J|A|C|S
ARTICLES
Published on Web 08/23/2002

Structure Activity Relationship by NMR and by Computer: A Comparative Study

Finton Sirockin,[†] Christian Sich,[§,||] Sabina Improta,[§] Michael Schaefer,[‡,⊥] Vladimir Saudek,[§,#] Nicolas Froloff,[§,¥] Martin Karplus,[‡] and Annick Dejaegere*,[†]

Contribution from the Laboratoire de Biologie et Génomique Structurales, UMR 7104, Ecole Supérieure de Biotechnologie de Strasbourg, Boulevard S. Brant, FR-67400 Illkirch, France, Laboratoire de Chimie Biophysique, ISIS, Rue Blaise Pascal, FR-67000 Strasbourg, France, and Department of Structural Biology and Cheminformatics, Department of Biotechnology, Sanofi~Synthélabo, Rue d'Ankara, FR-67000 Strasbourg, France

Abstract: There has recently been considerable interest in using NMR spectroscopy to identify ligand binding sites of macromolecules. In particular, a modular approach has been put forward by Fesik et al. (Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. *Science* **1996**, *274*, 1531−1534) in which small ligands that bind to a particular target are identified in a first round of screening and subsequently linked together to form ligands of higher affinity. Similar strategies have also been proposed for in silico drug design, where the binding sites of small chemical groups are identified, and complete ligands are subsequently assembled from different groups that have favorable interactions with the macromolecular target. In this paper, we compare experimental and computational results on a selected target (FKBP12). The binding sites of three small ligands ((2*S*)1-acetylprolinemethylester, 1-formylpiperidine, 1-piperidinecarboxamide) in FKBP12 were identified independently by NMR and by computational methods. The subsequent comparison of the experimental and computational data showed that the computational method identified and ranked favorably ligand positions that satisfy the experimental NOE constraints.

Introduction

Many different strategies are being applied in the search for new therapeutic agents, but an essential step in almost all strategies is the identification of an appropriate lead structure.[2,3] Leads are often identified by screening large libraries of compounds with an assay for the binding.[4] Virtual in silico calculations can be combined with physical screening to increase the efficiency of the procedure.[5] However, in some cases, no suitable leads are found by this approach. This is sometimes due to the sensitivity limit of the biochemical detection assay, which may miss weakly binding compounds in the micromolar to millimolar range. In some instances, a compound contains an active core structure that would bind to the target, but loses affinity for the target owing to sterically or electronically unfavorable side-chain contacts.[2] There is thus an interest in identifying small molecular weight compounds that could form the core of lead compounds but that bind to a target with low affinity. For identifying such compounds, screening assays using biophysical methods, such as NMR, have been proposed.[1,3,6−8] These core compounds can be used in different ways toward lead identification. For example, several core compounds can be linked together to form larger ligands, in a de novo ligand build.[1] Alternatively, such a core compound can be used to more effectively screen libraries.[3] They can also be used to design combinatorial libraries that are biased toward the target structure and can be assayed subsequently.[9]

The utility of binding data for a macromolecule−ligand complex is greatly enhanced if structural information on the complex is available. Obtaining experimental structural information by NMR or X-ray crystallography can be time-consuming. It can be severely hampered by problems linked to the solubility of the ligands, the size of the target (NMR), or the crystallization process. In silico docking methods are possible alternatives for obtaining structural information, and there exists vast literature

* To whom correspondence should be addressed. E-mail: annick@esbs.u-strasbg.fr.
[†] Laboratoire de Biologie et Génomique Structurales.
[‡] Laboratoire de Chimie Biophysique, ISIS.
[§] Sanofi~Synthélabo.
[||] Present address: Volkswagen AG, Brieffach 1617/0, D-38436 Wolfsburg, Germany.
[⊥] Present address: Syngenta Crop Protection, P.O. Box, CH-4002 Basel, Switzerland.
[#] Present address: Aventis Pharma, Paris Research Center, 13, quai Jules Guesde, F-94400 Vitry-sur-Seine, France.
[¥] Present address: L'Oréal Recherche, 1, avenue Eugène Schueller, BP 22, FR-93601 Aulnay sous Bois, Cedex, France.

(1) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. *Science* **1996**, *274*, 1531−1534.
(2) Boehm, H.-J.; Boehringer, M.; Bur, D.; Gmuender, H.; Huber, W.; Klaus, W.; Kostrewa, D.; Kuehne, H.; Luebbers, T.; Meunier-Keller, N.; Mueller, F. *J. Med. Chem.* **2000**, *43*, 2664−2774.
(3) Fejzo, J.; Lepre, C. A.; Peng, J. W.; Bemis, G. W.; Ajay; Murcko, M. A.; Moore, J. M. *Chem. Biol.* **1999**, *6*, 755−769.
(4) Silverman, L.; Campbell, R.; Broach, J. R. *Curr. Opin. Chem. Biol.* **1998**, *2*, 397−403.
(5) Walters, P.; Stahl, M. T.; Murcko, M. A. *Drug Discovery Today* **1998**, *3*, 160−178.
(6) Liepinsh, E.; Otting, G. *Nat. Biotechnol.* **1997**, *15*, 264−268.
(7) Dalvit, C.; Floersheim, P.; Zurini, M.; Widmer, A. *J. Biomol. NMR* **1999**, *14*, 23−32.
(8) Moy, F. J.; Haraki, K.; Mobilio, D.; Walker, G.; Powers, R.; Tabei, K.; Tong, H.; Siegel, M. M. *Anal. Chem.* **2001**, *73*, 571−581.
(9) Joseph-McCarthy, D.; Tsang, S. K.; Filman, D. J.; Hogle, J. M.; Karplus, M. *J. Am. Chem. Soc.* **2001**, *123*, 12758−12769.

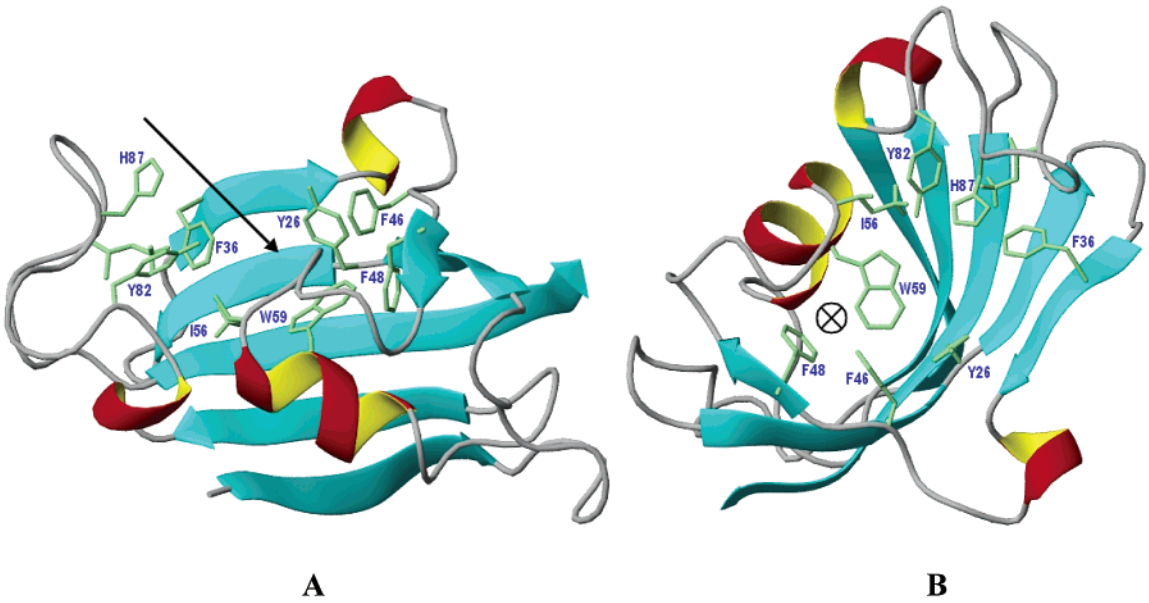


Figure 1. Ribbon drawing of FKBP12 showing the side chains of the residues lining the binding pocket. (A) Side view. The arrow indicates the entrance to the binding pocket. (B) Top view with TRP 59 and the arrow from (A) shown. This orientation will be used in subsequent figures that present ligands bound to FKBP12.

on the development and assessment of computational methods for ligand docking (see, for example, Nussinov et al.[10] and references therein). These methods have been mostly applied to docking of ligands with the full size of a drug molecule (MW 250−500[11]) rather than small compounds that can be used as fragments for developing larger optimized ligands. In this study, we determined the structures of three protein/small ligand complexes. The structure is determined independently by NMR and by a computational protocol. The docking procedure is a refinement of the force-field-based MCSS method;[12] that is, after finding local minima for the MCSS groups corresponding to the ligands, the relative binding free energies of the minima are postprocessed by a Poisson−Boltzmann continuum model for the solvent. The purpose of our study was two-fold. First, we wished to assess the performance of a physics-based docking method in the case of low-affinity compounds that have been much less studied than higher affinity ones. In addition, we thought it important to assess the performance of the computational procedure in a "real life" situation by comparing its predictions to NMR data that were generated independently and that were available only after the predictions had been made.

The target used for the present study is the protein FKBP12. FKBP12 is a peptidyl prolyl cis−trans isomerase (PPIase) protein belonging to the family of the immunophilins.[13,14] The FKBP12 structure is shown in Figure 1. FKBP12 binds FK506 (Figure 2)[15] and Rapamycin[16,17] with high affinity;[18,19] these are natural products that have activity as immunosuppressive agents. FKBP12 possesses a hydrophobic core containing six of the protein's nine aromatic side chains. The binding pocket is mostly hydrophobic, but it is lined on the outside by charged residues. The ligands chosen were (2*S*)1-acetylprolinemethylester (ACPM), 1-formylpiperidine (FOPI), and 1-piperidinecarboxamide (PICA) (Figure 2). These three polar ligands were commercially available and sufficiently soluble for the NMR study. They have common features with FK506 (in particular, the five- or six-membered heterocycle was chosen to mimic the pipecolinyl ring in FK506), and they were therefore expected to be weak but specific FKBP12 ligands.

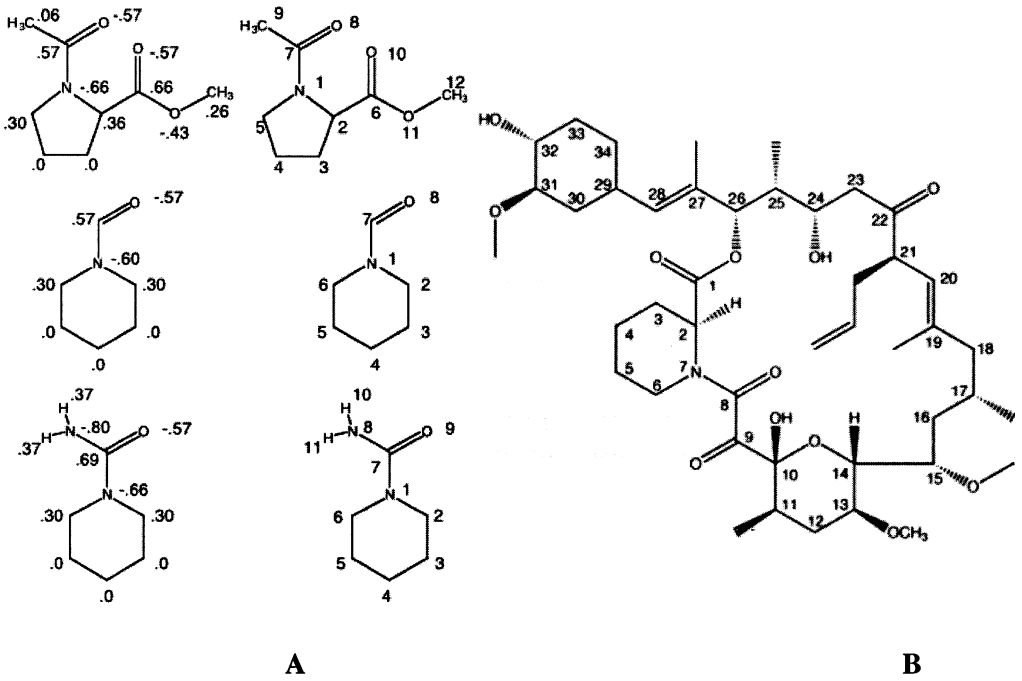
The multiple copy simultaneous search (MCSS) method[12] is employed to place the three small ligands in the FKBP12 binding pocket. The MCSS method was used to perform an extensive sampling of the possible positions of the ligands in the known binding pocket of FKBP12 using multiple copy energy minimization. Each position is subsequently evaluated by a physics-based scoring function that includes contributions from a molecular mechanics force field (CHARMM[20]) and implicit solvation.[21,22] The comparison of the computational and NMR results allows for a critical evaluation of the docking and scoring procedure.

Materials and Methods

NMR Spectroscopy. NMR measurements were performed on a Bruker DMX-600 spectrometer equipped with a 5-mm triple resonance probe and z-shielded gradient coils. All spectra were acquired at $T = 300$ K.

The FK506-binding protein (FKBP12) was produced as previously described.[23] Ac-Pro-OMe was purchased from Bachem, and 1-piperidinecarboxamide and 1-formylpiperidine were purchased from Sigma Aldrich.

(10) Halperin, I.; Ma, B.; Wolfson, H.; Nussinov, R. *Proteins* **2002**, *47*, 409−443.
(11) Bissantz, C.; Folkers, G.; Rognan, D. *J. Med. Chem.* **2000**, *43*, 4759−4767.
(12) Miranker, A.; Karplus, M. *Proteins* **1991**, *11*, 29−34.
(13) Wiederrecht, G.; Etzkorn, F. *Pespect. Drug Discovery Des.* **1994**, *2*, 57−84.
(14) Fischer, G. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 1425−1436.
(15) Tanaka, H.; Kuroda, A.; Marusawa, H.; Hatanaka, H.; Kino, T.; Hoto, T.; Hashimoto, M.; Taga, T. *J. Am. Chem. Soc.* **1987**, *109*, 5031−5033.
(16) Sehgal, S. N.; Baker, H.; Vezina, C. *J. Antibiot.* **1975**, *28*, 721−726.
(17) Vezina, C.; Kudzelski, A.; Sehgal, S. N. *J. Antibiot.* **1975**, *28*, 727−732.
(18) Siekierka, J. J.; Hung, S. H.; Poe, M. *Nature* **1989**, *341*, 755−757.
(19) Harding, M. W.; Galat, A.; Uehling, D. E. *Nature* **1989**, *341*, 758.
(20) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(21) Davis, M. E.; Madura, J. D.; Luty, B. A.; McCammon, J. A. *Comput. Phys. Commun.* **1991**, *62*, 187−197.
(22) Madura, J. D.; Davis, M. E.; Gilson, M. K.; Luty, B. A.; Wade, R. C.; McCammon, J. A. *Rev. Comput. Chem.* **1994**, *5*, 229−267.



A B

Figure 2. (A) Structure of (2*S*)1-acetylprolinemethylester (ACPM) (top), 1-formylpiperidine (FOPI) (middle), and 1-piperidinecarboxamide (PICA) (bottom). On the left the partial charges (see text) are given, and on the right the atom numbering is indicated. (B) Structure of FK506, a natural ligand of FKBP12.

Assignment of the ligand's resonances was achieved from 2D TOCSY and 2D NOESY spectra acquired on 100 mM solutions of Ac-Pro-OMe, 1-piperidinecarboxamide, and 1-formylpiperidine in aqueous buffer (0.1 M sodium phosphate, pH = 6.5).

Protein resonances and intermolecular NOEs were assigned from 2D TOCSY and 2D NOESY recorded on NMR samples containing 1 mM FKBP12 in aqueous buffer (100 mM sodium phosphate in 90% H_2O/10% D_2O and 0.01% NaN_3, pH 6.5) and an excess of ligand molecules. Ligand-to-protein ratios were 60/1 for ACPM, 120/1 for FOPI, and 120/1 for PICA. The NMR experiments employed a multiple solvent suppression scheme allowing for the detection of weak intermolecular NOE interactions.[7] An additional relaxation (T2) filtering showed to be not feasible due to significant losses of the signal caused by passive scalar couplings. This problem was linked to the ring structure of the ligands that led to complex scalar coupling networks. The assignment of protein resonances and structure determination of the complexes followed a procedure similar to that proposed by Dalvit et al.[7] in a study of complexes of DMSO with FKBP12.

NMR spectra were processed on Indigo SGI workstations using XWINMR (Bruker AG, Karlsruhe) and Felix (Accelrys Inc., San Diego) softwares.

Structure Analysis from NMR Data. Intermolecular NOE cross-peaks were identified in 2D NOESY spectra with a NOE mixing time of 100 ms and translated into distance restraints with an upper limit of 4.5 Å. Intermolecular restraints were introduced as ambiguous restraints if degenerate protons were involved or the protein resonance could not be unambiguously assigned.[24]

Structure calculations were performed using XPLOR software[25] from the MSI 97.2 package (Accelrys Inc., San Diego). Models of the three ligands were built in InsightII and were minimized using the CVFF force field. The program XPLO2D[26] was then used to generate the parameter and topology files for the three ligands. The files were edited manually to introduce protons. All dihedral angles of Ac-Pro-OMe were allowed to rotate freely except for the one defining the amide bond which was fixed in the trans conformation. The piperidine ring was fixed in a chair conformation, and the amide and urea bonds of 1-piperidinecarboxamide and 1-formylpiperidine were kept planar.

Protein coordinates were taken from the X-ray structure of FKBP12 complexed to a small ligand[27] (1FKG) (the rotamase inhibitor (1*R*)-1,3-diphenyl-1-propyl (2*S*)-1-(3,3-dimethyl-1,2- dioxopentyl)-2-piperidinecarboxylate) and kept fixed throughout the structural analysis. Starting structures for the complex were generated by placing the ligands in random orientations with respect to FKBP12. A standard simulated annealing protocol[28] was applied to generate ligand conformations that satisfied the intermolecular restraints.

Molecular Modeling. System Setup. The crystal structure of human FKBP12 bound to FK506 (PDB[29] entry code 1FKJ) was used for the MCSS calculations. We note that it is a different structure from that employed in the experimental analysis (see above). However, it is known that the conformation of FKBP12 is generally little altered by ligand binding. The RMSD for the backbone atoms between 1FKJ and 1FKG is 0.51 Å. In any case, use of slightly different structures is a test of the robustness of the method. The structures were solved at 1.7 and 2.0 Å resolution for 1FKJ and 1FKG, respectively, and refined to an *R* factor of 0.162 and 0.184 for 1FKJ and 1FKG, repectively.[27,30] The coordinates of the hydrogen atoms were added.[31] Experimental NMR data[32] indicate that the three histidines are singly protonated at pH 7, and the single proton was placed on Nδ. Other titrable residues were assigned the standard protonation states (i.e., all Glu, Asp, and the C-ter were deprotonated, and all Lys, Arg, and the N-ter were

(23) Sich, C.; Improta, S.; Cowley, D. J.; Guenet, C.; Merly, J.-P.; Teufel, M.; Saudek, V. *Eur. J. Biochem.* **2000**, *267*, 5342−5354.
(24) Nilges, M. *J. Mol. Biol.* **1995**, *245*, 645−660.
(25) Brünger, A. T. In ?, 3.1 ed.; Yale University Press: New Haven, CT, 1993.
(26) Kleywegt, G. J.; Jones, A. T. *Methods Enzymol.* **1997**, *277*, 208−230.
(27) Holt, D. A.; Luengo, J. L.; Yamashita, D. S.; Oh, H. J.; Konialian, A. L.; Yen, H. K.; Rozamus, L. W.; Brandt, M.; Bossard, M. J.; Levy, M. A.; Eggleston, D. S.; Liang, J.; Schultz, L. W.; Stout, T. J.; Clardy, J. C. *J. Am. Chem. Soc.* **1993**, *115*, 9925−9938.
(28) Nilges, M.; Gronenborn, A. M.; Clore, G. M. *FEBS Lett.* **1988**, *239*, 129−136.
(29) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J.; Meyer, E. E. J.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535−542.
(30) Wilson, K. P.; Yamashita, M. M.; Sintchak, M. D.; Rotstein, S. H.; Murcko, M. A.; Boger, J.; Thomson, J. A.; Fitzgibbon, M. J.; Navia, M. A. *Acta Crystallogr., Sect. D* **1995**, *51*, 511.
(31) Brunger, A. T.; Karplus, M. *Proteins* **1988**, *4*, 148−156.
(32) Yu, L.; Fesik, S. W. *Biochim. Biophys. Acta* **1994**, *1209*, 24−32.

protonated). The unliganded protein was minimized with the CHARMM program,[20] to an energy gradient tolerance of 0.01 kcal/mol Å to remove bad contacts. A constant dielectric and decreasing harmonic constraints on backbone and side-chain atoms were used. The RMS difference between the crystal and the minimized structure is 0.81 Å for the backbone atoms. The parameters from the polar hydrogen set (PARAM19)[33] were used for the protein. The minimized protein structure was fixed in the MCSS calculations; this corresponds to what was done in the NOE analysis. For small ligands such as the ones considered here, no major changes in the protein conformation are expected.

Multiple Copy Simultaneous Search (MCSS). The MCSS method determines energetically favorable positions and orientations (local minima of the potential energy) of functional groups on the surface of a protein or receptor of know three-dimensional structure.[12] The library of standard MCSS groups was supplemented with (2*S*)1-acetylprolinemethylester (ACPM), 1-formylpiperidine (FOPI), and 1-piperidinecarboxamide (PICA) (Figure 2).

Parameters for the ligands were obtained using the MMFF force field.[34] The MMFF partial charges used for each group are given in Figure 2. The van der Waals and internal energy parameters were adapted from the CHARMM force field. Each group was built in the 2DSketcher module in Quanta96, and minimized with the Quanta96 version of CHARMM using the MMFF force field. The minimized ligand structures were then used as starting geometries for MCSS.

For each ligand, 500 replicas were randomly distributed in a sphere of 10 Å radius with its center located in the middle of the binding pocket (defined as the mirror point of Leu74 Hδ13 with respect to Trp59 Cϵ2 atom). The sphere is sufficient to enclose the entire hydrophobic binding pocket and the surface residues near the binding pocket. A minimal distance of 1.2 Å between the atoms of the groups and those of FKBP12 was used for the random-distribution procedure to avoid steric clashes in the MCSS procedure. The CHARMM PARAM19 force field was used. We used a distance-dependent dielectric constant ($\epsilon = 4r$) for MCSS rather than the default vacuum value.[35] The 500 replicas were then simultaneously minimized, using 800 steps of steepest descent followed by 20 cycles of 500 steps of conjugate gradient each.[20,36] Positions were compared after the steepest descent minimization and every 500 steps of conjugate gradients to eliminate replicas converging toward a common minimum, so as to speed up the calculations. The criteria used to characterize a common minimum are an RMS deviation of 0.2 Å or less between two replicas and a decreasing RMS distance in the final 200 steps.[12] A convergence criterion of 0.0001 kcal/(mol Å) for terminating the minimization was used. Replicas with interaction energy with the protein above a given cutoff were discarded. This energy cutoff is set to a high value for the first minimizations (500 kcal/mol) and is reduced steadily during the following minimization cycles (in the last minimization, this cutoff was set to 3 kcal/mol); for 1-formylpiperidine (FOPI), for example, 7 groups out of 500 were discarded in the first series of minimizations by the energy cutoff, and 13 were eliminated in the last iteration. The procedure is then repeated with a new series of 500 randomly distributed minima with an identical minimization setup, except that minima converging to positions already obtained in previous cycles are eliminated. We repeated the procedure 20 times, so that a total of 10 000 replicas is used for each ligand. The MCSS docking procedure step took 604 min for ACPM, 283 min for FOPI, and 647 min for PICA on a Pentium III 667 MHz.

MCSS gives a ranking of the minima according to the sum of the internal energy of each group and the fragment−protein intermolecular interaction energy. This takes no account of solvation other than the crude correction introduced through the distance-dependent dielectric. In this study, MCSS was used to determine exhaustively the positions and orientations of groups in the binding pocket, and a postprocessing procedure was implemented to rank them in terms of an approximate relative free energy of binding (see below).

Postprocessing. The postprocessing of the MCSS minima consisted of the computation of the change in internal energy of the fragment on binding (ΔE^{fragm}), the van der Waals interaction between each minimum and the protein ($\Delta E^{\text{intern}}_{\text{vdW}}$), the loss in solvent-accessible surface area of both the protein and the functional group upon complexation ($\Delta G^{\text{complex}}_{\text{np}}$), and the electrostatic contribution to binding ($\Delta G_{\text{elect binding}}$), which includes the electrostatic interaction between each group and the protein ($\Delta G^{\text{interm}}_{\text{elect}}$), and the desolvation cost for the protein and each group ($\Delta G^{\text{protein}}_{\text{elect desolv}}$ and $\Delta G^{\text{fragm}}_{\text{elect desolv}}$), see eq 4. Thus, for every protein−MCSS minimum complex, the binding free energy was approximated by the use of the following equation:[37]

$$\Delta G_{\text{binding}} = \Delta E^{\text{fragm}} + \Delta E^{\text{interm}}_{\text{vdW}} + \Delta G^{\text{complex}}_{\text{np}} + \Delta G_{\text{elect binding}} \qquad (1)$$

The first term on the right side represents the change in internal energy of the fragment upon binding; it can be decomposed as follows:

$$\Delta E^{\text{fragm}} = \Delta E^{\text{fragm}}_{\text{bonded}} + \Delta E^{\text{fragm}}_{\text{vdW}} + \Delta E^{\text{fragm}}_{\text{el}} \qquad (2)$$

The CHARMM force field was used to compute ΔE^{fragm}, which is the sum of the change in bonding (bonds, angles, and torsion) energy terms ($\Delta E^{\text{fragm}}_{\text{bonded}}$), van der Waals ($\Delta E^{\text{fragm}}_{\text{vdW}}$), and Coulombic vacuum energies of the fragment ($\Delta E^{\text{fragm}}_{\text{el}}$). It was also used to compute $\Delta E^{\text{interm}}_{\text{vdW}}$, the van der Waals interaction energy between the protein and each MCSS minimum. The solvation free energy is expressed as a sum of separate electrostatic and nonpolar contributions.[38] The nonpolar contribution to the free energy of binding ($\Delta G^{\text{complex}}_{\text{np}}$) is assumed to be proportional to the loss in solvent-accessible surface area (SAS):[39]

$$\Delta G^{\text{complex}}_{\text{np}} = \gamma(SAS^{\text{complex}} - (SAS^{\text{protein}}_{\text{isolated}} + SAS^{\text{fragm}}_{\text{isolated}})) \qquad (3)$$

The constant γ, which may be interpreted as the vacuum−water microscopic surface tension, is assigned a value of 0.025 kcal/(mol Å^2).[40] The solvent-accessible surface area calculations for the protein alone, the ligands alone, and each MCSS minimum complexed with the protein have been performed with the CHARMM implementation of the Lee−Richards algorithm[39] by using a probe sphere of 1.4 Å radius.

The electrostatic contribution to the free energy of binding consists of shielded intermolecular interaction ($\Delta G^{\text{interm}}_{\text{elect}}$), protein desolvation ($\Delta G^{\text{protein}}_{\text{elect desolv}}$), and desolvation of the fragment ($\Delta G^{\text{fragm}}_{\text{elect desolv}}$). These energy values are calculated by solving the finite-difference LPB equation.[41] The calculation of the electrostatic contribution to the free energy of complex formation makes use of the program UHBD[21] that solves the Poisson−Boltzmann equation numerically on a three-dimensional Cartesian grid. The approach is based on the continuum model, where the solvent is described as a continuous dielectric with the macroscopic polarizability of bulk water, that is, a dielectric constant of $\epsilon_{\text{water}} = 80$. The interior of the protein plus the ligands is calculated as the volume that is inaccessible to a solvent probe sphere with radius 1.4 Å; this is the same as that used for the nonpolar contribution. This solvent-inaccessible volume is assigned the vacuum dielectric constant of $\epsilon_{\text{solute}} = 1$, which is consistent with the partial charges of the

(33) Neria, E.; Fischer, S.; Karplus, M. *J. Chem. Phys.* **1996**, *105*, 1902−1921.
(34) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490−519.
(35) Caflisch, A.; Schramm, H. J.; Karplus, M. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 161−179.
(36) Hestenes, M. R.; Stiefel, E. *J. Res. Natl. Bur. Stand.* **1952**, *49*, 409−436.
(37) Caflisch, A. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 372−396.
(38) Sitkoff, D.; Sharp, K. A.; Honig, B. *J. Phys. Chem.* **1994**, *98*, 1978−1988.
(39) Lee, B.; Richards, F. M. *J. Mol. Biol.* **1971**, *55*, 379−400.
(40) Chothia, C. *Nature* **1974**, *248*, 338−339.
(41) Luty, B. A.; Davis, M. E.; McCammon, J. A. *J. Comput. Chem.* **1992**, *13*, 768−771.

CHARMM force field.[33,42,43] The principal benefit of the use of the finite-difference Poisson−Boltzmann (FDBP) approach in the framework of ligand binding free-energy calculations, relative to the standard vacuum energy function, is that it accounts for the solvent screening of intramolecular charge−charge interactions and for the desolvation of polar and ionized groups upon binding. For simplicity, the linearized form of the Poissson-Boltzmann equation was solved, where the spatial distribution of ions in solution is approximated by a linear function of the solute potential.[44] The ionic strength was set to 145 mM.

For each protein−ligand complex, a series of FDPB calculations needs to be performed to calculate the electrostatic binding free energy as described previously:[35,37]

$$\Delta G_{\text{elect binding}} = \Delta G^{\text{interm}}_{\text{elect}} + \Delta G^{\text{protein}}_{\text{elect,desolv}} + \Delta G^{\text{ligand}}_{\text{elect,desolv}} \qquad (4)$$

In the FDPB calculations, it is necessary to employ a finite-difference grid with a grid constant of 0.3 Å or less to obtain results that are accurate to within ±0.5 kcal/mol.[45] For a large protein−substrate complex, for example, the system studied here, this means that on the order of 10 min of CPU time on a fast workstation is required per calculation of $\Delta G_{\text{elect binding}}$ for a single MCSS minimum.

To reduce the amount of computer time that is required for the treatment of several hundreds of MCSS minima, a novel approach was developed[45] that treats the energetically important short-range electrostatic interactions with a FDPB grid at high resolution (grid constant $a = 0.3$ Å) and weak long-range interactions at low resolution ($a = 1.0$ Å). A fully automated UNIX csh-script was written that uses the standard MCSS output files as input and that calculates a table of $\Delta G_{\text{elect binding}}$ for all minima. Details of the approach, its calculation parameters, and its accuracy are given elsewhere.[45]

It should be noted that eq 1 is designed to be valid for the relative ranking of the binding energies of the MCSS minima for a given functional group. It neglects terms, such as the configurational, rotational, and translational entropy loss on binding, that contribute to the relative binding free energies of different ligands and to the absolute binding energy, but which are expected to be very similar for a given ligand in different positions.

Clustering. The MCSS procedure produces a large number of minima for each ligand. To visualize these results more easily, the minima are clustered into families. This reduces the need to inspect hundreds of minima to the visualization of about a dozen clusters. Each cluster is represented by the minimum with the most favorable binding free energy as computed from eq 1. To classify the minima into families, we implemented an adaptation of the "minimum leap" algorithm.[46] The parameter used to classify the minima is the collection of van der Waals contacts that a given minimum makes with the protein residues. For each minimum, a van der Waals contact map or signature is produced, which consists of the amino acids that have significant van der Waals interactions with the minimum. Comparisons of van der Waals contacts maps are performed, and the minima with similar van der Waals contacts are clustered. The van der Waals energy is short ranged; thus minima with similar contact maps are also close in space. The number of clusters obtained depends on two parameters that are set by the user. One parameter is the energy threshold below which an interaction is considered significant. A large (in absolute value) energy threshold leads to a small number of clusters, and vice versa. The second parameter is the contact difference that can be varied if one wishes to allow a certain flexibility. By flexibility, we mean that if a given minimum is in contact with N protein residues, it can be placed in the same cluster as a minimum that makes contacts with a subset of these residues, typically N-1 or N-2; if all contacts are identical (i.e., contact difference equals zero), then no flexibility is allowed. When more differences are allowed in the contacts between residues of the same cluster, a smaller number of clusters is obtained. The optimal values of the energy threshold and contact differences used in the clustering are empirically determined. In this study, we allowed a difference in contacts maps of two ligand−protein interactions, and we used a van der Waals energy threshold of −1.0 kcal/mol for the three ligands.

The contact-energy-based clustering approach has several advantages over the more standard distance-based clustering in that it is trivial to extend to calculations where protein flexibility is explicitly taken into account. Also, the van der Waals clustering can be used in a straightforward manner to compare the positions of different ligands, though care has to be exercised if the molecules have different sizes.

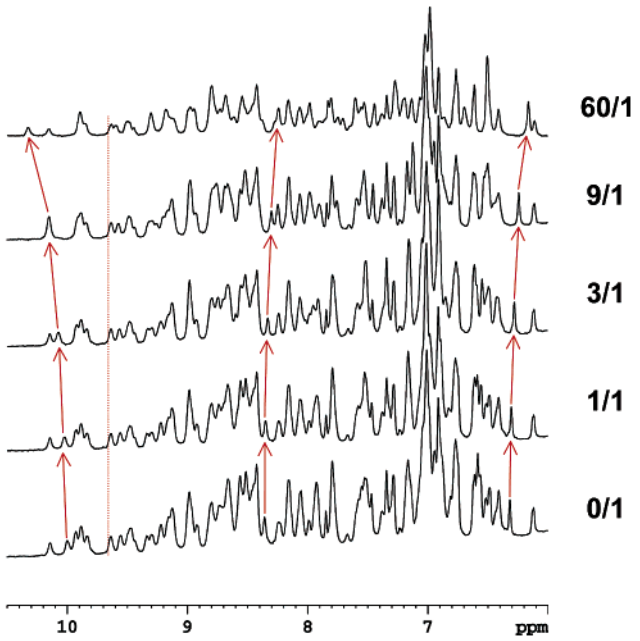


Figure 3. Changes in the downfield region of the proton spectrum in the course of titrating Ac-Pro-OMe to FKBP12. The ligand was first dissolved in aqueous buffer (0.1 M sodium phosphate, pH 6.5) and added to the protein solution as small aliquots (0.5−20 μL). Molar ratios between Ac-Pro-OMe and FKBP12 are given on the right of each spectrum. Progressive shifting of three isolated protein peaks is marked by red arrows. The dashed line serves as reference for a resonance that is not shifting.

Results and Discussion

NMR Results. Addition of the three ligands to the FKBP12 solution resulted in chemical shift changes for a limited number of the protein resonances (Figure 3). This suggests a specific interaction between FKBP12 and the three ligands.

Dissociation equilibrium constants (K_D) for Ac-Pro-OMe, 1-formylpiperidine, and 1-piperidinecarboxamide were determined by monitoring the protein chemical shift changes as a function of ligand concentration. Data were collected for a set of resolved FKBP12 ^{1}H NMR resonances and fitted to a single binding site model according to the following equation:

$$[\mathrm{L}]_0 = \frac{\Delta\delta_{\mathrm{M}}[\mathrm{P}]_0}{\Delta\delta} - K_{\mathrm{d}} \qquad (5)$$

where $[\mathrm{L}]_0$ and $[\mathrm{P}]_0$ are the total ligand and protein concentration, respectively, and $\Delta\delta_{\mathrm{M}}$ is the chemical shift difference of the fully bound form with respect to its position in the unbound

(42) Schlenrich, M.; Brickmann, J.; MacKerell, A. D., Jr.; Karplus, M. In *Biological Membranes. A Molecuar Perspective from Computation and Experiment*; Merz, K. M., Roux, B., Eds.; Birkhauser: Boston, 1996; pp 31−81.
(43) Foloppe, N.; MacKerell, A. D., Jr. *J. Comput. Chem.* **2000**, *21*, 86−104.
(44) Davis, M. E.; McCammon, J. A. *Chem. Rev.* **1990**, *90*, 509−521.
(45) Schaefer, M.; Zoete, V.; Karplus, M., manuscript in preparation.
(46) Lebart, L.; Morineau, A.; Fénelon, J.-P. *Traitement des données statistiques*; Dunod: Paris, 1982.

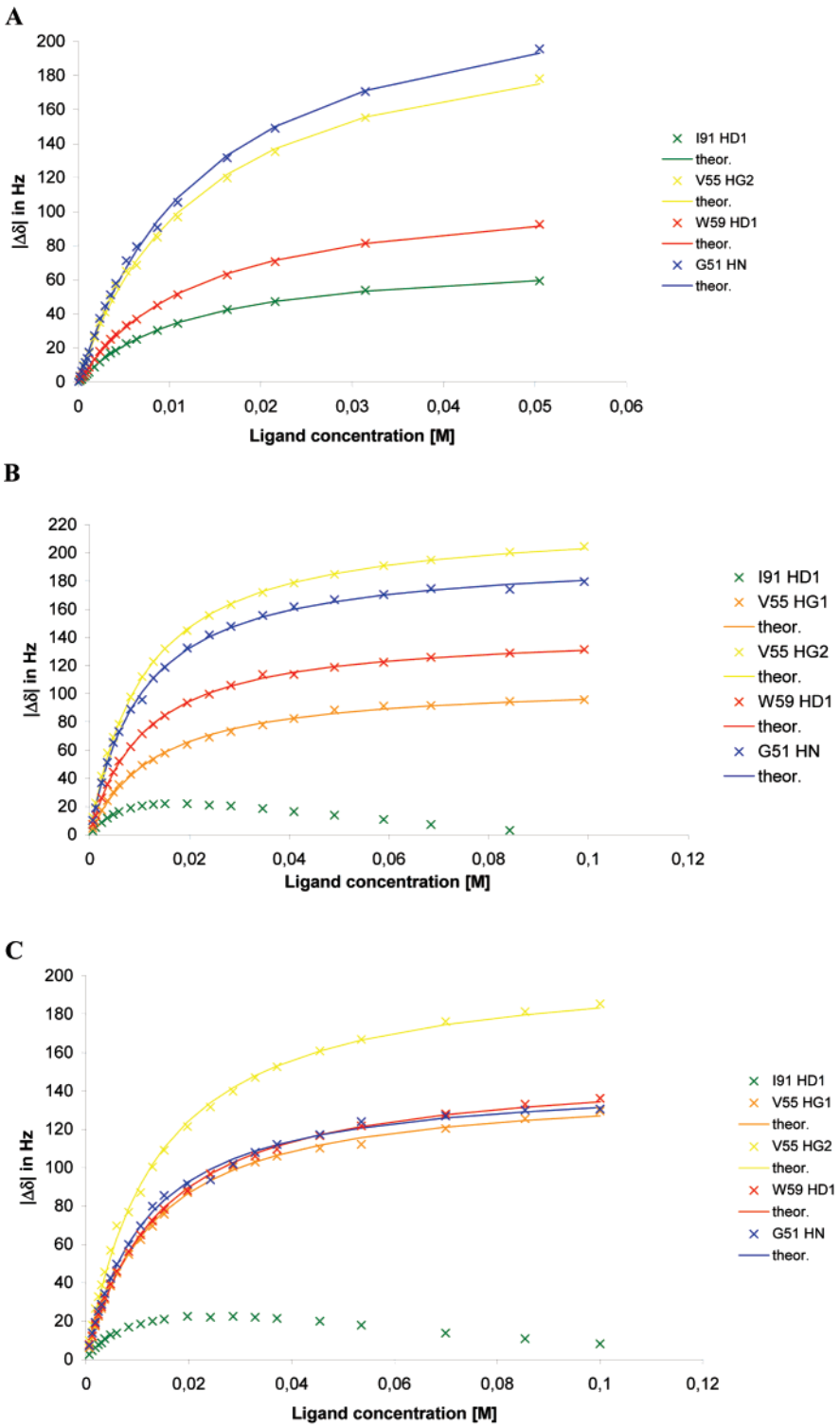


Figure 4. Chemical shift changes of FKBP12 resonances as a function of ligand concentration. The experimental data were fitted to eq 1, and the resulting curves are presented. (A) Titration of Ac-Pro-OMe. Individual fitting yielded a $K_D = 11.6$ mM, $\Delta\delta_M = 74$ Hz for Ile 91H$^{\gamma 1}$, 12.6 mM, 220 Hz for Val55 H$^{\gamma 2}$, 12.7 mM, 117 Hz for Trp59 H$^{\gamma 1}$, and 13.1 mM, 244 Hz for Gly51 HN. The average value of the dissociation equilibrium constant is $K_D = 12.5 \pm 0.6$ mM. (B) Titration of 1-formylpiperidine. Individual fitting yielded a $K_D = 12.2$ mM, $\Delta\delta_M = 108$ Hz for Val55 H$^{\gamma 1}$, 10.0 mM, 224 Hz for Val55 H$^{\gamma 2}$, 10.0 mM, 144 Hz for Trp59 H$^{\delta 1}$, and 9.5 mM, 198 Hz for Gly51 HN. The average value of the dissociation equilibrium constant is $K_D = 10.5 \pm 1.2$ mM. The experimental values of Ile91 H$^{\delta 1}$ could not be fitted to a single dissociation equilibrium constant. (C) Titration of 1-piperidinecarobxamide. Individual fitting yielded a $K_D = 12.6$ mM, $\Delta\delta_M = 143$ Hz for Val55 H$^{\gamma 1}$, 12.9 mM, 207 Hz for Val55 H$^{\gamma 2}$, 13.6 mM, 153 Hz for Trp59 H$^{\delta 1}$, and 11.1 mM, 146 Hz for Gly51 HN. The average value of the dissociation equilibrium constant is $K_D = 12.6 \pm 1.0$ mM. The experimental values of Ile91 H$^{\delta 1}$ could not be fitted to a single dissociation equilibrium constant.

form. As shown in Figure 4, most of the titration curves could be fitted to a single dissociation constant, giving average K_D values of 12.5 ± 0.6, 12.6 ± 1.0, and 10.5 ± 1.2 mM for Ac-Pro-OMe, 1-formylpiperidine, and 1-piperidinecarboxamide, respectively. However, eq 5 cannot reproduce the biphasic behavior exhibited by the chemical shift changes of Ile91 H$^{\delta 1}$

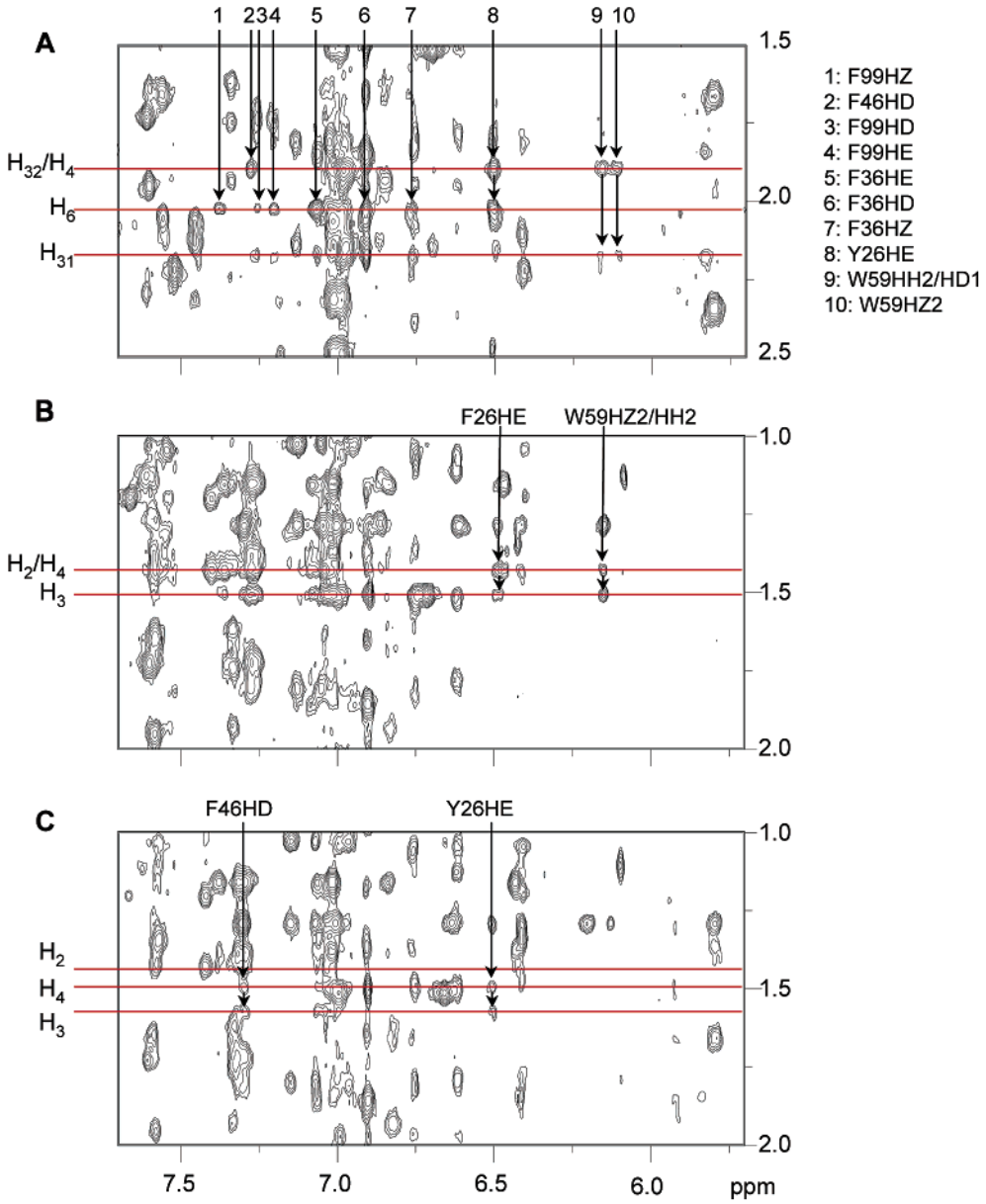


Figure 5. 2D NOESY spectra showing NOE cross-peaks between ligand resonances and aromatic resonances of FKBP12. The mixing time was 100 ms. Assignments of ligand resonances are given on the left and of protein resonances on the top of the spectra. (A) Ac-Pro-OMe FKBP12 NOEs, (B) 1-piperidinecarboxamide FKBP12 NOEs, and (C) 1-formylpiperidine FKBP12 NOEs.

upon titration with 1-piperidinecarboxamide and 1-formylpiperidine. These data suggest the existence of a second binding site close to Ile91.

The ^{1}H chemical shifts of Ac-Pro-OMe, 1-formylpiperidine, and 1-piperidinecarboxamide in aqueous solution were readily assigned. Two sets of resonances corresponding to the trans and cis isomers were identified for Ac-Pro-OMe; however, only the trans form was found to interact with the protein.

Assignment of the ^{1}H chemical shifts of free FKBP12 has been reported[47] and was extended to the resonances that exhibited significant changes in the presence of the ligands.

Figure 5 shows regions of NOESY spectra containing intermolecular cross-peaks observed for Ac-Pro-OMe, 1-formylpiperidine, and 1-piperidinecarboxamide. Because of extensive spectral overlap, intermolecular NOEs could be detected only with the most upfield shifted signals of FKBP12 and the aromatic resonances. These NOEs indicated ligand contacts with residues located in the active site of the protein. The majority of the distance restraints were consistent with one binding mode for each ligand. However, the observation of intermolecular NOEs with Ile91 $H^{\delta 1}$ seemed to identify a second binding site for 1-formylpiperidine and 1-piperidinecarboxamide (see also titration curves). In the absence of other NOEs that could locate a second ligand molecule, the restraints involving Ile91 $H^{\delta 1}$ were not included in the final structure calculations.

Views of Ac-Pro-OMe, 1-formylpiperidine, and 1-piperidinecarboxamide in the protein binding site are shown in Figure 6A, B, C, respectively. Six representative conformations consistent with the NMR data are displayed for each ligand. The three molecules bind in the hydrophobic pocket formed by Tyr26, Phe46, Val55, Ile56, Trp59, and Phe99. In particular, the location of the pyrrolidine and piperidine rings is identical

(47) Rosen, M. K.; Michnick, S. W.; Karplus, M.; Schreiber, S. L. *Biochemistry* **1991**, *30*, 4774−4789.

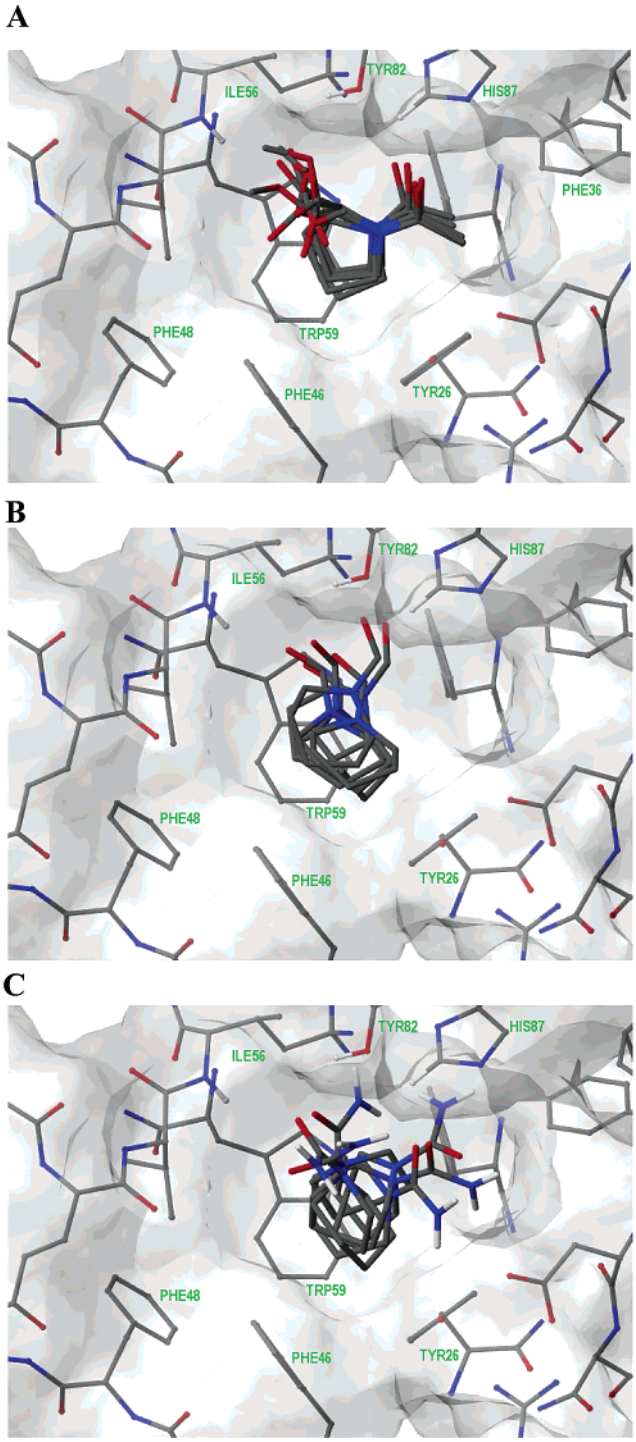


Figure 6. Minimum energy structures obtained by NMR; six structures are presented for each ligand. (A) Ac-Pro-OMe, (B) 1-formylpiperidine, (C) 1-piperidinecarboxamide bound to FKBP12.

to that of the pipecolinyl part of FK506 in the immunosuppressant/immunophilin complex.

The NMR structures of Ac-Pro-OMe exhibit very little dispersion, except for the ester group that cannot be constrained by NOEs. Despite its undetermined orientation, the O10 carbonyl oxygen was found within hydrogen-bonding distance of the amide group of Ile56 in many structures. Hydrogen-bonding interactions could also be formed between the O8 oxygen of Ac-Pro-OMe and the hydroxyl group of Tyr82. Interestingly, these interactions have their equivalents in the complex with the macrocyclic inhibitor.

The exact orientation of 1-piperidinecarboxamide in the FKBP12 binding site could not be determined due to the lack of restraints on the carboxamide group. However, in all NMR-derived structures, the hydrophilic part of the ligand points toward the outside of the protein cleft.

Two major orientations, parallel or perpendicular to the Trp59 side chain, were found for the piperidine ring of 1-formylpiperidine bound to FKBP12. The formyl group is preferentially oriented toward the surface of the binding pocket, being anchored by one intermolecular NOE between H7 and Ile56 $H^{\gamma 2}$.

Computational Results. In describing the performance of the computational predictions, two aspects need to be considered: the first is whether the docking algorithm is able to correctly position ligands in the macromolecule binding pocket, and, second, because the docking algorithm gave several positions for the ligand, an essential aspect is to identify the most favorable positions with an approximate free-energy function and compare the results with experiment. These two aspects are referred to as docking and scoring, respectively.

Docking Results. The MCSS method is devised to identify a very large number of putative binding sites of the ligand−macromolecule complex (see Table 1). These structures represent local energy minima determined with a vacuum force field and a rigid model for the protein. Both of those simplifications make it possible to rapidly screen a large number of ligands for possible docking sites.

In Figure 7, we summarize the docking results by presenting one docked position for each cluster of minima (see Materials and Methods), that is, the position with the most favorable binding free energy (from eq 1). One representative experimental NMR position is also displayed for comparison. For each of the calculated positions displayed in Figure 7, the violations of NOE restraints are given in Table 2. It can be seen that for all ligands, at least one docked position satisfies the NOE restraints. The MCSS method thus performs well in identifying correct ligand positions in the binding pocket.

Scoring Results: Relative Values of Binding Free Energies of MCSS Minima. By relative values of binding free energies, we mean the comparison of the values of free energies for the same ligand occupying different positions in the binding pocket and not the comparison of different ligands. Correct relative energies thus mean that the scoring procedure is able to identify a structure that is close to the experimentally determined one. As noted in the Materials and Methods section, the approximate binding free energy given by eq 1 neglects configurational entropy terms and is not devised to compute absolute values of binding free energies.

As we describe in more detail below, an essential aspect of the scoring free-energy function is the inclusion of the effect of solvation (cf. eq 1). Figure 7 shows that the computational results including solvation agree with the NMR data in placing the ligands in the hydrophobic pocket on top of the TRP 59, which is comparable to the position of the pipecolinyl ring of FK506 when bound to FKBP12.

The general features of the MCSS minima are summarized in Table 1. For ACPM, five clusters are situated inside the FKBP known binding pocket, and the binding free energies (from eq

Table 1. Functional Groups Used for MCSS[a]

group[b]	electrostatic solvation free energy[c]	CHARMM energy[d] lowest	CHARMM energy[d] highest	no. of minima[e]	no. of minima with favorable $\Delta G_{binding}$ (<0)[f]	$\Delta G_{binding}$[f] lowest	2nd	3rd	highest
ACPM	−14.4	−54.6	2.6	397	345	−19.2	−18.8	−17.4	22.7
FOPI	−9.2	−18.1	3.0	281	254	−17.0	−16.6	−15.9	13.2
PICA	−14.0	−10.0	2.9	58	51	−14.5	−14.1	−13.3	13.7

[a] All energy values are in kcal/mol. [b] ACPM = 1-acetylprolinemethylester; FOPI = 1-formylpiperidine; PICA = 1-piperidinecarboxamide. [c] Calculated by numerical solution of the linear Poisson−Boltzmann equation. [d] The CHARMM energy is the sum of intermolecular and intraligand energies. [e] With CHARMM energy lower than 3.0 kcal/mol (see text). [f] Calculated by use of eq 1.

1) for the best minima from these clusters are in the range from −19 to −10 kcal/mol. These binding energies are more favorable than those for the clusters situated outside the binding pocket (which are in the range from −10 to −5 kcal/mol). For FOPI, four clusters are situated inside the pocket (with binding energies for the best minima in the range from −18 to −14 kcal/mol), and four clusters are found outside the pocket (with binding energies in the range from −12 to −8 kcal/mol). For PICA, four clusters are found inside the pocket, with binding free energies for the best binder of each cluster in the range from −14 to −10 kcal/mol. Five clusters are found outside the pocket, with energies in the range from −9 to −6 kcal/mol (with the exception of one cluster situated at the entrance of the binding pocket for which the best binder has an energy of −12.5 kcal/mol). Thus, the minima representative of the clusters situated inside the binding pocket have better binding free energies than those of the minima situated outside of the binding pocket.

If solvation contributions are neglected (i.e., only the CHARMM force-field energy is used[12]), the scoring of the minima is significantly modified, and the most favorable positions are outside of the binding region identified by NMR (Figure 1 of the Supporting Information). The importance of solvation−desolvation effects in ligand binding has been demonstrated in a number of papers,[48−50] and this study is another striking example of its effect on the results of docking predictions.

We now examine in more detail the computational predictions for the ligands situated inside the binding pocket. We define a ligand as correctly docked if it satisfies the NOE constraints. It can be seen from Table 2 that, in all cases, the docked ligand position that satisfies the NOE constraints is ranked in the 10 top scorers (rank 7 for ACPM, 5 for FOPI, 1 and 3 for PICA). Table 2 also shows that the RMS deviation between the top scoring position and the position that satisfies the NOEs is 3.2 Å for ACPM and 2.3 Å for FOPI, so that even when the top scoring position violates the NOEs, it is still reasonably docked.[51−53]

As can be seen in Figure 2, the three ligands have a hydrogen-bonding moiety (H-bond acceptor carbonyl group for all three ligands plus two additional acceptor sites for ACPM and donor site for PICA) and a hydrophobic moiety. For all three ligands, both the top scoring position and the position that satisfies the NOEs form a hydrogen bond between one carbonyl group of the ligand and the main chain NH of isoleucine 56. This H-bond is similar to the one formed between the FK506 C1 ketone and Ile56-NH in the FK506−FKBP12 complex. Thus, the computational procedure identifies this important interaction.

For ACPM, the main difference between the top ranked position and the position that satisfies the NOEs constraints is that the former has a second H-bond with Tyr82 Hη, but has less favorable van der Waals interactions with the protein (the proline ring is perpendicular to the aromatic plane of Trp59, making van der Waals contacts with Val55, Ile56, and Trp59), while the position that satisfies the NOEs makes only one H-bond but has more favorable van der Waals interactions (the proline ring is parallel to Trp 59 making contacts to Ile 56, Trp 59, and Tyr 82 but also specifically with Tyr 26, Phe 46, and Phe 99). These differences in position correspond to a fairly large difference in binding energy: −19 kcal/mol for the top scorer versus −14 kcal/mol for the position that satisfies the NOEs.

For 1-formylpiperidine (FOPI), three clusters are found in the main binding pocket; they all form one H-bond to Ile 56 NH. The minima ranked best in these three clusters are close in energy (−17, −16, and −15 kcal/mol; cf. Table 2); they differ mainly in the orientation of the pipecolinyl ring (minima labeled 1, 5, 8, respectively, in Figure 7B). Contrary to what was seen for ACPM where one could clearly see a shift in the balance of interactions between minima 1 and 7, in the case of FOPI it is clear from Table 2 that all of the contributions to the calculated binding are similar for minima 1, 5, and 8. Given the approximate nature of the computed values, the three minima can thus be considered as equivalent from a computational standpoint, so that the scoring function has performed well in identifying the relevant position.

Piperidine carboxamide (PICA) can be either a hydrogen bond donor with its amide or a hydrogen bond acceptor with its carbonyl. It can also form hydrophobic interactions with the piperidine ring. The position of PICA is the least well defined by the NMR data, because of the lack of constraints on the carboxamide group. The MCSS calculations propose a series of orientations for the carboxamide, with different H-bonding patterns, which are equally compatible with the NMR constraints on the piperidine ring. This explains the larger number of minima satisfying the NOE constraints (cf. Table 2) for this ligand; three out of the four clusters situated inside the binding pocket satisfy the NOE constraints. The energetically most favorable minimum forms a hydrogen bond between its carbonyl group and Ile56 NH and also donates a hydrogen bond to Glu54 CO with the amide group. The best binders of the second cluster accept a hydrogen bond with their carbonyl from Tyr82 Hη and donate a hydrogen bond with one of their NH to Glu54

(48) Zou, X.; Sun, Y.; Kuntz, D. *J. Am. Chem. Soc.* **1999**, *121*, 8033−8043.
(49) Majeux, N.; Scarsi, M.; Caflisch, A. *Proteins* **2001**, *42*, 256−268.
(50) Caflisch, A.; Fischer, S.; Karplus, M. *J. Comput. Chem.* **1997**, *18*, 723−743.
(51) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.* **1997**, *267*, 727−748.
(52) Vieth, M.; Hirst, J. D.; Dominy, N.; Daigler, H.; Brooks, C. L., III. *J. Comput. Chem.* **1998**, *19*, 1623−1631.
(53) Vieth, M.; Hirst, J. D.; Kolinski, A.; Brooks, C. L., III. *J. Comput. Chem.* **1998**, *19*, 1612−1622.

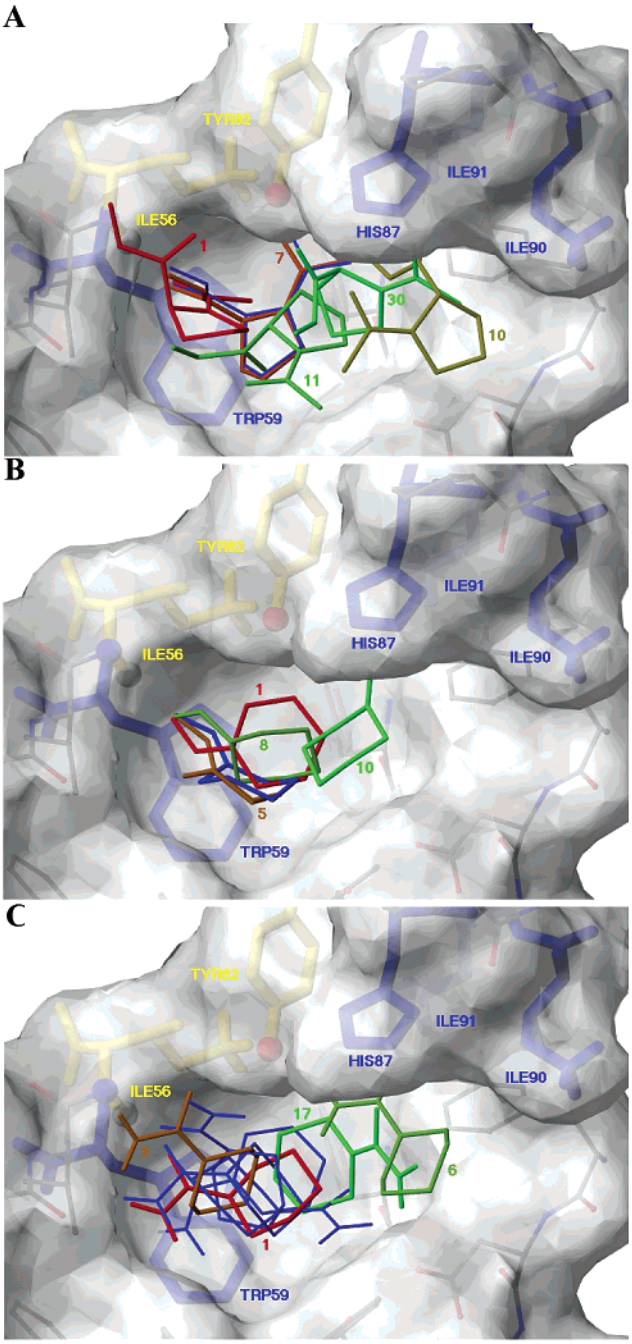


Figure 7. Minimum energy structures obtained by the MCSS procedure. The minima have been grouped in clusters (see text), and each cluster is represented by the structure with the most favorable free energy of interaction. The MCSS minima are colored from green (weak binding) to red (tight binding), and the overall ranking of each minimum is shown. For the sake of clarity, clusters situated outside of the binding pocket are not represented. The NMR position with the smallest NOE violations is shown in blue for comparison. (A) Ac-Pro-OMe (ACPM), (B) 1-formylpiperidine (FOPI), (C) 1-piperidinecarboxamide (PICA).

CO. Both clusters are situated on top of TRP59, with different orientations of the piperidine ring (cf., Figure 7C). These two clusters have similar calculated binding free energies and similar violations of the NOEs. A third cluster still satisfies the NOEs but shows poorer ranking. Its best minimum (#17; cf. Figure 7C) donates an H-bond to Asp37 Oδ2 with an amide hydrogen and is found between the main binding site and the side pocket. This position is less likely than the other two.

Scoring Results: Hydrophobic Subpocket. As can be seen from Table 2, for all ligands the docking procedure also identifies a cluster with favorable energies that shows large NOE violations and a large RMS deviation (>6 Å) with respect to the experimental position. These clusters are all situated in a hydrophobic subpocket, lined with Tyr82, His87, Ile90, and Ile91, and not in the main binding site. As the NMR structure refinement used only the NOEs from the main binding pocket, large violations are expected. In the case of FOPI and PICA, the variation of chemical shift upon addition of increasing quantities of the ligand exhibits biphasic behavior (cf., Figure 4B, C). This indicates the presence of two binding sites. Moreover, intermolecular NOEs with Ile91 H$^{\delta 1}$, which is situated in the side pocket, were observed for FOPI and PICA (but not used in the simulated annealing refinement, see above).

This subpocket is used by known ligands of FKBP12; it is filled by the C11 methyl group of the pyranose cycle of FK506[30] and by equivalent groups in rapamycin[54] and 28-*O*-methylrapamycin.[55] For FOPI and ACPM, the minima found in this pocket do not form H-bonds to the protein and would thus not have been found by algorithms that rely on H-bond complementarity to dock ligands. For PICA, the best binder in the subpocket donates one of its NH to Tyr82 Oη and makes hydrophobic contacts with Phe36, His87, Ile90, and Ile91.

From the computational results, it is clear that if a ligand is in the subpocket (for example, minimum #10 for FOPI, Figure 7B, or minimum #6 for PICA, Figure 7C), placing simultaneously a ligand in the main pocket is difficult for steric reasons, but still possible. For FOPI, it can be seen from Figure 7B that minimum #5, which satisfies the NOE constraints, does allow the positioning of a second molecule in the binding subpocket. Similarly, for PICA, it is possible to accommodate minimum #3 in the main pocket simultaneously to minimum #6 in the side pocket, but not minima #1 and #6, as they would overlap (cf., Figure 7C). Thus, the computational results identified a side pocket that is known experimentally from other ligands. They also indicate which positions can be occupied simultaneously in the binding pocket.

Use of Scoring Functions in NMR Structure Refinement. We also calculated binding free energies for the experimental NMR positions. It is not always possible to obtain enough NOE restraints to uniquely define a structure. In the present case, several structures that are equally compatible with the NOEs are proposed for each ligand (cf., Figure 6A, B, C). It is thus of interest to check whether a force-field minimization followed by an evaluation using eq 1 would be useful in ruling out some of the proposed structures.

We performed steepest descent energy minimizations in the field of the fixed protein, using the experimental NMR structures as starting geometries. The energy minimization used the same force field and dielectric constant as the MCSS calculations and was stopped when the gradient of the energy was less than 0.0001 kcal/(mol Å). After the minimization, all of the structures still satisfy the NOE experimental constraints. The resulting structures were then scored using eq 1, and, using this energy, they were ranked relative to the docked positions obtained by

(54) Wilson, K. P.; Yamashita, M. M.; Sintchak, M. D.; Rotstein, S. H.; Murcko, M. A.; Boger, J.; Thomson, J. A.; Fitzgibbon, M. J.; Navia, M. A. *Acta Crystallogr., Sect. D* **1995**, *51*, 511.
(55) Kallen, J. A.; Sedrani, R.; Cottens, S. *J. Am. Chem. Soc.* **1996**, *118*, 5857−5861.

Table 2. Clusters of Minima[a]

	intermolecular		desolvation								NOE	
					electrostatic				MCSS	H-bond	violation	RMSD[i]
rank[b]	strain[c]	vdWaals[d]	elect[e]	nonpolar[f]	protein[e]	ligand[e]	Σelectrostat	$\Delta G_{binding}$[g]	rank[h]	partners	(Å)	(Å)
						ACPM						
1	1.4	−11.7	−17.0	−10.6	10.3	8.4	1.6	−19.2	34	ILE56 HN O8 TYR82 Hη O10	1.4	3.3
7	2.2	−15.4	−10.1	−11.4	10.7	10.0	10.5	−14.1	99	ILE56 HN O10	0.2	
10	2.2	−9.3	−11.7	−9.8	9.4	6.5	4.2	−12.7	49		3.8	6.9
11	0.1	−12.4	−4.6	−11.1	8.0	7.4	10.8	−12.7	228		1.5	3.4
30	3.5	−5.0	−14.2	−10.2	8.4	7.3	1.5	−10.2	47		1.2	4.6
						FOPI						
1	0.1	−10.2	−10.1	−9.3	6.0	6.6	2.4	−17.0	24	ILE56 HN O8	0.9	2.3
5	0.1	−9.7	−9.3	−8.8	4.9	6.8	2.5	−16.0	27	ILE56 HN O8	0.2	
8	0.1	−10.0	−9.3	−9.4	6.7	6.6	4.0	−15.3	36	ILE56 HN O8	1.3	2.4
10	0.1	−8.7	−11.7	−8.3	9.0	6.6	3.9	−13.0	42		3.4	6.3
						PICA						
1	0.5	−13.1	−9.3	−9.4	5.8	11.0	7.6	−14.5	14	ILE56 HN O9 GLU54 CO H11	0.4	
3	0.3	−10.6	−8.9	−9.3	9.3	6.0	6.4	−13.3	10	TYR82 NH O9 GLU54 CO H10	0.4	2.9
6	0.4	−8.1	−15.3	−8.4	8.7	9.6	3.1	−13.1	26	TYR82 OH H11	2.7	6.5
17	0.5	−10.3	−7.1	−9.9	8.0	8.4	9.3	−10.5	30	ASP37 Oδ2 H10	0.5	5.5

[a] Energy values in kcal/mol are listed for MCSS minima that represent clusters of ACPM, FOPI, and PICA found inside the FKBP12 binding pocket. [b] Ranking (according to eq 1) of the minimum with the most favorable binding energy in each cluster. [c] Sum of intraligand energy terms is calculated with CHARMM (eq 2). [d] Calculated with CHARMM. [e] Calculated by numerical solution of the linear Poisson−Boltzmann equations as explained in text. [f] Calculated by use of eq 3. [g] Calculated by use of eq 1, that is, the sum of columns 3−8. [h] Ranked among the minima of the same functional group type according to total CHARMM energy, that is, the sum of intermolecular and intraligand energies. [i] RMSD (in Å) of the minima with respect to the best ranked MCSS minimum that satisfies the NOES.

Table 3. Minimized NMR Positions[a]

	intermolecular		desolvation							NOE
					electrostatic				H-bond	violation
rank[b]	strain[c]	vdWaals[d]	elect[e]	nonpolar[f]	protein[e]	ligand[e]	Σelectrostat	$\Delta G_{binding}$[g]	partners	(Å)
						ACPM				
5	1.7	−15.9	−9.8	−11.7	11.1	9.6	11.0	−14.9	ILE56 HN O10	0.2
26	3.8	−12.3	−11.5	−11.7	10.6	10.5	9.5	−10.8	ILE56 HN O10	0.3
113	1.6	−13.2	−0.1	−11.1	10.1	6.1	16.1	−6.6	TYR82 Hη O10	0.0
156	0.8	−11.8	1.5	−11.9	10.3	5.5	17.3	−5.6		0.4
168	2.2	−14.9	3.0	−11.3	10.9	4.8	18.7	−5.3	TYR82 Hη O11	0.0
						FOPI				
5	0.1	−9.7	−9.3	−8.7	4.9	6.8	2.5	−15.9	ILE56 HN O8	0.2
8	0.1	−10.1	−8.3	−9.0	5.1	6.6	3.4	−15.6	ILE56 HN O8	0.3
74	0.4	−9.0	−1.7	−8.8	7.5	2.8	8.6	−8.8	TYR82 Hη O8	0.3
						PICA				
5	0.3	−10.6	−8.9	−9.2	9.3	6.0	6.4	−13.2	TYR82 Hη O9 GLU54 CO H10	0.4
9	0.3	−13.7	−8.8	−9.5	6.1	12.8	10.1	−12.9	ILE56 HN O9	0.5
11	0.7	−8.9	−10.9	−9.3	6.4	9.6	5.1	−12.4	GLU54 CO H11	0.3
15	0.3	−12.3	−4.0	−9.3	5.7	8.1	9.8	−11.5	ASP37 Oδ2 H11	0.4
22	0.3	−11.4	−4.9	−9.4	9.1	5.8	10.0	−10.5		0.0
40	0.1	−11.8	−5.5	−9.5	6.9	12.1	13.5	−7.7	TYR82 OH H10	0.4
43	0.6	−13.9	0.2	−9.6	6.6	9.1	15.8	−7.1	ASP37 Oδ2 H11	0.3
47	1.3	−9.9	−4.5	−9.2	7.2	8.9	11.5	−6.3	TYR82 OH H11	0.1
57	0.2	−13.2	2.5	−9.8	7.9	9.9	20.3	−2.4		0.0

[a] Energy values in kcal/mol are listed for NMR positions of ACPM, FOPI, and PICA. [b] Ranked among the MCSS minima of the same functional group according to binding free energy. [c−g] cf. Table 2.

MCSS. The results of this procedure are given in Table 3; the number in the first column gives the indicated ranking.

In several cases, different starting experimental positions converged into one unique minimum: 5 minima are obtained out of 17 NMR structures for ACPM, 3 out of 6 for FOPI, and 9 out of 18 for PICA. The minimized NMR positions coincide in several cases with minima independently identified by the MCSS procedure, which further validates the docking procedure.

The computed free energy of binding for the minimized experimental positions is given in Table 3. It can be seen that, in some cases, the minimization plus evaluation by the function of eq 1 yields structures with very different energies and suggests that some of the proposed NMR structures are unlikely; for ACPM, three minima out of five have a much poorer ranking, and for FOPI one minimum out of three.

In the case of PICA, where the position of the ligand is the less well defined by the NMR data, the postprocessing of the minimized positions narrows down the range of possible structures, but several H-bonding patterns for the carboxamide

groups still yield positions that are both ranked favorably and satisfy the NOE constraints.

Conclusion

In this paper, we compared theoretically and experimentally (NMR) determined positions for three ligands in the binding pocket of FKBP12. The theoretical results used multiple copy minimization with the CHARMM program as implemented in the MCSS method to identify binding sites for the ligands. The most favorable positions were subsequently determined by using a free-energy function that included solvation contributions (see eq 1). In the three cases, the results from the theoretical ranking compare very well with the NMR experimental results. Ligand positions that satisfy the experimental NOE constraints are ranked in the top 5%. It must be recalled that eq 1 is based solely on physical contributions to the free energy of complexation and does not involve fitting to a training set of experimental binding free-energy data. The incorporation of desolvation effects in the free-energy function is an essential element for this good performance, as was shown by the comparison with the results obtained without the desolvation correction.

A second test of the docking/scoring function was performed using the experimental NMR structures of the protein−ligand complexes as a starting point. These structures were minimized in the vacuum force field (using a distance-dependent dielectric constant), and the resulting ligand positions were evaluated by eq 1. This test showed that experimental NMR positions were close to minima of the vacuum force field, which further validates the docking method. It also showed that the free-energy function can rule out several of the proposed experimental structures and thus is useful in NMR structure refinement of complexes, particularly when only a few NOE restraints are available.

This study shows that the multiple copy minimization coupled with the binding free-energy estimate from eq 1 performs well for the purpose of identifying interesting positions for ligand binding in a drug design scheme. In an incremental ligand construction approach, the important aspect is to discriminate interesting binding positions from irrelevant ones. Low free energy binding positions in a larger ligand do not necessarily correspond exactly to the most favorable binding site of the isolated building block. The current study showed that the procedure used in this paper identified and consistently ranked in the top scorers positions in the main binding pocket, which form hydrogen bonds known to be important for identified ligands such as FK506, and placed the experimental result among the 10 top scorers in all cases. The desolvation free energy correction eliminated irrelevant positions outside the binding pocket. It also identified positions in a binding subpocket shown to be a secondary binding site by the NMR study.

The results obtained in this study suggest that binding sites identified by MCSS and ranked by postprocessing can serve as a computational analogue of the structure activity relationship (SAR) by NMR.

Acknowledgment. This work was supported by the Centre National de la Recherche Scientifique, The Institut National de la Santé et de la Recherche Médicale, and the Université Louis Pasteur de Strasbourg. We thank IDRIS and CINES for generous allowance of computer time and the Région Alsace for their support of this work (Bourse de thèse régionale for F.S.). C.S. acknowledges the support by Marie Curie Research Training Grant under the Training and Mobility of Research (TMR) program of the European Commission. We thank Roland Stote, Vincent Zoete, Andrew Atkinson, and Jean-Marie Bernassau for helpful discussions, and Rebecca Pruss for her support of this work.

Supporting Information Available: Figure comparing minima ranking without and with solvation contribution (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA0265658