Rational Design of Protein-loaded Polymeric nanoparticles: A Computational and Experimental Approach

Christopher Wanjohi Nyambura

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2022

Reading Committee:

Jim Pfaendtner

Elizabeth Nance

Lilo Pozzo

Program authorized to Offer Degree:

Chemical Engineering

University of Washington

Abstract

Rational Design of Protein-loaded Polymeric nanoparticles:
A Computational and Experimental Approach

Chair(s) of the Supervisory Committee:
Jim Pfaendtner
Elizabeth Nance
Department of Chemical Engineering

Enzymes are a class of macromolecules that can catalyze a wide range of chemical reactions. Enzymes are critical in the function of complex biological processes and more recently, industrial processes that result in high-value products and services. Their large and complex structure, when compared to small molecules, allows for unique chemistries to occur, due to highly specific binding of substrates within the reaction site and subsequent conversion to products with high regio- and/or chemo-selectivity. However, the environment in which enzymes execute their function must be tailored to ensure preservation of protein structure and easy access to substrates. For therapeutic drug delivery, toxic side-effects such as protease recognition and development of neutralizing antibodies can occur when administrating a therapeutic enzyme in free form; hence, their targeted delivery is necessary to ensure high patient compliance and to reduce healthcare costs associated with medical care. Polymer nanoparticles (PNPs) can potentially address issues observed in enzymatic drug delivery due to their high surface area, ability to encapsulate a wide range of proteins, and ability to tailor the nanocarrier surface to minimize non-specific interactions. Nevertheless, ensuring high enzymatic loading while preserving protein activity when formulating protein-loaded PNPs remains a difficult task, due to a poor understanding of molecular-level mechanisms that allow for high encapsulation efficiencies and desirable PNP characteristics such as monodispersity, near neutral surface charge, and small size.

To aid in the rational design of enzyme-loaded PNPs, Molecular Dynamics provides a way to probe key interactions, in atomistic detail, that are present at many points during nanoparticle formulation, resulting in better methodologies for making highly effective nanocarriers. In this dissertation, I elucidated structure and dynamics of the poly(lactic-*co*-glycolic acid)–polyethylene glycol (PLGA-PEG) copolymer and its homopolymer constituents in solvents commonly used to form core-shell PNPs, resulting in key insights that are necessary to control polymer chain rigidity and shape. Next, I examined the role of polymer extension on protein-polymer interactions prior to and during formation of PLGA-PEG nanoparticles, to better understand how the choice of solvent could impact enzymatic loading. Lastly, I employed hydrophobic-ion pairing to reduce the hydrophilic nature of catalase to drive increased encapsulation within PLGA-PEG nanoparticles. This work demonstrates the advantages of using both computational and experimental tools to develop and rationally design enzyme-loaded PNPs.

Table of Contents

List of Figures

List of Tables

ACKNOWLEDGEMENTS

My PhD in chemical engineering has been the hardest thing I've done with my life. I will be the first scientist and engineer in my family and I couldn't have made it with my team: My mother and brother. When we first moved to the USA in 2005 with just our clothes in a suitcase, I never imagined I would end up here. We have gone through so much as immigrants. Leaving my friends and family at a young age to pursue a better life, in many ways, prepared me to do graduate school. My mother and brother were a constant source of inspiration and support, at all points of my academic journey. Without them, I'm not sure if I would have made it to the end. They showed me what it takes to achieve my dreams: immense sacrifice and consistency in effort. For that, I will be forever grateful to them.

Next, I would like to thank my advisors: Jim Pfaendtner and Elizabeth Nance. Engineering at the nanoscale is no joke. Nance helped me drastically improve my understanding of my thesis problem. She showed me how to clearly and effectively communicate my computational work to her experimental group. I will always admire how great of a scientific communicator she is. Her perspective and suggestions were important in my work, and I will always be grateful for that. Jim's patience with helping me get up to speed with running computer simulations, in addition to him asking me tough scientific questions and helping me think on the atom level, is statement to his mentorship. As my academic parents, I truly appreciate their effort and support during my PhD thesis work.

I would also like to acknowledge past members of the Pfaendtner and Nance research groups for helping me adjust to the grad school grind and helping me understand research expectations: These people are Arushi Prakash, Josh Smith, Sarah Alamdari, Janani Sampath, Rick Liao and Andrea Joseph. Thank you for answering my constant stream of questions and for helping me become a computational modeling and nanoparticle expert. I want to give a special shoutout to Janani Sampath: thank you for giving me the critical feedback I needed to publish my polymer/solvent work. You helped me understand the research field of polymer science and what I need to do to match the quality of my work to the field. Lastly, I want to acknowledge myself for not giving up at the toughest points of this journey, for learning how to eat right and exercise, for practicing self-love in a racist world, for becoming more self-aware of my feelings and emotions and for never losing my resolve. I know now what it takes to do something difficult and worthwhile; that experience/perspective is irreplaceable, and I will always be thankful to my PhD journey for showing me what it takes.

DEDICATION

For every immigrant who has big dreams: Always remember your worth and double down on yourself. You are stronger and smarter than you think you are. Keep pushing!

To every brilliant black person chasing their dreams: Don't believe the lies of racism and white supremacy. Practice self-love every day and show yourself grace. Remember your history because we have always been innovators, engineers and scientists (not just artists) who have shifted the trajectory of humanity.

1. Chapter 1: Introduction

Drugs, in order to be efficacious and effective in raising a patient's quality of life, have to overcome difficult biological barriers. Depending on the route of administration, an active pharmaceutical ingredient (API) has to overcome the hepatic first pass, avoid immune recognition and protease degradation; moreover, the API should provide a controlled therapeutic effect without resulting in off-target side effects[1]. Typically, most small molecule therapeutics have been delivered orally via a solid dosage form, due to higher patient compliance with this administration route. If crystallization is difficult or other macroscopic properties impede solid dosage processing, then other administration routes, such as intravenous or intranasal delivery, are used[1,2]. The choice of delivery method, however, is not as straightforward and the manner in which a drug is absorbed, metabolized, distributed, and excreted (ADME) by the human body will ultimately dictate whether the right administration method was chosen. With the rise of protein drugs in the pharmaceutical market, the means of delivery have become more important, since this class of APIs is typically high molecular weight, hydrophilic in nature and have charged surface moieties that can prevent high absorption into the bloodstream[2,3]. Oral delivery of unencapsulated proteins is limited, due to enzyme degradation in the gastrointestinal tract and inability to bypass intestinal epithelial cells[4,5]. With multiple biological barriers in place to hamper drug efficacy, biodegradable polymeric nanoparticles (PNPs) provide a way for protein and peptide drugs to stay within an optimal therapeutic window, by controlling delivery and release.

PNPs are colloidal particles that are 1-1000 nanometers (nm) in size and are made of synthetic polymers that can be biocompatible and biodegrade with no toxic byproducts. Over the course of my graduate work, I have focused on the poly(lactic-*co*-glycolic acid)–polyethylene glycol (PLGA-PEG) copolymer and its homopolymer constituents because PLGA and PEG are FDA approved and satisfy the aforementioned material requirements of a delivery platform[6,7]. Typically, the PLGA-PEG copolymer forms spherical core-shell NPs where the surface consists of a hydrophilic PEG shell and a hydrophobic PLGA core[8]. The hydrophilic PEG shell is advantageous because the therapeutic drug is shielded from neutralizing antibodies, in addition to limiting systemic clearance and particle-particle aggregation once administered *in vivo*[9]. Formulation methodologies that form this type of PNP include solvent evaporation such as nanoprecipitation and single or double emulsion techniques. Furthermore, while these two methods slightly differ in the processing conditions, both still produce drug-loaded PLGA-PEG

nanoparticles that can overcome therapeutic delivery issues. The extent of studies on these formulation methodologies have indeed yielded PNPs that have reached clinical trials, for the treatment of various cancers and other life-threatening diseases[10,11], showing that PNPs have great potential in increasing the market reach of protein therapeutics.

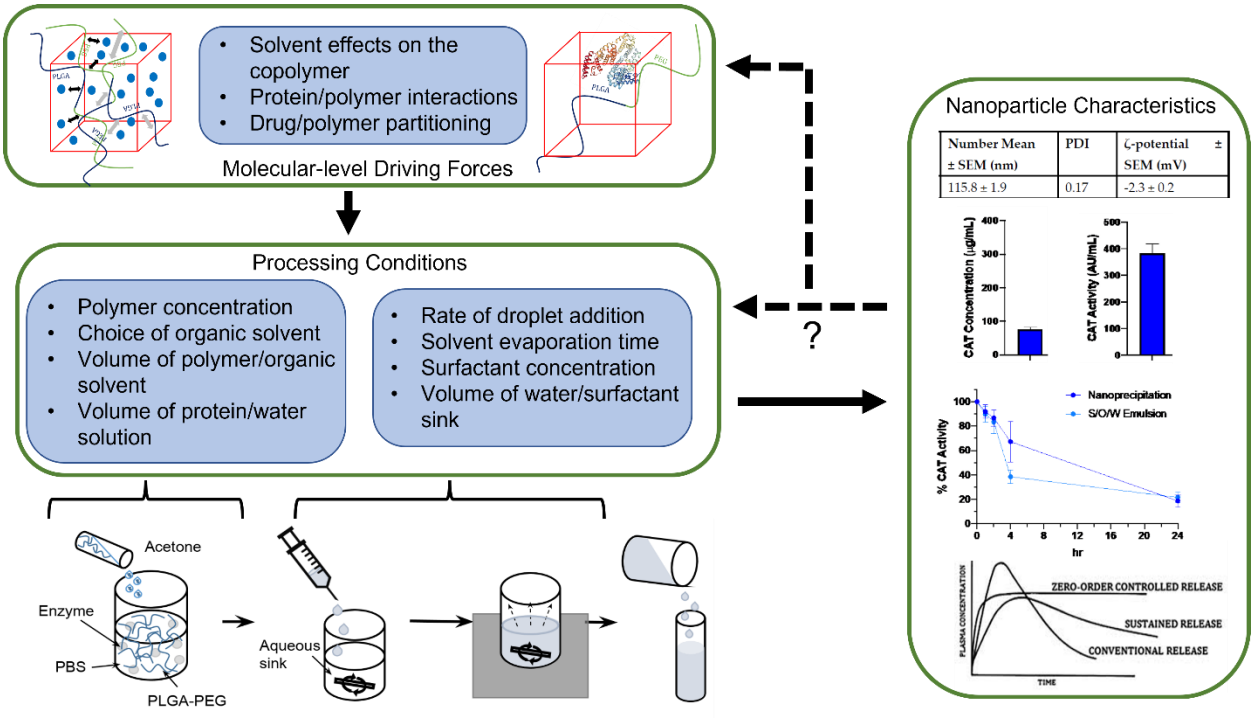

Figure 1-1: Schematic showing the processing conditions that can be altered during PLGA-PEG nanoparticle formulation and the metrics that are relevant during nanoparticle characterization. In addition, development of processing conditions that result in a specific nanoparticle drug loading and activity requires knowledge of dominant, molecular-level driving forces, present at various stages of nanoparticle formulation. Portions of this image were adopted from Joseph et al.[12] and Nautyal et al.[13]

Currently, tuning of nanoparticle characteristics such as size, drug loading or encapsulation efficiency involves changing formulation parameters until PNPs match the desired therapeutic profile *in vivo*. Using the nanoprecipitation methodology as an example, PLGA-PEG is dissolved in a water-miscible solvent at a given polymer concentration, in addition to the protein of interest being dissolved in an aqueous medium. Subsequently, those two solutions are rapidly mixed together, drawn into a glass syringe and is added dropwise to a water/surfactant solution. Once added, the sink is left mixing for a certain time period to allow for evaporation then moved to an ultracentrifuge for nanoparticle collection. As seen in **Figure 1-1**, certain processing conditions can be systematically explored at different stages within the

nanoprecipitation formulation, resulting in nanoparticle characteristics such as ~100 nm in size, monodisperse and preserved protein activity in proteolytic conditions. Using the data from nanoparticle characterization, empirical models have been developed that attempt to provide insights on the dominant thermodynamic and kinetic driving forces during nanoparticle formulation, for a certain combination of process parameters[14,15]. Researchers, however, have yet to perform inverse design of protein-loaded PNPs because the complete picture of atomic level interactions that govern protein-polymer self-assembly is still incomplete. Advancements in computational methods and hardware, over the past decades, have allowed modeling of key interactions, relevant to many technical challenges facing drug delivery. In tandem with experimental characterization techniques, rational design of protein loaded PLGA-PEG nanoparticles has become feasible, potentially revolutionizing the way PNP design and optimization is performed.

Experimental characterization of mesoscopic phenomena occurring during PNP formulation can be found in literature for a select number of drug/PLGA-PEG combinations[16–19] but atomic-level insights have been limited. Molecular Dynamics (MD), a computational technique that models the time evolution of the motion of a classical many-body system using Newton's equation of motion, is well-suited to provide such insights because it can elucidate interactions (e.g. hydrophobic, dipole-dipole, electrostatic) that can explain certain trends observed during PNP formulation. Typically, MD simulations are classical and happen at a specific thermodynamic ensemble. With further development in computational methodologies, nonequilibrium MD simulations can be also done in order to explore areas of phase space that would be inaccessible when doing classical MD. In this dissertation, MD simulations with explicit solvent, in addition to experimental characterization methods, were used to answer three fundamental questions:

1. What are the impacts of the organic solvent on copolymer structure and dynamics?
2. How does a computational scientist model and characterize relevant protein/polymer interactions at small length(< 10 nm) and timescales(< 500 ns) in a way that can result in broadly applicable insights important to "bottom-up" development of protein-loaded PLGA-PEG nanoparticles?
3. How important is the macroscopic phase behavior of the protein to drug/polymer partitioning during PNP formation?

2. Chapter 2: Methods

2.1 Computation Workflow

Determination of molecular driving forces required to perform rational design of protein loaded PNPs is non-trivial. The choice of the organic solvent used to dissolve common polymers, like PLGA-PEG, influences chain structure and dynamics. When this solution is combined with a therapeutic protein of interest, the complexity of interactions grows and are coupled to one another. Polymer/solvent and polymer/polymer interactions dictate structure and dynamics, which in turn influences protein/polymer interactions. In addition, the non-aqueous environment also influences protein-protein interactions and the overall stability of the protein. Literature contains numerous computational studies that have examined different parts of this interconnected web of key interactions that occur at different time and length scales[20,21]. Specifically, studies have either aimed to use atomistic and/or coarsed-grained MD to provide insights on polymer behavior that can validated by experimental methods or to further elucidate protein/polymer interactions that might promote or discourage stability of the modeled protein of interest[18,22–24]. Using those investigations as a starting point, my exploration of the complex interplay between solvent, polymer and protein has resulted in a unique computational workflow that will allow future researchers to further refine design principles of PLGA-PEG nanoparticles and better understand how to utilize both computational and experimental tools for PNP development and optimization.

2.1.1 Molecular Dynamics

MD is a computational technique that models the time evolution of the motion of a classical many-body system, which subsequently allows for the calculation of equilibrium and transport properties. In such a system, certain assumptions are made: no bonds are broken or formed, and each particle has a point mass and a point charge. Particles can be modeled as individual atoms on the periodic table with a Van der Waals (VdW) radius and volume (atomistic description) or as a collection of atoms (coarse-grained). All simulation work in this dissertation will mainly be atomistic MD. These atoms are characterized by their radial positions as a function of cartesian coordinates ($\vec{r_i}(x,y,z)$) and by their momenta ($\vec{p_i} = m_i \frac{d\vec{r_i}}{dt} = m_i v_i$). Atomic motion is then propagated by Newtonian mechanics, where

$$-\frac{\partial V}{\partial r_i} = F_i = m_i a_i \quad (1)$$

From equation 1, propagation of motion requires a differentiable potential energy ($V(r)$) function. In addition, a small timestep (2 femtoseconds) is used because of high frequency of bond vibrations. Periodic boundary conditions and the minimum image convention are also employed to eliminate surface effects for atoms close to the system boundary and ensure preservation of the thermodynamic ensemble[25]. These two choices are important for ensuring realistic dynamics and prevent large deviation from total system energy. Furthermore, recent advances in MD has resulted in the Hydrogen Mass Repartitioning method[26]; it allows for a 4 femtosecond timestep because a fraction of heavy atoms' mass, attached to hydrogens, is shifted to the given hydrogen, reducing the vibrational bond frequency and allowing for a faster timestep. This method was utilized for simulations after Chapter 4 and 5 of this dissertation. Most potential energy (PE) functions focus on two-body (pair) interactions and consist of different analytical functions that are used to calculate both bonded and nonbonded forces within a simulation box. Bonded forces come from bond, angle and torsional PE functions containing constants ascertained from quantum mechanical (QM) calculations and experiments, which are typically contained in a forcefield. For nonbonded forces, Lennard-Jones and coulomb's PE function are employed. Using these analytical functions, whose parameters are contained in a forcefield, allows one to simulate a given thermodynamic ensemble. For all simulations in my graduate work, the isobaric-isothermal (NPT) thermodynamic ensemble was used because key interactions during nanoparticle formulation is occurring at constant pressure and temperature.

Long-range electrostatic interactions were accounted for via Particle-Mesh Ewald (PME) summations[27] and the GROMACS MD engine[28,29] was used to set up and launch simulations. GROMACS Version 2016.3[30] was used for chapter 3 simulations and Version 2020.5[31] were used for chapter 4 and 5 simulations. For proteins, Amber99sb-ildn[32] and Amber19SB[33] were chosen for protein topological parameters used in simulations described in chapter 4 and 5. GLYCAM-06j-1 forcefield[34] was used for dextran sulphate topological parameters and the general amber forcefield[35,36] (GAFF) was used for sodium dodecyl sulphate (SDS) and taurocholic acid (TA) molecules simulated with bovine serum albumin (BSA) in chapter 4. Using an online MD preparation platform PlayMolecule[37], BSA structure at pH 3.7 was extracted, where the predicted net charge was +78. The TIP3P[38] was used to model explicit water for simulations in chapters 3 and 5 while OPC water model[39,40] was employed for chapter 4 simulations. Previously published acetone and DMSO partial charges and their associated

GAFF topological parameters were employed for simulations in chapter 3[41–44]. For any new molecules to be simulated, partial atomic charges, along with additional or topological parameters missing in the aforementioned forcefield, were extracted using the Gaussian09[45] for quantum mechanical (QM) calculation, the residual electrostatic potential fitting method[46] and the Antechamber/GAFF framework (ambermd.org/antechamber/ac.html) for generating residue topology files with transferable Amber parameters. After this step, a charge scaling procedure, seen in equation 2 below, was performed and ensured a net zero charge for all molecules or oligomers of varying monomer length. Scaled RESP charges used for polymer simulations in chapters 3 and 4 can be seen in **Tables B-1 – B-3**, along with representative oligomer structures used for QM calculation and RESP charge fitting shown in **Figures B-1 – B-3**. Lastly, nonequilibrium simulations, employed in chapter 4, used an additional harmonic PE function (equation 3) to explore the effect of polymer extension on protein/polymer interactions. *k* is the scaled force constant that was set at 140 $\frac{\text{kJ}}{\text{mol}}$. ε_o is the center of the radius of gyration restraint and was set at a chosen R_g value, for oligomers at a certain level of extension; w_ε is the width and had a default value of 1 nm^2. Temperature control was achieved using the Bussi-Donadio-Parinello thermostat[47], where the time constant was 0.1 ps. Pressure control was achieved using the Parrinello-Rahman Barostat[48].

$$q_{i,mod} = q_{i,RESP} - \frac{\sum_{i=1}^{N_{atoms}} q_{i,RESP}}{N_{atoms}} \quad (2)$$

$$V(\varepsilon) = \frac{1}{2} k \left(\frac{\varepsilon - \varepsilon_o}{w_\varepsilon}\right)^2 \quad (3)$$

Each simulation, following the packing of molecular species into a cubic box with a pre-specified side length using PACKMOL[49], first underwent a steepest descent energy minimization step and two equilibration periods at constant volume and temperature (NVT) followed by constant pressure and temperature (NPT), while holding system species in place. NVT and NPT equilibration period was set at 1 ns for all systems simulated. The leap-frog integrator was used for time-step integration. Coordinates, velocities, and energies were saved every 10 picoseconds. Short-range coulombic and Van der Waals interactions employed the Verlet cutoff-scheme, along with a cut-off radius of 1 nm. LINCS[50] was used as the main constraint algorithms for bonds. MDAnalysis[51] was the main python package used for analyzing all simulations, alongside in-house codes.

2.1.2 Methods for polymer structure and dynamics analysis

Calculation of structural properties for all simulations was performed in chapter 3, similar to previous computational studies[52,53]**.** Instead of considering each atom along the backbone when calculating key properties, we focus on the center-of-mass (COM) of each monomer. Hence, our analysis viewed each polymer chain as *N* beads connected by $N - 1$ bonds, where the bond length (L_{COM}) was the time-averaged distance between the 1st and 2nd monomer COM. Each oligomer can furthermore be characterized by its maximum length, at full extension, called the contour length (L_C). Calculation of the contour length requires knowledge of bond lengths, provides an upper length limit for each oligomer, and represents the theoretical maximum end-to-end distance. For the homopolymers, L_C is calculated as seen in equation 4 while for the copolymer, its calculation is described in equation 5. Bond length values were unique for the PLGA ($L_{COM} = 5.4 \pm 0.7$ Å) and PEG ($L_{COM} = 3.2 \pm 0.2$ Å*)* systems, regardless of chain length. As for the PLGA-PEG, the monomer's COM distance for the PLGA and PEG block were similar, with the exception of the LGA:EG linker with an intermediate value of $L_{COM} = 4.2 \pm 0.4$ Å.

$$L_C = (N - 1)L_{COM} \tag{4}$$

$$L_C = (N_{PLGA} - 1)L_{PLGA} + L_{linker} + (N_{PEG} - 1)L_{PEG} \tag{5}$$

Radius of gyration (R_g), end-to-end distance (R_{ee}), hydrodynamic radius (R_{hyd}), mean squared end-to-end distance (R_{ee}^2) and mean squared radius of gyration (R_g^2) were extracted from all polymer-solvent systems, ensemble-averaged (denoted by $\langle \dots \rangle$) and used to examine oligomer shape in solution. R_g was calculated using equation 6, where $\boldsymbol{r}_i$ is the position of the i^{th} monomer's COM, *M* is the mass of an oligomer, m_i is the mass of the i^{th} monomer and $\boldsymbol{R}_{CM}$ is the position of each oligomer's COM. R_{ee}^2 was calculated using equation 7, where $\boldsymbol{r}_1$ is the position of the 1^{st} monomer's COM and $\boldsymbol{r}_N$ is the position of the last monomer's COM for a given oligomer. R_{hyd} was calculated using equation 8, where r_{ij} is the distance between the *i* and *j* monomer's COM and *N* is the number of monomers for a given oligomer. Using equation 9, deviation from non-ideal chain behavior was studied, since an ideal linear chain is predicted to have a theoretical $\frac{\langle R_{ee}^2 \rangle}{\langle R_g^2 \rangle}$ value of 6, in dilute solution as $N_{monomer}$ goes to infinity. Equation 10 was also useful in understanding polymer shape in solution, instead of just plotting R_g or R_{ee} versus (vs.) time.

$$R_g = \sqrt{\frac{1}{M}\sum_{i=1}^{N} m_i(\boldsymbol{r}_i - \boldsymbol{R}_{cm})^2} \quad (6)$$

$$\langle R_{ee}^2 \rangle = \langle (\boldsymbol{r}_N - \boldsymbol{r}_1)^2 \rangle \quad (7)$$

$$\langle \frac{1}{R_{hyd}} \rangle = \frac{1}{N^2}\sum_{i=1}^{N-1}\sum_{j>i}^{N} \langle \frac{1}{\mathrm{r}_{ij}} \rangle \quad (8)$$

$$\frac{\langle R_{ee}^2 \rangle}{\langle R_g^2 \rangle} \; vs. \; L_c \quad (9)$$

$$\frac{\langle R_g \rangle}{\langle R_{hyd} \rangle} \; vs. \; L_c \quad (10)$$

Persistence (L_p) and Kuhn length (L_k) were determine as follows[54]: after $\langle R_{ee}^2 \rangle$ is calculated at each contour length value, a nonlinear fitting procedure, using equation 11 and the python package LMFIT[55] (https://lmfit.github.io/lmfit-py/), was used to arrive at one L_p value for each polymer/solvent system. LMFIT provides standard errors and relative errors, which were reported alongside the fitted L_p value. As for L_k, the relationship between $\langle R_{ee}^2 \rangle$ and L_C in equation 12, as provided from the freely jointed model[54] was used, where L_k is the slope of the line.

$$\langle R_{ee}^2 \rangle = 2L_p^2 \left[e^{-\frac{L_c}{L_p}} - 1 + \frac{L_c}{L_p} \right] \quad (11)$$

$$\langle R_{ee}^2 \rangle = L_k L_c \quad (12)$$

Polymer structure properties can provide an important baseline to understand copolymerization effects in different solvents, but evaluating dynamical behavior is important in learning how a polymer solution responds to stresses under certain timescales, ranging from nanoseconds to hours. In chapter 3, the positional bead autocorrelation, as seen in equation 13, is used to extract the time relaxation spectrum of each oligomer in solution; N is the number of beads, $\boldsymbol{r}_i$ denotes the position of bead i, $\boldsymbol{R}_{cm}$ denotes the position of the center of mass of the chain, and t denotes time. The initial value of $C^{pos}(t=0)$ is unity because $\boldsymbol{r}_i - \boldsymbol{R}_{cm}$ at time 0 and *t* is normalized, prior to taking the dot product. If $\boldsymbol{r}_i - \boldsymbol{R}_{cm}$ is not normalized, $C^{pos}(t=0)$ will yield R_g^2. A running average procedure was used for ensemble averaging in equation 13, since block averaging does not yield enough samples for noise reduction.

Experimental studies of dilute polymer solutions are interested in the longest relaxation time because this time contains information on important viscoelastic and rheological

properties[56–59]. With computational methods like atomistic MD, previous studies have used equation 13 to probe the relaxation times and hydrodynamic/excluded volume effects of polymer-solvent systems that were investigated[59,60]. Typically, the 1st relaxation time (τ_1) is of primary interest for polymers in dilute solution because sub-chains within a singular chain will have different relaxation modes following uniform deformation of the solution. The manner in which dilute polymer chains relax, however, depends on the relaxation mode of the entire chain, which tends to be the slowest; hence this leads τ_1 to be the longest of all times in the relaxation spectrum and provides insights into the elastic behavior of polymer solutions. Once the positional bead autocorrelation is calculated, the averaged data is fit to equation 14 by using a nonlinear fitting procedure contained within the Python LMFIT package[55]. Equation 14 can be fit to the data by changing τ_1 and h^*. However, h^* was held constant at $h^* = 0.26$ and only τ_1 was varied, similar to previous studies. Equation 15 shows the Zimm relaxation times; if equation 16 resulted in $\tilde{a}_i = a_i$ (where $\beta = 1$ and $\sigma = 0$), then eqn. 15 will yield the Rouse relaxation times.

$$C_{pos}(t) = \frac{1}{N}\sum_{i=1}^{N}\langle[\boldsymbol{r}_i(t) - \boldsymbol{R}_{cm}(t)] \cdot [\boldsymbol{r}_i(0) - \boldsymbol{R}_{cm}(0)]\rangle \quad (13)$$

$$C^{pos}(t) = \frac{6}{N^2-1}\sum_{i=1}^{N-1}\frac{1}{a_i}e^{-\frac{t}{\tau_i}} \quad (14)$$

$$\tau_i = \left(\frac{\tilde{a}_1}{\tilde{a}_i}\right)\tau_1 \quad (15)$$

$$\tilde{a}_i = a_i\beta\left(\frac{i}{N}\right)^{\sigma} \quad (16)$$

$$a_i = 4sin^2\left(\frac{i\pi}{2N}\right) \quad (17)$$

$$\beta = 1 - 1.66(h^*)^{0.78} \quad (18)$$

$$\sigma = -1.40(h^*)^{0.78} \quad (19)$$

Scaling of R_g and τ_1 with number of monomers was performed to determine the Flory exponent (v), which provides a measure of oligomer self-similarity, and the alpha exponent (α), which is used to evaluate hydrodynamic and excluded volume effects. They were determined by fitting a line to $log_{10}(R_g)$ vs. $log_{10}(N)$ and $log_{10}(\tau_1)$. vs. $log_{10}(N)$ data for each polymer/solvent system and extracting the slope (equations 20 and 21). Lastly, the arrangement of solvent around a given oligomer chain is important to seeing the main role of polymer-solvent interactions and was analyzed using the radial distribution function (RDF) coded into the GROMACS software[31].

$$\langle R_g \rangle \propto N^{\nu} \tag{20}$$

$$\tau_1 \propto N^{\alpha} \tag{21}$$

2.1.3 Methods for characterization of protein/polymer interactions

Prior to extracting the molecular driving forces behind protein-polymer binding, protein structure over the course of a MD simulation was examined to ensure that solvent accessible residues are the ones primarily interacting with the polymer. Root mean squared deviation (RMSD) of a protein's backbone from its crystal structure, ascertained from the RCSB protein databank, was the main metric used in simulations containing a protein. *N* is the total number of backbone heavy atoms, where δ_i indicates the distance between atom *i* and its reference position within the crystal structure (equation 22 below). Furthermore, all simulations containing a protein were simulated within a pure water medium only because the atomistic forcefields chosen for my graduate work have been optimized to reproduce experimental measurements of protein structures in aqueous environments[33,61,62].

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \delta_i^2} \tag{22}$$

Following a completed simulation of a singular protein and polymer oligomers solvated in explicit water, the distance between the center of mass (COM) of the protein and each polymer oligomer (d_i^{COM}, where i is the oligomer identifier) was used to understand the general extent of protein/polymer interactions, over the course of a trajectory. Similarly, the distance between the COM of each oligomer to another oligomer (d_{i-j}^{COM}, where i is the oligomer identifier and j is the next oligomer where $i \neq j$) was examined to understand the extent of polymer-polymer interactions. These distances, over simulation time, were extracted using open-source, community-developed PLUMED library[63], version 2.7.1[64] and used to establish four classes of interactions: protein/polymer interactions only (prot-poly int. only), polymer/polymer interactions only (poly-poly int. only), both protein/polymer and polymer/polymer interactions (both prot-poly/poly-poly int. only) and lastly, free in solution (free in soln). Prot-poly interactions were defined as $d_i^{COM} \leq 4.3$ nm, due to a protein's general shape and length of simulated oligomers. Protein/polymer simulations within my dissertation primarily focused on BSA and iduronate-2-sulfatase (ID2S), an enzyme involved in lysosomal storage disease[65], in addition to polymer

oligomers that have a contour length of ~10 nm. As an example, ID2S shape, seen in **Figure 2-1**, can be approximated as an ellipsoid using measuring the distance between the protein's COM and selected surface residues, where the largest semi-axial length is ~3.5 nm. Observations of multiple trajectories and examination of different cutoff distances showed that 4.3 nm is a reasonable distance because it accounts for transient prot-poly interactions that can occur between different parts of individual oligomers and ID2S.

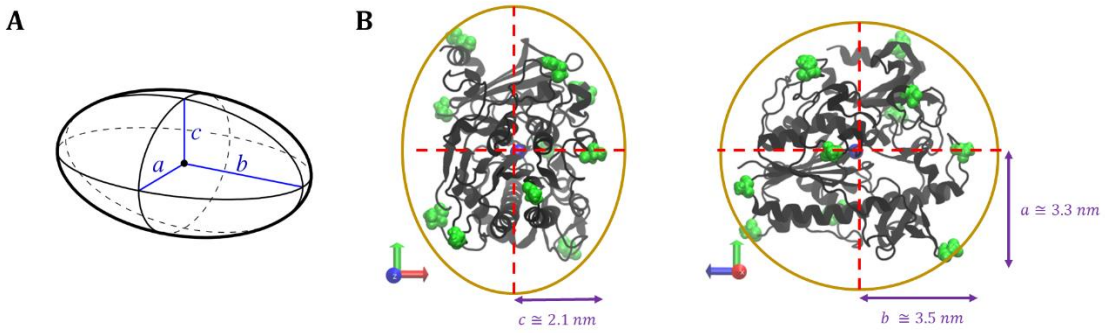

Figure 2-1: (A) General shape of a three-dimensional ellipsoid, *a, b* and *c* are the semi-axial lengths. **(B)** Iduronate-2-sulfatase's approximated ellipsoidal shape where $a \cong 3.3\ nm, b \cong 3.5\ nm$ and $c \cong 2.1\ nm$. The blue sphere in the center of ID2S is the protein's center of mass and surface residues colored in green are surface residues used to determine the semi-axial lengths.

Poly-poly interactions were defined as $d_{i-j}^{COM} \leq 2.3\ nm$ to account for interactions that occur over a wide range of oligomer conformations; collapsed conformations, at the contour lengths present in protein/polymer simulations in chapter 4, were measured to have R_g values less than 1.2 nm whereas extended conformations were measured to have R_g values greater than 1.8 nm. In addition, initial configurations placed oligomers a distance greater than 3.5 nm away from each other. Hence, this cutoff distance was chosen to ensure polymer-polymer interactions were occurring and allowed for analysis of poly-poly binding at different $R_g$values. Both prot-poly/poly-poly interactions were defined when $d_i^{COM} \leq 4.3$ and $d_{i-j}^{COM} \leq 2.3\ nm$ whereas free in soln. was defined when neither of those two distance cutoffs were met. By categorizing simulation trajectory observations into these 4 interaction classes, the general interplay between protein-polymer and polymer-polymer binding can be better understood and provides a path for data generation and compilation for machine learning algorithms to predict the most favorable protein/polymer combinations for the purpose of protein loading within PNPs.

To understand how often each protein-polymer simulation spent in each of the interaction classes, the time frequency (TF) of each oligomer's interaction class was calculated and used to examine how trial simulations with different initial configurations explored phase space. As seen in equation 23 and 24, i is the oligomer identifier and N is the total number of oligomers. This metric was extracted for each oligomer and averaged (TF_{avg}) to allow for easy data visualization and aid analysis. TF_{avg} values from each simulation trial with the same polymer type was averaged again to evaluate the differences in protein binding, across polymers with different chemical moieties (Equation 25).

$$TF_i = \frac{\text{No. of frames oligomer i spent in a specific interaction class}}{\text{No. of frames in a simulation trajectory}} \tag{23}$$

$$TF_{avg} = \frac{\sum_{i=1}^{N} TF_i}{N} \tag{24}$$

$$TF_{avg}^{poly} = \frac{\sum_{i=1}^{N_{trials}} TF_{avg}^{i}}{N_{trials}} \tag{25}$$

The strength of binding between either the protein and polymer or different polymer oligomers was examined using GROMACS 2018.5 analysis toolkit[66] to calculate the energy contributions of various nonbonded interactions. Typically, London dispersion and repulsive forces are captured by the Lennard-Jones potential energy function while electrostatic forces are described by Coulomb's law[28]. This information is valuable because it can show whether protein/polymer binding is more favorable than polymer/polymer binding, when varying polymer types or oligomer conformation. In addition, results from this analysis can help explain why a certain protein/polymer interaction interface undergoes more irreversible binding events across simulation trials.

In order to understand the residency time of polymer oligomers near a set of protein surface residues over the course of a simulation, percent occupancy was calculated, for each amino acid residue, by counting the number of frames a polymeric monomer was within 4 angstroms (Å) of a protein amino acid surface amino acid (AA), divided by the total number of frames within the entire trajectory (Equation 26).

$$\%\ \text{Occupancy} = \frac{\text{No. of frames a polymeric monomer was within 4 Å of an AA residue}}{\text{No. of frames in a simulation trajectory}} \times 100 \tag{26}$$

Results from this analysis was used to define a protein/polymer interaction interface from each simulation trial that had a different initial system configuration. Since some trials can result in unique binding modes, the prot-poly interface from each trial can be aggregated into a collapsed interface by sorting repeating and non-repeating residues with non-zero occupancies into a new dictionary and averaging occupancy values for repeating residues. This was done using Python 3.6[67], similar to most analyses done in this work, and allowed for a more streamlined analysis workflow. Subsequently, the elimination of low occupancy residues can be done to identify sets of surface amino acids that had long polymer residence times. The cutoff values of percent occupancy that were used for low occupancy AA elimination in chapter 4 analysis have been 50% and 90%. Residues that have occupancy values greater than 50% were observed to undergo favorable binding with the simulated polymer, while also accounting for reversible binding events. Residues with occupancy values greater than 90% allow for the identification of specific residues that experienced irreversible binding but ignore the local chemical surface environment near high residency AAs that likely contribute to protein-polymer binding. Once these filtered residues have been extracted, depending on the cutoff occupancy value, they can be grouped into 5 categories: negative, positive, polar, hydrophobic and aromatic. The number of residues in each grouping was then divided by the total number of residues that have occupancy values greater than the cutoff, in order to understand the AA distribution of the sub-selected protein-polymer interface (Equation 27).

$$\text{Fraction of residues with } >X\% \text{ occupancy} = \frac{\text{No. of residues in an AA grouping with}>X\% \text{ occupancy}}{\text{Total no. of residues with}>X\% \text{ occupancy}} \quad (27)$$

For rational design of protein-loaded nanoparticles to occur successfully, an in-depth knowledge of the protein surface chemistry is required but methodologies to ascertain such information have varied greatly and no standard has been established to perform this investigation systematically[21]. Filtering surface amino acids via calculation of percent occupancies is useful for identifying a binding interface but does not consider the effect of residue clusters on protein-polymer interactions. Taking inspiration from Jones et al.[68] in their analysis of interaction sites within protein-protein complexes, a surface patch analysis was used to more strictly examine the collapsed protein-polymer interface and determine: (1) What makes this interface different from the rest of the protein surface? (2) What is the AA composition and average number of residues per cluster involved in irreversible polymer binding? (3) Are high residency residue clusters unique in their physio-chemical properties (hydrophobic, polarizability, planarity, etc.), when compared to other patches on the protein surface?

Surface patch analysis applied to protein/polymer simulations consisted of generating N overlapping, contiguous surface patches with a varying number of surface AAs. Surface residues were determined using the VMD plugin[69] to measure the solvent accessible surface area (SASA) of all residues within a given protein. Residues with a SASA value $> 1\text{Å}^2$ were considered surface residues because water molecules have an approximate radius of 1.4 Å[70] and are likely to interact with these selected AAs**.** Moreover, these residues were used to calculate the protein-polymer contact surface area (CSA) as seen in equation 28 below, where $N_{interface}$ is the number of residues in the protein-polymer interface and $SASA_{AA,i}$ is the SASA value of an amino acid in the interface.

$$\text{Protein-Polymer Contact Surface Area} = \sum_{i=1}^{N_{interface}} SASA_{AA,i} \quad (28)$$

For each surface residue, a solvent vector ($\vec{V}_{sol}$) was calculated to ensure that residues on the opposite face or rings of residues are not selected in a surface patch. This vector is defined by first finding the C-alpha (C_α) atoms of the n nearest surface residues using a given surface AA's C_α atom as the center of the search. MDAnalysis Neighbor Search Wrapper[51] was utilized to perform atom and residue nearest neighbor search. Specifically, the MDAnalysis search wrapper could not always provide the same number of nearest neighbors for each residue, since the user has to provide a range of radial values from which the algorithm uses to execute its search. However, a range of radial values was determined such that most surface residues had 10 nearest neighbors, with the remaining AA's having 11 or 12 near neighbors (See **Figure 2-2**). Using these atoms, the COM was calculated and used to define a vector between the central surface residue and the COM of its n nearest neighbors ($\vec{V}_{CR-COM}$); this vector is normalized to aid determination of surface patches. The inverse of normalized $\vec{V}_{CR-COM}$ results in $\vec{V}_{sol}$, as seen in **Figure 2-2**.

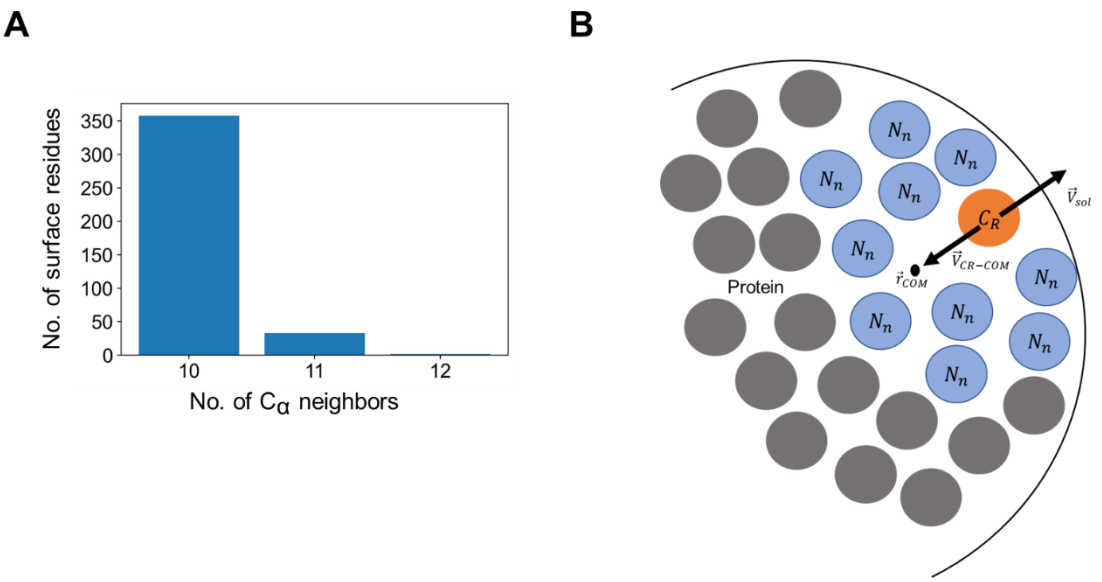

Figure 2-2: (A) Using the MDAnalysis Neighbor Search Wrapper with a search radius of 10 Å, the number of surface residues with 10, 11 or 12 C_α atom neighbors, for ID2S, was calculated and plotted **(B)** Schematic showing the determination of a solvent vector($\vec{V}_{sol}$) for each surface residue. Each sphere represents a C_α atom of a surface residues. C_R is the central surface residue whose solvent vector is being calculated and N_n is the n nearest atom neighbor. $\vec{r}_{COM}$ is the location of the center of mass of the n nearest neighbors and $\vec{V}_{CR-COM}$ is the vector used to define $\vec{V}_{sol}$

To define an overlapping surface patch for each surface residue, a nearest residue neighbor search is performed on the surface AAs, employing a search radius of 13 Å. This also resulted in the same distribution of nearest neighbors seen in **Figure 2-2**, where most surface residues also had 10 nearest residue neighbors. Using the solvent vectors and a given surface residue as the center of the patch, an adjacent nearest residue neighbor is included in a patch if the angle between its $\vec{V}_{sol}$ and the central residue's $\vec{V}_{sol}$ is less than a chosen cutoff value (θ_{cut}). For ID2S, θ_{cut} was chosen to be 125^o after examining the resulting distribution of the number of residues per patch while varying θ_{cut}. Furthermore, most surface patches with $\theta_{cut} = 125^o$ consisted of 9 residues, varying in morphology and AA composition (**Figure A-1**). Next, 4 physiochemical descriptors for all 20 natural amino acids, experimentally measured by Fauchere et al.[71], were used to characterize the physiochemical nature of each overlapping surface patch; they are hydrophobicity (HB_{AA}), polarizability (PB_{AA}), graph shape index (GSI_{AA}) and normalized Van der Waals ($nVdW_{AA}$) volume. HB_{AA} is described using water-octanol partition coefficients and GSI_{AA} encodes for complexity, branching and symmetry, providing a measure of the steric influence for a given patch. PB_{AA} is a measure of how easy a residue's electron cloud can be distorted, due to the presence of induced or permanent charge dipoles.

$nVdW_{AA}$ was useful for analyzing why an interface may experience more dispersion forces when comparing across polymer types since residues with larger $nVdW_{AA}$ can lead to greater VdW forces. Calculation of each surface patch physiochemical property can be seen in equations 28-31, where $N_{r,patches}$ is the number of residues in a surface patch. Using these patch properties that capture the local chemical nature of defined residue clusters and examining their overall distribution (**Figure 2-3**), the characterization of the collapsed protein-polymer interface can be performed and compared to the other surface patches.

$$HB_{patch\ i} = \frac{\sum_{i=1}^{N_{r,patches}} HB_{AA,i}}{N_{r,patches}} \tag{28}$$

$$PB_{patch\ i} = \frac{\sum_{i=1}^{N_{r,patches}} PB_{AA,i}}{N_{r,patches}} \tag{29}$$

$$GSI_{patch\ i} = \frac{\sum_{i=1}^{N_{r,patches}} GSI_{AA,i}}{N_{r,patches}} \tag{30}$$

$$nVdW_{patch\ i} = \frac{\sum_{i=1}^{N_{r,patches}} nVdW_{AA,i}}{N_{r,patches}} \tag{31}$$

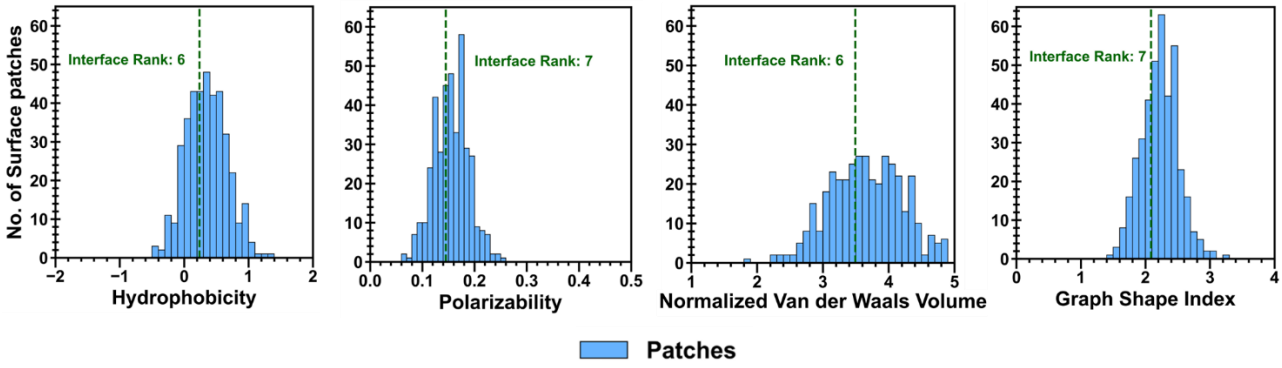

Figure 2-3: The distribution of the four physiochemical descriptors for ID2S surface patches (in blue) are plotted, in addition to the parameter ranking and descriptor values (green dashed line) for a protein/polymer interface. For example, hydrophobicity of the selected interface is ~0.22, meaning it is in the 60-70% range of the hydrophobicity patch distribution and therefore gets a rank of 6. This comparison of the physiochemical descriptors between a sample interface and the generated surface patches was valuable in characterizing differences in protein/polymer binding.

The same physiochemical descriptors, calculated for each surface patch, were also extracted for each protein/polymer interface ascertained from a simulation trial; they were

determined the same way as seen in equations 28-31, where $N_{r,patches}$ is replaced with $N_{interface}$, the number of residues in the collapsed interface. These descriptors were ranked on a scale from 1 to 10, relative to the overlapping surface patches. A rank of 1 means the prot-poly interface scores in the highest 10% range of a physiochemical descriptor distribution of all surface patches; whereas a rank of 10 means the prot-poly interface scores in the lowest 10% range[68]. This ranking provides a way to compare the prot-poly interface to other surface patches (**Figure 2-3**). The distribution of parameter interface ranks from multiple simulation trials allows for the overall evaluation of physiochemical nature of the protein/polymer interface for a given polymer type and level of extension, as each protein/polymer simulation explores a new region of phase space. Furthermore, such information can be utilized to understand how polymer chemistry may result in an observed interface with a certain parameter rank. As for the unique collapsed interface whereby residues with occupancy values greater than 50% were sub-selected, percent patch overlap was calculated to determine which surface patches had the largest intersection with the collapsed interface, which contain most of the AA's with high (>90%) occupancy or long residence time of the polymer; moreover, the effect of residue clusters on polymer binding can be examined using this metric. Equation 32 below shows how it was calculated, where $N_{r,patch\ i}$ is the number of residues in surface patch i.

$$\%\ \text{Patch Overlap} = \frac{\mathrm{N_{r,\ patch\ i} \cap N_{interface}}}{\mathrm{N_{interface}}} \times 100 \tag{32}$$

2.2 Experimental Workflow

2.2.1 Materials

PLGA (50:50, MW 45kDa) copolymerized to methoxy PEG (mPEG, MW 5kDa) (45k:5k PLGA-PEG), was purchased from Akina PolySciTech, and used for nanoparticle formulations. Acetone is purchased from Fisher Scientific. Polysorbate 80 (P80), 1x phosphate buffered saline, Catalase (Cat) from bovine serum, lysozyme (Lys) were ascertained from Sigma-Aldrich (St Louis, MO). Dextran sulphate (DS, Mr 5000) used for hydrophobic ion pairing, was purchased from Millipore Sigma. All reagents and solvents were used as received. Pierce Bicinchoninic Acid (BCA) Assay kit provided by Thermo Fisher Scientific to quantify protein mass. Pronase, a protease from *Streptomyces griseus*, was bought from Sigma Aldrich to

quantify nanoparticle protection, along with hydrogen peroxide and 50 mM phosphate buffer (PB) for protein activity measurements.

2.2.2 Formulation of Protein-Loaded PLGA-PEG nanoparticles

PNPs were formulated using the solvent-evaporation method called nanoprecipitation, as seen below in **Figure 2-4**. To start, PLGA-PEG copolymer is dissolved in acetone, a water-miscible solvent, at a concentration of 25 mg/mL. The protein of choice is dissolved in 1x phosphate buffered saline (PBS) at 1 mg/mL. 0.1 mL protein solution is added to 0.9 mL polymer solution, rapidly mixed, then is added dropwise into 25 mL 1% w/v P80 water/surfactant sink. The volumes of the protein and polymer solution can be adjusted but the total protein/polymer mixture volume is set to 1 mL. Once added, nanoparticles are formed spontaneously within the sink and stirred continuously for 3 hours at 500 RPM to allow for acetone evaporation. Nanoparticles were then collected via ultracentrifugation at 100,000 RCF for 1 hour at 4°C. 1 mL of remaining water/surfactant sink is collected for drug loading evaluation. The resulting pellet is resuspended in 10 mL de-ionizied (DI) H_2O and recollected at 100,000 RCF for 25 minutes at 4°C; this step is done twice to ensure removal of residual organic or surfactant. The final pellet is resuspended in 1 mL 1x PBS or DI H_2O and stored at 4°C for short term storage or lyophilized for long-term storage.

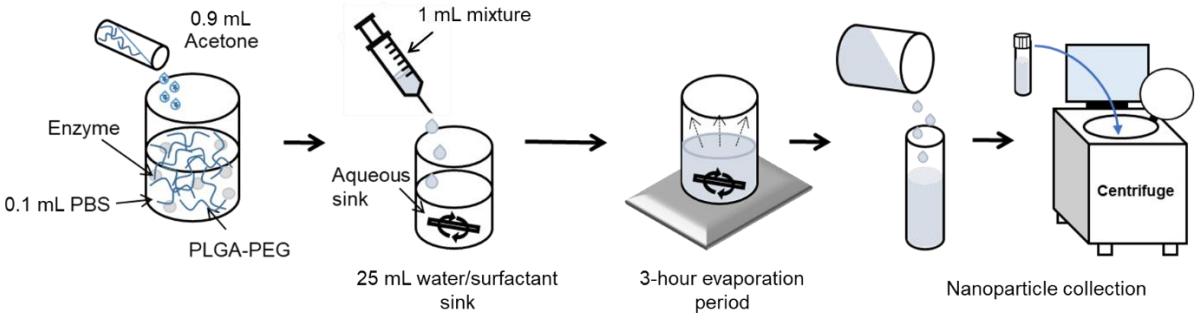

Figure 2-4: Schematic showing the nanoprecipitation formulation, starting from addition of polymer/acetone solution to the aqueous enzyme mixture and ending with the final collection of formed nanoparticles, via ultracentrifugation.

2.2.3 Characterization of NP size polydispersity and surface charge

After the protein-loaded nanoparticles have been resuspended, hydrodynamic diameter and polydispersity (PDI) were measured by dynamic light scattering, using the Zetasizer Nano ZS (Malvern Instruments, Malvern, UK). Surface charge was determined by measuring the zeta potential, using the same Zetasizer Nano ZS instrument. Prior to triplicate size and zeta potential measurements, nanoparticle samples undergo a 1000x dilution in 10 mM NaCl aqueous solution at pH 7.4. The same procedure was used to characterize hydrophobic ion complexes formed between cat and DS; methodologies to form these complexes is described in section 2.2.5

2.2.4 Activity measurement of free catalase and loaded within PLGA-PEG NPs in proteolytic conditions

Cat-loaded nanoparticles and Cat/DS complex coacervates are checked for protein activity. Specifically, UV-Vis spectrophotometer, along with the enzymatic assay of catalase provided by Sigma Aldrich, were used to determine how fast the protein is converting hydrogen peroxide to water and oxygen gas, by measuring the decrease of absorbance at 240 nm. All measurements take place in a 3.00 mL reaction mix within a cuvette reader, where the final concentrations are ~50 mM PB, 0.036% (w/w) hydrogen peroxide, and ~10 units of catalase/mL. The activity standard is done with free catalase in solution. The nanoparticle sample is not degraded, prior to measurement of protein activity. To evaluate protein protection over time from protease degradation, cat-loaded nanoparticles were first incubated with pronase (0.2 mg/mL concentration) at 37°C. At specific timepoints, 200 μL is collected from the pronase/nanoparticle solution and protein activity is measured through the same enzymatic assay.

2.2.5 Preparation and Construction of protein-DS complexation state point diagram

The method outlined in this section allows for the determination of the set of pH and molar ratio (MR) values that result in water-insoluble, polyelectrolyte coacervates formed between an enzyme and a poly-anionic or poly-cationic polymer. To clarify, hydrophobic ion pairing is typically used as an umbrella term describing complexation between a protein and charged surfactant, known as surfactant-based coacervation, or charged polymer, known as

macromolecular/polyelectrolyte coacervation. Both types of coacervation can result in water-insoluble complexes. DS was chosen as the poly-anionic polymer to be investigated while catalase and lysozyme were the enzymes employed for complexation. Protein solutions at 10 mg/mL (1 wt. %) were prepared by dissolving the protein solid in either acidic or basic DI water; DS solutions were made in a similar fashion. Knowledge of the pKa of the polymer's ionizable moieties and the protein's isoelectric point (pI) is critical for determining which set of pH values that should be explored. The ionizable groups for DS are sulphates, where their pKa is approximately 2. Catalase (240 kDa) has a reported pI of 5.4[72] whereas lysozyme (14 kDa) has pI of ~11[73]. Therefore, catalase/DS complexation exploration was restricted to pH 2-5.4 because DS sulphate groups will be de-protonated at pH values above 2 (pH > pKa) and catalase will have a net positive charge (pH < pI). However, lysozyme/DS complexation was restricted to pH 7-10.5 because lysozyme will have a net positive charge within that range. If a poly-cationic polymer with ammonium ionizable groups (pKa ~ 10.5), like poly-lysine, is used, then the range of pHs that can be explored is shifted. For example, catalase/poly-lysine complexation exploration would occur between pH 5.4-10 because ammonium groups will be protonated (pH < pKa) and positively charged, while catalase would have a net negative charge (pH > pI). For lysozyme, electrostatic complexation with poly-lysine would not occur because lysozyme would be negatively charged at pH above 11 (pH > pI) and poly-lysine's ionizable groups would have a net neutral charge at pH > 11, following the deprotonation of ammonium groups (pH > pKa). Prior to attempting complexation exploration, measurement of enzyme activity using activity assays and structure using circular dichroism, across a wide range of pHs, can be performed in order to determine which pH ranges result in water-insoluble coacervates that have preserved enzymatic activity, similar to free protein in aqueous solution.

The range of pH values that may be explored are either acidic or basic. Aqueous stock solutions at different pH values were made fresh and measured prior to dissolution of measured protein mass. pH measurement was performed using the Fisherbrand accumet AB150 pH benchtop meter, where a pH reading was stable for a time period greater than 5 seconds prior to recording the given value. Using $pH = -\log_{10}([H^+])$, water solution at a specific hydrochloric acid (HCl) concentration was made using 1 N HCl and DI water. On the other hand, using $pH = 14 - [-\log_{10}([OH^-])]$, water solution at a specific sodium hydroxide (NaOH) concentration was made using 1 N NaOH and DI water. To explore pH 2-7, HCl concentrations would range from 10-0.0001 mM (**Figure A-2**) whereas NaOH concentrations with the same range (10-0.0001 mM) is expected result in pH values 12-7. 1 mM HCl or DI water at pH 7 was primarily

used to lower or increase pH if the HCl stock solution was slightly above or below the desired acidic pH. Similarly, 1 mM NaOH or DI water at pH 7 was used to increase or decrease pH if NaOH stock solution was not at the desired basic pH. pH values closer near 7 were the most difficult to adjust and, depending on the day, DI water may be slightly acidic or basic.

Knowledge of the protein's and charged polymer's molecular weight (MW) was important for establishing which MR values should be explored. For example, catalase has a MW of 240 kDa whereas DS has MW of ~5000 Da, meaning most polymer chains are 12 monomers long. Using Microsoft Excel to perform conversions and setting the catalase concentration to 10 mg/mL and DS concentration to 50 mg/mL, different DS:Cat MR values were generated by determining the necessary volumes of DS solution to be added to a specified volume of protein solution, such that dilution effects were minimized. For lysozyme, DS concentration was increased to 500 mg/mL, while keeping the protein concentration at 10 mg/mL, to ensure DS volumes added to a 7 mL protein solution mixing at 500 RPM were small (**Table 2-1**). A 7 mL reaction volume was chosen and held constant across different MRs because mixing at a specific RPM using a stir plate and bar could be easily done in a 30 mL beaker. At high protein molecular weights, a lower DS concentration was required whereas low protein MWs require higher DS concentrations, assuming DS MW is 5000 Da. An increase in the charged polymer's MW would also influence what MRs can be explored. Following the addition of DS volume to the protein solution, the mixture was left to mix for 1 hour to ensure well-mixed conditions. In addition, 2 mL were saved for complex characterization in two 1.5 mL Eppendorf tubes and the remaining 5 mL were placed in a 20 mL test tube. 1 mL was immediately characterized by DLS and the remaining 6 mL were placed in a 4^{o}C refrigerator for 24 hours to allow solutions at different MRs to reach phase equilibrium. A catalase or lysozyme control was used to aid comparison of protein:DS mixtures to protein only solution. Smaller reaction volumes can be employed if protein mass is expensive but mixing conditions should be adjusted accordingly.

Table 2-1: Volumes of DS solution, added to a 7 mL protein solution at 10 mg/mL, are reported corresponding to a given DS/protein molar ratio. For lysozyme, DS concentration was 500 mg/mL while for catalase, DS concentration was 50 mg/mL. Even at the largest MR values, no more than 0.5 mL (500 µL) of DS solution was added to the protein solution, minimizing dilution effects.

DS:Cat MR	DS volume [uL]
1	29
3	88
5	146
7	204
10	292
13	379

DS:Lys MR	DS volume [uL]
0.5	24.5
1	48.4
3	146
5	243
7	341
9	438

Following the 24 hour resting period at 4ºC, the state point diagram between the protein and DS was constructed via visual observations of the equilibrated mixture, based on work done by Lan et al.[74] To distinguish different phase behaviors of the observed phase separations in the test tubes at certain pH and MR pairs, five different symbols were used; specifically, ○, □,▲, ■, ● were used to represent translucent solution (○), turbid solution (□), precipitation & cloudy solution (▲), precipitation & clear solution (■), and precipitation & clear solution (●) with higher volume of precipitation than ■, respectively[74]. **Figure 2-4** shows representative pictures of the protein:DS reaction mixture after 24 hours, corresponding to each of the five symbols. For nanoparticle formulations, pH and MR values that form precipitation & clear solution are of primary interest because they signify the regime in which water-insoluble polyelectrolyte coacervates are formed and isolated for nanoparticle encapsulation. Armed with the knowledge of protein activity and structure across a wide range of pHs, this methodology allows for the determination of pH and MR values that result in highly active enzyme:charged polymer complex coacervates that are water-insoluble.

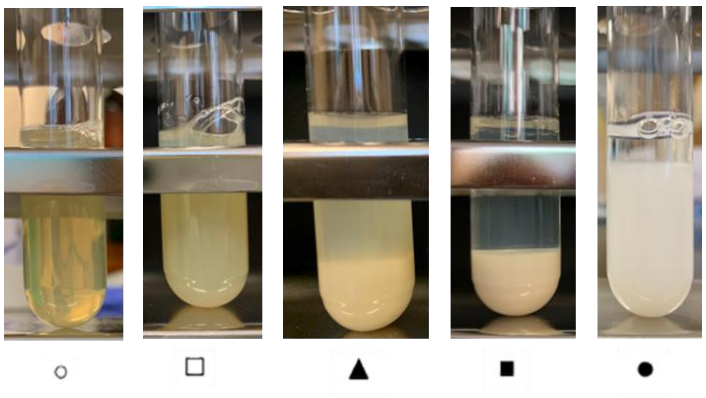

Figure 2-5: Visual inspection and categorization of 5 mL reaction mixture inside a test tube into 5 different symbols used for state point construction, following a 24 hour rest period at 4^{o}C. Not all protein:charged polymer pairs, across different pH and MR pairs, will exhibit all 5 phase states.

3. Chapter 3: Exploring structure and dynamics of the PLGA-PEG copolymer and its homopolymer constituents in various solvents using all-atom molecular dynamics[1]

3.1 Introduction

Poly(lactic-co-glycolic) acid (PLGA) is a biocompatible and biodegradable polymer that has gained popularity over the last few decades as a synthetic material for biomedical applications. Specifically, this polymer has been used to engineer wound dressings for burn victims[75], develop bone scaffolds ideal for bone regeneration[76] and improve therapeutic efficacy of certain drugs that have undesirable macroscopic properties[6,77]. PLGA spherical microparticles and nanoparticles have been extensively used for encapsulating a wide range of small molecule and macromolecular drugs, resulting in sustained release over time[15,78–80]. Furthermore, these drug-loaded nanoparticles can result in improved drug solubility and longer circulation half-life when administered *in vivo*[10,81]. When poly(ethylene-glycol) (PEG) is copolymerized to PLGA, the resulting diblock copolymer can be used to form core-shell nanoparticles capable of high drug loadings while reducing drug side effects and degradation by blood proteases[78]; hence, nanoparticles made with the PLGA–PEG diblock copolymer are highly desired for drug delivery applications.

Current formulation methodologies, such as solvent evaporation and double emulsion, can consistently produce polymeric nanoparticles with similar drug loading and release, but precise control of drug encapsulation during nanoparticle formulation is difficult. Formation of PLGA–PEG nanoparticles relies on dominant hydrophobic interactions of the PLGA domain in poor solvent conditions, resulting in a spherical PLGA core-PEG shell structure. Depending on the target of interest and hydrophilic nature of the drug, determination of the optimal drug loading and distribution within the nanoparticle is non-trivial and can be an intensive, iterative process. Moving towards a rational design-based approach could reduce the time to develop new nanoparticles with tailorable release rates. This, however, requires an understanding of polymer/solvent, polymer/polymer, and polymer/drug interactions that govern the formulation process. The interplay of polymer/solvent and polymer/polymer interactions dictates a PLGA–

[1] This chapter was reproduced with permission from Nyambura, C. W.; Sampath, J.; Nance, E.; Pfaendtner, J.* Exploring Structure and Dynamics of the PLGA-PEG Copolymer and its Homopolymer Constituents in Various Solvents Using All-Atom Molecular Dynamics. *J. Appl. Polym. Sci. 2022, e52732.*

PEG chain's structure and dynamics. Understanding these effects is critical to controlling drug encapsulation within such core-shell nanoparticles, since certain polymer chain structures and their diffusive motion in solution may promote highly favorable polymer/drug interactions. Both computational and experiment approaches have been used to examine polymer characteristics of PLGA, PEG, and their various multiblock polymers, building an important foundation for this study.

Experimentally, small-angle neutron scattering (SANS) and atomic force microscopy (AFM) have yielded insights on polymer structure and dynamics for dilute and semi-dilute solutions; they have enabled the extraction of parameters, such as mean squared end-to-end distance (R_{ee}^2), radius of gyration (R_g), Kuhn (L_k), and persistence (L_p) length. Le Coeur et al. studied PEG-conjugated hemoglobin with SANS and found that PEG chains <10 kDa have R_g similar to free PEG in water, whereas chains >10 kDa had a decreased R_g due to their compaction around the protein[82]. SANS has also been used to study PEG-PLGA–PEG triblock polymers[83] in aqueous solutions and other PEG-biologic complexes[84]. With AFM in dilute polymer solutions, Kienberger et al. covalently tethered PEG oligomers to their AFM tip and performed single molecule stretching experiments by measuring force-extension profiles[85]. They were able to extract PEG's persistence length and examined its mechanical response in water.

Computational techniques, like MD simulations, have also proven useful to understanding polymer structure and dynamics that take place during the nanoparticle formulation process. MD can provide estimates of useful polymer properties, such as surface hydrophobicity, H-bond donation and acceptability, and equilibrium chain conformations. For example, Oelmeier et al. performed a series of single PEG-water MD simulations at varying PEG chain length[52]; they showed a decrease in the number of hydrogen bonds and an increase in the average helicity with increasing PEG chain length. More recently, Andrews et al. simulated PLGA oligomers of different lengths and lactic acid stereoisomeric ratios, in water, ethyl acetate (EA) and a water-EA 2-phase system[53]. Their results showed PLGA highly favored an extended conformation in EA and a collapsed structure in water. Overall, experimental and computational studies have examined PLGA or PEG in dilute solutions, in addition to PEG-conjugated biologics. However, there are no studies, to our knowledge, that explore scaling relations in the structural and dynamic properties of the PLGA–PEG diblock copolymer in dilute solution.

Using atomistic MD with explicit solvent, the primary focus of this study was to understand the effect of PLGA–PEG copolymerization on polymer structure and dynamics compared to PLGA and PEG alone. We simulated three polymer types at various monomer lengths in three different solvents, in the dilute limit. Acetone, DMSO and water were examined because of their ubiquity in solvent evaporation and emulsion techniques used to make polymeric delivery vehicles. Using the results from the scaling analysis of key relaxation modes and the radius of gyration, we studied polymer self-similarity and hydrodynamic effects that are extensible to all chain lengths. Polymer stiffness and rigidity were explored via calculation of Kuhn and persistence length to further tease out the influence of copolymerization. Investigation of oligomer-solvent interactions and oligomer shape in solution also yielded insights on how the tested solvents have an impact on PLGA–PEG structure. Overall, this work seeks to bolster the "bottom-up" development of drug-loaded PLGA–PEG nanoparticles and other biomedical applications using the copolymer and solvents examined here.

3.2 Results and Discussion

3.2.1 Scaling analysis reveals nature of polymer self-similarity, along with presence of hydrodynamic and excluded volume effects

Oligomers of length 6, 8, 10, 16, 20 and 30 monomers were simulated for all polymer/solvent combinations. For all simulations, temperature and pressure were maintained around their reference values (1 bar, 298.15 K). Full time series of a subset of the simulations are provided in **Figures B-4 – B-6** and show that the systems quickly equilibrate their structural properties on timescales shorter than a few nanoseconds (ns). Dynamic properties (chain end-to-end autocorrelation) take between 10-20 ns to equilibrate.

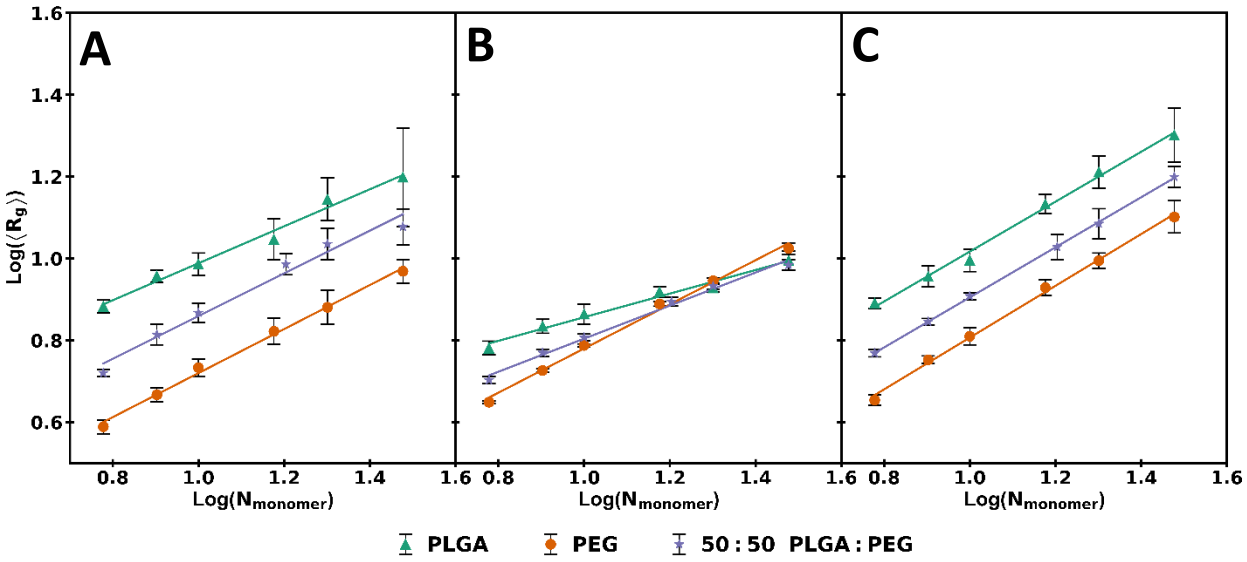

Figure 3-1: Scaling Relationship between the radius of gyration (R_g) and monomer length for PLGA, PEG and PLGA-PEG (50:50 LGA:EG ratio) in **(A)** acetone **(B)** water and **(C)** DMSO. Flory exponent (the slope), its standard deviation (SD) and the Pearson coefficient (r^2) are reported within the text below and in Table 3-1. SD in R_g comes from block averaging of trajectories.

Scaling analysis between R_g and N was then performed to extract the Flory exponent for each polymer/solvent system **(Figure 3-1)**. For the acetone systems, PLGA-PEG ($v_{PLGA-PEG} = 0.52 \pm 0.04,\ r^2 = 0.975$), PEG ($v_{PEG} = 0.54 \pm 0.02,\ r^2 = 0.996$) and PLGA ($v_{PLGA} = 0.45 \pm 0.03, r^2 = 0.985$) are near the theta point ($v = 0.5$), with no significant differences in their values. In pure DMSO systems, scaling values suggest good solvent conditions ($v > 0.5$) for both the homopolymers ($v_{PLGA} = 0.61 \pm 0.02,\ r^2 = 0.994$; $v_{PEG} = 0.63 \pm 0.02,\ r^2 = 0.997$) and the copolymer ($v_{PLGA-PEG} = 0.61 \pm 0.01,\ r^2 = 0.999$), indicating that polymer-solvent interactions are more dominant than polymer-polymer interactions for all 3 polymer types. Unsurprisingly, for good quality solvents, the PLGA-PEG copolymer behaves identically to the individual component behavior in the same solvent. Moreover, the PLGA-PEG Flory exponent in both acetone and DMSO can explain why these two solvents are ubiquitous in nanoparticle formulations. In both solvents, polymer-polymer interactions are less favorable, which could promote enhancement of monomer-drug interactions, resulting in better encapsulation of small molecules like doxorubicin[86] and trimethoprim[87] or hydrophobic-ion protein complexes[88,89].

PLGA ($v_{PLGA} = 0.29 \pm 0.02, r^2 = 0.980$), PEG ($v_{PEG} = 0.54 \pm 0.02,\ r^2 = 0.994$) and the copolymer ($v_{PLGA-PEG} = 0.40 \pm 0.034,\ r^2 = 0.992$) simulated in water show the most interesting

scaling behavior, providing insight into what happens when both blocks are different in their net attraction to the solvent. PEG had the largest Flory exponent and is very close to the previously reported data from Oelmeier et al.[52] ($v = 0.535$). PLGA had the smallest exponent, aligning with results from Andrews et al.[53] ($v = 0.4$). With 50:50 LGA:EG copolymer being the intermediate, a crossover in R_g is observed at N = 20 monomer length. PEG in water is near the theta point and PLGA is in poor solvent conditions ($v < 0.5$), resulting with the PLGA-PEG also being poor solvent conditions. This suggests that the PEG block within the copolymer compensates for the collapsed PLGA block, due to more favorable monomer-solvent interactions. At small monomer lengths, the copolymer structure is close to PEG because the number of LGA monomers is small; however, as the poor solvent conditions dominate with more LGA monomers, the PLGA block starts to affect PEG's random coil structure because of its attractive self-interactions.

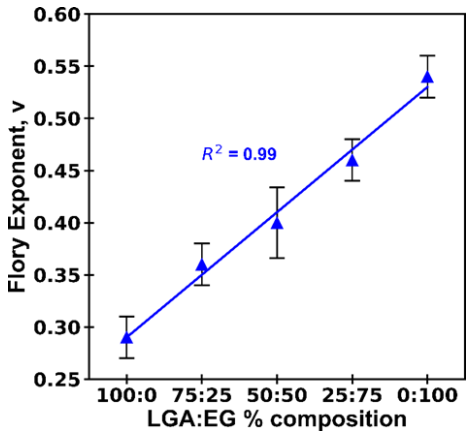

Figure 3-2: Linear relationship between the Flory exponent and LGA:EG composition for the copolymer/water systems. The Pearson coefficient (r^2) is reported within the plot above. Error in v comes from SD of the slope from R_g vs. monomer length scaling.

To further understand this behavior, additional compositions of the PLGA-PEG system were simulated in pure water including 75:25 LGA:EG ($v_{75\ LGA:25\ EG} = 0.36\ \pm 0.02, r^2 = 0.991$) and 25:75LGA:EG ($v_{25\ LGA:75\ EG} = 0.46\ \pm 0.02,\ r^2 = 0.994$) copolymer in pure water (**Figure B-7***)*. The relationship between the Flory exponent and LGA:EG composition, for the copolymer/water systems, were then examined (**Figure 3-2**). Linear regression was performed on this data, resulting in a high r^2 value, suggesting the deviation from theta point as LGA composition increases occurs linearly. Moreover, copolymer LGA composition greater than 25% yields Flory exponent values below 0.5, signifying poor solvent conditions. Thus, good solvent conditions for the PLGA block are likely more critical when compared to the PEG block. PEG in acetone and in water have the same average value for the Flory exponent, confirming the homopolymer's amphiphilic ability.

Figure B-8 & B-9 shows the positional bead autocorrelation data and relaxation times generated from the copolymer/solvent systems using equation 8 and fit to equation 9, in order to extract each oligomer's time relaxation spectrum. Visual inspection of such data is important for understanding how solvent conditions can influence decay and long-time fluctuation of this autocorrelation function. PLGA and PLGA-PEG oligomer autocorrelation decay, in DMSO, took nearly 10 ns or greater for all monomer lengths; however, in acetone and water, decay took less than or equal to 10 ns for all monomer lengths. The slower decay for polymers in good solvent conditions is indicative of expanded oligomer conformations, whereas rapid decay corresponds to collapsed conformations for polymers in poor solvent; this behavior was also observed for a 10-bead chain in good and poor solvent quality from Kong et al.[60] For PEG oligomers that were 10 monomers or longer in DMSO, their autocorrelation decay occurred in 10 ns or longer. Rapid decorrelation was also observed at all monomer lengths when PEG is in water or acetone. At large monomer lengths (N > 20), autocorrelation data fluctuations near zero for PLGA and PLGA-PEG systems in DMSO were relatively large (~100 ns), despite using a running average to reduce simulation noise.

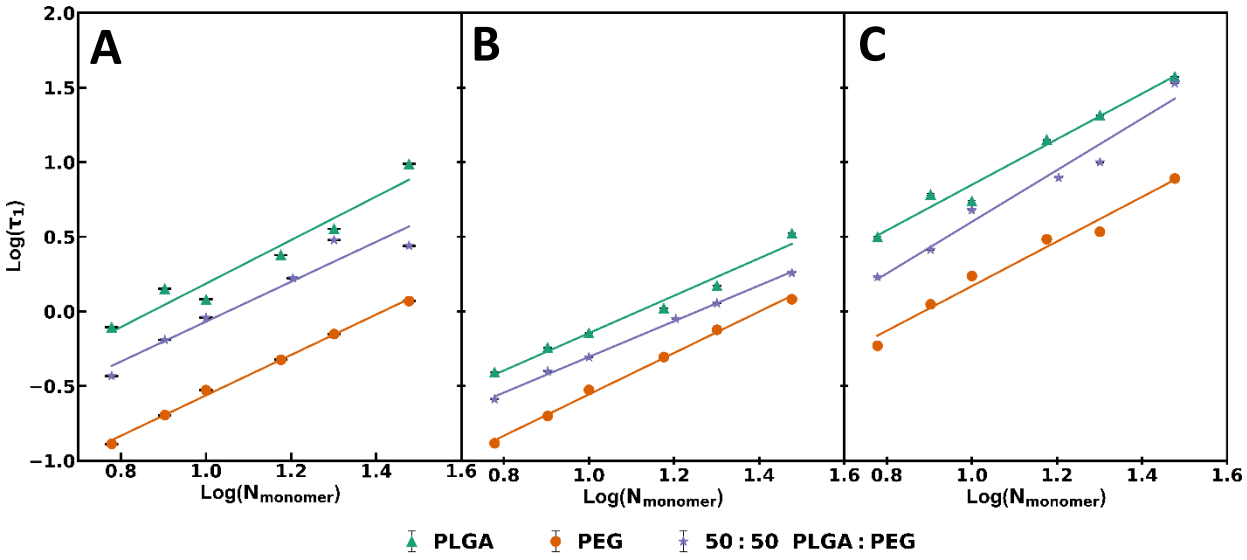

Figure 3-3: Scaling Relationship between the 1st relaxation time (τ_1) and monomer length ($N_{monomer}$) for PLGA, PEG and PLGA-PEG (50:50 LGA:EG ratio) in **A)** acetone **B)** water and **C)** DMSO. Alpha exponent is the slope of the line seen above. Standard error (SE) in τ_1 comes from the nonlinear fitting of positional bead autocorrelation data to eqn. 14.

Scaling analysis of the first principal relaxation time with monomer length was done to evaluate hydrodynamic and excluded volume effects. **Figure 3-3** show the dynamical scaling behavior, where the slope is represented by alpha (α). **Table 3-1** summarizes the Flory (v) and alpha (α) exponents. The Rouse model, which does not account for hydrodynamic and excluded volume effects, predicts $\alpha = 2$ and $v = 0.5$ for polymers in a theta solution[90] and best describes the dynamics of unentangled polymer melts. Zimm theory incorporates hydrodynamic interactions but not excluded volume effects[59,91], predicting $\alpha = 1.5$ and $v = 0.5$ and models dynamics of polymers in dilute solution well. De Gennes was successful in developing a model that accounts for both effects and yields values of alpha similar to Zimm theory, which showed polymers in good solvent condition have $\alpha = 1.77$ and $v = 0.59$, consistent with both experiment and theory[91].

Table 3-1: Flory and alpha exponents are summarized for PLGA, PEG and PLGA-PEG (50:50 LGA:EG ratio) in the 3 different pure solvents simulated in this study. SDs from linear regression are reported here. Pearson coefficients for the alpha exponents are reported in the text.

	Acetone		**Water**		**DMSO**	
	v	α	v	α	v	α
PLGA	0.45 ± 0.03	1.46 ± 0.18	0.29 ± 0.02	1.25 ± 0.10	0.61 ± 0.02	1.53 ± 0.12
PLGA-PEG	0.52 ± 0.04	1.34 ± 0.18	0.40 ± 0.04	1.20 ± 0.03	0.61 ± 0.01	1.73 ± 0.16
PEG	0.54 ± 0.02	1.35 ± 0.04	0.54 ± 0.02	1.39 ± 0.04	0.63 ± 0.02	1.49 ± 0.12

Using these models to interpret our data, analysis showed all polymer/solvent systems did not follow the Rouse model, as expected for polymers in dilute solution (**Table 3-1**). PLGA ($\alpha_{PLGA} = 1.46 \pm 0.18, r^2 = 0.942$) in acetone had v and α values that were closest to the ones predicted by the Zimm model, showing that hydrodynamic interactions were dominant and excluded volume effects were minimal. As with PEG ($\alpha_{PEG} = 1.35 \pm 0.04, r^2 = 0.997$) and 50:50 copolymer ($\alpha_{PLGA-PEG} = 1.34 \pm 0.18, r^2 = 0.934$) in acetone, values of α are below the Zimm prediction, despite $v \sim 0.5$. The weaker monomer length dependence on the first relaxation time for PEG/acetone, when compared to PLGA, can be attributed to a lower frictional resistance of the oligomers' diffusive motion through the solvent medium. PEG does not have methyl or carboxyl groups that can drag more of the surrounding solvent; therefore, the relaxation mode of the entire chain can relax more quickly than what is predicted by Zimm theory. This faster relaxation could also explain why the α value for the 50:50 copolymer/acetone systems match that of PEG, showing the impact of copolymerization on dynamics. As for pure DMSO, PLGA-

PEG ($\alpha_{PLGA-PEG} = 1.73 \pm 0.16, r^2 = 0.967$) has an α value that is within error of that predicted by the De Gennes scaling model. However, dynamical scaling for PEG ($\alpha_{PEG} = 1.49 \pm 0.12, r^2 = 0.975$) and PLGA ($\alpha_{PLGA} = 1.53 \pm 0.12, r^2 = 0.975$) resulted in alpha values closer to the Zimm model, despite $v \sim 0.6$. This suggests that hydrodynamic interactions and excluded volume effects both have a strong impact on the copolymer dynamics, whereas hydrodynamic interactions were more prevalent for the homopolymer/DMSO systems. This self-avoiding motion of the copolymer oligomers may be caused by the differences in each block's hydrophilicity. PEG is known to be the hydrophilic block while PLGA is known to be hydrophobic block. When copolymerized, polymer chain motion in dilute solution tries to minimize LGA-EG interactions, when dissolved in good solvent conditions.

Lastly, PEG ($\alpha_{PEG} = 1.39 \pm 0.04, r^2 = 0.997$) in water is close in value to PEG in acetone, showing that the amphiphilic nature of PEG is also observed in dynamical scaling analysis. No crossover was observed for polymer-water systems at the chain lengths considered (**Figure 3-3B**). PLGA ($\alpha_{PLGA} = 1.25 \pm 0.10, r^2 = 0.977$) and PLGA-PEG ($\alpha_{PLGA-PEG} = 1.20 \pm 0.03, r^2 = 0.998$) in water have the smallest α values when compared to all other polymer-solvent systems, resulting in τ_1 having the weakest dependence on N. This is likely due to strong LGA self-interactions, as a result of being in poor solvent conditions. The downstream effect is observed as a fast relaxation of the entire chain and aggregated LGA regions dictate most of the oligomers' motion. In poor solvent, the PLGA and PLGA-PEG have a negligible effect on viscoelastic solution properties, since the LGA monomers minimize their interactions with the solvent. This is likely why water is used to promote PLGA-PEG nanoparticle formation and stabilization for emulsion or solvent-evaporation methodologies.

3.2.2 Solvent structure around oligomer backbone provides insight into polymer-solvent interactions

The radial distribution function (RDF) was used to understand the strength of polymer-solvent interactions and evaluate the solvent structure around a given oligomer. **Figure 3-4** shows RDF profiles for the largest monomer length for each polymer type. For DMSO systems, the first peak for PEG is close to 1, whereas PLGA and PLGA-PEG have similar first and second peaks. For acetone, the first peak for PLGA and PEG are most similar, with PLGA-PEG being slightly below. The height of PEG's and the copolymer's first peak in DMSO, when

compared to acetone, shows slightly higher DMSO preference to those polymer chains. This shows that PEG-solvent interactions were stronger in DMSO, when compared to acetone and water. Moreover, PLGA-PEG in DMSO had the highest first peak, when compared to the other copolymer/solvent systems, likely due to good solvent conditions for PEG and PLGA block. Water solvent density for PLGA at radial distances below 1 nm is the smallest, followed by PLGA-PEG and lastly, PEG. Furthermore, the same trend is observed in polymer structure analysis, in water. Recent work by Dahal et al. showed via atomistic MD that PEG in water forms 1.2 hydrogen bonds per monomer unit, explaining why this system has the largest 1st peak when compared to PLGA and the copolymer[92]. Poor solvent conditions for PLGA and PLGA-PEG are reflected in lower RDF values at small radial distances, likely due to the collapsed PLGA structure that excludes solvent.

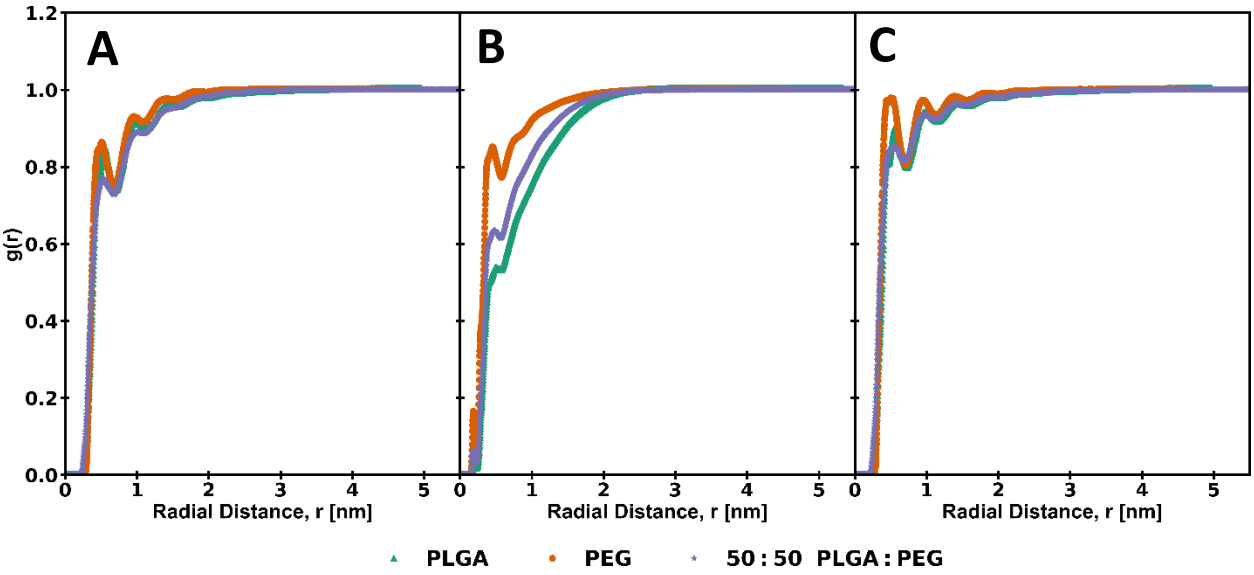

Figure 3-4: Radial Distribution Function profiles, where the references (or group A) atoms are the polymer backbone atoms for $N_{monomer} = 30$ only and the selection (or group B) atoms are **(A)** acetone heavy atoms, **(B)** water and **(C)** DMSO heavy atoms.

3.2.3 Copolymerization effects can be observed by analysis of polymer shape, stiffness and rigidity in different solvents

Scaling analysis developed from Flory theory has allowed us to examine the role of self-similarity, which has been useful in explaining polymer swelling in experiments and in understanding net interactions for each polymer/solvent system simulated here. The interactions that take place could result in polymer structures that can be classified in three regimes:

swollen, collapsed or ideal, as indicated by the scaling of R_g with chain length. The ratio of mean squared end-to-end distance over mean squared radius of gyration ($\langle R_{ee}^2 \rangle / \langle R_g^2 \rangle$) also provides a measure of whether or not an oligomer is behaving as a non-self-interacting chain with end-to-end distances that have a gaussian distribution[93]. When $\langle R_{ee}^2 \rangle / \langle R_g^2 \rangle = 6$, a polymer chain behaves as an ideal linear chain with gaussian distributed end-to-end distance. If $\langle R_{ee}^2 \rangle / \langle R_g^2 \rangle < 6$, then a polymer chain displays real chain behavior and is in a poor solvent condition, since $\langle R_{ee} \rangle$ is small in magnitude when compared to the ideal chain prediction. If $\langle R_{ee}^2 \rangle / \langle R_g^2 \rangle > 6$, then a polymer chain is extended, in good solvent conditions but also behaves non-ideal, with large R_{ee} distances likely having higher probabilities.

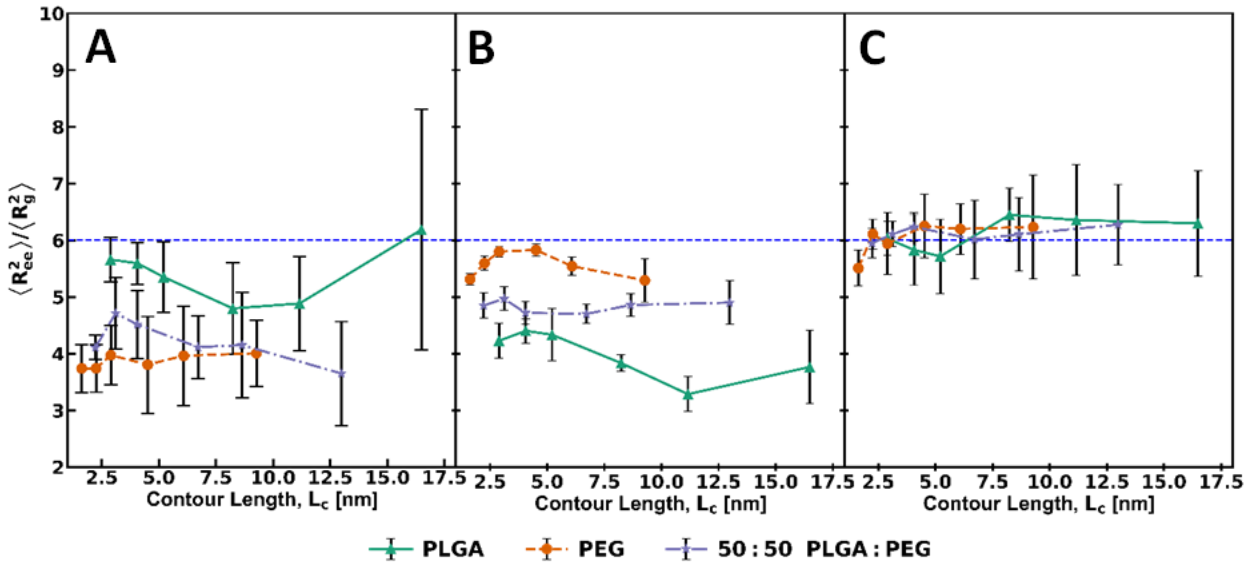

Figure 3-5: Mean squared end-to-end distance over mean squared radius of gyration ($\langle R_{ee}^2 \rangle / \langle R_g^2 \rangle$) for PLGA, PEG and PLGA-PEG (50:50 LGA:EG ratio) in **(A)** acetone **(B)** Water and **(C)** DMSO plotted against contour length. Horizontal dashed line in **A-C** is the predicted theoretical value of an ideal linear chain in dilute solution, as $N_{monomer}$ goes to infinity.

For polymer systems in DMSO, $\langle R_{ee}^2 \rangle / \langle R_g^2 \rangle$ was fluctuating near 6, for both homopolymers and the copolymer (**Figure 3-5**), showing how important good solvent conditions are for both blocks in PLGA-PEG in adopting ideal-chain conformations. PEG in water oligomers have $\langle R_{ee}^2 \rangle / \langle R_g^2 \rangle$ ratios closest to 6 whereas PLGA oligomers have $\langle R_{ee}^2 \rangle / \langle R_g^2 \rangle$ ratios furthest from 6, which is unsurprising given that water is a poor solvent for PLGA. Therefore, PLGA-PEG deviation from ideal-chain behavior can likely be explained by the collapsed PLGA block and semi-extended PEG block. For the acetone systems, $\langle R_{ee}^2 \rangle / \langle R_g^2 \rangle$ ratio for PEG and

PLGA-PEG were below 6 with errors considered. However, PLGA oligomers had ratios with average values closest to 6. Despite having the same Flory exponent as PEG in water and being at the theta point, PEG in acetone oligomers have non-gaussian distributed end-to-end distances, likely due to the presence of EG-EG monomer interactions. When PEG is copolymerized to PLGA, EG-EG interactions are still taking place in theta-point solvent conditions, resulting in departure of $\langle R_{ee}^2 \rangle / \langle R_g^2 \rangle$ away from the ideal-chain prediction.

The behavior of PLGA-PEG in acetone, when compared to DMSO, provides insight into the differences in PLGA-PEG nanoparticle structure, observed in Kaldybekov et al.[19] They found that PLGA-PEG nanoparticles precipitated in acetone had smaller cores and lower R_g values of PEG chains in the corona, when compared to DMSO. This observation could be due to more favorable LGA-LGA interactions and PEG's non-ideal chain behavior when the copolymer dissolved in acetone. In DMSO, strong polymer-solvent interactions result in corona PEG chains behaving ideally, leading to extended conformations. Moreover, these interactions might be why differences in the local structure between the PLGA core and PEG brush are observed, when compared to DMSO and acetone[19].

Table 3-2: Persistence (L_p) and Kuhn (L_k) lengths for PLGA, PEG and PLGA-PEG (50:50 LGA:EG ratio) in acetone, water and DMSO. Unit of Length is Angstrom. L_p values are reported alongside their standard error. L_p extraction resulted from non-linear fitting of eqn. 11 to simulation, whereas L_k came from linear regression using eqn. 12.

	Acetone		**Water**		**DMSO**	
	L_p [Å]	L_k [Å]	L_p [Å]	L_k [Å]	L_p [Å]	L_k [Å]
PLGA	4.90 ± 0.27	9.30 ± 1.04	1.37 ± 0.21	1.30 ± 0.29	8.05 ± 0.19	16.5 ± 0.69
PLGA-PEG	2.50 ± 0.22	3.90 ± 0.66	2.05 ± 0.15	3.13 ± 0.20	6.12 ± 0.16	12.6 ± 0.56
PEG	1.95 ± 0.02	3.80 ± 0.08	3.55 ± 0.17	6.25 ± 0.42	5.63 ± 0.07	11.4 ± 0.21

In order to get a better idea of how rigid and expanded these polymers are, **Table 3-2** shows persistence and Kuhn length values; **Figures B-10 – B-12** show the $\langle R_{ee}^2 \rangle$ vs. L_c data and the fits performed for L_p and L_k extraction. For the water systems, PLGA has L_p and L_k are nearly the same, with PLGA-PEG's L_k being one angstrom larger its L_p value. Persistence length for PEG in water is close to its experimental value of 3.8 ± 0.02 Å[85] and MD value from Lee et. al. of 3.7 Å[94]; in addition, this system has the largest L_k and value L_p, when compared to PLGA and PLGA-PEG. From these data, PLGA in water is the least rigid and has the smallest

freely jointed effective bond length out of all the systems. PLGA-PEG is not as rigid or as expanded in solution, when compared to PEG. The 50:50 copolymer structure in water can be explained by the impact of the extended PEG block on rigidity and Kuhn bond length, as seen in the slightly larger L_p value. Poor solvent conditions for PLGA impedes presence of rigid (large L_p) and expanded (large L_k) polymer structures, due to preferred intra-chain interactions. Furthermore, the PLGA block also could explain the drop in PLGA-PEG's L_k value.

For the acetone systems, PEG and PLGA L_k average values are twice their L_p average value. PLGA has the largest $L_p$and L_k values, when compared to PEG and the 50:50 copolymer. L_k average value for PLGA-PEG is the same as PEG, but L_k/L_p is below 2. Therefore, PLGA-PEG and PEG average effective bond lengths show PEG block has a significant effect on the 50:50 copolymer structure, despite scaling analysis showing PEG in acetone is near the theta point. This could be due to fewer PEG-solvent interactions and higher self-interactions. In addition, the PLGA block is in a theta solvent and could be responsible for the slight increase in PLGA-PEG rigidity. When examining the polymer/DMSO systems, PLGA has the largest L_p and L_k values out of all simulated systems, followed by PLGA-PEG and lastly, PEG. L_k average values for all 3 polymer systems are twice their L_p average values, suggesting that they are all behaving like a worm-like chain[54]; it predicts $2L_p = L_k$. This result further supports the additive-like behavior of the PLGA-PEG copolymer; if the homopolymers are rigid and expanded in good solvent conditions, then the 50:50 copolymer will also be rigid and expanded. The high level of extension for PLGA-PEG in DMSO, in stark contrast to acetone, indicates that using a mixture of these two water-miscible solvents during nanoparticle formulation is a viable choice. Specifically, varying the DMSO mole fraction within a PLGA-PEG/acetone solution can change polymer rigidity and expansion, which in turn influences drug interactions and subsequent loading.

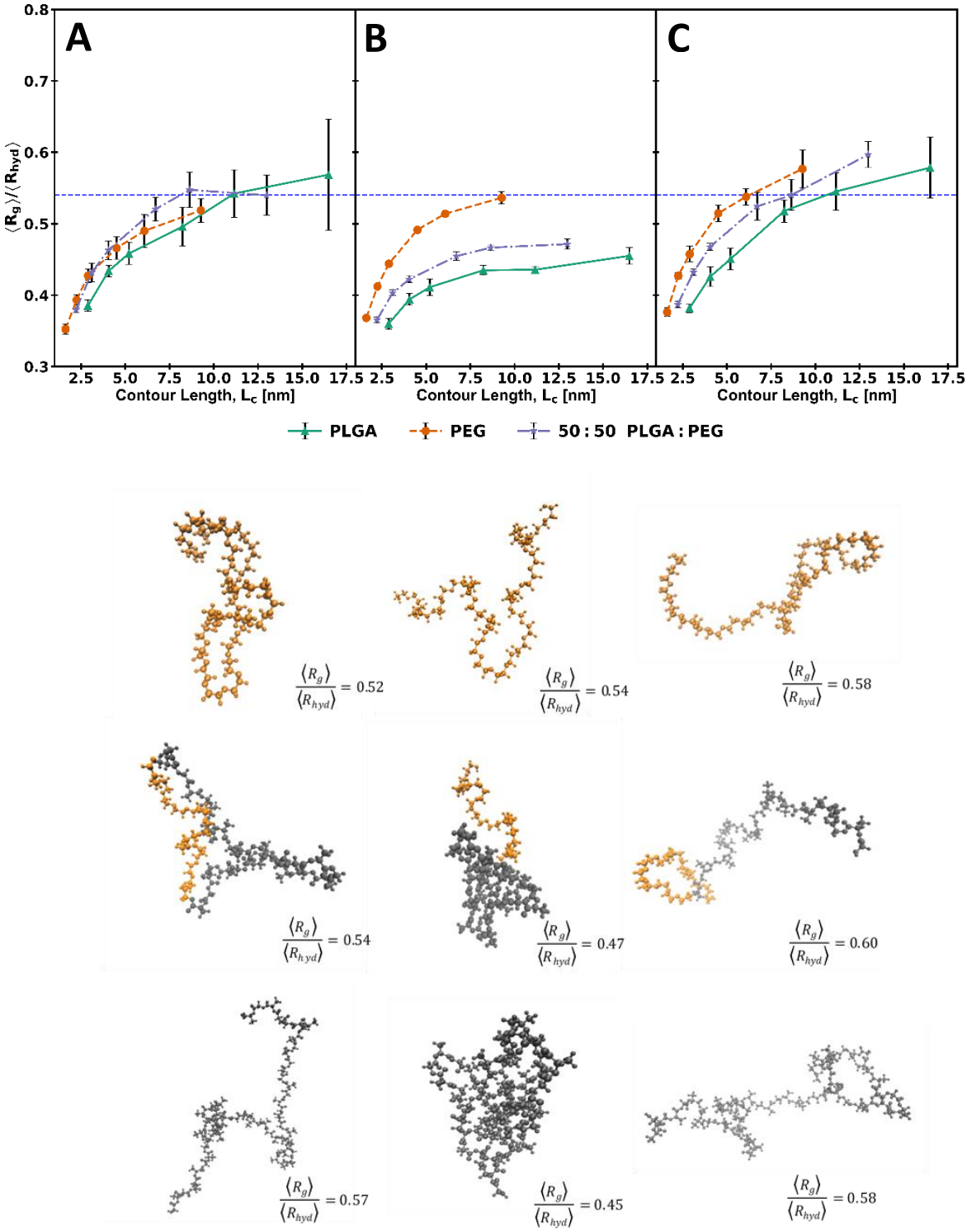

Figure 3-6: Radius of gyration over hydrodynamic radius ($\langle R_g \rangle / \langle R_{hyd} \rangle$) plotted against contour length for PLGA, PEG and PLGA-PEG (50:50 LGA:EG ratio) in the **(A)** acetone, **(B)** water and **(C)** DMSO. Representative oligomer snapshots of N = 30 systems at specific $\langle R_g \rangle / \langle R_{hyd} \rangle$ values are shown below the plots, where left column of images is the acetone systems, right column is DMSO, and middle column is water systems. Black is PLGA and orange is PEG.

Polymer shape in solution was examined by the radius of gyration over the hydrodynamic radius ratio ($\langle R_g \rangle / \langle R_{hyd} \rangle$). Previous studies have used this ratio to characterize globular proteins ($\frac{R_g}{R_{hyd}} \approx 0.77$) in solution[95] and understand PLGA oligomer sizes in different solvents. Moreover, Andrews et al.[53] used $\frac{R_{stokes}}{R_{hyd}} \approx 0.50 - 0.54$, based on oligomer sizes and a monomer's COM, to determine that extended conformations have $\frac{R_g}{R_{hyd}} > 0.54$ and collapsed structures have $\frac{R_g}{R_{hyd}} < 0.54$; the same approach is used in this study to understand copolymerization effects[16]. $\langle R_g \rangle / \langle R_{hyd} \rangle$ showed values below 0.77 for all 3 polymer types (**Figure 3-6**), despite the monomer length differences of simulated oligomers. For water systems, PEG oligomers have the largest $\langle R_g \rangle / \langle R_{hyd} \rangle$ ratios at increasing monomer lengths, whereas PLGA has smallest $\langle R_g \rangle / \langle R_{hyd} \rangle$ ratios. PLGA-PEG in water is the intermediate, for both ratios. At small contour lengths, PEG and the copolymer in acetone have similar $\langle R_g \rangle / \langle R_{hyd} \rangle$ values, but slightly deviate at large lengths. Asymptotic behavior is observed in all 3 polymer/solvent systems for the $\langle R_g \rangle / \langle R_{hyd} \rangle$ ratio, with distinct separation observed in the DMSO and water systems. $\langle R_g \rangle$ is less than $\langle R_{hyd} \rangle$ for all systems and is expected for dilute polymers at the simulated monomer lengths.

3.4 Conclusions

In this study, PLGA and PEG homopolymers, along with the PLGA-PEG copolymer, were simulated in 3 different pure solvents using atomistic MD. We showed that the Flory exponent was able to capture the impacts of each copolymer block in solution, which aided understanding of why the PLGA block in poor solvent conditions was responsible for collapsed PLGA-PEG oligomers in water; this is also observed in the RDF profile, showing lower water density near the PLGA and PLGA-PEG backbone at short distances when compared to acetone and DMSO. All three polymer types were found to be in good solvent conditions when dissolved in DMSO and in theta conditions when dissolved in acetone. Most solvent-evaporation formulations in literature first dissolve polymer and/or drug in acetone or DMSO, likely due to good/ideal solvent conditions promoting expanded, well-solvated oligomers. Subsequently, that solution is added to water/surfactant solution to promote formation of PLGA-PEG nanoparticles because dominant LGA-LGA interactions in poor solvent conditions results in polymer

partitioning into the well-known core-shell nanoparticle. Dynamical scaling of the 1st relaxation time with monomer length confirmed the presence of hydrodynamic and excluded volume effects for all simulated systems, indicating that PLGA-PEG in DMSO can be described by the De Gennes scaling model. Future rheological studies on polymer solutions, frequently used in nanoparticle formulation and in fields that employ PLGA-PEG or its homopolymer constituents, can be done to experimentally validate scaling analysis performed in this study.

The two ratios, $\langle R_g \rangle / \langle R_{hyd} \rangle$ and $\langle R_{ee}^2 \rangle / \langle R_g^2 \rangle$, were calculated and provided insights on PEG's non-ideal linear chain behavior when dissolved in acetone, despite Flory theory prediction of $v = 0.5$. PLGA oligomers in acetone were not all at ideal-chain conditions, suggesting that intrachain interactions can be more dominant than polymer-solvent interactions. As for the copolymer in acetone, the PEG block and favorable LGA self-interactions are likely responsible for the deviation away from ideal-chain behavior. Good solvent conditions for the homopolymers in DMSO aligned with a linear expanded structure for PLGA-PEG, as seen with extraction of L_p and L_k. When compared to acetone, PLGA-PEG is not as rigid and expanded. This contrast in copolymer structure indicates that PLGA-PEG dissolved in a mixture of acetone/DMSO will have some intermediate level of rigidity and expansion. Furthermore, the acetone mole fraction can be used to tune PLGA-PEG extension in solution, possibly impacting drug loading. Future formulations using acetone/DMSO mixtures should investigate the effect of mole fraction on nanoparticle characteristics and drug loading. In addition, deviation of PLGA-PEG and PEG from non-ideal behavior, when dissolved in acetone, should also be examined in future studies to bolster development of novel biomedical applications.

3.5 Acknowledgements

This work was funded by National Science Foundation Grant #1703438. This work was facilitated through the use of advanced computational, storage, and networking infrastructure provided by the Mox Hyak supercomputer system at the University of Washington

4. Chapter 4: Examining the effect of polymer extension on protein/polymer interactions that occur during formulation of poly(lactic acid-co-glycolic acid)-polyethylene glycol nanoparticles.

4.1 Introduction

Examination of poly(lactic acid-co-glycolic acid)-polyethylene glycol (PLGA-PEG) and its homopolymer constituents in different solvents, as seen in Chapter 3, revealed the extent of polymer/solvent interactions on chain conformation and rigidity, in addition to its impacts on the core-shell structure of the PLGA-PEG nanoparticle. Insights from this work indicated that finding the right polymer/solvent combination, prior to nanoparticle formation, that results in high protein drug loading and sustained release while maintaining desirable nanoparticle characteristics is not straightforward. Polymer chains in solution fluctuate between extended or collapsed conformations and their average size, measured using ensemble-averaged R_{ee} or R_g, can result in a wide range of polymer-drug binding modes that dictate the overall nanoparticle's release profile. Depending on the hydrophilic nature of the encapsulated drug, its localization within the nanoparticle matrix, the polymer-to-drug ratio and the interplay between drug-polymer, polymer-polymer and polymer-solvent molecular interactions may result in a sustained release profile but only after extensive iterations of varying formulation process parameters. Experimental and theoretical investigations of drug release from polymeric carriers are present in literature have resulted good analytical models to describe PLGA-PEG nanoparticles loaded with either hydrophilic or hydrophobic small molecules; they can show that hydrophobic therapeutics have a lowered degree of an initial burst phase, when compared to hydrophilic compounds[14]. Protein drugs, however, are typically 1-2 orders of magnitude larger than their small molecule counterparts, varying in their surface morphology and chemistry. With advances in computational capacity over recent years, knowledge of the most favorable binding modes between the polymer and protein can be garnered from computational tools, like atomistic MD, and allow for bottom-up design of protein-loaded PLGA-PEG nanoparticles.

Experimental and computational studies examining protein-polymer interactions in various drug delivery and biopharmaceutical applications have established a general baseline of how both techniques have been used to connect atomic-level observations to those seen on the mesoscale. Scanning transmission X-ray microscopy, employed by Leung et al. to investigate

human serum albumin (HSA) adsorption on phase separated polystyrene-b-poly(methyl methacrylate) (PS-b-PMMA) thin films, showed increased adsorption of HSA near inter-domain interfaces between PS and PMMA, in addition to other tested proteins showing preferred adsorption on hydrophobic PS domians[96]. Computationally, Yang et al. simulated an insulin/water/PEG system while varying polymer chain length by using a simulated annealing procedure to achieve faster convergence on the system's equilibrium state for high PEG molecular weights[97]. After an equilibrium state had been reached, they ran the systems for 10 ns and found that the PEGylated insulin structure was similar to that of free insulin. Moreover, they observed favorable hydrophobic interactions occur with the insulin surface at certain PEG chain lengths, likely resulting in a decrease in activity at high PEG molecular weight. Ultimately, combining atomistic MD and experimental studies, like Leung et al., will be critical for developing a deeper mechanistic understanding of dominant protein-polymer interactions during formulation of PNPs. Chapter 3 characterized polymer structure properties of PLGA-PEG and its homopolymer constituents; however, no study has investigated whether or not varying polymer's radius of gyration results in more favorable protein drug binding and influences the driving forces necessary for high protein drug loading

In this work, protein/polymer interactions were evaluated by simulating 3 polymer oligomers in the presence of a therapeutically relevant protein, Iduronate-2-sulphatase (ID2S), and pure water. This MD model was based on the experimental polymer and protein concentrations employed the nanoprecipitation formulation employing PLGA-PEG. ID2S is the protein of interest because it is the main cause of Hunter's syndrome, an inherited genetic disorder that results in accumulation of glycosaminoglycans within lysosomes, due to a ID2S deficiency[65]. Therefore, ID2S-loaded PLGA-PEG nanoparticles can potentially be used as an enzyme replacement therapy, thus providing patients a viable treatment option for this rare disease. Three polymers (PLGA-PEG, PLGA and PEG) and three levels of oligomer extension (high, medium and low) were tested to observe if there are differences in copolymer/protein interactions, when compared to its homopolymer constituents, in addition to understanding whether polymer conformation has a large or negligible impact on the protein/polymer interactions. We aim to use insights from this investigation to explain how polymer conformation, which is controlled by the choice of the organic solvent, affects protein/polymer interactions that occur at either the first step of the nanoparticle formulation process (polymer organic solution is mixed with protein aqueous solution), the very early stages of nanoparticle formation (polymer/protein mixture is added dropwise to large water sink) and during protein release out of PLGA-PEG nanoparticles while dissolved in an aqueous medium.

Table 4-1: Monomer length and contour length of the 3 oligomers, for each polymer type, simulated in this study. Each PLGA-PEG oligomer consisted of $N_{LGA} = 12$ and $N_{EG} = 13$. This was done to ensure that the simulated oligomers are still diblock copolymers with a nearly 1:1 LGA:EG ratio, at the given contour length.

Polymer	N_{mon}	L_c [Å]
PLGA	20	104.7
PLGA-PEG	25*	103.7
PEG	33	103.7

Table 4-1 below shows that, regardless of the polymer type, all simulated oligomers had nearly the same contour length. By holding L_c constant, the same R_g values corresponding to the three levels of extension can be used across the three tested polymers, since the range of possible oligomer sizes in solution, across polymer types, will approximately be the same and allow for a more systematic investigation. To ensure adequate sampling of different binding modes across polymer types and levels of extension, 7 different initial configurations were used, with the oligomers randomly oriented around the protein across each trial; this resulted in 63 simulations total (**Figure 4-1A**). A harmonic potential was used to restrain all 3 oligomers' R_g at high (2 nm), medium (1.5 nm) and low (1.1 nm) levels of extension for the first 100 ns (restraint ON phase) to assess the effect of varying solvent quality on protein/polymer binding. These R_g values were based on classical simulations of PLGA oligomers in various solvents used ubiquitously in nanoparticle formulations (**Figure C-1**). After 100 ns, the harmonic restraint is turned off (restraint OFF phase) and the system is allowed to propagate for another 100 ns to gain some insights on whether polymer binding is irreversible in the pure water medium, mimicking the step where protein/polymer mixture is added dropwise, and PLGA-PEG nanoparticle formation is beginning (**Figure 4-1B**). A total simulation time of 200 ns was chosen because diffusion of ID2S through a ~10 nm box would last ~200-300 ns, calculated using Stokes-Einstein diffusion equation to estimate the protein's motion in pure water. Furthermore, prior investigations into the solvent displacement method show that diffusion of the water-miscible solvent and other molecular species within the non-solvent medium is rapid, when compared to the rate of nanoparticle nucleation[98,99]. Therefore, we assume the pure water environment in the restraint OFF phase closely mimics the early stages of ID2S/PLGA-PEG cluster formation.

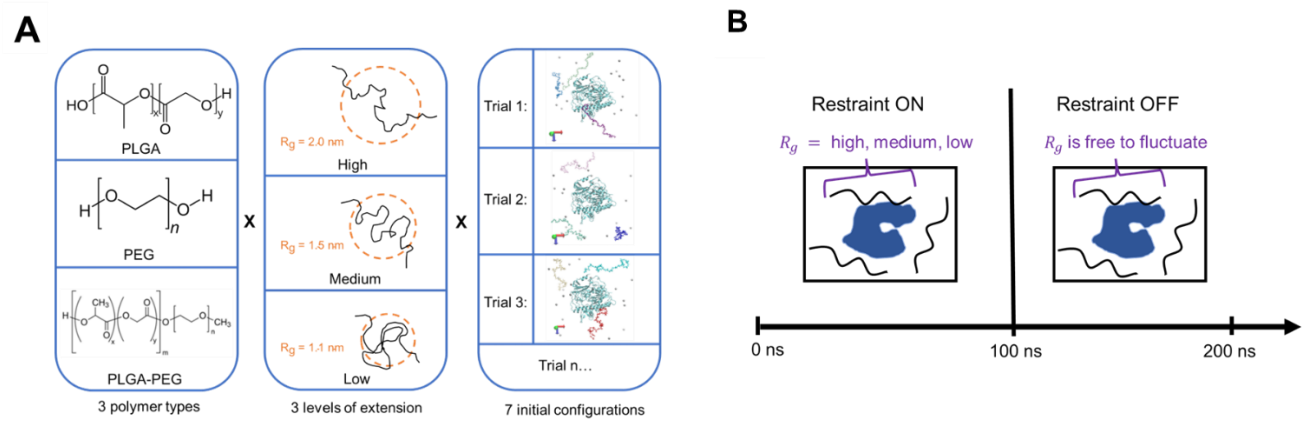

Figure 4-1: Layout of **A)** the experimental design implemented in this study and **B)** the progression of each simulation, where all oligomers in the simulation box are restrained to the same R_g value corresponding to a certain level of extension.

4.2 Results and Discussion

4.2.1 Copolymerization and extension effects on protein-polymer interactions

Protein stability across both restraint phases was evaluated by extracting ID2S backbone RMSD from each simulation trial as a function of time and comparing that time series to ID2S in water only (**Figure C-2**). Across the three polymer types tested, RMSD fluctuations during all simulation trials were lower or similar to ID2S alone, showing that interactions between the protein and oligomers were not resulting in protein unfolding and that the simulated polymers are primarily sampling the ID2S surface. In addition, temperature and pressure were maintained at the set values of 298.15 K and 1 bar in both phases for all trials, confirming that system configurations were well-equilibrated and final trajectories can be used for further analysis. We characterized and visualized three of the four classes of interactions described in the methods chapter: protein-polymer (prot-poly) interactions, polymer-polymer (poly-poly) interactions and both prot-poly and poly-poly interactions. The oligomer and time averaged time frequency, $\mathrm{TF}_{\mathrm{avg}}^{\mathrm{poly}}$, for those 3 interaction classes can be seen in **Figure 4-2** for PLGA-PEG and its homopolymer constituents.

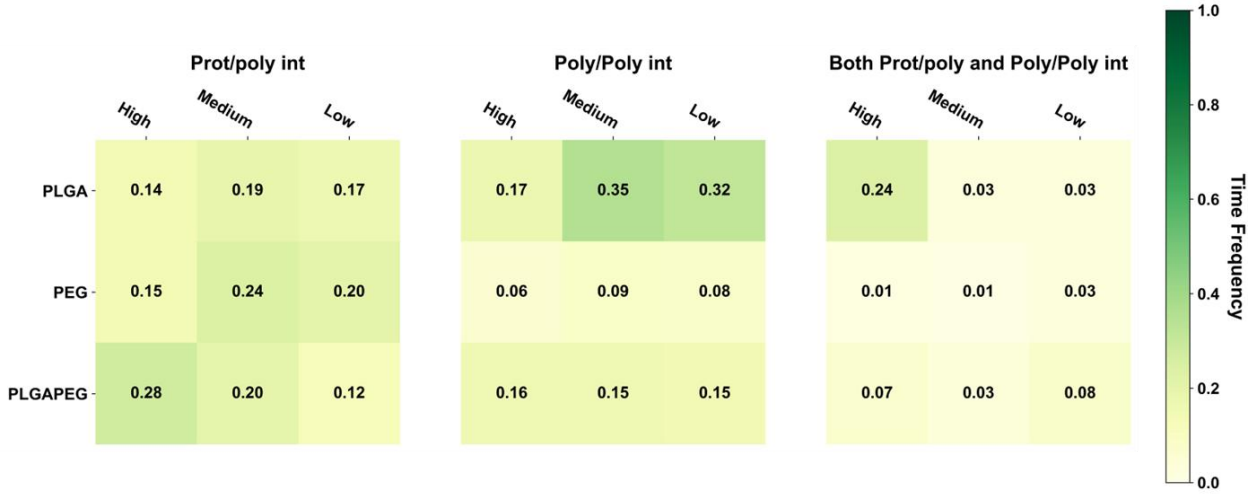

Figure 4-2: TF_{avg}^{poly} values for each polymer type and interaction class, across different levels of oligomer extension. Time frequency values from simulation trials (n = 7) and oligomers (n = 3/trial) were averaged to arrive at the observed values above.

For PLGA/ID2S systems, poly-poly interactions occurred more often than prot-poly interactions for oligomers at medium and low levels of extension whereas both prot-poly and poly-poly interactions were most frequent for the high extension systems. Moreover, prot-poly interactions only were still present for all levels of extension, showing that, in all system configurations, PLGA oligomers sampled the ID2S surface. Examination of TF across various configurations (**Figure C-3**) at high and medium levels of extension shows that three of the seven trials (trials 4-6) underwent significant poly-poly interactions, indicating that polymer oligomers primarily aggregated and were free in solution at medium extension. This is due to PLGA oligomers being in poor solvent conditions, resulting in more dominant LGA-LGA interactions. At low extension, trials 4 and 6 showed the highest frequency of prot-poly interactions only when compared to other trials, with little to no frequency observed for both prot-poly & poly-poly interactions across the 7 trials. This suggests that most times, collapsed PLGA oligomers would either individually contact the protein or be aggregated together in solution at low extension. At medium extension, trials 1-3 had a higher frequency of prot-poly interactions when compared to the other 2 classes, despite the driving force of PLGA oligomer aggregation; the effect of multiple trials is seen for PLGA/ID2S systems since four of the seven simulations were still able to undergo prot-poly binding without the oligomers rapidly aggregating.

As for the PEG/ID2S trials, poly-poly and both prot-poly & poly-poly interactions were minimal, where trials 3 and 5 were responsible for the TF values in these interaction classes.

Medium level of extension was found to have the largest averaged TF value for the prot-poly interaction class, resulting from 4 trials experiencing favorable binding. This observation may be explained by the PEG oligomers being in good solvent conditions. At high extension, polymer-solvent interactions can occur more easily and are likely more favorable than prot-poly interactions. The ease in which polymer-solvent interactions can occur, however, is reduced at medium and low levels of extension since opportunities for hydrogen bonding is limited; collapsed PEG oligomers interact still contact ID2S but TF values across trials were below 0.3 (6 of 7 trials), indicating that, for PEG/ID2S systems, the solvent had an appreciable impact on prot-poly interactions. Lastly, PLGA-PEG/ID2S systems showed that copolymerization reduces the extent of poly-poly interactions; furthermore, trial systems with oligomers at high extension resulted in more frequent prot-poly interactions (**Figure C-3**), when compared to medium and low extension systems. Trials 5-7 encountered had higher TF values for both prot-poly & poly-poly interactions while trials 1-4 encountered more poly-poly interactions. Overall, increasing PLGA extension promotes more prot-poly interactions. When PLGA is copolymerized with PEG, poly-poly interactions are reduced, allowing for stabilization of favorable binding modes. During nanoparticle formulation, PLGA-PEG chains will have polymer blocks differing in their interaction with the solvent. This meaning oligomers with collapsed LGA and extended EG domains may undergo less frequent protein contact whereas extended LGA and EG domains result in more frequent protein contact.

In order to understand how different levels of extension affects oligomer binding and extent of its interaction with ID2S, R_g time series and probability density data, were extracted for each oligomer and is shown in **Figure 4-3** for select simulation trials. Furthermore, the four interaction classes are overlaid over the R_g time series data to examine how transient protein-polymer binding is, when the restraint is on or off. For the high extension case, oligomers in PLGA/ID2S trial 2 were initially free in solution then 2 oligomers encountered ID2S surface separately when the restraint was on; once the restraint was off, poly-poly interactions started to occur between the oligomers, while they were in contact with ID2S. One oligomer underwent rapid collapse, due to being in poor contact with ID2S, while the other two took 40-60 ns to collapse to 1 nm in size. This slow collapse can also be seen in broader distribution of the probability density for those oligomers. This behavior was also observed in other trials whereby favorable PLGA oligomer binding with ID2S in the restrained phase resulted in slow decay of R_g when the restraint was turned off. In addition, poly-poly interactions in early simulation time impacted the frequency of prot-poly interactions since oligomer aggregation during the

restrained phase would persist in some trials and subsequently collapse together while contacting the protein or being free in solution.

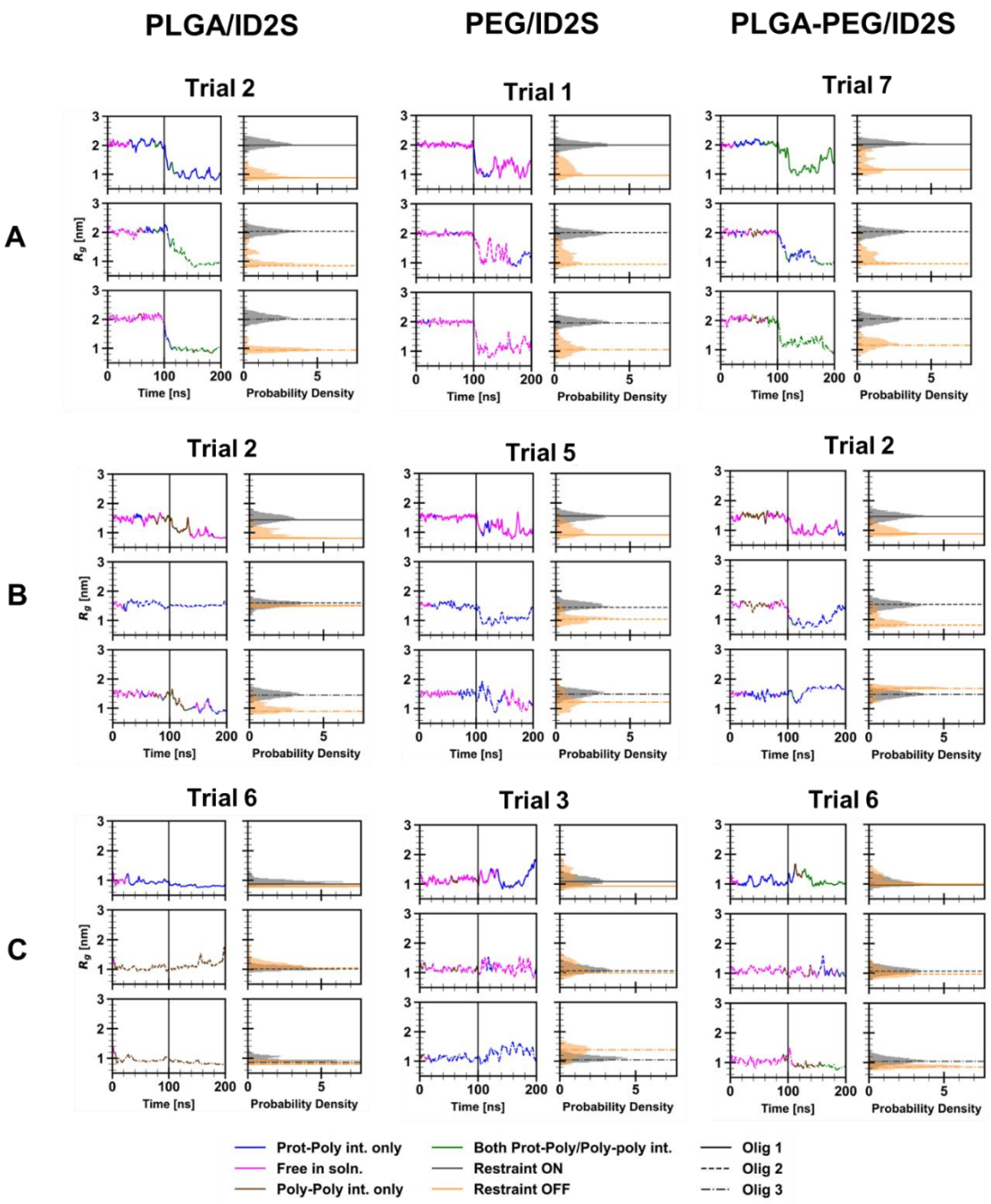

Figure 4-3: Radius of gyration (R_g), following a moving average calculation to reduce noise, is plotted against simulation time for each oligomer for the selected trials, in addition to the probability density of R_g in both restraint ON and OFF phase. **(A)** indicates high extension, **(B)** medium extension and **(C)** low extension for PLGA/ID2S **(left)**, PEG/ID2S **(middle)** and PLGA-PEG/ID2S **(right)** systems. The four interaction classes are overlaid on the R_g vs. time data to understand the overall behavior of protein/polymer binding.

PEG/ID2S trial 1 oligomers at high extension showed reversible protein binding in both restrained and unrestrained phases. Probability densities of the oligomers across all trials were broader, when compared to PLGA/ID2S trials after the restraint is turned off, indicating their return back to ideal-chain conformations due to being in good solvent condition. Across PEG/ID2S trials regardless to extension level, polymer-polymer interactions were infrequent, and oligomers were mostly free in solution in both phases; TF values extracted across trials also confirms this observation of PEG/ID2S binding. Trial 5 at medium extension and trial 3 at low extension, however, exhibited irreversible binding between ID2S and one oligomer, showing that favorable PEG-ID2S interactions do occur despite the strength of PEG-water interactions. Two oligomers in PLGA-PEG/ID2S trial 7 at high extension were mostly free and experienced transient prot-poly and poly-poly interactions in the restraint ON phase, whereas one oligomer underwent irreversible binding. In the unrestrained phase, poly-poly interactions increased while all 3 copolymer oligomers were contacting ID2S surface, resulting in a broader range of R_g values similar to the PEG/ID2S systems. 4 trials showed irreversible binding between ID2S and at least one oligomer in the restrained phase and persisted in the unrestrained phase, with poly-poly interactions occurring less frequently than PLGA/ID2S systems. Overall, reducing polymer-polymer interactions between PLGA domains during the early stages of nanoparticle formation is likely critical to promoting protein-polymer contact; however, strong polymer-solvent interactions are also not desired, as seen with the PEG/ID2S systems, because this leads to more reversible polymer binding events. Through PLGA copolymerization to PEG, poly-poly interactions are reduced and more polymer-solvent interactions are able to occur, resulting in an increase in protein-polymer interactions.

As for the low extension case, PLGA oligomers stayed collapsed across all trials after the restraint was turned off, whereby the frequency of both prot-poly and poly-poly interactions occurring together was much lower when compared to the other polymer/ID2S systems. Oligomer aggregation was persistent for some trials, but it did not prevent PLGA/ID2S systems from experiencing irreversible protein-polymer binding (trial 6 at low extension), as seen in three out of the seven trials. Moreover, collapsed oligomers did not always aggregate together and stayed in solution. On the contrary, PEG oligomers expanded their structure in the unrestrained phase across simulation trials, regardless of whether they were undergoing protein contact. Only trials 1 and 3 at low extension resulted in irreversible ID2S binding when the restraint was on, likely due to PEG oligomers being free in solution with minimal polymer-polymer interactions occurring. Most protein-polymer binding occurred in the unrestrained phase, since the oligomers

exhibit more expanded conformations as seen with the broader R_g probability density distribution. PLGA-PEG/ID2S systems at low extension encountered the lowest frequency of protein-polymer interactions across trials, whereby trials 5 and 6 contained oligomers that underwent irreversible binding events. Most oligomers were free in solution and underwent transient prot-poly and poly-poly interactions but, in the unrestrained phase, they did not expand to a great extent and stayed relatively collapsed similar to PLGA/ID2S systems.

As for the medium extension case, oligomers in three PLGA/ID2S trials experienced protein binding but, in other trials, they were mostly dominated by poly-poly interactions or were free in solution. Across all trials, oligomers' R_g decreased quickly to values representative of a collapsed state in the unrestrained phase, even if irreversible binding occurred. In trial 2 however, one oligomer irreversibly bound to ID2S and stayed at the same R_g value even after the restraint was turned off, confirming that favorable protein-polymer binding modes are dependent on a polymeric chain's level of expansion. PEG/ID2S systems showed oligomers more frequent prot-poly interactions only in both the restrained and unrestrained phases across trials, in addition to a slight drop in R_g and broader R_g distribution in the last 100 ns. Low frequency of prot-poly interactions was observed across 4 PLGA-PEG/ID2S trials, with oligomers being free in solution or undergoing poly-poly interactions. Interestingly, oligomer binding, similar to PLGA/ID2S trial 2, was also observed with PLGA-PEG/ID2S trial 2, whereby an oligomer stayed near the same R_g as it was in the restrained phase. Furthermore, some copolymer oligomers in the unrestrained phase did undergo collapse but this was not always the case across trials, unlike the PLGA/ID2S systems; some oligomer probability density distributions were similar to those of PEG oligomers at medium extension. At the simulated contours length used for this study, medium level of extension, for PEG/ID2S systems, seems to promote the most frequent prot-poly interactions when the restraint was on; for PLGA/ID2S and PLGA-PEG/ID2S systems, binding events without a drastic collapse in the oligomer structure are observed, indicating that these interactions at a medium level of extension are highly favorable and can overcome dominant LGA self-interactions in the unrestrained phase.

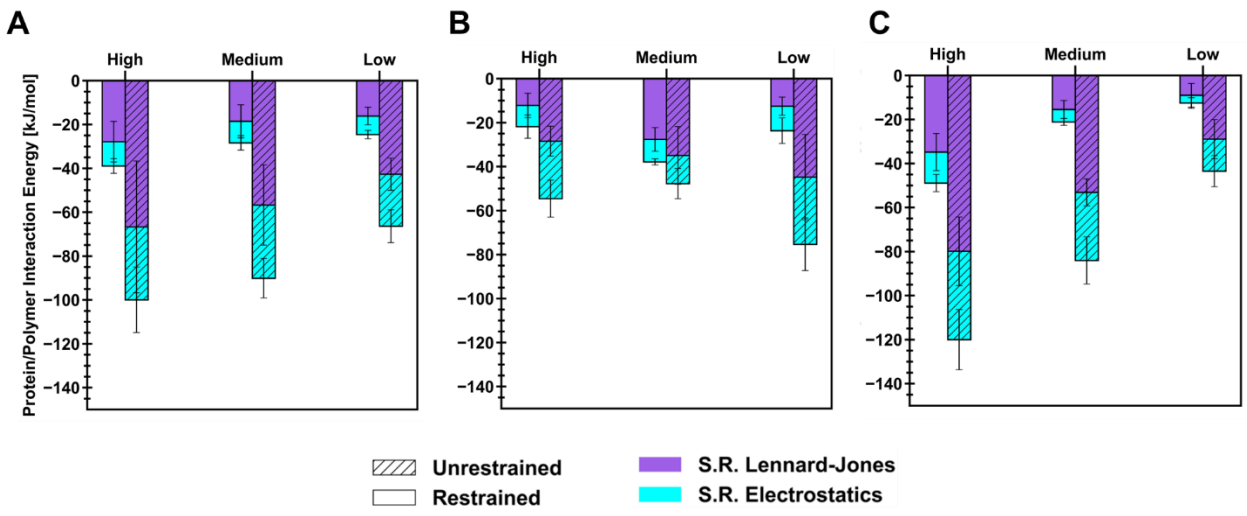

Figure 4-4: Protein/polymer interaction short-range interaction energies for each level of extension for **(A)** PLGA/ID2S, **(B)** PEG/ID2S and **(C)** PLGA-PEG/ID2S. From each stacked bar, the sum of Lennard-Jones and electrostatic interaction energies equal the total interaction energy. Mean and its standard error (SE) came from first averaging interaction energies across oligomers (n = 3/trial) then across simulation trials (n = 7).

Energetics of short range (SR) non-bonded interactions were extracted from polymer/ID2S simulations to examine the extent of favorable binding between the protein and oligomers, in addition to evaluating oligomer-oligomer behavior. As seen in **Figure 4-4**, the total interaction energy (E_{int}) across different levels of extension for each polymer type increases after the restraint is turned off but the magnitude of its increase is seen to be dependent on how extended the polymer is, in the restrained phase. This increase can be attributed to the polymer oligomers interacting with more ID2S surface AAs, while attempting to reach their equilibrium conformations when dissolved in pure water. For the PLGA/ID2S simulations, a linear decrease in the mean E_{int} is observed as the set oligomer conformation is shifted from high to low, with the high extension case showing the most favorable protein-polymer binding. SR Lennard-Jones interaction energies are larger than SR electrostatics across the different levels of extension in both phases, showing that dispersion forces were primarily responsible for PLGA binding. Polymer-polymer interaction energies (**Figure C-4**) were similar in magnitude to protein-polymer interaction energies at high and medium extension, showing that expanded structures was important to aiding protein-polymer binding despite the strong attraction between PLGA oligomers. As for the PEG/ID2S systems, a significant increase in E_{int} is seen at high and low extension in the unrestrained phase, when compared to the medium level of extension

system with only a slight increase in E_{int}. Furthermore, change in E_{int} in the unrestrained phase was observed to be the largest at the low level of extension whereas, for the high level of extension, E_{int} is similar in magnitude to the medium extension case. This suggests that collapsed PEG oligomers' return to good solvent conditions, which led to an expansion of chain structure, resulted in more favorable protein-polymer interactions, when compared to the other level of extension. Polymer-polymer interactions were minimally attractive in nature (**Figure C-3**), further confirming the prevalence of polymer-solvent interactions that are driving the increase in R_g in the unrestrained phase for PEG oligomers in the low extension case, leading to an increase in more favorable protein-polymer interactions. For highly extended PEG chains, their conformations shifted to lower R_g values matching those seen for PEG oligomers in water at the given contour length, allowing them to interact with ID2S surface at a similar magnitude as the medium level of extension.

The trend in E_{int}, observed for the PLGA/ID2S simulations, is the same for the PLGA-PEG/ID2S system. At the highest extension, protein-polymer interactions are most favorable in the restrained and unrestrained phases, when comparing to medium and low extension levels. The added effect of copolymerization is likely responsible for the observed differences since an expanded PLGA structure in the restrained phase, in addition to PEG domain being in good solvent conditions and PLGA relaxation around the ID2S surface in the unrestrained phase, resulted in overall stronger binding. Moreover, strong polymer-polymer interaction between copolymer oligomers were significantly reduced when compared to PLGA/ID2S systems (**Figure C-4**), further supporting the idea that the PEG domain mediates the extent of LGA interactions and allows for more favorable PLGA/protein binding. Such insights confirm the importance of choosing the right solvent when formulating protein-loaded PLGA-PEG nanoparticles because protein-polymer interactions are likely to be more favorable when PLGA and copolymer domains are highly extended prior to NP formation and growth. Furthermore, neutron scattering measurements of BSA-PEG aqueous mixtures by Abbott et al. showed that PEG had a slightly attractive interaction with BSA, despite having a net repulsive interaction[100]. PEG oligomers have been observed to undergo transient binding with ID2S that is generally less favorable when compared to the other polymer types, corroborating their experimental results.

4.2.2 Characteristics of the protein-polymer interface

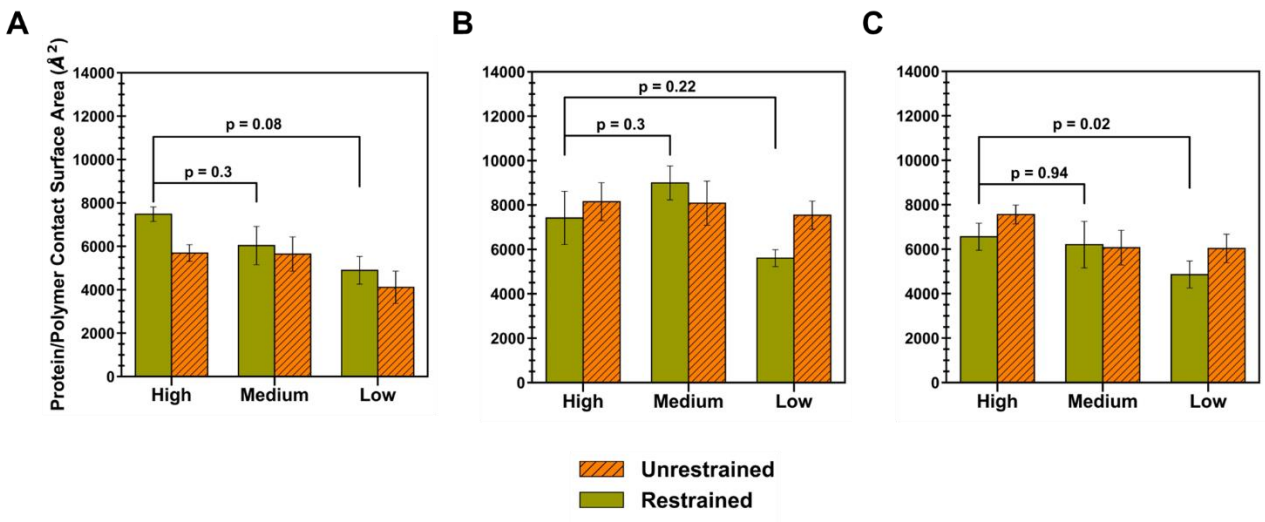

Figure 4-5: Protein/polymer contact surface area (CSA) at each level of extension for **(A)** PLGA/ID2S, **(B)** PEG/ID2S and **(C)** PLGA-PEG/ID2S. Mean values and their respective SE came from averaging CSA values of the protein/polymer interface across simulation trials (n = 7). Statistical significance was evaluated using the Wilcoxon signed-rank test, a non-parametric version of the paired T-test.

In order to understand how the mean interaction surface area varied across polymer types and levels of extension, the protein/polymer contact surface area (CSA) was calculated for each simulation trial and averaged, for both the restraint on and off phases (**Figure 4-5**). Briefly, the CSA is given by equation 28, where the interface residues are defined by whether a surface residue had a non-zero occupancy value. Out of a total ID2S surface area of 20,438 Å^2, oligomers at high extension, for PLGA/ID2S systems, had the largest average CSA in the restraint on phase, followed by medium and low extension systems; however, a drop in CSA was observed for the high extension case in the restraint off phase and across simulation trials, likely due to dominant LGA interactions causing oligomers to exhibit a more semi-collapsed conformation on the ID2S surface. For the restrained phases, the CSA for the high extension case was nearly statistically significant (p = 0.08) when compared to the low extension case. A negligible change in CSA was observed for the medium extension case while a slight drop in CSA was seen for the low extension, while in the restraint off phase. The number of trials with CSA values greater than 8000 Å^2 decreased from three in the medium extension case to zero in the low extension case (**Figure C-5**), showing that expanded PLGA chain conformation results in a larger protein/polymer contact area. As for the PEG/ID2S systems, medium extension of PEG oligomers led to the largest average CSA, when compared to other extension cases. As

the restraint is turned off, PEG oligomers contact more of the ID2S surface, as seen with the slight rise in CSA for the high extension case and a greater CSA increase for the low extension case; this was also observed across multiple trials, with the medium extension case resulting in trials with the largest CSA values (>8000 $Å^2$) compared to the other polymer types.

For the PLGA-PEG/ID2S systems, high and medium extension case showed similar CSA values ($p = 0.94$), while the low extension case had a statistically significant lower CSA ($p = 0.02$) when compared to the high extension case in the restrained phase. A majority of trials at high and medium extension had protein-polymer interfaces increase in CSA, in the unrestrained phase, but the overall magnitude of the CSA at low extension was much lower in both phases when compared to high extension. Furthermore, less contact was observed as copolymer oligomers went from medium to low extension, similar in trend as the PLGA/ID2S systems. This shows that PLGA-PEG chain expansion in good solvent conditions should allow for greater and more favorable contact with the protein. Moreover, PEG chains moving from a collapsed to semi-extended state results in a larger protein-polymer contact that is less favorable, when compared to PLGA domains. In the context of protein release from a PLGA-PEG nanoparticle in an aqueous environment, PLGA chain collapse, in nanoparticle regions where water absorption is significant and LGA interactions are dominant, PEG-protein interactions are likely not favorable enough to prevent the protein drug from diffusing out of the nanoparticle, reducing the therapeutic efficacy of the delivery vehicle. In addition, protein interactions with collapsed LGA domains would be weak and occur at less frequently with a small CSA, causing further drug leakage from the nanocarrier as water diffuses into its matrix. Future drug release studies should consider the relaxation time of the PLGA-PEG polymer at different molecular weights in solvent mediums of interest since a fast transition from a highly extended to collapsed conformation during nanoparticle formation or during drug release could have a critical impact on protein-polymer interactions, which in turn influences protein loading and *in vivo* release profiles.

Physiochemical descriptors of the protein-polymer interfaces, used to calculate CSA for each simulation trial, were ranked relative to those properties for all surface patches to examine how their chemical nature shifts across polymer types. Hydrophobicity, polarizability, normalized Van der Waals volume and graph shape index of all surface patches were distributed between -0.8 –1.4, 0 – 0.26, 2 – 5 and 1.4 – 3.4 respectively, as shown in **Figure 2-3** in the method chapter. Graph shape index interface ranks across polymer types and levels of extension were between rank 4-10, with no major discernable trends (**Figure C-6**). As for the hydrophobicity

interface rank distribution, most PLGA/ID2S interfaces from the restrained phase at high extension had a rank of 7 (Hydrophobicity ~0.10), which then shifted to rank 5 (Hydrophobicity ~0.16) in the unrestrained phase. This shows that PLGA collapse around the ID2S surface is resulting in interaction with hydrophobic residues, resulting in a more hydrophobic interface. For the low extension case for both restraint on and off phases, a broad interface rank distribution (rank 3-10) was observed while a slightly narrower spread (rank 4-7) was seen for the medium extension case. PEG/ID2S (rank 5-8) and PLGA-PEG/ID2S (rank 4-9) interfaces also showed the same wide spread and did not show any major differences across the levels of extension.

Polarizability and normalized Van der Waals volume (**Figure C-7**), on the other hand, were more valuable in understanding variations in the physiochemical properties of the protein-polymer interface. At high extension, PLGA/ID2S interface ranks for polarizability were centered at rank 7-8 (polarizability ~0.14) in the restrained phase but shifted to a broader rank spread between rank 4-8 in the unrestrained phase, indicating that oligomers were also interacting with more polarizable residues during their contraction around ID2S. At low and medium level of extension regardless of oligomer extension, polarizability ranks were between 4-8, with a peak at rank 6-7. For PEG/ID2S interfaces, polarizability ranks were centered between rank 6-8 across different extension levels whereas for PLGA-PEG/ID2S interfaces, polarizability ranks at high extension fell mostly on rank 6 in the restrained phase but shifted to ranks 3-10 in the unrestrained phase, similar in trend to PLGA/ID2S interface ranks; this could be due to PEG domains interacting with more ID2S surface AA in some simulation trials after the restraint is off. Lastly, normalized Van der Waals volume parameter ranking for PLGA/ID2S at high extension showed 6 out 7 interfaces, from the restrained phase, had a rank of 7 (normalized Van der Waals volume ~3.3), whereas some interfaces from the unrestrained phases changed to rank 5-6; such differences indicate that the PLGA oligomers contact more residues with slightly larger normalized Van der Waals volume when released from the restraint. At low and medium extension, normalized Van der Waals volume ranks are distributed between 4-8, regardless of restraint on or off phases. PEG/ID2S (rank 4-8) and PLGA-PEG (rank 4-9) interfaces shows similar spread in normalized Van der Waals volume ranking across levels of extension and in both phases, with peaks observed between rank 6-7. Overall, shifts in parameter ranking distribution for the protein-polymer interfaces have been useful to aiding identification of unique interface characteristics, in addition to understanding how the unrestrained phase results in PLGA/ID2S interfaces that are more polarizable, larger in their Van der Waals volume and slightly more hydrophobic.

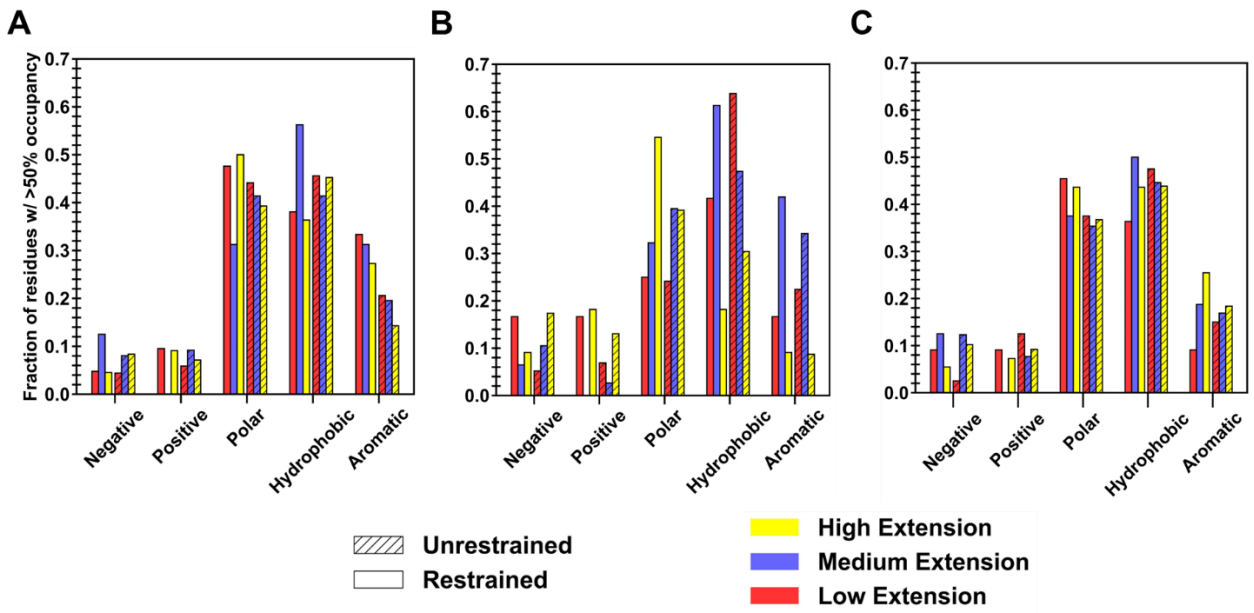

Figure 4-6: Fraction of residues, constituting the averaged collapsed interface (n = 7) from both the restraint ON and OFF phase, that have occupancy values greater than 50% at different levels of oligomer extension for **(A)** PLGA/ID2S **(B)** PEG/ID2S and **(C)** PLGA-PEG/ID2S systems. Amino acids are grouped into 5 respective categories: negative, positive, polar, hydrophobic and aromatic, where aromatic residues are also counted in the hydrophobic group

Surface residues that are responsible for favorable binding for each polymer type were extracted by first generating a collapsed protein/polymer interaction interface from all simulation trials in both restraint on and off phases (2 averaged interfaces per level of extension), containing AA's that have a non-zero percent occupancy. A cutoff of >50% was then applied to each collapsed interface to sub-select for high residency residues, whose AA composition can be seen in **Figure 4-6**. For clarity, high residency residues are AAs that have >90% occupancy while residues with occupancies between 50-90% are considered to be important in constructing the protein-polymer interface. Given that ID2S surface residues are composed of ~30% polar, ~48% hydrophobic, ~12% negatively and ~9% positively charged residues at neutral pH in its native state (**Figure C-8**), it is not surprising that a large fraction of residues are primarily polar and hydrophobic, with fractional values being less than 0.15 for charged surface residues across polymer types. The ID2S-PLGA filtered interface, in the restrained phase at high and low extension, was primarily composed of mostly polar then hydrophobic residues, where aromatic groups contribute a significant proportion to the fraction of nonpolar residues that driving binding. In the unrestrained phase, the fraction of hydrophobic residues increases while that of polar and aromatic residues decreases, suggesting that PLGA oligomers increased

their interactions with non-aromatic hydrophobic residues as they were collapsing around the ID2S surface. As for the medium extension case when the restraint is on, hydrophobic fraction of residues was the largest (~0.6) while polar and aromatic fractions were similar in value, indicating that this intermediate R_g value is highly favorable for hydrophobic interactions to occur between PLGA and ID2S. In the restraint off phase, an increase in the polar fraction was observed along with a decrease in hydrophobic and aromatic fraction, showing that both polar and hydrophobic residues are driving PLGA oligomer stabilization on the ID2S surface; the presence of methyl and carboxyl groups within polymer chain are likely mediating favorable contact with ID2S, resulting in large occupancy values and favorable polymer binding.

As for PEG/ID2S collapsed interface in the restrained phase at high extension, polar residues made up a large portion, followed by positively charged and hydrophobic residues; in the unrestrained phase, the fraction of polar and positively charged residues drops while an increase in the fraction of negatively charged and hydrophobic residues is observed. This could be due to the shift of PEG oligomers toward ideal chain conformations, resulting in a greater contact with non-aromatic hydrophobic residues. At low extension case in the restrained phase, charged and aromatic residues make up a similar fraction (~0.15), with the rest of the interface being composed of hydrophobic residues. Polar fraction of residues remains constant in the unrestrained phase while a drastic increase in the hydrophobic residues fraction and decrease in the charged residue fraction is observed. Such trends indicate the increase in ID2S-PEG CSA in the unrestrained phase is due to expanded PEG oligomers contacting a greater number of hydrophobic residues. At medium extension, polar and hydrophobic residues composed a large fraction of the collapsed interface in both restraint on and off phases, with their slight decrease and increase in the polar fraction being seen in the unrestrained phase; this can be explained by strength of polymer-solvent interactions that is pulling PEG oligomers away from the ID2S surface and allowing them to explore other polar residues.

At high extension for the PLGA-PEG/ID2S interfaces, polar and hydrophobic residue fractions combine to be ~80%, for the restrained phase, but polar and aromatic residue fraction decrease while the hydrophobic fraction stays constant for the unrestrained phase. This could be due to persistent PLGA-ID2S interactions and PEG's slight attraction to ID2S, across the various simulation trials. Polar and negatively charged fraction of residues decreases when compared interfaces from both phases in the low extension case while polar and hydrophobic fraction increases. For the medium extension case, a slight decrease for the hydrophobic and polar categories, while an increase in positively charged fraction is seen, indicating that

irreversible copolymer-ID2S binding could be attributed to polar residues, in addition to a mixture of aromatic and non-aromatic hydrophobic AAs. Overall, the interplay of chemical moieties on a polymer chain and its time averaged conformation in solution have a critical impact on whether strong protein-polymer binding events occur during nanoparticle formulation or degradation.

4.2.3 Surface patch characteristics at high and low extension

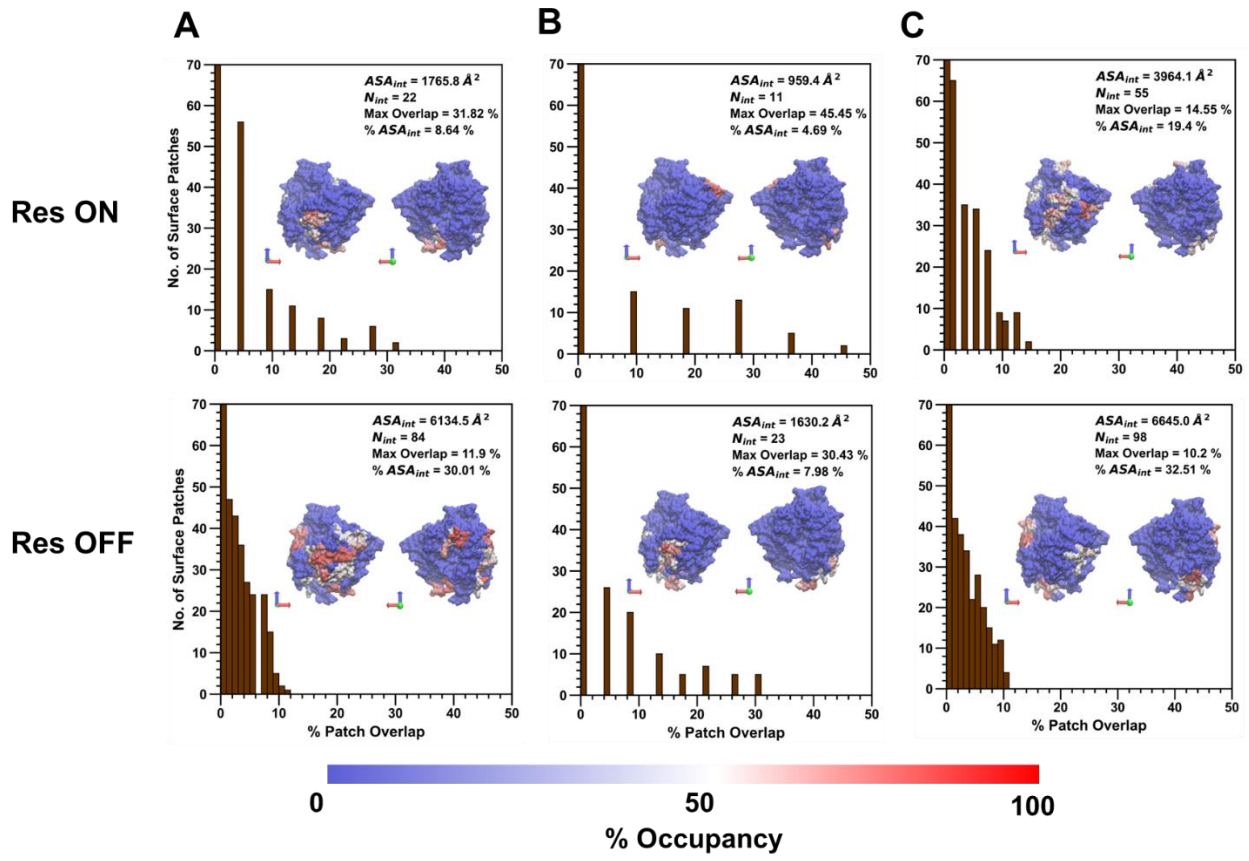

Figure 4-7: At high extension, the distribution of the number of contiguous surface patches that overlap with the collapsed, averaged protein-polymer interface containing residues with >50% occupancy, at various overlap percentages, is shown for **(A)** PLGA/ID2S, **(B)** PEG/ID2S and **(C)** PLGA-PEG/ID2S. Within each plot is the visualization of the collapsed interface whereby red indicates max occupancy and blue indicates no occupancy. ASA_{int} is the surface area of the collapsed prot-poly interface, N_{int} is the total number of residues in the collapsed interface and % ASA_{int} is the percentage of ID2S surface area that is composed of prot-poly collapsed interface. Max overlap value for each polymer/ID2S interface is also shown in each plot above.

Following the analysis of AA composition of the collapsed protein-polymer interface containing residues with >50% occupancy across polymer types, the percent patch overlap of generated surface patches with the polymer/ID2S interface (**Figure 4-7**) was calculated for the high and low extension cases to show how residues clusters containing high residency surface AAs can be identified using this patch analysis method by Jones et al.[68] Across all polymer/ID2S systems, most surface patches had an overlap with the collapsed interface between 0-1%, with ~8-15% of patches having an overlap greater than 5%; moreover, N_{int} was smaller in the restraint on phase, when compared to the restraint off phase, across all polymer types. This is likely due to the presence of the restraint ensuring that the polymer oligomers maintain the set R_g; thus, monomers that are weakly bound to the ID2S surface will be pulled off, resulting in a large majority of interacting surface residues having low occupancy values. This idea is also supported by analysis of non-bonded interaction energetics and the much smaller ASA_{int} of the filtered interface, when compared to the average CSA for polymer/ID2S systems at high extension for the restrained phase. As N_{int} increases for the unrestrained phases, a general drop in the percent max overlap for certain surface patches is observed. For the PLGA/ID2S filtered interface, the decrease in the % max overlap from 31.8% to 11.9% signifies that the increase in the interface area results in a fewer amount of contiguous surface patches that overlap with a high percentage of the protein/polymer interface. Regardless, most surface patches that had the largest % overlap typically contained high residency residues. As for the PEG/ID2S interfaces, ASA_{int} and N_{int} were smaller in magnitude than the other polymer/ID2S interfaces, providing more evidence that PEG-water interactions were dominant and PEG oligomers preferred to be solution; furthermore, % max overlap only dropped to 30.4%. The effect of high extension and copolymerization can also be seen in the PLGA-PEG/ID2S interface, whereby ASA_{int} and N_{int} were largest and % max overlap value were the smallest in both restraint on and off phase when comparing to the homopolymer/ID2S systems.

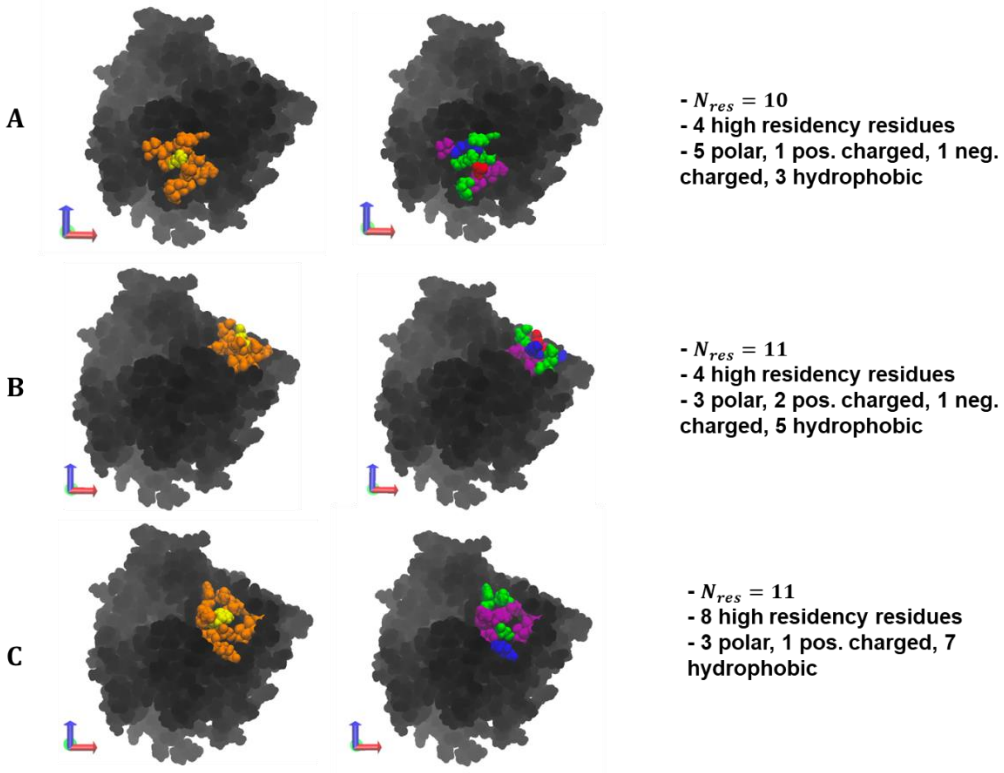

Figure 4-8: At high extension, an individual surface patch with the highest patch overlap percentage with the filtered interface is shown from **(A)** PLGA/ID2S Res OFF phase, **(B)** PEG/ID2S Res ON phase and **(C)** PLGA-PEG/ID2S Res ON phase. In the first column of images, yellow signifies the central residue from which the patch was generated and orange are residues that were included in the given patch by satisfying the θ_{cut} cutoff requirement. Amino acid composition of each patch can be seen in the second column of images, where blue residues are positively charged, red residues are negatively charged, green residues are polar and purple residues are hydrophobic. N_{res} is the number of residues in the selected patch.

Visualization of the max overlap patches and the AA composition can be found in **Figure 4-8**. Across the polymer types, selected residue clusters consisted of 10-11 residues, with a majority of them being polar and hydrophobic and varying in shape and morphology. The selected surface patches, across the polymer/ID2S systems, contained 4 and 8 high residency residues, showing that this method is capable of identifying residue clusters containing polymer binding hotspots. For the PLGA/ID2S patch, the distance between the 3 hydrophobic residues, filled with the 5 polar residues, could result in a binding mode where carboxyl groups on the oligomer chain first undergo dipole-dipole interactions then as the oligomer settles on the ID2S surface, methyl groups on either side of the chain experience hydrophobic interactions, leading to irreversible binding. The PEG patch has 3 charged and 3 polar residues, suggesting that this residue cluster would be relatively hydrophilic and would undergo hydrogen bonding or dipole-

dipole interactions with PEG oligomers; hydrophobic residues are interspersed within the patch and provide it with some hydrophobic character. 7 hydrophobic residues cover most of the PLGA-PEG patch, along with the 3 polar and 1 positively charged residues. The morphology of the residue cluster is likely to promote binding of the PLGA domain and allow for oligomer collapse, since multiple methyl groups can interact with the hydrophobic strip as the chain changes conformation. Overall, selected patches at the highest % overlap with the prot-poly interaction interface allows for systematic identification of binding zones that can be evaluated for their uniqueness and be engineered to increase nanoparticles drug loading or tune drug release profiles.

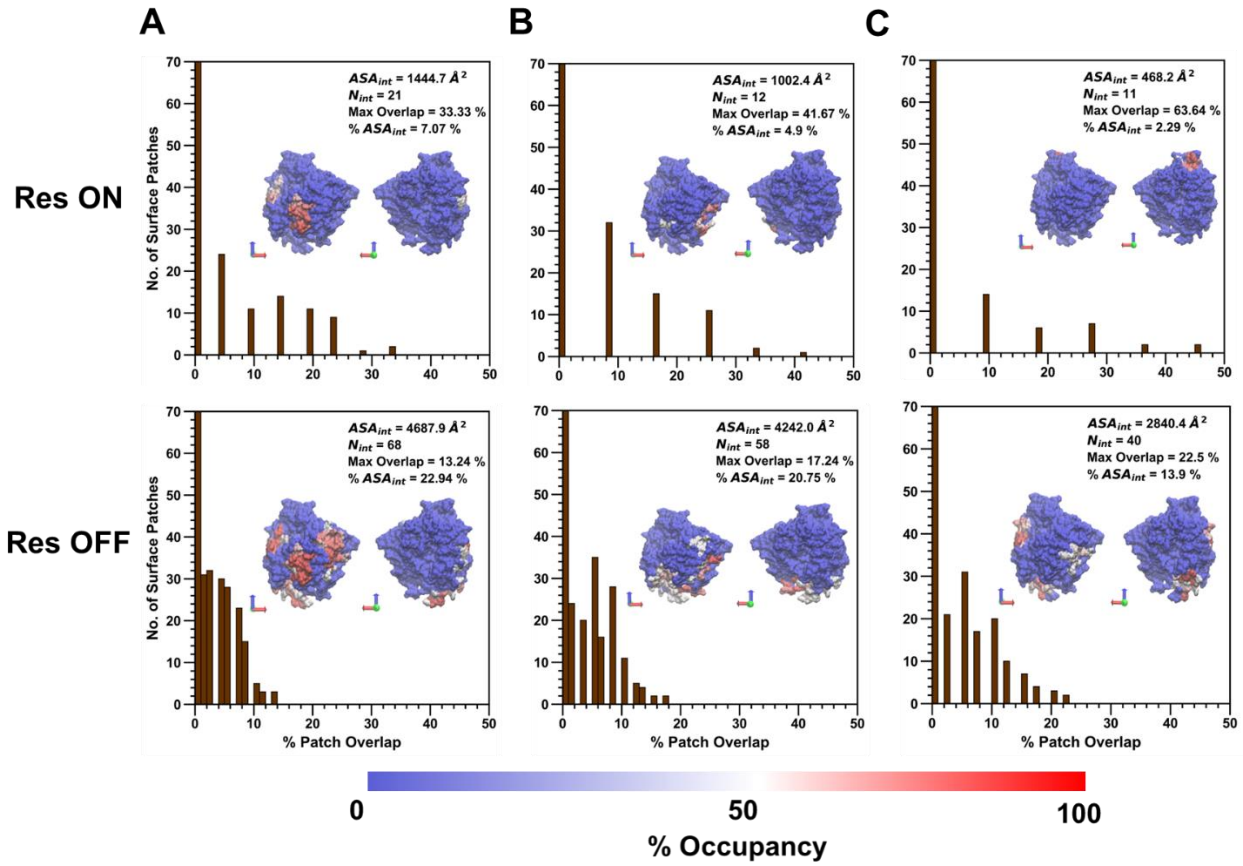

Figure 4-9: At low extension, the distribution of the number of contiguous surface patches that overlap with the collapsed, averaged protein-polymer interface containing residues with >50% occupancy, at various overlap percentages, is shown for (A) PLGA/ID2S, (B) PEG/ID2S and (C) PLGA-PEG/ID2S. Within each plot is the visualization of the collapsed interface whereby red indicates max occupancy and blue indicates no occupancy. ASA_{int} is the surface area of the collapsed prot-poly interface, N_{int} is the total number of residues in the collapsed interface and % ASA_{int} is the percentage of ID2S surface area that is composed of prot-poly collapsed interface. Max overlap value for each polymer/ID2S interface is also shown in each plot above.

As seen in **Figure 4-9**, PEG/ID2S and PLGA-PEG/ID2S filtered interfaces in the restrained phase have smaller ASA_{int} and N_{int} values, when compared to the PLGA interface. Moreover, the % max overlap of certain surface patches increases as the interface gets smaller, when going from PLGA, PEG then the copolymer. This shows that the PEG and copolymer oligomers at low extension undergo mostly reversible binding with ID2S and tend to be free in solution or interacting with another oligomer in solution. PLGA oligomers in both the restrained and unrestrained phased, however, are still able to undergo favorable and irreversible binding with 7% and 22.9 % of the protein surface area, even in the presence of strong polymer-polymer interactions. Comparing to the PLGA interface surface area at high extension from the restraint off phase, ASA_{int} and N_{int} are smaller, showing that expanded PLGA oligomers do lead to interactions with more surface residues. PLGA/ID2S interface % max overlap is lowest among the polymer types for interfaces from the restraint off phase. When the restraint is turned off, the large increase observed in $\mathrm{ASA_{int}}$ and N_{int}, in addition to the drop in % max overlap for the PEG/ID2S interface, further confirms the oligomers' return to good solvent conditions from a collapsed conformation results in an increased interaction with ID2S surface residues. A similar trend is also seen for PLGA and PLGA-PEG interfaces; for the copolymer, the PEG domain at ideal chain conditions and the collapsed PLGA domain interacting with mostly hydrophobic and polar residues are likely why there was an increase in CSA from the unrestrained phase.

For the PLGA 11-residue patch at low extension (**Figure 4-10**), hydrophobic residues surround polar residues and charged residues seem to be slightly less exposed to the surface. Collapsed oligomers with methyl groups on their surface were likely able to favorably contact the front facing hydrophobic residues, leading to irreversible binding. As for the PEG patch, 4 polar and 4 hydrophobic residues primarily make up this 8 residue cluster, whereby 7 AAs are high occupancy values. This may be due to the spread of polar residues that allow for multiple opportunities for hydrogen bonding, while hydrophobic residues interact with methylene groups on the PEG chain. The 7 residue patch for the copolymer interface is made up of 4 polar and 2 hydrophobic AAs, with 1 positively charged residues. Comparing to location of high residency residues for PLGA-PEG/ID2S interface in the restraint on phase, it is reasonable to conclude this patch and its arrangement of polar and hydrophobic residues resulted in favorable copolymer binding, driven by the presence of methyl, carboxyl and/or ester groups. It should be noted that these selected surface patches also promoted strong oligomer binding across the polymer types at both levels of extension. A stricter occupancy cutoff can be employed to result in a smaller protein-polymer interface, leading to larger % max overlap values for certain surface patches. Ultimately, examination of local chemical environment of the protein surface that

shows strong protein-polymer binding is a nontrivial task and this patch analysis provides a way to gain such information.

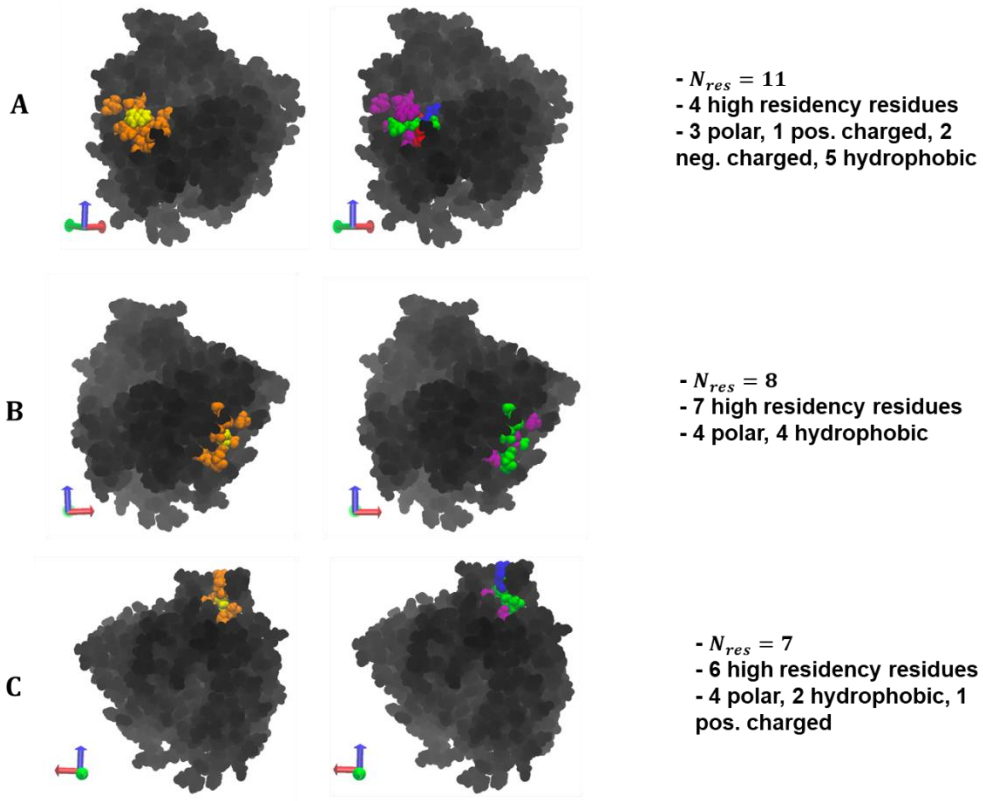

Figure 4-10: At low extension, an individual surface patch with the highest patch overlap percentage with the filtered interface is shown from **(A)** PLGA/ID2S Res OFF phase, **(B)** PEG/ID2S Res ON phase and **(C)** PLGA-PEG/ID2S Res ON phase. In the first column of images, yellow signifies the central residue from which the patch was generated, and orange are residues that were included in the given patch by satisfying the θ_{cut} cutoff requirement. Amino acid composition of each patch can be seen in the second column of images, where blue residues are positively charged, red residues are negatively charged, green residues are polar and purple residues are hydrophobic. N_{res} is the number of residues in the selected patch.

4.3 Conclusions

In this work, PLGA, PEG and PLGA-PEG oligomers were simulated in the presence of ID2S at various level of extension to examine the role of polymer conformation on protein-polymer interactions and provide mechanistic understanding into the molecular-level driving forces that are present during protein loading within PLGA-PEG nanoparticles and during nanoparticle dissolution leading to protein release. Highly expanded PLGA-PEG conformations were shown to lead to greater ID2S contact and subsequent movement to poor solvent conditions, by turning off the restraint, resulted in persistent binding even as strong LGA-LGA interactions dominate. Moreover, collapsed LGA domains can still favorably and irreversibly interact with ID2S. This suggests that relaxation times of the PLGA homopolymer and PLGA-PEG copolymer at different molecular weights in solvent mediums relevant to drug release studies should be considered. PEG oligomers' weak attraction to proteins, in addition to dominant LGA interactions, could explain why a burst release profile may be observed for some protein loaded PLGA-PEG nanoparticles. As water and ions infiltrate the nanoparticle matrix, PEG domains beginning to transition to ideal chain conformations. In addition, degradation of large PLGA domains into smaller oligomers results in their collapse into LGA-rich parts of the nanoparticle matrix, leading to decrease in protein-polymer interactions. Since the PEG shell prefers to be in the aqueous solvent and ions are promoting further hydration of proteins, released proteins will be able to diffuse into the surrounding solvent medium and be detected. Overall, successful encapsulation of protein drugs within PLGA-PEG nanoparticles and their administration *in vivo* will require protein-polymer interactions that can match or outcompete the strength of LGA self-interactions, in order to ensure proper shielding of the therapeutic protein from the surrounding environment.

4.5 Acknowledgements

This work was funded by National Science Foundation Grant #1703438. This work was facilitated through the use of advanced computational, storage, and networking infrastructure provided by the Klone Hyak supercomputer system at the University of Washington

5. Chapter 5: Phase behavior of polyelectrolyte complex coacervates and its implication on enzyme encapsulation within polymeric nanoparticles

5.1 Introduction

Synthesis of polymeric nanoparticles is a complex process that relies on key kinetic and thermodynamic behavior to form core-shell nanoparticles, when the PLGA-PEG copolymer is used. Depending on the process parameters and a protein hydrophilicity, a drug could end up either on the nanoparticle surface or within the nanoparticle core, resulting in a burst or sustained release profile *in vivo*. Control of drug localization is desired because of its effect on release but for this to occur, the extent of protein-polymer partitioning must be understood. First, organic solvents used during synthesis are non-ideal solvent environments for most proteins, due to their hydrophilic nature[2,3,12]. Nonpolar groups within the core of most proteins are likely attracted to hydrophobic organic solvent, possibly resulting in disruption of important tertiary structures required for their function. This, however, is not an issue when encapsulating hydrophobic small molecule drugs that have poor aqueous solubility, since the primary dominant interaction, hydrophobicity, can ensure entrapment within PLGA-PEG nanoparticles[14]. Using atomistic MD, Chapter 4 provided insights on whether PLGA/protein or PEG/protein interactions were more favorable when oligomers were in an extended conformation versus a collapsed state and revealed the dominance of LGA-LGA interactions. Nonetheless, past literature studies show that the hydrophilic nature of drugs prevents localization within the PLGA-PEG nanoparticle core, leading to surface adsorption; hence, low drug loading, little drug protection and a burst release have been observed[101,102]. Evaluation of catalase protection within PLGA-PEG nanoparticles was performed on particles made via the nanoprecipitation method, as seen in **Figure D-1**. Rapid catalase deactivation occurred within less than 2 hours, providing more evidence that protein-PLGA partitioning is not favorable. One mitigation strategy that has been used in past literature is to increase the amount of polymer used in the nanoparticle formulation[103,104]. Unfortunately, the consequence is that PLGA degradation products are likely to cause an acidic local environment, significant enough to impact protein structure or even denature the encapsulated drug[103]. Therefore, hydrophobic ion pairing (HIP)

complexation has gained popularity as a means to alter the protein surface to promote greater nanoparticle drug loading and increase drug protection in proteolytic conditions.

HIP complexation employs electrostatic interactions between ionizable surface amino acids and oppositely charged groups of ion-pairing (IP) agents (surfactants or polymers) to form a lipophilic protein/IP complex that can more easily partition into the PLGA-PEG nanoparticle matrix. Furthermore, the complexes can easily dissociate in the presence of excess ions of the same charge as the IP agent[103]. Past literature studies have shown how the type of IP agents used can impact HIP complex lipophilicity[89,103] and set the baseline for atomistic MD simulations to examine the key interactions involved in HIP. Patel et al. tested 3 different IP agents, sodium dodecyl sulfate (SDS), taurocholic acid and dextran sulphate (MW 5 kDa), on their ability to complex with IgG-Fab, a model antibody protein, and subsequently encapsulated the most hydrophobic Fab/IP complex within PLGA (85:15, MW 50-75 kDa) nanoparticles using a solid-in-oil-in-water (S/O/W) emulsion and a nanoprecipitation method[89]. Their results showed dextran sulphate formed the most hydrophobic complex with IgG-Fab at pH of 4 and the highest encapsulation efficiency was observed for S/O/W nanoparticles loaded with dextran sulphate/Fab complexes. Moreover, both methodologies formed nanoparticles less than 200 nm and had low polydispersity. Another study by Gaudana et al. prepared HIP complexes using bovine serum albumin (BSA), as their model protein and dextran sulphate (DS) as their IP agent, then formulated nanoparticles using PLGA (85:15, MW 50-75 kDa) as their polymer and the S/O/W emulsion as their method[105]. They showed that significant entrapment, even at low drug/polymer ratios; in addition, they used circular dichroism to show protein structure did not change through HIP complexation and nanoparticle formulation.

Despite the few experimental studies that applied HIP to drive protein loading within PLGA-based nanoparticles, no one has performed atomistic MD simulations that investigated how the complexation occurs when using a surfactant versus a polyelectrolyte. In addition, methodologies of exploring the phase behavior of complex coacervates have been primarily used for food science applications but not for therapeutic protein drug delivery[74,106,107]. Chapter 5 work uses MD simulations to examine protein stability and amino acid composition of favorable binding sites that promote electrostatic complexation. Using these computational insights, experimental investigation of which pH/Molar Ratio (MR) combinations results in water-insoluble polyelectrolyte coacervates was performed by constructing a phase state point diagram according to the methodology described by Lan et al.[106] The dual computational and experimental approach reveals why dextran sulphate is the ideal ion-pairing agent to form

complexes with therapeutic proteins and shows how the state point diagram can answer questions regarding the macroscopic and aggregation behavior of polyelectrolyte coacervates. Ultimately, HIP complexation provides a way to change a protein's surface characteristics without covalent bond formation and promote strong interaction with the PLGA domain that forms the core of the PLGA-PEG nanoparticle, resulting in greater protein encapsulation and protection in proteolytic conditions.

5.2 Results and Discussion

5.2.1 Molecular scale features of protein-IP complexes[2]

Three ion-pairing agents were chosen to be simulated with BSA, a model protein, in a pure water condition for 200 ns: dextran sulphate (DS), sodium dodecyl sulfate (SDS) and taurocholic acid (TA). In order to understand any potential differences in the behavior of DS, SDS and TA when binding to proteins, we performed MD simulations to evaluate the structure and dynamics of protein-ion complexes. BSA is well matched to catalase, having a similar profile of surface amino acid residues as shown in **Figure D-2**. As such, it is reasonable to expect that the nature and extent of protein-ion interactions we obtain from the BSA/ion-pairing agent simulations will provide useful insight to the behavior of other proteins with a similar surface profile.

Following completion of the MD simulations, we analyzed the structure of BSA and related conformational changes. We also analyzed the chemical interactions between ions and different types of residues at mole ratios of 128 (SDS/TA) and 5 (DS), which were selected to roughly control for a consistent number of anionic sulfate groups between DS (120 total) and SDS/TA (128 total). The MD simulation analyses for these systems are shown in **Figure 5-1**. As previously noted by Baler et al.[108], we expect BSA at pH 3.7 to undergo a conformational rearrangement of tertiary structure even on the timescale of MD simulation (usually hundreds of nanoseconds). **Figure 5-1A** shows BSA backbone RMSD from its crystal structure as a function of simulation time for the three ion-paired systems as well as a control system with only Cl-

[2] This section was reproduced with permission from Joseph, A.; Nyambura, C. W.; Bondurant, D.; Corry, K.; Beebout, D.; Wood, T.R.; Pfaendtner, J.; Nance, E.* Formulation and Efficacy of Catalase-Loaded Nanoparticles for the Treatment of Neonatal Hypoxic-Ischemic Encephalopathy. *Pharmaceutics* **2021**, *13*, 1131. https://doi.org/10.3390/pharmaceutics13081131

present for charge neuralization. The level of conformational change in the control and DS system (RMSD ~0.5 nm at 200 ns) corresponds well to the expected structure of BSA in the N-isoform. In contrast, BSA conformational changes in the SDS and TA systems (RMSD ~0.9 nm and ~1.1 nm at 200 ns) indicate that the protein is transitioning from the N to the F-isoform. This suggests that DS complexation is able to retain BSA's native state, unlike SDS and TA complexation. Snapshots of the final structures of each of the simulations are provided in **Figure D-2**. The individual domains of BSA do not undergo any significant unfolding for any of the systems (**Figure 5-1B**). Finally, we observe that SDS and TA behave similarly from the point of view of the dominant chemical interactions on the BSA surface (**Figure 5-1C**), showing significant interactions with hydrophobic and aromatic residues. In contrast, DS has comparatively very few interactions with these residue types.

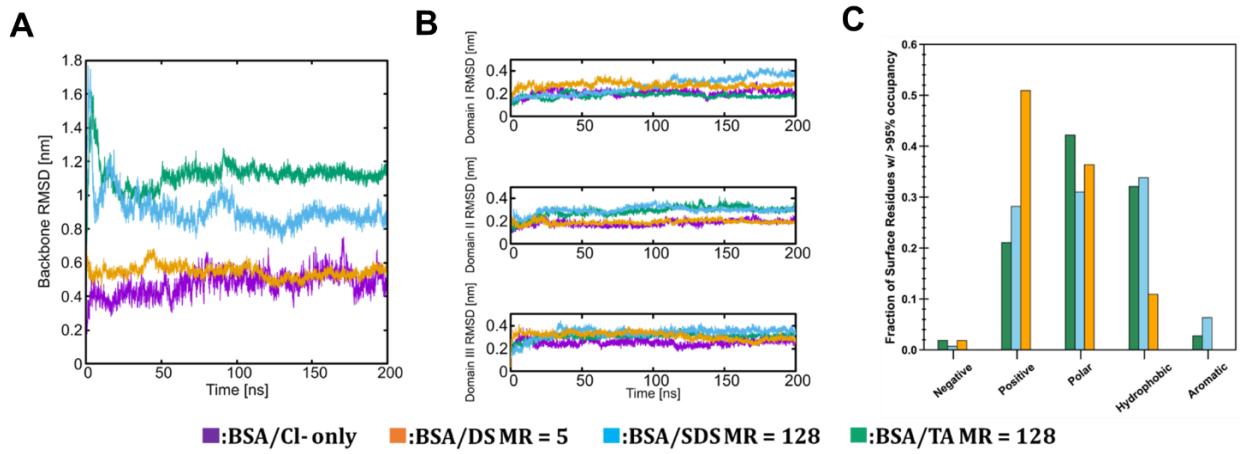

Figure 5-1: Characterization of BSA structure and IP agent interactions with surface amino acids. **(A)** BSA backbone RMSD (in nanometers [nm]) vs. time (in nanoseconds[ns]), **(B)** Backbone RMSD of each BSA domain vs time, **(C)** Fraction of surface residues with >95% occupancy vs. residue grouping. This figure was ascertained from Joseph et al.[12]

5.2.2 Elucidation of the state point phase diagrams of macromolecular complex coacervates formed between enzyme and dextran sulphate

Analysis of the catalase-DS state point diagram (**Figure 5-2**) reveals the pH/MR combination that result in formation of insoluble complexes. For pure catalase, clear and translucent solutions were observed from pH 2-2.5 then as pH increases past 2.5, turbid solutions, with and without precipitates, are formed. This behavior is likely indicative of catalase reducing its net positive surface charge, as the pH approaches its pI of 5.4. At very acidic pHs,

catalase enzymes have a strong net positive charge, meaning that electrostatic repulsion between adjacent proteins and further water hydration by counter-ions may be responsible for the transparency observed pH 2-2.5; at higher pHs, the net positive charge decreases in magnitude and protein-protein interactions become more dominant, leading to catalase aggregation and precipitation. With the addition of DS, the strength of electrostatic attraction with catalase can been seen at pH 2-2.5 because coacervate precipitates are formed across all MRs. The mixtures at pH 2/MR 7-13 and pH 2.5/MR 3 result in phase separation into a catalase-DS coacervate-rich and dilute phase, characteristic of insoluble complex formation. All other MRs between pH 2-2.5 results in partial phase separation, where some complexes remain stable in solution. At pH 3, precipitation and cloudy solutions are observed for catalase only and MR 1, while MR 3-13 only result in turbid mixtures; for pH >3, mixtures form a translucent solution with a slight level of opacity greater than turbid solutions. This indicates that pH 3 is the primary boundary transition pH across the tested MRs, whereby insoluble complex formation occurs below that pH and soluble complexes are made above that pH. In addition, DS addition to catalase solutions between pH 3.5-5.4 suggests that the polyanionic polymer is promoting solvation of catalase enzyme and reducing the extent of dominant protein-protein interactions. Such behavior may be due to DS increasing the hydration of the protein, eliminating the formation of precipitates.

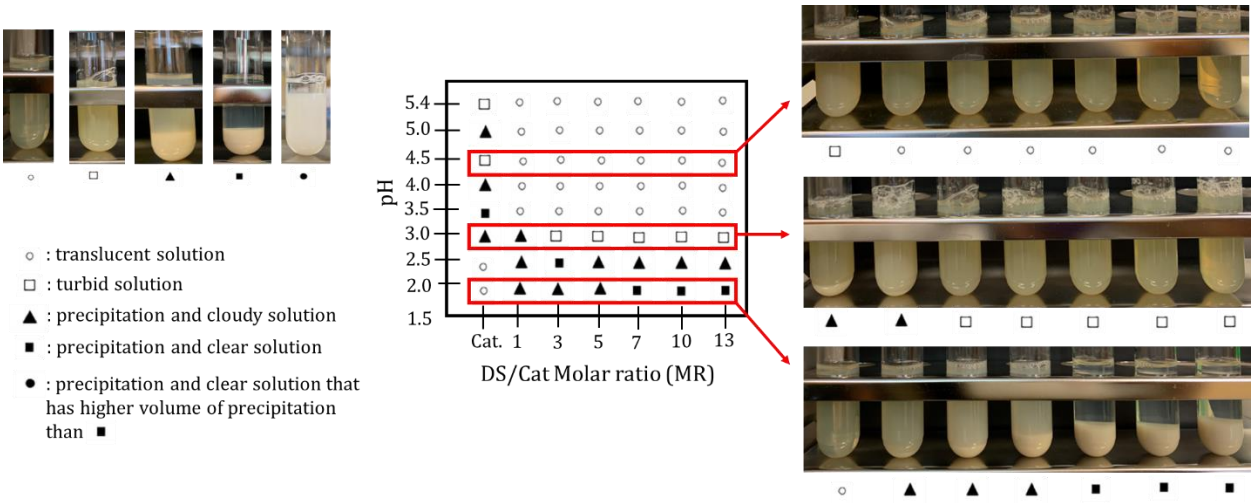

Figure 5-2: State point diagram of catalase and catalase-DS mixtures at different molar ratios. Prior to visual examination, these mixtures were placed in test tubes and allowed to rest for 24 hours at 4°C. Pictures used to construct the state points across all MRs for pH 2.0, 3.0 and 4.5 are shown to the right of the diagram. Five different symbols were used; specifically, ○, □,▲, ■, ● were used to represent translucent solution (○), turbid solution (□), precipitation & cloudy solution (▲), precipitation & clear solution (■), and precipitation & clear solution (●) with higher volume of precipitation than ■, respectively

In order to better understand the nature of electrostatic attraction between DS and catalase, zeta potential (ZP) measurements of the concentrated (**Figure 5-3A**) and diluted (**Figure 5-3B**) mixtures, prior to resting for 24 hours, were taken and plotted against DS:Cat molar ratio. Concentrated solutions with catalase only from pH 2-2.5 were found to have positive ZP values greater than +10 mV, in addition pH 2 having a diluted ZP value of -19.2 mV. This shows that DLS measurements can detect the large net positive charge of catalase that is likely driving electrostatic complexation with DS and resulting in insoluble complexes. As the DS:Cat MR increases, ZP values stay near ~+20 mV from MR 1-5 but then decays to ~-12 mV from MR 5-13 at pH 2 for the concentrated solutions; however, for the diluted solution, a gradual increase in ZP toward the neutral range (-10 to +10 mV) is seen, as MR increase. A similar trend is observed at pH 2.5, expect the decay in ZP for concentrated mixtures occurs between MR 1-5 and ZP values for the diluted solution gradually decreases to ~15 mV. Such an observation suggests that, with a strong net positive charge, catalase requires higher number of DS chains (MR >5) to saturate its surface before seeing a shift toward a net negative charge. It is likely that aggregation and precipitation is triggered by the reduction of short-range repulsive forces between complexes, due to the DS surface binding. Furthermore, as the net positive charge of catalase decrease (pH increasing toward pI), the number of DS chains that are needed to saturate the surface decreases, with the soluble complexes having a net negative charge (ZP ~-20 mV) in concentrated and near neutral charge (ZP ~-7 mV) in diluted solutions. The standard in food science literature has only to measure the ZP of concentrated solutions due to the ease of extracting trends on the nature of the electrostatic interaction but we propose that ZP measurements in diluted salt solutions are also necessary to get a more accurate measurement of the complexes' surface charge.

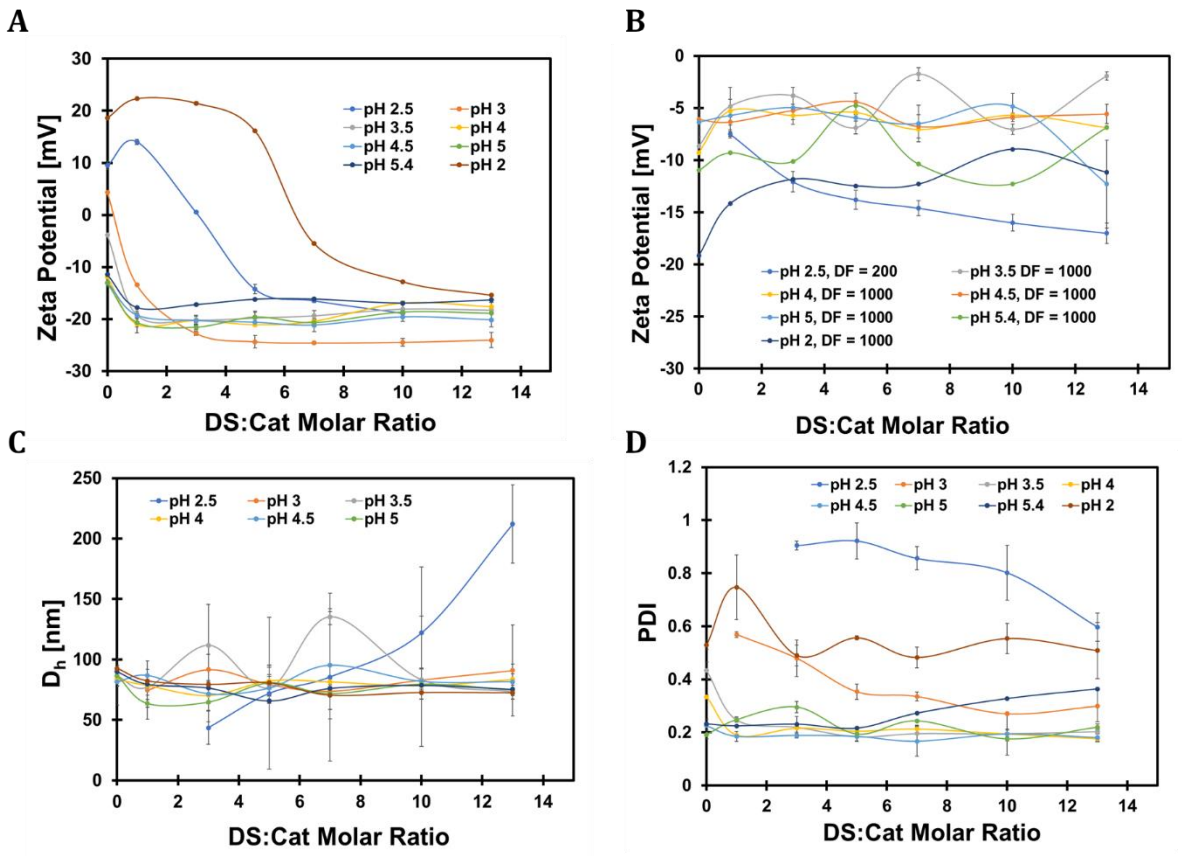

Figure 5-3: Catalase-DS coacervate characteristics, measured after the 1 hour mixing period, across the tested pH and MRs are shown, where **(A)** Zeta potential of coacervates in the undiluted mixture, **(B)** Zeta potential of coacervates diluted in 10 mM NaCl (1:1000 dilution), **(C)** Hydrodynamic Diameter (D_h) and **(D)** polydispersity of coacervates from dilution solutions were measured.

Hydrodynamic Diameter (D_h) (**Figure 5-3C**) and polydispersity (PDI) (**Figure 5-3D**) of coacervates were also evaluated by DLS to examine the colloidal characteristics at pH/MR ratios. Solutions containing only of catalase showed enzyme clusters were 70-90 nm in size across all pHs whereas PDI decreased from 0.6 to 0.2 as pH increased, possibly due to dynamics of protein-protein interactions. As DS:Cat MR increases at pH 2, size of insoluble complexes rose in value to ~200 nm while the PDI reduced from ~0.9 to ~0.6. All other pH values showed complexes' average size fluctuated between 50-150 nm while complexes from pH 3.5-5.4 had PDI values that fluctuated between 0.2-0.4 across DS:Cat MRs. pH 2 complexes stayed around a PDI of ~0.55 and pH 3 complexes decreased in PDI from ~0.6 to ~0.3. Soluble and insoluble complex formation between catalase and DS results in coacervates that <200 nm in size. As the pH transitions from 2 to 5, a wide size distribution of insoluble complexes which is

indicative of aggregation forces being present while for the soluble complexes, the size distribution narrows, showing that soluble complexes are monodisperse.

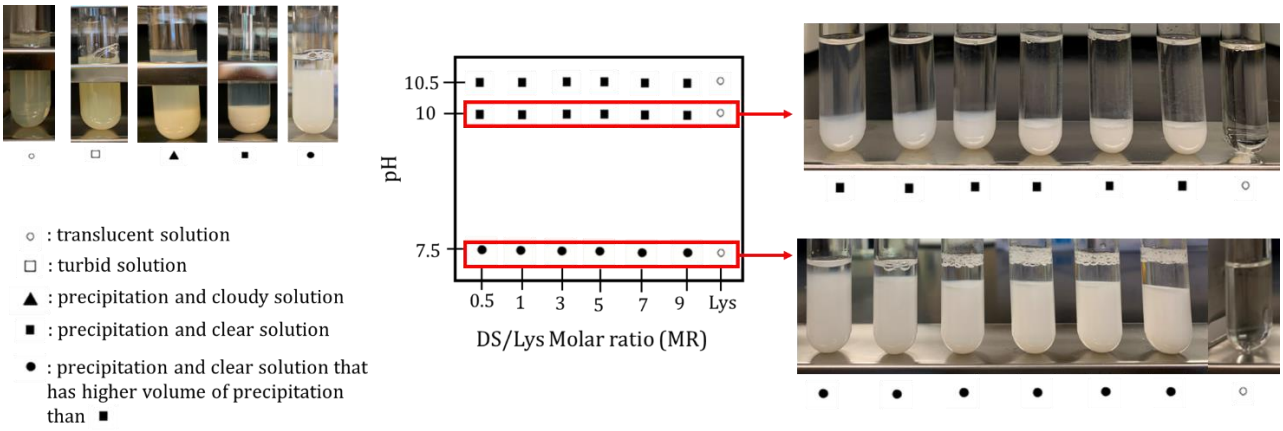

Figure 5-4: State point diagram of lysozyme and lysozyme-DS mixtures at different molar ratios. Prior to visual examination, these mixtures were placed in test tubes and allowed to rest for 24 hours at 4^{o}C, following mixing for 1 hour. Pictures used to construct the state points across all MRs for pH 2.0, 3.0 and 4.5 are shown to the right of the diagram. Two different symbols were used; specifically, ■ & ● were used to represent precipitation & clear solution(■), and precipitation & clear solution(●) with higher volume of precipitation than ■, respectively

State point diagram construction was also done with lysozyme (**Figure 5-4**), a smaller enzyme with a fewer number of ionizable surface amino acids, at pH 7.5, 10 and 10.5. Across the tested pHs, lysozyme only solutions were all translucent and showed no aggregation, unlike catalase. As DS was added, only two different phases (■ and ●) are observed, along with precipitate formation for all MR values. At pH 7.5, precipitate volumes were greater than those seen at pH 10 and 10.5, indicating that aggregation driving forces are weak and lysozyme complexes have higher colloidal stability in solution; in addition, those volumes seems to decrease as MR increases. This shows that the enzyme size and surface amino acid composition has a large impact on the macroscopic phase behavior that observed from the state point diagram. It is possible that a certain arrangement of ionizable groups, for proteins at specific molecular weights, promotes or discourages complex formation and precipitation. The potential for machine learning is high since such algorithms could employ atomic level descriptors of a protein's surface and the polyelectrolyte to see which pH/MR combinations result in insoluble complex formation.

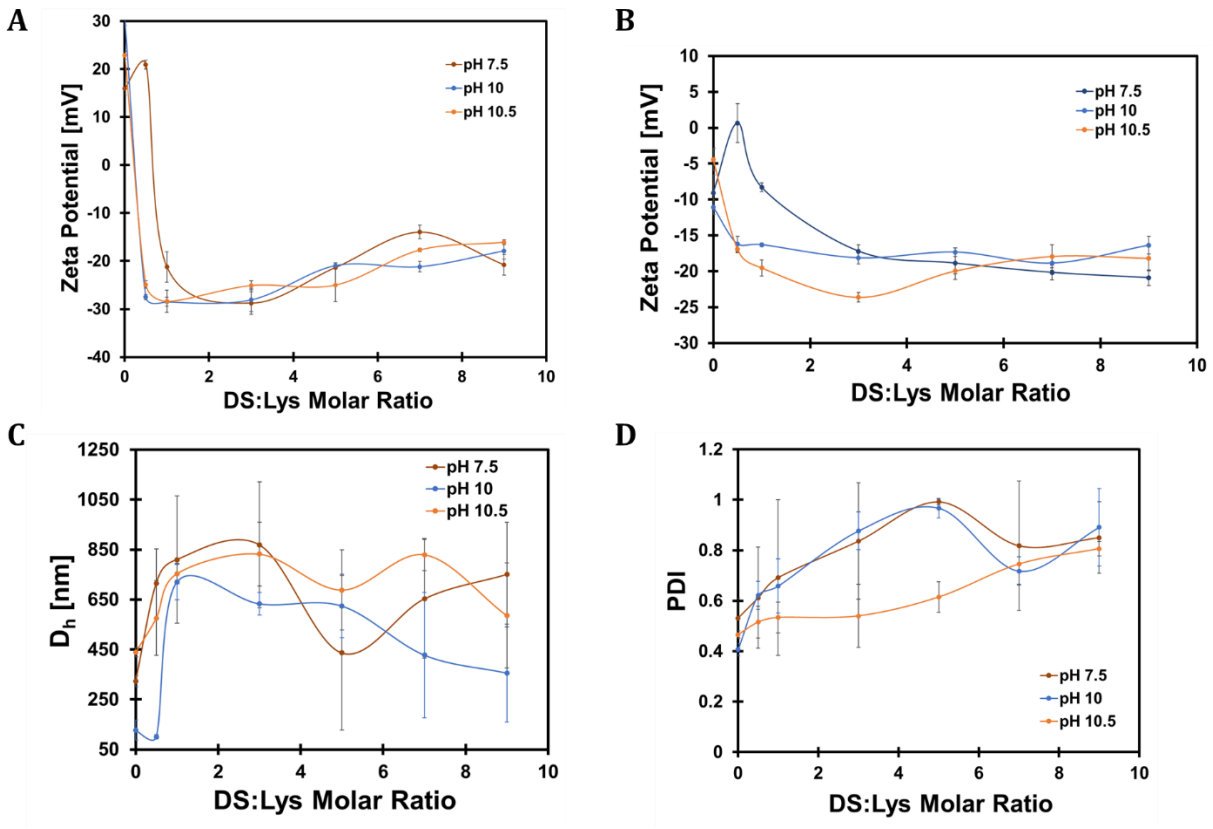

Figure 5-5: Lysozyme-DS coacervate characteristics, measured after the 1 hour mixing period, across the tested pH and MRs are shown, where **(A)** Zeta potential of coacervates in the undiluted mixture, **(B)** Zeta potential of coacervates diluted in 10 mM NaCl (1:1000 dilution), **(C)** Hydrodynamic Diameter (D_h) and **(D)** polydispersity of coacervates were measured.

ZP measurements of the concentrated (**Figure 5-5A**) and diluted (**Figure 5-5B**) lysozyme only solutions shows net positive ZP values across all tested pHs, providing evidence why insoluble complexes are being made; the presence of a net positive charge, even as the pH approaches lysozyme isoelectric point, is still strong enough to induce their formation. A sharp decrease in ZP is seen after DS:Lys MR 0.5 and stabilizes between -30 to -20 mV for concentrated and -15 to -25 mV for diluted solutions, as MR increases. The near neutral ZP value for complexes at pH 7.5, MR 0.5-1 may be the most optimal for further optimization, given that excess DS is likely to adsorbed to the complex surface. Coacervates' D_h (**Figure 5-5C**) and PDI (**Figure 5-5D**) were large across various pH and MRs, when compared to cat-DS complexes. PDI increased from 0.4 to 0.9 as pH increased, with a maximum observed at ~1 for pH 7.5 and 10. Average coacervate D_h increased to 650-850 nm between MR 0.5-3, where a slight size decrease is seen as MR increases. Overall, these data show that number of phases

that can observed and characterized, along with the colloidal behavior of coacervates, is dependent on the enzyme molecular weight and strength of the favorable electrostatic interaction between the protein and polyelectrolyte. Future investigations may reveal that if a protein is small (10-20 kDa), insoluble complex formation with DS will still occur, even if it has a weak net positive surface charge.

5.3 Conclusions

A dual computational and experimental approach was used in this work to examine the molecular level driving forces behind HIP and macroscopic phase behavior of water insoluble complexes. Atomistic MD was able to identify dextran sulphate as the most ideal to use ion-pairing agent for therapeutic proteins due to the high stability of BSA structure and its respective domains, when compared to BSA in water alone. Positively charged surface residues were primarily responsible for DS, whereas polar, hydrophobic and positively charged residues composed a major fraction of high residency residues for BSA/SDS and BSA/TA systems. From these simulations, future work should investigate how increasing the hydrophobicity of a polyanionic polymer like DS would influence a protein's structure, using both classical MD and circular dichroism measurements of coacervates. State point diagram construction of catalase-DS complexes showed that pH 3 was a boundary formation pH, whereby pH lower than 3 resulted in the formation of insoluble complexes while pH greater than 3 resulted in soluble complex formation; moreover, DS seems to promote solubilization of catalase for pH 3.5-5.4 since catalase only solutions were turbid, with some pHs showing precipitation. Zeta potential measurements showed that higher DS:Cat MRs were needed to neutralize the enzyme's positive charge and saturate its surface, at acidic pHs. As for lysozyme-DS complexes, insoluble complex formation was observed at all tested pHs of 7.5, 10 and 10.5, with pH 7.5 having precipitate volumes greater than pH 10; a wide variation in size (400-850 nm) and PDI (0.4-0.9) across all pH and MRs suggested that aggregation forces were dominant for this enzyme-DS pair. With more studies performing state point diagram construction, machine learning can be done to predict which physiochemical characteristics of the protein surface and polyanionic polymer results in insoluble complex formation.

6. Chapter 6: Significance and Perspective

Much of the foundational literature on drug delivery design and development has required massive effort and series of failures from researchers in order to arrive to today, where we have vaccines and other therapeutics avenues employing nanocarriers. Prior to increased access to advanced computational tools and resources, mechanistic understanding of the molecular level driving forces, necessary to control nanoparticle formation, was ascertained primarily from empirical models fit to experimental data. The difficulty of probing nanoscale phenomena with experimental characterization techniques has also been a major limitation; therefore, an iterative, trial-and-error approach has been the standard for the optimization of drug-loaded nanoparticles. Computational methods, like atomistic MD, offer a solution for investigating key molecular level interactions and can be used in tandem with experimental tools to develop a clear picture of the complex interplay of thermodynamics and kinetics that take place during the nanoparticle formulation process and allow for the bottom-up design of nanoparticles. A small number of research groups, like the Pfaendtner and Nance research groups, have attempted or are currently performing rational design of protein-loaded polymeric nanoparticles for therapeutic delivery, where much of the work is still in its early stages of development.

This graduate work lays the computational and experimental basis that will allow future researchers to answer more complex questions about nanoscale behavior between proteins and polymers, in addition to designing novel polymeric nanoparticles. PLGA-PEG diblock copolymer has the been the primary focus, since the polymer is biocompatible, and mechanism of particle formation is well-known[8]. Across well-known nanoparticle formulation methodologies using PLGA-PEG, organic solvents like acetone or DMSO are employed to promote polymer dissolution, but little is known the effects of copolymerization on chain structure and dynamics in solution. Chapter 3 fills this knowledge gap by using classical MD to provide insights into copolymer self-similarity and its associated hydrodynamical behavior that are applicable to experimentally relevant molecular weights; moreover, extraction of Kuhn and persistence lengths corroborated experimental measurements of PLGA-PEG nanoparticle structure from Kaldybekov et al.[19], showing the important role computational tools play in providing molecular level insights. Differences in structure and dynamics observed in this chapter showed the strong impact of polymer-solvent interactions on equilibrium conformations in solution, leading to an

important question that is explored in Chapter 4: What is the extent of protein/polymer interactions when oligomeric chains are highly extended versus collapsed?

The work done in Chapter 4 is the first time an investigation into the role of polymer conformation on protein-polymer interactions has been performed, either experimentally or computationally. As a result, we showed that PLGA-PEG oligomers undergo highly favorable protein binding when they are fully expanded, when compared to their collapse state. In addition, PEG's preference to be in water medium versus interacting with ID2S indicated how strong polymer-solvent interactions can lead to more transient, reversible polymer binding on the protein surface. Furthermore, the weak attraction of PEG with ID2S was also observed experimentally via neutron scattering by Abbott et al.[100], providing more reason why both rational design of protein-loaded PLGA-PEG nanoparticles will require both computational and experimental tools. These conclusions provided insights into the types of protein-copolymer interactions that might be occurring during protein release and resulted in an experimental recommendation, that being measurements of the PLGA-PEG first relaxation time in different solutions of interest, at various molecular weights, be done. The methods employed to extract residue clusters containing high residency amino acids (Chapter 2) are unique to the Pfaendtner research group and can be employed for future computational studies examining protein-polymer interactions from atomistic MD trajectories.

Despite the power of atomistic MD in being to investigate molecular level interactions, experimental measurements of the PLGA-PEG nanoparticle's ability to protect encapsulated catalase from proteolytic enzymes showed rapid deactivation in less than 2 hours. Changes in the choice of organic solvent to one that promoted PLGA-PEG expansion still resulted in poor nanoparticle protection. This observation, along with the insights from Chapter 4, strongly indicated that the protein surface needs to be altered in such a way that protein-polymer interactions persist through PLGA partitioning during nanoparticle formation and growth, resulting in core-shell nanoparticles with high encapsulation efficiencies; thus Chapter 5 used polyelectrolyte coacervation to alter the protein surface while preserving activity, in order to drive encapsulation and subsequently increase protection of the loaded therapeutic enzyme. Classical MD was used to screen which ion pairing agents did not result in protein destabilization and was the first study to examine the electrostatic complexation process between BSA and different IP agents with atomistic resolution. Work by Lan et al.[74,106] on pea protein isolates was then applied here to explore a wide range of pH and MR pair combinations and identify the range of pHs that result in water-insoluble polyelectrolyte coacervates

containing highly active therapeutic enzymes. This method of phase exploration, for the purpose of therapeutic protein drug delivery, has not yet been done in literature and work in Chapter 5 provides a critical pathway for future optimization of coacervates that allow for PLGA-PEG nanoparticles to be loaded with 2 or more proteins.

In conclusion, rational design of enzyme-loaded PLGA-PEG nanoparticles remains elusive but considerable efforts in both drug delivery and computational fields in recent years are making great progress in achieving this goal. This dissertation is simply the beginning of what is yet to be discovered and engineered, in order to solve society's most pressing medical problems. Future investigations should: (1) examine the stability and solubility of enzyme/dextran sulfate coacervates in organic solvents with different dielectric constants used for PLGA-PEG nanoparticle formulations via experimental techniques like UV-Vis spectrophotometry and High Performance Liquid chromatography, (2) Use atomistic MD and machine learning to uncover how polyelectrolyte coacervation changes the enzyme surface to promote favorable drug-polymer partitioning and employ molecular-level descriptors of the protein/polyelectrolyte complex surface to predict drug release or encapsulation efficiencies and lastly, (3) examine whether polyelectrolyte coacervates between dextran sulphate and two different enzymes can be co-loaded together into PLGA-PEG nanoparticles, while preserving their activity and desirable nanoparticle characteristics. Collaborations with polymer chemists could make way for chemical modification of dextran sulphate or another polyelectrolyte polymer as a way to tune the physiochemical nature of protein surface, hence tuning the protein release profile from PLGA-PEG nanoparticles. Ultimately, this work moves the drug delivery field a step closer to highly precise treatments that can cure the most complex of medical ailments.

A. Appendix A: Chapter 2

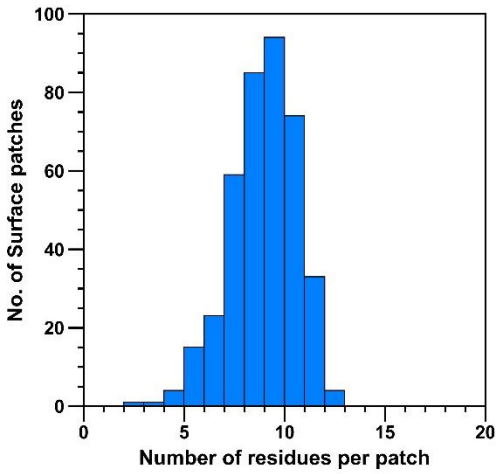

Figure A-1: Number of residues per surface patch when the angle cutoff for adjacent residue inclusion within a surface patch was 125^{o}

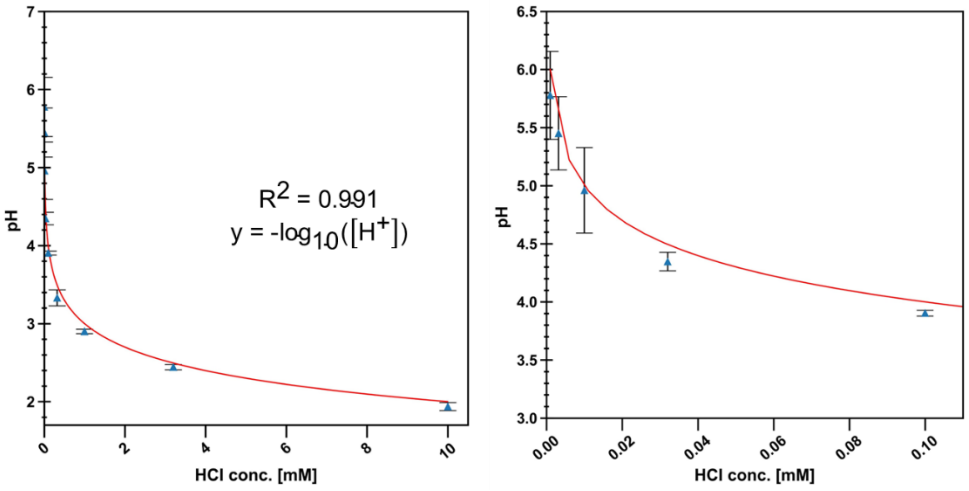

Figure A-2: pH as a function of HCl concentration is plotted above. Blue dots are measured pH values and the red line is the fit of pH = -$\log_{10}([H^+])$ to the experimental data points, showing that this equation can be used for preparing acidic solutions at a specified pH. The same procedure was done for basic solution, where pH = 14 – pOH and pOH = -$\log_{10}([OH^-])$

B. Appendix B: Chapter 3

Chapter 3**:** Exploring Structure and Dynamics of the PLGA-PEG Copolymer and its Homopolymer Constituents in Various Solvents Using All-Atom Molecular Dynamics

Authors: Chris W. Nyambura[1], Janani Sampath[2], Elizabeth Nance[1], *Jim Pfaendtner[1]

[1]Dept. of Chemical Engineering, University of Washington, Seattle 98195

[2]Dept. of Chemical Engineering, University of Florida, Gainesville 32611

*Corresponding author: jpfaendt@uw.edu

Table B-1: Partial atomic charges for a PEG trimer, resulting from a RESP calculation using Hartee Fork method with a 6-31G* basis set. Charges from quantum calculation were scaled to ensure oligomer charge neutrality, regardless of monomer length. sPEG is the starting monomer, PEG is the primary monomer and tPEG is the terminal monomer

sPEG		
Atom label	Amber atom-type	Scaled RESP charges
H1	HO	0.4111
O1	OH	-0.7889
C1	CT	0.2543
H2	H1	-0.0503
H3	H1	-0.0503
C2	CT	0.3114
H4	H1	-0.0436
H5	H1	-0.0436

PEG		
Atom label	Amber atom-type	Scaled RESP charges
O2	OS	-0.6070
C3	CT	0.3371
H6	H1	-0.0168
H7	H1	-0.0168
C2	CT	0.3371
H8	H1	-0.0168
H9	H1	-0.0168

tPEG		
Atom label	Amber atom-type	Scaled RESP charges
O3	OS	-0.5764
C5	CT	0.3834
H10	H1	0.0284
H11	H1	0.0284
C6	CT	0.3264
H12	H1	0.0217
H13	H1	0.0217
O4	OH	-0.7168
H14	HO	0.4832

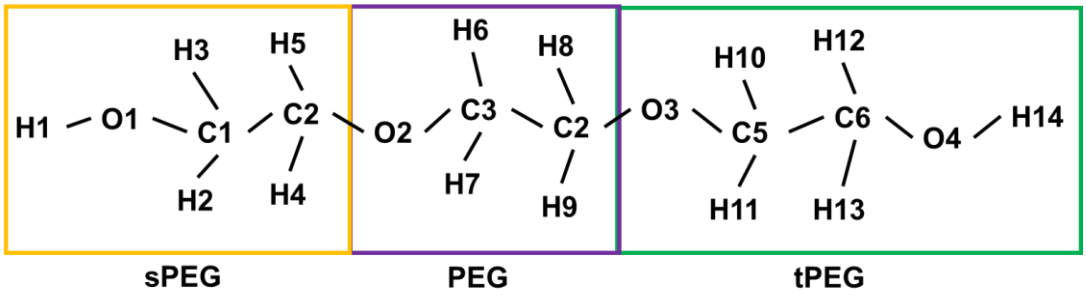

Figure B-1: Representative structure of PEG trimer, with atom labels

Table B-2: Partial atomic charges for a PLGA trimer, resulting from a RESP calculation using Hartee Fork method with a 6-31G* basis set. Charges from quantum calculation were scaled to ensure oligomer charge neutrality, regardless of monomer length. sPLG is the starting monomer, PLG is the primary monomer and tPLG is the terminal monomer. Representative structure with atom labels can be seen below the tables.

sPLG		
Atom label	Amber atom-type	Scaled RESP charges
O9	OS	-0.5052
C10	C	0.8992
O10	O	-0.5776
C11	CT	0.3129
O11	OS	-0.6526
C12	C	0.9528
O12	O	-0.5493
C13	CT	0.0926
O13	OH	-0.7006
H2	HO	0.4869
C14	CT	-0.2516
H12	H1	0.0430
H13	H1	0.0705
H14	H1	0.0705
H15	HC	0.1029
H16	HC	0.1029
H17	HC	0.1029

PLG		
Atom label	Amber atom-type	Scaled RESP charges
O5	OS	-0.5736
C6	C	0.8565
O6	O	-0.5962
C7	CT	0.3802
O7	OS	-0.6199
C8	C	0.9263
O8	O	-0.5910
C5	CT	-0.0671
C15	CT	-0.2957
H9	H1	0.0272
H10	H1	0.1284
H11	H1	0.1284
H18	HC	0.0988
H19	HC	0.0988
H20	HC	0.0988

tPLG		
Atom label	Amber atom-type	Scaled RESP charges
O1	OH	-0.7153
H1	HO	0.4766
C1	C	0.6278
O2	O	-0.5826
C2	CT	0.5797
C3	CT	-0.4571
O3	OS	-0.6980
C4	C	0.8779
O4	O	-0.6065
C5	CT	0.0368
H3	H1	-0.0300
H4	HC	0.1055
H5	HC	0.1055
H6	HC	0.1055
H7	H1	0.0870
H8	H1	0.0870

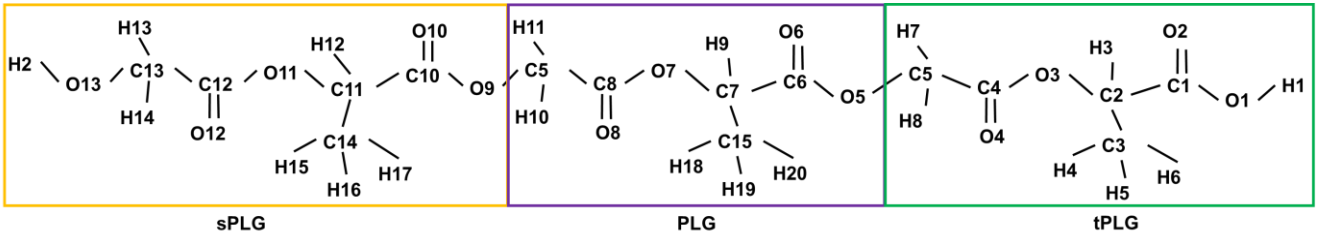

Figure B-2: Representative structure of PLGA trimer with atom labels

Table B-3: Partial atomic charges for a PLGA-PEG tetramer (2 LGA: 2 EG), resulting from a RESP calculation using Hartee Fork method, with a 6-31G* basis set. Charges from quantum calculation were scaled to ensure oligomer charge neutrality, regardless of monomer length. sbPLG is the starting LGA monomer, bPLG is the center LGA monomer, bPEG is the center PEG and sbPEG is the terminal PEG monomer. For the copolymer charges, bPEG and bPLG were used for linker, and homopolymer charges, PEG and PLG, were used for center residues in each respective block.

bPLG		
Atom label	Amber atom-type	Scaled RESP charges
O5	OS	-0.4917
C6	C	0.7982
O6	O	-0.5806
C7	CT	0.3795
O7	OS	-0.6530
C8	C	0.9138
O8	O	-0.5684
C5	CT	-0.0534
C15	CT	-0.2850
H9	H1	0.0095
H10	H1	0.1317
H11	H1	0.1317
H18	HC	0.0893
H19	HC	0.0893
H20	HC	0.0893

sbPEG		
Atom label	Amber atom-type	Scaled RESP charges
H1	HO	0.4102
O1	OH	-0.7037
C1	CT	0.2845
H2	H1	-0.0477
H3	H1	-0.0477
C2	CT	0.0785
H4	H1	0.0129
H5	H1	0.0129

sbPLG		
Atom label	Amber atom-type	Scaled RESP charges
O9	OS	-0.5341
C10	C	0.8968
O10	O	-0.5759
C11	CT	0.3360
O11	OS	-0.6572
C12	C	0.9484
O12	O	-0.5496
C13	CT	0.1020
O13	OH	-0.6980
H2	HO	0.4864
C14	CT	-0.2277
H12	H1	0.0413
H13	H1	0.0706
H14	H1	0.0706
H15	HC	0.0968
H16	HC	0.0968
H17	HC	0.0968

bPEG		
Atom label	Amber atom-type	Scaled RESP charges
O2	OS	-0.3553
C3	CT	0.1024
H6	H1	0.0270
H7	H1	0.0270
C2	CT	0.0427
H8	H1	0.0781
H9	H1	0.0781

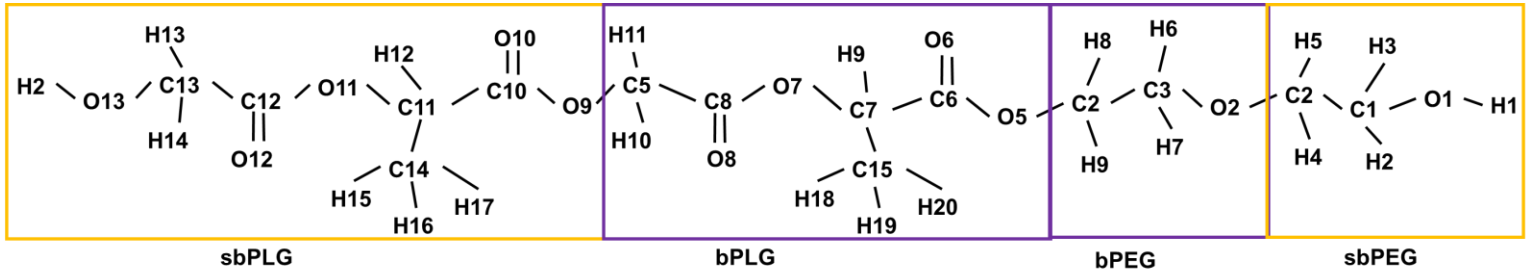

Figure B-3: Representative structure of PLGA-PEG tetramer with atom labels

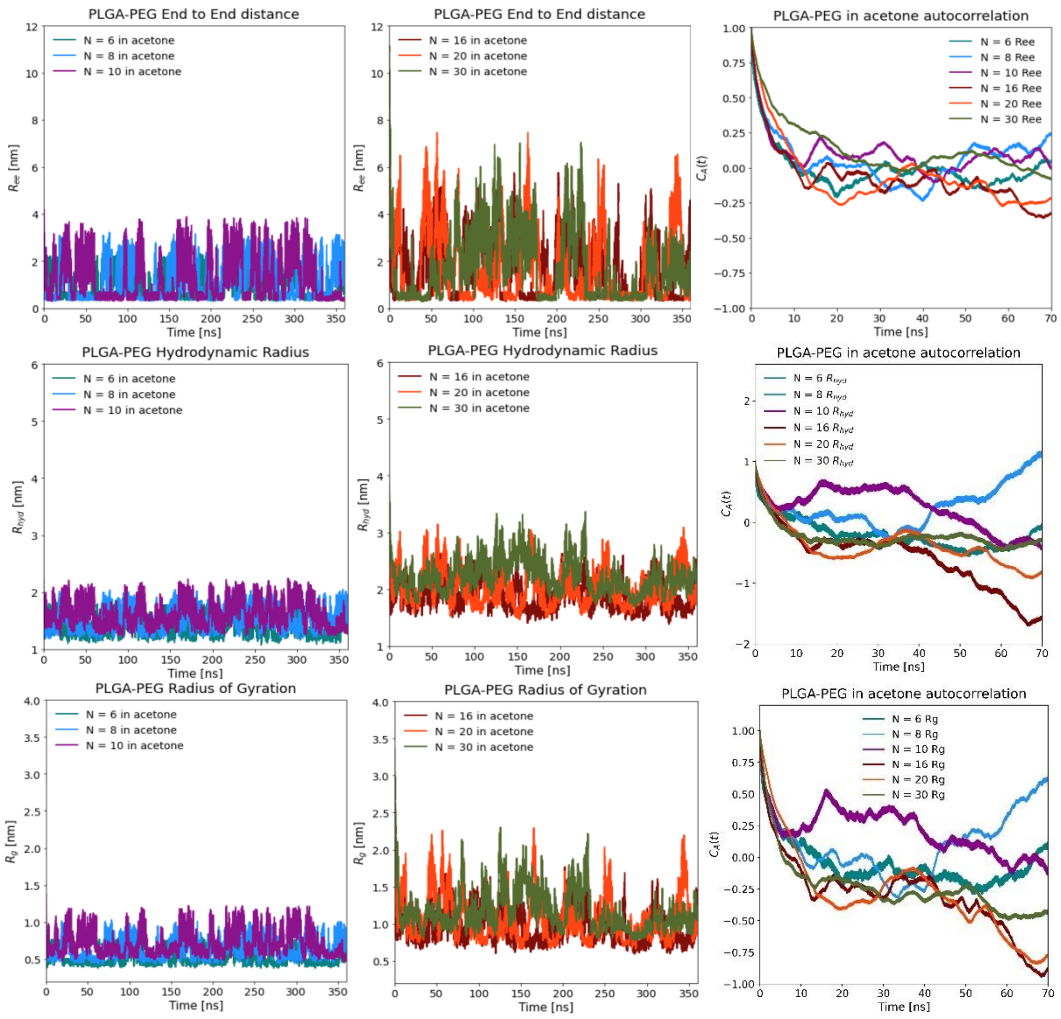

Figure B-4: Radius of gyration, hydrodynamic Radius and End-to-end distance time series, alongside their respective autocorrelations for 50:50 LGA:EG PLGA-PEG/acetone simulations. Autocorrelations were used to determine length of trajectory blocks used to generate averages and standard deviations for R_g and N scaling, $\langle R_{ee}^2 \rangle / \langle R_g^2 \rangle$ and $\langle R_g \rangle / \langle R_{hyd} \rangle$ vs L_c plots. A running average procedure was used to reduce noise in the autocorrelation data. Similar procedure was performed for the homopolymer/acetone systems

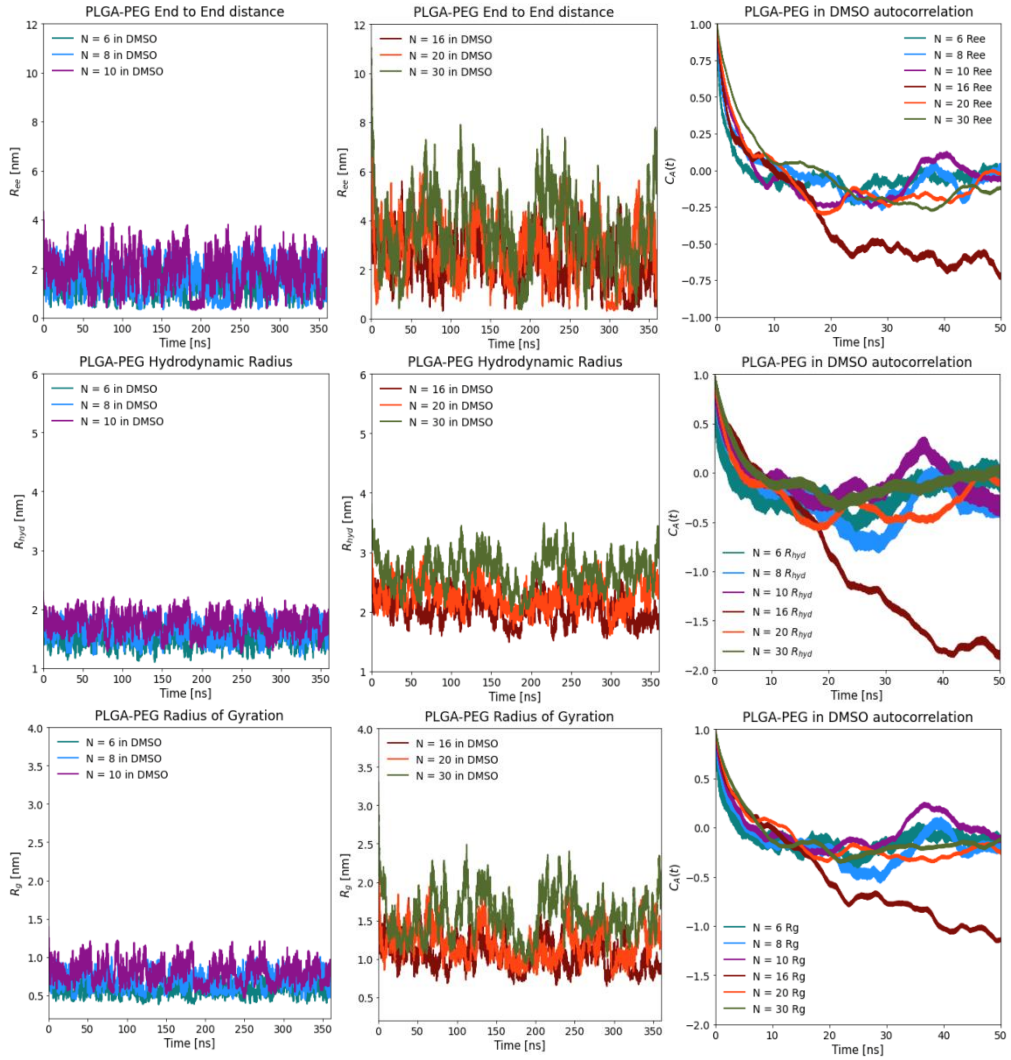

Figure B-5: Radius of gyration, hydrodynamic Radius and End-to-end distance time series, alongside their respective autocorrelations for 50:50 LGA:EG PLGA-PEG/DMSO simulations. Autocorrelations were used to determine length of trajectory blocks used to generate averages and standard deviations for $\mathrm{R_g}$ and N scaling, $\langle \mathrm{R_{ee}^2} \rangle / \langle \mathrm{R_g^2} \rangle$ and $\langle \mathrm{R_g} \rangle / \langle \mathrm{R_{hyd}} \rangle$ vs $\mathrm{L_c}$ plots. A running average procedure was used to reduce noise in the autocorrelation data. Similar procedure was performed for the homopolymer/DMSO systems

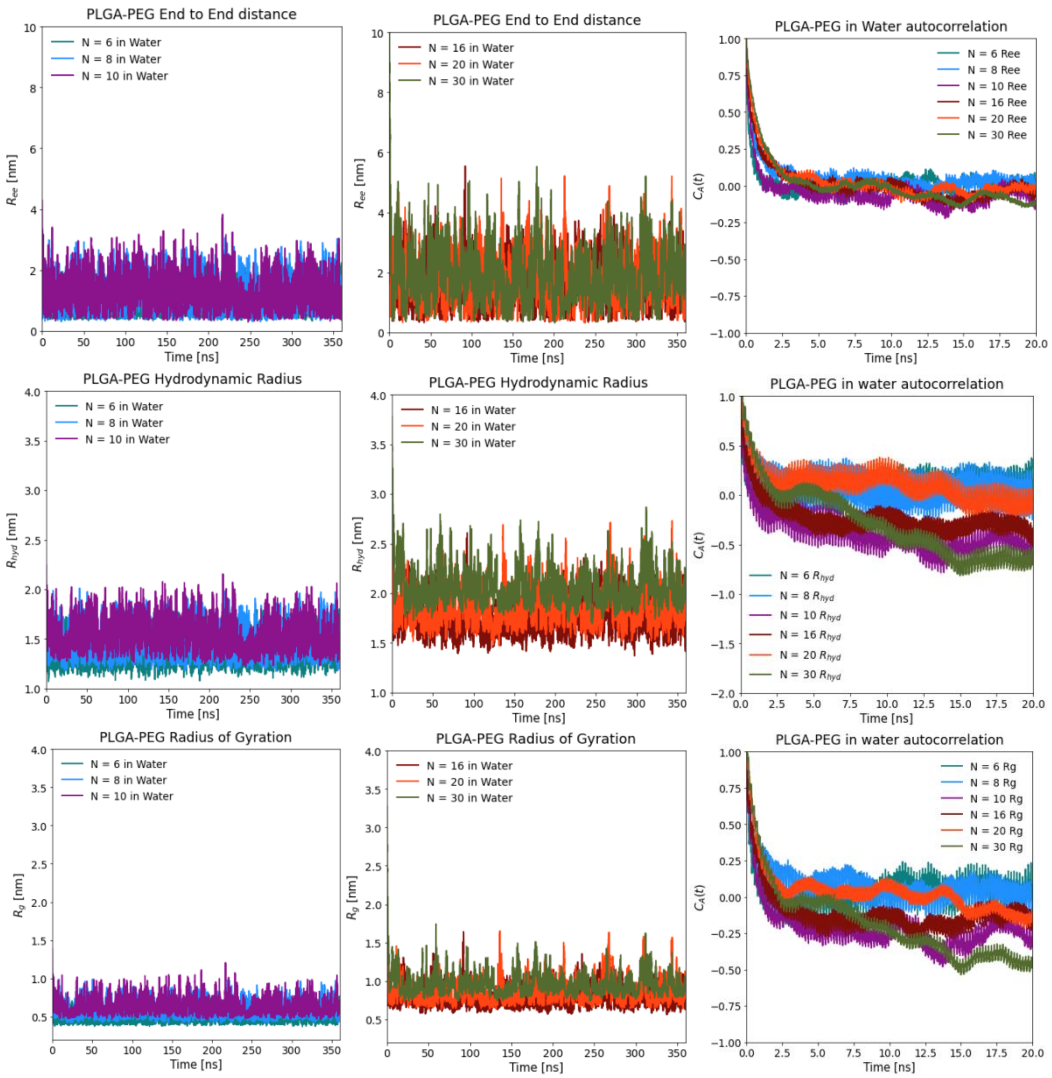

Figure B-6: Radius of gyration, hydrodynamic Radius and End-to-end distance time series, alongside their respective autocorrelations for 50:50 LGA:EG PLGA-PEG/water simulations. Autocorrelations were used to determine length of trajectory blocks used to generate averages and standard deviations for R_g and N scaling, $\langle R_{ee}^2 \rangle / \langle R_g^2 \rangle$ and $\langle R_g \rangle / \langle R_{hyd} \rangle$ vs L_c plots. A running average procedure was used to reduce noise in the autocorrelation data. Similar procedure was performed for the homopolymer/water systems

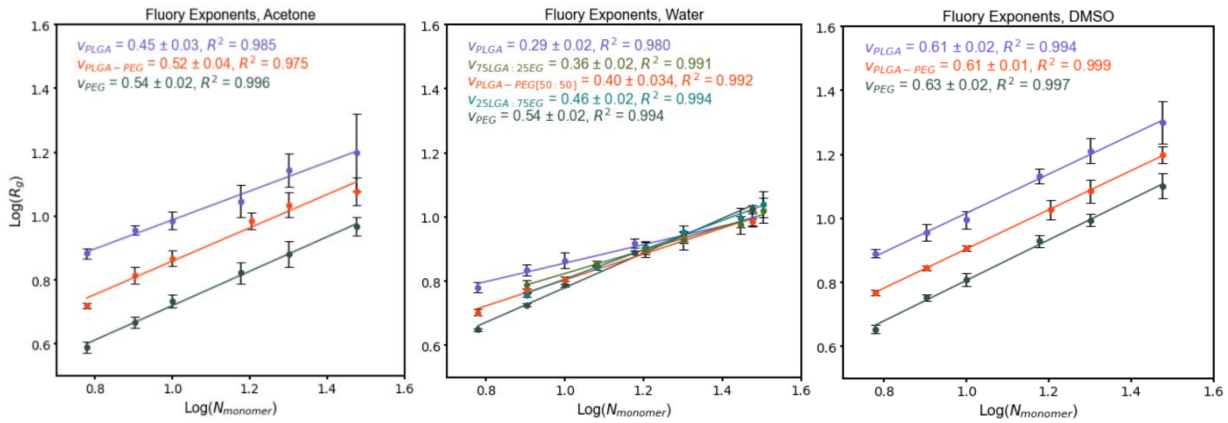

Figure B-7: Scaling Relationship between the radius of gyration (R_g) and monomer length for PLGA, PEG, 50:50 LGA:EG PLGA-PEG in acetone, water and DMSO, alongside 25:75 LGA:EG and 75:25 LGA:EG PLGA-PEG in water. Flory exponent (the slope) is reported, along with its standard error during linear regression. Std. deviation in R_g comes from block averaging of trajectories. Pearson coefficient is shown next to slope to indicate goodness of fit.

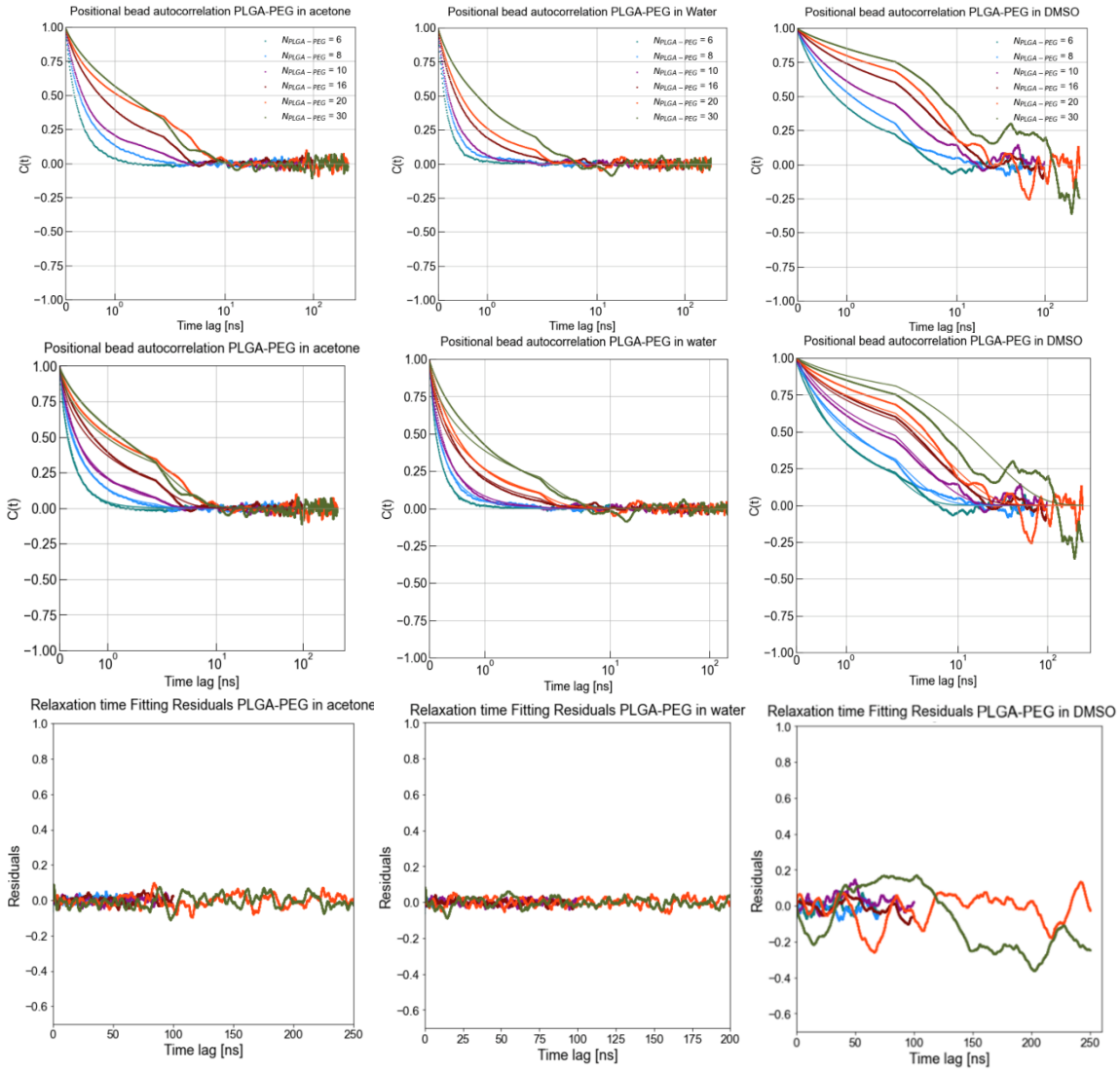

Figure B-8: Positional bead autocorrelation extracted from simulation in the 1st row, pos. bead autocorrelation with their nonlinear fit in the 2nd row, and the fit residuals on the 3rd row for 50:50 LGA:EG PLGA/solvent simulations. Similar behavior was observed for homopolymer systems.

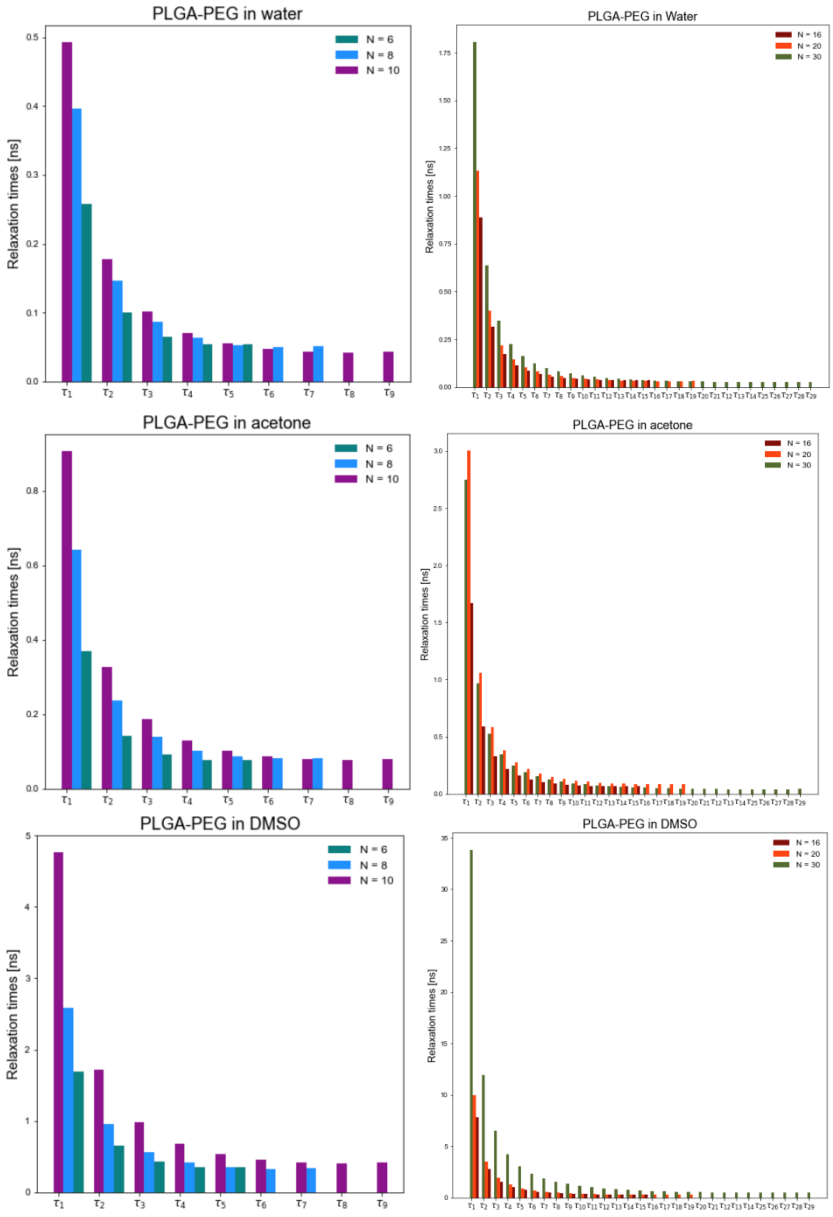

Figure B-9: Relaxation time spectrum for 50:50 LGA:EG PLGA-PEG/solvent simulations. Similar decay of relaxation times was observed for the homopolymer systems.

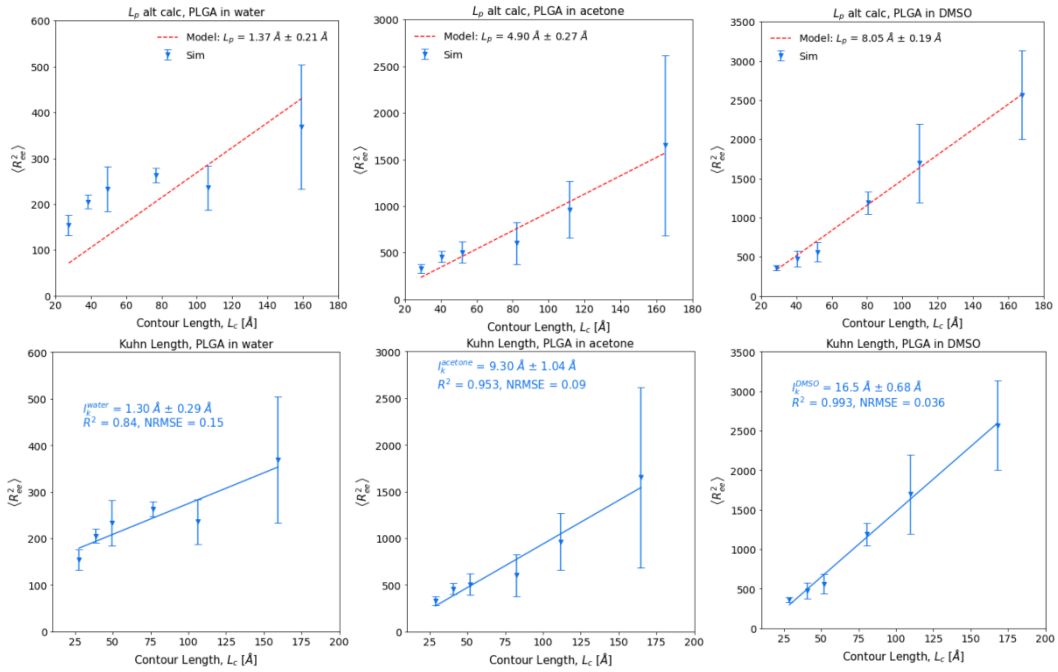

Figure B-10: Persistence length (1st row of plots) and Kuhn length (2nd row of plots) fits for PLGA/solvent simulations. Std. deviations and averages for $\langle R^2_{ee} \rangle$ values came from n = 5 block samples from each simulation trajectory. Pearson coefficient and the normalized root mean squared error are shown for the linear fits used to extract Kuhn lengths. Standard error of the mean is reported alongside persistence length values.

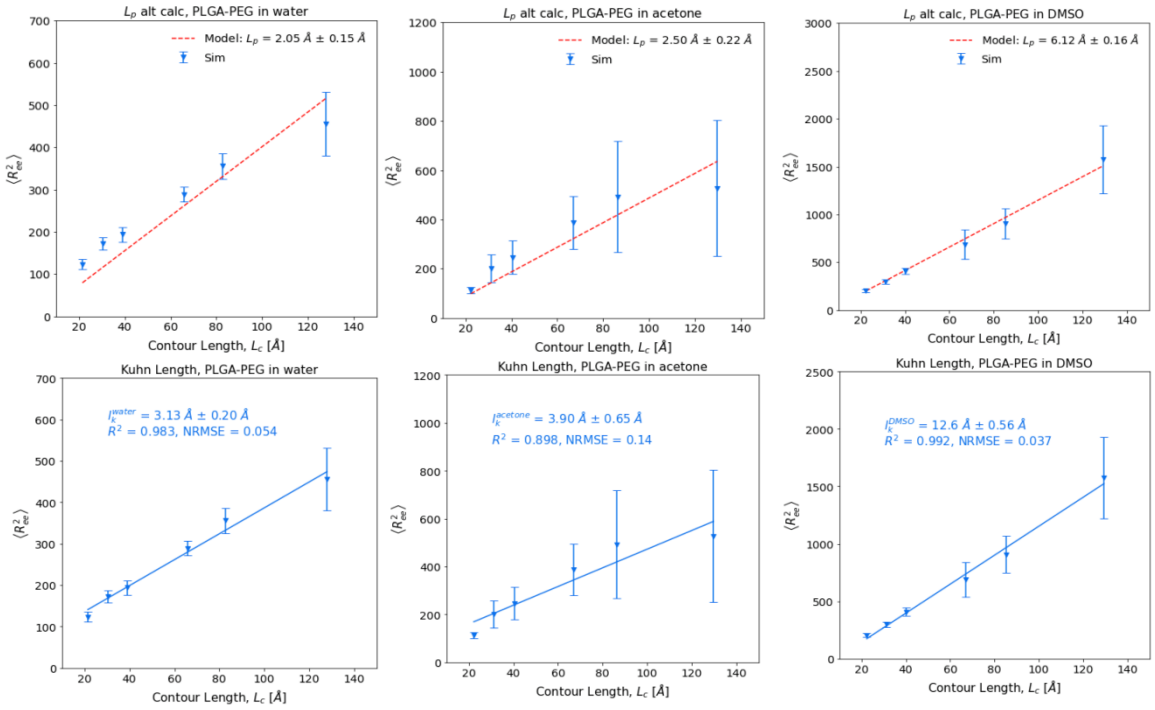

Figure B-11: Persistence length (1st row of plots) and Kuhn length (2nd row of plots) fits for 50:50 LGA:EG PLGA-PEG /solvent simulations. Std. deviations and averages for $\langle R_{ee}^2 \rangle$ values came from n = 5 block samples from each simulation trajectory. Pearson coefficient and the normalized root mean squared error are shown for the linear fits used to extract Kuhn lengths. Standard error of the mean is reported alongside persistence length values.

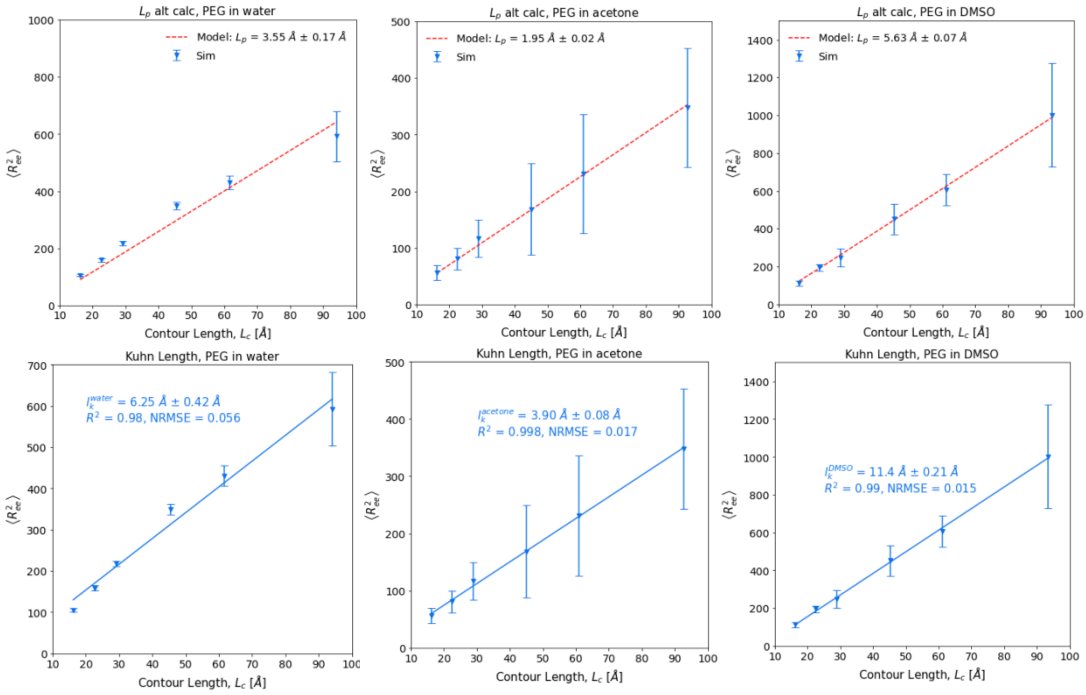

Figure B-12: Persistence length(1st row of plots) and Kuhn length(2nd row of plots) fits for PEG/solvent simulations. Std. deviations and averages for $\langle R_{ee}^2 \rangle$ values came from n = 5 block samples from each simulation trajectory. Pearson coefficient and the normalized root mean squared error are shown for the linear fits used to extract Kuhn lengths. Standard error of the mean is reported alongside persistence length values.

C. Appendix C: Chapter 4

Chapter 4**:** Examining the effect of polymer extension on protein/polymer interactions that occur during formulation of poly(lactic acid-co-glycolic acid)-polyethylene glycol nanoparticles. (In preparation for publication)

Authors: Chris W. Nyambura[1], Elizabeth Nance[1], *Jim Pfaendtner[1]

[1]Dept. of Chemical Engineering, University of Washington, Seattle 98195

*Corresponding author: jpfaendt@uw.edu

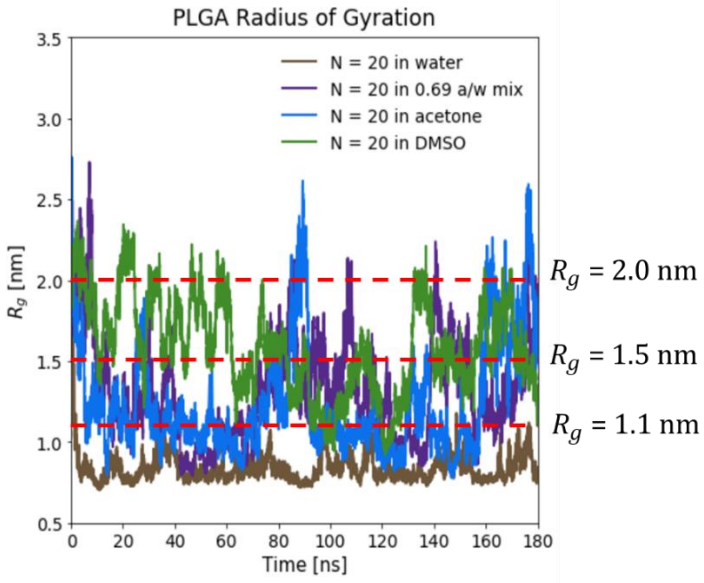

Figure C-1: Radius of gyration as a function of time for one PLGA oligomer in different solvent mediums, with a monomer length of 20. This information, ascertained during Chapter 3 analysis, was used to establish the values of R_g that were used as set points for high (R_g = 2.0 nm), medium (R_g = 1.5 nm) and low (R_g = 1.1 nm) levels of extension.

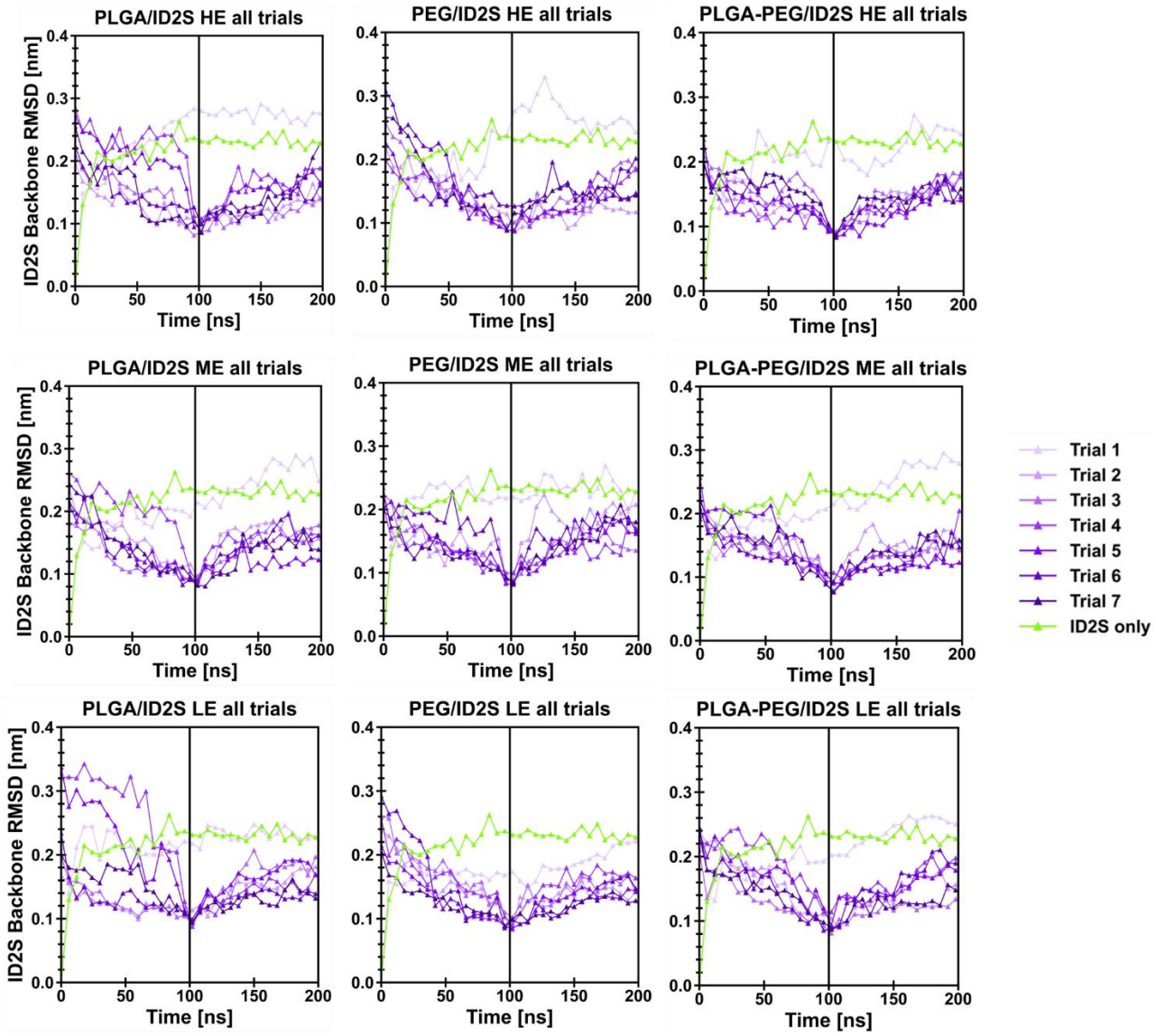

Figure C-2: Root mean squared deviation (RMSD) of ID2S from its crystal structure versus time is shown for all levels of extension for PLGA/ID2S (left column), PEG/ID2S (middle column) and PLGA-PEG/ID2S (right column). RMSD values are plotted every 6 ns. ID2S in water system served as a control and is shown in green.

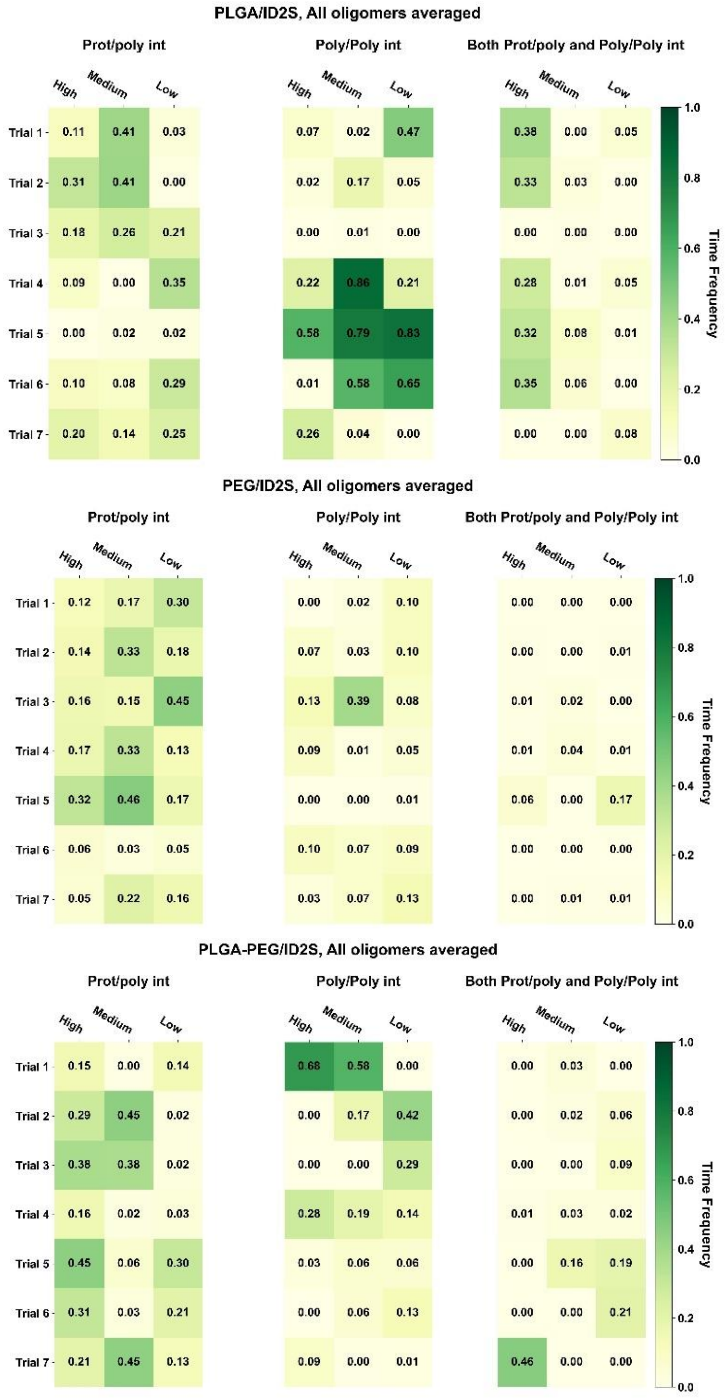

Figure C-3: Oligomer averaged time frequency of each interaction class, for each simulation trial across the different polymer types. Both restraint on and off phases are accounted for, in the calculation of TF. These mean values were used to calculate TF_{avg}^{poly}, seen in the Results section in Chapter 5

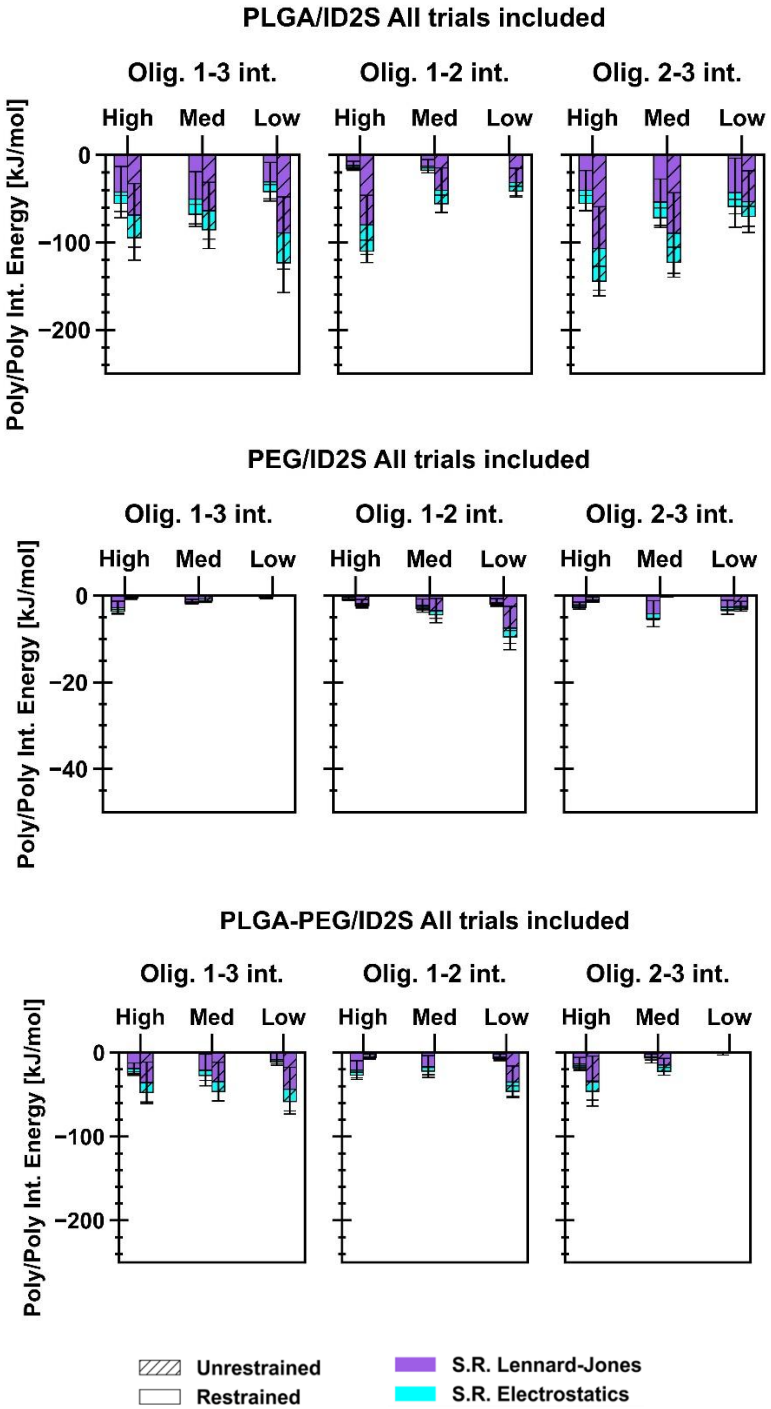

Figure C-4: Non-bonded interaction energies for polymer-polymer interactions across different levels of extension for PLGA/ID2S (top row), PEG/ID2S (middle row) and PLGA-PEG/ID2S (bottom row). Mean and SE values were calculated by averaging across trials for each level of extension and restraint phase. Since each simulation contains 3 oligomers, 1-3, 1-2 and 2-3 oligomer interactions were primarily extracted.

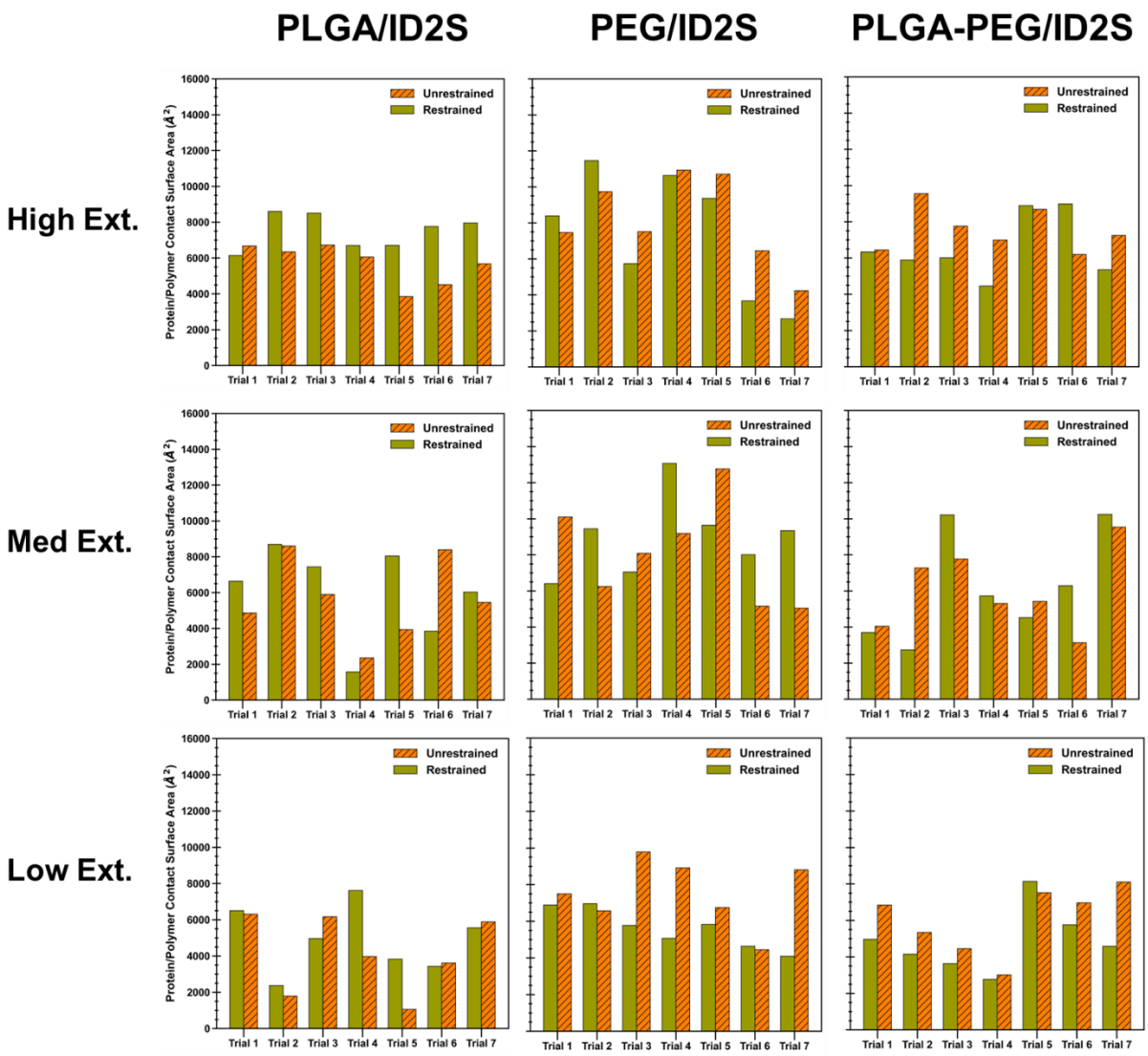

Figure C-5: Protein-polymer contact surface area from each simulation trial across the different levels of extension and restraint phases for PLGA/ID2S (left column), PEG/ID2S (middle column) and PLGA-PEG/ID2S (right column). For reference, ID2S total surface area is 20,438 $Å^2$.

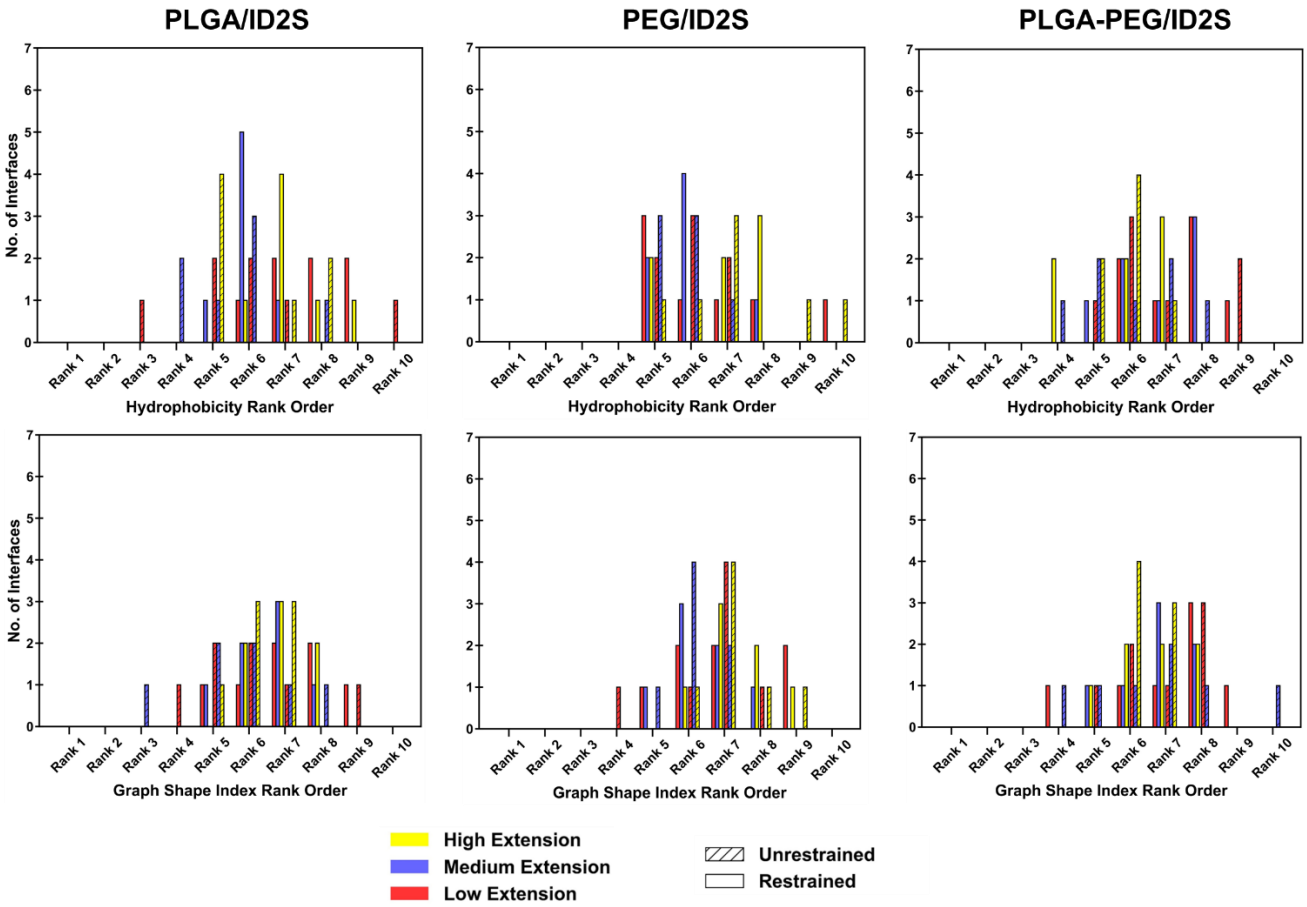

Figure C-6: Hydrophobicity and Graph Shape Index Rank distributions of the collapsed filtered protein-polymer interface for PLGA/ID2S (left column), PEG/ID2S (middle column) and PLGA-PEG/ID2S (right column) systems. Methodology detailing calculation of each parameter rank for each collapsed interface can be found in Chapter 2.

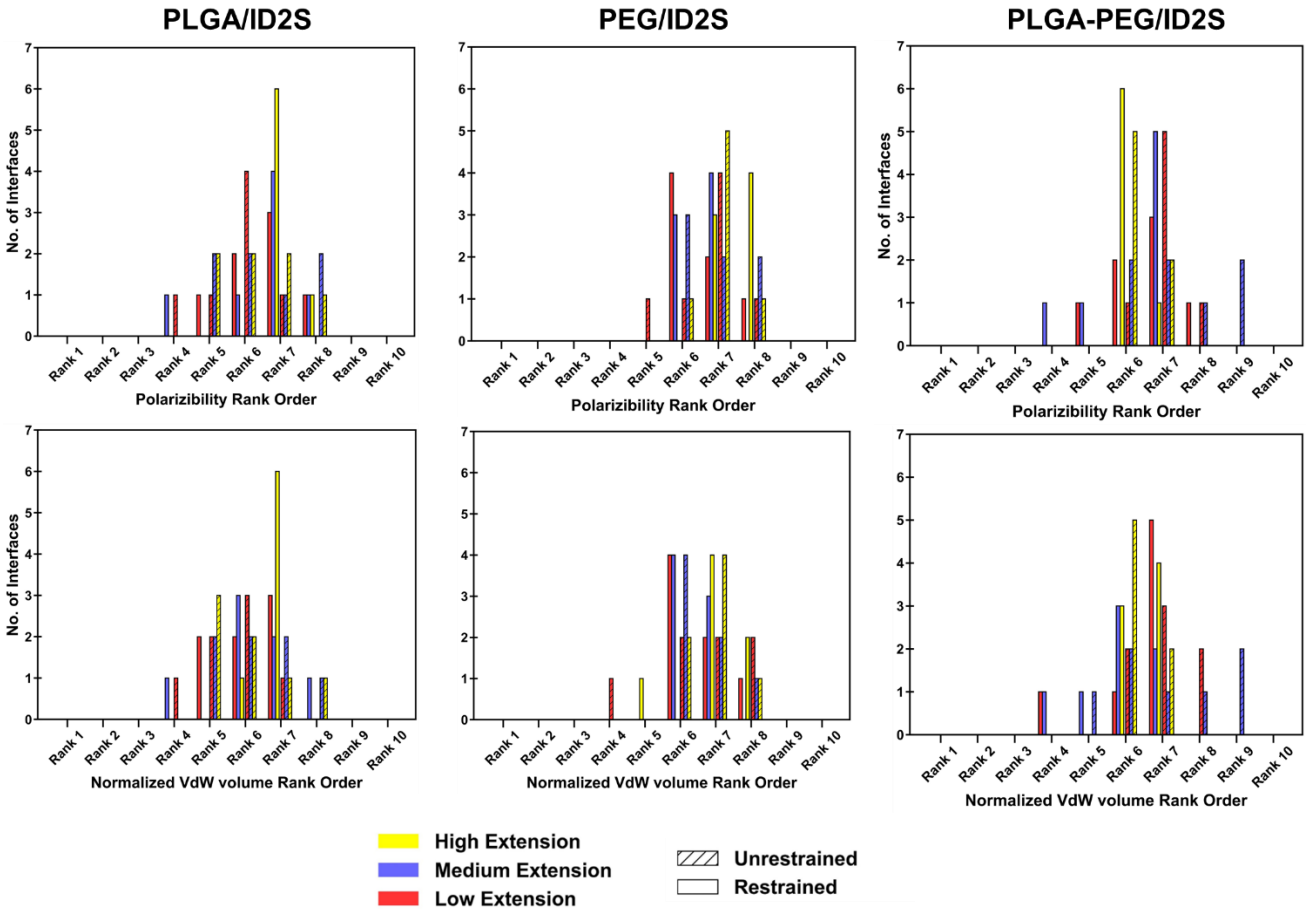

Figure C-7: Polarizability and Normalized Van der Waals Volume Rank distributions of the collapsed filtered protein-polymer interface for PLGA/ID2S (left column), PEG/ID2S (middle column) and PLGA-PEG/ID2S (right column) systems. Methodology detailing calculation of each parameter rank for each collapsed interface can be found in Chapter 2.

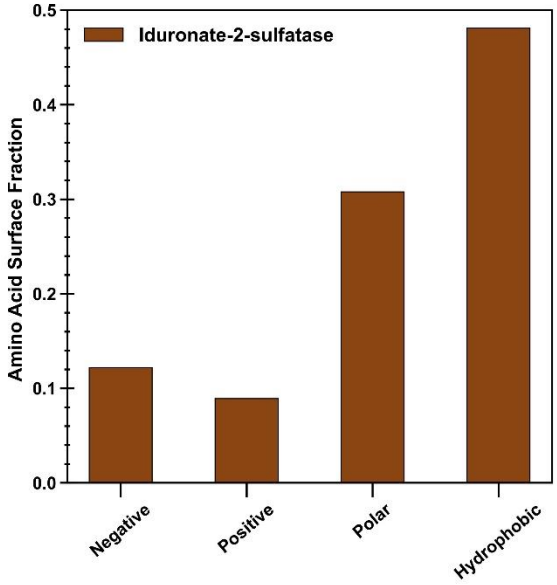

Figure C-8: Surface amino acid composition of iduronate-2-sulfatase at pH 7 in its native state. Surface residues consisted of 393 out of 516 total residues.

D. Appendix D: Chapter 5

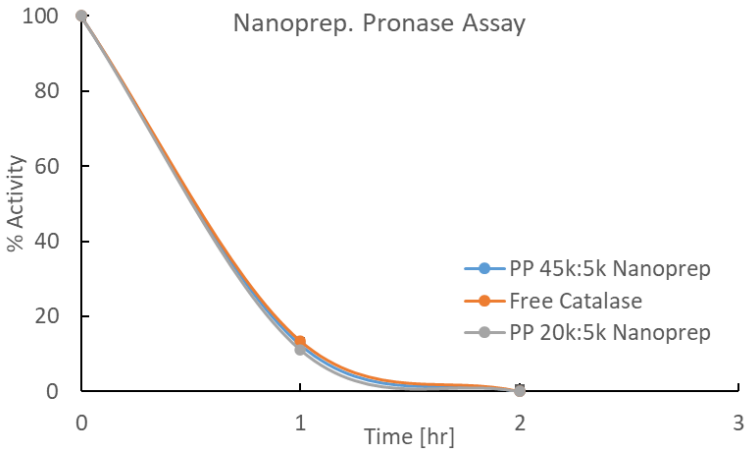

Figure D-1: Nanoparticle protection assay evaluated on catalase-loaded PLGA-PEG nanoparticles, formulated using the nanoprecipitation methodology and at different copolymer molecular weights described in this document's method section.

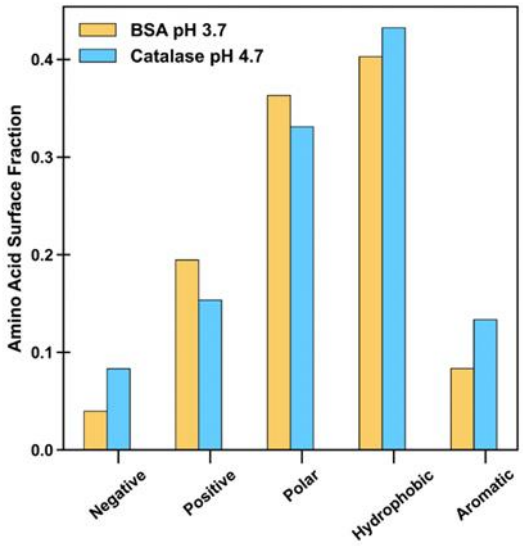

Figure D-2: Surface amino acid composition of BSA at pH 3.7 in its native state, compared to catalase at pH 4.7.

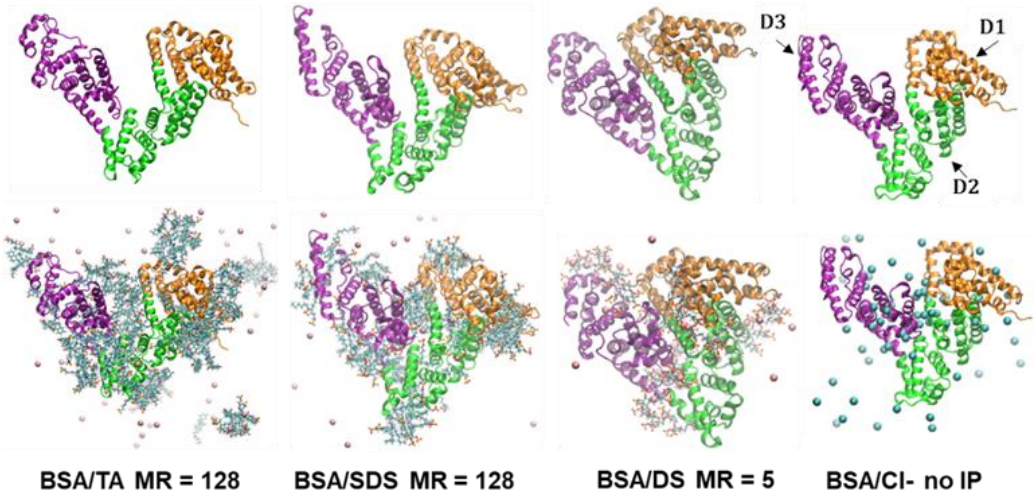

Figure D-3: End of trajectory snapshots (t = 200ns) of BSA without (top row) and with (bottom row) ion-pairing agents and neutralizing ions. Chlorine ions are represented by cyan particles and potassium ions are represented by brown particles. Domain I (D1) is in orange, Domain II (D2) is in green and Domain III (D3) is in purple

E. References

1. Wang, B., Hu, L. & Siahaan, T. J. *Drug Delivery: Principles and Applications*. (Wiley, 2016).

2. Leader, B., Baca, Q. J. & Golan, D. E. Protein therapeutics: a summary and pharmacological classification. *Nat. Rev. Drug Discov.* **7**, 21 (2008).

3. Bruno, B. J., Miller, G. D. & Lim, C. S. Basics and recent advances in peptide and protein drug delivery. *Ther. Deliv.* **4**, 1443–1467 (2013).

4. Bourganis, V., Karamanidou, T., Kammona, O. & Kiparissides, C. Polyelectrolyte complexes as prospective carriers for the oral delivery of protein therapeutics. *Eur. J. Pharm. Biopharm.* **111**, 44–60 (2017).

5. Mitragotri, S., Burke, P. A. & Langer, R. Overcoming the challenges in administering biopharmaceuticals: formulation and delivery strategies. *Nat. Rev. Drug Discov.* **13**, 655–672 (2014).

6. Kumari, A., Yadav, S. K. & Yadav, S. C. Biodegradable polymeric nanoparticles based drug delivery systems. *Colloids Surf B Biointerfaces* **75**, 1–18 (2010).

7. Gentile, P., Chiono, V., Carmagnola, I. & Hatton, P. V. An overview of poly(lactic-co-glycolic) acid (PLGA)-based biomaterials for bone tissue engineering. *Int J Mol Sci* **15**, 3640–3659 (2014).

8. Almoustafa, H. A., Alshawsh, M. A. & Chik, Z. Technical aspects of preparing PEG-PLGA nanoparticles as carrier for chemotherapeutic agents by nanoprecipitation method. *Int. J. Pharm.* **533**, (2017).

9. Ghosh Chaudhuri, R. & Paria, S. Core/Shell Nanoparticles: Classes, Properties, Synthesis Mechanisms, Characterization, and Applications. *Chem. Rev.* **112**, 2373–2433 (2012).

10. Anselmo, A. C. & Mitragotri, S. Nanoparticles in the clinic. *Bioeng Transl Med* **1**, 10–29 (2016).

11. Anselmo, A. C. & Mitragotri, S. Nanoparticles in the clinic: An update. *Bioeng. Transl. Med.* **4**, (2019).

12. Joseph, A. *et al.* Formulation and Efficacy of Catalase-Loaded Nanoparticles for the Treatment of Neonatal Hypoxic-Ischemic Encephalopathy. *Pharmaceutics* **13**, 1131 (2021).

13. Nautyal, U., . D. & Gupta, D. Oral Sustained Release Tablets: An Overview With A Special Emphasis On Matrix Tablet. *Int. J. Heal. Biol. Sci.* **3**, 6–13 (2020).

14. Xu, Y., Kim, C.-S., Saylor, D. M. & Koo, D. Polymer degradation and drug delivery in PLGA-based drug-polymer applications: A review of experiments and theories. *J. Biomed. Mater. Res. Part B Appl. Biomater.* **105**, 1692–1716 (2017).

15. Szlek, J., Paclawski, A., Lau, R., Jachowicz, R. & Mendyk, A. Heuristic modeling of macromolecule release from PLGA microspheres. *Int J Nanomedicine* **8**, 4601–4611 (2013).

16. Kang, E., Robinson, J., Park, K. & Cheng, J.-X. Paclitaxel distribution in poly(ethylene glycol)/poly(lactide-co-glycolic acid) blends and its release visualized by coherent anti-Stokes Raman scattering microscopy. *J. Control. Release* **122**, 261–268 (2007).

17. Vega, E. *et al.* PLGA nanospheres for the ocular delivery of flurbiprofen: Drug release and interactions. *J. Pharm. Sci.* **97**, 5306–5317 (2008).

18. Wilkosz, N. *et al.* Molecular Insight into Drug-Loading Capacity of PEG-PLGA Nanoparticles for Itraconazole. *J Phys Chem B* **122**, 7080–7090 (2018).

19. Kaldybekov, D. B., Filippov, S. K., Radulescu, A. & Khutoryanskiy, V. V. Maleimide-functionalised PLGA-PEG nanoparticles as mucoadhesive carriers for intravesical drug delivery. *Eur J Pharm Biopharm* **143**, 24–34 (2019).

20. Cho, D. H. & Hahm, J. Protein–Polymer Interaction Characteristics Unique to Nanoscale Interfaces: A Perspective on Recent Insights. *J. Phys. Chem. B* **125**, 6040–6057 (2021).

21. Taylor, P. A. & Jayaraman, A. Molecular Modeling and Simulations of Peptide–Polymer Conjugates. *Annu. Rev. Chem. Biomol. Eng.* **11**, 257–276 (2020).

22. Kurinomaru, T., Kuwada, K., Tomita, S., Kameda, T. & Shiraki, K. Noncovalent PEGylation through Protein–Polyelectrolyte Interaction: Kinetic Experiment and Molecular Dynamics Simulation. *J. Phys. Chem. B* **121**, 6785–6791 (2017).

23. Karanicolas, J. & Brooks, C. L. Improved Gō-like Models Demonstrate the Robustness of Protein Folding Mechanisms Towards Non-native Interactions. *J. Mol. Biol.* **334**, 309–325 (2003).

24. Hamed, E., Ma, D. & Keten, S. Multiple PEG Chains Attached onto the Surface of a Helix Bundle: Conformations and Implications. *ACS Biomater. Sci. Eng.* **1**, 79–84 (2015).

25. Allen, M. P., Allen, M. P., Tildesley, D. J., Tildesley, D. J. & ALLEN, T. *Computer Simulation of Liquids*. (Clarendon Press, 1989).

26. Hopkins, C. W., Le Grand, S., Walker, R. C. & Roitberg, A. E. Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning. *J Chem Theory Comput* **11**, 1864–1874 (2015).

27. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N ·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).

28. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**–**2**, (2015).

29. van der Spoel, D., van Maaren, P. J. & Caleman, C. GROMACS molecule & liquid database. *Bioinformatics* **28**, 752–753 (2012).

30. Abraham, M. J., van der Spoel, D., Lindahl, E. & team, G. development. GROMACS User Manual version 2016. (2018).

31. Lindahl, E., Abraham, M. J., Hess, B. & Van der Spoel, D. GROMACS 2020.5 Source code. *Zenodo* (2021) doi:10.5281/zenodo.3562495.

32. Aliev, A. E. *et al.* Motional timescale predictions by molecular dynamics simulations: case study using proline and hydroxyproline sidechain dynamics. *Proteins* **82**, 195–215 (2014).

33. Tian, C. *et al.* ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against

Quantum Mechanics Energy Surfaces in Solution. *J. Chem. Theory Comput.* **16**, 528–552 (2020).

34. Kirschner, K. N. *et al.* GLYCAM06: a generalizable biomolecular force field. Carbohydrates. *J Comput Chem* **29**, 622–655 (2008).

35. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).

36. Wang, J., Wang, W., Kollman, P. A. & Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model* **25**, 247–260 (2006).

37. Martinez-Rosell, G., Giorgino, T. & De Fabritiis, G. PlayMolecule ProteinPrepare: A Web Application for Protein Preparation for Molecular Dynamics Simulations. *J Chem Inf Model* **57**, 1511–1516 (2017).

38. Mark, P. & Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *J. Phys. Chem. A* **105**, 9954–9960 (2001).

39. Izadi, S., Anandakrishnan, R. & Onufriev, A. V. Building Water Models: A Different Approach. *J. Phys. Chem. Lett.* **5**, 3863–3871 (2014).

40. Izadi, S. & Onufriev, A. V. Accuracy limit of rigid 3-point water models. *J. Chem. Phys.* **145**, 074501 (2016).

41. Fischer, N. M., van Maaren, P. J., Ditz, J. C., Yildirim, A. & van der Spoel, D. Properties of Organic Liquids when Simulated with Long-Range Lennard-Jones Interactions. *J Chem Theory Comput* **11**, 2938–2944 (2015).

42. Caleman, C. *et al.* Force Field Benchmark of Organic Liquids: Density, Enthalpy of Vaporization, Heat Capacities, Surface Tension, Isothermal Compressibility, Volumetric Expansion Coefficient, and Dielectric Constant. *J Chem Theory Comput* **8**, 61–74 (2012).

43. Fox, T. & Kollman, P. A. Application of the RESP Methodology in the Parametrization of Organic Solvents. *J. Phys. Chem. B* **102**, 8070–8079 (1998).

44. Zhang, J., Tuguldur, B. & van der Spoel, D. Force Field Benchmark of Organic Liquids. 2. Gibbs Energy of Solvation. *J Chem Inf Model* **55**, 1192–1201 (2015).

45. Frisch, M. J. Gaussian 09. (2009).

46. Bayly, C. I., Cieplak, P., Cornell, W. & Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **97**, 10269–10280 (1993).

47. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J Chem Phys* **126**, 14101 (2007).

48. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).

49. Martinez, L., Andrade, R., Birgin, E. G. & Martinez, J. M. PACKMOL: a package for building initial configurations for molecular dynamics simulations. *J Comput Chem* **30**, 2157–2164 (2009).

50. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).

51. Michaud-Agrawal, N., Denning, E. J., Woolf, T. B. & Beckstein, O. MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J Comput Chem* **32**, 2319–2327 (2011).

52. Oelmeier, S. A., Dismer, F. & Hubbuch, J. Molecular dynamics simulations on aqueous two-phase systems - Single PEG-molecules in solution. *BMC Biophys.* **5**, 14 (2012).

53. Andrews, J. & Blaisten-Barojas, E. Exploring with Molecular Dynamics the Structural Fate of PLGA Oligomers in Various Solvents. *J Phys Chem B* **123**, 10233–10244 (2019).

54. Rubinstein, M. & Colby, R. H. *Polymer physics*. (Oxford University Press, 2003).

55. Newville, M., Stensitzki, T., Allen, D. B. & Ingargiola, A. LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python. (2014).

56. Del Giudice, F., Haward, S. J. & Shen, A. Q. Relaxation time of dilute polymer solutions: A microfluidic approach. *J. Rheol. (N. Y. N. Y).* **61**, 327–337 (2017).

57. Patel, S. S. & Takahashi, K. M. Polymer dynamics in dilute and semidilute solutions. *Macromolecules* **25**, 4382–4391 (1992).

58. Dinic, J., Zhang, Y., Jimenez, L. N. & Sharma, V. Extensional Relaxation Times of Dilute, Aqueous Polymer Solutions. *ACS Macro Lett.* **4**, 804–808 (2015).

59. Schlijper, A. G., Hoogerbrugge, P. J. & Manke, C. W. Computer simulation of dilute polymer solutions with the dissipative particle dynamics method. *J. Rheol. (N. Y. N. Y).* **39**, 567–579 (1995).

60. Kong, Y., Manke, C. W., Madden, W. G. & Schlijper, A. G. Effect of solvent quality on the conformation and relaxation of polymers via dissipative particle dynamics. *J. Chem. Phys.* **107**, 592–602 (1997).

61. Lindorff-Larsen, K. *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**, 1950–1958 (2010).

62. Barbosa, N. S. V, Zhang, Y., Lima, E. R. A., Tavares, F. W. & Maginn, E. J. Development of an AMBER-compatible transferable force field for poly(ethylene glycol) ethers (glymes). *J. Mol. Model.* **23**, (2017).

63. Bonomi, M. *et al.* PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Comput. Phys. Commun.* **180**, 1961–1972 (2009).

64. Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C. & Bussi, G. PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **185**, 604–613 (2014).

65. Tylki-Szymańska, A. Mucopolysaccharidosis type II, Hunter's syndrome. *Pediatr. Endocrinol. Rev.* **12 Suppl 1**, 107—113 (2014).

66. Abraham, M. J., van der Spoel, D., Lindahl, E., Hess, B. & team, G. development. GROMACS User Manual 2018. in (2018).

67. Python Software Foundation. Python Language Reference, version 3.6.

68. Jones, S. & Thornton, J. M. Analysis of protein-protein interaction sites using surface patches. *J Mol Biol* **272**, 121–132 (1997).

69. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.*

14, 33–38 (1996).

70. Chaplin, M. Water Structure and Science References. https://water.lsbu.ac.uk/water/h2o_molecule.html#g (2021).

71. FAUCHÈRE, J.-L., CHARTON, M., KIER, L. B., VERLOOP, A. & PLISKA, V. Amino acid side chain parameters for correlation studies in biology and pharmacology. *Int. J. Pept. Protein Res.* **32**, 269–278 (2009).

72. Sutay Kocabas, D., Bakir, U., Phillips, S. E. V., McPherson, M. J. & Ogel, Z. B. Purification, characterization, and identification of a novel bifunctional catalase-phenol oxidase from Scytalidium thermophilum. *Appl. Microbiol. Biotechnol.* **79**, 407–415 (2008).

73. Proctor, V. A., Cunningham, F. E. & Fung, D. Y. C. The chemistry of lysozyme and its use as a food preservative and a pharmaceutical. *C R C Crit. Rev. Food Sci. Nutr.* **26**, 359–395 (1988).

74. Lan, Y., Ohm, J.-B., Chen, B. & Rao, J. Phase behavior and complex coacervation of concentrated pea protein isolate-beet pectin solution. *Food Chem.* **307**, (2020).

75. Liu, S.-J. *et al.* Electrospun PLGA/collagen nanofibrous membrane as early-stage wound dressing. *J. Memb. Sci.* **355**, 53–59 (2010).

76. Webster, T. J. & Smith, T. A. Increased osteoblast function on PLGA composites containing nanophase titania. *J Biomed Mater Res A* **74**, 677–686 (2005).

77. Pinto Reis, C., Neufeld, R. J., Ribeiro, A. J. & Veiga, F. Nanoencapsulation I. Methods for preparation of drug-loaded polymeric nanoparticles. *Nanomedicine Nanotechnology, Biol. Med.* **2**, 8–21 (2006).

78. Chan, J. M. *et al.* PLGA–lecithin–PEG core–shell nanoparticles for controlled drug delivery. *Biomaterials* **30**, 1627–1634 (2009).

79. Tran, V.-T. *et al.* Protein-loaded PLGA–PEG–PLGA microspheres: A tool for cell therapy. *Eur. J. Pharm. Sci.* **45**, 128–137 (2012).

80. Fodor-Kardos, A., Kiss, Á. F., Monostory, K. & Feczkó, T. Sustained in vitro interferon-beta release and in vivo toxicity of PLGA and PEG-PLGA nanoparticles. *RSC Adv.* **10**, 15893–15900 (2020).

81. Wang, H. *et al.* Enhanced anti-tumor efficacy by co-delivery of doxorubicin and paclitaxel with amphiphilic methoxy PEG-PLGA copolymer nanoparticles. *Biomaterials* **32**, 8281–8290 (2011).

82. Le Coeur, C. *et al.* Conformation of the Poly(ethylene Glycol) Chains in DiPEGylated Hemoglobin Specifically Probed by SANS: Correlation with PEG Length and in Vivo Efficiency. *Langmuir* **31**, 8402–8410 (2015).

83. Khorshid, N. K. *et al.* Novel Structural Changes during Temperature-Induced Self-Assembling and Gelation of PLGA-PEG-PLGA Triblock Copolymer in Aqueous Solutions. *Macromol. Biosci.* **16**, 1838–1852 (2016).

84. Pai, S. S. *et al.* The Conformation of the Poly(ethylene glycol) Chain in Mono-PEGylated Lysozyme and Mono-PEGylated Human Growth Hormone. *Bioconjug. Chem.* **22**, 2317–2323 (2011).

85. Kienberger, F. *et al.* Static and Dynamical Properties of Single Poly(Ethylene Glycol)

Molecules Investigated by Force Spectroscopy. *Single Mol.* **1**, 123–128 (2000).

86. Yoo, H. S. & Park, T. G. Biodegradable polymeric micelles composed of doxorubicin conjugated PLGA–PEG block copolymer. *J. Control. Release* **70**, (2001).

87. Brauner, B., Schwarz, P., Wirth, M. & Gabor, F. Micro vs. nano: PLGA particles loaded with trimethoprim for instillative treatment of urinary tract infections. *Int. J. Pharm.* **579**, (2020).

88. Ristroph, K. D. & Prud'homme, R. K. Hydrophobic ion pairing: encapsulating small molecules, peptides, and proteins into nanocarriers. *Nanoscale Adv* **1**, 4207–4237 (2019).

89. Patel, A., Gaudana, R. & Mitra, A. K. A novel approach for antibody nanocarriers development through hydrophobic ion-pairing complexation. *J Microencapsul* **31**, 542–550 (2014).

90. Rouse, P. E. A Theory of the Linear Viscoelastic Properties of Dilute Solutions of Coiling Polymers. *J. Chem. Phys.* **21**, (1953).

91. De Gennes, P. G. Some physical properties of polymer solutions and melts. *Die Makromol. Chemie* **3**, (1979).

92. Dahal, U. R. & Dormidontova, E. E. The dynamics of solvation dictates the conformation of polyethylene oxide in aqueous, isobutyric acid and binary solutions. *Phys Chem Chem Phys* **19**, 9823–9832 (2017).

93. Forero-Martinez, N. C., Baumeier, B. & Kremer, K. Backbone Chemical Composition and Monomer Sequence Effects on Phenylene Polymer Persistence Lengths. *Macromolecules* **52**, 5307–5316 (2019).

94. Lee, H., Venable, R. M., MacKerell, A. D. & Pastor, R. W. Molecular Dynamics Studies of Polyethylene Oxide and Polyethylene Glycol: Hydrodynamic Radius and Shape Anisotropy. *Biophys. J.* **95**, (2008).

95. Wilkins, D. K. *et al.* Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry* **38**, 16424–16431 (1999).

96. Leung, B. O., Brash, J. L. & Hitchcock, A. P. Characterization of Biomaterials by Soft X-Ray Spectromicroscopy. *Materials (Basel).* **3**, 3911–3938 (2010).

97. Yang, C., Lu, D. & Liu, Z. How PEGylation Enhances the Stability and Potency of Insulin: A Molecular Dynamics Simulation. *Biochemistry* **50**, 2585–2593 (2011).

98. Bilati, U., Allémann, E. & Doelker, E. Development of a nanoprecipitation method intended for the entrapment of hydrophilic drugs into nanoparticles. *Eur. J. Pharm. Sci.* **24**, 67–75 (2005).

99. Lepeltier, E., Bourgaux, C. & Couvreur, P. Nanoprecipitation and the 'Ouzo effect': Application to drug delivery devices. *Adv Drug Deliv Rev* **71**, 86–97 (2014).

100. Abbott, N. L., Blankschtein, D. & Hatton, T. A. Protein partitioning in two-phase aqueous polymer systems. 3. A neutron scattering investigation of the polymer solution structure and protein-polymer interactions. *Macromolecules* **25**, 3932–3941 (1992).

101. Raman, C., Berkland, C., Kim, K. (Kevin) & Pack, D. W. Modeling small-molecule release from PLG microspheres: effects of polymer degradation and nonuniform drug distribution.

J. Control. Release **103**, 149–158 (2005).

102. Cabezas, L. I., Gracia, I., de Lucas, A. & Rodríguez, J. F. Validation of a Mathematical Model for the Description of Hydrophilic and Hydrophobic Drug Delivery from Biodegradable Foams: Experimental and Comparison Using Indomethacin as Released Drug. *Ind. Eng. Chem. Res.* **53**, 8866–8873 (2014).

103. Wibel, R., Friedl, J. D., Zaichik, S. & Bernkop-Schnurch, A. Hydrophobic ion pairing (HIP) of (poly)peptide drugs: Benefits and drawbacks of different preparation methods. *Eur J Pharm Biopharm* **151**, 73–80 (2020).

104. Dalwadi, G. & Sunderland, B. An ion pairing approach to increase the loading of hydrophilic and lipophilic drugs into PEGylated PLGA nanoparticles. *Eur. J. Pharm. Biopharm.* 231–242 (2009) doi:10.1016/j.ejpb.2008.08.004.

105. Gaudana, R., Khurana, V., Parenky, A. & Mitra, A. K. Encapsulation of Protein-Polysaccharide HIP Complex in Polymeric Nanoparticles. *J Drug Deliv* **2011**, 458128 (2011).

106. Lan, Y., Chen, B. & Rao, J. Pea protein isolate–high methoxyl pectin soluble complexes for improving pea protein functionality: Effect of pH, biopolymer ratio and concentrations. *Food Hydrocoll.* **80**, 245–253 (2018).

107. Liu, S., Low, N. H. & Nickerson, M. T. Effect of pH, salt, and biopolymer ratio on the formation of pea protein isolate-gum arabic complexes. *J Agric Food Chem* **57**, 1521–1526 (2009).

108. Baler, K. *et al.* Electrostatic unfolding and interactions of albumin driven by pH changes: a molecular dynamics study. *J Phys Chem B* **118**, 921–930 (2014).