Journal of Computer-Aided Molecular Design, **16:** 935–953, 2002.
KLUWER/ESCOM

A structure-based design approach for the identification of novel inhibitors: application to an alanine racemase

Gabriela Iurcu Mustata & James M. Briggs*
Department of Biology and Biochemistry, University of Houston, Houston, TX 77204-5001, USA

MS received 4 April 2002; Accepted in final form 19 February 2003

Key words: alanine racemase, dynamic pharmacophore model, LigBuilder, CATALYST

Summary

We report a new structure-based strategy for the identification of novel inhibitors. This approach has been applied to *Bacillus stearothermophilus* alanine racemase (AlaR), an enzyme implicated in the biosynthesis of the bacterial cell wall. The enzyme catalyzes the racemization of L- and D-alanine using pyridoxal 5′-phosphate (PLP) as a cofactor. The restriction of AlaR to bacteria and some fungi and the absolute requirement for D-alanine in peptidoglycan biosynthesis make alanine racemase a suitable target for drug design. Unfortunately, known inhibitors of alanine racemase are not specific and inhibit the activity of other PLP-dependent enzymes, leading to neurological and other side effects.

This article describes the development of a receptor-based pharmacophore model for AlaR, taking into account receptor flexibility (i.e. a 'dynamic' pharmacophore model). In order to accomplish this, molecular dynamics (MD) simulations were performed on the full AlaR dimer from *Bacillus stearothermophilus* (PDB entry, 1sft) with a D-alanine molecule in one active site and the non-covalent inhibitor, propionate, in the second active site of this homodimer. The basic strategy followed in this study was to utilize conformations of the protein obtained during MD simulations to generate a dynamic pharmacophore model using the property mapping capability of the LigBuilder program. Compounds from the Available Chemicals Directory that fit the pharmacophore model were identified and have been submitted for experimental testing.

The approach described here can be used as a valuable tool for the design of novel inhibitors of other biomolecular targets.

Abbreviations: MD: molecular dynamics; SD: steepest descent; MWT: molecular weight

Introduction

In recent years, several cases of successful applications of structure-based drug design have been reported [1–4]. Based on a binding site derived pharmacophore model or a pattern of putative interaction sites, the results consist of a collection of virtual ligands complementary to a three-dimensional structure of the binding pocket. In this article we describe a new structure-based design approach for the identification of novel inhibitors that we apply to alanine racemase, a broad-spectrum anti-bacterial drug target. Including receptor flexibility in structure-based inhibitor design strategies has been a substantial problem. This method offers a next step toward achieving this goal.

Alanine racemase (AlaR) is an enzyme universal to bacteria, including mycobacteria, which catalyzes the racemization of L- and D-alanine, and requires pyridoxal 5′-phosphate (PLP) as a cofactor attached to the enzyme via an internal Schiff's base linkage. In the L to D direction, the enzyme provides D-alanine, known to be an essential constituent of the peptidoglycan layer of bacterial cell walls. All of the known bacteria require D-alanine, while only L-alanine is used in eukaryotic protein synthesis. Therefore, alanine racemase has been recognized as a suitable target for drug

*To whom correspondence should be addressed: e-mail: jbriggs@uh.edu

design. A number of effective inhibitors of AlaR have been developed and some of them are of particular interest as agents against *Mycobacterium tuberculosis.* However, most of these compounds are suicide inhibitors that react with the cofactor itself, inhibiting the activity of many PLP containing enzymes due to their lack of target specificity. D-cycloserine (DCS) is the only inhibitor that has been marketed clinically, and is mainly used as a second-line anti-tuberculosis agent. Although it is an excellent inhibitor of tuberculosis growth, as well as of other pathogenic bacterial species, side effects, especially toxicity and neurological disorders have limited its use [5]. Therefore, a structure-based drug design approach focusing on **non-covalent** inhibitors will be of great interest and utility in the development of potent and more specific inhibitors of the enzyme.

Five crystallographic structures of AlaR from *Bacillus stearothermophilus* have been reported so far [6–9], including the complex between AlaR and the competitive inhibitor, propionate [8]. The enzyme is a homodimer, each monomer composed of two domains (Figure 1A). The N-terminal domain is made up of an eight-stranded α/β barrel while the C-terminal domain is mainly composed of β-strands. The PLP cofactor binds in the active site, situated at the C-terminal end of the first β-strand of the α/β domain.

Inspired by the work of Carlson *et al.* (2000) [10], although using a different strategy, we sought to develop a 'dynamic' pharmacophore model for AlaR, which would allow the determination of pharmacophore features that complement the protein's active site taking into account receptor flexibility. A 'dynamic' pharmacophore model is one in which the conformational flexibility of the protein is accounted for in the construction of the model. This is generally accomplished by analyzing small molecule binding sites or property volumes in multiple conformations of the protein to identify conserved pharmacophore elements. The main motivation in developing a dynamic pharmacophore is based on the fact that in the case of HIV-1 integrase, known inhibitors were identified from small molecule databases using a dynamic pharmacophore model, but not with the commonly used static model (i.e. a pharmacophore model based on a single receptor structure) [10]. In fact, the dynamic model achieved a phenomenal inhibitor prediction capacity of 25–33% [10].

The Carlson approach makes use of the BOSS Monte Carlo program [11] to carry out a Multi-Unit Search for Interacting Conformers (MUSIC) via, essentially, a Monte Carlo simulated annealing calculation. In this approach, small molecules or molecular fragments, invisible to one another, are simultaneously docked against a rigid receptor. The docking studies are carried out separately against selected protein conformations taken from a molecular dynamics simulation. The docked molecules are clustered against each protein conformation, resulting in static pharmacophore models. All of the protein structures and corresponding static pharmacophore models are overlain and clustered, resulting in a dynamic pharmacophore model. This study differs from that described by Carlson *et al.* (2000) [10] in the following ways. The very time consuming explicit molecule docking (MUSIC) studies have been replaced by a very rapid active site property mapping procedure. Besides being much faster, this approach has an additional advantage, which is that one can easily generate a heterogeneous pharmacophore model (i.e. one consisting of conserved elements sampled by a variety of probes, such as, hydrogen-bond acceptor or donor, positive or negative charge, and hydrophobic). The MUSIC approach can accomplish the same task but at much greater expense since separate docking studies are required for each type or small molecule probe. Additionally, in our approach, a formal statistical procedure is used for clustering and selecting significant clusters. Finally, we make use of Lipinski's *rule of 5* [12] to aid in filtering large lists of compounds into smaller ones that contain molecules that should have better 'drug-like' properties, as described later.

In order to develop a dynamic pharmacophore model for AlaR, molecular dynamics (MD) simulations were performed on the full length AlaR dimer from *Bacillus stearothermophilus* with a D-alanine molecule in one active site and the non-covalent inhibitor, propionate, in the second site. Different conformations of the protein, taken at 100 ps intervals from our molecular dynamics simulation were evaluated and superimposed, followed by the identification of the conserved pharmacophore elements using the property mapping feature of the LigBuilder program [13]. The advantage offered by our approach, compared to the one employed by Carlson *et al.* [10] is that, the generation of the 'dynamic' pharmacophore model, once the molecular dynamics simulation on the biological system has been performed, takes only a couple of minutes, as compared to many tens of hours of computer time for the original dynamic pharmacophore method.

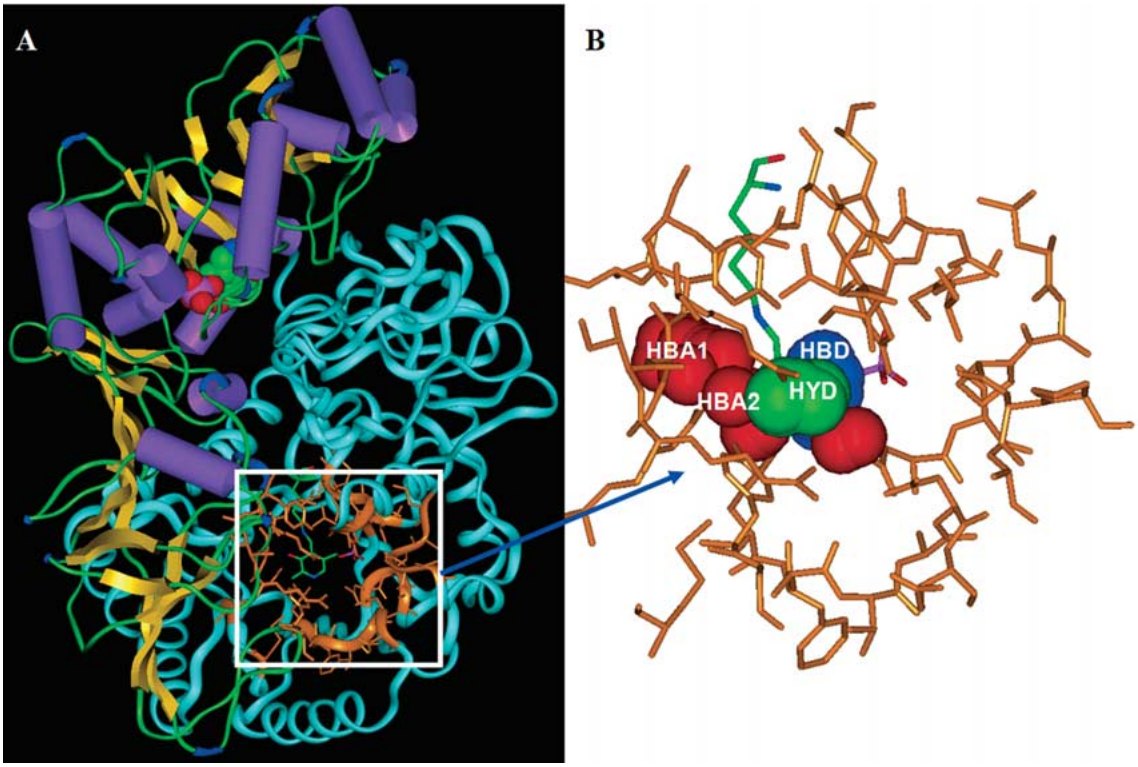

Figure 1. A) Pictorial view of the alanine racemase dimer. The first monomer is displayed in ribbon and colored in turquoise; the binding site region is indicated by a white box and residues are displayed in orange; the PLP cofactor is displayed in stick form and colored by atom. The second monomer is rendered following the Kabsch-Sander classification; the PLP cofactor is rendered in CPK. B) Superposition of the static pharmacophore models generated at 100 ps intervals (0 ps–1000 ps) with *hydrogen-bond donor/+ charge* grid points in blue (HBD), *hydrogen-bond acceptor/- charge* grid points in red (HBA1 and HBA2), and *hydrophobic* grid points in green (HYD). Residues of the AlaR binding site are the same as in Figure 1A within the white box and are displayed in stick form and colored in orange; the PLP cofactor is colored by atom.

Resulting dynamic pharmacophore models were used to search the Available Chemicals Directory (ACD) [14] using the CATALYST program [15], and have yielded a set of compounds satisfying the pharmacophore features. Many of these compounds are being assayed for activity.

Methods

The computational approach we describe here is based on the development of a 'dynamic' pharmacophore model for alanine racemase, which allows for the identification of pharmacophore features that complement the protein's active site while taking into account receptor flexibility. First, the protein of interest, in our case AlaR, is subjected a molecular dynamics simulation in order understand the dynamical behavior of the active site and to generate an ensemble of protein conformations that will be used in the dynamic pharmacophore development. The protein conformations are individually subjected to property mapping studies resulting in static pharmacophore models. The collection of the static pharmacophore models derived from each conformation/snapshot are then superimposed and statistically clustered resulting in a 'dynamic' pharmacophore model. The dynamic pharmacophore model is then used to search databases of commercially available compounds that match the pharmacophore model. A series of rules are used to filter the hits from the database searches.

System setup and MD simulation

The protein structure used for molecular dynamics studies corresponds to *Bacillus stearothermophilus* alanine racemase (Figure 1A) (pdb code: 1SFT) [6]. Considering the fact that the active sites are far apart (ca. 30 Å) [6] and that the active sites have been shown not to interact (i.e. there is no cooperativity between

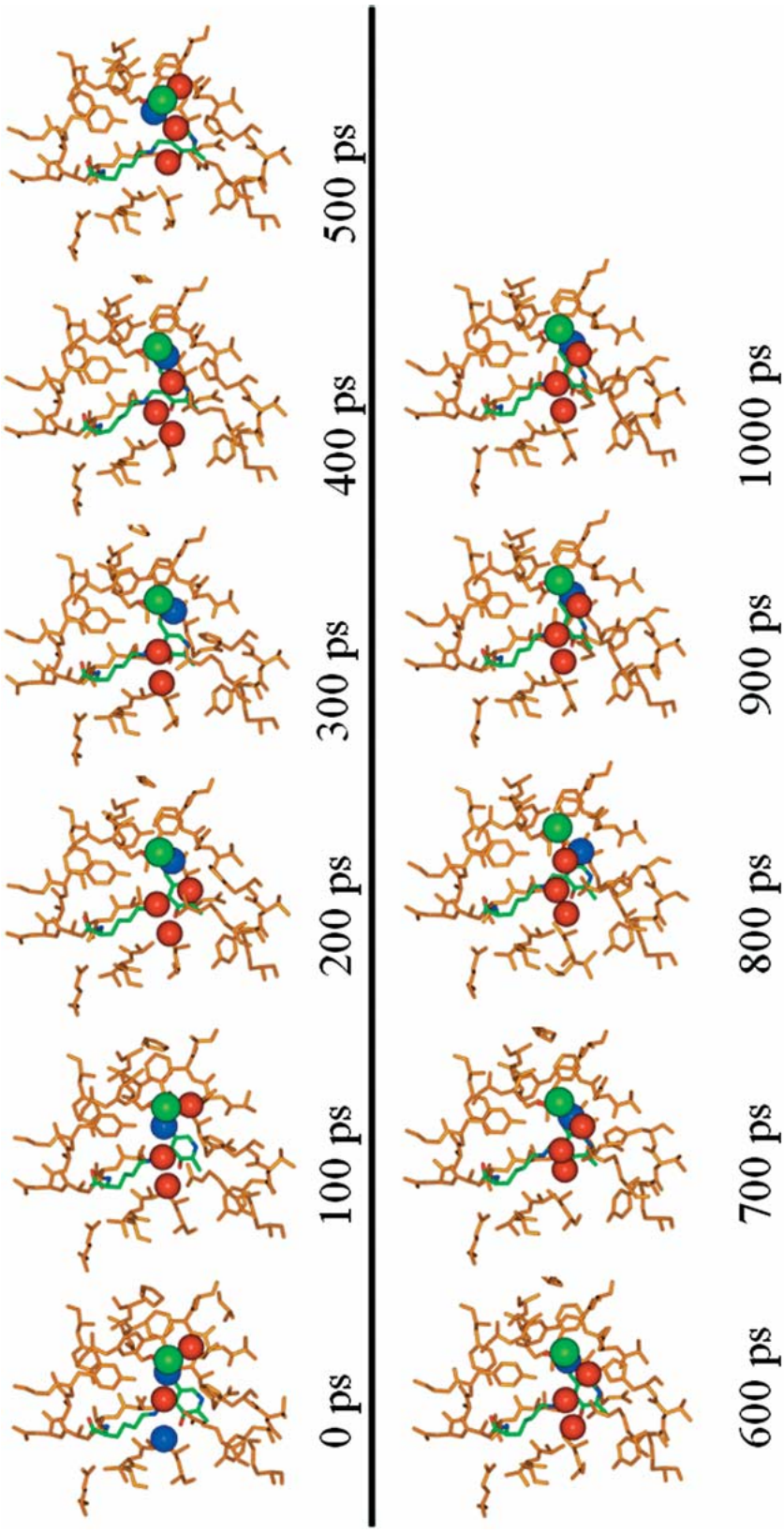

Figure 2. Static pharmacophore models generated at 100 ps intervals (0 ps–1000 ps) with *hydrogen-bond donor/+ charge* grid points in blue (HBD), *hydrogen-bond acceptor/- charge* grid points in red (HBA1 and HBA2), and *hydrophobic* grid points in green (HYD). Residues of the AlaR binding site are in the same orientation as in Fig. 1A and 1B; the PLP cofactor is colored by atom and displayed in stick form.

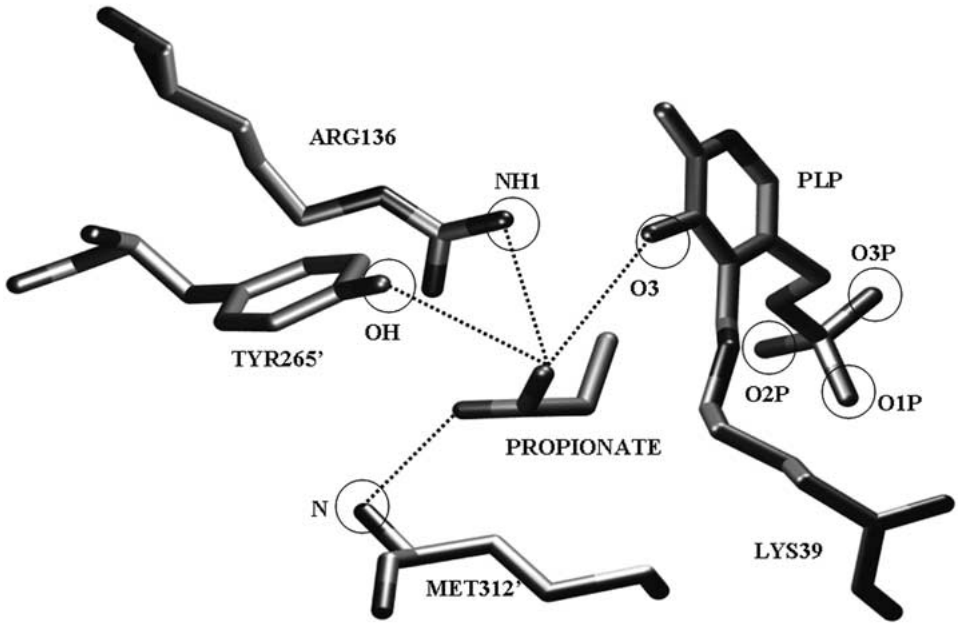

Figure 3. Schematic diagram of interactions made by the carboxylate group of the propionate inhibitor [8]. The residues Tyr265′ and Met312′ belong to the second monomer of the enzyme and are part of the active site.

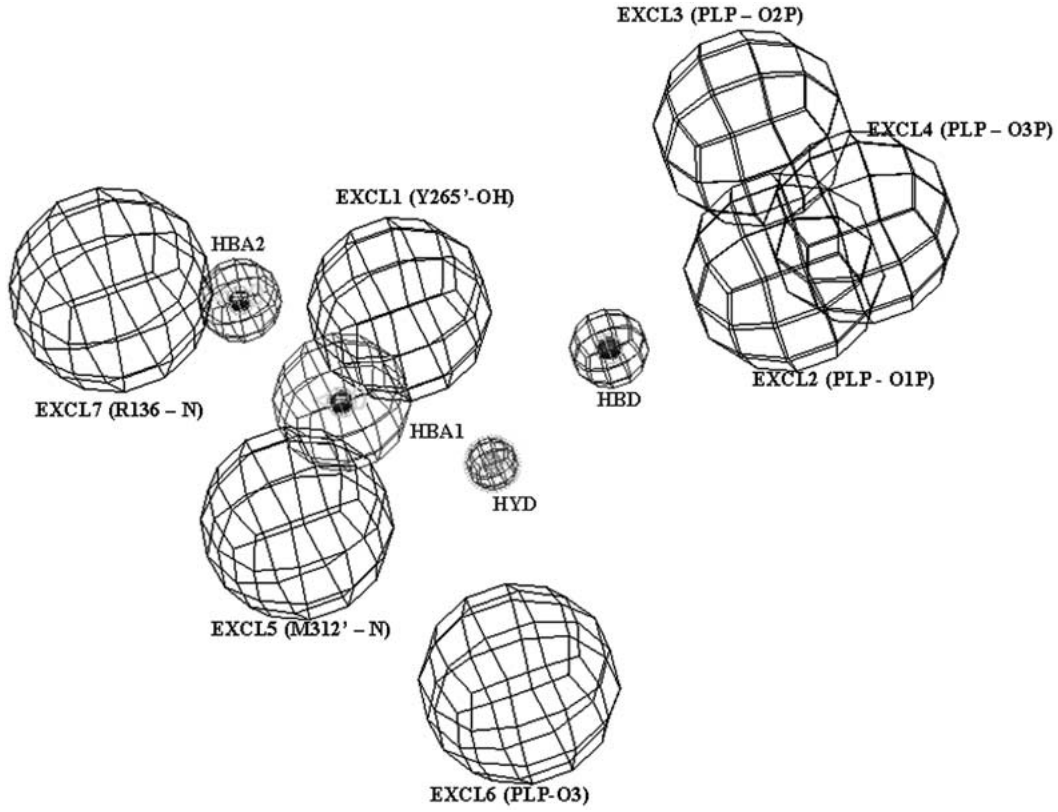

Figure 4. Characteristics of the dynamic feature-based pharmacophore model generated from the 11 protein conformations of the 1 ns MD simulation. The mesh spheres illustrate the location constraints of each feature in approximately the same orientation as in Figure 1B. The black spheres, labeled '**EXCL**', represent excluded volumes added to eliminate compounds with steric conflicts with the receptor.

the two sites) [16] we performed an MD simulation on the full dimer with the D-Ala substrate in one active site and propionate in the second one. In this way we could obtain twice the amount of information from a single simulation (i.e. that for an active site with an inhibitor and one with a substrate). A detailed description and analysis of how the active site region is affected by the presence of the substrate vs. inhibitor is given elsewhere (G. Iurcu Mustata *et al.*, in press). Propionate was modeled into the active site making use of the AlaR-propionate crystal structure (pdb code: 2SFP) [8], while D-Ala was modeled making use of the AlaR-1-aminoethylphosphonic acid complex structure (pdb code: 1BD0) [7], taking into account the predicted binding site of the D-Ala substrate suggested by Ringe *et al.* (1998).

Ab initio calculations were carried out in order to obtain starting force field parameters for the pyridoxal phosphate cofactor in its free and lysine cross-linked forms. The bonded parameters assigned to the pyridoxal phosphate molecule were derived from pyridine rings and similar functional groups from the CHARMM 22 force field [17]. The initial geometry was optimized at the MP2 level and atomic charges were calculated by fitting to the electrostatic potential obtained from an HF/6-31G* *ab initio* calculation with Gaussian98 [18]. The lysine to which the PLP is covalently bonded was defined as a regular residue, while the pyridine ring as well as the methyl phosphate group of the PLP was defined as two independent patches.

Before the MD simulation was started, protonation states were predicted for all of the ionizable residues in the protein using the pK_a [19] procedure implemented in the UHBD program [20], to determine whether any of the residues were likely to adopt nonstandard ionization states. Ondrechen *et al.* (2001) reported some unusual charge states for certain residues in presence or absence of the cofactor [21]. However, on the basis of the results of these calculations in the presence of crystal waters and with or without acetate, these residues were found to be close to their default ionization states and were modeled accordingly.

The system was prepared and energy minimized using the CHARMM program [22]. The thermalization, equilibration, and the subsequent production phases were carried out using the NAMD program [23], in the NPT ensemble (pressure = 1 bar) with explicit solvent and periodic boundary conditions, and a dielectric constant (ε) of 1. A constant temperature of 300 K was maintained by coupling to an external bath [24]. The protein and solvent interact via the CHARMM 22 force field where all protein atoms are explicitly represented [17] and where water is characterized by the TIP3 model [25]. The hydrogen atoms were added using the HBUILD routine in CHARMM. Bonds to hydrogen atoms were constrained using the SHAKE algorithm [26] and the equations of motion were solved using the Verlet algorithm [27]. The simulation consisted of a system defined by the protein (762 residues; 12 137 atoms) plus the crystal waters (206 HOH; 618 atoms) and 15 943 water molecules in a periodic box with dimensions $76 \times 106 \times 76$ Å. The total size of the system is 60 547 atoms. Long-range interactions were smoothly truncated at 10.5 Å with a shifting function for the electrostatic interaction and a switching function for the van der Waals interactions. A switching function was applied in the range from 9.5 to 10.5 Å, in order to smoothly reduce the van der Waals potential to zero at the cutoff distance.

The system was energy minimized with CHARMM with the following protocol. The whole system was constrained and subjected to 500 steps of steepest descent (SD) energy minimization. The solvent plus the substrate and the inhibitor were then relaxed and also subjected to 500 steps of SD energy minimization. Then the whole system was relaxed and minimized using SD for another 1000 steps. The entire system was then gradually heated, by means of temperature reassignment, from 50 K to 300 K for 60 ps. The whole system was then equilibrated for 60 ps of molecular dynamics at 300 K using a 2 fs time step, at which point the system achieved stability. The subsequent production phase was performed in the NPT ensemble for 1 ns with a time step of 2 fs. Analyses were performed on conformations collected every 0.1 ps.

Alanine racemase structure-based inhibitor design with LigBuilder

The POCKET module of LigBuilder [13] was applied to alanine racemase from which a binding-site-based pharmacophore model was derived. The POCKET module defines a box which covers the ligand (in our case, the propionate) and all neighboring residues from which it creates regular-spaced grid points within the box. Next, it places a hydrogen atom probe on each grid point to check for its accessibility and, finally, it derives key interaction sites within the binding pocket. Three types of probe atoms are used to screen the binding pocket; A positively charged sp^3 nitrogen atom (ammonium cation), representing a hydrogen-

bond donor/positive charge; a negatively charged sp^2 oxygen atom (carboxyl oxygen atom), representing a hydrogen-bond acceptor/negative charge; and finally an sp^3 carbon atom (methane), serving as a hydrophobic group. The methodology of how the pharmacophore model is generated is well described by Wang *et al.* (2000) [13], therefore we will not describe it here. However, we would like to emphasize that the generated pharmacophore model is a receptor-derived pharmacophore model that gives the pharmacophore elements of the candidate ligands in response to the protein's active site environment.

This procedure was applied sequentially to 11 different conformations of the protein extracted from our 1 ns MD trajectory. From each conformation, a pharmacophore model was generated, each of which we call 'static' pharmacophore models.

Dynamic pharmacophore development – clustering procedure

The FASTCLUS procedure from the SAS Statistical Package [28] was used to cluster the pharmacophore features derived from 11 conformations of the protein taken from snapshots of the MD simulation at intervals of 100 ps (e.g. from 0, 100, ..., 1000 ps). This clustering procedure (a partitional clustering method) finds disjoint clusters of observations using the *K*-means method applied to coordinate data (in our case each pharmacophore feature is characterized by the Cartesian coordinates). The result of a partitional clustering method generally consists of a set of clusters, each object belonging to one cluster. Each cluster is represented by a centroid, or a cluster representative that contains a summary description of all the objects contained in a cluster.

The criterion function is the average squared distance of the data items x_k from their nearest cluster centroids,

$$E_k = \sum_k \left\| x_k - m_{c(x_k)} \right\|^2$$

Where $c(x_k)$ is the index of the centroid that is closest to x_k. Minimizing the cost function begins by initializing a set of k cluster centroids denoted by m_i, $i = 1, \ldots, k$. The positions of the m_i are then adjusted iteratively by first assigning the data samples to the nearest clusters and then recomputing the centroids. The iteration is stopped when E no longer undergoes significant changes [28].

The main advantage of this method is its simplicity and rapidity but also the detailed information provided: cluster summary (frequency, RMS standard deviation, nearest cluster, distance between cluster centroids, cluster mean coordinates etc.). The result of clustering the 11 'static' pharmacophore models is what we name here, a 'dynamic' pharmacophore model.

CATALYST methodology

We used the 'dynamic' AlaR-receptor-derived pharmacophore model as the query structure to perform 3D database searching. This approach was used in order to find novel ligand molecules that fit to our flexible target protein, alanine racemase, as represented by the dynamic pharmacophore model.

The mean Cartesian coordinates of the clustered pharmacophore features were imported into CATALYST [15] and a 3D hypothesis was generated. For each feature, a radius was assigned in accordance with the RMS standard deviations across variables of the cluster. Each feature was associated with a weight of 1 (a measure of its proposed importance to the pharmacophore as a whole). The dynamic pharmacophore model was used to screen the Available Chemicals Directory [14] in order to identify molecules that fit the model.

Conformations of all training set molecules were generated using the 'best quality' option of CATALYST using an energy threshold of 20 kcal/mol from the local minimized structures. A maximum of 250 conformations of each molecule were generated to ensure maximum coverage of the conformational space, while conserving computer time.

The 'Best Compare/Fit' procedure was used next in order to evaluate the geometry of the conformers of the compounds to test how well they fit the hypothesis (i.e., the dynamic pharmacophore model). The fit value represents the quality of the mapping between the compound and the hypothesis; the better the overall superimposition of functional groups of the molecule to the appropriate features of the pharmacophore, the higher the score of the fit [29]. The fit value, as implemented in CATALYST, is computed as follows:

$$\text{Fit} = \sum \text{mapped features } f \text{ of weight} W(f) \times [1 - SSE(f)]$$

Where:
$SSE(f)$ = sum over location constraints c on feature f of $(D/T)^2$

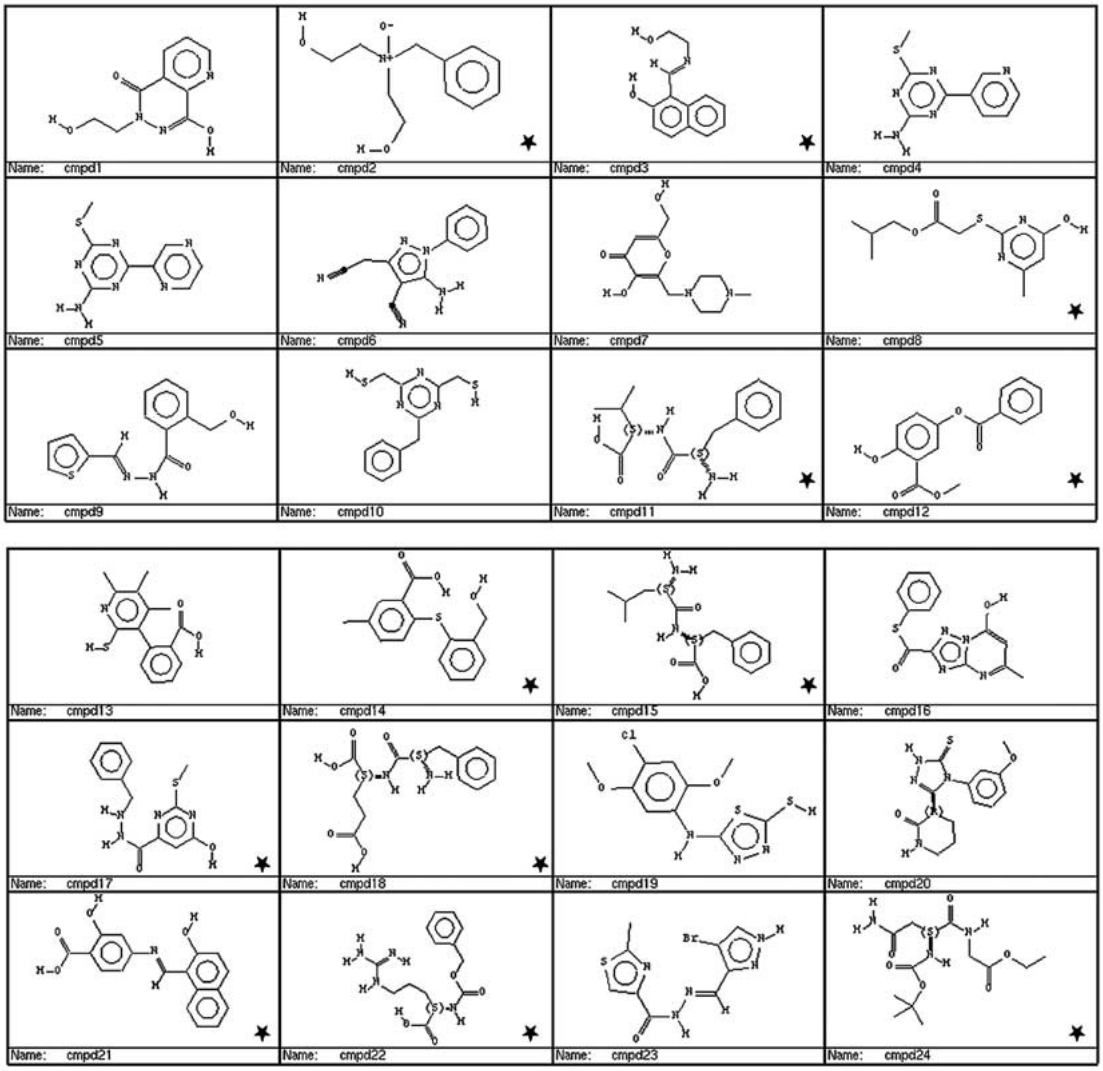

Figure 5. Structures of the 34 compounds, out of 43 identified from the ACD screening against the dynamic pharmacophore model. Compounds are drawn in the fashion used in CATALYST and groups are shown in their natural state as they appear in the ACD. The compounds labeled with a black star will be tested for AlaR inhibition. The compound names, as given in the ACD, are: cmpd1: MAYBRIDGE S 15408, cmpd2: benzyl-di-β-hydroxy-ethylamine-oxide, cmpd3: 1-(2-hydroxyethyl) iminomethyl-2-naphtol, cmpd4: MAYBRIDGE RF 02731, cmpd5: MAYBRIDGE KM 05093, cmpd6: 5-amino-4-cyano-1-phenyl-3-pyrazoleacetonitrile, cmpd7: 6-(4-methyl-piperazinomethyl)kojic acid, cmpd8: (4-hydroxy-6-methyl-pyrimidin)-2-ylsulfanyl-acetic-acid-isobutyl-ester, cmpd9: MAYBRIDGE JPV 00053, cmpd10: 2-benzyl-4,6-dimercapto-methyl-1,3,5-triazine, cmpd11: L-phenylalanyl-L-valine, cmpd12: methyl-5(benzoyloxy)salicylate, cmpd13: 2-(2-mercapto-4,5,6-trimethyl-3-pyridyl benzoic acid), cmpd14: MAYBRIDGE RJC 01860, cmpd15: D,L-leucyl-D,L-phenylalanine, cmpd16: *S*-phenyl-7-hydroxy-5-methyl-1,3,4-triazaindolizine-2-thiocarboxylate, cmpd17: 6-hydroxy-2-methylsulfanyl-pyrimidine-4-carboxylic acid-*N*-benzyl-hydrazide, cmpd18: H-Phe-Glu-OH, cmpd19: MAYBRIDGE S 03879, cmpd20: MAYBRIDGE BTB 11516, cmpd21: 4-(2-hydroxy-1-naphthylmethyleneamino-salicylic acid, cmpd22: Z-Arg-OH, cmpd23: MAYBRIDGE SPB 02869, cmpd24: BOC-ASN-GLY-OET, cmpd25: toluene-4-sulfonic acid-7-ethyl-3-oxo-2-aza-bicyclo-6-yl-ester, cmpd26: D,L-mandelic-acid-hemimagnesium salt, cmpd27: HO-Me-sulfanyl-pyrimidine-4-carboxylic acid (Di, Me aminobenzylidene)-hydrazide, cmpd28: MAYBRIDGE BTB 11516, cmpd29: MAYBRIDGE JP 00758, cmpd30: MAYBRIDGE RDR 03349, cmpd31: diethyl-3-oxo-2-(3-trifluoromethylphenyl-hydrazone-glutarate), cmpd32: Not disclosed, cmpd33: SPECS CIF3907, cmpd34: 4-(2,4-dichlorobenzyl-oxy-amino-methyl-phthalazin)-1-one.

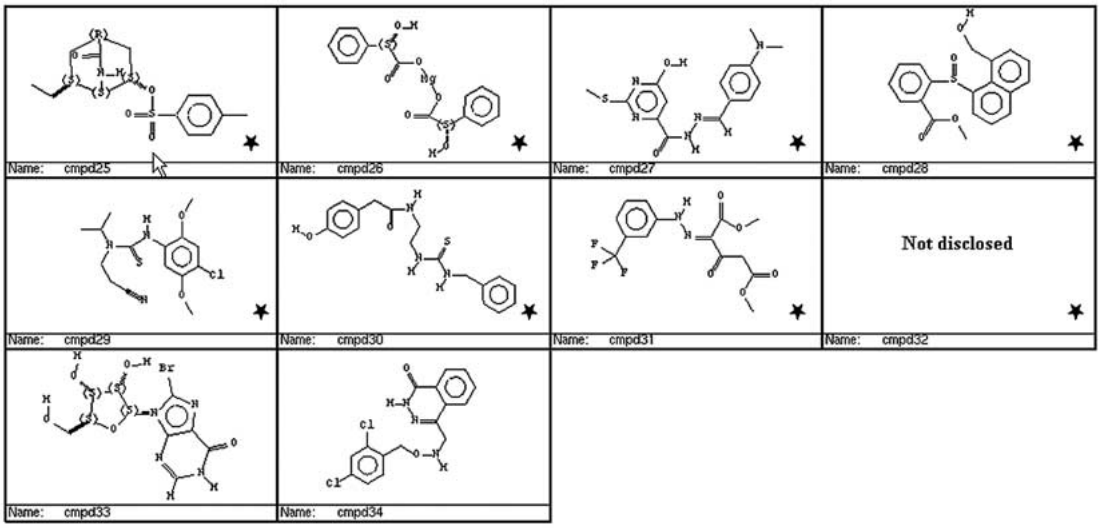

Figure 5. Continued.

W = the weight of the feature
D = the displacement of the feature from the center of the location constraint
T = the radius of the location constraint sphere for the feature (tolerance) [15].

Descriptors

There is a wide range of choices available for describing molecules. In order to rank the hits obtained from the ACD search we chose a small set of 1D descriptors, namely, molecular weight (MWT), number of hydrogen-bond donors (expressed as the sum of OHs and NHs), number of hydrogen-bond acceptors (expressed as the sum of oxygens and nitrogens), number of rotable bonds, and $\log P_{o/w}$ (o/w = octanol/water). $\text{Log}P_{o/w}$ was calculated using the XLOGP v2.0 [30] program, which calculates the $\log P_{o/w}$ value by summing the contributions of component atoms and correction factors derived from regression analyses of a large number of organic compounds with known experimental $\log P_{o/w}$ values.

Results and discussion

The 3D structure of *Bacillus stearothermophilus* alanine racemase (Figure 1A) (pdb code: 1SFT) [6] was used as the starting point for this study. In an effort to develop a 'dynamic' pharmacophore model for AlaR, a molecular dynamics simulation was performed on the full AlaR dimer with a D-alanine molecule in one active site and the non-covalent inhibitor propionate in the second active site of this homodimer. A detailed description and analysis of the 1 ns MD trajectory is given elsewhere (G. Iurcu Mustata *et al.*, in press). The strategy followed in this study was to use conformations of the protein from an MD simulation and develop a dynamic pharmacophore model using LigBuilder [13].

Dynamic pharmacophore model generation

We used 11 snapshots, taken at 100 ps intervals from our MD simulation, from which all solvent molecules were removed. The PLP cofactor was present in all cases, since its covalent linkage to the active site residue Lys39 is required for the proper functioning of the enzyme. We have introduced an approximation in our study by choosing snapshots from the MD trajectory (i.e. taking snapshots every 100 ps). However, this approximation was motivated by the fact that only small conformational changes take place in the propionate-binding region during the MD simulation. The fluctuations of the atoms in the active site region were analyzed from our 1 ns MD trajectory where the average atomic fluctuations averaged over all C_α atoms of the propionate binding region was found to be 0.7 Å (G. Iurcu Mustata *et al.*, in press). In addition to this, in order to gain another measure of the dynamics of the propionate binding region, distances across the active site have been calculated (G. Iurcu Mustata *et al.*, in press). These distances were monitored for the entire length of the simulation and showed no

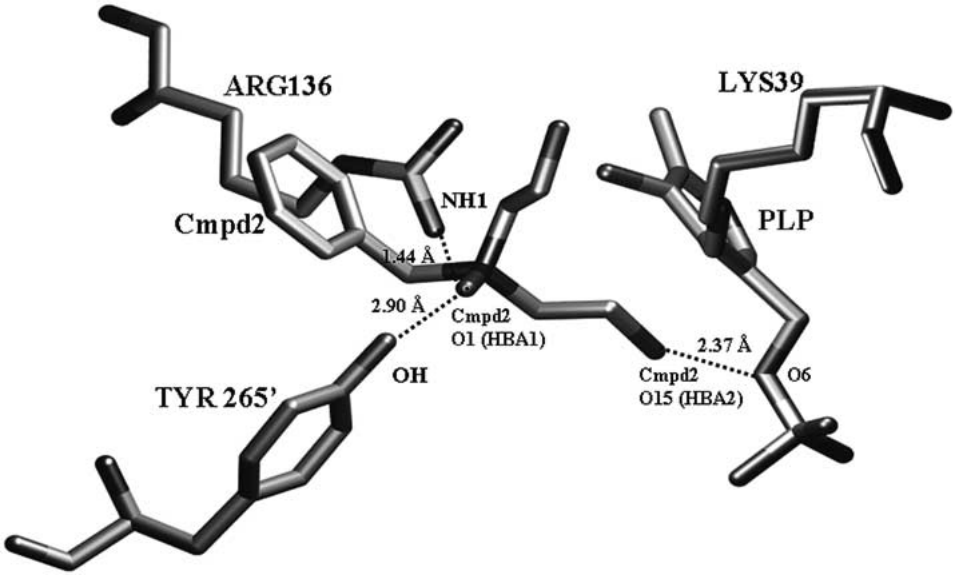

Figure 6. Mapping of the compound 2 (Cmpd2) onto the 11-point pharmacophore model. The pharmacophore features are labeled as HBD for hydrogen-bond donor, HBA1 and HBA2 for hydrogen-bond acceptor, and HYD for hydrophobic.

significant changes with the exception of the distance between Val 37 and Tyr 265′, which define the width of the active site. This change is mainly due to the side chain flexibility of Tyr 265′. No significant changes were observed for the other two distances. This reveals that changes occur in the active site region with the bound propionate. Since there was also no significant difference in the atomic fluctuations at for example 50 ps as compared to 100 ps, or 150 ps as compared to 200 ps, 11 snapshots were considered to be sufficient to sample the flexibility of the active site region. Finally, examination of the 11 static pharmacophore models (see Figure 2) reveals that selection of just one static model is likely to provide an insufficient description of the active site (e.g. snapshot 0 ps, the one that is most likely to be used by most researchers, is not representative of the rest). More rational approaches for the selection of relevant protein conformations and simpler methods for generating them are being developed in our group.

The LigBuilder software was used to generate pharmacophore models derived from 11 snapshots of the protein. Our MD studies indicate that the active site becomes more stabilized in the presence of the substrate, D-Ala, vs. the non-covalent inhibitor, propionate (G. Iurcu Mustata *et al.*, in press). The active site residues with the bound propionate are more flexible and the active site itself is larger as compared to the D-Ala bound active site. Consequently, focusing on the active site with the bound inhibitor (i.e. propionate) was considered to be more appropriate from a structure-based drug design point of view, since we were interested in taking into account protein flexibility upon inhibitor binding. All 11 conformations of the protein with the resulting 11 'static' pharmacophore models were superimposed using the backbone least-squares superposition in InsightII [31] to transform the coordinates into the same reference frame. Figure 1B shows the superposition of the static pharmacophore models generated at 100 ps intervals from our MD simulation. From the figure we can distinguish four clusters of which 2 represent hydrogen-bond acceptor features, one hydrogen-bond donor and one hydrophobic site. A fifth, smaller cluster, representing a hydrogen-bond acceptor feature, can be seen but, because its presence was observed just in three out of the 11 conformations of the protein, we did not consider this cluster to be statistically significant and we discarded it from our study.

On the basis of the Cartesian coordinates of the 11 superimposed pharmacophore models, we used the FASTCLUS procedure implemented in the SAS package [28] to analyze the conservation of the pharmacophore features in the 11 conformations of the protein. The advantage offered by using a statistical method in clustering our data is that we avoid the ambiguity that could have resulted through manual selection of sites by viewing of the overlaid structures.

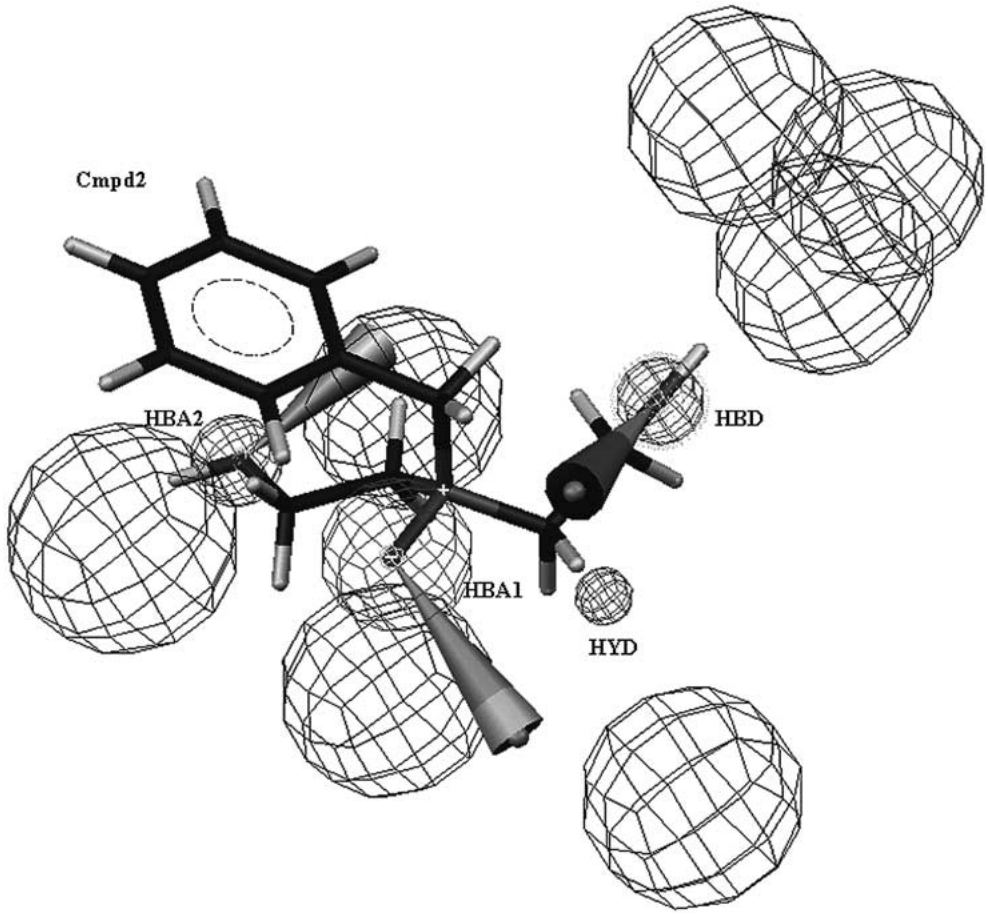

Figure 7. Modeled structure of compound 2 (Cmpd2) bound to the AlaR active site. Dashed lines indicate hydrogen-bonding to the inhibitor; the mapped pharmacophore features are indicated in parenthesis.

The 11 static pharmacophore models clustered using SAS [28] resulted in what we call a 'dynamic' pharmacophore model. This model indirectly takes into account the conformational flexibility of the protein under the influence of explicit solvent molecules and an inhibitor.

The crystal structure of the AlaR/propionate complex revealed that important interactions exist between the carboxylate group of the inhibitor and residues of the active site [8] (Figure 3). Taking into account these interactions as well as the positioning of the PLP in the active site, a 4-point pharmacophore model was supplemented with 7 excluded volume-sites (the circled atoms on Figure 4). These volume exclusions were added in order to prevent the identification of inhibitor leads that would fit the pharmacophore elements well but overlap with receptor atoms (Figure 4).

In the case of the excluded volumes (spaces which inhibitor atoms are not allowed to occupy), the average positions (computed from the 11 conformations of the protein) of the circled atoms in the Figure 3 were used as the geometric centers for the seven excluded volumes. The radii of the excluded volumes were set to 1.5 Å, approximately the van der Waals radii of the corresponding atoms (i.e. 1.52 for O and 1.55 for N [32]). The rationale for making this choice was because we wanted to take into account not only protein flexibility but also protein flexibility upon the binding of a candidate ligand. By using a 1.5 Å radii for the excluded volumes, we increased the effective size of the binding site cavity in order to partially compensate for the flexibility of both the receptor protein and the ligand.

This resulted in an 11-point dynamic pharmacophore model: 7 excluded volumes, 2 hydrogen-bond acceptors, 1 hydrogen-bond donor, and 1 hydrophobic feature (Figure 4). The characteristics of the dynamic

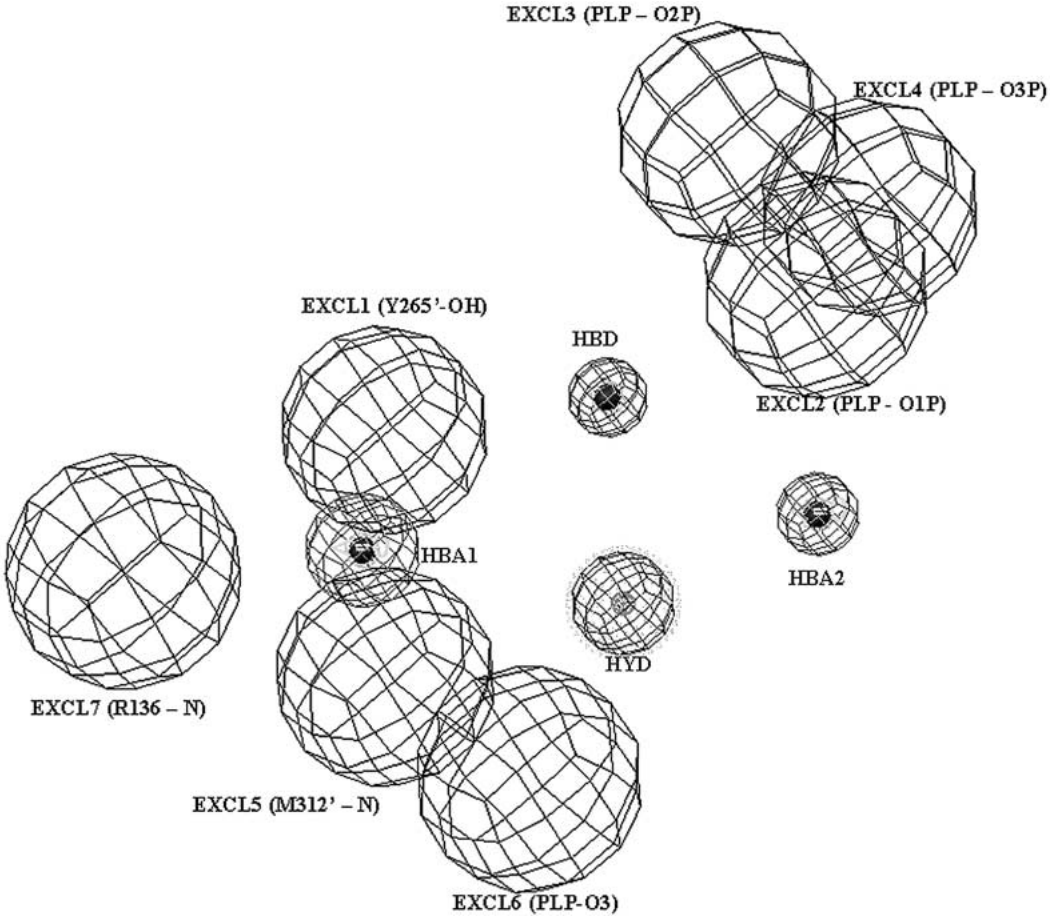

Figure 8. Characteristics of the static pharmacophore model generated from the crystal structure. The mesh spheres illustrate the location constraints of the static pharmacophore features. The black spheres, labeled '**EXCL**', represent excluded volumes added to eliminate compounds with potential steric conflicts with the receptor.

pharmacophore model are given in the Supporting Information.

Database searching and identification of potentially inhibitory compounds

The radii of the pharmacophore elements were set to 0.75R, 1R, 1.5R and 2R (R = RMS standard deviations across variables of the cluster) and dynamic pharmacophore models, with/without excluded volumes, were used to screen the Available Chemicals Directory (ACD). The spherical-excluded volumes were also scaled to 25%, 50% or 100% of their respective van der Waals (vdW) radii. As one would expect, the number of hits was modulated by the vdW scaling factor used on the excluded volume sites (the smaller the radius, the larger the number of hits – data not shown). Due to the impressive number of hits (see Table 1) we retained only those found **conserved** among the hits given by the ACD search using a different radius for the pharmacophore elements and with or without excluded volumes using a 1.5 Å radius.

Since the molecular weight of compounds increases by 1–200 during lead optimization, and corroborated by the fact that very few compounds entering Phase II clinical trials have a MWT of over 500 [12], a subset of the total hits with a MWT less than 350 narrowed our data set to 43 molecules.

A representative family of conformations was generated for each molecule using the Poling algorithm [33] and the 'best conformational analysis method' implemented in the CATALYST program. Conformations were selected within a 20 kcal/mol range above the lowest energy conformation identified. The 'Best Compare/Fit' procedure was used next in order to test how well the 43 compounds fit the pharmacophore

Table 1. Total number of hits obtained from the ACD screening against the dynamic pharmacophore model with or without excluded volumes.

	ACD HITS		
	Radius[a]	Total No. Hits	No. Hits (MWT < 350)
No excluded volumes	0.75R	5 438	1 085
	1R	12 704	309
	1.5R	37 088	2 867
	2R	58 500	25 430
Excluded volumes (R = 1.5 Å)	0.75R	364	142
	1R	1 630	794
	1.5R	10 450	5 454
	2R	26 828	14 700

[a]Radius of the pharmacophore feature represents the RMS standard deviation across cluster variables (see Methods).

model such that each successful molecule had to fit at least 3 of the 4 pharmacophore elements in the dynamic pharmacophore model. As a result, only 34 compounds out of 43 met the required criteria as shown in Figure 5. The fit values of the molecules against the dynamic pharmacophore model with different radii of the pharmacophore elements are given in the Supporting Information. In general, the fit value was higher as the radii of the pharmacophore elements were increased in size. However, no changes in the fit value were observed when the excluded volumes were scaled to 25%, 50% or 100% of their vdW radii. This is not surprising since these molecules resulted from the ACD screening with a 1.5 Å radius of the excluded volumes.

In Figure 6 we show the mapping of one of the compounds (Cmpd2) onto the 11-point pharmacophore model. We can see that this compound mapped all pharmacophore features, except the hydrophobic one. Interestingly, by importing this compound into the AlaR active site and by superimposition with the dynamic pharmacophore model, we can see that complementarity between ligand-receptor exist through hydrogen-bonding (see Figure 7); TYR265′ (OH) and PLP (OH) act as hydrogen-bond partners for O1 and O2 of compound 2. These two former atoms mapped the hydrogen-bond acceptor features of the dynamic pharmacophore model (see Figure 6).

Table 2. Characteristics of the 34 compounds in terms 1D descriptors.

Compound	MWT[a]	Rotlbonds[b]	Hbond acceptor[c]	Hbond donor[d]	$logP^{e}_{o/w}$
1	207.19	4	6	2	−0.09
2	211.26	8	4	2	−0.62
3	215.25	5	3	2	2.08
4	219.26	2	5	2	0.03
5	220.25	2	6	2	−1.12
6	223.23	4	5	2	0.43
7	254.28	5	6	2	−0.73
8	256.31	7	5	1	0.99
9	260.31	6	4	2	3.13
10	263.37	6	3	0	0.69
11	264.32	8	5	4	1.4
12	272.26	6	5	1	4.04
13	273.35	4	3	1	3.57
14	274.33	6	3	2	3.46
15	278.35	9	5	4	2.51
16	286.31	4	6	1	2.18
17	290.34	7	6	3	−0.02
18	294.31	11	7	5	1.07
19	303.78	5	5	1	0.22
20	304.36	3	5	2	1.98
21	307.31	6	5	3	2.49
22[f]	308.34	12	8	6[f]	4.22
23	314.16	4	6	2	1.66
24	317.34	12	9	4	1.69
25	323.41	4	5	1	0.24
26	326.58	10	6	2	1.07
27	331.39	7	7	2	2.8
28	340.39	6	4	1	2.58
29	341.85	9	5	1	3.24
30	343.44	11	5	4	2.72
31	346.26	8	7	1	4.35
32	346.39	10	8	1	1.46
33	347.12	5	9	4	−2.49
34[f]	350.2	5	5	2	5.09[f]

[a]Molecular weight; [b]Number of rotatable bonds; [c]Hydrogen-bond acceptors (expressed as the sum of nitrogens and oxygens); [d]Hydrogen-bond donors (expressed as the sum of OHs and NHs); [e]$LogP_{o/w}$ values as calculated by XLOGP 2.0 [30] [f]The significance of the gray boxes indicates a violation of Lipinski's rule [12].

Compound selection

In recent years, more and more attention has been paid to predicting the bioavailability of a molecule. Important criteria for 'drug-like' status is the Lipinski '*rule of 5*' [12] which states that: poor absorption or permeation are more likely to occur when there are more than 5 hydrogen-bond donors (expressed as the sum of OHs and NHs); the molecular weight is more than 500, the $\log P_{o/w}$ is more than 5, or there are more than 10 hydrogen-bond acceptors (expressed as the sum of nitrogens and oxygens). In addition to this, $\log P_{o/w}$, the partition coefficient for octanol/water solubility, has been shown to correlate with bioavailability [34]. Therefore, $\log P_{o/w}$ values are very important when we choose from a large number of candidate molecules for screening during drug design and discovery.

We have included the molecular weight, number of hydrogen-bond acceptors, number of hydrogen-bond donors, number of rotable bonds and $\log P_{o/w}$ as 1D ('one-dimensional') descriptors for 'rank-ordering' our candidates. The $\log P_{o/w}$ was computed using the XLOGP 2.0 program [30]. Compounds 22 and 34 were the only ones to violate Lipinski's rule (the number of hydrogen-bond donors was greater than 5, the $\log P_{o/w}$ value was greater than 5) (Table 2).

Finally, we chose to select from these compounds those that were most likely to be safe drugs, by identifying and avoiding structural elements with a high risk of metabolic toxicity. Examples include electron-rich π-systems, quinones, quinonimines, aromatic methylene-dioxy groups, aromatic nitro groups, primary aromatic amines and sulfhydryl groups [35].

A set of 19 out of 43 molecules has been finally selected (marked with a black star in Figure 5). Compound 22 was eliminated from our data set because of the Lipinski's rule violation. In the case of the last compound, even if its predicted $\log P_{o/w}$ value was to some extent a violation to the Lipinski's rule, we did not discard it from our data set because of the high fit values (see Supporting Information). These 19 compounds are being acquired and will be tested for AlaR inhibition.

Validation of the pharmacophore model

In order to validate our model, O-acetyl-L-serine, O-carbamoyl-L-serine and propionate, which are among the very few known reversible competitive inhibitors [36, 8], were mapped onto the generated dynamic pharmacophore model. From Table 3 we can see that the propionate, a weak non-covalent inhibitor of *Bacillus stearothermophilus* AlaR ($K_i = 20 \pm 3$mM [8]) did not fit any of the features of the dynamic pharmacophore model with excluded volumes but, did fit two features of the model with no excluded volumes. We should mention here that in the hypothesis generation with excluded volumes, at least three location constraints are necessary to orient compounds correctly [15]. However, this is not mandatory for hypothesis generation with no excluded volumes. In CATALYST the hypothesis matches a compound if the topology of the compound completely satisfies the specification for all atoms and bonds in the hypothesis, and if at least one conformer of the compound fulfills all geometric constraints (distance, angle, torsion, location, excluded volumes) in the hypothesis. The fact that propionate did not fit the model with excluded volumes is simply a result of its inability to orient onto this model. From Table 3 we can see that the compound matched only two features of the model with no excluded volumes, which was insufficient in order to have the compound oriented onto the model with excluded volumes. The other two non-covalent inhibitors did map three of the four features of the dynamic pharmacophore model with no excluded volumes. In the case of the model with excluded volumes, the O-acetyl-L-serine fitted only the 1.5R and 2R models, while O-carbamoyl-L-serine fitted all the models with excluded volumes, except the 0.75R model. Unfortunately, due to the paucity of experimental data for non-covalent inhibitors, we were not able to interpret which features of the dynamic pharmacophore model might be the most important in order to further filter our hits, as for instance it was possible in the case of HIV-integrase [10].

D-Cycloserine (DCS) was also tested against the dynamic pharmacophore model and did not fit a 3 out 4 site model well. This is mainly due to its small size and rigidity, therefore hampering it from orienting onto the pharmacophore model. In addition, DCS has 2–3 pharmacophore elements (CO- in deprotonated state of ring NH, NH_3+, and possibly some hydrophobicity from a portion of the ring) while our receptor-based pharmacophore models contain 4–5 elements (4 for the dynamic and 5 for some of the static models). DCS does fit, but the score is not as high as molecules containing 4–5 matching pharmacophore elements. In addition, DCS does not bind strongly to the active site of the enzyme (its K_i is ca. 0.5 mM) [37] and its initial mode of action is covalent attachment to the PLP and possibly later to both PLP and the

Table 3. Performance of the dynamic pharmacophore models with different radii[a] in fitting known reversible competitive inhibitors from the literature.

Competitive Inhibitors	Without excluded volumes								With excluded volumes							
	0.75R		1R		1.5R		2R		0.75R		1R		1.5R		2R	
	1F	2F	1F	2F	1F	2F	1F	2F	1F	2F	1F	2F	1F	2F	1F	2F
Propionate	–	+ HYD HBA1	–	+ HYD HBA1	–	+ HYD HBA1	–	+ HYD HBA1	–	–	–	–	–	–	–	–
O-Acetyl-L-serine	+ HBD HBA1 HBA2	+ HBD HBA1 HBA2	+ HBD HBA1 HBA2	+ HBD HBA1 HBA2	+ HBD HBA1 HBA2	+ HBD HBA1 HBA2	+ HBD HBA1 HBA2	+ HBD HBA1 HBA2	–	–	–	–	+ HBD HBA1 HBA2	+ HBD HBA1 HBA2	+ HBD HBA1 HBA2	+ HBD HBA1 HBA2
O-Carbamoyl-L-serine	+ HBD HBA1 HBA2	+ HBD HBA1 HBA2	+ HBD HBA1 HBA2	+ HBD HBA1 HBA2	+ HBD HBA1 HBA2	+ HBD HBA1 HBA2	+ HBD HBA1 HBA2	+ HBD HBA1 HBA2	–	–	+ HBD HBA1 HBA2	+ HBD HBA1 HBA2	+ HBD HBA1 HBA2	+ HBD HBA1 HBA2	+ HBD HBA1 HBA2	+ HBD HBA1 HBA2

[a]Radius of the pharmacophore feature represents the RMS standard deviation across cluster variables (see Methods). Compounds that fit the model with one/two omitted features ('1F/2F') out of 4 features are marked with '+', and those that do not fit are marked with '−'. The mapped features are listed; 'HYD' stands for hydrophobic, 'HBA' for hydrogen-bond acceptor and 'HBD' for hydrogen-bond donor feature.

Table A. Characteristics of the dynamic pharmacophore model based on the 11 snapshots taken at 100 ps intervals from the 1 ns MD simulation.

Pharmacophore Element No.	SITE[a]	R (Å)[b]	X (Å)	Y (Å)	Z (Å)
1	HYD	0.37	15.62	−3.92	−24.65
2	HBA1	1.05	17.16	−1.70	−20.98
3	HBA2	0.67	18.06	−5.76	−20.94
4	HBD	0.60	18.03	−0.29	−17.21
5	EXCL1	1.5	19.54	−2.53	−17.21
6	EXCL2	1.5	17.54	−8.66	−21.16
7	EXCL3	1.5	16.55	−8.92	−19.01
8	EXCL4	1.5	18.76	−9.97	−19.62
9	EXCL5	1.5	18.81	−0.31	−21.50
10	EXCL6	1.5	19.56	−2.96	−25.37
11	EXCL7	1.5	13.28	0.57	−21.05

[a]HYD stands for hydrophobic, HBA stands for hydrogen-bond acceptor and HBD stands for hydrogen-bond donor pharmacophore feature; EXCL represent the excluded volumes.
[b]Radius of the pharmacophore element.

enzyme [38]. This is also consistent with the plethora of clinical side effects of DCS, thought to be due to the non-specific inactivation of other PLP containing enzymes [38].

The static model

It was shown that in the case of HIV-integrase a dynamic pharmacophore model compared very well to known inhibitors of the target, particularly those with very low IC_{50} values [10]. On the contrary, none of the known inhibitors were found to fit the static model, derived from the crystal structure alone. Given this example, a dynamic model, which takes into account the flexibility of the target, should be more generally suitable than a static model. In addition, a receptor-based pharmacophore model derived from a single protein conformation is subject to substantial sensitivity to the exact sidechain conformations in the inhibitor binding site (e.g., a single sidechain that is in a non-optimal orientation can block the availability of a cavity for the binding of inhibitor functional groups). In our case, the static pharmacophore model (Figure 8), derived from the crystal structure used to initiate the MD was tested against the 34 **conserved** molecules that resulted from the ACD search using different radii for the dynamic pharmacophore elements and with/without excluded volumes.

The static pharmacophore model, representing the three-dimensional collection of property spheres given by LigBuilder [13], is based on the active site of the target alanine racemase at 0 ps of the MD simulation. The static pharmacophore model at 0 ps contains five features, as shown on Figure 2; one hydrophobic, two hydrogen-bond acceptor and two hydrogen-bond

Table B. Characteristics of the target ligands on the dynamic pharmacophore model with differently scaled radii[a] of the pharmacophore element volumes. In all models, excluded volumes with a radius of 1.5 Å were used[b].

Compound	0.75R		1R		1.5R		2R	
	BestFit[c]	Conf. En. (kcal/mol)[d]	BestFit[c]	Conf. En. (kcal/mol)[d]	BestFit[c]	Conf. En. (kcal/mol)[d]	BestFit[c]	Conf. En. (kcal/mol)[d]
1	6.57	5.78	7.19	5.80	7.61	5.80	7.79	5.81
2	7.94	8.67	7.96	8.67	8.26	0.02	8.59	0.02
3	7.29	2.67	7.60	2.67	7.82	2.67	7.9	2.67
4	6.34	0.07	5.65	0.07	7.51	0.07	8.16	0.07
5	6.62	0.18	7.23	0.20	7.65	0.20	7.93	0.20
6	5.7	0.03	6.47	0.03	7.39	0.03	8.09	0.03
7	7.41	3.72	7.67	3.72	7.89	1.26	8.38	1.26
8	7.93	0.25	8.32	0.25	8.69	0.25	8.83	0.25
9	7.48	15.86	7.71	15.86	8.16	2.17	8.53	2.17
10	7.71	4.97	8.14	6.95	8.61	6.95	8.78	6.95
11	7.45	15.11	7.70	15.71	7.86	15.71	8.35	5.77
12	6.3	17.99	7.04	17.99	7.57	17.99	7.76	17.99
13	7.61	1.43	7.78	1.43	7.90	1.43	7.90	1.43
14	7.48	3.85	7.71	3.09	7.97	4.60	8.42	4.60
15	7.27	5.22	7.59	5.22	7.81	5.22	7.94	9.45
16	6.04	0.03	7.25	0.03	8.22	0.03	8.56	0.03
17	8.07	4.19	8.47	4.19	8.76	4.19	8.33	15.57
18	7.77	3.80	7.87	3.79	7.94	3.79	7.87	3.71
19	7.11	3.71	7.50	3.71	7.78	3.71	7.87	3.71
20	7.62	16.65	7.79	16.64	7.91	16.64	7.95	16.65
21	5.53	2.63	6.61	2.63	7.38	2.63	7.65	2.63
22	7.81	10.79	7.89	10.79	7.95	10.79	8.24	12.47
23	6.79	3.06	7.32	3.06	7.70	3.06	7.99	0.01
24	7.65	0.03	7.80	0.02	7.91	0.02	7.95	0.02
25	6.32	2.85	5.78	2.57	7.40	1.66	8.10	1.66
26	5.28	0.02	6.40	0.02	7.29	0.02	7.60	0.02
27	7.52	0.03	7.73	0.01	7.88	0.01	7.98	0.84
28	7.04	2.69	7.46	2.68	7.76	2.69	8.21	14.72
29	5.86	18.79	6.79	18.79	7.46	18.79	7.69	18.79
30	7.46	15.08	7.70	15.07	7.87	15.07	7.92	15.08
31	6.66	7.32	7.24	7.32	7.66	7.32	8.18	18.71
32	6.93	3.80	7.40	3.80	7.73	3.80	8.19	6.98
33	7.56	1.32	7.75	1.32	7.89	1.32	8.07	0.01
34	7.91	6.62	7.94	6.62	8.07	3.80	8.48	3.80

[a] Radius of the pharmacophore feature represents the RMS standard deviation across cluster variables (see Methods).
[b] Excluded volumes were added as described in the text in order to prevent hits from overlapping with receptor atoms.
[c] Fit values are calculated by CATALYST as described in the Methods Section and indicate how well the features in the pharmacophore overlap with the chemical features in the ligand and normalized for the number of features in the hypothesis.
[d] ΔE (kcal/mol): Energy difference between the conformer used for mapping ('best fit') and the local minimum calculated by CATALYST.

Table C. Characteristics of the static pharmacophore model based on the crystal structure.

Pharmacophore Element No.	SITE[a]	R (Å)[b]	X (Å)	Y (Å)	Z (Å)
1	HYD	0.64	15.81	−4.17	−25.03
2	HBA1	0.76	16.31	−2.17	−21.03
3	HBA2	0.55	19.31	−7.17	−25.53
4	HBD	0.53	17.31	−0.71	−17.03
5	EXCL1	1.5	18.37	−1.48	−23.91
6	EXCL2	1.5	16.56	−8.88	−19.22
7	EXCL3	1.5	18.89	−9.64	−19.79
8	EXCL4	1.5	17.59	−8.48	−21.33
9	EXCL5	1.5	13.15	0.34	−21.26
10	EXCL6	1.5	18.52	−7.25	−19.48
11	EXCL7	1.5	18.20	−0.62	−21.13

[a]HYD stands for hydrophobic, HBA stands for hydrogen-bond acceptor and HBD stands for hydrogen-bond donor pharmacophore feature; EXCL represent the excluded volumes.
[b]Radius of the pharmacophore element.

donor features. Because one of the hydrogen-bond donor features was too distant (9.1 Å away from the hydrophobic pharmacophore feature), we were unable to include this site into CATALYST [15]. Therefore, our static model, as shown in Figure 8, based on the first conformation of the protein at 0 ps from our MD simulation is lacking one hydrogen-bond donor feature. The characteristics of the four remaining features of the static model used in CATALYST are given in the Supporting Information. The radius of each static pharmacophore feature was determined from the key interaction sites within the binding pocket, derived by LigBuilder [13]. These interaction sites were then clustered following the same procedure as described in the Methods section. The radii of the static pharmacophore elements were set to 0.75R, 1R, 1.5R and 2R (R = RMS standard deviations across variables in the cluster). Models were generated with and without excluded volumes, as in the case of the dynamic model.

The static model was tested against the **conserved** hits with only 27 molecules out of 34 fitting the static model. The fact that almost 20% of the ACD hits did not fit the static model implies that many of the hits identified from a database search using a dynamic model will be missed when using a static model for the 3D database searches. Overall, an 80% agreement between the static model and the dynamic one might not seem to be so poor but, as we mentioned before, the static model used for testing the **conserved** hits with CATALYST is incomplete, lacking one feature which was too distant from the other ones. This is an indication for us that a static model it is not a good staring point as compared to a dynamic model which gives a more accurate description of the binding site taking into account protein flexibility. Moreover, since the static model is based on a single conformation of the protein (the X-ray structure), errors in the starting structure will be magnified.

One could argue that a dynamic model could be derived from a static one by assigning to each pharmacophore feature a radius proportional to the B-factors of the neighboring atoms. However, crystallographic B-factors form an incomplete measure of the motion since they are a measure of the spread of the electron density as a function of position in the crystal, with no consideration to which particular atom is contributing to the electron density. In addition to this, the crystallographic B-factors are generally restricted to a given maximum value during structure refinement, whereas the simulated B-factors are true atomic positional fluctuations [39, 40]. Hence, atomic fluctuations derived from a molecular dynamics simulation provide a better tool to investigate protein flexibility.

Conclusion

A much faster structure-based dynamic pharmacophore method has been developed using the program LigBuilder and multiple conformations of the receptor. This new method has been applied to an alanine racemase, which is an important broad-spectrum antibacterial drug design target. The advantage provided by developing a dynamic pharmacophore model is offered by the fact that it takes into account the protein's active site flexibility, extracted from molecular dynamics simulations under the influence of explicit solvent molecules and possibly a substrate or inhibitor. Based on the generated pharmacophore model, a series of compounds that fit our model were identified as possible inhibitors of the enzyme. We intend to have these compounds tested for activity against alanine racemases. The goal of any such computer-aided inhibitor discovery effort is the identification of lead compounds that will be refined into clinical candidates. The refinement of such leads is the goal of computational studies using other more precise methods and state of the art medicinal chemistry. Including receptor flexibility in structure-based inhibitor design strategies has been a substantial problem. This method

offers a next step toward achieving this goal. Current work is focused on identifying rational approaches for selection of relevant protein conformations and simpler methods for generating them.

Acknowledgements

This research was supported in part by NSF cooperative agreement ACI-9619020 through computing resources provided by the National Partnership for Advanced Computational Infrastructure at the Texas Advanced Computing Center. We gratefully acknowledge the NIH for support of this work via Grant AI 46340, NIAID. GIM is also grateful to M.S. Mudadu for help in identifying the chemically reactive structures, Dr. R. Wang for support in the use of LigBuilder and Dr. H.C. Huang for valuable assistance and helpful discussions. Numerous members of the Center for Molecular Research in Infectious Diseases, University of Houston, are acknowledged for many useful conversations and advice. Gratitude is also expressed to Accelrys, Inc. for software made available to us through the Institute for Molecular Design at the University of Houston.

Supporting information

Characteristics of the dynamic pharmacophore model based on the 11 snapshots taken at 100 ps intervals from the 1 ns MD simulation (Table A); Characteristics of the target ligands on the dynamic pharmacophore model with differently scaled radii of the pharmacophore element volumes (Table B); Characteristics of the static pharmacophore model based on the crystal structure (Table C).

References

1. Roe, D.C. and Kuntz, I.D., J. Comp. Aid. Mol. Des., 9 (1995) 269.
2. Böhm, H.-J., J. Comp. Aid. Mol. Des., 10 (1996) 265.
3. Kick, E.K., Roe, D.C., Skillman, A.G., Liu, G.C., Ewing, T.J.A., Sun, Y., Kuntz, I.D. and Ellman, J.A., Chem. Biol., 4 (1997) 297.
4. Klebe, G., J. Mol. Med. 78 (2000) 269.
5. Helmy, B., Scand. J. Respir. Dis. 71S (1970) 220.
6. Shaw, J.P., Petsko, G.A. and Ringe, D., Biochemistry 36 (1997) 1329.
7. Stamper, C.G.F., Morollo, A.A. and Ringe, D., Biochemistry 37 (1998) 10438.
8. Morollo, A.A., Petsko, G.A. and Ringe, D., Biochemistry 38 (1999) 3293.
9. Watanabe, A., Yoshimura, T., Mikami, B., Hayashi, H., Kagamiyama, H. and Esaki, N., J. Biol. Chem. (2002) in press.
10. Carlson, H.A., Masukawa, K.M., Rubins, K., Bushman, F.D., Jorgensen, W.L., Lins, R.D., Briggs, J.M. and McCammon, J.A., J. Med. Chem. 43 (2000) 2100.
11. Jorgensen, W.L., BOSS Version 3.8 (1997), Yale University, New Haven, CT.
12. Lipinski, C.A., Lombardo, F., Dominy, B.W. and Feeney, P., Advanced Drug Delivery Reviews 23 (1997) 3.
13. Wang, R., Gao, Y. and Lai, L., J. Mol. Model. 6 (2000) 498.
14. ACD: Available Chemicals Directory; Version 2000, Accelrys Inc., http://www.accelrys.com.
15. CATALYST, Accelrys Inc., San Diego, CA, http://www.accelrys.com.
16. Strych, U., Benedik, M.J., J. Bact. 184, Vol. 1 (2002) 4321.
17. MacKerell, A.D., Bashford, D., Bellott, M., Dunbrack, R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F.T.K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodhom, B., Reiher, W.E., Roux, B., Schlenkrich, M., Smith, J.C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., Karplus, M., J. Phys. Chem. 102 (1998) 586.
18. Gaussian98, Gaussian, Inc., Pittsburgh, PA, 1998, http://www.gaussian.com.
19. Antosiewicz, J., Briggs, J.M., Elcock, A.H., Gilson, M.K., McCammon, J.A., J. Comput. Chem. 17 (1996) 1633.
20. Madura, J.D., Briggs, J.M., Wade, R.C., Davis, M.E., Lutty, B.A., Ilin, A., Antosiewicz, J., Gilson, M.K., Bagheri, B., Scott, L.R., and McCammon, J.A. Comp. Phys. Comm. 91 (1995) 57.
21. Ondrechen, M.J., Briggs, J.M. and McCammon, J.A., J. Am. Chem. Soc. 123 (2001) 2830.
22. Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. and Karplus, M., J. Comp. Chem. 4 (1983) 187.
23. Kalé, L., Skeel, R., Bhandarkar, M., Brunner, R., Gursoy, A., Krawetz, N., Phillips, J., Shinozaki, A., Varadarajan, K. and Schulten, K., J. Comput. Phys. 151 (1999) 283.
24. Berendsen, H.J.C., Postma, J.P.M., van Gusteren, N.F., Di Nola, A. and Haak, J.R., J. Chem. Phys. 81 (1984) 3684.
25. Jorgensen, W.L., Chandresekhar, J., Madura, J., Impey, R., Klein, M.J., Chem. Phys. 79 (1983) 926.
26. Ryckaert, J.P., Cicotti, G. and Berendsen, H.J.C., J. Comp. Phys. 23 (1977) 327.
27. Allen, M.P. and Tildesley, D.J., Computer Simulations of Liquids, Oxford University Press, New York, 1987.
28. SAS 8.1, SAS Institute Inc., Cary, NC 27513, USA, http://www.sas.com.
29. Barnum, D., Greene, J., Smellie, A., Sprague, P., J. Chem. Inf. Comput. Sci. 36 (1996) 563.
30. Wang, R., Gao, Y. and Lai, L., Perspectives in Drug Discovery and Design 19 (2000) 47.
31. InsightII, Accelrys Inc., http://www.accelrys.com.
32. Bondi, A., J. Phys. Chem 68 (1964) 441.
33. Smelie, A., Teig, S., Towbin, P., J. Comp. Chem. 16 (1995) 171.
34. Hansch, C., Bjorkroth, J.P. and Leo, A., J. Pharm. Sci. 76 (1987) 663.
35. Bozler, G. and Schmid, J, Martin, Y.C., Kutter, E. and Austel, V. (Eds.), Principles of Pharmacokinetics and Drug Metabolism, Modern Drug Research – Path to Better and Safer

Drugs, Marcel Dekker, Inc., New York, NY, U.S.A., 1989, pp. 77–160.
36. Wang, E. and Walsh, C., Biochemistry 17 (1978) 1313.
37. Roze, U., Strominger, J.L., Mol. Pharmacol. 2 (1966) 92.
38. Neuhas, F.C., Gotlich, P.L., Shaw (Eds.), Antibiotics, vol. 1, Springer-Verlag, Heidelberg, Germany, 1967, pp. 40.
39. Hünenberger, P.H., Mark, A.E., van Gunsteren, W.F., J. Mol. Biol. 252 (1995) 492.
40. Stocker, U., Spiegel, K., van Gunsteren, W.F, J. Biomol. NMR 18 (2000) 1.