



A comparative study of ligand-receptor complex binding affinity prediction methods based on glycogen phosphorylase inhibitors

Sung-Sau So^{a,*,**} & Martin Karplus^{a,b,*}

^aDepartment of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138, U.S.A.; ^bLaboratoire de Chimie Biophysique, Institut le Bel, Université Louis Pasteur, 4 rue Blaise Pascal, F-67000 Strasbourg, France

Received 24 June 1998; Accepted 1 October 1998

Key words: binding affinity prediction, CoMFA, genetic neural network, glycogen phosphorylase inhibitor, QSAR, structure-based drug design

Abstract

Finding an accurate method for estimating the affinity of protein ligands activity is the most challenging task in computer-aided molecular design. In this study we investigate and compare seven different prediction methods for a set of 30 glycogen phosphorylase (GP) inhibitors with known crystal structures. Five of the methods involve quantitative structure-activity relationships (QSAR) based on the 2D or 3D structures of the GP ligands alone. They are hologram QSAR (HQSAR), receptor surface model (RSM), comparative molecular field analysis (CoMFA), and applications of genetic neural network to similarity matrix (SM/GNN) or conventional descriptors (C2GNN). All five QSAR-based models have good predictivity and yield q^2 values ranging from 0.60 to 0.82. The other two methods, LUDI and a structure-based binding energy predictor (SBEP) system, make use of the structures of the ligand-receptor complexes. The weak correlation between biological activities and the LUDI scores of this set of inhibitors suggests that the LUDI scoring function, by itself, may not be a general method for reliable ranking of a congeneric series of compounds. The SBEP system is derived from a set of physical properties that characterizes ligand-receptor interactions. The final neural network model, which yields a q^2 value of 0.60, employs four descriptors. A jury method that combines the predictions of the five QSAR-based models leads to an increase in predictivity. A multi-layer scoring system that utilizes all seven prediction methods is proposed for the evaluation of novel GP ligands.

Introduction

One of the major challenges in computational approaches to pharmacological problems is the reliable prediction of the binding affinity of a ligand for a receptor. Quantitative structure-activity relationships (QSAR), an analysis that derives mathematical relationships between physico-chemical properties of molecules and their biological activities, has been the principal method for activity prediction for systems where the receptor structure is unknown [1]. Until 10 years ago, the majority of descriptors used in such

correlation studies were substituent parameters representing specific properties of chemical group [2] or topological indices encoding molecular connectivity using graph theoretical concepts [3, 4]. This type of model is classified as 2D QSAR since the descriptors do not capture any 3D information concerning the ligands, and furthermore their magnitudes are independent of the specific conformation or orientation of the molecules, i.e., they are spatially invariant. A new 2D technique called HQSAR has been introduced recently [5–9]. In this method, a chemical structure is converted to a characteristic molecular fingerprint based on enumerating the presence of certain type of molecular fragments. This numerical representation of molecules is used as the QSAR descriptor.

*To whom correspondence should be addressed.

**Present address: Hoffmann-La Roche, 340 Kingsland St., Nutley, NJ 07003, U.S.A.

With the availability of fast graphics workstations and the increasing popularity of commercial modeling packages, a growing number of QSAR studies now incorporate 3D information for the ligands, which, in principle, provides a more detailed analysis of putative ligand-receptor interactions. A caveat concerning 3D QSAR methods is that the conformation or spatial orientation of the molecules is an inherent, and often critical, element of the analysis. Thus, when the molecular orientation of a given system is not obvious, the uncertainty of the bioactive conformation may constitute a problem. Traditionally, the simplest way to incorporate spatial information is the inclusion of standard 3D descriptors, for example, molecular moments [10] or their Cartesian components. More advanced techniques attempt to model the receptor environment from the perspective of the ligand structures. The GRID [11] and the CoMFA [12] programs place the drug molecules on a lattice and sample their molecular interactions with a steric or an electrostatic probe at evenly spaced grid points. QSAR models can be obtained by application of multivariate partial least squares (PLS) method to the interaction field matrix. Alternatively, the field matrix can be compressed to a molecular similarity matrix (SM), for which QSARs can be derived using PLS [13, 14] or genetic neural network (GNN) methods [15–18]. Recently Hahn has introduced a 3D method called receptor surface model (RSM). The program creates a pseudo-receptor surface based on the putative binding conformations of a small number of active compounds [19, 20]. The interaction between the ligand and the modeled surface is then used to make the activity prediction.

Some prediction methods have been developed to take advantage of the actual binding site environment when the 3D structure of the ligand-bound receptor is known. For example, programs like LUDI [21, 22], SMOG [23, 24] and Hammerhead [25] utilize empirical scoring functions to predict free energy of binding. Though the complexity of the scoring functions may differ, this kind of prediction method tends to be fast. Other studies relate the binding affinity between a receptor and a ligand with their interaction energies. They may be obtained using the energy minimized structures of the complex [26], or from an ensemble average over a short molecular dynamics trajectory [27]. There are also approaches which decompose the free energy of binding into different contributions. Many of the methods have been discussed in a recent review by Ajay and Murcko [28].

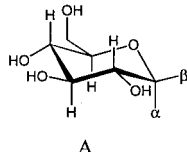
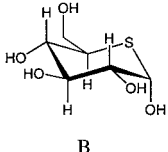
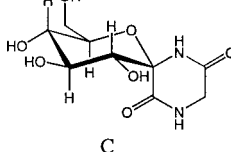
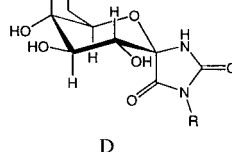
Free energy simulations provide, in principle, the most rigorous method for calculating the difference in ΔG_{bind} between two reasonably similar ligands, though absolute binding energies can be estimated as well [29]. Application of this approach to activity prediction has been limited by two factors. First, the calculations are computationally very demanding; and second, in many cases, new ligands are considerably different from the known set of molecules, so that correct converged results are difficult to achieve, in part because the atomic positions of the compounds may not be available and are difficult to calculate.

The aforementioned methods are a sample of the wide range of the activity prediction methods that are currently available. Some of the calculations can be performed using modest equipment, and some require supercomputing power. However, none of the approaches is satisfactory by itself (i.e. calculates accurate binding affinities for a wide range of molecules in a reasonable time). The purpose of the present study is twofold. First, by comparing a set of different methods for a well-defined problem, it gives some insights into their strong and weak points. For this purpose we chose to use existing commercial programs in most cases. Second, given the results a strategy is devised that makes the best collective use of a number of methods. We investigate seven methods for activity predictions and combine their predictions in a meaningful way. Since each prediction algorithm has limitations which arise from different sources, a jury decision that is based on several models is likely to give a more reliable prediction than does any one alone.

We apply the prediction methods to a set of glycogen phosphorylase (GP) inhibitors. GP is a large enzyme that plays a regulatory role in glycogen metabolism. Since GP is a potential therapeutic target relating to the treatment of diabetes, there has been considerable interest in designing a more potent GP inhibitor than α -D-glucose, the physiological regulator [30]. The binding constants of a large number of potential inhibitors have been determined by Johnson and co-workers, and many of the inhibitor-receptor complexes have high resolution X-ray structures [30–39]. For the present study, Prof. Johnson has made a set of 30 complexed structures available to us (Table 1).

The Methods section gives a brief overview of the different prediction methods used in this study. The results are presented and discussed in the Results section. The Conclusions section outlines the conclu-

Table 1. Structure and affinity data for the GP data set. The compounds are ranked from worst to best

															
A				B				C				D			
No.	Struct.	R	pK _i	No.	Struct.	R	pK _i								
1	A	α-CONHCH ₃	1.435	16	A	β-CONH-cyclopropyl	2.886								
2	A	β-SCH ₂ CONH ₂	1.676	17	A	β-NHCOCH ₂ NHCOCH ₃	3.004								
3	A	β-CH ₂ CONH-2,4-F ₂ -C ₆ H ₃	1.724	18	A	β-CONH ₂	3.357								
4	A	α-CONHC ₂ H ₅ OH	1.772	19	A	β-CONHNH ₂	3.398								
5	A	β-CH ₂ NH ₃ ⁺	1.775	20	A	α-CONH ₂	3.432								
6	A	β-O-(1-6)-D-glucose	1.788	21	A	β-NHCOCH ₂ NH ₂	3.432								
7	A	β-CH ₂ N ₃	1.818	22	A	β-CONHCH ₃	3.796								
8	A	α-CONH-4-OH-C ₆ H ₄	2.252	23	D	NH ₂	3.836								
9	A	β-CH ₂ OSO ₂ CH ₃	2.319	24	A	β-NHCONH ₂	3.854								
10	A	β-SCH ₂ CONHC ₆ H ₅	2.444	25	A	β-NHCOC ₃ H ₇	4.027								
11	A	α-CONHNH ₂	2.523	26	C	—	4.229								
12	A	β-CO ₂ CH ₃	2.553	27	A	β-NHCOCH ₂ Cl	4.347								
13	B	—	2.699	28	A	β-NHCOCH ₃	4.495								
14	A	β-CONHNHCH ₃	2.745	29	A	α-CONH ₂ , β-NHCO ₂ CH ₃	4.796								
15	A	α-OH	2.770	30	D	H	5.523								

sions. In the companion paper (to be submitted), the implementation of the multi-layer prediction system proposed in this study is described in detail, and it is applied to a number of novel GP ligands.

Methods

Most of the methods used in this study have been described elsewhere and therefore will not be discussed in detail. However, to aid in following the present approach a brief outline of each method is given. The first five methods, (a)–(e), are ligand-based QSAR techniques. The next method, (f), utilizes the empirical scoring function of the LUDI program. The last method, (g), is a structure-based prediction method where the parameters are tuned for the system under study. All calculations were done on a Silicon Graphics Octane workstation with dual 175 MHz R10000 processors.

In this study two variables are reported to give a measure of the statistical fit of the models. The Pearson correlation coefficient, r^2 , corresponds to the correlation between the *calculated* activities and the

observed biological activities of the 30 training compounds. The predictivity of the model is indicated by a q^2 value, which is derived from the *predicted* activities based on a leave-one-out (LOO) cross-validation procedure [18]:

$$q^2 = 1 - \frac{\sum_{i=1}^N (y_{i,\text{observed}} - y_{i,\text{predicted}})^2}{\sum_{i=1}^N (y_{i,\text{observed}} - \bar{y}_{i,\text{observed}})^2}$$

(a) HQSAR

Hologram QSAR (HQSAR) is an emerging QSAR technique that requires no explicit 3D information for the ligands. First, the molecule is hashed to a molecular fingerprint which encodes the frequency of occurrence of various molecular fragment types using a pre-defined set of rules. To construct a molecular 'hologram', this molecular fingerprint is cut into strings at a fixed interval as specified by a hologram length (HL) parameter. The strings are then aligned

and the sum of each column constitutes the individual component of the molecular hologram of a particular length. Figure 1 shows the transformation of the chemical representation of a molecule into its corresponding molecular hologram, which is enumerated by a string of integers. This numerical representation of molecules is exploited by a subsequent correlation analysis; typically a PLS QSAR model is constructed. The HQSAR module of SYBYL (version 6.3, Tripos Inc.) is used for the calculation. The hashed fingerprints encode the presence of all molecular fragment types containing between four and seven atoms. We include hydrogen atoms and information on chiral centers in the generation of the molecular holograms. The optimal HQSAR model is derived from screening through the 12 default HL values, which are a set of 12 prime numbers ranging from 53 to 401 [8].

(b) RSM

The receptor surface model (RSM) is a hypothetical molecular surface that characterizes a putative receptor environment based on the spatial arrangement of the ligands [19, 20, 40]. A few active compounds are used as templates for which a surface is constructed to represent their aggregate molecular shape. Physical attributes (i.e., charge, electrostatic potential, hydrogen bond propensity and hydrophobicity) that are associated with each surface point are parameterized using the same set of molecules and property values complementary to the neighboring ligand atoms are assigned. For example, if the receptor surface is based on a single molecule, each surface point is assigned a partial charge that is equal but opposite in sign to the charge of the nearest atom in the molecule. If multiple molecular templates are used in the surface construction, the assigned charge is complementary to the averaged partial atomic charges of the closest atoms of the set. To generate a number of descriptors for use in QSAR applications, a molecule is placed inside the surface and is energy minimized within the context of a receptor model by use of a simplified force field; for the details of the energetic calculations and the implementation of the *Clean* force field that is used in the computation, the readers are referred to the published paper [19]. Six signed energy measures are generated: E_{inter} , $E_{\text{inter,vdW}}$ and $E_{\text{inter,elec}}$ are the total, the van der Waals, and the electrostatic interaction energy between the ligand and the receptor surface, E_{inside} is the internal energy of the energy minimized ligand inside the receptor surface, E_{relax} is the corresponding internal

energy of the ligand when the receptor surface is removed, and E_{strain} , which is a measure of strain energy induced by the surface, is defined as the difference between E_{inside} and E_{relax} . The standard RSM module from Cerius² (version 3.0, Molecular Simulations Inc.) is used for this study. For the set of known active compounds, we extracted the experimental binding conformation of the ligands from the structures of the protein-ligand complex for the surface calculations. The surface fit parameter specifies how tight (or loose) the generated receptor model is relative to the vdW surface of the molecules used to create the surface. For this application we decided to set a larger value (1.0 Å) for this parameter so as to allow the resulting receptor surface to accommodate a more diverse set of compounds for later predictions. Similar to the first published QSAR application of RSM [20], we include both linear and spline functions to combine the descriptors, and use the genetic function approximation (GFA) [41] to build good predictors.

(c) CoMFA

Comparative molecular field analysis (CoMFA)[12] has been an important tool for drug design since its introduction 10 years ago, and is often regarded as the current standard in 3D QSAR modeling. In a typical CoMFA calculation, the ligands are mutually aligned and are placed in a common 3D lattice. The steric and electrostatic fields of each ligand are sampled at the various grid points of the lattice. The resulting field matrix is analyzed by the PLS method, from which a QSAR model can be constructed. In this study, the CoMFA module in SYBYL (version 6.3, Tripos Inc.) was used. The conformation and the alignment of the ligands were extracted from the crystal structures of the bound complexes. MOPAC AM1 Mulliken charges, which were used in the electrostatic calculations, were computed for each ligand using a single-point energy calculation. The field calculation sets up a $22 \times 22 \times 24$ Å³ lattice with a 1 Å grid spacing for both steric and electrostatic interactions. The default truncation cutoff of 30 kcal/mol was applied to both fields.

(d) SMGNN

SMGNN is a QSAR method that makes use of the concept of molecular similarity [42, 43]. The initial step of the calculation, the generation of 3D electrostatic and steric fields of the ligands, is similar to that of CoMFA. Molecular similarity indices between

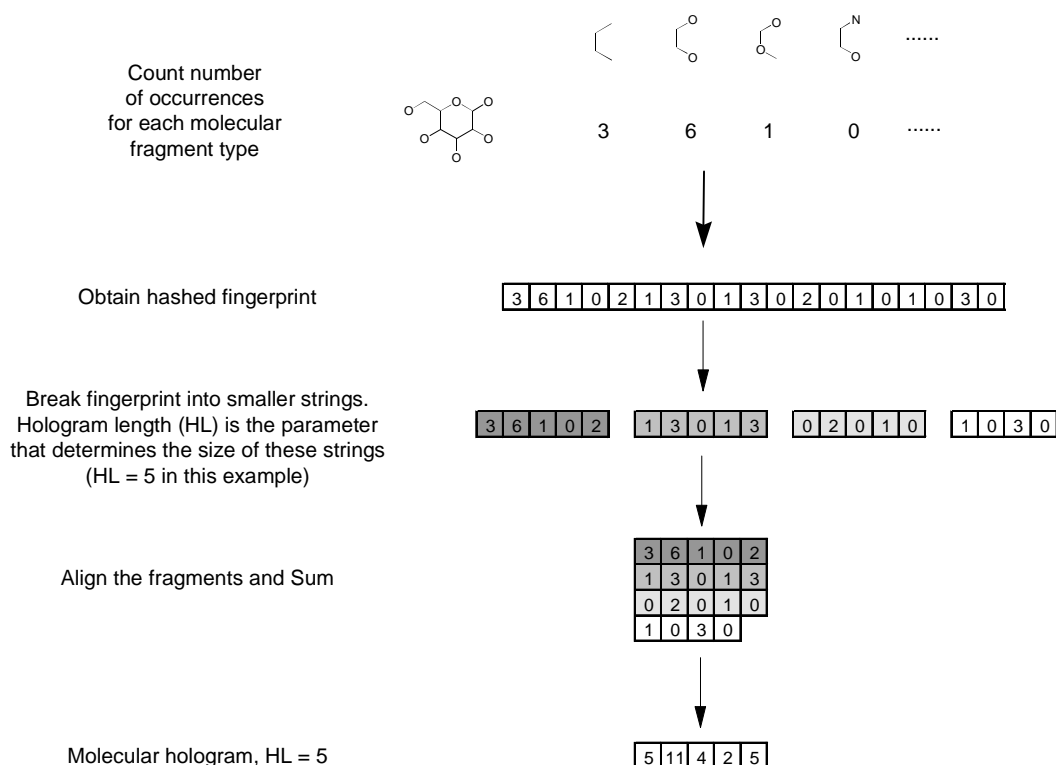


Figure 1. Transformation of a chemical structure to its characteristic molecular hologram. In this example, the glucose molecule is converted to a molecular hologram with a hologram length of 5. Subsequently, each of the values inside the 5 boxes serves as an independent descriptor for the later PLS analysis. For illustration purposes, a short hashed fingerprint is shown in this example. In a typical hologram generation, the hashed fingerprint contains information on all ($\sim 10^3$) molecular fragment types ranging from 4 to 7 atoms in length. In the standard implementation of HQSAR in SYBYL, 12 different hologram lengths ranging from 53 to 401 are tried.

two molecules are then evaluated using the Carbó [44] or the Hodgkin [45] formulae. By taking all pairwise comparisons between the molecules, the raw field matrix is transformed into a very compact molecular similarity matrix. Statistical techniques, such as PLS or GNN, can be applied to the similarity matrix. In this study, we took advantage of the X-ray alignment and the bioactive conformations of the ligands for the similarity calculations. The electrostatic and steric similarity matrices were generated by the Search_Compare module of the INSIGHT program [46], using the CFF91 force field [47, 48]. The standard GNN simulation protocol, as described in an earlier study [17], was applied to the combined matrix.

(e) C2GNN

Twenty-one 2D- and 3D-descriptors are selected from the standard QSAR database in the Cerius² program [49]. They represent a diverse set of descriptors that can be computed very rapidly. They include five electronic, seven spatial, four structural and five thermo-

dynamics descriptors (Table 2). The standard GNN algorithm [15, 16] was used for descriptor selection and activity correlation.

(f) LUDI

LUDI is a knowledge-based ligand design system with a set of geometric rules derived from a statistical analysis of a series of small molecule crystal structures [21, 22]. Its empirical scoring function, which was calibrated using the binding constants from a database of 45 ligand-protein complexes, can be used, in principle, to estimate the binding affinity of any complex. The expression for the LUDI energy function is given by [50]:

$$\text{LUDI score} = k \times \Delta G$$

Table 2. The initial pool of descriptors that are used to correlate biological activity in C2GNN calculations

Descriptor key	Descriptor type	Property
Apol	Electronic	Sum of atomic polarizabilities
Sr	Electronic	Superdelocalizability calculated by MOPAC
LUMO_MOPAC	Electronic	LUMO calculated by MOPAC
DIPOLE_MOPAC	Electronic	Dipole moment calculated by MOPAC
HOMO_MOPAC	Electronic	HOMO calculated by MOPAC
RoG	Spatial	Radius of gyration
Area	Spatial	Surface area
PMI	Spatial	Principal moment of inertia (magnitude)
PMI_X	Spatial	Principal moment of inertia (x-direction)
PMI_Y	Spatial	Principal moment of inertia (y-direction)
PMI_Z	Spatial	Principal moment of inertia (z-direction)
Vm	Spatial	Volume
MW	Structural	Molecular weight
RotBond	Structural	Number of rotatable bonds
Hbond_Acceptor	Structural	Number of hydrogen bonds acceptor
Hbond_Donor	Structural	Number of hydrogen bonds donors
AlogP	Thermodynamic	logP
Fh2o	Thermodynamic	Desolvation free energy for water
Foct	Thermodynamic	Desolvation free energy for 1-octanol
HF_MOPAC	Thermodynamic	Heat of formation calculated by MOPAC
MR	Thermodynamic	Molar refractivity

$$\begin{aligned}\Delta G = & \Delta G_0 + \Delta G_{hb} \sum_{h-bonds} f_1(\Delta R) f_2(\Delta \alpha) \\ & + \Delta G_{ion} \sum_{ionic} f_1(\Delta R) f_2(\Delta \alpha) \\ & + \Delta G_{lipo} A_{lipo} + \Delta G_{rot} N_{rot}\end{aligned}$$

where k takes the value of -73.33 mol/kcal and the projected potency of an inhibitor with a LUDI score of $S \times 100$ corresponds to a K_i value of 10^{-S} M at 300 K. ΔG and the other scalar parameters (ΔG_x) are in the unit of kcal/mol. ΔG_0 (set equal to 1.3 kcal/mol) represents the contribution to the binding energy from non-specific interactions with the receptor (e.g. loss of translational and rotational entropy of the ligand). ΔG_{hb} (-1.1 kcal/mol) and ΔG_{ion} (-2.0 kcal/mol) represent the contributions from an ideal hydrogen bond and an unperturbed ionic interaction, respectively. The functions $f_1(\Delta R)$ and $f_2(\Delta \alpha)$ scale the strength of these interactions according to the deviation from its ideal hydrogen bond length or angle. The ΔG_{lipo} (-0.040 kcal/mol) term represents the lipophilic contribution; it is proportional to the lipophilic contact surface, A_{lipo} , between the receptor and the ligand. The ΔG_{rot} (0.33 kcal/mol) term

represents the contribution due to the freezing of internal degrees of freedom in the ligand, and N_{rot} is the number of rotatable bonds.

Thus, the scoring function contains representative components describing three important factors of ligand binding; i.e., the ability of a given ligand to form hydrogen bonds in favorable orientations, the complementarity of lipophilic contacts with the receptor and the change in entropy upon binding. The relatively simple functional form allows for rapid assessment of the various terms and this efficiency is an important factor for performing searches on a large chemical database. In the present study, the LUDI score of the 30 GP inhibitors were calculated using the crystal structures.

(g) SBEP (Structure-based binding energy predictor)

A large number of approaches have been suggested for the estimation of binding free energies of compounds in binding sites of known structure [28]. In most cases some fitting parameters have been introduced to apply them to particular sets of ligands and protein pairs. LUDI is an approach of this type. Here we used a

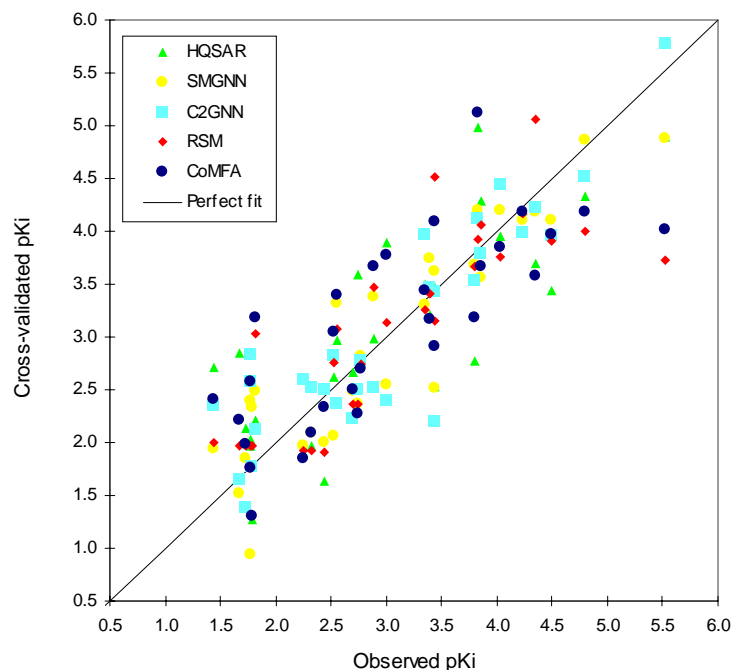


Figure 2. Plot of the cross-validated pKi values against the observed values for the five QSAR-based methods. (a) HQSAR predictions as green triangles; (b) RSM as red diamonds; (c) CoMFA as dark blue circles; (d) SMGNN as yellow circles; and (e) C2GNN as light blue squares. The diagonal line represents a perfect correlation between the observed and the predicted activity.

set of six quantities to generate a scoring function. The first two terms characterize the strength of the interaction between ligand and receptor in vacuum. They are the electrostatic and van der Waals interaction energies (Int_{elec} and Int_{vdW}), calculated using an atomistic potential energy function. The calculations were performed using the Discover program [51] with the CFF91 force field [47, 48]. The next two terms describe the solvent effect on the electrostatic and non-polar contributions to binding. The change in electrostatic energy (ΔG_{elec}) is calculated by the Delphi program using the parameter settings reported in a recent study [52]: an interior dielectric constant of 2 was used for the solute with a probe radius of 1.4 Å for water to define its molecular surface; a bulk dielectric of 80 and zero ionic strength was used for the solvent. The focusing method was used. The potential values derived from a preliminary run using a large, coarse grid (20 Å border space with a 2 Å resolution) provide the boundary values of the focused grid (1 Å border space with a 0.5 Å resolution) in the final calculation. The default vdW radius and charge templates of the Delphi program (version 95.0, Molecular Simulations Inc.) were used. The non-polar contribution to binding is estimated by the change in solvent-accessible

surface area (ΔSSA). This parameter was calculated by the Connolly surface generation program available in Cerius². A probe radius of 1.4 Å for water and a dot density of 10 per Å² were used to define the molecular surface. The final two terms, $T\Delta S_{\text{rec}}$ and N_{rot} , are related to the entropy loss of the receptor and the ligand upon binding. The loss of configuration entropy of the protein is estimated using the empirical scale reported by Pickett and Sternberg [53]. The main-chains or the side-chains of the protein were assumed to be immobile if they are within 3 Å of an atom of the ligand. The total loss of entropy is the sum of the contributions from all immobilized mainchain (−2.0 kcal/mol per residue) [52] and side-chain (the values are listed in Table 5 of [53]) units. The ligand entropy term was estimated by the number of rotatable bonds, as calculated by the LUDI program. The energy minimized structures of the complex were used to determine the two interaction energy terms (Int_{elec} and Int_{vdW}); the remaining descriptors (ΔG_{elec} , ΔSSA , $T\Delta S_{\text{rec}}$ and N_{rot}) were obtained using the original crystal structures coordinates.

Results

(i) Results from individual prediction models

We have examined the seven different approaches described in the Methods section for activity predictions of the 30 glycogen phosphorylase inhibitors. The first five methods consider only the 2D or 3D information of the ligand, except that the relative orientations required for all of these (except HQSAR) are based on the crystal coordinates, and the final two use the ligand-receptor complexed structures.

(a) HQSAR

Table 3 shows a summary of the results of the HQSAR calculation. The optimal HQSAR method is generated using a hologram length of 97. This PLS model contains four components, where each component consists of some linear combinations of the 97 descriptors encoded in the hologram (see Figure 1). It yields an r^2 value of 0.878 for the fitting of training data (all 30 compounds in Table 1), and a q^2 value of 0.636 for cross-validation (Table 4). The individual cross-validated predictions, shown as green triangles, are plotted against the experimental values in Figure 2. With the exception of the least active compound, all of the predictions appear to be quite accurate.

One attractive feature of the HQSAR program is that it does not require molecular superposition and that the descriptors can be generated very rapidly. This approach provides an efficient means for screening a large number of compounds from chemical databases, or products from a combinatorial library. In the companion paper (to be submitted), we describe the result of the application of the HQSAR model for screening the National Cancer Institute (NCI) database.

(b) RSM

A pseudo receptor surface was parameterized using the six most active compounds as templates (Table 1: 25–30). The best equation obtained from the GFA is:

$$\text{pK}_i = 1.95 + 1.92 \times \langle -2.68 - E_{\text{inter}} \rangle - 5.37 \times \langle -0.66 - E_{\text{inter, vdW}} \rangle$$

where $\langle x \rangle$ is a spline function that returns the value of argument x if x is positive and zero otherwise. The first non-constant term predicts that binding affinity would generally increase with a stronger interaction energy (electrostatic and van der Waals) between the ligand

Table 3. Results of the HQSAR calculations. For a given hologram length, the q^2 value of the QSAR model with the optimal number of components is given. The 12 HL values listed are the default prime numbers employed by the HQSAR module in SYBYL. The optimal HQSAR model, shown in bold typeface, is a four-component PLS model based on the molecular hologram generated using a length of 97

Hologram length (HL)	q^2	No. of components
53	0.534	3
59	0.443	2
61	0.484	2
71	0.496	2
83	0.532	2
97	0.636	4
151	0.543	2
199	0.491	2
257	0.499	3
307	0.494	3
353	0.506	2
401	0.501	2

Table 4. Statistical data for the five QSAR-based models and the combined model taking the average prediction values from the five

Model	r^2	q^2	Rms error
HQSAR	0.88	0.64	0.63
RSM	0.75	0.70	0.57
CoMFA	0.93	0.60	0.66
SMGNN	0.95	0.82	0.44
C2GNN	0.90	0.80	0.47
Combined	0.96	0.86	0.40

and the receptor surface. The second term appears to be a correction term that penalizes ligands with large van der Waals contributions, i.e., large ligands. This QSAR equation gives quite good performance in both correlation and prediction, with an r^2 value of 0.750, and a q^2 value of 0.704 for the 30 GP inhibitors (Table 4); it is interesting that the r^2 and q^2 values are more similar than for the other methods. The cross-validated predictions are shown as red diamonds in Figure 2. The other regression equations emerged from the GFA search are also dominated by the same two descriptors; most of them in fact have the same functional form as the one above but with a slightly different set of scalar parameters.

The receptor surface is shown in Figure 3. The brown region shown in the lower figure depicts the area where hydrophobic contacts are expected. The blue/red region on the surface in the upper figure depicts the region where atoms with negative/positive charges are expected. These surfaces give qualitative agreement with the binding site environment of GP. This is not unexpected since the ligands are oriented based on their bioactive conformations in the structure of the complex. There is a general weakness in using this method for activity prediction of novel analogs: new ligand that extends beyond the surface boundary will be predicted to be poor candidates, i.e., the model is over-constrained and tends to give more false negatives [19, 20, 40].

(c) CoMFA

CoMFA was applied to the 30 ligands to model their inhibitory activity on GP. The final CoMFA model is a PLS regression model using four components that are linear combinations of field descriptors. The model yields an r^2 value of 0.933 and a q^2 value of 0.600 (Table 4). The cross-validated predictions are shown as dark blue circles in Figure 2. The fraction of steric and electrostatic field contributions is 36.5% and 63.5%, respectively. Figure 4 shows a map that highlights the regions where the most important molecular interactions between the ligands and the receptor are expected to take place (the molecule shown in the figure is the most active analog, **30**). Sterically favorable areas are shown in green, and unfavorable areas in yellow. Regions where the presence of positive charges are favorable are labeled in red, and negative charges in blue. An experienced molecular modeler can utilize such plots for introducing better substituents. In the companion paper (to be submitted), we describe an automated procedure to construct new ligands by the systematic attachment of new functional groups.

Since RSM and CoMFA are both, in essence, receptor site mapping tools, it is interesting to compare them. In RSM, the visualization of receptor is based on a smooth molecular surface whereas the CoMFA map highlights only the most distinguishing features and is therefore more ‘pharmacophoric’ in character. Specifically, the key regions are positioned by CoMFA near the variable parts of the GP inhibitors (α - and β -substituents of C1) but not in the conserved molecular regions. Overall, the qualitative information extracted by the two methods is in fact very similar (e.g. the sites of important regions for polar interactions or hydrophobic contacts). In term of quantitative perfor-

Table 5. The total and the individual components of the LUDI score for the 30 GP inhibitors. Together with the three components, the sum also includes a score of -95 for every compound to account for the non-specific ligand interactions (Methods section (f))

Cpd	Total	Hydrogen bond	Lipophilicity	Entropy
1	379	232	293	-51
2	654	573	278	-102
3	930	647	480	-102
4	480	446	231	-102
5	637	576	207	-51
6	668	545	320	-102
7	512	486	198	-77
8	505	307	344	-51
9	580	496	281	-102
10	795	550	442	-102
11	424	413	157	-51
12	620	514	252	-51
13	557	477	201	-26
14	558	464	240	-51
15	609	564	166	-26
16	603	443	332	-77
17	737	635	299	-102
18	616	602	160	-51
19	566	534	178	-51
20	440	403	183	-51
21	564	511	225	-77
22	722	607	261	-51
23	732	684	169	-26
24	506	474	178	-51
25	754	604	347	-102
26	762	661	222	-26
27	632	526	278	-77
28	688	556	278	-51
29	685	647	210	-77
30	593	551	163	-26

mance, the CoMFA model appears to be more flexible in learning specific training patterns (i.e., a very high r^2 value), while its predictivity is surprisingly low (a relatively low q^2 value). Improvement in predictivity ($q^2 = 0.65$, unpublished data) can be obtained by screening out noisy input using a ‘region focusing’ option in Advanced CoMFA. However, we later encounter some inconsistency in the predictions based on the focused model for database screening and it was therefore not considered further.

It is also interesting to compare the current CoMFA result with an application of the related GRID program [11] to a larger set of 51 GP inhibitors [37]. It

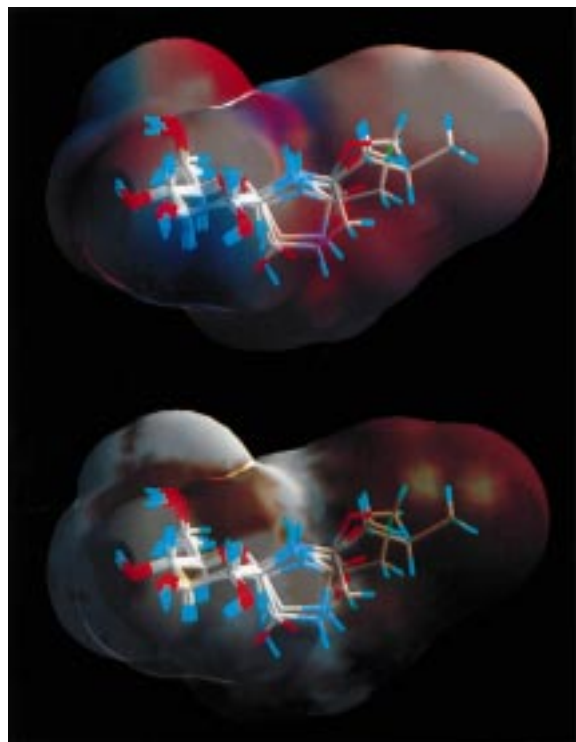


Figure 3. The RSM model constructed using the six most active inhibitors, shown inside the molecular surface. The blue/red region on the upper figure depicts the area where atoms with negative/positive charges are expected. The brown region on the lower figure depicts the area where hydrophobic contact is expected.

was found that without variable selection of the field descriptors, the GRID PLS model also yielded a low q^2 of 0.47. However, the reduction of variables from 7221 to 460 using the GOLPE routine led to a significant increase in predictivity ($q^2 = 0.76$). We expect that the use of variable selection, such as GOLPE, in conjunction with the CoMFA program can lead to a similar improvement.

(d) SMGNN

The optimal SMGNN model is a six-descriptor model, with an r^2 value of 0.950 and a q^2 value of 0.823 (Table 4). The individual predicted pK_i values, plotted as yellow circles, are shown in Figure 2. Five of the descriptors are electrostatic similarity measures using compounds **16**, **26**, **27**, **28** and **30** as templates; the remaining descriptor is a steric similarity index with respect to **29**. The use of similarity matrices that are calculated from INSIGHT has little effect on the model predictivity, relative to one previously reported ($q^2 = 0.82$) [18]. In the companion study, we describe

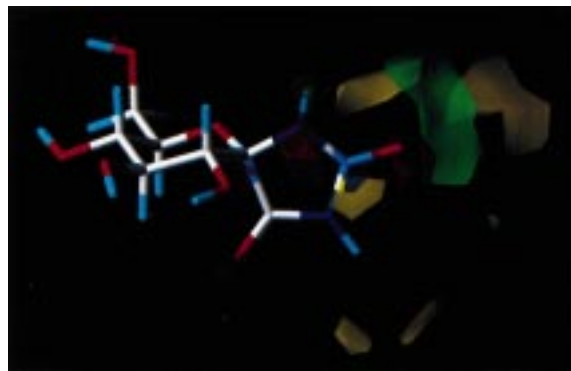


Figure 4. The CoMFA map highlighting the most important molecular interaction sites. Sterically favorable/unfavorable regions are shown in green/yellow. Regions where the presence of positive charges is preferable are labeled in red and negative charges in blue. Compound **30** is also shown in the figure.

the use of the SMGNN QSAR model as a 3D approach to screen the NCI database.

(e) C2GNN

The optimal GNN model from the pool of standard Cerius² descriptors is a five-descriptor model, with an r^2 value of 0.899 and a q^2 value of 0.797 (Table 4). The predicted pK_i is shown in Figure 2, plotted as light blue squares. The chosen descriptors are Sr, HOMO_MOPAC, PMI_X, PMI_Y and MW (see Table 2 for definitions).

(f) LUDI

The LUDI scores for the 30 GP inhibitors were evaluated using their respective ligand-receptor structures. A plot of the LUDI scores against the experimental pK_i values is shown in Figure 5. It is evident that the correlation between the two sets of values is weak ($r^2 = 0.04$). The origin of this discrepancy appears to be related to the formulation of the scoring function. In many systems used in the parameterization of the LUDI program, favorable lipophilic contacts between a ligand and a receptor play an important role in distinguishing the most potent inhibitors, so that the lipophilic factor becomes a significant portion of the overall score. Table 5 lists the total and the individual components of the LUDI scoring function. Most of the potent GP inhibitors currently under study are, in general, hydrophilic molecules and thus the lipophilic component of the LUDI score may have too great an importance for the present data. The predictions of several compounds (**3**, **10** and **30**) are particularly poor. The first two, both of which contain an aromatic

Table 6. Calculated values for the six descriptors that are considered for the SBEP calculations. The unit for the four energetic terms (Int_{elec} , Int_{vdW} , ΔG_{elec} and $T\Delta S_{\text{rec}}$) is kcal/mol, and ΔSSA is in \AA^2

	Int_{elec}	Int_{vdW}	ΔSSA	ΔG_{elec}	$T\Delta S_{\text{rec}}$	N_{rot}
1	-29.70	-11.34	-435.15	28.81	-30.24	2
2	-41.59	-7.80	-455.68	20.20	-31.12	4
3	-35.19	-17.28	-645.92	35.10	-35.09	4
4	-32.78	-15.14	-456.83	30.33	-35.12	4
5	-37.48	-2.82	-397.51	25.72	-27.93	2
6	-47.40	-15.36	-621.26	40.04	-35.39	4
7	-30.63	-8.63	-391.48	15.40	-27.18	3
8	-30.60	-16.12	-551.77	44.87	-29.85	2
9	-32.03	-15.22	-486.67	21.57	-32.14	4
10	-36.75	-17.85	-651.94	33.91	-35.12	4
11	-41.08	-7.58	-426.61	21.87	-25.55	2
12	-31.21	-11.87	-418.33	16.13	-29.12	2
13	-33.29	-6.67	-348.87	17.02	-26.30	1
14	-34.92	-10.77	-407.78	23.31	-31.12	2
15	-35.67	-5.84	-333.74	12.59	-26.30	1
16	-37.53	-12.29	-424.40	16.91	-32.14	3
17	-40.90	-13.81	-444.32	22.55	-32.14	4
18	-38.64	-5.28	-360.64	15.35	-25.93	2
19	-36.50	-7.54	-374.23	16.11	-25.93	2
20	-36.44	-10.72	-370.37	23.58	-25.55	2
21	-37.99	-10.24	-415.74	16.10	-31.18	3
22	-36.87	-8.46	-393.05	15.36	-27.18	2
23	-41.50	-8.88	-452.78	14.99	-29.18	1
24	-37.17	-9.98	-391.37	18.95	-26.30	2
25	-36.88	-11.46	-446.12	14.13	-29.18	4
26	-50.83	-10.72	-444.55	17.23	-29.18	1
27	-38.48	-10.37	-403.71	18.10	-31.18	3
28	-37.05	-11.35	-397.50	15.62	-29.18	2
29	-43.54	-13.52	-455.85	22.25	-29.18	3
30	-40.83	-10.12	-411.06	19.14	-25.93	1

Table 7. Correlation (r) matrix for the six SBEP descriptors listed in Table 4

	pK_i	Int_{elec}	Int_{vdW}	ΔSSA	ΔG_{elec}	$T\Delta S_{\text{rec}}$
Int_{elec}	-0.42					
Int_{vdW}	0.13	-0.07				
ΔSSA	0.35	0.13	0.79			
ΔG_{elec}	-0.51	0.09	-0.64	-0.78		
$T\Delta S_{\text{rec}}$	0.41	0.04	0.76	0.77	-0.58	
N_{rot}	-0.38	0.03	-0.60	-0.59	0.38	-0.75

ring, have their binding affinity substantially overestimated, whereas the last one, a small hydrophilic

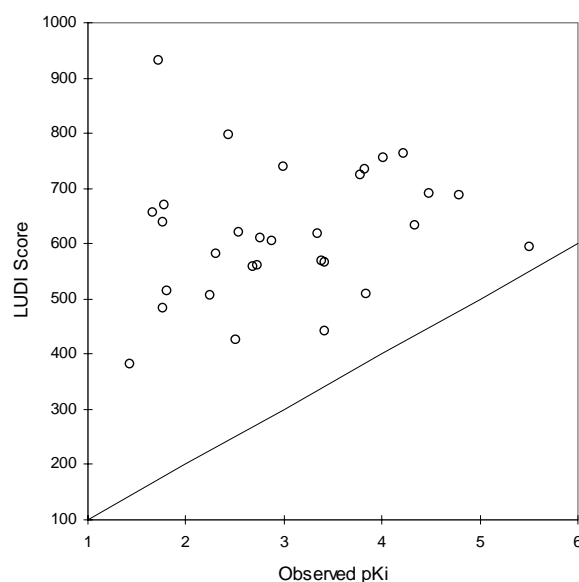


Figure 5. Plot of the calculated LUDI scores against the observed pK_i values for the 30 GP inhibitors. The diagonal line indicates perfect correlation between the observed value and the projected activities based on LUDI score. All compounds have a score over 300.

compound, has a much lower score than expected for the experimental value.

An attempt was made to re-calibrate the scalar coefficients for the various LUDI components and establish a scoring function that is specific to the GP system. Using the scores as listed in Table 5, the following equation is formulated:

$$\begin{aligned} \text{pK}_i = & 1.89 + 0.00456 \text{LUDI}_{\text{hb}} \\ & - 0.00156 \text{LUDI}_{\text{lipo}} \\ & + 0.0133 \text{LUDI}_{\text{entropy}} \end{aligned}$$

The r^2 of the reformulated equation has increased significantly from 0.04 to 0.35, though cross-validation still gives a low q^2 value (0.16). The coefficients of the individual terms further suggest that the hydrogen bond and the entropy components increase ligand-receptor affinity whereas an overly lipophilic ligand tends to weaken binding for the current set of compounds.

We are grateful to a reviewer who brought our attention to a study that reported an application of a LUDI-like scoring function to a set of 82 ligand-receptor complexes [54]. The prediction accuracy of this empirical function was, in general, less satisfactory on sugar-binding proteins relative to the other classes of proteins in the training set, and the data shown are in accord with this result.

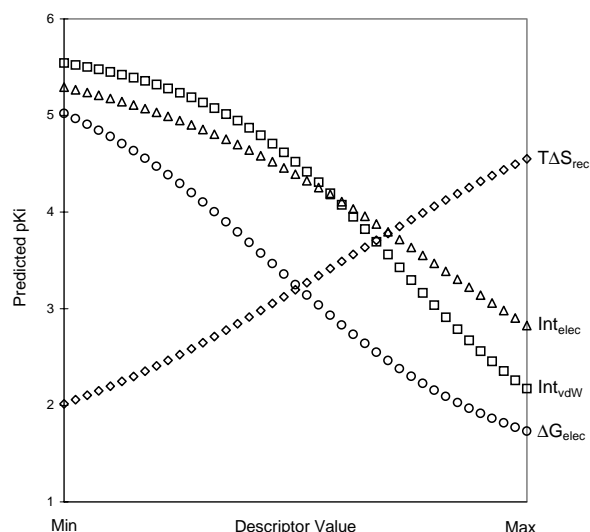


Figure 6. Functional dependence plots for the four SBEP descriptors. Their minimum and maximum values are Int_{elec} (−50.83 to −29.70), Int_{vdW} (−17.85 to −2.82), ΔG_{elec} (12.59 to 44.87) and $T\Delta S_{\text{rec}}$ (−35.39 to −25.55). The scale is linear between min and max for each descriptor.

(g) SBEP

The six descriptors were evaluated for each of the 30 ligand-receptor complexes, and their values are shown in Table 6. A correlation matrix of the six descriptors, together with the pK_i values collected from each one alone, is shown in Table 7. Some descriptors are strongly correlated. For example, Int_{vdW} and ΔSSA have a positive correlation because both measures are related to the formation of complementary steric overlaps between ligand and receptor; $T\Delta S_{\text{rec}}$ and N_{rot} are negatively correlated since a larger decrease in entropy of the receptor can be associated with a larger ligand which, in general, has more rotatable bonds. To investigate the optimal choice of descriptors, an exhaustive enumeration of all $2^6 - 1$ multiple linear regressions, using different combinations of descriptors, was made. The result is summarized in Table 8. The following four-descriptor equation yields the optimal q^2 value of 0.51 and an r^2 value of 0.68:

$$\text{pK}_i = 5.557 - 0.103 \text{Int}_{\text{elec}} - 0.238 \text{Int}_{\text{vdW}} - 0.076 \Delta G_{\text{elec}} + 0.246 T\Delta S_{\text{rec}}$$

The equation predicts that stronger intermolecular interactions (Int_{elec} and Int_{vdW}) between ligand and receptor will increase inhibitory activity. Since the two interaction energy terms are always negative, they contribute positively to the binding. On the other hand,

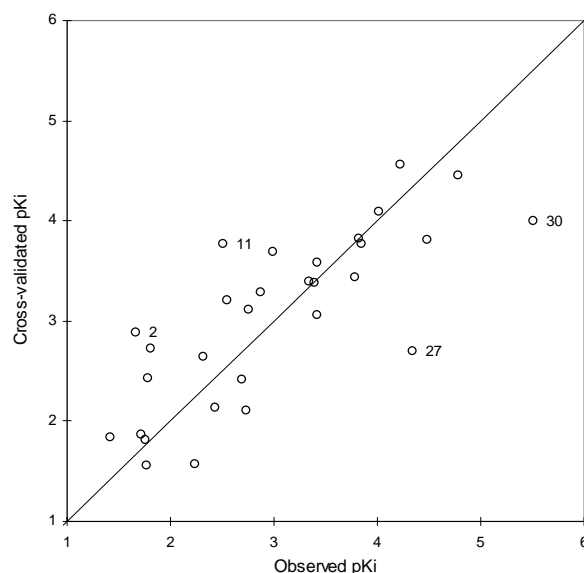


Figure 7. Plot of the cross-validated pK_i values against the observed values using the SBEP method. The four compounds with prediction errors greater than 1.0 in pK_i are labeled in the plot. The diagonal line shows a perfect correlation between the observed and the cross-validated pK_i value.

the ΔG_{elec} (always positive) and the $T\Delta S_{\text{rec}}$ (always negative) terms, as expected, weaken the binding of the ligand to a different extent. Thus, the signs for the four coefficients are consistent with the physical attributes of the binding.

A similar exhaustive enumeration was performed using a computational neural network to obtain activity correlation, and their q^2 values were determined. Interestingly, the neural network in the optimal model (Table 8) utilized the same four descriptors listed above. A functional dependence plot [55] for the four descriptors in question is shown in Figure 6. Qualitatively the slope for each descriptor is consistent with the signs of the coefficients in the regression equation, and as already discussed, describes the correct physical aspects of ligand binding. It appears that all four thermodynamics factors make important contributions to the binding process, as suggested by a similar range of the projected pK_i values when each descriptor is screened from its minimum to maximum value in the dependence plot.

Compared to the regression statistics, the optimal neural network model yields a small gain in r^2 (0.72) and a more significant increase in q^2 (0.61). Figure 7 shows the individual cross-validated pK_i values of the 30 ligands plotted against the experimental values. Most of the predictions are very accurate; 19 of them

Table 8. Statistical data for the optimal regression and neural network models using a specific number of SBEP descriptors. This analysis indicates that the four-descriptor regression or neural network models have the highest predictivity amongst all of the $2^6 - 1$ combinations being considered

No. of descriptors	Best regression model			Best neural network model		
	Choice of descriptors	r^2	q^2	Choice of descriptors	r^2	q^2
1	ΔG_{elec}	0.260	0.169	ΔG_{elec}	0.298	0.174
2	$\Delta G_{\text{elec}} \text{Int}_{\text{elec}}$	0.406	0.212	$\Delta G_{\text{elec}} \text{Int}_{\text{elec}}$ 0.465	0.278	
3	$\Delta G_{\text{elec}} \text{Int}_{\text{vdW}} \text{T}\Delta S_{\text{rec}}$	0.470	0.336	$\Delta G_{\text{elec}} \text{Int}_{\text{elec}} \text{Int}_{\text{vdW}}$	0.559	0.380
4	$\Delta G_{\text{elec}} \text{Int}_{\text{elec}} \text{Int}_{\text{vdW}} \text{T}\Delta S_{\text{rec}}$	0.676	0.511	$\Delta G_{\text{elec}} \text{Int}_{\text{elec}} \text{Int}_{\text{vdW}} \text{T}\Delta S_{\text{rec}}$	0.721	0.605
5	All but N_{rot}	0.693	0.507	All but N_{rot}	0.735	0.574
6	All six	0.712	0.480	All six	0.761	0.549

are within 0.5 log unit of the observed pK_i values. Only four compounds have prediction errors over 1.0: they are **2**, **11**, **27** and **30** (labeled in the figure). The overestimation in activity of the two relatively weak compounds (**2** and **11**) is mainly due to their very favorable electrostatic interactions with the receptor. On the other hand, **27**, a reasonably potent compound, has only average values for all four descriptors and is therefore predicted as moderately active. It is not unexpected that cross-validation results in a relatively large under-prediction for the most potent compound, **30**. Given that its pK_i value is more than 0.7 log unit higher than the next most potent analog, the SBEP prediction of this compound is satisfactory.

(ii) Results from a jury QSAR model

In the previous section we have presented the individual results of the seven activity prediction methods. Since different methods involve different sets of assumptions and limitations, some methods are expected to provide more accurate prediction for certain types of molecules than others. Thus, it has been suggested that combining a number of prediction models can improve the accuracy of the predictions [56]. It seems reasonable, if we can assume that the prediction errors amongst different methods are normally distributed, that taking the averaged outcome will give a better result. This phenomenon has been demonstrated in several QSAR studies [15, 16, 41] as well as in the prediction of protein secondary structure [57]. In this section, we investigate the effect of averaging results from different prediction methods. The five ligand-based QSAR models, (a)–(e), which have demonstrated various degrees of success in activity prediction, are combined in the current jury system.

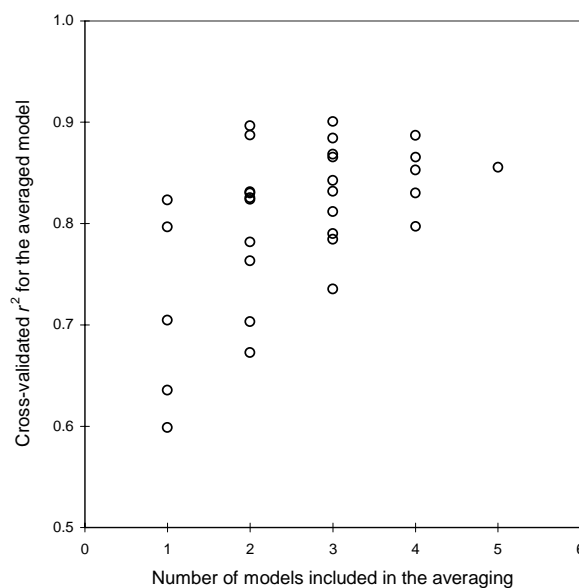


Figure 8. Cross-validated r^2 of the $2^5 - 1$ N -averaged jury models. Jury models are made by averaging the predictions from different combinations of QSAR models. Results are sorted into sets according to the number of models used in the average.

All possible combinations ($2^5 - 1$) of QSAR predictions were averaged and their q^2 values were evaluated. Figure 8 shows a plot of the q^2 value as a function of the number of models involved in the averaging, and the average and standard deviation in q^2 are listed in Table 9 (see also Table 4). The result shows that there is a general improvement in predictivity with an increase in the number of QSAR models. This further suggests that instead of employing the single method that yields the best q^2 value to predict activity, it will be useful to combine the various methods to create a fast scoring system that is more reliable.

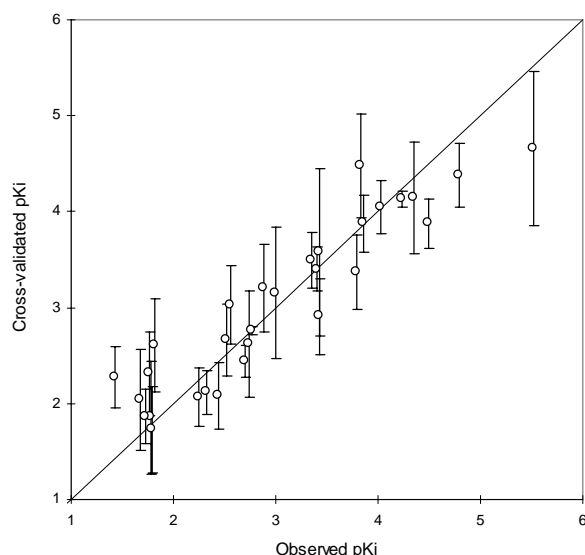


Figure 9. Plot of cross-validated pK_i values versus observed values. The predictions are made by averaging the results of the five QSAR models; the error bar corresponds to one standard deviation from the average value. The diagonal line depicts perfect correlation between the two sets of values.

Table 9. Statistical data for the N -averaged jury QSAR models. N_{set} is the number of distinct combinations in a set involving a specific number of averaged models. The average and the standard deviation of q^2 for each set are shown

	Number of models included in the averaging				
	1	2	3	4	5
N_{set}	5	10	10	5	1
q^2 average	0.71	0.80	0.83	0.85	0.86
q^2 std dev.	0.10	0.07	0.05	0.03	—

Conclusions

We have applied seven different approaches for activity predictions to a set of 30 GP inhibitors with known crystal structures and K_i values. For the present set of compounds, the LUDI score is not an accurate means for ranking though it may provide a crude estimate of absolute activity. Five classical QSAR methods were also applied to the set of compounds. One of them, HQSAR, is 2D in nature though it contains information about the chirality of the compounds. The other four QSAR methods make explicit use of 3D descriptors that are derived from molecular fields, molecular surfaces or other spatial properties. The q^2 values of these models range from 0.60 to 0.82. We have shown that a simple average using all five QSAR predictions

Table 10. Approximate CPU time (on a 175 MHz R10000 Silicon Graphics Octane workstation) required for the generation of each prediction model, and the processing time to make an activity prediction for each new compound

Model	Model generation time	Prediction time/compound
HQSAR	1 min	< 1 s
CoMFA	2 min	1 s
RSM	2 min	1 s
SMGNN	30 min	2 s
C2GNN	30 min	2 s
LUDI	2 min	3 s
SBEP	2–3 days	2 h

leads to a more reliable prediction. The jury prediction model has a q^2 value of 0.86. Its rms prediction error for pK_i is 0.40, which is less than one-tenth of the observed range. Finally, a neural network predictor was constructed that employs thermodynamics variables derived from atomistic calculations on the basis of the 3D structure of the bound complexes. This SBEP system yielded a q^2 value of 0.61. The approximate CPU usage for model generation and prediction of new analogs for each of the methods is given in Table 10.

An interesting result of this study is that the five ligand-based QSAR models perform better than the two structure-based methods which try to capture the essence of the molecular interactions. One point is that the compounds are closely related and have essentially the same orientation in the binding site (i.e., the glucose moiety has nearly the same position in all compounds) so that they are ideal for the QSAR methods used here. The LUDI scoring function, on the other hand, is calibrated on a wide set of protein-ligand complexes and it is therefore not surprising that it could not provide accurate predictions on the *relative* binding affinity of this set of congeneric ligands. LUDI does particularly badly here, probably because its significant lipophilic component fails to differentiate the present set of hydrophilic compounds. As to the SBEP approach, it is a very oversimplified model for the physical aspects of the binding with a small number of adjustable descriptors. The latter is important since it also reduces the chance of spurious correlation relative to the QSAR models. There is much effort now to develop more sophisticated structure-based methods [27, 29, 54, 58–60]. It is our conviction that the next generation of activity predictors will make better

use of structural information so that the generality and the predictivity of the physically based model will be vastly improved.

Given the results of the different methods tested in this study, we suggest that it is appropriate here to consider a multi-layer system for activity predictions. The first layer is a preliminary screening layer which evaluates a large number of compounds at minimal computational cost. For this part of the calculation, accurate numerical predictions of activity are not essential; all that one needs to do is to exclude compounds that have a high likelihood of being poor ligands. The LUDI program, which is commonly used for coarse screening of putative ligands with very diverse chemistries in database searches, provides the initial filter by discarding candidates that are below a threshold score. The second layer is an activity prediction layer, whose principal function is to provide a more accurate measure of activity which is subsequently used as a criterion for compound selection. A jury decision derived from the averaged predicted values of the five QSAR-based methods forms the basis of this prediction layer. Only a few of the most promising candidates can reach the final validation layer, where a computationally more demanding structure-based prediction procedure, such as SBEP, may be applied to make an additional activity assessment that is complementary to the ligand-based QSAR methods. The grouping of the various prediction methods is made based on the individual results and their associated computational cost. In the companion paper (to be submitted), the implementation of the proposed prediction system is presented. We will apply this new system to test novel GP ligands that are found through database searches and de novo design methods.

Acknowledgements

We are very grateful to Dr K. Watson and Prof. L. Johnson for providing the glycogen phosphorylase data set and for assistance with their use. We thank Molecular Simulations Inc. and Tripos Inc. for the software support. This work is supported in part by a grant from the National Science Foundation and a gift from Eli Lilly and Company.

References

1. Tokarski, J.S. and Hopfinger, A.J., *J. Chem. Inf. Comput. Sci.*, 37 (1997) 792.

2. Hansch, C. and Leo, A., *Substituent Constants for Correlation Analysis in Chemistry and Biology*, John Wiley & Sons, Inc., New York, NY, 1979.
3. Randic, M., *J. Am. Chem. Soc.*, 97 (1975) 6609.
4. Hall, L.H. and Kier, L.B., *J. Chem. Inf. Comput. Sci.*, 35 (1995) 1039.
5. Hurst, T. and Heritage, T., *Abstract of Papers of the Am. Chem. Soc.*, 213 (1997) 19.
6. Hurst, T., Heritage, T.W. and Clark, R.D., *Abstract of Papers of the Am. Chem. Soc.*, 215 (1998) 38.
7. Tong, W.D., Perkins, R., Sheehan, D.M., Welsh, W.J., Lewis, D.R. and Goddette, D.W., *Abstract of Papers of the Am. Chem. Soc.*, 214 (1997) 81.
8. HQSAR, Version 1.0, Tripos, Inc., St. Louis, MO.
9. Winkler, D.A. and Burden, F.R., *Quant. Struct.-Act. Relat.*, 17 (1998) 224.
10. Silverman, B.D. and Platt, D.E., *J. Med. Chem.*, 39 (1996) 2129.
11. Goodford, P.J., *J. Med. Chem.*, 28 (1985) 849.
12. Cramer III, R.D., Patterson, D.E. and Bunce, J.D., *J. Am. Chem. Soc.*, 110 (1988) 5959.
13. Good, A.C., So, S.-S. and Richards, W.G., *J. Med. Chem.*, 36 (1993) 433.
14. Good, A.C., Peterson, S.J. and Richards, W.G., *J. Med. Chem.*, 36 (1993) 2929.
15. So, S.-S. and Karplus, M., *J. Med. Chem.*, 39 (1996) 1521.
16. So, S.-S. and Karplus, M., *J. Med. Chem.*, 39 (1996) 5246.
17. So, S.-S. and Karplus, M., *J. Med. Chem.*, 40 (1997) 4347.
18. So, S.-S. and Karplus, M., *J. Med. Chem.*, 40 (1997) 4360.
19. Hahn, M., *J. Med. Chem.*, 38 (1995) 2080.
20. Hahn, M. and Rogers, D., *J. Med. Chem.*, 38 (1995) 2091.
21. Böhm, H.-J., *J. Comput.-Aided Mol. Design*, 6 (1992) 593.
22. Böhm, H.-J., *J. Comput.-Aided Mol. Design*, 8 (1994) 243.
23. DeWitte, R.S. and Shakhnovich, E.I., *J. Am. Chem. Soc.*, 118 (1996) 11733.
24. DeWitte, R.S., Ishchenko, A.V. and Shakhnovich, E.I., *J. Am. Chem. Soc.*, 119 (1997) 4608.
25. Welch, W., Ruppert, J. and Jain, A.N., *Chem. Biol.*, 3 (1996) 449.
26. Holloway, M.K., Wai, J.M., Halgren, T.A., Fitzgerald, P.M., Vacca, J.P., Dorsey, B.D., Levin, R.B., Thompson, W.J., Chen, L.J. and deSolms, S.J., *J. Med. Chem.*, 38 (1995) 305.
27. Åqvist, J., Medina, C. and Samuelsson, J.E., *Protein Eng.*, 7 (1994) 385.
28. Ajay and Murcko, M.A., *J. Med. Chem.*, 38 (1995) 4953.
29. Gilson, M.K., Given, J.A., Bush, B.L. and McCammon, J.A., *Biophys. J.*, 72 (1997) 1047.
30. Martin, J.L., Johnson, L.N. and Withers, S.G., *Biochemistry*, 29 (1990) 10745.
31. Martin, J.L., Veluraja, K., Ross, K., Johnson, L.N., Fleet, G.W., Ramsden, N.G., Bruce, I., Orchard, M.G., Oikonomakos, N.G. and Papageorgiou, A.C., *Biochemistry*, 30 (1991) 10101.
32. Johnson, L.N., Snape, P., Martin, J.L., Acharya, K.R., Barford, D. and Oikonomakos, N.G., *J. Mol. Biol.*, 232 (1993) 253.
33. Oikonomakos, N.G., Kontou, M., Zographos, S.E., Tsitoura, H.S., Johnson, L.N., Watson, K.A., Mitchell, E.P., Fleet, G.W., Son, J.C., Bichard, C.J., Leonidas, D.D. and Acharya, K.R., *Eur. J. Drug Metab. Pharmacokinet.*, 19 (1994) 185.
34. Watson, K.A., Mitchell, E.P., Johnson, L.N., Son, J.C., Bichard, C.J., Orchard, M.G., Fleet, G.W., Oikonomakos, N.G., Leonidas, D.D. and Kontou, M., *Biochemistry*, 33 (1994) 5745.

35. Oikonomakos, N.G., Kontou, M., Zographos, S.E., Watson, K.A., Johnson, L.N., Bichard, C.J., Fleet, G.W. and Acharya, K.R., *Protein Sci.*, 4 (1995) 2469.
36. Oikonomakos, N.G., Zographos, S.E., Johnson, L.N., Papa-georgiou, A.C. and Acharya, K.R., *J. Mol. Biol.*, 254 (1995) 900.
37. Watson, K.A., Mitchell, E.P., Johnson, L.N., Cruciani, G., Son, J.C., Bichard, C.J.F., Fleet, G.W.J., Oikonomakos, N.G., Kontou, M. and Zographos, S.E., *Acta Crystallogr.*, D51 (1995) 458.
38. Bichard, C.J.F., Mitchell, E.P., Wormald, M.R., Watson, K.A., Johnson, L.N., Zographos, S.E., Koutra, D.D., Oikonomakos, N.G. and Fleet, G.W.J., *Tetrahedron Lett.*, 36 (1995) 2145.
39. Krülle, T.M., Watson, K.A., Gregoriou, M., Johnson, L.N., Crook, S., Watkin, D.J., Griffiths, R.C., Nash, R.J., Tsitsanou, K.E., Zographos, S.E., Oikonomakos, N.G. and Fleet, G.W.J., *Tetrahedron Lett.*, 36 (1995) 8281.
40. Hahn, M., *J. Chem. Inf. Comput. Sci.*, 37 (1997) 80.
41. Rogers, D.R. and Hopfinger, A.J., *J. Chem. Inf. Comput. Sci.*, 34 (1994) 854.
42. Dean, P.M., In Kubinyi, H. (Ed.), *3D QSAR in Drug Design: Theory, Methods and Applications*, ESCOM, Leiden, The Netherlands, 1993, p. 150.
43. Richards, W.G., In Pullman, A., Jortner, J. and Pullman, B. (Eds.), *Modelling of Biomolecular Structures and Mechanisms*, Kluwer Academic Publishers, Dordrecht, the Netherlands, 1995, p. 365.
44. Carbó, R., Leyda, L. and Arnau, M., *Int. J. Quantum Chem.*, 17 (1980) 1185.
45. Hodgkin, E.E. and Richards, W.G., *Int. J. Quantum Chem., Quantum Biol. Symp.*, 14 (1987) 105.
46. *Search_Compare*, Version 95.0, Molecular Simulations Inc., San Diego, CA.
47. Maple, J.R., Hwang, M.-J., Stockfisch, T.P., Dinur, U., Waldman, M., Ewig, C.S. and Hagler, A.T., *J. Comput. Chem.*, 15 (1994) 162.
48. Hwang, M.-J., Stockfisch, T.P. and Hagler, A.T., *J. Am. Chem. Soc.*, 116 (1994) 2515.
49. *Cerius²*, Version 3.0, Molecular Simulations Inc., San Diego, CA.
50. *Ligand_Design*, Version 95.0, Molecular Simulations Inc., San Diego, CA.
51. *Discover*, Version 95.0, Molecular Simulations Inc., San Diego, CA.
52. Froloff, N., Windemuth, A. and Honig, B., *Protein Sci.*, 6 (1995) 1293.
53. Pickett, S.D. and Sternberg, M.J.E., *J. Mol. Biol.*, 231 (1993) 825.
54. Eldridge, M.D., Murray, C.W., Auton, T.R., Paolini, G.V. and Mee, R.P., *J. Comput.-Aided Mol. Design*, 11 (1997) 425.
55. So, S.-S. and Richards, W.G., *J. Med. Chem.*, 35 (1992) 3201.
56. Ajay, *Chemometr. Intell. Lab.*, 24 (1994) 19.
57. Chandonia, J.-M., *Neural network based algorithms for protein structure prediction*, Ph.D. Thesis, The Committee on Higher Degrees in Biophysics, Harvard University, 1997.
58. Head, R.D., Smythe, M.L., Oprea, T.I., Waller, C.L., Green, S.M. and Marshall, G.R., *J. Am. Chem. Soc.*, 118 (1996) 3959.
59. Hansson, T., Marelus, J. and Åqvist, J., *J. Comput.-Aided Mol. Design*, 12 (1998) 27.
60. Guo, Z. and Brooks, C.L.I., *J. Am. Chem. Soc.*, 120 (1998) 1920.