Aromatic Stacking

DOI: 10.1002/ange.201105857

Stacked Fluoroaromatics as Supramolecular Synthons for Programming Protein Dimerization Specificity**

*Christopher J. Pace, Hong Zheng, Ruben Mylvaganam, Diane Kim, and Jianmin Gao**

Supramolecular synthons that direct molecular associations are highly desirable for the design of self-assembled materials and foldamers that interact with biological systems.[1] With a few exceptions, much work in the areas of protein and peptide design utilizes polar groups as supramolecular synthons that afford structural specificity through hydrogen bonding and salt-bridge formation.[1e] Although prevalent in protein structures,[2] aromatic interactions have been rarely utilized in protein design,[3] presumably owing to the incomplete understanding of their interaction energetics. The aromatic residues are primarily considered to be hydrophobic, yet they are known to engage in electrostatic interactions.[4] One well-known example is the cation–π interaction, which is employed by numerous signaling proteins, such as acetylcholine receptors and chromodomains that recognize methylated histones.[5] Two stacked aromatic rings may also interact with each other through electrostatic mechanisms, often referred to as π–π interactions or quadrupole interactions.[6] Recent work from our group describes that a stacked phenyl and perfluorophenyl pair dictates the dimerization specificity of a helix-bundle protein, thereby showcasing the potential of stacked aromatics as supramolecular synthons in aqueous media.[7] Herein, we systematically examine the aromatic stacking energetics by introducing various stacked aromatic pairs into the model protein α_2D.[8] The results reveal a surprisingly large contribution of dipole–dipole and dipole–induced-dipole interactions to aromatic stacking. We further demonstrate that the stacked aromatic pairs effectively afford self-sorting of highly analogous peptide monomers to give specific dimeric species.

α_2D is a de novo designed protein reported by DeGrado and co-workers.[8] This 35-residue polypeptide folds into a dimeric helix bundle and displays a highly cooperative and reversible folding behavior, which makes it easy to characterize the thermodynamics of its folding and dimerization.[9] A prominent feature of α_2D is the aromatic core, which consists of two phenylalanine pairs stacking in the face-to-face geometry (Figure 1a). This unique aromatic core presents an ideal system for investigating the energetics of aromatic stacking interactions. We primarily used the fluorinated analogues of phenylalanine in this study because of the minimal steric perturbation caused by the hydrogen-to-fluorine substitutions. Although sterically conservative, fluorination can introduce rather large perturbations to the

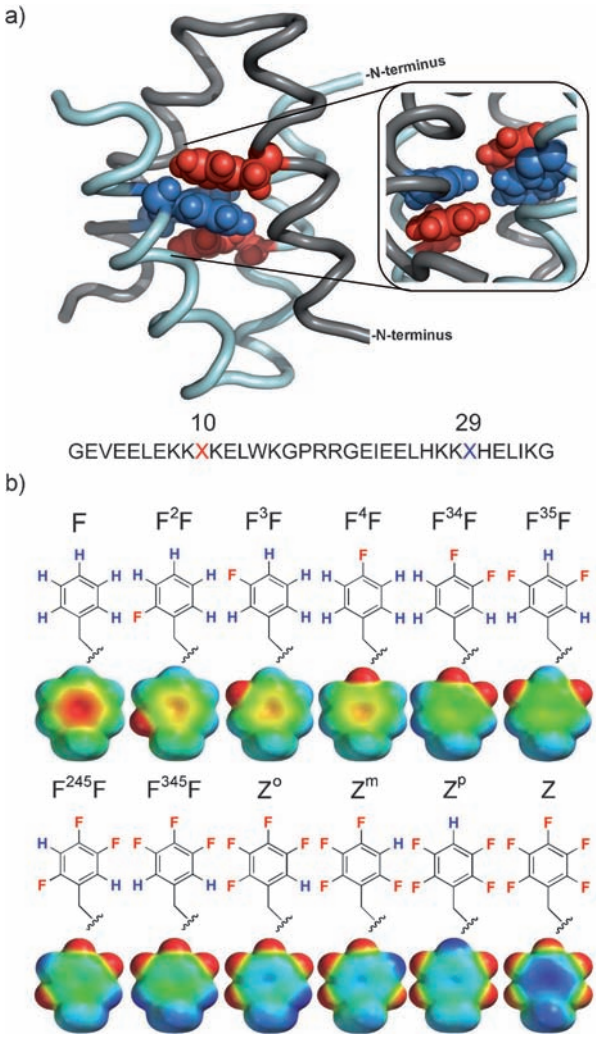

Figure 1. a) Schematic representation of the α_2D dimer (PDB 1PQ6). The monomers are colored in gray and light cyan. The two face–face stacking pairs in the core of the dimer are highlighted to the right, with F10 colored in red and F29 in blue (F = Phe = phenylalanine). b) Side chains of fluorinated phenylalanine derivatives incorporated into α_2D single and/or double mutants. Electrostatic potential maps (blue = positive, red = negative) were generated for the toluene derivatives with Spartan.

[*] C. J. Pace, H. Zheng, R. Mylvaganam, D. Kim, Prof. J. Gao
Department of Chemistry, Merkert Chemistry Center
Boston College
2609 Beacon street, Chestnut Hill, MA 02467 (USA)
E-mail: jianmin.gao@bc.edu

[**] We thank the Smith Family Foundation and Boston College for financial support. We also acknowledge the NSF (grant no DBI-0619576) for providing the financial support for the BC Mass Spectrometry Center.

Supporting information (including additional data and experimental details) for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201105857.

electronic properties of an aromatic ring.[4b,10] We have introduced phenylalanine analogues with varied numbers of fluorine atoms and substitution patterns into α_2D (Figure 1b). The resulting mutants were analyzed for their thermodynamic stability.

While a collection of the fluorinated phenylalanine derivatives is available from commercial sources, we synthesized the three tetrafluorinated phenylalanine derivatives through an improved protocol. The synthesis described herein involves alkylation of the Schöllkopf chiral auxiliary (see the Supporting Information for details), which is much cheaper than the Seebach's auxiliary used in our previous report.[11] The synthesis also avoids heating to reflux in concentrated NaOH for deprotection and consequently affords higher overall yields of the target amino acids. Through solid-phase peptide synthesis we prepared a series of α_2D double mutants with residues Phe10 and Phe29 replaced with the fluorinated analogues (Figure 1b). For ease of discussion, we named each peptide according to the identity of their residues 10 and 29. For example, the wild-type α_2D is named as (F,F). All peptides are purified by reversed-phase HPLC and subsequent gel filtration. The purity (higher than 95%) and identity were confirmed by analytical LC–MS (Table S1 in the Supporting Information).

All α_2D variants fold into homodimeric complexes as shown by size-exclusion chromatography. The thermodynamic parameters of dimerization are obtained through van't Hoff analysis of thermal melting curves of α_2D at varied concentrations (Table 1 and the Supporting Information).

Table 1: Summary of thermodynamic parameters for α_2D double mutants.

Peptide	T_m [°C][a]	ΔH_f [kcal mol^{-1}][b]	ΔG_f [kcal mol^{-1}][c]	$\Delta\Delta G_f$ [kcal mol^{-1}] [d]
F, F	28.9	-28.4 ± 2.6	-5.9 ± 0.1	0.0
F^2F, F^2F	39.7	-25.0 ± 3.9	-6.9 ± 0.1	1.0
F^3F, F^3F	46.9	-51.4 ± 5.9	-8.3 ± 0.2	2.4
F^4F, F^4F	39.8	-43.6 ± 5.2	-7.1 ± 0.1	1.2
F^{34}F, F^{34}F	57.8	-60.7 ± 7.4	-10.5 ± 0.5	4.6
F^{35}F, F^{35}F	55.2	-48.2 ± 5.8	-9.3 ± 0.3	3.4
F^{245}F, F^{245}F	63.7	-57.4 ± 2.4	-11.2 ± 0.2	5.3
F^{345}F, F^{345}F	72.1	-59.7 ± 5.0	-12.8 ± 0.5	6.9
Z^p, Z^p	52.2	-41.1 ± 2.3	-8.7 ± 0.1	2.8
Z^m, Z^m	65.0	-60.8 ± 3.3	-11.7 ± 0.3	5.8
Z^o, Z^o	75.9	-51.0 ± 3.5	-12.4 ± 0.4	6.5
Z, Z	78.3	-50.1 ± 3.1	-12.6 ± 0.4	6.7

[a] Peptide concentration 20 μM, all measurements within ±1 °C.
[b] Calculated by plotting $1/T_m$ vs. ln$K$ for different peptide concentrations and using the van't Hoff equation $\ln K = -\Delta H_f/RT + \Delta S/R$.
[c] Determined at 37 °C, 20 μM peptide concentration.
[d] $\Delta\Delta G_f = \Delta G_{f(WT)} - \Delta G_{f(Mutant)}$.

The peptide mutants display a large variation in their thermodynamic stabilities: the melting temperatures (T_m) of the α_2D homodimers vary from 29 to 78 °C and the folding free energies (ΔG_f) range from −5.9 to −12.8 kcal mol^{-1}. All fluorinated homodimers display improved thermal stabilities in comparison to the wild type. Impressively, the mutant incorporating two perfluorinated phenylalanines (Z,Z) gives a melting temperature of 78 °C, nearly 50 degrees higher than the wild type. This increase is perhaps expected given the fact that fluorocarbon compounds are generally more hydrophobic than the corresponding hydrocarbon compounds. To further analyze this panel of data, we plotted the folding free energies of the double mutants against the calculated Log$P$ values of the aromatic side chains (Figure 2a). Interestingly, a positive but poor correlation was observed with R^2 of 0.67. A similarly poor correlation was obtained between the ΔG_f values and the calculated surface areas of these fluoroaromatic side chains (Figure 2b). Collectively, these data indicate that factors other than hydrophobicity must contribute significantly to the stability of the α_2D homodimers.

A closer look at the data in Table 1 reveals surprisingly large stability differences among the α_2D variants harboring regioisomers of fluorophenylalanines. For example, the α_2D mutants incorporating tetrafluorinated phenylalanines display an order of $(Z^o,Z^o) > (Z^m,Z^m) > (Z^p,Z^p)$ in the stability of their dimeric structures, with the ΔG_f value of the mutant (Z^o,Z^o) favoring the folded form by nearly −4 kcal mol^{-1} more than that of mutant (Z^p,Z^p). Interestingly, the order of the mutants' folding stabilities agrees nicely with the magnitude of their dipole moments, thus indicating that dipole–dipole coupling may be an important stabilizing force of the homodimers. Similarly, comparison of the trifluorinated phenylalanine analogues shows that the homodimer of mutant (F^{345}F,F^{345}F) is more stable than that of mutant (F^{245}F,F^{245}F) by a large margin (−1.6 kcal mol^{-1} in ΔG_f) as well. In fact, the homodimer of mutant (F^{345}F,F^{345}F) exhibits the most favorable folding free energy of all α_2D variants investigated, presumably owing to the largest dipole moment of this unnatural amino acid.

Indeed, global analysis of all α_2D double mutants shows a positive correlation ($R^2 = 0.51$) between the folding free energy and the dipole moments of the corresponding fluoroaromatic rings (Figure 2c). The data of Figure 2a–c collectively suggest that the hydrophobic effect and the dipole–dipole coupling of the aromatic residues perhaps contribute equally to the stability of α_2D homodimers. Interestingly, a much improved correlation was observed when a synthetic parameter was considered (Figure 2d, $R^2 = 0.93$). As a linear combination of Log$P$ and dipole moment, this synthetic parameter reflects the contribution of both hydrophobicity and dipole moments of these aromatic rings. In other words, the combination of the hydrophobic effect and dipole–dipole interactions between the stacked phenylalanine pairs reliably predicts the overall stability of the α_2D homodimers.

Given the significance of dipole–dipole coupling between stacked aromatics, we further investigated the dipole–induced-dipole interactions by analyzing a series of α_2D single mutants (Table 2). In this (F,X) series, Phe29 is mutated to the fluorinated phenylalanine analogues, while Phe10 remains unchanged. Upon folding, a (F,X) homodimer positions the fluoroaromatic side chain to stack with that of Phe10 (Figure S5 in the Supporting Information). Therefore, the relative stability of these single mutants should reveal the best fluorinated phenylalanine analogues for targeting a native phenylalanine through face–face stacking interactions.

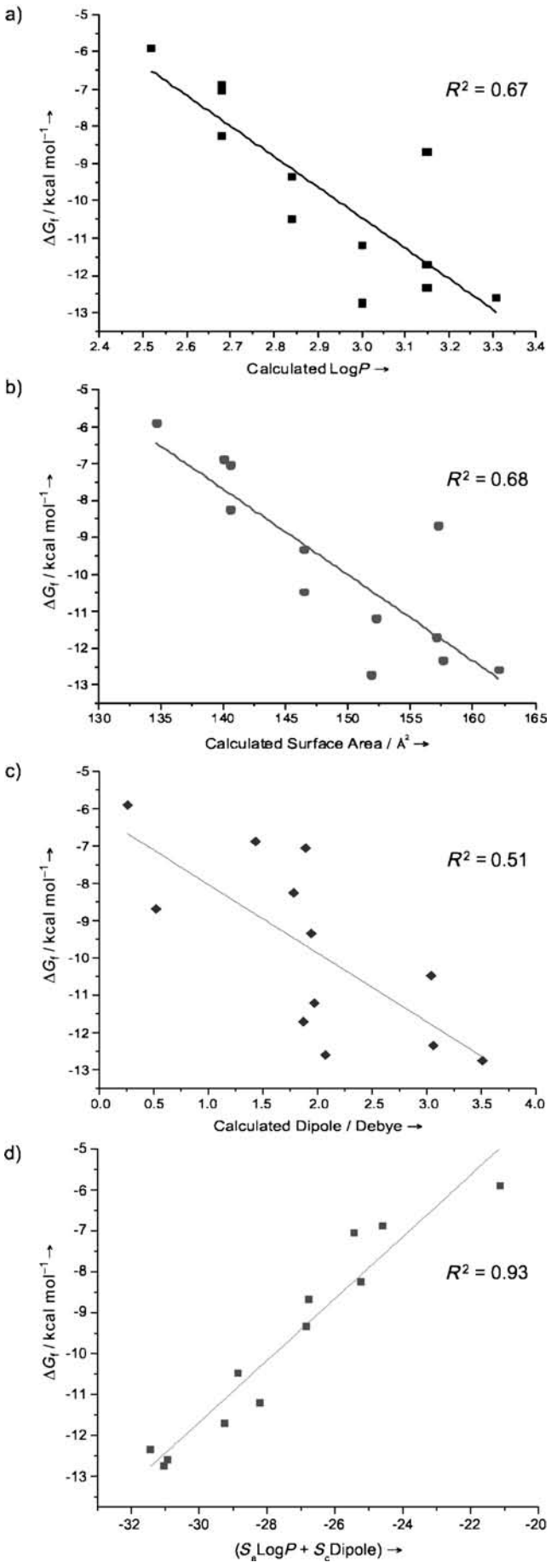

Table 2: Summary of thermodynamic parameters for α_2D single mutants.

Peptide	T_m [°C][a]	ΔH_f [kcal mol^{-1}][b]	ΔG_f [kcal mol^{-1}][c]	$\Delta\Delta G_f$ [kcal mol^{-1}][d]
F,F	28.9	-28.4 ± 2.6	-5.9 ± 0.1	0.0
F,F^{245}F	55.4	-61.5 ± 5.2	-10.1 ± 0.3	4.2
F,F^{345}F	63.1	-61.5 ± 1.3	-11.4 ± 0.1	5.5
F,Z^{p}	56.6	-51.3 ± 1.4	-9.7 ± 0.1	3.8
F,Z^{o}	68.6	-64.2 ± 2.8	-12.6 ± 0.3	6.7
F,Z	62.5	-53.7 ± 3.1	-10.9 ± 0.2	5.0

[a] Peptide concentration 20 μM, all measurements within ±1 °C.
[b] Calculated by plotting $1/T_m$ vs. ln$K$ for different peptide concentrations and using the van't Hoff equation $\ln K = -\Delta H_f/RT + \Delta S/R$.
[c] Calculated folding free energies at 37 °C.
[d] $\Delta\Delta G_f = \Delta G_{f(WT)} - \Delta G_{f(Mutant)}$.

Interestingly, comparison of the folding free energies reveals that the (F,Z^{o}) single mutant gives the most stable homodimer, which is even more stable than that of mutant (F,Z) by a large margin (-1.7 kcal mol^{-1}). This finding is particularly remarkable considering that perfluorinated phenylalanine Z is more hydrophobic and expected to have a more favorable quadrupole complementarity with phenylalanine. Analogous to the double-mutant series, comparison of single mutants (F,Z^{o}) and (F,Z^{p}) clearly demonstrates the significance of dipole–induced-dipole interactions in aromatic stacking. Specifically, the tetrafluorinated mutants (F,Z^{o}) and (F,Z^{p}), with the only difference being the position of one fluorine atom, differ in the folding free energy by approximately 3 kcal mol^{-1} (Table 2), presumably owing to the large difference of their dipole moment (3.05 Debye for Z^{o} and only 0.52 Debye for Z^{p}). This trend is also evident when mutants (F,F^{345}F) and (F,F^{245}F) are compared. When ΔG_f values are plotted against Log$P$ for the α_2D single mutants, there is no correlation ($R^2 = 0.05$); conversely, a positive correlation is observed when the folding free energy is plotted against the dipole moment ($R^2 = 0.66$, Figure S6 in the Supporting Information). Overall, the single-mutant data suggest that an optimal combination of hydrophobicity and dipole moment is necessary for the strongest stacking interaction with native aromatic residues. Consequently, the tetrafluorinated phenyl ring (Z^{o}) affords the strongest interaction.

Our data clearly demonstrate the range and hierarchy of aromatic stacking energetics. With the improved understanding, we hypothesize that aromatic interactions of different physical mechanisms (e.g. quadrupole interaction vs. dipole–dipole coupling) can direct orthogonal molecular assembly or self-sorting behavior of peptides. Toward this end, we tested the thermodynamic equilibrium of a three-component system, which consists of mutants (F,F), (F^{345}F,F^{345}F), and (Z,Z). Random dimerization of the three components will in

Figure 2. a) ΔG_f correlation with hydrophobicity Log$P$. b) ΔG_f correlation with surface area. c) ΔG_f correlation with dipole moment. d) ΔG_f correlation with the synthetic parameter (S_aLog$P$ + S_cDipole), where S_a and S_c represent the slope values from (a) and (c), respectively. See the Supporting Information for a detailed explanation of how these parameters were calculated.

principle give six dimeric species. The thermodynamic equilibrium of the three-component mixture was analyzed through a disulfide cross-linking experiment. Specifically, we mutated histidine residue His30 to a homocysteine (hC). The C_2 symmetry of the α_2D structure positions two hC side chains into close proximity for cross-linking (Figure S7 in the Supporting Information). The covalent dimers can be readily separated and identified through analytical LC–MS. Interestingly, out of the six possibilities, only two major peaks were observed (Figure 3): one corresponds to the heterodimer of

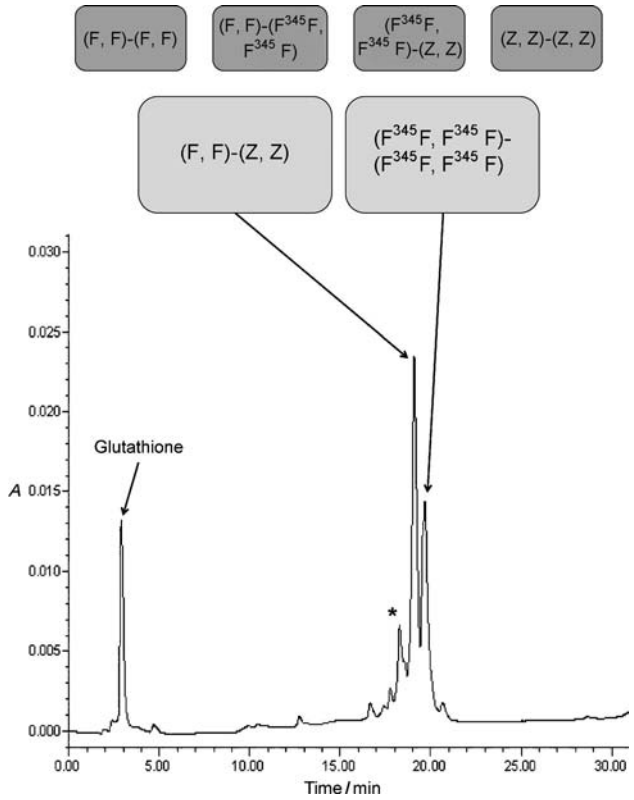

Figure 3. LC–MS chromatogramm of the disulfide cross-linking experiment with a three-component mixture of (F, F), (Z, Z), and ($F^{345}F$, $F^{345}F$). The six possible dimers are shown in boxes, and the two primary products, the (F, F)–(Z, Z) heterodimer and ($F^{345}F$, $F^{345}F$)–($F^{345}F$, $F^{345}F$) homodimer are highlighted. The third observable peak, denoted by an asterisk (*), is composed of monomer-glutathione adducts for (Z, Z) and ($F^{345}F$, $F^{345}F$), which are side products of the cross-linking reaction.

mutants (F, F) and (Z, Z); the other corresponds to the ($F^{345}F$, $F^{345}F$) homodimer. This self-sorting behavior is presumably due to the quadrupole interaction that stabilizes the (F, F)–(Z, Z) heterodimer and the dipole–dipole coupling that stabilizes the ($F^{345}F$, $F^{345}F$) homodimer. The specific assembly is particularly remarkable given that the three peptides are isosteric and differ only in the number of hydrogen-to-fluorine substitutions.

There is an increasing interest in using fluorinated amino acids in protein science and engineering because of the benefit of fluorination in NMR spectroscopy analysis, PET (positron emission tomography) imaging, and protein stabilization.[12] Herein we present, to our knowledge, the first systematic investigation of the energetics of aromatic stacking in the context of proteins. These data provide guidelines on the energetic considerations of incorporating fluorinated aromatic amino acids into target proteins. Furthermore, our results reveal the surprisingly large contribution of dipole–dipole and dipole–induced-dipole interactions to the association of aromatic pairs in the stacked geometry. These results are consistent with recent publications from both experimental[13] and theoretical[14] perspectives, which highlight the significance of dipole contributions to aromatic stacking. The comparable significance of hydrophobicity and dipole moment makes the tetrafluorophenylalanine Z° the best "warhead" to target native Phe residues through aromatic stacking. This finding is particularly important for the design of enzyme inhibitors where aromatic residues exist in the enzyme binding pocket.[15] Finally, we demonstrate that self-sorting of isosteric peptides can be achieved by solely exploiting stacked aromatics as supramolecular synthons. Future research will address the generality and scope of this approach in designing self-assembled materials and inhibitors of protein–protein interactions.

Received: August 18, 2011
Published online: November 21, 2011

Keywords: dipole–dipole interactions · protein–protein interactions · self-assembly · stacking interactions · supramolecular chemistry

[1] a) E. Gazit, *Chem. Soc. Rev.* **2007**, *36*, 1263 – 1269; b) E. H. C. Bromley, K. Channon, E. Moutevelis, D. N. Woolfson, *ACS Chem. Biol.* **2008**, *3*, 38 – 50; c) M. C. Branco, J. P. Schneider, *Acta Biomater.* **2009**, *5*, 817 – 831; d) S. Cavalli, F. Albericio, A. Kros, *Chem. Soc. Rev.* **2010**, *39*, 241 – 263; e) H. Robson Marsden, A. Kros, *Angew. Chem.* **2010**, *122*, 3050 – 3068; *Angew. Chem. Int. Ed.* **2010**, *49*, 2988 – 3005.

[2] a) S. K. Burley, G. A. Petsko, *Science* **1985**, *229*, 23 – 28; b) G. B. McGaughey, M. Gagne, A. K. Rappé, *J. Biol. Chem.* **1998**, *273*, 15458 – 15463.

[3] A limited number of examples include: a) M. Reches, E. Gazit, *Science* **2003**, *300*, 625 – 627; b) M. G. Woll, E. B. Hadley, S. Mecozzi, S. H. Gellman, *J. Am. Chem. Soc.* **2006**, *128*, 15932 – 15933; c) D. M. Ryan, T. M. Doran, B. L. Nilsson, *Langmuir* **2011**, *27*, 11145 – 11156.

[4] a) C. A. Hunter, *Angew. Chem.* **1993**, *105*, 1653 – 1655; *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 1584 – 1586; b) E. A. Meyer, R. K. Castellano, F. Diederich, *Angew. Chem.* **2003**, *115*, 1244; *Angew. Chem. Int. Ed.* **2003**, *42*, 1210; c) R. M. Johnson, K. Hecht, C. M. Deber, *Biochemistry* **2007**, *46*, 9208 – 9214; d) M. L. Waters, *Curr. Opin. Chem. Biol.* **2002**, *6*, 736 – 741; e) S. L. Cockroft, C. A. Hunter, K. R. Lawson, J. Perkins, C. J. Urch, *J. Am. Chem. Soc.* **2005**, *127*, 8594 – 8595.

[5] a) D. A. Dougherty, *Science* **1996**, *271*, 163 – 168; b) S. Mecozzi, A. P. West, Jr., D. A. Dougherty, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 10566 – 10571; c) J. P. Gallivan, D. A. Dougherty, *J. Am. Chem. Soc.* **2000**, *122*, 870 – 874; d) Y. Mo, G. Subramanian, J. Gao, D. M. Ferguson, *J. Am. Chem. Soc.* **2002**, *124*, 4832 – 4837; e) R. M. Hughes, K. R. Wiggins, S. Khorasanizadeh, M. L. Waters, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 11184 – 11188.

[6] a) C. R. Patrick, G. S. Prosser, *Nature* **1960**, *187*, 1021; b) C. A. Hunter, J. K. M. Sanders, *J. Am. Chem. Soc.* **1990**, *112*, 5525 –

5534; c) J. H. Williams, J. K. Cockcroft, A. N. Fitch, *Angew. Chem.* **1992**, *104*, 1666 – 1669; *Angew. Chem. Int. Ed. Engl.* **1992**, *31*, 1655 – 1657; d) J. H. Williams, *Acc. Chem. Res.* **1993**, *26*, 593 – 598; e) F. Cozzi, F. Ponzini, R. Annunziata, M. Cinquini, J. S. Siegel, *Angew. Chem.* **1995**, *107*, 1092 – 1094; *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1019 – 1020; f) R. E. Gillard, J. F. Stoddart, A. J. P. White, B. J. Williams, D. J. Williams, *J. Org. Chem.* **1996**, *61*, 4504 – 4505; g) A. P. West, Jr., S. Mecozzi, D. A. Dougherty, *J. Phys. Org. Chem.* **1997**, *10*, 347 – 350; h) G. W. Coates, A. R. Dunn, L. M. Henling, D. A. Dougherty, R. H. Grubbs, *Angew. Chem.* **1997**, *109*, 290 – 293; *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 248 – 251; i) S. M. Butterfield, P. R. Patel, M. L. Waters, *J. Am. Chem. Soc.* **2002**, *124*, 9751 – 9755.

[7] a) H. Zheng, J. Gao, *Angew. Chem.* **2010**, *122*, 8817 – 8821; *Angew. Chem. Int. Ed.* **2010**, *49*, 8635 – 8639; b) H. Robson Marsden, J. G. Fraaije, A. Kros, *Angew. Chem.* **2010**, *122*, 8750 – 8752; *Angew. Chem. Int. Ed.* **2010**, *49*, 8570 – 8572.

[8] R. B. Hill, W. F. DeGrado, *J. Am. Chem. Soc.* **1998**, *120*, 1138 – 1145.

[9] D. P. Raleigh, S. F. Betz, W. F. DeGrado, *J. Am. Chem. Soc.* **1995**, *117*, 7558 – 7559.

[10] L. M. Salonen, M. Ellermann, F. Diederich, *Angew. Chem.* **2011**, *123*, 4908 – 4944; *Angew. Chem. Int. Ed.* **2011**, *50*, 4808 – 4842.

[11] H. Zheng, K. Comeforo, J. Gao, *J. Am. Chem. Soc.* **2009**, *131*, 18 – 19.

[12] a) E. N. G. Marsh, *Chem. Biol.* **2000**, *7*, 153 – 157; b) N. C. Yoder, K. Kumar, *Chem. Soc. Rev.* **2002**, *31*, 335 – 341.

[13] a) J. S. Lai, J. Qu, E. T. Kool, *Angew. Chem.* **2003**, *115*, 6155 – 6159; *Angew. Chem. Int. Ed.* **2003**, *42*, 5973 – 5977; b) Y. Benitex, A. M. Baranger, *J. Am. Chem. Soc.* **2011**, *133*, 3687 – 3689.

[14] S. E. Wheeler, *J. Am. Chem. Soc.* **2011**, *133*, 10262 – 10274.

[15] For relevant examples of enzyme-inhibitor design, see: a) C. Y. Kim, J. S. Chang, J. B. Doyon, T. T. Baird, C. A. Fierke, A. Jain, D. W. Christianson, *J. Am. Chem. Soc.* **2000**, *122*, 12125 – 12134; b) C. Y. Kim, P. P. Chandra, A. Jain, D. W. Christianson, *J. Am. Chem. Soc.* **2001**, *123*, 9620 – 9627; c) J. B. Doyon, E. A. Hansen, C. Y. Kim, J. S. Chang, D. W. Christianson, R. D. Madder, J. G. Voet, T. A. Baird, C. A. Fierke, A. Jain, *Org. Lett.* **2000**, *2*, 1189 – 1192; d) A. X. Zhang, R. P. Murelli, C. Barinka, J. Michel, A. Cocleaza, W. L. Jorgensen, J. Lubkowski, D. A. Spiegel, *J. Am. Chem. Soc.* **2010**, *132*, 12711 – 12716.

