Pergamon

Bioorganic & Medicinal Chemistry Letters 10 (2000) 2129–2132

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

HINT Predictive Analysis of Binding Between Retinol Binding Protein and Hydrophobic Ligands

Anna Marabotti,[a] Lara Balestreri,[a] Pietro Cozzini,[b] Andrea Mozzarelli,[a,c,*] Glen E. Kellogg[d] and Donald J. Abraham[d]

[a]*Institute of Biochemical Sciences, University of Parma, 43100 Parma, Italy*
[b]*Department of General and Inorganic Chemistry, University of Parma, 43100 Parma, Italy*
[c]*National Institute for the Physics of Matter, University of Parma, 43100 Parma, Italy*
[d]*Department of Medicinal Chemistry & Institute for Structural Biology and Drug Discovery, School of Pharmacy, Virginia Commonwealth University, Richmond, Virginia 23298-0133, USA*

Received 14 March 2000; revised 3 July 2000; accepted 17 July 2000

Abstract—The interaction between the retinol binding protein and four ligands was evaluated using HINT, a software based on experimental Log*P* values of individual atoms. A satisfactory correlation was found between the HINT scores and the experimental dissociation constants of three of the ligands, fenretinide, *N*-ethylretinamide and all-*trans* retinol, despite their hydrophobic nature. A prediction is made for the binding affinity of the fourth ligand, axerophtene, not yet determined in solution.

The exponential increase of available three-dimensional structures of proteins determined by X-ray crystallography has opened the way to the search of new drugs for pharmaceutically relevant protein targets. This process requires both the identification of lead compounds and the estimation of their binding affinity. The first step is achieved by the use of several docking software programs based on a variety of approaches.[1,2] The evaluation of the binding free energy is a difficult problem due to water molecules bound both to the isolated ligand and the protein binding site and released upon the formation of the complex. Several strategies have been proposed, and some of them have been successful in describing the formation of complexes between a protein and polar ligands.[3–7] In the present study, we evaluate whether the program HINT[8] (Hydropathic INTeractions) is able to predict the binding free energy between the retinol binding protein and its hydrophobic ligands. Most other programs are based on Newtonian physics and parameterization to calculate atom–atom interaction forces, whereas HINT is based on hydropathic values derived for every atom from experimentally determined solvent partitioning data, Log*P*, between water and octanol. The program has correctly predicted the free energy changes associated with native and mutant hemoglobin dimer–tetramer assembly.[9,10]

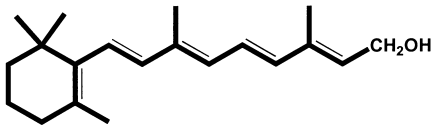
The retinol binding protein (RBP) is a 21 kDa protein acting as a carrier of vitamin A in the plasma. Vitamin A is delivered to the cell where it is bound by a cellular RBP, and transferred to an enzymatic system that oxidizes retinol to all-*trans* retinoic acid.[11] RBP belongs to the lipocaline family and exhibits a typical antiparallel β-barrel fold. The structures of bovine apo- and holo-RBP complexed with all-*trans* retinol, fenretinide, *N*-ethylretinamide and axerophtene (Table 1) have been determined at 1.8–1.9 Å resolution.[12–14] The ligands bind with the polar moiety pointing towards the exterior of the protein (Fig. 1), whereas in other members of the lipocaline family, as the cellular retinol binding proteins, the hydroxyl group is deeply buried in the protein central cavity. The dissociation constant of RBP ligands, measured using fluorimetric methods, ranges from about 70 to 170 nM (Table 2).[13,15,16] Some uncertainty on these values arises due to the reduced water solubility and photodamage. Retinol analogues are presently used in therapy for cancer and psoriasis.[17] Therefore, the evaluation of the strength of binding of RBP ligands is of interest for the design of new and potent drugs.

The first step of the procedure was the correction of atom type for the ligand and protein atoms to standard

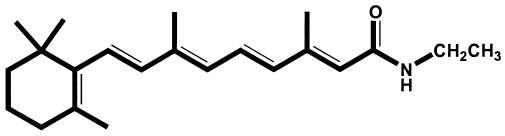
*Corresponding author. Tel.: +39-0521-905138; fax: +39-0521-905151; e-mail: biochim@unipr.it

0960-894X/00/$ - see front matter
PII: S0960-894X(00)00414-5

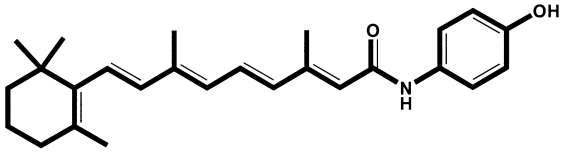
Table 1. Structures of the ligands of RBP



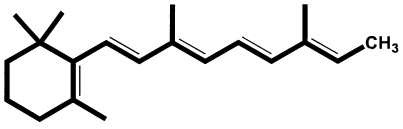
All-*trans* retinol



N-ethylretinamide



Fenretinide



Axerophtene

Sybyl[18] format from PDB. Hydrogens were then added to both the ligand and the protein. To reduce steric clashes caused by the added hydrogens, the protein structure with water molecules were energy minimized while keeping the coordinates of all non-hydrogens atoms fixed. No minimization was carried out for the hydrogens of the ligand. The calculated HINT scores of the complexes are reported in Table 2 (1), either counting all hydrogens of the ligand or only those bound to polar groups, called 'essentials'. In both cases, only the essential hydrogens of the protein were considered. The

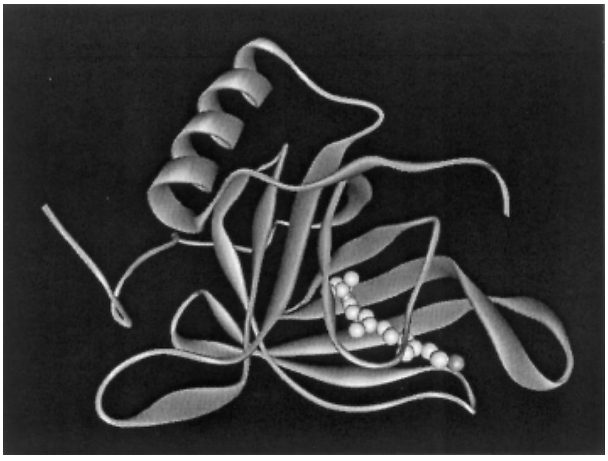


Figure 1. Three-dimensional structure of the RBP complexed with all-*trans* retinol.[12]

resulting HINT scores for all-*trans* retinol does not correlate with the experimental binding affinity. In order to understand this discrepancy, the four complexes were overlapped (Fig. 2). The terpene chain of retinol is not planar, as it should be, showing a clear distortion from a canonical 120° angle for the C20 methyl group. Therefore, the geometry of this ligand was optimized before hydrogen minimization. This step removed the distortion in the retinol conformation while no other parts of the molecule experienced structural changes. We considered this new structure as the correct one. For consistency, the same procedure was applied to the other ligands. The new calculated HINT scores are shown in Table 2 (2). The largest deviation of HINT scores was now observed for fenretinide. The crystallographic studies show that this ligand possesses three planes: the first contains C7–C14 and the methyl groups C19 and C20, the second is formed by the amidic group making an angle of 85° with the first plane, and the third is formed by the phenol ring making an angle of −45° with the terpenic chain.[14] This geometry is completely lost upon minimization and, therefore, we have to assume that the corresponding HINT score is not valid. Both the *N*-ethylretinamide and axerophtene undergo only small structural changes upon minimization indicating that their structures were already correct even before the optimization process. Therefore, in these cases, the HINT scores can be calculated directly from the original coordinates. Finally, due to the relevance of hydrophobic forces in the ligand–RBP interaction, the last step was to carry out the calculation of HINT scores considering all hydrogens for both protein

Table 2. HINT scores for the various RBP–retinoid complexes

Ligand	PDB code	K_d (nM)	Log*P* ligand	Hint score (1)		Hint score (2)		Hint score (3)
				H ligand		H ligand		H ligand
				All	Essential	All	Essential	All
All-*trans* retinol	1 hpb	70	6.16	639	168	811	287	933
Axerophtene	1 fen	—	8.15	2032	1400	2080	1428	921
N-Ethylretinamide	1 erb	90	6.31	981	700	1437	792	817
Fenretinide	1 fel	170	12.16	657	311	2466	1927	489

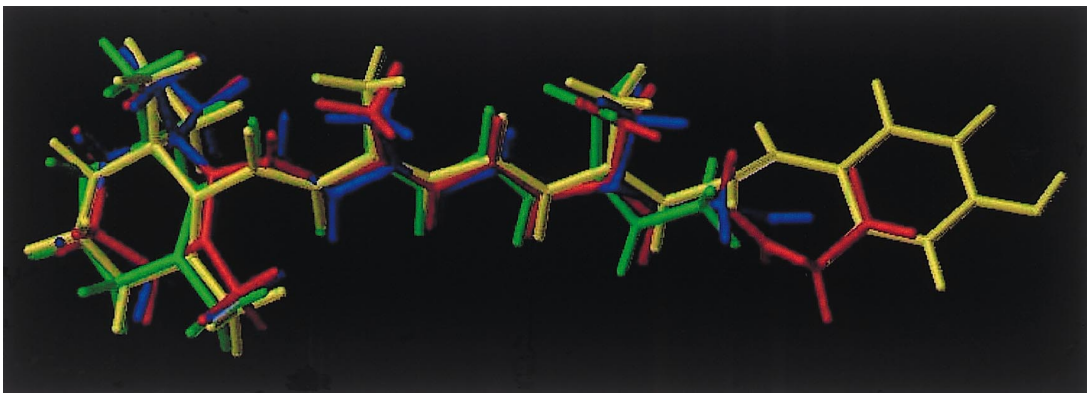


Figure 2. Superposition of the four ligands of RBP. Red: *N*-ethylretinamide; blue: retinol; yellow: fenretinide; green: axerophtene.

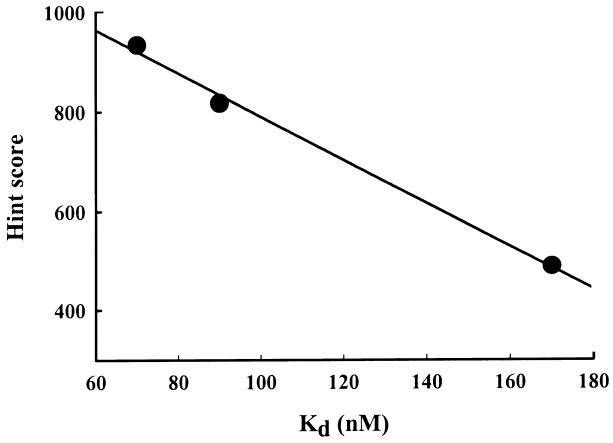


Figure 3. HINT scores (Table 2 (3)) versus dissociation constants. The straight line through data points exhibits a correlation coefficient r^2 of 0.99.

and ligands, after minimization of all hydrogens except those of the ligands, and geometry optimization of only the retinol structure. These values are reported in Table 2 (3). Their correlation with the dissociation constants is remarkably good (Fig. 3). On the basis of the regression analysis, the predicted dissociation constant of axerophtene, a compound for which the experimental binding affinity is not yet available, is close to 70 nm.

These results indicate that HINT correctly evaluates the predominant contribution to the binding free energy of the entropic term associated to water desolvation from the protein and from the ligand. An analysis using the scores obtained by LUDI and attributable to lipophilic contact area did not lead to any correlation. In spite of the good correlation found for RBP–ligand complexes, we are aware of the possibility that this success is due to the particular nature of the interaction in this system. In particular, the following two considerations are important in evaluating the correlation to experimental binding constants: (i) the HINT calculations are based upon crystallographically derived atomic heavy atom positions. In this regard, the X-ray structure of the complex with fenretinide has a closer than van der Waals contact for one pair of interacting heavy atoms (Val 61 CG1 to amide N of fenretinide) as well as a too close contact for a pair of hydrogens (amide NH of Leu 64 with a phenyl ring hydrogen *ortho* to the phenolic oxygen) that are in restricted positions (the corresponding heavy atom positions are not within van der Waals contact); (ii) the HINT calculations are also based upon the positions of hydrogens that are not determined or constrained crystallographically. Therefore, the results may be biased toward the final position of the hydrogen atoms that may not be optimal for hydrogen bonding. This methodology is being tested further with several other series of polar and apolar protein–ligand complexes to better evaluate HINT's capability to correlate calculated binding scores with experimental results. These studies are in preparation for publication.

Acknowledgements

The work was supported by the U.S. National Institutes of Health grant 5R01HL32793-15 to D.J.A. and G.E.K., and the Italy National Research Council grant 98.01117.CT14, the Italy Ministry of University and Scientific and Technological Research Programs of Research of National Interest 1999 grant and International University Cooperation 1998 grant to A.M.

References and Notes

1. Ajay; Murcko, M. A. *J. Med. Chem.* **1995**, *38*, 4953.
2. Böhm, H.-J.; Klebe, G. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 2588.
3. Åqvist, J. *J. Comp. Chem.* **1996**, *17*, 1587.
4. Böhm, H.-J. *J. CAMD* **1994**, *8*, 243–256.
5. Wang, J.; Dixon, R.; Kollman, P. A. *Proteins* **1999**, *34*, 69.
6. Wang, W.; Wang, J.; Kollman, P. A. *Proteins* **1999**, *34*, 395.
7. Shoichet, B. K.; Leach, A.; Kuntz, I. *Proteins* **1999**, *34*, 4.
8. The program HINT (eduSoft, LC, Ashland, VA, USA) was used as an add-on module within Sybyl. The options for the HINT score calculations are:

- *Partition method* (Log*P* calculation). The Calculate method was used for the ligand and the Dictionary method for the protein with the neutral condition for the solvent.
- *Interactions*. All the interactions were evaluated between the protein and the ligand.
- Other parameters were set as default.

9. Abraham, D. J.; Kellogg, G. E.; Holt, J. M.; Ackers, G. K. *J. Mol. Biol.* **1997**, *272*, 613.
10. Burnett, J. C.; Kellogg, G. E.; Abraham, D. J. *Biochemistry* **2000**, *39*, 1622.
11. Boyd, A. S. *Am. J. Med.* **1989**, *86*, 568.
12. Zanotti, G.; Berni, R.; Monaco, H. *J. Biol. Chem.* **1993**, *268*, 10728.
13. Zanotti, G.; Malpeli, G.; Berni, R. *J. Biol. Chem.* **1993**, *268*, 24873.
14. Zanotti, G.; Marcello, M.; Malpeli, G.; Folli, C.; Sartori, G.; Berni, R. *J. Biol. Chem.* **1994**, *269*, 29613.
15. Noy, N.; Xu, Z.-J. *Biochemistry* **1990**, *29*, 3878.
16. Berni, R.; Clerici, M.; Malpeli, G.; Cleris, L.; Formelli, F. *FASEB J.* **1993**, *7*, 1179.
17. Smith, M. A.; Parkinson, D. R.; Cheson, B. D.; Friedman, M. A. *J. Clin. Oncol.* **1992**, *10*, 839.
18. The program Sybyl version 6.5 (Tripos, Inc., St. Louis, MO, USA), used for this work, is installed on a Silicon Graphics OCTANE workstation. The following conditions have been used throughout the analysis:

- Hydrogens were added with Add hydrogens in the Biopolymer and Build/Edit menus.
- The hydrogens minimization was carried out with the Minimize option from the Compute menu, with the following parameters: Method Powell; Gradient 1; Charges Gasteiger-Huckel; Iterations 100. Other parameters were set as default.

.