European Journal of Pharmaceutics and Biopharmaceutics xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

European Journal of Pharmaceutics and Biopharmaceutics

journal homepage: www.elsevier.com/locate/ejpb

Research paper

The accurate measurement of second virial coefficients using self-interaction chromatography: Experimental considerations

A. Quigley [a], J.Y.Y. Heng [b], J.M. Liddell [b], D.R. Williams [a,*]

[a] Surfaces and Particle Engineering Laboratory, Department of Chemical Engineering, Imperial College London, UK
[b] Fujifilm Diosynth Biotechnologies, Billingham, UK

ARTICLE INFO

Article history:
Received 15 August 2012
Accepted in revised form 8 April 2013
Available online xxxx

Keywords:
Second virial coefficient
Self-interaction chromatography
B_{22}
Non-linear chromatography
Lysozyme
Peak asymmetry
Protein–protein interaction

ABSTRACT

Measurement of B_{22}, the second virial coefficient, is an important technique for describing the solution behaviour of proteins, especially as it relates to precipitation, aggregation and crystallisation phenomena. This paper describes the best practise for calculating B_{22} values from self-interaction chromatograms (SIC) for aqueous protein solutions. Detailed analysis of SIC peak shapes for lysozyme shows that non-Gaussian peaks are commonly encountered for SIC, with typical peak asymmetries of 10%. This asymmetry reflects a non-linear chromatographic retention process, in this case heterogeneity of the protein–protein interactions. Therefore, it is important to use the centre of mass calculations for determining accurate retention volumes and thus B_{22} values. Empirical peak maximum chromatogram analysis, often reported in the literature, can result in errors of up to 50% in B_{22} values. A methodology is reported here for determining both the mean and the variance in B_{22} from SIC experiments, includes a correction for normal longitudinal peak broadening. The variance in B_{22} due to chemical effects is quantified statistically and is a measure of the heterogeneity of protein–protein interactions in solution. In the case of lysozyme, a wide range of B_{22} values are measured which can vary significantly from the average B_{22} values.

1. Introduction

The production of protein based therapeutics is now one of the fastest growing sectors in the pharmaceutical industry, being currently used to treat patients suffering from many conditions including cancers, diabetes, heart attacks and cystic fibrosis [1,2]. Despite the increasing success in the discovery of protein based medicines, the manufacture of these protein based biotherapeutics in a cost effective and reliable fashion still remains a major international industrial challenge.

Due to their structural complexity and molecular size, proteins in solution are prone to physical instabilities. As a result, the formation of protein aggregates is an industry wide problem [3]. In the biotherapeutic industry, the presence of any kind of protein aggregate is generally deemed to be undesirable, and protein aggregation has been reported to occur at every stage of the production process [4]. The formation of aggregates is of particular concern during the final production filling operation. This stage is critical as aggregation here can lead to reduced product half-life, altered efficacy, decreased bioavailability and enhanced immunogenicity in the final product [4]. The concern that the presence of aggregates in the final formulation could result in an immunogenic reaction or may trigger adverse effects in patients during clinical administration is of particular note [5].

Modern biochemistry relies on the protein structural information as often provided by X-ray diffraction of protein crystals for protein 3D structure determinations. This knowledge in turn informs us about the fundamental mechanisms of biochemical action. Unlike small molecule crystallisation, protein crystallisation is a difficult and time-consuming process, with many important proteins proving to be thus far impossible to crystallise [6]. Knowledge of protein solubility and solution behaviour is crucial for understanding the nucleation and crystallisation of proteins [7]. Growing theoretical and experimental evidence suggests that protein crystallisation can be understood in terms of a phase transition in a system of weakly attractive particles [8]. Therefore, controlling these attractive interactions is essential for crystal nucleation and subsequent crystal growth. However, an *a priori computational* prediction of protein pair intermolecular potentials governing these attractions is a difficult task and remains a significant modelling challenge which could greatly assist in the determination of optimum crystallisation conditions [4]. Meanwhile, the current empirically based crystallisation trial and error approaches will remain the approach most commonly used.

Experimental methods which allowed the detailed physiochemical characterisation of the influence of solution conditions (pH, ionic strength, buffers and temperature) on protein solution

* Corresponding author. Surfaces and Particle Engineering Laboratory, Department of Chemical Engineering, Imperial College London, London SW7 2BY, UK. Tel.: +44 2075945611.
E-mail address: d.r.williams@imperial.ac.uk (D.R. Williams).

0939-6411/$ - see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.ejpb.2013.04.004

behaviour, including specifically aggregation and crystallisation, would provide the experimentalist with appropriate data sets upon which predictions about the best solution conditions to either prevent aggregation or encourage crystallisation could be established.

One of the key parameters for describing protein self-interactions is the second virial coefficient: B_{22}. The accurate and reliable experimental measurement of B_{22} values has great importance as a method of quantifying and potentially predicting protein–protein interactions in solution. B_{22} can be described as a thermodynamic parameter that characterises the overall net intermolecular forces exerted between two protein bodies in dilute solutions. These molecular interactions include contributions from hydrophobic interactions, hydration forces, van der Waal forces as well as long- and short-range electrostatic factors (attractive and repulsive). Positive B_{22} values correspond to net repulsive forces, whilst negative values represent net attractive forces [3,10]. B_{22} can be easily evaluated experimentally using self-interaction chromatography (SIC) which is the topic of this study. However, until now, a detailed assessment of SIC chromatograms analysis, SIC operational conditions and related experimental practise has not been reported. A specific focus of the current study is a detailed analysis of SIC chromatographic peak shape, and what it tells us about B_{22} values.

1.1. SIC theory

The SIC technique is based on weak affinity chromatography principles, in which the target protein serves as both the chromatographic ligand and ligate. The technique necessitates the covalent immobilisation of the protein of interest onto a suitable solid-state chromatographic stationary phase. Once the column has been packed with a protein immobilised stationary phase, a protein concentration front or pulse of the same protein in a suitable buffer is then injected and eluted through the column stationary phase under isocratic conditions. The column elutant is then monitored for protein concentration, with the retention or breakthrough time reported [11,12]. Under the assumption that the immobilised protein retains its native three-dimensional and secondary structure and that it is immobilised in a broad range of orientations (thus avoiding an unrepresentative side specific interaction), the resulting retention volume estimated from the peak shape, or from the inflection point in the case of a breakthrough curve, will be a measure of the ensemble average strength of protein–protein interaction energy of the two proteins under the solution conditions used.

These data can then be directly related to the B_{22}, second virial coefficient, via statistical mechanics based analysis [11–14] using the following equation:

$$B_{22} = B_{HS} - \frac{(k')}{(\rho_S \phi')} \quad (1)$$

where $k' = \frac{(V_r - V_o)}{(V_o)}$

The SIC retention volume measurements are used to calculate the retention factor k', where V_r is the retention volume of the protein and V_o is the dead volume. B_{HS} represents the excluded volume or hard sphere contribution. The total number of immobilised protein molecules per unit area is denoted by ρ_S and A_s is the total accessible surface area. The phase ratio ϕ is defined to be $\phi = A_s/V_o$, which is the total surface available to the mobile phase protein. The phase ratio can be interpreted using the work of DePhilips and Lenhoff [15].

However, SIC is not without its complexities as in common with most physiochemical chromatographic approaches. One of the core challenges is establishing whether linear chromatography theory can be applied to the data. In short, this assumption requires that the chromatographic results are independent of the solute concentration. This assumption forms the basis of the analysis given above in Eq. (1), and all work previously published on SIC.

1.2. Measuring B_{22}

B_{22} data are not regularly used in biopharmaceuticals, protein science or structural biology research even though these data can offer an important insight into the solution behaviour of bio-macromolecules. There has been limited take up of B_{22} as a routine experimental measurement for crystallisation screening or for aggregation studies. This observation is mainly because the classical B_{22} measurements techniques are quite inefficient in terms of the required amount of protein, experimental measurement time and their troubleshooting demands.

B_{22} was first measured using membrane osmometry (MO) [13,16,17]. But MO is prone to some serious experimental problems including membrane fouling and adsorption and can be particularly problematic for obtaining accurate results at high protein concentrations, where non-ideality is prevalent[18]. So, when Edsall et al. [19] showed that static light scattering (SLS) could produce comparable results to MO, it soon became a more popular method and is now the method most commonly described in the literature and is considered as the reference technique [9]. However, measurement using SLS is very slow because to make one B_{22} measurement requires investigating the scattered intensity versus protein concentration for five or more different concentrations of the protein of interest [20]. As it does not usually produce run-to-run consistent results, several replicates are often required to validate one data point. SLS also suffers from the disadvantages of being very sensitive to dust and impurities. Additionally, it is unsuitable for use with peptides as they do not scatter sufficient light [21]. Other techniques for measuring B_{22} include sedimentation equilibrium [27] and sedimentation velocity analytical ultracentrifugation [28], small angle X-ray scattering [29], small angle neutron scattering [20] and dynamic light scattering [30]. It is, however, important to note that all these methods suffer from some discrepancies compared to sedimentation equilibrium [31]. The main reason behind this discrepancy is sedimentation equilibrium measurements determine only protein–protein interactions, whereas the other methods include the combined contributions of protein–protein and protein-buffer interactions.

New methods have more recently been developed that improve on the limitations of the traditional MO and SLS techniques. In 1996, a new protein characterisation technique was introduced by Patro and Przybycien; self-interaction chromatography (SIC) which promised to overcome some of these disadvantages [12]. The technique was shown to deliver sensible B_{22} values for model proteins including lysozyme and BSA [11,22,23]. Tessier et al. [11] went further and elaborated on the initial method refining the equation relating retention time to B_{22} to consider the column structure more explicitly and introduced the experimental use of a 'dead column' for determining V_o, the dead volume value, used in the B_{22} calculation.

Using SIC to measure B_{22} requires 10 times less protein than SLS, even when the protein immobilised on the column is included and also has the potential to be miniaturised to a microchip level [24], significantly further reducing required amounts of protein. Considering that the same column can be used for months, giving reproducible results, and that after the initial time required to immobilise the protein and pack the column, a B_{22} value can be obtained through SIC every approximately every 30 min, and the SIC technique is significantly more efficient in terms of experimental time. It also has the added potential advantage that it can be used to measure cross-interactions (e.g. B_{23}) between different proteins which was not previously possible [23]. It is also important to note that SIC may also be more accurate than SLS as it is much less

Please cite this article in press as: A. Quigley et al., The accurate measurement of second virial coefficients using self-interaction chromatography: Experimental considerations, Eur. J. Pharm. Biopharm. (2013), http://dx.doi.org/10.1016/j.ejpb.2013.04.004