

Structure of full-length ERGIC-53 in complex with MCFD2 for cargo transport

Received: 22 August 2023

Accepted: 7 March 2024

Published online: 16 March 2024

Satoshi Watanabe^{1,2,3}✉, Yoshiaki Kise⁴, Kento Yonezawa^{5,8}, Mariko Inoue¹, Nobutaka Shimizu⁵, Osamu Nureki⁴ & Kenji Inaba^{1,2,3,6,7}✉

ERGIC-53 transports certain subsets of newly synthesized secretory proteins and membrane proteins from the endoplasmic reticulum to the Golgi apparatus. Despite numerous structural and functional studies since its identification, the overall architecture and mechanism of action of ERGIC-53 remain unclear. Here we present cryo-EM structures of full-length ERGIC-53 in complex with its functional partner MCFD2. These structures reveal that ERGIC-53 exists as a homotetramer, not a homohexamer as previously suggested, and comprises a four-leaf clover-like head and a long stalk composed of three sets of four-helix coiled-coil followed by a transmembrane domain. 3D variability analysis visualizes the flexible motion of the long stalk and local plasticity of the head region. Notably, MCFD2 is shown to possess a Zn²⁺-binding site in its N-terminal lid, which appears to modulate cargo binding. Altogether, distinct mechanisms of cargo capture and release by ERGIC-53 via the stalk bending and metal binding are proposed.

Secretory proteins and membrane proteins are newly synthesized in the endoplasmic reticulum (ER) and undergo oxidative folding and N-glycosylation with the help of luminal chaperones, glycosylation enzymes, and oxidoreductases^{1,2}. A large portion of folded secretory proteins diffuse into the ER exit site, and are then transported into the Golgi apparatus (GA) by a bulk flow process³. However, subsets of secretory and membrane proteins are captured by cargo receptors in the ER, and subsequently transported to the GA in an accelerated manner³.

ERGIC-53 (also named LMAN1, P58) is a cargo receptor that captures certain specific secretory proteins^{4,5}, including coagulation factor V (FV) and factor VIII (FVIII)^{6,7}, α 1-antitrypsin^{8–10}, cathepsin C¹¹, cathepsin Z¹², Mac-2 binding protein¹³, matrix metalloproteinase-9¹⁴, and IgM¹⁵. Mutations of the *ergic-53* gene cause a genetic bleeding disorder called combined deficiency of FV and FVIII (F5F8D)^{7,16}. Another causative gene of F5F8D is the multiple coagulation factor deficiency 2

gene (MCFD2)^{16,17}. MCFD2 is a small protein composed of an EF-hand domain with two Ca²⁺ binding sites, and functions as a binding partner of ERGIC-53 in a Ca²⁺-dependent manner^{17–20}.

ERGIC-53 has also been reported to promote the ER-to-GA transport of the chaperone protein Erp44²¹ and several membrane proteins such as γ -aminobutyric acid type A receptors (GABA_ARs)²² and NDST1²³. Furthermore, ERGIC-53 interacts with surface glycoproteins of infected RNA viruses and enhances their viral propagation^{24,25}.

From the amino acid sequence, ERGIC-53 is predicted to be a single-pass transmembrane protein that consists of a luminal carbohydrate recognition domain (CRD), a long stalk domain, a transmembrane (TM) helix and a short cytoplasmic tail (CT)⁴. The CRD is structurally and functionally similar to a plant I-type lectin and binds to high-mannose type glycans of the target cargo proteins in a pH- and Ca²⁺-dependent manner^{11,18,19,26,27}. The CT contains a C-terminal KKFF motif required for the ERGIC-53 cycling between the ER and GA^{28–30}.

¹Institute of Multidisciplinary Research for Advanced Materials, Tohoku University, Sendai, Miyagi 980-8577, Japan. ²Department of Molecular and Chemical Life Sciences, Graduate School of Life Sciences, Tohoku University, Sendai, Miyagi 980-8577, Japan. ³Department of Chemistry, Graduate School of Science, Tohoku University, Sendai, Miyagi 980-8578, Japan. ⁴Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan. ⁵Structural Biology Research Center, Institute of Materials Structure Science, High Energy Accelerator Research Organization (KEK), Tsukuba, Ibaraki 305-0801, Japan. ⁶Medical Institute of Bioregulation, Kyushu University, Fukuoka 812-8582, Japan. ⁷Core Research for Evolutional Science and Technology (CREST), Japan Agency for Medical Research and Development (AMED), Tokyo, Japan. ⁸Present address: Center for Digital Green-innovation, Nara Institute of Science and Technology, Ikoma, Nara 630-0192, Japan. ✉e-mail: satoshi.watanabe.c1@tohoku.ac.jp; kenji.inaba.a1@tohoku.ac.jp

The stalk domain is predicted to be composed of four long α -helices followed by a long loop and the single-pass TM helix. The three stalk α -helices except the most N-terminal helix are predicted to form a coiled-coil³¹. The long loop between the fourth stalk helix and TM domain contains two conserved Cys residues (Cys466 and Cys475), which form intermolecular disulfide bonds³¹.

After the discovery of ERGIC-53 in 1988, it was long believed that this cargo receptor existed as a covalent hexamer or covalent dimer via the interprotomer disulfide bonds between the stalk domain³². A subsequent study by Neve et al., however, proposed that ERGIC-53 existed in two forms, a covalent hexamer and a noncovalent hexamer, in which three of disulfide-bridged dimers of ERGIC-53 form the disulfide-linked hexamer (covalent hexamer) or are noncovalently assembled into the hexamer (noncovalent hexamer)³³. This preceding study also demonstrated that the cysteine mutant that lacks the intermolecular disulfide bonds retained the hexameric state and was localized in the ERGIC as was wild-type ERGIC-53. It is thus suggested that the intermolecular disulfide bonds are not essential for the maintenance of its hexameric structure and cellular localization³³.

Although the oligomeric states of ERGIC-53 were analyzed by biochemical approaches, more detailed structural and functional studies have been focused on the CRD and its complex with MCFD2^{19,34–38}. Therefore, it still remained elusive how full-length ERGIC-53 forms the putative hexamer and works as a cargo receptor in this oligomeric state. Here, we present cryo-electron microscopy (cryo-EM) structures of full-length ERGIC-53 in complex with MCFD2. Contrary to the previous reports, the present structures, in combination with size-exclusion chromatography combined with multiangle light scattering (SEC-MALS) and with small-angle X-ray scattering (SEC-SAXS) analysis, reveal a tetrameric rather than hexameric architecture, with a long flexible stalk domain. High-resolution structures of its headpiece elucidate the molecular basis of the tetramer formation of ERGIC-53 and the zinc ion (Zn^{2+})-binding to the N-terminal region of MCFD2 for likely regulation of cargo binding and release. Altogether, the present study provides insights into the mechanism of cargo transport by ERGIC-53 using the long flexible stalk domain and the cargo-binding head region regulated by Zn^{2+} .

Results

ERGIC-53 exists as a tetramer with a long extended structure

The previously proposed hexamer model of ERGIC-53 was based on PAGE analyses of crude purified samples³³. To verify the oligomeric states of ERGIC-53 by more conclusive methods, we performed SEC-MALS analysis of purified full-length ERGIC-53. Recombinant full-length human ERGIC-53 was successfully purified to high purity from cumate-inducible stable HEK293T cell lines (Supplementary Fig. 1a). The monodisperse peak fractions from SEC elution contained two bands in nonreducing SDS-PAGE at ~300 kDa and ~110 kDa, which were presumed to result from the covalent and noncovalent hexamers of ERGIC-53, respectively. However, SEC-MALS conjugated analysis showed that the molecular masses of purified ERGIC-53 and its complex with MCFD2 were determined to be 230 and 280 kDa, respectively (Supplementary Fig. 1b–d). The determined molecular masses correspond to a homotetramer of ERGIC-53 alone (221.4 kDa) and a complex of four ERGIC-53 protomers with four MCFD2 molecules (277 kDa). The tetrameric state of ERGIC-53 was retained at weakly acidic Golgi pH (pH 6.5) (Supplementary Fig. 2). Thus, full-length ERGIC-53 likely exists as a mixture of covalent and noncovalent homotetramers, and each protomer forms a 1:1 complex with MCFD2. The hexamer model (332 kDa) suggested by preceding works did not receive further experimental support.

To gain further insight into the overall structure of ERGIC-53 in solution, we next performed SEC-SAXS analysis (Supplementary Fig. 1e and Supplementary Table 1). The pair distance distribution function $P(r)$ of full-length ERGIC-53 showed two separate peaks and estimated its maximum molecular dimension (D_{max}) to be 347 Å, suggesting that

this cargo receptor adopts an overall dumbbell-like shape with an unusually long length (Supplementary Fig. 1f, h). Compared with isolated ERGIC-53, the ERGIC-53-MCFD2 complex displayed almost the same D_{max} and two slightly higher and sharper peaks in the $P(r)$ function (Supplementary Fig. 1f, h), suggesting that in the complex, the corresponding two regions are more compact in conformation than those in isolated ERGIC-53. In the dimensionless Kratky plots of ERGIC-53 alone, the main peak was shifted to the larger qR_g side (~3.5) and has a higher height (~2.5) than those of typical globular proteins, suggesting that the overall conformation of isolated ERGIC-53 is extended and in a highly asymmetric state³⁹. On the other hand, the complex showed not only a slightly narrower width of the peak but also lower values at the high qR_g region than ERGIC-53 alone, indicating that conformational flexibility or disorder was repressed upon complex formation with MCFD2 (Supplementary Fig. 1g). These results suggest that ERGIC-53 binds MCFD2 to adopt a more rigid or tightly folded structure than the isolated state.

Cryo-EM structures of the head region of ERGIC-53 in complex with MCFD2

Based on the solution structure of the more rigid conformation of the ERGIC-53-MCFD2 complex, we performed cryo-EM analysis of full-length ERGIC-53 in complex with MCFD2. In the motion-corrected micrographs, long dumbbell-like particles of the complex were observed (Supplementary Fig. 3, red square), consistent with the $P(r)$ function obtained in the SEC-SAXS analysis. However, conventional autoparticle picking programs failed to recognize the full-length ERGIC-53 particles due to their unusual shapes. Instead, the autopicking programs recognized two globular portions in the dumbbell as separate particles (green circles in Supplementary Fig. 3). 2D classification of these two globular portions of the full-length particles showed clear 2D average images of the N-terminal tetrameric head of the complex and the C-terminal portion composed of the TM domain (including detergent micelle) and a part of the stalk domain, respectively (Supplementary Fig. 4a). Consequently, we determined the cryo-EM structures of the head region of full-length ERGIC-53 complexed with MCFD2 in two forms (forms A and B) at 2.53 and 2.59 Å resolutions, respectively (Supplementary Fig. 4a–g and Supplementary Table 2). On the other hand, only low-resolution EM maps were obtained for the C-terminal portion, probably due to its high flexibility and heterogeneity in the shape of detergent micelles.

The present cryo-EM analysis reveals that the overall structure of the assembled head is composed of four ERGIC-53 protomers and four MCFD2 molecules, confirming that ERGIC-53 exists as the tetramer (Fig. 1a, b). Each head region consists of the CRD (residues 31–268), stalk helices 1 and 2 (S-H1 and S-H2), and loops connecting these domains (Fig. 1c, left). The tetrameric structure is stabilized by a central coiled-coil composed of the long S-H2 from each protomer (residues 326–367) (Fig. 1b and Supplementary Fig. 5a). The S-H2 forms a typical parallel four-helix coiled-coil. Hydrophobic residues are located at the heptad positions **a** and **d** in the coiled-coil, forming the hydrophobic core within it (Supplementary Fig. 5b–e). A polar residue Gln338 is also located in the **a2** layer, where water or some ion is bound to stabilize the polar interactions (Supplementary Fig. 5f). The local resolution of S-H1 and the loops between CRD and S-H1, and between S-H1 and S-H2 are relatively low, suggesting the high flexibility of these regions (Supplementary Fig. 4c). MCFD2 binds the CRD at a 1:1 molar ratio as observed in the crystal structures of the CRD-MCFD2 complex^{34–37}. In addition, the S-H1 weakly associates with a part of MCFD2 (Fig. 1c and Supplementary Fig. 6a, inset 3), likely affecting the conformation of its N-terminal region (also see below).

CRD head assembly

Although ERGIC-53 forms the homotetramer, the four CRDs are not related by a fourfold axis, but arranged in a C2 symmetry. Two CRDs

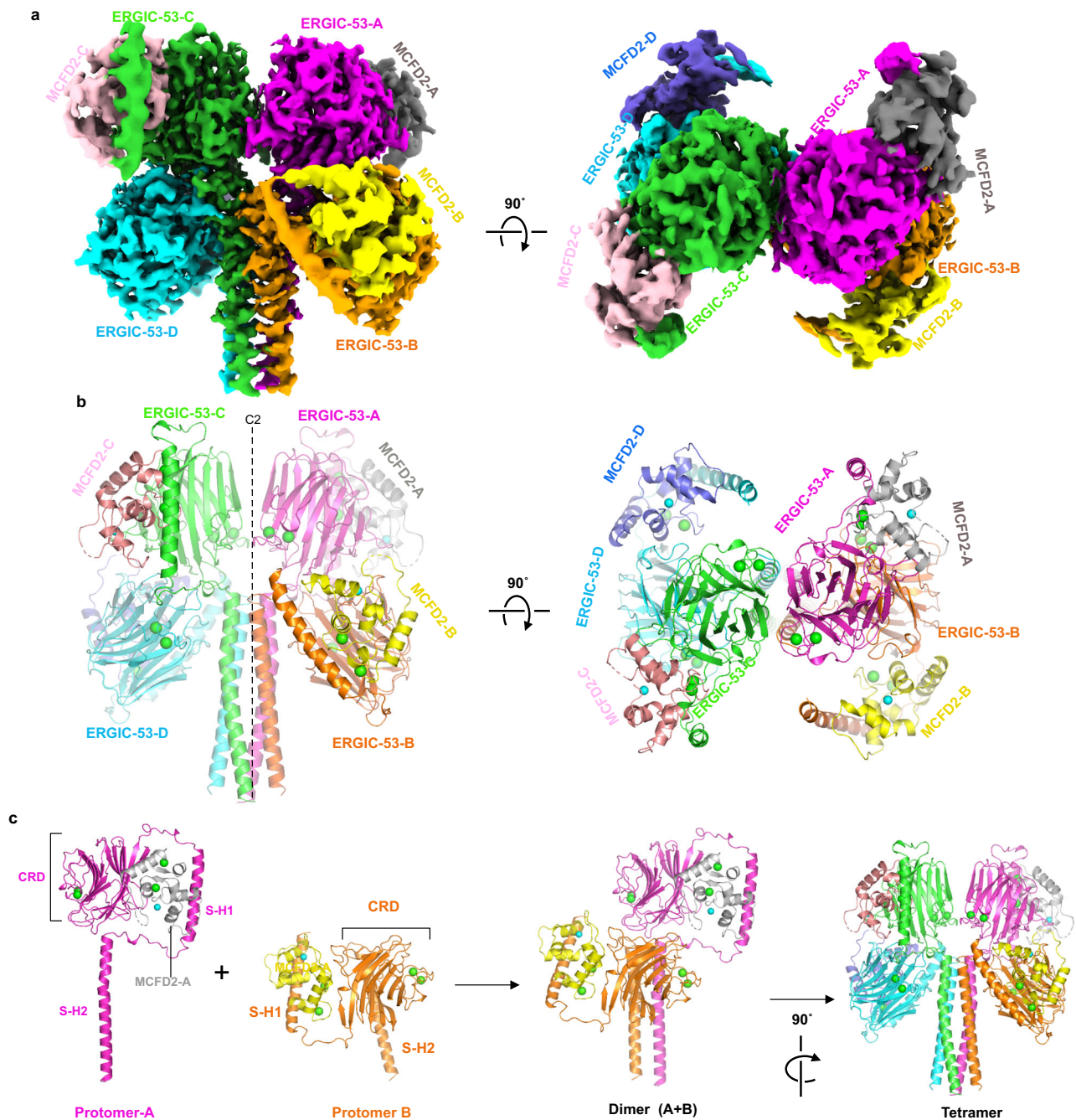


Fig. 1 | Cryo-EM structures of the head region of ERGIC-53 in complex with MCFD2. **a** Side and top views of the cryo-EM map of the head region of ERGIC-53 with MCFD2. Four ERGIC-53 protomers and four MCFD2 molecules are shown in magenta (ERGIC-53 protomer A), orange (protomer B), green (protomer C), cyan (protomer D), gray (MCFD2-A), yellow (MCFD2-B), pink (MCFD2-C), and slate blue (MCFD2-D), respectively. **b** Ribbon diagram of the structure of the

tetrameric head region of the ERGIC-53. The four ERGIC-53 protomers and four MCFD2 molecules are shown in the same colors as in 1a. Green and cyan spheres represent calcium and zinc ions, respectively. **c** Details of the tetrameric head formation. The left two panels show structures of the head region of the ERGIC-53 protomers in the upper-CRD configuration (protomer A) and lower-CRD configuration (protomer B).

are located in the upper layer, whereas the other two lie in the lower layer (Fig. 1b). In the protomers with the upper-CRD (protomers A and C), the head region adopts a “Hisayaku (Japanese ladle)”-like shape, in which both CRD and S-H1 are positioned above the S-H2 (Fig. 1c). MCFD2 is accommodated in the pocket formed by CRD, S-H1 and two long loops between CRD and S-H1 and between S-H1 and S-H2. In the protomers with the lower CRD (protomers B and D), CRD and S-H1 are rotated by -180° around the S-H1-S-H2 loop, and are located below the top of S-H2. The upper-CRD protomer vertically associates with the lower CRD protomer to form a dimer, and

then the two dimers are further assembled to form the four-leaf clover-like tetramer (Fig. 1b, c).

In the assembled head region, two different CRD-CRD interfaces are formed: a vertical interface (interface-V) between the vertically assembled CRDs (between protomers A and B and between protomers C and D) (Supplementary Fig. 6a) and the other interface (interface-H) between the horizontally assembled upper CRDs (between protomers A and C) (Supplementary Fig. 6b). Interface-V is formed by van der Waals contacts between residues in the β -strands and their flanking loops of each CRD (Supplementary Fig. 6a, inset 1). A similar interface is

also observed in the crystal packing of the CRD-MCFD2 complex (PDB code: 3WHU, 4YGE, etc.). The lower CRD (protomer B) also interacts with S-H2 from the counter protomer (protomer A) (Supplementary Fig. 6a, inset 2). On the other hand, interface-H is formed by only four hydrogen bonds, including that between Arg192 and Gln191¹ and two van der Waals contacts between the CRDs (Supplementary Fig. 6b).

As described above, two different forms (A and B) of the head region are determined, in which the positions and orientations of the CRDs relative to the S-H2 are different. The superimposition on the central S-H2 coiled-coil demonstrates that during the conversion from form A to form B, the upper and lower CRDs rotate by 4.8° as a rigid body (Supplementary Fig. 6c). Consequently, interface-H in form B is formed by interactions between different residues, such as van der Waals contact between Gln191 and Glu228¹ (Supplementary Fig. 6b right), while interface-V is almost maintained. These results suggest that interface-V involves tight binding, whereas interface-H is variable. As a result, the vertically assembled pair of CRDs can rotate to a certain extent (by 4.8°) as a rigid body.

Overall architecture of full-length ERGIC-53 with MCFD2

Next, we endeavored to determine the whole structure of the full-length ERGIC-53-MCFD2 complex. Manual picking of full-length particles was first conducted based on the positions of the head and TM regions. An initial picking model was trained by Topaz⁴⁰ using particles selected from 2D classification, and the picking model was further improved by repeating Topaz picking, 2D classification, and Topaz training. Finally, Topaz successfully recognized the full-length particles and enabled us to perform a single particle analysis of full-length ERGIC-53 (Supplementary Fig. 7).

2D classification of the collected full-length particles generated clear 2D average images of their four-leaf clover-like structures consisting of the head, long stalk, and TM regions (Fig. 2a). Some particles adopt nearly straight conformations, while others possess stalks largely bent at various angles. These results suggest that the stalk region of ERGIC-53 is highly flexible.

By using straight particles, the cryo-EM structure of full-length ERGIC-53 in complex with MCFD2 was determined at 6.8 Å resolution (Fig. 2b and Supplementary Table 2). The maximum length of the full-length ERGIC-53 structure is ~340 Å, which is almost comparable to the D_{\max} value (347 Å) obtained by SAXS analysis (Supplementary Fig. 1f). Based on the obtained cryo-EM map, we built a structure model of full-length ERGIC-53 by using AlphaFold2 multimer⁴¹ (Fig. 2c). Similar to S-H2, the stalk helices 3 (S-H3) and 4 (S-H4) also appear to form a four-helix coiled-coil. Thus, the three sets of long four-helix coiled-coils are vertically connected by relatively long loops to form the stable but flexible tetramer. Although the TM helices were not well resolved in the current cryo-EM map, they are also predicted to form a four-helix bundle and be vertically connected to S-H4. As a result, the full-length ERGIC-53 adopts a slender overall architecture. To our knowledge, the molecular height of full-length ERGIC-53 is taller than that of any other membrane protein of known structure, including the largest Ca²⁺ channel ryanodine receptor RyRs, V-type ATPase, and spike protein from COVID-19 (Supplementary Fig. 7c).

To assess the structural contribution of MCFD2 to the overall architecture of ERGIC-53, we also attempted the cryo-EM analysis of full-length ERGIC-53 alone. While 2D average images showed features of the C-terminal portion including S-H4 and the TM domain, clear 2D images were not observed for the tetrameric head regions (Supplementary Fig. 8a). In the full-length particle analysis by using Topaz, the head regions were highly blurred after 2D averaging (Supplementary Fig. 8b). Consistently, our SEC-SAXS analysis suggests that the structure of full-length ERGIC-53 is more flexible in the absence of MCFD2 (Supplementary Fig. 1f, g). Thus, MCFD2 binding seems likely to stabilize the structure of the tetrameric head of ERGIC-53.

Detailed molecular motions of full-length ERGIC-53

To gain insight into the molecular motions of full-length ERGIC-53, 3D variability analysis (3DVA) was performed in CryoSPARC⁴², resulting in the identification of two different continuous motions of this cargo receptor (components I and II) (Fig. 2d, e and Supplementary Movie 1). 3DVA suggests that the entire structure of ERGIC-53 can be divided into three rigid-body segments, which correspond to the head region (segment 1), S-H3 (segment 2), and S-H4 (segment 3), respectively, followed by the C-terminal TM helix. The hinges between the three segments correspond to loops between S-H2 and S-H3 (hinge A) and between S-H3 and S-H4 (hinge B). In component I, segments 1 and 2 undergo a slight bending and stretching motion around hinge A. Segment 3 also undergoes a swing movement, rotating about hinge B in the opposite direction to segments 1 and 2 (Fig. 2d). The rotation directions of these three segments are nearly parallel to the interface-H between the upper CRDs (Fig. 2d bottom). The density of the TM domain was largely variable during the conversion from the first to last frames in this component, suggesting that the TM domain also swings around the loop flanked by S-H4 and the TM domain (i.e., the third hinge) (Fig. 2d). In component II, the segments 1–3 undergo similar bending and stretching movements through the two hinges, in which their rotation directions are perpendicular to interface-H (Fig. 2e, bottom). In addition, other conformations with even more greatly bent stalks were also observed in the 2D class-average images of the intact particles (Fig. 2a lower panels). These results suggest that full-length ERGIC-53 continuously adopts diverse conformations by bending stalk and TM helices at the three hinges.

Molecular motion of the head region in full-length ERGIC-53

As mentioned above, the two different structures of the head region suggest its conformational variability (Supplementary Fig. 6c). To further assess the conformational variability of the head region, 3DVA was also performed by focusing exclusively on this region (Supplementary Fig. 9a and Supplementary Movie 2). As observed in forms A and B (Supplementary Fig. 6c), component I^{head} shows a swing movement of the CRDs, in which the CRD dimer between upper and lower protomers shows a side-to-side rotation as a rigid body dimer (Supplementary Fig. 9a left and Movie 2 left). The rotation direction of the dimer is nearly parallel to interface-H (side-to-side rotation) (Supplementary Fig. 9a lower left). In component II^{head}, two CRD dimers rotate along a vertical axis (i.e., nearly parallel to the central coiled-coil) in opposite directions (referred to as twisting motion) (Supplementary Fig. 9a lower right and Supplementary Movie 2, right).

Based on a clustering analysis of the 3DVA of the head region, four different substate structures (A–D) were reconstructed at 3.29–3.51 Å resolution (Supplementary Fig. 9b, c). Superimposition of the central coiled-coil helices from the substrates A and B reveals that the upper and lower CRD dimers are rotated by 6.2° as a rigid body, which is a larger rotation between form A and form B described above, suggesting that form A and form B represent more averaged structures of the head region individually (Supplementary Figs. 9d vs. 4c). In the superposition of substate structures C and D, S-H1 and MCFD2 in the upper layers and those in the lower layers rotate by 2.5° and 1.9°, respectively, breaking the C2 symmetry of the head region (Supplementary Fig. 9e). These rotations seem to be ascribed to the flexible loops between the CRD and S-H1 and between S-H1 and S-H2. Thus, the two loops within the head region likely generate the plasticity of the head region, allowing the CRD-MCFD2 complex to fluctuate continuously in the tetrameric structure.

Local plasticity of the head region independent of the global stalk bending

To investigate the linkage between the local plasticity of the head region and the global stalk bending, we constructed an ERGIC-53 ΔH34 mutant in which S-H3 and S-H4 within the stalk were deleted, and

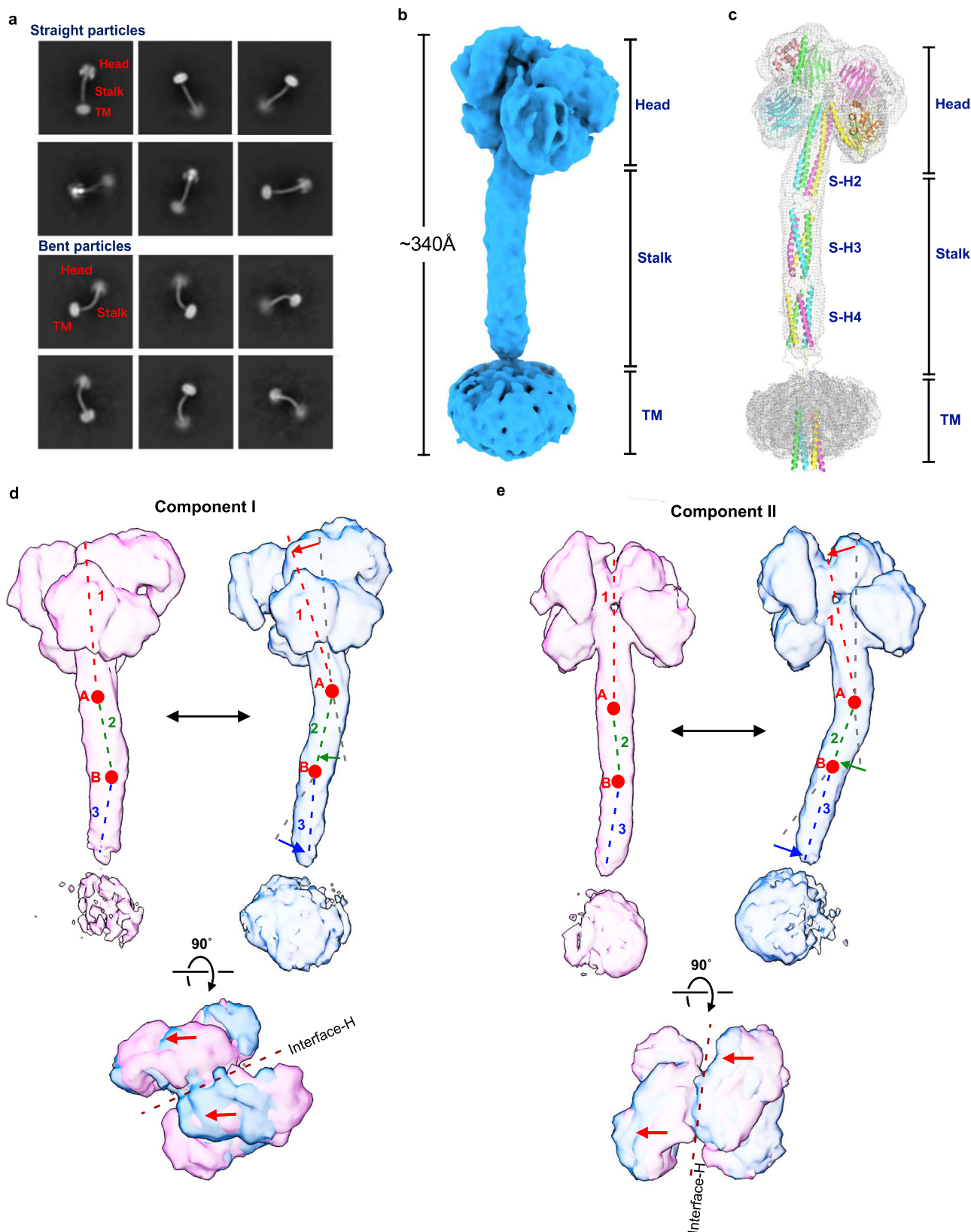


Fig. 2 | Cryo-EM structure of full-length ERGIC-53 in complex with MCFD2. **a** Representative 2D class-average image of the full-length particles. The upper and lower panels represent straight and bent particles, respectively. **b** Cryo-EM map of full-length ERGIC-53 in complex with MCFD2. **c** The overall architecture of full-length ERGIC-53 overlaid in the EM map shown in (b). **d, e** Results of 3DVA of the full-length ERGIC-53 with two variability components I (d) and II (e). The 3D

EM maps of the first (pink) and last frames (blue) of continuous conformational changes generated by 3DVA are displayed. Broken lines represent the center of three segments. Red circles represent hinges between the segments 1, 2, and 3. The lower panels show the top views of the two maps superimposed on each other. The brown broken line represents the interface-H between the upper CRDs.

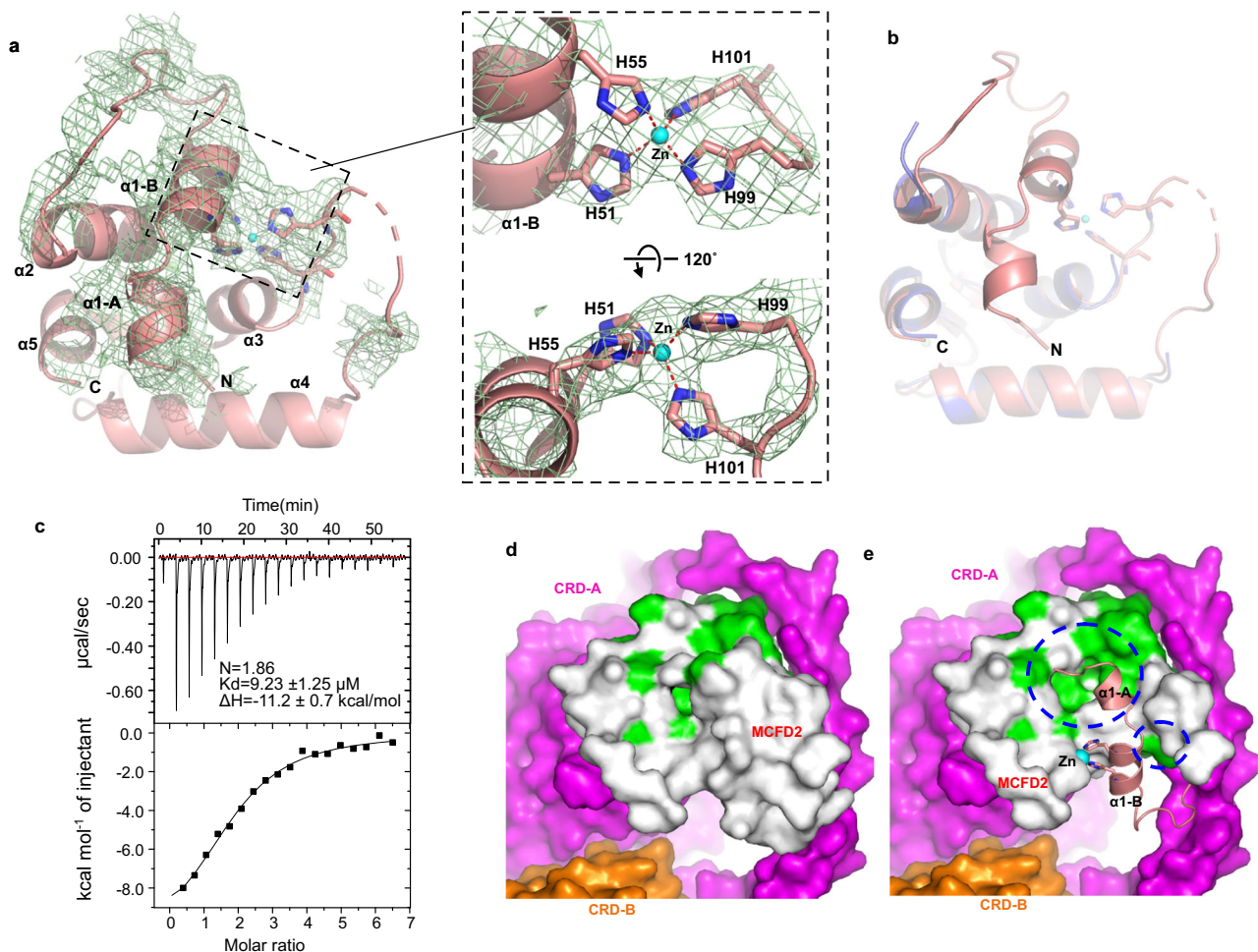


Fig. 3 | Visualized N-terminal segment of MCFD2 contains a Zn^{2+} binding site. **a** The overall structure of MCFD2 bound to full-length ERGIC-53. The EM density map around the region that is missing in the previous crystal structures is shown in a light green mesh. The inset is showing a close-up view of the His cluster, showing that Zn^{2+} is probably coordinated by four His residues. **b** Superposition of the Zn^{2+} -bound MCFD2 structure revealed by the present cryo-EM analysis (pink) with the previously reported structure of Zn^{2+} -free MCFD2 (blue, PDB ID: 4YGE). **c** ITC raw

data (upper) and binding isotherm data (lower) for titration of $ZnCl_2$ (500 μM) into the EDTA-treated MCFD2 sample. Errors represent fitting residuals (standard deviation). **d, e** Comparison of FV/FVIII binding sites in MCFD2 with (d) or without (e) the N-terminal lid. MCFD2 (white) and two CRDs (magenta and orange) are shown in the surface representation. Previously identified FV/FVIII binding sites in MCFD2 are colored in green. Blue dashed circles indicate the regions of the FV/FVIII binding site masked by the N-terminal lid.

determined its cryo-EM structure at 3.78 Å resolution (Supplementary Fig. 10a, b). The $\Delta H34$ mutant retains a tetrameric structure with the assembled head region, similar to wild-type. Additionally, the TM helices of the mutant were not well resolved in the density map, as was the case for the wild-type. 3DVA of this deletion mutant revealed its three different motions (Supplementary Fig. 10c and Supplementary Movie 3). In the component I ^{$\Delta H34$} , the TM domain shows a large swing movement, while the head region remains static (Supplementary Fig. 10c, left and Movie 3 left). On the other hand, components II ^{$\Delta H34$} and III ^{$\Delta H34$} show the side-to-side rotation and twisting motion of the CRD dimers, respectively (Supplementary Fig. 10c, center and right and Supplementary Movie 3, center and right), as observed in the 3DVA of the head region of full-length ERGIC-53. The TM domain shows only marginal rotation in either component II ^{$\Delta H34$} or III ^{$\Delta H34$} . Thus, the CRD dimers in the head region appear to move separately from the region below the S-H2, suggesting that there is no linkage between the head-domain plasticity and the stalk-domain bending.

Visualization of an N-terminal segment with a zinc-binding site in MCFD2

In the previously reported crystal structures of MCFD2 alone and its complex with CRD^{20,34–37}, the N-terminal region of MCFD2 (residues

27–66) and a loop between $\alpha 3$ and $\alpha 4$ helices (residues 100–111) were disordered and hence invisible (Supplementary Fig. 11a). However, the present cryo-EM map displays significant density for the regions preceding the $\alpha 2$ helix (Fig. 3a). In this context, AlphaFold2 prediction⁴¹ suggests that the disordered N-terminal region contains two short α -helices ($\alpha 1$ -A and $\alpha 1$ -B) (Supplementary Fig. 11b). Although we had difficulty in de novo modeling of the MCFD2 N-terminal region from the present map, the predicted N-terminal $\alpha 1$ -A, and $\alpha 1$ -B short helices and a loop between the $\alpha 1$ -B and $\alpha 2$ helices are nicely fitted onto the observed density map (Fig. 3a). Thus, the present cryo-EM structure of MCFD2 in complex with full-length ERGIC-53 is nearly identical to the original structures (RMSD of 0.53 Å) except the N-terminal segment (Fig. 3b). The difference in the maps of MCFD2 obtained by the crystal and cryo-EM analyses is likely due to the presence of S-H1 of ERGIC-53 in the cryo-EM sample. Indeed, this helical segment interacts with the N-terminal region of MCFD2, seemingly restricting its conformational flexibility (Fig. 2c and Supplementary Fig. 6a).

In the updated MCFD2 structure, four conserved His residues (His51, His55, His99, and His101) are clustered at the $\alpha 1$ -B helix and a loop between the $\alpha 3$ and $\alpha 4$ helices (Fig. 3a and Supplementary Fig. 11d). Notably, the geometry of the histidine cluster and the presence of density at the center of the cluster are consistent with

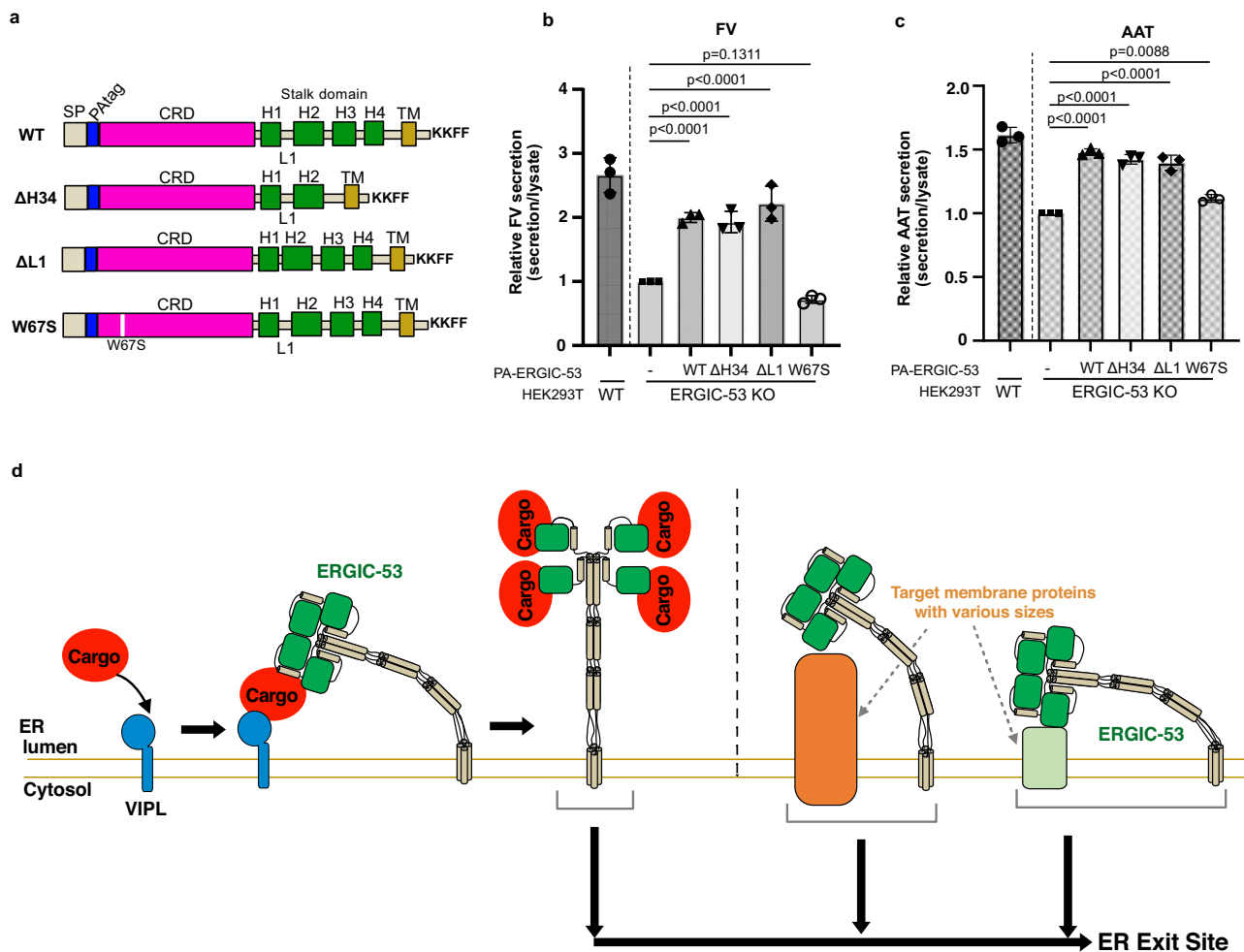


Fig. 4 | Role of the long stalk and the head assembly in cargo transport.

a Schematic representation of the ERGIC-53 mutants used for the rescue experiments. **b, c** Rescue experiments of FV-HiBiT (**b**) and AAT-HiBiT (**c**) secretion by coexpression of ERGIC-53 or its mutants in ERGIC-53 KO cells. The amounts of secreted FV/AAT relative to those of intracellular FV/AAT were quantified with a commercially available HiBiT-tagged protein detection reagent and normalized

to those from the KO cells transfected with an empty vector. Each symbol represents the individual data points. Error represents the standard deviation of the mean ($N=3$ biological replicates). Statistical significance was measured with one-way ANOVA followed by Dunnett's test. Exact P values are indicated in the graphs for each construct. **d** A working model of cargo capture by ERGIC-53.

coordination of a metal, most likely a zinc ion (Fig. 3a inset). In support of this, a colorimetric assay with 4-(2-pyridylazo) resorcinol (PAR), a chromophoric chelator for divalent metal ions, revealed that purified MCFD2 contained significant amounts of divalent metal ions such as Zn^{2+} (Supplementary Fig. 11c). Consistently, our isothermal titration calorimetry (ITC) analysis revealed that an EDTA-treated MCFD2 sample can bind up to two molar equivalents of Zn^{2+} with micromolar affinity (Fig. 3c). Thus, MCFD2 has been shown to be a zinc-binding protein with the conserved His cluster.

A previous NMR study revealed that the EF-hand helices of MCFD2 are involved in the interaction with FV/FVIII⁴³. In the present cryo-EM structures, the reported FV/FVIII binding site in MCFD2 is not fully exposed to the solvent but rather shielded by its N-terminal short helices (Fig. 3d, e). The N-terminal segment of MCFD2 appears to be hooked by the Zn^{2+} -bound His cluster formed between the α 1-B helix and α 3- α 4 loop (Fig. 3a). These structural features raise a possibility that Zn^{2+} binding to the His cluster may prevent the cargo binding to the head of ERGIC-53.

To explore how Zn^{2+} affects cargo transport, we performed an FV secretion assay using WT and ERGIC-53 KO 293T cells with or without Zn^{2+} supplementation. Intracellular Zn^{2+} concentration is known to be affected by extracellular Zn^{2+} levels in culture medium⁴⁴. When 10 μ M

$ZnCl_2$ was added to the serum-free Expi293 medium (Supplementary Fig. 11e), FV secretion from WT cells was slightly but significantly prevented. In contrast, the ERGIC-53 KO cells hardly displayed the Zn^{2+} -induced prevention of FV secretion. Altogether, these results suggest that the cargo transport mediated by ERGIC-53/MCFD2 can be inhibited by Zn^{2+} , consistent with the present observation that the cargo-binding site of ERGIC-53 is masked by the N-terminal segment of “ Zn^{2+} -bound” MCFD2.

Physiological significance of the long stalk and head assembly

Our cryo-EM structures reveal that full-length ERGIC-53 adopts the four-leaf clover-like structure with an unusually long and flexible stalk. To examine the physiological roles of the long stalk in the cargo transport, rescue experiments were performed by monitoring FV and alpha-anti-trypsin (AAT) secretion under coexpression of the cargo proteins and ERGIC-53 WT or its stalk-deletion mutant in the ERGIC-53 KO cells (Fig. 4a and Supplementary Fig. 12a). Upon cotransfection, the expression levels of WT and its mutant were comparable to that of endogenous ERGIC-53 in HEK293T cells (Supplementary Fig. 13a, d). As expected, the coexpression of ERGIC-53 WT enhanced the secretion of FV and AAT, compared to the empty vector (Fig. 4b, c and Supplementary Fig. 13b–e). Similarly, expression of the Δ H34 mutant with a

much shorter stalk (Fig. 4a) rescued the FV section at a level comparable to WT (Fig. 4b). A similar result was obtained for the AAT secretion (Fig. 4c). Thus, deletions of the S-H3 and S-H4 only marginally affected the cargo transport ability of ERGIC-53. To verify the validity of our ERGIC-53-mediated cargo secretion assay, we also investigated the FV and AAT secretion from the ERGIC-53 KO cells expressing ERGIC-53 W67S, a missense mutant deficient in both MCFD2 and mannose bindings⁴⁵. As reported previously^{45,46}, this negative control mutant failed to rescue the FV secretion while the AAT secretion was partially rescued by this mutant (Fig. 4b, c), probably because ERGIC-53 captures AAT in an MCFD2-independent manner⁴⁷. Collectively, it can be interpreted that the shorter version of ERGIC-53 (Δ H34) still retains the ability to enhance the secretion of FV and AAT, although the distance between the head and TM domain and the flexibility of the stalk are substantially reduced compared to those of wild-type.

Additionally, to examine the physiological significance of the head region assembly for cargo transport, we constructed a mutant lacking the L1 loop between S-H1 and S-H2 (residue Gln313 to Glu322) (Δ L1 mutant) (Fig. 4a and Supplementary Fig. 12b, c). In the head region, two S-H1s of protomers A and B are located \sim 65 Å apart by the L1 loops to form the vertically assembled CRD dimers. (Supplementary Fig. 12b). Hence, the deletion of the L1 loop is predicted to disrupt the interface-V between the upper and lower CRDs, as shown in the AlphaFold2 (AF2) predicted model (Supplementary Fig. 12c, d). The Δ L1 mutant displayed a similar migration pattern to the WT on a nonreducing SDS gel (Supplementary Fig. 13a lane 5), suggesting that this mutant also exists as the mixture of covalent and noncovalent tetramers, likely formed through the central four-helix coiled-coils. Coexpression of the Δ L1 mutant also rescued the FV and AAT secretion at a level similar to that of the WT (Fig. 4b, c). These results suggest that the head region assembly is not an essential factor for the efficient transport of soluble cargo proteins by ERGIC-53 and MCFD2.

Discussion

The present cryo-EM analysis reveals the tetrameric slender architecture of full-length ERGIC-53 in complex with MCFD2. The long stalk of ERGIC-53 is composed of three sets of four-helix coiled-coils that are vertically aligned. Full-length ERGIC-53 can adopt various conformations, including straight and largely bent conformations by bending the stalk domain. The CRDs within the head region also fluctuate independently of the stalk bending. These flexible motions of ERGIC-53 are most likely generated by the loops that function as hinges between the CRD, stalk helices, and TM domain.

The tetramer formation of ERGIC-53 revealed in this study is similarly observed with its yeast homolog Emp46p and Emp47p^{48,49}, suggesting that the tetrameric state may be widely conserved among the eukaryotic ERGIC-53 homologs. The heterotetramer complex of Emp46p and Emp47p was stable at pH 6–7, as observed for ERGIC-53 (Supplementary Fig. 2). However, the heterotetramer composed of Emp46p and Emp47p was reported to be dissociable at more acidic pH. In this transition, a glutamate residue Glu306 of Emp46p, which is located inside the hydrophobic core of its coiled-coils, is suggested to function as a pH sensor that regulates its oligomeric state. By contrast, ERGIC-53 contains no charged residues inside the hydrophobic core of the coiled-coils. Notably, the three stalk helices of ERGIC-53 (\sim 120 residues) are much longer than those of Emp46p/Emp47p (\sim 60 residues), suggesting tighter interactions between the coiled-coils in ERGIC-53. Furthermore, the ERGIC-53 tetramer is stabilized by intermolecular disulfide bonds. Thus, the tetrameric state of ERGIC-53 is likely robust under the environmental conditions of the early secretory pathway and retained during the cycling between the ER and GA.

In the prostate and several tissues, an ERGIC-53 homolog, ERGL (ERGIC-53-like, LMANIL), is highly expressed⁵⁰. A recent study reported that ERGL is also involved in the secretion of the FVIII section in HCT116 cells. However, ERGL is not expressed at a detectable level in

HEK293 cells⁵¹, suggesting that the functional contribution of endogenous ERGL may be negligible in our rescue experiments. The putative stalk domain of ERGL, which shows a low sequence similarity to that of ERGIC-53, is also predicted to form three sets of coiled-coils (Supplementary Fig. 14a). Moreover, AF2 multimer predicted a possible heterotetramer structure of ERGIC-53 and ERGL via three sets of four-helix coiled-coils, despite the low sequence similarity between these two (Supplementary Fig. 14b). Thus, ERGL may exit as a homotetramer or form a heterotetramer with ERGIC-53 to function as an additional or auxiliary cargo receptor in some tissues and cells.

ERGIC-53 has been proposed to cooperate with other luminal L-type lectins, such as VIP36(LMAN2), VIPL(LMAN2L), and ERGL (LMANIL), to maintain the homeostasis of secretory cargo proteins^{26,50,52}. In the proposed model, properly folded cargo proteins with M9 glycans are first recognized by VIPL, and then transferred to ERGIC-53. VIPL consists of a CRD and a single TM but lacks a stalk region. ERGIC-53 presumably receives cargo proteins from the small VIPL by utilizing the largely bent long stalk (Fig. 4d left). In this context, it is reasonable that the Δ H34 mutant with the much shorter stalk still retains the cargo transport activity at a similar level to WT. Our rescue experiments also revealed that the head assembly is not essential for efficient cargo transport. Given the docking model of dimannose bound to the CRD based on the previous crystal structure (4YGE), the 1-OH of the Man (4) moiety of Man₉(GlcNAc)₂ glycan is located near the L1 loop between S-H1 and S-H2 (Supplementary Fig. 11f). In the tetrameric head, the L1 loop may interfere with the accommodation of N-glycans of cargo proteins. Thus, the present CRD assembled structure possibly represents a post-cargo release state of the ERGIC-53-MCFD2 complex. Cargo binding may alter the relative positions of the CRDs to capture a target cargo protein in concert with MCFD2.

In addition to soluble secretory proteins, ERGIC-53 is known to facilitate the transport of certain specific membrane proteins, including neuroreceptors such as GABA_AR in neuronal cells²² and several Golgi resident membrane proteins such as GOLGA5, NDST1, and FKRP²³. ERGIC-53 is also associated with surface glycoproteins of arenavirus, hantavirus, coronavirus, and hepatitis B virus^{24,25}. Given that these target membrane proteins vary in their size and structures, the conformational flexibility of the stalk domain and the local rotation of the CRDs appear to help adjust the height and position of the head region of this cargo receptor to those membrane proteins (Fig. 4d right). Notably, although the CRD of ERGIC-53 is an essential element for the transport of these membrane proteins, its glycan binding ability is dispensable for the interaction with them^{24,25}. Collectively, ERGIC-53 is likely to recognize and bind target membrane proteins differently from soluble proteins.

Here, MCFD2 has been found to be a zinc-binding protein. The ITC analysis for Zn²⁺ binding to MCFD2 and purification of MCFD2 in a metal ion (Zn²⁺)-bound form from *E. coli* cells suggest that the Zn²⁺ binding ability of MCFD2 does not rely on the complex formation with ERGIC-53. While the overall structure of MCFD2 is not largely altered upon Zn²⁺ binding, its cargo-binding site is masked by its N-terminal helix in the present cryo-EM structure, and this local structure is probably stabilized by the Zn²⁺-bound His cluster (Fig. 3a). The relatively low local resolution of the N-terminal region of MCFD2 may suggest that this segment is readily moved away upon cargo binding. Thus, the N-terminal helix of MCFD2 appears to function as a lid to regulate the cargo binding/release in a Zn²⁺-dependent manner (Supplementary Fig. 15). It has been presumed that ERGIC-53 releases cargo proteins in the GA with lower Ca²⁺ concentration and lower pH^{11,18,19,26,27}. In contrast to Ca²⁺, the labile Zn²⁺ concentration has been reported to be much higher in the GA than in the ER and other organelles^{53–55}. Accordingly, these findings suggest that Zn²⁺ binds transiently to MCFD2 in the GA and thereby promotes the cargo release by closing its N-terminal lid. The proposed working hypothesis of Zn²⁺-dependent regulation of the ERGIC-53 MCFD2 system is relevant to our recent

finding that Zn²⁺ is vital to protein quality control mediated by ERp44 in the early secretory pathway (ESP)^{54,56}. In sharp contrast to the ERGIC-53-MCFD2 system, Zn²⁺ enhances the complex formation between ERp44 and its client proteins in the GA, whereas ERp44 releases its clients in the ER at much lower Zn²⁺ concentrations^{54,56}. Future studies will further elucidate the physiological roles of Ca²⁺, Zn²⁺, and other metal ions in protein homeostasis mediated by various molecular chaperones and cargo receptors in the ESP.

As an additional note, the revealed slender structure of full-length ERGIC-53 seems highly relevant to C-type lectins, which are also composed of a CRD, long stalk, and TM domain. Many of the C-type lectin family proteins are localized on the cell surface and function as receptors for a broad range of ligands, including pathogens^{57,58}. Intriguingly, a recent study reported that in dendritic cells and their related cells, such as airway epithelial cells, a fraction of ERGIC-53 is located on the cell surface and functions as a receptor for house dust mite (HDR) allergens⁵⁹. The molecular height of ERGIC-53, which is comparable to those of the C-type lectins, is likely advantageous for this receptor to capture target HDR and other potential ligands, without substantial interference from surrounding C-type lectins on the cell surfaces. It is still difficult to visualize cell surface receptors such as full-length C-type lectins due to their unusually long stalks and large flexibility. Nonetheless, as shown here, cryo-EM single particle analysis can visualize the overall structures of the flexible receptor not only in the 2D class-average images but also in 3D reconstruction at a middle resolution that enables the reliable identification of each domain. Flexibility analysis of the EM data, such as 3DVA, can visualize their global motion. Further structural analysis of long flexible receptors will be feasible and provide a deep understanding of their structures and mechanisms of action.

Methods

Cell cultures and plasmids

The cDNA of human ERGIC-53 was subcloned from a cDNA library from HeLa cells and inserted into the pcDNA3.1 vector (Invitrogen). Subsequently, the ERGIC-53 gene with a C-terminal PA tag was inserted into a PiggyBac-Cumate Switch Inducible vector (System Bioscience). The pcDNA-SP-PA-ERGIC-53 plasmid was generated using the In-Fusion cloning kit (Takara) with PCR fragments from the pcDNA3.1-ERGIC-53-PA plasmid. Mutants of ERGIC-53 were constructed by PCR-based site-directed mutagenesis. A PCR fragment of the ΔH34 mutant with a C-terminal FLAG was also inserted into a pEMulti-puro vector (Fujifilm-Wako). Sequences of the primers used in this study are summarized in Supplementary Table 3.

The cDNA of human MCFD2 (IRAK026G06) was provided by the RIKEN BRC through the National BioResource Project of the MEXT/AMED, Japan. The lumen domain of MCFD2 (residues 27–146) was subcloned into the pET15b vector with an N-terminal histidine tag. The cDNA of human Factor V (FV) (FXC01788) was purchased from Kazusa DNA Research Institute. The secreted region of FV (residues 29–2224) was inserted into the phlsec-vector (Addgene, 72348) with the C-terminal HiBiT-Tag (Promega), using the In-Fusion cloning kit. A codon-optimized cDNA fragment of the secreted region of human AAT was synthesized (Eurofins) and inserted into the phlsec-vector with the C-terminal HiBiT-Tag.

Human embryonic kidney (HEK) 293T cells were purchased from ATCC (American Type Culture Collection, CRL-3216). Human ERGIC-53 (LMAN1) KO 293T cell lines (ab266248), as well as wild-type cells (ab255449) were purchased from Abcam.

Protein expression and purification

The PiggyBac-cumate-ERGIC-53-PA tag plasmid, along with the Super PiggyBac Transposase Expression Vector (System Bioscience), was transfected to generate a stable HEK293T cell line with cumate-inducible ERGIC-53 expression. The established stable cells were

cultured in Dulbecco's modified Eagle's medium (DMEM, Nacalai Tesque) supplemented with 4–5% fetal calf serum (FCS; Thermo Fisher Scientific, or Nichirei Biosciences Inc) and 1% penicillin-streptomycin mixed solution (Nacalai Tesque) under 5% CO₂ at 37 °C. When cells were grown to 70–80% confluency in 15 cm dishes, the medium was replaced with serum-free DMEM supplemented with 6 mM sodium butyrate (Nacalai Tesque). After 6 h incubation at 37 °C, expression of ERGIC-53 was induced by the addition of cumate (240 μg/ml) and phorbol 12-myristate 13-acetate (PMA, 50 nM), and then cells were cultured at 30 °C in 5% CO₂ for another 2 days before harvesting. The cells were washed with PBS and stored at –80 °C. The pEB-multi-puro-ERGIC-53 ΔH34-FLAG was transfected to generate a stable HEK293T cell line for expression of the ERGIC-53 ΔH34-FLAG mutant. The established stable cells were cultured in DMEM supplemented with 4–5% FCS and 1% penicillin-streptomycin mixed solution (Nacalai Tesque) under 5% CO₂ at 37 °C. When cells were grown to 70–80% confluency in 15 cm dishes, sodium butyrate (6 mM) and PMA (50 nM) were added, and then cells were cultured at 30 °C in 5% CO₂ for another 2 days before harvesting. The cells were washed with PBS and stored at –80 °C.

All purification procedures were conducted at 4 °C or on ice. Cell pellets expressing ERGIC-53-PA from 3 L of medium or those expressing the ERGIC-53 ΔH34 mutant from 1 L of medium were resuspended in buffer A (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM CaCl₂, 10% glycerol) supplemented with 5 mM MgCl₂ and 1x protein inhibitor cocktail (Nacalai Tesque) and –50 μg/ml DNase I (Fujifilm-Wako). The cells were broken by brief sonication and were solubilized by the addition of 1% (v/v) n-dodecyl-β-D-maltoside (DDM; Nacalai Tesque) for 2 h. The solubilized fraction was clarified by ultracentrifugation (Hitachi CS100FNX, 150,000×g, 20 min) and then mixed with 7.5 ml anti-PA-tag antibody beads (Fujifilm-Wako) for 2 h (for ERGIC-53-PA) or 1 ml Anti-DYDDDK tag antibody beads (MBL) for 2 h (for the ERGIC-53 ΔH34 mutant). The beads were transferred into a gravity flow column (Econo-column, Bio-Rad) and washed with 20 CV of buffer A supplemented with 0.02% (w/v) GDN (Anatrace). The bound proteins of ERGIC-53 WT were eluted with overnight incubation of 2 CV elution buffer (buffer A supplemented with 0.02% GDN, and 0.2 mg/ml PA peptide), followed by repeating 5 min incubation with 1 CV elution buffer. For the ΔH34 mutant, the bound proteins were eluted by repeating the 5 min incubation of 1 CV elution buffer (buffer A supplemented with 0.02% GDN and 0.1 mg/ml DYKDDDDK peptide). The eluted protein was concentrated using Amicon Ultra-15 centrifugal filters (100 kDa cut-off, Merck-Millipore). After treatment with 1 mM diamide for 15 min to completely oxidize the protein, the concentrated protein was applied onto a Superose 6 Increase 10/300 GL column (GE Healthcare) equilibrated with SEC buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 10 mM CaCl₂ and 0.02% GDN). For the preparation of the ERGIC-53-MCFD2 complex, the ERGIC-53 sample was mixed with excess purified MCFD2 (described below) for 10 min, followed by diamide treatment for 15 min, and then the complex sample was injected into the same column. The peak fractions were concentrated using Amicon-Ultra-4 and –0.5 centrifugal filters (100 kDa cut-off) for cryo-EM samples.

The pET15b plasmid of MCFD2 was introduced into E-cos Escherichia coli BL21 (DE3) (Nippon gene). Gene expression was induced with 0.1 mM isopropyl β-D-1-thiogalactopyranoside (Fujifilm-Wako) at an OD₆₀₀ of ~0.5, with further cultivation for 24 h at 293 K before harvesting. Cells were washed with 50 mM Tris-HCl pH 8.5 and then stored at –80 °C. The cell pellet was resuspended in buffer A (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 5 mM CaCl₂) supplemented with a protease inhibitor cocktail (Nacalai Tesque) and was sonicated on ice for 5 min. After the removal of cell debris by centrifugation (17000×g × 30 min at 4 °C), the supernatant was applied to a Ni-NTA resin (Qiagen). After 10 CV washes with buffer A supplemented with 20 mM imidazole, the bound protein was eluted with buffer A supplemented with 200 mM imidazole. The eluted protein was concentrated and using Amicon Ultra-15 centrifugal filters (3 kDa cut-off,

Merck-Millipore), and then buffer was exchanged to buffer B (50 mM Tris-HCl pH 8.0, 5 mM CaCl₂). The protein was loaded onto an anion-exchange column (MonoQ 10/100 GL, GE Healthcare) equilibrated with buffer B, and eluted with a 0.1–0.4 M linear gradient of NaCl. The eluted protein was concentrated and treated with thrombin (Nacalai Tesque) at 4 °C overnight to cleave the His tag. The sample was further purified with a size-exclusion column equilibrated with 20 mM Tris-HCl, 150 mM and 5 mM CaCl₂. The eluted protein was concentrated and stored at –80 °C.

SEC-MALS/SAXS

Size-exclusion chromatography coupled with multiangle light scattering (SEC-MALS) analysis was performed by using a high-performance liquid chromatography (HPLC) system, Alliance 2695 (Waters) with DAWN HELEOS II (Wyatt Technology) and 2414 Refractive Index (RI) detector and 2489 UV/Visible detector (Waters). Samples (1.6–1.7 mg/ml × 10 μl) were loaded on a Superdex 200 increase 3.2/300 column (Cytiva) equilibrated with SEC buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 10 mM CaCl₂, and 0.02% GDN) under a flow rate at 0.1 ml/min. The conjugate analysis of the molecular masses by combining MALS-UV-RI signals was carried out by ASTRA 6.1 (Wyatt Technology) with a dn/dc value of 0.185 mL/g for protein and 0.17 mL/g for GDN, and an extinction coefficient (1 mg/ml) of 0.766 for protein and 0.01014 for GDN. Although the dn/dc value of GDN was arbitrarily used based on the other detergents, this value did not affect the calculation of the molecular masses of the protein.

Size-exclusion chromatography combined with small-angle X-ray scattering (SEC-SAXS) was performed on the BL-15A2 beamline at the photon factory, KEK (Tsukuba, Japan)⁶⁰. The sample (1.6–1.7 mg/ml × 50 μl) was loaded on a Superdex 200 increase 3.2/300 column equilibrated with the SEC buffer using an HPLC system Nexera-i (SHIMADZU) under the flow rate at 0.03 ml/min. A fiber spectrometer, QEpro (Ocean Insight), mounted at an angle of 45° to the sample cell, was also utilized to obtain the concentration for each frame. The sample-to-detector distance was set to 2650 mm, and the measured X-ray wavelength was adjusted to 1.0 Å. The 2D scattering intensity images were recorded using PILATUS3 2M (DECTRIS) with an exposure time of 3 s. UV-visible absorption spectra were also recorded at 3-s intervals with a 1-s integration time. All scattering images were azimuthally averaged to convert the one-dimensional scattering intensity data, and then the subtraction of the background profile was also calculated. The absolute intensity calibration was performed using water as a standard. These processes were carried out using SAngler⁶¹. Scattering profiles above half of the elution peaks were averaged by using MOLASS⁶². The radius of gyration (R_g) and the maximum molecular dimension (D_{max}) from the Guinier approximation, and the pair distribution function ($P(r)$) were calculated by AUTORG and PRIMUSqt from ATSAS, respectively (Supplementary Fig. 1h)⁶³.

Colorimetric analysis

To detect the metal ions contained in the purified sample of MCFD2, PAR analysis was conducted⁶⁴. The purified sample of MCFD2 (50 μM) was denatured with 4 M guanidine-HCl at pH 7.6 for 5 min at 37 °C. The denatured sample was mixed with a freshly prepared solution of (4-(2-pyridylazo) resorcinol (PAR) (Tokyo Chemical industry) (final 100 μM), and then the UV-vis spectrum was immediately measured with a Hitachi UV3900. Note: Due to the relatively weak affinity, the amount of the bound metal ions (probably Zn²⁺ or Ni²⁺) in the purified MCFD2 samples varied between each purification.

Isothermal titration calorimetry ITC

Isothermal titration calorimetry (ITC) experiments were carried out using the MicroCal™ iTC200 calorimeter (Malvern) at 293 K. After a 0.4 μL initial injection, 2 μL of ZnCl₂ solution (500 μM) was injected into the EDTA-treated MCFD2 solution (15 μM) in 20 mM Tris-HCl pH 7.5,

150 mM NaCl, and 0.2 mM CaCl₂, at 180 s intervals with stirring 750 rpm. The data analysis was performed with Microcal Origin Software (version 7.0) using a one-site binding model. The experiments were repeated at least twice with similar results.

Cryo-EM sample preparation and data collection

The purified samples of full-length ERGIC-53 with or without MCFD2 (4–5 mg/ml) were mixed with 3 mM Fluorinated Fos-Choline-8 (Anatrace), and then 3 μl aliquot of the sample was applied to glow-discharged Quantifoil R1.2/1.3 Au 300 mesh grids. The grids were blotted for 4 s and immersed in liquid ethane using Vitrobot Mark IV systems (FEI/Thermo Fisher) operated at 4 °C and 100% humidity. The grids were initially screened using Talos Arctica (FEI) with a K2 direct electron detector (Gatan). The best grids were loaded to a Titan Krios (FEI) microscope operated at 300 kV and equipped with a Gatan Quantum-LS Energy Filter (GIF) and a Gatan K3 BioQuantum direct electron detector. Movies were automatically collected using SerialEM software⁶⁵.

For the sample of the ERGIC-53 ΔH34-MCFD2 complex (1.7 mg/ml), two rounds of 3 μl aliquot loading and manual blotting were performed, and then a 3 μl aliquot of the sample was applied to glow-discharged Quantifoil R1.2/1.3 Au 200 mesh grids, as previously reported⁶⁶. The grids were blotted for 5 s and immersed in liquid ethane using Vitrobot Mark IV systems (FEI/Thermo Fisher) operated at 4 °C and 100% humidity. Grid observation and data collection were performed with a CRYO ARM™ 300II (JEOL) operated at 300 kV and equipped with a JEOL in-column Omega energy filter and a Gatan K3 BioQuantum detector. Movies were automatically collected using SerialEM software⁶⁵. The data collection parameters are summarized in Supplementary Table 2.

EM image processing

Movies were aligned using beam-induced motion correction implemented in RELION v3.1 and 4.0^{67,68}, and contrast transfer function parameters were estimated using Patch-based CTF estimation in CryoSPARC⁶⁹. The following image processing was mainly performed by using CryoSPARC. Bayesian polishing was performed by RELION v3.1 and 4.0 with the help of UCSF pyem tool⁷⁰ for data conversion from CryoSPARC to Relion.

For the sample of ERGIC-53 in complex with MCFD2, 6345 movies were collected. For analysis of the head and TM regions, blob-based autopicking and subsequent 2D classification were performed by using 800 movies to create templates for template-based autopicking in CryoSPARC. A total of 2,132,031 particles were automatically picked by the template picker in CryoSPARC and extracted at a pixel size of 3.32 Å followed by 2D classification. Good classes with clear 2D average images of the head region were selected and subjected to ab initio modeling and subsequent heterogeneous refinements. Particles belonging to forms A and B were re-extracted at a pixel size of 0.996 Å and subjected to non-uniform (NU) refinement⁷¹. The refined particles were subjected to Bayesian polishing, and then NU refinement with global CTF and particle defocus refinement. Consequently, the final maps of forms A and B were refined to 2.53 and 2.59 Å resolution, respectively. The maps were sharpened and locally filtered based on the local resolution in CryoSPARC.

For its full-length particle analysis, full-length ERGIC-53 particles were manually picked from 50 micrographs. Particles showing full-length-like structures were selected from 2D classification and subjected to Topaz training⁴⁰. A trained model of Topaz was further trained by repeating autopicking by Topaz and 2D classification using 50, 100, and 600 micrographs. Subsequently, full-length particles were picked by Topaz from all micrographs, and extracted at a pixel size of 2.49 Å. 2D classification generated clear 2D average images of full-length ERGIC-53 particles. Particles belonging to the straight conformation were selected and subjected to homogeneous refinement and local refinement. The final map was refined to 6.8 Å resolution.

For the ERGIC-53 Δ H34-MCFD2 complex, 6175 movies were collected, and CTF parameters were estimated with CTFFIND4. Particles were initially picked using Laplacian-of-Gaussian (LoG) based autopicking in RELION 4 to create 2D templates. A total of 127,9058 particles were picked and extracted at a pixel size of 3.142 Å and subjected to 2D classification. Good 2D classes were selected and subjected to ab initio modeling and subsequent heterogeneous refinements. The best particles were re-extracted at a pixel size of 1.3791 Å and subjected to homogenous refinement. The refined particles were subjected to Bayesian polishing, CTF refinement, defocus refinement, and homogenous refinement. The final map (consensus map) was reconstructed at 3.51 Å resolution.

To gain insight into the structural viability of all the obtained EM maps, the refined particles were subjected to 3D variability analysis in CryoSPARC⁴². In the head region analysis, particles were further classified by cluster analysis in 3DVA (12 classes), resulting in the determination of four substate structures at 3.3–3.4 Å resolution.

Model building, refinement, and visualization

For the CRD region and MCFD2, the previously published crystal structure of the CRD and MCFD2 complex (PDB ID 3WNX) was docked in the EM map. The S-H1 and S-H2 regions were manually built by using Coot⁷². The models were refined with Servalcat⁷³ and phenix.real_space_refine⁷⁴ and further manually corrected with Coot. For the building of the full-length model, partial tetrameric models of S-H3-SH4, and S-H4-TM were predicted by using Colabfold with AF2 multimer⁷⁵. Then partial models were assembled and manually docked in the EM map. The model was further refined with phenix_real_space refined as rigid bodies. Structural figures were prepared with PyMOL (<http://www.pymol.org>) and Chimera X⁷⁶. Sequence alignment was performed and visualized with Jalview⁷⁷. Prediction and analysis of coiled-coils was performed with Deepcoil2 and SamCC in MPI bioinformatics Toolkit⁷⁸.

Secretion assay

HEK293T WT or ERGIC-53 KO cells (6×10^5 cells) were plated on six-well plates and transfected with 100 ng of the plasmid expressing FV-HiBiT/AAT-HiBiT and 200 ng of the plasmid expressing ERGIC-53 WT or each of its mutant. After 24 h incubation, the medium was removed, and cells were washed with Expi293 medium and incubated in 1 ml of Expi293 medium for 4 h. In the Zn²⁺ supplement conditions, 10 μ M ZnCl₂ was added in the medium. After harvest of the conditioned mediums (CMs), the remained cells were washed with PBS and lysed in a lysis buffer (PBS, 1% Triton (Thermo Fisher), and protease inhibitor cocktail (Nakarai Tesque), and then briefly sonicated. The total volume of the cell lysates was adjusted to the same volume as the CMs. The amounts of secreted and intracellular HiBiT-tagged FV/AAT proteins were quantified by using the Nano-Glo HiBiT Extracellular Detection System or Nano-Glo HiBiT lytic assay system, respectively. Luminescence from HiBiT-tagged proteins was measured with GloMax Navigator (Promega). Luminescence from CMs and lysates of nontransfected cells were used for background correction.

Secreted and intracellular HiBiT-tagged FV or AAT proteins, as well as expressed ERGIC-53 were also detected by Western blotting with the Nano-Glo[®] HiBiT Blotting System (LgBiT protein) and anti-Factor V (Abcam, ab108614, 1:3300) or anti-LMAN1 (Abcam, ab125006, 1:5000) antibodies. Total proteins were detected by stain-free technology (Bio-Rad). The experiments were performed independently at least three times. Statistical analysis was performed with GraphPad prism 9.3, using a two-tailed unpaired *t*-test for the comparison of two groups or one-way ANOVA followed by Dunnett's test for the other comparisons.

Reporting summary

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

Data availability

The cryo-EM density maps have been deposited in the Electron Microscopy Data Bank (EMDB) under accession codes EMD-36467 (form A), EMD-36468 (form B), EMD-36469 (substate A), EMD-36470 (substate B), EMD-36471 (substate C), EMD-36472 (substate D), EMD-36479 (full length) and EMD-36482 (the Δ H34 mutant). The atomic coordinates have been deposited in the Protein Data Bank under accessions 8JP4 (form A), 8JP5 (form B), 8JP6 (substate A), 8JP7 (substate B), 8JP8 (substate C), 8JP9 (substate D), and 8JPG (full length). The following atomic coordinates were used in this study: 3WNX (the crystal structures of the CRD and MCFD2 complex), and 4YGE (Zn-free MCFD2 structure). Raw movies have been deposited in EMPIAR-11645 (the Δ H34 mutant) and EMPIAR-11646 (full-length ERGIC-53). Source data are provided with this paper.

References

- Barlowe, C. K. & Miller, E. A. Secretory protein biogenesis and traffic in the early secretory pathway. *Genetics* **193**, 383–410 (2013).
- Okumura, M., Kadokura, H. & Inaba, K. Structures and functions of protein disulfide isomerase family members involved in proteostasis in the endoplasmic reticulum. *Free Radic. Biol. Med.* **83**, 314–322 (2015).
- Barlowe, C. & Helenius, A. Cargo capture and bulk flow in the early secretory pathway. *Annu. Rev. Cell Dev. Biol.* **32**, 197–222 (2016).
- Zhang, Y. C., Zhou, Y., Yang, C. Z. & Xiong, D. S. A review of ERGIC-53: its structure, functions, regulation and relations with diseases. *Histol. Histopathol.* **24**, 1193–1204 (2009).
- Tang, V. T. & Ginsburg, D. Cargo selection in endoplasmic reticulum-to-Golgi transport and relevant diseases. *J. Clin. Invest.* **133**, e163838 (2023).
- Zheng, C. & Zhang, B. Combined deficiency of coagulation factors V and VIII: an update. *Semin. Thromb. Hemost.* **39**, 613–620 (2013).
- Nichols, W. C. et al. Mutations in the ER-golgi intermediate compartment protein ERGIC-53 cause combined deficiency of coagulation factors V and VIII. *Cell* **93**, 61–70 (1998).
- Nyfelner, B. et al. Identification of ERGIC-53 as an intracellular transport receptor of α 1-antitrypsin. *J. Cell Biol.* **180**, 705–712 (2008).
- Zhang, B. et al. Mice deficient in LMAN1 exhibit FV and FVIII deficiencies and liver accumulation of α 1-antitrypsin. *Blood* **118**, 3384–3391 (2011).
- Zhu, M. et al. Analysis of MCFD2- and LMAN1-deficient mice demonstrates distinct functions in vivo. *Blood Adv.* **2**, 1014–1021 (2018).
- Vollenweider, F., Kappeler, F., Itin, C. & Hauri, H.-P. Mistargeting of the lectin ERGIC-53 to the endoplasmic reticulum of HeLa cells impairs the secretion of a lysosomal enzyme. *J. Cell Biol.* **142**, 377–389 (1998).
- Appenzeller, C., Andersson, H., Kappeler, F. & Hauri, H.-P. The lectin ERGIC-53 is a cargo transport receptor for glycoproteins. *Nat. Cell Biol.* **1**, 330–334 (1999).
- Chen, Y., Hojo, S., Matsumoto, N. & Yamamoto, K. Regulation of Mac-2BP secretion is mediated by its N-glycan binding to ERGIC-53. *Glycobiology* **23**, 904–916 (2013).
- Duellman, T., Burnett, J., Shin, A. & Yang, J. LMAN1 (ERGIC-53) is a potential carrier protein for matrix metalloproteinase-9 glycoprotein secretion. *Biochem. Biophys. Res. Commun.* **464**, 685–691 (2015).
- Anelli, T. et al. Sequential steps and checkpoints in the early exocytic compartment during secretory IgM biogenesis. *EMBO J.* **26**, 4177–4188 (2007).
- Zhang, B. Combined deficiency of factor V and factor VIII is due to mutations in either LMAN1 or MCFD2. *Blood* **107**, 1903–1907 (2006).
- Zhang, B. et al. Bleeding due to disruption of a cargo-specific ER-to-Golgi transport complex. *Nat. Genet.* **34**, 6 (2003).

18. Kawasaki, N. et al. The sugar-binding ability of ERGIC-53 is enhanced by its interaction with MCFD2. *Blood* **111**, 1972–1979 (2008).
19. Zheng, C. et al. Structural characterization of carbohydrate binding by LMAN1 protein provides new insight into the endoplasmic reticulum export of factors V (FV) and VIII (FVIII). *J. Biol. Chem.* **288**, 20499–20509 (2013).
20. Guy, J. E., Wigren, E., Svård, M., Härd, T. & Lindqvist, Y. New insights into multiple coagulation factor deficiency from the solution structure of human MCFD2. *J. Mol. Biol.* **381**, 941–955 (2008).
21. Tempio, T. et al. A virtuous cycle operated by ERp44 and ERGIC-53 guarantees proteostasis in the early secretory compartment. *iScience* **24**, 102244 (2021).
22. Fu, Y.-L., Zhang, B. & Mu, T.-W. LMAN1 (ERGIC-53) promotes trafficking of neuroreceptors. *Biochem. Biophys. Res. Commun.* **511**, 356–362 (2019).
23. Huang, Y. et al. An in vitro vesicle formation assay reveals cargo clients and factors that mediate vesicular trafficking. *Proc. Natl Acad. Sci. USA* **118**, e2101287118 (2021).
24. Klaus, J. P. et al. The intracellular cargo receptor ERGIC-53 is required for the production of infectious arenavirus, coronavirus, and filovirus particles. *Cell Host Microbe* **14**, 522–534 (2013).
25. Zeyen, L., Döring, T., Prange, R. & Hepatitis, B. Virus exploits ERGIC-53 in conjunction with COPII to exit cells. *Cells* **9**, 1889 (2020).
26. Kamiya, Y. Molecular basis of sugar recognition by the human L-type lectins ERGIC-53, VIPL, and VIP₃₆. **283**, 1857–1861 (2008).
27. Itin, C., Roche, A. C., Monsigny, M. & Hauri, H. P. ERGIC-53 is a functional mannose-selective and calcium-dependent human homologue of leguminous lectins. *Mol. Biol. Cell* **7**, 483–493 (1996).
28. Itin, C., Schindler, R. & Hauri, H. P. Targeting of protein ERGIC-53 to the ER/ERGIC/cis-Golgi recycling pathway. *J. Cell Biol.* **131**, 57–67 (1995).
29. Kappeler, F., Klopfenstein, D. R. C., Foguet, M., Paccaud, J.-P. & Hauri, H.-P. The recycling of ERGIC-53 in the early secretory pathway ERGIC-53 carries a cytosolic endoplasmic reticulum-exit determinant interacting with COPII. *J. Biol. Chem.* **272**, 31801–31808 (1997).
30. Nufer, O. et al. Role of cytoplasmic C-terminal amino acids of membrane proteins in ER export. *J. Cell Sci.* **115**, 619–628 (2002).
31. Lahtinen, U., Svensson, K. & Pettersson, R. F. Mapping of structural determinants for the oligomerization of p58, a lectin-like protein of the intermediate compartment and cis-Golgi. *Eur. J. Biochem.* **260**, 392–397 (1999).
32. Schweizer, A., Fransen, J. A., Bächli, T., Ginsel, L. & Hauri, H. P. Identification, by a monoclonal antibody, of a 53-kD protein associated with a tubulo-vesicular compartment at the cis-side of the Golgi apparatus. *J. Cell Biol.* **107**, 1643–1653 (1988).
33. Neve, E. P. A., Lahtinen, U. & Pettersson, R. F. Oligomerization and intercellular localization of the glycoprotein receptor ERGIC-53 is independent of disulfide bonds. *J. Mol. Biol.* **354**, 556–568 (2005).
34. Satoh, T., Suzuki, K., Yamaguchi, T. & Kato, K. Structural basis for disparate sugar-binding specificities in the homologous cargo receptors ERGIC-53 and VIP36. *PLoS ONE* **9**, e87963 (2014).
35. Satoh, T. et al. Crystallographic snapshots of the EF-hand protein MCFD2 complexed with the intracellular lectin ERGIC-53 involved in glycoprotein transport. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **76**, 216–221 (2020).
36. Nishio, M. et al. Structural basis for the cooperative interplay between the two causative gene products of combined factor V and factor VIII deficiency. *Proc. Natl Acad. Sci. USA* **107**, 4034–4039 (2010).
37. Wigren, E., Bourhis, J.-M., Kursula, I., Guy, J. E. & Lindqvist, Y. Crystal structure of the LMAN1-CRD/MCFD2 transport receptor complex provides insight into combined deficiency of factor V and factor VIII. *FEBS Lett.* **584**, 878–882 (2010).
38. Velloso, L. M., Svensson, K., Schneider, G., Pettersson, R. F. & Lindqvist, Y. Crystal structure of the carbohydrate recognition domain of p58/ERGIC-53, a protein involved in glycoprotein export from the endoplasmic reticulum. *J. Biol. Chem.* **277**, 15979–15984 (2002).
39. Receveur-Bréchet, V. & Durand, D. How random are intrinsically disordered proteins? A small angle scattering perspective. *Curr. Protein Pept. Sci.* **13**, 55–75 (2012).
40. Bepler, T. et al. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat. Methods* **16**, 1153–1160 (2019).
41. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
42. Punjani, A. & Fleet, D. J. 3D variability analysis: resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *J. Struct. Biol.* **213**, 107702 (2021).
43. Yagi, H. et al. Improved secretion of glycoproteins using an N-glycan-restricted passport sequence tag recognized by cargo receptor. *Nat. Commun.* **11**, 1368 (2020).
44. Lo, M. N., Damon, L. J., Wei Tay, J., Jia, S. & Palmer, A. E. Single cell analysis reveals multiple requirements for zinc in the mammalian cell cycle. *eLife* **9**, e51107 (2020).
45. Yamada, T. et al. A novel missense mutation causing abnormal LMAN1 in a Japanese patient with combined deficiency of factor V and factor VIII. *Am. J. Hematol.* **84**, 738–742 (2009).
46. Zhang, Y., Liu, Z. & Zhang, B. Separate roles of LMAN1 and MCFD2 in ER-to-Golgi trafficking of FV and FVIII. *Blood Adv.* **7**, 1286–1296 (2023).
47. Zhang, Y. et al. LMAN1–MCFD2 complex is a cargo receptor for the ER–Golgi transport of α 1-antitrypsin. *Biochem. J.* **479**, 839–855 (2022).
48. Ishii, K. et al. pH-Dependent assembly and segregation of the coiled-coil segments of yeast putative cargo receptors Emp46p and Emp47p. *PLoS ONE* **10**, e0140287 (2015).
49. Kato, K., Furuhashi, T., Kato, K., Oda, A. & Kurimoto, E. The assembly mechanism of coiled-coil domains of the yeast cargo receptors Emp46p/47p and the mutational alteration of pH-dependency of complex formation. *J. Biochem.* **163**, 441–446 (2018).
50. Yerushalmi, N. et al. ERGL, a novel gene related to ERGIC-53 that is highly expressed in normal and neoplastic prostate and several other tissues. *Gene* **265**, 55–60 (2001).
51. Geiger, T., Wehner, A., Schaab, C., Cox, J. & Mann, M. Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying expression of most proteins. *Mol. Cell. Proteomics* **11**, M111.014050 (2012).
52. Nufer, O., Mitrovic, S. & Hauri, H.-P. Profile-based data base scanning for animal L-type lectins and characterization of VIPL, a novel VIP36-like endoplasmic reticulum protein. *J. Biol. Chem.* **278**, 15886–15896 (2003).
53. Kowada, T. et al. Quantitative imaging of labile Zn²⁺ in the golgi apparatus using a localizable small-molecule fluorescent probe. *Cell Chem. Biol.* **27**, 1521–1531.e8 (2020).
54. Amagai, Y. et al. Zinc homeostasis governed by Golgi-resident ZnT family members regulates ERp44-mediated proteostasis at the ER–Golgi interface. *Nat. Commun.* **14**, 2683 (2023).
55. Liu, R. et al. Organelle-level labile Zn²⁺ mapping based on targetable fluorescent sensors. *ACS Sens.* **7**, 748–757 (2022).
56. Watanabe, S. et al. Zinc regulates ERp44-dependent protein quality control in the early secretory pathway. *Nat. Commun.* **10**, 603 (2019).
57. Brown, G. D., Willment, J. A. & Whitehead, L. C-type lectins in immunity and homeostasis. *Nat. Rev. Immunol.* **18**, 374–389 (2018).
58. Mnich, M. E., van Dalen, R. & van Sorge, N. M. C-type lectin receptors in host defense against bacterial pathogens. *Front. Cell. Infect. Microbiol.* **10**, 309 (2020).

59. Miller, M. H. et al. LMAN1 is a receptor for house dust mite allergens. *Cell Rep.* **42**, 112208 (2023).
60. Takagi, H. et al. New high-brilliance small angle x-ray scattering beamline, BL-15A2 at the photon factory. *AIP Conf. Proc.* **2054**, 060038 (2019).
61. Shimizu, N. et al. Software development for analysis of small-angle x-ray scattering data. *AIP Conf. Proc.* **1741**, 050017 (2016).
62. Yonezawa, K., Takahashi, M., Yatabe, K., Nagatani, Y. & Shimizu, N. MOLASS: software for automatic processing of matrix data obtained from small-angle X-ray scattering and UV-visible spectroscopy combined with size-exclusion chromatography. *Biophys. Physicobiol.* **20**, e200001 (2023).
63. Manalastas-Cantos, K. et al. ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis. *J. Appl. Crystallogr.* **54**, 343–355 (2021).
64. McCall, K. A. & Fierke, C. A. Colorimetric and fluorimetric assays to quantitate micromolar concentrations of transition metals. *Anal. Biochem.* **284**, 307–315 (2000).
65. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
66. Snijder, J. et al. Vitrification after multiple rounds of sample application and blotting improves particle density on cryo-electron microscopy grids. *J. Struct. Biol.* **198**, 38–42 (2017).
67. Kimanius, D., Dong, L., Sharov, G., Nakane, T. & Scheres, S. H. W. New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochem. J.* **478**, 4169–4185 (2021).
68. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
69. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
70. Asarnow, D., Palovcak, E. & Cheng, Y. asarnow/pyem: UCSF pyem v0.5. <https://doi.org/10.5281/zenodo.3576630> (2019).
71. Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat. Methods* **17**, 1214–1221 (2020).
72. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
73. Yamashita, K., Palmer, C. M., Burnley, T. & Murshudov, G. N. Cryo-EM single-particle structure refinement and map calculation using Servalcat. *Acta Crystallogr. Sect. Struct. Biol.* **77**, 1282–1291 (2021).
74. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. Sect. Struct. Biol.* **74**, 531–544 (2018).
75. Mirdita, M. et al. ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).
76. Pettersen, E. F. et al. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci. Publ. Protein Soc.* **30**, 70–82 (2021).
77. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).
78. Gabler, F. et al. Protein sequence analysis using the MPI bioinformatics toolkit. *Curr. Protoc. Bioinformatics* **72**, e108 (2020).
- Medicine in Tohoku University; and Y. Amagai and M. Inoue for help in culturing cells and cell biology experiments. This work was supported by Grants-in-Aid for Scientific Research (C) (JP18K06075 to S.W.), Grant-in-Aid for Transformative Research Areas (JP21H05253 to K.I. and S.W.), Grant-in-Aid for Challenging Research (Exploratory) (JP23K18193 to S.W.) from the MEXT of Japan, a research funding from AMED-CREST (21gm1410006h0001) to K.I., a research grant from the Naito foundation (to S.W.), a research grant from Mochida memorial foundation for medical and pharmaceutical research (to S.W.), a research grant from AMED (JP22gm6410026 to Y.K.), and Platform Project for Supporting Drug Discovery and Life Science Research (Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS)) from AMED under grant number JP21am0101115 (support number 1025) JP21am0101071 (support number 2078), JP23ama121012 (support number 5756), JP21am0101095 and JP23ama121038.

Author contributions

S.W. designed the research and performed almost all experiments, cryo-EM data collection at Tohoku University, image processing, and model building and refinement. Y.K. performed grid preparation/optimization and cryo-EM data collection at the University of Tokyo under the supervision of O.N. K.Y. performed SEC-MALS/SAXS analysis at KEK under the supervision of N.S. M.I. performed plasmid construction and secretion assays. S.W. and K.I. wrote the manuscript with the help of all other authors. S.W. and K.I. supervised this work.

Competing interests

O.N. is a co-founder, board member, and scientific advisor for Curreio. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-46747-1>.

Correspondence and requests for materials should be addressed to Satoshi Watanabe or Kenji Inaba.

Peer review information *Nature Communications* thanks Koichi Kato and the other, anonymous, reviewer for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://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) 2024

Acknowledgements

We would like to thank staff scientists at the University of Tokyo's cryo-EM facility, especially, A. Tsutsumi, K. Kobayashi, H. Yanagisawa, M. Kikkawa, and R. Danev; T. We also thank T. Yokoyama, K. Nanatani, J. Inoue, S. Koshiba, and M. Yamamoto for the management of the cryo-EM facility at Advanced Research Center for Innovations in Next-Generation