Dynamic aspects of pressure and temperature-stabilized intermediates of outer surface protein A

Takuro Wakamoto[1], Soichiro Kitazawa[1], Tomoshi Kameda[2], and Ryo Kitahara[3]

[1]Ritsumeikan Daigaku - Biwako Kusatsu Campus
[2]National Institute of Advanced Industrial Science and Technology Tokyo Bay Area Center
[3]Ritsumeikan University - Biwako Kusatsu Campus

April 28, 2020

Abstract

Structural characterization of alternatively folded and partially disordered protein conformations remains challenging. Outer surface protein A (OspA) is a pivotal protein in Borrelia infection, which is the etiological agent of Lyme disease. OspA exists in equilibrium with intermediate conformations, in which the central and the C-terminal regions of the protein have lower stabilities than the N-terminal. Here, we characterize pressure- and temperature-stabilized intermediates of OspA by nuclear magnetic resonance spectroscopy combined with paramagnetic relaxation enhancement (PRE). We found that the C-terminal region of the intermediate was partially disordered; however, it retains weak specific contact with the N-terminal region, owing to a twist of the central β-sheet and increased flexibility in the polypeptide chain. The disordered C-terminal region of the pressure-stabilized intermediate was more compact than that of the temperature-stabilized form. Further, molecular dynamics simulation demonstrated that temperature-induced disordering of the β-sheet was initiated at the C-terminal region and continued through to the central region. An ensemble of simulation snapshots qualitatively described the PRE data from the intermediate and indicated that the intermediate structures of OspA may expose tick receptor-binding sites more readily than does the basic folded conformation.

Introduction

Proteins in solution fluctuate between folded and unfolded conformations, which is often related to their functions and toxic aggregate formation.[1-4] Outer surface protein A (OspA) is an immunogenic lipoprotein expressed by the spirochete *Borrelia* (e.g. *B. burgdorferi, B. garinii, and B. afzelii*), the etiologic agent of Lyme disease. It consists of 21 repeated β-strands and a short C-terminal α-helix. OspA also contains two globular domains, the N-terminal (β1–β7) and the C-terminal (β11–C-terminus) domains, as well as a single-layer β-sheet (central β-sheet, β8–β10) that connects the terminal domains.[5,6]

OspA binds to the tick receptor TROSPA when it enters the tick gut.[7] Since the receptor-binding sites are occluded in the interior of the C-terminal domain,[8] their exposure is believed to be required for receptor-binding. Koide et al employed native state hydrogen exchange nuclear magnetic resonance (NMR) spectroscopy to reveal that the C-terminal region, more specifically, the β9–C-terminus, including a portion of the single-layer β-sheet and the C-terminal domain, of this protein is less stable than the N-terminal domain. This study provided the first evidence of disordered intermediate conformations in the C-terminal regions of the protein.[9-11] They also reported that the intermediate became stabilized as temperature increased.[12,13] The temperature-stabilized intermediate was characterized using solution NMR,[12] differential scanning calorimetry,[12,13] and small angle X-ray scattering.[11] Moreover, our group found that the pressure-stabilized intermediate was nearly fully disordered in the entire C-terminal region of the polypeptide chain, from β9 to the C-terminus. The formation of this intermediate was caused by exposure of a large internal cavity in the C-terminal

domain.[14] Recently, Makabe et al reported similar identity between a kinetic intermediate and the equilibrium intermediate.[15] These results indicate that the pressure- and temperature-stabilized intermediates have similar structural characteristics to those existing under physiological conditions.

Transition into the intermediate *in vivo* may accelerate the binding of OspA to the tick receptor, as the intermediate exposes the receptor-binding sites (i.e. residues 236–237 and 242–244)[8] , and can interact with a receptor distant from OspA.[14] In addition, the intermediate may be important for efficient translocation of the protein through the outer membrane.[16] More importantly, OspA and its C-terminal fragment have been included as vaccine candidates to prevent *Borrelia* transmission,[17] some of which have been tested in Phase 3 clinical trials.[18] A recombinant vaccine using *B. burgdorferi* OspA was previously licensed for Lyme disease (LYMErix, SmithKlineBeecham, Pittsburgh, PA, USA) in 1998, however, the manufacturer voluntarily withdrew the product from the market 3 years later.[19] Although human vaccines for Lyme disease are not currently available, genetically modified OspA, which contain protective elements from two different OspA serotypes, continue to represent important vaccine candidates.[18]

Although previous studies have reported that the C-terminal half of the protein (β9–C-terminus) exhibits lower stability than the N-terminal half, our understanding of how the central β-sheet and C-terminal domain are disordered remains limited. Here, we further characterized the pressure- and temperature-stabilized intermediates of OspA using a paramagnetic relaxation enhancement (PRE)-assisted NMR.[20-24] The PRE effect arises from dipole-dipole interactions between unpaired electrons of the paramagnetic agent and the nucleus of interest, and thus, spin relaxation contributions show r^{-6} dependence of distance between the paramagnetic center and nucleus. Accordingly, when the paramagnetic agent was conjugated to a portion of the central domain, we were able to collect information on the distance between the central β-sheet and the N- and C-terminal domains.

Materials and Methods

Sample preparation

A soluble form of OspA,[25] consisting of residues 18–273 whereby the membrane-associated N-terminal region (residues 1–6) was truncated and the cysteine at residue 84 was substituted by serine, was used as the pseudo wild-type (C84S; WT*) form of OspA (BMRB Entry 4076). Uniformly ^{15}N-labeled OspA WT* and three variants (D118C, E128C, and A140C) were produced by conventional *E. coli* expression in M9 medium with ^{15}NH$_4$Cl as the sole nitrogen source. The protein was purified by a Ni^{2+}-affinity column (Bio-Rad Laboratories, Hercules, CA) and a HiLoad 26/60 Superdex 75 prep grade column in the AKTA explore 10S (GE Healthcare Life Sciences, Pittsburgh, PA). To cleave the His-tag, the protease thrombin (GE Healthcare Life Sciences) was used. The cysteine variants were dissolved in 20 mM Tris-HCl buffer at pH 7.0, mixed with *S* -([1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl]methyl) methanesulfonothioate (MTSL) (Toronto Research Chemicals, Ontario, Canada), and incubated for approximately 12 h in 20 mM Tris-HCl buffer at 277 K. The protein solution was filtered and concentrated using a Microcon (Merck Millipore, Billerica, MA, USA). The final protein solution was adjusted to a concentration of 0.3 mM for PRE experiments in 10 mM phosphate buffer (pH 5.9) containing 50 mM NaCl and 10%^{2}H$_2$O. The protein solution for temperature experiments contained 0.4 M guanidium chloride (GdmCl) in phosphate buffer. The spin-labeled MTSL was reduced with a 2-fold excess of ascorbic acid relative to the protein concentration.

2.2 NMR experiments

High-pressure NMR experiments were performed on an AVANCE3-950 (^{1}H, 950.33 MHz) spectrometer (Bruker, Billerica, MA, USA) with a cryo-probe system using a pressure-resistant NMR cell (Daedalus Innovations, Aston, PA, USA). For temperature-variable experiments, the slot NMR tube for high-concentration electrolyte samples (SHIGEMI, Hachioji, Japan) was used. To measure PRE, ^{1}H/^{15}N heteronuclear single-quantum correlation (HSQC) spectra of the MTSL-bound OspA and its variants were collected for the oxidized and reduced forms of MTSL at multiple pressures and temperatures. A 200 complex × 4096 real data matrix was acquired with spectral widths of 36.0 (^{15}N,F_1) and 12.98 ppm (^{1}H,F_2). NMR data were processed using Topspin version 3.2 (Bruker, Billerica, MA, USA), NMRPipe,[26] and NMRViewJ.[27] ^{1}H che-

mical shifts were referenced to the methyl signals of 2,2-dimethyl-2-silapentane-5-sulfonate (DSS), and ^{15}N chemical shifts were indirectly referenced to DSS (0 ppm for ^{1}H).

2.3 Thermodynamic analysis

The pressure-induced conformational transition from the native state (N) to the intermediate state (I), was determined by examining changes in NMR peak intensities, i.e. peak-area of L109 and V199 methyl groups to estimate an equilibrium constant K (= $[I]/[N]$). Assuming two-state exchange between N and I, the Gibbs free energy difference at any given pressure ΔG^{p} was obtained according to Equation 1. The Gibbs free energy difference ΔG^{0} and the partial molar volume difference ΔV^{0} between the two conformers at 0.1 MPa was estimated by assuming zero compressibility difference between the states as follows:

$$G^{p} = -RTln\ K = G^{0} + V^{0}\left(p - p^{0}\right)\ (1)$$

where R is the gas constant, T is the absolute temperature, p and p^{0} are high pressure and 0.1 MPa, respectively.

2.4 Analysis of PRE

The ratio of peak height of a paramagnetic sample to that of a diamagnetic sample, that is, I_{para}/I_{dia}, can be expressed as in Equation 2:[22]

$$\frac{I_{para}}{I_{dia}} = \frac{R_{2int}exp(-R_{2PRE}t)}{R_{2int}+R_{2PRE}}(2)$$

where R_{2int} and R_{2PRE} represent the intrinsic transverse relaxation rate and PRE effects on the transverse relaxation rate for each amide proton, respectively; and t is the total INEPT evolution time of the HSQC (11.5 ms). The R_{2int} for each amide proton was calculated from the half-height line-width of peaks of the diamagnetic sample. Using Equation 2, the R_{2PRE} was estimated by assigning the t and I_{para}/I_{dia} values. The R_{2PRE} was then converted into distance using Equation 3:

$$r^{6} = \frac{K}{R_{2PRE}}\left(4\tau_c + \frac{3\tau_c}{1+\varpi_H^2\tau_c^2}\right)(3)$$

where r is the distance between the unpaired electron and nucleus, ω_H is the Larmor frequency of protons, K is 1.23×10^{-32} cm^{6}s^{-2},[28] and τ_c is the correlation time given as $1/\tau_c = 1/\tau_r + 1/\tau_e$ (τ_r is the rotational correlation time of the electron–nucleus vector and τ_e is the electron spin relaxation time). Since τ_e (> 10^{-7} s) of the unpaired electron is much longer than τ_r, τ_c is approximately equal to τ_r. The τ_c value is approximately 12 ns for the folded conformation of OspA at 318 K and 0.1 MPa.[29] The half-height line-widths ($\Delta\nu$(hertz) = R_{2int}/π) were approximately 20 Hz for amide proton signals in the protein. The distances from the I_{para}/I_{dia} were estimated with a τ_r of 12 ns and R_{2int} (20 Hz). Back-calculation of I_{para}/I_{dia} for structural models of OspA was performed with τ_r = 12 ns and R_{2int} = 20 Hz.

Alternatively, estimation of PRE for the unfolded polypeptide chain is not as straightforward due to the fluctuations in both the orientation of the spins and its distance.[30] Therefore, the distance r was character-ized according to the probability distribution. Here, to simulate PREs of the OspA temperature-stabilized intermediate, we used the simplest restraining model, in which r is between the distance of closest approach for the two spheres (i.e. unpaired electron and nucleus). The r^{6} in Eq.3 can be substituted with $d_0^3L^3$, assuming that the time scale of the distance fluctuation is slow on the time scale of τ_{rot}.[30] d_0 and L are the distance of the minimum and the maximum separation, respectively. Further, the probability distribution of r was obtained from a trajectory of molecular dynamics simulation for heat unfolding of the protein (see *Molecular dynamics simulation*), and d_0 and L were estimated from the distance distribution. The τ_c of 12 ns, which is for native state OspA, was used for the intermediate conformations of the protein. More realistic restraining models were discussed in the literature.[30]

2.5 Molecular dynamics simulation

Molecular dynamics (MD) simulation of 1 μs was performed for OspA using the Amber18 simulator[31] with the AMBER ff14SBonlysc force field[31] and the generalized born (GB) solvent model.[32] The crystal structure

3

of the protein (PDB:1OSP) was used for the initial conformation. The simulation was conducted with the *NVT* ensemble, and temperature (350 K) was controlled using a Langevin thermostat with the solvent viscosity of water to 1.0 ps⁻¹. Noncovalent interactions, including electrostatic, van der Waals, and GB, were used without cutoffs. Covalent bonds, including hydrogen atoms, were constrained using the SHAKE method and the integration time step was 2 fs. Snapshots were recorded every 0.1 ns.

Results and Discussion

3.1 Pressure- and temperature-induced changes in ^{1}H-^{15}N HSQC spectra

^{1}H/^{15}N HSQC spectra of the three variants were measured with the oxidized and reduced forms of MTSL at 0.1, 200, and 250 MPa for 313 K, using an AVANCE3 950 MHz spectrometer. ^{1}H/^{15}N HSQC cross-peaks in the variant proteins were assigned to individual amino-acid residues with reference to the assignments of WT* OspA (BMRB code: 4076).[9] To evaluate the effects of residue substitution and the MTSL-tag, the ^{1}H and ^{15}N chemical shifts of individual cross-peaks in the HSQC spectra were compared between the WT* and MTSL-tagged cysteine variants (D118C, E128C, and A140C) at 313 K and 0.1 MPa. Chemical shift changes (i.e. $((\Delta\delta\,H)^2 + (\Delta\delta N/5)^2)^{0.5}$) are shown in Figure S1. Since chemical shift changes larger than 0.2 ppm were only observed at residues spatially proximal to the substituted residue, they were likely the result of local changes in magnetic shielding, and thus, the effects of residue substitution, and the MTSL-tag on the structure and conformational dynamics of the protein are expectedly limited.

Pressure effects on the ^{1}H/^{15}N-HSQC spectra of the E128C variant reduced form are shown in Figure 2. The HSQC cross-peaks corresponding to the folded conformation were well-dispersed, and the chemical shifts of the backbone amide proton and nitrogen were similar to those of WT* (Figure S1B); thus, the folded conformation of the variant was similar to that of WT*. Pressure-induced changes in the intensities (i.e., volumes) of the cross-peaks present in the 0.1 MPa spectrum (the original cross-peaks) are plotted in Figure S2. As pressure was increased to 150 MPa, the intensities of original cross-peaks showed minimal changes. However, several peaks corresponding to amide groups in the central and C-terminal regions preferentially and gradually decreased when pressure exceeded 150 MPa, after which relative intensities reached values of approximately 0.2. Additionally, many new peaks appeared in the central portion of the spectrum, where disordered polypeptide chains are typically observed (Figure 2). These results indicate that the locally disordered, intermediate conformation was partially stabilized under high pressure and occurred within 80% of the total protein.

Moreover, peak intensities (i.e., volumes) relative to those at atmospheric pressure are plotted along with the residue number at 250 MPa (Figure 3A), indicating that pressure-induced conformational changes occur at residues β8 to the C-terminus. A few peak also showed mountain-like intensity profiles (initial gain followed by a decrease in intensity with pressure) similar to those of WT* (Figure S2). As discussed previously[14], conformational dynamics on the microsecond-to-millisecond time scale may correlate with broadening and missing of the cross-peaks, even at 0.1 MPa.

When the temperature was increased from 303 to 318 K at 0.1 MPa, preferential decreases in cross-peaks were observed for residues between β8 and the C-terminus (Figure 3B), and relative intensities reached values of approximately 0.2 (Figure S3). Besides, many new peaks appeared in the central aspect of the spectrum (Figure S4), as observed in the pressure experiments.

Similar pressure- and temperature-induced decreases in peak intensities were observed in the ^{1}H NMR spectra for L109 and V199 methyl protons (Figure S5), which are located within the central β-sheet and C-terminal domain of the variant (data not shown). The transitions of the two side-chain signals, occurring in the same pressure and temperature ranges as those of backbone amides, serve as direct evidence for the cooperative transition of the central β-sheet and the C-terminal domain into the intermediate state. Based on pressure-induced changes in peak intensities of the methyl groups, the Gibbs free energy difference (ΔG^0) and the partial molar volume difference (ΔV^0) between the two conformers at 0.1 MPa and 313 K were estimated to be 17 ± 3 kJ/mol and $-80 +- 12$ mL/mol, respectively, which are slightly smaller than those of WT*.[14] As discussed previously, the large volume decrease observed by this transition may be the result of

4

collapse and hydration of the large hydrophobic cavity in the C-terminal domain (Fig. 1).[14,35]

An interesting observation is that cross-peak intensities of several amide groups in the N-terminal domain were relatively increased when the central and C-terminal domains were denatured under higher pressure (above 150 MPa) or temperature (higher than 313 K) (Figures S2 and S3). Although we did not delve deeper into the issue, a rotational motion of the folded N-terminal domain may have been enhanced, thereby producing more intense signals when the central and C-terminal domains were denatured.

3.2 Pressure- and temperature-stabilized intermediates

To further characterize the pressure- and temperature-stabilized intermediates, PRE effects on the ^{1}H-transverse relaxation rate constant, R_2 were measured for the variant. Rather than directly measuring the relaxation rates from a series of two-dimensional NMR experiments, we obtained the ratios of peak heights for a paramagnetic (oxidized) sample to those of a diamagnetic (reduced) sample, i.e., I_{para}/I_{dia} (Eq. 2).[22] Figure 4A shows the ratio for each amide group in the E128C variant along with the residue number (i.e., 25–110) at 0.1, and 250 MPa at 313 K. The I_{para}/I_{dia} of residues from 25 to 275 are shown in Figure S6. When the ratio was 1, no PRE effects were observed. In contrast, a lower ratio indicated greater PRE effects. Assuming an intrinsic line width at half-height ($\Delta\nu$(hertz) = R_{2int}/π) of 20 Hz and correlation time of 12 ns for the folded OspA, distances between the paramagnetic probe and nucleus of interest were calculated by a previously described method (Figure S7, Eq. 3).[22] Note that the multiplicity of MTSL positions was not taken into account in the calculation. At 0.1 MPa, ratios of less than 0.8, corresponding to a distance of 15~20 Å from the paramagnetic probe, were observed for residues 82, 83, 101, and 102, while values < 0.3 corresponded to a distance of 12~15 Å for residues 103–110. These results indicate that residues 103–110 are closer to residue 128 than they are to residues 82, 83, 101, and 102, consistent with the structure resolved by X-ray crystallography at atmospheric pressure.[6] At 250 MPa, I_{para}/I_{dia} showed both increased and decreased values at several sites. For example, the ratios for residues 82, 83, and 102 were increased by approximately 0.2, indicating that they drifted away from residue 128. In contrast, a remarkable decrease in the ratio for residue 59 indicated that it shifted toward residue 128. Figure 5A shows the locations of residues 59, 62, 82, 83, and 128, while the pressure dependence details for the ratios are shown in Figure 5B.

Figure 4B shows the intensity ratio, I_{para}/I_{dia}, for each amide group of the E128C variant for each residue at 303 and 318 K at 0.1 MPa. As the temperature increased, the ratios of residues 80, 82, and 83 increased by approximately 0.1–0.2, indicating they moved away from residue 128. Alternatively, slight decreases in the ratios for residues 33, 36, 37, 59, 62, and 105 indicated that these residues shifted closer to residue 128. Figure 5C shows the gradual increases and decreases in I_{para}/I_{dia} of residues 59, 62, 82, and 83 as temperature increases.

Different behaviors were observed at residues 36, 37, and 106 between increasing pressure and increasing temperature, as shown in Figures 6 and S8. As pressure increased, split cross-peaks were observed (Figure 6A). The second (new) cross-peak increased at the expense of the first (original) cross-peak with increasing pressure. In contrast, only a single peak was observed as the temperature was increased from 303 to 318 K (Figure 6B). These observations indicate that these amide groups have at least two folded conformations. Assuming a two-state exchange process, split peaks can be observed if the exchange rate, namely k_{ex}, is much smaller than the chemical shift difference ($\Delta\omega$) between two states; however, a single averaged peak is observed if k_{ex} is much larger than $\Delta\omega$. Figure 6C shows the I_{para}/I_{dia} values of original and new peaks for residues 36 and 37, and 106 at 200 MPa, where both original and new peaks were observed in the HSQC spectrum. For residues 36 and 37, new peaks showed 0.2–0.4-fold smaller values than the original peaks, indicating shorter distances in the intermediate than in the native conformations (Figure S6). Conversely, the new peak of residue 106 demonstrated a 0.6-fold larger value than the original peak, indicating much longer distance in the intermediate than in the native conformation. In the case of the averaged peak, the observed PRE effect is the weighted population average of the PRE effects for the two conformations, i.e., native and intermediate conformations. Since the average I_{para}/I_{dia} values of residues 36, 37, and 106 at 318 K and 0.1 MPa are similar to those of their new peaks (200 MPa and 313 K) (Figure 6D), the second conformation might be dominantly populated at 318 K.

Pressure and temperature effects on I_{para}/I_{dia} for each amide group were also investigated for the variants D118C and A140C, as shown in Figures S9 and S10, respectively. As pressure increased, the I_{para}/I_{dia} values of residues 59–61 of the D118C variant increased by approximately 0.1–0.2 (Figure S9A), indicating that D118 moved away from the β3–β4 turn in the pressure-stabilized intermediate. As the temperature increased, the values of residues 57–58 and 93–94 of the variant increased by approximately 0.2–0.3 (Figure S9B), indicating that residue 118 moved away from the β3–β4 and β6–β7 turns in the temperature-stabilized intermediate. Similarly, as pressure or temperature increased, the I_{para}/I_{dia} values of residues 94–97 of the A140C variant increased by 0.2–0.3, indicating that residue 140 moved away from the β7 in both of the intermediates (Figure S10A and B). I_{para}/I_{dia} values of residues 36, 37, and 106 for the A140C variant are summarized in Figure S11. A pair of cross-peaks was only observed for residue 36. Unlike E128C, the intensity ratio of the new residue 36 peak is similar to that of the original one (Fig. S11A). The ratios at higher temperature were significantly higher for residues 36 and 37, and slightly lower for residue 106 (Fig. S11B), showing opposite tendencies with the data from E128C. In addition, for the variants D118C and A140C, residues 33–39, including the β1–β2 turn, maintained lower I_{para}/I_{dia} values (i.e., 0–0.6) in both intermediates.

These results indicate that the pressure- and temperature-stabilized intermediates shared a common structural feature in which residue 128 moved away from the β5–β6 turn (consisting of residues 82 and 83) while becoming more accessible to the β3–β4 turn region, including residues 59 and 62. Moreover, residues 118 and 140 moved away from the N-terminal domain, yet remained accessible to the β1–β2 turn. As residues 118, 128, and 140 are in the turn region between β8–β9, β9–β10, and β10–β11, respectively, these PRE results suggest that the central β-sheet, i.e., β8–β10, is partially disordered, and thus is located a longer average distance from the N-terminal domain in the intermediates than it is in the native conformation (Figure 1). Alternatively, the β-sheet region may retain weak specific contact with the β1–β2 turn, owing to a twist of the central β-sheet and increased flexibility of the polypeptide chain. Such specific long-range contact has also been observed in intrinsically disordered proteins such as α-synuclein.[23,36]

New peaks appeared in the central aspect of the spectrum, an area in which disordered polypeptide chains are typically observed, as pressure or temperature was increased. The original signals corresponding to the N-terminal domain and the new NMR signals showed gradual decreases in intensity after incubating the protein for a few days at 250 MPa, likely owing to protein aggregation. Therefore, the assignment of new peaks using the conventional triple-resonance NMR experiments has not been completed. However, it is useful to analyze the I_{para}/I_{dia} ratios for new peaks that likely correspond to amide groups in the central and C-terminal domains. Figure 7A shows the frequency of I_{para}/I_{dia} values for all original cross-peaks that originated from the native protein conformation, as the paramagnetic probe was attached at residue 140 in the central domain. The ratios were widely distributed between 0 and 1; however, were primarily above 0.8, indicating that a folded conformation remained near the paramagnetic probe, and many amide groups were not accessible to the probe. Figures 7B and C show the frequencies of I_{para}/I_{dia} values of new peaks in the pressure-stabilized and temperature-stabilized intermediates, respectively, of the A140C variant. In the temperature-stabilized intermediate (Figure 7C), the frequency *vs.* I_{para}/I_{dia} plot had a Gaussian shape, with a center value of approximately 0.7; frequencies of less than 0.2 and more than 0.9 were not observed, indicating an entirely different shape from the native conformation. Since a Gaussian shape is typically obtained for random events, the PRE results may have resulted from an ensemble of disordered conformations of the polypeptide chain. The PRE provides $1/r^6$ ensemble-averaged distance information, and thus different conformations may show different PRE effects. Even amide groups more than 25 Å, apart from the paramagnetic probe in the native conformation, could transiently access the probe in the intermediates due to the increased flexibility of the polypeptide chain, while very few amide groups showed consistently strong PRE effects. Although the I_{para}/I_{dia}-frequency profile of the pressure-intermediate was similar to that of the temperature, the Gaussian shape was not as apparent, and the strong and weak I_{para}/I_{dia} ratios, i.e., > 0.9 and < 0.2, were still observed (Figure 7B). Since an intermediate population is retained in both conditions (Figure S2 and Figure S3), the differences in the I_{para}/I_{dia}-frequency profiles indicate that the pressure-intermediate has a more compact conformation. Similar frequency profiles were

observed for the D118C and E128C variants, as shown in Figures S12 and S13, respectively.

3.3 Molecular dynamics simulation

We sought to characterize the locally disordered intermediate conformation of OspA from PRE data. However, a conversion of R_{2PRE} into distance is not straightforward in this case since the protein exists as a mixture of different conformational states, including ~80% locally disordered, ~20% native, and minimal amounts of completely unfolded protein. Even for the completely unfolded state, it is not easy to reconstruct the probability distribution of the distance between the paramagnetic probe and each amide proton, as multiple sources of stochastic dynamics in the system must be considered.[30] Here, to produce potential structural models of the intermediate conformations of WT* OspA, we performed a 1 μs MD simulation with the NVT ensemble (350 K). The MD simulation demonstrated that temperature-induced disordering of the β-sheet began at the C-terminal region, followed by the central region.

Distance distributions were calculated between the paramagnetic probe and each amide proton for four different time-regions in the MD trajectory, corresponding to the different levels of disorder in β9–β11. Figure S14 presents the distance distribution between the C_β of residue 118/128/140 (a substitute for the paramagnetic center) and the amide proton of residue 37 for the folded, partially disordered, and completely disordered ensembles. A specific simulation snapshot was obtained in the folded and partially disordered ensembles, in which β11, β10, and β9 are disordered (Figure 8A). The distance for the closest approach d_0 and that of the maximum separation L for the ensembles were obtained from the distance distribution. In the partially disordered and completely disordered ensembles, the distance distribution became wider than that of the folded ensemble. I_{para}/I_{dia} profiles of the folded and the partially disordered ensembles, as predicted by the MD simulation, are shown in Figure 8B (see Methods). Increasing the ratios for residues 33, 93, and 94 of the D118C variant can be explained by the partially disordered ensemble 3 (PDE3, Figure S8). Meanwhile, the increased ratios of residues 80–83, 103, and 106 of the E128C may have resulted from partially disordered ensembles 1 and 2 (PDE1 and PDE2). Further, PDE1 and PDE2 can explain the incremental ratios for residues 36, 37, 94, and 95 of the A140C variant. However, no ensemble can explain the decreases observed in the ratios of residues 33-37, 59, and 62. In short, any single specific ensemble cannot explain the entire profile of I_{para}/I_{dia}. Moreover, there is no evidence to suggest that the new peaks corresponding to the residues in the central β-sheet appeared in the central portion of the spectrum. Hence, we do not deny broadening and missing of HSQC cross-peaks owing to conformational dynamics; however, we speculate that the central β-sheet does not become unfolded, but rather partially disordered, and contains heterogeneous conformations in the intermediates.

Thus, the intermediate state of the protein is likely a mixture of conformations, with different levels of partial disorder in the central β-sheet. This idea is consistent with the conformational heterogeneity of the equilibrium intermediate mapped by native state hydrogen exchange NMR and scanning mutagenesis.[10,11] To satisfy the entire profile more quantitatively, PRE should be measured under conditions in which the intermediate state is more dominantly stabilized, and larger numbers of snapshots should be generated via MD simulation of the partially disordered protein.

3.4 Functional significance of the intermediate conformations

The current PRE results provide further evidence of disorder in the central and C-terminal regions of the OspA intermediates and similarities between the effects of pressure and temperature on the protein. In addition, a possible ensemble of partially disordered conformations was suggested from the MD simulation. These data indicate that the intermediates more readily expose the tick receptor-binding sites (i.e., residues 236–237 and 242–244)[8] than does the basic folded conformation. Hence, the transition into the intermediates *in vivo* may accelerate the binding of OspA to the tick receptor. Although the folded OspA protein and its C-terminal fragment have been included as vaccine candidates to prevent *Borrelia* transmission,[17] the intermediates, in which the central and C-terminal domain are disordered, are unlikely to be recognized by antibodies targeted at the folded C-terminal portion, thereby reducing the effectiveness of the vaccine. The disordering of the C-terminal domain is also considered to be important when the protein is translocated

7

through the outer membrane in the C-terminal-to-N-terminal direction. However, destabilization of the C-terminal domain by site-specific mutation induced protein degradation.[16] Results from the present and previous studies indicate that a delicate interplay between localized conformational dynamics and thermodynamic stability may be required for OspA function.

Conclusion

Although structural characterization of protein intermediates remains challenging, the partial knowledge that we gathered here not only expands the current understanding on the molecular mechanism of protein function, but also facilitates novel drug design. The intermediate state of OspA consists of a mixture of conformations with different levels of partial disorder in the C-terminal region. The C-terminal region of the intermediate is partially disordered; it however retains weak specific contact with the N-terminal region, owing to a twist of the central β-sheet and increased flexibility of the polypeptide chain. These intermediates may escape antibody recognition. Information regarding localized conformational dynamics and thermodynamic stability might be useful for the rational design of effective OspA vaccines. Moreover, since PRE observation under high pressure is only limited in ubiquitin, the present is a confirmation of PRE under high pressure.[24] A combination of PRE-assisted high-pressure NMR and MD simulation is, in general, useful for characterizing the dynamic structures of functionally essential protein intermediates.

ASSOCIATED CONTENT

Supporting Information Figure S1-S14

Accession codes

OspA (*Borrelia burgdorferi*), UniProtKB-Q45040

Author contributions

The manuscript was written by R. K. NMR measurements and analyses were performed by T. W., S. K., and R. K. MD simulations were performed by T. K. All authors reviewed and approved the final manuscript.

Notes

The authors declare no competing interests.

REFERENCES

1. Boehr DD, McElheny D, Dyson HJ, and Wright PE. The dynamic energy landscape of dihydrofolate reductase catalysis. Science. 2006;313:1638-1642.
2. Boehr DD, Dyson HJ, and Wright PE. An NMR perspective on enzyme dynamics. Chem Rev. 2006;106:3055-3079.
3. Booth DR, Sunde M, Bellotti V, Robinson CV, Hutchinson WL, Fraser PE, Hawkins PN, Dobson CM, Radford SE, Blake CCF, and Pepys MB. Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. Nature. 1997;385:787-793.
4. Chiti F and Dobson CM. Amyloid formation by globular proteins under native conditions. Nat Chem Biol. 2009;5:15-22.
5. Li H, Dunn JJ, Luft BJ, and Lawson CL. Crystal structure of Lyme disease antigen outer surface protein A complexed with an Fab. Proc Natl Acad Sci U S A. 1997;94:3584-3589.
6. Makabe K, Tereshko V, Gawlak G, Yan S, and Koide S. Atomic-resolution crystal structure of Borrelia burgdorferi outer surface protein A via surface engineering. Protein Sci. 2006;15:1907-1914.
7. Pal U, Li X, Wang T, Montgomery RR, Ramamoorthi N, deSilva AM, BaoFK, Yang XF, Pypaert M, Pradhan D, Kantor FS, Telford S, Anderson JF, and Fikrig E. TROSPA. an Ixodues scapularis receptor for Borrelia burgdorferi. Cell. 2004;119:457-468.
8. Pal U, de Silva AM, Montgomery RR, Fish D, Anguita J, Anderson JF, Lobet Y and Fikrig E. Attachment of Borrelia burgdorferi within Ixodes scapularis mediated by outer surface protein A. J Clin Invest. 2000;106:561-569.

9. Pham TN, Koide A, and Koide SA. stable single-layer beta-sheet without a hydrophobic core. Nat Strct Biol. 1998;5:115-119.

10. Yan S, Kennedy SD and Koide S. Thermodynamic and kinetic exploration of the energy landscape of Borrelia burgdorferi OspA by native state hydrogen exchange. J Mol Biol. 2002;323:363-375.

11. Yan S, Gawlak G, Smith J, Silver L, Koide A and Koide S. Conformational heterogeneity of an equilibrium folding intermediate quantified and mapped by scanning mutagenesis. J Mol Biol. 2004;338:811-825.

12. Koide S, Bu ZM, Risal D, Pham TN, Nakagawa T, Tamura A and Engelman DM. Multistep denaturation of Borrelia burgdorferi OspA, a protein containing a single-layer beta-sheet. Biochemistry. 1999;38:4757-4767.

13. Nakagawa T, Shimizu H, Link K, Koide A, Koide S and Tamura A. Calorimetric dissection of thermal unfolding of OspA, a predominantly beta-sheet protein containing a single-layer beta-sheet. J Mol Biol. 2002;323:751-762.

14. Kitahara R, Simorellis AK, Hata K, Maeno A, Yokoyama S, Koide S and Akasaka K. A delicate interplay of structure, dynamics, and thermodynamics for function: a high pressure NMR study of outer surface protein A. Biophys J. 2012;102:916-926.

15. Makabe K, Nakamura T, Dhar D, Ikura T, Koide S and Kuwajima K. An Overlapping Region between the Two Terminal Folding Units of the Outer Surface Protein A (OspA) Controls Its Folding Behavior. J Mol Biol. 2018;430:1799-1813.

16. Schulze RJ, Chen S, Kumru OS and Zückert WR. Translocation of Borrelia burgdorferi surface lipoprotein OspA through the outer membrane requires an unfolded conformation and can initiate at the C-terminus. Mol Microbiol. 2010;76:1266-1278.

17. Koide S, Yang X, Huang X, Dunn JJ and Luft BJ. Structure-based design of a second-generation Lyme disease vaccine based on a C-terminal fragment of Borrelia burgdorferi OspA. J Mol Biol. 2005;350:290-299.

18. Livey I, O'Rourke M, Traweger A, Savidis-Dacho H, Crowe BA, Barrett PN, Yang X, Dunn JJ, and Luft BJ. A new approach to a Lyme disease vaccine. Clinical Infectious Diseases. 2011;52:S266-S270.

19. Nigrovic LE and Thompson KM. The Lyme vaccine: a cautionary tale. Epidemiol. Infect. 2007;135:1-8.

20. Clore GM and Iwahara J. Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. Chem Rev. 2009;109:4108-4139.

21. Liu Z, Gong Z, Jiang WX, Yang J, ZhuWK, Guo DC, Zhang WP, Liu ML and Tang C. Lys63-linked ubiquitin chain adopts multiple conformational states for specific target recognition. eLife. 2015;4:e05767.

22. Battiste JL and Wagner G. Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. Biochemistry. 2000;39:5355-5365.

23. Salmon L, Nodet G, Ozenne V, Yin GW, Jensen MR, Zweckstetter M and Blackledge M. NMR Characterization of Long-Range Order in Intrinsically Disordered Proteins. J Am Chem Soc. 2010;132:8407-8418.

24. Wakamoto T, Ikeya T, Kitazawa S, Baxter NJ, Williamson MP and Kitahara R. Paramagnetic relaxation enhancement-assisted structural characterization of a partially disordered conformation of ubiquitin. Protein Sci. 2019;28:1993-2003.

25. Dunn JJ., Lade BN and Barbour AG. Outer surface protein a (OspA) from the lyme disease spirochete, borrelia burgdorferi: High level expression and purification of a soluble recombinant form of OspA. Protein Exp Purif. 1990;1:159-168.

26. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J. and Bax A. Nmrpipe - a Multidimensional Spectral Processing System Based on Unix Pipes. J Biol NMR. 1995;6:277-293.

27. Johnson BA and Blevins RA. Nmr View - a Computer-Program for the Visualization and Analysis of Nmr Data. J Biol NMR. 1994;4:603-614.

28. Kosen PA. Spin labeling of proteins. Methods Enzymol. 1989;177:86-121.

29. Pawley NH, Koide S and Nicholson L. K. Backbone dynamics and thermodynamics of Borrelia outer

surface protein A. J Mol Biol. 2002;324:991-1002.

30. Xue Y, Podkorytov IS, Rao DK, Benjamin N, Sun H and Skrynnikov NR. Paramagnetic relaxation enhancements in unfolded proteins: Theory and application to drkN SH3 domain. Protein Sci. 2009;18:1401-1424.

31. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. J Chem Theory Comput. 2015;11:3696-3713.

32. Mongan J, Simmerling CA. McCammon JA, Case D, Onufriev A, Generalized born with a simple, robust molecular volume correction. J Chem Theory Comput. 2007;3:156-169.

33. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC and Ferrin TE. UCSF chimera - A visualization system for exploratory research and analysis. J Comput Chem. 2004;25:1605-1612.

34. Tian W, Chen C, Lei X., Zhao J. and Liang J. CASTp 3.0: computed atlas of surface topography of proteins. Nucleic Acids Res. 2018;46:W363-W367.

35. Kawamura T, Wakamoto T, Kitazawa S, Sakuraba S, Kameda T and Kitahara R, Analysis of O_2-binding sites in proteins using gas-pressure NMR spectroscopy: Outer surface protein A. Biophys J. 2017;112:1820-1828.

36. Bernado P, Bertoncini CW, Griesinger C, Zweckstetter M and Blackledge M. Defining long-range order and local disorder in native alpha-synuclein using residual dipolar couplings. J Am Chem Soc. 2005;127:17968-17969.

Figure legends

Figure 1. Three-dimensional structure of OspA. Residues substituted to cysteine and with the attached spin-labeled MTSL are depicted as spheres. OspA contains the N-terminal (β1–β7) and C-terminal (β11 to the C-terminus) globular domains, which are connected by a single-layer β-sheet (β8–β10). The image was prepared using Chimera.[33] Internal cavities were detected using CASTp server with 1.4 Å radius probe.[34]

Figure 2. ^{1}H/^{15}N-HSQC spectra of the reduced OspA E128C variant form as pressure was gradually increased from 0.1 to 250 MPa at 313 K (black, 0.1 MPa; blue, 50 MPa; cyan, 100 MPa; green, 150 MPa; purple, 200 MPa; red, 250 MPa). (Inset) Change in chemical shifts of V199 (not visible in the main plot).

Figure 3. Resonance intensities (i.e. peak volumes) of the OspA E128C variant at different pressures and temperature conditions. (A) Resonance intensities at 250 MPa relative to those at 0.1 MPa (313 K) as a function of residue number. (B) Resonance intensities at 318 K relative to those at 303 K (0.1 MPa) as a function of residue number.

Figure 4. Intensity ratios, I_{para}/I_{dia}, for each amide group of the E128C variant for each residue. (A) Pressure dependence of the I_{para}/I_{dia} value at 313 K. (B) Temperature dependence of the I_{para}/I_{dia} value at 0.1 MPa. Secondary structures are indicated at the top of the panels. Error bars for I_{para}/I_{dia} were estimated from the noise considering error propagation, i.e. $\left(\frac{I_{para}}{I_{dia}}\right) = \left(\frac{I_{para}}{I_{dia}}\right) \sqrt{\left(\frac{I_{para}}{I_{para}}\right)^2 + \left(\frac{I_{dia}}{I_{dia}}\right)^2}$.

Figure 5. Pressure and temperature dependence of peak height ratios of a paramagnetic sample (E128C) to those of a diamagnetic sample, i.e., I_{para}/I_{dia}. (A) Locations of residues 59, 62, 82, 83, and 128 in OspA. (B) Pressure-induced changes in the ratios of residues 59, 82, and 83 at 313 K. (C) Temperature-induced changes in the ratios of residues 59, 62, 82, and 83 at 0.1 MPa. Panel A was prepared using Chimera.[33]

Figure 6. Pressure- and temperature-dependent effects on selected HSQC cross-peaks. (A) Selected ^{1}H/^{15}N HSQC spectral regions for the reduced MTSL-E128C variant at multiple pressures (black, 0.1 MPa; blue, 50 MPa; cyan, 100 MPa; green, 150 MPa; purple, 200 MPa; red, 250 MPa). (B) Selected ^{1}H/^{15}N HSQC spectral regions of the protein at multiple temperatures (black, 303 K; blue, 308 K; green, 313 K; red, 318 K). Assignments for G36 and G106 peaks are depicted at the top of the 0.1 MPa-peak in panel A or 303 K-peak in panel B. The peaks of G36 and G106 are marked with asterisks. (C) I_{para}/I_{dia} values of original

and new peaks of residues 36, 37, and 106 at 200 MPa. (D)I_{para}/I_{dia} values of the averaged peaks of residues 36, 37, and 106 at 303 and 318 K.

Figure 7. Frequency of I_{para}/I_{dia} values calculated from HSQC cross-peaks for the A140C variant. (A) Frequency of I_{para}/I_{dia} of original cross-peaks at 0.1 MPa and 303 K, corresponding to the native state. (B) Frequency of I_{para}/I_{dia} of new cross-peaks at 250 MPa and 303 K, corresponding to the pressure-stabilized intermediate. (C) Frequency of I_{para}/I_{dia} of new cross-peaks at 0.1 MPa and 318 K, corresponding to the temperature-stabilized intermediate. Broken lines depict value 1.

Figure 8. I_{para}/I_{dia} profiles on the basis of MD simulation. (A) MD simulation snapshots of heat unfolding (snapshots after 0.2 ns, 62.7 ns, 73.1 ns, and 80.7 ns from top to bottom, respectively). (B) I_{para}/I_{dia} ratios for the conformational ensembles of different levels of disorder in β9-β11 (black, the folded ensemble (16 snapshots in the MD trajectory of 0.1 ns~1.6 ns); red, the partially disordered ensemble 1 (PDE1, 31 snapshots in the MD trajectory of 60 ns~63 ns); blue, the partially disordered ensemble 2 (PDE2, 121 snapshots in the MD trajectory of 63.1 ns~75.1 ns); green, the partially disordered ensemble 3 (PDE3, 65 snapshots in the MD trajectory of 75.2 ns~81.6 ns)). I_{para}/I_{dia} ratios are calculated from the distance distribution between the C_{β} of residue 118 (top)/128 (middle)/140 (bottom) (a substitute for the paramagnetic center) and the amide proton of each residue with τ_r of 12 ns (see Fig. S14). Panel A was prepared using Chimera.[33]
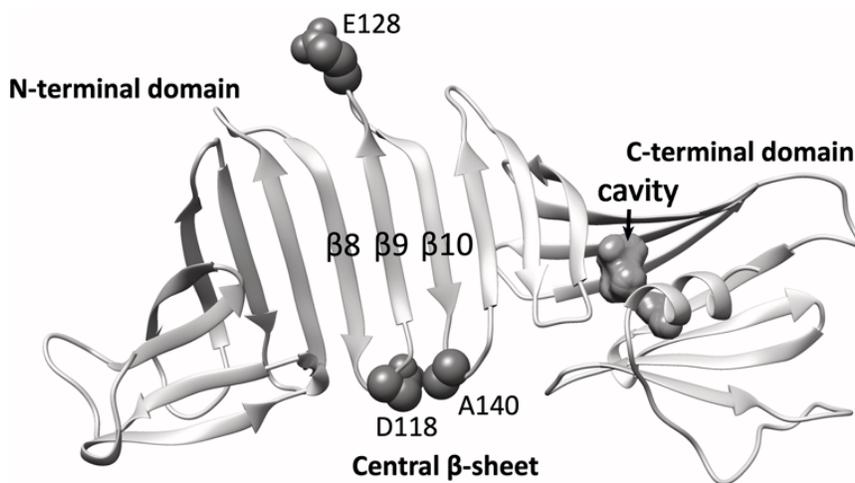
Figure 1

Figure 2

Figure 3

12

Figure 4

Figure 5

A
G106
G36

¹⁵N/ppm

110.0
110.5
111.0
111.5
112.0

¹H/ppm
9.0   8.8   8.6

B
G106
G36

¹⁵N/ppm

109.5
110.0
110.5
111.0
111.5
112.0

¹H/ppm
9.0   8.8   8.6

C

I(para)/I(dia)

1.2
1
0.8
0.6
0.4
0.2
0

■ Original (200 MPa)
■ New   (200 MPa)

G36   G36   E37   E37   G106   G106

D

I(para)/I(dia)

1.2
1
0.8
0.6
0.4
0.2
0

■ 303 K
■ 318 K

G36   G36   E37   E37   G106   G106

Figure 6

A

Frequency

35
30
25
20
15
10
5
0

0 0.2 0.4 0.6 0.8 1 1.2 1.4
I(para)/I(dia)

B

Frequency

35
30
25
20
15
10
5
0

0 0.2 0.4 0.6 0.8 1 1.2 1.4
I(para)/I(dia)

C

Frequency

35
30
25
20
15
10
5
0

0 0.2 0.4 0.6 0.8 1 1.2 1.4
I(para)/I(dia)

Figure 7

15

Figure 8