Med Chem Res
DOI 10.1007/s00044-014-1121-y

MEDICINAL CHEMISTRY RESEARCH

ORIGINAL RESEARCH

Quantitative structure activity relationship and binding investigation of *N-alkyl glycine amides* as inhibitors of Leukotriene A4 hydrolase

Paula B. Paz · Esteban G. Vega-Hissi · Matias F. Andrada · Mario R. Estrada · Juan C. Garro Martinez

Received: 11 March 2014 / Accepted: 19 June 2014

Abstract The leukotriene A4 hydrolase (LTA4H) is a bifunctional zinc enzyme that catalyzes the final (rate-limiting) step in the synthesis of leukotriene B4 (LTB4), which is involved in several diseases. Many pharmaceutical attempts to exploit the LTA4H/LTB4 pathway have been unsatisfactory, hence, the development of new inhibitory drugs is essential. This paper describes the generation of a quantitative structure–activity relationship (QSAR) model on a series of *50 N-alkyl glycine amides* with experimentally defined IC_{50}. In addition, the optimized molecular structures of the inhibitors were docked into the active site of the enzyme to identify the enzyme-ligand interactions and quantify the estimated free energy of binding (ΔG_{bind}). A simple four-descriptor QSAR model with high predictive capacity was obtained. The statistic parameters of the model are: regression coefficient (R_{test}) of 0.714 and a standard deviation (S_{test}) of 0.696. The predicted inhibitory activity of 85 new *N-alkyl glycine amides* compounds was obtained with this QSAR model and these compounds were docked into LTA4H. Ten of the compounds present predicted IC_{50} values lower than 10 nM and binding poses and affinity values similar to the natural ligand (leukotriene A4), turning them into suitable candidates for experimental assays.

Keywords Leukotriene A4 hydrolase · N-alkyl glycine amides · Inhibitory activity · Molecular docking · QSAR

Introduction

The leukotriene A4 hydrolase enzyme (LTA4H) is a bifunctional zinc metalloenzyme with epoxide hydrolase and aminopeptidase activities that catalyzes the rate-limiting step in the production of leukotriene B4 (LTB4). Its structure consists of three domains: a catalytic domain, an N-terminal, and C-terminal domains (Thunnissen *et al*., 2001). Crystallographic studies (Wunder *et al*., 2010; Thunnissen *et al*., 2002; Sandanayaka *et al*., 2010) reveal that the active site contains a Zn^{2+} cation coordinated by the amino acids His295, His299, and Glu318. Other studies indicate the zinc atom and residues Glu271, Arg563, Asp375, and Tyr378 are involved in the epoxide hydrolase mechanism. In addition, these studies showed that above amino acids and also Gln136, Lys565, and Tyr267 could interact with LTA4H inhibitors (Thunnissen *et al*., 2001; Rinaldo-Matthis and Haeggström, 2010; Rudberg *et al*., 2002).

The rate-limiting step of LTB4 biosynthesis is the stereospecific hydrolysis of an unstable epoxide, Leukotriene A4 (LTA4) (Haeggström, 2004). The LTB4 is a potent pro-inflammatory activator related to inflammatory conditions such as asthma, inflammatory bowel disease (IBD), chronic obstructive pulmonary disease (COPD), arthritis, psoriasis, and atherosclerosis (Goodarzi *et al*., 2003; Barnes, 2001; Back *et al*., 2005). It was also recently reported that increased production of LTB4 is associated with the increased risk for

Electronic supplementary material The online version of this article (doi:10.1007/s00044-014-1121-y) contains supplementary material, which is available to authorized users.

P. B. Paz · E. G. Vega-Hissi · M. F. Andrada · M. R. Estrada · J. C. Garro Martinez
Facultad de Química, Bioquímica y Farmacia, Universidad Nacional de San Luis, San luis, Argentina

E. G. Vega-Hissi · J. C. Garro Martinez (✉)
Centro Científico Tecnológico San Luis (CCT-CONICET), Facultad de Química, Bioquímica y Farmacia, Universidad Nacional de San Luis, Chacabuco 917, San Luis 5700, Argentina
e-mail: jcgarro@unsl.edu.ar

Published online: 17 July 2014

myocardial infarction, stroke (Sandanayaka *et al*., 2010), and cancer (Chen *et al*., 2004). The relevance of the LTA4H in human diseases has proved to be substantial, but pharmaceutical attempts to exploit 5-lipoxygenase pathway have been disappointing (Shim and Paige, 2012).

In the last years, an enough variety of compounds that inhibit the biosynthesis of LTB4 were presented (Kirkland *et al*., 2008; Çaliskan and Banoglu, 2013; Tanis *et al*., 2012; Ye *et al*., 2008; Penning, 2001). However, the research and the development of novel and more potent drugs are essential for the treatment and prevention of inflammatory diseases states. Several molecular modeling techniques including structure–activity relationship (SAR), quantitative structure–activity relationship (QSAR), molecular docking, and molecular similitude (or combined approaches) are used as potent tools for rational drug design. Recently, two QSAR and molecular docking studies on LTA4H inhibitors have been reported for the purpose of finding a relationship between the activity and their structures (Thangapandian *et al*., 2013; Thangapandian *et al*., 2011). In one of those studies, a Bayesian model was developed using a training set containing 26 compounds and nine molecular descriptors (Thangapandian *et al*., 2011). In other one, the same authors have employed an identical training set and the genetic function approximation (GFA) technique to develop a six-descriptor QSAR model.

According to the current interest in this topic, the main aim of this paper was to develop a new QSAR model in order to predict the activity of novel potential LTA4H inhibitors. For this purpose, the QSAR model was established using a series of 50 inhibitors of the *N-alkyl glycine amide* family that have not been used previously in any QSAR study. The multiple linear regression (MLR) variable selection approach, considered one of the most popular statistical techniques, was employed to develop the model, exploring more than a thousand theoretical molecular descriptors. In addition, we performed a molecular docking analysis to find the binding poses and to get insights into the interactions between the inhibitors and LTA4H.

Four previously synthesized compounds (Kirkland *et al*., 2008) and 85 new designed *N-alkyl glycine amides*, referred to as the tested compounds from now on, were docked into the active site of LTA4H to ensure their binding affinity and were evaluated as potential LT4H inhibitors with the developed QSAR.

Materials and methods

Biological experimental data

The 50 *N-alkyl glycine amides* and its associated experimental activity were extracted from the studies of Kirkland *et al*. (2008). In the cited studies the inhibitory activities were determined under identical experimental conditions leading to a regular data distribution. The IC_{50} (concentration of a compound required to inhibit 50 % of the hydrolase activity of LTA4H) is the biological activity data used in this study. The IC_{50} values, exhibiting a range of activity from 27 to 2,000 nM, were converted to the corresponding $\log_{10}IC_{50}$ and used as the dependent variable in all QSAR investigations. The inhibitory activities and the molecular structures of the 50 *N-alkyl glycine amides* are presented in Table 1 and Fig. 1S of supplementary material, respectively.

Computational details

The molecular structure of all compounds (50 *N-alkyl glycine amides* and the 89 tested compounds) was optimized at the semiempirical PM3 (parametric method-3) level of theory using the Polak-Ribiere algorithm and a gradient norm limit of 0.01 kcal $Å^{-1}$ with Hyperchem 7.0 package. Molecular Docking calculations were performed using Autodock Vina (Trott and Olson, 2010). The most stable docked conformation of each of the 50 *N-alkyl glycine amides* was used to calculate 1,497 molecular descriptors with Dragon software. The multiple lineal regression (MLR) calculations were carried out with Matlab 7.0.

Molecular docking approach

The molecular docking study was performed keeping the amino acid side chains rigid. The crystal structure of LTA4H was obtained from Protein Data Bank (Bernstein *et al*., 1977) (PDB accession code: 3CHO (Kirkland *et al*., 2008)). The co-crystallized ligand (2-amino-N-[4-(phenylmethoxy)phenyl]-acetamide) and all water molecules were removed from the crystal structure. The grid box was set to include completely the previously proposed active site (Paz *et al*., 2012) and standard docking parameters were used except the exhaustiveness for which a value of 100 was set. The free energy of binding (ΔG_{bind}) was estimated from the best docking results, *i.e.,* the conformation with the lowest energy.

Calculation of the molecular descriptors

The 1,497 calculated molecular descriptors include all types of descriptors such as Constitutional, Topological, Geometrical, Charge, GETAWAY (Geometry, Topology, and Atoms-Weighted AssemblY), WHIM (Weighted Holistic Invariant Molecular descriptors), 3D-MoRSE (3D-Molecular Representation of Structure based on Electron diffraction), Molecular Walk Counts, BCUT descriptors,

Table 1 Structure, experimental IC_{50}, predicted $\log_{10}IC_{50}$, and ΔG_{bind} (kcal mol^{-1}) for the 50 *N-alkyl glycine amides* derivatives

1Compounds	IC_{50} nM	$\log_{10}IC_{50}$ Exp	$\log_{10}IC_{50}$ Eq. 1	ΔG_{bind} kcal mol^{-1}
1	135	2.130	2.241	−7.8
2	350	2.544	2.750	−10.6
3[a]	1,500	3.176	2.829	−9
4	230	2.362	2.125	−8.4
5	430	2.634	2.353	−10.7
6	120	2.079	2.135	−10.7
7	160	2.204	2.188	−10.9
8	290	2.462	2.309	−10.7
9	340	2.532	2.402	−11
10	360	2.556	2.700	−10.9
11	440	2.644	2.549	−10
12	80	1.903	2.196	−10.6
13	250	2.398	2.142	−11.6
14	230	2.362	2.381	−10.8
15	70	1.845	2.180	−10.9
16	120	2.079	2.113	−11
17[a]	90	1.954	2.066	−10.9
18	210	2.322	2.277	−10
19[b]	60	1.778	2.092	−10.4
20	420	2.623	2.639	−10.5
21	70	1.845	1.706	−10.9
22[a]	140	2.146	2.197	−10.9
23	180	2.255	2.115	−11.4
24	240	2.380	2.524	−10
25	270	2.431	2.532	−10.4
26	260	2.415	2.603	−10.2
27	380	2.580	2.343	−11.4
28	220	2.342	1.803	−11.8
29[a,b]	60	1.778	0.524	−11.1
30[b]	30	1.477	1.808	−11
31[b]	60	1.778	2.028	−11.9
32	180	2.255	2.224	−10.7
33[b]	48	1.681	1.773	−9.3
34[a,b]	27	1.431	1.273	−9.8
35	85	1.929	2.084	−8.9
36	130	2.114	2.130	−9.8
37[a]	380	2.580	1.929	−8.9
38	110	2.041	2.078	−9.5
39[a]	180	2.255	2.261	−9.7
40	570	2.756	2.429	−9.6
41[a]	110	2.041	2.154	−9.8
42[a]	92	1.964	2.444	−0.2
43	72	1.857	2.171	−9
44	110	2.041	1.973	−8.9
45	440	2.644	2.405	−8.8
46[a]	2,000	3.301	2.561	−9.8
47[a]	1,000	3.000	2.330	−10.7
48[a]	1,000	3.000	2.986	−10.5
49	1,000	3.000	2.777	−10.4
50	1,000	3.000	3.003	−11

[a] Compounds of test set,

[b] Lead compounds for the design of novel inhibitors

2D-Autocorrelations, Aromaticity Indices, Randic Molecular Profiles, Radial Distribution Functions, Functional Groups, Atom-Centered Fragments, Empirical, and Properties (Todeschini and Consonni, 2009) among many others. The dipole moment (μ), E_{HOMO}, E_{LUMO} and GAP (obtained from semiempirical PM3 calculations), and ΔG_{bind} (computed from molecular docking study) were also added to the total set of molecular descriptors.

Development and validation of the QSAR model

The series of 50 *N-alkyl glycine amides* was partitioned in two different sets: the training set (38 compounds) and the test set (12 compounds). The elements of each set were carefully selected such that: (a) the compounds share similar structural features and (b) the experimental data of the test set are sufficiently representative of the whole span (Table 1).

A full search of optimal variables is impractical because it requires $D!/[d!(D-d)!]$ linear regressions. Where D is the total set of descriptor and d represents the number of selected descriptor. Therefore, an alternative method is necessary. We employed the replacement method (RM) as the molecular descriptor selection approach (Mercader *et al.*, 2010). This method is an efficient tool that produces MLR-QSAR models that are quite close to the full search methods but with lower computational cost. This technique approaches the minimum of S taking into account the relative errors of the coefficients of the least-squares model given by a set of d descriptors $d = \{X1,X2,\ldots,Xd\}$. In addition, the RM provides models with better statistical parameters than those from the forward stepwise regression procedure and similar to those from the more elaborated genetic algorithm approach (Duchowicz *et al.*, 2005).

The model was validated through four different approaches: (a) the leave-one-out (loo) and (b) the leave-more-out (l-%-o) cross-validation procedures, generating a million cases of random data removal for l-%-o, where the % is ≈20 (12 compounds); (c) a rigorous and more realistic validation procedure that involves the use of a set of molecules (test set) which do not form part of the training set; and (d) 10,000 cases of y-randomization (Wold and Eriksson, 1995), which consists in the interchange of the

experimental property such that the property value and the compound do not match.

Results and discussion

Molecular docking analysis

The 50 *N-alkyl glycine amides* were docked into LTA4H (Fig. 1 right panel). The spatial arrangement of the compounds within the enzyme demonstrates that they bind into the active site (compare to the position of the natural ligand in the left panel of Fig. 1). Also, the superposition of the inhibitors' molecular structures reveals that there is a large overlap between their structures.

Among the multiple weak interactions that take place between the inhibitors and LTA4H, four interacting regions can be distinguished: (a) an aromatic ring region, comprising the residues Tyr267, Tyr378, Tyr383, and Phe314 through π–π interactions; (b) the region of Pro266 whose backbone carbonyl group acts as hydrogen bond acceptor; (c) the Zn^{2+} cation zone with a coordinate interaction; and (d) the entry of the active site where Arg563 and Lys565 side chains allow the formation of hydrogen bonds.

The calculated free energies of binding (ΔG_{bind}) of the 50 *N-alkyl glycine amides* indicate that these compounds could present high binding affinity with the enzyme, Table 1, and some of them higher than the natural ligand itself (-9.5 kcal mol^{-1}) (Paz *et al*., 2012). It is well known that docking is not indicative of biological activity, however, these types of studies can be very useful to understand how ligands bind to the enzyme.

Table 2 Statistic parameters and molecular descriptors for the best 1–5 descriptor models

Number of descriptors	Molecular Descriptors	Calibration		Validation	
		R	S	R	S
1	H-051	0.332	0.343	0.539	0.597
2	ATS1m, ATS8m	0.498	0.320	0.739	0.464
3	MW, ATS8m, RDF080u	0.691	0.271	0.658	0.610
4	Sv, SIC0, ATS8m, RDF080u	0.814	0.221	0.714	0.697
5	AAC, Eig1p, ATS8m, ATS3p, RDF080u	0.857	0.199	0.676	0.849

Quantitative structure–activity relationship study

According to the number of compounds of the training set (N = 38), linear regression models containing from one to five descriptors (from a total of 1,497) would provide sufficient information about the relationship between the inhibitory activity and the structure of the compounds. The classic semiempirical "rule of thumb" indicates that at least six or seven data points (i.e., compounds) should be present by descriptor (Hansch, 1990). The search for these models was performed using the RM selection approach, which is able to select the most relevant and representative molecular descriptors for the training set. The five obtained models and the details of the selected molecular descriptors are summarized in Tables 2 and 3, respectively.

As can be seen in Table 2, the calibration statistical parameters did not improve significantly when five descriptors were utilized. In addition, the statistical

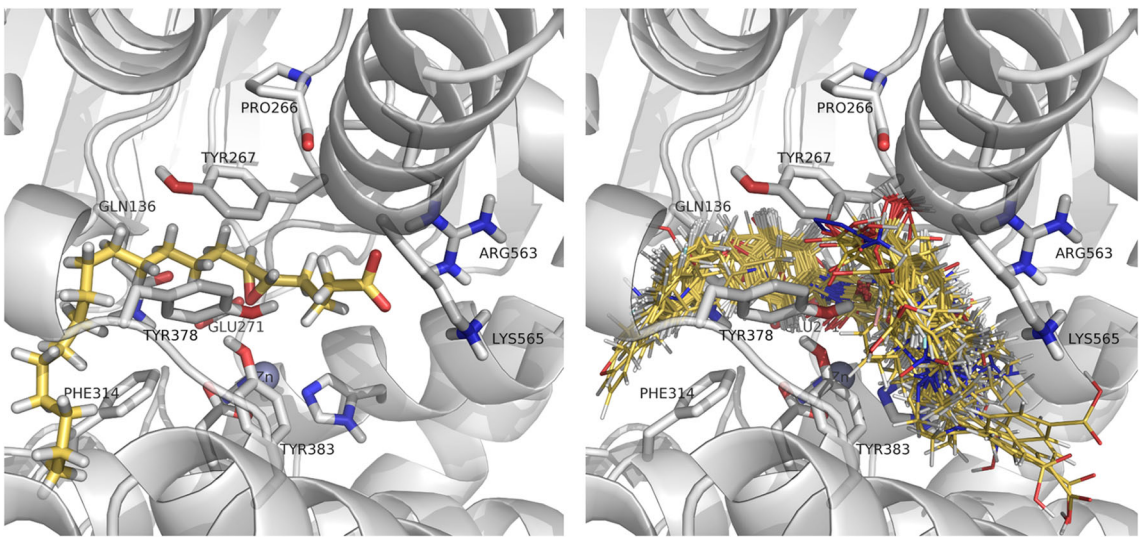

Fig. 1 Representation of the molecular structure of LTA4H (*gray* cartoon) active site with its natural ligand(*yellow sticks*, *left* panel) and 50 N-alkyl glycine amide inhibitors (*yellow sticks*, *right* panel). The amino acid side chains represented in *gray sticks* are involved in the interaction with inhibitors (Color figure online)

Table 3 Brief description of molecular descriptors used in the five models

Molecular Descriptors	Type of descriptor	Brief description
MW	Constitutional	Molecular weight
Sv	Constitutional	Sum of atom van der Waals volumes (scaled on Carbon atom)
AAC	Topologic	Mean information indexon atomic composition.
SIC0	Topologic	Structural information content (neighborhood symmetry of order-0)
Eig1p	Topologic	Leading eigenvalue from polarizability weighted distance matrix
ATS1m	2D autocorrelation	Broto-Moreau autocorrelation of topologic structure – lag 1/weighted by atomic masses
ATS8m	2D autocorrelation	Broto-Moreau autocorrelation of topologic structure – lag 8/weighted by atomic masses
ATS3p	2D autocorrelation	Broto-Moreau autocorrelation of topologic structure – lag 3/weighted by atomic polarizabilities
RDF080u	Radial distribution Function	Radial distribution Function – 8.0/unweighted
H-051	Atom-centered fragment	H attached to alfa-C

parameters of the test set only improved when d = 4. This suggests that the optimal number of descriptors is four, and that the model with five descriptors is possibly overfitted. The following equation and parameters belong to the four-descriptor QSAR model:

$$log_{10}IC_{50} = -18.041(2.84) + 0.569(0.07)\boldsymbol{Sv} + 37.731(6.14)\boldsymbol{SIC}0 - 0.242(0.03)\boldsymbol{ATS}8\boldsymbol{m} - 0.054(0.008)\boldsymbol{RDF}080\boldsymbol{u} \quad (1)$$

$$N = 38,\ S_{train} = 0.220,\ R_{train} = 0.814,$$
$$R^2 = 0.662,\ F = 16.237,\ p < 10^{-4}$$
$$R_{test} = 0.714,\ R^2_{test} = 0.510,$$
$$S_{test} = 0.696,\ S_{loo} = 0.257,\ R_{loo} = 0.740,$$
$$R_{l\%o} = 0.600\ S_{l\%o} = 0.351 S^{rand} = 0.300$$

where N is the number of molecules of the training set, R is the coefficient of correlation, S stands for the standard deviation, F is the Fisher parameter, train, and test subindex are applied to the training and test set, respectively, subindex loo and l-%-o stand for the leave-one-out and leave-%-out cross-validation techniques, respectively, and rand superindex stands for y-randomization.

Table 1 lists the predicted $\log_{10}IC_{50}$ values obtained from the four-descriptor model, Eq. 1. Figure 2 shows the plot of the predicted values as a function of the experimental values, for both the training and test sets. The data are grouped along the straight line of perfect fit which indicates that the established QSAR is an acceptable model. Despite compound number 29 (included in the test set) does not present a relevant structural feature which may differentiate it from the remainder of the set, the model did not predict properly its biological activity. The statistical parameters of the model improve slightly ($R_{test} = 0.766$ and $S_{test} = 0.551$) when this compound was considered outlier.

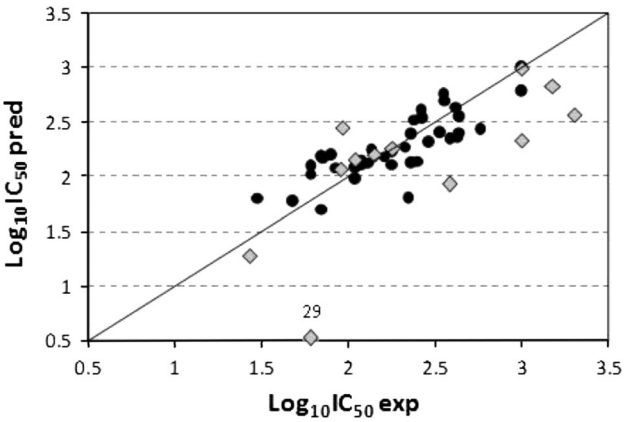

Fig. 2 Log–Log plot of the experimental and predicted biological activity according to equation 1. The training set is represented as *black filled circles* and the test set in *gray diamonds*. The *straight line* indicates the perfect fit

The validation was carried out through four different methods: leave-one-out, leave-%-out, employing a test set and y-randomization. As the name suggests, leave-one-out technique involves using a single data from the total set as the validation data, and the remaining data as the training set. This is repeated such that each data in the sample is used as validation. In leave-%-out, a percentage data is removed from the full set as validation data and the remaining data becomes the training set.

The regression coefficients of leave-one-out and leave-%-out cross-validations (R_{loo} and $R_{l\text{-}\%\text{-}o}$) exceed the accepted value of 0.50, showing that the QSAR model is predictive (Golbraikh and Tropsha, 2002). In addition, the smallest S_{rand} value ($S_{rand} = 0.300$) achieved through the analysis of 10,000 cases of y-randomization was greater than the value found ($S = 0.220$) when true calibration was

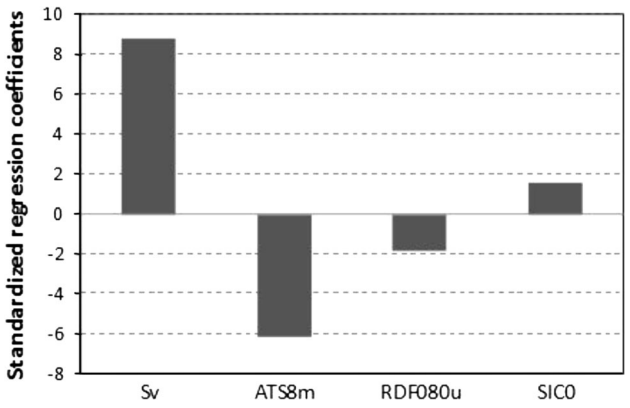

Fig. 3 Standardized regression coefficients of the molecular descriptors employed in equation 1. Refer to Table 3 for descriptors definition

considered. Therefore, the QSAR model found does not result from happenstance.

Thangapandian *et al.* (2013) have developed a six-descriptor QSAR model to characterize the activity of diverse LTA4H inhibitors with a training set of 26 compounds. In contrast to Thangapandian study, and in conformity to the classic semiempirical "rule of thumb" (Hansch, 1990), we have developed a four-descriptor model using a training set of 38 compounds. To the best of our knowledge, we think that a major number of data in the training set would supply more information to the model expanding the diversity of compounds that can be predicted correctly. In addition, a model with a smaller number of descriptors accounts for a simpler relationship with the activity. The statistic parameters indicate that both models have an excellent training, however, our model has greater predictive power ($R_{test} = 0.714$) than the model developed by Thangapandian ($R_{test} = 0.502$).

Molecular descriptors analysis

The four molecular descriptors participating in the QSAR are: Sv, SIC0, ATS8 m, and y RDF080u. The standardization of the regression coefficients of Eq. 1, presented in Fig. 3, allows assigning a greater importance to those molecular descriptors with larger absolute standardized coefficient values. The model descriptors significance decreases in the following order: Sv (8.846) > ATS8 m (6.104) > RDF080u (1.812) > SIC0 (1.573).

Sv descriptor is the sum of atomic van der Waals volumes (scaled on Carbon atom) and it is related to the molecular volume (Todeschini and Consonni, 2009). The positive sign indicates that $\log_{10}IC_{50}$ value is directly proportional to this descriptor. The ATS8 m descriptor (Broto–Moreau autocorrelation descriptor of lag 8 that is weighted by atomic mass) (Nekoei *et al.*, 2011) belongs to the 2D autocorrelation descriptors family. It has a negative sign, which indicates that $\log_{10}IC_{50}$ value is indirectly related to this descriptor. When these descriptors are analyzed together, it can be seen that the activity is adversely affected by the increase of molecular volume because of the restrictions imposed by the size of the active site. However, a very small molecule will not fit properly in the active site. The incorporation of the ATS8 m descriptor compensates this situation because the increase of the molecular mass (and therefore the value of this descriptor) causes a decrease of $\log_{10}IC_{50}$ value.

The RDF080u and SIC0 descriptors have less influence on the activity, Fig. 3. The Radial Distribution Function (RDF) is a kind of molecular descriptor defined for an ensemble of atoms, and may be interpreted as the probability distribution for finding an atom in a spherical volume of certain radius, incorporating different types of atomic properties in order to differentiate the nature and contribution of atoms to the property being modeled. In the case of RDF080u, the sphere radius is 8.0 Å and no atomic property is used, characterizing the molecular size. This descriptor also reveals an enthalpic contribution on activity (related to the interactions of hydrogen bond and van der Waals types) and it is important for hydrophobic interactions with an enzyme (Zarei and Atabati, 2009; Vicente *et al.*, 2010; Jain *et al.*, 2011). A high RDF080u value suggests that the compound has great capacity for establishing hydrophobic interactions increasing its biological activity. This is indicated by the negative sign of the descriptor coefficient (see Eq. 1 and Fig. 3).

Table 4 Predicted $\log_{10}IC_{50}$ values for the previously synthesized (Kirkland *et al.*, 2008)

Compound	$\log_{10}IC_{50}$ Eq.1	IC_{50} [nM]	ΔG_{bind} [kcal mol^{-1}]
51[a]	2.331	214.29	−10.7
52[a]	2.205	160.32	−11.1
53[a]	2.509	322.85	−10.7
54[a]	2.314	206.06	−10.8
100	0.507	3.21	−10.8
103	0.931	8.54	−10.3
104	0.716	5.20	−10.9
105	0.493	3.11	−9.8
129	0.889	7.75	−10.3
131	0.887	7.72	−11.5
134	0.206	1.61	−11.7
135	0.924	8.39	−9.7
138	0.352	2.25	−11.7
139	0.165	1.46	−10.2

Predicted $\text{Log}_{10}IC_{50}$ values and ΔG_{bind} for the ten designed compounds of lower IC_{50} values

[a] Extracted from Kirkland *et al.* (2008)

The SIC0 topological descriptor corresponds to the structural information content (neighborhood symmetry of 0-order) family. It measures the complexity of a compound as the diversity of elements present in its molecular structure, such as atoms, bonds, cycles, etc. and encodes the size and the degree of branching in the compound. It is known that the size and the shape of the molecule also affect the intermolecular interaction (Natarajan *et al*., 2008; Ramírez-Galicia *et al*., 2012). In this case, the sign suggests that an increase in SIC0 (compounds with high molecular complexity) has a negative effect on the biological activity.

Design of novel compounds

The information provided above, in conjunction with the molecular docking results, could be employed in the rational design of new inhibitors. We used this information alongside a molecular modulation technique and designed 85 new *N-alkyl glycine amides* compounds. The molecular modulation consists in varying the lead compound with limited modifications (the structure should maintain the initial characteristics) in order to find a better product with higher activity, improved bioavailability, reduced toxicity, and minimal secondary reactions. Despite these apparent limitations in the possibility of variation, it is common to find positive results in applying this technique, provided it is done rationally. In this work, some molecules with IC_{50} values $\leq$ 60 nM (compounds 19, 29, 30, 31, 33, and 34, see Table 1) were considered as lead compounds to propose new structures.

The IC_{50} values of the 85 new compounds, listed in Table 1S of supplementary material, were predicted by means of our QSAR model, eq. 1. In addition of these, four previously synthesized *N-alkyl glycine amides* (Kirkland *et al*., 2008) with unknown experimental IC_{50} values were also examined. We found that ten out of the 85 new designed structures show predicted IC_{50} values below 10 nM (Table 4), lower than the values of the lead compounds, while the four synthesized by Kirkland compounds

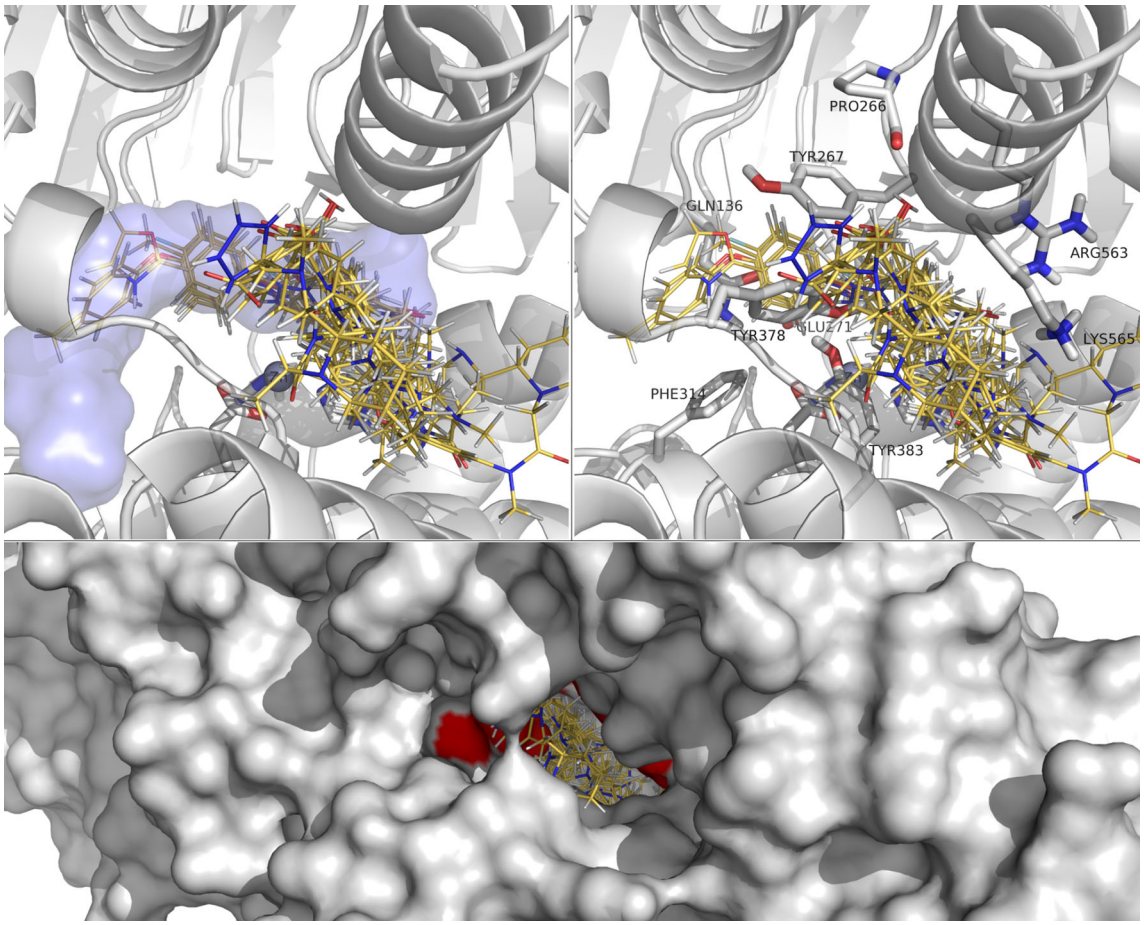

Fig. 4 Molecular docking of the new designed *N-alkyl glycine amides* with predicted IC_{50} lower than 10 nM (*yellow sticks*) into the active site of LTA4H (represented in either *gray cartoon* or *gray surfaces*). The LTA4 binding site, represented as a *light blue* surface, is partially occupied by the inhibitors (*top left panel*) which interact with the amino acid side chains represented in *gray sticks* (*top right panel*). The active site entry is sterically hindered (*bottom panel*) (Color figure online)

have a predicted activity greater than 160 nM. We think that the ten compounds presented here, could be candidates for future pharmacological studies.

Variations in the structure of the lead compounds generated these ten designed molecules. According to the descriptors Sv and ATS8m included in the QSAR model, the molecular volume must not be markedly modified. Hence, the incorporation of the ringed substituents in compounds 134, 135, and 138 could negatively influence on the activity. However, these substituents provide a marked possibility of interacting with the enzyme by means of hydrogen bond interactions such as indicated by the RDF080u descriptor. The substituents of the designed compounds have simple chemical structure relatively easy to be synthesized and tested experimentally.

To get insights into the binding mode of these ten compounds with LTA4H, a molecular docking analysis was performed. Fig. 4 presents the superposition of the best docking pose of each new compound into LTA4H. All the structures are located within the active site of the enzyme where LTA4 (natural ligand) binds, top left panel of Fig. 4. The interactions that take place involve the same regions of the enzyme described for the compounds of the training set, as can be seen comparing Fig. 1 (right panel) and Fig. 4 (top right panel). Although only a part of the LTA4 binding site is occupied by the inhibitors, the active site entry is sterically hindered, limiting the access to the natural ligand (Fig. 4, bottom panel). The calculated values of ΔG_{bind} (-9.6 to -11.7 kcal.mol^{-1}) are similar to the values of the lead compounds which means that they could have high affinity for the enzyme.

Conclusion

We carried out a quantitative structure activity relationship analysis in conjunction with a molecular docking on a series of the 50 *N-alkyl glycine amides*. The results obtained were used to elucidate the interaction with the enzyme and test the predicted inhibitory activity of four previously synthesized (Kirkland *et al*., 2008) and 85 designed compounds.

The docking of the 50 studied compounds reveals the high affinity of some them with the active site of the enzyme. These compounds interact with Lys565, Tyr378, and Tyr267, similar to the LTA4, and also with Pro266, Phe314, and Gln136. Then, we developed a four-descriptor QSAR model which have a good predictive power, $R_{train} = 0.814$ and $R_{test} = 0.714$. The descriptors of the model are related to the volume and molecular mass, the ability to make hydrophobic interactions and molecular diversity, important features in these ligand-enzyme interactions. The result is a simpler QSAR model and high predictive capability compared with other models found in the literature which discusses this biological activity.

The predictions of the tested compounds show that ten compounds have IC_{50} values lower than 10 nM and four of them have values lower than 5 nM, Table 4. The ΔG_{bind} values indicate that the affinity of these new compounds for the active site is similar to the 50 *N-alkyl glycine amides* series and the LTA4.

We consider that the information provided in this report can be used as useful, fast, and costless tool for future investigations and development of new potential LTA4H inhibitors.

Acknowledgments This work was supported by the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) project PIP11220100100151 and Universidad Nacional de San Luis (UNSL).

References

Back M, Bu DX, Branstrom R, Sheikine Y, Yan ZQ, Hansson GK (2005) Leukotriene B4 signaling through NF-{kappa}B-dependent BLT1 receptors on vascular smooth muscle cells in atherosclerosis and intimal hyperplasia. Proc Natl Acad Sci USA 102:17501–17506

Barnes PJ (2001) Future advances in COPD therapy. Respiration 68:441–448

Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M (1977) The Protein Data Bank: a computer-based archival file for macromolecular structures. J Mol Biol 112:535–542

Çalışkan B, Banoglu E (2013) Overview of recent drug discovery approaches for new generation leukotriene A4 hydrolase inhibitors. Expert Opin Drug Discov 8:49–63

Chen X, Wang S, Wu N, Yang CS (2004) Leukotriene A4 hydrolase as a target for cancer prevention and therapy. Curr Cancer Drug Targets 4:267–283

Duchowicz PR, Castro EA, Fernandez FM, González MP (2005) A new search algorithm for QSPR/QSAR theories: normal boiling points of some organic molecules. Chem phys lett 412:376–380

Golbraikh A, Tropsha A (2002) Beware of q2! J Mol Graph Model 20:269–276

Goodarzi K, Goodarzi M, Tager AM, Luster AD, von Andrian UH (2003) Leukotriene B4 and BLT1 control cytotoxic effector T cell recruitment to inflamed tissues. Nat Immunol 4:965–973

Haeggström JZ (2004) Leukotriene A4 hydrolase/aminopeptidase, the gatekeeper of chemotactic leukotriene B4 biosynthesis. J Biol Chem 279:50639–50642

Hansch C (1990) Comprehensive drug design. Pergamon Press, New York

Jain A, Sharma R, Chaturvedi SC (2011) Structural Insight for benzimidazole as angiotensin II AT1 receptor antagonist by using molecular property and biolgogical activity correlation: Qsar approach. Int J Pharm PharmSci 3:541–546

Kirkland TA, Adler M, Bauman JG, Chen M, Haeggstrçm JZ, King B, Kochanny MJ, Liang AM, Mendoza L, Phillips GB, Thunnissen M, Trinh L, Whitlow M, Ye B, Ye H, Parkinson J, Guilford WJ (2008) Synthesis of glutamic acid analogs as potent inhibitors of leukotriene A4 hydrolase. Bioorgan Med Chem 16:4963–4983

Mercader AG, Duchowicz PR, Fernandez FM, Castro EA (2010) Replacement method and enhanced replacement method versus

the genetic algorithm approach for the selection of molecular descriptors in QSPR/QSAR theories. J Chem Inf Model 50:1542–1548

Natarajan R, Basak SC, Mills D, Kraker JJ, Hawkinsb DM (2008) Quantitative structure-activity relationship modeling of mosquito repellents using calculated descriptors. Croat Chem Acta 81:333–340

Nekoei M, Salami M, Dolatabadi M, Mohammadhosseini M (2011) A quantitative structure–activity relationship study of tetrabutylphosphonium bromide analogs as muscarinic acetylcholine receptors agonists. J Serb Chem Soc 76:1117–1127

Paz PB, Vega-Hissi EG, Estrada MR, Garro Martinez JC (2012) In silico modeling of the molecular structure and minding of leukotriene A4 into leukotriene A4 hydrolase. Chem Biol Drug Des 80:902–908

Penning TD (2001) Inhibitors of Leukotriene A4 (LTA4) hydrolase as potential anti inflammatory agents. Curr Pharm Des 7:163–179

Ramírez-Galicia G, Garduño-Juárez R, Correa-Basurto J, Deeb O (2012) Exploring QSAR for inhibitory effect a set of heterocyclic thrombin inhibitors by multilinear regression refined by artificial neural network and molecular docking simulations. J Enzym Inhib Med Chem 27:174–186

Rinaldo-Matthis A, Haeggström JZ (2010) Structures and mechanisms of enzymes in the leukotriene cascade. Biochimie 92:676–681

Rudberg PC, Tholander F, Andberg M, Thunnissen MM, Haeggström JZ (2002) Leukotriene A4 hydrolase: identification of a common carboxylate recognition site for the epoxide hydrolase and aminopeptidase substrates. J Biol Chem 279:27376–27382

Sandanayaka V, Mamat B, Mishra RK, Winger J, Krohn M, Zhou LM, Keyvan M, Enache L, Sullins D, Onua E, Zhang J, Halldorsdottir G, Sigthorsdottir H, Thorlaksdottir A, Sigthorsson G, Thorsteinnsdottir M, Davies DR, Stewart LJ, Zembower DE, Andresson T, Kiselyov AS, Singh J, Gurney ME (2010a) Discovery of 4-[(2S)-2-{[4-(4-Chlorophenoxy)phenoxy]methyl}-1-pyrrolidinyl]butanoic Acid (DG-051) as a Novel Leukotriene A4 Hydrolase Inhibitor of Leukotriene B4 Biosynthesis. J Med Chem 53:573–585

Sandanayaka V, Mamat B, Mishra RK, Winger J, Krohn M, Zhou LM, Keyvan M, Enache L, Sullins D, Onua E, Zhang J, Halldorsdottir G, Sigthorsdottir H, Thorlaksdottir A, Sigthorsson G, Thorsteinnsdottir M, Davies DR, Stewart LJ, Zembower DE, Andresson T, Kiselyov AS, Singh J, Gurney ME (2010b) Discovery of 4-[(2S)-2-{[4-(4-Chlorophenoxy)phenoxy]methyl}-1-pyrrolidinyl]butanoic Acid (DG-051) as a Novel Leukotriene A4 Hydrolase Inhibitor of Leukotriene B4 Biosynthesis. J Med Chem 53:573–585

Shim YM, Paige M (2012) Leukotriene A4 hydrolase–an evolving therapeutic target. In: Khatami Mahin (ed) inflammatory diseases-immunopathology, clinical and pharmacological bases. InTech, west palm Beach, pp 254–278

Tanis VM, Bacani GM, Blevitt JM, Chrovian CC, Crawford S, De Leon A, Fourie AM, Gomez L, Grice CA, Herman K, Kearney AM, Landry-Bayle AM, Lee-Dutra A, Nelson J, Riley JP, Santillán A Jr, Wiener JJM, Xue X, Young AL (2012) Azabenzthiazole inhibitors of leukotriene A4 hydrolase. Bioorg Med Chem Lett 22:7504–7511

Thangapandian S, John S, Sakkiah S, Lee KW (2011) Pharmacophore-based virtual screening and Bayesian model for the identification of potential human leukotriene A4 hydrolase inhibitors. Eur J Med Chem 46:1593–1603

Thangapandian S, John S, Son M, Arulalapperumal V, Lee KW (2013) Development of Predictive QSAR Model and Its Application in the Discovery of Human Leukotriene A4 Hydrolase Inhibitors. Future Med Chem 5:27–40

Thunnissen MM, Nordlund P, Haeggström JZ (2001) Crystal structure of human leukotriene A(4) hydrolase, a bifunctional enzyme in inflammation. Nat Struct Biol 8:131–135

Thunnissen MM, Andersson B, Samuelsson B, Wong CH, Haeggstrom JZ (2002) Crystal structures of leukotriene A4 hydrolase in complex with captopril and two competitive tight-binding inhibitors. J FASEB 16:1648–1650

Todeschini R, Consonni V (2009) Molecular Descriptors for Chemoinformatics. Wiley, Weinheim

Trott O, Olson AJ (2010) AutoDock VINA: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. J Comp Chem 31:455–461

Vicente E, Duchowicz PR, Benítez D, Castro EA, Cerecettoc H, González M, Monge A (2010) Anti-T cruzi activities and QSAR studies of 3-arylquinoxaline-2-carbonitrile di-N-oxides. Bioorg Med Chem Lett 20:4831–4835

Wold S, Eriksson L (1995) Chemometrics methods in molecular design. Wiley, Weinheim

Wunder F, Tinel H, Kast R, Geerts A, Becker EM, Kolkhof P, Hutter J, Erguden J, Harter M (2010) Pharmacological characterization of the first potent and selective antagonist at the cysteinyl leukotriene 2 (CysLT2) receptor. Br J Pharmacol 160:399–409

Ye B, Bauman J, Chen M, Davey D, Khim SK, King B, Kirkland T, Kochanny M, Liang A, Lentz D, Maya K, Mendoza L, Phillips G, Selchau V, Schlyer S, Tseng JL, Wei RG, Ye H, Parkinson J, Guilford WJ (2008) Synthesis of N-alkyl glycine amides as potent inhibitors of leukotriene A4 hydrolase. Bioorg Med Chem Lett 18:3891–3894

Zarei K, Atabati M (2009) QSAR study of anti-HIV activities against HIV-1 and some of their mutant strains for a group of HEPT derivatives. J Chin Chem Soc 56:206–213