ICML
International Conference
On Machine Learning

HITS
Heidelberg Institute for
Theoretical Studies

Stockholm
University

MAX PLANCK INSTITUTE
FOR POLYMER RESEARCH

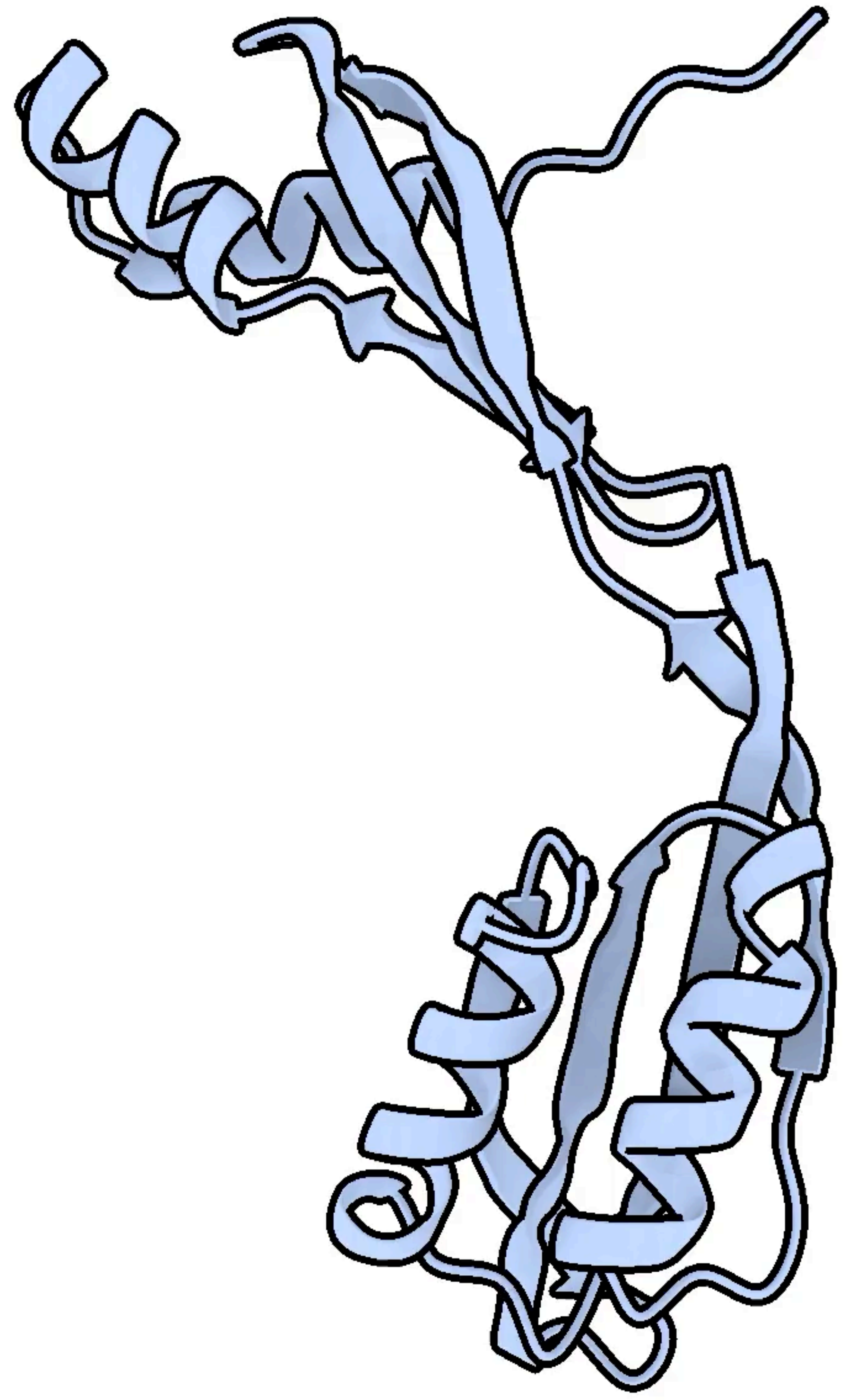
UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

KIT
Karlsruher Institut für Technologie

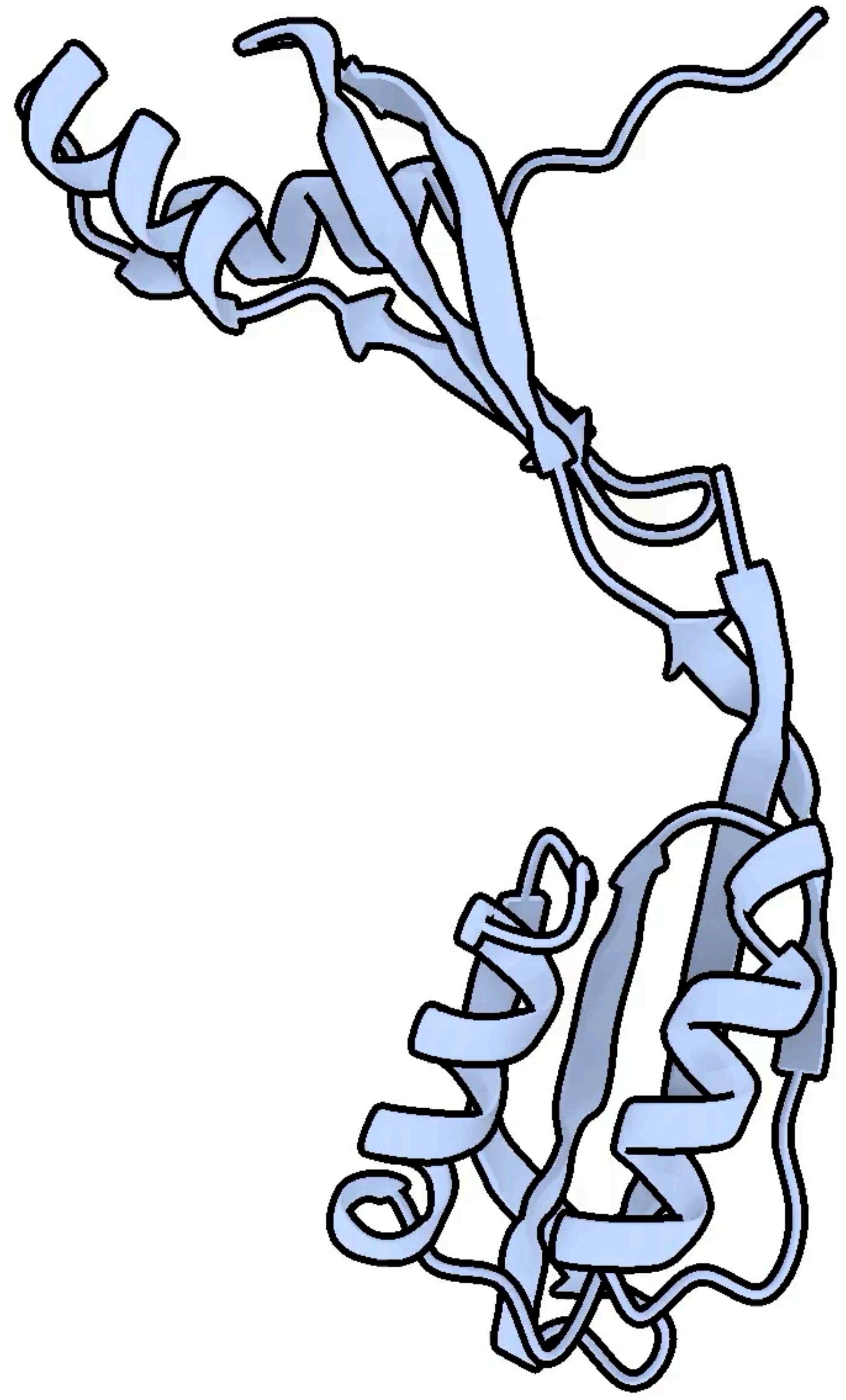
Flexibility-conditioned protein structure design with flow matching

Vsevolod Viliuga*, Leif Seute*, Nicolas Wolf, Simon Wagner, Arne Elofsson
Jan Stühmer, Frauke Gräter

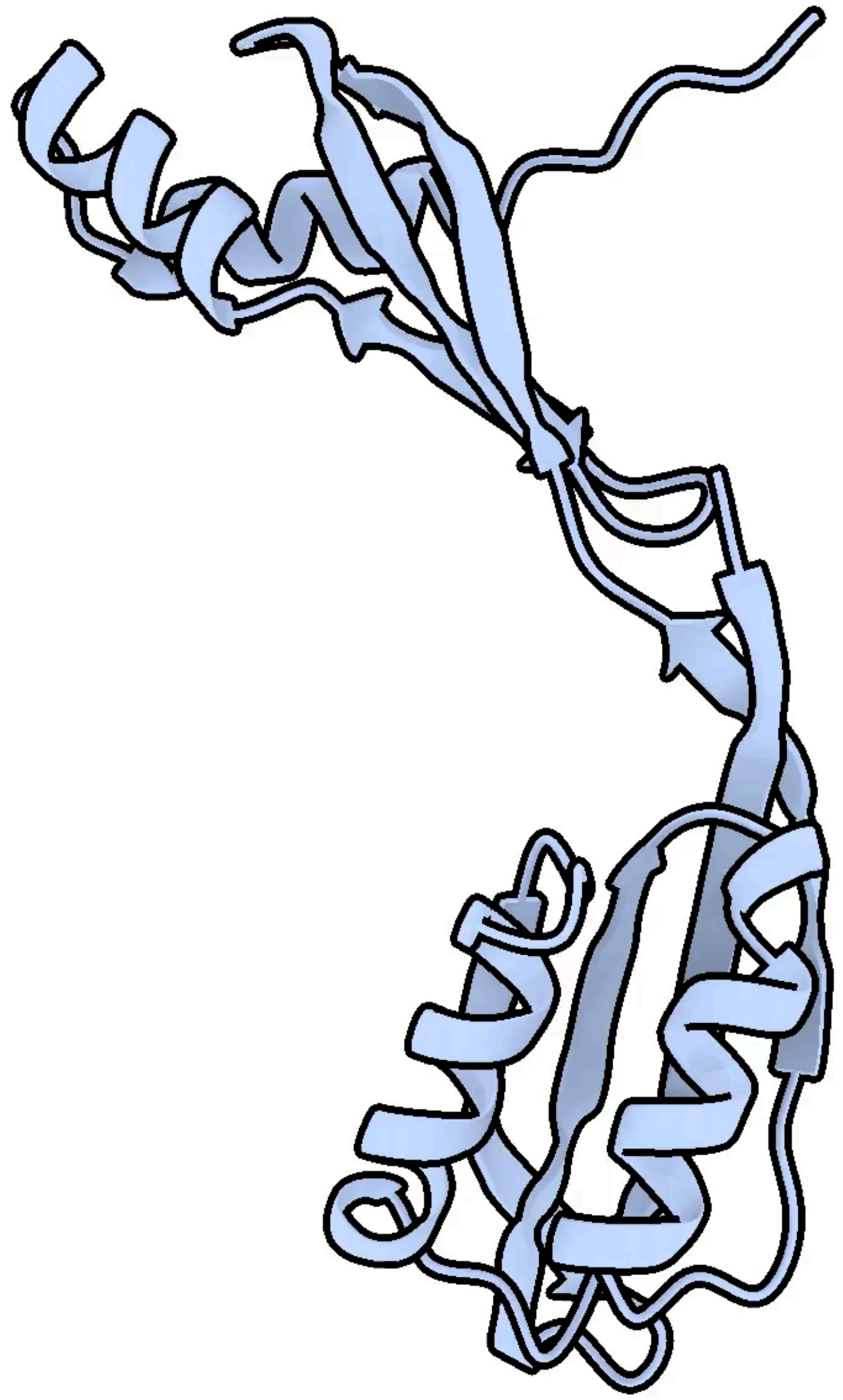
Functional protein structures are dynamic

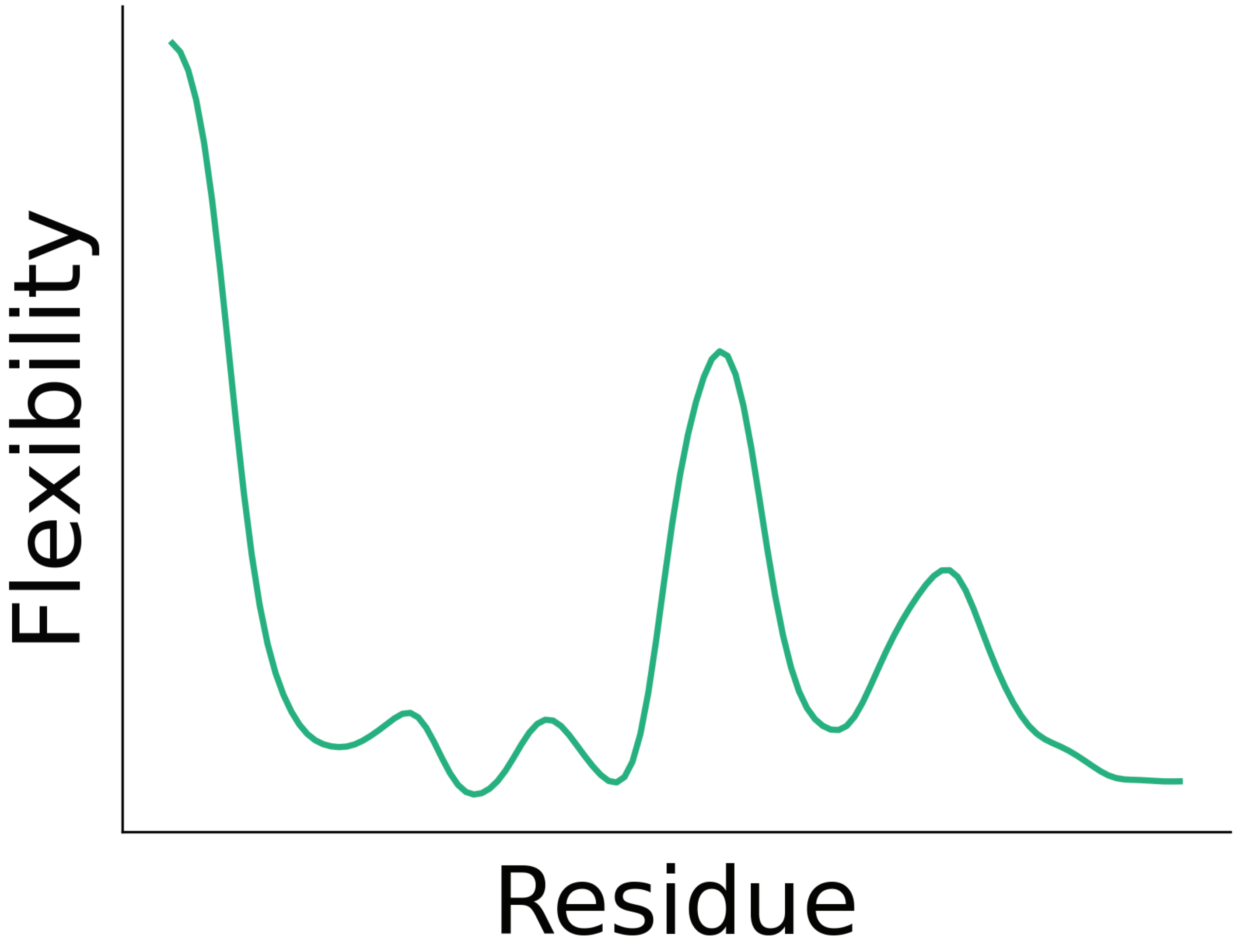


Functional protein structures are dynamic

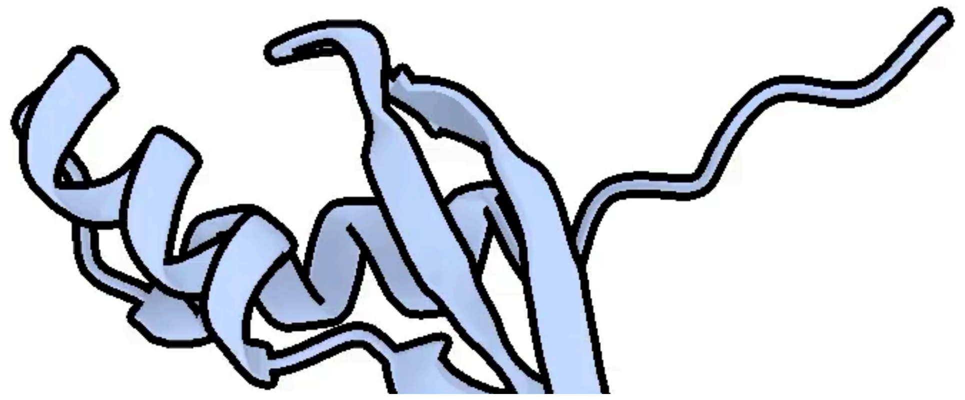


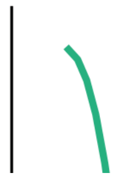
Functional protein structures are dynamic



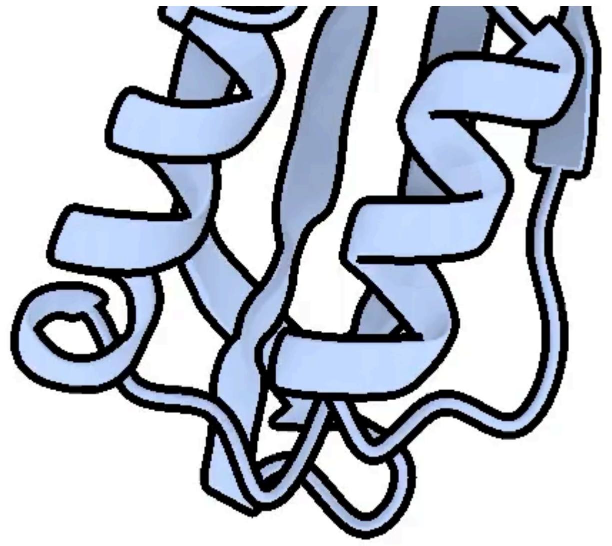


Functional protein structures are dynamic



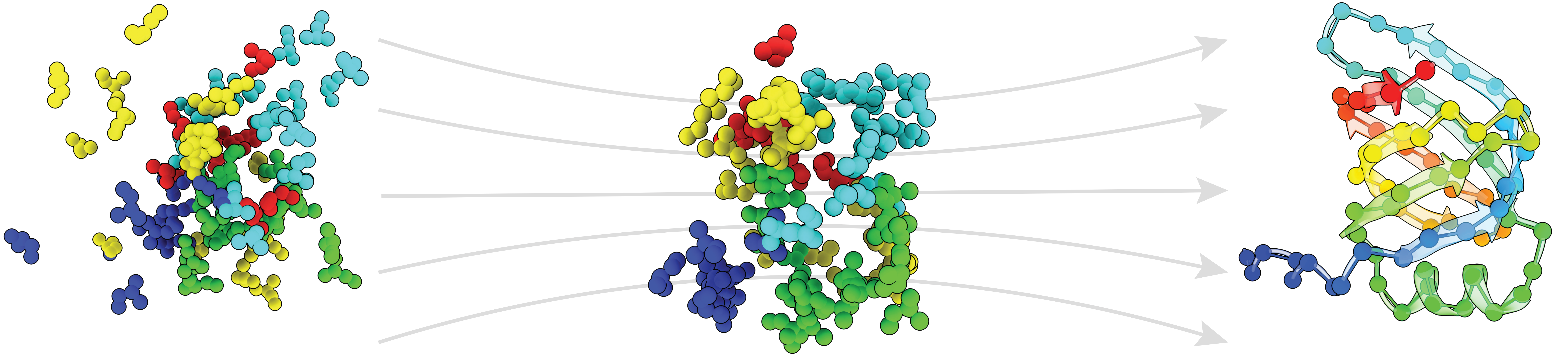


Can we condition protein structure generation on flexibility?



FliPS: Flexibility-conditioned protein structure design

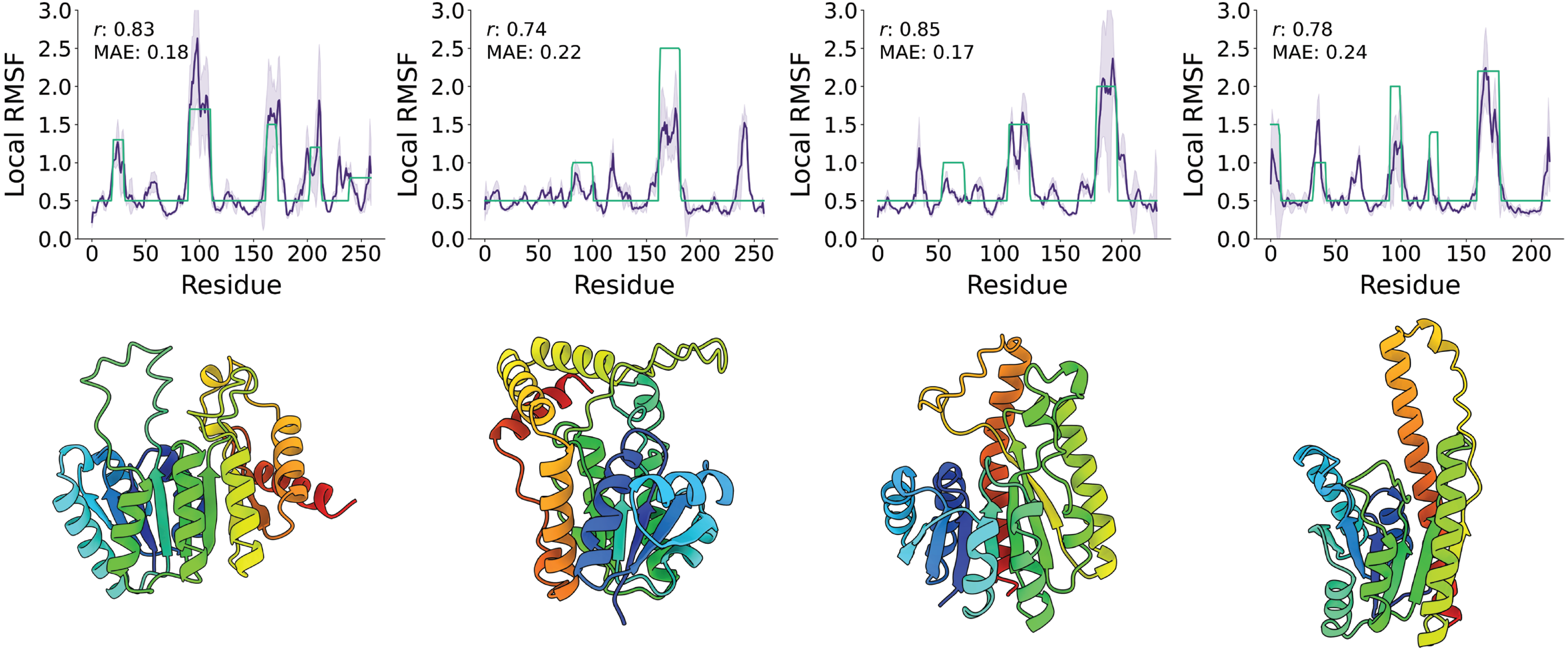
$$\mathbf{T}_{noised} \sim p_0 \qquad \phi(\mathbf{T}, t, \xi) = \mathbf{T} + \int_0^t v(\mathbf{T}, t', \xi)\, dt' \qquad \mathbf{T}_{denoised} \sim p_1$$



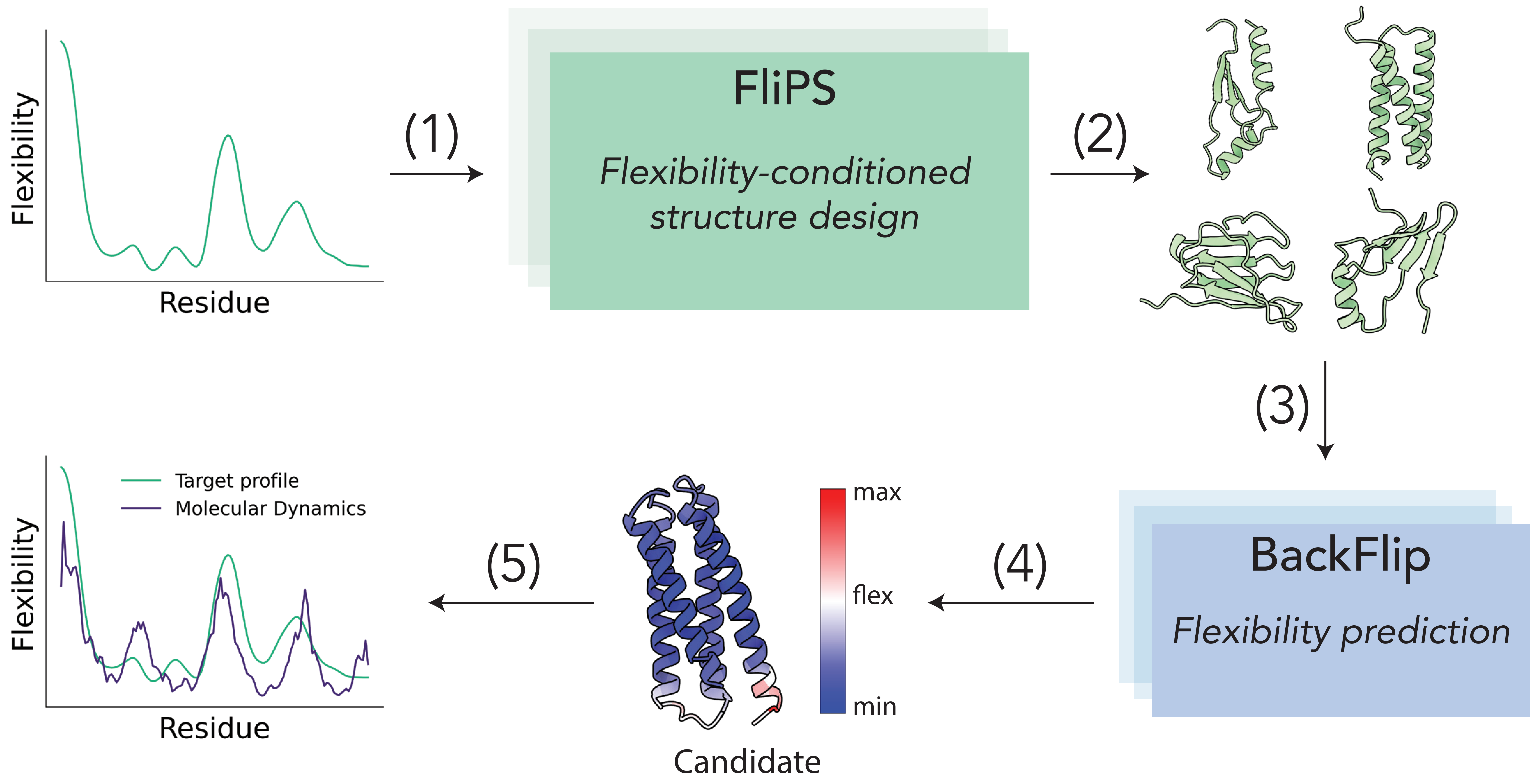
`t = 0` `t = 0.66` `t = 1`

$$\mathbf{T}_i = \{r_i, x_i\} \in \mathrm{SE}(3)$$

Structures generated by FliPS reflect target flexibility



Framework to design and screen protein structures for desired flexibility



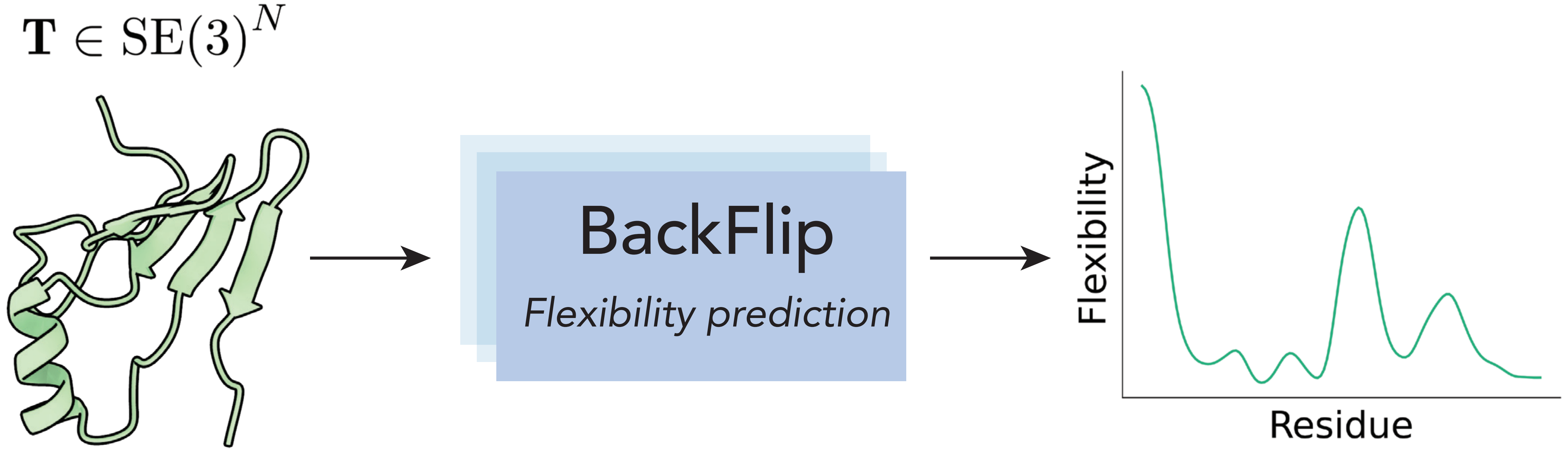
FliPS

Flexibility-conditioned structure design

BackFlip

Flexibility prediction

BackFlip: Backbone Flexibility predictor



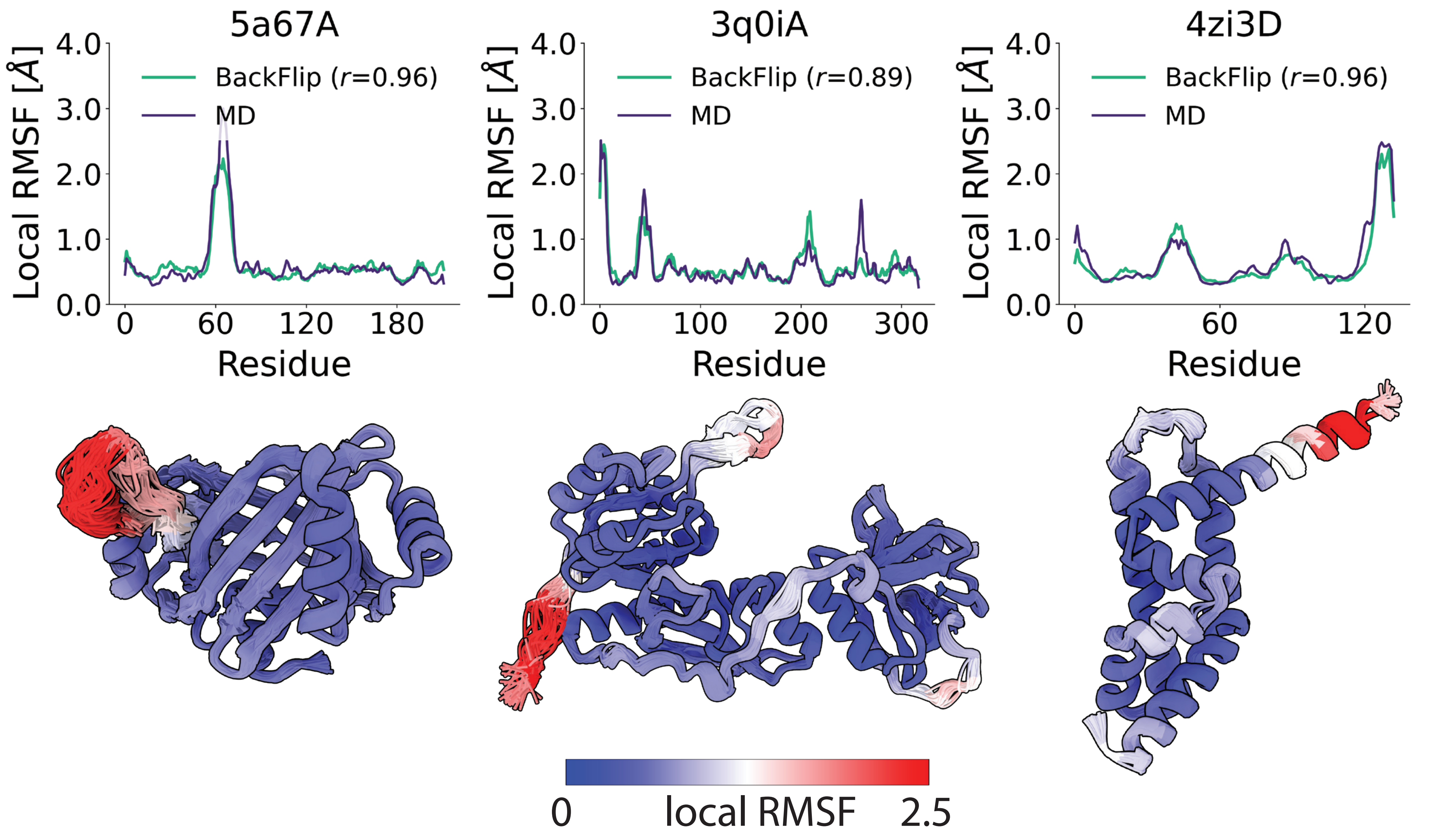
$\mathbf{T} \in \mathrm{SE}(3)^N$

BackFlip

Flexibility prediction

ATLAS Molecular Dynamics (MD) dataset of 1390 proteins

	ATLAS test set		**De novo proteins**		
	Global r (↑)	MAE [Å] (↓)	Global r (↑)	MAE [Å] (↓)	Time [s] (↓)
MD (Ground Truth)	0.84 (0.00)	0.14 (0.00)	0.80 (0.01)	0.10 (0.00)	$\mathcal{O}$(10,000)
BackFlip	0.80 (0.00)	0.17 (0.00)	0.73 (0.00)	0.11 (0.01)	0.6



BackFlip flexibility screening of RFdiffusion and FoldFlow2 yields samples with worse flexibility metrics

	MD of top samples		
	r (↑)	MAE [Å] (↓)	Novelty (↓)
4 hand-drawn profiles			
FliPS	**0.79** (0.01)	**0.21** (0.00)	0.61 (-)
RFdiffusion-BFS	0.41 (0.02)	0.31 (0.00)	0.57 (-)
FoldFlow2-BFS	0.35 (0.02)	0.34 (0.00)	**0.48** (-)

BackFlip flexibility screening (BFS)

$$s(\xi, \xi_{\text{ref}}) = w_{\text{corr}}\, r(\xi, \xi_{\text{ref}}) - w_{\text{mae}}\, \mathbf{MAE}(\xi, \xi_{\text{ref}})$$

References

[1] Joseph L. Watson et al. "De novo design of protein structure and function with RFdiffusion." Nature 620.7976 (2023): 1089-1100.
[2] Huguet, Guillaume et al. "Sequence-augmented SE(3)-flow matching for conditional protein backbone generation." arXiv preprint arXiv:2405.20313 (2024).

Summary

- We introduce a flexibility-conditioned flow matching generative model for protein structure **FliPS** and a flexibility predictor **BackFlip**
- We propose a framework to design de novo proteins with desired flexibilities
- Target flexibilities of proteins generated with FliPS are recapitulated in MD

Authors


Vsevolod Viliuga


Leif Seute


Nicolas Wolf


Simon Wagner


Arne Elofsson


Jan Stühmer


Frauke Gräter