Pharmaceutical Research, Vol. 22, No. 11, November 2005 (© 2005)
DOI: 10.1007/s11095-005-7624-y

Research Paper

Modeling of Human Corticosteroid Binding Globulin. Use of Structure–Activity Relations in Soft Steroid Binding to Refine the Structure*

Roy J. Little[1,2] and C. G. Rodríguez[1]

Received May 6, 2005; accepted July 25, 2005

Purpose. We propose a model for human corticosteroid binding globulin that is capable of explaining at the molecular level the experimentally observed binding affinities of four ligands. A new method of analyzing data from docking studies is proposed.
Methods. Displacement of radioactive ligand by competitive binding gives the experimentally determined binding affinities of the competitors. A theoretical model, based on homology with crystallographically determined structures, was studied in an automated docking procedure for the determination of theoretical affinities. The docking runs were analyzed by a hybrid principal component-clustering analysis.
Results. Of the two binding sites considered, only one—that in the vicinity of Cys 60—can reproduce the experimentally observed order of binding affinities although the lowest energies are found at the site in the vicinity of Cys 228.
Conclusions. Models proposed for proteins should be always conditioned to take into account experimentally observed results. In the current work, we have shown that an informed analysis of theoretical docking studies can lead to a more logical model of the protein, one that can explain and give a deeper understanding of the stereochemical requirements of binding.

KEY WORDS: antiinflammatory steroids; drug delivery; drug design; soft drugs; structure–activity relations.

INTRODUCTION

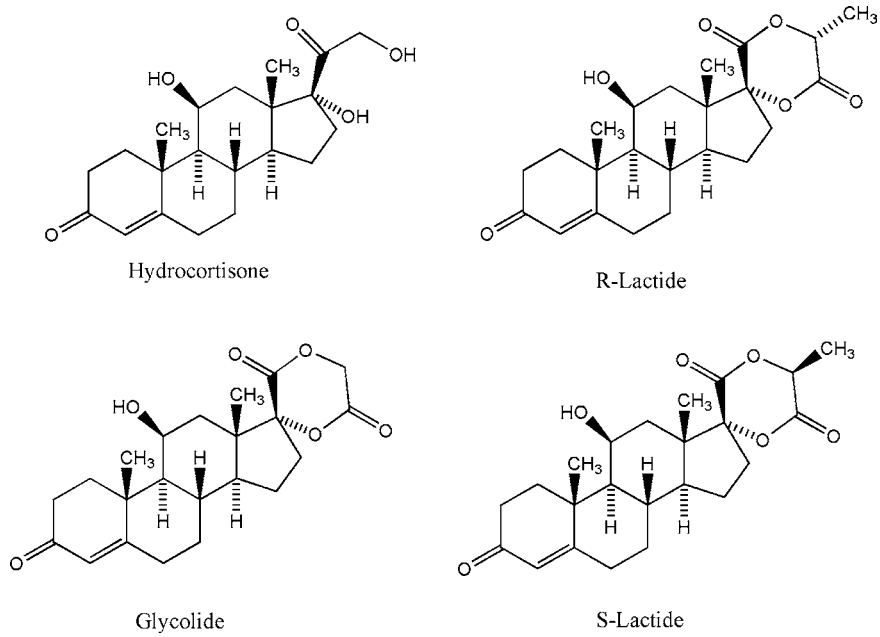
Modern drug design techniques continually emphasize the improvement of parameters governing the pharmaceutical phase (absorption, distribution, metabolism, and excretion) of drug action (1). This is in contrast with more traditional approaches in which the pharmacological phase (drug–receptor interaction) receives more attention. Specific delivery of an active drug to the site of action can be achieved through several means, including chemical delivery systems (2) as well as protein-based vehicles (3). The case of the plasma corticosteroid transport protein, corticosteroid binding globulin (CBG), is particularly interesting in this respect. Being a member of the serpin (serine protease inhibitor) class of proteins, although it is not itself an antiprotease, it has been postulated (4,5) that the protein does indeed interact with proteases found in elevated levels at an inflamed site, liberating its bound steroid, and thus serving as a specific, natural, antiinflammatory steroid delivery system. Application of this idea to the improvement of systemic antiinflammatory steroid therapy makes the investigation of the details of the binding of therapeutic corticosteroids to this protein of particular importance in the design of safer drugs.

Previously, we have reported on the development of soft drugs based on hydrocortisone (6). The current work was inspired by the soft drug approach (7), and three soft derivatives of cortienic acid, an inactive, nontoxic metabolite of the naturally occurring hormone, Hydrocortisone, were prepared so we could study their activity. Although none of the compounds demonstrated antiinflammatory activity in a number of biological assays, high affinity was observed in all three cases for CBG. In competitive binding studies in which the ability of the ligand to displace tritiated hydrocortisone from the plasma binding protein was determined, we also noted a marked stereochemical dependence in the affinity of binding in the epimeric pair of 17-spiro lactides, with one of the epimers having greater affinity (by 0.75 kcal/mol) over the other. Recently, we determined (8) that the epimer having the lower affinity has the absolute configuration of R at the epimeric position. Any proposed model of CBG must be able to explain this marked stereochemical dependence.

Experiments using the technique of electron spin resonance were used in the determination of several salient characteristics of the binding site of CBG (9). Affinity labeling experiments, with 6β-bromoprogesterone, established the importance of one of the two cysteinyl residues in the binding site (10). The structure–affinity relations observed in CBG

*A preliminary account of this work was presented at the 5th Retrometabolism Based Drug Design and Targeting Conference, Kowaki-En, Japan, May 8–11, 2005.

[1] Department Química, Facultad de Ciencias, Universidad de los Andes, Mérida, Venezuela.
[2] To whom correspondence should be addressed. (e-mail: rlittle@ula.ve)



Scheme 1. Chemical structures of synthetic ligands used in study.

binding were summarized by Westphal (11) in 1983. The primary structure of CBG was inferred from DNA sequence studies (12), at which time it was categorized as a serpin. The corroboration in this work of the existence of only two cysteinyl residues, at positions 60 and 228 of the 383-residue primary sequence, was another important result. Recent work on CBG using point mutation studies has provided strong evidence that the binding site is at Cys 60 (13). Pemberton *et al.* (4) discovered that the site of cleavage of CBG is similar to that found in nearly all serpins, near the carboxyl terminus, at the same time implying that the ligand binding site is also situated near this terminus. This has led to somewhat of a controversy about the ligand binding site of CBG. Based on a photoaffinity labeling study in which it was found that a tryptophan residue, #371, was selectively labeled by various ligands (14), the ligand binding site was proposed to involve residues near the carboxyl end of the protein. Finally, models have been proposed (15,16), in which, if there is any cysteine involved in the binding site, it would have to be Cys 228.

However, and perhaps of much greater significance, these reported models do not attempt to explain the relationship between the chemical structure of ligands and their observed biological activity in a structure–activity type of analysis. Thus, in the present work, we present the results of theoretical docking studies and propose a structural model of the protein that can partially explain the experimentally observed affinity of the ligands for the receptor. The model that we propose here, therefore, is not only a model of the structure of the protein, but of the observed binding behavior of the ligands as well.

Although the present study is only semiquantitative in terms of its analysis of structure–activity relationships, it was important for us to address one of the most fundamental assumptions in quantitative structure–activity relations, namely, that the binding modes of the members of a series of analogous compounds are indeed directly comparable and can therefore be used to establish a quantitative relation. It is one of the principal findings of this study that the results of a docking study can, and indeed should, be organized according to this principle. The problem of the significance of the individual geometries of the binding modes produced in the docking study (that is, which of them is to be chosen as representing reality) is addressed using a multivariate analysis.

The chemical structures of the four steroids used in the study are presented in Scheme 1.

MATERIALS AND METHODS

Laboratory Studies

Chemistry

The synthetic details of the preparations of two of the ligands studied in the current work (the lactides) have been previously reported (8). The glycolide was synthesized in a completely analogous fashion from 17α-bromoacetoxy cortienic acid. Hydrocortisone was purchased from Sigma (St. Louis, MO, USA). Labeled (^{3}H) hydrocortisone was purchased from New England Nuclear (Boston, MA, USA).

Binding to CBG

The procedure used was similar to that described by Berko and Pearlman (17). Whole human plasma was first precipitated with saturated $(NH_4)_2SO_4$; the supernatant was then removed and treated with 10% activated charcoal/dextran to remove both the free and bound endogenous hydrocortisone to produce a steroid-free receptor preparation. After filtering, the treated plasma was diluted to a concentration of 0.1% with Tris-molybdate buffer. Competitive binding studies were performed using ^{3}H-hydrocortisone

and the competitor in nanomolar concentrations. Free (unbound) ^{3}H-hydrocortisone was determined in a separation procedure (18) (using 1% activated charcoal/dextran) in which total performed and bound labeled steroid were quantified by scintillation counting in a Packard Tri-Carb LSC using the direct dpm option. This option was calibrated using a series of quenched ^{3}H standards supplied by New England Nuclear. Adjustments were made for nonspecific binding as well as the efficiency of the activated charcoal separation of bound vs. free ^{3}H-hydrocortisone, which were combined into one adjustable parameter in the nonlinear equation that was used. The inhibitory constant, K_I, and the total receptor concentration, $[R]_T$, were the other two parameters. The dissociation constant, K_D, for hydrocortisone was first determined in an RIA-type assay, and then used as an unadjustable parameter in the studies of the inhibitors.

Theoretical Studies

All calculations were performed on a Pentium-IV workstation using Debian Linux, version 3.1, and Linux kernel 2.4.25.

Modeling of CBG

We used the program Modeller, version 6v2 (19), to prepare a protein data bank-type structure, which was based on the templates 1ATU.pdb, 1QLP.pdb, and 1KCT.pdb, all structures of α_1-antitrypsin, which were obtained from the Protein Data Bank (20). The primary sequence used for human CBG was obtained from Swiss-Prot (21) (code number P08185). This sequence is of the CBG precursor and we removed the first 22 residues before starting the modeling procedure to obtain the 383-residue protein, starting at residue 1 (Met) in accordance with the numbering system introduced by Hammond (12). Two Modeller scripts (malign.top and get-model.top, included with the program as example scripts) were minimally modified to accommodate this study's purpose; three models were produced, the results for one of which are presented, the others being inconclusive (data not shown).

Preparation of Models of the Ligands

The crystallographically determined structure of the R-lactide served as the starting point in the construction of models of the four ligands. With the use of the programs Babel (22), version 1.6, and Ghemical (23), version 1.1, hydrogens were added to the crystallographic structure for R-lactide and the necessary modifications to the side chains were made. MOPAC, version 93 (24), was used to partially optimize the geometry of the molecules, leaving unchanged the structure based on the R-lactide, and to calculate the atomic charges on the atoms.

Automated Docking

The program Autodock (25,26), version 3.05, was used to model the docking of the four ligands to the macromolecule. The chosen macromolecule was put into final form for use with AutoDock by adding polar hydrogens (capable of forming hydrogen bonds), atomic charges, and solvation parameters to the .pdb file using the program AutoDockTools (27), which assigns charges and other parameters according to a model developed by workers at the Olson laboratory. Finally, the models of the ligands were put into final form for use with AutoDock, also using AutoDockTools (ADT),

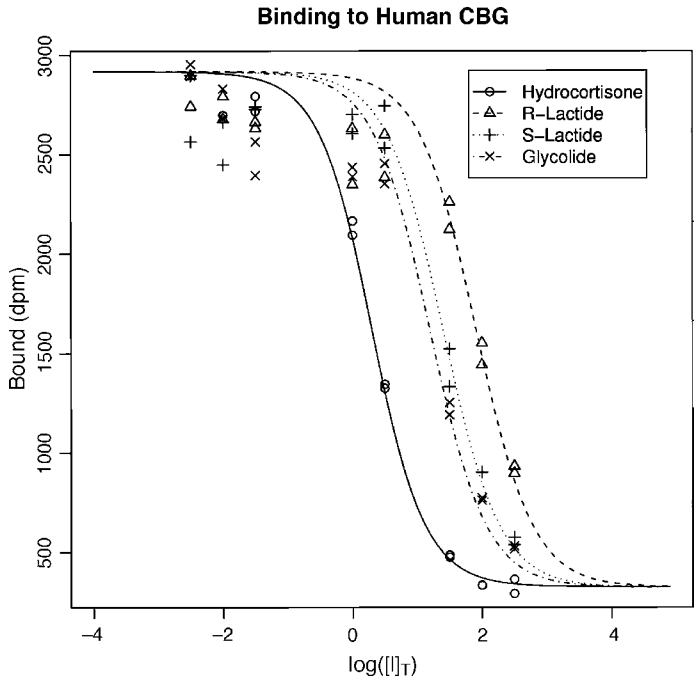


Fig. 1. Competitive displacement of ^{3}H hydrocortisone binding to human CBG. See text for details.

which assigned the rotatable groups, producing only a partially flexible ligand because ADT does not take into account torsions in rings. The charges that were calculated in the partial optimization in MOPAC93 were retained in the final models of the ligands.

The AutoDock grid was centered around the sulfur atom of either Cys 60 or Cys 228. Each of the x, y, and z dimensions of the grid was 27 Å, and the spacing of the grid points was 0.375 Å. Maps were prepared for the carbon and oxygen atoms of the ligand as well as two types of hydrogen maps, one for the polar hydrogens attached to oxygen atoms in the ligand as well as for the nonpolar hydrogens attached to carbon. Of the several available algorithms available to model the docking, the hybrid Genetic Algorithm–Local Search (GALS) was used, a decision based on our previous experience with the program. In each docking run, 300 geometries, or positions, of the ligand were produced, each of which represents a local minimum in the potential energy surface.

Statistical Analyses

Nonlinear Regression Analysis of Experimental Binding

The mathematical models used to describe the behavior of the binding of ligands to the receptor are nonlinear in parameters, and a computer program was written in the C language and compiled using the Gnu C compiler, version 3.3, to analyze the binding data. The source files and a makefile for the program are available for download (28). For the experiment in which the binding of ^{3}H-hydrocortisone was displaced by nonlabeled hydrocortisone, an RIA-type equation was used as follows.

$$[R]_T = (B(1+m) - mT)\left(\frac{K_D}{T-B} + \frac{[I]_T}{T} + 1\right) \qquad (1)$$

The adjustable parameters are K_D, the dissociation constant for hydrocortisone, $[R]_T$, the total nanomolar concentration of the receptor, and m, the nonspecific binding of the ligand in which the adjustment for the efficiency of the separation procedure is included. The data are B and T, denoting the bound and total radioactivity, respectively, converted to units of nanomolar concentrations by a conversion factor.

The corresponding equation for the competition experiments is as follows.

$$[R]_T = (B(1+m) - mT)\left(\frac{K_D}{T-B} + \frac{K_D[I]_T}{K_I(1+m)(T-B) + K_D(B(1+m) - mT)} + 1\right) \qquad (2)$$

In this equation, K_D, R_T, and m are nonadjustable parameters whose values were determined in the previous (RIA) analysis. In nonlinear regression analysis, it is necessary to use derivatives of the above model equations and in this case, analytical derivatives were used. As the derivatives were somewhat complex, they were checked for accuracy by using the symbolic algebra module of Mathematica, version 4.1.

Analysis of Docking Studies

Several utility programs were written in the Python (29) language for extraction of data from the docking runs as well as in the statistical programming language R (30), for the additional analyses (principal components and hierarchical clustering) that were necessary for this work. We used the RMSD (root mean square deviation) utility in the program VMD (Visual Molecular Dynamics) (31,32) to calculate the value of this variable.

As noted previously, results of the Autodock studies are the positions of the ligand, each representing a local minimum of the potential energy function, given in terms of pdb format records, which additionally contain the intermolecular and internal energies, as well as the van der Waals and electrostatic potentials of each atom of each geometry of the ligand. In the table, the results for one (the first) of the geometries (designated here as MODEL 1) of the hydrocortisone binding in the vicinity of Cys60 AutoDock run are partially shown. Entries for x, y, and z and other data continue up to atom #56, the number of atoms in hydrocortisone. The problem of which geometry to choose to represent the true binding of the ligand begins with the recognition that these results form clusters, or groups, of geometries, reflecting different modes and energies of binding. It is then a matter of deciding which group is representative of the true binding. It is the principal goal of the following procedure to find this representative group.

```
DOCKED: MODEL        1
DOCKED: USER    Run = 1
DOCKED: USER    DPF = kf-cbgE-cent.dpf
DOCKED: USER    Estimated Free Energy of Binding    =   -9.58 kcal/mol  [=(1)+(3)]
DOCKED: USER   Estimated Inhibition Constant, Ki   =   +9.58e-08  [Temperature = 298.15 K]
DOCKED: USER
DOCKED: USER    Final Docked Energy                 =   -9.26 kcal/mol  [=(1)+(2)]
DOCKED: USER
DOCKED: USER    (1) Final Intermolecular Energy     =   -9.58 kcal/mol
DOCKED: USER    (2) Final Internal Energy of Ligand =   +0.31 kcal/mol
DOCKED: USER    (3) Torsional Free Energy           =   +0.00e+00 kcal/mol
DOCKED: USER
DOCKED: USER                              x       y       z     vdW  Elec       q
DOCKED: USER
DOCKED: ATOM      1  C   UNK     1     -14.520   3.444  19.350 -0.56 -0.01    -0.101
DOCKED: ATOM      2  C   UNK     1     -13.008   3.483  19.459 -0.58 -0.04    -0.145
.
.
. etc.
```

Table I. RIA Study of Hydrocortisone Binding to Human CBG

Parameter	Value	SE (n = 16)
K_D	1.0036	0.138
$[R]_T$	0.490	0.053
m	0.035	0.0056

Values are expressed in nanomolar (nM).

Table II. Competitive Binding to Human CBG

	R-lactide	S-lactide	Glycolide
K_I	47.6 ± 8.4	14.6 ± 2.4	9.05 ± 1.31

Values are expressed in nanomolar (nM).

To extract information from and organize these data, a principal components analysis was first performed on five variables chosen to characterize the geometries: intermolecular binding energy (designated as E1); internal energy of the ligand (designated as E2) (E1 and E2 being directly taken from the docking run); and three root mean square deviations (rmsd). These rmsd variables were from: (1) the most stable geometry of the run of hydrocortisone, designated rmsdb, and (2 and 3) the atoms of the two residues presumed to have importance in binding, Cys 60 and Arg 64, designated as rmsdc and rmsdr, respectively. For the analogous analysis of binding near Cys 228, the atoms of Cys 228 and Trp 371, designated as rmsdc and rmsdw, respectively, were used. As an example, again for the case of hydrocortisone binding in the vicinity of Cys 60, loadings of these variables on the five principal components are shown. Naturally, each AutoDock run has its own set of PC coefficients. A hierarchical clustering analysis was subsequently performed based on the scores of the geometries on these five principal components. We cut the dendrogram for all cases at 20 clusters, which we found to be adequate for our purposes. In comparison, AutoDockTools also performs a cluster analysis based on deviations from the most stable base structure, but this analysis often gives from 70 to over 100 clusters.

RESULTS

Binding to CBG

The binding of the cold ligand, hydrocortisone, to human CBG in displacing the labeled compound obeys an RIA-type mathematical behavior. Results are presented in Fig. 1 and Table I. Sixteen data points were determined for each ligand, as indicated in the graph.

A description of the parameters, which are expressed in nanomolar (nM), is found in Materials and Methods. In the graph, $[I]_T$ is also expressed in nanomolar.

Binding of the test ligands in competition with labeled hydrocortisone obeys Eq. (2), which describes competitive inhibition, and the results are presented in Table II. The smooth curves are calculated based on experimentally determined parameters.

Modeling of CBG

This model is shown in Fig. 2, which was constructed to make comparison with α_1-antitrypsin [Fig. 1 of Gettins's (33) review article on serpins] easier for the identification of secondary structural elements. A comparison of the above structure to that of Gettins's shows a great deal of similarity. The three principal beta sheets are present as well as all but one of the helices (hJ). There are some subtle differences in the types of helices and some that are incomplete. In particular, helix hI is very ill defined and, as noted, helix hJ is not present. In the figure we have emphasized the locations of the two cysteines present in the protein. The importance of hydrophobic interactions is well known in steroid binding, and we have thus emphasized the helices surrounding the two cysteinyl groups: hB and hC around Cys 60, and hH and hG around Cys 228. The binding site proposed by analogy with that of thyroxin binding hormone (16) is in the vicinity of Cys 228, apparently formed by the β-strands B and C, although cleaved α_1-antitrypsin was used, in spite of the knowledge that binding to CBG is greatly reduced upon cleavage (4). In the model shown in Fig. 2, a consideration of the location of Trp 371 relative to that of Cys 228 is important. Helix H is interposed between them, making simultaneous interaction with the ligand improbable. Thus, this model can not explain the importance of Trp 371 as well as that of Cys 228.

Docking Studies

Binding to the Vicinity of Cys 60

Results of binding to this vicinity are presented graphically in Fig. 3. The geometries corresponding to the cluster chosen to represent binding of the steroid to the true binding site are shown as addition symbols ("+"). The other runs are indicated in the graphs by plotting the points as their cluster numbers. The salient feature of the graphs is that the lowest energy cluster for each of the first three ligands—hydrocortisone, glycolide, and S-lactide—is also the lowest with respect to the RMSD from the comparison base structure, which in all cases was the lowest energy geometry from the hydrocortisone run. This is in contrast with the

```
Importance of components:
                          PC1    PC2     PC3     PC4    PC5
Standard deviation     7.636 1.6538 0.67730 0.40993 0.0786
Proportion of Variance 0.945 0.0444 0.00744 0.00272 0.0001
Cumulative Proportion  0.945 0.9897 0.99718 0.99990 1.0000
              PC1          PC2          PC3          PC4          PC5
E1     0.119103581 -0.065165752  0.95924561 -0.24306542 -0.048320023
E2    -0.005712936  0.008250563 -0.04367962  0.02115922 -0.998771079
rmsdb  0.886991638 -0.441046546 -0.12786289  0.04868752 -0.002093589
rmsdc  0.250804563  0.473977826 -0.20660242 -0.81836645 -0.005821077
rmsdr  0.368950680  0.759281448  0.13750082  0.51805151  0.009123500
```

situation for the R-lactide, in which the lowest energy cluster has an average RMSD of about 7.

The proposed binding site for the corticosteroids is presented as a stereogram in Fig. 4, in which we have emphasized in blue five residues (in addition to Cys 60, the sulfur atom of which is represented as a yellow van der Waals sphere) that have important interactions with the ligands. These are the two arginines, Arg 64 and 311, Gly 58, Ser 120, and Glu 119, which have important interactions with the polar groups of the D- and E-rings, and Leu 57. The ligands are represented in the binding site as solid "licorice" style bonds, the hydrogen atoms not being represented for clarity. All of the heavy atoms of hydrocortisone are represented: carbon (black) and oxygen (red). Only the glycolide and S-lactide ring systems are represented, as gray and tan, respectively. Finally, two structures are presented for the R-lactide: the most similar to the reference geometry, in green, and the lowest energy, in red. It is readily apparent that the red structure is occupying an alternative site, sandwiched between Arg 64 and Glu 119. The more comparable green structure is slightly rotated with respect to the other steroids, thus incurring an unfavorable interaction with Leu 57 upon being obliged to avoid an even more unfavorable interaction (perhaps with Arg 64) through its epimeric methyl group on the E-ring.

Binding to the Vicinity of Cys 228

Results of this binding study are presented in Fig. 5, in analogous fashion to those for the Cys 60 site. In this case, the lowest energy geometry of the hydrocortisone run was also used as the standard of comparison for the calculation of RMSD in all four ligands. In contrast to the case of binding near Cys 60, all the lowest energy clusters are lowest in RMSD as well. Additionally, the binding energies are much lower than those observed in the previous case. The binding site with the four ligands in their lowest energy geometries is shown in Fig. 6. The discussion on the representations of ligands for Fig. 4 also applies in this case. There are, however, serious inconsistencies with this postulated binding site. The first and most obvious is the separation between Cys 228 and Trp 371: they are separated by a distance of 18 Å (the length of hydrocortisone is 12.7 Å from O-3 to 21-OH) and helix H is positioned between them, rendering impossible any putative simultaneous interaction between the steroid and the two residues. In two models that have been proposed (15,16) for h-CBG, the importance of cysteine is ignored. Thus, a choice must definitely be made between the two

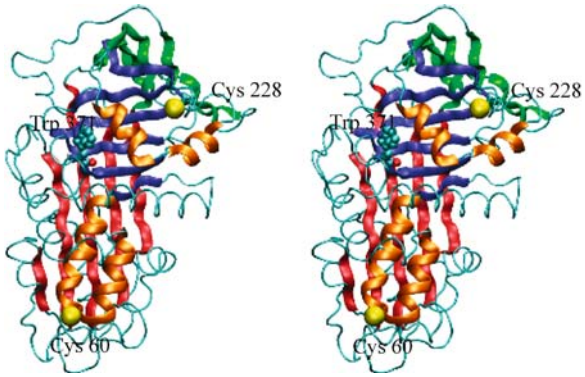


Fig. 2. Human CBG model. In addition to the two cysteines, sheets A, B, and C (red, blue, and green, respectively), and helices B and C (near Cys 60) and helices G and H (near Cys 228) are emphasized.

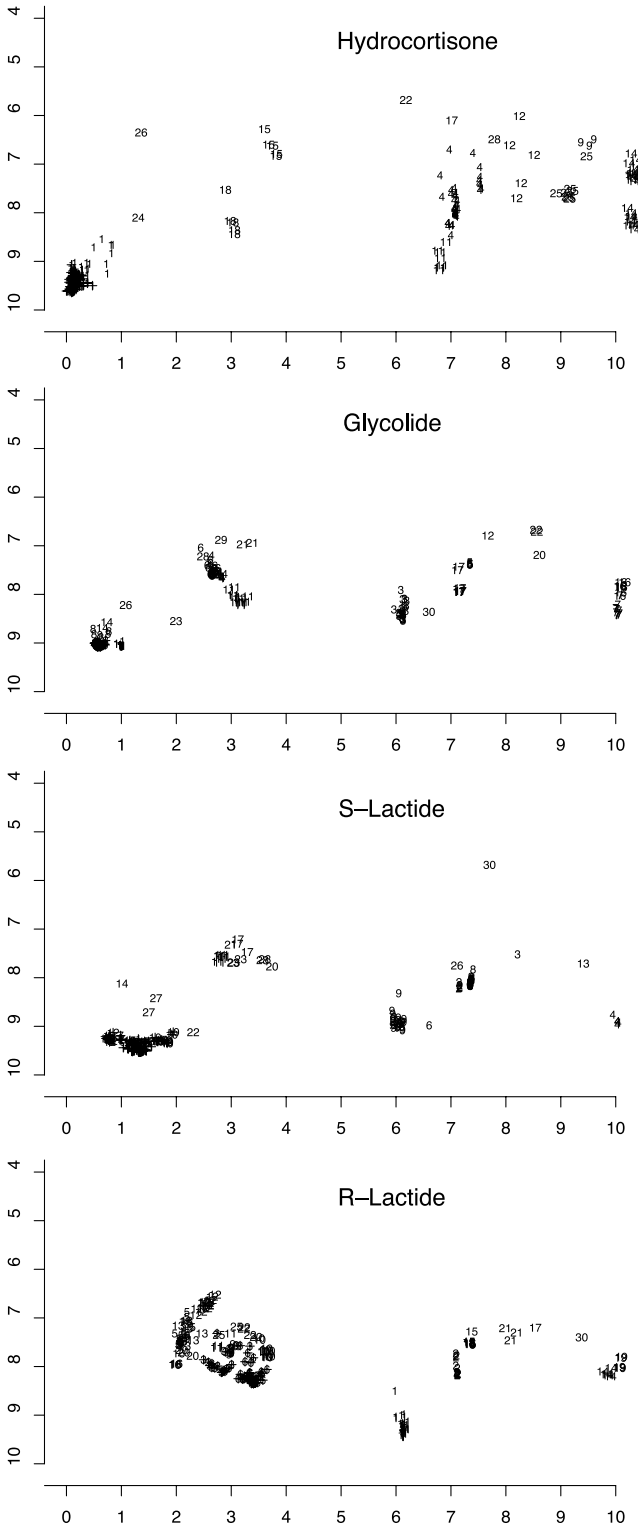


Fig. 3. Docking study of four ligands to the vicinity of Cys 60. Vertical axis is free energy of binding (in kcal/mol); horizontal axis is RMSD from the lowest energy hydrocortisone geometry. See text for details.

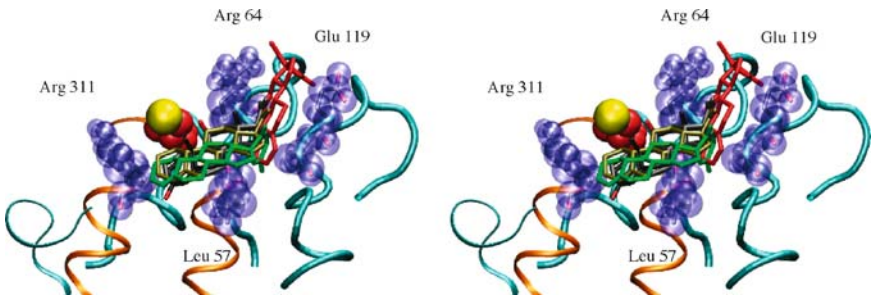


Fig. 4. Stereogram of postulated binding site in vicinity of Cys 60 with four ligands. Important residues in the protein are emphasized with van der Waals spheres. Two structures (red and green) are presented for the R-lactide.

residues, because corticosteroid cannot interact with both in the binding site. To further test the hypothesis that corticosteroids bind in the vicinity of Cys 228, we related the calculated binding energies to those observed experimentally. Results are presented in Fig. 7. These results conclusively contradict the hypothesis that the binding of corticosteroids is in the vicinity of Cys 228, because the correlation is negative. Thus, the steroid with the lowest affinity for the receptor, the R-lactide with its epimeric methyl group presumably exercising some steric inhibition that is not found in the other cases, has the greatest affinity at this site. In contrast, the correlation, although not as good in statistical terms, in the case of binding near Cys 60 is positive. The percentage of geometries that were chosen (the "+" symbol in Figs. 3 and 5) to represent the docking for ligand for comparison of the theoretical binding of the ligands is presented in Table III.

DISCUSSION

Experimental evidence concerning the mode of binding and the determination of the structural elements necessary for a compound to have high affinity for the protein has traditionally been based on structure–activity relationships. A model for the binding site was proposed for guinea pig CBG (34) based exclusively on structure–affinity relationships. Before the advent of techniques—above all, X-ray crystallography—allowing a view of the molecular interactions, structure–activity relations were the only means to attain this perspective. In time, the technique was largely superseded by more modern methods. To illustrate an example, the mechanism of chymotrypsin-catalyzed proteolysis was thoroughly studied via structure–inhibitor relations and was concluded to involve a histidine residue in its active site during hydrolysis (35). What could never really be concluded from this previous work was subsequently shown from the determination of the three-dimensional structure of α-chymotrypsin: a proton transfer chain involving the sequence Asp–His–Ser that could only be seen in the crystallographic structure (36). It is instructive to note that in the review article by Blow (36), not a single reference was made to the extensive list of previous works based on structure–activity relations. The objective of reaching the same level of understanding with respect to CBG binding and activity has been the implicit goal of considerable research. The affinity labeling study (10), in which the importance of one of two titratable Cys residues in the protein was discovered, was performed under conditions very similar to those used in the present study, and we thus consider it of principal relevance to our work. The use of a spin-labeled ligand in the electron spin resonance (ESR) study (9), although also corroborating the importance of Cys in the binding site, is not strictly comparable because of the size of the nitroxyl radicals coupled to the steroid to give a detectable ESR signal. The postulation of a "hydrophobic pocket" approximately 25 Å deep dates from this study. Introduction of the technique of site-directed mutagenesis to CBG has resulted in two contradictory studies. Ghose-Dastidar *et al.* (13) mutated Cys 228 to Ser and Ala and noted no reduction in binding affinity for corticosteroids, thus providing strong evidence that the important Cys is at position 60. In another site-directed mutagenesis study, Avvakumov and Hammond (37) prepared mutants in which they replaced the Trp residues of the protein and concluded that Trp 371 is important in binding. A photoaffinity labeling study also implicated Trp 371 in the binding of 6-dehydrocorticosteroids (14). The focus of subsequent research has continued to be on the carboxyl end of the protein, perhaps because this is the region in which cleavage occurs in the presence of, for example, neutrophil elastase. Thus, models have been proposed by analogy with cleaved α_1-antitrypsin (17) and homology with uncleaved (15) α_1-antitrypsin, which also emphasize the carboxyl terminus of the protein. In these later studies, not only has the carboxyl terminus been given greater importance, but the cysteine residue, so predominant in earlier research, has been largely forgotten. In our work, we have chosen to reemphasize this residue.

Use of a modeling method to gain insight into the energetics of the binding of a ligand to the active site of a protein necessarily involves many caveats. The principal exception in the present case is the restriction by AutoDock to a rigid receptor and a minimally flexible ligand. This limitation is counteracted by the great number of dockings that may be realized, giving a more global picture of the energetics of many possible binding sites. Thus, certain characteristics of the binding site may be, with a certain degree of confidence, be proposed. The Cys 60 group positioned near the center of the ligand supports our conclusion that the active site should possess a geometry similar to this, because in the affinity labeling experiment, 6α-cysteinylprogesterone was obtained from incubation of CBG with 6β-bromoprogesterone, the 6-position being near the center of the steroid molecule.

The clustering method chosen in this work for treatment of docking data results in clusters with many elements

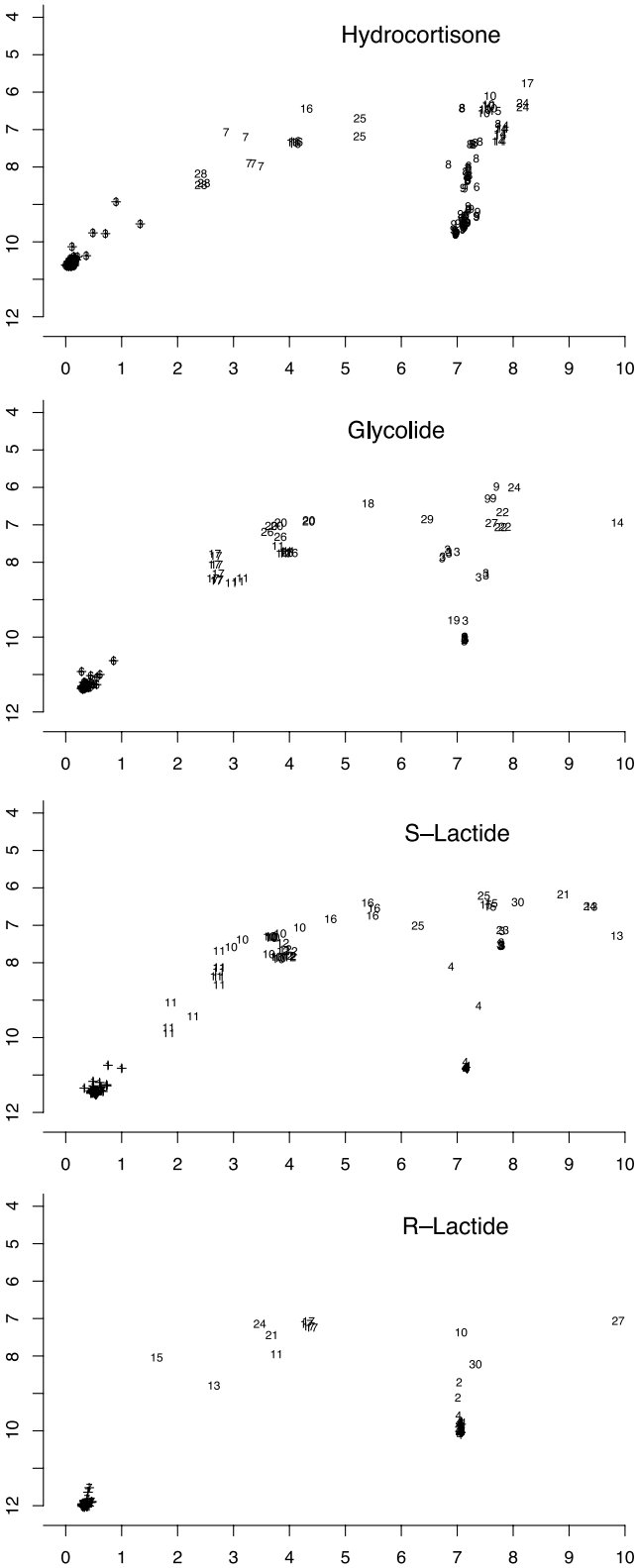


Fig. 5. Docking study of four ligands to the vicinity of Cys 228. Vertical axis is free energy of binding (in kcal/mol); horizontal axis is RMSD from the lowest energy hydrocortisone geometry. See text for details.

(geometries), suitable for statistical analysis. In this sense, structures in the clusters do not vary greatly in a translational sense nor do they flip their configuration in a rotational sense, a problem that is only resolved in other methods of clustering by making the number of clusters very large and, consequently, the number of elements per cluster small. This consideration is important, because in our current work we are using the least squares coefficients of the van der Waals potentials of atoms as molecular descriptors in a more extensive quantitative structure–activity relationship (QSAR) study, and it is necessary that each cluster contain at least seven members to carry out the partial least squares determination of the regression coefficients of the atoms.

Treatment of the docking data is new in the sense that all of the data are considered with respect to their ability to explain experimental results. This is in contrast with other studies having the unstated goal of finding the "true" binding site that would be identified by its being the "global" minimum. In all cases, we have tried to use the lowest energy cluster, in accordance with previous uses of the AutoDock program. Only in the case of binding of the R-lactide to the Cys 60 site have we chosen a different, but nonetheless, in our opinion, a comparable cluster.

The binding site near Cys 228 is not the "hydrophobic pocket" postulated (9) in the ESR study, and a careful reading of the work of Dey and Roychowdhury (15) (their Fig. 3) leads to the conclusion that that particular binding site is similar to the one found in this study in the vicinity of Cys 228.

CONCLUSION

In a theoretical study of the binding of ligand to receptor, the question of whether different binding activities reflect a continuous gradation in the orientation of the ligand in one binding site or whether there are fundamentally different modes of binding is inescapable. A mathematical analysis of the relative importance of structural descriptors of the chemical structure of the ligands is only permissible in the case of the former. In reference to the work of Mickelson and Westphal (34), the observation is made that, in comparing cortisol binding with that of cortisol reduced at the 20-oxo position, a nearly 300-fold reduction in affinity occurs. The argument is then made that there must be a hydrogen-bond donor in the vicinity of the 20-oxo group, because if there is to be a lessening in affinity in going to the reduced compound, this reduced compound *must* bind in the same fashion as cortisol itself. This is a statement of the fundamental assumption in all QSAR studies. The search for the "global minimum" does not necessarily equate with that of the "true" binding site, and both goals are perhaps unattainable. Thus, in comparing the results for binding in the vicinity of Cys 228 with those near Cys 60, one would naively conclude that, based on the lower energies of binding at that site, the binding site is indeed near Cys 228. However, our contention is that experimental binding data, in the form of structure–activity relations, must be taken into account when modeling a binding site, and even if one is dealing with the crystallographically determined three-dimensional structure of the protein, it still remains that this structure does not necessarily reflect the situation *in vivo*. Thus, the negative correlation between experimentally observed binding affinity and that

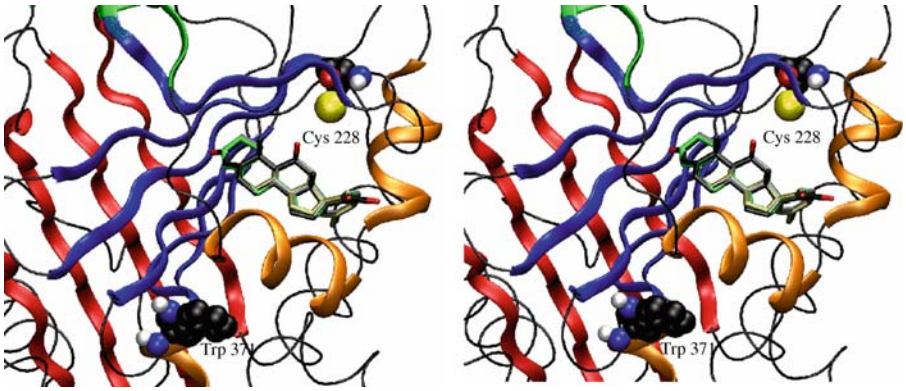


Fig. 6. Stereogram of postulated binding site in vicinity of Cys 228 with four ligands.

determined by AutoDock contradicts the conclusion reached upon when one injudiciously considers only the energies calculated by AutoDock, and prompts us to seriously doubt the binding site near the carboxyl terminus as the true binding site.

In the present work, we have a concrete example of the problem of which structures are to be compared in the case of the R-lactide. With respect to docking near Cys 60, the most comparable cluster is not the lowest energy cluster that is quite removed spatially from the lowest energy clusters of the three other ligands, and would seem to be occupying an alternative site, sandwiched between Arg 64 and Glu 119. Consequently, we can make a semiquantitative deduction concerning the important structural factors related to biological activity, namely, that the orientation of the methyl group on the spiro lactide ring in the R configuration impedes the formation of stabilizing interactions with Ser 120 (in the case of hydrocortisone) or Gly 58 (in the case of the glycolide and S-lactide), and more toward a destabilizing interaction with Leu 57. More importantly, although we have emphasized the qualitative aspects of the steroid-CBG binding, the goal of any structure–activity relation, be it quantitative or otherwise, is the same: a deeper understanding of the events occurring at the molecular level. With the techniques we have introduced here, the questions relating to molecular mechanisms of the reduction of affinity upon cleavage can now be studied in much greater detail. This work represents an effort to integrate structure–activity relations into studies of modeling of receptors.

Table III. Percent Occurrence of Representative Geometries in Docking Runs

	Cys 60	Cys 228
Hydrocortisone	38.3	17.0
Glycolide	25.7	11.7
S-Lactide	17.7	16.7
R-Lactide	25.7	13.0

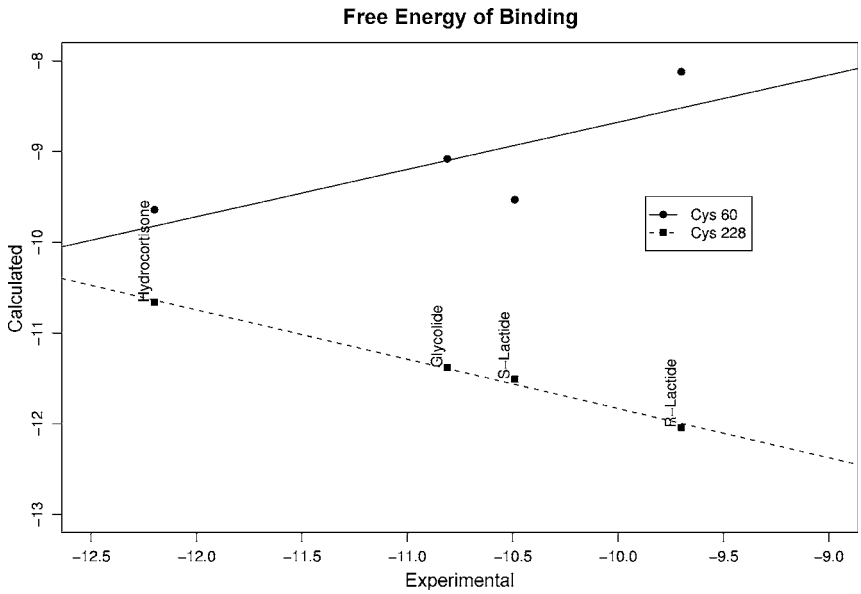


Fig. 7. QSAR plot showing calculated vs. observed binding affinities for the two models of the ligand binding site of human CBG.

ACKNOWLEDGMENTS

This work was financed in part by a grant from the Consejo de Desarrollo Científico of the Universidad de los Andes, project # C11890308B. The financial support of the University of Florida College of Pharmacy and of the Otsuka Pharmaceutical Company of Osaka, Japan, is gratefully acknowledged. We also wish to acknowledge the global effort of the Free Software Movement (Richard S. Stallman, founder), which in great part has made possible the current study. A great majority of the programs used in this work are freely available and modifiable by the user, who thereby becomes part of a global effort to improve the software. The corrections and helpful comments of Prof. Zaida Parra are also gratefully acknowledged.

REFERENCES

1. N. Bodor. Recent advances in retrometabolic drug design and targeting approaches. *Pharmazie* **55**:163–166 (2000).
2. N. Bodor. Retrometabolic drug design—novel aspects, future directions. *Pharmazie* **56**(Suppl 1), S67–S74 (2001).
3. F. A. De Wolf and G. M. Brett. Ligand-binding proteins: their potential for application in systems for controlled delivery and uptake of ligands. *Pharm. Rev.* **52**:207–232 (2000).
4. P. A. Pemberton, P. E. Stein, M. B. Pepys, J. M. Potter, and R. W. Carrell. Hormone binding globulins undergo serpin conformational change in inflammation. *Nature* **336**:257–258 (1988).
5. G. L. Hammond, C. L. Smith, N. A. Paterson, and W. J. Sibbald. A role for corticosteroid-binding globulin in delivery of cortisol to activated neutrophils. *J. Clin. Endocrinol. Metab.* **71**:34–39 (1990).
6. R. J. Little, N. Bodor, and T. Loftsson. Soft drugs based on hydrocortisone. The inactive metabolite approach and its application to steroidal antiinflammatory agents. *Pharm. Res.* **16**: 961–967 (1999).
7. N. Bodor and P. Buchwald. Designing safer (soft) drugs by avoiding the formation of toxic and oxidative metabolites. *Mol. Biotechnol.* **26**:123–132 (2004).
8. A. Mora, B. Ramírez, G. Delgado, and R. J. Little. The 17-spiro lactide of cortienic acid. A probe for studying the active sites of steroidal receptors. *Acta Crystallogr.* **C55**:1870–1873 (1999).
9. G. Defaye, M. Baset, N. Monnier, and E. M. Chabaz. Electron spin resonance study of human transcortin: thiol groups and binding site topography. *Biochim. Biophys. Acta* **623**:91–97 (1980).
10. M. S. Khan and W. Rosner. Investigation of the binding site of human corticosteroid-binding globulin by affinity labeling. *J. Biol. Chem.* **252**:1895–1900 (1977).
11. U. Westphal. Corticosteroid-binding globulin. A review of some recent aspects. *Mol. Cell. Biochem.* **55**:145–157 (1983).
12. G. L. Hammond, C. L. Smith, I. S. Goping, D. A. Underhill, M. J. Harley, J. Reventos, G. L. Gunsalus, and C. W. Bardin. Primary structure of human corticosteroid binding globulin, deduced from hepatic and pulmonary cDNAs, exhibits homology with serine protease inhibitors. *Proc. Natl. Acad. Sci. USA* **84**:5153–5157 (1987).
13. J. Ghose-Dastidar, R. Green, and J. B. Ross. Identification of the cysteine in the steroid-binding site of human corticosteroid binding globulin by site-directed mutagenesis and site-specific chemical modification. *J. Steroid Biochem. Mol. Biol.* **48**:139–144 (1994).
14. C. Grenot, T. Blachere, M. R. Deravel, E. Mappus, and C. Y. Cuilleron. Identification of Trp-371 as the main site of specific photoaffinity-labeling of corticosteroid-binding globulin using delta-6-derivatives of cortisol, corticosterone, and progesterone as unsubstituted photoreagents. *Biochemistry* **33**:8969–8981 (1994).
15. R. Dey and P. Roychowdhury. Homology model of human corticosteroid binding globulin: a study of its steroid binding ability and a plausible mechanism for steroid hormone release at the site of inflammation. *J. Mol. Model* **9**:183–189 (2003).
16. P. Edgar and P. Stein. Hormone binding site of corticosteroid binding globulin. *Nat. Struct. Biol.* **2**:196–197 (1995).
17. R. M. Berko and W. H. Pearlman. The influence of 21-acylation of corticosterone on its binding affinity for corticosteroid-binding globulin. *J. Steroid Biochem.* **14**:1035–1039 (1981).
18. W. D. Odell. Use of charcoal to separate antibody complexes from free ligand in radioimmunoassay. *Methods Enzymol.* **70**:274–279 (1980).
19. A. Sali and T. L. Blundell. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**:779–815 (1993).
20. The RCSB Protein Data Bank. http://www.rcsb.org/pdb/ (accessed 4/22/05).
21. ExPASy – Swiss-Prot and TrEMBL. http://us.expasy.org/sprot/ (accessed 4/22/05).
22. babel http://www.ccl.net/cca/software/UNIX/babel/index.shtml.
23. Ghemical Homepage http://www.uku.fi/~thassine/ghemical/ (accessed 4/22/05).
24. Quantum Chemistry Program Exchange. http://qcpe.chem.indiana.edu Program #689 (accessed 4/22/05).
25. G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, and A. J. Olson. Automated docking using a Lamarckian genetic algorithm and empirical binding free energy function. *J. Comput. Chem.* **19**:1639–1662 (1998).
26. AutoDock http://www.scripps.edu/mb/olson/doc/autodock/index.html (accessed 4/22/05).
27. Tools and UNIX shell scripts for AutoDock. http://www.scripps.edu/pub/olson-web/doc/autodock/tools (accessed 22/4/05).
28. Competitive Binding Assay Analysis. http://webdelprofesor.ula.ve/ciencias/rlittle/comp.tgz (Accessed 28/6/05).
29. Python Programming Language http://www.python.org/ (accessed 4/22/05).
30. R Development Core Team. R: A language and environment for statistical computing. http://www.R-project.org (accessed 4/22/05).
31. W. Humphrey, A. Dalke, and K. Schulten. VMD—visual molecular dynamics. *J. Mol. Graphics* **14**:33–38 (1996).
32. VMD—Visual Molecular Dynamics. http://www.ks.uiuc.edu/Research/vmd/ (accessed 4/22/05).
33. P. G. W. Gettins. Serpin structure, mechanism, and function. *Chem. Rev.* **102**:4751–4803 (2002).
34. K. E. Mickelson and U. Westphal. Steroid–protein interactions. Influence of steroid structure and temperature on the binding of steroids to guinea pig corticosteroid-binding globulin. *Biochemistry* **19**:585–590 (1980).
35. G. E. Hein and C. Niemann. Steric course and specificity of α-chymotrypsin-catalyzed reactions. II. *J. Am. Chem. Soc.* **84**: 4495–4503 (1962).
36. D. M. Blow. Structure and mechanism of chymotrypsin. *Acc. Chem. Res.* **9**:145–152 (1976).
37. G. V. Avvakumov and G. L. Hammond. Substitutions of tryptophan residues in human corticosteroid-binding globulin: impact on steroid binding and glycosylation. *J. Steroid Biochem. Mol. Biol.* **49**:191–194 (1994).