Affinity and Selectivity of Matrix Metalloproteinase Inhibitors: A Chemometrical Study from the Perspective of Ligands and Proteins[†]

Hans Matter* and Wilfried Schwab

Hoechst Marion Roussel, Chemical Research, D-65926 Frankfurt am Main, Germany

A novel strategy to understand affinity and selectivity for enzyme inhibitors using information from ligands and target protein 3D structures is described. It was applied to 2-arylsulfonyl-1,2,3,4-tetrahydro-isoquinoline-3-carboxylates and -hydroxamates as inhibitors of the matrix metalloproteinases MMP-3 (stromelysin-1) and MMP-8 (human neutrophil collagenase). As the first step, consistent and predictive 3D-QSAR models were derived using CoMFA, CoMSIA, and GRID/Golpe approaches, leading to the identification of binding regions where steric, electronic, or hydrophobic effects are important for affinity. These models were validated using multiple analyses using two or five randomly chosen cross-validation groups and randomizations of biological activities. Second, 3D-QSAR models were derived based on the affinity ratio IC_{50}(MMP-8)/IC_{50}(MMP-3), allowing the identification of key ligand determinants for selectivity toward one of both enzymes. In addition to this ligands' view, the third step encompasses a chemometrical approach based on principal component analysis (PCA) of multivariate GRID descriptors to uncover the major differences between both protein binding sites with respect to their GRID probe interaction pattern. The resulting information, based on the accurate knowledge of the target protein 3D structures, led to a consistent picture in good agreement with experimentally observed differences in selectivity toward MMP-8 or MMP-3. The interpretation of all three classes of statistical models leads to detailed SAR information for MMP inhibitors, which is in agreement with available data for binding site topologies, ligand affinities, and selectivities. Thus the combined chemical analyses provide guidelines and accurate activity predictions for designing novel, selective MMP inhibitors.

1. Introduction

Matrix metalloproteinases (MMPs) are a family of zinc-containing endopeptidases involved in turnover, maintenance, and remodeling of tissue. They show proteolytic activities against most constituents of the extracellular matrix, like fibronectin, laminin, basement membrane, and interstitial collagens.[1] This activity is controlled by endogenous tissue inhibitors of metalloproteinases (TIMPs) and nonspecific α2-macroglobulins,[2] while any disruption causes tissue degradation. Hence, MMPs are important therapeutic targets for treatment of cancer,[3] arthritis,[4] joint destruction,[5] and Alzheimer's disease,[6] leading to a considerable interest in non-peptidic inhibitors[7] for oral administration. In particular MMP-3 and MMP-8 attack proteoglycan aggregates (*aggrecans*) as major cartilage components at different cleavage sites under pathological conditions in arthritis.[4]

Available MMP-3[8] and MMP-8[9−12] X-ray structures earlier led us to a rational design of 2-arylsulfonyl-1,2,3,4-tetrahydro-isoquinolines with complementary substituents to the MMP S1' pocket and hydroxamates or carboxylates as Zn^{2+} binding groups. A detailed 3D-QSAR analysis of 90 analogues as MMP-8 inhibitors is reported in a parallel publication[13] using 3D structures

of the MMP-8 catalytic domain[12,14] as template for alignment.

Here we describe a novel strategy for understanding affinity and selectivity for MMP-3 and -8 inhibitors, taking 3D information from both ligands and protein 3D structures into account. Structural information for protein cavities of MMP-8 and MMP-3 and a consistent set of inhibitors active over several orders of magnitude at both targets stimulated this study to investigate the requirements for selectivity toward one or the other enzyme by 3D-QSAR techniques. First, predictive 3D-QSAR models were derived for MMP-3 ligands in addition to analyses for understanding their MMP-8 affinity.[13] Those models serve to identify essential binding regions, where steric, electronic, or hydrophobic effects are important to explain ligand affinity. Selectivity of potential drugs against only a single biological target, on the other hand, is a major requirement for the treatment of chronic diseases such as arthritis to avoid side effects. This caused additional 3D-QSAR analyses using the ratio IC_{50}(MMP-8)/IC_{50}(MMP-3) as selectivity measure, revealing key determinants for selectivity. Finally, this information from the perspective of ligands was complemented by a study to quantitatively describe the differences of the target biomolecules. This third analysis step encompasses a chemometrical approach based on molecular field interaction descriptors as computed by the program GRID[15,16] to uncover major differences of both protein binding sites with respect to their GRID probe interaction pattern.[17,18]

With accurate 3D structures of both proteins it becomes possible to quantitatively describe their differences for designing selective ligands using principal component analysis (PCA) as a statistical technique, resulting in a consistent picture in good agreement with experimental data.

Comparative molecular field analysis (CoMFA)[19−21] and related methods[22,23] are accepted techniques to correlate molecular property fields with biological activities. As those models provide insight into protein binding requirements, they are important to potentially enhance the potency of new analogues. The application of those models for a quantitative prediction of binding affinities is possible, if the new candidate molecules fall within the range of the model.[13] After defining an alignment rule, electrostatic and steric interaction energies between each ligand and a probe atom on a predefined grid were computed for CoMFA, while for CoMSIA fields based on similarity indices between probe atoms and molecule were utilized.[22] Finally the GRID/Golpe approach is based on the analysis of GRID molecular interaction fields using particular probes. The PLS method (*partial least squares*)[24] is used to derive a linear relationship for highly underdetermined matrices, while cross-validation[25] was used to probe consistency and predictiveness. The visual interpretation of 3D-QSAR results using contour maps enhances the understanding of electrostatic, hydrophobic, and steric requirements for ligand binding. However, those important regions are only identified if there is some structural variance in the underlying data set. Then any new inhibitor design can be focused to those regions, where alterations in property fields correlate to ligand affinity or selectivity.

2. Methods

2.1. Design of MMP-3/8 Inhibitors. The MMP inhibitor design was based on previous X-ray binding site geometries, revealing that metalloproteinase specificity mainly resides in the S1′ pocket,[26] close to the catalytic Zn^{2+} ion. Hydroxamic and carboxylic acids were selected to coordinate to zinc. They were connected to a 1,2,3,4-tetrahydro-isoquinoline scaffold, which is N-substituted by arylsulfonyl substituents proposed to bind in the S1′ pocket. These sulfonamides are available in many variations and provide ideal hydrogen bond geometries. The 2-arylsulfonyl-1,2,3,4-tetrahydro-isoquinoline-3-carboxylates and -hydroxamates as inhibitors of MMP-3 and MMP-8 have been synthesized and biologically tested as described elsewhere.[27,28] Their chemical structures, biological activities, and selectivities are summarized in Table 1.

2.2. General Computational Procedures. All model building was done using the program SYBYL[29] on SGI workstations. Conformations of ligands and protein−ligand complexes were energy minimized using a quasi-Newton−Raphson or conjugate gradient procedure based on the TRIPOS 6.0 force field[30] and Gasteiger−Marsili charges.[31] Docking studies were based on MMP-3 or MMP-8 X-ray structures. After analyzing key protein−ligand interactions using GRID,[15] candidate molecules were manually placed in the active site and minimized, treating all ligand atoms plus all protein residues within a sphere of 4 Å as flexible. Other compounds were built accordingly to obtain the 3D-QSAR alignment rule based on structure information for both MMPs, as described in ref 13.

2.3. 3D-QSAR Using CoMFA and CoMSIA. After defining a superposition rule by minimizing individual protein−ligand complexes,[13] the steric and electrostatic interaction energies between a probe atom and all compounds are computed at the surrounding points of a predefined grid, using a volume-dependent lattice with 1 or 2 Å grid spacing, a positively charged carbon atom, and a distance-dependent dielectricity constant for CoMFA. The magnitude of the regions was defined to extend the ensemble of superimposed conformers by 4.0 Å along the principal axes of a Cartesian coordinate system. The maximum field values were truncated to 30 kcal/mol for steric and ±30 kcal/mol for electrostatic interactions. For points "*inside*" a molecule (steric energy of 30 kcal/mol), no electrostatic energy was computed. This alignment was also used to compute steric, electrostatic, and hydrophobic[32] CoMSIA similarity index fields. CoMSIA has the advantage that no singularities occur at atomic positions due to a Gaussian-type distance dependence of the physicochemical properties. Similarity indices[33] were computed using a probe with charge +1, a radius of +1, a hydrophobicity of +1, and an attenuation factor α of 0.3 for the Gaussian-type distance dependence.

Equal weights for CoMFA or CoMSIA fields were assigned using the CoMFA_STD block scaling option.[34] Cross-validated analyses were run using the leave-one-out method in SAM-PLS[35] and cross-validation groups with random members selection, averaged over 100 runs. While CoMFA columns with a variance smaller than 2.0 were excluded prior to the PLS analysis (*minimum-sigma*), no column-filtering was used for CoMSIA. The overall quality of all PLS analyses was expressed using the cross-validated r^2 defined as

$$r^2(\text{cv}) = (\text{SD} - \text{PRESS})/\text{SD} \qquad (1)$$

where SD is the variance of biological activities around the mean values and PRESS refers to the sum of squared differences between predicted and target property values.
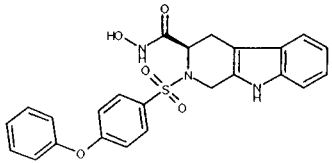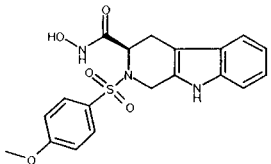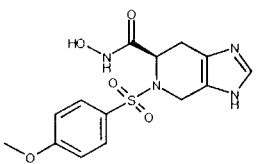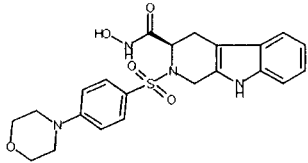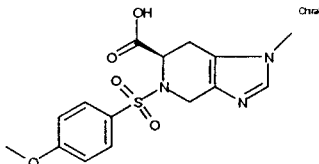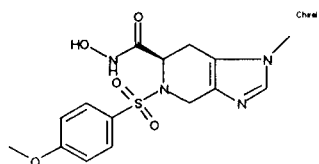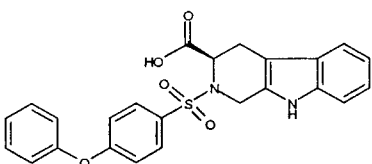
2.4. 3D-QSAR Using GRID/Golpe. To complement CoMFA and CoMSIA results, a matrix of interaction energies between a probe and all compounds is computed at surrounding points of a 1 Å grid using the GRID force field[15] and a phenolic OH probe. The same alignment as for CoMFA/CoMSIA was used. GRID probes are very specifically encoding spatial information about molecular interactions. The OH probe especially can donate and accept hydrogen bonds, while its electronic configuration allows it to interact with the π system of aromatic rings. Those interactions are of steric, electrostatic, and hydrophobic nature; they are added together to give a single value for each point.
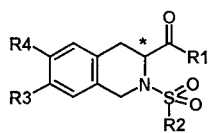
Data pretreatment and statistical analyses are done using the program Golpe.[36] A maximum cutoff of 1 kcal/mol rejects information from repulsive (positive) interaction values,[37,38] while attractive interactions decline smoothly with increasing distance. Columns where all but one ligand have similar values (2- and 3-level variables[23]) are excluded from PLS.

D-Optimal[39] preselection of variables and fractional factorial design (FFD) forms the basis for the Golpe algorithm (*generating optimal linear PLS estimations*.[23]). First the most informative variables are selected by D-optimal design from a PLS model with optimal number of components. In an iterative cycle, not more than 50% redundant variables are omitted in each cycle[23] and then a new PLS model is generated with all remaining columns, until the r^2 value decreases (typically after 3−4 cycles). The influence of each variable on the predictive ability is estimated by multiple cross-validated PLS analyses, in which variables are included according to a FFD design matrix[23] with the number of columns equal to the number of variables after D-optimal preselection. Thus the influence of each variable,[40] expressed as SDEP (*standard deviation error of predictions*[41]) can be computed. To identify only those variables that significantly improve the predictive power, some random columns are introduced in the design matrix, which by definition have no influence and thus serve to estimate confidence levels. While the D-optimal design works fast, the CPU time for the FFD selection is between 5 and 10 h on a SGI R10000 workstation.

Alternatively the SRD (*smart region definition*) method[42] for grouping individual descriptors into regions of neighboring 3D variables with similar statistical and chemical information was applied to enhance chemical relevance of the results. It

Table 1. 2-Arylsulfonyl-1,2,3,4-tetrahydro-isoquinoline-3-carboxylates and -Hydroxamates as Inhibitors of Matrix Metalloproteinases MMP-3 and MMP-8[a]

No.	Structure	Isomer	MMP-8 IC50 [nMol/l]	MMP-3 IC50 [nMol/l]	Selectivity
1		R	7	60	0.93
2		R	3	20	0.82
3		R	4	20	0.70
4		R	70	300	0.63
5		R	1000	10000	1.00
6		R	3	20	0.82
7		R	100	2000	1.30

No.	Isomer	R1	R2	R3	R4	MMP-8 IC 50 [nMol/l]	MMP-3 IC 50 [nMol/l]	Selectivity
8	R	NH-OH	4-Phenoxybenzene	H	H	2	20	1.00
9	R	NH-OH	Benzene	H	H	20	300	1.18
10	R	NH-OH	4-Acetamidobenzene	H	H	100	700	0.85
11	R	NH-OH	3-Ethoxycarbonylaminobenzene	H	H	300	6000	1.30
12	R	NH-OH	Dibenzofuran-2-yl	H	H	10	100	1.00
13	R	NH-OH	Phenylethyl	H	H	20	100	0.70
14	R	NH-OH	4-Ethoxycarbonylaminobenzene	H	H	30	500	1.22

Table 1. (Continued)
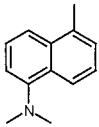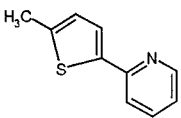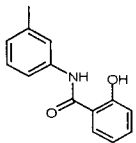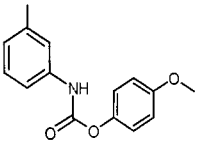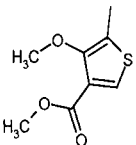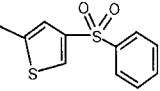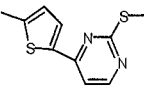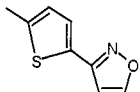
No.	Isomer	R1	R2	R3	R4	MMP-8 IC 50 [nMol/l]	MMP-3 IC 50 [nMol/l]	Selectivity
15	R	NH-OH	Hexadecyl	H	H	3000	10000	0.52
16	R	NH-OH	4-Methoxybenzene	NO$_2$	H	8	80	1.00
17	R	NH-OH	2-Naphtyl	H	H	70	400	0.76
18	R	NH-OH	Benzyl	H	H	3000	~10000	xxx
19	R	NH-OH	(structure)	H	H	200	400	0.30
20	S	NH-OH	4-Methoxybenzene	H	H	1000	7000	0.85
21	R	NH-OH	4-Methoxybenzene	H	H	4	20	0.70
22	R	NH-OH	4-(4-Dimethylaminophenoxy)-phenyl	H	H	2	30	1.18
23	R	NH-OH	4-Biphenyl	H	H	2	20	1.00
24	R	NH-OH	(structure)	H	H	2	20	1.00
25	R	NH-OH	4-Benzoylphenyl	H	H	3	30	1.00
26	R	NH-OH	(structure)	H	H	200	~10000	xxx
27	R	NH-OH	(structure)	H	H	200	3000	1.18
28	R	NH-OH	(structure)	H	H	80	700	0.94
29	R	COOH	4-Methoxybenzene	H	H	7000	10000	0.15
30	R	NH-OH	4-Methoxybenzene	NH$_2$	H	2	10	0.70
31	R	NH-OH	4-Methoxybenzene	tert.-Butoxycarbonylamino	H	3	20	0.82
32	R	NH-OH	4-Pyridylethyl	H	H	30	900	1.48
33	R	NH-OH	4-Butoxybenzene	H	H	2	30	1.18
34	R	NH-OH	(structure)	H	H	2000	10000	0.70
35	R	NH-OH	(structure)	H	H	20	2000	2.00
36	R	NH-OH	4-Morpholinobenzene	H	H	100	300	0.48
37	R	NH-OH	4-Methoxybenzene	Ethoxycarbonyloxy	H	2	20	1.00
38	R	NH-OH	4-Methoxybenzene	OH	H	1	10	1.00
39	R	NH-OH	4-Methoxybenzene	4-Methoxybenzene-sulfonyloxy	H	8	30	0.57
40	R	NH-OH	(structure)	H	H	8	50	0.80

Table 1. (Continued)
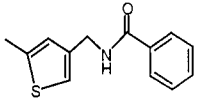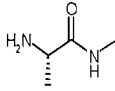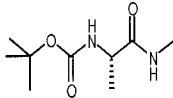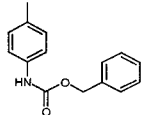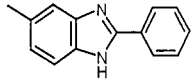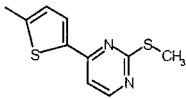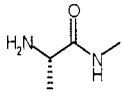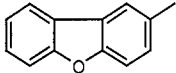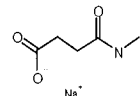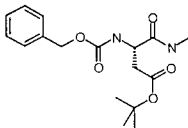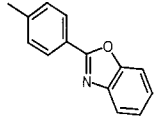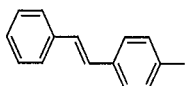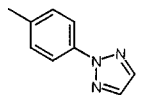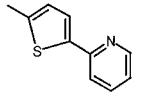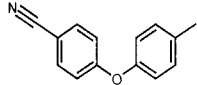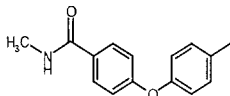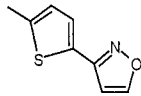
No.	Isomer	R1	R2	R3	R4	MMP-8 IC 50 [nMol/l]	MMP-3 IC 50 [nMol/l]	Selectivity
41	R	NH-OH		H	H	800	6000	0.88
42	R	NH-OH	4,5-Dibromthiophene2-yl	H	H	500	10000	1.30
43	(S)-Ala-(R)-Tic	NH-OH	4-Methoxybenzene		H	5	30	0.78
44	(S)-Ala-(R)-Tic	NH-OH	4-Methoxybenzene		H	2	20	1.00
45	R	NH-OH		H	H	800	400	-0.30
46	R	NH-OH	4-Trifluoromethylbenzene	H	H	50	500	1.00
47	R	NH-OH		H	H	40	500	1.10
48	R	COOH	4-Methoxybenzene	NH_2	H	2000	10000	0.70
49	R	COOH	4-Chlorobiphenyl	H	H	5	100	1.30
50	R	NH-OH	4-Chlorobiphenyl	H	H	3	20	0.82
51	R	COOH		H	H	1000	~10000	xxx
52	R	COOH	4-Phenoxybenzene	H	H	9	200	1.35
53	R	COOH	4-(4-Fluorobenzyloxy)-benzene	H	H	30	200	0.82
54	R	COOH	4-Benzoylphenyl	H	H	300	3000	1.00
55	(S)-Ala-(R)-Tic	COOH	4-Methoxybenzene		H	2000	10000	0.70
56	R	COOH		H	H	500	~10000	xxx
57	R	NH-OH	4-(4-Fluorobenzyloxy)-benzene	H	H	2	20	1.00
58	R	COOH	4-Chlorobiphenyl	NO_2	H	30	600	1.30
59	R	COOH	4-Chlorobiphenyl	H	NO_2	20	500	1.40
60	R	COOH	4-Chlorobiphenyl	NH_2	H	7	200	1.46
61	R	COOH	4-Phenylmercaptobenzene	H	H	200	2000	1.00
62	R	COOH	4-Biphenyl	H	H	10	500	1.70
63	R	COOH	4-(4-Dimethylaminophenyl)-benzene	H	H	1	100	2.00
64	R	COOH	4-Fluorobiphenyl	NO_2	H	40	1000	1.40
65	R	COOH	4-Fluorobiphenyl	H	NO_2	20	500	1.40
66	R	COOH	4-Methoxybenzene	OH	H	3000	10000	0.52
67	S	COOH	4-Chlorobiphenyl	H	H	3000	10000	0.52
68	R	COOH	4-(4-Chlorophenoxy)-benzene	H	H	20	500	1.40
69	R	COOH	4-Fluorobiphenyl	H	tert.-Butoxy-carbonyl-amino	30	400	1.12
70	R	NH-OH	4-Fluorobiphenyl	tert.-Butoxycarbonylamino	H	10	100	1.00
71	R	COOH	4-Fluorobiphenyl	H	H	20	400	1.30

Table 1. (Continued)

No.	Isomer	R1	R2	R3	R4	MMP-8 IC 50 [nMol/l]	MMP-3 IC 50 [nMol/l]	Selectivity
72	R	NH-OH	4-Methoxybenzene		H	40	20	-0.30
73	(S)-Asp-(R)-Tic	NH-OH	4-Methoxybenzene		H	10	40	0.60
74	R	COOH	4-Fluorobiphenyl	H	NH₂	80	200	0.40
75	S	COOH	4-Fluorobiphenyl	H	H	4000	10000	0.40
76	R	COOH		H	H	80	800	1.10
77	R	NH-OH	4-Fluorobiphenyl	NH₂	H	2	20	1.00
78	R	COOH		H	H	60	700	1.07
79	R	COOH	4-Chlorobiphenyl	OH	H	20	200	1.00
80	R	COOH		H	H	90	600	0.82
81	R	COOH	4-(4-Dimethylamino)-phenoxybenzene	H	H	30	600	1.30
82	R	COOH		H	H	90	2000	1.35
83	R	COOH	1-Styryl	H	H	300	1000	0.52
84	R	COOH		H	H	10	1000	2.00
85	R	COOH		H	H	40	300	1.00
86	R	COOH	(4-Biphenyl)-ethyl	H	H	20	50	0.40
87	R	COOH		H	H	600	2000	0.52
88	R	NH-OH	4-(4-Dimethylaminophenyl)-benzene	H	H	2	20	1.00
89	R	COOH	4-(4-Trifluoromethyl)-biphenyl	H	H	10	300	1.30
90	R	COOH	4-Bromophenyl	H	H	600	~5000	xxx

[a] Ninety MMP inhibitors from ref 13 are listed, while only those with entries in the column selectivity are used for 3D-QSAR analyses. Selectivity: computed using eq 3; xxx indicates entry not used for QSAR, as MMP-3 value is only approximated. No.: compound ID in agreement with ref 13.

works by extracting a subset of highly informative *X* variables and partitioning the space around the molecules among them. These regions have been shown to represent the structural variability within a series better than individual descriptor columns and are able to replace the latter in chemometrical analyses. The identified regions, containing single pieces of information, are subsequently used for the FFD-based variable selection, now replacing individual columns by groups of descriptors. Some alternative procedures to group variables in *X* space are described elsewhere.[43] The algorithm implemented in Golpe encompasses the following steps: First the most informative variables are selected as seeds using an initial PLS model. This selection works in the chemometrical space of the PLS weights; columns with high importance, i.e., high absolute weights, are selected using D-optimal design. Those seeds are placed in the Cartesian space around the molecules and used to group the remaining *X* variables to the nearest seed using Euclidean distance criterion, producing a number of Voronoi polyhedra. If neighboring regions now contain similar information, analyzed using the correlation of their averaged positive and negative grid energies per molecule, they are collapsed into larger regions. The criteria for collapsing are conservative and do not merge slighly different regions.[42a] As for the original Golpe procedure, a FFD-based

Table 2. Thirty-Eight Used GRID Probes for Multivariate Binding Site Characterization and Principal Component Analysis (Abbreviations and Chemical Meaning)

N3+	sp3 amine NH3 cation	N2+	sp3 amine NH_2 cation
N2:	sp3 NH_2 with lone pair	N2=	sp2 amine NH_2 cation
N2	neutral flat NH_2, e.g., amide	N1+	sp3 amine NH cation
N1:	sp3 NH with lone pair	N1=	sp2 amine NH cation
N1	neutral flat NH, e.g., amide	NH=	sp2 NH with lone pair
N1#	sp NH with one hydrogen	N:	sp3 N with lone pair
N:=	sp2 N with lone pair	N:#	sp N with lone pair
N−:	anionic tetrazole N	NM3	trimethyl-ammonium cation
O	sp2 carbonyl oxygen	O::	sp2 carboxy oxygen atom
O−	sp2 phenolate oxygen	O=	O of SO_4 or sulfonamide
OH	phenol or carboxy OH	O1	alkyl hydroxy OH group
OC2	ether or furan oxygen	OES	sp3 ester oxygen atom
ON	oxygen of nitro group	OS	O of sulfone/sulfoxide
OH2	water	H	neutral hydrogen atom
C3	methyl CH_3 group	C1=	sp2 CH aromatic or vinyl
PO4	PO_4 phosphate dianion	PO4H	PO_4H phosphate anion
S1	neutral SH group	F	organic fluorine atom
CL	organic chlorine atom	BR	organic bromine atom
I	organic iodine atom	DRY	hydrophobic probe

variable selection is carried out as the final step, now working on regions rather than individual variables, which speeds up CPU time (2−4 h).

2.5. Understanding Protein Selectivity. To understand selectivity toward one target protein, a recent method[17] using principal component analysis (PCA) of GRID descriptors was applied. Its objective is to quantitatively characterize ligand−protein interactions in order to uncover the main differences of both proteins' ligand binding sites and to identify the most selective chemical fragments for new ligands. This potentially useful approach could also be applied to ligand molecular fields, leading to the identification of their main structural differences, while here it was applied to fields derived from the corresponding receptor structures. First, 3D structures for each protein were retrieved from the PDB database and prepared by deletion of water and ligands. Counterions were added for charged functional groups using GRID routines MINIM and FILMAP.[16] For MMP-8 the PDB file 1jao[9] was superimposed with 1sln[8] for MMP-3 using an iterative fitting procedure on 157 residues using Pro86-Gly242 for MMP-8 and Pro90-Tyr223 plus Ser225-Gly247 for MMP-3. Cα pairs were rejected for subsequent iterations if they showed an rms deviation higher than 3 times the standard deviation, leading to an rmsd of 0.41 Å for 117 Cα atoms. The ligand binding sites within a box with edges of 16, 22, and 22 Å in *x, y, z* dimensions were considered. The binding site characterization was done using the GRID force field[15,16] with functional groups listed in Table 2 on a 1 Å grid and a static protein treatment (directive *MOVE*).

A maximum cutoff value of 0 kcal/mol was used to reject information from repulsive GRID interaction points in the **X** matrix. When using only favorable interaction energies, the information solely related to steric repulsion is removed and the analysis is focused on favorable protein−ligand interactions. All data were centered by subtracting the column average from each column and no further scaling was applied, as the physicochemical origin of the data is similar. Two probes, Mg^{2+} and Ca^{2+}, were rejected from the initial 40 probes in Table 2, as they were outliers in an initial PCA. Thus the data matrix finally consists of 8993 *x* variables × 38 probes × 2 proteins. Each row corresponds to a specific probe−target interaction, while columns are variables that describe interactions at regular grid.

A PCA[44] was carried out using Golpe[23] to contract the large number of collinear and multicollinear variables with redundant information to a few orthogonal principal properties. In PCA, the original data matrix **X** with interaction energies x_{ik} for *i* probe−target interactions and *k* grid points is decomposed to means (x_k), scores (t_{ia}), loadings (p_{ak}), and residuals (e_{ik}), with *a* denoting the number of model dimensions:

$$x_{ik} = x_k + \sum_{a=1}^{A} t_{ia}p_{ak} + e_{ik} \qquad (2)$$

The data matrix **X** is approximated by the product of two smaller matrixes, scores and loadings. The score matrix p_{ak} gives a simplified picture of the objects (i.e., probes interacting with proteins), represented by a few uncorrelated new variables. Those scores can be plotted to visualize differences in protein-interaction pattern. The first new principal component (PC) describes the maximum variance among all possible directions, the second one the next largest variation among all directions orthogonal to the first one, etc. The resulting eigenvalues represent the overall variance after extraction of each successive new factor. If most of the variation of the original data can be described by the first few factors, a much simpler data structure exists. Here the NIPALS algorithm was used, which calculates every component in a stepwise manner and is faster than diagonalization of the covariance matrix if only the first eigenvalues are desired. The principal component axes were not additionally rotated.

3. Results and Discussion

3.1. CoMFA and CoMSIA Models for MMP-3 Affinity. All 3D-QSAR models for MMP-3 affinity reveal a high degree of consistency. Using a 2 Å grid spacing, a CoMFA model with an r^2(cv) value of 0.563 for six relevant PLS components and a conventional r^2 of 0.944 was obtained (Table 3). The alignment for this and other models is obtained by docking all compounds into the MMP-3 binding site (Figure 1). The steric field explains 53.5% of the variance, showing a balance to the electrostatic field. With a reduced 1 Å grid spacing, a five-component PLS model results with an r^2(cv) of 0.432 and r^2 of 0.917 (Table 3). The fitted versus experimental biological activities[45] for this model are given in the Supporting Information.

Both models were subjected to several validation studies to assess their predictive power. First, the effect of the alignment relative to the grid position was investigated by consistently translating all compounds in increments of 0.5 Å in all three dimensions of the Cartesian space, while keeping the alignment intact. A mean r^2(cv) value of 0.42 with a range from 0.29 to 0.59 is obtained after several translations, revealing a significant dependence from the molecular orientation for CoMFA. Second, the biological activities were randomized and analyzed using PLS.[46] The mean r^2(cv) for 100 randomizations is −0.19 (SD 0.14) for a 2 Å grid, while for the 1 Å grid, a mean r^2(cv) of −0.08 (SD 0.06) was obtained, demonstrating the significance of the original

Table 3. Summary of 14 3D-QSAR Models for MMP-3/MMP-8 Affinity and Selectivity[a]

	r^2(cv)	SD	comp.	r^2	validation
MMP-3					
CoMFA (2 A)	0.563	0.629	6	0.944	LOO, 2 CV, randomize, grid var
CoMFA (1 A)	0.432	0.717	5	0.917	LOO, 2 CV, randomize
CoMSIA (2 A)	0.413	0.738	8	0.957	LOO, 2 CV, randomize
CoMSIA (1 A)	0.382	0.757	8	0.954	LOO
Golpe_FFD (1 A)	0.795	0.413	5	0.967	LOO, LTO, 5RG
Golpe_SRD (1 A)	0.789	0.419	5	0.964	LOO, LTO, 5RG
MMP-8					
CoMFA (2 A)	0.569	0.685	5	0.905	LOO, 2 CV, randomize, grid var
CoMFA (1 A)	0.516	0.726	5	0.911	LOO
CoMSIA (2 A)	0.478	0.763	7	0.924	LOO, 2 CV, randomize
CoMSIA (1 A)	0.447	0.786	7	0.924	LOO
Golpe_FFD (1 A)	0.729	0.512	4	0.934	LOO, LTO, 5RG
Golpe_SRD (1 A)	0.719	0.521	4	0.936	LOO, LTO, 5RG
selectivity					
Golpe_FFD (1 A)	0.532	0.280	3	0.831	LOO, LTO, 5RG
Golpe_SRD (1 A)	0.510	0.286	3	0.821	LOO, LTO, 5RG

[a] r^2(cv): cross-validated r^2 using leave-one-out; SD: standard deviation of error in leave-one-out; comp.: optimal number of PLS components, i.e., latent variables; r^2: non-cross-validated regression coefficient; validation 2 CV: cross-validation using two randomly chosen cross-validation groups 100 times; randomize: randomization of biological activities 100 times; grid var: shifting the alignment within fixed grid box; LOO: leave-one-out; LTO: leave-two-out; 5RG: cross-validation using five random groups 20 times. The used grid spacing is given in parentheses.

Figure 1. Superposition of 2-arylsulfonyl-1,2,3,4-tetrahydro-isoquinoline-3-carboxylates and -hydroxamates as MMP-3/-8 inhibitors for 3D-QSAR studies.

PLS models. Furthermore, PLS analyses with two randomly chosen cross-validation groups were repeated 100 times, leading to averaged r^2(cv) values of 0.362 (SD 0.11) for a 2 Å grid and 0.327 (SD 0.08) for the 1 Å grid. Histograms with the distribution of r^2(cv) over 100 runs for 2 cross-validation groups or randomized activities are given in the Supporting Information.

Similar results were obtained for CoMSIA: Using a 2 Å grid, an r^2(cv) value of 0.413 for eight components and a conventional r^2 of 0.957 was obtained (Table 3), while a 1 Å grid led to an eight-component model with an r^2(cv) of 0.382 and a conventional r^2 of 0.954. The steric fields in both models explain 17% of the variance, while 45% for electrostatic fields and 38% for the hydrophobic fields are observed, suggesting that the CoMFA steric field contribution is a balance between pure steric plus hydrophobic effects. Similar validations were run for the CoMSIA 2 Å model. A mean r^2(cv) value of −0.19 (SD 0.10) is observed for 100 analyses with randomized activities, while two cross-validation groups

lead to a mean r^2(cv) of 0.240 (SD 0.10), counting for a significant and predictive model.

For MMP-8, CoMFA and CoMSIA analyses produced models of similar significance, and additional validations count for predictive models, e.g. a CoMFA model with r^2(cv) of 0.569 (five component) and r^2 of 0.905. Details are given elsewhere,[13] while statistical results are summarized in Table 3.

3.2. Comparison to MMP-3 and MMP-8 Binding Site Topologies. The steric and electrostatic *std*coeff* CoMFA fields from the final PLS models with 1 Å grid spacing for MMP-8 and MMP-3 are displayed as contour maps in Figure 2 in combination with the inhibitor **8** containing a hydroxamate and a biphenyl ether-sulfonamide attached to the 1,2,3,4-tetrahydroisoquinoline scaffold (IC$_{50}$ = 20 nM for MMP-3, 2 nM for MMP-8). The interpretation of CoMSIA contour maps led to similar conclusions for both targets (see ref 13 for a discussion on MMP-8). Dark or green contours in Figures 2 and 3 on the left correspond to sterically favorable regions for biological affinity obtained from CoMFA (>85% contribution), while gray or yellow contours indicate regions affecting the biological activity in a negative way: a bulky substituent here reduces activity (<15% contribution).

Although the PLS results are derived from ligand information, the obtained contour maps are consistent with steric, electrostatic, and hydrophobic binding site requirements, as can be seen from inspection of Figure 3. Here the steric and electrostatic CoMFA contours were mapped onto the protein/**8** complex, showing that these maps correspond to steric and electrostatic requirements in both ligand binding sites, providing additional insights into key protein−ligand interactions.

Three main regions can be consistently identified for MMP-3 and -8, where steric bulk increases activity. A green contour at the distal biphenyl ether ring indicates a hydrophobic cleft within the S1′ pocket formed by the side chains of Tyr219, Leu193, and Val194, which is filled with water in some experimental MMP-inhibitor structures. Another green region highlights steric requirements in the ortho position of this ring close to the "Met-turn", a βI-turn formed by Ala213(*i*)-Leu214-
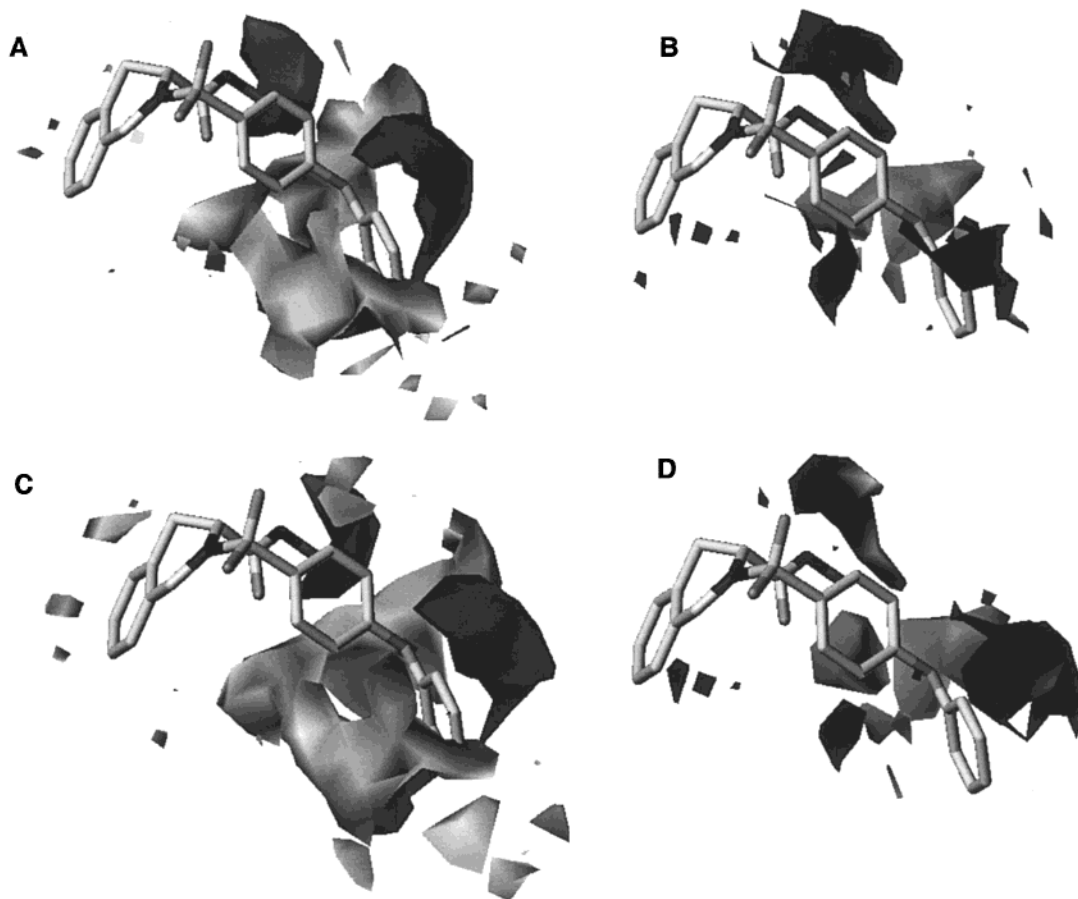
Figure 2. Steric (A, C) and electrostatic (B, D) *std*coeff* CoMFA derived contour maps from PLS models with 1 Å grid spacing for MMP-3 (A, B) and MMP-8 (C, D). For reference, the inhibitor **8** is shown (IC_{50} = 20 nM MMP-3, 2 nM MMP-8). For steric contour maps, dark contours indicate sterically favorable regions for biological affinity (>85% contribution), while gray contours show where a bulky substituent reduces activity (<15% contribution). In the electrostatic *std*coeff* contour maps, dark contours (>85% contribution) show positive charge favorable for affinity, while gray contours (<15% contribution) indicate regions where negative charge is favored or positive disfavored.

Met215-Tyr216(*i* + *4*). A third region at the zinc binding group indicates steric requirements at the metal coordination center, suggesting that the optimal distance of two zinc coordinating oxygens is better realized in hydroxamates than in carboxylates.

Steric regions connected to decreasing biological activity are also consistently identified for MMP-3 and MMP-8. A gray or yellow contour in Figures 2 and 3 at the first biphenyl ether ring indicates unfavorable steric interactions at the narrow entrance into the S1′ pocket, formed by Pro217, His207, His197, and Val194. Other yellow regions indicate sterical restrictions at the edge and bottom of the S1′ pocket, pointing to the Asn218-Tyr219 peptide backbone and Arg222 (MMP-8 numbering) located at the bottom of this pocket. Finally a region at the upper side of the aromatic isoquinoline ring indicates the preferred stereochemistry at carbon C3 for both enzymes, as the inversion of this chiral center orients this ring toward this yellow contour.

The electrostatic *std*coeff* contour maps are also similar for CoMFA (Figures 2 and 3, right panels) and CoMSIA. Dark or blue contours (>85% contribution) point to those regions, where positive charge is favorable to enhance biological affinity, while gray or red contours (<15% contribution) indicate regions where negative charge is favored. Blue CoMFA contours close to the hydroxamate N*H* indicates favorable hydrogen bond

interactions to the Ala161 backbone carbonyl oxygen and the Glu198 side chain functionality in agreement with X-ray crystallographic studies,[47] while red contours at the catalytic zinc indicate the optimal geometry for zinc complexation. Another blue contour next to the meta position of the first biphenyl ether ring corresponds to a favorable interaction to Val194-C=O and Leu193-C=O. In fact, the chemical interpretation of CoMFA and CoMSIA contour maps for both targets led to similar conclusions for MMP-3 and MMP-8 inhibition, raising the question of how to explain observed experimental selectivities in this set of tetrahydroisoquinolines.

3.3. GRID/Golpe Models for MMP-3 and MMP-8 Affinity. To identify additional binding requirements for MMP-3 and MMP-8, 3D-QSAR analyses were performed using an approach based on computed interaction energies from each individual ligand to a phenolic OH probe positioned on regular points of a predefined 1 Å grid. The same alignment as for CoMFA and CoMSIA was used for both analyses. Interaction energies were computed using the GRID force field[15,16] and analyzed with PLS and variable selection in Golpe.[23] The effect of individual variables on predictivity based on a FFD design matrix points exactly to relevant variables (Golpe_FFD, Table 3). The PLS model for MMP-3 after FFD-based variable selection contains
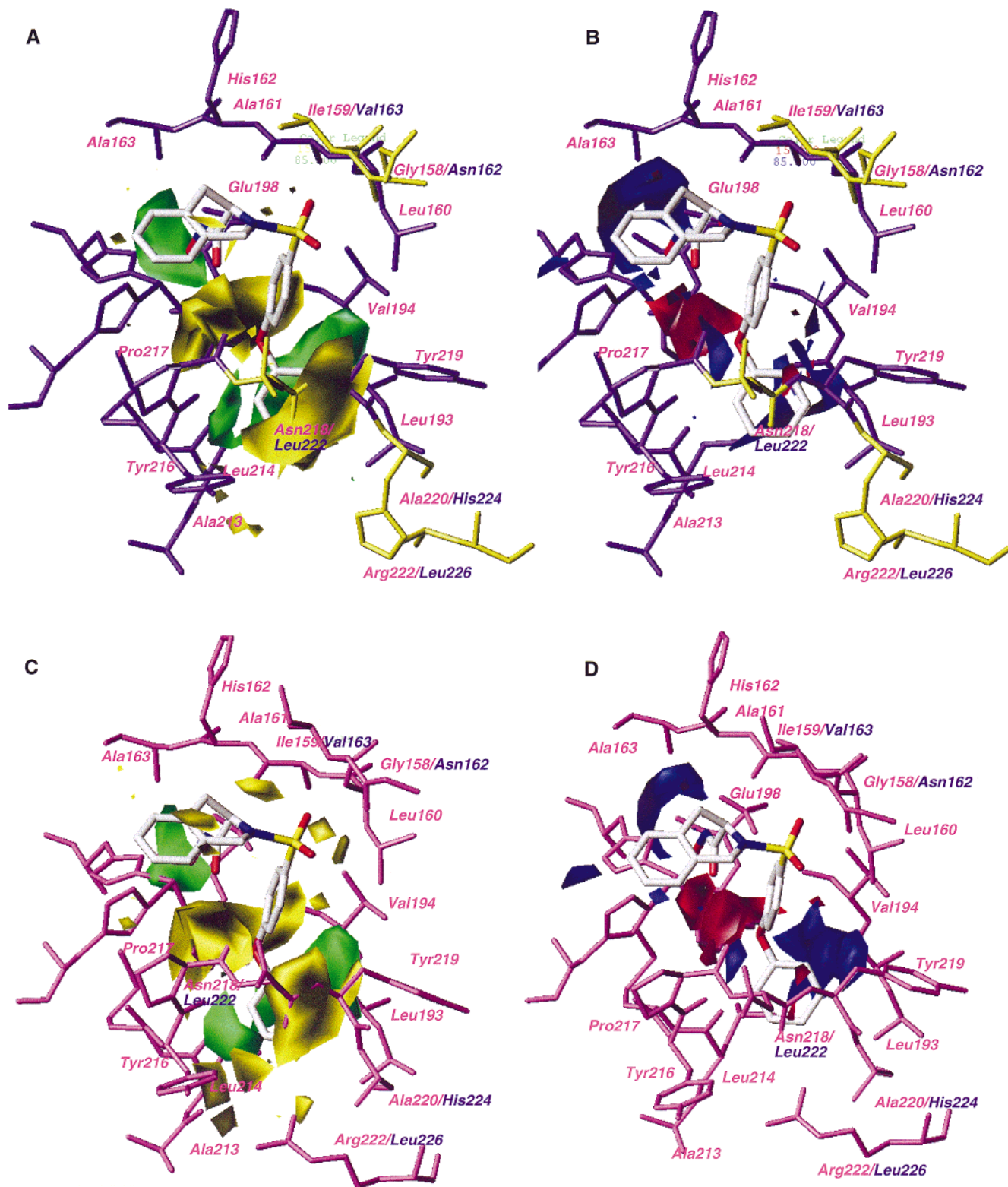
Figure 3. The model from Figure 2 with inhibitor **8** and protein topology (MMP-3 in A, B and MMP-8 in C, D). Only selected amino acid residues are shown. The numbering scheme is taken from MMP-8, while differences in MMP-3 are indicated at the second position. MMP-3 mutations and insertions (A, B) compared to MMP-8 (C, D) are highlighted in yellow, while similar residues are colored in purple. Green contours indicate sterically favorable regions, while yellow contours show where a bulky substituent reduces activity. In the electrostatic *std*coeff* contour maps, blue contours indicate positive charge favorable for affinity, while yellow contours indicate regions where negative charge is favored or positive disfavored.

1316 from 25 740 grid points, leading to an r^2(cv) of 0.795 with five PLS components and an r^2 of 0.967 (Table 3). Validation using the leave-two-out approach (r^2(cv) = 0.792) and five cross-validation groups (r^2(cv) mean = 0.749) demonstrate the significance of this

model. For MMP-8, a PLS model with an r^2(cv) of 0.729 and an r^2 of 0.934 for four components and 1382 variables after FFD-based variable selection was obtained (see Table 3). Here only 85 inhibitors with available MMP-3 data were included, in contrast to ref

13. This predictivity of this model is revealed by leave-two-out cross-validation (r^2(cv) = 0.728) and five cross-validation group analyses (r^2(cv) mean = 0.703).

In addition the "smart region definition" (SRD) method for grouping variables into regions of neighboring descriptors with similar statistical and chemical information was applied to enhance the chemical relevance of results. Here the final PLS model after SRD contains 1049 variables, and an r^2(cv) of 0.789 (five components) and an r^2 of 0.964 are obtained (Table 3). This model was validated by using the leave-two-out approach (r^2(cv) = 0.787) and five cross-validation groups (r^2(cv) mean = 0.743). The corresponding MMP-8-based four-component PLS model contains 1282 selected variables, and an r^2(cv) value of 0.719 and an r^2 of 0.936 are obtained (Table 3). Its significance is revealed by results from leave-two-out cross-validation (r^2(cv) = 0.713) and five cross-validation groups (r^2(cv) mean = 0.669). The FFD- or SRD-based models for MMP-3 and MMP-8 led to very similar results, providing consistent insights into favorable interactions by a complementary approach to CoMFA and CoMSIA.

The coefficient maps for the MMP-3 and MMP-8 models from SRD or FFD variable selection are compared in Figure 4 using the inhibitor **8** docked into MMP-3 or MMP-8. Similar residues in MMP-3 and MMP-8 are colored in purple in Figure 4E, while different amino acids are yellow. The numbering in Figure 6E,F follows the MMP-8 convention, while mutations in MMP-3 are indicated. The negative and positive PLS coefficient fields are shown in cyan (−0.003 in 6A and −0.007 in 6B−D) and orange (+0.003 in 6A and +0.007 in 6B−D), respectively. 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, orange contours indicate that repulsive interactions (space filling) increase affinity, while attractive interactions decrease activity.

In general, the chemical interpretation for the PLS models for MMP-3 or MMP-8 led to similar results, regardless of the method for variable selection (SRD in Figure 4A,C; FFD in Figure 4B,D). Corresponding regions with cyan contours can be identified, indicating favorable attractive interactions between selected ligands and polar binding site regions. The first cyan region indicates favorable interactions involving the catalytic zinc and the residues Ala165 and Glu202 in the binding site to the hydroxamate and carboxylate functional group, highlighting the optimal zinc binding geometry. The polar entrance into the S1′ pocket close to the sulfonamide is highlighted by a cyan contour, while favorable ligand interactions to the backbone carbonyl oxygens of Tyr220 and Pro221 pointing into S1′ are also indicated by cyan contours.

Unfavorable interactions are indicated by orange contour regions in Figure 4. One region is located close to the small hydrophobic cleft Leu197, Val198, and Tyr223 in the S1′ pocket, which agrees with CoMFA and CoMSIA results, revealing that hydrophobic, sterically demanding substituents filling the S1′ pocket increase affinity. Another orange contour region in all models in

Figure 4 indicates repulsive interactions to the catalytic zinc ion, which in combination with the neighboring cyan colored region provides detailed information about zinc binding. Thus the chemical interpretation of the GRID/Golpe derived models for both targets is in good agreement to CoMFA and CoMSIA, while their comparison does not allow the extraction of key regions responsible for selective interaction to only one of both MMPs.

3.4. 3D-QSAR Model for Ligand Selectivity. Although all previous 3D-QSAR models are designed to explain MMP-8 and MMP-3 affinities, no information was obtained to understand selectivities, as from the resulting contour maps in Figures 2, 3, and 4 no different interactions pattern to both MMPs are obvious. To this end, two complementary approaches were used. First a 3D-QSAR model was derived using

$$y = \log(1/IC_{50}(MMP-8) \times 100\ 000) - \log(1/IC_{50}(MMP-3) \times 100\ 000) \quad (3)$$

as a quantitative selectivity measure for the dependent variable in PLS.[48] Again GRID interaction energies to the phenolic OH probe on a 1 Å grid were used as descriptors. The PLS model after FFD-based variable selection with 1226 from 25 740 variables resulted in an r^2(cv) of 0.532 for three components and an r^2 of 0.831 (Table 3). This model was validated using leave-two-out (r^2(cv) = 0.529) and five cross-validation groups (r^2(cv) mean = 0.479). The corresponding model obtained after SRD-based grouping of variables containing 1127 variables produces an r^2(cv) of 0.510 (three components) and an r^2 of 0.821 (Table 3). The leave-two-out validation (r^2(cv) = 0.507) and repeated models with five cross-validation groups (r^2(cv) mean = 0.461) also support a statistically significant model. However, individual models for MMP-3 or MMP-8 affinity show higher r^2(cv) values, suggesting that models of higher consistency are obtained using a single biological observable rather than introducing a dependent variable as function of different IC_{50} values with associated experimental errors.[49]

Both analyses led to models of similar chemical interpretability, thus only contour maps from the SRD model are shown in Figure 5. Cyan contours now highlight regions where favorable interactions to the OH probe increase selectivity toward MMP-8 or decrease it for MMP-3, while a particular substituent oriented toward an orange region increases selectivity toward MMP-8 by repulsive or unfavorable hydrophobic OH interactions.

Four cyan selectivity regions in the S1′ pocket are seen in Figure 5B in agreement with binding site topologies. To further guide chemical interpretation, the selective MMP-8 inhibitor **84** (IC_{50} = 1000 nM MMP-3, 10 nM MMP-8) is shown in Figure 5, carrying a Zn binding carboxylate and a biphenyl ether P1′ moiety with a para-cyano substituent at the distal ring. Its observed selectivity can be rationalized by a cyan contour at the bottom of S1′ close to the cyano group, showing that polar interactions here are favorable for MMP-8 selectivity. This contour agrees with selective interactions to the Arg222 side chain at the bottom of S1′ in MMP-8. Furthermore, a cyano group close to a
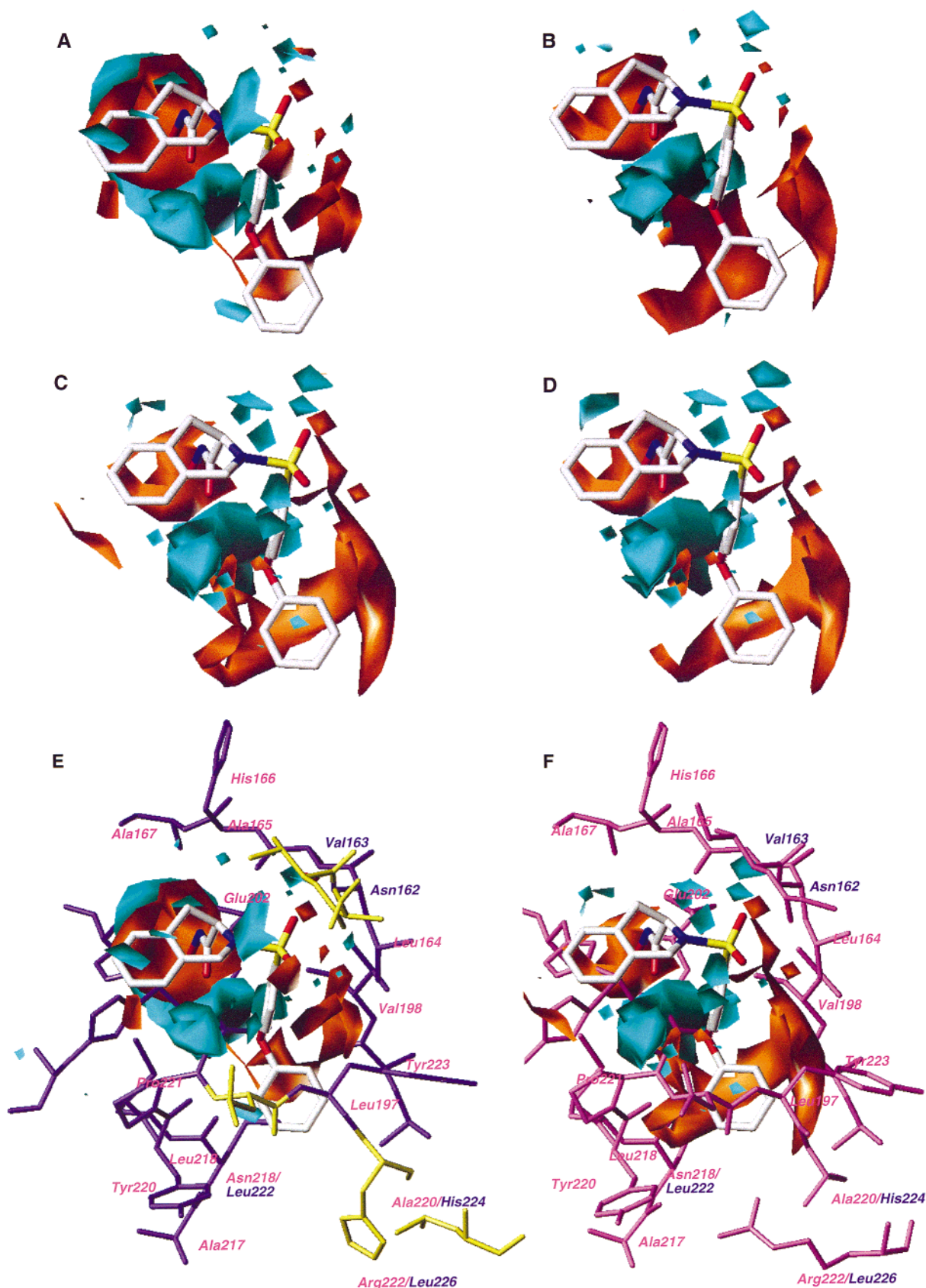
Figure 4. PLS coefficient maps for GRID/Golpe models with a phenolic OH probe and a 1 Å grid spacing for MMP-3 (A, B) and MMP-8 (C, D). The models on the left (A, C) were obtained by smart region definition (SRD)-based variable selection, while the models on the right (B, D) resulted from FFD-based variable selection. For reference, the inhibitor **8** is displayed. Cyan contours at −0.003 for A and −0.007 for B−D indicate regions where attractive OH interactions increase affinity, while repulsive interactions lower activity. Orange contours at +0.003 for A and +0.007 for B−D highlight repulsive interactions increasing biological affinity (e.g., space filling), while attractive interactions decrease activity. For comparison, the models A and B are shown in combination with the corresponding protein structure in E for MMP-3 and F for MMP-8. See Figure 3 for details.
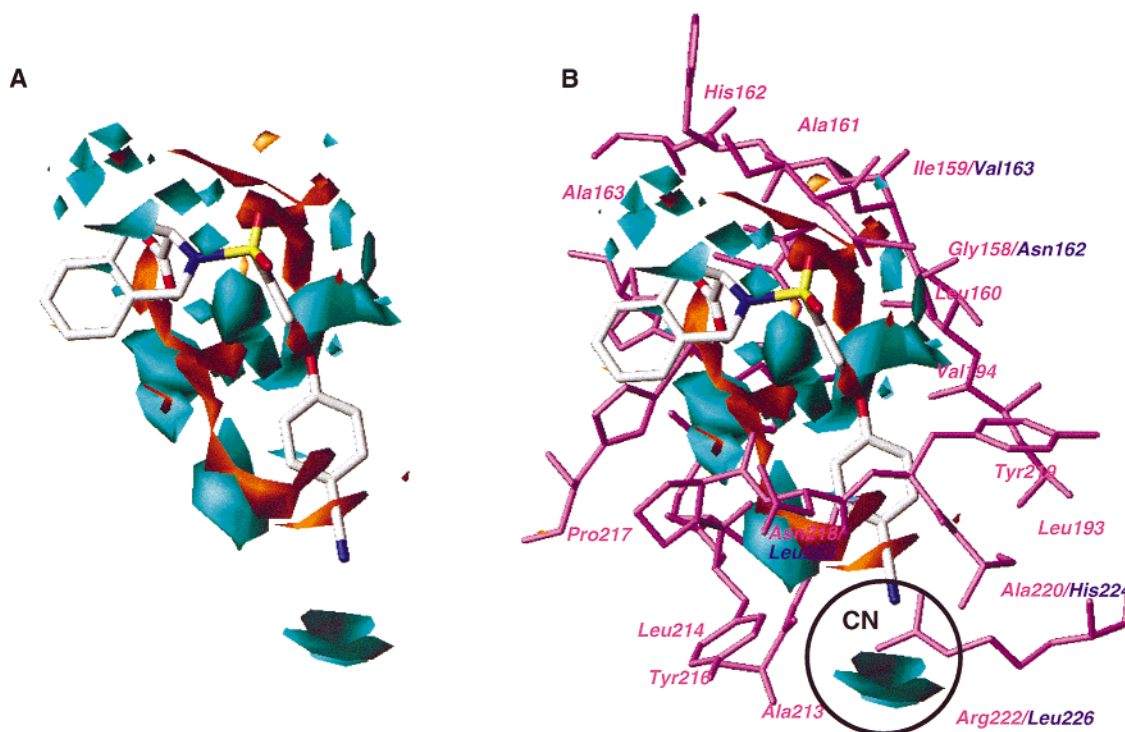
Figure 5. PLS coefficient maps for the GRID/Golpe selectivity model with an OH probe on a 1 Å grid without (A) and with MMP-8 binding site (B), derived using SRD-based variable selection. The inhibitor **84** (IC_{50} = 1000 nM MMP-3, 10 nM MMP-8) is shown. Cyan contours at -0.0025 indicate regions where favorable OH interactions increase selectivity for MMP-8, while orange contours at $+0.0035$ highlight regions where increased MMP-8 selectivity results from repulsive or unfavorable OH interactions.

guanidine is found four times when searching 8059 PDB complexes using the program ReliBase.[50] While aceto-nitril is found two times (1scb; subtilisin Carlsberg[51]), in 2cgr (IGG2B (κ) FAB fragment[52]) and 1hpo (HIV-1 protease mutant/U103265 complex[53]), typical ligands reveal this interaction.

In MMP-8, the Arg222 guanidino group is interacting via hydrogen bonds with Pro211-C=O, Gly212-C=O (via solvent) and Ala213−C=O. In many MMP-8 X-ray structures, the S1′ bottom is filled with solvent mol-ecules hydrogen-bonded to Arg222 and backbone amides at the wall of this hydrophobic pocket. Arg222 is replaced by the less hydrophilic Leu226 in MMP-3, while the bottom of S1′ is not occluded. This hydrophobic environment agrees with 3D-QSAR results.

A second cyan colored region indicating selectivity differences is pointing toward the polar residue Asn218, which is changed to the hydrophobic Leu222 in MMP-3. Here selective interactions to MMP-8 can be designed by adequate heteroatom substitution within or attached to the distal aromatic substituent. The third cyan region is pointing toward the polar entrance into S1′, which is formed by the residues from Gly158 to Ala161 plus Tyr219. The main differences to MMP-3 are replace-ments of Ile159 by Val163 and Gly158 by the polar Asn162.

Several polar groups at the S1′ wall anchor substrates and inhibitors by hydrogen bonds, like the side chains of Asn218, Glu198, the backbone amides of the wall-forming strands, and the catalytic Zn^{2+}. The replace-ment of Asn218 versus Leu222 causes minor conforma-tional shifts in its neighborhood, which are reflected by 3D-QSAR contours next to the sulfonamide and the first aromatic ring of selective inhibitors. Thus the chemical interpretation of 3D-QSAR results based on GRID/Golpe

contour maps is in good agreement with structural differences uncovered by visual inspection of both protein cavities.

3.5. Multivariate Characterization of Protein Selectivity. In addition to this view from the ligands' perspective, a final study was based on a chemometrical approach using a PCA of multivariate GRID descriptors to uncover differences of binding sites with respect to their GRID probe interaction pattern. The knowledge of structures from X-ray crystallography provide ex-tremely useful information to focus on ligand−protein interactions and selectivity regions. The results from this approach[17,18] led to a consistent picture in agree-ment with all other investigations.

After superimposing the 3D structures for MMP-8 (PDB code: 1jap) and MMP-3 (1sln), their binding sites were characterized by GRID interaction energies to several functional groups (Table 2). This data matrix was analyzed using a PCA on only favorable interac-tions, resulting in a significant two-component model for 38 GRID probes. The first principal component (PC) explains 33.5% of the variance, while the second adds 27.8%.

In the 2D score plot in Figure 6A, each point indicates a single object from the original **X** matrix, referring to interactions of a GRID probe with one of both targets. The clustering of points into two separate groups demonstrates that the first principal component PC1 (*x*-axis) discriminates between targets (MMP-3 negative PC1 scores, *mp3*; MMP-8 positive PC1 scores, *mp8*). The PC1 scores are related to the ability of GRID probes to selectively interact with only one target. The greater the horizontal spread of a probe for both targets is, the more relevant this probe is for discrimination. For selective compounds, one should insert groups with
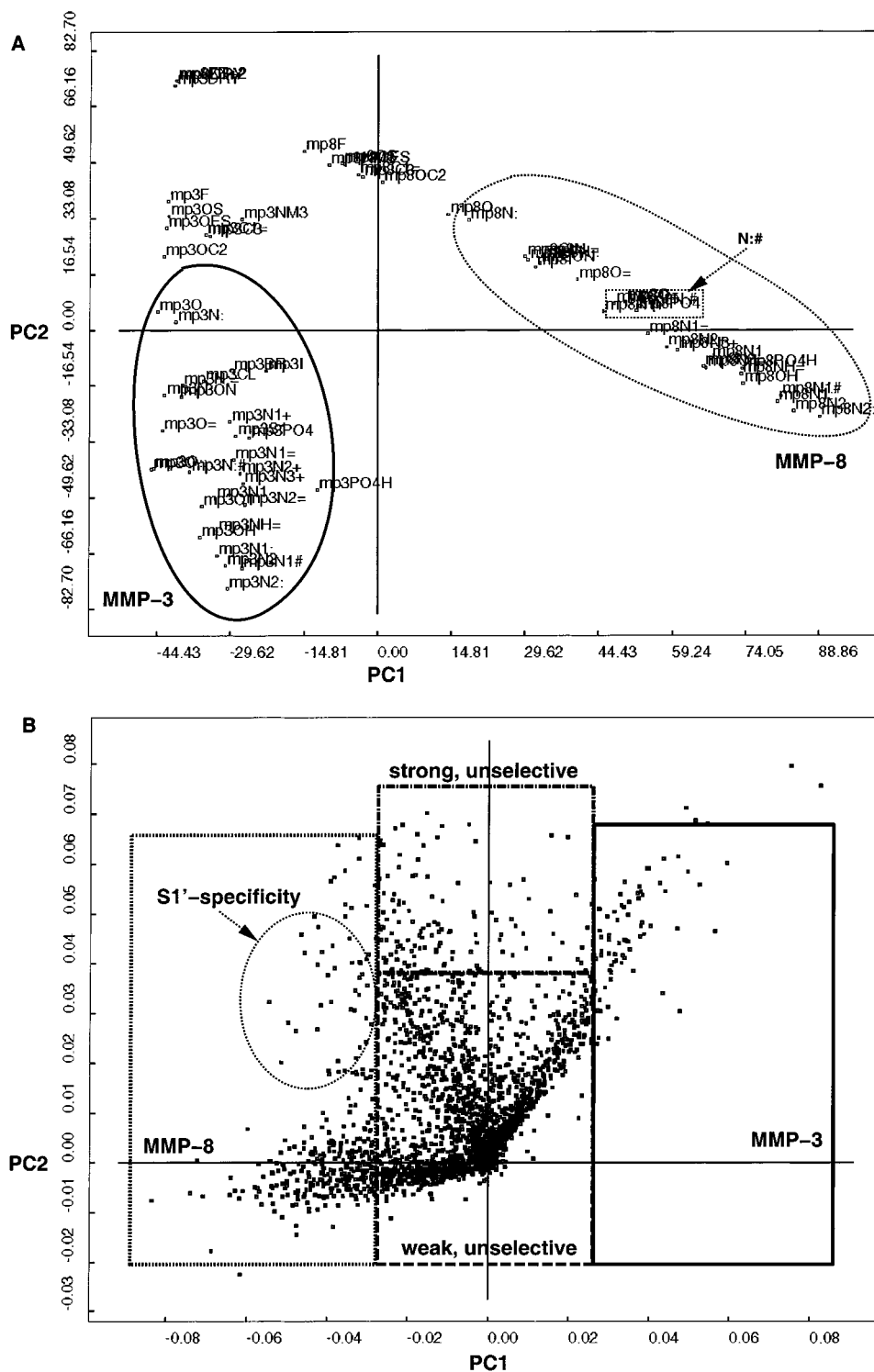
Figure 6. The 2D score (A) and loadings plot (B) of the first principal component (PC1) versus the second (PC2) for the PCA on grid interaction energies between 38 probes and both targets MMP-3 and MMP-8. The points in the score plot represent individual objects of the **X** matrix, i.e., interactions of a probe with a target. Points on the left represent MMP-3 interactions, while points on the right indicate MMP-8 interactions. The points in the loadings plot represent individual variables (grid points) of the **X** matrix. Indicated regions are discussed in the text.

higher PC1 score differences. In contrast, PC2 indicates nonselective ligand−protein interactions, ranking all 38 GRID probes by their ability to interact with common binding site regions. Higher PC2 scores indicate stronger unselective interactions to both MMPs.

The PCA loadings are displayed in Figure 6B, where main regions are indicated by different boxes. Each point represents the contribution to the PCs of each grid position with computed probe−protein interactions. Variables with high absolute PC1 loadings indicate binding regions with different interaction behavior, i.e., regions, where a group interacts loosely with one enzyme and tightly with the other. Several selectivity regions in the binding site can be identified: On the left the MMP-8 selectivity region with high PC1 and PC2 scores points to strong, selective interactions. The
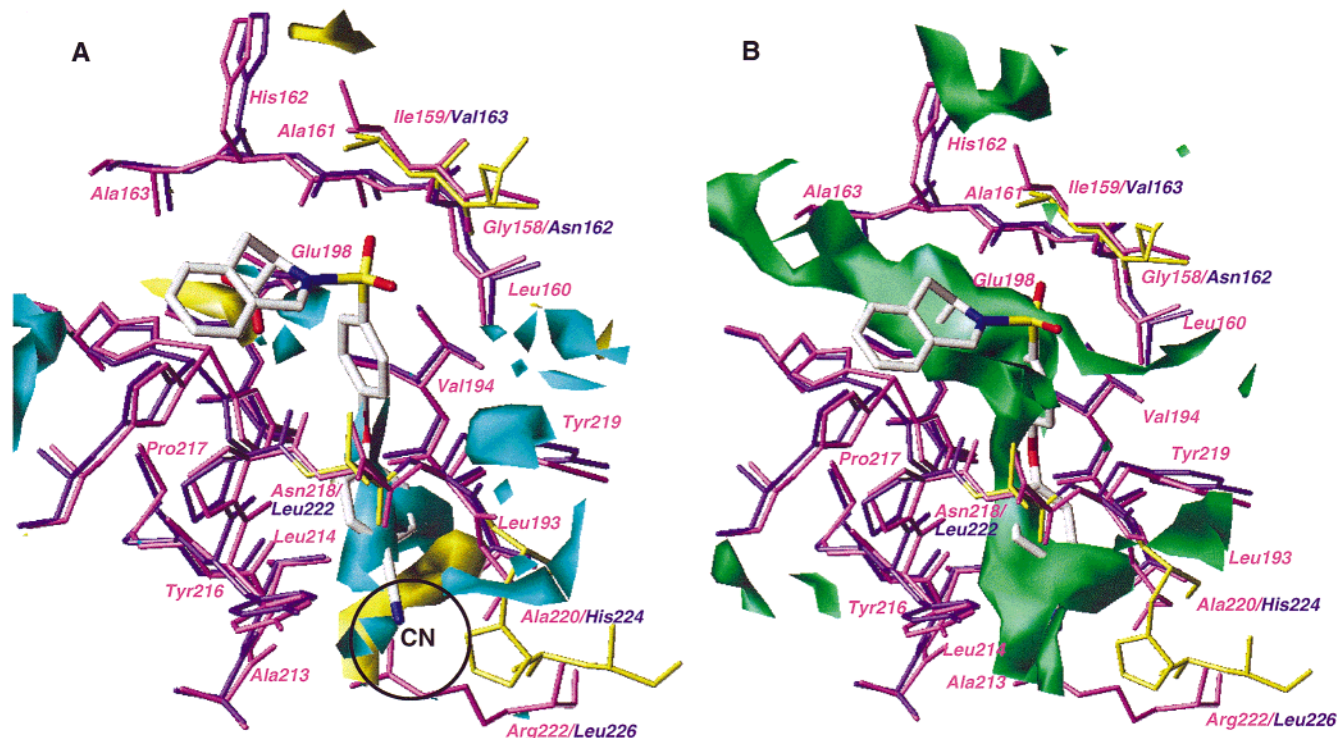
Figure 7. The 3D contour map of the PC1 (A) and PC2 (B) loadings for the PCA model using multivariate GRID descriptors. PC1 highlights selectivity regions (cyan: MMP-8, yellow: MMP-3); PC2 indicates unselective affinity regions for both targets. For reference, the inhibitor **84** is displayed. See Figure 3 for details.

MMP-8 selective variables circled in Figure 6B are located in the S1′ pocket, when plotting the loadings in 3D (Figure 7A). Similar variables in Figure 6B on the right-hand side indicate MMP-3 selectivity regions. In contrast, low values for PC1 and PC2 indicate regions exhibiting weak, unselective interactions, while high PC2 but low PC1 values refer to strong, unselective interactions, shown at the top of Figure 6B.

It should be noted that a PCA is a purely mathematical technique to simplify complex matrices of many thousand data points and to obtain an informative picture of the underlying data structure. This method does not force any discrimination between unspecific affinity and selectivity regions that is not clearly present in the underlying matrix.

In Figure 7A, 3D contour maps with PC1 loadings are indicating those selectivity regions where appropriate substituents are favorable to interact with only one of both MMPs. Selectivity regions for MMP-8 are indicated by cyan contours, while substitutions in yellow regions by appropriate substituents (probes) would improve selectivity toward MMP-3. The major protein differences were discussed above, like the replacement of Arg222 by Leu226 in MMP-3, the nonoccluded S1′ pocket, and the substitution of Ile159, Gly158 by Val, Asn in MMP-3 and Asn218, Ala220 of the upper rim of S1′ by Leu, His in MMP-3. The identified contour regions from this analysis are in agreement with S1′ differences, demonstrating its chemical significance to explain selective protein−ligand interactions.

For interpretation, the MMP-8 selective inhibitor **84** is shown. As suggested by the selectivity 3D-QSAR analysis, the para-cyano or other polar groups here are key determinants for selectivity, as it directly points to a cyan MMP-8 selectivity region at the bottom of S1′

close to Arg222. This preference is reflected by the position of the N:# probe atom (sp nitrogen with lone pair) in Figure 6A. Thus a discrimination between MMP-8 and MMP-3 is possible by placement of functional groups in those regions according to the ranking of probes in Figure 6A. The other cyan contour in the S1′ pocket is in agreement with heteroatom substitutions at the distal biphenyl ether ring, resulting in MMP-8 selective compounds with distal heterocycles or appropriate substituents. A selective example by inserting two aromatic nitrogen atoms into the distal aromatic ring is inhibitor **35**. A third MMP-8 selective example is only used for MMP-8 QSAR studies,[13] as its MMP-3 activity is too low (**26**, IC_{50} = ∼90000 nM MMP-3, 200 nM MMP-8). This molecule with a phenolic OH and a carbonyl oxygen linking two phenyl rings in S1′ highlights the additional MMP-8 S1′ selectivity region. There are other selective compounds, where characteristic functional groups are located toward corresponding protein selectivity regions, e.g., **63** (IC_{50} = 100 nM MMP-3, 1 nM MMP-8) with a para-dimethylamino group directed toward the cyan contour at the S1′ bottom. Inspection of Figure 7A explains the experimental finding that most compounds are MMP-8 selective, while there is only a limited possibility in the S1′ pocket to obtain MMP-3 selectivity. Another selectivity region is located in the S2 pocket, which is not probed by this set of inhibitors while it is occupied by an isobutyl side chain in the PDB file 1jap.[9] The majority of less selective inhibitors is characterized by S1′-directed side chains, interacting at common MMP-3 and -8 affinity regions. There is only a small number of slightly selective compounds for MMP-3 (e.g., **45**, IC_{50} = 400 nM MMP-3, 800 nM MMP-8) with a bulky group

too large for the occluded MMP-8 pocket, while the MMP-3 pocket tolerates more steric bulk.

In Figure 7B the regions for unselective strong ligand recognition, corresponding to positive PC2 loadings in Figure 6B, are shown as contour maps, revealing that almost the entire S1′ pocket plus the region on the left of the MMP-8 residue Ala161 are important for unspecific MMP affinity. This area corresponds to the binding region for the peptide Pro-Leu-Gly-NHOH, which occupies the unprimed subsites P1-P2-P3. All PC1 and PC2 maps correspond to 3D-QSAR results, and a consistent picture explaining affinity and selectivity of MMP-3 and MMP-8 inhibitors is obtained, which is in good agreement with experimental selectivity differences.

4. Conclusions

Understanding protein−ligand interactions is essential for designing novel synthetic candidates, while those interactions are difficult to describe. Here some useful insights for MMP inhibitors are presented by combining different computational approaches. A novel strategy to understand affinity and selectivity for metalloproteinase inhibitors using information from ligands plus protein 3D structures is successfully applied to a series of 2-arylsulfonyl-1,2,3,4-tetrahydro-isoquinoline-3-carboxylates and -hydroxamates as MMP-3 and -8 inhibitors. Although the combined interpretation of 3D-QSAR results for both targets highlights protein−ligand interactions, only limited information for selectivity were extracted. Hence, 3D-QSAR models were derived using the affinity ratio MMP-8/MMP-3, which led to key ligand determinants for selectivity. A consensus prediction using several QSAR models for all biological observables should lead to a reliable scoring to guide novel inhibitor design.

To complement this ligands' view on affinity and selectivity, a chemometrical approach using a principal component analysis uncovers differences of metalloproteinase binding sites with respect to their GRID probe interaction. A consistent picture in good agreement with experimental selectivity differences and 3D-QSAR models for selectivity is obtained, while all statistical models can be jointly interpreted and correspond to available data for binding site topologies, ligand affinities, and selectivities. The application of field-based 3D-QSAR techniques and other chemometrical approaches to extract knowledge from the protein binding site enhances our understanding of some aspects in protein−ligand interaction. From those models, potent and selective MMP inhibitors may emerge for several therapeutic approaches.

Supporting Information Available: Additional two figures with graphs showing experimental versus fitted biological activities for discussed PLS models and histograms from model validation studies. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Woessner, J. F., Jr. Matrix metalloproteinases and their inhibitors in connective tissue remodeling. *FASEB J.* **1991**, *5*, 2145−2154. (b) Birkedal-Hansen, H.; Moore, W. G. I.; Bodden, M. K.; Windsor, L. J.; Birkedal-Hansen, B.; DeCarlo, A.; Engler, J. A. Matrix metalloproteinases: a review. *Crit. Rev. Oral Biol. Med.* **1993**, *4*, 197−250. (c) Murphy, G.; Docherty, A. J. P. The matrix metalloproteinases and their inhibitors. *A. J. Res. Cell. Mol. Biol.* **1992**, *7*, 120−125. (d) Matrisian, L. M. Metalloproteinases and their inhibitors in matrix remodeling. *Trends Genet.* **1990**, *6*, 121−125.

(2) (a) Stetler-Stevenson, W. G.; Krutsch, H. C.; Liotta, L. A. Tissue inhibitor of metalloproteinase (TIMP-2). A new member of the metalloproteinase inhibitor family. *J. Biol. Chem.* **1989**, *264*, 17374−17378. (b) Docherty, A. J. P.; Lyons, A.; Smith, B. J.; Wright, E. M.; Stephens, P. E.; Harris, T. J. R.; Murphy, G.; Reynolds, J. J. Sequence of human tissue inhibitor of metalloproteinases and its identity to erythroid-potentiating activity. *Nature* **1985**, *318*, 66−69. (c) Enghild, J. J.; Salvesen, G.; Brew, K.; Nagase, H. Interaction of human rheumatoid synovial collagenase (matrix metalloproteinase 1) and stromelysin (matrix metalloproteinase 3) with human a2-macroglobulin and chicken ovostatin. Binding kinetics and identification of matrix metalloproteinase cleavage sites. *J. Biol. Chem.* **1989**, *264*, 8779−8785.

(3) (a) Murphy, G. J. P.; Murphy, G.; Reynolds, J. J. The origin of matrix metalloproteinases and their familial relationships. *FEBS Lett.* **1991**, 289, 4−7. (b) Stetler-Stevenson, W. G.; Liotta, L. A.; Kleiner, D. E., Jr. Extracellular matrix 6: role of matrix metalloproteinases in tumor invasion and metastasis. *FASEB J.* **1993**, *7*, 1434−1441. (c) Shingleton, W. D.; Hodges, D. J.; Brick, P.; Cawston, T. E. Collagenase: a key enzyme in collagen turnover. *Biochem. Cell Biol.* **1996**, *74*, 759−775.

(4) (a) Murphy, G.; Hembry, R. M. Proteinases in rheumatoid arthritis. *J. Rheumatol.* **1992**, *19*, 61−64. (b) Lohmander L. S.; Hoerrner L. A.; Lark M. W. Metalloproteinases, tissue inhibitor, and proteoglycan fragments in knee synovial fluid in human osteoarthritis. *Arthrit. Rheum.* **1993**, *36*, 181−189.

(5) Wernicke, D.; Seyfert, C.; Hinzmann, B.; Gromnica-Ihle, E. Cloning of collagenase 3 from the synovial membrane and its expression in rheumatoid arthritis and osteoarthritis. *J. Rheumatol.* **1996**, *23*, 590−595.

(6) Peress, N.; Perillo, E.; Zucker, S. Localization of tissue inhibitor of matrix metalloproteinases in Alzheimer's disease and normal brain. *J. Neuropathol. Exp. Neurol.* **1995**, *54*, 16−22.

(7) (a) Nishino, N.; Powers, J. C. Design of potent reversible inhibitors for thermolysin. Peptides containing zinc coordinating ligands and their use in affinity chromatography. *Biochemistry* **1979**, *18*, 4340−4347. (b) Powers, J. C.; Harper, J. W. Inhibitors of metalloproteinases. In *Proteinase Inhibitors*; Barrett, A. J., Salvesen, G., Eds.; Elsevier: Amsterdam, 1986; pp 219−298. (c) Johnson, W. H.; Roberts, N. A.; Borkakoti, N. Collagenase inhibitors: Their design and potential therapeutic use. *J. Enzyme Inhib.* **1987**, *2*, 1−22. (d) Zask, A.; Levin, J. I.; Killar, L. M.; Skotnicki, J. S. Inhibition of Matrix Metalloproteinases: Structure Based Design. *Curr. Pharm. Design* **1996**, *2*, 624−661. (e) Beckett, R. P.; Davidson, A. H.; Drummond, A. H.; Huxley, P.; Whittaker, M. Recent advances in matrix metalloproteinase inhibitor research. *Drug Discuss. Today* **1996**, *1*, 16−26. (f) Beckett, R. P. Recent advances in the field of matrix metalloproteinase inhibitors. *Exp. Opin. Ther. Patents* **1996**, *6*, 193−198. (g) Beckett, R. P.; Whittaker, M. Matrix metalloproteinase inhibitors 1998. *Exp. Opin. Ther. Patents* **1998**, *8*, 259−282.

(8) Becker, J. W.; Marcy, A. I.; Rokosz, L. L.; Axel, M. G.; Burbaum, J. J.; Fitzgerald, P. M. D.; Cameron, P. M.; Esser, C. K.; Hagmann, W. K.; Hermes, J. D.; Springer, J. P. Stromelysin-1: Three-dimensional structure of the inhibited catalytic domain and of the C-truncated proenzyme. *Protein Sci.* **1995**, *4*, 1966−1976.

(9) Stams, T.; Spurlino, J. C.; Smith, D. L.; Wahl, R. C.; Ho, T. F.; Qoronfleh, M. W.; Banks, T. M.; Rubin, B. Structure of human neutrophil collagenase reveals large S1′ specificity pocket. *Nature Struct. Biol.* **1994**, *1*, 119−123.

(10) Grams, F.; Reinemer, P.; Powers, J.; Kleine, T.; Pieper, M.; Tschesche, H.; Huber, R.; Bode, W. X-ray structures of human neutrophil collagenase complexed with peptide hydroxamate and peptide thiol inhibitors; implications for substrate binding and rational drug design. *Eur. J. Biochem.* **1995**, *228*, 830−841.

(11) Reinemer, P.; Grams, F.; Huber, R.; Kleine, T.; Schnierer, S.; Pieper, M.; Tschesche, H.; Bode, W. Structural implications for the role of the N-terminus in the "superactivation" of collagenases. *FEBS Lett.* **1994**, *338*, 227−233.

(12) Grams, F.; Crimmin, M.; Hinnes, L.; Huxley, P.; Pieper, M.; Tschesche, H.; Bode, W. Structure determination and analysis of human collagenase complexed with a hydroxamate inhibitor. *Biochemistry* **1995**, *34*, 14012–14020.

(13) Matter, H.; Schwab, W.; Barbier, D.; Billen, G.; Haase, B.; Neises, B.; Schudok, M.; Thorwart, W.; Schreuder, H.; Brachvogel, V.; Lönze, P.; Weithmann, K. U.Quantitative Structure–Activity Relationship of Human Neutrophil Collagenase (MMP-8) Inhibitors Using Comparative Molecular Field Analysis and X-ray Structure Analysis. *J. Med. Chem.* **1999**, *42*, 1908–1920.

(14) PDB files from Protein Database (National Brookhaven Laboratories): 1MMB, 1MNC, 1KBC, 1JAN, 1JAO, 1JAP, and 1JAQ (http://www.pdb.bnl.gov). Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **1977**, *112*, 535–542.

(15) (a) Goodford, P. J. A Computational Procedure for Determining Energetically Favorable Binding Sites on Biologically Important Macromolecules. *J. Med. Chem.* **1985**, *28*, 849–857. (b) Boobbyer, D. N. A.; Goodford, P. J.; McWhinnie, P. M.; Wade, R. C. New hydrogen-bond potentials for use in determining energetically favorable binding sites on molecules of known structure. *J. Med. Chem.* **1989**, *32*, 1083–1094. (c) Wade, R. C.; Clerk, 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. (d) 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.

(16) Goodford, *P. GRID version 16 for SGI*; Molecular Discovery Ltd.: Oxford, U.K., 1996.

(17) Pastor, M.; Cruciani, G. A Novel Strategy for Improving Ligand Selectivity in Receptor-Based Drug Design. *J. Med. Chem.* **1995**, *38*, 4637–4647.

(18) Cruciani, G.; Goodford, P. J. A search for specificity in DNA-drug interactions. *J. Mol. Graphics* **1994**, *12*, 116–129.

(19) Cramer, R. D.; Patterson, D. E.; Bunce, J. E. Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959–5967.

(20) Clark, M.; Cramer, R. D.; Jones, D. M.; Patterson, D. E.; Simeroth, P. E. Comparative Molecular Field Analysis (CoMFA). 2. Towards its use with 3D-Structural Databases. *Tetrahedron Comput. Methods* **1990**, *3*, 47–59.

(21) *3D-QSAR in Drug Design. Theory, Methods and Applications*; Kubinyi, H., Ed.; ESCOM: Leiden (NL), 1993.

(22) (a) Klebe, G.; Abraham, U.; Mietzner, T. Molecular Similarity Indices in a Comparative Analysis (CoMSIA) of Drug Molecules to Correlate and Predict Their Biological Activity. *J. Med. Chem.* **1994**, *37*, 4130–4146. (b) Klebe, G.; Abraham, U. Comparative molecular similarity index analysis (CoMSIA) to study hydrogen-bonding properties and to score combinatorial libraries. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 1–10. (c) Böhm, M.; Stürzebecher, J.; Klebe, G. Three-Dimensional Quantitative Structure–Activity Relationship Analyses Using Comparative Molecular Field Analysis and Comparative Molecular Similarity Indices Analysis To Elucidate Selectivity Differences of Inhibitors Binding to Trypsin, Thrombin, and Factor Xa. *J. Med. Chem.* **1999**, *42*, 458–477. (d) Klebe, G. Comparative molecular similarity indices analysis. CoMSIA. *Perspect. Drug Discovery Des.* **1998**, *12/13/14*, 87–104.

(23) Baroni, M.; Costantino, G.; Cruciani, G.; Riganelli, D.; Valigi, R.; Clementi, S. Generating Optimal Linear PLS Estimations (GOLPE): An Advanced Chemometric Tool for Handling 3D-QSAR Problems. *Quant. Struct.-Act. Relat.* **1993**, *12*, 9–20.

(24) (a) Wold, S.; Albano, C.; Dunn, W. J.; Edlund, U.; Esbenson, K.; Geladi, P.; Hellberg, S.; Lindberg, W.; Sjöström, M. In *Chemometrics: Mathematics and Statistics in Chemistry*; Kowalski, B., Ed.; Reidel: Dortrecht, The Netherlands, 1984; pp 17–95. (b) Dunn, W. J.; Wold, S.; Edlund, U.; Hellberg, S.; Gasteiger, J. Multivariate Structure–Activity Relationship Between Data from a Battery of Biological Tests and an Ensemble of Structure Descriptors: The PLS Methodology. *Quant. Struct.-Act. Relat.* **1984**, *3*, 31–137. (c) Geladi, P. Notes on the history and nature of Partial Least Squares (PLS) modelling. *J. Chemom.* **1988**, *2*, 231–246. (d) Geladi, P.; Kowalski, B. R. Partial least-squares regression: a tutorial. *Anal. Chem. Acta* **1986**, *185*, 1–17. (e) Jöreskog, K., Wold, H., Eds. *Chemical Systems under Indirect Observation*; North-Holland: Amsterdam, NL, 1982.

(25) (a) Wold, S. Cross-Validatory Estimation of the Number of Component in Factor and Principal Component Models. *Technometrics* **1978**, *4*, 397–405. (b) Diaconis, P.; Efron, B. Computer-Intensive Methods for Statistics. *Sci. Am.* **1984**, *116*, 96–117.

(c) Cramer, R. D.; Bunce, J. D.; Patterson, D. E. Crossvalidation, Bootstrapping and Partial Least Squares Compared with Multiple Regression in Conventional QSAR Studies. *Quant.-Struct.-Act. Relat.* **1988**, *7*, 18–25.

(26) Schlechter, I.; Berger, A. On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157–162.

(27) Schwab, W.; Thorwart, W.; Barbier, D.; Billen, G.; Haase, B.; Matter, H.; Neises, B.; Schudok, M.; Weithmann, K. U. Tetrahydroisoquinoline-3-carboxylate based Matrix Metalloproteinase Inhibitors: Design, Synthesis and Structure–Activity Relationship. *Bioorg. Med. Chem.*, to be submitted.

(28) Thorwart, W.; Schwab, W.; Schudok, M.; Haase, B.; Bartnik, E.; Weithmann, K.- Preparation of cyclic N-substituted α-iminohydroxamates as matrix metalloproteinase inhibitors. U. Ger. Offen., 17 pp. CODEN: GWXXBX. DE 19542189 A1 970515. CAN 127:50547.

(29) *SYBYL Molecular Modelling Package, Versions 6.4, 6.5*; Tripos: St. Louis, MO, 1997/8.

(30) Clark, M.; Cramer, R. D.; Van Opdenbosch, N. Validation of the general purpose Tripos 5.2 force field. *J. Comput. Chem.* **1989**, *10*, 982–1912.

(31) Gasteiger, J.; Marsili, M. Iterative partial equalization of orbital electronegativity: a rapid access to atomic charges. *Tetrahedron* **1980**, *36*, 3219–3228.

(32) (a) Ghose, A.; Crippen, G. Atomic Physicochemical Parameters for Three-Dimensional Structure-Directed Quantitative Structure–Activity Relationships. 1. Partition Coefficients as a Measure of Hydrophobicity. *J. Comput. Chem.* **1986**, *7*, 565–577. (b) Viswanadhan, V. N.; Ghose, A. K.; Revankar, G. R.; Robins, R. K. Atomic Physicochemical Parameters for Three-Dimensional Structure-Directed Quantitative Structure–Activity Relationships. 4. Additional Parameters for Hydrophobic and Dispersive Interactions and Their Application for an Automated Superposition of Certain Naturally Occurring Nucleoside Antibiotics. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 163–172.

(33) Kearsley, S. K.; Smith, G. M. An Alternative Method for the Alignment of Molecular Structures: Maximizing Electrostatic and Steric Overlap. *Tetrahedron Comput. Method* **1990**, *3*, 615–633.

(34) (a) Thibaut, U.; Folkers, G.; Klebe, G.; Kubinyi, H.; Merz, A.; Rognan, D. Recommendations for CoMFA Studies and 3D QSAR Publications. In *3D QSAR in Drug Design. Theory, Methods and Applications*; Kubinyi, H., Ed., ESCOM: Leiden, The Netherlands, 1993; pp 711–717. (b) Folkers, G.; Merz, A.; Rognan, D. CoMFA: Scope and Limitations. Ibid. pp 583–616. (c) Cramer, R. D.; DePriest, S. A.; Patterson, D. E.; Hecht, P. The Developing Practice of Comparative Molecular Field Analysis. Ibid. pp 443–485.

(35) Sheridan, R. P.; Nachbar, R. B.; Bush, B. L. Extending the trend vector: The trend matrix and sample-based partial least squares. *J. Comput.-Aided Mol. Design* **1994**, *8*, 323–340.

(36) Clementi, S. *Golpe 3.0, SGI version*; Multivariate Infometrics Analyses (MIA): Perugia, Italy, 1995.

(37) 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*, 2589–2601.

(38) Nilsson, J.; Wikström, H.; Smilde, A.; Glase, S.; Pugsley, T.; Cruciani, G.; Pastor, M.; Clementi, S. GRID/GOLPE 3D Quantitative Structure–Activity Relationship Study on a Set of Benzamides and Naphthamides, with Affinity for the Dopamine D3 Receptor Subtype. *J. Med. Chem.* **1997**, *40*, 833–840.

(39) (a) Mitchell, T. J. An Algorithm for the Construction of "D-optimal" Experimental Designs. *Technometrics* **1974**, *16*, 203–210. (b) Steinberg, D. M.; Hunter, W. G. Experimental Design Review and Comment. *Technometrics* **1984**, *26*, 71–76.

(40) Morgan, E. In *Chemometrics: Experimental Design*; Chadwick, N., Ed.; John Wiley and Sons Ltd.: Chichester, U.K., 1991.

(41) Cruciani, G.; Baroni, M.; Costantino, G.; Riganelli, D.; Skagerberg, B. Predictive Ability of Regression Models. Part 1: Standard Deviation of Prediction Errors (SDEP). *J. Chemom.* **1992**, *6*, 335–346.

(42) (a) Pastor, M.; Cruciani, G.; Clementi, S. Smart Region Definition: A New Way to Improve the Predictive Ability and Interpretability of Three-Dimensional Quantitative Structure–Activity Relationships. *J. Med. Chem.* **1997**, *40*, 1455–1464. (b) Clementi, S.; Cruciani, G.; Riganelli, D.; Valigi, R. GOLPE: Mertis and Drawbacks in 3D-QSAR. In *QSAR and Molecular Modelling: Concepts, Computational Tools and Biological Applications*; Sanz, F., Giraldo, J., Manaut, F., Eds.; Prous Science Publ.: Barcelona, 1996; pp 408–414.

(43) (a) Cho, S. J.; Tropsha, A. Cross-Validated R^2-Guided Region Selection for Comparative Molecular Field Analysis: A Simple Method to Achieve Consistent Results. *J. Med. Chem.* **1995**, *38*, 1060–1066. (b) Norinder, U. Single and Domain Mode Variable Selection in 3D QSAR Applications. *J. Chemometr.* **1996**, *10*, 95–105.

(44) (a) Dillon, W. R.; Goldstein, M. *Multivariate Analysis: Methods and Applications*; Wiley: New York, 1984. (b) Malinowski, E. R.; Howery, D. G. *Factor Analysis in Chemistry*; Wiley: New York, 1980. (c) Cramer, R. D. BC(DEF) Parameters. 1. The Intrinsic Dimensionality of Intermolecular Interactions in the Liquid State. *J. Am. Chem. Soc.* **1980**, *102*, 1837−1849. (d) Stahle, L.; Wold, S. Multivariate Data Analysis and Experimental Design in Biomedical Research. in *Progress in Medicinal Chemistry*; Ellis, G. P., West, G. B., Eds.; Elsevier: Amsterdam, 1988; pp 292−338. (e) Wold, S.; Albano, C.; Dunn, W. J., III; Edlund, U.; Esbensen, K.; Geladi, P.; Hellberg, S.; Johanson, E.; Lindberg, W.; Sjöström, M. Multivariate Data Analysis in Chemistry. In *Chemometrics: Mathematics and Statistics in Chemistry*; Kowalski, B. R., Ed.; NATO, ISI Series C 138; D. Reidel Publ. Co.: Dordrecht, Holland, 1984; pp 17−96.

(45) Biological activities are expressed as $\log(1/IC_{50} \times 100\,000)$.

(46) Clark, M.; Cramer, R. D. The Probability of Chance Correlation using Partial Least Squares (PLS). *Quant. Struct.-Act. Relat.* **1993**, *12*, 137−145.

(47) Betz, M.; Huxley, P.; Davies, S. J.; Mushtaq, Y.; Pieper, M.; Tschesche, H.; Bode, W.; Gomis-Rüth, F. X. 1.8 Å crystal structure of the catalytic domain of human neutrophil collagenase (mtarix metalloproteinase-8) complexed with a peptidomimetic hydroxamate primed-side inhibitor with a distinct selectivity profile. *Eur. J. Biochem.* **1997**, *247*, 356−363.

(48) Wong, G.; Koehler, K. F.; Skolnick, P.; Gu, Z.; Ananthan, S.; Schönholzer, P.; Hunkeler, W.; Zhang, W.; Cook, J. M. Synthetic and Computer-Assisted Analysis of the Structural Requirements for Selective, High Affinity Ligand Binding to Diazepam-Insensitive Benzodiazepine Receptors. *J. Med. Chem.* **1993**, *36*, 1820−1830.

(49) Cheng, Y.; Prusoff, W. H. Relationship between the inhibition constant (K_i) and the concentration of inhibitor which causes 50% inhibition (I_{50}) of an enzymatic reaction. *Biochem. Pharm.* **1973**, *22*, 3099−3108.

(50) (a) *ReliBase Version 3.2.* Febr. 24. 1999 (8959 PDB complexes) from M. Hendlich, University Marburg, Germany: http://www2.ebi.ac.uk: 8081/home.html. (b) Hendlich, M.; Rippmann, F.; Barnickel, G.; Hemm, K.; Aberer, K. RELIBase − an object-oriented comprehensive receptor−ligand database. *Folding Des.* **1996**, *1*, 30. (c) Hendlich, Manfred. Databases for protein−ligand complexes. *Acta Crystallogr.* **1998**, *D54*, 1178−1182.

(51) Fitzpatrick, P. A.; Steinmetz, A. C. U.; Ringe, D.; Klibanov, A. M. Enzyme crystal structure in a neat organic solvent. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8653−8657.

(52) Guddat, L. W.; Shan, L.; Anchin, J. M.; Linthicum, D. S.; Edmundson, A. B. Local and transmitted conformational changes on complexation of an anti-sweetener Fab. *J. Mol. Biol.* **1994**, *236*, 247−274.

(53) (a) Skulnick, H. I.; Johnson, P. A.; Aristoff, P. A.; Morris, J. K.; Lovasz, K. D.; Howe, W. J.; Watenpaugh, K. D.; Janakiraman, M. N.; Anderson, D. J.; Reischer, R. J.; Schwartz, T. M.; Banitt, L. S.; Tomich, P. K.; Lynn, J. C.; Horng, M.; Chong, K.; Hinshaw, R. R.; Dolak, L. A.; Seest, E. P.; Schwende, F. J.; Rush, B. D.; Howard, G. M.; Toth, L. N.; Wilkinson, K. R.; Kakuk, T. J.; Johnson, C. W.; Cole, S. L.; Zaya, R. M.; Zipp, G. L.; Possert, P. L.; Dalga, R. J.; Zhong, W.; Williams, M. G.; Romines, K. R. Structure-Based Design of Nonpeptidic HIV Protease Inhibitors: The Sulfonamide-Substituted Cyclooctylpyranones. *J. Med. Chem.* **1997**, *40*, 1149−1164. (b) Skulnick, H. I.; Johnson, P. D.; Howe, W. J.; Tomich, P. K.; Chong, K.; Watenpaugh, K. D.; Janakiraman, M. N.; Dolak, L. A.; McGrath, J. P. Structure-Based Design of Sulfonamide-Substituted Non-Peptidic HIV Protease Inhibitors. *J. Med. Chem.* **1995**, *38*, 4968−4971.