DOI: 10.1002/cmdc.200800214

The Use of Chemical Double-Mutant Cycles in Biomolecular Recognition Studies: Application to HCV NS3 Protease Inhibitors

Stephen H. Kawai,* Murray D. Bailey, Ted Halmos, Pat Forgione, Steven R. LaPlante, Montse Llinàs-Brunet, Julie Naud, and Natalie Goudreau[a]

Chemical double-mutant cycles (DMCs) have been elegantly used to quantify the energies of noncovalent interactions between specific pairs of functional groups in the molecular recognition of abiotic systems.[1] The method has also been employed to estimate the free energy of a novel noncovalent interaction in an intramolecular context.[2] The analysis assumes that there are no significant structural differences between the four components of the cycle and that the free energies of the individual interactions are additive. It can then be used to determine the energetic contribution of a specific interaction in the context of a more complex, global binding event in such a manner that secondary interactions are effectively cancelled out.

Here, we describe the novel use of a particular type of chemical DMC, which serves as a useful formalism to characterize the mutual effect of different substructures within a molecule on a biomolecular recognition process, namely, binding to an enzyme active site. This treatment varies from those described above in that two "deletion mutations" occupy different positions on the same chemical species while the binding partner is kept constant. Contrary to the DMC whose validity rests on the conditions of unchanging structure and the additive contributions of subsites to the total binding energy, the present analysis is useful in providing a clear indication of whether two portions of a ligand contribute to the total binding energy in an additive or cooperative (synergistic or antagonistic[3]) manner, thus affording insight into the binding phenomenon and/or aiding in the understanding of structure–activity relationships (SAR) when structural data is lacking. The treatment is exemplified by thermodynamic cycles constructed from peptidic inhibitors of the hepatitis C virus NS3 protease (HCV protease) where the binding event is competitive inhibition of the enzyme.

Examples of all three cases (additive, synergistic and antagonistic) are described. HCV protease represents an important target in the quest for a specific antiviral agent against hepatitis C infection, a serious global health problem for which the therapeutic need has not been met.[4] Of central importance in our early efforts was the discovery that this novel serine protease suffers product inhibition by oligopeptide N-terminal cleavage products.[5] Optimization of product-based hexapeptides entailed N-truncation,[6] elaboration of a novel P1[7] residue[8] and the incorporation of an aryloxy substituent onto the P2 residue, which led to potent tripeptides bearing a free C-terminal carboxylate. This series ultimately provided the first HCV protease inhibitor demonstrated to reduce viral load in infected patients.[9]

A molecule central to this optimization effort was inhibitor **4**, which has become a key standard and biological tool in the study of this class of ligands[10] and the enzyme target. As is generally observed for peptidyl inhibitors of serine proteases, compound **4** is bound to HCV protease in an extended conformation (Figure 1), canonical H-bonds involving the P1-NH, P3-NH and P3-carbonyl groups anchoring the ligand in an antiparallel β-sheet arrangement with the E2-strand of the protein.

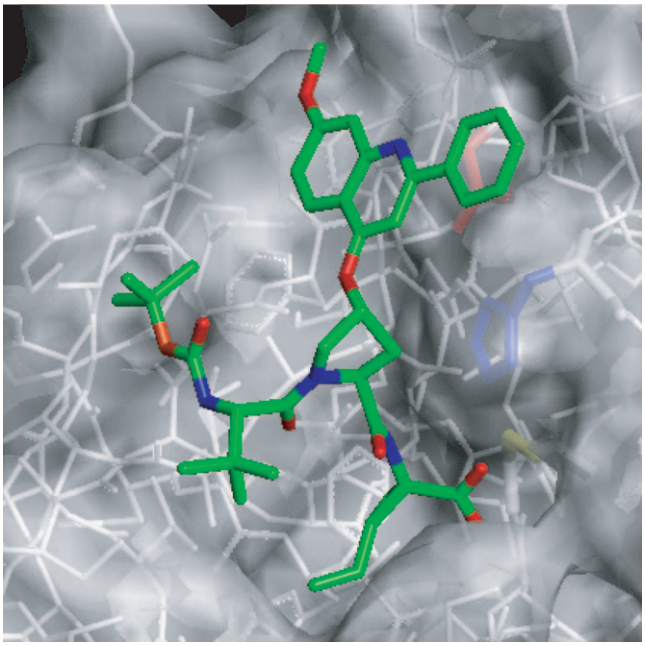
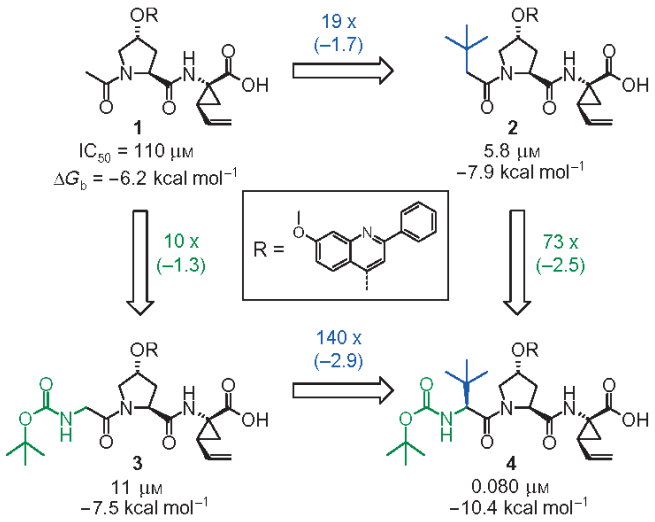


Figure 1. Representation of inhibitor **4** (green) bound in the active site of HCV protease.

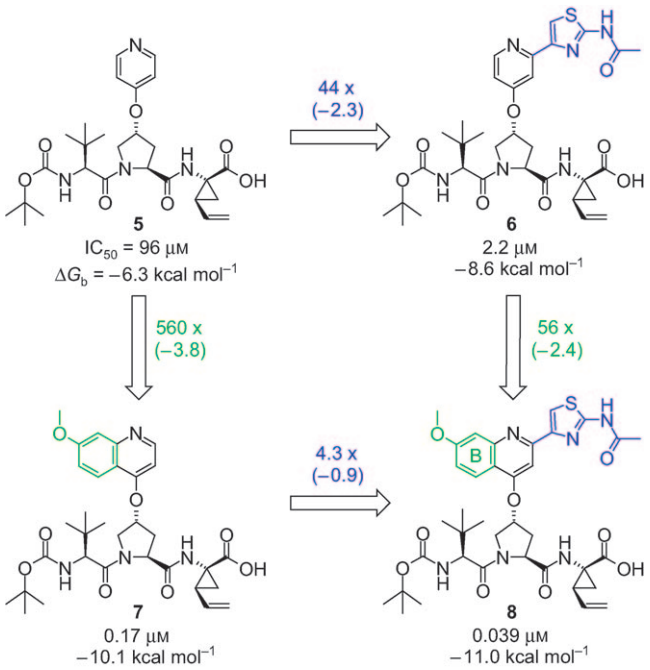
N-terminal truncation of inhibitor **4** to the acetylated dipeptide **1** results in a much weaker ligand, which serves as the starting point for the construction of a chemical DMC where the effects, both independently and combined, of the *tert*-butyl and Boc-amino moieties of molecule **4** can be assessed. As shown in Scheme 1, both the *tert*-butylacetylated dipeptide **2** and the Boc-glycine tripeptide **3** exhibit improved potency. If the binding contributions of these two extensions to inhibitor **1** were additive, one would expect to obtain a molecule exhib-

[a] *Dr. S. H. Kawai, Dr. M. D. Bailey, T. Halmos, Dr. P. Forgione, Dr. S. R. LaPlante, Dr. M. Llinàs-Brunet, J. Naud, Dr. N. Goudreau*
Department of Chemistry, Boehringer Ingelheim (Canada) Ltd., 2100, rue Cunard, Laval, Québec H7S 2G5 (Canada)
Fax: (+1) 450-682-8434
E-mail: stephen.kawai@boehringer-ingelheim.com

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cmdc.200800214.



Scheme 1. Chemical DMC used to evaluate the role/relationship of the P3 side chain and capping groups in inhibitor binding to HCV protease. Numbers next to arrows indicate fold improvement in IC_{50} values with the $\Delta\Delta G_b$ (kcal mol^{-1}) in parentheses.



Scheme 2. Chemical DMC used to evaluate the relationship between the P2 aryloxy B-ring and thiazolyl substituent of the tripeptide **8**. Numbers next to arrows indicate fold improvement in IC_{50} values with the $\Delta\Delta G_b$ (kcal mol^{-1}) in parentheses.

iting an IC_{50} value of 0.58 μM (ΔG_b of -9.2 kcal mol^{-1}). However, the activity of tripeptide **4** is significantly better (80 nM), reflecting a binding free energy of 1.2 kcal mol^{-1} in excess of the value expected if the Boc-amino and *tert*-butyl groups were noninteracting. We believe this *synergy* is primarily due to conformational/dynamic factors. Detailed studies[11] probing the behavior of the free ligands in solution indicate that the presence of both groups results in the global rigidification of peptide **4** to the bioactive conformation such that the Boc-P3 region is properly disposed in an extended fashion to form the two canonical H-bonds with Ala 157 of the protease.

The second chemical DMC is presented in Scheme 2 and allows us to analyze the contributions of the P2 quinoline B-ring and the thiazole moieties to the binding of inhibitor **8**, a close analogue of compound **4**, which is complexed by HCV protease in a similar fashion. Deletion of these two substructures to afford pyridine **5** results in a nearly 2500-fold decrease in binding affinity. Incorporation of either the methoxylated B-ring or the thiazolyl moiety into inhibitor **5** results in substantial gains in binding in both cases, however, chemical DMC analysis clearly reveals that they antagonize each other. If the two groups contributed linearly to the total binding energy, one would expect molecule **8** to exhibit an IC_{50} value of 3.9 nM or a ΔG_b value of -12.4 kcal mol^{-1}. The observed value is 10-fold less favorable (1.4 kcal mol^{-1}).

We hypothesized that the most likely explanation for this mutually negative effect on binding for the two extensions of pyridine **5** was an alternate binding mode for the biaryl inhibitor **6** in which the thiazole does not occupy the same position as the phenyl group in compound **4** (Figure 1). Previous studies involving an analogous biphenyl structure on the P2 Pro residue revealed such an alternate binding orientation.[12] Therefore, NMR studies were undertaken to determine the protease-bound conformation of the thiazolylpyridine ligand **9**.[13] Using distance constraints derived from the cross-peak volumes of transferred NOE (nuclear Overhauser effect) data, molecular modeling was employed to obtain a structure of bound inhibitor **9** (Figure 2). The results of the NMR studies are fully consistent with a different binding mode in which the thiazole ring is oriented towards the N-terminal end of the molecule rather than in the position occupied by the phenyl group of the P2-aryoxy substituent of inhibitor **4** (as shown in Figure 1). In this case, the observed *antagonism* is a result of a change in binding mode.

Finally, an example is provided in which two functionalities, the P2 phenyl and the terminal acyl sulfonamide, contribute to the free energy of binding in a linearly *additive* fashion, the ΔG_b for inhibitor **13** being only 0.1 kcal mol^{-1} more favorable than the value predicted from the single additions of the groups to compound **10** (Scheme 3).[14] While such cases, where structural modification at a particular position on a ligand leads to changes in binding that are quantitatively paralleled in analogous compounds that differ at remote subsites, are often observed, this is far from the rule. Nonetheless, the assumption of additivity[15] is often the only reasonable approach when trying to rationally optimize inhibitor binding, and has long been the mindset of medicinal chemists.

In spite of great advances in our understanding of biomolecular recognition processes, the optimization of the binding affinity of small molecules to biological receptors remains a challenging process. The use of the thermodynamic cycles described herein highlights the relevance of secondary interactions. When coupled to structural studies, the nonlinear SAR

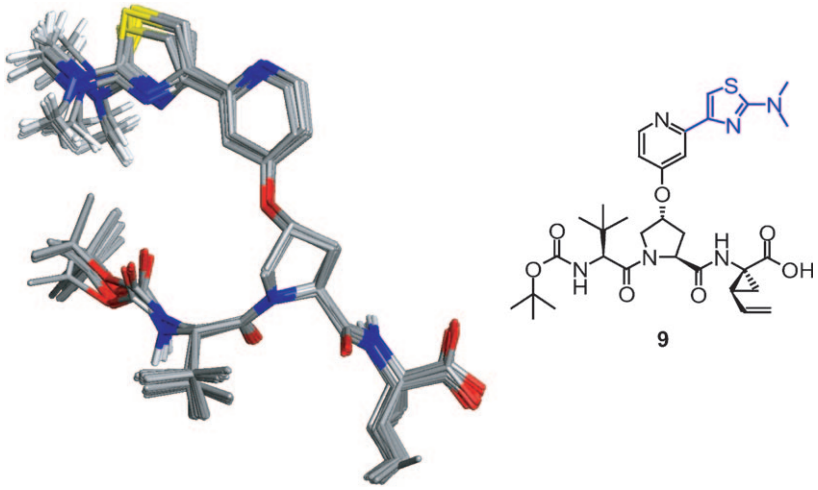
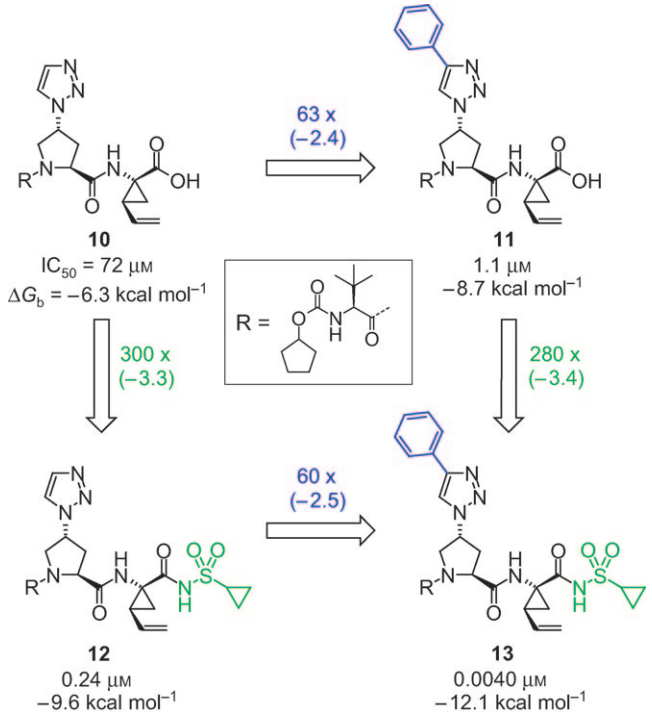


Figure 2. Structure of HCV protease-bound inhibitor **9** as determined by transferred NOE.



Scheme 3. Chemical DMC used to evaluate the relationship between the C-terminal acyl sulfonamide and P2 phenyl substituent of the tripeptide **13**. Numbers next to arrows indicate fold improvement in IC_{50} values with the $\Delta\Delta G_b$ (kcal mol^{-1}) in parentheses.

data may be understood in terms of conformational effects, alternate binding modes, chelate effects, etc. This can guide subsequent inhibitor optimization or identify new opportunities, taking advantage of novel binding orientations. Chemical DMCs are a useful tool in dissecting biomolecular recognition phenomena. We continue to utilize both chemical DMCs as well as TMCs (chemical triple-mutant cycles) to rationalize and further our optimization of the binding of small molecules to biological targets of therapeutic importance.

Experimental Section

Characterization of the inhibitors is provided in the Supporting Information. Binding energies (ΔG_b) are the free energy changes calculated using $\Delta G_b = -RT\ln(1/K_i)$ where K_i is the inhibition constant obtained from in vitro enzymatic assays. K_i values were calculated from IC_{50} values (average of at least four determinations) using the equation: $IC_{50} = 0.5\,[E_{total}] + K_i\,(1 + [S]/K_M)$. In the case of compounds **1**–**8**, a radiometric assay[16] was employed, whereas a fluorogenic assay[16] was used for inhibitors **10**–**13**. The protease-bound structure of **9** was determined by transferred NOE methods as described.[12] The model of the inhibitor **4**–HCV protease complex was generated from the co-crystal structure of a closely related complex.[17] Minor changes were made to replace the original co-crystallized substrate to inhibitor **4**, which was then docked and energy minimized using the MOE software (mmff94 s mod.ff forcefield, solvation included). The structure of inhibitor **4** in the complex obtained was fully consistent with transferred NOE data for bound inhibitor **4**, as well as published structures of similar complexes.

Acknowledgements

We thank D. Thibeault for carrying out the enzymatic assays, and Dr. P. Bonneau for helpful discussions.

Keywords: antiviral agents · structure–activity relationships · inhibitors · chemical double-mutant cycles · biomolecular recognition

[1] a) S. L. Cockroft, C. A. Hunter, *Chem. Soc. Rev.* **2007**, *36*, 172–188; b) H. Adams, F. J. Carver, C. A. Hunter, J. C. Morales, E. M. Seward, *Angew. Chem.* **1996**, *108*, 1628–1631; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1542–1544.

[2] a) F. R. Fischer, W. B. Schweizer, F. Diederich, *Angew. Chem.* **2007**, *119*, 8418–8421; *Angew. Chem. Int. Ed.* **2007**, *46*, 8270–8273; b) F. Hof, D. M. Scofield, W. B. Schweizer, F. Diederich, *Angew. Chem.* **2004**, *116*, 5166–5169; *Angew. Chem. Int. Ed.* **2004**, *43*, 5056–5059.

[3] We use the terms "antagonism" and "synergism" to denote the nonadditive contributions by substructures within a ligand to avoid the more widely used terms "negative-"and "positive cooperativity". See D. H. Williams, N. L. Davies, R. Zerella, B. Bardsley, *J. Am. Chem. Soc.* **2004**, *126*, 2042–2049.

[4] R. De Francesco, G. Migliaccio, *Nature* **2005**, *436(7053)*, 953–960.

[5] a) M. Llinàs-Brunet, M. Bailey, G. Fazal, S. Goulet, T. Halmos, S. Laplante, R. Maurice, M. Poirier, M.-A. Poupart, D. Thibeault, D. Wernic, D. Lamarre, *Bioorg. Med. Chem. Lett.* **1998**, *8*, 1713–1718; b) C. Steinkuhler, G. Biasiol, M. Brunette, A. Urbani, U. Koch, R. Cortese, A. Pessi, R. De Francesco, *Biochemistry* **1998**, *37*, 8899–8905.

[6] M. Llinàs-Brunet, M. Bailey, G. Fazal, E. Ghiro, V. Gorys, S. Goulet, T. Halmos, R. Maurice, M. Poirier, M.-A. Poupart, J. Rancourt, D. Thibeault, D. Wernic, D. Lamarre, *Bioorg. Med. Chem. Lett.* **2000**, *10*, 2267–2270.

[7] I. Schechter, A. Berger, *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157–162 describes the peptide/protease subsite nomenclature used, where

the amino acid residues are denoted …P3–P2–P1–P1′–P2′.., and the scissile bond is situated between P1 and P1′.

[8] J. Rancourt, D. R. Cameron, V. Gorys, D. Lamarre, M. Poirier, D. Thibeault, M. Llinàs-Brunet, *J. Med. Chem.* **2004**, *47*, 2511–2522.

[9] a) D. Lamarre, P. C. Anderson, M. Bailey, P. Beaulieu, G. Bolger, P. Bonneau, M. Boes, D. R. Cameron, M. Cartier, M. G. Cordingley, A.-M. Faucher, N. Goudreau, S. H. Kawai, G. Kukolj, L. Lagacé, S. R. LaPlante, H. Narjes, M.-A. Poupart, J. Rancourt, R. E. Sentjens, R. St. George, B. Simoneau, G. Steinmann, D. Thibeault, Y. S. Tsantrizos, S. M. Weldon, C.-L. Yong, M. Llinàs-Brunet, *Nature* **2003**, *426*, 186–189; b) H. Hinrichsen, Y. Benhamou, H. Wedemeyer, M. Reiser, R. E. Sentjens, J. L. Calleja, X. Forns, A. Erhardt, J. Crönlein, R. L. Chaves, C.-L. Yong, G. Nehmiz, G. G. Steinmann, *Gastroenterology* **2004**, *127*, 1347–1355.

[10] a) M. Llinàs-Brunet, M. D. Bailey, E. Ghiro, V. Gorys, T. Halmos, M. Poirier, J. Rancourt, N. Goudreau, *J. Med. Chem.* **2004**, *47*, 6584 –6594; b) F. Orvieto, Federica, U. Koch, Uwe, V. G. Matassa, E. Muraglia, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 2745–2748; c) R. Roenn, T. Gossas, Y. A. Sabnis, H. Daoud, E. Aakerblom, U. H. Danielson, A. Sandstroem, *Bioorg. Med. Chem.* **2007**, *15*, 4057–4068; d) G. Dahl, A. Sandstrom, E. Akerblom, U. H. Danielson, *FEBS J.* **2007**, *274*, 5979–5986.

[11] *Unpublished results*.

[12] N. Goudreau, D. Cameron, P. Bonneau, V. Gorys, C. Plouffe, M. Poirier, D. Lamarre, M. Llinàs-Brunet, *J. Med. Chem.* **2004**, *47*, 123–132.

[13] This close analogue of compound **6** was used due to its better solubility and peak separation.

[14] J. Naud, C. Lemke, N. Goudreau, E. Beaulieu, P. D. White, M. Llinàs-Brunet, P. Forgione, *Bioorg. Med. Chem. Lett.* **2008**, *18*, 3400–3404.

[15] S. M. Free, J. W. Wilson, *J. Med. Chem.* **1964**, *7*, 395–399.

[16] Y. S. Tsantrizos, G. Bolger, P. Bonneau, D. R. Cameron, N. Goudreau, G. Kukolj, S. R. LaPlante, M. Llinàs-Brunet, H. Nar, D. Lamarre, *Angew. Chem.* **2003**, *115*, 1394–1398; *Angew. Chem. Int. Ed.* **2003**, *42*, 1356–1360 (Supporting Information).

[17] A. Pause, G. Kukolj, M. Bailey, M. Brault, F. Do, T. Halmos, L. Lagacé, R. Maurice, M. Marquis, G. McKercher, C. Pellerin, L. Pilote, D. Thibeault, D. Lamarre, *J. Biol. Chem.* **2003**, *278*, 20374–20380.

Received: July 8, 2008
Revised: August 13, 2008
Published online on October 22, 2008

