

Experimental Determination of van der Waals Energies in a Biological System**

Martin A. Wear, Daphne Kan, Amir Rabu, and Malcolm D. Walkinshaw*

We herein present a crystallographic approach for determining and analyzing the free energy of ligand binding, which provides a detailed picture identifying, at atomic level, specific van der Waals interactions that contribute to the binding energy. The difference in binding energy ($\Delta\Delta G$) of two ligands differing only by a methyl group was measured in the crystal, giving a value of $-0.99 \text{ kcal mol}^{-1}$. The buried surface area of the CH_3 group (5.98 \AA^2) is equivalent to a contribution to binding (van der Waals) energy of $-0.17 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$. These crystallographically determined binding energies provide a means to experimentally determine van der Waals interactions in a biological system. They agree well with binding constants measured from enzyme-inhibition experiments, which suggests that these group contributions towards binding energy should be transferable between a wide range of biologically important intermolecular interactions.

The experimentally determined equilibrium dissociation constant (K_d) for bimolecular protein–ligand complex formation provides a measure of the overall interaction energy (ΔG), which is equal to $-RT\ln(1/K_d)$, where R is the gas constant and T is the absolute temperature. Such solution studies, however, tell us little about the three-dimensional specifics of the individual functional groups involved in the interaction. Attempts to predict and assign proportional energetic contributions to individual functional groups for protein–ligand dissociation constants have been made by correlating binding measurements in solution with chemical features.^[1–6] A systematic analysis of the binding affinity of a number of protein–drug complexes correlated to the type and number of functional groups^[1] gave an average contribution to the binding affinity for a methyl group of $-0.8 \text{ kcal mol}^{-1}$. A maximum contribution to affinity of $-1.5 \text{ kcal mol}^{-1}$ per non-hydrogen atom has been estimated by using a similar correlation of over 60 protein–ligand complexes.^[3,4] Converting these values to binding energy per unit of surface area

gives a range of between -0.08 to $-0.2 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ as the average ligand contribution to protein–ligand complex stability (for non-hydrogen atoms). Other experimental approximations obtained by the deletion of methyl groups in binding sites, both within proteins or drug–receptor complexes, give a range of values from -0.17 to $-0.26 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$.^[6,7] An estimated van der Waals contribution to the binding energy of $-0.12 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ for nonpolar interactions between ligands and proteins has also been determined.^[8] However, such estimated functional-group contributions are composite values and it is impossible to identify how changes in conformation and/or changes in solvent structure contribute to the overall binding energy. The crystallographic method presented herein provides both quantitative equilibrium binding data and a snapshot of the structural changes and interactions that take place during ligand binding and also allows the separation of mainly entropic effects caused by changes in solvent structure upon binding.

Crystals of *C. elegans* cyclophilin 3 (Cyp3) were soaked in various concentrations of Xaa-Pro dipeptide inhibitors (Gly-Pro, Ala-Pro, and Ser-Pro). Crystallographic values of “ligand occupancy” were used to calculate a value for the ligand-dissociation constant, which gives a value for ligand-binding energy^[9] (see the Supporting Information). The $(2F_o - F_c)$ electron-density maps (where F_o and F_c are the observed and calculated structure factors, respectively) of Cyp3 soaked in increasing concentrations of dipeptide, highlighting the ligand and the position of Arg62 and four water molecules (W_a – W_d) in the active site (Figure 1 A and B), illustrate the conversion from the native to the bound state. The guanidino group of Arg62 adopts a different orientation to that in the native structure and displaces water molecule W_d , which is present in the native structure. A further three water molecules (W_a , W_b , and W_c) that are present in the native structure are also displaced by the binding of dipeptide ligands. All other residues around the binding site, with the exception of Arg62, adopt an identical conformation to that found in the native structure.^[9–11] Refinement of the ligand occupancies of the native structure as well as intermediate dipeptide ligand concentrations allow for the calculation of a crystal equilibrium dissociation constant (K_{dc} ; Figure 1 C). K_{dc} values of $(50 \pm 11) \text{ mM}$, $(9.1 \pm 0.8) \text{ mM}$ (from reference [9]), and $(6.3 \pm 0.7) \text{ mM}$ were determined for Gly-*cis*-Pro, Ala-*cis*-Pro, and Ser-*cis*-Pro, respectively (Table 1).

We also assessed the binding and inhibition of Cyp3 by Gly-Pro, Ala-Pro, and Ser-Pro in solution by measuring their ability to inhibit Cyp3s peptidyl prolyl isomerase (PPIase) activity (for experimental details see the Supporting Information). The prolyl imide bond in most peptides adopts *cis* and *trans* conformations. The cyclophilin PPIase activity

[*] Dr. M. A. Wear,^[†] Dr. D. Kan,^[†] Dr. A. Rabu,
Prof. Dr. M. D. Walkinshaw
The Centre for Translational and Chemical Biology
The University of Edinburgh
Michael Swann Building
King's Buildings
Edinburgh EH9 3JR (UK)
Fax: (+44) 131-650-7055
E-mail: m.walkinshaw@ed.ac.uk

[†] These authors contributed equally to this work.

[**] This work was supported by the MRC, the BBSRC, and the Wellcome Trust.

Supporting information for this article is available on the WWW under <http://www.angewandte.org> or from the author.

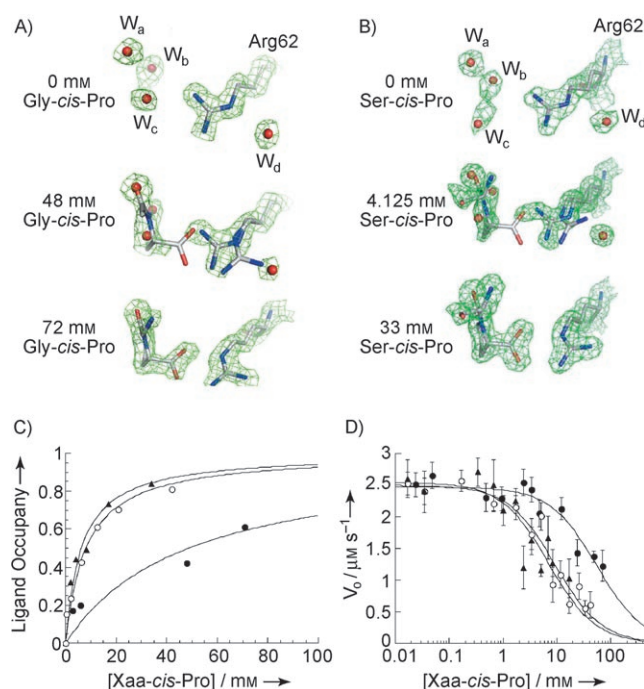


Figure 1. Determination of van der Waals energies in a biological system. Sections of the $(2F_o - F_c)$ difference electron-density maps of native crystals of Cyp3 soaked with Gly-Pro (A) or Ser-Pro (B) at the indicated concentration, highlighting the ligand, the position of Arg62, and the four water molecules (W_a – W_d) in the active site (●), and contoured at 1.5σ by using Pymol. The water molecules in the active site of free Cyp3 have been displaced by the Xaa-*cis*-Pro dipeptide ligand, and Arg62 adopts a different conformation compared with the native structure. C) Plot of refined fractional ligand occupancy versus the Gly-*cis*-Pro (●), Ala-*cis*-Pro (○), or Ser-*cis*-Pro (▲) concentration (mM).^[9] The solid line is a least-squares fit to Equation (1), which is shown in the Supporting Information, giving the values for the crystal equilibrium dissociation constants (K_{dc}) shown in Table 1. Ligand occupancy data for Ala-Pro was taken from reference [9] D) Plot of the initial background-corrected PPIase rate, V_0 ($\mu\text{M}^{-1} \text{s}^{-1}$), versus the Gly-*cis*-Pro (●), Ala-*cis*-Pro (○), or Ser-*cis*-Pro (▲) concentration (mM). The solid line is a least squares fit to Equation (2), which is shown in the Supporting Information, giving the mean solution equilibrium dissociation constants (K_{ds}) shown in Table 1.

catalyzes rotation about this imide bond, speeding up the attainment of the *cis*–*trans* equilibrium.^[12–14] However, the *cis* conformer of Xaa-Pro dipeptides acts as a weak cyclophilin inhibitor.^[9,11,15] All three peptides inhibit Cyp3s PPIase activity in the mM range. The mean solution equilibrium dissociation constants (K_{ds}) for Gly-*cis*-Pro, Ala-*cis*-Pro, and Ser-*cis*-Pro binding to Cyp3 were (49 ± 12) mM, (7.8 ± 1.5) mM, and (5.9 ± 1.7) mM, respectively (Figure 1D, Table 1). The K_{ds} value of (7.8 ± 1.5) mM for Ala-Pro determined in this study agrees well with a value of 8.2 mM, which was determined previously for Ala-Pro,^[9] when corrected for the amount of the *cis* conformer present in aqueous solution (35%,^[16]). Remarkably, the K_{dc} values, along with that reported for Ala-Pro binding to Cyp3,^[9] are, within experimental error, the same as those K_{ds} values determined in solution with the PPIase assay (Table 1). We thus have determined a series of binding affinities, which are in agreement with each other, from two separate experimental

methodologies. The data from the structural studies provides us with a unique set of modularized protein–peptide ligand interactions that allows for the dissection of the individual energetic components of the binding interactions.

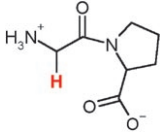
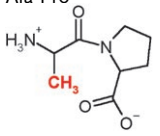
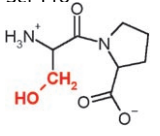
The X-ray crystal structures show Gly-*cis*-Pro, Ala-*cis*-Pro, and Ser-*cis*-Pro complexes with the ligands adopting identical poses (Figure 2). The proline residue and the peptidyl–prolyl bond in all three Xaa-*cis*-Pro structures are superimposable with a root mean square deviation (RMSD) fit for all comparable atoms of less than 0.2 Å. The RMSD values for Ala-*cis*-Pro, Ser-*cis*-Pro (in orientation), and Ser-*cis*-Pro (out orientation; see below), compared with Gly-*cis*-Pro, are 0.19 Å², 0.18 Å², and 0.16 Å², respectively. The RMSD values for Ser-*cis*-Pro (in orientation) and Ser-*cis*-Pro (out orientation), compared with Ala-*cis*-Pro, are 0.18 Å² and 0.13 Å², respectively (Figure 2). The binding of proline displaces three waters, W_a – W_c , and the reorientation of Arg62 displaces W_d (Figure 1A, B). There are two direct hydrogen bonds between the proline residue and Arg62/NH2 and Arg62/NE that are retained in all three structures as is the interaction between proline and Gln70/NE2. The hydration pattern, including the positions of the four water molecules (W_c , W_b , W_g , W_h) that mediate hydrogen bond interactions between Xaa-Pro and the protein, are also conserved.

As the ligand poses and the water positions in the three Xaa-*cis*-Pro structures are identical, the difference in the binding affinity between these ligands is only due to the side chain of the Xaa amino acid. We can therefore use the crystallographically determined interactions of the Gly, Ala, and Ser side chains to apportion individual components of the binding interaction in terms of the extra contacts, hydrogen bonds, and buried surface area. The difference in binding energy between Gly-*cis*-Pro and Ala-*cis*-Pro ($\Delta\Delta G_c^{\text{Gly}}$), which differ only by a methyl group, was calculated in the crystal to be $-0.99 \text{ kcal mol}^{-1}$ (Table 1). A very similar difference in binding energy was calculated in solution ($\Delta\Delta G_s^{\text{Gly}} = -1.02 \text{ kcal mol}^{-1} \text{ Å}^{-2}$, Table 1). The protein backbone and side chains as well as the “microscopic” water structure around the binding site in the Ala-*cis*-Pro structure is, within experimental error, the same as the Gly-*cis*-Pro structure (Figure 2). Thus, essentially only the van der Waals contact—the methyl group in the Ala-*cis*-Pro complex—has increased. The calculated buried surface area of the CH₃ group is 5.98 Å², which, when using the value determined above of $-0.99 \text{ kcal mol}^{-1}$, equates to a van der Waals energy contribution of $-0.17 \text{ kcal mol}^{-1} \text{ Å}^{-2}$.

The addition of an OH group in the Ser-*cis*-Pro complex only contributes a further $-0.27 \text{ kcal mol}^{-1}$ to the binding energy (Table 1). There are two orientations for serine oxygen—“in” (30% occupancy) and “out” (70% occupancy). As seen in Figure 2, there is a different hydrogen-bonding pattern in each orientation. The rather small contribution (maximally $-0.1 \text{ kcal mol}^{-1}$ per hydrogen bond) to the ligand-binding energy is presumably because, in both conformations, the serine hydroxy group points into the solvent and mainly forms hydrogen bonds with water rather than forming specific hydrogen bonds with the protein.

The individual ΔG values for Xaa-*cis*-Pro dipeptides binding to Cyp3, determined from solution assays or from

Table 1: Equilibrium dissociation constants (in [nM]) and free energies (in [kcal mol⁻¹]) for the Xaa-*cis*-Pro dipeptides binding to Cyp3.^[a]

| Dipeptide/Structure | K_{dc} | Crystal | | | K_{ds} | Solution | | |
|--|-----------|--------------|--------------------------|--------------------------|-----------|--------------|--------------------------|--------------------------|
| | | ΔG_c | $\Delta\Delta G_{c,Gly}$ | $\Delta\Delta G_{c,Ala}$ | | ΔG_s | $\Delta\Delta G_{s,Gly}$ | $\Delta\Delta G_{s,Ala}$ |
| Gly-Pro  | 50 ± 11 | -1.73 | N/A | N/A | 49 ± 12 | -1.67 | N/A | N/A |
| Ala-Pro  | 9.1 ± 0.8 | -2.72 | -0.99 | N/A | 7.8 ± 1.5 | -2.69 | -1.02 | N/A |
| Ser-Pro  | 6.3 ± 0.7 | -2.99 | -1.26 | -0.27 | 5.9 ± 1.7 | -2.85 | -1.18 | -0.16 |

[a] The values for K_{dc} , ΔG_c , $\Delta\Delta G_{c,Gly}$, and $\Delta\Delta G_{c,Ala}$ and K_{ds} , ΔG_s , $\Delta\Delta G_{s,Gly}$ and $\Delta\Delta G_{s,Ala}$ are shown for three Xaa-Pro dipeptides and were determined from crystal-soaking experiments (at 18 °C^[9]) and solution assays (at 6 °C^[12]), respectively. Dissociation binding constants for Ala-Pro was taken from^[9] and corrected for the concentration of the *cis* conformer in aqueous solution (35 % for Ala-Pro^[16]). The side chain of the Xaa amino acid is highlighted in red in the structural diagram of the compound. ΔG_x values were calculated by using the equation $\Delta G_x = -RT \ln(1/K_{dx})$, where R is the universal gas constant, T is the absolute temperature, and K_{dx} is the corresponding equilibrium dissociation constant, which was determined from the respective solution or crystal experiments. $\Delta\Delta G_{x,Gly}$ values are the difference between the ΔG_x values of Ala-*cis*-Pro or Ser-*cis*-Pro compared with that of Gly-*cis*-Pro and $\Delta\Delta G_{x,Ala}$ values are the difference between the ΔG_x values of Ser-*cis*-Pro compared with that of Ala-*cis*-Pro and were all determined from the respective solution or crystal experiments. Errors for K_{dc} are from the fitting program (Kaleidagraph, v4.0). K_{ds} values are the mean value ± standard error, with $n = 9$.

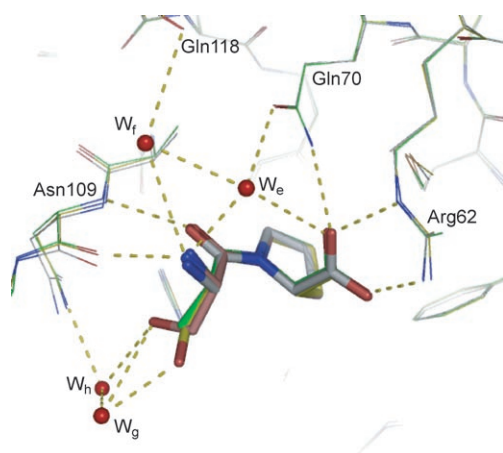


Figure 2. Superposition of the Xaa-*cis*-Pro-dipeptide-Cyp3 structures. The conserved water molecules (W_e – W_h) that are considered part of the composite protein surface in Cyp3–Xaa-*cis*-Pro interactions are colored as red spheres. All hydrogen bonds formed in the complexes are indicated as yellow dashed lines. Carbon atoms of the ligands Gly-*cis*-Pro, Ala-*cis*-Pro, Ser-*cis*-Pro (“in” form), and Ser-*cis*-Pro (“out” form) are shown in white, green, pink, and yellow, respectively.

crystal soaking data, are essentially the same (Table 1). This is a remarkable result considering the very different methods used to measure the affinity of the protein–ligand binding. The implication is that the interactions that can be accurately determined in the essentially static crystal structure are the same as those that provide the major contributions to the

binding in solution, which must also include the water molecules in contact with the ligand.

The difference of one methyl group between the Gly-Pro and Ala-Pro ligands results in a “microscopic change” between the crystal structures in which water positions, molecular conformations, and molecular poses are identical; only the van der Waals contact has increased. This observation allows us to apportion the $-0.17 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ binding energy to the difference in van der Waals interaction of the ligands in an aqueous environment compared with that of the protein-bound state. It is interesting that this value lies firmly in the range of previous estimates in which different binding poses, hydration patterns, protein conformations, and entropic effects must also contribute in varying and unaccountable ways to the binding energy. The method for determining and analyzing free energy presented herein not only provides a novel way of measuring binding energy, it also provides a detailed picture identifying, at the atomic level, which interactions are responsible for the changes in binding affinity.

Received: May 11, 2007

Published online: July 25, 2007

Keywords: cyclophilin · inhibitors · noncovalent interactions · van der Waals energies · X-ray diffraction

-
- [1] P. R. Andrews, D. J. Craik, J. L. Martin, *J. Med. Chem.* **1984**, 27, 1648.
 - [2] H. J. Bohm, *J. Comput.-Aided. Mater. Des.* **1998**, 12, 309.
 - [3] I. D. Kuntz, K. Chen, K. A. Sharp, P. A. Kollman, *Proc. Natl. Acad. Sci. USA* **1999**, 96, 9997.
 - [4] N. Brooijmans, K. A. Sharp, I. D. Kuntz, *Proteins* **2002**, 48, 645.
 - [5] S. Huo, J. Wang, P. Cieplak, P. A. Kollman, I. D. Kuntz, *J. Med. Chem.* **2002**, 45, 1412.
 - [6] D. H. Williams, E. Stephens, D. P. O'Brien, M. Zhou, *Angew. Chem.* **2004**, 116, 6760; *Angew. Chem. Int. Ed.* **2004**, 43, 6596.
 - [7] L. Serrano, J. L. Neira, J. Sancho, A. R. Fersht, *Nature* **1992**, 356, 453.
 - [8] P. Burkhard, P. Taylor, M. D. Walkinshaw, *J. Mol. Biol.* **2000**, 295, 953.
 - [9] S. y. Wu, J. Dornan, G. Kontopidis, P. Taylor, M. D. Walkinshaw, *Angew. Chem.* **2001**, 113, 602; *Angew. Chem. Int. Ed.* **2001**, 40, 582.
 - [10] J. Dornan, A. P. Page, P. Taylor, S. Wu, A. D. Winter, H. Husi, M. D. Walkinshaw, *J. Biol. Chem.* **1999**, 274, 34877.
 - [11] J. Dornan, P. Taylor, M. D. Walkinshaw, *Curr. Top. Med. Chem.* **2003**, 3, 1392.
 - [12] J. L. Kofron, P. Kuzmic, V. Kishore, E. Colon-Bonilla, D. H. Rich, *Biochemistry* **1991**, 30, 6127.
 - [13] D. Kern, G. Kern, G. Scherer, G. Fischer, T. Drakenberg, *Biochemistry* **1995**, 34, 13594.
 - [14] A. Galat, *Curr. Top. Med. Chem.* **2003**, 3, 1315.
 - [15] Y. Zhao, H. Ke, *Biochemistry* **1996**, 35, 7362.
 - [16] M. Brandsch, I. Knutter, F. Thunecke, B. Hartrodt, I. Born, V. Borner, F. Hirche, G. Fischer, K. Neubert, *Eur. J. Biochem.* **1999**, 266, 502.
-