

The Statistical Conformation of a Highly Flexible Protein: Small-Angle X-Ray Scattering of *S. aureus* Protein A

Jo A. Capp,¹ Andrew Hagarman,^{1,2} David C. Richardson,¹ and Terrence G. Oas^{1,*}

¹Department of Biochemistry, Duke University School of Medicine, Durham, NC 27710, USA

²Present address: KBI Biopharma, 1101 Hamlin Road, Durham, NC 27704, USA

*Correspondence: oas@duke.edu

<http://dx.doi.org/10.1016/j.str.2014.06.011>

SUMMARY

Staphylococcal protein A (SpA) is a multidomain protein consisting of five globular IgG binding domains separated by a conserved six- to nine-residue flexible linker. We collected SAXS data on the N-terminal protein-binding half of SpA (SpA-N) and constructs consisting of one to five domain modules in order to determine statistical conformation of this important *S. aureus* virulence factor. We fit the SAXS data to a scattering function based on a new polymer physics model, which provides an analytical description of the SpA-N statistical conformation. We describe a protocol for systematically determining the appropriate level of modeling to fit a SAXS data set based on goodness of fit and whether the addition of parameters improves it. In the case of SpA-N, the analytical polymer physics description provides a depiction of the statistical conformation of a flexible protein that, while lacking atomistic detail, properly reflects the information content of the data.

INTRODUCTION

Many depictions of proteins are single static structures. However, protein structures are highly dynamic, and the function of a protein depends on this dynamicism. Subtle changes in a protein conformation may be side chain movements that facilitate ligand binding in a domain. Another example of protein flexibility is the movement of two domains due to a flexible linker. Finally, the most extreme example of protein flexibility and conformational change is the protein folding-unfolding transition. All these events have been implicated in the function of various proteins (Janin and Sternberg, 2013; Teilum et al., 2009; Tokuriki and Tawfik, 2009).

Staphylococcus protein A (SpA) functions as a crucial *S. aureus* virulence factor through a wide array of intermolecular interactions (Palmqvist et al., 2002). It has been shown to bind to the F_c fragment of antibodies to inhibit host immune response (Deisenhofer, 1981; Moks et al., 1986). It can activate tumor necrosis factor alpha (TNF- α) receptors (Gómez et al., 2004),

leading to the inflammatory response, sepsis, and death of the host. It binds to von Willebrand factor (Hartleib et al., 2000), allowing *S. aureus* to adhere to platelets and withstand shear stress. In addition, it binds to C1qR inhibiting complement pathway activation and the host immune response (Nguyen et al., 2000). SpA also plays a role in biofilm formation (Merino et al., 2009), although the exact mechanism is unknown. As in other systems, this diverse range of SpA functions may be associated with its structural flexibility. A description of SpA's structural flexibility will facilitate a better understanding of the role this property plays in the protein's diverse functions. Within the definition of flexibility, there are two extremes of conformational flexibility to consider: intradomain local motions, consisting of side-chain flexibility and movement of secondary structure elements, and interdomain global motions, i.e., the movement of one domain relative to the others. This study focuses on a description of the interdomain flexibility in SpA.

SpA is a multidomain protein consisting of an N-terminal domain containing a signal sequence and five IgG binding domains (Löfdahl et al., 1983; Moks et al., 1986), and a C-terminal region used to target the protein to the cell surface via an LPXTG motif (Schneewind et al., 1995) (Figure 1A). The N-terminal half of this protein (SpA-N) interacts with the host-cell proteins, mediating the immune response. The five IgG binding domains in the N-terminal half are the functional portion of the protein and have a high degree of sequence identity (Figure 1B). The structure of the D domain has been determined in complex with the F_{ab} fragment of a human IgM antibody (Graille et al., 2000). The structures of E and B domains have been determined by NMR spectroscopy (Starovasnik et al., 1996; Zheng et al., 2004). The 58-residue domains consist of three almost-parallel alpha helices with N- and C-terminal flexible residues (Figure 1C). We have previously determined that the five domains fold independently of each other; there is no thermodynamic coupling between the domains (A.H., W. Franch, Y. Qi, and T.G.O., unpublished data). An NMR dynamics map indicates that there is a six-residue flexible linker between each domain (A.H., W. Franch, Y. Qi, and T.G.O., unpublished data).

It is unknown how the individual domains are structurally related to each other and how the presence of repeated domains structurally constrains the protein. It is these structural constraints that constrain the flexibility of the statistical conformation and affect the thermodynamics of the statistical conformation. So, by determining the statistical conformation

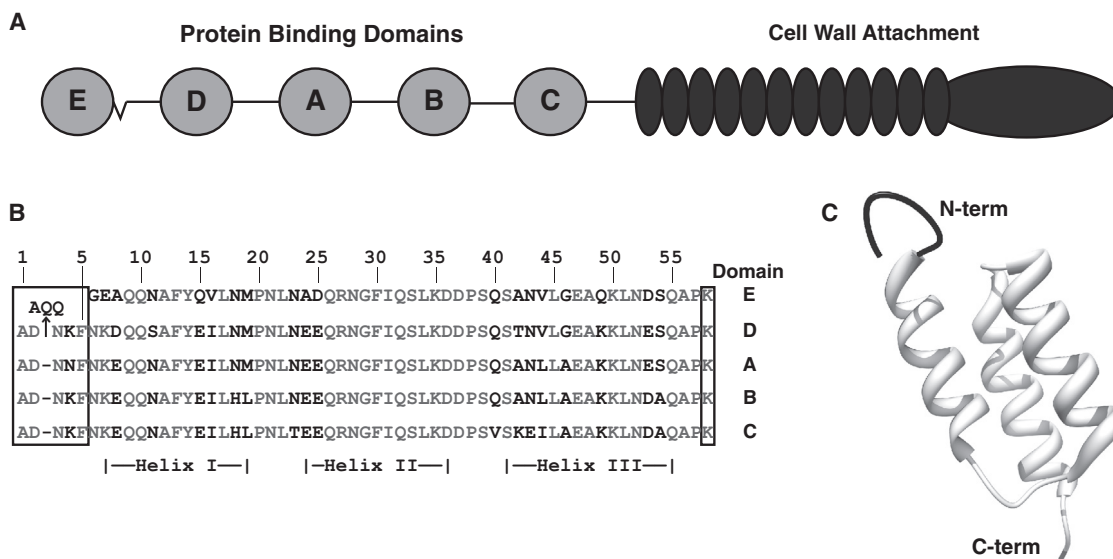


Figure 1. *S. aureus* Protein A

(A) A schematic of SpA shows its two major regions: an N-terminal protein binding region (SpA-N) and a C-terminal domain involved in cell wall attachment. There is a conserved linker (black) between each of the five protein binding domains, E, D, A, B, and C (light gray).

(B) The sequence alignment of the five nearly identical protein binding domains. The sequence identical in all domains is shown in gray. The linker region is boxed in black. There is a three-amino-acid insertion between domains E and D.

(C) The structure of Z-BdpA, a B domain homology. Protein Data Bank (PDB) ID: 1Q2N. Each SpA-N domain consists of a three-helix bundle and flexible N and C termini (shown in black).

we can gain important insights into how the statistical conformation contributes to the function of SpA-N.

In this study, we use small-angle X-ray scattering (SAXS) to determine and describe the statistical conformation of SpA-N. In order to study the statistical conformation of SpA-N and determine the structural relationship between individual domains, we constructed a series of proteins consisting of 1, 2, 3, 4, or 5 repeats of the B domain (BdpA). This simplification of SpA-N into five identical domains allows us to use a polymer physics approach to describe the statistical conformation. A polymer physics approach is one that seeks to describe the statistical properties of a polymer using simple mathematical models. From these simple polymer models, we can describe the “structure” of the thermodynamic state of the ensemble. Small-angle scattering is an ideal technique to analyze the statistical behavior of a polymer because it reports on the structure of the entire thermodynamic ensemble (Rambo and Tainer, 2013). In other words, it provides a structural description of the population-weighted average conformational ensemble—the statistical conformation.

RESULTS

SAXS data were obtained from five B domain protein fragments (BdpA, 2-BdpA, 3-BdpA, 4-BdpA, 5-BdpA) and the N-terminal region of SpA (SpA-N). This allowed us to use a combinatorial approach to study the statistical conformation of SpA-N and derive analytical models to describe that conformational space. Five data sets were collected for each protein fragment at a concentration range of 5–0.5 mg/ml. Data sets were screened for

concentration-dependent effects in the low- q region, and the data sets for each protein fragment that were free of concentration-dependent effects and had the highest signal-to-noise ratio were selected for further analysis.

The Radius of Gyration of SpA-N and n-BdpA Can Be Fit to an Excluded Volume Polymer Model

A Guinier analysis of the scattering data allows for direct estimation of the radius of gyration (R_g) of each protein construct (Koch et al., 2003). The Guinier plot is an algebraic transformation ($\ln(I)$ versus q^2) of the data that produces a linear q^2 dependence in the “Guinier region” found at very small scattering angles ($q < 0.05 \text{ \AA}^{-1}$). The slope of the data is directly proportional to the radius of gyration of the overall protein chain. The q range of the Guinier region is dependent on R_g and the globularity of the molecule (Hjelm, 1985). We determined the Guinier region and R_g of 1-BdpA, which is globular, with the limit of $qR_g < 1.3$, whereas the Guinier regions and R_g of (2-5)-BdpA and SpA-N were determined with $qR_g < 1.0$, the limit of the Guinier region for elongated or flexible macromolecules (Hjelm, 1985; Jacques et al., 2012) (Figure S1 available online). Excellent linear correlations within the Guinier regions are observed for 2-BdpA, 3-BdpA, 4-BdpA, 5-BdpA, and SpA-N data (Figure 2), indicating that each sample was free of self-association or interparticle interference, which might otherwise bias derived models. Interparticle interference is observed in the very low- q region of 1-BdpA, so only the data from $0.0005 < q^2 < 0.01$ were used to estimate R_g . The Guinier plots for the n-BdpA protein fragments and SpA-N show a systematic increase in R_g as domains are added (Figure 2). However, the R_g is not a linear function of the

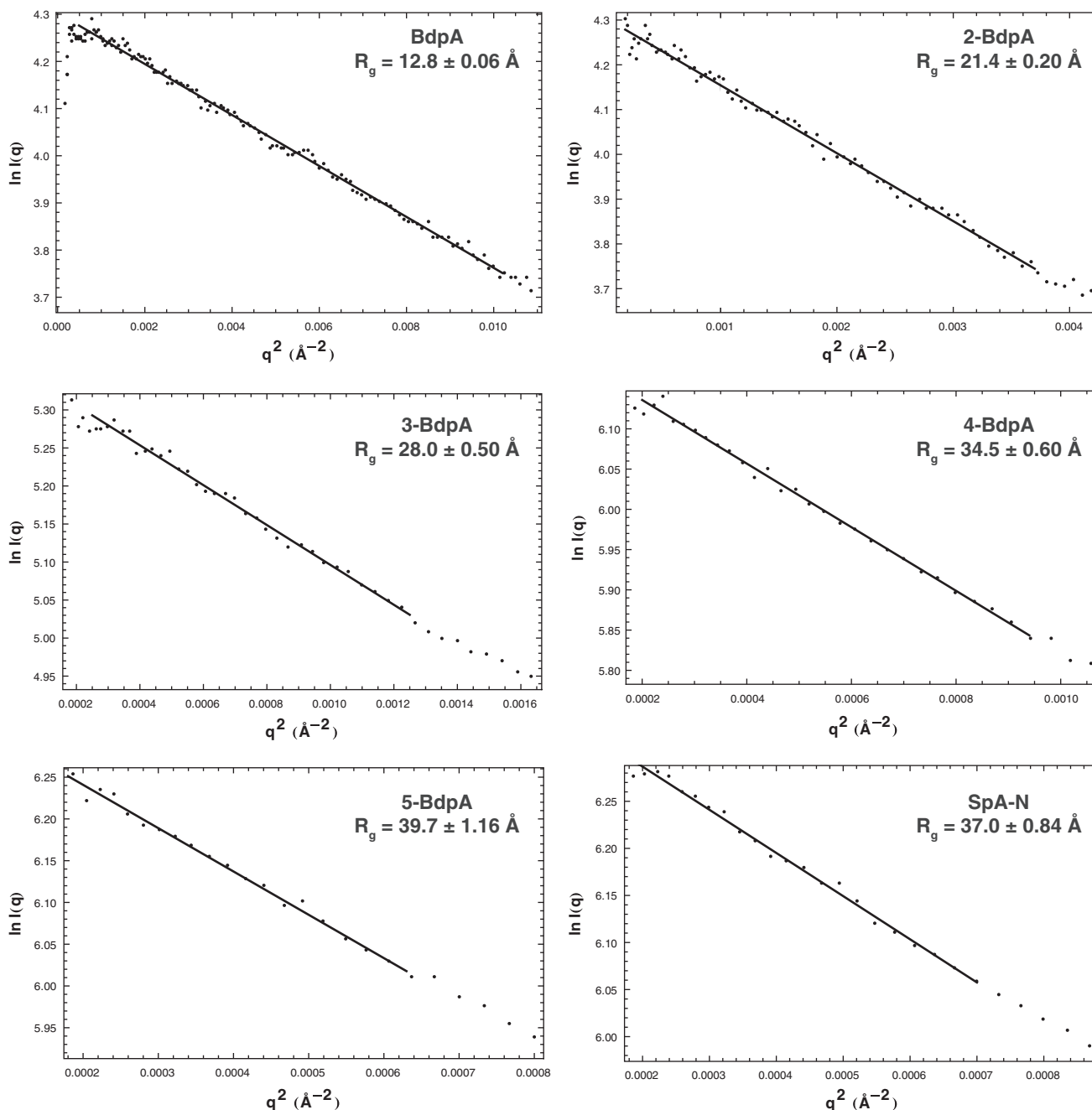


Figure 2. Guinier Analysis of n-BdpA and SpA-N Protein Fragments

The Guinier plots for all constructs show excellent linear correlations in the low- q regions (solid lines). The radius of gyration is not a linear function of monomer number. This indicates that n-BdpA and SpA-N do not explore a conformational space consistent with that of a rigid rod, but rather can be fit to a polymer model. Protein concentrations for each data set: BdpA, 1 mg/ml; 2-BdpA, 1 mg/ml; 3-BdpA, 5 mg/ml; 4-BdpA, 5 mg/ml; 5-BdpA, 5 mg/ml; SpA-N, 5 mg/ml; see also Figure S1.

number of domains, indicating that 5-BdpA and, by extension, SpA-N are not elongated rigid rods, but rather can be described by a polymer model (see below). The SpA-N R_g is 6.8% smaller than that of 5-BdpA, indicating that SpA-N is more compact than 5-BdpA, even though the two molecules have nearly identical molecular weights (32,571 Da for SpA-N versus 33,186 Da for 5-BdpA, 1.8% difference). This difference may be the result of

interdomain interactions that are more favorable in SpA-N, more unfavorable in 5-BdpA, or both.

The systematic increase in R_g among the n-BdpA protein fragments suggests that the relationship between monomer number and R_g can be fit to a polymer model. A simple model that describes the stiffness and conformational space of a polymer is the swollen Gaussian coil. For this model, the R_g is (Hammouda, 1993)

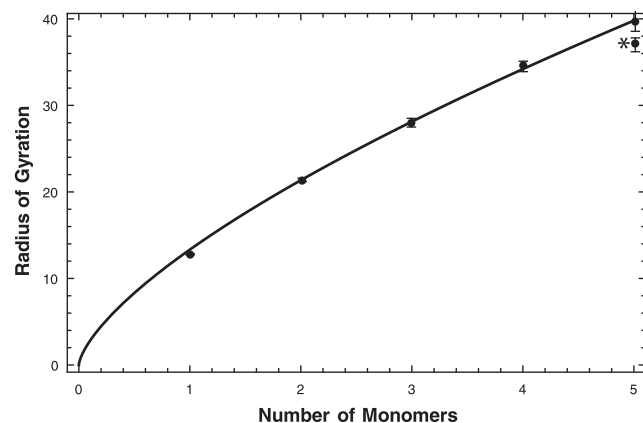


Figure 3. Nonlinear Least-Squares Fit of the Radius of Gyration of n-BdpA to Equation 1

The radius of gyration does not increase linearly with increasing monomer number. The excellent agreement of the data with Equation 1 suggests that n-BdpA behaves as a flexible polymer. The asterisk (*) indicates the radius of gyration for SpA-N. The error bars represent 95% confidence intervals derived from fits of the data depicted in Figure 2.

$$R_g = l_p \sqrt{\frac{N^{2\nu}}{(2\nu+1)(2\nu+2)}} \quad (1)$$

where l_p is the persistence length of the polymer and ν is the Flory coefficient. The persistence length gives the length scale of polymer flexibility, which is reflected in the Flory coefficient. A freely jointed chain has a persistence length on the same order of magnitude as the bond length and $\nu = 0.5$ (Flory, 1953). When the persistence length is on the same order as a bond length and $0.5 < \nu < 1$, then the polymer is said to be semi-flexible (Rubinstein and Colby, 2003). A rigid rod has a persistence length of ∞ and $\nu = 1.0$. It is important to note that this model allows long polymers to intersect with themselves, but the avoidance of short-range monomer intersections is accounted for in the parameters l_p and ν , which reflect chain stiffness, i.e., short-range excluded volume effects.

A nonlinear least-squares fit of the n-BdpA Guinier R_g versus monomer number to Equation 1 gives a persistence length of 37.5 Å (95% confidence intervals: 36.3–38.8 Å) and a Flory coefficient of 0.68 (95% confidence intervals: 0.64–0.72) (Figure 3). The persistence length is comparable to the length of a single BdpA (29.6 Å) and three flexible residues (10.2 Å, assuming 3.4 Å is the average distance between $C\alpha$ atoms in an unstructured polypeptide). The best-fit Flory coefficient is larger than the coefficient for a fully swollen Gaussian coil (0.6) (Flory, 1953), but smaller than that of a rigid rod (1.0), indicating that 5-BdpA and, by extension, SpA-N behave as semi-flexible excluded volume biopolymers, not as rigid rods or ideal Gaussian coils.

The Statistical Conformation of 5-BdpA Can Be Described by the Excluded Volume Pearl Necklace Polymer Model

The radius of gyration is a model-independent measure of the polymer's overall size. It is calculated from the very low- q region

of the scattering curve. The higher- q region of the scattering curve can be used to determine the overall shape and behavior of the polymer by comparing scattering curves of simple shape models to the scattering data. For example, the SAXS scattering curve of simple shapes, such as a sphere, cylinder, wormlike chain, and Gaussian coil, can be used to discriminate between polymer models. Even the scattering curves of a subset of simple shapes can be used to distinguish between polymer models.

One set of polymer models is the Gaussian coil models, where the Flory coefficient and statistical segment length describe the flexibility of the polymer, and the end-to-end distance distribution of the polymer is a Gaussian distribution. The swollen Gaussian coil model describes the monomers as points whose spatial arrangement is that of a random coil, and excluded volume interactions are taken into account by an increase in the Flory coefficient where $\nu > 0.5$ (the Flory coefficient for a random flight polymer). In the scattering function for this model (Hammouda, 1993), there is no term to account for the volume of the monomers except for the statistical segment length and the Flory coefficient.

A subset of Gaussian coil polymer models is the pearl necklace model (PNM). This model has been used to describe polyelectrolytes in various solvents (Dobrynin et al., 1996) and the statistical conformation of long repeat proteins like fibronectin (Pelta et al., 2000). In the pearl necklace model, the monomers are represented as spheres separated by a linker. Previous studies have used various implementations of this model to derive scattering functions for fitting scattering data. These variations include different equations for the relative positions of the “pearl” monomers (structure factor) and the contribution of the linker to the scattering (Table S1). In contrast to the swollen Gaussian coil model, pearl necklace models explicitly define the volume of each monomer and represent it as a sphere. The linear pearl necklace model describes the spheres joined by a rigid rod. The scattering function for this model (Dobrynin et al., 1996) does not include explicit terms for the scattering contribution of the rigid rod or the scattering interference between the spheres and the rigid rod. Thus, the rigid strings connecting the pearls are “invisible” in the scattering function. The parameters in this model are the radius of the spheres and the center-to-center distance between spheres. The random flight pearl necklace model describes spherical monomers connected by freely jointed rods; in other words, the spatial arrangement of the spheres is that of a freely jointed polymer chain. Because the spheres are joined to one another by a freely jointed chain, there are no excluded volume constraints on this model. In other words, two spheres may occupy the same volume. The scattering function for this model (Schweins and Huber, 2004) includes explicit terms for the scattering from the spheres, scattering from the rods, and a cross-term for the scattering interference between the spheres and rods. The fitted-for parameters in this model are the radius of the spheres and the center-to-center distance between spheres. As an addition to these implementations, we have developed a PNM scattering function that represents the spatial arrangement of spheres as a swollen Gaussian coil and explicitly includes terms for the linker and sphere-to-linker scattering (see the Supplemental Experimental Procedures). This model will be referred to hereafter as the excluded volume pearl necklace (EV-PNM) model. The fitted-for

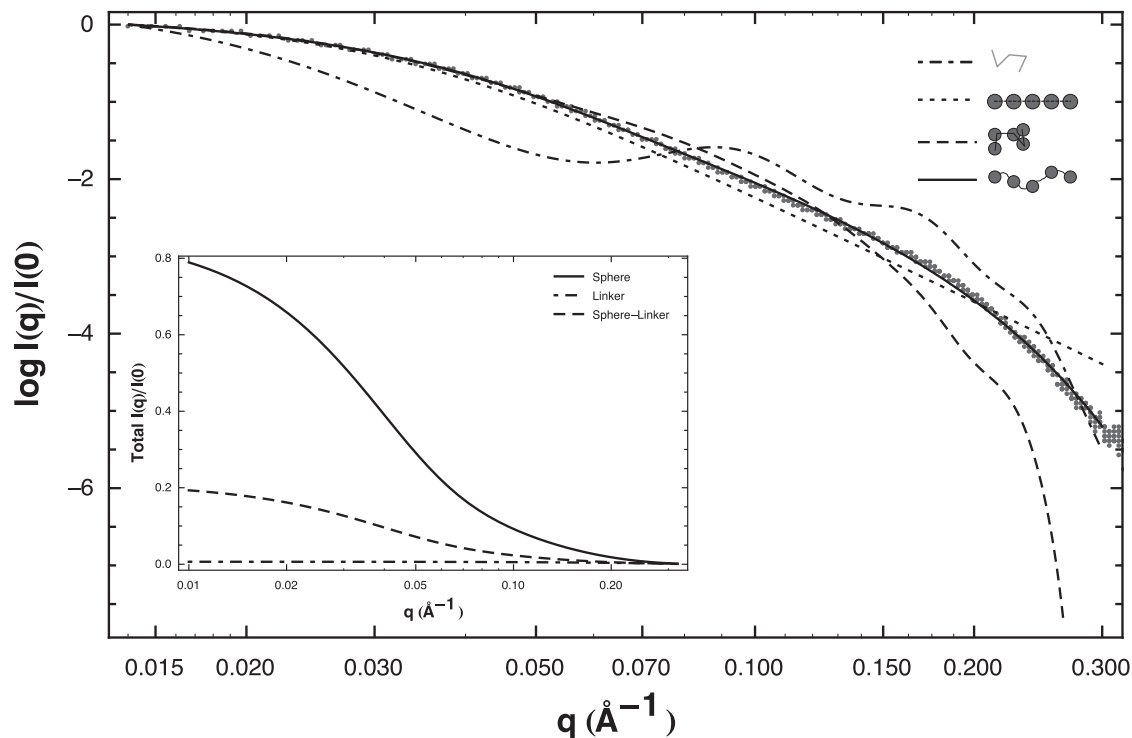


Figure 4. Polymer Model Comparison

Fit of 5-BdpA to polymer models. Points: SAXS data. Short dash: swollen Gaussian coil model. Long dash: linear pearl necklace model. Dash-dot: random flight pearl necklace model. Solid line: excluded volume pearl necklace model. For goodness-of-fit statistics, see Table 1. Inset: plot of $I(q)$ versus $\log q$ showing the component scattering functions of the EV-PNM scattering function ($I_{\text{EV-PNM}}(q)$). Short dash: sphere-sphere scattering and interference. Solid line: sphere-coil interference. Long dash: coil scattering. $I_{\text{EV-PNM}}(q)$ has been scaled to $I_{\text{EV-PNM}}(0) = 1$; see also Table S1.

parameters in this model are the radius of the spheres, the persistence length of the chain, and the Flory coefficient.

In order to select the most appropriate PNM to determine the statistical conformation of 5-BdpA, we performed a weighted nonlinear least-squares fit of the 5-BdpA SAXS data using the scattering functions described above. The adjustable parameters for each scattering function were constrained to be non-negative and nonzero. The results are shown in Figure 4A. The best-fit parameters for each model and a measure of the goodness of fit of the scattering curves to the experimental data are given in Table 1. In the low- q region (0.013 – 0.04 \AA^{-1}), three models fit the data equally well: the swollen Gaussian coil model (Hamouda, 1993), the linear PNM (Dobrynin et al., 1996), and the EV-PNM. The random flight PNM (Schweins and Huber, 2004) does not fit the data. At $q > 0.04 \text{ \AA}^{-1}$, the best-fit model to the scattering curve is the EV-PNM. This result suggests that the EV-PNM model is the most appropriate model to use when modeling the n-BdpA and SpA-N statistical conformations.

There are two simplifications in the scattering function derived from the EV-PNM model: the shape of the domains is approximated as a sphere and the structure factor for coil to coil interference is not included. A spherical form factor describing the shape of BdpA is a reasonable and minimally parameterized model for each domain. A more complicated model for the domains would require more parameters in the scattering function, which would not be supported by the data because the fit is excellent without these parameters. The small difference in the

data versus fit at high q could be due to the nonspherical shape of each domain. The scattering amplitude of each sphere is 8.17 times larger than the scattering amplitude of each coil, so the coil-to-coil interference is a very small component of the total scattered intensity and can be neglected (Figure 4, Inset).

The good fit of the 5-BdpA scattering data to the EV-PNM scattering function suggests that the scattering data from the n-BdpA and SpA-N protein fragments can be globally fit to the model, resulting in a comprehensive description of the conformational space of SpA-N.

Global Fit of the EV-PNM to the Scattering Data

A global fit of the scattering data from 3-BdpA, 4-BdpA, 5-BdpA, ((3-5)-BdpA), and SpA-N protein constructs to the EV-PNM scattering function using a weighted nonlinear least-squares fitting algorithm gives a monomer radius of 11.1 \AA (95.4% confidence intervals: 10.0 – 11.8 \AA), a persistence length of 35.6 \AA (95.4% confidence intervals: 30.2 – 38.0 \AA), and a Flory coefficient of 0.75 (95.4% confidence intervals: 0.72 – 0.80). The persistence length ($37.5 \pm 1.3 \text{ \AA}$) and Flory coefficient (0.68 ± 0.04) determined using the R_g data are within the 95.4% confidence intervals determined by the global fit of all the scattering data.

The fits of the data to the EV-PNM and the goodness of fit statistics are presented in Figure 5. The model captures the dominant features of the scattering data in the (3-5)-BdpA and SpA-N constructs. The best fit of the data to the model is in the 5-BdpA data (Figure 5D; $\chi^2 = 1.06$). The worst fit of the

Table 1. Fit of the 5-BdpA SAXS Data to Polymer Models

Polymer Model	Parameters	95% Confidence Intervals
Swollen Gaussian coil	R_g : 45.00 Å	44.30–45.67 Å
	ν : 0.68	0.67–0.69
	$\chi^2 = 2.48$	
Linear pearl necklace	distance between spheres: 28.32 Å	25.10–31.35 Å
	radius: 15.00 Å	15.97–16.00 Å
	$\chi^2 = 2.28$	
Random flight pearl necklace	distance between spheres: 57.13 Å	54.02–60.23 Å
	radius: 12.81 Å	12.70–12.91 Å
	$\chi^2 = 3.00$	
Excluded volume pearl necklace	persistence length: 36.30 Å	31.3–37.4 Å
	ν : 0.76	0.72–0.80
	radius: 10.5 Å	9.9–11.4 Å
	$\chi^2 = 1.06$	

data to the model is 3-BdpA (Figure 5A; $\chi^2 = 1.18$). The discrepancy may be a consequence of the low number of monomers ($N = 3$), which may limit the ability of the EV-PNM to depict the statistical conformation of 3-BdpA. There is good agreement between the model and data at $0.013 \text{ \AA}^{-1} < q < 0.1 \text{ \AA}^{-1}$, indicating that the model adequately describes the statistical ensemble at large length scales (62–500 Å). There is a systematic deviation between the model and the data over the range $0.1 \text{ \AA}^{-1} < q < 0.32 \text{ \AA}^{-1}$. This q region corresponds to real-space dimensions of 20–62 Å. In this region, deviations in local structure from the model can account for the discrepancy between the model and data, particularly deviations in the shape of the monomers. The BdpA domains are not spherical (discussed below). Our simple EV-PNM model does not take this refinement into consideration. Detailed information about the structure of the monomer units is outside the scope of these experiments because their purpose is to determine the statistical conformation of SpA-N. Simplifying the shape of the domains to a sphere is sufficient to determine the overall shape of SpA-N and (3-5)-BdpA. This minimally parameterized model describes the arrangement of the domains relative to each other at a level consistent with the information content of the data.

The best-fit radius of the monomer sphere is similar to the radius of a single BdpA domain (Figure 6). Figure 6 shows the solution structure of BdpA (Zheng, et al., 2004) (gray) and a sphere of 11.1 Å centered on its center of mass (black). Most of the mass of the globular portion of BdpA is contained within the sphere. There is volume in the sphere that is not occupied by BdpA, but some of the side-chain volume is outside the modeled sphere, which partially compensates for the empty space. These discrepancies confirm that a sphere is not a perfect model for an individual domain. Despite deviations in the atomic details, the agreement between the EV-PNM overall domain dimensions and those from the structure of BdpA supports the accuracy of the persistence length and Flory coefficient we obtained from the global fit of the (3-5)-BdpA data.

Fit of 2-BdpA to the Barbell Model

The scattering data from 2-BdpA cannot be modeled by the EV-PNM since this model assumes $N \geq 3$. These data can, however, be fit to a modified barbell model. This model describes two spheres separated by a line with no scattering mass. Our variation of the scattering model for this model explicitly includes a term for the scattering of the linker (see the [Supplemental Experimental Procedures](#)). We fit the 2-BdpA scattering data to this model and obtained an average distance between domains of 5 Å (± 0.24 Å) and a radius of 15.32 Å (± 0.10 Å). The radius of the spheres is larger than that determined by the global fit of (3-5)-BdpA and SpA-N to the EV-PNM, but it is still reasonable given the structure of BdpA.

DISCUSSION

Model of the SpA-N Statistical Conformation

The persistence length of a polymer can be used to describe the length scale of polymer flexibility. The statistical conformation of a polymer at length scales smaller than the persistence length is rigid, while at length scales larger than the persistence length it is flexible (Fujita, 1990). The persistence length obtained from the global fit of (3-5)-BdpA and SpA-N (35.6 Å) is similar to the length of a single BdpA domain (29.6 Å). This result indicates that SpA-N is a highly flexible biopolymer. To gain a more intuitive understanding of how flexible the biopolymer is, we can compare the persistence length of (3-5)-BdpA and SpA-N to the persistence lengths of other well-known polymers and biopolymers (Table 2). The persistence length of an unfolded polypeptide (6.6 Å) (Lairez et al., 2003) is an order of magnitude smaller than that of SpA-N. This result is expected since SpA-N is not an unfolded protein, but rather is composed of five globular domains. Surprisingly, the persistence length of SpA-N is quite similar to that of single-stranded DNA (22.2 Å) (Chi et al., 2013), even though the monomer dimensions are very different.

We must ask: what are the physical and chemical properties of SpA-N that confer this flexibility? The excluded volume constraint imposed by the dimensions of the monomers lends stiffness to the chain. The peptide bond geometry, sterics, and solvent-peptide interactions constrain the statistical conformation of the linker lending stiffness to the chain. However, if there were interdomain attraction or if the linker was completely rigid, the chain would have a much larger persistence length and much larger chain dimensions. We therefore hypothesize that the only constraints on the statistical conformation of SpA-N are excluded volume interactions between domains and the constraints imposed by the chemical properties of the linker. This hypothesis is consistent with our previous research on n -BdpA and SpA-N. Careful denaturation experiments comparing the stability of each isolated domain and the same domain within the SpA-N molecule showed that the folding of SpA-N domains is thermodynamically uncoupled (A.H., W. Franch, Y. Qi, and T.G.O., unpublished data). Similar studies showed that n -BdpA has the same denaturation curve for $n = 1-5$, again demonstrating the lack of thermodynamic interaction between domains (A.H., W. Franch, Y. Qi, and T.G.O., unpublished data). NMR relaxation studies of backbone ^{15}N - ^1H pairs showed that the order parameters of all residues in 5-BdpA are high except for those in the termini and a six-residue linker between each

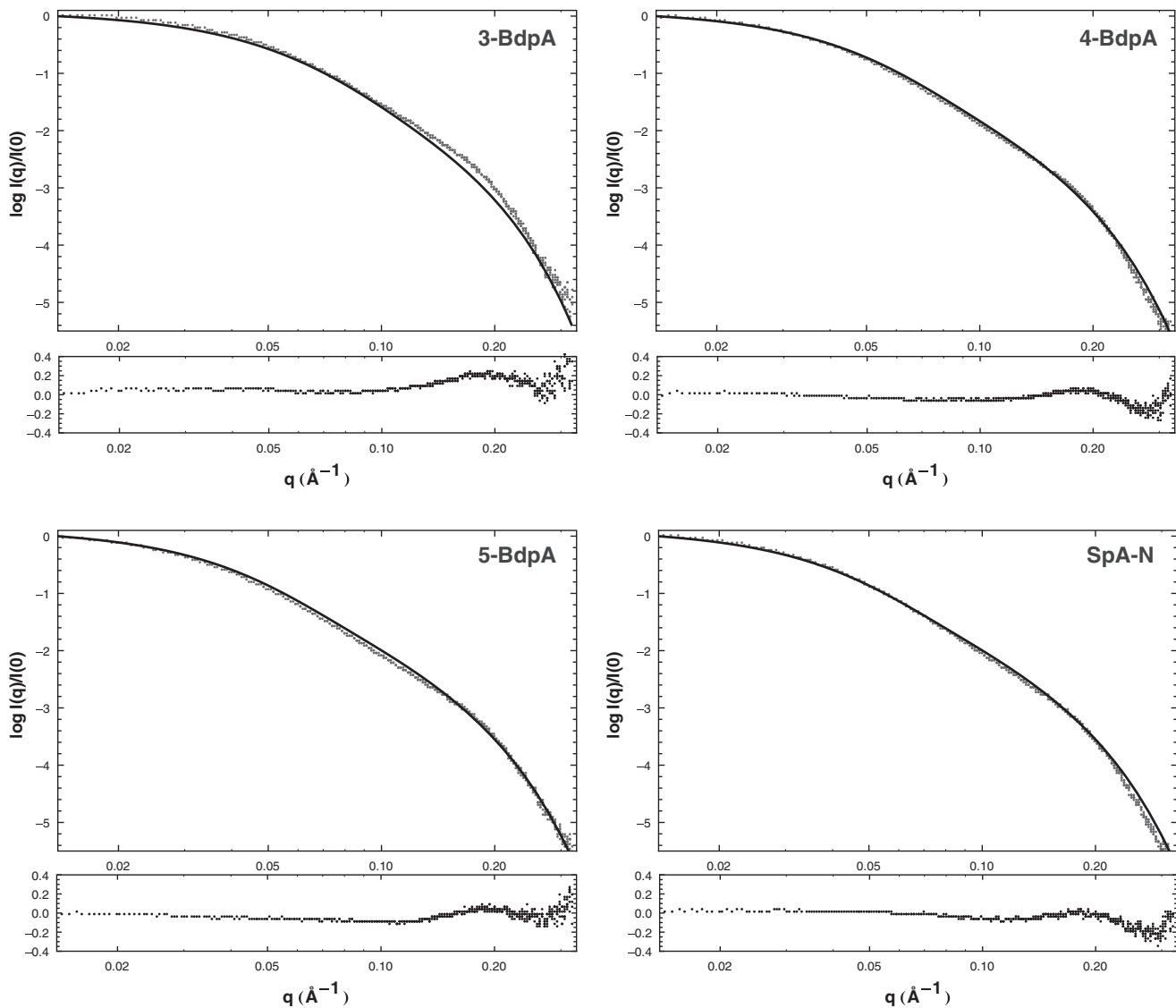


Figure 5. Global Fit of the EV-PNM Model to the (3-5)-BdpA and SpA-N SAXS Data

Points: data. Solid line: model. Bottom panel of each plot: residuals (data model). The χ^2 statistics for the global fit of the model to each data set are: 3-BdpA, 1.18; 4-BdpA, 1.12; 5-BdpA, 1.06; SpA-N, 1.07; see also [Figure S2](#) and [Table S2](#).

domain (A.H., W. Franch, Y. Qi, and T.G.O., unpublished data). These results demonstrated that the linker residues are almost as flexible as the corresponding residues in the termini. Taken together, both previous results and the present SAXS data strongly support the conclusion that SpA-N and n-BdpA are highly flexible chains of inflexible domains that lack any significant favorable interdomain interactions.

Describing the statistical conformation of SpA-N and n-BdpA with a polymer physics model provides a continuous description of conformational space—we do not discretize conformational space into an ensemble of unique conformers. Based on the SAXS data presented above, the global conformational space of SpA-N includes the fully extended conformations and quite compact conformations and all conformations between these two extremes, as long as they avoid steric overlap. The SpA-N

SAXS data are inconsistent with a single compact conformation, or a thermal blob (multiconformation compact conformers), or an elongated conformation. Instead, our EV-PNM fit of the SAXS data implies that the statistical conformation includes all these conformations and intermediate conformers as well.

It is important to note that this continuous description of the statistical conformation is consistent with our previous knowledge of the structure of SpA-N and is consistent with the information content of the SAXS data. Describing the statistical conformation of SpA-N to any higher “resolution” would overinterpret the SAXS data and overparameterize the model of allowed SpA-N conformations. Our SAXS data cannot provide us with any information about the atomistic detail of each domain or linker, cannot help us determine a “minimal ensemble” of SpA-N conformers (see below and the [Supplemental Experimental](#)

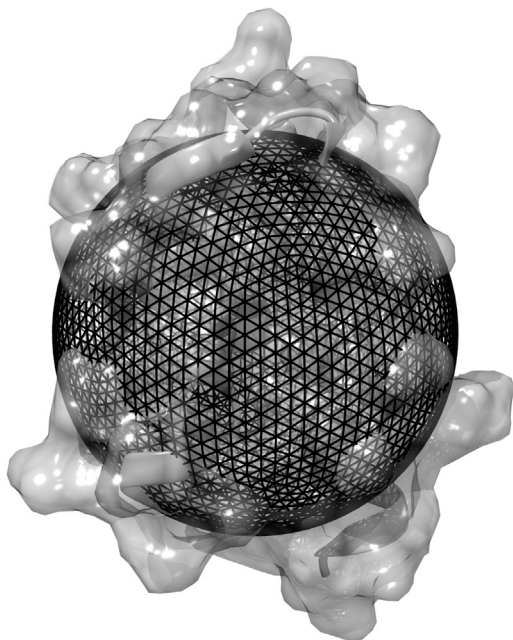


Figure 6. Comparison of a Single BdpA Structure to an EV-PNM Sphere

Space filling model of BdpA (gray) (PDB: 1Q2N); Zheng, et al., 2004, with the flexible residues not shown, superimposed with a sphere of 11.05 Å (black).

Procedures), and cannot provide information about any anisotropic domain-domain motion in the statistical conformation. It may be that such anisotropies exist because of the nonspherical shape of the domains, but given the excellent fits to the simplistic EV-PNM model, we conclude that the SAXS data do not contain any information regarding such anisotropies. Other biophysical techniques are needed to further limit the allowed conformational space. Most importantly, the SAXS data cannot provide us with any information about the distribution of conformational space, since SAXS data are a population-weighted average of all allowed conformational space. The continuous statistical description of conformational space provided by the EV-PNM most accurately describes the allowed conformational space of SpA-N and is consistent with the information content of the SAXS data. Therefore, we do not present a “structure” of the statistical conformation of SpA-N beyond the end-to-end distance distribution, persistence length, Flory coefficient, and radius of the identical spheres that represent the domains.

The Structural Flexibility of SpA May Contribute to Its Functional Plasticity and Allow for Maximum Binding of Ligands in the Extracellular Environment

An evolving view of the protein structure-function relationships is that conformationally dynamic proteins can exhibit functional promiscuity (Tokuriki and Tawfik, 2009). Some highly flexible proteins can recognize multiple ligands at a single binding surface. This structural flexibility allows for the accommodation of mutations as two proteins coevolve and allows a single molecule to interact with multiple structurally unique binding partners.

SpA exhibits both structural flexibility and functional plasticity. Its sequence has evolved to perform a wide array of functions

Table 2. The Persistence Length of Various Polymers

Polymer	Persistence Length (Å)	Reference
SpA-N	35.6	this work
Unfolded polypeptide	6.6	Lairez et al., 2003
Polystyrene	10	Wignall et al., 1974
Single stranded DN	22.2	Chi et al., 2013
Double stranded DNA	500	Hagerman, 1988

that confer virulence to *S. aureus* including binding to both F_c and F_{ab} fragments of antibodies (Deisenhofer, 1981; Moks et al., 1986); von Willebrand factor (Hartleib et al., 2000); and TNF- α receptor (Gómez et al., 2004). This panoply of binding partners no doubt requires corresponding structural plasticity on the part of SpA. One manifestation of this structural plasticity might be the flexibility we observe between domains. Because the SpA statistical conformation includes a large ensemble with a variety of interdomain orientations, it can accommodate multiple ligands binding in many different contexts. In principle this would give it the potential to rapidly evolve in response to changing environmental conditions, conferring resistance and adaptability to *S. aureus*.

The high flexibility of SpA assures that the surface available for interaction with cognate binding partners around each SpA attachment in the cell wall is maximized. The high abundance of SpA in the *S. aureus* cell wall (Sjöholm et al., 1972) suggests that, in aggregate, this interaction surface could represent a large fraction of the bacterial surface. This high surface availability may be a key determinant of SpA’s function as a virulence factor: both the abundance of SpA and its flexibility would maximize SpA-ligand interaction.

Because SpA flexibility is the consequence of a short, conserved six-amino-acid segment between domains, it should be feasible to fully explore the sequence dependence of flexibility. If such studies were to yield a set of sequences with a wide range of flexibilities, it would be possible to directly test the biological significance of SpA flexibility.

Modeling the Statistical Conformation of Highly Flexible Proteins

It is common to use small-angle scattering data to model the statistical conformations of highly flexible and unfolded proteins (Bernadó and Svergun, 2012b). Three modeling approaches—*ab initio* modeling, ensemble modeling, and polymer-physics-based modeling—have been employed to model the conformational space of a biopolymer based on SAXS data. *Ab initio* modeling and a combination of *ab initio* modeling and rigid-body modeling have been used quite successfully to interpret SAXS data from biological macromolecules that lack large-scale flexibility (Putnam et al., 2007). These approaches model a single 3D envelope or structure of the macromolecule that is consistent with the SAXS data. However, when the SAXS data are conformationally averaged due to large-scale flexibility, this approach yields unreliable models. Ensemble-based methods seek to overcome the aforementioned limitations in *ab initio* and rigid-body modeling by describing the conformational space as a discrete

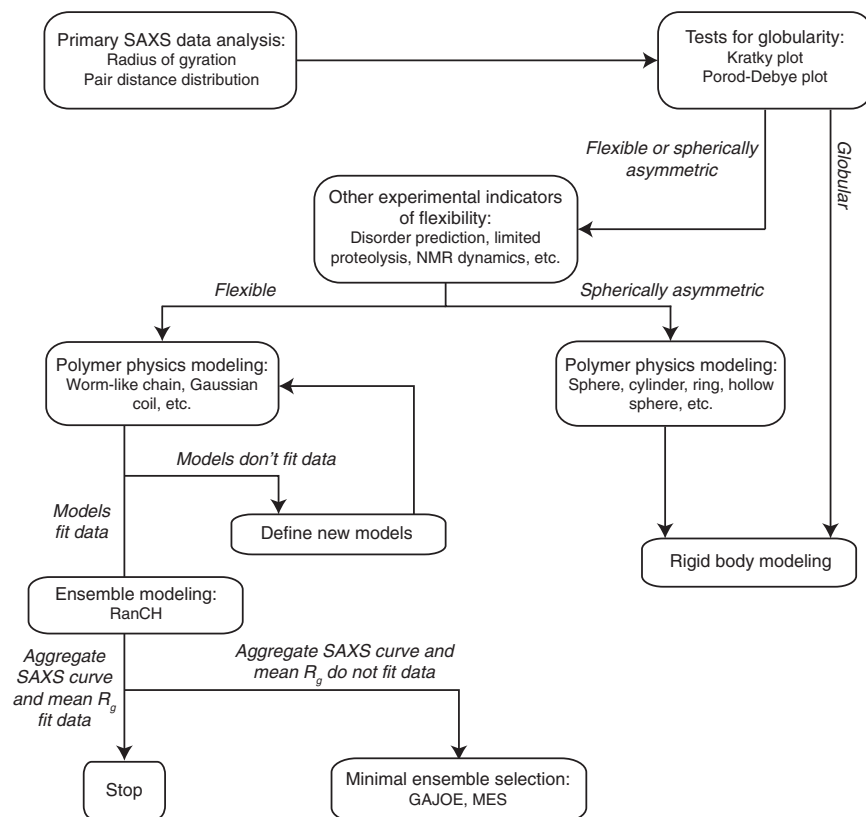


Figure 7. Flowchart of a Proposed Method to Determine and Describe the Statistical Conformation of Highly Flexible Proteins
See also Figure S3.

ensemble.) Following the EOM protocol, we used a modified RanCH algorithm that produces a structurally converged parent ensemble of atomic-resolution structures to generate an unconstrained parent ensemble of self-avoiding conformations. The aggregate scattering profiles from these unconstrained ensembles matched the data very well (Figure S2A), suggesting that no further information may be obtained through fitting. However, we continued the EOM protocol and ran the GAJOE algorithm multiple times to select 22–50 conformations from these parent ensembles (minimal ensembles) whose calculated aggregate SAXS curves fitted the observed data. The resulting distributions of the global structural parameters R_g and D_{max} were very similar to the distributions of the parent ensembles (Figures S2B and S2C). In the case of 2-BdpA, the best-fit minimal ensemble

ensemble of models (typically 10–50) (Bernadó et al., 2007). These methods yield distributions of R_g and D_{max} that are consistent with the SAXS data using conformational sampling (Bernadó and Svergun, 2012a). However, because the SAXS curves from conformationally averaged samples are often relatively featureless, in many cases these distributions are not unique. Indeed, we show below that EOM-based distributions of n-BdpA are not unique. The final approach, employed often by materials scientists and used in this study, but rarely utilized by structural biologists and used in this study, but rarely utilized by structural biologists, describes the conformational space of highly flexible proteins using polymer physics models. This approach eschews the use of atomic resolution models in favor of a continuous statistical chain description. While the polymer physics description of conformational space does not lend itself to pictorial depictions of the protein, it does describe the statistical conformation without discretizing conformational space, so it is not an overparameterized model of the statistical conformation.

In order to compare the description of conformational space resulting from the fit of the EV-PNM to our SAXS data to the description of conformational space resulting from ensemble-based methods, we used this method to analyze the n-BdpA SAXS data. We chose to use the EOM suite of programs to perform the ensemble analysis (Bernadó et al., 2007). EOM consists of two independent programs: RanCH (which generates a large number of protein models based on hard-sphere energy functions and is unconstrained by the data) and GAJOE (which uses a genetic algorithm to select a minimal ensemble that is consistent with the SAXS data from this parent

was bimodally or trimodally partitioned into more compact and more extended conformations, with no similarity between runs. In the case of (3-5)-BdpA, the minimal ensemble distributions were statistically identical to each other and nearly the same as those of the parent ensemble, although there appears to be a slight asymmetry favoring longer R_g and D_{max} in the minimal ensemble distributions, which we attribute primarily to instrument noise that biases the minimal ensemble selection process. In all instances, the fits of the minimal ensembles to the SAXS data were equally good. These results indicate that discrete atomistic models do not fit our SAXS data significantly better than the EV-PNM continuous model. Thus, in this case the models are highly overparameterized and only confirm that SpA is a self-avoiding chain of domains linked via six-residue flexible linkers, which is the conclusion derived from the three parameters of the EV-PNM polymer physics model.

A Protocol for Deciding Whether to Fit SAXS Data from Highly Flexible Proteins to Polymer Physics or Atomistic Ensemble Models

The approach we have applied to SpA-N and n-BdpA may be generally useful as a first conservative step in the interpretation of SAXS data from other multidomain proteins that might have similar flexibility. We propose the following five steps to determine whether an atomistic ensemble model of such systems is appropriate or overfit (Figure 7): (1) determine if the system is globular as indicated by a convergence in the Kratky plot ($q^2 I(q)$ versus q) and the existence of a plateau in the Porod-Debye plot ($q^4 I(q)$ versus q) (Rambo and Tainer, 2011). If either of

these tests fails, then it indicates that the protein may either be flexible or spherically asymmetric. Additional experimental evidence should be used to determine if the protein is globular, spherically asymmetric, or flexible, since SAXS analysis alone cannot distinguish between flexible and spherically asymmetric molecules. (2) If it is determined that the protein is flexible, fit the SAXS data to various polymer physics models: e.g., worm-like chain, flexible cylinder, Gaussian coil, and pearl necklace. These models are included in the standard fitting packages of programs such as SASfit (<https://kur.web.psi.ch/sans1/SANSSoft/sasfit.html>), SasView (<http://www.sasview.org>), or NIST Center for Neutron Research SANS (<http://www.ncnr.nist.gov/programs/sans/>). If no existing polymer model fits the SAXS data, a new model can be defined based on the characteristics of the system. A corresponding scattering function can be derived using an approach similar to the one used here (see the [Supplemental Experimental Procedures](#)). (3) As a comparison, use an atomistic model and sampling of conformational space to obtain an ensemble representing a coarse sampling of the statistical conformation, unconstrained by the SAXS data. We have developed a Python script, RanCHiterate (<https://www.dropbox.com/s/yn2kti2usesia4g/RanCHiterate.py>), that allows one to use RanCH (Bernadó et al., 2007) (part of the ATSAS suite of programs) to generate an ensemble of models where the ensemble size is dependent on the degrees of freedom of the molecule. This script determines the number of models necessary to produce converged R_g and D_{max} distributions and generates an aggregate scattering curve for the unconstrained ensemble. In some cases, differences between this calculated scattering curve and the experimental data would be random and comparable to noise (as they were for the data described in this paper). In this event, subsampling the ensemble would result in an effectively unconstrained minimal ensemble whose members inadequately represent the actual range of conformational space and that yields inaccurate distributions for R_g or D_{max} . (4) If the RanCH-generated ensemble produces an aggregate SAXS profile that does not match the experimental data, use GAJOE to select a minimal set of conformations from the ensemble generated in step 3 whose total predicted scattering curve better fits the observed SAXS data. Test this minimal ensemble for statistical significance: is the R_g distribution determined from the minimal ensemble repeatable using the same starting ensemble? Are the moments of the R_g distribution for each replicate calculation similar? If so, this analysis supports a higher resolution interpretation of the SAXS data, as represented by the minimal ensemble. If not, the SAXS data do not support such a highly parameterized model of the statistical conformation of the flexible system.

The protocol presented in [Figure 7](#) and described above can be used to determine if a polymer physics model adequately accounts for all variation in the statistical conformation observed in the SAXS data or if a more fine-grained ensemble description of the statistical conformation is supported by the data. We have demonstrated that in the case of (3-5)-BdpA and SpA-N that a simple polymer physics model is sufficient to accurately fit the SAXS curve, indicating that the protein is too flexible to contain interdomain structure detectable by SAXS. For systems with less flexibility, like the counterexample described in the [Supplemental Experimental Procedures](#) ([Figure S3](#)), ensemble-based

modeling can produce additional conformational insights that polymer physics cannot provide. However, for SpA-N and other highly flexible systems like it, a polymer physics description of protein conformation is the most accurate reflection of the information content of the SAXS data.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification

Plasmid constructs were transformed into *E. coli* BL21(DE3) cells using standard procedures. 1L LB media containing 100 mg/L ampicillin were then inoculated with a single colony of the transformed cells. The cells were grown at 37°C until an OD_{600} of 0.8–1.0 was reached. The culture was induced with isopropyl β -D-1-thiogalactopyranoside to a final concentration of 1 mM then harvested 4–6 hr post-induction and centrifuged. The cell pellet was resuspended in 50 mM Tris (pH 8.8), 1 mM EDTA, and protease inhibitors (4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride, pepstatin, bestatin, and E-64). The cells were lysed and insoluble material was cleared by centrifugation. The pH of the cleared lysate was adjusted to pH 9.0, and 10 μ l micrococcal nuclease was added to digest large DNA fragments. The resulting solution was brought to 4 M guanidinium HCl and 20 mM Tris(2-carboxyethyl)phosphine (TCEP) by the addition of solid guanidinium HCl (Bio-Basic) and 1 M TCEP. The solution was dialyzed into a 5% acetic acid solution, and insoluble materials were cleared by centrifugation. The soluble material was dialyzed into deionized water. The protein solution was loaded onto an SP Sepharose (GE Healthcare) column in 50 mM sodium acetate (pH 3.6). The protein was eluted from the column by a 600 ml 100–500 mM NaCl gradient in 8 ml fractions. The fractions were checked for purity by SDS-PAGE. The most pure fractions were pooled and dialyzed against deionized water. This protein solution was loaded onto a DEAE Sephacil (GE Healthcare) column in 50 mM sodium acetate (pH 3.6). The protein was eluted from the column by an 800 ml 0–250 mM NaCl gradient in 8 ml fractions. The fractions were checked for purity by SDS-PAGE. The most pure fractions were pooled and dialyzed into deionized water. The protein was then lyophilized and stored in a desiccator.

SAXS Sample Preparation

Lyophilized protein was resuspended in deionized water to make stock solutions. Sodium acetate (pH 5.5), sodium chloride, and glycerol were added to each stock solution to a final concentration of 50 mM sodium acetate (pH 5.5), 100 mM sodium chloride, and 1% glycerol. The protein samples were then dialyzed against 50 mM sodium acetate (pH 5.5), 100 mM sodium chloride, and 1% glycerol for 6 hr at room temperature using a 3,500 Da molecular weight cutoff microdialysis unit (Pierce).

For Advanced Light Source data collection, samples were centrifuged at 16,000 $\times g$ for 20 min, and then the concentration of each sample was calculated by A_{280} . Samples were diluted to a concentration of 5 mg/ml, 2.5 mg/ml, or 1.25 mg/ml using dialysate. The samples were stored at 4°C for no more than 24 hr.

For APS data collection, each sample was stored at 4°C for no more than 36 hr prior to data collection. Just prior to data collection each sample was centrifuged at 16,000 $\times g$ for 20 min and the concentration was calculated by A_{280} . Samples were diluted to a final concentration of 2 mg/ml, 1 mg/ml, or 0.5 mg/ml using dialysate.

SAXS Data Acquisition and Analysis

Data were collected at beamline 12.3.1 (SIBYLS) at the Advanced Light Source, Lawrence Berkeley National Laboratory, and at beamline 18-ID (Bio-CAT) at the Advanced Photon Source, Argonne National Laboratory.

At the SIBYLS beamline, 25 μ l of protein samples were loaded into a sample cell and then exposed for 0.5, 1, or 4 s at an energy of 12 keV, with a sample-to-detector distance of 1.5 M, corresponding to a q range of 0.01–0.32 \AA^{-1} . Data were collected from an identical buffer sample, using dialysate from the equilibrium dialysis, for each protein sample using identical data collection conditions. All data were collected at 10°C. Beamline specific software was used to reduce the data and subtract the buffer signal to generate final scattering data for each protein sample.

At the Bio-CAT beamline, 120 μ l protein samples were loaded into a sample capillary and the sample was oscillated in the beam to minimize radiation damage, such that no single protein molecule was exposed for more than 100 ms. Data were collected at an energy of 12 keV, and a sample-to-detector distance of 2,750 mm, corresponding to a q range of 0.008–0.29 \AA^{-1} . The flux of the beam was attenuated by using 18 foil attenuators. Data from identical buffer samples from the dialysate were collected for each protein sample. All data were collected at 10°C. Data were reduced using the Nika package for Igor Pro (<http://usaxs.xray.aps.anl.gov/staff/ilavsky/nika.html>). Fifteen individual data sets for each protein and buffer sample were averaged in PRIMUS (Konarev et al., 2003), and the buffer signal was subtracted from the data signal using PRIMUS to generate final scattering data for each protein sample.

Guinier analysis for each construct was performed using PRIMUS to determine the radius of gyration and $I(0)$. Polymer models were fit to the scattering data using a nonlinear least-squares fitting algorithm implemented in Mathematica 9. The 95% confidence intervals and standard errors were calculated in Mathematica 9. The 95.4% confidence intervals were calculated according to the method of Bevington and Robinson (2003).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, and two tables and can be found with this article online at <http://dx.doi.org/10.1016/j.str.2014.06.011>.

ACKNOWLEDGMENTS

We wish to thank Prof. Jane Richardson, Yang Qi, Lindsay Deis, and Prof. Gerhard Findenegg for their helpful discussions and Prof. David Goldenberg for many useful suggestions. This work was supported by PHS/NIH grant R01-GM081666 to T.G.O. and R01-GM073930 to D.C.R.

A portion of this work was conducted at the SIBYLS beamline at the Advanced Light Source (ALS), a national user facility operated by Lawrence Berkeley National Laboratory on behalf of the Department of Energy, Office of Basic Energy Sciences, through the Integrated Diffraction Analysis Technologies (IDAT) program.

A portion of this work was performed at beamline 18-ID of the Advanced Photon Source. Use of the Advanced Photon Source, an Office of Science User Facility operated for the U.S. Department of Energy (DOE) Office of Science by Argonne National Laboratory, was supported by the U.S. DOE under Contract No. DE-AC02-06CH11357.

Received: January 3, 2014

Revised: June 7, 2014

Accepted: June 11, 2014

Published: July 31, 2014

REFERENCES

- Bernadó, P., and Svergun, D.I. (2012a). Analysis of intrinsically disordered proteins by small-angle X-ray scattering. *Methods Mol. Biol.* **896**, 107–122.
- Bernadó, P., and Svergun, D.I. (2012b). Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. *Mol. Biosyst.* **8**, 151–167.
- Bernadó, P., Mylonas, E., Petoukhov, M.V., Blackledge, M., and Svergun, D.I. (2007). Structural characterization of flexible proteins using small-angle X-ray scattering. *J. Am. Chem. Soc.* **129**, 5656–5664.
- Bevington, P.R., and Robinson, D.K. (2003). *Data reduction and error analysis for the physical sciences*, Third Edition. (Boston: McGraw-Hill).
- Chi, Q., Wang, G., and Jiang, J. (2013). The persistence length and length per base of single-stranded DNA obtained from fluorescence correlation spectroscopy measurements using mean field theory. *Physica A* **392**, 1072–1079.
- Deisenhofer, J. (1981). Crystallographic refinement and atomic models of a human F_c fragment and its complex with fragment B of protein A from *Staphylococcus aureus* at 2.9- and 2.8- \AA resolution. *Biochemistry* **20**, 2361–2370.
- Dobrynin, A.V., Rubinstein, M., and Obukhov, S.P. (1996). Cascade of transitions of polyelectrolytes in poor solvents. *Macromolecules* **29**, 2974–2979.
- Flory, P.J. (1953). *Principles of polymer chemistry*. (Ithaca: Cornell University Press).
- Fujita, H. (1990). *Polymer solutions*. (Amsterdam: Elsevier).
- Gómez, M.I., Lee, A., Reddy, B., Muir, A., Soong, G., Pitt, A., Cheung, A., and Prince, A. (2004). *Staphylococcus aureus* protein A induces airway epithelial inflammatory responses by activating TNFR1. *Nat. Med.* **10**, 842–848.
- Graille, M., Stura, E.A., Corper, A.L., Sutton, B.J., Taussig, M.J., Charbonnier, J.B., and Silverman, G.J. (2000). Crystal structure of a *Staphylococcus aureus* protein A domain complexed with the Fab fragment of a human IgM antibody: structural basis for recognition of B-cell receptors and superantigen activity. *Proc. Natl. Acad. Sci. USA* **97**, 5399–5404.
- Hagerman, P.J. (1988). Flexibility of DNA. *Annu. Rev. Biophys. Biophys. Chem.* **17**, 265–286.
- Hammouda, B. (1993). SANS from homogenous polymer mixtures: A unified overview. *Adv. Polym. Sci.* **106**, 87–133.
- Hartleib, J., Köhler, N., Dickinson, R.B., Chhatwal, G.S., Sixma, J.J., Hartford, O.M., Foster, T.J., Peters, G., Kehrel, B.E., and Herrmann, M. (2000). Protein A is the von Willebrand factor binding protein on *Staphylococcus aureus*. *Blood* **96**, 2149–2156.
- Hjelm, R.P. (1985). The small-angle approximation of x-ray and neutron scatter from rigid rods of non-uniform cross-section and finite length. *J. Appl. Crystallogr.* **18**, 452–460.
- Jacques, D.A., Guss, J.M., Svergun, D.I., and Trehwella, J. (2012). Publication guidelines for structural modelling of small-angle scattering data from biomolecules in solution. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 620–626.
- Janin, J., and Sternberg, M.J. (2013). Protein flexibility, not disorder, is intrinsic to molecular recognition. *F1000 Biol. Rep.* **5**, 2.
- Koch, M.H., Vachette, P., and Svergun, D.I. (2003). Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. *Q. Rev. Biophys.* **36**, 147–227.
- Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J., and Svergun, D.I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **36**, 1277–1282.
- Lairez, D., Pauthe, E., and Pelta, J. (2003). Refolding of a high molecular weight protein: salt effect on collapse. *Biophys. J.* **84**, 3904–3916.
- Löfdahl, S., Guss, B., Uhlén, M., Philipson, L., and Lindberg, M. (1983). Gene for staphylococcal protein A. *Proc. Natl. Acad. Sci. USA* **80**, 697–701.
- Merino, N., Toledo-Arana, A., Vergara-Irigaray, M., Valle, J., Solano, C., Calvo, E., Lopez, J.A., Foster, T.J., Penadés, J.R., and Lasa, I. (2009). Protein A-mediated multicellular behavior in *Staphylococcus aureus*. *J. Bacteriol.* **191**, 832–843.
- Moks, T., Abrahmsén, L., Nilsson, B., Hellman, U., Sjöquist, J., and Uhlén, M. (1986). Staphylococcal protein A consists of five IgG-binding domains. *Eur. J. Biochem.* **156**, 637–643.
- Nguyen, T., Ghebrehewet, B., and Peerschke, E.I. (2000). *Staphylococcus aureus* protein A recognizes platelet $gC1qR/p33$: a novel mechanism for staphylococcal interactions with platelets. *Infect. Immun.* **68**, 2061–2068.
- Palmqvist, N., Foster, T., Tarkowski, A., and Josefsson, E. (2002). Protein A is a virulence factor in *Staphylococcus aureus* arthritis and septic death. *Microb. Pathog.* **33**, 239–249.
- Pelta, J., Berry, H., Fadda, G.C., Pauthe, E., and Lairez, D. (2000). Statistical conformation of human plasma fibronectin. *Biochemistry* **39**, 5146–5154.
- Putnam, C.D., Hammel, M., Hura, G.L., and Tainer, J.A. (2007). X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q. Rev. Biophys.* **40**, 191–285.
- Rambo, R.P., and Tainer, J.A. (2011). Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers* **95**, 559–571.
- Rambo, R.P., and Tainer, J.A. (2013). Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* **496**, 477–481.

- Rubinstein, M., and Colby, R.H. (2003). *Polymer physics*. (Oxford: Oxford University Press).
- Schneewind, O., Fowler, A., and Faull, K.F. (1995). Structure of the cell wall anchor of surface proteins in *Staphylococcus aureus*. *Science* 268, 103–106.
- Schweins, R., and Huber, K. (2004). Particle scattering factor of pearl necklace chains. *Macromol. Symp.* 211, 25–42.
- Sjöholm, I., Ekenäs, A.K., and Sjöquist, J. (1972). Protein A from *Staphylococcus aureus*. Acetylation of protein A with acetylimidazole. *Eur. J. Biochem.* 29, 455–460.
- Starovasnik, M.A., Skelton, N.J., O'Connell, M.P., Kelley, R.F., Reilly, D., and Fairbrother, W.J. (1996). Solution structure of the E-domain of staphylococcal protein A. *Biochemistry* 35, 15558–15569.
- Teilum, K., Olsen, J.G., and Kragelund, B.B. (2009). Functional aspects of protein flexibility. *Cell. Mol. Life Sci.* 66, 2231–2247.
- Tokuriki, N., and Tawfik, D.S. (2009). Protein dynamism and evolvability. *Science* 324, 203–207.
- Wignall, G.D., Schelten, J., and Ballard, D.G.H. (1974). Measurements of radius of gyration and persistence length in bulk atactic polystyrene by low-angle neutron-scattering. *J. Appl. Crystallogr.* 7, 190.
- Zheng, D., Aramini, J.M., and Montelione, G.T. (2004). Validation of helical tilt angles in the solution NMR structure of the Z domain of Staphylococcal protein A by combined analysis of residual dipolar coupling and NOE data. *Protein Sci.* 13, 549–554.