Calculated and Experimental Low-Energy Conformations of Cyclic Urea HIV Protease Inhibitors

C. Nicholas Hodge,*,‡ Patrick Y. S. Lam,‡ Charles J. Eyermann,‡ Prabhakar K. Jadhav,‡ Y. Ru,‡ Christina H. Fernandez,‡ George V. De Lucca,‡ Chong-Hwan Chang,‡ Robert F. Kaltenbach III,‡ Edward R. Holler,‡ Francis Woerner,‡ Wayne F. Daneker,‡ George Emmett,‡ Joseph C. Calabrese,† and Paul E. Aldrich‡

Contribution from the Chemical and Physical Sciences Department, Research Division, DuPont Merck Pharmaceutical Company, P.O. Box 80500, Wilmington, Delaware 19880-0500, and Central Research Department, DuPont Company, Wilmington, Delaware 19880-0328

Received July 14, 1997

Abstract: One important factor influencing the affinity of a flexible ligand for a receptor is the internal strain energy required to attain the bound conformation. Calculation of fully equilibrated ensembles of bound and free ligand and receptor conformations are computationally not possible for most systems of biological interest; therefore, the qualitative evaluation of a novel structure as a potential high-affinity ligand for a given receptor can benefit from taking into account both the bound and unbound (usually aqueous) low-energy geometries of the ligand and the difference in their internal energies. Although many techniques for computationally generating and evaluating the conformational preferences of small molecules are available, there are a limited number of studies of complex organics that compare calculated and experimentally observed conformations. To assess our ability to predict a priori favored conformations of cyclic HIV protease (HIV-1 PR) inhibitors, conformational minima for nine 4,7-bis(phenylmethyl)-2*H*-1,3-diazepin-2-ones **I** (cyclic ureas) were calculated using a high temperature quenched dynamics (QD) protocol. Single crystal X-ray and aqueous NMR structures of free cyclic ureas were obtained, and the calculated low-energy conformations compared with the experimentally observed structures. In each case the ring conformation observed experimentally is also found in the lowest energy structure of the QD analysis, although significantly different ring conformations are observed at only slightly higher energy. The 4- and 7-benzyl groups retain similar orientations in calculated and experimental structures, but torsion angles of substituents on the urea nitrogens differ in several cases. The data on experimental and calculated cyclic urea conformations and their binding affinities to HIV-1 PR are proposed as a useful dataset for assessing affinity prediction methods.

Introduction

The search for inhibitors of HIV protease with clinically useful antiviral properties has led to the discovery of a variety of highly potent compounds.[1] We recently described[2] a novel series of HIV protease inhibitors that were designed for specific binding to retroviral, as opposed to mammalian, aspartyl proteases by displacing a structural water molecule found only in retroviral proteases,[3] and to provide a compact, rigid structure that optimized tight binding interactions with the enzyme while minimizing unnecessary molecular weight and conformational freedom. X-ray crystallographic[2] and NMR[4] studies of the complexes have demonstrated that these compounds do in fact bind to HIV protease in the expected manner. Since our initial report we have identified cyclic ureas that are highly potent, selective, orally available inhibitors of a broad set of retroviral proteases.[5,6] As with other HIV protease inhibitors, even those now in clinical use,[7] virus has been identified with decreased sensitivity to first-generation cyclic ureas,[8] suggesting that the compounds may be of limited benefit over the long term. Our recent focus has been to apply structure-based design and medicinal chemistry to identify cyclic urea analogues with high

† Central Research Department, DuPont Company.

‡ Research Division, DuPont Merck Pharmaceutical Company.

(1) West, M. L.; Fairlie, D. P. *Trends Pharmacol. Sci.* **1995**, *16*, 67−74.

(2) Lam, P. Y. S.; Jadhav, P. K.; Eyermann, C. J.; Hodge, C. N.; Ru, Y.; Bacheler, L. T.; Meek, O. M. J.; Rayner, M. M.; Wong, Y. N.; Chang, C. H.; Weber, P. C.; Jackson, D. A.; Sharpe, T. R.; Erickson-Viitanen, E. V. *Science* **1994**, *263*, 380−384.

(3) Miller, M.; Schneider, J.; Sathyanarayana, B. K.; Toth, M. V.; Marshall, G. R.; Clawson, L.; Selk, L.; Kent, S. B. H.; Wlodawer, A. *Science* **1989**, *246*, 1149−1152.

(4) (a) Yamazaki, T.; Nicholson, L. K.; Wingfield, P.; Stahl, S. J.; Kaufman, J. D.; Eyermann, C. J.; Hodge, C. N.; Lam, P. Y. S.; Torchia, D. A.; et. al. *J. Am. Chem. Soc.* **1994**, *116*, 10791−10792. (b) Grzesiek, S.; Bax, A.; Nicholson, L. K.; Yamazaki, T.; Wingfield, P.; Stahl, S. J.; Eyermann, C. J.; Torchia, D. A.; Hodge, C. N.; Lam, P. Y.; Jadhav, P. K.; Chang, C. H. *J. Am. Chem. Soc.* **1994**, *116*, 1581−2. (c) Nicholson, L. K.; Yamazaki, T.; Torchia, D. A.; Grzesiek, S.; Bax, A.; Stahl, S. J.; Kaufman, J. D.; Wingfield, P. T.; Lam, P. Y. S.; Jadhav, P. K. Hodge, C. N.; Domaille, P. J.; Chang, C. H. *Nat. Struct. Biol.* **1995**, 2, 274−280.

(5) (a) Hodge, C. N.; Aldrich, P. E.; Bacheler, L. T.; Chang, C.-H.; Eyermann, C. J.; Garber, S.; Grubb, M.; Jackson, D. A.; Jadhav, P. K.; Korant, B.; Lam, P. Y. S.; Maurin, M. B.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; Reid, C.; Sharpe, T. R.; Shum, L.; Winslow, D. L.; Erickson-Viitanen, S. *Chem. Biol.* **1996**, *3*, 301−315. (b) Lam, P.; Ru, Y.; Jadhav, P. K.; Aldrich, P. E.; DeLucca, G. V.; Eyermann, C. J.; Chang, C. H.; Emmett, G.; Holler, E. R.; Daneker, W. F.; Li, L.; Confalone, P. N.; McHugh, R. J.; Han, Q.; Li, R.; Markwalder, J. A.; Seitz, S. P.; Sharpe, T. R.; Bacheler, L. T.; Rayner, M. R.; Klabe, R. M.; Shum, L.; Winslow, D. L.; Kornhauser, D. M.; Jackson, D. A.; Erickson-Viitanen, S.; Hodge, C. N. *J. Med. Chem.* **1996**, *39*, 3514−3525. (c) Nugiel, D. A.; Jacobs, K.; Worley, T.; Patel, M.; Kaltenbach, R. F. I.; Meyer, D. T.; Jadhav, P. K.; De Lucca, G. V.; Smyser, T. E.; Klabe, R. M.; Bacheler, L. T.; Rayner, M. M.; Seitz, S. P. *J. Med. Chem.* **1996**, *39*, 2156−2169.

S0002-7863(97)02357-3 CCC: $15.00
Published on Web 05/02/1998

potency against resistant variants that retain the favorable physicochemical properties of the initial series.[9]

In this project, as in most drug discovery efforts, once a new series of biologically active ligands with adequate in vitro properties is identified it is necessary to optimize the physical properties of the lead compound to obtain desired in vivo properties for clinical use; these properties include oral bioavailability, safety, stability, and ease of synthesis and formulation. In the process of modulating physical properties the target activity of the lead series is frequently lost or diminished, resulting in a time-consuming cycle of analogue synthesis and testing. Traditionally this process has been successfully addressed by applying medicinal chemistry and quantitative structure activity relationship (QSAR) analyses, but these methods do not adequately take into account the effect of modifications of the ligand on its three-dimensional shape, either free or bound to the enzyme. Computational methods that provide even a crude estimate of the binding affinity of suggested novel scaffolds and functional groups for HIV-1 PR can assist medicinal chemists in prioritizing the large number of possible candidate ligands conceived in their effort to balance potency against the target receptor with the required physical properties.[10]

Considerable progress toward this end—the qualitative assessment of the affinity of a hypothetical ligand for a receptor—has been made by the development of 3D QSAR and Molecular Field Analysis models,[11] wherein the activities of known analogues are used to infer a set of binding determinants or a receptor field, and the activity of a new analogue is predicted based on interactions with that field. Docking,[12] with and without flexibility of the ligand and the protein, has been used to evaluate modifications of known inhibitors and to screen three-dimensional structure databases for novel ligands. Free energy perturbation methods have been used to rationalize observed binding affinity changes in stereoisomeric ligands or isosteric replacements of small groups.[13] De novo design,[14] in which the active site of the receptor is used as a negative mold for the unbiased construction of novel ligands and fragments are fitted into the active site with fast scoring algorithms and in some cases connected to form potent inhibitors, has also been applied to HIV protease. Developing quantitative structure−activity data from both ligand and macromolecular structure has yielded interesting results in HIV protease[15] and phospholipase A_2.[16]

A less sophisticated but more rapid approach is to consider a potent, selective ligand with a defined manner of binding to a known receptor as a positive mold and to screen novel proposed ligands by their ability to attain the shape and electrostatic character of this known reference compound:[17] essentially a molecular field analysis in which the "field" is defined by the shape of a single bound ligand of known high affinity. If a proposed new compound cannot project the important recognition elements into the correct regions of space in an energetically favorable fashion, i.e., with internal strain energy at or near its global minimum, then it is rejected as a synthetic target. If the required shape is found among a large number of other low-energy conformations that do not provide a good match to the reference compound, then the target is intermediate in quality; and if the proposed structure displays a strong energetic bias toward the required shape, then it is, relatively speaking, an attractive candidate for experimental validation.[18]

In practice, many flexible compounds that are proposed as mimetics of a ligand whose binding conformation is known do

(6) (a) Erickson-Viitanan, S.; Klabe, R. M.; Cawood, P. G.; O'Neal, P. L.; Meek, J. L. *Antimicrob. Agents Chemother.* **1994**, *38*, 1628−1634. (b) Grubb, M. F.; Wong, Y. N.; Burcham, D. L.; Saxton, P. L.; Quon, C. Y.; Huang, S. M. *Drug Metab. Dispos*. **1994**, *22*, 709−712. (c) Rayner, M. M.; Cordova, B. C.; Meade, R. P.; Aldrich, P. E.; Jadhav, P. K.; Ru, Y.; Lam, P. Y. S. *Antimicrob. Agents Chemother*. **1994**, *38*, 1635−1640. (d) Wong, Y. N.; Burcham, D. L.; Saxton, P. L.; Erickson-Viitanen, S.; Grubb, M. F.; Quon, C. Y.; Huang, S.-M. *Biopharm. Drug Dispos*. **1994**, *15*, 535-44. (e) Ottto, M. J.; Reid, C. D.; Garber, S.; Lam, P. Y. S.; Scarnati, H.; Bacheler, L. T.; Rayner, M. M.; Winslow, D. L. *Antimicrob. Agents Chemother*. **1993**, *37*, 2606−2611.

(7) For reviews see: Lin, J. H. *Adv. Drug Delivery Rev.* **1997**, *27(2, 3)*, 774−778. Vacca, J. P.; Condra, J. H. *Drug Discovery Today* **1997**, **2**(7), 261−272.

(8) (a) Otto, M. J.; Garber, S.; Winslow, D. L.; Reid, C. D.; Aldrich, P.; Jadhav, P. K.; Patterson, C. E.; Hodge, C. N.; Cheng, Y. S. E. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 7543−7547. (b) Reference 6a. (c) Winslow, D. L.; Anton, E. D.; Horlick, R. A.; Zagursky, R. J.; Tritch, R. J.; Scarnati, H.; Ackerman, K.; Bacheler, L. T. *Biochem. Biophys. Res. Commun*. **1994**, *205*, 1651−1657. (d) Winslow, D. L.; Stack, S.; King, R.; Scarnati, H.; Bincsik, A.; Otto, M. J. *AIDS Res. Hum. Retroviruses* **1995**, *11*, 107−113. (e) King, R. W.; Garber, S.; Winslow, D. L.; Reid, C.; Bachelor, L. T.; Anton, E.; Otto, M. J. *Antiviral Chem. Chemother.* **1995**, *6*, 80−88.

(9) (a) Ala, P.; Huston, E.; Klabe, R.; McCabe, D.; Duke, J.; Rizzo, C.; Korant, B.; DeLoskey, R.; Lam, P. Y. S.; Hodge, C. N.; Chang, C.-H. *Biochemistry* **1997**, *36*, 6(7), 1573−1580. (b) Jadhav, P. K.; Ala, P.; Woerner, F. J.; Chang, C.-H.; Garber, S. S.; Anton, E. D.; Bacheler, L. T. *J. Med. Chem.* **1997**, *40*, 181−191.

(10) (a) Lybrand, T. P.; Kollman, P. A. *J. Chem. Phys.* **1985**, *83*, 2923−2933. (b) Kuntz, I. D. *Science* **1992**, *257(5073)*, 1078−1982. (c) Crippen, G. M. *J. Comput. Chem.* **1995**, *16*, 486−500. (d) Ajay, Murcko, M. *J. Med. Chem.* **1995**, *38*, 4953−4967.

(11) (a) Cho, S. J.; Tropsha, A. *J. Med. Chem.* **1995**, *38*, 1060−6. (b) Klebe, G.; Abraham, U. *J. Med. Chem.* **1993**, *36*, 70−80. (c) Yamada, M.; Itai, A. *Chem. Pharm. Bull.* **1993**, *41*, 1203−1205. (d) Cruciani, G.; Watson, K. A. *J. Med. Chem.* **1994**, *37*, 2589−2601.

(12) (a) Meng, E. C.; Gschwend, D. A.; Blaney, J. M.; Kuntz, I. D. *Proteins: Struct., Funct., Genet.* **1993**, *17*, 266−278. (b) Stoddard, B. L.; Koshland, D. E., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1146−1153. (c) Yamada, M.; Itai, A. *Chem. Pharm. Bull.* **1993**, *41*, 1203−5. (d) Cherfils, J.; Janin, J. *Curr. Opin. Struct. Biol*. **1993**, *3*, 265−269. (e) Di Nola, A.; Roccatano, D.; Berendsen, H. J. C. *Proteins: Struct., Funct., Genet*. **1994**, *19*, 174−82. (f) Goodsell, D. S.; Olson, A. J. *Proteins: Struct., Funct., Genet.* **1990**, *8*, 195−202. (f) Goodsell, D. S.; Lauble, H.; Stout, C. D.; Olson, A. J. *Proteins: Struct., Funct., Genet.* **1993**, *17*, 1−10. (g) Kearsley, S. K.; Underwood, D. J.; Sheridan, R. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 565−582. (h) Luty, B. A.; Wasserman, Z. R.; Stouten, P. F. W.; Hodge, C. N.; Zacharias, M.; McCammon, J. A. *J. Comput. Chem.* **1995**, *16*, 454−464. (i) Guarnieri, F.; Wilson, S. R. *J. Comput. Chem.* **1995**, *16*, 648−653. (j) Wasserman, Z. R.; Hodge, C. N. *Proteins: Struct., Funct., Genet.***1996**, *24*, 227−237.

(13) (a) Cieplak, P.; Kollman, P. A. *J. Comput.-Aided Mol. Des*. **1993**, *7*(3), 291−304. (b) DeBolt, S. E.; Pearlman, D. A.; Kollman, P. A. *J. Comput. Chem.* **1994**, *15*, 351−373. (c) Ferguson, D. M.; Radmer, R. J.; Kollman, P. A. *J. Med. Chem.* **1991**, *34*, 2654−2659. (d) Hansson, T.; Aaqvist, J. *Protein Eng.* **1995**, *8*, 1137−1144. (e) Rao, B. G.; Murcko, M. A. *J. Comput. Chem.* **1994**, *15*, 1241−1253. (f) Reddy, M. R.; Viswanadhan, V. N.; Weinstein, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 10287−10291. (g) Reddy, M. R.; Varney, M. D.; Kalish, V.; Viswanadhan, V. N.; Appelt, K. *J. Med. Chem.* **1994**, *37*, 1145−1152. (h) Varney, M. D.; Appelt, K.; Kalish, V.; Reddy, M. R.; Tatlock, J.; Palmer, C. L.; Romines, W. H.; Wu, B.-W.; Musick, L. *J. Med. Chem.* **1994**, *37*, 2274−2284.

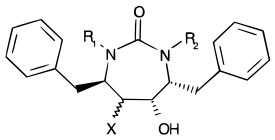
(14) (a) Clark, T. D.; Ghadiri, M. R. *J. Am. Chem. Soc*. **1995**, *117*, 12364−12365. (b) Pearlman, D. A.; Murcko, M. A. *J. Comput. Chem.* **1993**, *14*, 1184−1193. (c) Baca, M.; Alewood, P. F.; Kent, S. B. H. *Protein Sci.* **1993**, *2*, 1085−1091. (d) *ibid. idem*., 623−32. (e) Boehm, H.-J. *J. Comput.-Aided Mol. Des*. **1994**, *8*, 243−256. (f) Rotstein, S. H.; Murcko, M. A. *J. Med. Chem.* **1993**, *36*, 1700−1710. (g) Bures, M. G.; Hutchins, C. W.; Maus, M.; Kohlbrenner, W.; Kadam, S.; Erickson, J. W. *Tetrahedron Comput. Methodol.* **1990**, *3*, 673−680. (h) Roe, D. C.; Kuntz, I. D. *J. Comput.-Aided Mol. Des*. **1995**, *9*, 269−282. (i) Waszkowycz, B.; Clark, D. E.; Frenkel, D.; Li, J.; Murray, C. W.; Robson, B.; Westhead, D. R. *J. Med. Chem.* **1994**, *37*, 3994−4002.

(15) (a) Verkhivker, G. M.; Rejto, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 60−64. (b) Verkhivker, G.; Appelt, K.; Freer, S. T.; Villafranca, J. E. *Protein Eng*. **1995**, *8*, 677−691.

(16) Ortiz, A. R.; Pisabarro, M. T.; Gago, F.; Wade, R. C. *J. Med. Chem.* **1995**, *38*, 2681−2691.

(17) Hodge, C. N.; Straatsma, T. P.; McCammon, J. A.; Wlodawer, A. Rational Design of HIV Protease Inhibitors. In *Structural Biology of Viruses*; Chiu, W., Burnett, R. M., Garcea, R., Eds.; Oxford University Press: New York, 1997; pp 451−473.

Table 1



I

entry	R_1	R_2	X	Ki,[a] nm
1A	H	H	(*S*)-OH	4500
1B	methyl	methyl	(*S*)-OH	5700
1C	cyclopropylmethyl	cyclopropylmethyl	(*S*)-OH	1.9
1D	cyclopropylmethyl	cyclopropylmethyl	(*R*)-OH	1.1
1E	(2-naphthyl)methyl	H	(*S*)-OH	2.8
1F	4-fluorobenzyl	4-fluorobenzyl	(*S*)-OH	1.4
1G	(2-naphthyl)methyl	2-naphthylmethyl	(*S*)-OH	0.23
1H	3-aminobenzyl	3-aminobenzyl	(*S*)-OH	0.25
1H·HCl	3-aminobenzyl·$(HCl)_2$	3-aminobenzyl·$(HCl)_2$	(*S*)-OH	0.25
1I	H	H	(*R*)-OAc	$>10^3$
1J	allyl	allyl	(*S*)-OH	4.7

[a] Inhibition constant vs HIV-1 PR (see refs 2 and 5).

not attain a unique low-energy geometry that overlays well with the target shape, and considerable efficiency is added to the discovery process if these compounds are screened out prior to synthesis.[19] If desired, the ligand can then be evaluated further by modeling other factors contributing to enzyme affinity, such as the enthalpy of interaction with the receptor and of receptor reorganization on ligand binding; the free energy of desolvation of ligand; the free energy of de- or resolvation of the receptor site, if the new ligand is of different size than the reference ligand; and the entropic cost of restricting enzyme and ligand to their bound conformations.[20] If net energetic estimates are favorable relative to the reference ligand, then it is prioritized as a synthetic target based on these considerations and other factors such as synthetic accessibility and the likelihood of improved physical properties. But the key is that the initial shape evaluation is fast relative to synthesis and useful for crude screening of diverse novel structures for steric fit.

In all of the methods described above, irrespective of whether an explicit (e.g., crystal coordinates) or an implicit (e.g., CoMFA) receptor field is evaluated, an accurate evaluation of low-energy ligand conformations is required. This in turn requires a method of searching conformational space and evaluating the energies of stable conformations and is a slow- or rate-determining step in many procedures that take ligand conformational flexibility and internal strain energy into account. The importance of ligand flexibility in determining binding free energy is addressed in an recent study by Vajda and co-workers.[21] A more detailed discussion of other methods of determining ligand conformations is provided below.

The purpose of this report is to (a) describe parameters for a quenched dynamics simulation[22] that rapidly generates low-energy conformations of cyclic urea HIVPR inhibitors; (b) detail the synthesis, X-ray, and ^{1}H NMR structures of a series of these structures; and (c) compare the calculated structures with experimental conformations of both enzyme-bound and free cyclic ureas. Our results indicate that, with some caution, the force fields and protocols we describe are useful as one of several necessary steps in estimating ligand binding affinities.[23] In addition, our previously reported structures of several cyclic ureas bound to HIV protease are or will be available to the public via the Brookhaven Protein Data Bank: complexes related to this study include XK263, DMP450, XK216, and DMP323 bound to wild type and several mutants (refs 2 and 5). These combined data on affinity constants, calculated structures, and experimental structures of free ligands and complexes are proposed as useful known controls for testing methods of ligand-protein docking and affinity prediction.

Methods

Chemistry. The synthesis of the protected precursors to substituted cyclic ureas have been published.[5a,b] The compounds used in this study, shown in Table 1, were synthesized as described in the Experimental Section: treatment of urea **7** with various alkylating agents in the presence of sodium hydride in dry dimethylformamide, followed by aqueous workup, removal of the protecting groups, and purification yielded **1C** and **1E**−**1G**. Oxidation of **1A** to the hydroxyketone and reduction to the *R*,*S*,*R*,*R* diol provided **1D**; the absolute stereochemistry is confirmed by the X-ray crystal structure (see Supporting Information). Details of the synthesis will be described in a subsequent manuscript. **1H** and **1I** were prepared as described in ref 5a, and **1B** was prepared as described in ref 5b.

Single-Crystal X-ray Analysis. Structural characterization of compounds **1B**−**1I** and **1H**·$(HCl)_2$ were carried out using

(18) A fundamental assumption in implicit in this approach: structures that can obtain the binding geometry only at the expense of significant internal strain energy are not of primary interest as targets. Several arguments can be made to counter this assumption: (a) there are unpublished indications that many ligands in the Brookhaven Protein Data Bank appear to retain significant internal strain energy on binding; (b) the possibility that the net effect of binding to the receptor may be to stabilize ligand conformations that are less favorable in the unbound state; and (c) a preorganized ligand may not have a kinetic pathway to bind to its target. However, in the *prospective* design process, in which many good ideas must be prioritized into a few synthetic targets, it is a reasonable starting point to expect an energetic advantage for a ligand whose low energy conformation in the ambient milieu (in our case, aqueous pH 5.5 buffer) is very close to its optimal bound conformation in the target receptor.

(19) Yamada, M.; Itai, A. *Chem. Pharm. Bull.* **1993**, *41*, 1203−1205.

(20) Ajay; Murcko, M. A. *J. Med. Chem.* **1995**, *38*, 4953−67.

(21) Vajda, S.; Wheng, Z.; Rosenfeld, R. *Biochemistry* **1994**, *33*, 13977−13988.

(22) (a) Al-Obeidi, F.; Hadley, M. E.; Pettitt, B. M.; Hruby, V. J. *J. Am. Chem. Soc.* **1989**, *111*, 3413−3416. (b) O'Connor, S. D.; Smith, P. E.; Al-Obeidi, F.; Pettitt, B. M. *J. Med. Chem.* **1992**, *35*, 2870−2881.

(23) During preparation of this manuscript Hulten et al. reported on the behavior of DMP323 and some other P1/P1′ cyclic urea and cyclic sulfonamide analogues in molecular dynamics simulations in water and in the presence of HIV protease. A full conformational search was not undertaken nor are free ligand geometries analyzed experimentally, which is the focus of our research. The results do show weak correlation of calculated and experimental affinities to HIV-1 protease using the GROMOS force field. Hulten, J.; Bonham, N. M.; Nillroth, U.; Hansson, T.; Zuccarello, G.; Bouzide, A.; Aaqvist, J.; Classon, B.; Danielson, H.; Karlen, A.; Kvarnstrom, I.; Samuelsson, B.; Hallberg, A. *J. Med. Chem.* **1997**, *40*, 0(6), 885−897.

standard X-ray crystallographic techniques. Complete reports including crystallization conditions, atomic coordinates, thermal parameters, and interatomic distances and angles are available as Supporting Information.

Solution Structure Analysis. NMR spectra were taken on a solution containing approximately 30 mg of sample dissolved in 0.8 mL of the appropriate solvent. Proton and carbon chemical shifts were referenced to external TSP (TSP = sodium 2,2,3,3-d_4-trimethylsilylpropionate, D2O solvent) or the relevant organic solvent peaks. ^{1}H chemical shift and coupling constant data were derived from spin simulation calculations when needed. All spectra were obtained at 30 °C.

^{1}H and ^{13}C spectra were taken on either a Varian VXR-400S or a Varian Unity-400 NMR spectrometer (Palo Alto, CA) operating at 399.95 and 100.59 MHz, respectively. A standard switchable probe was used for the one-dimensional ^{13}C; all other data were acquired using an inverse detection probe equipped with a z-gradient coil. ^{1}H spectrum were recorded with a digital resolution of 0.25 Hz/pt (AT = 4.06s), a tip angle of 30°, and a relaxation delay of 2.0 s. ^{13}C spectra were recorded with a digital resolution of 0.76 Hz/pt (AT = 1.31 s), a 60° tip angle, and a 4.0 s relaxation delay. ^{1}H−^{13}C correlation spectra were obtained using a value of 140 Hz for $^1J_{C-H}$ and 7 Hz for $^3J_{C-H}$. NOESY spectra were acquired using a 2.0 s presaturation pulse, a 3.0 s predelay, and a 1.0 s mixing time.

Ki Determination. Inhibition constants of compounds **1A**−**1I** were determined as described earlier (ref 6a). This selection of compounds is of interest as the inhibition constants span more than four logs, and it may provide a useful dataset for benchmarking affinity prediction methods (see Discussion).

Conformational Search. Many routines are available for using empirical force fields to find low-energy conformations of drug-sized molecules[24] (during this discussion "energy" will refer to the calculated internal strain energy of a single conformation in the gas phase). The method we employed has been described as high-temperature quenched dynamics by Pettit and co-workers[22] and employs molecular dynamic simulations at temperatures high enough to overcome conformational barriers and minimization at fixed intervals along the trajectory to yield representative low energy structures. Due to the limited degrees of freedom of the ligands, techniques to avoid local minima such as simulated annealing are not necessary. The method is rapid and, as described here and elsewhere, appears to provide broad sampling of conformational space provided that the conformational barriers are lower than kT.[25] This study does not attempt to evaluate the various methods or compare them to quenched dynamics; however, there are clearly restrictions on which protocols would perform at the level of accuracy required in a time frame consistent with screening reasonable numbers of candidate ligands. The requirements we had were speed, the ability to handle diverse functional groups without excessive parameter development, and sufficient accuracy to generate the experimentally observed conformations as among the lowest energy calculated conformations. Systematic search methods[26] with even fairly coarse dihedral grids are prohibitively slow since eight significant dihedrals as well as ring flexibility need to be evaluated in our system. More accurate semiempirical AM1 optimization in the gas phase is reasonably fast for single conformations, but multiple conformations would need to be sampled. Metropolis Monte Carlo[27] (MC) and distance geometry[28] with penalty functions, followed by energy minimization, would be as accurate as the method employed here; for these latter two and quenched dynamics the slow (minimization) step would be identical, so speed would depend on the number of structures generated that minimized to the same family (a deterministic systematic search algorithm that is optimized to reduce the number of minimizations was reported recently in ref 26c; the speed of this method on small ring systems has not yet been described). Another recent report describes an alternating Monte Carlo random sampling and stochastic dynamics (SD) simulation strategy for conformational searching that appears to be both fast and accurate[29] and allows the calculation of free energies from ensembles of conformations as well as the use of a continuum solvation model. The comparison of a quenched dynamics protocol with mixed MC/SD has not been carried out to our knowledge. Other methods have been compared and discussed in refs 24a−b. A promising algorithm for the complete analysis of conformational free energies, termed "mining minima", was also published recently.[30]

In our work, the Biosym consistent valence force field, cff91,[31] was used for molecular dynamics with standard atom parameters (parameters and input files are available as Supporting Information). This force field has succeeded in reproducing experimental geometries of small organic molecules in several studies.[32] Comparative studies with other empirical and semiempirical methods have also been reported.[33] The quenched dynamics simulation was run using Discover 2.9 in stand-alone mode on a Silicon Graphics Challenge 8 processor R4400 server as follows: time step of 1 fs, dielectric ϵ, ***k*** femtoseconds of simulation at ***T*** °C, minimization of every 1000th frame using 200 steps of steepest descents to RMSD of <1.0 kcal/Å, 500 steps of conjugate gradients to RMSD <0.1 kcal/Å, and 1000 steps of VA90A (modified Newton-Rapheson) to RMSD <0.01 kcal/Å. The minimized structure was archived,

(24) Discussions with leading references can be found in (a) Gundertofte, K.; Liljefors, T.; Norrby, P.-o.; Pettersson, I. *J. Comput. Chem.* **1996**, *17*, 429−449. (b) Treasurywala, A. M.; Jaeger, E. P.; Peterson, M. *J. Comput. Chem.* **1996**, *17*, 1171−1182. (c) Ngo, J. T.; Karplus, M. *J. Am. Chem. Soc.* **1997**, *119*, 5657−5667.

(25) (a) Cheng, Y. K.; Pettitt, B. M. *J. Am. Chem. Soc.* **1992**, *114*, 4465−4473. (b) Burgess, K.; Ho, K.-K.; Pettitt, B. M. *J. Am. Chem. Soc.* **1994**, *116*, 799−800. (c) Burgess, K.; Ho, K.-K.; Pal, B. *J. Am. Chem. Soc.* **1995**, *117*, 3808−3819. (d) Taber, D. F.; Christos, T. E.; Hodge, C. N. *J. Org. Chem.* **1996**, *61*, 2081−2084.

(26) (a) Smellie, A.; Kahn, S. D.; Teig, S. L. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 285−294. (b) Leach, A. R.; Kuntz, I. D. *J. Comput. Chem.* **1992**, *13*, 730−748. (c) Saunders: M.; Houk, K. N.; Wu, Y.-D.; Still, W. C.; Lipton, M. Chang, G.; Guida, W. C. *J. Am. Chem. Soc.* **1990**, *112*, 1419−1427.

(27) (a) Beveridge, D. L.; Mezei, M.; Mehrotra, P. K.; Marchese, F. T.; Thirumalai, V.; Ravi-Shanker, G. *Ann. N.Y. Acad. Sci.* **1981**, *367*, 108−131. (b) Chang, G.; Guida, W. C.; Still, W. C. *J. Am. Chem. Soc.* **1989**, *111*, 4379−4386. (c) Kolossvary, I.; Guida, W. C. *J. Comput. Chem.* **1993**, *14*, 691−698. (d) Kolossvary, I.; Guida, W. C. *J. Am. Chem. Soc.* **1993**, *115*, 2107−19. (e) Liwo, A.; Tempczyk, A.; Oldziej, S.; Shendrovich, M. D.; Hruby, V. J.; Talluri, S. *Biopolymers* **1996**, *38*, 157−175.

(28) Blaney, J. M.; Dixon, J. S. *Rev. Comput. Chem.* **1994**, *5*, 299−335.

(29) (a) McDonald, D. Q.; Still, W. C. *J. Am. Chem. Soc.* **1994**, *116*, 11550−11553. (b) McDonald, D. Q.; Still, W. C. *J. Org. Chem.* **1996**, *61*, 1385−1391.

(30) Head, M. H.; Given, J. A.; Gilson, M. K. *J. Phys. Chem. A* **1997**, *101*(8), 1609−1618.

(31) (a) Rizo, J.; Koerber, S. C.; Bienstock, R. J.; Rivier, J.; Hagler, A. T.; Gierasch, L. M. *J. Am. Chem. Soc.* **1992**, *114*, 2852−2859. (b) Kitson, D. H.; Avbelj, F.; Moult, J.; Nguyen, D. T.; Mertz, J. E.; Hadzi, D.; Hagler, A. T. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8920−8924. The Biosym suite of programs is available from MSI, San Diego, CA.

(32) (a) Ross, R. B.; Hockswender, T. R.; Wilson, C. A. *Comput. Polym. Sci.* **1995**, *5*, 203−212. (b) Liang, C.; Yan, L.; Hill, J. R.; Ewig, C. S.; Stouch, T. R.; Hagler, A. T. *J. Comput. Chem.* **1995**, *16*, 883−897. (c) Lee, C. H.; Zimmerman, S. S. *J. Biomol. Struct. Dyn.* **1995**, *13*, 201−218. (d) Ramek, M. *Int. J. Quantum Chem., Quantum Biol. Symp.* **1995**, 22, 75-81.

(33) (a) Knopp, B.; Jung, B.; Wortmann, F. J. *Macromol. Theory Simul.* **1996**, *5*, 947−956. (b) Rudnicki, W. R.; Lesyng, B. *Comput. Chem.* **1995**, *19*, 253−258.

Table 2. Effect of Frame Length, *k*, on Search Efficiency and Conformational Sampling ($m = 200$ Frames, $T = 2500$ °C, $\epsilon = 1$)

simulation time, fs per frame	lowest energy observed, kcal	CPU time, min[a]
1000	−3.8	71
100	−3.8	40
50	−3.3	38
25	0.19	43

[a] Four processors on eight processor Challenge, R4400 CPU.

and the dynamics trajectory continued at high temperature from 1001th frame, i.e., from the last *unminimized* conformation. This was repeated ***m*** times, for a total simulation time of ***k*******m***. The bis-*N*-allyl substituted cyclic urea **XK216 (1J,** Table 1),[2] not a member of the experimental versus calculated comparison set, was used to determine the effect of parameter variation on simulation performance as follows.

Effect of Parameter Variation on Low-Energy Search of XK216. Parameters ϵ, ***k***, ***m***, and ***T*** were examined to select conditions for optimum efficiency and breadth of conformational search. The same "random" seed value was used in all comparative runs.

A search of XK216 with $\epsilon = 1$ and $\boldsymbol{T} = 2500$ °C was run, minimizing every 1 ps ($\boldsymbol{k} = 1000$). After 21 ps the lowest energy observed was 3.3 kcal; after 130 ps the lowest energy was 3.8 kcal. No lower energy structures were observed after 4 ns, indicating either complete sampling or repeated searching of the same regions of conformational space. Inspection of the high-temperature trajectory files show extreme contortion of bond angles and lengths, suggesting that the latter possibility is unlikely. A value of $\boldsymbol{m} = 200$ was therefore chosen for standard searching. Extended searches of several of the compounds shown in Table 1 also indicated complete search within 150 ps.

The effect of the simulation time per frame, ***k***, is shown in Table 2. Based on these results standard searches were run with $\boldsymbol{k} = 1000$ fs to ensure complete search. The last three entries demonstrate, as expected, that minimization is considerably slower than the dynamics calculations.

The most efficient search occurred at the highest temperature examined (2500 °C), but the minimum-energy structure was located down to 1000 °C within 200 ps. At 600 °C and below, the minimum energy structure was not found even after 2 ns simulation. Following this study simulations were run at 2500 °C; in the case of cyclopropyl-substituted compounds, energy conservation failures occurred and simulations were carried out at 1500 °C.

These data point out that very high temperatures are needed in the quenched dynamics protocol in order to locate conformational minima of cyclic ureas, since temperatures below 1000 °C were unable to locate the lowest energy structure. In some cases, simulation temperatures below 1000 °C were insufficient for chair to chair conformational inversion.

The effect of the dielectric on the conformational search was examined by running the simulation conditions shown in Table 2, with $k = 1000$ fs, and at $\epsilon = 1$, 40, 80 and a distant dependent dielectric, 4**r*. The same families of structures were observed regardless of dielectric value, but the grouping of low-energy structures differed as well as the absolute energy values. For the test system described here the "best" clustering, i.e., the greatest spread between the lowest and those close in energy was observed with $\epsilon = 80$; this was chosen for subsequent simulations. This choice is consistent with the dielectric of the aqueous medium in which the assay takes place but less so with crystallization conditions. However, it does not have a significant effect on the ranking of conformations by energy; all of the calculations discussed below were run with $\epsilon = 1$, 40, and 80, and no significant change in the energetic ordering of conformations was observed. AM1 calculations (data not shown) and experimental data (see below) also suggest that solvent does not play a significant role in determining conformation in this series of compounds.

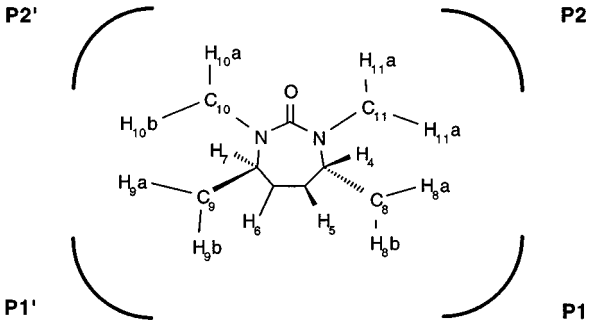


Figure 1. Numbering of hydrogen and carbon atoms.

Effect of Starting Structure on Low-Energy Conformational Search. Varying the starting structure by selecting frames randomly from the trajectory file and resubmitting to the above conditions resulted in identical low-energy conformations being generated, not necessarily within the same number of frames but always within the standard 200 ps time. The longest required time was 186 picoseconds, indicating the importance of adequate simulation times. These results suggest that longer simulation times should be employed if structures are being evaluated that vary significantly from the reference structures, as opposed to the homologous series reported here.

Inverting the configuration at each of the chiral ring carbons of XK216, to yield the *S*,*R*,*R*,*S* ax,eq,eq,ax molecule, and repeating the simulation also resulted in the set of low-energy conformations shown above, again suggesting limited dependence of results on initial conditions provided adequate temperatures and simulation times are used. The simulation does not invert the chiral centers.

Results

X-ray Crystal Structures of Cyclic Ureas. Free Ligands. To obtain information on the preferred low energy conformations of cyclic ureas in the solid state, single-crystal X-ray structures of the compounds shown in Table 1 were obtained as described above. These were chosen as a congeneric series that represent variations in ring substitution and stereochemistry and for which all relevant data were available, including the Ki values for enzyme inhibition. In the case of the urea with no nitrogen substituents, **1A**, we were unable to obtain an adequate crystal and report the structure of the acetate derivative **1I**, which according to NMR and calculated structures has a similar conformation to **1A**. The numbering scheme used in the following discussions is shown in Figure 1.

The crystal structures of **1B**−**1I** are shown in Figure 2. With the exception of **1I**, all of the derivatives crystallize with the ring conformation observed in the enzyme·CU complex (shown in Figure 3). The 4- and 7-benzyl groups project axially from the chairlike seven-membered ring, and the hydroxyls are equatorial, except for **1D**, in which the stereochemistry of one hydroxyl is inverted. **1D** is also the only structure, including **1I**, in which the four atoms of the urea group are out of plane (169°); the others are within 2−3° of a common plane. However, to provide the twist angle of the chair conformation, the O−C(2)−N(1)−C(7) and O−C(2)−N(3)−C(4) dihedrals

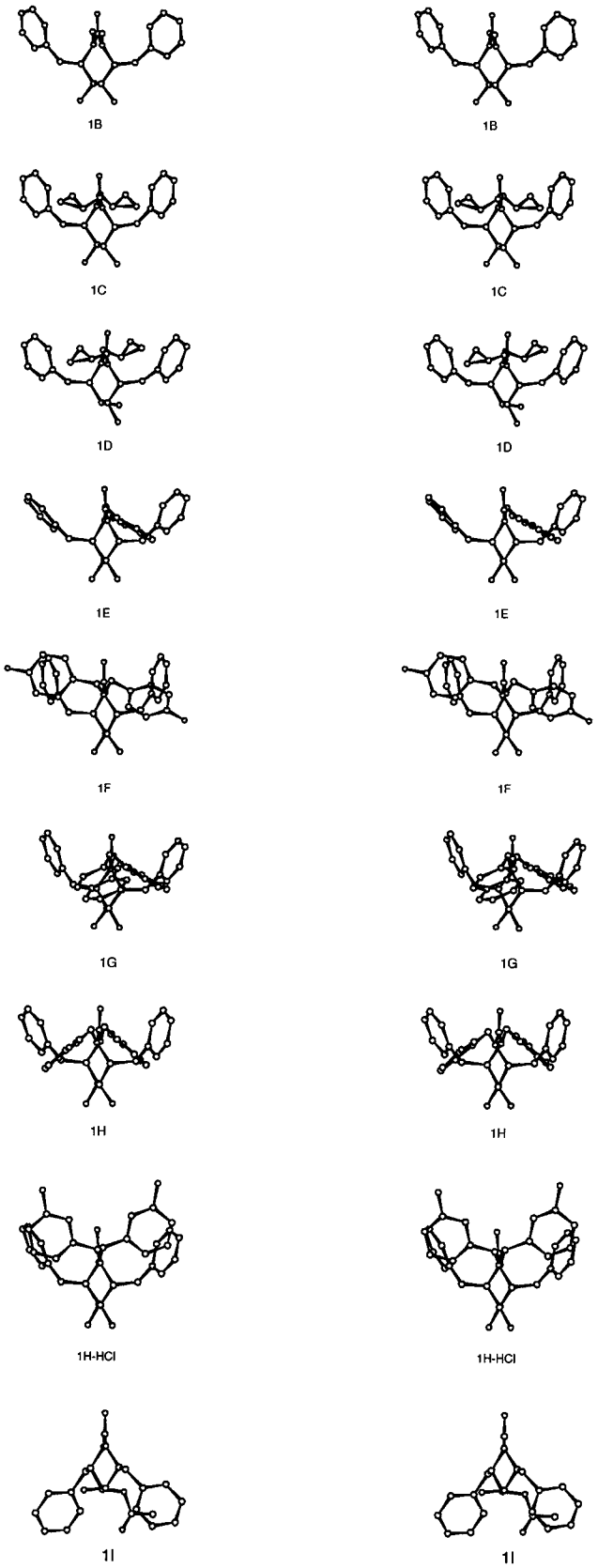


Figure 2. Stereoviews of the crystal structures of 1B−1I, viewed approximately along the axis of the hydroxyl-bearing ring carbons. ORTEP diagrams are available as Supporting Information.

range from −135 to −149°. That this angle arises in part due to strain relief in forming the optimum chairlike ring is supported by the observation that O−C(2)−N(1)−E* and O−C(2)−N(3)−E* dihedrals, where E* is the exocyclic nitrogen substituent, range from 10 to 20° when both nitrogens are substituted. Thus the exocyclic angles are closer to optimum (a search of the Cambridge Crystallographic Database shows that the most common dihedrals are between 0 and 10°) since no relief of ring strain compensates for the deviation. The same dihedrals are 8.0 and 8.7° for the substituted and unsubstituted nitrogens, respectively, of **1E**, and are 6.4 and −0.6 for the unsubstituted **1I**.

Thus with the exception of **1I**, the effect of the nitrogen substitution and hydroxyl inversion on ring conformation is fairly small, and all of the compounds are preorganized in the manner in which they were originally conceived. As discussed in the original report of cyclic urea HIV protease inhibitors[2] and is described in more detail in a more recent study,[5b] the alternate chair conformation occurs in **1I** to place the hydroxyls axial and the benzylic groups equatorial.

The 4- and 7-benzylic groups of all of the *N*,*N*-disubstituted crystal structures are remarkably uniform in orientation, showing the clockwise (looking down the O═C bond axis) screw or propellor shape that is observed on complexation with enzyme. The phenyl rings of monosubstituted **1E** splay outward but still have the same approximate torsions, while the unsubstituted **1I** projects the phenyls "downward", away from the carbonyl, a rotation of 180° relative to all of the other structures. A possible explanation is that the latter configuration is more favorable when the hydroxyls are axial, to avoid steric contact with the benzylic hydrogens; additionally when three or four hydrocarbon groups are arrayed around the central ring, (presumably) favorable van der Waals contacts can occur with neighboring substituents by placing the benzylic groups "upward" and the urea substituents "downward".

In fact a closely packed configuration is observed for the side chains in all of the structures. The urea substituents are tightly interdigitated with the benzylic groups in all cases but one, although several different angles are sterically allowed. The compounds with aromatic rings in the P2/P2′ position show a preference for a dihedral angle that allows edge to edge, rather than edge to face, contacts between the aryl groups. The only exception is the anilinium ring of DMP450, in which the benzylic torsion angles turn the charged nitrogens upward, presumably to interact with the extensively hydrogen-bonded water that is observed within and across the crystalline array (five ordered waters per unit cell in addition to the two chloride ions; see Supporting Information). Forces affecting crystal packing may also play a role in the observed conformations—hydrogen bonds between the hydroxyls and carbonyls of a neighboring molecule are observed in all of the structures. However, in nonpolar, polar aprotic, and aqueous solution the same geometry appears to occur (see below).

NMR Solution Structure. To gain insight into the preferred aqueous conformation of a water-soluble cyclic urea, spectra of the bis-methanesulfonate of DMP450 in deuterium oxide were obtained, and chemical shifts, coupling constants, and NOE cross-peaks were assigned from 2D NOESY and COSY experiments (see Supporting Information). The structures generated from the quenched dynamics simulation, above, were used to identify conformations consistent with the distances and angles observed in the NMR spectra. Three coupling constants and a single NOE were sufficient to unambiguously select a closely related family of structures and rule out all other conformations.

The 1D spectrum shows a single resonance for each proton and its symmetry-related counterpart about the C2 axis (see Supporting Information), suggesting a symmetrical structure on

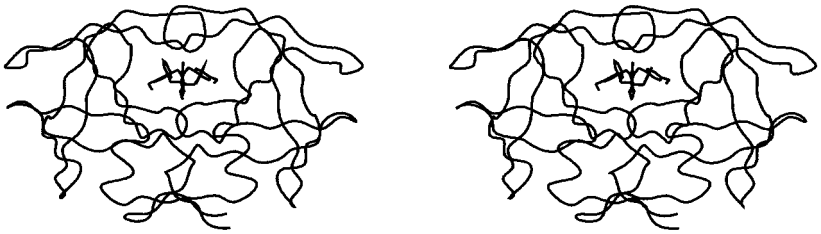


Figure 3. Stereoview of DMP323 in the active site of HIV-1 protease (represented as a ribbon structure). For details on the interactions between cyclic urea and protein, see refs 2 and 5.

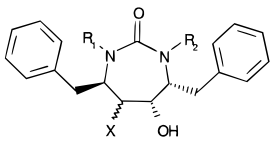
Table 3. Coupling Constants Used in Determining DMP450 Solution Conformation (See Figure 4)

proton	∂, ppm	*J*, Hz
H_4	3.74	12.1, 2.0, 1.3
H_5	3.93	1.3
H_{8a}, H_{8b}	2.78, 3.03	13.6, 12.1, 2.0

the NMR time scale. In the following discussion, referring to a single proton will include its symmetry partner, e.g., referring to H_{8a} will be understood to mean H_{8a} and H_{9a}. Also, P1 and P1′ are used to refer to the benzyl groups attached to seven-membered ring carbons C4 and C7 and P2 and P2′ refer to the substituents on N1 and N3, respectively. These terms reflect the protease binding pockets that are occupied by the cyclic urea substituents in the protein crystal structure.[2] The coupling constants used for conformational assignment are shown in Table 6. The modified Karplus equation allows a dihedral angle of between 60 and 90° for $H_4-C_4-C_5-H_5$ and $H_{8a}-C_8-C_4-H_4$ and of 160−180° for $H_{8a}-C_8-C_4-H_4$.[34] Of the 200 conformations generated in the QD analysis described above, 127 were eliminated by these requirements. A conformation consistent with these requirements (structure 168 in the QD protocol) is shown in Figure 4. The benzylic proton that is shifted upfield (H_{8B} in Figure 8) relative to its geminal counterpart (H_{8A}) also has the larger observed coupling constant, which supports the assigned structure, since H_{8B} projects into the deshielding cone of the P2 aromatic ring.

We also assumed that since one of the *N*-benzyl hydrogens, H_{11a} or H_{11b}, forms a strong cross-peak with H_4, but the other does not, then a conformation with a distance difference of ≥1 Å is likely (Figure 4, bottom). This value is proposed since, for example, the distance between H_4 and its vicinal partner H_5 can be no more than 2.5 Å, yet it gives rise to a weaker NOE than it does through space to one of the *N*-benzylic hydrogens. The two benzylic protons are in similar chemical surroundings, so it is reasonable to assume that the lower limit of the

Table 4. Comparison of Calculated and Experimental Conformations



entry	R_1	R_2	X	RMSD[a] all atoms[b]	RMSD[a] ring + 1 atoms[c]
1B	CH_3	CH_3	(*S*)-OH	0.28	0.13
1C	CH_2-cyclopropyl	CH_2-cyclopropyl	(*S*)-OH	0.49	0.22
1D	CH_2-cyclopropyl	CH_2-cyclopropyl	(*R*)-OH	0.49	0.13
1E	CH_2-2-naphthyl	H	(*S*)-OH	1.7	0.13
1F	4-fluorobenzyl	4-fluorobenzyl	(*S*)-OH	0.51	0.16
1G	CH_2-2-naphthyl	CH_2-2-naphthyl	(*S*)-OH	3.5	0.12
1H	3-aminobenzyl	3-aminobenzyl	(*S*)-OH	1.7	0.11
1H·$(HCl)_2$	3-aminobenzyl·$(HCl)_2$	3-aminobenzyl·$(HCl)_2$	(*S*)-OH	1.43	0.09
1I	H	H	(*R*)-OAc	3.5	0.39

[a] RMS difference (Å) between lowest energy calculated structure and crystal structure. [b] All heavy atoms. [c] Atoms in ring plus directly attached heavy atoms.

Table 5. Energy Distribution of Calculated Conformations

entry	total range,[a] kcal	unique,[b] <1 kcal: 4a, 5e, 6e, 7a chair[c]	unique,[b] <1 kcal: 4e, 5a, 6a, 7e chair[c]	unique,[b] <1 kcal: boat	unique,[b] <5 kcal: 4a, 5e, 6e, 7a chair[c]	unique,[b] <5 kcal: 4e, 5a, 6a, 7e chair[c]	unique,[b] <5 kcal: boat
1B	20	1	0	0	2	0	2
1C	21	4	0	0	10	0	4
1D	19	6	0	0	17	0	0
1E	32	2	0	1	8	0	12, 4[d]
1F	18	4	0	0	12	0	2
1G	21	2	0	0	17	0	2
1H	23	5	0	0	14	0	0
1H·$(HCl)_2$	20	3	0	0	13	0	0
1I	28	2	0	10	1	3	8, 11[e]

[a] Difference in internal strain energies between highest and lowest energy conformations from QD protocol, kcal. [b] Number of unique conformers within 1 or 5 kcal of lowest energy observed (see text). [c] a = axial, e = equatorial. [d] P1′ equatorial (12); P1 equatorial (4). [f] Acetate equatorial (8) acetate axial (11).

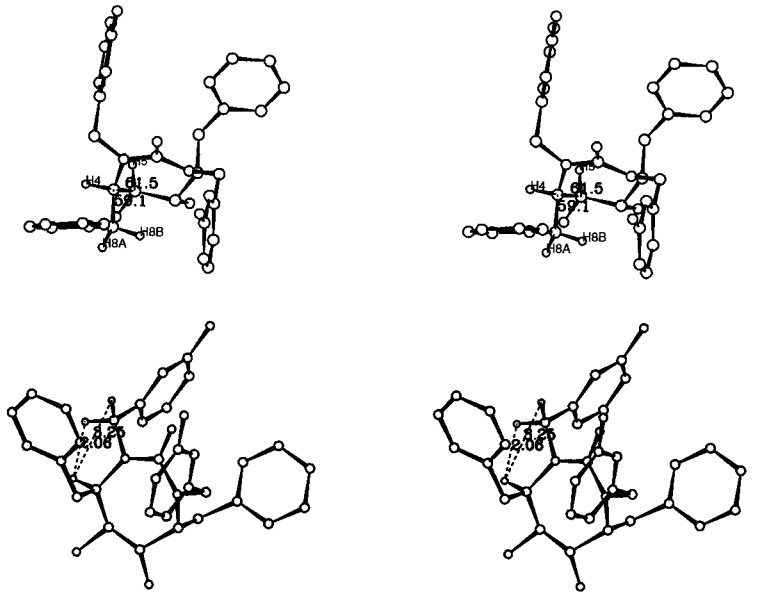


Figure 4. Top: Calculated structure of **1H**·$(HCl)_2$ consistent with NMR constraints showing the dihedral angle $H_4-C_4-C_5-H_5$ and H_{8a}-$C_8-C_4-H_4$. Not printed for clarity is the dihedral angle $H_{8a}-C_8-C_4-H_4$ (174.6°). Bottom: the distances $d(H_{11a}-H_4)$ and $(H_{11b}-H_4)$.

N-benzylic proton that does not give rise to a cross-peak with H_4 is 3 Å, and the upper limit of the proton that does give rise to a cross-peak is 2.5 Å. The upper limit proposed is consistent with recent studies on small molecules;[35] the validity of using unobserved NOEs in conformational assignments is also addressed in these references. This requirement eliminated 41 more conformations. The remaining conformations have identical ring geometry, that is, axial/equatorial/equatorial/axial, and differ only by dihedral angles defined by the benzylic carbons. These conformations fall into three families, where members of each family differ from other family members by less than 0.1 Å. All of the structures are consistent with the other observed cross-peaks in NOESY spectrum, although these peaks were not necessary for structural assignment. Structure 168 is most consistent with the pronounced upfield shift observed for one of the *N*-benzylic protons (3.30 vs 4.31 ppm for H_{11a} vs H_{11b}), since it has a strong interaction with the shielding region of the P1 aromatic ring. Structure 168 also has very similar ring and P1 dihedrals to those observed in the crystal structure and has the lowest calculated energy.

It is possible that low-energy conformations not identified in the QD analysis would also be consistent with the observed NOEs and coupling constants. Therefore the three angles and the single distance were used as restraints in a high-temperature constrained dynamics calculation (not shown); all of the structures generated minimized to the same families found in the original calculation. Thus, provided the force field employed is not grossly inaccurate, it is likely that no other stable conformations meet the simple NMR constraint requirements. Other force fields have been used in calculating low-energy cyclic urea conformations (data not shown) and also yield the axial,equatorial,equatorial,axial conformation as most stable for fully substituted cyclic ureas. It is also reasonable to suggest that the distances and angles used in the assignment represent averages of values that lie outside the assigned ranges; again, however, *low-energy* conformations that would give rise to such averages do not appear in the QD analysis, and the observed low-energy conformations are sufficient to explain the observed NMR.

Thus the NMR studies are consistent with the average conformation of DMP450 in water being close to that observed in the free crystal structure and the lowest energy calculated structure (see below). Data in organic solvents are not conclusive but are consistent with the axial,equatorial,equatorial,axialconformation for *N*,*N*-disubstituted cyclic ureas (see Experimental Section). The NMR solution structure of DMP323 complexed to HIV-1 PR also displays the axial,equatorial,-equatorial,axial conformation.

Conformations of Cyclic Ureas Bound to HIV−Protease. Published crystallographic analyses[2,5] have shown that CUs with large (> three atoms) substituents on both nitrogens bind to HIV-1 protease with the ring in the axial,equatorial,equatorial,-axial chair conformations. The NMR solution structure of DMP323 complexed with HIV-1 PR is consistent with the same average conformation.[4] With the exception of the mono- and unsubstituted analogues, **1E** and **1I**, the same ring conformation is observed in the free ligands. The P1/P1′ benzyl torsion angles differ slightly, and the P2/P2′ torsions differ significantly in some cases (**1F**, **1H**·$(HCl)_2$).

Calculated Conformations. To determine our ability to computationally generate a priori the experimentally observed low-energy conformations, the cyclic ureas in Table 1 were analyzed using the optimized quenched dynamics conditions described above. A summary of the single lowest-energy structures and their RMS difference from the experimental conformations is shown in Table 4. The column denoted "Ring+1 RMSD" refers to the RMS deviation of ring atoms as well as the exocyclic atoms attached to ring atoms—i.e., a total of 14 heavy atoms— and "all atoms" refers to all heavy atoms. The effects of symmetry are taken into account. Figure 5 summarizes the lowest energy conformation of each molecule observed in the QD simulation.

As described in the methods section, the low-energy conformations are calculated in the gas phase but with a dielectric $\epsilon = 80$, and no effort is made to reproduce the intermolecular contacts with solvent or neighboring molecules observed in the crystal structure. Therefore the calculated and observed struc-

(34) Altona, C.; Franke, R.; de Haan, R.; Ippel, J. H.; Daalmans, G. J.; Hoekzema, W. A. J. A.; van Wijk, J. *Magn. Reson. Chem.* **1994**, *32*, 670−678.

(35) (a) Brueschweiler, R.; Blackledge, M.; Ernst, R. R. *J. Biomolecular NMR* **1991**, *1*, 3−11. (b) Bean, J. W.; Kopple, K. D.; Peischoff, C. D. J. *J. Am. Chem. Soc.* **1992**, *114*, 5328−5334.

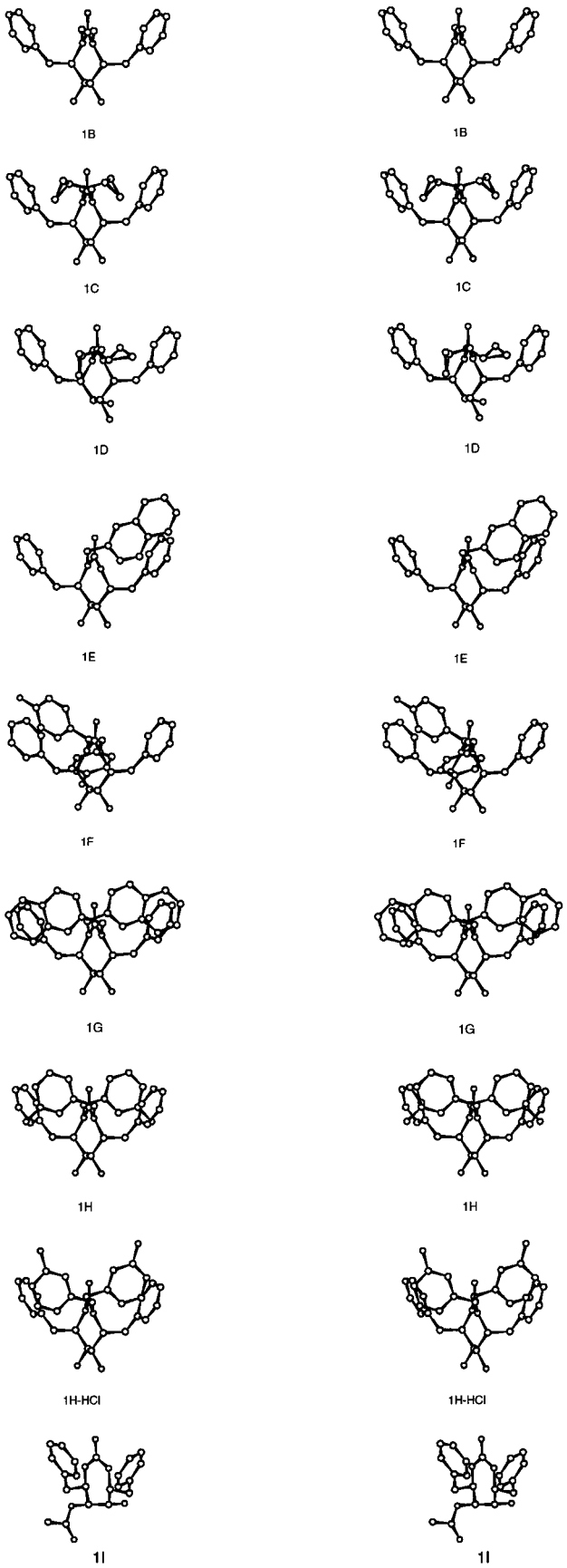


Figure 5. Stereoviews of lowest energy conformation of cyclic ureas from quenched dynamics simulation viewed approximately along the axis of the hydroxyl-bearing ring carbons (see also Table 5).

tures are not expected to be identical if these intermolecular forces are significant and in opposition to preferred intramolecular conformers. However, in many cases the structures match fairly closely, suggesting that crystal packing has limited effect at least on ring conformational preference (or, less likely, that inaccuracies in the force field balance out the missing intermolecular effects).

Lowest Energy Conformation. The low-energy QD-generated *ring* conformations for all of the disubstituted cyclic ureas are almost identical to the crystallographically observed free ligand conformations (Figure 5 and Table 4). The main variation between the observed and calculated structures occurs in the dihedral angles of the nitrogen substituents, and as the group increases in size, the relative weighting of small changes increases, causing larger RMSD values for all atoms. Compound **1B** has a very small all-atom RMSD, as even the small methyl group on nitrogen substituents appears to enforce the P1/P1′ and ring geometry. The bis-*N*-cyclopropylmethyl substituted compounds **1C** and **1D** display four and two distinct rotamers of equivalent energy in the calculated structures, respectively; these same atoms show high thermal motion in the X-ray structures (see Supporting Information), supporting the validity of several degenerate conformations. A 180° rotamer of C_4-C_8 is observed in the lowest 1 kcal only in **1D**, although this variant appears in other calculated structures at higher energy.

Mono-*N*-2-naphthylmethyl cyclic urea **1E** crystallizes with the plane of the naphthyl ring tucked inward between the two phenyl rings (Figure 2), similar to the geometry of P2 benzyl rings in complex with the HIVPR enzyme (see below), whereas the two calculated conformations in the lowest 1 kcal include a twist boat as well as the usual chair. (NMR NOESY experiments also suggest a twist boat for this compound in $CDCl_3$: Lim, M. personal communication.) The calculated chair conformer of **1E** has the naphthalene ring turned up and outward to expose its edge to "solvent" and its face to the neighboring phenyl ring. All of the calculated structures that contain an aromatic nitrogen substituent show a preference for this upturned ring, in which an face−edge, rather than an edge−edge, interaction is made with the P1 phenyl rings. This orientation differs from some of the experimental structures and may be in part an artifact of the force field, which will not capture the details of $\pi-\pi$ interactions. However, one of the benzyl rings of bis-4-fluorobenzyl derivative **1F**, and both of the rings of the anilinium derivative **1H** (but not aniline derivative **1G**) exhibit this preference in the crystal structure, and the solution structure of **1H** appears to be consistent with the calculated and crystallographic conformation. In all of the bis-P2-aryl (**1F**, **1G**, **1H**) cyclic ureas analyzed by quenched dynamics, between three and five equienergetic benzyl rotamers are observed; in these cases thermal motion is not observed in the crystal structures. Two possible explanations come to mind: the force field does not accurately balance the interaction energies with the conformational strain energy, or the hydrogen bonding between neighboring molecules observed crystallographically, but neglected in the simulation, stabilizes the observed conformation. In crystal structures of cyclic ureas co-complexed with HIV-1 PR, the hydrogen bonds from amino acid backbone nitrogens to the cyclic urea carbonyl also clearly favor the conformation in which the aromatic groups are tucked between the P1/P1′ phenyl rings and away from steric contact with other residues (Figure 3). The preferred arrangement of the four phenyl rings appears to be due to a delicate stereoelectronic balance that is reproduced in part, but not completely, by the force field. Thus observed differences between the calculated and observed benzyl rotamers appear to be due to small

inaccuracies in the force field's treatment of aryl−aryl interactions as well as to neglect of the crystal environment.

The QD-generated conformations of **1I** place the equatorial alcohols, axial alcohols, and a boat conformer at approximately equal energy, whereas the crystal structure shows a well-resolved chair conformer with axial alcohols (Figure 5i). Additionally, the P1/P1′ benzyl groups in the X-ray structure are oriented almost 180° from the calculated chair conformer. The QD trajectory in fact contains the experimentally observed conformation at 123 ps, with a relative energy only 1.6 kcal higher than lowest energy structure. Unlike the N-substituted cyclic ureas, the QD trajectory for **1I** contains many distinct conformers that are close in energy, due to the increased steric freedom of the P1/P1′ benzyl groups. NMR in organic solvent shows coupling constants consistent with either the eq,ax,ax,eq or ax,-eq,eq,ax chair conformation, but no conclusions on the orientation of the benzyl dihedral can be drawn (see Experimental Section). It should also be pointed out that the crystal structure of **1I** shows intermolecular hydrogen bonds between the urea nitrogens and carbonyl, so the poorer match between gas-phase calculations and the crystal structure is not surprising.

Conformations Observed in Lowest 5 kcal. Table 5 summarizes the energetic distribution of the two hundred minimum energy conformations from each simulation, and their distribution into +1 and +5 kcal energy bands. On average, each conformation was visited about two times during the 200 ps simulation. The second column indicates the difference between the highest and lowest energy conformers found in the simulation. The number of unique conformers in the lowest 1 kcal is shown in the third column. To find unique conformations, the full set of 200 were ranked by total strain energy and those within 1 kcal of the minimum were clustered based on the 14 atoms that make up the ring and their non-hydrogen attachments. Families were formed of clusters that differed from each other by more than 0.3 Å RMSD, which grouped ring conformers into chair, boats, twist boats, etc., uniquely. Within families, a conformation was considered unique if any of the side chain dihedrals differed from the others by >10°: that is, if all of the dihedrals were identical within ±10° of another member of the family, a conformer was discarded; if a single dihedral differed by >10° (see ref 26c) the conformer was considered unique. Compounds **1B** and **1I** therefore have four dihedrals, compound **1E** has six, and the others have eight for the purpose of this clustering.

With the exception of unsubstituted **1I**, all of the conformations could be classed as pseudo-chair or pseudo-boat; bond angles in each geometry were similar across all of the CUs studied. The number of conformers with each ring geometry is shown in the fourth and fifth columns of Table 5. Only **1E** and **1I** have boats among the lowest 1 kcal (Figure 5D and I). The same analysis was carried out for conformations occurring within the lowest 5 kcal (columns 6−8 of Table 5). At the higher energy the boat conformations begin to be more populated−interestingly, the *R*,*S*,*R*,*R* isomer **1D** does not adopt a boat conformation within this energy band, whereas **1C**, identical except for the stereochemistry of one hydroxyl, has a number of low-energy boat-side chain combinations.

Discussion

For the purposes of molecular design, the value of accurately reproducing free ligand low-energy conformations lies in being able to prospectively, as opposed to retrospectively, select ligands that are able to adopt a complementary conformation to the active site of a receptor of interest. If one analyzed N-substituted cyclic ureas with the procedure described here and selected only the lowest energy conformation, the result would be a close match with both the free and (where available) bound crystallographic conformation in all important geometric parameters except the P2/P2′ benzylic dihedrals, which have several low-lying minima. Although not the subject of this paper, these calculated conformations can easily be docked and minimized in the active site to match the bound ligand geometry within a few tenths of an angstrom.[17]

Perhaps just as important to both ligand design and the prediction of binding affinity is the observation that other, significantly different cyclic urea conformations are very close in energy to the enzyme-bound conformation, as seen in Table 5 and Figure 5. These results indicate that the initial conformational analyses[2,5b] are insufficient to allow detailed understanding of binding affinity, since the entropic contributions of other forms would be significant at ambient temperature. This in turn suggests that further increases in potency may be possible with increased conformational rigidity.

The observation that the calculated structures as well as the experimental structures of the free inhibitor in both crystalline and aqueous environments have stable conformations that are very close to enzyme-bound conformations supports the notion[2] that the potency of cyclic ureas relative to their acyclic counterparts arises in part from preorganization of the side chains, the diol, and the water-mimicking carbonyl for interaction with the respective enzyme residues.

The quenched dynamics method described here is able to sample conformational space efficiently and thoroughly for molecules such as cyclic ureas; although fairly simple structurally, the eight significant dihedral angles and the presence of a flexible ring cause difficulty in full conformational analysis by other means. Regarding the use of quenched dynamics as a method of prescreening putative ligands for the ability to attain the shape of a reference ligand: a clear advantage is that initially no simulation of enzyme is carried out; only conformational searching of the novel ligand and comparison to the reference ligand are required. The exponential scaling of computational time with increasing numbers of atoms that is characteristic of molecular mechanics is thus avoided, and any candidate ligand for which trustworthy parameters are available can be evaluated in a matter of minutes. The method obviously suffers from the fact that regions of the enzyme that do not interact with the reference ligand are not explored; also, potentially favorable adjustments in the enzyme active site that could increase inhibitor affinity are neglected. A more sophisticated analysis is needed for targets that pass the prescreening requirements. It should be noted, however, that attempts to even qualitatively rank order ligands in order of binding affinity using computational techniques has met with questionable success, even among highly congeneric series.[36] Thus in some cases, the crude assessment provided by matching low energy conformations of a prospective ligand to the known conformation of a protein-bound ligand may be the most accurate information that a chemist will have access to in a useful time frame. It is also probably safe to say that matching a prospective ligand to a bound conformation without carrying out a complete conformational analysis will often yield poor results if the ligands are flexible.

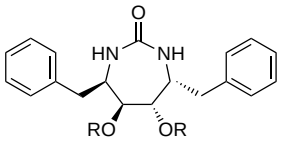
(36) (a) Head, R. D.; Smythe, M. L.; Oprea, T. I.; Waller, C. L.; Green, S. M.; Marshall, G. R. *J. Am. Chem. Soc.* **1996**, *118*, 3959−3969. (b) Reference 20. (c) Kauvar, L. M.; Higgins, D. L.; Villar, H. O.; Sportsman, J. R.; Engqvist-Goldstein, A.; Bukar, R.; Bauer, K. E.; Dilley, H.; Rocke, D. M. *Chem. Biol*. **1995**, *2*, 107−118. (d) Welch, W.; Ruppert, J.; Jain, A. N. *Chem. Biol.* **1996**, *3*, 449-462.

This report expands the set of complex organic molecules that have been searched for conformational minima rapidly and successfully (if the benchmark for success is finding experimentally observed conformations). Subtle differences with experiment are noted, particularly in the edge-interactions of phenyl rings. The calculated lowest energy structure of nitrogen-unsubstituted cyclic urea **1I** is significantly different from its X-ray crystal structure, but the correct structure does appear at 1.6 kcal higher in energy, and the discrepancy is not difficult to rationalize (see Results section). The structures generated by the quenched dynamics protocol are also useful as a simple means of identifying conformations consistent with observed NMR NOEs and coupling constants. Estimation of the internal strain energy imposed on cyclic ureas after binding to HIV1-PR will be reported separately.

Finally, these results add to the number of cyclic ureas reported in the literature for which experimentally determined free and enzyme-bound conformations and enzyme affinity constants are known. We have found these data to be useful in evaluating the performance of various conformational analysis, ligand−protein affinity prediction, and ligand−protein docking methods developed in our laboratories and elsewhere. The highly specific binding mode, the absence of unusual functionality, and the rigidity of these inhibitors make them particularly tractable for this purpose; these studies as well as other fully characterized congeneric series of interest will be reported shortly. Others in the chemistry community may find these data equally valuable in benchmarking their methods.

Experimental Section

Protease Inhibition Assays. Values for inhibition constant, Ki, were determined with a fluorescent peptide substrate at pH 5.5 with 1.0 M NaCl as described previously (ref 6a).



II

IIA: R = [(2-trimethylsilyl)methoxy]ethoxy (SEM)
IIB: R = 2-methoxyethoxy (MEM)

Chemistry. All procedures were carried out under inert gas in oven-dried glassware unless otherwise indicated. Proton NMR spectra were obtained on VXR or Unity 300 or 400 MHz instruments (Varian Instruments, Palo Alto) with TMS as an internal reference standard. Melting points were determined on a Mettler SP61 apparatus and are uncorrected. Elemental analyses were performed by Quantitative Technologies, Inc., Bound Brook, NJ. High-resolution mass spectra were carried out on a VG 70-VSE instrument with NH3 chemical ionization. Thin layer and column chromatography were carried out on plates or silica gel from E. Merck, Darmstadt, FRG. Separation of optical isomers was performed using supercritical fluid chromatography with a Chiracel OD (Daicel Chemical Ind. Ltd.) and 20% methanol modified CO2 mobile phase. Optical rotations were obtained on a Perkin-Elmer 241 polarimeter. Solvents and reagents were obtained from commercial vendors in the appropriate grade and used without further purification unless otherwise indicated.

[4*R*-(4a,5a,6b,7b)]Hexahydro-5,6-bis(hydroxy)-1,3-bis(cyclopropylmethyl)-4,7-bis(phenylmethyl)2*H*-1,3-diazepin-2-one (1C). A. Alkylation. SEM-protected urea **IIA**[5a] (700 mg, 1.19 mmol) in 2 mL of DMF was added to a flask containing NaH (11.9 mmol) that had been washed with dry hexane (2 × 10 mL) in 8 mL of DMF. The mixture was stirred at room temperature under N_2 for 10 min. Bromomethylcyclopropane (Aldrich, 0.81 mL, 9.33 mmol) was added, and the solution was heated at 70 °C for 5−7 h or until TLC (10% ethyl acetate/hexane) indicated loss of starting material. The reaction was cooled, quenched with methanol (1−2 mL), and partitioned between ethyl acetate (80 mL) and water (70 mL), and the upper layer removed. The lower layer was washed with ethyl acetate (2 × 30 mL). The organic extracts were combined and washed with water (2 × 50 mL) and brine (50 mL), dried over $MgSO_4$, and concentrated to a residue. The residue was chromatographed on silica gel and eluted with 10−15% ethyl acetate/hexane to afford the bis-cyclopropylmethyl cyclic urea (SEM-protected) as a colorless oil (710 mg, 86%). ^{1}H NMR (300 MHz, $CDCl_3$): 7.51−7.4 (m, 10H, Ph); 4.8−5 (d, 4H, OCH2O); 4.08(s, 2H); 3.5−3.9 (m, 14H); 0.8−0.9 (m, 4H). MS: (CDI) 696 (M + 1, 100%).

B. Deprotection. The above product was hydrolyzed in methanol (10 mL) and 1.0 M HCl in ether (10 mL) at room temperature under N_2 for 12 h. The solvent was evaporated, and the residue was chromatographed with 10−15% ethyl acetate/dichloromethane eluent to afford 377 mg (86%) bis-alkylated cyclic urea 1C: white solid, mp 210−212 °C. NMR (300 MHz, CDCl3): 7.0−7.4 (m, 10H); 4.8−5 (q, J = 7.5 Hz, 4H); 3.9 (s, 2H); 3.8 (m, 2H); 3.6 (m, 4H); 3.5 (q, J = 7.5 Hz, 2H); 3.1 (t, J = 11 Hz, 2H); 3.0 (t, J = 11 Hz, 2H); 1.9 (q, J = 7.5 Hz, 2H); 0.8−1.0 (m, 8H); 0.2−0.4 (m, 4H). MS. (CDI) 435 (M + 1, 100%). HRMS: calcd, 435.2647; found, 435.2636. Anal. Calcd for $C_{27}H_{34}N_2O_3$: C, 74.62, H, 7.89; N, 6.44. Found: C, 74.46; H, 7.81; N, 6.39. Crystals for X-ray analysis from toluene/CH_2Cl_2, colorless irregular cubes.

[4*R*-(4a,5a,6b,7b)]-Hexahydro-5-hydroxy-6-acetoxy-4,7-bis-(phenylmethyl)-2*H*-1,3-diazepin-2-one (1A). MEM-protected urea **IIB**[5b] was deprotected using the procedure described above to yield urea **1A**: mp 172−174 °C. MS(CDI) 327(M + 1, 100%). HRMS calcd: 327.1701; found: 327.1713. Anal. Calcd for $C_{19}H_{22}N_2O_3$: C, 69.92; N, 8.58; H, 6.79. Found: C, 69.39; N, 8.47; H, 6.73.

[4*R*-(4a,5a,6b,7b)]Hexahydro-5,6-bis(hydroxy)-1,3-bis[(4-fluorophenyl)methyl)-4,7-bis(phenylmethyl)-2*H*-1,3-diazepin-2-one (XL472, 1F). Following the two-step general procedure described above, *N*,*N*-bis(4-fluorobenzyl) cyclic urea **1F** was obtained from MEM-protected urea **IIA** using 4-fluorobenzyl bromide as the alkylating agent to give colorless crystals from butyl ether: mp 133 °C. MS (CDI): 583.2 (M + H, 99%). HRMS calcd, 583.2772; found: 583.2767. Anal. Calcd for $C_{36}H_{36}N_2O_3F_2$: C, 74.21; H, 6.23; N, 4.81. Found: C, 73.82; H, 6.21; N, 4.74.

[4*R*-(4a,5a,6b,7b)]Hexahydro-5,6-bis(hydroxy)-1,3-bis(2-naphthylmethyl)-4,7-bis(phenylmethyl)-2H-1,3-diazepin-2-one (XK263, 1G). Following the two-step general procedure described above, *N*,*N*-bis-(2-naphthylmethyl) cyclic urea **1G** was obtained from SEM-protected urea **IIA** using 2-bromomethylnaphthalene as the alkylating agent: white solid, 202−204 °C. NMR (300 MHz, CDCl3): 7.0−8.0 (m, 14H); 5.1 (d, 2H); 3.6 (m, 4H, CH); 3.2 (d, 2H); 3.1 (d, 4H); 2.2 (s, 2H). MS (CDI) *m*/*z* 607 (M + 1, 100%). HRMS calcd, 607.2961; found, 607.2960. Anal. Calcd for $C_{41}H_{38}N_2O_3$: C, 81.16; H, 6.31; N, 4.62. Found: C, 80.85; H, 6.20; N, 4.54. Crystals for X-ray analysis from CH_3OH, colorless hexagonal rods.

[4*R*-(4a,5a,6b,7b)]Hexahydro-5,6-bis(hydroxy)-1-(2-naphthylmethyl)-4,7-bis(phenylmethyl)-2H-1,3-diazepin-2-one (XK291, 1E). MEM-protected urea **IIB** (3.55 g, 7.06 mmol) in 20 mL DMF of was added to a flask containing NaH (60% in oil) (1.78 g, 44.5 mmol) that had been washed with dry hexane (2 × 10 mL) in 8 mL of DMF. The mixture was stirred at room temperature under N_2 for 10 min. 2-Bromomethylnaphthalene (2.01 g, 8.9 mmol) was added, and the solution was heated to 40 °C for 4h. The solution was quenched with methanol (5−8 mL), partitioned between ethyl acetate (180 mL), and water (170 mL) and the organic layer was removed. The aqueous layer was washed with ethyl acetate (2 × 80 mL). The organic extracts were combined and washed with water (2 × 80 mL) and brine (80 mL), dried over $MgSO_4$, and concentrated to a residue. The residue was chromatographed on silica gel and eluted with 10−40% ethyl acetate/hexane to separate the protected bis-substituted (1.4 g, 25%) and monosubstituted (0.6 g, 14%) cyclic ureas as colorless oils. The monosubstituted naphthalene (90 mg, 0.14 mmol) was hydrolyzed in methanol (10 mL) and 4.0 M HCl in dioxane (10 mL) by stirring at

room temperature under N_2 for 12 h. The solvent was evaporated, and the residue was chromatographed on silica gel and eluted with 40% ethyl acetate/dichloromethane to afford mono-*N*-naphthylmethyl cyclic urea **1E**: white solid, 49.5 mg, (76%); mp 210−213 °C. ^{1}H NMR (300 MHz, $CDCl_3$): 7.2−7.8 (m, 17H); 5.1 (d, J = 15 Hz, 2H); 3.9 (m, 1H); 3.6 (m, 2H); 3.4 (m, 1H); 3.1 (m, 4H); 2.9 (d, J = 15 Hz, 1H); 2.7 (br s, 1H); 2.4 (br s, 1H). MS (CDI) 467 (M + 1, 100%). HRMS calcd: 467.2335; found: 467.2330. Anal. Calcd for $C_{30}H_{30}N_2O_3 \cdot {}^1/_2H_2O$: C, 75.76; H, 6.57; N, 5.89. Found: C, 75.68; H, 6.30; N, 5.82. Crystals for X-ray analysis from CH_3OH/hexane, colorless parallelepipeds.

Supporting Information Available: X-ray crystal data, descriptions of data collection, treatment, solution, and refinement and tables of fractional coordinates, isotropic and anisotropic thermal parameters for compounds **1B**−**1I**; input file for the quenched dynamics protocol; and ^{1}H, ^{13}C, hsqc and COSY spectra of the bis-methanesulfonate salt of ^{1}H (79 pages, print/PDF). See any current masthead page for ordering and web access instructions.

JA972357H