Eur. J. Med. Chem. 36 (2001) 21–30

PII: S0223-5234(00)01185-5/FLA

Original article

Protein-based alignment in 3D QSAR of 26 indole inhibitors of human pancreatic phospholipase A_2

Marco Pintore[a], Philippe Bernard[a,b], Jean-Yves Berthon[b], Jacques R. Chrétien[a]*

[a]Laboratory of Chemometrics & BioInformatics, Faculty of Sciences, University of Orléans, BP 6759, 45067 Orléans Cedex 2, France
[b]Greentech S. A., Biopôle Clermont Limagne, 63360 Saint-Beauzire Cedex, France

Received 20 December 1999; revised 13 July 2000; accepted 2 August 2000

Abstract – An automated docking procedure was applied on a series of 26 reversible and competitive indole inhibitors of human pancreatic phospholipase A_2 (hp-PLA_2). X-ray data of this enzyme are not available and the structure was then reconstructed exploiting its protein sequence and the crystallographic data of a bovine pancreatic source. The docking data were used to build a three-dimensional quantitative structure–activity relationship (3D QSAR) model, established using the comparative molecular field analysis (CoMFA) method. This model, joined to the previous one developed for the indole inhibitors of human non-pancreatic secretory phospholipase A_2 (hnps-PLA_2), an enzyme involved in inflammation processes, will allow for the selection of new strong anti-inflammatory drugs with negligible side effects, at least at the level of hp-PLA_2. © 2001 Éditions scientifiques et médicales Elsevier SAS

pancreatic phospholipase A_2 / anti-inflammatory drugs / automated docking / 3D QSAR / CoMFA

1. Introduction

Phospholipase A_2 (PLA_2) corresponds to a class of enzymes that catalyze the hydrolysis of membrane glycerophospholipids at the *sn*-2 position to release fatty acids and lysophospholipids. When the fatty acid is arachidonic acid, a complementary metabolism leads to pro-inflammatory mediators such as prostaglandins, leukotrienes, thromboxanes and platelet activating factors (PAF) [1]. Thus, modulating the production of pro-inflammatory lipid mediators by inhibiting PLA_2 activity remains a potential target for the development of new drugs in the treatment of inflammatory diseases [2].

The enzyme is widespread in bacteria, plants, snake and bee venoms, mammalian cells, and secretions [3, 4]. Evidence exists that more than one form of PLA_2 might be found [5, 6]. Classes I, II and III of PLA_2 are small proteins of 119–143 amino acids, with molecular weights ranging between 12 and 15 kDa. Class I have been isolated from *Elapidae* and *Hydrophidae* snake venoms and mammalian pancreas, whereas class II PLA_2s are present in the snake venoms of *Crotalidae* and *Viperidae* species and in a variety of non-pancreatic mammalian tissues [7]. Class III enzymes have principally been isolated from lizard and bee venoms [8]. Finally, with the discovery of an 85 kDa cytosolic PLA_2, an additional group of PLA_2s, class IV, has been proposed [9].

In a previous paper concerning the research project to find new natural bioactive molecules, active against inflammation, by combining ethnopharmacology studies with bioinformatics, the human non-pancreatic secretory phospholipase A_2 (hnps-PLA_2), belonging to class II PLA_2s, was found to be the best target to reach this goal [10]. High concentrations of this enzyme were reported in septic shocks [11, 12], multiple injuries [13], and rheumatoid arthritis [14]. Thus a hnps-PLA_2 would be a potential drug against disor-

* Correspondence and reprints.
E-mail address: jacques.chretien@univ-orleans.fr (J.R. Chrétien).

ders generated by high levels of this enzyme. An automated docking [15] was performed on a series of 188 reversible and competitive hnps-PLA_2 inhibitors susceptible to be potential drugs in the modulation of the inflammation process [16–18]. The docking data were used to perform a comparative molecular field analysis (CoMFA) [19, 21] in order to build a three-dimensional quantitative structure–activity relationship (3D QSAR) model [19–22]. After having divided the 188 hnps-PLA_2 inhibitors into two sets, a training set and a test one, the CoMFA model with a protein-based alignment showed a good capacity to correctly predict the inhibitory capacities of the molecules of the test set. Finally, the pharmacophore of hnps-PLA_2 was elaborated.

A search for the best inhibitors against hnps-PLA_2 has to be associated with a high specificity of the molecules selected. In fact, to maximize the efficacy of the drug and to avoid possible human side effects, it is important that the inhibitor's activity against human pancreatic phospholipase A_2 (hp-PLA_2), which shows a negligible concentration in all tissues except in pancreas [23], should be kept extremely low.

The aim of this paper is to build, similarly to the case of hnps-PLA_2, a 3D QSAR model of hp-PLA_2. Pancreatic PLA_2 (pan-PLA_2) is synthesized and secreted by pancreatic cells into the pancreatic juice where it serves as a digestive enzyme [24]. In acute pancreatitis, increased concentrations of pan-PLA_2 can be found in patient sera [25]. The enzyme is synthesized as inactive prophospholipase A_2 [26] which is activated by tryptic hydrolysis in the lumen of the duodenum. PLA_2 is activated within the pancreas in acute pancreatitis [27].

Differently of the case of hnps-PLA_2, a more limited number of experimental results is available in literature, where only 26 indole hp-PLA_2 have been found [17, 18]. The task is complicated by the lack of crystallographic data concerning the pancreatic enzyme; but knowing the 2D structure and using the protein homology of the 3D bovine and porcine pan-PLA_2 might help to derive a 3D QSAR model of hp-PLA_2. In agreement with the results obtained in the hnps-PLA_2 case, in which it was evidenced the importance of a total docking to obtain a high quality CoMFA model, all the inhibitors were docked to hp-PLA_2 and the derived structures were used as a protein-based alignment to feed the CoMFA procedure [28, 29].

2. Methods

2.1. Automated docking

2.1.1. Enzyme

Many sources of bovine [30] and porcine [31] pancreatic phospholipase A_2 have been used and crystallized, but X-ray data concerning the enzyme present in the human pancreas do not exist. However, the proteine sequence of hp-PLA_2 is available on the Swiss-Prot Database, under the entry name PA21_HUMAN [32, 33]. Then, following the main lines of a procedure used by Christensen et al. in 1993 [34], the hp-PLA_2 structure was reconstructed exploiting the non-human pancreatic sources. The crystal structures of non-hpPLA_2 were obtained from the Brookhaven Protein Data Bank (PDB). Twenty-three hp-PLA_2 proteins are presently available from PDB, but only five are complexed with a ligand; four structures are of bovine origin (PDB code: 1BP2 [35], 1FDK [36], 1MKV [30] and 3BP2 [37]) and one of porcine origin (PDB code: 5P2P [31]). Only hp-PLA_2–ligand complexes were selected because they include a 'pocket', around the active site, able to receive the inhibitor ligands more easily. Porcine PLA_2 is a mutant protein and was immediately discarded because, compared to hp-PLA_2, it shows a wide range of different residues in the protein sequence. The four ligands associated to bovine structures are respectively, (i) 2-methyl-2,4-pentanediol (MPD) (PDB code: 1BP2), (ii) 1-decyl-3-trifluoro ethyl-sn-glycero-2-phopshomethanol (GLE) (PDB code: 1FDK), (iii) L-1-o-octyl-2-heptylphosphonyl-sn-glycero-3-phosphoethanolamine (GEL) (PDB code: 1MKV) and (iv) pyruvoil group (PVL) (PDB code: 3BP2). Amidst the bovine structures, the 1MKV complex was chosen to fit hp-PLA_2 because the GEL ligand is more similar to the structures of 26 indole inhibitors. Furthermore, it binds the Ca^{2+} ion of the active site as a bidentate ligand, with the same estimated mechanism as with the hp-PLA_2 inhibitors. The comparison between the protein sequence of the PA21_HUMAN and 1MKV enzymes showed that the two chains differ for 29 residues, but only five residues are included in the region delimited by a cut-off of 5 Å around all the atoms of the GEL ligand. Only the latter five residues were changed in the 1MKV structure (see *table 1*) to reproduce the PA21_HUMAN sequence, because the influence of the other peripheral residues on the inhibition properties should be negligible. To make the new structure (MKV_HUMAN) suitable for further modeling, hydrogen atoms were added to the modified

PDB source using the BIOPOLYMER module of SYBYL [38]. Then the geometry of the protein was optimized in two steps. Firstly a complete conformational search on the five modified residues was made using the SYBYL/SYSTEMATIC SEARCH; all rotable bonds were relaxed and the rotation step was also fixed at 20°. The total energy of the enzyme–ligand complex was calculated for each conformation generated and the lowest energy one was minimized using the Tripos force field. The lowest energy geometry of the MKV_HUMAN was aligned to the 1MKV bovine protein structure and the different enzymatic zones were found to superpose completely. The GEL ligand was removed from the MKV_HUMAN structure and the geometry of the protein was optimized using the Tripos force field.

Table I. Modified residues between the 1MKV protein and MKV-HUMAN in the region delimited by a cut-off of 5 Å around all the atoms of the GEL ligand.

Residue	1MKV	MKV-HUMAN
2	Leu	Val
6	Asn	Arg
19	Leu	Phe
22	Phe	Tyr
53	Lys	Asp

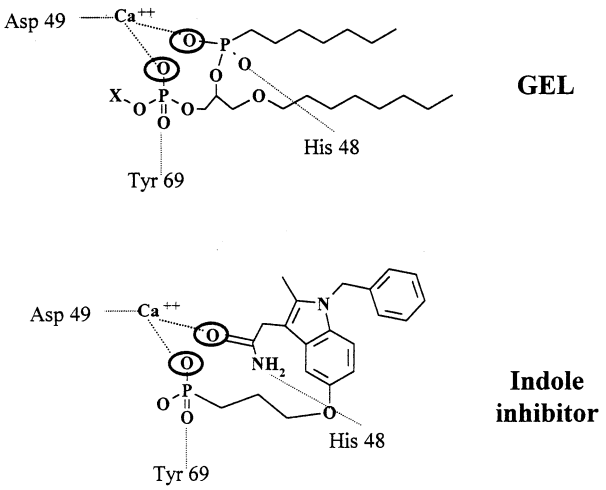


Figure 1. Schematic representation of an inhibitor binding to hp-PLA_2. Homology binding between the GEL ligand and compound **18**, an indole inhibitor. Atoms selected as anchor points for the automated docking procedure are circled with a bold line.

2.1.2. Ligands

Each of the 26 indole ligands was modeled using SYBYL 6.5 on a Silicon Graphics O_2 R10000 station coupled to a Silicon Graphics Origin 200 R10000×2 server. The ligands were considered in their neutral form. The starting conformations were optimized by molecular mechanics algorithm using the Tripos Force Field. The lowest energy conformations were found by means of the SYBYL/SEARCH option and then used as initial conformations for docking.

2.1.3. Ligand–enzyme docking

The FIT ATOMS option of SYBYL allowed to place the phosphoryl group or carbonyl group of each indole inhibitor into the enzyme, prepared as mentioned above, at the position corresponding to the GEL phosphoryl group. The two oxygen atoms linked to the calcium atom of the enzyme were successively used as the anchor point, as shown in *figure 1*. The molecule was then rotated around the three coordinate axes as previously described [37] with a step of 30°. For each orientation, a complete conformational search was made using the SYBYL/SYSTEMATIC SEARCH option. All rotable bonds were relaxed and the rotation step was also fixed at 30°. The total energy of the enzyme–ligand complex was calculated for each generated conformation. The geometries of the lowest energy conformations found for the enzyme–ligand complexes were then refined by further spatial rotations and conformational search with a step of 5° in the interval of ±30° around the torsion values obtained. These refined enzyme–ligand complexes were optimized using the Tripos force field. The lowest energy geometries of the inhibitors were then used in the CoMFA analysis.

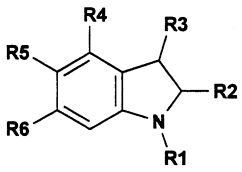
2.2. CoMFA

2.2.1. Data set

The 26 indole inhibitors and the corresponding biological data used in this study were selected from literature [17, 18]. The molecular structures and hnps PLA_2 inhibitory activity data for these 26 indole derivatives are summarized in *table II*. All the collected biological data (IC_{50}) were measured in vitro under the same experimental conditions [39].

The 26 compounds were divided into two subsets: 18 of them were used as a training set and eight were used as a test set. Every third compound was selected from the series and gathered in the test set. This choice raises

Table II. In vitro inhibition of hp-PLA_2 by 26 indole inhibitors.



Compound	R1	R2	R3	R4	R5	R6	IC_{50} (μM)	Log $1/IC_{50}$
1	$C_6H_5CH_2$	CH_3	$COCONH_2$	OCH_2CO_2H	H	H	0.761	0.12
2	2-$(C_6H_5)C_6H_4CH_2$	CH_3	$COCONH_2$	OCH_2CO_2H	H	H	0.364	0.44
3	3-$(C_6H_5)C_6H_4CH_2$	CH_3	$COCONH_2$	OCH_2CO_2H	H	H	0.57	0.24
4	4-$(C_6H_5)C_6H_4CH_2$	CH_3	$COCONH_2$	OCH_2CO_2H	H	H	1.09	−0.04
5	1-naphthylCH_2	CH_3	$COCONH_2$	OCH_2CO_2H	H	H	1.2	−0.08
6	n-C_8H_{17}	CH_3	$COCONH_2$	OCH_2CO_2H	H	H	0.78	0.11
7	$C_6H_5CH_2$	CH_3CH_2	$COCONH_2$	OCH_2CO_2H	H	H	0.228	0.64
8	2-$(C_6H_5CH_2)C_6H_4CH_2$	CH_3CH_2	$COCONH_2$	OCH_2CO_2H	H	H	0.062	1.21
9	3-$ClC_6H_4CH_2$	CH_3CH_2	$COCONH_2$	OCH_2CO_2H	H	H	0.390	0.41
10	$C_6H_5CH_2$	CH_3	CH_2CONH_2	OCH_2CO_2H	H	H	1.4	−0.15
11	$C_6H_5CH_2$	CH_3	CH_2CONH_2	$O(CH_2)_3CO_2H$	H	H	3.66	−0.56
12	$C_6H_5CH_2$	CH_3	CH_2CONH_2	H	$O(CH_2)_3CO_2H$	H	69	−1.84
13	$C_6H_5CH_2$	CH_3	CH_2CONH_2	H	(2-C_6H_4)CH_2	H	22.5	−1.35
14	$C_6H_5CH_2$	CH_3CH_2	CH_2CONH_2	OCH_2CO_2H	H	H	1.8	−0.26
15	$C_6H_5CH_2$	CH_3CH_2	CH_2CONH_2	H	$O(CH_2)_3CO_2H$	H	94	−1.97
16	$C_6H_5CH_2$	Br	CH_2CONH_2	H	$O(CH_2)_3CO_2H$	H	15.9	−1.20
17	$C_6H_5CH_2$	CH_3	CH_2CONH_2	$OCH_2PO(OH)_2$	H	H	73.5	−1.87
18	$C_6H_5CH_2$	CH_3	CH_2CONH_2	H	$O(CH_2)_3PO(OH)_2$	H	67	−1.83
19	$C_6H_5CH_2$	CH_3CH_2	CH_2CONH_2	H	$O(CH_2)_3PO(OH)_2$	H	91.1	−1.96
20	$C_6H_5CH_2$	Br	CH_2CONH_2	H	$O(CH_2)_3PO(OH)_2$	H	6.2	−0.79
21	3-$ClC_6H_4CH_2$	CH_3CH_2	CH_2CONH_2	H	$O(CH_2)_3PO(OH)_2$	H	46.2	−1.67
22	2-$(C_6H_5)C_6H_4CH_2$	CH_3	CH_2CONH_2	H	$O(CH_2)_3PO(OH)_2$	H	39	−1.59
23	$C_6H_5CH_2$	CH_3CH_2	CH_2CONH_2	H	$O(CH_2)_3SO_3H$	H	135	−2.13
24	$C_6H_5CH_2$	CH_3	CH_2CONH_2	H	$(CH_2)_3CO_2H$	H	94	−1.97
25	$C_6H_5CH_2$	CH_3CH_2	CH_2CONH_2	H	$S(CH_2)_3CO_2H$	H	16	−1.20
26	$C_6H_5CH_2$	Br	CH_2CONH_2	H	$O(CH_2)_3CO_2H$	Cl	3.2	−0.51

Table III. Statistics and cross-validation results of the three CoMFA models.

		Steric	Electrostatic	Steric and electrostatic
N			18	
'leave-one-out'	n_{PC}	4	3	4
	Q^2	0.73	0.66	0.70
'leave-some-out' (10 groups)	n_{PC}	4	4	4
	Q^2	0.72	0.65	0.71
R^2		0.99	0.96	0.98
F		368	109	199
sd		0.10	0.21	0.16

the problem of the separation of a series of compounds into a training set and a test set. The order of the compounds found in *table III* follows an order related to their organic synthesis. *Table II* is formed by the sum of chemical subseries. Thus, selecting every third compound corresponds in fact to making a selection inside each chemical subseries. This allows a good molecular diversity in the test set.

2.2.2. CoMFA method

A CoMFA study normally begins with searching a suitable alignment of the compounds under investigation by using a constrained reference compound. In the present study this problem was a priori solved by docking the compounds to the hp-PLA_2 derived structure.

Both the steric and electrostatic CoMFA fields were calculated using a sp^3 carbon atom as a probe, with a charge of +1 in the grid points around the molecules. The grid points were spaced 1.5 Å apart in all three dimensions. Partial atomic charges for the electrostatic field calculation were obtained by the MOPAC AM1 method [40]. The CoMFA region was chosen to include all the molecules with margins of at least 4.0 Å. The CoMFA region was defined as a cube with an edge of 24 Å. The field values were truncated at +30 kcal mol^{-1} for steric and ±30 kcal mol^{-1} for electrostatic interactions.

The partial least-squares method (PLS) [41] was used to linearly relate the CoMFA fields to the inhibitory activity values. The optimal number of PLS components was determined using the 'leave-one-out' and 'leave-some-out' (5 groups) cross-validation procedures [41]. The model quality is expressed in terms of Q^2 the cross-validated correlation coefficient, R^2 the conventional correlation coefficient, s the standard error, and F the Fisher test.

3. Results and discussion

3.1. Docking

In the present section, inhibitor structures will be discussed to illustrate the most characteristic ligand–protein interactions. The biological activity data for the 26 hp-PLA_2 inhibitors are homogeneously distributed on the activity scale, ranging from 62 nM for compound **8** up to 135 μM for compound **23**. Within this range, all the compounds were considered as being able to enter the active site of the enzyme and to interfere with it via the same mechanism.

The automated docking procedure was applied for all the compounds in the series. As X-ray studies on hp-PLA_2 are not available, the crystallographic data concerning hnps-PLA_2 [42] and 1MKV [30] proteins were exploited. These data showed two interactions between two oxygen atoms of the ligand and the calcium atom of the protein. The automated docking procedure was applied with these two oxygen atoms as anchor points. In all the lowest energy conformations, the calcium atom is bonded by the oxygen amide group and by one oxygen of the carboxyl group or the phosphonate group. Only in the case of compound **23**, the least active in the series, the calcium atom is bonded exclusively by the oxygen amide. The protein based alignment of all ligands allows for the definition of a volume with a size of about 14.5×16×11 Å.

Figure 2 presents compound **8**, the most active compound in the series, docked to the hp-PLA_2. This study revealed the two interactions already described between two oxygen atoms of the ligand and the calcium atom, and two additional interactions: i) between residue His48 and the NH moiety of the amide group; ii) between residue Phe5 and the carbonyl group connected to the amide group. The latter are hydrogen bonds with a distance of about 1.8 and 2.4 Å between the acceptor and the donor, respectively.

Electrostatic forces are very important in the substitutions occurring at the R2, R3, R4, and R6 positions on indole inhibitors to modulate the activity. In the R2 position the substitution of a Br atom for a CH_3 group increases the activity 11 times (see compounds **18** and **20)**, in the R6 position the substitution of a chlorine atom for an hydrogen atom increases the activity five times (see compounds **16** and **26**). In the R3 position, the substitution of a CH_2 group for a carbonyl group decreases the activity to the ninth when considering, for example, compounds **7** ($CO–CONH_2$) and **14** ($CH_2–CONH_2$). In the R4 position, the substitution of a sulfonate group for a carboxyl group decreases the activity by a factor 53 when considering compounds **10** and **17**.

As for the R1 position, the inhibition capacity is governed also by the electrostatic factors, but their contribution remains relatively weak. In fact, the substitution of a phenyl group or a chlorine atom for an hydrogen atom increases the activity 1.5/2 times (see, for example, compounds **18** and **22** or **19** and **21**).

Steric hindrance is an important factor in the activity modulation. In fact, the displacement of the same carboxylic chain from the R4 to the R5 position on indole inhibitors decreases the activity by a factor 19 when considering for example compounds **11** and **12**. The decrease in activity is due to the presence of residue Tyr69 near these positions, which probably causes also the non-binding of the large sulfonate group at the calcium atom in compound **23**.

3.2. CoMFA

In agreement with the results reported in previous papers [14, 15], in the present CoMFA study each of the 26 indole inhibitors studied was aligned in an independent way. No reference compound was used and the geometries of the inhibitors were only determined by interactions with the biological receptor. The automated docking study showed that, despite the flexibility of the molecules studied, the choice of the suitable conformations was not ambiguous. Indeed, in most cases the conformational search yielded only a single conformation (or a single conformational family) for the ligand–protein complex whose energy was much lower than that of the other candidate complexes. Hence, it may be suggested once more that the active site of the protein is quite selective and does not permit the conformational mobility of the bonded inhibitor.

The protein-based alignment may have an impact on the nature of the structural information that we expect to obtain from the CoMFA analysis. A good statistic and predictive quality of the CoMFA model normally suggests that the molecules in the real biological system are aligned in accordance with the initial alignment of the CoMFA procedure. In the present case, as the alignment was prepared using the structure of the receptor as a template, the good quality of the CoMFA model would suggest the validity of the proposed model of inhibitor–enzyme interactions obtained by the automated docking procedure.

The separation of the series into a training set and a test set was realized as described in the materials and methods section. For a better understanding of the factors which underlie the activity, three different CoMFA models were derived: (i) a model with the steric field only, (ii) a model with the electrostatic field and (iii) a model taking both fields into account. The results of these analyses are presented in *table III* and underline the similarity of the role played by the electrostatic and steric fields. This remark is in agreement with the previous discussion in the docking section, where the considerable importance of the electrostatic and steric contributions was shown. Indeed, the analysis including both fields shows that the relative contributions to the

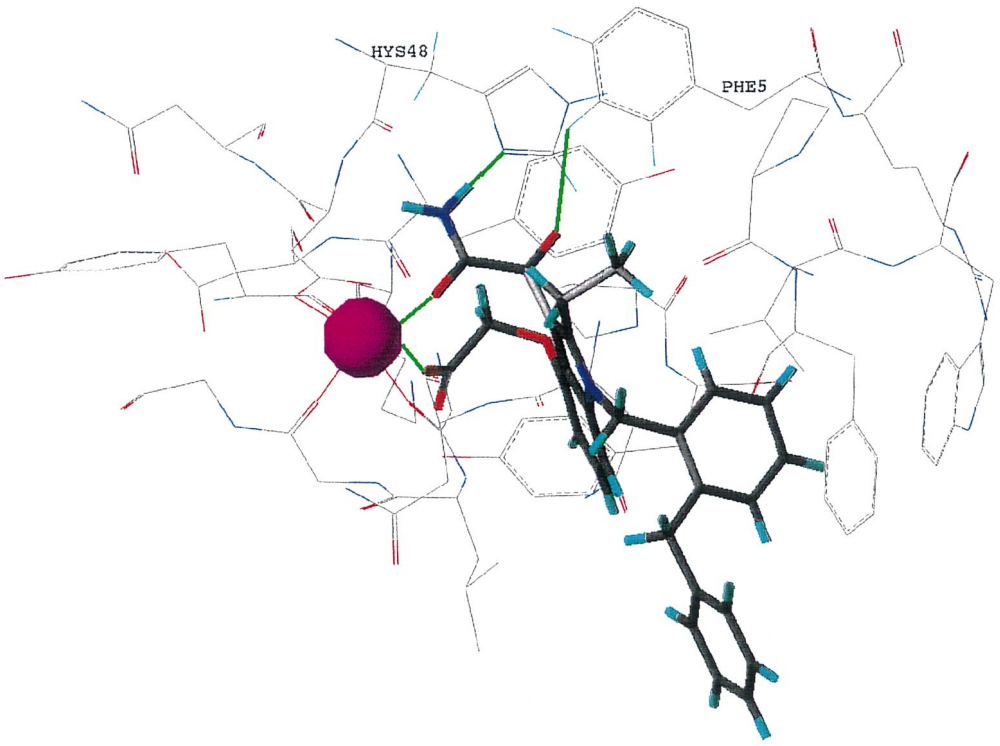


Figure 2. Enzyme–ligand complex obtained by automated docking between indole inhibitor **8** and hp-PLA_2. The violet ball represents the calcium ion. The oxygen, nitrogen and hydrogen atoms are shown in red, blue and cyan, respectively. The most important interactions between the ligand and the enzyme are reported with a green line.

Figure 3. CoMFA electrostatic STDEV*COEFF field plot. Increasing negative charges inside the red regions and increasing positive charges in the blue regions favor the inhibitory capacity. Compound **8** is shown.

model are 53 % for the steric field and 47 % for the electrostatic one. Moreover, the statistic and predictive criteria for the analyses including either the steric field or the electrostatic one are very similar to that including both fields. In the present study, to remain coherent with the previous model developed for hnps-PLA_2 [10], both the steric and electrostatic fields ($Q^2 = 0.70$) were finally taken to plot the CoMFA statistic fields and to predict the hp-PLA_2 inhibitory activity for new compounds.

Finally, a non-cross-validated PLS run was performed with four principal components in order to generate the final CoMFA plots. These plots outline a CoMFA statistic field which expresses the relationship between the variation of the steric and electrostatic fields and the variation of the biological activity. The values of the fields are calculated at each lattice intersection and are equal to the product of the descriptor coefficient by the corresponding standard deviation (STDEV*COEFF). Hence, an extremely low value of STDEV*COEFF indicates that the presence of the corresponding steric or electrostatic field at this point decreases the activity. A high value of STDEV*COEFF means that the presence of fragments produces a field favorable to the activity. Compound **8** in *figure 3* and compound **22** in *figure 4* illustrate the main features of the CoMFA plots. *Figure 3* shows that important electrostatic variations are registered in all the positions on indole inhibitors and, above all, that in several zones (R1, R3 and R4) significant positive and negative contributions exist. This behavior, together with the considerable importance of the steric hindrance, can explain that in the CoMFA analysis the total values are nearly the arithmetical mean of the steric and electrostatic contributions. Some of the colored regions on the plots mark essential ligand–protein interactions. For instance, the red zone facing two carboxyl oxygen atoms of the inhibitor indicates that an oxygen atom is able to form an electrostatic bond with the calcium atom whereas the second oxygen atom is

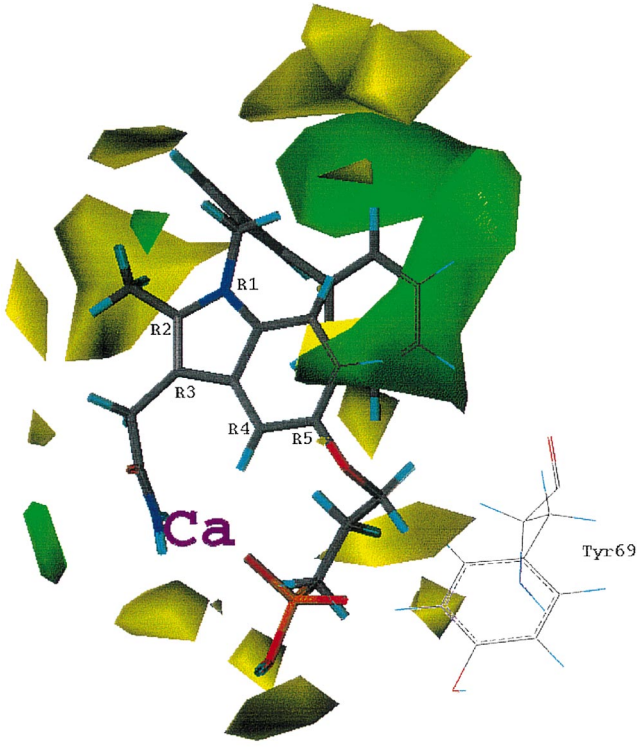


Figure 4. CoMFA steric STDEV*COEFF field plot. Increasing bulk effects inside the green regions and decreasing bulk effects in the yellow regions favor the inhibitory capacity. Compound **22** is shown.

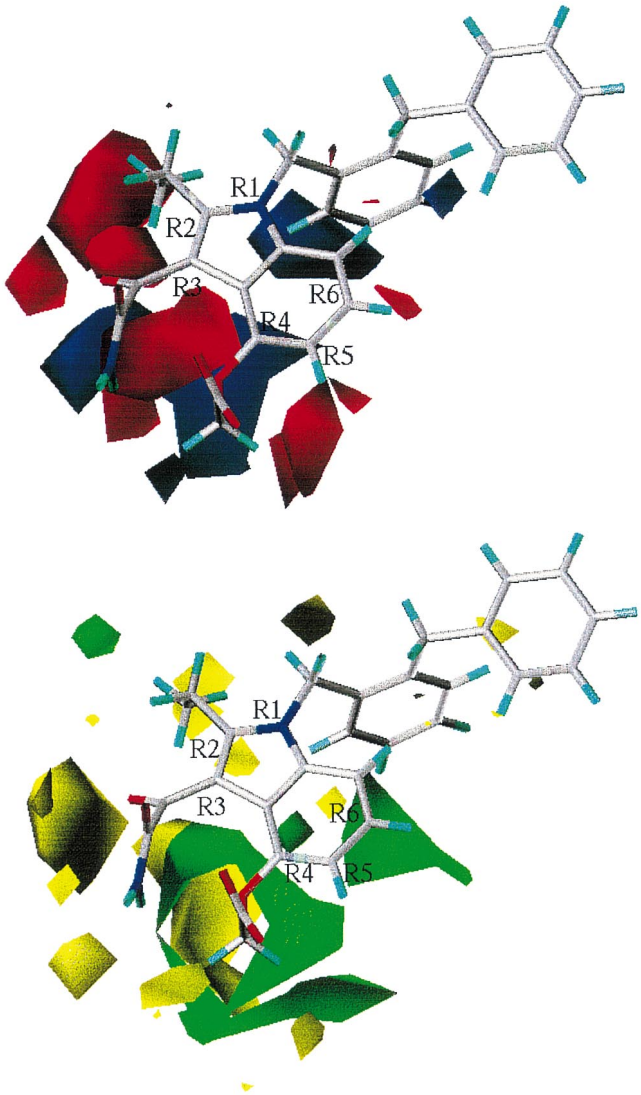


Figure 5. CoMFA steric and electrostatic STDEV*COEFF field plots in hnps-PLA_2 for compound **8** [10]. The colored regions represents the same relations between field and inhibitory capacity defined in *figures 3* and *4*.

directed toward Tyr69. The blue and largest volume covering the R3 amide group explains the importance of hydrogen bond donors with residue His48, whereas the red region represents the hydrogen bond acceptor with residue Phe5. In *figure 4*, the yellow volumes beside the R5 group indicate the position of Tyr69 where no bulky substitution is possible. The region around the R1 position is dipped between green and yellow zones and a bulky substitution is possible if the group is sufficiently flexible. Indeed, the substitution of a $(C_6H_5CH_2)C_6H_4$ group for a naphthyl group increases the activity about 20 times when considering compounds **5** and **8**. Substitutions exhibit unfavorable steric effects near the R2 position around the ethyl group. As the 26 compounds never include bulky substituents, it will be interesting to further investigate this R2 position, which looks very critical.

The comparison of the *figures 3* and *4* with the *figure 5*, in which the CoMFA electrostatic and steric field plots for the compound **8** in hnps-PLA_2 are represented [10], allows to explicate: i) the reasons of the differential binding ability of PLA_2 non-pancreatic and pancreatic; ii) how this different ability could be exploited to design new compounds with strong anti-inflammatory properties but negligible side effects, at least at level of hp-PLA_2.

For example, for compounds **23** [16–18], the key factor to explain the different activity between hp-PLA_2 and hnps-PLA_2 (with a rapport of 1:2700), consists in the presence in hp-PLA_2 of the residue Tyr69, absent in hnps-PLA_2 [10], that prevents the binding at the calcium atom of the inhibitors with bulky R5 position substituents.

Globally, from the CoMFA plot analysis, it is possible deriving that a high selective hnps-PLA_2 inhibition could be got: i) adding bulky groups in R5 position, ii) selecting ligands with positively charged groups in R6 and R7; iii) employing negatively charged groups in position R2. The positions R3 and R4 do not allow any selectivity, as implied in Ca binding and in the same hydrogen bond network for both the PLA_2 enzymes. Finally, the position R1 needs further investigation, but its contribution should be negligible, being dipped between opposite regions for both the electrostatic and steric fields.

3.3. Predictive aspect of the CoMFA model

After validating our model by means of cross-validation, the next step of the investigation consisted in applying the model to a test set of inhibitors independently from the training set with which the model was established.

The compound numbers in the test set are referring to *table II*, predicted and real activity values with the corresponding deviations for the above eight indole inhibitors are shown in *table IV*. The correlation coefficient R between the experimental and predicted data has a value of 0.96. It means that the relative inhibitory capacities are correctly predicted for the whole series of

Table IV. Predicted activity of the 8 hp-PLA_2 indole inhibitors with the 3D QSAR model.

Compound	Experimental Log $1/IC_{50}$	Predicted Log $1/IC_{50}$	Δ (Exp.-Pred.)
3	0.24	0.53	−0.29
6	0.11	0.06	0.05
9	0.41	0.11	0.30
12	−1.84	−1.42	−0.42
15	−1.97	−1.54	−0.43
18	−1.83	−1.45	−0.38
21	−1.67	−1.97	0.30
24	−1.97	−1.50	−0.47

eight molecules. The parameter which is widely used to estimate the quality of test set predictions, i.e. *PRESS*/*SSY* [43], has an excellent value of 0.07, which indicates no important deviations of the predicted values from the actual ones. Clementi and Wold suggested that: 'In a reasonable QSAR model, *PRESS*/*SSY* should be smaller than 0.4, whereas a value smaller than 0.1 for this ratio indicates an excellent model' [43].

Two conclusions may be derived from this part of the study. First, the tight correlation between the predicted and experimentally observed values in this study suggests that the present model is able to provide reliable predictions of the hp-PLA_2 inhibitory capacities in a set of new inhibitors. The second conclusion is that the good statistical parameters of the CoMFA model and its good predictive capacities allow to validate the inhibitor conformations selected by the docking procedure.

4. Conclusion

An automated docking study was performed on a series of 26 competitive hp-PLA_2 inhibitors, in order to build a model parallel to the one previously developed for hnps-PLA_2, a protein involved in inflammation processes. In this way, it will be possible to select new strong hnps-PLA_2 inhibiting drugs that have also a high specificity, that is negligible side effects at least at the level of hp-PLA_2.

No X-ray data concerning hp-PLA_2 are available, but a structure was reconstructed exploiting the protein sequence deposited in the Swiss-Prot Database and fitting the crystallographic data of a bovine pancreatic source. The docking study showed that the calcium atom of the protein is bonded by the oxygen amide group and by one oxygen of the carboxyl group or the phosphonate group of the ligand. Only for the least active compound of the series, the calcium atom is bonded exclusively by the oxygen amide. Furthermore, two additional interactions are present: the first between residue His48 and the NH moiety of the amide group and the other between residue Phe5 and the carbonyl group connected to the amide group.

The docking data were then used to perform a comparative molecular field analysis. The 26 hp-PLA_2 inhibitors were divided into two sets, one with 18 compounds for the model, the other with eight compounds for the validation of the model. The CoMFA model with the protein-based alignment has a good predictive capacity according to the cross-validation test ($Q^2 = 0.70$). The inhibitory capacities were correctly predicted for the eight molecules of the test set.

The CoMFA plots are in agreement with the docking results and could be used efficiently, coupled with the CoMFA model derived for hnps-PLA_2, to design new effective and selective inhibitors, to allow the development of a virtual screening of large databases. The efficacy of this virtual screening will be tested by experimental enzymatic analysis, comparing the experimental and the calculated inhibitory activities. Moreover, the new experimental results will be entered into the above database to improve the drug design model.

References

[1] Irvine R.F., Biochem. J. 204 (1982) 3–10.
[2] Mayer R.J., Marshall L.A., FASEB J. 7 (1993) 339–348.
[3] Bérézia G., Etienne J., Kokkinidis M., Oliver J.L., Pernas P., J. Lipid Mediators 2 (1990) 159–172.
[4] van den Bosch K., Biochem. Biophys. Acta 604 (1980) 191–246.
[5] Waite M., The Phospholipases: Handbook of Lipid Research, Plenum Press, New York, 1988.

[6] Arni R.K., Ward R.J., Toxicon 34 (1996) 827–841.

[7] Dennis E.A., Rhee S.G., Billah M.M., Hannun Y.A., FASEB J. 5 (1991) 2068–2077.

[8] Renetseder R., Brunie S., Dijkstra B.W., Drenth J., Sigler P.B., J. Biol. Chem. 260 (1985) 11627–11636.

[9] Clark J.D., Milona N., Knopf J.L., Proc. Natl. Acad. Sci 87 (1990) 7708–7712.

[10] Bernard P.P., Berthon J.-Y., Chrétien J.R., Eur. J. Med. Chem., (2000) in press.

[11] Green J.A., Smith G.M., Buchta R., Lee R., Ho K.Y., Rajkovic I.A., Scott K.F., Inflammation 15 (1991) 355–367.

[12] Vadas P., Pruzanski W., Circ. Shock 39 (1993) 160–167.

[13] Uhl W., Buchler M., Nevalainen T.J., Deller A., Beger H.G., J. Trauma. 30 (1990) 1285–1290.

[14] Hara S., Kudo I., Chang H.W., Matsuta K., J. Biochem. 105 (1989) 395–399.

[15] Bernard P., Kireev D.B., Chretien J.R., Fortier P.-L., Coppet L., J. Mol. Model. (2000) in press.

[16] Dillard R.D., Bach N.J., Draheim S.E., Berry D.R., Carlson D.G., Chirgadze N.Y., Clawson D.K., Hartley L.W., Johnson L.M., Jones N.D., McKinney E.R., Mihelich E.D., Olkowski J.L., Schevitz R.W., Smith A.C., Snyder D.W., Sommers C.D., Wery J.-P., J. Med. Chem. 39 (1996) 5119–5136.

[17] Dillard R.D., Bach N.J., Draheim S.E., Berry D.R., Carlson D.G., Chirgadze N.Y., Clawson D.K., Hartley L.W., Johnson L.M., Jones N.D., McKinney E.R., Mihelich E.D., Olkowski J.L., Schevitz R.W., Smith A.C., Snyder D.W., Sommers C.D., Wery J.-P., J. Med. Chem. 39 (1996) 5137–5158.

[18] Draheim S.E., Bach N.J., Dillard R.D., Berry D.R., Carlson D.G., Chirgadze N.Y., Clawson D.K., Hartley L.W., Johnson L.M., Jones N.D., McKinney E.R., Mihelich E.D., Olkowski J.L., Schevitz R.W., Smith A.C., Snyder D.W., Sommers C.D., Wery J.-P., J. Med. Chem. 39 (1996) 5159–5175.

[19] Ortiz A.R., Pastor M., Palomer A., Cruciani G., Gago F., Wade R.C., J. Med. Chem. 40 (1997) 1136–1148.

[20] Kuntz I.D., Meng E.C., Shoichet B.K., Acc. Chem. Res. 27 (1994) 117–123.

[21] Cramer R.D. III, Patterson D.E., Bunce J.D., J. Amer. Chem. Soc. 110 (1988) 5959–5967.

[22] Ortiz A.R., Pisabarro M.T., Gago F., Wade R.C., J. Med. Chem. 38 (1995) 2681–2691. Published erratum appears in additions and corrections, J. Med. Chem. 40 (1997) 4168.

[23] Nevalainen T.J., Haapanen T.J., Inflammation 17 (1993) 453–464.

[24] de Haas G.H., Postema N.M., Nieuwenhuizen W., van Deenen L.L.M., Biochim. Biophys. Acta 159 (1968) 103–117.

[25] Nevalainen T.J., Eskola J.U., Aho H.J., Havia V.T., Lövgren N.-E., Näntö V., Clin. Chem. 31 (1985) 1116–1120.

[26] de Haas G.H., Postema N.M., Nieuwenhuizen W., van Deenen L.L.M., Biochim. Biophys. Acta 159 (1968) 117–129.

[27] Nevalainen T.J., Scand. J. Gastroenterol. 15 (1980) 641–650.

[28] Bromidge S.M., Dabbs S., Davies D.T., Duckworth D.M., Forbes I.T., Ham P., Jones G.E., King F.D., Saunders D.V., Starr S., Thewlis K.M., Wyman P.A., Blaney F.E., Naylor C.B., Bailey F., Blackburn T.P., Holland V., Kennett G.A., Riley G.J., Wood M.D., J. Med. Chem. 41 (1998) 1598–1612.

[29] Matter H., Schwab W., Barbier D., Billen G., Haase B., Neises B., Schudok M., Thorwart W., Schreuder H., Schreuder H., Brachvogel V., Lönze P, Weithmann K.U., J. Med. Chem. 42 (1999) 1908–1920.

[30] Sekar K., Kumar A., Liu X., Tsai M.D., Gelb M.H., Sundarilingam M., Acta Crystallogr. D Biol. Crystallogr. 54 (1998) 334–341.

[31] Thunnissen M.M., Ab E., Kalk K.H., Drenth J., Dijkstra B.V., Kuipers O.P., de Haas G.H., Verheij H.M., Nature 347 (1990) 689–691.

[32] Grataroli R., Dijkman R., Dutilh C.E., van der Ouderaa F., de Haas G.H., Figarella C., J. Biochem. 122 (1982) 111–117.

[33] Verheij H.M., Westerman J., Sternby B., de Haas G.H., Biochim. Biophys. Acta 747 (1983) 93–99.

[34] Christensen I.T., Jørgensen F.S., Svensson L.Å, Högberg T., Drug Des. Discov. 10 (1993) 101–113.

[35] Dijkstra B.V., Kalk K.H., Hol H.W.J., Drenth J., J. Mol. Biol. 147 (1981) 97–123.

[36] Sekar K., Eswaramoorthy S., Jain M.K., Sundaralingam M., Biochemistry 36 (1997) 14186–14191.

[37] Dijkstra B.V., Kalk K.H., Hol H.W.J., Drenth J., de Haas G.H., Egmond M.R., Slotboom A.J., Biochemistry 23 (1984) 2759–2766.

[38] SYBYL is available from Tripos Associates, 1699 South Hanley Road, St Louis, MO 63144.

[39] Reynolds L.J., Hughes L.L., Dennis E.A., Anal. Biochem. 204 (1992) 190–197.

[40] Dewar M.J.S., Zoebich E.G., Healy E.F., Stewart J.J.P., J. Amer. Chem. Soc. 107 (1985) 3902–3909.

[41] Wold S., Eriksson L., in: Waterbeemd H. (Ed.), Chemometric Methods in Molecular Design, VCH, Weinheim, 1995, pp. 309–318.

[42] Schevitz R.W., Bach N.J., Carlson D.G., Chirgadze N.Y., Clawson D.K., Dillard R.D., Draheim S.E., Hartley L.W., Jones N.D., Mihelich E.D., Olkowski J.L., Snyder D.W., Sommers C., Wery J.-P., Nat. Struct. Biol. 2 (1995) 458–465.

[43] Clementi S., Wold S., in: van de Waterbeemd H (Ed.), Chemometric Methods in Molecular Design, VCH, Weinheim, 1995, pp. 319–338.