QRB Discovery

www.cambridge.org/qrd

Mass spectrometry integrates protein design into structural biology method development

Alexander Stevens[1], Hannah Osterholz[1], Vsevolod Viliuga[2,3], Thibault Vosselman[1], Surabhi Kokane[1], Erik G. Marklund[4], Arne Elofsson[2], Axel Leppert[1] and Michael Landreh[1,5]

[1]Department of Cell and Molecular Biology, Uppsala University, Sweden; [2]Science for Life Laboratory and the Department of Biochemistry and Biophysics, Stockholm University, Sweden; [3]Max Planck Institute for Polymer Research, Mainz, Germany; [4]Department of Chemistry for Life Sciences, Uppsala University, Sweden and [5]Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Sweden

Perspective

Cite this article: Stevens A, Osterholz H, Viliuga V, Vosselman T, Kokane S, Marklund EG, Elofsson A, Leppert A, Landreh M (2026). Mass spectrometry integrates protein design into structural biology method development. *QRB Discovery*, **7**: e4, 1–8
https://doi.org/10.1017/qrd.2025.10018.

Received: 23 October 2025
Revised: 08 December 2025
Accepted: 29 December 2025

Keywords:
biophysical methods; ion mobility mass spectrometry; machine learning; protein structure analysis; structure prediction

Corresponding author:
Michael Landreh;
Email: michael.landreh@icm.uu.se

Abstract

Recent advances in machine learning (ML) have transformed protein science, enabling engineering and *de novo* design of artificial proteins with novel structures and functions. However, experimental analysis of key design features, such as oligomerization, folding, ligand binding, and dynamic conformational changes, remains critical. Here, we outline how mass spectrometry (MS) complements protein design through its ability to corroborate a wide range of design objectives. Furthermore, engineered proteins have become valuable tools for exploring the use of MS in detecting structural features, charge effects, and weak interactions by serving as testbeds for method development. Integrating ML and native MS thus creates a feedback loop: new designs challenge analytical techniques, while improved methods provide richer data to guide and improve future predictions. This synergy is vital for expanding the capabilities of protein engineering, including toward applications in synthetic biology and artificial protocell development.

Introduction

Recent advances in machine learning (ML) have enabled breakthroughs in protein science. A range of ML-based computational tools has been developed as improved solutions to the protein structure prediction problem. These tools, including AlphaFold2 and 3, RosettaFold, ColabFold, and others, utilize deep learning to produce highly accurate structural models (Jumper *et al.*, 2021; Mirdita *et al.*, 2022; Baek *et al.*, 2023; Wohlwend *et al.*, 2025). Besides being able to predict structure from sequence, these advances also revolutionized protein design by being able to predict protein sequences that adopt a specific fold (Goverde *et al.*, 2023; Frank *et al.*, 2024; Pacesa *et al.*, 2024). ML-based tools have enabled co-design of protein structures and interactions with other biomolecules and ligands (Abramson *et al.*, 2024; Baek *et al.*, 2024; Krishna *et al.*, 2024; Butcher *et al.*, 2025). Beyond its impact on the analysis of native proteins and their complexes, these advances have enabled the development of artificial, that is, not naturally evolved, proteins in two ways:

(1) ML has expanded the horizons of protein engineering. The sequence of an existing scaffold can be modified to accommodate new structural or functional features (Nikolaev *et al.*, 2024; Sumida *et al.*, 2024), and their impact can be assessed using structure predictions. Alternatively, motif scaffolding poses another interesting application where a natural functional structure motif is fixed, and a *de novo* backbone accommodating it is generated.
(2) *De novo* design of proteins has expanded the protein fold universe, building proteins with new structures using physical, and not evolutionary, constraints. Here, a design objective is set, and new structures are generated using a computational pipeline, which usually has one or more foundational protein ML models at its core.

Both approaches hold great promise for protein science, ranging from the development of new biomaterials to environmental engineering to the targeted modulation of cellular pathways (Chu *et al.*, 2024; Kortemme, 2024; Wang *et al.*, 2024). However, the computational breakthroughs are intricately linked to insights from biophysical protein characterization. The underlying neural networks must be trained on large numbers of experimentally determined structures from the Protein Data Bank, using data from X-ray crystallography, nuclear magnetic resonance, and cryo-electron microscopy. Conversely, experimental methods are also required to corroborate the computationally designed proteins, usually by solving their high-resolution structures with NMR, X-ray crystallography, or cryo-EM.

The implication of the relationship between ML-based predictions and experiment is that the scope of what can be predicted computationally is limited by what can be observed experimentally in sufficiently large quantities to generate training data (Chakravarty *et al.*, 2025). This fact

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

has largely limited the prediction and design pipelines to stably folded proteins, although advances are being made toward including disorder and dynamics. Strategies such as introducing ambiguity in the structure predictions by clustering or sub-sampling the sequence alignments can help to uncover alternative states in folded proteins (Monteiro da Silva *et al.*, 2024; Wayment-Steele *et al.*, 2024; Kalakoti and Wallner, 2025). For intrinsically disordered regions (IDRs), large-scale molecular dynamics simulations can be used to generate network training data (Tesei *et al.*, 2024) and enable the design of IDRs with specific conformational preferences (Pesce *et al.*, 2024). Disordered and dynamic structures have been included in the design pipelines, such as the *de novo* design of folded binders that induce disorder-to-order transitions in their targets (Liu *et al.*, 2025; Wu *et al.*, 2025), and *de novo* design of enzymes using state-of-the-art all-atom methods (Butcher *et al.*, 2025). Together, these developments have fostered the prediction and design of protein structural dynamics. One example is the design of conformational switches that change their dynamics upon external stimuli such as ion binding (Praetorius *et al.*, 2023; Guo *et al.*, 2025), another is the design of disordered low-complexity sequences that facilitate controlled liquid–liquid phase separation (LLPS) (von Bülow *et al.*, 2025). Combining folded domains that engage in specific interactions with IDRs that mediate dynamic association is a potential step toward cell-like self-organization abilities (Mohanty *et al.*, 2022; von Bülow *et al.*, 2025).

Here, we describe how mass spectrometry (MS) can be used to corroborate protein designs, but also how the concept can be reversed by using designed proteins to test the limits of structural proteomics. Based on these insights, we suggest specifically designed or engineered proteins as a versatile tool to test and expand the limits of biophysical methods.

Corroborating protein design with MS

Considering the rapid pace of the field, corroborating the increasingly complex designs may become a potential bottleneck for future development. Since the three-dimensional fold predictions are already of high accuracy, experimental investigations are likely to focus on interactions and dynamics. We have previously demonstrated how MS can be integrated with ML-based structure predictions to monitor binding, folding, and stability of proteins (Allison *et al.*, 2022). Although MS generally does not provide sufficient constraints to solve a protein structure from scratch, it can corroborate predicted interactions. Electrospray ionization (ESI)-MS enables the transfer of proteins into the gas phase with minimal distortion of their structural features (Konermann *et al.*, 2013). This approach, termed 'native MS', can be used to measure the masses of intact complexes (Benesch *et al.*, 2007; Lössl *et al.*, 2016). Protein folded states, as well as, in some cases, conformational changes, can be detected by monitoring changes in ion charge state distributions (Konermann *et al.*, 2013; Chingin and Barylyuk, 2018). Native ion mobility MS (IM-MS) can be used to determine the collision cross sections (CCS) of protein ions, which provides insights into their architecture and dynamics (Christofi and Barran, 2023).

MS in combination with protein digestion and liquid chromatography can additionally detect local dynamics through crosslinking or hydrogen/deuterium exchange, which improves structure prediction accuracy (Stahl *et al.*, 2024; Wang *et al.*, 2025). In this manner, MS can capture complex structural processes such as enzyme dynamics, bridging the gap between structure and function (Busenlehner and Armstrong, 2005; Marklund and Benesch, 2019).

However, connecting structural features in the gas phase to predicted or experimental solution structures is not without its challenges. Desolvated protein complexes can self-solvate, where previously solvent-exposed residues form new bonds with neighboring residues or the backbone (Breuker and McLafferty, 2008). Unsupported domains, disordered loops, and cavities can undergo significant compaction when the solvent is removed (Hansen *et al.*, 2018; Rolland *et al.*, 2022). As the entire energy landscape is altered in a vacuum, proteins can adopt entirely new conformations that are not accessible in solution (Cropley *et al.*, 2024). It is important that these effects are considered when comparing MS data to protein structure predictions. Despite these caveats, it was recently found that the use of gas phase CCS data from IM-MS in the scoring function consistently improves the accuracy of AF2 models (Turzo *et al.*, 2022).

The capability to inform about interactions, stabilities, and conformations of proteins makes MS a complementary strategy to ML-based structure prediction. The same applies to *de novo* protein designs, but with one key difference: corroboration of predicted structures largely means confirming oligomeric states and architectures. Validation of a protein design, on the other hand, must confirm the central design objective(s), which can be far more diverse and range from the structural to the functional (Figure 1a, b). MS can detect many designed features: Oligomeric states and ligand binding can be studied with intact mass measurements, interface stabilities can be determined using gas-phase dissociation, the impact of covalent modifications and sequence variations can be validated using native top-down sequencing, and folded-unfolded transitions manifest themselves in changes in charge state distributions or cross sections in ion mobility experiments (Figure 1c). Intact mass measurements have been applied to show the desired stoichiometry of designed protein nanocages (Khmelinskaia *et al.*, 2025). Relative stabilities of hetero-trimers were determined by comparing signal intensities for each species in the mass spectra. Including a surface-induced dissociation (SID) step in the native MS workflow revealed connectivities within mixed oligomers (Chen *et al.*, 2019). By allowing a large number of designed subunits to associate freely in a single mixture and monitoring oligomerization with MS, it was possible to determine which subunits had complementary interfaces without having to solve their structures (Chen *et al.*, 2019; VanAernum *et al.*, 2020). Lastly, the action of a kinase-activated protein switch could be verified by correlating the mass shift from phosphorylation with protein fluorescence (Woodall *et al.*, 2021).

Common to the above examples is that MS is used to check the intended design features. Another advantage is its ability to interrogate multiple features simultaneously (Figure 1d). In this manner, one feature could be changed in the protein, and its impact on another feature can be determined from the resulting mass spectra. A recent example is the use of MS to assess the impact of single-residue substitutions – which is not always reliably predicted by ML (Allison *et al.*, 2022; Buel and Walters, 2022)- on protein oligomerization efficiency (Figure 1d, panel i). By introducing mutations into a designed scaffold and measuring their impact on protein unfolding or dissociation, it is possible to design pairwise interactions that improve stability (Figure 1d, panel i) (Liu *et al.*, 2025). We employed native MS to guide the design of a lipid-binding site in an artificial membrane protein. Here, lipid ligands were added to different scaffold designs, binding was monitored with exact mass measurements, and the impact of the bound lipids on complex stability was determined through gas-phase dissociation (Figure 1d, panel ii). From these experiments, we could delineate that local protein flexibility and the coordination of the lipid give rise to lipid-mediated stabilization of membrane proteins(Abramsson *et al.*, 2025). It is likely that even more

https://doi.org/10.1017/qrd.2025.10018 Published online by Cambridge University Press

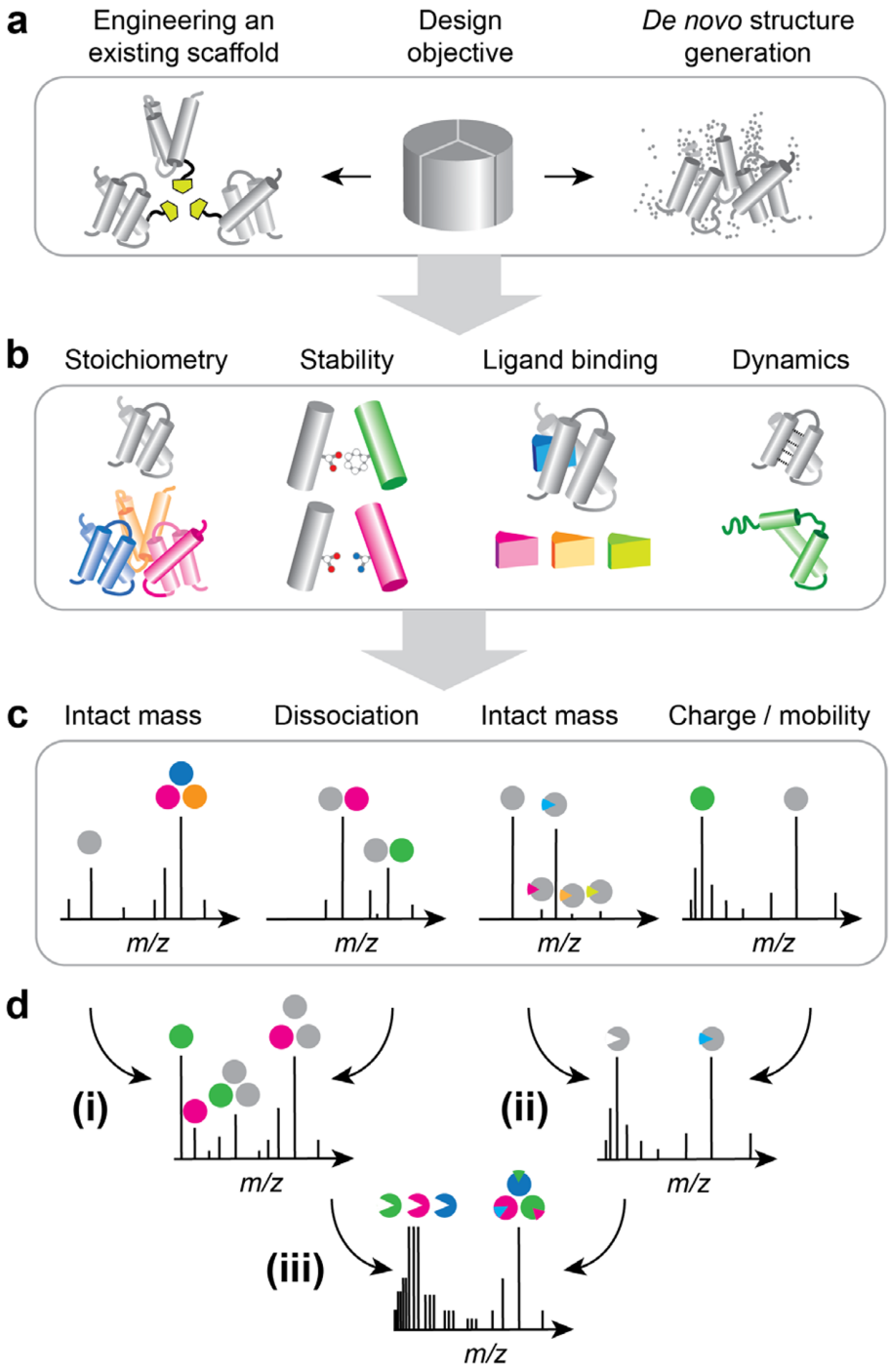


Figure 1. Applications of native MS to verify protein design objectives. (a) A design objective (e.g. trimer) is implemented through modification of an existing protein scaffold and confirmed with ML-based structure prediction, or through de novo generation of a structure using generative models. (b) Examples of design objectives that can be probed by MS. (c) Schematic mass spectra showing the manifestation of each feature in (b). (d) Multiple design objectives can be observed in individual mass spectra, such as the impact of mutations on the stability of oligomers (i), ligand-induced changes in protein flexibility (ii), and ligand-induced changes in oligomerization efficiency (iii).

complex combinations of design objectives can be observed by MS, connecting, for example, ligand-controlled changes in flexibility with self-assembly and stability (Figure 1d, panel iii).

Testing the limits of native MS with engineered proteins

The application of MS to corroborate designs of folded proteins is relatively straightforward, mostly because the abilities and limits of MS have been validated extensively for such systems. Yet some proteins have unusual or complex architectures that test the limits of the approach. Prominent examples are proteins that undergo LLPS, that is, self-assembly into dynamic structures such as stress granules. The underlying interactions are often heterogeneous and can involve an uneven distribution of charged residues, long disordered segments, and multivalent contacts (Wang *et al.*, 2018). MS analysis of these proteins raises multiple questions (Sahin *et al.*, 2023): Can we capture

https://doi.org/10.1017/qrd.2025.10018 Published online by Cambridge University Press

the presence of multiple folded and/or disordered domains in the same protein? Do extreme solution charges and charge distributions affect ionization? And can MS detect low-affinity interactions between IDRs? Several studies using protein engineering have shed light on how proteins with complex charges, architectures, and interactions can be studied with native MS (Figure 2a).

Multi-domain proteins. To find out whether MS can detect the presence of multiple folded domains, Ruotolo and colleagues engineered covalently linked chains of single-domain proteins (Zhong *et al.*, 2014). The resulting chains, containing one to four autonomously folded units, were subjected to solution and gas-phase unfolding and analyzed with IM-MS. Strikingly, the number of unfolding transitions correlated with the number of folded domains, regardless of the proteins that were used as building blocks. These data show that separately folded domains in single protein chains are retained in the gas phase, which suggests that complex protein architectures are amenable to MS analysis (Figure 2b).

Extreme solution charges. Native MS relies on ionization of proteins in their native state. For folded proteins, the resulting number of charges scales with their surface area, and charges are carried by ionizable side-chains (Kaltashov and Mohimen, 2005; Kebarle and Verkerk, 2009). Charge states thus provide structural insights (Hall and Robinson, 2012; Chingin and Barylyuk, 2018). However, folded domains in phase-separating proteins can contain extreme or uneven distributions of charged residues (Kapelner and Obermeyer, 2019; Kim *et al.*, 2024). To test whether the number and location of solution charge affects gas-phase structure, we engineered protein variants with positive and negative charges, and scrambled or partitioned their distributions on the surface. We found that surface area dictates electrospray charge in all cases, but partitioning the charge locations destabilizes the protein ions (Figure 2c). These results indicate that the structures and interactions of proteins with uneven or extreme surface charges can be investigated with native MS, provided that Coulombic effects are taken into account (Abramsson *et al.*, 2021).

Charge and disorder. Interactions involving highly charged IDRs are a common feature of LLPS (Liao *et al.*, 2024), raising the question of how they are impacted by the electrospray charging process. Barran and co-workers explored the role of charge patterning on IDR conformations in solution and in the gas phase by varying the locations of ionizable residues in an intrinsically disordered protein. Computational simulations and IM-MS showed that proteins with even charge distributions populate dynamic ensembles, whereas partitioning into positive and negative blocks results in increased intramolecular interactions and compaction (Figure 2d).

Mixed-domain architectures. Proteins that undergo LLPS often combine folded and disordered domains. We therefore expanded the charge engineering approach to multi-domain architectures, and engineered folded proteins with disordered tails with positive, negative, or no charges. ML was used to generate IDR sequences with comparable solution conformations (Janson *et al.*, 2023). In MS, all variants exhibited near-identical charge state distributions, which correlated with the presumed surface areas of the folded domain with compact, intermediate, and extended IDRs (Figure 2e). Although the correlations were not quantitative, the data suggest that native MS can inform about conformational preferences of IDRs in multi-domain proteins regardless of their charge in solution.

Artificial proteins as tools for method development

The rapid evolution of protein design opens the possibility of moving toward proteins with unusual structures and properties. However, such designs, which can combine non-specific and specific interactions, disordered and folded domains, and extreme sequence compositions and lengths, are likely to test the limits of experimental corroboration.

Based on the insights that engineered proteins have provided for native MS applications, we propose protein design as a universal tool for the development of biophysical methods. In short, proteins with specific features are designed to test the limits of a specific method, and comparisons between the predicted and experimental data serve as a gauge for method validity. By gradually changing one protein feature at a time and monitoring the impact on the experimental data, it is possible to assess, for example, the sensitivity of a method toward the feature, calibrate its effect on the experimental data, or detect experimental artefacts, paving the way for efficient fine-tuning strategies. The same iterative process can also be applied to test the limits of prediction confidence by defining the point at which prediction and experiment diverge. Some potential applications are described below:

Defining resolution limits for structural changes and local dynamics. A recent addition to the protein design toolbox is the generation of proteins with defined local and directional flexibility (Viliuga *et al.*, 2025). This development opens avenues for testing the sensitivity of a method toward local conformational changes or stability. For example, IM-MS analysis of a fixed protein scaffold with varying structural flexibility could be used to determine whether local flexibility affects the stability or CCS of proteins in the gas phase (Figure 3a). An outstanding challenge that could be addressed in this way is the interpretation of gas-phase unfolding data, and whether it can be correlated with structural features in solution (Eldrid *et al.*, 2022; Rider *et al.*, 2024). Going forward, designed protein dynamics could be employed to map the impact of specific thermodynamic properties on binding and folding in the gas phase, for example, using variable temperature (vT) ESI (Laganowsky *et al.*, 2022).

Model systems for gas-phase structure determination. Single-molecule structure analysis remains a frontier in biophysics. A promising approach is the use of free electron lasers to create diffraction patterns from individual protein complexes in the gas phase (Neutze *et al.*, 2000), for which (IM)MS is excellently suited to deliver intact or activated mass- and conformation-selected proteins into the X-ray interaction region (Kierspel *et al.*, 2023). This is a powerful concept in its own right, but it requires information about their orientations at the time of exposure. Relative orientations can, in theory, be inferred from the complete data set if the data amount and quality are high enough, but in reality, the algorithms do not always converge to a useful solution. This challenge can potentially be met through electric fields that act on a protein's dipole moment (Marklund *et al.*, 2017; Sinelnikova *et al.*, 2021), although computational simulations highlight the trade-off between field strength, orientation homogeneity, and protein stability (Wollter *et al.*, 2024; Agelii *et al.*, 2025). Additionally, electric fields could be used to control protein unfolding in a precise way (Sinelnikova *et al.*, 2021). Designed proteins with strong dipole moments and stable gas-phase conformations could be used to identify experimental conditions required for structure determination, as they would potentially retain a known structure at high field strengths, yet become sufficiently orientated to generate an interpretable diffraction pattern (Figure 3b).

'Frozen' intermediates for protein folding studies. High-resolution protein structures usually represent the native basin of protein folding. Structures of (un-)folding intermediates are short-lived or unstable and therefore considerably more challenging to capture, leading to limited structural information. Protein design potentially

https://doi.org/10.1017/qrd.2025.10018 Published online by Cambridge University Press

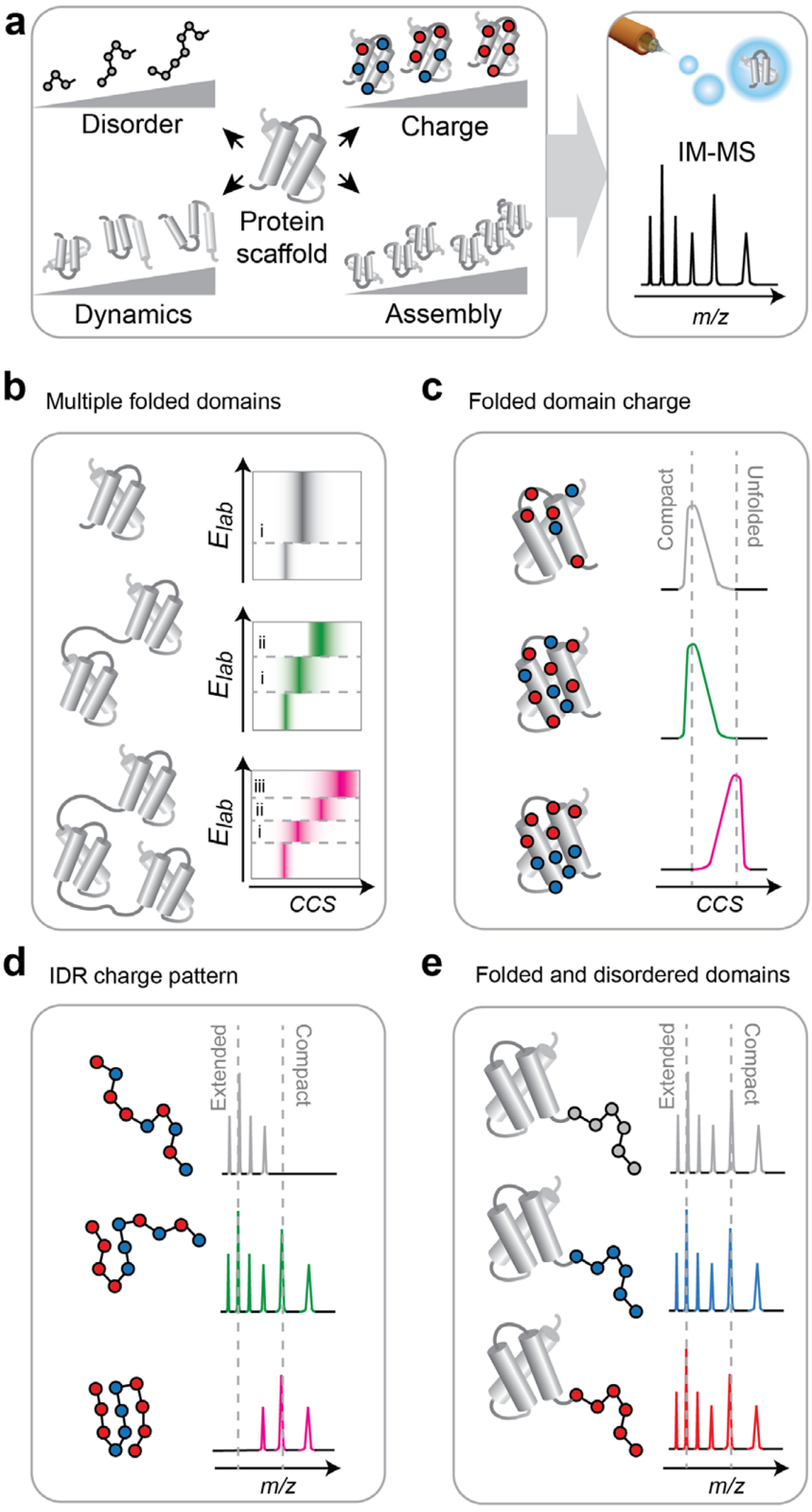


Figure 2. Engineered proteins reveal principles of native MS. (a) Protein engineering as a strategy for method testing. Features of interest, such as charged residues, disordered tails, additional domains, or flexible hinges, are added to a protein scaffold. Their effect on electrospray charge and gas-phase conformation is then assessed with native MS and IM-MS. (b) IM-MS captures the presence of multiple folded domains via their unfolding transitions. (c) Conformational stability in IM-MS is impacted by the location of charged residues. (d) IM-MS reveals how the patterning of charged residues in disordered proteins controls compaction in solution. (e) The presence of folded and disordered regions is reflected in ion charge states in MS regardless of protein solution charge.

https://doi.org/10.1017/qrd.2025.10018 Published online by Cambridge University Press

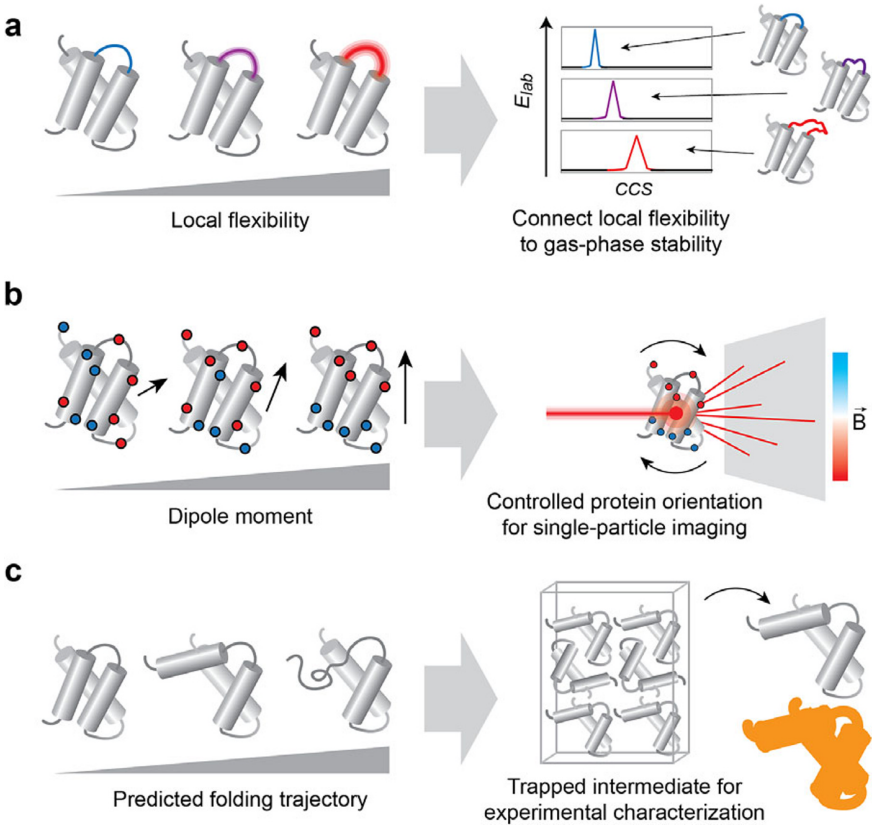


Figure 3. Potential applications of protein design for method development in experimental biophysics. (a) Designed proteins with defined local B-factors can be used to test the effect of local structural fluctuations on protein stability and CCS in IMMS experiments. (b) Designed proteins with defined dipole moments can be used to control gas-phase orientation for single-particle imaging with XFEL. (c) Predicted protein folding intermediates can be used as templates for the design of trapped states for high-resolution structure determination. The high-resolution 'trapped state'-structure can be compared to biophysical or computational data on short-lived intermediates (represented as orange shadow).

offers ways to investigate folding pathways. For example, predicted intermediate states could be used as design templates to create 'frozen' snapshots, or scaffold proteins could be generated to stabilize a specific intermediate for biophysical analysis. These reference data could be employed to corroborate experimental observations or to deconvolute overlapping states (Figure 3c).

Conclusions

The ever-increasing accessibility of protein design tools continues to expand their applications in biomaterial science, biotechnology, and biomedicine. Proteins can be designed for a variety of purposes, but also with previously inaccessible properties relating to their charge, stability, dynamics, size, and shape. As outlined here, their adaptability makes them excellent tools for testing the limits of MS. The fact that the pipeline from design to recombinant material is relatively short and uncomplicated means that proteins can be designed and produced locally in time- and resource-effective ways. They also offer a way to standardize or compare methods across instrument platforms or labs by potentially reducing issues with variability and availability that can be encountered with naturally occurring proteins. Therefore, we believe that proteins that have been engineered or designed with ML can help expand the biophysical toolbox by providing bespoke systems for method development. We anticipate that artificial proteins will emerge as standard reagents for a variety of structural biology applications.

Open peer review. To view the open peer review materials for this article, please visit http://doi.org/10.1017/qrd.2025.10018.

Acknowledgments. The authors thank their colleagues for the helpful discussions.

Author contribution. Conception: AS, AE, EGM, AL, ML. Original draft writing: AS, HO, VV, ML. All authors contributed to the final article.

Financial support. ML is supported by a KI faculty-funded Career Position, a Cancerfonden Project grant (22–2023 Pj), a Swedish Research Council (VR) Project Grant (2024–04483), and a Consolidator Grant from the Swedish Society for Medical Research (SSMF). AE is supported by a KAW Project grant (2022.0032). EGM is supported by a Swedish Research Council (VR) Project Grant (2020–04825) and a Röntgen-Ångström Cluster Grant (2021–05988).

Competing interests. The authors have no conflicts of interest to declare.

References

Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, Willmore L, Ballard AJ, Bambrick J, Bodenstein SW, Evans DA, Hung C-C, O'Neill M, Reiman D, Tunyasuvunakool K, Wu Z, Žemgulytė A, Arvaniti E, Beattie C, Bertolli O, Bridgland A, Cherepanov A, Congreve M, Cowen-Rivers AI, Cowie A, Figurnov M, Fuchs FB, Gladman H, Jain R, Khan YA, Low CMR, Perlin K, Potapenko A, Savy P, Singh S, Stecula A, Thillaisundaram A, Tong C, Yakneen S, Zhong ED, Zielinski M, Žídek A, Bapst V, Kohli P, Jaderberg M, Hassabis D and Jumper JM (2024) Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500.

Abramsson ML, Corey RA, Skerle JL, Persson LJ, Anden O, Oluwole AO, Howard RJ, Lindahl E, Robinson CV, Strisovsky K, Marklund EG, Drew D, Stansfeld PJ and Landreh M (2025) Engineering cardiolipin binding to an artificial membrane protein reveals determinants for lipid-mediated stabilization. *eLife* **14**, RP104237.

Abramsson ML, Sahin C, Hopper JTS, Branca RMM, Danielsson J, Xu M, Chandler SA, Österlund N, Ilag LL, Leppert A, Costeira-Paulo J, Lang L, Teilum K, Laganowsky A, Benesch JLP, Oliveberg M, Robinson CV, Marklund EG, Allison TM, Winther JR and Landreh M (2021) Charge engineering reveals the roles of Ionizable side chains in electrospray ionization mass spectrometry. *JACS Au* **1**, 2385–2393.

Agelii H, Jakobsson ELS, De Santis E, Elfrink G, Mandl T, Marklund EG and Caleman C (2025) Dipole orientation of hydrated gas phase proteins. *Physical Chemistry Chemical Physics PCCP* **27**, 10939–10948.

Allison TM, Degiacomi MT, Marklund EG, Jovine L, Elofsson A, Benesch JLP and Landreh M (2022) Complementing machine learning-based structure predictions with native mass spectrometry. *Protein Science : A Publication of the Protein Society* **31**, e4333.

Baek M, Anishchenko I, Humphreys IR, Cong Q, Baker D & DiMaio F (2023) Efficient and accurate prediction of protein structure using RoseTTAFold2. 2023.05.24.542179. Available at https://doi.org/10.1101/2023.05.24.542179v1 (accessed 23 May 2025).

Baek M, McHugh R, Anishchenko I, Jiang H, Baker D and DiMaio F (2024) Accurate prediction of protein–nucleic acid complexes using RoseTTAFoldNA. *Nature Methods* **21**, 117–121.

Benesch JLP, Ruotolo BT, Simmons DA and Robinson CV (2007) Protein complexes in the gas phase: Technology for Structural Genomics and Proteomics. *Chemical Reviews* **107**, 3544–3567.

Breuker K and McLafferty FW (2008) Stepwise evolution of protein native structure with electrospray into the gas phase, 10(−12) to 10(2) s. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 18145–18152.

Buel GR and Walters KJ (2022) Can AlphaFold2 predict the impact of missense mutations on structure? *Nature Structural & Molecular Biology* **29**, 1–2.

von Bülow S, Tesei G, Zaidi FK, Mittag T and Lindorff-Larsen K (2025) Prediction of phase-separation propensities of disordered proteins from sequence. *Proceedings of the National Academy of Sciences of the United States of America* **122**, e2417920122.

Busenlehner LS and Armstrong RN (2005) Insights into enzyme structure and dynamics elucidated by amide H/D exchange mass spectrometry. *Archives of Biochemistry and Biophysics* **433**, 34–46.

Butcher J, Krishna R, Mitra R, Brent RI, Li Y, Corley N, Kim P, Funk J, Mathis S, Salike S, Muraishi A, Eisenach H, Thompson TR, Chen J, Politanska Y, Sehgal E, Coventry B, Zhang O, Qiang B, Didi K, Kazman M, DiMaio F & Baker D (2025) De novo Design of All-atom Biomolecular Interactions with RFdiffusion3. 2025.09.18.676967. Available at https://doi.org/10.1101/2025.09.18.676967v1 (accessed 23 September 2025).

Chakravarty D, Lee M and Porter LL (2025) Proteins with alternative folds reveal blind spots in AlphaFold-based protein structure prediction. *Current Opinion in Structural Biology* **90**, 102973.

Chen Z, Boyken SE, Jia M, Busch F, Flores-Solis D, Bick MJ, Lu P, VanAernum ZL, Sahasrabuddhe A, Langan RA, Bermeo S, Brunette T, Mulligan VK, Carter LP, DiMaio F, Sgourakis NG, Wysocki VH and Baker D (2019) Programmable design of orthogonal protein heterodimers. *Nature* **565**, 106–111.

Chingin K and Barylyuk K (2018) Charge-state-dependent variation of signal intensity ratio between unbound protein and protein-ligand complex in electrospray ionization mass spectrometry: The role of solvent-accessible surface area. *Analytical Chemistry* **90**, 5521–5528.

Christofi E and Barran P (2023) Ion mobility mass spectrometry (IM-MS) for structural biology: Insights gained by measuring mass, charge, and collision cross section. *Chemical Reviews* **123**, 2902–2949.

Chu AE, Lu T and Huang P-S (2024) Sparks of function by de novo protein design. *Nature Biotechnology* **42**, 203–215.

Cropley TC, Liu FC, Chai M, Bush MF and Bleiholder C (2024) Metastability of protein solution structures in the absence of a solvent: Rugged energy landscape and glass-like behavior. *Journal of the American Chemical Society* **146**, 11115–11125.

Eldrid C, Cragnolini T, Ben-Younis A, Zou J, Raleigh DP and Thalassinos K (2022) Linking gas-phase and solution-phase protein unfolding via Mobile proton simulations. *Analytical Chemistry* **94**, 16113–16121.

Frank C, Khoshouei A, Fuβ L, Schiwietz D, Putz D, Weber L, Zhao Z, Hattori M, Feng S, de Stigter Y, Ovchinnikov S and Dietz H (2024) Scalable protein design using optimization in a relaxed sequence space. *Science* **386**, 439–445.

Goverde CA, Wolf B, Khakzad H, Rosset S and Correia BE (2023) De novo protein design by inversion of the AlphaFold structure prediction network. *Protein Science* **32**, e4653.

Guo AB, Akpinaroglu D, Stephens CA, Grabe M, Smith CA, Kelly MJS and Kortemme T (2025) Deep learning–guided design of dynamic proteins. *Science* **388**, eadr7094.

Hall Z and Robinson CV (2012) Do charge state signatures guarantee protein conformations? *Journal of the American Society for Mass Spectrometry* **23**, 1161–1168.

Hansen K, Lau AM, Giles K, McDonnell JM, Struwe WB, Sutton BJ and Politis A (2018) A mass-spectrometry-based modelling workflow for accurate prediction of IgG antibody conformations in the gas phase. *Angewandte Chemie (International Ed. in English)* **57**, 17194–17199.

Janson G, Valdes-Garcia G, Heo L and Feig M (2023) Direct generation of protein conformational ensembles via machine learning. *Nature Communications* **14**, 774.

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P and Hassabis D (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589.

Kalakoti Y and Wallner B (2025) AFsample2 predicts multiple conformations and ensembles with AlphaFold2. *Communications Biology* **8**, 373.

Kaltashov IA and Mohimen A (2005) Estimates of protein surface areas in solution by electrospray ionization mass spectrometry. *Analytical Chemistry* **77**, 5370–5379.

Kapelner RA and Obermeyer AC (2019) Ionic polypeptide tags for protein phase separation. *Chemical Science* **10**, 2700–2707.

Kebarle P and Verkerk UH (2009) Electrospray: From ions in solution to ions in the gas phase, what we know now. *Mass Spectrometry Reviews* **28**, 898–917.

Khmelinskaia A, Bethel NP, Fatehi F, Mallik BB, Antanasijevic A, Borst AJ, Lai S-H, Chim HY, John Wang JY, Miranda MC, Watkins AM, Ogohara C, Caldwell S, Wu M, Heck AJR, Veesler D, Ward AB, Baker D, Twarock R and King NP (2025) Local structural flexibility drives oligomorphism in computationally designed protein assemblies. *Nature Structural & Molecular Biology* **32**, 1050–1060.

Kierspel T, Kadek A, Barran P, Bellina B, Bijedic A, Brodmerkel MN, Commandeur J, Caleman C, Damjanović T, Dawod I, De Santis E, Lekkas A, Lorenzen K, Morillo LL, Mandl T, Marklund EG, Papanastasiou D, Ramakers LAI, Schweikhard L, Simke F, Sinelnikova A, Smyrnakis A, Timneanu N, Uetrecht C and **for the MS SPIDOC Consortium** (2023) Coherent diffractive imaging of proteins and viral capsids: Simulating MS SPIDOC. *Analytical and Bioanalytical Chemistry* **415**, 4209–4220.

Kim J, Qin S, H-X Z and Rosen MK (2024) Surface charge can modulate phase separation of multidomain proteins. *Journal of the American Chemical Society* **146**, 3383–3395.

Konermann L, Ahadi E, Rodriguez AD and Vahidi S (2013) Unraveling the mechanism of electrospray ionization. *Analytical Chemistry* **85**, 2–9.

Kortemme T (2024) De novo protein design-from new structures to programmable functions. *Cell* **187**, 526–544.

Krishna R, Wang J, Ahern W, Sturmfels P, Venkatesh P, Kalvet I, Lee GR, Morey-Burrows FS, Anishchenko I, Humphreys IR, McHugh R, Vafeados D, Li X, Sutherland GA, Hitchcock A, Hunter CN, Kang A, Brackenbrough E, Bera AK, Baek M, DiMaio F and Baker D (2024) Generalized biomolecular modeling and design with RoseTTAFold all-atom. *Science* **384**, eadl2528.

Laganowsky A, Clemmer DE and Russell DH (2022) Variable-temperature native mass spectrometry for studies of protein folding, stabilities, assembly, and molecular interactions. *Annual Review of Biophysics* **51**, 63–77.

https://doi.org/10.1017/qrd.2025.10018 Published online by Cambridge University Press

Liao J, Yeong V and Obermeyer AC (2024) Charge-patterned disordered peptides tune intracellular phase separation in bacteria. *ACS Synthetic Biology* **13**, 598–612.

Liu C, Wu K, Choi H, Han HL, Zhang X, Watson JL, Ahn G, Zhang JZ, Shijo S, Good LL, Fischer CM, Bera AK, Kang A, Brackenbrough E, Coventry B, Hick DR, Qamar S, Li X, Decarreau J, Gerben SR, Yang W, Goreshnik I, Vafeados D, Wang X, Lamb M, Murray A, Kenny S, Bauer MS, Hoofnagle AN, Zhu P, Knowles TPJ and Baker D (2025) Diffusing protein binders to intrinsically disordered proteins. *Nature* **644**, 809–817.

Liu X, Roberts DS, Bingman CA, Ge Y and Gellman SH (2025) Rational Design of a Foldon-Derived Heterotrimer Guided by quantitative native mass spectrometry. *Journal of the American Chemical Society* **147**, 23819–23826.

Lössl P, van de Waterbeemd M and Heck AJ (2016) The diverse and expanding role of mass spectrometry in structural and molecular biology. *The EMBO Journal* **35**, 2634–2657.

Marklund EG and Benesch JL (2019) Weighing-up protein dynamics: The combination of native mass spectrometry and molecular dynamics simulations. *Current Opinion in Structural Biology* **54**, 50–58.

Marklund EG, Ekeberg T, Moog M, Benesch JLP and Caleman C (2017) Controlling protein orientation in vacuum using electric fields. *Journal of Physical Chemistry Letters* **8**, 4540–4544.

Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S and Steinegger M (2022) ColabFold: Making protein folding accessible to all. *Nature Methods* **19**, 679–682.

Mohanty P, Kapoor U, Sundaravadivelu Devarajan D, Phan TM, Rizuan A and Mittal J (2022) Principles governing the phase separation of multidomain proteins. *Biochemistry* **61**, 2443–2455.

Monteiro da Silva G, Cui JY, Dalgarno DC, Lisi GP and Rubenstein BM (2024) High-throughput prediction of protein conformational distributions with subsampled AlphaFold2. *Nature Communications* **15**, 2464.

Neutze R, Wouts R, van der Spoel D, Weckert E and Hajdu J (2000) Potential for biomolecular imaging with femtosecond X-ray pulses. *Nature* **406**, 752–757.

Nikolaev A, Kuzmin A, Markeeva E, Kuznetsova E, Ryzhykau YL, Semenov O, Anuchina A, Remeeva A and Gushchin I (2024) Reengineering of a flavin-binding fluorescent protein using ProteinMPNN. *Protein Science* **33**, e4958.

Pacesa M, Nickel L, Schellhaas C, Schmidt J, Pyatova E, Kissling L, Barendse P, Choudhury J, Kapoor S, Alcaraz-Serna A, Cho Y, Ghamary KH, Vinué L, Yachnin BJ, Wollacott AM, Buckley S, Westphal AH, Lindhoud S, Georgeon S, Goverde CA, Hatzopoulos GN, Gönczy P, Muller YD, Schwank G, Swarts DC, Vecchio AJ, Schneider BL, Ovchinnikov S & Correia BE (2024) BindCraft: one-shot design of functional protein binders. Available at https://doi.org/10.1101/2024.09.30.615802 (accessed 26 August 2025).

Pesce F, Bremer A, Tesei G, Hopkins JB, Grace CR, Mittag T and Lindorff-Larsen K (2024) Design of intrinsically disordered protein variants with diverse structural properties. *Science Advances* **10**, eadm9926.

Praetorius F, Leung PJY, Tessmer MH, Broerman A, Demakis C, Dishman AF, Pillai A, Idris A, Juergens D, Dauparas J, Li X, Levine PM, Lamb M, Ballard RK, Gerben SR, Nguyen H, Kang A, Sankaran B, Bera AK, Volkman BF, Nivala J, Stoll S and Baker D (2023) Design of stimulus-responsive two-state hinge proteins. *Science* **381**, 754–760.

Rider RL, Lantz C, Fan L and Russell DH (2024) Structure and stabilities of solution and gas phase protein complexes. *Journal of the American Society for Mass Spectrometry* **35**, 3028–3036.

Rolland AD, Biberic LS and Prell JS (2022) Investigation of charge-state-dependent compaction of protein ions with native ion mobility-mass spectrometry and theory. *Journal of the American Society for Mass Spectrometry* **33**, 369–381.

Sahin C, Leppert A and Landreh M (2023) Advances in mass spectrometry to unravel the structure and function of protein condensates. *Nature Protocols* **18**, 3653–3661.

Sinelnikova A, Mandl T, Agelii H, Grånäs O, Marklund EG, Caleman C and De Santis E (2021) Protein orientation in time-dependent electric fields: Orientation before destruction. *Biophysical Journal* **120**, 3709–3717.

Sinelnikova A, Mandl T, Östlin C, Grånäs O, Brodmerkel MN, Marklund EG and Caleman C (2021) Reproducibility in the unfolding process of protein induced by an external electric field. *Chemical Science* **12**, 2030–2038.

Stahl K, Warneke R, Demann L, Bremenkamp R, Hormes B, Brock O, Stülke J and Rappsilber J (2024) Modelling protein complexes with crosslinking mass spectrometry and deep learning. *Nature Communications* **15**, 7866.

Sumida KH, Núñez-Franco R, Kalvet I, Pellock SJ, Wicky BIM, Milles LF, Dauparas J, Wang J, Kipnis Y, Jameson N, Kang A, De La Cruz J, Sankaran B, Bera AK, Jiménez-Osés G and Baker D (2024) Improving protein expression, stability, and function with ProteinMPNN. *Journal of the American Chemical Society* **146**, 2054–2061.

Tesei G, Trolle AI, Jonsson N, Betz J, Knudsen FE, Pesce F, Johansson KE and Lindorff-Larsen K (2024) Conformational ensembles of the human intrinsically disordered proteome. *Nature* **626**, 897–904.

Turzo SMBA, Seffernick JT, Rolland AD, Donor MT, Heinze S, Prell JS, Wysocki VH and Lindert S (2022) Protein shape sampled by ion mobility mass spectrometry consistently improves protein structure prediction. *Nature Communications* **13**, 4377.

VanAernum ZL, Busch F, Jones BJ, Jia M, Chen Z, Boyken SE, Sahasrabuddhe A, Baker D and Wysocki VH (2020) Rapid online buffer exchange for screening of proteins, protein complexes and cell lysates by native mass spectrometry. *Nature Protocols* **15**, 1132–1157.

Viliuga V, Seute L, Wolf N, Wagner S, Elofsson A, Stühmer J & Gräter F (2025) Flexibility-Conditioned Protein Structure Design with Flow Matching. Available at https://arxiv.org/abs/2508.18211 (accessed 1 September 2025).

Wang J, Choi J-M, Holehouse AS, Lee HO, Zhang X, Jahnel M, Maharana S, Lemaitre R, Pozniakovsky A, Drechsel D, Poser I, Pappu RV, Alberti S and Hyman AA (2018) A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. *Cell* **174**, 688–699, e16.

Wang J, Watson JL and Lisanza SL (2024) Protein design using structure-prediction networks: AlphaFold and RoseTTAFold as protein structure foundation models. *Cold Spring Harbor Perspectives in Biology* **16**, a041472.

Wang L, Tučs A, Ding S, Tsuda K and Sljoka A (2025) HDXRank: A deep learning framework for ranking protein complex predictions with hydrogen-deuterium exchange data. *Journal of Chemical Theory and Computation* **1**, 7173–7187.

Wayment-Steele HK, Ojoawo A, Otten R, Apitz JM, Pitsawong W, Hömberger M, Ovchinnikov S, Colwell L and Kern D (2024) Predicting multiple conformations via sequence clustering and AlphaFold2. *Nature* **625**, 832–839.

Wohlwend J, Corso G, Passaro S, Getz N, Reveiz M, Leidal K, Swiderski W, Atkinson L, Portnoi T, Chinn I, Silterra J, Jaakkola T & Barzilay R (2025) Boltz-1 Democratizing Biomolecular Interaction Modeling., 2024.11.19.624167. Available at https://doi.org/10.1101/2024.11.19.624167v4 (accessed 20 May 2025).

Wollter A, De Santis E, Ekeberg T, Marklund EG and Caleman C (2024) Enhanced EMC-advantages of partially known orientations in x-ray single particle imaging. *The Journal of Chemical Physics* **160**, 114108.

Woodall NB, Weinberg Z, Park J, Busch F, Johnson RS, Feldbauer MJ, Murphy M, Ahlrichs M, Yousif I, MacCoss MJ, Wysocki VH, El-Samad H and Baker D (2021) De novo design of tyrosine and serine kinase-driven protein switches. *Nature Structural & Molecular Biology* **28**, 762–770.

Wu K, Jiang H, Hicks DR, Liu C, Muratspahić E, Ramelot TA, Liu Y, McNally K, Kenny S, Mihut A, Gaur A, Coventry B, Chen W, Bera AK, Kang A, Gerben S, Lamb MY-L, Murray A, Li X, Kennedy MA, Yang W, Song Z, Schober G, Brierley SM, O'Neill J, Gelb MH, Montelione GT, Derivery E and Baker D (2025) Design of intrinsically disordered region binding proteins. *Science* **389**, eadr8063.

Zhong Y, Han L and Ruotolo BT (2014) Collisional and Coulombic unfolding of gas-phase proteins: High correlation to their domain structures in solution. *Angewandte Chemie (International Ed. in English)* **53**, 9209–9212.

https://doi.org/10.1017/qrd.2025.10018 Published online by Cambridge University Press