THE PROTEIN SOCIETY

pH-sensitive residues in the P19 RNA silencing suppressor protein from carnation Italian ringspot virus affect siRNA binding stability

Sean M. Law, Bin W. Zhang, and Charles L. Brooks III*

Department of Chemistry and Biophysics Program, University of Michigan, Ann Arbor, Michigan 48109

Received 18 December 2012; Revised 6 February 2013; Accepted 10 February 2013
DOI: 10.1002/pro.2243
Published online 1 March 2013 proteinscience.org

Charles L. Brooks III is the recipient of the Protein Society 2012 Hans Neurath Award.

Abstract: *Tombusviruses*, such as Carnation Italian ringspot virus (CIRV), encode a protein homodimer called p19 that is capable of suppressing RNA silencing in their infected hosts by binding to and sequestering short-interfering RNA (siRNA) away from the RNA silencing pathway. P19 binding stability has been shown to be sensitive to changes in pH but the specific amino acid residues involved have remained unclear. Using constant pH molecular dynamics simulations, we have identified key pH-dependent residues that affect CIRV p19–siRNA binding stability at various pH ranges based on calculated changes in the free energy contribution from each titratable residue. At high pH, the deprotonation of Lys60, Lys67, Lys71, and Cys134 has the largest effect on the binding stability. Similarly, deprotonation of several acidic residues (Asp9, Glu12, Asp20, Glu35, and/or Glu41) at low pH results in a decrease in binding stability. At neutral pH, residues Glu17 and His132 provide a small increase in the binding stability and we find that the optimal pH range for siRNA binding is between 7.0 and 10.0. Overall, our findings further inform recent experiments and are in excellent agreement with data on the pH-dependent binding profile.

Keywords: pH-dependence; CIRV p19; Constant pH molecular dynamics simulations; protein–RNA interactions

Introduction

RNA silencing (or RNA interference [RNAi])[1–3] is an evolutionarily conserved gene inactivation pathway in eukaryotes that involves the conversion of long double-stranded RNA (dsRNA) into 21–24 nucleotide-long short-interfering RNA (siRNA) or microRNA (miRNA) by DICER, an enzyme that is a part of the endoribonuclease family of proteins.[4] These small RNAs are then separated into individual strands, incorporated into a multiprotein complex called RNA-induced silencing complex (RISC),[5] and ultimately used to target the degradation of reasonably complementary messenger RNA (mRNA). In plants, RNA silencing has evolved into a mechanism that can respond to both endogenous and exogenous dsRNA, the latter of which helps to defend against transgenes, transposons, and infection by RNA viruses. As a result, *Tombusviruses* such as Tomato bushy stunt virus (TBSV) and Carnation Italian

Abbreviations: CIRV, carnation Italian ringspot virus; CPHMD, constant pH molecular dynamics; dsRNA, double-stranded RNA; MC, Monte Carlo; MD, molecular dynamics; miRNA, microRNA; RISC, RNA-induced silencing complex; RMSD, root mean squared deviation; RNAi, RNA interference; S_{unprot}, fraction of unprotonated states; siRNA, short-interfering RNA; TBSV, tomato bushy stunt virus

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: NIH; Grant number: GM057513.

*Correspondence to: Charles L. Brooks III, Department of Chemistry and Biophysics Program, University of Michigan, Ann Arbor, MI 48109. E-mail: brookscl@umich.edu

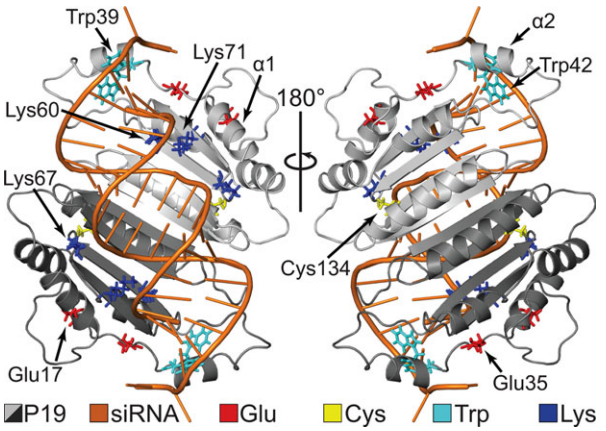

Figure 1. CIRV p19–siRNA complex. The complex has been rotated by 180° (PDBID: 1RPU).

ringspot virus (CIRV) have evolved an elegant 19 kDa protein, aptly named p19, which is capable of suppressing RNA silencing in its host by binding to and sequestering siRNA from a RISC.[6,7]

X-ray structures of p19 bound to a 21 nucleotide (nt) siRNA have been crystallized from TBSV[8] and CIRV.[9] In both structures, p19 exists as a homodimer and is composed of five α-helices and four β-strands in each monomer (Fig. 1). Key tryptophan residues, shown experimentally and computationally to be crucial for siRNA recognition, form end-capping calipers around the RNA by stacking with the exposed terminal base pairs.[8–10] Three conserved lysine residues (Lys60, Lys67, and Lys71) found in the β-sheet–RNA interface form important sequence-independent interactions with the siRNA phosphate backbone.[8,9] Mutations of Lys60 and Lys71 to alanine in TBSV have displayed decreases in the lethal necrosis phenotype.[11,12] Recent experiments have also identified two cysteine residues that appear to be responsible for maintaining the overall structural integrity of the p19 protein as modifications of these cysteines (Cys110 and Cys134 in CIRV) resulted in a reduction of siRNA binding activity.[13,14]

Over the past few years, p19 has been used in several systems to suppress RNAi[15–22] and has also emerged as a valuable tool for characterizing small RNAs.[8,9,23–26] Furthermore, as environmental factors (e.g., acidity/alkalinity, salt concentration, water levels, etc.) can vary significantly across different plant hosts, it has become increasingly important to understand how the protein environment can affect the function of p19. Recent investigations using fluorescence detection assays have revealed that CIRV p19 has the most significant affinity for 21-nt siRNA in the pH range from 6.0 to 9.0.[27] More specifically, it was shown that p19–siRNA binding is dependent on three apparent pK_a values, 7.1, 8.0, and 10.6, that were hypothesized to correspond to one or more histidine, cysteine, and lysine residues, respectively. However, due to the limited resolution of the experiment, the identity of these ionizable residues has remained unknown. Therefore, it is necessary to consider alternative approaches.

Computational methods using molecular dynamics (MD) simulations an/or Monte Carlo (MC) sampling have been developed with considerable success for predicting protein pK_a values (see reviews[28–30]). Often referred to as constant pH MD (CPHMD) simulations, the titration coordinate is typically implemented in either a discrete manner[31–43] where protonation states are modified with an MC step at some regular MD interval or using a continuous function[44–46] that describes the protonation state via the λ dynamics method developed by Brooks and coworkers.[47–49] Recent studies have shown that CPHMD is a reliable and robust method that is capable of predicting pK_a values in a variety of biomolecular systems.[50–56] Thus, to uncover the pH-dependent residues in the CIRV p19 protein involved in siRNA binding stability, we have carried out CPHMD simulations[44–46] of the p19 protein dimer in both holo (siRNA-bound) and apo (siRNA-free) forms and determined the pK_a values for all titratable residues. These results were then used to calculate the pH-dependent siRNA binding stability profile and corresponding pH-dissociation constant profile. Details of the conformational dynamics for important titratable residues at different pH conditions were also investigated and the results were compared with experiment.

Results

CPHMD simulation stability

CPHMD simulations ranging from pH 1 to 14 were performed for both apo and holo systems and the C_α-root-mean-square deviation (RMSD) relative to the crystal structure at different pHs is shown in Supporting Information Figure S1. In general, the apo simulations demonstrated larger average C_α-RMSDs than the holo simulations. The largest C_α-RMSD was ~5.3 Å among all apo simulations (at pH 2) and ~3.7 Å from all holo simulations (at pH 14). Visual inspection of the protein structure from the apo simulations revealed that the p19 core and RNA-binding interface (residues 55–152) were very stable. Instead, the elevated p19 C_α-RMSDs in the apo systems came from the increased dynamics of the α2-helix (residues 39–45) from both monomers (Fig. 1 and Supporting Information Fig. S2). In the holo system, α2 is connected to the N-terminal α1-helix (residues 9–17) by a long flexible linker (residues 18–38) and contains important tryptophan residues (Trp39 and Trp42) that form end-capping calipers around the terminal base pairs of the siRNA.[9] In the apo simulations, these base stacking

Table I. *pK_a Values[a] Calculated from CPHMD for the Holo and Apo States*

Residue	pK_a^{holo}	pK_a^{apo}	ΔpK_a[b]
Asp9	4.17	3.51	0.66
Glu12	5.29	4.30	1.00
Glu17	4.94	6.15	-1.20
Asp20	5.54	4.48	1.06
Asp34	3.56	2.41	1.15
Glu35	5.97	5.20	0.77
Glu41	4.75	4.24	0.51
Lys60	11.11	8.94	2.17
Lys67	11.43	10.26	1.18
Lys71	12.00	9.93	2.06
His132	5.76	6.39	-0.63
Cys134	10.93	10.05	0.88
Glu151	5.15	4.59	0.56

[a] Only pK_a values with $\Delta pK_a \geq 0.5$ are displayed. A full list pK_a values for all titratable residues is provided in Supporting Information Table S1.
[b] $\Delta pK_a = pK_a^{holo} - pK_a^{apo}$.

interactions are lost due to the absence of RNA which resulted in an increase in α2 dynamics and, to a lesser extent, α1 dynamics.

pK_a values calculated from holo and apo simulations

The pK_a values for each Glu, Asp, His, Cys, and Lys residue were obtained by fitting their simulated S_i^{unprot} values (combined from both monomers) to the Henderson–Hasselbalch equation (see Materials and Methods section). The calculated results are summarized in Table I and Supporting Information Table S1.

Effects of pH on p19–siRNA binding

The net charge of both holo and apo p19 as a function of pH was determined from the computed pK_a values [Fig. 2(A)]. Overall, both systems became progressively more negative as the pH was increased until all 26 titratable residues were completely deprotonated. The apo form of the protein was found to be more negatively charged than the holo form except between pH 6 and 8 where the total charge of both forms of the protein was nearly the same. Substituting the pK_a values into the Wyman–Tanford linkage equation, we calculated the pH-dependent changes in the total binding stability ($\Delta\Delta G^{holo \rightarrow apo}$) as well as the individual contributions from each titratable residue [Figs. 2(B) and 3]. At pH > 9, the deprotonation of several conserved lysine residues (Lys60, Lys67, and Lys71) and a nonconserved cysteine residue (Cys134) led to a large destabilization in siRNA binding by about 14.0 kcal/mol. Upon deprotonation at 4 < pH < 6, several acidic residues (Asp9, Glu12, Asp20, Glu35, and Glu41) destabilized the protein–siRNA complex by ~7 kcal/mol. Changes in the binding stability were smallest (<1.0 kcal/mol) between pH 6 and 8, which is consistent with the lack of difference in net charge between both apo and holo systems in the same pH range. Only deprotonation of Glu17 and His132 contributed significantly to stabilizing the siRNA-bound complex in this pH range.

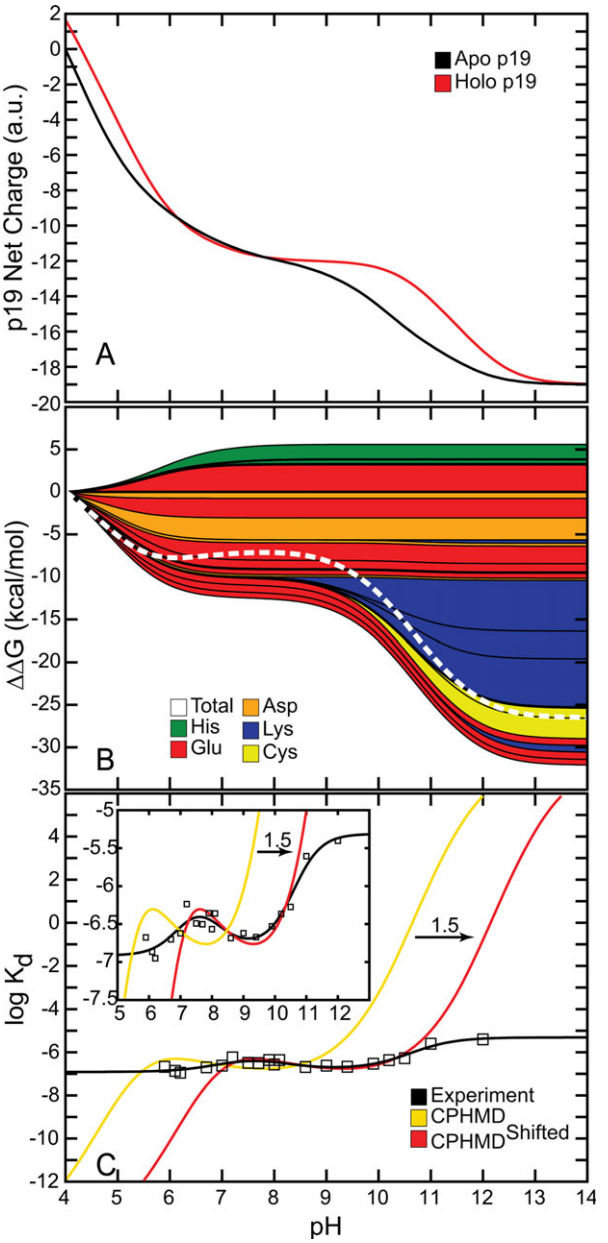

Figure 2. Effects of pH on p19–siRNA binding. (A) Net charge of the holo and apo p19. (B) Total binding stability (white) and per residue contributions to binding stability (colors). Positive $\Delta\Delta G$ values increase binding stability (i.e., favors holo form) while negative $\Delta\Delta G$ values destabilizes binding (i.e., favors apo form). (C) pH-dependent dissociation constant profiles compared with experiment. $CPHMD^{shifted}$ (red) is identical to CPHMD (yellow) except that it is shifted to the right by 1.5 pH units for comparison with experiment (black).

Figure 2(C) shows the dissociation constant profile and a modified profile in comparison with the experimentally determined profile from Koukiekolo *et al*.[27] The unmodified log K_d profile showed minimal change over the last 5 ns of production simulation (see Supporting Information Fig. S3) and was in close agreement with the experimental profile. A much better match with experiment was achieved in the modified profile ($CPHMD^{shifted}$), which was shifted by 1.5 pH units relative to the unmodified profile [see Fig. 2(C) and inset].

pH-dependent conformational dynamics

Various pH-dependent intraprotein and protein–RNA interactions were assessed from the holo simulations (Fig. 4). Deprotonation of Lys60 and Lys67 showed a drastic decrease in Lys–RNA salt bridge formation [Supporting Information Fig. S2 and Fig. 4(A, B)]. However, Lys71, in its neutral form, only displayed a moderate decrease in Lys–RNA interactions [Fig. 4(C)]. Next, measuring the solvent accessible surface area (SASA) for the Cys134 sulfur atom revealed that its side chain was essentially buried when in its reduced form and was much more solvent exposed when negatively charged [Fig. 4(D)]. Upon ionization, Glu17 and Glu35 formed more stable salt bridges with Arg72 and Arg85, respectively [Supporting Information Fig. S2 and Fig. 4(E, F)]. Formation of the Glu35–Arg85 salt bridge also appears to stabilize the Trp39–RNA base stacking interactions but has little to no effect on Trp42–RNA interactions [Supporting Information Fig. S2 and Fig. 4(G, H)].

Discussion

The primary goal of this study was to assess the overall p19–siRNA binding stability and to identify the important pH-sensitive residues that affect siRNA binding. Previously, Koukiekolo *et al*. hypothesized that p19–siRNA binding is dependent on the ionization of one or more histidine, cysteine, and lysine residues.[27] They determined this by fitting fluorescence data to an equation that represents the titration of three apparent pK_a values (found to be 7.1, 8.0, and 10.6) and then attributed these numbers to a particular type of residue based on each residue's reference pK_a value. However, the detailed resolution needed to pinpoint the residues associated with these experimental pK_a values is well beyond the capacity of their assays. Thus, using atomic-level resolution CPHMD simulations, we have computed pK_a values for 26 titratable residues from the p19 protein dimer in both holo and apo systems and compared our results with the current literature. As pointed out previously, the observed changes in the binding stability are likely the result of the deprotonation/protonation of titratable residues that interact with the siRNA and/or the result of local pH-sensitive changes that affect the structural integrity of the p19 protein.[27]

Overall, the 14 holo and 14 apo CPHMD simulations showed remarkable stability in the p19 structure as demonstrated by their C_α-RMSD in different pH environments (Supporting Information Fig. S1). This structural stability is consistent with traditional explicit solvent MD simulations of CIRV p19 with fixed protonation states.[10] The holo simulations displayed slightly lower C_α-RMSD values compared with the apo systems due to the presence of the siRNA. Visual inspection of the apo p19 simulations showed that the differences in RMSD were caused by the loss of Trp39/Trp42–RNA end-capping interactions which resulted in a significant increase in the α2-helix dynamics along with added mobility in the α1-helix (Fig. 1 and Supporting Information Fig. S2). The flexibility found in the N-terminal region of p19 is in line with a model where a negatively charged siRNA first binds to the positively charged p19 β-sheet surface and then the tryptophan residues act as calipers to measure the length of the bound dsRNA by stacking with the terminal base pairs.[9,10]

Using the pK_a values calculated from all 26 titratable residues in the holo and apo p19 simulations (Table I and Supporting Information Table S1), we computed the net charge of both p19 systems as well as the pH-dependent binding stability ($\Delta\Delta G^{holo\rightarrow apo}$) of the entire system and for each titratable residue [Fig. 2(A, B)]. $\Delta\Delta G^{holo\rightarrow apo} > 0$ increases binding stability while $\Delta\Delta G^{holo\rightarrow apo} < 0$ decreases binding stability. Overall, both systems became more negative as each titratable residue was deprotonated due to an increase in pH and, in general, the net charge of the holo system was more positive than the apo system [Fig. 2(A)]. This was expected as a more negative net charge would result in siRNA dissociation due to the charge–charge repulsion between the negatively charged siRNA and the protein in the holo system. Between pH 6 and pH 8, the total charge for both systems was nearly identical. It is also interesting to note that, with the exception of Cys134 (not conserved) and His132 (charged conserved as arginine), 11 of the 13 titratable residues found to contribute significantly to the binding stability (Table I) are well conserved across the *Tombusvirus* p19 family[9] which generally implies some level of functional importance.

At pH > 9, the side chains of Lys60, Lys67, and Lys71 become neutralized and, as a result, their direct interactions with the negatively charged siRNA backbone are reduced [Figs. 3 and 4(A–C)]. Deprotonation of these three conserved residues has the largest destabilizing effect on siRNA binding as reflected in the 14.0 kcal/mol drop in free energy [Fig. 2(B)]. This is consistent with past mutations of Lys60 and Lys71 that resulted in decreases in the

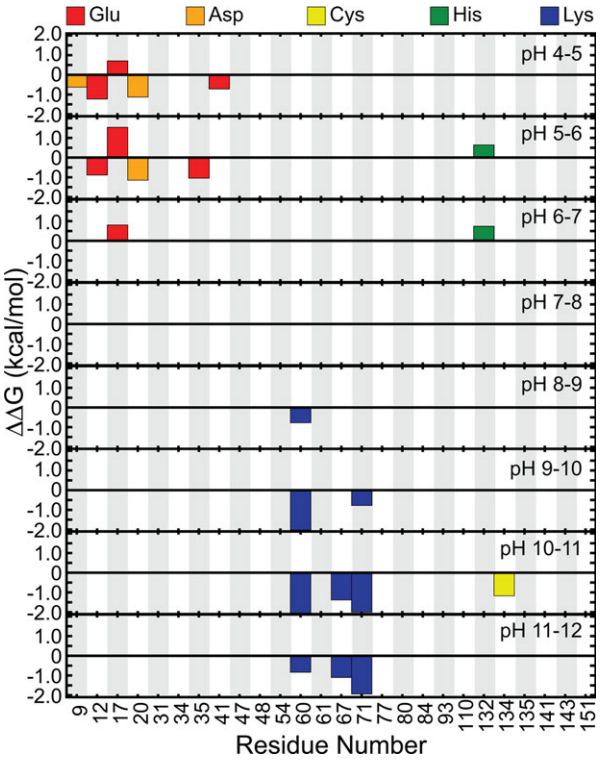

Figure 3. Individual contributions to binding stability at different pH ranges. Negative and positive contributions to binding stability will have $\Delta\Delta G < 0$ and $\Delta\Delta G > 0$, respectively, for a given titratable residue. Only residues with $|\Delta\Delta G| > k_B T$. are shown where k_B. is the Boltzmann constant and T is the temperature (298 K).

lethal necrosis phenotype.[11,12] Thus, we hypothesize that mutating one or more of these lysine residues to arginine, which has a much higher pK_a and therefore would remain positively charged, may help to increase the binding stability at higher pH ranges. In addition, Cys134 was found to be highly solvent exposed when it is deprotonated [Figs 3 and 4(D)]. Koukiekolo *et al*. suggested that the role of the cysteine amino acid is to preserve the structural integrity of the protein and that the deprotonation of cysteine (or other titratable residues) could lead to structural changes that could either destabilize the p19 dimer or the p19–siRNA complex.[13,14,27] From our simulations, we suggest that Cys134 becomes more solvent exposed to prevent having a buried charge that could affect the stability of the local p19 structure. Therefore, we proffer that mutation of Cys134 to its isosteric equivalent, serine, may be beneficial to the CIRV p19 binding affinity.

At 4 < pH < 6, deprotonation of Glu35 not only facilitates the formation of a salt bridge with Arg85 [Figs 3 and 4(F)] but it also leads to an increase in base stacking interactions between the nearby Trp39 and terminal RNA base pair [Fig. 4(G)]. Surprisingly, the Trp42–RNA base stacking interactions were not affected [Fig. 4(H)]. Glu12 and Asp20 are located near the p19–siRNA binding surface and likely destabilize the complex once they are deprotonated by conferring strong electrostatic repulsion with the siRNA (Fig. 3). Based on this observation, we hypothesize that replacing Glu12 and/or Asp20 to neutral glutamine and asparagine, respectively, would increase the overall binding stability.

At neutral pH, there was a small increase in $\Delta\Delta G^{holo \rightarrow apo}$ that was caused by the deprotonation of Glu17 and His132 [Figs. 2(B) and 3]. In the crystal structure, His132 is positioned beside a buried Arg117 which is expected to be structurally less stable when both residues are protonated.[9] We speculate that the neutralization of His132 reduces the local concentration of positive charges and ultimately stabilizes the p19 structure. In contrast, ionized Glu17 appears to facilitate the positioning of key lysine residues (Lys60, Lys67, and Lys71) along the protein–RNA interface by forming salt bridge interactions with Arg72 [Supporting Information Fig. S2 and Fig. 4(E)]. However, detailed correlation analyses revealed that the Glu17–Arg72 salt bridge

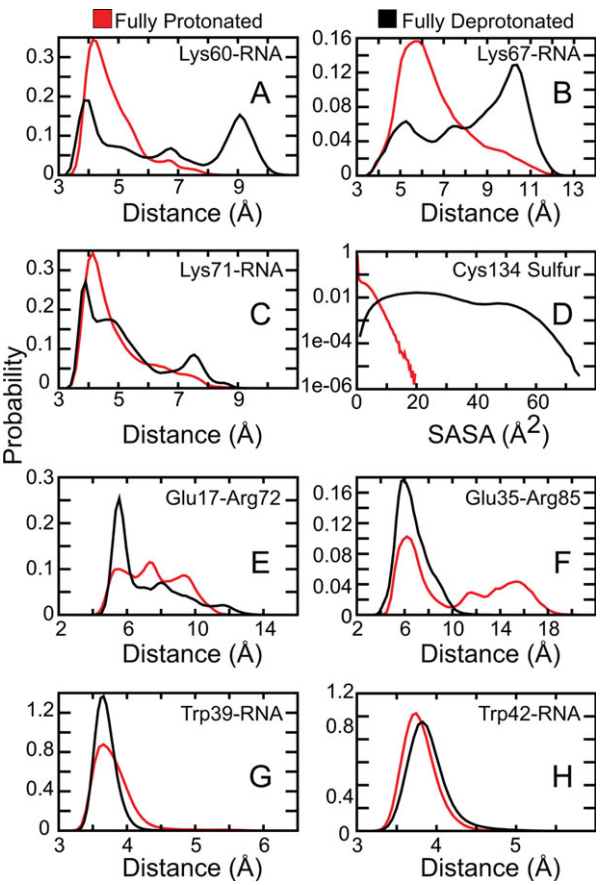

Figure 4. Conformational dynamics of key intraprotein and protein–RNA interactions. Lys–RNA and Glu–Arg distances correspond to salt bridge interactions. Trp–RNA distances and Cys SASA correspond to end-capping interactions and the solvent accessible surface area (SASA) for the cysteine sulfur atom, respectively. See Materials and Methods section for more detail.

formation and the Lys–RNA interactions are basically uncorrelated (with correlation coefficients, R, between -0.04 and 0.22). Alternatively, it has also been proposed that the Glu17–Arg72 salt bridge (and possibly Glu35–Arg85 and Glu41–Arg75 salt bridges) may play a role in positioning the end-capping tryptophan residues.[9,10] The importance of this specific salt bridge is further illustrated by previous studies that swapped Arg72 for glycine and found that the activity of p19 was severely compromised.[11,12,57,58]

The computed dissociation constant profile [Fig. 2(C)] showed little change after the first 15 ns of production simulation time (Supporting Information Fig. S3) and was in good agreement with experiment.[13] This suggests that the results from the CPHMD simulations were converged. Also, a much better correspondence with experiment was obtained when we shifted the dissociation constant profile horizontally by 1.5 pH units [see "CPHMD$^{\text{shifted}}$" in Fig. 2(C) and inset]. The shift in the log K_d profile can be attributed to an underestimation of the desolvation energy and has been shown to lead to a systematic underestimation of the pK_a values (from which the dissociation constant profiles are derived).[59] Henceforth, discussions of the computed dissociation constant profile will be in reference to the shifted profile (CPHMD$^{\text{shifted}}$).

The optimal range for siRNA binding appears to be between pH 7 and 10 where the binding stability fluctuates between $-7.0 \leq \log K_d \leq -6.5$ [Fig. 2(C)]. From our simulations, we can attribute the highest experimentally observed pK_a of 10.6[27] to residues Lys60, Lys67, Lys71, and Cys134 (Fig. 3). These residues demonstrated the largest reduction in binding stability at high pH [Fig. 2(B)] and are excellent candidates for further mutational studies (see discussion above). Similarly, the lowest experimental pK_a value, 7.1, can be assigned to multiple residues (i.e., Asp9, Glu12, Asp20, Glu35, and/or Glu41) (Fig. 3). Ionization of all these residues appeared to have a synergistic destabilizing effect on the p19–siRNA complex. Finally, the last observed apparent pK_a of 8.0 corresponded to residues Glu17 and His132. These two residues were the only groups that were beneficial for significantly increasing the binding affinity upon deprotonation and we found them to be important for maintaining the structural integrity of p19.

Conclusions

In summary, we presented CPHMD simulations of a large protein–RNA complex in implicit solvent. Overall, the results agree well with experiment. We identified several titratable residues that are highly pH-dependent and that could be assigned to experimentally observed pK_a values. Lys60, Lys67, Lys71, and Cys134 appear to affect binding stability at pH > 9 while several glutamic and aspartic acids destabilize the complex at pH < 6. These residues were found to be important for maintaining the stability of the protein structure and/or for siRNA binding. The optimal pH range for siRNA binding is from about 7.0 to 10.0 and is largely stabilized by Glu17 and His132.

The CPHMD method has developed into an accurate and powerful tool for predicting protein pK_a values[30] and for generating pH-dependent binding stability curves that can be directly compared with experiment. Ultimately, identifying the key pH-sensitive residues using the CPHMD approach would allow us to design p19 proteins that have a higher affinity for siRNA which could be used to characterize RNA silencing complexes, to manage cellular levels of siRNA levels, and for discriminating between single-stranded RNA and dsRNA, and so forth.[13,23,26,60] Furthermore, understanding the pH-dependence of the viral protein could enable us to engineer plants that can survive outside of the virulent pH range and avoid infection altogether. This study clearly illustrates the value of complementing experiment with theoretical techniques and offers results that can be further validated.

Materials and Methods

Simulation setup

The CIRV p19 X-ray crystal structure bound to a 21-nt siRNA (PDBID: 1RPU)[9] was used as our model. The unresolved short linker (residues 49–54) in each protein monomer was constructed using MODELLER[61,62] and the loop modeling facility in the MMTSB Tool Set.[63] Missing hydrogen atoms were added using the HBUILD algorithm in the CHARMM simulation package.[64] The holo (siRNA-bound) and apo (siRNA-free) systems contain 5889 and 4551 atoms, respectively. All simulations were performed in implicit solvent using the generalized born with a simple switching (GBSW) model[65,66] implemented in CHARMM along with the CHARMM27 protein–nucleic acid force field[67,68] and an energy correction map (CMAP).[69] A 50 ps^{-1} friction coefficient was used for Langevin dynamics and the experimental salt concentration of 0.1 M[27] was simulated using the Debye–Hückel model for screening charge–charge interactions.[70] Consistent with previous GBSW simulations, the nonpolar surface tension coefficient was set to 0.005 kcal/mol/Å^2 [53,65,71,72] and a switching cutoff that reduces the electrostatic solvation and van der Waals contributions to zero beginning from 20 Å to 24 Å was used. Optimized atomic radii for proteins[65,73] and nucleic acids[74] were used in place of the standard van der Waals radii for the GBSW calculations. The siRNA was harmonically restrained to its initial starting position using a 2 kcal/mol/Å^2 force constant to prevent large structural changes in the RNA. Both holo

and apo systems were energy minimized and heated up slowly to a final temperature of 298 K followed by 500 ps of equilibration using a 1 fs simulation time step. All molecular images were created using PyMOL.[75]

Constant pH MD simulations and pK$_a$ calculations

A total of 28 independent simulations (14 for siRNA-bound and 14 for siRNA-free) ranging from pH 1 to 14 were conducted using the CPHMD methodology.[44–46] Each simulation was 25 ns long and resulted in a cumulative simulation time of 0.7 μs. The first 5 ns of each simulation were discarded in the analysis so that all 28 production simulations were each 20 ns long. Atomic charges for protonated and unprotonated states of aspartic acid, glutamic acid, histidine, lysine, and cysteine have been described previously.[44,68,76,77] Arginines, whose model pK_a value is typically ~12,[78] were kept permanently in its protonated form because CIRV p19 was found experimentally to be unstable at pH > 12[27] All titratable residues were simulated following the CPHMD method originally developed in the Brooks research group where a continuous titration coordinate, $0 > \lambda_i \leq 1$, controls the protonation state for the ith titratable residue.[44–46] In that model, $\lambda_i = 1$ and $\lambda_i = 0$ correspond to the fully unprotonated and fully protonated states, respectively, and $N(\lambda_i)$ is the number of simulation snapshots with protonation state λ. However, to increase the number of times that a titratable residue is considered to be fully protonated (N_i^{unprot}) or fully unprotonated (N_i^{prot}), we have defined a more generous cutoff for λ:

$$\begin{aligned} N_i^{\text{unprot}} &= N(\lambda_i \geq 0.9) \\ N_i^{\text{prot}} &= N(\lambda_i \leq 0.1) \end{aligned} \tag{1}$$

Thus, the fraction of unprotonated states, S_i^{unprot}, is given by:

$$S_i^{\text{unprot}}(\text{pH}) = \frac{N_i^{\text{unprot}}}{N_i^{\text{unprot}} + N_i^{\text{prot}}} \tag{2}$$

and the pK_a of the ith titratable residue can be calculated by fitting a set of S^{unprot} (at different pH values) to the standard Hendersen–Hasselbalch equation:

$$S^{\text{unprot}} = \frac{1}{1 + 10^{n(pK_a - pH)}} \tag{3}$$

where n represents the Hill coefficient. It has been discussed previously that small deviations in the Hill coefficient away from 1 have a negligible effect on the free energy[53,54] and, indeed, we find only small differences in n during the curve fitting process. Thus, we have set $n = 1$ for all our calculations. Finally, due to the fact that the experimentally determined pK_a values were extracted from a homodimer, we have combined the data from both monomers to effectively double the sampling for calculating S_i^{unprot} and its corresponding pK_a.

pH-dependent binding stability and pH-dissociation constant profiles

The pH-dependent binding stability profile was calculated using the Wyman–Tanford linkage equation[79,80]:

$$\begin{aligned} \partial \Delta G / \partial \text{pH} &= \ln(10)\text{RT}\Delta Q(\text{pH})^{\text{holo}\rightarrow\text{apo}} \\ &= \ln(10)\text{RT}\left(Q(\text{pH})^{\text{apo}} - Q(\text{pH})^{\text{holo}}\right) \end{aligned} \tag{4}$$

where ΔG is the dissociation free energy, R is the gas constant, and T is the temperature in Kelvins. $\Delta Q(\text{pH})^{\text{holo}\rightarrow\text{apo}}$ is the difference in the net charge between the holo and apo states at a particular pH and is calculated from:

$$\Delta Q^{\text{holo}\rightarrow\text{apo}} = \sum_i \langle \rho(i) \rangle_{\text{apo}} - \sum_i \langle \rho(i) \rangle_{\text{holo}} \tag{5}$$

The average charge of the system, $\langle \rho(i) \rangle$, is obtained from:

$$\langle \rho(i) \rangle = -S_i^{\text{unprot}} + \frac{\gamma(i) + 1}{2} \tag{6}$$

where $\gamma(i)$, defined previously,[81] is equal to 1 or −1 for a basic and acidic group, respectively. Integrating Eq. (4) after substituting in Eqs. (5) and (6) gives the dissociation free energy at a given pH relative to a reference pH (pH$_{\text{ref}}$):

$$\begin{aligned} \Delta\Delta G^{\text{holo}\rightarrow\text{apo}} &= \Delta G^{\text{holo}\rightarrow\text{apo}}(\text{pH}) - \Delta G^{\text{holo}\rightarrow\text{apo}}(\text{pH}_{\text{ref}}) \\ &= \ln(10)\text{RT} \int_{\text{pH}_{\text{ref}}}^{\text{pH}} \Delta Q^{\text{holo}\rightarrow\text{apo}} \text{pH} \\ &= \text{RT} \sum_i \ln \frac{\left(1 + 10^{pK_a^{\text{holo}}(i) - \text{pH}}\right)\left(1 + 10^{pK_a^{\text{apo}}(i) - pH_{\text{ref}}}\right)}{\left(1 + 10^{pK_a^{\text{apo}}(i) - \text{pH}}\right)\left(1 + 10^{pK_a^{\text{holo}}(i) - pH_{\text{ref}}}\right)} \end{aligned} \tag{7}$$

where $pK_a^{holo}(i)$ and $pK_a^{apo}(i)$ are the pK_as for the holo and apo states, respectively. The final summation in Eq. (7) allows the binding stability to be decomposed into contributions from each titratable residue. The final dissociation constant profile was then computed by applying the basic relation:

$$\begin{aligned} \log K_{\text{d}} &= \ln(10)\frac{-\Delta G}{\text{RT}} \\ &= \ln(10)\frac{-\left[\Delta\Delta G^{\text{holo}\rightarrow\text{apo}} + \Delta G^{\text{holo}\rightarrow\text{apo}}(\text{pH}_{\text{ref}})\right]}{\text{RT}} \end{aligned} \tag{8}$$

where K_{d} is the dissociation constant. However, as Eq. (7) gives us $\Delta\Delta G^{\text{holo}\rightarrow\text{apo}}$, then the resulting

$\log K_d(\mathrm{pH})$ in Eq. (8) at any given pH depends on $\Delta\Delta G^{\mathrm{holo}\rightarrow\mathrm{apo}}(\mathrm{pH})$ and $\Delta G^{\mathrm{holo}\rightarrow\mathrm{apo}}(\mathrm{pH}_{\mathrm{ref}})$, the latter of which is a constant. Thus, $\Delta G^{\mathrm{holo}\rightarrow\mathrm{apo}}(\mathrm{pH}_{\mathrm{ref}})$ was chosen to allow the best match of $\log K_d(\mathrm{pH})$ to experiment.

Side chain conformational dynamics

Motivated by previous discussions,[8,9,27] the conformational dynamics of several different intraprotein and protein–RNA interactions were analyzed (from the holo simulations) by comparing the normalized probability of an interaction when a particular titratable side chain is either fully protonated ($\mathrm{pH} \gg pK_a$) or fully deprotonated ($\mathrm{pH} \ll pK_a$). Glu–Arg salt bridge distances were measured from the Glu-C_δ atom to the Arg-C_ζ atom. Trp–RNA base stacking distances were measured from the center-of-mass of the Trp side chain (not titrated) to the center-of-mass of the closest RNA base. Lys–RNA salt bridges were measured from the Lys-N_ζ atom to the closest RNA backbone phosphorus atom. The SASA was measured for the Cys-S_γ atom.

Acknowledgment

The authors would like to thank Garrett B. Goh for many valuable scientific discussions.

References

1. Fire A, Xu SQ, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature 391:806–811.
2. Hannon GJ (2002) Rna interference. Nature 418: 244–251.
3. Voinnet O (2001) RNA silencing as a plant immune system against viruses. Trends Genet 17:449–459.
4. Bernstein E, Caudy AA, Hammond SM, Hannon GJ (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 409:363–366.
5. Hutvagner G, Simard MJ (2008) Argonaute proteins: key players in RNA silencing. Nat Rev Mol Cell Bio 9: 22–32.
6. Voinnet O, Pinto YM, Baulcombe DC (1999) Suppression of gene silencing: a general strategy used by diverse DNA and RNA viruses of plants. Proc Natl Acad Sci USA 96:14147–14152.
7. Silhavy D, Molnar A, Lucioli A, Szittya G, Hornyik C, Tavazza M, Burgyan J (2002) A viral protein suppresses RNA silencing and binds silencing-generated, 21-to 25-nucleotide double-stranded RNAs. EMBO J 21:3070–3080.
8. Ye KQ, Malinina L, Patel DJ (2003) Recognition of small interfering RNA by a viral suppressor of RNA silencing. Nature 426:874–878.
9. Vargason JM, Szittya G, Burgyan J, Hall TMT (2003) Size selective recognition of siRNA by an RNA silencing suppressor. Cell 115:799–811.
10. Xia Z, Zhu ZH, Zhu J, Zhou RH (2009) Recognition mechanism of siRNA by viral p19 suppressor of RNA silencing: a molecular dynamics study. Biophys J 96: 1761–1769.
11. Chu M, Desvoyes B, Turina M, Noad R, Scholthof HB (2000) Genetic dissection of tomato bushy stunt virus p19-protein-mediated host-dependent symptom induction and systemic invasion. Virology 266:79–87.
12. Turina M, Omarov R, Murphy JF, Bazaldua-Hernandez C, Desvoyes B, Scholthof HB (2003) A newly identified role for Tomato bushy stunt virus P19 in short distance spread. Mol Plant Pathol 4:67–72.
13. Sagan SM, Koukiekolo R, Rodgers E, Goto NK, Pezacki JP (2007) Inhibition of siRNA binding to a p19 viral suppressor of RNA silencing by cysteine alkylation. Angew Chem Int Ed Engl 46:2005–2009.
14. Cheng J, Koukiekolo R, Kieliszkiewicz K, Sagan SM, Pezacki JP (2009) Cysteine residues of Carnation Italian Ringspot virus p19 suppressor of RNA silencing maintain global structural integrity and stability for siRNA binding. Biochim Biophys Acta 1794:1197–1203.
15. Papp I, Mette MF, Aufsatz W, Daxinger L, Schauer SE, Ray A, van der Winden J, Matzke M, Matzke AJM (2003) Evidence for nuclear processing of plant micro RNA and short interfering RNA precursors. Plant Physiol 132:1382–1390.
16. Lakatos L, Szittya G, Silhavy D, Burgyan J (2004) Molecular mechanism of RNA silencing suppression mediated by p19 protein of tombusviruses. EMBO J 23:876–884.
17. Chapman EJ, Prokhnevsky AI, Gopinath K, Dolja VV, Carrington JC (2004) Viral RNA silencing suppressors inhibit the microRNA pathway at an intermediate step. Gene Dev 18:1179–1186.
18. Dunoyer P, Lecellier CH, Parizotto EA, Himber C, Voinnet O (2004) Probing the microRNA and small interfering RNA pathways with virus-encoded suppressors of RNA silencing. Plant Cell 16:1235–1250.
19. Li WX, Li HW, Lu R, Li F, Dus M, Atkinson P, Brydon EWA, Johnson KL, Garcia-Sastre A, Ball LA, Palese P, Ding SW (2004) Interferon antagonist proteins of influenza and vaccinia viruses are suppressors of RNA silencing. Proc Natl Acad Sci USA 101:1350–1355.
20. Lecellier CH, Dunoyer P, Arar K, Lehmann-Che J, Eyquem S, Himber C, Saib A, Voinnet O (2005) A cellular MicroRNA mediates antiviral defense in human cells. Science 308:557–560.
21. Lu R, Maduro M, Li F, Li HW, Broitman-Maduro G, Li WX, Ding SW (2005) Animal virus replication and RNAi-mediated antiviral silencing in *Caenorhabditis elegans*. Nature 436:1040–1043.
22. Voinnet O, Rivas S, Mestre P, Baulcombe D (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. Plant J 33:949–956.
23. Cheng J, Sagan SM, Assem N, Koukiekolo R, Goto NK, Pezacki JP (2007) Stabilized recombinant suppressors of RNA silencing: functional effects of linking monomers of Carnation Italian Ringspot virus p19. Biochim Biophys Acta 1774:1528–1535.
24. Calabrese JM, Sharp PA (2006) Characterization of the short RNAs bound by the P19 suppressor of RNA silencing in mouse embryonic stem cells. RNA 12: 2092–2102.
25. Jin JM, Cid M, Poole CB, McReynolds LA (2010) Protein mediated miRNA detection and siRNA enrichment using p19. Biotechniques 48:Xvii–Xxiii.
26. Koukiekolo R, Jakubek ZJ, Cheng J, Sagan SM, Pezacki JP (2009) Studies of a viral suppressor of RNA silencing p19-CFP fusion protein: a FRET-based probe for sensing double-stranded fluorophore tagged small RNAs. Biophys Chem 143:166–169.
27. Koukiekolo R, Sagan SM, Pezacki JP (2007) Effects of pH and salt concentration on the siRNA binding activity of the RNA silencing suppressor protein p19. FEBS Lett 581:3051–3056.

28. Alexov E, Mehler EL, Baker N, Baptista AM, Huang Y, Milletti F, Nielsen JE, Farrell D, Carstensen T, Olsson MHM, Shen JK, Warwicker J, Williams S, Word JM (2011) Progress in the prediction of pK(a) values in proteins. Proteins 79:3260–3275.
29. Mongan J, Case DA (2005) Biomolecular simulations at constant pH. Curr Opin Struct Biol 15:157–163.
30. Wallace JA, Shen JK (2009) Predicting pKa values with continuous constant pH molecular dynamics. Methods Enzymol 466:455–475.
31. Burgi R, Kollman PA, van Gunsteren WF (2002) Simulating proteins at constant pH: an approach combining molecular dynamics and Monte Carlo simulation. Proteins 47:469–480.
32. Baptista AM, Teixeira VH, Soares CM (2002) Constant-pH molecular dynamics using stochastic titration. J Chem Phys 117:4184–4200.
33. Machuqueiro M, Baptista AM (2006) Constant-pH molecular dynamics with ionic strength effects: protonation-conformation coupling in decalysine. J Phys Chem B 110:2927–2933.
34. Machuqueiro M, Baptista AM (2008) Acidic range titration of HEWL using a constant-pH molecular dynamics method. Proteins 72:289–298.
35. Machuqueiro M, Baptista AM (2009) Molecular dynamics at constant pH and reduction potential: application to cytochrome c(3). J Am Chem Soc 131:12586–12594.
36. Dlugosz M, Antosiewicz JM (2004) Constant-pH molecular dynamics simulations: a test case of succinic acid. Chem Phys 302:161–170.
37. Dlugosz M, Antosiewicz JM, Robertson AD (2004) Constant-pH molecular dynamics study of protonation-structure relationship in a heptapeptide derived from ovomucoid third domain. Phys Rev E 69:021915.
38. Mongan J, Case DA, McCammon JA (2004) Constant pH molecular dynamics in generalized born implicit solvent. J Comput Chem 25:2038–2048.
39. Mongan J, Case DA, McCammon J (2004) Discrete-state constant pH molecular dynamics in generalized born implicit solvent. Protein Sci 13:219.
40. Meng YL, Roitberg AE (2010) Constant pH replica exchange molecular dynamics in biomolecules using a discrete protonation model. J Chem Theory Comput 6: 1401–1412.
41. Williams SL, de Oliveira CAF, McCammon JA (2010) Coupling constant pH molecular dynamics with accelerated molecular dynamics. J Chem Theory Comput 6: 560–568.
42. Messer BM, Roca M, Chu ZT, Vicatos S, Kilshtain AV, Warshel A (2010) Multiscale simulations of protein landscapes: using coarse-grained models as reference potentials to full explicit models. Proteins 78: 1212–1227.
43. Aqvist J, Warshel A (1993) Simulation of enzyme-reactions using valence-bond force-fields and other hybrid quantum-classical approaches. Chem Rev 93:2523–2544.
44. Lee MS, Salsbury FR, Brooks CL III (2004) Constant-pH molecular dynamics using continuous titration coordinates. Proteins 56:738–752.
45. Khandogin J, Brooks CL III (2005) Constant pH molecular dynamics with proton tautomerism. Biophys J 89: 141–157.
46. Khandogin J, Brooks CL III (2006) Toward the accurate first-principles prediction of ionization equilibria in proteins. Biochemistry 45:9363–9373.
47. Guo Z, Brooks CL III, Kong X (1998) Efficient and flexible algorithm for free energy calculations using the lambda-dynamics approach. J Phys Chem B 102: 2032–2036.
48. Knight JL, Brooks CL III (2009) lambda-dynamics free energy simulation methods. J Comput Chem 30:1692–1700.
49. Kong XJ, Brooks CL III (1996) lambda-Dynamics: a new approach to free energy calculations. J Chem Phys 105:2414–2423.
50. Wallace JA, Shen JK (2011) Continuous constant pH molecular dynamics in explicit solvent with pH-based replica exchange. J Chem Theory Comput 7:2617–2629.
51. Goh GB, Knight JL, Brooks CL III (2012) Constant pH molecular dynamics simulations of nucleic acids in explicit solvent. J Chem Theory Comput 8:36–46.
52. Zhang BW, Brunetti L, Brooks CL III (2011) Probing pH-dependent dissociation of HdeA dimers. J Am Chem Soc 133:19393–19398.
53. Shen JK (2010) Uncovering specific electrostatic interactions in the denatured states of proteins. Biophys J 99:924–932.
54. Shen JK (2010) A method to determine residue-specific unfolded-state pK(a) values from analysis of stability changes in single mutant cycles. J Am Chem Soc 132: 7258.
55. Wallace JA, Shen JK (2012) Unraveling a trap-and-trigger mechanism in the pH-sensitive self-assembly of spider silk proteins. J Phys Chem Lett 3:658–662.
56. Arthur EJ, Yesselman JD, Brooks CL III (2011) Predicting extreme pK(a) shifts in staphylococcal nuclease mutants with constant pH molecular dynamics. Proteins 79:3276–3286.
57. Uhrig JF, Canto T, Marshall D, MacFarlane SA (2004) Relocalization of nuclear ALY proteins to the cytoplasm by the tomato bushy stunt virus P19 pathogenicity protein. Plant Physiol 135:2411–2423.
58. Park JW, Faure-Rabasse S, Robinson MA, Desvoyes B, Scholthof HB (2004) The multifunctional plant viral suppressor of gene silencing P19 interacts with itself and an RNA binding host protein. Virology 323:49–58.
59. Wallace JA, Wang YH, Shi CY, Pastoor KJ, Nguyen BL, Xia K, Shen JK (2011) Toward accurate prediction of pK(a) values for internal protein residues: the importance of conformational relaxation and desolvation energy. Proteins 79:3364–3373.
60. Scholthof HB (2006) Timeline—the Tombusvirus-encoded P19: from irrelevance to elegance. Nat Rev Microbiol 4:405–411.
61. Fiser A, Do RKG, Sali A (2000) Modeling of loops in protein structures. Protein Sci 9:1753–1773.
62. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234:779–815.
63. Feig M, Karanicolas J, Brooks CL III (2004) MMTSB tool set: enhanced sampling and multiscale modeling methods for applications in structural biology. J Mol Graphics Model 22:377–395.
64. Brooks BR, Brooks CL III, Mackerell AD, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M (2009) CHARMM: the biomolecular simulation program. J Comput Chem 30:1545–1614.
65. Chen JH, Im WP, Brooks CL III (2006) Balancing solvation and intramolecular interactions: toward a consistent generalized born force field. J Am Chem Soc 128:3728–3736.
66. Im WP, Lee MS, Brooks CL III (2003) Generalized born model with a simple smoothing function. J Comput Chem 24:1691–1702.

67. Foloppe N, MacKerell AD (2000) All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. J Comput Chem 21:86–104.
68. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 102:3586–3616.
69. MacKerell AD, Feig M, Brooks CL III (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.
70. Srinivasan J, Trevathan MW, Beroza P, Case DA (1999) Application of a pairwise generalized Born model to proteins and nucleic acids: inclusion of salt effects. Theor Chem Acc 101:426–434.
71. Khandogin J, Chen JH, Brooks CL III (2006) Exploring atomistic details of pH-dependent peptide folding. Proc Natl Acad Sci USA 103:18546–18550.
72. Khandogin J, Brooks CL III (2007) Linking folding with aggregation in Alzheimer's beta-amyloid peptides. Proc Natl Acad Sci USA 104:16880–16885.
73. Nina M, Beglov D, Roux B (1997) Atomic radii for continuum electrostatics calculations based on molecular dynamics free energy simulations. J Phys Chem B 101: 5239–5248.
74. Banavali NK, Roux B (2002) Atomic radii for continuum electrostatics calculations on nucleic acids. J Phys Chem B 106:11026–11035.
75. Schrödinger L. The PyMOL Molecular Graphics System, Version 1.3. Schrödinger, LLC.
76. Foloppe N, Sagemark J, Nordstrand K, Berndt KD, Nilsson L (2001) Structure, dynamics and electrostatics of the active site of glutaredoxin 3 from Escherichia coli: comparison with functionally related proteins. J Mol Biol 310:449–470.
77. Foloppe N, Nilsson L (2004) The glutaredoxin -C-P-Y-C- motif: influence of peripheral residues. Structure 12: 289–300.
78. Nozaki Y, Tanford C, Examination of titration behavior. In: Hirs CHW, Ed. (1967) Methods enzymol. New York: Academic Press, Vol. 11, pp 715–734.
79. Tanford C (1970) Protein denaturation: part C. Theoretical models for the mechanism of denaturation. Adv Prot Chem 24:1–95.
80. Wyman J (1964) Linked functions and reciprocal effects in hemoglobin—a 2nd look. Adv Prot Chem 19: 223–286.
81. Yang AS, Honig B (1993) On the pH-dependence of protein stability. J Mol Biol 231:459–474.