Available online at www.sciencedirect.com
ScienceDirect
Reactive & Functional Polymers 67 (2007) 1561–1569
REACTIVE & FUNCTIONAL POLYMERS
www.elsevier.com/locate/react

Investigation of protein retention in hydrophobic interaction chromatographic (HIC) systems using the preferential interaction theory and quantitative structure property relationship models

Jie Chen [a,1], Ting Yang [a,1], Qiong Luo [b], Curt M. Breneman [b], Steven M. Cramer [a,*]

[a] Department of Chemical and Biological Engineering, RPI, NY 12180, United States
[b] Department of Chemistry, RPI, NY 12180, United States

Available online 25 July 2007

Abstract

The retention of proteins in hydrophobic interaction chromatographic (HIC) systems was investigated using the preferential interaction theory in concert with quantitative structure property relationship (QSPR) modeling. The numbers of water molecules and salt ions released upon protein binding were determined from $\ln(k')$ vs. salt concentration plots for a large number of proteins with a range of properties. The effect of salt type on protein binding was also studied by comparing the number of water molecules released in the presence of different salts. Quantitative structure property relationship (QSPR) models based on a support vector machine (SVM) approach were successfully generated for predicting the water molecules released values. These models employed protein crystal structure and primary sequence information as well as a set of hydrophobicity descriptors based on the solvent accessible surface area of the proteins. In addition to successfully predicting the water release values, the selected descriptors provide insights into the protein physicochemical properties which influence protein affinity in HIC system.

Keywords: Hydrophobic interaction chromatography (HIC); Preferential interaction theory; Quantitative structure property relationship (QSPR) modeling; Molecular descriptors; Protein retention

1. Introduction

Hydrophobic interaction chromatography (HIC) has been shown to be an important separation mode for the purification of biomolecules [1–4]. There have been considerable efforts toward understanding the mechanism of protein retention in HIC systems [5–11]. To date, two major theories have been widely accepted in explaining salt effects in HIC systems, the solvophobic theory and preferential interaction theory. The solvophobic theory [5], based on the association and solvation of the participating species, described retention in terms of the molal

* Corresponding author. Tel.: +1 (518) 276 6198; fax: +1 (518) 276 4030.
E-mail address: crames@rpi.edu (S.M. Cramer).
[1] Jie Chen and Ting Yang contributed equally to this work.

1381-5148/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.reactfunctpolym.2007.07.029

surface tension increment of the salt. Fausnaugh and Regnier [12] found that protein retention was also related to the specific interaction between some proteins and salts. The preferential interaction theory [6,9] relates the effect of salt on protein retention to the change in the distribution of salt ions and water molecules. It has been successfully applied in predicting protein adsorption behavior under linear [9] and non-linear conditions [13,14], as well as for studies on the effect of pH [15], ligand chain length and temperature [16] on protein binding in HIC systems. Protein properties, such as the quantity of aromatic amino acid residues, have also been used to explain resin selectivity in HIC systems [10].

Recently, alternative approaches have been developed which combine relatively large amounts of chromatographic data, with known aspects of protein structure and appropriate modeling for describing protein adsorption in HIC systems. Asenjo and co-workers have used a measure of the protein's average surface hydrophobicity to predict the retention in HIC systems [17–22]. They have recently shown that protein primary sequence alone can be used to predict protein retention time in HIC systems [23].

Predictive quantitative structure property relationship (QSPR) models were developed using state-of-the-art non-linear support vector machine (SVM) regression techniques to correlate the experimental results with specific physicochemical properties of the proteins. It was previously reported that models developed using these technique are successful for investigating salt type effects on protein retention [24–26] in ion-exchange systems. Recently, QSPR models have also been employed in HIC systems to evaluate resin effects on protein retention and selectivity [27]. However, investigations of the underlying thermodynamic mechanisms have yet to be performed in detail through the application and interpretation of QSPR models.

In this work, a set of proteins were selected based on their charge, hydrophobicity, molecular weight and functional diversity. Variations in protein retention with salt concentration and salt type were examined by carrying out isocratic experiments on a Butyl Sepharose 4 Fast Flow column. The number of water molecules released upon protein binding in the presence of two salts (ammonium sulfate and sodium chloride) were determined and analyzed through the use of the preferential interaction model. Finally, quantitative structure property relationship (QSPR) models which correlate protein properties with the water molecules released values were developed and analyzed.

2. Theory

2.1. Preferential interaction theory

In HIC systems, it is believed that protein binding is driven by the increase of entropy which is associated with the release of water molecules from the hydration layers of the protein and resin ligands. Thus, the number of released water molecules can be employed as a useful tool to evaluate protein binding affinity. Based on the preferential interaction theory, Perkins et al. [9] obtained the relationships for the capacity factor of a solute in the presence of salt as follows.

For non-electrolyte:

$$\ln k' = c - \frac{n \cdot \Delta\upsilon_1}{m_1 \cdot g} m_3 + \frac{(\Delta\upsilon_+ + \Delta\upsilon_-)}{g} \cdot \ln(m_3). \tag{1}$$

For electrolyte:

$$\ln k' = c + \left[\frac{(\Delta b_+ + \Delta b_-)}{g} - \frac{n \cdot \Delta\upsilon_1}{m_1 \cdot g}\right] m_3, \tag{2}$$

where $\upsilon_i = b_i \cdot m_3$, and Δb is the stoichiometrically weighted change in the ion binding coefficients; m_1 and m_3 are the molal concentration of water and salt respectively; $g = \left(\frac{\partial \ln m_3}{\partial \ln a_\pm}\right)_{T,P}$, a is the activity of ions; n is the valence of salt ions; $\Delta\upsilon_1$ is the number of water molecules released during the binding process; $\Delta\upsilon_+$ and $\Delta\upsilon_-$ are the number of cations and anions released during the binding, respectively. Both n and g will have characteristic values for different types of salts that are employed. The constant g can be calculated from the Debye–Huckel equation or from osmotic pressure experiments. In the present system, $n = 3$, $m = 55.15$, $g = 1.7$ for ammonium sulfate and n=2, $m = 55.15$ and $g = 1.6$ for sodium chloride [9,14].

Eq. (1) can be simplified to

$$\ln k' = \alpha + \beta \cdot C_{\text{salt}} + \gamma \cdot \ln(C_{\text{salt}}), \tag{3}$$

where β and γ are called the preferential interaction parameters, and are given as

$$\beta = -\frac{n \cdot \Delta\upsilon_1}{m_1 \cdot g} \tag{4}$$

$$\gamma = \frac{(\Delta\upsilon_+ + \Delta\upsilon_-)}{g}. \tag{5}$$

The number of released water molecules ($-\Delta\upsilon_1$) and salt ions $-(\Delta\upsilon_+ + \Delta\upsilon_-)$ during the binding process