Pharmacophore Modeling, Quantitative Structure–Activity Relationship Analysis, and in Silico Screening Reveal Potent Glycogen Synthase Kinase-3β Inhibitory Activities for Cimetidine, Hydroxychloroquine, and Gemifloxacin

Mutasem O. Taha,*[,†] Yasser Bustanji,[‡] Mohamed A. S. Al-Ghussein,[†] Mohammad Mohammad,[†] Hiba Zalloum,[†] Ihab M. Al-Masri,[†] and Naji Atallah[†]

Department of Pharmaceutical Sciences and Department of Biopharmaceutics and Clinical Pharmacy, Faculty of Pharmacy, University of Jordan, Amman, Jordan

Received August 6, 2007

The pharmacophoric space of glycogen synthase kinase-3β (GSK-3β) was explored using two diverse sets of inhibitors. Subsequently, genetic algorithm and multiple linear regression analysis were employed to select optimal combination of pharmacophores and physicochemical descriptors that access self-consistent and predictive quantitative structure–activity relationship (QSAR) against 132 training compounds (r^2_{123} = 0.663, F = 24.6, r^2_{LOO} = 0.592, r^2_{PRESS} against 29 external test inhibitors = 0.695). Two orthogonal pharmacophores emerged in the QSAR, suggesting the existence of at least two distinct binding modes accessible to ligands within GSK-3β binding pocket. The validity of the QSAR equation and the associated pharmacophores was established by the identification of three nanomolar GSK-3β inhibitors retrieved from our in-house-built structural database of established drugs, namely, hydroxychloroquine, cimetidine, and gemifloxacin. Docking studies supported the binding modes suggested by the pharmacophore/QSAR analysis. In addition to being excellent leads for subsequent optimization, the anti-GSK-3β activities of these drugs should have significant clinical implications.

1. Introduction

In recent years there has been a dramatic increase in the global prevalence of diabetes. Incidence is steadily increasing, creating a heavy burden on health care systems and a need for new methods of treatment.[1,2] Diabetes mellitus can result in frequent and serious complications including macrovascular and microvascular complications. Insulin independent diabetes mellitus (type 2 diabetes) accounts for more than 90% of diabetic cases.[2] Resistance to the biological actions of insulin in tissues like muscle, liver, and adipocytes is a major feature of the pathophysiology in type 2 diabetes.[3]

The central role of glycogen synthase kinase-3 (GSK-3[a]) in glucose metabolism makes it an exciting target for controlling hyperglycemia.[4] GSK-3 is a cytosolic serine/threonine protein kinase and one of many signaling components downstream from the insulin receptor.[5–9] It is constitutively active in resting cells and inhibited through the action of extracellular signals including insulin.[4] Insulin activates glycogen synthase (GS) via increasing protein phosphatase 1G activity and inhibiting protein kinases such as GSK-3.[3] Moreover, GSK-3 kinases have a direct effect on glucose transport by inactivating kinesin, which is involved in regulating the trafficking of glucose transport protein-4 (GLUT-4) vesicles to the plasma membrane.[10,11]

The negative regulation of GSK3 by insulin occurs at key cell types important for glycogen metabolism: hepatocytes, myocytes, and adipocytes.[12–14] Overexpression of GSK3 in these cells could lead to insulin resistance and contribute to the pathology of diabetes.[15,16] Therefore, inhibiting GSK3 in concert with insulin-induced signaling should increase the activity of GS and improve glycogen deposition in critical glucose-controlling tissues.[4]

Moreover, GSK-3 activity has been connected to obesity,[17] Alzheimer disease,[18,33] Down syndrome,[19] and colorectal cancer.[20,21] Interestingly, a homologous GSK-3 is expressed in *Plasmodium falciparum* and was evaluated as a potential screening target for the identification of antimalarial agents.[22]

In humans, GSK-3 exists in two forms: GSK-3α (51 kDa) and GSK-3β (47 kDa). Both isoforms have nearly identical biochemical functions and substrate affinities.[5,23–27] GSK-3 inhibitors have arisen as promising drugs for several diseases such as stroke, cancer, mood disorders, inflammationm and Alzheimer's disease.[5,28] However, inhibitors of GSK-3β have particularly useful antidiabetic properties because they improve insulin sensitivity, glycogen synthesis, and glucose metabolism in skeletal muscles of diabetic patients.[6,8,29] Currently, there are several GSK-3β small molecule inhibitors in clinical trials for the treatment of type II diabetes.[5–8,28–32] The vast majority of reported GSK-3β inhibitors compete with ATP for binding with GSK-3.[33–35]

The main focus of current efforts toward the development of new GSK-3β inhibitors concentrate on structure-based ligand design.[36] To date, 13 moderately resolved (resolution of 1.95–2.8 Å) GSK-3β X-ray complexes are documented in the Protein Data Bank (1GNG, 1H8F, 1I09, 1J1B, 1J1C, 1O9U, 1PYX, 1Q3D, 1Q3W, 1Q41, 1Q4L, 1Q5K, and 1UV5).[37–43] Nevertheless, although considered the most reliable structural information that can be used for drug design, crystallographic structures are limited by inadequate resolution[44] and crystallization-related artifacts of the ligand−protein complex.[45] Furthermore, crystallographic structures generally ignore structural heterogeneity related to protein anisotropic motion and discrete conformational

* To whom correspondence should be addressed. Telephone: 00962 6 5355000, extension 23305. Fax: 00962 6 5339649. E-mail: mutasem@ju.edu.jo.

† Department of Pharmaceutical Sciences.

‡ Department of Biopharmaceutics and Clinical Pharmacy.

[a] Abbreviations: GSK-3, glycogen synthase kinase-3; QSAR, quantitative structure–activity relationship; MLR, multiple linear regression; GFA:, genetic function algorithm; HBA, hydrogen bond acceptor; HBD, hydrogen bond donor.

10.1021/jm7009765 CCC: $40.75
Published on Web 03/07/2008

substates.[46] The latter factor is particularly significant in the case of GSK-3β because it was recently reported to exhibit significant induced fit flexibility.[47] Collectively, these factors may have limited the success in developing new GSK-3β inhibitors through structure-based design efforts.

The recent interest in designing new hypoglycemic agents based on GSK-3β inhibition, combined with the drawbacks of structure-based design, prompted us to explore the possibility of developing ligand-based three-dimensional (3D) pharmacophore(s), integrated within self-consistent QSAR model, for GSK-3β inhibitors. This approach avoids the pitfalls of structure-based techniques; furthermore, the pharmacophore model(s) can be used as 3D search query(ies) to mine 3D libraries for new GSK-3β inhibitors.

We employed the HYPOGEN module from the CATALYST software package[48–50] to construct plausible binding hypotheses for GSK-3β inhibitors. Subsequently, genetic function algorithm (GFA) and multiple linear regression (MLR) analyses were combined to search for optimal QSAR that merge high-quality binding pharmacophores with other molecular descriptors and are capable of explaining bioactivity variation across a collection of diverse GSK-3β inhibitors. The optimal pharmacophores were subsequently used as 3D search queries to virtually screen our in-house-built database of established drug molecules to identify GSK-3β inhibitory hits from drug space. Such hits are expected to be excellent leads for subsequent optimization because they have established toxicological and SAR profiles.

CATALYST models drug−receptor interaction using information derived only from the drug structure.[50–52] Molecules are described as a collection of chemical functionalities arranged in 3D space. The conformational flexibility of training ligands is modeled by creating multiple conformers, judiciously prepared to emphasize representative coverage over a specified energy range. It identifies a set of chemical features common to active training molecules. This 3D array of chemical features provides a relative alignment for each input molecule consistent with their binding to a proposed common receptor site. The chemical features considered can be hydrogen bond donors and acceptors (HBDs and HBAs), aliphatic and aromatic hydrophobes, positive and negative charges, positive and negative ionizable groups, and aromatic planes (RingArom). Successful examples involving the use of CATALYST have been reported, wherein the CATALYST derived pharmacophore has been used efficiently as a query for database searching and in 3D-QSAR studies.[53–58]

2. Results and Discussion

2.1. Data Mining and Conformational Coverage. The literature was extensively surveyed to collect a large group of diverse GSK-3β inhibitors (**1**–**152**; see Table 1 and Figure 1).[59–62] Their 2D structures were imported into CATALYST and converted automatically into plausible 3D single conformer representations, which were then used as starting points for conformational analyses and in the determination of various molecular descriptors for QSAR modeling.

The conformational space of each inhibitor was extensively sampled utilizing the poling algorithm employed within the CONFIRM module of CATALYST. Poling promotes conformational variation via employing molecular mechanical force field algorithm that penalizes similar conformers.[63] Conformational coverage was performed employing the "Best" module to ensure extensive sampling of conformational space (see section 4.1.3 under Experimental Section). Efficient conformational coverage guarantees minimum conformation-related noise during pharmacophore generation and validation stages. Pharmacophore generation and pharmacophore-based search procedures are known for their sensitivity to inadequate conformational sampling within the training compounds.[64]

2.2. Exploration of GSK-3β Pharmacophoric Space. CATALYST-HYPOGEN enables automatic pharmacophore construction by using a collection of at least 16 molecules with bioactivities spanning over 4 orders of magnitude.[49,50,63,65–67] Therefore, since we have an informative list of 152 GSK-3β inhibitors of evenly spread bioactivities over more than 4 orders of magnitude, we employed HYPOGEN to explore possible pharmacophoric binding modes assumed by different inhibitors within GSK-3β binding pocket. HYPOGEN implements an optimization algorithm that evaluates large number of potential models within the pharmacophoric space of a particular target (see generation of pharmacophoric hypotheses section in Experimental Section).[66] The extent of the evaluated space is indicated by the configuration (Config) cost of each modeling run. It is generally recommended that the Config cost of any HYPOGEN run not exceed 17 (corresponding to 2^{17} hypotheses to be assessed by CATALYST); otherwise, a thorough analysis of all models cannot be guaranteed.[63] The size of the investigated pharmacophoric space is a function of training compounds, selected input chemical features, and other CATALYST control parameters such as feature tolerances and weights.[63] Restricting the pharmacophoric space might improve the efficiency of optimization by allowing effective evaluation of limited number of pharmacophoric models.[54] On the other hand, severe restrictions imposed on the pharmacophoric space might reduce the possibility of discovering optimal pharmacophoric hypotheses, particularly if they occur outside the "boundaries" of the pharmacophoric space.[54]

Therefore, we decided to explore the pharmacophoric space of GSK-3β inhibitors under reasonably imposed "boundaries" via six HYPOGEN runs and two carefully selected training subsets from the collected compounds: subsets A and B in Tables 2 and 3. The training compounds in both subsets were selected in such away to guarantee maximal 3D diversity and continuous bioactivity spread over four logarithmic cycles. Furthermore, the training inhibitors were selected in such a way that differences in their anti-GSK-3β bioactivities are primarily attributable to the presence or absence of pharmacophoric features (e.g., HBA or HBD or Hydrophobic) rather than steric shielding and/or bioactivity-enhancing or -reducing auxiliary groups (e.g., electron donating or withdrawing groups). We gave special emphasis to the 3D diversity of the most active compounds in each training subset (Table 2) because they have significant influence on the extent of the evaluated pharmacophore space through the Constructive Phase of HYPOGEN algorithm (see generation of pharmacophoric hypotheses section under Experimental Section).[66] We restricted the software to explore pharmacophoric models incorporating from zero to three features of any particular selected feature type (i.e., HBA, HBD, Hydrophobic, and RingArom), i.e., instead of the default range of zero to five, as shown in Table 3. Furthermore, we instructed HYPOGEN to explore only four- and five-featured pharmacophores, i.e., ignore models of lesser number of features (Table 3). The later restriction has the dual benefit of narrowing the investigated pharmacophoric space and best-representing the feature-rich nature of GSK-3β binding pocket. CATALYST can produce pharmacophore hypotheses of a maximum of five features.[49–51]

The input pharmacophore features were selected in agreement with published SAR studies and crystallographic data (Table 3). For example, the fact that GSK-3β crystallographic informa-

Table 1. Structures of GSK-3b Inhibitors Utilized in Modeling[a]

compd	R1	R2	R3	X	Y	W	Z	IC_{50} (nM)
1	H	H	H					590
2	H	3-chloro	H					301
3	H	3-hydroxy	H					704
4[b]	H	3,5-dichloro-4-hydroxy	H					149
5	H	3-carboxy	H					291
6	H	4-chloro-3-carboxy	H					143
7	H	4-SCH_3	H					404
8	2-chloro	H	H					216
9[b]	2-chloro	3-chloro	H					195
10	2-chloro	3-hydroxy	H					374
11	2-chloro	3-chloro-4-hydroxy	H					152
12	2-chloro	3,5-dichloro-4-hydroxy	H					93
13	2-chloro	3-carboxy	H					136
14	2-chloro	4-chloro-3-carboxy	H					74
15	2-chloro	4-S-CH_3	H					161
16[b]	2-methoxy	H	H					216
17	2-methoxy	3-chloro	H					114
18	2-methoxy	3-hydroxy	H					259
19[b]	2-methoxy	3-chloro-4-hydroxy	H					139
20	2-methoxy	3,5-dichloro-4-hydroxy	H					82
21	2-methoxy	4-S-CH_3	H					110
22[b]	2-nitro	3-chloro	H					104
23	2-nitro	3-hydroxy	H					251
24	2-nitro	3-chloro-4-hydroxy	H					104
25	2-nitro	3,5-dichloro-4-hydroxy	H					52
26	2-nitro	4-chloro-3-carboxy	H					28
27	3-chloro	3-hydroxy	H					1478
28	3-chloro	3-chloro-4-hydroxy	H					94
29	3-chloro	3,5-dichloro-4-hydroxy	H					58
30	3-chloro	3-carboxy	H					134
31	3-chloro	4-chloro-3-carboxy	H					76
32	3-chloro	4-S-CH_3	H					532
33	3-methoxy	3-chloro	H					257
34[b]	3-methoxy	3-hydroxy	H					472
35	3-methoxy	3,5-dichloro-4-hydroxy	H					142
36	3-methoxy	3-carboxy	H					195
37	3-methoxy	4-chloro-3-carboxy	H					85
38	3-methoxy	4-S-CH_3	H					203
39	3-nitro	H	H					141
40[b]	3-nitro	3-chloro	H					70
41	3-nitro	3-hydroxy	H					236
42[b]	3-nitro	4-hydroxy	H					123
43	3-nitro	3-chloro-4-hydroxy	H					59
44	3-nitro	3,5-dichloro-4-hydroxy	H					20
45[b]	3-nitro	3-carboxy	H					79
46	3-nitro	4-chloro-3-carboxy	H					26
47	3-nitro	4-S-CH_3	H					152
48	4-chloro	H	H					514
49	4-chloro	3-chloro	H					447
50	4-chloro	3-hydroxy	H					407
51	4-chloro	4-hydroxy	H					317
52	4-chloro	3-chloro-4-hydroxy	H					173
53[b]	4-chloro	3,5-dichloro-4-hydroxy	H					91
54	4-chloro	3-carboxy	H					186
55	4-chloro	4-chloro-3-carboxy	H					109
56[b]	4-chloro	4-S-CH_3	H					529
57[b]	4-methoxy	H	H					390
58	4-methoxy	3-chloro	H					156
59	4-methoxy	3-hydroxy	H					481
60	4-methoxy	3,5-dichloro-4-hydroxy	H					83
61	4-methoxy	3-carboxy	H					214
62[b]	4-methoxy	4- S-CH_3	H					243
63	4-nitro	3,5-dichloro-4-hydroxy	H					71
64	4-nitro	4- S-CH_3	H					392
65	H	H	methyl					2613
66	3-nitro	H	methyl					1398
67[b]	4-chloro	H	methyl					2285
68	2-chloro							337
69	2-methoxy							187
70	2-nitro							131
71	3-chloro							460
72[b]	3-nitro							161
73	4-chloro							1412
74[b]	4-methoxy							694
75		phenyl	*n*-propyl	CH	CH		CH	99
76		phenyl	*n*-propyl	N	CH		CH	7
77		phenyl	*n*-propyl	N	CH		N	2697
78		phenyl	*n*-propyl	N	N		C-phenyl	691
79[b]		phenyl	$(CH_2)_3N(CH_3)_2$	N	N		CH	22
80		phenyl	$(CH_2)_3$pyrrolidine	N	N		CH	11

Table 1. Continued

compd	R1	R2	R3	X	Y	W	Z	IC_{50} (nM)
81		phenyl	$(CH_2)_3$-piperazinyl-N-(C_2H_5)	N	N		CH	7
82		phenyl	$(CH_2)_3$-morpholinyl	N	N		CH	5
83		phenyl	4-piperidine-N- CH_3	N	N		CH	9
84		phenyl	$(CH_2)_4$-piperazinyl-N-(C_2H_5)	N	N		CH	5
85		2,2-difluoro-phenyl	$(CH_2)_3N(CH_3)_2$	N	N		CH	5
86	phenyl	bromo	cyclopropyl	CH	N		CH	75
87	4-OH-phenyl	bromo	cyclopropyl	CH	N		CH	0.8
88	4-OH-phenyl	H	cyclopropyl	CH	N		CH	8
89[b]	3-bromo-4-OH-phenyl	H	cyclopropyl	CH	N		CH	5
90[b]	3-chloro-4-OH-phenyl	H	cyclopropyl	CH	N		CH	7
91	4-OH- phenyl	phenyl	cyclopropyl	CH	N		CH	24
92	4-OH- phenyl	bromo	$(CH_2)_3$-4-piperazinyl-N-(C_2H_5)	CH	N		CH	4
93	3-OH- phenyl	H	cyclopropyl	CH	N		CH	12
94	2-thienyl	bromo	cyclopropyl	CH	N		CH	39
95	2-thienyl	bromo	cyclopentyl	CH	N		CH	7
96	2-furyl	H	cyclopropyl	CH	N		CH	141
97	2-furyl	bromo	cyclopropyl	CH	N		CH	7
98	2-thiazoyl	bromo	cyclopropyl	CH	N		CH	99
99	2-thiazoyl	bromo	cyclopentyl	CH	N		CH	16
100	2-thienyl	bromo	CH_2-4-piperidine-N-CH_3	CH	N		CH	18
101	2-furyl	bromo	3- pyrrolidine-N-benzyl	CH	N		CH	14
102	H	phenyl	methyl	CH	N		CH	291
103	H	phenyl	ethyl	CH	N		CH	43
104	H	phenyl	*n*-propyl	CH	N		CH	56
105	H	phenyl	isopropyl	CH	N		CH	19
106	H	phenyl	cyclopentyl	CH	N		CH	5
107	H	phenyl	$NH(C_2H_5)$	CH	N		CH	2810
108	H	B(OC$(CH_3)_2$-C$(CH_3)_2$O)	*n*-propyl	CH	N		CH	356
109	H	H	*n*-propyl	CH	N		CH	2343
110[b]	H	2-fluoro-phenyl	*n*-propyl	CH	N		CH	18
111	H	3- fluoro-phenyl	*n*-propyl	CH	N		CH	20
112	H	2,3-difluoro-phenyl	*n*-propyl	CH	N		CH	7
113	H	2-chloro-phenyl	*n*-propyl	CH	N		CH	27
114	H	3-pyridyl	*n*-propyl	CH	N		CH	11
115	H	4-pyridyl	*n*-propyl	CH	N		CH	443
116	H	4-bisphenyl	*n*-propyl	CH	N		CH	851
117	H	2-naphthyl	*n*-propyl	CH	N		CH	169
118	H	1-naphthyl	*n*-propyl	CH	N		CH	241
119[b]	phenyl	H	cyclopropyl	CH	N		CH	425
120[b]	3,4-di-OH-phenyl	H	cyclopropyl	CH	N		CH	8
121	3-O(CH_3)-phenyl	H	cyclopropyl	CH	N		CH	125
122	2-OH-phenyl	H	cyclopropyl	CH	N		CH	36
123	2-O(CH_3)-phenyl	H	cyclopropyl	CH	N		CH	1593
124	4-OH-phenyl	chloro	cyclopropyl	CH	N		CH	1
125	4-OH-phenyl	methyl	cyclopropyl	CH	N		CH	6
126	phenyl	phenyl	cyclopropyl	CH	N		CH	415
127	phenyl	chloro	cyclopropyl	CH	N		CH	234
128	phenyl	CN	cyclopropyl	CH	N		CH	87
129[b]	phenyl	bromo	4-piperidine-N-(CH_3)	CH	N		CH	383
130[b]	4-OH- phenyl	H	4-piperidine-N-(CH_3)	CH	N		CH	12
131	4-OH- phenyl	bromo	4-piperidine-N-(CH_3)	CH	N		CH	1
132	3-OH- phenyl	H	(CH_2)-3-piperazinyl-N-(C_2H_5)	CH	N		CH	21
133	4-O(CH_3)-phenyl	H	cyclopropyl	CH	N		CH	23000
134	phenyl		cyclopropyl	CH	CH		CH	498
135	4-OH-phenyl		cyclopropyl	CH	CH		CH	15
136	5-indolyl		cyclopropyl	CH	CH		CH	42
137	phenyl-3-SO_2NH_2		cyclopropyl	CH	CH		CH	481
138	3-fluoro- phenyl		cyclopropyl	CH	CH		CH	828
139[b]	2-pyrrolyl		cyclopropyl	CH	CH		CH	320
140[b]	3-furanyl		cyclopropyl	CH	CH		CH	35
141	2-thienyl		cyclopropyl	CH	CH		CH	215
142	3-thienyl		cyclopropyl	CH	CH		CH	329
143	2,5-di fluoro-phenyl		cyclopropyl	CH	CH		CH	1000
144	phenyl	phenyl	NH_2	N	N	NH	N	250
145	phenyl	H	NH_2	N	N	NH	N	530
146	phenyl	H	NH_2	CH	N	NH	N	430
147[b]	phenyl	H	NH_2	N	CH	NH	N	1260
148	H	phenyl	NH_2	CH	N	NH	N	23000
149[b]	phenyl		$NHSO_2(CH_3)$	CH	N	NH	N	3572
150	phenyl		NH_2	CH	N	N(CH_3)	N	23000
151	phenyl		NH_2	CH	N	O	N	23000
152[b]	phenyl		NH_2	CH	N	NH	CH	23000

[a] The corresponding scaffolds are in Figure 1. [b] These compounds were employed as the external testing subset in QSAR modeling.

tion suggested the involvement of Val135 in hydrogen bonding with some ligands prompted us to select hydrogen bond donor and acceptor (HBA and HBD) functionalities as possible pharmacophoric features. In the same manner, aromatic rings (RingArom) and hydrophobic groups were fed into HYPOGEN as possible pharmacophoric features, as shown in Table 3.

Each HYPOGEN run generated 10 optimal pharmacophoric hypotheses, yielding 60 models from 6 automatic runs (Tables

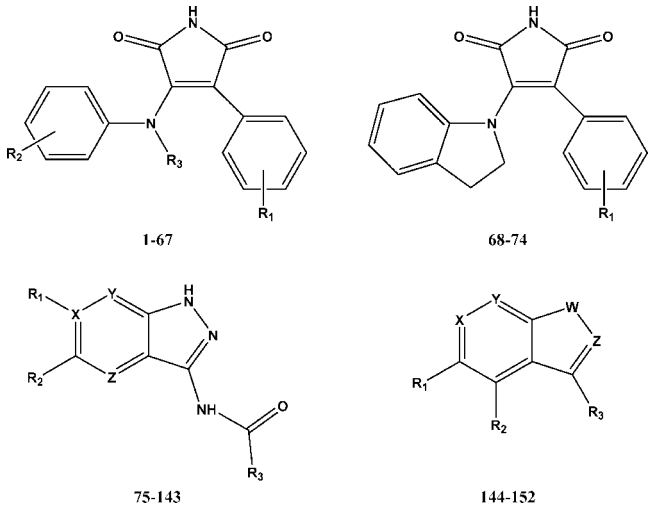

Figure 1. Chemical scaffolds of training compounds. The detailed structures are as in Table 1.

3 and 4). The binding hypotheses from each run were automatically ranked according to their corresponding "total cost" values. Total cost is defined as the sum of error cost; weight cost and configuration cost (see pharmacophore validation section under Experimental Section). Error cost provides the highest contribution to total cost, and it is directly related to the intrinsic capacity of the particular hypothesis in correlating the molecular structures to their corresponding biological responses.[49,63,65–67] HYPOGEN also calculates the cost of the null hypothesis, which presumes that there is no relationship in the data and that experimental activities are normally distributed about their mean. Accordingly, the greater the difference from the null hypothesis cost (residual cost, Table 4), the more likely that the hypothesis does not reflect a chance correlation, i.e., statistically signficant.[49,63,65–67]

An additional validation technique based on Fischer's randomization test was recently introduced into CATALYST, i.e., Cat.Scramble.[49,68] In this test the biological data and the corresponding structures are scrambled several times, and the software is challenged to generate pharmacophoric models from the randomized data. The confidence in the parent hypotheses (i.e., generated from unscrambled data) is lowered proportional to the number of times the software succeeds in generating binding hypotheses from scrambled data of apparently better cost criteria than the parent hypotheses. This approach minimizes the possibility of adopting fortuitous pharmacophores, which might happen because of the vast number of potential pharmacophores that CATALYST evaluates during a particular modeling cycle (up to 2^{17}; see above).

Table 4 shows the different success criteria of the generated pharmacophore hypotheses. Clearly from Table 4, all models shared comparable features and good overall success criteria; in particular they illustrated satisfactory Cat.Scramble confidence levels (95%). The emergence of numerous high-quality pharmacophore models is probably related to the ability of GSK-3β ligands to assume multiple pharmacophoric binding modes within the binding pocket. Therefore, it is quite challenging to select any particular pharmacophore hypothesis as a sole representative of the binding process.

2.3. QSAR Modeling. Although pharmacophore models provide great insights into ligand−macromolecule interactions, their predictive value as 3D-QSAR models is limited by steric shielding and bioactivity-enhancing or -reducing auxiliary groups.[69] This point combined with the fact that modeling GSK-3β inhibitors furnished numerous binding hypotheses of excellent statistical criteria (Table 4) prompted us to employ classical QSAR analysis to search for the best combination of pharmacophore model(s) and other structural descriptors capable of explaining bioactivity variation across the whole list of inhibitors (**1–152**, Table 1 and Figure 1). This task was performed via a combination of genetic function algorithm and multiple linear regression QSAR analysis (GFA-MLR-QSAR, see QSAR Modeling under Experimental Section).

Several descriptor categories were calculated for the collected compounds, including simple and valence connectivity indices, electrotopological state indices, κ shape indices, Shadow indices, JURS descriptors, semiempirical quantum mechanical energy terms (HOMO and LUMO), logarithm of partition coefficient, polarizability, dipole moment, molecular volume, molecular weight, and molecular surface area.[70] Furthermore, the training compounds were fitted against the generated pharmacophores and their fit values were added as additional descriptors. Pharmacophore fit values can be obtained automatically via eq 5 (see Assessment of the Generated Hypotheses under Experimental Section).[49]

GFA-MLR-QSAR selects optimal descriptor combinations based on the Darwinian concept of genetic evolution, via mating and mutation, whereby the statistical criteria of regression models from different descriptor combinations (chromosomes) are employed as fitness criteria.[70]

However, to avoid overloading GFA-MLR with independent variables, which may allow the emergence of less-than-optimal regression models,[70] we clustered the generated pharmacophores (60 models) into 10 groups and considered the best representatives only (i.e., based on their cost criteria; see Clustering of the Generated Pharmacophore Hypotheses under Experimental Section) for GFA-MLR-QSAR analysis (10 bolded models in Table 4). Accordingly, the fit values obtained by mapping the 10 representative pharmacophore models against 123 GSK-3β inhibitors (Table 1 and Figure 1) were enrolled as independent variables (genes) in a cycle of GFA-MLR-QSAR analysis over 30 000 iterations.[70,71] However, since it is essential to access the predictive power of the resulting QSAR models on an external set of inhibitors, we randomly selected 29 molecules (marked with asterisks in Table 1; see QSAR Modeling under Experimental Section) and employed them as external test molecules for validating the QSAR models (r^2_{PRESS}). Moreover, the models were cross-validated automatically using the leave-one-out cross-validation.[70,71]

Table 5 shows the statistical criteria of the top-ranking QSAR models. Model B (Table 5) exhibited the best statistical qualities. Interestingly, all optimal QSAR models required at least some of their explanatory variables to be transformed into spline forms (see below).[70] Furthermore, they shared most of their descriptors, highlighting their significance as well as the significance of their corresponding descriptors. Equation 1 shows the optimal QSAR model, and Figure 2 shows the scatter plots of experimental versus estimated bioactivities for the training and testing inhibitors subsets.

$$\log(1/\mathrm{IC}_{50}) = -5.52 + 0.21[\mathrm{Hypo4/5} - 4.40] + 0.45[\mathrm{Hypo6/3} - 5.67] - 1.31[\mathrm{S\_aasC} - 3.47] + 0.44\kappa_2 - 3.81[12.64 - \mathrm{S\_aaN}] - 1.31[5.65 - \kappa_2] - 2.14[\mathrm{Jurs\text{-}FPSA\text{-}2} - 1.13] + 3.69[12.90 - \mathrm{S\_aaN}] + 0.26[1.96 - \mathrm{S\_sCH3}] \quad (1)$$

$$r^2 = 0.663,\ F\text{-statistic} = 24.6,\ r^2_{\mathrm{BS}} = 0.662,\ r^2_{\mathrm{LOO}} = 0.592,\ r^2_{\mathrm{PRESS(29)}} = 0.695 \quad (1a)$$

Table 2. Training Subsets Employed in Exploring the Pharmacophoric Space of GSK-3β Inhibitors[a]

training set	most active[b]	moderately active	least active[b]
A	**44**, **46**, **88**, **91**, **94**, **95**, **101**, **140**	**1**, **10**, **12**, **14**, **22**, **27**, **49**, **73**, **75**, **78**, **86**, **109**, **116**, **121**, **134**, **142**, **145**−**147**, **149**	**148**, **151**, **152**
B	**87**, **92**, **95**, **125**	**1**, **2**, **26**, **43**, **73**, **88**, **109**, **118**, **119**, **126**, **129**, **140**, **144**, **146**	**107**, **148**, **151**, **152**

[a] Numbers correspond to compounds in Table 1 and Figure 1. [b] Potency categories as defined by eqs 2 and 3 in section 4.1.4, Exploration of the Pharmacophoric Space of GSK-3β, under Experimental Section.

Table 3. Training Sets and CATALYST Run Parameters Employed for Exploring GSK-3β Pharmacophoric Space

run number	training set[a]	no. of training compounds	selected input features: types and ranges[b]	interfeature spacing[c] (pm)
1	A	31	HBA (0–3), HBD (0–3), Hbic (0–3), RingArom (0–3)	5
2	A	31	HBA (0–3), HBD (0–3), Hbic (0–3), RingArom (0–3)	150
3	A	31	HBA (0–3), HBD (0–3), Hbic (0–3), RingArom (0–3)	300
4	B	22	HBA (0–3), HBD (0–3), Hbic (0–3), RingArom (0–3)	5
5	B	22	HBA (0–3), HBD (0–3), Hbic (0–3), RingArom (0–3)	150
6	B	22	HBA (0–3), HBD (0–3), Hbic (0–3), RingArom (0–3)	300

[a] The letters correspond to training sets in Table 2. [b] HBA: hydrogen bond acceptor. HBD: hydrogen bond doner. RingArom: aromatic ring. Hbic: hydrophobic. The allowed ranges of input features are in parentheses. [c] The number of output features was allowed to vary from 4 to 5 features. Other parameters were set to their default values.

where r_{123} is the correlation coefficient, r^2_{LOO} is the leave-one-out correlation coefficient, r^2_{BS} is the bootstrapping regression coefficient, and r^2_{PRESS} is the predictive r^2 determined for the 29 test compounds (via eq 6, see section QSAR Modeling under Experimental Section).[70,71] Hypo4/5 and Hypo6/3 are the fit values of the training compounds against the fifth and third pharmacophores from the fourth and sixth HYPOGEN runs, respectively (Tables 3 and 4), as calculated from eq 5 (see Assessment of the Generated Hypotheses under Experimental Section), κ_2 is the second order κ shape index, Jurs-FPSA-2 is the fraction of negatively charged solvent-accessible surface area (obtained by dividing the sum of the solvent-accessible surface areas of all negatively charged atoms by the total solvent accessible molecular surface area). S_sCH3, S_aaN, and S_aasC are the electrotopological sum descriptors for methyl, heteroaromatic nitrogens, and aromatic carbons, respectively.[70]

Several descriptors emerged in eq 1 in spline format. In fact, disallowing spline transformations of the explanatory terms caused at least 10% decline in the statistical criteria of the resulting models. The spline terms employed herein are "truncated power splines" and are denoted by bolded brackets ([ ]). For example, $[f(x) - a]$ equals zero if the value of $(f(x) - a)$ is negative; otherwise, it equals $(f(x) - a)$.[70]

Interestingly, the combination of Hypo4/5 and Hypo6/3 frequented in the highest ranking QSAR models (Table 5) suggesting they represent two complementary binding modes accessible to ligands within the binding pocket of GSK-3β. This proposition is further supported by the fact that the two pharmacophores are orthogonal (their cross-correlation $r^2 = 0.21$), indicating that one of the pharmacophores optimally explains the bioactivities of some training inhibitors while other inhibitors are more appropriately explained by the second pharmacophore. A similar conclusion was reached about the flexibility of factor Xa binding pocket based on the emergence of several orthogonal binding pharmacophores in the optimal QSAR model.[54]

Figures 3 and 4 show the two pharmacophores and how they map to inhibitors **87** ($IC_{50} = 0.80$ nM) and **44** ($IC_{50} = 20.0$ nM), while Table 6 shows the *X*, *Y*, and *Z* coordinates of the two pharmacophores.

Interestingly, both Hypo4/5 and Hypo6/3 emerged in eq 1 in spline format, indicating that each binding mode contributes to ligand/GSK-3β affinity only if the fit value of the particular ligand exceeds the corresponding spline threshold. For example, the ability of a certain ligand to map Hypo6/3 will impact its actual affinity to GSK-3β only if its fit value exceeds 5.67 (the spline intercept associated with this pharmacophore in eq 1). Since the two spline cutoffs (of both pharmacophores) resemble moderate overall ligand/pharmacophore mapping (the maximum value is 12.0; see section 4.1.5, Assessment of the Generated Hypotheses, in Experimental Section), it appears that ligand binding to GSK-3β is sensitive to moderate misalignments among the attracting moieties within the complex such that lowering the fits value below 4.40 and 5.67 for Hypo4/5 and Hypo6/3, respectively, nullifies any affinity gains from mapping the pharmacophores.

Emergence of topological and shape descriptors (i.e., $\kappa 2$, S_sCH3, S_aaN, and S_aasC) in eq 1 illustrate certain role played by the ligands' topology in the binding process. However, despite the predictive significance of these descriptors, their information content is quite obscure. Nevertheless, emergence of Jurs-FPSA-2 in eq 1, in association with a negative regression coefficient, suggests an inverse relationship between ligand/GSK-3β affinity and the ligands' negative charges. We believe this trend is explainable by the fact that ionic groups favor hydration instead of docking into the binding site.

2.4. In Silico Screening and Subsequent Experimental Evaluation. Pharmacophore models are suitable for screening 3D molecular databases because they can capture compounds that exhibit optimally oriented binding features complementary to a proposed binding site.[50–58] On the other hand, QSAR models can be employed to predict the bioactivities of the captured hits. Accordingly, the fact that Hypo6/3 is associated with a higher regression coefficient in QSAR eq 1 (i.e., compared to Hypo4/5) prompted us to use this pharmacophore as a 3D search query against our in-house-built 3D database of established drugs (1490 compounds). The search process captured 397 hit compounds. Hits are defined as those compounds that have their chemical moieties spatially overlap (map) with corresponding features in the pharmacophoric query. However, it is not always necessary for hits to optimally fit their corresponding pharmacophore query. A poor fit value for a hit molecule means that although the chemical functionalities of the compound overlap with the corresponding pharmacophoric features, the centers of its functional groups are displaced from the centers of the corresponding pharmacophoric features such that the term $\sum(\text{disp/tol})^2$ in eq 5 (section 4.1.5, Assessment of the Generated Hypotheses, in the Experimental Section) approaches 1.0 and the overall fit value approaches zero.[66,63,65]

Table 4. Performance of Different Pharmacophoric Hypotheses Generated for GSK-3β Inhibitors Employing Different Settings as in Table 3

hypotheses[b]	pharmacophoric features in generated hypotheses	r^2[c]	cost value		config[d]	total	cost of null hypothesis	residual cost[e]	Cat. Scramble (%)
			error	weight					
					Run[a] 1				
1[f]	**2 × HBA, HBD, Hbic**	**0.89**	**117.7**	**2.4**	**14.1**	**134.3**	**166.5**	**32.2**	**95**
2	2 × HBA, 2 × Hbic	0.85	121.3	1.5	14.1	136.9	166.5	29.6	95
3	HBA, HBD, Hbic, RingAro	0.84	123.1	1.6	14.1	138.8	166.5	27.7	95
4	HBA, HBD, Hbic, RingAro	0.83	123.3	1.5	14.1	138.9	166.5	27.6	95
5	HBA, HBD, Hbic, RingAro	0.83	123.9	1.5	14.1	139.6	166.5	26.9	95
6	HBA, 2 × HBD, Hbic	0.82	124.5	1.3	14.1	139.8	166.5	26.7	95
7	HBA, 2 × HBD, Hbic	0.82	125.2	1.2	14.1	140.4	166.5	26.1	95
8	HBA, HBD, Hbic, RingAro	0.82	125	1.4	14.1	140.5	166.5	26	95
9	2 × HBA, Hbic, RingAro	0.81	125.6	1.2	14.1	140.8	166.5	25.7	95
10	2 × HBA, HBD, Hbic	0.81	125.8	1.1	14.1	141	166.5	25.5	95
					Run[a] 2				
1	2 × HBA, 2 × Hbic	0.84	122.5	1.1	13.6	138.8	166.5	27.7	95
2	2 × HBA, HBD, Hbic	0.83	123.9	1.3	13.6	138.8	166.5	27.7	95
3	2 × HBA, HBD, Hbic	0.82	124.5	1.4	13.6	139.5	166.5	27	95
4	HBA, HBD, 2 × Hbic	0.83	124.1	1.9	13.6	139.7	166.5	26.8	95
5	2 × HBA, 2 × Hbic	0.83	123.6	2.8	13.6	140	166.5	26.5	95
6	HBA, 2 × HBD, Hbic	0.81	125.4	1.2	13.6	140.3	166.5	26.2	95
7	HBA, 2 × HBD, Hbic	0.81	125.5	1.2	13.6	140.4	166.5	26.1	95
8	2 × HBA, HBD, Hbic	0.81	125.6	1.9	13.6	141.1	166.5	25.4	95
9	2 × HBA, Hbic, RingAro	0.81	126.1	1.4	13.6	141.1	166.5	25.4	95
10	2 × HBA, HBD, Hbic	0.81	125.8	1.8	13.6	141.2	166.5	25.3	95
					Run[a] 3				
1	HBA, HBD, 2 × Hbic	0.84	122.8	1.2	12	135.9	166.5	30.6	95
2[f]	**HBA, 2 × Hbic, RingAro**	**0.84**	**122.9**	**1.1**	**12**	**136**	**166.5**	**30.5**	**95**
3	HBA, 2 × Hbic, RingAro	0.84	122.9	1.1	12	136	166.5	30.5	95
4	HBA, HBD, 2 × Hbic	0.84	122.7	1.8	12	136.5	166.5	30	95
5	HBA, HBD, 2 × Hbic	0.83	124.1	1.4	12	137.6	166.5	28.9	95
6	HBA, HBD, 2 × Hbic	0.81	125.6	1.3	12	138.9	166.5	27.6	95
7[f]	**HBD, 2 × Hbic, RingAro**	**0.81**	**125.5**	**1.8**	**12**	**139.4**	**166.5**	**27.1**	**95**
8	2 × HBA, 2 × Hbic	0.81	125.8	1.9	12	139.7	166.5	26.8	95
9	2 × HBA, 2 × Hbic	0.81	125.8	2.2	12	140.1	166.5	26.4	95
10	2 × HBA, 2 × Hbic	0.79	127.3	1.3	12	140.6	166.5	25.9	95
					Run[a] 4				
1	3 × HBA, Hbic	0.91	86.9	1.2	16.5	104.5	146.9	42.4	95
2[f]	**3 × HBA, Hbic**	**0.91**	**87**	**1.4**	**16.5**	**104.9**	**146.9**	**42**	**95**
3[f]	**2 × HBA, Hbic, RingAro**	**0.9**	**88.6**	**1.5**	**16.5**	**106.5**	**146.9**	**40.4**	**95**
4	3 × HBD, Hbic	0.89	89.6	1.1	16.5	107.2	146.9	39.7	95
5[f]	**2 × HBD, 2 × Hbic**	**0.89**	**89.8**	**1.1**	**16.5**	**107.4**	**146.9**	**39.5**	**95**
6	2 × HBA, HBD, Hbic	0.88	89.9	1.2	16.5	107.6	146.9	39.3	95
7	2 × HBA, HBD, Hbic	0.89	89.7	1.4	16.5	107.6	146.9	39.3	95
8	2 × HBA, HBD, Hbic	0.88	90.1	1.2	16.5	107.8	146.9	39.1	95
9	3 × HBD, Hbic	0.88	90.1	1.2	16.5	107.8	146.9	39.1	95
10	3 × HBA, Hbic	0.88	90.4	1.1	16.5	108	146.9	38.9	95
					Run[a] 5				
1	2 × HBA, HBD, Hbic	0.89	89.8	1.5	16.1	107.3	146.9	39.6	95
2	3 × HBA, Hbic	0.88	89.9	1.3	16.1	107.3	146.9	39.6	95
3	3 × HBA, Hbic	0.89	89.9	1.8	16.1	107.8	146.9	39.1	95
4	3 × HBA, Hbic	0.88	90.2	1.8	16.1	108.1	146.9	38.8	95
5	3 × HBD, Hbic	0.88	91.2	1.2	16.1	108.4	146.9	38.5	95
6	3 × HBA, Hbic	0.88	91.1	1.3	16.1	108.5	146.9	38.4	95
7	3 × HBD, Hbic	0.87	91.4	1.2	16.1	108.7	146.9	38.2	95
8	2 × HBA, HBD, Hbic	0.87	91.5	1.2	16.1	108.8	146.9	38.1	95
9	2 × HBA, Hbic, RingAro	0.87	91.6	1.2	16.1	108.8	146.9	38.1	95
10	2 × HBA, HBD, Hbic	0.87	91.6	1.1	16.1	108.9	146.9	38	95
					Run[a] 6				
1[f]	**2 × HBA, HBD, Hbic**	**0.88**	**90.8**	**1.1**	**14.6**	**106.5**	**146.9**	**40.4**	**95**
2	2 × HBA, HBD, Hbic	0.88	90.9	1.1	14.6	106.6	146.9	40.3	95
3[f]	**HBA, HBD, 2 × Hbic**	**0.88**	**90.9**	**1.1**	**14.6**	**106.7**	**146.9**	**40.2**	**95**
4	2 × HBA, 2 × Hbic	0.87	91.1	1.2	14.6	106.9	146.9	40	95
5	HBA, HBD, 2 × Hbic	0.87	91.3	1.2	14.6	107	146.9	39.9	95
6[f]	**HBA, HBD, 2 × Hbic**	**0.89**	**89.8**	**2.8**	**14.6**	**107.1**	**146.9**	**39.8**	**95**
7	2 × HBA, 2 × Hbic	0.87	92.2	1.1	14.6	108	146.9	38.9	95
8[f]	**HBA, 2 × HBD, Hbic**	**0.88**	**90.5**	**3.2**	**14.6**	**108.3**	**146.9**	**38.6**	**95**
9	2 × HBA, 2 × Hbic	0.88	90.8	3.1	14.6	108.5	146.9	38.4	95
10	2 × HBA, 2 × Hbic	0.87	91.6	3.1	14.6	109.3	146.9	37.6	95

[a] The automatic pharmacophore run employing the training sets, parameters, and conditions in Table 3. [b] GSK-3β inhibition hypotheses from the corresponding runs. The models are ranked by CATALYST-HYPOGEN according to their cost criteria. [c] The correlation coefficients between bioactivity estimates and bioactivities within each respective training subset (see Table 3). [d] Configuration costs. [e] The difference between the total cost and the cost of the corresponding null hypotheses. [f] These pharmacophores (bolded) were selected as the best representatives of their clusters (see text).

Table 5. Statistical Results of the Scanned QSAR Models

model	explanatory terms[a]	r^2_{123}[b]	F	r^2_{LOO}[c]	r^2_{BS}[d]	r^2_{PRESS}[e]	PRESS[f]
A	11	0.712	27.6	0.668	0.712	0.502	9.928
B[g]	**10**	**0.663**	**24.6**	**0.592**	**0.662**	**0.695**	**6.08**
C	9	0.664	28.3	0.615	0.666	0.552	8.923
D	8	0.629	27.9	0.572	0.630	0.451	10.935

[a] Including the intercept (the correlation constant). [b] Non-cross-validated correlation coefficient for 123 training compounds. [c] Cross-validation correlation coefficients determined by the leave-one-out technique. [d] Bootstrapping correlation coefficient. [e] Predictive r^2 determined for the 29 test compounds. [f] The sum of squared deviations between predicted and actual activity values for every molecule in the test set of 29 compounds. [g] This QSAR equation was selected to predict the GSK-3β inhibitory activities of the captured hits because it yielded the best statistical criteria.

The captured hits were fitted against Hypo6/3 and those of fit values <5.67 (the spline intercept associated with Hypo6/3 in QSAR eq 1) were excluded. The remaining hits, i.e., 57, were subsequently fitted against Hypo4/5 and Hypo6/3, and their fit values, together with their other relevant physicochemical descriptors, were substituted in QSAR eq 1 to determine their predicted bioactivities. However, in order to minimize the impact of any extrapolatory prediction errors (which are not unusual in QSAR models[81]) on decisions regarding which hits merit subsequent in vitro testing, we employed the predicted log(1/IC_{50}) values to rank the hits such that the highest ranking compounds were selected for experimental evaluation. Hydroxychloroquine, cimetidine, and gemifloxacin ranked highest, and therefore, they were evaluated in vitro and in vivo. Table 7 shows their estimated and experimental anti-GSK-3β bioactivities, while Figure 5 shows how the three compounds fit Hypo6/3 and Hypo4/5.

Figure 7 shows the responses of liver glycogen reserves in fasting Balb/C mice to ip administered hydroxychloroquine, cimetidine, and gemifloxacin at doses equivalent to 1, 2, and 4 times the human doses normalized to mice weights (hydroxychloroquine = 5.70 mg/kg; gemifloxacin = 4.57 mg/kg; cimetidine = 11.40 mg/kg). Clearly from the figure, all three drugs caused marked elevation in liver glycogen proportional to the administered doses. However, cimetidine illustrated the most potent glycogen-sparing properties.

On the basis of these findings, we believe that the potent anti-GSK-3β activity of hydroxychloroquine is at least partly responsible for its established anti-inflammatory and antimalarial pharmacological properties. Plasmodial GSK is highly homologous to human GSK-3β,[22] while human GSK-3β is involved in several inflammatory responses such as rheumatoid arthritis.[1] Furthermore, we believe that the glycogen-sparing effect of hydroxychloroquine is responsible for its reported hypoglycemic effects.[72]

Cimetidine and quinolone antibacterials (the family of gemifloxacin) were also reported to cause significant hypoglycemia,[73–75] which can be also attributed to their potent anti-GSK-3β effects.

2.5. Comparison of Hypo6/3 with the Binding Site of GSK-3β. Despite problems associated with crystallographic structures,[44–46] the pharmacophore features of Hypo6/3 and Hypo4/5 can be compared with the binding pocket of GSK-3β to identify probable residues important for inhibition. The features in Hypo6/3 and Hypo4/5, as well as the alignment of cimetidine (selected as representative example) as proposed by Hypo6/3 and Hypo4/5, were compared with the corresponding docked structure of cimetidine into the binding pocket of GSK-3β (PDB code 1Q5K, resolution = 1.94 Å), as in Figure 6. The docking experiment was performed employing the LIGANDFIT docking engine and through default docking parameters and a consensus scoring function.[76–78] A marked similarity was observed between the features proposed by the pharmacophore model and the ligand binding features in the docked structure.

The docking study shows that the two highest-ranking docked poses of cimetidine (based on consensus of six scoring functions PLP1, PLP2, Ligscore1, Ligscore2, JAIN, and PMF[76–78]) agree with how cimetidine fits Hypo6/3 and Hypo4/5 (Figures 5G,H and 6): In the case of Hypo6/3 (Figure 5G), mapping the cyano group of cimetidine by a hydrogen-bond acceptor feature seems to correspond to hydrogen-bonding with the amidic NH of VAL135 in one of the optimal docked poses (Figure 6A). On the other hand, results from mapping the guanidino methyl and sulfide linker of cimetidine by two hydrophobic features in Hypo6/3 (Figure 5G) apparently correspond to hydrophobic interactions with the isobutyl side chain of ILE62 (Figure 6A). However, mapping the imidazole NH of cimetidine with a hydrogen-bond donor in Hypo6/3 does not seem to be supported by any corresponding interaction in the corresponding docked structure. However, we believe this interaction resembles hydrogen-bonding to certain water molecules not explicit in the crystallographic structure and outside the binding pocket.

Similarly, fitting cimetidine against Hypo4/5 agrees with another optimally docked pose of cimetidine, as in Figures 5H and 6B. Mapping the imidazolic NH of cimetidine with a hydrogen-bond donor feature in Hypo4/5 (Figure 5H) corresponds to hydrogen-bonding to an explicit water molecule bound to the peptidic NH of ASP200 (Figure 6B). Furthermore, mapping the imidazolic methyl group of cimetidine with a hydrophobic feature in Hypo4/5 (Figure 5H) seems to correspond to how this group is inserted into a hydrophobic pocket comprising VAL110, LEU132, and ALA83 in the docked complex (Figure 6B). The apparent existence of hydrogen-bond interactions connecting the guanidino NHCN of cimetidine and the carbonyl oxygen of PRO136 (as in Figure 6B) agrees with mapping this moiety with the hydrogen-bond donor feature in Hypo4/5 (Figure 5H). Finally, results from mapping the guanidino CH_3 of cimetidine with the hydrophobic feature in Hypo4/5 (Figure 5H) agree with the apparent hydrophobic interactions connecting this group and the isobutyl side chain of ILE62 in the docked complex (Figure 6B).

3. Conclusion

Our results suggest that pharmacophore modeling of GSK-3β can be a useful tool for finding potential hypoglycemic agents. The exploration of the pharmacophoric space of different GSK-3β inhibitors was performed via two diverse sets of inhibitors and through the use of CATALYST-HYPOGEN to identify high-quality binding model(s). QSAR analysis identified an optimal combination of pharmacophoric models and 2D descriptors capable of explaining bioactivity variations across a large set of diverse inhibitors. Two orthogonal pharmacophores emerged in the QSAR equation, suggesting the existence of at least two distinct binding modes accessible to ligands within GSK-3β binding pocket.

The optimal QSAR equation and the associated pharmacophoric models were experimentally validated by the identification of three potent GSK-3β inhibitors retrieved from a structural database of established drug molecules, namely, hydroxychloroquine, cimetidine, and gemifloxacin. The three compounds illustrated potent anti-GSK-3β bioactivities both in vivo (hypoglycemic and glycogen reservation) and in vitro.

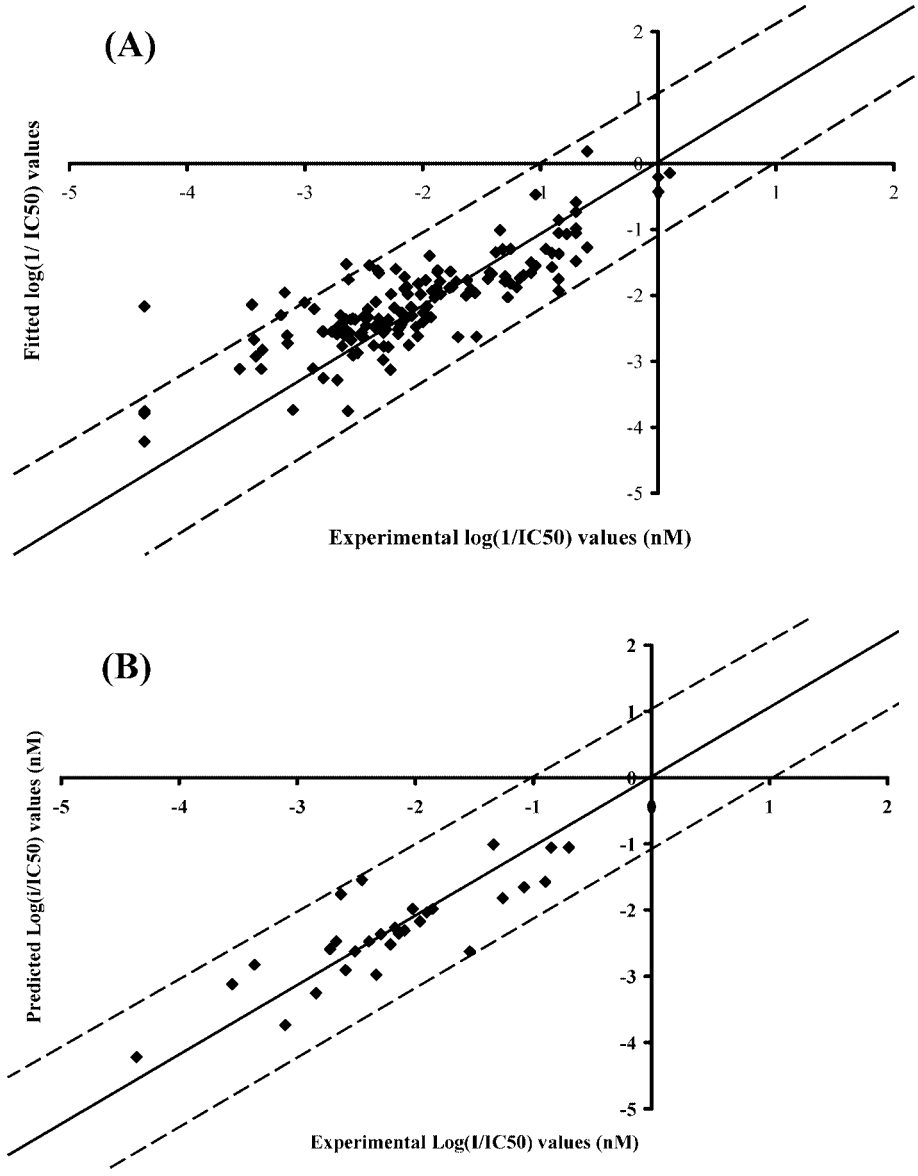

Figure 2. Experimental versus (A) fitted (123 compounds, $r^2_{LOO} = 0.592$) and (B) predicted (29 compounds, $r^2_{PRESS} = 0.695$) bioactivities calculated from the best QSAR model (eq 1). The solid lines are the regression lines for the fitted and predicted bioactivities of training and test compounds, respectively, whereas the dotted lines indicate the 1.0 log point error margins.

4. Experimental Sections

4.1. Molecular Modeling. 4.1.1. Software and Hardware. The following software packages were utilized: CATALYST (version 4.11) from Accelrys Inc. (www.accelrys.com); CERIUS2 (version 4.10) from Accelrys Inc. (www.accelrys.com); CS ChemDraw Ultra 7.01 from Cambridge Soft Corp. (http://www.cambridgesoft.Com).

Pharmacophore modeling, QSAR analysis, and docking studies were performed using CATALYST and CERIUS2 software suites (Accelrys Inc., San Diego, CA, www.accelrys.com) installed on a Silicon Graphics Octane2 desktop workstation equipped with a dual 600 MHz MIPS R14000 processor (1.0 GB RAM) running the Irix 6.5 operating system. Structure drawing was performed employing CS ChemDraw Ultra 7.01 installed on a Pentium 4 PC.

4.1.2. Data Set. The structures of 152 GSK-3β inhibitors were collected from recently published literature.[59–62] The in vitro bioactivities of the collected inhibitors were determined employing identical bioassay conditions and expressed as the concentration of the test compound that inhibited the activity of GSK-3β by 50% (IC_{50}). Figure 1 and Table 1 show the structures and IC_{50} values of the considered inhibitors. The logarithms of the IC_{50} values were used in pharmacophore modeling and QSAR analysis, thus correlating the data as linear to the free energy change.

The two-dimensional (2D) chemical structures of the inhibitors were sketched using ChemDraw Ultra and saved in MDL-molfile format. Subsequently, they were imported into CATALYST, converted into corresponding standard 3D structures, and energy-minimized to the closest local minimum using the molecular mechanics CHARMm force field implemented in CATALYST. The resulting 3D structures were utilized as starting conformers for CATALYST conformational analysis and were stored in SD format for calculation of a variety of physicochemical properties for QSAR analysis within CERIUS2. Two diverse sets of GSK-3β inhibitors were selected for pharmacophore modeling: sets A and B in Table 2. The whole set (**1–152**, Figure 1 and Table 1) was employed for QSAR analysis.

4.1.3. Conformational Analysis. Molecular flexibility was taken into account by considering each compound as a collection of conformers representing different areas of the conformational space accessible to the molecule within a given energy range. Accordingly, the conformational space of each inhibitor (**1**−**152**, Figure 1 and Table 1) was explored by adopting the "best conformer generation" option within CATALYST, which is based on the

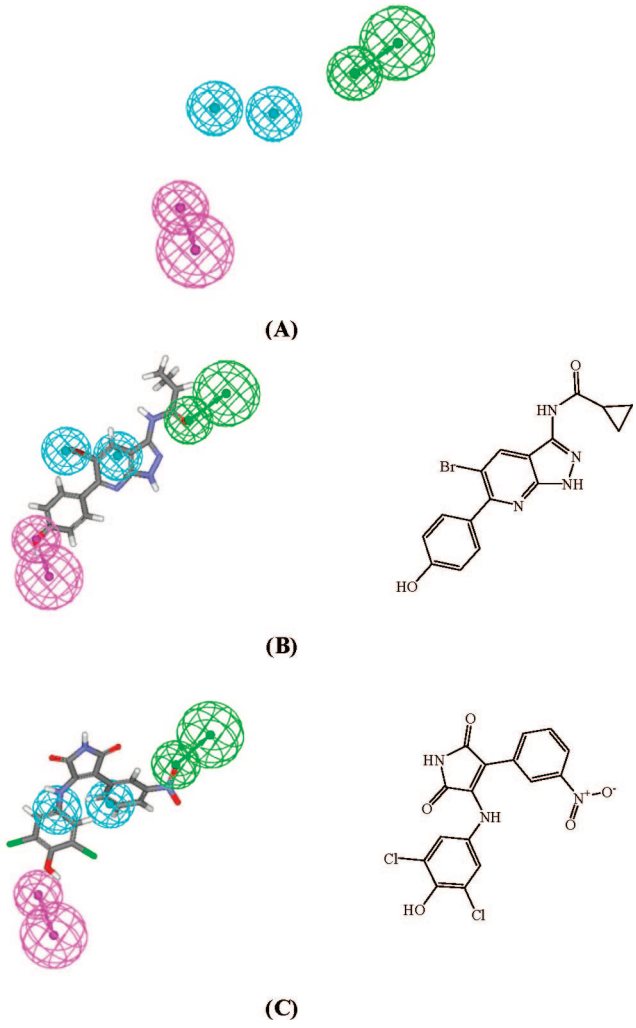

Figure 3. (A) Hypo6/3 pharmacophoric features of the binding model, where hydrogen bond acceptors are depicted as green vectored spheres, hydrophobic features as blue spheres, and hydrogen-bond donors as violet vectored spheres; (B) Hypo6/3 fitted against the most potent inhibitor **87** (Table 1 and Figure 1, IC_{50} = 0.80 nM); (C) Hypo 6/3 fitted against inhibitor **44** (Table 1 and Figure 1, IC_{50} = 20.0 nM).

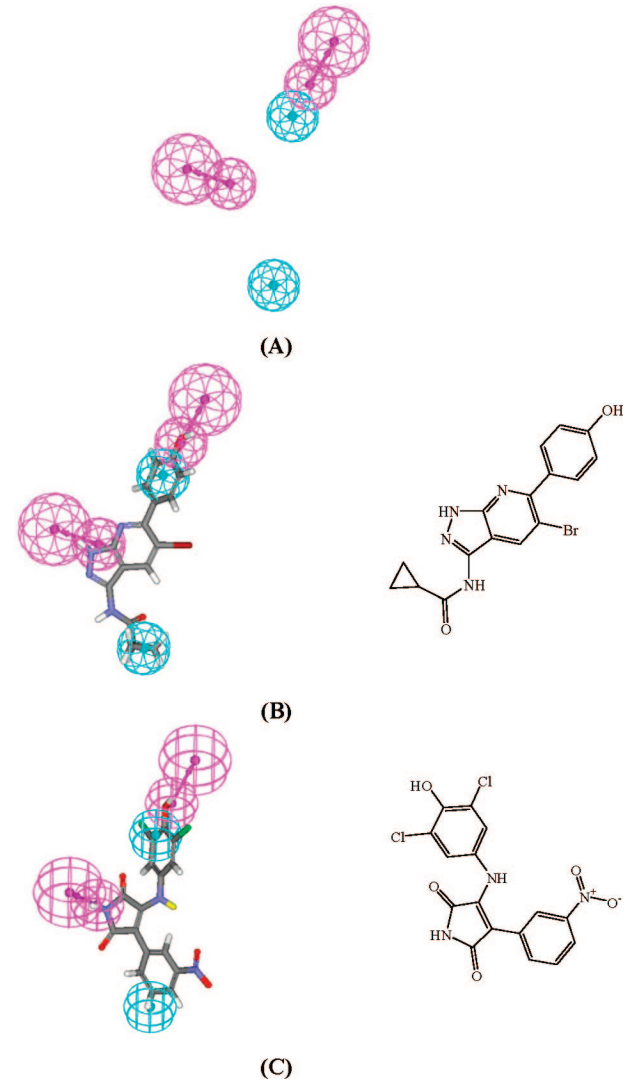

Figure 4. (A) Hypo4/5 pharmacophoric features of the binding model, where hydrogen bond donors are depicted as violet vectored spheres and hydrophobic features as blue spheres; (B) Hypo4/5 fitted against the most potent inhibitor **87** (Figure 1 and Table 1, IC_{50} = 0.80 nM); (C) Hypo4/5 fitted against inhibitor **44** (Figure 1 and Table 1, IC_{50} = 20.0 nM).

generalized CHARMm force field implemented in the program. Default parameters were employed in the conformation generation procedure; i.e., a conformational ensemble was generated with an energy threshold of 20 kcal/mol from the structure of the lowest energy level and a maximum limit of 250 conformers per molecule. This search procedure will probably identify the best three-dimensional arrangement of chemical functionalities explaining the activity variations among the training set.[49]

4.1.4. Exploration of the Pharmacophoric Space of GSK-3*β*. All 152 molecules with their associated conformational models were regrouped into a spreadsheet. The biological data of the inhibitors were reported with an "uncertainty" value of 3, which means that the actual bioactivity of a particular inhibitor is assumed to be situated somewhere in an interval ranging from one-third to 3 times the reported bioactivity value of that inhibitor.[65,66] Subsequently, two structurally diverse training subsets (Table 2) were carefully selected from the collection for pharmacophore modeling. Typically, CATALYST requires informative training sets that include at least 16 compounds of evenly spread bioactivities over at least 4 orders of magnitude. Lesser training lists could lead to chance correlation and thus faulty models.[65,66]

The selected training sets were utilized to conduct six modeling runs to explore the pharmacophoric space of GSK-3β inhibitors (Table 3). Different hypotheses were generated by altering the spacing parameter in the resulting pharmacophores (Tables 3 and 4). We allowed the number of features in the generated pharmacophore hypotheses to vary from 4 to 5.

Pharmacophore modeling employing CATALYST proceeds through three successive phases: the constructive phase, subtractive phase, and optimization phase.[65,66] During the constructive phase, CATALYST generates common conformational alignments among the most active training compounds. Only molecular alignments based on a maximum of five chemical features are considered. The program identifies a particular compound as being within the most active category if it satisfies eq 2.[65,66]

$$(\mathrm{MAct} \times \mathrm{UncMAct}) - (\mathrm{Act}/\mathrm{UncAct}) > 0.0 \qquad (2)$$

where "MAct" is the activity of the most active compound in the training set, "Unc" is the uncertainty of the compounds, and "Act" is the activity of the training compounds under question. However, if there are more than eight most-active inhibitors, only the top eight are used.

In the subsequent subtractive phase, CATALYST eliminates some hypotheses that fit inactive training compounds. A particular training compound is defined as being inactive if it satisfies eq 3:[65,66]

$$\log(\mathrm{Act}) - \log(\mathrm{MAct}) > 3.5 \qquad (3)$$

Table 6. Pharmacophoric Features and Corresponding Weights, Tolerances, and 3D Coordinates of Hypo6/3[a] and Hypo4/5[b]

	chemical feature					
parameter	HBA		HBD		Hbic	Hbic
	Model Hypo6/3[a]					
weight	2.02163		2.02163		2.02163	2.02163
tolerances	1.60	2.20	1.60	2.20	1.60	1.60
coordinate *X*	−4.83	−7.66	7.16	7.73	0.12	1.80
coordinate *Y*	−1.62	−2.18	0.10	−1.88	−0.06	2.84
coordinate *Z*	0.91	1.84	−2.67	−4.86	0.32	−1.40

	chemical feature					
	HBD		HBD		Hbic	Hbic
	Model Hypo4/5[b]					
weight	1.99602		1.99602		1.99602	1.99602
tolerance	1.60	2.20	1.60	2.20	1.60	1.60
coordinate *X*	−3.81	−5.95	−7.11	−8.80	0.65	−5.96
coordinate *Y*	0.39	1.48	−5.18	−7.24	4.65	−3.54
coordinate *Z*	3.37	5.16	−1.08	−2.47	0.04	−0.14

[a] Hypo6/3 is hypothesis number 3 generated in run number 6 (as in Table 4). [b] Hypo4/5 is hypothesis number 5 generated in run number 4 (as in Table 4).

However, in the optimization phase, CATALYST applies fine perturbations in the form of vectored feature rotation, adding a new feature and/or removing a feature, to selected hypotheses that survived the subtractive phase in an attempt to find new models of enhanced bioactivity/mapping correlation, i.e., improved 3D-QSAR properties. Eventually, CATALYST selects the highest-ranking models (10 by default) and presents them as the optimal pharmacophore hypotheses resulting from a particular automatic modeling run.

4.1.5. Assessment of the Generated Hypotheses. When generating hypotheses, CATALYST attempts to minimize a cost function consisting of three terms: weight cost, error cost, and configuration cost.[49,63,65,66,69] Weight cost is a value that increases as the feature weight in a model deviates from an ideal value of 2. The deviation between the estimated activities of the training set and their experimentally determined values adds to the error cost. The activity of any compound can be estimated from a particular hypothesis through eq 4.[49]

$$\log(\text{estimated activity}) = I + \text{Fit} \quad (4)$$

where I is the intercept of the regression line obtained by plotting the log of the biological activity of the training set compounds against the Fit values of the training compounds. The Fit value for any compound is obtained automatically by employing eq 5.[49]

$$\text{Fit} = \sum(\text{mapped hypothesis features}) \times W[1 - \sum(\text{disp/tol})^2] \quad (5)$$

where $\sum$(mapped hypothesis features) represents the number of pharmacophore features that successfully superimpose (i.e., map or overlap with) corresponding chemical moieties within the fitted compound and where W is the weight of the corresponding hypothesis feature spheres. This value is fixed to 2.0 in HYPOGEN-generated models. disp is the distance between the center of a particular pharmacophoric sphere (feature centroid) and the center of the corresponding superimposed chemical moiety of the fitted compound. tol is the radius of the pharmacophoric feature sphere (known as tolerance, equaling 1.6 Å by default). $\sum(\text{disp/tol})^2$ is the summation of $(\text{disp/tol})^2$ values for all pharmacophoric features that successfully superimpose corresponding chemical functionalities in the fitted compound.[49]

The third term, i.e., the configuration cost, penalizes the complexity of the hypothesis. This is a fixed cost, which is equal to the entropy of the hypothesis space. The more the numbers of features (a maximum of five) are in a generated hypothesis, the higher is the entropy with subsequent increase in this cost.

The overall cost (total cost) of a hypothesis is calculated by summing over the three cost factors. Error cost is the main contributor to total cost.

CATALYST also calculates the cost of the null hypothesis, which presumes that there is no relationship in the data and that experimental activities are normally distributed about their mean. Accordingly, the greater is the difference from the null hypothesis cost, the more likely that the hypothesis does not reflect a chance correlation. In a successful automatic modeling run, CATALYST ranks the generated models according to their total costs.[49]

An additional approach to assess the quality of CATALYST pharmacophores is to cross-validate them using the Cat.Scramble algorithm. This validation procedure is based on Fischer's randomization test.[68] In this validation test, we selected a 95% confidence level, which generates 19 random spreadsheets. Subsequently, CATALYST is challenged to use these random spreadsheets to generate hypotheses using the same features and parameters used in generating the initial unscrambled hypotheses. Success in generating pharmacophores of comparable cost criteria to those produced by the original unscrambled data reduces the confidence in the training compounds and the pharmacophore models from unscrambled data.

Table 4 shows the pharmacophoric features and success criteria of the generated binding hypotheses, including the cost values (error, configuration, weight, and total costs) of the resulting pharmacophores and the costs of the corresponding null hypotheses. The table also shows the corresponding Cat.scramble confidence for each pharmacophore.

4.1.6. Clustering of the Generated Pharmacophore Hypotheses. Because of the large number of generated pharmacophores (60 model) and their similar 3D features and properties (cost criteria, Cat.scramble confidence, etc.), it was decided to cluster them into 10 groups utilizing the hierarchical average linkage method available in CATALYST. Thereafter, the highest-ranking hypothesis within each cluster (i.e., of lowest cost) was selected to represent the corresponding cluster in subsequent QSAR modeling. Table 4 shows the representative pharmacophores (bolded).

4.1.7. QSAR Modeling. A subset of 123 compounds from the total list of inhibitors (Table 1 and Figure 1) was utilized as a training set for QSAR modeling. However, since it is essential to access the predictive power of the resulting models on an external set of inhibitors, the remaining 29 molecules (about 20% of the data set) were employed as an external test subset for validating the QSAR models. The test molecules were selected as follows: the inhibitors (**1**−**152**, Table 1 and Figure 1) were ranked according to their IC_{50} values, and then every fifth compound was selected for the test set starting from the high-potency end. This selection considers the fact that the test molecules must represent a range of biological activities similar to that of the training set. The selected test inhibitors are **4**, **9**, **16**, **19**, **22**, **34**, **40**, **42**, **45**, **53**, **56**, **57**, **62**, **67**, **72**, **74**, **79**, **89**, **90**, **110**, **119**, **120**, **129**, **130**, **139**, **140**, **147**, **149**, and **152** (numbers are as in Table 1 and Figure 1).

The logarithm of measured $1/IC_{50}$ (nM) values were used in QSAR, thus correlating the data as linear to the free energy change.

The chemical structures of the inhibitors were imported into CERIUS2 as standard 3D single conformer representations in SD format. Subsequently, different descriptors (100 terms) were calculated for each compound employing the C2.DESCRIPTOR module of CERIUS2. The calculated descriptors included simple and valence connectivity indices, electrotopological state indices, κ shape indices, Shadow indices, JURS descriptors, semiempirical quantum mechanical energy terms via AM1-MOPAC (HOMO and LUMO), logarithm of partition coefficient, polarizability, dipole moment, molecular volume, molecular weight, and molecular surface area.[70] Furthermore, the training compounds were fitted against the representative pharmacophores (using the best-fit option in CATALYST[49]) and their fit values were added as additional descriptors (see section 2.3, QSAR Modeling). The fit value for any compound is obtained automatically via eq 5.[49] However, we removed 20% of the descriptors based on their poor variance. Genetic function approximation (GFA) was employed to search

Table 7. Hit Molecules Captured by Hypo6/3 and Hypo4/5 and Their Corresponding QSAR Estimates from Eq *1* and Their in Vitro Bioactivities

	fit values against[b]		QSAR estimates[c]		exptl values	
drug[a]	Hypo6/3	Hypo4/5	log(1/IC_{50})	IC_{50} (nM)	log(1/IC_{50})	IC_{50} (nM)[d]
hydroxychloroquine	7.82	5.63	−0.2	1.6	−1.5	33
gemifloxacin	6.34	0	−2.8	631	−1.9	88
cimetidine	6.91	3.82	−2.0	100	−1.1	13
standard inhibitor[e]						1500

[a] Hit drugs shown in Figure 5. [b] Best-fit values against each binding hypothesis calculated by eq 4. [c] QSAR estimates from eq 1. [d] In vitro enzyme inhibition. [e] See ref 33.

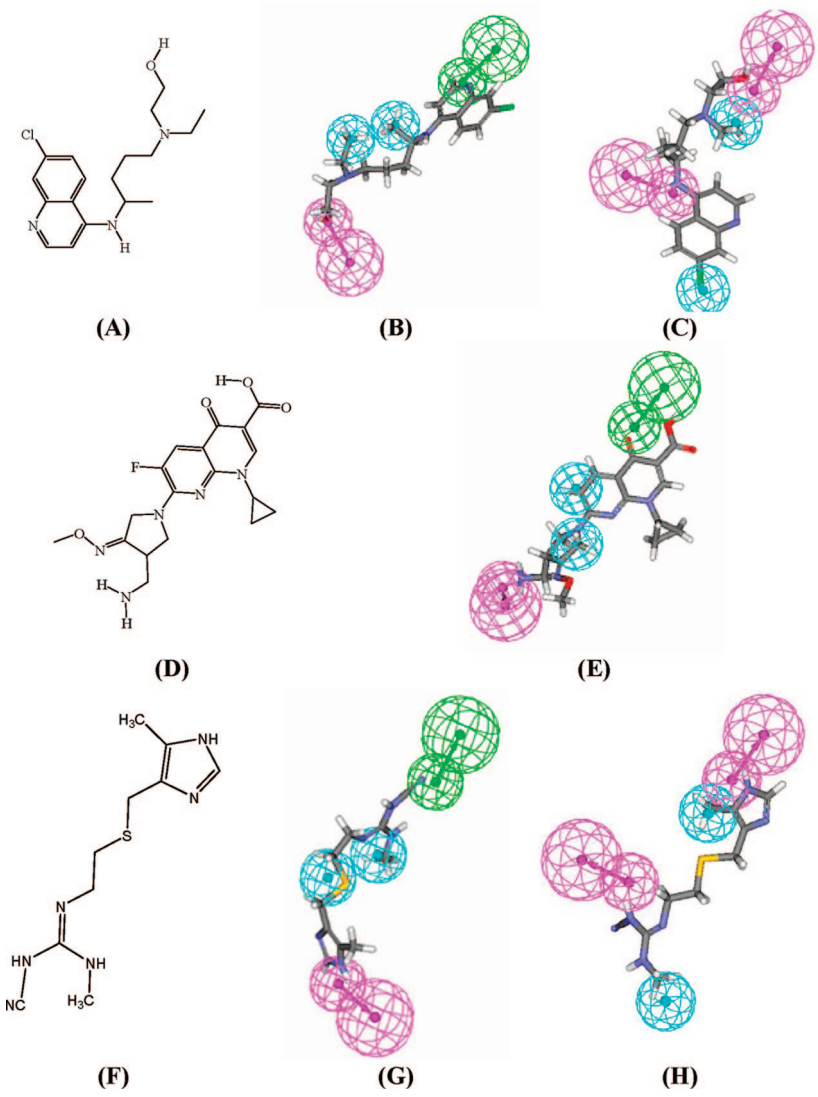

Figure 5. Chemical structures of the highest-ranking drug hits (as suggested by the best QSAR model) tested as GSK-3β inhibitors and how they map the associated pharmacophores: hydroxychloroquine (A); mapped to Hypo6/3 (B) and Hypo4/5 (C); gemifloxacin (D); mapped to Hypo6/3 (E) and cimetidine (F); mapped to Hypo6/3 (G) and Hypo4/5 (H).

for the best possible QSAR regression equation capable of correlating the variations in biological activities of the training compounds with variations in the generated descriptors, i.e., multiple linear regression modeling (MLR). GFA techniques rely on the evolutionary operations of "crossover and mutation" to select optimal combinations of descriptors (i.e., chromosomes) capable of explaining bioactivity variation among training compounds from a large pool of possible descriptor combinations. Each chromosome is associated with a fitness value that reflects how good it is compared to other solutions. The fitness function employed herein is based on Friedman's "lack-of-fit" (LOF).[70]

Our preliminary diagnostic trials suggested the following optimal GFA parameters: exploration of linear and spline equations at mating and mutation probabilities of 50%; population size = 500; number of genetic iterations = 30 000; lack-of-fit (LOF) smoothness parameter = 1.0. However, to determine the optimal number of explanatory terms (QSAR descriptors), it was decided to scan and evaluate all possible QSAR models resulting from 8 to 30 explanatory terms.

All QSAR models were validated employing leave one-out cross-validation (r^2_{LOO}), bootstrapping (r^2_{BS}), and predictive r^2 (r^2_{PRESS}) calculated from the test subsets. The predictive r^2_{PRESS} is defined as

$$r^2_{PRESS} = SD - PRESS/SD \quad (6)$$

where SD is the sum of the squared deviations between the biological activities of the test set and the mean activity of the training set molecules and where PRESS is the squared deviations between predicted and actual activity values for every molecule in

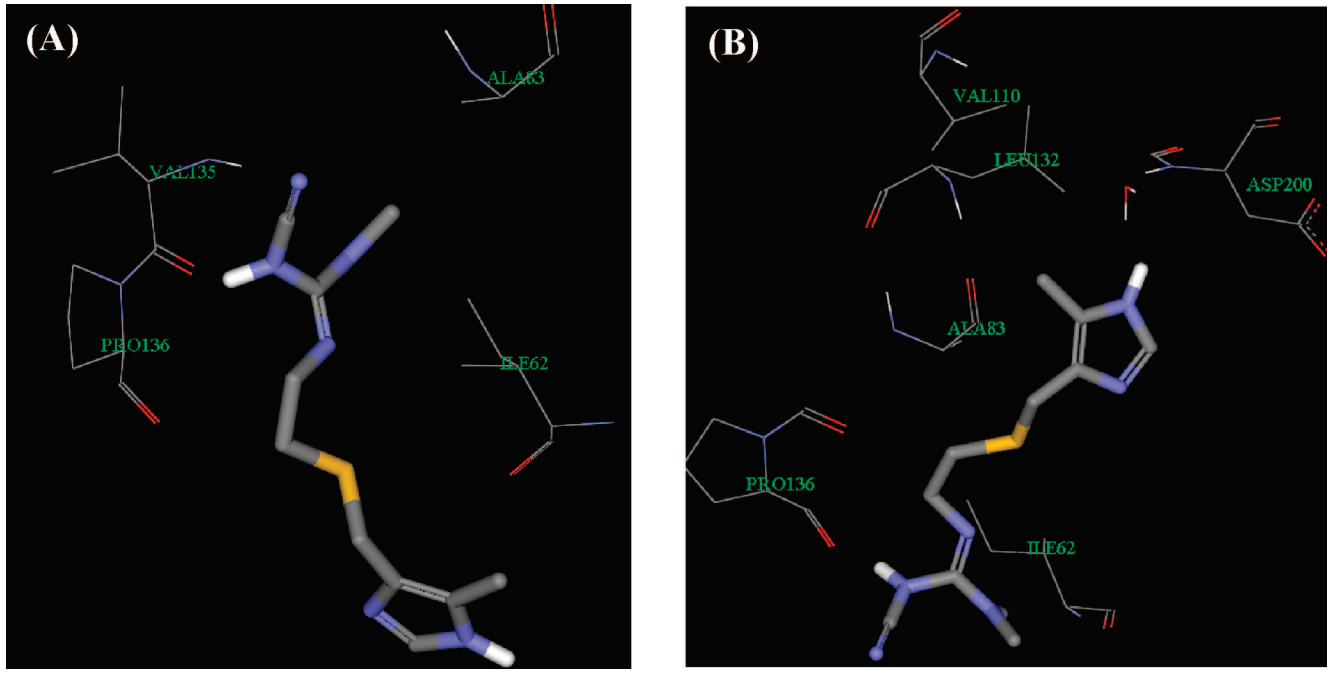

Figure 6. Successful docking poses of cimetidine resembling (A) Hypo6/3 and (B) Hypo4/5.

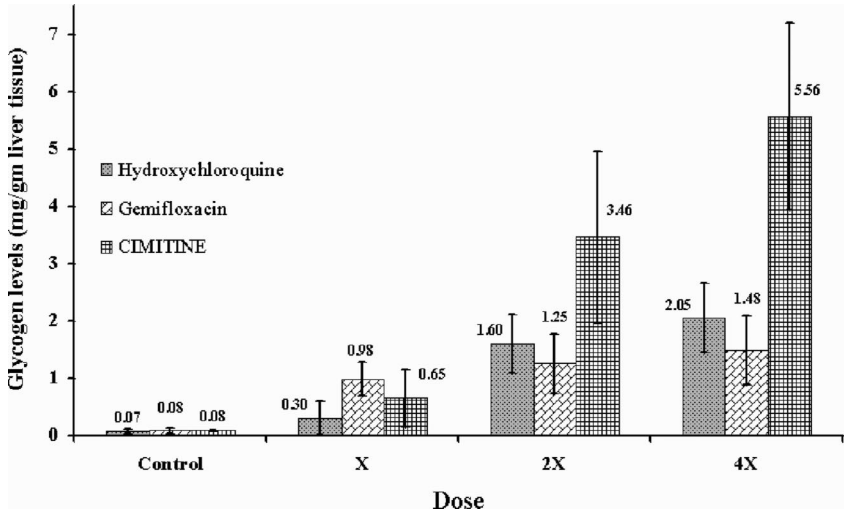

Figure 7. Liver glycogen reserves in fasting 12-week-old male Balb/C mice 3 h after ip administration of vehicle control (PBS buffer) and 1-, 2-, and 4-fold (1×, 2×, and 4×, respectively) of the human doses of hydroxychloroquine, gemifloxacin, and cimetidine normalized to mice weights (20 g). Each result represents the average glycogen level from five mice. Error bars indicate the standard deviations of the measurements.

the test set. Descriptor-scanning identified eq 1 (QSAR model B in Table 5) as the best equation. Figure 2 shows the plots of experimental versus fitted (training set) and predicted (testing set) GSK-3β inhibitory bioactivities calculated from the best QSAR equation. Two pharmacophore hypotheses emerged in this model and frequented in high-ranking QSAR equations (models A−D in Table 5), namely, Hypo6/3 and Hypo 4/5. Table 6 shows the three-dimensional coordinates of the two pharmacophores, while Figures 3 and 4 show the pharmacophoric features of the two models and how they map two potent inhibitors.

4.1.8. In Silico Screening for New GSK-3β Inhibitors. Hypo6/3 was employed as 3D search query to screen our in-house multiconformer database of established drug molecules via the "best flexible database search" approach within CATALYST. It captured 397 hits that were fitted against Hypo6/3 using the "best fit" option within CATALYST. The fact that Hypo6/3 appeared in QSAR eq 1 in combination with a spline intercept of −5.67 prompted the selection of hits that have fit values above 5.67 leaving 57 drug hits. Subsequently, these hits were fitted against Hypo4/5 and Hypo6/3 and their fit values together with their other relevant molecular descriptors were substituted in QSAR eq 1.

QSAR predictions suggested that three of the hits should possess potent GSK-3β inhibitory activities: hydroxychloroquine, cimetidine, and gemifloxacin. The three drugs were acquired and tested in vitro as GSK-3β inhibitors and in vivo for their glycogen-boosting bioactivities.

4.2. In Vitro Assays. 4.2.1. Preparation of Hit Compounds for in Vitro Enzymatic Assay. Each hit compound was dissolved in HEPES (pH 7.2) to give a final stock solution of 40 μM for subsequent enzymatic assays.

4.2.2. Determination of the Inhibitory IC_{50} Values. Recombinant GSK-3β (Upstate Biotechnology) was dissolved in a buffer solution (pH 7.2) containing the following: 40 mM HEPES, 5 mM $MgCl_2$, 5 mM EDTA, 100 μM ATP, and 50 μg/mL heparin to reach a final enzymatic solution of 10 pg/mL. Subsequently, 50 μL aliquots of the enzymatic solution were pipetted into 0.5 mL vials. Thereafter, appropriate volumes of the hits' stock solutions were

pipetted into the enzymatic solution to yield 10 μM, 1 μM, 100 nM, and 10 nM of each hit after completion to 75 μL with the buffer solution. The hits were incubated with the enzyme over 30 min at room temperature. Then an amount of 25 μL of 2000 pg/mL tau protein solution in HEPES was added to give a final tau protein concentration of 500 pg/mL. This mixture was incubated over 1 h at room temperature.

The detection of tau phosphorylation was performed as follows:[79] The GSK-3β reaction mixtures were diluted 1:1 with sodium azide aqueous solution (15 mM) to achieve a final tau protein concentration of 250 pg/mL. Then 100 μL aliquots of this solution were pipetted into the wells of the tau [pS396] phosphoELISA kit (Biosource). Subsequently, the wells were incubated for 2 h at room temperature. Then they were aspirated and washed (with the washing solution provided in the kit). Thereafter, 100 μL aliquots of rabbit detector antibody solution were pipetted in the wells and incubated for 1 h at room temperature. Thereafter, the wells were aspirated and washed with the wash buffer. Then 100 μL aliquots of goat (polyclonal) antirabbit IgG-HRP were added to the wells and incubated for 30 min at room temperature. Subsequently, the wells were aspirated and washed with the wash buffer. Finally, a total of 100 μL of TMB substrate chromogen solution aliquots was added to each well and incubated for 20–30 min. After the termination of the HRP reaction in each well, the solution absorbances were measured spectrophotometrically at λ of 450 nm. A standard GSK-3β inhibitor (4-benzyl-2-methyl-1,2,4-thiadiazolidine-3,5-dione, also known as TDZD-8, from Biosource) was employed as positive control.[33]

4.3. In Vivo Evaluation. 4.3.2. Preparation of Drug Hits. The effect of the three drugs on liver glycogen content of Balb/C mice was investigated. Each drug hit was dissolved in PBS (pH 7.2) and prepared in three different concentrations: (i) human dose normalized to mouse weight (about 20 g), (ii) double the normalized human dose, and (iii) 4 times the normalized human dose. All doses were injected ip as PBS solutions (0.20 mL). The human doses of the three drugs are as follows: hydroxychloroquine (5.70 mg/kg for antimalarial action), gemifloxacin (4.57 mg/kg for antibiotic action), cimetidine (11.40 mg/kg for H_2 receptor antagonist action).

4.3.3. Determination of Liver Glycogen. Twelve-week old male Balb/C mice with average weight of 20 g were used for the investigations. The animals were randomized and fed ad libitum with standard feed and water except when fasting was needed in the course of the study.

All animals were housed in the same conditions and separated randomly in five-mice groups. Four groups were used to investigate each drug: one group was injected ip with PBS as negative control. The other three groups were injected with the three escalating doses of each drug (mentioned earlier). On the day of the experiment, food and water were removed 4 h before the injection. The animals were scarified by cervical dislocation, and their livers were removed immediately for glycogen determination. Liver glycogen content was determined quantitatively following a reported procedure.[80] Briefly livers were removed immediately after scarifying the animals and were homogenized by a blender (Ultra-Turrax, Janke & Kunkel GmbH & CoKg, Germany) with appropriate volume of 5% trichloroacetic acid over 5 min. The homogenate was centrifuged (Hettich Zentrifugen, Germany) at 3000 rpm for 5 min. The supernatant fluid was taken and filtered using acid-washed filter paper, and the residues were homogenized again with another volume of 5% trichloroacetic over 1–3 min to ensure better extraction of glycogen. The glycogen of 1.0 mL of this filtrate was precipitated using ethanol (95%, 5 mL), incubated in water bath (Raypa water bath, Germany) at 37–40 °C for 3 h, and centrifuged at 3000 rpm for 15 min. The clear liquid is gently decanted from the packed glycogen, and the tubes were allowed to drain in an inverted position for 10 min. The glycogen was dissolved in distilled water (2 mL) and mixed with 10 mL of the anthrone reagent (0.05% anthrone, 1.0% thiourea in 72% sulfuric acid). The mixture was incubated in boiling water over 30 min, and subsequently, the absorbance was measured at $\lambda = 620$ nm by spectrophotometer (Spectroscan 80D- UV–vis spectrophotometer). Blank and standard solutions were prepared by adding 10 mL of anthrone reagent to 2 mL of water and to 2 mL of glucose solution containing 0.1 mg of glucose in saturated benzoic acid, respectively. The liver glycogen content is estimated using the following formula:

$$\text{amount (mg) of glycogen liver tissue} = \frac{\text{DU}}{\text{DS}} \times \frac{\text{volume of extract (mL)}}{\text{weight of liver tissue (g)}} \times 0.09 \quad (7)$$

where DU is the absorbance of the unknown sample and DS is the absorbance of the standard.

Acknowledgment. This project was partially sponsored by the Faculty of Graduate Studies (M.Sc. Thesis of Mohamed A. S. Al-Ghussein). The authors thank the Deanship of Scientific Research and Hamdi-Mango Center for Scientific Research at the University of Jordan for their generous funds.

References

(1) Wagman, A. S.; Nuss, J. M. Current therapies and emerging targets for the treatment of diabetes. *Curr. Pharm. Des.* **2001**, *7*, 417–450.
(2) Ceriello, A. Postprandial hyperglycemia and diabetes complications: is it time to treat? *Diabetes* **2005**, *54*, 1–7.
(3) Saltiel, A. R. New perspectives into the molecular pathogenesis and treatment of type 2 diabetes. *Cell* **2001**, *104*, 517–529.
(4) Wagman, A. S.; Johnson, K. W.; Bussiere, D. E. Discovery and development of GSK3 inhibitors for the treatment of type 2 diabetes. *Curr. Pharm. Des.* **2004**, *10*, 1105–1137.
(5) Meijer, L.; Flajolet, M.; Greengard, P. Pharmacological inhibitors of glycogen synthase kinase 3. *Trends Pharmacol. Sci.* **2004**, *25*, 471–480.
(6) Bullock, W. H.; Magnuson, S. R.; Choi, S.; Gunn, D. E.; Rudolph, J. Prospects for kinase activity modulators in the treatment of diabetes and diabetic complications. *Curr. Top. Med. Chem.* **2002**, *2*, 915–938.
(7) Nikoulina, S. E.; Ciaraldi, T. P.; Mudaliar, S.; Mohideen, P.; Carter, L.; Henry, R. R. Potential role of glycogen synthase kinase-3 in skeletal muscle insulin resistance of type 2 diabetes. *Diabetes* **2000**, *49*, 263–271.
(8) Dokken, B. B.; Sloniger, J. A.; Henriksen, E. J. Acute selective glycogen synthase kinase-3 inhibition enhances insulin signaling in prediabetic insulin-resistant rat skeletal muscle. *Am. J. Physiol. Endocrinol. Metab.* **2005**, *288*, E1188–E1194.
(9) Eldar-Finkelman, H. Glycogen synthase kinase 3: an emerging therapeutic target. *Trends Mol. Med.* **2002**, *8*, 126–132.
(10) Morfini, G.; Szebenyi, G.; Elluru, R.; Ratner, N.; Brady, S. T. Glycogen synthase kinase 3 phosphorylates kinesin light chains and negatively regulates kinesin-based motility. *EMBO J.* **2002**, *21*, 281–293.
(11) Emoto, M.; Langille, S. E.; Czech, M. P. A role for kinesin in insulin-stimulated GLUT4 glucose transporter translocation in 3T3-L1 adipocytes. *J. Biol. Chem.* **2001**, *276*, 10677–10682.
(12) Cross, D. A.; Alessi, D. R.; Vandenheede, J. R.; McDowell, H. E.; Hundal, H. S.; Cohen, P. The inhibition of glycogen synthase kinase-3 by insulin or insulin-like growth factor 1 in the rat skeletal muscle cell line L6 is blocked by wortmannin, but not by rapamycin: evidence that wortmannin blocks activation of the mitogen-activated protein kinase pathway in L6 cells between Ras and Raf. *Biochem. J.* **1994**, *303* (Part 1), 21–26.
(13) Welsh, G. I.; Proud, C. G. Glycogen synthase kinase-3 is rapidly inactivated in response to insulin and phosphorylates eukaryotic initiation factor eIF-2B. *Biochem. J.* **1993**, *294* (Part 3), 625–629.
(14) Moule, S. K.; Welsh, G. I.; Edgell, N. J.; Foulstone, E. J.; Proud, C. G.; Denton, R. M. Regulation of protein kinase B and glycogen synthase kinase-3 by insulin and beta-adrenergic agonists in rat epididymal fat cells: activation of protein kinase B by wortmannin-sensitive and -insensitive mechanisms. *J. Biol. Chem.* **1997**, *272*, 7713–7719.
(15) Eldar-Finkelman, H.; Argast, G. M.; Foord, O.; Fischer, E. H.; Krebs, E. G. Expression and characterization of glycogen synthase kinase-3 mutants and their effect on glycogen synthase activity in intact cells. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 10228–10233.
(16) Summers, S. A.; Kao, A. W.; Kohn, A. D.; Backus, G. S.; Roth, R. A.; Pessin, J. E.; Birnbaum, M. J. The role of glycogen synthase kinase 3beta in insulin-stimulated glucose metabolism. *J. Biol. Chem.* **1999**, *274*, 17934–17940.
(17) Tang, Q. Q.; Gronborg, M.; Huang, H.; Kim, J. W.; Otto, T. C.; Pandey, A.; Lane, M. D. Sequential phosphorylation of CCAAT enhancer-binding protein beta by MAPK and glycogen synthase kinase 3beta is required for adipogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 9766–9771.

(18) Li, X.; Lu, F.; Tian, Q.; Yang, Y.; Wang, Q.; Wang, J. Z. Activation of glycogen synthase kinase-3 induces Alzheimer-like tau hyperphosphorylation in rat hippocampus slices in culture. *J. Neural Transm.* **2006**, *113*, 93–102.
(19) Hilioti, Z.; Gallagher, D. A.; Low-Nam, S. T.; Ramaswamy, P.; Gajer, P.; Kingsbury, T. J.; Birchwood, C. J.; Levchenko, A.; Cunningham, K. W. GSK-3 kinases enhance calcineurin signaling by phosphorylation of RCNs. *Genes Dev.* **2004**, *18*, 35–47.
(20) Shakoori, A.; Ougolkov, A.; Yu, Z. W.; Zhang, B.; Modarressi, M. H.; Billadeau, D. D.; Mai, M.; Takahashi, Y.; Minamoto, T. Deregulated GSK3beta activity in colorectal cancer: its association with tumor cell survival and proliferation. *Biochem. Biophys. Res. Commun.* **2005**, *334*, 1365–1373.
(21) Haefner, B. A model for NF-kappa B regulation by GSK-3 beta. *Drug Discovery Today* **2003**, *8*, 1062–1063.
(22) Droucheau, E.; Primot, A.; Thomas, V.; Mattei, D.; Knockaert, M.; Richardson, C.; Sallicandro, P.; Alano, P.; Jafarshad, A.; Baratte, B.; Kunick, C.; Parzy, D.; Pearl, L.; Doerig, C.; Meijer, L. *Plasmodium falciparum* glycogen synthase kinase-3: molecular model, expression, intracellular localisation and selective inhibitors. *Biochim. Biophys. Acta* **2004**, *1697*, 181–196.
(23) Woodgett, J. R. cDNA cloning and properties of glycogen synthase kinase-3. *Methods Enzymol.* **1991**, *200*, 564–577.
(24) Ali, A.; Hoeflich, K. P.; Woodgett, J. R. Glycogen synthase kinase-3: properties, functions, and regulation. *Chem. Rev.* **2001**, *101*, 2527–2540.
(25) Frame, S.; Cohen, P. GSK3 takes centre stage more than 20 years after its discovery. *Biochem. J.* **2001**, *359*, 1–16.
(26) Kaidanovich, O.; Eldar-Finkelman, H. The role of glycogen synthase kinase-3 in insulin resistance and type 2 diabetes. *Expert Opin. Ther. Targets* **2002**, *6*, 555–561.
(27) Shaw, P. C.; Davies, A. F.; Lau, K. F.; Garcia-Barcelo, M.; Waye, M. M.; Lovestone, S.; Miller, C. C.; Anderton, B. H. Isolation and chromosomal mapping of human glycogen synthase kinase-3 alpha and -3 beta encoding genes. *Genome* **1998**, *41*, 720–727.
(28) Dorronsoro, I.; Castro, A.; Martinez, A. Inhibitors of glycogen synthase kinase-3: future therapy for unmet medical needs. *Expert Opin. Ther. Pat.* **2002**, *12*, 1527–1536.
(29) Bridges, A. J. Therapeutic challenges of kinase and phosphatase inhibition and use in anti-diabetic strategy. *Biochem. Soc. Trans.* **2005**, *33*, 343–345.
(30) Schultz, P.; Ring, D. B.; Harrison, S. D.; Bray, A. M. Preparation of Purines as Inhibitors of Glycogen Synthase Kinase 3 (GSK3). Patent 97-US19472 19971010, 1998; Chiron Corp.
(31) Naerum, L.; Norskov-Lauritsen, L.; Olesen, P. H. Scaffold hopping and optimization towards libraries of glycogen synthase kinase-3 inhibitors. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 1525–1528.
(32) Martinez, A.; Castro, A.; Dorronsoro, I.; Alonso, M. Glycogen synthase kinase 3 (GSK-3) inhibitors as new promising drugs for diabetes, neurodegeneration, cancer, and inflammation. *Med. Res. Rev.* **2002**, *22*, 373–384.
(33) Martinez, A.; Alonso, M.; Castro, A.; Perez, C.; Moreno, F. J. First non-ATP competitive glycogen synthase kinase 3 beta (GSK-3beta) inhibitors: thiadiazolidinones (TDZD) as potential drugs for the treatment of Alzheimer's disease. *J. Med. Chem.* **2002**, *45*, 1292–1299.
(34) Smith, D. G.; Buffet, M.; Fenwick, A. E.; Haigh, D.; Ife, R. J.; Saunders, M.; Slingsby, B. P.; Stacey, R.; Ward, R. W. 3-Anilino-4-arylmaleimides: potent and selective inhibitors of glycogen synthase kinase-3 (GSK-3). *Bioorg. Med. Chem. Lett.* **2001**, *11*, 635–639.
(35) Coghlan, M. P.; Culbert, A. A.; Cross, D. A.; Corcoran, S. L.; Yates, J. W.; Pearce, N. J.; Rausch, O. L.; Murphy, G. J.; Carter, P. S.; Roxbee Cox, L.; Mills, D.; Brown, M. J.; Haigh, D.; Ward, R. W.; Smith, D. G.; Murray, K. J.; Reith, A. D.; Holder, J. C. Selective small molecule inhibitors of glycogen synthase kinase-3 modulate glycogen metabolism and gene transcription. *Chem. Biol.* **2000**, *7*, 793–803.
(36) Jaramillo, C.; de Diego, J. E.; Hamdouchi, C.; Collins, E.; Keyser, H.; Sanchez-Martinez, C.; del Prado, M.; Norman, B.; Brooks, H. B.; Watkins, S. A.; Spencer, C. D.; Dempsey, J. A.; Anderson, B. D.; Campbell, R. M.; Leggett, T.; Patel, B.; Schultz, R. M.; Espinosa, J.; Vieth, M.; Zhang, F.; Timm, D. E. Aminoimidazo[1,2-*a*]pyridines as a new structural class of cyclin-dependent kinase inhibitors. Part 1: Design, synthesis, and biological evaluation. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 6095–6099.
(37) Bax, B.; Carter, P. S.; Lewis, C.; Guy, A. R.; Bridges, A.; Tanner, R.; Pettman, G.; Mannix, C.; Culbert, A. A.; Brown, M. J.; Smith, D. G.; Reith, A. D. The structure of phosphorylated GSK-3beta complexed with a peptide, FRATtide, that inhibits beta-catenin phosphorylation. *Structure* **2001**, *9*, 1143–1152.
(38) Dajani, R.; Fraser, E.; Roe, S. M.; Young, N.; Good, V.; Dale, T. C.; Pearl, L. H. Crystal structure of glycogen synthase kinase 3 beta: structural basis for phosphate-primed substrate specificity and autoinhibition. *Cell* **2001**, *105*, 721–732.
(39) Ter Haar, E.; Coll, J. T.; Austen, D. A.; Hsiao, H. M.; Swenson, L.; Jain, J. Structure of GSK3beta reveals a primed phosphorylation mechanism. *Nat. Struct. Biol.* **2001**, *8*, 593–596.
(40) Dajani, R.; Fraser, E.; Roe, S. M.; Yeo, M.; Good, V. M.; Thompson, V.; Dale, T. C.; Pearl, L. H. Structural basis for recruitment of glycogen synthase kinase 3beta to the axin-APC scaffold complex. *EMBO J.* **2003**, *22*, 494–501.
(41) Bertrand, J. A.; Thieffine, S.; Vulpetti, A.; Cristiani, C.; Valsasina, B.; Knapp, S.; Kalisz, H. M.; Flocco, M. Structural characterization of the GSK-3beta active site using selective and non-selective ATP-mimetic inhibitors. *J. Mol. Biol.* **2003**, *333*, 393–407.
(42) Bhat, R.; Xue, Y.; Berg, S.; Hellberg, S.; Ormo, M.; Nilsson, Y.; Radesater, A. C.; Jerning, E.; Markgren, P. O.; Borgegard, T.; Nylof, M.; Gimenez-Cassina, A.; Hernandez, F.; Lucas, J. J.; Diaz-Nido, J.; Avila, J. Structural insights and biological effects of glycogen synthase kinase 3-specific inhibitor AR-A014418. *J. Biol. Chem.* **2003**, *278*, 45937–45945.
(43) Meijer, L.; Skaltsounis, A. L.; Magiatis, P.; Polychronopoulos, P.; Knockaert, M.; Leost, M.; Ryan, X. P.; Vonica, C. A.; Brivanlou, A.; Dajani, R.; Crovace, C.; Tarricone, C.; Musacchio, A.; Roe, S. M.; Pearl, L.; Greengard, P. GSK-3-selective inhibitors derived from Tyrian purple indirubins. *Chem. Biol.* **2003**, *10*, 1255–1266.
(44) Beeley, N. R. A.; Sage, C. GPCRs: an update on structural approaches to drug discovery. *Targets* **2003**, *2*, 19–25.
(45) Waszkowycz, B. In *Advances in Drug Discovery Techniques*; Harvey, A. L., Ed.; John Wiley & Sons: Chchester, U.K., 1998; pp 150–153.
(46) DePristo, M. A.; de Bakker, P. I.; Blundell, T. L. Heterogeneity and inaccuracy in protein structures solved by X-ray crystallography. *Structure* **2004**, *12*, 831–838.
(47) Gadakar, P. K.; Phukan, S.; Dattatreya, P.; Balaji, V. N. Pose prediction accuracy in docking studies and enrichment of actives in the active site of GSK-3beta. *J. Chem. Inf. Model.*, in press.
(48) *Advanced Seminars in CATALYST*, Proceedings of the 9th European CATALYST User Group Meeting; Frankfurt, Germany, March 24, 2006; Accelrys Inc.: San Diego, CA, 2006.
(49) *CATALYST 4.11 Users' Manual*; Accelrys Software Inc.: San Diego, CA, 2005.
(50) Clement, O. O.; Mehl, A. T. Pharmacophore Perception, Development, and Use in Drug Design. In *IUL Biotechnology Series*; Guner, F. O., Ed.; International University Line: La Jolla, CA, 2000; pp 71–84.
(51) Sprague, P. W.; Hoffmann, R. In *Computer Assisted Lead Finding and Optimization. Current Tools for Medicinal Chemistry*; Van de Waterbeemd, H., Testa, B., Folkers, G., Eds.; VHCA: Basel, Switzerland, 1997; pp 230–240.
(52) Barnum, D.; Greene, J.; Smellie, A.; Sprague, P. Identification of common functional configurations among molecules. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 563–571.
(53) Singh, J.; Chuaqui, C. E.; Boriack-Sjodin, P. A.; Lee, W. C.; Pontz, T.; Corbley, M. J.; Cheung, H. K.; Arduini, R. M.; Mead, J. N.; Newman, M. N.; Papadatos, J. L.; Bowes, S.; Josiah, S.; Ling, L. E. Successful shape-based virtual screening: the discovery of a potent inhibitor of the type I TGFbeta receptor kinase (TbetaRI). *Bioorg. Med. Chem. Lett.* **2003**, *13*, 4355–4359.
(54) Taha, M. O.; Qandil, A. M.; Zaki, D. D.; AlDamen, M. A. Ligand-based assessment of factor Xa binding site flexibility via elaborate pharmacophore exploration and genetic algorithm-based QSAR modeling. *Eur. J. Med. Chem.* **2005**, *40*, 701–727.
(55) Keller, P. A.; Bowman, M.; Dang, K. H.; Garner, J.; Leach, S. P.; Smith, R.; McCluskey, A. Pharmacophore development for corticotropin-releasing hormone: new insights into inhibitor activity. *J. Med. Chem.* **1999**, *42*, 2351–2357.
(56) Karki, R. G.; Kulkarni, V. M. A feature based pharmacophore for *Candida albicans* MyristoylCoA: protein *N*-myristoyltransferase inhibitors. *Eur. J. Med. Chem.* **2001**, *36*, 147–163.
(57) Taha, M. O.; Al-Bakri, A. G.; Zalloum, W. A. Discovery of potent inhibitors of pseudomonal quorum sensing via pharmacophore modeling and in silico screening. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 5902–5906.
(58) Taha, M. O.; Bustanji, Y.; Al-Bakri, A. G.; Yousef, A. M.; Zalloum, W. A.; Al-Masri, I. M.; Atallah, N. Discovery of new potent human protein tyrosine phosphatase inhibitors via pharmacophore and QSAR analysis followed by in silico screening. *J. Mol. Graphics Modell.* **2007**, *25*, 870–884.
(59) Katritzky, A. R.; Pacureanu, L. M.; Dobchev, D. A.; Fara, D. C.; Duchowicz, P. R.; Karelson, M. QSAR modeling of the inhibition of glycogen synthase kinase-3. *Bioorg. Med. Chem.* **2006**, *14*, 4987–5002.
(60) Witherington, J.; Bordas, V.; Gaiba, A.; Garton, N. S.; Naylor, A.; Rawlings, A. D.; Slingsby, B. P.; Smith, D. G.; Takle, A. K.; Ward, R. W. 6-Aryl-pyrazolo[3,4-*b*]pyridines: potent inhibitors of glycogen synthase kinase-3 (GSK-3). *Bioorg. Med. Chem. Lett.* **2003**, *13*, 3055–3057.

(61) Witherington, J.; Bordas, V.; Gaiba, A.; Naylor, A.; Rawlings, A. D.; Slingsby, B. P.; Smith, D. G.; Takle, A. K.; Ward, R. W. 6-Heteroarylpyrazolo[3,4-*b*]pyridines: potent and selective inhibitors of glycogen synthase kinase-3 (GSK-3). *Bioorg. Med. Chem. Lett.* **2003**, *13*, 3059–3062.
(62) Witherington, J.; Bordas, V.; Garland, S. L.; Hickey, D. M.; Ife, R. J.; Liddle, J.; Saunders, M.; Smith, D. G.; Ward, R. W. 5-Arylpyrazolo[3,4-*b*]pyridines: potent inhibitors of glycogen synthase kinase-3 (GSK-3). *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1577–1580.
(63) Sutter, J.; Güner, O.; Hoffmann, R.; Li, H.; Waldman, M. In *Pharmacophore Perception, Development, and Use in Drug Design*; Güner, O. F., Ed.; International University Line: La Jolla, CA, 2000; pp 501–511.
(64) Sheridan, R. P.; Kearsley, S. K. Why do we need so many chemical similarity search methods? *Drug Discovery Today* **2002**, *7*, 903–911.
(65) Kurogi, Y.; Guner, O. F. Pharmacophore modeling and three-dimensional database searching for drug design using catalyst. *Curr. Med. Chem.* **2001**, *8*, 1035–1055.
(66) Li, H.; Sutter, J.; Hoffmann, R. In *Pharmacophor Perception, Development, and Use in Drug Design*; Güner, O. F., Ed.; International University Line: La Jolla, CA, 2000; pp 173–189.
(67) Poptodorov, K.; Luu, T.; Langer, T.; Hoffmann, R. In *Methods and Principles in Medicinal Chemistry. Pharmacophores and Pharmacophores Searches*; Hoffmann, R. D., Ed.; Wiley-VCH: Weinheim, Germany, 2006; Vol. 2, pp 17–47.
(68) Fischer, R. *The Principle of Experimentation Illustrated by a Psycho-Physical ExpeHafner Publishing Co.*, 8th ed.; Hafner Publishing: New York, 1966; Chapter II.
(69) Bersuker, I. B.; Bahçeci, S.; Boggs, J. E. In *Pharmacophore Perception, Development, and Use in Drug Design*; Güner O. F.; International University Line: La Jolla, CA, 2000; pp 457–473.
(70) *CERIUS2, QSAR Users' Manual*, version 4.10; Accelrys Inc.: San Diego, CA, 2005; pp 43–88, 221–235.
(71) Ramsey, L. F.; Schafer, W. D. *The Statistical Sleuth*, 1st ed.; Wadsworth Publishing Company: Belmont, CA, 1997.
(72) Shojania, K.; Koehler, B. E.; Elliott, T. Hypoglycemia induced by hydroxychloroquine in a type II diabetic treated for polyarthritis. *J. Rheumatol.* **1999**, *26*, 195–196.
(73) Stoa-Birketvedt, G.; Paus, P. N.; Ganss, R.; Ingebretsen, O. C.; Florholmen, J. Cimetidine reduces weight and improves metabolic control in overweight patients with type 2 diabetes. *Int. J. Obes. Relat. Metab. Disord.* **1998**, *22*, 1041–1045.
(74) Czyzyk, A.; Lao, B.; Szutowski, M.; Szczepanik, Z.; Muszynski, J. Enhancement of alcohol-induced hypoglycaemia by H2-receptor antagonists. *Arzneim.-Forsch.* **1997**, *47*, 746–749.
(75) Graumlich, J. F.; Habis, S.; Avelino, R. R.; Salverson, S. M.; Gaddamanugu, M.; Jamma, K.; Aldag, J. C. Hypoglycemia in inpatients after gatifloxacin or levofloxacin therapy: nested case-control study. *Pharmacotherapy* **2005**, *25*, 1296–1302.
(76) Venkatachalam, C. M.; Jiang, X.; Oldfield, T.; Waldman, M. LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *J. Mol. Graphics Modell.* **2003**, *21*, 289–307.
(77) *CERIUS2 4.10 LigandFit User Manual*; Accelrys Inc.: San Diego, CA, 2005; pp 3–48.
(78) Gehlhaar, D. K.; Verkhivker, G. M.; Rejto, P. A.; Sherman, C. J.; Fogel, D. B.; Fogel, L. J.; Freer, S. T. Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: conformationally flexible docking by evolutionary programming. *Chem. Biol.* **1995**, *2*, 317–324.
(79) Cho, J. H.; Johnson, G. V. Glycogen synthase kinase 3beta phosphorylates tau at both primed and unprimed sites. Differential impact on microtubule binding. *J. Biol. Chem.* **2003**, *278*, 187–193.
(80) Carroll, N. V.; Longley, R. W.; Roe, J. H. The determination of glycogen in liver and muscle by use of anthrone reagent. *J. Biol. Chem.* **1956**, *220*, 583–593.
(81) Cronin, M. T. D.; Schultz, T. W. Pitfalls in QSAR. *J. Mol. Struct.: THEOCHEM* **2003**, *622*, 39–51.

JM7009765