Overall structure of fully assembled cyanobacterial KaiABC circadian clock complex by an integrated experimental-computational approach

Yasuhiro Yunoki1,2,8,10, Atsushi Matsumoto3,10, Ken Morishima4, Anne Martel5, Lionel Porcar5, Nobuhiro Sato4, Rina Yogo1,2,9, Taiki Tominaga6, Rintaro Inoue4, Maho Yagi-Utsumi1,2, Aya Okuda4, Masahiro Shimizu4, Reiko Urade4, Kazuki Terauchi7, Hidetoshi Kono3✉, Hirokazu Yagi2✉, Koichi Kato1,2✉ & Masaaki Sugiyama4✉

In the cyanobacterial circadian clock system, KaiA, KaiB and KaiC periodically assemble into a large complex. Here we determined the overall structure of their fully assembled complex by integrating experimental and computational approaches. Small-angle X-ray and inverse contrast matching small-angle neutron scatterings coupled with size-exclusion chromatography provided constraints to highlight the spatial arrangements of the N-terminal domains of KaiA, which were not resolved in the previous structural analyses. Computationally built 20 million structural models of the complex were screened out utilizing the constrains and then subjected to molecular dynamics simulations to examine their stabilities. The final model suggests that, despite large fluctuation of the KaiA N-terminal domains, their preferential positionings mask the hydrophobic surface of the KaiA C-terminal domains, hindering additional KaiA-KaiC interactions. Thus, our integrative approach provides a useful tool to resolve large complex structures harboring dynamically fluctuating domains.

https://doi.org/10.1038/s42003-022-03143-z

In the cyanobacterial circadian clock system, KaiA, KaiB and KaiC periodically assemble into a large complex. Here we determined the overall structure of their fully assembled complex by integrating experimental and computational approaches. Small-angle X-ray and inverse contrast matching small-angle neutron scatterings coupled with size-exclusion chromatography provided constraints to highlight the spatial arrangements of the N-terminal domains of KaiA, which were not resolved in the previous structural analyses. Computationally built 20 million structural models of the complex were screened out utilizing the constrains and then subjected to molecular dynamics simulations to examine their stabilities. The final model suggests that, despite large fluctuation of the KaiA N-terminal domains, their preferential positionings mask the hydrophobic surface of the KaiA C-terminal domains, hindering additional KaiA-KaiC interactions. Thus, our integrative approach provides a useful tool to resolve large complex structures harboring dynamically fluctuating domains.

1 Exploratory Research Center on Life and Living Systems (ExCELLS) and Institute for Molecular Science (IMS), National Institutes of Natural Sciences, 5-1 Higashiyama, Myodai-cho, Okazaki 444-8787, Japan. 2 Graduate School of Pharmaceutical Sciences, Nagoya City University, 3-1 Tanabe-dori, Mizhoku, Nagoya 467-8603, Japan. 3 Institute for Quantum Life Science, National Institutes for Quantum Science and Technology (QST), Umemidai, Kizu, Kyoto 619-0215, Japan. 4 Institute for Integrated Radiation and Nuclear Science, Kyoto University, 2-1010 Asashironishi, Kumatori, Sennan-gun, Osaka 590-0494, Japan. 5 Institut Laue-Langevin, 71, avenue des martyrs, 38042 Grenoble, France. 6 Neutron Science and Technology Center, Comprehensive Research Organization for Science and Society (CROSS), Tokai, Ibaraki 319-1106, Japan. 7 Graduate School of Life Sciences, Ritsumeikan University, 1-1-1 Noji-higashi, Kusatsu, Shiga 525-8577, Japan. 8 Present address: Institute for Integrated Radiation and Nuclear Science, Kyoto University, 2-1010 Asashironishi, Kumatori, Sennan-gun, Osaka 590-0494, Japan. 9 Present address: Biomedical Research Centre, School of Biomedical Engineering, The University of British Columbia, 2222 Health Sciences Mall, Vancouver, BC V6T 1Z3, Canada. 10 These authors contributed equally: Yasuhiro Yunoki, Atsushi Matsumoto.✉ Email: kono.hidetoshi@qst.go.jp; hyagi@phar.nagoya-cu.ac.jp; kkatonmr@ims.ac.jp; sugiyama.masaaki.5n@kyoto-u.ac.jp

COMMUNICATIONS BIOLOGY | (2022) 5:184 | https://doi.org/10.1038/s42003-022-03143-z | www.nature.com/commsbio
Homoeostatic activities of biological systems are regulated through dynamically concerted assembly and disassembly of biomolecules. This is best exemplified by the circadian clock in cyanobacteria (Kai-clock), which is constituted of three proteins, KaiA, KaiB, and KaiC. These proteins undergo an association–dissociation cycle coupled with phosphorylation–dephosphorylation oscillation of KaiC in the presence of adenosine triphosphate (ATP). During the circadian cycle, the Kai-clock system generates three forms of complex, two binary KaiAC and KaiBC complexes and one ternary KaiABC complex, at specific clock phases.

KaiA consists of N-terminal domain (residues 1–161, referred to as NΔ), canonical linker (residues 162–181), and C-terminal domain (residues 182–284, referred to as CΔ) and forms a homodimer (A2) through CΔ and a double doughnut-like shape through the NΔ domain and forms a homotetramer (B4)12,13. KaiC consists of two domains (CI and CII) and forms a homohexamer (C6) with a double doughnut-like shape4–16. The KaiC hexamer can interact with one A2 dimer through the C-terminal tails of CI domains, giving rise to the C6A2 complex17–19. The KaiC hexamer can also bind six KaiB molecules, which are arranged in a hexameric ring on the top of the CI domains of C6, forming the B6C6 complex20,21. Regarding the ternary KaiA-KaiB-KaiC complex (ABC complex), not only supramolecular architecture but also stoichiometry in the complex have been controversial for a long time22–24. Recently, a cryo-EM study revealed the structure of ABC complex (more precisely, A12B6C6 complex) in which six A2 dimers were captured onto the KaiB ring in the B6C6 subcomplex25,26. In that structure, each A2 interacts with one KaiB protomer through one of two dimerized CΔ domains. Hereafter, CΔ domains will be termed C1Δ and C2Δ depending if it is bound or unbound to KaiB, respectively. In total, 6 C1Δ and 6 C2Δ domains were visualized as a ring structure like an Elizabethan collar on the top of B6C6 subcomplex. In contrast, the NΔ domains were missing in the cryo-EM structure, suggesting that the NΔ domains could dynamically fluctuate in the A12B6C6 complex. Accordingly, the overall structure of A12B6C6 complex still remained to be completely solved.

The ABC complex is critical in switching from positive to negative feedback in both KaiC phosphorylation and complex formation cycles20,27,28. In addition, previously reported mutational studies indicated that the NΔ domain is essential for generation of circadian rhythm29,30. Therefore, it is crucially important to solve the overall structure of A12B6C6 complex, including the locations of NΔ domains.

Small-angle X-ray and neutron scattering (SAXS and SANS) provide overall structural information of supramolecular complex in solution31–33 and can potentially be used for investigating a dynamic structure by combination with computational analysis34–36. However, there are three issues to be addressed in elucidating the structure of A12B6C6 complex with small-angle scattering. The first is how to eliminate contributions to scattering from undesirable components. The A12B6C6 complex stably exists in the forward scattering intensity, χ, as indicated in the last column (contribution ratio of Supplementary Table 2, it is 100% D2O buffer, we selectively observe the hA protomers in the A12B6C6 complex. However, our sample solution also included hA2, dB4, dB6dC6, and their aggregates. To overcome such problems, we recently developed a method based on the combined use of SEC-SANS with iCM-SANS (SEC-iCM-SANS)37.

The third issue is how to build a three-dimensional structural model and characterize conformational dynamics of the large complex. To address this issue, we developed a method combining computational and experimental approaches. A vast array of computational models of the A12B6C6 complex were generated based on the cryo-EM and X-ray crystallographic data and subjected to screening based on the SEC-SAXS and SEC-iCM-SANS data. Eventually, selected models were verified through molecular dynamics simulations.

By overcoming these challenges with the state-of-the-art solution scattering techniques, SEC-SAXS and SEC-iCM-SANS, in conjunction with the computational approach, the present study successfully provided information on the overall structure of A12B6C6 complex, highlighting spatial arrangements of the NΔ domains.

Results

Oligomeric state of the ABC complex. On the dephosphorylation process, KaiC hexamer interacts with six KaiBs and six KaiA dimers, thus forms the A12B6C6 complex (Supplementary Fig. 1a).25,26. We established a preparation method of the A12B6C6 complex under over-saturation conditions of A2, B4, and B6C6, thereby overcoming the instability of the ternary complex (see Materials and methods, Supplementary Figs. 1b, 2a). The sample was subsequently subjected to analytical ultracentrifugation (AUC), which confirmed that the major component (p4) was the A12B6C6 complex. To address this issue, we developed a method combining computational and experimental approaches. A vast array of computational models of the A12B6C6 complex were generated based on the cryo-EM and X-ray crystallographic data and subjected to screening based on the SEC-SAXS and SEC-iCM-SANS data. Eventually, selected models were verified through molecular dynamics simulations.

Second issue is how to selectively observe the scattering originating from the components of interest in a complex. In the previous works, the structures and dynamics of multi-domain proteins and protein complexes were investigated with SAXS and coarse-grained molecular dynamics (CG-MD) simulations40–42. However, a single SAXS profile is not enough to analyze the structure of the large A12B6C6 complex with the fluctuating NΔ domains and then it is required to edit scattering data focusing on the KaiA protomers in question. For this purpose, we applied inverse contrast-matching SANS (iCM-SANS), which enables selective observation of hydrogenated component(s) within a biomacromolecular complex consisting of hydrogenated and deuterated components by taking advantage of the isotope effect of hydrogen in neutron scattering43,44. When we measured SANS of the A12B6C6 complex consisting of hydrogenated KaiA (hA), 75%–deuterated forms of KaiB and KaiC (dB and dC) in 100% D2O buffer, we selectively observe the hA protomers in hA12dB6dC6 complex. However, our sample solution also included hA2, dB4, dB6dC6, and their aggregates. To overcome such problems, we recently developed a method based on the combined use of SEC-SANS with iCM-SANS (SEC-iCM-SANS)37.

The third issue is how to build a three-dimensional structural model and characterize conformational dynamics of the large complex. To address this issue, we developed a method combining computational and experimental approaches. A vast array of computational models of the A12B6C6 complex were generated based on the cryo-EM and X-ray crystallographic data and subjected to screening based on the SEC-SAXS and SEC-iCM-SANS data. Eventually, selected models were verified through molecular dynamics simulations.

By overcoming these challenges with the state-of-the-art solution scattering techniques, SEC-SAXS and SEC-iCM-SANS, in conjunction with the computational approach, the present study successfully provided information on the overall structure of A12B6C6 complex, highlighting spatial arrangements of the NΔ domains.
closer to the experimental one but still deviated largely ($\chi^2 = 366$). The calculated $R_g$ was 62.2 Å, still smaller than that derived from the SEC-SAXS experimental value. This discrepancy is presumably due to dislocation of the NA domains in the A12B6C6 complex in solution.

SEC-iCM-SANS of A12B6C6 complex. The iCM-SANS technique enables selective observation of hydrogenated components in a complex consisting of hydrogenated and 75%-deuterated components in 100% D2O solution (Supplementary Fig. 3a). We prepared the A12B6C6 complex with hydrogenated KaiA (hA), 75%-deuterated forms of KaiB and KaiC (dB and dC) (designated as hA12dB6dC6 complex), utilizing our established method (Supplementary Fig. 3b and see Supplementary note 3). The sample solution inevitably includes undesirable components, hA2, dB4, dB6dC6-complex and their aggregates. Consequently, we conducted SEC-iCM-SANS, which provided the scattering profile of hA protomers in the hA12dB6dC6 complex in the multi-component solution (Supplementary Fig. 3b).

Fig. 1 Scattering profiles and their Guinier plots of the A12B6C6 complex. a SAXS profiles and b their Guinier plots. Black circles show the SEC-SAXS profile and a green line does the SAXS profile calculated from the cryo-EM structure ($\chi^2 = 1195$). A cyan line expresses the SAXS profile of the overall A12B6C6 model ($\chi^2 = 366$). In this model, the missing NA domains were supplemented by superimposing six A2 dimers onto the cryo-EM structure. c SANS profiles of hA12dB6dC6 complex and d its Guinier plots. Small dots are the SAXS profiles and its Guinier plot for the reference. e SANS profiles of hA12dB6D6C6 complex and f its Guinier plots. A straight line in panel e indicates $I(Q) \sim Q^{-2}$. Red lines in panels b, d, and f show the results of the least-square fitting for the experimental data with Guinier formula. Error bars represent standard deviation of the mean.
Figure 1c, d shows a SANS profile of hA12:hB6:hC6 complex in D2O solution and its Guinier plot after SEC operation (Supplementary Fig. 2d). The SEC-iCM-SANS profile of fully hydrogenated hA12:hB6:hC6 complex and its $R_g$ of 69.9 ± 0.4 Å well agreed with those obtained by the SEC-SAXS measurement of A12B6C6 complex (dotted lines in Fig. 1c, d) as expected. In contrast, the SEC-iCM-SANS profile of hA12:dB6:dC6 complex in D2O solution (Fig. 1e and Supplementary Fig. 2e) and its Guinier plot (Fig. 1f) were drastically different from those of hA12:hB6:hC6 complex (Fig. 1c, d). The scattering profile decreased with $Q^{-2}$, indicating that a scatterer was a disk-like shape based on the classical interruption for a SAS profile. This observation suggested that the six KaiA dimers were arranged in a doughnut-like shape on top of the B6C6 subcomplex as reported in the cryo-EM study. In addition, its larger $R_g$ of 78.1 ± 1.0 Å supported this doughnut-like arrangement.

**A12B6C6 complex modeling.** We computationally built three-dimensional models of the overall structure of A12B6C6 complex, which reproduced the SAXS and iCM-SANS profiles. The models obtained were further examined and screened through MD simulations from a viewpoint of stability. Our modeling procedure is described below (Fig. 2).

**Fig. 2** Modeling procedure of the overall structure of A12B6C6 complex. Left column denotes the structural models and right one does step-by-step computational process.

**Step 1: Preparation of the initial model.** Since the cryo-EM structure (PDB:5n8y) lacks the side-chain information, we build an initial model with crystal structures providing them. There are currently three crystal structures available that can be used to build the structure model of A12B6C6 complex: a KaiC hexamer (C6, PDB code: 3dvl)16, a ternary complex (C2A-B-C2C, PDB code: 5jwr)26 consisting of two C-terminal domains of KaiA dimer (C2A), a KaiB monomer (B) and a CI domain of KaiC (C2C), and a full-length structure of KaiA dimer (A2, PDB code: 1r8j)11. Here, we outline the initial modeling procedure of A12B6C6 (for details, see Supplementary note 2). First, we placed six C2A-B-C2C to C6 by superposing the C2C domain (green) (Supplementary Fig. 4a), thereby modeling the A12B6C6 complex, referred to as Complex 1. Note that Complex 1 does not have any NA domains. Complex 1 well agreed with the previously reported cryo-EM structure (RMSD = 3.9 Å between them). Next, we placed six full-length A2 to Complex 1 by superposing the C1A and C2A domains (Supplementary Fig. 4b). Note that the N2A (blue) and N2A (red) domains are derived from one A2 dimer and connected to the C1A (cyan) and C2A (magenta) domains, respectively (inset of Supplementary Fig. 4b). Finally, we obtained an overall structure model of A12B6C6 complex, referred to as Complex 2. In this structure, each N2A domain (red) structurally overlapped with KaiB (yellow) (Supplementary Fig. 4c), indicating that the A2 dimer undergoes a conformational change in terms of the spatial arrangements of N2A domains upon formation of the A12B6C6 complex.

**Step 2: Generation of the structural candidates.** To remove the KaiA-KaiB structural overlap in Complex 2, we systematically altered the positions and orientations of individual N1A and N2A domains belonging to one A2 dimer and gave the same conformation for the remaining five A2 dimers (applying the C6 symmetry around the first axis defined by B6C6). At this stage, we ignored the linkers connecting NA and CA domains to reduce the computational cost. As a result, we obtained ~20 million models of A12B6C6 complex as initial structural candidates (Supplementary Fig. 4d and see Supplementary note 4).

**Step 3: Selection of models without linkers based on the SAXS and SANS data.** We calculated the scattering curve for each of the models for screening based on the criterion of $\chi^2_{\text{SAXS}} < 10.0$ to the experimental SAXS data. We obtained about 400,000 models from the candidates generated in Step 2. The selected models were classified into three types, i.e., Types 1, 2, and 3, based on the location of the N2A domains (Supplementary Fig. 5a). Type 1 holds both N1A (blue) and N2A (red) domains below the reference plane defined by the top plane of KaiB hexameric ring in the B6C6 subcomplex. Type 2 has one of N2A domains below the reference plane while the other upper. In Type 3, both N1A and N2A domains are located upper the reference plane. The numbers of models were 331,000, 3000, and 62,000 for Types 1, 2, and 3, respectively.

Next, we evaluated the SAXS-selected models based on the iCM-SANS data as source information on the KaiA protomer conformations in the A12B6C6 complex. We found that only Type 1 gave the small $\chi^2$ values ($\chi^2_{\text{SANS}} < 3.0$) among the three types (Supplementary Fig. 5b), therefore leaving the Type 1 models as candidates. Thus, the combining of multiple experimental data can compensate their low resolution, underscoring the importance of multilateral evaluation in structural modeling of a huge complex.

**Step 4: Linker formation.** Through step 3, we selected 29,809 models with small $\chi^2$ values ($\chi^2_{\text{SAXS}} < 6.0$ and $\chi^2_{\text{SANS}} < 1.5$) (structures within a white dotted box in Supplementary Fig. 5c) from the set of Type 1. Using the Rosetta program suite51,52, we attempted to complement these models with an NA-C A linker, which was ignored in the previous steps. Consequently, linker modeling was successful for about a quarter of the models (8608). For each of these models, about 100 multiple linker conformations were tested and the best model containing the linkers with the smallest $\chi^2$ value for the SAXS data was selected.

**Step 5: Second selection of models with linkers by SAXS and SANS.** We noticed that the linker addition affected their $\chi^2$ values for the SAXS and SANS (Supplementary Fig. 5d, e). We then re-evaluated and selected 1550 models for the overall A12B6C6 complex, which met the experimental SAXS and iCM-SANS profiles with $\chi^2_{\text{SAXS}} < 5.0$ and $\chi^2_{\text{SANS}} < 1.5$ (models within a white dotted square in Supplementary Fig. 5f).

Here we summarize the structural features of the selected models: although the NA domains occupy variable positions, we attempt to identify common features on positioning with respect to...
to the B6C6 subcomplex. For this purpose, we defined the coordinate as shown in Fig. 3a and examined the positions of the NA domains. The space is divided into cells considering the symmetry of the cryo-EM structure, Complex 1 (for details of the space division and grouping procedure, see Supplementary note 5). The positions of NA domains were classified into two distinct groups: one distributed on the upper (U) rings surrounding Complex 1, and the other one, on the lower (L) ring (Fig. 3a, b). We also found significantly preferred positions for NA domains. In one A2 dimer, when one NA domain was located at the U ring, the other one was always at the lower L ring (Supplementary Fig. 6a). In each ring, two sets of possible positions were available for the NA domain, i.e. U1, U2, L1, and L2. In the U ring, the six NA domains distribute into mutually exclusive locations, U1 (blue spheres) and U2 (green spheres), each of which follows a six-fold symmetry (Supplementary Fig. 6a, b). In the L ring, the six NA domains also distribute into mutually exclusive locations with six-fold symmetry, L1 (red spheres) and L2 (orange spheres) (Supplementary Fig. 6c). In addition, we considered the linker connections between NA and CA domains in one A2 dimer, i.e. N1A–C1A and N2A–C2A (Fig. 3c, d). Taken together, the structural models were classified into eight groups as shown in Fig. 3e–l. In Groups I, II, and III, the N1A and N2A domains in one A2 dimer are located at U1 and L1, respectively. In Group III’, the positions of N1A and N2A domains are swapped from those in Group III, i.e., N1A and N2A are at L1 and U1, respectively. In Groups IV and V, the N1A and N2A domains in one A2 dimer are at U1 and L2, respectively. In Group V’, N1A, and N2A domains are reversely arranged as compared with those in Group V. In Group VI, N1A, and N2A domains are located at U2 and L2, respectively. The A2 dimer exhibits distinct conformations among the different groups as clarified in Fig. 3e–l (surrounded by red line).

In summary, in the L ring, the NA domains tend to be located at L1 or L2 (θ ~ 200 or 220 in Fig. 3 and Supplementary Fig. 6c) with six-fold symmetry. Their positions are between C1A and C2A
domains belonging to two adjacent A_2, respectively. At the U ring, the N_A domains radially sit at the staggered positions with respect to its counterpart N_A domains in the L ring. This exclusive rule is true for all the groups except Group VI. The structural features of the eight groups are summarized in Supplementary Table 3.

Step 6: Verification by MD simulation. In the screening procedure of Steps 1–5, the protein domains were treated as rigid bodies and only the exclusive volume of the molecules was considered. Thus, we further checked structural stability of the obtained models using MD simulation. As the first quick test for stability, we randomly selected the 384 models with $\chi_{SAXS}^2 < 5.0$, $\chi_{SAXS}^2 < 1.25$ from the eight groups and performed 10-ns MD simulations. Figure 4a, b shows the relative locations of N_A domains with $\chi_{SAXS}^2$ after 10 ns. The values of $\chi_{SAXS}^2$ are expressed in colors, as green, yellow, and red dots correspond to the structures with $\chi_{SAXS}^2 < 5.0$, $5.0 \leq \chi_{SAXS}^2 < 10.0$ and $10.0 \leq \chi_{SAXS}^2$, respectively, and no model exceeds 1.5 of $\chi_{SANS}^2$ after 10 ns MD simulation. The results clearly indicate that the models belonging to Groups II, III, and III' maintained $\chi_{SAXS}^2 < 5.0$ but the model in the other groups yielded larger $\chi_{SAXS}^2$. This suggests that the L1 position is more suitable for N_A domains than L2. Moreover, Group I did not have any stable model, possibly because the stretching of the linker between G1_A and N1_A would make the structure unstable (Fig. 3e).

Considering the results above, we extended the simulation for 100 ns on 22 randomly selected models from Groups II, III', and III that maintained stability during the 10 ns MD simulations. The number of selected models is 13, 4, and 5 for Groups II, III, and III', respectively. For each model, we recorded trajectories every 20 ps and calculated the SAXS and SANS profiles of the 5000 snapshot structures. The averaged $\chi^2$ and $R_g$ fits to the SAXS and SANS profiles are shown in Fig. 4c, e are summarized in Table 1. To find the structural models reproducing SAXS and SANS over the simulation time, we marked values with asterisks in Table 1, where two asterisks, one asterisk, and no asterisk for $\chi^2$ denote $\chi^2 > 10$, $\chi^2 > 25$, and $\chi^2 < 25$ (SAXS) and $\chi^2 > 10$, $\chi^2 > 25$, and $\chi^2 < 25$ (SANS), respectively, and they for $R_g$ also denote $|\Delta R_g| > 2 \times \text{Error}$, $2 \times \text{Error} \geq |\Delta R_g| > \text{Error}$, and $|\Delta R_g| \leq \text{Error}$ ($\Delta R_g = R_{g,MD} - R_{g,EXP}$), respectively. The structures that best reproduce the averaged scattering profiles are Model II-12 in Group II and Model III-2 in Group III. We further examined the time evolutions of $\chi_{SAXS}^2$ and $\chi_{SANS}^2$ as shown in Fig. 4d, f. Model II-12 shows that $\chi_{SAXS}^2$ and $\chi_{SANS}^2$ were initially small but gradually increased (after 50 ns). On the contrary, both $\chi_{SAXS}^2$ and $\chi_{SANS}^2$ of Model III-2 remained stable for all the 100 ns. In addition, trajectories and root mean square fluctuations (RMSFs) of center of mass (COM) of C_A domains of Models II-12 and III-2 are calculated in the 100 ns MD simulations (Supplementary Fig. 7). In Model II-12, four C_A domains (C2-1, C2-3, C2-5, and C2-6) gives RMSFs of over 4.0 Å and the averaged value was also 4.0 Å. On the contrary, Model III-2 has only one C_A domain, which yielded a large fluctuation and then the averaged RMSF for the C_A domains was less than 4.0 Å (3.16 Å). This means if the fluctuation as seen in Model III-12 occurs, the structures of C_A domains would not be determined with a method like cryo-EM analysis. However, the
Although each NA domain could not be stabilized at a fixed position, we found their preferential positions in proposed models. Consequently, we successfully obtained a structural model, Model III-2, which does not only reproduce SEC-SAXS and SEC-ICM-SANS data but also remains structurally stable during the 100 ns-long MD simulation. Figure 5c–h and Supplementary Movies 1 and 2 show its structure and dynamic fluctuations.

### Table 1 Time averaged $\chi^2_{\text{SAXS}}$, $R_g$SAXS, $\chi^2_{\text{SANS}}$, and $R_g$SANS of 100 ns MD simulated models.

| Model | $\chi^2_{\text{SAXS}}$ | $R_g$SAXS | $\chi^2_{\text{SANS}}$ | $R_g$SANS |
|-------|-----------------------|-----------|-----------------------|-----------|
| II-1  | 6.6**                 | 69.0**    | 1.6*                  | 77.8      |
| II-2  | 18.3**                | 68.0**    | 1.9**                 | 77.4      |
| II-3  | 3.6                   | 69.2      | 1.7*                  | 78.0      |
| II-4  | 4.7                   | 69.1      | 1.6*                  | 77.9      |
| II-5  | 411**                 | 67.3**    | 2.1**                 | 76.0**    |
| I-6   | 8.7**                 | 68.6**    | 1.7*                  | 77.6      |
| I-7   | 4.9                   | 69.8      | 3.3**                 | 81.9**    |
| I-8   | 7.9**                 | 69.1**    | 1.6*                  | 78.7      |
| I-9   | 7.3**                 | 69.0      | 1.7*                  | 79.6*     |
| I-10  | 5.9**                 | 69.2**    | 1.7*                  | 77.6      |
| I-11  | 10.5**                | 68.3**    | 1.6*                  | 78.0      |
| I-12  | 4.7                   | 69.6      | 1.4                  | 78.7      |
| I-13  | 5.8**                 | 68.8**    | 1.8**                 | 75.5      |
| I-1-2 | 19.2**                | 67.7**    | 2.0**                 | 75.7**    |
| I-3   | 5.0                   | 69.2      | 1.5                  | 78.7      |
| I-4   | 21.3**                | 68.3**    | 1.6*                  | 77.5      |
| I-5   | 11.9**                | 68.4**    | 1.7*                  | 77.1      |
| I-6   | 5.2**                 | 69.0      | 1.6*                  | 78.4      |
| I-7   | 6.4**                 | 68.8**    | 1.8**                 | 78.2      |
| I-8   | 14.6**                | 68.6**    | 1.7*                  | 77.8      |
| I-9   | 5.8**                 | 68.9**    | 1.7*                  | 76.1**    |
| I-10  | 7.7**                 | 69.2      | 1.7*                  | 76.9*     |
| Experiment | 69.5 ± 0.3 | 78.1 ± 1.0 |

Two asterisks, one asterisk, and no asterisk for $\chi^2_{\text{SAXS}}$ denote $\chi^2_{\text{SAXS}}$ > 5.0, $\chi^2_{\text{SAXS}}$ > 4.0, and $\chi^2_{\text{SAXS}}$ ≤ 3.0, respectively. For $\chi^2_{\text{SANS}}$, one asterisk, two asterisks, and no asterisk denote $\chi^2_{\text{SANS}}$ ≤ 1.5, $\chi^2_{\text{SANS}}$ ≤ 1.6, and $\chi^2_{\text{SANS}}$ ≤ 1.7, respectively.

In summary, we considered the role of NA domains related with their positions. During the circadian cycle, the KaiC phosphorylation switches the interaction modes with KaiA. Namely, the dephosphorylated C6 does not bind KaiB but interacts with A2 through its C-terminal tails, thereby forming the A2C6 complex (Process 1 in Supplementary Fig. 1)22,56,57, which subsequently promotes direct binding of the KaiA dimer to the KaiB hexameric ring8,26, giving rise to the A12B6C6 complex (Process 3 in Supplementary Fig. 1)25, which has been characterized in this study. The KaiC dephosphorylation is accelerated in the ABC complex (Process 4 in Supplementary Fig. 1)20,27,28. In the A2C6 complex, a hydrophobic surface close to the dimeric interface of C6 domains (a black dot circle in Fig. 7a) accommodates the C-terminal tail of KaiC (a black string in Fig. 7b). This binding surface appears to be exposed to solvent in the cryo-EM structure of A12B6C6 complex (Fig. 7c). Intriguingly, this binding surface is masked by the N1A domain on the U ring in our structural model of the A12B6C6 complex (Model III-2), presumably hindering potential interaction via the KaiC C-terminal tail to form a larger complex (Fig. 7d). This is consistent with the AUC data indicating that the ABC complex did not form any complex larger than A12B6C6 complex (Supplementary Fig. 2b). Our findings provide a structural basis for the mechanism behind the precise circadian rhythm, that the formation of ABC complexes prevents the N1A domains on the U ring from additional interactions with any of the KaiC hexamers in the system, which would lead to infinite elongation of the complex.

During the 10 ns MD simulation, the N2A domains on the L ring traveled around the center between two adjacent CA domains (C1A and C2A domains, Fig. 7e). As each C1A domain bind one KaiB protomer in the B6C6 subcomplex, there exist six vacant spaces between adjacent CA domains (Fig. 3a). That is why the positioning of KaiA protomer in the complex appears unstable. However, the N2A domains of our models can be accommodated in these vacant spaces, which could stabilize the positioning of KaiA protomers.

In summary, we firstly succeeded in delineating the overall structure of A12B6C6 complex by the integration of experimental techniques such as SEC-SAXS, SEC-ICM-SANS, and AUC and randomly fluctuating, explaining why NA domains are invisible in the cryo-EM structure29.
computational modeling and simulations. The main issue to be resolved was to locate the N-terminus missing in the cryo-EM structure. For this purpose, we used two different sets of scattering data, from SEC-SAXS reflecting the overall shape and from SEC-iCM-SANS extracting the conformational information of KaiA domains for screening the structures generated by the computational modeling. We emphasize that the SEC-iCM-SANS could reject the models, Types 2 and 3, which could not be excluded on the basis of the SAXS data only (Supplementary Fig. 5a). In addition, we demonstrated that MD simulation can be used for further model selection. In fact, the SAXS profiles of Model II-7 gave a small averaged \( \chi^2 \) value and they were stable during the 100 ns MD simulation (Fig. 4c, d). However, the SANS profiles clearly showed the difference among the models with small \( \chi^2 \) (Fig. 4e, f). For example, the KaiA domains of Model II-7 were conformationally transformed from a six-fold symmetry to a pseudo-3-fold symmetry within a short time (20 ns). This latter conformation could well meet the experimental SAXS profiles but not the iCM-SANS profiles. Thus, this study demonstrates that the integrated approach of modern solution scattering methods, the SAXS and iCM-SANS techniques, and computational modeling and molecular dynamics simulation provide a powerful and generally applicable tool for resolving structures of supramolecular complexes harboring dynamically fluctuating domains/subunits like the KaiABC complex.

**Materials and methods**

**Expression and purification of hydrogenated and deuterated Kai proteins.** KaiA, KaiB, and KaiC from *Synechococcus* sp. PCC 7942 were expressed in *Escherichia coli*. KaiA was cloned into pET-28b according to the literature. KaiA was expressed as hexahistidine (his)-tagged recombinant protein and purified after the cleavage of the his-tag as described previously. KaiB was expressed as a glutathione S-transferase (GST)-tagged recombinant protein and purified after the cleavage of the GST-tag as described previously. KaiC was expressed and purified as a Strep-tagged recombinant protein as described previously. Here, we used the phosphorylation mimic KaiC in which one phosphorylation site (S431) was substituted with an aspartate residue, because of its high affinity for KaiB.

For preparation of the deuterated proteins, the bacterial cells were grown in M9 minimal media containing glucose as a mixture with varying ratios of isotopically natural and fully deuterated glucose (1,2,3,4,5,6-D7, 98%, Cambridge Isotope Laboratories, Inc.), along with varying ratios of H2O and D2O as previously described.

**Preparation for solution of A12B6C6 complex.** We established a two-steps procedure to prepare a fully assembled ABC complex, which was expected to be A12B6C6 complex. In the first step, we produced B6C6 complex by mixing of KaiB and KaiC with the ratio of 9:6 in the buffer, 20 mM Tris–HCl buffer (pH 8.0) containing 150 mM NaCl, 5 mM MgCl2, 0.5 mM EDTA, 1 mM DTT, and 1 mM ATP at 10 °C. Then, the BC complex was isolated from the mixture with SEC.
COMs of N1A, N2A, C1A, and C2A domains, respectively. Yellow dots express COMs of KaiB domains. Crosses show the initial positions in the 100 ns MD averaged RMSFs with the error bars. Overlayed images in 100 ns MD simulations for evolutionary sedimentation data with Lamm formula using SEDFIT software (http://sedfit.htm)63. The sedimentation coefficient was normalized to be the value at 20 °C in pure water, s20,w. In addition, the molecular weight for each component was calculated using the corresponding peak area in c(s20,w) and then the contribution ratio r in the forward scattering intensity of SASS was also estimated from the molecular weight65. The results are listed in Supplementary Table 2.

SEC-SAXS. SEC-SAXS experiment for the A12B6C6 complex was performed with Photon Factory BL-10C (Tsukuba, Japan) using UPLC ACQUITY (Waters Corp., Milford, MA, USA) integrated with a SAXS setup57. The wavelength of the injected X-ray and the detector were 1.50 Å and PILATUS3 2 M detector, respectively. The sample-to-detector distance were set to 3034.9 mm and then the covered Q-range was from 0.005 to 0.18 Å−1.

In the measurement, the sample solution of 250 μL of 11.7 mg/mL was loaded onto a Superose 6 increase 10/300GL (GE Healthcare, Chicago, IL, USA) pre-equilibrated with the buffer at a flow rate of 0.5 mL/min. During the elution of proteins, the flow rate was reduced to 0.05 mL/min. The SAXS and UV spectra at 280 nm were recorded every 20 and 10 s, respectively. The observed SAXS intensity was corrected for background, empty cell and buffer scatterings, and transmission factors and subsequently converted to the absolute scale using SAngler47. The unit of scattering intensity was converted to the absolute scale by referring to a standard scattering intensity of water at 293 K (1.632 × 10−7 cm−1)64.

There appeared a clear peak corresponding to A12B6C6 complex in the elution chart (Supplementary Fig. 2c). We selected appropriate time ranges (indicated with the pink zones in Supplementary Fig. 2c) and made the average of the scattering intensities over the time range, where the averaged concentration was 1.02 mg/mL. We obtained the timing-coincident integrated scattering intensities in the Q-range from 0.008 to 0.015 Å−1. (Supplementary Fig. 2c).

AUC-SAXS. AUC-SAXS experiment for A1 in solution was performed with ProteomeLab XL-1 (Beckman Coulter Inc., Brea, CA, USA) and NANOPIX (RIGAKU Co., Ltd., Tokyo, Japan). SV-AUC measurement was carried out using Rayleigh interference optics at 60,000 rpm. The optical path and the volume of the AUC cell were 12 mm and 400 μL, respectively. All measurements were performed using Rayleigh interference optics at 60,000 rpm at 30 °C. With this setting, we measured five samples (listed Supplementary Table 1), simple solutions of KaiA, KaiB, and KaiC, binary mixture solution of KaiB and KaiC, and ternary mixture solution of KaiA, KaiB, and KaiC. The first four sample were references and the last one was the sample solution including A12B6C6 complex as expected. The concentrations of the sample solutions of KaiA alone, KaiB alone, KaiC alone, KaiB-KaiC mixture, and KaiA-KaiB-KaiC mixture were 0.5, 0.5, 0.5, 0.6, and 1.0 mg/mL, respectively.

The AUC profile, the weight concentration distribution of particles in a solution c(s20,w) was obtained as a function of sedimentation coefficient by fitting the time evolution sedimentation data with Lamm formula using SEDFIT software (http://www.analyticalultracentrifugation.com/sedfit.htm)64. The sedimentation coefficient was normalized to be the value at 20 °C in pure water, s20,w. In addition, the molecular weight for each component was calculated using the corresponding peak value s20,w and the fraction ratio f/s, which was also provided in the data reduction with SEDFIT (summarized in Supplementary Table 1).

For each component included in the ternary mixture sample, the weight fraction r was calculated from the corresponding peak-area in c(s20,w) and then the contribution ratio r in the forward scattering intensity of SASS was also estimated from the molecular weight65. The results are listed in Supplementary Table 2.
was from 0.01 to 0.2 Å \(^{-1}\). We subjected 1.0 mg/mL of KaiA solution to these measurements at 25 °C. AUC-SAXS treatment was conducted to eliminate the effect of aggregates and make the scattering data precise according to the previous report\(^65\).

**SEC-iCM-SANS.** SEC-iCM-SANS experiments for hA\(_{12}\)dB\(_6\)dC\(_6\) and hA\(_{12}\)hB\(_6\)hC\(_6\) (as reference) were performed with the SEC system at D22 of the Institut Laue-Langevin (ILL), Grenoble, France. The neutron wavelength and the sample-to-detector distance were set to 6.0 Å and 5600 mm, respectively, and the covered Q-range was from 0.008 to 0.15 Å \(^{-1}\).

In the measurement, the sample solutions of 235 μL with 23.4 mg/mL (ternary mixture of hA-hB-hC) and 275 μL with 18.9 mg/mL (ternary mixture of hAhBhC) were loaded onto a Superose 6 Increase 10/300GL column (GE Healthcare, Chicago, IL, USA) with the D\(_2\)O buffer at a flow rate of 0.5 mL/min. During the elution of proteins, the flow rate was reduced to 0.07 mL/min. The SANS data were collected for every 30 s and UV absorbance at 260 nm were recorded every 1 s. The observed SANS intensity was corrected for background, empty cell and buffer scatterings, and transmission factors and subsequently converted to the absolute scale using GRASP software using incident beam flux\(^65\). There appeared clear peaks corresponding to A\(_1\), B\(_6\), C\(_6\) complexes in the elution charts (Supplementary Fig. 2d, e). We selected the time range of FWHM of the peak and made the average of the scattering intensities over the selected time range (indicated with the pink zones in Supplementary Fig. 2d, e), where the averaged concentrations were 1.34 mg/mL and 1.27 mg/mL for the hA\(_{12}\)hB\(_6\)hC\(_6\) complex and hA\(_{12}\)dB\(_6\)dC\(_6\) complexes, respectively. We obtained the timing-coincident integrated scattering intensities in the Q-range from 0.008 to 0.015 Å \(^{-1}\). (Supplementary Fig. 2d, e).

**MD simulations for model verification.** We performed MD simulations with the models that well reproduced the experimental SAXS and SANS profiles to examine whether they existed stably in solution. We performed conventional MD simulations (NVT) at a temperature of 300 K with no restraint using GROMACS\(^66\)-\(^72\) with the Amber 14SB force field\(^73\) and the TIP3P water model\(^74\). The temperature was controlled using the V-rescale method\(^75\). The Na\(^+\) and Cl\(^-\) ions were added to neutralize the system and maintain the salt concentration at 150 mM. The salt concentration was set at the same value as the scattering experiment (see Supplementary Table 4 for other details).

SAXS and iCM-SANS profiles of MD snapshot structures were calculated with CRYSOL\(^76\) and CRYSON\(^77\). The smearing effect of instrumental resolution on the dynamical cross-correlation map was considered with the resolution provided by GRASP\(^65\).

**Dynamical cross-correlation map.** The dynamical cross-correlation\(^53\) is calculated for the Ca atom pairs in the N\(_A\) domains during the 100 ns-long MD simulation. The correlation coefficient between i\(_{th}\) and j\(_{th}\) Ca atoms, whose positions are \(\mathbf{r}_i\) and \(\mathbf{r}_j\) respectively, is defined as \(c_{ij} = \frac{\langle (\mathbf{r}_i - \langle \mathbf{r}_i \rangle) \cdot (\mathbf{r}_j - \langle \mathbf{r}_j \rangle) \rangle}{\sigma_i \cdot \sigma_j}\), where \(\sigma_i = \langle (\mathbf{r}_i - \langle \mathbf{r}_i \rangle)^2 \rangle^{1/2}\) and \(\langle \rangle\) denotes the average during the simulation. Before the

---

**Fig. 7 Positioning sites of N\(_A\) domains of A\(_2\) protomers.** a Hydrophobic surface (black dot circle) close to the dimeric interface of c\(_A\) domains (cyan and magenta). **b** A complex of KaiA C-terminal domains (cyan and magenta) and disordered C-terminal segment of KaiC (black) (PDB code: 1suy). This is a part of A\(_2\)C\(_6\) complex. c Complex 1 (same as the cryo-EM structure). The hydrophobic surfaces to be exposed to solvent. d Model III-2. The hydrophobic surfaces are masked by the N\(_1\)A domains and the linkers. The coloring is as follows: N\(_1\)A (blue), N\(_2\)A (red), C\(_1\)A (cyan), C\(_2\)A (magenta), N\(_A\)-CA linker (black), KaiB (yellow), and KaiC (gray). e (upper) Number distribution map of the N\(_2\)A domain connecting the asterisked C\(_2\)A domain (see lower inset) along the L-IV as references. (Middle and Lower) The structural models of Groups II-IV as references. Colors of spheres for domains are the same as those in Fig. 3.
calculation, the atoms in each KaiA dimer are structurally aligned by the RMS-fitting so that all the $\alpha$ domains during the simulation overlap each other. In this way, the comparison of the motions of the six KaiA dimers is straightforward. The correlations between upper $\alpha$A domains and those between lower $\alpha$A domains are computed (Supplementary Fig. 8a, b).

Statistics and reproducibility. The fittings for Guinier formula to derive $I(0)$ and $R_f$ were performed with the linear least-square method. The errors were defined as the standard deviation.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
The datasets generated and analyzed during the current study are available from the corresponding authors on reasonable request. The SEC-SAXS and SEC-ICM-SANS data are deposited in SASDB under SASDINJ and SASDIN2, respectively\(^{58}\). The 108 ns MD trajectory of the representative model of the $\alpha_1\beta_3\gamma_2$ complex (II-3) are deposited in the Biological Structure Model Archive (BASM-Arc) under BSM-ID BM00030\(^{59}\).

Received: 22 October 2021; Accepted: 8 February 2022; Published online: 10 March 2022

References
1. Bonomi, M. & Vendruscolo, M. Determination of protein structural ensembles using cryo-electron microscopy. Curr. Opin. Struct. Biol. 56, 37–45 (2020).
2. Patil, N. K., Bohannon, J. K., Hernandez, A., Patil, T. K. & Sherwood, E. R. Regulation of leukocyte function by citric acid cycle intermediates. J. Leukoc. Biol. 106, 105–117 (2019).
3. Sora, V. et al. Structure and dynamics in the ATG8 family from experimental to computational techniques. Front. Cell Dev. Biol. 8, 420 (2020).
4. Ziegler, S. J., Mallinson, S. J., John, P. C. S. & Bomble, Y. J. Advances to computational techniques. Proc. Natl Acad. Sci. USA 110, 1124–1129 (2013).
5. Ishiura, M. et al. Expression of a gene cluster encoding a circadian clock protein in Synechococcus PCC 7942. Front. Cell Dev. Biol. 14, 2930–2939 (2002).
6. Mori, T. et al. Revealing circadian mechanisms of integration and resilience by visualizing clock proteins working in real time. Nat. Commun. 9, 3245 (2018).
7. Ishiura, M. et al. Mutations in KaiA, a clock protein, extend the period of circadian rhythm in the cyanobacterium Synechococcus elongatus PCC 7942. Proc. Natl Acad. Sci. USA 92, 19127 (1995).
8. Chang, Y. G. et al. Circadian rhythms. A protein fold switch joins the inverse contrast matching-small-angle neutron scattering. J. Biol. Chem. 279, 19127–19135 (2004).
9. Oyama, K., Azai, C., Nakamura, K., Tanaka, S. & Terauchi, K. Conversion of a post-translational circadian oscillator and its interactions with the output His kinase SasA. PLoS ONE 6, e23697 (2011).
10. Phong, C., Markson, J. S., Wilhoite, C. M. & Rust, M. J. Robust and tunable circadian rhythms from differentially sensitive catalytic domains. Proc. Natl Acad. Sci. USA 110, 1124–1129 (2013).
11. Sugiyma, M. et al. Structural characterization of the circadian clock protein complex composed of KaiB and KaiC by inverse contrast-matching small-angle neutron scattering. Sci. Rep. 6, 35567 (2016).
12. Kageyama, H. et al. Cyanobacterial circadian pacemaker: Kai protein complex dynamics in the KaiC phosphorylation cycle in vitro. Mol. Cell 23, 161–171 (2006).
13. Mori, T. et al. Elucidating the ticking of an intracellular circadian clockwork. Proc. Natl Acad. Sci. USA 110, 1124–1129 (2013).
14. Mori, T. et al. Circadian rhythms. A protein fold switch joins the inverse contrast matching-small-angle neutron scattering. J. Biol. Chem. 279, 19127–19135 (2004).
15. Bretschneider, C. et al. A sequestration feedback determines dynamics and temperature entrainment of the KaiABC circadian clock. Science 355, 1174–1180 (2017).
16. Pattanayek, R. et al. Visualizing a circadian clock protein: crystal structure of the cyanobacterial circadian oscillator KaiC. Proc. Natl Acad. Sci. USA 101, 1479–1484 (2004).
17. Pattanayek, R. et al. Combined SAXS/EM based models of the S. elongatus post-translational circadian oscillator and its interactions with the output His kinase SasA. PLoS ONE 6, e23697 (2011).
18. Ziegler, S. J., Mallinson, S. J., John, P. C. S. & Bomble, Y. J. Advances to computational techniques. Proc. Natl Acad. Sci. USA 110, 1124–1129 (2013).
19. Sugiyma, M. et al. Combined SAXS/EM based models of the S. elongatus post-translational circadian oscillator and its interactions with the output His kinase SasA. PLoS ONE 6, e23697 (2011).
49. Sato, N. et al. A feasibility study of inverse contrast-matching small-angle neutron scattering method combined with size exclusion chromatography using antibody injections as model systems. J. Biochem. 169, 701–708 (2021).

50. Morishima, K. et al. Integral approach to biomacromolecular structure by analytical-ultracentrifugation and small-angle scattering. Commun. Biol. 3, 1–7 (2020).

51. Mandell, D. J., Coutsias, E. A. & Kortemme, T. Sub-angstrom accuracy in protein loop reconstruction by robotics-inspired conformational sampling. Nat. Methods. 6, 551–552 (2009).

52. Huang, P.-s. et al. RosettaRemodel: a generalized framework for flexible backbone protein design. PLoS ONE 6, e24109 (2011).

53. Hünénberger, P. H., Mark, A. E. & van Gunsteren, W. F. Fluctuation and cross-correlation analysis of protein motions observed in nanosecond molecular dynamics simulations. J. Mol. Biol. 252, 492–503 (1995).

54. Iwasaki, H., Nishiwaki, T., Kitayama, Y., Nakajima, M. & Kondo, T. KaiA-stimulated KaiC phosphorylation in circadian timing loops in cyanobacteria. Proc. Natl Acad. Sci. USA 99, 15788–15793 (2002).

55. Kim, Y. I., Dong, G., Carruthers, C. W. Jr., Golden, S. S. & LiWang, A. The day/night switch in KaiC, a central oscillator component of the circadian clock of cyanobacteria. Proc. Natl Acad. Sci. USA 105, 12825–12830 (2008).

56. Kitayama, Y., Iwasaki, H., Nishiwaki, T. & Kondo, T. KaB functions as an attenuator of KaiC phosphorylation in the cyanobacterial circadian clock system. EMBO J. 22, 2127–2134 (2003).

57. Nishiwaki, T. et al. A sequential program of dual phosphorylation of KaiC as a basis for circadian rhythms in cyanobacteria. EMBO J. 26, 4029–4037 (2007).

58. Valentin, E., Kikkney, A. G., Previtali, G., Jefferies, C. M. & Svergun, D. J. SASBDB, a repository for biological small-angle scattering data. Nucleic Acids Res. 43, D357–D363 (2015).

59. Bekker, G.-J., Kawabata, T. & Kurisu, G. The Biological Structure Model Archive (BSM-Arc): an archive for in silico models and simulations. Biophys. Rev. 12, 371–375 (2020).

60. Chang, Y. G., Kuo, N. W., Tseng, R. & LiWang, A. Flexibility of the C-terminal, or CII, ring of KaiC governs the rhythm of the circadian clock of cyanobacteria. Proc. Natl Acad. Sci. USA 105, 12825–12830 (2008).

61. Van Der Spoel, D. & van Drunen, R. GROMACS: a message-passing parallel molecular dynamics implementation. Computer Phys. Commun. 91, 43–56 (1995).

62. Lindahl, E., Hess, B. & Van Der Spoel, D. GROMOS 3.0: a package for molecular simulation and trajectory analysis. Mol. Model. Annu. 7, 306–317 (2001).

63. Van Der Spoel, D. et al. GROMACS: fast, flexible, and free. J. Comput. Chem. 26, 1701–1718 (2005).

64. Hess, B., Kutzner, C., Van Der Spoel, D. & Lindahl, E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J. Chem. Theory Comput. 4, 435–447 (2008).

65. Pronk, S. et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics 29, 845–854 (2013).

66. Pál, S., Abraham, M. J., Kutzner, C., Hess, B. & Lindahl, E. Tackling exascale software challenges in molecular dynamics simulations with GROMACS. In Solving software challenges for exascale, (eds Markidis, S. & Laure, E.), 3–27, (Springer, 2015).

67. Abraham, M. J. et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 1, 19–25 (2015).

68. Maier, J. et al. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. J. Chem. Theory Comput. 11, 3696–3713 (2015).

69. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 79, 926–935 (1983).

70. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. J. Chem. Phys. 126, 014101 (2007).

71. Franke, D. et al. ATLAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. J. Appl. Crystallogr. 50, 1212–1225 (2017).

72. Svergun, D. et al. Protein hydration in solution: experimental observation by x-ray and neutron scattering. Proc. Natl Acad. Sci. USA 95, 2267–2272 (1998).

73. Moroder, H. et al. ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering experiments. Methods Enzymol. 517, 93–111 (2015).

74. Pronk, S. et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular dynamics implementation. J. Chem. Theory Comput. 11, 3696–3713 (2015).

75. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 79, 926–935 (1983).

76. Franke, D. et al. ATLAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. J. Appl. Crystallogr. 50, 1212–1225 (2017).