



Cryo-EM structure determination of small proteins by nanobody-binding scaffolds (Legobodies)

Xudong Wu^{a,b,1} and Tom A. Rapoport^{a,b,1}

^aHHMI, Harvard Medical School, Boston, MA 02115; and ^bDepartment of Cell Biology, Harvard Medical School, Boston, MA 02115

Contributed by Tom A. Rapoport, August 26, 2021 (sent for review August 13, 2021; reviewed by Yifan Cheng and Gabriel C. Lander)

We describe a general method that allows structure determination of small proteins by single-particle cryo-electron microscopy (cryo-EM). The method is based on the availability of a target-binding nanobody, which is then rigidly attached to two scaffolds: 1) a Fab fragment of an antibody directed against the nanobody and 2) a nanobody-binding protein A fragment fused to maltose binding protein and Fab-binding domains. The overall ensemble of ~120 kDa, called Legobody, does not perturb the nanobody–target interaction, is easily recognizable in EM images due to its unique shape, and facilitates particle alignment in cryo-EM image processing. The utility of the method is demonstrated for the KDEL receptor, a 23-kDa membrane protein, resulting in a map at 3.2-Å overall resolution with density sufficient for de novo model building, and for the 22-kDa receptor-binding domain (RBD) of SARS-CoV-2 spike protein, resulting in a map at 3.6-Å resolution that allows analysis of the binding interface to the nanobody. The Legobody approach thus overcomes the current size limitations of cryo-EM analysis.

small protein | cryo-EM | nanobody | Legobody | scaffold

Single-particle electron cryo-microscopy (cryo-EM) has become the method of choice for the determination of protein structures. Cryo-EM analysis has several advantages over X-ray crystallography or NMR (1), but the method becomes increasingly challenging for smaller proteins. Large molecules are relatively easy to identify in noisy low-dose images of vitrified samples and have sufficient contrast and features to determine their orientation and position for alignment and averaging. The structural analysis of small particles (~100 kDa or less) is much more difficult. Small targets often lack recognizable shape features that can facilitate initial image alignment at low resolution. Without symmetry, small particles require optimal conditions, such as a highly homogeneous sample, rigid protein conformation, and random particle distribution in thin ice, conditions that are difficult to achieve with most samples (2). However, structure determination of small proteins is of great interest, as most proteins have sizes below 100 kDa and ~50% are smaller than 50 kDa, including many membrane proteins and proteins of medical importance. It is thus a major goal in the field to expand the use of cryo-EM to the routine analysis of small proteins.

One approach to employ cryo-EM for small proteins is based on phase contrast methods, such as the use of Volta phase plates. This method has been used to determine the structure of streptavidin, a protein of 52 kDa, at 3.2-Å resolution (3). However, the structure of this protein could be determined even without phase plates (4), likely because streptavidin forms rigid tetramers and the particles display a near-perfect distribution in very thin ice, which greatly facilitates structural analysis.

An alternative strategy is to make the target protein larger, either by fusing it to another protein or by using a binding partner. In either case, high rigidity of the added scaffold itself and its rigid connection to the target protein are required to facilitate particle alignment and averaging in cryo-EM images.

The fusion approach has been tried with different scaffolds. For example, in a recent study, the BRIL domain was fused into a loop of a small GPCR protein by extending helices on both sides of the fusion point; the size of the scaffold was further increased by

a Fab directed against the BRIL domain (5). However, this approach is limited to proteins containing suitable α -helices; their extension has to be customized for each new target to generate a rigid connection, which is difficult to achieve without prior knowledge of the target structure.

More promising is the use of a binding partner that can be selected with a screening platform, such as modified ankyrin repeat proteins (DARPs), Fab fragments of antibodies, or nanobodies. In recent studies, DARPs selected against GFP were grafted onto large scaffolds and used to visualize GFP by cryo-EM (6, 7). However, the intrinsic conformational heterogeneity of DARPs limits their potential to achieve high-resolution structures of small proteins (7), and so far only a few DARPs have been selected against membrane proteins. Fab fragments can be used as a fiducial marker to facilitate image alignment in cryo-EM images (8), but they have been mainly used in X-ray crystallography. Only a few examples of their application for cryo-EM analysis have been reported (9–11), in part because the selection of appropriate Fabs is not trivial. In addition, the size of the Fabs (~50 kDa) and the existence of a somewhat flexible hinge region between the two subdomains still make structural analysis challenging.

Nanobodies, derived from single-chain antibodies of camelids, are also becoming popular as versatile binding partners of target proteins. Nanobodies have several attractive features. They form rigid structures that can bind to diverse shapes of target proteins, such as loops, convex surfaces, and cavities (12). They can bind to small exposed surfaces, which may not be accessible to Fab

Significance

Structure determination by cryo-EM is difficult or impossible to apply to proteins smaller than ~100 kDa, excluding many membrane proteins and proteins of pharmaceutical importance from the analysis. Here, we report on a general method that allows structure determination of small proteins. The method is based on the availability of a nanobody to a target protein. The nanobody is then rigidly attached to two scaffolds: 1) a Fab fragment of an antibody directed against the nanobody and 2) a nanobody-binding protein A fragment fused to maltose binding protein and Fab-binding domains. We call the overall ensemble Legobody. The method is demonstrated for two small proteins that have sizes of ~22 kDa.

Author contributions: X.W. conceived the idea; X.W. designed research with input from T.A.R.; X.W. performed research; X.W. analyzed data; X.W. and T.A.R. wrote the paper; and T.A.R. supervised the project.

Reviewers: Y.C., University of California San Francisco Medical Center; G.C.L., The Scripps Research Institute

Competing interest statement: Harvard University has filed a patent application on behalf of the authors regarding the nanobody-binding scaffold protein constructs.

This article is a PNAS Direct Submission.

This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

¹To whom correspondence may be addressed. Email: tom_rapoport@hms.harvard.edu or xudong_wu2@hms.harvard.edu.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2115001118/-DCSupplemental>.

Published October 7, 2021.

fragments. Nanobodies can be selected from immunized camels or from large in vitro libraries displayed by phages, yeast cells, or on ribosomes (12, 13), and can be produced in large quantities in a fairly short time. They often lock a protein into a fixed conformation, particularly in the case of membrane proteins, and have been used extensively to determine X-ray structures. The small size of nanobodies (~12 to 15 kDa) limits their direct application in cryo-EM, but the problem might be overcome if one could increase their size with the rigid attachment of a large scaffold. One reported approach is to fuse a scaffold into a loop of the nanobody, generating a “megabody” (14). However, the linker consisted of β -strands between the nanobody and scaffold, which caused some flexibility and limited the use of the scaffold for particle alignment in cryo-EM analysis.

Here, we describe a versatile method that allows cryo-EM analysis of even the smallest protein once a tightly binding nanobody is available. The size of the nanobody is increased to ~120 kDa by two rigidly attached scaffolds. The overall design is reminiscent of a Lego construction, so we propose to call the scaffolds/nanobody ensemble “Legobody.” The utility of the Legobody method is demonstrated by structures of two small proteins (22 kDa and 23 kDa) that are asymmetric monomers and have a size well below the estimated limit for direct cryo-EM single-particle analysis (~40 kDa) (15). The Legobody approach can easily be applied to any target protein and should greatly expand the use of cryo-EM single-particle analysis by overcoming the current size limitations.

Results

Generation of a Nanobody-Binding Fab. Our first nanobody-interacting scaffold is a Fab fragment of an antibody that is directed against a surface present in many nanobodies and not involved in target interaction. To generate such a Fab, we raised monoclonal antibodies in mice against a nanobody (Nb₀) that contains a framework sequence almost identical to that used in two libraries employed for rapid in vitro screening (12, 13). Amino acids in the antigen-interacting complementarity-determining regions (CDRs) of Nb₀ were chosen to minimize the immunogenicity of the antigen-interacting surface (see *SI Appendix, Table S1* for sequence). To select for monoclonals that do not perturb nanobody-antigen interaction, hybridoma clones were screened with a complex of a nanobody against MBP (Nb_{MBP}) and its antigen MBP (12). After several rounds of selection, hybridoma clone 8D3 was obtained, which produced monoclonal antibodies that strongly bind to both Nb₀ and the Nb_{MBP}/MBP complex.

Although it is possible to directly use the antibodies secreted by clone 8D3 to generate Fab fragments, future applications are greatly facilitated if the Fabs can be made recombinantly. To this end, we first determined the DNA sequences of the regions coding for the variable regions of the light and heavy chains of clone 8D3. These sequences were then combined with the sequences coding for the constant regions of murine IgG1, and both genes were expressed together in HEK293 cells. Because the yield of Fabs was rather low, the constant regions of the light and heavy chains were replaced with those of human IgG. The resulting Fab_{8D3} was expressed in HEK293 cells as a secreted protein and purified by nickel-nitrilotriacetic acid (Ni-NTA) chromatography on the basis of a His tag attached to the heavy chain. The yield is ~5 to 8 mg from 1 L of cell culture. Recombinantly purified Fab_{8D3} forms a stable complex with nanobody Nb₀, as shown by comigration of the proteins in size-exclusion chromatography (Fig. 1A).

To identify the exact interaction surface, we determined a crystal structure of the complex of Fab_{8D3} and Nb₀. After confirming that the crystals contained both components (*SI Appendix, Fig. S1A*), a structure of the complex was determined at 1.8-Å resolution (Fig. 1B and *SI Appendix, Table S2*). As expected, Fab_{8D3} binds to a surface of Nb₀ that is distal from the CDRs and contains conserved amino acids present in many nanobodies (Fig. 1C and *SI Appendix, Fig. S1B*). The Fab-interacting amino

acids of Nb₀ are located in loops between β -strands A and B, C and C', E and F, as well as in segments of the β -strands A and G. It should be noted that four amino acids (LEHH) introduced by the cloning of the His tag are also involved in the interaction (Fig. 1C). The extensive interactions between the Fab and nanobody generate a rigid interface, a conclusion supported by the B factor profile of the X-ray structure (*SI Appendix, Fig. S1C*).

Generation of an MBP-Based Scaffold Interacting with Nanobodies.

The second scaffold was developed on the basis of reports that protein A from *Staphylococcus aureus* can bind to nanobodies (16). Protein A contains five repeats of three-helical bundles (domains A–E). All these domains associate with the constant region of IgG antibodies, but also bind with different affinities to the variable region of the heavy chain of some antibodies (human V_{H3} family) (17), a region that is similar in sequence to the common framework of many nanobodies. Consistent with this sequence homology, protein A has been reported to interact with nanobodies in a similar way as with Fabs (18). To identify the strongest binding protein A domain, we fused domain D (PrAD) and the most divergent domains C and E (PrAC and PrAE) through a long, flexible linker to MBP (MBP_L_PrAC, MBP_L_PrAD, and MBP_L_PrAE) and tested these fusions for their interaction with a nanobody. Coelution of the proteins in size-exclusion chromatography showed that all three domains interact with the nanobody, but domain C forms the most stable complex (*SI Appendix, Fig. S2A*).

Next, we grafted domain C of protein A (PrAC; ~6 kDa) to MBP to generate a larger nanobody-binding partner. MBP is frequently used as an N-terminal fusion partner, as it can increase the solubility of its fusion partners. Although fusions can be designed as helical extensions of the C-terminal helix of MBP and have been extensively used in X-ray crystallography (19), such linkers are not rigid enough for cryo-EM analysis. To generate a more rigid connection between PrAC and MBP, we used the “shared helix” approach (20), applying it to a helix of domain C that is not involved in nanobody interaction (*SI Appendix, Fig. S2B*). Residues on one side of this helix were mutated to those of MBP's C-terminal helix that face the core of MBP. The resulting construct MBP_PrAC (Fig. 1D) could be expressed in *Escherichia coli* and purified in large quantities. Like the MBP fusion of PrAC containing a flexible linker (MBP_L_PrAC), MBP_PrAC interacted with the nanobody in pull-down experiments (Fig. 1E).

Legobody Assembly. The nanobody surfaces interacting with Fab_{8D3} and MBP_PrAC are not overlapping (Fig. 2A and B), providing an opportunity to use both scaffolds at the same time. To increase the stability of the ensemble, we fused two Fab-binding domains to the C terminus of MBP_PrAC. One is the domain D of protein A (PrAD), which has been shown to interact with the variable region of the heavy chain of Fabs (21). The other is protein G (PrG), which is known to strongly interact with the constant region of the heavy chain (22). All domains were connected by short linkers, generating a scaffold designated MBP_PrA/G (Fig. 2A and B). To allow the interaction of PrAD with Fab_{8D3}, some residues in the variable region of the heavy chain were mutated based on the crystal structure of a Fab/PrAD complex (PDB code 1DEE), generating Fab_{8D3_2}. Avidity effects increase the binding constants of both the MBP_PrAC and Fab scaffolds for the nanobody, making the overall assembly very stable.

The complex between Fab_{8D3_2} and MBP_PrA/G was assembled before adding a nanobody. All three components comigrated in size-exclusion chromatography (Fig. 2C). Negative-stain EM showed strong structural features for the different parts of the Legobody (Fig. 2D), suggesting overall rigidity of the assembly, an assumption confirmed by subsequent cryo-EM analysis (see below).

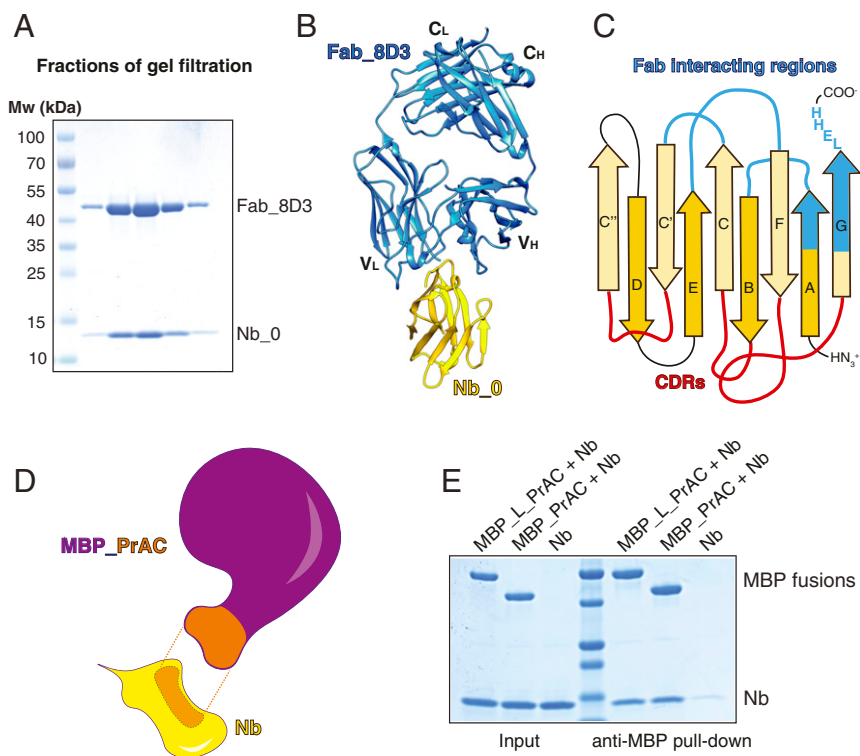


Fig. 1. Generation of two nanobody-interacting scaffolds. (A) Nanobody Nb_0 was mixed with Fab_8D3 and the complex was subjected to gel filtration. Fractions were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie blue staining. (B) Crystal structure of the Nb_0/Fab_8D3 complex shown in cartoon representation. Nb_0 is colored in yellow and Fab_8D3 in blue. (C) Scheme of the nanobody, with β -strands labeled according to convention. CDR loops are colored in red and regions interacting with the Fab in blue. (D) Scheme of the interaction of nanobody (Nb; yellow) with the domain C of protein A (PrAC; orange) grafted onto maltose-binding protein (MBP; purple). The interacting regions between Nb and MBP_PrAC are shown in orange with dashed lines in between. (E) PrAC fused to MBP through a flexible linker (MBP_L_PrAC) or grafted onto MBP (MBP_PrAC) was mixed with Nb. The mixture was incubated with MBP-interacting amylose resin, and the bound material analyzed by SDS-PAGE and Coomassie blue staining. The input corresponds to 50% of the material used for the pull-down.

Case Study I: KDEL Receptor (~23 kDa). To test the utility of the Legobody method, we first chose a small membrane protein, the KDEL receptor. This protein binds to the C-terminal KDEL sequence of luminal endoplasmic reticulum (ER) proteins that have escaped the ER, so that these proteins can be returned from the Golgi to the ER by vesicular transport (23). The KDEL receptor has seven transmembrane segments and a molecular weight of only ~23 kDa. A crystal structure of the KDEL receptor in complex with a tightly binding nanobody has been reported (24).

To generate a sample suitable for cryo-EM analysis, we devised a protocol that should be applicable to many other challenging membrane proteins (Fig. 3A). The solubilized KDEL receptor (KDELR), tagged with a streptavidin-binding peptide, was first immobilized on streptavidin beads. We employed the detergent decyl maltose neopentyl glycol (DMNG), as it resulted in a more homogeneous sample than n-dodecyl- β -D-maltoside (DDM) used for the crystal structure (24). The beads containing KDELR were then incubated with Legobody containing the reported nanobody (Nb_KR) against KDELR (24). Finally, to reduce aggregation during purification, the complex was reconstituted into a nanodisc on the beads. After elution from the beads with biotin, the complex of KDELR, Legobody, and the nanodisc was further purified by size-exclusion chromatography (Fig. 3B). Negative-stain EM showed features for the Legobody and the nanodisc (Fig. 3C).

We next analyzed the complex by cryo-EM. When placed directly onto EM grids, the particles showed severe aggregation and strong preferred orientation, likely caused by denaturation of the molecules at the water-air interface. To alleviate this problem, surface lysine residues were modified with low molecular weight

polyethylene glycol (PEG), a previously introduced method that makes the surface of proteins more hydrophilic and reduces their denaturation on the grids (25). Although the particles still showed some aggregation and preferred orientation (SI Appendix, Fig. S3), the cryo-EM analysis was straightforward, as the size and unique shape of the Legobody greatly facilitated particle picking and two-dimensional (2D) and three-dimensional (3D) classification. The final 3D refinement of the selected particles resulted in a 3D reconstruction with an overall resolution of 3.2 Å and little directional differences in resolution (SI Appendix, Fig. S3 and Table S3).

With the exception of the relatively flexible PrAD domain, all parts of the Legobody and KDELR had well-resolved structural features, allowing even the visualization of the maltose molecule bound to MBP (Fig. 4A and SI Appendix, Fig. S4A). The local resolution ranged from 3.0 to ~4.0 Å and showed that the interactions between the two scaffolds and the nanobody, as well as the connection between PrAC and MBP, are all rigid (Fig. 4B). Because the nanobody is tightly associated with the KDELR, and because the center of alignment of the Legobody is at the position of the nanobody, an excellent map was obtained for the KDELR (Fig. 4C). The local resolution of this part of the map ranged from 3.0 to ~3.5 Å and all amino acid side chains of the KDELR were clearly visible. In addition, several bound phospholipid molecules could easily be identified. The cryo-EM structure of the KDEL/nanobody complex is almost identical to that obtained by X-ray crystallography (SI Appendix, Fig. S4B) (24).

Case Study II: Receptor-Binding Domain of the SARS-CoV-2 Spike Protein (~22 kDa). Our second test protein for the Legobody method was the receptor-binding domain (RBD) of the SARS-CoV-2 spike protein.

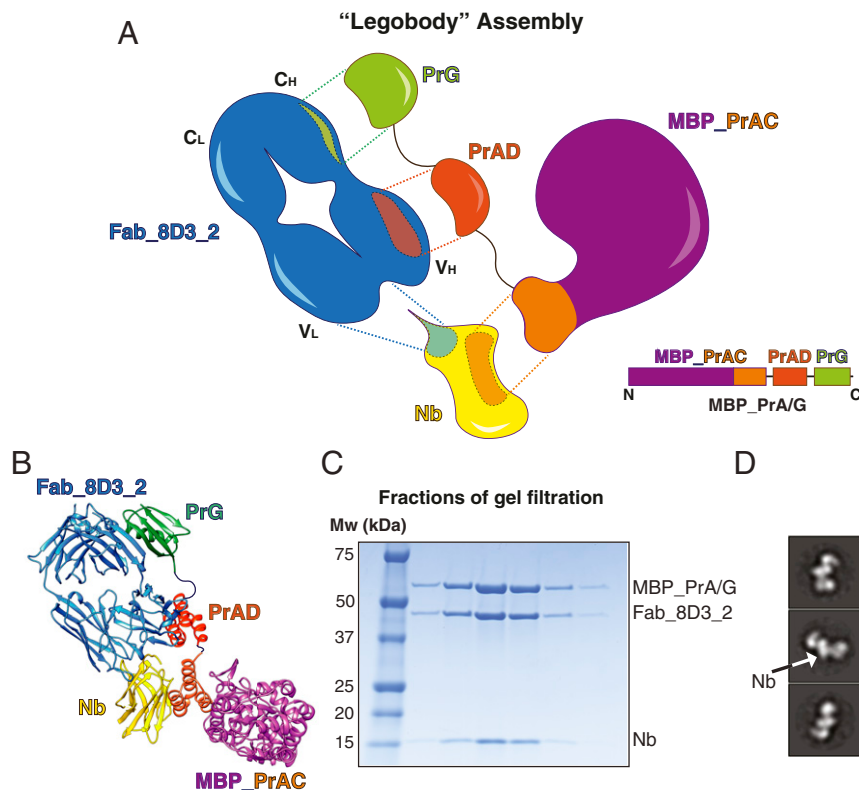


Fig. 2. Assembly and purification of the Legobody. (A) Scheme of the Legobody assembly. Fab_8D3-2, a derivative of Fab_8D3, and MBP_PrAC bind directly to the nanobody (Nb). The Fab-interacting D domain of protein A (PrAD) and protein G (PrG) were fused sequentially to the C terminus of MBP_PrAC via short linkers, generating MBP_PrA/G. Interacting surfaces are colored and highlighted by dashed lines. (B) Model of the assembled Legobody in cartoon representation. The model was assembled from structures of Nb_0/Fab_8D3 (Fig. 1B), Fab/PrAD (PDB code 1DEE), and Fab/PrG (PDB code 1IGC), and by manually building a model for MBP_PrAC from the structures of MBP (PDB code 1ANF) and PrAC (PDB code 4NPD). (C) Legobody assembled with a nanobody was subjected to gel filtration. Fractions were analyzed by SDS-PAGE and Coomassie blue staining. (D) Purified Legobody was analyzed by negative-stain EM. Shown are representative 2D class averages, with density for the nanobody highlighted by an arrow.

The RBD allows SARS-CoV-2 to bind to the ACE2 receptor and infect human cells (26). This interaction is of great medical interest, particularly during the current pandemic, and therefore many RBD-neutralizing nanobodies have been generated (27). The RBD has a molecular weight of only ~22 kDa.

The RBD was expressed in HEK293 cells as a secreted protein and purified by Ni-NTA chromatography on the basis of an attached His tag. The protein was mixed with the preassembled Legobody containing a reported nanobody (Nb_RBD) against the RBD (28). The complex was further purified by size-exclusion chromatography

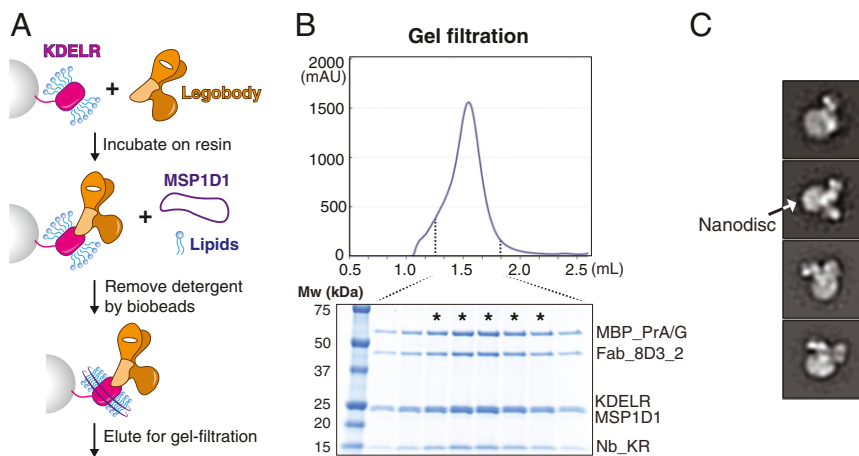


Fig. 3. Purification of a KDELR/Legobody complex. (A) Scheme for the purification of the KDELR/Legobody complex. KDELR solubilized in detergent was immobilized on streptavidin beads. The resin was incubated with purified Legobody. Nanodisc-scaffolding protein MSP1D1 and lipids were added, and the reconstitution of nanodiscs was initiated by the addition of BioBeads. After extensive detergent removal, the complex was eluted from the resin with biotin. (B) The eluted complex of KDELR, Legobody, and nanodisc was subjected to gel filtration and the elution of protein followed by the absorbance at 280 nm (Upper). Fractions between the indicated dashed lines were analyzed by SDS-PAGE and Coomassie blue staining (Lower). Fractions indicated by an asterisk were pooled and used for EM analysis. (C) The purified complex was analyzed by negative-stain EM. Shown are representative 2D class averages, with density for the nanodisc highlighted by an arrow.

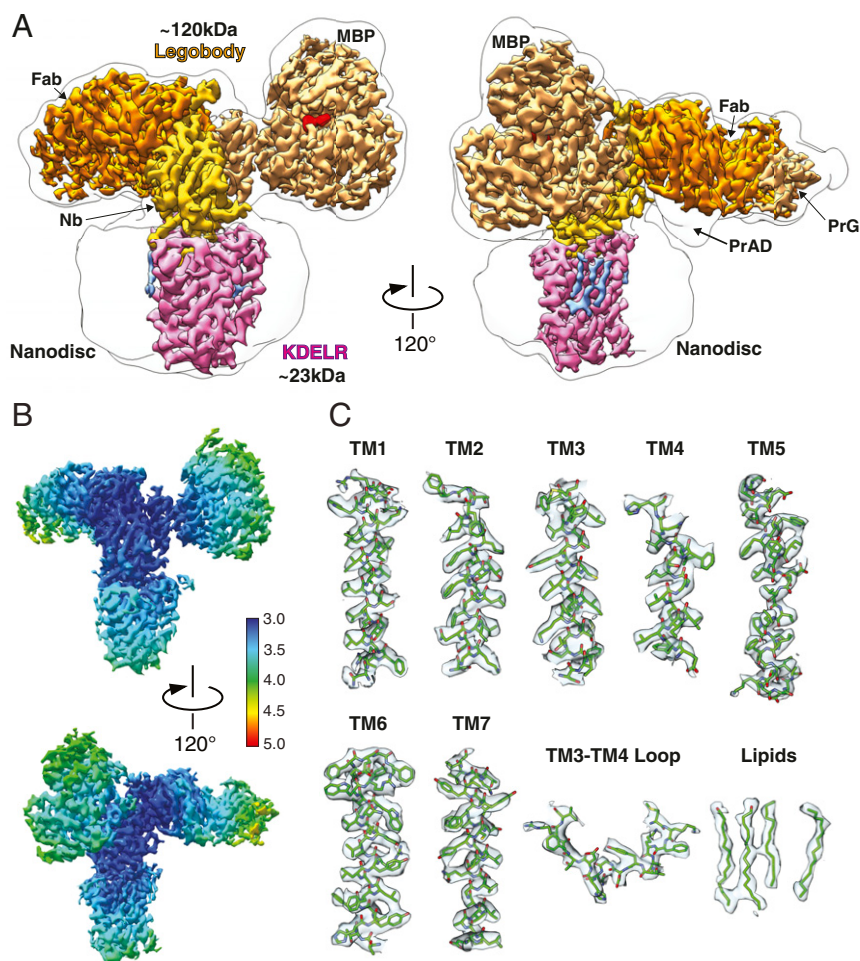


Fig. 4. Cryo-EM structure of KDELr/Legobody complex. (A) Cryo-EM map of the KDELr/Legobody complex in nanodiscs at 3.2-Å resolution, shown in two views. The region of the KDELr is in pink and bound phospholipids are in light blue. The regions of the Legobody components are labeled and shown in different shades of yellow and orange. The density for MBP-bound maltose is shown in red. The solid line shows the outline of the cryo-EM map filtered to 10 Å, which allows the visualization of the nanodisc and flexible PrAD domain (Right). (B) Two views of the final map colored according to local resolution (scale on the Right). (C) Map and fitted model for different segments of the KDELr and for bound phospholipids.

(Fig. 5A) and analyzed by cryo-EM. After 2D and 3D classification, followed by 3D refinement, a map with an overall resolution of 3.6 Å was obtained, again with little directional differences in resolution (Fig. 5B and *SI Appendix, Fig. S5 A–E* and Table S3). The local resolution ranged from ~3.4 to ~4.4 Å and showed good density for the central regions of the Legobody and target protein (Fig. 5C). Specifically, the RBD region showed good side-chain density for amino acids at the interface with the nanobody (Fig. 5D). The cryo-EM structure of the RBD/nanobody complex is almost identical to that obtained by X-ray crystallography (*SI Appendix, Fig. S5F*) (29). Some polypeptide loops distal to the binding interface were invisible in the cryo-EM map (*SI Appendix, Fig. S5F*). The biased particle orientation suggests that this may be caused by the partial denaturation of the RBD at the water/air interface, i.e., by issues unrelated to the Legobody method. The moderate preferred particle orientation observed with the KDELr and the somewhat stronger bias seen with the RBD are likely not caused by the Legobody itself, as the populated angles were different.

General Applicability of the Legobody Method. The Legobody method was developed using nanobodies that have a common framework similar to that used in two *in vitro* libraries (12, 13). The surfaces interacting with the Fab and MBP_PrAC are essentially identical

in these libraries and therefore all selected nanobodies can be used directly for the Legobody method. However, nanobodies generated in alpaca often have a different framework. For example, a nanobody directed against the ALFA tag peptide, which is frequently used to purify or visualize fusion proteins (30), differs significantly in its framework sequence (sequence identity ~75%) and would therefore not allow interaction with our scaffolds. Specifically, nine amino acid residues in the Fab- and PrAC-interacting regions are different from the ones in the common framework. However, when these residues are mutated, the resulting Nb_ALFA can be assembled into a Legobody (*SI Appendix, Fig. S6 A and B*). These mutations do not affect antigen binding, as GST-tagged ALFA peptide was able to pull down the preassembled Legobody containing the modified nanobody (*SI Appendix, Fig. S6B*). We therefore believe that all nanobodies can be used, regardless of whether they are obtained from *in vitro* libraries or from animal species.

Discussion

Here we describe a general method that allows cryo-EM structures to be determined for small proteins. Our Legobody approach thus overcomes current limitations of cryo-EM analysis and greatly expands its use. The method can be applied to any target protein once a tightly binding nanobody is available. The nanobody is assembled into a Legobody by the binding of two scaffolds, a Fab fragment

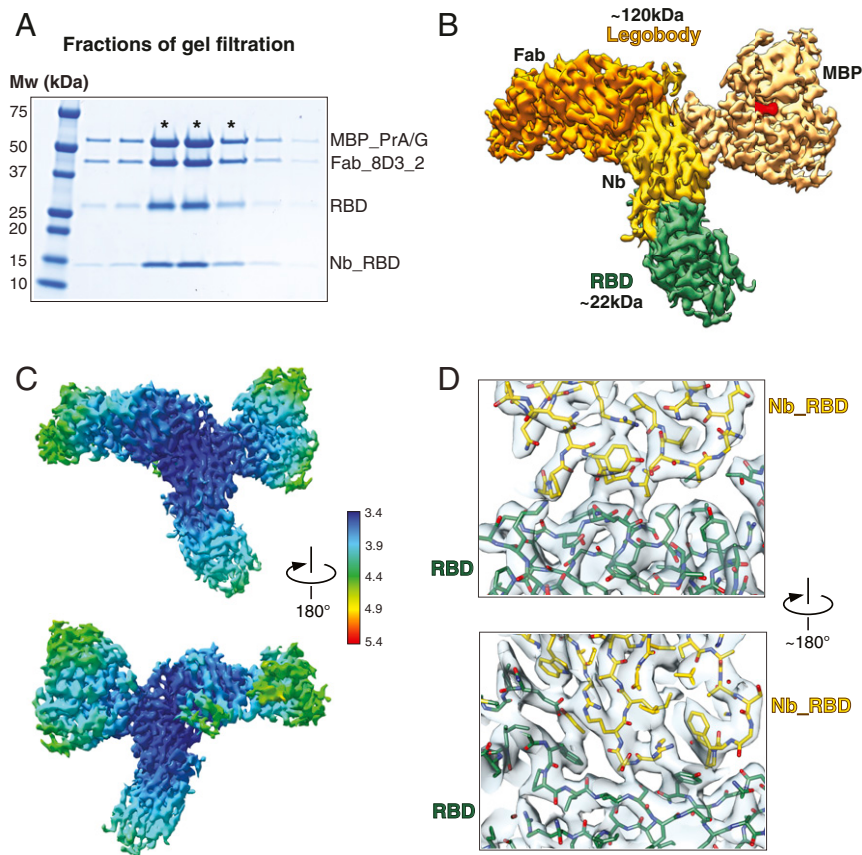


Fig. 5. Cryo-EM structure of a complex of the RBD domain of the SARS-CoV-2 spike protein and Legobody. (A) Purified RBD was mixed with Legobody containing a nanobody directed against the RBD (Nb_RBD). The sample was subjected to gel filtration and fractions were analyzed by SDS-PAGE and Coomassie blue staining. Fractions labeled with an asterisk were pooled and used for EM analysis. (B) Cryo-EM map of the RBD/Legobody complex. The region of the RBD is shown in green. The regions of the Legobody components are labeled and shown in different shades of yellow and orange. The density for MBP-bound maltose is shown in red. (C) Two views of the final map colored according to local resolution (scale on the *Right*). (D) Map and fitted model for the interface between the RBD and nanobody.

and a MBP molecule to which domain C of protein A domain has been grafted (MBP_PrAC). All interactions were designed to be rigid. In addition, Fab-interacting domains were fused to MBP_PrAC to further solidify the complex. The Legobody has a characteristic shape, consisting of two lateral arms, formed by the two scaffolds, and a central lobe, contributed by the nanobody. The overall size (~120 kDa) and shape of the Legobody, and the center of alignment at the position of the nanobody, greatly facilitate all steps of cryo-EM analysis, from particle picking, classifications, to final refinement. We demonstrate the utility of the Legobody method with two examples of small target proteins (KDEL [23 kDa] and the RBD [22 kDa] of the SARS-CoV-2 spike protein).

The membrane protein KDEL poses a particular challenge for cryo-EM analysis, as it is small, has no domains outside membrane, and no symmetry to facilitate particle alignment in EM images. The protein tends to aggregate during purification and on cryo-EM grids at the water–air interface of thin ice. To determine its structure, we not only used the Legobody approach, but also employed two other tricks, which likely are applicable to other challenging membrane proteins. First, we used a purification strategy, in which the KDEL/Legobody complex was incorporated into a nanodisc while bound to beads (Fig. 3A). This strategy reduces aggregation of the receptor and increases its stability in solution. Second, before applying the sample to EM grids, we modified surface lysines with low-molecular-weight PEG. We have used this protocol routinely for other proteins (25, 31, 32), and it often dramatically reduces particle aggregation and preferred particle orientation. Using

standard cryo-EM data analysis, we were able to obtain a high-quality map for the KDEL, with density visible for all amino acid side chains. The map would have allowed straightforward de novo model building.

The KDEL is representative of a large group of membrane proteins, which are of small size and pose similar challenges for cryo-EM analysis. Examples include G protein-coupled receptors, solute carrier transporters, and membrane-embedded enzymes, many of which are of great interest for drug development. The Legobody method now makes all these proteins accessible to cryo-EM analysis. Of course, the target proteins might adopt several physiologically relevant conformations and a nanobody might select only one of them.

The RBD of the SARS-CoV-2 spike protein also presents a challenge for cryo-EM analysis, as it is small and consists mainly of β -strands and extended polypeptide segments, which are more difficult to model into a map than α -helices. The map obtained with the Legobody approach was of good quality, especially at the RBD/nanobody interface, with side-chain density for all interacting amino acids. Because this interface is the region of medical interest, our results show that cryo-EM can be used to optimize RBD-neutralizing nanobodies, which may be important for the quick response to future virus pandemics. By comparison with X-ray crystallography, cryo-EM requires only small amounts of protein and can be performed in a significantly shorter time period.

A possible limitation of the Legobody method could arise from steric clashes between the scaffolds and targets. To test whether this is a serious problem, crystal structures of complexes of nanobody with small monomeric target proteins were aligned with the Legobody structure on the basis of the nanobody. Observed steric clashes are listed in *SI Appendix, Table S4*. For soluble proteins, only three examples were found in which the Fab or MBP_PrAC would clash with the target. For membrane proteins embedded into detergent micelles, nanodiscs, or lipid cubic phase, clashes would sometimes be caused by the PrAD domain or the MBP_PrAC scaffold. No clashes were observed for the Fab, but the number of available structures is too small to exclude their existence in other cases. Clashes with the PrAD domain can be avoided by deleting this domain and connecting MBP_PrAC directly to PrG via a suitable linker. In fact, this domain bound only weakly to its intended binding site on the Fab and should therefore be dispensable even in the original design. Tests for the compatibility of the scaffolds with the nanobody/target interaction are straightforward (similar to the pull-down experiments in *SI Appendix, Fig. S6B*), which could also be used to quickly screen for nanobodies that are compatible with both scaffolds, avoiding nanobodies that would cause steric clashes.

Because of the modular design of the Legobody, only one of the two scaffolds can be attached to the nanobody. For example, for certain targets the Fab fragment might be sufficient. If only the MBP_PrAC scaffold is used, we recommend fusing the nanobody to the N terminus of MBP_PrAC via a flexible linker. However, the use of both scaffolds together not only increases the size of the target but also provides a unique shape that is easily recognizable in EM images. In addition, the both scaffolds may increase the solubility and monodispersibility of the target protein. For example, the KDELR tends to aggregate during purification if either the Fab or the MBP_PrAC scaffold are omitted.

The Legobody used in this study could easily be modified to further increase its molecular mass, stability, and rigidity, which would help to further improve the resolution of the cryo-EM maps. For example, the three-helix bundle of PrAC could be engineered to increase its binding affinity to the nanobody or it could be grafted onto other large proteins. In addition, fusions could be generated with protein L (33) or a modified version of protein M (34), which bind to the Fab at different sites than the fusion partners used in the current Legobody design. We believe that the current Legobodies and their possible variations will make cryo-EM structure determination of small proteins a routine method.

Methods

Purification of Nanobodies. Genes for His-tagged or Strep II-tagged nanobodies were cloned into the pET 26b vector (Novagen). The expression and purification of all His-tagged nanobodies have been described previously (13). For immunization of mice, the His tag was removed by treating purified nanobody Nb_0 with carboxypeptidase A (Sigma) and B (Roche) overnight at 4 °C. The treated nanobodies were passed through a Ni-NTA column (Thermo Fisher) and the flow-through fraction was further purified by size-exclusion chromatography on a S75 increase 10/300 GL column (Cytiva) in 25 mM Hepes pH 7.4, 150 mM NaCl, 5% glycerol. The nanobody against MBP (Nb_MBP) was derived from Sb_MBP#1 (12). Strep II-tagged nanobody Nb_MBP was purified using StrepTactin resin (IBA). The beads were washed with 25 mM Hepes pH 7.4, 150 mM NaCl, and the protein was eluted in 25 mM Hepes pH 7.4, 150 mM NaCl, 2 mM desthiobiotin. For screening of hybridoma cell clones, a complex of Strep II-tagged nanobody Nb_MBP and MBP was used. The eluted Strep II-tagged Nb_MBP protein was mixed with separately purified MBP protein at a molar ratio of 3:1. The mixture was subjected to size-exclusion chromatography on a S200 increase 10/300 GL column (Cytiva) in 25 mM Hepes pH 7.4, 150 mM NaCl, 5% glycerol. The peak fractions of the complex were stored for future use. The nanobodies against KDELR receptor (Nb_KR) and against SARS-CoV-2 spike RBD domain (Nb_RBD) were derived from Syb37 (24) and Sb#45 (28), respectively.

Generating Antibodies Against Nanobodies. Monoclonal antibodies were generated by immunizing mice with purified tagless nanobody Nb_0 at the VGTI Monoclonal Core. Nanobody Nb_0 used for immunization and hybridoma

screening also contained two mutations in the common framework (S7Y and S17Y) intended to increase the chances of selecting Fabs binding near that region, but the crystal structure showed that they were not involved in Fab binding. Hybridoma clones were screened under non-denaturing condition using the purified complex consisting of Strep II-tagged Nb_MBP and MBP. Antibodies secreted by clone 8D3 bound both Nb_0 and the Nb_MBP/MBP complex with high affinity. The 8D3 clone was expanded for further characterization. The sequences of the variable light (V_L) and heavy (V_H) chain regions of the monoclonal antibody were determined by Syd Labs.

Recombinant Expression of Fabs. To increase the yield of recombinant expression of the Fabs in HEK293 cells, the constant regions of the light and heavy chains were replaced by sequences from human Fabs (for sequences, see *SI Appendix, Table S1*). The resulting chimera genes for the light chain and the His-tagged heavy chain of the Fabs were separately cloned into the pCAGEN vector (a gift from Connie Cepko, Harvard Medical School, Boston, MA) (Addgene plasmid #11160) (35). For cotransfection of a 1-L HEK293-*freestyle* (Thermo Fisher) culture, 0.5 mg of both plasmids were incubated with 3 mg of Linear PEI 25K (Polysciences) in 100 mL of Opti-MEM (Thermo Fisher) medium at room temperature for 30 min. The mixture was then added dropwise into the medium containing HEK293-*freestyle* cells to reach a final cell density of 2 million/mL. The cells were cultured at 37 °C for ~12 to 16 h before addition of 10 mM sodium butyrate to boost expression. Medium containing secreted Fabs was harvested ~48 to 62 h posttransfection. Purification of the Fabs was carried out as follows: Harvested medium free of cells was supplemented with 50 mM Tris pH 8.0, 200 mM NaCl, 20 mM imidazole, and 1 μ M NiSO₄. His-tagged Fabs were purified by Ni-NTA chromatography. The beads were washed extensively with 25 mM Hepes pH 7.4, 200 mM NaCl, 20 mM imidazole. Fabs were eluted in 25 mM Hepes pH 7.4, 200 mM NaCl, 300 mM imidazole. Eluted Fabs were concentrated and buffer was exchanged into 25 mM Hepes pH 7.4, 150 mM NaCl, 5% glycerol. Aliquots were snap frozen for future use.

Based on reports that Fabs interact preferentially with domain D of protein A (17) and on a crystal structure of a complex of Fab and this domain (PDB code 1DEE), we introduced several mutations in the variable region of the heavy chain of the partially humanized Fab_8D3 (see above), resulting in Fab_8D3_2 (mutations: G16K, R18L, K19R, I58K, F80Y, and T84N; the sequence is shown in *SI Appendix, Table S1*). Fab_8D3_2 was purified in the same way as the original Fab_8D3.

Determination of a Crystal Structure of the Nb_0/Fab_8D3 Complex. Purified His-tagged Fab_8D3 was mixed with purified His-tagged Nb_0 nanobody at a molar ratio of 1:3. The mixture was treated with carboxypeptidase A (Sigma) overnight at 4 °C to remove the His tags. The sample was subjected to size-exclusion chromatography on a S200 increase 10/300 GL column in 25 mM Hepes pH 7.4, 150 mM NaCl. The peak fractions of the complex were pooled and concentrated to 10 mg/mL and used to set up crystal screens.

Purified Nb_0/Fab_8D3 complex (0.2 μ L of a 10 mg/mL solution) was mixed with 0.2 μ L of mother liquor containing 0.2 M ammonium formate, 20 to 22% wt/vol PEG 3350 using a Mosquito robot (TTP LabTech). Crystals were grown at 4 °C with the hanging drop method over a reservoir of 100 μ L mother liquor and reached full size in about 2 wk. Crystals were cryoprotected before harvest in a solution containing mother liquor supplemented with 25 mM Hepes 7.5 and 18% ethylene glycol. X-ray diffraction data were collected on the 24-ID-E beamline at the Advanced Photon Source (APS). Initial phases were obtained by molecular replacement using crystal structures of Fabs and nanobodies with similar amino acid sequences as search models. In both search models, the CDRs were removed.

Purification of MBP Fusions. The sequences of all MBP fusion proteins are given in *SI Appendix, Table S1*. Based on crystal structures and modeling, we predicted residue Ala405 of MBP_PrAC-His6 to be close to the Fab and therefore mutated it to histidine to boost the interaction. All variants of MBPs were purified as follows: The genes were cloned into the pET28b vector (Novagen) with either an N- or C-terminal His6 tag. The expression was induced by addition of 1 mM isopropylthio- β -galactoside (IPTG) for 4 h at 37 °C. The cells were lysed by sonication in 25 mM Hepes pH 7.4, 400 mM NaCl, 20 mM imidazole. The proteins were purified by Ni-NTA chromatography using lysis buffer as the washing buffer. After elution with imidazole, proteins were subjected to size-exclusion chromatography on a S200 increase 10/300 GL column in 25 mM Hepes pH 7.4, 150 mM NaCl, 5% glycerol. The peak fractions were stored for future use.

Purification of GST-Fused ALFA Peptide. GST-fused ALFA peptide was purified as follows: The gene for the ALFA peptide was cloned into the pGEX6p1

vector (Cytiva). The expression was induced by addition of 1 mM IPTG for 5 h at 30 °C. The cells were lysed by sonication in 25 mM Hepes pH 7.4, 400 mM NaCl. The proteins were purified by GST resins using lysis buffer as the washing buffer. After elution with reduced glutathione (GSH), proteins were subjected to size-exclusion chromatography on a S200 increase 10/300 GL column in 25 mM Hepes pH 7.4, 150 mM NaCl, 5% glycerol. The peak fractions were stored for future use.

Purification of Legobodies. Legobodies were assembled by first incubating purified MBP_Pra/G with Fab_8D3_2 at a molar ratio of 1:1.1 in 25 mM Hepes pH 7.4, 150 mM NaCl. Then, the mixture was incubated with a chosen nanobody added at a threefold molar excess over MBP_Pra/G. The sample was applied to an amylose resin and the complex was eluted with 25 mM Hepes pH 7.4, 150 mM NaCl, 20 mM maltose. The Legobodies were further purified by size-exclusion chromatography on a S200 increase 10/300 GL column in 25 mM Hepes pH 7.4, 150 mM NaCl, 5% glycerol, 2 mM maltose. The peak fractions of the complexes were concentrated and stored for future use.

Purification of a Complex of KDELR and Legobody. The codon-optimized gene for the full-length KDELR with a SBP tag at its C terminus was cloned into the pS425-Gal1 vector (ATCC 87331) (36). The expression of the receptor and preparations of the membrane fractions were carried out as previously described (25). Membranes from 15 g of INVSc1 (Invitrogen) cells expressing the receptor were solubilized in 30 mL of 25 mM Hepes pH 7.4, 400 mM NaCl, 1% DMNG (Anatrace) for 1 h. After removing insoluble material by ultracentrifugation, the lysate was incubated with 250 μ L streptavidin resin (Thermo Fisher) for 1.5 h. The beads were collected and an excess of purified Legobody was added to the bound KDELR receptor to promote complex formation on the resin. After 1 h of incubation, the resin was washed with eight column volumes of 25 mM Hepes pH 7.4, 150 mM NaCl, 2 mM maltose, 0.02% DMNG. Nanodiscs were assembled on the resin by adding 1.25 mM lipids (POPC/DOPE [Avanti] at a 4:1 ratio in DDM [Anatrace]) and 25 μ M nanodisc-scaffolding protein MSP1D1 in 700 μ L of washing buffer. After 30 min of incubation, the detergents were removed by the addition of two aliquots of 40 mg of Bio-Beads and overnight incubation. The next day, the streptavidin resins were separated from Bio-Beads SM-2 (Bio-Rad), taking advantage of their different rates of sedimentation by gravity. The streptavidin resins were washed by 25 mM Hepes pH 7.4, 150 mM NaCl, 2 mM maltose, and bound material was eluted with biotin. The KDELR receptor/Legobody complex in a nanodisc was then purified by size-exclusion chromatography on a S200 increase 5/150 GL column in 25 mM Hepes pH 7.4, 150 mM NaCl, 2 mM maltose. The peak fractions of the complex were concentrated, snap frozen, and stored for cryo-EM analysis.

Purification of a Complex of the SARS-CoV-2 Spike RBD Domain and Legobody. The codon-optimized gene for the RBD domain (residues 334 to 526) of SARS-CoV-2 spike protein with an N-terminal Flag tag and a C-terminal His8 tag was cloned into the pCAGEN vector. The RBD was expressed and purified in the same way as the Fabs. After elution from Ni-NTA beads, the protein was subjected to size-exclusion chromatography on a S75 increase 10/300 GL column in 25 mM Hepes pH 7.4, 150 mM NaCl. Peak fractions were mixed with Legobody at a molar ratio of 3:1. The mixture was subjected to size-exclusion chromatography on a S200 increase 5/150 GL column in 25 mM Hepes pH 7.4, 150 mM NaCl, 2 mM maltose. The peak fractions of the complex were concentrated, snap frozen, and stored for cryo-EM analysis.

Cryo-EM Sample Preparation and Data Acquisition. The KDELR/Legobody complex at 0.8 mg/mL was PEGylated by incubation with MS(PEG)12 methyl-PEG-NHS-ester (Thermo Fisher) at a 1:40 molar ratio for 2 h on ice to reduce preferred particle orientation on the grids. The chosen ratio allows a maximum of 1/3 of the total lysins to be modified, which minimizes effects of PEG modification on the stability of the complex. The PEGylated sample was then applied to a glow-discharged Quantifoil gold grid (1.2/1.3, 400 mesh). The grids were blotted for ~6 to 7 s at 100% humidity and plunge frozen in liquid ethane using a Vitrobot Mark IV instrument (Thermo Fisher). The RBD/Legobody complex at 2.5 mg/mL was incubated with MS(PEG)12 methyl-PEG-N-hydroxysuccinimide (NHS)-ester (Thermo Fisher) at a ~1:25 to 1:28 molar ratio for 2 h on ice. Right before plunge freezing, the PEGylated samples were diluted, using the gel-filtration buffers supplemented with detergent IGEPAL CA-630 (Sigma), so that the final protein and detergent concentrations were 1.2 mg/mL and 0.005%, respectively. The grids were frozen in the same way as described for the KDELR/Legobody sample.

Cryo-EM data for all samples were collected on a Titan Krios electron microscope (FEI) operated at 300 kV and equipped with a K3 direct electron

detector (Gatan) at Harvard Cryo-EM Center for Structural Biology. A Gatan Imaging filter with a slit width of 25 eV was used to remove inelastically scattered electrons. All cryo-EM movies were recorded in counting-mode using SerialEM. For the KDELR/Legobody sample, the nominal magnification of 81,000 \times corresponds to a calibrated pixel size of 1.06 Å on the specimen. The exposure rate was 23.38 electrons/Å²/s. The total exposure time was 2.2 s, resulting in a total electron exposure of 51.44 electrons/Å², fractionated into 50 frames (44 ms per frame). For the RBD/Legobody sample, the calibrated pixel size was 1.06 Å. The exposure rate was 23.3 electrons/Å²/s. The total exposure time was 2.164 s, resulting in a total electron exposure of 50.42 electrons/Å², fractionated into 50 frames (44 ms per frame). The defocus range for both samples was between -1.0 and -2.6 μ m.

Cryo-EM Image Processing. For the KDELR/Legobody complex, dose-fractionated movies were subjected to motion correction using the program MotionCor2 (37) with dose weighting. The program CtfFind4 (38) was used to estimate defocus values of the summed images from all movie frames. During data collection, the particles (close to 1 million) picked by YOLO (39) by “on-the-fly” analysis using an automatic workflow at Harvard Medical School. The particles were then subjected to 2D classification (T2, 80 classes, 30 iterations) in Relion 3.1 (40). For 3D classification, an initial model was generated ab initio in Relion 3.1. After one round of 3D classification (T4, 5 classes, 50 iterations), there was only one class with clear protein secondary structure features. Particles of this class were selected for 3D refinement, resulting in an initial reconstruction at 3.8-Å overall nominal resolution. This initial 3D reconstruction was used as 3D template to perform autopick in Relion 3.1 on the entire dataset, resulting in 2,532,161 particles. After 2D classifications (T2, 100 classes, 30 iterations), repicked particles were “seeded” with particles used in the previous 3D refinement for 3D classification. After 3D classification (T4, 5 classes, 35 iterations), only the class showing clear protein secondary structure features of the whole complex was selected. After removing duplicates, the particles were subjected to 3D refinement, followed by polishing, Contrast Transfer Function (CTF) refinement, and another round of 3D refinement. Local 3D classification without image alignment (T20, 5 classes, 25 iterations) was performed using a mask including only the nanobody and KDELR. A total of 246,878 particles were finally selected for 3D refinement using a mask excluding the nanodisc and the more flexible D domain of protein A.

For the RBD/Legobody complex, data analysis was performed in a similar way, except that particles from the on-the-fly analysis were not refined. Local resolution calculation and map sharpening were both performed in Relion 3.1. All reported resolutions are based on gold-standard refinement procedures and the Fourier Shell Correlation (FSC) = 0.143 criterion. Histograms of directional FSC curves and sphericity values were calculated with the 3DFSC server (41).

Model Building. All model building was done in Coot. For the crystal structure of Nb_0/Fab_8D3, the initial phases were obtained by molecular replacement using the Phaser module in Phenix (42). The search models contained a nanobody, as well as the variable and constant regions of a Fab of similar framework sequence. After obtaining an initial density map, the model was refined with rigid bodies and then modified manually. The model was further refined using the Phenix.refine module with simulated annealing, XYZ, TLS, and individual B factors. For the cryo-EM structures, initial models were based on the crystal structure of the Nb_0/Fab_8D3 complex, modified to account for the mutations in Fab_8D3_2, and the crystal structures of the KDELR receptor (6I6J), RBD (7KGJ), the manually grafted MBP_PraC (1ANF and 4NPD), the D domain of Protein A (1DEE), and protein G (1IGC). These structures were docked into the maps and manually modified based on the cryo-EM density map. Models were then refined using the Phenix real-space refinement module with minimization_global, local_grid_search, and Atomic Displacement Parameters (ADP). For all refinements, secondary restraints, model restraints, and Ramachandran restraints were used.

Data Availability. All study data are included in the article and/or supporting information. The coordinates of the atomic models of Nb_0/Fab_8D3, KDELR/Legobody, and RBD/Legobody have been deposited in the Protein Data Bank with accession codes 7R9D (43), 7RXC (44), and 7RXD (45), respectively. The cryo-EM maps of the KDELR/Legobody and RBD/Legobody have been deposited with accession codes EMD-24728 (46) and EMD-24729 (47), respectively. Plasmids for the Legobody have been deposited to Addgene [176075 (48), 176076 (49), 176077 (50)].

ACKNOWLEDGMENTS. We thank S. Sterling, R. Walsh, and M. Mayer at the Harvard Cryo-EM Center for Structural Biology for help in microscope operation and data collection. The X-ray structural work is based on research

conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences (NIGMS) from the NIH (P30 GM124165). The Eiger 16M detector on 24-ID-E is funded by an NIH-Office of Research Infrastructure Programs High-End Instrumentation grant (S10OD021527). This research used resources from the Advanced Photon Source, a US Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne

National Laboratory under Contract DE-AC02-06CH11357. We thank C. Bahl for initial efforts to use Rosetta for the design of fusions between MBP and Protein A domains, T.-C. Hu for help with the generation of Fabs, the SBGrid team for software and workstation support, and Maofu Liao and Susan Shao for comments on the manuscript. This work was supported by a Jane Coffin Child fellowship to X.W. and an NIGMS award to T.A.R. (R01GM052586).

1. X.-C. Bai, G. McMullan, S. H. W. Scheres, How cryo-EM is revolutionizing structural biology. *Trends Biochem. Sci.* **40**, 49–57 (2015).
2. M. A. Herzik Jr, M. Wu, G. C. Lander, High-resolution structure determination of sub-100 kDa complexes using conventional cryo-EM. *Nat. Commun.* **10**, 1032–1039 (2019).
3. X. Fan *et al.*, Single particle cryo-EM reconstruction of 52 kDa streptavidin at 3.2 Å resolution. *Nat. Commun.* **10**, 2386 (2019).
4. Y. Han *et al.*, High-yield monolayer graphene grids for near-atomic resolution cryo-electron microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 1009–1014 (2020).
5. N. Tsutsumi *et al.*, Structure of human Frizzled5 by fiducial-assisted cryo-EM supports a heterodimeric mechanism of canonical Wnt signaling. *eLife* **9**, e58464 (2020).
6. Y. Liu, D. T. Huynh, T. O. Yeates, A 3.8 Å resolution cryo-EM structure of a small protein bound to an imaging scaffold. *Nat. Commun.* **10**, 1864–1867 (2019).
7. Q. Yao, S. J. Weaver, J.-Y. Mock, G. J. Jensen, Fusion of DARPin to aldolase enables visualization of small protein by Cryo-EM. *Structure* **27**, 1148–1155.e3 (2019).
8. S. Wu *et al.*, Fabs enable single particle cryoEM studies of small proteins. *Structure* **20**, 582–592 (2012).
9. Y. Lee *et al.*, Cryo-EM structure of the human L-type amino acid transporter 1 in complex with glycoprotein CD98hc. *Nat. Struct. Mol. Biol.* **26**, 510–517 (2019).
10. J. S. Bloch *et al.*, Structure and mechanism of the ER-based glucosyltransferase ALG6. *Nature* **579**, 443–447 (2020).
11. J. Kim *et al.*, Structure and drug resistance of the *Plasmodium falciparum* transporter PfCRT. *Nature* **576**, 315–320 (2019).
12. I. Zimmermann *et al.*, Synthetic single domain antibodies for the conformational trapping of membrane proteins. *eLife* **7**, e34317 (2018).
13. C. McMahon *et al.*, Yeast surface display platform for rapid discovery of conformationally selective nanobodies. *Nat. Struct. Mol. Biol.* **25**, 289–296 (2018).
14. T. Uchański *et al.*, Megabodies expand the nanobody toolkit for protein structure determination by single-particle cryo-EM. *Nat. Methods* **18**, 60–68 (2021).
15. R. Henderson, The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Q. Rev. Biophys.* **28**, 171–193 (1995).
16. L. G. Frenken *et al.*, Isolation of antigen specific llama VHH antibody fragments and their high level secretion by *Saccharomyces cerevisiae*. *J. Biotechnol.* **78**, 11–21 (2000).
17. B. Jansson, M. Uhlén, P. A. Nygren, All individual domains of staphylococcal protein A show Fab binding. *FEMS Immunol. Med. Microbiol.* **20**, 69–78 (1998).
18. P. C. Fridy, M. K. Thompson, N. E. Ketaren, M. P. Rout, Engineered high-affinity nanobodies recognizing staphylococcal Protein A and suitable for native isolation of protein complexes. *Anal. Biochem.* **477**, 92–94 (2015).
19. D. S. Waugh, Crystal structures of MBP fusion proteins. *Protein Sci.* **25**, 559–571 (2016).
20. S.-J. Youn *et al.*, Construction of novel repeat proteins with rigid and predictable structures using a shared helix method. *Sci. Rep.* **7**, 2595 (2017).
21. M. Graille *et al.*, Crystal structure of a *Staphylococcus aureus* protein A domain complexed with the Fab fragment of a human IgM antibody: Structural basis for recognition of B-cell receptors and superantigen activity. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5399–5404 (2000).
22. J. P. Derrick, D. B. Wigley, The third IgG-binding domain from streptococcal protein G. An analysis by X-ray crystallography of the structure alone and in a complex with Fab. *J. Mol. Biol.* **243**, 906–918 (1994).
23. J. C. Semenza, K. G. Hardwick, N. Dean, H. R. Pelham, ERD2, a yeast gene required for the receptor-mediated retrieval of luminal ER proteins from the secretory pathway. *Cell* **61**, 1349–1357 (1990).
24. P. Bräuer *et al.*, Structural basis for pH-dependent retrieval of ER proteins from the Golgi by the KDEL receptor. *Science* **363**, 1103–1107 (2019).
25. X. Wu *et al.*, Structural basis of ER-associated protein degradation mediated by the Hrd1 ubiquitin ligase complex. *Science* **368**, eaaz2449 (2020).
26. M. Hoffmann *et al.*, SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271–280.e8 (2020).
27. T. F. Czajka, D. J. Vance, N. J. Mantis, Slaying SARS-CoV-2 one (single-domain) antibody at a time. *Trends Microbiol.* **29**, 195–203 (2021).
28. J. D. Walter *et al.*, Synthetic nanobodies targeting the SARS-CoV-2 receptor-binding domain. *bioRxiv*, 2020.04.16.045419 (2020).
29. J. Ahmad, J. Jiang, L. F. Boyd, K. Natarajan, D. H. Margulies, Synthetic nanobody–SARS-CoV-2 receptor-binding domain structures identify distinct epitopes. *bioRxiv*, 2021.01.27.428466 (2021).
30. H. Götzke *et al.*, The ALFA-tag is a highly versatile tool for nanobody-based bioscience applications. *Nat. Commun.* **10**, 4403–4412 (2019).
31. X. Wu, C. Cabanos, T. A. Rapoport, Structure of the post-translational protein translocation machinery of the ER membrane. *Nature* **566**, 136–139 (2019).
32. J. Chung *et al.*, LDAF1 and seipin form a lipid droplet assembly complex. *Dev. Cell* **51**, 551–563.e7 (2019).
33. M. Graille *et al.*, Complex between *Peptostreptococcus magnus* protein L and a human antibody reveals structural convergence in the interaction modes of Fab binding proteins. *Structure* **9**, 679–687 (2001).
34. R. K. Grover *et al.*, A structurally distinct human mycoplasma protein that generically blocks antigen-antibody union. *Science* **343**, 656–661 (2014).
35. T. Matsuda, C. L. Cepko, Electroporation and RNA interference in the rodent retina in vivo and in vitro. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 16–22 (2004).
36. D. Mumberg, R. Müller, M. Funk, Regulatable promoters of *Saccharomyces cerevisiae*: Comparison of transcriptional activity and their use for heterologous expression. *Nucleic Acids Res.* **22**, 5767–5768 (1994).
37. S. Q. Zheng *et al.*, MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
38. A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
39. T. Wagner *et al.*, SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Commun. Biol.* **2**, 218 (2019).
40. J. Zivanov *et al.*, New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
41. Y. Z. Tan *et al.*, Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods* **14**, 793–796 (2017).
42. D. Liebschner *et al.*, Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877 (2019).
43. X. Wu, T. A. Rapoport, Crystal structure of Nb_0 in complex with Fab_8D3. RCSB. <https://www.rcsb.org/structure/7R9D>. Deposited 23 August 2021.
44. X. Wu, T. A. Rapoport, CryoEM structure of KDELr with Legobody. RCSB. <https://www.rcsb.org/structure/7RXC>. Deposited 23 August 2021.
45. X. Wu, T. A. Rapoport, CryoEM structure of RBD domain of COVID-19 in complex with Legobody. RCSB. <https://www.rcsb.org/structure/7RXD>. Deposited 23 August 2021.
46. X. Wu, T. A. Rapoport, CryoEM structure of KDELr with Legobody. EMD. <https://www.emdataresource.org/EMD-24728>. Deposited 23 August 2021.
47. X. Wu, T. A. Rapoport, CryoEM structure of RBD domain of COVID-19 in complex with Legobody. EMD. <https://www.emdataresource.org/EMD-24729>. Deposited 23 August 2021.
48. X. Wu, T. A. Rapoport, pCAG-Fab_8D3_2_H-H6. Addgene. <https://www.addgene.org/176075/>. Deposited 26 August 2021.
49. X. Wu, T. A. Rapoport, pCAG-Fab_8D3_2_L. Addgene. <https://www.addgene.org/176076/>. Deposited 26 August 2021.
50. X. Wu, T. A. Rapoport, pET28 MBP_PrA/G-H6. Addgene. <https://www.addgene.org/176077/>. Deposited 26 August 2021.