J Mol Model (2012) 18:4417–4426
DOI 10.1007/s00894-012-1434-z

ORIGINAL PAPER

Exploring the substructural space of indole-3-carboxamide derivatives binding to renin: a novel active-site spatial partitioning approach

Tao Jing · Jian Feng · Yumei Zuo · Boli Ran · Jianping Liu · Guoxiang He

Received: 19 January 2012 / Accepted: 16 April 2012 / Published online: 16 May 2012

Abstract Renin has recently attracted much attention in the antihypertensive community, since this enzyme starts the angiotensin-converting cascade and forms the rate-limiting step in this cascade. In the present study, we describe a new method called active-site spatial partitioning (ASSP) for quantitatively characterizing the nonbonding interaction profile between renin and the substructures of indole-3-carboxamide derivatives—a novel class of achiral renin inhibitors that exhibit both high affinity and strong specificity for renin, thus blocking its active state—on the basis of structural models of protein–ligand complexes. It is shown that the ASSP-derived potential parameters are highly correlated with the experimentally measured activities of indole-3-carboxamides; the statistical models linking the parameters and activities using a sophisticated partial least squares regression technique show much promise as an effective and powerful tool for generalizing and predicting the pharmaceutical potencies and the physicochemical properties of other modified derivatives. Furthermore, by visually examining substructure-color plots generated by the ASSP procedure, it is found that the relative importance of nonbonding contributions to the recognition and binding of a ligand by renin is as follows: steric < hydrophobic < electrostatic. The polar and charged moieties that float on the surface of the ligand molecule play a critical role in conferring electrostatic stability and specificity to renin–ligand complexes, whereas the aromatic rings embedded in the core region of the ligand are the main source of hydrophobic and steric potentials that lead to substantial stabilization of the complex architecture.

Keywords Antihypertensive drug · Renin · Active-site spatial partitioning · Quantitative structure–activity relationship

Electronic supplementary material The online version of this article (doi:10.1007/s00894-012-1434-z) contains supplementary material, which is available to authorized users.

T. Jing · J. Feng · Y. Zuo · B. Ran · J. Liu (✉) · G. He (✉)
Department of Cardiology, Southwest Hospital,
Third Military Medical University,
Chongqing 400038, China
e-mail: ljp1022@mail.tmmu.com.cn
e-mail: yyxnk@yahoo.com.cn

Introduction

Persistent hypertension is one of the risk factors for stroke, myocardial infarction, heart failure, and arterial aneurysm, and is the leading cause of mortality in the Western world [1]. The renin–angiotensin system (RAS) plays a key role in the regulation of blood pressure and in the maintenance of sodium and volume homeostasis, which can be activated when there is a loss of blood volume or a drop in blood pressure (such as in hemorrhage). Alternatively, a decrease in the filtrate NaCl concentration and/or a decreased filtrate flow rate will stimulate the macula densa to release renin. Renin stimulates the production of angiotensin I from angiotensinogen, an α_2-globulin that is produced constitutively and released into the circulation mainly by the liver. Angiotensin I by itself is inactive. However, when acted upon by angiotensin-converting enzyme (ACE), it gets converted to angiotensin II, which is active and is responsible for most of the pressor effects [2]. Since the action of renin represents the rate-limiting step in this cascade and angiotensinogen is its only known substrate, inhibiting this step would be a very effective antihypertensive strategy [3].

Direct renin inhibitors are the newest type of medicine for high blood pressure. These block renin from entering the active state, thus preventing the production of angiotensin I from angiotensinogen. As a consequence, blood vessels

relax and widen, making it easier for blood to flow through the vessels. Monoclonal antibodies were used as early inhibitors of renin, and these proved to be excellent probes of enzyme function. However, they were in no way suitable for use as a medication, as most were immunogenic and had to be administered parenterally [4]. Transition state analogs were then synthesized and found to be potent inhibitors of renin [5], but they had drawbacks because of their peptide-like nature and their lack of oral bioavailability. Further modifications of these lead compounds led to the development of aliskiren (Fig. 1a), the only direct renin inhibitor that is clinically used as an antihypertensive drug [6]. A review of six large-scale clinical trials of aliskiren was published in the May 2007 issue of the *American Journal of Hypertension*. The authors report that, because of reactive renin secretion, this drug is not any more effective than those already widely used to control hypertension [7]. In addition, the four chiral centers and high molecular weight associated with the aliskiren molecule make it rather difficult to synthesize and separate this compound. In order to realize new antihypertensive drugs with low structural complexity and high biological activity and oral bioavailability, several new classes of direct renin inhibitors have been developed recently [8, 9], among which nonpeptide indole-3-carboxamide derivatives (Fig. 1b) seem to be good candidates. Indole-3-carboxamide derivatives have no chiral centers and possess low molecular weights and considerable rigidity [9]. In addition, the crystal structures of three indole-3-carboxamide compounds in complex with renin have been solved at a high level of resolution, and unexpected binding modes were observed in these complex systems. These are particularly interesting from the perspectives of elucidating the physicochemical basis of inhibitor recognition and association with the target protein as well as optimizing the ligand structure to obtain high-affinity and high-specificity molecular entities.

Given the known complex crystal structures and the quantitatively measured affinities between renin and indole-3-carboxamides [9], we would like to answer questions like: how do the substructures of ligands in the complexed state affect their interactions with renin, and which nonbonding properties associated with a given substructure can substantially contribute to the binding potency of ligand molecules? Previously, Politi et al. successfully developed a docking-based scoring strategy aimed at the precise prediction of binding affinity between renin and its inhibitors [10].

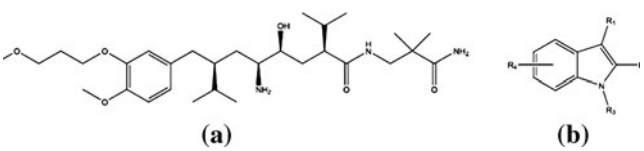

Fig. 1 **a**–**b** Comparison of structural complexity between aliskiren (**a**) and indole-3-carboxamide derivatives (**b**) (R_1, R_2, and R_3 are six-membered ring derivatives while R4 is a short chain)

Alternatively, the quantitative structure–activity relationship (QSAR) approach appears to be a promising way to tackle these problems, as this can correlate the structural properties of ligands to biological activity statistically, and it can explain physicochemical aspects underlying biological activity in a straightforward manner. The long development of QSAR methodology has ranged from the use of the Hansch–Fujita [11] and Free–Wilson [12] methods to topological indices [13], quantum parameters [14], 3D-QSAR [15], hologram QSAR [16], 4D-QSAR [17] (or 5D- [18] or 6D-QSAR [19]), and receptor-based QSAR [20], although 3D-QSAR seems to be the most popular because it can accurately describe the conformational information for drug ligands, and also because a number of powerful 3D-QSAR methods such as CoMFA [21], CoMSIA [22] and SOMFA [23] have been readily available over the past few decades. 3D-QSAR has demonstrated that it is good at predicting the biological activities of ligands and explaining the contributions of molecular structures to this activity, but it mainly concentrates on the ligands rather than the interaction between the ligands and their common receptor. 3D-QSAR also has some other problems that make this method far from perfect (see, e.g., Doweyko's critical review of 3D-QSAR methodology [24]). Since then, Hopfinger and Vedani et al. have attempted to develop post-3D-QSAR approaches such as 4D-, 5D-, and even 6D-QSAR [17–19], all of which aimed to incorporate receptor–ligand interaction information into the modeling process. However, these efforts can only be regarded as indirect approaches to solving this problem, because they do not use real complex structure data (i.e., crystal or NMR structures) to derive QSAR models. With the rapid increase in the number of solved protein–ligand 3D complex structures deposited in the PDB database in recent years, it has been possible to use this direct information to develop receptor-based QSAR (RB-QSAR) approaches. Several RB-QSAR methods have already been proposed to characterize the nonbonding interactions of a ligand with a receptor. However, most of these methods do not split the total interaction of a ligand into the contributions from each structural fragment of the ligand, and they do not provide spatial color plots to straightforwardly explain how the structural fragments influence the biological activity. This is because RB-QSAR methods commonly compute atom-pair nonbonding interactions between ligand and receptor, and these interactions cannot be assigned to a specific region in the active site of the receptor.

In an attempt to (at least partially) solve these problems, we have developed a novel QSAR method called active-site spatial partitioning (ASSP), which dissects the nonbonding profile around the active site of renin and associates the substructural properties of bound ligands with their biological activity. In this procedure, structural models of complexes of

renin bound to indole-3-carboxamide derivatives were obtained computationally, and the space containing the entire active site of renin was then partitioned into thousands of small subregions. Subsequently, the nonbonding potentials between renin and the ligand moieties of each region were calculated using an empirical or semi-empirical formula and correlated with the experimentally determined activity of the ligand molecule using a statistical modeling approach. This allowed several structure-based statistical models to be built which could be used to quantitatively predict the biological activities of unknown indole-3-carboxamide derivatives and to qualitatively explain the nonbonding contributions of bound ligand substructures to the biological activity of the ligand. In this paper, we describe work done to evaluate the potential of ASSP as a predictive and analytical tool. Furthermore, utilizing substructure-color plots generated by the ASSP procedure, the structural foundations of renin–indole-3-carboxamide recognition and the binding process are discussed in detail with respect to their potential application to lead modification and optimization.

Materials and methods

Data set

Very recently, Scheiper et al. reported that a new class of indole-3-carboxamide compounds were potent and achiral renin inhibitors [9]. In their study, the first active candidate for an indole analog was discovered via high-throughput screening (HTS) of an internal collection, and this was rapidly optimized, resulting in phenoxy and benzyl derivatives at the 2 position of the indole. These optimized compounds displayed increased binding activity, and were used as a starting point to create a series of indole-3-carboxamide derivatives. The molecular structures and biological activities of 44 derivatives as well as the first active candidate are tabulated in Table S1 of the "Electronic supplementary material" (ESM), in which the in vitro activity values, pIC_{50}, are expressed as the negative logarithm of IC_{50}, which is the median effect concentration of a given renin inhibitor. Briefly, to calculate this IC_{50} value, the activity of the inhibitor was evaluated in duplicate in a ten point concentration range, performing 2–4 independent experiments. The IC_{50} value was then obtained as the geometric mean of these replicated assays [25].

Setting up the structural models of renin–indole-3-carboxamide complexes

Recently, the crystal structures of the complexes of renin with three indole-3-carboxamide ligands were solved by Scheiper et al. (PDB entries: 3oqk, 3oqf, and 3oot, respectively) [9]. We selected the complex formed by the compound with the highest activity, compound 41 (IC_{50}= 0.002 μM), as the template for constructing structural models of complexes of the other 44 compounds bound to renin. A combination strategy was employed to generate a target complex model from the template. Briefly, the ligand molecule in the template complex was modified manually into a target ligand in the environment of HyperChem 8.0 (Hypercube, Inc.), and then the modified ligand as well as renin residues within 6 Å of the ligand were optimized in an unconstrained manner using the CVFF force field as implemented in Insight II (Accelrys, Inc.). The superposition of the virtually constructed ligands 2 and 3 onto their crystal structures in the active pocket of renin is shown in Fig. 2. As can be seen, the crystal and the constructed structures were fairly well aligned, with corresponding root-mean-square deviations (RMSD) of 0.307 and 0.438 Å for ligands 2 and 3, respectively. This suggests that the combination protocol described above is capable of reproducing the native conformations of the indole-3-carboxamide series in the active site of renin with significant reliability.

Spatial partitioning of the active site of renin

The active-site spatial partitioning (ASSP) proposed in this work is described in detail in the following four steps (Fig. 3):

In *step 1*, as mentioned above, structural models of the complexes of 44 ligands with renin were constructed from the crystal structure of renin–compound 41. The constructed ligands were therefore automatically aligned with compound 41. However, in order to further improve the alignment of the constructed ligands in the renin active pocket so that we could use the subregions with the most unused space ("least full subregions") to accommodate all of the atoms of the ligand, a re-superposition procedure was conducted. Briefly, all 44 constructed ligands as well as the renin residues within a distance of 6 Å from the ligands were re-superposed on the template complex (i.e., containing compound 41) using the popular minimizing RMSD approach, which is also applied in the famous CoMFA approach [21], as implemented in Insight II (Accelrys, Inc.). It is worth noting that, although in our work the alignment was carried out on ligands as well as their surrounding residues, only heavy ligand atoms were considered in the alignment (i.e., residue atoms did not contribute to the alignment; they were just passively translated and rotated with the ligand atoms). This was done because the main goal of the re-superposition was to concentrate ligand atoms, not residue atoms, in the least full subregions.

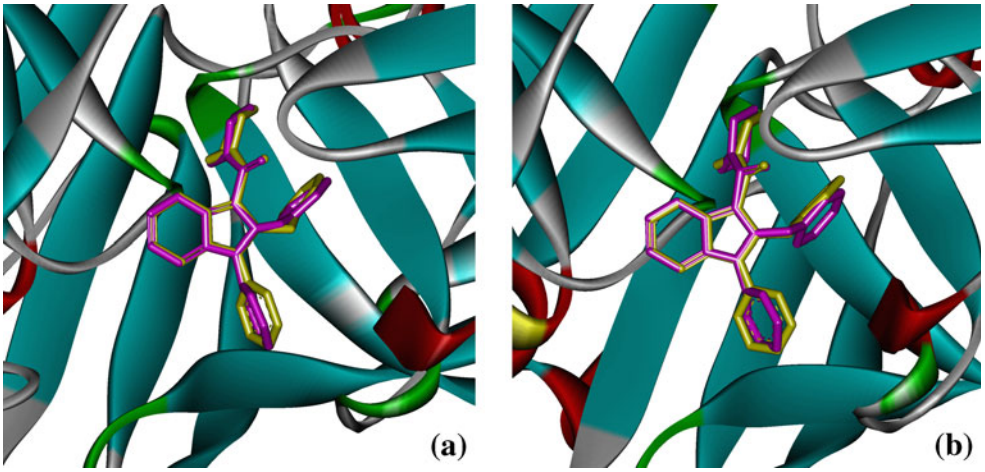

Fig. 2 **a**–**b** Superposition of virtually constructed ligands 2 (**a**) and 3 (**b**) onto their crystal structures in the active pocket of renin. The ligands colored *yellow* and *purple* are native and constructed, respectively

In *step 2*, the consensus space around these superposed ligands was partitioned into thousands of small subregions. This space should contain all of the atoms in the 45 ligands and the dimensions of each subregion in the *x*, *y*, and *z* directions should be substantially less than the common bond length in an organic molecule (~1–1.5 Å) to ensure that each subregion can only contain at most one atom of a specific ligand. Following the principle described above, a space of length (*x*) : width (*y*) : height (*z*)=12 Å : 8 Å : 8 Å around the re-superposed ligands was defined, and 0.8 Å was used as the *x*, *y*, and *z* dimensions of each subregion. Thus, the active-site space was partitioned into 1500 subregions.

In *step 3*, all of the atoms of the ligand were assigned to these 1500 subregions based on the locations of atoms in the active site. In other words, the number of occupied subregions was equal to the number of atoms in the ligand. However, since there are always intrinsic differences between the locations of any two ligands, the occupied subregions should differ to some degree among all 45 superposed ligands. Therefore, the subregions considered in the model were those that contained at least one atom from these 45 ligands. In this way, 347 subregions were selected, and the remaining 1153 empty subregions were then ignored during the subsequent modeling steps. Further, the three types of nonbonding potentials that dominate ligand binding to a host protein, i.e., electrostatic (E), steric (S), and hydrophobic (H) interactions, were calculated between renin and the ligand atoms in each subregion in turn using empirical formulae. The electrostatic potential between an atom of renin and a ligand atom was computed using the classical Coulomb law: $EP_{ij} = \kappa \frac{q_i q_j}{\varepsilon_0 d_{ij}}$, where q_i is the partial charge on atom i and d_{ij} is the

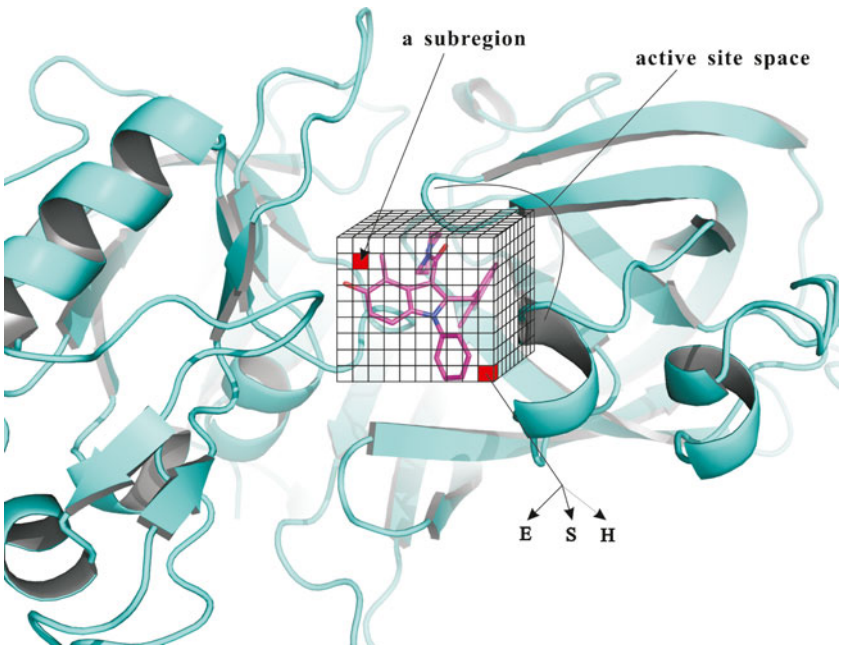

Fig. 3 Schematic representation of the active-site spatial partitioning (ASSP) of a renin–ligand complex. The space around the ligand molecule is partitioned into thousands of small subregions, and three kinds of nonbonding potentials—electrostatic (*E*), steric (*S*), and hydrophobic (*H*) interactions—between renin and the ligand atoms in each subregion are calculated in turn using empirical formulae

distance between atoms i and j. The partial charges on the protein and ligand atoms were assigned using AMBER [26] and Gasteiger [27] charges, respectively. The Lennard–Jones 12–6 potential ($LP_{ij} = \ell_{ij}\left[\left(\frac{D^*_{ij}}{d_{ij}}\right)^{12} - 2\left(\frac{D^*_{ij}}{d_{ij}}\right)^{6}\right]$, where ℓ_{ij} and D^*_{ij} are the potential well and the contact distance between atoms i and j, which were retrieved from the AMBER force field [26]) was used to describe the steric interaction between the renin and ligand atoms. An empirical equation suggested by Zhou et al. [28] was employed to characterize the interatomic hydrophobic potential: $HP_{ij} = -\left(S_i\rho_i + S_j\rho_j\right)e^{-d_{ij}}$, where ρ represents the inherent hydrophobicity of the atoms, which can be measured using Eisenberg's scale [29], and S is the atomic solvent-accessible surface area computed by Sanner's strategy [30]. A detailed description of these empirical potentials can be found in our previous publication [31], where they were successfully employed to quantitatively characterize the nonbonding profile at the binding interface between the SH3 domain and its peptide ligands, and they were considered to be useful for exploring the renin–indole-3-carboxamide complexes in the present work too.

It worth noting that all of the nonbonding potential calculations performed here were done using atomic coordinates. However, different ligands have different numbers of atoms, so directly using atoms to assign parameters would result in different numbers of independent variables for different ligands. Therefore, the strategy of partitioning the ligand space into 1500 consensus subregions was used to obtain a consistent number of variables for different ligands.

In *step 4*, noting that three interaction potentials were calculated for each subregion of a specific ligand, then a total of 3×1500=4500 nonbonding parameters wwere generated for each ligand to characterize its binding profile with the renin receptor. The nonbonding parameters of a series of ligand molecules were combined into an independent variable matrix **X**, which was then linearly correlated with the dependent matrix **Y**, which represented the biological activity pIC_{50}, using a sophisticated partial least squares (PLS) regression technique [32]. This chemometric technique has been widely used to linearly correlate numerous variables with a few experimental points. The PLS regression procedure was divided into three steps. (i) The 347×3=1041 nonzero variables (the other 3459 that were zero were not considered) were recombined and transformed into 1041 latent variables, where each latent variable was the result of linearly combining all 1041 original variables. (ii) A select few of the most informative and statistically significant latent variables were selected as the so-called significant latent variables. (iii) These significant latent variables were linearly correlated with the biological activities of the 45 ligands. In this study, the significant latent variables were selected by leave-one-out cross-validation (LOOCV); that is, latent variables were added to the model one at a one, in order of their relative importance, and the LOOCV q^2 statistic of each model was computed. The optimal number of significant latent variables was thus the number that resulted in the largest q^2 value.

More details on the PLS algorithm can be found in [33–35].

Results and discussion

Development of PLS models

In the areas of statistical modeling and QSAR, validation is one of the most important aspects for developing reliable and predictable models. Although LOOCV has been broadly used for this purpose, in a highly cited paper, Golbraikh and Alexande pointed out that a high LOOCV performance appears to be a necessary but not a sufficient condition for a highly predictive model, and they further emphasized that external validation is the only way to establish a reliable model [36]. In this regard, we randomly selected ten samples from the 45 investigated indole-3-carboxamide derivatives in order to define an external test set that could be used to validate the PLS models built from the internal training set defined by the other 35 unselected compounds. LOOCV was also employed in the present work to perform a preliminary test of the model's stability and generalization ability.

The statistics resulting from the stepwise development of PLS models based upon different combinations of nonbonding components are listed in Table 1. As can be seen, the model's quality is generally improves as the number of nonbonding components used increases, and the final model, which utilizes all three components, shows both the greatest fitting ability (indicated by r^2=0.766) and the highest stability (given by r_{cv}^2=0.704), suggesting that none of the three nonbonding aspects independently dominates the binding behavior of indole-3-carboxamides to renin, and that various properties have significant effects on the binding. Based on the contributions of the different nonbonding components to the PLS models, it seems that the relative importance of these components increases in the order: steric < hydrophobic < electrostatic (for example, the contributions are, respectively, 22.1 % : 32.5 % : 45.4 % in the final model). This is expected, because a number of charged salt bridges and

Table 1 Stepwise development of PLS models based on different combinations of nonbonding components

Components	NL[a]	No validation		LOOCV		Fraction		
		r^{2b}	SD[c]	r_{cv}^2	SD_{cv}	Electrostatic	Steric	Hydrophobic
Electrostatic	4	0.681	0.628	0.594	0.712	1.000	0.000	0.000
Steric	3	0.630	0.671	0.483	0.810	0.000	1.000	0.000
Hydrophobic	4	0.694	0.618	0.560	0.739	0.000	0.000	1.000
Electrostatic + steric	4	0.731	0.560	0.654	0.658	0.745	0.255	0.000
Electrostatic + hydrophobic	5	0.755	0.546	0.678	0.631	0.663	0.000	0.337
Steric + hydrophobic	4	0.727	0.571	0.610	0.694	0.000	0.390	0.610
Electrostatic + steric + hydrophobic	6	0.766	0.529	0.704	0.580	0.454	0.221	0.325

[a] *NL* number of significant latent variables.
[b] r^2 coefficient of determination.
[c] *SD* standard deviation.

polar hydrogen bonds have been observed at the binding interfaces of crystallized renin–ligand complexes [9], and there is a considerable hydrophobic effect due to intense contact between nonpolar renin residues and the aromatic rings of the indole-3-carboxamides (this point will be discussed further in the next section). Furthermore, the optimal model, which incorporates all three nonbonding components, was used to fit the 35 internal training samples and to carry out predictions for ten external test samples. The scatter in the fitted or predicted pIC_{50} with respect to experimentally measured values for these samples are shown in Fig. 4. It is evident that the scatter is fairly low; all of the points are close to the lines of best fit, and there are no obvious outliers that deviate significantly from the lines. Specifically, the coefficient of determination (r_{prd}^2=0.638) and the standard deviation (SD=0.673) obtained for the external test is acceptable, which confirms that this model is statistically reliable and can be exploited as an effective and powerful predictive tool for generalizing and interpreting the biological activities and physicochemical properties of other modified derivatives (vide post).

Analysis of the substructure-color plots

A significant advantage of our ASSP method over traditional molecular analyses is that ASSP not only provides reliable predictions for candidate molecules with unknown activities, but it also provides a series of so-called substructure-color plots (SCPs) that visually show the contributions of diverse substructural segments of the studied compound to the biological activity or binding affinity of that compound. As mentioned above, in the ASSP procedure, the space around the active site of renin is partitioned into 1500 subregions, and each of these is associated with three nonbonding properties, i.e., electrostatic, steric, and hydrophobic aspects. Consequently, 4500 variables characterizing the nonbonding profile between renin and the ligand atoms that belong to different subregions of the active site were calculated and correlated with the biological activities of the ligands using PLS regression. The standardized PLS coefficients provide direct insight into the contribution of each subregion (associated with each nonbonding aspect) to the activity. Given a specific nonbonding aspect, we define a subregion as a favorable one if its corresponding coefficient value satisfies $CC_{\min} + \frac{2}{5} \times (C_{\max} - C_{\min})$ (attractive interactions are always associated with negative potential), an unfavorable one if $CC_{\max} - \frac{2}{5} \times (C_{\max} - C_{\min})$ (repulsive interactions are always associated with positive potential), or an insignificant one (if $C_{\max}$ and $C_{\min}$ are the maximum and minimum values among all coefficients, respectively). The moieties of a ligand can be assigned to different subregions and then marked with distinct colors to indicate the nature of their contributions (favorable, unfavorable, or insignificant) to the ligand's activity. In this way, three SCPs that present the electrostatic, steric, and hydrophobic aspects, respectively, are generated for a studied ligand. In this work, the compound with the highest activity, 41, was used as a paradigm to illustrate the use and significance of SCPs (Fig. 5), and stereoview and schematic representations of the nonbonding interactions between renin and compound 41 are also provided for interpretive purposes in Fig. 6. Before we begin this discussion, if is worth bearing in mind that—as recently pointed out by Bissantz et al. [37]—the contributions of the different parts of a ligand molecule to its affinity are highly dependent on each other, even when these parts are separated by a significant distance.

Electrostatic SCP

The electrostatic interaction is known to be a critical influence on the stability and specificity of the renin–indole-3-carboxamide architecture, since this series of ligands all have a protonated secondary amine in the piperazine ring at the R_1 position (see Table S1) that can form two solid salt bridges with the negatively charged Asp32 and Asp215 residues of renin (Fig. 6a, d). Also, three strong hydrogen bonds were observed between the >C=O/–OH groups of compound 41 and the polar Thr77/His287 residues of renin (Fig. 6a, d). If we compare these to the electrostatic SCP (Fig. 5a), these molecular moieties are always found to provide a favorable

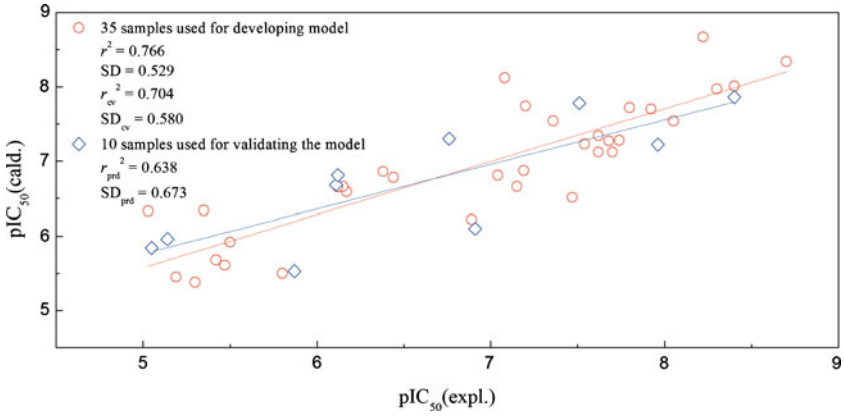

Fig. 4 Plot of calculated against experimental pIC_{50} values for the 35 training samples (shown as *red circles*) and the ten test samples (represented by *blue diamonds*). The PLS model used in this plot incorporated all three nonbonding components

contribution (red) to the biological activity, indicating that the formation of salt bridges and hydrogen bonds by these moieties is crucial to the high activity of compound 41. In addition, two $–CH_3$ groups and an –F group on the indole ring and at the R_2 position also have favorable electrostatic effects on the ligand's activity. This was confirmed by the fact that removing any of these three groups resulted in a substantial decrease in the activity. For example, cutting the indole $–CH_3$ leads to a decrease in the pIC_{50} value from 8.70 (for compound 41) to 8.40 (for compound 40), and cutting R_2–CH_3 and –F leads to a pIC_{50} value of 8.05 (for compound 15). By visually analyzing the crystal structure, we found that there was intense contact between the $–CH_3$ and the polar atoms of renin, and particularly between the –F and the nonpolar H atoms of renin (Fig. 6b), which created a complicated electrostatic network across the ligand-binding interface. Furthermore, all of the electrostatically unfavorable segments in the compound 41 (blue) seem to be close to favorable regions. However, we would expect this given that the strong electronic effect of the polar moieties of the favorable segments would fundamentally influence the electrostatic properties of neighboring unfavorable regions, leading to significant coupling among them.

Steric SCP

An initial glance of the steric SCP (Fig. 5b) suggests that most of the sterically favorable segments (green) are located on the three aromatic rings (indole, R_2, and R_3) of compound 41, which are surrounded by a number of unfavorable moieties (yellow). This can be explained by the nature of the origination of steric potential between the receptor and ligand; the electron-rich aromatic rings of the ligand are easy to polarize and hence are strongly affected by dispersion due to their surroundings, which significantly contributes to the attractive aspect of the steric potential. On the other hand, the atoms and groups that float on the surface of the ligand molecule prefer to penetrate into the body of renin, which leads to minor steric collisions at the interface. A further survey of the steric SCP revealed that the molecular segments associated with the formation of salt bridges, hydrogen bonds, and electrostatic interactions with renin residues are most likely to be the sterically unfavorable regions of the ligand. In fact, it is widely recognized that a molecular moiety tends to clash with its surroundings if this moiety is participating in direct polar interactions with its interacting partner [38]. In this

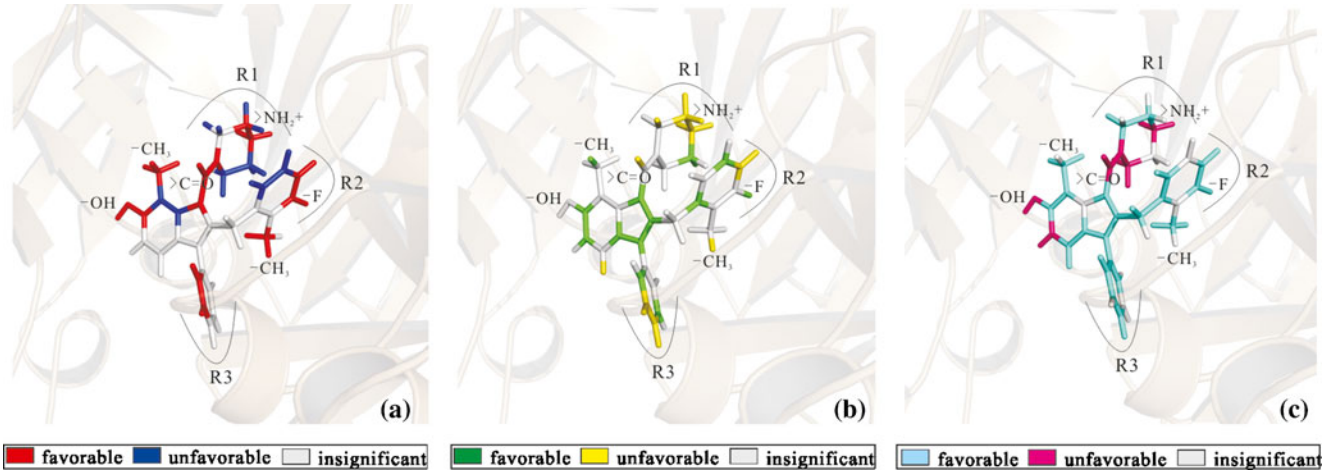

Fig. 5 **a**–**c** Substructure-color plots: **a** electrostatic plot; **b** steric plot; **c** hydrophobic plot. The substructural segments that contribute to the molecular activity in different ways are marked in different colors

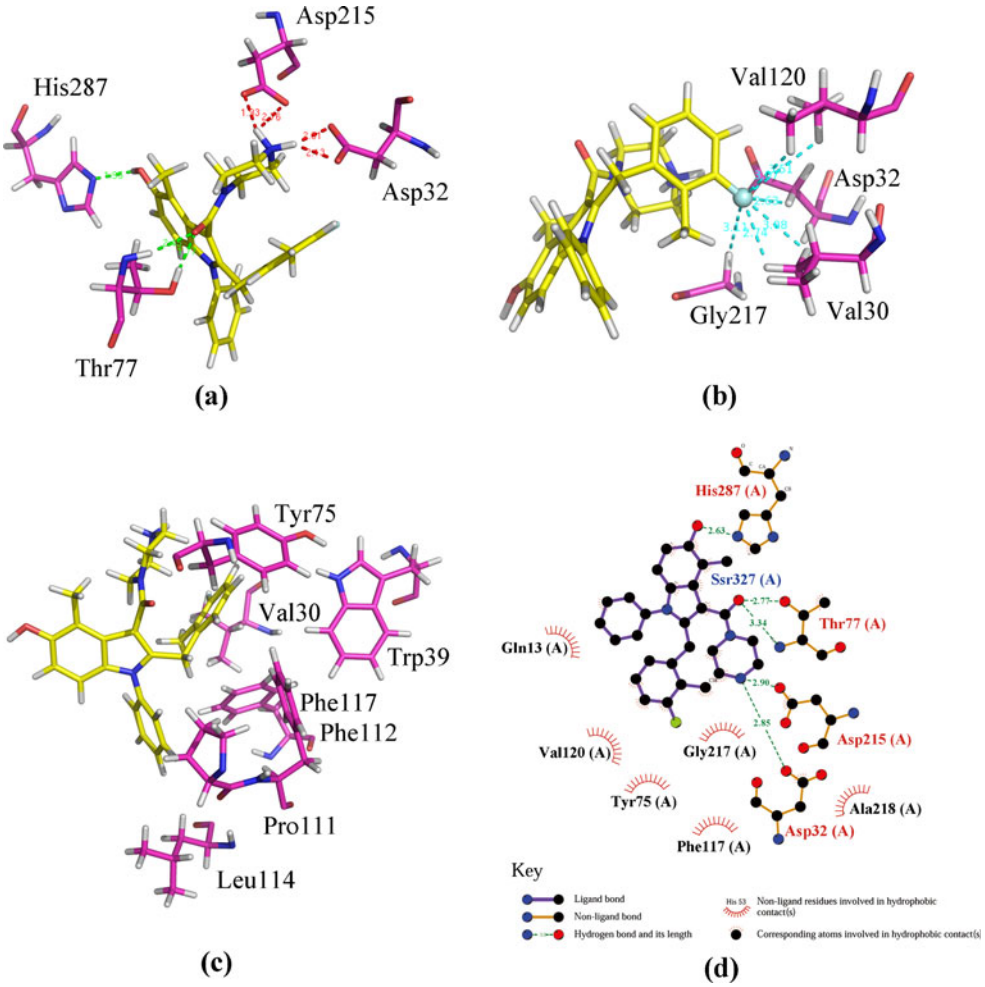

Fig. 6 **a** Close-up of hydrogen bonds and salt bridges between renin and a ligand; **b** close-up of nonbonding interactions between renin and a ligand that involve fluorine; **c** close-up of the distribution of nonpolar residues around a ligand; **d** schematic representation of the nonbonding interactions between renin and a ligand (generated with LIGPLOT [39]). All of these plots were prepared using the crystal structure of renin bound to the compound with the highest activity, compound 41 (PDB entry: 3oot)

way, the unfavorable collisions are offset by the favorable polar interactions.

Hydrophobic SCP

The hydrophobic force is known to be a major factor that promotes the binding of indole-3-carboxamide ligands to renin, since the ligands possess a number of aromatic rings that are highly suited to the hydrophobic pocket near the active site of renin. As can be seen from the hydrophobic SCP (Fig. 5c), hydrophobically favorable segments (cyan) are common in these aromatic regions, leading to strong interactions with the vicinal nonpolar residues Val30, Trp39, Tyr75, Pro111, Phe112, Leu114, and Phe117 of renin (Fig. 6c, d). By contrast, the polar and charged segments (such as >C=O, –OH, and $>NH_2^+$) of compound 41 are, as might be anticipated, colored pink, indicating that these groups have an unfavorable hydrophobic effect. This is reasonable because the packing of the polar groups that occurs upon binding is an endothermic process that is accompanied by the loss of water molecules from the surface. This incurs a significant desolvation penalty and is thus energetically unfavorable for binding.

Possible ASSP limitations

In contrast to traditional grid-based QSAR methods such as CoMFA [21], ASSP only considers subregions containing at least one atom of one ligand, so most of the subregions are empty. This feature means that this method overlooks the contributions of the subregions that are only filled with new ligand atoms. Thus, we surveyed the 347 filled regions in the active site of renin, and found that very few (just 12) of these regions were filled with only one ligand atom; indeed, half of them had more than 20 atoms. This implies that the occurrence of substructures of the new ligands in unfilled subregions is a relatively rare phenomenon, and may not have a significant influence on the predicted results of ASSP. In addition, one would expect that ASSP is very sensitive to slight conformational changes in ligands. We therefore examined the effect of the re-superposition of the 45 ligands on the QSAR modeling, and found that there is

essentially no difference between the modeling results obtained using the original and re-superposed ligands. Nevertheless, the ASSP method is currently far from perfect, and we will modify and improve it in future.

Conclusions

Statistical analysis and visual examination of the nonbonding profile between protein receptors and their cognate ligands are effective ways of enhancing our understanding of the molecular basis for and the structural foundations of ligand recognition and binding by host proteins. In this study, we proposed a new method called active-site spatial partitioning (ASSP), which was employed to dissect the nonbonding profile around the active site of renin and associate the substructural properties of bound ligands with the biological activities of the ligands. The investigations revealed that:

(a) The binding behavior of indole-3-carboxamide derivatives to renin is primarily dominated by the electrostatic component and secondarily by hydrophobic forces, whereas the steric aspect only has a relatively mild effect on the binding.

(b) The polar and charged groups (such as >C=O, –OH, and $>NH_2^+$) of ligand molecules are the most important sources of electrostatic attraction and steric penetration, since they can form complicated hydrogen-bond and salt-bridge networks, but they also undergo intense collisions with their surrounding renin residues. In contrast, the ligand's aromatic rings appear to be the main providers of hydrophobic potential and dispersion interactions with the renin receptor.

(c) The fluorine and nonpolar hydrogen atoms were observed to participate in a number of weakly attractive interactions with their surroundings, and can therefore considerably enhance the biological activities of ligands, provided that their presence does not lead to severe steric hindrance and collisions.

References

1. Kearney PM, Whelton M, Reynolds K, Muntner P, Whelton PK, He J (2005) Global burden of hypertension: analysis of worldwide data. Lancet 365:217–223
2. Paul M, Poyan MA, Kreutz R (2006) Physiology of local renin–angiotensin systems. Physiol Rev 86:747–803
3. Wood JM, Stanton JL, Hofbauer KG (1987) Inhibitors of renin as potential therapeutic agents. J Enzyme Inhib Med Chem 1:169–185
4. Dzau VJ, Devine D, Mudgett-Hunter M, Kopelman RI, Barger AC, Haber E (1983) Antibodies as specific renin inhibitors: studies with polyclonal and monoclonal antibodies and F_{ab} fragments. Clin Exp Hypertens A 5:1207–1220
5. Boger J, Payne LS, Perlow DS, Lohr NS, Poe M, Blaine EH, Ulm EH, Schorn TW, LaMont BI, Lin TY (1985) Renin inhibitors. Syntheses of subnanomolar, competitive, transition-state analogue inhibitors containing a novel analogue of statine. J Med Chem 28:1779–1790
6. Gradman A, Schmieder R, Lins R, Nussberger J, Chiang Y, Bedigian M (2005) Aliskiren, a novel orally effective renin inhibitor, provides dose-dependent antihypertensive efficacy and placebo-like tolerability in hypertensive patients. Circulation 111:1012–1018
7. Sealey JE, Laragh JH (2007) Aliskiren, the first renin inhibitor for treating hypertension: reactive renin secretion may limit its effectiveness. Am J Hypertens 20:587–597
8. Corminboeuf O, Bezencon O, Grisostomi C, Remen L, Richard-Bildstein S, Bur D, Prade L, Hess P, Strickner P, Fischli W, Steiner B, Treiber A (2010) Design and optimization of new piperidines as renin inhibitors. Bioorg Med Chem Lett 20:6286–6290
9. Scheiper B, Matter H, Steinhagen H, Stilz U, Bocskei Z, Fleury V, McCort G (2010) Discovery and optimization of a new class of potent and non-chiral indole-3-carboxamide-based renin inhibitors. Bioorg Med Chem Lett 20:6268–6272
10. Politi A, Durdagi S, Moutevelis-Minakakis P, Kokotos G, Mavromoustakos T (2010) Development of accurate binding affinity predictions of novel renin inhibitors through molecular docking studies. J Mol Graph Model 29:425–435
11. Hansch C, Fujita T (1964) ρ-σ-π analysis. A method for the correlation of biological activity and chemical structure. J Am Chem Soc 86:1616–1626
12. Free SM, Wilson JW (1964) A mathematical contribution to structure–activity studies. J Med Chem 7:395–399
13. Balaban AT, Devillers J (2000) Topological indices and related descriptors in QSAR and QSPAR. CRC, Boca Raton
14. Karelson M, Lobanov VS (1996) Quantum-chemical descriptors in QSAR/QSPR studies. Chem Rev 96:1027–1043
15. Verma J, Khedkar VM, Coutinho EC (2010) 3D-QSAR in drug design—a review. Curr Top Med Chem 10:95–115
16. Lowis DR (1997) HQSAR, a new, highly predictive QSAR technique. Tripos Tech Notes 1:1–10
17. Andrade CH, Pasqualoto KFM, Ferreira EI, Hopfinger AJ (2010) 4D-QSAR: perspectives in drug design. Molecules 15:3281–3294
18. Vedani A, Dobler M (2002) 5D-QSAR: the key for simulating induced fit? J Med Chem 45:2139–2149
19. Vedani A, Dobler M, Lill MA (2005) Combining protein modeling and 6D-QSAR. Simulating the binding of structurally diverse ligands to the estrogen receptor. J Med Chem 48:3700–3703
20. Lushington GH, Guo JX, Wang JL (2007) Whither combine? New opportunities for receptor-based QSAR. Curr Med Chem 1:1863–1877
21. Cramer RD, Patterson DE, Bunce JD (1988) Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. J Am Chem Soc 110:5959–5967
22. Klebe G, Abraham U, Mietzner T (1994) Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. J Med Chem 37:4130–4146
23. Robinson DD, Winn PJ, Lyne PD, Richards WG (1999) Self-organizing molecular field analysis: a tool for structure–activity studies. J Med Chem 42:573–583
24. Doweyko AM (2004) 3D-QSAR illusions. J Comput Aided Mol Des 18:587–596
25. Steinhagen H, Scheiper B, Matter H, Stilz HU, McCort G (2009) PCT Int Appl. WO 2009095163
26. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM Jr, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J Am Chem Soc 117:5179–5197

27. Gasteiger J, Marsili M (1980) Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges. Tetrahedron 36:3219–3228
28. Zhou P, Tian F, Li Z (2007) A structure-based, quantitative structure-activity relationship approach for predicting HLA-A*0201-restricted cytotoxic T lymphocyte epitopes. Chem Biol Drug Des 69:56–67
29. Eisenberg D, McLachlan AD (1986) Solvation energy in protein folding and binding. Nature 319:199–203
30. Sanner MF, Olson AJ, Spehner JC (1996) Reduced surface: an efficient way to compute molecular surfaces. Biopolymers 38:305–320
31. He P, Wu W, Wang H, Yang K, Liao K, Zhang W (2010) Toward quantitative characterization of the binding profile between the human amphiphysin-1 SH3 domain and its peptide ligands. Amino Acids 38:1209–1218
32. Wold S, Sjöström M, Eriksson L (2001) PLS regression: a basic tool of chemometrics. Chemom Intell Lab Syst 58:109–130
33. Helland IS (1990) PLS regression and statistical models. Scand J Stat 17:97–114
34. Haenlein M, Kaplan AM (2004) A beginner's guide to partial least squares analysis. Underst Stat 3:283–297
35. Vinzi VE, Chin WW, Henseler J, Wang H (2010) Handbook of partial least squares. Springer, Berlin
36. Golbraikh A, Tropsha A (2002) Beware of q^2! J Mol Graph Model 20:269–276
37. Bissantz C, Kuhn B, Stahl M (2010) A medicinal chemist's guide to molecular interactions. J Med Chem 53:5061–5084
38. Steiner T, Desiraju GR (1998) Distinction between the weak hydrogen bond and the van der Waals interaction. Chem Commun 8:891–892
39. Wallace AC, Laskowski RA, Thornton JM (1995) LIGPLOT: a program to generate schematic diagrams of protein–ligand interactions. Protein Eng 8:127–134