European Journal of Pharmaceutics and Biopharmaceutics 96 (2015) 282–290

Contents lists available at ScienceDirect

European Journal of Pharmaceutics and Biopharmaceutics

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

Research Paper

The second virial coefficient as a predictor of protein aggregation propensity: A self-interaction chromatography study

A. Quigley, D.R. Williams *

Surfaces and Particle Engineering Laboratory, Department of Chemical Engineering, Imperial College London, London SW7 2BY, UK

ARTICLE INFO

Article history:
Received 31 March 2015
Revised 26 July 2015
Accepted in revised form 27 July 2015
Available online 7 August 2015

Keywords:
Second virial coefficient
b_2
Protein aggregation
Protein–protein interactions
Self-interaction chromatography
Lysozyme
Lactoferrin
Catalase
Concanavalin A
Protein stability

ABSTRACT

The second osmotic virial coefficients (b_2) of four proteins – lysozyme, recombinant human lactoferrin, concanavalin A and catalase were measured by self-interaction chromatography (SIC) in solutions of varying salt type, concentration and pH. Protein aggregate sizes based on the initial hydrodynamic radius of the protein solution species present were measured using dynamic light scattering, and the relationship between b_2 and protein aggregate size was studied. A linear correlation was established between b_2 values and protein aggregate hydrodynamic size for all proteins, and for almost all solution conditions. Aggregate sizes of $<\sim$10 nm, indicative of non-aggregated protein systems, were consistently observed to have b_2 values >0. The observed b_2 trends as a function of solution conditions were very much protein dependent, with notable trends including the existence of attractive interactions (negative b_2 values) at low ionic strengths for catalase and concanavalin A, and the highly positive b_2 values observed for lactoferrin over a wide range of solution conditions, reflecting lactoferrin's innately high stability. It is concluded that the quantification of protein–protein interactions using SIC based b_2 data is a potentially valuable screening tool for predicting protein aggregation propensity.

1. Introduction

Protein aggregation is known to occur at every stage in the production, formulation, storage, shipping and even during administration of protein-based therapeutics. As such protein aggregation is a problem of significant magnitude for the biopharmaceutical industry, and despite enormous technical advances in recent years it continues to be a major obstacle to development [16]. Therefore, the ability to predict, minimise, restrict and/or reverse protein aggregation is crucial to the viable manufacture and formulation of biotherapeutics. Unfortunately the control of aggregation is a considerable challenge because the mechanisms of aggregation follow numerous pathways, and although much knowledge of aggregation mechanisms has been accumulated it is still not currently possible to robustly predict a protein's propensity to aggregate [46]. However, the current models of aggregation have identified two factors that govern stability; one is colloidal and the other is conformational. Colloidal stability is determined by the balance of repulsive and attractive intermolecular interactions between protein molecules in solution. Conformational stability is defined as the difference in free energy between the folded and unfolded states of a protein molecule. Current techniques for predicting protein aggregation propensity are therefore based on the assessment of conformational and colloidal stabilities. These include *in silico* sequence/structure based predictions [11] and determination of melting temperature (T_m) as indicators of conformational stability [35] and the determination of the osmotic second virial coefficient (B_{22}) as a measure of colloidal stability [47,10].

B_{22} can be determined experimentally using static light scattering (SLS) [63], self-interaction chromatography (SIC) [59], membrane osmometry (MO) [32] and analytical ultracentrifugation (AUC) [3]. SIC has established itself as an important experimental technique for the measurement of the B_{22} with comparable results and several advantages over the more established SLS methodology, including reduced amounts of sample and shorter experimental times. B_{22} quantifies the magnitude and direction of protein–protein interactions in dilute solution. Measurement of B_{22} values has been identified as a method of great potential that could have a significant role in the prediction of protein aggregation where attractive protein–protein colloidal interactions are dominant. Negative B_{22} values denote net attractive protein–protein interactions whilst positive values represent overall repulsive interactions.

* Corresponding author. Tel.: +44 2075945611.
E-mail address: d.r.williams@imperial.ac.uk (D.R. Williams).

http://dx.doi.org/10.1016/j.ejpb.2015.07.025
0939-6411/

The use of B_{22} data as a semi-quantitative tool for predicting optimal solution conditions for crystallisation is now a well-documented approach [22,23,63,7,58,19]. Of course a precursor stage, and indeed crucial step, to the crystal growth process is the formation of critical nuclei in solution. Such nucleation events are intrinsically related to aggregation so it is not surprising that B_{22} data could potentially be a useful screening tool for predicting protein aggregation propensity. A number of authors have reported on the potential utility of B_{22} as a predictor of protein aggregation propensity. Published work in this area has shown that protein aggregation behaviour is frequently well correlated to B_{22} values determined under the same conditions [60,29,13]. As such, screening for positive B_{22} values could be used for a rapid determination of high stability solution conditions for proteins.

It is perhaps unsurprising that B_{22} does not always reflect aggregation rates and propensities given that proteins probably belong to the most complex colloidal systems encountered, considering the possible variations in size, morphology-structure, surface charge and surface chemistry. A paper on the pH dependence of B_{22} and aggregation propensity of 3 monoclonal antibodies (mAbs) [49] reported that the correlation between aggregation propensities and B_{22} became insignificant when B_{22} values were negative, as such, even if the B_{22} values were the same the three mAbs exhibited different aggregation propensities. It should also be noted that conformational stability plays an important role in aggregation propensity, with partially unfolded conformational intermediates often considered as the main cause of aggregate formation. When a conformational change is responsible for the onset of aggregation, B_{22} does not always correlate with measured rates of aggregation [12]. Furthermore, a study on ovalbumin and a mAb conducted by Bajaj and co-workers [6], concluded that it was unlikely B_{22} would correlate with long term aggregation because the aggregation-prone structurally perturbed state could be present in a small fraction compared to the native species, yet the structural changes could be significant enough to lead to aggregation in the long term.

The present study investigates the relationship between the B_{22} and aggregation to gain a better understanding of the potential utility of this parameter to predict the propensity of a protein to undergo aggregation. Four different model proteins were used in this study; lysozyme (pI = 11.0, M_W 14.3 kDa), catalase (pI 5.4, M_W 250 kD), concanavalin A (con A) (pI 4.5–5.5, M_W 104–112 kDa), and recombinant human lactoferrin (lactoferrin) (pI 9.5, M_W 82 kDa). Instead of reporting B_{22} values which are dependent on the molecular weight of the protein, it is more appropriate for the comparison of different proteins that data are presented as the reduced or dimensionless osmotic second virial coefficient (b_2) for which b_2 is normalised by the excluded volume contribution B_2^{HS}. B_2 can easily be converted to B_{22} through the following equation [8]:

$$b_2 = \frac{B_2}{B_2^{HS}} = \frac{3B_2}{2\pi\sigma^3} = \frac{B_{22}M_w^2}{N_a B_2^{HS}} \quad (1)$$

All proteins studied were subjected to solution conditions intended to rapidly induce aggregation as well as those in which they were stable. b_2 values were measured under similar conditions using the improved (first moment) method to determine retention times from SIC data recently reported [42]. It has been recently shown that b_2 values obtained in this manner show more accurate correlation with protein aggregation and that peak shape may be itself an important indicator for conformational changes in protein samples. These observations regarding peak shape complement earlier work discussing the possibility that retention peak data contain information not only on the average B_{22} values typically reported, but for a range of B_{22} values reflecting the heterogeneity of protein solution interactions [42].

2. Materials and methods

2.1. Materials

Chicken egg white lysozyme (62971), catalase from bovine liver (C9322) and concanavalin A (con A) from *Canavaliaensiformis* (L7647) were purchased from Sigma Aldrich. Recombinant human lactoferrin (lactoferrin) was produced at Fujifilm Diosynth using *Aspergillus niger* as the expression system and purified by cation exchange chromatography. The concentration of lactoferrin was 100.7 mg/mL in phosphate buffered saline solution pH 7.5. Potassium phosphate, sodium cyanoborohydride, dibasic sodium phosphate, MES, *N*-(3-dimethylaminopropyl)-*N*′-ethylcarbodiimide-hydrochloride (EDC), *N*-hydroxysuccinimide (NHS), ethanolamine, HCl and NaOH were all purchased from Sigma–Aldrich (ACS or BioXtra grade). NaCl, Sodium acetate trihydrate, glacial acetic acid and acetone were purchased from Fisher Scientific and were AR grade. Toyopearl AF-Formyl-650M and AF-Amino-650M chromatography particles (08004 and 08002) were obtained from Tosoh Biosep. Deionised water used for preparing all buffer and protein solutions was processed by a Centra ELGA system. The pH was measured using a Mettler Toledo FiveEasy pH meter. All solutions were filtered prior to use using 0.22 μm filters from Millipore. Protein concentrations were determined by BCA protein assay using a kit obtained from Pierce and a Lambda 4B spectrophotometer from Perkin–Elmer.

2.2. Protein immobilisation

The immobilisation of lysozyme, lactoferrin and con A to Toyopearl AF-Formyl-650M particles was based on the method by Tessier et al. [59] as detailed here [42]. Catalase was immobilised to Toyopearl AF-Amino-650M particles using a method described by Dumetz et al. [18]. Between 65 mg and 110 mg of each protein were dissolved in 10 mL buffer solution (lysozyme in 0.1 M potassium phosphate at pH 7.5, lactoferrin in 20 mM sodium phosphate pH 7, catalase in 5 mM MES pH 6.5 containing 0.1 M NaCl, and con A in 20 mM sodium acetate pH 4.5). The coupling was catalysed using sodium cyanoborohydride for lysozyme, lactoferrin and con A and with a mixture of EDC and NHS for catalase. Any remaining active sites on the media were capped using ethanolamine. The protein loaded stationary phase was then slurry packed (at a flow rate of no more than 3 mL/min) into the column and washed *in situ*. Samples were collected from the initial protein solution and each of the washes in order to calculate the net amount of protein immobilised on the stationary phase by BCA protein assay. When not in use the columns were stored in a pH 7 50 mM sodium phosphate buffer at 4 °C. A column without protein immobilised, referred to as the dead column was also prepared in order to calculate the dead volume of the column as described by Tessier et al. [59]. The choice of resin for these experiments was based on the highest levels of immobilisation achieved for each protein to be studied.

2.3. Self-interaction chromatography

SIC measurements were performed using an Agilent 1100 series liquid chromatograph (Agilent Technologies, Cheshire, UK) consisting of a binary pump, degassex, autosampler, column temperature control unit, Phenomenex Degassex model DG-4400 vacuum four-channel on-line degassex (Phenomenex, Torrance, CA) and two variable wavelength detectors – one before and one after the column. The LC system was controlled and data were collected using Chemstation software version Rev.A.10.02 for LC systems (Agilent Technologies). The protein loaded stationary phase was