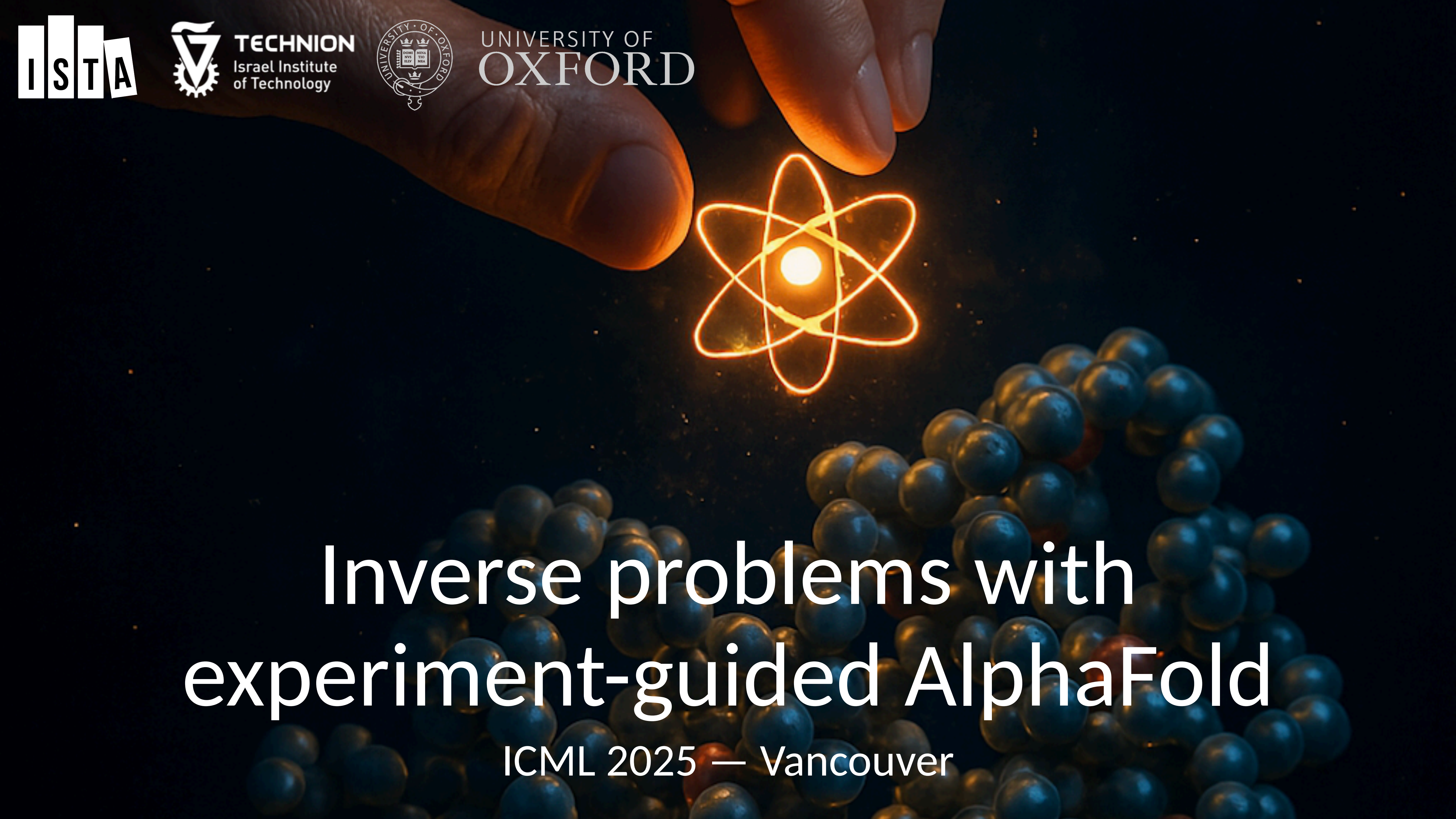
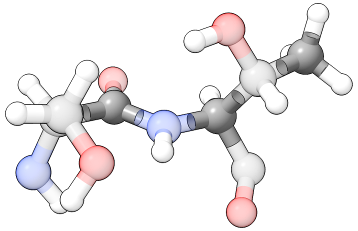


ISTA
TECHNION Israel Institute of Technology
UNIVERSITY OF OXFORD

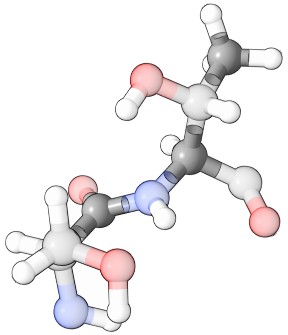
# Inverse problems with experiment-guided AlphaFold

ICML 2025 — Vancouver

# Problem definition

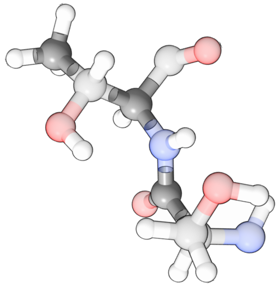


$X_1$



$X_2$

...



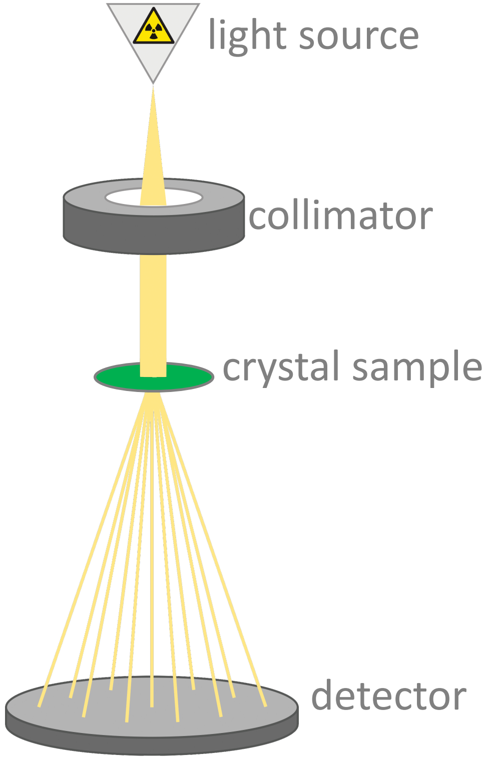
$X_n$

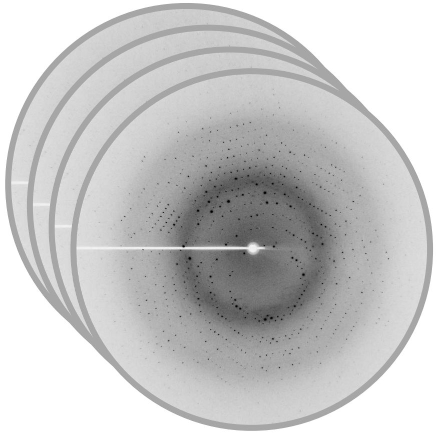
$$X_1, \ldots X_n \sim p(X_1, \ldots X_n | \text{amino acid sequence})$$

$$s.t$$

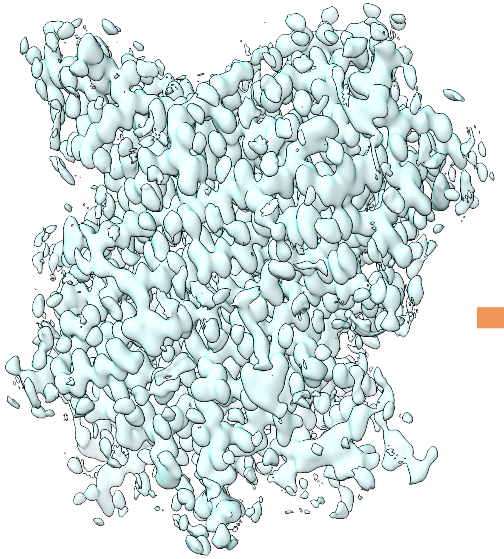
$$X_1, \ldots X_n = \arg\max_{X_1, \ldots X_n} p(\text{experimental observation} | X_1, \ldots X_n)$$

# X-ray crystallography

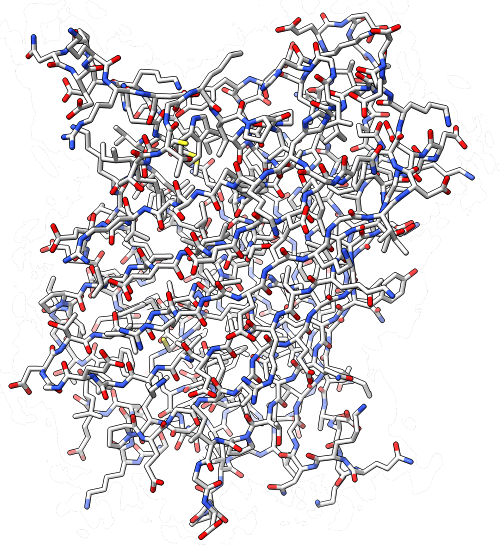


2D slices of $|\mathscr{F}F_o|^2$
discrete on $\mathcal{L}^*$

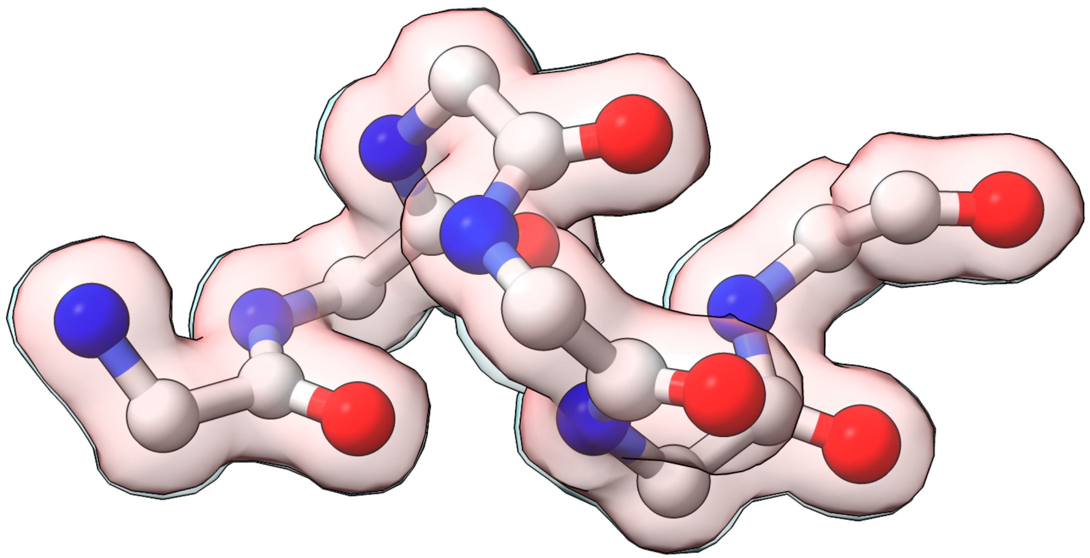


Electron density $F_\mathrm{o}$
periodic on $\mathcal{L}$



fitted molecular
model

# Electron density forward model



Single atom contributes to calculated electron density at point $\boldsymbol{\xi}$

$$\underbrace{\sum_{i=j}^{5} a_{ij} \cdot \left(\frac{4\pi}{b_{ij}+B_j}\right)^{\frac{3}{2}} \exp\left(-\frac{4\pi^2\|\boldsymbol{\xi}-\mathbf{x}_j\|^2}{b_{ij}+B_j}\right)}$$

$$= K(\boldsymbol{\xi}-\mathbf{x}_j; B_j)$$ kernel

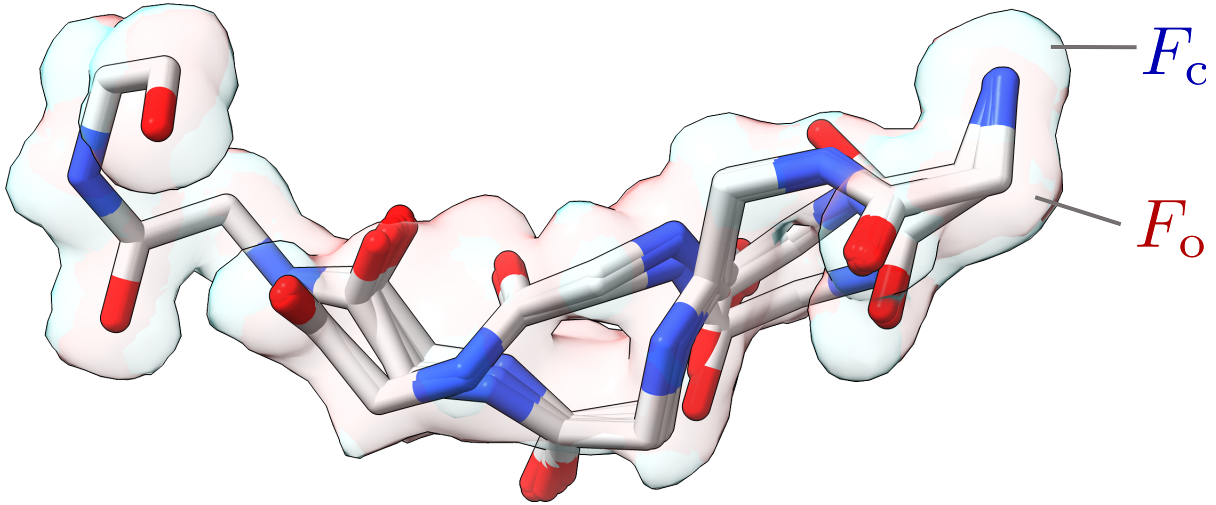
**Output:** calculated electron density

$$F_{\mathrm{c}}(\boldsymbol{\xi}) = \sum_j K(\boldsymbol{\xi}-\mathbf{x}_j; B_j)$$

**Model fitting:** $F_{\mathrm{c}} \approx F_{\mathrm{o}}$

# Likelihood = forward model

Given an ensemble, what is the probability of observed ED?

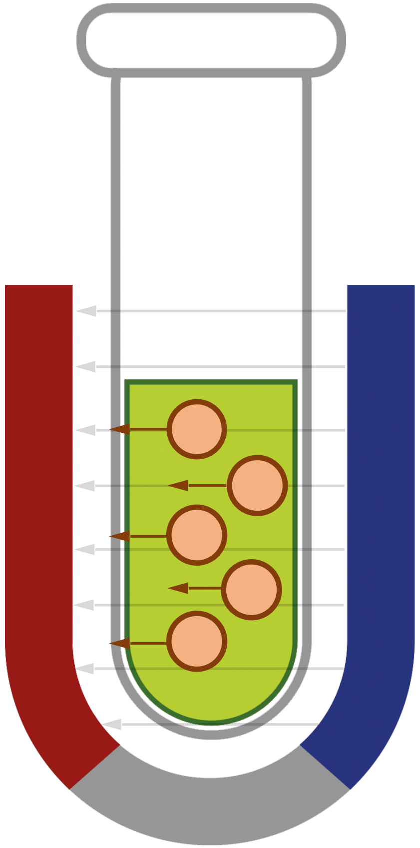


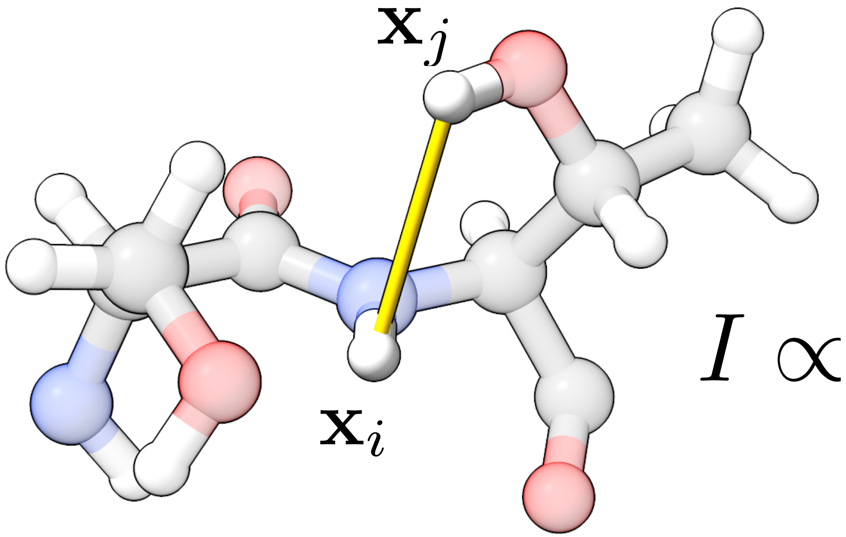

$$\log p(F_o | \underbrace{\mathbf{X}_1, \ldots, \mathbf{X}_n}_{\mathcal{X}}) = - \left\| F_o(\cdot) - \underbrace{\frac{1}{n} \sum_k F_c(\cdot | \mathbf{X}_k)}_{} \right\|^2$$

$$F_c(\cdot | \mathcal{X}) = \sum_j K(\boldsymbol{\xi} - \mathbf{x}_j; B_j)$$

*symmetry operators omitted for clarity

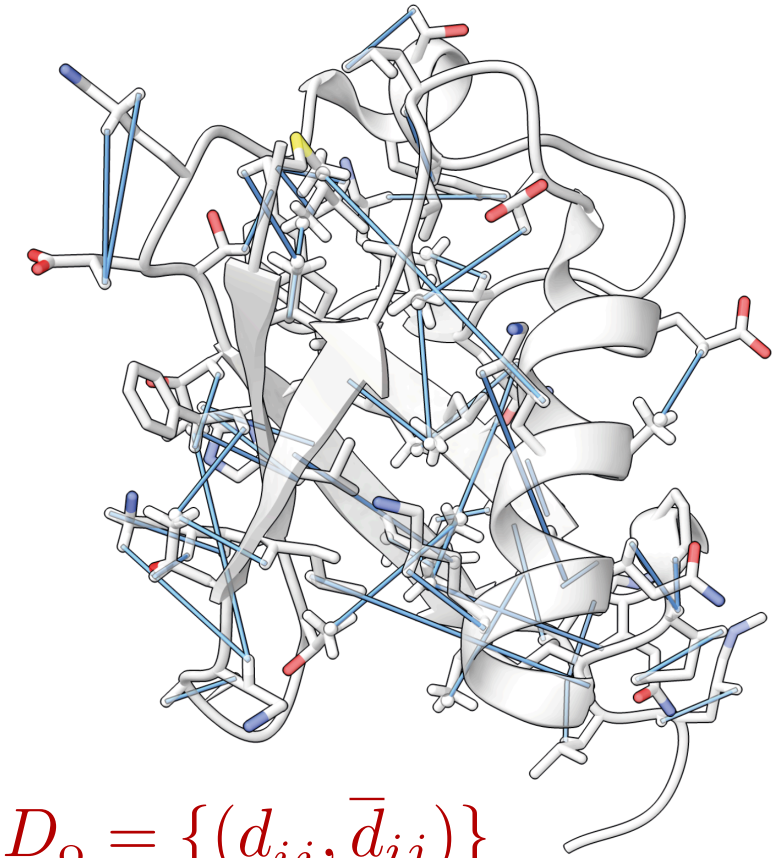
# NMR Nuclear Overhauser effect spectroscopy






$$I \propto \|\mathbf{x}_i - \mathbf{x}_j\|^{-6}$$

# Likelihood = forward model



$D_\text{o} = \{(\underline{d}_{ij}, \overline{d}_{ij})\}$

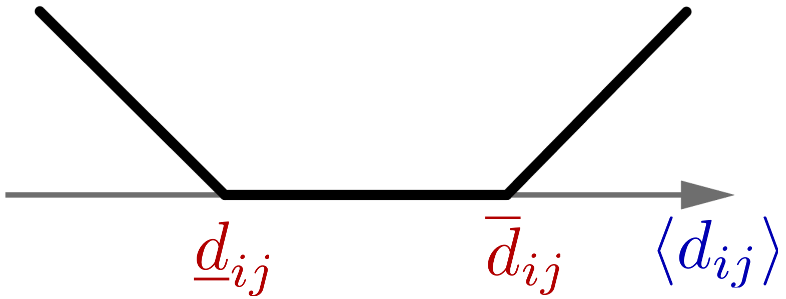
upper bound ($\overline{d}_{ij}$)

lower bound ($\underline{d}_{ij}$)

Given an ensemble, what is the probability of observed distance constraints?

$$\log p(D_\text{o} | \underbrace{\mathbf{X}_1, \ldots, \mathbf{X}_n}_{\mathcal{X}}) =$$

$$- \sum_{ij} \text{viol} \left( \frac{1}{n} \sum_k d_{ij}(\mathbf{X}_k), \underline{d}_{ij}, \overline{d}_{ij} \right)^2$$




Maddipatla, Sellam, Bojan, Vedula, Schanda, Marx, Bronstein. *ICML* **2025**

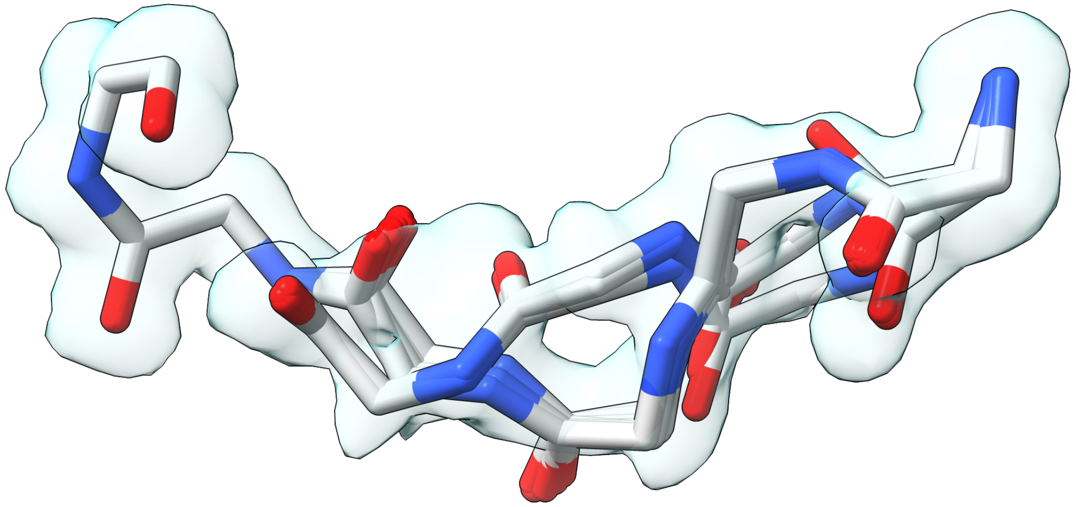
# Sampling from the posterior

Given the sequence and an experimental observation, give me an ensemble

$$\mathcal{X} \sim \underbrace{p(\mathbf{X}_1, \ldots, \mathbf{X}_n | \mathbf{a}, \mathbf{y})}_{\text{posterior}}$$

$$\propto \underbrace{p(\mathbf{y}|\mathbf{X}_1, \ldots, \mathbf{X}_n)}_{\text{likelihood}} \cdot \underbrace{p(\mathbf{X}_1, \ldots, \mathbf{X}_n|\mathbf{a})}_{\text{prior}}$$

$$= \underbrace{p(\mathbf{y}|\mathbf{X}_1, \ldots, \mathbf{X}_n)}_{\text{inseparable}} \cdot \underbrace{p(\mathbf{X}_1|\mathbf{a}) \cdots p(\mathbf{X}_n|\mathbf{a})}_{\text{separable}}$$



# Gluing everything together

$$d \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_n \end{bmatrix} = -\left( \frac{1}{2} \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_n \end{bmatrix} + \begin{bmatrix} \nabla_{\mathbf{X}} \log p_t(\mathbf{X}_1 | \mathbf{a}) \\ \vdots \\ \nabla_{\mathbf{X}} \log p_t(\mathbf{X}_n | \mathbf{a}) \end{bmatrix} \right) \beta_t dt + \sqrt{\beta_t} \begin{bmatrix} \mathbf{N}_1 \\ \vdots \\ \mathbf{N}_n \end{bmatrix}$$

ensemble

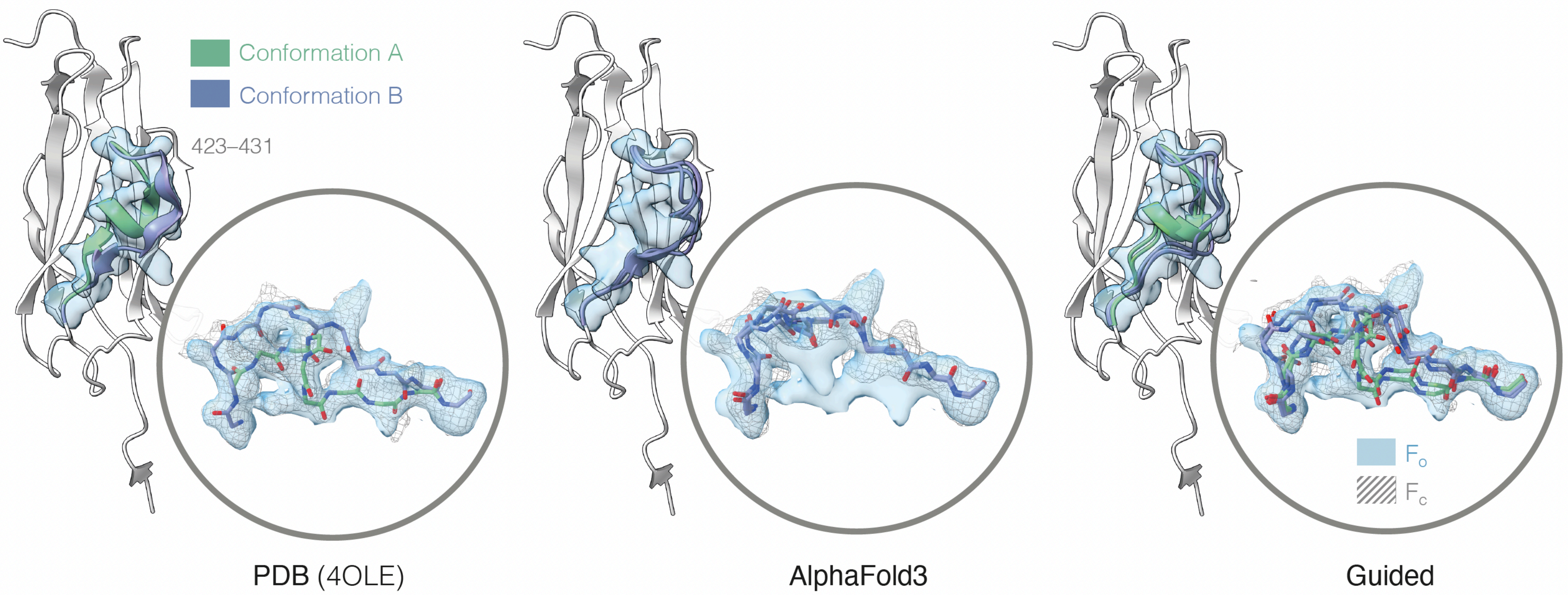
prior

$$+ \nabla_{\mathbf{X}} \log p\left(\mathbf{y} \,|\mathbf{X}_1, \ldots, \mathbf{X}_n\right)$$

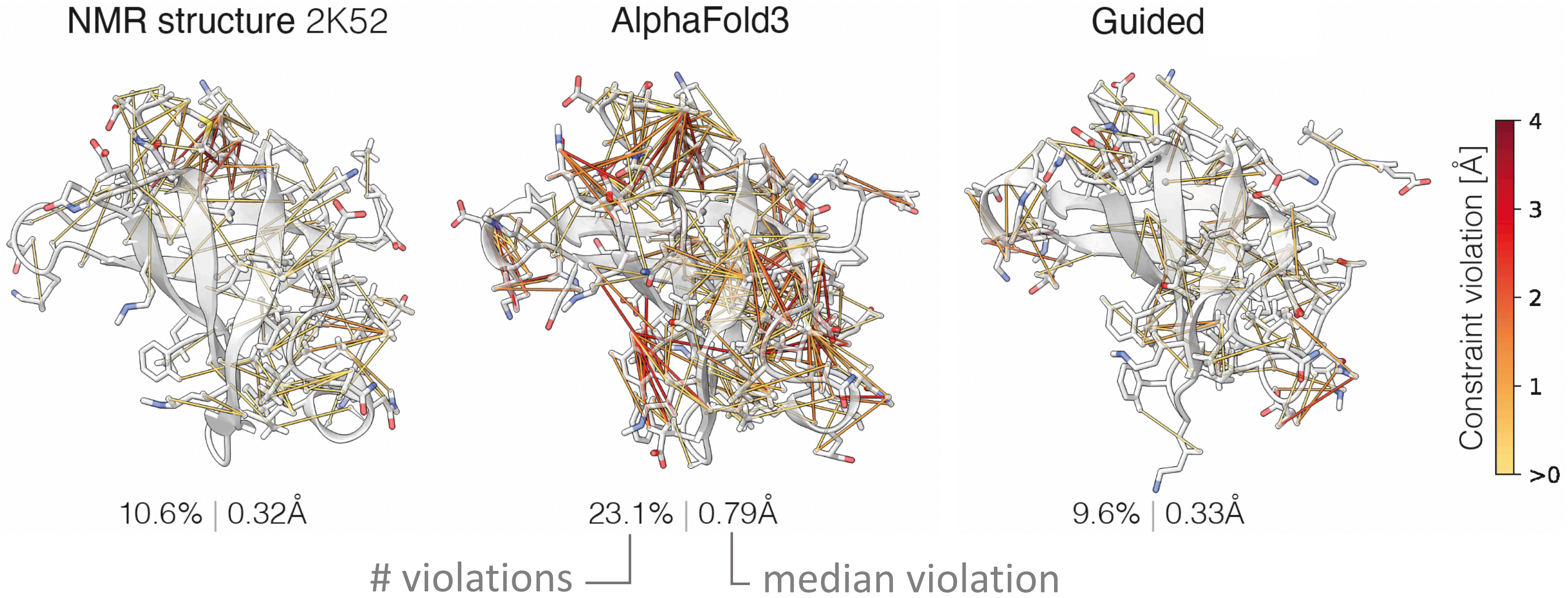
experimental guidance

Maddipatla, Sellam, Vedula, Marx, Bronstein. *NeurIPS MLSB* **2024**

# One crystal, two conformations



# Guiding with NOE constraints



ICML
International Conference
On Machine Learning

# Collaborators

Nadav Bojan Sellam

Meital Bojan

Sanketh Vedula

Ailie Marx

Paul Schanda

Alex Bronstein

Thank you!