



OPEN

Solution structure of multi-domain protein ER-60 studied by aggregation-free SAXS and coarse-grained-MD simulation

Aya Okuda^{1,2}, Masahiro Shimizu^{1,2}, Ken Morishima¹, Rintaro Inoue¹, Nobuhiro Sato¹, Reiko Urade^{1✉} & Masaaki Sugiyama^{1✉}

Multi-domain proteins (MDPs) show a variety of domain conformations under physiological conditions, regulating their functions through such conformational changes. One of the typical MDPs, ER-60 which is a protein folding enzyme, has a U-shape with four domains and is thought to have different domain conformations in solution depending on the redox state at the active centres of the edge domains. In this work, an aggregation-free small-angle X-ray scattering revealed that the structures of oxidized and reduced ER-60 in solution are different from each other and are also different from those in the crystal. Furthermore, structural modelling with coarse-grained molecular dynamics simulation indicated that the distance between the two edge domains of oxidized ER-60 is longer than that of reduced ER-60. In addition, one of the edge domains has a more flexible conformation than the other.

Multi-domain proteins (MDPs) constitute one of the major categories of ubiquitous proteins, presumably accounting for over half of all human proteins^{1,2}. MDPs can have several domain conformations and dynamically transit between them under physiological conditions, in solution. The domain conformation plays a role in the association with other proteins, nucleic acids, small molecules, etc., therefore determining a specific MDP function. Thus, elucidating the domain conformations and their dynamics in solution is important.

We have focused on ER-60, also called ERp57, GRP58, or PDIA3^{3–13}, which is a protein folding enzyme belonging to the protein disulphide isomerase family. ER-60 is also a typical MDP consisting of four thioredoxin-like domains, **a**, **b**, **b'**, and **a'**^{14,15}. Two edge domains, **a** and **a'**, have catalytically active Cys–Gly–His–Cys (CGHC) motifs, which interconvert between oxidized (disulphide) and reduced (dithiol) forms. The oxidized form catalyses the oxidation reaction of dithiol on a substrate (an unfolded protein), whereas the reduced form catalyses the reduction or isomerization of disulphide bonds on a substrate. In this catalytic process, ER-60 must have appropriate domain conformations to perform the formation and rearrangement of disulphide bonds on the substrate¹⁶. Therefore, it is assumed that ER-60 can change its domain conformation according to the redox state of the CGHC motif to either oxidize or reduce the substrate.

The atomic-scale structure of full-length human ER-60 has only been revealed in complex with tapasin in an X-ray crystallographic study¹⁸. As shown in Supplementary Fig. S1a, in the complex, the four domains of ER-60 are arranged in a twisted U-shape and the N-terminal domain of tapasin is sandwiched between the two edge domains, **a** and **a'**. This suggested that, under physiological conditions, namely in solution, the domains of ER-60 could also form a similar U-shape and associate with a substrate using the **a** and **a'** domains, and that the conformation of the **a** and **a'** domains could be changed according to the redox state of the CGHC motif. To confirm this, the structure of ER-60 in solution, especially the conformation of the four domains, and its dependence on the redox state of the CGHC motif, must be revealed.

For this purpose, small-angle X-ray scattering (SAXS) is a powerful tool that can be used to observe the structure of biomacromolecules in solution. However, SAXS is sensitive to unfavourable contaminations and aggregates in solution, especially large molecules when even in small amounts^{18–21}. Therefore, we carefully prepared highly purified oxidized and reduced ER-60, examined the quality of the samples, and conducted state-of-the-art SAXS methods combined with size-exclusion chromatography (SEC) (SEC-SAXS)²² and with analytical

¹Institute for Integrative Radiation and Nuclear Science, Kyoto University, Kumatori, Sennan-gun, Osaka 590-0494, Japan. ²These authors contributed equally: Aya Okuda and Masahiro Shimizu. ✉email: urade.reiko.8w@kyoto-u.ac.jp; sugiyama@rri.kyoto-u.ac.jp

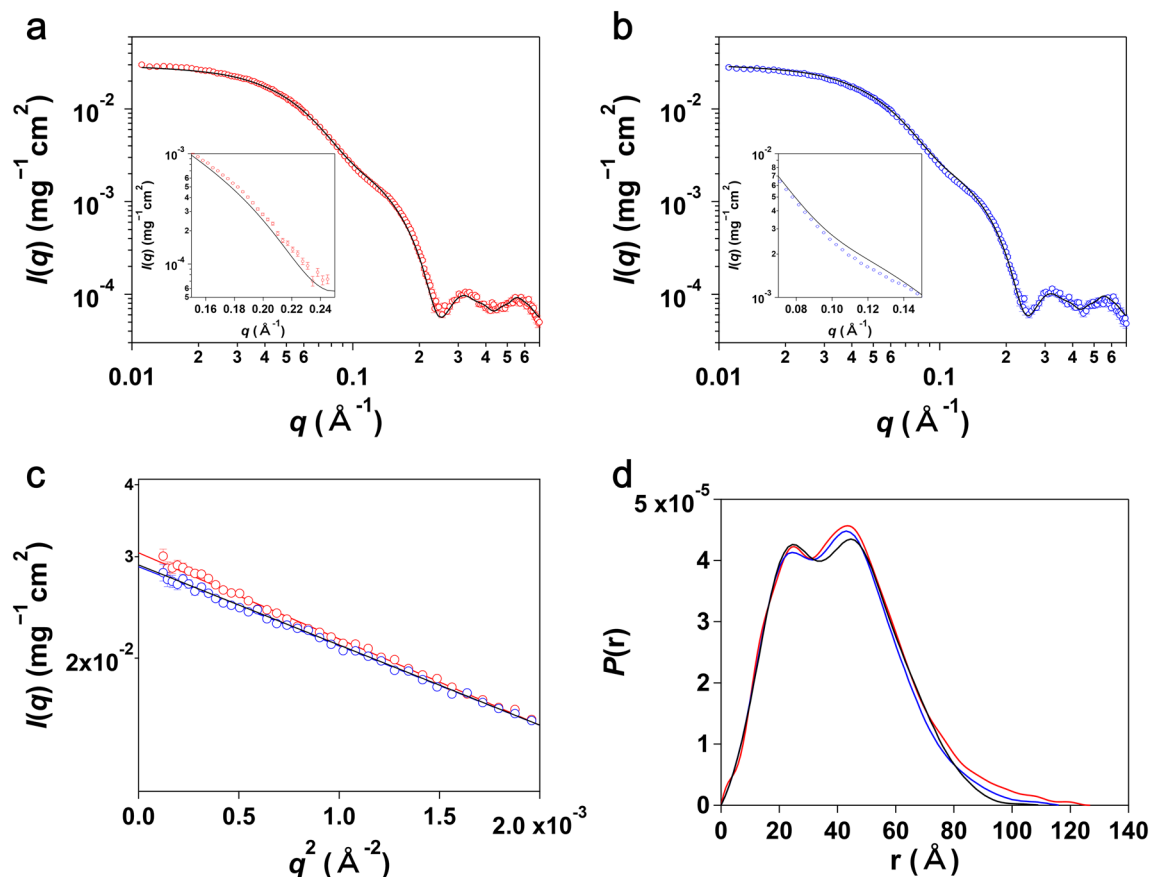


Figure 1. Solution scattering of ER-60. **(a, b)** Small-angle X-ray scattering (SAXS) profiles of oxidized **(a, red circles)** ER-60 and reduced **(b, blue circles)** ER-60, respectively. Black lines indicate the profiles calculated from the crystal structure (3F8U). Inset panels are expanded views of them in a q range of 0.17–0.25 \AA^{-1} **(a)** and 0.07–0.15 \AA^{-1} **(b)**, respectively. **(c)** Guinier plots of oxidized (red circles) and reduced (blue circles) ER-60. The red and blue straight lines indicate the results of the least square fitting using the Guinier formula: $I(q) = I(0) \exp(-R_g/3)^2 q^2$, where R_g is the gyration radius and $I(0)$ is the forward scattering intensity. The black line indicates the profiles of the crystal structure (3F8U) in the Guinier plot. **(d)** Pair distance distribution function, $P(r)$, of oxidized (red line) and reduced (blue line) ER-60 calculated using GNOM software³⁶. The black line indicates the $P(r)$ of the crystal structure (3F8U).

ultracentrifugation (AUC) (AUC-SAS)²¹. The integrative organization of these methods allowed us to achieve high-precision aggregation-free SAXS profiles.

To visually understand domain conformation, we built three-dimensional structures of oxidized and reduced ER-60 in solution by computational modelling using the SAXS data. Prior to the computational modelling, we compared two crystal structures of the **b–b'** domains: one is the **b–b'** domains of full-length ER-60 in the tapasin complex (3F8U) and the other is only **b–b'** domains, ER-60 truncated the **a** and **a'** domains²³ (2H8L) (Supplementary Fig. S1b). Since they were almost the same as indicated by the calculated root mean square deviation of *ca.* 0.5 (Supplementary Fig. S1b), we assumed that the conformation of the **b** and **b'** domains are almost fixed, while that of the two edge domains, **a** and **a'**, can be flexible. Under this assumption, we designed the proper interaction potential of intra- and inter-domains and then performed coarse-grained molecular dynamics (CG-MD) simulations to search for the possible structural space of ER-60. Compared to standard atomistic MD simulations, the CG-MD approach drastically accelerates structural space sampling^{24,25}. The sampled models were screened with the aggregation-free SAXS profiles of the oxidized and reduced ER-60 forms.

Finally, we succeeded in deriving two sets of structural models from simulation snapshots, which represented the solution structures of oxidized and reduced ER-60, respectively. From the structural models, we examined the differences in their domain conformations.

Results and discussion

Solution scattering of ER-60. Figure 1a,b show the aggregation-free SAXS profiles of oxidized and reduced ER-60, respectively, which were obtained by the integrative organization of the methods described in the “Aggregation-free SAXS data” subsection of the “Methods” section. First, we compared the observed SAXS profiles of ER-60 with that calculated using the crystal structure (3F8U). As emphasized in the inset panels of Fig. 1a,b, the SAXS profiles of oxidized and reduced ER-60 in solution were different from that calculated using

	R_g (Å)	D_{max} (Å)	Measured average mass (Da)	Calculated average mass (Da)	Peak top of elution volume (mL)
ER-60					
Oxidized	32.1 ± 0.2	126.7 ± 4.9	54,259.3	54,260.5	13.48
Reduced	30.8 ± 0.2	117.1 ± 4.2	54,264.0	54,264.6	13.60
Crystal	31.0	109.0	–	52,807.8	–

Table 1. The values obtained from the SAXS measurement, SEC, and ESI-TOF MS of ER-60.

the crystal structure. Furthermore, the calculated values of χ^2 between the observed SAXS profiles and that calculated using the crystal structure were 13.5 and 10.7 for oxidized and reduced ER-60, respectively. This means that the complex-free structures in solution are different from the crystal structure of the complexed protein.

Next, we examined the structural differences between oxidized and reduced ER-60. Their Guinier plots and distance distribution functions are displayed in Fig. 1c,d, respectively. The gyration radii R_g with Guinier analyses and the longest distance D_{max} from the distance distribution functions are presented in Table 1. In addition, the elution volumes for oxidized and reduced ER-60 on SEC are also described in Table 1. The elution charts are shown in Supplementary Fig. S2. The average mass values of the two samples, oxidized and reduced ER-60, were measured using LC-ESI-TOF mass spectrometry. The measured average mass values were the same as those calculated from their amino acid sequences (Table 1 and Supplementary Fig. S3). These results and SDS-PAGE of the samples (Supplementary Fig. S4d,e) showed that the oxidized and reduced samples were intact without any truncation or degradation. Therefore, the larger R_g , the longer D_{max} , and the smaller elution volume of the oxidized ER-60 compared to reduced ER-60 clearly indicated that oxidized ER-60 forms a more expanded structure than reduced ER-60 in solution.

In summary, in terms of χ^2 , there are no large differences between the three structures: the oxidized and reduced ER-60, and the crystal structures of the complexed protein. However, we have succeeded to find the small but certain differences with our careful sample preparation and high-precision SAXS measurements.

Structural modelling of ER-60. We built structural and dynamical models for both forms, which reflect their differences. Considering the structural differences between two forms and the crystal form are not large, it is assumed that the solution structures for both forms could distribute homogeneously around the crystal structure. Therefore, we extensively sampled structures with various domain conformations by CG-MD. Subsequently, the structures that reproduced the experimental SAXS profiles for both forms were collected from the sampled structures.

First, we took 800,000 snapshots with CG-MD (see the details in the “Supplemental methods”). Then, SAXS profiles were calculated for all snapshots. The structural models reproducing the observed SAXS profile of oxidized ER-60 (ox-models) were extracted from all snapshots using two evaluation criteria: $\chi^2 < 7.0$ and $31.9 \text{ Å} < R_g < 32.3 \text{ Å}$. The total number of ox-models was 230. The structural models reproducing the observed SAXS profile of reduced ER-60 (red-models) were extracted with two evaluation criteria: $\chi^2 < 7.0$ and $30.6 \text{ Å} < R_g < 31.0 \text{ Å}$. The total number of red-models was 1572. Figure 2 shows typical extracted structures, indicating that all ox- and red-models have a twisted U-shape.

Subsequently, we investigated the differences in domain conformation between the ox- and red-models. Three angles and one distance, which characterise domain conformation, were calculated for the ox- and red-models. As shown in Supplementary Fig. S5, $\theta_{a-b-b'}$ and $\theta_{b-b'-a'}$ are defined as angles formed by the **a**, **b**, **b'** and **b**, **b'**, **a'** domains, respectively, while $\varphi_{a-b-b'-a'}$ is defined as the dihedral angle formed by all four domains and $D_{a-a'}$ is the distance between the centre of mass (COM) of the **a** and **a'** domains. As shown in Fig. 3a,b, $\theta_{a-b-b'}$ showed similar ranges distributed around 75° in both models, whereas $\theta_{b-b'-a'}$ showed relatively different ranges, which were 110° – 155° and 95° – 150° for ox- and red-models, respectively. This suggested that the **a** domains of the ox- and red-models have almost the same position, while the **a'** domains of the ox-models locate at a more open position than those of the red models. Moreover, as shown in Fig. 3c, $\varphi_{a-b-b'-a'}$ had a similar range distribution in both models, meaning that the degrees of twisting of the U-shape were almost the same. The effect of the more open conformation in the ox-models is clearly recognized in the $D_{a-a'}$. As shown in Fig. 3d, the $D_{a-a'}$ of the ox-models was distributed over a longer range than that of the red models.

The structural images are shown in Fig. 4. For visualisation, the **b** and **b'** domains of all models were superimposed, and then the positions of the COM of every domain were averaged over the ox- and red-models, respectively. The COM-positions were expressed with small spheres and the subscripts for **a** and **a'** domains, o and r, corresponding to ox- and red-model(s), respectively. As shown in Fig. 4, **a**_o and **a**_r are almost the same, but **a'**_o is slightly shifted to be more open than **a'**_r. The distance between averaged **a**_o position and averaged **a**_r position was 3.4 Å, whereas that between averaged **a'**_o and averaged **a'**_r was 4.7 Å. This also suggests that the structures of oxidized and reduced ER-60 mainly differ in the position of the **a'** domain. The results are summarized in Table 2.

The details of these conformational features were clarified by referring to the relations between the parameters. Figure 5 shows the contour maps of the frequency distribution as functions of two parameters in ox- and red-models and the difference between them. Here, we focused on the relations concerning the conformation of the **a'** domain. As shown in Fig. 5d,e, the **a'** domains in both the ox- and red-models were distributed keeping the distances $D_{a-a'}$. This implied that the **a'** domain rotates around the **a** domain, which results in the twisting of the U-shape because the **a** domain is almost fixed. This motion is shown in Fig. 5m,n. In the case of the ox-models,

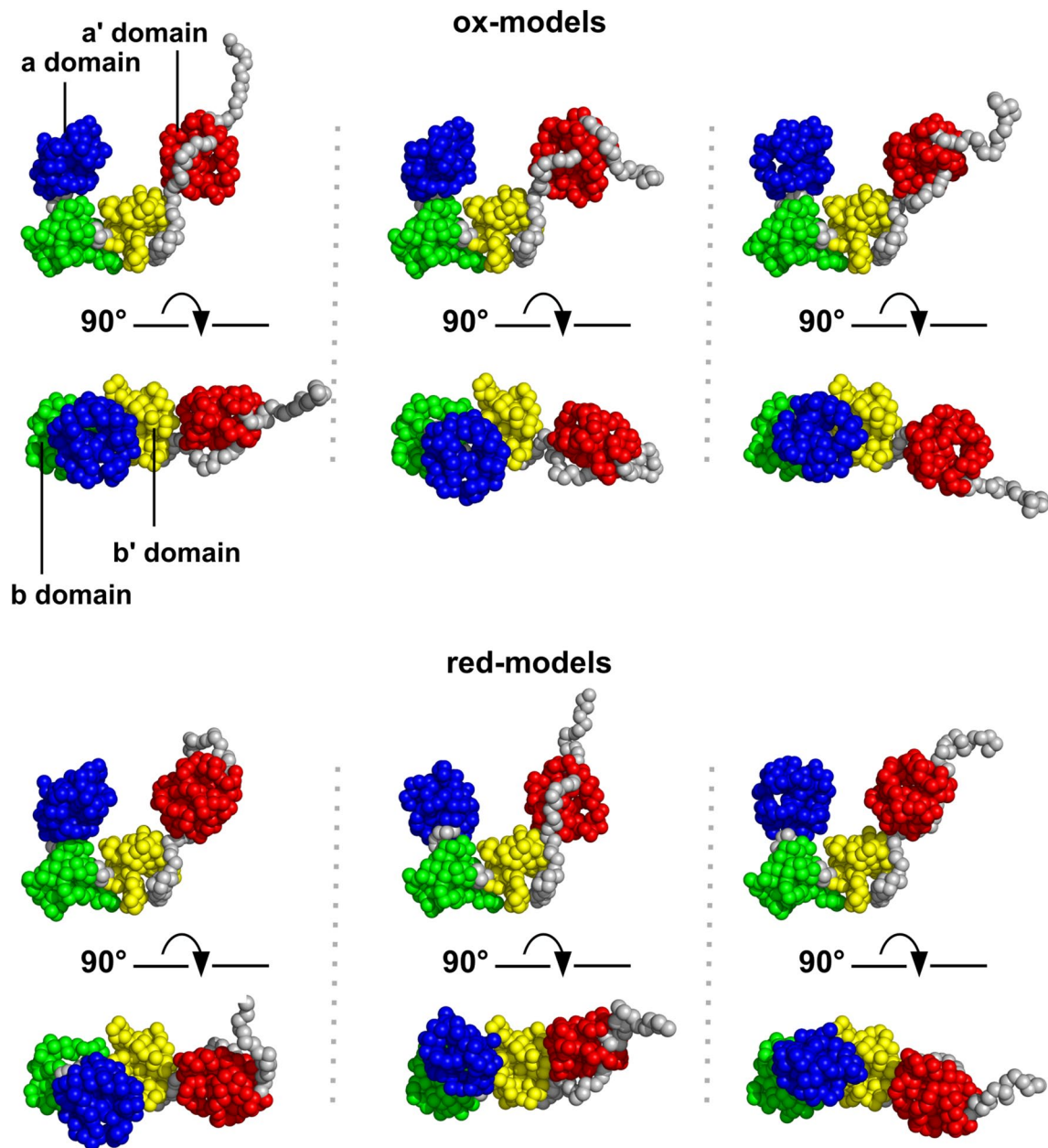


Figure 2. Oxidized ER-60 (ox-models) and reduced ER-60 (red-models) models. A total of six models were taken for oxidized and reduced ER-60 (three each) from the extracted ox- and red-models and are shown from two viewpoints. Here, the **a**, **b**, **b'**, **a'** domains and the inter-domain regions are coloured in blue, green, yellow, red, and gray, respectively.

when the **a'** domain shifts from $\theta_{b-b'-a'} = 140^\circ$ to $\theta_{b-b'-a'} = 110^\circ$, meaning that the **a'** domain rises up to the **a** domain, the dihedral angle changes from $\varphi_{a-b-b'-a'} = 0^\circ$ to $\varphi_{a-b-b'-a'} = 60^\circ$, meaning that the flat U-shape is being twisted (Fig. 5m). More correctly, the slightly fluctuated **a** domain ($\Delta\theta_{a-b-b'} = 10^\circ$) correlated with the position shift of the **a'** domain ($\Delta\theta_{b-b'-a'} = 30^\circ$). As shown in Fig. 5j,m, triple angle relations ($\theta_{a-b-b'}$, $\theta_{b-b'-a'}$, $\varphi_{a-b-b'-a'}$) were 80° , 140° , 0° and 70° , 110° , 60° , respectively. The difference in conformational flexibility between the **a** and **a'** domains could be ascribed to the linkers between the **b** and **b'** domains. In the red-models, same conformational relations such as triple angle relations ($\theta_{a-b-b'}$, $\theta_{b-b'-a'}$, $\varphi_{a-b-b'-a'}$) were observed from 80° , 140° , 0° to 70° , 110° , 60° . The difference from the ox-models is the shorter $D_{a-a'}$, as described. These structural pictures are shown in Fig. 6. It should be noted that no correlations were found between the conformation of C-terminal loop of ER-60 and $\theta_{a-b-b'}$, $\theta_{b-b'-a'}$, $\varphi_{a-b-b'-a'}$, or $D_{a-a'}$ (Supplementary Fig. S6).

We also compared the ox- and red-models with the crystal structure. In Fig. 3a–d, the four angles of the crystal structure are indicated by black lines. Briefly, in the tapasin complex form, the **a** domain slightly turns to the **b** and **b'** domains, and the **a'** domain has the most open conformation and rotates by inserting tapasin between the **a** and **a'** domains, making their distance longer. As a result, the U-shape of the crystal structure is more twisted than those of the two complex-free structures, ox- and red-models, in solution. This is illustrated in Fig. 7. In

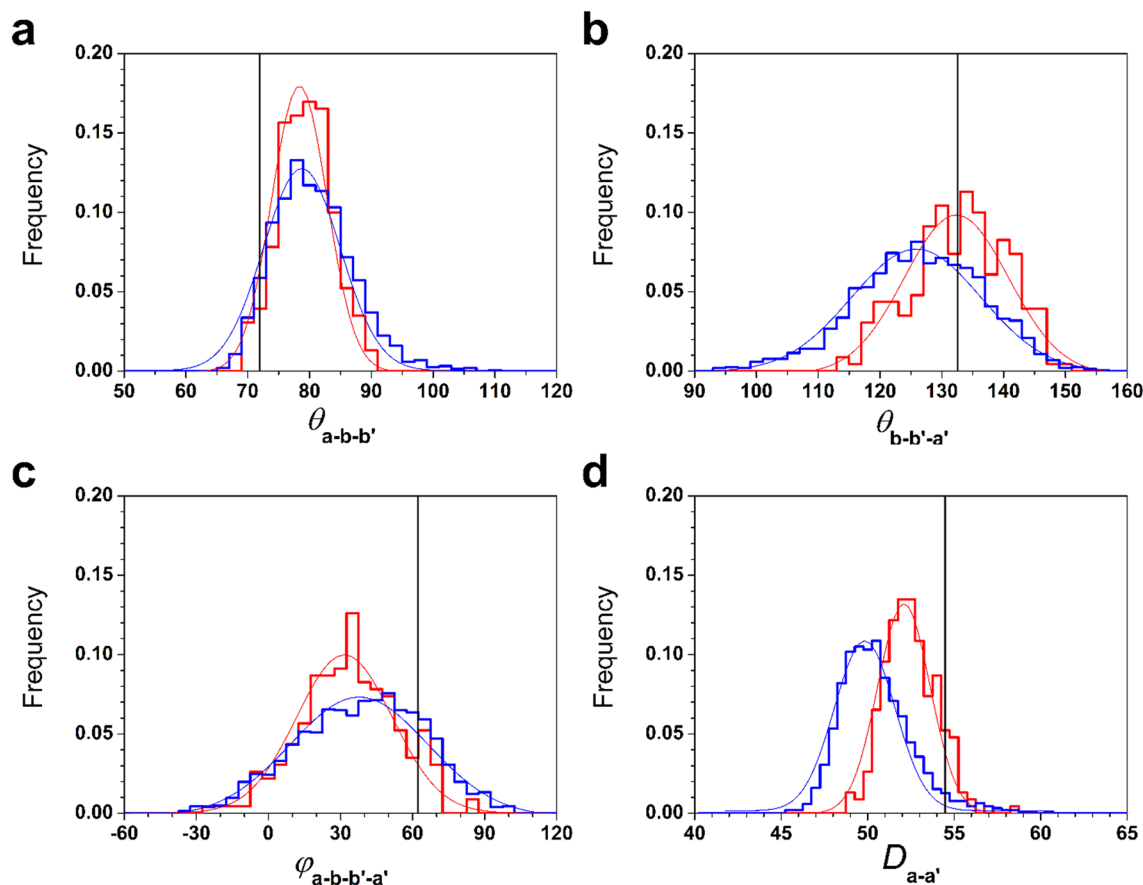


Figure 3. Structural analysis of the ER-60 models. Distributions of $\theta_{a-b-b'}$ (a), $\theta_{b-b'-a'}$ (b), and $\varphi_{a-b-b'-a'}$ (c), and the $D_{a-a'}$ (d) are shown with red and blue histograms. The frequency is the fraction of the count in each bin divided by the total number of models. The values of the crystal structure are also shown by black lines. Histograms were fit by Gaussians for visibility. The eye-guide curves are shown with red and blue lines for ox- and red-models, respectively.

addition, the conformational flexibility of the **a'** domain also contributes to the twisted structure in the crystal. In conclusion, complex-free ER-60 in solution acquires a less twisted structure than the complexed protein by releasing tapasin. The twisted form acquired by the conformational flexibility of the **a'** domain could contribute to capture a variety of proteins for the formation (oxidation) and cleavage (reduction) of disulphide bonds.

In summary, the combination of aggregation-free SAXS and CG-MD simulations is a powerful tool to analyse the structure of MDPs in solution, especially for domain conformation fluctuation.

Aggregation-free SAXS data of oxidized and reduced ER-60 were acquired by integrating the following procedures: highly purified sample preparation, step-by-step sample quality checks, SEC-SAXS, and AUC-SAS. The SAXS data indicated that both solution structures are different from the crystal structure of the tapasin-ER-60 complex and that the structure of oxidized and reduced ER-60 differs; the former is slightly more expanded than the latter. To analyse the structures and fluctuations of their domain conformations in more detail, three-dimensional structures were constructed with CG-MD simulation, which allows searching for a large phase space within a relatively short time compared with the all-atom MD simulation. CG-MD analysis revealed that the U-shape in the crystallographic structure is held in solution, but the distances between the **a** and **a'** domains are different (longer in the crystal, then in oxidized ER-60, and finally in reduced ER-60). Interestingly, the **a'** domain could fluctuate more than the **a** domain, may allow twisting of the U-shape. We speculate that the twisting resulting from the fluctuation of the **a'** domain may allow ER-60 to capture a variety of substrates. To probe this will be the next research target.

Methods

Aggregation-free SAXS. We developed several methods to obtain aggregation-free SAXS data. Figure 8 shows the procedure, which consists of sample preparation, its purification (SEC), quality checks, SAXS measurement, radiation damage check, and data refinement. The entire procedure was performed as follows.

Sample preparation. Recombinant ER-60 was expressed as described previously²⁶. Briefly, *E. coli* BL21(DE3) cells (Takara Bio, Inc., Shiga, Japan) transformed with expression plasmids encoding mature human ER-60 (Ser²⁵-Leu⁵⁰⁵) without signal peptide were precultured in LB broth at 37 °C for 5 h. The expression was induced

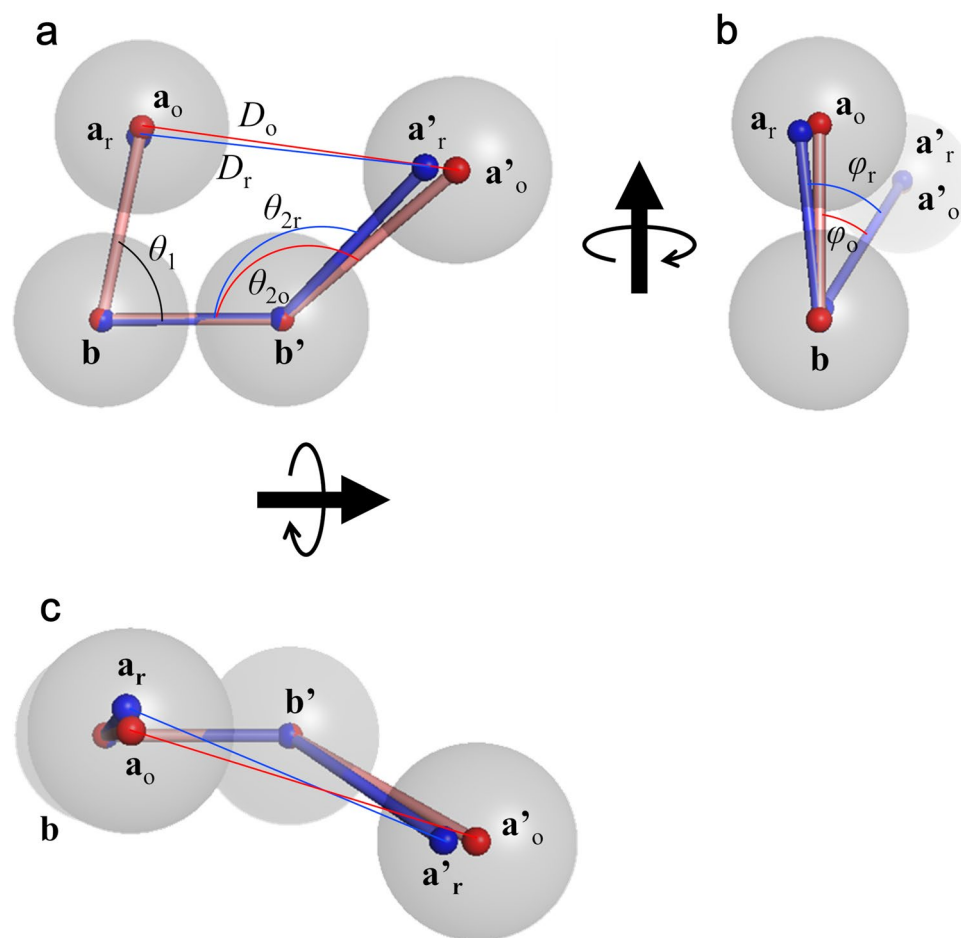


Figure 4. Ball-and-stick models of ER-60. The averaged position of each domain is shown with red ball-and-stick models and with blue ball-and-stick models for ox- and red-models, respectively. The ox- and red-models are aligned by the **b** and **b'** domains. Gray balls roughly indicate domains as an eye guide. θ_1 is the $\theta_{a-b-b'}$. θ_{2o} and θ_{2r} indicate the $\theta_{b-b'-a'}$ of the ox- and red-models, respectively. φ_o and φ_r indicate the $\varphi_{a-b-b'-a'}$ of the ox- and red-models, respectively.

	$\theta_{a-b-b'}$ [°]	$\theta_{b-b'-a'}$ [°]	$\varphi_{a-b-b'-a'}$ [°]	$D_{a-a'}$ [Å]
Oxidized	78.4 ± 4.3	131.8 ± 7.6	31.4 ± 20.0	52.2 ± 1.5
Reduced	79.6 ± 6.4	125.4 ± 10.3	36.1 ± 25.8	50.3 ± 2.1
Crystal	72.0	132.6	62.2	54.5

Table 2. The values from the structural analysis of the ER-60 models.

in the presence of 0.4 mM isopropyl- β -D-thiogalactopyranoside at 25 °C for 48 h. Purification procedures of the recombinant ER-60 were carried out at 4 °C. *E. coli* cells suspended in buffer A (20 mM HEPES buffer, pH 6.8, containing 50 mM KCl, 5 mM EDTA, and 1 mM phenylmethylsulphonyl fluoride) were disrupted by sonication and centrifuged at $10,000 \times g$ at 4 °C. The expressed recombinant ER-60 was purified from the cell lysates using affinity column chromatography on AF heparin toyopearl 650 M (Tosoh, Tokyo, Japan) with buffer A at 4 °C. Recombinant ER-60 was eluted from the AF heparin column with 0.4 M KCl in buffer A and subsequently purified by ion-exchange column chromatography on Resource Q (GE Healthcare, IL, USA) with 20 mM Tris-HCl, pH 7.4 buffer containing 0.2 mM EDTA and 0.5 mM phenylmethylsulphonyl fluoride, followed by SEC on Superdex 200 Increase 10/300GL (GE Healthcare) with SEC buffer (20 mM Tris-HCl buffer, pH 7.4, containing 150 mM NaCl and 1 mM CaCl_2) at 4 °C. A sample from each purification step was resolved by SDS-PAGE (Supplementary Fig. S4a–c).

To prepare oxidized or reduced ER-60, the purified samples were incubated in 0.1 mM GSSG or 10 mM DTT in SEC buffer for 2 h at 4 °C. Then, SDS-PAGE was performed to evaluate the oxidized or reduced state of the active centre cysteines in the ER-60 samples. The ER-60 solutions were incubated in 80 mM Tris/HCl buffer at pH 6.8, with or without 25 mM 4-acetamido-4-maleimidylstilbene-2,2-disulphonate (AMS) at 4 °C for 30 min.

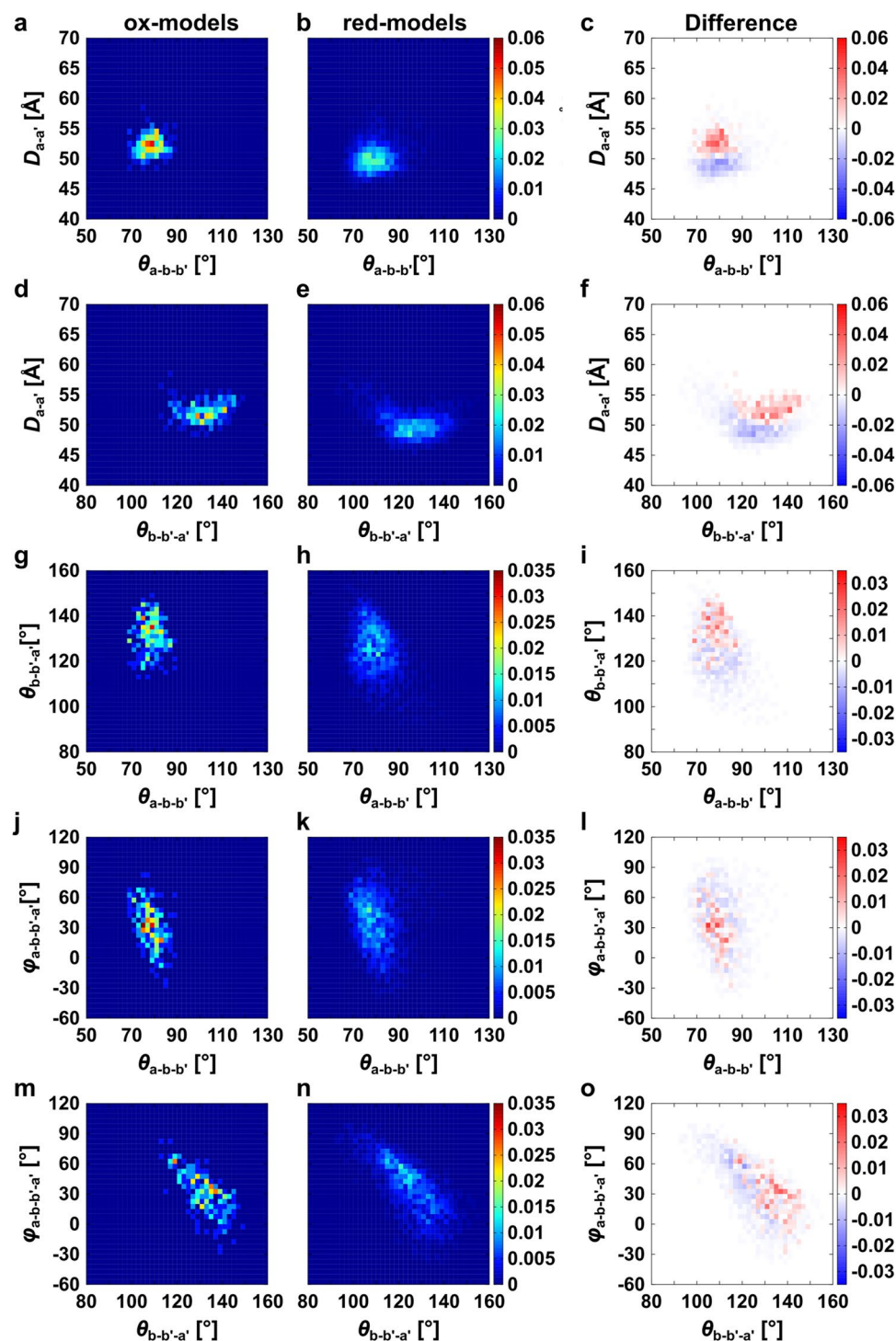


Figure 5. Two-dimensional mapping of the model structures around the **a** and **a'** domains. Distributions of correlations between two parameters, $\theta_{a-b-b'}$ — $D_{a-a'}$ (**a**, **b**), $\theta_{b-b'-a'}$ — $D_{a-a'}$ (**d**, **e**), $\theta_{a-b-b'}$ — $\theta_{b-b'-a'}$ (**g**, **h**), $\theta_{a-b-b'}$ — $\varphi_{a-b-b'-a'}$ (**j**, **k**), and $\theta_{b-b'-a'}$ — $\varphi_{a-b-b'-a'}$ (**m**, **n**) of the ox- and red-models. Panels (**c**, **f**, **i**, **l**, **o**) show the subtraction of the frequency in the red-models from that in the ox-models.

Then, they were diluted with $2 \times$ Laemmli SDS sample buffer²⁷ without reducing agent and incubated at 37 °C for 2 h. The samples were resolved by SDS-PAGE (Supplementary Fig. S4d). Free thiols of oxidized and reduced ER-60 were modified with three and seven AMS molecules, respectively, indicating that the dithiols of the active cysteine pair were fully oxidized and reduced as described previously²⁶.

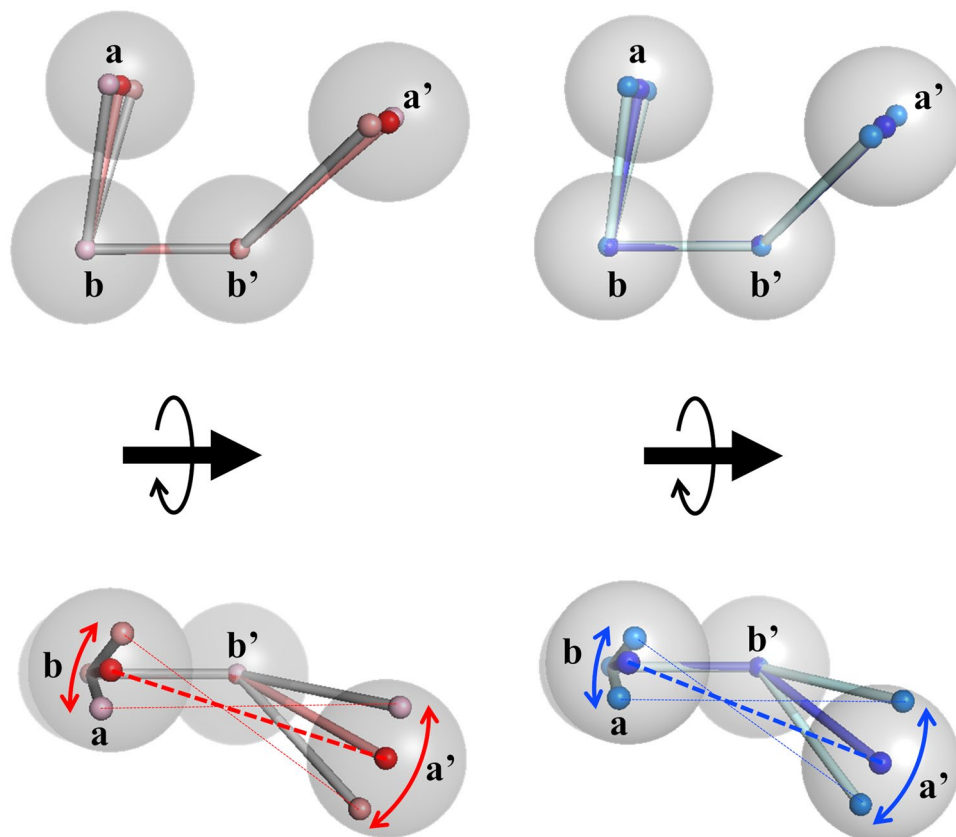


Figure 6. Conformational flexibility between the **a** and **a'** domains of ER-60. The conformational distribution of the domains is shown with red ball-and-stick models and with blue ball-and-stick models for ox- and red-models, respectively. The ox- and red-models are aligned by the **b** and **b'** domains. Gray balls roughly indicate domains as an eye guide. Coloured arrows indicate the expected ranges of flexibility in the **a** and **a'** domains.

SEC procedure in SEC-SAXS. To reduce contamination and aggregates in the samples for SAXS measurements, SEC-SAXS was applied to the samples. In this measurement, a stopping mode in a laboratory-based SEC-SAXS system was adapted to increase the statistics of scattering intensity (see the “SAXS” section)²².

For SEC, Superdex 200 Increase 10/300GL was used as the separation column. The buffers that were used for the preparation of oxidized and reduced samples were flowed through the column at 25 °C. The initial injection volume and concentration of both samples were 500 μL and 5.0 mg mL^{-1} , respectively, and the flow rate was 0.02 mL min^{-1} . The eluted solution at the elution peak of the monomer was divided into two portions (shown in Table 1 and Supplementary Fig. S2). One portion was immediately provided for the SAXS measurement, and the other was used for its quality check.

Sample quality check. We examined the quality of the samples before the SAXS measurements.

SDS-PAGE. The first examination was SDS-PAGE. The samples were diluted with $2\times$ Laemmli SDS sample buffer with or without reducing agent (2-mercaptoethanol) and incubated at 37 °C for 2 h. The samples were resolved by SDS-PAGE (Supplementary Fig. S4e). There were no bands other than the main ones, demonstrating that the samples were neither truncated nor degraded.

Analytical ultracentrifugation 1. Next, AUC was conducted to determine the abundance of monomers, dimers, trimers, other aggregations, and contaminations in the sample solution. The AUC experiments were performed at 60,000 rpm at 25 °C with a sedimentation velocity method using Rayleigh interference optics (ProteomeLab XL-I, Beckman Coulter). The weight concentration distribution $c(s_{20,w})$ as a function of the sedimentation coefficient was obtained by fitting the time evolution of sedimentation data with Lamm formula²⁸ using SEDFIT software²⁹. The sedimentation coefficient was also normalized to the value at 20 °C in pure water ($s_{20,w}$). The molecular weight for each component was calculated using $s_{20,w}$ and the friction ratio f/f_0 .

Supplementary Fig. S7a shows the weight concentration distribution $c(s_{20,w})$. There was one major peak around $s = 3.5$, and two minor peaks around $s = 5.2$ and $s = 7.0$ in the AUC profiles of both oxidized and reduced samples. This means that the sample solutions were highly purified, but there were very small aggregates ($s = 5.2$ and $s = 7.0$). The detailed values are listed in Supplementary Table S1.

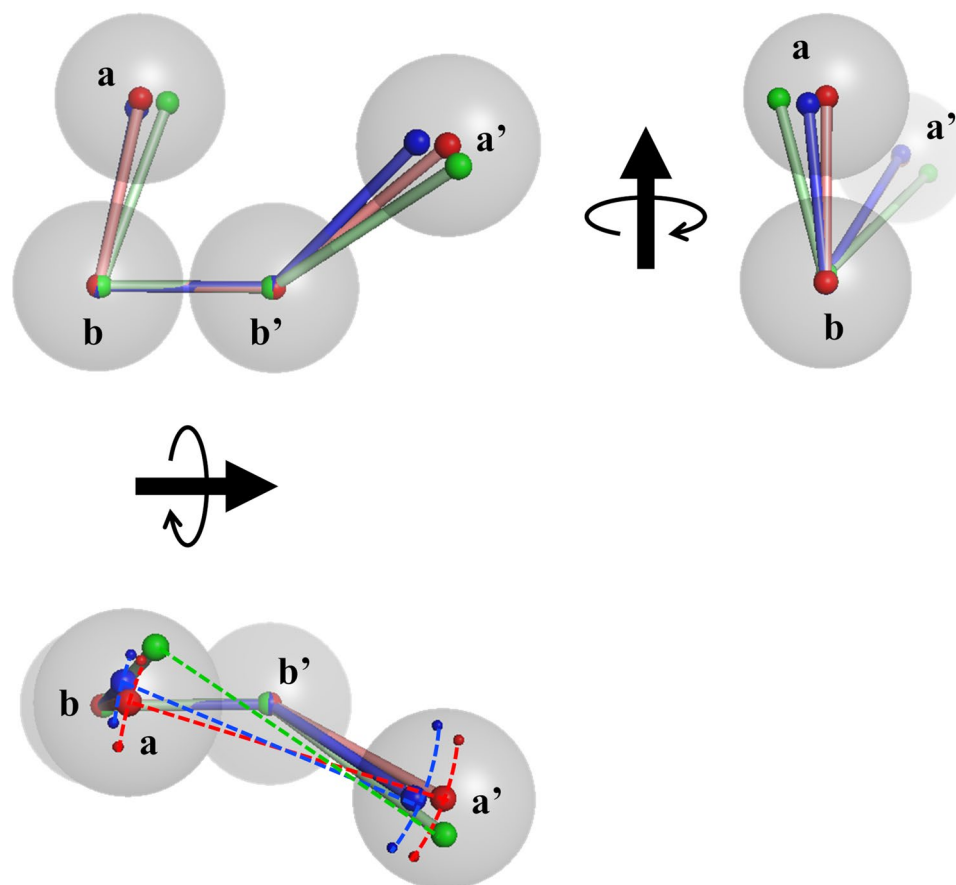


Figure 7. Comparison of the ball-and-stick models for ox- and red-models with that for the crystal structure. The averaged position of each domain is shown with red ball-and-stick models and with blue ball-and-stick models for ox- and red-models, respectively. The ox- and red-models and the crystal structure are aligned by the **b** and **b'** domains. The corresponding positions of the crystal structure are shown in the green ball-and-stick model. Gray balls roughly indicate domains as an eye guide. Dashed lines between smaller balls indicate the ranges of flexibility in the **a** and **a'** domains.

MALDI-TOF mass spectrometry 1. MALDI-TOF MS was used to confirm the aggregation numbers of the above-mentioned aggregates in the sample solutions. The samples and a saturated solution of sinapinic acid (Bruker Daltonics, Bremen, Germany) in TA30 (30% acetonitrile in 70% of 0.1% TFA in water) were mixed at a ratio 1:9 and dropped on the ground steel MALDI plate (Bruker Daltonics). After drying and crystallisation of the droplets, the measurement was conducted on a microflexLT MALDI-TOF mass spectrometer (Bruker Daltonics) in positive ion mode. External calibrations were performed using protein standards II. Mass spectra data were recorded with flexControl and analysed using FlexAnalysis (Bruker Daltonics).

Supplementary Fig. S8a shows the mass spectra, and details of the detected peaks are shown in Supplementary Table S2. There were three components in the oxidized and reduced sample solutions whose molecular weights corresponded well to the monomer, dimer, and trimer of ER-60. Therefore, the three components observed in the AUC measurements were identified as the monomer, dimer, and trimer in the order of the *s*-value.

LC-ESI-TOF mass spectrometry. As the final sample check, the average mass values of the samples were measured rigorously using LC-ESI-TOF mass spectrometry to determine whether ER-60 samples were truncated or degraded. One hundred-fifty mM ammonium acetate (pH 6.8) was flowed at $2 \mu\text{L min}^{-1}$ through a capillary column (0.3 mm, $\varphi \times 550$ mm, Chemco Scientific, Osaka, Japan) packed with resin (TSK G3000 SW gel bed, Tosoh Corporation, Tokyo, Japan). Two microliters of ER-60 solutions were loaded into the column by a sample injector (Rheodyne, Cotati, CA, USA) with a $2 \mu\text{L}$ sample loop. The samples were directly analysed using a microTOF II mass spectrometer with an electrospray ionisation (ESI) source (Bruker Daltonics). The pressure of the nebuliser gas, temperature of the dry heater, flow of the dry gas, and capillary voltage were set at 0.5 bar, 250°C , 3 L min^{-1} , and 4000 V, respectively. The mass spectra were obtained in the positive ion mode with a scanning range of m/z 500–10,000. ESI-L Low Concentration Tuning Mix (Agilent Technologies, Santa Clara, USA) was used for internal calibration. Mass spectrum data were analysed using Bruker Compass Data Analysis 4.2.

Supplementary Fig. S3 shows the mass spectra, indicating that the measured average mass values of the oxidized and reduced samples (54,259.3 and 54,264.0, respectively) were in agreement with those calculated

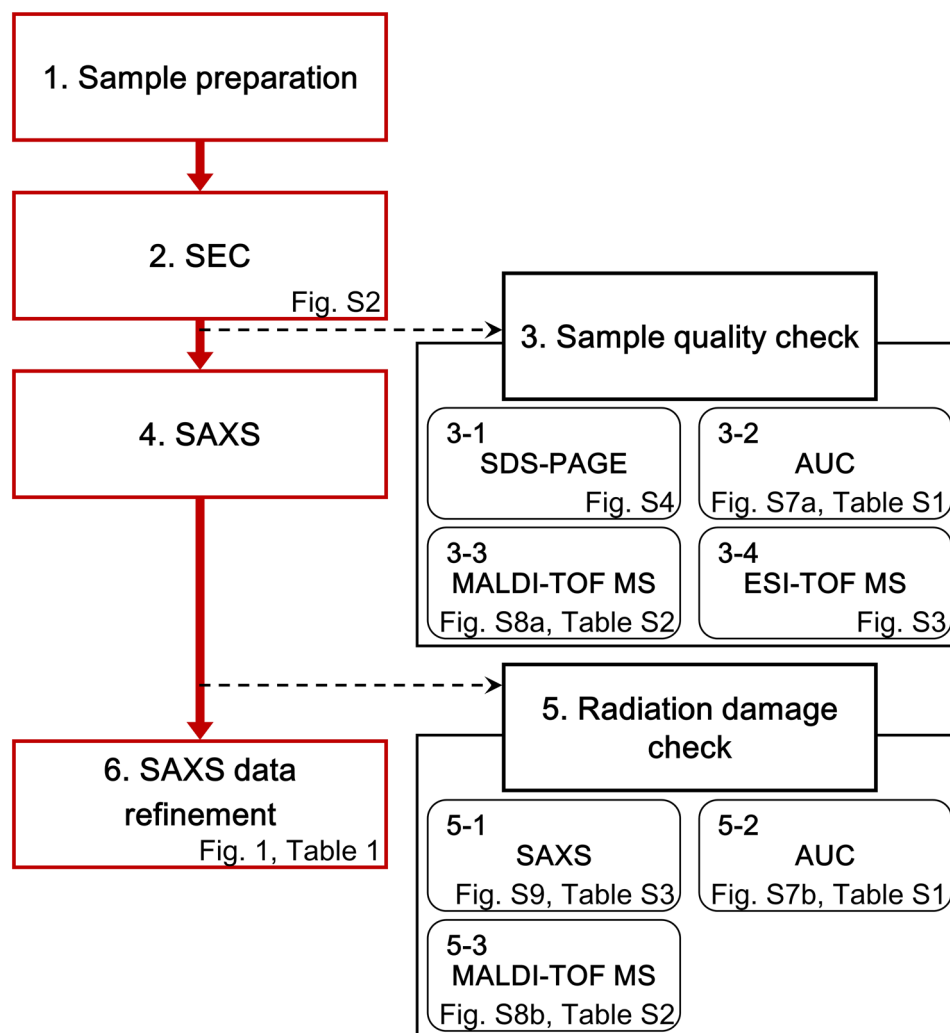


Figure 8. Flowchart of the work steps for aggregation-free SAXS measurement of ER-60. To obtain aggregation-free SAXS data, several methods were assembled. The procedure consists of sample preparation, purification with SEC, quality checks, SAXS measurement, radiation damage check, and data refinement.

from their sequences (54,260.5 and 54,264.6, respectively). The values are summarized in Table 1. This clearly indicated that both samples were neither truncated nor degraded.

SAXS. SAXS measurements were performed using a NANOPIX (Rigaku, Tokyo, Japan). X-rays from a high-brilliance point-focused X-ray generator (MicroMAX-007HF, Rigaku, Tokyo, Japan) were focused with a confoal mirror (OptiSAXS) and collimated with the lower parasitic scattering slit system, “ClearPinhole”. The scattered X-rays were detected with a two-dimensional semiconductor detector (HyPix-6000, Rigaku, Tokyo, Japan) with a spatial resolution of 100 μm . Two sample-to-detector distance (SDD) conditions, 1330 mm and 300 mm, were used to cover a wide q range (0.01–0.70 \AA^{-1}), where, $q = (4\pi/\lambda)\sin(\theta/2)$, with $\lambda (= 1.54\text{\AA})$ and θ representing the X-ray wavelength and scattering angle, respectively.

In the lower- q range measurement (0.01–0.07 \AA^{-1} , SDD = 1330.0 mm), the SEC-SAXS measurement with “stopping mode”²² was employed to remove the unfavourable aggregates from the sample solution. Then, the scattering data were collected for 1120 min (10 min \times 112). In the higher- q range measurement (0.04–0.70 \AA^{-1} , SDD = 300.0 mm), standard SAXS measurements were applied. The scattering data were collected for 500 min (10 min \times 50).

The obtained two-dimensional scattering patterns were converted to one-dimensional scattering profiles by radial averaging. Subsequently, one-dimensional scattering profiles were corrected for the intensity of the incident beam and the transmission. Then, the scattering profile of a protein in solution was obtained by subtracting that of the buffer. Finally, the unit of the scatter intensity was converted to an absolute scale by comparison with the scattering intensity of water ($I_{\text{water}} = 1.632 \times 10^{-2} \text{ cm}^{-1}$). All reductions were processed with SAngler³⁰. The low- q and high- q profiles were connected between $q = 0.061 \text{ \AA}^{-1}$ and 0.064 \AA^{-1} .

Radiation damage check. The SAXS measurements with the stopping and normal modes took a long time. Therefore, we examined the radiation damage on the samples with the SAXS data in the first and last 100 min and measured the amount of the remaining aggregates in the samples immediately after the SAXS measurements utilizing AUC and MALDI-TOF mass spectrometry.

SAXS profiles in the first and last 100 min. Supplementary Fig. S9 shows the Guinier plots of the time-averaged SAXS profiles of the oxidized and reduced samples in the first and last 100 min of the experiments. There was no deviation between the Guinier plots of the time-averaged SAXS profiles for both samples. In addition, there was no difference in the calculated R_g values, as listed in Supplementary Table S3. These values were different from those shown in Table 1 because these were calculated from SAXS profiles before AUC-SAS treatment. These data suggested that further aggregation could not occur during the SAXS measurements.

Analytical ultracentrifugation 2. Using the same procedure as Analytical ultracentrifugation 1, we identified the abundance of monomers, dimers, trimers, other aggregates, and contaminants in the sample solutions immediately after the SAXS measurements. Supplementary Fig. S7b shows the weight concentration distribution $c(s_{20,w})$ after the SAXS measurement. The amount of aggregates did not change before and after SAXS, indicating that there were no further aggregates due to radiation damage. This is also clear from the peak positions and areas on $c(s_{20,w})$ listed in Supplementary Table S1.

MALDI-TOF mass spectrometry 2. Using the same procedure as that used for MALDI-TOF MS 1, we confirmed the aggregation numbers of the above-mentioned aggregates in the sample solutions immediately after the SAXS measurements. Supplementary Fig. S9b also shows the mass spectra of the sample solutions immediately after the SAXS measurements. As expected, there were no changes in the peak positions, indicating that there were no additional aggregates except for those observed in the solution before the SAXS measurements.

SAXS data refinement. As indicated in Supplementary Table S1, whereas the major component (96 wt%) in both sample solutions was our targeting ER-60 monomer, there was still an unremovable small amount of aggregates (4%), mainly dimers and trimers. For data refinement, we conducted the AUC-SAS method to eliminate the effect of these aggregates and make the scattering data precise. AUC-SAS extracts the scattering profile of a target component from a polydispersed system²⁰. Supplementary Fig. S10 shows the SAXS data refinement by the AUC-SAS method with the observed AUC data. The gyration radii were refined from 33.4 ± 0.2 Å and 32.0 ± 0.2 Å to 32.1 ± 0.2 Å and 30.8 ± 0.2 Å for oxidized and reduced ER-60, respectively.

Finally, we obtained aggregation-free SAXS profiles for detailed structural analysis.

Coarse-grained molecular dynamics (CG-MD) simulation. We used the atomic-interaction-based coarse-grained model (AICG2+), which is the so-called “structure-based model”³¹. In this model, each amino acid is represented by a single particle at the C_α atom position. The reference structure for the structure-based potential was prepared by replacing the mutated Ala⁶⁰ of the crystal structure of ER-60 (3F8U) with Cys using the Phyre2 modelling server¹⁷. The C-terminal region of ER-60 was modelled using PyMOL³². In order to make the structural modulation, we also added both Debye-Hückel-type electrostatic interactions with an ionic strength of 150 mM and hydrophobic interactions to the standard AICG2+ potential³³.

We regarded the three peptides as flexible loops: Gly¹³³ and Leu³⁶⁵, which are located between the **a–b** and **b'–a'** domains, respectively, and the region from Gln⁴⁹⁰ to Leu⁵⁰⁵, which is a flexible C-terminal region (Supplementary Figs. S1 and S11). The flexible loop is determined based on following two points: (1) the reported crystal structures and solution NMR data (2ALB, 2DMM), most of inter-domain linkers are parts of folded regions. (2) The observed SAXS profiles can be reproduced without denaturing them from folded domains. First, we did not apply any structure-based potential for these loops to make them flexible. Second, we assigned +1 charges to six Lys particles and –1 charges to one Asp and three Glu particles in the flexible loops, respectively. Third, we treated the other regions as folded; the intra-domain native-contact potential was included. It should be noted that the **b–b'** domains were regarded as one domain. For these folded regions, the charge of each amino acid was determined based on the surface charges of the reference structure³⁴.

We performed 40 individual Langevin dynamics simulations for 10^8 MD steps at 300 K using the simulation software CafeMol³⁵. The simulation snapshots were collected every 5000 MD steps and a total of 800,000 snapshots were recorded.

Data availability

The relevant data generated and analysed in this study are available from the corresponding authors upon appropriate request.

Code availability

All software used in this study is publicly available at the URLs below. SAngler 2.0.0: <http://pfwww.kek.jp/saxs/SAngler.html>. SEDFIT 15.01c: <http://www.analyticalultracentrifugation.com/sedfit.htm>. Pepsi-SAXS-NMA: <https://files.inria.fr/NanoDFiles/Website/Software/Pepsi-SAXS/MacOS/Pepsi-SAXS-NMA>. GNOM: <https://www.embl-hamburg.de/biosaxs/gnom.html>. The Phyre2 modelling server: <http://www.sbg.bio.ic.ac.uk/~phyre/>. The PyMOL Molecular Graphics System, Version 1.8: <https://pymol.org/2/>. CafeMol: <https://www.cafemol.org/>. BBQ: <http://biocomp.chem.uw.edu.pl/tools/bbq>. SCWRL4: <http://dunbrack.fccc.edu/SCWRL3.php/#availability>.

Received: 18 December 2020; Accepted: 25 February 2021

Published online: 11 March 2021

References

1. Apic, G., Gough, J. & Teichmann, S. A. Domain combinations in archaeal, eubacterial and eukaryotic proteomes. *J. Mol. Biol.* **310**, 311–325 (2001).
2. Orengo, C. A. & Thornton, J. M. Protein families and their evolution—A structural perspective. *Annu. Rev. Biochem.* **74**, 867–900 (2005).
3. Mann, M. & Jensen, O. N. Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* **21**, 255–261 (2003).
4. Melero, J. A. & Smith, A. E. Possible transcriptional control of three polypeptides which accumulate in a temperature-sensitive mammalian cell line. *Nature* **272**, 725–727 (1978).
5. Lee, A. S. The accumulation of three specific proteins related to glucose-regulated proteins in a temperature-sensitive hamster mutant cell line K12. *J. Cell Physiol.* **106**, 119–125 (1981).
6. Bennett, C. F., Balcarek, J. M., Varrichio, A. & Crooke, S. T. Molecular cloning and complete amino-acid sequence of form-I phosphoinositide-specific phospholipase C. *Nature* **334**, 268–270 (1988).
7. Martin, J. L. *et al.* A metabolite of halothane covalently binds to an endoplasmic reticulum protein that is highly homologous to phosphatidylinositol-specific phospholipase C- α but has no activity. *Biochem. Biophys. Res. Commun.* **178**, 679–685 (1991).
8. Srivastava, S. P., Chen, N., Liu, Y. & Holtzman, J. L. Purification and characterization of a new isozyme of thiol:protein-disulfide oxidoreductase from rat hepatic microsomes. *J. Biol. Chem.* **266**, 20337–20344 (1991).
9. Urade, R. *et al.* Protein degradation by the phosphoinositide-specific phospholipase C- α family from rat liver endoplasmic reticulum. *J. Biol. Chem.* **267**, 15152–15159 (1992).
10. Srivastava, S. P., Fuchs, J. A. & Holtzman, J. L. The reported cDNA sequence for phospholipase C α encodes protein disulfide isomerase, isozyme Q-2 and not phospholipase-C. *Biochem. Biophys. Res. Commun.* **193**, 971–978 (1993).
11. Mazzarella, R. A. *et al.* Erp61 is GRP58, a stress-inducible luminal endoplasmic reticulum protein, but is devoid of phosphatidylinositol-specific phospholipase C activity. *Arch. Biochem. Biophys.* **308**, 454–460 (1994).
12. Hirano, N. *et al.* Molecular cloning and characterization of a cDNA for bovine phospholipase C- α : proposal of redesignation of phospholipase C- α . *Biochem. Biophys. Res. Commun.* **204**, 375–382 (1994).
13. Hirano, N. *et al.* Molecular cloning of the human glucose-regulated protein ERp57/GRP58, a thiol-dependent reductase. Identification of its secretory form and inducible expression by the oncogenic transformation. *Eur. J. Biochem.* **234**, 336–342 (1995).
14. Kozlov, G., Määttä, P., Thomas, D. Y. & Gehring, K. A structural overview of the PDI family of proteins. *FEBS J.* **277**, 3924–3936 (2010).
15. Ellgaard, L. & Ruddock, L. W. The human protein disulphide isomerase family: Substrate interactions and functional properties. *EMBO Rep.* **6**, 28–32 (2005).
16. Lu, J. & Holmgren, A. The thioredoxin superfamily in oxidative protein folding. *Antioxid. Redox Signal.* **21**, 457–470 (2014).
17. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
18. Dong, G. *et al.* Insights into MHC class I peptide loading from the structure of the tapasin-ERp57 thiol oxidoreductase heterodimer. *Immunity* **30**, 21–32 (2009).
19. Jeffries, C. M. *et al.* Preparing monodisperse macromolecular samples for successful biological small-angle X-ray and neutron scattering experiments. *Nat. Protoc.* **11**, 2122–2153 (2016).
20. David, G. & Pérez, J. Combined sampler robot and high-performance liquid chromatography: A fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. *J. Appl. Cryst.* **42**, 892–900 (2009).
21. Morishima, K. *et al.* Integral approach to biomacromolecular structure by analytical-ultracentrifugation and small-angle scattering. *Commun. Biol.* **3**, 294 (2020).
22. Inoue, R. *et al.* Newly developed Laboratory-based Size exclusion chromatography Small-angle X-ray scattering System (La-SSS). *Sci. Rep.* **9**, 12610 (2019).
23. Kozlov, G. *et al.* Crystal structure of the bb' domains of the protein disulfide isomerase ERp57. *Structure* **14**, 1331–1339 (2006).
24. Takada, S. *et al.* Modeling structural dynamics of biomolecular complexes by coarse-grained molecular simulations. *Acc. Chem. Res.* **48**, 3026–3035 (2015).
25. Shimizu, M. *et al.* Near-atomic structural model for bacterial DNA replication initiation complex and its functional insights. *Proc. Natl. Acad. Sci. USA* **113**, E8021–E8030 (2016).
26. Urade, R. *et al.* ER-60 domains responsible for interaction with calnexin and calreticulin. *Biochemistry* **43**, 8858–8868 (2004).
27. Laemmli, U. Cleavage of structural proteins during assembly of head of bacteriophage-T4. *Nature* **227**, 680–685 (1970).
28. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.* **78**, 1606–1619 (2000).
29. Sawicki, M. SEDfit: Software for spectral energy distribution fitting of photometric data. *Publ. Astron. Soc. Pac.* **124**, 1208–1218 (2012).
30. Shimizu, N. *et al.* Software development for analysis of small-angle X-ray scattering data. *AIP Conf. Proc.* **1741**, 50017 (2016).
31. Li, W., Wang, W. & Takada, S. Energy landscape views for interplays among folding, binding, and allostery of calmodulin domains. *Proc. Natl. Acad. Sci. USA* **111**, 10550–10555 (2014).
32. The PyMOL Molecular Graphics System, Version 1.8, Schrödinger, LLC.
33. Tanaka, T., Hori, N. & Takada, S. How co-translational folding of multi-domain protein is affected by elongation schedule: Molecular simulations. *PLoS. Comput. Biol.* **11**, e1004356 (2015).
34. Terakawa, T. & Takada, S. RESPAC: Method to determine partial charges in coarse-grained protein model and its application to DNA-binding proteins. *J. Chem. Theory. Comput.* **10**, 711–721 (2014).
35. Kenzaki, H. *et al.* Cafemol: A coarse-grained biomolecular simulator for simulating proteins at work. *J. Chem. Theory. Comput.* **7**, 1979–1989 (2011).
36. Semenyuk, A. V. & Svergun, D. I. GNOM – a program package for small-angle scattering data processing. *J. Appl. Cryst.* **24**, 537–540 (1991).

Acknowledgements

This study was supported by MEXT/JSPS KAKENHI Grants (JP20K22629 to M. Shimizu; JP19K16088 to K. M.; JP17K07361, JP19KK0071, and JP20K06579 to R. I.; JP17K07816 to N. S.; JP18H05229, JP18H05534, and JP18H03681 to M. Sugiyama). and the Sasakawa Scientific Research Grant from The Japan Science Society assigned to A. O. The study was also partially supported by a project for the construction of the basis for advanced materials science and analytical study by the innovative use of quantum beams and nuclear sciences at the Institute for Integrated Radiation and Nuclear Science, Kyoto University (KURNS) and a Grant for research promotion in KURNS to A. O., M. Shimizu, and K. M. SAXS measurements using the SAXS at KURNS were

performed under proposal No. R2121 to N. S. We thank Prof. Shoji Takada at Kyoto University for providing the computational resources.

Author contributions

A.O. and R.U. assembled and performed the sample preparation. A.O., K.M., R.I., and N.S. performed SAXS measurements and analysed the SAXS profiles. A.O. and K.M. performed the AUC measurements and analysed the profiles. A.O. performed MALDI-TOF MS and LC-ESI-TOF MS and analysed the data. M. Shimizu performed CG-MD simulations and analysed the simulation data. R.U. and M.S. designed the research, and all authors wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-85219-0>.

Correspondence and requests for materials should be addressed to R.U. or M.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021