Short communication

Prediction of retention time of cutinases tagged with hydrophobic peptides in hydrophobic interaction chromatography

M.E. Lienqueo [a,*], O. Salazar [a], K. Henriquez [a], C.R.C. Calado [b,c], L.P. Fonseca [b], J.M.S. Cabral [b]

[a] *Centre for Biochemical Engineering and Biotechnology, University of Chile, Santiago, Chile*
[b] *Centro de Engenharia Biológica e Química, Instituto Superior Técnico, 1049-001 Lisboa, Portugal*
[c] *Faculdade de Engenharia da Universidade Católica Portuguesa, Rio de Mouro, Portugal*

Abstract

Hydrophobic interaction chromatography (HIC) is an important technique for protein purification, which exploits the separation of proteins based on hydrophobic interactions between the stationary phase ligands and hydrophobic regions on the protein surface. One way of enhancing the purification efficiency by HIC is the addition of short sequences of peptide tags to the target protein by genetic engineering, which could reduce the need for extra and expensive chromatographic steps. In the present work, a methodology for predicting retention times of cutinases tagged with hydrophobic peptides in HIC is presented. Cutinase from *Fusarium solani pisi* fused to tryptophan–proline (WP) tags, namely $(WP)_2$ and $(WP)_4$, and produced in *Saccharomyces cerevisiae* strains, were used as model proteins. From the simulations, the methodology based on tagged hydrophobic definition proposed by Simeonidis et al. (Φ_{tagged}), associated to a quadratic model for predicting dimensionless retention times, showed small differences (RMSE < 0.022) between observed and estimated retention times. The difference between observed and calculated retention times being lower than 2.0% (RMSE < 0.022) for the two tagged cutinases at three different stationary phases, except for the case of cut_$(wp)_2$ in octyl sepharose–2 M ammonium sulphate. Therefore, we consider that the proposed strategy, based on tagged surface hydrophobicity, allows prediction of acceptable retention times of cutinases tagged with hydrophobic peptides in HIC.

Keywords: Hydrophobic interaction chromatography; Protein hydrophobicity; Retention time prediction; Hydrophobic peptide tags

1. Introduction

Hydrophobic interaction chromatography (HIC) has become widely used as a bioseparation technique for the laboratory and industrial-scale purification of biomolecules [1,2]. HIC exploits the separation of proteins based on hydrophobic interactions between the stationary phase ligands and hydrophobic regions on the protein surface.

Since hydrophobic interactions could be very selective, small differences in surface hydrophobicities between proteins can be used as an efficient means to perform protein purification easily [3]. The hydrophobicity of a protein can be modified with genetic engineering, such as site-directed mutagenesis or fusion of hydrophobic peptide tags. Examples of short hydrophobic tags that presented a strong effect on the relative hydrophobicity of the tagged protein are $(WP)_2$, $(WP)_4$, T_3, $(TP)_3$, T_3P_2, T_4, $(TP)_4$, T_6, T_6P_2, T_8 [4]. Of these tags, the most commonly used are tryptophane-containing tags (e.g. $(WP)_2$, $(WP)_4$) [5–9]. The advantage of fusion of a tag over site-direct mutagenesis is that the structure/function changes are minimised in relation to the original structure/function of the native protein. Furthermore, if necessary, the fused tag could be enzymatically removed after purification. An important advantage of hydrophobic polypeptide tags over traditional affinity tags is the possibility of exploring simple and much less expensive bioseparation materials.

A mathematical model that predicts the chromatographic behaviour in HIC of the tagged-protein could be very useful. By defining the relevant parameters that influence the chromatographic behaviour of the tagged-protein in relation to the native

* Corresponding author at: Department of Chemical and Biotechnology Engineering, University of Chile, Beauchef 861, Santiago, Chile.
Tel.: +56 2 9784709; fax: +56 2 6991084.
E-mail address: mlienque@ing.uchile.cl (M.E. Lienqueo).

one, the model can constitute a tool to design the optimal tag for the protein purification process. In this way, it is possible to design tags focused on improving the protein purification process, while the possible negative effect of a random design on the function or the production of the target protein is minimised.

The mechanism of protein binding to and elution from HIC adsorbents has been studied with the goal to increase the recovery and resolution [10–13]. Several authors have developed models to predict the chromatographic behaviour of proteins in HIC based on the surface hydrophobicity of proteins [14–20]. There are many ways to estimate the surface hydrophobicity of tagged proteins [14]. Bergreen et al. [21] proposed that each amino acid in the protein has a relative contribution to surface hydrophobicity $\Phi_{surface}$, and the amino acids in the tag presents a full exposed surface. The average surface hydrophobicity of proteins is calculated by Eq. (1).

$$\Phi_{surface} = \sum_{i=1}^{20} \left(\frac{S_{aai}}{S_p} \times \phi_{aai} \right) \quad (1)$$

where i ($i = 1, \ldots, 20$) indicates the 20 different amino acids, ϕ_{aai} the value of the hydrophobicity assigned to amino acid "i" using the Miyazawa-Jernigan scale [22], s_{aai} the total exposed area of the amino acid residue "i" in the tagged protein and s_p is the total surface of the tagged protein. These values were calculated using the Graphical Representation and Analysis of Structural Properties (Grasp) program [23].

The second method to estimate the surface hydrophobicity, designated tagged surface hydrophobicity (Φ_{tagged}), was proposed by Simeonidis et al. [24]. It calculates the surface hydrophobicity of the tagged protein as an average surface hydrophobicity of the original protein (without the tag) plus the hydrophobicity of the peptide tag. A fully exposed surface of the amino acids in the tag is assumed [25]. Then, the tagged surface hydrophobicity of proteins is calculated by Eq. (2).

$$\Phi_{tagged} = \sum_{i=1}^{20} \left(\frac{s_{aai}}{s_p} \times \phi_{aai} \right) + \sum_{k=1}^{20} \left(\frac{s_{tag\_aak} \times n_k}{s_p + \sum (s_{tag\_aak} \times n_k)} \times \phi_{aak} \right) \quad (2)$$

where n_k is the number of amino acids "k" in the tag and s_{tag_aak} is the fully exposed surface of amino acid "k" in the tag; [24,25]. In cases where this is not applicable, selecting to place the peptide tag on the other terminus (the N-terminus instead of the C-terminus of the protein product or vice versa) can solve this problem [24].

On the other hand, the methodology for predicting dimensionless retention time (DRT) of single proteins [26] and mixtures of proteins [27] in HIC, includes three steps: (i), obtaining the 3D structure of the original proteins using the Protein Data Bank File (PDB); (ii), determine the average surface hydrophobicity of the protein, $\Phi_{surface}$, considering that each amino acid has a relative contribution to surface properties, i.e., using Eq. (1); (iii), finally, a quadratic model is used to predict the dimensionless retention time (DRT) of proteins using the proteins average surface hydrophobicity, $\Phi_{surface}$. The model can be written as follows:

$$DRT = A\Phi_{surface}^2 + B\Phi_{surface} + C \quad (3)$$

where DRT is defined as:

$$DRT = \frac{t_R - t_0}{t_f - t_0} \quad (4)$$

where t_R is the time corresponding to the peak of the chromatogram, t_0 the time corresponding to the start of the salt gradient, and t_f is the time corresponding to the end of the salt gradient. The values of A, B and C, for several stationary phases are summarized in Table 1.

In previous publications [26,27] this methodology was tested and validated with an individual, standard, and recombinant mixture of proteins with well known three-dimensional structure. In this work, we extend the methodology for predicting dimensionless protein retention times in HIC of tagged proteins. Using this definition, we analyse which method of hydrophobic calculation describes more adequately the tagged-protein behaviour in hydrophobic interaction chromatography.

2. Experimental

2.1. Microorganism

The wild type (wt) cutinase, cutinase-$(WP)_2$ (fusion peptide composed of two tryptophan residues interspersed with two proline residues) and cutinase-$(WP)_4$ (fusion peptide composed of four tryptophan residues interspersed with four proline residues) producing *Saccharomyces cerevisiae* MM01 strains (Mata, leu2-3, *ura3*, *gal1*: URA3, *MAL-8*, *MAL3*, *SUC3*), containing the expression vectors pUR7320, pUR807, and pUR806, respectively, were constructed and provided by Unilever Research Laboratory, Vlaardingen, The Netherlands within the European Union project: Integrated bioprocess design for large scale production and isolation of recombinant proteins [28] (BIO4-CT96-0435).

Table 1
Constants and correlation coefficients for predicting the dimensionless retention time in HIC (adaptation of Lienqueo et al. [27][a])

Operating conditions	A	B	C	r^2
Butyl sepharose–2 M ammonium sulphate	−3.64	7.33	−1.17	0.97
Octyl sepharose–2 M ammonium sulphate	−11.81	11.59	−1.72	0.92
Phenyl sepharose–2 M ammonium sulphate	−26.95	19.40	−2.50	0.97

In this adaptation DRT is equal to 1 for an extremely hydrophobic protein, in this case the Neisserial surface protein A (NspA), which showed bigger surface hydrophobicity than protein ankyrin, protein previously used as the most hydrophobic protein.

[a] The quadratic model is:

$$DRT = A\phi^2 + B\phi + C. \quad (3)$$

Table 2
Comparison between predicted and observed protein retention times in HIC at different operating conditions

Protein	$DRT_{observed}$[a]	Prediction based on average surface hydrophobicity, $\Phi_{surface}$			Prediction based on tagged surface hydrophobicity, Φ_{tagged}		
		$\Phi_{surface}$[b]	$DRT_{calculated}$[c]	Deviation[d] (%)	Φ_{tagged}[e]	$DRT_{calculated}$[f]	Difference[d] (%)
Butyl sepharose (4FF) 2 M ammonium sulphate							
Cut_$(wp)_2$	0.88	0.317	0.79	10.36	0.338	0.89	2.04
Cut_$(wp)_4$	0.98	0.335	0.88	10.39	0.379	1.00	1.75
Average deviation				10.37			1.46
RMSE[g]				0.135			0.022
Octyl sepharose (4FF) 2 M ammonium sulphate							
Cut_$(wp)_2$	0.97	0.317	0.77	20.6	0.338	0.85	12.20
Cut_$(wp)_4$	1.00	0.335	0.84	15.9	0.379	0.89	1.90
Average deviation				18.2			7.10
RMSE[g]				0.256			0.163
Phenyl sepharose (6FF, high sub) 2 M ammonium sulphate							
Cut_$(wp)_2$	1.00	0.317	0.94	6.0	0.338	0.98	2.00
Cut_$(wp)_4$	1.00	0.335	0.98	2.0	0.379	0.99	1.00
Average deviation				4.0			1.50
RMSE[g]				0.063			0.022

[a] Observed dimensionless retention times.
[b] Average surface hydrophobicity calculated using Eq. (1).
[c] Estimated dimensionless retention times calculated using average surface hydrophobicity, Eq. (3) and constants of Table 1.
[d] Difference $= \frac{|DRT_{observed} - DRT_{calculated}|}{|DRT_{observed}|} \times 100$.
[e] Tagged hydrophobicity calculated using Eq. (2).
[f] Estimated dimensionless retention times calculated by using tagged hydrophobicity, Eq. (3) and constants of Table 1.
[g] RMSE: root mean squared error.

2.2. Cultivation and cell harvesting

Cultivation, cell harvesting and activity assay was carried out as described in Calado et al. [29].

2.3. Hydrophobic interaction chromatography

HIC was performed in a Pharmacia Fast Protein Liquid Chromatography (FPLC) system. HR 5 columns (50 mm × 5 mm i.d.) were filled with approximately 1 ml sorbent (butyl sepharose 4FF, octyl sepharose 4FF and phenyl sepharose 6FF by Amersham-Pharmacia). The experiments were performed at room temperature, using a flow rate equal to 0.75 ml/min. The absorbance of the effluent was monitored with an UV detector at 280 nm. Five-hunderd microliter fractions were collected and the enzyme activity and protein concentration was measured over the entire chromatogram. Retention time of cutinases (RT) was recorded and dimensionless retention time (DRT) was calculated using Eq. (4).

Elution was obtained by a decreasing concentration gradient of analytical-reagent grade ammonium sulphate for 10 ml (i.e. 10 column volumes). The initial eluent was 20 mM Bis–Tris, pH 7.0, plus 2 M ammonium sulphate (solvent A). The final eluent was 20 mM Bis–Tris, pH 7.0 (solvent B). The gradient steepness used was 7.5% B/min.

2.4. Determination of the hydrophobicity of proteins

Determination of the hydrophobicity of proteins was carried out as described in Lienqueo et al. [27].

3. Results and discussion

The proposed methodology was validated using two different mutants of extra cellular cutinases: cutinase-$(WP)_2$ and cutinase-$(WP)_4$ in three different resins: butyl sepharose–2 M ammonium sulphate, octyl sepharose–2 M ammonium sulphate, and phenyl sepharose–2 M ammonium sulphate. Table 2 shows observed dimensionless retention times; predicted dimensionless retention times calculated using the average surface hydrophobicity and the tagged surface hydrophobicity models, and the variation between observed and calculated dimensionless retention times for two tagged cutinases at three different stationary phases.

It was observed that both hydrophobicity definitions result in DRT estimations inside the 95% confidence intervals (data not shown) in all the sorbents under study. However, the difference from the estimated DRT using the tagged hydrophobicity designation in relation to the experimental DRT were under 2.0% (RMSE < 0.022) in all stationary phases under study, except for the case of cut_$(wp)_2$ in octyl sepharose–2 M ammonium sulphate, where the deviation was 12.2% (RMSE < 0.163). On the other hand, when using the average surface hydrophobicity definition, the difference between the experimental and calculated DRT was between 2.0% and 20.6%; where the biggest divergence was for the case of cut_$(wp)_2$ in octyl sepharose–2 M ammonium sulphate.

Since the tagged hydrophobicity definition considers primarily the influence of the tag (the second term of Eq. (2)), and based on the results obtained from the comparison of both hydrophobicity definitions in different resins, we suggest that

the interaction between tagged-protein and the chromatographic ligand occurs in or near the tagged zone.

4. Conclusions

The methodology for predicting dimensionless retention time in hydrophobic interaction chromatography, using hydrophobic descriptions (Φ) and a "quadratic model" was applied to tagged-cutinase. In general small differences (RMSE < 0.022) between observed and estimated retention times was obtained when applying both the average surface hydrophobicity and the tagged surface hydrophobicity definitions. However, the methodology based on tagged hydrophobic definition (Φ_{tagged}) proposed by Simeonidis et al. [28], which assumes that amino acids in the tag have a fully exposed surface, proved to be more adequate, since it presented a lower divergence between predicted and experimental retention times (RMSE < 0.022) for the two tagged cutinases evaluated in three different resins. Therefore, we consider that the proposed strategy, based on tagged surface hydrophobicity, allows prediction of acceptable retention times of cutinases tagged with hydrophobic peptides in HIC.

Acknowledgments

The authors would like to acknowledge the financial support of Fondecyt 1030668, the Programa de Cooperación Científica Internacional GRICES/CONICYT 2002-6-152 and Proyecto Enlace ENL06/14 of the Departamento de Investigación, Universidad de Chile. Cecília Calado acknowledges a fellowship (SFRH/BPD/11626/2002) from program POCTI-Formar e qualificar-Medida 1.1, Ministério da Ciência e Tecnologia of Portugal. The authors also wish to thank Drs. M. Egmond, A. Fellinger, and M. Mannesse of Unilever Research Laboratory for providing the yeast strains.

References

[1] J.A. Queiroz, C.T. Tomaz, J.M. Cabral, J. Biotechnol. 87 (2001) 143.
[2] M.M. Diogo, J.A. Queiroz, G.A. Monteiro, S.A. Martins, G.N. Ferreira, D.M. Prazeres, Biotechnol. Bioeng. 68 (2000) 576.
[3] S. Fexby, A. Nilsson, G. Hambraeus, F. Tjerneld, L. Bulow, Biotechnol. Prog. 20 (2004) 793.
[4] S. Fexby, L. Bulow, Trends Biotechnol. 22 (2004) 511.
[5] K. Berggren, A. Veide, P.A. Nygren, F. Tjerneld, Biotechnol. Bioeng. 62 (1999) 135.
[6] N. Bandmann, E. Collet, J. Leijen, M. Uhlen, A. Veide, P.A. Nygren, J. Biotechnol. 79 (2000) 161.
[7] A. Rodenbrock, K. Selber, M.R. Egmond, M.R. Kula, Bioseparation 9 (2000) 269.
[8] S. Fexby, L. Bulow, Protein Expr. Purif. 25 (2002) 263.
[9] A. Collen, M. Ward, F. Tjerneld, H. Stalbrand, J. Biotechnol. 87 (2001) 179.
[10] J. Chen, Y. Sun, J. Chromatogr. A 992 (2003) 29.
[11] H. Jennissen, Meth. Mol. Biol. 305 (2005) 81.
[12] C. Machold, K. Deinhofer, R. Hahn, A. Jungbauer, J. Chromatogr. A 972 (2002) 3.
[13] R. Hahn, K. Deinhofer, C. Machold, A. Jungbauer, J. Chromatogr. B: Anal. Technol. Biomed. Life Sci. 790 (2003) 99.
[14] M. Lienqueo, A. Mahn, Chem. Eng. Technol. 28 (2005) 1326.
[15] M. Lienqueo, A. Mahn, G. Navarro, J. Salgado, T. Perez-Acle, I. Rapaport, J. Asenjo, J. Mol. Recogn. 19 (2006) 260.
[16] H.P. Jennissen, A. Demiroglou, J. Chromatogr. A 1109 (2006) 197.
[17] A. Mahn, G. Zapata-Torres, J.A. Asenjo, J. Chromatogr. A 1066 (2005) 81.
[18] A. Ladiwala, F. Xia, Q. Luo, C.M. Breneman, S.M. Cramer, Biotechnol. Bioeng. 93 (2006) 836.
[19] F. Xia, D. Nagrath, S.M. Cramer, J. Chromatogr. A 1079 (2005) 229.
[20] A. Mahn, M.E. Lienqueo, J.A. Asenjo, J. Chromatogr. B (2006).
[21] K. Berggren, A. Wolf, J.A. Asenjo, B.A. Andrews, F. Tjerneld, Biochim. Biophys. Acta 1596 (2002) 253.
[22] S. Miyazawa, R. Jernigan, Macromolecules 18 (1985) 534.
[23] A. Nicholls, K.A. Sharp, B. Honig, Proteins 11 (1991) 281.
[24] E. Simeonidis, J.M. Pinto, M.E. Lienqueo, S. Tsoka, L.G. Papageorgiou, Biotechnol. Progr. 21 (2005) 875.
[25] C. Chothia, J. Mol. Biol. 105 (1976) 1.
[26] M.E. Lienqueo, A. Mahn, J.A. Asenjo, J. Chromatogr. A 978 (2002) 71.
[27] M.E. Lienqueo, A. Mahn, L. Vasquez, J.A. Asenjo, J. Chromatogr. A 1009 (2003) 189.
[28] C.R. Calado, M. Mannesse, M. Egmond, J.M. Cabral, L.P. Fonseca, Biotechnol. Bioeng. 78 (2002) 692.
[29] C.R.C. Calado, M.A. Taipa, J.M.S. Cabral, L.P. Fonseca, Enzyme Microb. Technol. 31 (2002) 161.