CoMFA and Docking Studies on Glycogen Phosphorylase *a* Inhibitors as Antidiabetic Agents[#]

Philip Prathipati, Gyanendra Pandey, and Anil K. Saxena*

Medicinal Chemistry Division, Central Drug Research Institute, Chatter Manzil Palace,
Lucknow-226 001, India

Glycogen phosphorylase (GP_a) is a specific target for the design of inhibitors and may prevent glycogenolysis under high glucose conditions in type II diabetes. The carboxamides first reported by Hoover D. J. et al. (*J. Med. Chem.* **1998**, *41*, 2934−2938) are one of the major classes of GP_a inhibitors other than glucose derivatives. The recent, X-ray crystallographic analyses (Oikonomakos et al. *Biochim. Biophys. Acta* **2003**, *1647*, 325−332) have revealed a distinct mechanism of action for these inhibitors, which bind at a new allosteric site away from the inhibitory and catalytic sites. To elucidate the essential structural and physicochemical requirements responsible for binding to the GP_a enzyme and to develop predictive models, CoMFA and docking studies have been carried out on a series of indole-2-carboxamide derivates. The CoMFA model developed using pharmacophoric alignments and hydrogen-bonding fields demonstrated high predictive ability against the training ($r^2 = 0.98$, $q^2 = 0.68$) and the test set ($r^2_{pred} = 0.85$). Further the superimposition of PLS coefficient contour maps from CoMFA with the GP_a active site (PDB: 1lwo) has shown a high level of compatibility.

INTRODUCTION

Diabetes mellitus is characterized by chronic elevated blood glucose levels. Noninsulin dependent diabetes mellitus (NIDDM) affects 3−6% of adults in most industrialized countries and accounts for over 80% of all cases of diabetes.[1] It emerges in middle or late life and is associated with many side effects, viz. arteriosclerosis; strokes and peripheral vascular diseases; kidney disease (nephropathy); contributes to ulcers, and increases lower limb amputations by more than 4-fold.[2,3]

Although the cause of the commonly encountered form of type II diabetes is yet to be identified, it is well established that it is a polygenic disease characterized by multiple defects in insulin action in muscles, adipose tissues, and liver and in pancreatic insulin secretion.[4] In the excessive case, hepatic glucose production (HGP) is a significant contributor to diabetic hyperglycemia. The liver is the major regulator of plasma glucose levels in the post absorptive state, and in type II diabetes HGP is significantly elevated relative to nondiabetes.[5,6] The type II diabetes is treated with diet, exercise, and drugs that promote insulin release from the beta cells of the pancreas and with insulin. However the use of available hypoglycemic agents do not provide the complete tight control of plasma glucose by insulin therapy for a long term which often leads to the fear of dangerous hypoglycemic episodes in addition to the development of tolerance. Hence such treatments are not satisfactory, and there is a continued search for new chemical entities to treat diabetes.

Glycogen phosphorylase *a* (GP_a) has been exploited as a specific target of inhibitors that might prevent glycogenolysis under high glucose conditions in type II diabetes.[7−15] GP_a is an archetypical control enzyme and exists in two interconvertible states, a dephosphorylated form (GP_b, low activity and low substrate affinity) and the Ser14 phosphorylated form (GP_a, high activity and high substrate affinity form). The allosteric effectors can promote equilibrium between both forms, a less active T state and a more active R state.[16−18] In the liver the major control mechanism is through reversible phosphorylation that converts the inactive *b* form of the enzyme (predominantly in T state) to the active *a* form (predominantly in R state). In the muscle control, the phosphorylase is excreted both by reversible phosphorylation in response to nervous or hormonal signals and through noncovalent bonding of allosteric effectors.[16,17,19] As the inhibitors of glycogen phosphorylase may help the shift in the balance between glycogen synthesis and glycogen degradation in favor of glycogen syntheses in both muscle and liver, and so GP_a is considered a possible target, for drugs that aim to prevent unwanted glycogen breakdown and to stimulate glycogen synthesis in noninsulin dependent diabetes.[19−25]

GP_a contains at least six potential regulatory sites: (i) the ser14 phosphate recognition site, (ii) the AMP activatory and the glucose-6-P-inhibitory allosteric sites, (iii) the catalytic site that binds glycogen and glucose-1-P, (iv) the inhibitor site located 12 Å from the catalytic site which binds caffeine and related compounds, (v) the glycogen storage site, and (vi) the allosteric site at the dimer interface of the protein as revealed by a recent X-ray study. This dimer interface is the binding site of indole-2-carboxamides, a major class of GP_a inhibitors other than glucose derivatives, as reported by J. H. Dennis et al.[26] Several QSAR studies[27−36] have been

---

GLYCOGEN PHOSPHORYLASE *A* INHIBITORS

J. Chem. Inf. Model., Vol. 45, No. 1, 2005 **137**

reported for glucose derivatives, which bind at the AMP activatory and the glucose-6-P-inhibitory allosteric site.

Computer aided drug design involves either of the two most commonly used techniques: the indirect/ligand based design[37-45] or the direct design also known as structure based design. The latter has assumed increased importance due to rapid advances in the fields of structural and molecular biology. However it appears relatively difficult to find a predictive model based on the calculated interaction energies obtained by docking.[46,47] To overcome this problem, highly predictive QSAR and CoMFA models are being developed using the technique of structure-based alignments. These models explain the protein−ligand interactions and are found consistent with the crystal structure of protein with ligand.

Since its introduction several years ago CoMFA (Comparative Molecular Field Analysis)[48] has become one of the most powerful tools for QSAR and drug design. In fact, CoMFA has pioneered a new paradigm of three-dimensional QSAR[49] where the shapes, properties, etc. of molecules are related to specific molecular features (substituents, etc.) and their spatial relationship. The qualitative information of CoMFA contour maps gives key features on not only the ligand−receptor interaction but also the topology of the receptor. However comparative molecular field analysis (CoMFA) is critically dependent on the conformation and alignment criteria for molecules, and thus among the plethora of methods available for superimposition, the pharmacophoric alignments are considered to be the best.[50-55] A pharmacophoric hypothesis collects common features distributed in three-dimensional space representing groups in a molecule that participate in important interactions between drugs and their active sites. Hence a pharmacophore model provides crucial information about how well the common features of a subject molecule overlaps with the hypothesis model. In addition, it also helps in the identification of the energetically reasonable conformation, which fits to the active site. Such characterized three-dimensional models convey important information in an intuitive manner.

In addition to the standard steric and electronic field used in CoMFA, several types of fields have been calculated in advanced CoMFA. These are, the hydrogen bonding fields, which are created by assigning energies equal to the steric cutoff energy to lattice points that are close to H-bond accepting or donating atoms, and the indicator fields, which are created by squaring the original field at each lattice point, with the retention of the sign of the original field.

Also as an alternative approach to the computation of molecular potential, the comparative molecular similarity indices analysis (CoMSIA) has been developed.[56] Since the form of the distant functions in the standard Lennard-Jones and coulomb type potentials generate unrealistically extreme values as the surface of the molecules under examination is approached, this results in a drastic change in the shape of the potential functions. To circumvent this problem, Klebe et al. implemented the steric, electrostatic, and hydrophobic similarity indices which replace the distance functions of the Lennard-Jones and coulomb-type potentials with Gaussian-type functions (i.e. exp [-{alpha} r2]). The {alpha} term describes a 'local smearing' effect which places more weight on interactions close to the molecular surface with a smooth transition to more distal points and no arbitrary cutoff values. The hydrophobic potential used in this case is an atom-based

method developed by Viswanadhan et al.[57] The net result of this field is to lessen the effect of changes in the field descriptors associated with minor variations in molecular superposition or conformation.

As the pharmacophore derived from ligand based approaches describes the complementary active site of the receptor, the rigid docking in which precomputed molecular structures are placed at the target binding site may be useful in substantiating the 3D QSAR models for their utility in virtual screening. In this method the best-mapped conformer (to a pharmacophoric hypothesis) out of several hundreds of possible conformers for each molecule may be chosen to be overlaid on a previously cocrystallized ligand for docking studies. Such a docking methodology should comparatively be quicker than the commonly used docking approaches.[58]
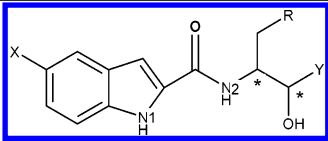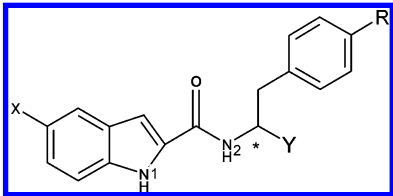
In view of the above, the CoMFA/CoMSIA model for Gp_a inhibitors derived using pharmacophoric alignment provided by the catalyst pharmacophoric hypothesis has been assessed for its complementarity with the active site of Gp_a by rigid body docking studies using the $HLGP_a$ enzyme crystal structure and using the scoring routines available in LU-DI.[59,60]

MATERIALS AND METHODS

Data Set. The QSAR analysis, CoMFA, and docking of glycogen phosphorylase-*a* inhibitors as anti-diabetic agents were performed on a series of 25 indole-2-carboxamides to quantitatively correlate the variation in structure and physicochemical properties with the in vitro human liver phosphyorylase *a* inhibitory activity (Table 1).[26] Though the size and the range of the data set used in the present analysis may be considered relatively small (25 compounds and 2 orders of log units) for the derivation and validation of a QSAR model in conventional statistical terms, however, in the realm of statistical model development, the limitations in the data set reflect a real world situation where no model is universal however ideal the data set used therein maybe. Thus, the validly and the universality of the ensuing QSAR models may be analyzed with respect to their statistical robustness and their complimentary (mirror image) with respect to the protein−ligand interactions as observed through X-ray crystallography derived complex.

Conformational Analysis, Pharmacophore Modeling, and Alignment Definition. The 3D structures of the compounds were built and optimized for their geometry using the modified CharmM force field, and the conformations were generated using the maximum limit of 250 conformations within a 20 Kcal cutoff for the common feature pharmacophore generation using the HIPHOP module[61,62] of Catalyst 4.5. The crystal structure of the compound CP320626 in its enzyme bound form and the most active molecule of this series were taken as a template (principal value = 2). All of the other 23 molecules were allowed to partially map (principal = 1) to the generated hypothesis. Finally the alignment rule was defined by a chemical function-mapping method and therefore based on a geometric fit of the chemical functions to the chemical features of the pharmacophore. The conformer selected for each compound corresponds to the conformation, which best fits the pharmacophore. These aligned conformers were then exported to Sybyl 6.9 for CoMFA and CoMSIA studies and to insightII for rigid fit docking and scoring studies using LUDI.

Table 1. Structures and Activities of the Molecules in the Training and Test (t) Sets and Predictions of the Best CoMFA Model

compound	X	R	Y	*	$-\log(\text{rHLGP}_a)$	CoMFA estimations
1^t	Cl	Ph	CONMe$_2$	3S,2R	−2.041	−2.137
2^t	Cl	Ph	CONMe$_2$	3R,2S	−3.255	−2.890
3^t	Cl	Ph	CONMe$_2$	3R,2R	−3.230	−2.939
4	Cl	Ph	CONMe$_2$	3S,2S	−3.940	−3.992
5	F	Ph	CONMe$_2$	3S,2R	−2.230	−2.292
6	Br	Ph	CONMe$_2$	3S,2R	−1.987	−2.137
7	H	Ph	CONMe$_2$	3S,2R	−2.643	−2.596
8	Cl	Ph	CONHMe	3S,2R	−2.833	−2.787
9	Cl	Ph	CONH$_2$	3S,2R	−3.230	−3.231
10	Cl	Ph	COOMe	3S,2R	−2.322	−2.329
11	Cl	Ph	COOH	3S,2R	−2.519	−2.421
12	Cl	Ph	CH$_2$OH	3S,2R	−2.833	−3.800
13	Cl	Cy	CONMe$_2$	3S,2R	−3.813	−3.859

compound	X	R	Y	*	$-\log(\text{rHLGP}_a)$	CoMFA estimations
14^t	Cl	H	CONMe$_2$	S	−1.913	−2.264
15	Cl	H	CONMe$_2$	R	−2.414	−2.389
16^t	Cl	H	CONHMe	S	−2.041	−2.524
17	Cl	H	CONHMe	R	−2.342	−2.222
18	Br	H	CONMe$_2$	S	−2.041	−2.249
19	F	H	CONMe$_2$	S	−2.633	−2.634
20	H	H	CONMe$_2$	S	−2.602	−2.592
21	OMe	H	CONMe$_2$	S	−3.672	−3.621
22	Cl	H	CO(1-piperidin-4-ol)	S	−2.414	−2.270
23	Cl	F	CO(1-piperidin-4-ol)	S	−2.311	−2.370
24	Cl	H	COOMe	S	−2.146	−2.110
25	Cl	H	COOH	S	−3.230	−3.249

Rational Selection of Training Set and Test Set. In view of the several recent findings that q^2 appears to be the necessary but not the sufficient condition for the model to have a high predictive power, the emphasis has been given for the validation on an external test set.[63−66] Among the various methods available to classify the data set into training and test sets, viz. random sampling, activity ranked clustering, and descriptor based clustering, the descriptor based clustering has been shown to be the most effective method for selection of a representative test set (both in terms of activity and structural diversity). So, the hierarchical cluster analysis using topological descriptors was used to rationally design representative and structurally diverse training set as reported in our previous study.[66] In brief, the study of structural dissimilarity was based on most contributing topological descriptors that efficiently encode structural information and are currently used in diversity analysis for combinatorial library design. The 20 clusters for GP$_a$ inhibitory activity were determined using the Euclidean distances between the descriptors. The distances were calculated using the Ward linkage algorithm as implemented in *Systat* version 7.[67] One representative molecule was chosen to represent each cluster. These representative molecules were taken for the training set, and the rest were considered for the test set. Thus the entire set of the molecules was divided into a training set consisting of 20 molecules and the test set consisting of five molecules spanning the activity

data (log units) in 1.93 and 1.341 orders of magnitude, respectively, knowing full well that the 25 compound data set is too small for such an analysis. It was thus done in a methodical way to select a training set in the above manner so that it has full information of the entire data set. Further this method may not yield a biased training and test set, because both the training and test sets span an almost equal range in terms of dependent parameters.

CoMFA. All of the CoMFA experiments were performed on a Silicon Graphics Octane workstation with the molecular modeling software *Sybyl 6.9*.[68] The molecular atomic charges were calculated using the Gasteiger-Hückel protocol. The steric and electronic field energies were calculated using fields viz. standard CoMFA field, H-bonding fields, indicator fields, and parabolic fields[69,70] for the CoMFA calculation. All models were investigated using the full cross-validated partial least-squares (PLS) (leave-one-out) method with CoMFA standard options for the scaling of variables. Initially the minimum-sigma (column filtering) was set to 2.0 kcal/mol to improve the signal-to-noise ratio by omitting those lattice points whose energy variation is below this threshold. The statistical significance and predictive ability of the resulting models was assessed using leave-one-out cross-validated r^2, also called q^2. The conventional r^2 was considered as a measure of the predictive ability within the training set, while the q^2 has been considered as a measure of predictive ability outside the training set. The q^2 values

GLYCOGEN PHOSPHORYLASE *A* INHIBITORS

J. Chem. Inf. Model., Vol. 45, No. 1, 2005 **139**

Table 2. Results of the Common-Feature Hypothesis Generation Study for HLGP$_a$ Inhibitors Belonging to Indole-2-carboxamide Class of Compounds

hypo	features[a]	rank scores	direct hit[b]	partial hit[b]
1	RHDA	282.346	1111111111111111111111111111	0000000000000000000000000000
2	RHDA	282.346	1111111111111111111111111111	0000000000000000000000000000
3	RHDD	281.756	1111111111111111111111111111	0000000000000000000000000000
4	RHDD	277.007	1111111111111111111111111111	0000000000000000000000000000
5	HHDA	270.574	1111111111111111111111111111	0000000000000000000000000000
6	HHDA	269.594	1111111111111111111111111111	0000000000000000000000000000
7	RHDA	268.352	1111111111111111111111111111	0000000000000000000000000000
8	RHDA	268.099	1111111111111111111111111111	0000000000000000000000000000
9	HHDA	268.044	1111111111111111111111111111	0000000000000000000000000000
10	HHDD	267.876	1111111111111111111111111111	0000000000000000000000000000

[a] D, hydrogen-bond donor; A, hydrogen-bond acceptor; R, ring aromatic; and H, hydrophobic. [b] Direct hit, all the features of the hypothesis are mapped. Direct hit = 1 means yes and direct hit = 0 is no; partial hit, partial mapping of the hypothesis. Partial hit = 1 means yes and partial hit = 0 means no.

were further optimized in the models exhibiting $q^2 > 0.2$ by the use of region focusing, which increases the resolution (decrease the lattice spacing) of CoMFA models without sacrificing predictive ability. Herein, the lattice points in a CoMFA region were weighed to enhance or attenuate their contribution to subsequent analyses by carefully selecting the weighting parameters. Further, region focusing also removes clutter from visual representations and greatly enhances the clarity of presentations by emphasizing the relevant field regions. Thus the PLS region focusing weighs the lattice points in a field using parameters derived from the CoMFA model and serves to enhance signal-to-noise ratio in the analysis by suppressing data from lattice points which do not contribute to the model. Region focusing uses some parameters drawn from the CoMFA PLS model—which draws on information from all grid point—to assign a weight to each grid point. Although any grid point-specific PLS parameter may be used, but the most robust is the discriminant power, which is the fraction of the variance in a model's components derived from each grid point. Thus region focusing of the hydrogen bonding field was accomplished in advanced CoMFA module by decreasing the lattice point spacing from 2 Å to 1 Å and selectively reweighing the discriminant power of the grid points in a region and creating a new CoMFA column using the focused region file and then rederiving the model.

CoMSIA. The CoMSIA technique exhibits greater robustness with respect to both region shifts and small shifts within the alignments with no application of arbitrary cutoffs and more intuitively interpretable contour maps. The standard settings (probe with charge +1, radius 1 Å and hydrophobicity +1, hydrogen-bond donating +1, hydrogen-bond accepting +1, attenuation factor *a* of 0.3, and grid spacing 2 Å) were used in CoMSIA to calculate seven different fields viz. steric, electrostatic, hydrophobic, hydrogen bond acceptor, hydrogen bond donor, steric & electrostatic, and hydrogen bond acceptor & donor.

Docking. All docking studies described herein were performed on a SGI O2 workstation using the *insightII* molecular modeling software. For the current study we have used the recently reported crystal structure of GP$_a$ enzyme (PDB: 1lwo). The inhibitor, CP320626, was unmerged from its monomer complex with the protein. The protein and the ligand were then automatically assigned the bond order and hydrogens were added. The other 24 molecules in their common feature pharmacophore aligned conformers were
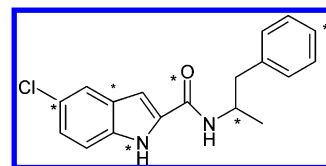
Figure 1. The structure of the molecule # 23 'CP320626' that was used as a template on which all the other 24 molecules were superimposed using the rigid RMS fitting procedure, the stars indicate the atoms selected as the fitting centers.

manually docked into the active site by superimposing them onto the crystal structure of CP320626. Rigid RMS fitting implemented in the Search Compare module of *insightII* was used to align the molecules, and the points considered for the alignment are shown in Figure 1. The activity of the molecules is considered to be due to the stabilization attained by the ligand—receptor complex, which is directly proportional to the magnitude of the interaction energy between them. These interaction energies principally scored in terms of hydrophobic (*Lipo score*), hydrogen bonding (*HB score & No. HB*), and unfavorable steric contacts (*Cont*) have been calculated using the Ludi module of *insightII* and the regression models explaining the combined effect of the various scores were generated using *Systat*.

RESULTS

Common Feature Hypothesis. The common feature hypothesis analysis using one or maximum combination of any five features, viz. hydrogen bond acceptor (A), hydrogen bond donor (D), hydrophobic (H), positive ionizable (P), and ring aromatic (R) led to the generation of 10 hypotheses containing four features with a good range of ranking scores (282.346 to 267.876) (Table 2). Four hypotheses viz. 1, 2, 7, and 8 consist of the same common feature hypothesis functions RHDA, while hypotheses 3, 4 and 7, 8 contain common feature hypothesis functions RHDD and RHDA, respectively. Finally, hypothesis 9 and 10 contained HHDA and HHDD, respectively, as the common feature functions. All 25 molecules belonging to the training and test sets mapped to every feature of all the hypotheses as indicated by the direct hit mask (Table 2) and the further the higher the ranking scores of the hypotheses indicate that it is less likely that the molecule in the training set fits the hypothesis by a chance correlation.

CoMFA and CoMSIA Models. The four CoMFA and seven CoMSIA models developed by pharmacophore-based

Table 3. Summary of Various CoMFA and CoMSIA Models of HLGP$_a$ Inhibitors Belonging to Indole-2-carboxamide Class of Compounds

model no.	CoMFA type[a]	comp[b]	q^{2c}	Spress[d]	r2[e]	S[f]
1	COMFA_H	6	0.261	0.652	0.977	0.115
2	COMFA_H (reg. foc.)	5	0.681	0.414	0.983	0.097
3	COMFA_S	1	0.040	0.640	0.980	0.120
4	COMFA_I	2	−0.150	0.720	0.980	0.120
5	COMFA_P	2	0.030	0.660	0.980	0.120
6	COMSIA_SE	1	−0.080	0.680	0.980	0.120
7	COMSIA_DA	2	−0.100	0.700	0.980	0.120
8	COMSIA_H	8	0.080	0.800	0.980	0.120
9	COMSIA_S	2	−0.020	0.680	0.980	0.120
10	COMSIA_E	1	0.000	0.650	0.980	0.120
11	COMSIA1_D	1	−0.100	0.680	0.980	0.120
12	COMSIA_A	5	0.060	0.710	0.980	0.120

[a] CoMFA_H, CoMFA_S, CoMFA_I, and CoMFA_P represent the CoMFA derived using the hydrogen bonding, standard steric and electrostatic fields, indicator, and parabolic fields, respectively, while the CoMSIA_SE, CoMSIA_DA, CoMSIA_H, CoMSIA_H, CoMSIA_D, and CoMSIA_A fields represent the CoMSIA fields derived using the steric and electrostatic, hydrogen bond donor and acceptor, hydrophobic, hydrogen bond donor, and hydrogen bond acceptor fields, respectively. [b] Optimum number of components. [c] Leave-one-out cross validated r^2. [d] Root-mean-square error of all target predictions. [e] Proportion of the original variance explained. [f] Standard error of estimate.
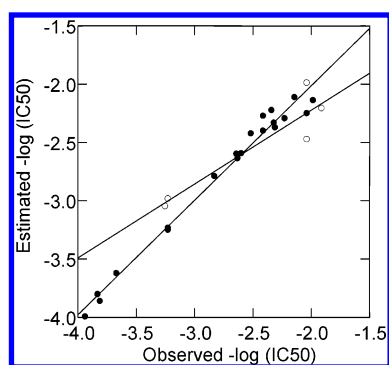
Figure 2. Graph of observed activity (*x*-axis) versus predicted activity (*y*-axis) of training (●) and test set (○) molecules from the CoMFA model.

alignments and using various field types resulted in 11 models with q^2 ranging from 0.261 to −0.15 (Table 3). Later various combinations of the CoMFA and CoMSIA models were generated using combinations of two, three, four, and all fields, but none of these models lead to the significant improvement in the q^2 values of the models. Hence the CoMFA model (model number 1) developed using hydrogen bonding field which exhibited the highest q^2 value (0.261) was optimized using the PLS region focused hydrogen bonding field. It led to the development of a new CoMFA model with hydrogen bond donor (51.3%) and acceptor (48.7%) contributions exhibited superior statistical parameters (q^2 = 0.681, Spress = 0.414, r^2 = 0.983, r$^2_{bootstrap}$ = 0.976, s = 0.097 and F = 157.76 and r$^2_{pred}$ = 0.862) both for the training and the test set (Table 3, Figure 2).

Docking. The rigid body docking was performed manually using the common feature hypothesis aligned conformers, and the interactions were scored using the LUDI module of the molecular modeling package *insightII* (Table 4). The combined effect of the various interactions in terms of *No. HB*, *Sum* (sum of all the scores), *HB_score*, *Lipo_score*, *No_Rot_Bonds*, and *Cont* was modeled using MLR and the

Table 4. Scores of Various GP$_a$-Ligand Interaction Energies as Estimated by LUDI

molecule	no. HB[a]	sum[b]	HB score[c]	lipo score[d]	no. rot. bonds[e]	cont[f]
GP1	9	170	72	320	5	−126
GP10	9	182	81	323	5	−126
GP11	9	190	80	332	5	−126
GP12	5	186	8	400	5	−126
GP13	9	178	71	329	5	−126
GP14	5	263	51	409	4	−101
GP15	5	209	0	406	4	−101
GP16	4	241	83	355	4	−101
GP17	5	188	0	385	4	−101
GP18	5	275	51	421	4	−101
GP19	3	250	47	400	4	−101
GP2	8	171	49	344	5	−126
GP20	3	218	45	370	4	−101
GP21	3	298	45	450	4	−101
GP22	7	429	140	486	4	−101
GP23	7	435	152	480	4	−101
GP24	5	278	134	341	4	−101
GP25	5	278	131	344	4	−101
GP3	5	253	90	385	5	−126
GP4	4	276	33	465	5	−126
GP5	8	249	92	379	5	−126
GP6	4	155	84	293	5	−126
GP7	4	188	84	326	5	−126
GP8	4	161	84	299	5	−126
GP9	8	208	86	344	5	−126

[a] No. HB = number of hydrogen bonds. [b] Sum = sum of all the scores. [c] HB score = hydrogen bonding score. [d] No. rot. bonds = number of rotatable bonds. [e] Cont = steric contact scores.

best model with *No. HB*, *Lipo_score*, and *Cont* exhibited a poor r^2 value of 0.317 (eq 1); however, the removal of two outliers led to a significant improvement in terms of r^2 (0.578) (eq 2).

$$-\log(IC50) = 2.662 \ (\pm 1.848) \ CONSTANT + 0.085 \ (\pm 0.063) \ No. \ HB - 0.005 \ (\pm 0.002) \ Lipo \ score - 0.035 \ (\pm 0.012) \ Cont \quad (1)$$

$$n = 25, \ r^2 : \ 0.317, \ s = 0.564, \ F = 3.247 \ (P = 0.03)$$

$$-\log(IC50) = 4.244 \ (\pm 1.504) \ CONSTANT + 0.133 \ (\pm 0.049) \ No. \ HB - 0.007 \ (\pm 0.002) \ Lipo \ score - 0.043 \ (\pm 0.009) \ Cont \quad (2)$$

$$n = 23, \ r^2 = 0.578, \ s = 0.427, \ F = 8.554 \ (P = 0.001)$$

DISCUSSION

Common Feature Hypothesis. Among the 10 hypotheses that demonstrated the good scores,[62] the hypothesis that was complementary with the ligand-GP$_a$ enzyme interactions as well as demonstrated the highest score was chosen to be the best hypothesis, considering the dual validation criteria for statistical robustness and the ligand−receptor interaction complementarity. Hence the first hypothesis (Figure 3) with the highest score value '282.346' was selected for the generation of bioactive conformations and molecular alignment for 3D QSAR studies using CoMFA and CoMSIA. This pharmacophoric model (hypothesis) consisted of four features, viz. hydrophobic, hydrogen bond donor (HBD), hydrogen bond acceptor (HBA), and ring aromatic, and its mapping to the molecule no. 23 (CP320626) is shown in Figure 3 well corresponded to the observed major interactions

GLYCOGEN PHOSPHORYLASE *A* INHIBITORS

J. Chem. Inf. Model., Vol. 45, No. 1, 2005 **141**
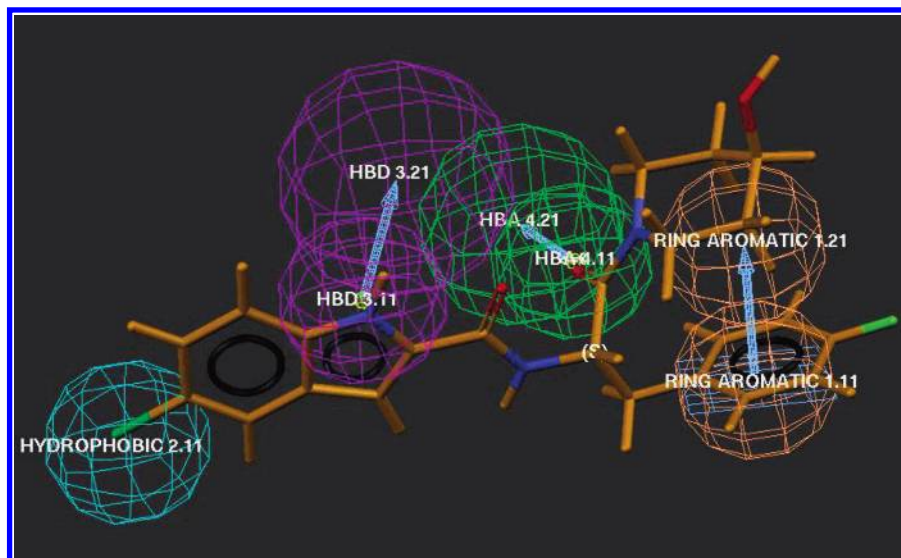
Figure 3. The structure of the molecule # 23 'CP320626' that was used as a template on which all the other 24 molecules were superimposed using the rigid RMS fitting procedure, the stars indicate the atoms selected as the fitting centers.

between the ligand and the GP$_a$ enzyme. The major interactions between the ligand and the GP$_a$ enzyme are as follows: the hydrophobic interactions between the indolyl aromatic ring and hydrophobic residues such as Pro229, Trp189, and Trp67 which is represented by a hydrophobic feature on the indolyl ring; the hydrogen bonding interactions between N1 of indolyl and O of the main chain represented as a hydrogen bond donor on N of indolyl ring; the other hydrogen bonding interaction observed between the CO of the piperidyl oxygen to the Lys191 is represented by the HBA feature on the CO of piperidyl oxygen; and the ring aromatic feature representing the aromatic $\pi-\pi$ staking between the 4-F phenyl ring and subunit residues His57′, Thr38′, and Pro188′. Hence it may be useful as an initial filter for virtual screening apart from providing the required bioactive conformation and molecular alignment for 3D QSAR studies.

CoMFA Model. In the present study, out of the eleven CoMFA and CoMSIA models developed initially, only the CoMFA hydrogen bonding field model was taken up for region focusing because the initial q^2 value from the best CoMSIA model was too low (q^2 = 0.08 with 8 components) to be considered for optimization as compared to the best CoMFA model (q^2 = 0.28 with 6 components) in the first few runs and also showed complimentarity with the active site information derived from X-ray structural studies.

The region focusing technique of CoMFA was used to optimize the best PLS model (CoMFA_H). This technique allows for an increase in the resolution by decreasing the lattice spacing of CoMFA models (from default 2.0 Å to 1.0 Å, in this case) without sacrificing predictive ability. It reweight the lattice points in a CoMFA region to enhance or attenuate the contribution of the lattice points in subsequent analyses as the careful selection of weighting parameters and sharpness of focusing leads to the improvement in the PLS calculations even at a much finer grid spacing. It thus improves the q^2 by reducing the cross-correlated "brown" noise in the data, hence the region focusing on the above best CoMFA model led to a significant enhancement in the quality of the model in terms of q^2.

Though this CoMFA model with 6 components for a 20 compound data set may not go well with the generally accepted [n(# observations)/k (# variables) = 5] norm for QSAR models based on MLR, however, as of now, there are there are no studies that specify the limits on the ratio of the number of samples of a data set to the number of components in PLS. Most importantly, the link between the numbers of components relative to the number of observations is mainly valid for MLR and not for PLS where this "rule" was adopted as a protection for singular matrices when the inverse (X′X) is calculated. Since PLS has no such algebraic requirement, in theory, n-1 components are permissible in a PLS model. However, as PLS is a dimension reduction technique, it is the norm to determine the optimal number of components and a common method to determine the same is to use cross-validation, wherein the optimal number of the PLS component is the one which exhibits the highest q$^2_{LOO.}$

Hence this model with the highest q^2 value (0.685) was further validated by bootstrapping as well as external test set predictions (r$^2_{pred}$ = 086).

The PLS analysis of this model, using the energy fields as descriptors and the binding affinity as the dependent variable, highlights the relative importance of different chemical groups for affinity to the binding sites for interaction. The contour maps (Figure 4) obtained using the pharmacophoric alignment are very different in nature than those obtained using a common substructure alignment in normal COMFA, because in the former case the variation in activity is explained not merely by the physicochemical descriptors in the form of fields but also influenced by the different conformational preferences of the compounds for interacting with the active site of the enzyme. As the structure of the glycogen phosphorylase *a* is known and the PLS coefficient contour maps provides an opportunity to interpret features indicated in the contour maps with respect to the protein environment, so the PLS contour maps obtained by the 3D QSAR analysis using advanced CoMFA were compared with the geometry and properties of the binding pocket (Figure 5). In general, there was an agreement

142 *J. Chem. Inf. Model., Vol. 45, No. 1, 2005*

PRATHIPATI ET AL.
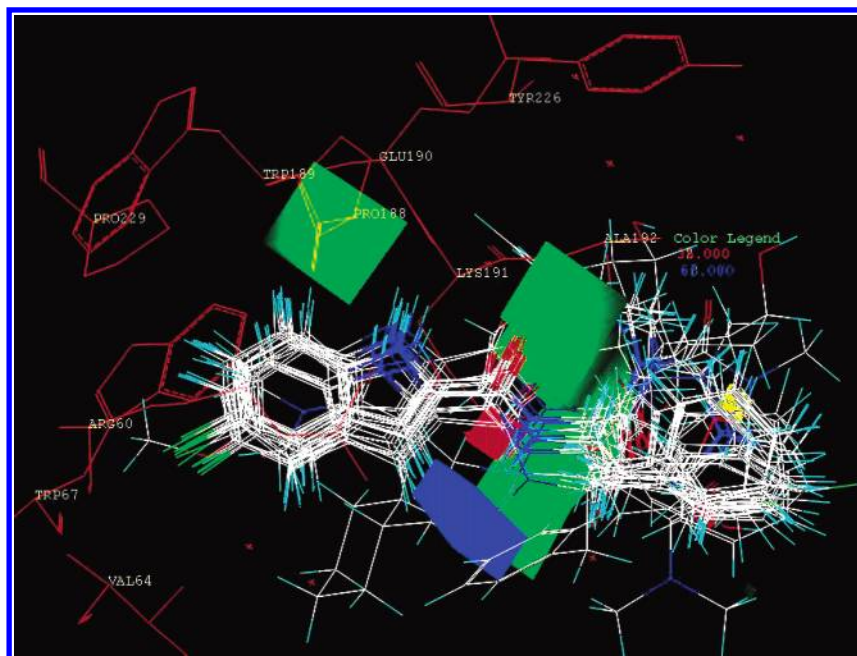
Figure 4. The superimpositions of all the molecules (colored by atom) along with the CoMFA "StDev*Coeff" contours and binding site residues in the background (colored in orange red).
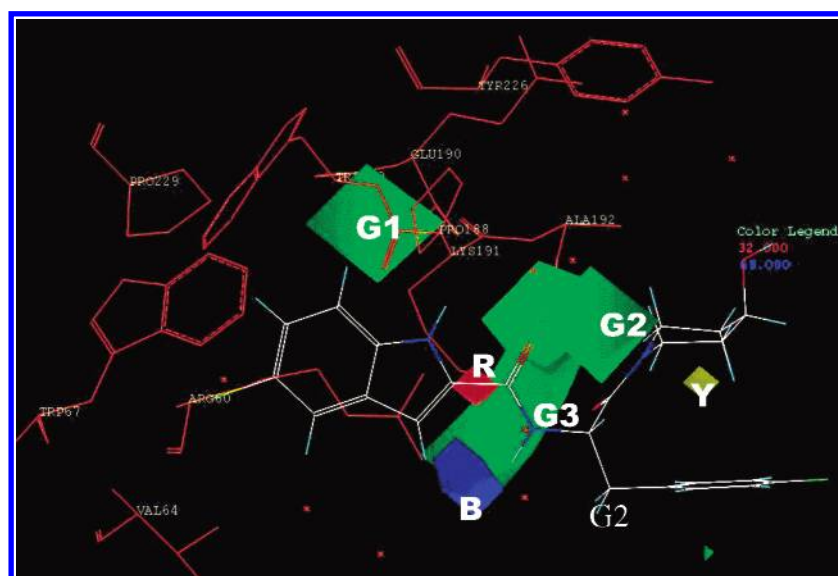


Figure 5. Comparison between the PLS coefficient "StDev*Coeff" maps and the amino acid residues located close to the binding pocket (the protein shown in orangered, CP320626 colored by atom). The regions G1−G3 indicate green contours (hydrogen bond acceptor favorable region), Y indicates an yellow contour (hydrogen bond acceptor disfavorable region), B indicates a blue contour (hydrogen bond donor favorable region), and R indicates a red contour (hydrogen bond donor disfavorable region).

between the nature of the "StDev*Coeff" plots generated using the default contour levels (80% and 20% for the favorable and unfavorable contributions respectively) and the positions and type of the amino acid residues implicated in interactions with the ligand at the active site. The green contour field (polyhedra) 'G1' located in a hydrogen-bonding region close to Trp189, Pro188, and Glu190 was consistent with the observation that the H-bonding between N1 of the indole ring and the main chain oxygen of Glu190 is important for enhancing the activity. The other green polyhedra 'G3' located in a region close to Lys191 and a water molecule which in turn hydrogen bonds to the main chain carbonyl of Ala192 is consistent with the observation that O1 of the

indolyl carbonyl forms a hydrogen bond through water molecules to Ala192 and O2 of the piperidyl carbonyl forms a hydrogen bond to 'N' of Lys191. However no amino acids were found in the vicinity of the G2 green polyhedra, which may correspond to the residues of the other subunit whose crystal coordinates were not available in PDB (pdb1lwo) at the time of this study. The blue polyhedra (B) located in the vicinity of a water molecule of the protein which may in turn hydrogen bond to the residues of other subunits of the dimer is consistent with the interactions of N2 of indole carboxamide with other residues of other subunits through water as observed in the crystal structure. The red polyhedra in the vicinity of N2 of indole carboxamide corresponds to

GLYCOGEN PHOSPHORYLASE *A* INHIBITORS

J. Chem. Inf. Model., Vol. 45, No. 1, 2005 **143**
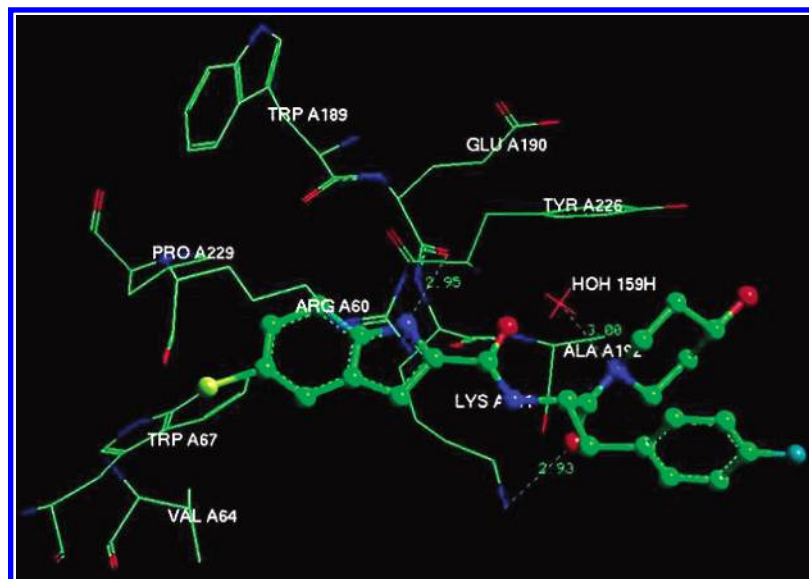
Figure 6. The binding mode of molecule # 23 [CP320626] (ball-and-sticks) in the active site of GP$_a$. The active site residues in the binding pocket are shown in color by element and the hydrogen bonds are indicated as broken green lines.

the unfavorable hydrogen bond donor region, and the yellow polyhedra 'Y' corresponds to the unfavorable hydrogen-bond acceptor region in the vicinity of N of piperidyl ring.

Docking. Some of the major interactions of molecule # 23 CP320626 with the active site of the enzyme are shown in Figure 6. The MLR model explaining the combined contribution of various interaction energies reveals that primarily hydrogen bonding energies contribute for binding, while hydrophobic and steric forces parametrized as Lipo_score and Cont, respectively, negatively contribute toward binding affinity. These results are also consistent with the CoMFA models wherein hydrogen-bonding fields explained well the variation in the activity. However the model derived by quantitatively correlating the scores and the variation in activity showed relatively poor correlation ($r^2 = 0.57$) as compared to those observed in CoMFA ($r^2 = 0.98$ and $r^2_{pred} = 0.84$). There may be several reasons for the observed low correlation of the interaction-energies versus binding affinity as observed elsewhere,[46] which may be due to the major lacunae of accommodating small deviations of atomic coordinates in the existing scoring protocol. Such deviations are likely to occur may be due to the the small errors in atomic coordinates even in well-refined X-ray structures, which have an influence on the energy landscape of ligand binding. Further there may be errors in the correct placement of the ligand in the binding pocket and in the optimization of its noncovalent interactions with the protein. In addition the problem is acquainted with our limitations in understanding of the physics and thermodynamics of ligand binding; especially solvation and entropic effects, which are at the moment fully neglected in interaction-energy based models.[47]

CONCLUSION

The present studies were aimed at deriving predictive models capable of elucidating the structural requirements for GP$_a$ inhibitory activity consistent with the information derived from the X-ray crystal structure. The common feature hypothesis study indicates the importance of the hydrophobic region in the vicinity of the position-5 of the indolyl aromatic

ring that could establish hydrophobic interactions with Pro229, Trp189, and Trp67, a hydrogen bond donor region in the vicinity of N of the indolyl ring which may establish hydrogen bond interactions with the residues Glu190, and a hydrogen bond acceptor region at the CO of the piperidyl ring for favorable hydrogen bonding interactions with the residues His57′, Thr38′, Leu39′, and Pro188′. The CoMFA study with hydrogen bonding fields demonstrates green contours, for the favorable hydrogen bond acceptor region around the indolyl aromatic ring and the indolyl carbonyl group, the blue polyhedra located around the N2 of indole carboxamide, for the favorable hydrogen bond donor region, and the yellow contour in the vicinity of N of piperidyl ring, for unfavorable hydrogen bonding interactions. These results are also consistent with the inferences from X-ray crystallographic studies and our pharmacophore analysis. Further the manual docking and scoring results using *LUDI* also illustrate the positive contribution of hydrogen bonding groups and the negative contribution of lipophilic and steric contacts.

REFERENCES AND NOTES

(1) Kenny, S. J.; Aubert, R. E.; Geiss, L. S. *Diabetes in America*, 2nd ed.; Harris, M., Ed.; NIH publication: 1995; pp 47-67.

(2) Panzram, G. Mortality and survival in Type 2 (Noninsulin-Dependent) Diabetes mellitus. *Diabetotogia* **1987**, *30*, 123−31.

(3) Nathan, D. M.; Long-term Complications of diabetes mellitus. *New Engl. J. Med.* **1993**, *328*, 1676−85.

(4) DeFronzo, R. A.; Bonadonna, R. C.; Ferrannini, E. Pathogenesis of NIDDM. A balanced overview. *Diabetes Care* **1992**, *15*, 8−369.

(5) Consoli, A. Role of liver in pathophysiology of NIDDM. *Diabetes Care* **1992**, *15*, 430−441.

(6) Gerich, J. E. Role of liver and muscle in type II diabetes. *Horm. Metab. Res.* **1992**, *26*, 18−21.

(7) Martin, J. L.; Veluraja, K.; Johnson, L. N.; Fleet, G. W. J.; Ramsden, N. G.; Bruce, I.; Oikonomakos, N. G.; Papageorgiou, A. C.; Leonidas, D. D.; Tsitoura, H. S. Glucose analogue inhibitors of glycogen

phosphorylase: the design of potential drugs for diabetes. *Biochemistry* **1991**, *30*, 10101−10116.

(8) Watson, K. A.; Mitchell, E. P.; Johnson, L. N.; Son, J. C.; Bichard, C. J. F.; Orchard, M. G.; Fleet, G. W. J.; Oikonomakos, N. G.; Leonidas, D. D.; Kontou, M.; Papageorgiou, A. C. Design of inhibitors of glycogen phosphorylase: a study of alpha- and beta-C-glucosides and 1-thio-beta-D-glucose compounds. *Biochemistry* **1994**, *33*, 5745−5758.

(9) 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.; Zographos, S. E. Glucose Analogue Inhibitors of Glycogen Phosphorylase: From Crystallographic Analysis to Drug Prediction Using GRID Force-Field and GOLPE Variable Selection. *Acta Crystallogr. Sect.* **1995**, *D51*, 458−472.

(10) 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.; Fleet, G. W. J. Potent inhibition of glycogen phosphorylase by a spirohydantoin of glucopyranose: first pyranose analogues of hydantocidin. *Tetrahedron Lett.* **1995**, *36*, 2145−2148.

(11) Oikonomakos, N. G.; Kontou, M.; Zographos, S. E.; Watson, K. A.; Johnson, L. N.; Bichard, C. J. F.; Fleet, G. W. J.; Acharya, K. R. *N*-acetyl-{beta}-D-glucopyranosylamine: A potent T-state inhibitor of glycogen phosphorylase. A comparison with {alpha}-D-glucose. *Protein Sci.* **1995**, *4*, 2469−2477.

(12) Zographos, S. E.; Oikonomakos, N. G.; Tsitsanou, K. E.; Leonidas, D. D.; Chrysina, E. D.; Skamnaki, V. T.; Bischoff, H.; Goldman, S.; Schram, M.; Watson, K. A.; Johnson, L. N. The structure of glycogen phosphorylase b with an alkyldihydropyridine-dicarboxylic acid compound, a novel and potent inhibitor. *Structure* **1997**, *5*, 1413−1425.

(13) Gregoriou, M.; Noble, M. E. M.; Watson, K. A.; Garman, E. F.; Krulle, T. M.; Fuente, C.; Fleet, G. W. J.; Oikonomakos, N. G.; Johnson, L. N. The structure of a glycogen phosphorylase glucopyranose spirohydantoin complex at 1.8 A resolution and 100 K: The role of the water structure and its contribution to binding. *Protein Sci.* **1998**, *7*, 915−927.

(14) Oikonomakos, N. G.; Tsitsanou, K. E.; Zographos, S. E.; Skamnaki, V. T.; Goldmann, S.; Bischoff, H. Allosteric inhibition of glycogen phosphorylase a by the potential antidiabetic drug 3-isopropyl 4-(2-chlorophenyl)-1,4-dihydro-1-ethyl-2-methyl-pyridine-3,5,6-tricarboxylate. *Protein Sci.* **1999**, *8*, 1930−1945.

(15) Oikonomakos, N. G.; Skamnaki, V. T.; Tsitsanou, K. E.; Gavalas, N. G.; Johnson, L. N. A new allosteric site in glycogen phosphorylase b as a target for drug interactions. *Structure* **2000**, *8*, 575−984.

(16) Johnson, L. N.; Hajdu, J.; Acharya, K. R.; Stuart, D. I.; McLaughlin, P. J.; Oikonomakos, N. G.; Barford, D. *Allosteric Enzymes*; CRC Press: Boca Raton, FL, 1989; pp 81−127.

(17) Johnson, L. N. Glycogen phosphorylase: control by phosphorylation and allosteric effectors. *FASEB J.* **1992**, *6*, 2274−2282.

(18) Oikonomakos, N. G.; Acharya, K. R.; Johnson, L. N. Rabbit muscle glycogen phosphorylase b. The structural basis of activation and catalysis (review). *Post-translational Modification of Proteins*; CRC Press: Boca Raton, FL, 1992; pp 81−151.

(19) Newgard, C. B.; Hwang, P. K.; Fletterick, R. J. The family of glycogen phosphorylase, structure and function. *Crit. Rev. Biochem. Mol. Biol.* **1989**, *24*, 69−99.

(20) Board, M.; Johnson, L. N. Effects of N-Acetyl-b−D-glucopyranosylamine on glycogen metabolism by isolated hepatocytes. *Diabetes Res.* **1995**, *28*, 95−109.

(21) Cori, C. F.; Cori, G. T. Mechanism of formation of hexosemonophosphate in muscle and isolation of a new phosphate ester. *Proc. Soc. Exp. Biol. Med.* **1936**, *34*, 720−705.

(22) Hers, H. Etudes enzymatiques sur fragments hepatiques; application a la classification des glycogenoses. *Rev. Int. Hepatol* **1959**, *9*, 35−36.

(23) Gorin, F. A.; Mullinax, R. L.; Ignacio, P. C.; Neve, R. L.; Kurnit, D. M. McArdle's and Hers' disases: glycogen phosphorylase transcriptional expression in human tissues. *J. Neurogenet.* **1987**, *4*, 293−308.

(24) van den Berghe, G. The role of the liver in metabolic homeostasis: implications for inborn errors of metabolism. *J. Inherit. Metab. Dis.* **1991**, *14*, 407−20.

(25) Puhakainen, I.; Koivisto, V. A.; Yki-Jarvinen, H. No reduction in total hepatic glucose output by inhibition of gluconeogenesis with ethanol in NIDDM patients. *Diabetes* **1991**, *40*, 1319−27.

(26) Hoover, D. J.; Lefkowitz-Snow, S.; Burgess-Henry, J. L.; Martin, W. H.; Armento, S. J.; Stock, I. A.; McPherson, R. K.; Genereux, P. E.; Gibbs, E. M.; Treadway, J. L. Indole-2-carboxamide inhibitors of human liver glycogen phosphorylase. *J. Med. Chem.* **1998**, *41*, 2934−2938.

(27) Pan, D.; Liu, J.; Senese, C.; Hopfinger, A. J.; Tseng, Y. Characterization of a ligand−receptor binding event using receptor-dependent four-dimensional quantitative structure−activity relationship analysis. *J. Med. Chem.* **2004**, *47*(12), 3075−88.

(28) Pan, D.; Tseng, Y.; Hopfinger, A. J. Quantitative structure-based design: formalism and application of receptor-dependent RD-4D-QSAR analysis to a set of glucose analogue inhibitors of glycogen phosphorylase. *J. Chem. Inf. Comput. Sci.* **2003**, *43*(5), 1591−607.

(29) Zamora, I.; Oprea, T.; Cruciani, G.; Pastor, M.; Ungell, A. L. Surface descriptors for protein−ligand affinity prediction. *J. Med. Chem.* **2003**, *46*(1), 25−33.

(30) Gohlke, H. Klebe DrugScore meets CoMFA: adaptation of fields for molecular comparison (AFMoC) or how to tailor knowledge-based pair-potentials to a particular protein. *J. Med. Chem.* **2002**, *45*(19), 4153−70.

(31) Marchand-Geneste, N.; Watson, K. A.; Alsberg, B. K.; King, R. D. New approach to pharmacophore mapping and QSAR analysis using inductive logic programming. Application to thermolysin inhibitors and glycogen phosphorylase B inhibitors. *J. Med. Chem.* **2002**, *45*(2), 399−409.

(32) So, S. S.; Karplus, M. Evaluation of designed ligands by a multiple screening method: application to glycogen phosphorylase inhibitors constructed with a variety of approaches. *J. Comput.-Aided Mol. Des.* **2001**, *15*(7), 613−47.

(33) Venkatarangan, P.; Hopfinger, A. J. Prediction of ligand−receptor binding thermodynamics by free energy force field three-dimensional quantitative structure−activity relationship analysis: applications to a set of glucose analogue inhibitors of glycogen phosphorylase. *J. Med. Chem.* **1999**, *42*(12), 2169−79.

(34) So, S. S.; Karplus, M. A comparative study of ligand−receptor complex binding affinity prediction methods based on glycogen phosphorylase inhibitors. *J. Comput.-Aided Mol. Des.* **1999**, *13*(3), 243−58.

(35) Pastor, M.; Cruciani, G.; Clementi, S. Smart region definition: a new way to improve the predictive ability and interpretability of three-dimensional quantitative structure−activity relationships. *J. Med. Chem.* **1997**, *40*(10), 1455−64.

(36) Cruciani, G.; Watson, K. A. Comparative molecular field analysis using GRID force-field and GOLPE variable selection methods in a study of inhibitors of glycogen phosphorylase b. *J. Med. Chem.* **1994**, *37*(16), 2589−601.

(37) Kashaw, S. K.; Rathi, L.; Mishra, P.; Saxena, A. K. Development of 3D-QSAR models in cyclic ureidobenzenesulfonamides: human 3-Adrenergic receptor agonist. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 2481−2484.

(38) Saxena, A. K.; Ram, S.; Saxena, M.; Singh, N.; Prathipati, P.; Jain, P. C.; Singh, H. K.; Anand, N. QSAR studies in substituted 1,2,3,4,6,7,12,12a-octa-hydropyrazino[2′,1′:6,1]pyrido[3,4-b]indoles− a potent class of neuroleptics. *Bioorg. Med. Chem.* **2003**, *11*, 2085−2090.

(39) Gupta, M. K.; Mishra, P.; Prathipati, P.; Saxena, A. K. 2D-QSAR in hydroxamic acid derivatives as peptide deformylase inhibitors and antibacterial agents. *Bioorg. Med. Chem.* **2002**, *10*, 3713−3716.

(40) Babu, M. K.; Shakya, N.; Prathipati, P.; Kaskhedikar, S. G.; Saxena, A. K. Development of 3D-QSAR models for 5-Lipoxygenase antagonists: chalcones. *Bioorg. Med. Chem.* **2002**, *10*, 4035−4041.

(41) Desai, B.; Sureja, D.; Naliapara, Y.; Shah, A.; Saxena, A. K. Synthesis and QSAR Studies of 4-Substituted phenyl-2,6-dimethyl-3,5-bis-N-(substituted phenyl) carbamoyl-1,4-dihydropyridines as potential antitubercular agents. *Bioorg. Med. Chem.* **2001**, *9*, 1993−1998.

(42) Saxena, A. K.; Pandey, S. K.; Seth, P.; Singh, M. P.; Dikshit, M.; Carpy. A. Synthesis and QSAR Studies in 2-(N-aryl-N-aroyl) amino-4,5-dihydrothiazole Derivatives as Potential Antithrombotic Agents. *Bioorg. Med. Chem.* **2001**, *9*, 2025−2034.

(43) Saxena, A. K.; Pandey, S. K.; Tripathi, R. C.; Raghubir, R. Synthesis, molecular modeling and QSAR studies in chiral 2,3-disubstituted-1,2,3,4-tetrahydro-9H-pyrido(3,4-b)indoles as potential modulators of opioid antinociception. *Bioorg. Med. Chem.* **2001**, *9*, 1559−1570.

(44) Pandya, T.; Pandey, S. K.; Tiwari, M.; Chaturvedi, S. C.; Saxena, A. K. 3-D QSAR studies of triazolinone based balanced AT1/AT2 receptor antagonists. *Bioorg. Med. Chem.* **2001**, *9*, 291−300.

(45) Rathi, L.; Kashaw, S. K.; Dixit, A.; Pandey, G.; Saxena, A. K. Pharmacophore identification and 3D-QSAR studies in *N*-(2-benzoyl phenyl)-L-tyrosines as PPARγ agonists. *Bioorg. Med. Chem.* **2004**, *12*, 63−69.

(46) Tame J. R. H. Scoring functions: A view from the bench. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 99−108.

(47) Donini O. A. T.; Kollman P. A. Calculation and prediction of binding free energies for the matrix metalloproteinases. *J. Med. Chem.* **2000**, *43*, 4180−4188.

(48) Cramer, R. D.; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959−5967.

(49) *3D QSAR in Drug Design, Theory, Methods, and Applications*; Kubyini, H., Ed.; ESCOM Science Publishers B.V.: Leiden, 1993.

(50) Norinder, U. The Alignment problem in 3D-QSAR: A combined approach using Catalyst and a 3D-QSAR technique. *QSAR and Molecular Modelling: Concepts, Computational Tools and Biological*

GLYCOGEN PHOSPHORYLASE *A* INHIBITORS

J. Chem. Inf. Model., Vol. 45, No. 1, 2005 **145**

Applications; Sanz, F., Giraldo, J., Manaut, F., Eds.; Prous Science Publishers: Barcelona, 1995; pp 433−438.

(51) Hoffmann, R. D.; Langer, T. Use of the Catalyst program as a new alignment tool for 3D-QSAR. *QSAR and Molecular Modelling: Concepts, Computational Tools and Biological Applications*; Sanz, F., Giraldo, J., Manaut, F., Eds.; Prous Science Publishers: Barcelona, 1995; pp 466−469.

(52) Semus, S. F. CoMFA: A field of dreams? http://www.netsci.org/Science/Compchem/feature11.html.

(53) Langer, T.; Hoffmann, R. D. On the Use of Chemical Function-Based Alignments as Input for 3D-QSAR. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 325−330.

(54) Bureau, R.; Daveu, C.; Baglin, I.; Sopkova-De Oliveira Santos, J.; Lancelot, J. C.; Rault, S. Association of Two 3D QSAR Analyses. Application to the Study of Partial Agonist Serotonin-3 Ligands. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 815−823.

(55) Zhu, L. L.; Hou, T. J.; Chen, L. R.; Xu, X. J. 3D QSAR Analyses of Novel Tyrosine Kinase Inhibitors Based on Pharmacophore Alignment. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1032−1040.

(56) Klebe, G.; Abraham, U.; Mietzner, T. Molecular Similarity Indices in a Comparative Analysis (CoMSIA) of Drug Molecules to Correlate and Predict Their Biological Activity. *J. Med. Chem.* **1994**, *37*, 4130−4146.

(57) Viswanadhan, V. N.; Ghose, A. K.; Revankar, G. R.; Robins, R. K. Atomic Physicochemical Parameters for Three-Dimensional Structure Directed Quantitative Structure−Activity Relationships. 4. Additional Parameters for Hydrophobic and Dispersive Interactions and Their Application for an Automated Superposition of Certain Naturally Occurring Nucleoside Antibiotics. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 163−172.

(58) Joseph-McCarthy, D.; Thomas, B. E., IV; Belmarsh, M.; Moustakas, D.; Alvarez, J. C. Pharmacophore-Based Molecular Docking to Account for Ligand Flexibility. *Proteins: Struct., Funct., Genet.* **2003**, *51*, 172−188.

(59) (a) Ludi manual; version 2.3; December 1993. (b) Bohm, H. J. LUDI: rule-based automatic design of new substituents for enzyme inhibitor leads. *J. Comput.-Aided Mol. Des.* **1992**, *6*, 593−606.

(60) Wang, R.; Lu, Y.; Wang, S. Comparative evaluation of 11 scoring functions for molecular docking. *J. Med. Chem.* **2003**, *46*, 2287−303.

(61) Catalyst tutorial; version 4.5; August 1999.

(62) Hirashim, A.; Morimoto, M.; Kuwano, E.; Taniguchi, E.; Eto, M. Three-dimensional common-feature hypotheses for octopamine agonist arylethanolamines. *J. Mol. Graphics Modell.* **2002**, *21*, 81−7.

(63) Golbraikh, A.; Tropsha, A. Beware of q2. *J. Mol. Graphics Modell.* **2002**, *20*, 269−276.

(64) Wold, S.; Martens, H.; Wold, H. 1983. The Multivariate Calibration Problem in Chemistry solved by the PLS Method. Proc. Conf. Matrix Pencils, *Lecture Notes in Mathematics 973*; Ruhe, A., Kågström, B., Eds.; Springer-Verlag: Heidelberg, 1982; pp 286−293.

(65) Wold, S. Cross validatory estimation of the number of components in factor and principal components models. *Technometrics* **1978**, *20*, p 397.

(66) Saxena A. K.; Prathipati P. Comparison of MLR, PLS and GA-MLR in QSAR analysis. *SAR QSAR Environ. Res.* **2003**, *14*, 433−446.

(67) Systat; version 7.0.

(68) Sybyl getting started manual; version 6.8; October 2001.

(69) Bohacek, R. S.; McMartin, C. Definition and Display of Steric, Hydrophobic, and Hydrogen-Bonding Properties of Ligand Binding Sites in Proteins Using Lee and Richards Accessible Surfaces: Validation of a High-Resolution Tool for Drug Design. *J. Med. Chem.* **1992**, *35*, 1671−1684.

(70) Kroemer, R. T.; Hecht, P. Replacement of steric 6−12 potential-derived interaction energies by atom-based indicator variables in CoMFA leads to models of higher consistency. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 205−212.