J Mol Model (2003) 9:88–98
DOI 10.1007/s00894-002-0115-8

ORIGINAL PAPER

Rowyna Kueh · Noorsaadah Abdul Rahman · Amir Feisal Merican

Computational docking of L-arginine and its structural analogues to C-terminal domain of *Escherichia coli* arginine repressor protein (ArgRc)

Received: 12 July 2002 / Accepted: 20 November 2002 / Published online: 1 March 2003

Abstract The arginine repressor (ArgR) of *Escherichia coli* binds to six L-arginine molecules that act as its co-repressor in order to bind to DNA. The binding of L-arginine molecules as well as its structural analogues is compared by means of computational docking. A grid-based energy evaluation method combined with a Monte Carlo simulated annealing process was used in the automated docking. For all ligands, the docking procedure proposed more than one binding site in the C-terminal domain of ArgR (ArgRc). Interaction patterns of ArgRc with L-arginine were also observed for L-canavanine and L-citrulline. L-Lysine and L-homoarginine, on the other hand, were shown to bind poorly at the binding site.

Keywords Arginine repressor · L-Arginine structural analogues · Computational docking

Introduction

The *Escherichia coli* K-12 arginine repressor protein (ArgR) is involved in two disparate biological functions: regulation of transcription of genes involved in the biosynthesis of L-arginine and as an obligate accessory protein in Xer site-specific recombination at *cer* sites on multicopy plasmids. [1] One unusual feature of the arginine system was that, in contrast to the genes of other biosynthetic pathways, its genes were scattered over the *E. coli* chromosome (coined regulon by [2]) and controlled by the same regulator protein, ArgR. ArgR is a 156-residue homohexameric protein consisting of N-terminal (residues 1–71) and C-terminal (residues 82–156) domains. The N-terminal domain (ArgRn) is involved in DNA binding, while the C-terminal domain (ArgRc) is involved in L-arginine binding as well as oligomerization. [3, 4, 5, 6, 7] Being a hexameric protein, ArgR is unique among regulatory proteins as most are dimeric and tetrameric. ArgR homologues have been identified in various groups of bacteria that include *Bacillus stearothermophilus*, [8, 9] *Haemophilus influenzae*, [10] *Salmonella typhimurium* [11] and *Bacillus subtilis*. [12] ArgR is an aporepressor that has been shown to be dependent on its co-repressor L-arginine to bind to DNA. It binds to six L-arginine molecules. The ArgR C-terminal domain (ArgRc)–L-arginine complex has been crystallized at 2.2 Å resolution. [6]

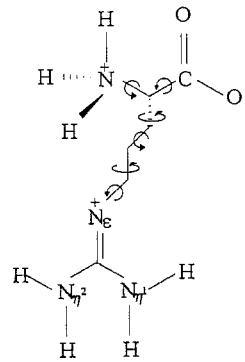
The L-arginine molecules act to stabilize ArgR as hexamers in solution. In the absence of L-arginine, a mixed solution containing full length ArgR and ArgRc was able to produce mixed hexamers, containing full length ArgR trimer and ArgRc trimer. [6] The full length ArgR and ArgRc trimers were able to exchange partners freely. In the presence of L-arginine, this trimer exchange between full length ArgR and ArgRc trimers was prevented and yielded only hexamers of full length ArgR as well as ArgRc. L-Canavanine (Fig. 1b), a structural analogue of L-arginine (Fig. 1a), was found to be able to act similarly to L-arginine to prevent the trimer exchange in solution but with less binding affinity. The other analogues (L-citrulline, L-lysine and L-homoarginine; Fig. 1c–e) were described as not being able to prevent the trimer exchange.

There are various approaches and programs currently available for studying protein–ligand interactions by computational docking. For example, geometric shape descriptors are used in the DOCK program, [13] while atomic representations of the structure are used in the AutoDock program. [14] A short overview of the variety

R. Kueh · A. F. Merican (✉)
Institute of Biological Sciences, Faculty of Science,
University of Malaya, 50603 Kuala Lumpur, Malaysia
e-mail: merican@um.edu.my
Tel.: +603-79676745
Fax: +603-79676749

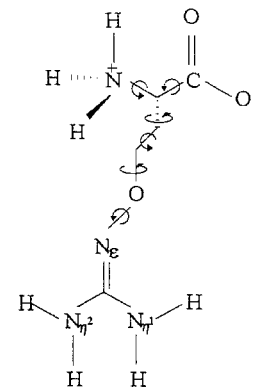
N. A. Rahman
Chemistry Department, Faculty of Science,
University of Malaya,
50603 Kuala Lumpur, Malaysia

L-arginine



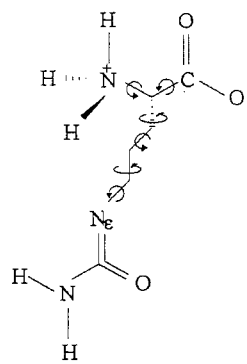
a

L-canavanine



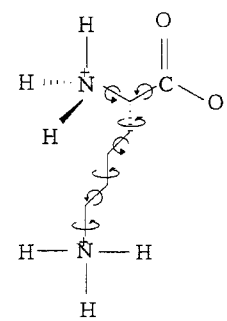
b

L-citrulline



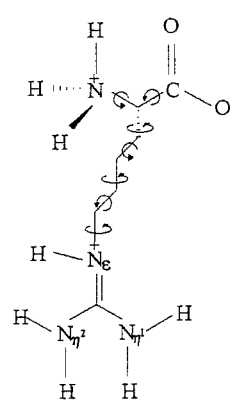
c

L-lysine



d

L-homoarginine



e

Fig. 1 Structure of L-arginine and its structural analogues. **a** L-Arginine. **b** L-Canavanine. **c** L-Citrulline. **d** L-Lysine. **e** L-Homoarginine. The arrows indicate the bonds defined as rotatable in the flexible docking runs

of methods that have been developed to deal with docking has been published. [15] AutoDock was first developed to provide an automated method of predicting the interactions of small flexible ligands and rigid biomacromolecular targets. It combines fast energy evaluation through a precalculated grid of affinity potentials with a Monte Carlo simulated annealing search algorithm. AutoDock has been used successfully in docking studies of various ligands with proteins and carbohydrates such as that of industrial glucohydrolase with monosaccharides [16] and antibody IgE Lb4 with several amino acid compounds. [15] In this study, AutoDock version 3.0 [17] was used. This version of the AutoDock program has been developed with new scoring functions as compared to the previous versions.

Here we report a computational docking study to compare the binding ability of L-arginine and its structural analogues to the C-terminal domain of the wild-type arginine repressor protein (ArgRc).

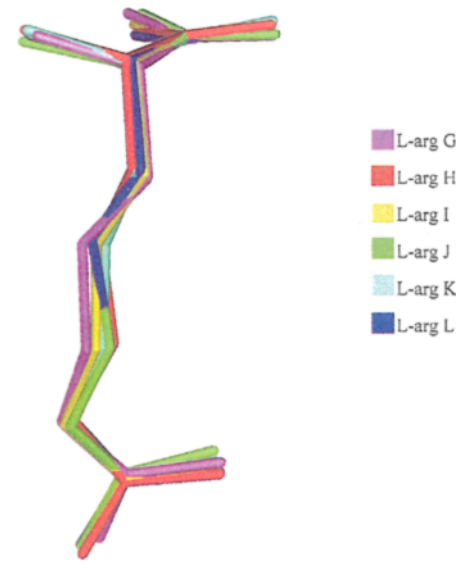


Fig. 2 All six crystallographic L-arginine molecules overlaid with each other (rendered as sticks)

Methods

ArgRc structure

The crystal structure of ArgRc (PDB entry 1xxa) was obtained from the Protein Databank (PDB). The structure 1xxa is the ArgRc complex bound with six L-arginine molecules (G, H, I. J, K and L) at six identical binding sites (see Fig. 2). The six molecules of L-arginine bind to the C-terminal domain at the interface between trimers. Three L-arginine molecules of one trimer are opposite the other three on another trimer. Each of the binding sites is made up of the interface residues of three ArgRc subunits (two subunits of one trimer and one subunit in the opposing trimer).

In preparation of the protein for docking, all the bound ligands and heteroatoms were removed from the PDB file. Hydrogen atoms were added to the 1xxa structure using the Builder module in Insight II (Accelrys Inc.) on a Silicon Graphics (SGI) O2 R5000 with the Irix 6.5 operating system. The positions of the hydrogen atoms were optimized using the Discover module in InsightII.

The six crystallographic L-arginine molecules do not show significant differences in their torsional angles (Fig. 3; Table 1). Therefore, the best structure chosen for docking was selected based on the suitability of the binding site. An assumption was made that any of the six L-arginine molecules has an equal chance of binding first to any of the binding sites. To select just one binding site, the B-factor values (obtained from the crystallographic PDB file of 1xxa) of the subunits that make up each binding site were evaluated and three ArgRc subunits with the lowest B-factor values were selected. Subunits A, B and F were found to have lower B-factor values than the other subunits. The corresponding binding site from these three subunits was bound by L-arginine H. Therefore, ligand H was selected for docking into its crystallographic binding site H.

Ligand structures

Under physiological conditions, the guanidino group of L-arginine has a pK_a of 12.48, resulting in a net positive charge for L-arginine. [18] L-Canavanine is a structural analogue of L-arginine, whereby the δ-methylene group in the carbon skeleton is replaced by oxygen (Fig. 1b). The presence of the oxygen atom in the side chain reduces the pK_a value of the guanidinooxy group to about 7.04. [19] Hence, at neutral pH, the L-canavanine side chain is not protonated. L-Citrulline differs from L-arginine in the fact that the guanidino group in the side chain of L-arginine is replaced by a carbamate group in L-citrulline (Fig. 1c). L-Lysine, on the other hand, has a basic side chain with a pK_a of around 10 and is positively charged at physiological pH (Fig. 1d). L-Homoarginine, a higher homologue of L-arginine, contains an additional methyl group in the carbon backbone (Fig. 1e).

Substitution of the δ-carbon atom in L-arginine with an oxygen atom using the Builder module in InsightII gave the L-canavanine structure. Similarly, L-citrulline was also obtained by substituting one of the terminal $-NH_2$ groups in the side chain of the L-arginine structure with an oxygen atom. For L-lysine, the Nϵ was substituted with a methylene group and a terminal $-NH_3^+$. The structure for L-homoarginine used in the calculation was extracted from a bovine endothelial nitric oxide synthase heme domain complex (PDB entry 1dm7). All structures were optimized with the Consistent Valence Force Field (CVFF). [20]

Docking calculations

AutoDock version 3.0 [17] was used for all docking calculations. Automated docking was performed using a Monte Carlo simulated

Table 1 Side chain dihedral angles of the six L-arginine crystal structure. The dihedral angles were obtained using the Biopolymer module in InsightII (Accelrys Inc.)

L-Arginine	χ_1 (CA-CB-CG-CD)	χ_2 (CB-CG-CD-Nε)	χ_3 (CG-CD-Nε-CZ)	χ_4 (CD-Nε-CZ-NH1)
G	47.2	−171.0	−155.0	99.7
H	60.4	178.8	170.7	109.3
I	54.5	−169.3	−171.5	93.1
J	48.5	−173.8	−178.4	91.5
K	68.1	−171.4	175.1	91.2
L	51.7	176.7	−176.6	111.5

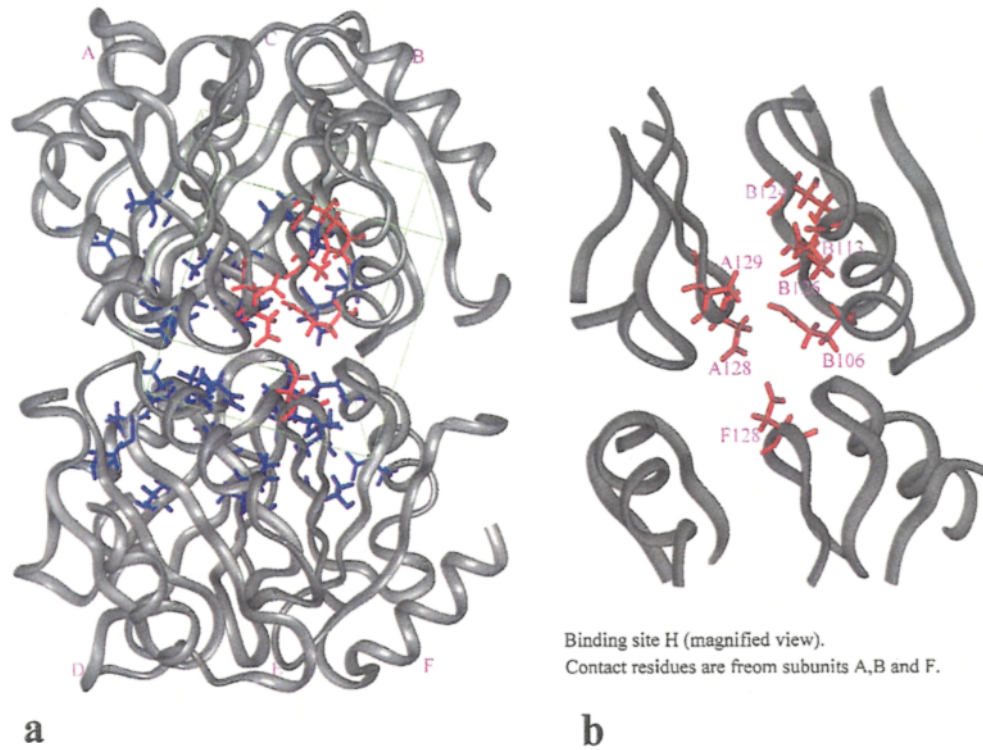


Fig. 3 **a** The crystal structure of ArgRc (subunits A, B, D, E and F) and the six binding sites of L-arginine in ArgRc. The ArgRc structure is represented by *grey* ribbons and the six binding sites are rendered as coloured sticks. *Red* coloured sticks: residues in binding site H that was selected for docking. *Blue* coloured sticks: all other binding sites. *Green* box: grid box centred in the binding site H. **b** The binding site H (right) consists of residues from subunits A, B and F. [6]

annealing process combined with a grid-based energy evaluation method. AutoDock version 3.0 uses an empirical binding free energy force field to estimate the free energy change upon binding. [17] The root mean square deviations (RMSDs) (over all atoms) for L-arginine and L-homoarginine were calculated with reference to the coordinates of the crystallographic structures. These values were used to compare the deviation of their docked structures from the crystallographic structure. For other ligands, the RMSD values were compared to their respective optimized structures, which were modified from the L-arginine structure.

For the docking calculations, the grid box generated was centred at binding site H (Fig. 3). A 20 Å side grid map was used for rigid and flexible L-arginine docking with 0.25 Å spacing. For the docking of the L-arginine structural analogues, a 25 Å side grid map was used with 0.25 Å spacing. The L-arginine H molecule used as the initial starting structure for docking was placed at one corner of the grid box in order to evaluate whether AutoDock is able to find the original binding site based on the placement of the grid box. However, the starting location of the ligand is not critical, as AutoDock will randomize the initial location and orientation at the beginning of each docking run. The other analogues were subjected to the same initial starting coordinates.

The simulated annealing process was carried out at a temperature corresponding to RT=1200 cal mol^{-1} and reduced by a factor of 0.90 after each cycle. One hundred runs with 100 cycles per run were performed, where each cycle consisted of a maximum of 30,000 accepted or rejected steps. Each step in a cycle corresponded to a random change in translational, rotational and torsional degrees of freedom of the ligand. Autotors, a utility program in AutoDock, was used to define rotatable torsion angles of the ligands. Six active torsions were assigned separately for L-arginine, L-canavanine and L-citrulline and seven active torsions for L-lysine and L-homoarginine (Fig. 1).

Results and discussion

Rigid-body docking of L-arginine

Rigid-body docking of L-arginine was performed initially to evaluate AutoDock's docking algorithm. In rigid-body docking, the docked configurations had the same conformation as the crystal structure of L-arginine H. Two sets of configurations at different binding sites were obtained from the docking calculation. One set contained configurations located in the crystallographic binding site H, while another set of configurations was located near the core of the protein structure. In order to differentiate the locations of the docked configurations, the binding site H was designated as Site 1 and the alternative site as Site 2 (Fig. 4). Docked energies of configurations in Site 1 ranged between −13.54 kcal mol^{-1} and −12.10 kcal mol^{-1} (Table 2), while docked energies of configurations in Site 2 ranged between −7.70 kcal mol^{-1} and −6.77 kcal mol^{-1} (Table 3). The resulting configurations in Site 2 showed higher docked energy values. Thus, Site 1 was chosen as the focus in this study, since such configurations showed similar binding positions and interactions to that of the crystal structure.

Configurations in Site 1 showed RMSD values of less than 0.81 Å from the crystallographic L-arginine H, while larger deviations (>5 Å) were calculated for the configurations in Site 2. The configurations in Site 1 were able to reproduce the hydrogen-bond interactions as seen in the crystal structure [6] and computational model of the ArgRc–L-arginine complex [21] (Fig. 5). A single L-arginine molecule formed ten hydrogen bonds with

Table 2 Summary of the docked energies, estimated free energy of binding and average RMSD produced from Site 1 in docking L-arginine and its structural analogues. The RMSD calculated in AutoDock is between docked conformations in each docking run with the corresponding initial starting structure. L-Arginine is compared to the crystal structure while the analogues were compared with their corresponding initial ligand structure used for docking

Ligands	No. of atoms	No. of free torsions	Docked energy in Site 1 (kcal mol^{-1})			Estimated free energy of binding in Site 1 (kcal mol^{-1})			Average RMSD (Å) calculated from conformations located in Site 1
			Min	Max	Average	Min	Max	Average	
Rigid docking									
L-Arginine	27	0	−13.54	−12.10	−12.99	−13.54	−12.10	−12.99	0.472
Flexible docking									
L-Arginine	27	6	−18.84	−14.71	−16.67	−11.63	−2.59	−6.15	2.677
L-Canavanine	27	6	−18.59	−13.10	−16.07	−8.91	−1.95	−6.30	1.404
L-Citrulline	25	6	−19.80	−15.83	−17.25	−8.88	−3.79	−6.97	1.470
L-Lysine	24	7	−13.55	−8.90	−10.34	−7.68	−0.94	−3.36	4.023
L-Homoarginine	30	7	−19.94	−16.37	−17.59	−5.82	−1.75	−3.77	3.712

Table 3 Summary of the docked energies, estimated free energy of binding and average RMSD produced from sites 2 to 6 in docking L-arginine and its structural analogues. The RMSD calculated in AutoDock is between docked conformations in each docking run with the corresponding initial starting structure. L-Arginine is compared to the crystal structure while the analogues were compared with their corresponding initial ligand structure used for docking. The large RMSD values obtained were due to changed orientation from alternative sites (Sites 3 to 6)

Ligands	No. of atoms	No. of free torsions	Docked energy in Sites 2 to 6 (kcal mol^{-1})			Estimated free energy of binding in Sites 2 to 6 (kcal mol^{-1})			Average RMSD (Å) calculated from conformations in Sites 2–6
			Min	Max	Average	Min	Max	Average	
Rigid docking									
L-Arginine[a]	27	0	−7.77	−6.77	−7.16	−7.77	−6.77	−7.16	7.003
Flexible docking									
L-Arginine[a]	27	6	−17.60	−13.91	−16.08	−5.18	−2.37	−3.47	8.077
L-Canavanine[b]	27	6	−15.74	−12.38	−14.08	−8.32	−1.23	−3.30	8.157
L-Citrulline[c]	25	6	−17.07	−13.02	−14.74	−4.51	−0.55	−2.47	7.757
L-Lysine[d]	24	7	−11.57	−2.04	−8.76	−4.62	−0.04	−1.57	8.900
L-Homoarginine[e]	30	7	−19.94	−11.88	−16.40	−6.28	0.74	−2.60	7.735

[a] Set of conformations located in Site 2
[b] Sets of conformations located in Sites 2, 3 and 4
[c] Sets of conformations located in Sites 2, 3 and 4
[d] Sets of conformations located in Sites 2, 3, 4, 5 and 6
[e] Sets of conformations located in Sites 2, 3 and 4

ArgRc, where the carboxylate and amino groups interact with two subunits of a trimer while the guanidino group forms a pair of hydrogen bonds with the side chain of Asp128 from one subunit in the opposing trimer [6] (Fig. 5a).

Flexible docking of ligands

The minimum energies of the docked conformations obtained from flexible docking of L-arginine and the analogues were lower than that of rigid-body docking with L-arginine (Table 2). This is because more degrees of freedom are involved with flexible ligands.

The minimum docked energy from the flexible docking of L-arginine produced fewer hydrogen-bond interactions than were obtained with the rigid-body docking of L-arginine. In this study, it was observed that the conformation with the lowest estimated free energy of binding was able to produce more interactions that are similar to the crystal structure and the modelled structure. [6, 21] Thus, for all analogues, conformations with the lowest estimated free energy of binding were used in addition to the minimum docked energy conformation in the analysis of the docking results. The estimated free energy of a docked ligand includes the intermolecular energy and the torsional free energy. The intramolecular interaction energy of the ligand is not included in the calculation of the estimated free energy of binding. However, the total docked energy of any docked conformation includes the intermolecular and intramolecular interaction energies.

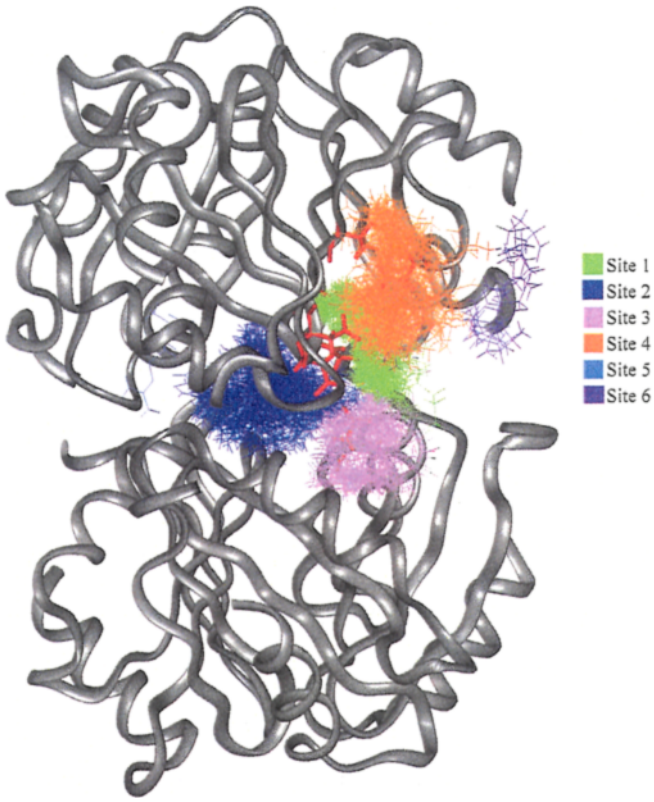


Fig. 4 A general overview of the sites found from docking the various ligands into ArgRc (*grey* ribbons). *Red* coloured sticks: residues in binding site H that was selected for docking.

L-Arginine

In the docking of flexible L-arginine, the docked conformations were more varied, although, in general, two sets of docked conformations were obtained (Site 1 and Site 2), similar to the ones described earlier for rigid docking of L-arginine (Fig. 4). Low docked energy conformations as well as high docked energy conformations could be found in both sites, but there were fewer low docked energy conformations in Site 2 than in Site 1, indicating that Site 1 is a preferable binding site for L-arginine H (Table 3). Although it was expected that the conformation with the lowest docked energy and RMSD value would be the closest to the native crystal structure of L-arginine, it was found from docking calculations that the lowest energy conformation did not have the lowest RMSD value. Instead, the conformations with low estimated free energy of binding appeared to reproduce the binding positions of the native crystal structure.

The lowest docked energy conformation showed an energy of −18.84 kcal mol^{-1} and a deviation of 2.43 Å from the crystallographic structure (Table 2). This conformation showed hydrogen-bond interactions with residues Asp128 and Asp129 from subunit A and Gln106 from subunit B (Fig. 6). There were two conformations that shared the same estimated free energy of binding of −11.63 kcal mol^{-1}. Both conformations also showed the

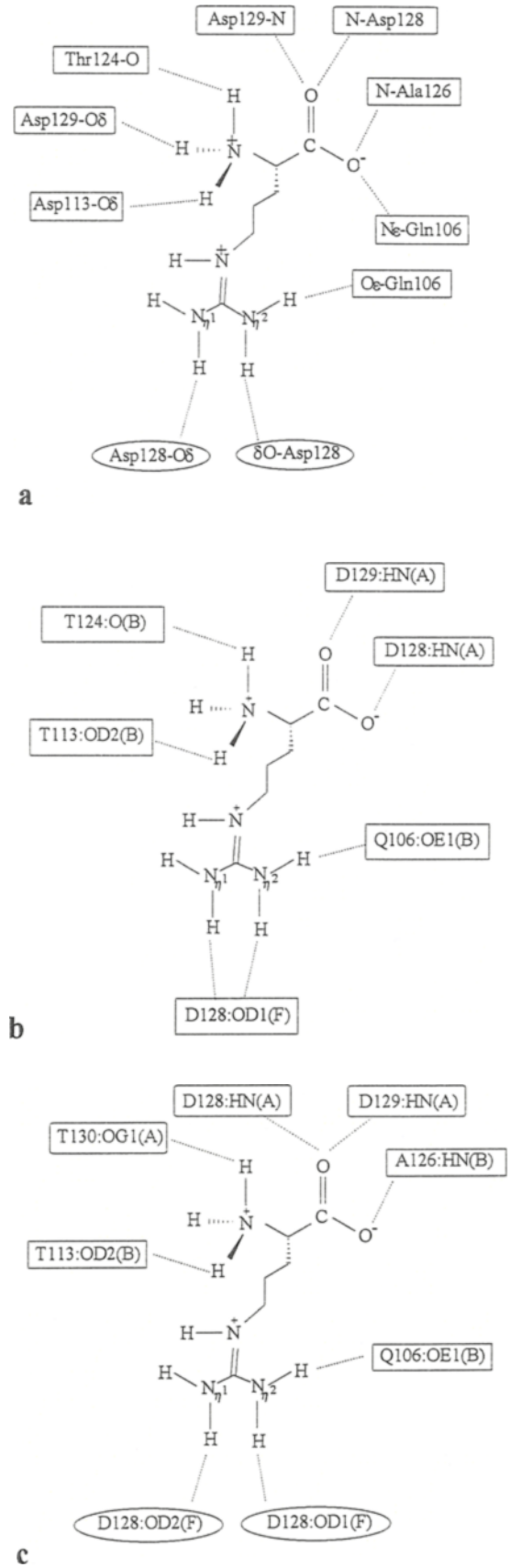


Fig. 5 Hydrogen-bond interactions of L-arginine in ArgRc. **a** Crystallographic data of L-arginine. [6] **b** Computational model. [21] **c** Lowest docked energy conformation in rigid-body docking (this study)

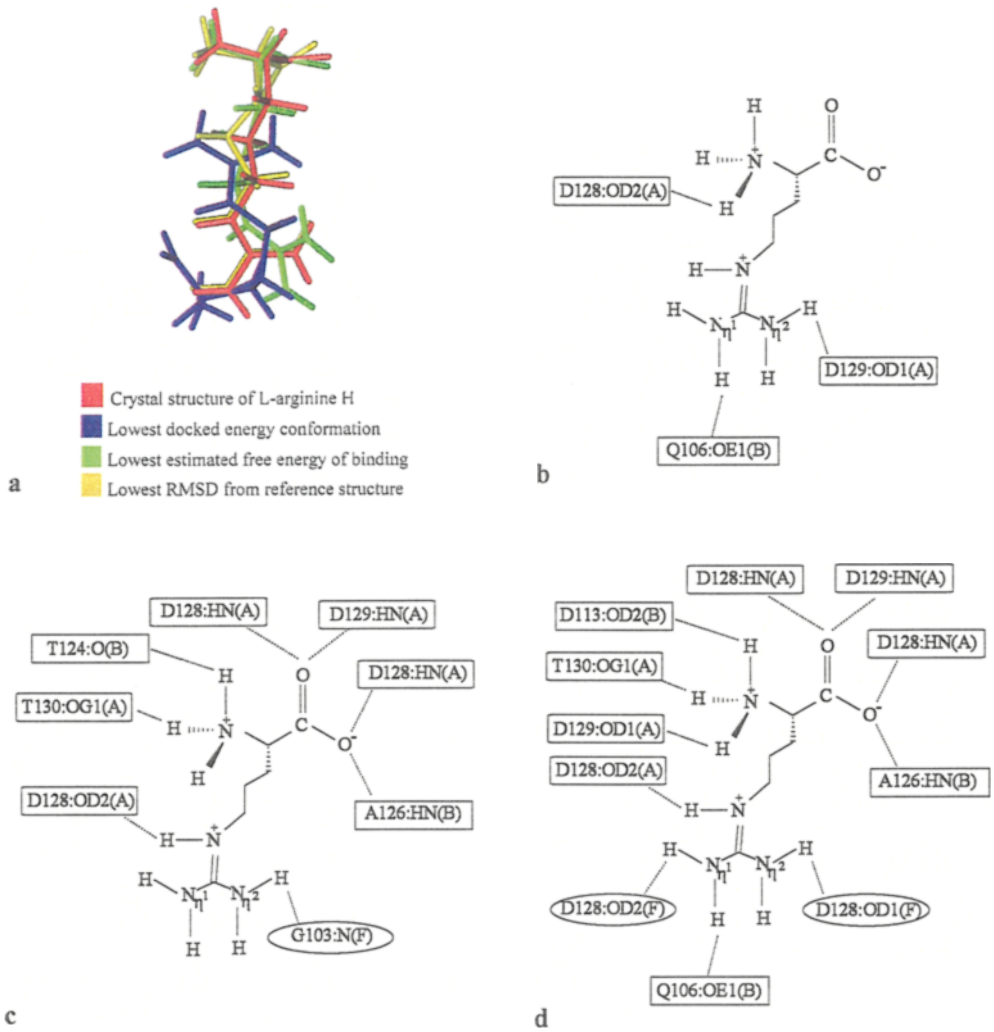


Fig. 6 Selected conformations from flexible docking of L-arginine and the hydrogen-bond interactions in ArgRc. **a** Orientations of three selected conformations with respect to the crystal structure of L-arginine. **b** Conformation with lowest docked energy. **c** Conformation with lowest estimated free energy of binding. **d** Conformation with lowest RMSD value with respect to the L-arginine crystal structure

docked energies of −17.36 kcal mol^{-1} with lower RMSD values (0.88 Å and 1.03 Å respectively). These two conformations showed interactions with ArgRc via nine hydrogen bonds to six residues from subunits A, B and F. The common residues for both conformations are Asp128, Asp129, Thr124 and Ala126 (Fig. 6). Additional residues include Thr130 and Gly103. In both these conformations, the guanidino group showed different orientations that enabled some hydrogen interactions with either Gly103 or Asp128 from subunit F. Their conformations were more similar to the crystal structure (Fig. 6).

Although L-arginine interacts preferentially with Asp128 and Asp129 of subunit F, [6, 21] some of the docked conformations showed that the guanidino side chain of L-arginine was able to orientate towards Pro102 or Gly103 to form a hydrogen bond. This suggests the possibility of L-arginine having more than one binding mode. Since a conformational shift occurred in the β-turn-α motif involving residues 100 to 112 when L-arginine is bound, this turn would have been pulled towards the centre of the hexamer. [6] This would have placed residues 102 and 103 closer to the binding site, making it possible for L-arginine to interact in the docking run with these residues through hydrogen bonding. Being a positively charged ligand, it would seem more likely that L-arginine would interact with Asp128 from the opposing trimer, which carries a net negative charge on the side chain. It is also interesting to note that Pro102 and Gly103 are conserved in ArgR homologues in *B. subtilis, H. influenzae, M. tuberculosis, S. clavuligerus and B. stearothermophilus*. [22]

L-Canavanine

Docking of L-canavanine produced more non-specific binding sites than for L-arginine. Apart from Site 1 and Site 2, several conformations produced were clustered around Site 1 (Fig. 4). To be consistent with the docking analysis of L-arginine, only docked conformations in Site 1 are taken into consideration for analysis. As expected, L-canavanine displayed similar binding positions to that of L-arginine and interacted with similar residues in ArgRc.

The docked conformations of L-canavanine in Site 1 appeared similar to each other where the docked energies

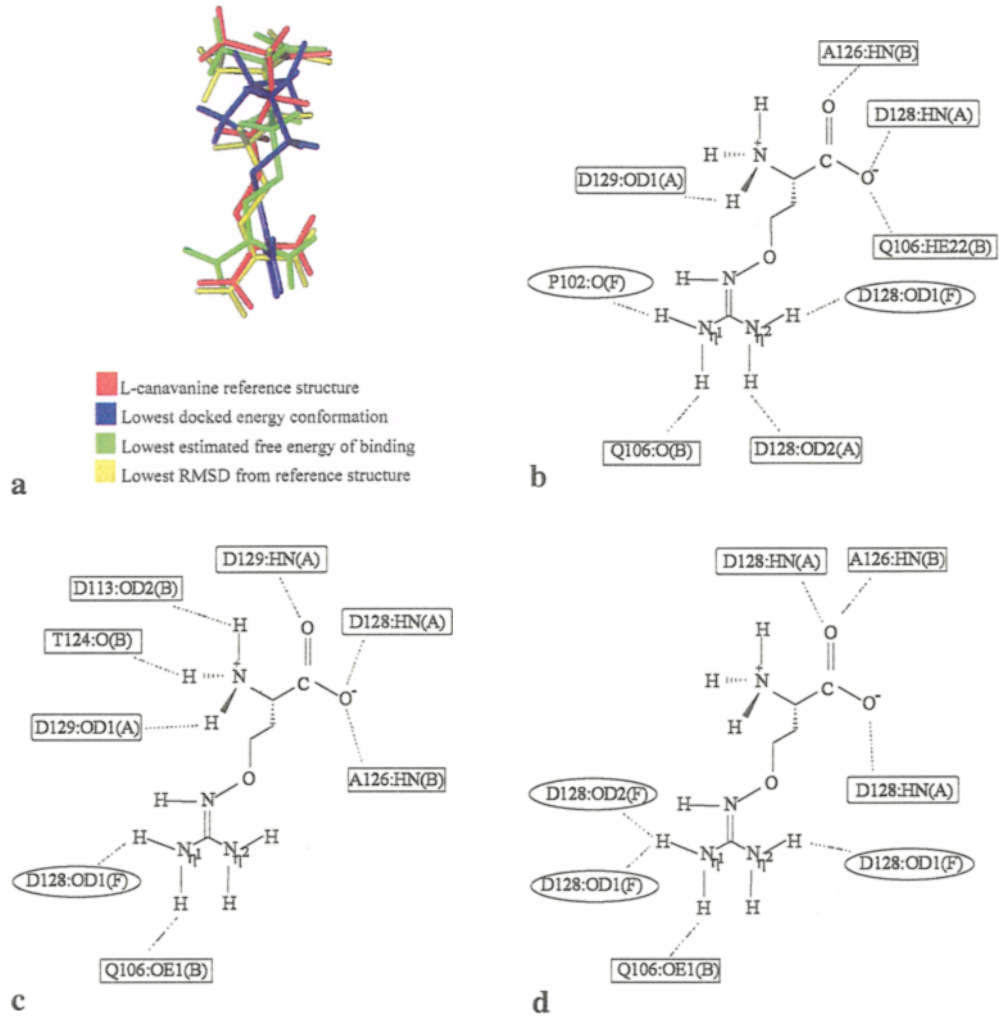


Fig. 7 Selected conformations from flexible docking of L-canavanine and the hydrogen-bond interactions in ArgRc. **a** Orientations of three selected conformations with respect to the reference structure of L-canavanine. **b** Conformation with lowest docked energy. **c** Conformation with lowest estimated free energy of binding. **d** Conformation with lowest RMSD value with respect to L-canavanine reference structure

are found to be lower than the docked energies found in the other sites (Tables 2 and 3). The RMSD of the docked conformations with respect to the initial L-canavanine starting structure were obtained from the AutoDock output file. The conformations located in Site 1 showed smaller deviations (0.75 to 4 Å) from the initial starting structure, while the docked energies ranged from −18.59 to -13.10 kcal mol^{-1}. The estimated free energy of binding from the docking of L-canavanine was higher than that of L-arginine (Table 2). In general, the lower docked energy conformations were able to produce interactions with contact residues in ArgRc similar to those of L-arginine. Several conformations in L-canavanine docking were able to form hydrogen bonds with three subunits of ArgRc (subunits A, B and F). The interactions with residues from subunit F were Pro102, Gly103 or Asp128. These interactions were made through the guanidinooxy side group of L-canavanine. There is no specific pattern as to whether conformations with lower docked energies or binding affinities showed preferences for interacting with 2 or 3 subunits of the protein. However, the docking results suggest that docked conformations of L-canavanine were able to interact with ArgRc through similar interacting residues as observed in L-arginine.

Our results may explain why L-canavanine was able to prevent the trimer exchange in solution, as described earlier in the introduction section. The structural similarity of L-canavanine to L-arginine enables it to satisfy the hydrogen-bonding requirements of the crystallographic ArgRc–L-arginine complex. [6] As such, L-canavanine can also act in place of L-arginine to form a stable hexamer of ArgR or ArgRc. As can be seen from the results (Fig. 7), the presence of two terminal amino groups made the interactions with the side chain carboxylic group of Asp128 from the opposing trimer possible. Docking results also showed that L-canavanine was able to interact with Gly103. This is possibly due to the fact that the position of Gly103 (as seen in the crystal structure of ArgRc) was within hydrogen-bond distance of docked L-canavanine.

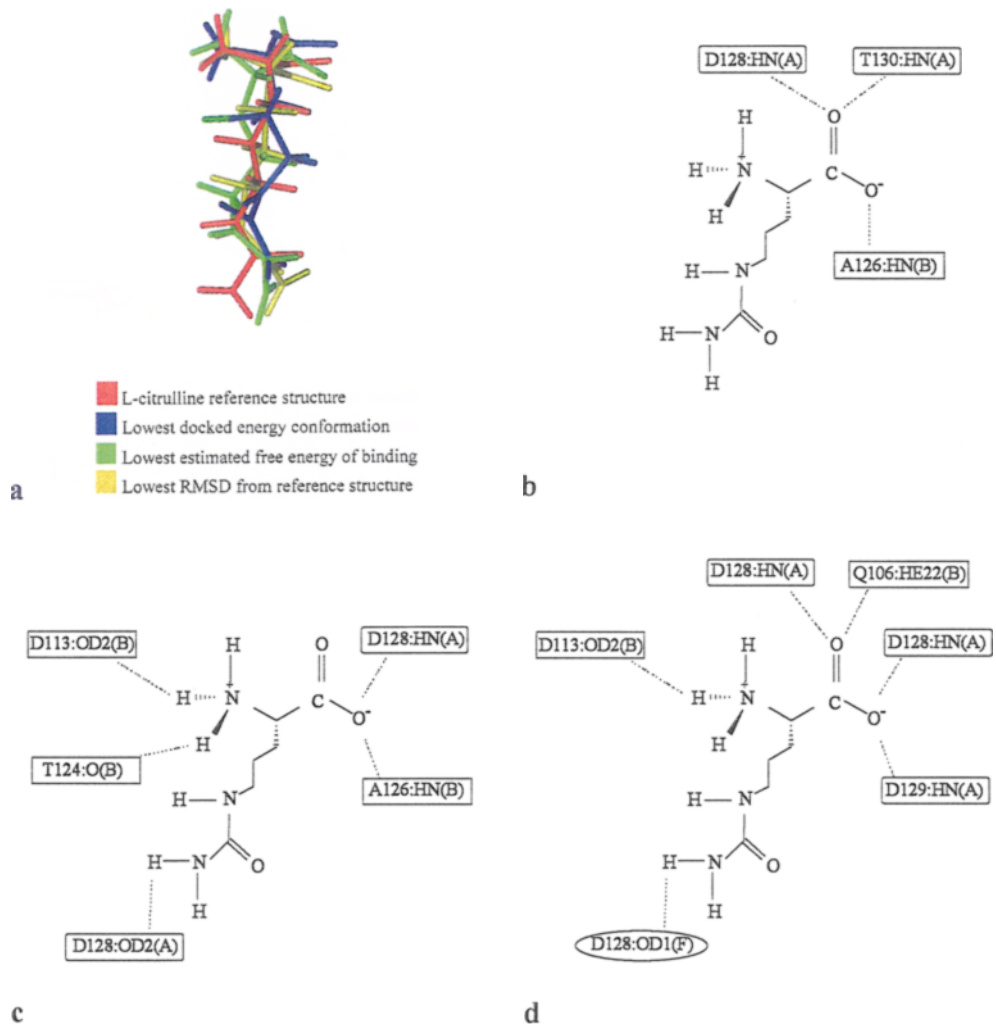


Fig. 8 Selected conformations from flexible docking of L-citrulline and the hydrogen-bond interactions in ArgRc. **a** Orientations of three selected conformations with respect to the reference structure of L-citrulline. **b** Conformation with lowest docked energy. **c** Conformation with lowest estimated free energy of binding. **d** Conformation with lowest RMSD from L-citrulline reference structure

L-Citrulline

Similar sets of conformations seen in L-canavanine docking were produced in L-citrulline docking. Several conformations were located around Sites 1 and 2 and their other locations can be generalized into Sites 3 to 6 (Fig. 4). As with the previous docking results, lower energy conformations were found in Site 1, where the energy ranged from -19.80 kcal mol^{-1} to −15.07 kcal mol^{-1} with RMSD values of less than 2.5 Å from the initial L-citrulline starting structure (Table 2). The results from L-citrulline docking seemed to indicate that docked conformations of L-citrulline were able to bind in positions similar to that of L-arginine (Fig. 8).

The conformation with minimum energy showed an estimated free energy of −8.35 kcal mol^{-1} and RMSD value of 1.32 Å from the initial starting structure. This conformation showed possible hydrogen-bond interactions with three residues of subunits A and B (Fig. 8). The docked conformation with the lowest estimated free energy of binding (−8.88 kcal mol^{-1}) showed more interactions with four residues of subunits A and B (Fig. 8). This conformation was similar to the initial starting structure with an RMSD value of 1.15 Å. There were some conformations where the terminal $-NH_2$ group of L-citrulline was able to form hydrogen bonds with Pro102, Gly103 or Asp128 of subunit F in the opposing trimer. However, the estimated free energy of binding was not as low as that of L-arginine. In addition, the frequency of such occurrence was not very high (14%) as compared to L-canavanine (26%).

As with L-canavanine, there was no specific pattern that indicates the preferences of L-citrulline to bind to two or three subunits. In the case of L-citrulline, only the amino group in the L-citrulline side chain was capable of forming hydrogen bonds with the side chain carboxylic group of Asp128 in the opposing trimer. Docking of L-citrulline produced several conformations that were able to form hydrogen bonds through its terminal $-NH_2$ group with Asp128 or Gly103 within subunit F of the opposing ArgRc trimer. However, conformations with lower binding energies showed L-citrulline interacting with two ArgRc subunits instead of three subunits. This may suggest that L-citrulline does not bind as tightly to binding site H of ArgRc as L-arginine and L-canavanine, possibly due to the mode of interaction exhibited by L-citrulline

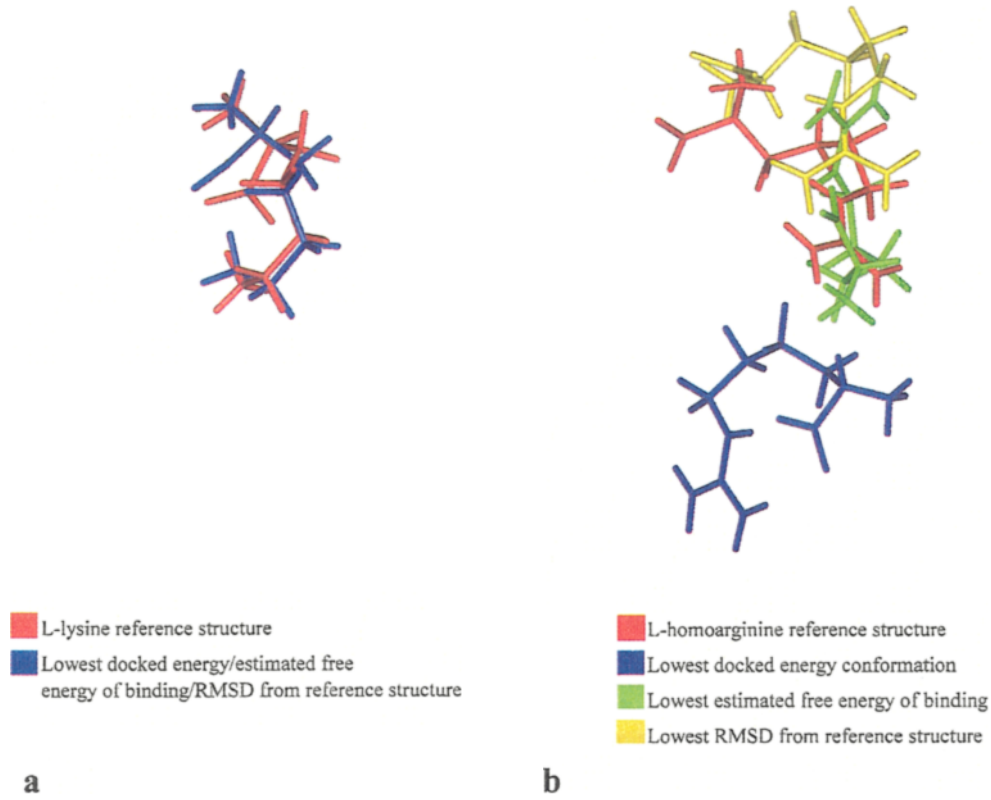


Fig. 9 Selected conformations from flexible docking of **a** L-lysine and **b** L-homoarginine

upon binding to ArgRc. Unlike L-arginine and L-canavanine, only one $-NH_2$ group is present in L-citrulline that is able to interact with residues from the opposing trimer. The estimated free energies of binding from L-citrulline docking were lower than that of L-arginine and L-canavanine.

L-Lysine

Computational docking results for L-lysine indicated that this ligand binds poorly to the binding site H in ArgRc. There were only a few docked conformations in Site 1, with docked energies ranging from −13.55 kcal mol^{-1} to −8.90 kcal mol^{-1}. L-Lysine showed a more "curled-up" conformation with a greater tendency for intramolecular hydrogen bonding of the side chain $-NH_3^+$ with carboxylate group rather than the more "linear" conformation required to form hydrogen bonds with the protein (Fig. 9a). The minimum energy conformation of L-arginine also showed the lowest estimated free energy of binding with the lowest deviation from the initial L-lysine starting structure (Table 2). However, only two-subunit interactions can be seen from this conformation. It would appear that L-lysine does not share a binding position similar to L-arginine.

L-Homoarginine

As with L-lysine, docking results for L-homoarginine also showed that this ligand interacted differently with binding site H. The docked conformations were scattered within Site 1 as well as around it (Sites 2 to 6; Fig. 4). The lowest docked energy conformation was not located in Site 1. This was also the case for the conformation with lowest estimated free energy of binding. The docked conformations located in Site 1 did not show similar binding conformations to that of L-arginine nor were there interactions that suggested L-homoarginine was able to interact with three ArgRc subunits. Selected conformations of L-homoarginine are shown in Fig. 9b. L-Homoarginine showed preferences of binding to other sites beyond the binding site of ArgRc.

Further discussion

In addition to binding site H (Site 1), our docking experiment also showed more than one possible binding site for the ligands (Fig. 4). However, since most of the docked conformations with minimum energy were located in Site 1, it indicates that the docking procedure used was able to distinguish the real binding site from other non-specific ones. This is true for L-arginine and perhaps for L-canavanine and L-citrulline as well. However, in the case of L-lysine and L-homoarginine, the docked conformations do not appear to have favourable docking

energies in Site 1. Thus, L-lysine and L-homoarginine were both shown to be poor ligands for ArgRc.

To study the ligand interactions in ArgRc, ideally it would be best to dock all six molecules of the same ligand together to a flexible protein The use of an automated docking procedure here presents possible binding modes of L-arginine and its structural analogues to ArgRc. Computational docking studies using AutoDock have been reported with good success rates at reproducing crystal structure interactions of different protein–ligand complexes, even without prior knowledge of the binding site. [23] Ideally, the conformations with low RMSD should be predicted as having the lowest docked energy as well. Initially, analyses were performed on the conformations having low docked energies with L-arginine docking. However, it is apparent that the crystal structure interactions were not reproducible with the lowest docked energy structure. Instead, the outcome of this study suggested that the conformations with the lowest estimated free energy of binding are able to reproduce the binding position of the crystal structure. Taking such consideration into account, the structures with both lowest energy docked and lowest free energy of binding were used for the analysis.

There are common interaction patterns exhibited by the docked ligands, which can be presented as follows:

1. Lowest docked energy conformations were located in Site 1 (except for L-homoarginine).
2. Conformations with the lowest estimated free energy of binding were located in Site 1 (except for L-homoarginine).
3. At least one alternative or non-specific binding site was obtained. For L-arginine docked to wild-type ArgRc, two sites were obtained, while for the other analogues more than two sites were obtained. Other studies using AutoDock have also reported docked positions in sites other than the original binding site. [17, 24].

The choice of force field also plays an important role in docking. A docking study of artemisinin to heme [25] found that the atomic charges of heme had a significant effect on the docking configurations. In our study, the lowest energy obtained can be assumed to be near the global minimum of the structure. The lowest docked energy found with the CVFF force field was ranked with low RMSD from the experimental position but not the lowest RMSD values.

Conclusion

We have presented computational docking results of L-arginine and its structural analogues to wild-type ArgRc. Out of the four L-arginine structural analogues studied, L-canavanine was able to show the most similarity of L-arginine binding to ArgRc. Results from docking L-citrulline to ArgRc showed some possibilities as well. L-Lysine and L-homoarginine were shown to be poor ligands to bind to the binding site of ArgRc.

Acknowledgements We would like to acknowledge Dr Francois Dupradeau of Université de Picardie Jules Verne, France for his critical comments. This study is supported by the Biotechnology TopDown Grant under the National Biotechnology Program of the Ministry of Science, Technology and Environment Malaysia (29-02-03-0001).

References

1. Maas WK (1994) Microbiol Rev 58:631–640
2. Mass WK, Clark AJ (1964) J Mol Biol 8:365–370
3. Tian G, Mass WK (1994) Mol Microbiol 13:599–608
4. Burke M, Merican AF, Sherratt DJ (1994) Mol Microbiol 13:609–618
5. Grandori R, Lavoie TA, Pflumm M, Tian G, Niersbach H, Mass WK, Fairman R, Carey J (1995) J Mol Biol 254:150–162
6. Van Duyne GD, Ghosh G, Mass WK, Sigler PB (1996) J Mol Biol 256:377–391
7. Sunnerhagen M, Nilges M, Otting G, Carey J (1997) Nat Struct Biol 4:819–829
8. Dion M, Charlier D, Wang H, Gigot D, Scavchenko A, Hallet JN, Glansdorf N, Sakanyan V (1997) Mol Microbiol 25:385–398
9. Jianping N, Sakanyan V, Charlier D, Glansdorf N, Van Duyne GD (1999) Nat Struct Biol 6:427–432
10. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM, McKenny K, Sutton G, Fitzhugh W, Field C, Gocayne JD, Scott J, Shirley R, Liu LI, Glodek A, Kelley JM, Weidman JF, Phillips CA, Spriggs T, Hedblom E, Cotton MD, Utterback TR, Hanna MC, Nguyen DT, Saudek DM, Brandon RC, Fine LD, Fritchman JL, Geoghagen NSM, Gnehm CL, McDonald LA, Small KV, Fraser CM, Smith HO, Venter JC (1995) Science 269:496–432
11. Lu CD, Houghton JE, Abdelal AT (1992) J Mol Biol 225:11–24
12. North AK, Smith MC, Baumberg S (1989) Gene 80:29–38
13. Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE (1982) J Mol Biol 161:269–288
14. Goodsell DS, Olson AJ (1990) Proteins: Struct Func Genet 8:195–202
15. Sotriffer CA, Flader W, Winger RH, Rode BM, Liedl KR, Varga JM (2000) Methods 20:280–291. DOI 10.1006/meth.1999.0922
16. Coutinho PM, Dowd MK, Reilly PJ (1997) Carbohydrate Res 297:309–324
17. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998) J Comput Chem 19:1639–1662
18. Greenstein JP, Winitz M (1961) Chemistry of the amino acids, vol 3. Wiley, New York
19. Boyar A, Marsh RE (1982) J Am Chem Soc 104:1995–1998
20. Dauber-Osguthorpe P, Roberts VA, Osgutthorpe DJ, Wolff J, Genest M, Hagler AT (1988) Proteins: Struct Func Genet 4:31–47
21. Merican AF, Asi AM, Rahman NA (2000) Asia Pacific J Mol Biol Biotechnol 8:13–25
22. Ni J, Sakanyan V, Charlier D, Glansdorff N, Van Duyne GD (1999) Nat Struct Biol 6:427–432
23. Hetényi C, van der Spoel D (2002) Protein Sci 11:1729–1737. DOI 10.1110/ps.0202302
24. Bitomsky W, Wade RC (1999) J Am Chem Soc 121:3004–3013
25. Tonmunphean S, Parasuk V, Kokpol S (2001) J Mol Model 7:26–33. DOI 10.1007/S008940100013