REVIEW ARTICLE

Review of the fundamental theories behind small angle X-ray scattering, molecular dynamics simulations, and relevant integrated application

Lauren Boldon*, Fallon Laliberte and Li Liu*

Department of Mechanical Aerospace and Nuclear Engineering, Rensselaer Polytechnic Institute, Troy, NY, USA

Abstract

In this paper, the fundamental concepts and equations necessary for performing small angle X-ray scattering (SAXS) experiments, molecular dynamics (MD) simulations, and MD-SAXS analyses were reviewed. Furthermore, several key biological and non-biological applications for SAXS, MD, and MD-SAXS are presented in this review; however, this article does not cover all possible applications. SAXS is an experimental technique used for the analysis of a wide variety of biological and non-biological structures. SAXS utilizes spherical averaging to produce one- or two-dimensional intensity profiles, from which structural data may be extracted. MD simulation is a computer simulation technique that is used to model complex biological and non-biological systems at the atomic level. MD simulations apply classical Newtonian mechanics' equations of motion to perform force calculations and to predict the theoretical physical properties of the system. This review presents several applications that highlight the ability of both SAXS and MD to study protein folding and function in addition to non-biological applications, such as the study of mechanical, electrical, and structural properties of non-biological nanoparticles. Lastly, the potential benefits of combining SAXS and MD simulations for the study of both biological and non-biological systems are demonstrated through the presentation of several examples that combine the two techniques.

Keywords: *small angle X-ray scattering*; *molecular dynamics*; *protein folding*; *nanoparticles*; *MD-SAXS*; *atomistic simulation*; *ab initio*; *radius of gyration*; *pair distribution function*; *Newtonian equations of motion*

Fallon Laliberte is a Ph.D. student studying under Dr. Li (Emily) Liu at Rensselaer Polytechnic Institute. Her interests are the use of small angle X-ray scattering (SAXS) and small angle neutron scattering (SANS) techniques to study the effects of radiation damage on materials, in particular semiconductors. Ms. Laliberte earned her M.Sc. in physics from the University of Rhode Island in 2013 and her B.A. in physics from the College of the Holy Cross in Worcester, MA in 2011.
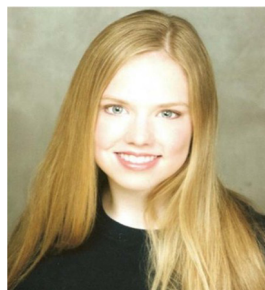
Lauren Boldon is a Ph.D. student studying under Dr. Li (Emily) Liu at Rensselaer Polytechnic Institute (RPI) with a Department of Energy Nuclear Energy University Program Graduate Fellowship. She graduated with her B.S. and M.Eng in Nuclear Science and Engineering from RPI in 2012. Lauren is a member of the Diversity, Women's Affairs and Outreach Committee and was Vice-Chair of the Student Advisory Council for the Mechanical, Aerospace, and Nuclear Engineering Department. Her research areas include small angle X-ray scattering and future energy systems incorporating small modular reactors.

Li (Emily) Liu, *Associate Professor*, Nuclear Engineering and Engineering Physics Program, Rensselaer Polytechnic Institute

As a Physicist and Nuclear Engineer by training, Prof. Liu's research is focused on solving high-impact problems

The theory of small angle X-ray scattering (SAXS), molecular dynamics (MD) simulation, and the combination of the two techniques is reviewed along with potential applications in this article. SAXS is an experimental technique that has become very popular in the biological community with many substantial advancements made over the last few decades, particularly

1

in data collection and analysis (1, 2). SAXS provides detailed structural analysis and physical information for a variety of 1–100 nm and beyond particle systems by characterizing average particle sizes and shapes (3–10). In a SAXS experiment, the sample is exposed to X-rays of a specific wavelength, which scatter elastically between 0 and 5 degrees to produce a spatially averaged intensity distribution (3, 4, 11, 12). *In situ*, static, and dynamical experiments may be performed to analyze samples with solid, liquid, and even gaseous components with crystalline, partially ordered, or randomly oriented structures (3). Data such as the pore size, specific inner surface, surface to volume ratio, solution structure factor, and lattice type and dimensions may be determined depending on the structure being analyzed (3).

Common methods of SAXS spatial or spherical averaging include the Debye formula, multipole expansion, numerical or spherical quadrature, Monte-Carlo sampling, Cubature formula, or Zernike polynomials, utilizing atomic, grid, and/or course grained models of the structure (11, 13–25). The final intensity distribution provides significant information regarding the size, shape, and general structure of the sample through direct calculations of the radius of gyration, volume of correlation, and the Porod invariant in the Guinier, Fourier, and Porod regions, respectively (3, 12, 26–28). The volume, mass, and pair distribution distances are just a few properties that may then be determined; the pair distribution, for example, is determined from an indirect Fourier transformation. Three-dimensional models, such as the *ab initio* model discussed in later sections, may be approximated to represent the most accurate fit to the SAXS profile (3, 11, 29–33).

The spatial averaging required for a SAXS experiment results in low-resolution imaging, as it reduces structural information down to one or two dimensions dependent on the experiment. As a result, the extraction of three-dimensional structural information may be difficult (3, 29–32, 34, 35). Other high-resolution techniques, such as X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, or even electron microscopy (EM), have limitations concerning the analysis of complex or dynamical configurations (11, 36). Despite the low resolution of SAXS imaging, it may be utilized to determine the structure of protein or macromolecular assemblies and to model the kinetics of the system over time (11, 37–39). Furthermore, its application to polymers, nanoparticles, proteins, and so on with different organizational properties and physical states makes it very useful for samples that are otherwise difficult to examine experimentally. High-resolution techniques such as X-ray crystallography, NMR spectroscopy, and EM have limitations concerning the analysis of the complex conformational changes.

associated with energy and the environment through fundamental investigations into the structure-function relationships of materials. For this purpose, she is developing a variety of experimental and computational tools based on neutron, X-ray, and light scattering as well as molecular dynamics (MD) simulations. More importantly, her work focuses on direct nanoscale experimental validation of simulation results as well as the integration of simulation, experiments, and theories at various length scales.

Developing higher resolution SAXS profiles requires supplemental geometric information, which may be derived from MD simulations. MD is a computer simulation technique, which models complex systems at the atomic level. An MD program simulates the motion of atoms by dividing the trajectory of the atoms into states and recording the velocity and position of each atom over time (40). The acting forces and the displacement of the particles are calculated for each time step to determine the new position and state of the particles in the system (40). To model systems of particles, MD simulations employ classical Newtonian mechanics to determine the forces acting on the system, which in turn provide information on the kinetic and thermodynamic properties of the system (40). The force field calculations provide information on the various features of the system at a particular time (41, 42) using the defined position, momentum, charge, bond information, and the potential energy functions (41, 42). Since most systems that are examined in an MD simulation are complex (i.e. more than one particle), it is necessary to calculate the potential functions and resulting forces for both the non-bonded atoms and bonded atoms comprising the system. From each of these potentials, a respective force is derived for the particle at every time step taken throughout the MD simulation.

Due to the complexity of both biological and non-biological systems, MD simulations are becoming increasingly popular for their power to predict and verify experimental results. They provide an opportunity to study the physical characteristics of systems that are not easily examined in the laboratory (43). For example, there is active research aimed toward enhancing the MD algorithms, so they may simulate protein folding and unfolding (44–48). In addition to biological applications, MD simulations have been used to study the physical characteristics of non-biological nanoparticles (49). MD simulation is advantageous in the fields of biology, chemistry, physics, and engineering due to its ability to provide information on system dynamics at the atomic scale, yet has been sparingly utilized for non-biological applications.

Coupling SAXS and MD simulation, or MD-SAXS, holds tremendous potential, especially in the structural and

mechanical analysis of complex particles like proteins and macromolecules with a multitude of conformational changes. SAXS techniques detail the folding patterns of these structures, while MD simulations model the movement between states. MD-SAXS may hold the key to furthering the study of nanoparticles and helping to overcome informational losses in SAXS experimentation due to spherical averaging and determining the theoretical profile, while retaining the inherent experimental flexibility of structures and states. Furthermore, advances in SAXS, such as time-resolved kinetic studies and the potential for 'super-resolution' will only further enhance the capabilities of MD-SAXS (50, 51). This review article will provide theory on SAXS, MD simulation, and MD-SAXS methods and illustrate several examples of each to demonstrate how the uses of MD-SAXS may be expanded to non-biological applications.

SAXS theory and methods

Calculating the theoretical intensity profile

In a SAXS experiment, the intensity is expressed as a function of the scattering vector q resulting from a photon of wavelength λ scattering off the sample at an angle 2θ (Equation 1) (3). For a macromolecule in solution, the intensity distribution of the macromolecule must be obtained by subtracting the buffer or solution profile from the total profile (3, 11).

$$q = \frac{4\pi \sin(\theta)}{\lambda} \tag{1}$$

The following parameters will directly alter the experimental intensity distribution of the system: particle size, volume, contrast (electron density), sample to detector distance, resolution, and beam collimation (3, 28). For calculating the theoretical intensity profile, all of these factors combined with the particle shape will result in greatly different profiles. To better detail the process, spherical and cylindrical particle profile equations are described followed by a flow chart. The electron contrast is calculated from the difference in the mean electron densities for the system and medium (12, 28).

The amplitude of the scattering intensity $A(q)$ for a particle is a function of the electron radial density $\rho(r)$ over particle volume V_p (Equations 2–3) (12, 28). For non-spherical shapes, there are amplitude components in each direction, such as the radial (Equation 4) and longitudinal directions for the cylindrical particle (Equations 5), where L is the length of the cylinder, R is the radius of the sphere or radial cylinder cross section, a is the cosine of the angle formed between the scattering vector and the z axis, and J_1 is the first-order Bessel function (4, 10, 12, 28, 52, 53). For homogeneous

spherical or cylindrical particles, the electron radial density $\rho(r)$ is equal to unity (28, 52, 54).

$$A_{sphers}(q) = \int_V \rho(r)e^{-iqr}dr = 4\pi \int_0^R \rho(r)\left(\frac{\sin(qr)}{qr}\right)r^2dr$$
$$= \frac{4}{3}\pi R^3\left(\frac{3(\sin(qR) - qR\cos(qR))}{qR^3}\right) \tag{2}$$

$$A_{cyl}(q) = L(q) * R(q) \tag{3}$$

$$L(q) = L\left(\frac{\sin\left(\frac{Lq\alpha}{2}\right)}{\frac{Lq\alpha}{2}}\right) \tag{4}$$

$$R(q) = \pi R^2\rho(r)\left[\frac{\mathcal{J}_1\left(qR\sqrt{1-\alpha^2}\right)}{qR\sqrt{1-\alpha^2}}\right] \tag{5}$$

The atomic form factor $P(q)$ represents the particle shape or the interference pattern and is directly related to the size of the particle (1). $P(q)$ is a function of the amplitude $A(q)$ and the particle volume V_p (Equation 6) (10, 12, 50). For a sphere, the equation is straight forward, but for a cylinder, it becomes more complex (Equation 7) (12, 52, 54).

$$P_{sphers}(q) = \frac{A_{sphers}(q)^2}{V_p} \tag{6}$$

$$P_{cyl}(q) = \left[\int_0^1 A_{cyl}(q) * 2\sin\alpha \; d\alpha\right]^2 \tag{7}$$

The theoretical intensity for a particle is a function of the form factor, electron contrast, and particle volume (Equation 8) (12). The total intensity profile is then the sum of each particle intensity for monodisperse particles (Equation 10) or a weighted, normalized summation for polydisperse particles (Equations 9 and 10) (12).

$$I_p(q) = [A(q)]^2 = (\Delta\rho)^2 V_p^2 P(q) \tag{8}$$

$$I(q) = \sum I_p(q) \tag{9}$$

$$d_i(r) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(r-R)^2}{2\sigma^2}\right] \tag{10}$$

$$I(q) = \sum d_i(r)I_p(q) \tag{11}$$

The overall process of calculating the intensity profile for comparison is outlined in a flow chart (Fig. 1). The equations may directly be utilized for spherical or disk/cylinder shaped particles.

Comparing the theoretical and experimental intensity profiles

For an accurate fit between the experimental profile $I_{exp}(q_i)$ and the computed theoretical profile $I(q_i)$, the square root of the chi-squared term X must be minimized
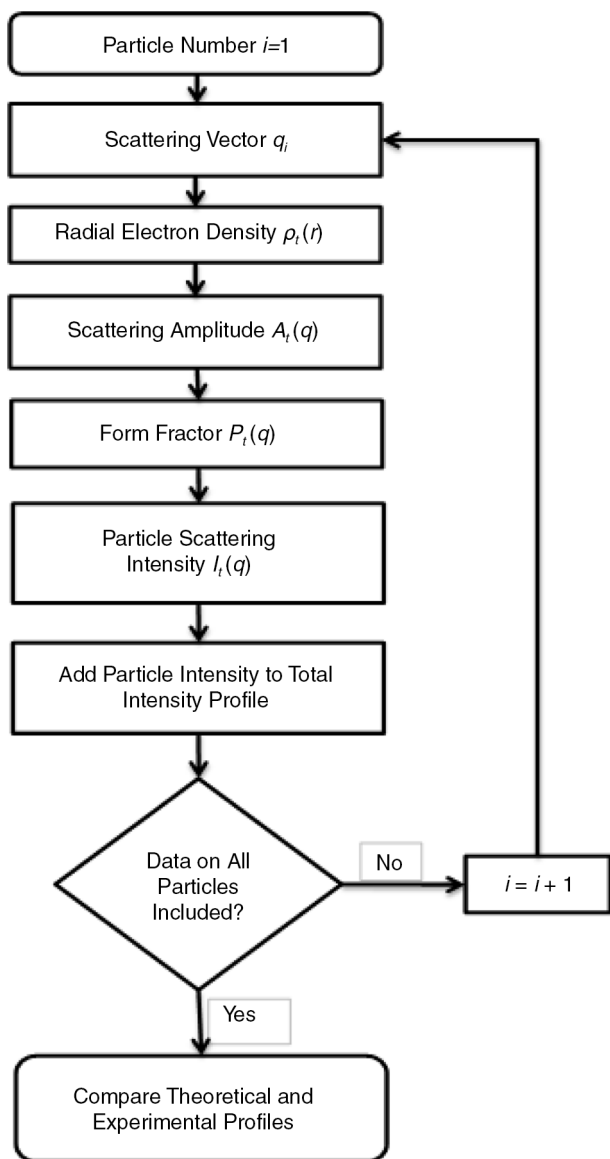
Fig. 1. SAXS flow chart.

over the total number of data points or particles M (Equation 12) (11, 15):

$$X = \sqrt{\frac{1}{M} \sum_{i=1}^{M} \left(\frac{I_{\exp}(q_i) - cI(q_i)}{\sigma(q_i)} \right)^2} \qquad (12)$$

where c is a scaling factor and $\sigma(q_i)$ is the experimental error. Different methods of calculating the theoretical intensity distribution will result in different statistical fits. For macromolecules in solution, several methods of spherical averaging with distinct representations – atomic, grid, or course grained – have been developed (5, 19, 20). As previously mentioned, variations on the Debye formula are commonly utilized (13, 14, 16, 22). Other methods, such as multipole expansion, may be used to speed up the calculation time (15). A balance must be met

between the computational time and accuracy required for a specific application, resulting in many additional methods: Monte-Carlo sampling, numerical quadrature, cubature, and Zernike polynomial expansion (17, 18, 21, 23, 24).

The excluded volume and treatment of the solvent and hydration layer are significant for information on the shape of the particle in the SAXS profile calculation and are also noted within the brief description of each computational method (11). The excluded volume may be assumed to have an electron density equal to that of the bulk solvent (15, 55, 56); utilize explicit water molecules (21); or it may be adjusted based on the fit to the experimental data (15, 22, 23). Similarly, the hydration layer will also effect the profile and may be represented by explicit water molecules (19, 21, 57); solvent density maps (23, 58); or an implicit layer (15, 17, 22). Putnam (37) reviewed methods of constructing a theoretical SAXS profile, providing additional information regarding the equations presented in this paper and significant information on assessing different spherical averaging methods for macromolecular structures. Rambo and Tainer provide a more extensive review of the SAXS theories, including analytical and computational methods as they have developed since the early 1900s (51).

Extracting structural data from the SAXS experimental profile

The SAXS profile has three distinct regions from which information may be extracted: Guinier, Fourier, and Porod (Fig. 2).

In the Guinier region, the experimental radius of gyration R_g may be approximated by fitting a line to the natural log of the intensity as a function of the square of the scattering vector q^2 (3, 12, 26, 28). The radius of gyration will be greatly affected by aggregation of particles, polydispersity, and improper subtraction of the background or buffer.

In the Fourier region, the pair distribution function may be determined by an indirect Fourier transformation
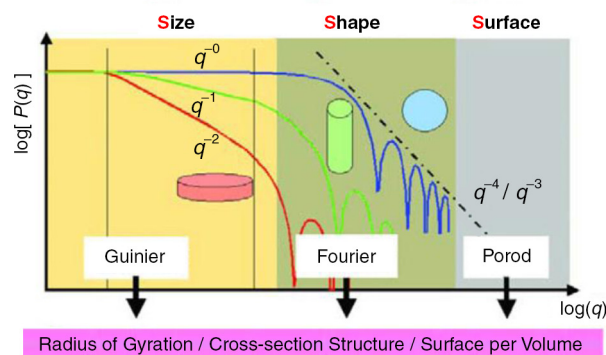

Fig. 2. Regions of SAXS profile and data that may be extracted from each (3).

of the experimental form factor, providing significant information regarding the particle shape $P(q)$ (Equation 13). It refers to the distribution of electrons averaged over a radius r. $\rho(r)$ curves are used to determine the general particle shape, provided all the particles in the sample are of similar shape (3, 12, 26, 28, 59).

$$\rho(r) = \frac{1}{2\pi^2 r} \int_0^\infty q P(q) \sin(qr) dq \qquad (13)$$

In the Porod region, the Porod invariant Q (Equation 14) may be determined, providing surface information such as the surface to volume ratio (Equation 15) and specific surface estimation for compact particles. The log of the form factor is proportional to q^{-4} for particles with uniform density (3, 4, 27, 28).

$$Q = \int_0^\infty q^2 I(q) dq \qquad (14)$$

$$\frac{S}{V} = \frac{\pi(\lim_{q\to\infty} I(q)q^4)}{Q} \qquad (15)$$

A Porod plot of $q^4 I(q)$ *vs* q may provide other useful information such as the Porod volume and molecular weight for compact particles at high q values where asymptotic behavior dominates (3, 28). The Kratky plot of $q^2 I(q)$ *vs* q is another useful tool particularly for protein analysis and determining whether disordered states are present in the sample or whether a protein is unfolded (28).

Ab initio *shape determination*
In small angle scattering, the three-dimensional shape may be reconstructed from the scattering patterns determined by sample geometries (3, 60). *Ab initio* shapes are primarily developed by two models, the envelope function and the bead modeling method. In the first method, the envelope function is used to represent the shape, $r = F(w)$, where r and w are spherical coordinates of the particle, which must be completely contained within the envelope. Spherical harmonics are used to provide details of the envelope, where L is the maximum harmonic order (Equation 16). The envelope model provides a unique solution of the shape. However, it has limitations in accuracy if the shape is overly complicated (29, 33, 60, 61).

$$F(w) = \sum_{l=0}^{L} \sum_{m=-l}^{l} f_{lm} Y_{lm}(w) \qquad (16)$$

In the bead modeling method, Monte-Carlo methods are used to best fit the scattering profile as determined by one of the previous methods described, such as the adapted Debye formula, where $f(q)$ is the form factor for each bead and r_{ij} is the distance between beads. (Equation 17) (13, 60). Densely packed beads or dummy atoms must fill a sphere of radius D_{max} as determined by the scattering profile. The radius r of these beads must be less than D_{max}

and the particle must fit entirely within this sphere. The particle shape is represented by string X of M beads, such that when $X_i = 1$, the bead belongs to the particle, and if $X_i = 0$, it belongs to the solvent (31, 60, 62, 63).

$$I(q) = f^2(q) \sum_{i=1}^{M} \sum_{j=1}^{M} X_i X_j \left(\frac{\sin(qr_{ij})}{qr_{ij}} \right) \qquad (17)$$

Methods of combining simple harmonics and bead modeling have also been performed in order to speed up the computational time required (60).

SAXS applications
SAXS has been fundamental in the study of a plethora of materials and structures both biological and non-biological from complex and large proteins in physiological conditions to crystalline and semi-crystalline materials containing large lattice structures. It may be applied to multi-phase systems as well as porous materials. Furthermore, reaction kinetics or time-resolved SAXS allow for the study of conformational changes. Progress in time-resolved SAXS has improved tremendously with many new techniques and advanced software. Kirby and Cowieson detail the current state of time-resolved SAXS as it applies to the study of dynamic biomolecules (50). Stawski and Benning studies *in situ* and time-resolved SAXS as it applies to precipitation reactions that are difficult to analyze with other techniques (64). Sinko et al. (65) studied the use of SAXS to analyze porous nanoparticles, explicating how the general theory calculations would differ. Doniach and Lipfert provide background information on how SAXS is useful for understanding ligand binding, the folding of structures like RNA, and even in analyzing intermediate folding states (66).

Protein and macromolecular characterization
Understanding key macromolecular and protein functions within the body holds tremendous promise for future medical advancements. With the number of possible protein functions and conformations, being able to predict when and why a protein is in a particular state or how it reacts to environmental changes may hold clues to inducing or inhibiting undesired reactions. Furthermore the ability to analyze the behavior of macromolecules, which make up DNA, RNA, and even cell networks, is significant in understanding biological functions. Petoukhov and Svergun detailed modern techniques of obtaining low-resolution SAXS images, such as methods to characterize macromolecules in solution, identify the conditions in which the macromolecule is close to native conformation, and assess the oligomeric state and quaternary structure (67).

An unfolded protein is randomly oriented with an unpredictable configuration. Once stimulated, the protein will fold into its native conformation, which serves an important biological purpose. Several processes, such as

dehydration and polypeptide collapse, occur to facilitate the folding. Barnase has two distinct folding phases, a burst intermediate phase and hidden intermediate phase, which are dependent on the surrounding conditions. Konuma et al. (68) studied the coupling of these processes and the time-resolved compaction that occurs during folding. Prior investigation into the barnase phases detailed how the burst intermediate phase is the denatured state of the protein when no denaturing substance is present. Kinetic SAXS was used to provide structural information for this phase. The Guinier approximation predicted the radius of gyration over time as barnase folding occurred. The intensity distribution as a function of vector space and the pair distribution function was also determined over time, starting from the unfolded state until the native state was reached. The results obtained align with the proposed folding mechanisms, made viewable with time-resolved SAXS (68).

SAXS has been utilized to study the temperature and pH dependence of three forms of hemoglobin *Glossoscolex paulistus* (HBGp) in order to determine what conditions result in denaturation and/or aggregation and whether these changes are reversible. The hemoglobin protein consists of heme and non-heme structures, linkers, monomers, and trimers (69). The three oxidation states analyzed by SAXS were the oxy, cyanomet, and met HbGp states. The resulting intensity and pair distribution information was utilized to determine critical temperatures and pH values of the hydrodynamic radius for the three oxidation states at various concentrations at which dissociation and denaturation occur (69). The radius of gyration, intensity, and maximum diameter were all determined for the three oxidation states and demonstrate when the hemoglobin protein structure will change. Oxy-HbGp, for example, will retain its native conformation at a pH of 5.0 for temperatures in the range 20–60°C (69). However, at a pH of 7.0, increasing the temperature above 50°C causes an increase in the radius of gyration and the maximum diameter, represented by protein enlargement during denaturation and aggregation. This information on changes to hemoglobin was made clear with SAXS analysis (69). Identifying how factors such as pH and temperature affect the human body and what specific results will occur to particular systems, such as the hemoglobin levels in the blood, provides greater insight into very specific biological functions. Macromolecular states are significant in the stability and action of cell networks. Multiple imaging techniques are required to develop models of the many macromolecular configurations. A process for identifying the flexibility of macromolecules in solution using SAXS profiles in combination with NMR and MX was presented, in which it is possible to differentiate between rigid and flexible particle conformations (Fig. 3) (70). A review of this process was applied to several macro-

molecule structures, including an Mre11-Rad50 complex, for which it was determined that flexibility increases when ATP is not present (70). Once again, it is clear that understanding the conditions in which macromolecular properties change will undoubtedly enhance one's ability to induce or inhibit these changes by affecting specific conditions.

Nucleic acids may be characterized by coupling SAXS and NMR. Burke et al. (71) studied all-atom models for U2/U6 ribonucleic acids in this manner. SAXS provides significant information on the shape of particles in solution and thus proves useful in the low-resolution characterization of nucleic acids. RNA has a highly regular structure with many electrons, making SAXS an excellent method of analysis. Burke et al. (71) detailed how NMR (or SAXS) profiles may be used to refine the results obtained from all-atom modeling. The ability to analyze such significant macromolecular structures will prove useful in future studies of both RNA and DNA which are key to the functioning of all cells in the body.

Surface structural properties and organization
Many materials such as composites may be analyzed with SAXS to determine how surface characteristics are influenced. TiO_2/CNT photocatalytic composites, for example, have been analyzed by SAXS to assess the charge transfer effects on surface roughness and micropore structure (72). These composites, created from hydration/dehydration methods, may theoretically be utilized to extract pollutants from contaminated water and air (72). The homogeneity of these composites, with different size TiO_2 nanoparticles strengthened by single-wall carbon nanotube (CNT) supports, was analyzed with SAXS to view differences in surface roughness and pore structure that might explain prior imaging results, in which all samples except one displayed low removal efficiencies. SAXS demonstrated that this one sample contained large micropores of approximately 0.8–0.9 nm, which was indistinguishable by other methods (72). The additional absorption by these micropores resulted in discrepancies in the interfacial interactions and increased the photocatalytic efficiencies. Multiple imaging techniques, including SAXS, demonstrated that the method of hydration/dehydration is unable to provide TiO_2/CNT composites with the required parameters, due to the distinct phases present in each of the materials. These phases must be combined in a way that promotes synergetic effects from the transfer of charges at the interface (72).

SAXS has often been coupled with a variety of techniques to obtain a more complete picture of organizational structure of a material. Yao Jin et al. (73) developed a method of combining SAXS with Ultrasound-coupled filtration cells to determine changes in the colloidal organization of concentrated Laponite layers. Laponite particles have concentrated layers that were analyzed and
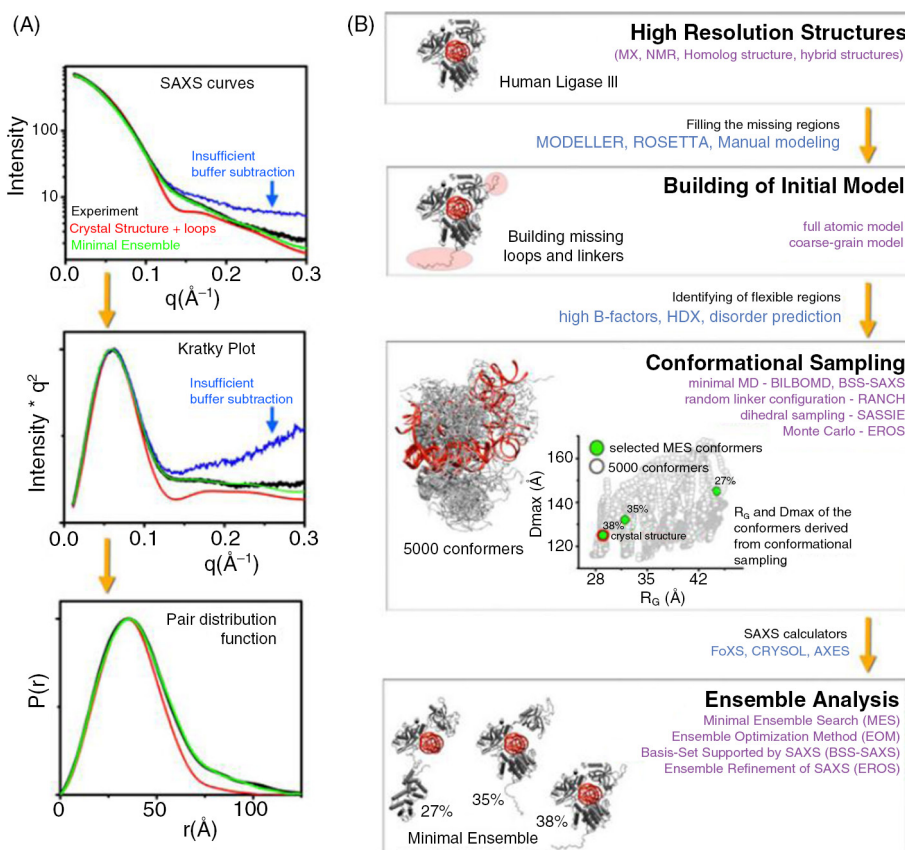
Fig. 3. Flow chart for the method of validating flexibility of macromolecules with SAXS profiles (a) and rigid body modeling (b). DNA Ligase III experimental SAXS profiles (black and blue) were compared to the theoretical profiles for crystal (red) and dynamic models (green) (70).

viewed for the first time with *in situ* time-resolved SAXS-Ultrasound method, demonstrating that the ultrasonic process reduces the concentrated layer when applied during the ultrafiltration process (73). A linear relationship between the volume fractions of Laponite dispersions and the scattering intensity was derived from the SAXS results (73). This type of study demonstrates the usefulness of SAXS and the possibility of coupling it with many other distinct methods to provide enhanced structural analyses.

Mechanical properties such as strain dependence on temperature may also be assessed through SAXS. Polyethylene samples have been tested with *in situ* SAXS to study the effects of temperature and microstructure on strain (74). As strain increases, the polyethylene structure contains both crystalline and amorphous regions. Four different polyethylene samples, PEA, PEB, PEC, and PED, were isothermally crystallized and annealed. The SAXS profiles were then determined and corrected with a Lorentz correction to study the polar and equatorial lamellae where the azimuthal angles are 0° and 90° (74). The strain distribution was analyzed and demonstrated heterogeneity, which could only be explained by interactions between the crystalline and amorphous regions in the polar and equatorial directions (74). In materials, the effects of stresses are significant. A method of analyzing changes between crystalline and amorphous structures and their interfaces in polyethylene will help determine how the plastic is affected when under applied stress. Improved behavior under stress through future material development, while maintaining other necessary properties, may be possible once the structural changes are fully understood. This will only help facilitate the development of improved materials in the future.

MD simulation methods

The following sections are meant to serve as a brief overview of MD theory and provide an example of an MD algorithm. The review by Bernardi et al. (75) addresses sampling methods and techniques for MD simulations of biological systems. Methods to apply MD simulations to small peptides and biological macromolecules are reviewed by Doshi et al. (76). The use of short time trajectory fragments to enhance sampling of kinetics is discussed in the review by Elber et al. (77). These three review articles may be referred to for further information regarding methods, techniques, and theory of how one may apply MD simulations to biological systems.

MD simulation theory: atomic force field

The atomic force field describes the system under study as a collection of atoms, which are held together by interatomic forces. To determine the interatomic force one must know the potential energy $U(\vec{r}_i, \ldots, \vec{r}_N)$ of the N atoms comprising the system as a function of their respective position $\vec{r}$ (43). The position of each of the N atoms is given by the vector $\vec{r}$ (Equation 18).

$$\vec{r} = (x_i, y_i, z_i) \tag{18}$$

One of the first challenges for MD simulation is to determine a realistic potential energy function that will describe the system. To create an accurate potential energy function, the interactions between both bonded and non-bonded atoms must be accounted for. For non-bonded atoms, the two main interactions are the van der Waals force and the electrostatic attraction or repulsion. To describe the van der Waals force, the Lennard-Jones potential is commonly used, while Coulomb's law is used to describe the electrostatic interaction. Thus, the potential energy function for non-bonded atoms in a biological system will take the form of $U(\vec{r}_i, \ldots, \vec{r}_N)_{non-bonded}$, where the value of ε_{ij} and σ_{ij} may be varied depending upon the environment of the system (43). This intermolecular potential energy function implicitly describes the geometric shapes of individual molecules or more specifically, their electron clouds. Therefore, when a potential energy function is defined, molecular behavior, such as rigidity or flexibility, the number of interaction sites per molecule, and so on are also determined (52). Note that the definition for non-bonded, bonded, and total potential energy is typical force field equations that are commonly used to describe biological systems. The terms included in the definition for the potential energy function may vary depending on the particular system (Equation 19).

$$U(\vec{r}_1, \ldots, \vec{r}_N)_{non-bonded} = \underbrace{\sum 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^6\right]}_{\substack{Lennard-Jones\\Potential}}$$
$$+ \underbrace{\sum k\frac{q_i q_j}{r_{ij}}}_{\substack{Coulomb\\Electrostatic\\Potential}} \tag{19}$$

For bonded atoms, the bond length potential, bond angle potential, and torsional potential must be determined. The additive potential of these three interactions is described by $U(\vec{r}_i, \ldots, \vec{r}_N)_{bonded}$ (Equation 20) (43). The first term describes the energies of deformation in the bond length, l_i, and bond angle, θ_i, from the equilibrium position, l_{i0} and θ_{i0}, respectively (43). The third term describes the energy of deformation resulting from rotations around the chemical

bond, where n describes the periodicity of energy terms for rotation (43).

$$U(\vec{r}_1, \ldots, \vec{r}_N)_{bonded} = \sum \frac{a_i}{2}(l_i - l_{i0})^2 + \sum \frac{b_i}{2}(\theta_i - \theta_{i0})^2$$
$$+ \sum \frac{c_i}{2}[(1 + \cos(n\omega_i - \gamma_i)] \tag{20}$$

The atomic force field model considers the potentials for bonded and non-bonded atoms to be additive, thus the force field that is used to describe a biological system will take the form of $U(\vec{r}_i, \ldots, \vec{r}_N)$(Equation 21) (41–49, 78). Once the potential energy of the system is known, the interatomic force $\vec{F}$ is found by taking the gradient of the potential energy with respect to the atomic displacements (Equation 22).

$$U(\vec{r}_1, \ldots, \vec{r}_N) = \sum \frac{a_i}{2}(l_i - l_{i0})^2 + \sum \frac{b_i}{2}(\theta_i - \theta_{i0})^2$$
$$+ \sum \frac{c_i}{2}[(1 + \cos(n\omega_i - \gamma_i)]$$
$$+ \underbrace{\sum 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^6\right]}_{\substack{Lennard-Jones\\Potential}} \tag{21}$$
$$+ \underbrace{\sum k\frac{q_i q_j}{r_{ij}}}_{\substack{Coulomb\\Electrostatic\\Potential}}$$

$$\vec{F} = -\left(\frac{\partial U}{\partial x_i}, \frac{\partial U}{\partial y_i}, \frac{\partial U}{\partial z_i},\right) \tag{22}$$

MD algorithm

In an MD simulation, the forces acting on the system evolves with time. Thus, one can write the position vector as a function of time, where the position of the ith particle is given by $\vec{r}_i(t)$ (Equation 23).

$$\vec{r}_i(t) = (x_i(t), y_i(t), z_i(t)) \tag{23}$$

Since MD simulations utilize classical Newtonian mechanics equations to study the time evolution of the system, one can write the force acting on a particular particle in the system at time t with mass m_i as $\vec{F}_i(t)$ (43). Note that $\vec{F}_i(t)$ is simply Newton's second law of motion (Equation 24).

$$\vec{F}_i(t) = m_i\frac{d^2\vec{r}_i(t)}{dt^2} \tag{24}$$

Integrating the force (Equation 26) yields the atomic momenta, while integrating a second time yields the atomic positions (79). Repeatedly integrating the force several 1,000 times produces individual atomic trajectories from which time averages $\langle A \rangle$ may be computed for macroscopic properties (Equation 25). At equilibrium, the time average does not depend on the initial time t_0.

The time average represents both static properties, such as thermodynamics, and dynamic properties such as transport coefficients (79).

$$\langle A \rangle = \lim_{t \to \infty} \frac{1}{t} \int_{t_0}^{t_0+t} A(\tau) d\tau \qquad (25)$$

The term 'particle' typically refers to a particular atom; however 'particle' could correspond to a larger entity. To perform the MD simulation, one must know all of the forces acting on each of the particles at every time step throughout the simulation, in addition to the initial positions and velocities of all particles. This creates what is called the 'many body problem'. The many body problem states that the quantum Schrödinger equation for any atom other than hydrogen or the classical equations of motion for a system of more than two point masses may only be solved approximately. Therefore, exact solutions are unavailable for more complex systems such as proteins or nanoparticles suspended in a solution. However, the inclusion of parameters from quantum mechanical calculations may better approximate the quantum mechanical result. Furthermore, due to the many body problem, the classical equations of motion are discretized and solved numerically. The position and velocity vectors describe the evolution of the system in phase space as they are propagated for a finite interval by a numerical integrator. One common numerical integrator used in MD simulations is the Verlet algorithm, $\vec{r_i}(t + \Delta t)$, which can be derived from the Taylor expansion of the position $\vec{r_i}(t)$ (Equation 26) (80, 81).

$$\vec{r_i}(t + \Delta t) \cong 2\vec{r_i}(t) - \vec{r_i}(t + \Delta t) + \frac{\vec{F_i}(t)}{m_i} \Delta t^2 \qquad (26)$$

The position and forces acting on each of the particles in the system are continuously updated in the neighbor list. Once the simulation is complete, the desired physical quantities are calculated and the results of the simulation are displayed.

Hard sphere MD simulation algorithm example

There are two types of MD algorithms: soft body algorithms and hard body algorithms. For soft bodies, the acting forces are continuous functions of the distances between molecules (79). For hard bodies, like the classic case of two billiard balls, the discontinuity in the force extends to the intermolecular potential. This section will provide an example of a hard body algorithm. Hard spheres with a diameter of σ interact via a potential energy function $U(r)$ (Equation 27) (79).

$$U(r) = \begin{cases} \infty & r \leq \sigma \\ 0 & r > \sigma \end{cases} \qquad (27)$$

The potential function describes a situation in which the spheres exert a force on one another only when they collide. In between collisions the spheres travel along straight lines at constant velocities (via Newton's laws). Therefore, the simulation algorithm computes the times of the collisions (82). The algorithm calculations are algebraic since the collisions between the two spheres are assumed to be purely elastic. Thus, collisions do not affect the conservation of linear momentum or kinetic energy. The conservation of linear momentum and kinetic energy allow for the MD simulation algorithm to calculate collisions times (82).

The hard sphere algorithm is divided into three main tasks: initialization, equilibration, and production of the result. First, it is important to note that MD programs use a system of units in which the dimensional quantities are unitless. The fundamental dimensions are mass, length, energy, and time (82). In the case of a hard sphere with a diameter of σ, the number density ρ is defined by (82):

$$\rho = \frac{N\sigma^3}{V} \qquad (28)$$

where N and V are the number and hard spheres and the volume of the primary cube enclosing the system, respectively. The packing fraction is then defined by (82):

$$\eta = \frac{V_{spheres}}{V} = \frac{N\sigma^3}{V6} = \frac{\pi\rho}{6} \qquad (29)$$

The total energy of the system $\langle E \rangle$ and the kinetic energy kT are made unitless in the following way (82):

$$\frac{\langle E \rangle}{NkT} = \frac{3}{2} \qquad (30)$$

MD simulations can be used to study a variety of motions and structures. Table 1 (83) below displays the times scales and length over which certain biological processes occur. The following flow chart outlines the process for each of

Table 1. Time scale and length over which molecular motions are calculated in a typical MD simulation program (83).

Motion	Length scale	Time
Local motions, i.e. atomic fluctuations, side-chain motions, loop motions	5 Angstrom	$10^{-15} - 10^{-1}$ sec
Rigid body motions, i.e. helix motions, domain motion, subunit motions	10 Angstrom	$10^{-9} - 1$ sec
Large scale motions, i.e. helix-coil transitions, dissociation/association, folding/unfolding	>5 Angstrom	$10^{-7} - 10^{4}$ sec

the main tasks (initialization, equilibration, production of result) for an MD simulation (Fig. 4).

MD simulation applications

MD simulations allow one to accurately model and simulate a variety of biological and non-biological nanoscale to microscale systems and behavior. The results from MD simulation can be used for its predictive power in experimental work, or can be used to serve as a comparison to experimental results. Currently, MD simulations are being widely utilized to study protein folding and unfolding. One of the greatest challenges for MD simulation concerning the study of protein folding/unfolding to/from the natural state is the time scale and the statistical error. Future research in this area will therefore be aimed at improving the algorithms to better suit the time scale of the protein processes, thereby reducing the statistical error of the calculation. In addition to furthering our understanding of protein characteristics and behavior, future work will focus of the application of MD simulation to study complex nanoscale systems, such as colloidal solutions with nanoparticles. MD simulation has demonstrated its ability to accurately model and predict mechanical, structural, and electrical properties of nanoparticles. In sum, MD simulation has proved its ability to serve as both a powerful predictive tool and method for studying complex systems that are not easily examined in an experimental system.

Protein characterization

One biological application of MD simulations is to study protein folding and unfolding. Proteins fold during a time scale ranging from microseconds to seconds (84). Complications arise due to the length of time required for the folding process. For example, if one were to use the atom-based model for potential energy and solve the time-discretized Newtonian equation of motion for folding from the denatured to native state, the simulation would require years for a protein composing of 100 residues. The experimental transition to this folded state, however, takes only 1 millisecond (84). During an MD simulation, some large proteins will not fold to their native states due to systematic error and reduced stability in the native state (84). Thus, the greatest challenges for MD simulations concerning protein folding are the time scale and statistical error.

On the other hand, protein unfolding is a much simpler process, which can be simulated on shorter time scales of 1–100 nanoseconds (84). Understanding protein folding at the atomic level of detail is important, as protein folding is a complicated reaction that cannot be easily studied from experimentation. However, experimental outcomes are significant when for validation or comparison with simulation results. Experimental results may

also be used to improve the force equations in a simulation by providing more realistic parameters.

Ferrara and Caflisch utilized MD simulation to study the reversible folding and free energy surface of two designed 20-residue sequences called beta3s (TWIQNGS TKWYQNGSTKIYT) and ^{D}PG (Ace-VFITSDPGKTY-TEVDPG-Orn- KILQ-NH). Both of these sequences have a three-stranded antiparallel β-sheet topology (44, 84). The beta3s have been examined previously by Nuclear Overhauser enhancement spectroscopy and chemical shift data, both of which showed that at 10°C, beta3s populates a single structured form-the three-stranded antiparallel β-sheet conformation with turns at Gly6-Ser7 and Gly14-Ser15 (44, 85).

Chemical shift data have shown that the designed amino acid sequence ^{D}PG adopts the three-stranded β-sheet conformation at 24°C in aqueous solution (86). Ferrara and Caflisch performed MD simulations at 300 K and found that both peptides met most of the Nuclear Overhauser enhancement spectroscopy distance restraints (84). Furthermore, they determined that the average effective energy and free energy landscape were similar for both peptides at 360 K despite sequence dissimilarity (84). The average effective energy for the peptides showed a downhill profile at and above the melting temperature (330 K); hence, Ferrara and Caflisch determined that the free energy barriers resulted from entropic losses due to the formation of a β-hairpin (84). They concluded that the free energy surface of the β-sheet peptides is different from a helical peptide, as the helical peptides' folding free energy barrier was much closer to the fully unfolded state than that of the β-sheet peptides (84). In summary, the topology of the peptide determines the free energy surface of the folding mechanism (84). This study demonstrates how MD simulation can be utilized to determine important physical characteristics of biological samples, such as the free energy barrier.

At physiological temperatures, small molecules can exhibit Arrhenius temperature dependence, meaning that they have a faster folding rate at high temperatures (84). However, at even higher temperatures, the folding rate can deviate and become non-Arrhenius (84). Ferrara and Apostolakis investigated the kinetics of folding for an α-helical structure and for a β-hairpin with MD simulations (87). They ran 862 simulations over a time period of four μs for the peptides (87). The negative activation enthalpy at high temperatures was an important feature of folding for both the α-helical and a β-hairpin structures (84). The folding rate increases with temperature, reaching a maximum and then decreases (87).

The observations made by Ferrara and Apostolakis are in agreement with experimental data, thereby validating that at high temperatures, non-Arrhenius behavior may arise despite the interactions being temperature independent (84). This might be explained by the temperature
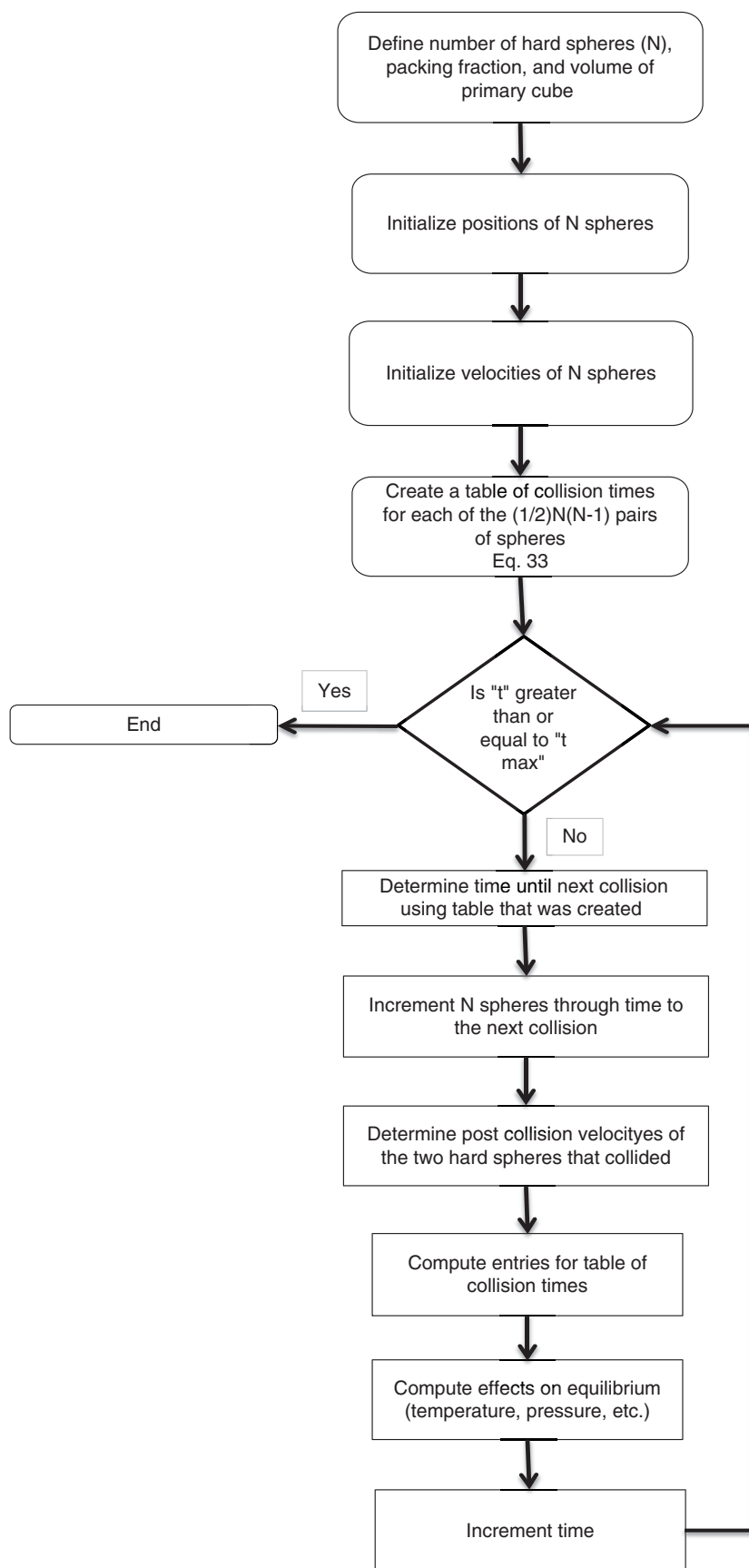
Define number of hard spheres (N), packing fraction, and volume of primary cube

Initialize positions of N spheres

Initialize velocities of N spheres

Create a table of collision times for each of the (1/2)N(N-1) pairs of spheres
Eq. 33

Is "t" greater than or equal to "t max"

Yes

End

No

Determine time until next collision using table that was created

Increment N spheres through time to the next collision

Determine post collision velocityes of the two hard spheres that collided

Compute entries for table of collision times

Compute effects on equilibrium (temperature, pressure, etc.)

Increment time

Fig. 4. Illustrates the overall process of an MD simulation for hard sphere collisions (82).

dependence on configuration space, meaning that at high temperatures, a larger portion of configuration space is accessible, resulting in a reduction in the folding rate (84). This study illustrates how environmental conditions, such as temperature can affect the rate of biological processes (i.e. folding/unfolding). It further demonstrates how MD simulation can strengthen experimental results through theoretical verification.

In addition to studying free energy surfaces and temperature dependence of folding rates, MD simulations have been valuable in studying the function of particular proteins such as the tumor suppressor protein, p53. Upon detection of DNA damage, p53 can either cause cell cycle arrest, allowing for DNA repair, or induce apoptosis. Therefore, the study of the function of p53 is important because this cell regulator protein is often inhibited in many forms of cancer. There is interest in drug development concerning how one can restore the p53 function. MD simulations play an important role because they allow for the study of structural changes and interactions at an atomic scale (87). To study the kinetics of p53, MD simulation packages like AMBER, CHARMM, and GROMACS can be used (87). The binding free energy from these programs may be utilized to study the binding affinities of certain p53 inhibitors, such as MDM2 (87). For example, one paper studied the role of electrostatic interactions concerning the formation of the p53-MDM2 complex by changing the distance of separation between p53 and MDM2 (85). It was found that electrostatic interactions dominate the p53-MDM2 complex formation at long distances while van der Waals interactions of

three residues in p53 (Phe19, Trp23, and Leu26) control the formation of p53-MDM2 at short range (88).

The main challenges that must be overcome for MD simulations of protein folding are building accurate force fields, providing sufficient sampling, and providing robust data analysis (Fig. 5) (89). Concerning the force field models, more work is needed to ensure that the force fields are capable of folding proteins into their native state in agreement with experimental results (89). Sufficient sampling is required to overcome force field inaccuracies. Since MD simulations must integrate Newton's equations of motion with femtosecond time steps, a typical folding simulation requires approximately 10^{12} time steps to reach a millisecond time scale (89). Furthermore, the average system is composed of many atoms (typically $\geq 10^5$ atoms), thereby requiring significant computing power to maintain a neighbor list, which will track the forces acting on and positions of each of the atoms in the system (89). Today, millisecond simulations are now possible due to improvements in computer hardware, software, and sampling techniques (89).

Another challenge for MD simulation is providing the physical knowledge of the system. Data analysis can be very complex, and depending upon the analysis method used, may result in different conclusions. Currently, there are two analysis techniques for protein folding, reaction coordinate, and associated transition states methods and Markov State Models (MSM) (Fig. 6) (89). For the reaction coordinate method, a single coordinate which is capable of describing the process from unfolded to folded is used, and then a kinetics model is built using that
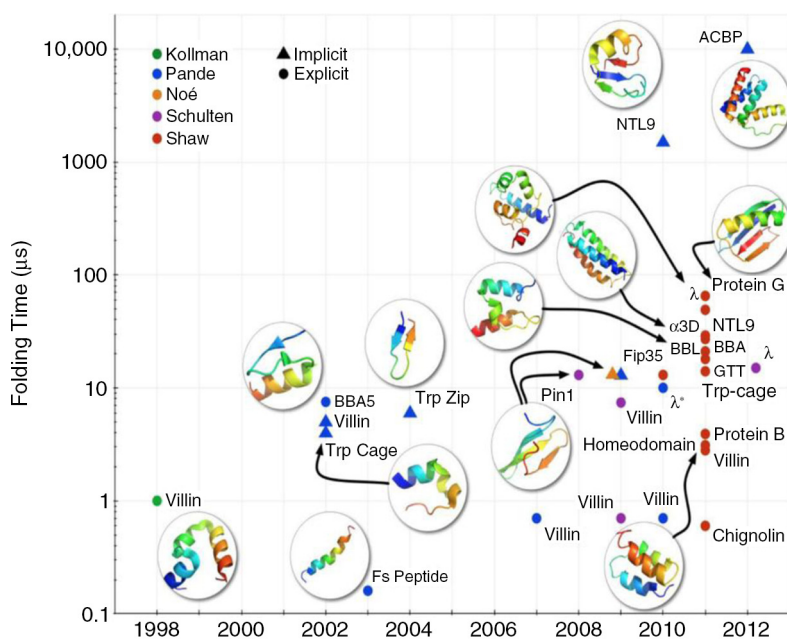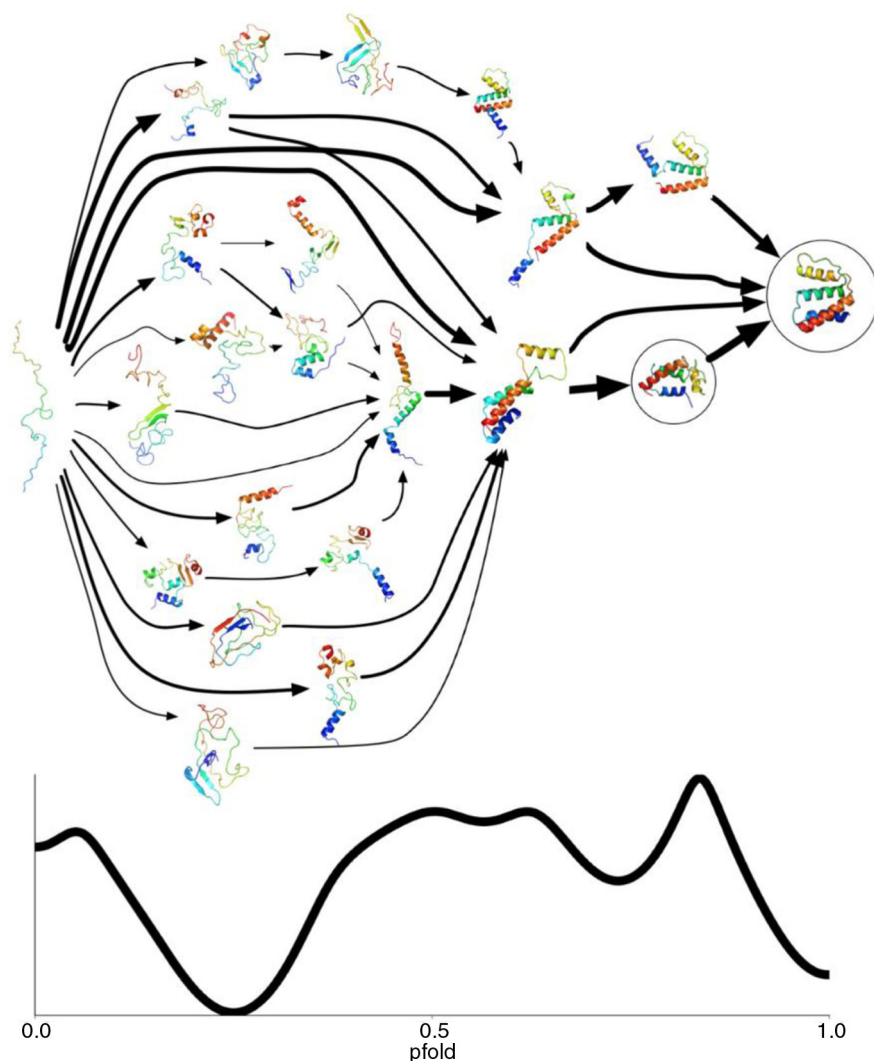

Fig. 5. Protein folding simulations conducted using unbiased, all-atom MD in empirical force fields reported in literature. Some folding times for the same protein differ, due to mutations. For lambda with a (*), the longest timescale seen in that simulation, which was not the folding time, occurred on the order of 10 ms (89).

Fig. 6. Two methods of data analysis, the MSM (top) and reaction coordinate (bottom), shown for the same system (ACBP). The MSM represents folding as interconversion between structurally similar states, and can be illustrated as flow through a network. Reaction coordinates attempt to depict folding as progress along a single degree of freedom, such as the committers (p_{fold}, shown). The MSM picture is more detailed, can capture parallel paths, has tunable resolution, and connects naturally to experiment – all advantages over the coordinate-based approach (89).

coordinate (86, 90, 91–93). The benefit of this method is that it reduces information to a single coordinate and finds one or more transition state ensembles that have equal probability of folding or unfolding (87).

MSM represents folding as first-order kinetics between a set of discrete states, making the data analysis simpler by discarding dynamics below a defined 'lag time' (89). By changing the defined lag time, one can achieve different resolutions (89). Furthermore, MSM can track parallel paths of folding and is generally closer to the experimental results (89). The problem is that the results obtained using the reaction coordinate method and MSM do not always agree. Thus, there is a need to develop a general data analysis method for MD simulations of protein folding and unfolding.

Structural, mechanical, and electrical properties

Anisotropic conductive adhesives (ACAs) are of interest for microelectronic devices, specifically for producing ultra-thin liquid-crystal displays. ACAs in liquid-crystal displays are subjected to compressive stresses during manufacturing and operation; thus, it is important to understand how these mechanical stresses affect their performance (94). There has been experimental research performed on the mechanical response of polymers ranging in size from 2.6 to 25.1 μm (86). He et al. (93) observed that decreasing particle diameters resulted in increasing stiffness of the polymer material. This size effect is very important for the production of electronics.

Zhao et al. (86) utilized a coarse-grained MD model to verify the observed size dependence for the particles.

A coarse-grained MD simulation was utilized because they can be simulated with less time than an all-atom model and because of the availability of coarse-grained potentials for surface tension (86). The coarse-grained model was performed in LAMMPS, an MD simulation source code, for linear polyethylene. As a result, they obtained an entangled molecular model (full atomic), which was then converted to a coarse-grained model where each bead represented three monomer units of polyethylene (86).

Five different polymer particles with different diameters from 5 to 40 nm were constructed (86). For comparison, a bulk coarse-grained model of linear polyethylene was developed using the same potential function (86). Compression loadings were simulated for each of the constructed particles to determine the influence of particle size on mechanical response. Using coarse-grained MD simulations, the nominal stress and strain curves were obtained. The continuum model represents a particle subjected to compressive loading between the two plates, as evaluated by finite element analysis and was included because the size effect does not occur; therefore, this model served as a control reference for comparison other models (86). In addition to studying the loading behavior of different sized particles, MD simulations were also performed for compression unloading of these particles.

From the compression loading and unloading simulations, it was determined that there is a size effect in polymer particles (86). In addition to studying loading and unloading properties of the particles, Zhao et al. (86) examined the surface energy vs. ratio of surface area to volume. Amorphous polymers have a mass density, which is higher at the surface than in the bulk of the material. This higher surface density can be explained by surface tension whereby polymer molecules at the surface are pulled by attractive forces toward the bulk, resulting in surface densification (86). The thickness of this dense region on the surface is constant; thus, for decreasing particle sizes, the volume fraction of the material will increase (86). This property gives smaller polymer particles stiffer mechanical responses to load. It was found that there is a linear relationship between the surface energy and the relative surface area (86). The increases in surface energy and particle stiffness result from the increase in the mass density of the material at the surface (86). This example illustrates how MD simulation can be utilized to study particular characteristics of a system that are difficult to examine in a laboratory setting.

In another study, Torres-Vega et al. (95) determined the threshold for the behavior of copper nanoparticles by studying their structural and electronic properties using MD simulations. Depending upon the size of the copper nanoparticles, they contained between 13 and 8,217 atoms (95). Metallic nanoparticles are of interest because their physical and chemical properties are strongly dependent on their size, structure, composition, and defects (95). MD simulations are crucial for the study of nanoparticles, because experiments can be very difficult to perform and the MD simulations allow one to study different properties at the atomic scale. Torres-Vega et al. (95) determined the threshold above which copper nanoparticles behave like their solid counterpart, in addition to the structural and electronic properties of small and large copper nanoparticles.

The MD simulations were performed using a force field derived from the Johnson potential for copper, which is based on the embedded atom method (95). The Johnson potential has been previously used to study the binding energy and atomic structure of copper nanoparticles and nanowires (95). Torres-Vega defined the total energy of the system E_{total} as the summation of $F(\rho_i)$, the embedded energy of atom i, and $\varphi_{ij}(r_{ij})$, the two-body potential between atoms i and j (Equation 31) (95). The electron density for an atom due to all other atoms ρ_i is defined as the summation of $f(r_{ij})$, the electron density at atom i due to atom j as a function of distance r_{ij} between the two atoms (Equation 32) (95).

$$E_{total} = \sum_i F(\rho_i) + \frac{1}{2} \sum_i \sum_j \varphi_{ij}(r_{ij}) \tag{31}$$

$$\rho_i = \sum_j f(r_{ij}) \tag{32}$$

The Newtonian equations of motion were integrated using a fifth order predictor − corrector algorithm with a time step of 2 femtoseconds (95). For fewer than 2,000 atoms, Torres-Vega et al. (95) observed that the potential energy decreased drastically with increases in the size of the nanoparticle. However, when there were greater than 2,000 atoms comprising the copper nanoparticles, the potential energy stabilized around −3.4 eV/atom, which is approximately the cohesive energy for bulk copper, −3.49 eV/atom (95).

For structural analysis, Torres-Vega et al. (95) noted that nanoparticles with fewer than 2,000 atoms presented a spherical external form while nanoparticles with greater than 600 atoms but less than 2,000 atoms presented an irregular form. The different structures for the copper nanoparticles based upon the number of atoms occurred because the system seeks a structure that will result in a minimum energy. The cohesive energy of bulk copper was described by the power law fit function, U, where $U_0 = -3.52$ eV/atom, $\alpha = 2.31$ eV/atom and $\alpha = -1/3$ (Equation 33) (85, 95). Torres-Vega et al. (95) found that when the number of atoms composing the nanoparticle approached infinity, the value of U_0 approached the value for bulk copper.

$$U = U_0 + aN^{\alpha} \tag{33}$$

For electronic analysis the influence of size on the electronic density of states (EDOS) was analyzed utilizing

a recursion method to calculate the EDOS. The results demonstrated that for small nanoparticles (those with less than 1,000 atoms), the total EDOS presents fluctuations which are caused by large surface to volume ratios; large nanoparticles (greater than 2,000 atoms), on the other hand, present negligible fluctuations (95). To compare the total EDOS for a nanoparticle of size N, the total DOS of crystal line copper model with half a million fcc structured η_T^{fcc}, a parameter Δ_T was defined (Equation 34) (95):

$$\Delta_T = \int [\eta_T^{fcc}(E) - \eta_T(E)]^2 dE \tag{34}$$

where η_T is the total EDOS of a nanoparticle of size N (Equation 35) (95).

$$\eta_T(E) = \sum_{\vec{R}} \eta_{\vec{R}}(E) \tag{35}$$

Furthermore, $\eta_{\vec{R}}(E)$, the local DOS at position $\vec{R}$, is defined by (Equation 36) (95):

$$\eta_{\vec{R}}(E) = \lim_{\xi \to 0}\left(-\frac{1}{\pi}\mathrm{Im}G_{\vec{R},\vec{R}}(E + i\xi)\right) \tag{36}$$

where $G(E + i\xi) = (E + i\xi - H)^2$ is the Green function and ξ is a positive infinitesimal (85). Using Equation 23 for Δ_T, the total EDOS of the copper nanoparticles verses the size was plotted to illustrate the relationship between EDOS and Δ_T.

In summary, using MD simulations and the Johnson potential for copper, Torres-Vega et al. (95) studied the influence of nanoparticle size on the structural and electrical properties and determined the threshold for copper nanoparticle behavior. Small nanoparticles (fewer than 1,000 atoms) exhibit a strong dependence on the surface and have an icosahedral shape, while nanoparticles greater than 3.5 nm (more than 2,000 atoms) exhibit a spherical shape (95). The threshold size predicted by MD simulation was close to the experimentally determined size of 3.8 nm. Finally, for nanoparticles with greater than 2,000 atoms present, an electronic character similar to the corresponding macroscopic counterpart was observed.

Integration of SAXS and MD theory

General process overview for biomolecules

The following process outlines the general steps required to perform a restrained MD simulation where the SAXS profile provides constraints for protein analysis. A requirement for this process is some knowledge of the structure. For the case outlined by Kojima et al. (36) a crystalline structure was used. SAXS theoretical profiles were determined for incrementing atomic coordinates which were then used in the regular MD process.

Additionally, in restrained MD, an extra step is added in which the potential energy is minimized by including a constrained energy value.

Kojima et al. (36) detailed the Restrained MD-SAXS process for proteins in solution using an AMBER united-atom force field model (95). The first stage is to calculate the theoretical intensity profile $I_{calc}(q)$ based on the scattering amplitudes in both a vacuum $A_v(q)$ and in the excluded volume $A_o(q)$ with an average bulk density ρ_o (Equations 37) (34, 96). The average bulk density for this case was calculated with the modified cube model (97). The vacuum and excluded volume intensity profiles $I_v(q)$ and $I_c(q)$ are calculated in the same manner as a typical SAXS profile (Equations 38 and 39) (11, 12, 36). It should be noted that all calculated intensities and gradients are averaged over all potential configurations (36).

$$I_{calc}(q) = |A_v(q) - \rho_o A_o(q)|^2 = I_v(q) + \rho_o I_{vc}(q) + \rho_o^2 I_c(q) \tag{37}$$

$$I_v(q) = A_v(q) * A_v^*(q) = [A_v(q)]^2 \tag{38}$$

$$I_c(q) = A_o(q) * A_o^*(q) = [A_o(q)]^2 \tag{39}$$

The combined intensity profile $I_{vc}(q)$ is calculated from the vacuum and excluded volume scattering amplitudes and their conjugates (Equation 40).

$$\begin{aligned} I_{vc}(q) &= -A_v(q)A_o^*(q) - A_v^*(q)A_0(q) \\ &= -2[Re(A_v(q))Re(A_o(q)) + Im(A_v(q))Im(A_o(q))] \end{aligned} \tag{40}$$

In order to solve for the potential energy, the gradients of each of the scattering intensity profiles with respect to the atomic coordinate must be determined (Equations 41–45) (36). It should be noted that Kojima et al. (36) set the gradient of the excluded volume intensity profile to zero for calculation purposes.

$$\frac{\partial I_v(q)}{\partial r_i} = 2|A_v(q)|\frac{\partial |A_v(q)|}{\partial r_i} \tag{41}$$

$$\frac{\partial I_c(q)}{\partial r_i} = 2|A_o(q)|\frac{\partial |A_o(q)|}{\partial r_i} \tag{42}$$

$$\begin{aligned} \frac{\partial I_{vc}(q)}{\partial r_i} = -2q[&Re(A_v(q))Im(A_0(q)) \\ &- Im(A_v(q))Re(A_0(q))] \end{aligned} \tag{43}$$

$$\frac{\partial I_{calc}}{\partial r_i} = \frac{\partial I_v(q)}{\partial r_i} + \rho_o\frac{\partial I_{vc}(q)}{\partial r_i} + \rho_o^2\frac{\partial I_c(q)}{\partial r_i} \tag{44}$$

$$\frac{\partial E_{const}}{\partial r_i} = -2kE_{const}\left(\frac{\partial I_{calc}(q)}{\partial r_i}\right) \tag{45}$$

Based on the specific system, the values for the scaling factor k and the relative weight $w(q_i)$ will vary. Kojima et al. (36) used specific formulas for these two variables to help reduce the variation between their observed and

theoretical profiles (Equations 46 and 47). The normalization factor N_A is also shown (Equation 58) (36).

$$k = \frac{\sum I_{calc}(q)I_{observed}(q)w(q)}{\sum I_{calc}(q)^2 w(q)} \tag{46}$$

$$w(q_i) = q_i^2 \tag{47}$$

$$N_A = \sum_{q_i=1}^{n} I_{exp}(q_i)w(q_i) \tag{48}$$

Finally, the constraint potential energy and total potential energy may be calculated (Equations 49 and 50) (36, 98, 99). The normal potential energy is derived from the AMBER united-atom force field (96).

$$E_{const} = \frac{1}{N_A} \sum_{q_i=1}^{n} [I_{exp}(q_i) - kI_{calc}(q_i)]^2 w(q_i) \tag{49}$$

$$E = E_{norm} + W_{const} \, E_{const} \tag{50}$$

The weighting factor w_{const} is chosen to balance the potential energy contributions from the constrained and normal potential energies per the energy minimization (36). The potential energy is then used in the regular MD process. A general method of applying these equations and feeding them into the MD simulation is shown in a flow chart (Fig. 7).

MD-SAXS method for proteins

The MD-SAXS method detailed is for protein structural studies, in which the MD simulation trajectories are used to simulate the SAXS data. Typically, continuum solvent models are applied to do so; however, MD-SAXS utilizes explicit water molecule hydrating, where MD simulation is performed for both the pure solvent and the solution. Then the pure solvent intensity profile is subtracted from the solution intensity profile to calculate the excess intensity $I(q)$ where U represents the solution and V represents the pure solvent. (Equation 51) (100).

$$I(q) = I^U(q) - I^V(q) \tag{51}$$

In this integrative MD-SAXS method, two regions for the solvation and bulk layer are created in order to facilitate the necessary calculations. Region O must enclose the entire solvation layer and part of the bulk, so that the experimental and theoretical SAXS profiles will be similar. Region b contains the remaining bulk of the sample. Any particle shape may be analyzed as long as it is fully contained in region O (Fig. 8) (100). The SAXS results are low resolution. Thus, all-atom MD simulation is performed to extract additional information.

To obtain the solution and pure solvent intensities, the form factor for the pure solvent $P_v(q)$ and electron radial densities, ρ_o and $\tilde{\rho}(r')$, of the bulk in solution and the pure solvent with respect to position r, must be determined over region O (Equation 52) (100). The brackets represent configurational averages over the possible protein ensembles that may exist in either the solution

or pure solvent. The bracket for Ω_q represents the orientational average. The tilde over a symbol represents an instantaneous value. Multipole expansion may be used for faster calculation of the intensity (100).

$$I(q) = \left\langle \left\langle \int_o \int_o (\tilde{\rho}(r') - \rho_0 P_v(q))(\tilde{\rho}(r) - \rho_0 P_v(q))e^{-iq(r-r')}d^3r'd^3r \right\rangle \Omega_q \right\rangle n \tag{52}$$

In MD simulations of spherical proteins, the spherical simulation box must include the entire protein and few water molecules, so the water molecules do not increase the computation time. If the shape is not spherical, then the box will contain more water molecules and the computational time will increase. This would happen, for instance, if the protein had an elongated shape (100).

SAXS-MD applications

Protein characterization

The integration of SAXS and MD simulation has already been used in several applications to analyze the more complex structures that are not amenable to the other conventional imaging methods (11, 101). Oroguchi et al. (101) studied intrinsically disordered proteins (IDPs) as well as multi-domain proteins, developing an MD-SAXS method to obtain low-resolution protein complexes in solution using SAXS and to analyze the dynamics of the proteins using MD simulation. The MD simulation helped overcome the loss of information that results in SAXS spatial averaging. The MD-SAXS method was utilized to characterize the structure of Endonuclease EcoO1091. The MD structural ensemble was validated with the SAXS profile (101). This work along with other advancements in structural analysis methods will help further the study of proteins.

SAXS combined with MD simulation is being used to study the structure of proteins in solution. To do this at an atomic resolution, it is necessary to develop a SAXS profile including scattering due to the hydration layer. Calculations for the hydration layer may be performed with MD simulations using either uniform density layers or explicit water molecules. Changes in ionic strength and the resulting changes in the SAXS profile were studied by performing MD simulations. Lysozymes present in hen egg whites were modeled over varying concentrations of NaCl (102). The SAXS profile demonstrated large fluctuations when the concentration of NaCl was varied from 0 to 100 mM and did not converge to the MD simulation predicted profile. Thus, the ionic concentrations and strength will affect the ability of SAXS and MD to be coupled and for the MD-SAXS method to produce high-resolution profiles (102). Once again, it is clear that MD-SAXS holds the potential to unlock new information on protein structures in solution and the effects on protein properties; however, there are limitations in the
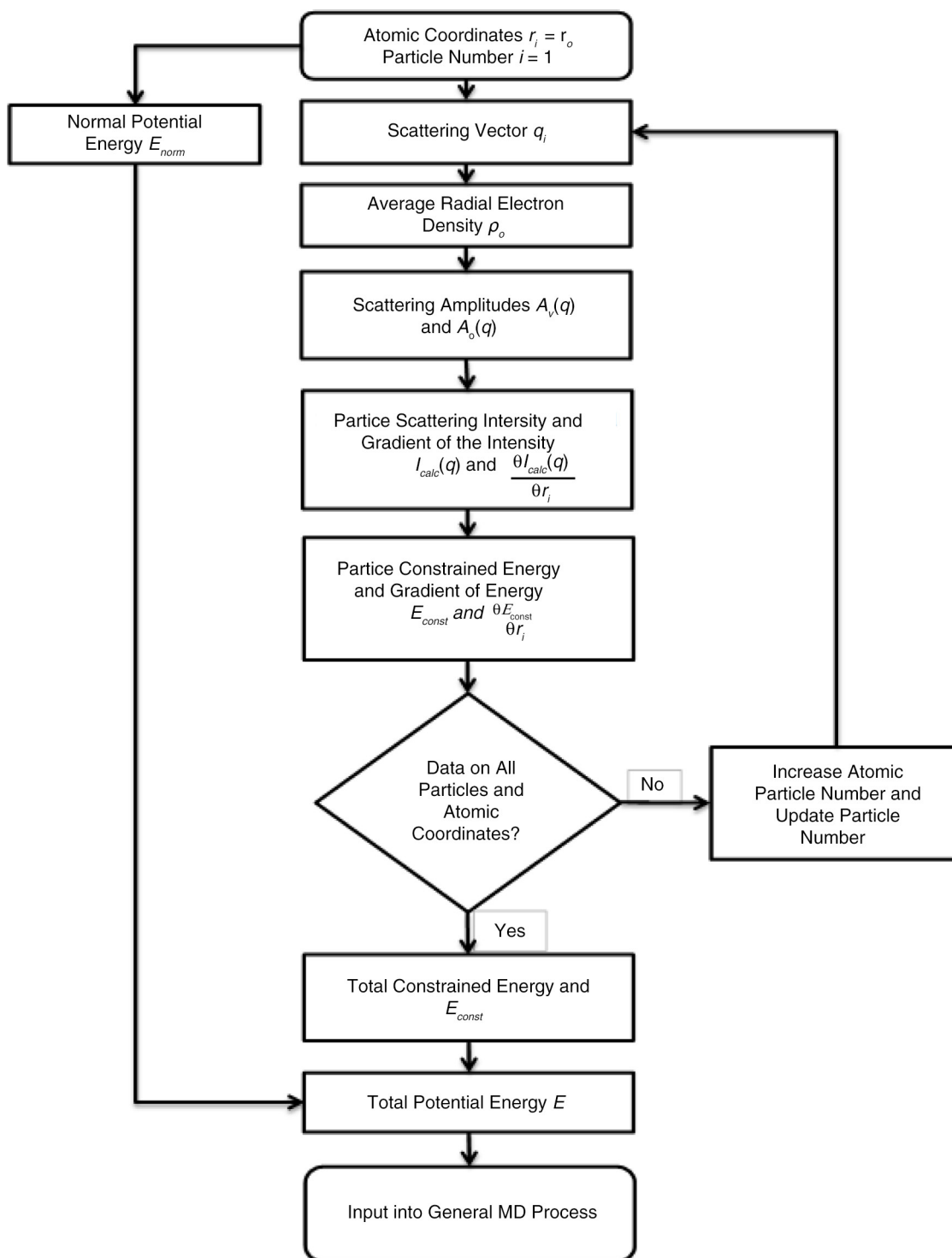
Fig. 7. Integrated SAXS and MD flow chart.

ability to provide the necessary input information in the MD simulation so the results converge with the SAXS experimental data.

Stabilizing membrane proteins, such as G-protein coupled receptors (GPCRs), are pivotal in the study of membrane protein structure. Midtgaad et al. (103)

utilized SAXS and Small Angle Neutron Scattering to study self-assembling peptides, which combine with the membrane phospholipids to yield nanodiscs over time and increase the membrane protein stability (103). The solution structure of ApoA1 peptides in phospholipids before and after assembly into a disc structure was viewed
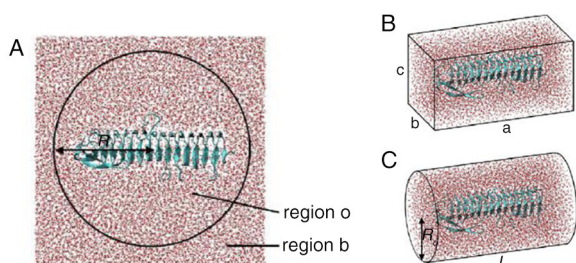
Fig. 8. Spherical (a), rectangular (b), and cylindrical (c) representations of regions o and b for MD-SAXS protein calculations (100).

in detail. Coarse-grained MD simulation of the change to disc shape and resulting dynamics to stabilize the membrane show clear aggregation of these peptides (103). Cell membrane protein behavior and stabilization methods are crucial to cell function and should be studied further.

These three examples represent the potential uses of combining distinct types of analysis – both simulation and experimental practices – and hold future promise for better understanding complex protein processes.

Structural and mechanical properties and conformational changes

There are a variety of properties such as charge that will impact electrical, mechanical, and surface characteristics. Often these characteristics can be analyzed with SAXS and MD coupling. For example, charge imbalance in archaeal lipid bilayers in cell membranes has been studied. Charge imbalance in cell membranes results in high transmembrane voltages, which affects the structure of archaeal lipid bilayers (104). MD simulations for DPPC bilayers mixed with archaeal lipids were performed to investigate these changes based on different charge imbalances and the resulting effects on electrical and pore properties. The electron density profiles, or pair distribution profiles, were compared to SAXS profiles for validation (104). The ability to accurately analyze the behavior of cell membranes under different conditions will greatly enhance our understanding of cells.

Dendrimers are another structure with complex conformations that has been effectively studied with SAXS and MD simulation. These highly branched structures have potential applications in biological and industrial fields and have been difficult to analyze in the past due to their many possible conformations. MD simulation to determine structural properties, such as the radius of gyration, monomer density distribution, molecular surface area and volume, and spatial array of branch points, was performed to analyze the PETIM dendrimer (generations G2–G4) in water at each branch point (105). SAXS imaging was also performed as validation for the MD results. It was determined that the radius of gyration

follows the scaling law $R_g \sim N^{\frac{1}{3}}$ for lower generations and $R_g \sim N^{28}$ for higher generations, where N is the number of monomers present in the dendrimer (105). The MD and SAXS monomer density distribution also showed that back folding was present and the dendrimer had flexible repeat units as expected. The PETIM structure combined with previous PAMAM studies indicate that dendrimers tend toward flexible folding non-spherical conformations (105). MD and SAXS demonstrate how complex and highly folded structures with many conformations may be accurately modeled, providing significant structural information. The behavior of dendrimers in different environments is crucial to many distinct fields and holds untold potential. The physical characteristics of cycloisomaltooligosaccharides (CIs) in solution were analyzed by three methods: viscosity measurements, SAXS, and MD simulation. SAXS was performed over multiple concentrations to determine whether the radius of gyration varied with concentration and how the degree of polymerization affected the shapes of CIs. The radius of gyration for CI was not directly affected by concentration. CIs with different degrees of polymerization resulted in similar SAXS profiles. Thus, there shapes are likely similar, but the radius of gyration will vary as a result of changes in the degree of polymerization. The experimental SAXS profile was compared to theoretical flexible Gaussian ring and rigid ring models. MD simulation was performed to find a ring model that better fit the SAXS data (106). The combination of MD and SAXS to determine a ring model fit demonstrates how simulation and experimental data may be incorporated in an effective manner. It holds tremendous promise in future applications for both biological and non-biological systems.

Conclusion

SAXS studies have been performed for a wide array of materials and structures, providing extremely useful low-resolution one or two-dimensional profiles meant to describe three-dimensional structures. Moving toward higher resolution profiles requires the use of MD simulation to extract relevant three-dimensional geometrical information. This coupled MD-SAXS method combines the advantages of both SAXS and MD simulation and provides a tremendous opportunity for advancement in the materials field for both biological and non-biological applications. Capitalizing on the ease of use and flexibility of sample type for SAXS and the ability of MD simulations to capture significant information on the motion of particles will allow for significant improvement on structural analysis and conformational changes of larger particles. For now, it is clear that MD-SAXS provides a solution to viewing complex protein and macromolecular changes, with SAXS allowing one to view folding and unfolding patterns and MD to track the

movement. Furthermore, MD-SAXS may be the key to revolutionizing nanoparticle studies, as SAXS is well suited for imaging nanoscale materials, and MD is suited for tracking the trajectories of complex nanoparticle components, such as branch points in dendrimers. In the future, the use of MD-SAXS should be expanded to include the analysis of non-biological structures, which is rarely done with MD simulation alone.

Acknowledgements

Conflict of interest and funding

There is no conflict of interest in the present study for any of the authors.

References

1. Perez J, Nishino Y. Advances in X-ray scattering: from solution SAXS to achievements with coherent beams. Curr Opin Struct Biol 2012; 22: 670–8.
2. Graewert M, Svergun D. Impact and progress in small and wide angle X-ray scattering (SAXS and WAXS). Curr Opin Struct Biol 2013; 23: 748–54.
3. Schnablegger H, Singh Y. The SAXS guide: getting acquainted with the principles, third ed. Austria: Anton Paar GmbH; 2013.
4. Li L. Structural analysis of cylindrical particles by small angle X-ray scattering, Dissertation, University of Bayreuth; 2005.
5. Glatter O, Kratky O. Small angle X-ray scattering. London: Academic Press; 1982.
6. Feigin LA, Svergun DI, Taylor GW. Structure analysis by small-angle X-ray and neutron scattering. New York: Plenum Press; 1987.
7. Lindner P, Zemb T. Neutron, X-ray and light scattering introduction to an investigative tool for colloidal and polymetric system. Proceedings of the European Workshop on Neutron, X-ray and Light Scattering as an Investigative tool for Colloidal and Polymeric Systems, May 27–June 2, 1990, Bombannes, France; Holland, 1991.
8. Brumberger H. Modern aspects of small-angle scattering. New York: Springer Science and Business Media; 1995.
9. Hunter RJ. Foundations in colloid science, Volume 2. Oxford: Clarendon Press; 1989.
10. Guinier A, Fournet G. Small-angle scattering of X-rays. New York: John Wiley & Sons, Inc.; 1955.
11. Schneidman-Duhovny D, Kim S, Sali A. Integrative structural modeling with small angle X-ray scattering profiles. BMC Struct Biol 2012; 12: 17.
12. Rochette C. Structural analysis of nanoparticles by small angle X-ray scattering, Dissertation. University of Bayreuth; 2011.
13. Debye P. Zerstreuung von Rontgenstrahlen. Annalen der Physik 1915; 351: 809–23.
14. Perkins SJ, Bonner A. Structure determinations of human and chimaeric antibodies by solution scattering and constrained molecular modelling. Biochem Soc Trans 2008; 382: 1089–106.
15. Svergun D, Barberato C, Koch MHJ. CRYSOL – a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. J Appl Crystallogr 1995; 28: 768–73.
16. Zuo X, Cui G, Merz KM Jr, Zhang L, Lewis FD, Tiede DM. X-ray diffraction "fingerprinting" of DNA structure in solution for quantitative evaluation of molecular dynamics simulation. Proc Natl Acad Sci USA 2006; 103: 3534–9.
17. Tjioe E, Heller WT. ORNL_SAS: software for calculation of small-angle scattering intensities of proteins and protein complexes. J Appl Crystallogr 2007; 40: 782–5.
18. Bardhan J, Park S, Makowski L. SoftWAXS: a computational tool for modeling wide-angle X-ray solution scattering from biomolecules. J Appl Crystallogr 2009; 42: 932–43.
19. Yang S, Park S, Makowski L, Roux B. A rapid coarse residue-based computational method for X-ray solution scattering characterization of protein folds and multiple conformation states of large protein complexes. Biophys J 2009; 96: 4449–63.
20. Stovgaard K, Andreetta C, Ferkinghoff-Borg J, Hamelryck T. Calculation of accurate small angle X-ray scattering curves from coarse-grained protein models. BMC Bioinformatics 2010; 11: 429.
21. Grishaev A, Guo L, Irving T, Bax A. Improved fitting of solution X-ray scattering data to macromolecular structures and structural ensembles by explicit water modeling. J Am Chem Soc 2010; 132: 15484–6.
22. Schneiderman-Duhovny D, Hammel M, Sali A. FoXS: a web server for rapid computation and fitting of SAXS profiles. Nucleic Acids Res 2011; 38: W540–4.
23. Poitevin F, Orland H, Doniach S, Koehl P, Delarue M. AquaSAXS: a web server for computation and fitting of SAXS profiles with non-uniformly hydrated atomic models. Nucleic Acids Res 2011; 39: W184–9.
24. Liu H, Morris RJ, Hexemer A, Grandison S, Zwart PH. Computation of small-angle scattering profiles with three-dimensional Zernike polynomials. Acta Crystallogr A 2012; 68: 278–85.
25. Virtanen JJ, Makowski L, Sosnick TR, Freed KF. Modeling the hydration layer around proteins: applications to small- and wide-angle X-ray scattering. Biophys J 2011; 101: 2061–9.
26. Guinier A. La diffraction des rayons X aux tres petits angles; application a l'etude de phenomenes ultramicroscopiques. Ann Phys (Paris) 1939; 12: 161–237.
27. Porod G. The x-ray small-angle scattering of close-packed colloid systems II. Kolloid Zeitschrift 1952; 125: 51–7 and 109–22.
28. Roessle M. Basics of X-ray scattering presentation. Luebeck University of Applied Science; 2009.
29. Svergun DI. Small-angle scattering studies of macromolecular solutions. J Appl Cryst 2007; 40: s10–17.
30. Svergun DI, Petoukhov MV, Koch MHJ. Determination of domain structure of proteins from X-ray solution scattering. Biophys J 2001; 80: 2946–53.
31. Chacon P, Moran F, Diaz JF, Pantos E, Andreu JM. Low-resolution structure of proteins in solution retrieved from X-ray scattering with a genetic algorithm. Biophys J 1998; 74: 2760–75.
32. Walther D, Cohen FE, Doniach S. Reconstruction of low-resolution three-dimensional density maps from one-dimensional small-angle X-ray solution scattering data for biomolecules. J Appl Cryst 2000; 33: 350–63.
33. Stuhrmann HB. Ein neues Verfahren zur Bestimmung der Oberflaechenform und der inneren Struktur von geloesten globularen Proteinen aus Roentgenkleinwinkelmessungen. Z Phys Chem. Neue Folge 1970; 72: 177–198.

34. Moore PB. Small-angle scattering. Information content and error analysis. J Appl Cryst 1980; 13, 168–75.

35. Taupin D, Luzzati V. Information content and retrieval in solution scattering studies. I. Degrees of freedom and data reduction. J Appl Cryst 1982; 15: 289–300.

36. Kojima M, Timchenko A, Higo J, Ito K, Kihara H, Takahashi K. Structural refinement by restrained molecular-dynamics algorithm with small-angle X-ray scattering constraints for a biomolecule. J Appl Cryst 2004; 37: 103–9.

37. Putnam D. Small angle X-ray scattering profiles to assist protein structure, Dissertation, Vanderbilt University; 2013.

38. Fergusen N, Fersht AR. Early events in protein folding. Curr Opin Struct Biol 2003; 13: 75–81.

39. Dill KA, Ozkan SB, Shell MS, Weikl TR. The protein folding problem. Annu Rev Biophys 2008; 37: 289–316.

40. Meller J. Molecular dynamics. Encyclopedia of life sciences. Nature Publishing Group; 2001. Available from: www.els.net [cited 20 May 2014].

41. Frenkel D, Smit B. Understanding molecular simulation. From algorithms to applications. San Diego: Academic Press; 1996.

42. Heile JM. Molecular dynamics simulations: elementary methods. New York: Wiley; 1992.

43. Salonen E. Introduction to molecular dynamics simulations. Presentation. Ruhr University, Bochum, 23–27 October 2006.

44. Caflisch A, Paci E. Molecular dynamics simulation to study protein folding and unfolding. In: Buchner J, Kiefhaber T, eds. Protein folding handbook part 1. Weinheim, Germany: Wiley-VCH Verlag GmbH; 2005, pp. 1143–69.

45. Mirny L, Shakhnovich E. Protein folding theory: from lattice to all-atom models. Annu Rev Biophys Biomol Struct 2001; 30: 361–96.

46. Creighton TE. Protein folding. New York: W. H. Freeman; 1992.

47. Merz KM Jr., Le Grand SM. The protein folding problem and tertiary structure prediction. Boston: Birkhäuser; 1994.

48. Shea JE, Brooks III, C. L. From folding theories to folding proteins: a review and assessment of simulation studies of protein folding and unfolding. Annu Rev Phys Chem 2001; 52: 499–535.

49. Tian P. Molecular dynamics simulation of nanoparticles. Annu Rep Prog Chem Sect. C 2008; 104: 142–64.

50. Kirby N, Cowieson N. Time-resolved studies of dynamic biomolecules using small angle X-ray scattering. Curr Opin Struct Biol 2014; 28: 41–6.

51. Rambo R, Tainer J. Super-resolution in solution X-Ray scattering and its applications to structural systems biology. Annu Rev Biophys 2013; 42: 415–41.

52. Pedersen JS. Form and structure factors: modeling and interactions presentation. Denmark: Department of Chemistry, iNANO Center, University of Aarhus. Available from: http://www.embl-hamburg.de/biosaxs/courses/embo2012/slides/form-structure-factors-pedersen.pdf [cited 15 April 2014].

53. Fournet G. Theoretical and experimental study of the diffusion of X-rays by dense aggregates of particles. Francaise de Z Mineralogie et de Cristallographie 1951; 74: 37–17.

54. Small-angle X-ray scattering (SAXS). Presentation. Helsinki. Available from: http://www.helsinki.fi/~serimaa/xray-luento/luento-SAXS-particle.pdf [cited 10 May 2014].

55. Fraser RDB, MacRae TP, Suzuki E. An improved method for calculating the contribution of solvent to the X-ray diffraction pattern of biological molecules. J Appl Cryst 1978; 11: 693–4.

56. Lattman EE. Rapid calculation of the solution scattering profile from a macromolecule of known structure. Proteins 1989; 5: 149–55.

57. Park S, Bardhan JP, Roux B, Makowski L. Simulated X-ray scattering of protein solutions using explicit-solvent models. Chem Phys 2009; 130: 134114.

58. Ballauff M. SAXS and SANS studies of polymer colloids. Curr Opin Colloid Interface Sci 2001; 6: 132–9.

59. Svergun D, Koch M. Small-angle scattering studies of biological macromolecules in solution. Rep Prog Phys 2003; 66: 1735–82.

60. Volkov V, Svergun D. Uniqueness of *ab initio* shape determination in small-angle scattering. J Appl Cryst 2003; 36: 860–4.

61. Svergun DI, Stuhrmann HB. New developments in direct shape determination from small-angle scattering. 1. Theory and calculations. Acta Crystallogy 1991; A47: 736–44.

62. Chacon P, Moran F, Diaz JF, Pantos E, Andreu JM. Reconstruction of protein form with X-ray solution scattering and a generic algorithm. J Mol Biol 2000; 299: 1289–302.

63. Svergun DI. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophys J 1999; 76: 2879–86.

64. Stawski T, Benning L. Chapter 5: SAXS in inorganic and bioinspired research. Methods Enzymol 2013; 532: 95–127.

65. Sinko K, Torma V, Kovacs A. SAXS investigation of porous nanostructures. J Non-Crystalline Solids 2008; 354: 5466–74.

66. Doniach S, Lipfert J. Chapter 11: use of small angle X-ray scattering (SAXS) to characterize conformational states of functional RNAs. Methods Enzymol 2009; 469: 237–51. ISSN 0076-6879.

67. Petoukhov M, Svergun D. Applications of small-angle X-ray scattering to biomacromolecular solutions. Int J Biochem Cell Biol 2013; 45: 429–37.

68. Konuma T, Kimura T, Matsumoto S, Goto Y, Fujisawa T, Fersht AR, et al. Time-resolved small-angle X-ray scattering study of the folding dynamics of Barnase. J Mol Biol 2011; 405: 1284–94.

69. Carvalho J, Santiago P, Batista T, Salmon C, Barbosa L, Itri R, et al. On the temperature stability of extracellular hemoglobin of *Glossoscolex paulistus* at different oxidation states: SAXS and DLS studies. Biophys Chem 2012; 163–4: 44–55.

70. Hammel M. Validation of macromolecular flexibility in solution by small-angle X-ray scattering (SAXS). Eur Biophys J 2012; 41: 789–99.

71. Burke J, Butcher S. Nucleic acid structure characterization by small angle X-ray scattering (SAXS). Curr Protoc Nucleic Acid Chem 2012; Chapter: Unit 7.18.

72. Miranda S, Romanos G, Likodimos V, Marques R, Favvas E, Katsaros F, et al. Pore structure, interface properties and photocatalytic efficiency of hydration/dehydration derived TiO2/CNT composites. Appl Catal B Environ 2014; 147: 65–81.

73. Jin Y, Hengl N, Baup S, Pignon F, Gondrexon N, Sztucki M, et al. Effects of ultrasound on colloidal organization at nanometer length scale during cross-flow ultrafiltration probed by *in-situ* SAXS. J Membr Sci 2014; 453: 624–35.

74. Xiong B, Lame O, Chenal J-M, Rochas C, Seguela R, Vigier G. In-situ SAXS study of the mesoscale deformation of polyethylene in the pre-yield strain domain: influence of microstructure and temperature. Polymer 2014; 55: 1223–7.

75. Bernardi RC, Melo MCR, Schulten K. Enhanced sampling techniques in molecular dynamics simulations of biological systems. Biochim Biophys Acta 2014. In Press.

76. Doshi U, Hamelberg D. Towards fast, rigorous and efficient conformational sampling of biomolecules: advances in accelerated molecular dynamics. Biochim Biophys Acta 2014; 1850: 587–858 (April 2015).

77. Elber R, Cardenas AE. Molecular dynamics at extended timescales. Isr J Chem 2014; 54: 1302–10.

78. Miller WH (Ed.). Dynamics of molecular collisions. Parts A and B. New York: Plenum; 1976.

79. Hoover WG. Molecular dynamics. Lecture notes in physics, 258. Berlin: Springer-Verlag; 1986.

80. Kalos MH, Whitlock PA. Monte Carlo methods, Vol. 1: basics. New York: Wiley; 1986.

81. Haile JM. Molecular dynamics simulation, elemental methods. New York: Wiley-Interscience Publication, John Wiley & Sons, Inc.; 1997.

82. Alder BJ, Wainwright TE. Studies in molecular dynamics. I. General method. J Chem Phys 1959; 31: 459.

83. Stote RH. Introduction to molecular dynamics simulations. Presentation. Institute of Chemistry, Universite Louis Pasteur, Strasbourg, France, 18 May 2014.

84. Ferrara P, Caflisch A. Folding simulations of a three- stranded antiparallel b-sheet peptide. Proc Natl Acad Sci USA 2000; 97: 10780–5.

85. Hummer G. From transition paths to transition states and rate coefficients. J Chem Phys 2004; 120: 516–23.

86. Schenck HL, Gellman SH. Use of a designed triple- stranded antiparallel b-sheet to probe b-sheet cooperativity in aqueous solution. J Am Chem Soc 1998; 120: 4869–70.

87. Ferrara P, Apostolakis J, Caflisch A. Thermodynamics and kinetics of folding of two model peptides investigated by molecular dynamics simulations. J Phys Chem B 2000; 104: 5000–10.

88. Dastidar SG, Madhumalar A, Fuentes G, Lane DP, Verma CS. Forces mediating protein-protein interactions: a computational study of p53 "approaching" MDM2. Theor Chem Acc 2010; 125: 621–35.

89. Jane TJ, Shukla D, Beauchamp KA, Pande VS. To milli-seconds and beyond: challenges in the simulation of protein folding. Curr Opin Struct Biol 2013; 23: 58–65.

90. van Kampen N. Stochastic processes in physics and chemistry. Amsterdam, The Netherlands: Elsevier; 2007.

91. Best RB, Hummer G. Coordinate-dependent diffusion in protein folding. Proc Natl Acad Sci USA 2010; 107: 1088–93.

92. Best RB, Hummer G. Diffusion models of protein folding. Chem Phys 2011; 13: 16902.

93. He JY, Zhang ZL, Midttun M, Fonnum G, Modahl GI, Kristiansen H, et al. A size effect on mechanical properties of micron-sized PS-DVB polymer particles. Polymer 2008; 49: 3993–9.

94. Zhao J, Nagao S, Odegard G, Zhang Z, Kristiansen H, He J. Size dependent mechanical behavior of nanoscale polymer particles through coarse grained molecular dynamics simulation. Nanoscale Res Lett 2013; 8: 541.

95. Torres-Vega JJ, Medrano LR, Landauro CV, Rojas-Tapia J. Determination of the threshold of nanoparticle behavior: structural and electronic properties study of nano-sized copper. Elsevier Physica B 2014; 436: 74–9.

96. Weiner SJ, Kollman PA, Case DA, Singh UC, Ghio C, Alagona G, et al. A new force field for molecular mechanical simulation of nucleic acids and proteins. J Am Chem Soc 1984; 106: 765–84.

97. Pavlov M, Federov BA. Improved technique for calculating X-ray scattering intensity of biopolymers in solution: evaluation of the form, volume, and surface of a particle. Biopolymers 1983; 22: 1507–22.

98. Brunger AT, Karplus M, Petsko GA. Crystallographic refine-ment by simulated annealing: application to a 1.5 å resolution structure of crambin. Acta Cryst 1989; A45: 50–61.

99. Jack A, Levitt M. Refinement of large structures by simulta-neous minimization of energy and R factor. Acta Cryst 1978; A34: 931–5.

100. Oroguchi T, Ikeguchi M. MD-SAXS method with nonsphe-rical boundaries. Chem Phys Lett 2012; 541: 117–21.

101. Oroguchi T, Ikeguchi M, Sato M. Towards the structural characterization of intrinsically disordered proteins by SAXS and MD simulation. J Phys Conf 2011; 272: 012005.

102. Oroguchi T, Ikeguchi M. Ionic effect on MD-SAXS profile. Biophys J 2010; 98.

103. Midtgaad S, Pedersen M, Kirkensgaard J, Sørensen K, Mortensen K, Jensen KJ, et al. Self-assembling peptides form nanodisks that stabilize membrane proteins. Soft Matter 2014; 10: 738.

104. Polak A, Tarek M, Tomšič M, Valant J, Ulrih N, Jamnik A, et al. Electroporation of archaeal lipid membranes using MD simulation. Bioelectrochemistry 2014. In Press.

105. Jana C, Jayamurugan G, Ganapathy R, Maiti P, Jayaraman N, Sood AK. Structure of poly(propyl ether imine) (PETIM) dendrimer from fully atomistic molecular dynamics simulation and by small angle X-ray scattering. J Chem Phys 2006; 124: 204719, doi: 10.1063/1.2194538.

106. Suzuki S, Yukiyama T, Ishikawa A, Yuguchi Y, Fuane K, Kitamura S. Conformation and physical properties of cycloi-somaltooligosaccharides in aqueous solution. Carbohydr Polymer 2014; 99: 432–7.

*Li Liu
110 8th St, JEC 5046
Troy, NY 12180
USA
Email: liue@rpi.edu

*Lauren Boldon
110 8th St, JEC 5046
Troy, NY 12180
USA
Email: boldol@rpi.edu