Modeling Flexible Protein Structure With AlphaFold2 and Crosslinking Mass Spectrometry

Authors

Karen Manalastas-Cantos, Kish R. Adoni, Matthias Pfeifer, Birgit Märtens, Kay Grünewald, Konstantinos Thalassinos, and Maya Topf

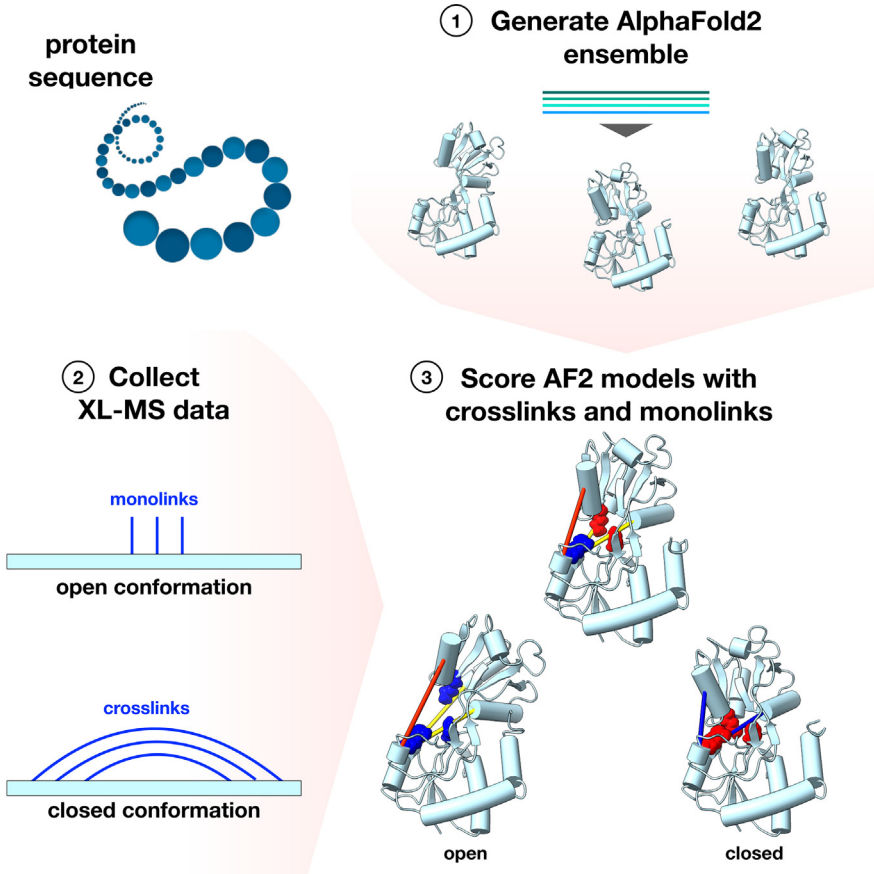
Correspondence

maya.topf@cssb-hamburg.de

In Brief

Many proteins can be described by more than one structure, each of which has a functional significance. Crosslinks and monolinks from crosslinking mass spectrometry (XL-MS) can help probe a protein's conformation under specific physiological conditions. We present a pipeline to predict a protein's specific conformation. We employ a combination of ensemble generation using AlphaFold2 and ensemble scoring using condition-specific XL-MS data. Both crosslinks and monolinks were found to be useful in finding the correct protein conformation.

Graphical Abstract



Highlights

- Multiple conformations of proteins can be predicted with AlphaFold2.
- Crosslinks and monolinks can be used to select the relevant protein conformation.
- Tools for data analysis are available at https://gitlab.com/topf-lab/xlms-tools.

2024, Mol Cell Proteomics *23*(*3*), 100724

https://doi.org/10.1016/j.mcpro.2024.100724

Modeling Flexible Protein Structure With AlphaFold2 and Crosslinking Mass Spectrometry

Karen Manalastas-Cantos[1,2], Kish R. Adoni[3,4], Matthias Pfeifer[2,5], Birgit Märtens[2,5], Kay Grünewald[2,6], Konstantinos Thalassinos[3,4], and Maya Topf[2,5,*]

We propose a pipeline that combines AlphaFold2 (AF2) and crosslinking mass spectrometry (XL-MS) to model the structure of proteins with multiple conformations. The pipeline consists of two main steps: ensemble generation using AF2 and conformer selection using XL-MS data. For conformer selection, we developed two scores—the monolink probability score (MP) and the crosslink probability score (XLP)—both of which are based on residue depth from the protein surface. We benchmarked MP and XLP on a large dataset of decoy protein structures and showed that our scores outperform previously developed scores. We then tested our methodology on three proteins having an open and closed conformation in the Protein Data Bank: Complement component 3 (C3), luciferase, and glutamine-binding periplasmic protein, first generating ensembles using AF2, which were then screened for the open and closed conformations using experimental XL-MS data. In five out of six cases, the most accurate model within the AF2 ensembles—or a conformation within 1 Å of this model—was identified using crosslinks, as assessed through the XLP score. In the remaining case, only the monolinks (assessed through the MP score) successfully identified the open conformation of glutamine-binding periplasmic protein, and these results were further improved by including the "occupancy" of the monolinks. This serves as a compelling proof-of-concept for the effectiveness of monolinks. In contrast, the AF2 assessment score was only able to identify the most accurate conformation in two out of six cases. Our results highlight the complementarity of AF2 with experimental methods like XL-MS, with the MP and XLP scores providing reliable metrics to assess the quality of the predicted models. The MP and XLP scoring functions mentioned above are available at https://gitlab.com/topf-lab/xlms-tools.

AlphaFold2 (AF2) has revolutionized structural biology with unprecedented accuracy in protein structure prediction, even on sequences for which related structures are unavailable (1). AF2 has been used to exhaustively predict the structures of more than 200 million UniProt sequences (2), which is both an invaluable resource to the structural biology community, as well as a test of AF2's current structure prediction capacity. In particular, it has been shown that AF2 predicts ordered protein domain structures well but performs less well on proteins with predicted flexibility or disorder (3, 4). A possible cause is that AF2 and other protein structure prediction approaches have been trained with the assumption that one protein sequence corresponds to one structure, something we know to be untrue for a wide variety of proteins, such as molecular switches which transition between two different conformations as part of their function (4, 5), as well as proteins that are either fully or partially disordered. The issue is compounded by the fact that the Protein Data Bank (PDB) (6)—the protein structure database used to train structure prediction models—is biased toward proteins that are relatively ordered, since disordered regions are usually recalcitrant to X-ray crystallography, and are thus poorly represented (7). In addition, proteins that have multiple possible states resulting from domain rearrangements (*e.g.*, upon binding to other proteins) may not have all of their conformations documented in the PDB.

To its credit, AF2 outputs error estimates for its structure predictions and usually assigns low confidence to intrinsically disordered regions of a protein. In the case of molecular switches that have relatively rigid domains that undergo interdomain movements, AF2's predicted aligned error is sometimes predictive of boundaries between domains, but not always (Fig. 1*A*).

Thus, experimental methods that can capture protein dynamics remain critical in a post-AF2 world. For example, in crosslinking mass spectrometry (XL-MS), proteins are incubated at near-physiological conditions with chemical

From the [1]Center for Data and Computing in Natural Sciences, Universität Hamburg, Hamburg, Germany; [2]Department of Integrative Virology, Leibniz-Institut für Virologie (LIV), Centre for Structural Systems Biology (CSSB), Hamburg, Germany; [3]Institute of Structural and Molecular Biology, Division of Biosciences, University College London, London, UK; [4]Institute of Structural and Molecular Biology, Birkbeck College, University of London, London, United Kingdom; [5]Universitätsklinikum Hamburg Eppendorf (UKE), Hamburg, Germany; [6]Department of Chemistry, Universität Hamburg, Hamburg, Germany

*For correspondence: Maya Topf, maya.topf@cssb-hamburg.de.

https://doi.org/10.1016/j.mcpro.2024.100724

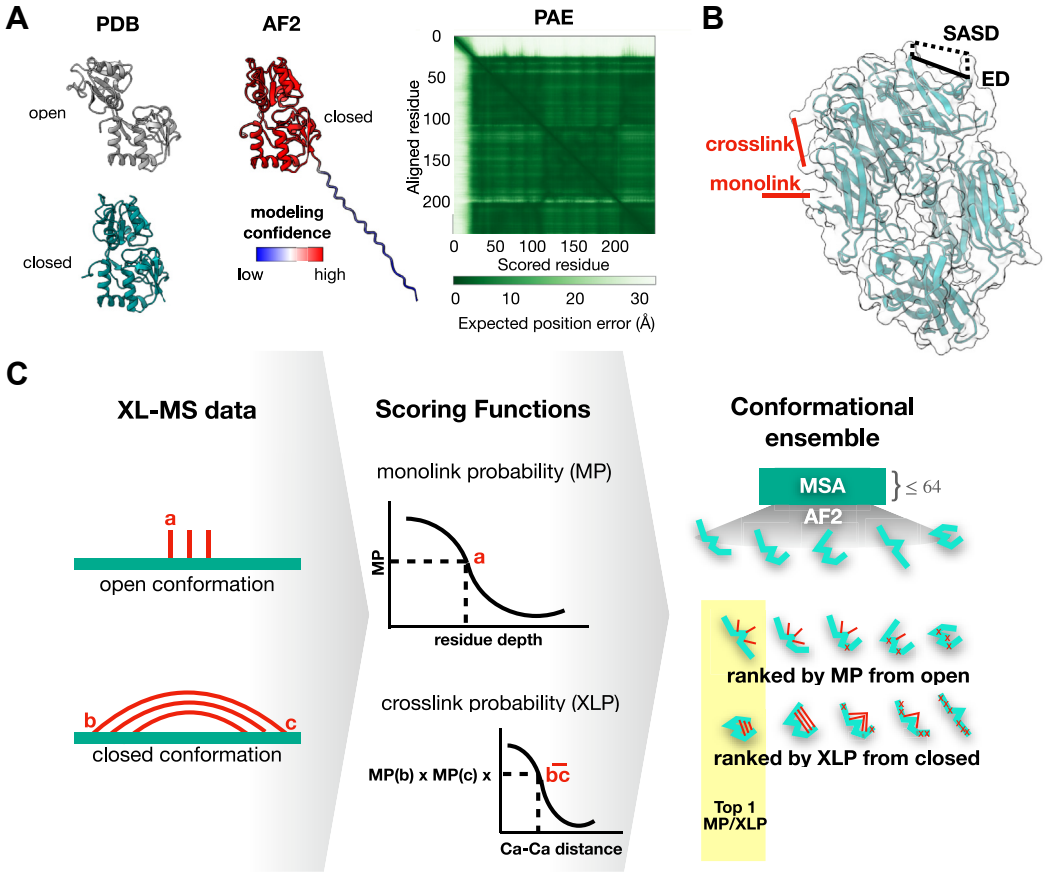


FIG. 1. **AlphaFold2 and XL-MS can be combined to model conformations of flexible proteins.** *A*, the glutamine-binding periplasmic protein (QBP) is known to have an open and closed conformation in the PDB, but AlphaFold2 (AF2) only predicts the closed conformation by default, with high confidence. Blocks of high predicted aligned error (PAE) typically indicate boundaries between domains, but QBP was predicted as a single, low PAE block, with the higher PAE region (*white on plot*) caused by the unfolded N terminus. *B*, crosslinks occur when both ends of a chemical crosslinker (*red*) bond covalently with the protein, while monolinks involve only one attachment site. The Euclidean distance (ED) is the shortest path between the crosslinked Cα atoms, while the solvent-accessible surface distance (SASD) is the shortest path between the crosslinked Cα atoms along the surface of the protein. *C*, our approach for flexible protein modeling involves acquiring XL-MS data in the form of monolinks (a) and crosslinks (bc) from the open and closed conformations, while predicting an ensemble of models using AlphaFold2 with shallow multiple sequence alignments (≤64 sequences). The AF2 models are scored by how well they match XL-MS data from each conformation, using the monolink (MP) and crosslink probability (XLP) scores described in Experimental Procedures. XL-MS, crosslinking mass spectrometry.

crosslinkers such as the most commonly used reagents disuccinimidyl suberate (DSS) and bissulfosuccinimidyl suberate (BS3), which both bind with highest affinity to lysine residues on the protein surface (8). The linker may attach at only one end, giving rise to a monolink, or both ends of the same linker may attach to different parts of the protein, giving rise to a crosslink (Fig. 1*B*). Analysis of the linker attachment sites by mass spectrometry (MS) gives information about what parts of the protein are on the surface. Additionally, crosslinks can constrain the maximum distance between the two sites on the protein where they are attached, since the linker has a defined length (11.4 Å for DSS and BS3) (9). The maximum Cα–Cα distance spanned by a particular crosslinker has thus been described as the sum of the crosslinker length and the side chain lengths of the two linked residues (10), along with a tolerance of 3 Å to account for side chain flexibility and atomic model resolution, resulting in a maximum distance of around 30 Å (11). Crosslinks have been used to evaluate protein structure models by either flagging maximum distance violations, where the maximum Cα–Cα Euclidean distance (ED) that can be spanned by the linker is exceeded (12–15), or by more sophisticated scoring functions that are based on the solvent-accessible surface distance (SASD) (Fig. 1*B*). The SASD, while more time-consuming to compute, is a better approximation of physical conditions and was shown to perform better than ED in distinguishing near-native protein structures from a group of decoys (16–20).

The primary tradeoff of methods like XL-MS that can be performed in solution is that they typically do not give atomic resolution information. For this reason, a common

experimental pipeline when studying flexible proteins is to combine both high-resolution structural information from X-ray crystallography, for example, with measurements in solution such as small-angle scattering to quantify flexibility and XL-MS or FRET to give distance constraints that might not be present in the crystal conformation. In the absence of experimentally determined structures, homology-based models of proteins have also been used in combination with distance constraints measured in solution in order to model the structure of flexible proteins. However, with the advent of deep learning-based protein structure prediction methods such as AF2, this pipeline could be performed on proteins that have no sequence homology to structures in the PDB or proteins for which only some of the conformations are known.

In this paper, we present how XL-MS data can be used to supplement structure predictions from AF2, particularly on flexible proteins (Fig. 1*C*). It has previously been shown that AF2 generates a wider variety of conformations when using a shallow multiple-sequence alignment (MSA) than when not limiting alignment depth (21). We leverage this observation to generate an ensemble of conformations for a set of flexible proteins that have an "open" and a "closed" conformation deposited in the PDB. We screened these AF2 predictions using experimental XL-MS data, with two functions we have developed to score crosslink and monolink information, that make use of the probabilities of spanning distance between crosslinked residues and the depth of the monolinked/crosslinked residue from the surface of the protein. We show that our scores are able to select for models that are near the open and closed conformations from among the AF2 ensembles, provided that AF2 was able to predict these conformations. Furthermore, we present a case in which monolinks are more informative than crosslinks in selecting the experimentally relevant conformer from an ensemble. Thus, we propose a general pipeline to model flexible proteins by first generating an ensemble with AF2, then identifying the conformer present in a given experimental condition using crosslinks and monolinks from XL-MS.

EXPERIMENTAL PROCEDURES

Defining Scoring Functions Using XL-MS Data

We defined two new scoring functions to distinguish near-native structures from a set of decoy protein structures, based on distributions of crosslinks and monolinks obtained from XL-MS experiments. Experimentally determined BS3 and DSS crosslinks and monolinks that were mapped to PDB structures were first obtained from the XlinkAnalyzer database (https://www.embl-hamburg.de/XlinkAnalyzer/database.html) (22). A total of 831 unique lysine residues across 43 PDB structures were found to be tagged with either BS3 or DSS. The depths of both tagged and untagged lysines were computed in the respective PDB structures using the program EDTSurf (https://zhanggroup.org/EDTSurf/), which is a fast algorithm for generating protein surfaces, as well as calculating depths to this surface, using Euclidean distance transform (23, 24). The following parameters were used to run EDTSurf: probe radius = 2.5, scale = 0.5.

The residue depths computed by EDTSurf were compared to those obtained from DEPTH (http://cospi.iiserpune.ac.in/depth/), which is a more accurate but slower method that places the protein structure in a random orientation in a box of explicit solvent for several iterations, while computing the distance of each residue from the nearest bulk solvent (25). Using the XlinkAnalyzer dataset, we show that the faster EDTSurf was able to approximate DEPTH to a Pearson cross-correlation (CC) coefficient of 0.97 (supplemental Fig. S1). Thus, unless otherwise stated, residue depths mentioned in this work were computed with the Euclidean distance transform algorithm.

The depth distribution of tagged lysines is shown in Figure 2*A*, along with the depth distribution of all lysines across the 43 PDB structures assessed. In addition, the EDs between Cα atoms of crosslinked lysines were computed, with the distribution of distances shown in Figure 2*B*. The depth distributions were fitted with a negative power function, while the Cα–Cα distance distribution was fitted with a sigmoidal function using SciPy v1.5.2 (https://scipy.org/) (26) (Fig. 2).

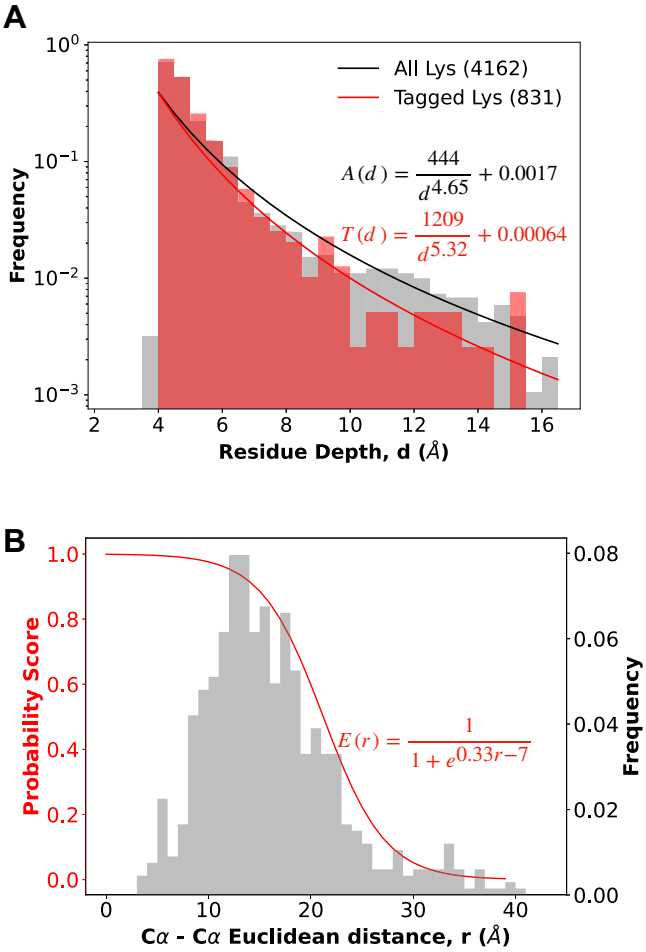


FIG. 2. **Empirical distributions of depths and Cα–Cα Euclidean distances of DSS- and BS3-crosslinked lysine residues.** Across 43 PDB structures, 831 unique lysine residues were found with an attached linker out of a total of 4162 lysines. *A*, depth distributions of tagged and untagged lysines were fit with a negative power function, while (*B*) the Cα–Cα distance distribution was fit with a sigmoidal function as shown. BS3, bissulfosuccinimidyl suberate; DSS, disuccinimidyl suberate; PDB, Protein Data Bank.

The monolink probability (MP) score for a monolink i in a given structure is thus defined as follows:

$$MP_i = \frac{T(d_i)}{A(d_i)}\delta_{d_i} - \frac{d_i}{\max(d)}(1-\delta_{d_i}), \text{where } \delta_{d_i} = \begin{cases} 1, & \text{if } d_i \le 15, \\ 0, & \text{if } d_i > 15. \end{cases} \quad (1)$$

with the first term corresponding to the ratio between tag probability T and lysine occurrence A at a given depth d_i in the structure (Fig. 2*A*), and the second term being a penalty for monolinks that are a depth of more than 15 Å in the structure, with the 15 Å cut-off set empirically from the observed residue depths. The MP score of a given model is the weighted average MP score of all monolinks found in the structure:

$$MP = \frac{\sum w_i MP_i}{\sum w_i} \quad (2)$$

The weight w_i of each monolink can be 1, if we are considering only the presence or absence of monolinks, but can also reflect relative abundance. In particular, we use occupancy o_i defined as follows:

$$w_i = o_i = \frac{I_{mono,i}}{I_{mono,i} + I_{unmod,i}} \quad (3)$$

where $I_{mono,i}$ is the monolinked peptide precursor ion intensity, and $I_{unmod,i}$ is the corresponding unmodified peptide ion intensity.

The crosslink probability (XLP) score for a crosslink spanning residues i and j in a given structure is then defined as:

$$XLP_{ij} = E(r_{ij})MP(d_i)MP(d_j)\omega_{r_{ij}}\delta_{d_i}\delta_{d_j} - \frac{r_{ij}(1-\omega_{r_{ij}})}{2*max(r)} - \frac{d_i(1-\delta_{d_i}) - d_j(1-\delta_{d_j})}{4*max(d)}$$

$$\text{where } \omega_{r_{ij}} = \begin{cases} 1, & \text{if } r_{ij} < 33, \\ 0, & \text{if } r_{ij} \ge 33. \end{cases} \quad (4)$$

with the first term corresponding to the likelihood E of the Cα–Cα distance r_{ij} being spanned by a BS3/DSS crosslink (Fig. 2*B*), weighted by the monolink probability MP at both residues i and j. The second term is a penalty for Cα–Cα distances equal to or exceeding 33 Å, while the third term is a penalty for either residue exceeding a maximum depth of 15 Å, as in the MP score. A threshold of 33 Å was set, as this is the maximum spanning distance for DSS and BS3 previously observed (13, 19). The weights for the penalty terms were set to 0.5 for Cα–Cα maximum distance violations and 0.25 for maximum residue depth violations. Similar to the MP score, XLP is averaged across all crosslinks found in the structure.

$$XLP = \frac{\sum w_{ij} XLP_{ij}}{\sum w_{ij}} \quad (5)$$

The weight w_{ij} of each crosslink can be 1, if we are considering only the presence or absence of crosslinks, but can also reflect relative abundance (*e.g.*, the product of the occupancies of each crosslinked position, $w_{ij} = o_i o_j$).

Unlike the monolink (27) and crosslink (19) scores we previously developed, both MP and XLP were formulated to scale smoothly from −1 to 1, with a score of 1 indicating that the given structure provides the highest probability that the given monolink or crosslink can be formed. Having such a defined scale means that the magnitude of the scores of a single model is independent of the distribution of other predicted models, and thus does not depend on the assumption that the models are normally distributed nor that the ensemble contains the correct model, unlike in other methods of score normalization, such as the computation of a *z*-score.

Benchmarking the MP and XLP Scores

The MP and XLP scores were evaluated on the structure decoy dataset 3DRobot, a set of 200 nonhomologous proteins randomly selected from the PDB, each with 300 structural decoys with RMSD from the native structure ranging from 0 to 12 Å (28). Sets of "simulated" crosslinks and monolinks were generated at different recovery rates using the native structures in the 3DRobot dataset. Similar to our previous work (27), we defined simulated monolinks as lysine residues in the native structure with a depth of 6.25 Å or less, as computed by DEPTH (25). Simulated crosslinks were defined as pairs of lysine residues in the native structure with a depth of at most 6.25 Å, and with a Cα–Cα SASD of at most 33 Å, computed from the structure with Jwalk (19). We simulated a range of recovery rates (*i.e.*, the fraction of detected crosslinks from all possible crosslinking events) from 10 to 100%, in increments of 10%. At each recovery rate, all possible crosslinking events were randomly sampled 1000 times, to simulate variation that may occur in XL-MS experiments. Similarly, we generated simulated monolinks in this bootstrapped manner, by randomly sampling all possible monolinks 1000 times at a range of recovery rates of 10 to 100%. To determine the XLP score's robustness to false crosslinks, over length crosslinks (SASD ≥ 33 Å) were also added in increasing fractions, from 10 to 100%, at each crosslink recovery rate.

At each recovery rate, the MP and XLP scores were calculated for each 3DRobot structure, using Equations 2 and 5 (with all $w_i = 1$ and $w_{ij} = 1$), bootstrapped across the 1000 random samplings. These bootstrapped MP and XLP scores were compared with a similarly bootstrapped MNXL (matched and nonaccessible crosslinks) score for each model in the 3DRobot dataset.

MNXL is defined as follows:

$$MNXL(i,j) = \begin{cases} N_{(18.62,35.94)}(r_{ij}), & \text{if } r_{ij} \le 33, \text{and i and j are solvent - accessible,} \\ -1, & \text{else.} \end{cases} \quad (6)$$

where i and j are the crosslinked residues, r_{ij} is the SASD between them, and $N_{(18.62,35.94)}$ is the normal distribution calculated from all SASDs ≤ 33 Å from the crosslink database XLdb (19).

A well-performing scoring function assigns higher ranks to structures that closely resemble the native structure and can effectively distinguish near-native structures from less accurate decoys. To assess ranking performance, we calculated the Spearman CC between the scores and the RMSD from the native structure. In addition, we measured each score's ability to identify near-native structures by computing the area under the receiver operating characteristic curve (AUC). The receiver operating characteristic curve plots the relationship between the number of true and false positives at different scoring thresholds. In our case, true positives were defined as near-native models (template modeling-score ≥ 0.9) scoring above the threshold, while false positives were non-native models (template modeling-score <0.9) also scoring above the threshold (29).

To determine if there were significant performance differences between XLP and MNXL, we compared the Spearman CC and AUC at each recovery rate using the Wilcoxon signed-rank test implemented in SciPy v1.5.2 (26).

Collating an Experimental XL-MS Dataset of Flexible Proteins

We compiled a test dataset of three proteins with experimental BS3 and/or DSS XL-MS data from the literature (supplemental Table S1). Each protein in the dataset had at least one "open" and one "closed" conformation in the PDB, which were defined based on their radius of

TABLE 1
Flexible proteins with PDB structures and experimental XL-MS data

Protein name (UniProt ID)	RMSD between conformations (Å)	PDBID, open	R_g, open (Å)	PDBID, closed	R_g, closed (Å)	Experimental XL-MS data
Complement component 3, C3 (P01024)	26.9	2i07 (54)	46.2	2a73 (55)	42.9	crosslinks (open and closed) (30)
Luciferase (P08659)	5.4	1lci (56)	24.6	4g36 (57)	23.3	crosslinks (open and closed) (32)
Glutamine-binding periplasmic protein (P0AEQ3)	5.3	1ggg (58)	19.0	1wdn (59)	17.5	crosslinks (closed) (31) monolinks (open, this study)

PDB, Protein Data Bank; R_g, radius of gyration; XL-MS, crosslinking mass spectrometry.

gyration, R_g (30–32). Two of the proteins—luciferase and glutamine-binding periplasmic protein (QBP)—consist of a single chain, while C3 is a protein complex consisting of two chains, one of which undergoes a large conformational change (>20 Å). Details of the dataset are summarized in Table 1. For QBP, no crosslinks or monolinks were reported for the open conformation. Simulating crosslinks and monolinks using our criteria from the previous section, we found that QBP has one lysine that is exposed only in the open conformation. To test whether this lysine is sufficient to distinguish QBP open conformation, QBP was produced and purified for subsequent XL-MS experiments, as described below.

Cloning, Expression and Purification of QBP

To obtain a C-terminal 6xHis-tagged QBP variant for purification, the ORF of QBP was amplified from *Escherichia coli* K12 DNA and cloned into bacterial expression vector pET28a-LIC (Addgene #26094). Transformed *E. coli* BL21 (DE3), cells were cultured in Terrific Broth (ROTH) at 37 °C to an A_{600} = 0.8, followed by IPTG (500 μM) induction of QBP expression for 5 h at 30 °C. Subsequently, cells were harvested, resuspended in lysis buffer (50 mM NaH2PO4; 300 mM NaCl; 10 mM imidazole; pH 8: supplemented with protease inhibitor cocktail [cOmplete, Roche], DNaseI [ITW REAGENTS], lysozyme [Sigma-Aldrich]), and sonicated. QBP was purified from whole-cell lysate in a three-step process using an ÄktaGo chromatography system (Cytiva). First, cleared lysate was loaded onto a HisTrapHP column (Cytiva) and washed with buffer A (50 mM NaH2PO4; 300 mM NaCl; 20 mM imidazole; pH 8). QBP was eluted in a buffer B (50 mM NaH2PO4; 300 mM NaCl; 500 mM imidazole; pH 8). Second, for release of bound glutamine guanidine hydrochloride was added to a final concentration of 6 M. Lastly, QBP was gradually dialysed in size-exclusion buffer (20 mM Hepes pH 7.5; 150 mM NaCl) and purified through size-exclusion chromatography (Superdex200, Cytiva). Samples of *E. coli* cultures were harvested and resuspended in *E. coli* lysis reagent (NEB). Lysates and purified proteins were separated by 4 to 12% Bis-Tris SDS-PAGE (Thermo Fisher Scientific) and visualized by Quick Coomassie stain (SERVA, supplemental Fig. S2).

Sample Crosslinking and Preparation for MS

QBP was buffer-exchanged into 20 mM Hepes, pH 7.5, at 4 °C (repeated for ten cycles) to remove salt using an Amicon Ultra-0.5 Centrifugal 10 kDa filter unit (Merck), before dilution to 5.00 μM in 20 mM Hepes. For glutamine-bound QBP, 5 mM glutamine was added prior to crosslinking. Next, 50 mM BS3-d0 crosslinker was added to glutamine-bound QBP, while 50 mM BS3-d4 crosslinker was added to glutamine-free QBP. Samples were then incubated at room temperature for 120 min before quenching with 20 mM ammonium bicarbonate. The glutamine-bound QBP with BS3-d0 and glutamine-free QBP with BS3-d4 were pooled before MS analysis. Samples were reduced with 8 mM DTT and incubated at 37°C for 45 min before alkylation using 20 mM iodoacetamide with incubation in darkness for 45 min, followed by a 1:5 dilution with 50 mM ammonium bicarbonate. Trypsin (Promega) was then added at a Trypsin:QBP mass ratio of 1:50 and incubated overnight at 37 °C. The digestion was then quenched by addition of trifluoroacetic acid to a final concentration of 0.50%. Samples were ZipTip desalted (Merck) before reconstitution in 1% formic acid at a concentration of 1.0 μg/μl.

LC-MS Setup and Data Collection

An UltiMateTM 3000 RSLCnano liquid chromatography system (Thermo Fisher Scientific), with a 50 cm μPAC Neo HPLC analytical column (COL-NANO050NEOB) and a 0.075 mm × 20 mm trap cartridge (Acclaim PepMap C18 100 Å, 3 μm), was connected to a SilicaTip emitter. Column temperature was set at 45 °C and column flow rate was set to 300 nl/min. Mobile phase A (0.1% formic acid, 3.2% acetonitrile) and mobile phase B (96.8% acetonitrile/0.1% formic acid) were applied with an elution gradient from 5.0 to 35.0% mobile phase B over 20 min. Total run time per sample was 30 min.

A Thermo Fisher Scientific Orbitrap Eclipse Tribrid mass spectrometer was used for this work. The mass spectrometer was externally calibrated using PierceTM FlexMixTM calibration solution. nanoESI was performed by the application of a voltage to a SilicaTip emitter, *via* a HPLC liquid junction cross. Spray stability and intensity was optimized by varying the SilicaTip electrospray voltage (1 kV–10 kV) and varying SilicaTip positioning (in the *x, y,* and *z* dimensions). Transfer capillary temperature was set to 275 °C, RF lens was set to 40%, precursor ion mass spectrum was acquired at a resolution of 120,000 with a mass range of 380 to 1400 *m/z* and a precursor ion charge state range of 3^+-8^+. MS1 spectra were recorded with a data-dependent analysis Top20 method, automatic gain control was set to a target of 400,000 (100%), maximum injection time mode was set to Auto, precursor ions were isolated with a 1.6 *m/z* window using the quadrupole mass filter and monoisotopic precursor selection was enabled. For Orbitrap mass analysis based MS2 acquisition, stepped higher energy collision induced dissociation was applied with a normalized energy of 25, 35%. Orbitrap resolution was set to 30,000, mass range was set to normal, automatic gain control was set to 200%, and maximum injection time was set to 54 ms.

Experimental Design and Statistical Rationale

For each condition (with *versus* without glutamine), two experimental replicates were performed, with each experimental replicate injected into the LC-MS/MS in technical triplicate. All validated

monolinked peptides were present in all three technical replicates of both replicates. This high stringency validation criterion ensured good reproducibility and statistical robustness within the experimental results. Data analysis was performed using Thermo Fisher Scientific Proteome Discoverer 2.5 (https://www.thermofisher.com/uk/en/home/industrial/mass-spectrometry/liquid-chromatography-mass-spectrometry-lc-ms/lc-ms-software/multi-omics-data-analysis/proteome-discoverer-software), with the Sequest HT search engine node used to search the RAW files generated from LC-MS/MS. Default settings were selected unless otherwise stated. RAW files were searched against a bespoke FASTA file containing the QBP sequence (UniProt accession P0AEQ3) and 200 randomly selected protein sequences from *Homo sapiens* UniProt version July 2022. Fixed value peptide-spectrum match validator node was used with false discovery rate (FDR): 0.05 (target-decoy method). Modifications included dynamic methionine oxidation (+15.995 Da), static cysteine carbamidomethylation (+57.021 Da), dynamic N-terminal acetylation (+42.0106 Da), dynamic lysine DSS-d0 amidated monolink (+155.095 Da), and dynamic lysine DSS-d4 amidated monolink (+159.120 Da) (note that search for DSS modification is identical to BS3). Precursor ion mass tolerance was set to 10 ppm, and fragment ion mass tolerance was set to 0.02 Da. Threshold score for accepting individual spectra was set to XCorr (Sequest search engine): equal to or greater than 3.00. For each validated monolinked lysine, its corresponding occupancy was determined using Equation 3. For all precursor ions used for this purpose, ion intensity was acquired over the liquid chromatography retention time range within which the precursor ion was eluted into the mass spectrometer. Precursor ion intensity for each peptide was obtained directly from XCalibur, and monolinked and corresponding nonmodified peptide precursor ions were always selected from the same MS RAW file. The resulting monolinks and occupancy values are listed in supplemental Table S1, which includes the lysine K:137 that was predicted to be exposed in the open conformation but occluded in the closed conformation of QBP (annotated MS/MS fragmentation spectrum in supplemental Fig. S3). All peptide and protein identifications information has been deposited on ProteomeXchange, Proteomics Identifications Database (PRIDE) dataset identifier: PXD046392.

Predicting an Ensemble of Conformations with AF2 and Selecting Near-native Models with XLP and MP

Conformational ensembles were predicted for luciferase and QBP using a modified version of AlphaFold v2.1.1 (1), which limits the MSA depth, resulting in a greater variation in predicted conformers than the default conditions (21). Alignment depths of 8, 16, 32, and 64 were used for template-free AF2 prediction for a total of 200 predicted models (50 per alignment depth).

For C3, which consists of two chains, AlphaFold-Multimer (AF-MM) v2.1.1 (33) was employed to model the complex. In this case, ten conformers were predicted per model, resulting in 50 AF-MM models. Alignment depth was not limited for C3, as the feature was not yet implemented for AF-MM. To avoid bias toward PDB-deposited structures, the max-template-date parameter was set to a date prior to the first deposited related structure, according to PDB's structural similarity query function (34).

The generated AF2 models were filtered based on their predicted template modeling (pTM) scores, an internal measure of modeling confidence in AF2. The distributions of pTM scores for each protein in the test dataset can be found in supplemental Fig. S4, along with the minimum score threshold used to remove models from further analysis.

Subsequently, XLP and MP scores were computed for each AF2 model and conformation. The AF2 models were ranked by decreasing XLP and MP scores, and the top-scoring models for each protein and conformation were visualized using UCSF ChimeraX (35).

Interpolating conformational ensembles for C3 and luciferase

As a theoretical exercise, to address the potential impact of limited sampling by AF2, we created structural ensembles for C3 and luciferase by morphing the structures from the open to closed conformation. This was accomplished using the morph function in ChimeraX, which generates intermediate conformations by interpolating along internal coordinates. The resulting ensembles, consisting of 50 models each, were subjected to energy minimization using GROMACS v2018.8 (https://www.gromacs.org/) with the AMBER94 force field (36, 37). Subsequently, the XLP score was applied to these morphed ensembles using experimentally determined crosslinks obtained from both the open and closed conformations. The top-scoring models were visualized using ChimeraX for further analysis.

RESULTS

Benchmarking monolink and crosslink scores with simulated XL-MS data

We first evaluated the performance of our model assessment scores, MP and XLP, using simulated monolink and crosslink data, by measuring the ranking accuracy (Spearman CC of the scores *versus* RMSD) and near-native selection (AUC) of each score on the 3DRobot dataset.

Figure 3*A* illustrates the performance of MP and XLP scores, compared to the matched and nonaccessible crosslink score (MNXL). All scoring functions demonstrated improved ranking accuracy (CC) and near-native structure selection (AUC) with increasing recovery rates. This is expected, as more crosslinks or monolinks provide additional constraints, narrowing down the range of possible conformations. Notably, XLP outperformed MNXL in both ranking accuracy and near-native selection in the low crosslink recovery regime. Removing depth information from XLP, which leaves only the ED, resulted in decreased performance compared to both XLP and MNXL, indicating the importance of depth information.

The monolink score MP showed the lowest performance among the assessed scores. However, it still demonstrated near-native selection accuracy comparable to XLP without depth information, particularly at higher recovery rates.

Adding increasing fractions of over length crosslinks to simulate an increasing FDR reduces XLP score performance proportionally. At low crosslink recovery rates (<0.4), XLP remains moderately discerning (AUC >0.7) if a maximum of 10% of the crosslinks are over length. At higher recovery rates, XLP remains moderately discerning at up to 30% FDR (Fig. 3*B*).

Conformational Variability Predicted by AF2 Using Shallow MSAs

We employed AF2 to predict ensembles of models for C3, luciferase, and QBP using shallow MSAs as described in the Experimental Procedures. Figure 4*A* displays the resulting AF2 ensembles, illustrating each model's RMSD from its

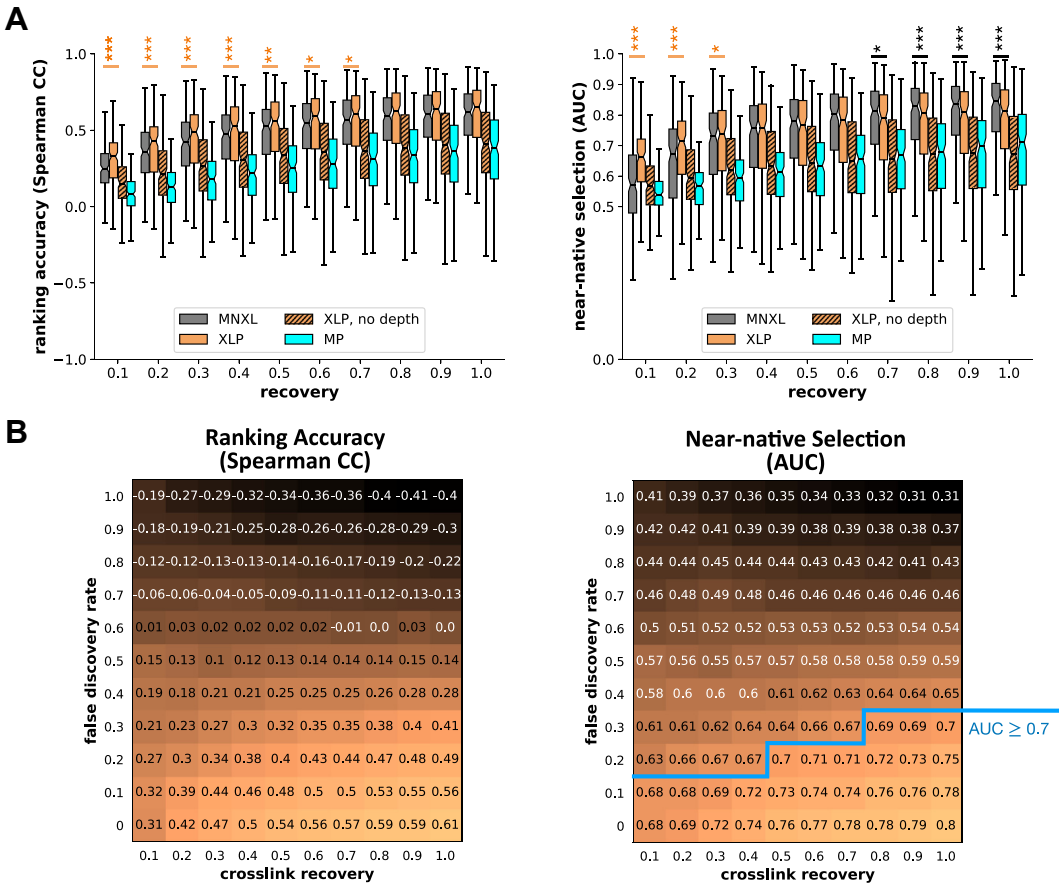


FIG. 3. **Crosslink and monolink-based scores can distinguish near-native structures in a simulated benchmark, with some tolerance for error.** *A*, performance of crosslink and monolink scoring functions in terms of ranking accuracy and near-native structure selection in the 3DRobot dataset across a range of 10 to 100% croslink or monolink recovery. Overall, XLP performs best among the scores, particularly at low crosslink recovery rates (≤30%). *Orange asterisks* denote recovery rates at which XLP significantly outperforms MNXL, while *black asterisks* denote the reverse (***$p < 0.001$, **$p < 0.01$, and *$p < 0.05$). Removing depth information from XLP degrades performance to below MNXL, but above the MP score. *B*, adding increasing fractions of false crosslinks to simulate a false discovery rate (FDR) results in worse XLP performance in both median ranking accuracy and median near-native selection across the 200 proteins in the 3DRobot dataset. The *blue line* delineates combinations of FDR and crosslink recovery, for which the AUC is moderately discerning of near-native structure (AUC ≥ 0.7). At low crosslink recovery rates, XLP remains moderately discerning at up to 10% FDR, and can tolerate up to 30% FDR at higher crosslink recovery rates. AUC, area under the receiver operating characteristic curve; MNXL, matched and nonaccessible crosslink score; XLP, crosslink probability.

respective open and closed PDB conformations. Notably, near-native models were observed for all cases except for the closed conformation of C3 and the open conformation of luciferase. Restricting MSA depth for luciferase and QBP resulted in an increase in model variability, compared to running AF2 under default conditions, where only the closed conformation was predicted (supplemental Fig. S5).

An interesting observation was that the models for C3 and QBP exhibited a linear distribution between the open and closed conformations (Fig. 4*A*). This suggests that these ensembles encompass a range of conformations, including near-open and near-closed states, as well as intermediate configurations between these "endpoints." However, for luciferase, the majority of model variation occurred away from the open-to-closed axis, albeit closer to the closed conformation. This indicates the presence of potentially misfolded structures. This observation is further supported by the 3D plot in supplemental Fig. S6, which demonstrates relatively low pTM scores off this axis for luciferase and QBP, indicating that AF2 is less confident about these off-axis models.

Filtering conformational variability from AF2 ensembles using experimental crosslinks and monolinks

We calculated XLP and MP scores for each AF2 model using experimental XL-MS data from either the open or closed conformation of C3, luciferase, and QBP. In all cases, the model with the highest XLP or MP score corresponded to a near-native structure, or, if a near-native model was not predicted by AF2, to a conformation within 1 Å of the most accurate model in the ensemble (Figs. 5 and 6). The near-native selection accuracy, as measured by the AUC, ranged from 0.69 for the closed conformation of QBP to ≥ 0.85 for the

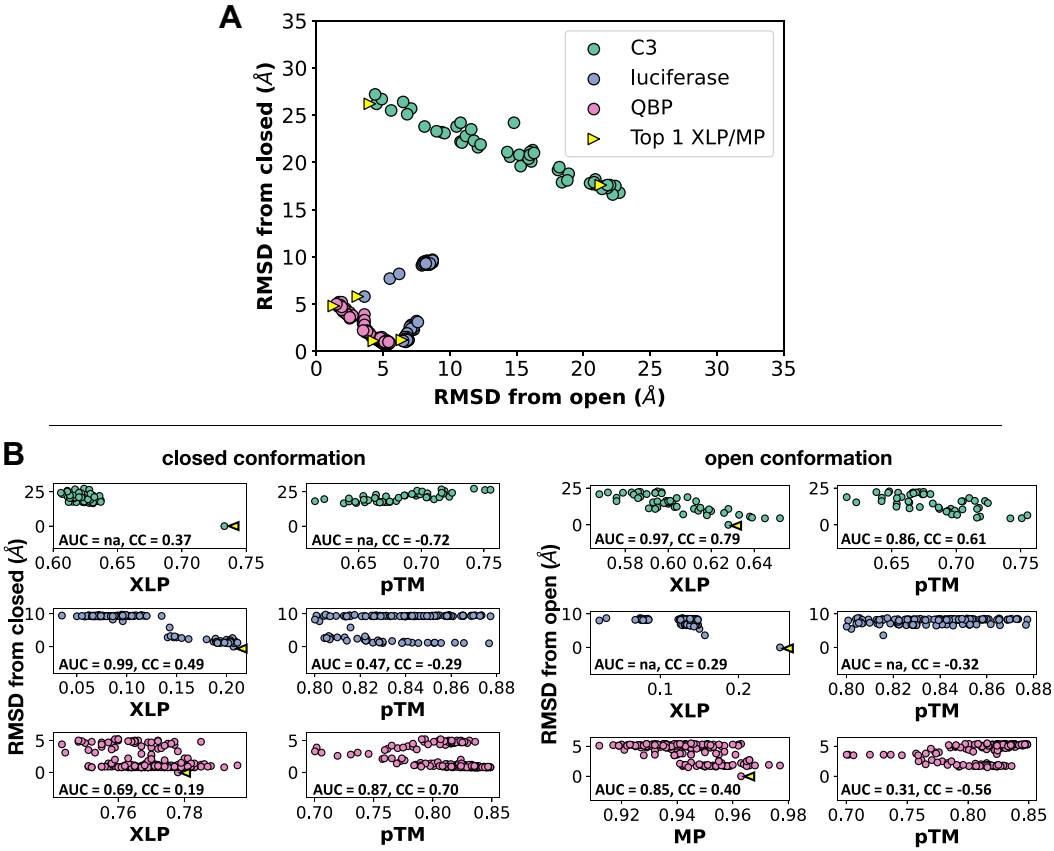


FIG. 4. **The XLP and MP scores are indicative of near-nativeness in an AF2-generated structural ensemble.** *A*, variation across predicted AF2 models for C3, luciferase, and QBP, expressed as the RMSD from their respective open and closed PDB conformations. Each point represents an AF2 model. *Yellow arrows* point to the Top 1-XLP/MP scoring models. In all cases, XLP or MP was either able to select a near-native model or in the absence of near-native structures, to select one of the lower RMSD structures from the set of AF2 predictions. For C3 and QBP, the AF2 structures lie along a *straight line* between the open and closed conformations, indicating that the AF2 ensemble might represent physically relevant intermediate structures between the two conformations. *B*, the XLP and MP scores have a roughly proportional relationship with model accuracy, and are thus inversely proportional to RMSD from the closed (*left*) and open (*right*) conformations. The corresponding RMSD *versus* pTM plots are shown for comparison. The PDB conformations are indicated by *yellow arrows*. Near-native selection (AUC) and ranking accuracy (CC) are indicated along each case (na, stands for no near-native structures found). The XLP and MP scores had better AUC and CC than the pTM score from AF2 in all cases except QBP closed conformation. AF2, AlphaFold2; AUC, area under the receiver operating characteristic curve; CC, cross-correlation; MP, monolink probability; PDB, Protein Data Bank; pTM, predicted template modeling; QBP, glutamine-binding periplasmic protein; XLP, crosslink probability.

other cases where AF2 predicted a near-native model (Fig. 4*B*). In contrast, when using the pTM score, the AUC varied across different conformations, showing selective behavior for the open conformation of C3 and the closed conformation of QBP (AUC ~0.9), antipredictive behavior for the open conformation of QBP (AUC = 0.31), and near-random performance for the closed conformation of luciferase (AUC = 0.47). XLP and MP scores also exhibited more accurate ranking of models compared to pTM, as measured by Spearman CC, except for the closed conformation of QBP (Fig. 4*B*). These findings demonstrate that the XLP and MP scores enable the identification of the conformer represented by a given set of XL-MS data from an ensemble of structural models. In contrast, the use of pTM alone is, at best, only selective for one conformation.

For all cases in our test dataset, we considered the monolinks and crosslinks as equally weighted events (all $w_i = w_{ij} = 1$ when calculating the MP and XLP scores). In the case of QBP open conformation, we also collected quantitative data and computed the occupancy of each monolink (Equation 3). Adding the occupancy values as weights improved the near-native selection and ranking accuracy from AUC = 0.80 and CC = 0.32 for the equally weighted version (Fig. S7), to AUC = 0.85 and CC = 0.40 for the occupancy-weighted MP score (Fig. 4*B*). This suggests that including occupancy might have a helpful effect on the scores but whether this effect generalizes will have to be benchmarked using a larger experimental XL-MS dataset.

It is self-evident, however, that our scoring functions cannot select for the correct model if it was not sampled by AF2. To

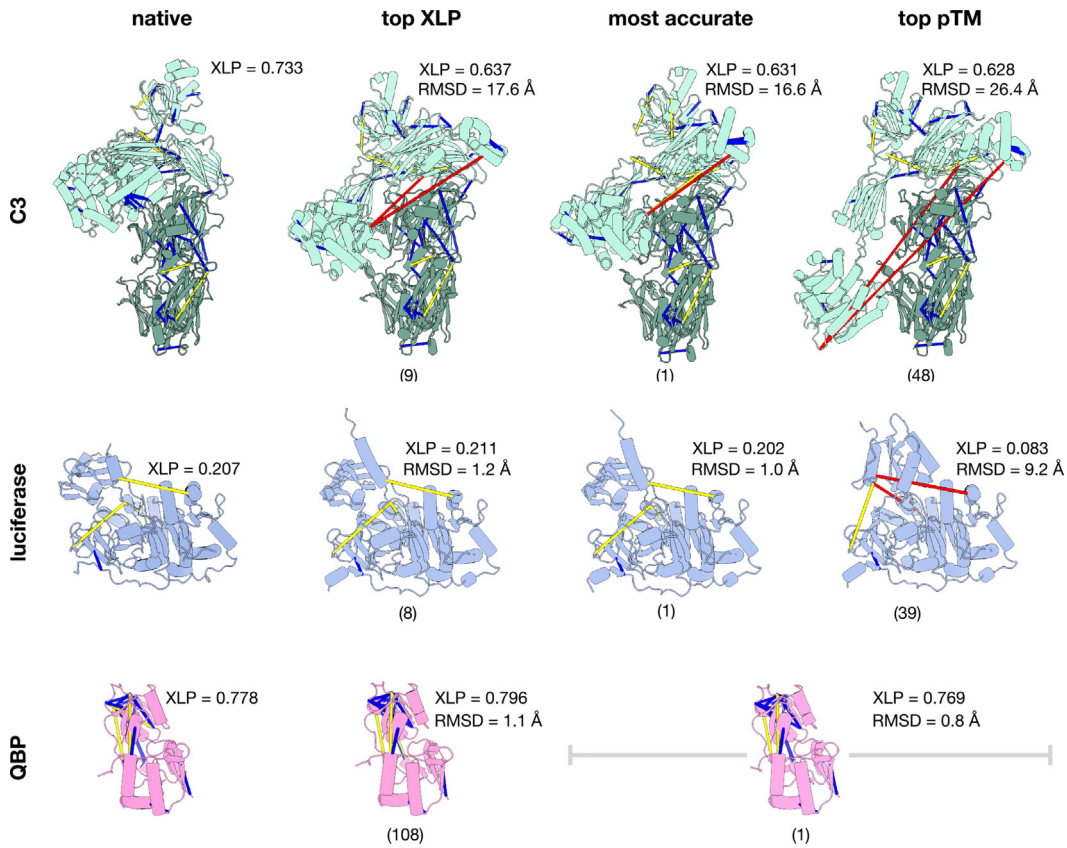


FIG. 5. **The XLP score selects for the closed conformation from among the AF2 ensembles of C3, luciferase, and QBP.** The PDB models are shown as reference, along with three representative structures from the AF2 ensembles: the *top* XLP or MP scorer, the most accurate (lowest RMSD from PDB structure), and the model with the highest pTM score. The corresponding XLP scores and RMSD from the target conformation are shown alongside each model, with the rank of the model in terms of accuracy shown below each structure in *brackets*. In each case, the experimentally derived crosslinks are shown, colored by Cα-Cα distance (*blue*: <21 Å, *yellow*: 21–33 Å, *red* >33 Å). AF2, AlphaFold2; MP, monolink probability; PDB, Protein Data Bank; pTM, predicted template modeling; QBP, glutamine-binding periplasmic protein; XLP, crosslink probability.

address this issue in the case of C3 and luciferase, we disentangled the effect of poor AF2 sampling by building an interpolated ensemble consisting of the open and closed conformation and intermediate structures between them (supplemental Fig. S8*A*). Using these interpolated ensembles resulted in very good performance of XLP in ranking (CC > 0.89), and near-native selection (AUC >0.95) for C3 open and closed, and luciferase open conformations, and reasonable ranking and near-native selection for luciferase closed conformation (CC = 0.68, AUC = 0.73) (supplemental Fig. S8*B*). For all cases, the top-XLP scoring models were near-native and within the top-10 models in terms of RMSD to the target (supplemental Fig. S8*C*).

DISCUSSION

AF2 represents a paradigm shift in structural biology in many ways, but it does not preclude the use of experimental methods, particularly in the case of flexible proteins. We show here that AF2 can fit very well in an integrative modeling pipeline for flexible proteins by providing an ensemble of structures that can be evaluated using distance constraints and solvent accessibility information from XL-MS.

In developing scoring functions based on XL-MS data, we wanted one that was more nuanced than just a count of maximum distance violations (12–15), to rank structures that are all possible within the relatively large distance constraint imposed by BS3 and DSS (13, 19). Within the umbrella of existing crosslink-based scoring functions, those based on SASD were found to perform better than ED (17–20). However, SASD computation is more time-consuming, as it involves tracing the possible paths of the crosslinker along the protein surface, whereas ED only requires a single distance computation. This results in around a tenfold decrease in runtime when using ED instead of SASD (supplemental Fig. S9).

According to our benchmark using simulated crosslinks, the performance of the SASD-based score MNXL can be approximated, and even exceeded, by our ED-based scoring function XLP, thereby circumventing the need for more time-consuming SASD computations and opening up the possibility of using XLP to assess large conformational ensembles.

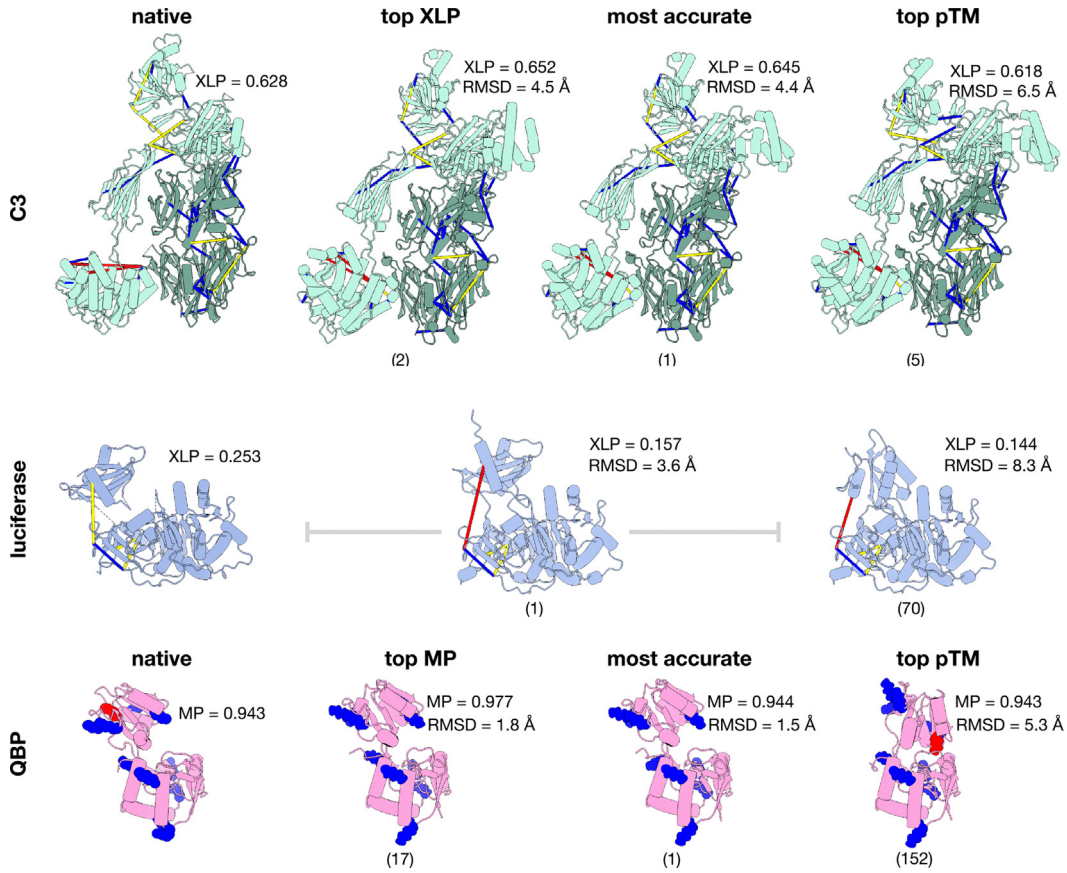


FIG. 6. **The XLP and MP scores select for the open conformation from among the AF2 ensembles of C3, luciferase and QBP.** The PDB models are shown as reference, along with three representative structures from the AF2 ensembles: the *top* XLP or MP scorer, the most accurate (lowest RMSD from PDB structure), and the model with the highest pTM score. The corresponding XLP or MP scores and RMSD from the target conformation are shown alongside each model, with the rank of the model in terms of accuracy shown below each structure in *brackets*. In each case, the experimentally derived crosslinks (for C3 and luciferase) or monolinks (for QBP) are shown. Crosslinks are colored by Cα–Cα distance (*blue*: <21 Å, *yellow*: 21–33 Å, *red* >33 Å), while monolinks are colored by residue depth (*blue*: ≤ 6.25 Å, *red*: >6.25 Å). AF2, AlphaFold2; QBP, glutamine-binding periplasmic protein; PDB, Protein Data Bank; pTM, predicted template modeling; XLP, crosslink probability.

These results are consistent with our previous findings that show that including solvent accessibility can improve the performance of an ED-based scoring function to almost the level of the SASD-based scoring function MNXL (19, 38). Instead of solvent accessibility, in our new scoring function we used residue depth, resulting in improved performance of XLP over MNXL, notably at low crosslink recovery rates of 30% or less. Performance improvements in the low crosslink recovery regime are particularly impactful, since experimental crosslink recovery rates are typically reported to be in this range (17, 19, 38–40), particularly in the noisier conditions found in *in situ* crosslinking (41, 42). We also showed that the XLP score remains robust even when 10 to 30% of the crosslinks are over length. This robustness to error is a useful attribute, especially if there is conformational heterogeneity in the experimental system, such that crosslinks from more than one conformation are detected. Based on the result, it is most likely safe to speculate that as long as the crosslink heterogeneity does not exceed approximately 30%, the XLP score can be used to select the dominant conformation without the need to partition the crosslinks into mutually exclusive sets. The XLP score was also shown to distinguish between different conformations of the proteins in our test dataset using experimentally determined crosslinks, with the top scoring model in each case either being the most accurate one or within 1 Å of the most accurate model.

We also conceptualized and benchmarked a scoring function based on monolinks, the MP score, which is a measure of the likelihood that an amino acid at a certain depth from the protein surface will be tagged with the crosslink reagent. Previously, we have demonstrated that although monolinks contain less information than crosslinks, they can also be used to select for near-native structures from a set of decoys, albeit at higher recovery rates of at least 50% (27). Our current work has confirmed these previous results. Furthermore, the MP score was needed to select for the open conformation of QBP as there were no crosslinks found exclusive to this conformation. This is due

to the small size of QBP, such that all possible crosslinking events can be spanned by BS3/DSS whether in open or closed conformation. What was informative in this case is the reduced residue depth that occurs when previously inaccessible regions of the protein are exposed to solvent when it is in the open conformation. Surprisingly, even one differentially exposed lysine residue was sufficient to distinguish the open conformation of QBP from the rest of the conformational ensemble (Fig. 6 and supplemental Table S1), indicating the potential of this approach in cases where more numerous conformation-specific monolinks could be found. Differential residue depth data can be obtained from experiments such as hydrogen-deuterium exchange MS (43), but one can also get this "for free" in XL-MS experiments in the form of monolinks, which are currently not frequently used for data analysis. We hope this demonstration of the importance of monolinks will encourage experimentalists to use monolink data more often and to deposit them in publicly available databases such as the PRIDE (44).

We also tested the effect of adding quantitative data to our scoring function by adding a weight (occupancy, defined in Equation 3) to each monolink found in QBP open conformation, resulting in a marked improvement in both ranking and near-native selection accuracy, compared to if the monolinks were considered equally weighted events. While it may be more physically accurate to model monolinking and crosslinking events in this differentially probable, quantitative manner, it remains to be seen whether this improvement will generalize to a larger, experimental XL-MS dataset. Nonetheless, it is a promising avenue that warrants future investigation.

We have also observed, during both the benchmarking and testing of XLP, that crosslink recovery is directly proportional to the ranking accuracy and near-native selection. It has also been previously shown that the location of crosslinkable sites in the protein also limits their potential information content (42, 45). For example, if both ends of a crosslink span a rigid region of an otherwise flexible protein, then it would not distinguish between different conformations of the protein. This position dependence can be ameliorated by increasing crosslink recovery with the use of crosslinking reagents with different reactive end groups to increase the likelihood of getting informative crosslinks in an XL-MS experiment. Another way to increase the discerning power of XL-MS based scoring functions is to use shorter crosslinkers, thereby providing narrower ranges of distance constraints (18, 46). This is particularly important to resolve the differences in conformation in a small protein like QBP in our dataset. In particular, zero-length crosslinkers have been described to not add any atoms of the crosslinked species, such that distance restraints cannot exceed salt-bridge distances. However, zero-length crosslinkers have also been shown to exceed their theoretical distance constraints and so care must be taken when applying stringent distance constraint criteria (47). Additionally, the inherent clustering of acidic residues may lead to ambiguous site localization of residues spanned by zero-length crosslinkers, further adding some uncertainty to the distance constraint. To reconcile these issues, a tolerance could be added to the maximum allowed Cα–Cα distance (similar to nonzero length crosslinks), as well as more stringent data acquisition and peptide validation to add confidence to the localizations of crosslinked residues (*e.g.* setting a XCorr threshold of 3.00 or greater). Nonetheless, taking the crosslink spacer atoms out of consideration would still result in around a 10 Å reduction in maximum allowed Cα–Cα distance compared to BS3/DSS, thereby providing stricter modeling constraints in principle. It should be noted, however, that since our proposed scoring function is based on empirical distributions of spanned Cα–Cα distances, this distribution would have to be computed separately for different crosslinking reagents, zero-length or otherwise, thus further highlighting the need to deposit experimental XL-MS data in PRIDE.

Thus we have quantified the performance of the XLP and MP scores on a large simulated dataset at varying crosslink and monolink recovery rates, and the robustness of the XLP score in the presence of increasing fractions of false crosslinks. We also showed that the scores are also able to select near-native or near-best models when using experimental crosslink and monolink information. In practice, however, our current approach has two main limitations. Firstly, while the XLP and MP scores are a good measure of relative model quality within an ensemble, it is difficult to set a threshold score value that has both good precision and recall. Case in point, setting the XLP score threshold to 0.6 would result in the PDB model of luciferase closed conformation to be considered incorrect (Fig. 5), while lowering the threshold to 0.2 to include this model would result in too many false positives for other proteins like C3 closed conformation, wherein all of the XLP scores are above 0.6. This suggests that other factors, such as linker length relative to protein size, might play a role in determining which models can be considered as high-confidence models, given a certain XLP or MP score. Determining what these factors are is a logical next step to this work.

Secondly, we are limited by the ability of structure prediction software to predict conformational ensembles well. In this case, running AF2 with a shallow MSA allowed the endpoint PDB structures of QBP to be predicted, as well as plausible intermediate structures. This is less the case for C3, for which only models near the open conformation were predicted, as well as some putative intermediate structures but no conformations at or near the PDB closed conformation. The lack of AF2 models near the closed conformation is hinted at by the two maximum distance violations in the top-scoring model (in red in Fig. 5), indicating that the model can be further refined. These refinements can be done through steered molecular dynamics or normal mode analysis (48, 49)

but only up to the point that the maximum distance violations are no longer present, as any more optimizations are probably not supported by the spatial resolution of the crosslinker. For luciferase, the structural variation captured by AF2 occurs along a completely different axis than the endpoint structures documented in the PDB, suggesting that these predicted conformations might not be physically relevant. This trade-off between modeling accuracy and conformational variability as a function of alignment-depth has been discussed in greater depth in the publication detailing the technique of using shallow MSAs for AF2 (21). In any case, structure prediction that is natively trained to predict ensembles instead of single models would probably perform better than modifications of AF2, which was not trained for this purpose. Alternatively, AF2 is widely documented to predict domain structures well, as seen in relatively low predicted aligned error, within domains (supplemental Fig. S5). These initial domain predictions can be used as starting structures, with different domain–domain orientations sampled through molecular dynamics or normal mode analysis, while keeping the domains rigid in a process analogous to protein–protein docking.

Though not explicitly benchmarked on a large dataset, the XLP and MP scores can be used to evaluate not only protein monomers but also complexes, as shown with C3. This is because the underlying distributions of residue depths and Cα–Cα distances that were used to compute monolinking and crosslinking probabilities (Fig. 2) came from empirical XL-MS data from both protein monomers and complexes. Crosslinks have been widely used to screen large sets of protein–protein docking models (45, 50, 51), but in the case of C3, AF-MM correctly predicted the binding interface for all models, only differing in the conformation of the flexible chain. As of current writing, the problem of protein complex prediction is not as well solved as the structure prediction of single protein chains, but we might extrapolate that the former problem will soon be solved with additional algorithmic refinements and/or protein complex structures in the PDB. Indeed, with the emergence of AF-MM, the use case for crosslinks is shifting away from the evaluation of large sets of protein–protein docking poses, to instead serving as a prior indication of which protein–protein interactions to model with AF-MM or as experimental validation of AlphaFold models (52). Nonetheless, whichever part of the modeling pipeline the XL-MS data are to be applied—whether to guide modeling (53), to evaluate a set of models, as in this current work, or as experimental validation of a predicted model—the XLP and MP scores are simple and modular enough to be used.

DATA AVAILABILITY

The source code of xlms-tools is available as a git repository at https://gitlab.com/topf-lab/xlms-tools. The AlphaFold and interpolated models can be accessed at Zenodo (https://doi.org/10.5281/zenodo.8304594). XL-MS data collected for this work were deposited in ProteomeXchange (PRIDE dataset identifier PXD046392).

Supplemental data—This article contains supplemental data (23–25).

Funding and additional information—This work was supported by Universität Hamburg and Hamburg-X grant LFF-HHX-03 to the Center for Data and Computing in Natural Sciences (CDCS) from the Hamburg Ministry of Science, Research, Equalities, and Districts (BWFGB), the Leibniz Institute of Virology as part of Leibniz ScienceCampus InterACt (funded by the BWFGB Hamburg and the Leibniz Association) (to M. T. and K. G.), a Wellcome Collaborative Award in Science (209250/Z/17/Z) (to K. G., K. T., and M. T.), and a Wellcome Trust Multiuser Equipment grant (221521/Z/20/Z) (to K. T.).

Author contributions—K. M.-C. and M. T. conceptualization; K. M.-C. software; K. M.-C. and K. R. A. formal analysis; K. M.-C. data curation; K. M.-C., K. R. A., and M. P. writing-original draft; K. M.-C. visualization; K. R. A., M. P., and B. M. investigation; K. G., K. T., and M. T. supervision; K. G. funding acquisition; K. T. and M. T. resources; K. G., K. T., and M. T. writing-review and editing.

Conflict of interest—The authors declare no competing interests.

Abbreviations—The abbreviations used are: AF2, AlphaFold2; AF-MM, AlphaFold-Multimer; AUC, area under the receiver operating characteristic curve; CC, cross-correlation; BS3, bissulfosuccinimidyl suberate; DSS, disuccinimidyl suberate; ED, Euclidean distance; FDR, false discovery rate; MS, mass spectrometry; MNXL, matched and nonaccessible crosslink score; MP, monolink probability; MSA, multiple-sequence alignment; PDB, Protein Data Bank; PRIDE, Proteomics Identifications Database; pTM, predicted template modeling; QBP, glutamine-binding periplasmic protein; SASD, solvent-accessible surface distance; XL-MS, crosslinking mass spectrometry; XLP, crosslink probability.

Received September 13, 2023, and in revised form, December 23, 2023 Published, MCPRO Papers in Press, January 22, 2024, https://doi.org/10.1016/j.mcpro.2024.100724

REFERENCES

1. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., *et al*. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589
2. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., *et al*. (2022) AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439–D444

3. Perrakis, A., and Sixma, T. K. (2021) AI revolutions in biology: the joys and perils of AlphaFold. *EMBO Rep.* **22**, e54046
4. Chakravarty, D., and Porter, L. L. (2022) AlphaFold2 fails to predict protein fold switching. *Protein Sci.* **31**, e4353
5. Ha, J.-H., and Loh, S. N. (2012) Protein conformational switches: from nature to design. *Chemistry* **18**, 7984–7999
6. Burley, S. K., Bhikadiya, C., Bi, C., Bittrich, S., Chao, H., Chen, L., *et al*. (2023) RCSB Protein Data Bank (RCSB.org): delivery of experimentally-determined PDB structures alongside one million computed structure models of proteins from artificial intelligence/machine learning. *Nucleic Acids Res.* **51**, D488–D508
7. Le Gall, T., Romero, P. R., Cortese, M. S., Uversky, V. N., and Dunker, A. K. (2007) Intrinsic disorder in the Protein Data Bank. *J. Biomol. Struct. Dyn.* **24**, 325–342
8. Rappsilber, J. (2011) The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *J. Struct. Biol.* **173**, 530–540
9. O'Reilly, F. J., and Rappsilber, J. (2018) Cross-linking mass spectrometry: methods and applications in structural, molecular and systems biology. *Nat. Struct. Mol. Biol.* **25**, 1000–1008
10. Liu, F., and Heck, A. J. R. (2015) Interrogating the architecture of protein assemblies and protein interaction networks by cross-linking mass spectrometry. *Curr. Opin. Struct. Biol.* **35**, 100–108
11. Merkley, E. D., Rysavy, S., Kahraman, A., Hafen, R. P., Daggett, V., and Adkins, J. N. (2014) Distance restraints from crosslinking mass spectrometry: mining a molecular dynamics simulation database to evaluate lysine-lysine distances. *Protein Sci.* **23**, 747–759
12. Young, M. M., Tang, N., Hempel, J. C., Oshiro, C. M., Taylor, E. W., Kuntz, I. D., *et al*. (2000) High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 5802–5806
13. Chen, Z. A., Jawhari, A., Fischer, L., Buchen, C., Tahir, S., Kamenski, T., *et al*. (2010) Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J.* **29**, 717–726
14. Kalisman, N., Adams, C. M., and Levitt, M. (2012) Subunit order of eukaryotic TRiC/CCT chaperonin by cross-linking, mass spectrometry, and combinatorial homology modeling. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 2884–2889
15. Tüting, C., Iacobucci, C., Ihling, C. H., Kastritis, P. L., and Sinz, A. (2020) Structural analysis of 70S ribosomes by cross-linking/mass spectrometry reveals conformational plasticity. *Sci. Rep.* **10**, 12618
16. Kahraman, A., Malmström, L., and Aebersold, R. (2011) Xwalk: computing and visualizing distances in cross-linking experiments. *Bioinformatics* **27**, 2163–2164
17. Kahraman, A., Herzog, F., Leitner, A., Rosenberger, G., Aebersold, R., and Malmström, L. (2013) Cross-link guided molecular modeling with ROSETTA. *PLoS One* **8**, e73411
18. Hofmann, T., Fischer, A. W., Meiler, J., and Kalkhof, S. (2015) Protein structure prediction guided by crosslinking restraints–A systematic evaluation of the impact of the crosslinking spacer length. *Methods* **89**, 79–90
19. Matthew Allen Bullock, J., Schwab, J., Thalassinos, K., and Topf, M. (2016) The Importance of non-accessible crosslinks and solvent accessible surface distance in modeling proteins with restraints from crosslinking mass spectrometry. *Mol. Cell. Proteomics* **15**, 2491–2500
20. Degiacomi, M. T., Schmidt, C., Baldwin, A. J., and Benesch, J. L. P. (2017) Accommodating protein dynamics in the modeling of chemical cross-links. *Structure* **25**, 1751–1757.e5
21. Del Alamo, D., Sala, D., Mchaourab, H. S., and Meiler, J. (2022) Sampling alternative conformational states of transporters and receptors with AlphaFold2. *Elife* **11**, e75751
22. Kosinski, J., von Appen, A., Ori, A., Karius, K., Müller, C. W., and Beck, M. (2015) Xlink Analyzer: software for analysis and visualization of cross-linking data in the context of three-dimensional structures. *J. Struct. Biol.* **189**, 177–183
23. Xu, D., and Zhang, Y. (2009) Generating triangulated macromolecular surfaces by euclidean distance transform. *PLoS One* **4**, e8140
24. Xu, D., Li, H., and Zhang, Y. (2013) Protein depth calculation and the use for improving accuracy of protein fold recognition. *J. Comput. Biol.* **20**, 805–816
25. Tan, K. P., Varadarajan, R., and Madhusudhan, M. S. (2011) DEPTH: a web server to compute depth and predict small-molecule binding cavities in proteins. *Nucleic Acids Res.* **39**, W242–W248
26. Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., *et al*. (2020) SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. *Methods* **17**, 261–272
27. Sinnott, M., Malhotra, S., Madhusudhan, M. S., Thalassinos, K., and Topf, M. (2020) Combining information from crosslinks and monolinks in the modeling of protein structures. *Structure* **28**, 1061–1070.e3
28. Deng, H., Jia, Y., and Zhang, Y. (2016) 3DRobot: automated generation of diverse and well-packed protein structure decoys. *Bioinformatics* **32**, 378–387
29. Zhang, Y., and Skolnick, J. (2004) Scoring function for automated assessment of protein structure template quality. *Proteins* **57**, 702–710
30. Chen, Z. A., Pellarin, R., Fischer, L., Sali, A., and Nilges, M. (2016) Structure of complement C3 (H2O) revealed by quantitative cross-linking/mass spectrometry and modeling. *Mol. Cell Proteomics* **15**, 2730–2743
31. Ding, Y.-H., Gong, Z., Dong, X., Liu, K., Liu, Z., Liu, C., *et al*. (2017) Modeling protein excited-state structures from "over-length" chemical cross-links. *J. Biol. Chem.* **292**, 1187–1196
32. Walzthoeni, T., Joachimiak, L. A., Rosenberger, G., Röst, H. L., Malmström, L., Leitner, A., *et al*. (2015) xTract: software for characterizing conformational changes of protein complexes by quantitative cross-linking mass spectrometry. *Nat. Methods* **12**, 1185–1190
33. [preprint] Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., *et al*. (2022) Protein complex prediction with AlphaFold-Multimer. *bioRxiv*. https://doi.org/10.1101/2021.10.04.463034
34. Guzenko, D., Burley, S. K., and Duarte, J. M. (2020) Real time structural search of the Protein Data Bank. *PLoS Comput. Biol.* **16**, e1007970
35. Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., *et al*. (2021) UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82
36. Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., *et al*. (2015) GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19–25
37. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., *et al*. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **117**, 5179–5197
38. Bullock, J. M. A., Sen, N., Thalassinos, K., and Topf, M. (2018) Modeling protein complexes using restraints from crosslinking mass spectrometry. *Structure* **26**, 1015–1024.e2
39. Herzog, F., Kahraman, A., Boehringer, D., Mak, R., Bracher, A., Walzthoeni, T., *et al*. (2012) Structural probing of a protein phosphatase 2A network by chemical cross-linking and mass spectrometry. *Science* **337**, 1348–1352
40. Schweppe, D. K., Zheng, C., Chavez, J. D., Navare, A. T., Wu, X., Eng, J. K., *et al*. (2016) XLinkDB 2.0: integrated, large-scale structural analysis of protein crosslinking data. *Bioinformatics* **32**, 2716–2718
41. McCafferty, C. L., Pennington, E. L., Papoulas, O., Taylor, D. W., and Marcotte, E. M. (2023) Does AlphaFold2 model proteins' intracellular conformations? An experimental test using cross-linking mass spectrometry of endogenous ciliary proteins. *Commun. Biol.* **6**, 421
42. O'Reilly, F. J., Graziadei, A., Forbrig, C., Bremenkamp, R., Charles, K., Lenz, S., *et al*. (2023) Protein complexes in cells by AI-assisted structural proteomics. *Mol. Syst. Biol.* **19**, e11544
43. Masson, G. R., Burke, J. E., Ahn, N. G., Anand, G. S., Borchers, C., Brier, S., *et al*. (2019) Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat. Methods* **16**, 595–602
44. Perez-Riverol, Y., Bai, J., Bandla, C., García-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S., *et al*. (2022) The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552
45. Leitner, A., Walzthoeni, T., Kahraman, A., Herzog, F., Rinner, O., Beck, M., *et al*. (2010) Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol. Cell. Proteomics* **9**, 1634–1649
46. Piersimoni, L., Kastritis, P. L., Arlt, C., and Sinz, A. (2022) Cross-linking mass spectrometry for investigating protein conformations and protein-protein Interactions—A method for all seasons. *Chem. Rev.* **122**, 7500–7531
47. Rivera-Santiago, R. F., Sriswasdi, S., Harper, S. L., and Speicher, D. W. (2015) Probing structures of large protein complexes using zero-length cross-linking. *Methods* **89**, 99–111
48. Gniewek, P., Kolinski, A., Jernigan, R. L., and Kloczkowski, A. (2012) Elastic network normal modes provide a basis for protein structure refinement. *J. Chem. Phys.* **136**, 195101

49. Brodie, N. I., Popov, K. I., Petrotchenko, E. V., Dokholyan, N. V., and Borchers, C. H. (2017) Solving protein structures using short-distance cross-linking constraints as a guide for discrete molecular dynamics simulations. *Sci. Adv.* **3**, e1700479
50. Sinz, A. (2006) Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. *Mass Spectrom. Rev.* **25**, 663–682
51. Vreven, T., Schweppe, D. K., Chavez, J. D., Weisbrod, C. R., Shibata, S., Zheng, C., *et al*. (2018) Integrating cross-linking experiments with ab initio protein-protein docking. *J. Mol. Biol.* **430**, 1814–1828
52. Bartolec, T. K., Vázquez-Campos, X., Norman, A., Luong, C., Johnson, M., Payne, R. J., *et al*. (2023) Cross-linking mass spectrometry discovers, evaluates, and corroborates structures and protein-protein interactions in the human cell. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2219418120
53. Stahl, K., Graziadei, A., Dau, T., Brock, O., and Rappsilber, J. (2023) Protein structure prediction with in-cell photo-crosslinking mass spectrometry and deep learning. *Nat. Biotechnol.* **41**, 1810–1819
54. Janssen, B. J. C., Christodoulidou, A., McCarthy, A., Lambris, J. D., and Gros, P. (2006) Structure of C3b reveals conformational changes that underlie complement activity. *Nature* **444**, 213–216
55. Janssen, B. J. C., Huizinga, E. G., Raaijmakers, H. C. A., Roos, A., Daha, M. R., Nilsson-Ekdahl, K., *et al*. (2005) Structures of complement component C3 provide insights into the function and evolution of immunity. *Nature* **437**, 505–511
56. Conti, E., Franks, N. P., and Brick, P. (1996) Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes. *Structure* **4**, 287–298
57. Sundlov, J. A., Fontaine, D. M., Southworth, T. L., Branchini, B. R., and Gulick, A. M. (2012) Crystal structure of firefly luciferase in a second catalytic conformation supports a domain alternation mechanism. *Biochemistry* **51**, 6493–6495
58. Hsiao, C. D., Sun, Y. J., Rose, J., and Wang, B. C. (1996) The crystal structure of glutamine-binding protein from Escherichia coli. *J. Mol. Biol.* **262**, 225–242
59. Sun, Y. J., Rose, J., Wang, B. C., and Hsiao, C. D. (1998) The structure of glutamine-binding protein complexed with glutamine at 1.94 A resolution: comparisons with other amino acid binding proteins. *J. Mol. Biol.* **278**, 219–229