

**MECHANISTIC INSIGHTS INTO THE ROLE OF PROTEIN
INTERACTIONS ON THE AGGREGATION BEHAVIOR OF
ANTI-STREPTAVIDIN IMMUNOGLOBULIN GAMMA-1**

by

Gregory V. Barnett

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemical Engineering

Fall 2015

© 2015 Gregory V. Barnett
All Rights Reserved

ProQuest Number: 10014766

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10014766

Published by ProQuest LLC (2016). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

MECHANISTIC INSIGHTS INTO THE ROLE OF PROTEIN
INTERACTIONS ON THE AGGREGATION BEHAVIOR OF
ANTI-STREPTAVIDIN IMMUNOGLOBULIN GAMMA-1

by

Gregory V. Barnett

Approved: _____
Abraham M. Lenhoff, Ph.D.
Chair of the Department of Chemical Engineering

Approved: _____
Babatunde Ogunnaike, Ph.D
Dean of the College of Engineering

Approved: _____
Ann L. Ardis, Ph.D.
Interim Vice Provost for Graduate and Professional Education

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Christopher J. Roberts, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

David W. Colby, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Eric M. Furst, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Vladimir I. Razinkov, Ph.D.
Member of dissertation committee

ACKNOWLEDGEMENTS

I would like to thank many people for their support during my time at the University of Delaware. First and foremost, I want to thank my advisor, Chris Roberts, for his guidance and support over the years. Chris encouraged many collaborative opportunities that has enriched this work and my experience here. He challenges students to think deeper about protein interactions and behavior. I will always value the rigorous approach Chris brings to chemical engineering and protein biophysics, and I will strive towards that standard in my career.

I want to thanks the many collaborators and colleagues who invited me to work in their labs. A special thanks to Amgen for supporting this work and Vladimir Razinkov who meet biweekly for almost four years. I thank you for your patience, helpfulness and practical industrial perspective. Thank you to Tom Laue for guidance on electrophoresis measurements, collaborators at Malvern Instruments, Wei Qi, Samiul Amin, and Neil Lewis, who allowed me to visit their lab and work with the DLS-Raman prototype. I enjoyed many scientific discussions with Wei and Samiul. Thanks to Yun Liu and Paul Butler at the NIST Center for Neutron Research who where always friendly and helpful discussing SANS experiments or data analysis.

I want to thank the current and past members of the Roberts group, especially Nayoung Kim, Marco Blanco, Chris O'Brien, and Ranendu Ghosh for their support, patience, and help with many long hours fixing and troubleshooting experiments and instruments. I could always count on open doors and thorough scientific discussions.

Lastly I want to thanks my friends and family. Jon, Ian and friends at Delaware have made the last four years fun and exciting. Most importantly, I owe a special thanks to my Mom and Dad for encouragement and education throughout the years and always having a welcoming home nearby.

TABLE OF CONTENTS

LIST OF TABLES	x
LIST OF FIGURES	xi
ABSTRACT	xviii

Chapter

1 INTRODUCTION	1
1.1 Motivation	1
1.2 Objectives	5
1.3 Non-native aggregation	6
1.3.1 Rate determining steps for non-native aggregation	7
1.4 Aggregation mechanisms	8
1.5 Modeling aggregation kinetics	8
1.5.1 Colloidal aggregation	9
1.5.2 Lumry-eyring nucleation polymerization (LENP) model	10
1.6 Factors affecting protein aggregation	10
1.6.1 Solution pH	11
1.6.2 Salts	12
1.6.3 Neutral osmolytes	13
1.6.4 Other osmolytes	15
1.7 Predicting aggregation	16
1.7.1 Non-arrhenius behavior	17
1.8 Anti-streptavidin IgG1	18
1.8.1 IgG1 structure	18

1.8.2	Prior work with AS-IgG1	19
1.9	Organization of the dissertation	20
2	SPECIFIC-ION EFFECTS ON IGG1 AGGREGATION BEHAVIOR	22
2.1	Introduction	22
2.2	Materials and methods	23
2.2.1	Sample preparation	23
2.2.2	Size exclusion chromatography with inline light scattering (SEC-MALS)	24
2.2.3	Protein-protein interactions via laser light scattering	25
2.2.4	Determination of radius of gyration for small oligomers with SAXS	26
2.2.5	IgG1 aggregate hydrodynamic radius from dynamic light scattering	26
2.2.6	IgG1 net charge (valence) via electrophoretic light scattering	27
2.2.7	Aggregate morphology using small angle neutron scattering (SANS)	28
2.3	Aggregation mechanism(s) from SEC-MALS	29
2.4	Aggregate mass to size scaling	34
2.5	Aggregate morphology using SANS and SAXS.	36
2.6	Protein-protein interactions based on static light scattering	40
2.7	Protein net echarge determined by electrophoretic light scattering (ELS)	42
2.8	Role of electrostatic protein-protein interactions on IgG1 aggregation mechanims	46
2.9	Aggregate morphology from scattering	49
2.10	G_{22}^* : semi-quantitative tool to predict aggregation mechanism	51
2.11	Summary and Conclusions	53
3	PARALLEL TEMPERATURE INITIAL RATES: PH, AND COUNTERION EFFECTS ON IGG1 AGGREGATION RATES	55
3.1	Introduction	55
3.2	Materials and Methods	57
3.2.1	Differential scanning calorimetry (DSC)	57
3.2.2	Quantifying Aggregation Rates	57

3.2.3	Parallel Temperatures Initial Rates (PTIR)	58
3.3	AS-IgG1 thermal unfolding using differential scanning calorimetry (DSC)	60
3.4	Aggregation Rates from PTIR and Standard Isothermal Approaches	62
3.5	Effects of pH, buffer, and NaCl on Temperature-Dependent Rates	69
3.6	Summary and Conclusions	73
4	OSMOLYTE EFFECTS ON MONOCLONAL ANTIBODY STABILITY AND CONCENTRATION-DEPENDENT PROTEIN INTERACTIONS WITH WATER AND COMMON OSMOLYTES	74
4.1	Introduction	74
4.2	Materials and methods	77
4.2.1	Sample preparation	77
4.2.2	Partial specific volume via densimetry	78
4.2.3	Model prediction of preferential interaction via solvent accessible surface area (ASA)	79
4.2.4	AS-IgG1 unfolding via differential scanning calorimetry	81
4.3	Partial specific volume via density measurements	81
4.4	Partial specific volume of AS-IgG1 in neutral osmolytes	83
4.5	Kirkwood-Buff integrals for protein-water and protein-osmolyte interactions	86
4.6	AS-IgG1 native state chemical potential	89
4.7	AS-IgG1 unfolding via DSC	92
4.8	Preferential interactions and their effect on protein unfolding and stability	95
4.9	Summary and conclusions	100
5	AGGREGATE STRUCTURAL CHANGES AND MECHANISMS AT ELEVATED CONCENTRATION	102
5.1	Introduction	102
5.2	Materials and Methods	103
5.2.1	Sample Preparation	103
5.2.2	<i>Ex-situ</i> monomer loss kinetics and light scattering using SEC-MALS	104
5.2.3	Circular Dichroism	104
5.2.4	Second derivative UV absorption	105

5.2.5	<i>In-situ</i> dynamic light scattering with Raman spectroscopy (DLS-Raman)	106
5.2.6	Small angle neutron scattering (SANS)	106
5.3	AS-IgG1 aggregation kinetics and growth mechanism	107
5.4	AS-IgG1 structural changes during aggregation	110
5.5	IgG aggregation monitored with <i>in-situ</i> DLS-Raman	114
5.6	Scaling of Aggregate M_w^{agg} and R_h	118
5.7	IgG1 monomer structure factor using small angle neutron scattering (SANS)	120
5.8	<i>In-situ</i> IgG aggregation using SANS	122
5.9	Structural changes during aggregation	126
5.10	Aggregate morphology from Kratky plots	127
5.11	Sucrose affects rates but not qualitative features of aggregation	129
5.12	Summary and Conclusion	130
6	SUMMARY AND FUTURE WORK	132
6.1	Summary	132
6.2	Future work	138
6.2.1	Refrigerated and room temperature predictions of protein aggregation rates	138
6.2.2	Mechanistic investigation of aggregate coalescence and condensation	139
6.2.3	Extended characterization of protein-osmolyte interactions . .	141
	BIBLIOGRAPHY	142
	Appendix	
A	COMPARISON OF PROTEIN-PROTEIN INTERACTIONS FROM DLS AND SLS, AS-IGG1 NET CHARGE MEASUREMENT DETAILS, AND AN IN-DEPTH CHARACTERIZATION OF AS-IGG1 UNDER AGGREGATION RESISTANT SOLUTION CONDITIONS	163
A.1	A comparison of protein-protein interactions via dynamic light scattering and static light scattering	163
A.2	Electrophoretic light scattering with phase analysis light scattering (PALS) using the Zetasizer	167
A.3	AS-IgG1 aggregation resistant condition: pH 4, 10 mM acetate	170

B	DETAILED DENSIMETRY RESULTS AND ANALYSIS	175
B.1	Density as a function of AS-IgG1 weight fraction	175
B.2	Statistics for linear regression of $\hat{V}_2(c_3\hat{V}_3)$	175
B.3	Preferential interaction model predictions	180
B.4	Osmolyte-water thermodynamics	181
C	REPRINT PERMISSION LETTERS	185

LIST OF TABLES

2.1	SANS fitting parameters I	38
2.2	SANS fitting parameters II	38
2.3	Buffer anion concentration as a function of pH. Concentration of acetate ion is the first row followed by citrate ions.	46
3.1	DSC apparent T_m values. Error bars are ± 0.2 °C	62
3.2	PTIR device and method reproducibility	66
4.1	Osmolyte partial specific volume. Error bars correspond to 95 percent confidence intervals	85
5.1	Solution conditions and aggregation mechanisms used to create aggregates	108
A.1	Comparison of G_{22}^* values for AS-IgG1 with 10 mM acetate	166
A.2	Comparison of G_{22}^* values for AS-IgG1 with 5 mM citrate	166
B.1	Regression analysis for IgG1 and sucrose	180
B.2	Regression analysis for IgG1 and trehalose	180
B.3	Regression analysis for IgG1 and sorbitol	180
B.4	Regression analysis for IgG1 and PEG	180

LIST OF FIGURES

1.1	Schematic of a non-native protein aggregation mechanism	6
1.2	IgG1 structure from (A) a front view and (B) in a block diagram. Arginine, lysine, and histidine amino acids are colored in red and aspartic acid and glutamic acid are colored blue.	18
2.1	SEC-MALS results distinguishing different aggregation mechanism. (A) Representative chromatograms for IgG1 monomer, the nucleation-dominated (ND), chain polymerization (CP), and association polymerization (AP) growth mechanisms. M_w^{tot} vs $(1 - m)^2$ plots showing aggregation through (B) ND, (C) CP, and (D) AP growth mechanisms. Blue open symbols represent aggregates from 5 mM citrate buffer and red solid symbols are for 10 mM acetate buffer.	31
2.2	AS-IgG1 aggregation state diagrams for (A) 10 mM acetate and (B) 5 mM citrate (Reproduced from reference [1]). Predominant aggregation mechanism is shown for no aggregation (open diamonds), ND (open circles), CP (filled circles), AP (filled squares), mix of CP and AP (open squares), and PS (closed triangles). The dotted curved lines illustrate the ionic strength of the (A) acetate and (B) citrate buffer component(s).	33
2.3	Scaling of average aggregate size with molecular weight based on (A) R_h or (B) R_g . ND (circles), CP (squares), and AP (triangles) depict different aggregate growth regimes. The black lines show power-law scaling relationship is shown for slopes of 1,2, and 3. Open blue (closed red) symbols represent aggregates created in citrate (acetate) buffer.	35
2.4	SANS intensity profiles for purified aggregates from (main panel A) ND and (inset) CP and (panel B) AP growth mechanism Closed red symbols are shown for 10 mM acetate buffer and open blue symbols for 5 mM citrate. The black lines are best fits to form factor models from the NCNR software.	37

2.5	Normalized SAXS Kratky plots for aggregates from ND growth mechanism created in (A) 10 mM acetate and (B) 5 mM citrate. Scattering intensity was normalized to the low Q limit and the Q was normalized to R_g	39
2.6	(A) Excess Rayleigh ratio vs. IgG1 concentration for pH 4, 5, and 6 in 5 mM citrate and 10 mM acetate. Refer to the text on the panel for exact pH-buffer conditions. Colored curves are fits to Eq. 2.3 and the black curve is the Rayleigh ratio for a hard sphere with equivalent diameter to the IgG1.	41
2.7	G_{22}^* values determined as a function of protein concentration for pH 4, 0 mM NaCl in (A) 10 mM acetate and (B) 5 mM citrate. At other pH and NaCl concentrations, G_{22} is not a function of concentration in the range of protein concentration tested. G_{22}^* values are shown for (C) 10 mM acetate and (D) 5 mM citrate buffer, for 0 mM NaCl (blue), 20 mM NaCl (red) and 200 mM NaCl (green). Error bars represent 95% confidence intervals.	43
2.8	IgG1 net charge determined with from ELS analysis. (A) IgG1 velocity plotted vs electric field for pH 4, 5 mM citrate for 20 mM NaCl (blue squares), 40 mM NaCl (red circles), and 100 mM NaCl (green triangles). (B) PEG 35kDa velocity plotted vs electric field for pH 6, 5 mM citrate for 20 mM NaCl (blue squares), 40 mM NaCl (red circles), and 100 mM NaCl (green triangles). PEG was used as a tracer to detect electro-osmotic flow (EOF). The net effective charge was determined from ELS analysis for 10 mM acetate (C) and 5 mM citrate (D) for 20 mM NaCl (blue), 40 mM NaCl (red) and 100 mM NaCl (green). Error bars are 95% confidence intervals.	45
2.9	Schematic depiction of electrostatic colloidal interactions causing shifts in observed aggregation mechanisms. (A) Potential mean force W_{22}/k_bT plotted as a function of center-to-center distance for two charged hard spheres with the same hydrodynamic radius as an IgG1. The blue arrow shows the effect of decreasing Z_{eff} (changing pH towards the pI) and increasing ionic strength (increasing NaCl concentration). (B) Schematic describing monomer-monomer (M-M), monomer-aggregate (M-A), and aggregate-aggregate (A-A) association (cf., main text).	47

2.10	Overlay of the IgG1 aggregation state diagrams with G_{22}^* surfaces from multivariate regression for (A) 10 mM acetate and (B) 5 mM citrate. Labels are G_{22}^* values determine from the multivariate regression. Symbols have same representation as Fig 2.	52
3.1	Schematic of the PTIR device. The device has ten independently controlled peltier heating block each capable of holding two 1.5 mL HPLC vials. The device allows one to accurately and efficiently determine liquid state polymeric or protein reaction rates.	58
3.2	DSC for IgG1 (solid) and Fc-IgG1 (dashed) formulated in 10 mM acetate buffer at pH 4 (A), pH 5 (B), and pH 6 (C) with no added salt or 100 mM NaCl added salt (offset 150 kcal/mol). Curves are offset vertically for easier visualization	61
3.3	Illustrative isothermal and PTIR results for determining IgG1 aggregation rates at pH 5, (5 mM citrate buffer), and with the addition of 100 mM NaCl. (A) Isothermal monomer loss versus time at 325.5 K (triangles), 330 K (circles), and 332 K (squares shown in the inset). (B) PTIR monomer loss at 2 hours (triangles, circles, diamonds, and red squares) or 24 hours (blue squares) as a function of incubation temperature. Multiple 2-hr PTIR data sets were repeated to illustrate experiment-to-experiment variability. (C) $\ln(k_{obs})$ from PTIR data and Eq. 3.2. Symbol correspond to the closed symbol points at the same temperature in panel B. Error bars are smaller than the size of the symbols unless visible in either panel.	64
3.4	Arrhenius plot: natural logarithm of the aggregation rate plotted as a function of inverse temperature. Accelerated aggregation rates were determined using PTIR method for IgG1 at pH 5 and 0 mM NaCl concentration (closed circles), pH 5 and 100mM NaCl concentration (closed squares), pH 6 and 0 mM NaCl concentration (closed triangles), and pH 6 and 100 mM NaCl concentration (closed diamonds). Open symbols correspond to aggregation rates determined using traditional isothermal incubations.	67
3.5	. (A) Arrhenius plot using PTIR for 2 hour and 24 hour incubations. (B) Arrhenius plot rescaling incubation temperature by the DSC peak temperatures for a given formulation. See main text for symbol notation.	69

3.6	IgG1 effective aggregation activation energy determined from PTIR data in Figure 3.5 for (A) 10 mM acetate (B) 5 mM citrate. Red bars are shown for no added NaCl and blue bars are shown for 100 mM NaCl concentration. The question mark on the red bar for pH 4 in panel A indicates that E_a is unknown for that condition because aggregation was too slow to measure. Error bars are shown for 95% confidence intervals.	71
4.1	Binary solutions. (A) water-osmolyte solutions: density as a function of osmolyte weight fraction for sucrose (squares), trehalose (circles), sorbitol (triangles), and polyethylene glycol $M_n = 6,000$ g/mol (diamonds). (B) water-protein solutions: reciprocal of the density plotted as a function of AS-IgG1 weight fraction. (C) Conversion of $c_3\hat{V}_3$ to osmolyte molarity.	82
4.2	Ternary solutions. AS-IgG1 partial specific volume as a function of $c_3\hat{V}_3$ for (A) sucrose, (B) trehalose, (C) sorbitol, and (D) PEG. The dashed black line is the prediction from the preferential interaction models [2, 3], using a homology model of AS-IgG1 based on the human IgG1 crystal structure (PDB 1HZH). Error bars correspond to 95 percent confidence intervals. Solid curves are empirical fits to interpolate the experimental values.	84
4.3	Average protein-water interactions (G_{23} , solid curves) and protein-osmolyte interactions (G_{12} , dashed curves) as a function of $c_3\hat{V}_3$ for (A) sucrose, (B) sorbitol or trehalose, and (C) PEG.	88
4.4	AS-IgG1 native state chemical potential as function of $c_3\hat{V}_3$ for (A) sucrose, (B) sorbitol or trehalose, and (C) PEG. dashed curves correspond to predictions using the preferential interaction models.	90
4.5	AS-IgG1 differential scanning calorimetry. DSC thermograms with increasing concentrations of (A) sucrose, (B) trehalose, (C) sorbitol, and (D) PEG. In each panel, curves are offset to help distinguish profiles.	93
4.6	AS-IgG1 differential scanning calorimetry. (A) C _H 2-Fab and (D) C _H 3 T_m values as a function of osmolyte concentration. Symbols correspond to trehalose (squares), sucrose (diamonds), sorbitol (circles), and PEG (triangles).	94

4.7	Schematic of (A) classic theory and (B) concentration dependent preferential interaction. Panels illustrate qualitative behavior of the protein chemical potential for the native (N) and unfolded (U) states as a function of osmolyte concentration	97
5.1	(A) Monomer fraction remaining as a function of time at elevated temperature for pD 5.3 heated at 69 °C (squares) and (inset) pD 5.1, 100 mM NaCl heated at 53 °C (circles) and pD 4.6, 100 mM NaCl heated at 58 °C (triangles). (B) M_w^{tot} vs $(1 - m)^2$ plot. In both panels, open symbols correspond to solutions with 0.15 M added sucrose.	109
5.2	(panel A) Illustrative CD spectra over time AS-IgG1 pD 5.3 heated at 69 °C. Net aggregate contribution ($[\theta_{agg}]$) to the CD spectra for aggregate growth via (panel B) ND , (panel C) CP, and (panel D) AP mechanism. (panel E) Average $[\theta_{agg}]$ for each mechanism and the unheated AS-IgG1 monomer spectra.	111
5.3	(A) Illustrative 2 nd derivative UV-vis spectra over time for AS-IgG1 via ND growth. (B) Tyrosine and (C) Tryptophan center of mass shift (COM) as a function of monomer consumption, $(1 - m)$. Symbols correspond to same solution conditions as Fig. 5.1	113
5.4	(A) AS-IgG1 <i>in-situ</i> DLS-Raman-I. (A) Illustrative Raman spectra for three time points at elevated temperature for IgG1 via ND growth. Raman markers plotted as a function of time for (B) Amide I region, (C) Trp at 1550 cm ⁻¹ , and (D) Ratio of Raman peak intensity at 510 to 540 cm ⁻¹ . Symbols correspond to same solution conditions as Figure 5.1	115
5.5	(A). AS-IgG1 <i>in-situ</i> DLS-Raman II.(A) Tyr COM at 830 cm ⁻¹ (B) Tyr COM at 850 cm ⁻¹ (C) Trp COM at 880 cm ⁻¹ and (D) Amide III region. Symbols correspond to same conditions as in Figure 5.1.	117
5.6	(A) <i>in-situ</i> R_h over time. (B) Scaling between M_w^{agg} and R_h . Symbols correspond to same solution conditions as Figure 5.1. Black lines in panel B correspond to slopes of 1, 2, and 3	119
5.7	(A) SANS intensity as a function of Q for unheated (monomer) AS-IgG1 at pD 5.3 and protein concentration of 5, 10, and 30 mg/mL at 25 °C. Structure factor as a function of Q for (B) pD 5.3, 0 mM NaCl, (C) pD 4.6, 100 mM NaCl, and (C) pD 5.1, 100 mM NaCl. Closed symbols correspond to buffer only, and open symbols to buffer with 0.15 M sucrose.	121

5.8	AS-IgG1 aggregation with <i>in-situ</i> SANS. (A) pD 4.6 with 100 mM NaCl and protein concentration of 30 mg/mL heated at 53 °C. Kratky plots for AS-IgG1 over time at elevated temperature. (panel B) pD 5.3 heated at 69 °C; (panel C) pD 4.6 with 100 mM NaCl heated at 53 °C; (panel C) pD 5.1 with 100 mM NaCl heated at 58 °C. Black curves correspond to buffer only, while blue curves are for 0.15 M sucrose.	123
5.9	SANS Normalized Kratky plots (NKP) for AS-IgG1 over time at elevated temperature. Plots in Fig. 5.8-C-D were normalized to $I(Q \rightarrow 0)$ and R_g . (panel A) NKP for AS-IgG1 at pD 4.6 with 100 mM NaCl heated at 53 °C; (panel C) pD 5.1 with 100 mM NaCl heated at 58 °C. Black curves correspond to buffer only, while blue curves are for 0.15 M sucrose.	125
A.1	IgG1 collective diffusion coefficient as a function of protein concentration. for pH 4, 10 mM acetate with 0 mM NaCl (blue), 20 mM NaCl (red), and 200 mM NaCl (green). Colored curves are fits to Eq.A.1 and the black curve is the expected trend for a hard sphere with an equivalent diameter to the IgG1.	165
A.2	IgG1 electrophoretic light scattering. (A) Average distance moved by the IgG1 from diffusion (red) and electrophoresis (blue). (B) Applied voltage vs time for a single run using the Zetasizer. (C) Phase shift plotted vs time for IgG1 pH, 4 20 mM NaCl. Each curve represents single electric field strength. (D) Phase sift plotted as a function of time for PEG at pH 6, 40 mM NaCl exhibiting electro-osmotic flow. Each curve corresponds to a different electric field strengths	168
A.3	AS-IgG1 unfolding under aggregation resistant solution conditions. (A) AS-IgG1 DSC thermogram at pH 4, 10 mM acetate showing three unfolding transitions. (B) AS-IgG1 SEC chromatogram after heating at 85 °C for one hour. (C and D) <i>In-situ</i> AS-IgG1 temperature scanning SAXS.	172
A.4	DLS auto-correlation functions as a function of temperature for IgG1 at pH 4, 10 mM acetate during heating and cooling.	174
B.1	Reciprocal of the solution density as a function of AS-IgG1 weight fraction at various sucrose concentrations.	176
B.2	Reciprocal of the solution density as a function of AS-IgG1 weight fraction at various trehalose concentrations.	177

B.3	Reciprocal of the solution density as a function of AS-IgG1 weight fraction at various sorbitol concentrations.	178
B.4	Reciprocal of the solution density as a function of AS-IgG1 weight fraction at various PEG concentrations.	179
B.5	Comparison of ASA algorithms determine backbone (top) and side chain (bottom) fractional exposure	182
B.6	Osmolyte-water binary systems. Water activity data for aqueous solutions of (Panel A) PEG (data from [4, 5]) and (Panel B) trehalose (data from [6]). Derivative of the osmolyte chemical potential with respect to osmolyte concentration as a function of osmolyte concentration for (C) PEG and (D) trehalose.	184

ABSTRACT

Monoclonal antibodies (mAbs) are one of the leading protein-based drug candidates in the biopharmaceutical industry. The formation of irreversible, non-native protein clusters (hereby called aggregates) is a common degradation route during manufacturing of mAbs and other therapeutic proteins. Aggregation may potentially decrease drug efficacy and jeopardize patient safety. A patient's immune system may mount a response against the therapeutic protein making future treatments ineffective and potentially dangerous. As such, controlling unwanted aggregation is an ongoing and crucial challenge in the development of protein-based therapeutics. The process(es) by which aggregates form or the aggregation mechanism(s) influence the aggregate size distribution, concentration, and structure, all of which may potentially impact drug potency and safety. A deeper understanding of the aggregation process may foster strategies to predict and control aggregation.

Thermally induced aggregation of anti-streptavidin immunoglobulin gamma-1 (AS-IgG1), a monoclonal antibody, is first mapped over a range of pH and NaCl concentration typical in formulated protein products. Aggregation mechanisms are influenced by low-concentration anions, such as acetate and citrate that are common buffer species. The relationships among monomeric protein-protein interactions, aggregate morphology, and aggregation mechanisms are explored. Colloidal interactions (i.e. potential of mean force) arguments are discussed in terms of AS-IgG1 aggregation mechanisms. Static light scattering and observed aggregation mechanisms suggest a citrate/acetate specific-ion-effect that cannot be explained with simplified colloidal interactions. The Kirkwood-Buff integral for protein-protein interactions semi-quantitatively predicts how changes in solution conditions change aggregation mechanisms, and may offer a practical tool to capture specific ion effects.

IgG1 aggregation rates are strongly affected by solution pH, ionic strength, and counter-ion species. Additionally, the rates are highly temperature dependent. For example, an increase of 5°C often increases aggregation rates multiple orders of magnitude. As such, it is challenging to compare aggregation rates across a broad range of solution conditions and temperatures. A Parallel Temperature Initial Rates (PTIR) device and method are introduced to accurately and efficiently determine liquid-state degradation rates. IgG1 aggregation rates determined using PTIR compare well to the traditional isothermal approach. The PTIR method has the advantage of being more sample sparing and providing temperature dependent rates with more statistical certainty. Thermal unfolding and aggregation rates of AS-IgG1 were determined across multiple pH values and NaCl concentrations using differential scanning calorimetry (DSC) or PTIR. AS-IgG1 aggregation rate coefficients partially collapse to a common profile upon rescaling the incubation temperature by the midpoint unfolding temperature (T_m) determined via DSC. However, the effective activation energies (E_a) depend on solution pH, NaCl concentration, and citrate/acetate buffer species. Acetate vs. citrate specific-ion-effects manifest themselves in T_m and E_a values. The roles of protein-protein interactions and protein unfolding are discussed in the context of AS-IgG1 aggregation.

Osmolytes, such as sucrose, trehalose, and sorbitol, are commonly added to protein solutions to prevent unfolding and aggregation. These additive molecules are thought to compete with water for interactions with proteins resulting in so-called preferential interactions with proteins that can increase or decrease the chemical potential of proteins. However, direct measurements of protein preferential interactions remain challenging. AS-IgG1 protein-water and protein-osmolyte interactions are determined using precise density measurements of AS-IgG1 with aqueous solutions of sucrose, trehalose, sorbitol, and PEG ($M_n = 6,000$ g/mol). AS-IgG1 with sucrose and PEG show preferential interactions that depend on osmolyte concentration, which contradicts conventional wisdom for how these molecules interact with proteins in solution. Preferential interactions are compared to protein thermal unfolding using DSC.

AS-IgG1 T_m values increase linearly as a function of sucrose, trehalose, and sorbitol concentrations, and decrease linearly as a function of PEG concentration. Results are compared to available models based on protein solvent exposed surface area (ASA) and discussed in terms of the classical preferential interaction theory.

mAb formulations require intermediate to high protein concentrations due to dosing requirements of intravenous administration or subcutaneous injections. Elevated concentrations may promote non-ideal protein-protein and protein-osmolyte interactions. The framework for AS-IgG1 aggregation developed in earlier sections under dilute protein concentrations is extended to AS-IgG1 at 30 mg/mL in deuterated water (D_2O) and in the presence and absence of 0.15 M (~ 5 w/w %) sucrose. Protein structural techniques are monitored with circular dichroism, second derivative UV absorbance, and Raman spectroscopy. Results are compared to the aggregate morphology, which is monitored with size exclusion chromatography with in-line laser light scattering, dynamic light scattering, and small angle neutron scattering (SANS).

Structural changes suggest the nucleation-dominated mechanism (at low pD and ionic strength) has significantly larger structural perturbations, as observed in the disulfide bond conformation, secondary structure, and tyrosine and tryptophan environments. The addition of 0.15 M sucrose decreases aggregation rates, particularly for the nucleation dominated mechanism, but does not alter aggregation mechanisms. AS-IgG1 monomer SANS structure factor ($S(Q)$) remains unchanged with the addition of sucrose, suggesting sucrose does not alter protein-protein interactions. Aggregate morphology is monitored with SANS, and the Kratky plot analysis shows a unique scattering profile for each mechanism regardless of the presence of sucrose. These results are consistent with the current framework that suggests protein-protein interactions mediate aggregation mechanisms. Overall, results in this dissertation illustrate the effect of protein-protein, protein-osmolyte, and protein-water interactions on the protein stability and aggregation behavior of AS-IgG1. Many tools and approaches utilized in this dissertation can be applied to degradation processes of other soft matter systems.

Chapter 1

INTRODUCTION

1.1 Motivation

Monoclonal antibodies (mAbs) are one of the fastest growing class of therapeutic proteins in the pharmaceutical industry [7]. Currently, 39 marketed monoclonal antibodies have been approved in the US or EU [8]. In 2013, mAbs grossed over \$75 billion globally [7]. These large labile proteins are often of commercial interest due to their long half-lives *in-vivo* and their affinity and specificity for targeted antigens [9]. Monoclonal antibodies are versatile; different classes can be utilized to elicit B-cell recognition or cell apoptosis, induce signal transduction or downregulate certain pathways [10]. Over the past three decades, the development of mAbs has lead to a number of treatments for previously unmet needs, including therapies for many cancers and auto-immune diseases [11].

Among protein-based therapeutics, mAbs are biological in nature and endogenous if created from humanized cell lines. However, mAbs and other protein-based therapeutics may have potential immunogenicity associated with treatment. This side-effect may occur when a patient's immune system mounts a response against the therapeutic protein. Often it is difficult to assess if and when immunogenicity will occur as the anti-drug immune response depends on patient physiology, use of concurrent medications, and length of treatment, among other factors [12, 13, 14]. Immunogenicity is also influenced by drug product quality attributes such as the presence of foreign impurities, final formulation, heterogeneous patterns of glycosylation, and the presence of non-native protein aggregates [12]. Here, and throughout this dissertation, a non-native aggregate is defined as an irreversible collection of protein monomers.

Regulatory agencies require a protein's critical quality attribute (CQA) to be extensively investigated before market approval [15]. CQAs may include aggregation, deamidation, fragmentation, oxidation or other degradation pathways [16]. Typically the protein is degraded using different stresses, such as agitation, heating, or exposure to UV light, and the resulting structure and morphology is investigated [16, 17]. In particular, aggregates with similar structure or repeating epitopes to the therapeutically active protein may have an increased immunogenic risk associated with treatment [18, 19]. In this case, future treatments will be ineffective and potentially jeopardize patient safety.

The aggregate size, concentration, and structure have been linked to immunogenicity [12]. Unfortunately, at least some level of aggregation is ubiquitous during manufacturing of protein-based therapeutics [20]. During the expression of mAbs in cell culture, aggregates may form before the protein is fully folded, as mAbs and many other proteins require chaperone proteins to properly fold [21]. During protein purification, low pH conditions are required for viral deactivation which may decrease protein stability and allow the protein to partially or fully unfold near room temperature. Often unfolding exposes buried hydrophobic patches, which may potentially be aggregation prone. During chromatographic purification, protein concentrations may reach 100 mg/mL or higher locally within the media pores. High protein concentrations and the presence of high salt concentrations (~ 2 M) typically in used in wash buffers may promote protein self-associations and later lead to irreversible aggregates. The final formulation, which includes the choice of pH, buffer, and any osmolytes (added macromolecules), may affect product shelf-life. During the fill-finish process, silicone oils used as lubricants in syringes or the process equipment may contaminate the drug product. Even the quality of the glass vials may promote protein adsorption and potentially lead to unfolding at the glass-liquid interface [22, 23, 24]. As there is no accepted or predetermined safe threshold of aggregate size and concentration in solution, minimizing and controlling aggregation is necessary for the development of protein-based therapeutics.

Controlling protein aggregation is challenging as the aggregate state(s) are expected to be lower in free energy compared to the native protein. The native monomeric protein may exist in a local free energy minimum. However, chemical, thermal, or mechanical stresses during manufacturing can increase partially unfolded states. Partially or fully unfolded proteins often have hydrophobic patches or “hotspots” that promote aggregation and potentially lead to the formation of strong intermolecular β -sheets. Thermodynamics dictate the free energy of unfolding and are strongly determined by the primary sequence of proteins. Solution conditions, such as pH, salt, buffer, and the addition of osmolytes, may also shift the protein chemical potential. However, thermodynamics alone may not offer strategies to eliminate aggregation, and therefore it is important to control the kinetics of aggregation [21].

Aggregation mechanisms directly affect the aggregate size, concentration, and resulting secondary and tertiary structure. Mass-action kinetic models have shown the aggregate size and concentration are not independent quantities [25, 26]. Combining mass-action models with scattering and monomer loss measurements has been shown to be a useful tool. Currently, *a priori* predictions of aggregation mechanisms and resulting aggregate sizes and concentrations have been an outstanding challenge. For a given protein, changes in the solution pH, salt concentration, buffer, or the addition of other osmolytes often mediate aggregation mechanisms [1, 27].

Often protein stability is thought to involve both protein-protein interactions and conformational stability, but the balance between these two factors remains unclear [28]. In the case of conformational stability, the free energy of unfolding is often not experimentally accessible for mAbs and many proteins. If unfolding is irreversible and/or aggregation occurs during the unfolding transition, the thermodynamic analysis may not be applicable as discussed previously [29]. Many studies have focused on interpreting protein unfolding with aggregation and have allowed one to draw at least semi-quantitative conclusions about relative increases in stability [30, 29].

On the other hand, protein-protein interactions or protein charge measurements

are accessible for the folded state, but often measurements are limited to lower temperatures to avoid aggregation. However, the interactions between partially or fully unfolded proteins are arguably the relevant parameter in the non-native aggregation process. It remains unclear if these surrogate biophysical parameters correlate or may be predictive of aggregation behavior. Many studies have pointed to the importance of conformational stability but it remains unclear what role protein-protein interactions play in aggregation [28].

It is challenging to accurately quantify mAb aggregation rates across various formulation parameters, such as pH, salt concentration, buffers, or other osmolytes. Aggregation is a temperature sensitive degradation process and an increase of even a few degrees Celsius may increase aggregation rates by at least an order of magnitude [31]. As such, comparing aggregation rates and behavior across multiple formulations is challenging. Temperature scanning based methods such as differential scanning calorimetry offer approaches to probe the relative stability of many formulations and or drug candidates. However, these techniques do not offer a practical means to accurately quantify aggregation rates. Currently, international guidelines for drug stability may require at most three stability temperatures at 5, 25, and 40 °C for assessing product storage and shelf-life [32]. Stability testing at these temperatures may require 6-18 months for mAbs and other proteins. Often, stability testing is continually in progress and shelf life is assessed as data is processed. Robust strategies are needed to collect many more temperature dependent rates and over a larger temperature and time range with the goal of predicting long-term stability. These strategies may enhance the “developability” of new drug candidates thereby increase the speed of new therapies to patients [33].

As mentioned above, surrogates for protein conformational stability and protein-protein interactions are used to guide formulations. Aggregation are often monitored with spectroscopic techniques (e.g. circular dichroism, infrared, or Raman) and aggregate sizing techniques (e.g scattering, microscopy, or micro-flow imaging). However,

aggregation rates based on monomer loss measured with chromatography based methods are considered a gold standard of stability testing for regulatory filings. Despite growing interest in developing techniques to monitor aggregation, particularly at high protein concentration, it remains unclear if spectroscopic or particle sizing techniques will correlate quantitatively with monomer loss rates [34].

1.2 Objectives

This dissertation addresses many of the current challenges mentioned above using a therapeutically relevant protein provided by Amgen, anti-streptavidin immunoglobulin gamma-1 (AS-IgG1). This dissertation is meant to extend the current understanding of mAb aggregation. A range of techniques are utilized in this dissertation and offer insights into the measurement capabilities and the advantages and disadvantages associated with them.

Specifically, the objectives of this dissertation are to

1. Investigate the relationships among AS-IgG1 protein-protein interactions, aggregation mechanisms, and aggregate morphology. Provide insights into molecular level interactions and offer strategies to predict aggregation behavior.
2. Develop a strategy to accurately and efficiently determine aggregation rates. Examine the balance of conformational stability and protein-protein interactions on protein stability. Provide a proof of concept method to measure formulation dependent aggregation rates over many orders of magnitude with the aim of applications in an industrial setting.
3. Investigate the influence of protein-osmolyte and protein-water interactions on protein stability. More specifically, examine the ability of protein interactions to predict protein stability. Evaluate *a priori* and semi-quantitative approaches to predict protein stability in the presence of osmolytes and offer insights to improve predictions.
4. Examine if newly developed hypotheses of AS-IgG1 aggregation hold for a new formulation, which includes elevated protein concentration (30 mg/mL), the presence and absence of sucrose, and heavy water (D_2O). Specifically, evaluate if semi-quantitative trends in aggregation mechanisms and osmolyte behaviors hold at elevated protein concentration and in a deuterated solvent. Compare how IgG1 structure and morphology changes for different aggregation mechanisms. Additionally, investigate the effect of sucrose on the aggregation mechanisms and rates and resulting aggregate morphology and structure.

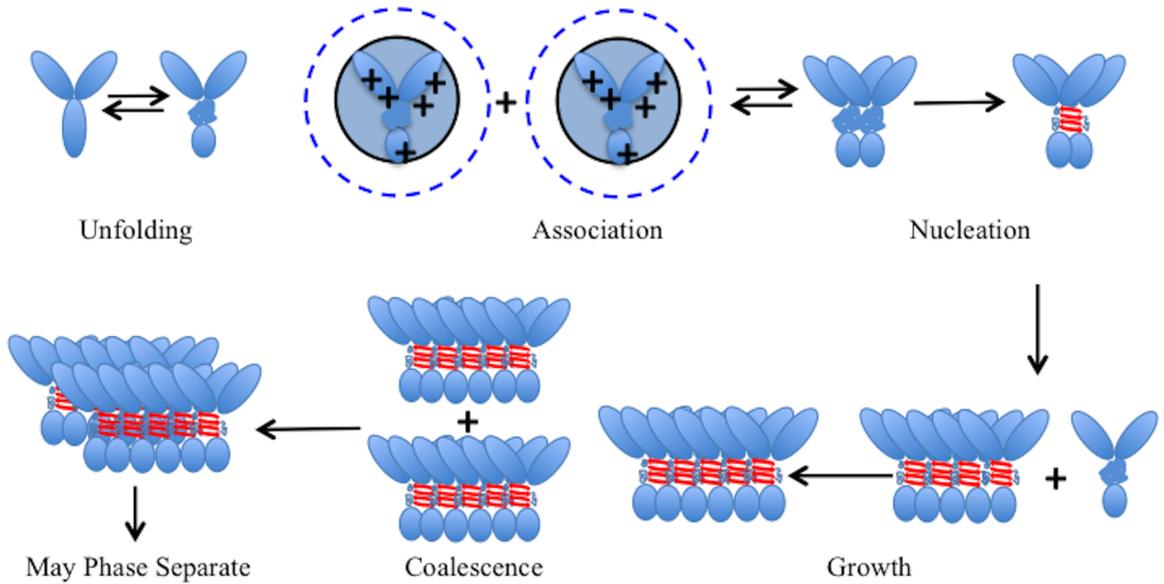


Figure 1.1: Schematic of a non-native protein aggregation mechanism

1.3 Non-native aggregation

Aggregation is a multi-step process by which a folded native protein irreversibly incorporates into an aggregate. A schematic for a possible aggregation pathway is presented in Figure 1.1. Starting with a folded IgG, depicted as a blue tripod, the monomer undergoes an unfolding transition of one or more domain(s). Following unfolding, IgG monomers may self-associate with neighboring molecule(s). In Fig. 1.1, the blue dotted line around IgG molecules in the association step denotes the range of the screening length. The Debye-Hückel screening length is a function of the ionic strength and the effective range of electrostatic interactions felt by neighboring protein molecules [35]. The black line surrounding an IgG molecule in the association step represents an effective hydrodynamic diameter of each molecule. The plus signs on the IgG molecules represent the effective charge on the protein surface (while there may be patches of positive and negative charged areas). The protein net charge and range of electrostatic interactions mediate the energy barrier for two or more protein molecules to associate. In the nucleation step, mAb monomers putatively rearrange

to form an irreversible nucleus. The red ribbons connecting monomers within the aggregates denote the irreversible “locked in” structural change, which may potentially be intermolecular β -sheet. Given a particular set of solution conditions (i.e. pH and or NaCl concentration), the aggregation may stop at this nucleus size or grow through monomer-addition to form larger chain-like aggregates [1, 36]. In this step, another partially unfolded IgG monomer adds directly to the aggregate. Depending on solution conditions, aggregates may also coalesce with one another. As aggregates grow larger, they may remain soluble or phase separate [37].

1.3.1 Rate determining steps for non-native aggregation

Any of the steps in Figure 1.1 has the potential to be a rate determining step (RDS). Weiss et al. summarized the parameters for each possible RDS [38]. Pre-equilibrated steps in the reaction mechanism add additional terms in the rate expression. For the simplest case of unfolding being the RDS, unfolded monomers will be directly converted to aggregates. The rate coefficient for folding and unfolding will be the key parameter governing aggregation kinetics. However, if self-association is the RDS, then unfolding will be pre-equilibrated and the rate coefficient for dimerization would determine the aggregation rates. This case resembles polymerization reactions in which an initiator (such as a free radical) starts the reaction, and the polymer usually grows to very large lengths. Finally, if nucleation is the RDS, the aggregation rate will include a timescale for rearrangement process to lock in the final aggregate conformation. For protein aggregation, the resulting intermolecular structure forms during the rearrangement process. Schmit provided a statistical mechanical approach for amyloid fibrillization found in tau aggregation and incorporated the equations into a kinetic model for aggregation [39, 40]. However, for many proteins, the intermolecular contacts within the aggregate and the process by which the structure forms are difficult to measure experimentally for a variety of reasons [41].

1.4 Aggregation mechanisms

Previous work with a model globular protein, α -chymotrypsinogen (aCgn) [42], and other proteins [43, 1, 44] has shown that if the incubation temperature is more than a few degrees Celsius below the melting temperature for protein unfolding, the RDS are almost invariably the nucleation and growth steps. Interestingly, the ratio of rate coefficients (or inverse timescales) for nucleation and growth determine the size and concentration of aggregates. Often the timescale for nucleation is orders of magnitude larger compared to growth and the aggregates grow to sizes consisting of 10-100 monomers [45, 27]. In this case, aggregation predominantly follows growth by monomer addition or chain-polymerization (CP). Surprisingly, prior work has shown for IgG aggregation at low pH (e.g. pH 4) and low ionic strength, aggregates effectively stopped growing past dimer and trimer sizes and the mechanism was categorized as nucleation-dominated (ND)[36, 1]. However, depending on solution conditions, once the aggregate population reaches a sufficiently high concentration, or the monomer pool is sufficiently depleted, aggregate coalescence or association polymerization (AP) may predominate. Once this growth mechanism predominates, large (1000+ monomers per aggregate) may remain soluble or phase separate (PS) and become visible particles.

1.5 Modeling aggregation kinetics

Mathematical models offer a useful tool to gain insight into the aggregation process. A given aggregate stoichiometry, which consists of a particular number of monomers, can be written in terms of mass-action kinetic equations. Integrating these equations results in the time-dependent concentration of each aggregate stoichiometry. This population balance approach was first introduced by Smoluchowski in 1916 and has been used extensively in colloidal systems [46, 47]. Typically, the population balance approach utilizes an experimental data set, which may consist of monomer fraction remaining, the molecular weight and or particle size from laser light scattering and time points over which the data were collected. A non-linear regression simultaneously integrates differential equations and fits the model parameters to experimental

data. Historically, population balance approaches have not accounted for steps such as unfolding, and most published models are not consistent with the long timescales involved with experimental aggregation rates for therapeutic proteins[48].

1.5.1 Colloidal aggregation

Traditionally, colloidal aggregation involved the stability of colloidal suspension, such as polystyrene beads in water. The aggregation rate and resulting particle morphology would be a function of the monomer size, functionalized surface, and solution pH and ionic strength. Derjaguin-Landau-Verwey-Overbeek (DLVO) theory provides the general framework for the intermolecular interactions relevant for colloidal stability [49]. Given a combination of short-range attractions and long-range repulsions, the aggregation rate and size can be reasonably well predicted [49]. Diffusion limited aggregation is the simplest colloidal aggregation model. The rate coefficients for aggregate stoichiometries are determined by the collision rate of particles and are inversely proportional to particles sizes. Colloidal aggregation may also follow reaction limited aggregation. This model introduces a probability that collisions result in a sticking event [50].

For particle systems with intermolecular interactions there may be an energy barrier for two particles approaching each other. For spherical, homogeneous colloidal particles, the energy barrier is well described using the DLVO potential of mean force. However, DLVO colloidal interactions are only valid when the particle size is much larger than solvent and ion molecules. Colloidal interactions treat the solvent as a continuous dielectric medium and salt ions as point charges. As the length scales of colloidal particles (e.g proteins) decrease, these simplifications become less accurate. Interestingly, the aggregate particles exhibit mass to size scaling or fractal dimension which depends on the growth process. For DLCA, the fractal dimension is ~ 1.86 , while for RLCA the value is ~ 2.1 [51].

A population balance approach was used to study IgG aggregation at temperatures near or above midpoint temperatures for unfolding [47, 48], where a large fraction

of protein molecules are unfolded and colloidal aggregation may be expected. Even under these extreme conditions, the fitted parameters for the aggregation energy barrier were very large compared to values one would expect based on DLVO theory [48]. However, one may anticipate that colloidal models would fail to capture aggregation events that are more complex than collide-and-stick events that are assumed in such models. Many studies have shown aggregates may form intermolecular β -sheets and rearrange after initial associations to “lock in” a final structure [52]. However, it is interesting to consider the applicability and limitations of simple colloidal models in context of protein aggregation.

1.5.2 Lumry-eyring nucleation polymerization (LENP) model

The Lumry-Eyring nucleation polymerization (LENP) model has been applied to a range of different proteins that undergo aggregation [53, 45, 43, 54]. One main difference between LENP and colloidal models is the treatment of “reactive” monomer species, which for the case of proteins requires an unfolding event. Additionally, the model treats aggregate growth as a separate process from coalescence. The first generation LENP model used monomer loss kinetics along with the light scattering total molecular weight to fit parameters for nucleation and growth [53]. The subsequent LENP model accounted for aggregate coalescence and recast the full set of differential equations into differential equations for the moments of the aggregate size distribution [26]. Interestingly, analysis using second generation LENP found that multiple possible models for the coalescence process will fit the data equally well [26, 27]. Therefore, additional information regarding the aggregates size distribution is needed to draw conclusions about coalescence processes.

1.6 Factors affecting protein aggregation

Most mAb therapeutics are formulated in the liquid state as they are administered to patients intravenously or through subcutaneous injections. Recent mAb products do not need to be lyophilized to attain sufficient stability at refrigerated

conditions. The solution conditions, such as pH, ionic strength, and the addition of osmolytes (e.g. salts, sugars, surfactants, or complex stabilizers) all have the potential to affect aggregation behavior. Each factor may affect multiple steps in the aggregation process and in a different manner. While the protein primary sequence dictates the final three dimensional structure of the fold protein, often a drug candidate sequence is determined prior to formulation. Therefore, the solution conditions and added osmolytes are the key factors one has the freedom to alter to maximize the stability of protein drug products. The following subsections provide an overview of typical factor affecting protein aggregation. Additionally, the factors below are organized from largest to smallest effects on protein stability.

1.6.1 Solution pH

The solution pH is arguably the most significant factor affecting protein stability, from the perspective of solution parameters that can easily be adjusted. For a given protein sequence and structure, the pH dictates the net charge on the protein surface, which is a combination of positively charged (arginine, lysine, histidine) and negatively charged (aspartic acid, glutamic acid) residues. Given the solution pH, a protein's raw net charge can be calculated from the number of charged residues and the pKa for each of the titratable side groups. Additionally, the local environment around each titratable residue may shift the pKa [55]. The raw charge calculation is based on the primary sequence and is a reasonable estimate of protein net charge for globular proteins [56]. However, for IgGs, the calculated raw charge usually overestimates the measured net charge [36, 57].

As the solution pH moves closer to the isoelectric point (pI) of protein, the free energy of unfolding and midpoint temperature of unfolding often increase. For IgGs, the pH has been shown to affect the stability of different regions of the IgG molecule differently. For example, differential scanning calorimetry revealed that one domain in the fragment crystallizable (Fc) region is often significantly more affected by pH than the fragment antigen-binding (Fab). As the pH moves away from the

pI, the protein becomes more charged, which increase the electrostatic protein-protein interactions. The protein solubility is also partly determined by the surface area that contains charged amino acids, as more patches of charged surface promote hydrogen bonding with water molecules. Prior work has hypothesized that if aggregates grow sufficiently large, phase separation may readily occur and a phase boundary exists for aggregates in solution [37]. This work showed that titrating aggregate solutions from pH 4 to 6 or increasing the ionic strength with a chaotropic salts led to precipitated aggregates. Interestingly, when solutions were titrated back to low pH or diluted with buffer aggregates became soluble.

1.6.2 Salts

The solution ionic strength is a significant factor in aggregation. Salts and other osmolytes are added to formulations to maintain an isotonic concentration and thereby prevent hyper(hypo)tonicity [58]. Typically 200-300 mM of osmolytes is added, but this value depends on the site and route of administration. Ions in solution create a double layer around the charged protein surface, effectively shielding electrostatic forces felt by neighboring proteins. The Poisson-Boltzmann equation (PBE) allows one to describe the spatial dependence of the electrostatic potential around a protein [35]. The characteristic length-scale for the linearized PBE is the Debye-Hückel screening length, which gives a numerical measure of the effective distance the double layer extends and the range of electrostatic forces felt by neighboring proteins [35].

Debye-Hückel theory treats all ions equally, but ions have different sizes and charge, which changes how they are solvated by water molecules [59]. Putatively, this could lead to different ions being more or less effective at shielding the net charge of a protein and thereby altering electrostatic protein-protein interactions. The Hofmeister effect was discovered in 1888 and describes the ability of certain ions to “salt in” or “salt out” proteins [60]. Typically the Hofmeister effect was observed for salt concentrations in 0.2-2 M range. Interestingly, anions and cations follow an order from ions that “salt in”, (chaotropes), to ions that “salt out” (kosmotropes). For many

decades, the Hofmeister effect has guided the choice of solution conditions for protein crystallization and formulation [60, 61]. While the Hofmeister phenomenon has been observed extensively for over 125 years, the molecular interactions that give rise to the effect remain controversial. Traditionally, chaotropes (kosmotropes) were thought of as structure breakers (makers), as they disrupt (stabilize) the structure of water molecules and lead to increased (decreased) interaction between water and proteins. However, one may hypothesize that the Hofmeister effect involves a more direct mechanism.

Ion solvation is thought to disrupt the hydration layer near the protein surface [59]. This hydration layer is thought to be important in protein stability as the hydrogen bonds between water molecules and the protein surface maintain a protein conformation. Notably, specific-ion effects have been observed with many proteins and influence protein unfolding and aggregation rates [61, 62, 63]. Chaotropic salts have been observed to more effectively decrease electrostatic repulsive protein-protein interactions and the protein's net charge [64]. Notably, these effects manifest at low salt concentration (0-100 mM) and follow the Hofmeister series [65, 66]. However, one may hypothesize that the specific-ion effect (or Hofmeister effect) may involve a direct interaction between the protein and ion(s) as these effects manifest themselves at low salt concentrations and directly affect electrostatic inter-protein interactions and net protein charge.

1.6.3 Neutral osmolytes

Interactions between proteins, water, and osmolytes mediate changes in protein stability and phase behavior [67, 68, 69, 70]. Neutral osmolytes do not increase the solution ionic strength, which may weaken electrostatic protein-protein interactions and reduce conformational stability. Notably, chemical denaturants such as urea or guanidinium are added to measure unfolding transitions and determine the free energy of unfolding. Similar to the hypothesis regarding the mechanism of the Hofmeister effect, it was unclear if chemical denaturants interacted directly with the protein or indirectly through a similar type of water structure mechanism. However, the urea and

water interaction is nearly ideal [71], and there is a lack of evidence for the structure breaker hypothesis mentioned above [72, 73].

Other osmolytes, such as sugars are often added to mAb formulations to stabilize proteins. Timasheff and coworkers proposed a mechanism by which sucrose stabilizes proteins [74]. They hypothesized that as sucrose is added to the solution, water becomes preferentially accumulated near a protein’s surface while sucrose becomes preferentially excluded from the surface [75]. Thermodynamics dictate that competing protein-water protein-osmolyte interactions alter the chemical potential of the protein. Timasheff and others proposed that the denatured state of a protein is expected to have a more extended conformation compared to the folded state. As such, the preferential exclusion hypothesis is expected to have a larger effect on the denatured state and increase the chemical potential of the denatured state of the protein more so than that of the native state. Therefore, osmolyte preferential exclusion is expected to increase the free energy difference between the native and denatured states and stabilize the protein [76, 77].

Sugars and other polyhydroxy compounds e.g., sucrose or trehalose are often observed to increase the unfolding free energy or midpoint unfolding temperature for proteins *in-vitro* [74, 78, 79]. Similarly, the addition of sulfate anions or polyethylene glycol (PEG) is expected to increase the chemical potential of folded proteins, and this motivates the use of these compounds to promote protein “salting out” or phase separation [80, 81]. Historically, protein solubility data in PEG or sulfate protein and has been used to predict protein solubility in the absence of such solutes [81, 82, 83]. The underlying basis for these effects is often summarized in terms of the preferential interactions framework that was a major focus of the seminal work of Timasheff, Schellman, and others [74, 84, 85, 86, 87, 88, 89].

In a slightly different context, sugars are also added (in relatively large amounts) during lyophilization or freeze-drying to stabilize proteins [90, 91, 92]. At these high sugar concentrations, sugar molecules are thought to replace water molecules near a protein’s surface, thereby preventing protein unfolding. However, this water-replacement

theory appears paradoxical to the above mentioned preferential exclusion of sugars. Trehalose, sucrose, glycerol, and sorbitol are commonly added in stabilizing proteins. In particular, the larger sugar molecules (trehalose and sucrose) are thought to have enhanced stabilizing effects due to excluded volume [84] at lower sugar concentrations and due to the water replacement mechanism at higher concentrations. This apparent discrepancy between protein preferential interactions has not been reconciled [93, 78].

1.6.4 Other osmolytes

Additionally, surfactants and amino acids (e.g. proline, arginine, or glycine) are often added to stabilize protein formulations. Many formulations contain some level of surfactants as these molecules are known to migrate towards hydrophobic interfaces and potentially reduce protein adsorption and unfolding at solid-liquid and air-liquid interfaces. Surfactants have been effective in reducing aggregation during freeze-thaw cycles (ice-liquid interface) [94] and in the presence of silicone oil microdroplets (oil-liquid interface) [23]. Additionally, surfactants stabilize proteins during agitation (air-liquid). However, prior work has shown some surfactants decrease while others increase stability when the protein is exposed to thermal or light stresses [95].

Arginine is a charged amino acid commonly added to protein formulations. However, this amino acid has mixed stability effects as it also been shown to destabilize the protein under certain conditions [91]. Arginine is thought to preferentially accumulate near hydrophobic patches on protein surfaces and thereby reduce self-associations [96, 97]. As such, the addition of arginine may reduce aggregation and solution viscosity if at high protein concentrations. However, arginine is a charged amino acid and increases the ionic strength of an aqueous solution. Therefore, at pH values far from the pI where the protein has a larger net charge, arginine will also have a positive charge. Often it is reported that arginine may reduce increased solution viscosity at pH values closer to the protein's pI [96, 97]. Many other osmolytes have been added to formulations, and may stabilize or destabilize proteins, and often these effects depend

on the specific protein and potentially other formulation parameters. For a detailed review consult the following references [98, 99]

1.7 Predicting aggregation

In the biopharmaceutical industry, often the optimal formulation for protein stability is determined using intermediate to high throughput biophysical measurements to screen across pH, salt, and osmolytes [100, 101, 102]. However, considering all possible combinations of factors is impractical from an industrial or academic perspective. Often, heuristics guide the choice of solution pH, ionic strength, and the addition of osmolytes when screening formulations.

Predicting protein aggregation rates has been a longstanding challenge. Currently there are limited tools to determine aggregation rates or mechanisms *a priori*. Aggregation prediction calculators offer one tool to predict aggregation propensity from a protein’s primary sequence. Protein aggregation studies utilizing site directed mutagenesis may find aggregation propensity calculators useful for comparing relative stability among point mutations with a given protein [103, 104]. However, these calculators often provide limited insight into how “aggregation propensity” alters aggregation mechanisms and rates. Also it often does not effectively predict how “aggregation propensity” will change with solution pH, ionic strength, and the addition of osmolytes.

Often more practical approaches utilize biophysical techniques that may potentially correlate with aggregation rates and mechanisms. However, characterization techniques only offer surrogates for relevant factors as mentioned above [38]. Often midpoint temperatures of unfolding from differential scanning calorimetry correlate well with relative stability of formulations at elevated temperatures [1, 27]. However, scanning-based techniques lack the ability to predict lower temperature aggregation rates. One such method highlighted here is Temperature Scanning Monomer Loss (TSML) [31]. It is a temperature scanning method developed to efficiently determine temperature dependent rates and potentially predict longer time aggregation rates. However, aggregation rates measured using scanning method may be scanrate

dependent. Isothermal incubations remain the industry standard, though only three temperatures are required for regulatory filing and product shelf life.

Predicting aggregation rates at refrigerated, room temperature, or even 40 °C, has been a long-standing challenge in the biopharmaceutical industry. Often the factors affecting aggregation may influence multiple steps in the aggregation mechanism at once and lead to highly non-linear changes in aggregation rates. Often a 5 °C increase in temperature can increase aggregation rates multiple orders of magnitude.

1.7.1 Non-arrhenius behavior

Aggregation rates are challenging to predict at low temperature as they often exhibit non-Arrhenius behavior in 4-80°C temperate window [31]. Aggregation rates at lower temperatures (25-40°C) are often much higher than rates extrapolated from higher temperatures [105]. Some non-Arrhenius behavior may be expected if one considers protein unfolding thermodynamics. The free energy of unfolding is expected to be temperature dependent as previously discussed [106]. Particularly, differences between the temperature dependent heat capacity of the native and unfolded state contribute to temperature dependent rates [106]. However, these differences between the fold/unfolded heat capacity are difficult to accurately measure and for mAbs does not account for highly non-Arrhenius behavior observed across many proteins and formulation conditions.

It difficult to measure temperature-dependent aggregation rates from accelerated conditions to refrigerated temperatures over a relatively short temperature window. Other degradation mechanism, such as fragmentation have been shown to occur at intermediate temperatures [107] (30-50 °). At these temperatures fragmentation rates are often faster than aggregation rates, but at lower temperatures aggregation rates become faster again. At room temperature and refrigerated conditions, aggregation may not follow a non-native pathway. Prior work has suggested that IgG1 at high protein concentrations and under refrigerated conditions may aggregate through the formation of reversible native dimers [80]. If this is the case, the non-native and

native mechanisms may have different pathways and intermediate states, which may help explain non-Arrhenius behavior.

1.8 Anti-streptavidin IgG1

1.8.1 IgG1 structure

This dissertation investigates the aggregation behavior of anti-streptavidin immunoglobulin gamma-1 (AS-IgG1) monoclonal antibody provided by Amgen. Gamma denotes the family of heavy chains and one denotes the subclass. IgG1 is the most common antibody, however other subclasses or families can be found *in-vivo*. Figure 1.2 shows the structure of IgG1 from a space-filled model (A) and block diagram (B). In panel A of Fig. 1.2, positive charged amino acids are colored in red, while negatively charged amino acids are shown in blue. The net charge will be positive in conditions explored in this work as the solution pH is always below the pI for this IgG1. However, as mentioned above the IgG1 net charge will be a function of pH.

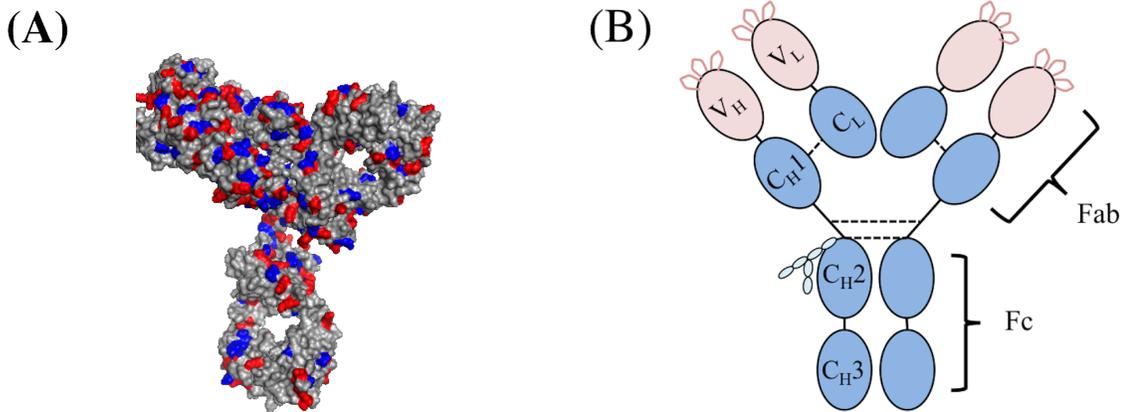


Figure 1.2: IgG1 structure from (A) a front view and (B) in a block diagram. Arginine, lysine, and histidine amino acids are colored in red and aspartic acid and glutamic acid are colored blue.

Panel B of Figure 1.2 illustrates the regions and domains of IgG1. The protein is comprised of a fragment crystallizable (Fc) region and two fragment antigen-binding

(Fab) regions. The Fc region includes the constant heavy chain (C_H) 2 and 3, while the Fab includes the C_H1 domain, variable light (V_L), and variable heavy (V_H) chains. The variable chains of the Fab contain three complementary-determining regions (CDRs), which are depicted as light red loops in Fig. 1.2-B. CDRs regions is responsible for the strong and specific antigen binding. The sequence and structure the IgG is highly conserved within a subclass, which enables new IgG molecules to be created by grafting different CDR loops [108].

The heavy chains and light chains are held together by disulfide bonds, which are shown as dashed lines in Fig. 1.2-B. Additionally, the two heavy chains are held together by disulfide bonds in the hinge region of the molecule. IgGs are unique in that this region is flexible and may contribute to non-ideal behavior and interactions of these macromolecules[109]. Fig. 1.2-B also depicts the glycosylation of the IgG1 as smaller light gray chains attached to the C_H2 domain. Typically IgGs include glycans attached to an asparagine residue in the C_H2 domain. Glycosylation is a post-translational modification. Deglycosylated mAbs have been shown to lower thermal unfolding transitions and increased aggregation [110]. In addition, the glycan profile may affect therapeutic efficacy as well as influence pharmacodynamics and pharmacokinetics in patients [111].

1.8.2 Prior work with AS-IgG1

AS-IgG1 has a molecular weight of 142.2 kDa and pI of ~ 9 . A number of studies have investigated the behavior of this same molecule with different stresses. Notably, under acidic conditions ($\text{pH} \leq 4$), AS-IgG1 has been shown to unfold through the C_H2 domain, which has a midpoint temperature of unfolding near room temperature [112]. Kim et al. investigated the aggregation behavior of AS-IgG1 in bulk solution with NaCl concentrations of 0-500 mM NaCl and pH range of 4-6 with 5 mM citrate buffer [1]. Again, at acidic conditions, differential scanning calorimetry showed the C_H2 domain unfolded at temperature significantly below Fab or C_H3 domain and mediated aggregation at low pH. In another study, Franey et al. compared AS-IgG1 to AS-IgG2 and concluded IgG1 subclass exhibited increased stability, which was primarily due to

reduced free thiols [113]. Additionally, the surface-mediated aggregation behavior of AS-IgG1 was investigated by Carpenter and coworkers, which showed pH and NaCl mediated agitation induced aggregation, and the addition of a surfactant dramatically decrease aggregation. Another study found AS-IgG1 preferred to adsorbed very strongly to a hydrophobic surface in a flat orientation [114]. Prior work has investigated AS-IgG1 with many different stresses and formulations. However, the motivation of the current work is not only to add to existing body of work, but provide a systematic approach to understand AS-IgG1 under thermal stress.

1.9 Organization of the dissertation

This dissertation provides a mechanistic understanding of how pH, salt, buffer, and osmolytes affect AS-IgG1 aggregation. The current work offers semi-quantitative tools to predict protein stability and aggregation behavior. The intent of this dissertation is to offer insights into the role of protein-protein and protein-osmolyte interactions involved in the aggregation process. Providing these tools may help guide predictions of protein behavior.

Chapter 2 investigates the relationship among protein-protein interactions, aggregation mechanism and resulting aggregate morphology. In particular, this work explores specific-ion-effects between AS-IgG1 and the citrate or acetate anions. Colloidal interactions are discussed in the context of protein-protein interactions and aggregation. The aggregate morphology is explored in terms of the aggregate mass to size ratio and small angle scattering profiles. We offer a semi-quantitative tool to predict aggregation mechanisms from protein-protein interactions.

Chapter 3 exhibits the development of the Parallel Temperatures Initial Rate device and method, which accurately and efficiently determines aggregation rates. Results highlight the influence of solution pH, NaCl concentration, and buffer species on protein stability. We provide insights into the balance between protein-protein interactions and conformational stability. The PTIR device and method enable others to study different polymer or protein-based systems and degradation processes.

Chapter 4 determines AS-IgG1 protein-osmolyte and protein-water interactions using densimetry. Here we explore AS-IgG1 preferential interactions with sucrose, trehalose, sorbitol and PEG. Kirkwood-Buff analysis provides a model-free approach to explore preferential interactions of osmolytes and water. These preferential interactions are directly related to AS-IgG1 native state chemical potential. Results are compared to available models, which predict preferential interactions based on AS-IgG1 solvent exposed surface area. AS-IgG1 thermal unfolding with differential scanning calorimetry is compared to preferential interaction models. We discuss classical preferential interaction theories in the context of current results.

Chapter 5 extends AS-IgG1 aggregation to elevated protein concentration, in the presence of sucrose, and in heavy water (D_2O). We evaluate how hypotheses developed in Chapter 2 and 4 apply to a new formulation. We examine differences in the aggregate structure and morphology for different aggregation mechanisms using various orthogonal spectroscopic techniques and evaluate how aggregation changes at higher protein concentration. Results also explore morphology changes during aggregation. In particular, *in-situ* small angle neutron scattering and combined dynamic light scattering with Raman spectroscopy are exhibited with different aggregation mechanisms and in the presence and absence of sucrose.

Chapter 2

SPECIFIC-ION EFFECTS ON IGG1 AGGREGATION BEHAVIOR

2.1 Introduction

As discussed in Section 1.7, there are no *a priori* methods available to predict protein aggregation mechanisms. The mechanism directly links the aggregate size and concentration, which may potentially impact immunogenicity. As illustrated in Fig.1.1 the protein aggregation reaction pathway is complex and many possible intermediate states may exist. Unfortunately, often these intermediates are experimentally inaccessible as the concentration of a particular species may be very dilute and their life-times may be very short [115]. As mentioned in Section 1.3.1, the RDS is important in determining the mechanism. While an *a priori* prediction of aggregation mechanisms from the protein sequence is a lofty goal, engineering based and even semi-quantitatively based approaches to predict aggregation mechanisms would enable one to strategically formulate proteins with a limited set of experiments.

Previous work with aCgn [27] and AS-IgG1 [1] showed pH and NaCl concentration mediated protein aggregation mechanisms. However, it did not consider that the type of ions in solution might affect the relevant protein-protein interactions and aggregation mechanisms. Specific-ion effects, sometimes referred to as Hofmeister effects (cf. Section 1.6.2), might be anticipated when considering different options for buffer components [66], or the choice of added salts [62] to control properties such as viscosity [64, 37]. While measurable Hofmeister effects typically require relatively large salt concentrations (~ 0.2 -2 M), [60, 116] it was hypothesized in this work that ion-specific effects may manifest even at low salt conditions that are more typical of therapeutic protein products [65].

The present chapter provides a systematic comparison of AS-IgG1 aggregation mechanisms in sodium acetate and sodium citrate buffers. Parts of this chapter are reproduced with permission from reference [117], copyright 2015 American Chemical Society. Citrate and acetate specific-ion effects were characterized via differences and similarities illustrated in the AS-IgG1 aggregation state diagrams, mass-to-size scaling behaviors that change with aggregate size, and the net protein-protein interactions and protein charge (valence). The results also illustrate the potential utility and limitations of using colloidal models, measured protein-protein interactions and effective surface charges for monomeric proteins to predict how changes in solution conditions will affect the dominant aggregation mechanism(s) and resulting aggregate sizes and morphology. Additionally, Appendix A provides further discussion on protein-protein interactions using DLS, details regarding AS-IgG1 net charge measurement using electrophoretic light scattering, and an in-depth investigation of AS-IgG1 in an aggregation resistant solution condition.

2.2 Materials and methods

2.2.1 Sample preparation

Anti-streptavidin IgG1 solution (98 % monomer) was provided by Amgen at a concentration of 30 mg/mL. The protein was dialyzed using Spectra/Por 7 tubing (10 kDa MWCO, Spectrum Laboratories, Santa Clara, CA) as previously reported [42] and filtered after dialysis [1]. The dialysate solutions were prepared with distilled, deionized water (Milli-Q filtration system, Millipore, Billerica, MA), in either 5 mM sodium acetate (Fisher Scientific) or 10 mM sodium citrate (Fisher Scientific), with pH adjusted using NaOH (Fisher Scientific). NaCl (Fisher Scientific) was also added gravimetrically to stock solutions after dialysis for conditions that included additional NaCl. The protein concentration was checked after dialysis using UV-Vis absorbance at 280 nm (Agilent 8453 UV-Vis, Agilent Technologies, Santa Clara, CA) and the IgG1 extinction coefficient was 1.586 mL/mg cm [1]. All solutions were diluted gravimetrically to working concentrations for later use.

2.2.2 Size exclusion chromatography with inline light scattering (SEC-MALS)

IgG1 samples prepared at 1 mg/mL for a given choice of solution conditions and were incubated isothermally in hermetically sealed HPLC vials (Waters, Milford, MA) for specified incubation times. Once removed from the water bath, a given vial was immediately quenched on ice to rapidly arrest aggregation and was subsequently held at room temperature (20-23 °C) prior to analysis with SEC-MALS. An Agilent 1100 HPLC (Agilent Technologies, Santa Clara, CA) was connected in-line to a Tosoh TSK-Gel 3000xL size-exclusion-column (Montgomeryville, PA), a Multi-angle-light-scattering (MALS) DAWN-HELEOS II (Wyatt, Santa Barbara, CA) and a Wyatt Optilab rEX refractive index detector. Samples were injected with an autosampler, and the monomer was separated from the aggregate portion with the SEC column. As the sample eluted from the column, a variable wavelength detector (VWD Agilent technologies, Santa Clara, CA) measured the protein concentration from the absorbance at 280 nm.

The sample monomer fraction was determined by integrating chromatograms as previously reported[42]. Following the VWD, the MALS measured light scattering intensity and the RI detector measured the differential refractive index. The latter determined the eluting protein concentration for the monomer and aggregate peaks. All data were collected using ASTRA software and analyzed as previously reported[27, 43]. The overall or total weight average molecular weight (M_w^{tot}) of a given sample across both the monomer and aggregate peaks was calculated using Equation 2.1, as previously described [118]. Here the molecular weight and concentration of the i th slice is given by (M_i) and (c_i) respectively. The values of M_w^{tot} were also confirmed to be equal to those obtained from batch MALS, suggesting larger aggregates did not remain on the SEC column. For aggregates with radius of gyration (R_g) larger than ~ 15 nm, there was statistically significant angular dependence over the Q range from light scattering in SEC-MALS and the R_g was determined using the ASTRA software [119].

$$\frac{M_w^{tot}}{M_0} = \frac{\sum_{i=1}^n M_i c_i}{\sum_{i=1}^n c_i} \quad (2.1)$$

2.2.3 Protein-protein interactions via laser light scattering

Static light scattering was performed using MALS with a Microcuvette accessory (Wyatt Technologies, Santa Barbara). Calibration was performed with filtered toluene at 90° scattering angle. The scattered intensity showed no angular dependence for IgG1 monomer, as expected because the laser wavelength (658.9 nm) is much larger than the size of the IgG1 ($R_g \sim 4.7$ nm) [120]. The excess Rayleigh ratio was calculated according to Equation 2.2. A_{inst} is the configuration-specific constant generated by the MALS calibration procedure, which is based on the scattering geometry so as to recover the correct value for the known value of the Rayleigh ratio of toluene for the laser wavelength (λ) and temperature of interest. V and V_0 are the 90° scattering voltages for the sample and buffer respectively. V_{laser} and V_{dark} are incident laser voltage and its dark offset. n is the refractive index of the solvent. Eq. 2.2 is based on a ray-tracing algorithm designed for the Microcuvette Assembly [121].

$$\frac{R_{ex}^{90}}{K} = \frac{A_{inst}(V - V_0)}{(V_{laser} - V_{dark})} n^{1.983} \quad (2.2)$$

The Kirkwood Buff integral for protein-protein interactions (G_{22}) was determined from a plot of excess Rayleigh ratio and protein concentration (see Eq.2.3). K is the canonical light scattering calibration constant which includes dn/dc (0.186 mL/g for this IgG1) and the laser wavelength (658.9 nm). M_2^{app} is the apparent molecular weight and c_2 is the protein concentration [122].

$$\frac{R_{ex}^{90}}{K} = M_2^{app} c_2 + G_{22}(c_2) c_2^2 \quad (2.3)$$

In the limit of low c_2 (or more rigorously, when the absolute value of $c_2 G_{22}$ is less than approx. 0.1) one recovers the canonical expression for R_{ex}^{90} in terms of the second osmotic virial coefficient (B_{22})[122]. Under such dilute or weak-interaction conditions, $B_{22} = -2G_{22}$ [122]. At higher protein concentrations or when net protein interactions

are larger in magnitude, the canonical expression is incorrect and Eq. 2.3 should be used. For cases where interactions are so large as to cause G_{22} to change with c_2 , Eq. 2.3 was used to fit R_{ex}^{90}/K as a function of c_2 with sliding windows of c_2 to obtain $G_{22}(c_2)$ [122]. As described elsewhere, this procedure allows one to obtain numerical values of $G_{22}(c_2)$ without needing to assume the mathematical form for $G_{22}(c_2)$ [122].

2.2.4 Determination of radius of gyration for small oligomers with SAXS

Aggregates with a radius of gyration larger than approximately 15 nm showed angular dependence with static light scattering by MALS (see above). For smaller aggregates, the radius of gyration was determined from the Guinier analysis (Equation 2.4) of small angle x-ray scattering (SAXS) intensity profiles. $I(0)$ is the scattering intensity in the limit of Q approaching zero, R_g is the radius of gyration, and Q is the magnitude of the scattering vector defined in Equation 2.5 using constants above and the scattering angle, θ [119].

$$I(Q) = I(0) \exp\left(-\frac{(QR_g)^2}{3}\right) \quad (2.4)$$

In Eq. 2.4, SAXS experiments were performed on the Cornell High Energy Synchrotron Source (CHESS) beamline G2 with a dual Pilatus 100K-S SAXS/WAXS detector. Samples were oscillated during exposure to minimize radiation damage during the measurement. The data were reduced and analyzed with BioXTAS RAW software [123]

$$Q = \frac{4\pi n}{\lambda} \sin(\theta/2) \quad (2.5)$$

2.2.5 IgG1 aggregate hydrodynamic radius from dynamic light scattering

Dynamic light scattering (DLS) was performed using the WyattQELS accessory (Wyatt Technologies, Santa Barbara) installed within the HELEOS II instrument, to analyze aggregated samples at low protein concentrations where structure-factor contributions are minimized. The correlation function was collected by the autocorrelator

and exported from the ASTRA software. Data were analyzed using non-linear regression (in Matlab) to fit each correlation function to a cumulant expansion given by Eq. 2.6 [124, 125].

$$g_2(\tau) = \alpha + \beta \exp(-Q^2 D_0 \tau) \left(1 + \frac{\mu_2}{2} \tau^2\right)^2 \quad (2.6)$$

α is a constant for the short delay-time baseline, β is an instrument specific constant, D_0 is the self diffusion coefficient (when protein concentrations are low), τ is the decay time, μ_2 is the second cumulant and is related to the sample polydispersity index, and Q is the scattering vector defined above.

The average hydrodynamic radius (R_h) was determined from the Stokes-Einstein relation, Eq. 2.7, where k_b is the Boltzmann constant, T is the absolute temperature, and η is the solution viscosity.

$$R_h = \frac{k_b T}{6\pi\eta D_0} \quad (2.7)$$

2.2.6 IgG1 net charge (valence) via electrophoretic light scattering

Electrophoretic mobility was measured with a Malvern Zetasizer Nano ZS (Malvern Instruments, Malvern, UK) using laser doppler velocimetry (LDV) and phase analysis light scattering (PALS). IgG1 samples at 10 mg/mL were prepared along with corresponding buffers. Folded capillary cells (Malvern Instruments, Part # DTS 1061) were filled with buffer and 90 microliters of protein sample was pipetted with a gel electrophoresis tip to the bottom of the cell near the scattering volume. As the IgG1 molecules move in the electric field (E), the scattered light produces a phase shift as a result of the Doppler effect [126, 127, 128]. Data were collected using the multimodal standard operating procedure in the Zetasizer DTS software. Polyethylene glycol (PEG) with a molecular weight of 35 kDa was used as an electro-osmotic flow tracer. PEG 35 kDa was chosen as it is neutral molecule and has a hydrodynamic radius comparable to that of the IgG1 [129]. The phase shift (ϕ) as a function of time (t) was collected for a minimum of four different electric field strengths with IgG1 and

PEG. Velocity (v) was determined from linear regression of the phase shift plotted against time, using Eq.2.8[128].

$$\frac{d\phi}{dt} = Q \cdot v \quad (2.8)$$

Electrophoretic mobility (μ) and 95 % confidence intervals were determined from linear regression (in Matlab) of IgG1 velocity defined as ($v = v_{sample} - v_{PEG}$) vs. E . Multiple field strengths were used in an attempt to provide more accurate determinations of v , given that proteins are much smaller than typical colloidal particles for which the PALS method was originally developed [56]. IgG1 effective charge (Z_{eff}) was determined from the mobility using Equation 2.9 [56]. D_0 is the self-diffusion coefficient for IgG1 monomer, which is determined from separate DLS measurements as described above, and e is the charge of an electron. Additional details regarding electrophoretic mobility measurements are provided in Appendix A.

$$Z_{eff} = \mu \frac{k_b T}{D_0 e} \quad (2.9)$$

2.2.7 Aggregate morphology using small angle neutron scattering (SANS)

Small angle neutron scattering was performed on the NGB 10 m beamline at the NIST Center for Neutron Research (Gaithersburg, MD). All samples were measured using three configurations: 1 m sample-to-detector distance (SDD) with 5 Å neutrons, 4 m SDD with 5 Å neutrons, and 4 m with 16 Å neutrons. The data were corrected for detector background and sensitivity, as well as the scattering contribution from empty titanium cells. The protein scattering profiles were also normalized by incident beam flux and the raw intensities were placed on an absolute scale using direct beam measurements. All data were reduced and analyzed using Igor Pro NCNR software with standard methods [130].

Protein samples were prepared for conditions that formed aggregates via the different growth mechanisms in both citrate and acetate buffers as described above. Aggregates were separated from the monomer on a semi-prep SEC column (Waters,

Milford, MA). Aggregate fractions eluting from the column were concentrated and buffer exchanged to conditions in D₂O rather than H₂O so as to achieve better scattering contrast and reduce incoherent background scattering. For each sample, the buffer was matched to the original NaCl concentration, buffer concentration, and the pD was matched to the pH of the original solutions [131]. The total scattering is given by Equation 2.10. $I(Q)$ is the scattered intensity, $\Delta\rho$ is the neutron scattering-length-density difference between the protein and buffer, V is the partial specific volume of the protein, $P(Q)$ is the particle form factor, and $S(Q)$ is the structure factor [119]. Scattering was performed on samples with aggregate concentrations between 3 to 6 mg/mL to minimize contributions from the structure factor. In subsequent analysis, $S(Q)$ was assumed to be effectively one.

$$I(Q) = \frac{(\Delta\rho)^2 \hat{V}^2}{N_a} c_2 M_w P(Q) S(Q) \quad (2.10)$$

2.3 Aggregation mechanism(s) from SEC-MALS

Differences between aggregation mechanism(s) were evaluated using SEC-MALS. For each pH, NaCl concentration, and choice of buffer species, isothermal incubations were performed at temperatures to achieve monomer loss half-lives of approximately two hours. These incubation conditions allowed for adequate temperature control when removing samples from the incubator for multiple time points, which reduced artificial lag times due to sample heating at much shorter time scales [31]. SEC-MALS analysis was performed to quantify monomer fraction (m) and total molecular weight (M_w^{tot}) of each sample (which includes all aggregate species and monomer), as a function of the sample incubation time at elevated temperature. Monomer fraction was determined from a given SEC chromatogram. It is defined as the monomer peak area for a given sample, divided by that for the initial, unheated sample. M_w^{tot} was determined from a weighted summation over M_i slices in from the chromatogram as described by Eq. 2.1.

Differential scanning calorimetry (DSC) measurements were also performed as a function of pH and NaCl concentration to provide reference temperatures for selecting

incubation temperatures, as aggregation rates under accelerated conditions are well known to be sensitive to changes in calorimetric peak temperatures (T_m) [132]. DSC results will be presented in Section 3.3. However, it is important to note all incubation temperatures used here were at or below the lowest T_m value for a given pH and NaCl concentration. As such, the concentration of “reactive” monomers will be pre-equilibrated during aggregation. While aggregate nucleation, growth, and coalescence may change with temperature, the categorized aggregation mechanism is expected to be qualitatively similar at different incubation temperatures. The DSC profiles and detail IgG1 aggregation kinetics will be explored in Chapter 3.

Figure 2.1-A shows illustrative SEC chromatograms; the relative concentration (right y-axis) and normalized molecular weight profiles (left axis) are plotted as a function of retention volume for different types of aggregation mechanisms. Each profile is an SEC chromatogram for either an unheated sample or a selected incubation time for a heated sample that illustrates different qualitative behavior in terms of SEC-MALS after a significant amount of monomer loss. The black curve (unheated sample) shows IgG1 monomer eluting as a single peak at a retention volume of approximately 8.2 mL. The blue curve illustrates the ND mechanism, which produces a mixture of dimers, trimers, and small oligomers. The green curve illustrates the CP mechanism, which is a combination of what is observed for ND, along with high molecular weight (HMW) soluble species that elute in the exclusion volume (near 5.5 mL) and have much larger M_w values. The red curve illustrates a SEC-MALS profile when the AP growth mechanism predominates, as determined by the kinetic and mass-balance arguments below. Essentially only monomer and HMW species are observed, although one must perform the analysis below to unambiguously conclude that AP growth is prevalent. The presence of large aggregates is not sufficient to conclude an AP mechanism is relevant [1, 31, 26, 27].

The mechanisms can be distinguished qualitatively by relating M_w^{tot} to increases in the amount of monomer consumed by aggregation; $(1 - m)$ denotes the fraction of the initial monomer population that has been consumed. Previous work has shown

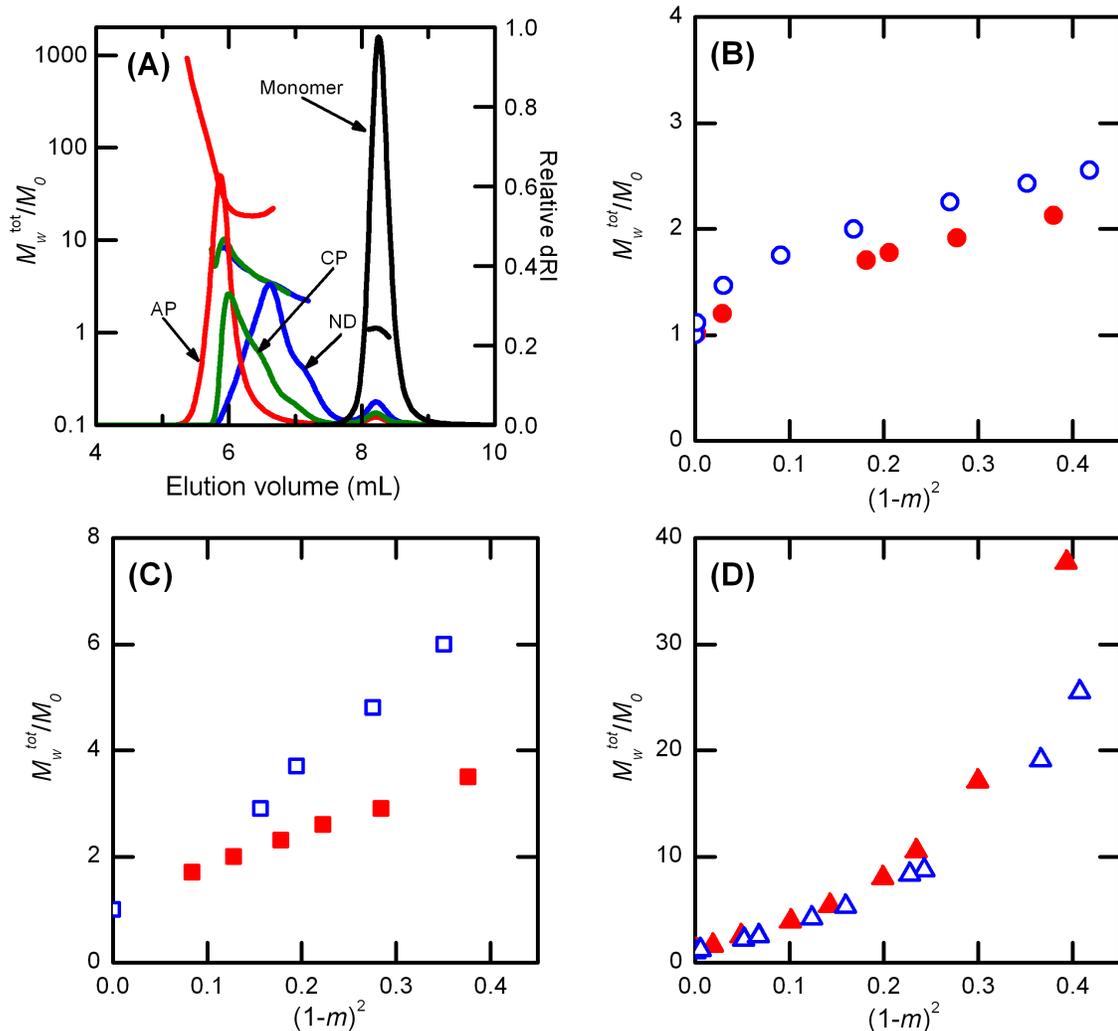


Figure 2.1: SEC-MALS results distinguishing different aggregation mechanism. (A) Representative chromatograms for IgG1 monomer, the nucleation-dominated (ND), chain polymerization (CP), and association polymerization (AP) growth mechanisms. M_w^{tot} vs $(1 - m)^2$ plots showing aggregation through (B) ND, (C) CP, and (D) AP growth mechanisms. Blue open symbols represent aggregates from 5 mM citrate buffer and red solid symbols are for 10 mM acetate buffer.

the relationship between M_w^{tot} and $(1 - m)^2$ depends on the relative rates of aggregate nucleation (e.g., dimerization in the present case), aggregate growth by chain polymerization or monomer addition, and growth via aggregate-aggregate coalescence or association polymerization [1, 26]. One can attempt to deduce mechanisms by regressing mathematical mass-action models to time-dependent data for $m(t)$ and $M_w^{tot}(t)$, but once significant growth occurs via AP then the quantitative fitting parameters become suspect because multiple models can fit the data equally well [26].

More generally, it has been shown that one can deduce qualitative mechanisms without a need for assuming a particular mathematical model simply by plotting $M_w^{tot}(t)$ and $(1 - m(t))^2$ parametrically, as shown in Figures 2.1-B-D. Each data point in Fig. 2.1-B-D corresponds to a time point from isothermal incubations at a given temperature, pH, salt concentration, and buffer type. This analysis also allows one to plot results from different solution conditions or time scales on the same scale for comparison.

The profiles in Fig. 2.1-B-D correspond to each of the representative aggregation mechanisms that were observed in 5mM citrate solutions (blue open symbols) and 10mM acetate solutions (red symbols). Figure 2.1 illustrates the ND mechanism for citrate and acetate buffer conditions. M_w^{tot}/M_0 increases essentially linearly with $(1 - m)^2$ but only reaches values less than three ($M_0 = \text{monomer } M_w$), even though the extent of monomer loss is almost 70 % ($m = 0.3$ gives $(1 - m)^2 = 0.49$ on the x axis). Figure 2.1-C illustrates aggregation through a CP mechanism. In this case, M_w^{tot} also increases linearly with $(1 - m)^2$ after an initial small non-linear region as the first dimers are formed. For growth via CP, M_w^{tot}/M_0 continues to increase significantly beyond that for small oligomers. Figure 2.1-D illustrates growth via AP, and possibly also CP at low extents of monomer loss. That is, there is a linear scaling between M_w^{tot} and $(1 - m)^2$ at early time points (low values of $1 - m$), but then M_w^{tot} increases nonlinearly in Figure 2.1-D at later times. The nonlinear behavior indicates growth via AP because aggregate coalescence does not consume monomer, and therefore M_w^{tot}/M_0 vs. $(1 - m)^2$ must curve upward, because there is an increase in M_w^{tot} without a

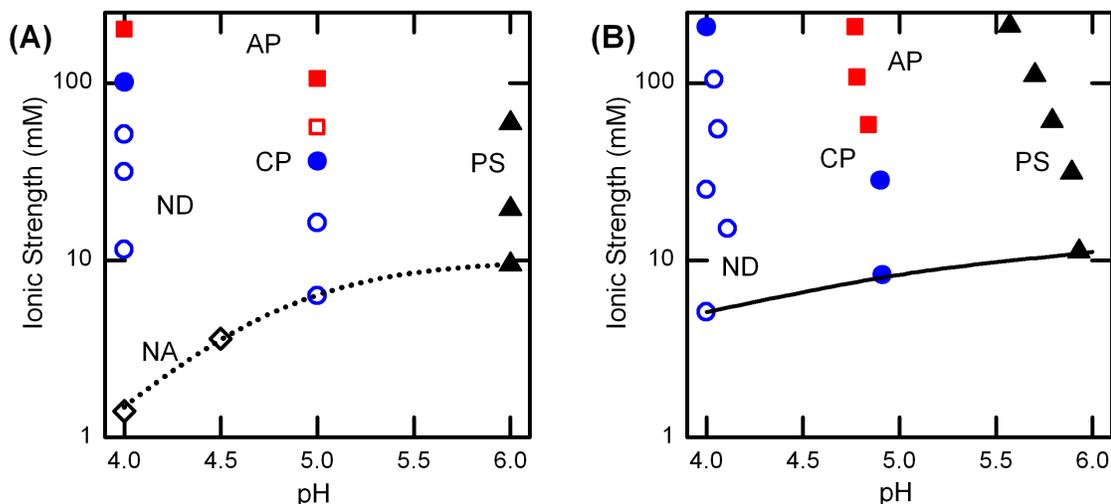


Figure 2.2: AS-IgG1 aggregation state diagrams for (A) 10 mM acetate and (B) 5 mM citrate (Reproduced from reference [1]). Predominant aggregation mechanism is shown for no aggregation (open diamonds), ND (open circles), CP (filled circles), AP (filled squares), mix of CP and AP (open squares), and PS (closed triangles). The dotted curved lines illustrate the ionic strength of the (A) acetate and (B) citrate buffer component(s).

concomitant decrease in monomer concentration[38, 1, 118].

The different mechanism(s) of aggregate growth are mapped in Figure 2.2 as aggregation state diagrams as a function of pH and ionic strength for 10 mM acetate buffer (panel A) and 5 mM citrate buffer (panel B).

Each data point on the state diagram represents aggregation behavior from a time course determined through SEC-MALS analysis illustrated in Figures 2.1-B-D. The data in Figure 2.2-B were reported previously and confirmed separately here (data not shown).²⁰ The dotted lines in Figure 2.2 show the lowest possible ionic strength that is practically achievable, given that all conditions include either 5 mM citrate or 10 mM acetate. As pH increases, the acetate or citrate ion(s) become deprotonated as one titrates the buffer with NaOH, and the net ionic strength increases because of the higher sodium and buffer anion concentrations that must result.

In terms of general trends in Figure 2.2, at low pH and ionic strength the mechanism is ND, but as the pH or ionic strength increase the mechanism shifts to CP then AP. At higher pH and NaCl concentrations, the aggregates are ultimately insoluble and phase separate (PS). There are also a number of differences between the two state diagrams in Fig. 2.2. One of the most pronounced is at pH 4 and low ionic strength. The acetate system exhibits no aggregation despite extreme heat treatment under those solution conditions. For example, IgG1 solutions heated at 1 mg/mL over multi-hour time scales at temperatures above T_m for this pH and ionic strength in acetate buffer did not result in detectable aggregation in SEC-MALS or by inspection (data not shown). Another pronounced difference between the acetate and citrate systems was a shift of aggregation mechanism at pH 5 and low ionic strength. The acetate system at pH 5 exhibits ND behavior, while aggregates grow readily via CP in citrate buffer. More subtle effects are the general shift in the boundaries between ND and CP, or CP and AP mechanisms. When moving from acetate to citrate buffer under conditions without significant added NaCl, inspection of Figure 2.2 shows that one typically finds that growth is less favored in acetate than in citrate.

2.4 Aggregate mass to size scaling

Aggregate average molecular weight and size scaling by laser light scattering For aggregates that were sufficiently large to display angle-dependent scattering, aggregate average molecular weight and size scaling was investigated with static and quasi-elastic light scattering. SEC separates the monomer from the pool of aggregates and allows one to characterize the average aggregate molecular weight and R_g via MALS. The weight-average molecular weight of the pooled aggregate population (M_w^{agg}) was determined from SEC-MALS using the aggregate peak when it was resolvable from the monomer peak. R_g was determined as described in Methods via the angular dependence of the scattering intensity. For aggregates created under ND conditions, MALS could not be used effectively to determine reliable values because the aggregates were too small. In those cases, SAXS was measured on aggregated samples to determine the

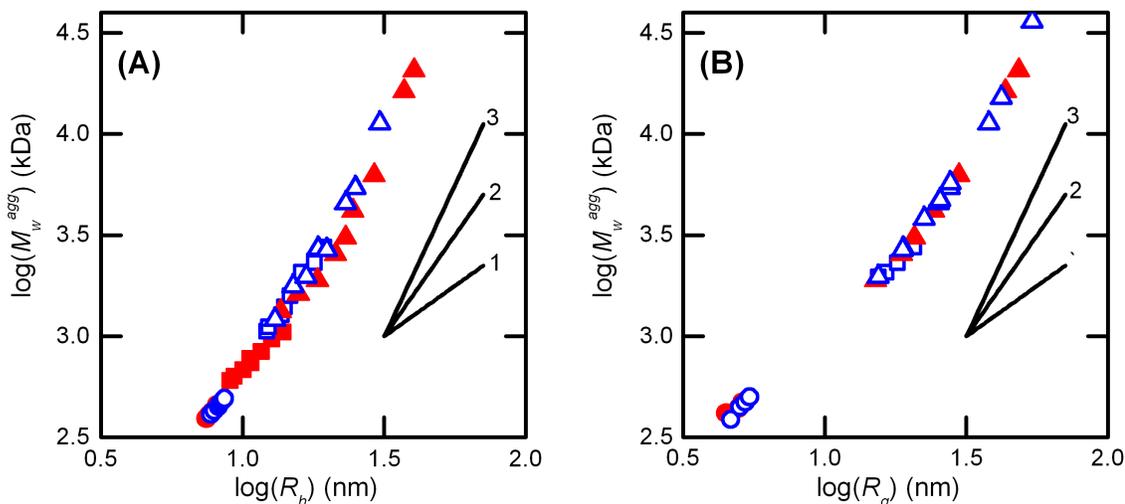


Figure 2.3: Scaling of average aggregate size with molecular weight based on (A) R_h or (B) R_g . ND (circles), CP (squares), and AP (triangles) depict different aggregate growth regimes. The black lines show power-law scaling relationship is shown for slopes of 1,2, and 3. Open blue (closed red) symbols represent aggregates created in citrate (acetate) buffer.

average radius of gyration using Guinier analysis (not purified with SEC). Values for the average aggregate R_h at a given incubation time were determined from the quasi-elastic light scattering collected simultaneously with SEC-MALS as described in the Methods section (cf 2.3).

The scaling relationship between the aggregate size (characteristic dimension, R_g or R_h) and molecular weight provides insight into how mass is distributed within aggregates. Figure 2.3-A (2.3-B) shows power-law scaling relationships between M_w^{agg} and R_h (R_g). Each data point corresponds to a time point for an isothermal incubation at a given temperature, pH, NaCl concentration, and buffer species. The symbols are labeled as ND (circles), CP (squares), AP (triangles), and symbol colors are labeled as blue (5 mM citrate) and red (10 mM acetate). Power-law scaling relationships with slopes of 1, 2, and 3 are also shown as straight lines in the inset of Figure 2.3 for reference.

Inspection of Figure 2.3 shows that data from different aggregation mechanisms tend to overlap with one another and fall on a common curve, which indicates the scaling regimes in Figure 2.3 correspond to different size regimes or degree of polymerization (M_w^{agg}/M_0) for the aggregates, rather than being dictated by the growth mechanism. The one exception is that aggregates that reached the largest sizes ($M_w \geq 100$ times that of monomer) only did so after AP (i.e., coalescence) became prominent, as observed previously in other systems [1, 133, 134]. Qualitatively, the scaling of M_w with R_g is not strictly linear, and shows slight upward curvature when viewed globally across all of the values from solution conditions tested here. Within the statistical scatter in the data, it is not possible to assign separate, distinct linear regimes. Visual inspection of the data in Figure 2.3 indicates that the values of the local slopes for M_w vs. R_g or R_h fall with the range of ~ 1.5 to slightly larger than 2 for smaller aggregates but the slope increase to values larger than 2 for largest aggregates. This suggests that aggregates become more compact, or have a net higher fractal dimension as they become larger. However, additional details regarding the average shape of the aggregates is difficult to discern from just the scaling behavior of M_w vs. R_g or R_h [135, 136].

2.5 Aggregate morphology using SANS and SAXS.

The average morphology of the aggregates was also characterized with small angle neutron or x-ray scattering, SANS or SAXS, which provides a complementary technique to MALS and scaling behaviors from laser light scattering. Figure 2.4-A (main panel) shows log-log plots of SANS scattering intensity profiles for the ND growth regime in acetate versus citrate buffer, after purification to remove residual monomer (cf., Methods). The profiles are shown after background subtraction and normalization to the total protein concentration. The points in Figure 2.4-A are the measured data, and the curves are best fits to geometric models for simple shapes that can be programmed analytically in available software so as to allow efficient data regression [130]. The inset of Figure 2.4-A shows SANS profiles for aggregates created

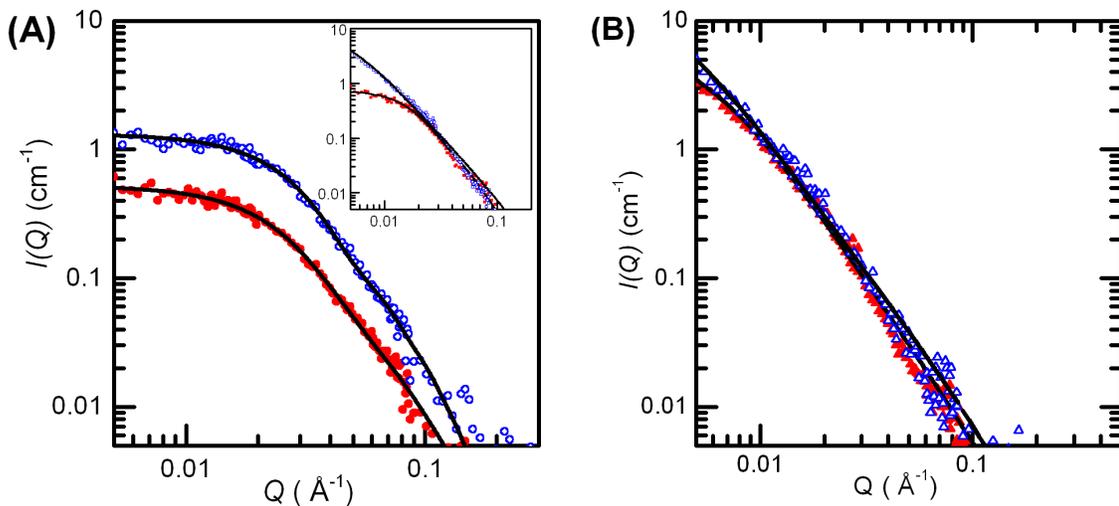


Figure 2.4: SANS intensity profiles for purified aggregates from (main panel A) ND and (inset) CP and (panel B) AP growth mechanism. Closed red symbols are shown for 10 mM acetate buffer and open blue symbols for 5 mM citrate. The black lines are best fits to form factor models from the NCNR software.

in the CP regime for acetate and citrate buffer conditions.

Aggregates in the ND regime have qualitatively similar morphologies, as both conditions display relatively simple, sigmoidal profiles. Aggregates from the ND regime and CP regime (acetate) had best fits to triaxial ellipsoid form factors. Aggregates from the CP regime (citrate) and AP regime had best fits to a fractal with flexible cylinder subunits.

SANS intensity profiles were fit to a variety of form factors available in the NCNR analysis macros [130]. Aggregates created in ND and CP regimes are shown in the main text. Aggregates created in AP regime are shown in Figure 2.4-B. These large aggregates were outside the SANS Q -range to determine reliable estimates on size and morphology. The SANS profiles for AP regime were fit to a model of a fractal with flexible cylinder subunits. Tables 2.1 and 2.2 reports the shape dependent parameters for the form factor with the best fit as indicated by the lowest chi squared parameter.

Figure 2.5 illustrates SAXS results, which are presented as normalized Kratky

Table 2.1: SANS fitting parameters I

Model parameter	ND-10 mM acetate	ND-5 mM citrate	CP-10 mM acetate
Semi-axis A (smallest) (\AA)	13 ± 1	16 ± 0.8	18 ± 1
Semi-axis C (middle) (\AA)	75 ± 4	79 ± 4	105 ± 4
Semi-axis C (largest) (\AA)	128 ± 6	109 ± 5	188 ± 7

Table 2.2: SANS fitting parameters II

Model parameter	CP-5 mM citrate	AP-10 mM acetate	AP-5 mM citrate
Fractal dimension	13 ± 1	16 ± 0.8	18 ± 1
Correlation length (\AA)	2.2 ± 0.1	2.34 ± 0.12	2.2 ± 0.1
Contour length (largest) (\AA)	229 ± 5	170 ± 3	304 ± 9
Kuhn length (largest) (\AA)	7.0 ± 0.1	7.6 ± 0.7	8 ± 0.3

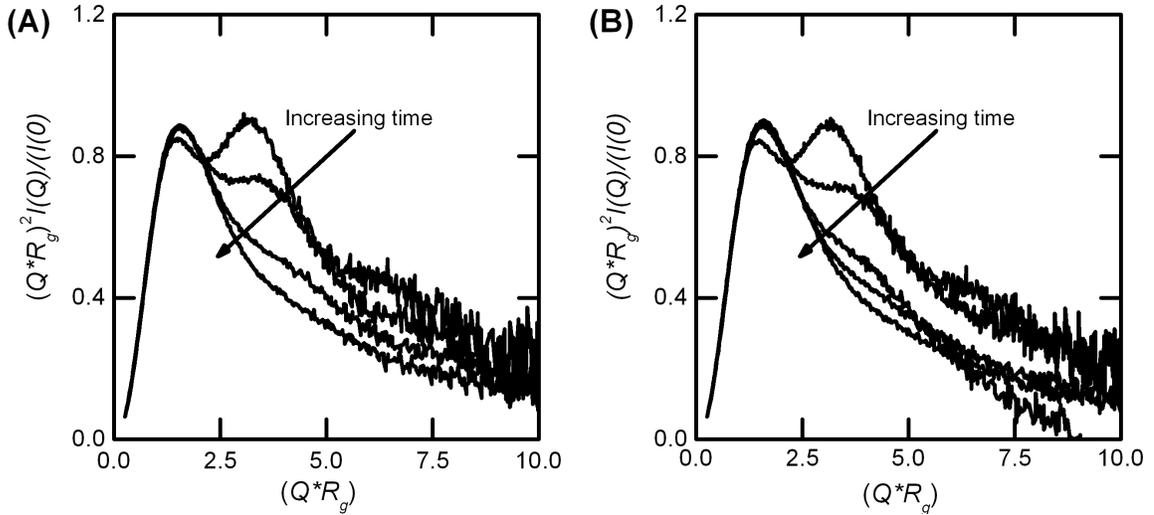


Figure 2.5: Normalized SAXS Kratky plots for aggregates from ND growth mechanism created in (A) 10 mM acetate and (B) 5 mM citrate. Scattering intensity was normalized to the low Q limit and the Q was normalized to R_g

plots for the ND mechanism in 10 mM acetate (A) and 5 mM citrate (B). Each curve is a quenched sample from an isothermal incubation at the temperature(s) used to generate the data in Figures 2.5. SAXS Kratky plots for different isothermal time points were compared by normalizing the intensity $I(Q)$ to the low Q intensity $I(Q = 0)$, and scaling Q by the R_g value for a given sample based on the Guinier analysis. $I(Q)$ curves plateau at low Q because the aggregates do not grow to be large in the ND mechanism, and the plateau corresponds to the weight average molecular weight of the sample. As noted elsewhere, scaling the SAXS data in this way allows one to normalize SAXS Kratky plots by size and molecular weight and compare differences in morphology[137].

Interestingly, ND growth mechanisms both citrate and acetate produce similar morphologies. The curve for IgG1 monomer in Figure 2.5 has two peaks; one at $QR_g = 1.5$, which corresponds to a length scale similar to the ensemble averaged size; and another peak at $QR_g = 3.1$, which corresponds to length scales smaller than the average size. The secondary peak at these shorter length scales has previously

been attributed to correlations in the positions of the Fc and Fab domains due to the flexibility in the hinge region of the native IgG1 [109, 138].

As aggregation proceeds, the left-most peak in Figure 2.5 increases and shifts to slightly larger $Q * R_g$ (~ 1.55). Interestingly, aggregated samples for different time points all have the same location of the peak maximum when plotted versus QR_g , regardless of the amount of monomer remaining or how much aggregate growth has occurred. This suggests the average aggregate morphology from the ND regime is compact and remains qualitatively similar, regardless of how the aggregates are distributed amongst dimer, trimer, tetramer, etc.

2.6 Protein-protein interactions based on static light scattering

Average protein-protein interactions for AS-IgG1 monomer were investigated using static light scattering and Kirkwood Buff analysis. Illustrative results for the excess Rayleigh ratio as a function of protein concentration are shown in Figure 2.6 for each pH and buffer combination investigated. The Kirkwood Buff integral for protein-protein interactions, G_{22} , was determined using linear regression to Eq. 2.3 when a single value of G_{22} provided a good description of the data. When G_{22} changed appreciably with protein concentration, a sliding-window method was instead more appropriate. pH and buffer dependent scattering results are provided as Fig. 2.6, with details of the regression method provided elsewhere [122].

Physically, G_{22} provides a measure of net attractions or repulsions between IgG1 molecules in the solution, averaged over all neighboring proteins around a given “central” protein. It includes simultaneous contributions from multiple “neighboring” proteins, and as such is a valid and rigorously defined measure of protein-protein interactions in both dilute and concentrated solutions, and for both weak and strong protein-protein interactions [122]. G_{22} was determined at pH 4, 5, and 6 with the addition of 0, 20, and 200 mM NaCl concentrations in 5 mM citrate buffer and in 10 mM acetate buffer. Salt concentrations were chosen to provide a range of Debye-Hückel screening lengths within the same ranges as those for the aggregation state diagrams in

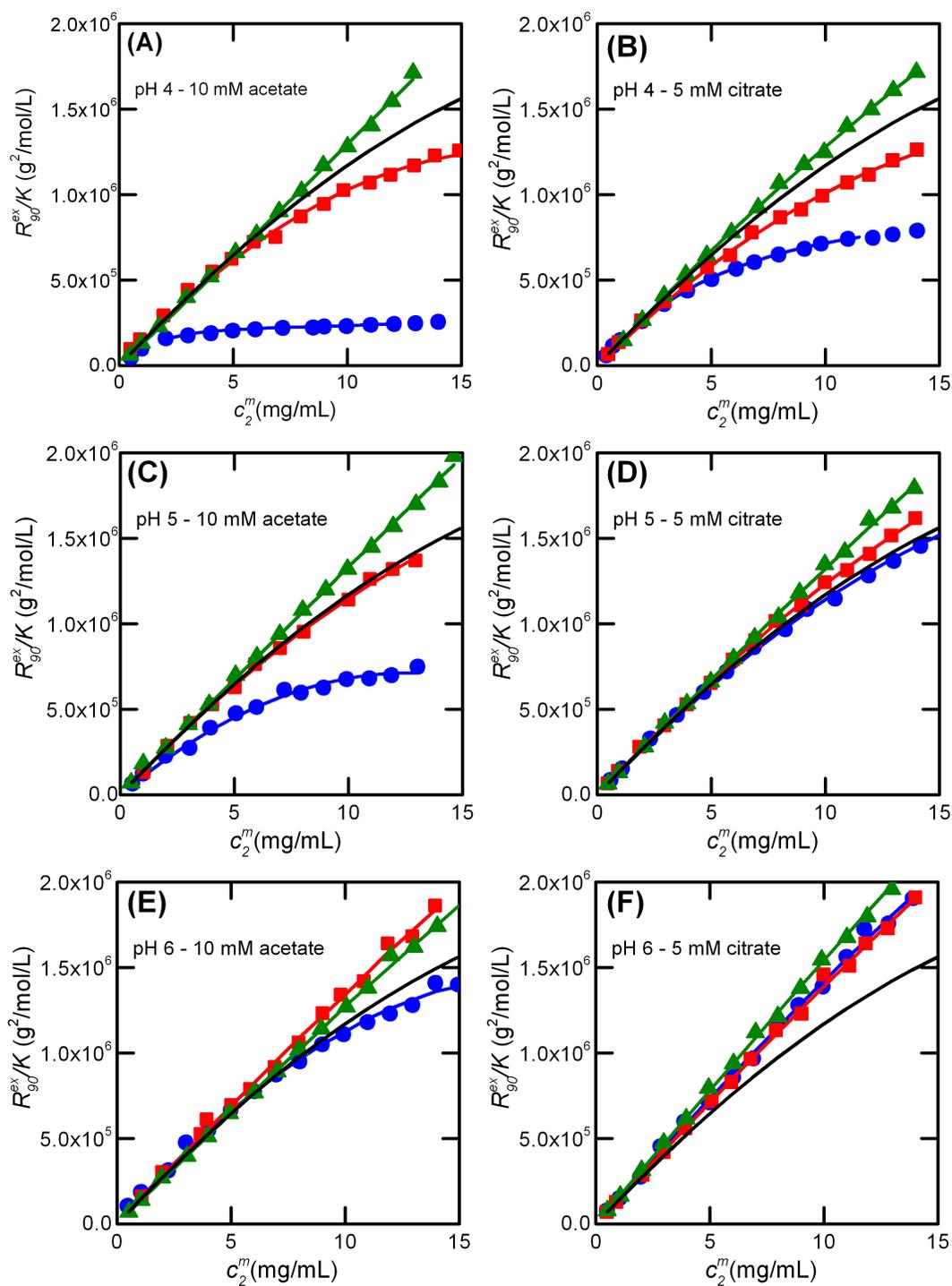


Figure 2.6: (A) Excess Rayleigh ratio vs. IgG1 concentration for pH 4, 5, and 6 in 5 mM citrate and 10 mM acetate. Refer to the text on the panel for exact pH-buffer conditions. Colored curves are fits to Eq. 2.3 and the black curve is the Rayleigh ratio for a hard sphere with equivalent diameter to the IgG1.

Fig. 2.2. The values of G_{22} are normalized by the value of B_{22} that corresponds to hard sphere or steric-only interactions, denoted as B_{22}^{HS} , as suggested by previous theoretical arguments [139]. The reduced or normalized G_{22} is defined as $G_{22}^* = -G_{22}/(2B_{22}^{HS})$. Using this definition, positive values of G_{22}^* correspond to net repulsive protein-protein interactions relative to an ideal (non-interacting) system, and vice versa for negative values. A G_{22} value of unity is expected for purely steric repulsions at low protein concentrations.

Figures 2.7-C-D show G_{22}^* values measured in 10 mM acetate buffer and 5 mM citrate respectively. The pI of this IgG1 is ~ 9 , and the molecule is expected to have a net positive charge at all solution conditions tested here. As pH is decreased farther away from the pI, the net charge on the protein surface will increase and lead to larger electrostatic repulsions, consistent with the results shown in Figure 2.7. With the addition of 200 mM NaCl, charge screening might be expected to be sufficient to minimize electrostatic repulsions between protein molecules. The results in Figure 2.7 show that G_{22}^* does not become significantly less than zero, but is less than one, with the addition of 200 mM NaCl. This suggests weak attractive non-electrostatic interactions are present, even if electrostatic repulsions are not completely screened by 200 mM NaCl.

2.7 Protein net echarge determined by electrophoretic light scattering (ELS)

IgG1 net charge (valence) was inferred from electrophoretic mobility measurements as a function of pH, NaCl concentration, and buffer species using electrophoretic light scattering. Electrophoretic mobility was determined by measuring the IgG1 velocity in solution under the influence of a steady electric field, as a function of electric field strength, and the friction coefficient or protein diffusion coefficient was determined from DLS (cf. Methods). These quantities were combined to give values for Z_{eff} via Eq. 2.9.

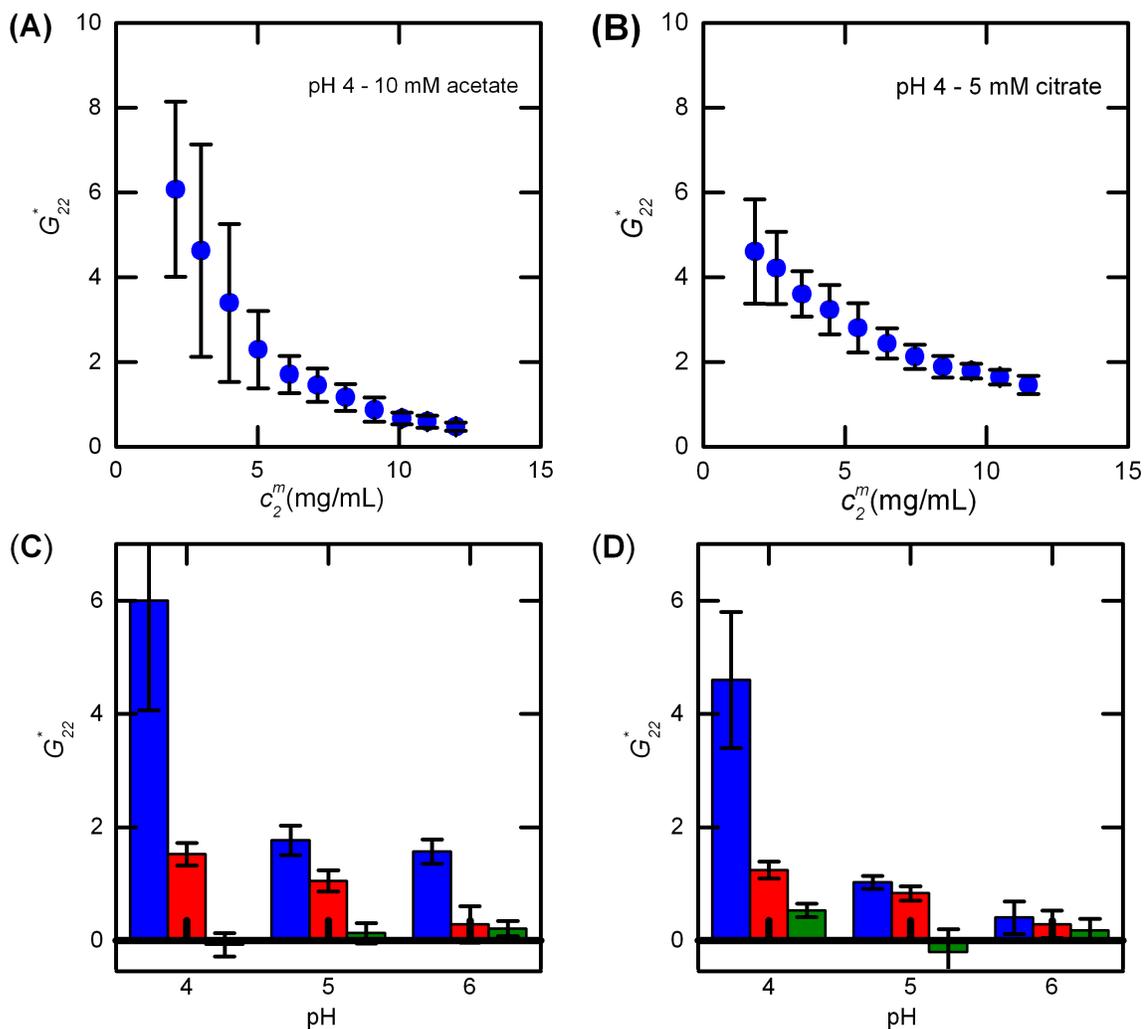


Figure 2.7: G_{22}^* values determined as a function of protein concentration for pH 4, 0 mM NaCl in (A) 10 mM acetate and (B) 5 mM citrate. At other pH and NaCl concentrations, G_{22} is not a function of concentration in the range of protein concentration tested. G_{22}^* values are shown for (C) 10 mM acetate and (D) 5 mM citrate buffer, for 0 mM NaCl (blue), 20 mM NaCl (red) and 200 mM NaCl (green). Error bars represent 95% confidence intervals.

Figure 2.8-A shows illustrative results for monomer IgG1 velocity determined using PALS in pH 4, 10 mM acetate as a function of electric field strength for 20 mM NaCl (squares), 40 mM NaCl (circles), and 100 mM NaCl (triangles). No detectable electro-osmotic flow (EOF) was observed using PEG as a neutral control. Figure 2.8-B shows PEG velocity in pH 6, 5mM citrate as a function of electric field strength for 20 mM NaCl (squares), 40 mM NaCl (circles), and 100 mM NaCl (triangles). EOF was observed in pH 6, 5 mM citrate conditions. In cases where EOF was significant, the PEG velocity was used to quantify the EOF and the net electrophoretic mobility for the protein was determined by subtracting the EOF contribution from the IgG1 velocity. See Appendix A for more details regarding PALS.

Figures 2.8-C-D show Z_{eff} of monomer IgG1 as a function of pH and added NaCl concentration, based on fits of electrophoretic velocity vs. electric field strength. Panel A (B) corresponds to acetate (citrate) buffer conditions. In all cases, PEG was used as a neutral control to detect electro-osmotic flow (EOF). Measurable EOF was only observed at pH 6 in 5 mM citrate conditions and was corrected for, as described in the Methods section.

Given the inherently large statistical uncertainties in the measured Z_{eff} values, only qualitative trends can be drawn with confidence from the results in Fig. 2.8. For both buffer types the net charge was positive at pH 4, but was statistically indistinguishable as a function of NaCl concentration. At pH 5 in acetate buffer, the net charge was positive. However, at pH 5 in citrate buffer, the net charge was statistically insignificant from zero; except at high NaCl concentration, where the data were consistent with charge inversion due to anion binding. The much lower net charge at pH 5 with citrate might be a result of preferential accumulation of the divalent citrate anions at the protein surface.

Table 2.3 reports calculated concentrations of the different buffer anion species in solution at pH 4, 5, and 6 based on pKa(s) of acetate and citrate,[140] illustrating the shift from predominantly monovalent anions for citrate at low pH, to divalent and

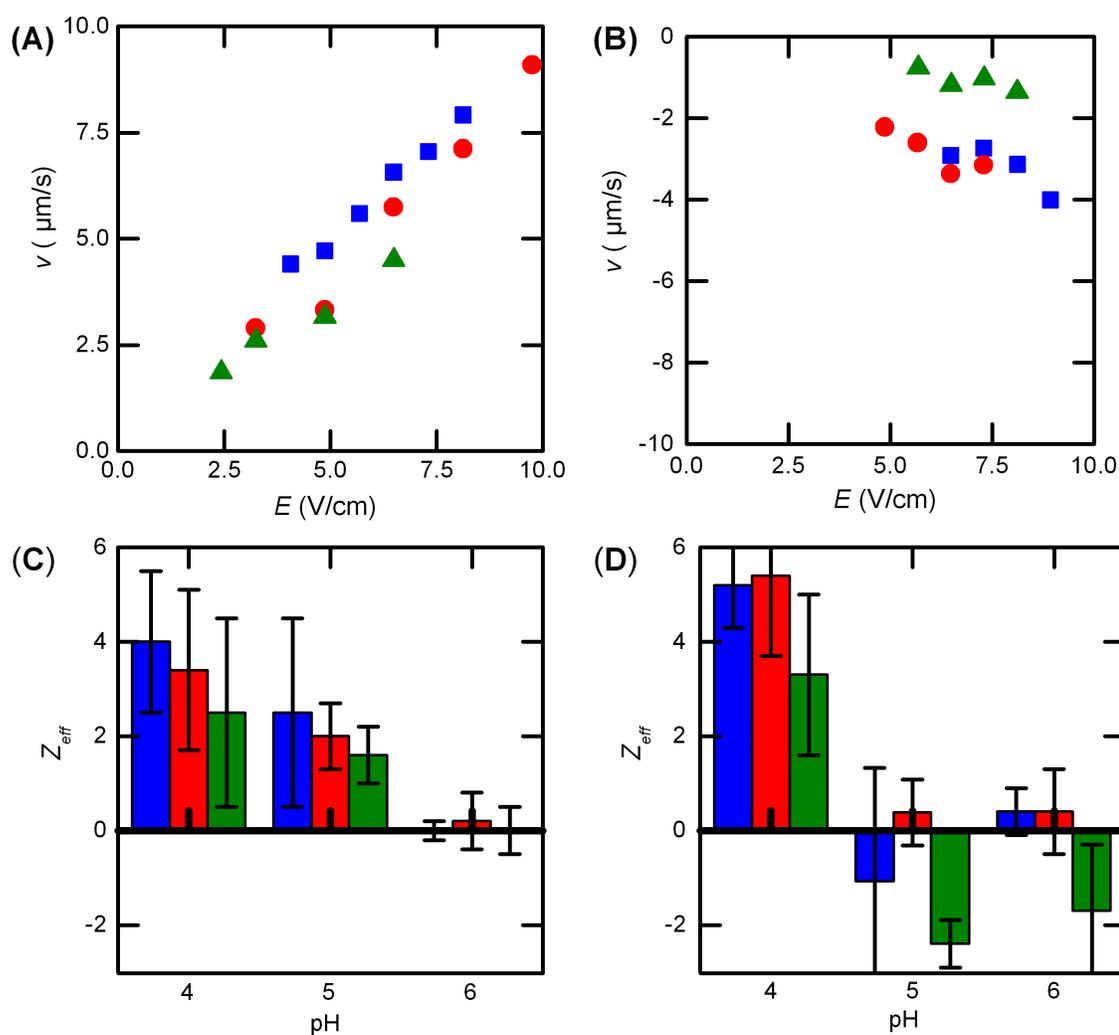


Figure 2.8: IgG1 net charge determined from ELS analysis. (A) IgG1 velocity plotted vs electric field for pH 4, 5 mM citrate for 20 mM NaCl (blue squares), 40 mM NaCl (red circles), and 100 mM NaCl (green triangles). (B) PEG 35kDa velocity plotted vs electric field for pH 6, 5 mM citrate for 20 mM NaCl (blue squares), 40 mM NaCl (red circles), and 100 mM NaCl (green triangles). PEG was used as a tracer to detect electro-osmotic flow (EOF). The net effective charge was determined from ELS analysis for 10 mM acetate (C) and 5 mM citrate (D) for 20 mM NaCl (blue), 40 mM NaCl (red) and 100 mM NaCl (green). Error bars are 95% confidence intervals.

Table 2.3: Buffer anion concentration as a function of pH. Concentration of acetate ion is the first row followed by citrate ions.

Buffer anion	pH 5	pH 5	pH 6
$\text{C}_2\text{H}_3\text{O}_2^{-1}$ (mM)	1.5	6.4	9.5
$\text{C}_6\text{H}_7\text{O}_6^{-1}$ (mM)	3.8	1.8	0.19
$\text{C}_6\text{H}_6\text{O}_6^{-2}$ (mM)	0.66	3.1	3.4
$\text{C}_6\text{H}_5\text{O}_6^{-3}$ (mM)	0.0026	0.12	1.4

trivalent anions at higher pH conditions. While preferential accumulation of multivalent citrate anions may also be anticipated at pH 6 due to both divalent and trivalent anions, this may be too weak of an affect to be apparent experimentally if the shift to pH 6 causes the inherent surface charge on the protein to be greatly reduced.

2.8 Role of electrostatic protein-protein interactions on IgG1 aggregation mechanisms

As Figures 2.1 and 2.2 illustrate, aggregation of ASA-IgG1 proceeds via different pathways that depend on the solution conditions. Depending on the pH, NaCl concentration, and identity of the buffer species, different mechanisms predominate. In all cases the aggregates were effectively irreversible; no aggregate dissociation was observed upon dilution or upon standing over multi-day time scales for any of the conditions tested. The solution conditions effectively dictated: (i) whether aggregates would remain small (i.e., ND, meaning minimal growth of aggregates); (ii) or would grow by CP or by AP to create much larger soluble, aggregates; (iii) or would effectively phase separate (PS) as macroscopic particles that sediment easily on the bench or under mild centrifugation.

As pH increases towards the pI, the IgG1 surface charge decreases. The pH will also affect the concentration of different buffer ions in solution. The solution ionic

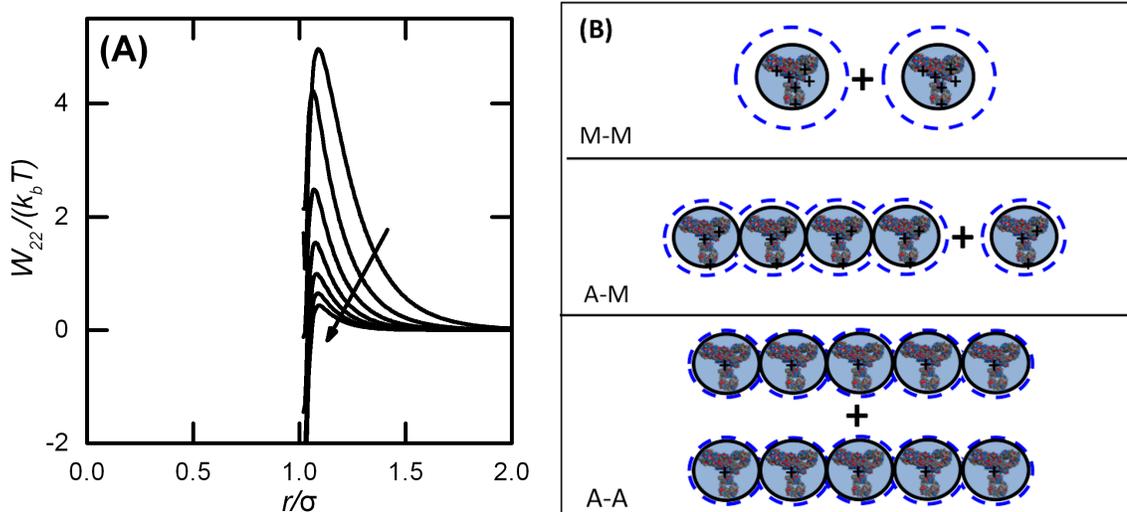


Figure 2.9: Schematic depiction of electrostatic colloidal interactions causing shifts in observed aggregation mechanisms. (A) Potential mean force $W_{22}/k_b T$ plotted as a function of center-to-center distance for two charged hard spheres with the same hydrodynamic radius as an IgG1. The blue arrow shows the effect of decreasing Z_{eff} (changing pH towards the pI) and increasing ionic strength (increasing NaCl concentration). (B) Schematic describing monomer-monomer (M-M), monomer-aggregate (M-A), and aggregate-aggregate (A-A) association (cf., main text).

strength will be affected by the buffer pKa value(s) and the addition of NaCl (0 - 200 mM in the present work). At a mean-field level, ions in solution create a Debye double layer, which screens the electrostatic repulsions felt between IgG1 monomers, as well repulsions between monomers and aggregates, and those between aggregates.

Figure 2.9-A schematically illustrates the effect of pH and ionic strength on a hypothetical potential of mean force, W_{22} , between two IgG1 molecules. $W_{22}/k_b T$, is plotted versus the center-to-center distance scaled by the effective diameter of IgG1 monomer, r/σ . For illustration, the electrostatic interactions between IgG1 molecules are described simply using a screened Coulomb potential [141]. Z_{eff} values based on a multivariate fit to the experimental values in Fig. 2.8 were input into the colloidal

model along with screening lengths calculated using Debye-Hückel theory and the relevant ionic strength.

The arrow indicates the shifts in W_{22}/k_bT for increasing pH and NaCl concentration. As the value of Z_{eff} is reduced with increasing pH, and/or the screening length decreases with increasing ionic strength, the electrostatic repulsions felt between two proteins decrease. While the W_{22}/k_bT profiles in Fig. 2.9-A are for two IgG1 interacting monomers, analogous profiles could be created for an IgG1 monomer interacting with an aggregate, or for interactions between two aggregates. In these latter cases, the aggregate geometry will affect the energy barriers, as illustrated below.

Figure 2.9-B depicts other scenarios for association of different species using simplified geometries to qualitatively illustrate each aggregate growth mechanism (ND, CP, and AP). In each case, two species (monomer or aggregate) come together as part of the process to form a larger species. IgG monomers are shown simply with black circles, and when circles are touching it denotes an aggregate that has more than one constituent monomer. The blue dotted circle traced around each black circle represents the thickness of the Debye double layer. Pairs are shown as monomer-monomer (M-M), monomer-aggregate (M-A) in an end-to-end orientation, and aggregate-aggregate (A-A) in the same plane with a side-to-side orientation.

M-M association is the dominant aggregation event for the ND growth regime observed at low pH and ionic strength. The protein surface has large positive charge, which is represented pictorially as a series of positive signs on the IgG1 surface. At low ionic strength, the Debye double layer extends well beyond the protein surface. At these solution conditions the electrostatic contributions to the potential mean force for M-M interactions will be strongly repulsive. Colloidal interaction for other geometries (M-A and A-A) would have even larger repulsive forces because as M approaches A, or A approaches another A, repulsions will be “felt” simultaneously by multiple constituent proteins within an aggregate. As a result, growth beyond small aggregates (e.g., dimers, trimers) is greatly disfavored.

As the pH and ionic strength are increased, association between species presumably becomes more favorable because the protein surface charge will be decreased, and the Debye double layer thickness will decrease with addition of ionized buffer species. As a result, M-A association becomes more viable from a kinetic perspective, as the repulsive interactions have decreased. Aggregate-aggregate association (A-A) will also become more viable at higher pH and ionic strength. At pH values approaching the pI, and or at elevated ionic strength, M-M and M-A interactions will also be expected to have insignificant electrostatic energy barriers, and therefore all mechanisms will occur simultaneously.

These simple geometries and colloidal arguments based on the potential of mean force are consistent with the qualitative and semi-quantitative behavior of the aggregation state diagrams in Fig. 2.2. However, one needs to bear in mind that aggregation is a multi-step pathway with many possible rate-determining steps. As shown in Fig. 2.3 and as discussed below, the average molecular weight and size scaling of aggregates appears to depend primarily on size (i.e., extent of polymerization). The results suggest that electrostatic interactions cause changes in the relative rates of the different growth mechanisms, rather than the morphology of the resulting aggregates. The colloidal arguments above can only explain energy barriers smaller than $\sim 10 k_b T$. One needs to include non-colloidal factors, such as aggregation prone regions or “hospots” that are specific to particular sequences in the protein [142] if one is to explain why monomers aggregate despite such large barriers [1, 143]. The colloidal model cannot explain specific-ion effects observed in the aggregation mechanism at low the NaCl concentration. If an anion preferentially accumulates, such as citrate, then the Debye double layer will effectively be shorter than what mean field arguments would predict, and this would result in greater screening of electrostatic repulsions.

2.9 Aggregate morphology from scattering

A combination of laser light scattering, SAXS, and SANS was used to characterize aggregate morphology for each mechanism. Using the power-law scaling relationship

between the aggregate molecular weight and radius of gyration one can gain insight into how the mass is distributed around the aggregate center of mass. The exponent in Equation 2.11 is defined as the fractal dimension [136].

$$M_w^{agg} \sim R_g^{d_f} \quad (2.11)$$

Illustrative values of the fractal dimensions for simple objects include $d_f=1$ for a long rod; $d_f=1.86$ for Diffusion-Limited Colloidal Aggregates (DLCA) [144]; $d_f=2$ for flat disks; $d_f=2.1$ for Reaction-Limited Colloidal Aggregates (RLCA) [145], and $d_f=3$ for uniform spheres. As the fractal dimension increases, the mass becomes more uniformly distributed around the object's center of mass. The concept of a fractal is best applied for large disordered architectures with self-similarities where the subunit is much smaller than the overall object [146]. Fractal behavior occurs for small aggregates, but the fractal concept works well for aggregates composed of 20 subunits and is a reasonable approximation for even smaller aggregates [147].

In the present case, Fig. 2.3 shows that d_f increases as the aggregates become larger. The value of d_f appears to primarily be determined by the size (or molecular weight) of the aggregates, rather than the solution conditions or presence of citrate versus acetate anions. However, once one considers that larger aggregates may have multiple "branch" points that would lead to a higher d_f but with aggregate shapes that are highly non-uniform, it becomes clear that one should not interpret the scaling results as being more than qualitative or semi-quantitative in the present context.

The measured SANS and SAXS profiles do not provide a means to further differentiate morphologies, but are qualitatively consistent with the results from MALS. For dimers and similarly small aggregates, it is difficult to define how compact or extended they are in terms of a value for d_f based on Fig. 2.3. However, the SAXS and SANS data are consistent with dimers being more globular than the folded IgG. As aggregates grow to intermediate sizes by CP or AP, they adopt relatively low fractal dimension structures ($d_f \sim 2$ in Fig. 2.3). The upward curvature in Fig. 2.3 shows that

aggregates become somewhat more compact (higher d_f) at the largest sizes. Preliminary results from cryo-transmission electron microscopy (not shown) are qualitatively consistent with higher fractal-dimension structures for larger aggregates. More quantitative structural conclusions are difficult to justify because of known limitations with scattering-based techniques, including bias towards scattering by large objects, and the fact that aggregate growth necessarily increases the polydispersity of the aggregate population [26].

Finally, one must bear in mind that aggregate nucleation and growth by addition of monomers both require some degree of unfolding to allow aggregation-prone “hot spots” to be exposed. This necessarily means that there are only a select number of geometric ways that proteins can form strong “connections” with one another as they form aggregates of different sizes. As such, one might anticipate that aggregates formed through different mechanisms (e.g., CP vs. AP) should have significantly different d_f values. The present data do not provide a clear assessment of this hypothesis. This is due, at least in part, to the fact that CP and AP mechanisms occur simultaneously, just with different net rates. As such, many of state points and time points upon which Figures 2.3-2.4 are based likely include aggregates from both mechanisms.

2.10 G_{22}^* : semi-quantitative tool to predict aggregation mechanism

Finally, we consider whether G_{22}^* is semi-quantitatively predictive for which AS-IgG1 aggregation mechanisms predominate for a given solution condition. G_{22} is a measure of net PPI, including repulsions due to screened electrostatic and steric interactions, and any electrostatic and non-electrostatic attractions. G_{22} is rigorously related to the integral over the potential of mean force between proteins, and determination of its experimental value does not require one to assume an underlying model for those interactions or extrapolate from a narrow set of solution conditions [148, 122]. The potential of mean force, $W_{22}(r)$, integrated over the intermolecular center-to-center distance, r , can be related to G_{22} , as shown in Equation 2.12 [122].

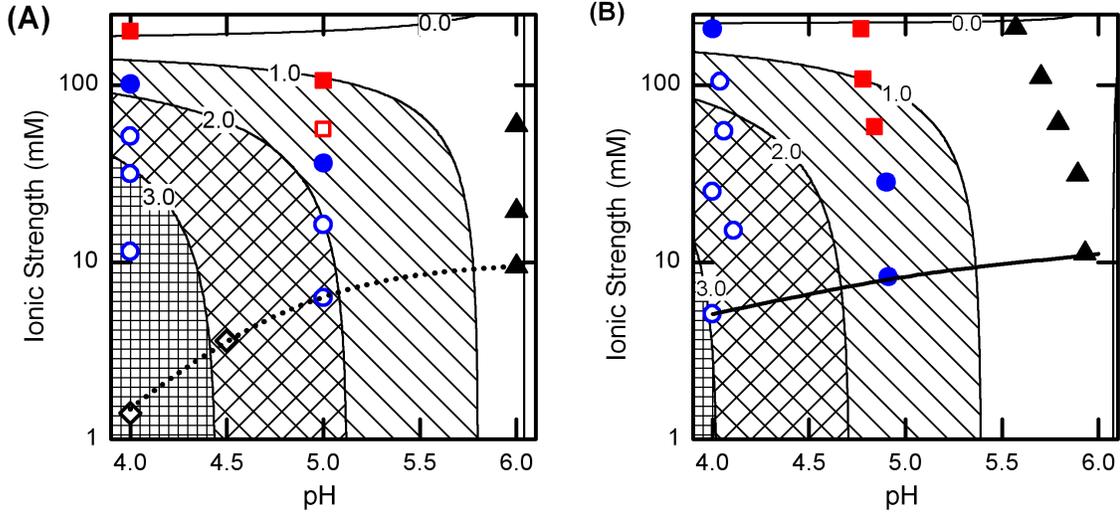


Figure 2.10: Overlay of the IgG1 aggregation state diagrams with G_{22}^* surfaces from multivariate regression for (A) 10 mM acetate and (B) 5 mM citrate. Labels are G_{22}^* values determine from the multivariate regression. Symbols have same representation as Fig 2.

However, unlike what was used in Fig. 2.9, W_{22} can have multi-body contributions that are not accounted for by dilute-solution treatments.

$$G_{22} = \int_0^\infty \exp(-(W_{22}(r))/(k_bT)) - 1) 4\pi r^2 dr \quad (2.12)$$

In the arguments above regarding the simple colloidal model for why the state diagrams have the qualitative features that are observed, no non-idealities were accounted for in the protein-protein interactions beyond those based on mean-field colloidal arguments (e.g., net charge and screening lengths compared to protein dimensions). However, G_{22} values (Figure 2.7 and 2.6) clearly show that there are differences in protein-protein interactions when one considers the acetate and citrate systems at lower NaCl concentrations. While the colloidal model above illustrate the role of electrostatic protein-protein interaction, it fails to capture specific-ion effects. However, the measured values for G_{22} do not suffer from this limitation.

A multivariate surface was fit to experimental G_{22}^* values determined from SLS (additional details, including fitted coefficients and p-values, are provided in supporting information). Figure 2.10 superimposes the G_{22}^* response surfaces as a function of pH and ionic strength with the corresponding aggregation state diagrams for acetate (panel A) and citrate (panel B) conditions. G_{22}^* correlates semi-quantitatively with the aggregation mechanism in citrate and acetate. When net repulsive interactions are sufficiently large, aggregation proceeds through ND (G_{22}^* between approximately 2 and 3 for both acetate and citrate systems). As G_{22}^* decreases with increasing pH and ionic strength, the aggregation mechanism shifts to CP growth. G_{22}^* values ~ 1 to 2 correspond to CP growth and transitioning to AP growth. Finally, aggregation occurs via AP and PS as G_{22}^* shifts to values less than unity i.e., when electrostatic repulsions are not sufficient to outweigh (non-electrostatic) attractions. Overall, G_{22}^* appears to be a reasonable predictor of IgG1 aggregation mechanism. While only phenomenological, it offers an experimental measure of net protein-protein interactions, which inherently captures specific-ion effects and does not rely on model regression or simplifying model assumptions

2.11 Summary and Conclusions

Changing the buffer from citrate to acetate had significant effects with respect to shifting the aggregation mechanism(s) or state diagrams, and altering protein-protein interactions for AS-IgG1 as a function of solution conditions (pH, [NaCl]) that are typical for therapeutic proteins to experience during purification, drug product manufacturing, and final product storage. Shifts of aggregate average molecular weight and size scaling appeared to be determined primarily by the size of the aggregates, rather than aggregate mechanism or protein-protein interactions. Simple models qualitatively illustrate how changes in electrostatic repulsions can explain the features of state diagrams, and suggest preferential interactions of proteins with ions can also alter rates of competing aggregation pathways. Those same models cannot predict specific-ion effects such as the difference between citrate and acetate for the AS-IgG1 system.

Independent of the ability to predict specific-ion effects, experimental G_{22}^* values may offer a phenomenological, semi-quantitatively means to predict which aggregate growth mechanism(s) will predominate.

Chapter 3

PARALLEL TEMPERATURE INITIAL RATES: PH, AND COUNTERION EFFECTS ON IGG1 AGGREGATION RATES

3.1 Introduction

A priori predictions of aggregation rates for a given protein formulation remain an outstanding challenge for a variety of fundamental and practical reasons [31, 132, 105]. The solution pH, choice of buffer species, and addition of salt and other excipients may affect conformational stability and/or protein-protein interactions, while temperature changes can dramatically effect conformational stability [44, 31, 1]. Prior work has indicated that conformational stability is a key factor affecting aggregation rates in solution, as the midpoint temperature of thermal unfolding from differential scanning calorimetry, or the onset temperature of aggregation from scanning techniques, is often at least qualitatively predictive of aggregation rates across different formulations [44, 149, 150, 151]. However, there can also be a competing effect between changes in conformational stability and protein-protein interactions as one changes solution conditions such as pH [38, 100].

A number of temperature-scanning techniques have been developed to at least qualitatively or semi-quantitatively monitor aggregation [31, 152, 153]. An inherent issue with temperature-scanning techniques is thermal history. For example, in the process of scanning through lower temperatures, one creates aggregates that can act as “seeds” to accelerate aggregation at subsequent (higher) temperatures, and thus overestimate aggregation rates [12]. It is difficult to predict when this will or will not be the case, as simple changes in the formulation pH and ionic strength can alter aggregation mechanisms and “seeding” effects [27, 107].

A large majority of biophysical techniques that are currently used to rapidly monitor aggregation use an indirect measure of monomer loss rates and are only surrogate measures of aggregation. A direct measurement of monomer concentration necessitates a separation of monomer from aggregate species or the ability to measure a monomer-specific marker. For example, in spectroscopic techniques such as circular dichroism, ThT dye-binding or intrinsic fluorescence, the spectra are ensemble averages. Therefore, they have contributions from monomer and aggregate species, and the spectral changes may or may not correlate with monomer consumption [27].

An indirect measure of monomer loss rates may also have a bias based on the measurement technique. For example, aggregation rates monitored using scattering techniques have a bias towards larger sized particles [154]. pH and ionic strength changes can alter aggregation mechanisms, and produce large and heterogeneous aggregate populations that provide much larger scattering intensities compared to smaller-sized aggregates at an identical monomer loss rates. These challenges are compounded if fragmentation occurs, as is relatively common for mAbs [155, 47, 156] and other proteins [27].

Parts of this chapter have been included in a manuscript submitted to a peer-reviewed publications [157]. This chapter introduces a Parallel-Temperature-Initial-Rates (PTIR) method to accurately and efficiently determine degradation rates as a function of temperature, as well as an apparatus to allow this to be done more effectively. PTIR is compared to rates determined using traditional isothermal incubations, and the method shows good quantitative agreement with aggregation rates for an anti-streptavidin (AS) immunoglobulin gamma 1 (IgG1) that has been reported previously [1, 80, 112, 158, 159]. Aggregation rates from accelerated (high temperature) to near-room temperature conditions are reported across multiple values of pH and NaCl concentration, as well as different buffer species. The results highlight conformational stability as a dominant factor in determining accelerated aggregation rates, but also illustrate the contributions of electrostatic colloidal interactions to the effective or net activation energy values for aggregation. The PTIR method is also illustrated as

a means to to more precisely determine activation energies. This enables additional mechanistic factors to be discernible experimentally, whereas traditional approaches would not have been able to distinguish them.

3.2 Materials and Methods

3.2.1 Differential scanning calorimetry (DSC)

DSC was performed using a VP-DSC (Microcal, Northhampton, MA) for solutions at a given pH and salt concentration (1 mg/mL IgG1 or 0.33 mg/mL Fc-IgG1). Scans were performed from 20 to 90 °C at a 1 °C /min scan rate. If precipitation did not occur after the scan, as indicated by the lack of a large exotherm, a rescan was performed to check for reversibility. None of the conditions that were tested exhibited reversibility upon a rescan. The absolute heat capacity was calculated from the buffer-subtracted DSC scans, as previously reported [42].

3.2.2 Quantifying Aggregation Rates

IgG1 stock solutions were prepared at 1 mg/mL at a given pH, NaCl concentration, and buffer type, and aliquot into hermetically sealed deactivated borosilicate glass HPLC vials (Waters, Milford, MA). Isothermal incubations were performed by heating multiple samples in a water bath or custom-built PTIR device at a given temperature and removing samples at pre-determined incubation times. Incubation temperatures were chosen such that multiple time points could be taken during the early periods of monomer loss ($m = 1 - 0.8$, m is defined as the concentration of monomer divided by the initial monomer concentration, as measured by SEC peak area). At each time point, a given vial was immediately quenched by immersion in an ice-water bath to arrest aggregation, and was subsequently held at room temperature (20-23 °C) prior to analysis with SEC.

Aggregation rates were determined by monitoring the monomer fraction remaining as a function of incubation time. The monomer fraction was quantified using SEC, described above. Over approximately the first ten to twenty percent monomer loss, the

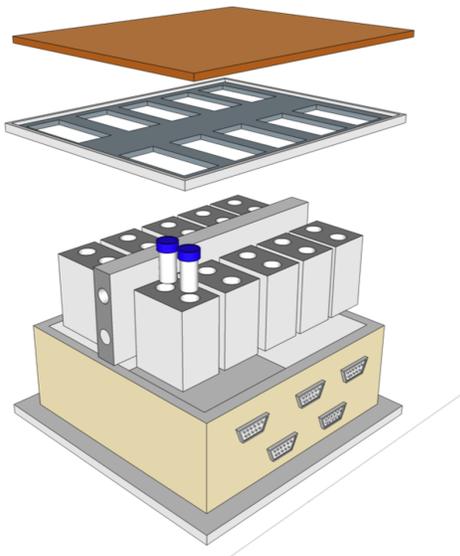


Figure 3.1: Schematic of the PTIR device. The device has ten independently controlled peltier heating block each capable of holding two 1.5 mL HPLC vials. The device allows one to accurately and efficiently determine liquid state polymeric or protein reaction rates.

rate of change of m remains nearly constant and the observed rate law can therefore be well described as zeroth order without the need to assume an underlying rate law [160]. Therefore, the monomer fraction was regressed with Equation 3.1 to obtain the aggregation rate coefficient (units of inverse time), k_{obs} , from the regime where m was between approximately 1 and 0.8.

$$m = 1 - k_{obs}t \quad (3.1)$$

3.2.3 Parallel Temperatures Initial Rates (PTIR)

Figure 3.1 illustrates a schematic of a device to more easily implement the Parallel Temperatures Initial Rates (PTIR) approach. The device was built to simultaneously and independently incubate samples at multiple temperatures; for concreteness, the version shown uses ten independent temperatures [161]. It is composed of 10

peltier-controlled metal blocks; each block has two machined holes capable of holding one 1.5 mL HPLC vial. The peltier-controlled blocks are cooled with a circulating water to maintain stable temperature control. The blocks are insulated from each other and the surroundings.

Incubations were performed by heating a single sample at a given temperature using a water bath or the parallel temperature device in Figure 3.1. The key difference between the PTIR approach and conventional approaches is that for conventional approaches one selects a small number of temperatures (sometimes only one, e.g., 40 °C) and measures multiple samples over a predetermined time course (e.g., as proscribed by ICH guidelines [16]); in the PTIR approach, one instead measures a small number of samples (e.g., one sample in the extreme example below) at multiple temperatures for the same incubation time. That is, rather than choose multiple time points at a given temperature, one chooses multiple temperatures with a given incubation time.

In the present examples, incubation temperatures were chosen so that samples quenched after 2 or 24 hours would have monomer loss values that fell in the initial-rate regime. The incubation time was selected to be not less than 2 hours so as to allow for sufficient temperature equilibration and elimination of artificial lag times at shorter incubation time scales when samples were heating to the set-point temperature. Sample temperatures were confirmed independently with a separately calibrated thermocouple. The longer incubation timescale of 24 hours was chosen to achieve initial rates approximately one order of magnitude slower than 2 hours [36, 27]. For some examples, 10-day incubations were also performed. Once a sample was removed from incubation it was quenched on ice as described above, prior to analysis with SEC.

Aggregation rates using the PTIR approach are based on Eq. 3.2, which is derived by rearranging Eq. 3.1 and solving for k_{obs} .

$$k_{obs}(T) = \frac{(1 - m(T))}{t} \quad (3.2)$$

In Eq. 3.2, it has been shown explicitly that temperature (T) is the variable of interest, as the incubation time (t) is held constant for a given experiment. The PTIR

analysis method is valid for initial-rate conditions, where the rate of degradation remains approximately constant. This is expected to hold for other degradation processes (e.g., chemical degradation) not tested here, as the principle of initial rates in reaction kinetics is more general than just the example shown here [162].

3.3 AS-IgG1 thermal unfolding using differential scanning calorimetry (DSC)

DSC was performed as a qualitative and semi-quantitative measure of IgG1 thermal stability. It is only a surrogate for the true conformational stability (free energy of unfolding, ΔG^{un}), as unfolding in DSC was found to be irreversible and the absolute heat capacity (C_p) was convoluted by a combination of unfolding and aggregation. Figure 3.2-A-C illustrates thermograms for AS-IgG1 and the corresponding Fc fragment at pH 4 (3.2-A), pH 5 (3.2-B), and pH 6 (3.2-C) in 10 mM acetate. Previous work reported DSC thermograms for AS-IgG1 at the same pH and NaCl concentrations, but in 5 mM citrate buffer [1]. In Fig. 3.2, profiles for conditions with 100 mM added NaCl are offset vertically to distinguish them from those with no added NaCl. The peaks of the Fc-IgG1 thermograms overlay with the smaller peaks or shoulders of the full IgG1 thermogram in panels A to C. As expected based on previous reports [1, 161, 163], there only two, relatively small, transitions for the Fc-IgG1 when compared with the full IgG1; the peak at lower (higher) temperature is assigned to the C_{H2} (C_{H3}) domain of the Fc. For the full IgG1, the peak for the Fab domains overlaps with one or both peaks from the Fc domains, depending on the solution pH. The DSC profiles are consistent with the pH dependent thermograms reported previously for a range of other IgG1 molecules [1, 161, 163].

Visual inspection of Figure 3.2 shows that increasing the value of the pH from 4 to 6 increases the temperatures for the calorimetric maxima (T_m values) for all of the peaks for the full IgG1 and for the Fc fragment. The addition of 100 mM NaCl decreases the T_m values in each case. At pH 4, the calorimetric transition for the C_{H2} peak occurs at a significantly lower temperature than the Fab, and the IgG1 thermogram shows three distinguishable peaks. At low pH, previous reports concluded that unfolding of

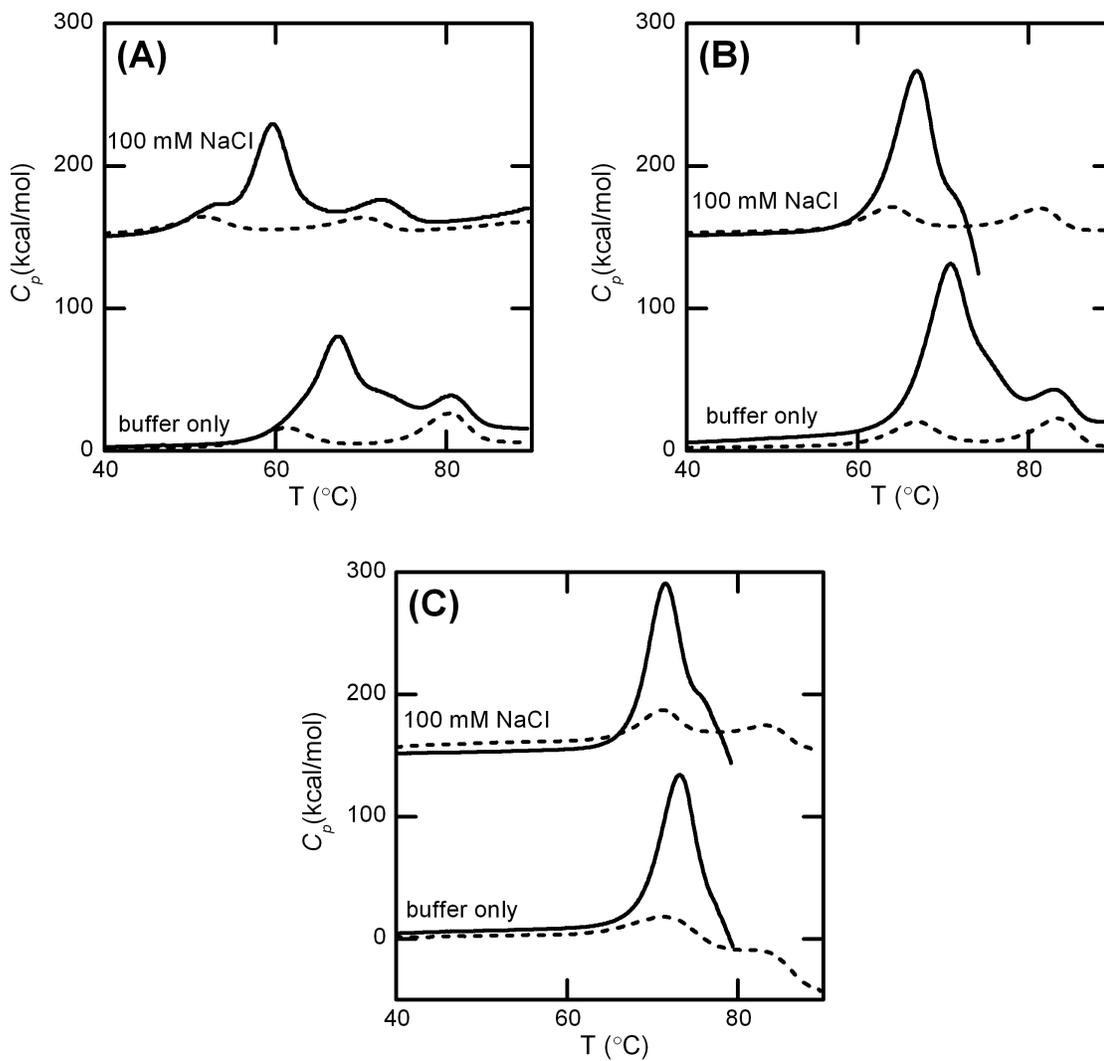


Figure 3.2: DSC for IgG1 (solid) and Fc-IgG1 (dashed) formulated in 10 mM acetate buffer at pH 4 (A), pH 5 (B), and pH 6 (C) with no added salt or 100 mM NaCl added salt (offset 150 kcal/mol). Curves are offset vertically for easier visualization

Table 3.1: DSC apparent T_m values. Error bars are ± 0.2 °C

Formulation	C _{H2}	Fab	C _{H3}
pH 4, 0 mM NaCl	61.3	67.3	80.4
pH 4, 100 mM NaCl	51.5	59.6	69.8
pH 5, 0 mM NaCl	67.0	70.3	83.4
pH 5, 100 mM NaCl	63.9	66.9	81.2
pH 6, 0 mM NaCl	71.0	73.1	N/A
pH 6, 100 mM NaCl	71.2	71.5	83.1

the C_{H2} was a primary step in exposing aggregation-prone sequences for monoclonal antibodies [1, 161, 163]. Recent results indicate that both the C_{H2} domain and the Fab domain(s) within the same protein can contain highly-aggregation prone sequences that become exposed upon unfolding [161].

The results in Figure 3.2 at pH 6 (with and without added NaCl) and pH 5 with 100 mM NaCl display large exotherms (i.e., decreases in C_p , and ultimately exotherms) at higher temperatures. Previous work has shown that this is indicative of irreversible processes such as protein aggregation and precipitation when using this particular instrument configuration [107]. From these results in isolation, it is unclear for pH 5 and 6 whether unfolding of the C_{H2} or the Fab region is most important with regards to promoting aggregation. However, prior work [1] showed that Fab unfolding was likely primarily responsible for aggregation of AS-IgG1 at elevated temperatures.

3.4 Aggregation Rates from PTIR and Standard Isothermal Approaches

The DSC thermograms guided the initial choices for incubation temperatures for accelerated aggregation rates. All incubation temperatures were selected to be below the DSC Fab peak temperatures for a given solution condition, based on the discussion above. Aggregation rates were determined using the PTIR method and quantitatively

compared to those determined by canonical isothermal-rate experiments. Briefly, for the PTIR approach one sample was incubated for a set incubation time at a given temperature, and many temperatures were used in parallel. For the standard isothermal approach, multiple samples were held for a series of incubation times at a single temperature. After quenching to cold temperature to arrest aggregation, aggregation rates or initial-rate coefficient (k_{obs}) values were calculated based on Eq. 3.1 or 3.2 (see Section 3.2.2 and 3.2.3 respectively). Using a single sample at each temperature for the PTIR approach provides a “worst” case example. One could easily supplement this with more than one time point or replicate samples at a given time point for each temperature. The results below indicate that this may not be necessary if one has sufficiently high-precision results with the assay of choice (e.g., SEC in the present case).

Figure 3.3 illustrates the results one obtains from the two different approaches. The standard isothermal monomer loss kinetics (panel A) and the PTIR method (panels B and C) are illustrated using 1 mg/mL IgG1 in pH 5 buffer (5 mM citrate) with 100 mM added NaCl. Figure 3.3-A shows isothermal monomer loss as a function of incubation time (t) for 325.5 K, 330 K, and 332 K. Visual inspection of Figure 3.3-A illustrates that monomer loss is linear versus t over the experimental range tested ($m = 1-0.8$). Previous work also showed linear kinetics during initial periods of aggregation [31], which is expected based on general mass action kinetic arguments when the extent of reactant consumption is small [160]. The initial-rate regime does not require one to know or assume the mechanism. However if one considers much larger extents of monomer loss ($m \ll 1$), the monomer loss profile is expected to become non-linear, and accurately quantifying the net or observed rate coefficient for monomer loss (k_{obs}) requires one to determine or assume the underlying aggregation mechanism [31, 38, 26, 53].

Figure 3.3-B illustrates aggregation rates implementing the PTIR approach. The closed symbols show m as a function of incubation temperature for 2-hour (circles, diamonds, triangles and red squares,) and 24-hour (blue squares) incubation times.

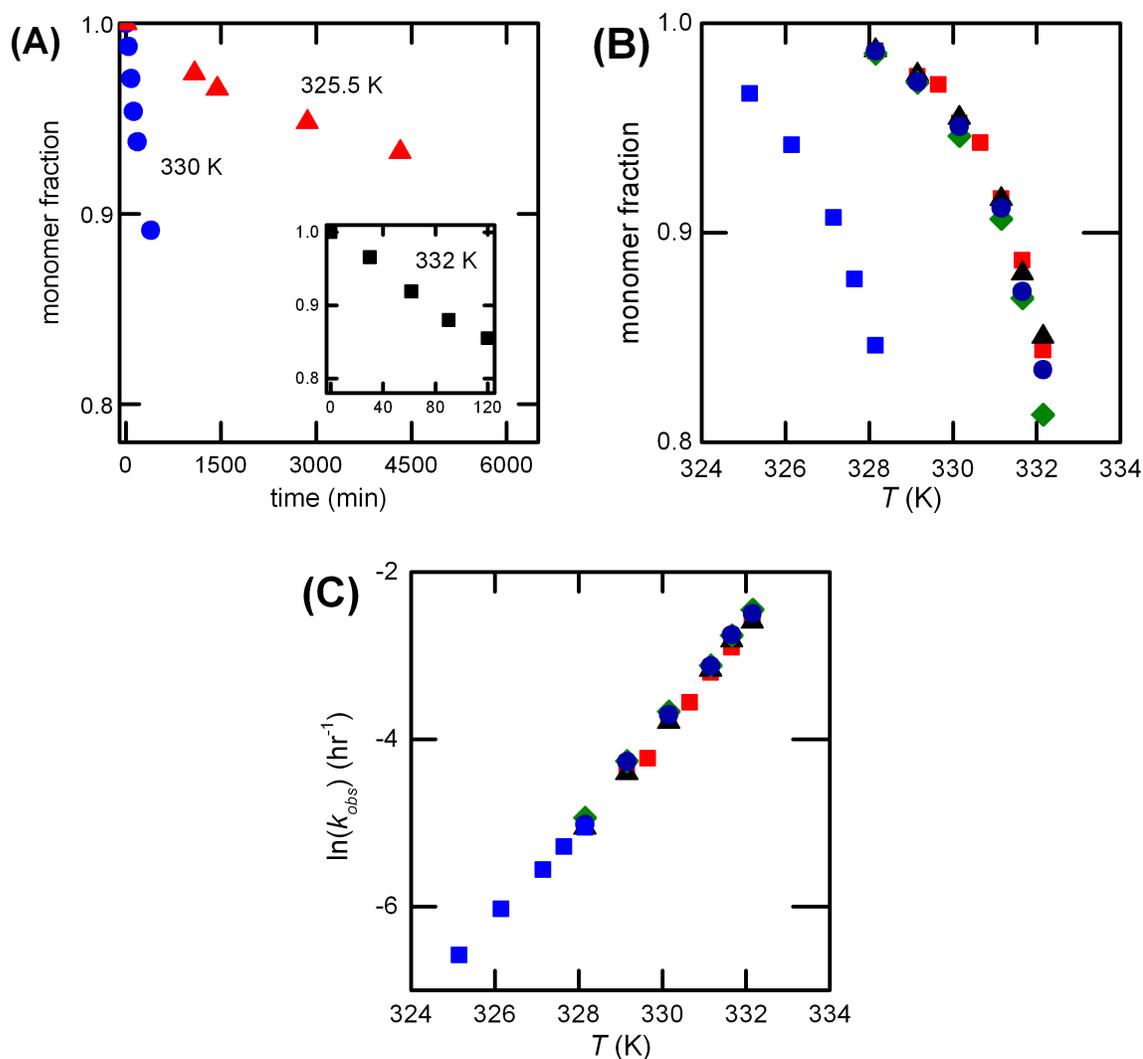


Figure 3.3: Illustrative isothermal and PTIR results for determining IgG1 aggregation rates at pH 5, (5 mM citrate buffer), and with the addition of 100 mM NaCl. (A) Isothermal monomer loss versus time at 325.5 K (triangles), 330 K (circles), and 332 K (squares shown in the inset). (B) PTIR monomer loss at 2 hours (triangles, circles, diamonds, and red squares) or 24 hours (blue squares) as a function of incubation temperature. Multiple 2-hr PTIR data sets were repeated to illustrate experiment-to-experiment variability. (C) $\ln(k_{obs})$ from PTIR data and Eq. 3.2. Symbol correspond to the closed symbol points at the same temperature in panel B. Error bars are smaller than the size of the symbols unless visible in either panel.

For the 2-hour experiments, three separate protein stocks were prepared and the experiment was repeated on separate days to provide a simple assessment of variability. Scatter in the data in Figure 3.3-B-C illustrates typical error expected from PTIR approach. Additional details are included in Supporting Information. The value of m for each symbol in Figure 3.3-B was converted to k_{obs} using Eq. 23.2, with the corresponding values of $\ln(k_{obs})$ given in Figure 3.3-C. As the PTIR approach is valid in the initial rate regime ($m = 1$ to approx. 0.8), aggregation rates measured at longer times necessarily correspond to incubations at lower temperature(s). The 24-hour time-scale experiments were chosen to extend the range of accessible k_{obs} values by at least an order of magnitude. 10-day incubations were also performed (data not shown in Figure 3.3), but in many cases significant fragmentation occurred, and this convolutes the interpretation and analysis to properly determine monomer loss rates for monoclonal antibodies (see below) [107, 164].

The reproducibility of the PTIR device and method were accessed using three independent AS-IgG1 stock solutions created in a protein concentration of 1 mg/mL at pH 5 100 mM NaCl and 5 mM citrate buffer. PTIR was performed using a constant temperature water bath with heating and cooling temperature control. The water bath was allowed to equilibrate and the final temperature (T) was calibrated ($T \pm 0.2$ °C). Additionally, PTIR was performed for three data sets using a water bath and one data set using a custom-built PTIR device. The fitted values and 95 percent confidence intervals are reported in Table 3.2. The reported values in Table 3.2 show good quantitative agreement between data sets and reported values have relatively small 95 percent confidence intervals, which support reproducibility using PTIR to determine aggregation rates.

The results in Figure 3.4 compare the values of k_{obs} versus inverse temperature (i.e., an Arrhenius diagram) for the PTIR and isothermal methods, and illustrate that the PTIR approach allows one to efficiently and accurately measure temperature-dependent aggregation rates across a range of temperatures. Each data set corresponds to a different solution condition with 5 mM sodium citrate buffer: pH 5, no added NaCl

Table 3.2: PTIR device and method reproducibility

Heating device	k_0 [hr ⁻¹]	E_a [kcal/mol]
Water Bath-1	3 ± 1	130 ± 10
Water Bath-2	2.9 ± 0.6	134 ± 5
Water Bath-2	2.9 ± 0.6	133 ± 6
PTIR device	3 ± 1	135 ± 9
All Data	2.9 ± 0.5	134 ± 5

(circles); pH 5, 100 mM added NaCl (squares); pH 6, no added NaCl (triangles); and pH 6, 100 mM added NaCl (diamonds). Open symbols correspond to aggregation rates determined from traditional isothermal incubation such as in Figure 3.4-A, while closed symbols are for the PTIR method as illustrated in Figure 3.3.

The aggregation rates from the PTIR approach are comparable in accuracy to those from the standard isothermal aggregation method that uses many samples at the same temperature, but the PTIR method provides rates for many more temperatures, with comparable consumption of protein material and user time. As noted above, the current results are a “worst case” example, in that only a single time point was used for a given temperature in the PTIR method. The method could easily be extended to use a small number of time points at a given temperature to provide even more robust values of k_{obs} vs. T .

An important use for values of k_{obs} vs. T is to determine an accurate value of the effective activation energy (E_a) of aggregation, so that accelerated aggregation rates may be more effectively extrapolated to lower temperatures e.g., for predicting room temperature shelf life [165, 105]. Intuitively, having k_{obs} values at more T values will allow one to regress E_a values with much better statistical confidence intervals and will provide greater utility for extrapolation of k_{obs} to lower temperatures. The data were regressed using the Arrhenius equation (Eq. 3.3), where k_{obs} is the experimentally

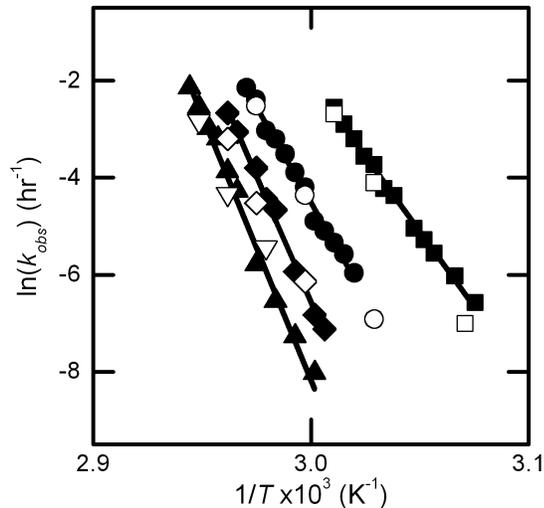


Figure 3.4: Arrhenius plot: natural logarithm of the aggregation rate plotted as a function of inverse temperature. Accelerated aggregation rates were determined using PTIR method for IgG1 at pH 5 and 0 mM NaCl concentration (closed circles), pH 5 and 100mM NaCl concentration (closed squares), pH 6 and 0 mM NaCl concentration (closed triangles), and pH 6 and 100 mM NaCl concentration (closed diamonds). Open symbols correspond to aggregation rates determined using traditional isothermal incubations.

determined value for reaction rate coefficient (units of inverse time), E_a is defined above, k_0 is value of k_{obs} at an arbitrarily chosen temperature, T_0 . In each case below, k_0 was a fitting parameter and T_0 was selected as 333.15 K because that is near the median of all incubations temperatures. Choosing different values for T_0 shifted the fitted value and confidence interval for k_0 , but not for E_a .

$$k_{obs} = k_0 \exp(E_a/R(1/T - 1/T_0)) \quad (3.3)$$

As anticipated above, the 95% confidence intervals for the E_a values from the PTIR data are much smaller than those from a traditional approach. Notably, current guidelines from regulatory agencies for accelerated stability tests for pharmaceutical

products require even fewer than three incubation temperatures [31, 105, 38]. Even with three temperature values, the fitted E_a values from the canonical isothermal approach are statistically insignificant and essentially worthless for extrapolating rates to lower temperatures. In contrast, the E_a values from the PTIR approach provide much improved precision and confidence intervals (see also, below). For the examples shown here, both the PTIR methods and the traditional method utilize comparable amounts of protein material and user time.

Although not shown in this chapter, in principle this PTIR approach can be extended to longer incubation times (multiple days to months) to yield results that may be predictive of rates at even lower temperatures. In the present case, aggregation rates were measured over reasonably small temperature windows (i.e., net change in rates on the order of 10^2) and therefore an Arrhenius equation is expected to be valid [31, 105, 38]. However, when extrapolating rates over a broader range of time scales, non-Arrhenius behavior may become significant for non-native aggregation, as discussed elsewhere [38]. Therefore, the present approach is expected to provide quantitative rate data for a range of temperatures but should be adjusted to lower temperatures if one seeks to minimize such non-Arrhenius behavior in the case of protein aggregation [105]. If one were to apply this approach to chemical degradation (e.g. deamidation [166, 167, 168]), one would more likely expect Arrhenius behavior over a broad range of temperatures.

Additionally, IgG1 and other IgG molecules are prone to fragmentation, and fragmentation depends strongly on pH and may be catalyzed by buffer and/or impurities [155]. Fragmentation will result in a loss of monomer and it is not always clear whether fragments contribute to aggregation [107]. The current example of the PTIR method focuses on the rate of monomer loss, and therefore fragmentation convolutes the analysis. As such, no rates are reported here for conditions where fragmentation occurs to a significant degree compared to monomer loss, as determined by analysis of the growth of fragment peaks in the SEC profiles (data not shown).

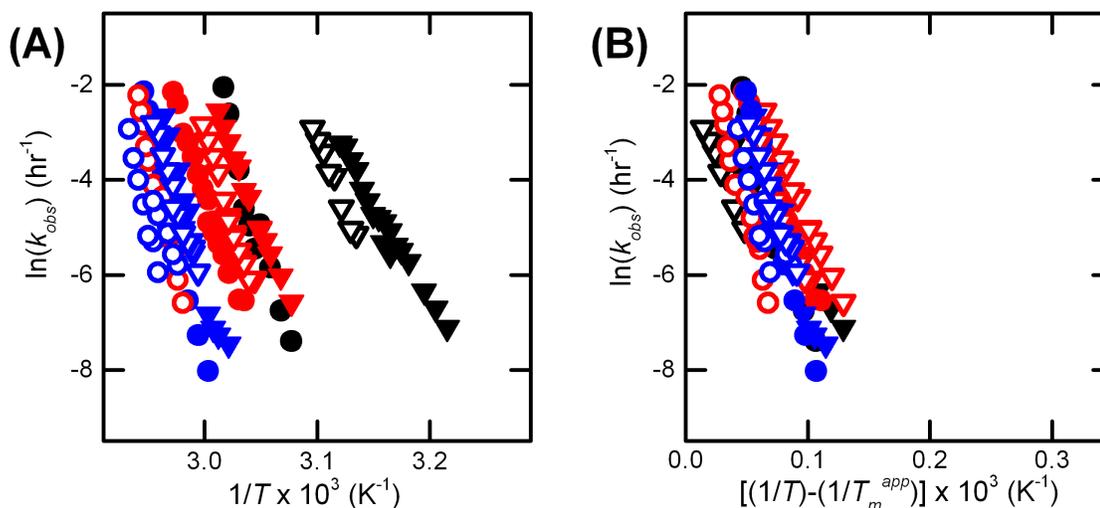


Figure 3.5: . (A) Arrhenius plot using PTIR for 2 hour and 24 hour incubations. (B) Arrhenius plot rescaling incubation temperature by the DSC peak temperatures for a given formulation. See main text for symbol notation.

Finally, many non-IgG proteins that are of interest from a biotechnology perspective do not exhibit fragmentation under typical product conditions. Smaller proteins are typically less conformationally stable than IgGs at a given temperature, and may be expected to aggregate more quickly than the present example protein. With this in mind, the PTIR approach could be useful to quantify and predict aggregation rates at much lower temperatures than what was done here.

3.5 Effects of pH, buffer, and NaCl on Temperature-Dependent Rates

$k_{obs}(T)$ was determined as a function of pH (4, 5, 6), added NaCl concentration (0 mM or 100 mM), and buffer species (citrate or acetate). Figure 3.5-A illustrates an Arrhenius diagram based on the PTIR method for all solution conditions that were tested. Symbols shown in Figure 3.5-A-B correspond to pH 4 (black), pH 5 (red), pH 6 (blue), 0 mM NaCl (circles), 100 mM NaCl (triangles). The two different buffer conditions are distinguished as: 10 mM acetate buffer (open symbols); 5 mM citrate buffer (closed symbols). Chapter one qualitatively showed that changing buffer species

could significantly alter aggregation rates. Additionally, IgG1 formulated at pH 4 in 10 mM acetate buffer with no added salt resulted in no aggregation even after heating at 85 °C for one hour.

From visual inspection of Figure 3.5-A, it is clear that no single incubation temperature would be practical to achieve aggregation rates on a comparable timescale (hours to weeks) for all solution conditions. For example, if one selected a temperature to achieve a rate corresponding to $\ln(k_{obs}) = -7$ for solution conditions indicated with the closed black circles, then the rates for solution conditions depicted with closed black triangles would be so large as to be impractical to measure (and vertically far off-scale in Figure 3.5-A). This highlights another utility of the PTIR approach. It allows one to obtain $k_{obs}(T)$ profiles for head-to-head comparison between solution conditions or different proteins that would otherwise be untenable to quantitatively compare if one had to select a common temperature for measuring aggregation rates.

Qualitatively, the shifts in $k_{obs}(T)$ with changing solution conditions correlate well with the trends for T_m values of the Fab domain from DSC in Table 3.1. Figure 3.5-B rescales the Arrhenius diagram based on the Fab peak temperatures, and this collapses the k_{obs} values onto more of a common profile. This illustrates the importance of conformational stability in determining aggregation rates, in that k_{obs} is strongly influenced by how close the incubation temperature is to the T_m for unfolding of the domain involved in exposing aggregation-prone sequences of the protein [38, 1]. However, differences in conformational stability cannot explain all of the trends in aggregation rates, as there are significant differences in the slopes (i.e., E_a values) over the different solution conditions.

Figure 3.6 shows E_a values and 95% confidence intervals determined from fitting 2-hr PTIR data to Eq. (3.3) for formulation conditions prepared in 10 mM acetate buffer (panel A) and 5 mM citrate buffer (panel B). Formulations prepared without added NaCl are shown in red; those with 100 mM NaCl concentration are shown in blue. Notably, pH 4 with no added NaCl and 10 mM acetate had an unmeasurably large E_a value, as IgG1 heated in this formulation did not aggregate (as mentioned

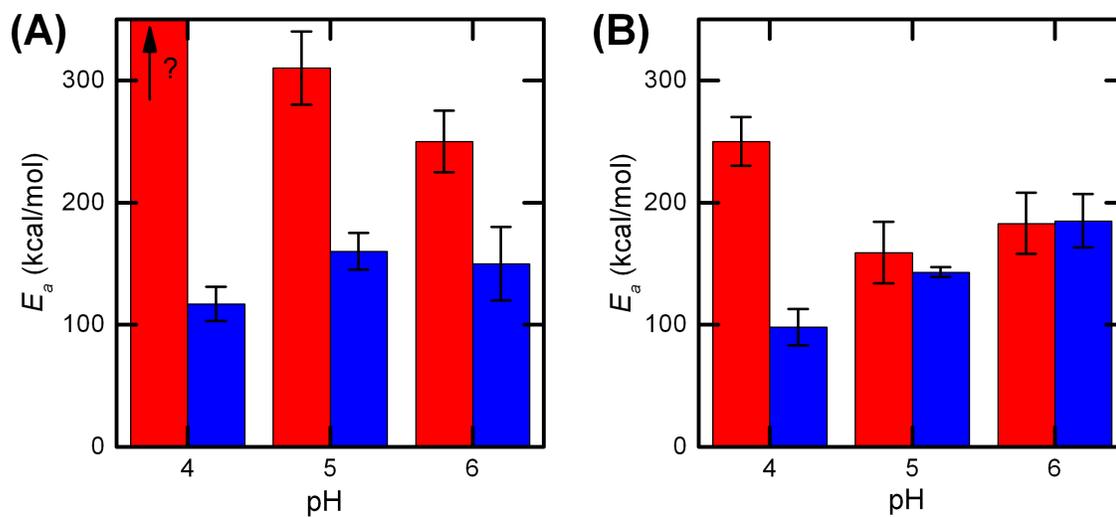


Figure 3.6: IgG1 effective aggregation activation energy determined from PTIR data in Figure 3.5 for (A) 10 mM acetate (B) 5 mM citrate. Red bars are shown for no added NaCl and blue bars are shown for 100 mM NaCl concentration. The question mark on the red bar for pH 4 in panel A indicates that E_a is unknown for that condition because aggregation was too slow to measure. Error bars are shown for 95% confidence intervals.

above).

At 100 mM added NaCl, E_a increases with increasing pH, which is expected as conformational stability (i.e. T_m) increases with pH. Based on thermodynamic arguments and qualitative mechanistic arguments for non-native aggregation [38, 169, 170], larger T_m values imply increased unfolding enthalpy values, and therefore higher E_a values. However, E_a values at low ionic strength conditions (no added NaCl) for acetate buffer show the opposite behavior. That is, T_m values decrease as one decreases pH, but E_a values increase substantially; in the extreme, at the lowest pH value tested there is no aggregation over multiple hours at temperatures close to boiling (i.e., effectively infinite E_a in the present context). This is despite the fact that the Fab T_m value (and all T_m values in Figure 3.2 for pH 4) are much lower than the selected incubation temperatures. As discussed in Section 2.9, large electrostatic repulsions between proteins under these pH conditions presumably helps prevent monomers from coming into contact.

Previous work has highlighted increased conformational stability as the pH increases towards the pI of a protein [44]. Often this results in slower aggregation rates at a given temperature, because reduced electrostatic repulsions between proteins are more than offset by changes in the conformational stability of the protein molecules [44, 132]. Alternatively, it has been argued that electrostatic repulsions explain differences in aggregation rates [48, 47]. However, the resulting activation barriers that are inferred from colloidal models are unphysically large. They also do not account for the importance of protein conformational changes that are needed to explain the stability of the resulting aggregates and the long time scales involved in nucleating such aggregates.

The present results highlight that both conformational stability and inter-protein interactions can play a discernable role in determining aggregation rates (monomer loss). The effects of conformational stability are evident under essentially all conditions, while those for electrostatic repulsions are most prevalent under conditions of low ionic strength and high net charge on the protein. While these conditions are

not typical of in vivo conditions for most proteins, they are potentially relevant for proteins under manufacturing conditions for biotechnology products [171]. The PTIR approach can be adapted to monitor or predict protein stability at 40 °C, which is becoming increasingly important, as the product shelf life is partially determined at this temperature [32].

3.6 Summary and Conclusions

A Parallel Temperatures Initial Rates (PTIR) device and method were introduced and validated against the traditional approach. The PTIR approach was applied to determine IgG1 aggregation rates across a range of solution conditions that covered more than three orders of magnitude for the initial-rate coefficient (k_{obs}) of monomer loss. The results for IgG1 aggregation highlight the importance of conformational stability, as well as electrostatic protein-protein interactions, for mediating k_{obs} and its effective activation energy. PTIR allows one to efficiently and accurately determine the temperature dependence of degradation kinetics when the initial-rates regime is relevant, and is straightforward to extend to temperature-sensitive reaction processes other than protein aggregation.

Chapter 4

OSMOLYTE EFFECTS ON MONOCLONAL ANTIBODY STABILITY AND CONCENTRATION-DEPENDENT PROTEIN INTERACTIONS WITH WATER AND COMMON OSMOLYTES

4.1 Introduction

Parts of this chapter have been included in a manuscript submitted to a peer-reviewed publication [172]. As discussed in Section 1.6.3, interactions between proteins, water, and osmolytes mediate changes in protein stability and phase behavior [67, 68, 69, 70]. Inside cells, the concentration of proteins and other macromolecules can reach volume fractions of 30-40 percent, and non-ideal interactions between proteins, water, and osmolytes are expected to be the norm rather than the exception [173, 174]. Often, the addition of sugars, polymers, and other osmolytes to protein solutions alters the protein chemical potential of the native and unfolded states [69, 175, 3]. This also causes changes in folding/unfolding equilibria when a given osmolyte alters the chemical potential of the native state to a different extent than for the unfolded state [76, 77].

Briefly, differences between competing protein-water and protein-osmolyte interactions gives rise to preferential interactions and leads to changes in protein chemical potential and solution behavior [176]. For a canonical case where the net protein-osmolyte interactions are less favorable (or more unfavorable) than water-protein interactions, the osmolyte is said to be preferentially excluded from the protein, or the protein is said to be preferentially hydrated [84]. As a result, the chemical potential of the protein necessarily increases upon increasing osmolyte concentration (c_3) at fixed pressure and protein concentration (c_2). The opposite occurs if the net protein-osmolyte interactions are more favorable than those with water, and the osmolyte instead is preferentially accumulated or “weakly bound” near the protein surface [87].

Kirkwood and Buff (KB), and others such as Hall, developed statistical mechanical theory that relates molecular radial distribution functions and spatial fluctuations in number density, to thermodynamics of multi-component solutions without the need to assume an underlying model about molecular shape or interactions [177, 178]. As shown previously by Ben-Naim and others, the thermodynamic relation given by Equations 4.1-4.3 establishes a means to directly relate preferential interactions and changes in protein chemical potential via the KB integrals [176].

$$\frac{1}{RT} \left(\frac{\partial \mu_3}{\partial c_3} \right)_{(T,P)} = \frac{1}{c_3(1 - c_3(G_{13} + G_{33}))} \quad (4.1)$$

$$\frac{1}{RT} \left(\frac{\partial \mu_2}{\partial c_3} \right)_{(T,P)} = \frac{(G_{12} - G_{23})}{(1 - c_3(G_{13} + G_{33}))} \quad (4.2)$$

$$\left(\frac{\partial \mu_2}{\partial c_3} \right)_{(T,P)} = (G_{12} - G_{23})c_3 \left(\frac{\partial \mu_3}{\partial c_3} \right)_{(T,P)} \quad (4.3)$$

In terms of statistical mechanics, one can express the KB integral for interactions between components i and j via Equation 4.4, where $g_{ij}(r)$ is the ensemble-averaged radial distribution function for component i with respect to component j in an open ensemble where all of the degrees of freedom of all components in the mixture have been Boltzmann averaged; r is the center-to-center distance between an i molecule and j molecule, and the factor of $4\pi r^2$ accounts for integration over a differential annulus.

$$G_{ij} = 4\pi \int_0^\infty (g_{ij}(r) - 1)r^2 dr \quad (4.4)$$

Using the nomenclature of Scatchard, as well as the arguments by Casassa and Eisenberg, water and (neutral) osmolytes are denoted as components 1 and 3, respectively [179, 180]. Protein with any counterions needed to exactly balance its net charge is denoted as component 2 [178, 176]. Eq. 4.3 holds in the limit of infinite dilution of protein, and the limit $c_2 \rightarrow 0$ is implicit in the derivatives and for evaluation of the KB integrals [176]. μ_2 is the protein chemical potential, μ_3 is the osmolyte chemical potential, T and P denote temperature and pressure, respectively, and the

other symbols were defined above. The derivative on the right hand side of Eq 4.3 is independent of the identity of the protein for $c_2 \rightarrow 0$ and is necessarily positive for any equilibrium system, and is available in the literature for a number of osmolyte-water binary system [181].

Biophysical characterization of protein properties such as the effective charge, second osmotic virial coefficient, or structure factor from scattering techniques do not provide direct measure of protein-water and protein-osmolyte interactions. In fact, scattering techniques, such as small angle scattering structure factor and laser light scattering second osmotic virial coefficient are biased towards largest molecules in solution and therefore of protein-protein interactions are easier to elucidate compared to protein-water or protein-osmolyte interactions.

Casassa and Eisenberg provided one of the first unified discussions on elucidating thermodynamics and connections to molecular interactions for solutions of biological macromolecules using techniques such as densimetry, light scattering, analytical ultracentrifugation, and osmometry [179]. Ben-Naim subsequently developed the “inverse KB” theory to evaluate KB integrals from thermodynamic measurements of partial specific volume, isothermal compressibility, and phase behavior or activity coefficients [176, 182]. Using the inversion process to obtain all of the possible KB integrals for three component systems has been challenging for a variety of reasons [183, 184].

Preferential interaction parameters have been measured with densimetry and osmometry. Timasheff and coworkers deduced preferential interaction parameters from density and dialysis equilibrium measurements of protein solutions with different osmolytes [74]. Record and coworkers have measured closely related, but not exactly equivalent, preferential interaction parameters for proteins with osmolytes using vapor pressure osmometry [185]. Recently, a KB approach was applied with a peptide additive transfer free energy model (ATFM) to capture preferential interaction of osmolytes with different amino acid groups by combining solubility measurements with a version of inverse KB theory [2]. Assuming the ATFM assumption holds, and that the contributions scale with the solvent accessible surface area for a given peptide group,

this allows one to calculate the free energy for transferring a protein from water to a given osmolyte solution for any protein with a known three dimensional structure [76, 186, 187].

In this chapter, densimetry and a version of inverse KB theory [176] are used to determine G_{12} , G_{23} , and the net preferential interactions for a monoclonal antibody with common neutral osmolytes as a function of osmolyte concentration. The osmolytes are sucrose, trehalose, sorbitol, and polyethylene glycol (number average mol. wt., $M_n = 6,000$). Results are compared to *a priori* predictions based on preferential interaction models [2, 3]. The results highlight quantitative and qualitative limitations of existing models and standard interpretations for some cases. Lastly these are combined with calorimetry measurements to assess whether the changes in native state chemical potential or preferential interactions may be predictive of thermal stability. The results raise questions of the validity of a number of common assumptions that underlie standard interpretations and simplified expectations regarding how some of these osmolytes improve the stability of antibodies, as well as raising concerns as to the validity of ATFM approaches for estimating how protein chemical potentials change with the addition of osmolytes.

4.2 Materials and methods

4.2.1 Sample preparation

AS-IgG1 stock solution were dialyzed against distilled-deionized Milli-Q water in the same manner as described in Section 2.2.1 Osmolyte stock solutions were prepared by dissolving known masses of sucrose ($\geq 99.5\%$ (HPLC), Sigma Aldrich), D - (+)- trehalose, dihydrate ($\geq 98\%$ Fisher Scientific), D-sorbitol ($\geq 99.5\%$ (HPLC), Sigma Aldrich) or polyethylene glycol (PEG average molecular weight 6000 Da, Sigma Aldrich) with distilled-deionized Milli-Q water. Final solutions were titrated with sodium hydroxide solution or hydrochloric acid stock to pH 6. This pH was chosen as it required minimal titration, and because prior work with this protein showed that the net charge

at pH 6 was close to zero, and protein-protein interactions were minimal (cf. Section 2.6. All masses and volume were measured gravimetrically.

4.2.2 Partial specific volume via densimetry

For each osmolyte of interest, a binary osmolyte stock solution (osmolyte and water) and a ternary protein stock solution (osmolyte, water, and protein) were prepared gravimetrically at constant osmolyte molality using a calibrated analytical microbalance (Denver Instruments) with ± 0.02 mg precision. All protein concentrations were below 5 mg/mL to minimize the potential effects of protein-protein interactions at higher concentrations. Ten or more protein solutions (~ 1 -1.25 mL each) ranging from zero to the highest protein concentration were prepared from the two above-mentioned solutions created at constant molality of osmolyte. Samples were gently mixed and centrifuged at 5000 g for 5 minutes after preparation. All solutions were measured within 24 hours of preparation.

The density of each protein solution was measured using a DMA 4500 (Anton-Paar, Ashland, VA). As all solutions were in the dilute protein concentration range, the density was linear as a function of protein concentration. The Scatchard notation is used in all equations below: component 1 is water; component 2 is protein; and component 3 is osmolyte [180]. The partial specific volume of the protein (\hat{V}_2) was determined from density as a function of protein weight fraction using Equation 4.5 as previously described [188, 189].

$$\hat{V}_2 = \frac{1}{\rho_0} + \frac{d\rho^{-1}}{dw_2} \quad (4.5)$$

ρ is the solution density for a given protein concentration, and ρ_0 is the solution density as the weight fraction of protein (w_2) approaches zero. A linear regression of the reciprocal of density as a function of w_2 gives an intercept and slope, which are the first and second terms, respectively, on the right hand side of Eq. 4.5. The 95 % confidence interval for the value of \hat{V}_2 was determined from the corresponding t-value and standard error of the slope and intercept [189].

Protein-water and protein-osmolyte interactions were determined from the behavior of \hat{V}_2 as a function of osmolyte concentration. Ben-Naim was the first to provide the relevant expressions for a three-component system with infinite dilution of one component (protein in the present case) in a convenient form to deduce G_{12} and G_{23} from densimetry [176].

$$\hat{V}_2 = \kappa_T k_b T - G_{12} + (G_{12} - G_{23}) c_3 \hat{V}_3 \quad (4.6)$$

κ_T is the isothermal compressibility of the solution, k_b is the Boltzmann constant, T is absolute temperature, c_3 is the osmolyte concentration, and \hat{V}_3 is the partial specific volume of the osmolyte, which is also determined experimentally in the same manner as described above. As aqueous solutions far from the critical point have small values for the isothermal compressibility ($\sim 0.1 - 1 \text{ GPa}^{-1}$), the first term of the right hand side of Eq. 4.6 is negligible compared to the values of the other terms [190]. The change of \hat{V}_2 as a function of osmolyte concentration then provides G_{12} and G_{23} based on slope and intercept, or based on the local tangent when \hat{V}_2 changes non-linearly with c_3 .

4.2.3 Model prediction of preferential interaction via solvent accessible surface area (ASA)

Predictions for values of $(\frac{\partial \mu_2}{\partial c_3})_{(T,P)}$ and therefore $(G_{12} - G_{23})$ for sucrose, trehalose, and sorbitol were calculated using the additive tripeptide preferential interaction model developed by Auton and Bolen [2]. The transfer free energy, $\Delta G_{(tr,N)}^{(0 \rightarrow c_3)}$ was calculated using Equation 4.7 [2]. It is defined as the free energy for moving the native protein from water to osmolyte solution with concentration (c_3). n_i is number of amino acids of type i ; $\Delta g_{(i,tr)}$ is the transfer free energy for moving a given amino acid from water to osmolyte solution with concentration c_3 ; α_i , the fraction of exposed surface area for a given amino acid, is calculated from the solvent accessible surface area (ASA) for each type of amino acid side chain (sc) or backbone (bb) as described previously [76]. ASA was determined using GET AREA,[191] Surface Racer,[192] or ProtSA,[193, 194] algorithms. These algorithms calculate the exposed surface area of a given residue or atom in the protein, which can be thought of as “rolling” a spherical

probe with radius of 0.14 nm over the surface of protein. In the present case, the crystal structure of a homology model of AS-IgG1 (1HZH)[195] was used in the ASA algorithms. The probe mimics the protein surface accessibility of a water molecule. For the unfolded state, ASA calculations were performed using ProtSA.

$$\Delta G_{(tr,N)}^{(0 \rightarrow c_3)} = \sum n_i \alpha_i^{sc} \Delta g_{tr,N}^{sc} + g_{tr,N}^{bb} \sum n_i \alpha_i^{bb} \quad (4.7)$$

As described previously, the protein chemical potential, with the addition of a given osmolyte concentrations, was calculated using Eq 4.8 [2].

$$\left(\frac{\partial \mu_2}{\partial c_3} \right)_{(T,P)} = \frac{\Delta G_{(tr,N)}^{(0 \rightarrow c_3)}}{c_3} \quad (4.8)$$

Next, predictions for the difference in Kirkwood-Buff integrals for protein-water and protein-osmolyte, $(G_{12} - G_{23})^{pred}$, were calculated using $\left(\frac{\partial \mu_2}{\partial c_3} \right)_{(T,P)}$ and Eq. 4.3. Concentration dependent values for $\left(\frac{\partial \mu_3}{\partial c_3} \right)_{(T,P)}$ were determined from analytical expressions for a given activity coefficient as a function of osmolyte concentration (sorbitol and sucrose)[181]. For trehalose, this quantity was determined from available thermodynamic data (See Appendix B), which measured the water activity of trehalose-water solutions using the isopiestic method [6].

Predictions for the AS-IgG1 and PEG preferential interaction were determined using the additive atom transfer preferential interaction model developed by Record and coworkers [3]. In this model, values for $\left(\frac{\partial \mu_2}{\partial c_3} \right)_{(T,P)}$ were calculated from the ASA of each type of atom (e.g. aliphatic carbon, amide nitrogen, etc.), which were determined using the available ASA algorithms mentioned above. Additionally, predictive values for $(G_{12} - G_{23})^{pred}$ were determined using Eq. 4.3 and concentration dependent values of $\left(\frac{\partial \mu_3}{\partial c_3} \right)_{(T,P)}$, which were determined from available water activity of water-PEG solutions [4, 5]. See Appendix B for details regarding $\left(\frac{\partial \mu_3}{\partial c_3} \right)_{(T,P)}$ calculations.

To compare the predictive values of $\left(\frac{\partial \mu_3}{\partial c_3} \right)_{(T,P)}$ and $(G_{12} - G_{23})^{pred}$ to densimetry results gathered in the present work, a predictive values for the protein partial specific volume was calculated as a function of osmolyte concentration via Eq 4.9. Starting at the measured value for \hat{V}_2^0 with no osmolyte present, the value for partial specific

volume of protein at next step, \hat{V}_2^{n+1} , is calculated given the previous value for the partial specific volume, \hat{V}_2^n , and the step size for osmolyte concentration, Δc_3

$$\hat{V}_2^{n+1} = \hat{V}_2^n + c_3 \hat{V}_3 (G_{12} - G_{23}) \quad (4.9)$$

In a similar manner, the protein chemical potential was calculated by integrating $(\frac{\partial \mu_2}{\partial c_3})_{(T,P)}$, which was calculated from densimetry measurements or predictive models.

4.2.4 AS-IgG1 unfolding via differential scanning calorimetry

Capillary differential scanning calorimetry (MicroCal cap-DSC, Malvern Instruments, Malvern, UK) was performed on AS-IgG1 solutions at different osmolyte concentrations. A buffer-buffer scan and subsequent protein-buffer scans (1 mg/mL AS-IgG1) were performed at 1 °C/min from 20-95 °C. With no osmolyte present, protein scan (after first scan) showed no reversibility (data not shown) and was not expected for AS-IgG1 based on past work [1, 36]. After subtracting the buffer scan, the absolute heat capacity, C_p was calculated from the DSC thermograms using the standard expressions [53, 196]. The mid-point unfolding temperature (T_m) and Van't Hoff enthalpy, ΔH , of the AS-IgG1 unfolding transitions were determined using the Peak Analyzer function in Origin Pro (Origin Lab Corporation, Northampton, MA). For a given osmolyte, the change in the protein chemical potential from native to unfolded state with respect to a given osmolyte concentration, $(\frac{\partial \Delta \mu_2^{(N \rightarrow U)}}{\partial c_3})_{(T,P)}$, also known as the m-value, [197, 198] was calculated using Equation 4.10 [30]. T_m^o and ΔH^o denotes the mid-point unfolding temperature and Van't Hoff enthalpy, respectively, of the reference state, which would be the case of AS-IgG1 unfolding with no osmolyte present.

$$\left(\frac{\partial \Delta \mu_2^{(N \rightarrow U)}}{\partial c_3} \right)_{(T,P)} = \frac{\Delta H^o}{T_m^o} \left(\frac{dT_m}{dc_3} \right) \quad (4.10)$$

4.3 Partial specific volume via density measurements

The protein and osmolyte partial specific volumes were determined from a series of density measurements at constant osmolyte molality. Figure 4.1 depicts illustrative

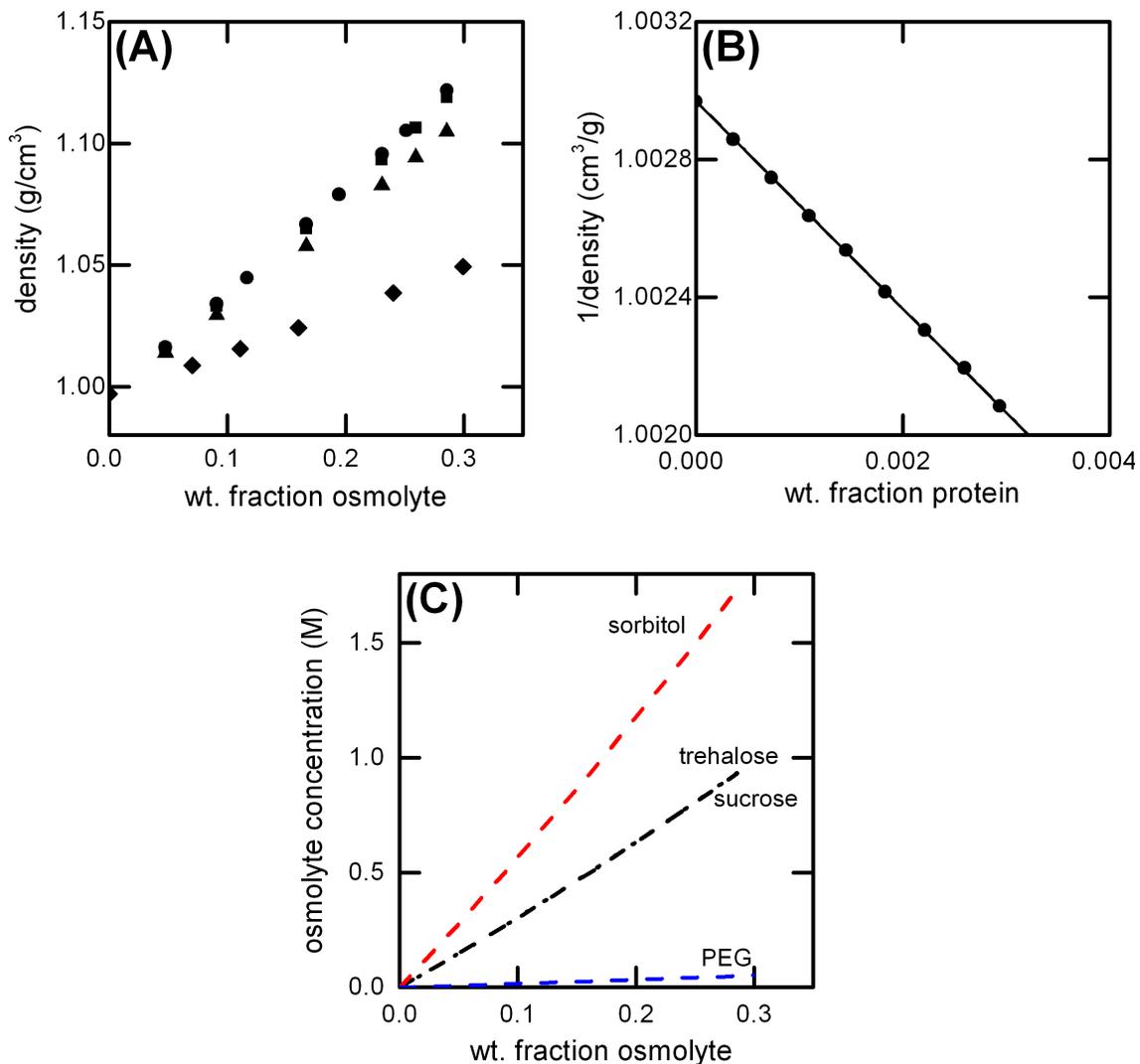


Figure 4.1: Binary solutions. (A) water-osmolyte solutions: density as a function of osmolyte weight fraction for sucrose (squares), trehalose (circles), sorbitol (triangles), and polyethylene glycol $M_n = 6,000$ g/mol (diamonds). (B) water-protein solutions: reciprocal of the density plotted as a function of AS-IgG1 weight fraction. (C) Conversion of $c_3\hat{V}_3$ to osmolyte molarity.

results for density measurements for the binary system of water and a given osmolyte. Figure 4.1-A shows the solution density as a function of osmolyte weight fraction for sucrose (squares), trehalose (circles), sorbitol (triangles), and polyethylene glycol (diamonds) with an average $M_n = 6,000$ g/mol (diamonds). For the binary system of AS-IgG1 and water in Figure 4.1-B, the reciprocal of the solution density is plotted as a function of weight fraction of protein. The maximum protein concentration was limited to ~ 5 mg/mL to avoid potential effects of protein-protein interactions (i.e., to assure the $c_2 \rightarrow 0$). Both panels in Figure 4.1 show linear increases in solution density as a function of weight fraction of osmolyte or protein consistent with high quality fits from regression to a linear function.

The partial specific volume of the protein in a given osmolyte is determined from a linear regression of the reciprocal of density as a function of protein weight fraction (e.g., as shown in Figure 4.1-B) and calculated using Eq. 4.5. The partial specific volume for each osmolyte (\hat{V}_3) in the absence of protein was calculated using the data in Figure 4.1-A in the same manner. Table 4.1 reports the fitted values and 95 percent confidence intervals for (\hat{V}_3) of each osmolyte in water. All data with protein present are in the dilute protein limit ($c_2 \rightarrow 0$); therefore, \hat{V}_3 is independent of protein concentration. The data in Figure 4.1-A illustrate that \hat{V}_3 is also independent of c_3 for the conditions tested here, and therefore the product $c_3\hat{V}_3$ in Eq. 4.6 is a linear function of c_3 . Figure 4.1-C shows this relationship graphically for each osmolyte in Table 4.1.

4.4 Partial specific volume of AS-IgG1 in neutral osmolytes

Figure 4.2 plots the values for the partial specific volume of protein as a function of $c_3\hat{V}_3$ for AS-IgG1 with sucrose (A), trehalose (B), sorbitol (C), and PEG (D). Each data point and 95 percent confidence interval were calculated from linear regression of a set of density measurements on solutions created at constant osmolyte molality. The solid curves in Fig. 4.2 correspond to best fits of \hat{V}_2 vs. $c_3\hat{V}_3$, and dashed curves are *a priori* predictions from available preferential interaction models (cf. Methods).

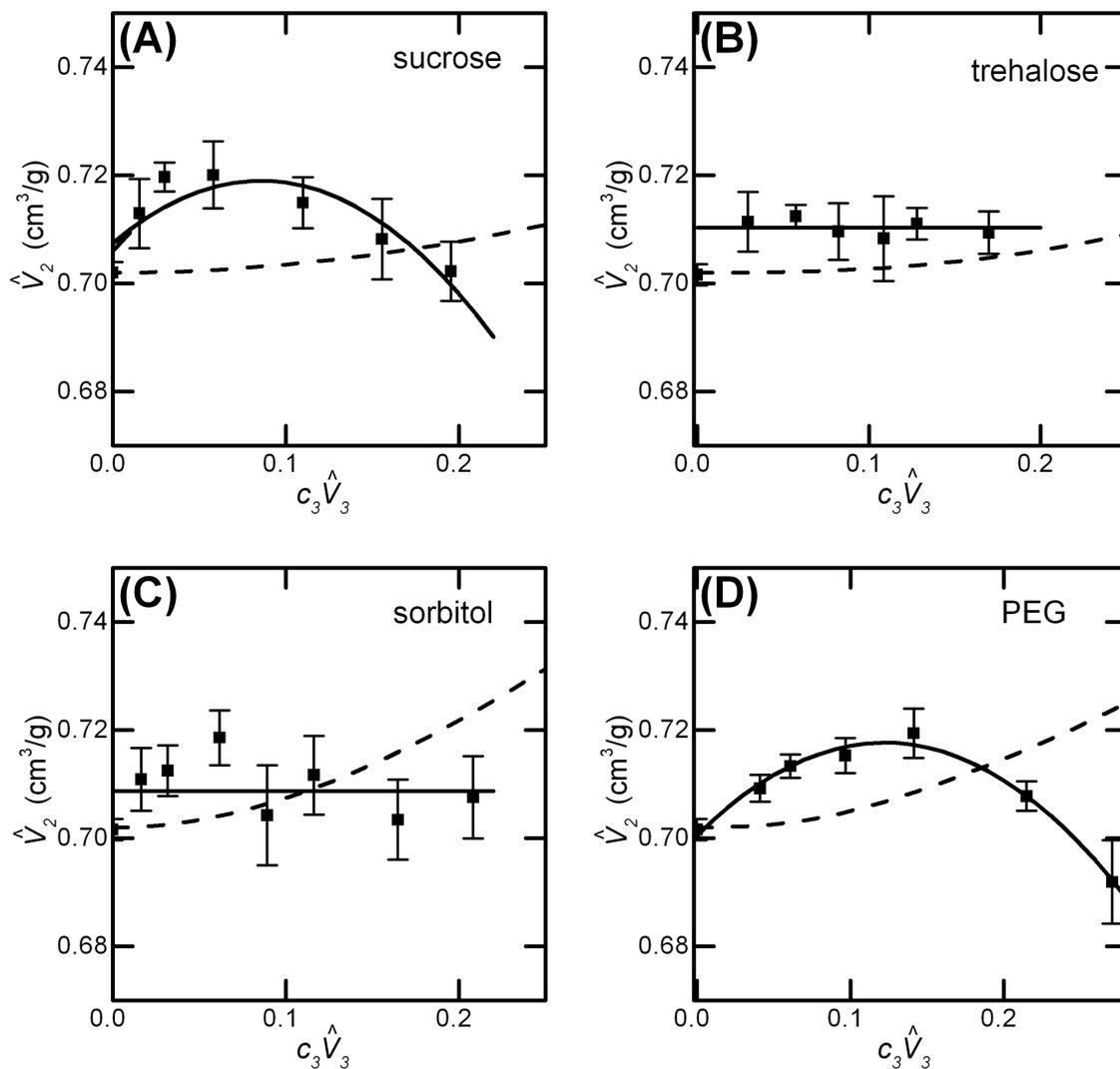


Figure 4.2: Ternary solutions. AS-IgG1 partial specific volume as a function of $c_3\hat{V}_3$ for (A) sucrose, (B) trehalose, (C) sorbitol, and (D) PEG. The dashed black line is the prediction from the preferential interaction models [2, 3], using a homology model of AS-IgG1 based on the human IgG1 crystal structure (PDB 1HZH). Error bars correspond to 95 percent confidence intervals. Solid curves are empirical fits to interpolate the experimental values.

Table 4.1: Osmolyte partial specific volume. Error bars correspond to 95 percent confidence intervals

Osmolyte	\hat{V}_3 [mL/g]
Sucrose	0.6206 ± 0.0007
Trehalose	0.6127 ± 0.0014
Sorbitol	0.660 ± 0.002
PEG ($M_n = 6k$)	0.8362 ± 0.0010

Inspection of Figure 4.2-A shows values of \hat{V}_2 increase initially at low sucrose concentrations. However, as the concentration of sucrose increases further, the value of \hat{V}_2 reaches a maximum then begins to decrease. Based on Eq. 4.6, this shows that $(G_{12} - G_{23})$ is greater than one at low c_3 , but decreases with increasing c_3 and eventually changes sign at c_3 greater than approximately 10 w/w % sucrose. This indicates that sucrose is preferentially excluded from the protein at lower sucrose concentrations but becomes preferentially accumulated near the protein at higher osmolyte concentrations. The dashed curve represents an *a priori* prediction of preferential interactions using available models[2, 3] and shows a shallow upward curvature in the values for \hat{V}_2 with increasing osmolyte concentration. As such, the model under predicts the preferential exclusion of sucrose at low osmolyte concentration, and is qualitatively incorrect at higher osmolyte concentrations.

Figures 4.2-B and 4.2-C illustrate the behavior of the partial specific volume of AS-IgG1 as a function of $c_3 \hat{V}_3$ for trehalose and sorbitol, respectively. There does not appear to be a significant trend for values of \hat{V}_2 over these osmolytes concentrations, indicating that magnitude of preferential interactions for this IgG and either of these osmolytes is too small to be directly observed via this experimental approach. Trehalose does not show a similar or greater degree of preferential exclusion compared to sucrose, contrary to what has been concluded for some proteins [78]. At higher

trehalose concentrations, no preferential accumulation was observed, within the sensitivity of the measurement. However, preferential interaction models (dashed curves in Fig. 4.2-B-C) predict significant upward curvature for AS-IgG1-sorbitol solutions, and qualitatively similar results for trehalose. Figure 4.2-D plots \hat{V}_2 as a function of $c_3\hat{V}_3$ for AS-IgG1 with PEG ($M_n = 6,000$ g/mol). Initially, values of \hat{V}_2 increase before reaching a maximum at value of $c_3\hat{V}_3$ of approximately 0.15. However, at higher PEG concentrations, values of \hat{V}_2 decrease. Interestingly, the partial specific volume of AS-IgG1 as a function of PEG concentration behaves similarly to sucrose; preferential exclusion of PEG at low osmolyte concentrations followed by preferential accumulation at higher osmolyte concentrations. Model predictions (dashed lines in Fig. 4.2-D) are qualitatively opposite to the experimental behavior as one increases PEG concentrations to levels that are similar to those used in common practice[85].

4.5 Kirkwood-Buff integrals for protein-water and protein-osmolyte interactions

Values of the protein-water KB integral (G_{12}) and the protein-osmolyte KB integral (G_{23}) were determined from linear regression of the data in Fig. 4.2 for the cases where values of \hat{V}_2 changed with osmolyte concentration (i.e., sucrose and PEG). Data for \hat{V}_2 as a function of $c_3\hat{V}_3$ were fit to a quadratic equation (See Appendix for detailed statistics) and the fitted coefficients were used to calculate the KB integrals. Briefly, the tangent of \hat{V}_2 as a function of $c_3\hat{V}_3$ is equal to $(G_{12} - G_{23})$, and the y-intercept of the tangent line is equal to $-G_{12}$ (cf., Eq 4.6) [176]. Therefore, if G_{12} and G_{23} are independent of osmolyte concentration then one expects \hat{V}_2 to be linear with respect to $c_3\hat{V}_3$. The slope is positive ($G_{12} \geq G_{23}$) for preferential exclusion of osmolytes, or vice-versa for preferential accumulation of the osmolyte near the protein.

The units for G_{ij} are volume (or volume per molecule) when one uses concentration units of molecules/volume. G_{ij} are reported here in terms of more experimentally conventional units (volume/mass), which can be converted to those in Eq. 4.4 by use of the protein molecular weight and Avogadro's number. Positive (negative) values

for KB integrals imply a net attractive (repulsive) interaction for the two components relative to an ideal gas, when averaged over all center-to-center distances between a molecule of component i and that of component j . As noted above, significant preferential interactions exist for the protein when the difference in KB integrals (i.e. $G_{12} - G_{23}$) is non-negligible.

Figure 4.3 shows G_{12} (dashed curves) and G_{23} (solid curves) for sucrose (panel A), sorbitol and trehalose (panel B), and PEG (panel C). Inspection of Figure 4.3 shows all values for Kirkwood-Buff integrals are negative for both protein-water and protein-osmolyte. This follows from the large steric contribution to the KB integrals, as the osmolyte or water molecules cannot overlap with the protein. From Eq. 4.4, configurations that correspond to overlapping molecules cause negative values in the integrand. Therefore, purely steric interactions will result in negative values of any G_{ij} , although they will not necessarily equal those from simple gas-phase or implicit solvent estimates [199].

For Figures 4.3-A and 4.3-D, one observes G_{ij} values that depend strongly on osmolyte concentration. In the absence of a given osmolyte, G_{12} is equal to \hat{V}_2 in the limit of low protein concentration, [176] and the initial value for G_{23} follows from the initial slope for \hat{V}_2 in Figure 2. For sucrose, G_{23} starts at lower values than G_{12} , as the osmolyte is initially preferentially excluded and also has a larger steric contribution to the KB integrals compared to that of water. In Fig. 4.3-A, one observes that the value of G_{12} decreases with increasing sucrose concentration, suggesting a decrease in the relative concentration of water molecules near the protein surface, while the value of G_{23} increase, suggesting an increase in the relative concentration of sucrose near the protein surface. The net result is that $G_{12} \geq G_{23}$ (i.e., preferential exclusion of sucrose) at low osmolyte concentration, and this switches to the opposite behavior above a sucrose concentration of approximately 0.3 M (~ 10 wt. percent). A qualitatively similar behavior is observed for PEG solutions (Figure 4.3-C), but with the switch-over from preferentially excluded PEG to preferentially accumulated PEG occurring at a higher PEG concentration. In contrast, Figure 4.3-B shows G_{12} (dashed line)

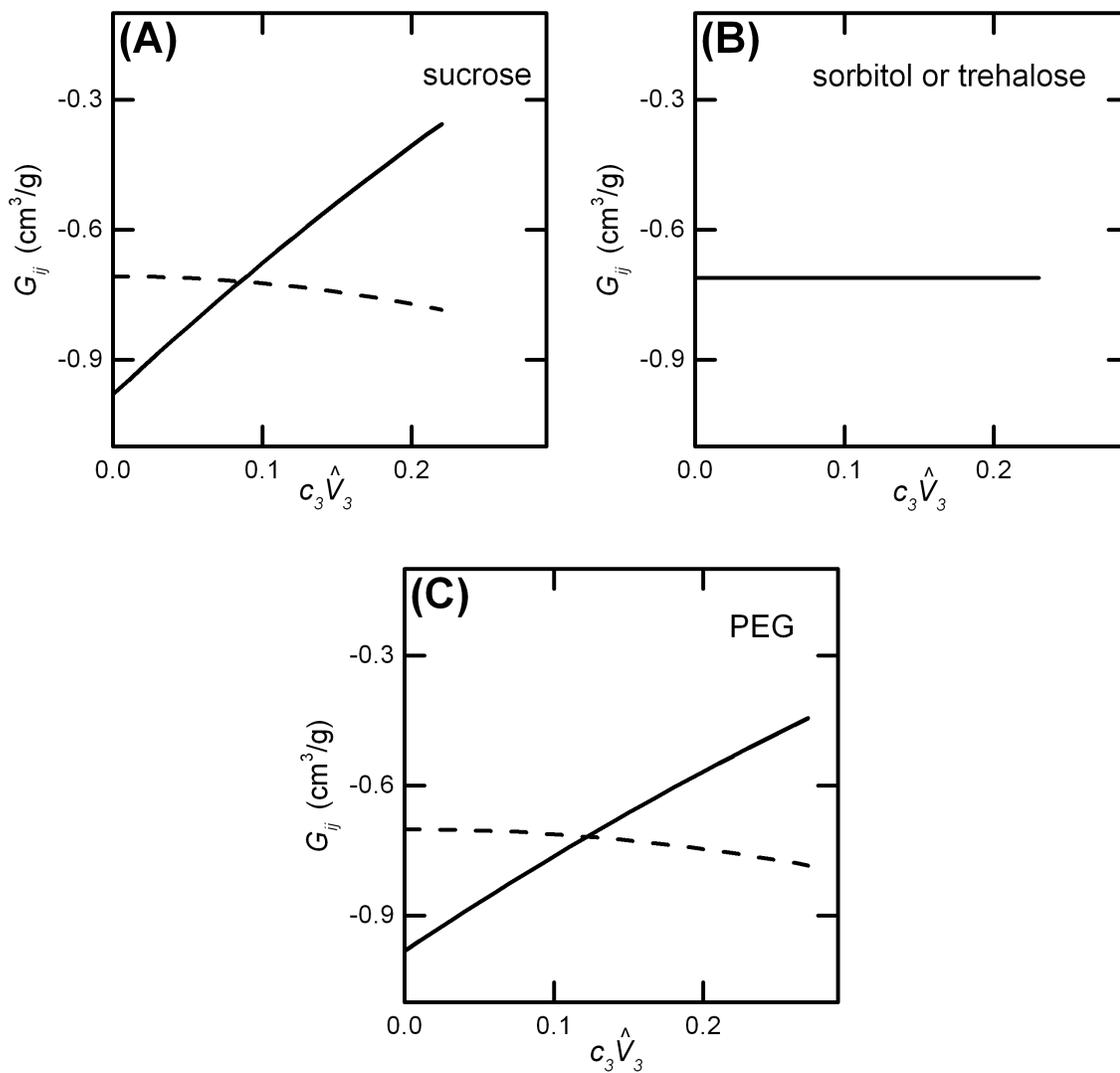


Figure 4.3: Average protein-water interactions (G_{23} , solid curves) and protein-osmolyte interactions (G_{12} , dashed curves) as a function of $c_3 \hat{V}_3$ for (A) sucrose, (B) sorbitol or trehalose, and (C) PEG.

and G_{23} (solid line) for trehalose and sorbitol osmolytes. As there were not significant changes in \hat{V}_2 , vs. c_3 in Fig. 4.2-B-C, G_{12} is approximately equal to G_{23} , implying preferential interactions were not large enough to be detected within the sensitivity of the densimetry measurements.

4.6 AS-IgG1 native state chemical potential

As noted above and shown by Eq. 4.3, thermodynamics establishes a direct relation between preferential interactions ($G_{12} - G_{23}$) and changes in chemical potential of the protein with respect to osmolyte concentration at fixed temperature and pressure for dilute protein solutions. The change in osmolyte chemical potential with respect to osmolyte concentration, $\left(\frac{\partial\mu_3}{\partial c_3}\right)_{(T,P)}$ is needed for this conversion between KB integrals and the derivative of the protein chemical potential with respect to c_3 . As mentioned above, $\left(\frac{\partial\mu_3}{\partial c_3}\right)_{(T,P)}$ does not depend on the identity of the protein because it is evaluated in the limit of $c_2 \rightarrow 0$. In the present case, values of $\left(\frac{\partial\mu_3}{\partial c_3}\right)_{(T,P)}$ as a function of osmolyte concentration were determined from literature values for the binary systems of water and a given osmolyte. However, at much higher protein concentrations, one must exercise caution as $\left(\frac{\partial\mu_3}{\partial c_3}\right)_{(T,P)}$ may not be synonymous to that of the binary system [176], as G_{11} , G_{13} , and G_{33} quantities may be indirectly affected by the protein. Using values of $(G_{12} - G_{23})$ determined from densimetry, and literature values of $\left(\frac{\partial\mu_3}{\partial c_3}\right)_{(T,P)}$, [181, 6, 4, 5] the quantity $\left(\frac{\partial\Delta\mu_2}{\partial c_3}\right)_{(T,P)}$ was determined via Eq. 4.3 for a given osmolyte concentration. Values of $\Delta\mu_{2,tr}^N$ were then determined as a function of osmolyte concentration by numerically integrating the derivative with respect to osmolyte concentration.

Figure 4.4 shows the results for AS-IgG1 $\Delta\mu_{2,tr}^N$ as a function of $c_3\hat{V}_3$ for sucrose (A), sorbitol and trehalose (B), and PEG (C), based on the results in Figure 4.3. In Figure 4.4, solid curves correspond to values determined using densimetry data, and dashed curves are based on preferential interaction model(s) used to produce the dashed curves in Figure 4.2 (cf. Methods). Inspection of Figure 4.4-A shows the value of $\Delta\mu_{2,tr}^N$

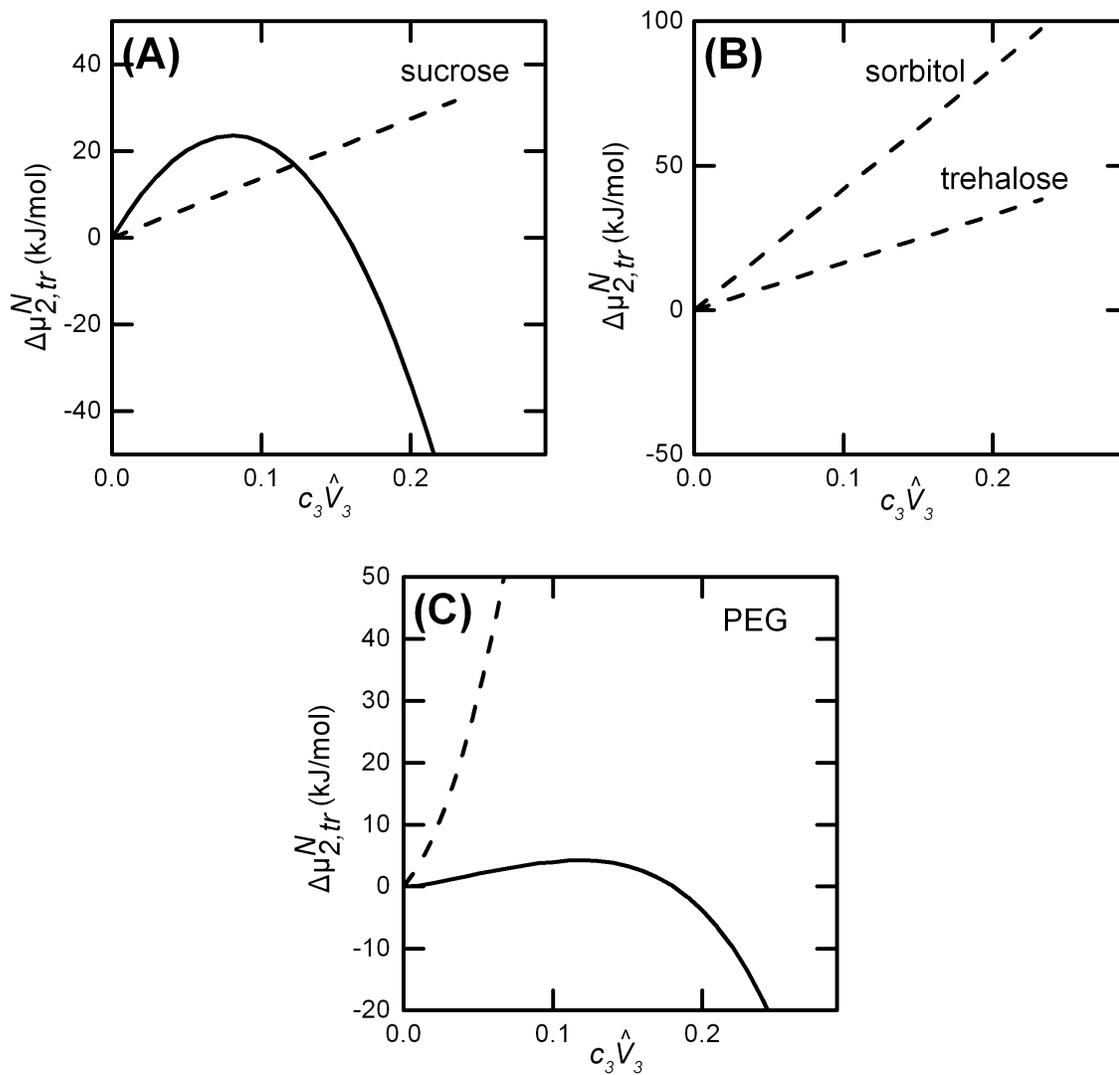


Figure 4.4: AS-IgG1 native state chemical potential as function of $c_3 \hat{V}_3$ for (A) sucrose, (B) sorbitol or trehalose, and (C) PEG. dashed curves correspond to predictions using the preferential interaction models.

increases at low sucrose concentration, reaches a maximum at approximately 20 kJ/mol at $c_3\hat{V}_3$ just less than 0.1, then decreases at much higher osmolyte concentrations.

In Figure 4.4-B, values of $\Delta\mu_{2,tr}^N$ from densimetry for sorbitol and trehalose are not plotted, as these osmolytes did not appear to show any preferential interaction within the sensitivity of the densimetry measurement. However, the preferential interaction model developed by Bolen and coworkers predicts a linear increase in chemical potential for sorbitol and trehalose [2]. The larger magnitude for sorbitol is due to the fact that the molarity is approximately double that for sucrose or trehalose at the same $c_3\hat{V}_3$ (see Fig. 4.1-C). Notably, comparing the magnitude of the predicted changes in transfer free energies for panel A and B shows that the models predict similar orders of magnitude for the changes due to sucrose, trehalose, and sorbitol. However, the densimetry data clearly shows that if there were preferential interactions for trehalose or sorbitol with AS-IgG1, then they are much less pronounced than those of sucrose. Therefore, the model predicts preferential interactions that are best a large overestimate when compared to experimental results.

From Fig 4.4-C, one observes a quadratic behavior of $\Delta\mu_{2,tr}^N$ with increasing PEG concentration and the value of $\Delta\mu_{2,tr}^N$ reaches a maximum value at $c_3\hat{V}_3$ just greater than 0.1. Notably, preferential interaction models predict an upward curvature in $\Delta\mu_{2,tr}^N$ with increasing PEG concentration [3].

Preferential exclusion (accumulation) of osmolytes leads to increases (decreases) in chemical potential of the protein. Perhaps it is not surprising that AS-IgG1 has concentration dependent preferential interactions as it is a labile molecules rather than a globular protein and also contains glycosylations. The preferential exclusion of sucrose and trehalose is thought to originate from the formation of an excluded network of sugar molecules that strengthen water-protein interaction near the surface of the protein [78, 90]. However, one may expect this mechanism to be dependent on a protein's 3-D geometry and a large flexible IgG with non-globular geometry may promote different trehalose network than a small globular protein. Additionally, at very high osmolyte concentrations (and larger osmolytes), depletion interactions may be expected

[49]. In this case, simply due to steric arguments, preferential accumulation may be expected. However, it is interesting to consider if the preferential interaction and chemical potential trends will be predictive of protein stability at high temperature.

4.7 AS-IgG1 unfolding via DSC

Given the large changes in chemical potential ($\sim 1-5$ kJ/mol) observed in Fig 4.4, one would expect the biophysical properties of the IgG, such as the free energy of unfolding and solubility to be significantly altered in the presence of osmolytes. Differential scanning calorimetry was performed on AS-IgG1 solutions over similar concentration ranges of sucrose, trehalose, sorbitol, and PEG as those used in the densimetry measurements. Figures 4.5 illustrate DSC profiles with as a function of sucrose (A), trehalose (B), sorbitol (C), and PEG (D) concentration. DSC profiles are vertically offset for easier viewing, and the arrow denotes the direction of increasing osmolyte concentration. As shown in Fig. 3.2 and elsewhere [36, 1, 161], three unfolding transition endotherms (peaks or shoulders) were commonly observed with mAbs. At low pH (e.g., pH 4) three transitions were observed, and the unfolding transitions corresponding to the following order, from lowest to highest transition temperatures: C_{H2} , (Fab), (C_{H3}). However, at pH 6, the unfolding transitions for the C_{H2} and Fab were found to overlap, and therefore only two transitions were observed by visible inspection [36, 1, 161].

From inspection of Figure 4.5-A, one observes two unfolding transitions, consistent with prior work at pH 6. The larger peak at lower temperatures corresponds to overlapping unfolding transitions of the C_{H2} and Fab, while the smaller peak at higher temperatures corresponds to that of the C_{H3} . Additionally, the apparent mid-point unfolding temperature (T_m) increases with increasing trehalose concentration (see Figure 4.5-C). However, for the conditions tested, the unfolding transitions were not fully reversible and may be convoluted with aggregation. As such, the transition temperature are referred to only as apparent T_m values. For trehalose, sucrose, and sorbitol, the C_{H2} /Fab unfolding transitions occurred together even with the addition of more

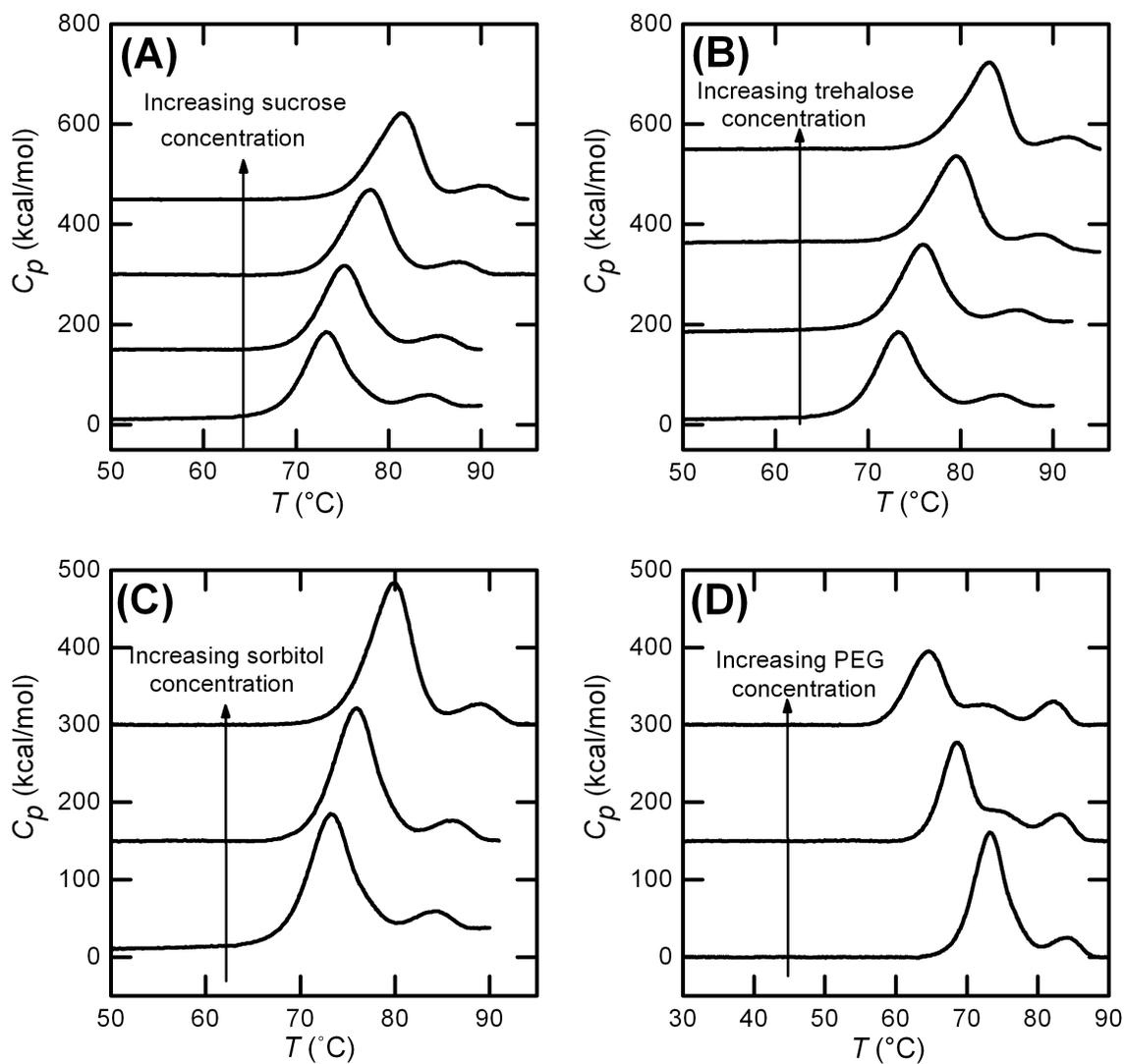


Figure 4.5: AS-IgG1 differential scanning calorimetry. DSC thermograms with increasing concentrations of (A) sucrose, (B) trehalose, (C) sorbitol, and (D) PEG. In each panel, curves are offset to help distinguish profiles.

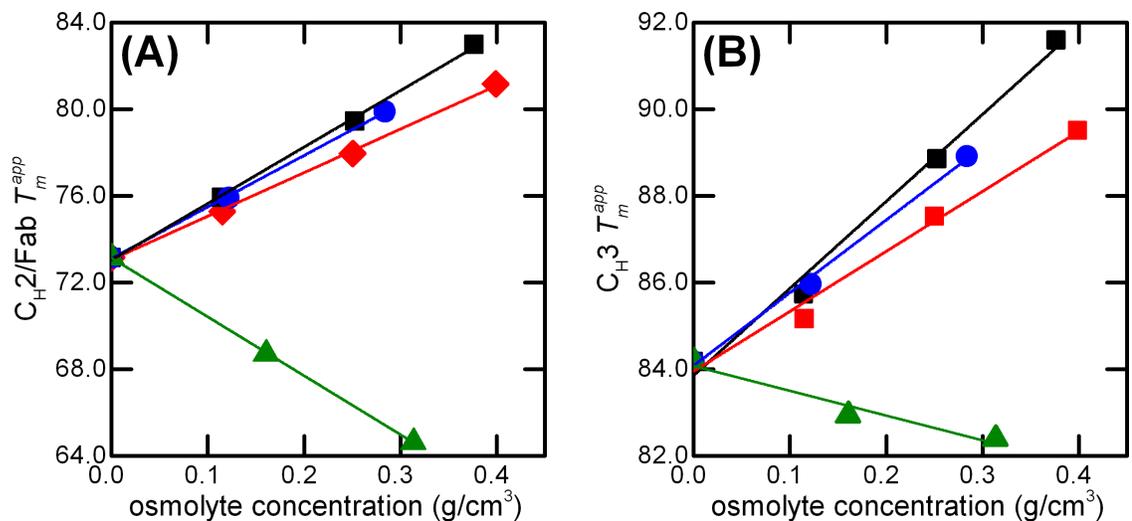


Figure 4.6: AS-IgG1 differential scanning calorimetry. (A) C_{H2}-Fab and (D) C_{H3} T_m values as a function of osmolyte concentration. Symbols correspond to trehalose (squares), sucrose (diamonds), sorbitol (circles), and PEG (triangles).

osmolyte. However, as one notes from Figure 4.5-B, as PEG concentration increases, C_{H2} and Fab peaks no longer overlap. Additionally, the Fab unfolding transition decreases more so than the C_{H2} transition, suggesting PEG preferentially interacts with the Fab more so than the C_{H2}.

T_m values for C_{H2}/Fab peaks and C_{H3} peaks were interpolated numerically (see Methods). Fig. 4.6-A (Fig. 4.6-B) plots T_m values as a function of osmolyte concentration for C_{H2}/Fab (C_{H3}) unfolding transition. For PEG, Fab T_m values are reported. In both figures, symbols correspond to sucrose (diamonds), trehalose (squares), sorbitol (circles), and PEG (triangles). Solid lines in Figure 4.6 show linear fits determined from regression for each data set. Interestingly, one notes from Fig. 4.6 that sucrose, trehalose, and sorbitol increase AS-IgG1 unfolding transitions, while PEG decreases the transitions. The m-value for the C_{H2}/Fab and C_{H3} unfolding transitions were calculated via Eq. 4.10 using the Van't Hoff enthalpy and T_m of AS-IgG1 unfolding profile with no osmolyte present, and values for $\left(\frac{dT_m}{dc_3}\right)$ from the slopes in Fig. 4.6.

4.8 Preferential interactions and their effect on protein unfolding and stability

The free energy of unfolding is determined by the difference in the protein chemical potential of the native and unfolded state. The m -value represents the difference in $\left(\frac{\partial\Delta\mu_2^{(N\rightarrow U)}}{\partial c_3}\right)_{(T,P)}$ between the native and unfolded states of the protein, and can be estimated experimentally via multiple techniques [198, 200]. By this sign convention used below, positive (negative) m -values indicate stabilization (destabilization) of the folded state relative to the unfolded one. In the present case, m -values were estimated from DSC and the analysis of Bechtel and Schellman (cf. Methods) [30]. The linear dependence of T_m values in Fig. 4.6, combined with Eq. 4.10, results in m -values are independent of osmolyte concentration.

The preferential interaction model of Bolen and coworkers also can be used to predict m -values for sucrose, trehalose, and sorbitol [2]. However, doing so requires one to assume a magnitude for the change in the ASA between the native and unfolded states of the protein, which is highly nontrivial [193]. In addition, the DSC thermograms are not at equilibrium because unfolding is convoluted with aggregation, and it is not clear whether one should consider the lower T_m transition to be simply an overlap of C_H2 and Fab transitions in the DSC signal, or a single transition that requires both domains to unfold as group.

Those quantitative limitations notwithstanding, it is possible to at least derive qualitative conclusions by comparing the results in Figures 4.2-4.4 with those in Figures 4.5-4.6. From the densimetry data is clear that for sucrose and PEG as osmolytes, the native state chemical potential for AS-IgG1 increases within increasing c_3 at low c_3 , reaches a maximum, then decreases with further increases in c_3 . For trehalose and sorbitol, it is possible that the native state chemical potential of AS-IgG1 increases with increasing c_3 , but this was not discernable experimentally. In what follows below, we assume that it was simply too weak of an effect to be detected with these experimental methods.

The DSC results clearly show that the chemical potential difference upon unfolding of the C_H3 domain of AS-IgG1 increases with increasing c_3 for sucrose, trehalose, and sorbitol; while it decreases for PEG as the osmolyte. A similar argument holds for the overlapping transition of the Fab and C_H2 domains. These results were unexpected based on the densimetry results and canonical models for protein unfolding in the presence of osmolytes [84]. That is, canonical expectations are that osmolytes such as sucrose, trehalose, and PEG are always preferentially excluded from the native state, and excluded even more so from the unfolded state [74, 75]. As a result, the chemical potential of both the folded and unfolded states are expected to increase by adding such osmolytes, but the increase is greater for the unfolded state. This has the net effect of increasing the unfolding free energy, and T_m values, upon addition of such osmolytes. If one were to consider the DSC results for sucrose or trehalose in isolation, they would normally be interpreted mechanistically in that way [74, 78, 79]. That would also commonly be done for sorbitol, although it has been found in some cases that it can change from being preferentially excluded to preferentially accumulated (e.g., as a function of temperature) [201].

The DSC results for PEG are unexpected for multiple reasons. PEG is not expected to be preferentially accumulated at the protein surface unless one deals with much smaller PEG molecules,⁸ and addition of PEG is not expected to decrease T_m values [85]. Furthermore, the results here show that PEG interacts more strongly (greater preferential accumulation) with the Fab domains of AS-IgG1 than it does with the C_H2 domain. This is evidenced by the observation that addition of PEG causes the peak for the Fab domain to become discernable from the C_H2 domain in Figure 4.5-D (top).

While the DSC results for sorbitol, sucrose, and trehalose would suggest the canonical models or interpretations hold for AS-IgG1, the densimetry results show this is not the case. Sucrose is preferentially excluded from the native state at low c_3 but it is accumulated at high c_3 . If the same were to hold for the native state, then the T_m values should increase with increasing c_3 at low c_3 , but decrease at higher c_3 . From

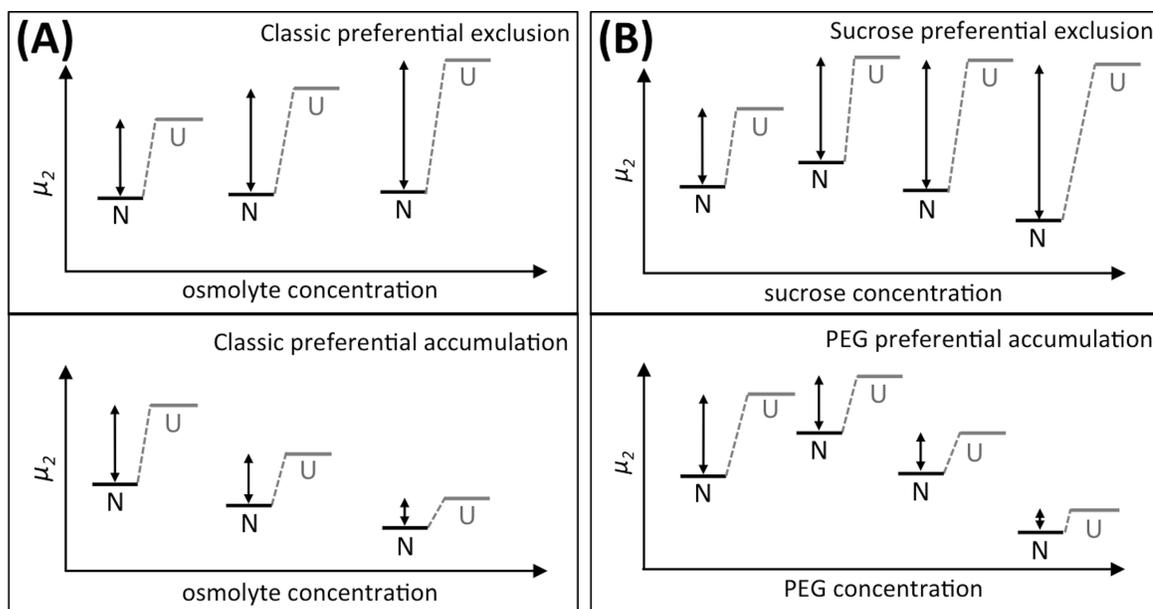


Figure 4.7: Schematic of (A) classic theory and (B) concentration dependent preferential interaction. Panels illustrate qualitative behavior of the protein chemical potential for the native (N) and unfolded (U) states as a function of osmolyte concentration

Figure 4.5, this is clearly not the case. For PEG, it is similarly complex. By similar reasoning to that for sucrose, one would expect the T_m values to first increase and then decrease with increasing PEG concentration. Instead, PEG is a destabilizer at all PEG concentrations tested here. Finally for sorbitol and trehalose, the canonical model might be correct, but the data here cannot discriminate whether this is the case.

To help summarize the key results and their interpretation graphically, Figure 4.7 illustrates the qualitative behavior expected from the theory of preferential interactions by plotting the protein chemical potential as a function of osmolyte concentration, for native (N) and unfolded (U) states. A simple two-state unfolding example is used for ease of depiction. For a multi-state protein such as an antibody, one could generalize to have unfolding intermediates in addition the N and U states.

Figure 4.7-A (top) shows the canonical example described above for preferential exclusion. The opposite behavior is shown in the bottom panel, and is what is expected

for classic chemical denaturants such as urea [81, 75]. The black bar denotes the native state (N) and the gray bar denotes the unfolded state (U). Units on the y-axis are arbitrary for the chemical potential of a given state, double-arrows are included to more easily see how the unfolding free energy ($\Delta\mu_2^{(N\rightarrow U)}$) increases or decreases with a change in osmolyte concentration. The results in Figure 4.5 for sorbitol and trehalose are consistent with Figure 4.7-A, although with relatively small changes to the native state chemical potential with increasing osmolyte concentration.

Figure 4.7-B shows a schematic that is more appropriate for the cases of AS-IgG1 with sucrose (top) and PEG (bottom). The chemical potential of the native state change non-monotonically with sucrose concentration, but the value of $\Delta\mu_2^{(N\rightarrow U)}$ increases linearly with the addition of sucrose. As such, preferential interactions of sucrose with the unfolded state must also depend on sucrose concentration in a non-trivial way. For PEG, it is a similar but reversed scenario to sucrose, in that $\Delta\mu_2^{(N\rightarrow U)}$ decreases linearly with the addition of PEG. Overall, these results indicate that preferential interactions of osmolytes with the native state should not be generally assumed to be predictive of interactions with the unfolded state(s). Osmolytes that are preferentially excluded in the native state (e.g. PEG or sucrose) may increase or decrease protein stability.

While it is common in current practice to assume that the osmolytes obey one of the two scenarios in Figure 4.7-A, Timasheff and others provided cautionary examples to the contrary [201, 96]. Recent data highlights that PEG interactions depend on the size of the PEG molecules in question, and that PEG preferential interactions are not simply the steric repulsions expected for an idealized hydrophilic polymer [3, 202]. However, sucrose and high-molecular weight PEG are canonical examples of osmolytes that not only are preferentially excluded, but also do not show significant changes in their behavior as a function of osmolyte concentration [3, 75, 85, 202]. The results here indicate that caution is needed in assuming such simplified models hold for a given protein, and illustrate experimental means to test such assumptions.

Finally, the results for sucrose may be of interest in another context. Sucrose

and trehalose have long been used as lyoprotectants and cryoprotectants for proteins during freezing and drying processes [90, 92]. One qualitative mechanism by which they putatively operate to stabilize proteins in that context is by serving as “water replacers”, in that they are able to hydrogen bond to the protein as the protein “loses” water due to water removal by drying or by freezing of ice [78, 90]. Such water-replacement would be expected to manifest as accumulation of the sugar molecules at the protein surface, at least for sufficiently high sugar concentrations [78, 90]. The present results extend only to sugar concentrations on the order of 30 w/w%, which is a factor of 2 or more lower than sugar concentrations in freeze-concentrated or freeze-dried states [203]. However, from a qualitative perspective the results for sucrose clearly indicate that a switchover from preferential exclusion to preferential accumulation is possible, and even indicate that it occurs more readily for sucrose than for trehalose in the case of AS-IgG1. While it is beyond the scope of this report, preliminary results with a canonical globular protein show qualitatively similar behavior and suggest this observation is more general than what is shown here for AS-IgG1; that study will be the topic of a future report.

Returning to question of AS-IgG1 unfolding and stability in solutions, it may be useful to note that AS-IgG1 is a large, flexible, non-globular protein with three cooperative unfolding transitions. Additionally, the unfolded state of AS-IgG1 may have significant residual structure from the 16 disulfide bonds, and accurately determining the ASA of the unfolded state remains especially challenging and may be similar to unfolded state of globular proteins. For AS-IgG1 unfolding with the addition of PEG shown in Fig 4.5-D revealed decreases in the temperature of Fab unfolding more so than C_H2. In contrast, trehalose, sorbitol, and trehalose, appeared to increase the temperature of the Fab and C_H2 similarly.

It remains an open question why trehalose stabilizes AS-IgG1 more so than sucrose. Given their very similar molecular structure and size, one would expect excluded volume contributions and chemical interactions with the protein surface to be similar. However, trehalose has been shown to increase stability over sucrose for many proteins

[78]. Many theories have been proposed for the enhanced stability of trehalose. One theory suggests the enhanced preferential exclusion of trehalose increases the chemical potential of the unfolded state more so than the folded state (cf. Fig 4.7-A). Another theory, which originates to explain protein lyophilization, suggest the sugar replaces the water molecules during the freeze-drying process at high sugar concentrations and thereby better maintains hydrogen bonds needed for protein secondary/tertiary structure. Lastly, a vitrication theory has been proposed, which suggests trehalose creates a glassy network around the protein, thereby reducing conformational fluctuations and propensity to unfold [78, 92]. The present results with sucrose suggest preferential exclusion at low osmolyte concentrations and preferential accumulation of sucrose at higher osmolyte concentrations appear to be consistent with both behaviors. However, current results do explain the enhanced stability of proteins with trehalose that has been observed [78, 92].

The mechanism by which PEG destabilizes AS-IgG1 remains unclear. PEG was shown to be preferentially excluded from AS-IgG1 and classical theory would suggest enhanced preferential exclusion from the unfolded state. The glycosylated chain present in the hinge region (Asn 297) of AS-IgG1 may add another layer of complexity [111]. Perhaps, one may expect osmolyte preferential interaction with the glycan, where hydrogen bonding is expected, may affect stability. Prior work has shown that de-glycosylated Fc region has lower stability compare to the glycosylated Fc [110]. While beyond the scope of this work, a more detailed investigation of the mechanism by which osmolytes decreases stability and the effect of glycosylation may provide insight into preferential interactions of osmolyte and glycoproteins such as monoclonal antibodies.

4.9 Summary and conclusions

Interactions of AS-IgG1 with water and common neutral osmolytes were quantified in terms of Kirkwood-Buff integrals, determined from protein partial specific volume measurements. Sucrose and PEG were preferential excluded from the native

AS-IgG1 at low osmolyte concentrations but, surprisingly, became preferentially accumulated at higher concentrations. In contrast, sorbitol and trehalose showed no significant interaction (within the sensitivity of the measurements). Determining Kirkwood-Buff integrals (G_{12} and G_{23}) offers a direct link to changes in the protein chemical potential with osmolytes, beyond what can be obtained from preferential-interaction measurements that can only probe the difference between G_{12} and G_{23} . The resulting transfer free energy ($\Delta\mu_{2,tr}^N$) for the transfer of a native protein from water to water plus a given osmolyte may therefore depend strongly on osmolyte concentration, and can be large compared to typical unfolding free energies (ΔG_{un}). We considered whether AS-IgG1 native chemical potential or *a priori* preferential model based on ASA are indicative of changes in ΔG_{un} as a function of osmolyte type and concentration. Thermal denaturation via DSC illustrates net stabilization with trehalose, sorbitol, and sucrose, but destabilization with PEG. Particularly, PEG destabilizes the Fab domain(s) more so than the C_H2 or C_H3 domains, suggesting preferential interactions may be specific to different regions of AS-IgG1. This result along with concentration dependent values of $\Delta\mu_{2,tr}^N$ suggest additional complexities to the classical preferential interaction model for interpreting osmolyte effects on protein unfolding. In addition, results illustrate pitfalls when adopting approaches that assume that $\Delta\mu_{2,tr}^N$ will be predictive of protein stability as a function of osmolyte type and concentration.

Chapter 5

AGGREGATE STRUCTURAL CHANGES AND MECHANISMS AT ELEVATED CONCENTRATION

5.1 Introduction

As discussed in Chapter two, aggregation mechanisms may change depending on solution conditions. While prior work has focused primarily on categorizing aggregation mechanisms based on how the average aggregate molecular weight changes as the process proceeds, it remains unclear if the underlying structure and aggregate morphology must also change as one alters solution conditions to mediate the aggregate growth mechanism(s) [1].

Kim et al. reported increased anti-parallel β -sheet content (via FTIR) for AS-IgG1 aggregation via the ND mechanism, when compared to the CP or AP mechanisms at low protein concentration [1]. However, recent results have also suggested that IgG1 aggregation at higher protein concentrations (i.e. ≥ 100 mg/mL) may occur first through native, reversible complexes that then unfold/rearrange to form irreversible aggregates. Elevated protein concentrations promote weak but non-negligible native self-association [80, 204, 205] that may lead to aggregation. *Ex-situ* measurements cannot discern between a mechanism in which unfolding occurs prior to self-association, rather than after it. Therefore, monitoring aggregation via *in-situ* measurements at elevated protein concentrations may provide insights into aggregation mechanisms that may be less kinetically important at lower protein concentrations.

As mentioned in Chapter one, high protein concentrations (≥ 100 mg/mL) are needed for dosing requirements of subcutaneous injections and intermediate protein concentrations are required for intravenous administration (10-50 mg/mL)). However,

Chapters two and three investigated the aggregation behavior at dilute protein concentrations (1 mg/mL) and changes in the protein stability with the addition of common osmolytes were highlighted in Chapter four. It remains unclear if and how osmolytes will alter protein-protein interactions, the aggregation growth mechanisms, and resulting aggregate structure and morphology at elevated protein concentrations.

This chapter provides a systematic characterization of AS-IgG1 aggregation at an intermediate protein concentration (30 mg/mL) with deuterated water and in the presence and absence of 0.15 M (~ 5 w/w %) sucrose. Aggregate morphology and structure are monitored for ND, CP, and AP growth mechanisms in the presence and the absence of sucrose using a series of complementary techniques. Structural changes during aggregation are monitored with traditional *ex-situ* techniques, such as circular dichroism and UV absorption. *In-situ* aggregation is monitored with a combination of dynamic light scattering, Raman scattering, and small-angle neutron scattering (SANS). The results highlight the effect of pD, NaCl, and sucrose on the growth mechanism. We consider if aggregation behavior under dilute protein concentration (Chapters 2-4) is applicable to new solution conditions at elevated protein concentration and with (without) sucrose. Parts of this chapter are reproduced with permission from [34], copyright 2015 American Chemical Society.

5.2 Materials and Methods

5.2.1 Sample Preparation

AS-IgG1 solutions were buffer exchanged into deuterium oxide (D₂O) (99% D atom, Sigma Aldrich) solutions. Buffers were prepared in 10 mM acetic acid (Fisher Scientific), and in some cases with the addition of 100 mM NaCl (Fisher Scientific) and/or 0.15 M sucrose ($\geq 99.5\%$ (HPLC grade), Sigma Aldrich). Solutions were adjusted to desired pD with 1 M sodium hydroxide (Fisher Scientific) stock solution in D₂O using the standard expression (Equation 5.1), which relates the pD value to pH*, the apparent pH value read from the instrument [131].

$$pD = pH^* + 0.4 \quad (5.1)$$

Each protein solution was buffer exchanged with its corresponding buffer using Amicon centrifuge tubes (Millipore, Billerica, MA) a minimum of four times. Following buffer exchange, the protein concentration was determined as described in Section 2.2.1, and values for the final protein concentration and pD were checked and reported in solution conditions below.

5.2.2 *Ex-situ* monomer loss kinetics and light scattering using SEC-MALS

Solutions were prepared in the following solution conditions based aggregation mechanism determined in Fig. 2.2: pD 5.3 and 10 mM acetate; pD 4.6 with 100 mM NaCl and 10 mM acetate; and pD 5.1 with 100 mM NaCl and 10 mM acetate. The final protein concentration was 30 mg/mL in all cases. Solutions with the same conditions as above, but also containing 0.15 M sucrose were also prepared. Aliquots of each protein solution were placed into HPLC vials (Waters, Milford, MA) and hermetically sealed before incubation. Vials were heated isothermally, and at a series of preselected time points, vials were quenched in an ice bath to arrest aggregation.

Quenched protein solutions were diluted to a final concentration of 0.5-1.5 mg/mL and left at room temperature (20-23 °C) before injecting onto the HPLC. For each diluted protein solution, the monomer fraction and molecular weight were determined with size exclusion chromatography with in-line multi-angle light scattering (SEC-MALS) as previously described in Section 2.3.

5.2.3 Circular Dichroism

Quenched protein solutions were diluted to 0.5 mg/mL, and far-UV circular dichroism (CD) spectra were recorded using a Jasco J-810 spectrophotometer (Jasco, Easton, MD) at 25 °C. CD spectra were measured from 200 to 250 nm at a scan rate of 20 nm/min using 1x10 mm Hellma cuvettes (Plainview, NY). Ten spectra were collected and averaged for each measurement. As previously described, the buffer spectrum was subtracted and the mean residue ellipticity, $[\theta]$, was calculated using the standard expression (Equation 5.2)[206]. Here, M_0 is the monomer molecular weight,

c_0 is the protein concentration, d is the cuvette path length, and n_r is the number of amino acids making up AS-IgG1.

$$[\theta] = \frac{\theta M_0}{10c_0dn_r} \quad (5.2)$$

Using the known monomer fraction, (m), the monomer contribution to a given CD spectrum was subtracted at a given wavelength by subtracting the product of m and the measured CD spectra for a pure monomeric solution, $[\theta_m]$ via Equations 5.3 and 5.4 as previously described [42, 27].

$$[\theta] = m[\theta_m] + (1 - m)[\theta_{agg}] \quad (5.3)$$

$$[\theta_{agg}] = \frac{[\theta] - m[\theta_m]}{1 - m} \quad (5.4)$$

The representative aggregate contribution to the measured ellipticity for a given sample, $[\theta_{agg}]$, was calculated from Eq. 5.4 for each wavelength. This analysis allows one to distinguish between changes in spectra that are due to different amounts of monomer present for samples that are incubated for different times to promote aggregation. In Eq. 5.3 and 5.4, m depends only on the incubation time for a given solution conditions, while the ellipticity values depend on both time and wavelength.

5.2.4 Second derivative UV absorption

Quenched samples were diluted to 1.5-2 mg/mL and the absorbance spectra from 200-400 nm were recorded using a UV-vis spectrometer with a photodiode array (Agilent Technologies). Spectra were collected over a 1 second integration time using a 2x10 mm quartz cuvette (Hellma, Plainview, NY) and measurements were blanked using corresponding buffer solution. The data were exported and analyzed in Matlab to give the second derivative spectrum calculated using a nine point data filter and third order Savitzky-Golay polynomial, as described elsewhere [207]. The wavelength shift of either the tyrosine or tryptophan peak were determined from shifts in the local minima of the second derivative spectra.

5.2.5 *In-situ* dynamic light scattering with Raman spectroscopy (DLS-Raman)

Dynamic light scattering (DLS) with Raman spectroscopy were performed on a Zetasizer Helix system (Malvern Instruments, Malvern, UK), which combines dynamic light scattering (DLS) with Raman spectroscopy. Raman scattering was excited by a 785 nm laser with approximately 280 mW power, while DLS was collected at the 173 degree backscattering angle from a 632 nm laser. Temperature dependent viscosities for heavy water were determined from available literature and used for calculation R_h using the Stoke-Einstein relation (Eq. 2.7). For a typical experiment, $\sim 50 \mu\text{L}$ sample was loaded into a titanium cuvette with 120 μm quartz window (Malvern Instruments, Malvern, UK) and placed into a Peltier temperature-controlled sample compartment. Raman and DLS data were acquired *in-situ* and semi-continuously, as the instrument allows one to alternate between Raman and DLS measurements while a sample is held continuously at a given temperature. To properly process the data, Raman spectra of corresponding buffer samples were acquired under the identical conditions with the same experimental configuration. Unless otherwise noted, Raman spectra were collected with 20 co-additions of a 15 second exposure. Detailed information regarding the Raman instrument and collection can be found elsewhere [208, 209, 210].

5.2.6 Small angle neutron scattering (SANS)

Neutron scattering was performed on the 30 m NGB beamline at the NIST Center for Neutron Research (Gaithersburg, MD). For conditions at ambient temperatures, three configurations were used to collect Q -dependent scattering over multiple orders of magnitude of Q : 1 m sample-to-detector distance (SDD) with 6 Å neutrons, 5 m SDD with 6 Å neutrons and 25 cm detector offset, and 13 m SDD with 6 Å neutrons. For *in-situ* conditions at elevated temperatures, scattering was collected using the 5 m SDD with 6 Å neutrons and 25 cm detector offset. Scattering data were corrected for detector background and sensitivity, and the scattering contribution from empty cells. The protein scattering profiles were also normalized by the incident beam flux and the

raw intensities were placed on an absolute scale using direct beam measurements. All data were reduced and analyzed using Igor Pro NCNR software with standard methods and analyzed according to Eq. 2.10 as previously described [130].

SANS was performed on unheated (monomer) AS-IgG1 at protein concentrations of 5, 10, and 30 mg/mL for each of the solution conditions and analyzed according to Equation 2.10. The structure factor ($S(Q)$) was calculated for 10 and 30 mg/mL solutions by assuming that form factor ($P(Q)$) is given by scattering intensity for 5 mg/mL. In a second set of experiments, SANS was performed on AS-IgG1 at each of the solution conditions while the protein was heated at elevated temperature.

For given solution conditions with and without 0.15 M sucrose, samples heated using a custom-built 10-cell sample chamber holder pre-equilibrated to the desired temperature. The temperature was controlled with an external circulator of ethylene glycol flowing through the 10-cell sample chamber holder. A titanium cell filled with water was inserted in the sample position adjacent to the protein sample. A calibrated thermocouple was inserted into the water cell and temperature was logged during scattering experiments to assure proper temperature control for the protein samples.

5.3 AS-IgG1 aggregation kinetics and growth mechanism

Table 5.1 lists the solution conditions, aggregation mechanisms (determined below), and incubation temperature used to create aggregates.²⁹ In subsequent sections, the nomenclature for solution conditions follows the categorized growth mechanism in far right column of the Table 5.1 (e.g. ND, CP, CP/AP), the category CP/AP denotes that both CP and AP occur together on similar timescales as noted previously for a number of proteins [27, 36, 1]

Aggregation was monitored at each solution condition and incubation temperature in the presence and absence of 0.15 M sucrose. For a given solution condition (i.e. pD, NaCl concentration, and sucrose concentration) AS-IgG1 was heated at a selected elevated temperature, and vials were quenched on ice at different time points to

Table 5.1: Solution conditions and aggregation mechanisms used to create aggregates

Solution Conditions	Aggregation Mechanism	Inc. Temp. (°C)
pD 5, 0 mM NaCl	Nucleation Dominated (ND)	69
pD 4.6, 100 mM NaCl	Chain Polymerization (CP)	53
pD 5.1, 100 mM NaCl	Association Polymerization (CP/AP)	58

effectively halt aggregation. The solution was diluted and SEC was performed to quantify the monomer fraction, which is defined as the fraction of the monomer peak area remaining for a heated sample relative to the monomer peak for unheated AS-IgG1. MALS inline with SEC was used to determine the molecular weight for eluting peaks. Incubation temperatures were selected to provide practically reasonable half-lives for monomer loss, based on short test experiments. Prior work showed that AS-IgG1 aggregates via unfolding of the Fab domains for the range solution conditions relevant here [1]. The particular solution conditions were selected to provide different aggregation mechanisms, based on Chapter two AS-IgG1 state diagrams at 1 mg/mL (Fig. 2.2).

Figure 5.1 shows monomer fraction remaining, m , as a function of incubation time (t) for AS-IgG1 at pD 5.3 with 10 mM acetate (closed squares) and with 10 mM acetate, 0.15 M sucrose (open squares). The inset in Fig. 5.1-A shows $m(t)$ for AS-IgG1 pD 4.6 with 100 mM NaCl and 10 mM acetate (red circles), and pD 5.1, 100 mM NaCl, 10 mM acetate (black triangles). In all cases, open symbols correspond to conditions with 0.15 M sucrose also present. AS-IgG1 aggregation growth mechanism for these conditions is inferred from inspection of Fig. 5.1-B, which plots M_w^{tot}/M_0 as a function of $(1 - m)^2$. Symbol types correspond to the same conditions as Fig. 5.1-A. As described in Section 2.3 plotting aggregation kinetics parametrically in this fashion allows one to determine the aggregation mechanism. Growth by monomer addition will result in a linear increase in M_w^{tot}/M_0 with $(1 - m)^2$. Growth via aggregate-aggregate

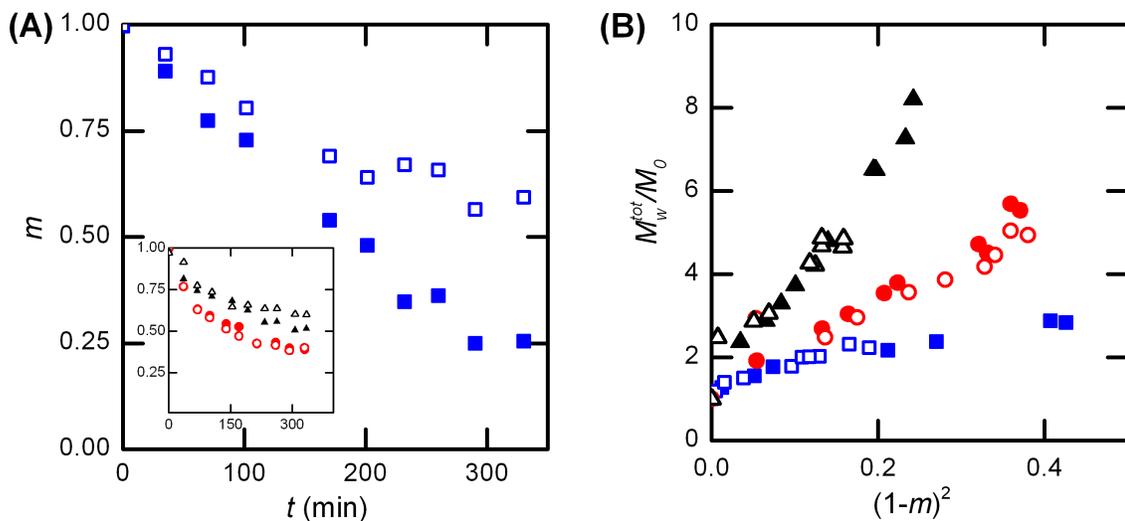


Figure 5.1: (A) Monomer fraction remaining as a function of time at elevated temperature for pD 5.3 heated at 69 °C (squares) and (inset) pD 5.1, 100 mM NaCl heated at 53 °C (circles) and pD 4.6, 100 mM NaCl heated at 58 °C (triangles). (B) M_w^{tot}/M_0 vs $(1 - m)^2$ plot. In both panels, open symbols correspond to solutions with 0.15 M added sucrose.

coalescence will result in non-linear increases in M_w^{tot}/M_0 , that show upward curvature when plotted versus $(1 - m)^2$, as coalescence does not consume monomer.

Inspection of Figure 5.1-B shows aggregation at pD 5.3, 10 mM acetate (blue squares) produces a mix of dimers and trimers and follows the ND mechanism (i.e. minimal growth). At pD 4.6, 100 mM NaCl, 10 mM (red circles), M_w^{tot} is linear with $(1 - m)^2$ and consistent with growth by monomer addition (i.e the CP mechanism). AS-IgG1 at pD 5.1, 100 mM NaCl, 10 mM (black triangles) also follows CP growth and possibly AP growth at later stages as M_w^{tot}/M_0 appears to be curving slightly upwards at larger values of $(1 - m)^2$. Comparison of Figures 5.1-A-B shows that the addition of 0.15 M sucrose (open symbols) slows the overall aggregation rate(s) in some cases, but has minimal or no change on the growth mechanism. For instance, for ND growth one observes slower monomer loss rates for sucrose compared to the buffer-only condition in Figure 5.1-A, but sucrose and buffer-only conditions have identical M_w^{tot}/M_0 profiles

in Figure 5.1-B. Interestingly, sucrose does not dramatically alter monomer loss rates or the growth mechanism for the CP or CP/AP conditions.

5.4 AS-IgG1 structural changes during aggregation

For the samples corresponding to the data in Figure 5.1, AS-IgG1 aggregation was quenched (cold temperature) after a given incubation time at elevated temperature, and samples were diluted to low protein concentration to be within the instrument working range for structural characterization using circular dichroism and second derivative UV absorption. Figure 5.2-A illustrates circular dichroism spectra for the case of ND growth over time from Figure 5.1. Previous work with AS-IgG1 [1] and other proteins [34, 27] have illustrated that CD spectra obtained for such quenched samples include contributions from the residual folded monomers and from the average over all aggregates in solution. Figures 5.2-B-D show $[\theta_{agg}]$ as a function of wavelength for ND (B), CP (C), and CP/AP (D) growth condition. Solid curves correspond to conditions with buffer only, while dotted spectra are for those with the addition of 0.15 M sucrose. Each spectra corresponds quenched time point from an isothermal incubation. As IgG1 was incubated at 30 mg/mL, quenched samples were diluted to 0.5-1 mg/mL, as mentioned above. From Figure 5.2B-D, one observes that for a given aggregation mechanism $[\theta_{agg}]$ spectra remains constant in the presence or absence of 0.15 M sucrose.

As described in the Section 5.2.3, the CD spectra in Figure 5.2-A were combined with the measured monomer fractions (cf. Figure 5.1-A) and separately measured CD spectra for monomer (unheated) solutions, to estimate the contribution to the CD spectra due to the aggregates via Eq. 5.3 and 5.4 by subtracting the contribution from unaggregated monomer. The curves of $[\theta_{agg}]$ as a function of wavelength was found to be independent of incubation time for a given mechanism. This supports the “two-state” treatment of the CD data using Eq.5.3 and 5.4, and is consistent with prior treatments[34, 27, 1]. Figure 5.2-E shows the average $[\theta_{agg}]$ versus wavelength for ND (blue dotted), CP (red solid), and CP/AP (black dashed) growth mechanisms.

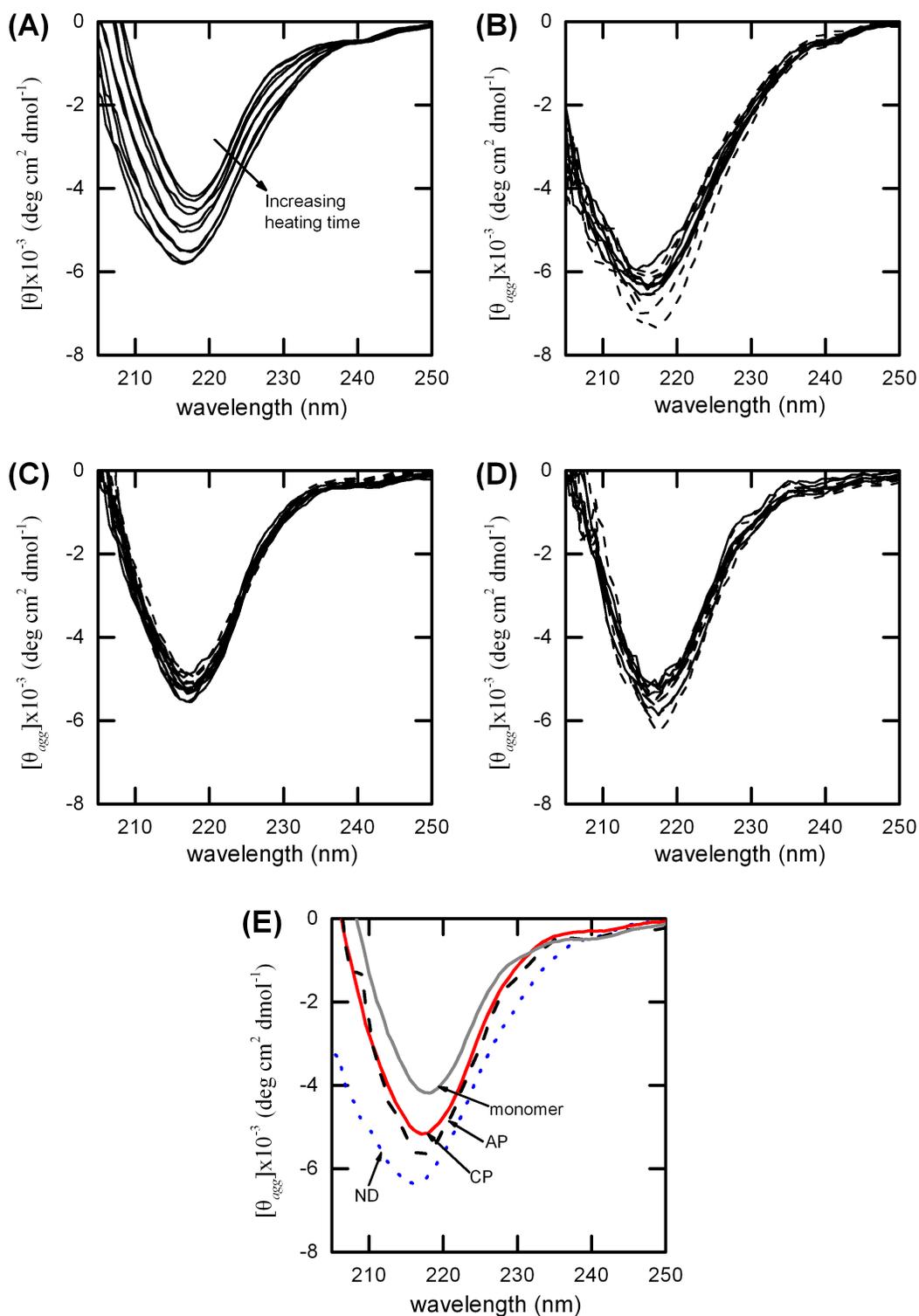


Figure 5.2: (panel A) Illustrative CD spectra over time AS-IgG1 pD 5.3 heated at 69 °C. Net aggregate contribution ($[\theta_{agg}]$) to the CD spectra for aggregate growth via (panel B) ND , (panel C) CP, and (panel D) AP mechanism. (panel E) Average $[\theta_{agg}]$ for each mechanism and the unheated AS-IgG1 monomer spectra.

These spectra correspond to averages of spectra in Fig 5.2-B-D. The CD spectrum for unheated AS-IgG1 monomer (solid gray line) is also shown. In Figure 5.2-E, one observes easily distinguishable differences in the CD spectra when comparing the different growth mechanisms. This suggests the underlying aggregate structure may depend on growth mechanism. All of the aggregate spectra are qualitatively consistent with increased β -sheet content compared to the monomer structure, which is expected based on results for other aggregation-prone proteins [45, 211, 133].

Figure 5.3-A shows illustrative second derivative UV absorption spectra for the conditions that show ND aggregation mechanism from Figure 5.1. The minima at 284 and 292 nm correspond to the tyrosine (Tyr) and tryptophan (Trp) residues, respectively, and shift to lower wavelengths as aggregation proceeds. Prior work has shown that such peak shifts correspond to changes in π - π interactions of the Tyr and Trp aromatic side chains, which are sensitive to the Trp and Tyr microenvironment and solvent accessibility [212]. Figure 5.3-B(C) quantifies how much the Tyr (Trp) peak shifts for different AS-IgG1 growth mechanism as a function of the extent of monomer loss, $(1 - m)$.

From Figure 5.3-B-C, the ND mechanism appears to have a much larger blue shift in the Tyr or Trp peak compared to CP or CP/AP growth. For CP and CP/AP growth mechanisms one observes a linear decrease in Tyr or Trp peak shift, while ND growth has notable downward curvature. The former is consistent with a relatively constant microenvironment during the aggregation process, while the latter suggests the Tyr/Trp microenvironment is changing as aggregation proceeds since the change in signal is not simply linear in the extent of mass converted from monomer to aggregate. Protein unfolding is expected to exhibit blue shift as the protein unfolds and aromatic side chains become exposed [207]. However, conditions were not identified that allowed only unfolding to occur for AS-IgG1, that would have allowed a quantitative comparison between the changes in second derivative UV spectra for aggregates and unfolded monomers. Interestingly, Tyr/Trp microenvironments of the aggregate are not affected by presence of sucrose, as conditions with and without sucrose overlap in Fig. 5.3-B-C.

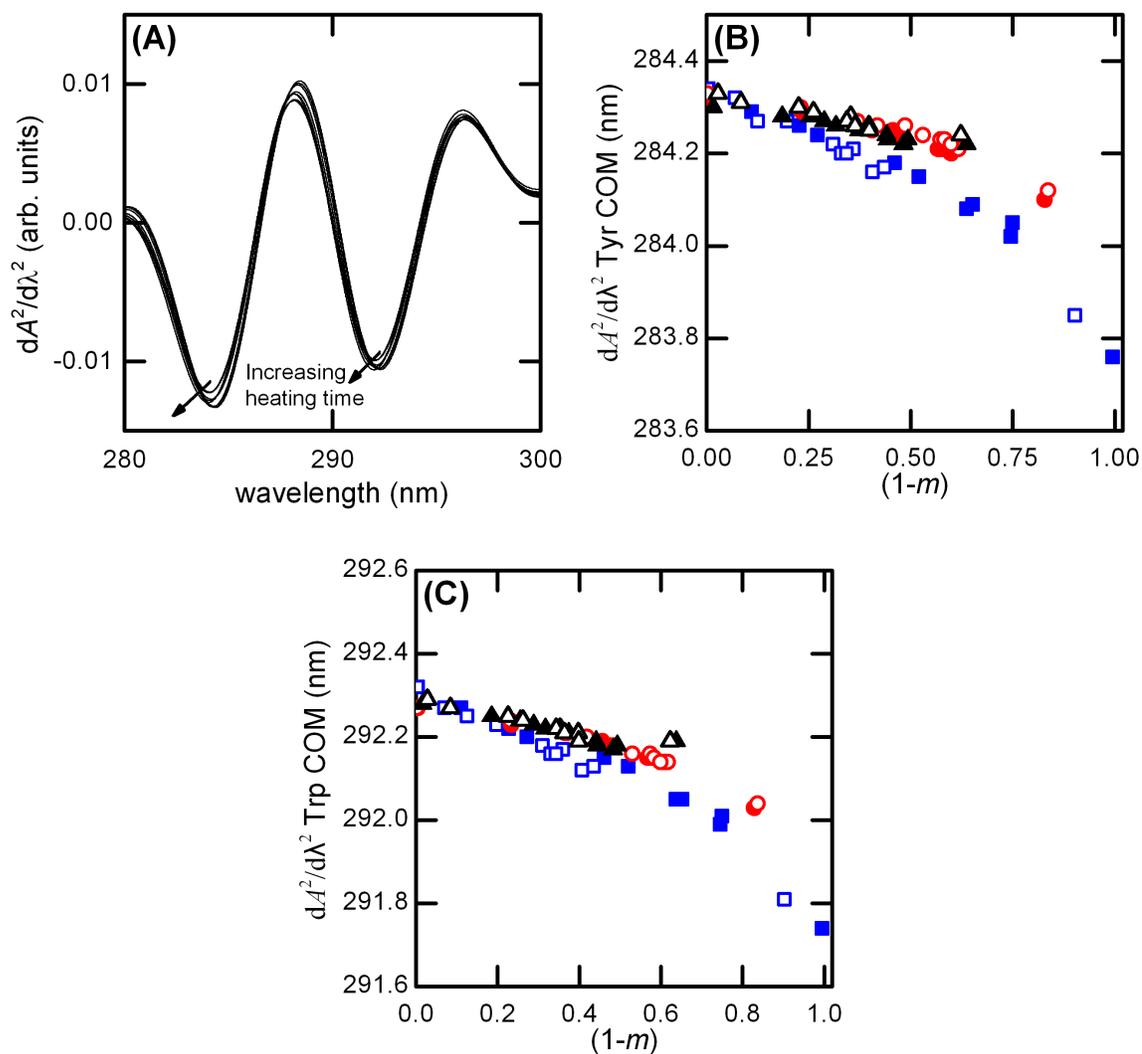


Figure 5.3: (A) Illustrative 2^{nd} derivative UV-vis spectra over time for AS-IgG1 via ND growth. (B) Tyrosine and (C) Tryptophan center of mass shift (COM) as a function of monomer consumption, $(1 - m)$. Symbols correspond to same solution conditions as Fig. 5.1

5.5 IgG aggregation monitored with *in-situ* DLS-Raman

In-situ DLS-Raman measurements were performed at a protein concentration of 30 mg/mL for each of the solution conditions illustrated in Figure 5.1. In what follows, the symbol shapes and colors correspond to the same conditions as in Figure 5.1. The bottom portion of Figure 5.4-A illustrates typical Raman spectra for AS-IgG1. The bottom spectrum corresponds to initial, unheated sample, while the middle and top spectra correspond to an intermediate and long incubation times, respectively. The Raman spectra contain many different structural markers that may be monitored during aggregation; a few are indicated with vertical lines that are added in Figure 5.4-A as guides to the eye. In particular, Tyr (dashed lines) and Trp (dotted lines) markers are sensitive to the microenvironment around the aromatic side chains, and Amide I and Amide III regions (solid lines) are pertinent to changes in the structure of the protein backbone (e.g., helix vs. sheet structures). Additionally, the top half of Figure 5.4-A depicts the Raman spectra for a subset of the range of wavenumbers (400 to 600 cm^{-1}), which contains information regarding the disulfide bonds.

Figure 5.4-B plots the Amide I center of mass (COM) as a function of time at elevated temperature for each condition from Figure 5.1. The Amide I region shows a red shift as the protein aggregates and follows a roughly exponential decay vs. time. To a first approximation, 80% of the Amide I signal corresponds to the carbonyl vibration mode, 10% corresponds to N-H bond bending and the remaining 10% to C-N bond stretching [213]. The Amide I region provides information with respect to changes in the secondary structure because the carbonyl and amide groups on the protein backbone participate in the hydrogen bonding that stabilize α -helices, β -sheets, and/or β -barrels. Inspection of Figure 5.4-B indicates that ND aggregation (squares) results in the largest change in the Amide I region, followed by CP/AP (triangles), and then CP (circles). The symbols that correspond to sucrose (open symbols) and buffer-only (closed symbols) solution conditions overlap which each other for CP/AP and CP cases. However, this is not the case for ND, which is consistent with the aggregation rates in Figure 5.4-A. Sucrose slows aggregation for the ND case but did not significantly slow

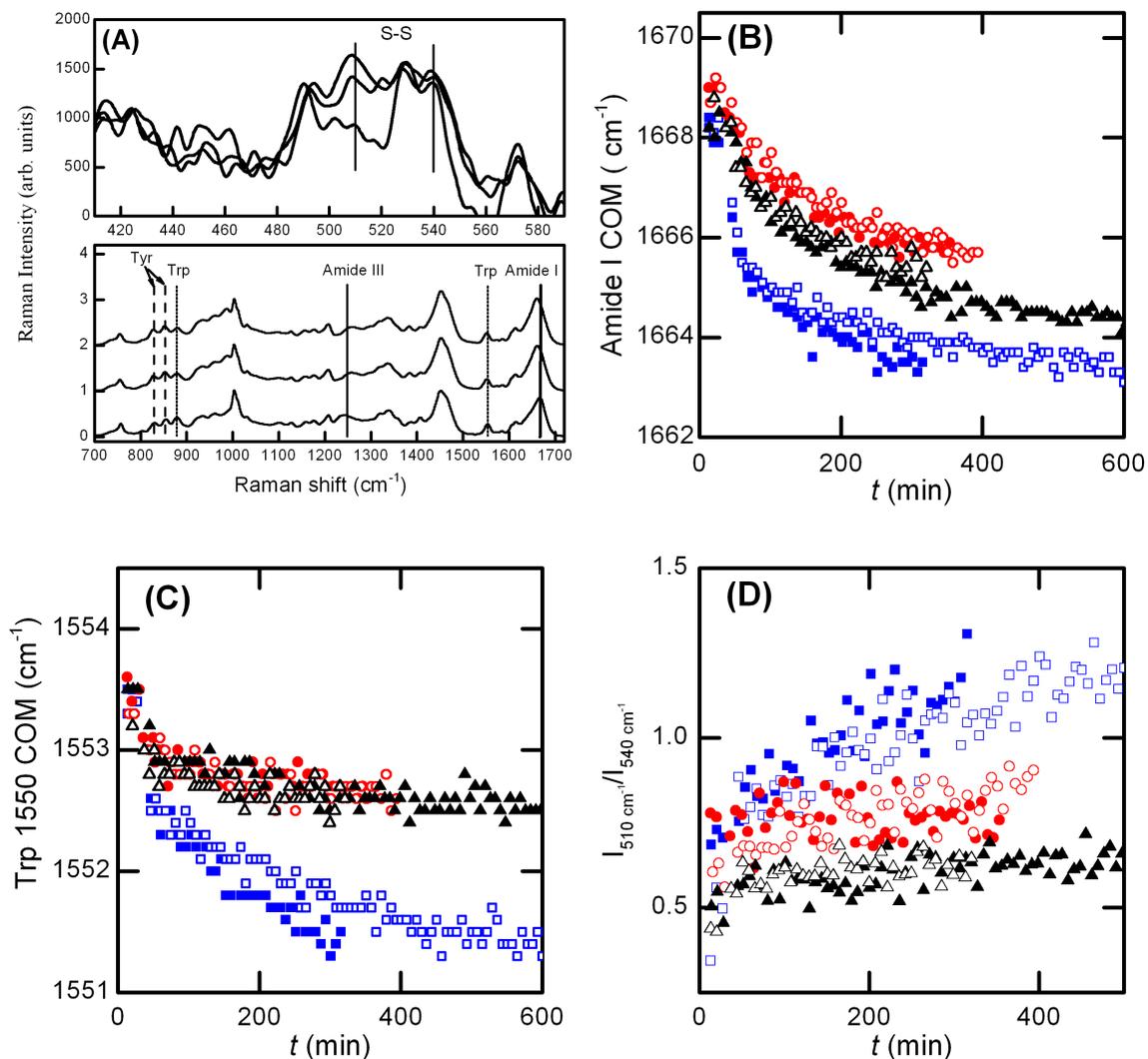


Figure 5.4: (A) AS-IgG1 *in-situ* DLS-Raman-I. (A) Illustrative Raman spectra for three time points at elevated temperature for IgG1 via ND growth. Raman markers plotted as a function of time for (B) Amide I region, (C) Trp at 1550 cm^{-1} , and (D) Ratio of Raman peak intensity at 510 cm^{-1} to 540 cm^{-1} . Symbols correspond to same solution conditions as Figure 5.1

aggregation for the other cases tested.

Typically, a blue shift in the Amide I region is observed with protein unfolding or aggregation using FTIR or Raman [213]. However, in the present case, AS-IgG1 solutions were prepared in D₂O, and the protein backbone may exchange a hydrogen atom for that of deuterium as the protein unfolds and aggregates. The heavier D-atoms are expected to slow the vibrational modes and lead to a red shift in the Raman Amide I region, as has been observed previously [214, 215, 216].

Raman spectra offers multiple markers to monitor aggregation that may or may not be convoluted by H-D exchange. The Tyr markers exhibit changes as protein was heated, however the markers at 830 cm⁻¹ and ~857 cm⁻¹ are sensitive to the hydrogen bonding of the phenoxy group in the Tyr side chain and may be convoluted with H-D exchange (see Figure 5.5 below)[215]. In contrast, Figure 5.4-C shows how the Trp marker at 1550 cm⁻¹ changes over time at elevated temperature. This peak arises from vibration modes of the aromatic side chain and the peptide bond plane. It gives a relative measure of the protein tertiary structure, and is not expected to be convoluted by H-D exchange. Inspection of Figure 5.4-C shows a red shift for the Trp marker as the protein is heated and aggregates. Interestingly, the red shift is more pronounced for the ND case compared to AP and CP cases. Also notable, the CP and CP/AP cases have Raman shifts that are indistinguishable. Again, one observes that conditions with sucrose are similar to conditions with buffer only.

Raman spectroscopy also provides insights into the conformation of disulfide bonded side-chains. Figure 5.4-D shows the ratio of the Raman intensity at 510 cm⁻¹ to 540 cm⁻¹, which is sensitive to the conformation, and packing density of disulfide bonds [213, 217]. Particularly, the disulfide region is sensitive to the gauche (G) and trans (T) conformers of the CCSSCC group [213]. The Raman intensity at 510 cm⁻¹ is attributed to the S-S stretching of the GGG conformer while the intensity at 540 cm⁻¹ corresponds to the S-S stretching of the TGT conformer. Interestingly, as aggregation proceeds, the ND case has larger perturbation in this disulfide marker compared to CP or CP/AP cases.

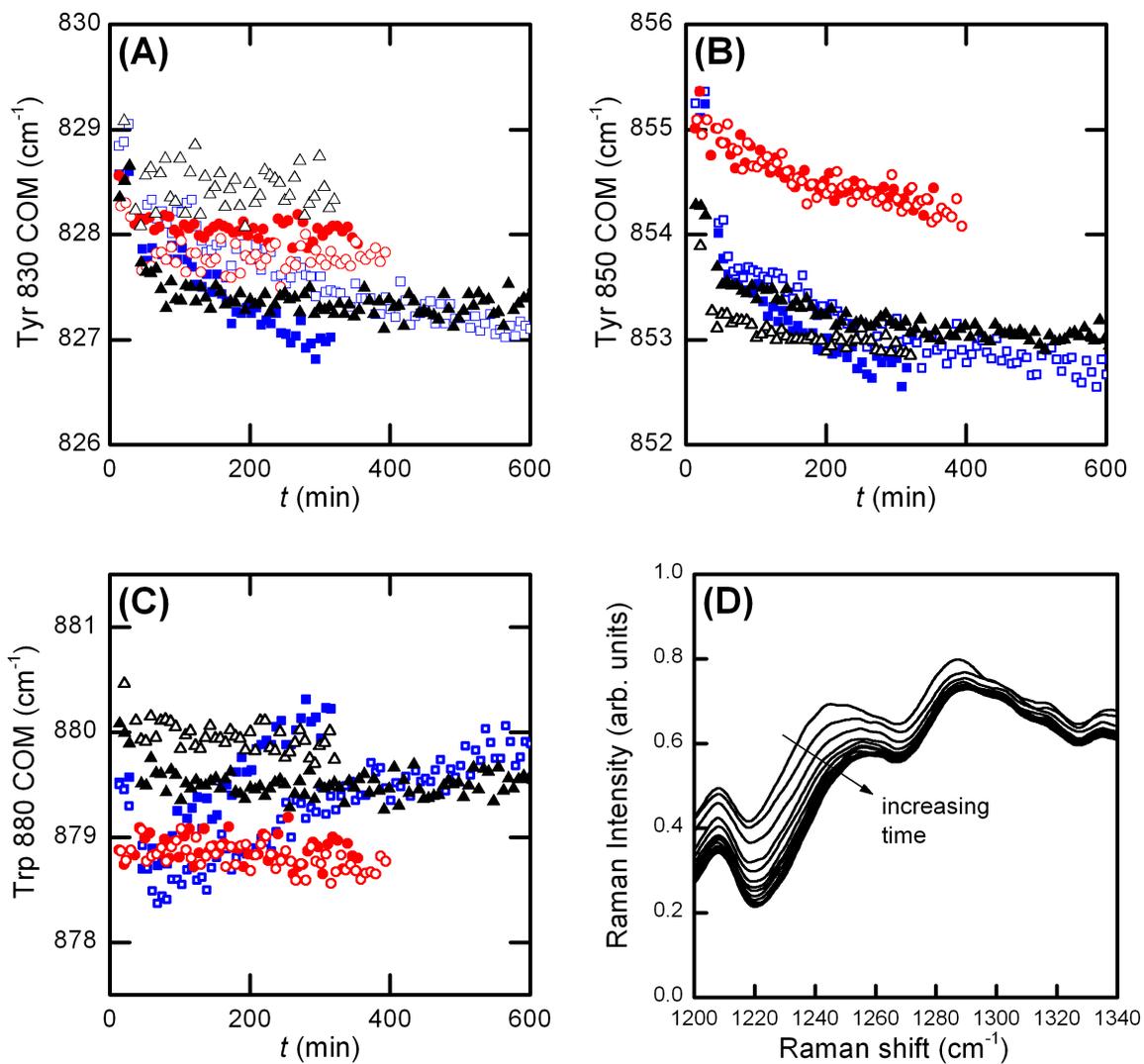


Figure 5.5: (A). AS-IgG1 *in-situ* DLS-Raman II. (A) Tyr COM at 830 cm^{-1} (B) Tyr COM at 850 cm^{-1} (C) Trp COM at 880 cm^{-1} and (D) Amide III region. Symbols correspond to same conditions as in Figure 5.1.

In the present case, H-D exchange may convolute the conformational changes occurring during aggregation. Figure 5.5 shows a few Raman markers that may be influenced by these two phenomenon. In particular, Fig. 5.5-D shows significant changes in the Amide III region. This result is attributed to H-D exchange on the protein backbone as Amide III is predominantly influenced by the N-H bending mode [213]. However, it was not possible to deconvolute hydrogen-deuterium exchange from protein structural changes in the present case. Detectable H-D exchange is expected to occur for regions of the IgG1 that become exposed at elevated temperature, because all protein solutions were prepared and equilibrated in D₂O buffer for multiple days before incubation. The trends in the Raman shifts observed in the Amide I region in Figure 5.4 are consistent with aggregation convoluted with H-D exchange and will a subject of future work. However, a comparative study with these spectroscopic techniques in H₂O was not performed.

While protein solutions were pre-equilibrated in D₂O buffers many days before, H-D exchange may occur on buried regions of the protein that become exposed during unfolding. Figure 5.5 shows Raman markers that may have been convoluted with H-D exchange. Figure 5.5A-B plots the tyrosine Raman markers which are attributed to hydrogen bonding of the phenoxyl side group [213]. Figure 5.5-C plots values for Trp 880 cm⁻¹ peak over time, which is sensitive to changes in the hydrogen bonding of the indole ring of the Trp side chain and may be convoluted with H-D exchange [213]. Additionally, Figure 5.5-D shows illustrative Raman spectra in the Amide III over time for AP/CP conditions at 58 °C. This region corresponds predominantly to N-H bending mode of the protein backbone and large shift in the spectra at 1235-1250 cm⁻¹ is attributed to N-H and C-H bending of the β -sheet [213].

5.6 Scaling of Aggregate M_w^{agg} and R_h

In SEC-MALS, samples are necessarily very dilute and the aggregate peak(s) are separated from the monomer. Therefore the average aggregate molecular weight (M_w^{agg}) and hydrodynamic radius (R_h) determined using SEC-MALS do not have any

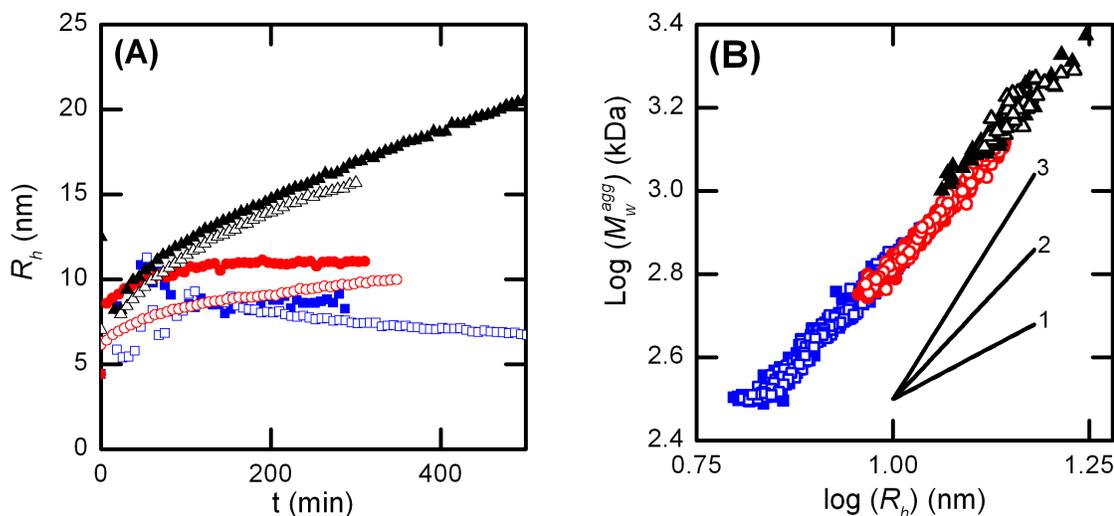


Figure 5.6: (A) *in-situ* R_h over time. (B) Scaling between M_w^{agg} and R_h . Symbols correspond to same solution conditions as Figure 5.1. Black lines in panel B correspond to slopes of 1, 2, and 3

contributions from the monomer and are not expected to be convoluted with protein-protein interactions. *In-situ* DLS was performed along with Raman scattering for the samples and conditions in Figure 5.4-A-C. Figure 5.6-A plots the z-averaged apparent hydrodynamic radius as a function of time at elevated temperature, which was calculated from the diffusion coefficient using the Stoke-Einstein relation (cf. Eq. 2.7). The increase in solution viscosity due to D₂O and sucrose at elevated temperature were accounted for in determining R_h [218, 219]. The presence of aggregate may also increase viscosity as has been shown previously, but such increases were not considered here and are expected to be relatively small contribution because aggregates did not grow as large as in previous studies [34].

Inspection of Fig. 5.6-A shows that the ND case produces aggregates that are relatively small, on average (i.e., less than ~ 10 nm for the effective R_h), while the CP and CP/AP cases reach values of R_h near and well above 10 nm, respectively. These differences in R_h are qualitatively consistent with the sizes (molecular weight values) in Figure 5.1-B. As DLS was collected *in-situ* at 30 mg/mL, protein-protein interactions

and hydrodynamic effects may convolute the z-averaged diffusion coefficient and the corresponding apparent R_h values. This may be one possible explanation for why the apparent R_h decreases with increasing incubation time at later stages for the ND case.

Figure 5.6-B shows the scaling of M_w^{agg} and R_h , as this type of scaling plot gives a qualitative measure of aggregate morphology [220, 135]. Each data point in Fig. 5.6-B corresponds to a slice of the aggregate peak eluting in SEC-MALS for a given sample condition, with samples quenched prior to injection for SEC. The symbols correspond to the same condition as Figure 5.1. Interestingly, all of the profiles follow a similar linear, almost overlapping, trend. This is reasonable, given that the range of M_w^{agg} and R_h data for the ND case overlaps with that for the CP case, but not the CP/AP case, while the CP data overlap with both ND with CP/AP cases. The solid lines in Figure 5.6-B provide guides to the eye for slopes of 1, 2, and 3. Overall, the data show a linear slope with a value between 1 and 2. This is a similar size-to-mass scaling as observed in Fig 2.3 at small aggregates sizes. Although those aggregates were created at lower protein concentrations and H₂O-based solution conditions, which resulted in larger sized aggregates.

5.7 IgG1 monomer structure factor using small angle neutron scattering (SANS)

Small angle neutron scattering was performed on unheated AS-IgG1 solutions at concentrations of 5, 10, and 30 mg/mL for each condition. Figure 5.7-A illustrates scattering intensity, $I(Q)$, after subtracting the background and normalizing for the protein concentration. As is common practice for protein solutions, the corresponding structure factor, $S(Q)$, was calculated by assuming the scattering profile for the lowest concentration (5 mg/mL) was dominated by the form factor because proteins are relatively weak scatters and $S(Q) \rightarrow 1$ at low protein concentrations [120].

Figure 5.7-B-D shows $S(Q)$ for 10 mg/mL and 30 mg/mL with buffer only and with the addition of 0.15 M sucrose for pD 5.3, 0 mM NaCl (panel B), pD 4.6, 100 mM NaCl (panel C), and pD 5.1, 100 mM NaCl (panel D). $S(Q)$ is a measure of

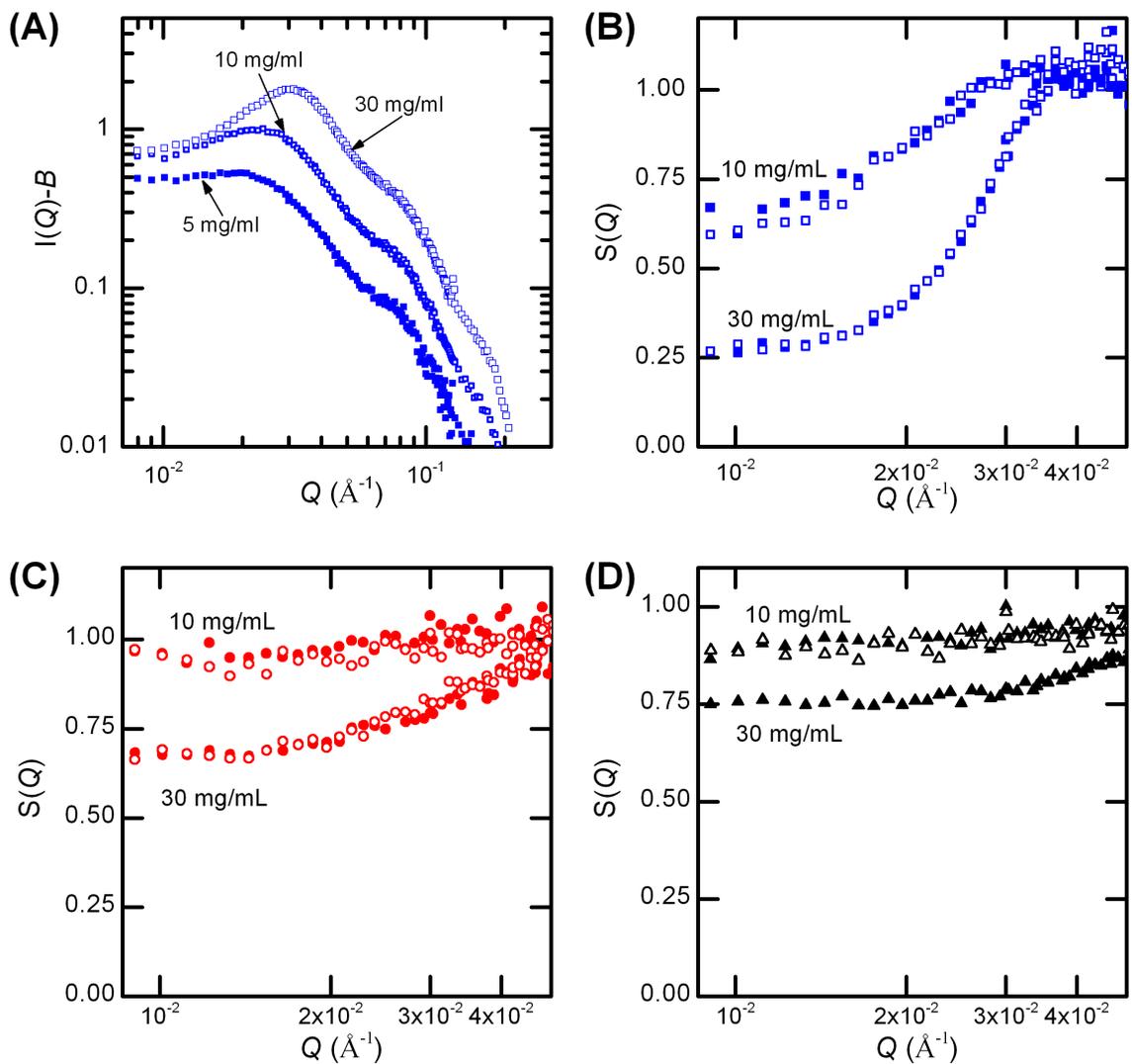


Figure 5.7: (A) SANS intensity as a function of Q for unheated (monomer) AS-IgG1 at pD 5.3 and protein concentration of 5, 10, and 30 mg/mL at 25 °C. Structure factor as a function of Q for (B) pD 5.3, 0 mM NaCl, (C) pD 4.6, 100 mM NaCl, and (C) pD 5.1, 100 mM NaCl. Closed symbols correspond to buffer only, and open symbols to buffer with 0.15 M sucrose.

the net protein-protein interactions (PPI) in solution. Values less (greater) than one correspond to net repulsive (attractive) PPI. $S(Q)$ as Q approaches zero is related to the osmotic compressibility [119] and the Kirkwood-Buff integral for protein-protein interactions [148]. However, at Q values larger than $\sim 0.03 \text{ \AA}$, $S(Q)$ approaches a value of one, and $I(Q)$ at that range of Q is dominated by monomer or aggregate morphology. For example, Figure 5.7-B indicates net repulsive PPI at 10 and 30 mg/mL. Based on this analysis, the solution conditions at pD 5.3, 10 mM acetate have the strongest repulsive PPI, followed by pD 4.6 with 100 mM NaCl, 10 mM acetate, and followed by pD 5.1, 100mM NaCl in 10 mM acetate. Interestingly, the addition of sucrose does not affect the Q -dependent PPI in any of these cases. As discussed in Chapter two, repulsive electrostatic PPI mediate IgG1 aggregation mechanisms at elevated temperature, suggesting sucrose may not alter mechanism. This result is consistent with the observation that sucrose does not alter the aggregation mechanisms (Fig. 5.1)

5.8 *In-situ* IgG aggregation using SANS

For each solution condition in Figure 5.1, protein aggregation was monitored *in-situ* using SANS at elevated temperatures. Figure 5.8-A illustrates the SANS $I(Q)$ profiles over time at elevated temperature for the CP case in buffer only (black curves) and with the addition of 0.15 M sucrose (blue curves). As aggregation occurs, the intensity at low Q increases, indicating an increase in total weight averaged molecular weight. The scattering at intermediate to high Q decreases with time, which is consistent with monomer changing conformation as it unfolds and with the consumption during the aggregation process. For the buffer-only conditions in Figure 5.9-A, aggregates grew to larger sizes (for a given time point) compared to same condition with added sucrose, which is consistent with Figure 5.1-B because aggregation rates were slower in the presence of sucrose.

In-situ SANS profiles were analyzed using so-called Kratky plots, which show $I(Q) * Q^2$ plotted against Q [119]. Traditionally, Kratky plots are used to investigate

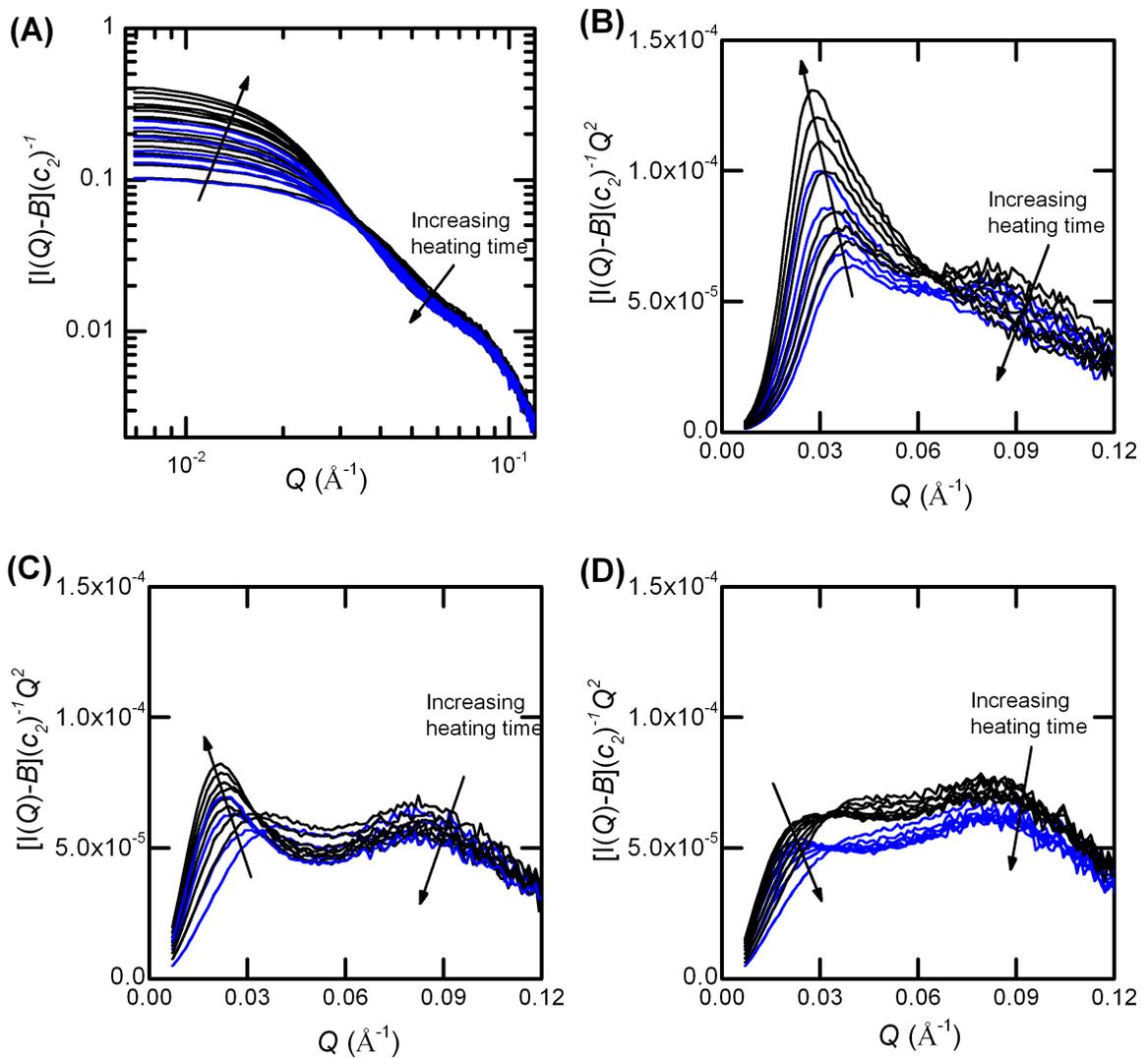


Figure 5.8: AS-IgG1 aggregation with *in-situ* SANS. (A) pD 4.6 with 100 mM NaCl and protein concentration of 30 mg/mL heated at 53 °C. Kratky plots for AS-IgG1 over time at elevated temperature. (panel B) pD 5.3 heated at 69 °C; (panel C) pD 4.6 with 100 mM NaCl heated at 53 °C; (panel D) pD 5.1 with 100 mM NaCl heated at 58 °C. Black curves correspond to buffer only, while blue curves are for 0.15 M sucrose.

polymer conformational changes. For semi-flexible polymers, the profile in a Kratky plot plateaus at high Q values, while for polymers in a globular conformation the profiles will reach a maximum at intermediate Q values then decrease towards a value of zero at high Q values. Additionally, normalizing the scattering intensity to $I(Q \rightarrow 0)$, and normalizing Q to the radius of gyration or the volume-of-correlation provides a useful tool to compare the conformation of various polymers or proteins with different M_w or R_g [137].

For the ND case, a significant structure factor contribution was observed and it was not possible to determine reliable $I(0)$ or R_g from a traditional Guinier analysis (not shown). Figure 5.8-B shows Kratky plots for ND growth. One observes from Figure 5.8-B that as aggregation proceeds, the peak at higher Q values decreases, suggesting monomer consumption leads to a loss in scattering in this high- Q region. Additionally, the peak at lower Q values increases over time at elevated temperatures. However, as observed in Figure 5.7, there is a significant influence of the structure factor at Q values less than 0.03 \AA^{-1} and therefore it is difficult to separate changes in the SANS profiles at these larger length scales where long-ranged inter-particle electrostatic interactions convolute with aggregate morphology. As mentioned above, for Q values greater than about 0.03 \AA^{-1} the value for the structure factor is essentially one and changes in the scattering intensity can be attributed to changes in the IgG1 monomer conformation [119].

For CP and CP/AP cases, there were not significant structure factor effects (cf. 5.7) and reliable values for $I(0)$ and R_g were determined from the Guinier analysis. In these cases, SANS Kratky plots were normalized using the values for $I(0)$ and R_g . Figure 5.9-A-B shows the normalized Kratky plots (NKP) for CP (panel A) and CP/AP (panel B) cases. Normalized Kratky plots allow one to compare changes in average aggregate morphology for aggregates across various sizes and for a wider range of the extent of monomer loss. Inspection of Figures 5.9-A-B shows that as aggregation proceeds, the first peak (at low values of QR_g) decreases slightly, while the second peak (higher QR_g values) tends to decrease and move towards even higher QR_g values.

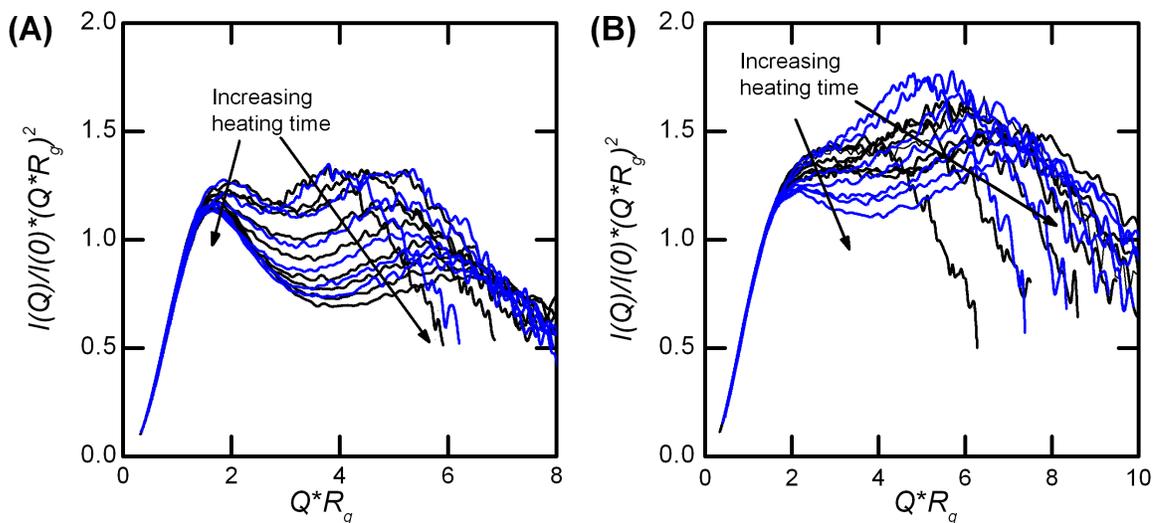


Figure 5.9: SANS Normalized Kratky plots (NKP) for AS-IgG1 over time at elevated temperature. Plots in Fig. 5.8-C-D were normalized to $I(Q \rightarrow 0)$ and R_g . (panel A) NKP for AS-IgG1 at pD 4.6 with 100 mM NaCl heated at 53 °C; (panel C) pD 5.1 with 100 mM NaCl heated at 58 °C. Black curves correspond to buffer only, while blue curves are for 0.15 M sucrose.

The first peak is indicative of the average mass and radius of the overall particle size distribution, while the secondary peak is related to average mass and size at much smaller length scales than the average R_g , and in the present case corresponds to monomeric protein.

The profiles in Figure 5.8-B have a peak at low QR_g values, which is indicative of the average particle mass and size. In the case of a fully unfolded protein that behaves similarly to an extended polymer, the Kratky plot is expected to increase at low Q values and reach a plateau [137]. In contrast, the CP/AP case in Figure 5.9-A does not have as well defined peak at low QR_g , possibly indicating increased flexibility of extended structures. Aggregates in the CP/AP case have larger average M_w^{tot} and R_g values compared to ND and CP cases. As aggregates grow and the peaks develop in Figure 5.9-B, the difference in length scales between the aggregates and monomers

increases (i.e., the scaling values of $I(0)$ and R_g are different for each curve in Figure 5.9-B). Overall, the time-dependent profiles in Figures 5.9-A-B are qualitatively different for each of the aggregation mechanisms tested here. Interestingly, the addition of sucrose slows aggregation, and after normalizing the scattering profiles, profiles with and without sucrose collapse again suggesting that sucrose does not play a role in the underlying growth process.

5.9 Structural changes during aggregation

The CD, second derivative UV absorption and Raman spectra all show that the underlying protein structure within aggregates is most perturbed, compared to unheated monomer structures, for aggregates created under the ND condition. The second derivative UV results not only showed larger structural changes for the ND condition, but also exhibited a non-linear decrease in Tyr peak position with monomer consumption. This suggests the protein structure within aggregates for the ND case is changing as aggregates grow from dimer to trimer. CP and CP/AP conditions in this work and prior work with a globular protein have shown that shifts in CD and intrinsic fluorescence spectra were linear with respect to the loss of monomer. That was interpreted as an indication that each monomer that was incorporated into an aggregate, no matter the size of the aggregate, underwent a similar structural change as part of that process; aggregate-aggregate coalescence did not result in significant changes in those spectroscopic signals because the constituent monomers had already undergone the structural change(s) needed for them to incorporate into the aggregates.

ND conditions result in little or no aggregate growth, and there are much larger concentration of dimers and small oligomer than for the CP and AP cases. It is speculated that dimer formation may involve a larger structural change for a constituent monomer compared to when larger aggregates incorporate monomers as they grow via monomer addition during CP conditions. Alternatively, it may simply be that the solution conditions that promote ND behavior require such repulsive inter-protein

interactions (e.g., strong electrostatics repulsions) that this also promotes larger intra-protein repulsions that lead to larger structural perturbations upon unfolding and aggregation. The observations here cannot reconcile which of these interpretations should be afforded a “causal” relationship to the observed results, but they do support a correlation between ND behavior and large structural perturbations that was noted previously at lower protein concentrations for two different mAbs [36, 1].

Raman scattering provided an orthogonal technique to probe aggregate structure. Consistent with CD and second derivative UV results, ND growth resulted in larger structural changes in the disulfide bonded side chains and Trp markers, which were not convoluted with H-D exchange. Additionally disulfide bonds were much more perturbed for ND growth compared to the other mechanisms, suggesting larger tertiary structural changes may occur while ND growth proceeds. While beyond the scope of this work, one may hypothesize that aggregates created via CP and CP/AP growth, which have less structural changes and grow to larger sizes, may have increased immunogenicity concerns because they more closely resemble the native IgG structure [12, 221].

5.10 Aggregate morphology from Kratky plots

The scaling of aggregate M_w^{agg} and R_h in Figure 5.6 shows a linear relationship between aggregate mass and size. It is difficult to discern if dimer or trimer sub-populations under AP conditions produce aggregates with a similar mass-to-size scaling for ND conditions because the scattering under AP conditions is biased towards the much larger particles in solution. However, SANS provides a more detailed technique to monitor aggregate morphology. SANS profiles at higher Q values can be attributed to IgG1 monomer contributions while the low- Q scattering is predominantly attributed to the aggregates. Normalizing the SANS profiles by the average aggregate mass and radius of gyration in the normalized Kratky plots (NKP) allows one to monitor the evolution of the qualitative aggregate morphology. Interestingly, each of the

growth mechanisms produces a discernably different qualitative NKP. As observed previously for this IgG1 and other proteins, the solution conditions (i.e. pH, NaCl, buffer) mediate the aggregation mechanism. As mentioned above, the addition of sucrose does not affect the aggregation mechanism in terms of the M_w^{tot} profiles in Figure 5.1-B, and normalized Kratky plots in Figure 5.9 appear identical for a given mechanism in buffer-only conditions or with sucrose present. Similarly, the spectroscopic profiles over time or compared to the amount of monomer consumed were unaffected by the presence of sucrose. Taken together, the results indicate that sucrose acts to slow the process of aggregation in some cases, but does not change the underlying mechanism(s) of aggregation. Presumably, sucrose slows aggregation by increasing the free energy of unfolding for the structural changes within the monomer that are precursors to nucleation, and this is consistent with previous reports [26] and with the observation that sucrose had the greatest effect on AS-IgG1 aggregation rates for ND conditions.

The SANS data and Kratky plot analysis provide insight into the compactness or flexibility of aggregates that is not afforded by other techniques. As discussed in Chapter two, Fig. 2.5 in dilute protein concentrations, where negligible structure factor contributions were observed, showed ND conditions produces a relatively compact morphology. While in the present case the Kratky plot for ND conditions (Figure 5.8-B) is convoluted with structure factor effects, results are consistent with a compact morphology. Additionally, in Fig. 5.9, the NKP for CP conditions initially shows similar compact morphology (with a well defined peak at low Q), but shows that the average aggregate flexibility increases as growth proceeds. This may be consistent with formation of a dimer and trimer mixture. Trimers and dimers may be expected to both have a compact morphology, but larger sized aggregates may have increased flexibility, and this would influence the SANS signal as aggregates grow. The NKP for CP/AP growth does not have a well defined peak at lower QR_g values, suggesting that process produces a more flexible aggregate morphology. One can speculate that this would be consistent with low fractal-dimension aggregates that result from association of multiple aggregates during the AP process.

NKP also illustrate the secondary peak in the SANS profile located at intermediate Q values that has been previously attributed to the flexibility of the IgG1 hinge region [109]. As aggregation proceeds and monomer is consumed, this intermediate- Q peak decreases, which may simply be indicative of the loss of monomer, given that the spectroscopic techniques generally show that the secondary and tertiary structure of protein chains in the aggregated state(s) are significantly different from that in the monomer state. The Kratky plot analysis also shows a decrease in the secondary peak at higher QR_g values, which is especially pronounced for ND conditions. This suggests that proteins lose much of their native structure as they incorporated into aggregates, consistent with the spectroscopic results for ND behavior discussed above. Overall, SANS provides a complementary tool to investigate aggregate morphology even at relatively high protein concentrations, and the analysis above illustrates an established means to interpret the results without the need to fit assumed models for structure factor or form factor.

5.11 Sucrose affects rates but not qualitative features of aggregation

Average aggregate structure and morphology were monitored as aggregation proceeded, with and without the addition of 0.15 M sucrose for each solution conditions that gave rise to different mechanisms. As observed in Fig. 5.1, the addition of sucrose tends to decrease the rate of aggregation rather than alter the qualitative growth mechanism/behavior. Additionally, sucrose markedly affects the rates in ND conditions, but minimally affects rates for the CP or AP conditions. The CD, second derivative UV absorption and Raman spectra also indicate that sucrose has little effect on aggregation rates for CP or AP conditions here. Perhaps surprisingly, sucrose did not change the protein-protein interactions determined from SANS (i.e., $S(Q)$). Sucrose did alter how large aggregates grew, but normalizing the SANS profiles by size and mass revealed each growth mechanism overlaps with and without sucrose, indicating that this again is simply a question of the net rate of aggregation. This suggests that the sucrose does not alter the aggregate growth process or mechanism. Rather, it

only decreases the rate of aggregation, consistent with its presumed role in decreasing the amount of unfolded monomer present in solution [74].

As discussed in Chapter four, sucrose and other saccharides are often added to protein solution as stabilizers. Sucrose is thought to be preferentially excluded from the protein surface as observed previously [222, 2]. The preferential exclusion mechanism is expected to have a more pronounced effect on the unfolded state compared to that of the native state and therefore increase the free energy of unfolding [223]. As a result, the concentration of unfolded “reactive” monomers will decrease in the presence of sucrose, which will decrease the aggregation rate [80]. Additionally, if a mechanism is dominated by nucleation, such as ND mechanism, one should observe a larger dependence on the monomer loss rate when one adds sucrose [74]. In contrast, if nucleation is dramatically slower than growth, such as CP and AP mechanism, sucrose will have less of an effect on monomer loss rates. However, available models would predict that the rate of monomer loss would still be affected appreciably by the addition of sucrose. As such, the relatively small effect of sucrose on AS-IgG1 aggregation rates for CP and AP/CP conditions merits additional consideration.

5.12 Summary and Conclusion

The aggregate structure and morphology were monitored for ND, CP, and CP/AP aggregation mechanisms for AS-IgG1 with and without sucrose present at typical formulation conditions (~ 5 w/w %). Sucrose did not affect the aggregation mechanism(s), or the resulting aggregate structure or morphology. Sucrose decreased monomer loss rates, which was most pronounced for the nucleation-dominated mechanism. Protein structural characterization with CD, Raman, and second derivative UV suggests aggregates via ND mechanism result in significantly larger structural perturbation compared to other growth mechanisms. *In-situ* DLS-Raman provided additional structural characterization that included Tyr, Trp environments disulfide bonding, and Amide I and III markers along with the z -averaged R_h . SANS Kratky plots provided a tool to monitor aggregate morphology and monomer loss. Each growth

mechanism showed a unique evolution of aggregate morphology with time in SANS, which complemented the structural changes detected with the other techniques. In general, aggregates that grew larger tended to exhibit increased flexibility. Overall, the results provide a comparison of aggregate structure and morphology using various orthogonal techniques aggregates created via ND, CP and AP growth mechanisms, and suggest that solution conditions that promote only small aggregates may also promote aggregates composed of monomers that are more structurally perturbed than conditions that promote larger aggregates.

Chapter 6

SUMMARY AND FUTURE WORK

6.1 Summary

This work focused on understanding how solution conditions, such as pH, NaCl concentration, buffer species, and the addition of common osmolytes (e.g sucrose, trehalose, sorbitol, PEG) mediated AS-IgG1 aggregation behavior. The dissertation investigated the role of protein-protein, protein-water and protein-osmolyte interactions on the protein stability and aggregation behavior of AS-IgG1. Changes in solution conditions and temperature led to highly non-linear changes in protein behavior. Results provided a mechanistic framework on the effect of protein-protein interaction on the aggregation mechanisms. Many techniques were utilized throughout the dissertation and provided insights into strengths and weakness associated with them. Many tools and methods can be applied to other protein (or polymer/colloid) systems in a similar fashion to gain mechanistic insight into their degradation or aggregation behavior.

The relationships among protein-protein interactions, aggregation mechanisms, and aggregate morphology were explored over a pH and NaCl concentration window typical of commercial mAb formulations. At low pH (well below the pI of AS-IgG1) and ionic strength, aggregates only grew to dimer/trimer/tetramer sizes and this behavior was categorized as nucleation domination growth. At pH 4 with 10 mM acetate, AS-IgG1 did not aggregate when heated at temperatures beyond mid-point unfolding temperatures for extended periods of time. As pH and NaCl concentration were increased, aggregates grew large through monomer addition to larger sizes. Further increases in pH and NaCl led to association polymerization dominated growth, whereby aggregates formed even larger sizes and potentially led to phase separation or condensation.

The pH-NaCl dependence on aggregation mechanisms suggested the importance of repulsive electrostatic protein-protein interactions on the aggregate growth mechanisms. Colloidal interactions between proteins (i.e. potential of mean force) putatively explained pH-NaCl effects on the aggregation mechanism. However, these simplified potential of mean force arguments failed to account for highly temperature dependent aggregation rates typical of mAbs and many proteins, and differences in acetate and citrate specific-ion-effects with AS-IgG1. Observed acetate/citrate specific-ion-effects were most pronounced at low ionic strength and particularly at pH 4, where the observed aggregation mechanism in citrate followed ND growth, but in acetate did not aggregate. A closer analysis of the Kirkwood-Buff integrals for monomeric protein-protein interactions showed that G_{22} semi-quantitative correlated with aggregation mechanisms. Results suggested that static light scattering measurements and Kirkwood-Buff analysis offers a potential tool to predict aggregation behavior.

Each of the aggregation mechanisms were categorized by how monomer consumption contributed to aggregate growth. Small angle x-ray and neutron scattering techniques were employed to investigate aggregate morphology. Normalized SAXS Kratky plots revealed a compact morphology of ND growth. Also Kratky plots showed the loss of mAb conformation and flexibility as the monomer was consumed and added to aggregates. The Kratky plots analysis also suggested the morphology did not change as aggregates grew from dimer to trimer and perhaps tetramer.

Laser light scattering provided a scaling relationship between average aggregate molecular weight and radius of gyration or hydrodynamic radius. The mass-to-size scaling exponent, or fractal dimension, increased as aggregates grew. This result suggested the aggregate morphology may change as aggregates grow. However, this result does not suggest aggregation mechanisms affect morphology. As scattering is always biased towards largest objects in solution, it is not clear the aggregate morphology for smaller sized aggregates changes based on mechanism.

Solution conditions also drastically change aggregation rates and comparing protein stability across many formulations can be challenging. A Parallel Temperature

Initial Rate device and method are introduced to accurately and efficiently determine temperature dependent degradation rates. The PTIR method is compared to the traditional isothermal approach and rate data show good agreement between methods. PTIR allows one to gather many more temperature dependent rates compared to the standard isothermal approach and determine activation energies with increased statistical confidence.

AS-IgG1 aggregation rates were collected over multiple orders of magnitude of timescales and solution conditions. AS-IgG1 thermal unfolding in 10 mM acetate via DSC showed pH and NaCl concentrations significantly altered protein conformational stability. Fc-IgG1 unfolding was also performed allowing one to clearly distinguish between Fab, C_{H2}, and C_{H3} unfolding transitions. At pH 4, the C_{H2} peak unfolded first followed by the Fab and C_{H3} peaks. As pH was increased from a value of 4 to 6, the C_{H2} and Fab unfolding transitions shifted to higher temperatures and overlapped, while the C_{H3} peak also shifted to higher temperature but remained indistinguishable. The addition of 100 mM NaCl decreased T_m values, but the decrease was more pronounced at pH 4 than at pH 6. This result is typical of mAbs and other proteins and presumably due to effect of electrostatic interactions of charge amino acids becoming screened with the addition of salt, which lowers the free energy of unfolding.

Aggregation rates changed dramatically across a broad range of pH, NaCl concentrations and no single temperature could be practically used to investigate isothermal rates. The rate coefficient data tended to collapse upon a common plot when rescaling incubation temperatures by T_m values. This result highlighted the primary effect of AS-IgG1 unfolding and conformational stability on the aggregation rates. With the increased accuracy of temperature dependent rates using PTIR, aggregation rates across solution conditions exhibited differences in effective activation energies. E_a values may have contributions from the free energy of unfolding, energy barrier of associating “reactive species (i.e. partially unfolded monomer or aggregates species), and energy associated with rearrangement to form irreversible intermolecular contacts

typical of non-native protein aggregation. In the present case, trends in E_a values suggested contributions from protein unfolding and strong electrostatic protein-protein interactions. At 100 mM added NaCl, E_a values followed expectations; as pH values increased from 4 to 6, E_a values increased as free energy of unfolding and T_m increased. This result highlighted the effect of conformation stability on E_a . However, with no added NaCl conditions, effective activation energies illustrated the effect of conformational stability and repulsive protein-protein interactions. This was most apparent at pH 4 and 5 in 10 mM acetate and pH 4 in 5 mM citrate, where E_a values were larger than those at pH 6. Particularly, pH 4 with 10 mM acetate showed unique behavior of not aggregating when heated well above T_m for multiple hours. Overall, results highlight the primary role of conformation stability on aggregation rates and the secondary effect that strong repulsive protein-protein interactions can play.

From a protein formulation context, the dissertation has shown that citrate anions preferentially accumulated near the AS-IgG1 surface more so than acetate anions. Also, the addition of NaCl decreases protein conformational stability and electrostatic protein-protein interaction, destabilized AS-IgG1. However, neutral osmolytes are commonly added to protein formulations to maintain solution isotonicity without increasing the solution ionic strength and decreasing protein stability. In addition, many sugars and hydrophilic poly-hydroxyl compounds actually stabilize proteins.

Preferential interactions of common neutral osmolytes with AS-IgG1 were determined at near neutral pH where AS-IgG1 has little or no net charge. Neutral osmolytes investigated include sucrose, trehalose, sorbitol, and PEG ($M_n = 6,000$ g/mol). Precise densimetry measurements were performed on AS-IgG1 ternary solutions (water, protein, and osmolyte) at varying levels of osmolyte molalities. The behavior of AS-IgG1 partial specific volume as a function of osmolyte concentration was directly related to protein-osmolyte and protein-water Kirk-wood Buff integrals. At low osmolyte concentrations, sucrose and PEG were preferentially excluded from AS-IgG1, but at higher osmolyte concentrations became preferentially accumulated. Preferential interactions

of AS-IgG1 with trehalose or sorbitol were not detectable within the densimetry measurement.

Differences between competing protein-water and protein-osmolyte interactions ($G_{12} - G_{23}$) were directly related to the changes in the chemical potential of the protein with respect to the osmolyte concentration. This conversion to protein chemical potential required knowledge of osmolyte-water interactions in the bulk solution. In the present case, for which protein concentrations were kept in the dilute limit, osmolyte-water binary interactions were utilized from available literature. In the case of aqueous solutions of either sucrose or sorbitol, analytical expressions for the molar activity coefficients were available in literature. However for binary systems of water with PEG or trehalose, water activity coefficients were found in literature and thermodynamic relation were subsequently used to determine changes in osmolyte chemical potential with osmolyte concentration.

From densimetry measurements and the Kirkwood-Buff analysis, changes in the chemical potential of the native state of AS-IgG1 were determined. The addition of sucrose or PEG increased the chemical potential of AS-IgG1 at low osmolyte concentrations and decreased the chemical potential of AS-IgG1 at higher concentrations. Densimetry results were compared to available models based on a proteins solvent exposed surface area (ASA). ASA was determined from the three dimensional structure of AS-IG1 homology model using available algorithms to calculate ASA. Models predicted linear changes in AS-IgG1 chemical potential with increasing osmolyte concentrations. This work showed osmolyte concentration dependent preferential interactions and preferential exclusion of sucrose and PEG and model predictions were are inconsistent with current results.

The effect of osmolytes on AS-IgG1 stability was determined using DSC. T_m values of the C_H2/Fab and C_H3 unfolding transitions increased linearly with sucrose, trehalose, and sorbitol, but decreased linearly with the addition of PEG. IgG1 was most stabilized by trehalose, followed by sorbitol, then sucrose. Additionally, PEG destabilized the Fab region more so than the C_H2 or C_H3, which suggested PEG may

interact with different regions of AS-IgG1 more so than others.

Overall, preferential interactions determined via densimetry showed that AS-IgG1 native-state chemical potential may be dependent on osmolyte concentration. However, AS-IgG1 unfolding with osmolytes showed linear changes in stability with osmolyte concentration. These results together preferential interactions of osmolytes with proteins may be more complicated than convectional classical theory would suggest. The classical preferential interaction theory suggests small changes in preferential interaction of native state of a protein and preferential interactions of the unfolded state predominantly contribute to increases free energy of unfolding. Results in this work suggested chemical potential of the native state may be non-linearly altered with increasing osmolytes concentrations. Results also suggested the chemical potential of the native state may not be predictive of chemical potential of the unfolded state (or protein stability). Particularly, both PEG and sucrose displayed concentration dependent preferential interactions, but upon AS-IgG1 thermal unfolding, sucrose stabilized the protein while PEG did the opposite.

The mechanistic framework for the role of protein-interactions on AS-IgG1 stability developed throughout the dissertation was under dilute protein concentrations. However, practical therapeutic formulations require intermediate to high protein concentrations ($\sim 10\text{-}150$ mg/mL) and non-ideal protein interactions may be expected. AS-IgG1 aggregation was extended to 30 mg/mL in deuterated solvent and in the presence and absence of 0.15 M (~ 5 w/w %) sucrose. In depth structural characterization was performed using *in-situ* Raman spectroscopy, second derivative UV absorbance, circular dichroism and compared to aggregate particle and morphology characterization performed using a size-exclusion chromatography with in-line light scattering, dynamic light scattering, and small angle neutron scattering.

Nucleation dominated growth mechanism exhibited large structural changes in the secondary structure, tyrosine and tryptophan microenvironments, and disulfide bond conformation. Interestingly, ND growth resulted in the smallest sized aggregates. The addition of sucrose decreased aggregation rates for ND growth mechanism, but

not so for CP/AP growth mechanisms. Sucrose did not alter growth mechanisms as M_w^{tot}/M_0 vs $(1 - m)^2$ profiles overlapped for conditions with and without sucrose. Also, aggregate morphology monitored *in-situ* with SANS Kratky plot analysis displayed unique scattering profile that was independent of the presence of sucrose, which also suggest sucrose does not alter aggregation mechanisms. AS-IgG1 monomer SANS structure factor was not altered by the presence of sucrose suggesting this osmolyte does not alter protein-protein interactions.

These results at elevated protein concentration are consistent with the current AS-IgG1 framework suggesting electrostatic-protein-protein interaction mediate growth mechanism. Together, results suggest sucrose and potentially other neutral osmolytes may only alter the concentration of unfolded protein and thereby decrease aggregation rates. Presumably, ND growth shows the largest decrease in aggregation rates with sucrose because this growth mechanism has a significant portion of dimers and based on mass action kinetics is expected to have larger change in monomer-loss rates.

6.2 Future work

This dissertation provides mechanistic insights into the role of protein-protein, protein-osmolyte, and protein-water interactions on the protein stability and aggregation behavior of AS-IgG1. Current work acts as a starting point for many different future investigations of protein aggregation, which are discussed below.

6.2.1 Refrigerated and room temperature predictions of protein aggregation rates

Current ICH guidelines require at most three stability temperatures (4, 25, and 40 °C). As such, these stability testing requires multiple months and data is typically gathered during development phases and even after regulatory filings. Furthermore, as observed in Chapter two, collecting aggregation rate data at three temperatures are essentially useless for predicting stability at other temperatures. Utilizing a more

predictive approach to protein aggregation in the biopharmaceutical industry would decrease the risk of developing unstable molecules and reduce uncertainty associated with time-intensive stability testing. Predictive methods requiring minimal experiments or ideally *in-silico* methods would enable rational design of stable drug candidates. Stable drug candidates are safer for patients and are typically easier to manufacture as they will have higher yields and require less intensive purification steps. Also, more stable drug candidates may potentially have faster regulatory approval and therefore increase speed to patients.

The PTIR approach allows one to accurately and efficiently determine temperature dependent aggregation rates. Future work may include extending PTIR capabilities to low temperature aggregation rates. PTIR may be a useful tool allowing one to determine low temperature and long-time aggregation rates. Additionally, an in-depth analysis of non-Arrhenius behavior of aggregation rates and the effects of protein interactions on non-Arrhenius behavior may provide mechanistic insights into long-term protein stability.

As discussed in Chapter two, one must exercise caution as to be sure other degradation pathways are not occurring simultaneously and artificially altering aggregation rates. Particularly, fragmentation has been observed at intermediate temperatures and may be faster than aggregation [107]. Additionally, deamidation (particularly at elevated pH) and oxidation may occur during incubation intermediate temperatures for extending timescales (e.g. weeks to months). These alternate degradation routes may or may not have a synergistic effect with aggregation.

6.2.2 Mechanistic investigation of aggregate coalescence and condensation

Aggregate coalescence leads to largest aggregate sizes and is a concern from an immunogenic and regulatory perspective (if visible particles). A better understanding of the mechanism of this process and aggregate phase separation would be useful from an immunogenic perspective. Aggregates that reach these large sizes almost invariably have undergone association polymerization. However, unanswered questions remain

regarding the thermodynamics and kinetics of association polymerization and phase separation (or condensation). Prior work has demonstrated that pH and salt titrations alter quasi-phase boundary of aggregates [37]. Interestingly, titrations showed hysteresis suggesting a complex thermodynamic/kinetic event may be occurring.

One potential strategy may involve mapping pH-salt phase diagrams of aggregate condensation with various sized aggregates. Prior work with aCgn has shown that under CP aggregation, the initial protein concentration alters the aggregate nucleation to growth timescales and the effect size of aggregates. Also, aCgn CP aggregation produces linear polymer-like aggregates with low polydispersity [45]. Additionally, titrations with PEG may provide useful insights into the coalescence mechanism. Traditionally, protein solubility measurement are often performed with the addition of PEG. In a similar fashion, aggregate solubility measurements with the addition of PEG may allow one semi-quantitatively determine aggregate chemical potential over different aggregate sizes and as a function of pH, salt and temperature. One may hypothesize that a combination of solution conditions (i.e. pH, salt, and temperature) and aggregate chemical potential may play a role in coalescence and condensation processes.

One unanswered discrepancy regarding aggregate coalescence centers around temperature dependent rates. Coalescence is thought to be a non-specific colloidal aggregation event as it does not consume monomer, but coalescence rates are strongly temperature dependent. Often quenching incubations originally at 60°C to 4°C or near room temperature effectively halts monomer-loss and coalescence rates. An in-depth investigation of temperature dependent coalescence rates may help elucidate this discrepancy. Additionally, monitoring of coalescence *in-situ* rather than *ex-situ* may be a useful tool as temperature strongly affects not only rates and also solubility/phase behavior. As such, quenching coalescing aggregates may not accurately reflect the “state” they were created under.

6.2.3 Extended characterization of protein-osmolyte interactions

Protein-osmolyte interactions are important from a biological and industrial perspective. However, current understanding of protein-water and protein-osmolyte interactions are limited. As mentioned in Chapter 4, direct measurement of protein-osmolyte interactions are challenging. However, characterizing these interactions may potentially allow one to predict protein chemical potential. Recently, hydrogen-deuterium exchange with mass spectroscopy has been used to investigate protein-osmolyte interactions. One study found that sucrose (which is thought to be preferentially excluded) decreased the H-D exchange rates of recombinant human granulocyte colony stimulating factor, while benzyl alcohol (thought to be preferentially accumulated) increase H-D exchange rates. These results suggest the conformational dynamics as probed by HD-MS may be sensitive to preferential interactions. Recently, terahertz absorption spectroscopy and low frequency Raman spectroscopy has also been used to investigate protein hydration and osmolyte preferential interactions. While these techniques show potential, current state-of-the-art of these technologies suggest these techniques may not yet have the required sensitivity to detect subtle changes in protein preferential interactions. However, future efforts may find these techniques useful.

BIBLIOGRAPHY

- [1] Nayoung Kim, Richard L. Remmele, Dingjiang Liu, Vladimir I. Razinkov, Erik J. Fernandez, and Christopher J. Roberts. Aggregation of anti-streptavidin immunoglobulin gamma1 involves Fab unfolding and competing growth pathways mediated by pH and salt concentration. *Biophysical Chemistry*, 172:26–36, February 2013.
- [2] Matthew Auton, D. Wayne Bolen, and Jrg Rsgen. Structural thermodynamics of protein preferential solvation: Osmolyte solvation of proteins, aminoacids, and peptides. *Proteins: Structure, Function, and Bioinformatics*, 73(4):802–813, May 2008.
- [3] D. B. Knowles, Irina A. Shkel, Noel M. Phan, Matt Sternke, Emily Lingeman, Xian Cheng, Lixue Cheng, Kevin OConnor, and M. Thomas Record. Chemical Interactions of Polyethylene Glycols (PEGs) and Glycerol with Protein Functional Groups: Applications to Effects of PEG and Glycerol on Protein Processes. *Biochemistry*, 54(22):3528–3542, June 2015.
- [4] L Ninni, M. S Camargo, and A. J. A Meirelles. Water activity in poly(ethylene glycol) aqueous solutions. *Thermochimica Acta*, 328(12):169–176, March 1999.
- [5] Alireza Salabat and Karamat Nasirzadeh. Measurement and prediction of water activity in PEG + (NH₄)₂so₄ + H₂o systems using polymer scaling laws. *Journal of Molecular Liquids*, 103104:349–358, March 2003.
- [6] M. V. Galmarini, J. Chirife, M. C. Zamora, and A. Prez. Determination and correlation of the water activity of unsaturated, supersaturated and saturated trehalose solutions. *LWT - Food Science and Technology*, 41(4):628–631, May 2008.
- [7] Saurabh Rob Aggarwal. Whats fueling the biotech engine2012 to 2013. *Nature*, 201:4, 2014.
- [8] Dawn M Ecker, Susan Dana Jones, and Howard L Levine. The therapeutic monoclonal antibody market. *mAbs*, 7(1):9–14, January 2015.
- [9] A. D. Nelson, M. M. Hoffmann, C. A. Parks, S. Dasari, A. G. Schrum, and D. Gil. IgG Fab Fragments Forming Bivalent Complexes by a Conformational Mechanism That Is Reversible by Osmolytes. *Journal of Biological Chemistry*, 287(51):42936–42950, December 2012.

- [10] Trevor T. Hansel, Harald Kropshofer, Thomas Singer, Jane A. Mitchell, and Andrew J. T. George. The safety and side effects of monoclonal antibodies. *Nature Reviews Drug Discovery*, 9(4):325–338, April 2010.
- [11] Vinochani Pillay, Hui K. Gan, and Andrew M. Scott. Antibodies in oncology. *New Biotechnology*, 28(5):518–529, September 2011.
- [12] Wei Wang, Satish K. Singh, Ning Li, Maria R. Toler, Kevin R. King, and Sandeep Nema. Immunogenicity of protein aggregates: Concerns and realities. *International Journal of Pharmaceutics*, 431(1-2):1–11, July 2012.
- [13] Huub Schellekens. Immunogenicity of therapeutic proteins: clinical implications and future prospects. *Clinical therapeutics*, 24(11):1720–1740, 2002.
- [14] Matthew P Baker, Helen M Reynolds, Brooke Lumicisi, and Christine J Bryson. Immunogenicity of protein therapeutics. *Self Nonself*, 1(4):314–322, 2010.
- [15] Amy S. Rosenberg. Effects of protein aggregates: an immunologic perspective. *The AAPS journal*, 8(3):E501–E507, 2006.
- [16] Andrea Hawe, Michael Wiggenhorn, Marco van de Weert, Joerg H. O. Garbe, Hanns-christian Mahler, and Wim Jiskoot. Forced degradation of therapeutic proteins. *Journal of Pharmaceutical Sciences*, 101(3):895–913, March 2012.
- [17] Rajsekhar Paul, Alexandra Graff-Meyer, Henning Stahlberg, Matthias E. Lauer, Arne C. Rufer, Hermann Beck, Alexandre Briguet, Volker Schnaible, Thomas Buckel, and Sabine Boeckle. Structure and Function of Purified Monoclonal Antibody Dimers Induced by Different Stress Conditions. *Pharmaceutical Research*, 29(8):2047–2059, April 2012.
- [18] M. K. Joubert, M. Hokom, C. Eakin, L. Zhou, M. Deshpande, M. P. Baker, T. J. Goletz, B. A. Kerwin, N. Chirmule, L. O. Narhi, and V. Jawa. Highly Aggregated Antibody Therapeutics Can Enhance the in Vitro Innate and Late-stage T-cell Immune Responses. *Journal of Biological Chemistry*, 287(30):25266–25279, July 2012.
- [19] Juliana Bessa, Sabine Boeckle, Hermann Beck, Thomas Buckel, Sonja Schlicht, Martin Ebeling, Anna Kiialainen, Atanas Koulov, Björn Boll, Thomas Weiser, Thomas Singer, Antonius G. Rolink, and Antonio Iglesias. The Immunogenicity of Antibody Aggregates in a Novel Transgenic Mouse Model. *Pharmaceutical Research*, 32(7):2344–2359, July 2015.
- [20] Mara Vázquez-Rey and Dietmar A. Lang. Aggregates in monoclonal antibody manufacturing processes. *Biotechnology and Bioengineering*, 108(7):1494–1508, July 2011.

- [21] F. Ulrich Hartl and Manajit Hayer-Hartl. Converging concepts of protein folding in vitro and in vivo. *Nature Structural & Molecular Biology*, 16(6):574–581, June 2009.
- [22] Tatiana Perevozchikova, Hirsh Nanda, Douglas P. Nesta, and Christopher J. Roberts. Protein Adsorption, Desorption, and Aggregation Mediated by Solid-Liquid Interfaces. *Journal of Pharmaceutical Sciences*, 104(6):1946–1959, June 2015.
- [23] Renuka Thirumangalathu, Sampathkumar Krishnan, Margaret Speed Ricci, David N. Brems, Theodore W. Randolph, and John F. Carpenter. Silicone oil- and agitation-induced aggregation of a monoclonal antibody in aqueous solution. *Journal of Pharmaceutical Sciences*, 98(9):3167–3181, September 2009.
- [24] Jared S. Bee, Theodore W. Randolph, John F. Carpenter, Steven M. Bishop, and Mariana N. Dimitrova. Effects of surfaces and leachables on the stability of biopharmaceuticals. *Journal of Pharmaceutical Sciences*, 100(10):4158–4170, October 2011.
- [25] Samiul Amin, Gregory V. Barnett, Jai A. Pathak, Christopher J. Roberts, and Prasad S. Sarangapani. Protein aggregation, particle formation, characterization & rheology. *Current Opinion in Colloid & Interface Science*, 19(5):438–449, October 2014.
- [26] Yi Li and Christopher J. Roberts. LumryEyring Nucleated-Polymerization Model of Protein Aggregation Kinetics. 2. Competing Growth via Condensation and Chain Polymerization. *The Journal of Physical Chemistry B*, 113(19):7020–7032, May 2009.
- [27] Yi Li, Babatunde A. Ogunnaike, and Christopher J. Roberts. Multi-variate approach to global protein aggregation behavior and kinetics: effects of pH, NaCl, and temperature for alpha-chymotrypsinogen A. *Journal of Pharmaceutical Sciences*, 99(2):645–662, February 2010.
- [28] Eva Y. Chi, Sampathkumar Krishnan, Brent S. Kendrick, Byeong S. Chang, John F. Carpenter, and Theodore W. Randolph. Roles of conformational stability and colloidal stability in the aggregation of recombinant human granulocyte colony-stimulating factor. *Protein Science*, 12(5):903–913, May 2003.
- [29] Jose M. Sanchez-Ruiz. Theoretical analysis of Lumry-Eyring models in differential scanning calorimetry. *Biophysical journal*, 61(4):921, 1992.
- [30] W. J. Becketl and J. A. Schellman. Protein stability curves. *Biopolymers*, 26(11):1859–1877, November 1987.

- [31] Rebecca K. Brummitt, Douglas P. Nesta, and Christopher J. Roberts. Predicting accelerated aggregation rates for monoclonal antibody formulations, and challenges for low-temperature predictions. *Journal of Pharmaceutical Sciences*, 100(10):4234–4243, October 2011.
- [32] US Department of Health, Human Services FDA Center for Drug, and others. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health and Quality of Life Outcomes*, 4:79, 2006.
- [33] Alexander Jarasch, Hans Koll, Joerg T. Regula, Martin Bader, Apollon Papadimitriou, and Hubert Kettenberger. Developability Assessment During the Selection of Novel Therapeutic Antibodies. *Journal of Pharmaceutical Sciences*, 104(6):1885–1898, June 2015.
- [34] Gregory V. Barnett, Wei Qi, Samiul Amin, E. Neil Lewis, and Christopher J. Roberts. Aggregate structure, morphology and the effect of aggregation mechanisms on viscosity at elevated protein concentrations. *Biophysical Chemistry*, 207:21–29, July 2015.
- [35] Donald Allan McQuarrie. *Statistical Mechanics*. University Science Books, 2000.
- [36] Rebecca K. Brummitt, Douglas P. Nesta, Liuquan Chang, Susan F. Chase, Thomas M. Laue, and Christopher J. Roberts. Nonnative aggregation of an IgG1 antibody in acidic conditions: Part 1. Unfolding, colloidal interactions, and formation of high-molecular-weight aggregates. *Journal of Pharmaceutical Sciences*, 100(6):2087–2103, June 2011.
- [37] Andrew M. Kroetsch, Erinc Sahin, Hsiang-yun Wang, Sintia Krizman, and Christopher J. Roberts. Relating particle formation to salt- and pH-dependent phase separation of non-native aggregates of alpha-chymotrypsinogen a. *Journal of Pharmaceutical Sciences*, 101(10):3651–3660, October 2012.
- [38] William F. Weiss, Teresa M. Young, and Christopher J. Roberts. Principles, approaches, and challenges for predicting protein aggregation rates and shelf life. *Journal of Pharmaceutical Sciences*, 98(4):1246–1277, April 2009.
- [39] Jeremy D. Schmit. Kinetic theory of amyloid fibril templating. *The Journal of Chemical Physics*, 138(18):185102, 2013.
- [40] Jeremy D. Schmit, Kingshuk Ghosh, and Ken Dill. What Drives Amyloid Molecules To Assemble into Oligomers and Fibrils? *Biophysical Journal*, 100(2):450–458, January 2011.
- [41] Aming Zhang, Jacob L. Jordan, Magdalena I. Ivanova, William F. Weiss, Christopher J. Roberts, and Erik J. Fernandez. Molecular Level Insights into Thermally

- Induced α -Chymotrypsinogen A Amyloid Aggregation Mechanism and Semiflexible Protofibril Morphology. *Biochemistry*, 49(49):10553–10564, December 2010.
- [42] Jennifer M. Andrews and Christopher J. Roberts. Non-Native Aggregation of α -Chymotrypsinogen Occurs through Nucleation and Growth with Competing Nucleus Sizes and Negative Activation Energies. *Biochemistry*, 46(25):7558–7571, June 2007.
- [43] Rebecca K. Brummitt, Douglas P. Nesta, Liuquan Chang, Andrew M. Kroetsch, and Christopher J. Roberts. Nonnative aggregation of an IgG1 antibody in acidic conditions, part 2: Nucleation and growth kinetics with competing growth mechanisms. *Journal of Pharmaceutical Sciences*, 100(6):2104–2119, June 2011.
- [44] Erinc Sahin, William F. Weiss, Andrew M. Kroetsch, Kevin R. King, R. Kendall Kessler, Tapan K. Das, and Christopher J. Roberts. Aggregation and pH-temperature phase behavior for aggregates of an IgG2 antibody. *Journal of Pharmaceutical Sciences*, 101(5):1678–1687, May 2012.
- [45] William F. Weiss, Travis K. Hodgdon, Eric W. Kaler, Abraham M. Lenhoff, and Christopher J. Roberts. Nonnative Protein Polymers: Structure, Morphology, and Relation to Nucleation and Growth. *Biophysical Journal*, 93(12):4392–4403, December 2007.
- [46] Marco Lattuada, Peter Sandkhler, Hua Wu, Jan Sefcik, and Massimo Morbidelli. Aggregation kinetics of polymer colloids in reaction limited regime: experiments and simulations. *Advances in colloid and interface science*, 103(1):33–56, 2003.
- [47] Paolo Arosio, Simonetta Rima, Marco Lattuada, and Massimo Morbidelli. Population Balance Modeling of Antibodies Aggregation Kinetics. *The Journal of Physical Chemistry B*, 116(24):7066–7075, June 2012.
- [48] Lucrce Nicoud, Paolo Arosio, Margaux Sozo, Andrew Yates, Edith Norrant, and Massimo Morbidelli. Kinetic Analysis of the Multistep Aggregation Mechanism of Monoclonal Antibodies. *The Journal of Physical Chemistry B*, 118(36):10595–10606, September 2014.
- [49] W. B. Russel, D. A. Saville, and W. R. Schowalter. *Colloidal Dispersions*. Cambridge University Press, March 1992.
- [50] A. Moncho-Jord, G. Odriozola, F. Martnez-Lpez, A. Schmitt, and R. Hidalgo-Ivarez. The DLCA-RLCA transition arising in 2d-aggregation: simulations and mean field theory. *The European Physical Journal E*, 5(4):471–480, 2001.
- [51] S. Tang, J. M. Preece, C. M. McFarlane, and Z. Zhang. Fractal morphology and breakage of DLCA and RLCA aggregates. *Journal of colloid and interface science*, 221(1):114–123, 2000.

- [52] Annette Eva Langkilde and Bente Vestergaard. Methods for structural characterization of prefibrillar intermediates and amyloid fibrils. *FEBS Letters*, 583(16):2600–2609, August 2009.
- [53] Jennifer M. Andrews and Christopher J. Roberts. A LumryEyring Nucleated Polymerization Model of Protein Aggregation Kinetics: 1. Aggregation with Pre-Equilibrated Unfolding. *The Journal of Physical Chemistry B*, 111(27):7897–7913, July 2007.
- [54] Mnica Rosa, Carlos Lopes, Eduardo P. Melo, Satish K. Singh, Vitor Geraldés, and Miguel A. Rodrigues. Measuring and Modeling Hemoglobin Aggregation below the Freezing Temperature. *The Journal of Physical Chemistry B*, 117(30):8939–8946, August 2013.
- [55] Mats H. M. Olsson, Chresten R. Sndergaard, Michal Rostkowski, and Jan H. Jensen. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *Journal of Chemical Theory and Computation*, 7(2):525–537, February 2011.
- [56] Jennifer A. Durant, Chuanying Chen, Thomas M. Laue, Thomas P. Moody, and Stuart A. Allison. Use of T4 lysozyme charge mutants to examine electrophoretic models. *Biophysical chemistry*, 101:593–609, 2002.
- [57] Sandeep Yadav, Steven J. Shire, and Devendra S. Kalonia. Viscosity behavior of high-concentration monoclonal antibody solutions: Correlation with interaction parameter and electroviscous effects. *Journal of Pharmaceutical Sciences*, 101(3):998–1011, March 2012.
- [58] Wei Wang, Satish Singh, David L. Zeng, Kevin King, and Sandeep Nema. Antibody structure, instability, and formulation. *Journal of Pharmaceutical Sciences*, 96(1):1–26, January 2007.
- [59] Kim D. Collins. Ion hydration: Implications for cellular function, polyelectrolytes, and protein crystallization. *Biophysical Chemistry*, 119(3):271–281, February 2006.
- [60] Robert L. Baldwin. How Hofmeister ion interactions affect protein stability. *Biophysical Journal*, 71(4):2056–2063, 1996.
- [61] Xavier Tadeo, Miquel Pons, and Oscar Millet. Influence of the Hofmeister Anions on Protein Stability As Studied by Thermal Denaturation and Chemical Shift Perturbation . *Biochemistry*, 46(3):917–923, January 2007.
- [62] Jian Zhang-van Enk, Bruce D. Mason, Lei Yu, Le Zhang, Wael Hamouda, Gang Huang, Dingjiang Liu, Richard L. Remmele, and Jifeng Zhang. Perturbation of Thermal Unfolding and Aggregation of Human IgG1 Fc Fragment by Hofmeister Anions. *Molecular Pharmaceutics*, 10(2):619–630, February 2013.

- [63] Peter J. Marek, Vadim Patsalo, David F. Green, and Daniel P. Raleigh. Ionic Strength Effects on Amyloid Formation by Amylin Are a Complicated Interplay among Debye Screening, Ion Selectivity, and Hofmeister Effects. *Biochemistry*, 51(43):8478–8490, October 2012.
- [64] Paolo Arosio, Baptiste Jaquet, Hua Wu, and Massimo Morbidelli. On the role of salt type and concentration on the stability behavior of a monoclonal antibody solution. *Biophysical Chemistry*, 168-169:19–27, July 2012.
- [65] Yatin R. Gokarn, R. Matthew Fesinmeyer, Atul Saluja, Vladimir Razinkov, Susan F. Chase, Thomas M. Laue, and David N. Brems. Effective charge measurements reveal selective and preferential accumulation of anions, but not cations, at the protein surface in dilute salt solutions: Effective Charge Measurements Reveal Direct Anion-Protein Interactions. *Protein Science*, 20(3):580–587, March 2011.
- [66] D. Roberts, R. Keeling, M. Tracka, C. F. van der Walle, S. Uddin, J. Warwicker, and R. Curtis. Specific Ion and Buffer Effects on Protein-Protein Interactions of a Monoclonal Antibody. *Molecular Pharmaceutics*, 12(1):179–193, January 2015.
- [67] Jason K. Cheung. Insights Into Crowding Effects on Protein Stability From a Coarse-Grained Model. *Journal of Biomechanical Engineering*, 131(7):071002, June 2009.
- [68] Vincent K. Shen, Jason K. Cheung, Jeffrey R. Errington, and Thomas M. Truskett. Coarse-Grained Strategy for Modeling Protein Stability in Concentrated Solutions. II: Phase Behavior. *Biophysical Journal*, 90(6):1949–1960, March 2006.
- [69] Onofrio Annunziata, Neer Asherie, Aleksey Lomakin, Jayanti Pande, Olutayo Ogun, and George B. Benedek. Effect of polyethylene glycol on the liquid-liquid phase transition in aqueous protein solutions. *Proceedings of the National Academy of Sciences*, 99(22):14165–14170, 2002.
- [70] Michele D'Amico, Samuele Raccosta, Marco Cannas, Vincenzo Martorana, and Mauro Manno. Existence of Metastable Intermediate Lysozyme Conformation Highlights the Role of Alcohols in Altering Protein Stability. *The Journal of Physical Chemistry B*, 115(14):4078–4087, April 2011.
- [71] George Scatchard, W. J. Hamer, and S. E. Wood. Isotonic Solutions. I. The Chemical Potential of Water in Aqueous Solutions of Sodium Chloride, Potassium Chloride, Sulfuric Acid, Sucrose, Urea and Glycerol at 251. *Journal of the American Chemical Society*, 60(12):3061–3070, December 1938.
- [72] Seishi Shimizu. The effect of urea on hydrophobic hydration: Preferential interaction and the enthalpy of transfer. *Chemical Physics Letters*, 517(1-3):76–79, November 2011.

- [73] Beate Moeser and Dominik Horinek. Unified Description of Urea Denaturation: Backbone and Side Chains Contribute Equally in the Transfer Model. *The Journal of Physical Chemistry B*, 118(1):107–114, January 2014.
- [74] J. C. Lee and S. N. Timasheff. The stabilization of proteins by sucrose. *Journal of Biological Chemistry*, 256(14):7193–7201, July 1981.
- [75] Serge N. Timasheff. Protein Hydration, Thermodynamic Binding, and Preferential Hydration. *Biochemistry*, 41(46):13473–13482, November 2002.
- [76] Matthew Auton and D. Wayne Bolen. Predicting the energetics of osmolyte-induced protein folding/unfolding. *Proceedings of the National Academy of Sciences of the United States of America*, 102(42):15065–15068, 2005.
- [77] Emanuel Schneck, Dominik Horinek, and Roland R. Netz. Insight into the Molecular Mechanisms of Protein Stabilizing Osmolytes from Global Force-Field Variations. *The Journal of Physical Chemistry B*, 117(28):8310–8321, July 2013.
- [78] Nishant Kumar Jain and Ipsita Roy. Effect of trehalose on protein structure. *Protein Science : A Publication of the Protein Society*, 18(1):24–36, January 2009.
- [79] T. Arakawa and S. N. Timasheff. Stabilization of protein structure by sugars. *Biochemistry*, 21(25):6536–6544, December 1982.
- [80] Douglas D. Banks, Ramil F. Latypov, Randal R. Ketchem, Jon Woodard, Joanna L. Scavezze, Christine C. Siska, and Vladimir I. Razinkov. Native-state solubility and transfer free energy as predictive tools for selecting excipients to include in protein formulation development studies. *Journal of Pharmaceutical Sciences*, 101(8):2720–2732, August 2012.
- [81] Donald H. Atha and Kenneth C. Ingham. Mechanism of precipitation of proteins by polyethylene glycols. Analysis in terms of excluded volume. *Journal of Biological Chemistry*, 256(23):12108–12117, 1981.
- [82] T. Arakawa and S. N. Timasheff. Preferential interactions of proteins with salts in concentrated solutions. *Biochemistry*, 21(25):6545–6552, December 1982.
- [83] T. Arakawa and S. N. Timasheff. Mechanism of poly(ethylene glycol) interaction with proteins. *Biochemistry*, 24(24):6756–6762, November 1985.
- [84] Serge N. Timasheff. The control of protein stability and association by weak interactions with water: how do solvents affect these processes? *Annual review of biophysics and biomolecular structure*, 22(1):67–97, 1993.
- [85] R. Bhat and S. N. Timasheff. Steric exclusion is the principal source of the preferential hydration of proteins in the presence of polyethylene glycols. *Protein Science: A Publication of the Protein Society*, 1(9):1133–1143, September 1992.

- [86] John A. Schellman. Protein Stability in Mixed Solvents: A Balance of Contact Interaction and Excluded Volume. *Biophysical Journal*, 85(1):108–125, July 2003.
- [87] John A. Schellman. Selective binding and solvent denaturation. *Biopolymers*, 26(4):549–559, April 1987.
- [88] J A Schellman. Temperature, stability, and the hydrophobic interaction. *Biophysical Journal*, 73(6):2960–2964, December 1997.
- [89] J. A. Schellman. The thermodynamics of solvent exchange. *Biopolymers*, 34(8):1015–1026, August 1994.
- [90] Liuquan Lucy Chang, Deanna Shepherd, Joanna Sun, David Ouellette, Kathleen L. Grant, Xiaolin Charlie Tang, and Michael J. Pikal. Mechanism of protein stabilization by sugars during freeze-drying and storage: native structure preservation, specific interaction, and/or immobilization in a glassy matrix? *Journal of Pharmaceutical Sciences*, 94(7):1427–1444, July 2005.
- [91] Peter Strtzel, Henning Gieseler, Margit Gieseler, Ahmad M. Abdul-Fattah, Michael Adler, Hanns-Christian Mahler, and Pierre Goldbach. Freeze Drying of l-Arginine/Sucrose-Based Protein Formulations, Part I: Influence of Formulation and Arginine Counter Ion on the Critical Formulation Temperature, Product Performance and Protein Stability. *Journal of Pharmaceutical Sciences*, pages n/a–n/a, May 2015.
- [92] Alain Hedoux, Laurent Paccou, Samira Achir, and Yannick Guinet. Mechanism of protein stabilization by trehalose during freeze-drying analyzed by *in situ* micro-Raman spectroscopy. *Journal of Pharmaceutical Sciences*, 102(8):2484–2494, August 2013.
- [93] James D. Andya, Chung C. Hsu, and Steven J. Shire. Mechanisms of aggregate formation and carbohydrate excipient stabilization of lyophilized humanized monoclonal antibody formulations. *Aaps PharmSci*, 5(2):21–31, 2003.
- [94] Aming Zhang, Satish K. Singh, Michael R. Shirts, Sandeep Kumar, and Erik J. Fernandez. Distinct Aggregation Mechanisms of Monoclonal Antibody Under Thermal and Freeze-Thaw Stresses Revealed by Hydrogen Exchange. *Pharmaceutical Research*, 29(1):236–250, January 2012.
- [95] Meera Agarkhed, Courtney ODell, Ming-Ching Hsieh, Jingming Zhang, Joel Goldstein, and Arvind Srivastava. Effect of Polysorbate 80 Concentration on Thermal and Photostability of a Monoclonal Antibody. *AAPS PharmSciTech*, 14(1):1–9, March 2013.
- [96] Curtiss P. Schneider and Bernhardt L. Trout. Investigation of CosoluteProtein Preferential Interaction Coefficients: New Insight into the Mechanism by

- Which Arginine Inhibits Aggregation. *The Journal of Physical Chemistry B*, 113(7):2050–2058, February 2009.
- [97] Curtiss P. Schneider, Diwakar Shukla, and Bernhardt L. Trout. Arginine and the Hofmeister Series: The Role of Ion-Ion Interactions in Protein Aggregation Suppression. *The Journal of Physical Chemistry B*, 115(22):7447–7458, June 2011.
- [98] Wei Wang. Advanced protein formulations. *Protein Science*, pages n/a–n/a, April 2015.
- [99] Tim J. Kamerzell, Reza Esfandiary, Sangeeta B. Joshi, C. Russell Middaugh, and David B. Volkin. Protein-excipient interactions: Mechanisms and biophysical characterization applied to protein formulation development. *Advanced Drug Delivery Reviews*, 63(13):1118–1159, October 2011.
- [100] Feng He, Gerald W. Becker, Jennifer R. Litowski, Linda O. Narhi, David N. Brems, and Vladimir I. Razinkov. High-throughput dynamic light scattering method for measuring viscosity of concentrated protein solutions. *Analytical Biochemistry*, 399(1):141–143, April 2010.
- [101] Vladimir I. Razinkov, Michael J. Treuheit, and Gerald W. Becker. Methods of high throughput biophysical characterization in biopharmaceutical development. *Current drug discovery technologies*, 10(1):59–70, 2013.
- [102] Yi Li, Henryk Mach, and Jeffrey T. Blue. High throughput formulation screening for global aggregation behaviors of three monoclonal antibodies. *Journal of Pharmaceutical Sciences*, 100(6):2120–2135, June 2011.
- [103] Erinc Sahin, Jacob L. Jordan, Michelle L. Spataro, Andrea Naranjo, Joseph A. Costanzo, William F. Weiss, Anne Skaja Robinson, Erik J. Fernandez, and Christopher J. Roberts. Computational Design and Biophysical Characterization of Aggregation-Resistant Point Mutations for D Crystallin Illustrate a Balance of Conformational Stability and Intrinsic Aggregation Propensity. *Biochemistry*, 50(5):628–639, February 2011.
- [104] Amit Shtainfeld, Tania Sheynis, and Raz Jelinek. Specific Mutations Alter Fibrillation Kinetics, Fiber Morphologies, and Membrane Interactions of Pentapeptides Derived from Human Calcitonin. *Biochemistry*, 49(25):5299–5307, June 2010.
- [105] Wei Wang and Christopher J. Roberts. Non-Arrhenius protein aggregation. *The AAPS journal*, 15(3):840–851, 2013.
- [106] P. L. Privalov. Thermodynamic problems of protein structure. *Annual review of biophysics and biophysical chemistry*, 18(1):47–69, 1989.

- [107] Christopher J. Roberts, Douglas P. Nesta, and Nayoung Kim. Effects of Temperature and Osmolytes on Competing Degradation Routes for an IgG1 Antibody: Competing Degradation Routes for an IgG1 Antibody. *Journal of Pharmaceutical Sciences*, 102(10):3556–3566, October 2013.
- [108] Benny KC Lo. Antibody humanization by CDR grafting. In *Antibody Engineering*, pages 135–159. Springer, 2004.
- [109] Nicholas J. Clark, Hailiang Zhang, Susan Krueger, Hyo Jin Lee, Randal R. Ketchem, Bruce Kerwin, Sekhar R. Kanapuram, Michael J. Treuheit, Arnold McAuley, and Joseph E. Curtis. Small-Angle Neutron Scattering Study of a Monoclonal Antibody Using Free-Energy Constraints. *The Journal of Physical Chemistry B*, 117(45):14029–14038, November 2013.
- [110] Xiaoling Wang, Sandeep Kumar, Patrick M. Buck, and Satish K. Singh. Impact of deglycosylation and thermal stress on conformational stability of a full length murine igG2a monoclonal antibody: Observations from molecular dynamics simulations. *Proteins: Structure, Function, and Bioinformatics*, 81(3):443–460, March 2013.
- [111] Liming Liu. Antibody Glycosylation and Its Impact on the Pharmacokinetics and Pharmacodynamics of Monoclonal Antibodies and Fc-Fusion Proteins. *Journal of Pharmaceutical Sciences*, 104(6):1866–1884, June 2015.
- [112] Sanjay B. Hari, Hollis Lau, Vladimir I. Razinkov, Shuang Chen, and Ramil F. Latypov. Acid-Induced Aggregation of Human Monoclonal IgG1 and IgG2: Molecular Mechanism and the Effect of Solution Composition. *Biochemistry*, 49(43):9328–9338, November 2010.
- [113] Heather Franey, Stephen R. Brych, Carl G. Kolvenbach, and Rahul S. Rajan. Increased aggregation propensity of IgG2 subclass over IgG1: Role of conformational changes and covalent character in isolated aggregates: Characterization of Aggregates in IgG1 and IgG2 Subclasses. *Protein Science*, 19(9):1601–1615, September 2010.
- [114] Meredith E. Wiseman and Curtis W. Frank. Antibody adsorption and orientation on hydrophobic surfaces. *Langmuir: the ACS journal of surfaces and colloids*, 28(3):1765–1774, January 2012.
- [115] Lisa J. Lapidus. Understanding protein aggregation from the view of monomer dynamics. *Mol. BioSyst.*, 9(1):29–35, 2013.
- [116] Jonathan Rubin, Lars Linden, Wayne M. Coco, Andreas S. Bommarius, and Sven H. Behrens. Salt-induced aggregation of a monoclonal human immunoglobulin G1. *Journal of Pharmaceutical Sciences*, 102(2):377–386, February 2013.

- [117] Gregory V. Barnett, Vladimir I. Razinkov, Bruce A. Kerwin, Thomas M. Laue, Andrea H. Woodka, Paul D. Butler, Tatiana Perevozchikova, and Christopher J. Roberts. Specific-Ion Effects on the Aggregation Mechanisms and Protein-Protein Interactions for Anti-streptavidin Immunoglobulin Gamma-1. *The Journal of Physical Chemistry B*, 119(18):5793–5804, May 2015.
- [118] Erinc Sahin and Christopher J. Roberts. Size-Exclusion Chromatography with Multi-angle Light Scattering for Elucidating Protein Aggregation Mechanisms. In Vladimir Voynov and Justin A. Caravella, editors, *Therapeutic Proteins*, volume 899, pages 403–423. Humana Press, Totowa, NJ, 2012.
- [119] Th Zemb and P. Lindner, editors. *Neutron, X-rays and Light. Scattering Methods Applied to Soft Condensed Matter*. North Holland, Amsterdam ; Boston, 1 edition edition, November 2002.
- [120] Eric J. Yearley, Isidro E. Zarraga, Steven J. Shire, Thomas M. Scherer, Yatin Gokarn, Norman J. Wagner, and Yun Liu. Small-Angle Neutron Scattering Characterization of Monoclonal Antibody Conformations and Interactions at High Concentrations. *Biophysical Journal*, 105(3):720–731, August 2013.
- [121] Wyatt Technology Corporation. *ASTRA VI User's Guide*. Santana Barbara, CA, 2012.
- [122] Marco A. Blanco, Erinc Sahin, Yi Li, and Christopher J. Roberts. Reexamining protein-protein and protein-solvent interactions from Kirkwood-Buff analysis of light scattering in multi-component solutions. *The Journal of Chemical Physics*, 134(22):225103, 2011.
- [123] S. S. Nielsen, K. Noergaard Toft, D. Snakenborg, M. G. Jeppesen, J. K. Jacobsen, B. Vestergaard, J. P. Kutter, and L. Arleth. *BioXTAS RAW*, a software program for high-throughput automated small-angle X-ray scattering data reduction and preliminary analysis. *Journal of Applied Crystallography*, 42(5):959–964, October 2009.
- [124] Barbara J. Frisken. Revisiting the method of cumulants for the analysis of dynamic light-scattering data. *Applied Optics*, 40(24):4087–4091, 2001.
- [125] Dennis E. Koppel. Analysis of Macromolecular Polydispersity in Intensity Correlation Spectroscopy: The Method of Cumulants. *The Journal of Chemical Physics*, 57(11):4814–4820, December 1972.
- [126] Jason C. W. Corbett, Malcolm T. Connah, and Kevin Mattison. Advances in the measurement of protein mobility using laser Doppler electrophoresis - the diffusion barrier technique. *Electrophoresis*, 32(14):1787–1794, July 2011.

- [127] Jan Gimsa, Peter Eppmann, and B. Prger. Introducing phase analysis light scattering for dielectric characterization: measurement of traveling-wave pumping. *Biophysical journal*, 73(6):3309–3316, 1997.
- [128] Klaus Schtzel John F Miller. The determination of very small electrophoretic mobilities in polar and nonpolar colloidal dispersions using phase analysis light scattering. *Journal of Colloid and Interface Science*, (2):532–554, 1991.
- [129] Yatin R. Gokarn, Matthew McLean, and Thomas M. Laue. Effect of PEGylation on Protein Hydrodynamics. *Molecular Pharmaceutics*, 9(4):762–773, April 2012.
- [130] Steven R. Kline. Reduction and analysis of SANS and USANS data using IGOR Pro. *Journal of Applied Crystallography*, 39(6):895–900, December 2006.
- [131] Arthur K. Covington, Maya Paabo, Robert Anthony Robinson, and Roger G. Bates. Use of the glass electrode in deuterium oxide and the relation between the standardized pD (paD) scale and the operational pH in heavy water. *Analytical Chemistry*, 40(4):700–706, April 1968.
- [132] Christopher J. Roberts, Tapan K. Das, and Erinc Sahin. Predicting solution aggregation rates for therapeutic proteins: Approaches and challenges. *International Journal of Pharmaceutics*, 418(2):318–333, October 2011.
- [133] Paolo Arosio, Simonetta Rima, and Massimo Morbidelli. Aggregation mechanism of an IgG2 and two IgG1 monoclonal antibodies at low pH: from oligomers to larger aggregates. *Pharmaceutical Research*, 30(3):641–654, March 2013.
- [134] Jens Feder, Torstein Jssang, and Einar Rosenqvist. Scaling Behavior and Cluster Fractal Dimension Determined by Light Scattering from Aggregating Proteins. *Physical Review Letters*, 53(15):1403–1406, October 1984.
- [135] Walther Burchard. Solution properties of branched macromolecules. In *Branched polymers II*, pages 113–194. Springer, 1999.
- [136] W. Burchard, M. Schmidt, and W. H. Stockmayer. Information on polydispersity and branching from combined quasi-elastic and intergrated scattering. *Macromolecules*, 13(5):1265–1272, 1980.
- [137] Vronique Receveur-Brchot and Dominique Durand. How Random are Intrinsically Disordered Proteins? A Small Angle Scattering Perspective. *Current Protein & Peptide Science*, 13(1):55–75, February 2012.
- [138] Xinsheng Tian, Annette E. Langkilde, Matthias Thorolfsson, Hanne B. Rasmussen, and Bente Vestergaard. Small-Angle X-ray Scattering Screening Complements Conventional Biophysical Analysis: Comparative Structural and Biophysical Analysis of Monoclonal Antibodies IgG1, IgG2, and IgG4. *Journal of Pharmaceutical Sciences*, 103(6):1701–1710, June 2014.

- [139] Alexander Grnberger, Pin-Kuang Lai, Marco A. Blanco, and Christopher J. Roberts. Coarse-Grained Modeling of Protein Second Osmotic Virial Coefficients: Sterics and Short-Ranged Attractions. *The Journal of Physical Chemistry B*, 117(3):763–770, January 2013.
- [140] Vincent S. Stoll and John S. Blanchard. Chapter 6 Buffers. In *Methods in Enzymology*, volume 463, pages 43–56. Elsevier, 2009.
- [141] John B. Hayter and Jeff Penfold. An analytic structure factor for macroion solutions. *Molecular Physics*, 42(1):109–118, January 1981.
- [142] Joseph M. Perchiacca, Moumita Bhattacharya, and Peter M. Tessier. Mutational analysis of domain antibodies reveals aggregation hotspots within and near the complementarity determining regions. *Proteins*, 79(9):2637–2647, September 2011.
- [143] Alexander K. Buell, Anne Dhulesia, Duncan A. White, Tuomas P. J. Knowles, Christopher M. Dobson, and Mark E. Welland. Detailed Analysis of the Energy Barriers for Amyloid Fibril Growth. *Angewandte Chemie International Edition*, 51(21):5247–5251, May 2012.
- [144] T. A. Witten and L. M. Sander. Diffusion-Limited Aggregation, a Kinetic Critical Phenomenon. *Physical Review Letters*, 47(19):1400–1403, November 1981.
- [145] null Weitz, null Huang, null Lin, and null Sung. Limits of the fractal dimension for irreversible kinetic aggregation of gold colloids. *Physical Review Letters*, 54(13):1416–1419, April 1985.
- [146] J. Teixeira. Small-angle scattering by fractal systems. *Journal of Applied Crystallography*, 21(6):781–785, 1988.
- [147] Marco Lattuada, Hua Wu, and Massimo Morbidelli. Hydrodynamic radius of fractal clusters. *Journal of Colloid and Interface Science*, 268(1):96–105, December 2003.
- [148] Marco A. Blanco, Tatiana Perevozchikova, Vincenzo Martorana, Mauro Manno, and Christopher J. Roberts. Protein-Protein Interactions in Dilute to Concentrated Solutions: -Chymotrypsinogen in Acidic Conditions. *The Journal of Physical Chemistry B*, 118(22):5817–5831, June 2014.
- [149] Deborah S. Goldberg, Steven M. Bishop, Ambarish U. Shah, and Hasige A Sathish. Formulation development of therapeutic monoclonal antibodies using high-throughput fluorescence and static light scattering techniques: Role of conformational and colloidal stability. *Journal of Pharmaceutical Sciences*, 100(4):1306–1315, April 2011.

- [150] Weiqiang Cheng, Sangeeta B. Joshi, Feng He, David N. Brems, Bing He, Bruce A. Kerwin, David B. Volkin, and C. Russell Middaugh. Comparison of high-throughput biophysical methods to identify stabilizing excipients for a model IgG2 monoclonal antibody: Conformational stability and kinetic aggregation measurements. *Journal of Pharmaceutical Sciences*, 101(5):1701–1720, May 2012.
- [151] Akhilesh Bhambhani, Julian M. Kissmann, Sangeeta B. Joshi, David B. Volkin, Ramesh S. Kashi, and C. Russell Middaugh. Formulation design and high-throughput excipient selection based on structural integrity and conformational stability of dilute and highly concentrated IgG1 monoclonal antibody solutions. *Journal of Pharmaceutical Sciences*, 101(3):1120–1135, March 2012.
- [152] Elisabeth Ablinger, Stefan Leitgeb, and Andreas Zimmer. Differential scanning fluorescence approach using a fluorescent molecular rotor to detect thermostability of proteins in surfactant-containing formulations. *International Journal of Pharmaceutics*, 441(1-2):255–260, January 2013.
- [153] Tim Menzen and Wolfgang Friess. Temperature-Ramped Studies on the Aggregation, Unfolding, and Interaction of a Therapeutic Monoclonal Antibody. *Journal of Pharmaceutical Sciences*, 103(2):445–455, February 2014.
- [154] Michael F. Drenski, Mark L. Brader, Roy W. Alston, and Wayne F. Reed. Monitoring protein aggregation kinetics with simultaneous multiple sample light scattering. *Analytical Biochemistry*, 437(2):185–197, June 2013.
- [155] Josef Vlasak and Roxana Ionescu. Fragmentation of monoclonal antibodies. *mAbs*, 3(3):253–263, May 2011.
- [156] Li Yi, Nia Beckley, Benson Gikanga, Jennifer Zhang, Y. John Wang, Hung-Wei Chih, and Vikas K. Sharma. Isomerization of Asp-Asp motif in model peptides and a Monoclonal Antibody Fab Fragment. *Journal of Pharmaceutical Sciences*, 102(3):947–959, March 2013.
- [157] Gregory V. Barnett, Vladimir I. Razinkov, Bruce A. Kerwin, Alexander Hillsley, and Christopher J. Roberts. Acetate and Citrate Specific-Ion-Effects on Unfolding and Temperature-Dependent Aggregation Rates of Anti-Streptavidin IgG1. *Journal of Pharmaceutical Sciences*, (submitted), 2015.
- [158] Ramil F. Latypov, Sabine Hogan, Hollis Lau, Himanshu Gadgil, and Dingjiang Liu. Elucidation of acid-induced unfolding and aggregation of human immunoglobulin IgG1 and IgG2 Fc. *Journal of Biological Chemistry*, 287(2):1381–1396, 2012.
- [159] Bruce D. Mason, Christian Schneich, and Bruce A. Kerwin. Effect of pH and Light on Aggregation and Conformation of an IgG1 mAb. *Molecular Pharmaceutics*, 9(4):774–790, April 2012.

- [160] Keith J. Laidler. *Chemical Kinetics*. Prentice Hall, New York, 3 edition edition, January 1987.
- [161] Haixia Wu, Rachel Kroe-Barrett, Sanjaya Singh, Anne S. Robinson, and Christopher J. Roberts. Competing aggregation pathways for monoclonal antibodies. *FEBS Letters*, 588(6):936–941, March 2014.
- [162] James B. Rawlings and John G. Ekerdt. *Chemical Reactor Analysis and Design Fundamentals*. Nob Hill Pub, Madison, Wisconsin, 2 edition edition, November 2013.
- [163] Roxana M. Ionescu, Josef Vlasak, Colleen Price, and Marc Kirchmeier. Contribution of variable domains to the stability of humanized IgG1 monoclonal antibodies. *Journal of Pharmaceutical Sciences*, 97(4):1414–1426, April 2008.
- [164] Tingting Wang, Ozan S. Kumru, Li Yi, Y. John Wang, Jennifer Zhang, Jae Hyun Kim, Sangeeta B. Joshi, C. Russell Middaugh, and David B. Volkin. Effect of ionic strength and pH on the physical and chemical stability of a monoclonal antibody antigen-binding fragment. *Journal of Pharmaceutical Sciences*, 102(8):2520–2537, August 2013.
- [165] Veysel Kayser, Naresh Chennamsetty, Vladimir Voynov, Bernhard Helk, Kurt Forrer, and Bernhardt L. Trout. Evaluation of a non-arrhenius model for therapeutic monoclonal antibody aggregation. *Journal of Pharmaceutical Sciences*, 100(7):2526–2542, July 2011.
- [166] Amanda L. Pace, Rita L. Wong, Yonghua Taylor Zhang, Yung-Hsiang Kao, and Y. John Wang. Asparagine deamidation dependence on buffer type, pH, and temperature. *Journal of Pharmaceutical Sciences*, 102(6):1712–1723, June 2013.
- [167] Bei Li, Eric M. Gorman, Kimberly D. Moore, Todd Williams, Richard L. Schowen, Elizabeth M. Topp, and Ronald T. Borchardt. Effects of acidic N+1 residues on asparagine deamidation rates in solution and in the solid state. *Journal of Pharmaceutical Sciences*, 94(3):666–675, March 2005.
- [168] Aditya A. Wakankar and Ronald T. Borchardt. Formulation considerations for proteins susceptible to asparagine deamidation and aspartate isomerization. *Journal of Pharmaceutical Sciences*, 95(11):2321–2336, November 2006.
- [169] Christopher J. Roberts. Kinetics of Irreversible Protein Aggregation: Analysis of Extended LumryEyring Models and Implications for Predicting Protein Shelf Life. *The Journal of Physical Chemistry B*, 107(5):1194–1207, February 2003.
- [170] Christopher J. Roberts, Richard T. Darrington, and Maureen B. Whitley. Irreversible aggregation of recombinant bovine granulocyte-colony stimulating factor (bG-CSF) and implications for predicting protein shelf life. *Journal of pharmaceutical sciences*, 92(5):1095–1111, 2003.

- [171] Wei Wang and Christopher J. Roberts. *Aggregation of Therapeutic Proteins*. John Wiley & Sons, December 2010.
- [172] Gregory V. Barnett, Vladimir I. Razinkov, Bruce A. Kerwin, Steven Blake, QI Wei, Robin A. Curtis, and Christopher J. Roberts. Osmolyte Effects on Monoclonal Antibody Stability and Concentration-Dependent Protein Interactions with Water and Common Osmolytes. *The Journal of Physical Chemistry B*, (submitted), 2015.
- [173] R. John Ellis. Macromolecular crowding: an important but neglected aspect of the intracellular environment. *Current Opinion in Structural Biology*, 11(1):114–119, February 2001.
- [174] S. B. Zimmerman and S. O. Trach. Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of *Escherichia coli*. *Journal of Molecular Biology*, 222(3):599–620, December 1991.
- [175] Thomas M. Scherer. Cosolute Effects on the Chemical Potential and Interactions of an IgG1 Monoclonal Antibody at High Concentrations. *The Journal of Physical Chemistry B*, 117(8):2254–2266, February 2013.
- [176] Arieh Ben-Naim. *Statistical Thermodynamics for Chemists and Biochemists*. Springer, S.I., softcover reprint of the original 1st ed. 1992 edition edition, October 2013.
- [177] John G. Kirkwood and Frank P. Buff. The Statistical Mechanical Theory of Solutions. I. *The Journal of Chemical Physics*, 19(6):774–777, June 1951.
- [178] D. G. Hall. Kirkwood-Buff theory of solutions. An alternative derivation of part of it and some applications. *Transactions of the Faraday Society*, 67:2516–2524, 1971.
- [179] E. F. Casassa and H. Eisenberg. THERMODYNAMIC ANALYSIS OF MULTICOMPONENT SOLUTIONS. *Advances in Protein Chemistry*, 19:287–395, 1964.
- [180] George Scatchard. Physical Chemistry of Protein Solutions. I. Derivation of the Equations for the Osmotic Pressure1. *Journal of the American Chemical Society*, 68(11):2315–2319, November 1946.
- [181] Jrg Rsgen, B. Montgomery Pettitt, and David Wayne Bolen. Protein folding, stability, and solvation structure in osmolyte solutions. *Biophysical journal*, 89(5):2988–2997, 2005.
- [182] Anil Kumar Nain. Inversion of the KirkwoodBuff Theory of Solutions: Application to Tetrahydrofuran + Aromatic Hydrocarbon Binary Liquid Mixtures. *Journal of Solution Chemistry*, 37(11):1541–1559, November 2008.

- [183] Arieh Ben-Naim. *Molecular Theory of Solutions*. Oxford University Press, October 2006.
- [184] Paul E. Smith. On the KirkwoodBuff inversion procedure. *The Journal of Chemical Physics*, 129(12):124509, 2008.
- [185] E. S. Courtenay, M. W. Capp, C. F. Anderson, and M. T. Record. Vapor Pressure Osmometry Studies of OsmolyteProtein Interactions: Implications for the Action of Osmoprotectants in Vivo and for the Interpretation of Osmotic Stress Experiments in Vitro . *Biochemistry*, 39(15):4455–4471, April 2000.
- [186] Matthew Auton, Luis Marcelo F. Holthausen, and D. Wayne Bolen. Anatomy of energetic changes accompanying urea-induced protein denaturation. *Proceedings of the National Academy of Sciences*, 104(39):15317–15322, 2007.
- [187] Peng Wu and D.W. Bolen. Osmolyte-induced protein folding free energy changes. *Proteins: Structure, Function, and Bioinformatics*, 63(2):290–296, February 2006.
- [188] P. E. Pjura, M. E. Paulaitis, and A. M. Lenhoff. Molecular thermodynamic properties of protein solutions from partial specific volumes. *Aiche Journal*, 41(4):1005–1009, 1995.
- [189] GILBERT NEWTON LEWIS RANDALL, MERLE. *THERMODYNAMICS*. 1961.
- [190] Don Green and Robert Perry. *Perry’s Chemical Engineers’ Handbook, Eighth Edition*. McGraw Hill Professional, November 2007.
- [191] Robert Fraczkiewicz and Werner Braun. Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *Journal of Computational Chemistry*, 19(3):319–333, 1998.
- [192] Oleg V. Tsodikov, M. Thomas Record, and Yuri V. Sergeev. Novel computer program for fast exact calculation of accessible and molecular surface areas and average surface curvature. *Journal of Computational Chemistry*, 23(6):600–609, April 2002.
- [193] Jorge Estrada, Pau Bernad, Martin Blackledge, and Javier Sancho. ProtSA: a web application for calculating sequence specific protein solvent accessibilities in the unfolded ensemble. *BMC Bioinformatics*, 10(1):104, April 2009.
- [194] Pau Bernad, Martin Blackledge, and Javier Sancho. Sequence-Specific Solvent Accessibilities of Protein Residues in Unfolded Protein Ensembles. *Biophysical Journal*, 91(12):4536–4543, December 2006.

- [195] E. O. Saphire. Crystal Structure of a Neutralizing Human IgG Against HIV-1: A Template for Vaccine Design. *Science*, 293(5532):1155–1159, August 2001.
- [196] Peter L. Privalov and Sergey A. Potekhin. [2] Scanning microcalorimetry in studying temperature-induced changes in proteins. *Methods in enzymology*, 131:4–51, 1986.
- [197] Hailong Fu, Gerald Grimsley, J. Martin Scholtz, and C. Nick Pace. Increasing protein stability: Importance of Cp and the denatured state. *Protein Science*, 19(5):1044–1052, March 2010.
- [198] J. K. Myers, C. N. Pace, and J. M. Scholtz. Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Science : A Publication of the Protein Society*, 4(10):2138–2148, October 1995.
- [199] Daniel W. Siderius, William P. Krekelberg, Christopher J. Roberts, and Vincent K. Shen. Osmotic virial coefficients for model protein and colloidal solutions: Importance of ensemble constraints in the analysis of light scattering data. *The Journal of Chemical Physics*, 136(17):175102, 2012.
- [200] Luis Marcelo F. Holthauzen, Matthew Auton, Mikhail Sinev, and Jrg Rsgen. Protein stability in the presence of cosolutes. *Methods in Enzymology*, 492:61–125, 2011.
- [201] G. Xie and S. N. Timasheff. Temperature dependence of the preferential interactions of ribonuclease A in aqueous co-solvent systems: thermodynamic analysis. *Protein Science : A Publication of the Protein Society*, 6(1):222–232, January 1997.
- [202] Irina A. Shkel, D. B. Knowles, and M. Thomas Record. Separating chemical and excluded volume interactions of polyethylene glycols with native proteins: Comparison with PEG effects on DNA helix formation. *Biopolymers*, 103(9):517–527, September 2015.
- [203] M Starzak and M Mathlouthi. Temperature dependence of water activity in aqueous solutions of sucrose. *Food Chemistry*, 96(3):346–370, June 2006.
- [204] V. Rajanikanth, Shanti Swaroop Srivastava, Aditya K. Singh, M. Rajyalakshmi, Kousik Chandra, Penmatsa Aravind, Rajan Sankaranarayanan, and Yogendra Sharma. Aggregation-Prone Near-Native Intermediate Formation during Unfolding of a Structurally Similar Nonlenticular -Crystallin Domain. *Biochemistry*, 51(43):8502–8513, October 2012.
- [205] Francesco Bemporad, Alfonso DeSimone, Fabrizio Chiti, and Christopher M. Dobson. Characterizing Intermolecular Interactions That Initiate Native-Like Protein Aggregation. *Biophysical Journal*, 102(11):2595–2604, June 2012.

- [206] Sharon M. Kelly, Thomas J. Jess, and Nicholas C. Price. How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1751(2):119–139, August 2005.
- [207] Ultraviolet Absorption Spectroscopy of Peptides - Springer. Number 1088 in Methods in Molecular Biology. Humana Press, 2014.
- [208] Chen Zhou, Wei Qi, E. Neil Lewis, and John F. Carpenter. Concomitant Raman spectroscopy and dynamic light scattering for characterization of therapeutic proteins at high concentrations. *Analytical Biochemistry*, 472:7–20, March 2015.
- [209] E. Neil Lewis, Wei Qi, Linda H. Kidder, Samiul Amin, Stacy M. Kenyon, and Steven Blake. Combined dynamic light scattering and Raman spectroscopy approach for characterizing the aggregation of therapeutic proteins. *Molecules (Basel, Switzerland)*, 19(12):20888–20905, 2014.
- [210] S. Amin, S. Blake, S. M. Kenyon, R. C. Kennel, and E. N. Lewis. A novel combination of DLS-optical microrheology and low frequency Raman spectroscopy to reveal underlying biopolymer self-assembly and gelation mechanisms. *The Journal of Chemical Physics*, 141(23):234201, December 2014.
- [211] Alex P. S. Brogan, Kamendra P. Sharma, Adam W. Perriman, and Stephen Mann. Isolation of a Highly Reactive -Sheet-Rich Intermediate of Lysozyme in a Solvent-Free Liquid Phase. *The Journal of Physical Chemistry B*, 117(28):8400–8407, July 2013.
- [212] Laura H. Lucas, Baran A. Ersoy, Lisa A. Kuelto, Sangeeta B. Joshi, Duane T. Brandau, Nagarajan Thyagarajapuram, Laura J. Peek, and C. Russell Middaugh. Probing protein structure and dynamics by second-derivative ultraviolet absorption analysis of cation- π interactions. *Protein Science: A Publication of the Protein Society*, 15(10):2228–2243, October 2006.
- [213] Zai-Qing Wen. Raman spectroscopy of protein pharmaceuticals:. *Journal of Pharmaceutical Sciences*, 96(11):2861–2878, November 2007.
- [214] Roman Tuma, Peter E. Prevelige, and George J. Thomas. Mechanism of capsid maturation in a double-stranded DNA virus. *Proceedings of the National Academy of Sciences*, 95(17):9885–9890, 1998.
- [215] Roman Tuma. Raman spectroscopy of proteins: from peptides to large assemblies. *Journal of Raman Spectroscopy*, 36(4):307–319, April 2005.
- [216] Roman Tuma and George J. Thomas Jr. Theory, design, and characterization of a microdialysis flow cell for Raman spectroscopy. *Biophysical journal*, 71(6):3454, 1996.

- [217] Andreas Barth and Christian Zscherp. What vibrations tell us about proteins. *Quarterly Reviews of Biophysics*, 35(4):369–430, November 2002.
- [218] Frank J. Millero, Roger. Dexter, and Edward. Hoff. Density and viscosity of deuterium oxide solutions from 5-70.deg. *Journal of Chemical & Engineering Data*, 16(1):85–87, January 1971.
- [219] J. F. Swindells and United States National Bureau of Standards. *Viscosities of sucrose solutions at various temperatures: tables of recalculated values*. For sale by the Supt. of Docs., U.S. G.P.O., 1958.
- [220] V. Trappe, J. Bauer, M. Weissmüller, and W. Burchard. Angular dependence in static and dynamic light scattering from randomly branched systems. *Macromolecules*, 30(8):2365–2372, 1997.
- [221] Sandeep Kumar, Mark A. Mitchell, Bonita Rup, and Satish K. Singh. Relationship between potential aggregation-prone regions and HLA-DR-binding T-cell immune epitopes: Implications for rational design of novel and follow-on therapeutic antibodies. *Journal of Pharmaceutical Sciences*, 101(8):2686–2701, August 2012.
- [222] Brent S. Kendrick, Byeong S. Chang, Tsutomu Arakawa, Brian Peterson, Theodore W. Randolph, Mark C. Manning, and John F. Carpenter. Preferential exclusion of sucrose from recombinant interleukin-1 receptor antagonist: role in restricted conformational mobility and compaction of native state. *Proceedings of the National Academy of Sciences*, 94(22):11917–11922, 1997.
- [223] Jun Zhang, Douglas D. Banks, Feng He, Michael J. Treuheit, and Gerald W. Becker. Effects of Sucrose and Benzyl Alcohol on GCSF Conformational Dynamics Revealed by Hydrogen Deuterium Exchange Mass Spectrometry. *Journal of Pharmaceutical Sciences*, 104(5):1592–1600, May 2015.
- [224] Wieland Weise, Andreas Sobotta, Martin Drewel, and others. Electroosmosis in an oscillating field: avoiding distortions in measured electrophoretic mobilities. *Journal of colloid and interface science*, 143(1):287–293, 1991.
- [225] Matthew Auton and D. Wayne Bolen. Additive Transfer Free Energies of the Peptide Backbone Unit That Are Independent of the Model Compound and the Choice of Concentration Scale . *Biochemistry*, 43(5):1329–1342, February 2004.
- [226] B. Lee and F. M. Richards. The interpretation of protein structures: estimation of static accessibility. *Journal of Molecular Biology*, 55(3):379–400, February 1971.
- [227] Jrg Rsgen, Bernard Montgomery Pettitt, and David Wayne Bolen. Uncovering the Basis for Nonideal Behavior of Biological Molecules. *Biochemistry*, 43(45):14472–14484, November 2004.

Appendix A

COMPARISON OF PROTEIN-PROTEIN INTERACTIONS FROM DLS AND SLS, AS-IGG1 NET CHARGE MEASUREMENT DETAILS, AND AN IN-DEPTH CHARACTERIZATION OF AS-IGG1 UNDER AGGREGATION RESISTANT SOLUTION CONDITIONS

A.1 A comparison of protein-protein interactions via dynamic light scattering and static light scattering

In Chapter two, G_{22} was determined as a function of pH, NaCl in 5 mM citrate and 10 mM acetate buffer. These results utilized static light scattering and the excess Rayleigh ratio plotted as a function of concentration. However, dynamic light scattering was simultaneously collected on identical protein solutions. In a similar way, protein-protein interactions can be determined from DLS. D_c , the collective diffusion coefficient is determined from the a non-linear regression of the auto-correlation function and described in Eq. 2.6 [119]. Prior work has shown the D_c is related to protein-protein interactions through the parameter k_D through Eq.A.2. h_0 is a hydrodynamic term which approaches zero in the dilute limit of protein concentration. D_0 is the protein self-diffusion coefficient.

$$D_c = D_0(1 + k_D c_2) \quad (\text{A.1})$$

$$D_c = D_0(1 + (h_0 + 2B_{22})c_2) \quad (\text{A.2})$$

Therefore, k_D values can be determined by fitting D_c as a function of c_2 in the limit of low protein concentration. DLS is much less sensitive to dust than SLS as it utilizes fluctuation in the scattered intensity rather than the absolute scattered intensity. Additionally, DLS is much more friendly to industrial applications as many

relatively high-throughput DLS instruments are commercially available. In contrast, SLS often requires costly quartz cuvette and precise optical configurations and is much more sensitive to dust.

Figure A.1 plots D_c as a function of c_2 . Only the low concentration linear portion of the data was used in fitting to Eq. A.1. At higher protein concentrations, the hydrodynamic factor is expected to have a significant contribution. Prior work has shown that even at low relatively low protein concentrations the hydrodynamic contribution may be significant [148]. This section compares values for k_D and G_{22} determined via DLS and SLS respectively.

Table A.2 reports the values and 95% confidence interval for G_{22} determined via SLS and DLS in both 5 mM citrate and 10 mM acetate buffers. There is reasonable agreement between the values determined from each scattering technique. Additionally, the DLS measurement results in tighter error bars. Perhaps this results is not unexpected as DLS is less sensitive to dust. These results offer a potential method for determined protein-protein interactions from laser light scattering, which may be used for consistency in comparing SLS to DLS. However, one must exercise caution in using DLS to interpret protein-protein interactions, as mentioned above hydrodynamic interactions may be convoluted with protein-protein interactions and often is is not easy to determine when this will occur [148].

Currently, the statistical mechanics is straight forward for determined protein-protein interactions from SLS [122]. The fluctuations in the refractive index of the system are unambiguously related to protein-protein interactions via Kirkwood-Buff theory. On the other hand, the physics underlying protein collective diffusion coefficient are less straightforward. Protein diffusion is a convolution of non-equilibrium thermodynamic (i.e. protein-protein interaction) and transport. As such, more rigorous and computationally intensive approaches may be need to fully understand the effect of protein-protein interaction and hydrodynamics at intermediate to high protein concentrations.

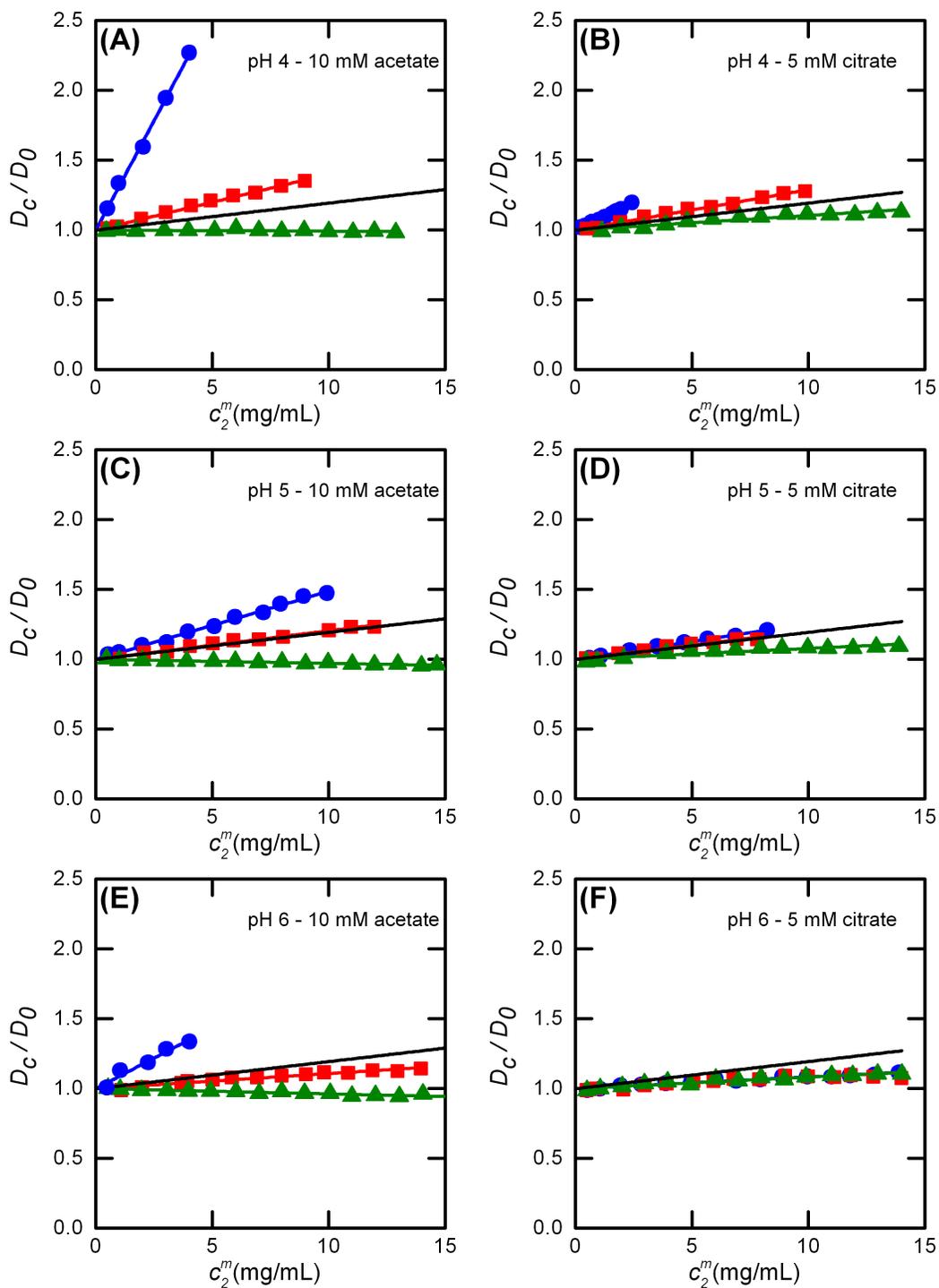


Figure A.1: IgG1 collective diffusion coefficient as a function of protein concentration. for pH 4, 10 mM acetate with 0 mM NaCl (blue), 20 mM NaCl (red), and 200 mM NaCl (green). Colored curves are fits to Eq.A.1 and the black curve is the expected trend for a hard sphere with an equivalent diameter to the IgG1.

Table A.1: Comparison of G_{22}^* values for AS-IgG1 with 10 mM acetate

$G_{22}/(2 * B_{22}^{HS})$ via SLS	0 mM NaCl	20 mM NaCl	200 mM NaCl
pH 4	6±2	1.5±0.2	-0.1±0.2
pH 5	1.8±0.3	1.0±0.2	-0.13±0.18
pH 6	1.6±0.2	0.3±0.3	-0.21±0.13
$-k_D/(2 * B_{22}^{HS})$ via DLS			
pH 4	16.2±0.7	2.07±0.07	-0.04±0.03
pH 5	2.52±0.08	1.06±0.05	-0.14±0.012
pH 6	4.5±0.7	0.55±0.3	-0.21±0.02

Table A.2: Comparison of G_{22}^* values for AS-IgG1 with 5 mM citrate

G_{22}^* via SLS	0 mM NaCl	20 mM NaCl	200 mM NaCl
pH 4	4.6±1.2	1.2±0.2	0.53±0.12
pH 5	1.03±0.12	0.83±0.13	-0.2±0.4
pH 6	0.4±0.3	0.3±0.2	0.2±0.2
$-k_D/(2 * B_{22}^{HS})$ via DLS			
pH 4	4.03±0.13	1.5±0.1	0.54±0.05
pH 5	1.32±0.04	1.06±0.09	0.41±0.06
pH 6	0.45±0.03	0.40±0.06	0.44±0.04

A.2 Electrophoretic light scattering with phase analysis light scattering (PALS) using the Zetasizer

Electrophoresis is a non-equilibrium thermodynamic process in which an applied electric field leads to a movement of molecules in solution. The advent of phase analysis light scattering has improved electrophoresis detection sensitivity, [128, 127] but the electrophoretic mobility measurement may be limited by molecular motion due to other processes such as diffusion [128].

During the electrophoretic light scattering measurement the IgG1 moves due to electrophoresis and diffusion. mAbs have a hydrodynamic diameter of about 10 nm and self-diffusion coefficient of 4.4×10^{-7} cm²/s (measured from DLS). With a large diffusion coefficient, the movement due to the diffusion may be comparable or greater than electrophoresis. While the displacement of the IgG1 molecule due to diffusion will average to zero over large timescales, as diffusion is a random walk, the displacement due to electrophoresis lies in the same direction as the applied electric field. As such, the average distance moved by electrophoresis scales linearly with time field is applied. However, the average distance moved by diffusion scales with \sqrt{t} .

Figure A.2-A shows the distance moved by diffusion (red dotted line) and electrophoresis (blue solid line) in an electric field of 20 V/cm and mobility of 1×10^{-4} cm²/s V. Given these conditions, the lines intersect at about one second, which indicates the distance moved by the protein due to electrophoresis is the same as diffusion. As Figure A.2-A indicates, measuring electrophoretic mobility over longer time produces more accurate results. Figure A.2-B shows the voltage applied during a measurement run using a standard operating procedure from Malvern Zetasizer Nano ZS (Malvern Instruments, Westborough, MA). In the first 0-1.2 seconds a fast field reversal (FFR) is used, in which the electric field is rapidly oscillated back and forth from positive to negative voltage. From 1.2 to 2.6 seconds a slow field reversal (SFR) is used. Typically this measurement is repeated multiple times to improve statistics. In principle, the fast field reversal leads to a true measure of the sample electrophoretic mobility as the electro-osmotic flow (EOF) will not develop on the short timescales [224]. The

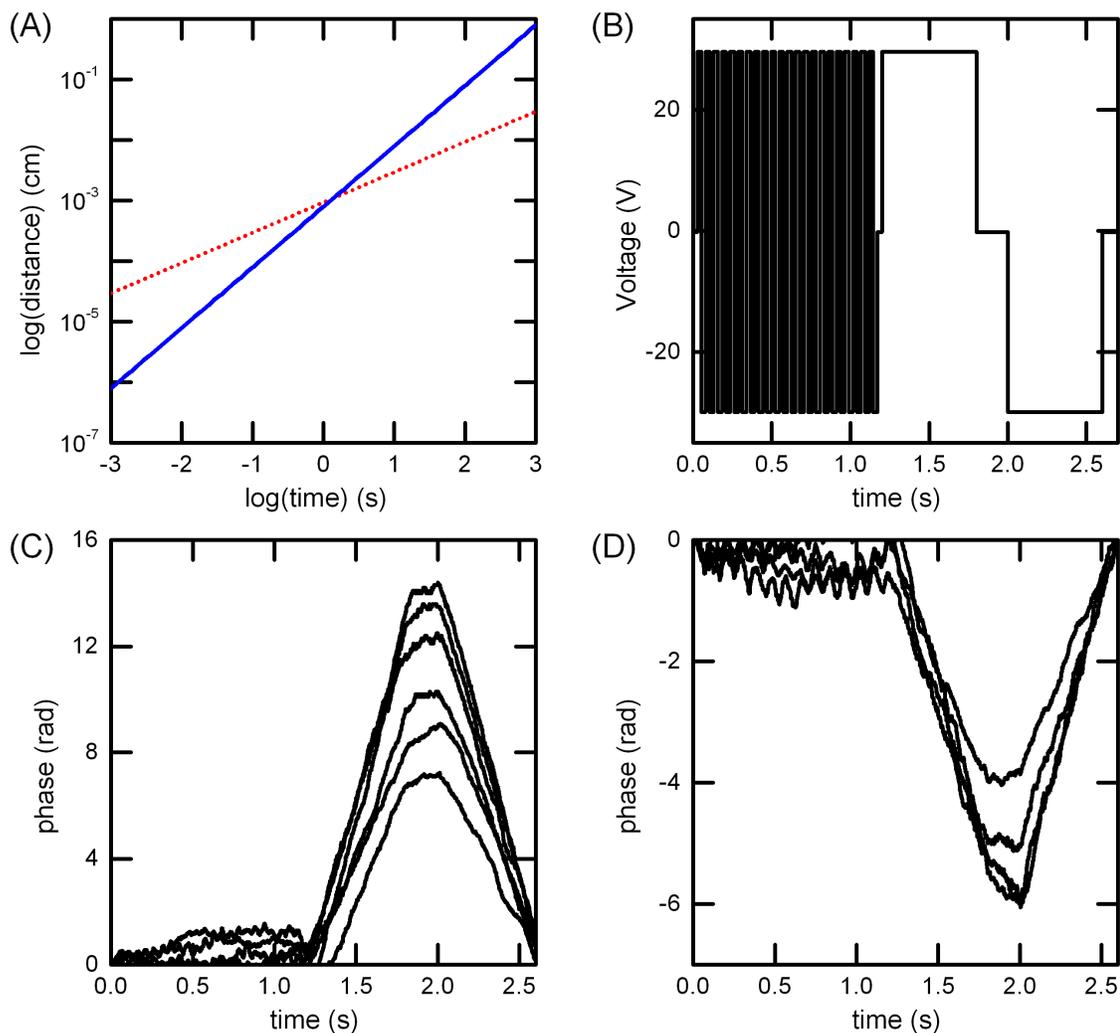


Figure A.2: IgG1 electrophoretic light scattering. (A) Average distance moved by the IgG1 from diffusion (red) and electrophoresis (blue). (B) Applied voltage vs time for a single run using the Zetasizer. (C) Phase shift plotted vs time for IgG1 pH, 4 20 mM NaCl. Each curve represents single electric field strength. (D) Phase shift plotted as a function of time for PEG at pH 6, 40 mM NaCl exhibiting electro-osmotic flow. Each curve corresponds to a different electric field strengths

slow field reversal is used to determine statistics regarding the electrophoretic mobility distribution, but may be convoluted with electro-osmotic effects. EOF occurs due to electrophoresis of solution ions, which hydrated water molecules and create flow [224]. The concentration and type of ions present affect EOF, but predicting the undesirable EOF effect is challenging. While measurements in the FFR will eliminate the possibility of EOF, the signal to noise ratio of the displacement in the direction of the electric field will be small as the distance moved to diffusion will be order of magnitude larger than movement due to electrophoresis (as mentioned above). However, the signal to noise ratio in the slow field reversal will be higher as the mobility was measured over a longer time.

Figure A.2-C shows the measured phase shift as a function of time for different electric field strengths using the IgG1 at pH 4 20 mM NaCl 5 mM citrate. In the FFR region of Fig. A.2-C there is a small zigzag pattern, but reliably interpreting this portion of the results is challenging as mentioned above. However, in the SFR region the phase changes linearly with the applied electric field. As mentioned in the main text the positive and negative slopes in the SFR region are proportional to the velocity and described by Equation 2.8 in the main text.

While the measured phase shift over time can be related to a velocity, the samples velocity may be convoluted by EOF. In the present case, EOF was monitored with a separate sample of polyethylene glycol with molecular weight of 35kDa (Sigma Alrich) in identical solution conditions as the IgG1. The only samples to exhibit EOF were pH 6 in 5 mM citrate buffer at 20 mM, 40 mM, and 100 mM NaCl. Figure A.2-D shows phase plot for PEG at pH 6 20 mM NaCl 5mM citrate with different values of the applied electric field. As the phase is negative, the PEG is moving in a direction towards the cathode. This may indicate multivalent citrate ions at pH 6 may be responsible to the EOF effect in this case. The electrophoretic mobility was determined from the slope of each data set in Fig. A.2 and the IgG1 effective charge determined using Eq. 2.8. In principle, measuring velocity over multiple electric field strengths improves accuracy of reported electrophoretic mobility and Z_{eff} values. There are not large differences

in the slopes of the data sets for different NaCl concentration in Fig. A.2. a possible explanations for this result may be the small NaCl concentration window chosen for measurements did not have a large change in Z_{eff} . One also has to bear in mind that the electrophoretic light scattering measurements is challenging as mentioned above.

The addition of 20 mM NaCl was arbitrary chosen as the lower limit in electrophoresis measurements to eliminate possible EOF. 100 mM NaCl was arbitrary chosen as the upper limit in mobility measurements as higher sample conductivity leads to larger currents, which may degrade the protein and sample cell. To ensure sample quality, dynamic light scattering was performed on samples before and after the electrophoresis measurement and the cell and sample were visually inspected for electrode degradation or precipitation.

A.3 AS-IgG1 aggregation resistant condition: pH 4, 10 mM acetate

As observed in the aggregation state diagram in Chapter two, AS-IgG1 at pH 4, 10 mM acetate heated at temperatures above the mid-point unfolding transitions do not appear to aggregate on timescale of minutes to hours. This section investigates an in-depth characterization of this aggregation resistant behavior using in-situ DLS and SAXS.

DLS was performed as described in Section 2.2.5. Small angle x-ray scattering was performed on the 5.2 beamline of Elettra Synchrotron in Trieste, Italy. A custom-made sample holder was used to perform temperature ramping studies [6]. AS-IgG1 at 10 mg/mL was heated from room temperature to 90°C while SAXS profile were collected. Higher protein concentrations were need to for adequate scattering signal, though at this high of a concentration significant structure factor effects are expected. See Chapter five and particularly, Fig. 5.7 for related discussion.

The DSC provides a qualitative measure of the mAb unfolding and conformational stability. As shown in Fig. 3.2-A and shown here for clarity, DSC profiles of full IgG1 and Fc fragment in the are are shown in Figure A.3-A. See Chapter two for more detailed discussion of DSC. For AS-IgG1 at pH 4, 10 mM acetate, the C_H2 domain

of the Fc region unfolds first followed by the Fab region and finally the C_H3 domain of the Fc. During unfolding process, hydrophobic patches of the protein may become exposed and interact with other protein hydrophobic patches resulting in nonnative aggregation. As discussed in Chapter two, this IgG1 at different formulations conditions showed irreversible aggregation occurs during heating below transition temperatures.

The right graph in Fig. A.3-B shows the size exclusion chromatogram of AS-IgG1 after heating for one hour at 85°C and subsequently quenched on ice. The chromatogram shows no detectable formation of aggregates. Further characterization was needed to confirm no aggregation was present and to characterize the structural changes occurring during heating. SEC analyzes aggregates that are incubated and quenched on ice but does not provide an *in-situ* characterization of the protein aggregate.

In-situ SAXS was collected on AS-IgG1 as it was heated from 25 to 90 °C. SAXS was used to investigate structural changes that may occur as AS-IgG1 unfolds and loses conformation. Also SAXS will be sensitive to aggregates that may potentially form *in-situ*. Figure A.3-C shows AS-IgG1 SAXS intensity profiles as a function of temperature. One observes from this plot that during temperatures of 34-58°C the scattering profile remains unchanged and has the characteristic shoulder seen commonly for IgG1 [138]. At these temperatures the IgG1 remains in its native form. From 62-77°C the scattering signal from $Q = 0.05 - 0.1 \text{ \AA}^{-1}$ monotonically decreases. This indicates IgG1 structural changes are occurring at these temperatures. This Q range corresponds to length-scales ranging from sizes similar to an IgG1 molecule to sub-nanometer sizes. As the IgG1 structure is changing, the intensity at low Q is simultaneously increasing. While the SAXS configuration used for these measurements did not reach low enough Q to measure these larger length scales, the low Q upturn is indicative of aggregation.

Figure A.3-D show SAXS profiles at 34°C, 89°C, and at 25°C post-scan. One observes in Fig. A.3-D that AS-IgG1 is aggregating at high temperature, but these aggregates dissociate when the protein is cooled to room temperature. Interestingly the SAXS profile post-scan is significantly altered at intermediate to high Q , suggesting AS-IgG1 did not refold to its native structure upon cooling.

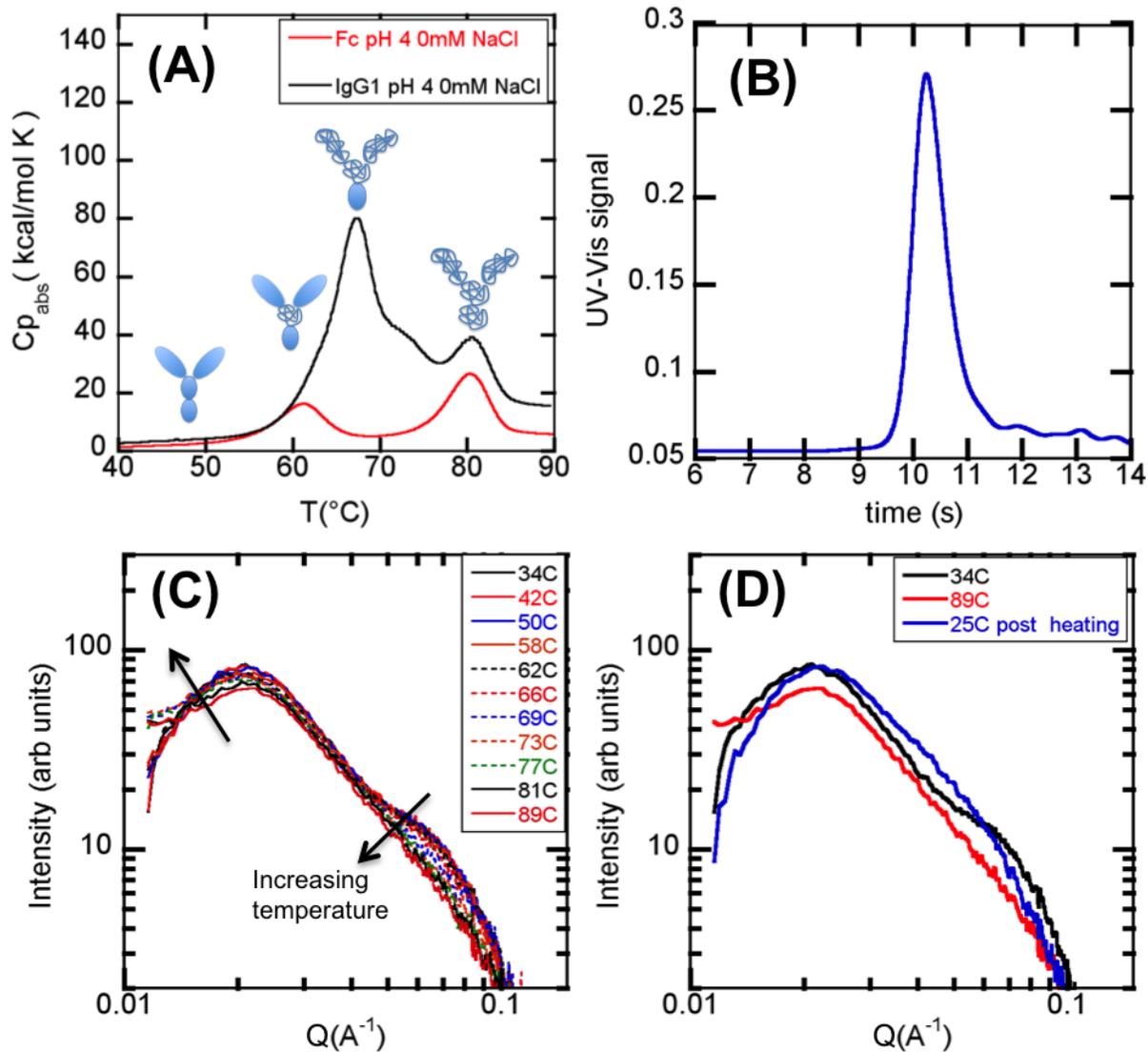


Figure A.3: AS-IgG1 unfolding under aggregation resistant solution conditions. (A) AS-IgG1 DSC thermogram at pH 4, 10 mM acetate showing three unfolding transitions. (B) AS-IgG1 SEC chromatogram after heating at 85 °C for one hour. (C and D) *In-situ* AS-IgG1 temperature scanning SAXS.

DLS provides an orthogonal *in-situ* technique complementing SAXS that is sensitive to small quantities of aggregates. Figure A.4 shows auto-correlation functions for AS-IgG1 heated from 25 to 90°C. At each temperature, the sample was allowed to equilibrate for five minutes to ensure stable temperature control. One observes from Fig. A.4-A that no significant changes are noticeable in the auto-correlation functions during heating from 25 to 55°C. However Fig A.4-B shows as AS-IgG1 is heated at higher temperatures, the protein begins to aggregate. At 75°C a significant population of aggregates has formed. Interestingly, the aggregate sizes from the auto-correlation function correspond to around 100 nm, which suggest these aggregates are composed of many monomers. Fig. A.4-C shows that heating AS-IgG1 further only increases aggregation. Finally, Fig. A.4-D shows the auto-correlation function upon cooling from 90 to 25°C. As AS-IgG1 cools, aggregates dissociate.

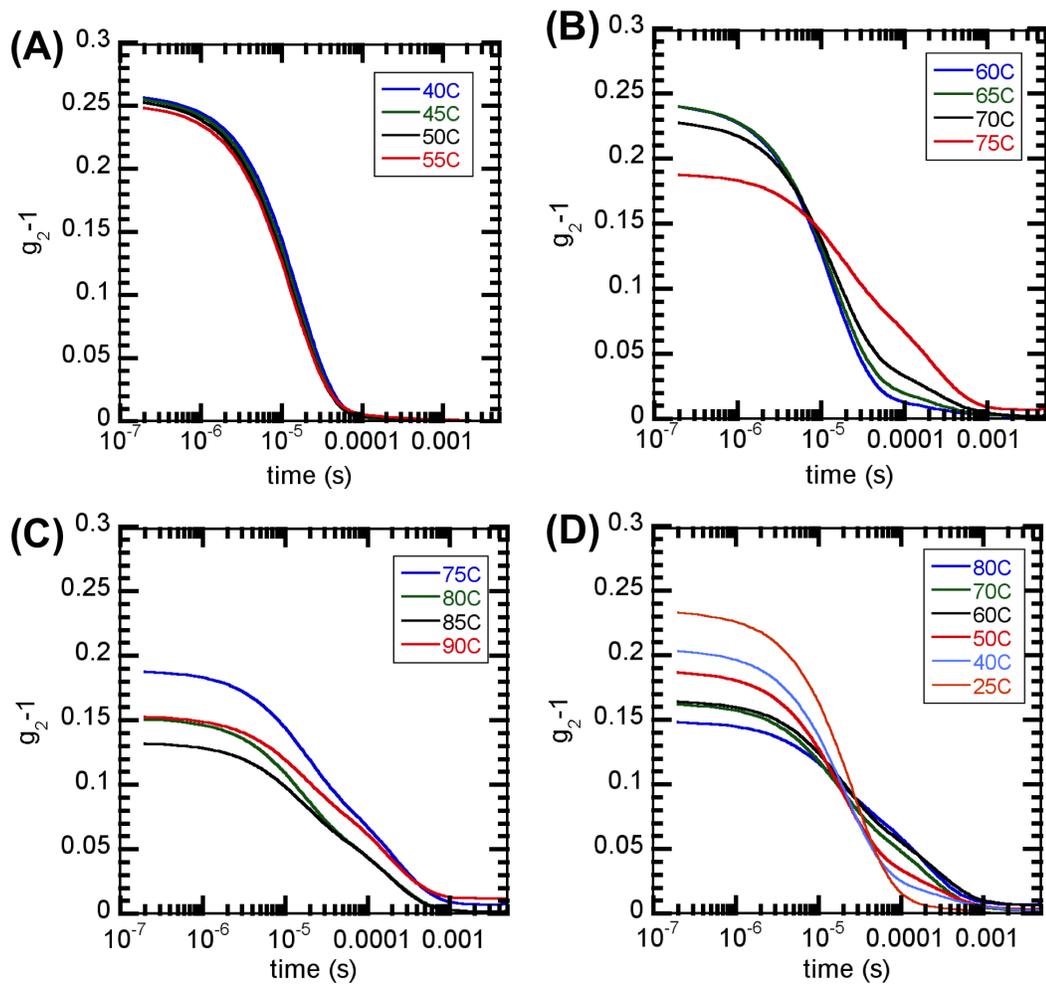


Figure A.4: DLS auto-correlation functions as a function of temperature for IgG1 at pH 4, 10 mM acetate during heating and cooling.

Appendix B

DETAILED DENSIMETRY RESULTS AND ANALYSIS

B.1 Density as a function of AS-IgG1 weight fraction

In Chapter four, values for \hat{V}_2 were determined from a set of density measurements of AS-IgG1 and a given osmolyte. Density measurements were performed at near neutral pH, where AS-IgG1 has little to no net charge, and solutions prepared between 0-5 mg/mL so as to minimize potential effects of protein-protein interactions. Values for the partial specific volume of the protein (\hat{V}_2) were determined from Equation 4.5 of the main text. Lewis and Randal derived Eq. 4.5 for a binary system and in a similar manner it was derived and extended to a ternary system. The following figures show raw densimetry data for sucrose (Figure B.1), trehalose (Figure B.2), sorbitol (Figure B.3), and PEG (Figure B.4).

B.2 Statistics for linear regression of $\hat{V}_2(c_3\hat{V}_3)$

For each protein-osmolyte data set consisting of $\hat{V}_2(c_3)$, a linear regression was performed using Equation B.1. The coefficients of A, B, and C are fitted values, \hat{V}_2 is the protein partial specific volume, \hat{V}_3 is the osmolyte partial specific volume, and c_3 is the osmolyte concentration. In some subsequent cases, higher order terms in Eq B.1 were omitted (based on statistical analysis).

$$\hat{V}_2 = A + B(c_3\hat{V}_3) + C(c_3\hat{V}_3)^2 \quad (\text{B.1})$$

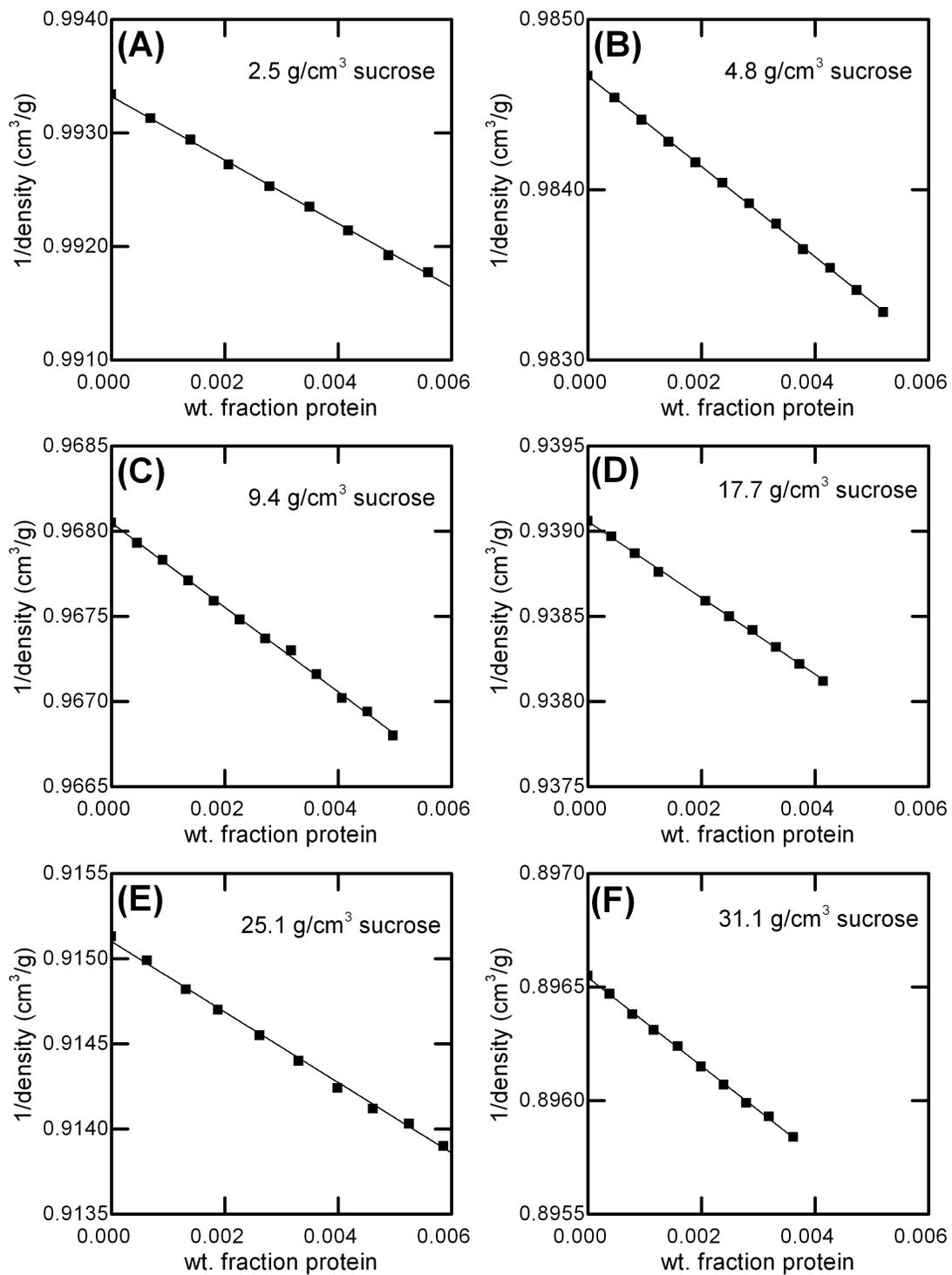


Figure B.1: Reciprocal of the solution density as a function of AS-IgG1 weight fraction at various sucrose concentrations.

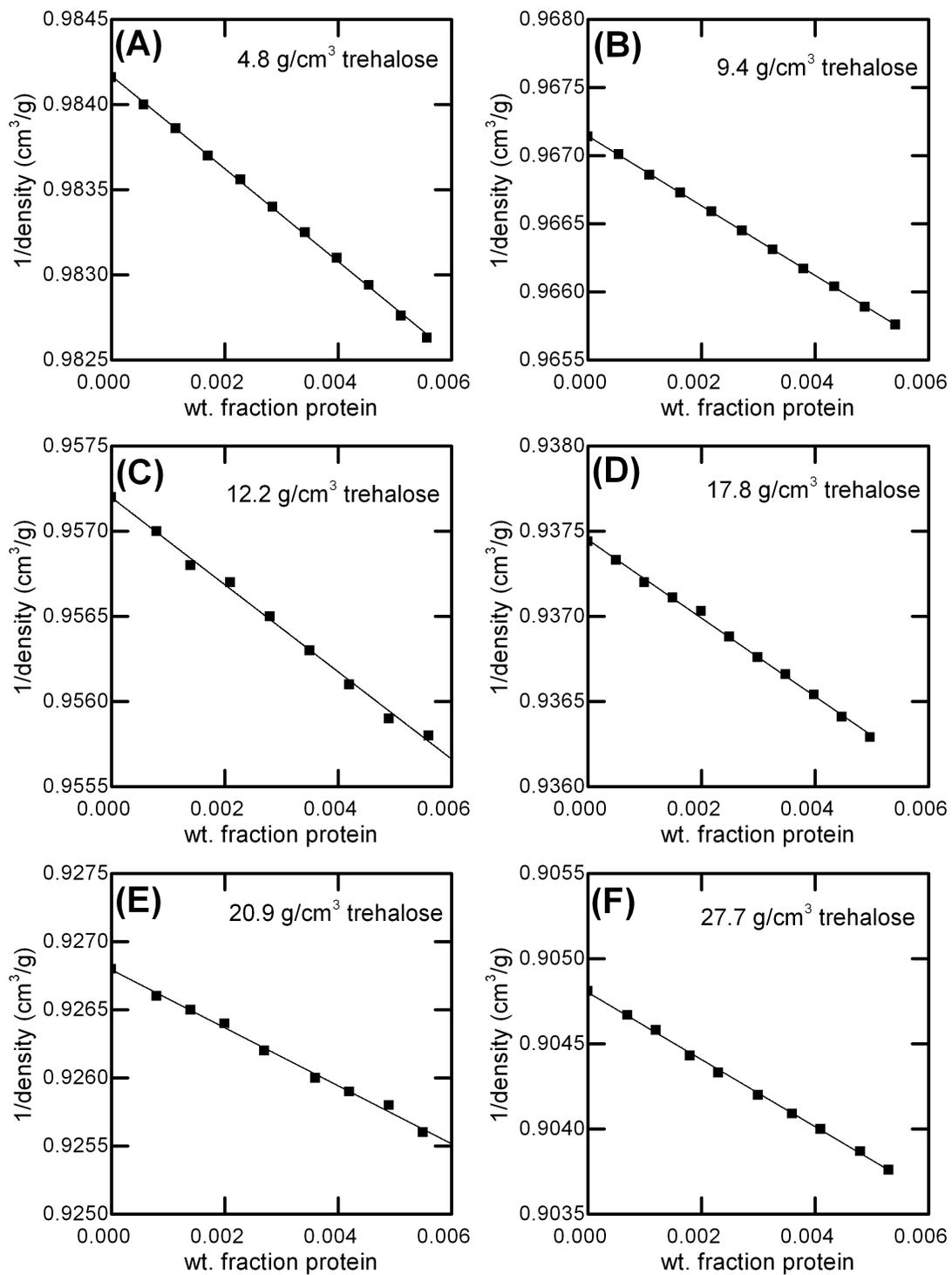


Figure B.2: Reciprocal of the solution density as a function of AS-IgG1 weight fraction at various trehalose concentrations.

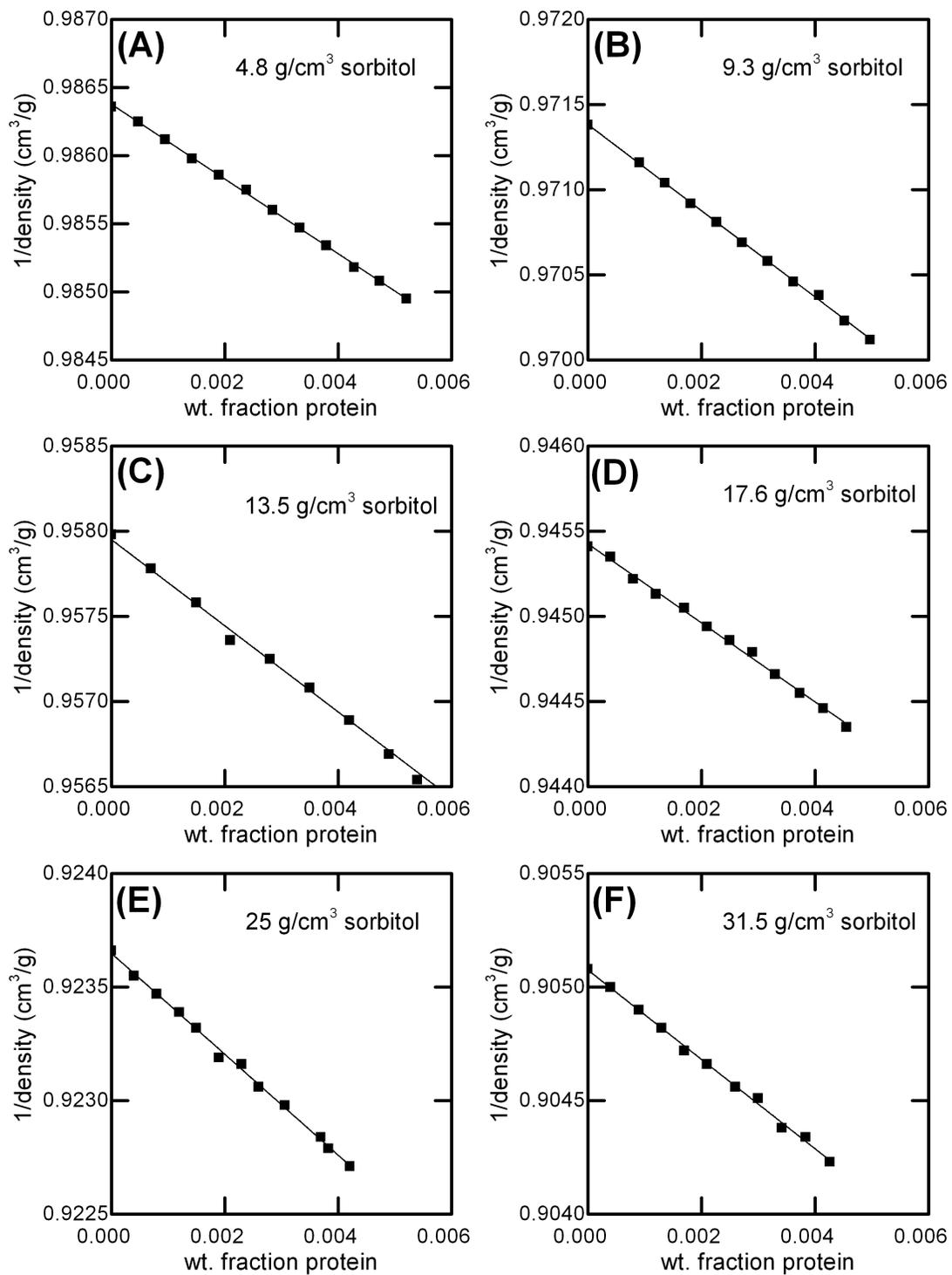


Figure B.3: Reciprocal of the solution density as a function of AS-IgG1 weight fraction at various sorbitol concentrations.

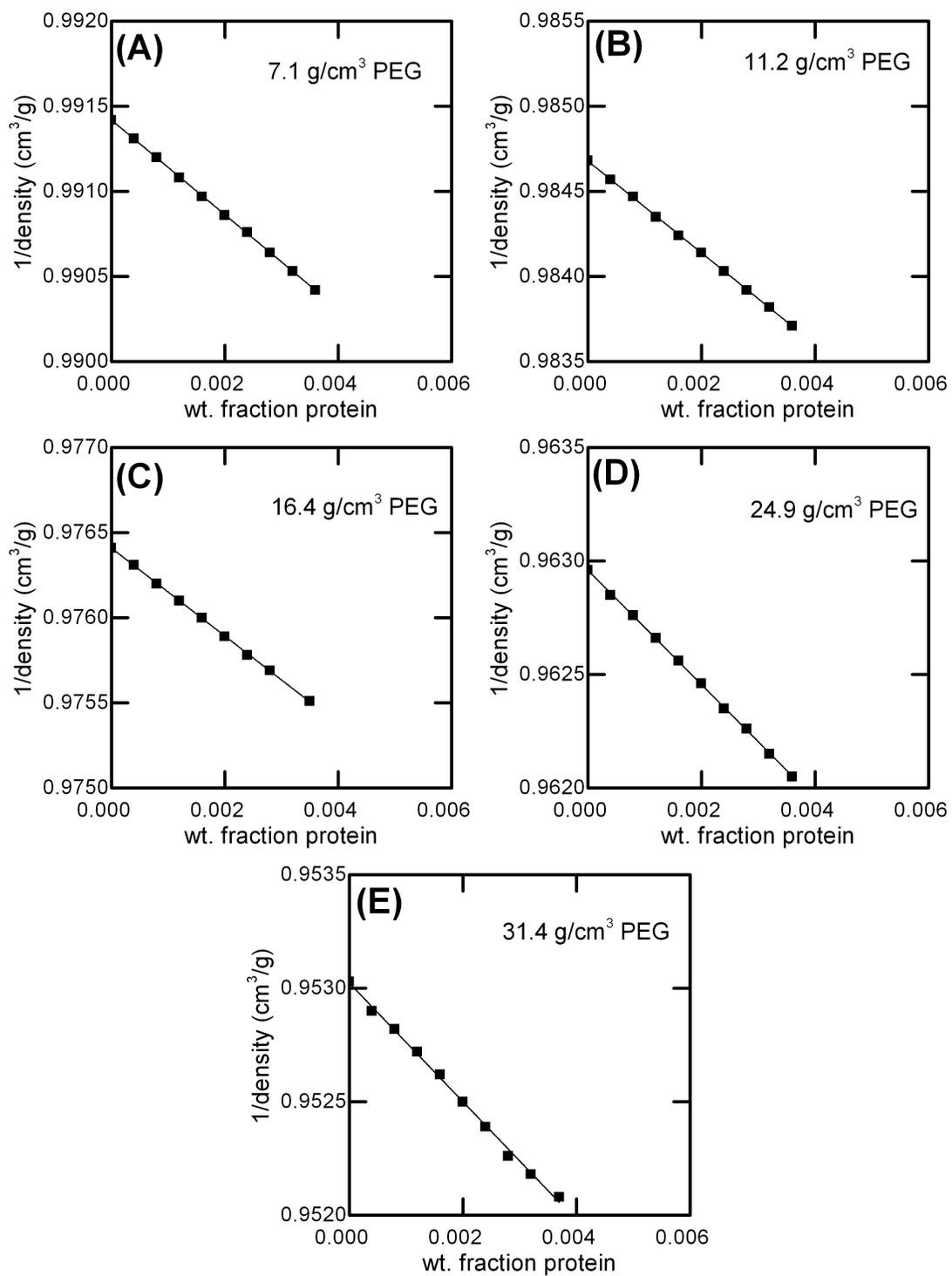


Figure B.4: Reciprocal of the solution density as a function of AS-IgG1 weight fraction at various PEG concentrations.

Table B.1: Regression analysis for IgG1 and sucrose

Coefficient	Value Standard error	p-value
A	0.706 ± 0.004	0.001
B	0.27 ± 0.011	0.06
C	-1.6 ± 0.5	0.04

Table B.2: Regression analysis for IgG1 and trehalose

Coefficient	Value Standard error	95 % confidence interval
A	0.7103 ± 0.0006	0.0016

Table B.3: Regression analysis for IgG1 and sorbitol

Coefficient	Value Standard error	95 % confidence interval
A	0.711 ± 0.004	0.011

Table B.4: Regression analysis for IgG1 and PEG

Coefficient	Value Standard error	p-value
A	0.7005 ± 0.0013	0.001
B	0.28 ± 0.02	0.001
C	-1.15 ± 0.08	0.001

B.3 Preferential interaction model predictions

Preferential interaction models were used to predict the AS-IgG1 native state chemical potential and m-value for a given osmolyte. Model calculations for sucrose, sorbitol, and trehalose utilized the additive tripeptide model developed by Auton and Bolen [2]. This model is based on the free energy required to transfer a given amino

acid from water to an aqueous osmolyte solution (typically 1 M osmolyte). The transfer free energy values were developed based on a series of peptide solubility measurements in water and the corresponding osmolyte solution [2, 225].

For PEG, an atom additive preferential interaction model was recently developed by Record and coworkers [3, 202]. This model parses out chemical interaction of functional group (e.g. Amide nitrogen, aliphatic carbon) with PEG and also has an additive excluded volume terms based on polymer theory. These two models utilize solvent accessible surface area of the protein, which can readily be determined for the native state using algorithms that determine the area accessible to a spherical probe (0.14 nm diameter) on the three dimensional (3D) surface of the protein. The 3D structure is available from NMR or x-ray crystal structure and available for many protein in the protein data bank(PDB) file. In the present case, a homology model was built of based on the crystal structure of a human IgG1 (PDB:1HZ1) [195].

ASA estimates for the folded state of the protein are available using many algorithms. In the present case, three algorithms used: GET AREA,[191] Surface Racer,[192] or ProtSA,[193, 194]. Figure B.5 shows comparison ASA determined from Surface Racer and GET AREA. The backbone (top) or side chain (bottom) fractional exposure of each amino acid was determined by normalizing the total ASA for the side chain or backbone by the number of residues and the ASA values for the fully exposed residue, which was determined previously using Gly-X-Gly tripeptides [226]. Here X represent each of the other amino acids. Gly does not have a side chain and therefore the exposed surface area for Gly-X-Gly tripeptides corresponds to fully exposed surface area for a given amino acid.

B.4 Osmolyte-water thermodynamics

The conversion between $(G_{12} - G_{23})$ and $\left(\frac{\partial \mu_2}{\partial c_3}\right)_{(T,P)}$ requires knowledge of the water-osmolyte G_{ij} 's as shown in Eqs. 4.1-4.2. For Kirkwood-Buff analysis of dilute protein concentrations, G_{ij} 's for osmolyte and water are known from thermodynamic

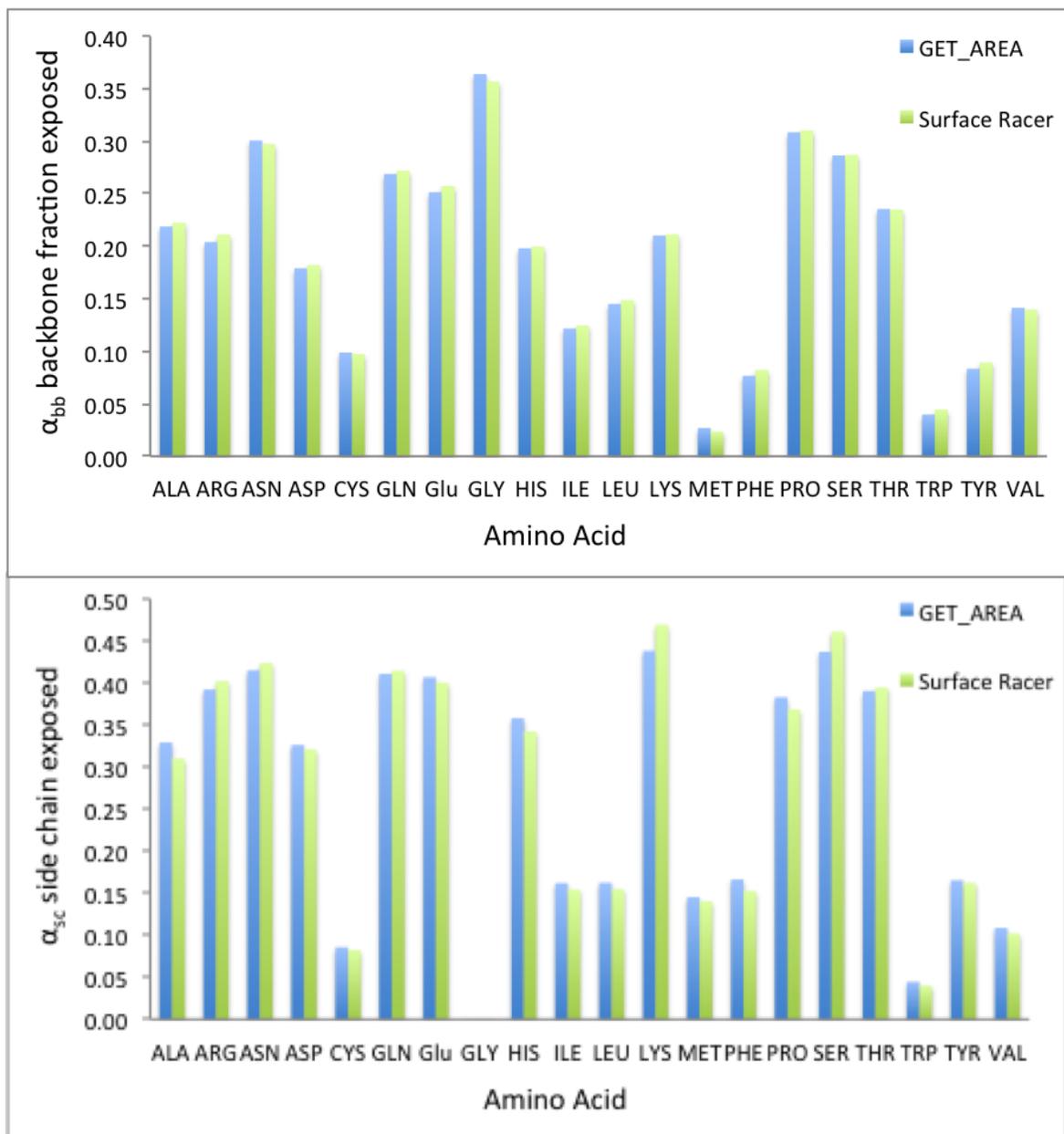


Figure B.5: Comparison of ASA algorithms determine backbone (top) and side chain (bottom) fractional exposure

analysis of osmolyte-water solutions. Particularly, analytical solutions utilizing osmolyte activity coefficients are found in the literature for sucrose and sorbitol [227]. However, for trehalose and PEG osmolyte-water aqueous thermodynamic data was needed to determine $\left(\frac{\partial\mu_3}{\partial c_3}\right)_{(T,P)}$.

Water activity data for aqueous solutions of trehalose [6] or PEG [4, 5] were available in literature and used to determine osmolyte non-ideality value $\left(\frac{\partial\mu_2}{\partial c_3}\right)_{(T,P)}$ as a function of trehalose concentration. The Gibbs-Duhem relation for a binary systems is given by Eq. B.2. It relates changes in chemical potential of component one to changes in the other component. c_i is the concentration of component i and μ_i is the chemical potential of component i . Differentiating Eq. B.2-B.3 with respect to c_3 and algebraic rearrangement results in Eq B.3, which is a useful form for analyzing water activity data of aqueous osmolyte solutions.

$$c_3\partial\mu_3 + c_1\partial\mu_1 = 0 \quad (\text{B.2})$$

$$\left(\frac{\partial\mu_3}{\partial c_3}\right)_{(T,P)} = -\frac{c_1}{c_3}\frac{\partial\mu_1}{\partial c_3} \quad (\text{B.3})$$

Figure B.6-A-B plots the water activity as a function of PEG (panel A) and trehalose (panel B) concentrations taken from literature. The symbols correspond to data taken from literature and black line to a polynomial fit. Osmolyte concentration dependent values of $\left(\frac{\partial\mu_2}{\partial c_3}\right)_{(T,P)}$ were calculated using Eq. B.3 and plotted in Figure B.6-C-D for PEG and trehalose respectively.

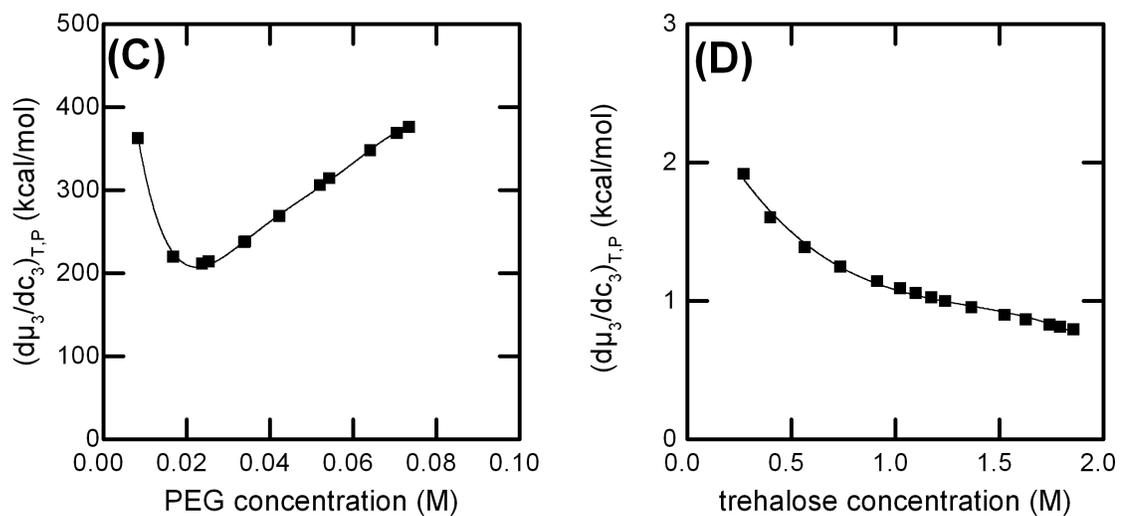


Figure B.6: Osmolyte-water binary systems. Water activity data for aqueous solutions of (Panel A) PEG (data from [4, 5]) and (Panel B) trehalose (data from [6]). Derivative of the osmolyte chemical potential with respect to osmolyte concentration as a function of osmolyte concentration for (C) PEG and (D) trehalose.

Appendix C
REPRINT PERMISSION LETTERS

**RightsLink®**[Home](#)[Account Info](#)[Help](#)**ACS Publications**
Most Trusted. Most Cited. Most Read.

Title: Specific-Ion Effects on the Aggregation Mechanisms and Protein-Protein Interactions for Anti-streptavidin Immunoglobulin Gamma-1

Logged in as:
Gregory Barnett
Account #:
3000875223

[LOGOUT](#)

Author: Gregory V. Barnett, Vladimir I. Razinkov, Bruce A. Kerwin, et al

Publication: The Journal of Physical Chemistry B

Publisher: American Chemical Society

Date: May 1, 2015

Copyright © 2015, American Chemical Society

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

[BACK](#)[CLOSE WINDOW](#)

Copyright © 2015 [Copyright Clearance Center, Inc.](#) All Rights Reserved. [Privacy statement.](#) [Terms and Conditions.](#)
Comments? We would like to hear from you. E-mail us at customer@copyright.com



RightsLink®

[Home](#)[Account Info](#)[Help](#)ACS Publications
Most Trusted. Most Cited. Most Read.

Title: Structural Changes and Aggregation Mechanisms for Anti-Streptavidin IgG1 at Elevated Concentration

Author: Gregory V. Barnett, Wei Qi, Samiul Amin, et al

Publication: The Journal of Physical Chemistry B

Publisher: American Chemical Society

Date: Nov 1, 2015

Copyright © 2015, American Chemical Society

Logged in as:
Gregory Barnett
Account #:
3000875223

[LOGOUT](#)

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

[BACK](#)[CLOSE WINDOW](#)

Copyright © 2015 [Copyright Clearance Center, Inc.](#) All Rights Reserved. [Privacy statement.](#) [Terms and Conditions.](#)
Comments? We would like to hear from you. E-mail us at customercare@copyright.com