BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Role of Electrostatic Interactions in Binding of Peptides and Intrinsically Disordered Proteins to Their Folded Targets. 1. NMR and MD Characterization of the Complex between the c-Crk N-SH3 Domain and the Peptide Sos

Yi Xue,[†,∥] Tairan Yuwen,[†] Fangqiang Zhu,[‡] and Nikolai R. Skrynnikov*[,†,§]

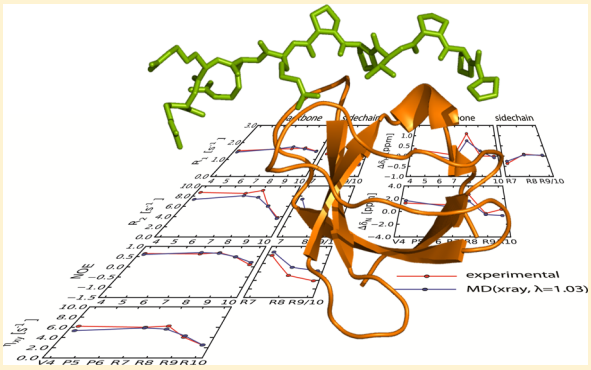
[†]Department of Chemistry, Purdue University, West Lafayette, Indiana 47907, United States

[‡]Department of Physics, Indiana University-Purdue University Indianapolis, Indianapolis, Indiana 46202, United States

[§]Laboratory of Biomolecular NMR, St. Petersburg State University, St. Petersburg 199034, Russia

Supporting Information

ABSTRACT: Intrinsically disordered proteins (IDPs) often rely on electrostatic interactions to bind their structured targets. To obtain insight into the mechanism of formation of the electrostatic encounter complex, we investigated the binding of the peptide Sos (PPPVPPRRRR), which serves as a minimal model for an IDP, to the c-Crk N-terminal SH3 domain. Initially, we measured ^{15}N relaxation rates at two magnetic field strengths and determined the binding shifts for the complex of Sos with wild-type SH3. We have also recorded a 3 μs molecular dynamics (MD) trajectory of this complex using the Amber ff99SB*-ILDN force field. The comparison of the experimental and simulated data shows that MD simulation consistently overestimates the strength of salt bridge interactions at the binding interface. The series of simulations using other advanced force fields also failed to produce any satisfactory results. To address this issue, we have devised an empirical correction to the Amber ff99SB*-ILDN force field whereby the Lennard-Jones equilibrium distance for the nitrogen−oxygen pair across the Arg-to-Asp and Arg-to-Glu salt bridges has been increased by 3%. Implementing this correction resulted in a good agreement between the simulations and the experiment. Adjusting the strength of salt bridge interactions removed a certain amount of strain contained in the original MD model, thus improving the binding of the hydrophobic N-terminal portion of the peptide. The arginine-rich C-terminal portion of the peptide, freed from the effect of the overstabilized salt bridges, was found to interconvert more rapidly between its multiple conformational states. The modified MD protocol has also been successfully used to simulate the entire binding process. In doing so, the peptide was initially placed high above the protein surface. It then arrived at the correct bound pose within ∼2 Å of the crystallographic coordinates. This simulation allowed us to analyze the details of the dynamic binding intermediate, i.e., the electrostatic encounter complex. However, an experimental characterization of this transient, weakly populated state remains out of reach. To overcome this problem, we designed the double mutant of c-Crk N-SH3 in which mutations Y186L and W169F abrogate tight Sos binding and shift the equilibrium toward the intermediate state resembling the electrostatic encounter complex. The results of the combined NMR and MD study of this engineered system will be reported in the next part of this paper.



The interplay between order and disorder is an essential feature of any proteome. Those proteins in which disorder prevails are known as intrinsically disordered proteins (IDPs). They can be identified with relative ease on the basis of primary sequence; proteins with significant net charge and a low proportion of hydrophobic residues tend to be disordered.[1−3] Additional parameters, such as flexibility and bulkiness of amino acids, can also be incorporated into prediction algorithms.[4] Estimates obtained along these lines suggest that as many as 50% of eukaryotic proteins contain long stretches of disordered residues,[5] and more than 20% can be described as predominantly disordered.[4] These statistics demonstrate that IDPs cannot be dismissed as a rare quirk of nature; on the contrary, they represent one of the broad and fundamentally important classes of proteins.

Of interest, the proportion of IDPs in eukarya appears to be much higher than in archaea or prokarya (in the case of archaea, only 2% of proteins are predicted to have long disordered regions).[6] Thus, IDPs can be viewed as an advanced

Received: July 23, 2014
Revised: September 9, 2014
Published: September 10, 2014

ACS Publications © 2014 American Chemical Society

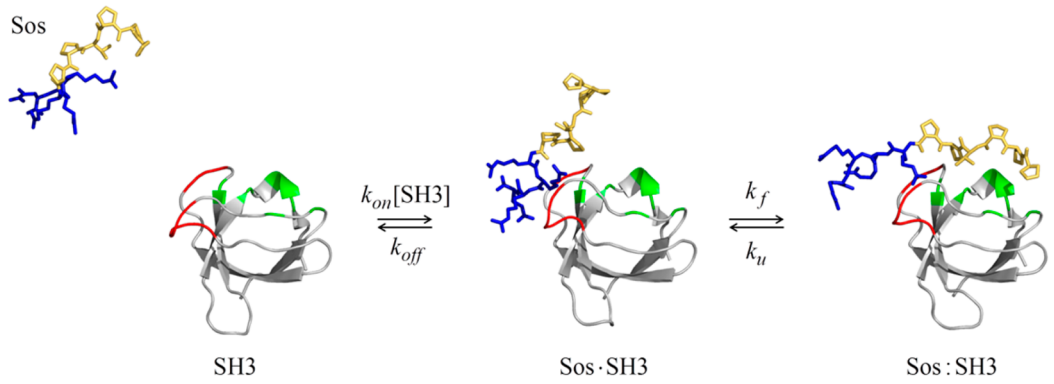


Figure 1. Schematic representation of the binding mechanism between the 10-residue peptide Sos and c-Crk N-SH3 domain. The images are from MD trajectories and are for the purpose of illustration only. The proteins are colored as follows: blue for positively charged Arg residues in Sos, red for negatively charged Glu and Asp residues that form an acidic patch at the periphery of the peptide binding site on the surface of the SH3 domain, gold for hydrophobic residues in Sos, and green for hydrophobic residues lining up the peptide binding site on the surface of the SH3 domain. Note that in reality free Sos and Sos·SH3 are mixtures of highly diverse conformational species.

type of molecular machinery that evolved in higher organisms. As it happens, disordered proteins are typically involved in cell signaling and regulation.[7,8] In this context, the unfoldedness of IDPs confers a number of functional advantages. IDPs can literally wrap themselves around their binding targets, thus making use of multiple binding sites.[9−13] This mechanism creates new possibilities in terms of binding affinity and specificity. For example, an intrinsically flexible protein may be able to bind several structurally different partners. Such multitasking helps in the development of efficient, highly integrated signaling networks, in which every protein serves in more than one role. In addition, IDPs offer the benefit of economy because they present large binding interfaces but do not require large supporting scaffolds.[14]

What are the functional requirements for an IDP binding to its structured target in the context of cell signaling? First, sufficiently high specificity is to be expected of any such interaction. This implies a fairly extensive pattern of contacts between an IDP and its folded target. To form these contacts, IDP needs to adopt an appropriate spatial conformation. Given that intrinsically disordered proteins normally go through a multitude of different conformations, this requirement should effectively slow the binding process and lead to low on rates (k_{on}). At the same time, there is an evolutionary pressure to maintain a high off rate (k_{off}). This is necessary to ensure a rapid response to changing external stimuli. The combination of a low k_{on} and a high k_{off} is expected to produce a large dissociation constant ($K_d = k_{off}/k_{on}$). In a cellular milieu, where the concentrations of proteins are typically low, this should render the entire mechanism unsustainable.

As it appears, nature has found a way to resolve this conundrum. As already pointed out, IDPs usually carry a significant net charge (which often includes a contribution from post-translational modifications, such as phosphorylation).[15−17] This fundamental characteristic of IDPs makes it possible to rely on the so-called electrostatic steering mechanism to boost the on rate.[18] Briefly, long-range electrostatic forces pull an IDP toward the relevant acidic (or basic) patch on the surface of a target protein, thus establishing contact in the vicinity of the binding site. The resulting state, in which IDP is loosely anchored at the periphery of the binding site, has been described as "electrostatic encounter complex".[19] This mechanism effectively increases the local concentration of the ligand in the vicinity of the binding site and thus promotes binding. The on rates in excess of 10^7 M^{-1} s^{-1} are invariably associated with this mechanism.[20,21]

After the electrostatic encounter complex is formed, it quickly evolves into the final complex, which is structured and has a (reasonably) high affinity. During this transition, the IDP adopts a suitable conformation to make multiple binding contacts. This process has been described as "folding upon binding".[22] (In general, binding of IDPs to their folded target falls under the rubric of "induced fit", although the elements of "conformational selection" have also been noted.[23,24]) It has also been observed that the individual segments of the IDP "coalesce" around the multiple attachment points on the surface of the target protein.[25,26] The electrostatic interactions continue to play a significant role at this late stage in the binding process. It has been pointed out that electrostatic complementarity is essential for binding specificity in the case of IDP complexes.[27] It has also been suggested that the folding is greatly accelerated by a pattern of charged residues surrounding the binding site.[28,29]

In this paper and the following paper, we focus on the details of the electrostatic interactions involved in the formation of the electrostatic encounter complex. We use the experimental nuclear magnetic resonance (NMR) data as well as molecular dynamics (MD) simulations to obtain a realistic model for these interactions. Recent advances in computer hardware and MD algorithms made it possible to simulate binding of a small IDP to its folded target using regular (unbiased) MD simulations in explicit solvent. This has been demonstrated for binding of peptides to modular domains such as SH2 and SH3 (where the peptides can be viewed as a minimal model for an IDP).[30,31] At the same time, Ubbink's laboratory recently presented an NMR study of low-affinity electrostatic complexes formed by short peptides with the sequence XKKKK and electron transfer protein plastocyanin.[32] In our work, we pursue a synergistic approach, seeking to combine experimental NMR data with MD simulations. Of note, we have used a rigorous MD protocol with a state-of-the-art force field, as opposed to coarse-grained simulations, replica exchange simulations, or Monte Carlo modeling, which were all successfully applied in the context of peptide docking.[29,32−34] The use of bona fide MD allows us to draw a direct comparison with the NMR measurables.

For the purpose of this study, we have selected a system consisting of the N-terminal SH3 domain from adaptor protein

dx.doi.org/10.1021/bi500904f | *Biochemistry* 2014, 53, 6473−6495

c-Crk and a 10-residue peptide from Ras activator protein Sos. Cellular Crk is ubiquitously expressed in a wide range of cells and tissues.[35,36] In its role as an adaptor protein, c-Crk appears at the center of a vast and complex signaling network; the function of c-Crk is to bring together various proteins that bind to its several modular domains (SH2 and one or two SH3 domains).[37−39] In particular, the N-terminal SH3 domain of c-Crk binds a number of guanine nucleotide exchange factors; these interactions have been directly linked to the high motility of cancerous cells.[40−42] One of the nucleotide exchange factors that binds to c-Crk N-SH3 is the Ras activator protein Sos.[43−45] The proline-rich sequence that is responsible for this binding belongs to the disordered carboxyl-terminal segment of Sos.[46] The corresponding 10-residue peptide, PPPVPPRRRR, has been crystallized in complex with murine c-Crk N-SH3.[47] The resulting crystallographic structure was used in the recent MD simulation study featuring a shorter version of the Sos peptide, PPPVPPRR.[30]

In the first part of our study (this paper, hereafter termed paper 1), we focus on the interaction between the wild-type murine c-Crk SH3 (wtSH3) and the 10-residue Sos peptide (Sos). The presumed binding mechanism is illustrated in Figure 1, which shows the transition from the free state of the peptide (Sos) to the putative electrostatic encounter complex (Sos·wtSH3) and further to the fully formed complex (Sos:wtSH3). The formation of the hypothetical encounter complex is underpinned by the interaction between four arginines at the C-terminus of the Sos peptide (blue) and the Glu/Asp-rich patch on the surface of wtSH3 (red). The final complex is then formed as the side chains of the two conserved Sos prolines (gold) are tucked into the shallow hydrophobic grooves on the surface of wtSH3 (green). As a matter of convention, the formation of Sos·wtSH3 is classified as binding (rate constants k_{on} and k_{off}) and the subsequent transition to Sos:wtSH3 is viewed as folding (rate constants k_f and k_u).

For the Sos:wtSH3 complex, which has a relatively high affinity, we have experimentally measured backbone and side-chain ^{15}N relaxation rates in the bound Sos as well as spectral frequency shifts caused by the binding. We have also recorded a long MD trajectory of Sos:wtSH3 in explicit solvent under the Amber ff99SB*-ILDN force field,[48−50] as well as another long MD trajectory representing the free Sos peptide. The MD data have been used to predict NMR measurables, and the calculated results were subsequently compared with the experimental data. This comparison revealed a systematic problem with the MD simulation—as it turned out, the salt bridge interactions in Sos:wtSH3 were consistently over-emphasized. This is a known problem with the most popular force fields.[51,52] To address this problem, we have introduced an ad hoc correction to the force field. Specifically, the parameters of the pairwise Lennard-Jones potential between the ionized groups in the Arg and Glu/Asp side chains have been slightly adjusted. Using this modified version of the force field, we were able to successfully reproduce the experimental data from the Sos:wtSH3 complex. Furthermore, we were able to simulate the entire binding process beginning from free Sos, progressing to the Sos·wtSH3 encounter complex, and ultimately undergoing the transition to the final Sos:wtSH3 complex. This MD trajectory was also consistent with the experimental NMR data in the part pertaining to the Sos:wtSH3 state.

As a next step, we focused on the investigation of the electrostatic encounter complex between Sos and c-Crk N-SH3 (second part of this work, termed paper 2, to be submitted for publication). The experimental characterization of encounter complexes of IDPs poses a serious challenge. First, they tend to be sparsely populated and short-lived. Second, they are comprised of many interconverting conformational species. Third, they are typically found in rapid exchange with the heavily populated free and/or bound states. NMR spectroscopy, which is the most important experimental technique in this arena, is affected by all of the complications mentioned above. In principle, encounter complexes involving IDPs can be targeted using those experiments that are highly sensitive to the presence of minor species: relaxation dispersion,[53,54] chemical exchange saturation transfer,[55,56] or paramagnetic relaxation enhancements.[57] Yet it remains extremely difficult to capture the three-state exchange process such as that illustrated in Figure 1. For instance, a recent relaxation dispersion study of the Fyn SH3 domain binding 12-residue peptide failed to produce any direct evidence of the encounter complex.[58]

In this situation, we resort to an alternative strategy. Specifically, we have modified the interaction of Sos with c-Crk N-SH3 such as to mimic the electrostatic encounter complex. Toward this goal, we introduced two point mutations in the area of two shallow hydrophobic grooves in the SH3 domain (green ribbon in Figure 1). These two mutations, Y186L and W169F, are supposed to abrogate the hydrophobic component of binding that is responsible for formation of the final complex. At the same time, these mutations preserve the electrostatic Arg-to-(Asp/Glu) interactions. The idea is to shift the thermodynamic balance away from Sos:SH3 and toward the Sos·SH3 electrostatic encounter complex.

The altered system is designed to boost the population of the electrostatic encounter complex. This opens the door for standard NMR experiments, which do not need to be sensitive to weakly populated species. In particular, in the case of the electrostatic encounter complex between Sos and the double mutant SH3 domain (dmSH3), we have recorded HSQC titration data reporting on k_{on} and k_{off} rates, backbone and side-chain ^{15}N relaxation rates reporting on the mobility of the Sos peptide, the corresponding ^{1}H and ^{15}N chemical shifts, and the temperature coefficients pertaining to these chemical shifts. Using our adaptation of the Amber force field, we were able to produce an MD model that proved to be in good agreement with the experimental data from the Sos·dmSH3 system. All of the results pertaining to Sos·dmSH3 will be discussed in paper 2.

MATERIALS AND METHODS

Protein and Peptide Expression. A primer for producing c-Crk N-SH3 was designed in house using mouse cDNA (GenBank accession number BC031149) and GST plasmid pGEX-4T-1. The protein was expressed and purified as described elsewhere[59] to produce unlabeled, ^{15}N-labeled, or ^{15}N- and ^{13}C-labeled material. The ^{15}N-labeled Sos peptide was bacterially expressed as a fusion construct with a solubility tag protein, the GB1 domain from protein G.[54,60,61] The Sos sequence was subcloned into GB1-containing plasmid GEV 2, with a single spacer residue (Asp). Following the carefully optimized expression and purification procedure, GB1-Sos was cleaved at the Asp-Pro site using a treatment with a 50% (v/v) solution of formic acid (24 h at 50 °C). The products were separated using C18 column chromatography. The sample was concentrated using a lyophilizer or SpeedVac apparatus; the buffer exchange was conducted over a period of at least 24 h

dx.doi.org/10.1021/bi500904f | Biochemistry 2014, 53, 6473−6495

using a dialyzer device with a 0.5 kDa membrane cutoff. The yield from this procedure was 1−2 mg of ^{15}N-labeled Sos from 1 L of M9 medium. Unlabeled Sos, as well as Sos selectively labeled with ^{15}N at the V4 position, was purchased commercially (Peptide 2.0 Inc.).

Sample Conditions. Unless indicated otherwise, the samples used in this work have been prepared in a buffer containing a 90% H_2O/10% D_2O buffer, 20 mM phosphate, and 0.02% NaN_3 (pH 6.0). The concentrations of the peptide and protein were determined using the NMR method by Wider and Dreier, which ensures an accuracy on the order of 3%.[62,63] Specifically, we determined the intensities of isolated methyl resonances from one-dimensional (1D) proton spectra (V4 in Sos and V184 in c-Crk N-SH3) and compared them with the external reference (10 mM sample of valine). For c-Crk N-SH3, the results were consistent with the standard approach based on UV absorption by aromatic residues, A_{280}. All NMR data were collected at 23.4 °C.

NMR Measurements. NMR measurements were conducted using a Varian Inova 600 MHz spectrometer equipped with an HCN probe; additionally, ^{15}N relaxation parameters were measured using a Bruker Avance DRX 500 MHz spectrometer equipped with a TXI cryoprobe. The backbone ^{15}N relaxation measurements were conducted using updated versions of the standard relaxation experiments,[64−66] including the recently corrected heteronuclear NOE sequence.[67] ^{15}N CSA-dipolar transverse cross-correlation rate η_{xy} has been measured by means of the IPAP sequence[68] with an added relaxation period with a length that is a multiple of $1/^1J_{NH}$.[69] For arginine side-chain $^{15}N^{\varepsilon}$ relaxation measurements, 1D versions of the same experiments were used, with the nitrogen rf carrier set to 80 ppm. The data have been processed using the nmrPipe suite of programs,[70] including the *autofit* script, and Sparky.[71]

Spectral Assignments. The resonance assignments of wtSH3 are from ref 72. The resonance assignment for backbone amides in Sos is obtained from standard NOESY and TOCSY experiments performed on a 10 mM sample of unlabeled Sos.[73] The assignment has been confirmed by solvent exchange data: at pH 7.0, the resonances from R8 and R9 are broadened beyond detection, which is consistent with expectations.[74] They were also confirmed using a triple-resonance experiment correlating $^1H^{\varepsilon}$,$^{15}N^{\varepsilon}$ spins with $^1H^N$,^{15}N spins of the subsequent residue (unpublished). The spectrum of the free peptide also contains a number of minor peaks arising from the *cis* conformation of adjacent proline residues;[75] some of these minor peaks have also been assigned. Four arginine $^1H^{\varepsilon}-^{15}N^{\varepsilon}$ correlations in the spectrum of the free Sos peptide are overlapped, resulting in a single unresolved peak. Two of these resonances become resolved and can be assigned in the complex with wtSH3. $^1H^{\eta}-^{15}N^{\eta}$ correlations were found to be severely broadened and have not been used.

K_d and k_{on} for Binding of Sos to wtSH3. Isothermal calorimetry measurements were conducted at 23.4 °C in NMR buffer using an iTC200 instrument (MicroCal). The solution of wtSH3 (concentration 0.1 mM, volume 300 μL) was titrated with Sos (concentration 10 mM, injection volume 2 μL, total of 18 injections). The data were fit using MicroCal Origin, assuming two-state exchange:

$$\text{Sos} + \text{wtSH3} \underset{k_{OFF}}{\overset{k_{ON}[\text{wtSH3}]}{\rightleftharpoons}} \text{Sos:wtSH3} \quad (1)$$

which is the reduction of the three-state scheme shown in Figure 1. (The notations k_{ON} and k_{OFF} are reserved specifically for the two-state scheme in eq 1; they are distinct from k_{on} and k_{off} used in the three-state scheme in Figure 1.) The fitting suggests single-site binding with a K_d of 1.5 ± 0.1 μM (see Figure 2).

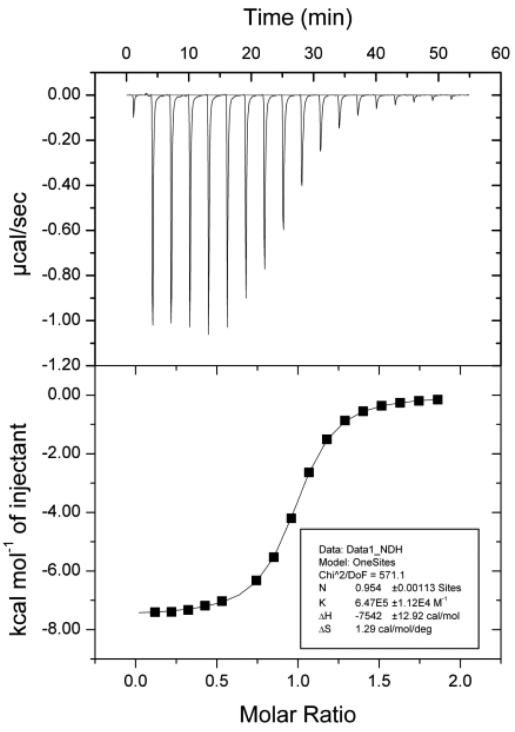


Figure 2. Calorimetric titration of wt c-Crk N-SH3 with Sos. The obtained affinity constant is 1.5 μM.

NMR titration was conducted by adding unlabeled wtSH3 to the sample of ^{15}N-labeled Sos (see Figure 3A). Large titration shifts are observed for residues V4, R7, and R8; for all these sites, the on−off exchange falls in the intermediate regime. The line shape analysis of V4 titration yields k_{ON} values of 1.6×10^9 and 2.0×10^9 M^{-1} s^{-1} (^{15}N and ^{1}H dimensions, respectively; see Figure 3B). At the same time, the joint treatment of all titration data from V4, R7, and R8 produces the global k_{ON} value of 1.5×10^9 M^{-1} s^{-1}.

Tumbling Time of wtSH3. The rotational correlation time of the protein needs to be known to complement MD simulations and facilitate the calculation of relaxation rates. For free wtSH3 in an aqueous solution, the calculation using HYDRONMR[79] predicts a nearly isotropic diffusion tensor with a τ_{rot} of 3.91 ns and a $D_{\parallel}/D_{\perp}$ of 1.05 (see Table S1 of the Supporting Information for details). To experimentally determine the parameters of overall tumbling, we relied on the ^{15}N R_2/R_1 data as analyzed by the program *r2r1_diffusion*.[80] The residues included in the analyses were selected according to the prescription of Tjandra et al.[81] Using the ^{15}N relaxation data from a dilute sample of free wtSH3 (protein concentration of 0.2 mM) and assuming that diffusion tensor is axially symmetric, we obtained τ_{rot} = 4.08 ns and a $D_{\parallel}/D_{\perp}$ = 1.14, in close agreement with expectations. However, as we increase the concentration of the protein to 2 mM, the experimentally determined τ_{rot} value shoots up to 6.64 ns. At the same time, the tumbling anisotropy remains nearly

dx.doi.org/10.1021/bi500904f | *Biochemistry* 2014, 53, 6473−6495

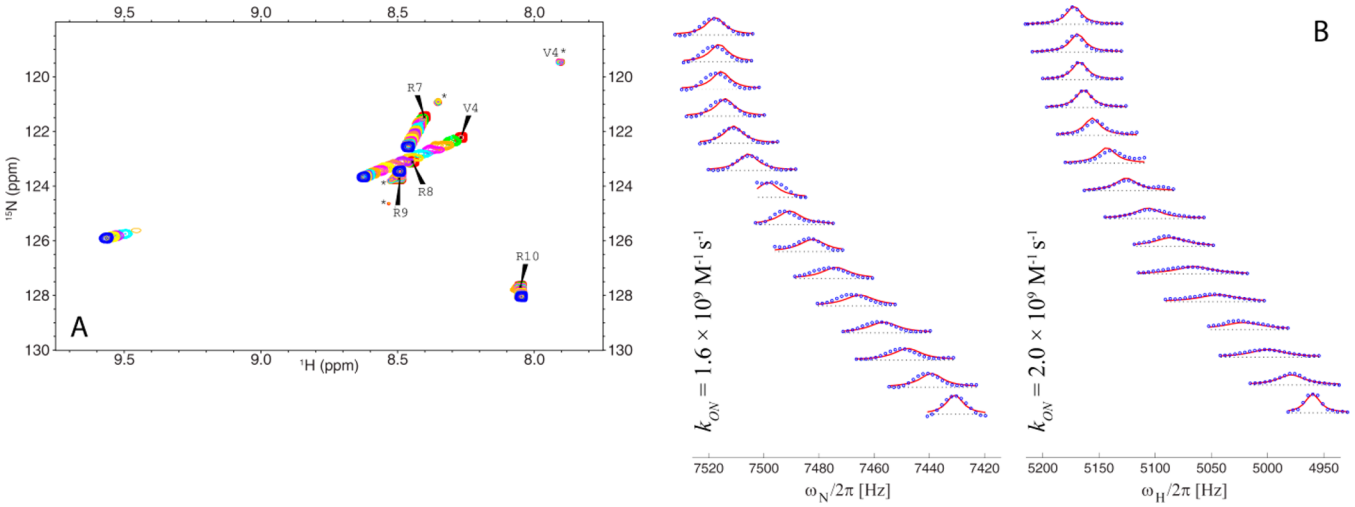


Figure 3. (A) $^{1}H^{N}-^{15}N$ HSQC titration of ^{15}N-labeled Sos with unlabeled wtSH3. The starting point corresponds to 0.20 mM Sos without wtSH3 (red contour in the plot); the final point corresponds to 0.18 mM Sos, after accounting for dilution, and 0.35 mM wtSH3 (blue contour in the plot). During the course of the titration, the peak from residue R8 disappears and then reappears far downfield. The minor peaks, labeled with asterisks, represent distinct conformational species arising from *cis*–*trans* isomerization of proline residues in Sos;[76,77] one of them has been identified as belonging to the valine residue (V4*). Minor peaks do not shift upon titration because the alternative conformational species Sos* have a much lower affinity for wtSH3. At the same time, these peaks gradually disappear as the protein:peptide ratio approaches 1:1. Indeed, with the addition of the protein, the Sos* ⇄ Sos ⇄ Sos:wtSH3 equilibrium becomes shifted to the right, thus depleting the population of the minor conformational species Sos*. The titration of the arginine side-chain N^{ε} sites has been recorded separately using the appropriate ^{15}N carrier setting (not shown). (B) Titration of spectral peak V4 (nitrogen and proton dimensions). The traces from HSQC spectra were fit using LineShapeKin.[78] Briefly, this program uses a fixed value of K_d as obtained from the ITC measurements; the line widths and intrinsic resonance frequencies are determined from the spectra of the free and protein-saturated Sos (bottom and top traces, respectively). The data from nitrogen and proton dimensions have been fit separately; the resulting k_{ON} values are indicated in the plot. The code of LineShapeKin has been slightly modified to obtain the desired graphics.

unchanged ($D_{\|}/D_{\perp}$ = 1.18), and the orientation of the long axis of the diffusion tensor changes only insignificantly, by 16° (Table S1 of the Supporting Information). This kind of behavior points toward low-affinity nonspecific (or weakly specific) dimerization, as has been previously observed for a variety of small globular proteins.[82−86]

Weak Dimerization of wtSH3. To quantitatively characterize the self-association effect in wtSH3, we have conducted a $^{1}H^{N}-^{15}N$ HSQC titration over the range of protein concentrations from 0.2 to 2.0 mM. A large number of signals are seen to titrate, shifting by a small amount (<0.06 and <0.3 ppm in the proton and nitrogen dimensions, respectively). The data have been fit assuming that the protein undergoes fast exchange between the monomeric and dimeric forms. The results of the fitting are presented in Figure S1 of the Supporting Information. Briefly, a dimerization site with an affinity constant in the range of 3−8 mM is found in the RT loop. In addition, there are multiple other sites with affinities of >10 mM that are broadly distributed over the surface of the protein. This is consistent with the notion that at a concentration of 2 mM roughly half of the protein is in the dimeric form, as can be deduced from ^{15}N relaxation data. The dimerization appears to be only weakly specific, which also agrees with the ^{15}N relaxation analyses (see above).

Tumbling Time of Sos:wtSH3. The τ_{rot} of the wtSH3:Sos complex has been determined using the same sample conditions that were used in the main series of relaxation measurements: 2 mM SH3 and 0.5 mM Sos. Addition of the peptide resulted only in a small increase in τ_{rot}, ca. 2% (see Table S1 of the Supporting Information). Considering that only one-fourth of the protein molecules are loaded with the peptide under the current experimental conditions, we estimate that the τ_{rot} of the Sos:wtSH3 complex is 7.2 ns (assuming an isotropic tumbling model).

Importantly, this result lends itself to direct experimental verification. The sample used to collect R_2/R_1 data was prepared with both wtSH3 and Sos peptide labeled with ^{15}N. Several of the Sos resonances are well-resolved in the spectra, including those from V4 and R8 that are in rigid contact with the protein (see the crystallographic structure Protein Data Bank entry 1CKB[47]). The R_2/R_1 ratio for these two residues translates into a tumbling time of 6.8 ns for the Sos:wtSH3 complex. More sophisticated model-free analysis using data collected at two magnetic field strengths yields values of 7.1 and 6.8 ns for the two peptide residues at hand. We conclude that the τ_{rot} value of 7.0 ± 0.2 ns represents an accurate and reliable result for the Sos:wtSH3 complex. For the purpose of subsequent MD analyses, it is reasonable to assume that the tumbling is isotropic; the modest anisotropy effect can be safely neglected.

MD Simulations. The main series of MD simulations has been conducted using the Amber 11 package with the ff99SB force field,[48,87] including ILDN side-chain correction[50] and backbone helical propensity correction.[49] The initial coordinates of the system were generated using two different approaches.

In the first approach, the initial coordinates were based on the crystallographic structure 1CKB.[47] The missing Sos residues R9 and R10 were added with random conformation using the program MODELLER.[88] The SH3 domain was simulated without three terminal residues, G132, S133, and R191, which were present in our NMR sample. We also conducted control simulations, in which these three residues were added to the crystallographic structure; the results proved to be unchanged.

dx.doi.org/10.1021/bi500904f | *Biochemistry* 2014, 53, 6473−6495

In the second approach, the peptide was placed over the surface of the protein in a random fashion. Specifically, in the crystal structure 1CKB, we defined a vector connecting the centers of mass of wtSH3 and Sos. The length of this vector was doubled from 12.6 to 25.2 Å, and the obtained point was used to place the center of mass of the Sos peptide. The initial conformation of peptide was chosen at random from a 2.1 μs MD simulation of the free Sos. The initial orientation of Sos was assigned randomly. Finally, a random rotation was applied to wtSH3 (the amplitude of this rotation was normally distributed with a variance of 30°).

All Glu and Asp residues in the SH3 domain were assumed to be deprotonated under the conditions of our experimental study (pH 6.0). This is consistent with the calculation employing the program PROPKA[89] and the crystallographic structure 1CKB. These calculated pK_a values are all sufficiently low (the highest value, 5.0, is predicted for residue E149). To verify this result, we have prepared a sample of Sos:wtSH3 with an elevated pH (7.0) where it is safe to assume that all Asp and Glu side chains are deprotonated. We have found that the chemical shifts of Sos, including $^1H^{\eta}$ and $^{15}N^{\eta}$ shifts, remain essentially unchanged in this sample. Therefore, we can conclude that Asp and Glu residues at the Sos binding interface are already deprotonated at pH 6.0.

The protein−peptide system thus generated was solvated using the explicit TIP3P solvent. A truncated octahedral water box was constructed such that the boundary of the water box was at least 12 Å away from any of the peptide or protein atoms. Given that the peptide is initially placed at considerable distance from the protein, the resulting water box contains a generous amount of water. As a result, the effective concentration of the protein in our simulations was relatively modest, 6−8 mM. This is not too far from the experimental NMR concentration. (The above description pertains to the setup with random placement of Sos. In the case when the initial configuration was based on 1CKB, we used the water box with a 14 Å margin. Because the 1CKB complex is sufficiently compact, the water box proved to be smaller, corresponding to an effective protein concentration 13−15 mM.)

The hydrated system was neutralized by adding one Na^+ ion (for the protein missing R191) and subjected to energy minimization (1000 steps using harmonic restraints with a force constant of 500 kcal mol^{-1} Å^{-2}, followed by 1000 steps with no restraints), then heated from 0 to 296 K, and equilibrated for 1 ns at 296 K. The production-stage MD simulations were conducted at 296 K using the NPT ensemble. During the simulations, all bonds involving hydrogen atoms were constrained using the SHAKE algorithm. The nonbonded cutoff was set to 8 Å. The integration step was 2 fs, and the coordinates were stored every 1 ps. The simulations were conducted using two GPU workstations, one equipped with four NVIDIA GeForce GTX480 cards and the other with four GTX580 cards (assembled by Electronics Nexus, Binghamton, NY, and Colfax International, Sunnyvale, CA). The production rate using the CUDA version of the pmemd program was in the range from 30 to 60 ns per day per card (depending on the size of the water box).

To adjust the strength of the salt bridge interactions in the Amber ff99SB*-ILDN force field, we have implemented one change in the original force field. The Lennard-Jones energy for a pair of atoms separated by distance r is parametrized in Amber as follows:

$$E_{\mathrm{L_J}} = \varepsilon\left[\left(\frac{r_0}{r}\right)^{12} - 2\left(\frac{r_0}{r}\right)^{6}\right] \quad (2)$$

where ε is the well depth and r_0 is the equilibrium distance. We have altered the r_0 pertaining only to certain specific atom pairs, namely the nitrogen−oxygen pairs involving N^{η} atoms in Arg side chains and $O^{\varepsilon}/O^{\delta}$ atoms in Glu/Asp side chains. The original Amber-generated value of r_0 for these atomic pairs has been multiplied by the scaling factor λ = 1.03. Because the GPU version of the Amber 11 pmemd program does not use look-up tables, this was accomplished by making a simple modification to the source code (kNLCPNE.h).

The three main MD trajectories discussed in this paper are as follows:

- (*i*) The 3.06 μs trajectory of Sos:wtSH3 starting from the crystallographic coordinates 1CKB recorded with the standard force field, λ = 1.00. This trajectory is termed MD(xray,λ = 1.00).
- (*ii*) The 3.20 μs trajectory of Sos:wtSH3 starting from the crystallographic coordinates 1CKB recorded with the altered force field, λ = 1.03. This trajectory is termed MD(xray,λ = 1.03).
- (*iii*) The 2.28 μs trajectory that represents the binding of Sos to wtSH3 starting with the random configuration as described above. This trajectory has been recorded using the altered force field, λ = 1.03, and is termed MD(rand, λ = 1.03).

In addition, to quantitate NMR resonance shifts that occur in Sos upon binding to wtSH3, we have also recorded the 2.1 μs trajectory of the free Sos (because the peptide contains no Asp or Glu residues, there is no difference between the original and altered force fields in this case). Other MD trajectories recorded as a part of this study are discussed where appropriate in the text.

Processing of the MD Trajectories and Prediction of NMR Parameters. The MD trajectories were manipulated using in-house C++ package trjtool. This package supports both text and binary formats for coordinate files and various output files, includes a convenient batch file processing mode, and offers a range of functions (extracting atomic coordinates and torsional angles, superimposing structures, evaluating correlation functions and order parameters, etc.). Some of the functionalities have been adapted for use with the GPU. The companion package pytrj, programmed in python, serves to calculate and plot various relaxation parameters (including ^{15}N relaxation rates and PREs). Both packages are available upon request.

To calculate chemical shifts, we have extracted every 10th frame from a given trajectory, processed these frames using SHIFTX2 version 1.07,[90] and then averaged the results. All computations were conducted using the SHIFTX+ module, which deals with the conformational dependence of chemical shifts. To investigate the effect of sampling, we repeated the calculations using all MD frames; the results were found to be virtually identical. Our focus is on the differential chemical shifts, $\Delta\delta_{\mathrm{Sos}} = \delta_{\mathrm{Sos}}^{\mathrm{bound}} - \delta_{\mathrm{Sos}}^{\mathrm{free}}$, which are termed binding shifts. In the calculations, the $\delta_{\mathrm{Sos}}^{\mathrm{bound}}$ component has been obtained from Sos:wtSH3 trajectories whereas $\delta_{\mathrm{Sos}}^{\mathrm{free}}$ has been obtained from the simulation of the free Sos. The effect of weak self-association of wtSH3 on $\Delta\delta_{\mathrm{Sos}}$ is inconsequential and has been ignored in these calculations.

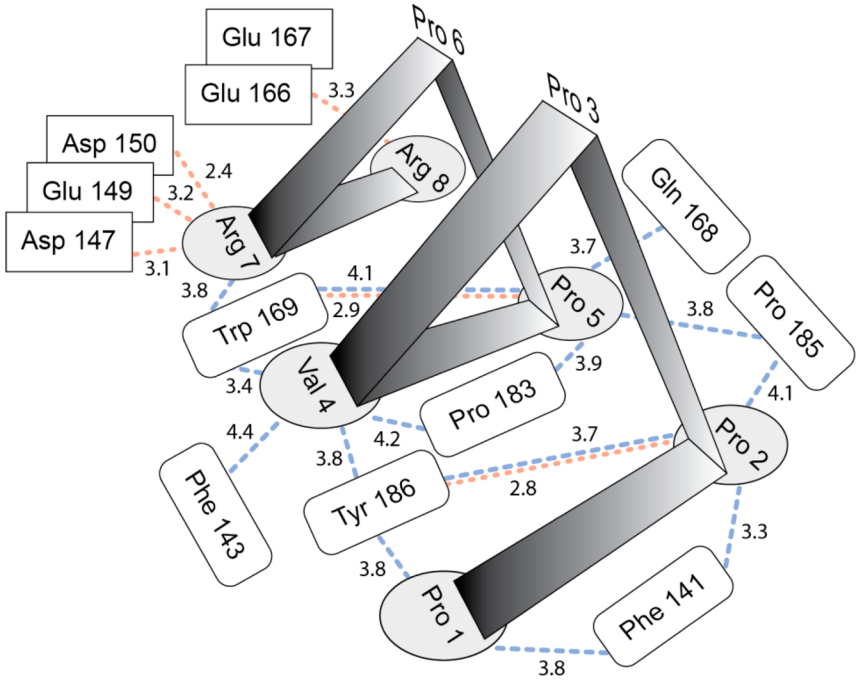


Figure 4. Binding of the Sos peptide to wild-type c-Crk N-SH3 as seen in the X-ray structure 1CKB.[47] The intermolecular distances indicated in the plot correspond to the nearest pairs of carbon atoms from hydrophobic residues (blue dashed lines) and nitrogen−oxygen pairs in the salt bridges or hydrogen bonds (pink dotted lines).

To calculate the relaxation rates, all MD frames were aligned with the crystallographic coordinates 1CKB by superimposing the protein coordinates (those of C^{α} atoms from the secondary structure residues). Note that only the protein coordinates have been superimposed; the motion of the peptide relative to the protein surface is fully preserved. The overall tumbling of the Sos:wtSH3 complex was then reintroduced using the experimentally determined correlation time, τ_{rot} = 7.0 ns (this value accounts for the effect of weak self-association as observed for wtSH3, see above). The autocorrelation functions have been computed for NH vectors using a sparse nonuniform time grid as described previously;[91] they were subsequently fit with six-exponential functions, multiplied by $\exp(-\tau/\tau_{rot})$, and then converted into spectral densities and further into relaxation rates.[92] In the case of CSA-dipolar cross-correlated cross-relaxation rate η_{xy}, the result was multiplied by $P_2(\cos\theta)$ assuming that the long axis of the ^{15}N CSA tensor makes an angle θ = 20° with the NH bond.[93] The length of the NH bond was assumed to be 1.02 Å, and the magnitude of the nitrogen chemical shift anisotropy was taken to be −172 ppm.

The in-house python script has been employed to analyze the appearance of salt bridges and hydrogen bonds in the MD trajectory. In the case of salt bridges, at least one nitrogen atom from the Arg side chain has been required to fall within 4 Å of the oxygen from the Glu/Asp side chain. In addition, the distance between the centroids of the two charged groups should not exceed 5 Å. This definition covers both high- and medium-strength ion−ion contacts, which mainly play a stabilizing role.[94] In the case of hydrogen bonds, we have required that the nitrogen−oxygen distance be <3.2 Å and the deviation from linearity not exceed 30°.[95]

Finally, the solvent-accessible surface area has been calculated on a per frame basis using the program POPS with the default parameter settings.[96]

RESULTS AND DISCUSSION

Overview of the Sos:wtSH3 Complex. The crystal structure of the complex between c-Crk N-SH3 and the Sos-derived peptide, PPPVPPRRRR, was determined by Wu and co-workers.[47] A schematic diagram of the binding interface is shown in Figure 4. The peptide is a class II ligand with consensus sequence PxhPx+ (x is any amino acid, h is a hydrophobic amino acid, and + is arginine or lysine).[97,98] Over most of its length, the peptide adopts a left-handed polyproline type II helix conformation. The helix has a triangular cross section; residues at the base of the triangle interact with the SH3 domain. The side chains of the two conserved prolines, P2 and P5, are tucked into the shallow hydrophobic grooves on the surface of SH3 (buried surface areas of 487 and 377 Å^2 for the peptide and protein, respectively). The grooves are defined by the side chains of conserved hydrophobic residues in the SH3 domain (represented by rounded boxes in Figure 4).

The positively charged C-terminus of the Sos peptide is delocalized and, therefore, partially absent from the crystallographic structure (no electron density for R10, R9, and the side chain of R8). The side chain of R7 is observed in the electron density map, even though crystallographic refinement reveals a substantial amount of variability at this site; the R7 guanidinium group forms salt bridges with carboxyl groups from D150, D147, and E149 (represented by rectangular boxes in Figure 4). One may expect that the side chain of R8 similarly interacts with E166 and E167,[47] although it must be significantly impacted by conformational disorder. Recall that electrostatic interactions involving the stretch of Arg residues in the Sos peptide anchor the putative encounter complex Sos·wtSH3. Given the dynamic nature of these interactions, this complex cannot be productively studied by X-ray crystallography. This situation calls for the development of new NMR-based methods.

dx.doi.org/10.1021/bi500904f | Biochemistry 2014, 53, 6473−6495

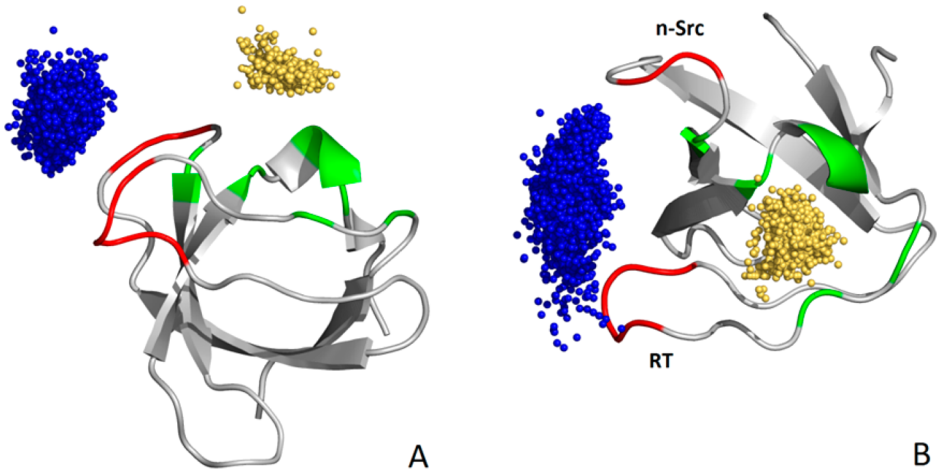


Figure 5. Position of the peptide Sos on the surface of wtSH3 according to the data from the 3.06 μs simulation MD(xray,λ = 1.00). To generate this plot, MD frames were first aligned with 1CKB by superimposing the SH3 coordinates (one frame per nanosecond). The peptide was then represented by the center of mass of its N-terminus (residues P1−P6, gold spheres) and C-terminus (residues R7−R10, blue spheres). In wtSH3, the charged residues at the binding interface are colored red (D147, E148, E149, and D150 in the RT loop and E166 and E167 in the n-Src loop) and the hydrophobic residues are colored green (F141, F143, W169, P183, P185, and Y186). Shown are the side view (A) and the top view (B) of the complex.

The dissociation constant (K_d = 5.2 ± 0.2 μM) has been previously reported for the binding of Sos to wtSH3 in the presence of 150 mM NaCl.[99] We were able to reproduce this result using the intrinsic fluorescence of residue W169 that resides at the ligand binding interface and thus is sensitive to Sos binding. Under identical conditions, our fluorescence titration data (not shown) yield a dissociation constant of 5.9 ± 0.2 μM. On the other hand, the ITC data collected under low-salt conditions, as also used in our NMR measurements, yield a K_d of 1.5 ± 0.1 μM (see Materials and Methods). As expected, the decrease in the ionic strength of the solution leads to stronger electrostatic attraction and, therefore, tighter binding.[100] The affinity of 0.1−10 μM is typical for binding of proline-rich peptides to SH3 domains and, more generally, for binding of peptides to their protein targets.[99,101] Such moderately strong binding is consistent with the signaling function that requires, on the one hand, a meaningful level of complex formation and, on the other, a rapid release of ligand in response to changing stimuli.

We have also analyzed the peak shapes from the NMR titration of the ^{15}N-labeled Sos with unlabeled wtSH3 (see Materials and Methods). The global fitting of the titration data using the fixed value of K_d as obtained in the ITC measurement allowed us to determine the on rate (k_{ON} = 1.5 × 10^9 M^{-1} s^{-1}). The off rate is then simply $k_{OFF} = K_d k_{ON}$. The approach in which ITC data are combined with NMR data to obtain a complete kinetic characterization of the system is similar to that employed by Demers and Mittermaier.[58]

The extremely high on rate obtained in our experiments is consistent with the prominent role of electrostatic interactions in bringing together Sos and wtSH3.[100,102−104] In fact, proteins that associate with rates >10^7 M^{-1} s^{-1} always seem to utilize the electrostatic steering mechanism. Previously, k_{ON} values of up to 5 × 10^9 M^{-1} s^{-1} have been measured for binding of barnase mutants to barstar.[100] An on rate as high as 1.7 × 10^{10} M^{-1} s^{-1} was recently found for the complex of the AD2 fragment from the intrinsically disordered N-terminal transactivation domain of tumor suppressor p53 with the zinc finger 2 domain of the CREB-binding protein.[105] These are extreme cases in which the proteins are propelled toward each other by strong electrostatic interactions.

Significant progress has been made in predicting k_{ON} based on atomic coordinates. In particular, the program TransComp produces results that are accurate to within 1 order of magnitude.[20] Although this program has been originally designed for folded proteins, it can also generate reasonable predictions for IDPs.[26] We used the selected MD frames as input for the TransComp calculations. Specifically, we have chosen the frames from the MD trajectory that represents the approach of Sos to wtSH3 and subsequent binding (see below). The k_{ON} rate estimated in this manner, ∼0.5 × 10^9 M^{-1} s^{-1}, is in good agreement with the experimental result.

The binding affinity is important from the perspective of NMR characterization. Our ^{15}N relaxation measurements and chemical shift determinations have been conducted on a sample containing 0.5 mM ^{15}N-labeled Sos and 2.0 mM unlabeled wtSH3. It is easy to verify that under these conditions 99.9% of Sos is bound. Therefore, the experimental NMR data pertain solely to the bound peptide.

Simulation of Sos:wtSH3 Using Amber ff99SB*-ILDN. Comparison with Experimental Data. The trajectory of Sos:wtSH3 has been recorded using the Amber 11 package with the ff99SB force field.[48,87] A number of recent comparative studies, in particular those based on the experimental NMR data, favor this force field over others.[106−111] In addition, recent corrections for selected backbone[49] and side-chain[50] torsional potentials have been employed. The revised force field is known as ff99SB*-ILDN.[112] The initial coordinates for the MD simulation have been taken from the crystallographic structure 1CKB. The C-terminal residues in Sos that are missing from the crystallographic model have been rebuilt using MODELLER.[88] The system was hydrated using TIP3P water.[113] Given that the C-terminal portion of Sos is dynamic and occasionally becomes lifted off the surface of wtSH3, we opted for a large water box (minimum 14 Å separation between any of the peptide or protein atoms and the water box boundaries). The resulting MD trajectory has a length 3.06 μs. In what follows, this simulation is termed MD(xray,λ = 1.00). The label indicates that the initial coordinates have been taken from the

dx.doi.org/10.1021/bi500904f | Biochemistry 2014, 53, 6473−6495

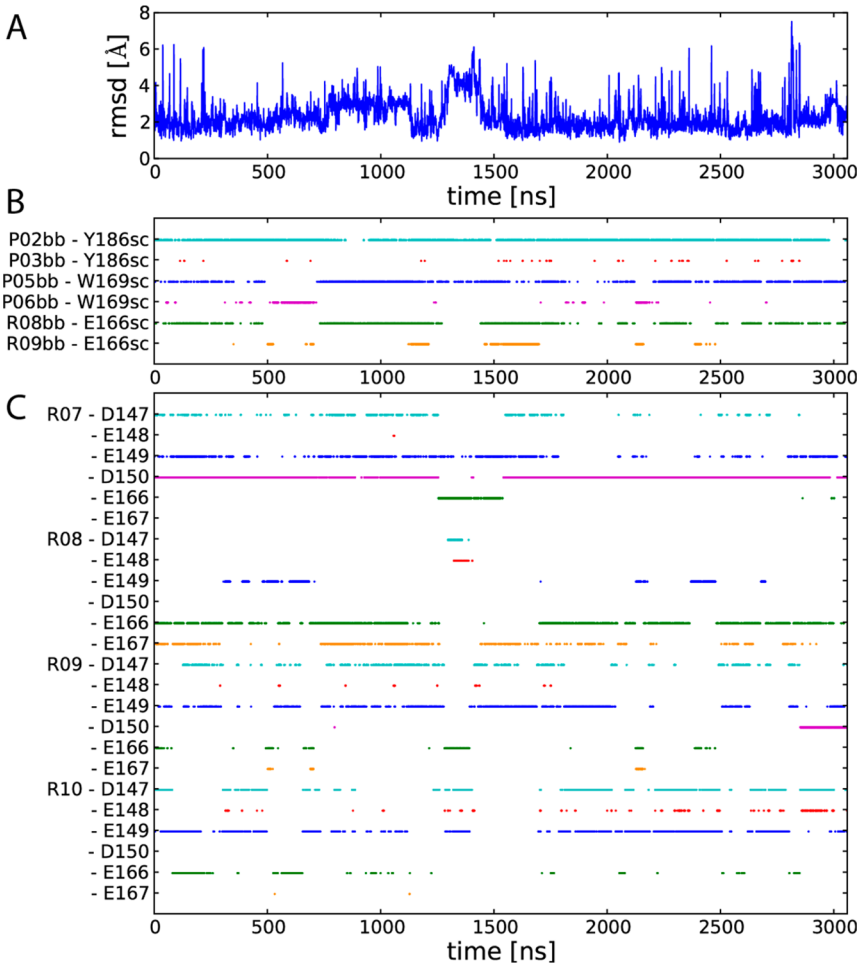


Figure 6. (A) rmsd of the peptide atomic coordinates relative to the crystallographic structure 1CKB as calculated for a 3.06 μs simulation of Sos:wtSH3 MD(xray,λ = 1.00). To generate this plot, MD frames were first aligned with 1CKB by superimposing the SH3 coordinates, and then the rmsd was computed for Sos atoms. Included in the rmsd calculation were all of the Sos atoms for which the crystallographic coordinates are available. (B) Peptide−protein hydrogen bonds and (C) salt bridges in the same Sos:wtSH3 trajectory (see Materials and Methods for definitions). The following procedure was used to generate these plots. The trajectory has been divided into 1 ns intervals, and the content of hydrogen bonds (salt bridges) has been determined for each interval. If the hydrogen bond (salt bridge) is present for more than half of the time during the 1 ns interval, then the dot is placed in the corresponding position in the graph. In generating panel B, we included only those hydrogen bonds that originate at the backbone of Sos and are populated at the level of at least 3%. In generating panel C, we have taken into consideration all possible combinations between the four Arg residues from the Sos peptide and the six Glu/Asp residues from the complementary acidic patch on the surface of the wtSH3 domain.

X-ray structure and the force field parameters have been used as is without any modifications (λ is the scaling coefficient that is discussed in detail below).

Figure 5 illustrates the dynamics of the Sos peptide bound to the surface of the wtSH3 domain. The gold spheres indicate the instantaneous position of the center of mass of the hydrophobic N-terminal segment of Sos (PPPVPP). The blue spheres correspond to the center of mass of the charged C-terminal segment (RRRR). It is apparent that the N-terminal portion of the peptide is well-localized, whereas the C-terminal portion moves a great deal. The data shown in the graph can be used to calculate the root-mean-square fluctuation for the respective mass centers, which can be conveniently expressed in a form of crystallographic *B* factors.[114] The values calculated for N- and C-terminal segments are 35 and 209 Å^2, respectively, underscoring the different dynamic status of the two regions.

The top panel in Figure 6 shows the root-mean-square deviation (rmsd) trace of the MD trajectory as calculated for the atoms of the Sos peptide using the crystal structure 1CKB as a reference. These results pertain to the ordered portion of the peptide for which the atomic coordinates are listed in the PDB file. The peptide clearly maintains the correct binding pose for the entire duration of the trajectory. For almost half of the time the rmsd remains below 2.0 Å, while the average rmsd value is just under 2.3 Å. The pattern of hydrogen bonds and salt bridges as seen in the crystal structure is also preserved in the MD simulation (Figure 6B,C). In particular, the signature hydrogen bonds P2−Y186 and P5−W169, as well as the R7−D150 salt bridge, persist throughout the simulation.

At the same time, occasional departures from the crystallographic structure are also observed. For example, at the point in time ca. 500 ns the peptide undergoes a certain amount of twist. The resulting conformation is stabilized by the P6−W169 hydrogen bond that replaces the canonical P5−W169 bond (see Figure 6B). This conformation survives for approximately 200 ns, before reverting to the original state. Another notable transition occurs at approximately 1300 ns. At this point in time, residue R8 moves away from E166 and E167 and instead

dx.doi.org/10.1021/bi500904f | *Biochemistry* 2014, 53, 6473−6495

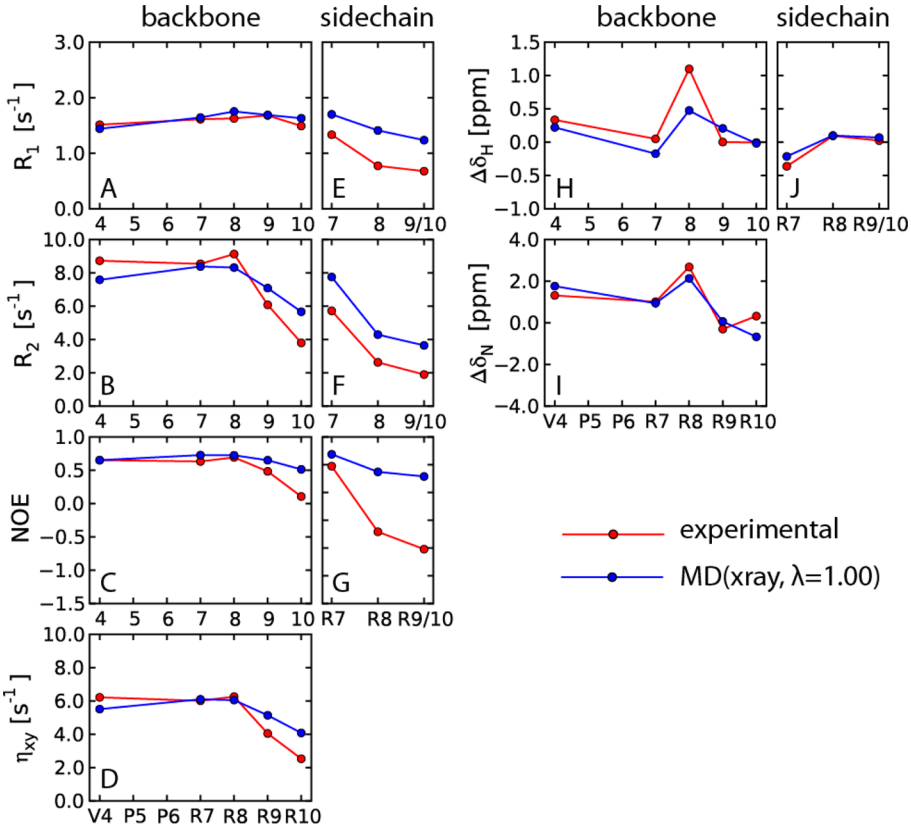


Figure 7. Comparison of the experimental (red) and simulated (blue) ^{15}N relaxation parameters and ^{1}H,^{15}N binding shifts for the Sos peptide in complex with wtSH3. Simulation MD(xray,λ = 1.00) uses the original Amber ff99SB*-ILDN force field and starts from the crystallographic coordinates 1CKB. Shown are the data from the backbone amide sites in the five non-proline residues in Sos, as well as the data from four arginine side-chain ^{1}H$^{\epsilon}$,^{15}N$^{\epsilon}$ sites. The ^{1}H$^{\epsilon}$,^{15}N$^{\epsilon}$ signals from R9 and R10 are overlapped in the NMR spectra; therefore, the corresponding experimental data effectively represent an average with respect to these two residues. To allow a comparison with experiment, the simulated data from these two sites have also been averaged. SHIFTX2 software used in our work does not predict ^{15}N$^{\epsilon}$ chemical shifts; therefore, the corresponding binding shifts have not been included in the comparison. The magnitude of the experimental errors is less than or comparable to the size of the symbols. To investigate the convergence properties of MD-based calculations, we recorded an additional 1.7 μs MD trajectory of Sos:wtSH3. The results (not shown) are similar to those plotted above. However, in the case of the binding shifts, the agreement is less than perfect because the simulated values are affected by rare events, such as the rupture of the R7−D150 salt bridge (cf. Figure 6).

engages D147 and E148. At the same time, the side chain of R7 shifts in the opposite direction, i.e., away from D150 and toward E166. This rearrangement leads to an appreciable increase in the rmsd, reaching ∼4 Å (see Figure 6A).

Of major interest for us is the dynamic behavior of the charged C-terminal tail of Sos, which is central to the formation of the electrostatic encounter complex. This arginine-rich tail is accommodated in the cleft between the RT loop and the n-Src loop, both of which carry stretches of negatively charged amino acids (see Figure 5B). The tail moves mostly in a sideways direction, sweeping over the surface of the protein. In doing so, it sporadically forms hydrogen bonds with wtSH3 as well as multiple salt bridges (see Figure 6B,C). Some of these interactions are stable and exist for a significant fraction of time, while others make only a brief appearance. Importantly, at each point in time, the C-terminal tail of Sos is tied to wtSH3 via a significant number of interactions (on average, 4.8 interactions, counting both salt bridges and the relevant hydrogen bonds in panels B and C of Figure 6). After a while, the tail rearranges itself, breaking some of its interactions with wtSH3 and establishing a new pattern of contacts.

As indicated above, the main focus of this study is on dynamics-modulated electrostatic interactions involving the C-terminal tail of Sos. Accordingly, the NMR data have been measured using the sample of ^{15}N-labeled Sos in the presence of unlabeled wtSH3. The sample was prepared with a 4-fold excess of wtSH3 to ensure that the spectral signals originate entirely from the bound Sos (see above). Using the standard ^{1}H,^{15}N experiments, we observed five resolved peaks from the non-proline backbone amide sites in Sos (see Figure 3A). In addition, we were also able to record the signals from four ^{1}H$^{\epsilon}$,^{15}N$^{\epsilon}$ sites in the arginine side chains, although two of the side-chain resonances are overlapped (R9 and R10). The experimental data from all these sites are plotted in Figure 7 as a function of residue number. Panels A−D and E−G of Figure 7 show ^{15}N relaxation data for the backbone and side-chain sites, respectively. Panels H−I and J show the binding shifts.

The standard set of ^{15}N relaxation data, comprised of R_1 and R_2 rates as well as ^{1}H−^{15}N saturation transfer NOEs, has been recorded as described in Materials and Methods. Initially, we reckoned with the possibility that R_2 rates may contain an exchange contribution R_{ex}, although we have not identified any exchange mechanisms that could potentially cause exchange broadening in Sos:wtSH3. From the perspective of MD modeling, the R_{ex} term represents a problem because it normally cannot be simulated (unless an ultralong trajectory is available).[115] To address this issue, we chose to measure transverse cross-correlated relaxation rates η_{xy} that carry the

dx.doi.org/10.1021/bi500904f | Biochemistry 2014, 53, 6473−6495

same dynamic information as R_2 but are free from μs−ms exchange effects.[116,117] All relaxation data have been collected at two static magnetic field strengths, 500 and 600 MHz. For the sake of clarity, only the latter data set is illustrated in Figure 7 (the discussion of field dependence is deferred until later).

In addition, we have also quantified ^{1}H and ^{15}N peak shifts that occur in the spectrum of Sos upon binding to wtSH3 (cf. Figure 3A). A number of prominent shifts have been observed. In particular, a large downfield shift of the amide resonance from R8 (+1.1 ppm on $^1H^N$ and +2.7 ppm on ^{15}N) is due to the hydrogen bond[118] with the side-chain carboxyl group of E166. A sizable upfield shift of the signal from the R7 side chain (−0.4 ppm on $^1H^\varepsilon$ and −1.7 ppm on $^{15}N^\varepsilon$) is caused by a ring current in W169, with which R7 forms a cation−π contact. Likewise, the downfield shift of the amide signal from V4 (+0.3 ppm on $^1H^N$ and +1.3 ppm on ^{15}N) is caused by the ring current from Y186. At the same time, residues R9 and R10 show very small binding shifts; apparently, these two residues remain extended into solvent and highly disordered. Consistent with crystallographic data, two terminal arginines are only transiently involved in the interactions with wtSH3, although they contribute to the kinetics of Sos−wtSH3 association through the initial electrostatic attraction.

This array of experimental data offers a good opportunity to put our MD model to a rigorous test. There is a long history of validating MD simulations via spin relaxation rates.[106,119−121] Briefly, the MD data are used to directly compute the dipolar (as well as CSA) temporal correlation functions for the individual amide sites. These correlation functions are subsequently converted into spectral densities, which are in turn used to calculate ^{15}N relaxation parameters. Spin relaxation is directly influenced by fast internal dynamics, i.e., the motions with correlation times on the order of τ_{rot} or shorter, and indirectly by slower motional processes. For example, ^{15}N relaxation can be indirectly affected by rare conformational transitions as seen in Figure 6. Indeed, formation and dissolution of salt bridges are expected to have an effect on the fast local dynamics of Sos bound to the surface of wtSH3. From this perspective, it is important that the MD simulation is sufficiently long, 3.06 μs, to sample such relatively rare events.

To calculate binding shifts, we used the knowledge-based prediction program SHIFTX2.[90] The program was first applied to the MD trajectory representing free Sos and then to the trajectory in which Sos is bound to wtSH3, thus allowing calculation of $\Delta\delta_{Sos} = \delta_{Sos}^{bound} - \delta_{Sos}^{free}$. A number of studies have appeared recently where chemical shift predictors have been applied to the frames from the MD simulations.[110,122,123] Moreover, a new generation of predictors has been developed specifically for this purpose.[124,125] Note, however, that these previous applications have a somewhat different focus—namely, they seek to improve the prediction of chemical shifts by taking into consideration the local dynamics of proteins. In our work, the calculated shifts are used to validate the dynamic scenario seen in the MD simulation.

In particular, $\Delta\delta$ values are sensitive to the (dynamically modulated) hydrogen bonds, cation−π contacts, and salt bridges such as those formed by several residues in the C-terminal portion of Sos. Hence, the comparison of the calculated and experimental shifts should shed light on these interactions and their dynamic status (degree of motional averaging). Of note, SHIFTX2 employs well-established and reliable algorithms for calculating shifts caused by hydrogen bonds and ring currents. This makes it well-suited for the purpose of our study, i.e., to probe the interactions between Sos and wtSH3.

We have also measured binding shifts in the reverse direction, i.e., the shifts in ^{15}N-labeled wtSH3 upon binding of unlabeled Sos. The results are graphically represented in Figure S2 of the Supporting Information. As expected, the most significant chemical shift perturbations are localized in the vicinity of the binding site. In principle, it should be possible to use $\Delta\delta_{SH3}$ to further validate the MD data, along the same lines as what has been done with $\Delta\delta_{Sos}$. In practice, however, we have found that such analysis is of no value. None of the backbone amide groups in wtSH3 forms a hydrogen bond with Sos. Furthermore, Sos contains no aromatic rings. Thus, two of the largest sources of binding shifts are eliminated in this case. As a consequence, the measured $\Delta\delta_{SH3}$ values are small to moderate; none of the shifts exceeds 0.5 ppm in the $^1H^N$ dimension or 1.3 ppm in the ^{15}N dimension. Both the MD simulation and the SHIFTX2 algorithm are unable to properly capture and process the subtle structural and dynamic changes that cause shifts of this magnitude. As a result, the computed $\Delta\delta_{SH3}$ values display little or no correlation with the experimental data (r = 0.40 for $^1H^N$ shifts, and r = 0.19 for ^{15}N shifts). This cannot be construed as evidence against the MD model but rather reflects the limitations of the method.

Figure 7 presents the summary of Sos ^{15}N relaxation data and the binding shifts, as measured experimentally (red symbols) and calculated on the basis of the MD trajectory (blue symbols). Let us first focus on the backbone relaxation data from residues V4, R7, and R8. These data are reasonably well reproduced by molecular dynamics. Although certain discrepancies are seen in the R_2 plot (panel B), the agreement is nearly perfect for all other parameters, including transverse cross-correlated relaxation η_{xy} (panel D). As already mentioned, η_{xy} is one of the more accurate measures of protein dynamics that is free of the exchange effect; in terms of information content, η_{xy} is equivalent to R_2. Generally, the backbone amide sites of V4, R7, and R8 can be categorized as tightly bound. They are essentially integrated with the rigid wtSH3 scaffold; their spin relaxation is controlled by the overall tumbling motion of the complex and, to a certain degree, small local fluctuations.

On the other hand, the backbone relaxation data for residues R9 and R10 as well as side-chain data for R7, R8, R9, and R10 show clear systematic deviations between the experiment and the MD model (panels A−G). Specifically, the simulated R_2 and η_{xy} rates are higher than the corresponding experimental values. Likewise, the simulated $^1H^N$−^{15}N saturation transfer NOEs are higher than the corresponding experimental values. This clearly indicates that MD simulation underestimates the amount of local mobility for all of these sites.

According to the MD model, the C-terminal tail of Sos is attached to wtSH3 through the network of salt bridges. While this network is occasionally rearranged, it still has a strong restraining effect on the tail (see Figures 5 and 6). As a result, the relaxation properties of the tail are rather similar to those of the rest of the complex, which is well-structured and can be viewed as rigid (cf. blue profiles in panels A−G). The experimental data, however, paint a rather different picture. It appears that the pair of C-terminal residues in Sos and all the arginine side chains enjoy a degree of motional freedom much greater than that suggested by the MD simulation. This dynamic scenario is also supported by the crystallographic evidence (reviewed above). It is also worth noting that the MD simulation fails to accurately reproduce the large binding shift

of $^{1}H^{N}$ in residue R8 (panel H). This observation implicates the side chain of residue E166, which forms an on−off hydrogen bond with this amide as well as transient interactions with several other sites in Sos.

Surveying the results shown in Figure 7, we hypothesize that the root cause of the observed discrepancies is the deficiency in the MD force field, specifically, the exaggerated strength of the salt bridge interactions.[51,52] This issue is discussed at length below.

Simulations of Sos:wtSH3 Using Alternative MD Setups. Two groups of force field parameters are of principal significance for ionic (salt bridge) interactions: partial charges and Lennard-Jones parameters (see the next section). As far as partial charges are concerned, there is a considerable amount of variability between different families of force fields, as well as different generations in the same family. For instance, the charge assigned to N^{η} atoms in the arginine side chain is −0.62 in CHARMM22*, −0.80 in CHARMM27, −0.86 in Amber ff99SB, −0.69 in Amber ff03, and −0.26 in GROMOS53A6. There are also significant variations in the Lennard-Jones parameters between different force fields; these variations are largely responsible for dramatic differences in the radial distribution functions associated with salt bridge interactions.[51] A very recent study by Debiec et al. used small molecules (guanidinium and acetate) to model Arg−Asp/Glu salt bridges.[52] It has been found that six popular force fields, including CHARMM22*, CHARMM27, Amber ff99SB, Amber ff03, and the newest Amber ff13α force field, all overestimate the population of such salt bridges by a factor of roughly 2. This result essentially does not depend on the choice of water model (TIP3P, SPC/E, TIP4P-Ew, etc.).

While the results of Debiec and co-workers strongly suggest that there is a problem with parametrization of salt bridge interactions, these results are relevant for small molecules and thus may not accurately reflect the situation in peptides and proteins. Indeed, it is conceivable that in protein simulations salt bridge interactions are partially offset by the interactions between charged side chains and backbone amides. If true, this may have an effect of partial error cancellation. Therefore, it should be instructive to investigate this problem directly in the context of protein MD simulation employing different force fields.

With this goal in mind, we have tested AMBER ff03 (Amber 11 platform), CHARMM22* (Desmond platform), GROMOS53A6 (GROMACS 4.5.5 platform), and AMOEBA polarizable force field (TINKER 6.0.05 platform using the OpenMM library to support the GPU acceleration).[126−130] In each case, we have recorded MD trajectories of Sos:wtSH3 using the crystallographic structure as a starting point; the durations of the simulations were 445, 345, 108, and 19 ns, respectively. These trajectories were subsequently analyzed in the same fashion as described above. None of these trials produced the results that we had hoped for. AMBER ff03 shows less propensity to form salt bridges. However, three native hydrogen bonds connecting the backbone of Sos with side chains of wtSH3 are almost completely lost in this simulation, leading to poor predictions of chemical shifts. Similar behavior has been noted for this force field previously.[131] In the CHARMM22* trajectory, two of the native hydrogen bonds are greatly weakened, while arginine side chains remain over-restrained. Chemical shift predictions using this trajectory also prove to be poor. Finally, GROMOS 53A6 significantly underestimates the strength of electrostatic interactions: all salt bridges, including the native R7−D150 bridge, virtually disappear, whereas hydrogen bonds become transient and weak. As a consequence, the bound peptide experiences rapid dynamic fluctuations; in turn, this leads to large systematic deviations between the predicted and experimental ^{15}N relaxation rates.

Of special interest are the results from the 19 ns trajectory of Sos:wtSH3 recorded with the AMOEBA polarizable force field. The polarizability effects are deemed to be important in the context of salt bridge formation,[132−134] and polarizable force fields have demonstrated their usefulness in the context of protein−ligand binding and protein structure refinement.[135−139] Although there have been reports of the protein instability in AMOEBA simulations,[130] in our case the protein structure remains intact (within 1.5 Å of the crystal coordinates viz. the C^{α} rmsd). However, three native hydrogen bonds linking Sos and wtSH3 are absent in this simulation, and there are only two salt bridges, R7−D150 and R7−D147. Chemical shift predictions have been only partially successful, and there are indications that the C-terminal segment of Sos remains over-restrained. Given that the simulation is short on the absolute scale, these observations should not be over-interpreted; because of the high computational cost, we have not attempted to extend this AMOEBA trajectory.

In a further attempt to improve the situation, we experimented with different water models. The choice of water model is known to affect protein hydration[140−142] as well as conformational preferences and dynamics of disordered peptides.[143,144] Using the Amber ff99SB*-ILDN force field, we have tested TIP3P, SPC/E, and TIP4P-EW models.[113,145,146] In addition, we have also tested a polarizable model SWM4-NDP (supported in the GROMACS package).[147] No significant improvement has been observed in any of these trial trajectories, consistent with the recent findings.[52] We have also addressed several other details pertaining to the MD procedure. For example, phosphate ions have been included in the MD setup to reflect the presence of 20 mM phosphate in the NMR buffer. In principle, phosphate ions can form strong interactions with Arg side chains,[148] thus competing with the Sos:wtSH3 salt bridges. However, this tactic also failed to produce any improvements. Likewise, increasing the non-bonded cutoff distance from 8 to 10 Å had no effect on the simulation results. While none of these trials can be considered in any way definitive, we note that our results are consistent with the recent small molecule studies.[51,52]

Finally, we have tested the so-called MDEC (molecular dynamics in electrostatic continuum) model.[149−151] This model focuses on electrostatic interactions involving ionized side chains. In all conventional force fields, ionized side chains are assigned a net charge of +1 or −1 and their electrostatic interactions are calculated accordingly. The authors of the MDEC model argue that this approach fails to take into consideration the effect of electronic dielectric screening. To remedy this defect, the authors suggest to scale the respective point charges with the factor $\alpha = 1/(\varepsilon_{el})^{1/2}$, where the electronic screening constant for a protein in aqueous solution is estimated to be $\varepsilon_{el} = 2$. We have implemented the MDEC strategy in the Amber ff99SB force field and subsequently recorded a number of trial trajectories with different settings of α. The best results were obtained for $\alpha = 0.9$, which is significantly different from the prescribed value ($\alpha = 0.7$). The MDEC approach is admittedly empirical; a rigorous implementation of this concept would require consistent reparamet-

dx.doi.org/10.1021/bi500904f | Biochemistry 2014, 53, 6473−6495

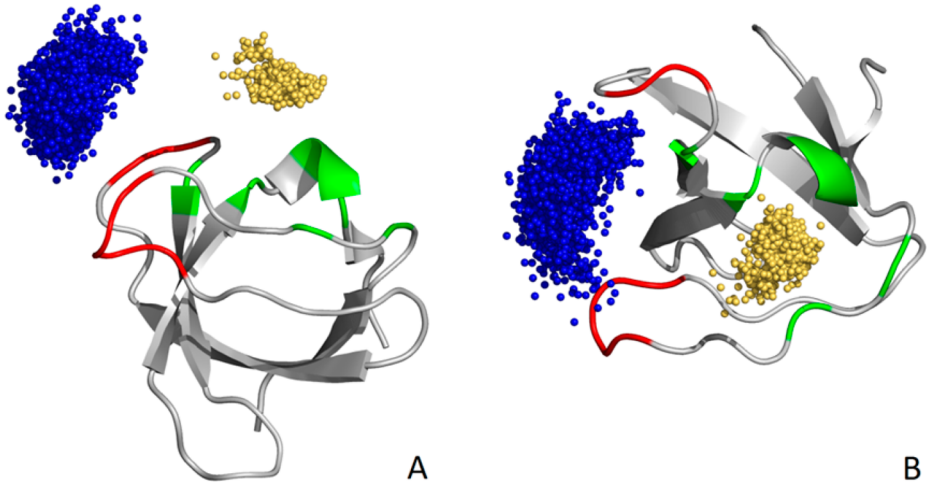


Figure 8. Position of the peptide Sos on the surface of wtSH3 according to the data from the 3.20 μs simulation using the modified force field, MD(xray,λ = 1.03). The plotting conventions are the same as in Figure 5.

rization of all force field parameters.[149] The rescaling of side-chain point charges should have an effect on many aspects of the system, including side-chain solvation, in a manner that is not easily predictable. Therefore, we abandoned this strategy in favor of a more targeted approach, as described in the next section.

Adjustment to Amber ff99SB*-ILDN. Fine-Tuning Salt Bridge Interactions. The interaction between two charged atoms is quantum mechanical in nature. When the two atoms are far apart, the interaction energy can be well approximated by the classical Coulomb's law. In contrast, when they approach each other such that the electron clouds start to overlap, the interaction becomes much more complex. In particular, there is an effect of induced polarization that gives rise to attractive dispersion forces; at short distances, there is also a strong repulsive force due to Pauli interaction. In classical molecular dynamics, all deviations from the standard Coulomb's law are empirically modeled by the Lennard-Jones (LJ) potential, eq 2. For the pair of atoms indexed i and j, the potential is defined by two parameters: the interatomic distance at which the LJ energy is at its minimum, r_0^{ij}, and the depth of the corresponding energy well, ε^{ij}. After the atomic charges are set, the LJ parameters determine how close the two atoms can approach each other and are thus strongly related to the magnitude of their optimal interaction energy.

In conventional MD force fields, the pairwise LJ parameters are normally determined by simple combination rules. In particular, each atom type is assigned a van der Waals radius, and the sum of the two radii for any two atoms is taken to be the distance r_0^{ij} for this particular pair. It has been increasingly recognized, however, that in certain cases the simple combination rules require additional fine-tuning to reproduce the desired interatomic interactions. For example, a slight increase in r_0^{ij} between Na^+ (K^+) and Cl^- (over the default value derived from the standard combination rule) made it possible to correct for the otherwise excessive ion pairing in the simulations of simple electrolytes.[152] The LJ parameters for the cation − carbonyl oxygen pairs have also been refined to better reproduce the energetics of K^+ ions in a protein channel.[153] Similarly, a small adjustment in r_0^{ij} has been used to avoid excessively strong binding of the cations to the negatively charged lipid headgroups, leading to an excellent agreement with experimental data.[154] The same approach has been employed to fine-tune the interactions between various metal ions and DNA phosphates, thus avoiding clustering artifacts in the DNA simulations.[155] Other very recent examples include interaction of Mg^{2+} with di- and triphosphate nucleotides,[156] as well as interaction between Ca^{2+} and acetate (viewed as an analogue of Asp and Glu side chains).[157] Finally, pairwise-specific LJ parameters have been implemented for aqueous solutions of alkanes, alcohols, and ethers in the CHARMM Drude polarizable force field, leading to accurate values of the corresponding hydration free energies.[158]

Here we use this philosophy to correct for the effect of overstabilized Arg−Asp/Glu salt bridges in Amber ff99SB*-ILDN simulations. Toward this end, a number of trial trajectories have been recorded where the Lennard-Jones r_0 distance for $N^{\eta}-O^{\delta/\varepsilon}$ pairs has been rescaled (see Materials and Methods). Specifically, we have tested scaling factors λ = 1.01, 1.02, 1.03, 1.04, 1.05, 1.10, 1.15, and 1.20. In each case, the MD data have been processed along the same lines as shown above (Figures 5−7). Unsurprisingly, large corrections to r_0 led to dramatic attenuation of salt bridges. A good indicator is the native salt bridge between R7 and D150, which is observed in the X-ray structure 1CKB. This salt bridge is present for 86% of the time in the original simulation, λ = 1.00 (trajectory length 3.06 μs), 70% of the time in the altered simulation with λ = 1.03 (3.2 μs), 24% of the time in the altered simulation with λ = 1.05 (0.8 μs), and is completely lost in the simulations with λ = 1.10 or higher. This result, as well as direct comparison of the simulated and experimental ^{15}N relaxation rates and binding shifts, led us to identify λ = 1.03 as the optimal scaling factor. While this is clearly an ad hoc choice, we provide an independent validation for it in paper 2, where we investigate the fuzzy complex between Sos and the double mutant of c-Crk N-SH3 (Sos·dmSH3).

It is worth noting that LJ distance parameters have been reoptimized together with partial charges in the new Amber force field ff13α. In particular, the r_0 distance for $N^{\eta}-O^{\delta/\varepsilon}$ pairs increased by 6%.[159] This new force field remains a work in progress; in its current state, Amber ff13α still shows a tendency to significantly overestimate the strength of salt bridge interactions.[52] Nevertheless, it is clear that any solution to this problem is likely to involve the Lennard-Jones parameters.

dx.doi.org/10.1021/bi500904f | *Biochemistry* 2014, 53, 6473−6495

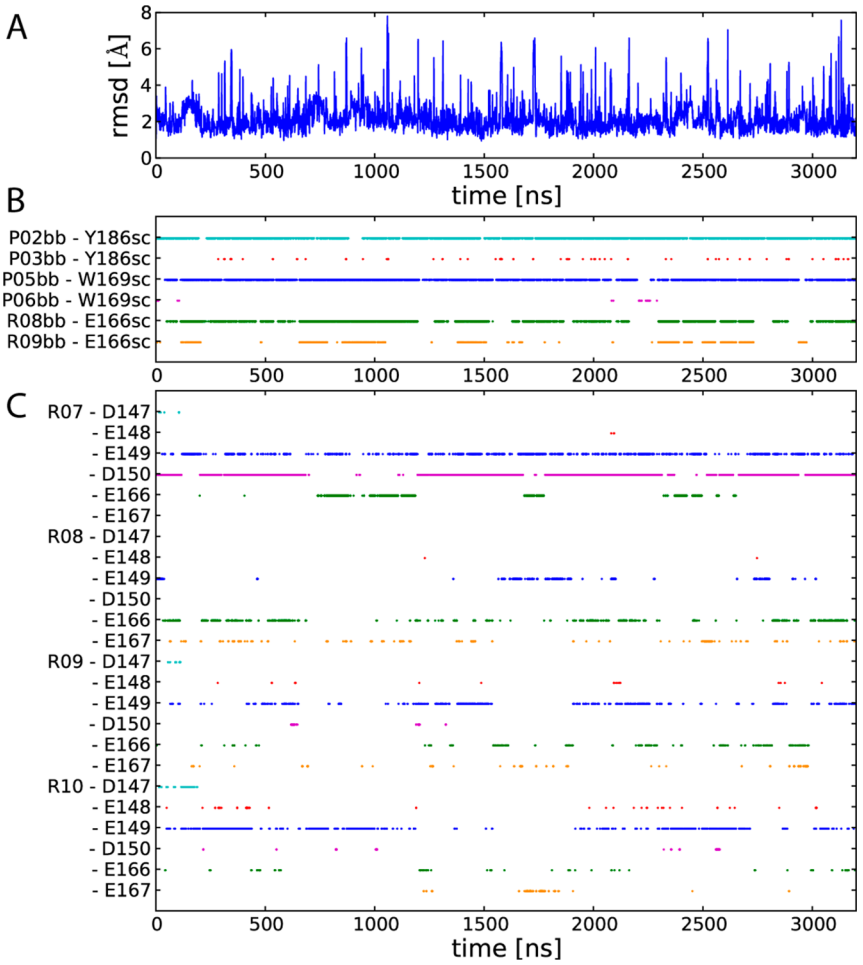


Figure 9. (A) rmsd of the peptide atomic coordinates relative to the crystallographic structure 1CKB as calculated for the 3.20 μs simulation of Sos:wtSH3 MD(xray,λ = 1.03). (B and C) Time traces representing the occurrence of peptide−protein hydrogen bonds and salt bridges in the trajectory. All conventions are the same as in Figure 6. Prominent hydrogen bonds and salt bridges are the same as those found in the crystallographic structure 1CKB (see Figure 4).

Finally, it is worth mentioning that MD simulations of proteins using implicit solvent are also faced with the problem of overstabilized salt bridges. This problem has been solved by adjusting the Born radii of the atoms that form the salt bridge.[160−162] Conceptually, this approach is similar to that described above.

Simulations of Sos:wtSH3 Using a Modified Version of Amber ff99SB*-ILDN. Comparison with Experimental Data. The calibration of Lennard-Jones parameters for N^{η}−$O^{\delta/\varepsilon}$ pairs, as discussed in the previous section, favors the distance r_0^{ij} that is 3% longer compared to the original setting. The new r_0^{ij} value has been coded into the ff99SB*-ILDN force field, thus overriding the standard combination rule for these particular off-diagonal LJ terms. Using the amended force field, we have recorded a 3.20 μs trajectory of Sos:wtSH3 beginning from the crystal coordinates. In what follows, this simulation is referred to as MD(xray,λ = 1.03) according to the choice of scaling factor λ. Aside from the change in the force field, all other details of the MD protocol were exactly the same as in the original 3.06 μs trajectory, MD(xray,λ = 1.00). The new simulation is analyzed using the same format as before; the results are shown in Figures 8−10.

Figure 8 illustrates the fluctuations of the Sos peptide when attached to the canonical binding site on the surface of wtSH3. Gold and blue spheres in the graph represent the centers of mass of the hydrophobic N-terminal segment of Sos and the arginine-rich C-terminal segment, respectively. The results are rather similar to those previously obtained from the conventional simulation (cf. Figure 5), although one can also notice certain differences. In the new trajectory, the hydrophobic portion of Sos is somewhat better localized in the binding site. The scatter of the gold spheres in Figure 8 corresponds to a temperature factor of 31 Å^2, which is slightly lower than the previously obtained value of 35 Å^2. At the same time, the arginine-rich portion of the peptide acquires a somewhat greater degree of motional freedom, 229 Å^2 vs 209 Å^2.

These observations lead us to conclude that there is a certain amount of strain in the Sos:wtSH3 complex. Specifically, formation of strong salt bridges by the C-terminal portion of Sos has a certain destabilizing effect on the binding of the N-terminal portion of Sos. Conversely, weakening of the salt bridges in the λ = 1.03 simulation creates conditions for tighter binding of the N-terminal segment. This kind of behavior is also manifested in Figures 6A and 9A. Shown in these graphs is the rmsd trace of the Sos peptide, calculated for the structured portion of Sos (those atoms for which the crystallographic coordinates are available). The graph from MD(xray,λ = 1.03) simulation clearly displays less dynamic fluctuations; in particular, there are no large-amplitude long-lived fluctuations such as those observed in the MD(xray,λ = 1.00) trajectory.

dx.doi.org/10.1021/bi500904f | Biochemistry 2014, 53, 6473−6495

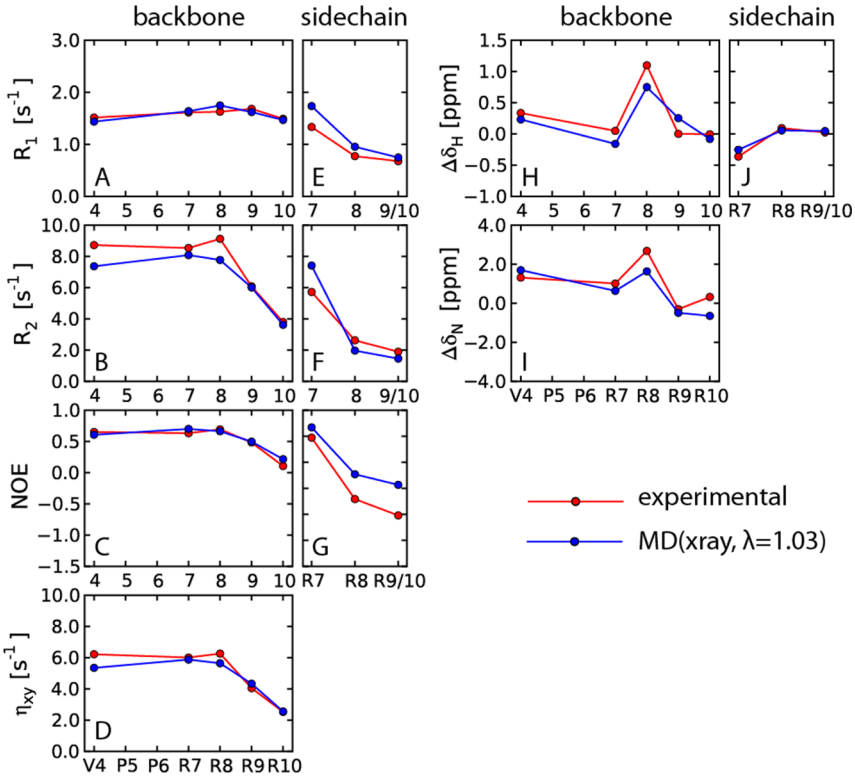


Figure 10. Comparison of the experimental (red) and simulated (blue) ^{15}N relaxation parameters and ^{1}H,^{15}N binding shifts for the Sos peptide in complex with wtSH3. The MD(xray,λ = 1.03) simulation uses the modified version of the Amber ff99SB*-ILDN force field in which the strength of salt bridge interactions involving Arg and Glu/Asp side chains has been adjusted.

The average Sos rmsd in the new simulation is 2.1 Å, somewhat lower than that registered previously, 2.3 Å. Hence, reducing the strength of salt bridges appears to enhance peptide binding, which is associated primarily with the well-structured N-terminal portion of the peptide.

An interesting picture emerges from these observations. The electrostatic attraction associated with Arg−Glu/Asp pairing pulls the peptide toward its target, thus efficiently increasing the k_{on} rate. However, once the peptide is bound the same electrostatic interactions involving the C-terminal portion of Sos may no longer be conducive to binding because these interactions are not necessarily compatible with hydrophobic contacts and hydrogen bonds formed by the N-terminal portion of the peptide. In other words, the propensity to form salt bridges may cause an increase in k_u (see Figure 1), thus slightly degrading the binding affinity. This highlights the dual role of electrostatic interactions in the context of peptide binding; they are responsible for dramatic gains in the on rate, but they may also have a certain destabilizing effect on the bound state.

The traces shown in panels B and C of Figure 9 illustrate the time dependence of hydrogen bonds and salt bridges connecting Sos and wtSH3 throughout the course of the MD(xray,λ = 1.03) simulation. Obviously, the propensity of the C-terminal tail to form salt bridges is greatly diminished. On average, at any given point in time, the C-terminus forms 2.3 salt bridges. This is a 2-fold decrease compared to the original simulation, where the corresponding number was 4.3. At the same time, the hydrogen bond contacts are somewhat strengthened, 2.1 in the altered simulation versus 1.7 in the original simulation. As already discussed above, the reduced propensity to form salt bridges removes some of the steric strain, which, in turn, leads to better hydrogen bonding.

Another way of looking at it would be to determine the number of contacts formed by the structured portion of the peptide (for which the crystallographic coordinates are available) and the remaining unstructured segment. For the structured portion of Sos, the average number of contacts is 3.0 as determined from the MD(xray,λ = 1.03) trajectory. This includes the native salt bridge R7−D150, which is largely preserved in the new simulation, see Figure 9C. The result is quite similar to the one obtained from the original trajectory MD(xray,λ = 1.00), where the average number of contacts is found to be 3.1. At the same time, the data from the unstructured portion of Sos display some significant differences. The altered MD simulation predicts, on average, 1.3 contacts in this area (salt bridges as well as hydrogen bonds), whereas the original trajectory predicts 2.9 such contacts.

Similar conclusions can be drawn from the analysis of the solvent accessible surface areas. Considering the N-terminal segment of Sos, the average buried surface area is essentially the same in the two simulations, 546 and 544 Å^2. At the same time, there is a significant difference in how the C-terminal segment is packed against the protein surface—the packing is less tight in the modified simulation protocol, 439 Å^2 vs 574 Å^2.

Finally, the most important element of the validation scheme is the direct comparison between the MD-derived NMR parameters and the experimentally measured data (Figure 10). A brief survey of these results indicates a big improvement compared to the standard implementation of molecular dynamics (Figure 7). In particular, the relaxation rates of the backbone amide groups belonging to residues R9 and R10 are now in perfect agreement with the experiment. The arginine

dx.doi.org/10.1021/bi500904f | Biochemistry 2014, 53, 6473−6495

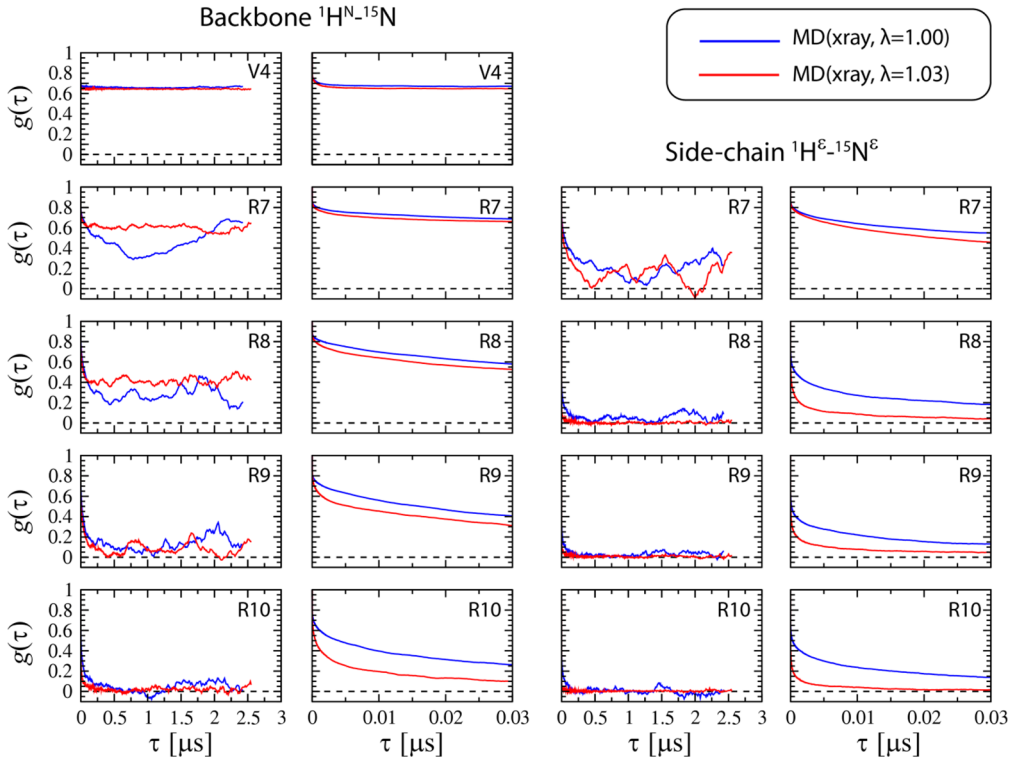


Figure 11. $^1H-^{15}N$ dipolar correlation functions $g(\tau)$ for backbone and arginine side-chain sites in Sos as extracted from MD(xray,λ = 1.00) and MD(xray,λ = 1.03) trajectories of Sos:wtSH3 (blue and red curves, respectively). The correlation functions have been calculated to 80% of the trajectory length. The expanded view of the initial portion of $g(\tau)$, which is relevant for spin relaxation, is shown in the second and fourth columns. The results reflect internal dynamics of the Sos:wtSH3 complex, with the overall reorientational motion (tumbling) factored out as described in Materials and Methods. Although it is difficult to tell with certainty, the plateau of the correlation function for backbone amides in residues R7 and R8 appears to be *lower* in the MD(xray,λ = 1.00) simulation than in the MD(xray,λ = 1.03) simulation. If this is true, this means that weakening of the salt bridges in the latter trajectory actually has a partial stabilizing effect on those two residues (similar to what has been discussed in relation to Figure 8). However, for the sake of discussion, we simply suggest here that the plateau values are similar within the uncertainty of calculations. Finally, note that the MD(xray, λ=1.03) trajectory, which features faster dynamics, shows better convergence properties than its MD(xray, λ=1.00) counterpart.

side-chain results for residues R8 and R9/R10 are also greatly improved. Furthermore, there is also a noticeable improvement in binding shifts; specifically, the large shift experienced by $^1H^N$ from residue R8 is now successfully reproduced by the MD-based calculations. At the same time, for those sites that belong to the well-structured portion of the peptide, e.g., V4, the level of agreement between the simulations and experiment remains unchanged (cf. Figures 7 and 10).

What is the structure/dynamics basis for the improvement observed in Figure 10? To answer this question, we turn to dipolar correlation functions underlying the ^{15}N relaxation rates. The full complement of these correlation functions, for both backbone and Arg side-chain sites in Sos, is shown in Figure 11. The plots in the first and third columns show the correlation functions $g(\tau)$ computed to 80% of the trajectory length. In each case, $g(\tau)$ are supposed to converge to a plateau, reflecting the long-term averaging properties of the particular $^1H-^{15}N$ vector. Although the convergence behavior is far from perfect, the data are generally sufficient to judge the amount of disorder at each site. As far as we can tell, the plateau values derived from MD(xray,λ = 1.00) and MD(xray,λ = 1.03) trajectories are similar within the uncertainty associated with incomplete convergence (cf. blue and red curves, first and third columns in Figure 11). Consider, for instance, the data for residues R9 and R10. The backbone and side-chain correlation functions in R9 and R10 essentially decay to zero. This means that in both trajectories the respective proton−nitrogen vectors experience complete orientational averaging.

The observations concerning the plateaus of the correlation functions (i.e., the order parameters) lead us to conclude that the amplitudes of internal motions in MD(xray,λ = 1.00) and MD(xray,λ = 1.03) trajectories are actually similar. This conclusion agrees with the analysis illustrated in Figures 5 and 8, where the differences between the two trajectories are relatively minor. Therefore, it is not the motional amplitudes that cause the differences in the simulated ^{15}N relaxation rates. Rather, it is the time scale of internal dynamics that plays the key role. The panels in the second and fourth columns of Figure 11 illustrate the expansion of the simulated correlation functions to 30 ns. This is the time interval that is relevant for spin relaxation in the Sos:wtSH3 complex where the tumbling time is τ_{rot} = 7 ns. From these expansion plots, it is obvious that the internal dynamics in Sos is faster in the altered MD(xray,λ = 1.03) simulation than it is in the original MD(xray,λ = 1.00) simulation. This observation is relevant for C-terminal residues, R8−R10; the effect is particularly striking for arginine side-chain sites (fourth column in Figure 11).

The analysis described above suggests the following picture of the simulated Sos dynamics. In the MD(xray,λ = 1.00) trajectory, the C-terminal segment of Sos samples different conformations, where it often becomes trapped for extended periods of time because of the overstabilized salt bridges. In

dx.doi.org/10.1021/bi500904f | *Biochemistry* 2014, 53, 6473−6495

contrast, in the MD(xray,λ = 1.03) simulation, Sos tends to interconvert between different conformations more rapidly. Consequently, this model is more successful in reproducing the experimental ^{15}N relaxation rates.

Finally, valuable insight into internal dynamics can be obtained from relaxation data collected at multiple magnetic field strengths. With this goal in mind, we have measured the ^{15}N relaxation in Sos:wtSH3 also at a proton Larmor frequency of 500 MHz. The experimental results are summarized in the left column of Figure 12, where solid and dashed lines represent

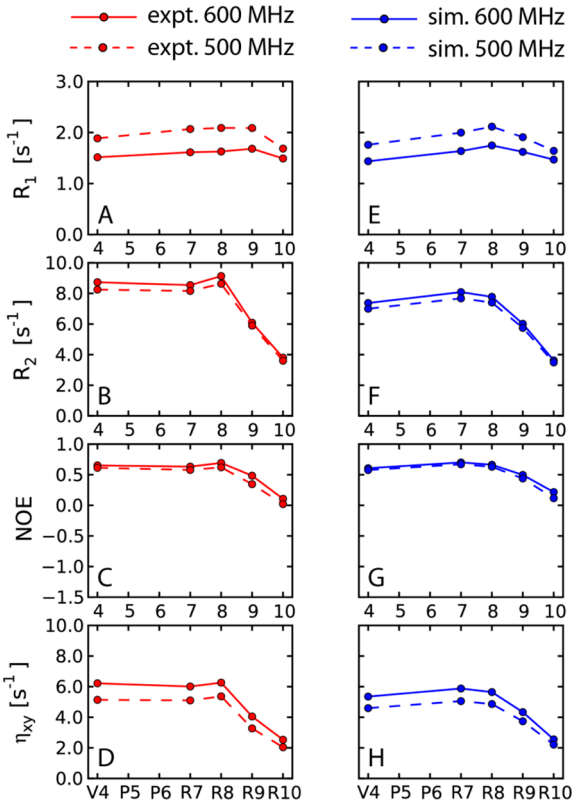


Figure 12. Experimental and simulated backbone ^{15}N relaxation rates in the peptide Sos bound to wtSH3 at 600 and 500 MHz spectrometer frequencies (solid and dashed lines, respectively). The simulated results are from the MD(xray,λ = 1.03) trajectory.

600 and 500 MHz data, respectively. The right column shows the equivalent data calculated on the basis of the MD(xray,λ = 1.03) trajectory. Insofar as the field dependence is concerned, the agreement between the experiment and the simulation is very convincing. For relatively rigid sites such as V4, the outcome is dictated mainly by the overall tumbling time, τ_{rot} = 7 ns. However, in the case of flexible residues such as R10, the results reflect a shorter correlation time, corresponding to the internal peptide dynamics (cf., for example, panels D and H in Figure 12).

Simulations of the Binding Process between Sos and wtSH3. Several years ago, Ahmad, Gu, and Helms[30] showed that it is possible to use MD simulations in explicit solvent to reproduce peptide binding in a system that is very similar to Sos:wtSH3 (the only difference is that these authors used a shorter peptide containing only two C-terminal arginines instead of four). Here we set out to reinvestigate this problem using the modified ff99SB*-ILDN force field (λ = 1.03) and significantly longer simulations. For this purpose, we have implemented a special protocol (see Materials and Methods). The initial conformation of the peptide was chosen randomly from a 2.1 μs trajectory of free Sos. The peptide was then placed at a certain distance from the surface of the protein (~15 Å considering the distance from the center of mass of Sos to the surface of wtSH3). The initial orientation of Sos was assigned randomly. The resulting system was used to construct a water box, allowing for a generous amount of water. Illustrated in Figure 13A is the initial setup that has been

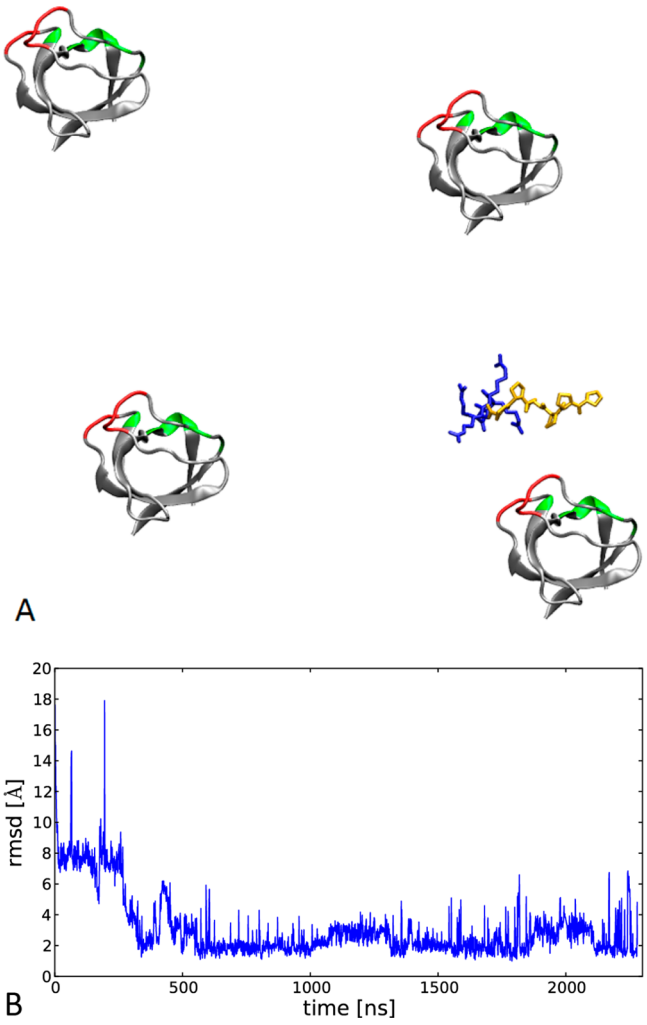


Figure 13. (A) Initial placement of the peptide and the protein in the 2.28 μs MD(rand,λ = 1.03) trajectory. Along with the actual simulated protein molecule, the plot also shows three periodic images. The space between the protein molecules is filled by TIP3P solvent. Even though the proline-rich sequence in Sos is sterically constrained, it remains sufficiently flexible in the MD simulation. (B) rmsd of the peptide atomic coordinates relative to the crystallographic structure 1CKB (same plotting conventions as in Figure 6). The lowest rmsd value is just under 1 Å. The simulation is 10-fold longer than those reported by Ahmad et al.[30]

used to record one of the MD(rand,λ = 1.03) trajectories. Along with the single simulated wtSH3 molecule, the graph also shows three periodic images of wtSH3. This representation makes it possible to appreciate the amount of water in the system (the empty space in the graph is filled with TIP3P water). In this particular trajectory, the effective concentration of the protein is 7.8 mM, which is not too far from the

dx.doi.org/10.1021/bi500904f | Biochemistry 2014, 53, 6473−6495

experimentally accessible concentrations (up to 4 mM in our experimental measurements, see the Supporting Information). Consequently, there is enough room in this simulation for the Sos peptide to fully explore the conformational space available to it, without being squeezed by the protein molecules.

Using this type of procedure, we have recorded thirty trajectories modeling the binding of Sos to wtSH3, with the trajectory length ranging from 20 to 265 ns. In nine of these trajectories the Sos peptide and wtSH3 formed a complex that was found to be similar, although not identical, to the crystal structure 1CKB. Of these nine trajectories, we have selected one that was subsequently extended to 2.28 μs. This trajectory, termed MD(rand,λ = 1.03), was used to simulate the NMR observables, which were subsequently compared with the experimental values.

In the starting frame of the MD(rand,λ = 1.03) trajectory, the Sos peptide is sufficiently far removed from the surface of wtSH3, see Figure 13. Initially, the displacement of the Sos peptide (i.e., rmsd relative to the coordinates of the bound Sos as found in the crystal structure 1CKB) amounts to 17.6 Å. As the Sos·wtSH3 encounter complex is formed, the rmsd drops rapidly to ca. 8 Å. The first salt bridge, R8−E167, is observed already after 4 ns. However, this turns out to be a transient interaction that immediately dissipates and subsequently makes only a few brief appearances throughout the trajectory. It is not before 45 ns that the two salt bridges are established in a sufficiently stable manner, R8−D150 and R8−D147. Both of these salt bridges are non-native. The resulting conformational state deserves a brief discussion. The pose of the peptide at this point generally resembles the binding pose found in 1CKB. In particular, the N-terminal portion of Sos hovers over the hydrophobic grooves where it is supposed to bind. However, the conformation of the peptide is not conducive to binding (there is a certain amount of twist); the peptide lies high above the surface of the protein and fails to form two signature hydrogen bonds, P2−Y186 and P5−W169. The hydrophobic buried surface area remains relatively small, and the peptide occasionally swings away from the surface of wtSH3 (large spikes in the initial portion of the graph in Figure 13B).

After some time, this labile state is transformed into a bona fide complex. At around 270 ns, several conformational transitions take place, proline rings P2 and P5 slide into the shallow hydrophobic grooves on the surface of wtSH3, and the native hydrogen bonds P2−Y186 and P5−W169 are established. At around 290 ns, the (weak) native salt bridge R7−E149 is formed. At 320 ns, the non-native salt bridge R8−D150 is dissolved and replaced with the native salt bridge R7−D150. As a result of these transformations, by 340 ns the coordinate rmsd of the Sos peptide drops to 1.3 Å. However, large fluctuations continue to occur until 550 ns when the complex is stabilized by the hydrogen bond between the R8 amide group and the E166 side chain, as well as certain additional salt bridges. After that, the peptide mostly remains within 2 Å of the crystallographic structure, although it still experiences substantial fluctuations involving the rearrangements of the C-terminal tail.

The MD(rand,λ = 1.03) trajectory can be used to simulate relaxation data and binding shifts that can be subsequently compared with the experimental data, in the same way as demonstrated previously. For this purpose, we have chosen the portion of the trajectory that begins at 400 ns (the point at which the Sos:wtSH3 complex is fully formed). The results are presented in Figure S3 of the Supporting Information. Briefly, they are virtually identical to those shown in Figure 10 with one exception—backbone shifts of residue R8 are significantly smaller than those observed experimentally (and those predicted on the basis of the MD(rand,λ = 1.03) trajectory). This deviation is caused by effective weakening of the hydrogen bond between R8 and the side chain of E166. In the selected time interval from 400 to 2280 ns, this hydrogen bond is identified in only 35% of all frames. By comparison, in the MD(xray,λ = 1.03) trajectory, this proportion reaches 50%. Most likely, the difference can be attributed to the less-than-perfect convergence of the MD analyses. As already indicated, 2−3 μs simulations are sufficient to reliably reproduce ^{15}N relaxation rates, but not necessarily the binding shifts. However, in principle, we cannot rule out another possibility. It is conceivable that formation of the Sos:wtSH3 complex has not been fully completed in the 2.28 μs MD(rand,λ = 1.03) simulation. Indeed, there are certain systematic differences between MD(rand,λ = 1.03) and MD(xray,λ = 1.03) trajectories—for example, the former features three intermittent salt bridges involving D147 that are virtually absent in the latter.

Finally, it is worthwhile to discuss the characteristic time scales as seen in the MD(rand,λ = 1.03) simulation. Using the values of k_{ON} and k_{OFF} experimentally determined in this study and the effective concentration of wtSH3 in the MD simulation, we can estimate that the Sos:wtSH3 complex should be formed with a time constant of ca. 100 ns and dissolve with a time constant of ca. 500 μs. These estimates are consistent with the actual simulation, in which the final complex is formed after ca. 400 ns and the peptide never becomes separated from the protein over the entire duration of the trajectory. Note that we have been successful in observing the peptide binding only due to electrostatic interactions; diffusion-limited binding is much slower and thus most likely cannot be reproduced in the relatively short MD simulations.

CONCLUSION

MD modeling combined with experimental NMR studies has been enormously popular in the context of (modestly sized) globular proteins. Such combined studies produced a wealth of information, greatly advancing our understanding of protein dynamics and at the same time stimulating the development of force fields. However, much of local dynamics in globular proteins is somewhat generic; e.g., small-amplitude fluctuations of peptide planes can be accurately predicted on the basis of protein structure, thus to some degree obviating the need in specialized MD studies. In contrast, intrinsically disordered proteins display a rich range of dynamic behaviors that are uniquely suited for MD and NMR studies. Yet until recently the progress in this area has been impeded by a lack of adequate computational resources. Indeed, proper modeling of an IDP requires a very large water box to accommodate the more extended protein conformations; it also requires a very long trajectory to sample the vast conformational space available to an IDP. All of this has become possible only with the advent of new computers and the suitably adapted MD algorithms.

In our study, we investigate the protein−peptide system that features a substantial element of dynamic disorder. For example, in the MD(rand,λ = 1.03) simulation, we model the flexible 10-residue peptide, which moves in a stochastic fashion toward its protein target and then forms a highly dynamic intermediate before making a transition to the bona fide complex. Furthermore, even in the bound state, the peptide

retains a substantial amount of motional freedom, especially the C-terminal residues. Although the presence of disorder in this system clearly raises the bar for MD studies (in terms of both the simulation length and the size of the water box), the simulation has been successful, reproducing the experimental NMR data with near-quantitative accuracy. We envisage that in the near future this methodology will progress from peptides to small disordered proteins and eventually to large IDPs, improving our understanding of many aspects of their function, including the disorder-to-order transition upon binding to their folded targets.

In this study, we focus on the moderately high affinity (~1 μM) complex between Sos and wild-type SH3. Comparing the experimental data (specifically, ^{15}N relaxation rates) with the results from the state-of-the-art Amber ff99SB*-ILDN simulation, we have found that molecular dynamics overemphasizes salt bridge interactions between the peptide and the protein. The tendency to misjudge the strength of salt bridges appears to be a common affliction of modern fixed-charge force fields.

To address this problem, we have devised a targeted correction involving the Lennard-Jones potential between the charged moieties of the Arg and Asp/Glu side chains. This is clearly an empirical solution; the proper way of approaching this problem would be to systematically reparameterize the entire force field. Nevertheless, there is little doubt that off-diagonal Lennard-Jones parameters associated with salt bridges need to be reassessed in Amber ff99SB as well as other widely used force fields. In this sense, we have likely identified an important target for future optimization of non-polarizable force fields.

The MD trajectory of Sos:wtSH3 recorded with the altered force field, Amber ff99SB*-ILDN (λ = 1.03), showed good agreement with the experiment, specifically, with respect to ^{15}N relaxation rates and binding shifts (in the latter case, the outcome is to some degree affected by convergence). As it turns out, weakening of the salt bridges in the C-terminal portion of Sos strengthens the binding at the N-terminal region. This implies that there is a certain amount of conflict between, on one hand, the salt bridge interactions implicated in formation of the electrostatic encounter complex and, on the other hand, hydrogen bonds and hydrophobic packing indicative of the tight complex. In other words, the interactions associated with the electrostatic encounter complex do not necessarily constitute a subset of the interactions found in the fully formed complex. As it appears, the weakening of the salt bridges in the λ = 1.03 simulation has an effect not so much on the amplitudes of motion of the C-terminal residues but on the respective correlation times. Compared to the original simulation, the C-terminal tail visits the same conformations but does not remain in these conformations for quite as long. Faster conformational exchange is reflected in the spin relaxation rates, leading to improved agreement with the experiment.

In this study, we have found the way to probe, via both MD simulation and experiment, the interactions that are key to the electrostatic encounter complex. In the second part of this work, we use the approach developed here to directly characterize the system that mimics the electrostatic encounter complex. Specifically, we have designed the double mutant of c-Crk N-SH3 (Y186L/W169F) in which the binding of the N-terminal portion of Sos is largely abrogated and the complex is held together mainly by the salt bridges, thus resembling the electrostatic encounter complex. The Sos·dmSH3 model has also been successfully characterized via NMR measurements and modeled with the help of the (modified) MD protocol. The results of this work will be described in paper 2.

■ ASSOCIATED CONTENT

Supporting Information

Table with rotational diffusion parameters for wild-type c-Crk N-SH3 in the free form and in the presence of Sos peptide, concentration dependence of chemical shifts reflecting weak self-association of wtSH3, chemical shift perturbation maps for wtSH3 self-association and for wtSH3 binding Sos, comparison of the experimental and simulated ^{15}N relaxation parameters and ^{1}H,^{15}N binding shifts for Sos:wtSH3 using the MD(rand,λ = 1.03) trajectory. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: nikolai@purdue.edu.

Present Address

[‖]Y.X.: Department of Biochemistry, Duke University, Durham, NC 27710.

Author Contributions

Y.X. and T.Y. contributed equally to this work.

Funding

This work has been supported by National Science Foundation Grant MCB 1158347.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful to Dr. Lake Paul for helping us set up the ITC experiment.

■ REFERENCES

(1) Williams, R. J. (1978) The conformational mobility of proteins and its functional significance. *Biochem. Soc. Trans. 6*, 1123−1126.

(2) Uversky, V. N., Gillespie, J. R., and Fink, A. L. (2000) Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins: Struct., Funct., Genet. 41*, 415−427.

(3) Romero, P., Obradovic, Z., Li, X. H., Garner, E. C., Brown, C. J., and Dunker, A. K. (2001) Sequence complexity of disordered protein. *Proteins: Struct., Funct., Genet. 42*, 38−48.

(4) Oldfield, C. J., Cheng, Y., Cortese, M. S., Brown, C. J., Uversky, V. N., and Dunker, A. K. (2005) Comparing and combining predictors of mostly disordered proteins. *Biochemistry 44*, 1989−2000.

(5) Uversky, V. N., and Dunker, A. K. (2008) Controlled Chaos. *Science 322*, 1340−1341.

(6) Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F., and Jones, D. T. (2004) Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol. 337*, 635−645.

(7) Iakoucheva, L. M., Brown, C. J., Lawson, J. D., Obradovic, Z., and Dunker, A. K. (2002) Intrinsic disorder in cell-signaling and cancer-associated proteins. *J. Mol. Biol. 323*, 573−584.

(8) Xie, H. B., Vucetic, S., Iakoucheva, L. M., Oldfield, C. J., Dunker, A. K., Uversky, V. N., and Obradovic, Z. (2007) Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. *J. Proteome Res. 6*, 1882−1898.

(9) Dames, S. A., Martinez-Yamout, M., De Guzman, R. N., Dyson, H. J., and Wright, P. E. (2002) Structural basis for Hif-1α/CBP recognition in the cellular hypoxic response. *Proc. Natl. Acad. Sci. U.S.A. 99*, 5271−5276.

(10) Bochkareva, E., Kaustov, L., Ayed, A., Yi, G. S., Lu, Y., Pineda-Lucena, A., Liao, J. C. C., Okorokov, A. L., Milner, J., Arrowsmith, C.

H., and Bochkarev, A. (2005) Single-stranded DNA mimicry in the p53 transactivation domain interaction with replication protein A. *Proc. Natl. Acad. Sci. U.S.A. 102*, 15412−15417.
(11) Mohan, A., Oldfield, C. J., Radivojac, P., Vacic, V., Cortese, M. S., Dunker, A. K., and Uversky, V. N. (2006) Analysis of molecular recognition features (MoRFs). *J. Mol. Biol. 362*, 1043−1059.
(12) Dancheck, B., Nairn, A. C., and Peti, W. (2008) Detailed structural characterization of unbound protein phosphatase 1 inhibitors. *Biochemistry 47*, 12346−12356.
(13) Ragusa, M. J., Dancheck, B., Critton, D. A., Nairn, A. C., Page, R., and Peti, W. (2010) Spinophilin directs protein phosphatase 1 specificity by blocking substrate binding sites. *Nat. Struct. Mol. Biol. 17*, 459−464.
(14) Gunasekaran, K., Tsai, C. J., Kumar, S., Zanuy, D., and Nussinov, R. (2003) Extended disordered proteins: Targeting function with less scaffold. *Trends Biochem. Sci. 28*, 81−85.
(15) Marsh, J. A., and Forman-Kay, J. D. (2010) Sequence determinants of compaction in intrinsically disordered proteins. *Biophys. J. 98*, 2383−2390.
(16) Iakoucheva, L. M., Radivojac, P., Brown, C. J., O'Connor, T. R., Sikes, J. G., Obradovic, Z., and Dunker, A. K. (2004) The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res. 32*, 1037−1049.
(17) Borg, M., Mittag, T., Pawson, T., Tyers, M., Forman-Kay, J. D., and Chan, H. S. (2007) Polyelectrostatic interactions of disordered ligands suggest a physical basis for ultrasensitivity. *Proc. Natl. Acad. Sci. U.S.A. 104*, 9650−9655.
(18) Schreiber, G., Haran, G., and Zhou, H. X. (2009) Fundamental aspects of protein-protein association kinetics. *Chem. Rev. 109*, 839−860.
(19) Janin, J. (1997) The kinetics of protein-protein recognition. *Proteins: Struct., Funct., Genet. 28*, 153−161.
(20) Qin, S., Pang, X. D., and Zhou, H. X. (2011) Automated prediction of protein association rate constants. *Structure 19*, 1744−1751.
(21) Zhou, H. X., and Bates, P. A. (2013) Modeling protein association mechanisms and kinetics. *Curr. Opin. Struct. Biol. 23*, 887−893.
(22) Wright, P. E., and Dyson, H. J. (2009) Linking folding and binding. *Curr. Opin. Struct. Biol. 19*, 31−38.
(23) Petros, A. M., Nettesheim, D. G., Wang, Y., Olejniczak, E. T., Meadows, R. P., Mack, J., Swift, K., Matayoshi, E. D., Zhang, H. C., Thompson, C. B., and Fesik, S. W. (2000) Rationale for Bcl-x(L)/Bad peptide complex formation from structure, mutagenesis, and biophysical studies. *Protein Sci. 9*, 2528−2534.
(24) Iesmantavicius, V., Dogan, J., Jemth, P., Teilum, K., and Kjaergaard, M. (2014) Helical propensity in an intrinsically disordered protein accelerates ligand binding. *Angew. Chem., Int. Ed. 53*, 1548−1551.
(25) Lacy, E. R., Filippov, I., Lewis, W. S., Otieno, S., Xiao, L. M., Weiss, S., Hengst, L., and Kriwacki, R. W. (2004) p27 binds cyclin-CDK complexes through a sequential mechanism involving binding-induced protein folding. *Nat. Struct. Mol. Biol. 11*, 358−364.
(26) Zhou, H. X., Pang, X. D., and Lu, C. (2012) Rate constants and mechanisms of intrinsically disordered proteins binding to structured targets. *Phys. Chem. Chem. Phys. 14*, 10466−10476.
(27) Wong, E. T. C., Na, D., and Gsponer, J. (2013) On the importance of polar interactions for complexes containing intrinsically disordered proteins. *PLoS Comput. Biol. 9*, e1003192.
(28) Ganguly, D., Otieno, S., Waddell, B., Iconaru, L., Kriwacki, R. W., and Chen, J. H. (2012) Electrostatically accelerated coupled binding and folding of intrinsically disordered proteins. *J. Mol. Biol. 422*, 674−684.
(29) Ganguly, D., Zhang, W. H., and Chen, J. H. (2013) Electrostatically accelerated encounter and folding for facile recognition of intrinsically disordered proteins. *PLoS Comput. Biol. 9*, e1003363.
(30) Ahmad, M., Gu, W., and Helms, V. (2008) Mechanism of fast peptide recognition by SH3 domains. *Angew. Chem., Int. Ed. 47*, 7626−7630.
(31) Giorgino, T., Buch, I., and De Fabritiis, G. (2012) Visualizing the induced binding of SH2-phosphopeptide. *J. Chem. Theory Comput. 8*, 1171−1175.
(32) Guan, J. Y., Foerster, J. M., Drijfhout, J. W., Timmer, M., Blok, A., Ullmann, G. M., and Ubbink, M. (2014) An ensemble of rapidly interconverting orientations in electrostatic protein-peptide complexes characterized by NMR spectroscopy. *ChemBioChem 15*, 556−566.
(33) Dagliyan, O., Proctor, E. A., D'Auria, K. M., Ding, F., and Dokholyan, N. V. (2011) Structural and dynamic determinants of protein-peptide recognition. *Structure 19*, 1837−1845.
(34) Staneva, I., Huang, Y. Q., Liu, Z. R., and Wallin, S. (2012) Binding of two intrinsically disordered peptides to a multi-specific protein: A combined Monte-Carlo and molecular dynamics study. *PLoS Comput. Biol. 8*, e1002682.
(35) Matsuda, M., Tanaka, S., Nagata, S., Kojima, A., Kurata, T., and Shibuya, M. (1992) Two species of human Crk cDNA encode proteins with distinct biological activities. *Mol. Cell. Biol. 12*, 3482−3489.
(36) Reichman, C. T., Mayer, B. J., Keshav, S., and Hanafusa, H. (1992) The product of the cellular Crk gene consists primarily of SH2 and SH3 regions. *Cell Growth Differ. 3*, 451−460.
(37) Kobashigawa, Y., Sakai, M., Naito, M., Yokochi, M., Kumeta, H., Makino, Y., Ogura, K., Tanaka, S., and Inagaki, F. (2007) Structural basis for the transforming activity of human cancer-related signaling adaptor protein CRK. *Nat. Struct. Mol. Biol. 14*, 503−510.
(38) Sarkar, P., Reichman, C., Saleh, T., Birge, R. B., and Kalodimos, C. G. (2007) Proline cis-trans isomerization controls autoinhibition of a signaling protein. *Mol. Cell 25*, 413−426.
(39) Camarero, J. A., Fushman, D., Sato, S., Giriat, I., Cowburn, D., Raleigh, D. P., and Muir, T. W. (2001) Rescuing a destabilized protein fold through backbone cyclization. *J. Mol. Biol. 308*, 1045−1062.
(40) Hirata, T., Nagai, H., Koizumi, K., Okino, K., Harada, A., Onda, M., Nagahata, T., Mikami, I., Hirai, K., Haraguchi, S., Jin, E. J., Kawanami, O., Shimizu, K., and Emi, M. (2004) Amplification, up-regulation and over-expression of C3G (CRK SH3 domain-binding guanine nucleotide-releasing factor) in non-small cell lung cancers. *J. Hum. Genet. 49*, 290−295.
(41) Jarzynka, N. J., Hu, B., Hui, K. M., Bar-Joseph, I., Gu, W. S., Hirose, T., Haney, L. B., Ravichandran, K. S., Nishikawa, R., and Cheng, S. Y. (2007) ELMO1 and Dock180, a bipartite Rac1 guanine nucleotide exchange factor, promote human glioma cell invasion. *Cancer Res. 67*, 7203−7211.
(42) Birge, R. B., Kalodimos, C., Inagaki, F., and Tanaka, S. (2009) Crk and CrkL adaptor proteins: Networks for physiological and pathological signaling. *Cell Commun. Signaling 7*, 7−13.
(43) Matsuda, M., Hashimoto, Y., Muroya, K., Hasegawa, H., Kurata, T., Tanaka, S., Nakamura, S., and Hattori, S. (1994) Crk protein binds to two guanine nucleotide-releasing proteins for the Ras family and modulates nerve growth factor-induced activation of Ras in PC12 cells. *Mol. Cell. Biol. 14*, 5495−5500.
(44) Feller, S. M., Knudsen, B., and Hanafusa, H. (1995) Cellular proteins binding to the first Src homology 3 (SH3) domain of the protooncogene product c-Crk indicate Crk-specific signaling pathways. *Oncogene 10*, 1465−1473.
(45) Okada, S., and Pessin, J. E. (1996) Interactions between Src homology (SH) 2/SH3 adapter proteins and the guanylnucleotide exchange factor SOS are differentially regulated by insulin and epidermal growth factor. *J. Biol. Chem. 271*, 25533−25538.
(46) Rojas, J. M., Oliva, J. L., and Santos, E. (2011) Mammalian son of sevenless guanine nucleotide exchange factors: Old concepts and new perspectives. *Genes Cancer 2*, 298−305.
(47) Wu, X. D., Knudsen, B., Feller, S. M., Zheng, J., Sali, A., Cowburn, D., Hanafusa, H., and Kuriyan, J. (1995) Structural basis for the specific interaction of lysine-containing proline-rich peptides with the N-terminal SH3 domain of c-Crk. *Structure 3*, 215−226.
(48) Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple Amber force fields and

development of improved protein backbone parameters. *Proteins 65*, 712−725.
(49) Best, R. B., and Hummer, G. (2009) Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. *J. Phys. Chem. B 113*, 9004−9015.
(50) Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., and Shaw, D. E. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins 78*, 1950−1958.
(51) Andrews, C. T., and Elcock, A. H. (2013) Molecular dynamics simulations of highly crowded amino acid solutions: Comparisons of eight different force field combinations with experiment and with each other. *J. Chem. Theory Comput. 9*, 4585−4602.
(52) Debiec, K. T., Gronenborn, A. M., and Chong, L. T. (2014) Evaluating the strength of salt bridges: A comparison of current biomolecular force fields. *J. Phys. Chem. B 118*, 6561−6569.
(53) Tolkatchev, D., Xu, P., and Ni, F. (2003) Probing the kinetic landscape of transient peptide-protein interactions by use of peptide N-15 NMR relaxation dispersion spectroscopy: Binding of an antithrombin peptide to human prothrombin. *J. Am. Chem. Soc. 125*, 12432−12442.
(54) Sugase, K., Dyson, H. J., and Wright, P. E. (2007) Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature 447*, 1021−1027.
(55) Fawzi, N. L., Ying, J. F., Ghirlando, R., Torchia, D. A., and Clore, G. M. (2011) Atomic-resolution dynamics on the surface of amyloid-β protofibrils probed by solution NMR. *Nature 480*, 268−272.
(56) Vallurupalli, P., Bouvignies, G., and Kay, L. E. (2012) Studying "invisible" excited protein states in slow exchange with a major state conformation. *J. Am. Chem. Soc. 134*, 8148−8161.
(57) Wu, K.-P., and Baum, J. (2010) Detection of transient interchain interactions in the intrinsically disordered protein α-synuclein by NMR paramagnetic relaxation enhancement. *J. Am. Chem. Soc. 132*, 5546−5547.
(58) Demers, J. P., and Mittermaier, A. (2009) Binding mechanism of an SH3 domain studied by NMR and ITC. *J. Am. Chem. Soc. 131*, 4355−4367.
(59) Terpe, K. (2003) Overview of tag protein fusions: From molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol. 60*, 523−533.
(60) Huth, J. R., Bewley, C. A., Jackson, B. M., Hinnebusch, A. G., Clore, G. M., and Gronenborn, A. M. (1997) Design of an expression system for detecting folded protein domains and mapping macromolecular interactions by NMR. *Protein Sci. 6*, 2359−2364.
(61) Lindhout, D. A., Thiessen, A., Schieve, D., and Sykes, B. D. (2003) High-yield expression of isotopically labeled peptides for use in NMR studies. *Protein Sci. 12*, 1786−1791.
(62) Wider, G., and Dreier, L. (2006) Measuring protein concentrations by NMR spectroscopy. *J. Am. Chem. Soc. 128*, 2571−2576.
(63) Dreier, L., and Wider, G. (2006) Concentration measurements by PULCON using X-filtered or 2D NMR spectra. *Magn. Reson. Chem. 44*, S206−S212.
(64) Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and a phosphopeptide-complexed src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry 33*, 5984−6003.
(65) Korzhnev, D. M., Skrynnikov, N. R., Millet, O., Torchia, D. A., and Kay, L. E. (2002) An NMR experiment for the accurate measurement of heteronuclear spin-lock relaxation rates. *J. Am. Chem. Soc. 124*, 10743−10753.
(66) Hansen, D. F., and Kay, L. E. (2007) Improved magnetization alignment schemes for spin-lock relaxation experiments. *J. Biomol. NMR 37*, 245−255.
(67) Ferrage, F., Cowburn, D., and Ghose, R. (2009) Accurate sampling of high-frequency motions in proteins by steady-state ^{15}N-{^{1}H} nuclear Overhauser effect measurements in the presence of cross-correlated relaxation. *J. Am. Chem. Soc. 131*, 6048−6049.
(68) Ottiger, M., Delaglio, F., and Bax, A. (1998) Measurement of *J* and dipolar couplings from simplified two-dimensional NMR spectra. *J. Magn. Reson. 131*, 373−378.
(69) Hall, J. B., and Fushman, D. (2003) Direct measurement of the transverse and longitudinal ^{15}N chemical shift anisotropy: Dipolar cross-correlation rate constants using ^{1}H-coupled HSQC spectra. *Magn. Reson. Chem. 41*, 837−842.
(70) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on Unix pipes. *J. Biomol. NMR 6*, 277−293.
(71) Goddard, T. D., and Kneller, D. G. (2002) *SPARKY 3*, University of California, San Francisco.
(72) Anafi, M., Rosen, M. K., Gish, G. D., Kay, L. E., and Pawson, T. (1996) A potential SH3 domain-binding site in the Crk SH2 domain. *J. Biol. Chem. 271*, 21365−21374.
(73) Wüthrich, K. (1986) *NMR of proteins and nucleic acids*, Wiley Interscience, New York.
(74) Bai, Y. W., Milne, J. S., Mayne, L., and Englander, S. W. (1993) Primary structure effects on peptide group hydrogen exchange. *Proteins: Struct., Funct., Genet. 17*, 75−86.
(75) Song, J. N., Burrage, K., Yuan, Z., and Huber, T. (2006) Prediction of cis/trans isomerization in proteins using PSI-BLAST profiles and secondary structure information. *BMC Bioinf. 7*, 124.
(76) O'Neal, K. D., Chari, M. V., McDonald, C. H., Cook, R. G., YuLee, L. Y., Morrisett, J. D., and Shearer, W. T. (1996) Multiple cis-trans conformers of the prolactin receptor proline-rich motif (PRM) peptide detected by reverse-phase HPLC, CD and NMR spectroscopy. *Biochem. J. 315*, 833−844.
(77) Wu, W. J., and Raleigh, D. P. (1998) Local control of peptide conformation: Stabilization of cis proline peptide bonds by aromatic proline interactions. *Biopolymers 45*, 381−394.
(78) Kovrigin, E. L. (2012) NMR line shapes and multi-state binding equilibria. *J. Biomol. NMR 53*, 257−270.
(79) de la Torre, J. G., Huertas, M. L., and Carrasco, B. (2000) HYDRONMR: Prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J. Magn. Reson., Ser. B 147*, 138−146.
(80) Lee, L. K., Rance, M., Chazin, W. J., and Palmer, A. G. (1997) Rotational diffusion anisotropy of proteins from simultaneous analysis of ^{15}N and $^{13}C^{\alpha}$ nuclear spin relaxation. *J. Biomol. NMR 9*, 287−298.
(81) Tjandra, N., Feller, S. E., Pastor, R. W., and Bax, A. (1995) Rotational diffusion anisotropy of human ubiquitin from ^{15}N NMR relaxation. *J. Am. Chem. Soc. 117*, 12562−12566.
(82) Pfuhl, M., Chen, H. A., Kristensen, S. M., and Driscoll, P. C. (1999) NMR exchange broadening arising from specific low affinity protein self-association: Analysis of nitrogen-15 nuclear relaxation for rat CD2 domain 1. *J. Biomol. NMR 14*, 307−320.
(83) Bernado, P., Akerud, T., de la Torre, J. G., Akke, M., and Pons, M. (2003) Combined use of NMR relaxation measurements and hydrodynamic calculations to study protein association. Evidence for tetramers of low molecular weight protein tyrosine phosphatase in solution. *J. Am. Chem. Soc. 125*, 916−923.
(84) Noguera, V., Walker, O., Rouhier, N., Jacquot, J. P., Krimm, I., and Lancelin, J. M. (2005) NMR reveals a novel glutaredoxin-glutaredoxin interaction interface. *J. Mol. Biol. 353*, 629−641.
(85) Baryshnikova, O. K., and Sykes, B. D. (2006) Backbone dynamics of SDF-1α determined by NMR: Interpretation in the presence of monomer-dimer equilibrium. *Protein Sci. 15*, 2568−2578.
(86) Liu, Z., Zhang, W. P., Xing, Q., Ren, X. F., Liu, M. L., and Tang, C. (2012) Noncovalent dimerization of ubiquitin. *Angew. Chem., Int. Ed. 51*, 469−472.
(87) Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The Amber biomolecular simulation programs. *J. Comput. Chem. 26*, 1668−1688.
(88) Eswar, N., Eramian, D., Webb, B., Shen, M.-Y., and Sali, A. (2008) Protein structure modeling with MODELLER. In *Methods in Molecular Biology* (Kobe, B., Guss, M., and Huber, T., Eds.) pp 145−159, Humana Press, Totowa, NJ.

(89) Bas, D. C., Rogers, D. M., and Jensen, J. H. (2008) Very fast prediction and rationalization of pK_a values for protein-ligand complexes. *Proteins 73*, 765−783.
(90) Han, B., Liu, Y. F., Ginzinger, S. W., and Wishart, D. S. (2011) SHIFTX2: Significantly improved protein chemical shift prediction. *J. Biomol. NMR 50*, 43−57.
(91) Xue, Y., Pavlova, M. S., Ryabov, Y. E., Reif, B., and Skrynnikov, N. R. (2007) Methyl rotation barriers in proteins from ^2H relaxation data. Implications for protein structure. *J. Am. Chem. Soc. 129*, 6827−6838.
(92) Bremi, T., Brüschweiler, R., and Ernst, R. R. (1997) A protocol for the interpretation of side-chain dynamics based on NMR relaxation: Application to phenylalanines in antamanide. *J. Am. Chem. Soc. 119*, 4272−4284.
(93) Cornilescu, G., and Bax, A. (2000) Measurement of proton, nitrogen, and carbonyl chemical shielding anisotropies in a protein dissolved in a dilute liquid crystalline phase. *J. Am. Chem. Soc. 122*, 10143−10154.
(94) Kumar, S., and Nussinov, R. (2002) Relationship between ion pair geometries and electrostatic strengths in proteins. *Biophys. J. 83*, 1595−1612.
(95) Jeffrey, G. A., and Saenger, W. (1991) *Hydrogen bonding in biological structures*, Springer-Verlalg, New York.
(96) Cavallo, L., Kleinjung, J., and Fraternali, F. (2003) POPS: A fast algorithm for solvent accessible surface areas at atomic and residue level. *Nucleic Acids Res. 31*, 3364−3366.
(97) Rickles, R. J., Botfield, M. C., Weng, Z. G., Taylor, J. A., Green, O. M., Brugge, J. S., and Zoller, M. J. (1994) Identification of Src, Fyn, Lyn, Pi3k and Abl SH3 domain ligands using phage display libraries. *EMBO J. 13*, 5598−5604.
(98) Cheadle, C., Ivashchenko, Y., South, V., Searfoss, G. H., French, S., Howk, R., Ricca, G. A., and Jaye, M. (1994) Identification of a Src SH3 domain binding motif by screening a random phage display library. *J. Biol. Chem. 269*, 24034−24039.
(99) Knudsen, B. S., Zheng, J., Feller, S. M., Mayer, J. P., Burrell, S. K., Cowburn, D., and Hanafusa, H. (1995) Affinity and specificity requirements for the first Src homology 3 domain of the Crk proteins. *EMBO J. 14*, 2191−2198.
(100) Schreiber, G., and Fersht, A. R. (1996) Rapid, electrostatically assisted association of proteins. *Nat. Struct. Biol. 3*, 427−431.
(101) Wafer, L. N., Streicher, W. W., McCallum, S. A., and Makhatadze, G. I. (2012) Thermodynamic and kinetic analysis of peptides derived from CapZ, NDR, p53, HDM2, and HDM4 binding to human S100B. *Biochemistry 51*, 7189−7201.
(102) Getzoff, E. D., Cabelli, D. E., Fisher, C. L., Parge, H. E., Viezzoli, M. S., Banci, L., and Hallewell, R. A. (1992) Faster superoxide dismutase mutants designed by enhancing electrostatic guidance. *Nature 358*, 347−351.
(103) Northrup, S. H., and Erickson, H. P. (1992) Kinetics of protein-protein association explained by Brownian dynamics computer simulation. *Proc. Natl. Acad. Sci. U.S.A. 89*, 3338−3342.
(104) Frisch, C., Fersht, A. R., and Schreiber, G. (2001) Experimental assignment of the structure of the transition state for the association of barnase and barstar. *J. Mol. Biol. 308*, 69−77.
(105) Arai, M., Ferreon, J. C., and Wright, P. E. (2012) Quantitative analysis of multisite protein-ligand interactions by NMR: Binding of intrinsically disordered p53 transactivation subdomains with the TAZ2 domain of CBP. *J. Am. Chem. Soc. 134*, 3792−3803.
(106) Showalter, S. A., and Bruschweiler, R. (2007) Validation of molecular dynamics simulations of biomolecules using NMR spin relaxation as benchmarks: Application to the AMBER99SB force field. *J. Chem. Theory Comput. 3*, 961−975.
(107) Penev, E., Ireta, J., and Shea, J. E. (2008) Energetics of infinite homopolypeptide chains: A new look at commonly used force fields. *J. Phys. Chem. B 112*, 6872−6877.
(108) Aliev, A. E., and Courtier-Murias, D. (2010) Experimental verification of force fields for molecular dynamics simulations using Gly-Pro-Gly-Gly. *J. Phys. Chem. B 114*, 12358−12375.
(109) Cerutti, D. S., Freddolino, P. L., Duke, R. E., and Case, D. A. (2010) Simulations of a protein crystal with a high resolution X-ray structure: Evaluation of force fields and water models. *J. Phys. Chem. B 114*, 12811−12824.
(110) Li, D. W., and Bruschweiler, R. (2010) Certification of molecular dynamics trajectories with NMR chemical shifts. *J. Phys. Chem. Lett. 1*, 246−248.
(111) Beauchamp, K. A., Lin, Y. S., Das, R., and Pande, V. S. (2012) Are protein force fields getting better? A systematic benchmark on 524 diverse NMR measurements. *J. Chem. Theory Comput. 8*, 1409−1414.
(112) Piana, S., Lindorff-Larsen, K., and Shaw, D. E. (2011) How robust are protein folding simulations with respect to force field parameterization? *Biophys. J. 100*, L47−L49.
(113) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys. 79*, 926−935.
(114) Chandrasekhar, I., Clore, G. M., Szabo, A., Gronenborn, A. M., and Brooks, B. R. (1992) A 500-ps molecular dynamics simulation study of interleukin-1β in water: Correlation with nuclear magnetic resonance spectroscopy and crystallography. *J. Mol. Biol. 226*, 239−250.
(115) Xue, Y., Ward, J. M., Yuwen, T. R., Podkorytov, I. S., and Skrynnikov, N. R. (2012) Microsecond time-scale conformational exchange in proteins: Using long molecular dynamics trajectory to simulate NMR relaxation dispersion data. *J. Am. Chem. Soc. 134*, 2555−2562.
(116) Kroenke, C. D., Loria, J. P., Lee, L. K., Rance, M., and Palmer, A. G. (1998) Longitudinal and transverse ^1H-^{15}N dipolar ^{15}N chemical shift anisotropy relaxation interference: Unambiguous determination of rotational diffusion tensors and chemical exchange effects in biological macromolecules. *J. Am. Chem. Soc. 120*, 7905−7915.
(117) Hall, J. B., and Fushman, D. (2003) Characterization of the overall and local dynamics of a protein with intermediate rotational anisotropy: Differentiating between conformational exchange and anisotropic diffusion in the B3 domain of protein G. *J. Biomol. NMR 27*, 261−275.
(118) Shen, Y., and Bax, A. (2007) Protein backbone chemical shifts predicted from searching a database for torsion angle and sequence homology. *J. Biomol. NMR 38*, 289−302.
(119) Brüschweiler, R., Roux, B., Blackledge, M., Griesinger, C., Karplus, M., and Ernst, R. R. (1992) Influence of rapid intramolecular motion on NMR cross-relaxation rates. A molecular dynamics study of antamanide in solution. *J. Am. Chem. Soc. 114*, 2289−2302.
(120) Pfeiffer, S., Fushman, D., and Cowburn, D. (2001) Simulated and NMR-derived backbone dynamics of a protein with significant flexibility: A comparison of spectral densities for the β ARK PH domain. *J. Am. Chem. Soc. 123*, 3021−3036.
(121) Nederveen, A. J., and Bonvin, A. M. J. J. (2005) NMR relaxation and internal dynamics of ubiquitin from a 0.2 μs MD simulation. *J. Chem. Theory Comput. 1*, 363−374.
(122) Markwick, P. R. L., Cervantes, C. F., Abel, B. L., Komives, E. A., Blackledge, M., and McCammon, J. A. (2010) Enhanced conformational space sampling improves the prediction of chemical shifts in proteins. *J. Am. Chem. Soc. 132*, 1220−1221.
(123) Robustelli, P., Stafford, K. A., and Palmer, A. G. (2012) Interpreting protein structural dynamics from NMR chemical shifts. *J. Am. Chem. Soc. 134*, 6365−6374.
(124) Lehtivarjo, J., Tuppurainen, K., Hassinen, T., Laatikainen, R., and Perakyla, M. (2012) Combining NMR ensembles and molecular dynamics simulations provides more realistic models of protein structures in solution and leads to better chemical shift prediction. *J. Biomol. NMR 52*, 257−267.
(125) Li, D. W., and Bruschweiler, R. (2012) PPM: A side-chain and backbone chemical shift predictor for the assessment of protein conformational ensembles. *J. Biomol. NMR 54*, 257−265.
(126) Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G. M., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J. M., and Kollman, P. (2003) A point-charge force field for molecular

mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem. 24*, 1999−2012.
(127) MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B 102*, 3586−3616.
(128) Mackerell, A. D., Feig, M., and Brooks, C. L. (2004) Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem. 25*, 1400−1415.
(129) Oostenbrink, C., Villa, A., Mark, A. E., and Van Gunsteren, W. F. (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem. 25*, 1656−1676.
(130) Ponder, J. W., Wu, C. J., Ren, P. Y., Pande, V. S., Chodera, J. D., Schnieders, M. J., Haque, I., Mobley, D. L., Lambrecht, D. S., DiStasio, R. A., Head-Gordon, M., Clark, G. N. I., Johnson, M. E., and Head-Gordon, T. (2010) Current status of the AMOEBA polarizable force field. *J. Phys. Chem. B 114*, 2549−2564.
(131) Trbovic, N., Kim, B., Friesner, R. A., and Palmer, A. G. (2008) Structural analysis of protein dynamics by MD simulations and NMR spin-relaxation. *Proteins 71*, 684−694.
(132) Frigyes, D., Alber, F., Pongor, S., and Carloni, P. (2001) Arginine-phosphate salt bridges in protein-DNA complexes: A Car-Parrinello study. *J. Mol. Struct.: THEOCHEM 574*, 39−45.
(133) Jiao, D., Golubkov, P. A., Darden, T. A., and Ren, P. (2008) Calculation of protein-ligand binding free energy by using a polarizable potential. *Proc. Natl. Acad. Sci. U.S.A. 105*, 6290−6295.
(134) Ng, A. H., and Snow, C. D. (2011) Polarizable protein packing. *J. Comput. Chem. 32*, 1334−1344.
(135) Liang, T., and Walsh, T. R. (2006) Molecular dynamics simulations of peptide carboxylate hydration. *Phys. Chem. Chem. Phys. 8*, 4410−4419.
(136) Kaminsky, J., and Jensen, F. (2007) Force field modeling of amino acid conformational energies. *J. Chem. Theory Comput. 3*, 1774−1788.
(137) Jiao, D., Zhang, J. J., Duke, R. E., Li, G. H., Schnieders, M. J., and Ren, P. Y. (2009) Trypsin-ligand binding free energies from explicit and implicit solvent simulations with polarizable potential. *J. Comput. Chem. 30*, 1701−1711.
(138) Fenn, T. D., Schnieders, M. J., Brunger, A. T., and Pande, V. S. (2010) Polarizable atomic multipole X-ray refinement: Hydration geometry and application to macromolecules. *Biophys. J. 98*, 2984−2992.
(139) Jiang, J. L., Wu, Y. B., Wang, Z. X., and Wu, C. (2010) Assessing the performance of popular quantum mechanics and molecular mechanics methods and revealing the sequence-dependent energetic features using 100 tetrapeptide models. *J. Chem. Theory Comput. 6*, 1199−1209.
(140) Wong, S. E., Bernacki, K., and Jacobson, M. (2005) Competition between intramolecular hydrogen bonds and solvation in phosphorylated peptides: Simulations with explicit and implicit solvent. *J. Phys. Chem. B 109*, 5249−5258.
(141) Hess, B., and van der Vegt, N. F. A. (2006) Hydration thermodynamic properties of amino acid analogues: A systematic comparison of biomolecular force fields and water models. *J. Phys. Chem. B 110*, 17616−17626.
(142) Paschek, D., Day, R., and Garcia, A. E. (2011) Influence of water-protein hydrogen bonding on the stability of Trp-cage miniprotein. A comparison between the TIP3P and TIP4P-Ew water models. *Phys. Chem. Chem. Phys. 13*, 19840−19847.
(143) Florova, P., Sklenovsky, P., Banas, P., and Otyepka, M. (2010) Explicit water models affect the specific solvation and dynamics of unfolded peptides while the conformational behavior and flexibility of folded peptides remain intact. *J. Chem. Theory Comput. 6*, 3569−3579.
(144) Nerenberg, P. S., and Head-Gordon, T. (2011) Optimizing protein-solvent force fields to reproduce intrinsic conformational preferences of model peptides. *J. Chem. Theory Comput. 7*, 1220−1230.
(145) Berendsen, H. J. C., Grigera, J. R., and Straatsma, T. P. (1987) The missing term in effective pair potentials. *J. Phys. Chem. 91*, 6269−6271.
(146) Horn, H. W., Swope, W. C., Pitera, J. W., Madura, J. D., Dick, T. J., Hura, G. L., and Head-Gordon, T. (2004) Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys. 120*, 9665−9678.
(147) Lamoureux, G., Harder, E., Vorobyov, I. V., Roux, B., and MacKerell, A. D. (2006) A polarizable model of water for molecular dynamics simulations of biomolecules. *Chem. Phys. Lett. 418*, 245−249.
(148) Woods, A. S., and Ferre, S. (2005) Amazing stability of the arginine-phosphate electrostatic interaction. *J. Proteome Res. 4*, 1397−1402.
(149) Leontyev, I., and Stuchebrukhov, A. (2011) Accounting for electronic polarization in non-polarizable force fields. *Phys. Chem. Chem. Phys. 13*, 2613−2626.
(150) Leontyev, I. V., and Stuchebrukhov, A. A. (2010) Electronic continuum model for molecular dynamics simulations of biological molecules. *J. Chem. Theory Comput. 6*, 1498−1508.
(151) Zhu, K., Shirts, M. R., and Friesner, R. A. (2007) Improved methods for side chain and loop predictions via the protein local optimization program: Variable dielectric model for implicitly improving the treatment of polarization effects. *J. Chem. Theory Comput. 3*, 2108−2119.
(152) Luo, Y., and Roux, B. (2010) Simulation of osmotic pressure in concentrated aqueous salt solutions. *J. Phys. Chem. Lett. 1*, 183−189.
(153) Berneche, S., and Roux, B. (2001) Energetics of ion conduction through the K^+ channel. *Nature 414*, 73−77.
(154) Venable, R. M., Luo, Y., Gawrisch, K., Roux, B., and Pastor, R. W. (2013) Simulations of anionic lipid membranes: Development of interaction-specific ion parameters and validation using NMR data. *J. Phys. Chem. B 117*, 10183−10192.
(155) Yoo, J. J., and Aksimentiev, A. (2012) Improved parametrization of Li^+, Na^+, K^+, and Mg^{2+} ions for all-atom molecular dynamics simulations of nucleic acid systems. *J. Phys. Chem. Lett. 3*, 45−50.
(156) Simonson, T., and Satpati, P. (2013) Simulating GTP:Mg and GDP:Mg with a simple force field: A structural and thermodynamic analysis. *J. Comput. Chem. 34*, 836−846.
(157) Kahlen, J., Salimi, L., Sulpizi, M., Peter, C., and Donadio, D. (2014) Interaction of charged amino-acid side chains with ions: An optimization strategy for classical force fields. *J. Phys. Chem. B 118*, 3960−3972.
(158) Baker, C. M., Lopes, P. E. M., Zhu, X., Roux, B., and MacKerell, A. D. (2010) Accurate calculation of hydration free energies using pair-specific Lennard-Jones parameters in the CHARMM Drude polarizable force field. *J. Chem. Theory Comput. 6*, 1181−1198.
(159) Cerutti, D. S., Rice, J. E., Swope, W. C., and Case, D. A. (2013) Derivation of fixed partial charges for amino acids accommodating a specific water model and implicit polarization. *J. Phys. Chem. B 117*, 2328−2338.
(160) Geney, R., Layten, M., Gomperts, R., Hornak, V., and Simmerling, C. (2006) Investigation of salt bridge stability in a generalized born solvent model. *J. Chem. Theory Comput. 2*, 115−127.
(161) Okur, A., Wickstrom, L., and Simmerling, C. (2008) Evaluation of salt bridge structure and energetics in peptides using explicit, implicit, and hybrid solvation models. *J. Chem. Theory Comput. 4*, 488−498.
(162) Shang, Y., Nguyen, H., Wickstrom, L., Okur, A., and Simmerling, C. (2011) Improving the description of salt bridge strength and geometry in a generalized Born model. *J. Mol. Graphics Modell. 29*, 676−684.