Small-Angle Neutron Scattering Study of a Monoclonal Antibody Using Free-Energy Constraints

Nicholas J. Clark,[†] Hailiang Zhang,[‡] Susan Krueger,[†] Hyo Jin Lee,[§] Randal R. Ketchem,[‖] Bruce Kerwin,[§] Sekhar R. Kanapuram,[§] Michael J. Treuheit,[‖] Arnold McAuley,[§] and Joseph E. Curtis*,[†]

[†]NIST Center for Neutron Research, National Institute of Standards and Technology, 100 Bureau Drive, Mail Stop 6102, Gaithersburg, Maryland 20899, United States

[‡]Institute for Research and Applied Physics, University of Maryland, College Park, Maryland 20742, United States

[§]Department of Drug Product Development, Amgen Inc., One Amgen Center Drive, Thousand Oaks, California 91320, United States

[‖]Therapeutic Discovery and Department of Drug Product Development, Amgen Inc., 1201 Amgen Court West, Seattle, Washington 98119, United States

ABSTRACT: Monoclonal antibodies (mAbs) contain hinge-like regions that enable structural flexibility of globular domains that have a direct effect on biological function. A subclass of mAbs, IgG2, have several interchain disulfide bonds in the hinge region that could potentially limit structural flexibility of the globular domains and affect the overall configuration space available to the mAb. We have characterized human IgG2 mAb in solution via small-angle neutron scattering (SANS) and interpreted the scattering data using atomistic models. Molecular Monte Carlo combined with molecular dynamics simulations of a model mAb indicate that a wide range of structural configurations are plausible, spanning radius of gyration values from ~39 to ~55 Å. Structural ensembles and representative single structure solutions were derived by comparison of theoretical SANS profiles of mAb models to experimental SANS data. Additionally, molecular mechanical and solvation free-energy calculations were carried out on the ensemble of best-fitting mAb structures. The results of this study indicate that low-resolution techniques like small-angle scattering combined with atomistic molecular simulations with free-energy analysis may be helpful to determine the types of intramolecular interactions that influence function and could lead to deleterious changes to mAb structure. This methodology will be useful to analyze small-angle scattering data of many macromolecular systems.
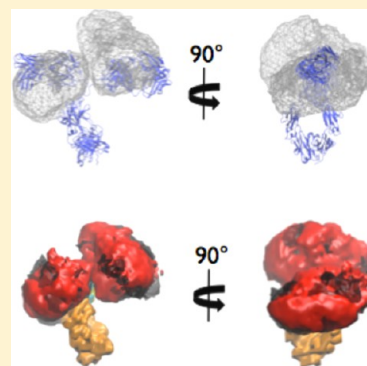
INTRODUCTION

Monoclonal antibodies (mAbs) are vital proteins that make up a major component of vertebrate immunity and are a class of protein therapeutics being developed commercially to treat a large number of diseases. The physics of mAbs is interesting due to structural and dynamical features that affect function and retention of biological activity. mAbs are glycoproteins with intrinsic domain flexibility[1−5] that has been correlated with function.[6] A predominant subclass of mAb proteins, known as immunoglobulins (IgG), have molecular masses of ~150 000 Da and are composed of two identical light chains (LC) and two identical heavy chains (HC). These chains are held together via covalent inter- and intrachain disulfide bonds and stabilized by intermolecular non-bonding interactions to form two Fab domains and one Fc domain. The Fab and Fc domains are separated by a short polypeptide hinge region of 23−24 amino acids in the two HC.[6] A cartoon representation of mAb structure is shown in Figure 1. Mobility of Fab and Fc domains has a direct effect on the dual function of mAbs.[6] The Fab domains contain regions of amino acids, named Fv, that are responsible for binding antigens. The mobility of Fab domains is required for antigen recognition and binding. Mobility of the Fc domain is responsible for modulating effector functions that

potentially lead to the elimination of foreign bodies by the immune system. IgG's have subtle differences in amino acid sequence that result in different structural, dynamical, and functional characteristics. IgG1 and IgG2 are two common subclasses of IgG mAbs. In their respective hinge regions, the position of a single interchain disulfide bond in IgG1 mAbs results in a more rigid structure,[7] although significant local and domain flexibility exists.[6−8] Removal[9] and genetic mutation[5] of the hinge region can eliminate and modulate function. Complete three-dimensional structures of mAbs have been determined by X-ray crystallography for both IgG1[7,9−11] and IgG2[12] subclasses. Regardless, structural heterogeneity has been observed by a variety of experimental techniques[1−5,8] and therefore crystal structures are thought to be single structures that do not represent the large number of configurations that exist in solution.[6,7,9,11,12] The hinge region is considered a tether that connects Fab and Fc domains to allow the motions required for function.
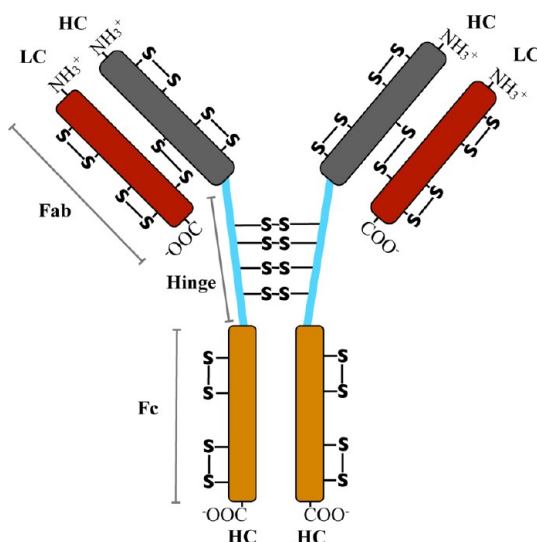
Figure 1. Schematic representation of an immunoglobulin G (IgG) subclass 2, isotype A protein. IgG is composed of two light chains (LC) and two heavy chains (HC) that are held together through several inter- and intrachain disulfide bonds (S—S) and non-bonding interactions. The LC and part of the HC form two identical Fab domains. Part of each HC also forms a single FC domain. The Fab and Fc domains are connected via a flexible hinge region (blue) that itself has four interchain disulfide bonds.

Small-angle neutron scattering (SANS) probes structure and interactions on length scales from 10 Å to greater than 1000 Å, making it a well-suited technique for the study of the low-resolution shape of proteins in solution. Small-angle X-ray scattering (SAXS) and SANS have been used to study the structure of various mAbs in solution.[13−17] Given their size and inherent flexibility, there are few computational approaches to exhaustively sample configuration space using atomistic models to generate ensembles to compare to experimental SAS data. Regardless, there are successful studies in the literature.[14−16] Recently developed algorithms and programs developed to study intrinsically disordered proteins can be applied to study SAS data of mAbs.[18−21] Several of these programs use heuristic algorithms to reduce the number of structures that represent the experimental data using linear combinations of scattering profiles. They do not take into account intermolecular interactions in either the generation or the evaluation of configurations in the ensemble.[18,22,23] These methods are valuable and widely used because they provide a low barrier to rapidly generate and/or evaluate models by comparison of theoretical scattering data to experimental data. Regardless, there is a need to develop new methods to incorporate more rigorous computational physical chemistry algorithms to generate and evaluate ensembles of structures by comparison to scattering data.

We have carried out SANS on a monodisperse solution of a human antistreptavidin monoclonal antibody class 2 molecule (ASA-IgG2). Experimental scattering profiles were analyzed by comparison to theoretical data taken from the atomic coordinates of sequential Monte Carlo (MC) and molecular dynamics (MD) simulations. The results are refined further by analyzing the ensemble via molecular mechanical and solvation free energies. The addition of energetic analysis to evaluate molecular models derived by comparison to scattering data could be helpful to determine relevant structures for a large number of soft-matter systems.

MATERIALS AND METHODS

Protein Preparation and Characterization. Concentrated stock solutions of ASA-IgG2 at 30 mg/mL were thawed at room temperature and transferred to Slide-a-lizer 10 000 Da molecular weight cutoff dialysis cassettes (Pierce) and equilibrated in 10 mM sodium acetate (pH 5.2) in D_2O overnight. Certain commercial equipment, instruments, materials, suppliers, or software are identified in this paper to foster understanding. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose. Buffers were filtered using Millipore 0.22 μm filter membranes. Samples were then centrifuged at 16 000g for 10 min prior to the absorbance measurement. A dilution series consisting of protein diluted into dialysis buffer was made in triplicate using the following dilution range: 1:30, 1:20, 1:10. Concentration was determined using the following equation: concentration (mg/mL) = $(A_{280} - A_{350})/\varepsilon$, where A_{280} is the absorbance at 280 nm and A_{350} is the absorbance at 350 nm, using an extinction coefficient, ε, of 1.6 mL mg^{-1} cm^{-1} at 280 nm for the ASA-IgG2 protein. Absorbance was measured using a Nanodrop spectrophotometer (Thermo).

Size-exclusion high performance liquid chromatography (SE-HPLC) was performed using an Agilent 1100 (Palo Alto, CA, USA) equipped with a quaternary pump, diode array detector, and refrigerated autosampler. Two Tosoh G3000SW TSK columns, 5 μm particle size, 7.8 mM ID × 30 cm (Montgomeryville, PA, USA), were used in series. 50 μg of sample was injected and eluted over 30 min using 25 mM sodium phosphate, 125 mM sodium chloride, pH 6.9, as the running buffer at 0.6 mL/min. Absorbance was monitored at 215 nm and at 280 nm. Data was collected using Agilent Chemstation software. ASA-IgG2 was studied by SE-HPLC at 1.0, 2.0, and 5.0 mg/mL at pH 5.2 in 20 mM sodium acetate ± 0.15 M NaCl. Post-SEC analysis, samples were analyzed by an online Dawn Heleos II (Wyatt Technologies) multiangle light scattering instrument followed by an REx refractive index detector (Wyatt Technologies). A dn/dc value of 0.185 was used to determine the concentration. Data analysis was carried out with Astra Software version 5.3.4. Reported values are the mass averaged molecular weights.

Small-Angle Neutron Scattering. Scattering measurements were performed on the 30 m SANS instruments[24] at the NIST Center for Neutron Research (NCNR) in Gaithersburg, MD. The neutron wavelength, λ, was 6 Å, with a wavelength spread, $\Delta\lambda/\lambda$, of 0.15. Scattered neutrons were detected with a 64 cm × 64 cm two-dimensional position-sensitive detector with 128 × 128 pixels at a resolution of 0.5 cm/pixel. The data were reduced using the IGOR program with SANS macro routines developed at the NCNR.[25] Raw counts were normalized to a common monitor count and corrected for empty cell counts, ambient room background counts, and nonuniform detector response. Data from the samples were placed on an absolute scale by normalizing the scattered intensity to the incident beam flux. Finally, the data were radially averaged to produce scattered intensity, $I(q)$, versus q curves, where $q = 4\pi \sin(\theta)/\lambda$ and 2θ is the scattering angle. Scattering was measured using sample-to-detector distances of 13.0, 5.0, and 1.5 m in order to cover the range 0.007 Å$^{-1} \leq q \leq 0.3$ Å^{-1}. The scattered intensities from the samples in the liquid state were corrected for buffer scattering and incoherent
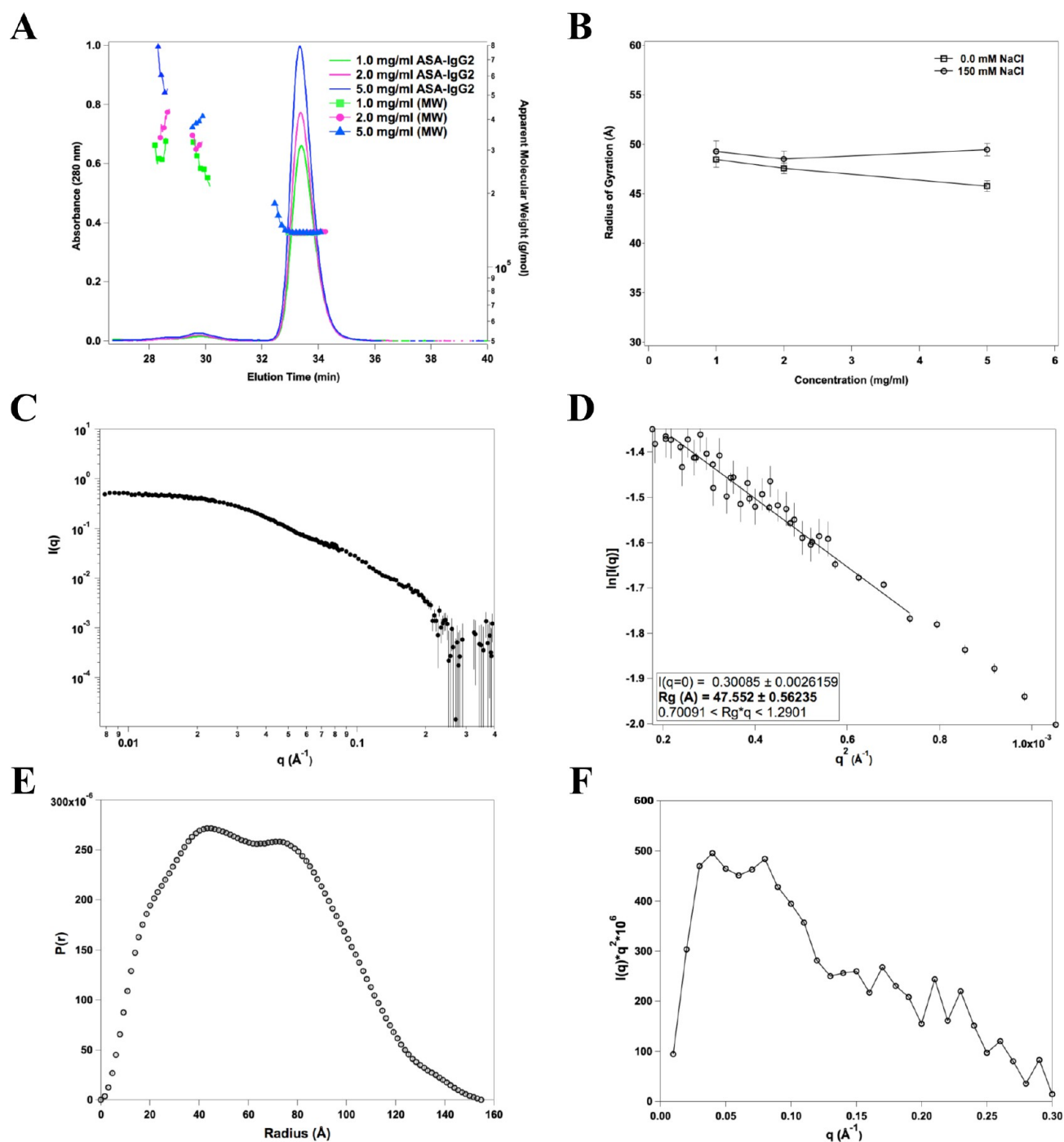
Figure 2. Small-angle scattering studies of anti-streptavidin IgG class 2 (ASA-IgG2). (A) Size-exclusion chromatography and multiangle light scattering of ASA-IgG2. (B) Plot of radius of gyration (R_g) versus concentration with and without the addition of 150 mM NaCl. (C) Small-angle neutron scattering (SANS) profile of dilute ASA-IgG2. (D) Guinier analysis of the 2 mg/mL SANS data results with an R_g of 47.5 Å for the molecule. (E) Plot of pairwise distribution ($P(r)$) for the ASA-IgG2 molecule. The molecule has an average maximum dimension (D_{max}) of 150 Å. Additionally, the apparent peak at ~43 Å indicates the presence of a subdomain or flexible region of roughly 42 Å in length (roughly the size of the Fab arms). (F) Kratky plot of ASA-IgG2. Error bars represent ±1 standard deviation.

scattering from hydrogen in the samples. Pair distribution analysis was performed using GNOM.[26] In order to optimize $P(r)$, several input Dmax values were used (ranging from 150 to 160 Å) to find the best-fit Dmax value of 155 Å. Guinier analysis[27] was carried out using $I(Q)/I(0) \approx \exp[-q^2R_g^2/3]$ to estimate the radius of gyration (R_g).

Atomistic Modeling. *Homology Modeling.* The initial homology model of ASA-IgG2 mAb was generated using the

Molecular Operating Environment from the Chemical Computing Group[28] using the Antibody Modeler tool to build the ASA-IgG2 Fv coordinates. This structure was then used in the Homology Model tool in MOE in conjunction with Protein Data Bank[29] entry 1IGT[12] to build the remainder of the ASA-IgG2 model. Disulfide bonding was chosen to represent the IgG2 type A isotype.[30,31] The homology model of ASA-IgG2, that contained 19 668 atoms, was energy minimized for

10 000 steps with the program NAMD[32] using the CHARMM-22 force field.[33] Subsequently, the structure was immersed in a previously equilibrated 200 Å cubic box of water (using the TIP3P water model[34]) and overlapping waters were removed and a neutralizing number of ions were added. The system was equilibrated at 300 K and 1 bar for 1 ns in the isothermal–isobaric ensemble. ASA-IgG2 coordinates were then used as the starting structure for further molecular simulations and are depicted as structure number 4 in Figures 3B and 5 and illustrated in Figure 4A, B, and D.

Molecular Simulation and Analysis Protocol. The equilibrated coordinates of ASA-IgG2 described in the previous section were used as the starting structure for molecular MC studies. The program SASSIE[21] was used to generate 56 511 nonoverlapping configurations by sampling backbone dihedral angles, ϕ or ψ, of three amino acids on each HC in the upper hinge of ASA-IgG2. Energetics of the specific dihedral angle to sample configurations was derived from the energy of a given ϕ or ψ angle that was calculated from the specific atomic (and thereby amino acid residue specific) composition about the given angle. The energy term was calculated from $V_{\text{dihedral}} = V(\theta) = k_\theta(1.0 + \cos(n\theta - \delta))$, where the angular force constant (k_θ), multiplicity (n), and δ are values from the CHARMM 22 all-atom protein force field and are specific for the atom types for the given dihedral angle of interest $(\theta = \phi$ or $\psi)$. In addition, non-bonded terms were included in the potential $U = V_{\text{dihedral}} + V_{\text{vdW}} + V_{\text{elec}}$ and calculated using CHARMM 22 parameters. For well-depths ε_{ij} and radii σ_{ij} for pairs of atoms i and j, the van der Waals potential energies were calculated using

$$V_{\text{vdW}} = 4\varepsilon_{ij} \sum_{i<j} \left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6} \right] \qquad (1)$$

V_{vdW} energies were smoothed to zero using a polynomial function for distances between 10 and 12 Å. Using atomic charges q_i, q_j, relative permittivity, ε_r, and a Debye screening length, L, the electrostatic potential energies were calculated using

$$V_{\text{elec}} = \sum_{i<j} \frac{q_i q_j}{\varepsilon_r r_{ij}} \exp(-r_{ij}/L) \qquad (2)$$

with a screening length of 25 Å. The full potential, U, was used to carry out the Metropolis sampling methodology at 300 K.

Following MC sampling, each configuration was energy minimized for 10 000 steps followed by 10 ps of generalized Born implicit solvent MD simulation[35] to relax each structure prior to a final 10 000 steps of energy minimization. SANS profiles were calculated from the atomic coordinates using Xtal2sas.[36,37] Comparisons of experimental to theoretical SANS profiles were done using reduced χ^2 calculated using

$$\chi^2 = \frac{1}{(N-1)} \frac{\sum Q(I_{\text{exp}}(Q) - I_{\text{calculated}}(Q))^2}{\sigma_{\text{exp}}(Q)^2} \qquad (3)$$

where $I_{\text{exp}}(Q)$ is the experimentally determined SANS profile, $I_{\text{calculated}}(Q)$ is the profile obtained using Xtal2Sas, $\sigma_{\text{exp}}(Q)$ is the experimentally determined Q-dependent variance, and the sum was taken over $N = 50$ grid points of momentum transfer, Q.

Molecular mechanical energies (ME) were taken from the final output of energy minimization. Polar solvation free-energy (ΔG_S) calculations were carried out using the adaptive

Poisson–Boltzmann solver (APBS) software.[38] APBS calculations for all configurations were carried out using the same grid size. An ionic radius of 1.62 Å at a concentration of 0.01 M wa s used to simulate experimental conditions. The protonation state of each ASA-IgG2 was set to model pH 5.2 using the programs pdb2pqr[39,40] and propka.[41] Energy minimization, MC samping, MD simulation, SANS profile calculations, comparisons to experimental data, creation of iso-density plots, and APBS calculations were carried out within SASSIE. Iso-density plot visualization was carried out using VMD.[42]

RESULTS AND DISCUSSION

Characterization and SANS of ASA-IgG2. Size-exclusion chromatography multiangle light scattering (SEC-MALS)

Table 1. Comparison of the Difference in Mean Square Atomic Displacement before and after 10 000 Steps of Energy Minimization for 1160 Structures with $\chi^2 < 2^a$

region	average	σ	min	max
all atoms	2.47×10^{-3}	3.29×10^{-4}	1.98×10^{-3}	4.9×10^{-3}
flexible atoms	0.129	3.31×10^{-2}	5.75×10^{-2}	0.237

^{a}Units for all values are Å^2. Statistical uncertainty is noted by the standard deviation σ.

enables the evaluation of solution polydispersity. This is accomplished by separating different molecular weight species according to their size and shape using size-exclusion chromatography followed by light scattering and refractive index measurements. Post-separation, the scattering signal and the difference in refractive index (dRI) are evaluated for each eluting species. The resulting scattering signal and dRI are combined to obtain the absolute molecular weight of each species. ASA-IgG2 was evaluated for the presence of higher molecular weight (MW) species and overall monodispersity by SEC-MALS (Figure 2A). Three protein concentrations and two salt conditions were analyzed in order to detect the presence of possible concentration dependent aggregates. At the highest concentration measured, ASA-IgG2 exists as 97.7% monomer in 0.0 mM NaCl and 97.8% monomer in a buffer containing 150.0 mM NaCl. There were no aggregates larger than ~400 kDa detected. At 5.0 mg/mL ASA-IgG2, larger molecular weight species (MW values >150 kDa) contributed 2.3 and 2.2% of the total protein in solution for the 0.0 and 150.0 mM NaCl samples, respectively. SEC-MALS results for only the 0 M NaCl are shown.

The average solution conformation of ASA-IgG2 was evaluated using SANS. In order to rule out concentration dependent interference due to interacting mAb molecules, ASA-IgG2 was evaluated in the same protein concentration and buffer conditions as used in the SEC-MALS experiments (see Figure 2B). SANS scattering profiles for each of the six conditions were fit using the Guinier approximation.[27] The Guinier plot derived radius of gyration for each protein concentration and NaCl condition are plotted in Figure 2B. While there is evidence of interparticle interference at 5.0 mg/mL ASA-IgG2, noted by decreasing R_g with increasing protein concentration, the value of R_g at 2.0 mg/mL was the same for both the high- and low-salt concentrations. This indicates that at 2.0 mg/mL ASA-IgG2 represents an infinitely dilute noninteracting system. Thus, 2.0 mg/mL ASA-IgG2 with 0 M NaCl SANS data shown in Figure 2C was used in subsequent
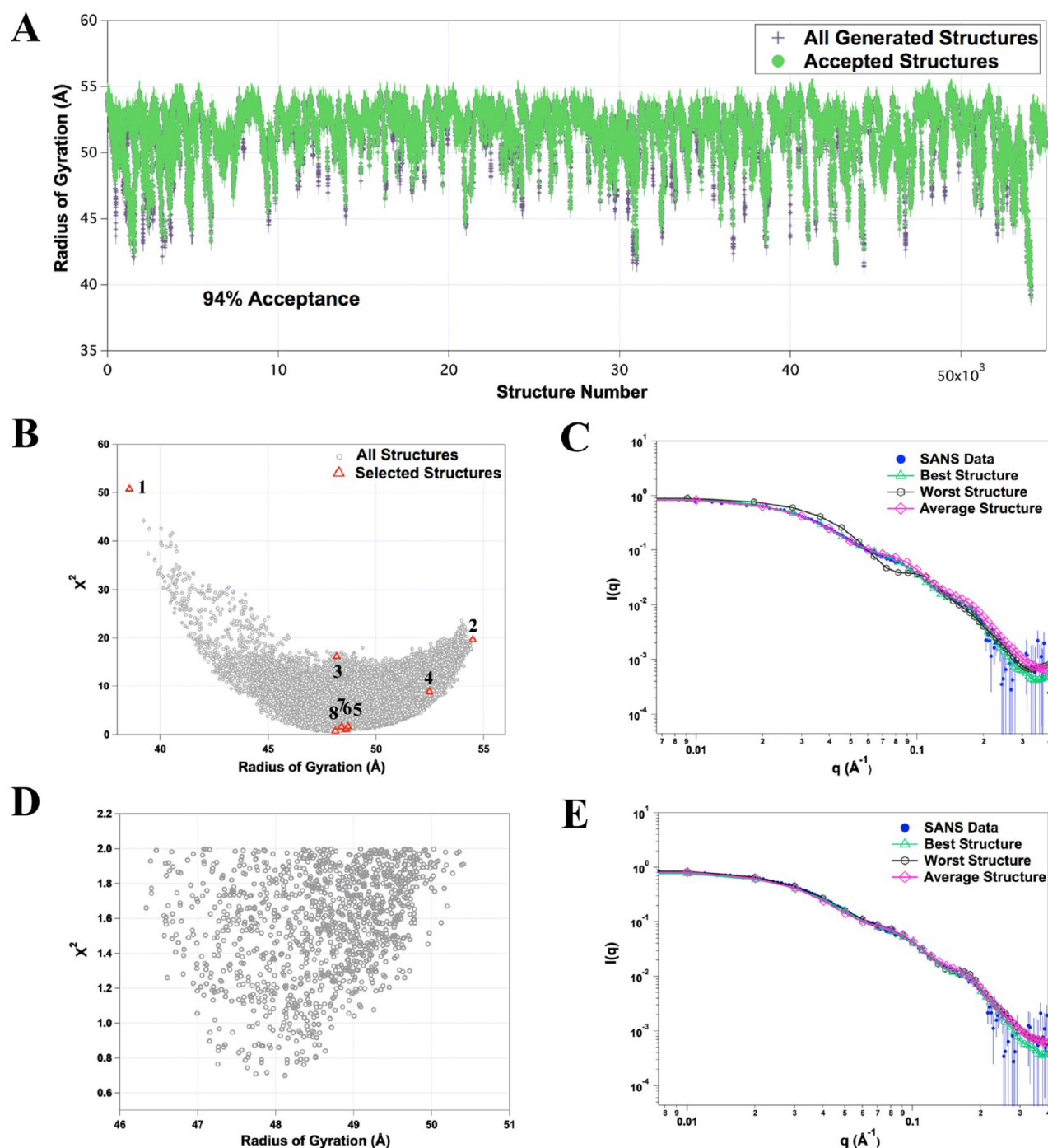
Figure 3. Molecular Monte Carlo modeling of ASA-IgG2 and comparison to SANS Data. (A) R_g values for each of 56 511 configurations of ASA-IgG2 generated by molecular Monte Carlo simulation are plotted versus their structure number. The SANS profile for each configuration was calculated using the programs xtal2sas. The scattering profiles were then compared to experimental data, and the χ^2 value for each configuration was determined. Plot of χ^2 versus R_g and SANS profiles (single best, single worst, and average structure for the entire ensemble (B and C) and for the subensemble of structures with χ^2 values <2.0 (D and E)). In part B, representative structures discussed in the text and in Figure 5 are itemized with enumerated red triangles. Error bars represent ±1 standard deviation.

analysis. The Guinier plot is shown in Figure 2D. The radius of gyration of ASA-IgG2 is 47.5 ± 0.6 Å. The pair distribution, $P(r)$, plot of ASA-IgG2 is shown in Figure 2E. The maximum dimension of the average ASA-IgG2 molecule is ~155 Å. The resulting real-space R_g value was ~48 Å and is in agreement with that calculated using the Guinier approximation. Note that there are two major peaks within the $P(r)$ plot. The peak at ~43 Å is consistent with distribution of scattering centers from

within the Fab regions, while the second peak at ~80 Å most likely represents the distribution of scattering centers from the entire antibody. A Kratky plot, shown in Figure 2F, indicates that ASA-IgG2 has the characteristics of a globular protein despite its inherent flexibility. This has been previously noted for antibodies.[13]

Molecular Simulation of ASA-IgG2. MC simulation using reduced degrees of freedom samples a rugged energy surface.
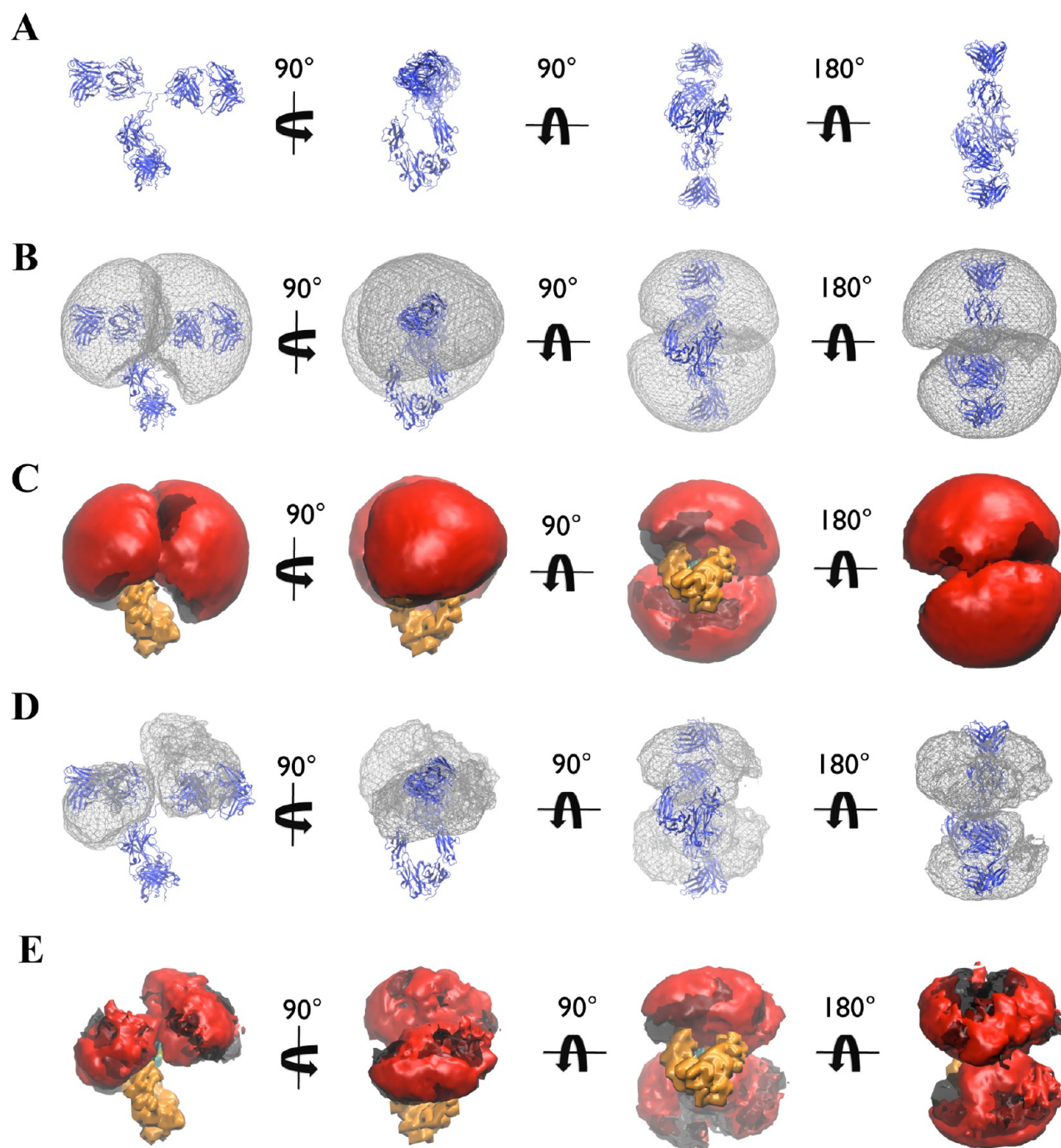
Figure 4. Iso-density plots of ASA-IgG2 from molecular Monte Carlo simulation. (A) Starting structure. (B) Mesh representation and (C) solid surface representation of all structures. (D) Mesh representation and (E) solid surface representation for structures with χ^2 values <2.0. Structures were aligned on their Fc region (brown) in order to demonstrate the configuration space covered by the flexible Fab arms (Fab light chain, red; Fab heavy chain, gray).

Therefore, energy minimization and relaxation using MD simulation is necessary to relax degrees of freedom (bond, angle, and improper) that were held fixed in the MC simulation to adequately represent the physics of the ensemble. Regardless, as shown in Table 1, the structure of each configuration overall and within the flexible upper-hinge region does not change dramatically upon the initial energy minimization. MC simulation using a maximum stepwise angular variation of $\delta\theta \sim 10°$ used in the dihedral sampling yielded a large acceptance percentage covering a large range of R_g values, as shown in Figure 3A. Theoretical SANS profiles for

each structure were calculated, and the individual χ^2 determined from eq 3 as a function of R_g is shown in Figure 3B. Even though the upper hinge of ASA-IgG2 has only three amino acids that can be used to vary the relative configuration of Fab to Fc domains, the R_g values spanned from ~39 to ~55 Å. A comparison of the SANS profiles of the best and worst single structures as well as the average SANS profile for the entire ensemble is shown in Figure 3C. A subset of structures with χ^2 values <2 (Figure 3D) compares favorably with the experimental data, as shown in Figure 3E.

Figure 5. Comparison of structures as a function of χ^2 and R_g. Numbering is based on the itemized and enumerated structures in Figure 3B.

Structure Number	Representation	χ^2	R_g (Å)
1		50.76	38.55
2		19.65	54.49
3		15.76	48.18
4		8.85	52.47
5		1.67	48.70
6		1.62	48.38
7		1.10	48.62
8		0.69	48.11

Figure 6. Thermodynamic cycle for relative configurational free-energy calculations. White boxes represent configurations in a vacuum, and blue shaded boxes represent configurations in water. Numbers depicted between states are described in the text.

The relative arrangement of Fab and Fc domains in mAbs has been correlated to the functional behavior in solution. However, the configuration space available to the domains has not been studied extensively. To determine the extent of configuration space sampled, the ensemble was analyzed using iso-density plots. These plots are generated by averaging the occupancy of ASA-IgG2 backbone alpha-carbon atoms on a three-dimensional grid. The intensity of each voxel is determined by the sum of backbone alpha-carbon atoms found in that voxel normalized by the maximum occupancy over the entire ensemble in order to normalize the data to a common reference. As shown in Figure 4, the range of configuration space sampled by the Fab domains is large for the entire ensemble (Figure 4B and C) and reduced for structures with $\chi^2 < 2$ (Figure 4D and E).

While ensemble and subensemble fits of the data are informative, it is also instructive to evaluate the quality of fit for individual struct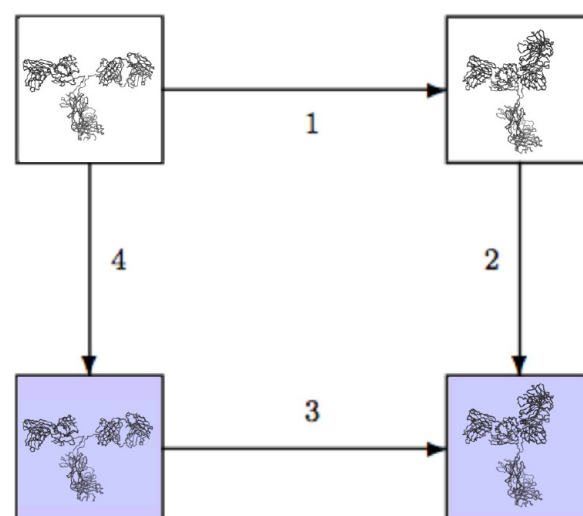ures. Eight structures are singled out for comparison, as enumerated in Figure 3B. These structures highlight the range of R_g values (structures 1 and 2), the initial homology model (4), and several structures with R_g values close to the experimentally determined value (3, 5−8). Structure 3 represents a configuration with an R_g near the experimental value yet with a rather poor χ^2 value. A comparison of the eight structures is shown in Figure 5. Generally, compact structures (1) and extended structures (2 and 4) have large χ^2 values, while intermediate structures are more favorable (5−8). Structure 3, while having an R_g close to the experimental value, is bent with one of the Fab domains parallel to the Fc domain to a larger extent than that of the individual more favorable structures (5−8). Thus, there are single structures and ensembles of structures that have SANS profiles that reproduce the experimental data. The number of relevant constraints inherent in a SAS profile is largely unknown, but it is generally accepted that the mathematical problem is underdetermined. Such highly underdetermined problems can have an infinite number of solutions.[43] Thus, one should be careful in considering a single structure, or a single linear combination of structures, as the definitive solution to the problem. Qualitatively, our conservative analysis mimics heuristic algorithms used to study intrinsically disordered proteins. There are many experimental measurements that could potentially be used to provide constraints to refine the ensemble including analytical ultracentrifugation, NMR (nuclear Overhauser effect and residual dipolar coupling constraints), electron microscopy, Förster energy resonance transfer, and chemical cross-linking among others.

Lacking such experimental constraints, our approach is to apply a free-energy analysis to evaluate the relative configurational energies of the subensemble of ASA-IgG2. This approach implements the widely used MM-PBSA method[44] that, for example, has used to study RNA/DNA stability,[45] the effect of amino-acid mutations in proteins,[46] protein−protein binding,[47] and ligand binding affinities.[48] Briefly, the approach used defines the standard Gibbs free energy of a particular molecular configuration as

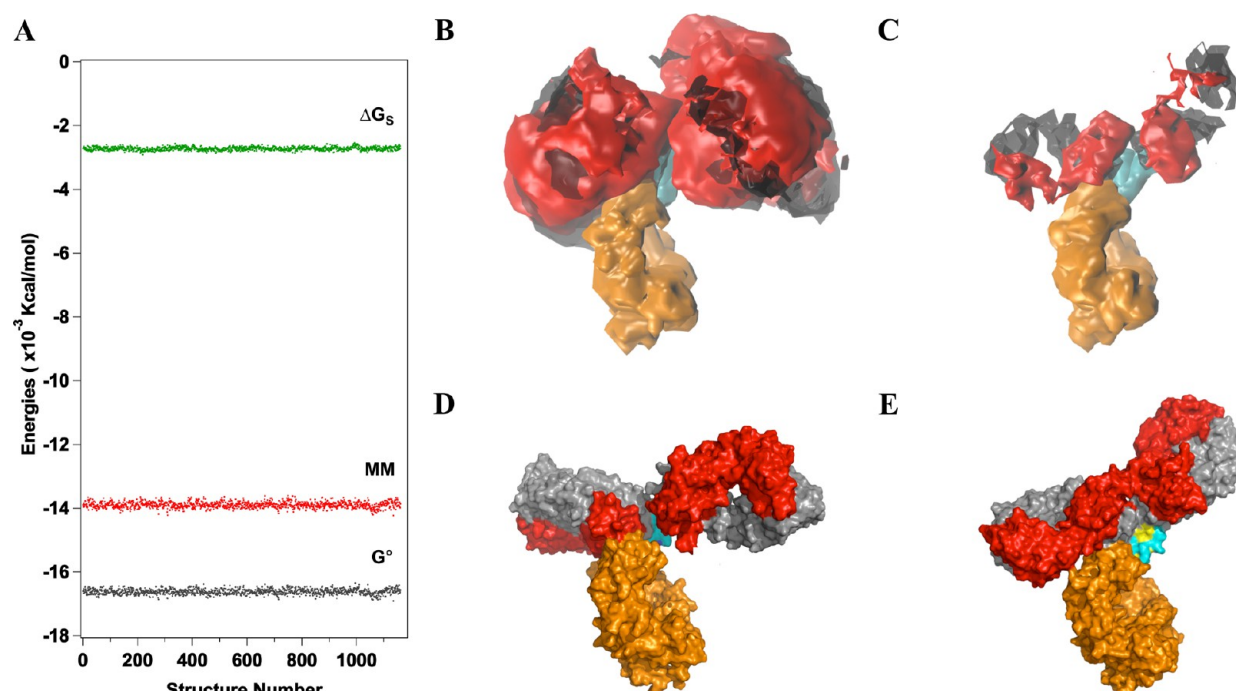$$G^\circ = G^\circ_{internal} + G^\circ_{solvation} \tag{4}$$

Figure 7. Relative configurational free-energy analysis of structures with $\chi^2 < 2$. (A) Molecular mechanical energy (MM), non-polar solvation energy (ΔG_S), and total free energy (G°) for 1060 structures (see Figures 3D and 4E). (B) Iso-density plot of ASA-IgG2 (reproduced from Figure 4E). (C) Iso-density plot of structures of lowest free energy within 1 standard deviation of the lowest single free-energy value in the ensemble. Parts D and E show structures with the single highest and lowest free energy, respectively.

The values of G° will be used to determine ΔG° between the various thermodynamic states defined in Figure 6. We estimate $\Delta G^\circ_{\text{internal}}$ to be the difference in intermolecular energy

$$\Delta G^\circ_{\text{internal}} \approx \Delta H^\circ_{\text{internal}} = \Delta U_{\text{internal}} \qquad (5)$$

where

$$\Delta U_{\text{internal}} = \Delta U_b + \Delta U_a + \Delta U_d + \Delta U_i + \Delta U_e + \Delta U_v \qquad (6)$$

with the subscripts defined as b = bond, a = angle, d = dihedral, i = improper, e = electrostatic, and v = van der Waals corresponding to the molecular mechanical energies taken from the final minimization step after MC and MD. We are concerned with the relative configurational free energy $\Delta\Delta G$ that can be calculated from the thermodynamic cycle that can be determined by

$$\Delta\Delta G = \Delta G_1 - \Delta G_3 = \Delta G_4 - \Delta G_2 \qquad (7)$$

The various ΔG components are calculated for

$$\Delta G^\circ_{\text{internal}} = \Delta U_{\text{internal}} \qquad (8)$$

and

$$G^\circ_{\text{solvation}} \approx \Delta G_{\text{polar solvation}} \qquad (9)$$

where $\Delta G_{\text{solvation}}$ was calculated using APBS.[38] Non-polar free-energy and entropic contributions contribute far less than the terms that we consider;[44,47] therefore, they are not included in our analysis.

As shown in Figure 7a, the relative free energy, ΔG°, of the subensemble (structures with $\chi^2 < 2$) shows that fewer configurations may be energetically favorable. The standard deviation of G values for the 1160 structures was ≈ 80 kcal/mol, close to the theoretical value of $G^\circ(N)^{1/2} \approx 78$ kcal/mol. Evaluation of structures near the energetic minimum could

allow a substantial reduction of the number of plausible, and hence representative, structures. Typically, to numerically converge energetic values, one can sample configurations from an extended MD simulation. There are 1160 structures with $\chi^2 < 2$ out of the entire 56 511 structures. Separate multi-nanosecond simulations of either of these ensembles and subsequent free-energy analysis is a monumental task. Attempts to discriminate energetic differences based on relative arrangements of Fab and Fc domains were inconclusive (not shown).

Alternatively, it is instructive to investigate the energetic fluctuations and converged averages for the highest and lowest energy structures in the subensemble described in Figure 7a. These individual structures are shown in Figure 7d and e. The two structures were separately equilibrated and simulated with explicit solvent for 100 ns as described above. Individual structures were extracted from each system every 100 ps to calculate the energies shown in Figure 7. The mean values and statistical fluctuations were -16400 ± 326 and -16800 ± 406 kcal/mol, respectively, where fluctuations are defined as 1 standard deviation. This limited analysis indicates that the lower energy structure is more favorable. Our analysis is not exhaustive, and there are limits of the accuracy of computationally derived values of free energies. Further development of computational calculation of free energies and validation to experimental data is needed.

The energetic analysis, while computationally demanding and inherently limited, includes thermodynamic constraints that have the potential to improve upon the results of conservative linear combinations of scattering profiles, as shown in Figures 4 and 5. This applies to any heurisitic method used to make linear combinations of scattering profiles to refine the fit to the scattering data. Furthermore, this analysis could be useful to evaluate specific atomistic contacts between Fab and Fc regions which could be used to determine potential

disulfide shuffling, oxidation mechanisms, and other deleterious events. The incorporation of energetic constraints will undoubtedly benefit the analysis of small-angle scattering data of proteins, nucleic acids, their complexes, and other soft-matter systems.

■ SUMMARY

The upper-hinge region of mAbs has long been recognized to provide a large degree of flexibility to Fab domains. Human IgG2 mAbs contain a contiguous region of ∼3 amino acids that potentially have the ability to enable this flexibility. The amino-terminal amino acid of the hinge region has steric limitations, and the carboxy-terminal is anchored by a disulfide bond. We have found via MC−MD simulations that a large number of configurations are energetically feasible covering a R_g range of at least 16 Å. The average of the scattering profiles suggest, as one possible solution, that the entire ensemble contributes equally and thus by itself represents the solution configurations of ASA-IgG2. A subensemble with lower average deviation from the scattering data is a second equally valid solution, and thus, a linear combination of these structures by itself, separately, can represent the solution configurations. To understand the ensemble on an energetic basis, a free-energy analysis of the subensemble was carried out. This analysis indicates that a reduced subset of structures have lower average free energies and thus using this constraint more physically reasonable structures were determined. This type of analysis could be useful where specific interactions that are known to effect function and or chemical stability are of interest. The software to create, manage, and evaluate such ensembles of flexible biomacromolecules starting from a single starting structure to the analysis via open-source solvation energy calculators[38] is available for general use.[21]

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: joseph.curtis@nist.gov.

Notes
The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Noelken, M. E.; Nelson, C. A.; Buckley, C. E.; Tanford, C. Gross Conformation of Rabbit 7 S Gamma-Immunoglobulin and its Papain-Cleaved Fragments. *J. Biol. Chem.* **1965**, *240*, 218−224.

(2) Wrigley, N. G.; Brown, E. B.; Skehel, J. J. Electron Microscopic Evidence for the Axial Rotation and Inter-Domain Flexibility of the Fab Regions of Immunoglobulin G. *J. Mol. Biol.* **1983**, *169*, 771−774.

(3) Roux, K. H. Direct Demonstration of Multiple Vh Allotopes on Rabbit Ig Molecules: Allotope Characteristics and Fab Arm Rotational Flexibility Revealed by Immunoelectron Microscopy. *Eur. J. Immunol.* **1984**, *14*, 459−464.

(4) Wade, R. H.; Taveau, J. C.; Lamy, J. N. Concerning the Axial Rotational Flexibility of the Fab Regions of Immunoglobulin. *J. Mol. Biol.* **1989**, *206*, 349−356.

(5) Schneider, W. P.; Wensel, T. G.; Stryer, L.; Oi, V. T. Genetically Engineered Immunoglobulins Reveal Structural Features Controlling Segmental Flexibility. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 2509−2513.

(6) Davies, D. R.; Chacko, S. Antibody Structure. *Acc. Chem. Res.* **1993**, *26*, 421−427.

(7) Harris, L. J.; Skaletsky, E.; McPherson, A. Crystallographic Structure of an Intact IgG1 Monoclonal Antibody. *J. Mol. Biol.* **1998**, *275*, 861−872.

(8) Kim, H.; Matsunaga, C.; Yoshino, A.; Kato, K.; Arata, Y. Dynamical Structure of the Hinge Region of Immunoglobulin G as Studied by 13C Nuclear Magnetic Resonance Spectroscopy. *J. Mol. Biol.* **1994**, *236*, 300−309.

(9) Guddat, L. W.; Herron, J. N.; Edmundson, A. B. Three-Dimensional Structure of a Human Immunoglobulin with a Hinge Deletion. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 4271−4275.

(10) Silverton, E. W.; Navia, M. A.; Davies, D. R. Three-Dimensional Structure of an Intact Human Immunoglobulin. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 5140−5144.

(11) Saphire, E. O.; Parren, P. W.; Pantophlet, R.; Zwick, M. B.; Morris, G. M.; Rudd, P. M.; Dwek, R. A.; Stanfield, R. L.; Burton, D. R.; Wilson, I. A. Crystal Structure of a Neutralizing Human IgG Against HIV-1: A Template for Vaccine Design. *Science* **2001**, *293*, 1155−1159.

(12) Harris, L. J.; Larson, S. B.; Hasel, K. W.; McPherson, A. Refined Structure of an Intact IgG2a Monoclonal Antibody. *Biochemistry* **1997**, *36*, 1581−1597.

(13) Ashish; Solanki, A. K.; Boone, C. D.; Krueger, J. K. Global Structure of HIV-1 Neutralizing Antibody IgG1 b12 is Asymmetric. *Biochem. Biophys. Res. Commun.* **2010**, *391*, 947−951.

(14) Perkins, S. J.; Okemefuna, A. I.; Nan, R.; Li, K.; Bonner, A. Constrained Solution Scattering Modelling of Human Antibodies and Complement Proteins Reveals Novel Biological Insights. *J. R. Soc., Interface* **2009**, *6* (Suppl.5), S679−96.

(15) Boehm, M. K.; Woof, J. M.; Kerr, M. A.; Perkins, S. J. The Fab and Fc Fragments of IgA1 Exhibit a Different Arrangement From That in IgG: A Study by X-ray and Neutron Solution Scattering and Homology Modelling. *J. Mol. Biol.* **1999**, *286*, 1421−1447.

(16) Lilyestrom, W. G.; Shire, S. J.; Scherer, T. M. Influence of the Cosolute Environment on IgG Solution Structure Analyzed by Small-Angle X-ray Scattering. *J. Phys. Chem. B* **2012**, *116*, 9611−9618.

(17) Gabel, F.; Wang, D.; Madern, D.; Sadler, A.; Dayie, K.; Daryoush, M. Z.; Schwahn, D.; Zaccai, G.; Lee, X.; Williams, B. R. G. Dynamic Flexibility of Double-Stranded RNA Activated PKR in Solution. *J. Mol. Biol.* **2006**, *359*, 610−623.

(18) Bernado, P.; Mylonas, E.; Petoukhov, M. V.; Blackledge, M.; Svergun, D. I. Structural Characterization of Flexible Proteins Using Small-Angle X-ray Scattering. *J. Am. Chem. Soc.* **2007**, *129*, 5656−5664.

(19) Hammel, M. Validation of Macromolecular Flexibility in Solution by Small-Angle X-ray Scattering (SAXS). *Eur. Biophys. J.* **2012**, *41*, 789−799.

(20) Różycki, B.; Kim, Y. C.; Hummer, G. SAXS Ensemble Refinement of ESCRT-III CHMP3 Conformational Transitions. *Structure* **2011**, *19*, 109−116.

(21) Curtis, J. E.; Raghunandan, S.; Nanda, H.; Krueger, S. SASSIE: A Program to Study Intrinsically Disordered Biological Molecules and Macromolecular Ensembles Using Experimental Scattering Restraints. *Comput. Phys. Commun.* **2012**, *183*, 382−389.

(22) Pelikan, M.; Hura, G. L.; Hammel, M. Structure and Flexibility within Proteins as Identified Through Small Angle X-ray Scattering. *Gen. Physiol. Biophys.* **2009**, *28*, 174−189.

(23) Fisher, C. K.; Huang, A.; M., S. C. Modeling Intrinsically Disordered Proteins with Bayesian Statistics. *J. Am. Chem. Soc.* **2010**, *132*, 14919−14927.

(24) Glinka, C. J.; Barker, J. G.; Hammouda, B.; Krueger, S.; Moyer, J. J.; Orts, W. J. The 30-Meter Small Angle Neutron Scattering Instruments at the National Institute of Standards and Technology. *J. Appl. Crystallogr.* **1998**, *31*, 430−445.

(25) Kline, S. R. Reduction and Analysis of SANS and USANS Data Using IGOR Pro. *J. Appl. Crystallogr.* **2006**, *39*, 895−900.

(26) Svergun, D. I. Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria. *J. Appl. Crystallogr.* **1992**, *25*, 495−503.

(27) Guinier, A.; Fournet, G. *Small Angle Scattering of X-rays*; John Wiley and Sons Inc.: New York, 1955.

(28) Chemical Computing Group. Molecular Operating Environment (MOE). 2011; 1010 Sherbooke St. West, Suite 910, Montreal, QC, Canada, H3A 2R7, 2011.

(29) Berman, H.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.; Weissig, H.; Shindyalov, I.; Bourne, P. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.

(30) Wypych, J.; Li, M.; Guo, A.; Zhang, Z.; Martinez, T.; Allen, M. J.; Fodor, S.; Kelner, D. N.; Flynn, G. C.; Liu, Y. D.; et al. Human IgG2 Antibodies Display Disulfide-Mediated Structural Isoforms. *J. Biol. Chem.* **2008**, *283*, 16194−16205.

(31) Dillon, T. M.; Ricci, M. S.; Vezina, C.; Flynn, G. C.; Liu, Y. D.; Rehder, D. S.; Plant, M.; Henkle, B.; Li, Y.; Deechongkit, S.; et al. Structural and Functional Characterization of Disulfide Isoforms of the Human IgG2 Subclass. *J. Biol. Chem.* **2008**, *283*, 16206−16215.

(32) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781−1802.

(33) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem.* **1983**, *4*, 187−217.

(34) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.

(35) Tanner, D. E.; Chan, K.-Y.; Phillips, J. C.; Schulten, K. Parallel Generalized Born Implicit Solvent Calculations with NAMD. *J. Chem. Theory Comput.* **2011**, *7*, 3635−3642.

(36) Heidorn, D. B.; Trewhella, J. Comparison of the Crystal and Solution Structures of Calmodulin and Troponin C. *Biochemistry* **1988**, *27*, 909−915.

(37) Krueger, S.; Gorshkova, I.; Brown, J.; Hoskins, J.; McKenney, K. H.; Schwarz, F. P. Determination of the Conformations of cAMP Receptor Protein and its T127L,S128A Mutant with and without cAMP from Small Angle Neutron Scattering Measurements. *J. Biol. Chem.* **1998**, *273*, 20001−20006.

(38) Baker, N. A.; D., S.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of Nanosystems: Application to Microtubules and the Ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10037−10041.

(39) Dolinsky, T. J.; Czodrowski, P.; Li, H.; Nielsen, J. E.; Jensen, J. H.; Klebe, G.; Baker, N. A. PDB2PQR: Expanding and Upgrading Automated Preparation of Biomolecular Structures for Molecular Simulations. *Nucleic Acids Res.* **2007**, *35*, W522−W525.

(40) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. PDB2PQR: An Automated Pipeline for the Setup, Execution, and Analysis of Poisson-Boltzmann Electrostatics Calculations. *Nucleic Acids Res.* **2004**, *32*, W665−W667.

(41) Li, H.; Robertson, A. D.; Jensen, J. H. Very Fast Empirical Prediction and Rationalization of Protein pKa Values. *Proteins* **2005**, *61*, 704−721.

(42) Humphrey, W.; Dalke, A.; Schulten, K. VMD - Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.

(43) Golub, G. H.; Loan, C. F. V. *Matrix Computations*, 3rd ed.; The Johns Hopkins University Press: Baltimore, MD, 1996.

(44) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; et al. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.* **2000**, *33*, 889−897.

(45) Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate-DNA Helices. *J. Am. Chem. Soc.* **1998**, *120*, 9401−9409.

(46) Huo, S.; Massova, I.; Kollman, P. A. Computational Alanine Scanning of the 1:1 Human Growth Hormone-Receptor Complex. *J. Comput. Chem.* **2002**, *23*, 15−27.

(47) Gohlke, H.; Kiel, C.; Case, D. A. Insights into Protein-Protein Binding by Binding Free Energy Calculation and Free Energy Decomposition for the Ras-Raf and Ras-RalGDS Complexes. *J. Mol. Biol.* **2003**, *330*, 891−913.

(48) Massova, I.; Kollman, P. Combined Molecular Mechanical and Continuum Solvent Approach (MM-PBSA/GBSA) to Predict Ligand Binding. *Perspect. Drug Discovery Des.* **2000**, *18*, 113−135.