how artificial intelligence is reshaping protein structure prediction and therapeutic design

Jens Meiler

Humboldt Professor and Director, Institute for Drug Discovery, University Leipzig, Leipzig, Saxony, Germany

Distinguished Professor, Department of Chemistry and Pharmacology, Vanderbilt University, Nashville, TN, USA

UNIVERSITÄT LEIPZIG

and

VANDERBILT UNIVERSITY

are partner universities: www.leipzig.vanderbilt.edu

Nobelpreis Chemie 2024

The Nobel Prize in Chemistry 2024 was divided, one half awarded to David Baker “for computational protein design”, the other half jointly to Demis Hassabis and John M. Jumper “for protein structure prediction”.

Ill. Niklas Elmehed © Nobel Prize Outreach

David Baker

Prize share: 1/2

Ill. Niklas Elmehed © Nobel Prize Outreach

Demis Hassabis

Prize share: 1/4

Ill. Niklas Elmehed © Nobel Prize Outreach

John M. Jumper

Prize share: 1/4

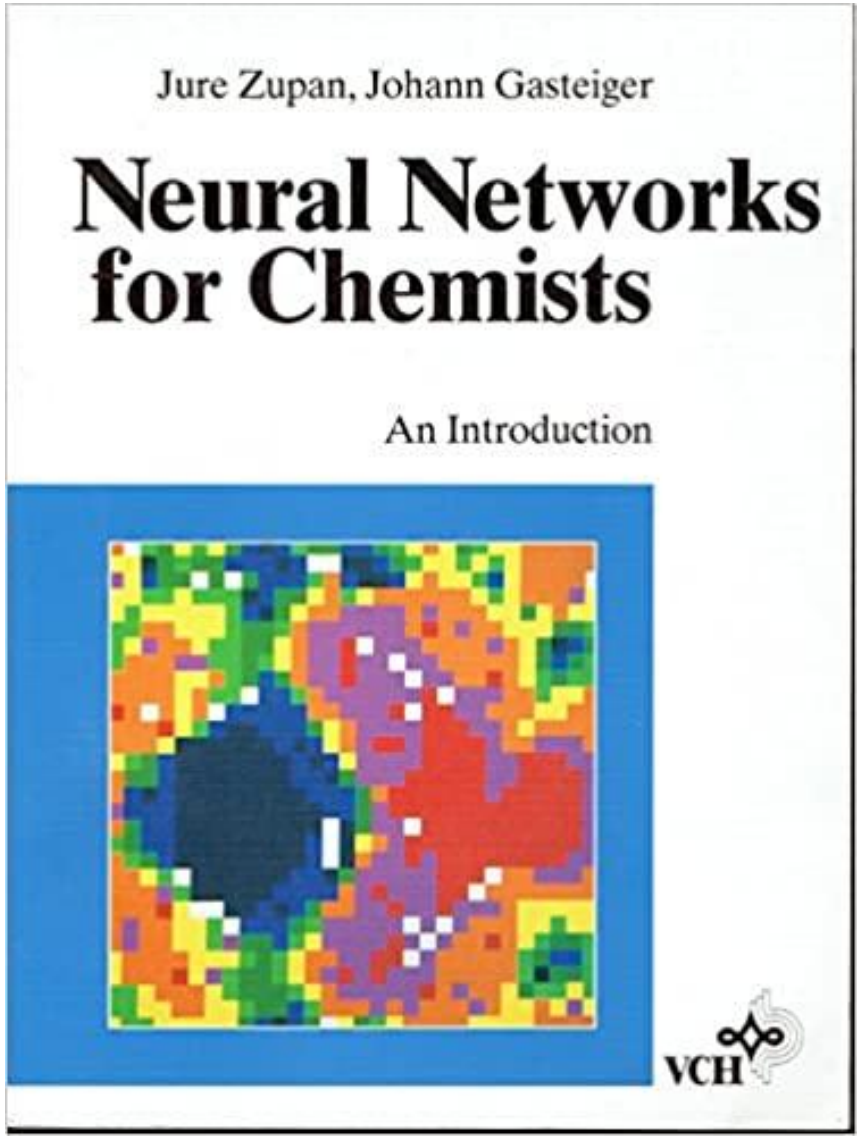
Nobelpreis Chemie 2024

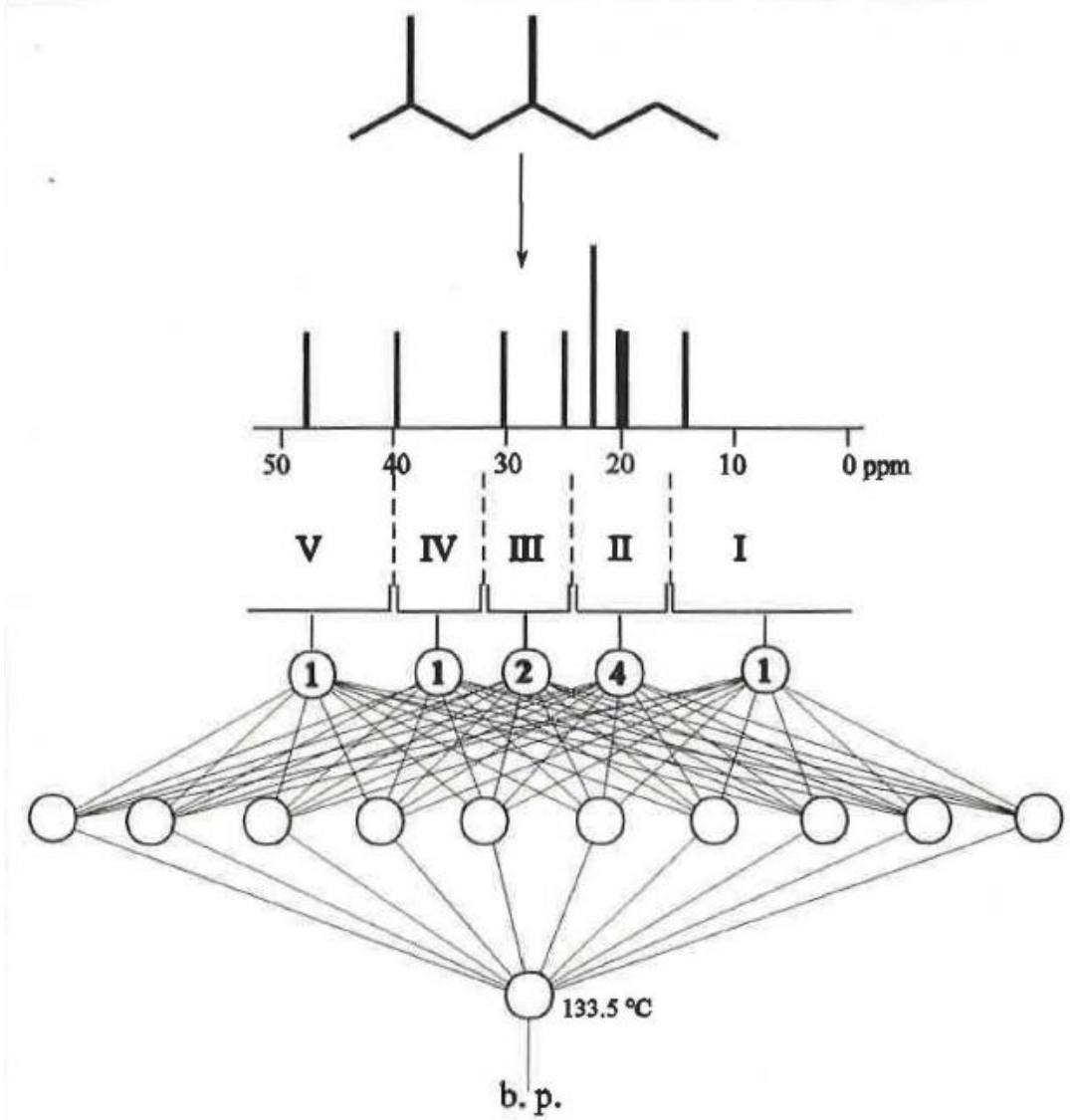
The Nobel Prize in Chemistry 2024 was divided, one half awarded to David Baker “for computational protein design”, the other half jointly to Demis Hassabis and John M. Jumper “for protein structure prediction”.

Nobelpreis Chemie 2024

The Nobel Prize in Chemistry 2024 was divided, one half awarded to David Baker “for computational protein design”, the other half jointly to Demis Hassabis and John M. Jumper “for protein structure prediction”.

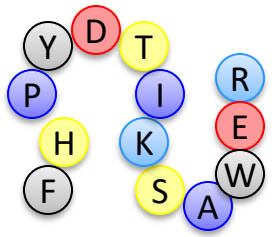
Since 1993(!) – Neural Networks for Chemists

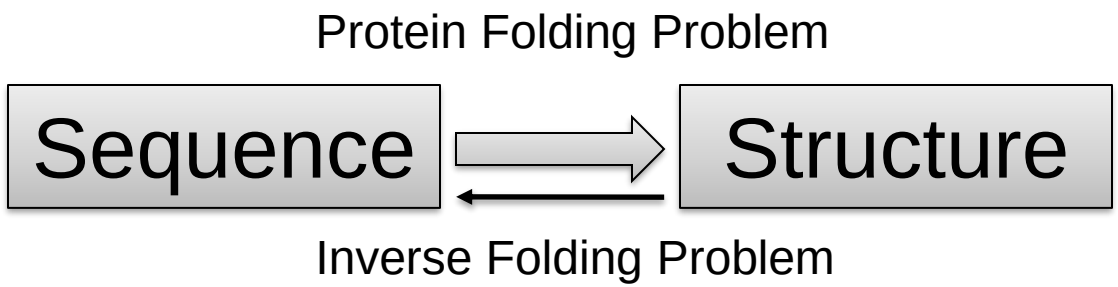


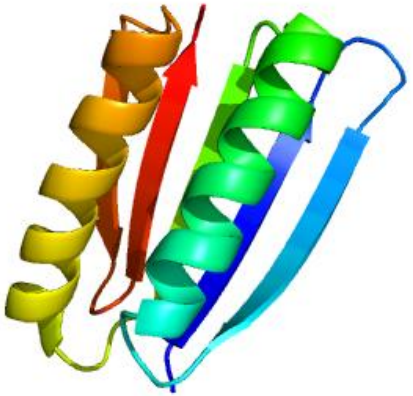


J. Meiler and R. Meusinger "Use of Neural Networks to Determine Properties of Alkanes from their 13C-NMR Spectra" in *Software - Entwicklung in der Chemie*; Gasteiger, J., Ed. Gesellschaft Deutscher Chemiker: Frankfurt am Main; **1995**; Vol. 10: p. 259-263.

The (Inverse) Protein Folding Problem Holy Grail of Comp. Structural Biology



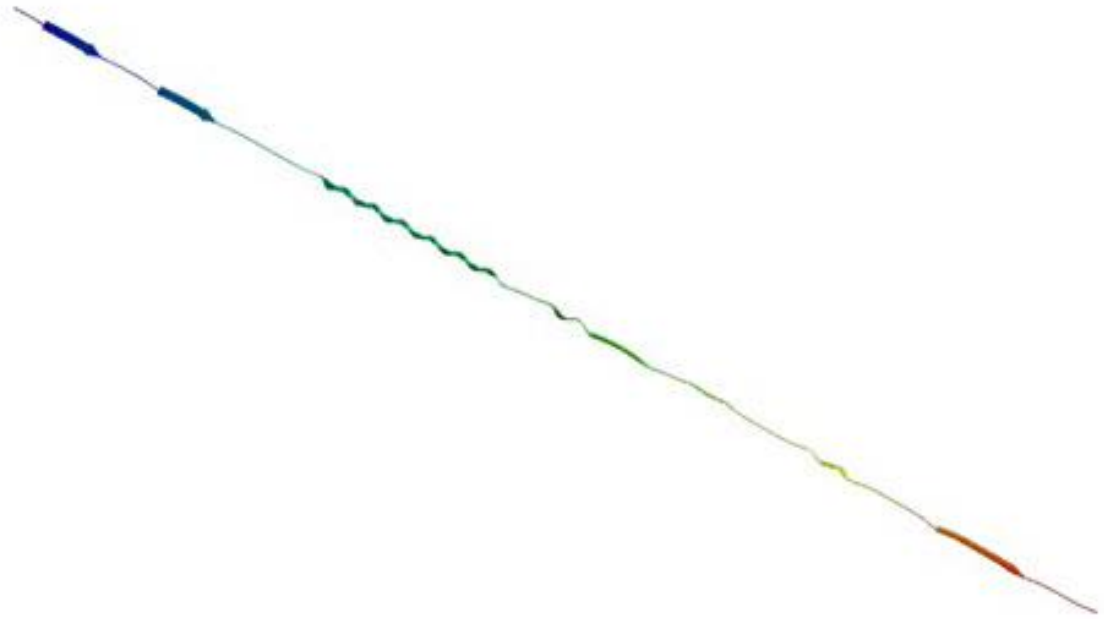


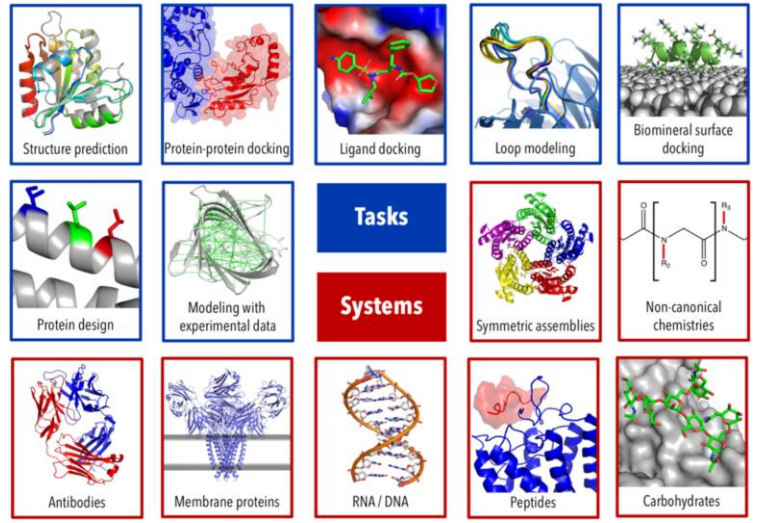


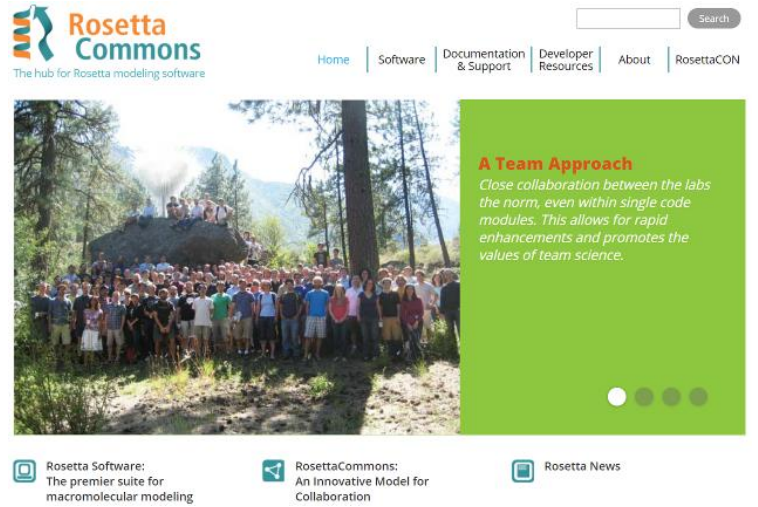
- Given a protein's AA sequence, what is its 3-dimensional fold , and how does it get there?
- Assume 100 conformations for each amino acid in a 100 amino acid protein ⇨ 10^{200} possible conformations!
- Cyrus Levinthal's paradox of protein folding,1968.

Rosetta: A Unified Framework for Protein Structure Prediction and Design

meilerlab

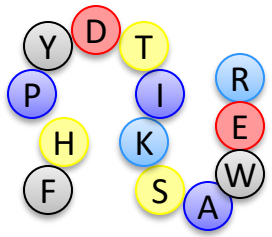






A. Leaver-Fay, et al.; "ROSETTA3: an object-oriented software suite …"; Methods Enzymol; 2011; Vol. 487 p. 545-74.
J. K. Leman, et al.; "Macromolecular modeling and design in Rosetta: recent methods and frameworks"; Nat Methods; 2020; Vol. 17 (7): p. 665-680.

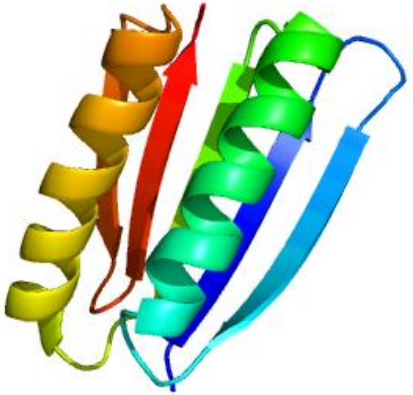
The (Inverse) Protein Folding Problem Holy Grail of Comp. Structural Biology



Protein Folding Problem

Sequence ⇌ Structure

Inverse Folding Problem

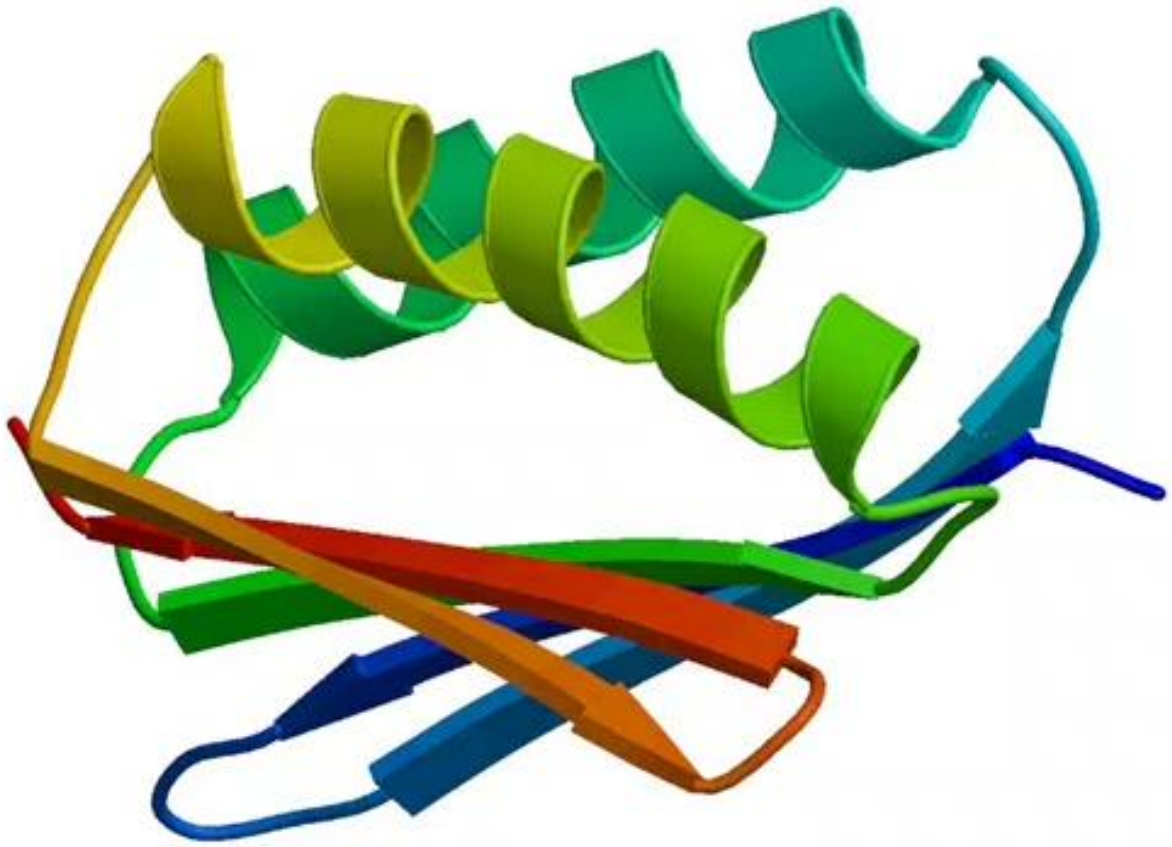


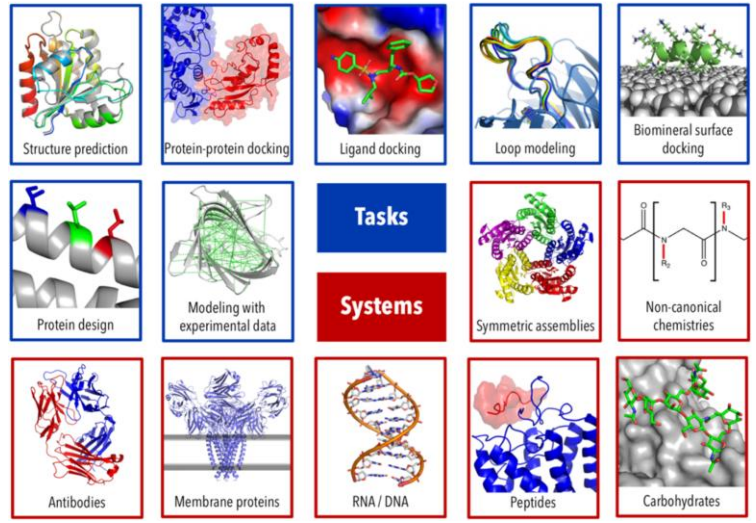
- Given a protein's AA sequence, what is its 3-dimensional fold , and how does it get there?
- Assume 100 conformations for each amino acid in a 100 amino acid protein ⇨ 10^{200} possible conformations!
- Cyrus Levinthal's paradox of protein folding,1968.

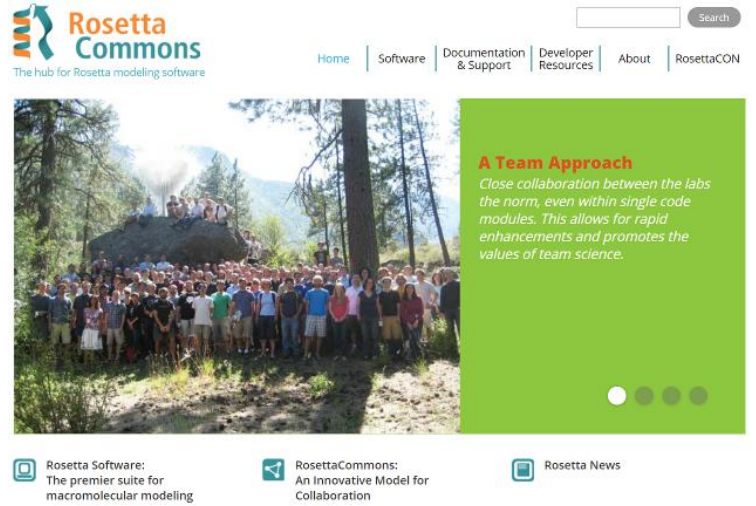
- Given a protein fold, which primary sequence(s) fold into it?
- Assume a total of 100 conformations for all 20 natural occurring amino acids side chains in a 100 amino acid protein ⇨ 10^{200} possible conformations!
- Earth is less than 10^{10} years old.

Rosetta: A Unified Framework for Protein Structure Prediction and Design

meilerlab

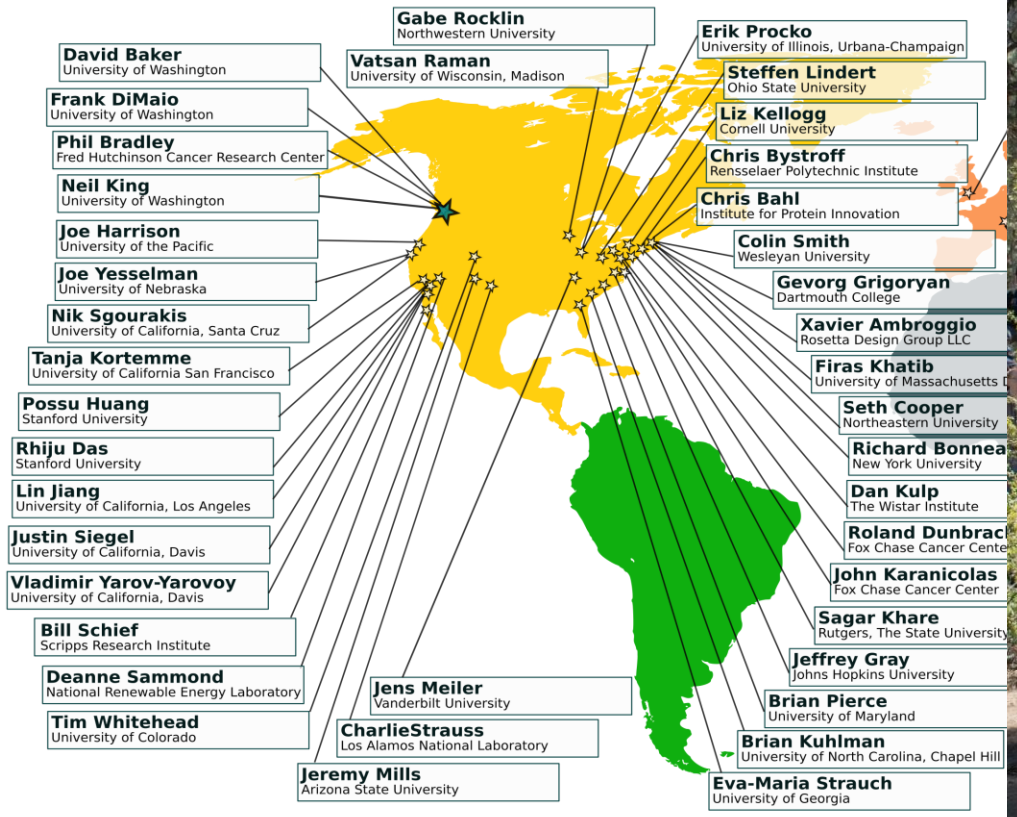






A. Leaver-Fay, et al.; "ROSETTA3: an object-oriented software suite …"; Methods Enzymol; 2011; Vol. 487 p. 545-74.
J. K. Leman, et al.; "Macromolecular modeling and design in Rosetta: recent methods and frameworks"; Nat Methods; 2020; Vol. 17 (7): p. 665-680.

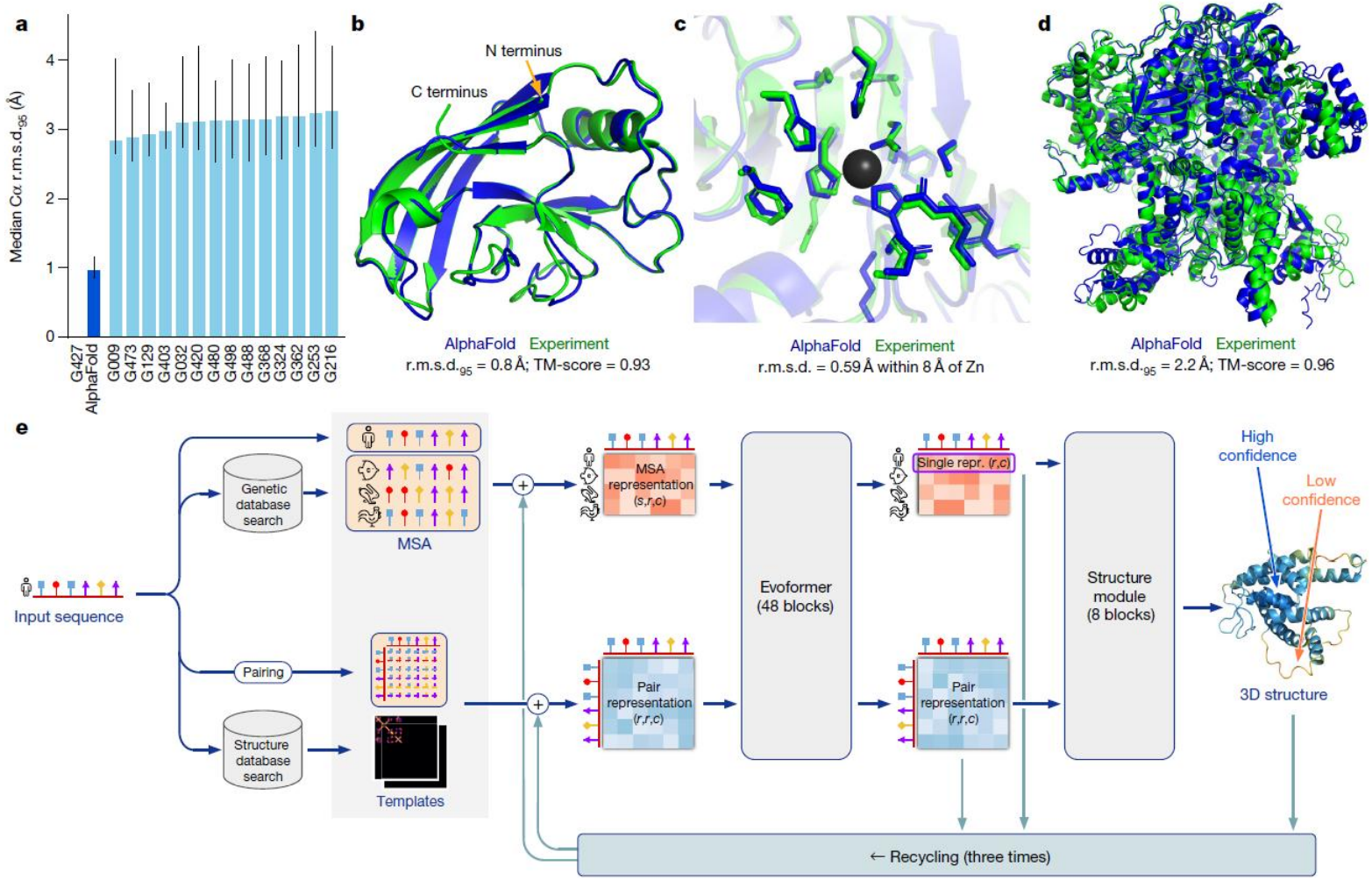
Rosetta: A Unified Framework for Protein Structure Perdiction and Design



A. Leaver-Fay, et al.; "ROSETTA3: an object-oriented software suite …"; Methods Enzymol; 2011; Vol. 487 p. 545-74.

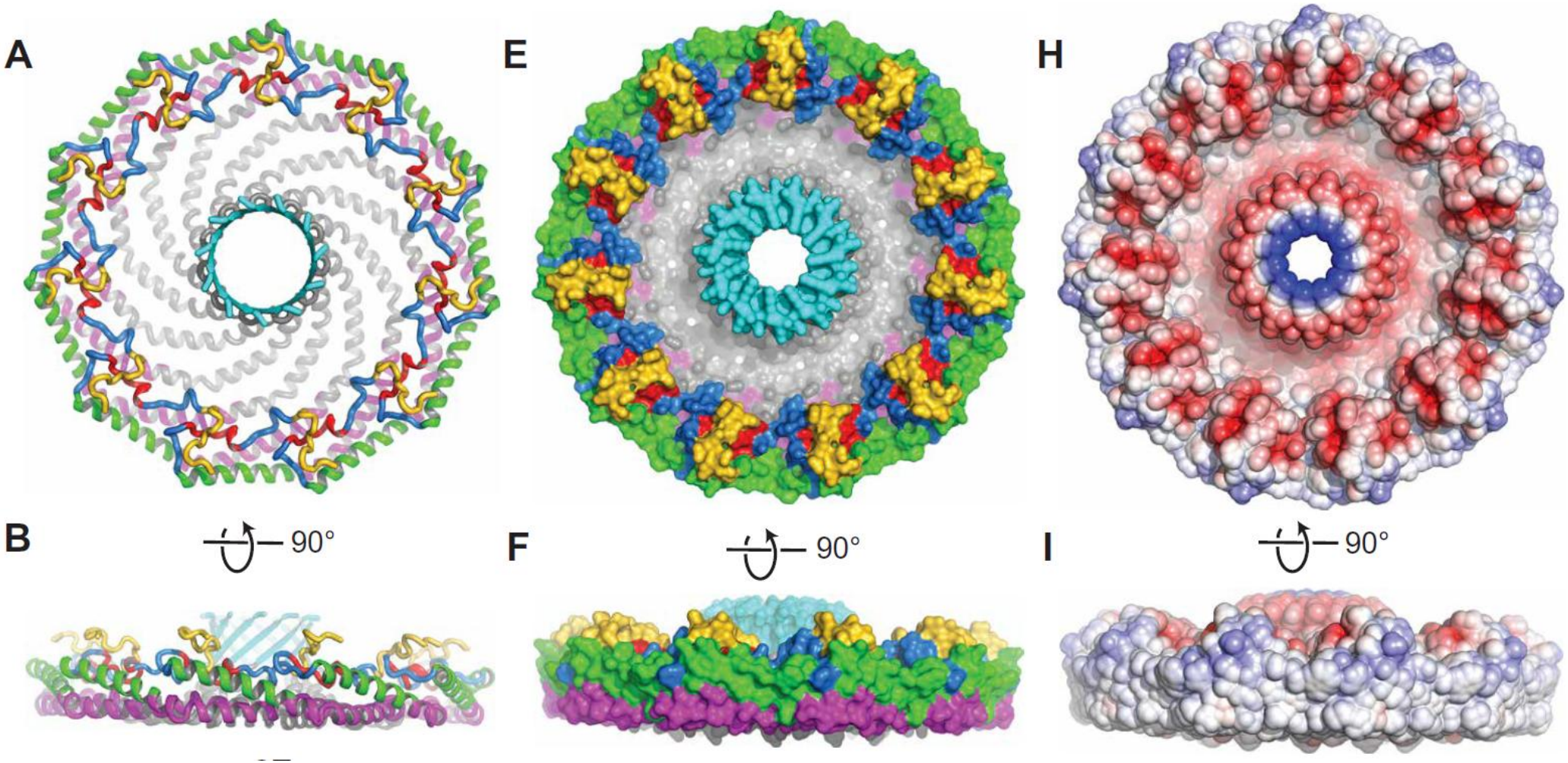
J. K. Leman, et al.; "Macromolecular modeling and design in Rosetta: recent methods and frameworks"; Nat Methods; 2020; Vol. 17 (7): p. 665-680.

Highly accurate protein structure prediction with AlphaFold2



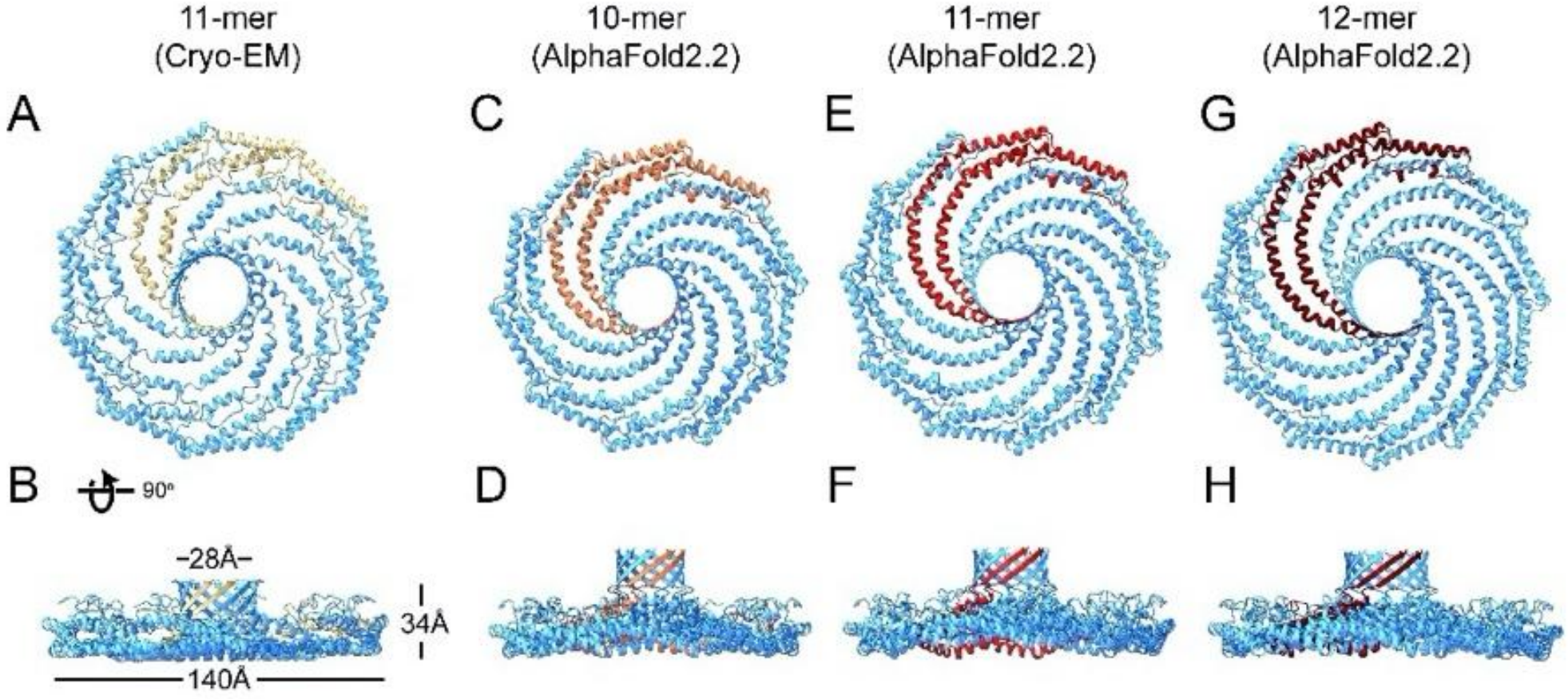
J. Jumper, et al.; "Highly accurate protein structure prediction with AlphaFold"; *Nature*; **2021**; Vol. 596 (7873): p. 583-589.

Molecular Architecture of the Human Caveolin-1 Complex



J. C. Porta, B. Han, A. Gulsevin, J. Chung, Y. Peskova, S. Connolly, H. S. Mchaourab, J. Meiler, E. Karakas, A. K. Kenworthy and M. D. Ohi; "Molecular architecture of the human caveolin-1 complex"; *Science Advances*; **2022**; Vol. p.

Molecular Architecture of the Human Caveolin-1 Complex with AlphaFold2

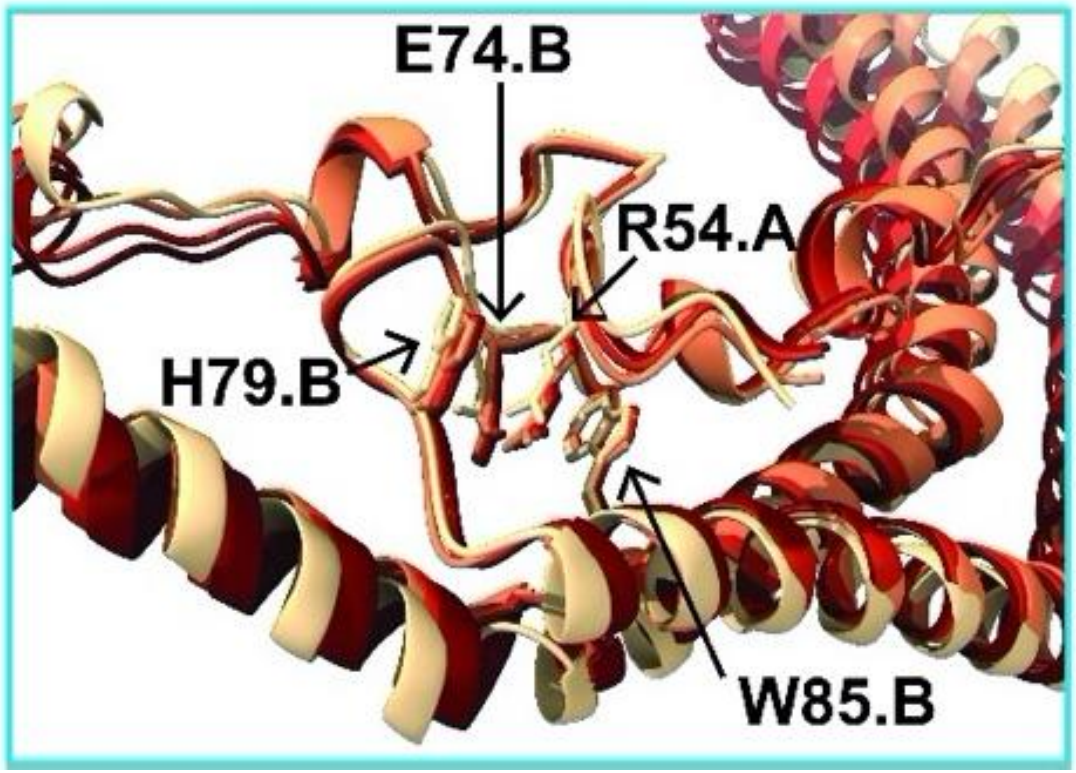


A. Gulsevin, B. Han, J. C. Porta, H. S. McHaourab, J. Meiler and A. K. Kenworthy; "Template-free prediction of a new monotopic membrane protein fold and assembly by AlphaFold2"; *Biophys J;* ***2022; Vol. p.***

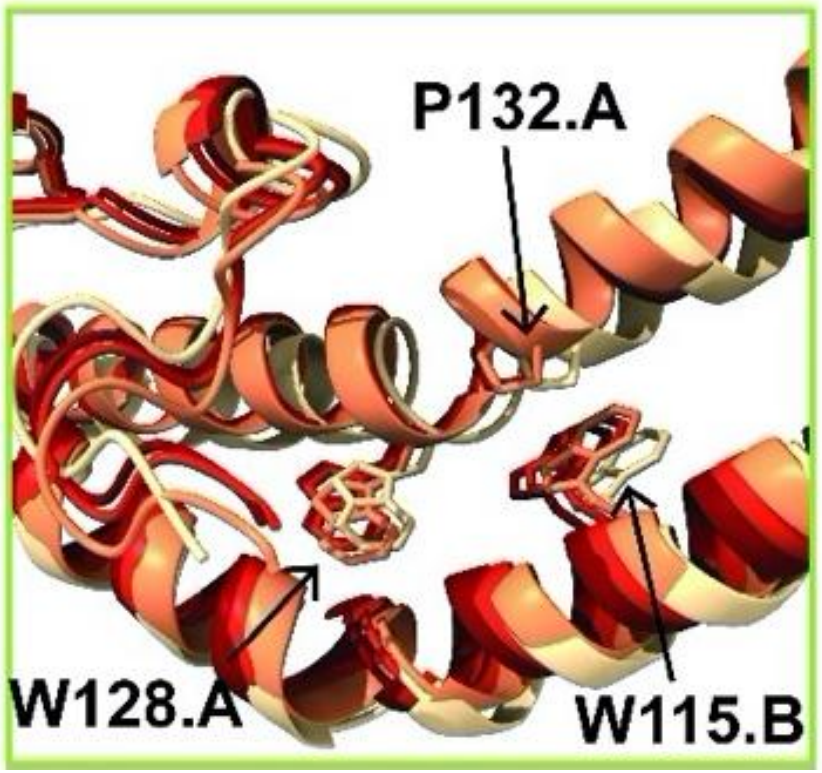
J. C. Porta, B. et al. "Molecular architecture of the human caveolin-1 complex"; *Science Advances*; **2022**; Vol. p.

Molecular Architecture of the Human Caveolin-1 Complex with AlphaFold2

K



L



A. Gulsevin, B. Han, J. C. Porta, H. S. McHaourab, J. Meiler and A. K. Kenworthy; "Template-free prediction of a new monotopic membrane protein fold and assembly by AlphaFold2"; *Biophys J;* ***2022; Vol. p.***

J. C. Porta, B. et al. "Molecular architecture of the human caveolin-1 complex"; *Science Advances*; **2022**; Vol. p.

What does the Nobel Prize mean for 2025+? Integrating Chemical and Structural Biology!

meilerlab

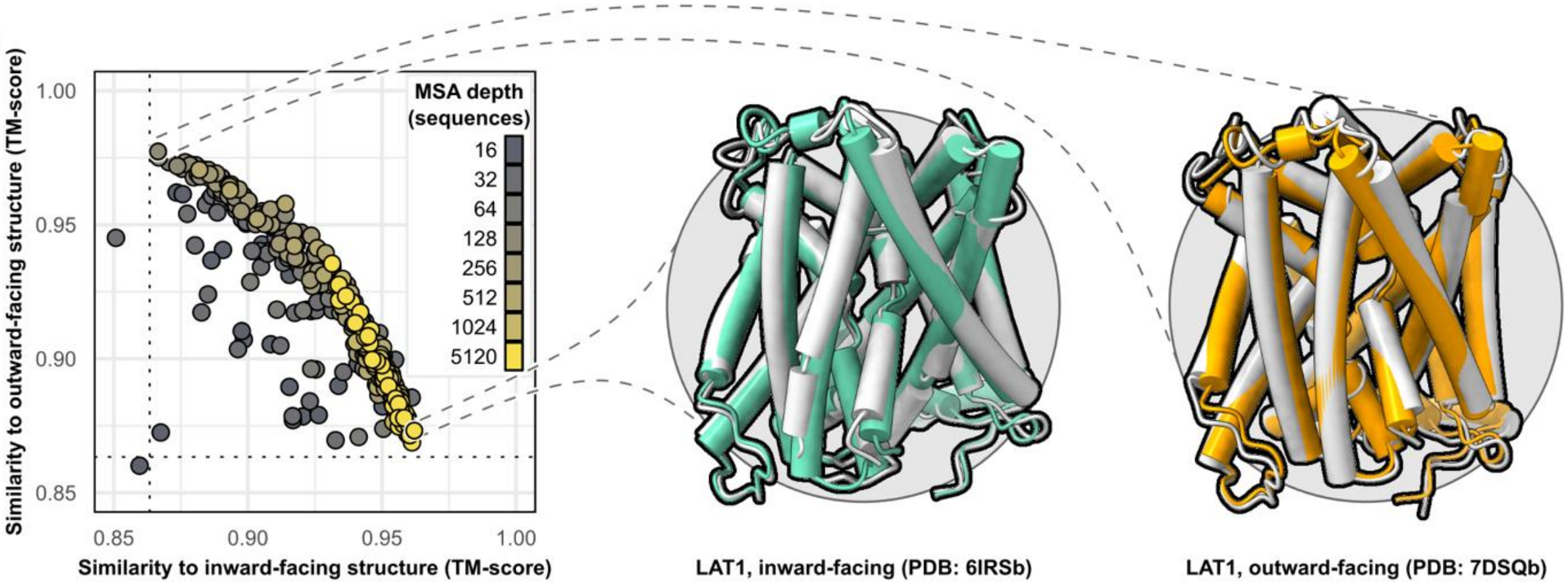
Technology

- Expand the chemical space of AI design
- Learn to design protein dynamics for function
- Control at 0.1Å instead of 1.0Å

Application

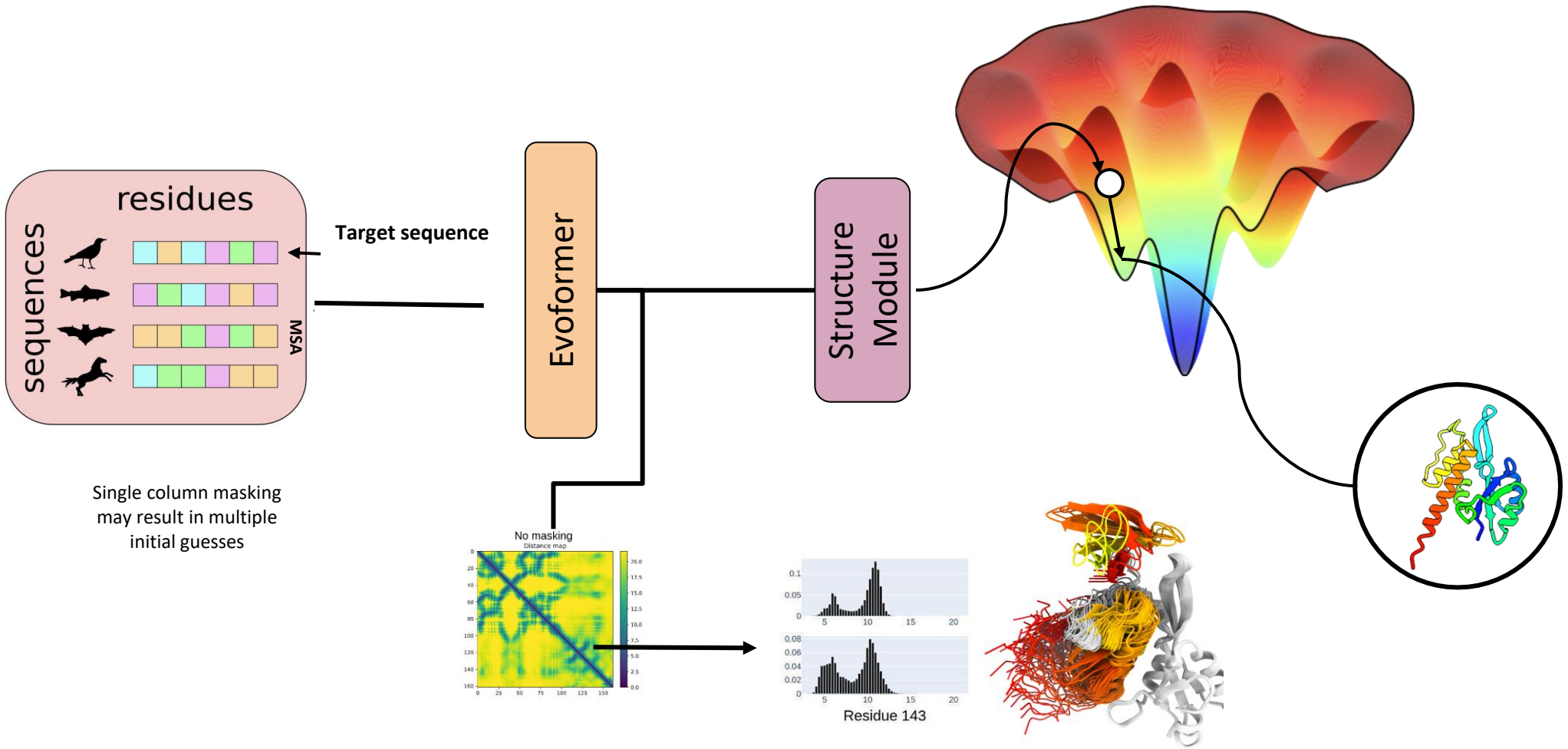
- Medicine & Therapeutics
- Technology and Materials
- Sustainability

Sampling Alternative Conformational States with AlphaFold2

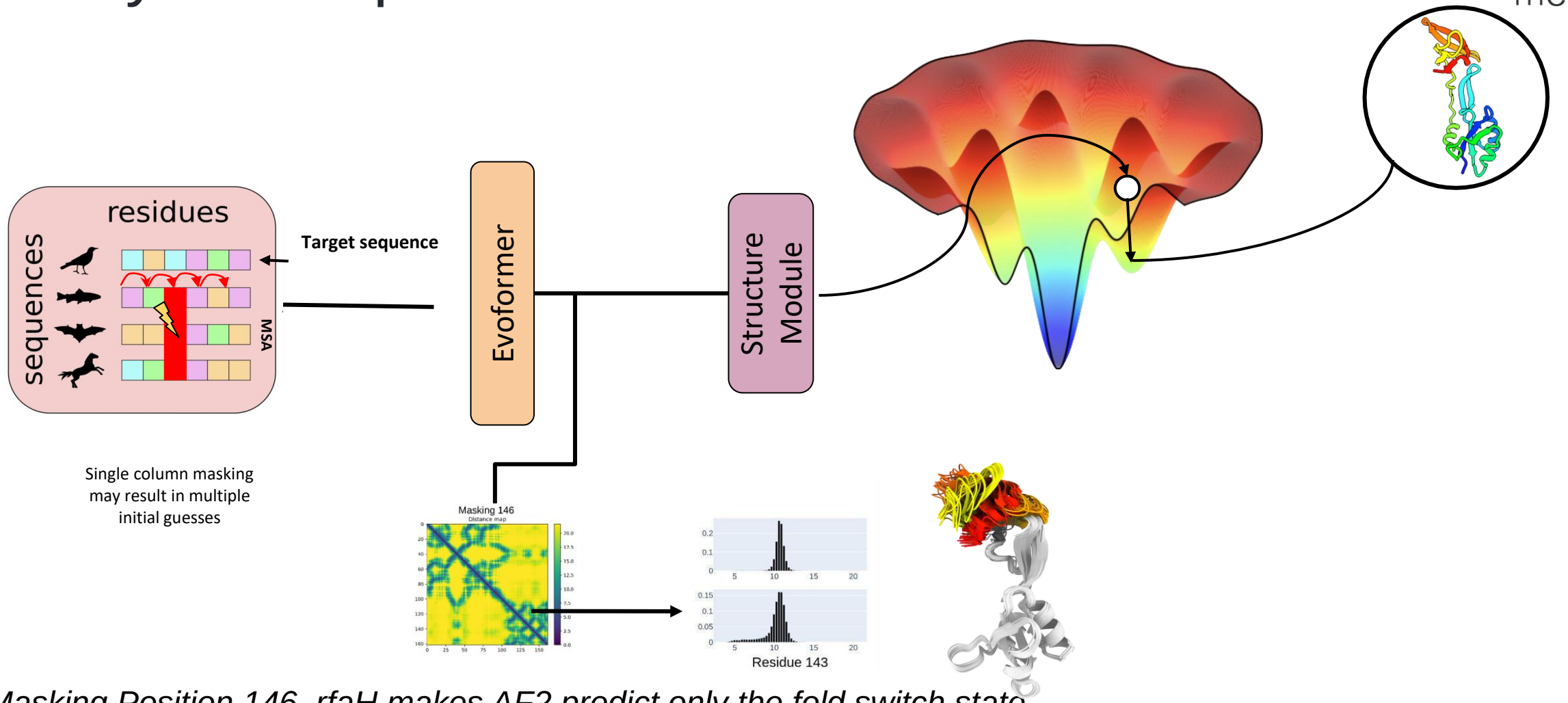


D. Del Alamo, D. Sala, H. S. Mchaourab and J. Meiler; "Sampling alternative conformational states of transporters and receptors with AlphaFold2"; *Elife*; **2022**; Vol. 11 p.

Explaining by removing: Demystifying AF2 ability to sample conformations – rfaH

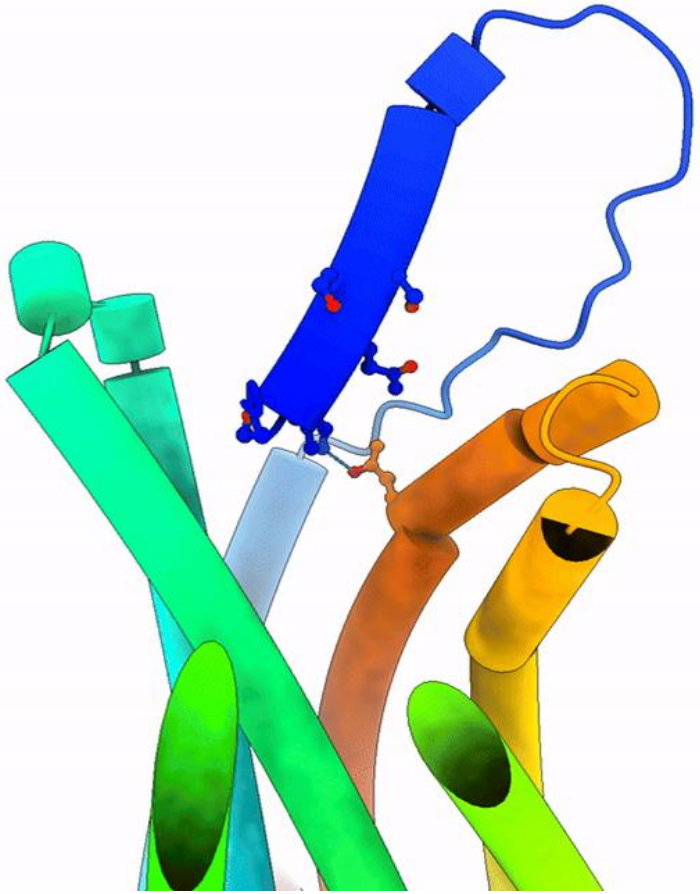


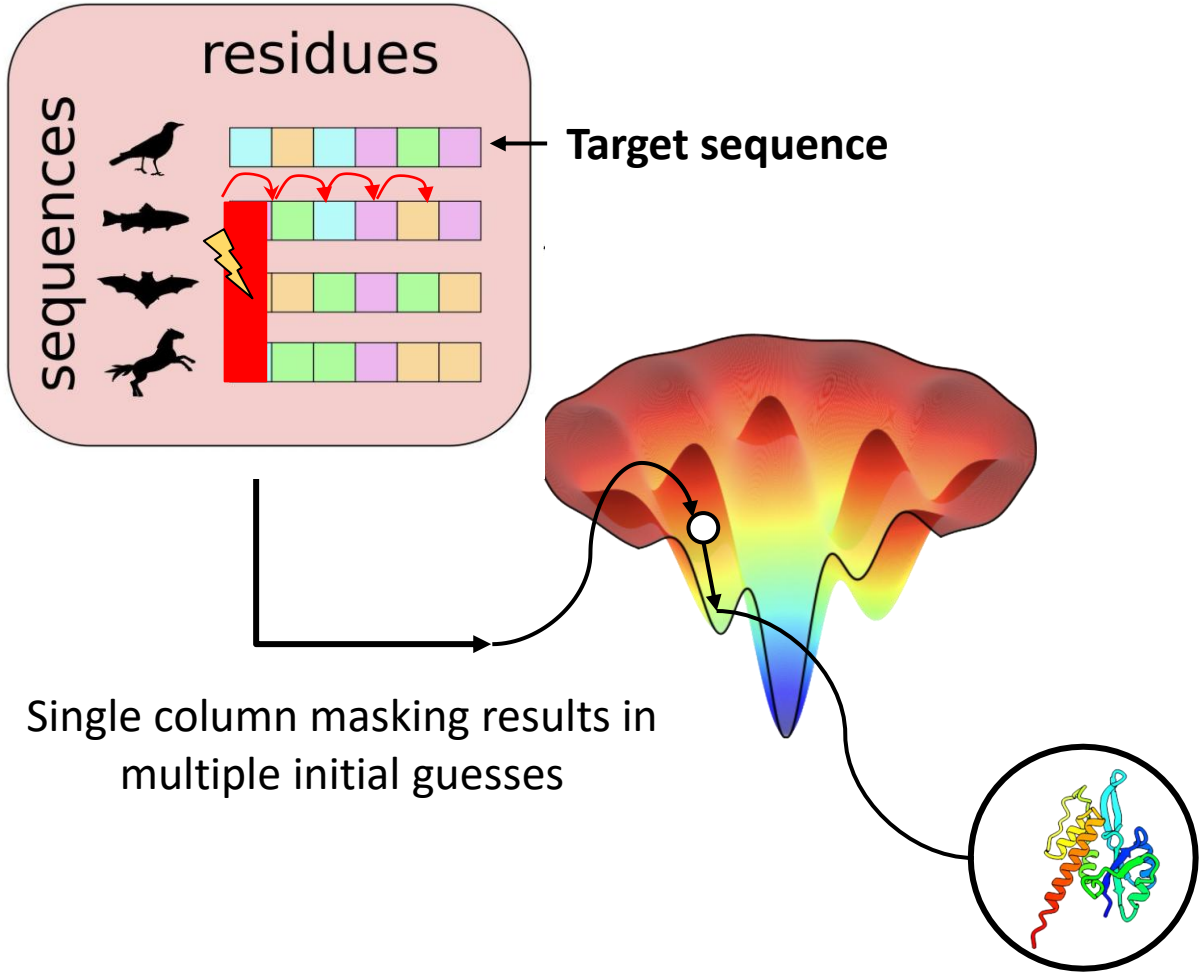
Explaining by removing: Demystifying AF2 ability to sample conformations – rfaH



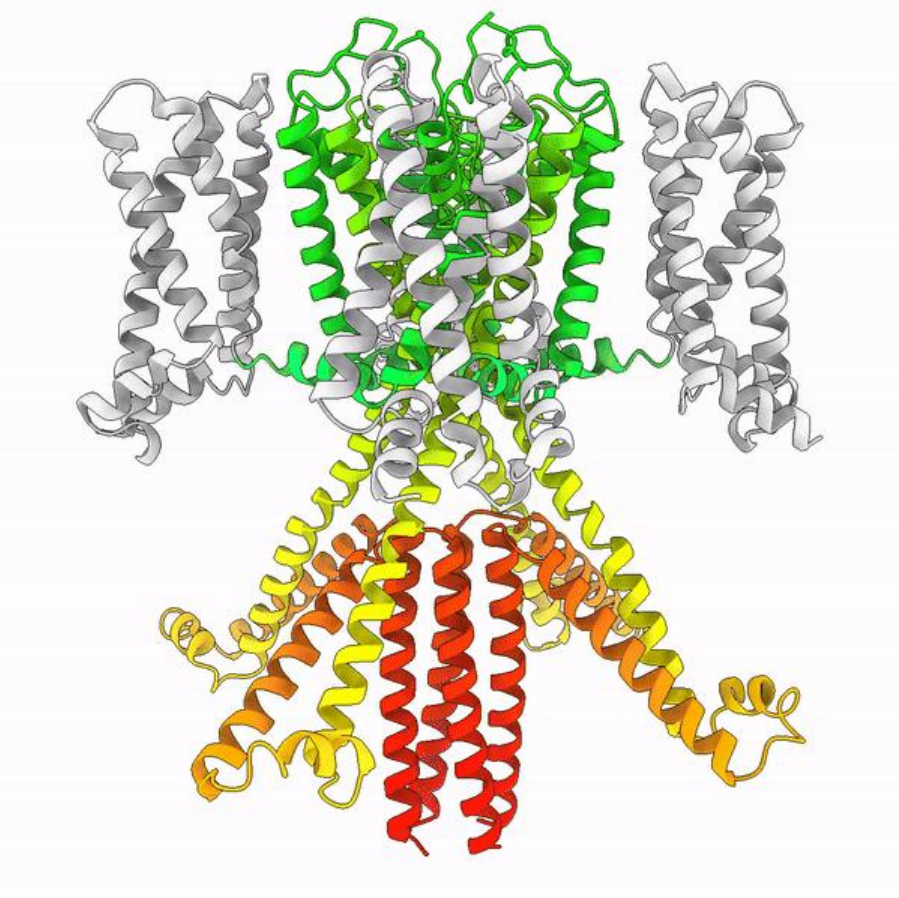
Masking Position 146 rfaH makes AF2 predict only the fold switch state

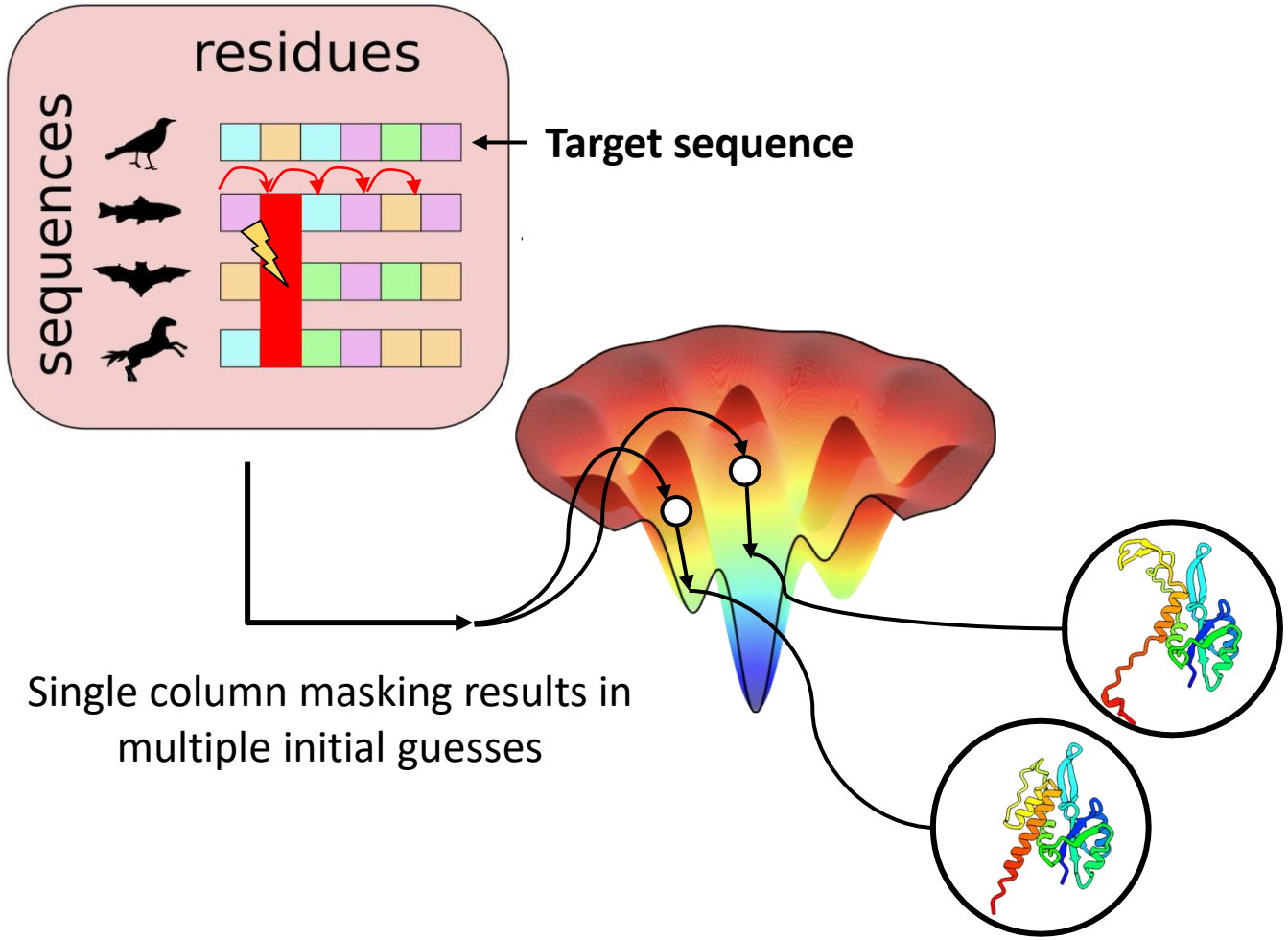
Demystifying AF2 ability to sample conformations
MC4R auto inhibiting peptide



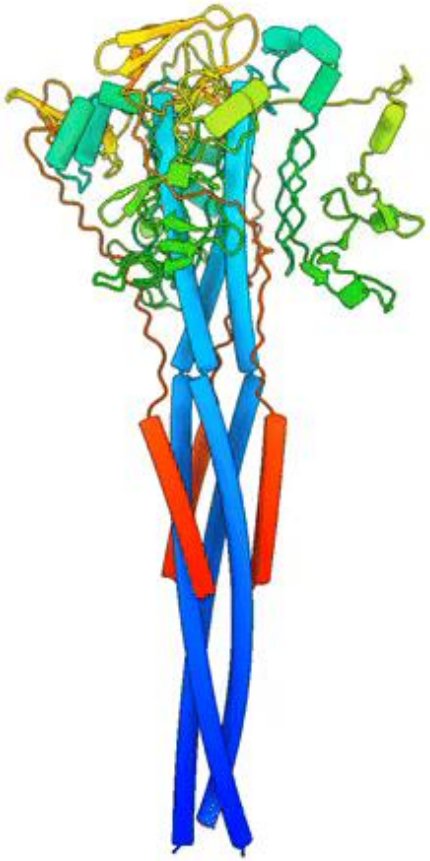


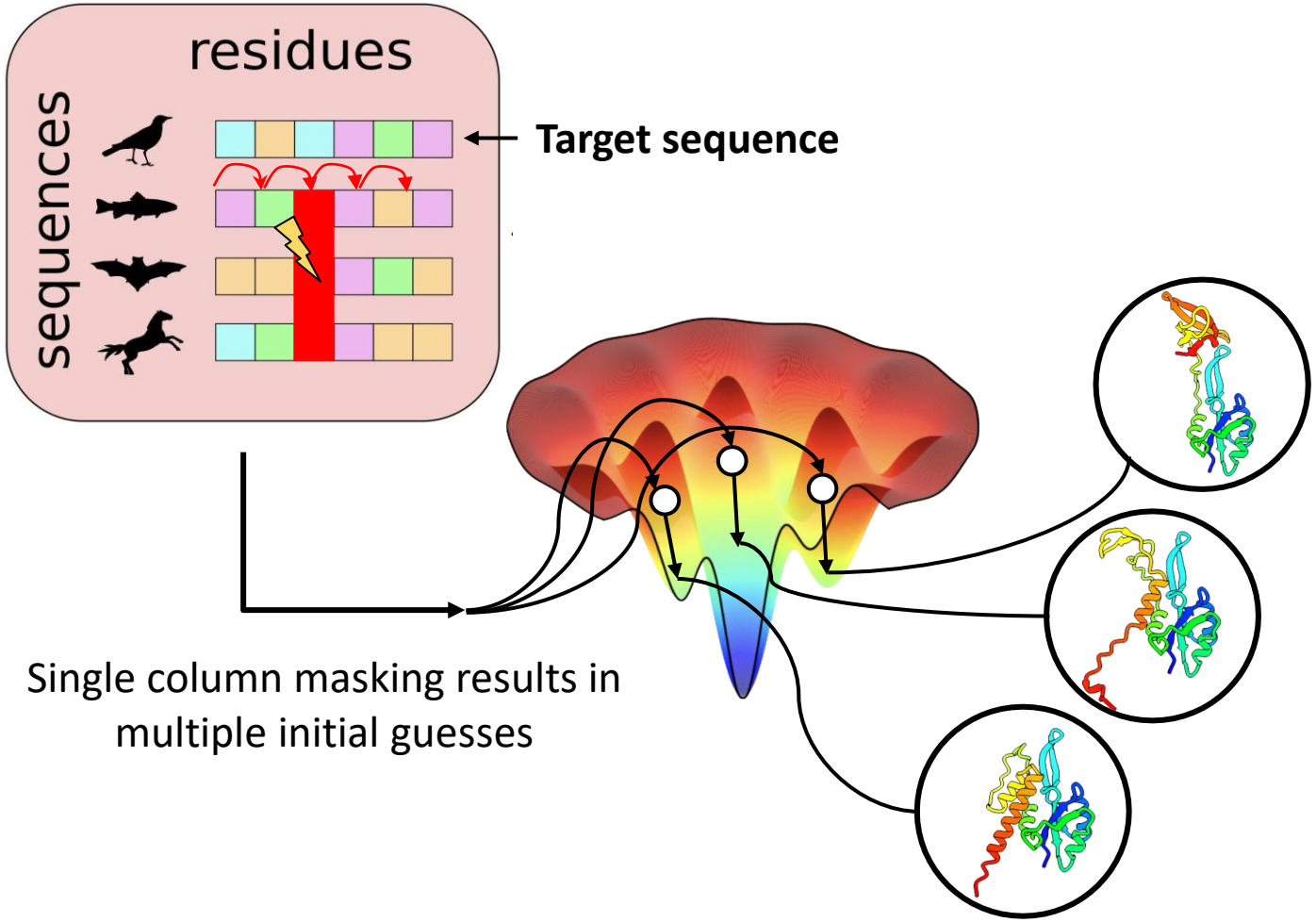
Demystifying AF2 ability to sample conformations
KCNQ1 potassium channel



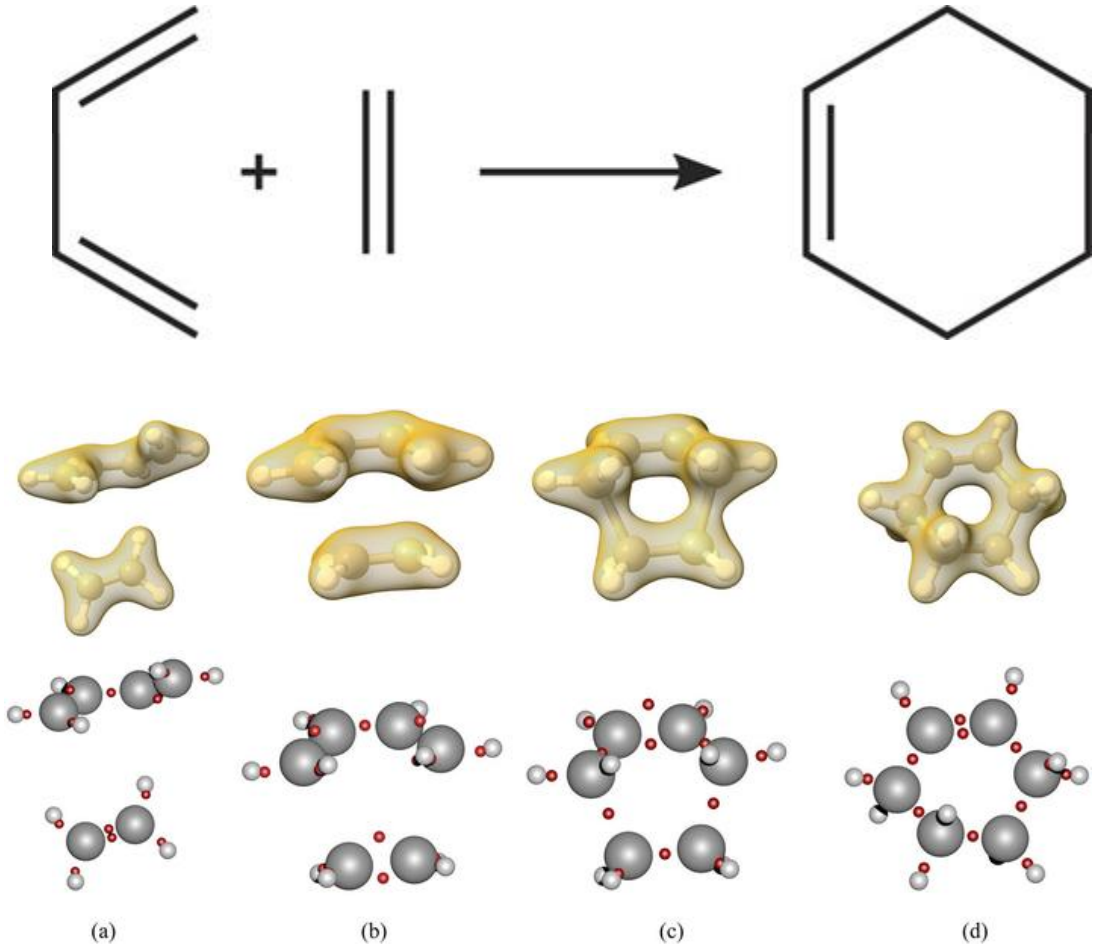


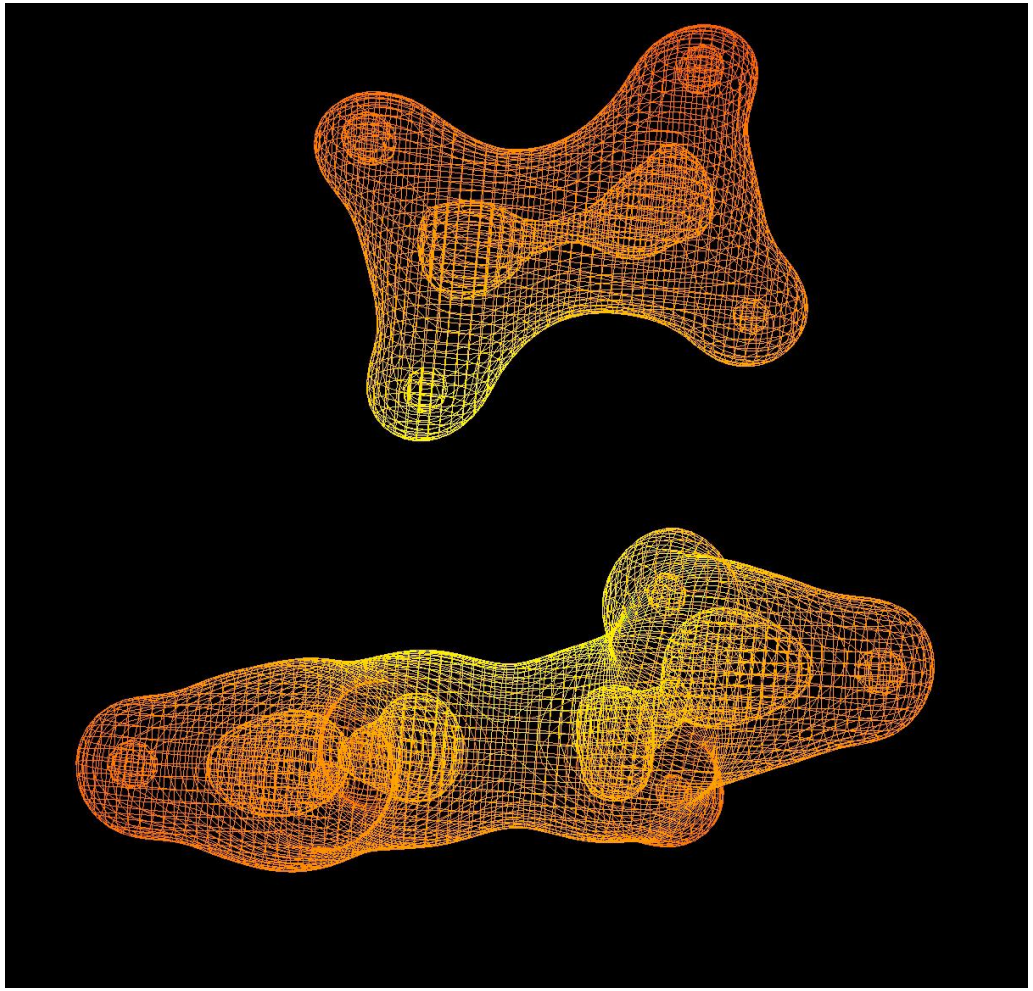
Demystifying AF2 ability to sample conformations *Respiratory Syncytial Virus Glycoprotein*



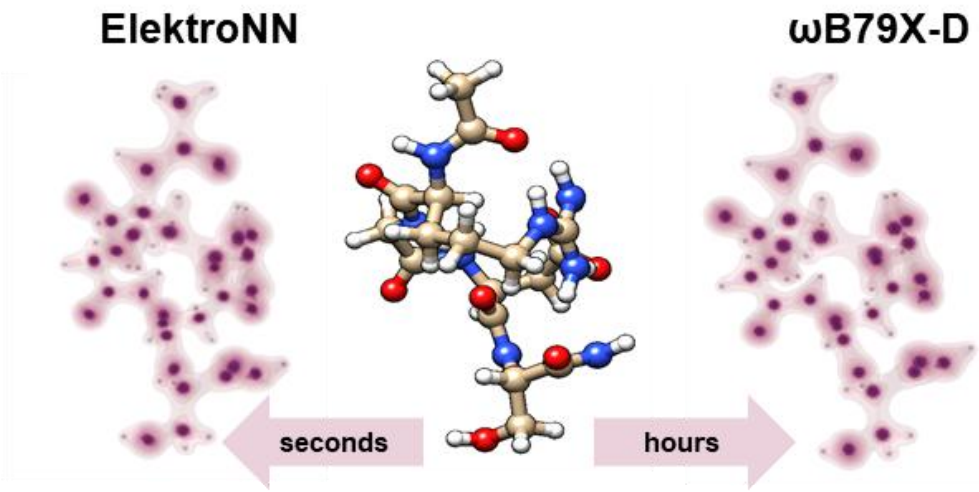


ElektroNN: Predicting Electron Density changes for the Diels-Alder Reaction

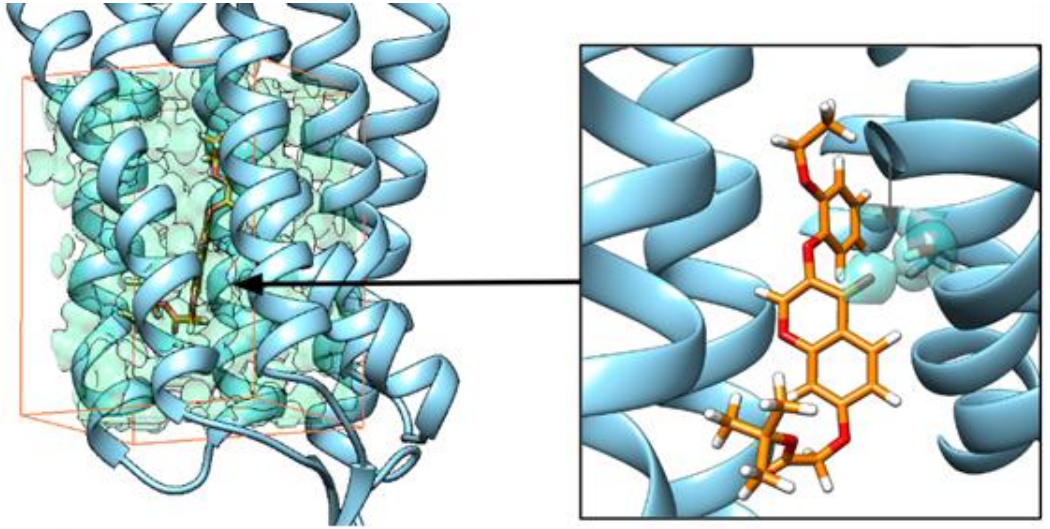


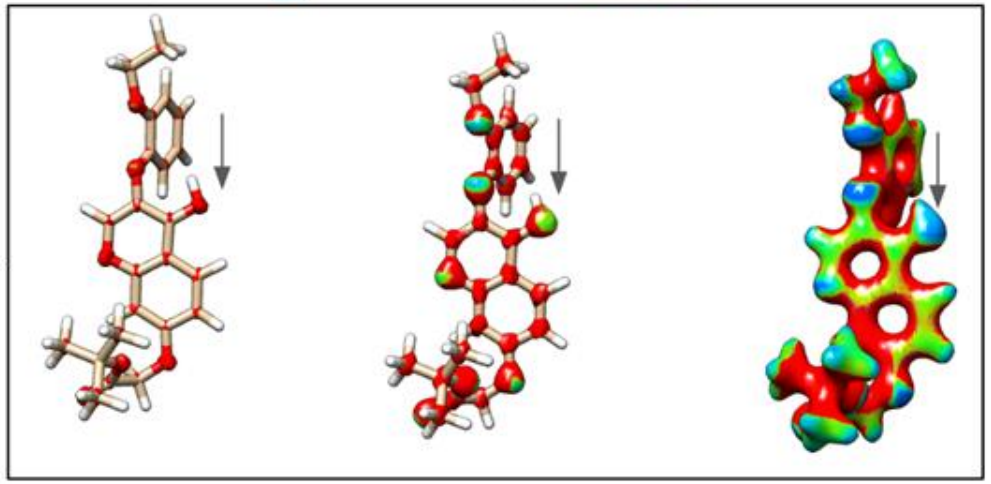


ElektroNN: Generalization to non-covalent Interactions in Ligands, Peptides, and Proteins



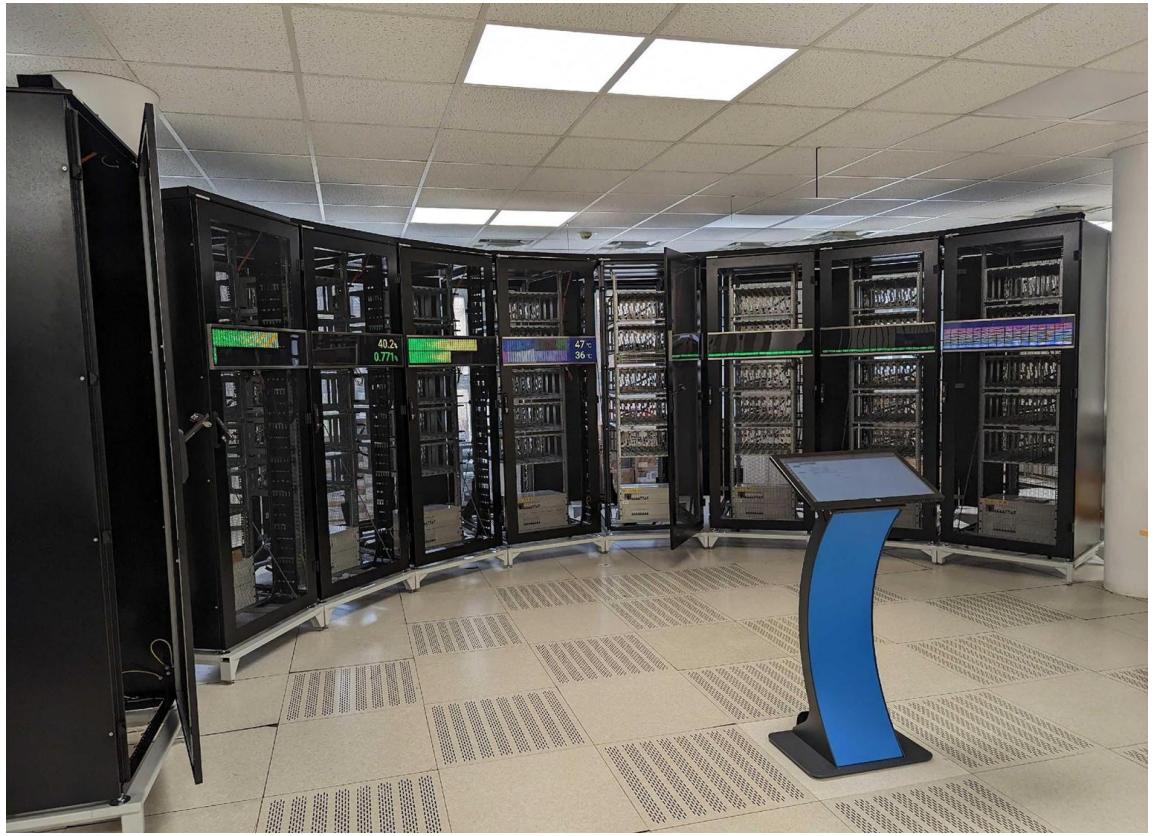
Peptide (5 conformations each)	Charge	Error per Atom (%)
VRN	1	0.30
KGD	0	0.35
DGEA	-2	0.29
APGL	0	0.25
RGD	0	0.30

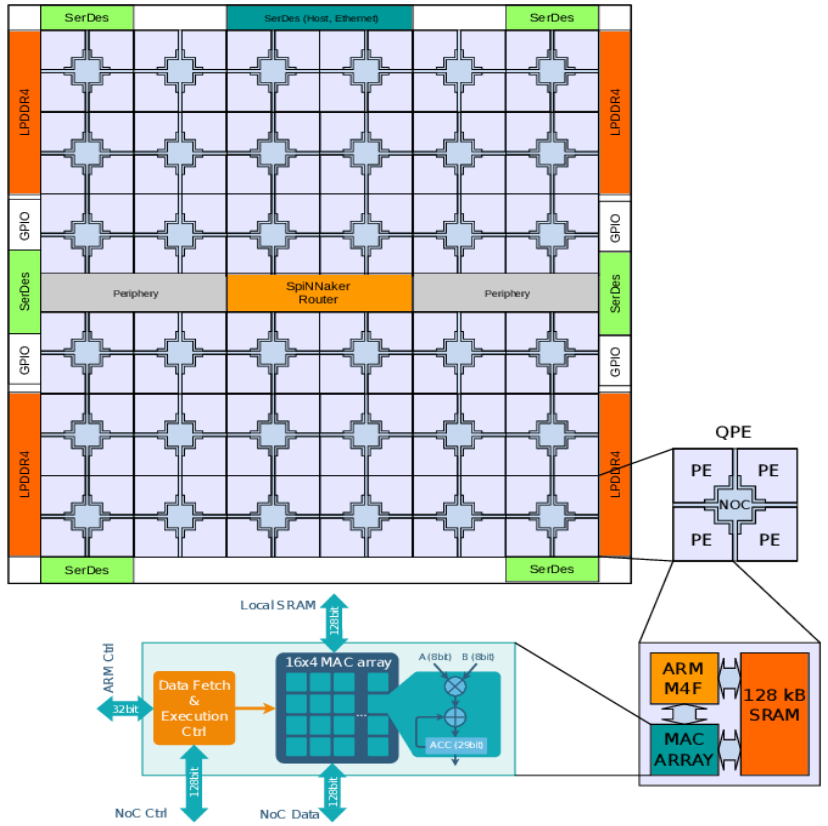




The SpiNNaker2 Platform: Between GPUs and QC
10^{10} Cores and 10^{14} Connections

meilerlab





Yan, Stewart, Choo, Vogginger, Partzsch, Höppner, Kelber, Eliasmith, Furber, and **Mayr**: “Comparing Loihi with a SpiNNaker 2 prototype on low-latency keyword spotting and adaptive robotic control”; Neuromorph. Comput. Eng. **2021** || Höppner, Yan, Dixius, Scholze, Partzsch, Stolba, Kelber, Vogginger, Neumärker, Ellguth, Hartmann, Schiefer, Hocker, Walter, Liu, Garside, Furber, **Mayr**: “The SpiNNaker 2 Processing Element Architecture for Hybrid Digital Neuromorphic Computing”; preprint arXiv:2103.08392 [cs.AR]; **2021**

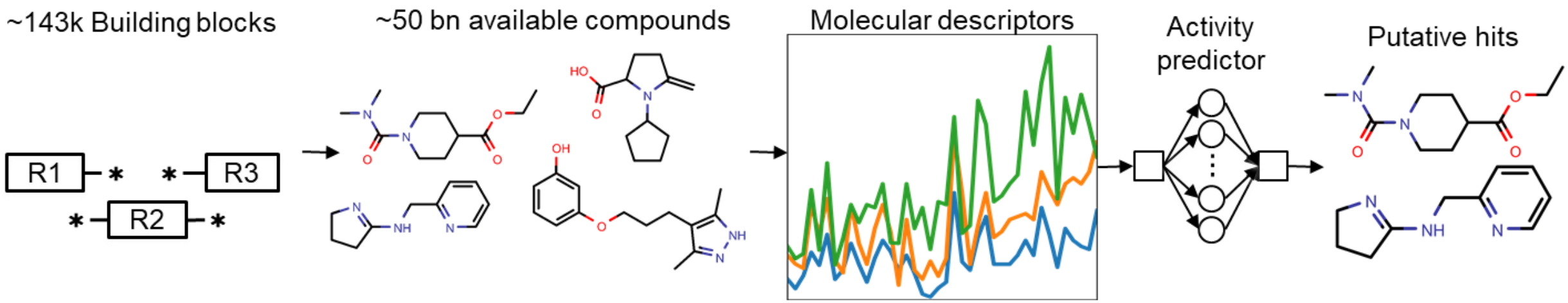
Advantages of SpiNNaker2 for Non-cognitive/non-AI Numerical Problems

meilerlab

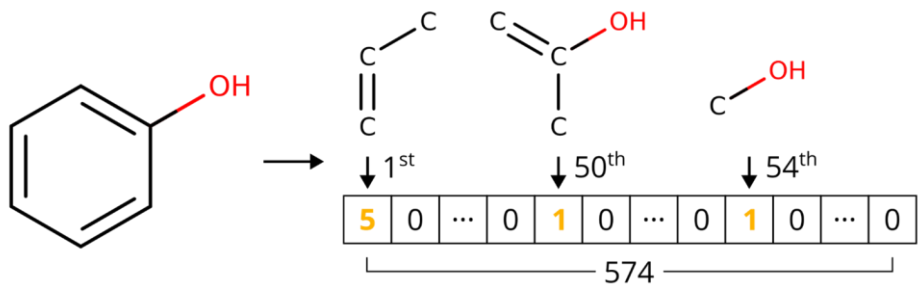
- Extreme parallelism of simple operations (think neurons...)
- (Search for) Sparse solutions in high-dimensional numerical rooms
- Stochastic computation/stochastic state representations
- Solving systems of locally coupled differential equations in a mesh/network topology (e.g. Neuron models, but also FEM and similar)

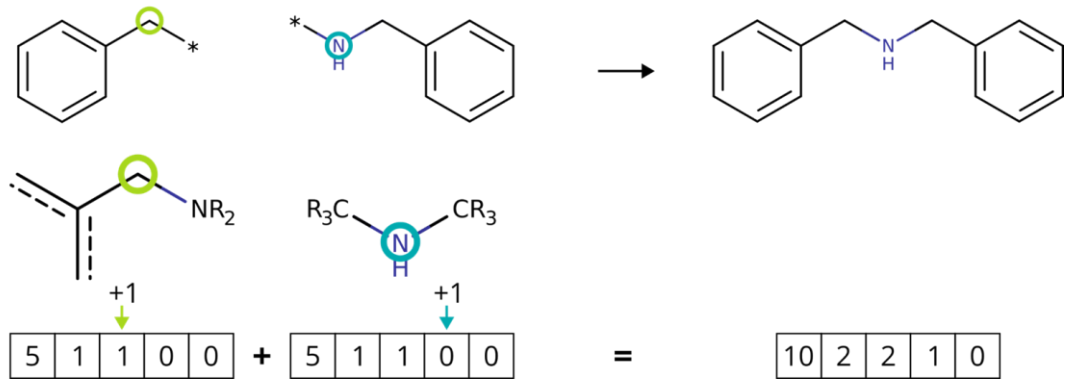
	HPC	SpiNNaker2	Quantum Computing
Parallelism	10^5 cores	10^{14} synaptic updates/msec	$>10^{25}$ quantum entanglements
Stochastic Computation	Only in software, 10^{10} stochastic decisions/sec	Hardware accelerators, 10^{17} stochastic decisions/sec	Inherent in Qubits, $>10^{30}$ stochastic decisions/sec
Sparsity in high-dimensional spaces	Not supported	Fully supported	Fully supported
FEM-type tesselations	10^5 elements, boundary condition updates µs to ms	10^7 elements in torus, boundary condition updates <10µs	Potentially very fast convergence, but tessalation limited to #Qubits: 10^2-10^3

Ultra-Large Library Screening on the SpiNNaker2 Platform



BCL::Mol2D descriptors from each atom's immediate neighborhood

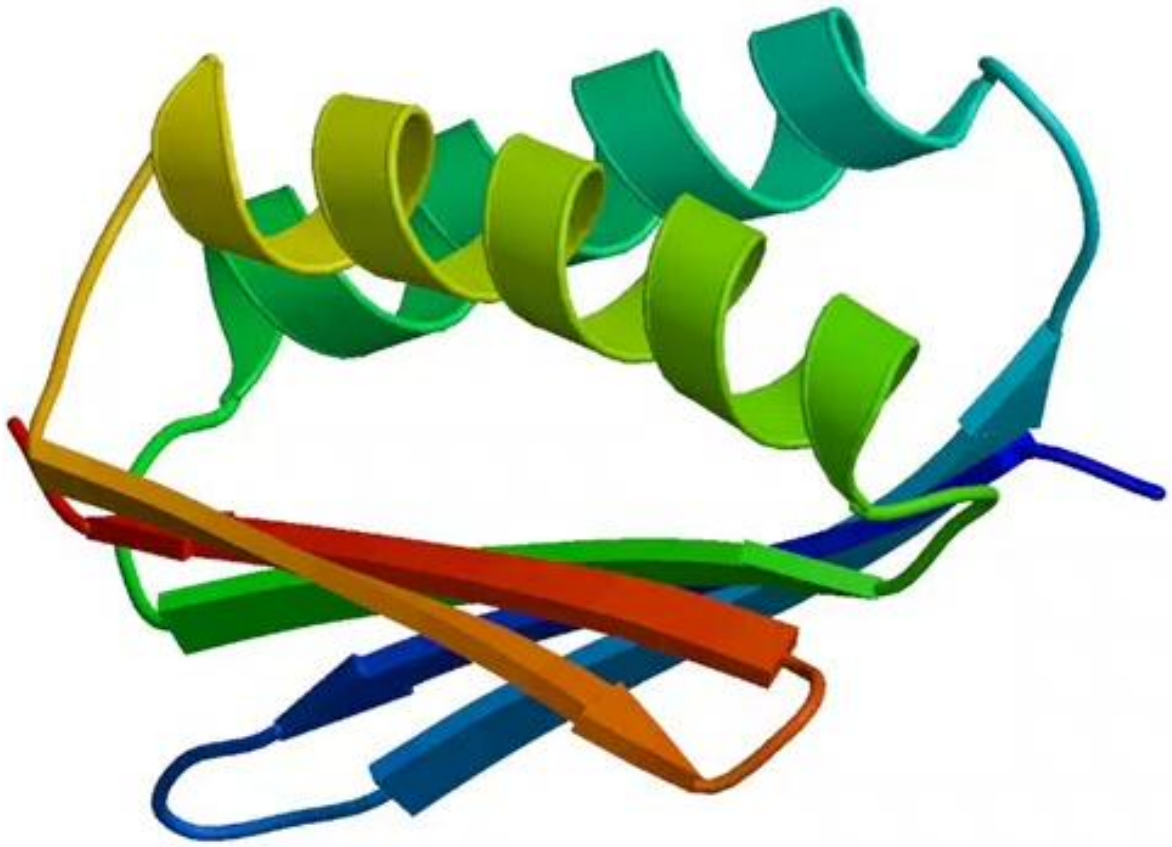


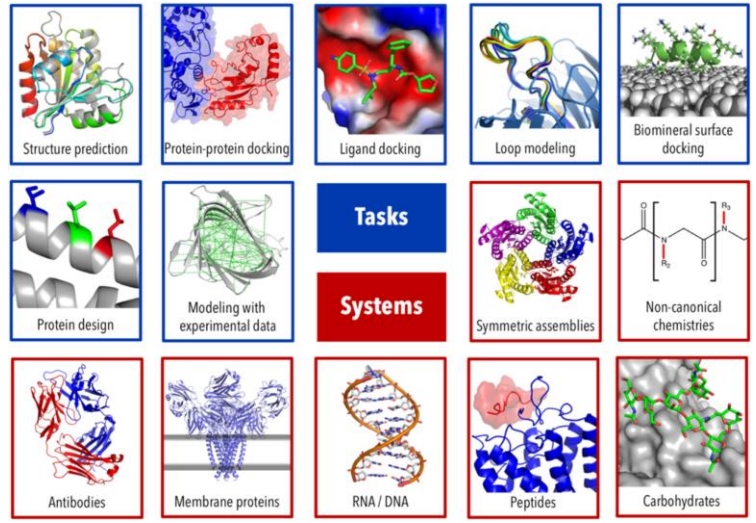


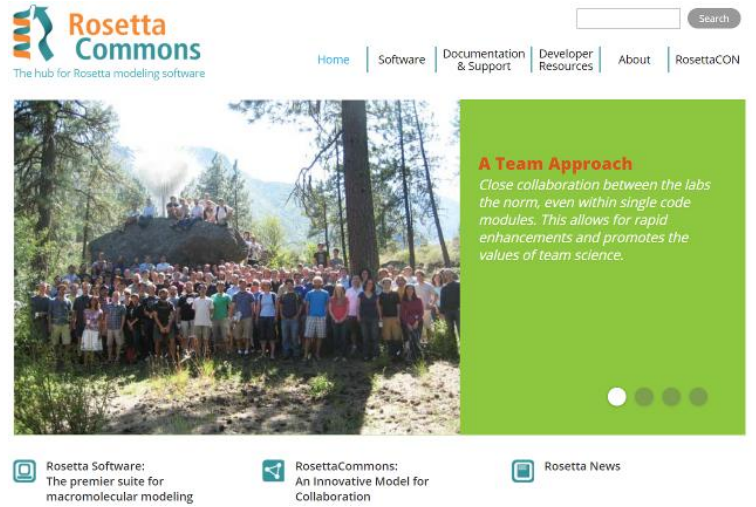
Vu, O., Mendenhall, J., Altarawy, D., & Meiler, J. (2019). BCL::Mol2D-a robust atom environment descriptor for QSAR modeling and lead optimization. *J Comput Aided Mol Des, 33(5), 477-486. https://doi.org/10.1007/s10822-019-00199-8*

Rosetta: A Unified Framework for Protein Structure Prediction and Design

meilerlab

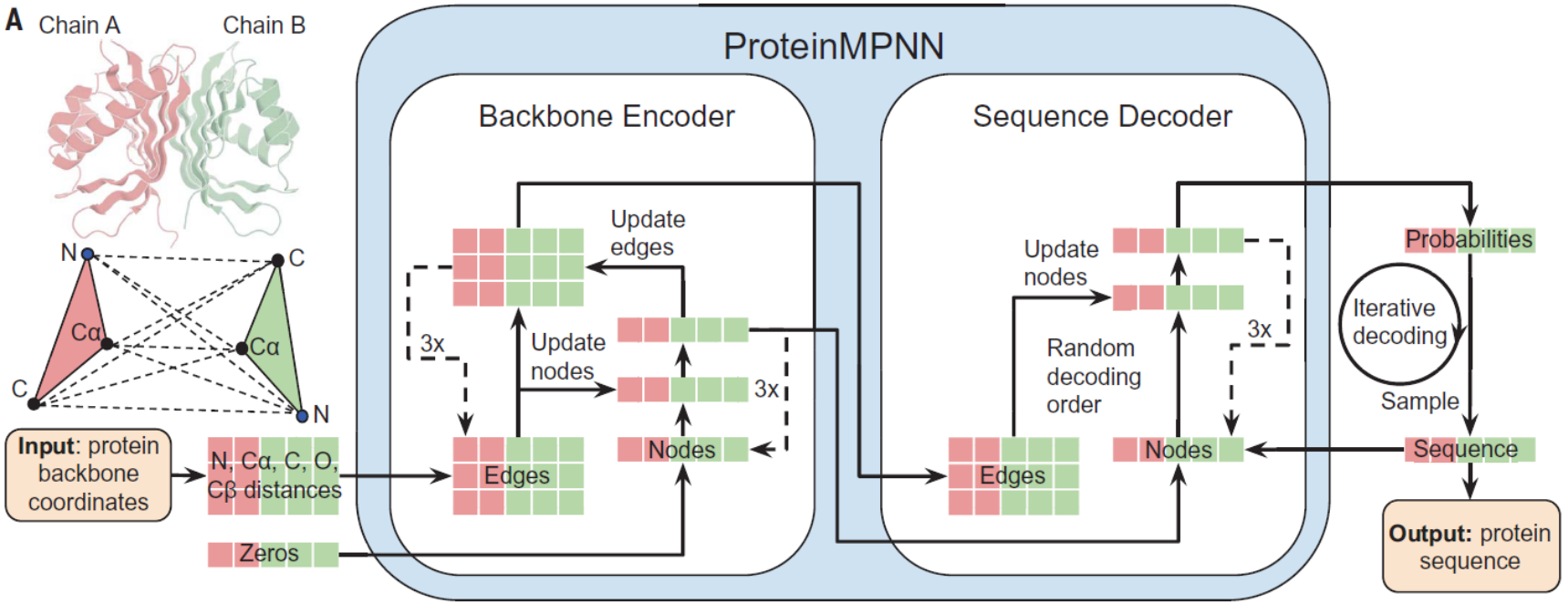






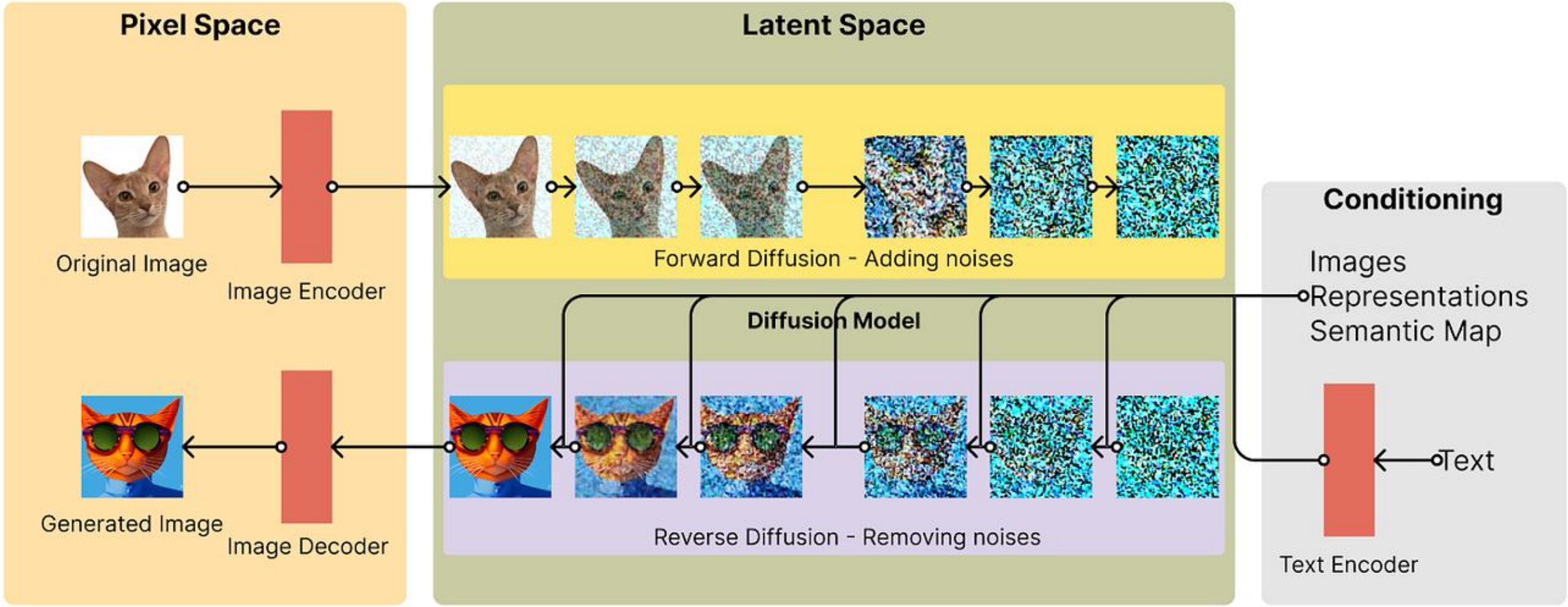
A. Leaver-Fay, et al.; "ROSETTA3: an object-oriented software suite …"; Methods Enzymol; 2011; Vol. 487 p. 545-74.
J. K. Leman, et al.; "Macromolecular modeling and design in Rosetta: recent methods and frameworks"; Nat Methods; 2020; Vol. 17 (7): p. 665-680.

Robust deep learning-based protein sequence design using ProteinMPNN

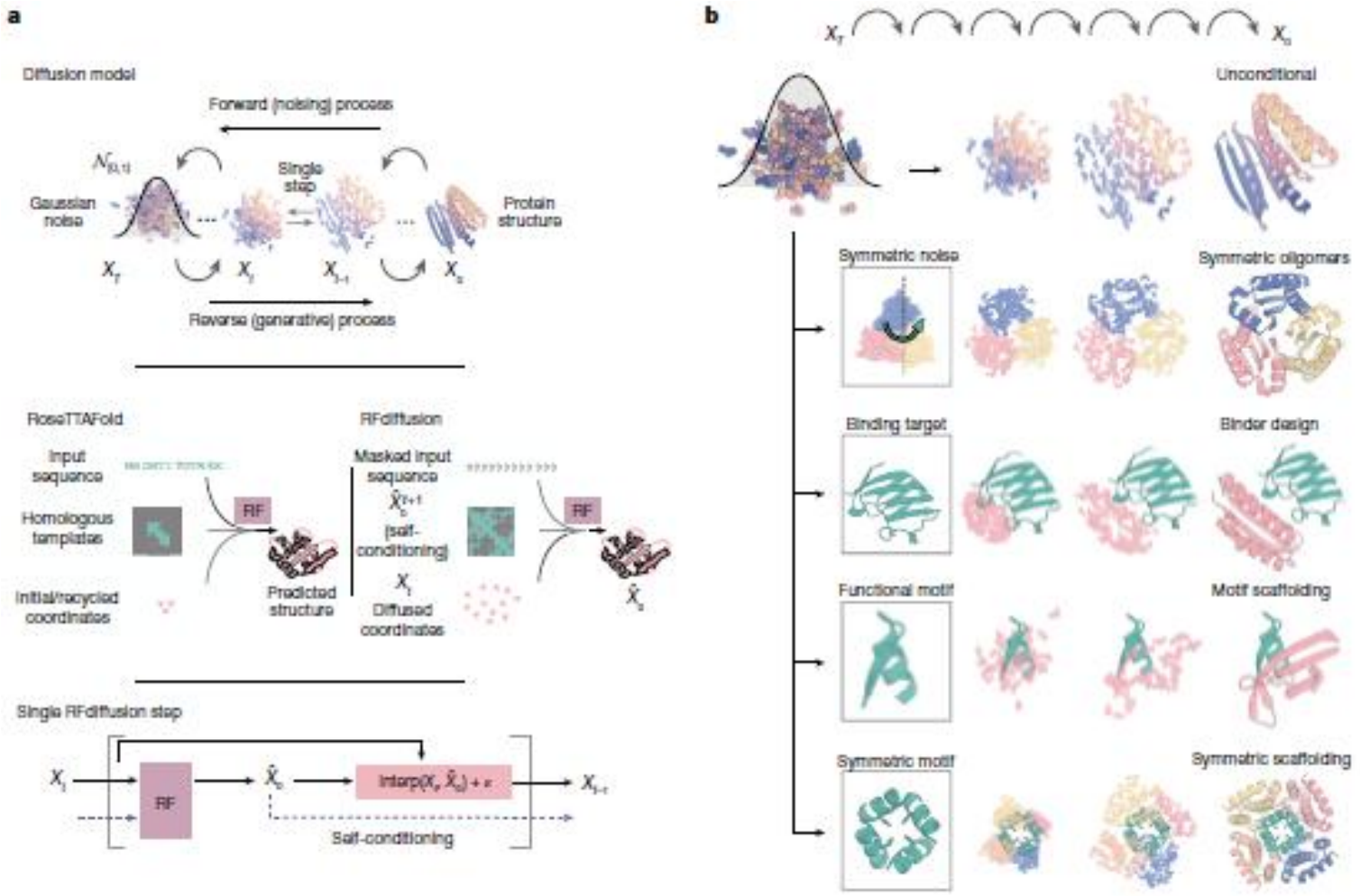


J. Dauparas, I. Anishchenko, N. Bennett, H. Bai, R. J. Ragotte, L. F. Milles, B. I. M. Wicky, A. Courbet, R. J. de Haas, N. Bethel, P. J. Y. Leung, T. F. Huddy, S. Pellock, D. Tischer, F. Chan, B. Koepnick, H. Nguyen, A. Kang, B. Sankaran, A. K. Bera, N. P. King and D. Baker; "Robust deep learning-based protein sequence design using ProteinMPNN"; *Science;* ***2022; Vol. 378 (6615): p. 49-56.***

How Stable Diffusion works in a Nutshell



De novo design of protein structure and function with RFDiffusion



Watson, J. L., Juergens, D., Bennett, N. R., Trippe, B. L., Yim, J., Eisenach, H. E., Ahern, W., Borst, A. J., Ragotte, R. J., Milles, L. F., Wicky, B. I. M., Hanikel, N., Pellock, S. J., Courbet, A., Sheffler, W., Wang, J., Venkatesh, P., Sappington, I., Torres, S. V., . . . Baker, D. (2023). De novo design of protein structure and function with RFdiffusion. *Nature, 620(7976), 1089-1100.*

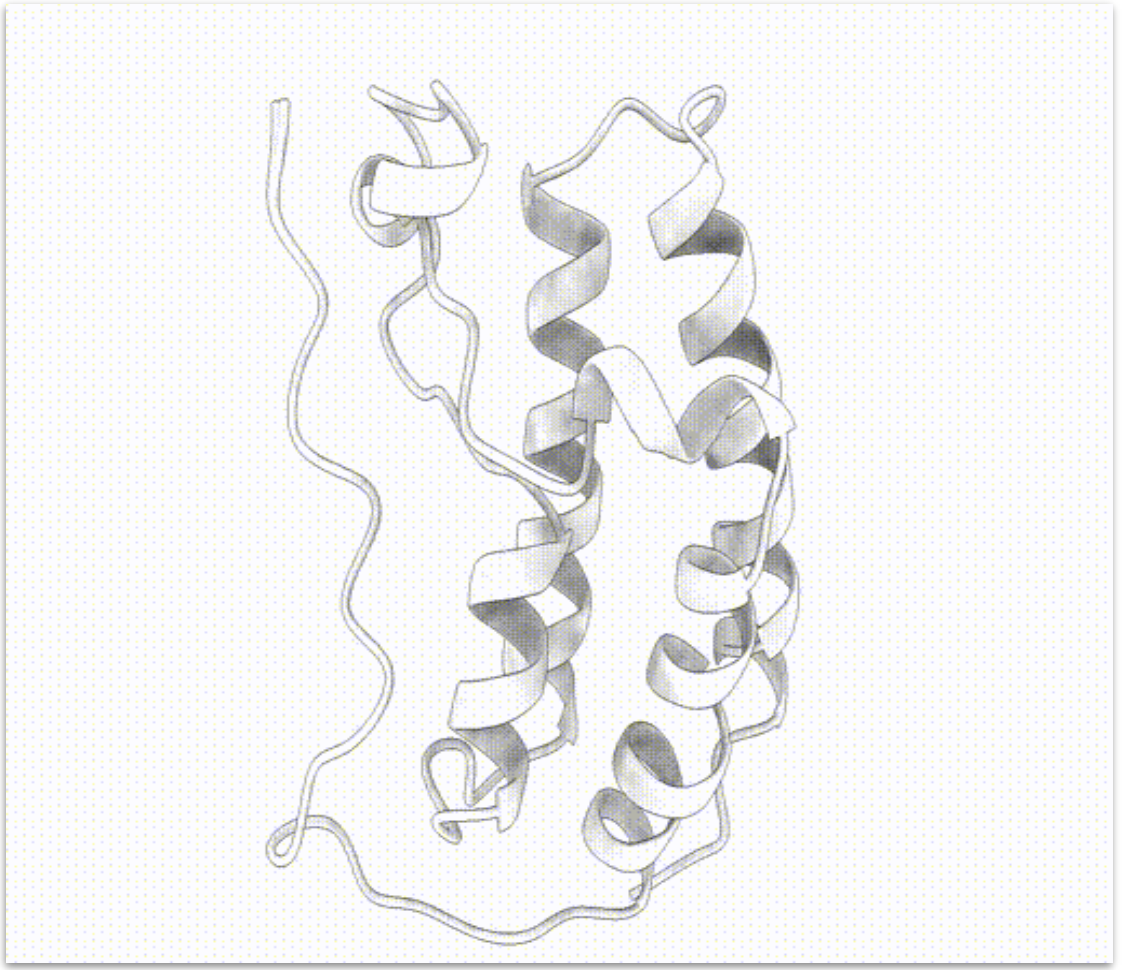
Modified Version of RFDiffusion for (Cyclic) Peptides

Normal positional embedding

0	1	2	3	4	5	6	7
-1	0	1	2	3	4	5	6
-2	-1	0	1	2	3	4	5
-3	-2	-1	0	1	2	3	4
-4	-3	-2	-1	0	1	2	3
-5	-4	-3	-2	-1	0	1	2
-6	-5	-4	-3	-2	-1	0	1
-7	-6	-5	-4	-3	-2	-1	0

Cyclic positional embedding

0	1	2	3	-4	-3	-2	-1
-1	0	1	2	3	-4	-3	-2
-2	-1	0	1	2	3	-4	-3
-3	-2	-1	0	1	2	3	-4
-4	-3	-2	-1	0	1	2	3
3	-4	-3	-2	-1	0	1	2
2	3	-4	-3	-2	-1	0	1
1	2	3	-4	-3	-2	-1	0



What does the Nobel Prize mean for 2025+? Integrating Chemical and Structural Biology!

meilerlab

Technology

- Expand the chemical space of AI design
- Learn to design protein dynamics for function
- Control at 0.1Å instead of 1.0Å

Application

- Medicine & Therapeutics
- Technology and Materials
- Sustainability

Acknowledgements

Current Members:

Alan Perez-Ratzke
Alexander Fürll
Alexander Zlobin
Ali Kanso
Ana Chang-Gonzalez
Andrea Pokrajcic
Anja Landsmann
Carie Fortenberry
Chris Jurich
Chris Moth
Christina Mercado
Claiborne Tydings
Clint Holt
Cristina Martina
David Jordan
Dieter Hoffmann
Dominic Rieger
Drew Neufer
Emma Web
Felipe Engelberger
Gustavo Araiza
Hannes Junker
Ivan Yushin
Jakob Riccabona
Jan Rossa
Jannis De Riz
Joe DeCorte
Johannes Klier
Johnny Siegert
Josuel Morel
Julia Beliaeva
Kaitlyn Ledwitch
Kelley Colopietro
Kuniko Hunter
Kyle Brown
Lance Liu
Lorenz Beckmann
Luisa Kaermer
Lukas v. Bredow
Marija Sorokina
Mateusz Sklodowski
Max Beining
Maximilian Hofmann
Paul Eisenhuth
Paul Kluge
Qianzhen Shao
Rajarshi SinhaRoy
Ratna Wolfram
Rocco Moretti
Sean Murray
Sevilay Gülesen
Shan Jiang
Sophia Hamza
Sophie Maclure
Te-Yu Kao
Thomas Scott
Tom Schlegel
Toni Oesterreich
Onur Elaldi
Vivian Ehrlich

Selected Alumni:

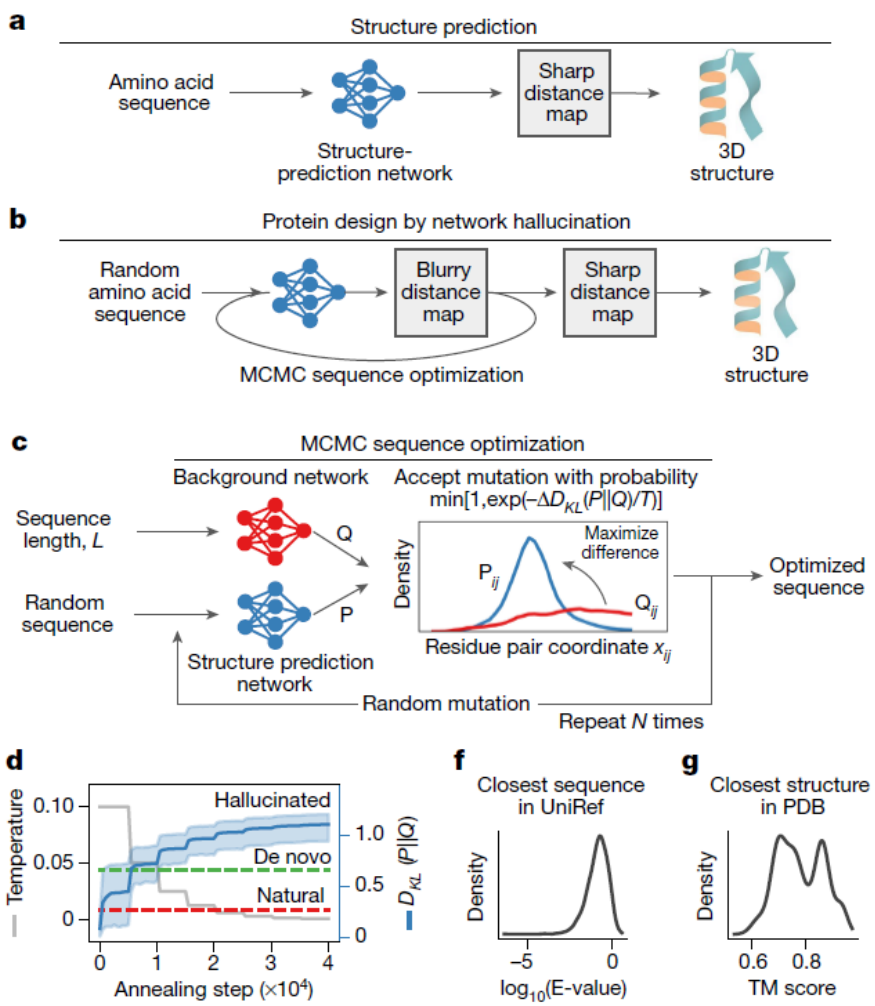
Abdullah Al Mamun
Alican Gulsevin
Annalen Bleckmann
Alex Sevy
Axel Fischer
Benjamin Brown
Brett Kroncke
Brian Bender
Brian Weiner
Clara Schoeder
David Nannemann
Diego del Alamo
Eli McDonald
Fabian Liessmann
Georg Kuenze
Hope Woods
Liz Dong Nguyen
Moritz Ertelt
Nils Woetzel
Steffen Lindert

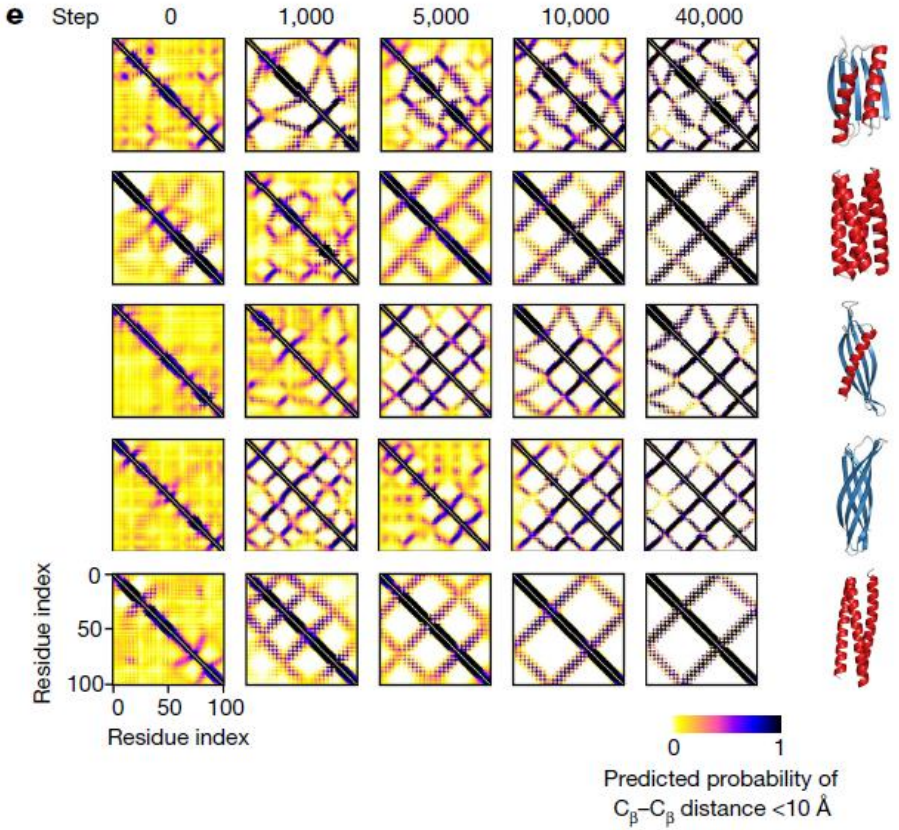
Selected Collaborators:

Jim Crowe, Craig Lindsley, Brian Shoichet, Hassane Mchaourab, Chuck Sanders,Al George, Lars Plate, Jonathan Schlebach, Annette Beck-Sickinger, Clara Schoeder, …

Funding: NIH NIAID U01 AI150739, NIH NILBI R01HL122010, NIH NIDA R01DA046138, NIH NLM LM013434, ARPA-H, DFG CRC 1423, DFG CRC 1052, DFG TRR 386, DFG SPP 2363, DFG SPP 3465, DAAD 57616814, BMBF, CEPI, Alexander von Humboldt Foundation
BAYER, BIONTECH, BASF

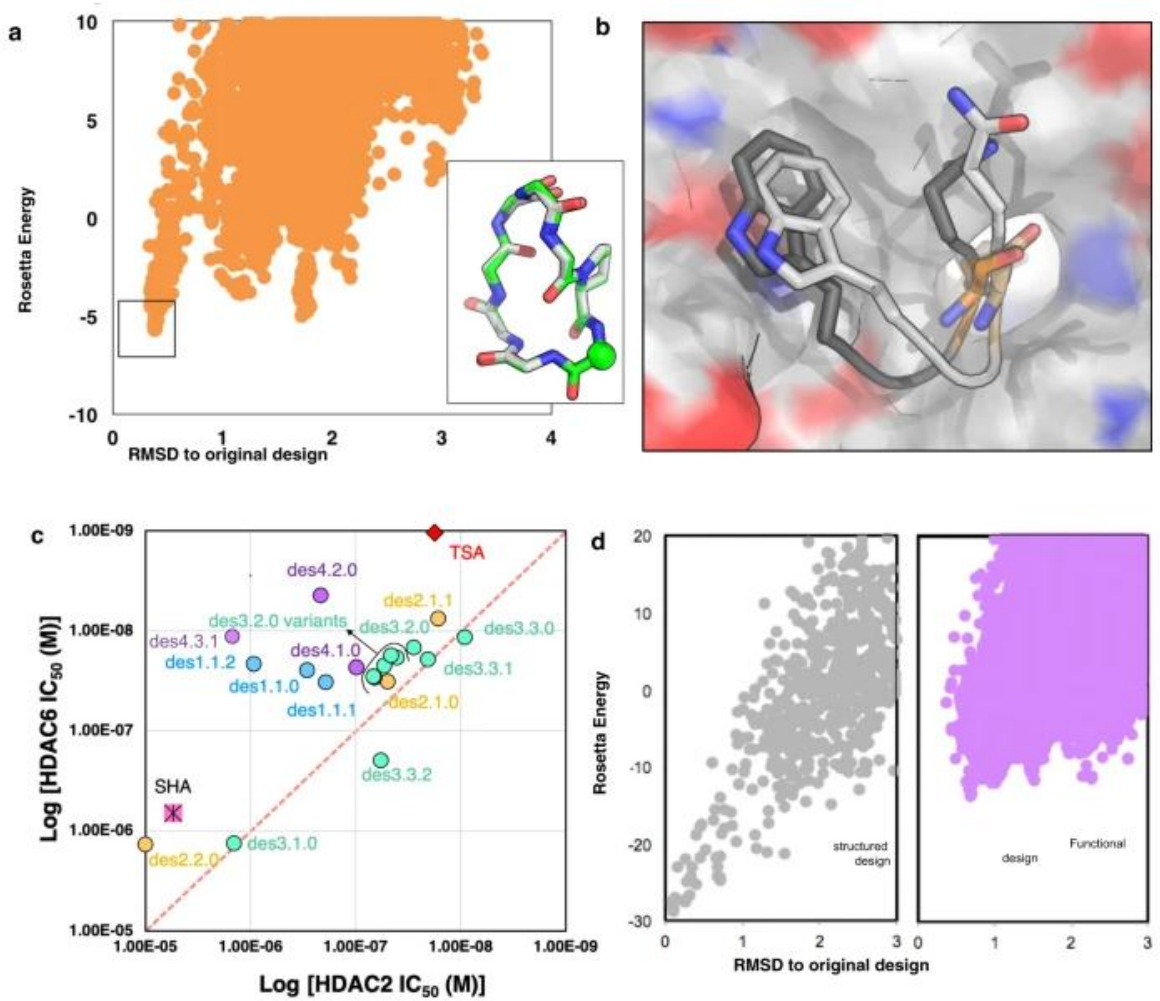
De novo Protein Design by Deep Network Hallucination





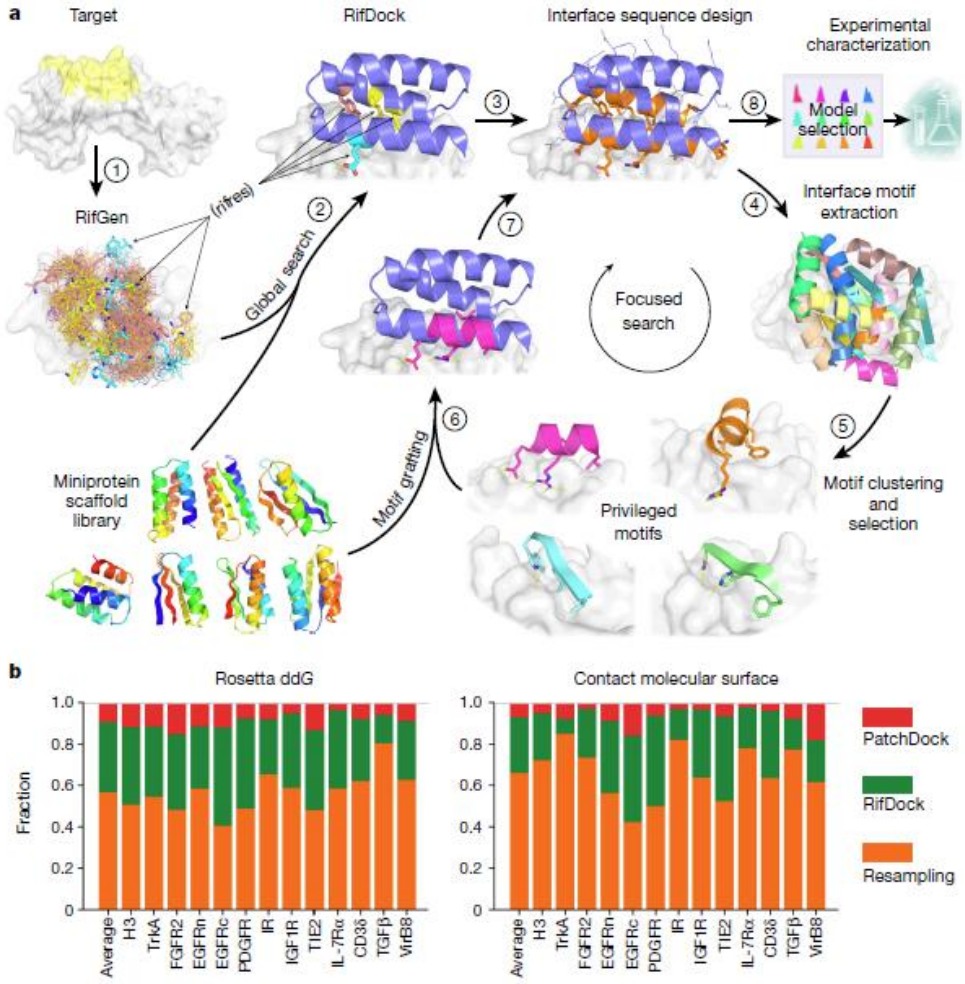
I. Anishchenko, S. J. Pellock, T. M. Chidyausiku, T. A. Ramelot, S. Ovchinnikov, J. Hao, K. Bafna, C. Norn, A. Kang, A. K. Bera, F. DiMaio, L. Carter, C. M. Chow, G. T. Montelione and D. Baker; "De novo protein design by deep network hallucination"; *Nature;* ***2021; Vol. 600 (7889): p. 547-552.***

Anchor extension: a structure-guided approach to design cyclic peptides



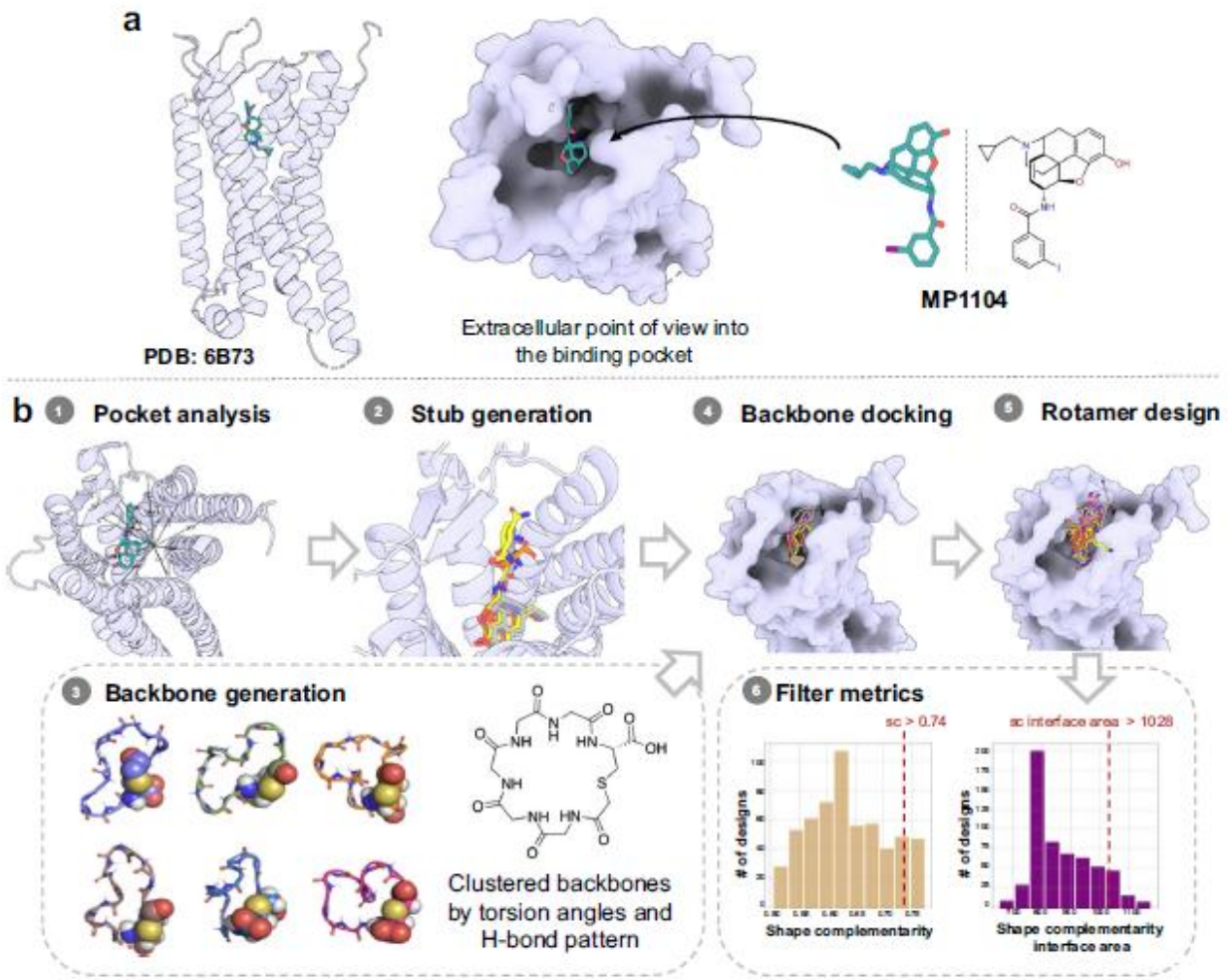
P. Hosseinzadeh, P. R. Watson, T. W. Craven, X. Li, S. Rettie, F. Pardo-Avila, A. K. Bera, V. K. Mulligan, P. Lu, A. S. Ford, B. D. Weitzner, L. J. Stewart, A. P. Moyer, M. Di Piazza, J. G. Whalen, P. J. Greisen, D. W. Christianson and D. Baker; "Anchor extension: a structure-guided approach to design cyclic peptides targeting enzyme active sites"; *Nat Commun;* ***2021; Vol. 12 (1): p. 3384.***

Design of protein-binding proteins from the target structure alone



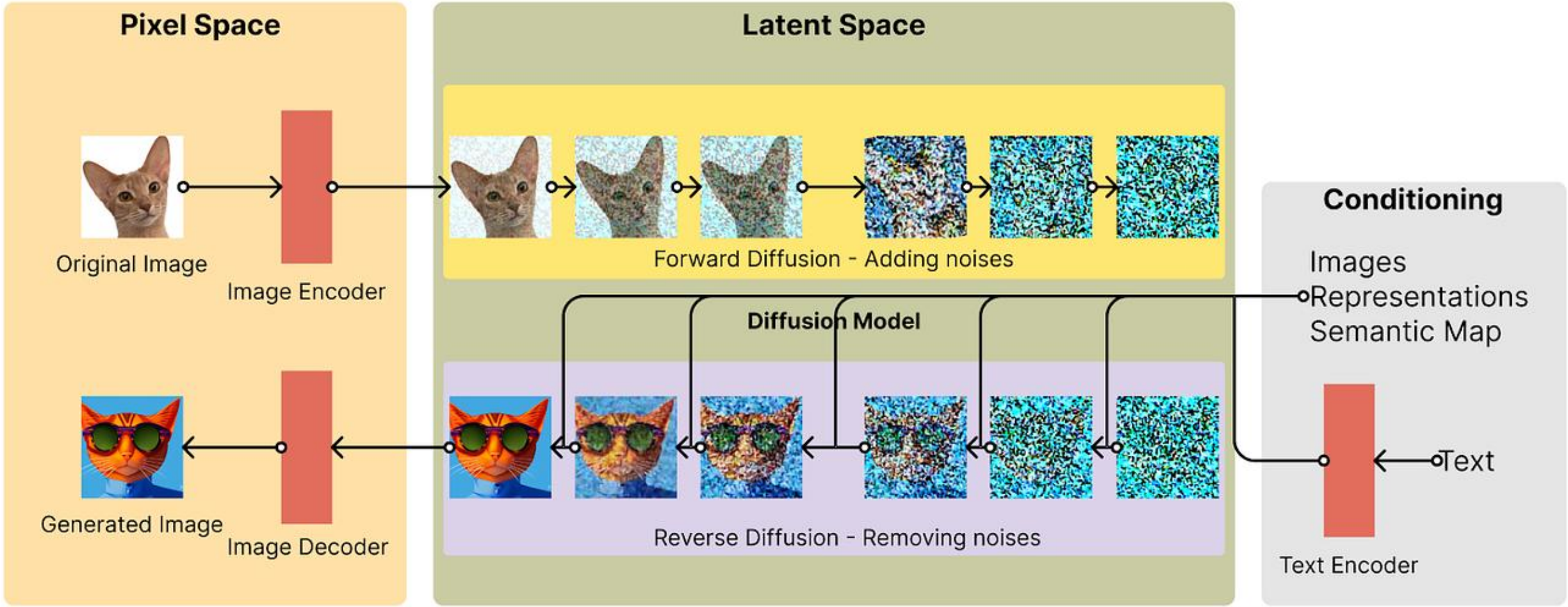
Cao, L. X., Coventry, B., Goreshnik, I., Huang, B. W., Sheffler, W., Park, J. S., Jude, K. M., Markovic, I., Kadam, R. U., Verschueren, K. H. G., Verstraete, K., Walsh, S. T. R., Bennett, N., Phal, A., Yang, A., Kozodoy, L., DeWitt, M., Picton, L., Miller, L., . . . Baker, D. (2022). Design of protein-binding proteins from the target structure alone. *Nature, 605(7910), 551-+.*

Design of peptide-drug conjugate ligands of the kappa-opioid receptor



Muratspahic, E., Deibler, K., Han, J., Tomasevic, N., Jadhav, K. B., Olive-Marti, A. L., Hochrainer, N., Hellinger, R., Koehbach, J., Fay, J. F., Rahman, M. H., Hegazy, L., Craven, T. W., Varga, B. R., Bhardwaj, G., Appourchaux, K., Majumdar, S., Muttenthaler, M., Hosseinzadeh, P., . . . Gruber, C. W. (2023). Design and structural validation of peptide-drug conjugate ligands of the kappa-opioid receptor. *Nat Commun, 14(1), 8064.*

How Stable Diffusion works in a Nutshell



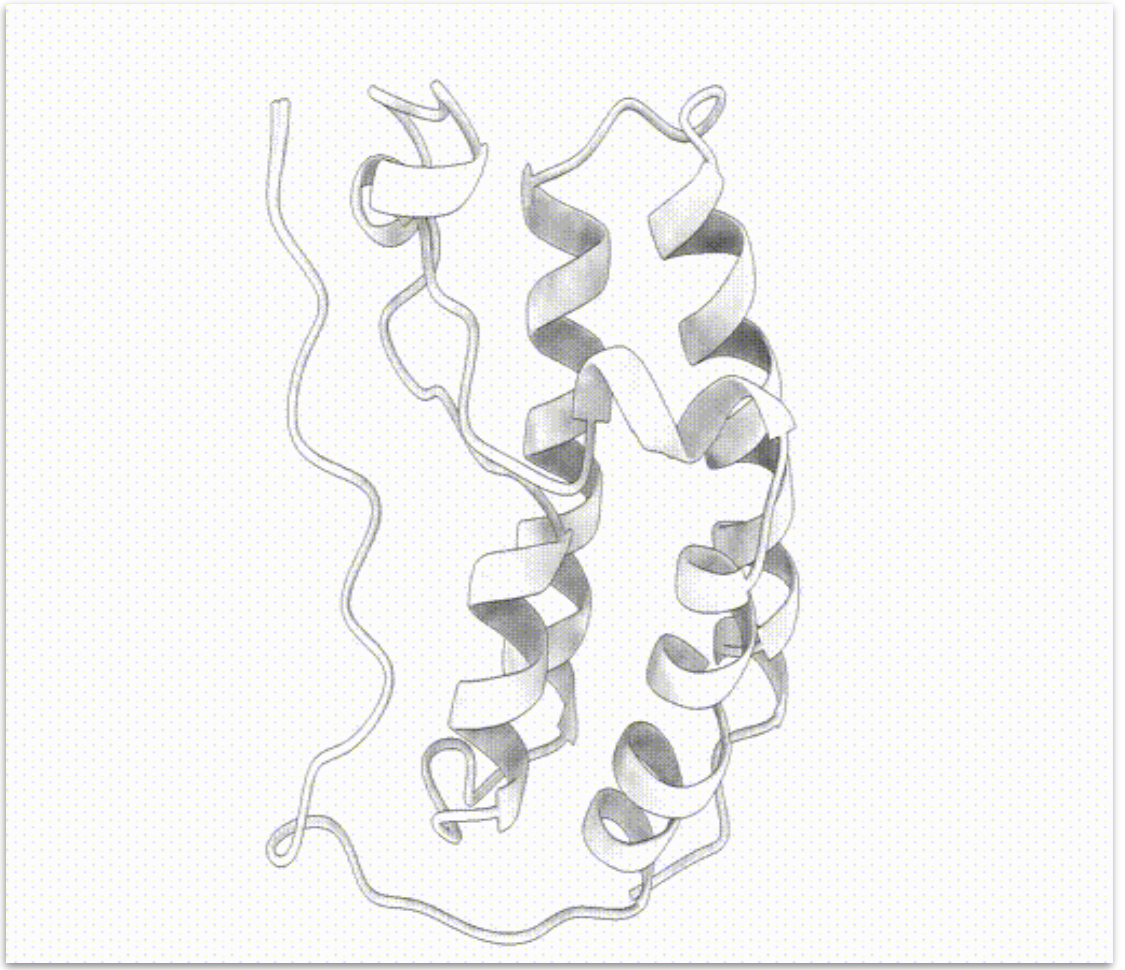
Modified Version of RFDiffusion for (Cyclic) Peptides

Normal positional embedding

0	1	2	3	4	5	6	7
-1	0	1	2	3	4	5	6
-2	-1	0	1	2	3	4	5
-3	-2	-1	0	1	2	3	4
-4	-3	-2	-1	0	1	2	3
-5	-4	-3	-2	-1	0	1	2
-6	-5	-4	-3	-2	-1	0	1
-7	-6	-5	-4	-3	-2	-1	0

Cyclic positional embedding

0	1	2	3	-4	-3	-2	-1
-1	0	1	2	3	-4	-3	-2
-2	-1	0	1	2	3	-4	-3
-3	-2	-1	0	1	2	3	-4
-4	-3	-2	-1	0	1	2	3
3	-4	-3	-2	-1	0	1	2
2	3	-4	-3	-2	-1	0	1
1	2	3	-4	-3	-2	-1	0



Acknowledgements

Current Members:

Alan Perez-Ratzke
Alexander Fürll
Alexander Zlobin
Ali Kanso
Ana Chang-Gonzalez
Anja Landsmann
Amy (Minh) Tran
Carie Fortenberry
Chris Jurich
Chris Moth
Christian Schmidt
Christina Mercado
Claiborne Tydings
Cristina Martina
Dieter Hoffmann
Dominik Rieger
Drew Neufer
Emma Web
Eric Bell
Fabian Liessmann
Felipe Engelberger
Gustavo Araiza
Hannes Junker
Ivan Yushin
Jakob Riccabona
Jan Rossa
Jannis De Riz
Joe DeCorte
Johannes Klier
Johnny Siegert
Jordan Davis
Julia Beliaeva
Kaitlyn Ledwitch
Kelley Colopietro
Kuniko Hunter
Kyle Brown
Lance Liu
Lorenz Beckmann
Luisa Kaermer
Lukas v. Bredow
Marija Sorokina
Markus Fischer
Mateusz Sklodowski
Max Beining
Maximilian Hofmann
Paul Eisenhuth
Paul Kluge
Qianzhen Shao
Rajarshi SinhaRoy
Ratna Wolfram
Rocco Moretti
Sean Murray
Sevilay Gülesen
Shan Jiang
Sophie Maclure
Stefan Arnold
Te-Yu Kao
Thomas Scott
Tom Schlegel
Toni Oesterreich
Tracy (Yidan) Tang
Onur Elaldi
Victoria Most
Vivian Ehrlich
Zhaoqian Su

Selected Alumni:

Abdullah Al Mamun
Alican Gulsevin
Annalen Bleckmann
Alex Sevy
Axel Fischer
Benjamin Brown
Brett Kroncke
Brian Bender
Brian Weiner
Clara Schoeder
David Nannemann
Diego del Alamo
Eli McDonald
Georg Kuenze
Hope Woods
Jordon Willis
Julia Koehler Leman
Liz Dong Nguyen
Moritz Ertelt
Nils Woetzel
Steffen Lindert

Selected Collaborators:

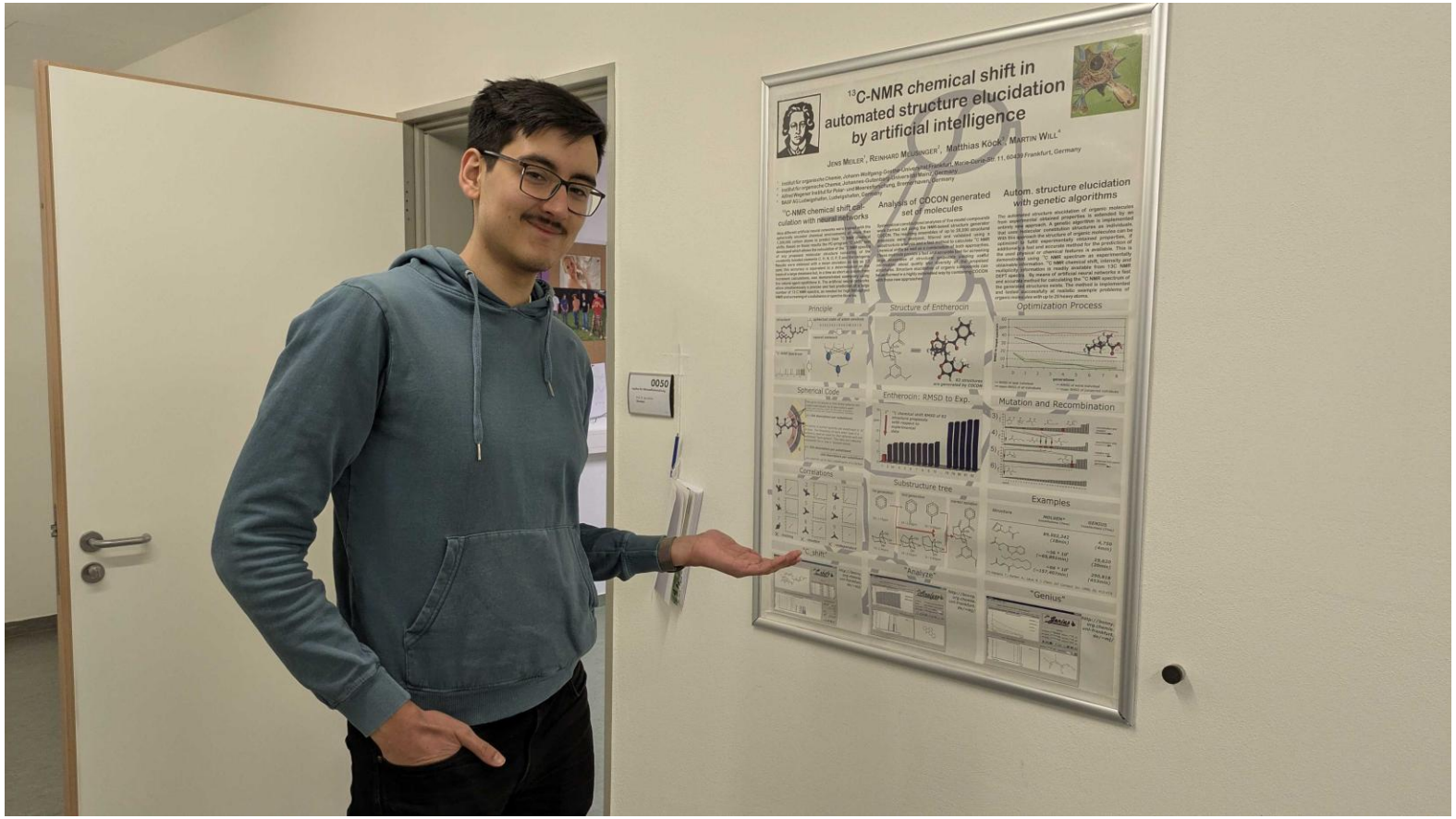
Jim Crowe, Heidi Hamm, Craig Lindsley, Brian Shoichet, Hassane Mchaourab, Chuck Sanders, Lars Plate, Jonathan Schlebach, Annette Beck-Sickinger, Clara Schoeder, …

Funding: NIH NIAID U19 AI117905, NIH NILBI R01HL122010, NIH NIDA R01DA046138, NIH NLM LM013434, ARPA-H

DFG CRC 1423, DFG CRC 1052, DFG TRR 386, DAAD 57616814, BMBF, CEPI, Alexander von Humboldt Foundation

BAYER, BIONTECH, BASF

We are doing this for 30+ Years
The Technology is ready for new applications



Medicine & Therapeutics

- pandemic preparedness
- cancer & immunotherapy
- neurodegeneration

Technology and Materials

- chemical sensors
- bio-electronic devices
- drug delivery

Sustainability

- breaking down plastic
- artificial photosynthesis
- nanoscale manufacturing