Journal of Computer-Aided Molecular Design, **17:** 1–11, 2003.
KLUWER/ESCOM

Enantioselectivity of epoxide hydrolase catalysed oxirane ring opening: a 3D QSAR Study

Joachim Paier, Thomas Stockner, Andreas Steinreiber, Kurt Faber & Walter M.F. Fabian*
Institut für Chemie (IfC), Karl-Franzens Universität Graz, Heinrichstr. 28, A-8010 Graz, Austria

Received 31 July 2002; accepted in final form 18 February 2003

Key words: Comparative Molecular Field Analysis (CoMFA), Comparative Molecular Similarity Indices Analysis (CoMSIA), epoxide hydrolase, *Rhodococcus ruber*, substrate model

Summary

A 3D QSAR analysis (quantitative structure activity relationships) of a set of 2,2-disubstituted epoxides, substrates for epoxide hydrolases originating from four different organisms, was conducted by CoMFA (comparative molecular field analysis) and CoMSIA (comparative molecular similarity indices analysis), with respect to the enantioselective ring opening to the corresponding *vicinal* diol. Structural variations of the substrates include alkyl chains of different lengths, unsaturated moieties ((*E*)- and (*Z*)-alkenyl, alkinyl, aryl) and electronegative groups (ether oxygens, halogen atoms) at different locations within the 2-substituent group. Generally, all four organisms, namely *Rhodococcus ruber* NCIMB 11216, *Rhodococcus ruber* DSM 43338, *Rhodococcus ruber* DSM 44540 and *Rhodococcus ruber* DSM 44539, preferentially react with the (*S*)-enantiomer of the epoxide. Enantioselectivities (enantiomeric ratio, ln E values) show a rather large variation, ranging from almost no ($\ln E < 1$) to nearly complete selectivity ($\ln E > 5.3$). In addition, the response of the epoxide hydrolases stemming from the four organisms towards structural modifications of the substrate is different. Models for the enantioselectivity (enantiomeric ratio, ln E values) obtained by CoMFA and CoMSIA are of different but reasonable predictive power, e.g., $q^2_{CV} = 0.701$ and $r^2 = 0.937$ for the CoMFA model of *Rhodococcus ruber* DSM 43338. Enantiomeric ratios for the test molecules can be well predicted. Plots of steric and electrostatic CoMFA (CoMSIA) fields allow conclusions to be drawn for the choice of the most suitable organism for a specific type of substrate.

Abbreviations

QSAR	=	Quantitative structure-activity relationship
CoMFA	=	Comparative molecular field analysis
CoMSIA	=	Comparative molecular similarity indices analysis
PLS	=	partial least squares
SAMPLS	=	Sample-distance partial least squares
LOO-CV	=	leave-one-out crossvalidation
SDEP	=	Standard error of prediction
SEE	=	Standard error of estimate
EH	=	Epoxide hydrolase

Introduction

Epoxide hydrolases catalyse the hydrolysis of epoxides to furnish the corresponding *vic*-diols [1]. Due to its inherent asymmetry, this reaction is of considerable importance for the synthesis of enantiopure compounds and has recently attracted much interest [2]. Despite of numerous (detoxification) studies on epoxide hydrolases from mammalian sources [3], the

*Corresponding author. Email: walter.fabian@uni-graz.at URL: http://www-ang.uni-graz.at/~fabian/

preparative-scale application of these catalysts became only feasible after several microbial sources (in particular among fungi [4], red yeasts [5] and bacteria [6]) were identified, which ensured their supply in sufficient amounts by fermentation. Bacterial epoxide hydrolases proved to be extremely flexible and exhibit exquisite stereoselectivities over a wide substrate pattern, particularly on 2,2-disubstituted oxiranes [7–15].

In order to broaden the applicability of these enzymes and to circumvent a tedious substrate-testing by trial-and-error using epoxide hydrolases from various bacterial strains for optimisation of stereoselectivities [16], we envisaged to develop a more general (substrate-) model [17], based on a set of data obtained with bacterial epoxide hydrolase preparations using a series of lipophilic 2,2-disubstituted oxiranes as substrates. Only recently X-ray structures of epoxide hydrolases of microbial origin have become available (Agrobacterium radiobacter, Aspergillus niger) and two of the enzymes active in the organisms being studied have been biochemically characterised. However, no sequence data are available to allow the construction of protein structures using homology modelling or some sort of structure based design. Therefore, as a first step towards the goal of a substrate model, a 3D QSAR (CoMFA [18], CoMSIA [19, 20]) investigation of substrate-selectivity relationships was done. For this purpose, four biocatalytically active organisms (*Rhodococcus ruber* NCIMB 11216, *Rhodococcus ruber* DSM 43338, *Rhodococcus ruber* DSM 44540 and *Rhodococcus ruber* DSM 44539) were chosen. Although several epoxide hydrolases from microbial sources have been purified to homogeneity [6], the majority of preparative-scale reactions have been performed using whole cells or crude cell-free extracts [6, 16]. A set of about twenty chiral 2,2-disubstituted oxiranes of different electrostatic and steric properties with known experimental enantiomeric ratios of the oxirane ring opening catalysed by the bacterial epoxide hydrolase was selected. Generally, the substrates were derived from a 2,2-disubstituted oxirane ring bearing both a small (methyl or ethyl) and a large group (n-alkyl chains of variable length (n = 4–8)). Further variations of this basic structures consisted in the introduction of a multiple carbon-carbon bond (E- or Z-configured alkene, alkyne and the sterically demanding aryl moiety) at different distances from the chiral centre. In addition, some of the carbon atoms were replaced by electronegative (oxygen, bromine) atoms (Chart 1).

Computational details

Generation of the molecular structures

Starting structures were generated with the molecular modelling program SYBYL [21] and preoptimised by the TRIPOS force field [22]. Geometry optimisation and calculation of the partial atomic charges were performed using MOPAC 6.0 [23] (AM1-Hamiltonian [24], keywords EF [25] and PRECISE). Initially all side chains were generated in the lowest energy stretched, i. e. *all-trans* conformation. Various other structures resulting from rotation around the C_2–C_5 bond (for atom numbering see Figure 1), described by the dihedral angle $\delta(O_1–C_2–C_5–C_6)$, were used to derive QSAR models. Additionally, in compounds containing an aryl ring, the orientation of the aromatic ring was also varied to check the corresponding influence on the resulting statistical model. For each conformation produced in this way, MOPAC partial charges were recalculated. Depending on the respective organism, different orientations of the side chain, i.e., different values of $\delta(O_1–C_2–C_5–C_6)$, led to the best respective model.

Alignment

The alignment is a critical step in a CoMFA/CoMSIA-study. As no crystallographic data were available for any enzyme-substrate complex of the investigated organisms, the crystal structure of a murine soluble epoxide hydrolase-inhibitor complex (PDB: 1EK1) served as a model for the substrate orientation. In analogy to the inhibitor molecule present in this crystal structure, all substrates were generated in a stretched, i.e. *all-trans* conformation. Since murine sEH is only very distantly related to that of *R. ruber*, this approach necessarily constitutes only a first guess. As detailed above, different orientations and folds of the side chain were tested and only those resulting in the best statistical models (largest q^2 values) were further analysed. Fortunately, most important for statistically significant QSAR models is not the best fitting to the correct active site geometry but rather self-consistency of the alignment [26]. Manual alignment was accomplished by using the epoxide ring and the carbon atom C_4 attached to C_2 as the template structure. Since manual fitting easily can lead to overfitted models, the field fit procedure [27] as implemented in SYBYL, was applied as an alternative approach. The field fit procedure minimises the root mean squares (RMS) difference in the sum of steric and/or electrostatic interaction energies, averaged across all grid points, between each considered molecule and a template

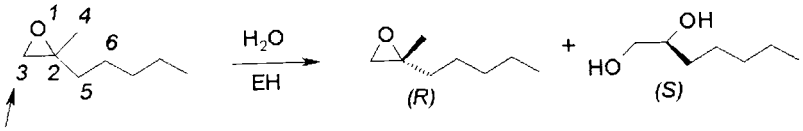

Figure 1. Ring opening reactions of oxiranes catalysed by epoxide hydrolases.

molecule thereby altering the position of the former. In this study, both template fields (steric/electrostatic; not weighted) were used. The internal geometry is kept invariant during this procedure. The (*S*) enantiomer of compound **13** was selected as template molecule because this compound is cleaved by all used organisms with rather high enantioselectivity.

CoMFA field calculations

The preferentially hydrolysed enantiomers of the substrates are those with (*S*)-configuration at position 2 of the oxirane ring. Steric and electrostatic CoMFA fields were calculated for the favoured (*S*) enantiomer and used for the correlation (see section Regression Analyses). The fields were calculated with the TRIPOS force field [22], using a C_{sp3} atom bearing a charge of $+1$ within a box of $28 \times 28 \times 28$ Å and a grid spacing of 2 Å. The influence of using a grid spacing of 1 Å was checked for all organisms. The energy truncation values for both interactions were kept with 30 kcal/mol (default value). The smoothing function of the CoMFA fields implemented in SYBYL was applied.

CoMSIA field calculations

The steric, electrostatic and the combined fields have been checked. The attenuation factor was set to 0.3. Box dimensions and grid spacing are given in Table 3. A considerable influence on the crossvalidated q^2 by alternation of the box dimensions was realised; the best results of the q^2 values were obtained by application of automatically created boxes.

Experimental data

The experimental E-values were taken from the literature [6]. It has to be emphasised, that the measured selectivity values are obtained from the reaction of epoxides with whole organisms or crude cell extracts rather than isolated enzymes (see section Introduction). In the correlation, ln E values were used as the target variable, because of the possibility to interpret E as an equilibrium constant [28]. As a consequence, ln E mathematically stands for the difference of the reaction rate of enantiomers and, thus, can be regarded as a difference of binding enthalpy between the enantiomers. The measurements of the upper selectivity-frame's margin is experimentally difficult. Extremely selective ($E \approx 200$) hydrolases do not allow for an accurate determination of the corresponding E-value. An E-value of 200 means almost 100% selectivity and represents the upper limit of the experimentally measurable E value. Obviously then, the CoMFA- and CoMSIA-models, derived with inclusion of these data have to be interpreted with some care, especially with respect to possible outliers. Despite the ambiguities introduced into the model thereby, we decided not to omit compounds showing very high ln E values. It ought to be pointed out that the enantiomeric ratios could not be measured for all of the compounds shown in Chart 1 for each of the four investigated organisms. Oxiranes **11**, **12** and **20** are missing in all but the *Rhodococcus ruber* NCIMB 11216's set of compounds. Besides these three compounds, oxiranes **4** and **22** are missing in the set of *Rhodococcus ruber* DSM 43338 and oxirane **21** is missing in the set of *Rhodococcus ruber* DSM 44539. No additional oxiranes are absent in the set of *Rhodococcus ruber* DSM 44540. Finally, the data set for NCIMB does not contain compounds **7** and **21**.

Regression analyses

As is usual with CoMFA/CoMSIA analyses, the partial least squares method PLS [29–31] was used. To obtain the optimum number of components to be extracted in the PLS method, leave-one-out crossvalidation (LOO-CV) was applied. For this, the SAMPLS method [32], implemented in SYBYL was used. The LOO-CV correlation coefficient (q^2) and the standard error of prediction (SDEP) were considered. Besides, the conventional crossvalidation procedure, with groups of five and seven, respectively, was checked. The number of the models' components was derived by choosing those with maximum values of q^2 and minimum values of SDEP. To avoid complex models, including a lot of noise, in some cases less than the suggested optimal number of components was

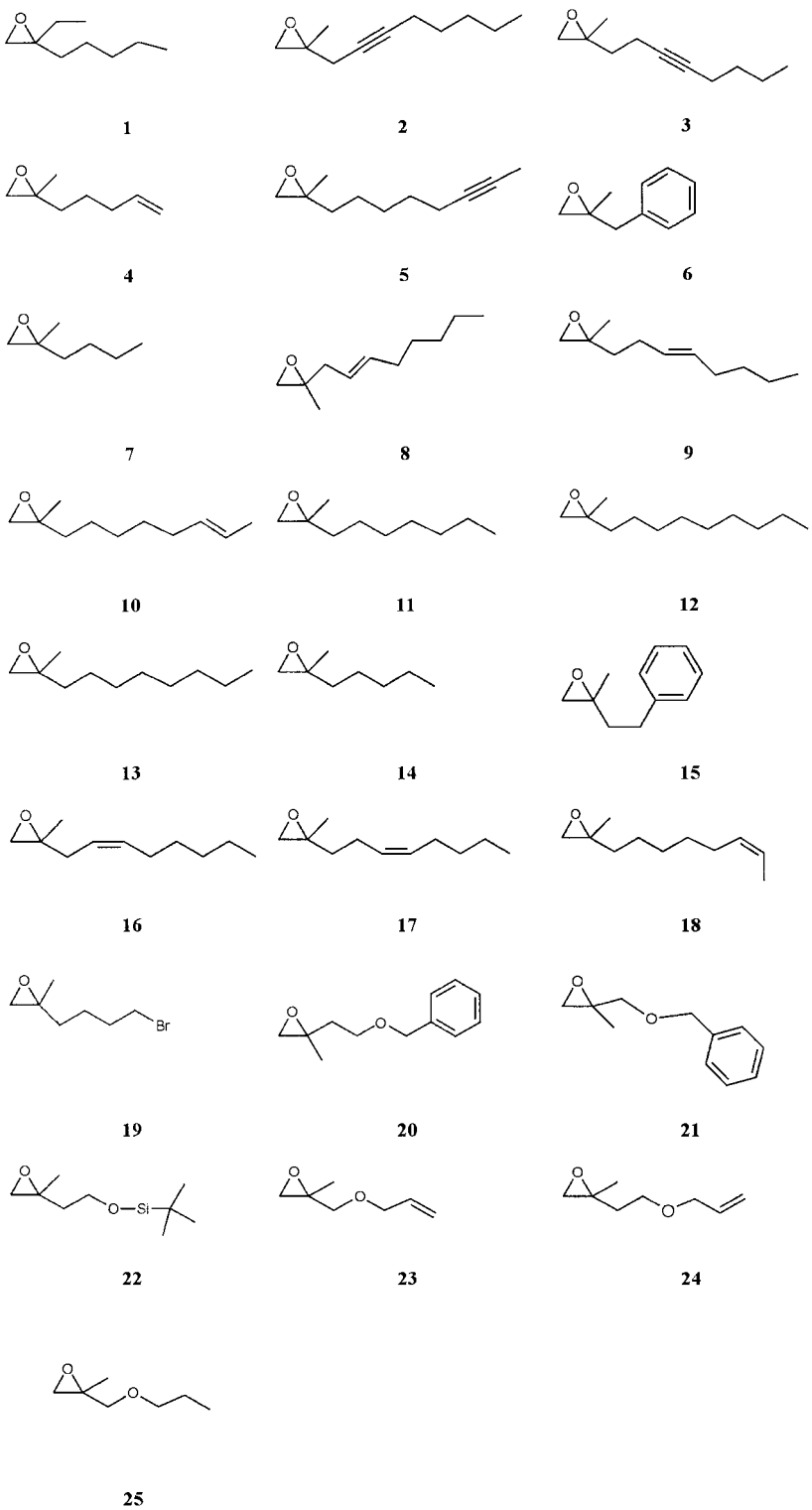

Chart 1. Chemical formulas of the epoxides treated in the QSAR-studies.

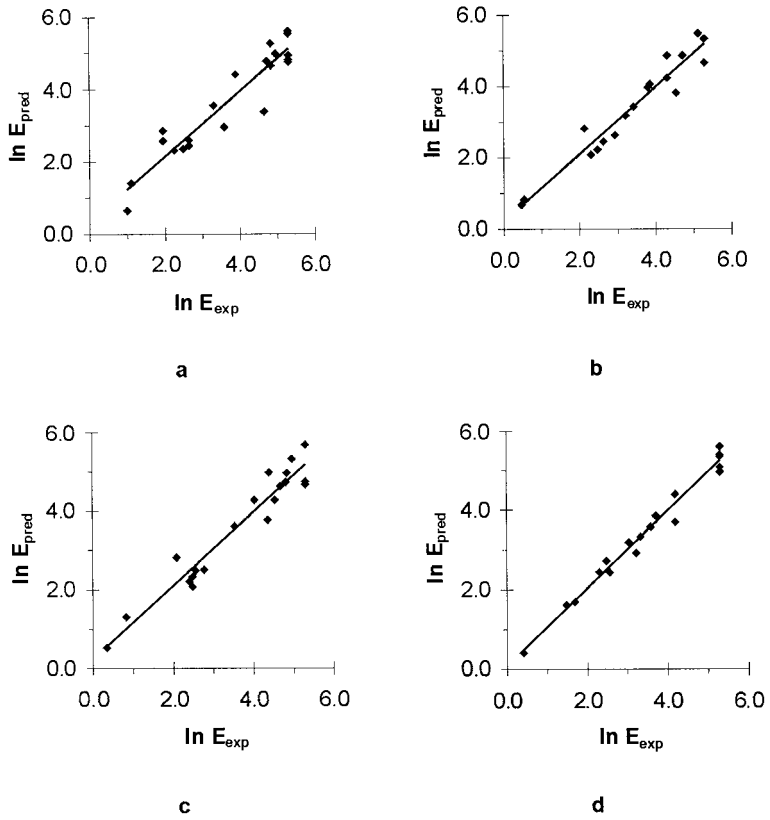

Figure 2. Predicted *vs* actual ln E values (CoMFA models with manual alignment). **a**: *R. ruber* NCIMB 11216; **b**: *R. ruber* DSM 43338; **c**: *R. ruber* DSM 44540; **d**: *R. ruber* DSM 44539.

taken for the final model. Leave-one out q^2 values are about 5% higher than the q^2 gained by groups of five or seven. This is in line with the general finding that LOO frequently is too optimistic [33].

Results and discussion

Comparative molecular field analysis – Comparative molecular similarity indices analyses

In the following, first, results with respect to different alignment procedures for the CoMFA method will be presented. Second, the two 3D QSAR methods (CoMFA and CoMSIA) will be compared concerning the respective predictive quality of the statistical models derived thereby. Third, the similarities and/or differences of the 3D QSAR models between the four investigated epoxide hydrolases will be discussed individually. As a last point, conclusions which can be derived from the CoMFA (CoMSIA) fields with respect to the choice of the proper substrate for a given organism and/or the choice of the most appropriate epoxide hydrolase for a given substrate, will be presented.

Results of the statistical analysis are given in Table 1 (manual alignment) and Table 2 (field fit alignment). As can be seen from the data presented in Table 1 and Table 2, both manual alignment and the field fit procedure lead to similar statistical results. However, ln E values for the test molecules **24** and **25** predicted by the model derived from field fit alignment are in significantly better agreement with the experimental data than those obtained by manual alignment (see below).

The analogous results obtained with the aid of the CoMSIA procedure are collected in Table 3. Generally, both crossvalidated q^2 as well as conventional correlation coefficients r^2 are comparable to those provided by CoMFA (compare Table 1 and 3), although the CoMSIA of *Rhodococcus ruber* NCIMB 11216 has led to an additional component.

A comparison between experimental ln E values and those obtained by the CoMFA/CoMSIA models

Table 1. Numerical results concerning the CoMFA-study[(1)]

	R. ruber NCIMB 11216	*R. ruber* DSM 43338	*R. ruber* DSM 44540	*R. ruber* DSM 44539
q^2	0.478	0.701	0.511	0.664
SDEP	1.180	0.885	1.150	0.952
r^2	0.913	0.937	0.932	0.979
SEE	0.481	0.405	0.427	0.237
F-test	59.2 (3,17)	69.8 (3,14)	72.6 (3,16)	219.1 (3,14)
No. of comp.	3	3	3	3
fraction				
steric	0.555	0.438	0.441	0.602
electrostatic	0.445	0.562	0.559	0.398

[(1)] A box size 28 × 28 × 28 Å was used throughout; manual alignment, grid spacing: 2 Å, probe atom Csp^3

Table 2. Numerical results concerning the CoMFA-study [(1)]

	R. ruber NCIMB 11216	*R. ruber* DSM 43338	*R. ruber* DSM 44540	*R. ruber* DSM 44539
q^2	0.453	0.565	0.541	0.717
SDEP	1.204	1.068	1.114	0.874
r^2	0.954	0.964	0.952	0.969
SEE	0.351	0.308	0.362	0.287
F-test	116.7 (3,17)	124.1 (3,14)	104.9 (3,16)	148.3 (3,14)
No. of comp.	3	3	3	3
fraction				
steric	0.504	0.488	0.424	0.501
electrostatic	0.496	0.512	0.576	0.499

[(1)] A box size 28 × 28 × 28 Å was used throughout; field fit alignment, grid spacing: 2 Å, probe atom Csp^3

Table 3. Numerical results concerning the CoMSIA-study [(1)]

	R. ruber NCIMB 11216	*R. ruber* DSM 43338	*R. ruber* DSM 44540	*R. ruber* DSM 44539
q^2	0.454	0.703	0.326	0.723
SDEP	1.204	0.882	1.350	0.864
r^2	0.833	0.894	0.884	0.957
SEE	0.665	0.526	0.559	0.341
F-test	28.3 (3,17)	39.5 (3,14)	40.8 (3,16)	103.4 (3,14)
No. of comp.	3	3	3	3
fraction				
steric	0.247	0.329	0.326	0.337
electrostatic	0.753	0.671	0.674	0.663

[(1)] Box sizes of 24 × 15 × 14 Å (*R. ruber* NCIMB 11216), 21 × 13 × 13 Å (*R. ruber* DSM 43338, 22 × 15 × 14 Å (*R. ruber* DNM 44540; *R. ruber* DNM 44539) were used. Grid spacing: 2 Å, probe atom Csp^3

Table 4. Observed and calculated ln E values of enzyme *R. ruber* NCIMB 11216

Compound-No.	actual	CoMFA[(a)] calc.	CoMFA[(b)] calc.	CoMSIA calc.
1	1.95	2.86	1.69	2.85
2	2.48	2.35	2.66	1.99
3	0.99	0.65	0.44	0.69
4	1.95	2.57	2.57	3.33
5	5.30	5.61	5.17	5.59
6	4.71	4.78	4.81	4.78
8	3.89	4.41	4.15	4.57
9	3.30	3.56	3.35	3.64
10	4.83	5.27	5.09	4.69
11	4.83	4.65	4.54	4.04
12	5.30	5.53	5.35	5.53
13	5.30	4.81	5.02	4.76
14	4.65	3.38	4.13	3.46
15	2.25	2.31	2.54	2.27
16	2.64	2.59	2.50	2.61
17	2.64	2.45	3.27	3.39
18	4.96	4.98	4.89	4.90
19	5.30	4.75	5.03	4.49
20	1.10	1.41	1.30	1.10
21	5.30	4.93	5.40	4.90
23	3.58	2.96	3.34	3.68
24	1.64	2.32	1.65	3.87
25	2.30	3.61	1.85	3.39

[(a)] manual alignment [(b)] field fit alignment

Table 5. Observed and calculated ln E values of enzyme *R. ruber* DSM 43338

Compound-No.	actual	CoMFA[(a)] calc.	CoMFA[(b)] calc.	CoMSIA calc.
1	4.73	4.85	5.15	4.77
2	0.53	0.83	0.46	0.39
3	0.47	0.68	0.41	0.88
5	5.14	5.47	5.12	5.17
6	3.43	3.42	3.37	3.20
7	4.32	4.85	4.29	5.28
8	2.64	2.45	2.90	2.79
9	2.13	2.81	2.54	2.50
10	4.32	4.22	4.36	4.27
13	4.55	3.81	3.84	3.53
14	5.30	4.65	4.82	4.42
15	2.48	2.22	2.52	2.81
16	2.30	2.08	2.39	2.18
17	2.94	2.63	2.68	2.75
18	3.87	4.05	4.10	3.81
19	5.30	5.32	5.33	5.14
21	3.81	3.96	3.98	3.66
23	3.22	3.17	3.23	3.94
24	1.06	4.20	2.82	1.63
25	2.29	3.26	1.19	4.76

[(a)] manual alignment [(b)] field fit alignment

are given in Tables 4–7. In Figure 2 plots of $\ln E_{pred}$ vs $\ln E_{exp}$ (CoMFA model) are shown for the four organisms.

As can be seen especially with organism *Rhodococcus ruber* NCIMB 11216, the quality of the measured E-values has a significant influence on the respective CoMFA/CoMSIA models (Figure 2). Substrates with a measured E-value of >200 result in broadly deviating predictions. With *Rhodococcus ruber* NCIMB 11216, nearly a quarter of the substrates show E-values of about 200. Consequently the models of this organism are of lower quality. According to the CoMFA model, both fields have similar contributions. With CoMSIA the fractional contribution of the electrostatic field is clearly higher (∼1:2 ratio) than that given by CoMFA. On the other hand, this study shows that both 3D QSAR methods lead to similar results with respect to both predictive ability as well as the relative contributions of steric and electrostatic effects. Oxirane ring cleavage by *Rhodococcus ruber* DSM 43338 and *Rhodococcus ruber* DSM 44540, which both have less substrates showing almost complete enantioselectivities than *Rhodococcus ruber* NCIMB 11216, can be described by considerably better statistical models (Table 1 and 3). Interestingly, however, *Rhodococcus ruber* DSM 44539 – containing the same number of possibly problematic experimental E-values as *Rhodococcus ruber* NCIMB 11216 – gives an excellent statistical model. Obviously, then, the QSAR model derived for this particular epoxide hydrolase is less sensitive towards our high selectivity approximation than that originating from *Rhodococcus ruber* NCIMB 11216. To test the predictive power of the statistical models described above, predictions for the ln E values of two additional molecules (**24** and **25**, for structures, see Chart I) are performed. Predicted ln E values are also listed in Tables 4–7. Our results show that prediction of the ln E values for the test molecules **24** and **25** by the CoMFA model derived from field fit clearly is superior compared to manual alignment or the CoMSIA procedure. In Figure 3 and 4, the electrostatic and steric contour plots of the CoMSIA analyses, which are claimed to be easier to interpret [19], are shown. At first glance, the contour maps of the 3D QSAR models for the enantioselectiv-

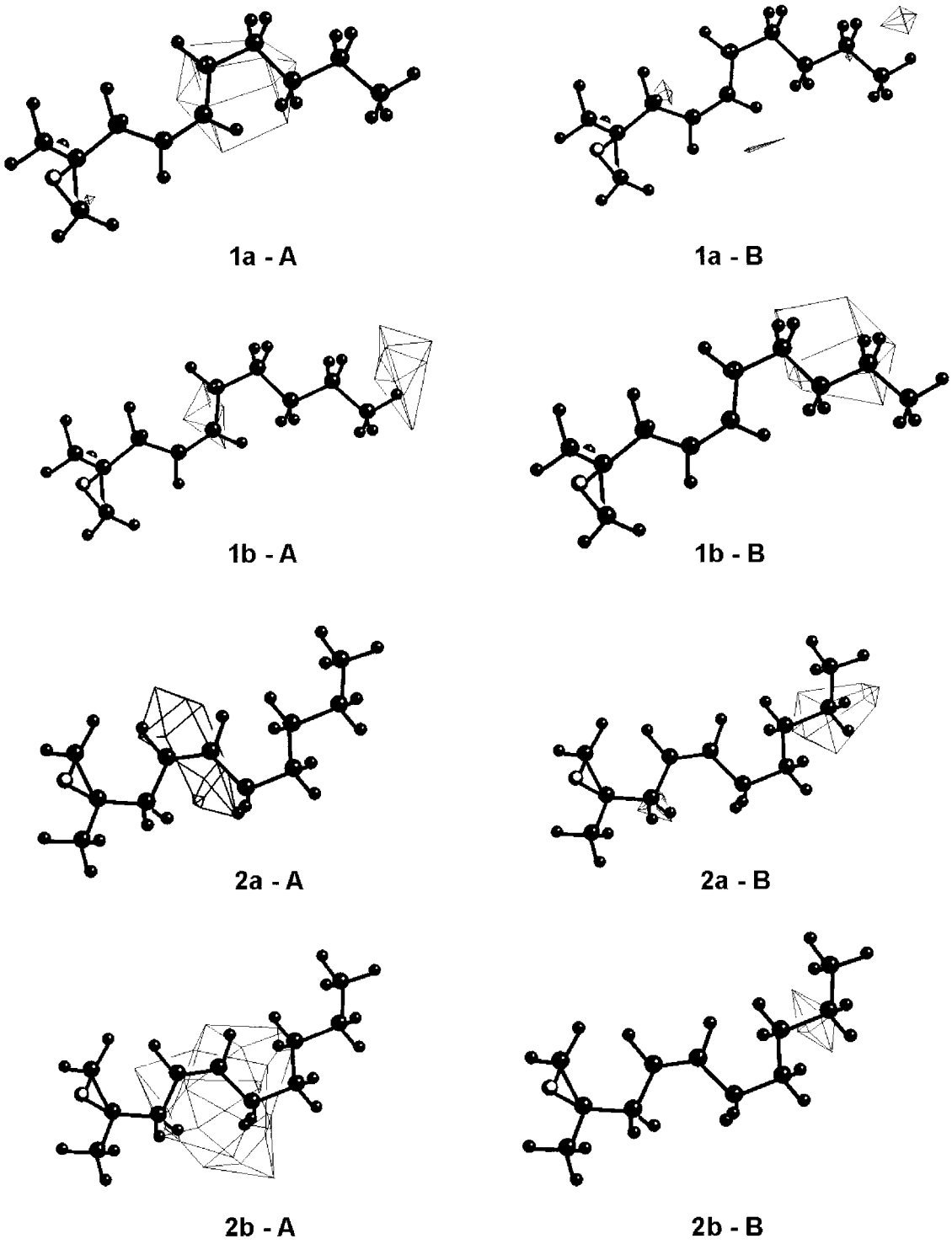

Figure 3. Electrostatic **(a)** and steric **(b)** CoMSIA contour maps (standard deviation × coefficients contoured by contribution) of substrates for the organisms ***R. ruber*** **NCIMB 11216 (1)** and ***R. ruber*** **DSM 44540 (2)**. **1a-A** and **2a-A** indicate regions where positive charge and **1b-A** and **2b-A** those where steric bulk enhances selectivity. **1a-B** and **2a-B** indicate regions where positive charge and **1b-B** and **2b-B** those where steric bulk decreases selectivity.

Table 6. Observed and calculated ln E values of enzyme *R. ruber* DSM 44540

Compound-No.	actual	CoMFA[(a)] calc.	CoMFA[(b)] calc.	CoMSIA calc.
1	4.84	4.96	5.25	4.69
2	0.83	1.31	0.57	0.76
3	0.34	0.53	0.37	0.60
4	5.30	4.75	4.60	4.54
5	4.97	5.33	5.16	5.97
6	4.81	4.73	4.71	4.25
7	4.39	4.98	4.50	5.05
8	2.77	2.51	2.91	2.80
9	2.08	2.82	2.37	2.42
10	4.54	4.28	4.34	4.42
13	4.36	3.77	3.73	3.73
14	5.30	4.67	4.75	4.52
15	2.48	2.34	2.57	2.93
16	2.48	2.09	2.42	2.40
17	2.56	2.49	2.55	2.57
18	4.03	4.28	4.14	4.25
19	5.30	5.69	5.44	5.21
21	4.68	4.63	4.81	4.60
22	3.53	3.61	4.02	2.96
23	2.40	2.22	2.80	3.32
24	1.13	4.59	3.24	0.12
25	2.16	3.55	1.58	4.03

[(a)] manual alignment [(b)] field fit alignment

Table 7. Observed and calculated ln E values of *R. ruber* DSM 44539

Compound-No.	actual	CoMFA[(a)] calc.	CoMFA[(b)] calc.	CoMSIA calc.
1	5.30	5.40	5.45	5.10
2	1.48	1.62	1.42	1.78
3	0.41	0.40	0.55	0.73
4	5.30	4.97	5.16	5.37
5	4.19	4.40	4.19	4.51
6	3.58	3.58	3.45	3.35
7	5.30	5.36	5.27	5.68
8	2.30	2.44	2.52	2.72
9	2.48	2.72	2.57	2.47
10	4.19	3.70	3.65	3.80
13	3.04	3.19	3.03	3.07
14	5.30	5.09	4.98	4.89
15	2.56	2.43	2.59	2.27
16	3.22	2.92	2.80	2.82
17	1.69	1.69	1.90	1.88
18	3.71	3.86	3.65	3.52
19	5.30	5.61	5.78	5.72
22	3.33	3.33	3.74	3.04
24	1.16	3.99	0.86	3.83

[(a)] manual alignment [(b)] field fit alignment

ity of oxirane ring opening are more or less different for the four organisms under investigation. However, closer inspection reveals some similarities between *Rhodococcus ruber* NCIMB 11216 and *Rhodococcus ruber* DSM 44540 on the one hand and *Rhodococcus ruber* DSM 43338 and *Rhodococcus ruber* DSM 44539 on the other hand, especially with respect to the electrostatic fields.

A common feature of all four epoxide hydrolases is a predicted enhancement of the enantiomeric ratio by increasing the negative charge at the end of the side chain. This is particularly important for the epoxide hydrolases originating from *Rhodococcus ruber* DSM 43338 and *Rhodococcus ruber* DSM 44539 (see Figure 4). An increase of positive charge (or decrease of negative one) at locations 2–4 atoms away from the oxirane ring should lead to an enhanced enantiomeric ratio in the case of *Rhodococcus ruber* NCIMB 11216, *Rhodococcus ruber* DSM 44540 and also *Rhodococcus ruber* DSM 43338. In contrast, for *Rhodococcus ruber* DSM 44539, almost no such effect is indicated. As the most obvious difference with respect to electrostatic fields between *Rhodococcus ruber* NCIMB 11216 and *Rhodococcus ruber* DSM 44540 on the one hand, and *Rhodococcus ruber* DSM 43338 and *Rhodococcus ruber* DSM 44539 on the other hand, is the increase of ln E of the epoxide hydrolases from the latter two sources by an increase of negative charge in the vicinity of the epoxide moiety. With respect to steric effects, the epoxide hydrolase from *Rhodococcus ruber* NCIMB 11216 apparently is the most demanding: only rather small regions at the end of the side chain allow for an increase of the steric bulk. According to the model derived for the EH from *Rhodococcus ruber* DSM 44540 enhanced enantiomeric ratios towards substrates with increased steric bulk in regions extending from the oxirane ring up to 4–5 atoms along the side chain can be expected. Both *Rhodococcus ruber* DSM 43338 as well as *Rhodococcus ruber* DSM 44539 should tolerate or even react with increased ln E values with sterically demanding substituents at distant regions of the side chain.

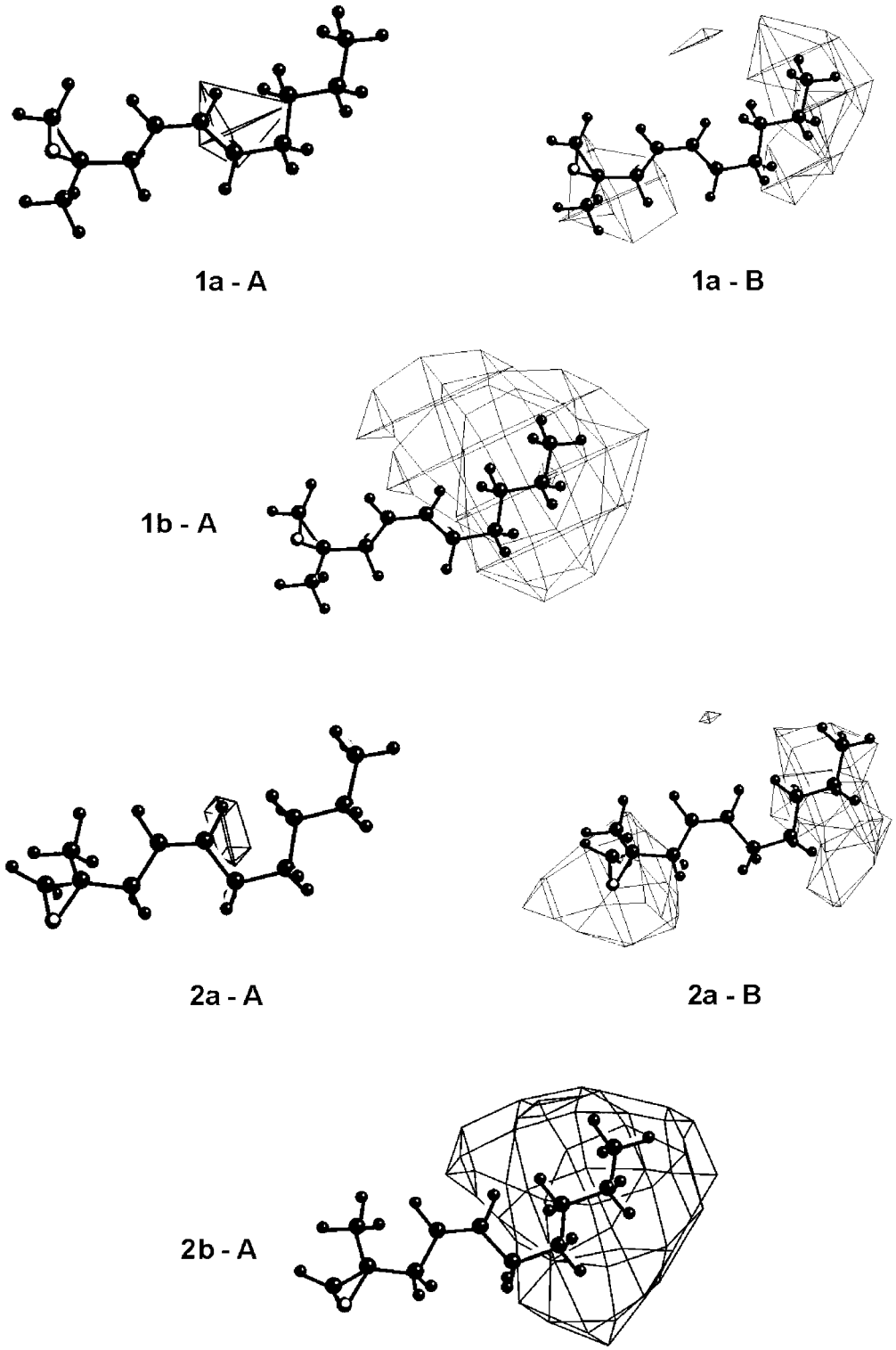

Figure 4. Electrostatic **(a)** and steric **(b)** CoMSIA contour maps (standard deviation × coefficients contoured by contribution) of substrates for the organisms ***R. ruber*** **DSM 43338 (1)** and ***R. ruber*** **DSM 44539 (2)**. **1a-A** and **2a-A** indicate regions where positive charge and **1b-A** and **2b-A** those where steric bulk enhances selectivity. **1a-B** and **2a-B** indicate regions where positive charge decreases selectivity.

Conclusions

We have presented 3D QSAR models obtained by the CoMFA as well as the CoMSIA procedure for the enantioselective oxirane ring opening catalysed by four different epoxide hydrolase containing bacteria, namely *Rhodococcus ruber* NCIMB 11216, *Rhodococcus ruber* DSM 43338, *Rhodococcus ruber* DSM 44540 and *Rhodococcus ruber* DSM 44539. Except for *Rhodococcus ruber* NCIMB 11216, where rather low crossvalidated q^2 and conventional r^2 values were obtained, statistically meaningful models could be derived. Thus, prediction of enantiomeric ratios ($\ln E$ values) with sufficient accuracy should be possible. Moreover, and even more important, an interpretation of the respective steric and electrostatic field makes it possible to draw conclusions concerning the most appropriate epoxide hydrolase for a given substrate.

References

1. Thomas, H., Oesch, F., ISI Atlas Sci.: Biochem., (1988) 287.
2. (a) For the asymmetric chemo-hydrolysis: Ready, J. M., Jacobsen, E. N., Angew. Chem., Int. Ed. Engl., 41 (2002) 1347. Schaus, S. E., Brandes, B. D., Larrow, J. F., Tokunaga, M., Hansen, K. B., Gould, A. E., Furrow, M. E., Jacobsen, E. N., J. Am. Chem. Soc., 124 (2002) 1307.
   (b) for the biohydrolysis see: Faber, K., Orru, R. V. A., in: Enzyme Catalysis in Organic Synthesis, K. Drauz, H. Waldmann (Eds.), Wiley-VCH, Weinheim, 2^{nd} edn., vol. II, 2002, pp. 579-608.
3. Oesch, F., Xenobiotica, 3 (1972) 305.
4. Archelas, A., J. Mol. Catal. B., 5 (1998) 79.
5. Weijers, C. A. G. M., Bont, de, J. A. M., J. Mol. Catal. B., 6 (1999) 199.
6. Steinreiber, A., Faber, K., Curr. Opinion Biotechnol., 12 (2001) 552.
7. Wandel, U., Mischitz, M., Kroutil, W., Faber, K., J. Chem. Soc., Perkin Trans. 1, (1995) 735.
8. Mischitz, M., Kroutil, W., Wandel, U., Faber, K., Tetrahedron: Asymmetry, 6 (1995) 1261.
9. Kroutil, W., Osprian, I., Mischitz, M., Faber, K., Synthesis, (1997) 156.
10. Osprian, I., Kroutil, W., Mischitz, M., Faber, K., Tetrahedron: Asymmetry, 8 (1997) 65.
11. Orru, R. V. A., Mayer, S. F., Kroutil, W., Faber, K., Tetrahedron, 54 (1998) 859.
12. W. Krenn, I. Osprian, W. Kroutil, G. Braunegg, K. Faber, Biotechnol. Lett., 21 (1999) 687.
13. Steinreiber, A., Osprian, I., Mayer, S. F., Orru, R. V. A., Faber, K., Eur. J. Org. Chem., (2000) 3703.
14. Steinreiber, A., Hellström, H., Mayer, S. F., Orru, R. V. A., Faber, K., Synlett, (2001) 111.
15. Hellström, H., Steinreiber, A., Mayer, S. F., Faber, K., Biotechnol. Lett., 23 (2001) 169.
16. Osprian, I., Stampfer, W., Faber, K., J. Chem. Soc., Perkin Trans. 1, (2000) 3779.
17. Hult, K., Faber, K.; Horvath, I. T. (Ed.), Encyclopedia of Catalysis, Wiley, New York, 2002, in press.
18. Cramer III, R. D., Patterson, D. E., Bunce, J. D., J. Am. Chem. Soc., 110 (1988) 5959.
19. Klebe, G., Abraham, U., Mietzner, T., J. Med. Chem., 37 (1994) 4130.
20. Kubinyi, H., Folkers, G., Martin, Y. C., (Eds.), 3D QSAR in Drug Design, Vols. 2 and 3, Kluwer, Dordrecht, 1998.
21. SYBYL, Molecular Modeling Software, Tripos Inc., 1699 S. Hanley Road, St. Louis, MO 63944, USA.
22. Clark, M., Cramer III, R. D., Opdenbosch, van, N., J. Comp. Chem., 10 (1989) 982.
23. Stewart, J. J. P., J. Comput.-Aided Mol. Design, 4 (1990) 1.
24. Dewar, M. J. S., Zoebisch, E. G., Healy, E. F., Stewart, J. J. P., J. Am. Chem. Soc., 107 (1985) 3902.
25. Baker, J., J. Comp. Chem., 7 (1986) 385.
26. Cramer III, R. D., DePriest, S. A., Patterson, D. E., Hecht, P., in Kubinyi, H. (Ed), 3D QSAR in Drug Design. Theory Methods and Applications, ESCOM, Leiden, 1993, 443.
27. Clark, M., Cramer III, R. D., Jones, D. M., Patterson, D. E., Simeroth, P. E., Tetrahedron Comput. Methodol., 3 (1990) 47.
28. Faber, K., Griengl H., Hönig, H., Zuegg, J., Biocatalysis, 9 (1994) 227.
29. Lindberg, W., Persson, J.-A., Wold, S., Anal. Chem., 55 (1983) 643.
30. Wold, S., Sjöström, M., Eriksson, L., Chemom. Int. Lab. Syst., 58 (2001) 109.
31. Cramer III, R. D., Bunce, J. D., Patterson, D. E., Frank, I. E., Quant. Struct.-Act. Relat. Pharmacol., 7 (1988) 18.
32. Bush, B. L., Nachbar, R. B., Jr., J. Comput.-Aided Mol. Design, 7 (1993) 587.
33. Kubinyi, H., Schleyer, P. v. R. (Eds.), Encyclopedia of Computational Chemistry, Wiley, Chichester, New York, 1998, 448.