Title

Structure and Dynamics of a Compact State of a Multidomain Protein, the Mercuric Ion Reductase

Journal

Biophysical Journal, 107(2)

ISSN

0006-3495

Authors

Hong, Liang
Sharp, Melissa A
Poblete, Simón
et al.

Publication Date

2014-07-01

DOI

10.1016/j.bpj.2014.06.013

Peer reviewed

Article

Structure and Dynamics of a Compact State of a Multidomain Protein, the Mercuric Ion Reductase

Liang Hong,[1,2,10] Melissa A. Sharp,[3,4] Simón Poblete,[5] Ralf Biehl,[6] Michaela Zamponi,[7] Noemi Szekely,[7] Marie-Sousai Appavou,[7] Roland G. Winkler,[5] Rachel E. Nauss,[8] Alexander Johs,[9] Jerry M. Parks,[1] Zheng Yi,[1,2] Xiaolin Cheng,[1] Liyuan Liang,[9] Michael Ohl,[7,*] Susan M. Miller,[8,*] Dieter Richter,[6,*] Gerhard Gompper,[5,*] and Jeremy C. Smith[1,2,*]

[1]Center for Molecular Biophysics, Oak Ridge National Laboratory, Tennessee; [2]Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville, Tennessee; [3]European Spallation Source ESS AB, Lund, Sweden; [4]Jülich Center of Neutron Science, Outstation at the Spallation Neutron Source (SNS), Oak Ridge, Tennessee; [5]Institute of Complex Systems & Institute for Advanced Simulation, Forschungszentrum Jülich, Jülich, Germany; [6]Jülich Center of Neutron Science & Institute of Complex Systems, Forschungszentrum Jülich, Jülich, Germany; [7]Jülich Centre for Neutron Science JCNS, Forschungszentrum Jülich GmbH Outstation at MLZ, Garching, Germany; [8]Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, California; [9]Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee; and [10]Department of Physics and Institute of Natural Sciences, Shanghai Jiao Tong University, Shanghai, China

ABSTRACT The functional efficacy of colocalized, linked protein domains is dependent on linker flexibility and system compaction. However, the detailed characterization of these properties in aqueous solution presents an enduring challenge. Here, we employ a novel, to our knowledge, combination of complementary techniques, including small-angle neutron scattering, neutron spin-echo spectroscopy, and all-atom molecular dynamics and coarse-grained simulation, to identify and characterize in detail the structure and dynamics of a compact form of mercuric ion reductase (MerA), an enzyme central to bacterial mercury resistance. MerA possesses metallochaperone-like N-terminal domains (NmerA) tethered to its catalytic core domain by linkers. The NmerA domains are found to interact principally through electrostatic interactions with the core, leashed by the linkers so as to subdiffuse on the surface over an area close to the core C-terminal Hg(II)-binding cysteines. How this compact, dynamical arrangement may facilitate delivery of Hg(II) from NmerA to the core domain is discussed.

INTRODUCTION

Flexible linkers are commonly found in modular proteins. These linkers play an important role in colocalizing functional domains, thus greatly increasing the corresponding effective molarity (1). However, the functional efficacy will depend on the length and flexibility of the linker and other structural aspects (2,3). Moreover, the linker flexibility and system compaction may be dependent on external conditions or on binding of a ligand or substrate to one or more domains. The characterization of the functional structure and dynamics of linker-connected protein domains in aqueous solution presents considerable technical challenges, arising in part from a dearth of experimental techniques capable of direct measurement of interdomain motion and timescales concerned being out of the reach of most atomic-detail simulation approaches.

The previous issues are highly relevant to the function of the enzyme mercuric ion reductase, MerA, which is central to the mercury resistance pathway found in many aerobic bacteria. Heavy metal ions, such as mercury, are highly toxic to cells, due to their ability to bind to functional groups, especially cysteine thiols, within the cell (4). To mitigate these effects and provide resistance against these metal ions, some microorganisms have evolved detoxification pathways (4–6). An example in bacteria is the set of gene products encoded by the tightly regulated *mer* operon, which provides specific response mechanisms to degrade organomercurials and remove oxidized, inorganic mercury from the cell (4). Central to this process is MerA, which catalyzes the reduction of the highly toxic Hg(II) cation to the relatively inert Hg(0) (4).

Several variants of MerA possess metallochaperone-like N-terminal domains (NmerA), which are tethered to the homodimeric catalytic core by ~30 residue linkers (Fig. 1) (4,7). NmerA contains a pair of cysteines in a GMTCXXC sequence motif that is conserved in those soft metal ion trafficking proteins sharing the common $\beta\alpha\beta\beta\alpha\beta$ structural fold (8). MerA binding and processing of Hg(II) is schematically illustrated in Fig. 1. Hg(II) binds first to a pair of cysteines (C11 and C14) in NmerA, these being the most solvent-accessible cysteines in the protein (8,9). The ion is then transferred to another pair of cysteines (C561′ and C562′) located on the mobile C-terminal segment of the core. The mobile segment then moves the C-terminal Hg(II)-bound

Submitted November 11, 2013, and accepted for publication June 10, 2014.

*Correspondence: m.ohl@fz-juelich.de or smiller@cgl.ucsf.edu or d.richter@fz-juelich.de or g.gompper@fz-juelich.de or smithjc@ornl.gov

Liang Hong, Melissa A. Sharp, Simón Poblete, and Ralf Biehl contributed equally to this work.

Editor: Hummer Gerhard.

0006-3495/14/07/0393/8 $2.00

http://dx.doi.org/10.1016/j.bpj.2014.06.013

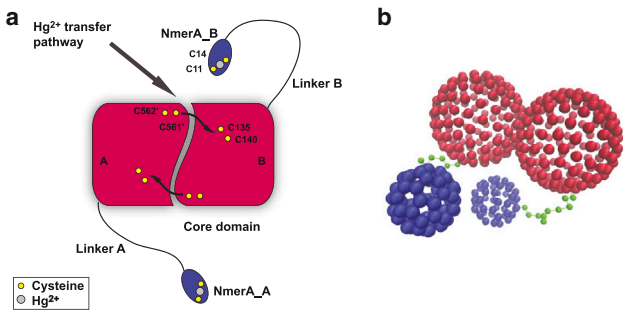

FIGURE 1 (*a*) Schematic illustration of the full-length MerA homodimer and the scheme for the transfer of Hg(II) from the solvent into the catalytic sites of the MerA core. Hg(II) is first bound by the NmerA cysteines (C11 and C14) and is then delivered to the C-terminal cysteines of the other monomer (C561′ and C562′). Finally, Hg(II) is transferred to a pair of cysteines in the catalytic site of the core (C135 and C140) to be reduced to Hg(0). (*b*) Simplified CG model for fl-MerA with a rigid double sphere core (*red*) and small spheres for the NmerA domains (*blue*), which are connected to the core by linkers (*green*). To see this figure in color, go online.

cysteines from the surface to the interior of the protein, from which the ions are further delivered to the buried, active-site cysteine pair (C135 and C140) where the reduction occurs (10,11). In vivo experiments showed that the presence of NmerA domains significantly enhances cell survival in the presence of Hg(II), especially under conditions of low cellular thiol concentration (9), in which the high-affinity Hg(II)-chelating NmerA domains are essential for rapid acquisition, localization, and transfer of Hg(II) to the core for reduction and detoxification (9).

Key to understanding the mechanism of how the NmerA domains and linkers bind and transfer Hg(II) to the catalytic core is the characterization of the conformations and dynamics of the full-length protein complex in solution. Although crystallography has provided static structures of the core domain (9), the flexibility of the linker and the dynamics of the NmerA domains relative to the core have not been elucidated. Two extreme scenarios are conceivable for the interdomain dynamics in the substrate-free MerA. In one case, the linker is completely flexible, and NmerA can dissociate readily from the core and move freely in the aqueous environment. In the other extreme, NmerA is noncovalently associated with the core at a position proximal to the C-terminal cysteine thiols, leashed by a relatively stiff linker. The former scenario significantly enlarges the volume over which NmerA can extend for binding Hg(II), but would be expected to prolong the process of delivering Hg(II) to the core, whereas the second extreme is just the opposite.

Here, we employ a combination of complementary techniques—molecular dynamics (MD) and coarse-grained (CG) simulations and small-angle neutron scattering (SANS), neutron spin-echo spectroscopy (NSE) and dynamic light scattering (DLS) experiments—to address these questions. Two systems are characterized. One is the catalytic core domain without linkers or NmerA domains, referred as the core, and the other is the full-length MerA protein including the core, linkers, and NmerA domains, denoted as fl-MerA. SANS is widely used to explore shapes of biological molecules in aqueous solution. Here, SANS is used to probe the global configurational distribution of the system, and NSE and DLS are applied to provide information on the dynamics. CG simulation enables examination of models with linkers of different stiffness (12,13), whereas MD probes timescales up to ~1 μs and furnishes a more detailed picture of microscopic mechanisms involved. The combination of these experimental and simulation techniques leads to a consistent picture in which MerA adopts a compact shape in solution with the NmerA domains in direct contact with the core while exploring the core surface close to the C-terminal Hg(II)-binding cysteines.

MATERIALS AND METHODS

Protein samples for neutron scattering experiments

Both the MerA core and fl-MerA from the Tn*21 mer* operon were prepared from *Escherichia coli* BL21(DE3) pLysS cells (see details in the Supporting Material). To avoid potential changes in structure due to oxidation of cysteine residues during the lengthy time frame required for NSE data collection, the six redox-sensitive cysteines involved in Hg(II) binding and transfers were mutated to alanines. These include C11, C14, C135, C140, C561, C562 in full-length and the latter four in the core. Immediately before the neutron measurements, the MerA samples were thawed at room temperature and exchanged into a deuterated buffer solution (50 mM NaH_2PO_4, 50 mM NaCl, D_2O), adjusted to pD 7.0 for full-length MerA and pD 7.2 for MerA-core.

SANS

SANS measurements were performed at instrument KWS-1 (fl-MerA) and KWS-2 (MerA core) at the FRM-II reactor in Garching (Germany). A q-range of ~0.006 – 0.33 $Å^{-1}$ was obtained by combining data collected at different wavelengths and sample-detector distances. The temperature used during the measurements was 10 °C, to be consistent with the NSE measurements. A series of four concentrations (fl-MerA: 2.7, 5.6, 10.8, and 53.9 mg/ml, core: 2.7, 5.0, 9.6, 46.9 mg/ml) was background corrected using the appropriate buffer solution. The background-corrected data were scaled by concentration and the form factor was extracted by extrapolating to zero concentration for individual q values, resulting in the form factor per unit mass for infinite dilution, $F(q)$. The structure factor, $S(q)$, describing correlations between protein molecules at the higher concentration was calculated by division of the background-corrected and concentration-scaled data by the form factor $F(q)$. For both fl-MerA and the core, SANS measurements were performed twice, immediately before and after the NSE measurements. Exactly the same samples were used to permit direct comparison of the results.

NSE spectroscopy

NSE is an inelastic scattering technique that measures the intermediate scattering function $I_{coh}(q,t)$ in the time domain with a high resolution in energy (14). NSE measurements were performed at J-NSE of the JCNS at the FRM-II reactor in Garching (Germany). Quartz cells of 30 × 30 mm^2

with a 4 mm path length were used. Three different wavelengths (6 Å, 8 Å, 12.8 Å) were used to cover a Fourier time from ~0.05 to 90 ns over a q-range from 0.03 to 0.19 $Å^{-1}$. Graphite powder was used to determine the instrument resolution function, and the buffer was measured to allow for background correction. Samples at two concentrations were measured: 10.8 and 53.9 mg/ml for fl-MerA and 9.6 and 46.9 mg/ml for the core. The experiments were conducted at 10 °C to preserve the samples because the data collection times for NSE are relatively long (several days).

DLS

DLS was applied to determine the long-range translational diffusion coefficient of the protein. DLS measurements on fl-MerA and the core were performed using a standard setup with an Ar+ laser (5140 Å) and an avalanche photo diode (APD, ALV, Germany) mounted on an ALV-125 compact goniometer together with an ALV5000E digital correlator (both ALV, Germany). All measurements were performed at 10°C in a temperature-controlled environment. As a second instrument, a Zetasizer ZS (Malvern, Germany) was used with a wavelength of 6320 Å at a scattering angle of 173°. An inverse Laplace transform algorithm known as CONTIN (15) was used for the analysis.

MD simulation

The protein sample characterized experimentally is Tn*21*, the structure of which is not available. However, an x-ray structure is available for the homodimeric catalytic core of Tn*501* MerA (9) and NMR solution structures are available for Tn*501* NmerA (8). The sequences of Tn*21* and Tn*501* MerA are ~90% identical (16). On the basis of the high degree of sequence similarity, we constructed a model of full-length Tn*501* MerA for MD simulation and the results were then compared to experiments conducted on Tn*21* MerA (16). The ~30-residue linkers connecting the NmerA domains with the catalytic core domain were modeled as extended, random coils in the initial structure. The solvent-exposed cysteines, C11, C14, C558, and C559 on both monomers were mutated to alanine to be consistent with the SANS and NSE experiments. Crystallographic water molecules were retained, and all ionizable side chains were modeled in their standard, pH 7.0 protonation states.

The MD simulation of the fl-MerA in explicit water solvent was performed using the TITAN supercomputer at Oak Ridge National Laboratory. The MD simulation was performed with GROMACS using the CHARMM27 all-atom force field (17) and the TIP3P water model (18). The mass ratio of protein to water was ~3%, similar to the NSE experiment, leading to a total of ~740,000 atoms. More details of the MD protocol can be found in the Supporting Material.

Mesoscopic CG simulations

In the CG model, the catalytic core was represented by two rigid, touching triangulated spheres of the same radius, and the NmerA domains by spheres of smaller radius, as depicted in Fig. 1 *b*. The radius of gyration of the core and the NmerA domains, as well as the contour length of the linkers, were chosen in proportion to the crystal structures (Protein Data Bank (PDB) entries 1ZK7 (9) and 2KT2 (8)). Four systems were studied: a core without linkers or NmerA domains, a complete protein with flexible linkers, i.e., the flexible model, and two rigid models that are both full molecules with rigid linkers. One rigid model assumes the average conformation of the flexible model obtained from the CG simulation, resulting in a semiextended structure, and the other takes the minimum-R_g structure of fl-MerA acquired from MD simulation, i.e., a highly compact structure.

The fluid was modeled by the multiparticle collision dynamics approach (13,19), i.e., a hydrodynamic simulation method that incorporates thermal fluctuations and provides hydrodynamic interactions.

RESULTS

Structure

SANS experiments were performed in D_2O buffer solution for both samples (fl-MerA and the core) to extract the form factor per unit mass at infinite dilution, $F(q)$, shown in Fig. 2 *a*, from which the overall shape of the protein molecule at low resolution ~1 nm can be extracted. The intensity of the forward scattering, i.e., $F(q \rightarrow 0)$, should be linearly dependent on the molecular weight of the molecule studied (20). It was found that the ratio of $F(q \rightarrow 0)$ between fl-MerA and the core, 1.21, is almost the same as the ratio of the corresponding molecular weights, 1.19, confirming

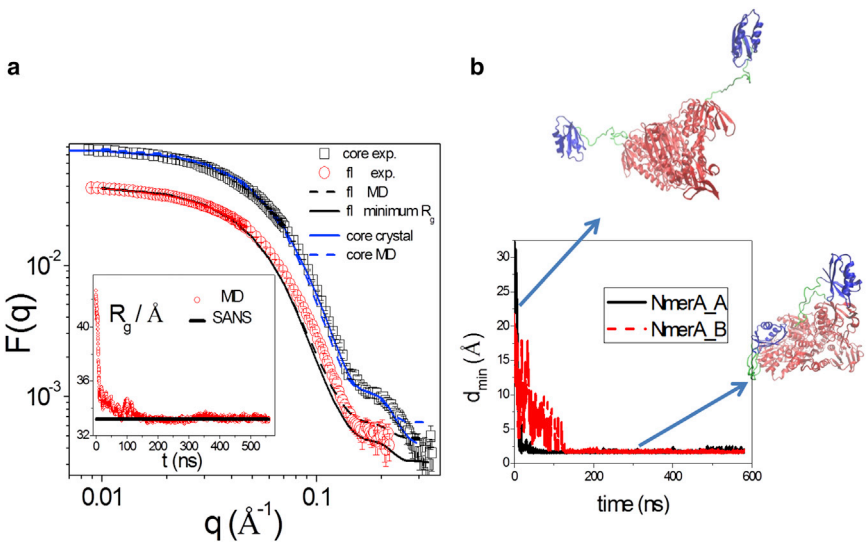

FIGURE 2 Structure profile of MerA measured by SANS and derived from MD simulation. (*a*) Form factor $F(q)$ describing the shape of full-length MerA and the core. The form factor of fl-MerA is shifted by a factor of 0.5 for clarity. Symbols are experimental data; dashed lines are the results derived from the last 400 ns of the MD trajectory; the solid blue curve denotes $F(q)$ calculated from the crystal structure of the core (PDB entry 1ZK7 (9)); and the solid black curve corresponds to the result calculated from the minimum-R_g structure of the fl-MerA obtained from MD simulation. The inset shows the time evolution of $R_{g,fl}$ over the entire 550 ns MD trajectory, where the value obtained from the SANS measurement is also plotted for comparison. The incoherent background contributes 0.00044 cm^{-1} for fl-MerA, and 0.00026 for the core, and these values were added into the simulation-derived $F(q)$ for comparison to the experimental data. (*b*) Time evolution of the minimum distance between each NmerA domain and the core obtained in MD. The two snapshots of the protein structure correspond to an extended conformation at $t = 0$ ns and a compact conformation at $t = 300$ ns, respectively. To see this figure in color, go online.

that the full-length MerA remains intact in solution, with the NmerA domains, linkers, and core contributing coherently to the scattering signal.

The radius of gyration of the protein molecule, R_g, was determined using the Guinier approximation in the range $0.009\ \text{Å}^{-1} < q < 0.03\ \text{Å}^{-1}$ by a least-square fit to the form factor $F(q) = F(0)\cdot\exp(-q^2R_g^2/3)$ (21). R_g of the core is 32.1 ± 0.3 Å, which is only 4% smaller than the value of fl-MerA ($R_{g,fl} = 33.2 \pm 0.4$ Å). Given that the mass of fl-MerA is ~20% larger than the core, this small difference in R_g suggests that fl-MerA adopts a very compact structure.

To further investigate the conformation of fl-MerA, all-atom MD simulation of the protein in aqueous solution was conducted. As seen in Fig. 2 *a*, $F(q)$ of the core domain calculated from the crystal structure and that derived from the MD trajectory are in quantitative agreement with the corresponding SANS data, if a reasonable background for incoherent scattering is assumed (see *details in legend* of Fig. 2).

The inset of Fig. 2 *a* shows the time evolution of the radius of gyration of the fl-MerA, $R_{g,fl}$, derived from the MD simulation. The initial conformation of the fl-MerA linker in MD was set to be a relatively extended random coil (Fig. 2 *b*). However, after ~50 ns the protein gradually collapses into a compact structure (Fig. 2 *b*), during which the minimum distance between NmerA and the core decreases from ~25 to 2 Å, at which point the NmerA domains are in direct contact with the core. A comparison between the experimental and MD-derived $R_{g,fl}$ (Fig. 2 *a*, *inset*) shows that the radius of gyration of the compact fl-MerA observed in the MD simulation agrees with the experimental result.

A direct comparison of $F(q)$ of fl-MerA derived from MD to that obtained experimentally is presented in Fig. 2 *a*. Two simulation-derived results are displayed: one calculated using the minimum-$R_{g,fl}$ protein structure obtained in MD, and the other calculated by averaging all the protein snapshots generated over the last 400 ns of the MD trajectory, during which MerA remained in a compact state. The simulation-derived $F(q)$ is in quantitative agreement with the experimental data with a small difference in the midrange of q (0.05–0.1 Å^{-1}). Hence, MD simulation and SANS experiments provide a consistent picture in which full-length MerA adopts a compact conformation with the NmerA domains in direct contact with the catalytic core.

The question arises as to what type of interaction stabilizes fl-MerA in the compact form (Fig. 2 *b*). Fig. 3 identifies the types of contacts between NmerA and core observed in the MD trajectory. A detailed list of the directly contacting residue pairs is presented in Table S3 in the Supporting Material. Most of these contacts are between oppositely charged residue pairs, suggesting that the electrostatic attraction between NmerA and the core stabilizes the collapsed form of fl-MerA.

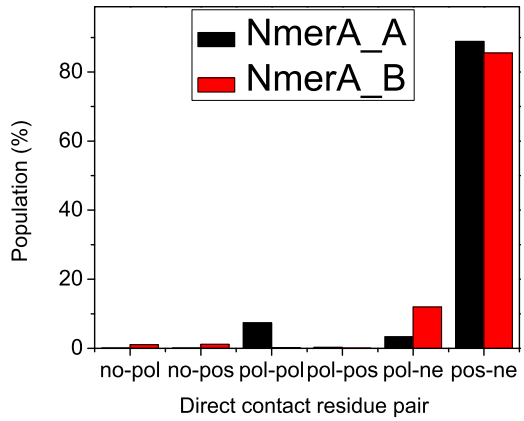

FIGURE 3 Populations of different types of residue pairs that are in direct contact between NmerA and the core observed in MD, no = nonpolar; pol = polar; pos = positive; and ne = negative. The charge of the core is −32 and each NmerA domain has a charge of +2. To see this figure in color, go online.

Dynamics

To obtain further insight into the compact state of fl-MerA, we compared the dynamics of core and fl-MerA using NSE spectroscopy, which directly probes correlated molecular motions on a timescale of 100 ps–100 ns and a length scale of several nm, i.e., spatiotemporal regions expected for the protein interdomain motion (22–24). The measured quantity is the coherent intermediate scattering function,

$$I_{coh}(q,t) = \sum_{i}^{Z}\sum_{j}^{Z} b_i b_j \left\langle \exp\left[-i\vec{q}\cdot\vec{R}_i(0)\right]\exp\left[i\vec{q}\cdot\vec{R}_j(t)\right]\right\rangle, \quad (1)$$

where $\vec{q}$ is the scattering wave vector, Z is the total number of atoms, b_j is the coherent scattering length, and $\vec{R}_j$ is the position of a given atom j, whereas the brackets denote averaging over the time origin and the orientation of $\vec{q}$. $I_{coh}(q,0)$ is the static scattering intensity as measured in a SANS experiment. The intermediate scattering function can also be calculated directly from simulation trajectories and analyzed in the same way as the experimental data.

$I_{coh}(q,t)/I_{coh}(q,0)$ measured for fl-MerA at different scattering wave vectors is presented in Fig. 4 *a*. Earlier studies on other multidomain proteins, such as phosphoglycerate kinase and alcohol dehydrogenase, showed that the decay of $I_{coh}(q,t)$ can be often described by two exponential contributions: the global motion (rotation and translation) of the protein molecule as a rigid body at long times, and the internal dynamics (mostly interdomain fluctuations) resulting in an additional relaxation at shorter times (22,23). However, for both fl-MerA and the core, only a single exponential decay, resulting from the global motion, exists in $I_{coh}(q,t)$ (Fig. 4 *a*). This indicates a relatively rigid structure for both samples with interdomain dynamics that is either of small amplitude or much slower than the accessible time scale.

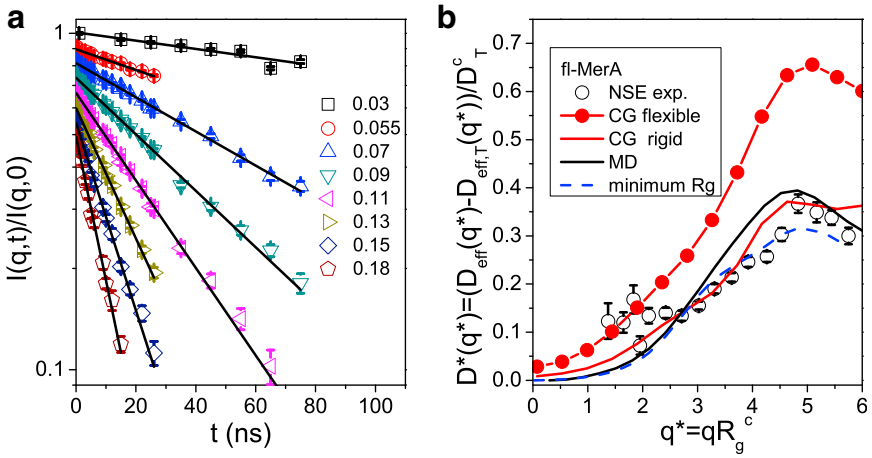

FIGURE 4 NSE results of MerA. (*a*) Intermediate scattering function measured for fl-MerA at a set of scattering wave vectors, q in Å^{-1}. The solid lines correspond to single exponential fits. Data are shifted by a consecutive factor of 0.9 for clarity. (*b*) Reduced effective diffusion coefficient $D^*(q^*) = (D_{eff}(q^*)\text{-}D_{eff,T}(q^*))/D^c{}_T$ vs. $q^* = qR_g^c$ for fl-MerA. Points represent NSE data. The dashed blue line was estimated using the analytical model (Eq. S1) based on the minimum-$R_{g,fl}$ protein structure obtained from MD (22,23), the solid black curve was derived from the last 400 ns MD trajectory, and the solid and dashed red lines were calculated from CG simulations of the rigid and flexible models, respectively. R_g^c and $D^c{}_T$ are the radius of gyration and the translational diffusion coefficient of the core, respectively, estimated using the HYDROPRO program (28). $D^*(q^*)$ derived from the CG models was averaged over eight independent simulations. The three lowest-q experimental data points in Fig. 4 *b* differ slightly from both MD and CG simulations. These small q data are dominated by overall (*whole-molecule*) diffusion. These three data points belong to the same detector image: the measurements were made with an area detector of ~30 cm × 30 cm and grouped into three slices for three individual q. It is possible that there was an operational error during this specific measurement. To see this figure in color, go online.

From the best estimates for the first cumulant of the coherent intermediate scattering function, $\lim_{t\to 0} d/dt \ln[I_{coh}(q,t)/I_{coh}(q,0)] = -\Gamma(q) = -q^2 D_{eff}(q)$, the relaxation rate $\Gamma(q)$ and the effective diffusion coefficient $D_{eff}(q)$ can be extracted, the latter carrying the essential information on the observed dynamics (22–24). In a nondilute solution, such as in the present work, the experimentally measured $D_{eff}(q)$ is a complex quantity, with components arising from self-diffusion of the protein molecule, the interparticle structure factor, $S(q)$, characterizing the interprotein scattering, and the hydrodynamic function, $H(q)$, due to the interprotein hydrodynamic interaction (see Eq. S1). $S(q)$ and $H(q)$ can be estimated independently as shown in the Supporting Material and (25). As seen in Fig. S3, when taking interprotein scattering and hydrodynamic interactions into account the experimental $D_{eff}(q)$ for the core can be well reproduced by both a CG simulation of a rigid MerA core and by the analytical model (Eq. S1) assuming diffusion of a rigid protein core, details of which are given in the Supporting Material (22,26). This again indicates that there is no significant internal dynamics in the core on the NSE time and length scales.

Fig. 4 *b* displays the NSE experimental data for fl-MerA, where $D^*(q^*) = (D_{eff}(q^*) - D_{eff,T}(q^*))/D_T{}^C$ and $q^* = q\,R_g{}^C$. Here, D_{eff} and $D_{eff,T}$ are the overall and translational diffusion coefficients of fl-MerA, whereas $D_T{}^C$ and $R_g{}^C$ are the translational diffusion coefficient and the radius of gyration of the core, respectively. $D^*(q)$ thus presented contains contributions from both the global rotation and internal dynamics of the protein. The long-range translational diffusion coefficient of the protein, measured using DLS, was found to agree with D_{eff} measured by NSE in the low-q limit and used to derive experimental $D_{eff,T}$ (q^*) (see Supporting Material and (23)).

$D^*(q)$ extracted from CG models with flexible and rigid linkers are also displayed in Fig. 4 *b*. In the flexible model the linker is highly dynamic, permitting the NmerA domains to detach from the core and explore the surroundings on the timescale accessible to the NSE, whereas the rigid model assumes the average conformation of the flexible model, i.e., a semiextended structure, with stiff linkers that are restrained by a strong harmonic potential (see the Supporting Material). A control analysis on a compact structure with the same rigid linkers was also performed, and the dynamics involved was found to be similar to that of the semiextended rigid model (Fig. S7).

The comparison between the flexible and rigid CG models highlights the direct contrast between the two previously mentioned extreme scenarios. There is a significant difference between the two profiles. D^* of the rigid model is in good agreement with experiment, whereas that of the flexible model is larger by approximately a factor of two. A fully flexible structure with large-scale motions of NmerA relative to the core, such as the previous flexible model, would significantly increase the D^* to well above the experimental error on the time and length scales probed by NSE. Hence, the scenario of complete flexibility can be excluded. In addition, the contribution to D^* from the global rotation of fl-MerA as a rigid body using the minimum-$R_{g,fl}$ protein structure obtained from MD was estimated according to the analytical model, Eq. S1 (22,23), and is presented in Fig. 4 *b*. This contribution is again similar to the experimental data, confirming that the global motion dominates the dynamical signal of fl-MerA probed by NSE.

Finally, D^* was also calculated from the MD-derived I_{coh} (q,t) by using the last 400 ns MD trajectory in which fl-MerA adopted a compact state. As shown in Fig. 4 *b*, the obtained D^* is again in good agreement with experiment. Furthermore, D^* derived from MD differs only

slightly from that calculated from the minimum-$R_{g,fl}$ protein structure and that from the CG rigid model, also indicating that the global motion dominates the NSE signal of fl-MerA. As discussed in (27), the finite box size of the MD simulation does affect the absolute value of the diffusion coefficient. However, the diffusion coefficients presented in Fig. 4 *b* are scaled by the translational diffusion coefficient, and it is expected that different components of the diffusion coefficient should be affected by the size of the simulation box in a similar manner. As a result, the rescaled diffusion coefficient is unlikely to be strongly influenced by the box size.

In summary, NSE, CG simulation, analysis of the rigid body motion using the analytical model (Eq. S1) and all-atom MD simulation all provide a consistent picture in which fl-MerA exhibits only low-amplitude internal dynamics relative to the global motion of the whole protein on the timescale investigated (~100 ps to 100 ns).

One approach that has been successful in earlier work, e.g. (29), is to use ensemble refinement to successfully locate the most likely ensemble of structures derived from the simulation that fits best to the small-angle scattering profile. However, such approach is beyond the scope of this work as the main focus here is the time evolution of the protein structure.

Detailed analysis of the MD simulation shows that, after collapsing onto the core, the NmerA domains are not completely immobilized. Instead, each domain diffuses short distances over a restricted area on the surface of the core (*gray zone* in Fig. 5 *a*) and is trapped intermittently at different positions through the interactions between the core and NmerA noted previously. Correspondingly, the mean-square atomic displacement (MSD) of the NmerA relative to the core (Fig. 5 *b*) exhibits a power-law time dependence ~$t^{1/4}$, indicating subdiffusive dynamics (30), with an amplitude of ~5 Å on the NSE timescale.

Further analysis shows that the solvent accessible surface area of the Hg(II)-binding cysteines on the NmerA domains remains almost constant throughout the entire 550 *ns* MD trajectory (Fig. 5 *c*). Thus, in addition to restricting the lateral movement of NmerA over the core surface, the interactions between NmerA and the core restrict the orientation of NmerA such that the Hg(II)-binding cysteines remain exposed to solvent even when NmerA collapses onto the core. Furthermore, the solvent interfacial area explored by the NmerA cysteines after NmerA collapses to the core is found to be ~10^3 Å^2, orders of magnitude larger than the solvent accessible surface area of the C-terminal cysteines (~40 Å^2). Hence, the presence of NmerA significantly increases the accessibility of MerA to Hg(II).

A key step in the function of fl-MerA is the diffusion of NmerA to deliver Hg(II) to the C-terminal cysteines. Due to the limited simulation time, this diffusive process was not observed to completion in the present MD simulation. Fig. 5 *a* labels the surface residues on the core (*cyan spheres*) that form direct contacts with NmerA. As seen in this figure, during the MD simulation NmerA diffuses in a rather localized space (*dark gray zone*) close to the C-terminal cysteines

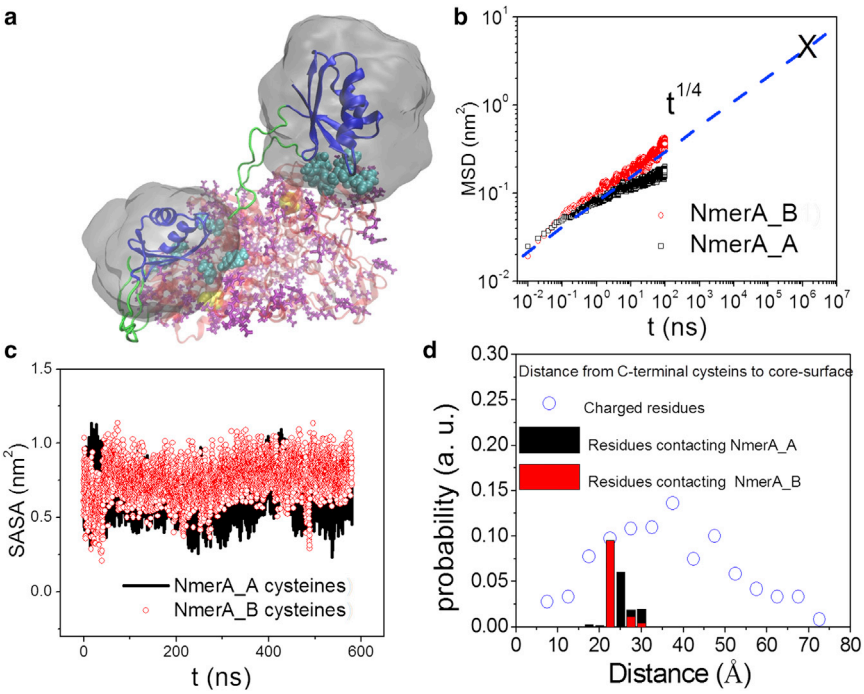

FIGURE 5 Interdomain dynamics in MerA observed in MD. (*a*) Protein structure at $t = 300$ ns, which has highly contracted linkers typical for compact fl-MerA. The dark gray regions denote the volume sampled by the NmerA domains after they collapse onto the core. Cyan spheres represent the core-surface residues, which form direct contacts with NmerA in MD, yellow spheres highlight the C-terminal cysteines, purple rods label the charged residues at the surface of the core, and the linkers are shown in green. (*b*) MSD of the NmerA domains relative to the core derived from the last 400 ns of the MD trajectory, with the coordinates of the core superimposed onto the frame at $t =$ 150 ns to eliminate the motion of the core. The symbol × is an estimate of the time required for NmerA to overcome a distance of 25 Å and reach the C-terminal cysteines by extrapolating the power-law dependence. Here, MSD is a three-dimensional property as normally defined. (*c*) Solvent-accessible surface area of the Hg(II)-binding cysteines on the NmerA domains. (*d*) Empty circles represent the distribution of the distance between the C-terminal cysteines and the charged residues on the surface of the core. The solid bars denote the distribution of the distances between the C-terminal cysteines and the core-surface residues forming direct contact with NmerA domains observed in the MD trajectory. To see this figure in color, go online.

(*yellow spheres*), over a distance of 25 ± 5 Å (Fig. 5 *d*). As a rough estimation, extrapolating the power-law relation in Fig. 5 *b*, i.e., MSD ~$t^{1/4}$, to longer times shows that NmerA would need ~1 ms to travel a further 20 Å to reach the C-terminal cysteines, i.e., to the mark × in Fig. 5 *b*. This function-related diffusive process is extremely sensitive to the distance between NmerA and the C-terminal cysteines, e.g., increasing the distance from 20 to 40 Å prolongs the time for NmerA to reach C-terminal cysteines by more than two orders of magnitude. Therefore, the restraint of NmerA close to the C-terminal cysteines, as demonstrated in this work, may be crucial for the efficient functioning of MerA.

Fig. 5 *a* marks all the charged residues on the surface of the core in purple. Most of these residues stay much further away from the C-terminal cysteines than the core residues visited by NmerA observed in the MD simulation (Fig. 5 *d*). Thus, if the interactions were determined solely by electrostatics, without the presence of the linkers, NmerA would, in principle, be able to interact anywhere on the surface of the core as the charged residues are uniformly distributed over the surface (Fig. 5 *a*). Therefore, the linkers may play an important role in restricting NmerA to a region close to the C-terminal cysteines (Fig. 5, *a* and *d*). Fig. 6 *a* shows that the end-to-end distance (d_{ee}) of each linker decreases ~50% after MerA collapses in the MD when $t >$ 50 ns, i.e., the linkers adopt contracted structures. Fig. 5 *a* displays a typical example of the linker structure (*green chain*) in the compact fl-MerA, which is highly contracted. Moreover, as seen in Fig. 6 *a*, the fluctuation of d_{ee} for the extended fl-MerA ($t <$ 50 ns) is much larger than that in the compact protein ($t >$ 50 ns). This feature is consistent with the finding that the linker region exhibits much larger root mean-square atomic fluctuations when fl-MerA is in an extended state than in the compact form (Fig. 6, *b* and *c*). Hence, the linkers, by adopting contracted, relatively rigid structures, restrict the motion of the NmerA domains to regions close to the C-terminal cysteines of the core.

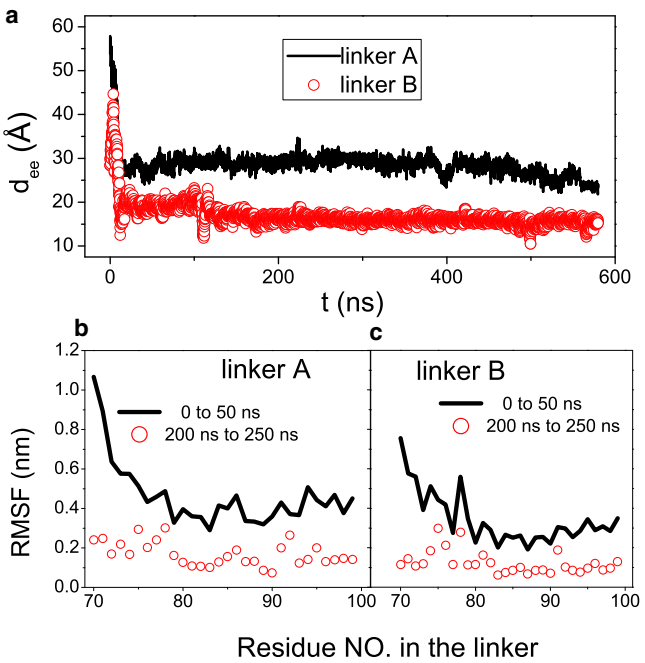

FIGURE 6 Linker flexibility derived from MD simulation. (*a*) End-to-end distance (d_{ee}) of each linker observed as a function of time. MD-derived root mean-square fluctuation of residues in (*b*) linker A and (*c*) linker B before (0–50 ns) and after (200–250 ns) the protein collapses into a compact state. Each linker consists of 29 residues and the corresponding residue numbers are displayed in (*b*) and (*c*) as abscissa. To see this figure in color, go online.

DISCUSSION

Here, we employ a combination of complementary methods—SANS, NSE spectroscopy, DLS, CG and all-atom MD simulations—to determine the shape and function-related internal dynamics of a multidomain protein in aqueous solution. The comparison of the SANS-derived radius of gyration of fl-MerA and MerA-core indicates that the full-length MerA assumes a compact structure. The NSE experiment furnishes information about the MerA internal dynamics, which is significantly suppressed and of small amplitude relative to the diffusion of the whole protein molecule (Fig. 4 *b*). Comparison of the NSE experimental results with the CG simulation indicates that the linkers are relatively rigid.

The all-atom MD, which is in agreement with the SANS and NSE experiments (Figs. 2 *a* and 4 *b*), provides additional insights. The NmerA domain interacts electrostatically with the core, such that the NmerA is in close contact with the core and conducts a subdiffusive motion over its surface. As the NmerA domains diffuse over the surface of the core, the Hg(II)-binding cysteines remain exposed to the solvent, potentially facilitating access of MerA to Hg(II), and the NmerA domains are held in the neighborhood of the C-terminal cysteines, leashed by the linkers, potentially enabling fast delivery of Hg(II) to the core. In this way, the presence of NmerA may appreciably enhance the functional efficiency of MerA, consistent with in vivo experiments demonstrating that the presence of the NmerA domains significantly improves cell survival (9). The present results are also in agreement with a proposed two-step process suggesting that NmerA is able to capture Hg(II) from other proteins and deliver it to the C-terminal cysteines faster than the direct transfer of Hg(II) from the proteins (9).

CONCLUSIONS

The determination of structure and dynamics of multisubunit protein complexes such as mercuric ion reductase, in aqueous solution is critical to developing a thorough understanding of their function but presents considerable technical challenges. The present approach, integrating multiple simulation and experimental methods, enables a

coherent, multiscale picture to be established of the relationship between shape, interdomain dynamics, and atomic-detail interactions. The present approach may be of general use in the elucidation of the functional dynamics of multidomain protein complexes.

SUPPORTING MATERIAL

Supporting Material, seven figures, three tables, and three equations are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(14)00620-1.

We acknowledge support from National Science Foundation (NSF) grant MCB-0842871 and grants ER65062 and ER65063 from the Subsurface Biogeochemical Research Program, Office of Biological and Environmental Research, U.S. Department of Energy (DOE). This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under contract No. DE-AC05-00OR22725.

REFERENCES

1. Kuriyan, J., and D. Eisenberg. 2007. The origin of protein interactions and allostery in colocalization. *Nature.* 450:983–990.
2. Krishnamurthy, V. M., V. Semetey, …, G. M. Whitesides. 2007. Dependence of effective molarity on linker length for an intramolecular protein-ligand system. *J. Am. Chem. Soc.* 129:1312–1320.
3. Anthis, N. J., and G. M. Clore. 2013. The length of the calmodulin linker determines the extent of transient interdomain association and target affinity. *J. Am. Chem. Soc.* 135:9648–9651.
4. Barkay, T., S. M. Miller, and A. O. Summers. 2003. Bacterial mercury resistance from atoms to ecosystems. *FEMS Microbiol. Rev.* 27:355–384.
5. Finney, L. A., and T. V. O'Halloran. 2003. Transition metal speciation in the cell: insights from the chemistry of metal ion receptors. *Science.* 300:931–936.
6. Silver, S., and L. T. Phung. 1996. Bacterial heavy metal resistance: new surprises. *Annu. Rev. Microbiol.* 50:753–789.
7. Barkay, T., K. Kritee, …, G. Geesey. 2010. A thermophilic bacterial origin and subsequent constraints by redox, light and salinity on the evolution of the microbial mercuric reductase. *Environ. Microbiol.* 12:2904–2917.
8. Ledwidge, R., B. Hong, …, S. M. Miller. 2010. NmerA of Tn501 mercuric ion reductase: structural modulation of the pKa values of the metal binding cysteine thiols. *Biochemistry.* 49:8988–8998.
9. Ledwidge, R., B. Patel, …, S. M. Miller. 2005. NmerA, the metal binding domain of mercuric ion reductase, removes Hg2+ from proteins, delivers it to the catalytic core, and protects cells under glutathione-depleted conditions. *Biochemistry.* 44:11402–11416.
10. Engst, S., and S. M. Miller. 1998. Rapid reduction of Hg(II) by mercuric ion reductase does not require the conserved C-terminal cysteine pair using HgBr2 as the substrate. *Biochemistry.* 37:11496–11507.
11. Engst, S., and S. M. Miller. 1999. Alternative routes for entry of HgX2 into the active site of mercuric ion reductase depend on the nature of the X ligands. *Biochemistry.* 38:3519–3529.
12. Kapral, R. 2008. Multiparticle collision dynamics: simulation of complex systems on mesoscales. *Adv. Chem. Phys.* 140:89–146.
13. Gompper, G., T. Ihle, …, R. G. Winkler. 2009. Multi-particle collision dynamics: a particle-based mesoscale simulation approach to the hydrodynamics of complex fluids. *Adv. Polym. Sci.* 221:1–87.
14. Monkenbusch, M., R. Schatzler, and D. Richter. 1997. The Julich neutron spin-echo spectrometer—design and performance. *Nucl. Instr. Meth. Phys. Res. A.* 399:301–323.
15. Provencher, S. W. 1982. Contin: a general-purpose constrained regularization program for inverting noisy linear algebraic and integral-equations. *Comput. Phys. Commun.* 27:229–242.
16. Johs, A., I. M. Harwood, …, S. M. Miller. 2011. Structural characterization of intramolecular Hg(2+) transfer between flexibly linked domains of mercuric ion reductase. *J. Mol. Biol.* 413:639–656.
17. MacKerell, A. D., D. Bashford, …, M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* 102:3586–3616.
18. Jorgensen, W. L., J. Chandrasekhar, …, M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.
19. Malevanets, A., and R. Kapral. 1999. Mesoscopic model for solvent dynamics. *J. Chem. Phys.* 110:8605–8613.
20. Mylonas, E., and D. I. Svergun. 2007. Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. *J. Appl. Cryst.* 40:S245–S249.
21. Guinier, A., and G. Fournet. 1955. Small Angle Scattering of X-Rays. Wiley, New York.
22. Biehl, R., B. Hoffmann, …, D. Richter. 2008. Direct observation of correlated interdomain motion in alcohol dehydrogenase. *Phys. Rev. Lett.* 101:138102.
23. Inoue, R., R. Biehl, …, D. Richter. 2010. Large domain fluctuations on 50-ns timescale enable catalytic activity in phosphoglycerate kinase. *Biophys. J.* 99:2309–2317.
24. Bu, Z., R. Biehl, …, D. J. E. Callaway. 2005. Coupled protein domain motion in Taq polymerase revealed by neutron spin-echo spectroscopy. *Proc. Natl. Acad. Sci. USA.* 102:17646–17651.
25. Biehl, R., M. Monkenbusch, and D. Richter. 2011. Exploring internal protein dynamics by neutron spin echo spectroscopy. *Soft Matter.* 7:1299–1307.
26. García De La Torre, J., M. L. Huertas, and B. Carrasco. 2000. Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys. J.* 78:719–730.
27. Smolin, N., R. Biehl, …, J. C. Smith. 2012. Functional domain motions in proteins on the ~1–100 ns timescale: comparison of neutron spin-echo spectroscopy of phosphoglycerate kinase with molecular-dynamics simulation. *Biophys. J.* 102:1108–1117.
28. Ortega, A., D. Amorós, and J. García de la Torre. 2011. Prediction of hydordynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys. J.* 101:892–898.
29. Francis, D. M., B. Różycki, …, R. Page. 2011. Resting and active states of the ERK2:HePTP complex. *J. Am. Chem. Soc.* 133:17138–17141.
30. Hong, L., N. Smolin, …, J. C. Smith. 2011. Three classes of motion in the dynamic neutron-scattering susceptibility of a globular protein. *Phys. Rev. Lett.* 107:148102.