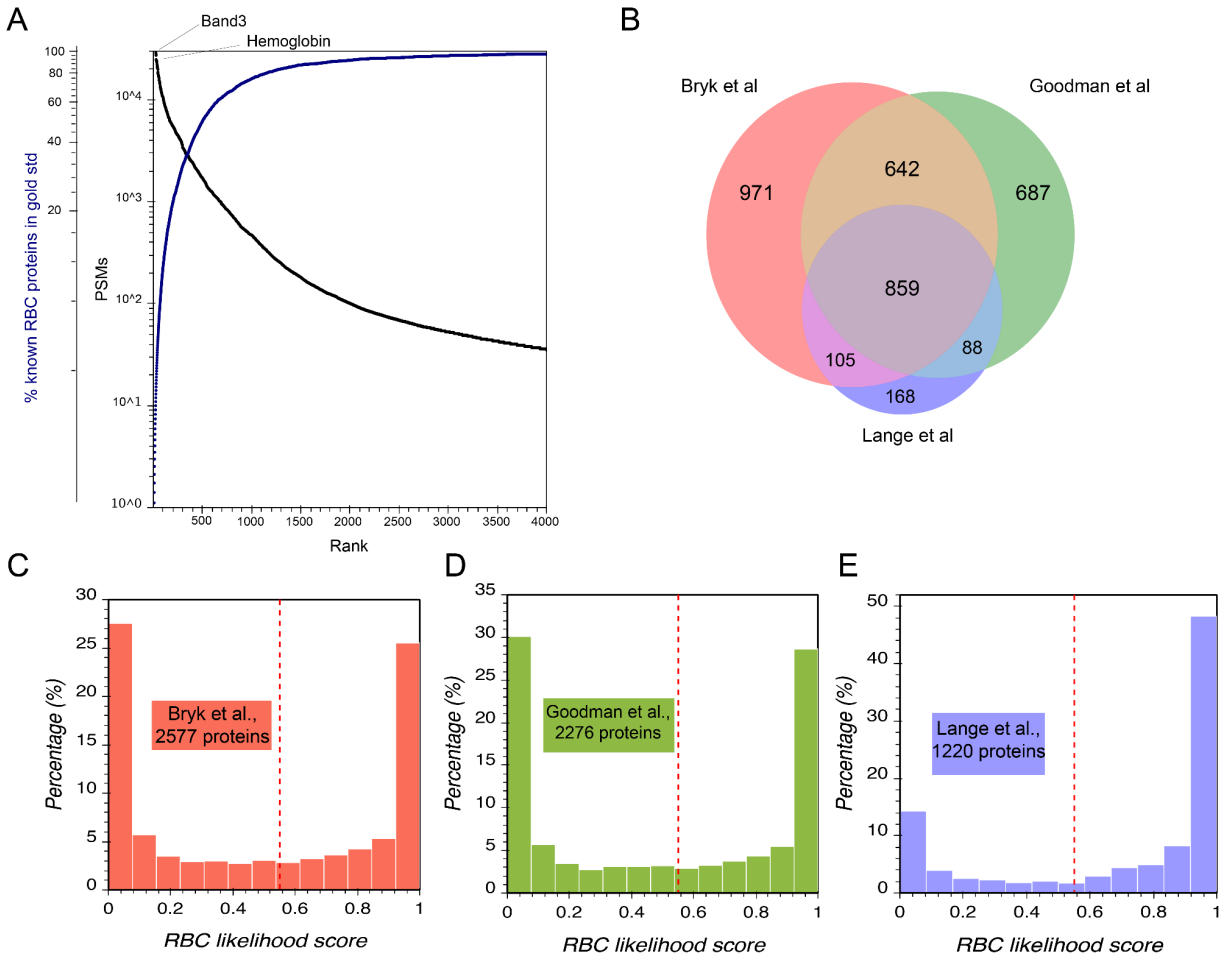
A
Band3
Hemoglobin
% known RBC proteins in gold std
PSMs
Rank
B
Bryk et al
Goodman et al
Lange et al
971
642
687
859
105
88
168
C
Bryk et al.,
2577 proteins
Percentage (%)
RBC likelihood score
D
Goodman et al.,
2276 proteins
Percentage (%)
RBC likelihood score
E
Lange et al.,
1220 proteins
Percentage (%)
RBC likelihood score

**Figure S1. Coverage and Validation of RBC proteome**.

**(A)** RBC proteins were ranked based on abundance (black plot). While hemoglobin is the most abundant protein in RBCs, band 3 is the most abundant in our experiment because of hemoglobin depletion steps in our protocol. Blue plot shows >90% of the RBC proteins in the gold standard set of 859 proteins were detected within the most 2,000 abundant proteins in our experiment, showing a good coverage of RBC proteome.

**(B-E)** A Venn diagram shows the overlap and differences of proteins detected in the three prior proteomic studies (Bryk and Wiśniewski, 2017; Goodman et al., 2007; Lange et al., 2014). Proteins in the intersection were considered well-supported RBC training examples; human proteins not observed by all 3 prior studies were considered non-RBC proteins. Proteins observed by only 1 or 2 prior studies were considered candidate RBC proteins to be scored by the classifier. The remaining panels plot the percentage of proteins in each of the 3 prior studies as a function of the confidence scores assigned by the classifier.  Vertical red dashed lines indicate the 1% FDR threshold.

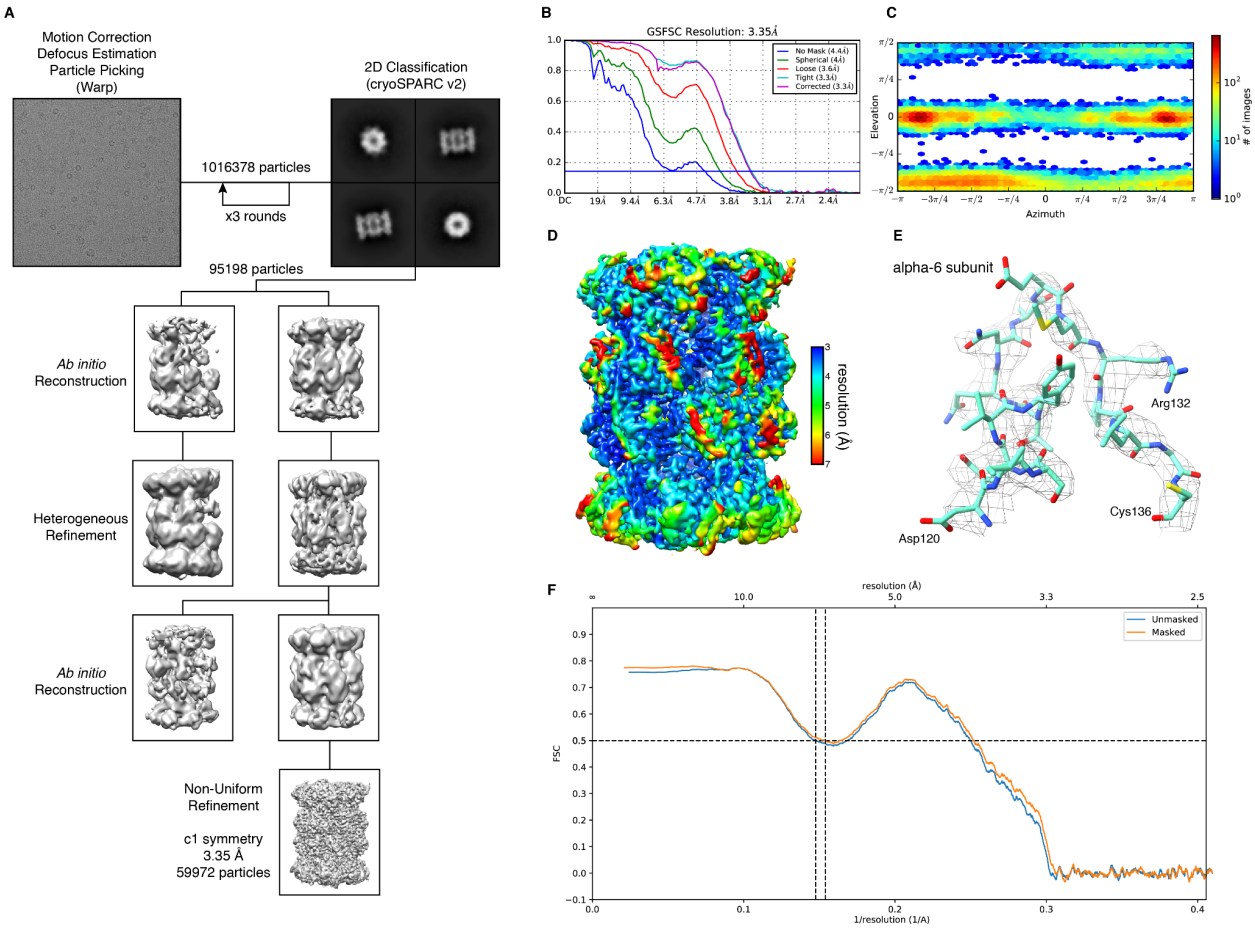


**Figure S2. Cryo-EM data processing workflow and structure validation for 20S proteasome**

- **(A)** Cryo-EM data processing workflow for 20S proteasome.
- **(B)** Fourier Shell Correlation (FSC) curves for the 20S proteasome based on the gold-standard between two independent half maps.
- **(C)** Euler angle distribution plot for the 20S proteasome.
- **(D)** Local resolution map of the 20S proteasome reconstruction.
- **(E)** Region from the alpha-6 subunit of the 20S proteasome reconstruction and model showing a more highly resolved portion of the map.
- **(F)** Map-to-model FSC.

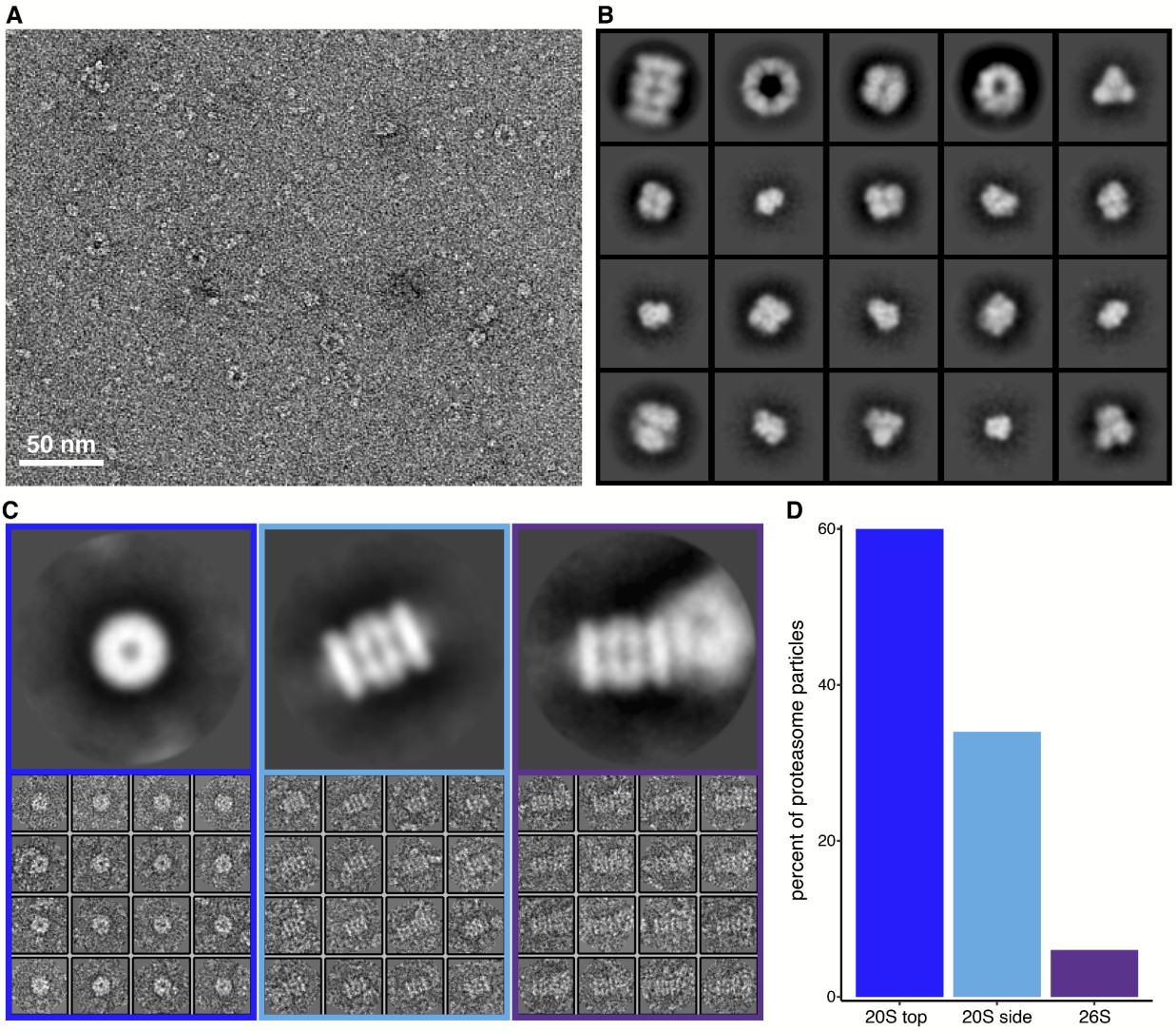
A
50 nm
B
C
D
60
40
20
0
percent of proteasome particles
20S top
20S side
26S

**Figure S3. Assessment of proteasomes from negative stain electron microscopy of RBC hemolysate shows a majority in the 20S form.**

**(A)** Representative micrograph of hemolysate after being passed through a 100 kDa filter.

**(B)** Representative reference-free 2D class averages from filtered hemolysate showing macromolecular assemblies of distinct sizes and shapes. Box length is 254 Å.

(**C**) Reference-free 2D class averages and aligned raw particles for 20S proteasome (top view), 20S proteasome (side view) and 26S proteasome (single-capped) from left to right. Box length corresponds to 459 Å.

(**D**) Distribution of observed proteasome states from negative stain EM of hemolysate. The total number of proteasome particles classified was 1,510.

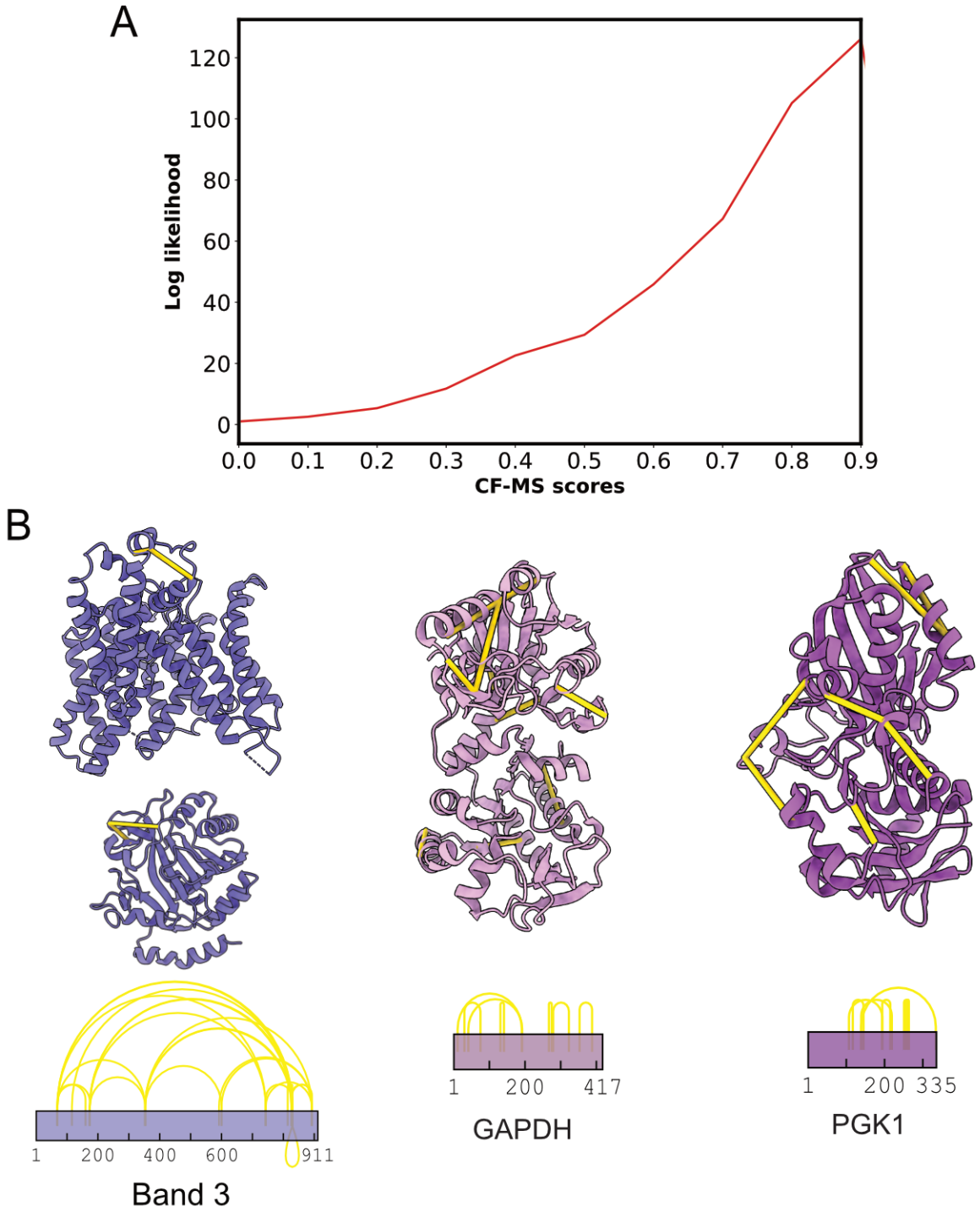
A
Log likelihood
120
100
80
60
40
20
0
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
CF-MS scores
B
1
200
400
600
911
Band 3
1
200
417
GAPDH
1
200
335
PGK1

**Figure S4. 3D models of individual membrane/cytoskeletal proteins prior to integration into larger complexes.**

(A) Global validation of crosslinks shows that crosslinked pairs are likely to have high CF-MS scores, suggesting that our crosslink experiment captures true physical interactions.

(B) Yellow lines in the protein structures represent intramolecular crosslinks detected in our crosslinking experiments. The yellow lines above rectangular boxes show the amino acid residues with detected crosslinks. Intramolecular crosslinks of individual proteins agreed with the existing x-ray crystal structures of these proteins.

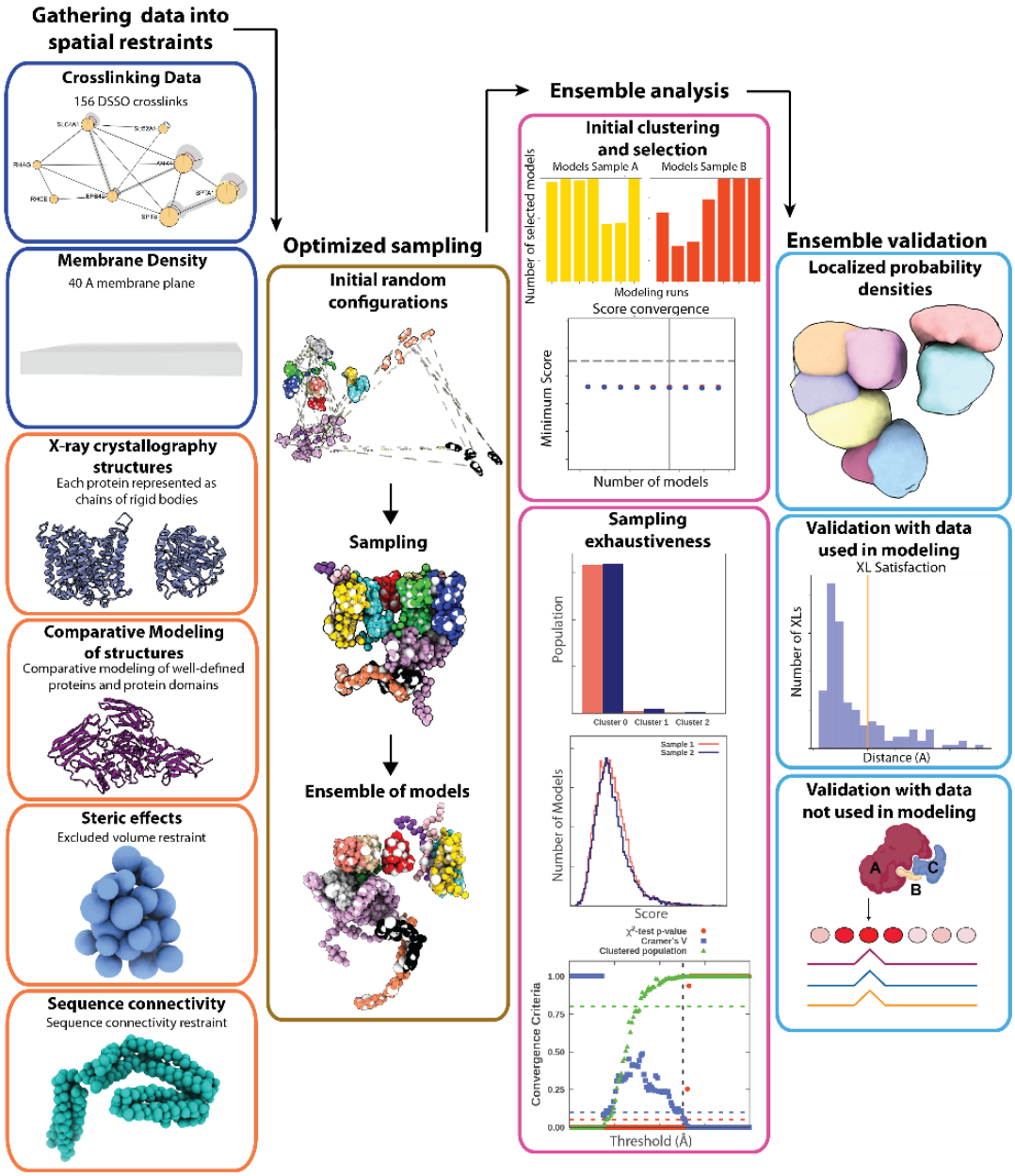
Gathering data into spatial restraints
Crosslinking Data
156 DSSO crosslinks
Membrane Density
40 A membrane plane
X-ray crystallography structures
Each protein represented as chains of rigid bodies
Comparative Modeling of structures
Comparative modeling of well-defined proteins and protein domains
Steric effects
Excluded volume restraint
Sequence connectivity
Sequence connectivity restraint
Optimized sampling
Initial random configurations
Sampling
Ensemble of models
Ensemble analysis
Initial clustering and selection
Models Sample A
Models Sample B
Number of selected models
Modeling runs
Score convergence
Minimum Score
Number of models
Sampling exhaustiveness
Population
Cluster 0
Cluster 1
Cluster 2
Sample 1
Sample 2
Number of Models
Score
χ²-test p-value
Cramer's V
Clustered population
Convergence Criteria
Threshold (Å)
Ensemble validation
Localized probability densities
Validation with data used in modeling
XL Satisfaction
Number of XLs
Distance (A)
Validation with data not used in modeling
A
B
C

**Figure S5. Integrative Modeling Platform (IMP) workflow.**

Data in the blue box are converted into spatial restraints. The orange boxes show x-ray crystal structures, statistical inferences, and physical properties. The gold box displays the sampling and scoring of the models.The models are analyzed via an initial clustering to select a high scoring group of models from various simulation runs followed by sampling exhaustiveness (pink boxes). The light blue boxes show the ensemble validation against the data used in the modeling, the probability density for each subunit, and a comparison against the known structure.

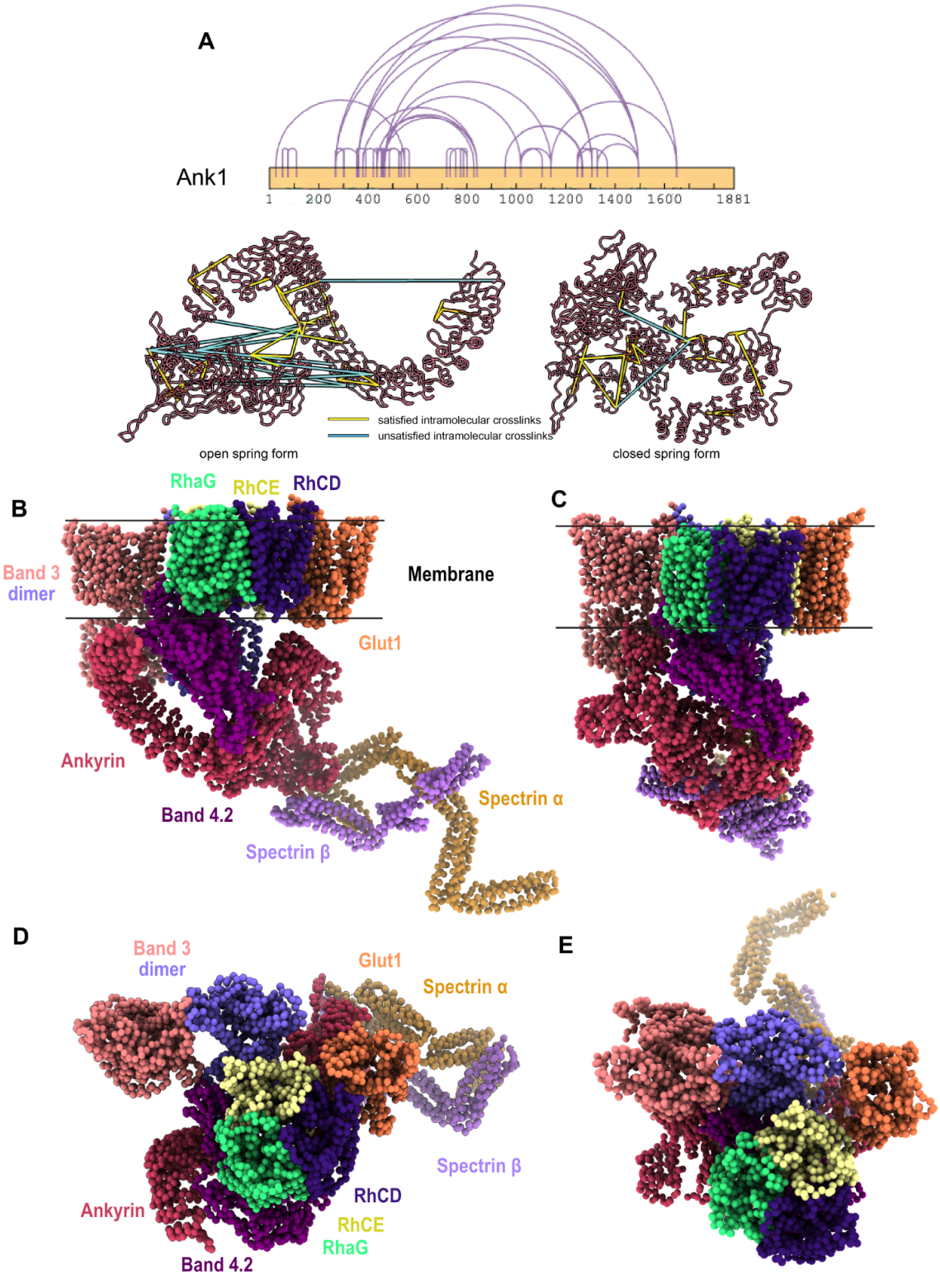
A
Ank1
1
200
400
600
800
1000
1200
1400
1600
1881
satisfied intramolecular crosslinks
unsatisfied intramolecular crosslinks
open spring form
closed spring form
B
RhaG
RhCE
RhCD
Band 3
dimer
Membrane
Glut1
Ankyrin
Band 4.2
Spectrin β
Spectrin α
C
D
Band 3
dimer
Glut1
Spectrin α
Spectrin β
RhCD
Ankyrin
RhCE
RhaG
Band 4.2
E

**Figure S6. Correlated cross-linker violations suggest ankyrin may adopt a highly compressed structure within the band 3-Ank1 complex.**

**(A)** 30% of crosslinks exceed the inter-amino acid distances expected for the DSSO crosslinker (30 Å) when mapped onto the ANK1 structure in its open conformation (as observed crystallographically by (Wang et al., 2014), but these crosslinks can be satisfied by a model of ANK1 in a closed conformation, suggesting that ANK1 in the band 3-Ank1 complex may adopt both open and closed conformations.

**(B&C)** Integrative 3D model of band 3-Ank1 complex with Ank1 in open (left) and closed (right) conformations. Side views of the band 3-Ank1 complex.

**(D&E)** Top views of the band 3-Ank1 complex with Ank1 in either open (left) or closed (right) conformation.

**Table S1. Cryo-EM data collection and reconstruction statistics.**

| *EM data collection and reconstruction statistics* | |
|---|---|
| Protein | 20S Proteasome |
| EMDB | EMD-24822 |
| Microscope | FEI Titan Krios |
| Voltage (kV) | 300 |
| Detector | Gatan K3 |
| Magnification (nominal) | 22,500 |
| Pixel size (Å/pix) | 1.045 |
| Flux ($e^-$/pix/sec) | 15.5 |
| Frames per exposure | 20 |
| Exposure ($e^-$/Å$^2$) | 42.58 |
| Defocus range ($\mu$m) | 1.09 - 2.5 |
| Micrographs collected | 6,606 |
| Particles extracted/final | 1016378/59972 |
| Symmetry imposed | none (C1) |
| Map sharpening B-factor | 122.3 |
| Unmasked resolution at 0.143 FSC (Å) | 4.4 |
| Masked resolution at 0.143 FSC (Å) | 3.35 |

**Table S2: Statistical validation for integrative modeling of protein complexes.**

| Protein complex | Cluster precision | Cluster size | % XLs satisfied | CCC |
|---|---|---|---|---|
| Ank1-open | 41.739 Å | 6,672 | 87% | 0.9785 |
| Ank1-closed | 43.884 Å | 4,452 | 95% | 0.9718 |
| Ank1-open w/ enzymes | 52.747 Å | 4,361 | 90% | 0.9921 |
| Ank1-closed w/ enzymes | 52.849 Å | 2,506 | 98% | 0.9442 |
| band 4.1-spectrin | 25.937 Å | 4,237 | 95% | 0.9866 |

**Table S3: Individual protein modeling source and representation.**

| Protein (gene) name | Uniprot ID | C-Score or PDB ID | Representation |
|---|---|---|---|
| ANK1 | P16157 | N/A | Chain of rigid bodies & flexible beads |
| EPB41 | P11171 | -0.54 | Rigid body |
| EPB42 | P16452 | 1.88 | Rigid body |
| GYPA | P02724 | 1AFO(81-120) | Rigid body & flexible beads |
| RhAG | Q02094 | 1.07 | Rigid body |
| RhCE | P18577 | 1.55 | Rigid body |
| SLC2A1 | P11166 | 1.45 | Rigid body |
| SLC4A1 | P02730 | 4YZF(381-887),1HYN(56-355) | Chain of rigid bodies & flexible beads |
| SPTA1 | P02549 | N/A | Chain of rigid bodies |
| SPTB | P11277 | N/A | Chain of rigid bodies |
| GAPDH | P04406 | 1u8f | Rigid body |
| PGK1 | P00558 | 4o33 | Rigid body |