Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography A, 1009 (2003) 189–196
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

Methodology for predicting the separation of proteins by hydrophobic interaction chromatography and its application to a cell extract

M. Elena Lienqueo*, Andrea Mahn, Liliana Vásquez, Juan A. Asenjo

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

Abstract

Hydrophobic interaction chromatography (HIC) is widely used in the downstream processing of proteins. Resolution of HIC is very good, but sometimes not as high as expected. Resolution values could be increased if good operating conditions were selected. In this paper we present a methodology for selecting good operating conditions. First, it is necessary to predict the dimensionless retention time (DRT) of each protein in the mixture. We use a correlation such as $DRT = A + B\phi + C\phi^2$, where ϕ is the superficial hydrophobicity of the protein, which is calculated considering the hydrophobicity of the superficial amino acids using the Miyazawa–Jernigan scale. Considering that there was little interaction amongst proteins in a mixture at the concentrations investigated (2 g/l of each protein), the behaviour of the proteins in the mixture was considered to be similar to that of the individual proteins. Using simulations it was possible to test different operating conditions for the purification of a target protein from a mixture of proteins and it was possible to select ideal conditions. The methodology developed was also tested for the purification of a recombinant protein from a fermentation extract of yeast producing human superoxide dismutase and the results have been satisfactory.

Keywords: Hydrophobic interaction chromatography; Retention time prediction; Optimisation; Hydrophobicity; Proteins

1. Introduction

Hydrophobic interaction chromatography (HIC) is an important technique for protein purification; it exploits the hydrophobic nature of the solute (proteins) and is often used in combination with ion-exchange chromatography and gel permeation chromatography. HIC involves the separation of protein molecules owing to the differential interaction of these molecules with hydrophobic sites on the surface of a solid support [1]. In the separation process, hydrophobic patches on the protein interact with hydrophobic molecules immobilized on the hydrophilic solid-phase surface [2].

HIC adsorbents have been used for purifying a variety of biomolecules such as serum proteins [3] membrane-bound proteins [4], nuclear proteins [5], receptors [6], cells [7], and recombinant proteins [8,9] in research and industrial laboratories.

Protein binding to HIC adsorbents is promoted by moderately high concentrations of anti-chaotropic salts, which also have a stabilising influence on protein structure. Elution is achieved by a linear or stepwise decrease in the concentration of salt in the adsorption buffer. Recovery and resolution levels of

*Corresponding author. Fax: +56-2-699-1084.
E-mail address: mlienque@cec.uchile.cl (M.E. Lienqueo).

0021-9673/03/$ – see front matter
doi:10.1016/S0021-9673(03)00924-5

HIC are often satisfactory. One way of increasing resolution is by testing different operating conditions on an experimental basis. Jennissen [2] proposed to use the critical hydrophobicity method; this method involves three basic steps: (i) selection of an appropriate alkyl chain length; (ii) determination of the critical surface concentration of alkyl residues (critical hydrophobicity); and (iii) determination of the minimal salt concentration necessary for a complete adsorption of the protein.

Another way is by trying to predict the protein retention time under different operating conditions and choosing the best option ''in silico''. We have proposed a methodology to predict the behaviour of a single protein in HIC [10] and in this paper we present the application of this methodology by testing the purification of recombinant proteins from a real broth. In addition, we evaluated the possible interactions between proteins when a mixture of proteins is adsorbed on a HIC matrix.

2. Theory

2.1. Dimensionless retention time

Dimensionless retention time (DRT) was defined as [11]:

$$\mathrm{DRT} = \frac{t_R - t_0}{t_f - t_0} \quad (1)$$

where t_R is the time corresponding to the peak of the chromatogram, t_0 is the time corresponding to the start of the salt gradient, and t_f is the time corresponding to the end of the salt gradient.

DRT is equal to 1 for an extremely hydrophobic protein (in this case the protein ankyrin, a membrane protein, which showed the maximum surface hydrophobicity).

2.2. Prediction of DRT in HIC

In previous work we proposed a methodology for predicting DRT of single proteins in HIC [10]. The methodology has three steps:

(1) It is necessary to know the three-dimensional structure of the target protein, using protein data bank (PDB) files.

(2) Then, it is necessary to calculate the surface hydrophobicity of the protein. It was evaluated using the methodology proposed by Berggren et al. [12]. This methodology assumes that each amino acid on the surface of a protein has a relative contribution to the surface properties, then:

$$\phi_{surface} = \sum(\phi_{aai} r_{aai}) \quad (2)$$

where $\phi_{surface}$ is the calculated value of the surface hydrophobicity for a given protein, i ($i = 1, \ldots, 20$) indicates the 20 different amino acids, ϕ_{aai} is the value of the hydrophobicity assigned to amino acid ''i'' using the Miyazawa–Jernigan scale [13]. r_{aai} is the relative surface area exposed for each amino acid ''i'' on the surface, and it is defined as:

$$r_{aai} = \left(\frac{s_{aai}}{s_p}\right) \quad (3)$$

where s_{aai} is the total exposed area of the amino acid residue ''i'' in the protein and s_p is the total surface of the protein. These values were calculated using the graphical representation and analysis of structural properties (Grasp) [14] program.

(3) Finally, by using a simple quadratic model we are able to predict chromatographic behaviour of proteins in HIC starting from the protein's hydrophobicity. The model is [10]:

$$\mathrm{DRT} = A\phi_{surface}^2 + B\phi_{surface} + C \quad (4)$$

where $\phi_{surface}$ is the hydrophobicity value calculated using Eq. (2), and A, B and C are constants for each set of operating conditions.

Table 1 shows the correlations and correlation coefficients of the models using two different operating conditions. The methodology proposed is summarised in Table 2.

In the previous paper [10] this methodology was tested with individual standard proteins, in the current work we tested this methodology with mixtures of standard proteins and a cell extract.

3. Experimental

3.1. Materials

Six proteins of known three-dimensional structure

Table 1
Correlations and correlation coefficients for predicting the single protein retention time in HIC [10,17]

Operating conditions	Correlation		Model deviation (%)
Phenyl Sepharose 6FF–2.0 *M* ammonium sulfate	$DRT_{PS\text{-}2M} = -12.14*\phi^2_{surface} + 12.07\phi_{surface} - 1.74$	(5)	7.3
Butyl Sepharose 4FF–2.0 *M* ammonium sulfate	$DRT_{BS\text{-}2M} = 10.02*\phi^2_{surface} + 0.54\phi_{surface} - 0.38$	(6)	7.9

Note: $\phi_{surface}$ is the surface hydrophobicity of the protein calculated using Eq. (2).

were used: conalbumin (1OVT), ovalbumin (1OVA), chymotrypsinogen A (2CHA), chicken lysozyme (2LYM), α-lactalbumin (1A4V), from Sigma (St. Louis, MO, USA). Thaumatin (1THV) was a gift of 4F Nutrition (Northallerton, UK). Water, prepared from a Milli-Q water cleaning system (Millipore, Bedford, MA, USA) and analytical-reagent grade ammonium sulfate were used in the preparation of the eluent. Protein solutions were prepared to contain approximately 2.0 mg/ml dissolved in the initial eluent. All protein solutions were filtered through 0.22-μm Millipore filters.

3.2. Equipment

The high-performance liquid chromatography system employed consisted of a fast protein liquid chromatography (FPLC) system (Pharmacia, Uppsala, Sweden) equipped with a 500-μl injection loop. The chromatographic columns were 1 ml Phenyl Sepharose 6FF and Butyl Sepharose 4FF (donated by Amersham Pharmacia Biotech, Uppsala, Sweden). The experiments were performed at room temperature, using a flow-rate equal to 0.75 ml/min and 10 column volumes (CVs). Finally, retention times (t_R) were recorded.

3.3. Buffer

Elution was obtained by a decreasing gradient of ammonium sulfate. The initial eluent was 20 m*M* Bis–Tris, pH 7.0 plus 1 or 2 *M* ammonium sulfate (solvent B). The final eluent was 20 m*M* Bis–Tris, pH 7.0 (solvent A). The gradient steepness used was 7.5% B/min. All buffers were filtered through 0.22-μm Millipore filters after preparation, and degassed with helium for 10 min.

3.4. Fermentation conditions of human superoxide dismutase (SOD) from Saccharomyces cerevisiae rhSOD 2060 411 SGA 122 (MATα, leu2)

Recombinant *S. cerevisiae* rhSOD 2060 411 SGA 122 (MATα, leu2) were cultivated following the methodology described by Gonzalez et al. [15]. The inoculum was prepared by initially transferring colonies from a solid medium into a 125-ml flask containing 25 ml of a complex medium. Flasks were incubated for 14 h, and the culture was harvested using the pellet in the inoculation of the 500-ml flask containing 150 ml of a mineral salt medium with leucine and copper.

Human superoxide dismutase purification involved

Table 2
Methodology for predicting a protein's retention time in HIC [10]

Step
(1) Know the three-dimensional structure of the target protein
(2) Calculate the surface hydrophobicity of the protein, using the equation: $\phi_{surface} = \Sigma(\phi_{aai} r_{aai})$
ϕ_{aai} is the value of the hydrophobicity assigned to amino acid "*i*" using the Miyazawa–Jernigan scale
r_{aai} is the relative surface area exposed for each amino acid "*i*" on the surface
(3) Using a simple quadratic model to predict the chromatographic behaviour of proteins in HIC. The model is: $DRT = A\phi^2_{surface} + B\phi_{surface} + C$

cell disruption by a bead mill. Then the cell extract was centrifuged and the proteins were concentrated by precipitation with 80% ammonium sulfate. After that, as the first chromatographic step, the proteins were injected in a 1 ml Phenyl Sepharose 6FF column with a decreasing salt gradient, from 2.0 to 0 *M* ammonium sulfate in 10 ml (called Phenyl Sepharose 6FF, 2 *M* ammonium sulfate). Fractions were collected and analysed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE, 12%), where molecular mass markers and human superoxide dismutase standards were included. Human superoxide dismutase spots were analysed using SigmaGel (Jandel Scientific, Sausalito, CA, USA).

3.5. Determination of the hydrophobicity of proteins

The proteins were characterised based on their amino acid sequences. The program, Grasp [14], was used to visualise protein surfaces and to calculate the accessible surface area of single residues in a protein. The program takes a PDB file as input (http://www.rcsb.org/pdb) and a probe radius value, the default value 1.4 Å representing a water molecule. Then, the hydrophobicity of each standard monomeric protein (ϕ) was calculated using the superficial area of each amino acid in the protein and the Miyazawa–Jernigan hydrophobicity scale for the amino acids.

4. Results and discussion

4.1. Study of the possible interactions between proteins present in a mixture adsorbed on HIC

We have previously proposed a methodology for predicting a single protein retention time in HIC [10]. Presently, our aim is to extend this methodology to a mixture of proteins. Hence, interactions between adsorbed proteins, at a relatively low protein concentration, 2 g/l, were studied. There is no evidence from literature that adsorption on HIC of multicomponent mixtures could affect the behaviour of each component at the concentrations used in the present work. At higher concentrations possible overloading and/or displacement effects might be seen.

The retention time of more than 20 mixtures of standard proteins were investigated. The mixtures were selected considering that peaks do not overlap too much. Most of them (those that do not overlap much) are shown in Table 3 for the matrix Phenyl Sepharose 6FF, 1 *M* ammonium sulfate concentration (0.6% average difference, maximal difference value 1.5%). For Phenyl Sepharose 6FF, 2 *M* ammonium sulfate the average difference was much lower (0.6% average, maximal value 1.5%) and also for Butyl Sepharose 4FF, 1 *M* ammonium sulfate (0.7% average, maximal value 1.8%) as shown in Table 4.

The possible interactions between proteins, at a concentration of 2 g/l, were evaluated by determining the differences in the retention time between the protein injected alone and in a mixture. On average, the deviation was 1.6% (see Table 4). A typical average situation is shown in Fig. 1. The biggest deviation is shown in Fig. 2 (lysozyme and chymotrypsinogen A); where the deviation was equal to 8.7% for lysozyme (only three values where above 4%). For this last case (lysozyme), it is possible that some conformational changes occur on lysozyme as it has been reported in the literature [16]; however, this situation only happens at Phenyl Sepharose 6FF, 1 *M* ammonium sulfate.

On the other hand, the reproducibility of the DRT measurement was 2%. Bearing in mind that the global average deviation is 1.6%, we considered that the possible interaction between adsorbed proteins does not have a significant effect in the retention time of each protein in the mixture under the conditions studied, Phenyl Sepharose 6FF, 1 and 2 *M* ammonium sulfate and Butyl Sepharose 4FF, 1 *M* ammonium sulfate.

4.2. Validation of the methodology of purification of recombinant proteins

We tested the methodology described previously (Table 2) for the purification by HIC of recombinant human SOD from *S. cerevisiae* extract as the first chromatographic step.

The human superoxide dismutase extract was

Table 3
Comparison between observed retention times of proteins injected in a mixture and individually in Phenyl Sepharose 6FF, 1 *M* ammonium sulfate

Mixture of proteins	t_R in mixture (min)	Individual t_R (min)	Difference (%)
Lysozyme	9.8	10.3	5.3*
Thaumatin	14.5	14.6	0.8
Lysozyme	9.4	10.3	8.7*
Chymotrypsinogen A	15.6	15.8	1.0
Ovalbumin	9.9	10.2	3.6
α-Lactalbumin	19.2	19.9	3.6
Chymotrypsinogen A	15.7	15.8	0.6
Conalbumin	9.3	8.6	7.0*
Conalbumin	8.4	8.6	3.2
α-Lactalbumin	20.4	19.9	2.5
Conalbumin	8.5	8.6	2.0
Thaumatin	14.5	14.6	0.7
Ovalbumin	10.1	10.2	1.4
Thaumatin	14.4	14.6	1.7
Average deviation			3.6

* Value above 4%.

loaded on Phenyl Sepharose 6FF, 2 *M* ammonium sulfate. Elution was achieved with a linear gradient of decreasing ammonium sulfate over 10 column volumes. The retention time (t_R) was registered, DRT was calculated using Eq. (1), and fractions were collected and analysed by SDS–PAGE. The chromatogram obtained is shown in Fig. 3 and that corresponding to the SDS gel is shown in Fig. 4. As seen in Fig. 4, the human superoxide dismutase extract contains several different proteins; however, the main proteins are human superoxide dismutase and one main contaminant (M_r approximately 40 000 Da).

On the other hand, the hydrophobicity of the human superoxide dismutase was calculated using the methodology proposed, considering that the PDB file of human superoxide dismutase is 1AZV. After that, using Eq. (5), we predicted the DRT at this operating condition. The predicted DRT of 0.57 (Table 5) corresponds to 10.7 ml in Fig. 4. The

Table 4
Deviation in retention time between proteins injected individually and in a mixture

Operating conditions	Average difference (%)	Maximum difference (%)
Phenyl Sepharose 6FF, 1 *M* ammonium sulfate	3.6	8.7
Phenyl Sepharose 6FF, 2 *M* ammonium sulfate	0.6	1.5
Butyl Sepharose 4FF, 1 *M* ammonium sulfate	0.7	1.8
Global average deviation	1.6	
Reproducibility of DRT measurements*	2.0	

* Reproducibility of DRT measurement was calculated as an average of the deviation between two DRT measurements under the same conditions (DRT_1, DRT_2), then:

$$\text{Reproducibility of DRT measurement} = \frac{\left[\sum_{i=1}^{N}\left(\frac{|DRT_1 - DRT_2|}{\frac{DRT_1 + DRT_2}{2}}\right)_i\right]}{N} \cdot 100$$

where N is the number of different DRT measurements analysed.

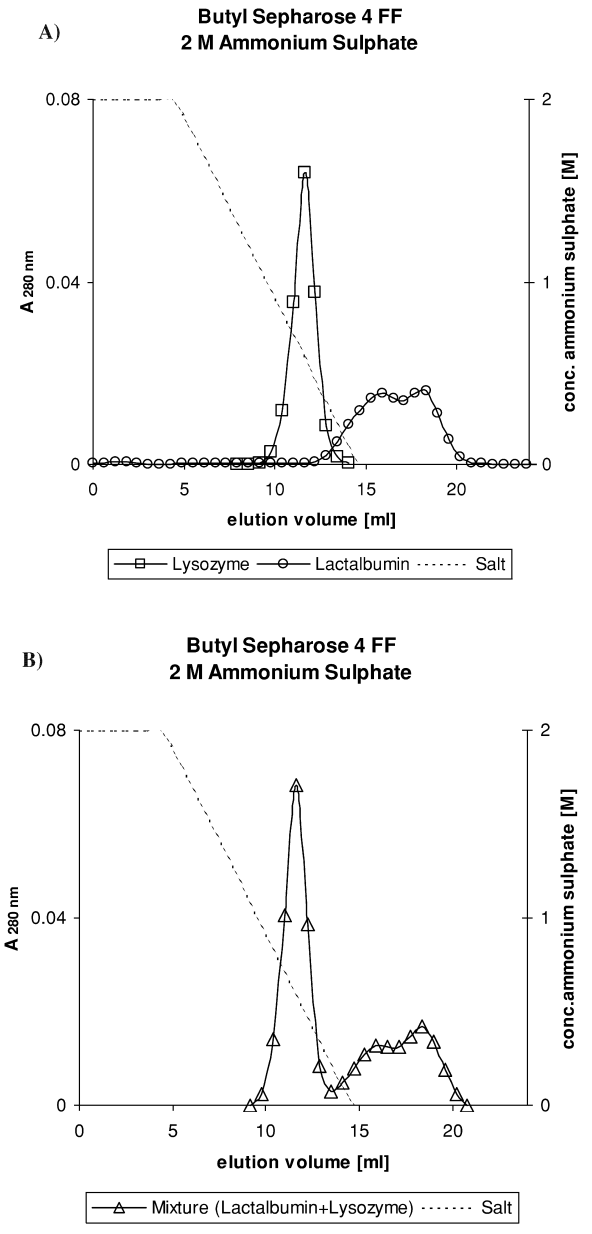

Fig. 1. Comparison between the retention time of the protein injected alone and in a mixture on Butyl Sepharose 4FF; Bis–Tris 20 m*M*, pH 7.0, 2 *M* ammonium sulfate showing the average situation. (A) Proteins injected alone, (B) proteins injected in a mixture. (-○-) α-Lactalbumin, (-□-) lysozyme, (-△-) α-lactalbumin + lysozyme, (- -) salt gradient.

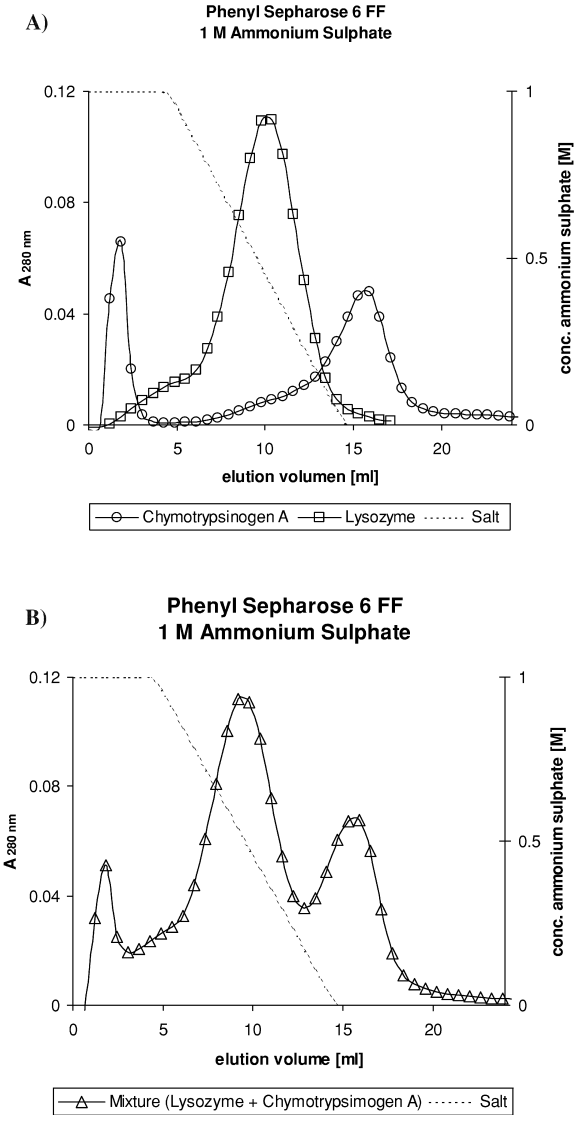

Fig. 2. Comparison between the retention time of the protein injected alone and in a mixture on Phenyl Sepharose 6FF; Bis–Tris 20 m*M*, pH 7.0, 1 *M* ammonium sulfate showing the biggest difference (8.7%). (A) Proteins injected alone, (B) proteins injected in a mixture. (-○-) Chymotrypsinogen A, (-□-) lysozyme, (-△-) chymotrypsinogen A + lysozyme, (- -) salt gradient.

comparison between predicted and observed retention times for the purification of human superoxide dismutase is shown in Table 5 and Fig. 5. This table shows that the deviation value between predicted and observed retention times (3.2%) was less than the

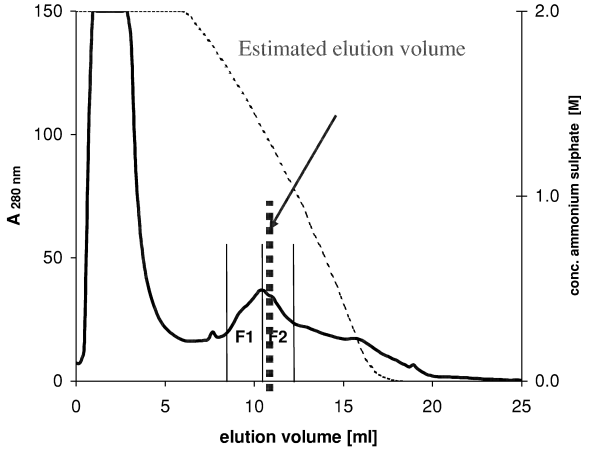

Fig. 3. Chromatogram of human superoxide dismutase from *S. cerevisiae* extract on Phenyl Sepharose 6FF; buffer A, Bis–Tris 20 m*M*, pH 7.0, 2 *M* ammonium sulfate. (—) UV line 280 nm, (---) concentration ammonium sulfate (*M*), (–––) predicted elution volume. F1, F2 human superoxide dismutase elution fractions shown in Fig. 4.

average model deviation shown in Table 4 (7.3%). Hence, the prediction obtained for this dimeric protein was considered satisfactory. In addition, confidence intervals and prediction limits were determined, and the results show that the predicted retention time was inside the confidence interval (shown in Fig. 5).

Thus, it has been possible to predict protein retention time in a practical purification of human superoxide dismutase from cell extract of *S. cerevisiae*.

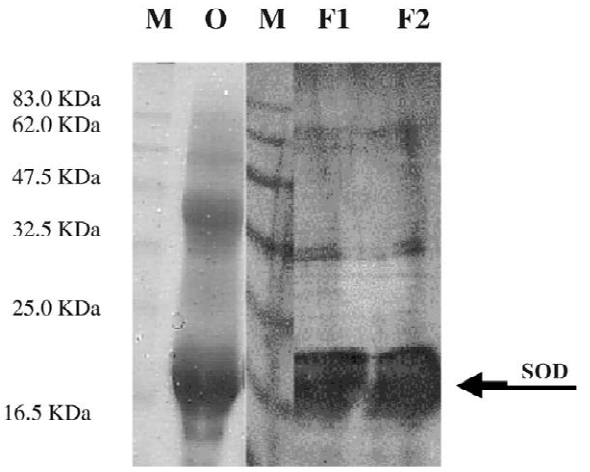

Fig. 4. SDS–PAGE of the collected fractions and original sample from purification of human superoxide dismutase from *S. cerevisiae* extract on Phenyl Sepharose 6FF, 2 *M* ammonium sulfate. M, Molecular mass marker proteins, O, original extract (Coomassie Blue staining), F1, F2, human superoxide dismutase elution fractions (silver staining). kDa = Kilodaltons.

4.3. Selection of the operating conditions in HIC for a mixture of known proteins

Considering the previous results: (i) methodology for predicting the protein retention time for a single protein in HIC; (ii) interactions between proteins adsorbed in HIC are minimal at the protein concentrations used, it would be possible to select the operating condition for purifying a target protein from a mixture of known proteins by HIC. It would consider the following steps:

(1) Predict the retention time of the target protein using the methodology proposed under one operating condition.

(2) Predict the retention time of the main contaminant proteins using the methodology proposed under the same operating conditions.

(3) If separation is adequate, use it on an experimental basis.

(4) If separation is not adequate, evaluate other operating conditions and repeat the first step, until obtaining a good separation.

5. Conclusions

We found that there is no apparent interaction between different proteins adsorbed in the HIC matrix under the conditions studied (at a relatively low protein concentration of 2 g/l). In addition, the methodology proposed for single proteins has been used satisfactorily in a practical example, namely the purification of human superoxide dismutase from an extract of *S. cerevisiae* with good results. Thus, we consider that it is possible to predict the retention time of proteins in a mixture by using the methodology proposed and the application of this methodology for the purification of proteins from cell extracts.

A methodology is also proposed for selecting

Table 5
Comparison between predicted and observed retention times for human superoxide dismutase

Operating conditions	Observed DRT	Predicted DRT	Error[a] (%)	Model deviation (%)
Phenyl Sepharose 6FF, 2 *M* ammonium sulfate	0.59	0.57	3.2	7.3

[a] $\text{Error} = |\text{DRT}_{\text{observed}} - \text{DRT}_{\text{predicted}}| \cdot 100/\text{DRT}_{\text{observed}}$.

separation conditions to purify a target protein from a mixture of known proteins by HIC.

Acknowledgements

This project was supported by the Fondecyt project 1010702. The chromatographic materials donated by Amersham Pharmacia Biotech are gratefully acknowledged.

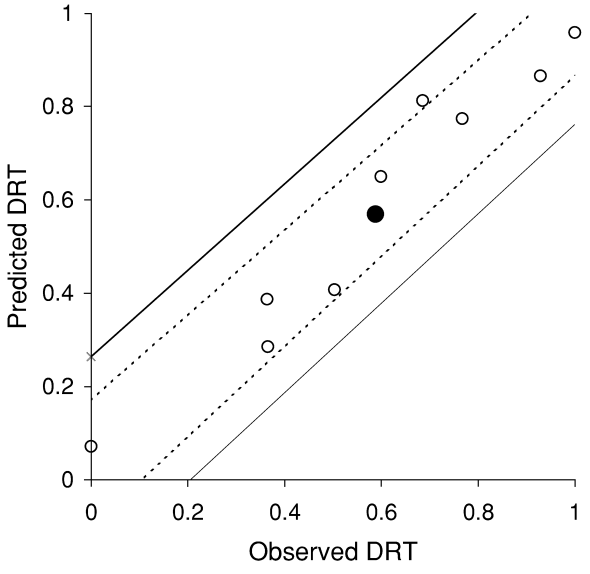

Fig. 5. Comparison between predicted and observed retention times for purification of human superoxide dismutase from recombinant *S. cerevisiae* extract on Phenyl Sepharose 6FF, 2 *M* ammonium sulfate. (○) Standard proteins, (●) human superoxide dismutase, (---) confidence intervals and (—) prediction limits.

References

[1] J.A. Queiroz, C.T. Tomaz, J.M.S. Cabral, J. Biotechnol. 87 (2001) 143.
[2] H.P. Jennissen, Int. J. Bio-Chromatogr. 5 (2000) 131.
[3] Z. Hrkal, J. Rejnkova, J. Chromatogr. 242 (1982) 385.
[4] R.D. McNair, A.J. Kenny, Biochem. J. 179 (1979) 379.
[5] D.E. Comings, A.G. Miguel, H.H. Lesser, Biochim. Biophys. Acta 563 (1979) 25.
[6] B.S. Warren, P. Kusk, R.G. Wolford, G.L. Hager, J. Biol. Chem. 271 (1996) 11434.
[7] S. Hjertén, in: D. Glick (Ed.), Methods of Biochemical Analysis, Wiley, New York, 1981, p. 89.
[8] M. Belew, M. Yafang, L. Bin, J. Berglöf, J.-C. Janson, Bioseparation 1 (1991) 397.
[9] S. Lefort, P. Ferrara, J. Chromatogr. 361 (1986) 209.
[10] M.E. Lienqueo, A. Mahn, J.A. Asenjo, J. Chromatogr. A 978 (2002) 71.
[11] E.W. Leser, M.E. Lienqueo, J.A. Asenjo, Ann. N. Y. Acad. Sci. 782 (1996) 441.
[12] K. Berggren, A. Wolf, J.A. Asenjo, B.A. Andrews, F. Tjerneld, Biochim. Biophys. Acta 1596 (2002) 253.
[13] S. Miyazawa, R. Jernigan, Macromolecules 18 (1985) 534.
[14] A. Nicholls, K. Sharp, B. Honig, Proteins Struct. Funct. Gen. 11 (1991) 281.
[15] R. Gonzalez, B.A. Andrews, J. Molitor, J.A. Asenjo, Biotechnol. Bioeng. 82 (2003) 152.
[16] J.L. Mc Nay, E.J. Fernández, J. Chromatogr. A 849 (1999) 135.
[17] A. Mahn, personal communication, 2003.