Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry 12 (2004) 3607–3617
Bioorganic & Medicinal Chemistry

GRIND/ALMOND investigations on $CysLT_1$ receptor antagonists of the quinolinyl(bridged)aryl type

Paolo Benedetti,[a] Raimund Mannhold,[b] Gabriele Cruciani[a,*] and Giorgio Ottaviani[a]

[a]*Dipartimento di Chimica, Laboratorio di Chemiometria, Università di Perugia, Via Elce di Sotto, 10, I-06123 Perugia, Italy*
[b]*Department of Laser Medicine, Molecular Drug Research Group, Heinrich-Heine-Universität, Universitätsstr. 1, D-40 225 Düsseldorf, Germany*

Received 30 January 2004; accepted 16 April 2004
Available online 18 May 2004

Abstract—One of the current routes in developing antiasthmatics is $CysLT_1$ receptor antagonism. For a training set of 54 $CysLT_1$ receptor antagonists of the quinolinyl(bridged)aryl type we developed chemometric QSAR models applying GRID independent descriptors (=GRIND). PLS analysis resulted in a two-component model explaining 67% of the variance for $CysLT_1$ receptor binding ($r^2 = 0.67$, SDEC = 0.47, $q^2 = 0.54$). GRIND variables 11-50 and 22-55 are responsible for high-affinity binding; variable 11-62 is detrimental. The predictivity of the above chemometric model is tested with a set of 69 $CysLT_1$ receptor antagonists, exhibiting varying chemical similarity to the training set. Nearly 50% of the test set are quite well predicted. The quality of prediction coincides in part with chemical subclassification: phenylene bridged compounds are quite well predicted; for structures with bridging heterocycles predictions are rather poor. For explaining the outlier behavior, a PLS discriminant analysis including the training set and the strongest outliers of the test set was performed. The scores plot of discriminant PLS shows an almost complete separation between the two subsets. A PLS coefficients plot explains which GRIND variables are important for the discrimination between the training set and the outliers of the test set.

1. Introduction

Asthma is an inflammatory condition of the airways characterized by bronchial hyperresponsiveness, airways obstruction, and epithelial damage.[1–3] Pathological changes are produced by the actions of various mediators, being released from inflammatory cells as a result of antigen recognition by IgE receptors. The inflammatory mediators have a variety of effects on target cells in the airways and may induce many of the pathological features found in asthma.[4] Mediators produce their effects on target cells by the activation of specific cell surface receptors.[2] Traditional asthma therapy mainly focused on relief of symptoms by the administration of mast cell stabilizers or bronchodilators.[5,6] The attention for new therapeutic strategies has shifted to suppression of the inflammatory component, sometimes in combination with bronchodilation.[7,8] For future development of antiasthmatics, one direction currently followed is blocking the actions of multiple inflammatory mediators. Leukotrienes were identified as inflammatory mediators in the late 1970s.

Specific modulation of the effects of leukotrienes is possible by intervention of the biosynthesis or by blocking the specific leukotriene receptors. Three types of leukotriene receptors have been identified in humans: BLT, $CysLT_1$, and $CysLT_2$. The $CysLT_1$ receptor, formerly LTD_4,[9] has been most thoroughly characterized.[10–13] Radioligand binding studies classified this receptor as G-protein-coupled and photoaffinity labeling in guinea pig lung identified it as a single polypeptide with a molecular mass of 45 kDa.[14] In addition to respiratory tissues,[11] $CysLT_1$ receptors are located in ileum, gall bladder, stomach, cardiovascular system, and gastrointestinal tract.[10,11,13] Pharmacological studies revealed that the bronchoconstriction, mucus secretion, and possibly also the inflammatory condition of the airways are caused by an activation of the $CysLT_1$ receptor.

Several highly potent $CysLT_1$ receptor antagonists with large structural diversity have been developed.[15] Beyond antagonists with structures analogous to cysteinyl

Keywords: $CysLT_1$ receptor antagonists; Grid independent descriptors (GRIND); PLS analysis; PLS discriminant analysis.
* Corresponding author. Tel./fax: +39-07545646; e-mail: gabri@chemiome.chm.unipg.it

0968-0896/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2004.04.018

leukotrienes, other subclasses comprise: (a) hydroxyacetophenones, (b) indoles and indazoles, (c) quinolines, benzothiazoles, and thiazoles. On the basis of structure–activity analyses several authors developed pharmacophore models for $CysLT_1$ receptor antagonists.[16–18]

The aim of the present study is to develop valid and predictive chemometric QSAR models for a set of $CysLT_1$ receptor antagonists, to demonstrate the applicability of GRID independent descriptors (=GRIND) for this purpose and to compare our GRIND/ALMOND approach with previous pharmacophore studies.

2. Datasets and methodology

2.1. Datasets

The training set (Table 1) comprises 54 $CysLT_1$ receptor antagonists of the quinolinyl(bridged)aryl type, described by Palomer et al.[18] Mainly, the bridging ring system is a naphthalene or benzene moiety. Test set structures ($n = 69$) were described by Zwaagstra and co-workers;[17,19] they exhibit varying chemical relation to the training set and are given in Table 2. In 39 cases the bridging ring moiety is benzene; bridging rings in the remaining molecules are piperazine, homopiperazine, or piperidine.

2.2. Structural descriptors

GRID independent descriptors, GRIND,[20] were generated, analyzed, and interpreted using the software ALMOND, version 3.0.[21] GRIND have been designed mainly to represent pharmacodynamic properties; they start from molecular interaction fields (MIF) computed with the program GRID.[22] When MIF are computed for the database molecules, the regions showing favorable energies of interaction represent positions where groups of a receptor would interact favorably with a database molecule. Using different probes, one can obtain a set of such positions, which defines a virtual receptor site (VRS). GRIND are based on the concept of VRS. Basically, GRIND are a small set of variables representing the geometrical relationships between relevant regions of the VRS. The procedure for obtaining GRIND involves three steps: (a) computing a set of MIF, (b) filtering the MIF to extract the most relevant regions that define the VRS, and (c) encoding the VRS into the GRIND variables. The alignment-independent GRIND, obtained this way, can be used directly for the chemometric analysis and can be interpreted with the appropriate software,[21] using graphical representations of the pharmacophoric regions and their interactions, together with the molecular structures, in interactive 3D plots.

2.3. Statistical analysis

Partial least squares (PLS) analysis[23] and discriminant PLS analysis were performed within the software ALMOND. No scaling was applied. The optimal dimensionality of the PLS model was chosen according to the results of crossvalidation. All computations were run on Silicon Graphics SGI O2 R10 000 and R12 000 workstations.

3. Results and discussion

3.1. Chemometric model building

Affinities of the 54 training set molecules of the quinolinyl(bridged)aryl type were determined by Palomer et al.[18] via binding studies in guinea pig lung membranes using $[^3H]$-LTD_4 as radioligand. Binding affinities, listed in Table 3, range from a pK_i-value of 9.0 for the most potent compound **1** to values of less than 6.0 for the rather weakly active compounds **46**, **49**, and **54** and cover a spectrum of more than 3 log units.

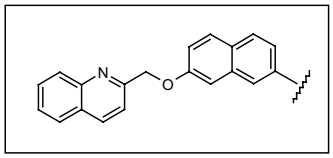
The 3D structures of the training set were generated in their neutral form and minimized with the software SYBYL[24] using the MMFF force field parameterized in vacuo. For developing QSAR models we performed chemometric analyses using the PLS option of the software ALMOND. As chemical descriptors the alignment-independent GRIND (=grid independent descriptors) were applied. For deriving GRIND the hydrophobic probe DRY, the N1 probe (amide nitrogen, representing HB acceptor interactions), and the O probe (carbonyl oxygen, representing HB donor interactions) were used; following default parameters were applied: 0.5 Å grid spacing, 250 nodes, 65% of weight assigned to the distance, smoothing window width 0.8. A total of 129 descriptors of the X-space was derived after variable selection via fractional factorial design. In preliminary calculations, the O probe failed to contribute to the model. For a better graphical representation, it was neglected in the final calculations and documentation. PLS analysis resulted in a two-component model explaining 67% of the variance for $CysLT_1$ receptor binding; a correlation coefficient of 0.82 and a standard deviation of the error of calculation (SDEC) of 0.47 was found. Validation of this model was performed using five random groups (100 SDEP calculations); the q^2 of the validated model was 0.54.

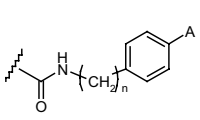
The PLS coefficients plot in Figure 1 allows to locate those chemical descriptors (X variables), which are directly or inversely correlated to the biological activity (Y variable): a positive value of a coefficient corresponds to a direct correlation to the Y; negative ones indicate an inverse correlation; large values mean a strong and low values a marginal impact on the Y variable. Coding of the GRIND variables is as follows: the first two digits indicate the probes used, digit 1 labels the hydrophobic probe DRY and digit 2 the hydrogen bond acceptor probe N1; the second number corresponds to distances between grid nodes representing favorable interactions with one of the probes. For an estimate of this distance in Å, the corresponding number has to be multiplied by the grid spacing (0.5); for details see Ref. 20. To give an example, variable 11-62 characterizes hydrophobic

interactions with the DRY probe at two grid nodes separated roughly 31 Å from each other. The PLS coefficients in Figure 1 show that GRIND variables 11-50 and 22-55 are responsible for high-affinity $CysLT_1$ receptor binding, whereas variable 11-62 is detrimental for binding.

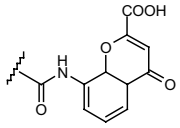
To elucidate the involved interacting regions in an easily understandable way, they are graphically shown in Figures 2 and 3 for a highly potent (no **17**) and a weakly active training set molecule (no **45**). Figure 2 demonstrates favorable hydrophobic interaction fields (11-50, distance 25 Å) due to the quinoline and the benzene ring,

Table 1. Structures of the training set molecules

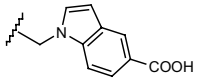




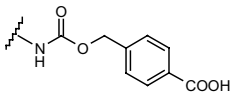
n = 2, A = COOH, **1**
n = 0, A = $(CH_2)_3COOH$, **2**
n = 1, A = Tetrazole, **3**
n = 2, A = Tetrazole, **4**
n = 3, A = Tetrazole, **5**
n = 2, A = $CONHSO_2Ph(o\text{-}CH_3)$, **6**



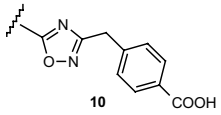
7



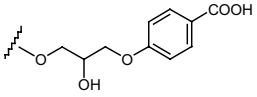
8



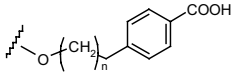
9



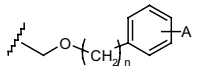
10



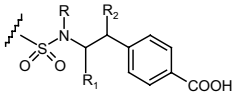
11



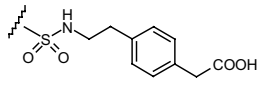
n = 2, **12**
n = 3, **13**



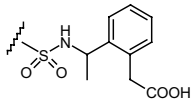
n = 2, A = p-COOH, **14**
n = 1, A = p-COOH, **15**
n = 2, A = m-COOH, **16**



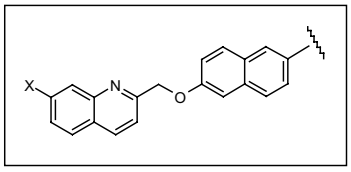
R = R_1 = R_2 = H, **17**
R = Me, **18**
R_1 = Me, **19**
R_2 = $(Me)_2$, **20**

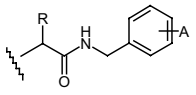


21

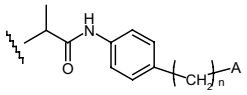


22

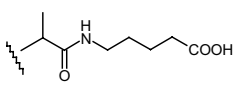




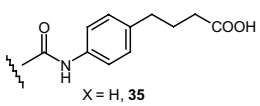
R = Me, A = p-COOH, X = H, **23**
R = Me, A = p.Tetrazole, X = H, **24**
R = H, A = p-Tetrazole, X = H, **25**
R = Me, A = o-Tetrazole, X = H, **26**
R = Me, A = m-COOH, X = H, **27**
R = Me, A = m-Tetrazole, X = H, **28**



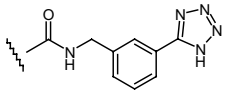
n = 0, A = COOH, X = H, **29**
n = 1, A = COOH, X = H, **30**
n = 1, A = Tetrazole, X = H, **31**
n = 3, A = COOH, X = H, **32**
n = 3, A = COOH, X = Cl, **33**



X = H, **34**



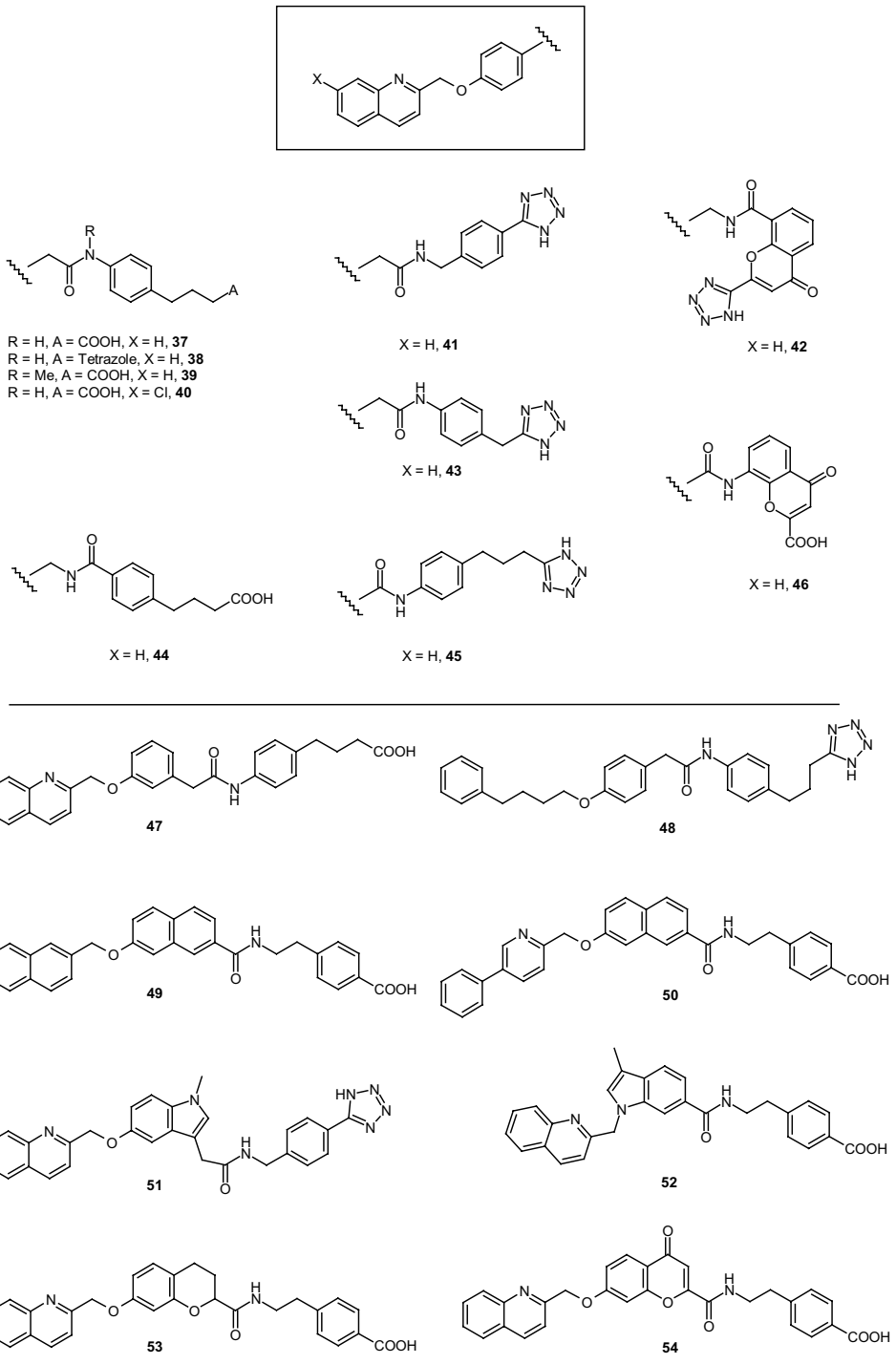
X = H, **35**



X = H, **36**

(*continued on next page*)

Table 1 (*continued*)



as well as detrimental hydrophobic interaction fields (11-62, distance 31 Å), due to the quinoline and the tetrazole moiety. The correlograms in the left-hand part show that the favorable interactions (11-50) are strong and the detrimental interactions (11-62) are weak for the highly potent compound **17**. The opposite holds for the weakly active compound **45**. To interpret node–node interactions in terms of binding strength, one has to consider positioning of the nodes. Positioning in the center of the interaction field indicates strong binding, whereas positioning in the periphery indicates weak binding. Thus, in the case of the highly active compound **17** nodes for the favorable GRIND variable 11-50 are positioned in the center of the interaction field, whereas they are positioned in the periphery for the weakly active compound **45**.

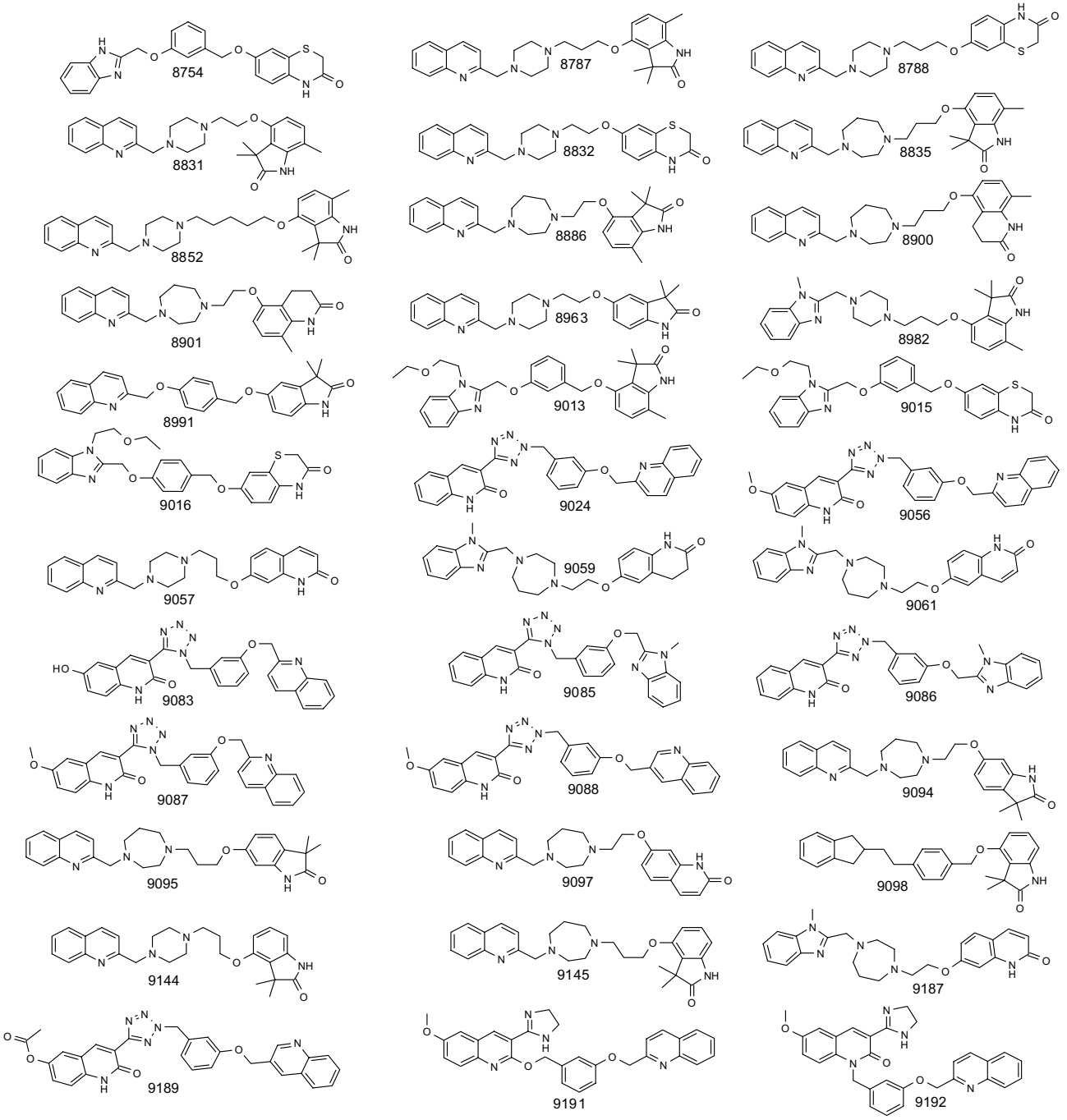
Figure 3 reflects hydrogen bond (HB) interaction fields for HB accepting moieties in the same molecules, represented by the quinolinic nitrogen, the carboxyl (structure **17**), respectively, the tetrazole moiety (structure **45**). As already shown in the PLS coefficients plot in

Figure 1, the GRIND variable 22-55 is mainly indicative for high-affinity receptor binding. The correlograms in the left-hand part of Figure 3 immediately indicate that this interaction is present for the highly potent compound **17**, whereas it is absent for the weakly active compound **45**.

As already mentioned, the training set molecules used here were previously applied by Palomer et al.[18] for deriving pharmacophore and CoMFA models. Pharmacophoric features from their study can be summarized as follows: (a) an acidic or negative ionizable function, (b) three hydrophobic regions, represented by the three aromatic ring systems, and (c) a hydrogen bond acceptor moiety, as given in the quinolinic nitrogen. Comparison of the pharmacophoric features defined by Palomer et al.[18] with the results of our GRIND/ALMOND approach indicates a high similarity between these completely different methods. The exception is that we detect two favorable hydrophobic regions, represented by the quinoline and the benzene moieties (see also Fig. 4).

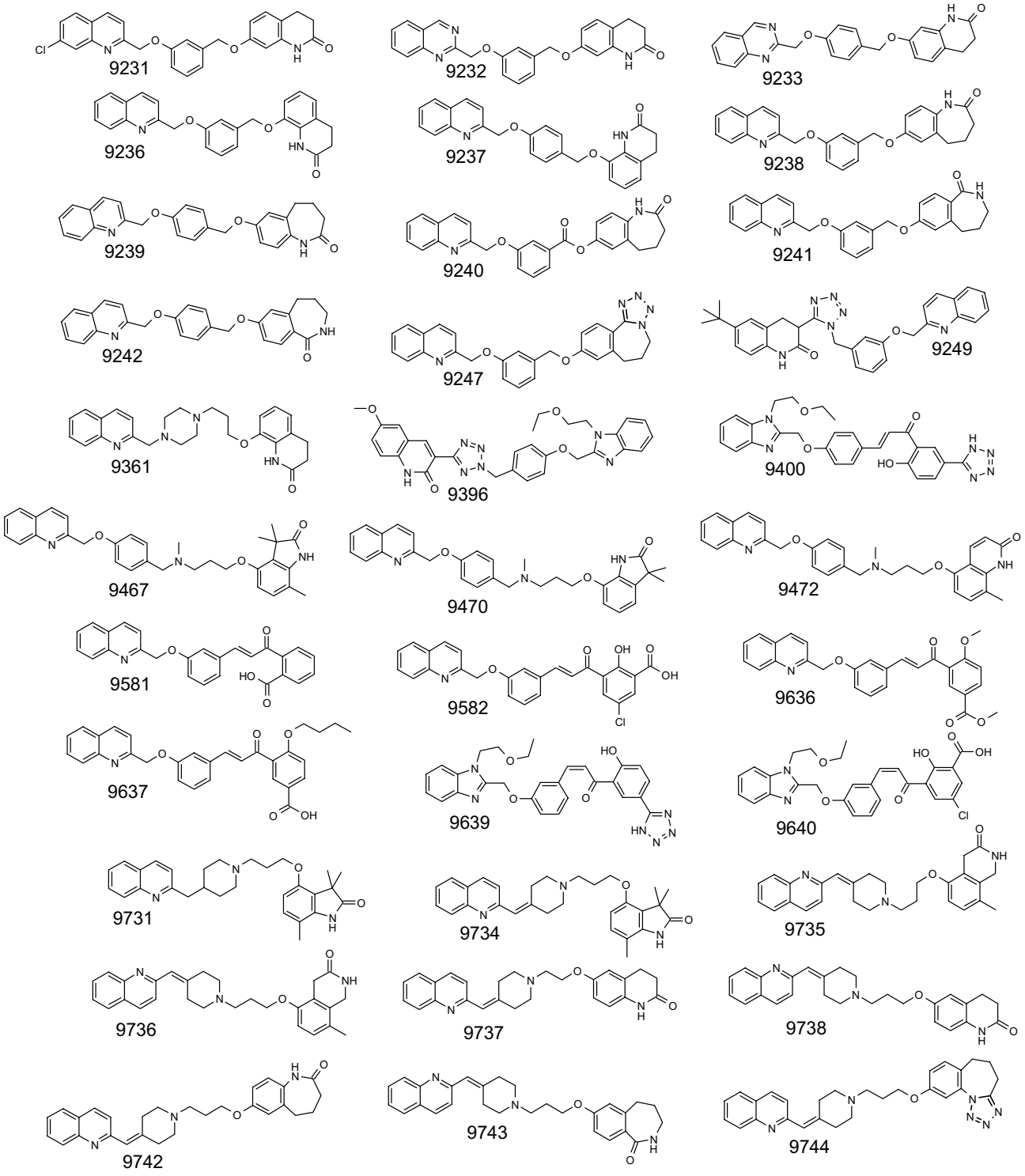
Table 3 allows a view on the predictive performance of the two approaches. It should be noticed that the data of Palomer et al.[18] represent their combined training and

Table 2. Structures of the test set molecules



(*continued on next page*)

Table 2 (*continued*)



test set, whereas we used their entire database as training set. For both approaches Table 3 lists the calculated pK_i-values as well as their differences ($= \Delta$) to the experimental pK_i. For the GRIND/ALMOND approach calculations fit the experimental data in most cases. Fourteen out of 54 training set molecules deviate by more than 0.5 log units; strongest outlier is compound **27** with a Δ of 1.1 log unit. As a criterion for the overall quality of calculations we find an averaged absolute residual sum of 0.38, which amounts to 0.40 in the case of the calculations of Palomer et al.[18] However, the latter authors had to omit four molecules due to missing fit to their models, whereas the GRIND/ALMOND approach satisfactorily calculates the entire set. It is not the scope of this comparison to judge a competitive quality of the two approaches rather than to show the good overall performance of the GRIND/ALMOND approach, which has simple computational

Table 3. Experimental and calculated $CysLT_1$ receptor binding data for the training set

No	Exp. K_i	Exp. pK_i	Palomer et al.[18]		This study	
			Calcd pK_i	Δ	Calcd pK_i	Δ
1	1.05 ± 0.08	9.0	8.8	−0.2	8.4	−0.6
2	1.5 ± 0.25	8.8	8.7	−0.1	8.5	−0.3
3	7.8 ± 2	8.1	8.2	+0.1	8.0	−0.1
4	3.4 ± 1.9	8.5	8.0	−0.5	8.2	−0.3
5	13 ± 2	7.9	7.9	0	7.9	0
6	15 ± 1	7.8	8.4	+0.6	8.5	+0.7
7	30 ± 5	7.5	7.7	+0.2	7.2	−0.3
8	30 ± 6	7.5	7.2	−0.3	7.6	+0.1
9	9.0 ± 3	8.0	8.0	0	8.3	+0.3
10	15 ± 2	7.8	8.2	+0.4	7.9	+0.1
11	4.0 ± 0.6	8.4	8.2	−0.2	8.2	−0.2
12	5.6 ± 0.4	8.3	8.4	+0.1	8.2	+0.1
13	21 ± 8	7.7	7.5	−0.2	7.6	−0.1
14	2.7 ± 0.4	8.6	8.7	+0.1	8.4	−0.2
15	7.3 ± 0.4	8.1	8.2	+0.1	7.4	−0.7
16	4.6 ± 2.7	8.3	7.8	−0.5	8.0	−0.3
17	2.6 ± 0.1	8.6	8.7	+0.1	8.5	−0.1
18	13 ± 2	7.9	7.8	−0.1	8.6	+0.7
19	4.1 ± 0.4	8.4	8.6	+0.2	8.3	−0.1
20	3.8 ± 0.3	8.4	8.5	+0.1	8.6	+0.2
21	2.4 ± 1.2	8.6	7.8	−0.8	8.4	−0.2
22	8.4 ± 3	8.1	7.3	−0.8	7.0	−1.1
23	70 ± 12	7.2	7.3	+0.1	7.0	−0.2
24	9.2 ± 2	8.0	7.4	−0.6	7.5	−0.5
25	5.5 ± 0.5	8.3	7.5	−0.8	8.1	−0.2
26	12 ± 3	7.9	Inactive		7.3	−0.6
27	68 ± 5	7.2	Omitted		6.9	−0.3
28	29 ± 3	7.5	8.0	+0.5	7.0	−0.5
29	118 ± 1	6.9	6.8	−0.1	6.9	0
30	36 ± 13	7.4	7.4	0	7.1	−0.3
31	34 ± 5	7.5	8.0	+0.5	7.0	−0.5
32	38 ± 6	7.4	7.7	+0.3	7.5	+0.1
33	190 ± 21	6.7	6.7	0	7.5	+0.8
34	360 ± 63	6.4	7.2	+0.6	6.9	+0.5
35	30 ± 1	7.5	7.4	−0.1	8.2	+0.7
36	28 ± 3	7.6	7.1	−0.5	7.7	+0.1
37	7.3 ± 3.6	8.1	7.1	−1.0	7.3	−0.8
38	88 ± 7	7.1	7.8	+0.7	6.8	−0.3
39	297 ± 79	6.6	6.5	−0.1	7.4	+0.8
40	4.8 ± 0.01	8.3	7.7	−0.6	7.9	−0.4
41	250 ± 90	6.6	7.0	+0.4	7.6	+1.0
42	86 ± 21	7.1	Inactive		7.2	+0.1
43	108 ± 43	7.0	7.4	+0.4	7.0	0
44	44 ± 15	7.4	8.1	+0.7	7.6	+0.2
45	595 ± 119	6.2	6.7	+0.5	6.7	+0.5
46	168 ± 58	6.8	6.6	−0.2	6.3	−0.5
47	40 ± 4	7.4	6.9	+0.7	7.2	−0.2
48	2590 ± 848	5.6	6.3	+0.7	6.4	+0.8
49	Inactive		Omitted		6.0	0
50	79 ± 18	7.1	7.7	+0.6	7.5	+0.4
51	1054 ± 220	6.0	Omitted		6.0	0
52	Inactive	6.0	6.2	+0.2	7.0	+1.0
53	237 ± 50	6.6	Omitted		7.2	+0.6
54	Inactive	6.0	8.5	+2.5	6.6	+0.6

Δ = difference between experimental and calculated pK_i.

demands and is correspondingly fast. One of its prime advantages is that GRIND descriptors are alignment independent, which significantly contrasts to 3D-QSAR analyses such as CoMFA. It should be underlined that the results from Palomer et al.[18] and our data obtained with the GRIND/ALMOND approach do not represent general models for $CysLT_1$ receptor antagonists, but refer to the subgroup of the quinolinyl(bridged)aryl type.

3.2. Predictivity testing

In the second part of our study we tested the predictivity of the above chemometric model with a test set of 69 $CysLT_1$ receptor antagonists, synthesized and first described by Zwaagstra;[19] for structures see Table 2. Molecular structures of the test set differ in a main aspect: the bridging ring in one subgroup is a phenylene

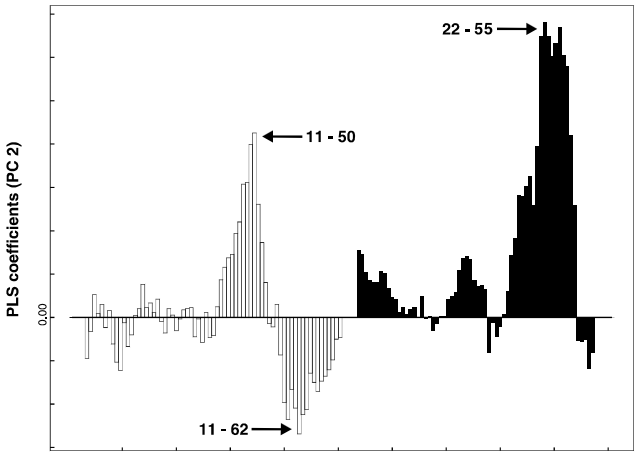


Figure 1. The PLS coefficients plot indicates the GRIND variables, which are directly (positive values) or inversely (negative values) correlated to the biological activity (*Y* variable), namely: 11-50: (DRY–DRY) interaction fields between benzene and quinoline, separated by 25 Å, 11-62: (DRY–DRY) interaction fields between quinoline and tetrazole, separated by 31 Å, and 22-55: (N1–N1) interaction fields between quinolinic nitrogen and tetrazole, separated by 27 Å. This plot also shows that variables 11-50 and 22-55 are responsible for high-affinity $CysLT_1$ receptor binding, whereas variable 11-62 is detrimental for binding. For the sake of a better graphical differentiation, white boxes correspond to GRIND variables derived from (DRY–DRY) interaction fields, whereas black boxes correspond to GRIND variables derived from (N1–N1) interaction fields.

moiety, whereas another subgroup is bridged by heterocycles like piperazine, homopiperazine, or piperidine. Thus, the former subgroup stronger resembles the structures of the training set than the latter. We intentionally included test set compounds with a varying chemical relation to the training set just to prove the general limits of predictivity. Expectedly, predictive power should correspond to the extent of training and test set similarity.

Affinities of the test set molecules were determined in binding studies on guinea pig lung membranes using [^{3}H]-LTD_4 as radioligand; they are listed in Table 4. Binding affinities range from a pK_i-value of 8.23 for compound 9087 to a pK_i of 3.06 for compound 8886 and cover a spectrum of more than 5 log units.

The plot of the experimental binding affinities versus the data calculated with our model is given in Figure 5. The test set molecules are labeled by open circles. From the plot one can easily derive that roughly 50% of the test set are quite well predicted, whereas the remaining structures behave as significant outliers. For the sake of comparison also the database molecules from Palomer et al.,[18] which served as our training set, are given; they are labeled by black circles.

According to the above expectations, the quality of prediction significantly corresponds with the chemical subclassification. Phenylene bridged compounds are quite well predicted; for the subgroup with bridging heterocycles predictions are rather poor. However, the coincidence between chemical subgrouping and the

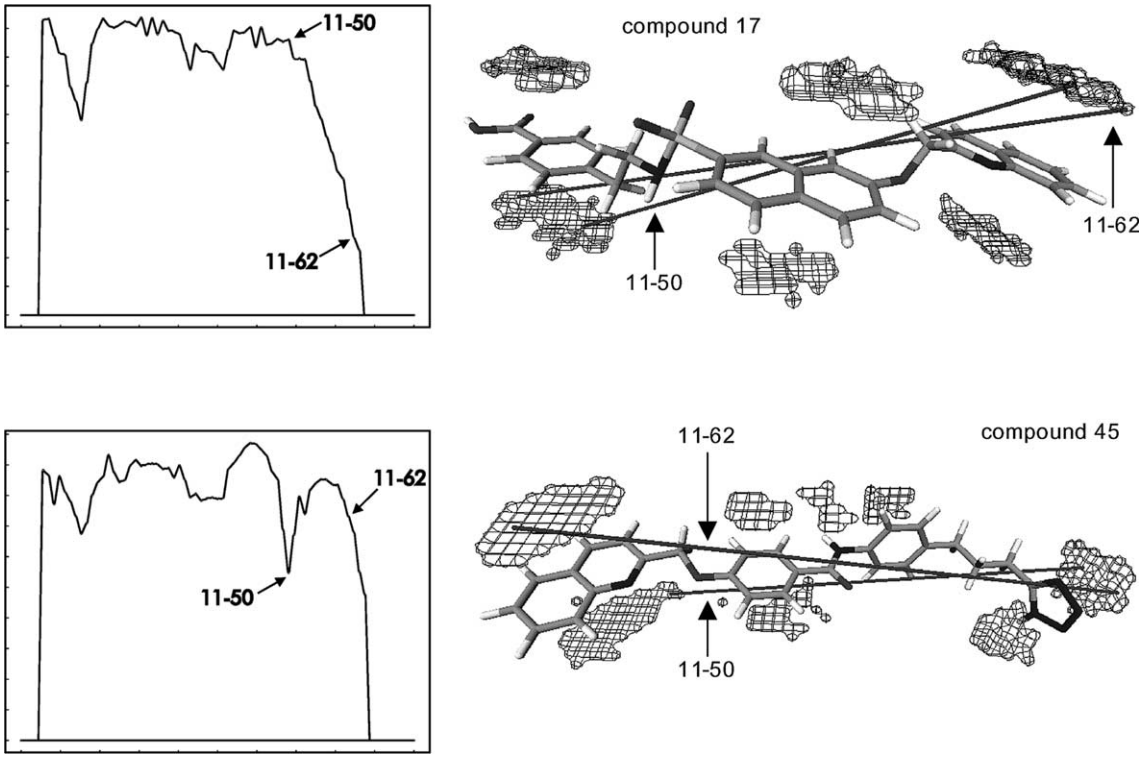


Figure 2. Graphical display of GRIND variables (DRY probe) with high impact on $CysLT_1$ receptor affinity. Favorable hydrophobic interaction fields (11-50, distance 25 Å), due to the quinoline and the benzene ring, as well as detrimental hydrophobic interaction fields (11-62, distance 31 Å), due to the quinoline and the tetrazole moiety, are shown. The correlograms in the left-hand part prove that the favorable interactions (11-50) are strong and the detrimental interactions (11-62) are weak for the highly potent compound **17**. The opposite holds for the weakly active compound **45**.

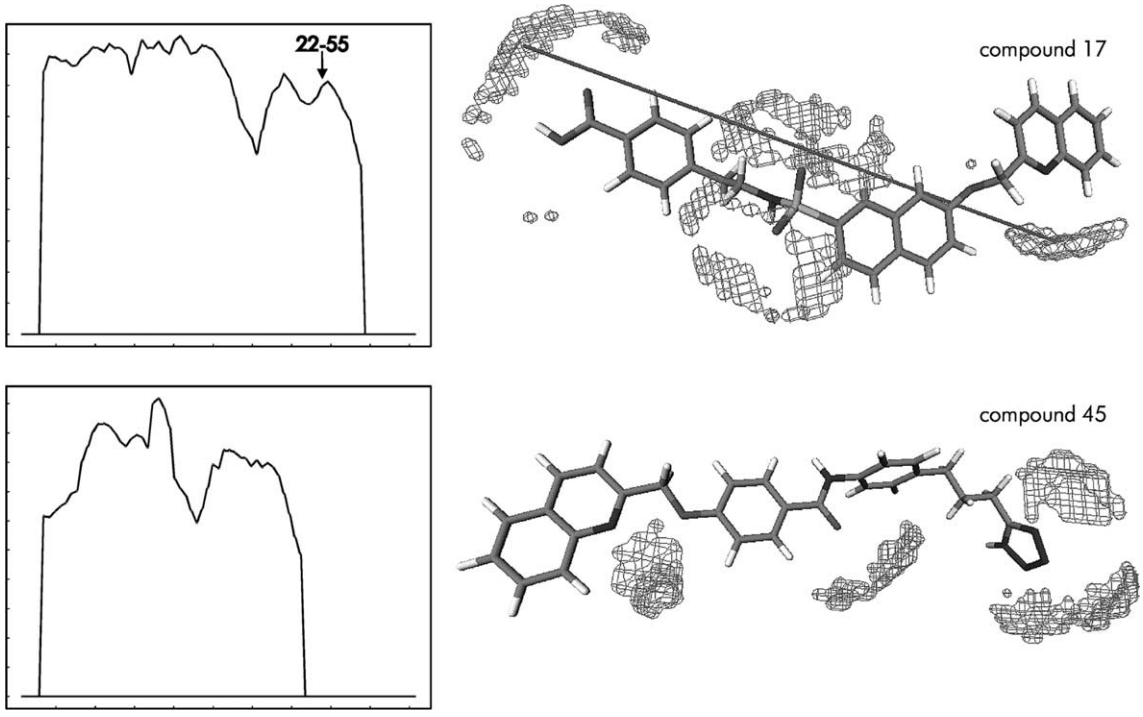


Figure 3. Graphical display of GRIND variables (N1 probe) with high impact on $CysLT_1$ receptor affinity. Variable 22-55 indicates favorable HB acceptor properties of the quinolinic nitrogen and the tetrazole (compound **45**), respectively, carboxyl moiety (compound **17**), separated by 27 Å. The correlograms (left-hand part) immediately indicate that this interaction is present for the highly potent compound **17**, whereas it is absent for the weakly active compound **45**.

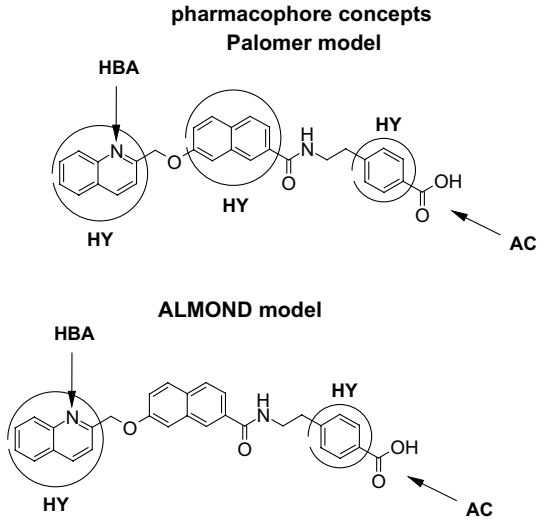


Figure 4. Comparison of the pharmacophoric features defined by Palomer et al.[18] with results of the GRIND/ALMOND approach, presented in this paper. The pharmacophore from the Palomer study includes (a) an acidic or negative ionizable function (AC), (b) three hydrophobic regions (HY), represented by the three aromatic ring systems, and (c) a hydrogen bond acceptor moiety (HBA), as given in the quinolinic nitrogen. Comparison with results of the GRIND/ALMOND approach indicates a high similarity between the two models. The only exception is that we detect two instead of three favorable hydrophobic regions, represented by the quinoline and the benzene moieties.

quality of prediction is not complete. Compounds 8991, 9013, 9015, 9016, 9396, 9467, 9470, 9472, 9639, and 9640 are phenylene bridged, yet poorly predicted. On the other hand, compound 9361 is bridged by a piperazine moiety, but well predicted. There exist chemical commonalities among these outliers, which might favor their deviating behavior. Compounds 9013, 9015, 9016, 9396, 9639, and 9640 exhibit an ethyl-oxy-ethyl substituent within the benzimidazole moiety; compounds 9467, 9470, and 9472 contain a central tertiary amine, which is lacking in the remaining structures. Presence of a tetrazole ring, which is absent in the training set structures, is presumably responsible for the poor prediction for the most potent test set compound 9087.

In order to better interpret the poor predictivity for nearly half of the test set compounds we performed a PLS discriminant analysis including the training set compounds as well as the strongest outliers of the test set. The scores plot of this discriminant PLS (Fig. 6) nicely shows an almost complete separation between the above mentioned two subsets, indicating some inherent chemical features present in the training set but missing in the outliers of the test set.

The PLS coefficients plot in Figure 7 indicates how the GRIND variables influence the clustering of the subset compounds. For example, the DRY coefficients 11-26 and 11-58 are important for the discrimination between training set compounds and outliers of the test set. These pharmacophoric features represent DRY–DRY interactions at distances of about 13 and 29 Å, which are missing in the outliers of the test set. On the other hand, the outliers set show DRY–DRY interactions at 18 Å distance (11-36), which are absent in the training set

Table 4. Experimental and calculated $CysLT_1$ receptor binding data for the test set

Compound	pK_i		Compound	pK_i	
	Exp.	Calcd		Exp.	Calcd
8754	6.87	6.60	9192	6.40	6.19
8787	5.37	7.34	9231	6.52	7.02
8788	5.03	7.31	9232	6.71	7.09
8831	5.23	6.97	9233	6.68	6.99
8832	4.42	6.94	9236	6.49	6.44
8835	4.80	7.37	9237	6.52	6.51
8852	4.30	7.99	9238	6.83	7.49
8886	3.06	6.93	9239	6.89	7.41
8900	5.06	7.62	9240	6.07	7.29
8901	5.79	7.26	9241	7.59	6.98
8963	5.49	6.97	9242	6.94	7.04
8982	5.48	6.74	9247	7.94	7.79
8991	5.96	7.34	9249	6.29	6.61
9013	5.26	7.13	9361	5.70	6.79
9015	5.20	7.81	9396	5.09	7.12
9016	5.19	7.79	9400	6.28	7.39
9024	6.75	6.81	9467	6.45	8.10
9056	6.94	6.71	9470	6.33	8.06
9057	5.70	7.10	9472	6.27	8.09
9059	5.13	6.97	9581	7.51	6.80
9061	3.96	6.99	9582	6.98	7.51
9083	7.87	7.05	9636	6.45	6.39
9085	6.49	6.09	9637	6.52	7.02
9086	6.84	6.72	9639	5.05	7.06
9087	8.23	6.67	9640	4.98	7.32
9088	6.69	7.24	9731	5.09	7.16
9094	5.17	6.79	9734	5.04	6.71
9095	5.13	7.92	9735	4.70	6.94
9097	5.06	7.08	9736	4.26	7.03
9098	4.74	6.42	9737	4.90	6.45
9144	5.52	7.32	9738	4.26	7.63
9145	4.45	7.41	9742	4.80	7.71
9187	6.39	6.28	9743	4.72	6.98
9189	7.42	7.17	9744	4.00	7.84
9191	5.70	6.42			

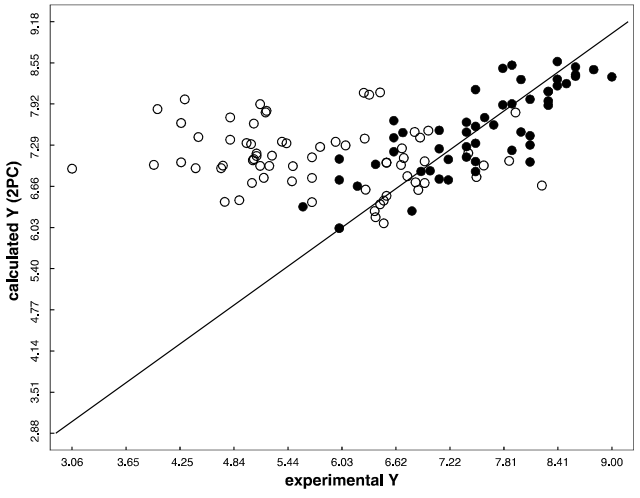


Figure 5. Plot of the experimental binding affinities (experimental *Y*) versus the data predicted with our model (calculated *Y*). The plot comprises both the training set molecules from Palomer et al.,[18] labeled by black circles and the test set molecules from Zwaagstra and co-workers,[17,19] labeled by open circles. The plot shows that rather 50% of the test set molecules are quite well predicted, whereas the remaining structures behave as significant outliers.

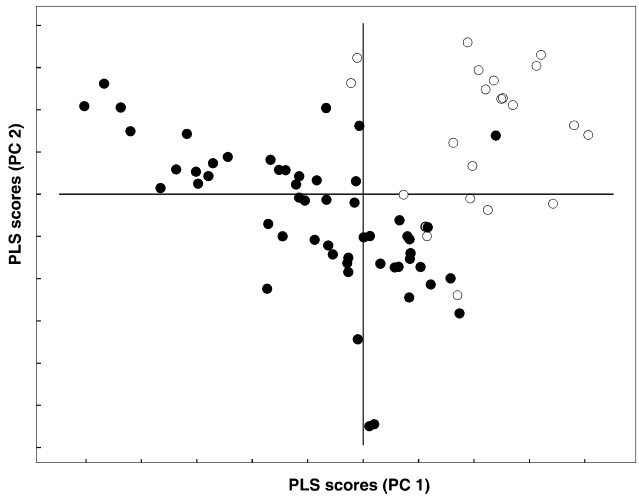


Figure 6. Scores plot of a PLS discriminant analysis including the training set compounds as well as the strongest outliers of the test set. This plot demonstrates an almost complete separation between the members of the two subsets indicating some inherent chemical features present in the training set but missing in the outliers of the test set.

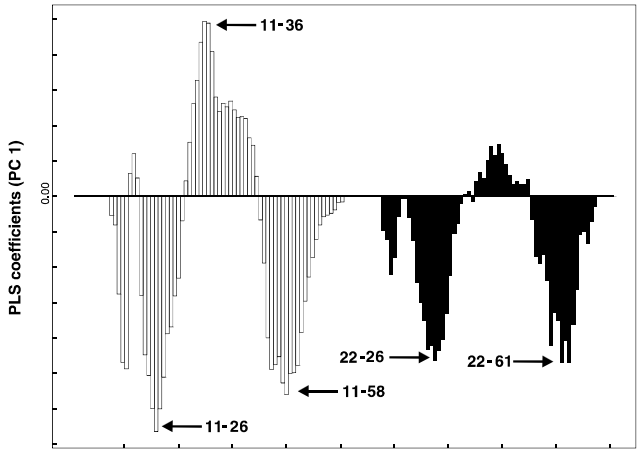


Figure 7. This PLS coefficients plot indicates which GRIND variables are important for the discrimination between the training and test set compounds. For example, the GRIND variables 11-26 and 11-58 (= DRY–DRY interactions at distances of 13 and 29 Å) are missing in the test set compounds. On the other hand, GRIND variable 11-36 (= DRY–DRY interaction at 18 Å distance) is absent in the training set compounds. Similar results are obtained for acceptor–acceptor patterns. The training set compounds show at least two acceptor regions at 30 Å distance (variable 22-61), which are lacking in the test set compounds, or are reported to be closer in the space. All these features explain the partial failure of the predictions. The negative peak of variable 22-26 represents HB acceptor interactions of the molecules at a distance of about 13 Å. Whereas this variable is almost generally present in the training set molecules, it only partially occurs in the test set molecules. This is due to involvement of the central molecular part in this interaction being more rigid in the case of the training set molecules and more flexible in the test set molecules resulting in the presence or absence of this variable depending on conformation. Thus, due to limited chemical interpretability we did not include it in the dataset discrimination.

compounds. Similar results are obtained for acceptor–acceptor patterns. The training set compounds show at least two acceptor regions at 30 Å distance (22-61). Those regions are not present in the test set compounds,

or are reported to be closer in the space. All these features explain the partial failure of the predictions.

4. Conclusion

The present study describes the development of valid and predictive QSAR models for a set of 54 $CysLT_1$ receptor antagonists of the quinolinyl(bridged)aryl type using the chemometric GRIND/ALMOND approach, which is alignment-independent, has simple computational demands and is correspondingly fast. Via the GRIND/ALMOND approach the entire training set could be modeled without any outlier behavior. Considering that conformational sampling can affect the quality of the model, we generated different conformational ensembles, but did not succeed in obtaining better models than presented here. PLS analysis resulted in a two-component model explaining 67% of the variance for $CysLT_1$ receptor binding. GRIND variables 11-50 and 22-55 are responsible for high $CysLT_1$ receptor affinity; variable 11-62 is detrimental.

External predictivity of the GRIND/ALMOND model was tested for a set of 69 $CysLT_1$ receptor antagonists with varying chemical similarity to the training set. Via inclusion of test set compounds with varying chemical similarity we intended to prove the general limits of external predictivity. The quality of prediction mainly coincides with chemical subclassification: phenylene bridged compounds are well predicted; for structures with bridging heterocycles predictions are rather poor. Thus, predictive power significantly corresponds to the extent of training and test set similarity.

References and notes

1. Kemp, J. P. *Arch. Int. Med.* **1993**, *153*, 805.
2. Barnes, P. J.; Chung, K. F.; Page, C. P. *Pharmacol. Rev.* **1988**, *40*, 49.
3. Barnes, P. J. *Brit. Med. J.* **1993**, *307*, 814.
4. Chung, K. F.; Barnes, P. J. In Barnes, P. J., Ed., *Asthma. Br. Med. Bull.* **1992**, *48*, 135–148.
5. Cockroft, D. W.; Murdock, K. Y. *J. Allergy Clin. Immunol.* **1987**, *79*, 734.
6. Barnes, P. J. *Ann. Rev. Med.* **1993**, *44*, 229.
7. Barnes, P. J. *Eur. Respir. J.* **1992**, *5*, 1126.
8. Gianaris, P. G.; Golish, J. A. *Drugs* **1993**, *46*, 1.
9. Coleman, R. A.; Eglen, R. M.; Jones, R. L.; Narumiya, S.; Shimizu, T.; Smith, W. L. *Adv. Prostaglan. Thrombox. Leukot. Res.* **1995**, *23*, 283.
10. Metters, K. M.; Frey, E. A.; Ford-Hutchinson, A. W. *Eur. J. Pharmacol.* **1991**, *194*, 51.
11. Frey, E. A.; Nicholson, D. W.; Metters, K. M. *Eur. J. Pharmacol.* **1993**, *244*, 239.
12. Norman, P.; Abram, T. S.; Cuthbert, N. J.; Tudhope, S. R.; Gardiner, P. J. *Eur. J. Pharmacol.* **1994**, *271*, 73.
13. Metters, K. M.; Gareau, Y.; Lord, A.; Rochette, C.; Sawyer, N. *J. Pharmacol. Exp. Ther.* **1994**, *270*, 399.
14. Metters, K. M.; Zamboni, R. J. *J. Biol. Chem.* **1993**, *268*, 6487.
15. Brooks, C. D. W.; Summers, J. B. *J. Med. Chem.* **1996**, *39*, 2629.
16. Hariprasad, V.; Kulkarni, V. M. *J. Comput.-Aid. Mol. Des.* **1996**, *10*, 284.
17. Zwaagstra, M. E.; Schoenmakers, S. H.; Nederkoorn, P. H.; Gelens, E.; Timmerman, H.; Zhang, M. Q. *J. Med. Chem.* **1998**, *41*, 1439.
18. Palomer, A.; Pascual, J.; Cabré, F.; Lluisa Garcia, M.; Mauleón, D. *J. Med. Chem.* **2000**, *43*, 392.
19. Zwaagstra, M. E. Molecular recognition of the leukotriene CysLT1 receptor. Ph.D. thesis, Faculty of Science, Free University of Amsterdam, 1997.
20. Pastor, M.; Cruciani, G.; McLay, I.; Pickett, S.; Clementi, S. *J. Med. Chem.* **2000**, *43*, 3233.
21. ALMOND version 3.0, Molecular Discovery Ltd, 4 Chandos Street, London W1A3BQ, 2002 (http://www.moldiscovery.com).
22. GRID version 20, Molecular Discovery Ltd, 4 Chandos Street, London W1A3BQ, 2002 (www.moldiscovery.com).
23. Wold, S.; Esbensen, K.; Geladi, P. *Chem. Intell. Lab. System.* **1987**, *2*, 37.
24. SYBYL 6.5 Tripos Inc., 1699 South Hanley Road, St. Louis, Missouri 63144, USA.