Preorganization in biological systems: Are conformational constraints worth the energy?*

Stephen F. Martin

Department of Chemistry and Biochemistry and The Institute of Cellular and Molecular Biology, The University of Texas, Austin, TX 78712, USA

Abstract: It is generally assumed that preorganizing a flexible ligand in the three-dimensional shape it adopts when bound to a macromolecular receptor will provide a derivative having an increased binding affinity, primarily because the rigidified molecule is expected to benefit from a lesser entropic penalty during complexation. We now provide the first experimental evidence that demonstrates this common belief is not universally true. Indeed, we find that ligand preorganization may be accompanied by an unfavorable entropy of binding, even when the constrained ligand exhibits a higher binding affinity than its flexible control. Thus, the effects that ligand preorganization have upon energetics and structure in protein–ligand interactions must be reevaluated.

Keywords: ligand preorganization; molecular recognition; protein–ligand interactions; thermodynamics; structure; enzyme inhibitors.

INTRODUCTION

Restricting the flexibility of peptides and other small molecules has proven to be an effective paradigm for discovering selective and bioavailable drug leads and candidates. One of the guiding principles in such endeavors is that preorganizing mobile ligands in a manner that corresponds to the three-dimensional structures they adopt upon binding to their biomacromolecular targets (i.e., their biologically active conformations) will lead to higher association constants [1]. It is often stated that this increased affinity is the consequence of a lower entropic penalty, which should be approximately 0.7–1.6 kcal/mol for each restricted rotor [2], provided both ligands interact similarly with protein and solvent. This simplistic view ignores, however, the important fact that preorganization has both enthalpic and entropic components. Unfortunately, it is difficult to assess to what extent preorganization itself affects energetics in protein–ligand interactions because there are few cases where the consequence of introducing a constraint has been explicitly assessed and compared with a flexible control having the same number and type of heavy atoms, the same functional groups, and the same number of hydrogen bond donors and acceptors [3]. Even when there is a close structural correspondence between constrained and flexible ligands, the specific contributions to ΔS and ΔH of binding associated with preorganizing the ligand are rarely determined. Structural information for the complexes of both ligands with the target protein is also often lacking. There is thus surprisingly little compelling scientific evidence to support the commonly espoused belief that favorable binding entropies attend ligand preorganization in biological

systems. Indeed, recent work in our laboratories reveals that preorganizing flexible protein-binding ligands does not necessarily lead to the expected entropic advantage.

RESULTS AND DISCUSSION

Ligand design

Toward developing new strategies for structure-based ligand design, we have long been interested in evaluating the thermodynamic and structural consequences of introducing selected conformational constraints into biologically relevant peptides. In this context, we invented a novel class of peptide mimics wherein cyclopropane rings serve as rigid replacements for the C_α and C_β carbon atoms and the NH groups of an amino acid residue **1** as shown in **2** [4]. The *trans* relationship of the substituents at C(1) and C(3) in **2** was predicted to locally stabilize a β-strand. Modeling shows that the R^1 group at C(2) in **2** is oriented so it occupies approximately the same region of space relative to the backbone that it would if the χ_1–angle in the corresponding amino acid residue in **1** were gauche(–). Although the cyclization of **1** to give **2** removes an amide N–H and potential hydrogen-bonding capability from the resultant pseudopeptide, we reasoned that the favorable energetic advantage that arises from restricting at least two rotors should approximately offset the energetic cost of losing an intermolecular hydrogen bond. Having designed cyclopropanes **2** as peptide replacements, it remained to evaluate their efficacy.
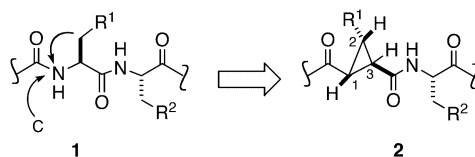

Some early results

Before we could determine the biological activity of pseudopeptides containing replacements such as **2**, it was necessary to develop general methods for the enantioselective synthesis of trisubstituted cyclopropanes. We discovered that the chiral rhodium(II) carboxamide catalyst Rh_2(5*S*-MEPY), which was originally developed by Doyle for bimolecular cyclopropanations, could be used to induce efficient cyclizations of diazoacetates **3** to give lactones **4** with high enantioselectivities, except when R^1 = H, R^2 = alkyl or aryl (Scheme 1) [5].
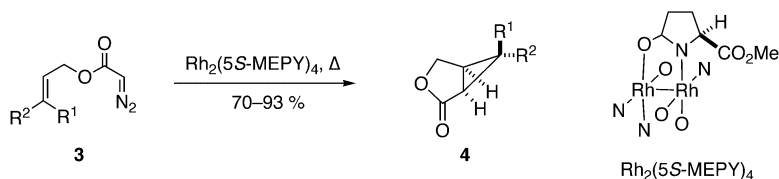

Scheme 1

Having developed methodology for the enantioselective synthesis of lactones **4**, tactics for elaborating these intermediates into cyclopropane-derived peptide mimics **2** were developed so we could evaluate the efficacy of such replacements in biologically relevant systems. In early experiments, we discovered that **5** and the flexible pseudopeptide **6** were approximately equipotent as subnanomolar inhibitors of renin [6]. Even though the constrained analog did not benefit from the expected higher affinity, this was an exciting initial result because introducing peptide mimics frequently results in signifi-

cant loss of activity. Moreover, this finding supported our hypothesis that the substituents on the cyclo-propane ring in **5** were positioned in a manner that correlated well with the three-dimensional orienta-tion of the corresponding groups in the bound structure of **6**. We then reasoned that cyclopropanes re-lated to **2** might serve as topographical probes of the biologically active conformations of peptides.

5   **6**
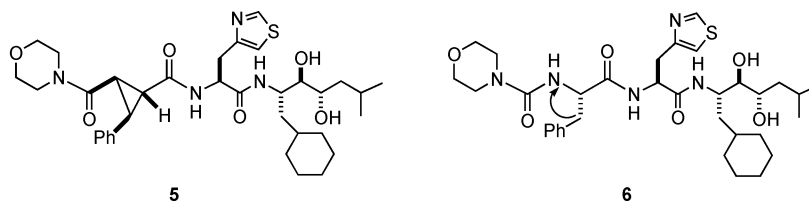
In subsequent studies, we found that **7** and **8** were both subnanomolar inhibitors of HIV-1 pro-tease and comparable in potency to the more flexible analog **9** [7]. The structure of **7** in solution was determined by NMR spectroscopy, and the structure of **8** bound to HIV-1 protease was established by X-ray crystallography. With the exception of the terminal benzyl groups that are freely rotating in so-lution, the solution structure of **7** and the bound structure of **8** were quite similar. Moreover, the bound structure of the bis-cyclopropane core of **8** corresponded closely to the structure of the Val-diol-Val subunits in the bound conformations of other more flexible HIV-1 protease inhibitors such as **9**. The two *trans*-cyclopropane rings in **7** had thus performed their assigned role of stabilizing an extended β-strand structure in solution that admirably mimicked the biologically active conformation of more flexible HIV-1 protease inhibitors.


7: $R^1 = R^2 = Me$
8: $R^1 = Me; R^2 = H$

9
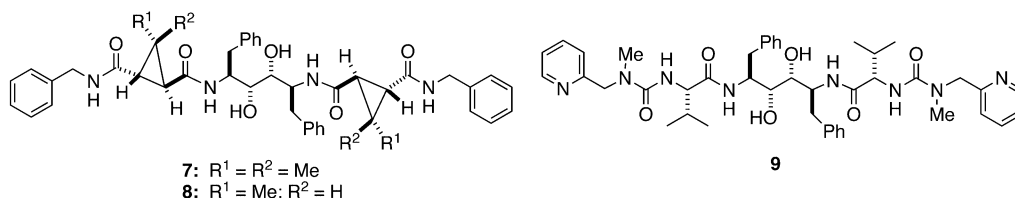
These early studies thus established the viability of using peptide mimics related to **2** in the syn-thesis of conformationally constrained pseudopeptides having high potencies. However, the renin in-hibitors **5** and **6** and the HIV-1 protease inhibitors **7–9** differ in the number and type of heavy atoms, so the energetic and structural effects associated explicitly with preorganization could not be ascertained in these systems. In order to address this deficiency, another testing ground was sought.
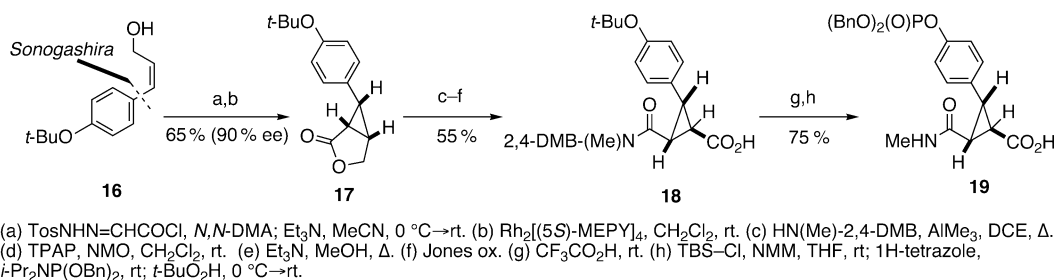
SH2 binding ligands

Interactions of SH2 domains with proteins containing a phosphorylated tyrosine residue (pTyr or pY) play key roles in numerous signal transduction pathways [8]. Compounds that selectively block bind-ing to SH2 domains have thus been pursued for studying mechanisms of signal transduction and for their potential as drug candidates to treat cancer and other diseases. Structural studies of complexes of Src SH2 domains with phosphopeptides related to pTyr–Glu–Glu–Ile (pYEEI) (**10**) reveal that these an-tagonists bind to the Src SH2 domain in *extended* conformations in which two well-defined binding pockets on the SH2 domain accommodate the pTyr residue and the hydrophobic residue (e.g., Ile, etc.) at the pY+3 position. On the other hand, phosphotyrosine peptides having sequences related to pTyr–Val–Asn (pYVN) (**11**) bind in *turned* conformations to the Grb2 SH2 domain.

Despite the different binding modes of phosphotyrosine-derived peptides to the Src and Grb2 SH2 domains, the pY moieties invariably bind with a high degree of similarity. Modeling the interactions in complexes of **10** with the Src SH2 domain and of **11** with the Grb2 SH2 domain suggested that **12** and **14** would serve as excellent constrained analogs of **10** and **11**, respectively. In particular, the cyclopropane ring in **12** and **14** orients the aromatic ring in the correct orientation relative to the backbone, and it locally rigidifies the backbone at the pY position in an extended conformation. Compounds **13** and **15** would then serve as appropriate flexible controls. Because the amide N–H at the pY–1 position of **10** and **11** is not involved in any hydrogen-bonding interactions with their respective SH2 domains, introducing the phosphotyrosine replacements in **12–15** was not expected to adversely affect the number of hydrogen-bonding and polar contacts within the binding site. We thus reasoned that ligands **12** and **13** would be nicely suited for probing the energetic and structural effects of ligand preorganization of Src SH2 binding ligands, whereas **14** and **15** would serve the same purpose for Grb2 SH2 binding ligands.
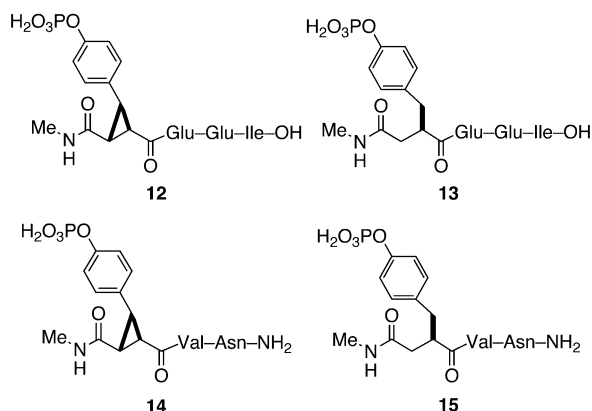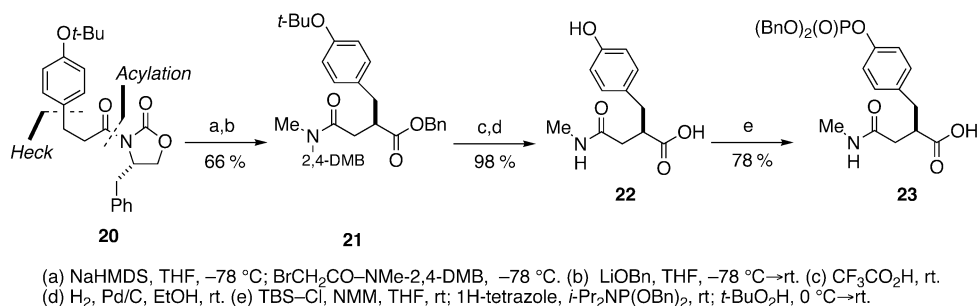

Synthesis of SH2 binding ligands

It was first necessary to prepare the constrained and flexible phosphotyrosine replacements **19** and **23**, respectively, which are present in **12–15**. This task was readily accomplished according to the reactions outlined in Schemes 2 and 3. The starting materials **16** and **20** were each prepared in three and four steps, respectively, from *p*-iodophenol. Amide bond formation between **19** and **23** and protected Glu–Glu–Ile tripeptides using standard peptide-coupling procedures followed by global deprotection then delivered **12** and **13** [9]. Pseudopeptides **14** and **15** were similarly prepared from **19** and **23** and a Val–Asn–NH₂ dipeptide [10].

(a) TosNHN=CHCOCl, *N,N*-DMA; Et₃N, MeCN, 0 °C→rt. (b) Rh₂[(5*S*)-MEPY]₄, CH₂Cl₂, rt. (c) HN(Me)-2,4-DMB, AlMe₃, DCE, Δ. (d) TPAP, NMO, CH₂Cl₂, rt. (e) Et₃N, MeOH, Δ. (f) Jones ox. (g) CF₃CO₂H, rt. (h) TBS–Cl, NMM, THF, rt; 1H-tetrazole, *i*-Pr₂NP(OBn)₂, rt; *t*-BuO₂H, 0 °C→rt.

Scheme 2

(a) NaHMDS, THF, −78 °C; BrCH$_2$CO–NMe-2,4-DMB, −78 °C. (b) LiOBn, THF, −78 °C→rt. (c) CF$_3$CO$_2$H, rt.
(d) H$_2$, Pd/C, EtOH, rt. (e) TBS–Cl, NMM, THF, rt; 1H-tetrazole, *i*-Pr$_2$NP(OBn)$_2$, rt; *t*-BuO$_2$H, 0 °C→rt.

Scheme 3

Thermodynamic and structural studies of Src SH2 binding ligands

The thermodynamic parameters (ΔG, ΔH, ΔS) for forming complexes of **12** and **13** with the Src SH2 domain were determined by isothermal titration microcalorimetry (ITC) [11]. Consistent with our expectations, an entropic advantage of ~9 eu/mol arose from preorganizing **13**. However, the more favorable entropy of binding for the constrained ligand **12** was offset by an enthalpic penalty that resulted in approximately equal binding affinities of **12** and **13**. The heat capacity change, ΔC_p was about −220 cal mol^{-1} K^{-1} for both **12** and **13**, so differences in entropic contributions do not appear to arise from solvation/desolvation or hydrophobic effects [12]. We also prepared three pseudopeptide derivatives in which the Glu residues at the pY+1 and pY+2 positions were sequentially replaced with an Asp residue, and the Ile at the pY+3 position was replaced with a Val residue. In every case, constrained and flexible ligand pairs bound with approximately the same affinities, and each of the constrained ligands exhibited a more favorable entropy of binding than its flexible counterpart; however, this entropic advantage was always balanced by an enthalpic penalty.

In a preliminary attempt to elucidate the origin of the observed enthalpic penalty in this series of test ligands, the structure of the complex of **12** with the Src SH2 domain was determined [11], and this structure was compared with that of an 11-mer peptide derivative bound to the Src SH2 domain [13]. The pYEEI segment of both ligands bound in similar extended conformations, and the interatomic distances between the domain and the ligands in each of the two structures were in close agreement. It was thus not possible on the basis of these structural data to identify the basis for the enthalpic differences in the binding energetics of **12** and **13**. A structure of the complex of **13** with the Src SH2 domain is needed, and we are presently working toward this objective.

Thermodynamic and structural studies of Grb2 SH2 binding ligands

In a similar series of studies, ΔG, ΔH, and ΔS for forming complexes of **14** and **15** with the Grb2 SH2 domain were determined by ITC [14]. All compounds exhibited low micromolar affinities. Although the constrained ligand **14** bound approximately two-fold better than its flexible counterpart **15**, *it bound with a binding entropy that was approximately 4 eu/mol less favorable than **15***. Similar results have been observed with analogs of **14** and **15** in which the Val residue was replaced with Leu and Ile residues. *These findings are unprecedented and completely unexpected based upon the conventional wisdom regarding the putative energetic effects associated with ligand preorganization.* In these cases, the slight enhancement in binding affinities of constrained ligands results from more favorable enthalpies, not entropies, of binding.

We solved the structures of the Grb2 SH2 domain complexed with the cyclopropane **14** and its flexible analog **15**. Inspection of the structures of complexes of **14** and **15** with the Grb2 SH2 domain shows the protein backbones in the two structures overlay closely (average rmsd of 0.5 Å) with the most noteworthy variations being in the flexible BC loop, which is displaced significantly in the Grb2 SH2/**14** complex relative to the Grb2 SH2/**15** complex. The direct interactions between the SH2 domain and the Val–Asn residues of **14** and **15** are virtually identical, but the displacement of the BC loop in

the former complex results in significant differences in the interactions between the phosphate groups of **14** and **15** with the domain (Fig. 1a). Namely, the C_α and nitrogen backbone atoms of Ser90 and Glu89 are approximately 2 Å *closer* to the phosphate group of the phosphotyrosine replacement in **14** than in the corresponding complex with **15**. This conformational change in the BC loop results in additional and closer hydrogen-bonding and polar contacts in the Grb2 SH2/**14** complex.

These structures were then compared with those of Grb2 SH2 complexed with the macrocycle **24** [3b] and the linear nonapeptide Ala–Pro–Ser–pTyr–Val–Asn–Val–Gln–Asn (**25**) [15]. The polar interactions and respective contact distances between the domain and the phosphate groups of **14** and **24** (Fig. 1b) were comparable as were those of **15** and **25** (Fig. 1d). These comparisons suggest that the phosphotyrosine replacements in **14** and **15** are good mimics of the phosphotyrosine residues in **24** and **25**, respectively. Hence, these replacements do not themselves appear to be uniquely responsible for inducing structural changes in the Grb2 SH2 domain.
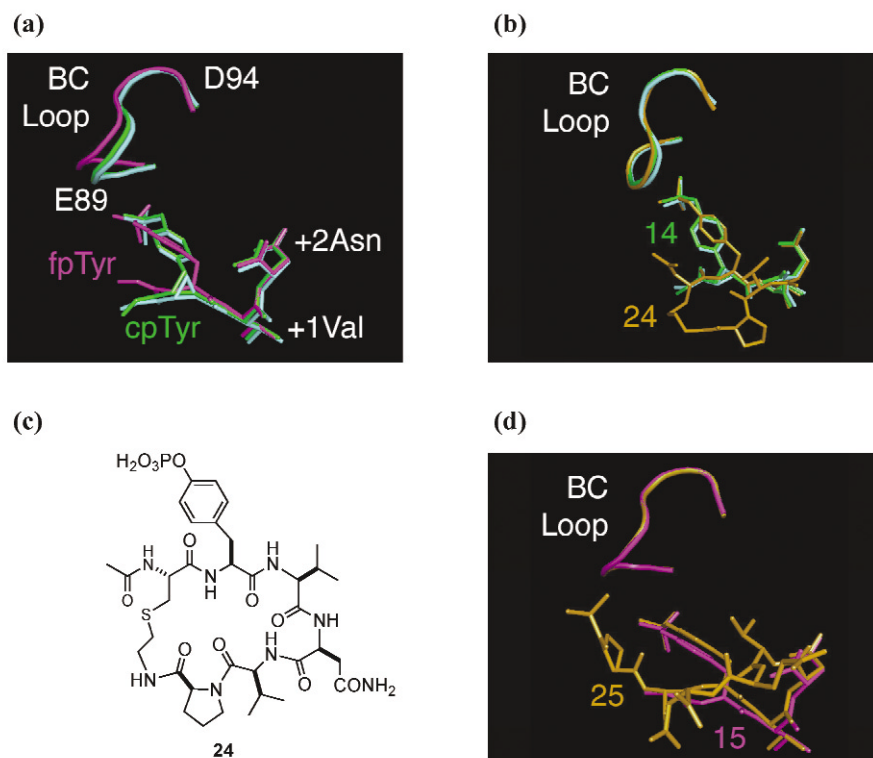

Fig. 1 Superimposition of the structures of the complexes of **14**, **15**, **24**, and **25** with the Grb2 SH2 domain. (**a**) Alignment of the structures of the two molecules of **14** from the two complexes in the asymmetric unit (green and cyan) with the structure of **15** (magenta) showing residues Ser88–Asp94 in the BC loop. The constrained and flexible phosphotyrosine replacements are labeled cpTyr and fpTyr, respectively. (**b**) Alignment of the two complexes in the asymmetric unit of **14** (green and cyan) with structure of the complex of **24** (orange) in ref. [3b] showing the BC loops. (**c**) Chemical structure of **24**. (**d**) Overlay of structure of **15** (magenta) with published structure of the linear nonapeptide **9** (orange) in ref. [15] showing the BC loops.

The greater number and closer contact distances between the phosphate group of the phosphotyrosine replacement in **14** and the BC loop of the Grb2 SH2 domain relative to those interactions in the complex with **15** are consistent with the more favorable enthalpy of binding observed for **14**. However, one cannot easily identify the specific origin(s) of either the enthalpic advantage or the entropic disadvantage associated with complexation of **14**. For example, forming a protein–ligand complex may result in an overall decrease or increase in protein flexibility that will be accompanied by changes in non-bonded interactions, which will have enthalpic consequences, and order, which will have entropic consequences, throughout the complex [16]. Although crystallographic B-factors are sometimes used to assess protein flexibility [17], these data do not provide information that is as reliable as dynamic NMR experiments, which are underway. Nevertheless, given the "tighter" packing of the BC loop in the Grb2 SH2/**14** complex, one might predict that there would be less flexibility in the BC loop of this complex relative to the Grb2 SH2/**15** complex. However, analysis of B-factors for the two complexes reveals that the thermal motions of the atoms in the BC loop of the Grb2 SH2/**14** complex are *greater than those in* the Grb2 SH2/**15** complex, suggesting that *there is more, not less, flexibility in the binding pocket of the energetically and enthalpically favored complex.*

CONCLUSIONS

Preorganized ligands may bind to proteins with higher affinities than their flexible counterparts. However, we are aware of no convincing experimental evidence that this process must be entropically favored as is widely purported. Indeed, we have shown that the *entropies of binding of preorganized ligands may be disfavored relative to their flexible controls* even though the constrained ligands may bind with more favorable free energies. This study suggests that the enhanced enthalpy of binding of a constrained ligand can arise from an *unexpected* increase in the number and proximity of polar contacts between the protein and ligand that do not directly involve the constraining atoms. However, one cannot focus solely on the interactions between the ligand and the protein, because changes in individual non-bonded interactions and order throughout the complex will contribute to the binding free energies. The prevailing view of the energetic consequences of ligand preorganization, which focuses primarily upon the ligand itself, in protein–ligand interactions must thus be modified. Future models must include explicit consideration of the enthalpic and entropic contributions arising from changes in nonbonded interactions and order that occur throughout the protein upon complex formation and how these changes vary as a function of ligand structure and flexibility. And then there is water....

ACKNOWLEDGMENTS

REFERENCES

1. A. Mann. In *The Practice of Medicinal Chemistry*, 2nd ed., C. G. Wermuth (Ed.), pp. 233–250, Academic Press, London (2003).
2. For example, see: (a) U. Gerhard, M. S. Searle, D. H. Williams. *Bioorg. Med. Chem. Lett.* **3**, 803 (1993) and refs. therein; (b) A. R. Khan, J. C. Parrish, M. E. Fraser, W. W. Smith, P. A. Bartlett, M. N. G. James. *Biochemistry* **37**, 16839 (1998).
3. For some examples, see: (a) B. P. Morgan, D. R. Holland, B. W. Matthews, P. A. Bartlett. *J. Am. Chem. Soc.* **116**, 3251 (1994); (b) P. Ettmayer, D. France, J. Gounarides, M. Jarosinski, M. S. Martin, J. M. Rondeau, M. Sabio, S. Topiol, B. Weidmann, M. Zurini, K. W. Bair. *J. Med. Chem.* **42**, 971 (1999); (c) Y. S. Tsantrizos, G. Bolger, P. Bonneau, D. R. Cameron, N. Goudreau, G. Kukolj, S. R. LaPlante, M. Llinas-Brunet, H. Nar, D. Lamarre. *Angew. Chem., Int. Ed.* **42**, 1356 (2003).
4. For a review, see: A. Reichelt, S. F. Martin. *Acc. Chem. Res.* **39**, 433 (2006).
5. For a leading reference, see: M. P. Doyle, R. E. Austin, A. S. Bailey, M. P. Dwyer, A. B. Dyatkin, A. V. Kalinin, M. M. Y. Kwan, S. Liras, C. J. Oalmann, R. J. Pieters, M. N. Protopopova, C. E. Raab, G. H. P. Roos, Q.-L. Zhou, S. F. Martin. *J. Am. Chem. Soc.* **117**, 5763 (1995).
6. S. F. Martin, R. E. Austin, C. J. Oalmann, W. R. Baker, S. L. Condon, E. deLara, S. H. Rosenberg, K. P. Spina, H. H. Stein, J. Cohen, H. D. Kleinert. *J. Med. Chem.* **35**, 1710 (1992).
7. S. F. Martin, G. O. Dorsey, T. Gane, M. C. Hillier, H. Kessler, M. Baur, B. Matha, J. W. Erickson, T. N. Bhat, S. Munshi, S. V. Gulnick, I. A. Topol. *J. Med. Chem.* **41**, 1581 (1998).
8. For reviews, see: (a) J. M. Bradshaw, G. Waksman. *Adv. Protein Chem.* **61**, 161 (2002); (b) K. Machida, B. J. Mayer. *Biochim. Biophys. Acta* **1747**, 1 (2005).
9. J. P. Davidson, S. F. Martin. *Tetrahedron Lett.* **41**, 9459 (2000).
10. H. R. Plake, T. B. Sundberg, A. R. Woodward, S. F. Martin. *Tetrahedron Lett.* **44**, 1571 (2003).
11. J. P. Davidson, O. Lubman, T. Rose, G. Waksman, S. F. Martin. *J. Am. Chem. Soc.* **124**, 205 (2002).
12. J. M. Sturtevant. *Proc. Natl. Acad. Sci. USA* **74**, 2236 (1977).
13. G. Waksman, S. E. Shoelson, N. Pant, D. Cowburn, J. Kuriyan. *Cell* **72**, 779 (1993).
14. A. P. Benfield, M. G. Teresk, H. R. Plake, J. E. DeLorbe, L. E. Millspaugh, S. F. Martin. *Angew. Chem., Int. Ed.* **45**, 6830 (2006).
15. P. Nioche, W.-Q. Liu, I. Broutin, F. Charbonnier, M.-T. Latreille, M. Vidal, B. Roques, C. Garbay, A. Ducruix. *J. Mol. Biol.* **315**, 1167 (2002).
16. For a discussion of effects of complexation upon changes in protein flexibility, see: D. H. Williams, E. Stephens, D. P. O'Brien, M. Zhou. *Angew. Chem., Int. Ed.* **43**, 6596 (2004).
17. For a recent discussion of using B-factors to compare protein flexibility in protein–ligand complexes, see: C.-Y. Yang, R. Wang, S. Wang. *J. Med. Chem.* **48**, 5648 (2005).