

Supplementary Information

Supplementary Texts

Supplementary Text 1. Benchmark set using synthesized density maps

To evaluate DEMO-EM, we collected a set of 357 non-redundant proteins from the PDB with the domain boundary assigned by DomainParser¹ (or SCOPE² and CATH³ when available). The length of these proteins ranges from 99 to 1,693 residues with 2 to 12 domains which covers the majority of single-chain multi-domain proteins in the PDB. The density maps of these proteins are simulated according to the experimental structures by EMAN2⁴(**Supplementary Text 2**), with the resolution randomly selected from 2 to 10 Å and a grid spacing of 1 Å/voxel (**Supplementary Fig. 2a**). Two separate tests are performed to assemble experimental domain structures extracted from the full-length target structures and domain models predicted by D-I-TASSER⁵. For experimental domain structure assembly, all domains were randomly rotated and translated as rigid bodies before assembly. When using D-I-TASSER to model the domain structures, all homologous templates of a sequence identity >30% to the query have been excluded; this resulted in domain models with variable quality and the TM-score⁶ ranging from 0.22 to 0.96 (see **Supplementary Fig. 2b** for a histogram of TM-score distribution). The average TM-score of these predicted domain models is 0.77, and there are 46 out of 357 proteins have at least one domain with incorrect global topology (TM-score <0.5).

Supplementary Text 2. Implementation of EMAN2 to generate synthesized density maps

EMAN⁷ is a scientific image processing package, which mainly focuses on single particle reconstruction from transmission electron microscopy images. EMAN2⁴ is the successor to EMAN1, with a fully refactored image processing library, and a wide range of features to make it much more flexible and extensible than EMAN1. For the 357 benchmark proteins and the 425 training proteins, we used the *e2pdb2mrc.py* program from the EMAN2 package (version 2.31) to generate synthesized EM density maps.

The following command was used to create the density maps:

```
e2pdb2mrc.py inputfile.pdb outputfile.mrc -A voxelsize -R resolution
```

where “inputfile.pdb” is PDB file to convert, and it is the native structure in our experiment; “outputfile.mrc” is the output file for the generated density map; “voxelsize” is the Angstroms per voxel in the output, and it is set to 1 in our experiment; “resolution” is the resolution of the created density map, which is randomly selected from 2 to 10 Å for each protein in our experiment.

Supplementary Text 3. Description of TM-score

TM-score⁶ is a metric for evaluating the topological similarity between protein structures, which can be calculated by

$$\text{TM-score} = \max \left[\frac{1}{L_{\text{target}}} \sum_{i=1}^{L_{\text{aligned}}} \frac{1}{1 + \left(\frac{d_i}{d_0(L_{\text{target}})} \right)^2} \right] \quad (S1)$$

where L_{target} is the amino acid sequence length of the target protein, L_{aligned} is the length of the aligned residues to the reference (native) structure, d_i is the distance between the i -th pair of aligned residues, $d_0(L_{\text{target}}) = 1.24^3 \sqrt{L_{\text{target}} - 15} - 1.8$ is a scale to normalize the match difference, and ‘max’ refers to the optimized value selected from various rotation and translation matrices for structure superposition. The value of TM-score ranges in (0,1], where 1 indicates that the two structures are identical. Stringent statistics showed that TM-score >0.5 corresponds to a similarity with two structures having the same fold and/or domain orientations⁸. Our assessment will be mainly on TM-score because TM-score is protein length-independent and its value is more sensitive to the fold and domain orientation similarities of the predicted model relative to the native, compared to RMSD.

It should be noted that TM-score can be discrepant to the widely used root-mean-square deviation (RMSD) for some protein structure pairs. This is mainly because by definition, RMSD ($=\sqrt{\frac{1}{N}\sum_{i=1}^N d_i^2}$) is calculated as an average of distance error (d_i) with equal weight over all residue pairs. Therefore, a big local error on a few residue pairs can result in a quite large RMSD. On the other hand, by putting d_i in the denominator of Eq. (S1), TM-score naturally weights smaller distance errors stronger than the larger distance errors. Therefore, TM-score value is more sensitive to the global structural similarity rather than to the local structural errors, compared to RMSD. Another advantage of TM-score is the introduction of the scale $d_0(L_{\text{target}}) = 1.24\sqrt[3]{L_{\text{target}} - 15} - 1.8$ which makes the magnitude of TM-score length-independent for random structure pairs, while RMSD is a length-dependent metric⁶. Due to these reasons, our discussion of modeling results is mainly based on TM-score. Since RMSD is intuitively more familiar to most readers, however, we also list RMSD values when necessary in the manuscript.

Supplementary Text 4. Analyses of cases with domain accuracy decreased after flexible assembly

As shown in Fig 2d, TM-score of 810 out of 890 individual domain models were improved after the DEMO-EM flexible assembly. The remaining 80 (=890-810) domain models with TM-score decreased after the flexible assembly and refinement by DEMO-EM come from 72 test targets. Detailed analyses (**Supplementary Fig. 3**) show that 34.7% of them include ≥ 1 initial domain model(s) with incorrect fold, where the incorrect initial domain structures can result in incorrect domain-map match which misguides the subsequent domain assembly and refinement simulations. For other 36.6% of these cases, although all domains were correctly modelled by D-I-TASSER with TM-score >0.5 , the full-length domain models created by DEMO-EM rigid-body assembly obtain a TM-score <0.7 due to the incorrect position and orientation of some domains in the density map, which also affect the algorithm to detect the unreasonably fold regions for remodeling. For the rest cases, the decreased TM-score of domain models was mainly caused by the low-resolution density map as 18 out of 21 cases employed density maps with resolution $>6\text{\AA}$.

Supplementary Text 5. Performance of DEMO-EM for large proteins, cases with discontinuous domains, and proteins with incorrect domain models

Although the degrees of freedom and searching space usually increases with the number of domains for both domain assembly and full-length model refinement, the quality of the final models by DEMO-EM does not significantly decrease with an increasing number of domains for both experimental and predicted domain model assemblies (**Supplementary Tables 1 and 2**). In particular, the full-length models assembled using predicted domain structures obtain an average TM-score of 0.84 for proteins with 3 or more domains, 91.6% of which have a final model TM-score >0.5 . This data demonstrates the ability of DEMO-EM for handling large complex structures of multi-domain proteins and relative long sequence proteins. In addition, proteins with discontinuous domains are usually difficult to model as they have several parts separated in sequence which increases the difficulty in individual domain modeling and inter-domain distances prediction. However, DEMO-EM correctly assembled nearly all the cases of discontinuous domains using experimental domain models, with an average TM-score=0.99 identical to that of the continuous domain model assembly (**Supplementary Table 1**). This is probably due to the constraints brought by the additional linkers which help guide the discontinuous domain structure assembly (see **Eq. 4** in Methods and **Supplementary Fig. 22**). When assembling predicted domain models, however, DEMO-EM achieves an average TM-score of 0.80 lower than the continuous domain assembly results (TM-score=0.86, **Supplementary Table 2**), due to the incorrect domain topologies (19.8% of the domains) and the modeling errors of predicted models in the tail and linker regions which affect the packing the segment structures. Furthermore, longer proteins are usually more difficult to model for most of the protein structure prediction methods due to the larger conformational space to search. However, when splitting the sequence into domains for independently modeling, the domain models created by D-I-TASSER obtains an average TM-score of 0.79 for the large proteins with sequence length ranges from 500 to 1,693, where 94.1% of the domains has correct fold with TM-score >0.5 . The full-length models assembled by DEMO-EM using these predicted domain models achieve an average TM-score of 0.81, 93.5% of which have a final model TM-score >0.5 .

These results suggest that the DEMO-EM pipeline can effectively handle the proteins with relative long sequences. Finally, incorrect domain models probably lead to poor final models as they can negatively affect the domain assembly and refinement simulations. However, DEMO-EM successfully assembled full-length models for 94.1% of the test proteins while only 87.1% of the test cases have all initial domain models with correct fold. We found that 60.9% of the cases which have ≥ 1 domain(s) with incorrect fold resulted in final full-length models with correct global fold. This data indicates that DEMO-EM has the ability to assemble correct full-length models even starting from low-quality domain structures for some targets.

Supplementary Text 6. Implementation of MDFF, Rosetta, and MAINMAST programs

MDFF is cryo-EM density-map guided protein structure fitting and refinement program through a combined search process of Monte Carlo simulation, conjugate-gradients minimization, and simulated annealing molecular dynamics simulations^{9,10}. Starting from the initial full-length model built by matching each domain model into density maps using Situs¹¹, the direct MDFF⁹ and cMDFF¹⁰ programs were applied for proteins with cryo-EM density maps at resolution $\geq 5\text{\AA}$ and $< 5\text{\AA}$, respectively. The direct MDFF was carried out by a domain rigid-body assembly with domain restraints for up to 200 ps by setting “-numsteps 200000”, followed by a three-step flexible fitting which includes two molecular dynamics (MD) simulations for up to 500 ps by setting “-numsteps 500000” and an final energy minimization by setting “-minsteps 2000”. The settings were determined according to the second section named “A simple MDFF example” and the fourth section named “MDFF with Domain Restraints” in the tutorial (March 28, 2019). In the domain rigid-body assembly and each MD simulation of the flexible fitting, the energy minimization was also performed by setting “-minsteps 2000”. According to the parameters suggested previously¹², the electron density term with a weight of 1.0, 1.0, and 0.3 was set for the rigid-body assembly and the two MD simulations of the flexible fitting, respectively, and a weight of 0.3 was used in the final energy minimization¹³. In cMDFF, Gaussian blurred maps were created using half-width σ from 5 Å to 0 Å with a decreased step of 1 Å, thus yielding 6 maps, including the original maps. The direct MDFF with up to 200 ps for each MD simulation of the flexible fitting were performed for each of the 6 maps to achieve convergence. In our experiment, the MDFF simulations were performed using NAMD (version 2.14b1), and the necessary files for MDFF and MDFF simulations were generated/analyzed using VMD (version 1.9.4a38).

Taking the target PDB named as “example.pdb” which includes 2 domains as an example, the necessary input files were generated through VMD according to the PDB file and the cryo-EM data:

1. Perform the following command to generate a PSF file (example_autopsf.psf) containing all the connectivity information and partial atomic charges required by NAMD and a PDB file (example_autopsf.pdb) through the AutoPSF plugin (version 1.8):

```
mol new example.pdb
package require autopsf
resetpsf
autopsf -protein example.pdb
```

2. Run the following command to generate the DX file (example_autopsf-grid.dx) defining the cryo-EM density restraint:

```
package require mdff
mdff griddx -i density_map.mrc -o example_autopsf-grid.dx
```

where “density_map.mrc” is the cryo-EM density map.

3. Run the following command to generate a PDB file (example_autopsf-grid.pdb) containing the per-atom scaling factors of the cryo-EM restraint:

```
mdff gridpdb -psf example_autopsf.psf -pdb example_autopsf.pdb -o example_autopsf-grid.pdb
```

4. Perform the following command to define restraints for ϕ and ψ dihedral angles for amino acid residues in helices or sheets, as well as restraints for hydrogen bonds involving backbone atoms from the same residues:

5. Generate restraints to prevent cis/trans peptide transitions and chirality errors through the following command:

```
mol new example_autopsf.psf
mol addfile example_autopsf.pdb
cispeptide restrain -o example-extrabonds-cispt.txt
chirality restrain -o example-extrabonds-chaira.txt
```

6. Select each group of atoms and set their beta column to define the domain through the following command:

```
set sel [atomselect top "all"]
$sel set beta 0
$sel set occupancy 0
set sel1 [atomselect top "segname 1 and name CA"]
$sel1 set beta 1
$sel1 set occupancy 1
set sel2 [atomselect top "segname 2 and name CA"]
$sel2 set beta 2
$sel2 set occupancy 1
$sel writepdb domain.pdb
```

Here, two domains were assigned (beta 1 and 2) and kept rigid during the MDFF simulations.

After generating the input files through the above commands, the following command was used to create the NAMD configuration file to run the MDFF domain rigid-body assembly:

```
mdff setup -o domain -psf example_autopsf.psf
-pdb example_autopsf.pdb
-griddx example_autopsf-grid.dx
-gridpdb example_autopsf-grid.pdb
-extrab {example-extrabonds.txt example-extrabonds-cispt.txt example-extrabonds-chaira.txt}
-gscale 1.0
-minsteps 2000
-numsteps 200000
```

Edit the configuration file created by the above command by adding the following lines anywhere:

```
tmd on
tmdfile domain.pdb
tmdk 500.
tmdfirststep 2001
tmdlaststep 202000
tmdoutputfreq 1000
```

Run NAMD to perform the MDFF simulation using the configuration file, i.e., run the following command:

```
namd2 xxx.namd > xxx.log
```

where “xxx.namd” is the NAMD configure file generated by VMD, and “xxx.log” is the log file of the simulation.

When the domain rigid-body assembly was completed, the necessary input files for the flexible refinement were generated according to the above commands for the domain rigid-body assembly, and the following command was employed to generate the NAMD configuration file for the flexible refinement in the direct MDFF:

```
mdff setup -o adk -psf xxx_autopsf.psf
-pdb xxx_autopsf.pdb
-griddx xxx_autopsf-grid.dx
-gridpdb xxx_autopsf-grid.pdb
-extrab {xxx-extrabonds.txt xxx-extrabonds-cispt.txt xxx-extrabonds-chaira.txt}
-gscale 1.0
-minsteps 2000
-numsteps 500000
mdff setup -o adk -psf xxx_autopsf.psf
-pdb xxx_autopsf.pdb
-griddx xxx_autopsf-grid.dx
-gridpdb xxx_autopsf-grid.pdb
-extrab {xxx-extrabonds.txt xxx-extrabonds-cispt.txt xxx-extrabonds-chaira.txt}
-gscale 0.3
-minsteps 2000
-numsteps 500000
-step 2
mdff setup -o adk -psf xxx_autopsf.psf
-pdb xxx_autopsf.pdb
-griddx xxx_autopsf-grid.dx
-gridpdb xxx_autopsf-grid.pdb
-extrab {xxx-extrabonds.txt xxx-extrabonds-cispt.txt xxx-extrabonds-chaira.txt}
-gscale 0.3
-minsteps 2000
-numsteps 0
-step 3
```

where “-numsteps” of the first two steps was set to 200000 in cMDFF, and the MDFF simulation of each step was performed by run NAMD using the corresponding configuration file. To avoid the MDFF web tutorial changing, we have put the current version at https://zhanggroup.org/DEMO-EM/data_set/benchmark51/tutorial_mdff.pdf. It should be noted that the process of MD simulation in MDFF is sensitive to the quality of initial models and may fail or quit prematurely without reaching the given number of steps if starting from a poor-quality initial model. In our experiment, MDFF had failed on

many targets when directly starting from the models created by Situs on the I-TASSER domains because of the broken linkers between two domains with C α -C α bond-length significantly higher than 3.8 Å.

To solve this problem, we used the following procedure to quickly rebuild the domain linkers for the Situs models: When a C α -C α bond gap (>4.0 Å) is identified, both sides of residues near the gap are gradually released until the two anchor residues could be connected (i.e., with distance < 3.5(l + 1), where l is the number of released residues)¹⁴. Next, the C α trace of the released residues is regenerated by self-avoiding random walks of C α -C α vectors with bond-length=3.8 Å, where other backbone (N, C, O) and sidechain atoms are added using FASPR¹⁵. Finally, a short Metropolis Monte Carlo simulation is performed to refine the linker conformation under the guidance of a potential containing a C α clash term, a statistical torsion-angle potential from Ramachandran plots^{16,17}, an orientation-dependent sidechain contact potential, and a statistical N-C α -C bond angle potential, where the linker model with the lowest energy is finally selected and refined with PULCHRA¹⁸. After the linker reconstruction, it was found that MDFF can be performed for all the cases, although the domain rigid-body assembly step may still quit prematurely for few cases due to the bad local structures inherited from the I-TASSER models.

Rosetta builds cryo-EM models also through Monte Carlo simulations and Limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) as guided by the Rosetta all-atom force field¹⁹. Rosetta ver. 3.12 (rosetta_bin_linux_2020.50.61505_bundle) was used. Started with the same initial models constructed from Situs¹¹, we first use the following protocol to further optimize the position and orientation of each domain in the density map by providing the flag “-edensity::realign min”:

```
$ROSETTA/source/bin/score_jd2.linuxgccrelease \
  -database $ROSETTA/database/ \
  -in::file::s pose.pdb pose.pdb \
  -ignore_unrecognized_res \
  -edensity::mapfile densityMap.mrc \
  -edensity::mapreso 5.0 \
  -edensity::grid_spacing 2.0 \
  -edensity::fastdens_wt 35.0 \
  -edensity::cryoem_scatterers \
  -edensity::realign min \
  -out::pdb \
  -crystal_refine
```

All domain models are then combined together to get an initial full-length model, and the above protocol is employed again to optimize the position and orientation of the initial full-length model in the density map. Finally, the following protocol described in https://faculty.washington.edu/dimaio/files/rosetta_density_tutorial_aug18_2.pdf is used to flexibly refine the full-length model according to the density map:

```
$ROSETTA/source/bin/rosetta_scripts.linuxgccrelease \
  -database $ROSETTA/database/ \
  -in::file::s pose.pdb \
  -parser::protocol A_asymm_refine.xml \
  -parser::script_vars denswt=35 rms=1.5 reso=3.4 map=half1.mrc testmap=half2.mrc \
  -ignore_unrecognized_res \
  -edensity::mapreso 3.4 \
  -default_max_cycles 200 \
  -edensity::cryoem_scatterers \
  -beta \
  -out::suffix _asymm \
  -crystal_refine
```

where “A_asymm_refine.xml” is shown below:

```
<ROSETTASCRIPITS>
<SCOREFXNS>
  <ScoreFunction name="cen" weights="score4_smooth_cart">
    <Reweight scoretype="elec_dens_fast" weight="20"/>
  </ScoreFunction>
  <ScoreFunction name="dens_soft" weights="beta_soft">
    <Reweight scoretype="cart_bonded" weight="0.5"/>
    <Reweight scoretype="pro_close" weight="0.0"/>
    <Reweight scoretype="elec_dens_fast" weight="%%denswt%%"/>
  </ScoreFunction>
  <ScoreFunction name="dens" weights="beta_cart">
    <Reweight scoretype="elec_dens_fast" weight="%%denswt%%"/>
    <Set scale_sc_dens_byres="R:0.76,K:0.76,E:0.76,D:0.76,M:0.76,C:0.81,Q:0.81,H:0.81,N:0.81,T:0.81,
      S:0.81,Y:0.88,W:0.88,A:0.88,F:0.88,P:0.88,I:0.88,L:0.88,V:0.88"/>
  </ScoreFunction>
</SCOREFXNS>
</ROSETTASCRIPITS>
```

```

</ScoreFunction>
</SCOREFXNS>
<MOVERS>
  <SetupForDensityScoring name="setupdens"/>
  <LoadDensityMap name="loaddens" mapfile="%%map%%"/>
  <SwitchResidueTypeSetMover name="tocen" set="centroid"/>
  <MinMover name="cenmin" scorefxn="cen" type="lbfgs_armijo_nonmonotone"
    max_iter="200" tolerance="0.00001" bb="1" chi="1" jump="ALL"/>
  <CartesianSampler name="cen5_50" automode_scorecut="-0.5" scorefxn="cen"
    mcscorefxn="cen" fasticsearch="dens_soft" strategy="auto" fragbias="density"
    rms="%%rms%%" ncycles="200" fullatom="0" bbmove="1" nminsteps="25" temp="4" fraglens="7"
    nfrags="25"/>
  <CartesianSampler name="cen5_60" automode_scorecut="-0.3" scorefxn="cen"
    mcscorefxn="cen" fasticsearch="dens_soft" strategy="auto" fragbias="density"
    rms="%%rms%%" ncycles="200" fullatom="0" bbmove="1" nminsteps="25" temp="4" fraglens="7"
    nfrags="25"/>
  <CartesianSampler name="cen5_70" automode_scorecut="-0.1" scorefxn="cen"
    mcscorefxn="cen" fasticsearch="dens_soft" strategy="auto" fragbias="density"
    rms="%%rms%%" ncycles="200" fullatom="0" bbmove="1" nminsteps="25" temp="4" fraglens="7"
    nfrags="25"/>
  <CartesianSampler name="cen5_80" automode_scorecut="0.0" scorefxn="cen"
    mcscorefxn="cen" fasticsearch="dens_soft" strategy="auto" fragbias="density"
    rms="%%rms%%" ncycles="200" fullatom="0" bbmove="1" nminsteps="25" temp="4" fraglens="7"
    nfrags="25"/>
  <ReportFSC name="report" testmap="%%testmap%%" res_low="10.0" res_high="%%reso%%"/>
  <BfactorFitting name="fit_bs" max_iter="50" wt_adp="0.0005" init="1" exact="1"/>
  <FastRelax name="relaxcart" scorefxn="dens" repeats="1" cartesian="1"/>
</MOVERS>
<PROTOCOLS>
  <Add mover="setupdens"/>
  <Add mover="loaddens"/>
  <Add mover="tocen"/>
  <Add mover="cenmin"/>
  <Add mover="relaxcart"/>
  <Add mover="cen5_50"/>
  <Add mover="relaxcart"/>
  <Add mover="cen5_60"/>
  <Add mover="relaxcart"/>
  <Add mover="cen5_70"/>
  <Add mover="relaxcart"/>
  <Add mover="cen5_80"/>
  <Add mover="relaxcart"/>
  <Add mover="relaxcart"/>
  <Add mover="report"/>
</PROTOCOLS>
<OUTPUT scorefxn="dens"/>
</ROSETTASCRIPTS>

```

MAINMAST^{20,21} is a de novo method for main-chain modeling from density maps at resolution $<5\text{\AA}$ by capturing the distribution of salient density points in the map. Following the instructions in their published papers^{20,21} and the newest tutorial in <https://kiharalab.org/emsuites/mainmast.php>, a set of C_{α} trace models are first generated by MAINMAST using different combinations of parameters. Then the top 500 models are selected based on the threading score to construct other atoms by PULCHRA¹⁸. Finally, the full-atom models are refined by MDFF^{9,10}, and the final model is determined according to the energy score of MDFF.

Supplementary Text 7. Reasons for the better performance of DEMO-EM over other methods

There are several reasons for the better performance of DEMO-EM over MDFF, Rosetta, and MAINMAST. First, the quick quasi-Newton searching process in combination with a space enumeration algorithm as taken by DEMO-EM (see Methods) can correctly match individual domains into the density map and thus generate optimal initial full-length models for the majority of proteins. As shown in **Supplementary Fig. 5a**, the average TM-score of initial full-length models constructed by the quasi-Newton search were 0.97 and 0.62 (where 99.2% and 68.8% had a TM-score >0.5) when starting with experimental and predicted domains, respectively, which are 14.1% and 17.0% higher than those by the start-of-the-art structure-map docking program Situs. This is probably because Situs is designed for the density map fitting of the macromolecular complexes mainly on the contour information extracted by the Laplacian filter¹¹. Compared to the big-size complexes, the structure of individual domains is relatively small and some of the domains are buried inside the density map

without high correlation to the contour of the density map, which makes it difficult to correctly fit the domains into the density map through the contour match. Meanwhile, many of the predicted domain models have poor local structures including surfaces, although they may have a correct fold, which also impacts the accuracy of the contour-based model fitting. In contrast, DEMO-EM takes a L-BFGS based searching algorithm which enables the model fitting based on the density correlation between the domain model and density map; this has a more robust correlation with the global fold and is less sensitive to the local structure and contour errors. In **Supplementary Fig. 6**, we present two examples, one being a 4-domain protein (PDBID: 1z1wA) with model built from experimental domain structure and another being a 3-domain protein (PDBID: 1zy9A) from D-I-TASSER predicted domain structure. Due to the large domain size and the low-quality domain models, respectively, Situs created incorrect fits for several domains and resulted in poor RMSDs (28.2 Å and 18.0 Å) of the overall models, while the DEMO-EM domain search on model-map correlations generated much closer matches with RMSD=0.1 and 3.3 Å, respectively, for the two proteins. The quality of the initial models created at this step can have an important impact on the final refined models since the energy landscape of the refinement force field usually has a funnel shape and can only refine the models close to the correct positions²². For example, most of the final models with TM-score >0.7 refined by MDFF and Rosetta are from these cases with starting models also have TM-score >0.7 (**Supplementary Fig. 7**). As MDFF, Rosetta, and MAINMAST mainly focus on the high-resolution density map refinement or modeling, we also show a head-to-head comparison of TM-scores between DEMO-EM and MDFF, Rosetta, and MAINMAST for the cases with resolutions of density maps ranging from 2 to 5 Å in **Supplementary Figs. 4f-4h**. These figures indicate that DEMO-EM outperforms MDFF, Rosetta, and MAINMAST on >89% of the proteins with the relatively high-resolution density maps.

Second, DEMO-EM took a hierarchical process of rigid-body and flexible model assembly simulations to progressively refine the multi-domain structures. In particular, the rigid-body domain assembly process can quickly adjust domain poses based on density maps. When starting with the predicted domains, for example, the average TM-scores of full-length models were improved from 0.62 to 0.75 where the number of cases with TM-score >0.5 increases from 68.8% to 85.4% after the rigid-body assembly step (**Supplementary Fig. 5b**). This domain model improvement helps the subsequent DEMO-EM steps to detect unreasonably folded regions in initial full-length models for more efficient atomic-level structural assembly and refinement simulations. Compared to *de novo* methods, the rigid-body assembly of DEMO-EM using the predicted domain models created an initial model with correct topology for most of the proteins, even for the cases with low resolution density maps, while modeling of low resolution density maps is usually hard for *de novo* methods²³. For example, most of models constructed by MAINMAST with TM-score >0.5 are the targets with resolutions of density maps <5.0 Å (**Fig. 2e**).

The last important advantage of DEMO-EM comes from the flexible assembly and refinement stage, which showed a high ability to improve full-length models and individual domain models simultaneously with the assistance of deep-learning based inter-domain distances predictions when coupled with density-map correlation and inherent DEMO-EM force field. To examine the efficiency of the flexible assembly and refinement process, we feed the same full-length models assembled by the DEMO-EM rigid-body assembly step on the predicted domains into MDFF and Rosetta. Although the better starting model quality resulted in a considerable improved final model for both MDFF and Rosetta, which have the average TM-score of 0.81 and 0.79 respectively, the overall quality is still worse than that of DEMO-EM with an average TM-score 0.85 (**Supplementary Table 4**). Compared to the rigid-body assembled models, the last stage of DEMO-EM simulations improved the TM-score by 13.3%, which is 66.7% higher than that by MDFF (8.0%) and 150% higher than that by Rosetta (5.3%), and DEMO-EM has more models with TM-score >0.9 (**Supplementary Fig. 5c**). The average structure refinement time spent by DEMO-EM is 2.40 h, which is comparable to Rosetta (2.36 h) but more than two-fold shorter than MDFF (5.51 h) (see **Supplementary Table 4**). These results demonstrate the efficiency of the atomic-level domain structure refinement of DEMO-EM even starting from the same full-length models.

Supplementary Text 8. Examples showing the construction process of DEMO-EM

In **Figs. 3b** and **3c**, we present two illustrative examples showing the construction process of DEMO-EM. For the *cation-independent mannose 6-phosphate receptor* (PDBID: 1q25A), a protein with 3 domains using a simulated density

map with a low resolution of 9.9 Å (**Fig. 3b**), two domains from N- and C-terminus were initially assigned into the same map space as they shared a similar fold with TM-score of 0.88; this resulted in a low TM-score=0.63 of full-length model. Guided by the global model-density correlation and particle movements implemented in the rigid-body assembly step, the incorrect domain fit was corrected with the full-length TM-score improved to 0.87. After the flexible assembly and refinement simulations, almost all wrong folding regions in the model were corrected which resulted in a global TM-score=0.96 and RMSD=1.9 Å. Again, the average TM-score/RMSD of the individual domains were significantly improved in this example from 0.82/3.0Å to 0.95/1.0Å, respectively.

Figure 3c shows another example from *human xanthine oxidoreductase mutant F3* (PDBID: 2e1qC), a complex protein with 8 continuous domains and 2 discontinuous domains (one of them has three discontinuous segments) using a simulated density map with a medium resolution of 5.3 Å. One of the domains was initially docked into an incorrect region in the quasi-Newton based search, resulting in a suboptimal full-length TM-score of 0.89 and RMSD=5.4 Å. After the second step of rigid-body assembly, the domain was moved to the correct map space but still with some regions stretched outside the maps in several domains and have the model with a TM-score to 0.95 and RMSD=3.8 Å. At the last step, the flexible simulations refined the overall quality including drawing the exposed loops into the density map, which resulted in a further improved model with TM-score=0.98 and RMSD=2.7 Å. In this case, model quality of the 10 individual domains is also improved with average TM-score/RMSD improved from 0.84/3.2Å to 0.95/1.6Å, respectively.

In **Supplementary Figure 8**, we also present the full-length models created by MDFF, Rosetta, and MAINMAST for these two examples, which have a TM-score/RMSD equal to 0.16/39.1Å, 0.09/83.2Å, and 0.18/36.2 Å for 2e1qC, and 0.38/21.2Å, 0.36/21.6Å, and 0.10/58.3Å for 1q25A, respectively; the low-quality models are mainly due to the initial full-length models with incorrect domain orientations for MDFF and Rosetta and low resolution maps for MAINMAST to distinguish the main chain. The results of these case studies reinforce the advantage of the DEMO-EM pipeline for assembling multi-domain protein complex structures.

Supplementary Text 9. Benchmark set with experimental density maps

To further examine the use of DEMO-EM on practical density maps, we collected a set of 51 non-redundant multi-domain proteins from EMDB that have experimental density map with resolution ranging from 2.9 to 10 Å (**Supplementary Fig. 9a** and **Supplementary Table 5**). The size of these proteins ranges from 144 to 1,664 residues with the number of domains ranging from 2 to 8. To emulate the common real-life scenarios where the domain structures of target proteins are unknown, we predict the domain boundaries from sequence by a deep-learning contact-based program FUpred²⁴ and a threading template-based method ThreaDom²⁵ (see Methods). The individual domain structures modelled by D-I-TASSER with all homologous templates of a sequence identity >30% to the query have been excluded.

Supplementary Text 10. Impact of the domain assignment and map segmentation on the accuracy of the final model

To further study the impact of the map segmentations on the accuracy of the final model, we also reassembled all 51 test cases by DEMO-EM using the full density maps. As shown in **Supplementary Fig. 11a**, the average TM-score (0.86) is comparable to that using segmented density maps (0.88), where the difference corresponds to a Student's t-test p-value of 0.3. For the domain-level models, the change on the average TM-score is also not significant (0.82 vs. 0.84), corresponding to a Student's t-test p-value of 0.2 (**Supplementary Fig. 11b**). This result is largely because the domain-map matching and domain rigid-body assembly procedure in DEMO-EM can correctly determine the position of each domain in the density map. Meanwhile, the comprehensive energy function from the inter-domain distance profile and the domain structure restraints can further prevent the global topology deviating too much away from the correct domain orientations even with full density maps. There are indeed a few cases (including 4vImC and 6n89A) with TM-score of domain model and full-length model obviously decreased (**Supplementary Fig. 11**). For all these cases, there are at least one domain whose structure was incorrectly predicted by D-I-TASSER, which makes the domain-map match more difficulty in a larger map, where the incorrect domain-map match further impacts the cryo-EM based domain assembly and refinement simulations.

Although full density map-based search may help amend the domain-map match for some of these cases, this will increase the overall running time of the DEMO-EM.

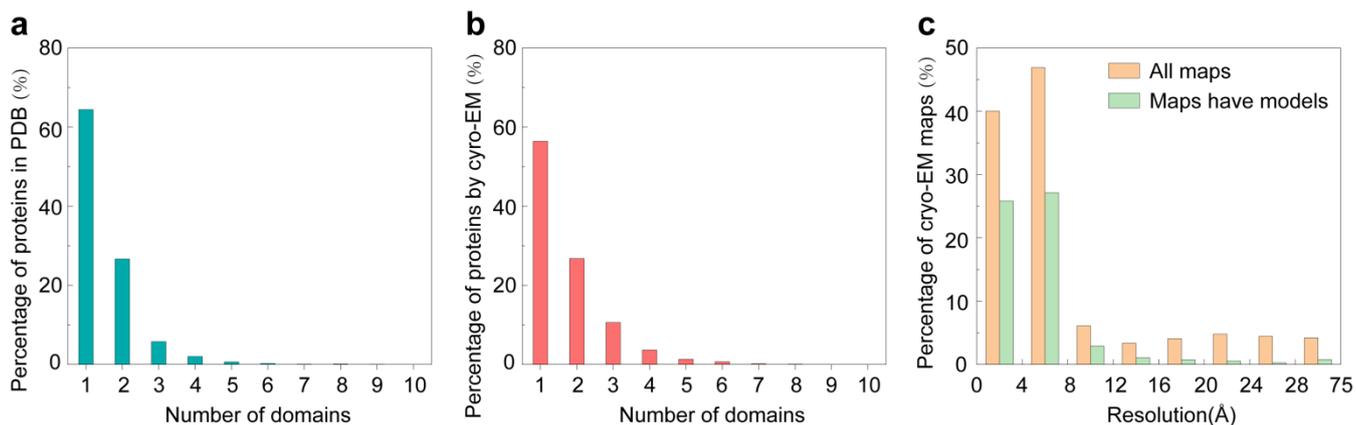
We also investigated the impact of the incorrect domain assignment on the final model by reconstructing the model of the 8 cases, which have the predicted number of domains not consistent with that of DomainParser, using the domain definitions determined by DomainParser. The results show that the use of correct domain definition can indeed improve the modeling accuracy with TM-score increased from 0.836 to 0.845 for full-length models and from 0.769 to 0.772 for domain-level models; but the difference is not dramatic. One reason is that the domain boundary definition is not always absolute and some of the DEMO-EM predictions have provided a reasonable alternative (see **Supplementary Figs. 12a-c**). On the other hand, the highly efficient rigid-body domain reassembly step of DEMO-EM can often amend the incorrectly assigned initial domain positions under the guidance the experimental density maps (see **Supplementary Figs. 12d-g**).

For example, the protein (3j1fA) shown in **Supplementary Fig. 12a** was split into 2 domains by DomainParser, while one of the domains was further divided into 2 domains by DEMO-EM and results in 3 domains (**Supplementary Fig. 12b**). However, both the experimental and predicted inter-residue contact map show that these two domain definitions are reasonable since there are very few inter-domain contacts (**Supplementary Fig. 12c**). In another example, the protein (3jb9B) shown in **Supplementary Fig. 12d** was split into 4 domains by DomainParser, while two of the domains determined by DomainParser (**Supplementary Fig. 12d**) were further split into two small domains (**Supplementary Fig. 12e**) by DEMO-EM because there are few inter-domain contacts in the predicted contact map (**Supplementary Fig. 12f**). **Supplementary Fig. 12g** shows the three stages of the DEMO-EM process for this case. Although D-I-TASSER was able to create correct domain models with an average TM-score=0.88, the three small domains were fit into the wrong positions, resulting in an initial model with TM-score =0.42, because positions of small domains are usually more difficult to identify in the whole density map compared to large domains. However, the incorrect domain fit was corrected in the density map guided rigid-body domain assembly step, resulting in a full-length model with TM-score improved to 0.92; the TM-score was increased to 0.95 after the flexible domain assembly and refinement.

Supplementary Text 11. Validation scores of DEMO-EM models for proteins from SARS-CoV-2 genome

We also compared the validation scores of DEMO-EM models with that of the deposited models for all proteins in SARS-CoV-2 genome. On average, DEMO-EM models contain more ‘Ramachandran favored’ residues, less rotamer outliers and clashes (**Supplementary Table 8**), which result in a lower MolProbity score (0.92) compared to that of deposited models (1.71). Looking at the results on each case, the MolProbity score of DEMO-EM model is lower than that of deposited model on all 6 cases, partly because DEMO-EM model includes much less clashes. In terms of EMringer score, DEMO-EM has 4 out of 6 cases higher than the deposited model, indicating DEMO-EM model achieves better model geometry and density-fit at side-chain level. However, the overall iFSC of DEMO-EM model is dropped compared to that of the deposited model on 3 out of 6 cases. To investigate the possible reason, we calculated the Q-score and FSC-Q²⁷ to evaluate the local quality of all proteins. The result indicates that the DEMO-EM model achieves higher average Q-score than the deposited model for all cases. Particularly for the 3 cases which have the iFSC dropped in the DEMO-EM model, the DEMO-EM models have >72% residues obtaining higher Q-score than deposited models (**Supplementary Figs. 16a-c**). For the case (6vsbC) that has deposited half maps, most of the residues (68.9%) of the DEMO-EM model obtain better FSC-Q than that of the deposited model (**Supplementary Fig. 16d**). For example, **Supplementary Fig. 16e** shows an example residue (THR-998) of 6vsbC, in which all atoms of the deposited model were fit into the density map but obtain a negative FSC-Q score. Therefore, we envision that the higher iFSC scores for the deposited models in these cases, for which DEMO-EM model have better model geometry and local quality, are probably because of the overfitting in some local regions of the deposited models. The DEMO-EM models for all the SARS-CoV-2 proteins are downloadable at https://zhanggroup.org/DEMO-EM/data_set/model_sarscov2.tar.gz.

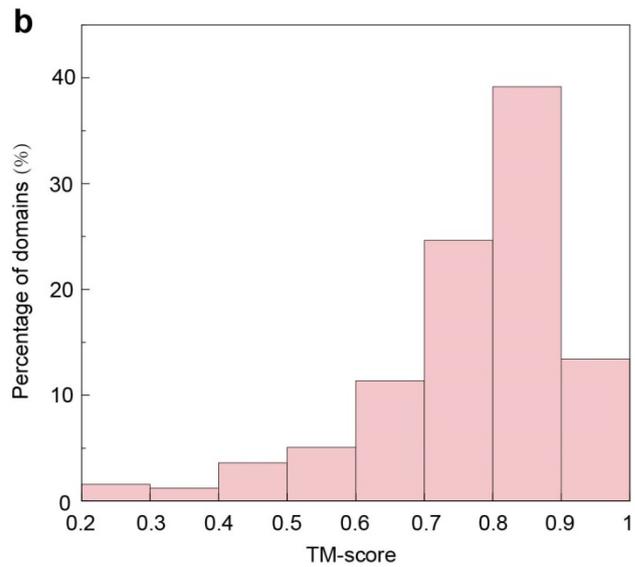
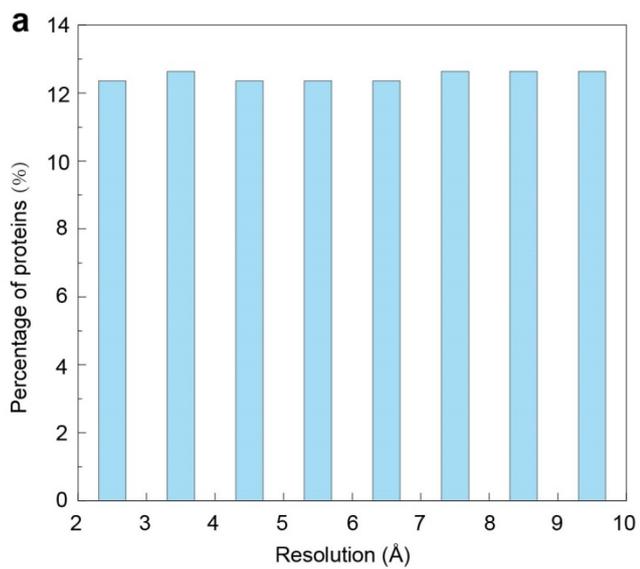
Supplementary Figures



Supplementary Figure 1

The proportion of multi-domain proteins in protein chains in the PDB and the proportion of released cryo-EM maps with associated atomic models in EMDB.

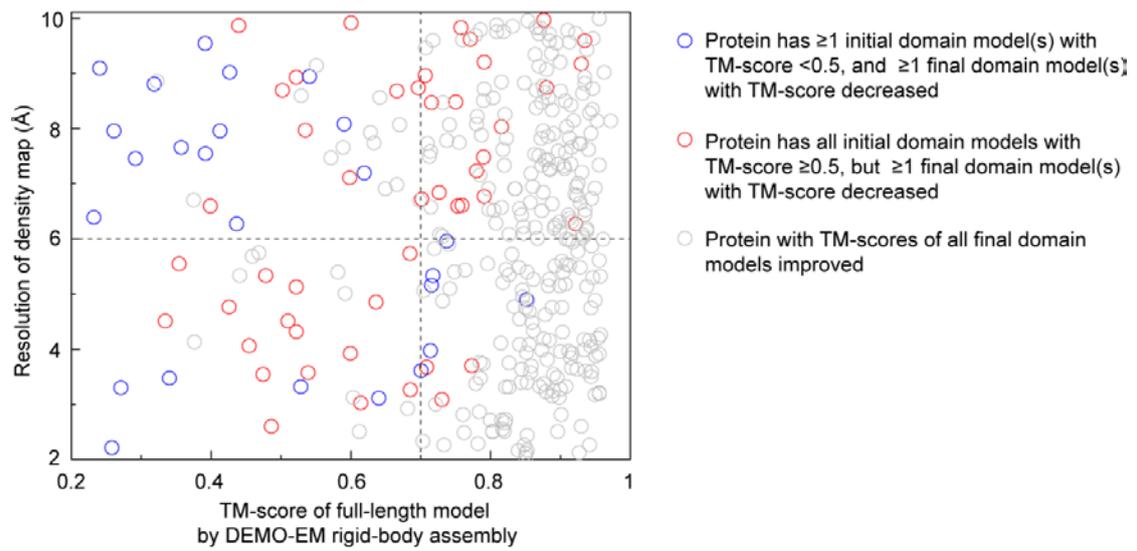
(a) Distribution of multi-domain proteins in protein chains in the entire PDB. (b) Distribution of multi-domain proteins in protein chains determined by cryo-EM, where domain boundaries determined by CATH database³ and DomainParser¹. Here, we just show proteins with number of domains less than 10. (c) Distribution of resolutions for all released cryo-EM density maps and maps have corresponding atomic structures in EMDB.



Supplementary Figure 2

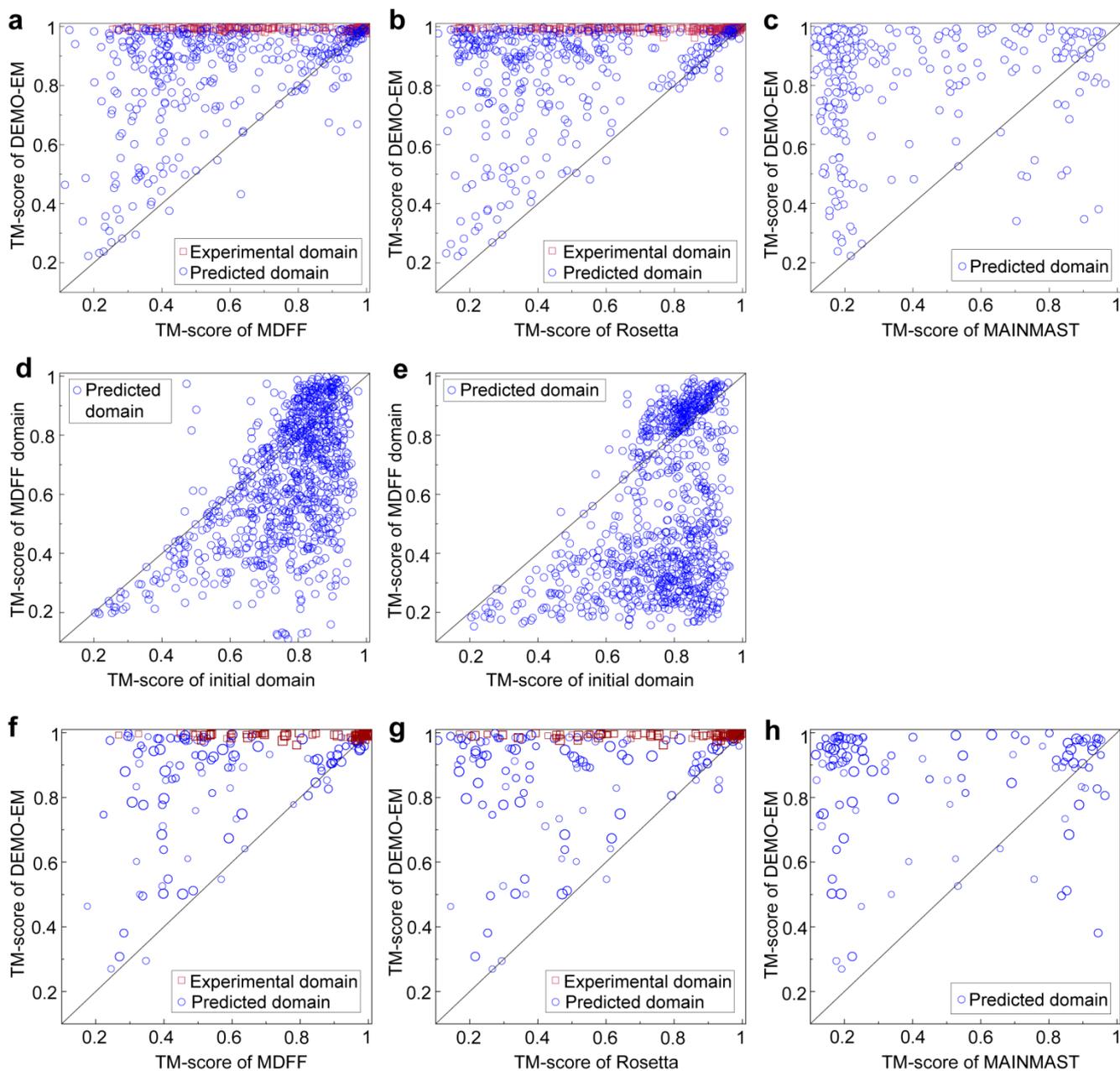
Initial information of the 357 test multi-domain proteins

(a) Distribution of resolutions for simulated density maps. (b) Distribution of TM-scores for all domain models predicted by D-I-TASSER.



Supplementary Figure 3

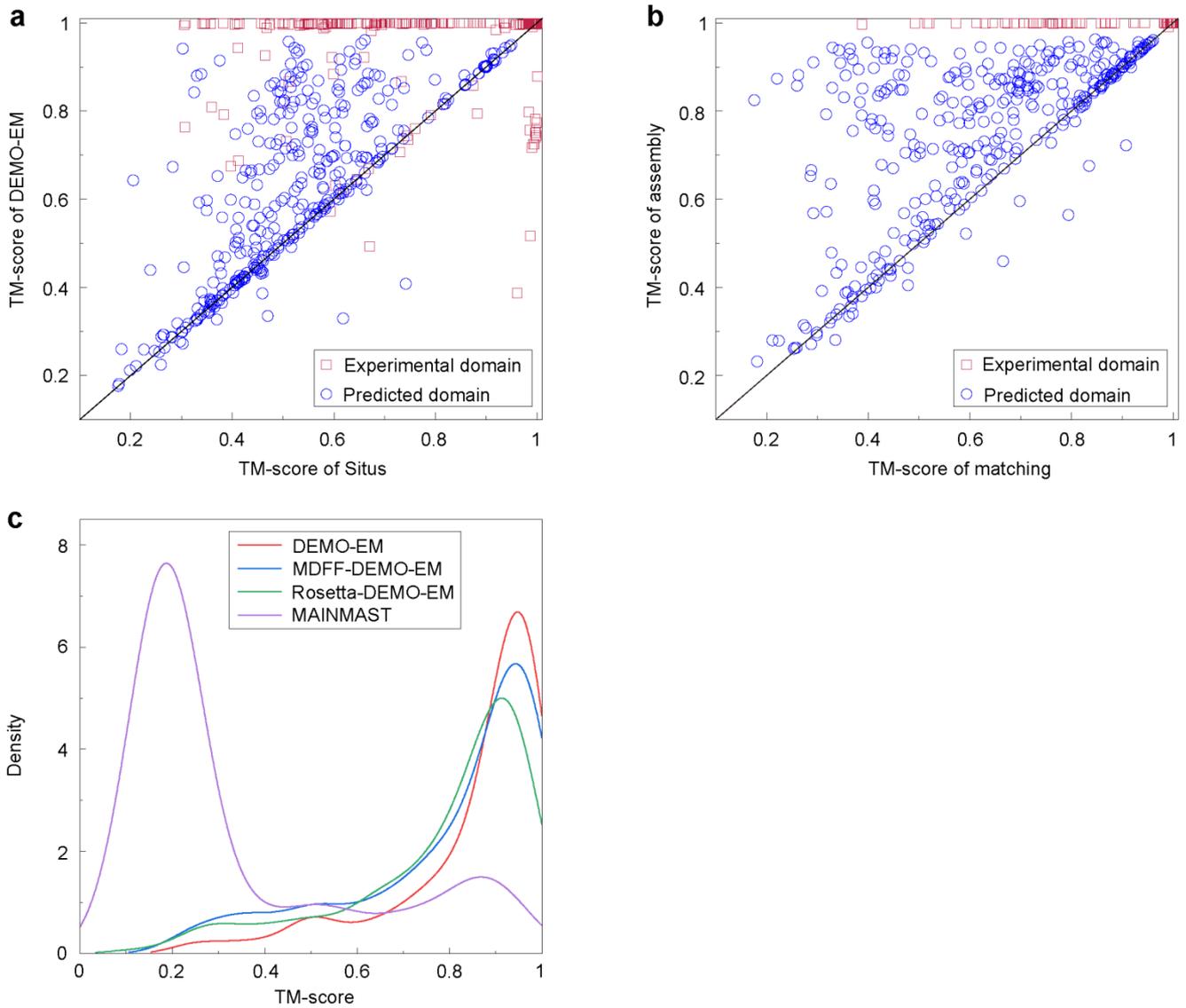
The relationship between the quality of initial data (TM-score of model by rigid-body assembly and map resolution) and the improvement of individual domain models by DEMO-EM flexible assembly for the 357 test multi-domain proteins.



Supplementary Figure 4

Summary of models generated by MDFF⁹, Rosetta¹⁹, and MAINMAST^{20,21} using synthesized density maps for the 357 proteins, where initial full-length models of MDFF and Rosetta are created by Situs¹¹.

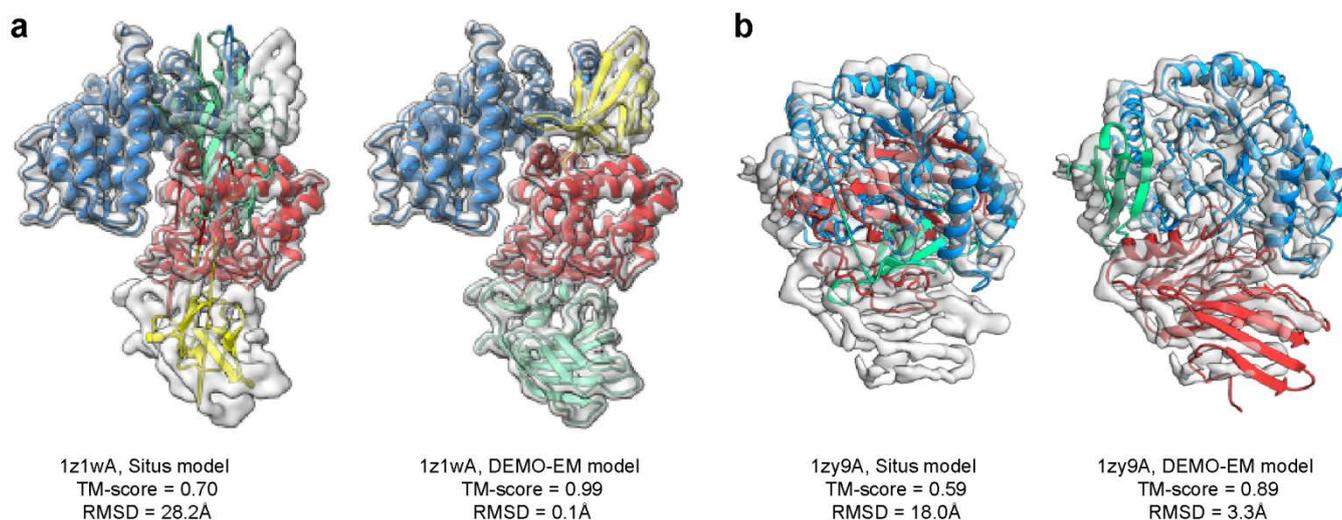
(a) Head-to-head comparison between TM-scores of final models built by DEMO-EM and MDFF. (b) Head-to-head comparison between TM-scores of final models built by DEMO-EM and Rosetta. (c) Head-to-head comparison between TM-scores of final models built by DEMO-EM and MAINMAST. (d) Head-to-head comparison between TM-scores of initial individual domain models and the corresponding domain models in final full-length models by MDFF. (e) Head-to-head comparison between TM-scores of initial domain individual models and the corresponding domain models in final full-length models by Rosetta. (f) Head-to-head comparison between TM-scores of final models built by DEMO-EM and MDFF for cases with resolutions of density maps ranging from 2 to 5 Å. (g) Head-to-head comparison between TM-scores of final models built by DEMO-EM and Rosetta for cases with resolutions of density maps ranging from 2 to 5 Å. (h) Head-to-head comparison between TM-scores of final models built by DEMO-EM and MAINMAST for cases with resolutions of density maps ranging from 2 to 5 Å. In (f-h), larger size symbols indicate cases with lower resolution density maps.



Supplementary Figure 5

Summary of models generated by domain matching, rigid-body assembly of DEMO-EM, and refinement or constructed by different methods for the 357 test proteins.

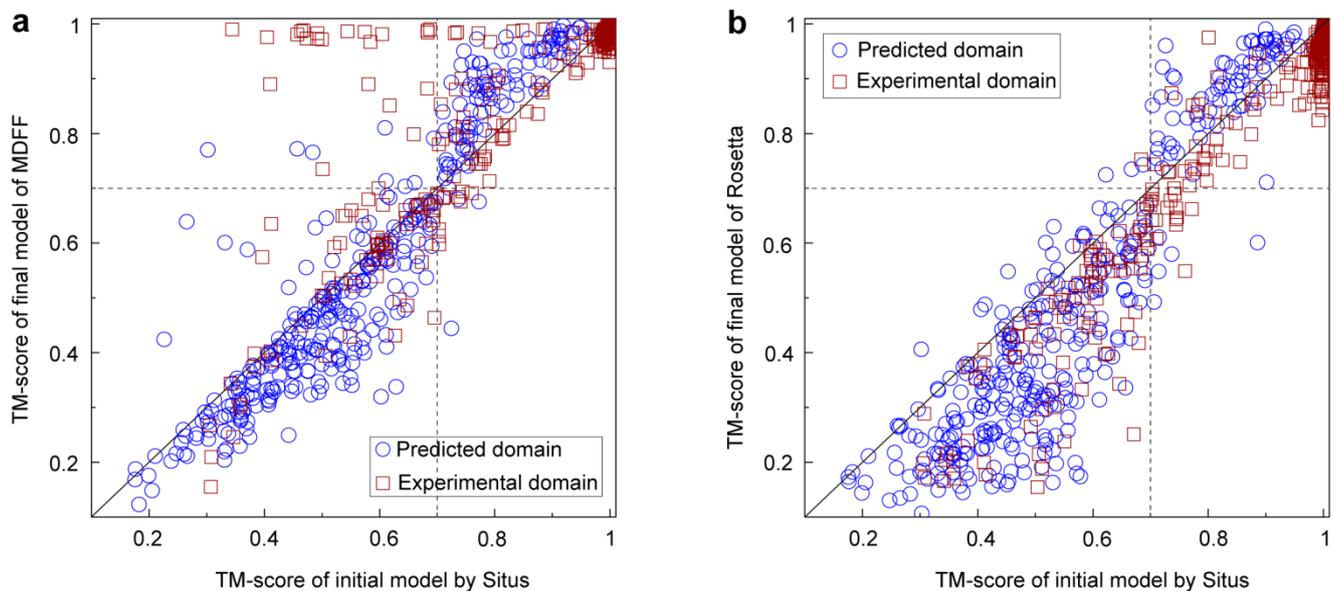
(a) Head-to-head comparison between TM-scores of initial full-length models generated by independently fitting each domain into density maps using DEMO-EM versus that created by Situs using the same domain models. (b) Head-to-head comparison between TM-scores of initial full-length models from domain-map fitting versus that of rigid-body assembled models by DEMO-EM. (c) Distribution of the TM-scores of models refined by DEMO-EM, MDFF, and Rosetta from the full-length models assembled by DEMO-EM in rigid-body using the predicted domain models, and distribution of the TM-scores of models created by MAINMAST.



Supplementary Figure 6

Examples showing initial full-length models constructed by independently matching each domain into the density map by Situs and DEMO-EM, where different domains are labeled with different colors.

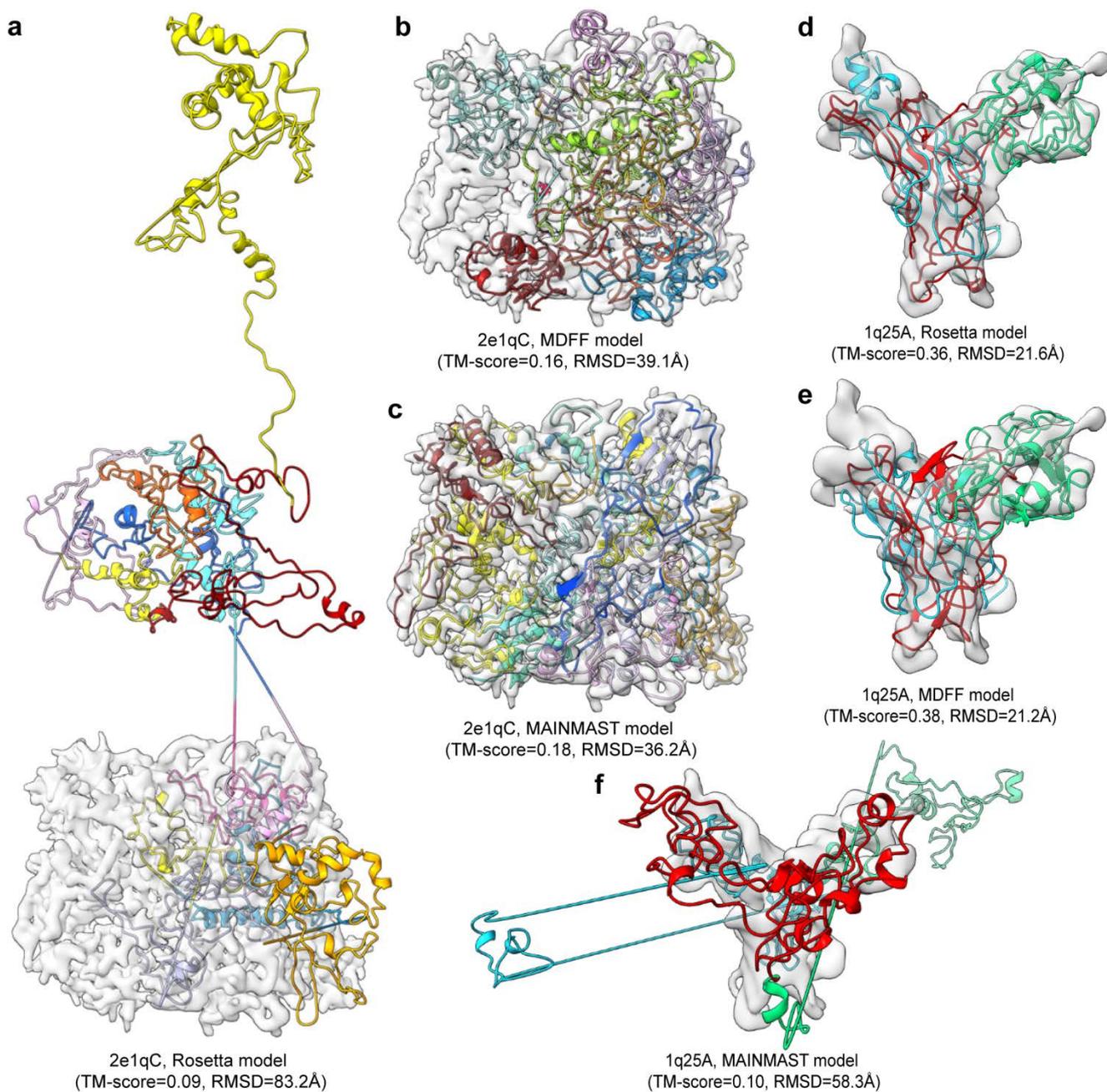
(a) Initial models generated by domain matching through Situs and DEMO-EM for a 4-domain protein (PDBID: 1z1wA) using experimentally determined domain structures. (b) Initial models generated by domain matching through Situs and DEMO-EM for a 3-domain protein (PDBID: 1zy9A) using D-I-TASSER predicted domain models, where the TM-scores are 0.76, 0.88, and 0.60 for red, blue, and green domains, respectively.



Supplementary Figure 7

The correlation between the TM-score of initial model and that of final model by MDFF and Rosetta for the 357 test multi-domain proteins.

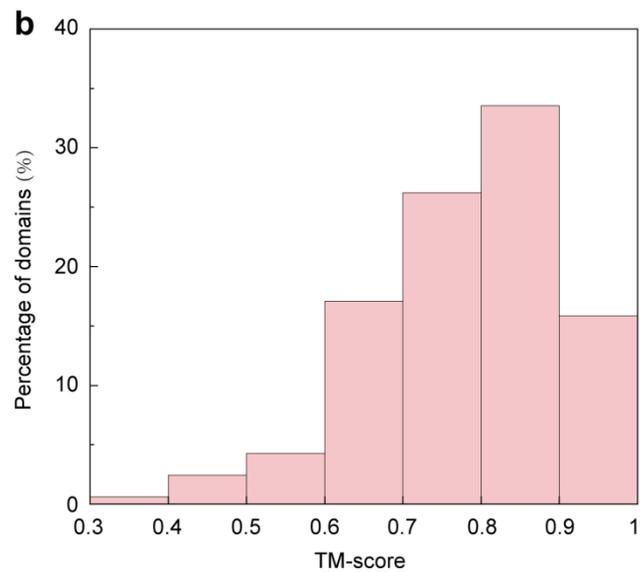
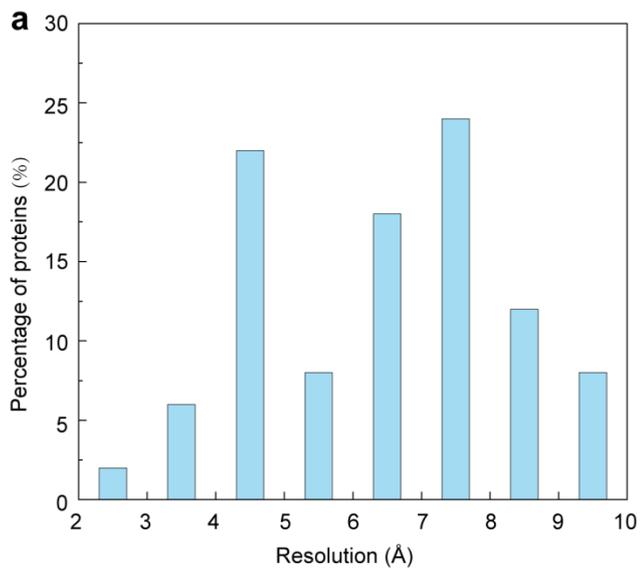
(a) Head-to-head comparison between TM-scores of final models generated by MDFF and TM-scores of its starting models by Situs. (b) Head-to-head comparison between TM-scores of final models generated by Rosetta and TM-scores of its starting models by Situs.



Supplementary Figure 8

Examples to show models refined by MDFF and Rosetta starting from full-length models generated by Situs and the *de novo* model created by MAINMAST.

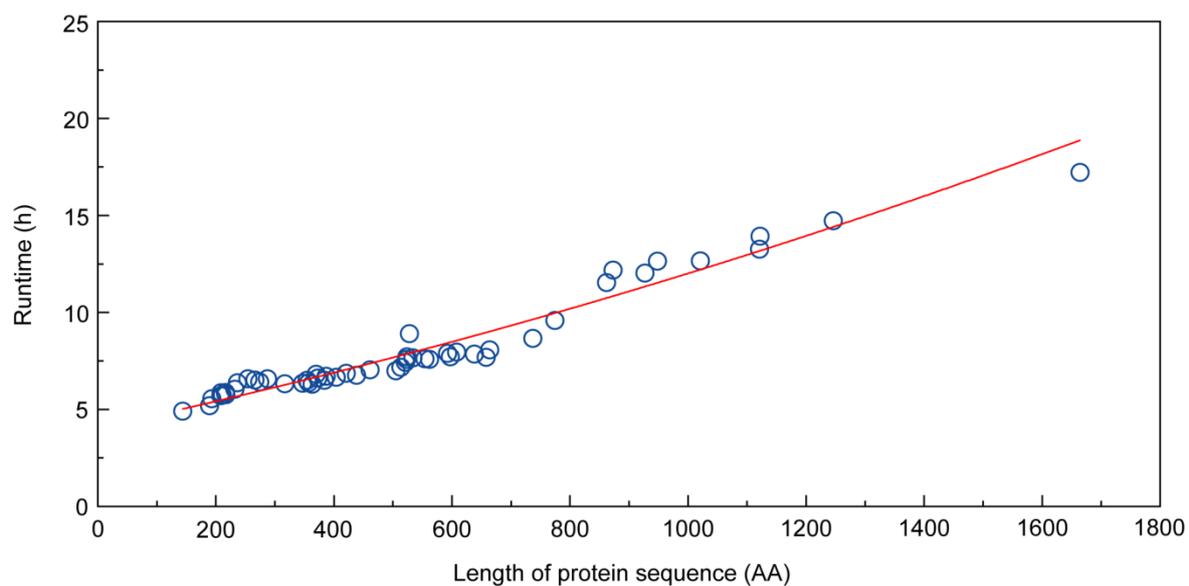
(a), (b), and (c) Models of 2e1qC created by Rosetta, MDFF, and MAINMAST, respectively. (d), (e), and (f) Models of 1q25A created by Rosetta, MDFF, and MAINMAST, respectively.



Supplementary Figure 9

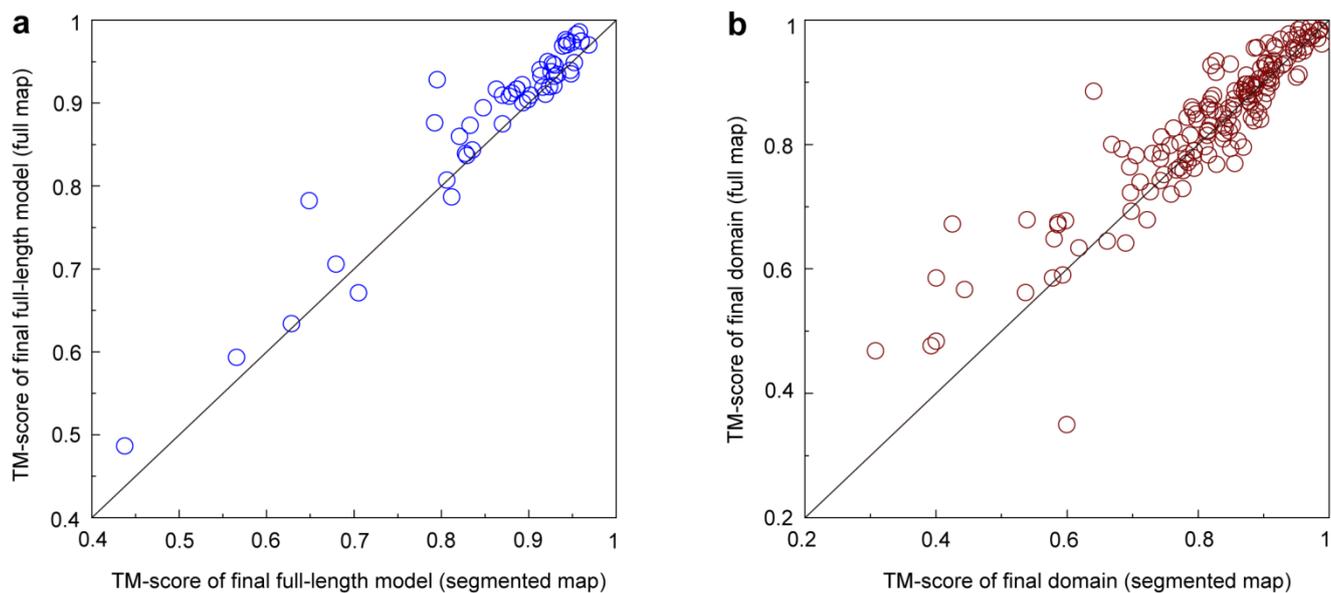
Initial information for the 51 cases with experimental density maps.

(a) Distribution of resolutions for density maps. (b) Distribution of TM-scores for domain models generated by D-I-TASSER.



Supplementary Figure 10

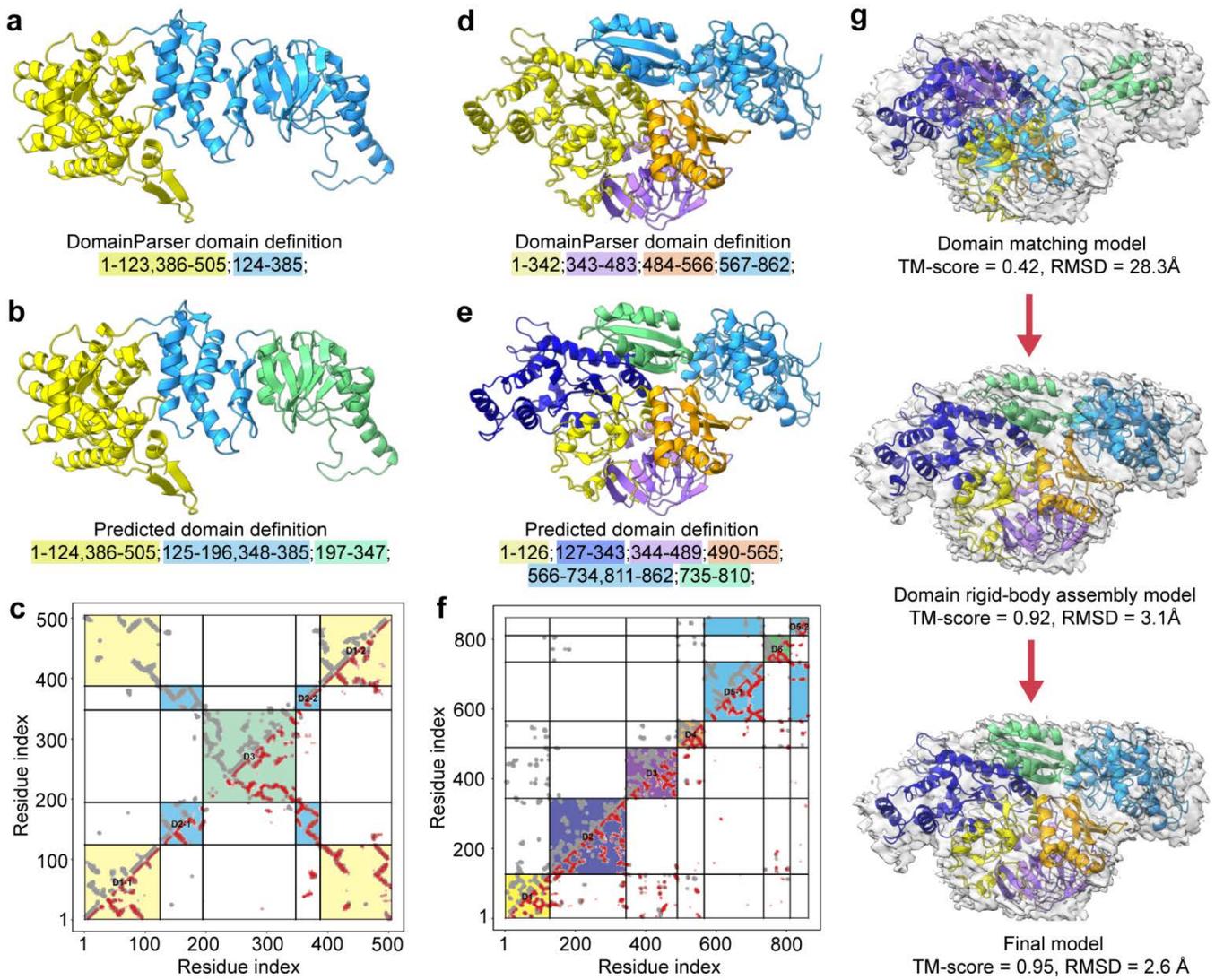
Runtime of DEMO-EM for the 51 cases with experimental density maps. Here, the runtime contains processing time of the whole pipeline of DEMO-EM when starting from the sequence and cryo-EM density map.



Supplementary Figure 11

Comparison of the results of DEMO-EM for the 51 cases when using segmented density maps and full density maps.

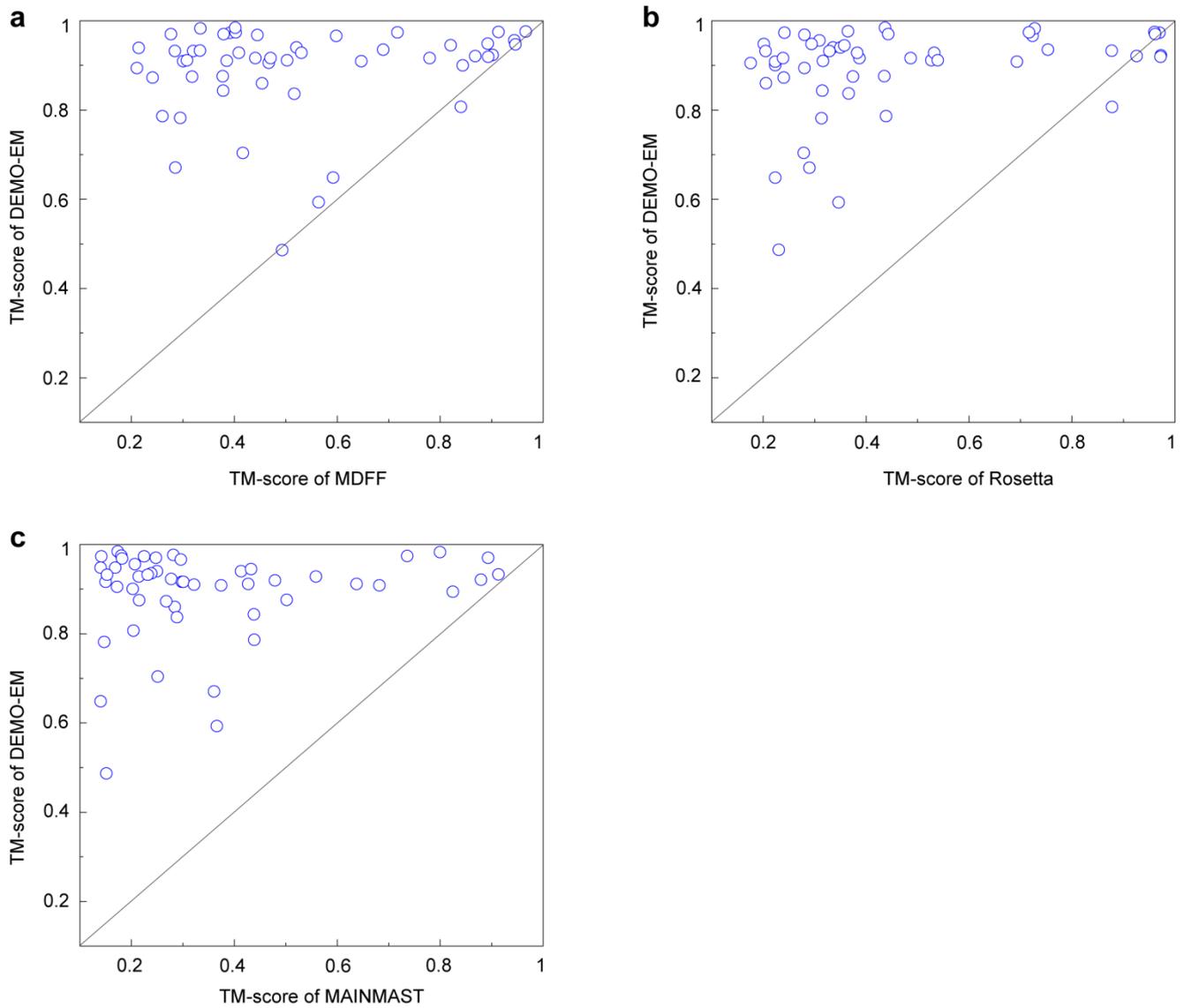
(a) TM-scores of final full-length models generated by DEMO-EM. (b) TM-scores for domain models of final full-length models created by DEMO-EM.



Supplementary Figure 12

Representative examples that the predicted number of domains is inconsistent with that determined by DomainParser using native model.

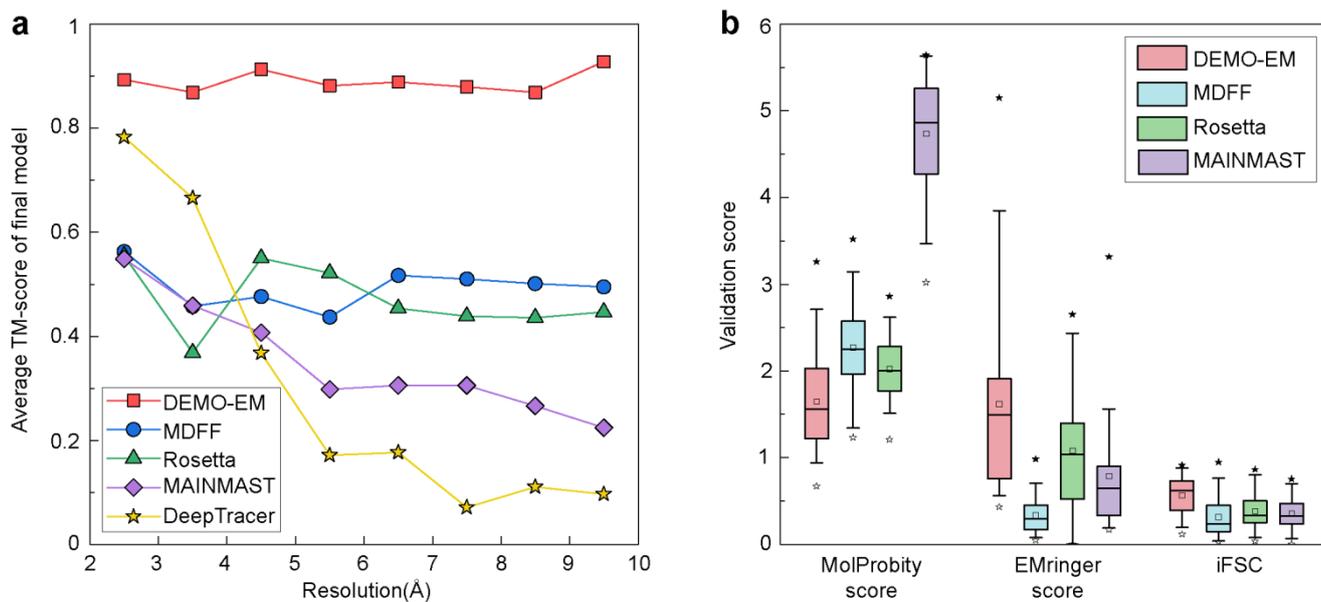
(**a, d**) Domain definitions determined by DomainParser and shown in the native structure for protein 3j1fA (**a**) and 3jb9B (**d**), where different domains represented by different colors. (**b, e**) DEMO-EM predicted domain definitions shown in the native structure for protein 3j1fA (**b**) and 3jb9B (**e**). (**c, f**) Native (gray) and predicted (red) contact maps for protein 3j1fA (**c**) and 3jb9B (**f**), where colored solid squares indicate the predicted domain boundaries by DEMO-EM. (**g**) The process of the DEMO-EM assembly for protein 3jb9B.



Supplementary Figure 13

Comparison of final models generated by different methods for the 51 cases with experimental density maps.

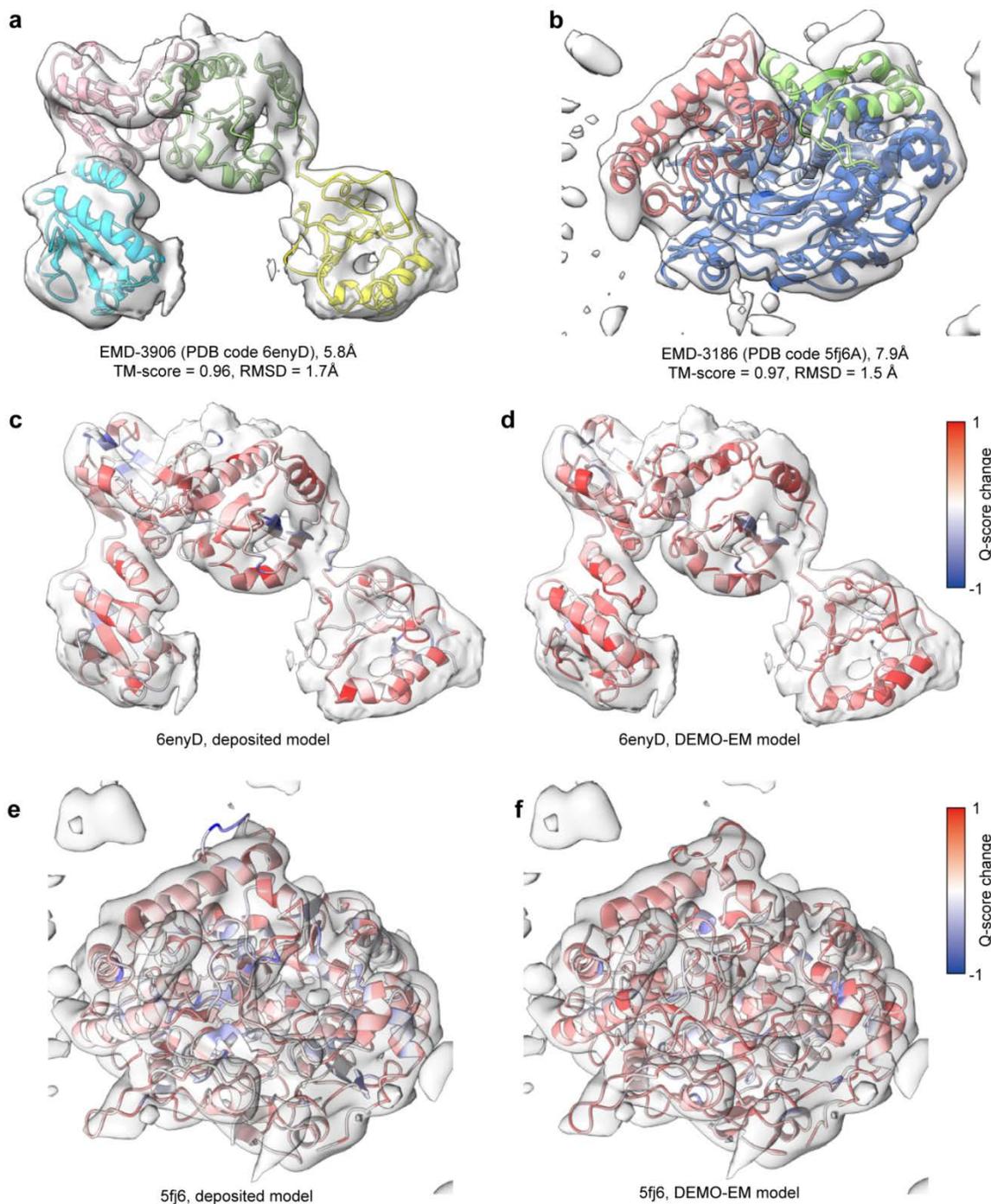
(a) Head-to-head comparison between TM-scores of final model generated by DEMO-EM and that by MDFF using D-I-TASSER predicted domain models. (b) Head-to-head comparison between TM-scores of final model generated by DEMO-EM and that by Rosetta using D-I-TASSER predicted domain models. (c) Head-to-head comparison between TM-scores of final model generated by DEMO-EM and that by MAINMAST.



Supplementary Figure 14

Performance analysis of different methods on the 51 cases with experimental density maps.

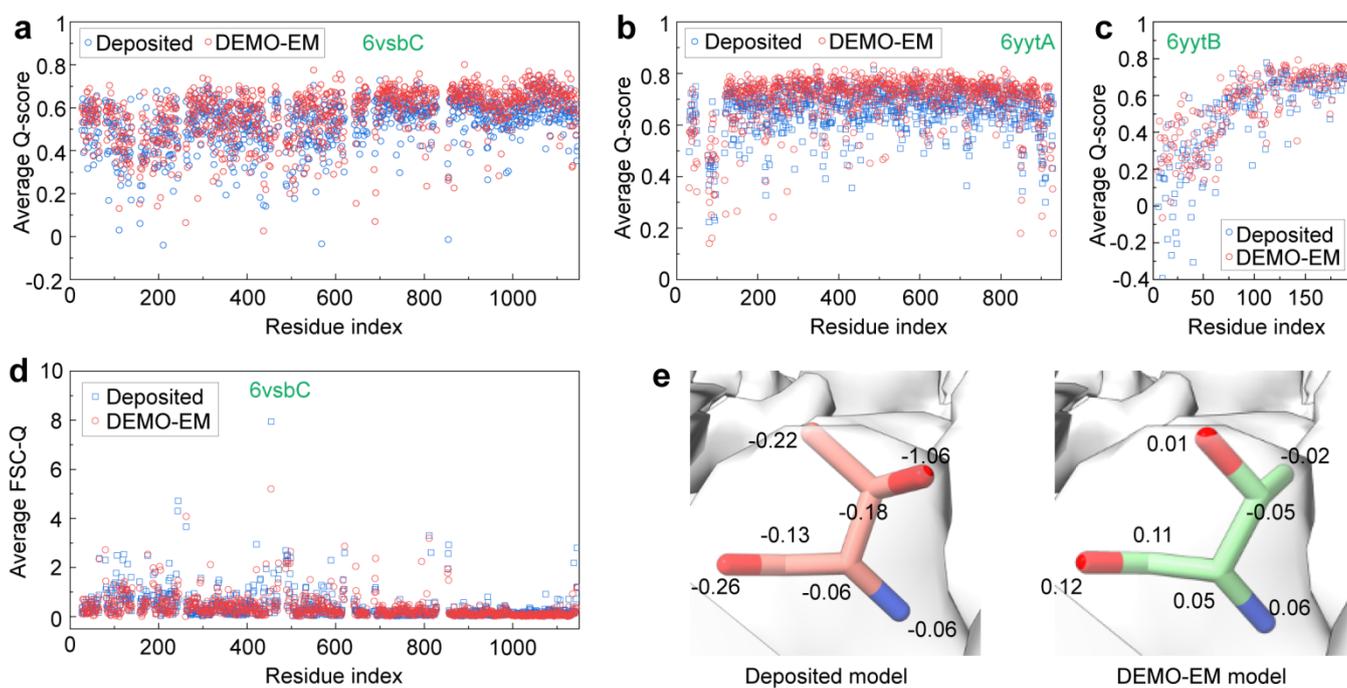
(a) Correlation between the TM-score of the final model and the resolution of density map for different methods. (b) Validation scores of final models generated by different methods. The box represents the lower to upper quartiles of the scores; the horizontal line and square in the box represent the median and mean, respectively; the whiskers indicate the 5th and 95th percentiles; the solid star and hollow star are the maximum and minimum values.



Supplementary Figure 15

Representative examples to show DEMO-EM constructed models using experimental density maps.

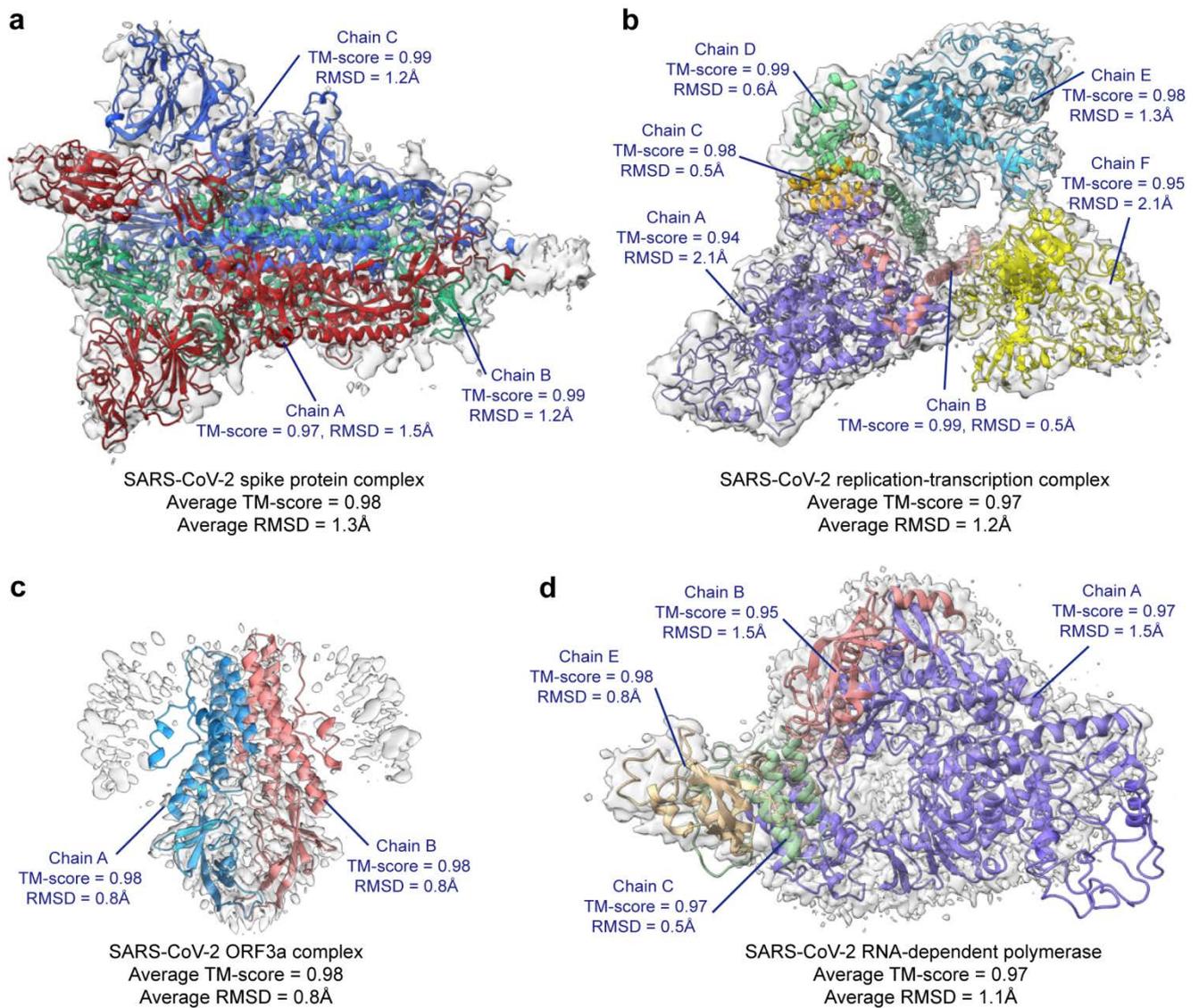
(a) 6enyD (EMD-3906), a protein with 4 continuous domains, where different domains represented by different colors. (b) 5fj6A (EMD-3186), a protein with two continuous domains and one discontinuous domain (green). (c, d) Q-score of the backbone atom showing in the deposited model (c) and DEMO-EM model (d) for 6enyD, where blue and red indicate the minimum (-1) and maximum (1) Q-score, respectively. (e, f) Q-score showing in the deposited model (e) and DEMO-EM model (f) for 5fj6A. It should be noted that only the Q-score of the backbone atoms of 6enyD was calculated because the deposited model only has the backbone atoms. The DEMO-EM models for all the test proteins are available at https://zhanggroup.org/DEMO-EM/data_set/benchmark51.



Supplementary Figure 16

FSC-Q and Q-score of the cases that DEMO-EM models have lower iFSC than deposited models.

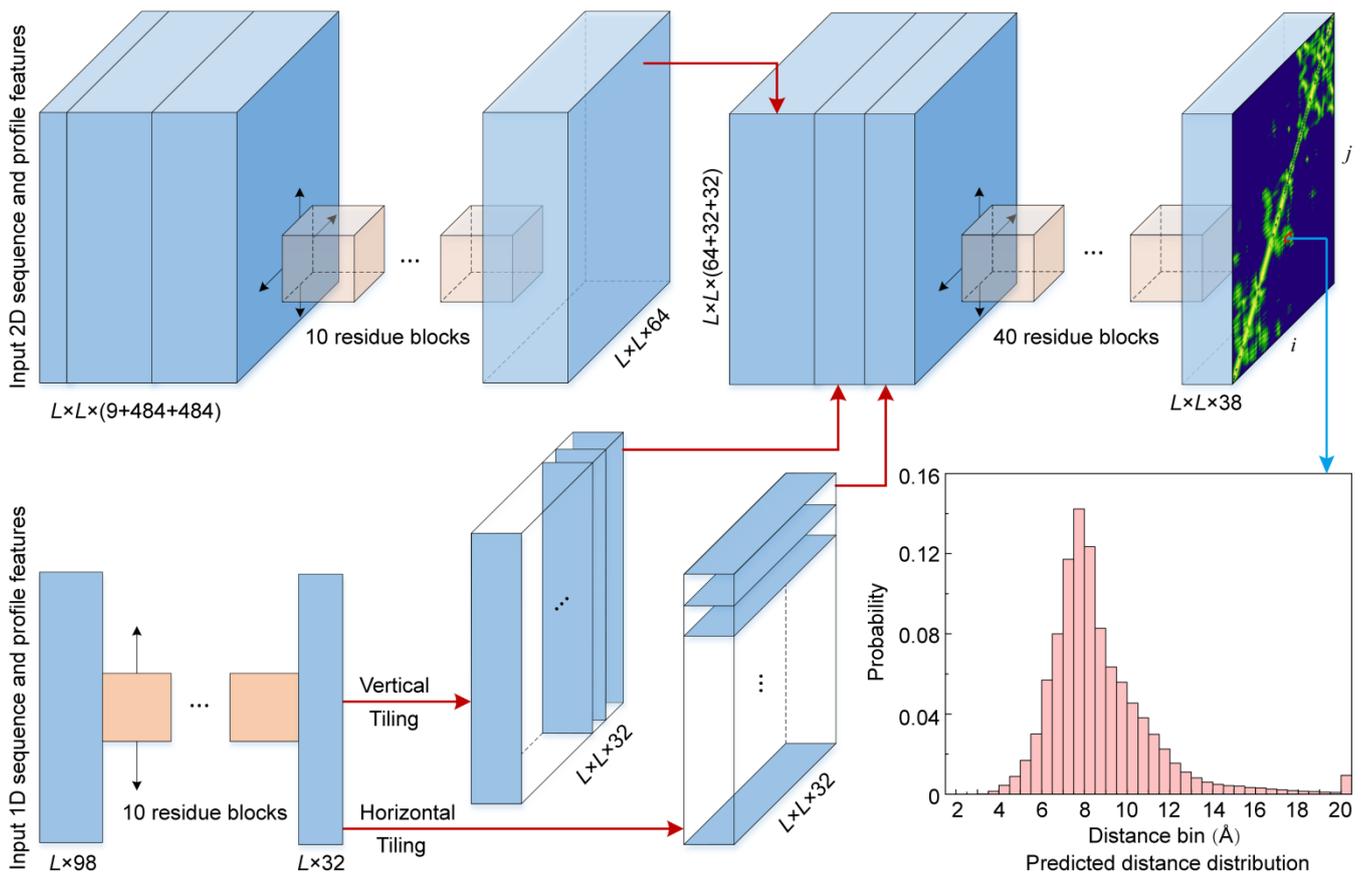
(a, b, c) The average Q-score of each residue of the deposited model (blue) and DEMO-EM model (red) for 6vsbC (a), 6yytA (b), and 6yytB (v). (d) The average absolute FSC-Q of each residue of the deposited model (blue) and DEMO-EM model (red) for 6vsbC. The lower FSC-Q_R indicates better fitting, while the higher Q-score corresponds better fitting. (e) Representative residue (THR-998) of 6vsbC, in which all atoms were fitted into the density map but get a negative FSC-Q. Here, we only report the FSC-Q of the protein with deposited half maps or effective half maps for FSC-Q calculation.



Supplementary Figure 17

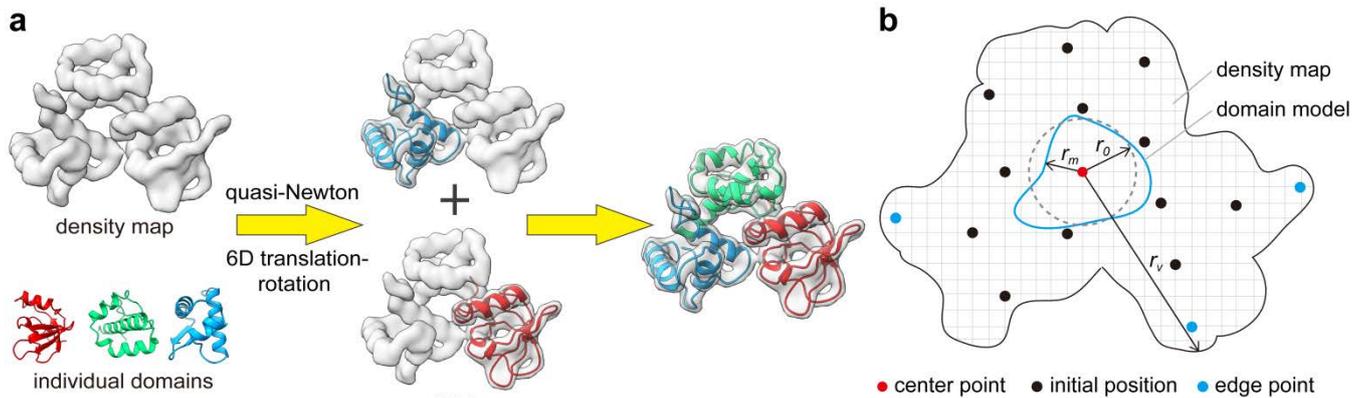
Representative examples to show structures of protein complexes constructed by DEMO-EM

(a) SARS-CoV-2 spike protein complex (EMD-21375), where different colors indicate different chains. (b) SARS-CoV-2 replication-transcription complex (EMD-22160). (c) SARS-CoV-2 ORF3a complex (EMD-22136). (d) SARS-CoV-2 RNA-dependent RNA polymerase (EMD-11007). It should be noted that some regions are not included in the map because we build the structure of the full-length chain even the density data is missed for some residues.



Supplementary Figure 18

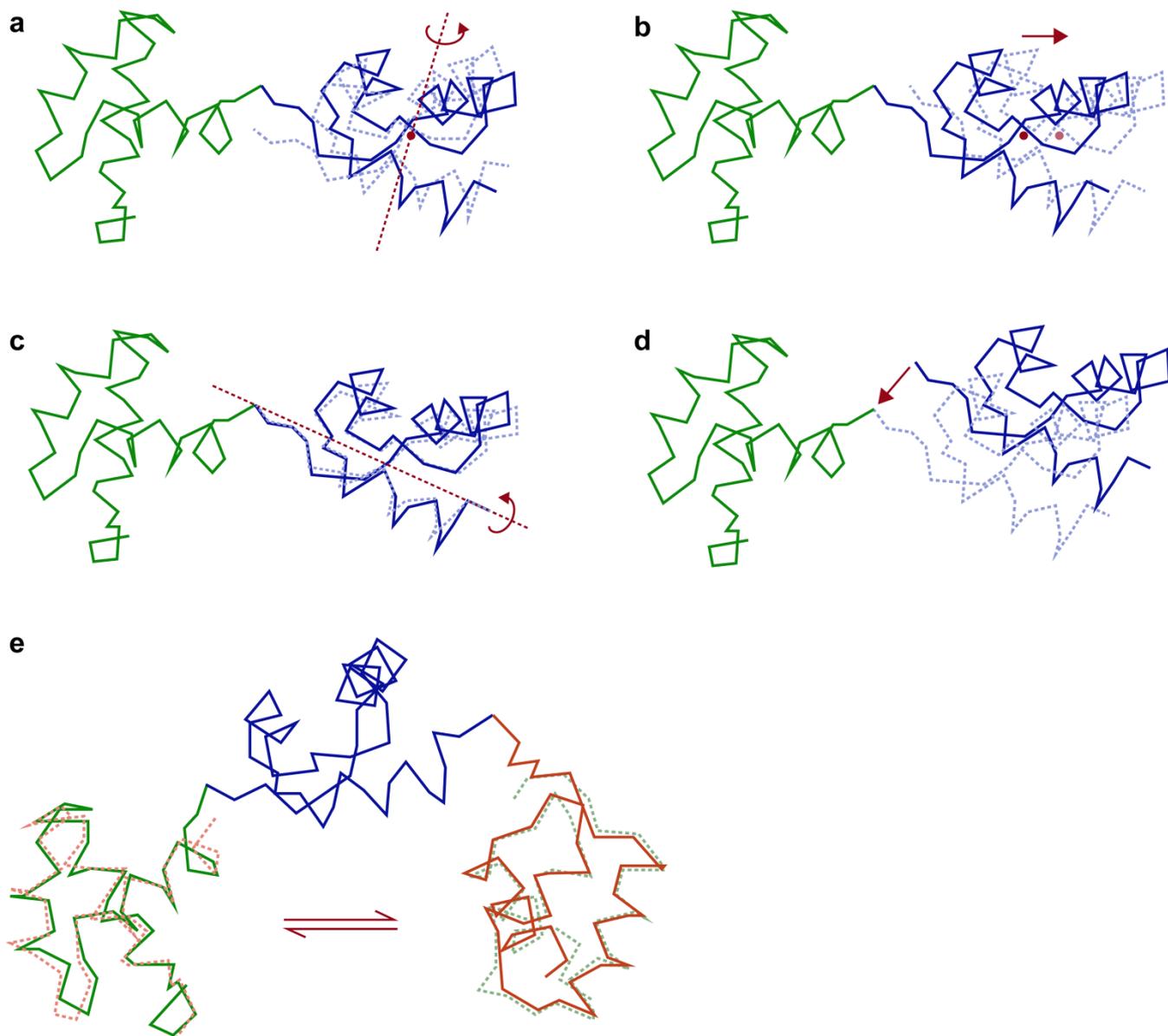
Outline of DomainDist for predicting histograms of inter-domain and intra-domain residue-pair distances using deep-learning based neural networks.



Supplementary Figure 19

Matching domain models into cryo-EM density maps

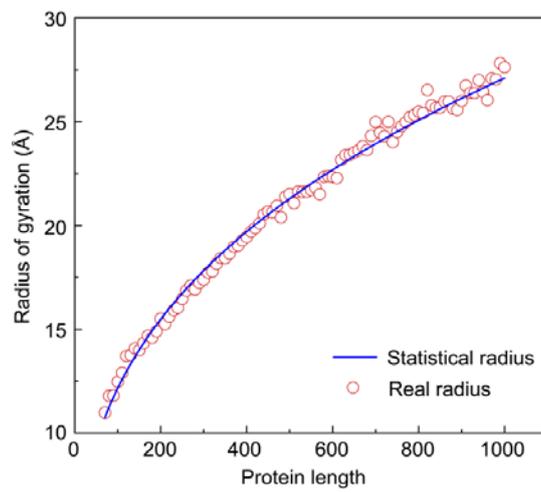
(a) The pipeline of domain structure matching into cryo-EM density maps in DEMO-EM. **(b)** Enumeration of initial position determination for a domain model. To eliminate redundant positions, we set the minimum distance between two neighboring positions, $r = \max(0.5r_m, r_0)$, where r_m is the radius of gyration of the model, and $r_0 = 5\text{\AA}$ is the minimum distance between two initial positions. To remove the edge positions, the maximum distance between each initial position and the center point of the density map is set as $R_{\text{eg}} = \min(\max(1.1(r_v - r_m), r_0), r_v)$, where $r_v = \sqrt{(\sum_{i=1}^{N'_{\text{vol}}} (\mathbf{v}_i - \mathbf{v}_{\text{center}})^2) / N'_{\text{vol}}}$ is the radius of gyration of the density map calculated by the N'_{vol} voxels with density ≥ 0.05 after normalizing density values to the range of 0 and 1, and $\mathbf{v}_{\text{center}} = \frac{1}{N'_{\text{vol}}} \sum_{j=1}^{N'_{\text{vol}}} \mathbf{v}_j$ is the center point of these voxels.



Supplementary Figure 20

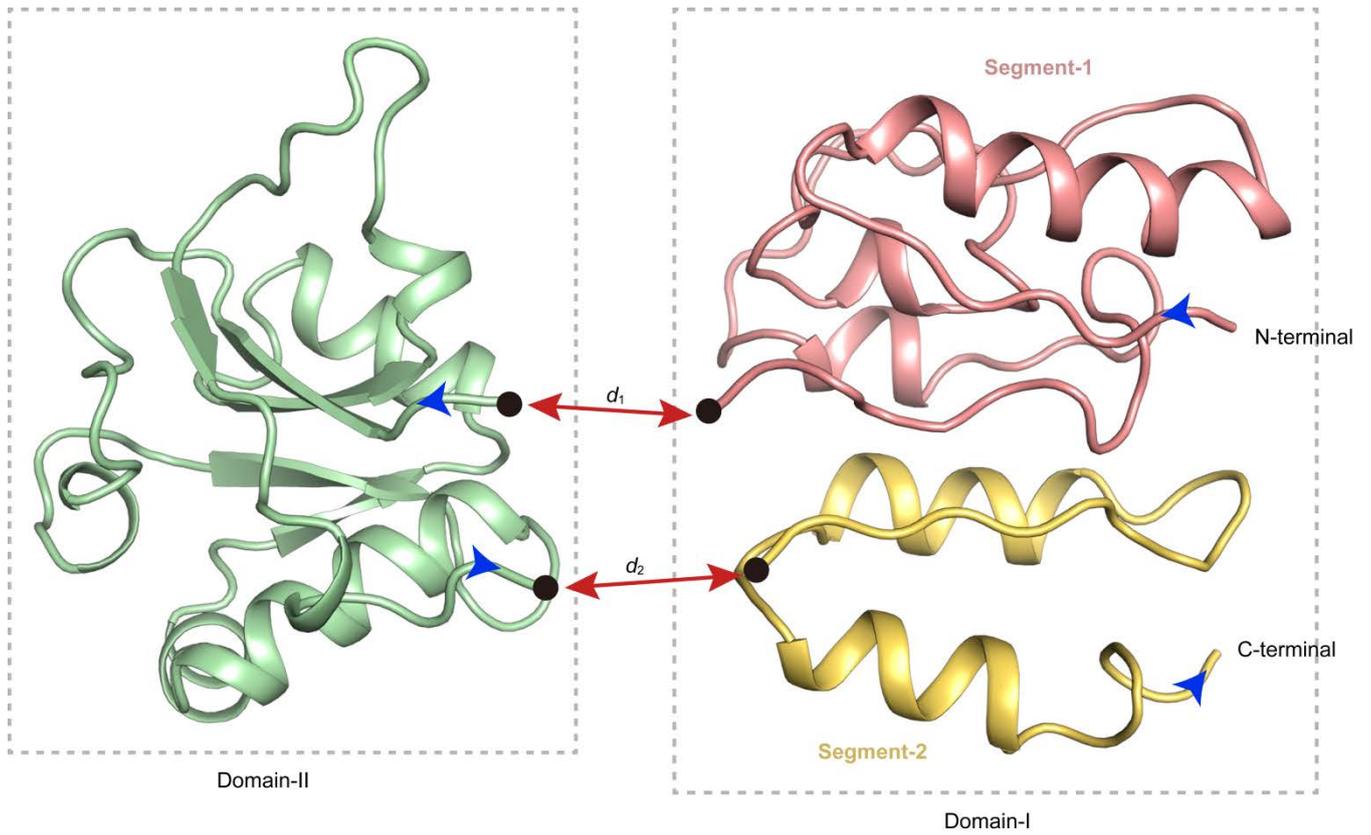
Movements for the rigid-body domain assembly.

(a) Random rigid-body rotation around the domain's center of mass. (b) Random rigid-body translation of the domain's center of mass. (c) Random rigid-body rotation around the axis connecting the domain's N- and C-terminal C_{α} atoms. (d) Rigid-body translation along the axis connecting two domains which are neighboring in sequence. (e) Pose exchange between two domains with similar structures.



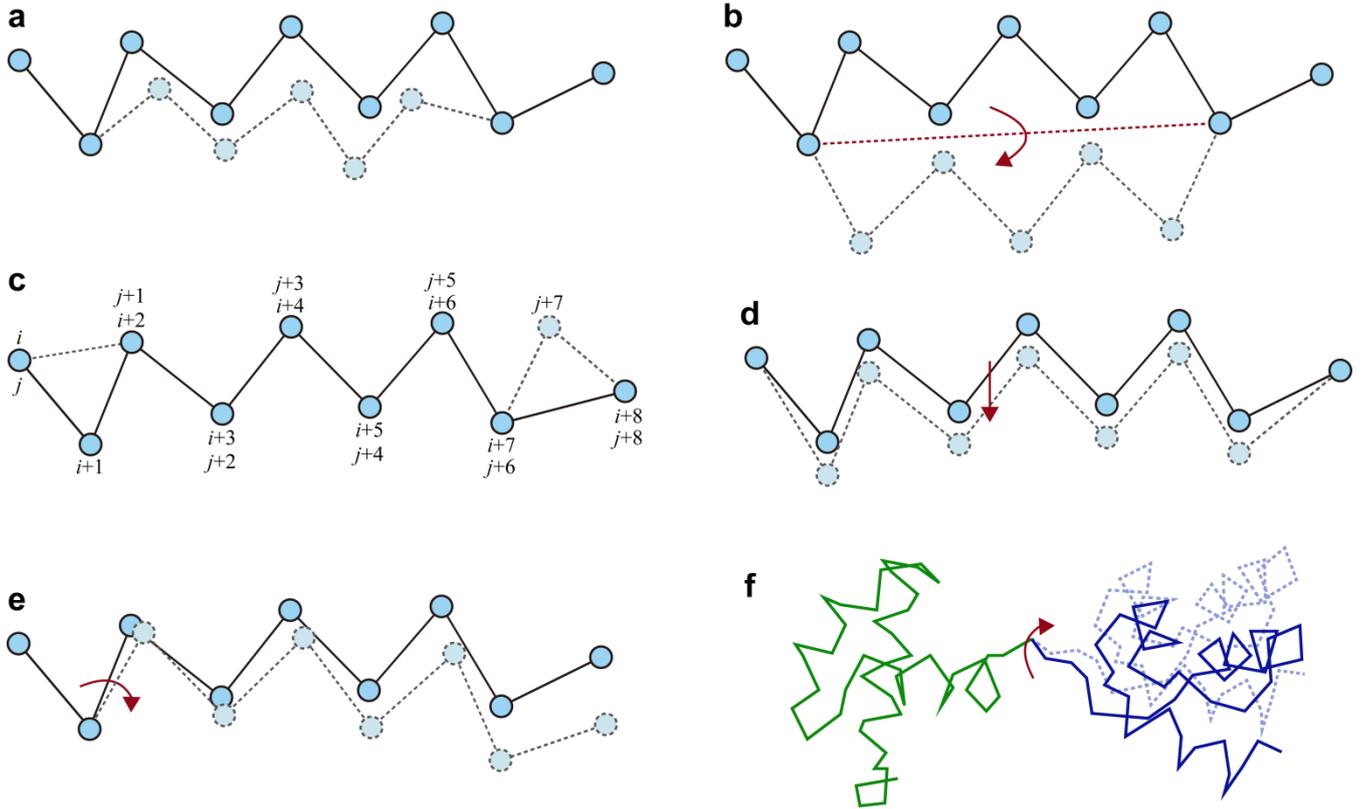
Supplementary Figure 21

Comparison the statistical minimum radius of gyration with the real values.



Supplementary Figure 22

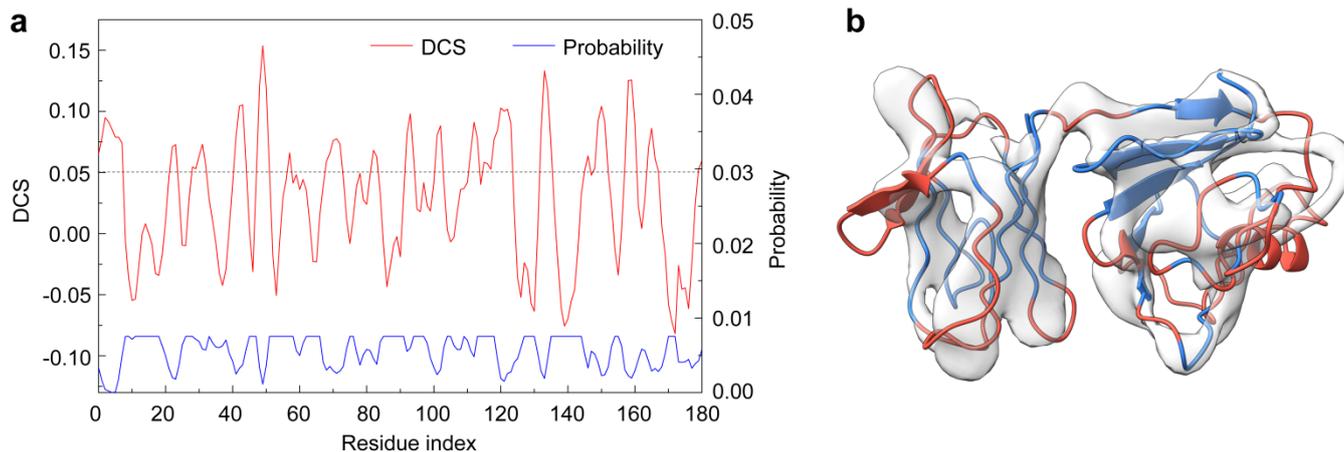
Illustration of domain boundary distance potential for a two-domain protein with discontinuous domains. The discontinuous domain (Domain-I) is split into two segments due to the insertion of the continuous domain (Domain-II). d_1 and d_2 are C_α -distances.



Supplementary Figure 23

Movements for the flexible domain assembly and refinement.

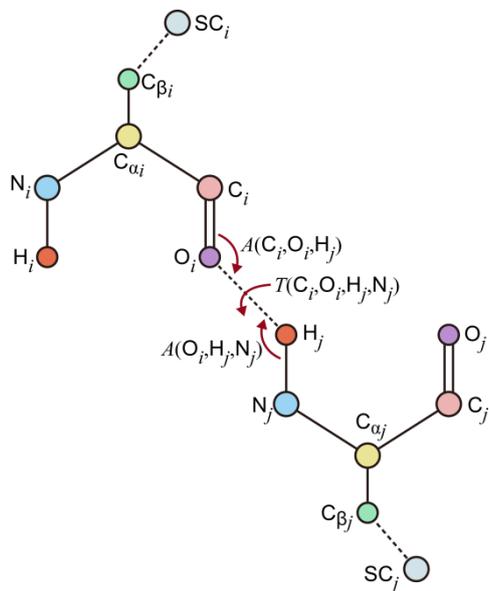
(a) LMProt²⁸ perturbation. (b) Segment rotation around the axis connecting two termini of the segment. (c) Conformational shift of segments along the sequence. (d) Rigid-body segment translation. (e) rigid-body tail rotation. (f) Rigid-body domain-level translation and rotation.



Supplementary Figure 24

Example of regions determination for remodeling and refinement.

(a) The local density correlation score of each residue in an example protein (PDB 1wv3A) and the corresponding probability of each residue to be selected for remodeling. (b) The 3D structure of 1wv3A superposed into the density map, where red regions are residues with local density score < 0.05 .



Supplementary Figure 25

Definition of H-O distance, inner angles, and torsion angle for hydrogen-bond.

Supplementary Tables

Supplementary Table 1. Results of final models constructed from experimental domain structures on 357 test proteins using simulated density maps.

	Method	TM-score	RMSD (Å)
2dom ^a (N=166)	MDFP	0.91(0.14)	4.6(6.8)
	Rosetta	0.86(0.18)	4.8(6.7)
	DEMO-EM	0.99(0.01)	0.5(0.3)
2dis ^b (N=81)	MDFP	0.89(0.17)	4.3(5.7)
	Rosetta	0.85(0.17)	5.4(6.6)
	DEMO-EM	0.99(0.01)	0.6(0.4)
3dom ^c (N=69)	MDFP	0.79(0.20)	10.1(9.0)
	Rosetta	0.72(0.25)	11.0(9.5)
	DEMO-EM	0.99(0.01)	0.7(0.3)
m4dom ^d (N=41)	MDFP	0.67(0.23)	17.9(11.0)
	Rosetta	0.55(0.31)	21.9(14.1)
	DEMO-EM	0.99(0.00)	0.8(0.5)
All (N=357)	MDFP	0.86(0.20)	7.1(9.5)
	Rosetta	0.79(0.23)	8.1(9.9)
	DEMO-EM	0.99(0.00)	0.6(0.3)

The former value is the average, while the value in parentheses is the standard deviation. Bold font highlights the best results from each category.

^a protein with 2 domains.

^b protein with discontinuous domain which contains 2 or more segments from separate regions of the query sequence.

^c protein with 3 domains.

^d protein with 4 or more domains.

Supplementary Table 2. Results of final models constructed from D-I-TASSER predicted domain models using simulated density maps.

	Method	TM-score	RMSD(Å)	TM-score (domain) ^a	RMSD(Å) (domain) ^b
2dom (N=166)	MDFP	0.57(0.23)	15.3(8.1)	0.65(0.19)	5.5(3.4)
	Rosetta	0.52(0.21)	16.9(8.6)	0.52(0.23)	7.9(5.9)
	MAINMAST	0.38(0.28)	16.4(8.2)	0.35(0.19)	13.1(7.0)
	DEMO-EM	0.87 (0.16)	4.8 (5.0)	0.85 (0.16)	3.5 (3.5)
2dis (N=81)	MDFP	0.56(0.21)	15.4(7.6)	0.63(0.18)	6.9(4.7)
	Rosetta	0.49(0.20)	18.6(7.3)	0.47(0.21)	10.1(5.8)
	MAINMAST	0.32(0.25)	20.2(9.0)	0.30(0.24)	15.8(7.6)
	DEMO-EM	0.80 (0.18)	7.9 (5.8)	0.78 (0.19)	5.9 (5.2)
3dom (N=69)	MDFP	0.50(0.20)	17.5(8.2)	0.64(0.19)	5.3(3.2)
	Rosetta	0.36(0.20)	25.6(10.6)	0.40(0.20)	10.7(7.4)
	MAINMAST	0.35(0.25)	18.2(7.7)	0.31(0.24)	12.9(6.6)
	DEMO-EM	0.87 (0.15)	4.7 (4.4)	0.85 (0.15)	2.9 (2.6)
m4dom (N=41)	MDFP	0.39(0.25)	23.2(9.8)	0.57(0.21)	6.2(3.8)
	Rosetta	0.27(0.23)	36.1(10.3)	0.47(0.27)	10.9(9.0)
	MAINMAST	0.31(0.23)	22.2(9.0)	0.25(0.21)	13.1(5.9)
	DEMO-EM	0.83 (0.21)	8.5 (9.2)	0.83 (0.18)	3.2 (3.4)
All (N=357)	MDFP	0.53(0.23)	16.6(8.2)	0.63(0.22)	5.9(3.9)
	Rosetta	0.45(0.21)	21.2(10.9)	0.48(0.26)	9.3(7.1)
	MAINMAST	0.35(0.26)	18.3(8.6)	0.32(0.25)	13.7(6.9)
	DEMO-EM	0.85 (0.18)	5.9 (6.1)	0.83 (0.16)	3.9 (3.8)

The former value is the average, while the value in parentheses is the standard deviation. Bold font highlights the best results from each category.

^a TM-score of individual domain models in full-length models.

^b RMSD of individual domains models in full-length models.

Supplementary Table 3. Validation scores of final models constructed from D-I-TASSER predicted domain models using simulated density maps.

	Method	Ramachandran favored (%)	Rotamer outliers (%)	Clash score	Molprobrity score	EMringer score	iFSC
2dom (N=166)	MDFP	66.98(10.87)	19.01(8.31)	106.30(55.08)	4.34(0.52)	1.20(0.92)	0.49(0.16)
	Rosetta	85.75(5.50)	0.95 (2.07)	15.04(10.33)	2.28(0.45)	1.23(0.84)	0.37(0.18)
	MAINMAST	38.2(14.76)	41.3(16.09)	587.86(572.19)	5.15(0.56)	1.35(0.87)	0.44(0.16)
	DEMO-EM	91.25 (3.69)	1.05(0.93)	2.48 (3.30)	1.47 (0.52)	1.44 (0.71)	0.67 (0.15)
2dis (N=81)	MDFP	65.70(7.38)	19.33(7.93)	105.89(52.08)	4.35(0.52)	1.18(1.05)	0.52(0.14)
	Rosetta	83.95(6.07)	2.65(6.24)	42.86(106.55)	2.53(0.74)	1.26(1.06)	0.39(0.20)
	MAINMAST	38.48(12.44)	45.86(8.80)	474.57(335.75)	5.19(0.56)	0.95(0.73)	0.47(0.14)
	DEMO-EM	90.17 (4.55)	1.67 (1.17)	4.84 (4.81)	1.78 (0.60)	2.02 (1.34)	0.65 (0.18)
3dom (N=69)	MDFP	63.57(10.17)	23.01(8.26)	107.36(53.74)	4.42(0.41)	1.07(0.93)	0.42(0.16)
	Rosetta	84.64(5.33)	0.73 (0.61)	21.39(52.93)	2.32(0.44)	1.19(0.87)	0.27(0.17)
	MAINMAST	35.04(10.07)	39.36(16.96)	781.56(838.44)	5.22(0.47)	1.03(0.62)	0.42(0.17)
	DEMO-EM	91.77 (3.73)	1.04(0.79)	2.89 (2.27)	1.62 (0.51)	1.20 (0.86)	0.69 (0.13)
m4dom (N=41)	MDFP	61.78(8.16)	23.22(7.89)	107.13(41.67)	4.53(0.23)	0.99(0.55)	0.36(0.17)
	Rosetta	83.95(3.43)	0.86 (0.66)	136.97(227.34)	2.62(0.70)	0.84(0.51)	0.20(0.12)
	MAINMAST	45.29(18.87)	35.42(20.98)	841.53(642.82)	5.15(0.60)	0.86(0.53)	0.42(0.19)
	DEMO-EM	90.52 (4.84)	1.09(0.87)	3.99 (5.58)	1.84 (0.64)	1.05 (0.45)	0.65 (0.20)
All (N=357)	MDFP	65.44(10.88)	20.30(8.20)	106.51(52.88)	4.38(0.48)	1.15(1.06)	0.47(0.16)
	Rosetta	84.92(7.07)	1.29(3.38)	36.58(50.69)	2.38(0.58)	1.18(0.91)	0.34(0.19)
	MAINMAST	38.45(14.45)	41.28(16.17)	628.73(611.72)	5.17(0.72)	1.14(0.76)	0.44(0.16)
	DEMO-EM	91.02 (4.07)	1.19 (0.99)	3.27 (3.96)	1.61 (0.57)	1.48 (0.98)	0.67 (0.16)

The former value is the average, while the value in parentheses is the standard deviation. Bold font highlights the best results from each category.

Supplementary Table 4. Refinement results for the full-length models generated by the rigid-body assembly of DEMO-EM using D-I-TASSER predicted domain structures.

Method	TM-score	RMSD(Å)	TM-score (domain)	RMSD(Å) (domain)	Molprobability score	EMringer score	iFSC	Time(h)
Initial Model ^a	0.75(0.19)	8.4(6.8)	0.77(0.14)	4.6(3.6)	2.78(0.69)	1.05(0.68)	0.38(0.17)	NA
MDFP-DEMO-EM ^b	0.81(0.21)	7.5(7.6)	0.79(0.19)	4.3(4.0)	2.18(0.57)	1.32 (1.15)	0.61(0.18)	5.51(5.95)
Rosetta-DEMO-EM ^c	0.79(0.20)	7.4(7.7)	0.76(0.20)	4.5(4.4)	1.96(0.52)	1.26(0.84)	0.63(0.15)	2.36 (3.02)
DEMO-EM	0.85 (0.18)	5.9 (6.1)	0.83 (0.16)	3.9 (3.8)	1.61 (0.57)	1.48 (0.98)	0.67 (0.16)	2.40(3.05)

The former value is the average, while the value in parentheses is the standard deviation. Bold font highlights the best results from each category.

^a Full-length models generated by the rigid-body assembly of DEMO-EM.

^b MDFP refinement for the full-length models by the rigid-body assembly of DEMO-EM.

^c Rosetta refinement for the full-length models by the rigid-body assembly of DEMO-EM.

Supplementary Table 5. The 51 proteins with experimental cryo-EM density maps from EMDB

EMDB code	Resolution(Å)	EMDB code	Resolution(Å)
EMD-20749	2.9	EMD-9389	6.7
EMD-20959	3.3	EMD-5245	6.7
EMD-0782	3.4	EMD-7093	7.2
EMD-21377	3.94	EMD-2183	7.2
EMD-0346	4	EMD-3202	7.3
EMD-10215	4.1	EMD-7461	7.4
EMD-10215	4.1	EMD-0367	7.5
EMD-10100	4.15	EMD-20724	7.5
EMD-6413	4.28	EMD-5784	7.5
EMD-20784	4.3	EMD-1180	7.7
EMD-8002	4.3	EMD-6102	7.7
EMD-2867	4.3	EMD-20114	7.8
EMD-9109	4.5	EMD-6207	7.8
EMD-9163	4.71	EMD-3186	7.9
EMD-3048	4.9	EMD-1999	8
EMD-4441	5.2	EMD-4156	8
EMD-7343	5.8	EMD-0773	8.1
EMD-3906	5.8	EMD-0491	8.5
EMD-3906	5.8	EMD-0293	8.6
EMD-3049	6	EMD-2598	8.75
EMD-2484	6	EMD-4671	9.1
EMD-5396	6.2	EMD-1569	9.1
EMD-2845	6.47	EMD-2597	9.15
EMD-2009	6.6	EMD-0876	9.58
EMD-2009	6.6	EMD-4232	10
EMD-2784	6.6		

Supplementary Table 6. Comparison of the DEMO-EM models with the full-length models predicted by AlphaFold2, and comparisons of different refinement methods starting from the AlphaFold2 predicted models.

Methods	Quality for full-length models		Quality for domain-level models	
	TM-score	RMSD (Å)	TM-score	RMSD (Å)
AlphaFold2	0.84(0.16)	4.4(4.2)	0.89(0.09)	2.0(1.3)
DEMO-EM	0.88(0.09)	4.2(3.2)	0.84(0.13)	3.2(2.8)
MDFP-AlphaFold2 ^a	0.89(0.14)	3.6(4.5)	0.86(0.12)	2.4(1.3)
Rosetta-AlphaFold2 ^b	0.88(0.15)	3.8(4.7)	0.85(0.11)	2.5(1.2)
DEMO-EM-AlphaFold2 ^c	0.93 (0.07)	2.5 (1.7)	0.90 (0.09)	1.8 (1.1)

The former value is the average, while the value in parentheses is the standard deviation. Bold font highlights the best results from each category.

^a MDFP refinement using initial full-length models generated by AlphaFold2.

^b Rosetta refinement using initial full-length models built by AlphaFold2.

^c DEMO-EM refinement using initial full-length models built by AlphaFold2.

Supplementary Table 7. Domain definitions of the 6 proteins in SARS-CoV-2 genome by FUpred²⁴ and ThreaDom²⁵. Different domains are separated by semicolons, and different segments of the discontinuous domain are separated by commas.

EMD code	Resolution(Å)	PDB code	Domain definition
EMD-21375	3.46	6vsbC	1-290;291-320,591-700;321-327,529-590;328-528; 701-717, 1072-1146; 718-1071;
EMD-11007	2.9	6yytB	1-125;126-198;
EMD-22160	3.5	6xezE	1-100;101-235;236-439;440-605;
EMD-22136	2.9	6xdcA	1-141;142-284;
EMD-11007	2.9	6yytC	1-83;
EMD-11007	2.9	6yytA	1-265;266-388;389-407,445-812;408-444,813-932;

Supplementary Table 8. Comparison of validation scores between models generated by DEMO-EM and deposited models for the 6 proteins in SARS-CoV-2 genome.

PDB code	Rama favored (%)	Rotamer outliers (%)	Clashscore	MolProbity score	EMringer score	iFSC	Q-score
6vsbC	95.87 (94.50)	1.41(0.64)	0.07 (12.11)	0.93 (1.97)	2.04(2.35)	0.42(0.45)	0.57 (0.53)
6yytB	95.11(97.83)	0.63(0.00)	0.00 (3.23)	0.84 (1.15)	1.67 (1.39)	0.45(0.49)	0.53 (0.50)
6xezE	89.56 (87.04)	3.07 (9.96)	0.32 (24.48)	1.55 (3.26)	1.08(1.21)	0.34 (0.33)	0.46 (0.41)
6xdcA	96.30 (96.30)	1.15 (1.15)	0.32 (4.15)	0.91 (1.49)	3.60 (3.58)	0.57 (0.57)	0.64 (0.61)
6yytC	100 (100)	0.00 (0.00)	0.00 (3.47)	0.50 (1.14)	1.95 (1.77)	0.55 (0.54)	0.63 (0.60)
6yytA	96.81 (96.69)	0.40 (0.53)	0.30 (2.44)	0.81 (1.24)	3.29 (3.22)	0.44(0.50)	0.69 (0.66)
Average	95.52 (95.39)	1.11 (2.05)	0.17 (8.31)	0.92 (1.71)	2.27 (2.25)	0.46(0.48)	0.59 (0.55)

The numbers in parentheses are for deposited models, and bold font highlights the best results. It should be noted that all scores of DEMO-EM are calculated based on those residues that are identical to deposited models.

Supplementary Table 9. Statistic values to determine coordinates of O, C_β, H, and side-chain center of mass (SC) according to their relative positions to the three backbone atoms (N, C_α, and C), where *D*, *T*, and *A* indicate the distance, torsion angle, and inner angle, respectively.

O_i $i \in [1, L - 1]$	$D(O_i, C_i)$ (Å)	1.229	H_i $i \in [2, L]$	$D(N_i, H_i)$ (Å)	0.987
	$T(C_{\alpha_i}, C_i, O_i, N_{i+1})$ (°)	179.672		$T(C_{i-1}, N_i, C_{\alpha_i}, H_i)$ (°)	179.817
	$A(C_{\alpha_i}, C_i, O_i)$ (°)	120.098		$A(H_i, N_i, C_{\alpha_i})$ (°)	119.255
O_i $i = L$	$D(O_i, C_i)$ (Å)	1.244	H_i $i = 1$	$D(N_i, H_i)$ (Å)	0.987
	$T(C_{\alpha_i}, C_i, O_i, N_i)$ (°)	0		$T(C_i, N_i, C_{\alpha_i}, H_i)$ (°)	60
	$A(C_{\alpha_i}, C_i, O_i)$ (°)	119.494		$A(H_i, N_i, C_{\alpha_i})$ (°)	116.345
C _β	Residue type specific: https://zhanggroup.org/DEMO-EM/potential/CB_position.txt				
SC	Residue type specific: https://zhanggroup.org/DEMO-EM/potential/SG_position.txt				

Supplementary Table 10. Van der Waals radius parameters from CHARMM22²⁹ are used to count for excluded volume interaction.

	C _α	N	C	O	C _β	SC
C _α	3.6	2.3	3.7	2.9	3.5	1.0
N	3.7	2.5	3.5	2.5	3.5	1.0
C	2.3	1.2	2.7	2.7	2.3	1.0
O	2.5	2.1	2.4	2.3	2.6	1.0
C _β	3.5	2.3	3.5	2.8	3.3	1.0
SC	1.0	1.0	1.0	1.0	1.0	1.0

Supplementary Table 11. Mean and standard deviation of four hydrogen-bond features in α -helix and β -sheet structures.

	$D(O_i, H_j)$ (Å)	$A(C_i, O_i, H_j)$ (°)	$A(O_i, H_j, N_j)$ (°)	$T(C_i, O_i, H_j, N_j)$ (°)
Helix, $j=i+4$	2.00/0.53	147/10.58	159/11.25	160/25.36
Helix, $j=i+3$	2.85/0.32	89/7.70	111/8.98	-160/7.93
Parallel	2.00/0.30	155/11.77	164/11.29	180/68.96
Antiparallel	2.00/0.26	151/12.38	163/11.02	-168/69.17

Supplementary References

1. Xu, Y., Xu, D. & Gabow, H.N. Protein domain decomposition using a graph-theoretic approach. *Bioinformatics* **16**, 1091-1104 (2000).
2. Chandonia, J.-M., Fox, N.K. & Brenner, S.E. SCOPe: manual curation and artifact removal in the structural classification of proteins—extended database. *Journal of molecular biology* **429**, 348-355 (2017).
3. Lam, S.D. *et al.* Gene3D: expanding the utility of domain assignments. *Nucleic acids research* **44**, D404-D409 (2015).
4. Tang, G. *et al.* EMAN2: an extensible image processing suite for electron microscopy. *Journal of structural biology* **157**, 38-46 (2007).
5. Zheng, W. *et al.* Protein structure prediction using deep learning distance and hydrogen-bonding restraints in CASP14. *Proteins: Structure, Function, Bioinformatics* **89**, 1734-1751 (2021).
6. Zhang, Y. & Skolnick, J. Scoring function for automated assessment of protein structure template quality. *Proteins: Structure, Function, and Bioinformatics* **57**, 702-710 (2004).
7. Ludtke, S.J., Baldwin, P.R. & Chiu, W. EMAN: semiautomated software for high-resolution single-particle reconstructions. *Journal of structural biology* **128**, 82-97 (1999).
8. Xu, J. & Zhang, Y. How significant is a protein structure similarity with TM-score = 0.5? *Bioinformatics* **26**, 889-95 (2010).
9. Topf, M. *et al.* Protein structure fitting and refinement guided by cryo-EM density. *Structure* **16**, 295-307 (2008).
10. Singharoy, A. *et al.* Molecular dynamics-based refinement and validation for sub-5 Å cryo-electron microscopy maps. *Elife* **5**, e16105 (2016).
11. Chacón, P. & Wriggers, W. Multi-resolution contour-based fitting of macromolecular structures. *Journal of molecular biology* **317**, 375-384 (2002).
12. Trabuco, L.G., Villa, E., Schreiner, E., Harrison, C.B. & Schulten, K. Molecular dynamics flexible fitting: a practical guide to combine cryo-electron microscopy and X-ray crystallography. *Methods* **49**, 174-180 (2009).
13. Pintilie, G. & Chiu, W. Validation, analysis and annotation of cryo-EM structures. *Acta Crystallographica Section D: Structural Biology* **77**, 1142-1152 (2021).
14. Zhou, X.G., Hu, J., Zhang, C.X., Zhang, G.J. & Zhang, Y. Assembling multidomain protein structures through analogous global structural alignments. *Proceedings of the National Academy of Sciences* **116**, 15930-15938 (2019).
15. Huang, X., Pearce, R. & Zhang, Y. FASPR: an open-source tool for fast and accurate protein side-chain packing. *Bioinformatics* **36**, 3758-3765 (2020).
16. Xu, D. & Zhang, Y. Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins* **80**, 1715-1735 (2012).
17. Ramachandran, G.T. & Sasisekharan, V. Conformation of polypeptides and proteins. in *Advances in protein chemistry*, Vol. 23 283-437 (Elsevier, 1968).
18. Rotkiewicz, P. & Skolnick, J. Fast procedure for reconstruction of full-atom protein models from reduced representations. *Journal of computational chemistry* **29**, 1460-1465 (2008).
19. Wang, R.Y.-R. *et al.* Automated structure refinement of macromolecular assemblies from cryo-EM maps using Rosetta. *Elife* **5**, e17219 (2016).
20. Terashi, G. & Kihara, D. De novo main-chain modeling for EM maps using MAINMAST. *Nature communications* **9**, 1-11 (2018).
21. Terashi, G., Kagaya, Y. & Kihara, D. MAINMASTseg: Automated Map Segmentation Method for Cryo-EM Density Maps with Symmetry. *Journal of chemical information modeling* **60**, 2634-2643 (2020).
22. Zhang, B., Zhang, X., Pearce, R., Shen, H.-B. & Zhang, Y. A new protocol for atomic-level protein structure modeling and refinement using low-to-medium resolution cryo-EM density maps. *Journal of molecular biology* **432**, 5365-5377 (2020).
23. Wang, R.Y.-R. *et al.* De novo protein structure determination from near-atomic-resolution cryo-EM maps. *Nature*

- methods* **12**, 335-338 (2015).
24. Zheng, W. *et al.* FUPred: Detecting protein domains through deep-learning based contact map prediction. *Bioinformatics* **36**, 3749-3757 (2020).
 25. Wang, Y. *et al.* ThreaDomEx: a unified platform for predicting continuous and discontinuous protein domains by multiple-threading and segment assembly. *Nucleic acids research* **45**, W400-W407 (2017).
 26. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).
 27. Ramírez-Aportela, E. *et al.* FSC-Q: A CryoEM map-to-atomic model quality validation based on the local Fourier Shell Correlation. *Nature Communications* **12**, 1-7 (2021).
 28. da Silva, R.A., Degève, L. & Caliri, A. LMProt: an efficient algorithm for Monte Carlo sampling of protein conformational space. *Biophysical journal* **87**, 1567-1577 (2004).
 29. MacKerell Jr, A.D. *et al.* All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The journal of physical chemistry B* **102**, 3586-3616 (1998).