Molecular Recognition

DOI: 10.1002/ange.200600844

Ligand Preorganization May Be Accompanied by Entropic Penalties in Protein–Ligand Interactions**

*Aaron P. Benfield, Martin G. Teresk, Hilary R. Plake, John E. DeLorbe, Laura E. Millspaugh, and Stephen F. Martin**

A prevailing hypothesis in the field of molecular recognition in chemistry and biology is that the preorganization of flexible hosts and their guests in a manner corresponding to the three-

[*] A. P. Benfield, M. G. Teresk, Dr. H. R. Plake, J. E. DeLorbe, L. E. Millspaugh, Prof. S. F. Martin
Department of Chemistry and Biochemistry
and The Institute of Cellular and Molecular Biology
University of Texas at Austin
1 University Station–A5300, Austin, TX 78712-1167 (USA)
Fax: (+1) 512-471-4180
E-mail: sfmartin@mail.utexas.edu
Homepage: http://research.cm.utexas.edu/smartin/

[**] Acknowledgment is made to the National Institutes of Health, The Robert A. Welch Foundation (F-652), the Texas Advanced Research Program, Pfizer, Inc., and Merck Research Laboratories for their generous support of this research. We are also grateful to Professor Gabriel Waksman (University College London, Birkbeck College), Andrew Prongay (Schering-Plough Research Institute), and Dr. John H. Clements (The University of Texas) for helpful discussions.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

dimensional structures they adopt upon complex formation will lead to higher association constants. In simple host–guest complexes, changes in free energy ΔG that result from introducing such constraints can be highly favorable, an observation that led Cram to characterize preorganization as "a central determinant of binding power".[1,2] Ligand preorganization has also emerged as a strategy for structure-based drug design, and restricting the flexibility of peptides and other small molecules has indeed led to the discovery of selective and bioavailable drug leads that may exhibit increased potencies.[3–5]

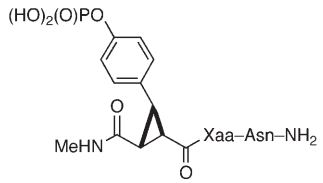
Although the energetic consequences of ligand preorganization on protein binding affinities are necessarily both enthalpic and entropic, preorganization seems universally regarded as having a favorable entropic component. Enhanced affinities arising from preorganization are thus typically, albeit simplistically, regarded as being primarily entropic in origin, provided that the constrained and flexible molecules make the same pairwise interactions with the protein and solvent. Fully restricting each independent rotor in a flexible ligand should accordingly be accompanied by an entropic advantage of about 0.7–1.6 kcal mol^{-1}.[6,7] However, increases in potency accompanying ligand preorganization are commonly much less than those predicted based upon reducing the entropic binding penalties that are associated with conformational flexibility.

Identification of the specific origin of differences in the protein binding affinities of preorganized ligands relative to their flexible analogues is problematic for several reasons. Firstly, there are few cases where association constants are determined for a pair of constrained and flexible ligands having the same number and type of heavy atoms, the same functional groups, and the same number of hydrogen-bond donors and acceptors. Appropriate controls are thus generally absent, although there are reports where the ligands either meet the above criteria[8–10] or differ only slightly in elemental composition and functionality.[7,11–15] Secondly, structural information for complexes of the constrained and flexible ligands with the biological target is typically lacking, so whether both ligands interact similarly with the biomacromolecule is unknown. Ligands also frequently bind to proteins in conformations that are higher in energy than their global minima in solution,[16] and the analysis of how differential conformational strain energies affect affinities of similar compounds is a complex matter. Finally, the specific contributions to ΔS and ΔH of binding are rarely determined, so the extent to which entropic and enthalpic factors play a role in the relative potencies of constrained and flexible ligand pairs cannot be assessed.[15] There is thus little compelling scientific evidence supporting the widely asserted hypothesis that ligand preorganization will lead to more favorable binding entropies in protein–ligand interactions.

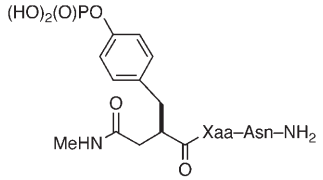
Toward explicitly elucidating the energetic and structural consequences of ligand preorganization in protein–ligand interactions, we have prepared pseudopeptides in which cyclopropane rings serve as rigid replacements for the Cα and Cβ carbon atoms and the NH groups of an amino acid residue, thereby constraining the backbone and orienting the side chain.[17] In one study, constrained and flexible derivatives of the tetrapeptide Ac-pTyr-Glu-Glu-Ile-OH, the consensus sequence in Src SH2 domain binding ligands,[18,19] were prepared that contained substituted cyclopropane and succinyl replacements for the phosphotyrosine residue.[20,21] The thermodynamic parameters for complex formation of these ligands with the Src SH2 domain were determined by isothermal titration calorimetry (ITC). Although the constrained ligands bound with more favorable entropies of binding than their flexible counterparts, this expected entropic advantage was always offset by a corresponding enthalpic penalty that resulted in comparable binding affinities of the various ligand pairs. There was thus no net energetic advantage associated with ligand preorganization because of balancing enthalpy–entropy compensation,[22,23] a ubiquitous phenomenon in host–guest and protein–ligand systems. Comparison of the X-ray crystal structures of the Src SH2 domain complexed with the cyclopropane-derived ligand and an 11-mer peptide containing the pTyr-Glu-Glu-Ile sequence revealed that both bound similarly, and all interatomic distances between the domain and the ligand in each of the two structures were in close agreement. Hence, the origin of the observed enthalpic disadvantage attending ligand preorganization was enigmatic, and investigations of other protein–ligand complexes would be required to gain further insights.

Phosphotyrosine peptides with the consensus sequence of Ac-pTyr-Xaa-Asn-NH_2, wherein Xaa is typically a hydrophobic residue, bind to the SH2 domain of the growth receptor binding protein 2 (Grb2).[18,19] The pseudopeptides **2** and **3** (see Table 1) were thus prepared as constrained and

Table 1: Thermodynamic parameters for complex formation between the Grb2 SH2 domain and pseudopeptides **2**–**7**.[a]



2: Xaa = Val
4: Xaa = Ile
6: Xaa = Leu



3: Xaa = Val
5: Xaa = Ile
7: Xaa = Leu

Cmpd.	K_a [M^{-1}]	ΔG [kcal mol^{-1}]	ΔH [kcal mol^{-1}]	ΔS [cal mol^{-1} K]
2	$(1.0\pm0.1)\times10^6$	-8.2 ± 0.1	-7.0 ± 0.1	4.2 ± 0.1
3	$(4.4\pm0.4)\times10^5$	-7.7 ± 0.1	-5.3 ± 0.1	8.2 ± 0.2
4	$(1.6\pm0.1)\times10^6$	-8.5 ± 0.1	-6.7 ± 0.1	6.0 ± 0.1
5	$(3.7\pm0.1)\times10^5$	-7.6 ± 0.1	-4.9 ± 0.1	9.1 ± 0.2
6	$(2.3\pm0.1)\times10^5$	-7.3 ± 0.1	-5.6 ± 0.1	5.9 ± 0.1
7	$(1.7\pm0.1)\times10^5$	-7.1 ± 0.1	-4.6 ± 0.1	8.6 ± 0.2

[a] ITC experiments were conducted at 25 °C in duplicate with the same batch of ligand and Grb2 SH2 domain in 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES, 50 mM) with NaCl (150 mM) at pH 7.5. Uncertainties in K_a, ΔG, ΔH, and ΔS values represent deviations from the average. Similar experiments conducted with **4** and **5** in Tris buffer gave comparable results.

flexible analogues of Ac-pTyr-Val-Asn-NH_2 (**1**).[24] As in our previous studies of Src SH2 binding ligands, the cyclopropane ring in **2** serves as a rigid mimic of the pTyr residue in **1**, whereas a benzyl succinyl moiety is a flexible pTyr replacement in the control **3**. The thermodynamic parameters (ΔG, ΔH, ΔS) for binding of **2** and **3** to the Grb2 SH2 domain were determined by ITC (Table 1), and the constrained ligand **2** was observed to bind approximately twofold better than its flexible counterpart **3**. However, in an unprecedented finding that is inconsistent with the conventional wisdom regarding the putative effects of preorganization, the entropy of binding for **2** was significantly less favorable than for **3**.

In an extension of these studies, the amino acid at the pY + 1 position of **2** and **3** was changed from Val to Ile and Leu, thus giving the corresponding ligand pairs **4**–**7** to assess the influence that variations of this hydrophobic residue might have on the energetics of binding. The thermodynamic parameters for binding of **4**–**7** to the Grb2 SH2 domain were determined (Table 1), and the constrained pseudopeptides **4** and **6** were observed to bind with slightly higher affinities than their corresponding flexible controls **5** and **7**. However, the increased affinities for the preorganized ligands **4** and **6** were again the consequence of more favorable binding enthalpies, and the binding entropies for **4** and **6** were unfavorable relative to those of their flexible counterparts **5** and **7**. The ΔC_p values for **4** and **5** were determined to be −88 and −94 cal mol^{-1} K, respectively, so differences in binding free energies do not appear to arise from solvation/desolvation or hydrophobic effects.[25] Based upon these unprecedented findings, it is now apparent that ligand preorganization may be accompanied by unfavorable binding entropies, yet more favorable binding free energies.

To explore whether structural factors might be responsible for the differential binding enthalpies of the constrained and flexible ligands, complexes of the monomeric Grb2 SH2 domain with **2** and **3** were co-crystallized for X-ray analysis. Data for the complexes of **2** and **3** with the Grb2 SH2 domain were collected to 1.9 and 1.7 Å resolution, respectively, and the structures were solved by molecular replacement using a known structure.[26,27] Similar to other Grb2 SH2 binding ligands,[26,28,29] **2** and **3** bind to the domain in a β-turn-like conformation. There are two Grb2 SH2–**2** complexes in the asymmetric unit, and the backbone atoms of the domain in these align with a root-mean-square deviation (rmsd) of 0.3 Å; atoms of **2** in the two complexes align with an rmsd of 0.2 Å. An overlay of the two Grb2 SH2–**2** complexes with that of the Grb2 SH2-**3** complex is shown in Figure 1a. The backbone atoms of the Grb2 SH2 domain in the complexes of **2** align with those in the complex of **3** with rms deviations of 0.5 and 0.6 Å. The variations occur primarily in the flexible loops, but the most notable difference, which is revealed by the close-up superimposition in Figure 1b, is the position of the BC loop that forms part of the phosphate binding pocket. In particular, the Cα and nitrogen backbone atoms of Ser90 and Glu89 in the BC loop in each of the complexes of **2** are displaced toward the ligand so they are about 2.0 Å *closer* to the phosphate group of **2** than they are in the corresponding complex with **3**. The atoms of the Val and Asn residues in the complexes of **2** and **3** superimpose with an rmsd of 0.1 Å, and the direct interactions between the domain and the Val-Asn dipeptide subunits of **2** and **3** are virtually identical. However, there are significant variations in the positions of the atoms in the constrained and flexible phosphotyrosine replacements of **2** and **3**, as these align with an rmsd of 0.6 Å. Dissimilarities in the positioning of the N-terminal methyl amide moieties of **2** and **3** appear inconsequential, whereas variations in the orientations of the phosphate groups lead to significant differences in the interactions of the phosphate oxygen atoms of **2** and **3** with the SH2 domain.

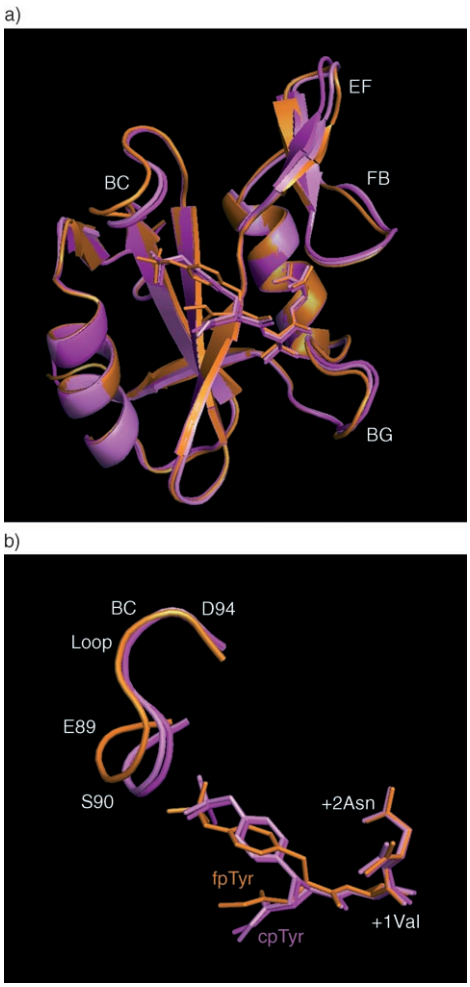


Figure 1. Superimposition of the structures of the complexes of **2** and **3** with the monomeric Grb2 SH2 domain. The protein is shown as a ribbon representation, whereas the ligands are shown as stick representations. The complex between **2** and the Grb2 SH2 domain crystallized with two molecules in the asymmetric unit, and the structure was refined to a R_{cryst} of 19.4% and a R_{free} of 22.5%. The complex between the Grb2 SH2 domain and **3** crystallized with one molecule in the asymmetric unit, and the structure was refined to a R_{cryst} of 20.7% and a R_{free} of 23.4%. a) Alignment of the two complexes in the asymmetric unit of **2** (magenta and violet) with the complex of **3** (orange). b) Close-up overlay of the structures of the two molecules of **2** from the two complexes in the asymmetric unit (magenta and violet) with the structure of **3** (orange) showing their relationship with residues Ser88-Asp94 in the BC loop. The constrained and flexible phosphotyrosine replacements are labeled cpTyr and fpTyr, respectively.

The detailed interactions between the Grb2 SH2 domain, ordered water molecules, and the phosphotyrosine replacements in **2** and **3** are presented in Figure 2a and b. An analysis of these interactions reveals that the aforementioned displacement of the BC loop and the different orientations of the phosphate group in the phosphotyrosine replacements in **2** and **3** result in a number of closer contacts between the ligand and the Grb2 SH2 domain in the complex with **2**. For example, the backbone amide N-H of Glu89 is directly hydrogen-bonded to a nonbridging oxygen atom of the phosphate group in the Grb2 SH2–**2** complex, whereas a water molecule mediates this interaction in the complex with **3**. The amide N-H of Ser90 in the Grb2 SH2–**2** complex forms a direct hydrogen bond with a nonbridging oxygen atom of the phosphate group, but there are no comparable direct interactions in the complex with **3**. The side chain of Glu89 in the complex of **2** is oriented toward the ligand, so the carboxy group interacts via a water molecule with a nonbridging oxygen atom of the phosphate group. There is no interaction between the carboxy group of Glu89 and the phosphate group in the complex of **3**.

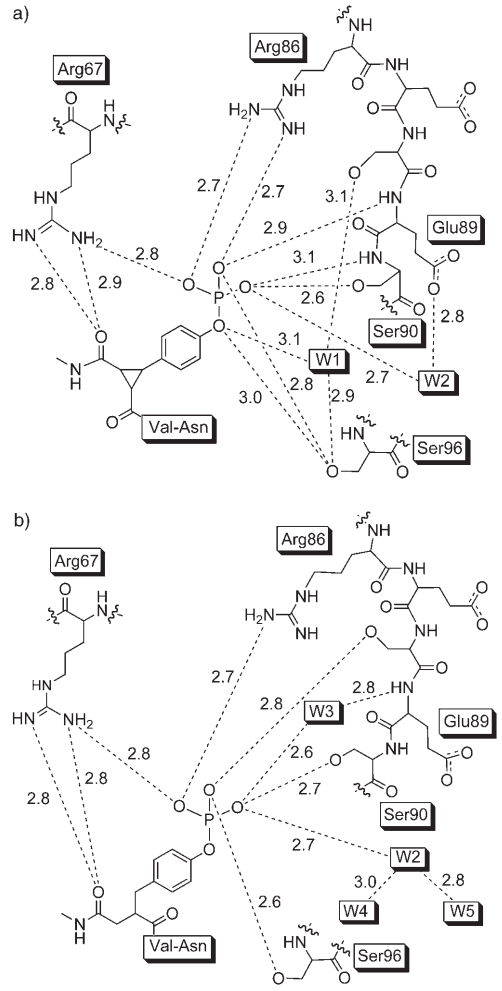


Figure 2. Hydrogen-bonding and polar interactions in the phosphotyrosine binding site of complexes of **2** and **3** with the Grb2 SH2 domain showing those water molecules, W, involved in binding networks. Polar contact distances between heavy atoms of ≤ 3.1 Å are indicated, but some hydrogen atoms involved in these interactions are omitted for clarity. a) One complex in an asymmetric unit of **2**; the corresponding pairwise distances in the other complex of **2** in the asymmetric unit are ± 0.1 Å. b) Complex of **3**.

The spatial dispositions of the atoms in the constrained and flexible phosphotyrosine replacements in **2** and **3** are different, so we queried whether these replacements might be uniquely responsible for the observed differences in the pairwise interactions with the Grb2 SH2 domain and the conformational perturbation of the BC loop. These complexes were thus compared with the structures of **8**, in which a 22-membered macrocycle constrains the pTyr-Val-Asn sequence in its biologically active conformation,[9] and the linear nonapeptide Ala-Pro-Ser-pTyr-Val-Asn-Val-Gln-Asn (**9**) bound to the Grb2 SH2 domain (Figure 3a, b).[26] The backbone atoms of the domain in the two complexes of **2** and the complex of **8** align with an rmsd of 0.3–0.4 Å; the corresponding atoms in **2** and **8** align with an rmsd fit of about 0.6 Å (Figure 3b). The polar interactions and their respective contact distances between the domain and the phosphate groups of **2** and **8** in these complexes are comparable. Hence, two different tactics for introducing a conformational constraint into a pYVN-derived ligand result in similar displacement of the BC loop toward the ligand. When the complexes of the Grb2 SH2 domain with **3** and the nonapeptide **9** are superimposed, the backbone atoms of the domain in the two complexes align with an rmsd of 0.5 Å; the corresponding atoms in **3** and **9** align with an rmsd of 0.3 Å (Figure 3c). There are slight variations in the pairwise polar interactions and contact distances between the ligands and the SH2 domain in these two complexes, but the differences are similar to those observed when comparing the structures of other phosphotyrosine ligands bound to the Grb2 SH2 domain.[9,26,28,29] These comparisons suggest that the cyclopropane ring in **2** is a good mimic of the phosphotyrosine residue in the macrocycle **8**, and that the succinyl moiety in **3** serves as a reasonable replacement of the phosphotyrosine in the linear peptide **9**. Hence, the specific phosphotyrosine replacements in **2** and **3** do not themselves appear to induce any unusual structural changes in the Grb2 SH2 domain upon binding.

The greater number and closer contact distances associated with the pairwise interactions in the Grb2 SH2–**2** complex relative to those in the Grb2 SH2–**3** complex are consistent with the more favorable enthalpy of binding observed for **2**. However, identification of the specific origins of this enthalpic advantage is problematic because of the complexity of these interactions, some of which involve bridging water molecules. Further complicating the analysis is the flexible nature of proteins and the energetic consequences thereof. Namely, the formation of a protein–ligand complex can result in an overall decrease or increase in protein flexibility that is accompanied by corresponding changes in nonbonded interactions and order throughout the complex.[23,30] The effect these changes have upon the relative magnitudes of the compensating enthalpies and entropies of

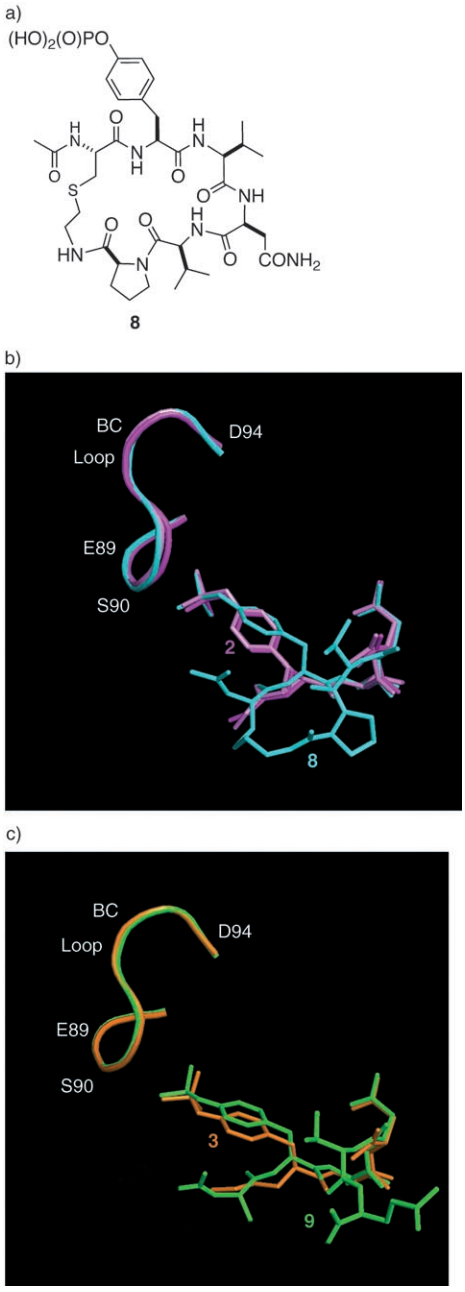


Figure 3. Complexes of different ligands with the Grb2 SH2 domain, with the protein shown in ribbon representation and the ligands shown in stick representation. a) Chemical structure of **8**. b) Alignment of the two complexes in the asymmetric unit of **2** (magenta and violet) with the published structure (2.1 Å resolution) of the complex of macrocycle **8** (cyan) in reference [9] (1BM2), which shows their relationship with residues Ser88-Asp94 in the BC loop of the respective Grb2 SH2 domains. c) Overlay of structures of **3** (orange) with the published structure (1.55 Å resolution) of the linear nonapeptide **9** (green) in reference [26] (1JYR), which has been truncated for clarity and shows their relationship with residues Ser88-Asp94 in the BC loop.

binding cannot be evaluated, because crystallographic data do not provide a reliable measure of protein flexibility in solution. Nevertheless, it is interesting in this context that a comparison of B-factors in the complexes of the Grb2 SH2 domain with **2** and **3** suggests there is *more* thermal motion in the BC loop of the Grb2 SH2–**2** complex. This observation is opposite to that which might be expected from the tighter interactions that are evident between this loop and the phosphate moiety of **2** relative to those in the Grb2 SH2–**3** complex. It is also inconsistent with the relative binding entropies of **2** and **3**.

Although preorganized ligands may bind to proteins with higher affinities than those of their flexible counterparts, we are not aware of any convincing evidence that the process must be entropically favored, as is widely assumed. Indeed, we have shown that the entropies of binding of preorganized ligands may be disfavored relative to their less potent, flexible controls. Increased affinities arising from ligand preorganization may thus result solely from more favorable enthalpies of binding. This study suggests that the enhanced enthalpy of binding can arise from an unexpected increase in the number and proximity of polar contacts between the protein and ligand; however, changes in individual nonbonded interactions throughout the complex will also contribute. These results clearly indicate that the prevailing view of the energetic consequences associated with ligand preorganization, which focuses largely if not exclusively on the ligand itself, in protein–ligand interactions must be modified. Future models must include explicit consideration of the enthalpic and entropic contributions arising from changes in nonbonded interactions and order that occur in the protein upon complex formation, and how these changes vary as a function of ligand structure and flexibility. Interactions with water molecules must also be taken into account.

Received: March 4, 2006
Revised: August 3, 2006
Published online: September 26, 2006

Keywords: ligand effects · molecular recognition · peptides · proteins · thermodynamics

[1] D. J. Cram, *Angew. Chem.* **1986**, *98*, 1041 – 1060; *Angew. Chem. Int. Ed. Engl.* **1986**, *25*, 1039 – 1057.
[2] K. N. Houk, A. G. Leach, S. P. Kim, X. Zhang, *Angew. Chem.* **2003**, *115*, 5020 – 5046; *Angew. Chem. Int. Ed.* **2003**, *42*, 4872 – 4897.
[3] H. Nakanishi, M. Kahn in *The Practice of Medicinal Chemistry*, 2nd ed. (Ed.: C. G. Wermuth), Academic Press, London, **2003**, pp. 477 – 500.
[4] D. F. Veber, S. R. Johnson, H. Y. Cheng, B. R. Smith, K. W. Ward, K. D. Kopple, *J. Med. Chem.* **2002**, *45*, 2615 – 2623.
[5] A. Mann in *The Practice of Medicinal Chemistry*, 2nd ed. (Ed.: C. G. Wermuth), Academic Press, London, **2003**, pp. 233 – 250.
[6] U. Gerhard, M. S. Searle, D. H. Williams, *Bioorg. Med. Chem. Lett.* **1993**, *3*, 803 – 808.
[7] A. R. Khan, J. C. Parrish, M. E. Fraser, W. W. Smith, P. A. Bartlett, M. N. G. James, *Biochemistry* **1998**, *37*, 16839 – 16845.
[8] B. P. Morgan, D. R. Holland, B. W. Matthews, P. A. Bartlett, *J. Am. Chem. Soc.* **1994**, *116*, 3251 – 3260.
[9] P. Ettmayer et al., *J. Med. Chem.* **1999**, *42*, 971 – 980; see Supporting Information.
[10] Y. S. Tsantrizos et al., *Angew. Chem.* **2003**, *115*, 1394 – 1398; *Angew. Chem. Int. Ed.* **2003**, *42*, 1356 – 1360; see Supporting Information.

[11] R. W. Marquis et al., *J. Med. Chem.* **2001**, *44*, 1380 – 1395; see Supporting Information.
[12] S. K. Sia, P. A. Carr, A. G. Cochran, V. N. Malashkevich, P. S. Kim, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 14664 – 14669.
[13] F. J. Dekker, N. J. de Mol, M. J. E. Fischer, J. Kemmink, R. M. J. Liskamp, *Org. Biomol. Chem.* **2003**, *1*, 3297 – 3303.
[14] A. K. Ghosh, L. M. Swanson, H. Cho, S. Leshchenko, K. A. Hussain, S. Kay, D. E. Walters, Y. Koh, H. Mitsuya, *J. Med. Chem.* **2005**, *48*, 3576 – 3585.
[15] T. Li, D. Saro, M. R. Spaller, *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1385 – 1388.
[16] E. Perola, P. S. Charifson, *J. Med. Chem.* **2004**, *47*, 2499 – 2510.
[17] A. Reichelt, S. F. Martin, *Acc. Chem. Res.* **2006**, *39*, 433 – 442.
[18] J. M. Bradshaw, G. Waksman, *Adv. Protein Chem.* **2002**, *61*, 161 – 210.
[19] K. Machida, B. J. Mayer, *Biochim. Biophys. Acta* **2005**, *1747*, 1 – 25.
[20] J. P. Davidson, O. Lubman, T. Rose, G. Waksman, S. F. Martin, *J. Am. Chem. Soc.* **2002**, *124*, 205 – 215.
[21] J. P. Davidson, PhD thesis, The University of Texas, **2001**.
[22] J. D. Dunitz, *Chem. Biol.* **1995**, *2*, 709 – 712.
[23] D. H. Williams, E. Stephens, D. P. O'Brien, M. Zhou, *Angew. Chem.* **2004**, *116*, 6760 – 6782; *Angew. Chem. Int. Ed.* **2004**, *43*, 6596 – 6616.
[24] H. R. Plake, T. B. Sundberg, A. R. Woodward, S. F. Martin, *Tetrahedron Lett.* **2003**, *44*, 1571 – 1574.
[25] J. M. Sturtevant, *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 2236 – 2240.
[26] P. Nioche, W.-Q. Liu, I. Brouitin, F. Charbonnier, M.-T. Latreille, M. Vidal, B. Roques, C. Garbay, A. Ducruix, *J. Mol. Biol.* **2002**, *315*, 1167 – 1177.
[27] X-ray crystallographic data are deposited in the Protein Data Bank under the accession codes 2HUW and 2HUY.
[28] J. Rahuel et al., *Nat. Struct. Biol.* **1996**, *3*, 586 – 589; see Supporting Information.
[29] J. Rahuel, C. Garcia-Echeverria, P. Furet, A. Strauss, G. Caravatti, H. Fretz, J. Schoepfer, B. Gay, *J. Mol. Biol.* **1998**, *279*, 1013 – 1022.
[30] C.-Y. Yang, R. Wang, S. Wang, *J. Med. Chem.* **2005**, *48*, 5648 – 5650.

