Determination of Protein–Ligand Binding Modes Using Complexation-Induced Changes in ^{1}H NMR Chemical Shift

Marina Cioffi,[†] Christopher A. Hunter,*[,†] Martin J. Packer,[‡] and Andrea Spitaleri[†]

Department of Chemistry, Centre for Chemical Biology, Krebs Institute for Biomolecular Science, University of Sheffield, Sheffield S3 7HF, U.K., and AstraZeneca, Alderley Park, Cheshire SK10 4TG, U.K.

Received September 24, 2007

A new method for determining three-dimensional solution structures of protein–ligand complexes using experimentally determined complexation-induced changes in ^{1}H NMR chemical shift (CIS) is introduced. The method has been validated using the complex formed between the protein antitumor antibiotic neocarzinostatin (NCS) and a synthetic chromophore analogue. The X-ray crystal structure of the unbound protein and the backbone amide proton CIS were the input data used in the determination of the three-dimensional structure of the complex. The experimental CIS values were used in a continuous direct structure refinement process based on genetic algorithms to sample conformational space. The calculated structure of the complex agrees well with the NMR solution structure, indicating the potential of this approach for structure determination.

Introduction

Determination of the three-dimensional (3D) structures of the complexes formed between proteins and small molecules (ligands) is important in drug discovery and in understanding protein function in general.[1,2] Many approaches have been developed to study interactions of ligands with biological receptors,[3–6] but obtaining structural information at an atomic level remains a challenge. X-ray crystallography is the most powerful tool but requires high-quality crystals and stable complexes.[7,8] The main advantage of NMR methods is that they can be used to study equilibria in solution, providing information on all species present.[9] NMR measurements, such as nuclear Overhauser effects (NOEs) and residual dipolar couplings (RDCs), provide structural restraints that can be used to calculate 3D structures of complexes.[10] Amide hydrogen–deuterium (H/D) exchange rates can give information on which parts of the protein are buried on complexation to define an interaction interface.[11] Another important parameter is the complexation-induced change in chemical shift (CIS[a]), which can be used to map the interfaces of complexes and locate binding pockets in proteins.[12]

CIS methods exploit the sensitivity of proton chemical shifts to their environment, and accurate data are relatively straightforward to acquire.[13] Experimental CIS data are used to determine a protein–protein or protein–ligand interaction site using a surface mapping approach. In CIS mapping, perturbations are mapped on the protein van der Waals surface, and if there are no significant ligand-induced conformational changes in the protein, the changes observed can be assigned to direct contacts with the ligand (e.g., hydrogen bonds) or to indirect effects such as ring current shifts produced by aromatic groups on the ligand. Residues on the protein that are perturbed are therefore assumed to be proximal to the ligand interaction site. Examination of CIS mapping forms the basis of the structure–activity relationship (SAR) by NMR approach that has become an important tool in structure-based drug discovery.[14] The method can be used to screen for low affinity ligands, providing information on both the location of the ligand binding site and the orientation of the ligand with respect to the binding site.[15] Comparison of CIS mapping data for closely related analogues can allow determination of the binding modes for a range of ligands.[16,17]

Exploitation of the experimental CIS data is usually applied qualitatively, but recently attempts have been made to use the data in a more quantitative manner to provide more detailed information on the structures of complexes.[18–20] The availability of NMR prediction tools to produce estimates of the CIS based on a chemical structure allows comparison of experimental data with expectations based on three-dimensional structural information. There are a number of methods available,[21,22] but the most successful methods for the prediction of CIS values for large molecules are based on semiempirical methodology.[13,23] These approaches have been used in conjunction with structure optimization algorithms to study protein–ligand docking. CIS data are usually introduced as part of a scoring function or as more qualitative ambiguous restraints to restrict the location of the ligand during sampling. Morelli et al.[24,25] incorporated a filter based on CIS data in the program BiGGER (biomolecular complex generation with global evaluation and ranking).[26] BiGGER produces a large set of binding configurations, which are filtered and scored on the basis of geometric surface complementarity, electrostatic potential, and solvation energy. These configurations are then further filtered and scored using CIS data, NOE information, or H/D exchange data to identify the best solutions. Other docking programs have implemented similar approaches, using a CIS scoring function to rank solutions obtained from a conventional conformational search.[27,28] In the program Jsurf, observed CIS values are assumed to be ring current effects, and this approach has been used to introduce structural restraints in combination with RDCs for solving the structures of protein–ligand and protein–protein complexes.[29,30] Bonvin et al.[31,32] developed the program HADDOCK (high ambiguity driven docking), which uses biochemical and bio-

* To whom correspondence should be addressed. Phone: +44 114 2229476. Fax: +44 114 2229346. E-mail: C.Hunter@shef.ac.uk.
[†] University of Sheffield.
[‡] AstraZeneca.

[a] Abbreviations: NCS, neocarzinostatin; CIS, complexation-induced changes in ^{1}H NMR chemical shift; *n*-rmsd$_{CIS}$, normalized root-mean-square difference between experimental and calculated CIS values; GA, genetic algorithm.

10.1021/jm701194r CCC: $40.75
Published on Web 03/27/2008

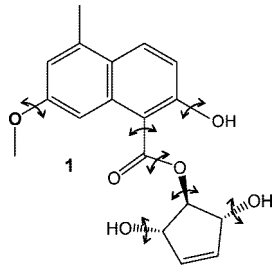

Figure 1. Structure of ligand **1**. Torsion angles that were allowed to vary during the structure determination process are indicated.

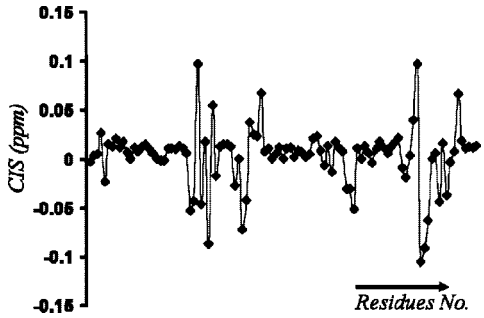

Figure 2. Experimental CIS values for the backbone amide protons of the protein NCS on binding ligand 1.[37]

physical data in order to drive the docking process, and CIS data can be introduced as ambiguous interaction restraints (AIRs). Similar approaches have been described by Clore and Schwieters[33] and by Schwalbe et al.[34]

We have previously described an approach that uses CIS to directly determine the 3D structures of synthetic supramolecular complexes.[35] The program generates a large number of conformations, calculates the CIS values for each, and then compares this information with experimentally determined values. The differences between the calculated and experimental CIS values are used by a genetic algorithm in a continuous refinement of the relative orientation, position, and conformation of the molecules. This method has been used successfully in the determination of 3D structures of intermolecular complexes in solution and the conformations of synthetic foldamers.[23,36] In this paper, we demonstrate the extension of the method to protein–ligand complexes.

Approach

CIS values are most strongly influenced by local functional group contacts and are therefore relatively short-range effects. This greatly simplifies the problem of the conformational search to dock the ligand and protein, since we only need to consider locations of the ligand that are in the binding site, and the location of the binding site is usually illuminated in a straightforward manner by CIS mapping. The method we have developed therefore has three key stages: (a) definition of the receptor binding site using the backbone amide CIS values; (b) generation of a set of ligand conformations and orientations for introduction into the receptor binding site (poses); (c) optimization of each pose based on comparison of the experimental and calculated CIS values for the amide backbone protons.

The system used to evaluate the potential of this approach is the complex formed by the chromoprotein antitumor antibiotic neocarzinostatin (NCS) and a synthetic chromophore (Figure 1), for which the three-dimensional structure of the complex in solution has been determined by conventional NMR methods using NOE restraints for structure refinement.[37] The structure of the complex reveals that the ligand is buried deep in the binding cleft of the protein, and the X-ray crystal structure of the unbound protein shows no large scale conformational changes on ligand binding. Analysis of the experimental CIS values shows that the largest perturbations occur for residues nearest the binding site (Figure 2). However, the observed CIS values are relatively small, with an average of 0.003 ppm and a largest value of 0.105 ppm, because the protein is ap-

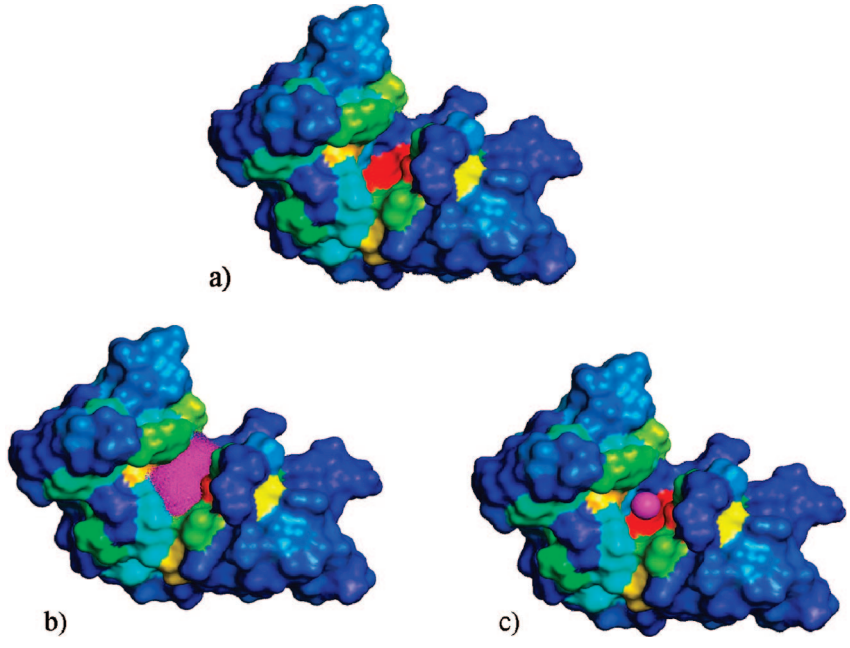

Figure 3. (a) CIS mapped onto the protein VDW surface. The surface is colored according to the magnitude of the CIS values. The largest absolute changes are represented in red, moderate changes in yellow, and green and smallest changes in blue. (b) *j*-surface (magenta) representing potential locations for the centre of the ligand. (c) Centre of the *j*-surface (magenta point).

proximately 40% bound under the experimental conditions used to acquire the spectra.

Results and Discussion

To ensure that our results were not biased by the details of the experimental solution structure of the protein–ligand complex, we used the unbound X-ray crystal structure of the protein (PDB code 1nco[38]), and the ligand was built using molecular mechanics and energy minimized.[39]

Stage 1: Location of the Binding Site. In the first step, the experimental CIS values were mapped onto the surface of the X-ray crystal structure of the protein using Jsurf.[29] When the protein surface was colored according to the magnitude of the CIS values (Figure 3a), the largest changes were clearly clustered around a cleft in the protein surface, and no significant perturbations were identified elsewhere. This demonstrates that the CIS are mainly due to the binding of the ligand in a single well-defined binding site, and there are minimal conformational changes in the rest of the protein. Using dot-density representations, spheres were constructed centered on each perturbed proton to create a *j*-surface representation of the binding site. Each dot represents a possible location for the center of the ligand. The Jsurf standard deviation parameter was manually varied from 0.1 to 4.0 until the *j*-surface was a small well-defined area in the middle of the binding site. A value of 2.5 gave a single *j*-surface that did not contact the protein backbone and was centered within the binding site cleft (Figure 3b). All points in the *j*-surface less than 2.5 Å from the protein backbone were removed, and the remaining points were averaged to give the coordinates of a single point identifying the center of the binding site (Figure 3c). This provides a good starting point for location of the center of the ligand, obviating the need for searching a large amount of redundant conformational space remote from the binding site.

Stage 2: Generation of a Diverse Ensemble of Sterically Accessible Poses. The location of the center of the binding site from Jsurf was used to generate a set of starting conformations of the ligand in the NCS binding site using GOLD.[40] This program uses genetic algorithms (GA) to generate and optimize structures of protein–ligand complexes according to a range of possible scoring functions. In our case, we are simply interested in generating a large number of possible structures that are sterically accessible and are not perturbed by any of the empirical scoring functions. The scoring function was therefore set to consider only van der Waals energies and intramolecular strain energy, and the molecular interaction terms were switched off to generate a set of 100–1000 poses based on shape complementarity only. Figure 4 shows an overlay of 20 poses generated by this method. The ligand is clearly located in the binding site, but the conformation and orientation of the ligand in each pose are very different showing that this method allows us to sample a wide range of different possible structures for the complex.

Stage 3: Structure Refinement Using CIS Values. The poses generated by GOLD were then optimized using the experimental CIS data in conjunction with the program Shifty.[35] This program calculates the CIS of each backbone amide using a semiempirical function (described in detail elsewhere[35]) and a GA to perform a conformational search in which the protein is considered as a rigid body and the ligand is allowed to move and change conformation (see Figure 1). In order to use a GA, it is necessary to define a fitness function which is that parameter that is maximized in the optimization process. The agreement between the calculated and experimental CIS values can be assessed using the normalized root-mean-square difference, and the fitness is therefore defined as the inverse of this parameter ($1/n\text{-rmsd}_{CIS} = \Delta\delta_{exp}/\Delta\delta_{cal}$).

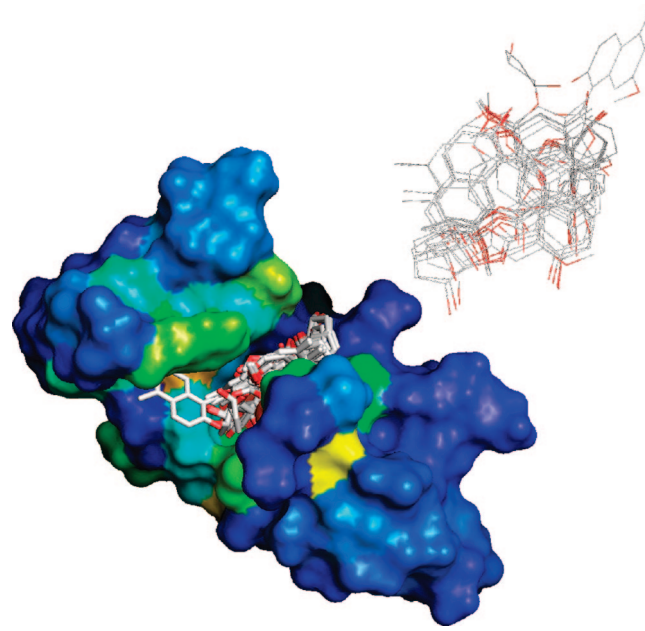

Figure 4. Overlay of 20 of the poses generated using GOLD.

$$\Delta\delta_{cal} = \frac{1}{N}\sqrt{\sum_{i=1}^{N}\frac{(CIS^{i}_{obs} - CIS^{i}_{calc})^2}{|CIS^{i}_{obs}|}} \quad (1)$$

$$\Delta\delta_{exp} = \frac{1}{N}\sqrt{\sum_{i=1}^{N}\left|CIS^{i}_{obs}\right|} \quad (2)$$

$\Delta\delta_{cal}$ is the root-mean-square difference between the calculated and experimental CIS values normalized by the absolute experimental CIS value. In cases where the absolute experimental CIS value was less than 0.1 ppm, this value was fixed at 0.1 ppm to reduce the significance of very small chemical shift changes that are subject to a large relative error. $\Delta\delta_{exp}$ is the root mean absolute experimental CIS value.

In addition, an association constant scaling factor was included as a variable to scale the experimental CIS values to allow for ambiguity in the extent to which the protein is bound in the NMR experiment. In principle, this correction could be made using an experimentally determined association constant, but the extra variable has little impact on the structure calculation. In addition, this strategy removes the need for experimental determination of the association constant and errors that could be introduced through inaccuracies in this value, which would affect every single CIS value used as input to the calculation. In effect, this approach optimizes the structure based on the relative changes in chemical shift rather than absolute values. For the complex described here, the association constant scaling factor for the optimized structures was around 0.3, which is consistent with the 40% bound protein determined experimentally.

Each of the poses generated by GOLD was used as an independent starting point for a structure optimization calculation. After optimization, the highest fitness obtained was 2.1, and most of the optimized structures gave an $n\text{-rmsd}_{CIS}$ of less than 0.5. Figure 5 illustrates the agreement obtained between the experimental and calculated backbone amide CIS values for the binding site residues of the top ranked pose. The quality of the structures of the complex obtained was evaluated by

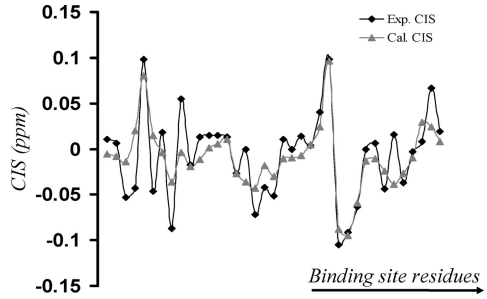

Figure 5. Comparison between calculated and experimental CIS values for the backbone amide protons in the binding site. The other backbone amide protons that are remote from the ligand have negligible CIS values, which are accurately reproduced by the calculation.

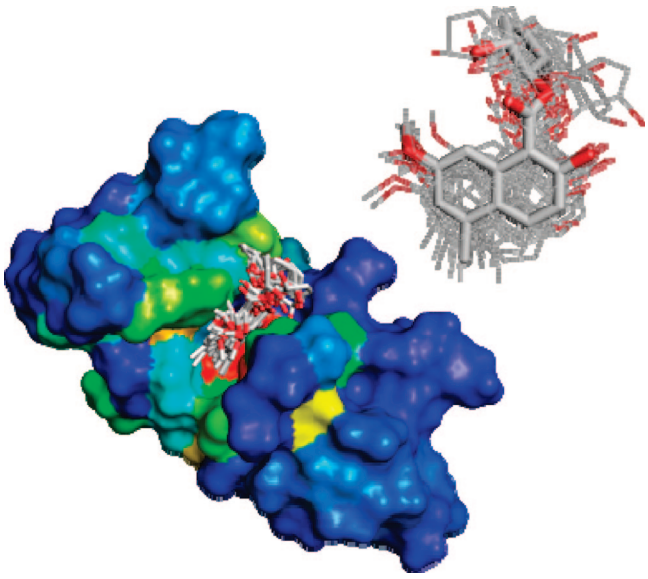

Figure 6. Overlay of the 20 highest fitness structures after optimization using the CIS values.

overlaying the backbone of the protein from the calculated structure with the backbone of the protein from the experimental solution NMR structure and determining the root-mean-square deviation (rmsd) of the ligand atoms in the two structures. The results in Figure 6 show that the optimization process converges to a clearly defined structure compared with the diverse set of starting structures in Figure 4. In the optimized structures, the naphthoate moiety of the ligand is located in a well-defined geometry which is the same in all of the structures, and the only variation is in the orientation of the ester side chain. This portion of the ligand has little magnetic anisotropy, and therefore, the geometry is less well-defined by the CIS values than the aromatic portion which is associated with large ring current effects. An overlay of the optimized structure with highest fitness and the NMR solution structure shows good agreement (Figure 7). The rmsd of the coordinates of the heavy atoms of the ligands in the two structures in Figure 7 is 2.9 Å. The difference is mainly due to the uncertainty in the position of the ester side chain of the ligand, and the rmsd of the naphthoate moiety alone is 1.0 Å.

Note that the protein is not allowed to move during the structure determination process, but the two protein frames in Figure 7 differ because one comes from the X-ray crystal structure of the unbound protein and the other comes from the solution NMR structure of the complex. Closer inspection of two protein frames reveals some movement of the loops flanking the binding site. In our calculation, the unbound protein structure

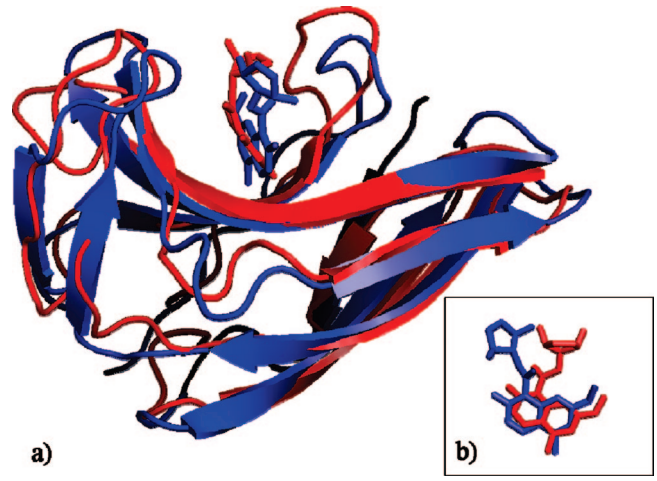

Figure 7. (a) Overlay of the CIS-optimized ligand orientation in the X-ray crystal structure of the unbound protein (blue) and the NMR solution structure of the complex (red). (b) Closer view of the orientation of the ligand in the two complexes.

was treated as a rigid body, and the small differences found in the orientation of the ligand are probably due to movement of the walls of the binding site on complexation. Thus, even though there are changes in the structure of the protein on binding, the method outlined above is sufficiently robust to accurately locate the ligand binding geometry because the CIS values are more sensitive to local intermolecular contacts that are relatively short-range in nature.

When structure optimizations were carried out using a set of 1000 starting poses, the structures of the top ranked poses were similar to those obtained using only 100 starting poses (Figure 8a). Similarly, increasing the number of generations and population size used in the optimization from 50 to 100 had little effect on the outcome (Figure 8b). Thus, relatively small fast searches can be used to effectively sample the conformational space. The results in Figure 8 show the distribution of NMR n-rmsd$_{CIS}$ versus structural rsmd values for each individual optimization. There is a clear correlation between the quality of the fit to the NMR data (measured by n-rmsd$_{CIS}$) and the quality of the structure (measured by rmsd). In other words, the NMR CIS data provide a useful tool for structure optimization. The n-rmsd$_{CIS}$ is a good measure of the quality of a structure of a complex so that optimizations based on this parameter converge toward the experimentally observed 3D structure.

Conclusions

A new protocol for determining the 3D structures of protein–ligand complexes in solution using NMR CIS data is described. The potential of this approach has been demonstrated using the complex formed by the chromoprotein antitumor antibiotic neocarzinostatin (NCS) and a synthetic ligand. The procedure requires only a 3D structure of the protein and the backbone amide CIS values and so could be relatively straightforward to implement in a high-throughput format for studying protein–ligand interactions. For a 105 residues−protein (11 kDa) with a small ligand (39 atoms), the chromophore of the ligand was located in the binding site with an accuracy of better than 1 Å compared with the experimental structure of the complex, and the calculated CIS values agree well with the experimental data. The main differences between the calculated and experimental structures are related to some conformational rearrangement of protein loops flanking the binding site and to the

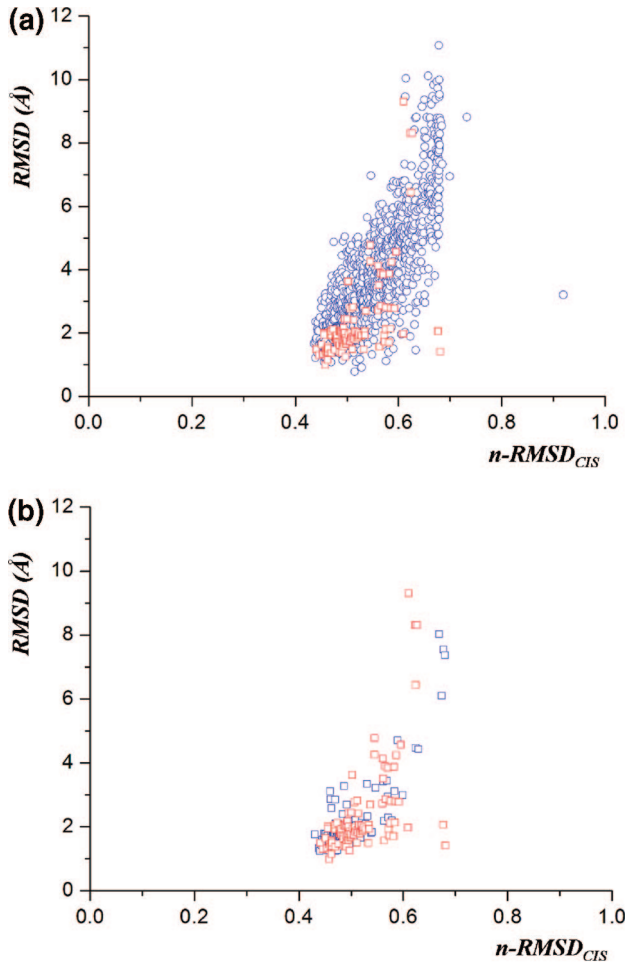

Figure 8. (a) Comparison between the ligand rmsd (considering only the coordinates of the heavy atoms of the rigid naphthoate moiety) and *n*-rmsd$_{CIS}$ for the set of 1000 poses (blue circles) and the set of 100 poses (red squares). (b) Comparison between the ligand rmsd and *n*-rmsd$_{CIS}$ for smaller (red squares) and larger (blue squares) conformational searches using increased generation and population sizes.

orientation of the flexible aliphatic side chain of the ligand. The aliphatic side chain carries little magnetic anisotropy and so is effectively invisible to the CIS method. Thus, the position and orientation of the ligand aromatic ring can be located with high accuracy in the binding pocket, but the ester side chain cannot. The method clearly shows some promise for determining the three-dimensional structures of protein–ligand complexes with relatively high precision, and we are currently investigating the general viability of CIS structure determination in different protein–ligand systems.

Experimental Procedure

The coordinates of the unbound protein were obtained from the protein data bank (PDB code 1nco[38]). The structure of the ligand was created with the XED 6.1.0 software using standard bond lengths and angles and energy-minimized. Our computational approach consists of a set of Perl scripts that implement the three main software packages used in the protocol and analyze the results.

Jsurf. In the first stage of the protocol, the location of the binding site is obtained using the program Jsurf. All points on the *j*-surface less than 2.5 Å from the protein backbone were removed, and the remaining points were averaged to give the coordinates of a single point: the center of the binding site.

GOLD. In the second stage, the generation of an ensemble of poses located in the binding site was carried out using the GOLD, version 2.2, software. The center of the ligand was located at the center of the binding site determined using Jsurf, giving a starting point for GOLD to generate a set of conformations of the ligand in the NCS binding site. The GOLD scoring function was modified by setting the contribution of the hydrogen bond energy term (H_BOND_WT) to 0.001 in gold.parms. This allows us to use this software as a rapid method to generate a set of poses based on shape complementary only. GOLD was used to generate structures using 10 runs with population sizes of 100–1000 for 10000–100000 generations. In this way, ensembles of 100 to 1000 poses could be obtained. A parameter used by GOLD in the default setting is "early termination". This option instructs the program to terminate runs as soon as a specified number of runs have given essentially the same answer, so this parameter was switched off in our calculations in order to ensure that GOLD generated a diverse sample of structures.

Shifty. In the final stage, structure determination was carried out using Shifty. The conformational search in Shifty was divided into two steps, each with population sizes of 50–100 run for 50–100 generations. The size of the search space was set to ensure that all possible complex conformations could be sampled. In the first step, the intermolecular distance limit was set to 1.5 Å and the range of allowed rotations of one molecule relative to the other was set to ±10°. Intramolecular torsions were allowed to change within the full range of ±180°. In the second step, these parameters were reduced to 1 Å, 5°, and 90°, respectively. To reduce the conformational space, a steric clash penalty was added for distances of less than 2 Å for intermolecular clashes and distances less than 1 Å for intramolecular clashes for non-hydrogen atoms.

Acknowledgment. We thank Dr. G. Moyna for making available the source code of the Jsurf program, Professor M. P. Williamson for some of the code used to develop Shifty, and Professor D. N. Woolfson for providing the experimental CIS data for the complex. We thank AstraZeneca for funding.

References

(1) Orengo, C. A.; Todd, A. E.; Thorton, J. M. From protein to structure to function. *Curr. Opin. Struct. Biol.* **1999**, *9*, 374–382.
(2) Amzel, L. M. Structure-based drug design. *Curr. Opin. Biotechnol.* **1998**, *9*, 366–369.
(3) Curto, E. V.; Moseley, H. N.; Krishna, N. R. CORCEMA evaluation of the potential role of intermolecular transferred NOESY in the characterization of ligand-receptor complexes. *J. Comput.-Aided Mol. Des.* **1996**, *10* (5), 361–371.
(4) Messina, P.; Prieto, G.; Dodero, V.; Ruso, J. M.; Schulz, P.; Sarmiento, F. Ultraviolet-circular dichroism spectroscopy and potentiometric study of the interaction between human serum albumin and sodium perfluorooctanoate. *Biopolymers* **2005**, *79* (6), 300–309.
(5) Rodger, A.; Marrington, R.; Roper, D.; Windsor, S. Circular dichroism spectroscopy for the study of protein−ligand interactions. *Methods Mol. Biol.* **2005**, *305* (Protein−Ligand Interactions), 343–363.
(6) Delfini, M.; Bianchetti, C.; Di Cocco, M. E.; Pescosolido, N.; Porcelli, F.; Rosa, R.; Rugo, G. An NMR spectroscopy study of bendaline−albumin interactions. *Bioorg. Chem.* **2003**, *31* (5), 378–388.
(7) Villar, H. O.; Yan, J.; Hansen, M. R. Using NMR for ligand discovery and optimization. *Curr. Opin. Chem. Biol.* **2004**, *8* (4), 387–391.
(8) Lepre, C. A.; Moore, J. M.; Peng, J. W. Theory and applications of NMR-based screening in pharmaceutical research. *Chem. Rev.* **2004**, *104* (8), 3641–3675.
(9) Peng, J. W.; Moore, J.; Abdul-Manan, N. NMR experiments for lead generation in drug discovery. *Prog. Nucl. Magn. Reson. Spectrosc.* **2004**, *44* (3–4), 225–256.
(10) Tolman, J. R.; Flanagan, J. M.; Kennedy, M. A.; Prestegard, J. H. Nuclear magnetic dipole interactions in field-oriented proteins: information for structure determination in solution. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 9279–9283.
(11) Takahashi, H.; Nakanishi, T.; Kami, K.; Arata, Y.; Shimada, I. A novel NMR method for determining the interfaces of large protein−protein complexes. *Nat. Struct. Biol.* **2000**, *7* (3), 220–223.
(12) Zuiderweg, E. R. P. Mapping protein−protein interactions in solution by NMR spectroscopy. *Biochemistry* **2002**, *41* (1), 1–7.
(13) Szilagyi, L. Chemical shifts in proteins come of age. *Prog. Nucl. Magn. Reson. Spectrosc.* **1995**, *27* (4), 325–443.
(14) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **1996**, *274* (5292), 1531–1534.

(15) McCoy, M. A.; Wyss, D. F. Alignment of weakly interacting molecules to protein surfaces using simulations of chemical shift perturbations. *J. Biomol. NMR* **2000**, *18* (3), 189–198.
(16) Wyss, D. F.; Arasappan, A.; Senior, M. M.; Wang, Y.-S.; Beyer, B. M.; Njoroge, F. G.; McCoy, M. A. Non-peptidic small-molecule inhibitors of the single-chain hepatitis C virus NS3 protease/NS4A cofactor complex discovered by structure-based NMR screening. *J. Med. Chem.* **2004**, *47* (10), 2486–2498.
(17) Medek, A.; Hajduk, P. J.; Mack, J.; Fesik, S. W. The use of differential chemical shifts for determining the binding site location and orientation of protein-bound ligands. *J. Am. Chem. Soc.* **2000**, *122* (6), 1241–1242.
(18) Helgaker, T.; Jaszunski, M.; Ruud, K. Ab initio methods for the calculation of NMR shielding and indirect spin−spin coupling constants. *Chem. Rev.* **1999**, *99* (1), 293–352.
(19) Vaara, J.; Jokisaari, J.; Wasylishen, R. E.; Bryce, D. L. Spin−spin coupling tensors as determined by experiment and computational chemistry. *Prog. Nucl. Magn. Reson. Spectrosc.* **2002**, *41* (3–4), 233–304.
(20) Contreras, R. H.; Barone, V.; Facelli, J. C.; Peralta, J. E. Advances in theoretical and physical aspects of spin−spin coupling constants. *Annu. Rep. NMR Spectrosc.* **2003**, *51*, 167–260.
(21) Xu, X.-P.; Case, D. A. Automated prediction of ^{15}N, ^{13}Ca, ^{13}Cb and ^{13}C′ chemical shifts in proteins using a density functional database. *J. Biomol. NMR* **2001**, *21* (4), 321–333.
(22) Neal, S.; Nip Alex, M.; Zhang, H.; Wishart David, S. Rapid and accurate calculation of protein ^{1}H, ^{13}C and ^{15}N chemical shifts. *J. Biomol. NMR* **2003**, *26* (3), 215–240.
(23) Hunter, C. A.; Packer, M. J.; Zonta, C. From structure to chemical shift and vice-versa. *Prog. Nucl. Magn. Reson. Spectrosc.* **2005**, *47* (1–2), 27–39.
(24) Morelli, X.; Dolla, A.; Czjzek, M.; Palma, P. N.; Blasco, F.; Krippahl, L.; Moura, J. J. G.; Guerlesquin, F. Heteronuclear NMR and soft docking: an experimental approach for a structural model of the cytochrome c553-ferredoxin complex. *Biochemistry* **2000**, *39* (10), 2530–2537.
(25) Morelli, X. J.; Palma, P. N.; Guerlesquin, F.; Rigby, A. C. A novel approach for assessing macromolecular complexes combining soft-docking calculations with NMR data. *Protein Sci.* **2001**, *10* (10), 2131–2137.
(26) Palma, P. N.; Krippahl, L.; Wampler, J. E.; Moura, J. J. G. BiGGER: a new (soft) docking algorithm for predicting protein interactions. *Proteins: Struct., Funct., Genet.* **2000**, *39* (4), 372–384.
(27) Kohlbacher, O.; Burchardt, A.; Moll, A.; Hildebrandt, A.; Bayer, P.; Lenhof, H.-P. Structure prediction of protein complexes by an NMR-based protein docking algorithm. *J. Biomol. NMR* **2001**, *20* (1), 15–21.
(28) Dobrodumov, A.; Gronenborn, A. M. Filtering and selection of structural models: combining docking and NMR. *Proteins: Struct., Funct., Genet.* **2003**, *53* (1), 18–32.
(29) McCoy, M. A.; Wyss, D. F. Spatial localization of ligand binding sites from electron current density surfaces calculated from NMR chemical shift perturbations. *J. Am. Chem. Soc.* **2002**, *124* (39), 11758–11763.
(30) McCoy, M. A.; Wyss, D. F. Structures of protein-protein complexes are docked using only NMR restraints from residual dipolar coupling and chemical shift perturbations. *J. Am. Chem. Soc.* **2002**, *124* (10), 2104–2105.
(31) Dominguez, C.; Boelens, R.; Bonvin, A. M. J. J. HADDOCK: a protein−protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **2003**, *125* (7), 1731–1737.
(32) van Dijk, A. D. J.; Boelens, R.; Bonvin, A. M. J. J. Data-driven docking for the study of biomolecular complexes. *FEBS J.* **2005**, *272* (2), 293–312.
(33) Schwieters, C. D.; Kuszewski, J. J.; Tjandra, N.; Marius Clore, G. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **2003**, *160* (1), 65–73.
(34) Schieborr, U.; Vogtherr, M.; Elshorst, B.; Betz, M.; Grimme, S.; Pescatore, B.; Langer, T.; Saxena, K.; Schwalbe, H. How much NMR data is required to determine a protein−ligand complex structure. *ChemBioChem* **2005**, *6* (10), 1891–1898.
(35) Hunter, C. A.; Packer, M. J. Complexation-induced changes in 1H NMR chemical shift for supramolecular structure determination. *Chem.−Eur. J.* **1999**, *5* (6), 1891–1897.
(36) Spitaleri, A.; Hunter, C. A.; McCabe, J. F.; Packer, M. J.; Cockroft, S. L. A ^{1}H NMR study of crystal nucleation in solution. *CrystEngComm* **2004**, *6*, 489–493.
(37) Urbaniak, M. D.; Muskett, F. W.; Finucane, M. D.; Caddick, S.; Woolfson, D. N. Solution structure of a novel chromoprotein derived from apo-neocarzinostatin and a synthetic chromophore. *Biochemistry* **2002**, *41* (39), 11731–11739.
(38) Kim, K. H.; Kwon, B. M.; Myers, A. G.; Rees, D. C. Crystal structure of neocarzinostatin, an antitumor protein−chromophore complex. *Science* **1993**, *262* (5136), 1042–1046.
(39) Vinter, J. G. Extended electron distributions applied to the molecular mechanics of some intermolecular interactions. II. Organic complexes. *J. Comput.-Aided Mol. Des.* **1996**, *10* (5), 417–426.
(40) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267* (3), 727–748.

JM701194R