit to a monomer-*m*-mer-*n*-mer equilibrium model using Heteroanalysis (Fig. S2b and S3b). Fixed parameters for Acid-1Y^{FF} were the same as above, with the additional parameter *m* set to either 2 or 4 (Fig. S2b). For $m = 2$: ln $K_1 = 8.143 \pm 0.4$; ln $K_2 = 68.0663 \pm 0.7$; RMSD = 0.00735; and baseline deviation < 0.02. For $m = 4$: ln $K_1 = 27.600 \pm 0.2$; ln $K_2 = 68.150 \pm 0.3$; RMSD = 0.00714; and baseline deviation < 0.02. Fixed parameters for Acid-1Y^{FF}★ were also the same as above, and *m* was set to either 2 or 4 (Fig. S3b). For $m = 2$: ln $K_1 = 11.560 \pm 0.4$; ln $K_2 = 79.707 \pm 1.3$; RMSD = 0.00640; and baseline deviation < 0.02. For $m = 4$: ln $K_1 = 29.742 \pm 0.2$; ln $K_2 = 72.644 \pm 0.2$; RMSD = 0.00634; and baseline deviation < 0.02. When *m* was allowed to float, it fit best to 4.14 for Acid-1Y^{FF} and 4.02 for Acid-1Y^{FF}★. The F-test was used to determine whether the lower RMSD values observed for the monomer-tetramer-octamer fits were significant relative to the monomer-octamer fits, or whether they represented an artifact caused by sampling more parameters. The F-test was conducted using the GraphPad software (available online at http://www.graphpad.com/quickcalcs/AIC1.cfm). The following equation was used to calculate the sum-of-squares (SS), where n is the number of data points: $SS = (RMSD^2)(n)$. For both Acid-1Y^{FF} and for Acid-1Y^{FF}★, the probability that the monomer-tetramer-octamer equilibrium was the correct model was greater than 1,000,000:1, corresponding to $P < 10^{-6}$.

1-Anilino-8-naphthalenesulfonate (ANS) Binding. Stock solutions of 800 μM Acid-1Y^{FF}, 1000 μM Acid1Y^{FF}★ and 20 μM ANS were prepared in phosphate buffer (10 mM phosphate, 200 mM NaCl, pH 7.1). Binding reactions were performed by mixing 80 μL of the ANS stock solution with an appropriate volume of the Acid-1Y^{FF} or Acid1Y^{FF}★ stocks and diluting with phosphate buffer to a final volume of

160 μL to produce 0, 12.5, 25, 50, 100, 150, and 200 μM solutions of Acid-1Y^{FF} and 0, 1.5625, 3.125, 6.25, 12.5, 25, 50, 100, 150, and 200 μM solutions of Acid1Y^{FF*}. The fluorescence intensity of each solution (counts/s) was measured at 25 °C using a Photon Technology International (Lawrenceville, NJ) Quantamaster C-60 spectrofluorimeter and a 1 cm path length Hellma (Mullheim, Germany) cuvette. Each sample was excited at 350 nm (4 nm slit width) and the emission measured at 1 nm intervals between 400 and 600 nm (Fig. S4).

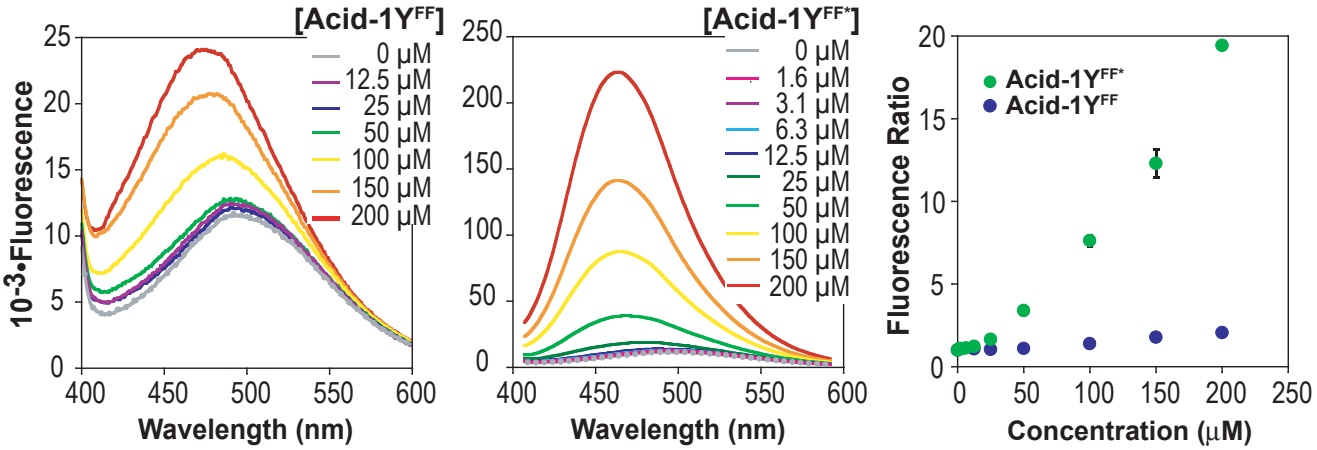

Fig. S4. Fluorescence emission of 1-anilino-8-naphthalenesulfonate (ANS) in the presence of Acid-1Y^{FF} and Acid1Y^{FF*}. Fluorescence is in counts•s^{-1}. The fluorescence ratio is defined as the maximal ANS fluorescence at each β-peptide concentration divided by the maximal ANS fluorescence in the absence of added β-peptide. The final [ANS] = 20 μM.

NMR analysis of hydrogen/deuterium exchange. The rate of exchange of amide N-H protons within the Acid-1Y^{FF} and Acid1Y^{FF*} bundles were determined at 25 °C by NMR using the WATERGATE solvent suppression pulse sequence ZZPWG (see www.bruker.com for pulse sequence information) and procedures described previously.[5, 19] In brief, we first acquired ^{1}H spectra of Acid-1Y^{FF} and Acid1Y^{FF*} at concentrations of 750 and 549 μM, respectively, in a 9:1 mixture of H_2O and D_2O, and the location of the H_2O resonance was identified.[19] The transmitter offset (^{01}P) was set to this frequency for ZZPWG acquisitions. In order to acquire deuterium exchange data, the Acid-1Y^{FF} sample was dissolved in phosphate buffer (10 mM NaH_2PO_4, 200 mM NaCl, pH adjusted to 7.1 with NaOH) spiked with 1 mM TMSP and lyophilized. The lyophilized sample was resuspended in D_2O at

a concentration of 750 μM just prior to insertion into the instrument. Any peptide that could not be resolubilized was removed from the solution by centrifugation at 5000 rpm for 1 minute before addition to the NMR tube. Prior to loading the Acid-1Y^{FF} sample, the spectrometer was locked and shimmed on a sample of Zwit-1F in 9:1 H_2O/D_2O in phosphate buffer. Due to the lack of amide protons, the data for Acid-1Y^{FF} was not further processed. For Acid1Y^{FF*}, a standard spectrum was acquired using a 500 μM sample of Acid1Y^{FF*} in the aqueous buffer. This standard was also used to lock and shim the instrument prior to running the H/D NMR acquisition. A separate sample was prepared in D_2O as described above, resulting in a concentration of 549 μM. Spectra were measured every 10 minutes, for 120 total spectra, with 16 scans per spectrum. Mestrelab Mnova (Mestrelab Research, Escondido, CA) was used for data processing, with 30 Hz line broadening and Bernstein polynomial baseline correction. Integrated areas were extracted from the same regions in each spectrum of the time series. We measured the pD of the sample after exchange was complete using a pH meter (Thermo Orion Model 410, Beverly, MA). The sample was found to be at pH 5.77, corresponding to a pD of 6.17.[20] Protection factors were calculated as previously described.[19]

References

1. B. Kuhlman, G. Dantas, G. C. Ireton, G. Varani, B. L. Stoddard and D. Baker, *Science*, 2003, **302**, 1364-1368.
2. J. J. Gray, S. Moughon, C. Wang, O. Schueler-Furman, B. Kuhlman, C. A. Rohl and D. Baker, *J. Mol. Biol.*, 2003, **331**, 281-299.
3. C. A. Rohl, C. E. M. Strauss, K. M. S. Misura and D. Baker, in *Numerical Computer Methods, Pt D*, Academic Press Inc, San Diego, 2004, vol. 383, pp. 66-+.
4. R. Das and D. Baker, *Annu Rev Biochem*, 2008, **77**, 363-382.
5. J. L. Goodman, E. J. Petersson, D. S. Daniels, J. X. Qiu and A. Schepartz, *J Am Chem Soc*, 2007, **129**, 14746-14751.
6. P. S. Huang, Y. E. Ban, F. Richter, I. Andre, R. Vernon, W. R. Schief and D. Baker, *PLoS One*, 2011, **6**, e24109.

7. A. Leaver-Fay, M. Tyka, S. M. Lewis, O. F. Lange, J. Thompson, R. Jacak, K. W. Kaufmann, P. D. Renfrew, C. A. Smith, W. Sheffler, I. W. Davis, S. Cooper, A. Treuille, D. J. Mandell, F. Richter, Y.-E. Ban, S. J. Fleishman, J. E. Corn, D. E. Kim, S. Lyskov, M. Berrondo, S. Mentzer, Z. Popovic, J. J. Havranek, J. Karanicolas, R. Das, J. Meiler, T. Kortemme, J. J. Gray, B. Kuhlman, D. Baker and P. Bradley, *Meth. Enzymol.*, 2011, **487**, 545-574.

,8. G. Dantas, C. Corrent, S. L. Reichow, J. J. Havranek, Z. M. Eletr, N. G. Isern, B. Kuhlman, G. Varani, E. A. Merritt and D. Baker, *J Mol Biol*, 2007, **366**, 1209-1221.

9. P. D. Renfrew, E. J. Choi, R. Bonneau and B. Kuhlman, *PLoS One*, 2012, **7**, e32637.

10. G. Dantas, C. Corrent, S. L. Reichow, J. J. Havranek, Z. M. Eletr, N. G. Isern, B. Kuhlman, G. Varani, E. A. Merritt and D. Baker, *J. Mol. Biol.*, 2007, **366**, 1209-1221.

11. T. Kortemme, D. E. Kim and D. Baker, *Sci. STKE*, 2004, **pl2**.

12. S. J. Shandler, M. V. Shapovalov, J. R. L. Dunbrack and W. F. DeGrado, *J. Am. Chem. Soc.*, 2010, **132**, 7312-7320.

13. D. Seebach, M. Overhand, F. N. M. Kuhnle, B. Martinoni, L. Oberer, U. Hommel and H. Widmer, *Helv Chim Acta*, 1996, **79**, 913-941.

14. J. A. Kritzer, J. Tirado-Rives, S. A. Hart, J. D. Lear, W. L. Jorgensen and A. Schepartz, *J Am Chem Soc*, 2005, **127**, 167-178.

15. S. A. Hart, A. B. Bahadoor, E. E. Matthews, X. J. Qiu and A. Schepartz, *J Am Chem Soc*, 2003, **125**, 4022-4023.

16. W. F. Degrado and J. D. Lear, *J Am Chem Soc*, 1985, **107**, 7684-7689.

17. V. Nanda and R. L. Koder, *Nat Chem*, 2010, **2**, 15-24.

18. H. Durchschlag and P. Zipper, in *Ultracentrifugation*, ed. M. Lechner, Springer Berlin / Heidelberg, 1994, vol. 94, pp. 20-39.

19. E. J. Petersson, C. J. Craig, D. S. Daniels, J. X. Qiu and A. Schepartz, *J Am Chem Soc*, 2007, **129**, 5344-5345.

20. P. K. Glasoe and F. A. Long, *J Phys Chem-Us*, 1960, **64**, 188-190.