Small-Molecule Interferon Inducers. Toward the Comprehension of the Molecular Determinants through Ligand-Based Approaches†

Ira Musmuca, Silvia Simeoni,* Antonia Caroli, and Rino Ragno*

Istituto Pasteur-Fondazione Cenci Bolognetti, Dipartimento di Chimica e Tecnologie del Farmaco, Sapienza Università di Roma, P. le A. Moro 5, 00185, Rome, Italy

Received February 23, 2009

Hepatitis C is becoming an increasingly common cause of mortality especially in the HIV-coinfected group. Due to the efficacy of interferon (IFN) based therapy in the treatment of hepatitis C, various compounds possessing IFN-inducing activity have been hitherto reported. In the present study, we describe how steric, electrostatic, hydrophobic, and hydrogen-bonding interactions might influence the biological activity of a published set of IFN inducers, using a three-dimensional quantitative structure−activity relationship (3-D QSAR) approach. Analyses were conducted evaluating different series of compounds structurally related to 8-hydroxyadenines and 1*H*-imidazo[4,5-*c*]quinolines. A ligand-based alignment protocol in combination with the GRID/GOLPE approach was applied: 62 3-D QSAR models were derived using different GRID probes and several training sets. Performed 3-D QSAR investigations proved to be of good statistical value displaying r^2, $q^2_{CV\text{-}LOO}$, and cross-validated SDEP values of 0.73, 0.61, 0.61 and 0.89, 0.64, 0.58 using the OH or the DRY probe, respectively. Additionally, the predictive performance was evaluated using an external test set of 20 compounds. Analyses of the resulting models led to the definition of a pharmacophore model that can be of interest to explain the observed affinities of known compounds as well as to design novel low molecular weight IFN inducers (IFNIs). To the best of our knowledge, this is the first 3-D QSAR application on IFN-inducing agents.

INTRODUCTION

Hepatitis C virus (HCV) is a small enveloped RNA virus, member of the Flaviviridae family, and is now recognized as the major cause of both parenterally transmitted and community-acquired non-A non-B hepatitis.[1,2]

HCV is the viral agent ascribed to induce the infection. Contact with blood products, the use of IV drugs, and/or sharing needles as well as getting a blood transfusion are the most common causes of transmission.[3] Chronic infection may progress to end-stage diseases such as liver cirrhosis and hepatocellular carcinoma within 10−15 years.[4] All current treatment protocols for hepatitis C are based upon interferon alpha (IFN-α) alone or in combination with Ribavirin.[5,6] Despite progress in therapeutic strategies such as the development of pegylated interferon,[7] sustained viral response (SVR) rates are still typically poor for genotype-1 infections that predominate in Europe, Japan, and the United States.[8] Additionally, the aforementioned therapy is often accompanied by significant problems:[6] (i) Recombinant IFNs, in fact, are genetically engineered, resulting in recognition by the body's immune system as "foreign". Consequently, the benefit of the treatment is strongly affected by the immune responses and the formation of neutralizing antibodies. (ii) The cost of treatment is very high because of the inordinately expensive IFN preparations. (iii) Administration of IFN by intramuscular or subcutaneous injection frequently causes pain and irritation at the site of injection. Due to the above-mentioned issues, the enhancement of the endogenous IFN release by the administration of orally available small molecular weight compounds is a promising approach in the anti-HCV therapy. Compounds possessing IFN-inducing activity have been already reported. Among them, small molecule IFN inducers (IFNIs) include tilorone,[9,10] BL-20803,[11] atabrine,[12] CP-28888,[13] ABMP,[14] DRB,[15] 10-carboxy-methyl-9-acridone,[16] bropirimine,[17] and imiquimod[3,18] (Chart 1).

Imiquimod (**1**) is a well-known IFN inducer (IFNI), clinically used in the United States for treatment of exophytic warts caused by the human papillomavirus.[3] As reported by Akira and co-workers,[19] imiquimod acts via Toll-like receptor 7 (TLR7) signaling pathway inducing the endogenous IFN release. However, due to the serious adverse effects that arose during clinical trials, imiquimod-based strategy against HCV[20] was abandoned. Even if analogues were discovered to be more effective than imiquimod, up to now no IFN inducers are clinically employed in antihepatitis C therapy.[21]

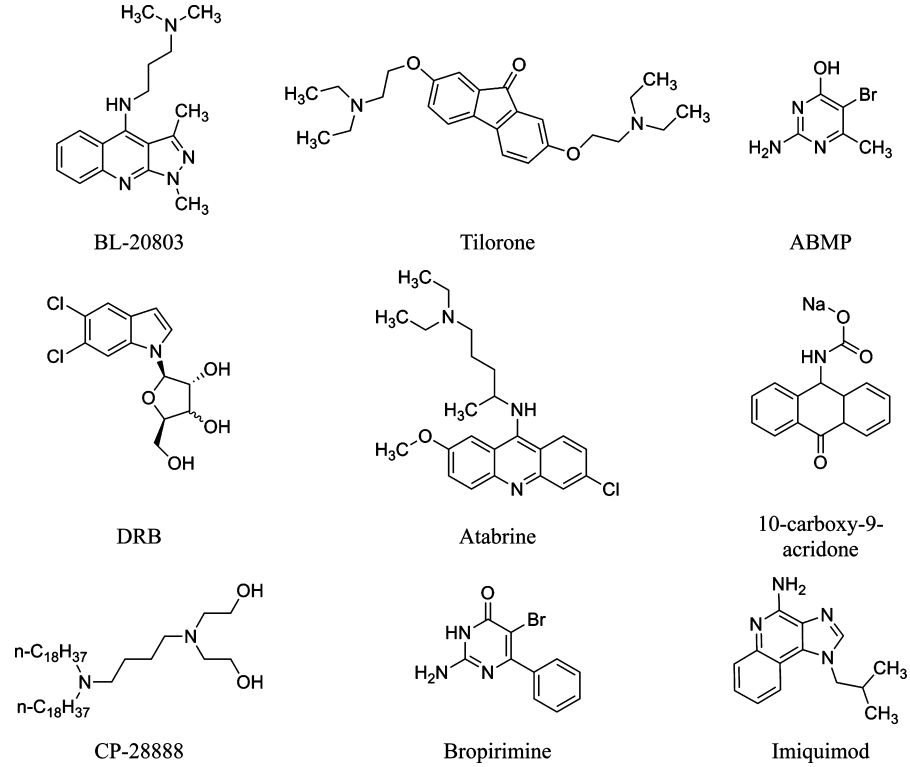
Three-dimensional quantitative structure−activity relationships (3-D QSAR) continues to be a vigorous field and an increasingly attractive discipline in the scientific community. 3-D QSAR methods such as comparative molecular field analysis (CoMFA)[22−24] and the GRID/GOLPE[25−28] combination have been successfully applied in many instances to guide the design of bioactive molecules.[29,30] Although several series of IFNIs have been disclosed, to the best of our knowledge, no 3-D QSAR studies have been performed on these classes of compounds. The present study is the application of the GRID/GOLPE procedure to different series

† Dedicated to Prof. Dr. Marino Artico on the occasion of his 75th birthday.

* Corresponding author phone +396-4991-3937 (R.R.), +396-4991-3152 (S.S.); fax: +396-4991-3627 (R.R.), +396-491-491 (S.S.); e-mail: rino.ragno@uniroma1.it (R.R.), silvia.simeoni@uniroma1.it (S.S.).

10.1021/ci900065a CCC: $40.75
Published on Web 06/05/2009

Chart 1. Chemical Structures of Small Molecular Weight IFNIs



of IFNIs. Robust and predictive 3-D QSAR models were built in order to gain insights into how steric, electrostatic, hydrophobic, and hydrogen-bonding interactions might influence their biological activity. The aim of this work was the development of useful models for future designing and forecasting the activity of untested IFNIs.

RESULTS

Structure−activity data for 176 IFNIs were considered in the present study. The whole set of adenines[3,31−34] and 1*H*-imidazo[4,5-*c*]quinoline[35] related structures (Figure 1) were selected from the literature. Compounds and biological activities are detailed in Tables S1−S6 in the Supporting Information.

IFNIs of Tables S1−S6 in the Supporting Information were first collected into five distinct groups according to both molecular scaffolds (1*H*-imidazo[4,5-*c*]quinolines and adenines) and substituents at positions 2, 8, and 9 on the adenine ring, providing the following: 1*H*-imidazo[4,5-*c*]-quinolines (group G1), 2-hydrogen/alkyl/cycloalkyl/aryl-8-hydroxy/mercapto-9-(substituted)arylalkyl adenines (group G2), 2-(alkyl)amino-8-hydroxy-9-benzyladenines (group G3), 2-alkylamino-8-hydroxy-9-(substituted)arylalkyl/alkyl/alkylhydroxy (group G4), and 2-alkoxy/alkylthio-8-hydroxy-9-benzyladenines (group G5). G1, G2, G3, G4, and G5 group chemical structures and biological activities are reported in Tables 1, 2, 3, 4, and 5, respectively. Activity data correspond to the minimum effective micromolar concentration required to observe IFN induction, as reported in the original papers. All derivatives[35] with activity values reported in micrograms per milliliter (μg/mL) were transformed into micromolar (μM) to have uniformity of data.

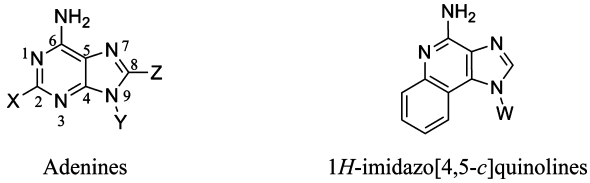
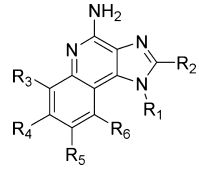


Figure 1. Adenine and 1*H*-imidazo[4,5-*c*]quinoline molecular scaffolds.

Compounds were then submitted to 3-D QSAR analyses with the aim of identifying and understanding structural properties and features able to influence their biological activity. In this context, starting with two-dimensional text formats as input and using the stand-alone version of PRODRG program[36,37] in conjunction with the molecular mechanics of GROMACS,[38] coordinates and molecular topologies of the ligands were generated.

In 3-D QSAR, a crucial role is played by the relative spatial orientation of structures within the grid.[39] To this purpose, all data set compounds have to be properly aligned. The generation of 3-D conformations and alignment of compounds is often a difficult procedure particularly when the ligands are flexible and/or the accurate position/orientation in the binding site is unknown. In our case, experimental bound conformation of ligands (G1−G5 groups) are not yet available since no crystal structure of the putative receptor TLR7 cocrystallized with an IFNI has been reported. To overcome this lack of knowledge, the crystal structure of imiquimod[40] was used as a starting point for the alignment procedure. The first step was to superimpose the most active compound (**175**, MEC = 0.001 μM) together with the most rigid analogue (**17**) to the crystal structure of imiquimod by means of Surflex-Sim,[41] which generates molecular alignments and hypotheses of bioactive ligand conformations for 3-D approaches. Using a morphological similarity[42] algorithm based on the shapes of molecules, H-bonding, and electrostatic properties, Surflex optimizes the alignment of molecules and maximizes their three-dimensional similarity.

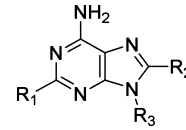
Table 1. Group G1: 1*H*-Imidazo[4,5-*c*]quinolines



entry	R_1	R_2	R_3	R_4	R_5	R_6	MEC[a]
1	$CH_2CH(CH_3)_2$	H	H	H	H	H	2.08
2[c]	$CH_2CH(OH)CH_2OH$	H	H	H	H	H	9.68
3	CH_3	H	H	H	H	H	2.52
4	CH_2CH_3	H	H	H	H	H	2.36
5	$CH_2CH_2CH_3$	H	H	H	H	H	2.21
6	$CH(CH_3)_2$	H	H	H	H	H	2.21
7	$CH_2CH_2CH_2CH_3$	H	H	H	H	H	2.08
8	$CH_2(CH_2)_3CH_3$	H	H	H	H	H	1.97
9	$CH_2(CH_2)_4CH_3$	H	H	H	H	H	1.86
10	$CH_2(CH_2)_5CH_3$	H	H	H	H	H	1.77
11	CH_2CH_2OH	H	H	H	H	H	2.19
12[c]	$CH_2CH(CH_3)OH$	H	H	H	H	H	1.03
13	$CH_2C(CH_3)_2OH$	H	H	H	H	H	1.95
14	CH_2Ph	H	H	H	H	H	0.36
15	CH_2CH_2Ph	H	H	H	H	H	1.73
16	$CH_2CH_2CH_2CH_2Ph$	H	H	H	H	H	0.32
17[b]	H	H	H	H	H	H	2.71
18	$CH_2CH(CH_3)_2$	H	H	H	H	H	0.2
19	$CH_2CH(CH_3)_2$	CH_2CH_3	H	H	H	H	0.19
20	$CH_2CH(CH_3)_2$	$CH_2CH_2CH_3$	H	H	H	H	0.18
21[b]	$CH_2CH(CH_3)_2$	$CH_2CH_2CH_2CH_3$	H	H	H	H	0.03
22	$CH_2CH(CH_3)_2$	$CH_2(CH_2)_3CH_3$	H	H	H	H	0.16
23	$CH_2CH(CH_3)_2$	$CH_2(CH_2)_4CH_3$	H	H	H	H	1.54
24	$CH_2CH(CH_3)_2$	$CH_2(CH_2)_5CH_3$	H	H	H	H	2.95
25[b]	$CH_2CH(CH_3)_2$	CH_2OPh	H	H	H	H	0.29
26[b]	$CH_2CH(CH_3)_2$	CH_2Ph	H	H	H	H	0.15
27	$CH_2CH(CH_3)_2$	CH_2CH_2Ph	H	H	H	H	2.9
28	$CH_2C(CH_3)_2OH$	CH_3	H	H	H	H	0.18
29	$CH_2C(CH_3)_2OH$	CH_2CH_3	H	H	H	H	0.18
30	$CH_2C(CH_3)_2OH$	$CH_2CH(CH_3)_2$	H	H	H	H	0.03
31	Ph	$CH_2CH_2CH_2CH_3$	H	H	H	H	1.58
32	$CH_2CH(CH_3)_2$	H	OH	H	H	H	3.9
33	$CH_2CH(CH_3)_2$	H	H	CH_3	H	H	1.97
34	$CH_2CH(CH_3)_2$	H	H	OCH_3	H	H	3.7
35	$CH_2CH(CH_3)_2$	H	H	OH	H	H	1.95

[a] Minimum effective concentration (μM). [b] Compounds used in the test set. [c] Compounds containing a stereogenic center.

Table 2. Group G2: Adenine Derivatives Substituted at Positions C-2, C-8, and N-9



entry	R_1	R_2	R_3	MEC[a]
36	H	OH	CH_2-Ph	10.0
37	H	SH	CH_2-Ph	10.0
38	H	OH	α-naphthylphenyl	10.0
39	H	OH	$CH_2C_6H_4$(4-F)	10.0
40	H	OH	$CH_2C_6H_4$(4-Cl)	10.0
41	H	OH	$CH_2C_6H_4$(4-CH_3)	10.0
42	H	OH	$CH_2C_6H_4$(2-OCH_3)	10.0
43	H	OH	$CH_2C_6H_4$(3-OCH_3)	10.0
44	H	OH	$CH_2C_6H_4$(4-OCH_3)	10.0
45	CH_3	OH	CH_2-Ph	1.0
46	C_2H_5	OH	CH_2-Ph	1.0
47	C_3H_7	OH	CH_2-Ph	1.0
48[b]	C_4H_9	OH	CH_2-Ph	0.01
49	Ph	OH	CH_2-Ph	10.0
50	Pen	OH	CH_2-Ph	0.03
51	Hep	OH	CH_2-Ph	1.0
52	*i*-Pr	OH	CH_2-Ph	0.1
53[b]	c-Pr	OH	CH_2-Ph	0.1
54	*i*-Bu	OH	CH_2-Ph	1.0
55	c-Hex	OH	CH_2-Ph	0.3
56	CH_2-c-Pen	OH	CH_2-Ph	0.1
57	CH_2-c-Hex	OH	CH_2-Ph	0.1
58[b]	3-pyridyl	OH	CH_2-Ph	10.0
59	benzyl	OH	CH_2-Ph	0.1
60	4-F-benzyl	OH	CH_2-Ph	0.1
61	CF_3	OH	CH_2-Ph	0.1
62	CHF_2	OH	CH_2-Ph	0.1
63	CF_2CF_3	OH	CH_2-Ph	0.3
64	$(CF_2)_2CF_3$	OH	CH_2-Ph	0.3
65	$(CH_2)_2CF_3$	OH	CH_2-Ph	0.1

[a] Minimum effective concentration (μM). [b] Compounds used in the test set.

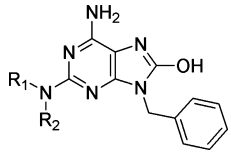
Aligned Imiquimod, **175**, and **17** were therefore used as a template for the alignment of the other molecules (Figure 2).

The next step was the calculation and application of the molecular interaction fields (MIFs) describing the spatial variation of the interaction energy between a molecular target and a chosen probe. This task was accomplished by the use of the GRID[43−45] program. Briefly, the protocol involves the computation of MIFs for positions of the probe at points on a rectilinear grid superimposed on the target. Thus, the MIFs provide information on where the favorable and the unfavorable locations for the probe around the target are. To this aim, OH2, OH, DRY, O, N1, and combined DRY + OH GRID probes were in turn employed.

The most widespread application of MIFs is likely the derivation of 3-D QSARs by diverse approaches such as CoMFA[22−24] or GRID/GOLPE.[25−28] Since in 3-D QSAR the biological activity may be seen as a function of the physicochemical characteristics of the compounds of interest, the need to convert such numerical data to useful information has led to the development of methodologies that rely on statistics and applied mathematics. Among them, an efficient tool deriving from the principal component regression is the partial least squares (PLS) methodology,[25,46] in which the original variables are replaced by a small set of linear combinations thereof. Generated latent variables are then used for multivariate regression in order to maximize the communality of predictor and response variable blocks. Additional attractive features of PLS are the ability to correctly work even when interdescriptor correlations exist and/or to reduce the risk of chance correlations.

To this point, due to the presence of asymmetric carbon atoms bearing compounds in groups G1, G3, and G5 (Tables 1, 3, and 5), our work proceeded considering both *R* and *S* enantiomers. As a consequence, eight training sets (G1-*R*, G1-*S*, G2, G3-*R*, G3-*S*, G4, G5-*R*, and G5-*S*) were obtained. The self-consistencies of training sets G1−G5(*R*, *S*) were evaluated developing[47] 48 3-D QSAR submodels (see models A−AV in Table S7 in the Supporting Information) distinguished by values of the correlation coefficient (r^2), cross-validated correlation coefficient (q^2), and cross-validated standard deviation errors of prediction[48] ($SDEP_{CV}$) falling in the ranges 0.65−0.99, 0.40−0.87, and 0.23−0.56, respectively. Then, to estimate the predictive power of each submodel, the training sets of the other submodels were in turn used as external test sets. Thus, whereas a good self-consistency was observed for each submodel, some of them were affected by a low prediction capability ($SDEP_{ext}$) (see Table S8 in the Supporting Information). The latter results highlight the intrinsic limitation of any statistical modeling or machine learning approach (extrapolation problem) and thus apply also to the above A−AV 3-D QSAR models.

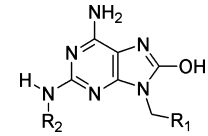
Table 3. Group G3: 2-(Alkyl)amino-8-hydroxy-9-benzyladenine Derivatives



entry	R_1	R_2	MEC[a]
66	*i*-Pr	H	1.0
67	*i*-Bu	H	0.1
68[c]	*s*-Bu	H	0.1
69	cyclohexyl	H	0.1
70	$(CH_2)_2OH$	H	1.0
71	$(CH_2)_3OH$	H	1.0
72	$(CH_2)_4OH$	H	1.0
73	$(CH_2)_3NMe_2$	H	10.0
74	$(CH_2)_2OMe$	H	0.1
75	$(CH_2)_3OMe$	H	0.1
76[b]	Bn	H	0.01
77[b]	Ph	H	1.0
78	$(CH_2)_2Ph$	H	1.0
79	3-MeO-Bn	H	1.0
80	4-MeO-Bn	H	1.0
81	4-Me-Bn	H	0.1
82	4-NMe_2-Bn	H	1.0
83	4-F-Bn	H	1.0
84	2-pyridylmethyl	H	1.0
85	3-pyridylmethyl	H	1.0
86[b]	4-pyridylmethyl	H	0.1
87[b]	Bn	Me	10.0
88	$-(CH_2)_2NMe(CH_2)_2-$		1.0
89	$-(CH_2)_2O(CH_2)_2-$		1.0
90	$-(CH_2)_2CHMe(CH_2)_2-$		1.0
91	H	H	10.0
92	CH_3	H	1.0
93	C_2H_5	H	1.0
94	C_3H_7	H	0.1
95	C_4H_9	H	0.1
96	C_5H_{11}	H	0.1

[a] Minimum effective concentration (μM). [b] Compounds used in the test set. [c] Compounds containing a stereogenic center.

Table 4. Group G4: 2-Alkylamino-8-hydroxy-9-Substituted Derivatives



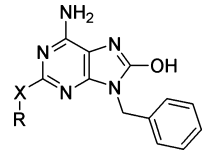
entry	R_1	R_2	MEC[a]
97	2-Cl-Ph	Bu	0.03
98	3-Cl-Ph	Bu	0.03
99	4-Cl-Ph	Bu	0.01
100	2-Me-Ph	Bu	0.03
101	3-Me-Ph	Bu	0.03
102	4-Me-Ph	Bu	0.01
103	4-F-Ph	Bu	0.01
104	4-CF_3-Ph	Bu	0.01
105	4-NO_2-Ph	Bu	0.01
106	4-OMe-Ph	Bu	0.01
107	4-OBn-Ph	Bu	0.1
108	4-*n*-Bu-Ph	Bu	0.1
109	4-*t*-Bu-Ph	Bu	0.03
110	4-Ph-Ph	Bu	0.1
111	4-OH-Ph	Bu	0.01
112[b]	3,4-Cl_2-Ph	Bu	0.01
113	3,4-F_2-Ph	Bu	0.01
114	2-Cl-4,5-OCH_2O-Ph	Bu	0.1
115	1-naphthyl	Bu	0.01
116[b]	2-naphthyl	Bu	0.03
117[b]	5-thiophen(2-Cl)	Bu	0.01
118	CMe_2OH	Bu	0.1
119	*i*-Pr	Bu	0.1
120	cyclohexyl	Bu	0.03
121	2-Py	Bu	0.03
122	3-Py	Bu	0.003
123	4-Py	Bu	0.01
124	2-pyrazine	Bu	0.03
125	4-F-Ph	$MeO(CH_2)_2$	0.03
126	4-F-Ph	$PhCH_2$	0.1
127	4-F-Ph	4-Py-CH_2	0.01
128	2-Me-3-Py	Bu	0.01
129	5-Me-3-Py	Bu	0.003
130	6-Me-3-Py	Bu	0.003
131	2-Cl-3-Py	Bu	0.01
132	4-Cl-3-Py	Bu	0.1
133	6-Cl-3-Py	Bu	0.003
134	6-OMe-3-Py	Bu	0.003
135	6-NMe-3-Py	Bu	0.003
136	2-Cl-6-Me-3-Py	Bu	0.01
137	2,6-Cl-3-Py	Bu	0.03

[a] Minima effective concentration (μM). [b] Compounds used in the test set.

In order to overcome extrapolation problems, two global molecular sets (GMSs) were obtained (GMS-*R* and GMS-*S*) by merging the G1-*R*, G2, G3-*R*, G4, and G5-*R* and G1-*S*, G2, G3-*S*, G4, and G5-*S* series. GMSs were therefore split into a training set of 156 inducers and a test set of 20 inducers (see Experimental Section). Both training and test set compounds were selected to cover the range of bioactivities and structures of the GMSs. By the use of aforementioned probes and performing the same procedure, further 12 3-D QSAR models resulted (M1-*R*−M6-*R*, M1-*S*−M6-*S*). In the early stage, the internal predictive power of each individual 3-D QSAR model was checked by "leave-one-out" (LOO) cross-validations (CV). The latter indicated that the number of optimal principal components was two for all models except for the DRY-based ones (see Table 6, models M3-*R* and M3-*S*). Although statistical coefficients were not as good as those obtained for the previous 48 submodels (A−AV in Table S7 in the Supporting Information), the r^2, $q^2_{CV\text{-}LOO}$, and $SDEP_{CV\text{-}LOO}$ values resulted in the ranges 0.62−0.90, 0.47−0.64, and 0.58−0.71, respectively, and can be still considered statistically significant as the $q^2_{CV\text{-}LOO}$ values are greater than the threshold values of 0.25, as suggested by Cramer[49] and in the majority of cases satisfy the more stringent rules depicted by Tropsha.[50] Moreover, different cross-validated PLS analyses were carried out to establish the models' robustness (leave-two-out, LTO; leave-some-out, LSO-5; leave-half-out, LHO), confirming the LOO principal component optimal number. It could be argued that the conventional r^2 values for models M1-*R*, M1-*S*, M2-*R*, M2-*S*, M4-*R*, M4-*S*, M5-*R*, and M5-*S* are not fully satisfying, ranging between 0.62 (M1-*R*) and 0.73 (M2-*R* and M2-*S*). Nevertheless, by definition of the PLS paradigm, they are highly containing informative and relatively highly robust models in terms of cross-validation. As a general trend, the (*R*) configuration training set showed similar to or slightly better statistical profiles than models containing the (*S*) enantiomers.

Analyzing values in Table 6, it is not surprising that the DRY model shows larger r^2, as the number of latent dimensions is five versus two for the other models. The DRY model has the largest r^2, the highest dimensionality, but not significantly better q^2 or SDEPs. However, information contained in the DRY-based models is of peculiar interest

Table 5. Group G5: 2-Alkoxy/2-alkylthio-8-hydroxy-9-benzyladenine Derivatives



entry	X	R	MEC[a]
138	S	*i*-Pr	0.1
139	S	*i*-Bu	0.1
140[c]	S	*s*-Bu	0.05
141	S	cyclohexyl	0.1
142	S	cyclohexylmethyl	0.1
143	S	Ph	1.0
144	S	Bn	0.1
145	S	$(CH_2)_2Me$	0.1
146	S	3-chlorobenzyl	1.0
147[b]	S	4-chlorobenzyl	1.0
148	S	3-methoxybenzyl	0.1
149	S	4-methoxybenzyl	0.1
150	S	CH_2OMe	0.01
151	S	$(CH_2)_2OMe$	0.1
152	S	$(CH_2)_2OEt$	0.1
153	S	$(CH_2)_3OEt$	0.1
154	S	$(CH_2)_2NMe_2$	10.0
155	S	$(CH_2)_3NMe_2$	0.1
156	S	$(CH_2)_2OH$	0.1
157	S	$(CH_2)_3OH$	0.1
158	S	$(CH_2)_4OH$	0.1
159	O	$(CH_2)_2OMe$	0.01
160	O	$(CH_2)_2OEt$	0.1
161	O	$(CH_2)_3OEt$	1.0
162[b]	O	$(CH_2)_2OH$	1.0
163	O	$(CH_2)_3OH$	1.0
164	S	H	10.0
165	S	CH_3	0.1
166	S	C_2H_5	0.1
167	S	C_3H_7	0.01
168[b]	S	C_4H_9	0.01
169	S	C_5H_{11}	10.0
170	S	$(O)C_4H_9$	0.01
171	O	H	0.1
172	O	CH_3	0.1
173	O	C_2H_5	0.1
174	O	C_3H_7	0.01
175	O	C_4H_9	0.001
176	O	C_5H_{11}	0.01

[a] Minimum effective concentration (μM). [b] Compounds used in the test set. [c] Compounds containing a stereogenic center.

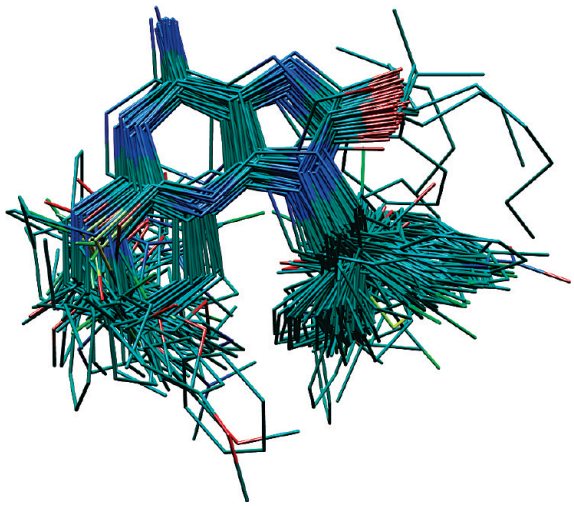


Figure 2. Superimposition of the full data set compounds according to Surflex-Sim alignment protocol. For the sake of clarity, hydrogen atoms are omitted.

and can be very useful for detecting hydrophobic patches; thus we decided to use it for interpretations.

Achiral models were also investigated. Compounds **2**, **12**, **68**, and **140** were kept out from the training set (152 IFNIs), which was in turn analyzed with OH and DRY probes. Resulting 3-D QSAR models did not show statistically significant differences when compared with models M1-*R*, M1-*S*, M3-*R*, and M3-*S*. The omitted compounds were added to the original test set and used for external validation.

It is likely that the tiny differences observed for the (*R*) and (*S*) models stem from the independent generation of conformers/alignments, or maybe slightly different positions of the aligned molecules in the grid. Statistical results for achiral models are detailed in Table S10 in the Supporting Information.

Prediction Capability of the 3-D QSAR Models. Along with the internal validation (cross-validation), the predictive capabilities of models (M1-*R*−M6-*R*) and (M1-*S*−M6-*S*) were estimated by the preselected test set consisting of 20 compounds (see Tables 1−5). The predictive ability of models containing (*R*) enantiomers showed a slightly better profile than those including (*S*) enantiomers. As represented in Table 6, the $SDEP_{ext}$ values were comprised between 0.98 and 1.09. Due to the higher statistical profile and the lower $SDEP_{ext}$, models obtained via OH and DRY probes (M2-*R* and M3-*R* in Table 6) were selected for a deeper analysis and graphical interpretations.

DISCUSSION

3-D QSAR Model Interpretations. One of the most important features of 3-D QSAR is the graphical representation of the model, making its interpretation easier. In fact, by using GOLPE[25−28] software, one may display the final model by several available options. Among them, PLS pseudocoefficient, present field, and activity contribution plots are very useful. The basic plot is the PLS pseudocoefficient plot, which allows the grid point visualization of a determined molecule/probe interaction energy level, providing whole training set graphical information. As a result, the contour coefficient maps indicate those areas in which the model has found a high correlation between the ligand−probe interaction energy and the bioactivity.

In Figure 3 are shown the PLS coefficient plots for models M2-*R* (OH probe) and M3-*R* (DRY probe). The most important areas are located in four regions (herein referred to as A, B, C, and D) which encompass the adenine and quinoline scaffolds. In particular, a positive PLS coefficient area (colored in yellow) is located around the C-2 position of the adenine scaffold corresponding to the quinoline 5a carbon atom (region A). Due to the GRID force field characteristics of the OH probe, A reflects areas in which the ligands could make either electrostatic or steric interactions. To properly allocate the nature of region A, an additional 3-D QSAR model was built using the C3 probe (mainly steric information).[25] Statistical results of the latter model (see Table S9 in the Supporting Information) were shown to be in good agreement with the previous models. In fact, comparing the C3 probe PLS coefficients with those of M2-*R*, a full overlap was detected (compare Figure S1 in the Supporting Information with Figure 3), highlighting as mainly steric or hydrophobic the characteristics of region

Table 6. Statistical Results for (M1-*R*−M6-*R*) and (M1-*S*−M6-*S*) 3-D QSAR Global Models Obtained from Diverse GOLPE PLS Analysis[a]

			LOO			LTO			LSO-5			LHO			external validation
M	P	r^2	q^2	SDEP	PC	q^2	SDEP	PC	q^2	SDEP	PC	q^2	SDEP	PC	$SDEP_{ext}$
M1-*R*	OH2	0.62	0.47	0.71	2	0.46	0.71	2	0.43	0.73	2	0.39	0.76	2	1.08
M1-*S*	OH2	0.62	0.48	0.70	2	0.48	0.70	2	0.47	0.70	2	0.44	0.72	2	1.08
M2-*R*	OH	0.73	0.61	0.61	2	0.61	0.61	2	0.60	0.61	2	0.56	0.64	2	1.05
M2-*S*	OH	0.73	0.62	0.60	2	0.62	0.60	2	0.61	0.61	2	0.57	0.64	2	1.04
M3-*R*	DRY	0.89	0.64	0.58	5	0.64	0.58	5	0.61	0.60	5	0.55	0.65	5	0.98
M3-*S*	DRY	0.90	0.64	0.58	5	0.64	0.58	5	0.61	0.60	5	0.54	0.66	5	1.04
M4-*R*	N1	0.69	0.58	0.65	2	0.56	0.65	2	0.55	0.65	2	0.52	0.67	2	1.09
M4-*S*	N1	0.69	0.58	0.64	2	0.56	0.65	2	0.55	0.65	2	0.52	0.67	2	1.06
M5-*R*	O	0.72	0.60	0.61	2	0.60	0.61	2	0.60	0.62	2	0.57	0.64	2	1.05
M5-*S*	O	0.69	0.55	0.65	2	0.55	0.65	2	0.55	0.65	2	0.52	0.67	2	1.08
M6-*R*	DRY + OH	0.70	0.56	0.64	2	0.56	0.64	2	0.55	0.65	2	0.53	0.67	2	1.09
M6-*S*	DRY + OH	0.69	0.57	0.64	2	0.56	0.64	2	0.56	0.64	2	0.53	0.66	2	1.08

[a] M, model name; P, GRID probe; LOO, leave-one-out cross-validation; LTO, leave-two-out cross-validation; LSO-5, leave-some-out cross-validation using five groups; LHO, leave-half-out cross-validation; r^2, conventional square correlation coefficient; q^2, cross-validation correlation coefficient; SDEP, cross-validated standard error of prediction; PC, optimal number of principal components; $SDEP_{ext}$, standard error of prediction for the external test set.

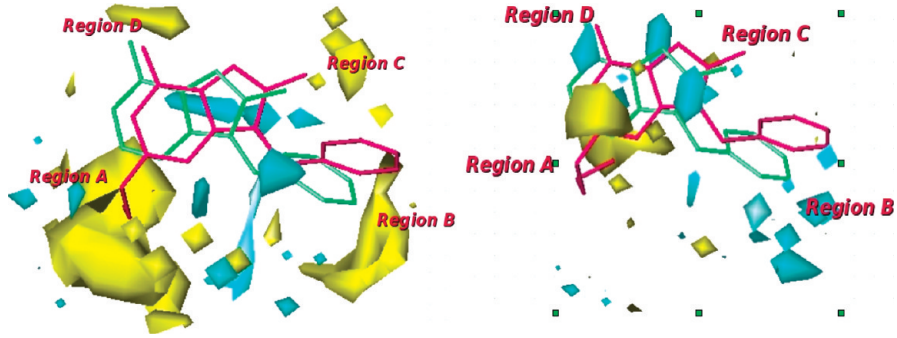


Figure 3. GRID/GOLPE PLS coefficient contour maps for the M2-*R* and M3-*R* 3-D QSAR models (M2-*R* contour levels: 0.0049 yellow, −0.0049 cyan; M3-*R* contour levels: 0.004 52 yellow, −0.004 52 cyan). To aid interpretation only the highest active compound (**175** in red) and one of the lowest active compounds (**36** in green) are shown. For the sake of clarity hydrogen atoms are omitted.

A. Moreover, for chemical interpretation one must remember that substituents being able to favorably interact with an OH probe led to negative GRID energies. Hence, cyan contours indicate attractive interactions, which will increase biological affinity, while repulsive, unfavorable interactions will lower activity. In contrast, yellow contours in coefficient maps (see Figure 3, OH probe) suggest that repulsive interactions (space filling) increase affinity, while attractive interactions decrease activity. While the PLS coefficient plot provides whole training set graphical information, the activity contribution plots (resulting from the multiplication of PLS coefficients by the actual field values of each object) help focusing on spatial regions that are individually important for the selected molecules. In the bottom of Figure 4 (model M2-*R* activity contribution plot for **175**, the most active IFNI), yellow polyhedrons appear in the region surrounding the 2-butyloxy group, suggesting that this substituent might increase IFN-inducing activity. On the other hand, in the activity contribution plot of compound **36**, one of the less active IFNIs lacking any substituents in position C-2, in region A some yellow polyhedrons are converted into cyan-colored ones (negative activity contribution). This confirms the steric nature of region A that somehow has to be filled with neutral groups.

Region B is located in proximity of substituents on adenine N-9 or imidazoquinoline N-1 positions. The majority of compounds are characterized by hydrophobic bulky alkyl, aryl or heteroaryl groups at this position. For instance, the most active compound of our training set (**175**) displays a benzyl substituent at this position. As shown in the right side of Figure 3, some cyan polyhedrons (negative PLS coefficients) are close to the hydrophobic moieties, and since DRY probe gives rise only to negative field values (Figure S2 in the Supporting Information), the presence of a hydrophobic group in this region positively contributes to the biological activity. Furthermore, to properly explain the influence of the regiochemistry of the substituents at various positions on the benzene ring, we examined PLS coefficient and activity contribution plots of para-, meta-, and ortho-substituted compounds (Figure 5). As clearly depicted, para-substituted compound **99** (colored in orange) displays a yellow polyhedron in proximity of the 4-chloro group even in the PLS coefficient plot or in the activity contribution one (see also Figure 6). In contrast, compounds **97** and **98**, respectively, ortho- and meta-chloro substituted, did not orient their groups into this yellow contour (PLS coefficient plots). Moreover, with reference to the activity contribution plots of both derivatives **97** and **98** (Figure 6), the deficiency of the yellow polyhedron (positive activity contribute) in the vicinity of the chloro groups is obvious, somehow confirming[31] the importance for the biological activity of a correct regiochemistry of the substituents in this area.

Regarding regions C and D of Figure 3, as previously reported[3,35] a hydroxyl group at adenine C-8 position (region C) and an amino group at adenine C-6 or imidazoquinoline C-4 position (region D) are essential for IFN-inducing activity. Despite the reported structure−activity relationships (SARs), in the PLS coefficient plots of Figure 3 only some

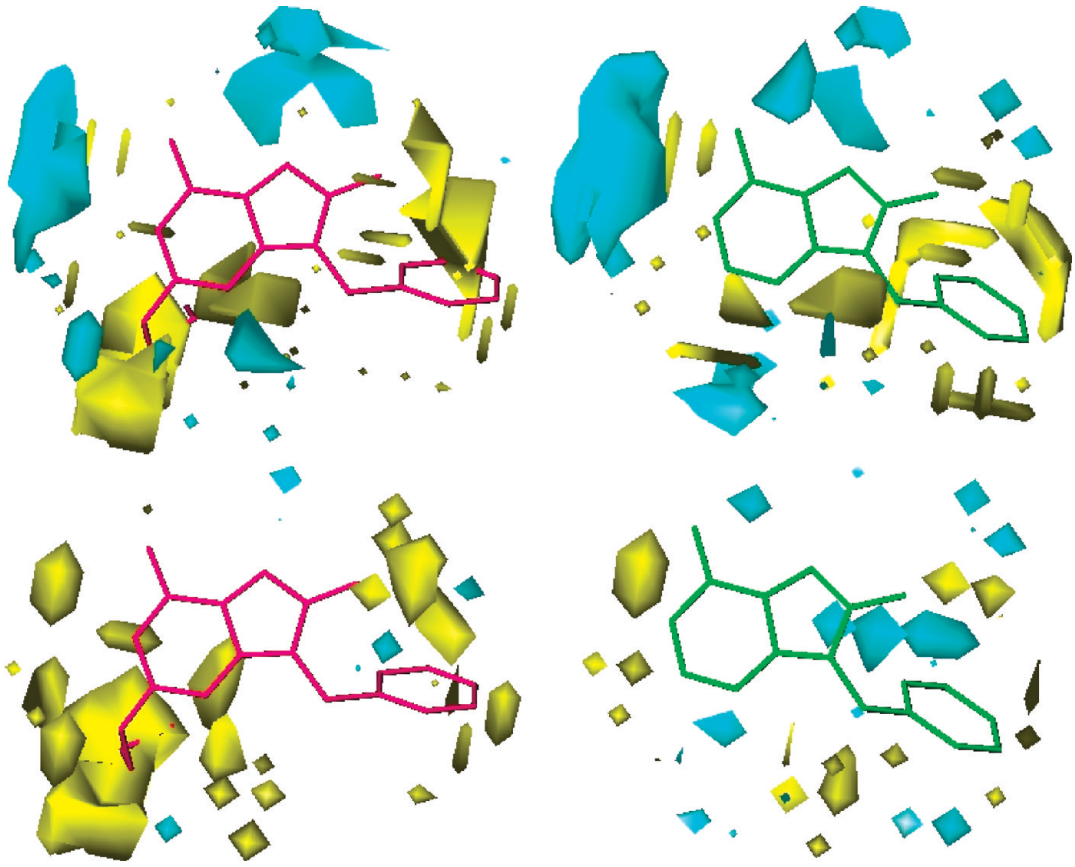


Figure 4. Top: graphical representation of present field contour maps (contour levels 3.7787 yellow; −2.3306 cyan) for compounds **175** (in red) and **36** (in green). Bottom: activity contribution plots (contour levels 0.0230 yellow; −0.0230 cyan) for the same compounds.

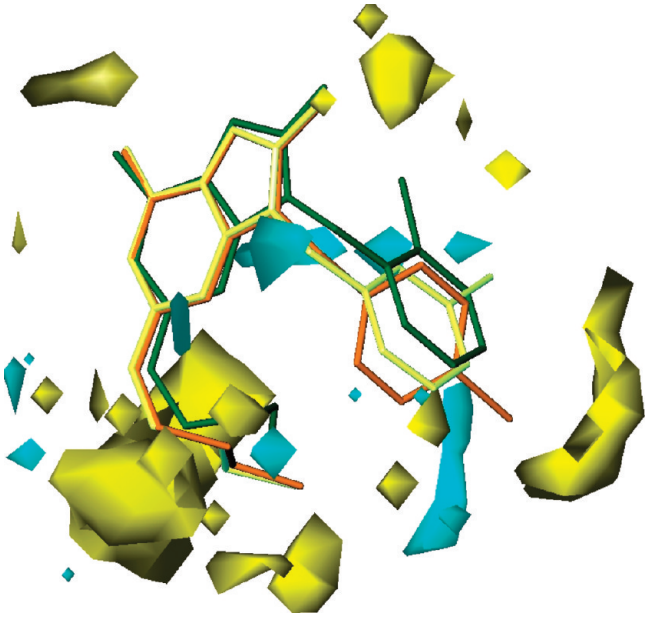


Figure 5. GRID/GOLPE PLS coefficient contour maps for the M2-*R* 3-D QSAR model (contour levels 0.0049 yellow, −0.0049 cyan). Compounds **97**, **98**, and **99** are depicted in green, yellow, and orange, respectively. For the sake of clarity hydrogen atoms are omitted.

small yellow polyhedrons are present in proximity to these positions. However, due to the constant presence of both the C-8 hydroxyl substituent and the C-6/C-4 amino groups the fractional factorial design (FFD) selection led to the exclusion of the majority of grid points around these groups, resulting in a small number of polyhedrons. The fewness of selected lattice points in these areas does not mean that these groups have no influence on the biological activity, but that the majority of compounds included in the training set exert similar influences in that particular area. Nevertheless, since visible in Figure 3 only, small polyhedrons suggest the influence of hydrogen bonding in these positions.

CONCLUSIONS

In this work we report the first 3-D QSAR study on a training set of 156 IFN inducers, including 8-hydroxyadenines and 1*H*-imidazo[4,5-*c*]quinoline derivatives. Twelve 3-D QSAR final models were built and analyzed. All models displayed good statistical coefficient values, and using an external test set, model M2-*R* proved to be predictive showing low standard error of predictions. Results from 3-D QSAR model interpretation let us individuate the most important areas around the molecules and are also in full agreement with the previously reported SARs. Analysis in terms of principal interaction points between the most actives and the GRID probes led to the definition of a plausible pharmacophoric model (Figure 7). To this, around the adenine and quinoline scaffold four regions seem particularly important for the modulation of the IFN-inducing activity: (a) a fillable steric pocket (S); (b) a hydrophobic area (Hy); (c) an acceptor hydrogen bonding region close to the adenine C-6/quinoline C4 amino groups (HA); (d) a polarized area, possibly with donator hydrogen bonding characteristics close to adenine C-8. These features are fully in agreement with several anti-HCV derivatives able to stimulate interferon

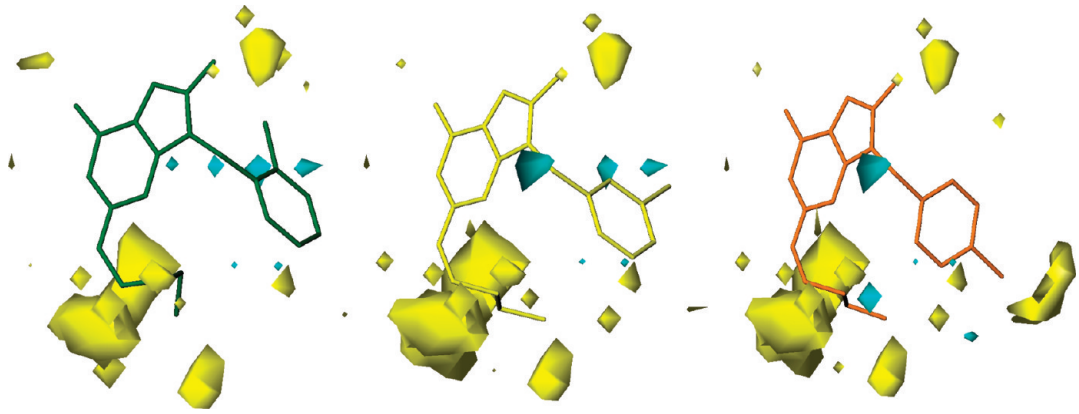


Figure 6. Graphical representation of the activity contribution plots (contour levels 0.0270 yellow, −0.0270 cyan) for compounds **97**, **98**, and **99** are depicted in green, yellow, and orange, respectively. For the sake of clarity hydrogen atoms are omitted.

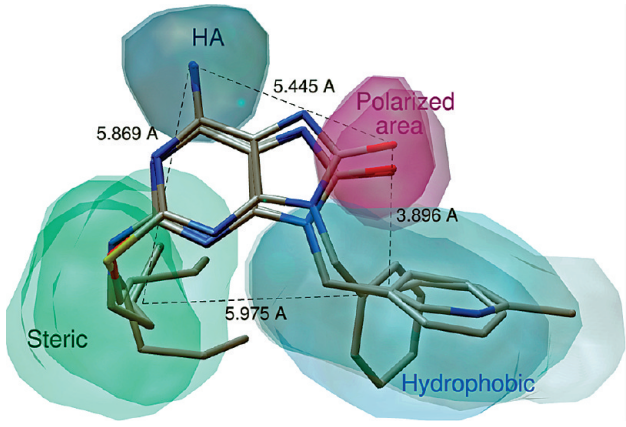


Figure 7. Graphical representation of the pharmacophoric model obtained from the analysis of the 3-D QSAR models.

release in PBMC (peripheral blood mononuclear cell), recently reported by Czarniecki[51] and Chong in a 2008 patent.[52]

EXPERIMENTAL SECTION

Data Sets. For the investigation, 221 IFNIs were selected from the literature.[3,31−35] According to the original papers, in vitro bioactivities were expressed as minimum effective concentration (MEC, μM) the concentration of compounds required for more than 1 IU/mL induction of interferon. These numbers were then transformed to the decadic log values of the molar MEC which were used as dependent variables in the present 3-D QSAR analyses. All IFNIs with no defined MEC values were excluded from the work, leading to a global set of 176 compounds (data set, Tables S1−S6 in the Supporting Information).

IFNIs of Tables S1−S6 in the Supporting Information were first collected into five distinct groups according to both molecular scaffolds (1*H*-imidazo[4,5-*c*]quinolines and adenines) and substituents at positions 2, 8, and 9 on the adenine ring, providing the following: 1*H*-imidazo[4,5-*c*]-quinolines (group G1), 2-hydrogen/alkyl/cycloalkyl/aryl-8-hydroxy/mercapto-9-(substituted)arylalkyl adenines (group G2), 2-(alkyl)amino-8-hydroxy-9-benzyladenines (group G3), 2-alkylamino-8-hydroxy-9-(substituted)arylalkyl/alkyl/alkylhydroxy (group G4), and 2-alkoxy/alkylthio-8-hydroxy-9-benzyladenines (group G5). G1−G5 group chemical structures and biological activities are reported in Tables 1−5.

For the final 3-D QSAR models, the training set comprises a series of 156 interferon inducers, whose chemical structures belong to the 8-hydroxyadenine and 1*H*-imidazo[4,5-*c*]-quinoline classes. The remaining 20 compounds were used as an external test set. Either training or test sets were selected so that they span the whole descriptor (chemical) space occupied by the entire data set and cover the same range of bioactivities. Even if the number of test set compounds is only 20, they do contain the maximal diversity in terms of molecular structures and bioactivities. All derivatives[35] with activity values reported in micrograms per milliliter were transformed into micromolar to have uniformity of data. Attempts to use concentrations of IFN induced by compounds as dependent variables for our 3-D QSAR models were unsuccessful (data not shown).

Molecular Modeling and Alignments. All molecules were built starting from ASCII text using the stand-alone version of PRODRG[36,37] in conjunction with GROMACS[38] suite and UCSF Chimera[53] as graphical user interface on a 3 GHz AMD CPU equipped IBM compatible workstation with the SUSE 9.0 version of the Linux operating system. SURFLEX[41,42] software was employed for the alignment procedure. Molecular graphics images were produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco.

GRID Calculations. The GRID[43] program (version 22) was used to describe the aligned molecular structures. Due to the presence of asymmetric carbon atoms in compounds of groups G1, G3, and G5 (Tables 1, 3, and 5), both (*R*) and (*S*) enantiomers were considered. As a result, eight initial training sets (G1-*R*, G1-*S*; G2; G3-*R*, G3-*S*; G4; G5-*R* and G5-*S*) were obtained. Since no X-ray data for the TLR7 receptor are yet available, a random probe selection was accomplished. Calculations on the 48 obtained submodels were carried out using probes OH2 (water), OH (hydroxyl group bonded to an aromatic system), DRY (hydrophobic), O (sp^2 carbonyl oxygen), N1 (neutral flat NH), and combined DRY + OH. The energy calculations were implemented using a 1 Å spacing between the grid points. Each set of calculated interaction energies contained in the resulting three-dimensional matrix, arranged as one-dimensional vector, was used as input for the program GOLPE.[43] The *xyz* coordinates (in angstroms) of the grid rectangular box used

for the computation are $X_{min}/X_{max} = -6.57/10.95$, $Y_{min}/Y_{max} = -6.83/13.24$, and $Z_{min}/Z_{max} = -10.82/10.19$.

The model graphical interpretation was conducted considering OH and DRY probes. To aid graphical interpretation, further 3-D QSAR models were built using the C3 GRID probe.

GOLPE Analyses. PLS models were calculated with GOLPE 4.6.0 running on PC IBM compatible hardware running the Linux operating system. To measure the goodness of the model, the statistical indices r^2, q^2, and SDEP were employed using different cross-validation methods.

$$r^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - Y_{fit})^2}$$

$$q^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - Y_{CV})^2}$$

$$SDEP = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - Y_{CV})^2}{N}}$$

where Y_i = experimental value, Y_{fit} = recalculated value; $\bar{Y}$ = mean value, Y_{CV} = predicted value, and N = number of experiments.

Variable Preselection. The resulting probe−target interaction energies for each compound were first unfolded to produce one-dimensional vector variables and then assembled in the so-called **X** matrix. The matrix was pretreated using a cutoff of 5 kcal/mol in order to produce a more symmetrical distribution of energy values. To obtain a model with better predictive capability, variable selection was operated by zeroing values with absolute values smaller than 0.01 kcal/mol and removing variables with standard deviation below 0.05. In addition, variables taking only two and three distributions were also removed. Tentative use of other settings in the variable preselection did not lead to any improvement of the models.

Smart Region Definition (SRD). A number of seeds (1000) were selected using a D-optimal design criterion in the weight space. Structural differences between different molecules in the series will be reflected in groups of variables, and therefore groups were generated around each seed in the 3-D space. Variables with a critical distance cutoff of 2 Å to the seeds were included in the groups. If two neighboring groups (with a distance smaller than 10 Å) contained the same information, the groups were collapsed. The groups were used in the variable selection procedure replacing the original variables. The effect of the groups on the predictability was evaluated, and groups instead of individual variables were removed from the data file. Tentative use of other settings in the SRD did not lead to any improvement of the models.

Variable Selection. The effect of the grouped variables on the predictive ability was evaluated using a fractional factorial design (FFD) procedure. A number of reduced models (twice the number of variables) were built removing some of the variables according to the FFD design. The effect of dummy variables (20%) on the predictive ability was calculated, and only variables with a positive effect on the predictability larger than the effect of the average dummy variable were included in the final model. During FFD the predictive ability of the generated matrices was evaluated by cross-validation, using five random groups and 100 randomizations. Weights were recalculated after object exclusion. Variables which were found to be uncertain were retained. The FFD selection was repeated until the r^2 and q^2 values did not increase significantly for all models. In the FFD selection the cross-validation was conducted using five random groups and a maximum of five principal components.

Cross-Validation. All 3-D QSAR models were tested by cross-validation using LOO (leave-one-out), LTO (leave-two-out), LSO-5 (leave-some-out with five random groups), and LHO (leave-half-out) methods with 100 randomizations. The optimum number of principal components used in any 3-D QSAR model was chosen on the basis of the different cross-validated PLS analyses, setting to five the maximum number of principal components to be extracted. For all 3-D QSAR models in the present study, the optimal number of principal components (PC = 2 for OH2, OH, N1, O, and DRY + OH based models and PC = 5 for DRY based model) refers to the ones where $SDEP_{CV}$ assumes the minimum value and the cross-validation correlation coefficient q^2 assumes the maximum value. Attempts to increase the maximum principal components during either CVs or in the above FFD variable selection did not lead to any substantially different model.

ACKNOWLEDGMENT

Many thanks are due to Prof. Gabriele Cruciani and Prof. Sergio Clementi (Molecular Discovery and MIA srl) for the use of GOLPE program in their chemometric laboratory (University of Perugia, Italy) and for having provided the GRID program.

Supporting Information Available: Molecular structures and biological activities of the compounds as reported in the original papers (Tables S1−S6), various statistical results (Tables S7−S10), and C3-based 3-D QSAR GOLPE contour maps. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Bartenschlager, R.; Lohmann, V. Replication of hepatitis C virus. *J. Gen. Virol.* **2000**, *81*, 1631–1648.
(2) Penin, F.; Dubuisson, J.; Rey, F. A.; Moradpour, D.; Pawlotsky, J. M. Structural biology of hepatitis C virus. *Hepatology* **2004**, *39*, 5–19.
(3) Hirota, K.; Kazaoka, K.; Niimoto, I.; Kumihara, H.; Sajiki, H.; Isobe, Y.; Takaku, H.; Tobe, M.; Ogita, H.; Ogino, T.; Ichii, S.; Kurimoto, A.; Kawakami, H. Discovery of 8-hydroxyadenines as a novel type of interferon inducer. *J. Med. Chem.* **2002**, *45*, 5419–5422.
(4) Saito, I.; Miyamura, T.; Ohbayashi, A.; Harada, H.; Katayama, T.; Kikuchi, S.; Watanabe, Y.; Koi, S.; Onji, M.; Ohta, Y.; et al. Hepatitis C virus infection is associated with the development of hepatocellular carcinoma. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 6547–6549.
(5) Koike, K. Antiviral treatment of hepatitis C: present status and future prospects. *J. Infect. Chemother.* **2006**, *12*, 227–232.
(6) Chevaliez, S.; Pawlotsky, J. M. Interferon-based therapy of hepatitis C. *Adv. Drug Delivery Rev.* **2007**, *59*, 1222–1241.
(7) Hayashi, N.; Takehara, T. Antiviral therapy for chronic hepatitis C: past, present, and future. *J. Gastroenterol.* **2006**, *41*, 17–27.

(8) Dymock, B. W. Emerging therapies for hepatitis C virus infection. *Expert Opin. Emerging Drugs* **2001**, *6*, 13–42.
(9) Stringfellow, D. A.; Glasgow, L. A. Tilorone hydrochloride: an oral interferon-inducing agent. *Antimicrob. Agents Chemother.* **1972**, *2*, 73–78.
(10) Mayer, G. D.; Krueger, R. F. Tilorone hydrochloride: mode of action. *Science* **1970**, *169*, 1214–1215.
(11) Siminoff, P.; Bernard, A. M.; Hursky, V. S.; Price, K. E. BL-20803, a new, low-molecular-weight interferon inducer. *Antimicrob. Agents Chemother.* **1973**, *3*, 742–743.
(12) Glaz, E. T.; Szolgay, E.; Stoger, I.; Talas, M. Antiviral activity and induction of interferon-like substance by quinacrine and acranil. *Antimicrob. Agents Chemother.* **1973**, *3*, 537–541.
(13) Hoffman, W. W.; Korst, J. J.; Niblack, J. F.; Cronin, T. H. N. N-dioctadecyl-N', N'-bis(2-hydroxyethyl) propanediamine: antiviral activity and interferon stimulation in mice. *Antimicrob. Agents Chemother.* **1973**, *3*, 498–502.
(14) Nichol, F. R.; Weed, S. D.; Underwood, G. E. Stimulation of murine interferon by a substituted pyrimidine. *Antimicrob. Agents Chemother.* **1976**, *9*, 433–439.
(15) Tamm, I.; Sehgal, P. B. A comparative study of the effects of certain halogenated benzimidazole ribosides on RNA synthesis, cell proliferation, and interferon production. *J. Exp. Med.* **1977**, *145*, 344–356.
(16) Taylor, J. L.; Schoenherr, C. K.; Grossberg, S. E. High-yield interferon induction by 10-carboxymethyl-9-acridanone in mice and hamsters. *Antimicrob. Agents Chemother.* **1980**, *18*, 20–26.
(17) Wierenga, W.; Skulnick, H. I.; Stringfellow, D. A.; Weed, S. D.; Renis, H. E.; Eidson, E. E. 5-substituted 2-amino-6-phenyl-4(3H)-pyrimidinones. Antiviral- and interferon-inducing agents. *J. Med. Chem.* **1980**, *23*, 237–239.
(18) Dockrell, D. H.; Kinghorn, G. R. Imiquimod and resiquimod as novel immunomodulators. *J. Antimicrob. Chemother.* **2001**, *48*, 751–755.
(19) Hemmi, H.; Kaisho, T.; Takeuchi, O.; Sato, S.; Sanjo, H.; Hoshino, K.; Horiuchi, T.; Tomizawa, H.; Takeda, K.; Akira, S. Small antiviral compounds activate immune cells via the TLR7 MyD88-dependent signaling pathway. *Nat. Immunol.* **2002**, *3*, 196–200.
(20) Strominger, N. L.; Brady, R.; Gullikson, G.; Carpenter, D. O. Imiquimod-elicited emesis is mediated by the area postrema, but not by direct neuronal activation. *Brain Res. Bull.* **2001**, *55*, 445–451.
(21) Kurimoto, A.; Tobe, M.; Ogita, H.; Ogino, T.; Takaku, H.; Ichii, S.; Kawakami, H.; Isobe, Y. Prodrugs of 9-benzyl-8-hydroxy-2-(2-hydroxyethylthio)adenine: potent interferon inducing agents in monkeys. *Chem. Pharm. Bull. (Tokyo)* **2004**, *52*, 466–469.
(22) Clark, M.; Cramer, R. D., III; Jones, D. M.; Patterson, D. E.; Simeroth, P. E. Comparative molecular field analysis (CoMFA). 2. Toward its use with 3D-structural databases. *Tetrahedron Comput. Methodol.* **1990**, *3*, 47–59.
(23) Cramer, R. D., III; Patterson, D. E.; Bunce, J. D. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959–5967.
(24) Cramer, R. D., III; Wold, S. B. Comparative molecular field analysis (CoMFA). PCT Int. Appl. US 88-237491, 1991.
(25) 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*, 2589–25601.
(26) Pastor, M.; Cruciani, G.; Watson, K. A. A strategy for the incorporation of water molecules present in a ligand binding site into a three-dimensional quantitative structure−activity relationship analysis. *J. Med. Chem.* **1997**, *40*, 4089–4102.
(27) Ortiz, A. R.; Pastor, M.; Palomer, A.; Cruciani, G.; Gago, F.; Wade, R. C. Reliability of comparative molecular field analysis models: effects of data scaling and variable selection using a set of human synovial fluid phospholipase A2 inhibitors. *J. Med. Chem.* **1997**, *40*, 1136–1148.
(28) Cruciani, G.; Clementi, S.; Pastor, M. GOLPE-guided region selection. *Perspect. Drug Discovery Des.* **1998**, *12/13/14*, 71–86.
(29) Sippl, W.; Contreras, J. M.; Parrot, I.; Rival, Y. M.; Wermuth, C. G. Structure-based 3D QSAR and design of novel acetylcholinesterase inhibitors. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 395–410.
(30) Audouze, K.; Nielsen, E. O.; Peters, D. New series of morpholine and 1,4-oxazepane derivatives as dopamine D4 receptor ligands: synthesis and 3D-QSAR model. *J. Med. Chem.* **2004**, *47*, 3089–3104.
(31) Isobe, Y.; Kurimoto, A.; Tobe, M.; Hashimoto, K.; Nakamura, T.; Norimura, K.; Ogita, H.; Takaku, H. Synthesis and biological evaluation of novel 9-substituted-8-hydroxyadenine derivatives as potent interferon inducers. *J. Med. Chem.* **2006**, *49*, 2088–2095.
(32) Isobe, Y.; Tobe, M.; Ogita, H.; Kurimoto, A.; Ogino, T.; Kawakami, H.; Takaku, H.; Sajiki, H.; Hirota, K.; Hayashi, H. Synthesis and structure-activity relationships of 2-substituted-8-hydroxyadenine derivatives as orally available interferon inducers without emetic side effects. *Bioorg. Med. Chem.* **2003**, *11*, 3641–3647.
(33) Kurimoto, A.; Ogino, T.; Ichii, S.; Isobe, Y.; Tobe, M.; Ogita, H.; Takaku, H.; Sajiki, H.; Hirota, K.; Kawakami, H. Synthesis and structure-activity relationships of 2-amino-8-hydroxyadenines as orally active interferon inducing agents. *Bioorg. Med. Chem.* **2003**, *11*, 5501–5508.
(34) Kurimoto, A.; Ogino, T.; Ichii, S.; Isobe, Y.; Tobe, M.; Ogita, H.; Takaku, H.; Sajiki, H.; Hirota, K.; Kawakami, H. Synthesis and evaluation of 2-substituted 8-hydroxyadenines as potent interferon inducers with improved oral bioavailabilities. *Bioorg. Med. Chem.* **2004**, *12*, 1091–1099.
(35) Gerster, J. F.; Lindstrom, K. J.; Miller, R. L.; Tomai, M. A.; Birmachu, W.; Bomersine, S. N.; Gibson, S. J.; Imbertson, L. M.; Jacobson, J. R.; Knafla, R. T.; Maye, P. V.; Nikolaides, N.; Oneyemi, F. Y.; Parkhurst, G. J.; Pecore, S. E.; Reiter, M. J.; Scribner, L. S.; Testerman, T. L.; Thompson, N. J.; Wagner, T. L.; Weeks, C. E.; Andre, J. D.; Lagain, D.; Bastard, Y.; Lupu, M. Synthesis and structure-activity-relationships of 1H-imidazo[4,5-c]quinolines that induce interferon production. *J. Med. Chem.* **2005**, *48*, 3481–3491.
(36) Schuttelkopf, A. W.; van Aalten, D. M. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 1355–1363.
(37) van Aalten, D. M.; Bywater, R.; Findlay, J. B.; Hendlich, M.; Hooft, R. W.; Vriend, G. PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 255–262.
(38) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **1995**, *91*, 43–56.
(39) Folkers, G.; Merz, A.; Rognan, D. CoMFA: Scope and Limitations. In *3D QSAR in Drug Design: Theory Methods and Applications*; Kubinyi, H., Ed.; ESCOM Science Publishers B.V.: Leiden, The Netherlands, 1993; pp 583−618.
(40) Zhao, B.; Rong, Y. Z.; Huang, X. H.; Shen, J. S. Experimental and theoretical study on the structure and electronic spectra of imiquimod and its synthetic intermediates. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 4942–4946.
(41) Jain, A. N. Ligand-based structural hypotheses for virtual screening. *J. Med. Chem.* **2004**, *47*, 947–961.
(42) Jain, A. N. Morphological similarity: a 3D molecular similarity method correlated with protein-ligand recognition. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 199–213.
(43) Goodford, P. J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* **1985**, *28*, 849–857.
(44) Wade, R. C.; Clark, K. J.; Goodford, P. J. Further development of hydrogen bond functions for use in determining energetically favorable binding sites on molecules of known structure. 1. Ligand probe groups with the ability to form two hydrogen bonds. *J. Med. Chem.* **1993**, *36*, 140–147.
(45) Wade, R. C.; Goodford, P. J. Further development of hydrogen bond functions for use in determining energetically favorable binding sites on molecules of known structure. 2. Ligand probe groups with the ability to form more than two hydrogen bonds. *J. Med. Chem.* **1993**, *36*, 148–156.
(46) Geladi, P.; Kowalski, B. R. Partial least-squares regression: a tutorial. *Anal. Chim. Acta* **1986**, *185*, 1–17.
(47) Ragno, R.; Simeoni, S.; Valente, S.; Massa, S.; Mai, A. 3-D QSAR Studies on Histone Deacetylase Inhibitors. A GOLPE/GRID Approach on Different Series of Compounds. *J. Chem. Inf. Model.* **2006**, *46*, 1420–1430.
(48) Cruciani, G.; Baroni, M.; Clementi, S.; Costantino, G.; Riganelli, D.; Skagerberg, B. Predictive ability of regression models. Part I. Standard deviation of prediction errors (SDEP). *J. Chemom.* **1992**, *6*, 335–346.
(49) Clark, M.; Cramer, R. D., III. The probability of chance correlation using partial least squares (PLS). *Quant. Struct.-Act. Relat.* **1993**, *12*, 137–145.
(50) Golbraikh, A.; Tropsha, A. Beware of q^2! *J. Mol. Graph. Model.* **2002**, *20*, 269−276.
(51) Czarniecki, M. Small Molecule Modulators of Toll-like Receptors. *J. Med. Chem.* **2008**, *51*, 6621–6626.
(52) Chong, L. S.; Desai, M. C.; Gallagher, B.; Graupe, M.; Halcomb, R. L.; Yang, H.; Zhang, J. R. Preparation of purine and thiadeazapurine phosphonate derivatives as modulators of toll-like receptor 7. PCT Int. Appl. WO 2008005555, 2008.
(53) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera−a visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–12.

CI900065A