Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography A, 1066 (2005) 81–88
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

A theory of protein–resin interaction in hydrophobic interaction chromatography

A. Mahn*, G. Zapata-Torres, J.A. Asenjo

Centre for Biochemical Engineering and Biotechnology, Millennium Institute for Advanced Studies in Cell Biology and Biotechnology, Department of Chemical Engineering, University of Chile, Beauchef 861, Santiago, Chile

Received 18 June 2004; received in revised form 23 November 2004; accepted 5 January 2005
Available online 21 January 2005

Abstract

Docking simulations were performed in order to investigate surface area of interaction between several ribonucleases and a reduced model for the hydrophobic moiety used in Phenyl Sepharose™ using the program AutoDock 3.0. For each ribonucelase, 80 independent simulations with populations consisting of 100 random structures were performed and from these the most probable docked protein–ligand conformations were obtained. A new methodology was used to select the most probable conformations, based on qualitative and quantitative considerations. The interacting amino acids in each protein were identified. The average surface hydrophobicity of the interfacial zone (local hydrophobicity, LH) was determined. The LH showed a high correlation level ($r^2 = 0.99$) with the "hydrophobic contact area" (HCA) experimentally determined for the different ribonucleases as well as with the dimensionless retention time ($r^2 = 0.90$). This study allowed us to identify the zones on the protein surface most probably involved in protein retention in HIC, without tedious experimental work. Given the good correlation level obtained, this new methodology may constitute a novel approach that could be used to predict protein behavior in HIC.

Keywords: Molecular docking; Local hydrophobicity; Dimensionless retention time; Hydrophobic accessible area

1. Introduction

Hydrophobic interaction chromatography (HIC) is a powerful technique for protein separation, based on the reversible interaction between the hydrophobic zones of a protein's surface and the hydrophobic ligands of a chromatographic resin [1]. HIC is widely used in the downstream processing of proteins, as it provides separation properties complementary to other protein purification techniques such as ion-exchange chromatography, affinity chromatography or gel filtration chromatography [2]. The main protein property determining retention in HIC is hydrophobicity, which can be estimated as "average surface hydrophobicity" starting from the protein 3D-structure data and considering the hydrophobic contribution of the exposed amino acids [3,4]. A novel methodology to predict protein retention time in HIC starting from protein's average surface hydrophobicity has been proposed, with a reasonable degree of success [4,5]. However, some proteins did not follow the expected behavior. This was attributed to a heterogeneous distribution of the hydrophobic zones on the protein surface [5]. The surface hydrophobicity distribution of proteins related to retention in HIC has been investigated by Mahn et al. [6]. Based on a classical thermodynamic model [7], the contact area between the hydrophobic ligands of the HIC matrix and the protein when adsorbed (hydrophobic contact area, HCA) was experimentally determined. This parameter was found to give an idea of the surface hydrophobicity distribution of proteins. HCA correlated extremely well with the dimensionless retention time showed by different ribonucleases in HIC with salt gradient elution. However, a high number of experiments are necessary to determine this parameter.

* Corresponding author. Fax: +56 2 6991084.
E-mail address: amahn_2000@yahoo.es (A. Mahn).

0021-9673/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2005.01.016

The aim of this work was to identify the zone of a protein's surface which most probably interacts with the hydrophobic ligands of resins used in HIC. We carried out molecular docking simulations to identify the interaction zone. This identification would allow us to find any parameter that could be used to predict chromatographic behavior of proteins in HIC, reducing in this way the experimental work.

1.1. Molecular docking

Computational methods are increasingly being used in the identification and characterization of protein–ligand interactions. The ligand docking and the selection algorithms are commonly used in drug design as well as in biochemical process elucidation [8]. In the protein–ligand docking process, a huge number of degrees of freedom has to be taken into account for both molecules, as well as the combination of energetic forces that acts on them. Molecular docking consists of a conformational sampling procedure, in which different protein–ligand conformations are examined to find the correct one. The sampling procedure is normally based on methods such as genetic algorithms and Monte Carlo simulation, among others. Besides, the conformational sampling involves an energy function ("score function") used to evaluate the fitness between the protein and the ligand [9]. The molecular docking has three steps: identification of the binding sites, a search algorithm to efficiently perform the conformational sampling in the search space, and a score function [8].

In this work, we simulated the interaction between different ribonucleases of known three-dimensional structure and the hydrophobic ligand used in the resin Phenyl Sepharose^{TM}, using the program AutoDock 3.0.5^{TM} [10]. For each protein eight simulations were carried out, each of them consisting of ten grids, obtaining eighty possible conformations of the protein–ligand complex for each ribonuclease. Based on qualitative (location of the interaction zone) and quantitative (free energy of the complex) considerations, the most probable protein–ligand conformations were chosen. Once the interaction zone was identified, the local hydrophobicity (LH) was determined, considering the amino acid residues that belong to that zone and their exposure level, using a methodology similar to that proposed before [4].

1.2. The main factors that affect protein retention in hydrophobic interaction chromatography (HIC)

In previous work [5], we have demonstrated that the main factor affecting protein retention in HIC is a protein's hydrophobicity. On reference to the chromatographic conditions, we have investigated the effect of different chromatographic conditions on protein retention in HIC. A linear correlation has been found between a protein's retention time using different hydrophobic matrixes (butyl and phenyl sepharose) or different initial salt concentration in the elution buffer. It was concluded that the types of matrix do not affect the elution order of proteins, despite the hydrophobic moieties in these matrixes interact with protein in a different way [11]. A correction factor was obtained (Eq. (5)), which can be used to estimate a protein's retention time using phenyl sepharose, starting from that obtained using butyl sepharose [5].

In addition, it has been possible to compare protein retention time using ammonium sulphate and sodium chloride at similar ionic strength [5]. Proteins showed a very different behaviour. Selectivity was reduced when using sodium chloride and the elution order of proteins was indeed affected. It was concluded that ammonium sulphate allows a much more predictable behaviour of proteins in HIC, because this salt stabilises a protein's structure in solution [12].

Then, if ammonium sulphate is used to build the elution gradient, the dimensionless retention time of a protein in HIC using the matrix phenyl sepharose, can be estimated from that obtained with butyl sepharose.

In this work we propose to investigate if it is possible to identify the zone of a protein's surface that interacts with a hydrophobic resin used in HIC. The results obtained in the docking simulations will be analysed based on the chromatographic behaviour of proteins, using the experimental conditions that favour a protein's structural stability.

2. Experimental

2.1. Proteins

In the simulations we used the crystal structures of four different ribonucleases, which have been used in our previous paper [6]: RNAse A (PDB code 1AFU), RNAse S (PDB code 1RBC), RNAse T1 (PDB code 1RGC) and a variant of RNAse T1 (PDB code 1TRP). The spatial coordinates were retrieved from The Protein Data Bank [13].

In the chromatographic runs we used the ribonucleases mentioned before. Ribonuclease T1 wild type (1RGC) and the variant Y45W/W59Y (1TRP) were obtained by expressing both enzymes in *E. coli* strain DH5α. Competent cells were transformed with the corresponding plasmids. RNase variants were produced and purified after the protocol published by Grunert and coworkers [14]. Both plasmids were kindly donated by Prof. Dr. Ulrich Hahn (University of Hamburg, Germany). Ribonuclease A (1AFU), Ribonuclease S (1RBC) and Tris buffer were purchased from Sigma Chemical Co. (St. Louis, Mo, USA). Water prepared from a Milli-Q water cleaning system (Millipore, Bedford, MA, USA) and analytical-reagent grade ammonium sulfate (Merk) was used in the preparation of the buffers.

2.2. Hydrophobic ligand

The interaction between the crystal structure of the different RNAses and the hydrophobic ligand used in Phenyl Sepharose^{TM} was studied. This ligand consists of a phenyl group linked to the hydrophobic resin through a three-carbon

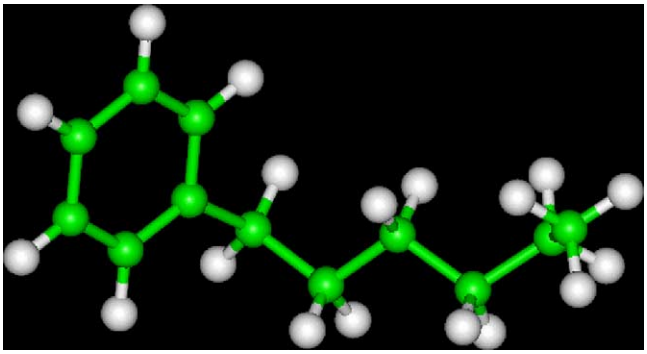

Fig. 1. Spatial coordinates of the hydrophobic ligand used in the phenyl sepharose HIC resin [14] as a simplified model, built using Insight II™.

chain and an ester bound. The ligand structure is shown in Fig. 1 [11]. The spatial coordinates of the ligand were elucidated using the program Insight II™, which gives a PDB file similar to the one that gives the spatial coordinates of the proteins under study.

2.3. Chromatographic runs

Gradient elution HIC experiments were carried out using a decreasing salt gradient, with a steepness of 7.5% B/min (a 10 column volume gradient). The initial eluent was Tris 20 mM pH 8.0 plus a maximum salt concentration of 2 M ammonium sulfate. The final eluent was Tris 20 mM pH 8.0 (buffer A). Samples, which contained 0.5 mg/mL protein, were injected through a 200 μL loop. Retention volume was recorded. All the runs were made in triplicate. All buffers were filtered through 0.22 μm Millipore filters after preparation and degassed with helium for 10 min.

The high-performance liquid chromatography system employed consisted of a FPLC (Pharmacia, Uppsala, Sweden). The chromatographic matrix used was Butyl Sepharose Fast Flow (a gift of Amersham Pharmacia Biotech, Uppsala, Sweden) packed in a 1 mL column. The experiments were performed at room temperature (23.5 °C), using a flow rate equal to 0.75 mL/min.

The chromatographic behavior of proteins was characterized by the parameter "dimensionless retention time" (DRT), using Eq. (1). In Eq. (1), t_R is the time corresponding to the peak maximum in the chromatogram, t_0 is the time corresponding to the start of the elution gradient, and t_f is the time corresponding to the end of the salt gradient.

$$\mathrm{DRT} = \frac{t_R - t_0}{t_f - t_0} \tag{1}$$

2.4. Determination of the hydrophobic contact area (HCA)

The hydrophobic contact area of each protein was determined starting from their isocratic retention factors obtained using different ammonium sulfate molalities [6]. The HCA was obtained from Eq. (2), where R is the universal gas constant and T is the absolute temperature. The value of the parameter C was obtained with linear regression with the respective data; the value of σ_s, which corresponds to the molal surface tension increment of ammonium sulfate, was obtained from literature, and was equal to 2.6×10^3 [dyn g/cm mol] [15].

$$C = \frac{(\mathrm{HCA})\sigma_s}{2.3RT} \tag{2}$$

2.5. Methods

Docking calculations were carried out using Autodock version 3.0.5 [10]. Three binding energy terms were taken into account in the docking step: the van der Waals interaction represented as a Lennard-Jones 12-6 dispersion–repulsion term, the hydrogen bonding represented as a directional 12-10 term, and the Coulombic electrostatic potential. Non-polar hydrogens were removed from the ligand, and their partial atomic charges were united with the bonded carbon atoms. The ligand was arbitrarily positioned at the protein's surface (using an Autodock script kindly provided by Dr. Jose Jaime Arbaldua) which in turn was divided in eight grids. Then the docking runs were performed using the Lamarckian genetic algorithm (LGA) [10] with grid sizes of $20 \times 20 \times 20$ (grid spacing 0.375 A), yielding 10 docked conformations by simulation. During the docking computation, free rotation was allowed about the side chain of the phenyl ligand. Parameters in Autodock were assigned default values. The resulting 10,000 docking orientations for each ligand produced by AutoDock 3.0.5 were analyzed by grouping them into clusters with similar ligand orientations. We employed a cluster analysis that classified structures as similar in 1 Å intervals of the ligand root mean square deviation.

The simulations were carried out in vacuum, but the program is able to estimate the binding energy of the complexes in solution, based on a thermodynamic cycle and using the Hess' law [16].

Given the high number of possible conformations obtained, it was necessary to make a selection of the most adequate one, for each protein. A new methodology was used to select the most probable conformation, based on qualitative and quantitative considerations.

2.6. Local hydrophobicity (LH) calculation

Once the most adequate protein–ligand conformations were chosen, the interaction zone was identified. Then the average surface hydrophobicity of this zone was estimated (local hydrophobicity, LH) using a methodology similar to that used before by our group [4,5] with the modified Eq. (3).

$$\phi_{\mathrm{surface}} = \frac{\sum(s_{\mathrm{aai}} \times \phi_{\mathrm{aai}})}{s_p} \tag{3}$$

In Eq. (3), ϕ_{surface} is the surface hydrophobicity, "i" indicates the different standard amino acids, s_{aai} is the solvent

accessible area occupied by amino acid "*i*", ϕ_{aai} is the hydrophobicity value assigned to amino acid "*i*", and s_p is the total solvent accessible area of the protein.

First, the residues belonging to the interaction zone were identified, considering a radii equal to 5 Å. To do this, the program Insight II™ was used. Starting form the PDBQS file given by AutoDock (converted to a PDB file), containing the spatial coordinates of each selected complex, the solvent accessible area of the interaction zone (s_{IZ}) and the solvent accessible areas of each residue in that zone (s_{aai}) were calculated, using the program GRASP [17]. Finally, the normalized amino acid hydrophobicity scale reported by Miyazawa and Jernigan [18,4] was used to assign a hydrophobicity value to each amino acid residue.

The modified Eq. (3) can then be expressed as follows:

$$\mathrm{LH} = \frac{\sum(s_{aai} \times \phi_{aai})}{s_{IZ}} \quad (4)$$

where LH is the average surface hydrophobicity of the interaction zone of the protein with the hydrophobic ligand (local hydrophobicity). s_{aai} is the solvent accessible area of each residue in the interaction zone. ϕ_{aai} is the amino acid hydrophobicity given by the normalized scale reported by Miyazawa and Jernigan [18], and s_{IZ} is the solvent accessible area of the interaction zone.

In this work we investigated if there was any correlation between the local hydrophobicity and the chromatographic behavior of the RNAses in HIC using salt gradient elution.

3. Results and discussion

In this work we carried out docking studies among different RNAses and the hydrophobic ligand used in phenyl sepharose. The purpose was to identify the most probable interaction zone of the proteins with a ligand. Because of the high number of possible orientations and conformations obtained for each protein, we selected the most representative according to the criterion described below.

3.1. Simulations and conformation selection

To select the most probable conformation of the complexes given by AutoDock we used quantitative and qualitative considerations. First, we choose the highest free energy conformations since our interest was focused on hydrophobic interactions. Then the chosen conformation was analyzed qualitatively based on the characteristics of interaction zone. In the case that in the chosen conformation the ligand was located in a pocket or concave zone, the docked complex was discarded and the conformation with the next free energy value was analyzed. This procedure was repeated until we found a conformation in which the interaction zone was located in a convex zone on the protein surface, preferably opposite to the protein's active site. This selection criterion was based on the knowledge that hydrophobic patches located in convex zones of a protein are more accessible to the hydrophobic ligands of the HIC resins [6]. Those hydrophobic patches located in opposite zone to the active site would have a greater effect on protein retention in HIC [19]. Fig. 2 shows the block diagram of the selection methodology proposed in this work.

One conformation of the protein–ligand complex was chosen for each of the eight grid simulations performed for each RNAse, based on the above criterion. In this way eight probable conformations were obtained for each complex, considering the whole protein surface. Finally, the same criterion was used to choose only one conformation for each protein.

The protein's interaction zone of the selected conformation for each RNAse is shown in Fig. 3. The closest amino acid residues that belong to this zone are identified with the three-letter codes. The amino acid residues that belong to each hydrophobic interaction zone are given in Table 1.

3.2. Calculation of local hydrophobicity (LH)

Once the most probable interaction zone of each RNAse was identified, the amino acid residues that belong to this zone were individualized. The program Insight II™ was used to visualize and isolate the interaction zone. PDB files were generated; they contain the spatial coordinates of the interest zone. Starting from these files and using the program GRASP [17], the solvent accessible area of the interaction zone (s_{IZ}) and the partial solvent accessible areas (s_{aai}) that correspond to each residue in that zone were determined. Then, using Eq. (4) the average surface hydrophobicity of the interaction zone (local hydrophobicity, LH) was estimated, considering the normalized Miyazawa and Jernigan [18,4] amino acid hydrophobicity scale.

Table 2 shows the local hydrophobicity obtained for each RNAse. For RNAse S and RNAse T1 variant the calculated

Table 1
Amino acid residues that belong to the selected interaction zone of the ribonucleases, within a radius equal to 5 Å from the center of the ligand

RNAse A (1AFU)	RNAse S (1RBC)	RNAse T1 (1RGC)	Rnase T1 variant (1TRP)
Asn 34	Glu 2	Asn 9	Ser 17
Leu 35	Ala 6	Tyr 11	Thr 18
Lys 37	Glu 9	Ile 61	Gln 20
Asp 38	Arg 10	Leu 62	Ala 21
Arg 39	Gln 11	Ser 63	Ala 22
Lys 41	His 12	Ser 64	Val 67
	Ala13	Asp 66	Asn 84
	Asp 14	Tyr 68	
	Ser 15	Ser 69	
	Met 25	Gly 70	
	Arg 29	Gly 71	
	Leu 47	Ser 72	
		Pro 73	
		Gly 74	
		Ala 75	
		Asp 76	

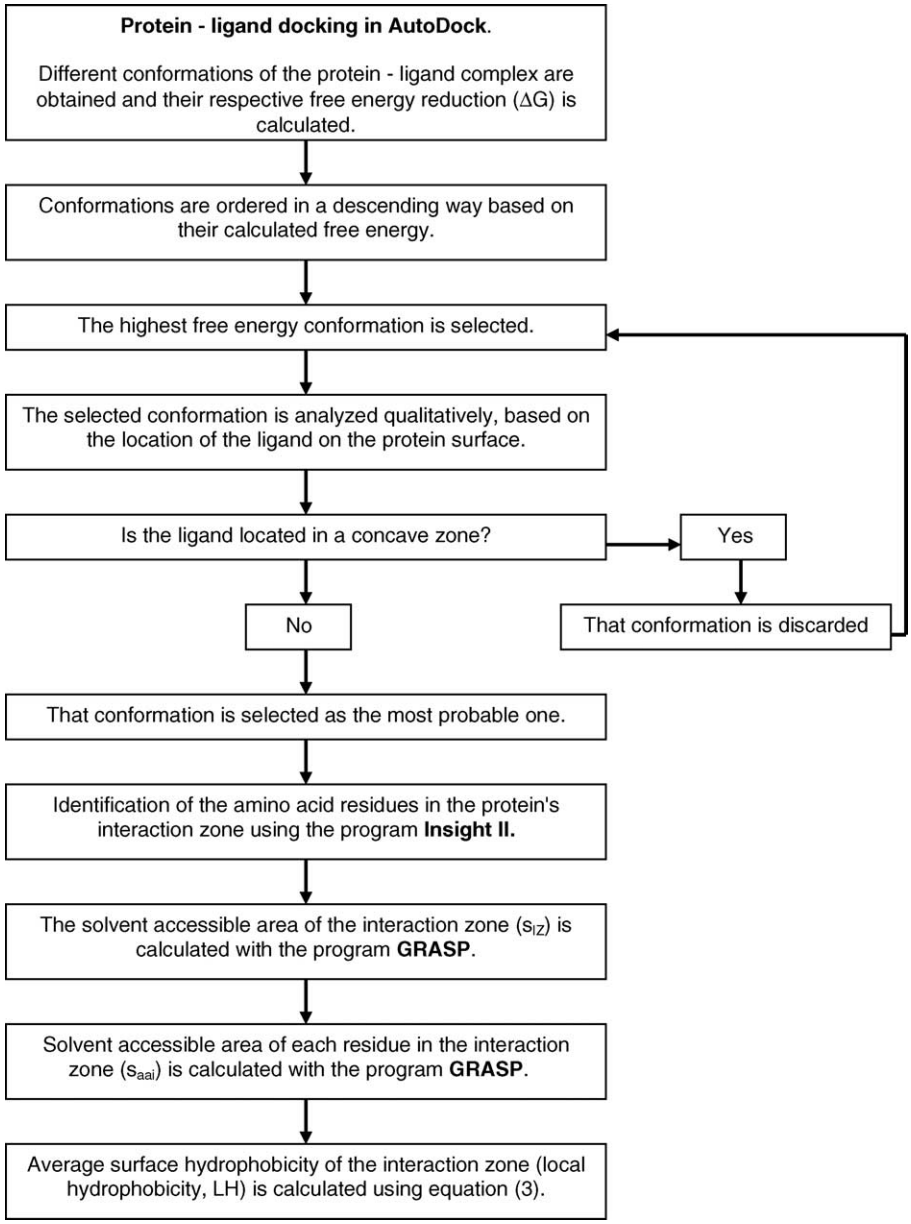

Fig. 2. Block diagram of the methodology used in the selection of the most probable protein–ligand complex conformation and calculation of local hydrophobicity (LH).

Table 2
Average surface hydrophobicity ($\phi_{surface}$) estimated after the methodology proposed before [4–6] and local hydrophobicity (LH) of the RNAses calculated with the methodology proposed in this paper

Protein	PDB ID[a]	$\phi_{surface}$ (−)[b]	LH (−)[c]
RNAse A	1AFU	0.230	0.160
RNAse S	1RBC	0.222	0.295
RNAse T1 *wild type*	1RGC	0.265	0.237
RNAse T1 variant	1TRP	0.269	0.335

[a] PDB ID is the file code given by The Protein Data Bank [13].
[b] $\phi_{surface}$ is the average surface hydrophobicity of proteins estimated by Eq. (3) [6].
[c] LH is the local hydrophobicity estimated by Eq. (4).

local hydrophobicity was higher than the average surface hydrophobicity ($\phi_{surface}$). This situation would account for the heterogeneous distribution of the hydrophobic patches on these proteins' surface, because there would be a zone where the hydrophobic residues are concentrated. Also, this zone would be located in an accessible zone for the ligands of a HIC resin.

On the other hand, RNAse A and RNAse T1 wild type showed a local hydrophobicity lower than their respective average surface hydrophobicity. This situation could be attributed to the less accessible location of the most hydrophobic zones of these proteins. Complex conformations which located the hydrophobic ligand on the most hydrophobic zones

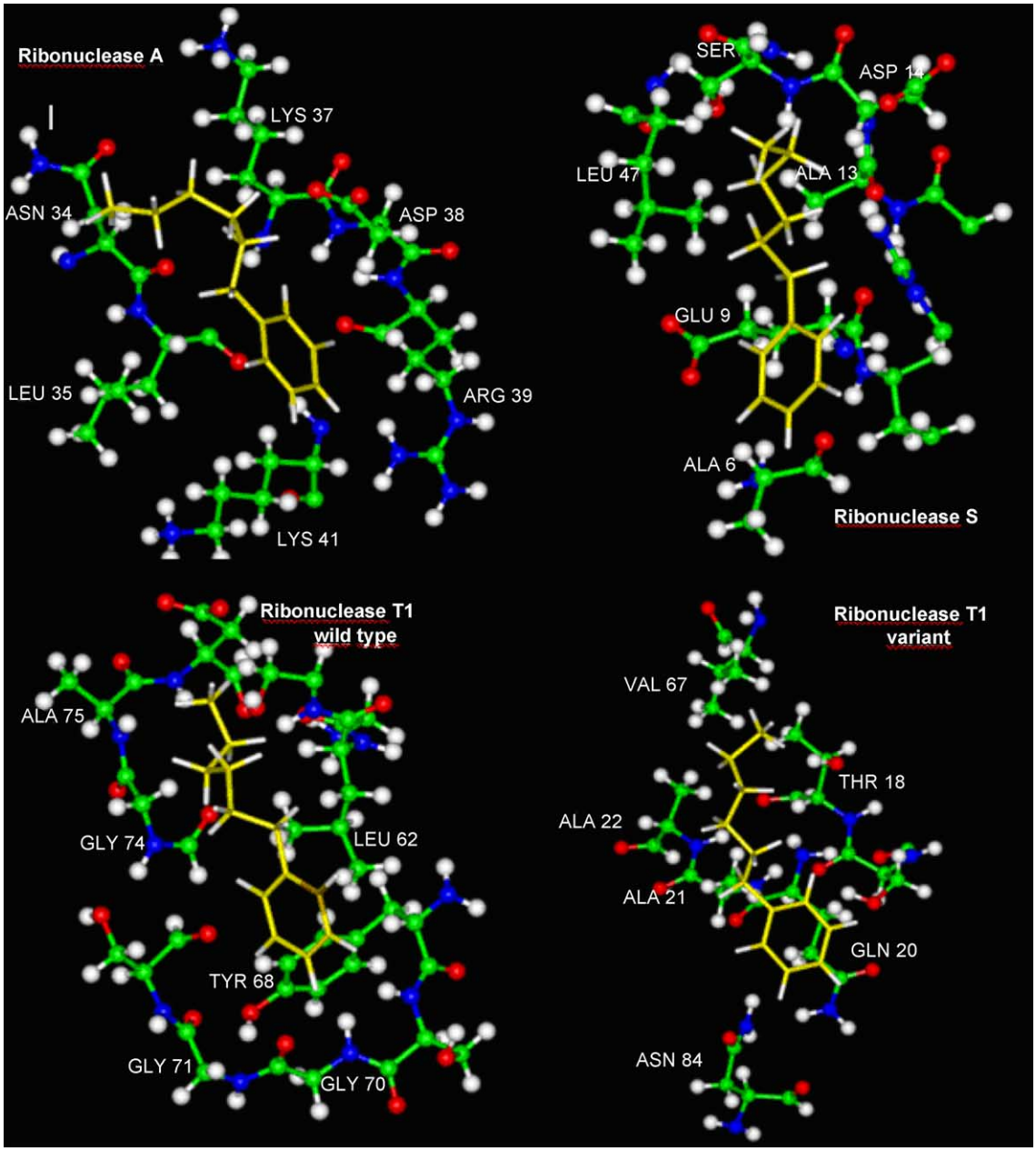

Fig. 3. Three dimensional diagrams of the docking between the hydrophobic ligand (colored in yellow) and the different ribonucleases, obtained with the program Insight II^{TM}. For simplicity, only the closest amino acid residues that belong to the selected interaction zone are indicated.

of these proteins were probably discarted during the qualitative analysis in the seletion procedure.

3.3. Estimation of protein retention time in HIC

In this work we investigated if it would be possible to predict the dimensionless retention time of the RNAses in HIC with salt gradient elution, starting form the knowlege of the proteins's structure. We considered the local hydrophobicity estimated for each protein to elucidate if there was any correlation between this magnitude and the DRT and the hydrophobic accessible area determined experimentally. The DRT of the RNAses in phenyl sepharose was estimated starting from that shown in butyl sepharose, using Eq. (5):

$$\mathrm{DRT}_{\mathrm{Phenyl\ sepharose}} = 0.92 \times \mathrm{DRT}_{\mathrm{Butyl\ sepharose}} + 0.13 \quad (5)$$

Eq. (5) corresponds to a "correction factor" which has been obtained by correlating the dimensionless retention time of 14 well-known proteins using two different hydrophobic matrixes: phenyl sepharose and butyl sepharose, under the same experimental conditions [5].

The dimensionless retention time of the RNAses in butyl sepharose have been determined experimentally [6].

Table 3
Local hydrophobicity (LH), hydrophobic contact area (HCA) and dimensionless retention time in butyl sepharose (DRT_B) and in phenyl sepharose (DRT_P) of the different RNAses

Protein	PDB ID	LH (−)[a]	HCA[b]	DRT_B[c]	DRT_P[d]
RNAse A	1AFU	0.160	64.5	0.239	0.348
RNAse S	1RBC	0.295	257.3	0.760	0.826
RNAse T1 *wild type*	1RGC	0.237	112.4	0.260	0.367
RNAse T1 variante	1TRP	0.335	386.8	0.382	0.479

[a] Local hydrophobicity estimated by Eq. (4).
[b] Hydrophobic contact area of the RNAses and the HIC resin butyl sepharose, reported before [6].
[c] Dimesionless retention time of the RNAses in butyl sepharose obtained experimentally [6].
[d] Dimensionless retention time of the RNAses in phenyl sepharose estimated from that obtained in butyl sepharose using Eq. (5) [5].

The hydrophobic accessible area (HCA) of a protein corresponds to the contact area between the stationary phase and the protein when attached to the HIC resin. In this paper we used the RNAses' HCA values reported before [6], obtained from isocratic HIC experiments. The HCA values have been determined based on a classical thermodynamic model, which describes protein retention due to electrostatic and/or hydrophobic interaction [7,6]. Table 3 shows LH, HCA and DRT of the four RNAses used in this paper.

Fig. 4 shows the correlation level between the local hydrophobicity of the RNAses and their HCA experimentally determined. A high correlation level between LH and HCA was obtained, with a determination coefficient (r^2) equal to 0.90 (Fig. 4a). It has to be noted that the HCA values come from experimental measurements, while the LH values are based exclusively on theoretical estimations starting from computer simulations. Additionally, the HCA values were obtained from HIC experiments using the resin butyl sepharose, while the hydrophobic ligand used in the simulations was that of phenyl sepharose. The type of matrix should not affect the LH of a protein, because the surface zone that most probably interacts with a hydrophobic ligand does not depend on the type of ligand. However, the type of matrix would indeed affect the intensity of the hydrophobic interaction, and thus the protein retention time, as we have demonstrated in previous work [5].

When RNAse A was not considered (Fig. 4b), the correlation level became higher ($r^2 = 0.99$), showing a close relation between LH and HCA. RNAse A slightly moved away from the expected behavior, possibly due to it's homogeneous surface hydrophobicity distribution [20]. In this case, the chromatographic behavior of RNAse A would be more affected by the average surface hydrophobicity than by a higher concentration of hydrophobic residues in a certain zone. Besides, the surface hydrophobicity of a homogeneous protein would be well represented by the average surface hydrophobicity. It could be expected that the parameter LH would represent adequately the chromatographic behavior in HIC of proteins with a heterogeneous surface hydrophobicity distribution.

These results let us suppose that it would be possible to estimate a protein's retention time in HIC starting form the local hydrophobicity, because of the close relation between HCA and DRT. It has been reported that it would be possible to predict a protein's retention time in HIC starting from the HCA value of that protein, but the problem was that a very high number of experiments are necessary to determine HCA [6]. Then, the main contribution of the present work is to give a first approach for the estimation of protein retention time using molecular docking tools and starting from knowledge of the crystal structure of proteins, reducing considerably the experimental work.

Fig. 5 shows the correlation level between LH and the dimensionless retention time of the RNAses. Considering the four RNAses (data not shown), a low correlation was found, with r^2 equal to 0.31. However, excluding RNAse S the correlation level improved considerably, obtaining a r^2 equal to

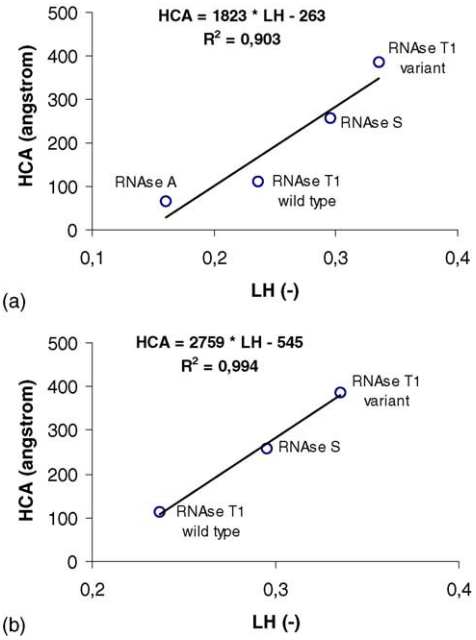

Fig. 4. Correlation level between local hydrophobicity (LH) and hydrophobic contact area (HCA) of the RNAses (a) considering the four RNAses; (b) excluding RNAse A. Individual points are the experimental data and the continuous line is the tendence line.

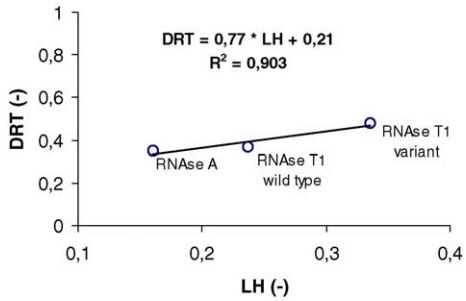

Fig. 5. Correlation level between local hydrophobicity (LH) and dimensionless retention time (DRT) in HIC of the RNAses excluding RNAse S. Individual points are the experimental data and the continuous line is the tendence line.

0.90 (Fig. 5). RNAse S presents an anomalous behavior in HIC [6] that has been attributed to its low stability and high dynamic flexibility [21,22]. It has to be noted that this protein would be an exceptional case, because similar structures are unusual [13].

Although the correlation between LH and DRT was not excellent, it is noticeable that two properties obtained independently in such different ways present an acceptable correlation level. The DRT was obtained experimentally, while LH is a magnitude that comes from theoretical estimations based on computer simulations and on knowledge of the three dimensional structure of the crystallized proteins. Elucidation of the most probable interaction zone of a protein with a hydrophobic ligand and the estimation of LH would probably allow an approximate estimation of protein retention time in HIC. This estimation does not consider the type of salt, making it less accurate because this may affect the protein conformation in solution and thus the chromatographic behavior [1,12]. To our empirical knowledge, we expect that the use of ammonium sulfate to build the elution gradient would allow better estimations using the methodology proposed in this paper.

On the other hand, LH estimation would be affected only in a minor way by the type of hydrophobic matrix used, as well as the estimation of HCA, conferring some generality to the methodology proposed in this work.

The results obtained in the present work let us propose that it would be possible to identify the zone on a protein's surface most probably involved in protein retention in HIC, reducing in this way the experimental work. Because of the good correlation level obtained between LH and DRT, this new methodology would constitute first approach to predict chromatographic behavior of proteins in HIC using molecular docking and considering the crystal structure of proteins.

Because of the low number of proteins used in the present study, we think that the proposed methodology should be tested with a higher number of proteins in a future. In addition, more experimental conditions should be analyzed in order to complete the present theory and to give more generality to this novel methodology.

4. Conclusion

Molecular docking simulations were carried out in order to investigate the hydrophobic interaction between different RNAses of known three-dimensional structure and the hydrophobic moiety in phenyl sepharose. A methodology was proposed to select the most probable interaction zone on a protein's surface. It was possible to identify the amino acid residues that belong to the selected interaction zone. A new parameter was found, named "local hydrophobicity", which correlated well with experimental parameters that represent chromatographic behavior of proteins in HIC: dimensionless retention time and hydrophobic contact area. We think that the methodology proposed in this paper could probably be used to predict protein retention time in HIC, avoiding tedious experimental work and thus facilitating a purification process design.

References

[1] W. Melander, Cs. Horvath, Arch. Biochem. Biophys. 183 (1977) 200.
[2] A.A. Shukla, K.M. Sunasara, R.G. Rupp, S.M. Cramer, Biotechnol. Bioeng. 68 (2002) 672.
[3] K. Berggren, A. Wolf, J.A. Asenjo, B.A. Andrews, F. Tjerneld, Biochim. Biophys. Acta, Protein Struct. M 1596 (2002) 253.
[4] M.E. Lienqueo, A.V. Mahn, J.A. Asenjo, J. Chromatogr. A 978 (2002) 71.
[5] A. Mahn, M.E. Lienqueo, J.A. Asenjo, Biotechnol. Bioeng., submitted for publication.
[6] A. Mahn, M.E. Lienqueo, J.A. Asenjo, J. Chromatogr. A 1043 (2004) 47.
[7] W. Melander, Z. El Rassi, Cs. Horvath, J. Chromatogr. 469 (1989) 3.
[8] B. MacConkey, V. Sobolev, M. Edelman, Curr. Sci. 83 (2002) 845.
[9] R. Wang, Y. Lu, S. Wang, J. Med. Chem. 46 (2003) 2287.
[10] G. Morris, D. Goodsell, R. Halliday, R. Huey, W. Hart, R. Belew, A. Olson, J. Comput. Chem. 19 (1998) 1639.
[11] Amersham Pharmacia Biotech. Hydrophobic Interaction Chromatography: Principles and Methods, Uppsala, Sweden, 2000, p. 278.
[12] T. Arakawa, S. Timasheff, Biochemistry 21 (1982) 6545.
[13] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, Nucleic Acids Res. 28 (2000) 235.
[14] H.P. Grunert, A. Zouni, M. Beineke, R. Quaas, Y. Georgallis, W. Saenger, U. Hahn, Eur. J. Biochem. 197 (1991) 203.
[15] A. Katti, Y.F. Maa, Cs. Horvath, Chromatographia 24 (1987) 646.
[16] G. Morris, D. Goodsell, R. Huey, W. Hart, S. Halliday, R. Belew, A. Olson, Autodock version 3.0.5 User's Guide (1998), p. 86 (http://www.scripps.edu/pub/olson-web/doc/Autodock/).
[17] A. Nicholls, K. Sharp, B. Honing, Proteins Struct. Funct. Genet. 11 (1991) 281.
[18] S. Miyazawa, R. Jernigan, Macromolecules 18 (1985) 534.
[19] J.L. Fausnaugh, F.E. Regnier, J. Chromatogr. 359 (1986) 131.
[20] D.D. Leonidas, R. Shapiro, L.I. Irons, N. Russo, K.R. Acharya, Biochemistry 36 (1997) 5578.
[21] R. Varadarajan, F.M. Richards, Biochemistry 31 (1992) 12315.
[22] P. Blackburn, S. Moore, Enzymes 15 (1982) 317.