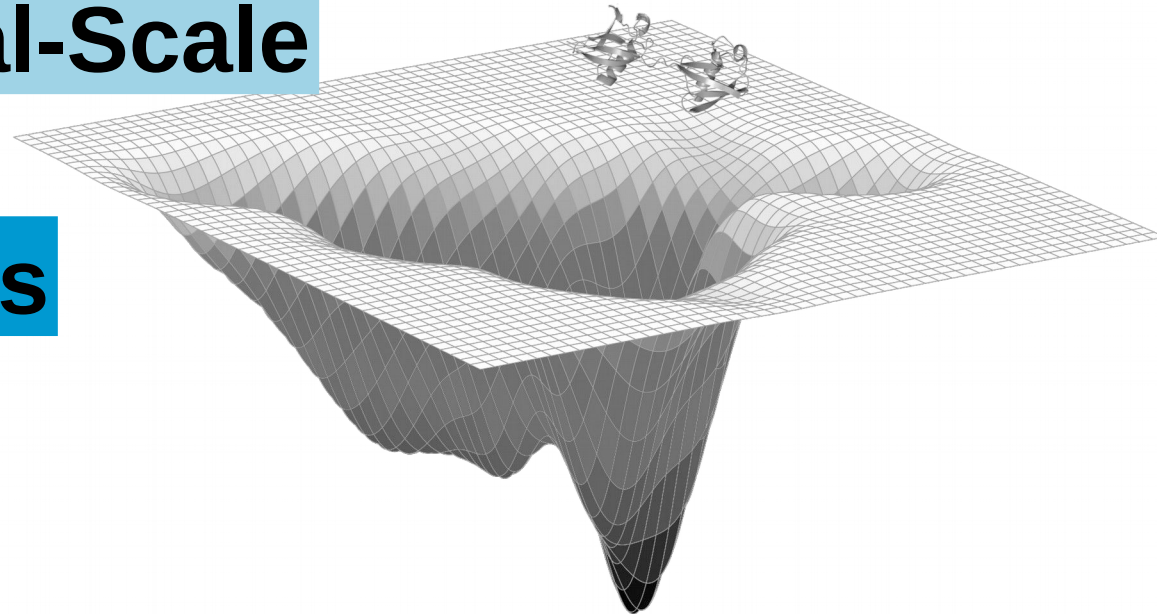


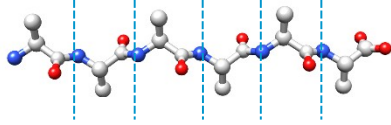
Conformational Analysis of Dual-Scale Simulations of Ubiquitin Chains



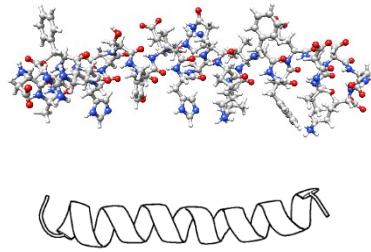
Andrej Berg

27.02.2020

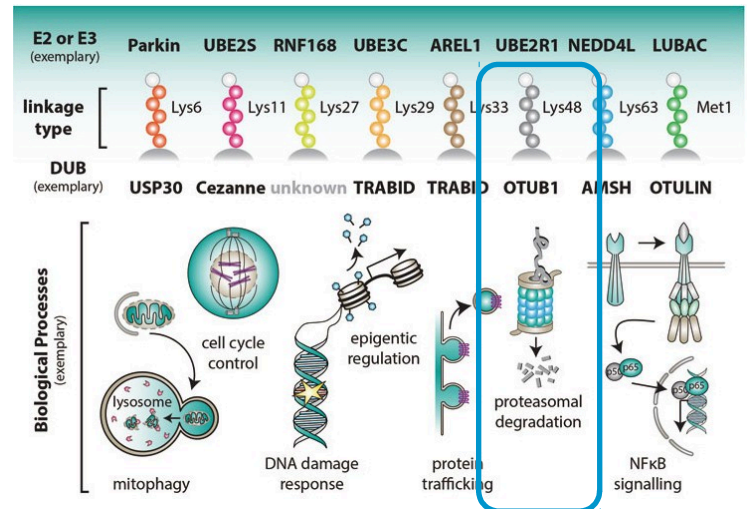
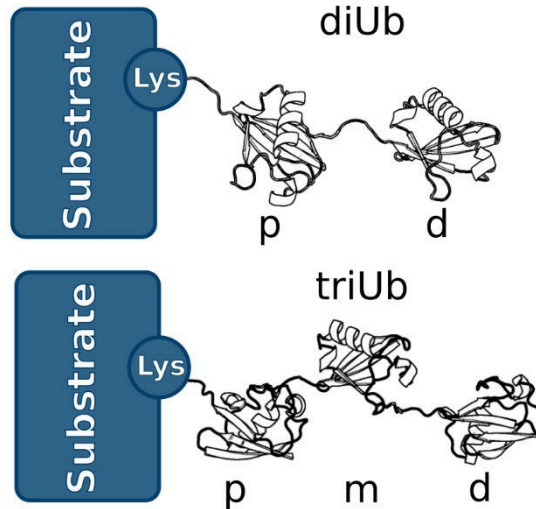
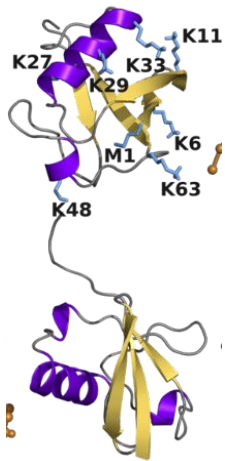
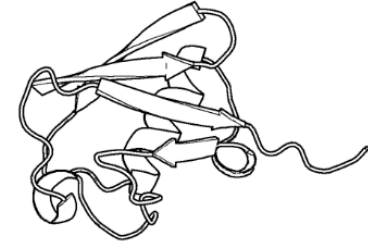
6x alanine



peptide (28 AS)

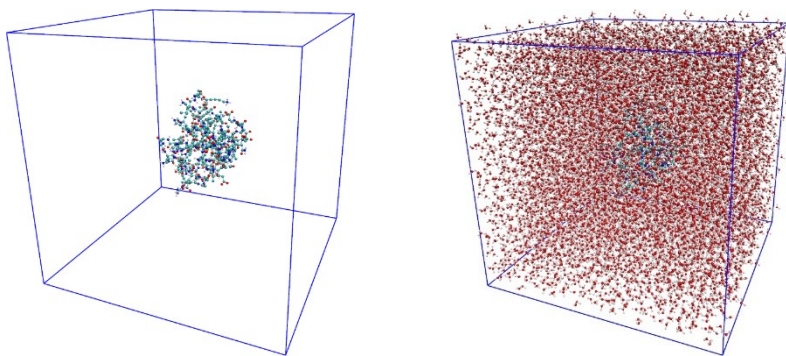


ubiquitin (76 AS)



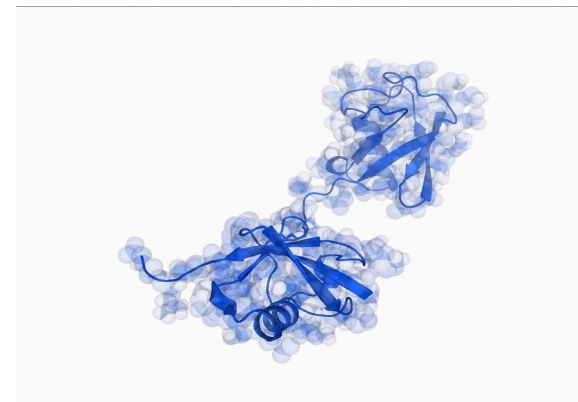
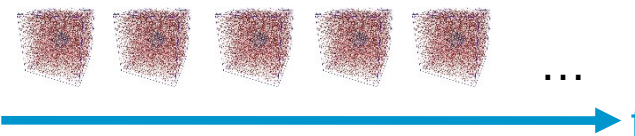
Swatek et al., Cell Research (2016)

initial (x, v)

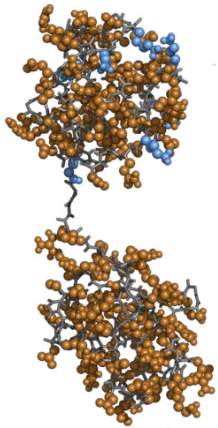


MD

trajectory

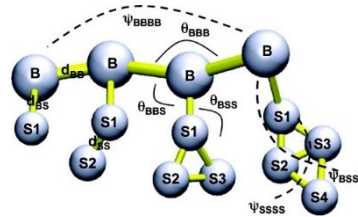
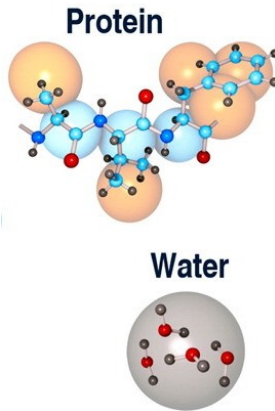


GROMOS54a7

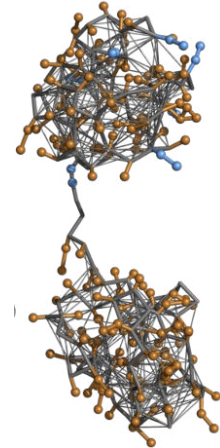


$N = 1521$
(+60 000)

MARTINI CG model



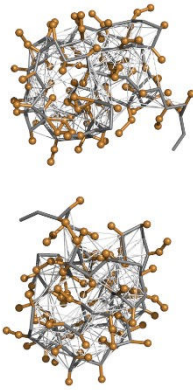
polyUb CG model



$N = 326$
(+5 000)

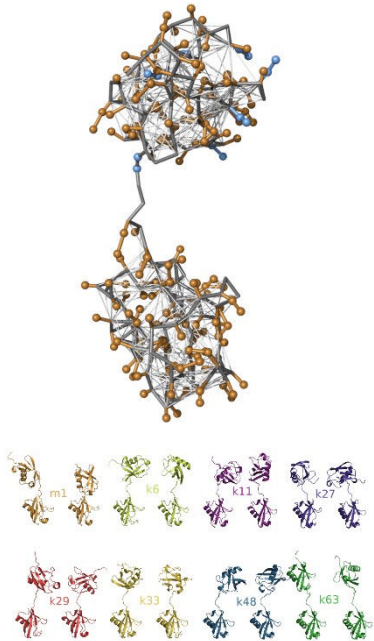
Monticelli et al., JCTC (2008)
Kmiecik et al., Chemical Reviews (2016)
Berg et al., PLOS Comp. Bio. (2018)

2xUb
(unlinked)



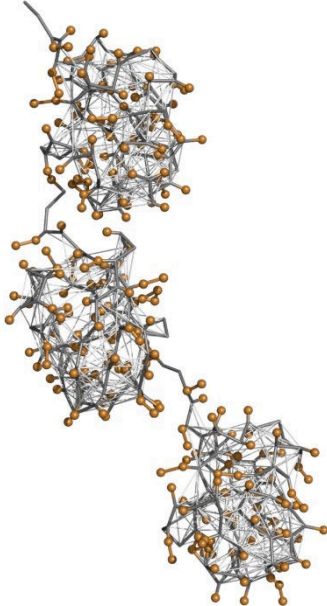
80 x 10 μ s
8*10¹⁰ steps

diUb

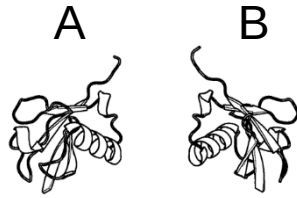


8 x 20 x 10 μ s
16*10¹⁰ steps

triUb



8 x 20 x 100 μ s
1.6*10¹² steps

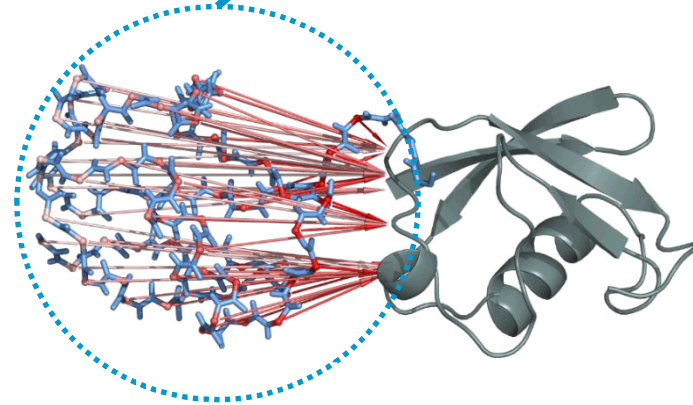


$$D_{A,B} = \begin{matrix} & x_1 & \dots & x_m \\ y_1 & \left(\begin{matrix} d_{a_1,b_1} & \dots & d_{a_1,b_m} \\ \vdots & & \vdots \\ d_{a_n,b_1} & \dots & d_{a_n,b_m} \end{matrix} \right) \\ \vdots & & & \\ y_n & & & \end{matrix}$$

$$A_B = [\min(y_1), \dots, \min(y_n)]$$

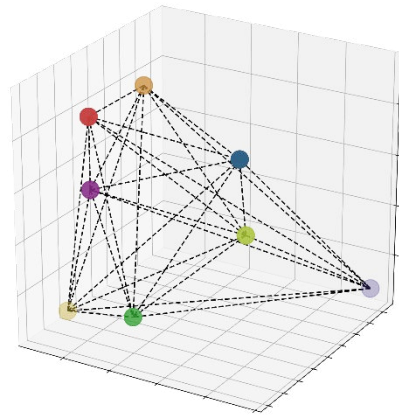
$$B_A = [\min(x_1), \dots, \min(x_m)]$$

$$RMD_{n+m} = A_B B_A$$



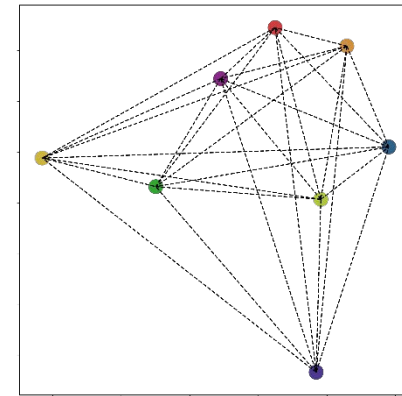
2xUb/diUb: **144** dimensions

MDS (multi-dimensional scaling)



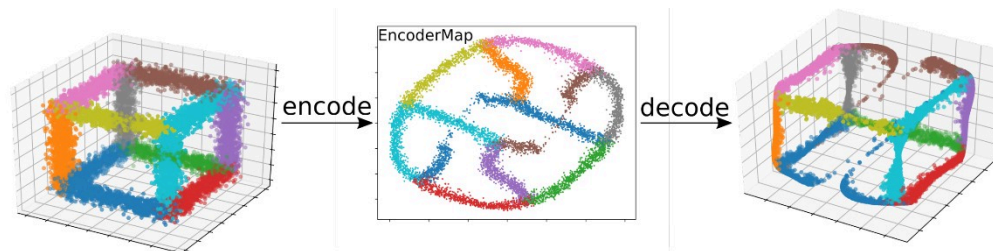
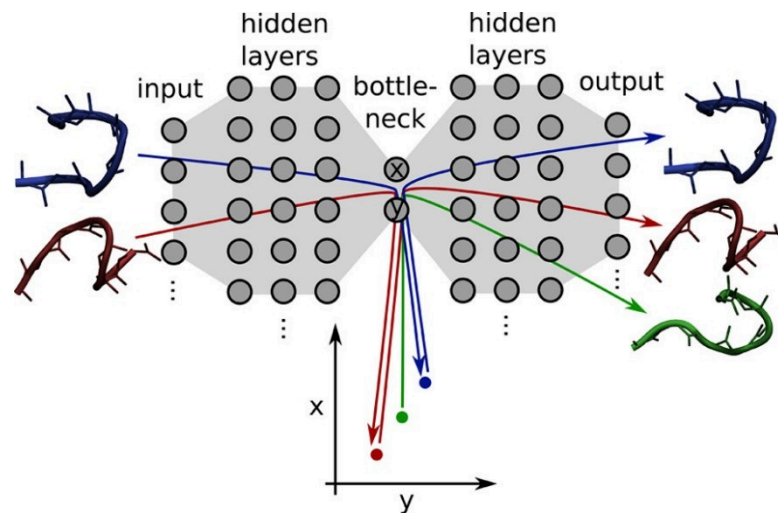
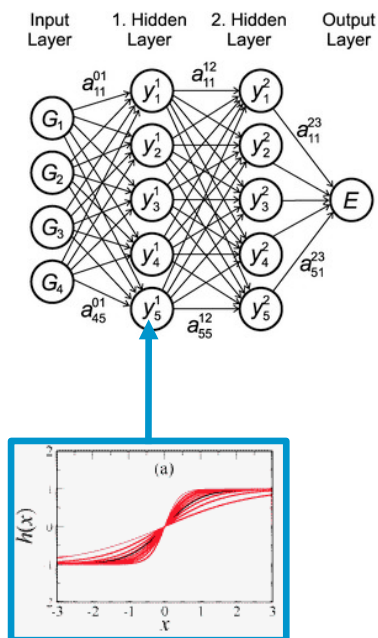
3D

$$S(D, d) = \sum_{i>j} (D_{ij} - d_{ij})^2$$



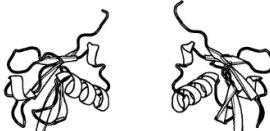
2D

Pedregosa et al., JMLR (2011)

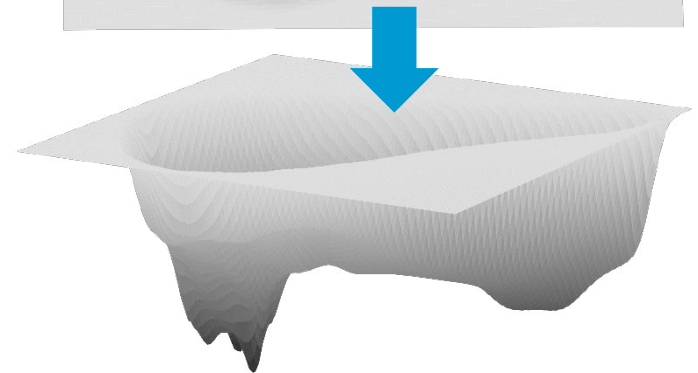
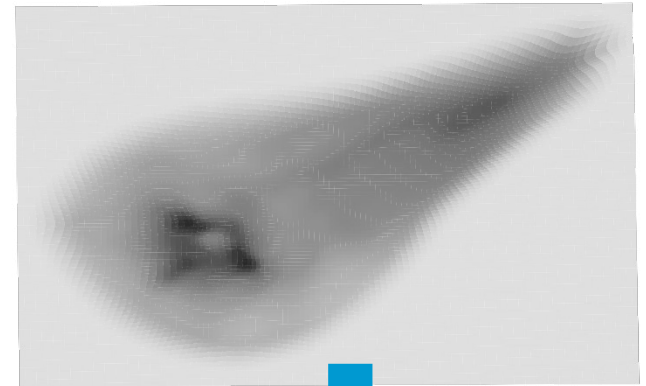
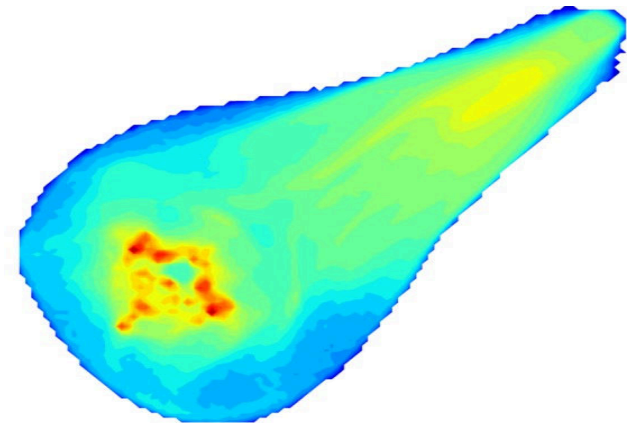
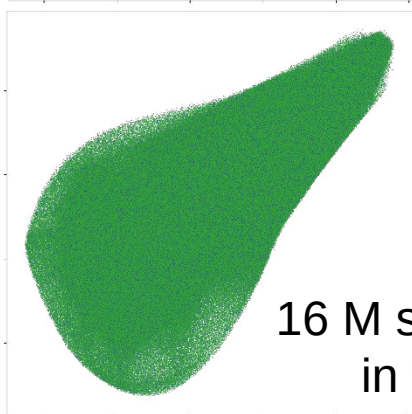
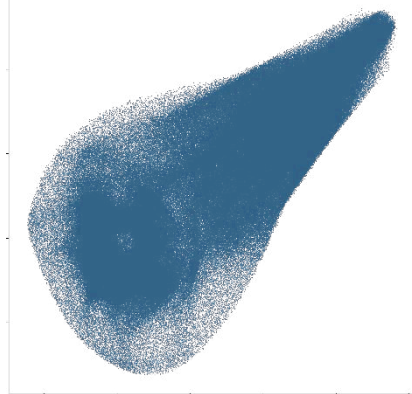
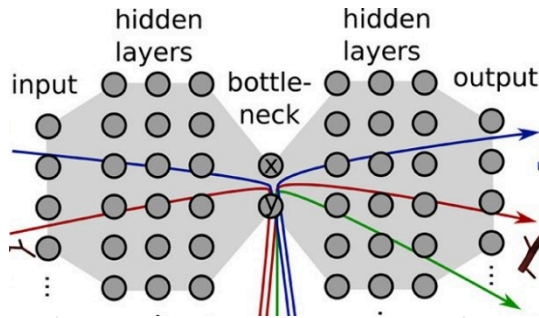


Behler, JCP (2011)
 Lemke & Peter, JCTC (2019)
 Lemke et al., JCIM (2019)
<https://github.com/AG-Peter/encodemap>

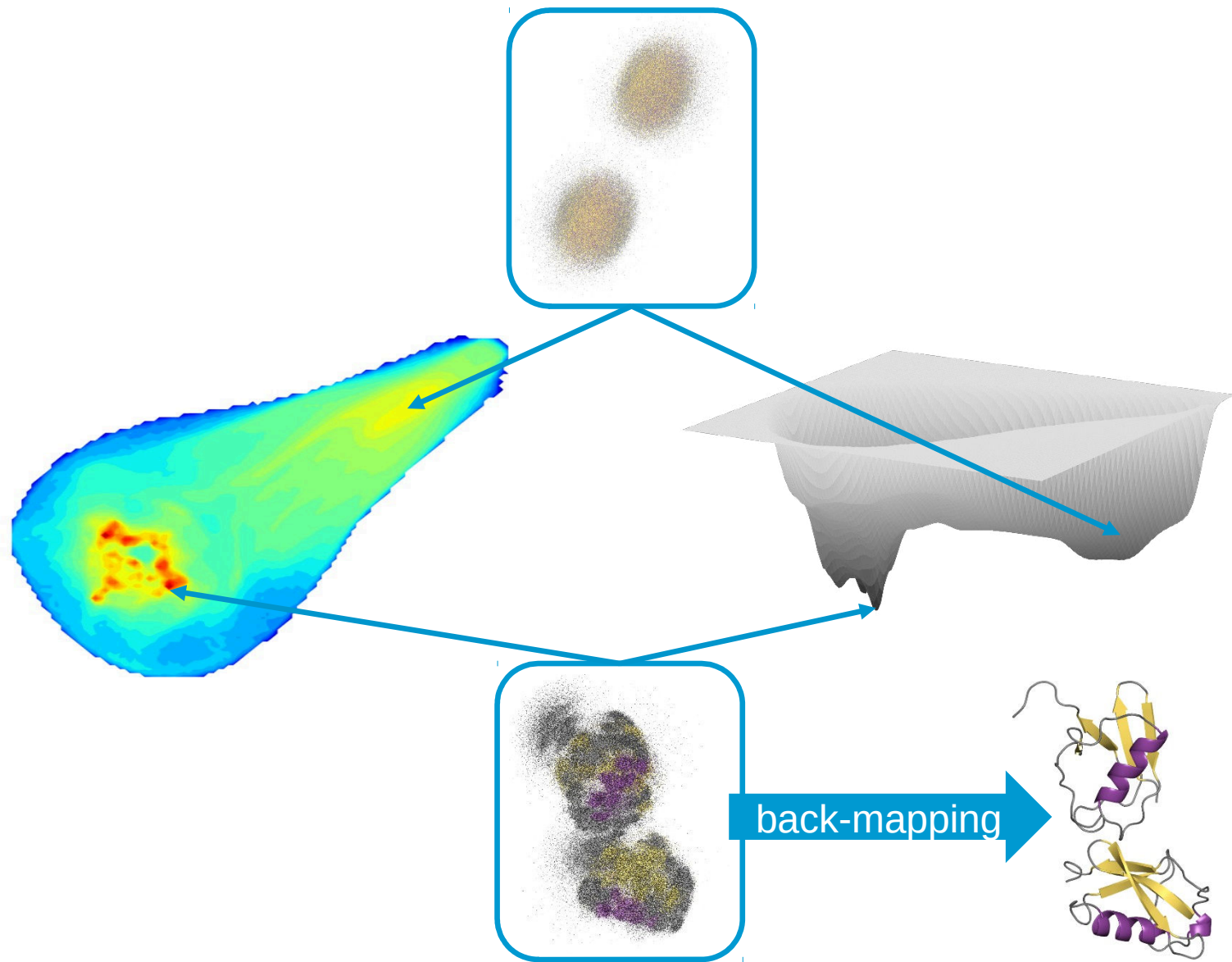
training set

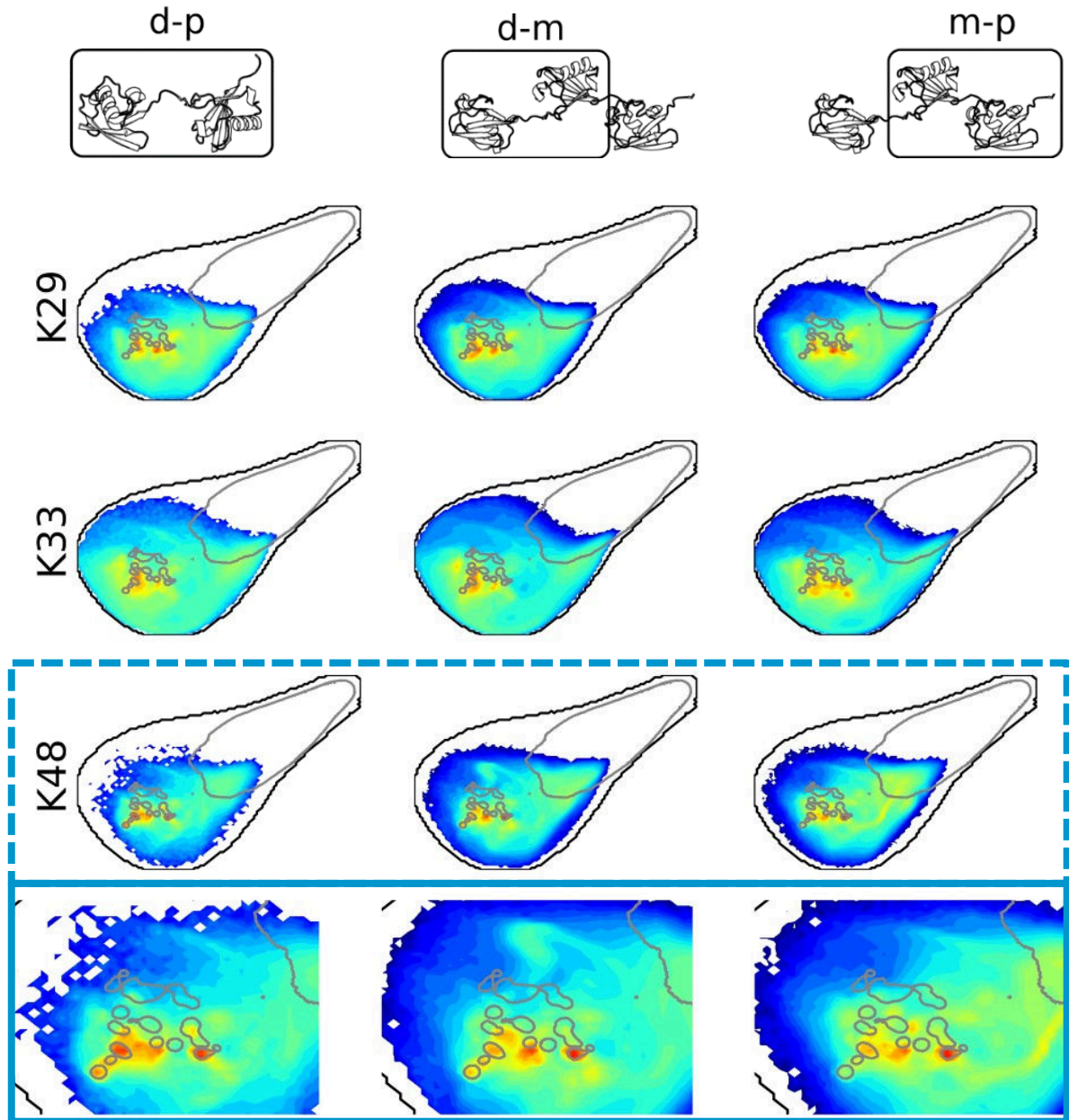
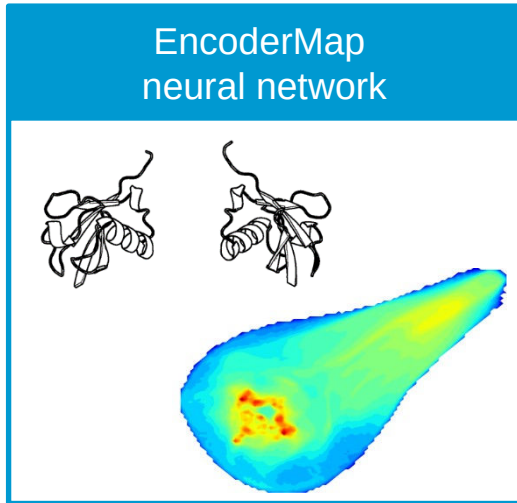


RMD 144D
5 M samples

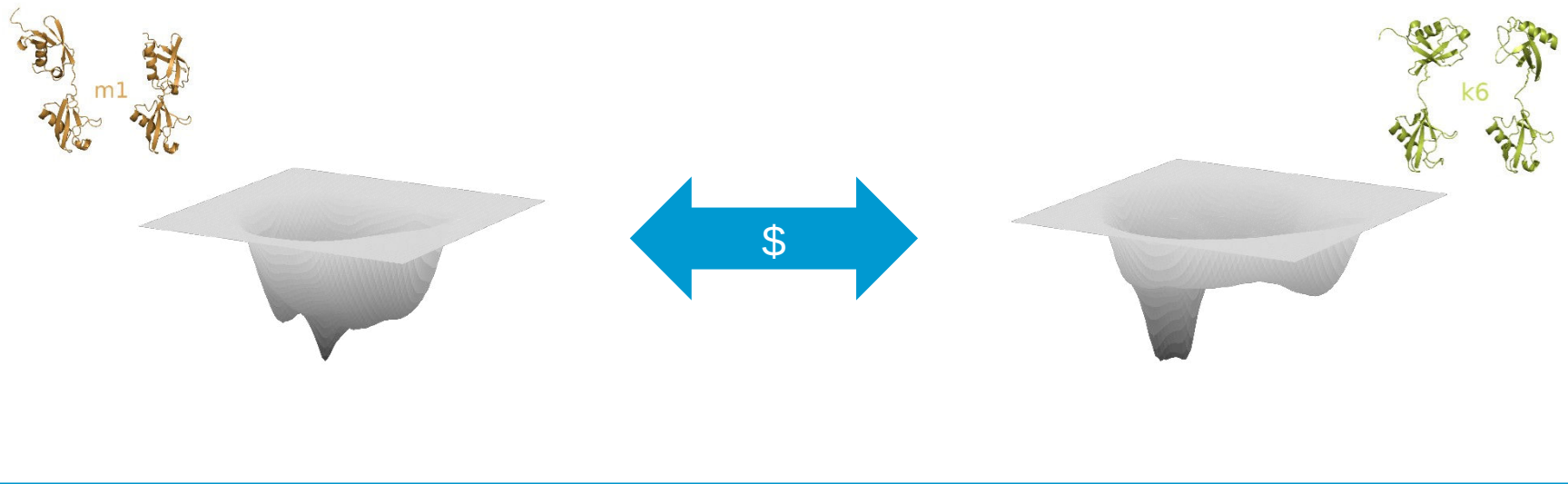


Berg et al., manuscript in revision

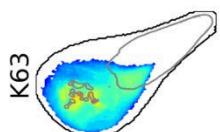
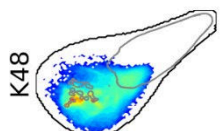
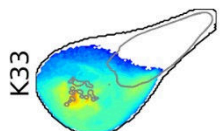
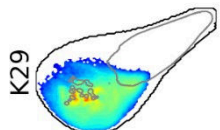
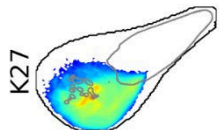
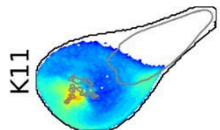
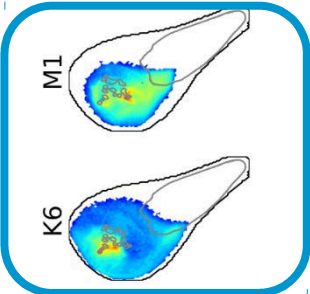




Earth-Mover-Distance (EMD)



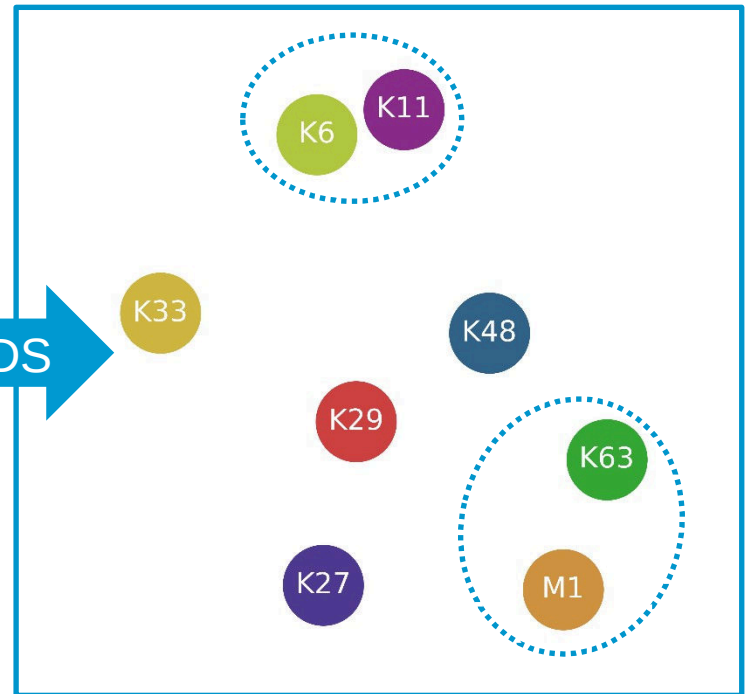
Rubner et al., IJCV (2000)
<https://github.com/garydoranjr/pyemd>

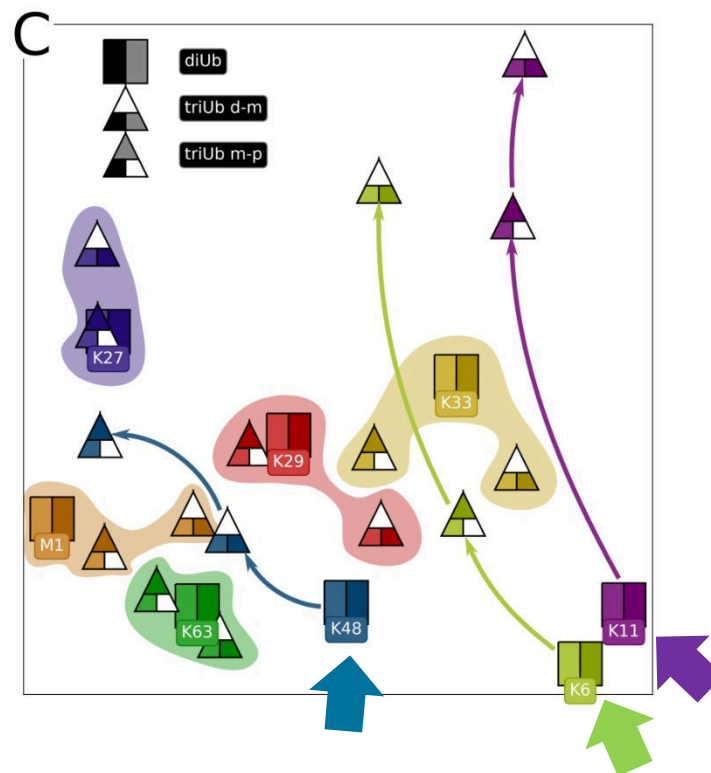
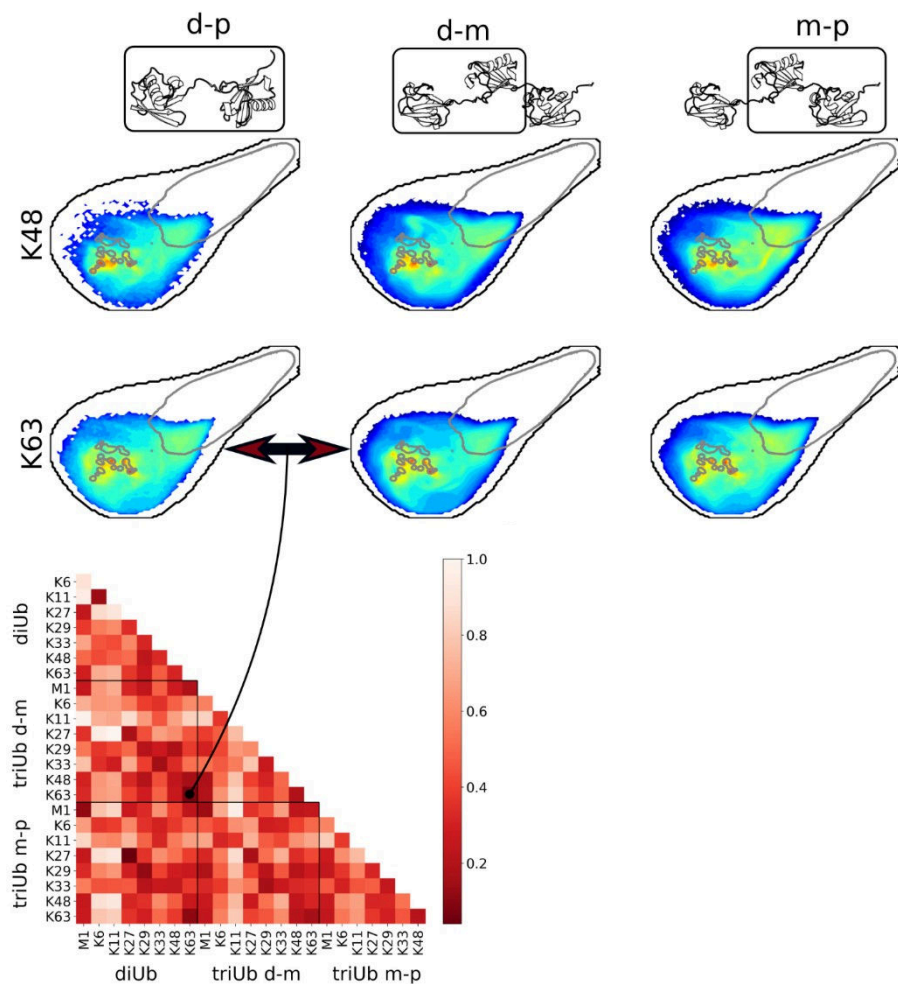


k6	0.9						
k11	1.0	0.1					
k27	0.3	0.9	1.0				
k29	0.4	0.6	0.6	0.4			
k33	0.7	0.5	0.5	0.6	0.3		
k48	0.5	0.4	0.5	0.5	0.2	0.3	
k63	0.3	0.7	0.8	0.4	0.3	0.5	0.3
	M1	K6	K11	K27	K29	K33	K48

relative pair-wise **EMD**

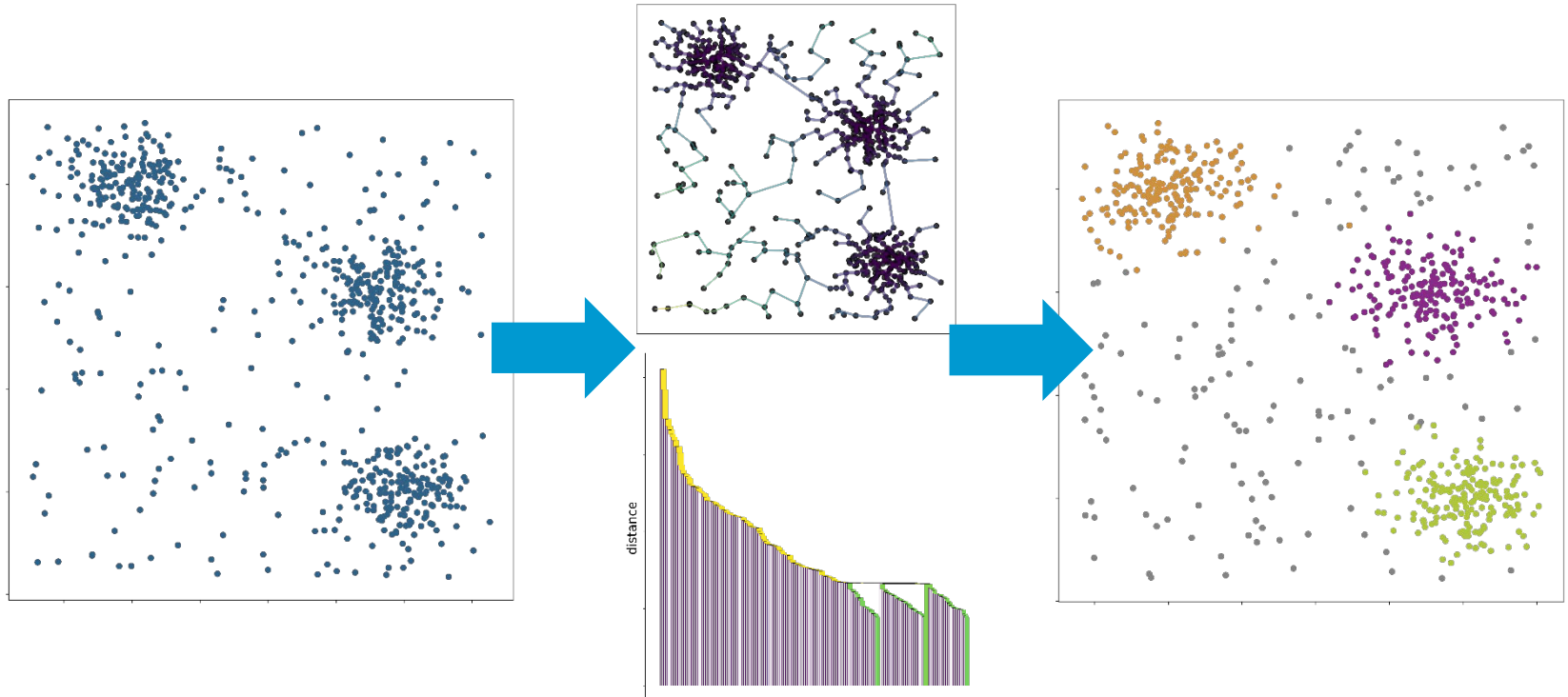
MDS

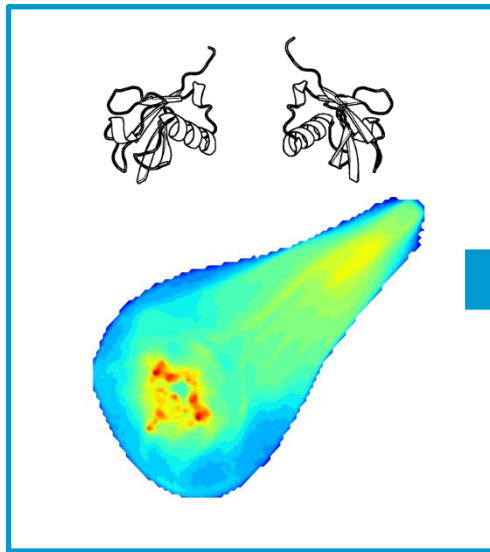




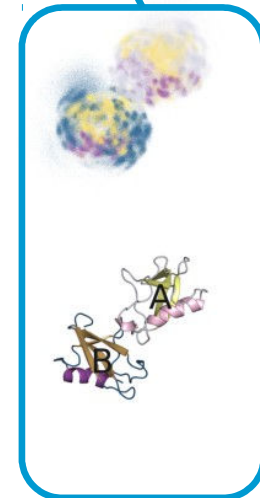
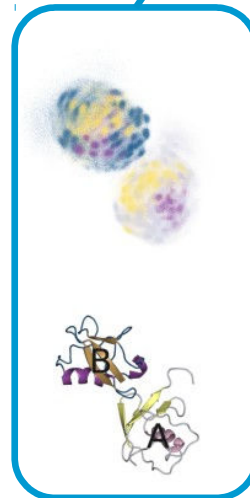
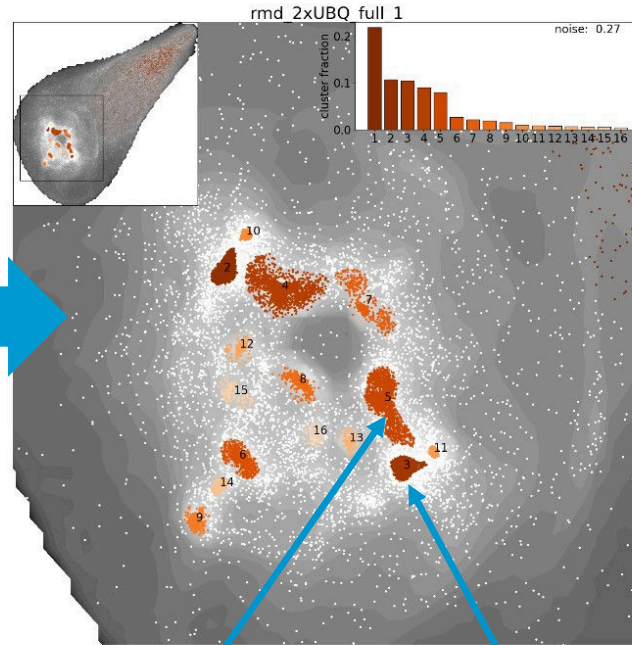
Clustering

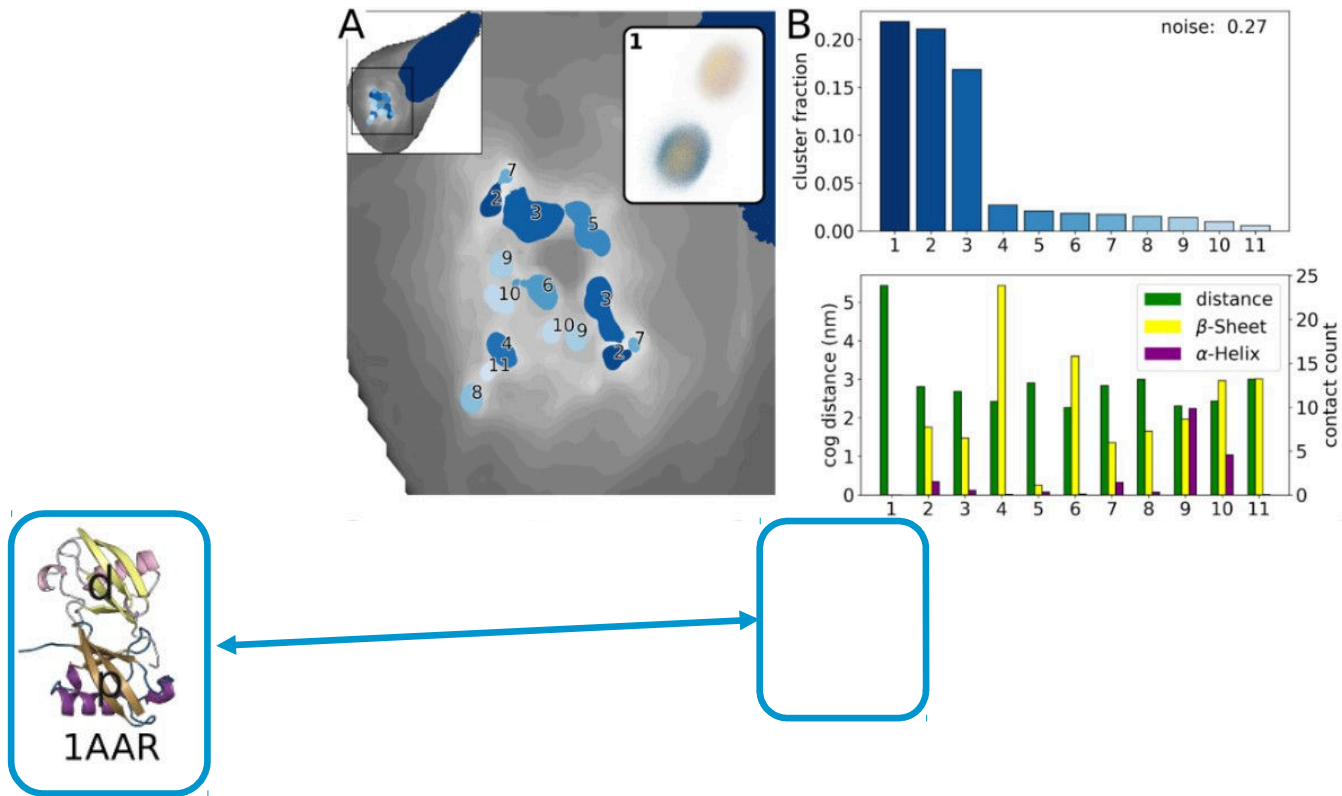
HDBSCAN

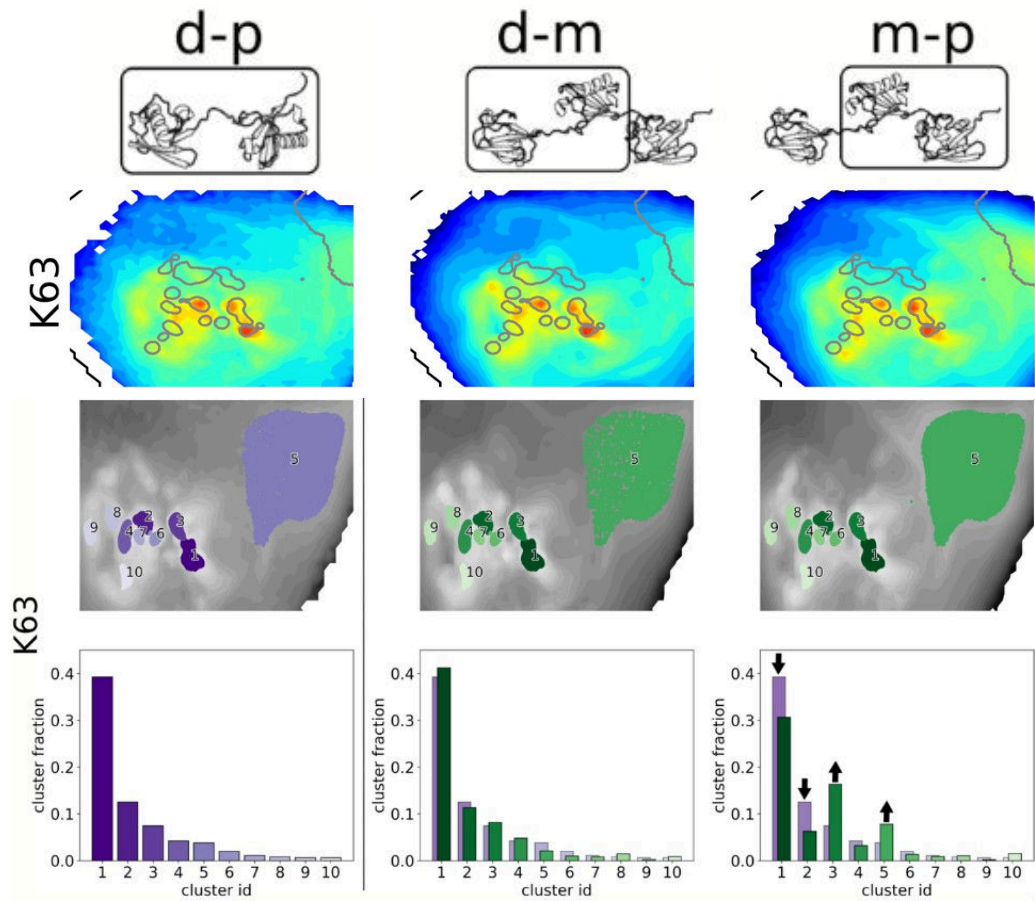
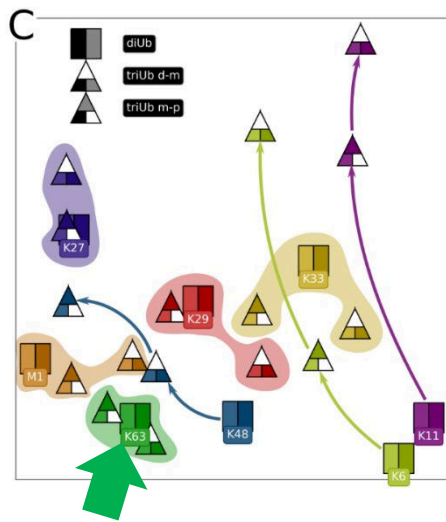


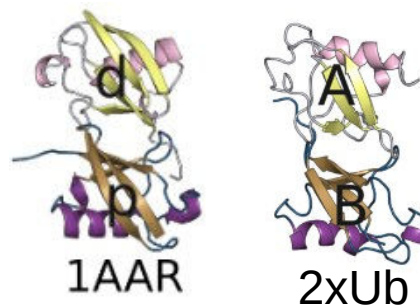
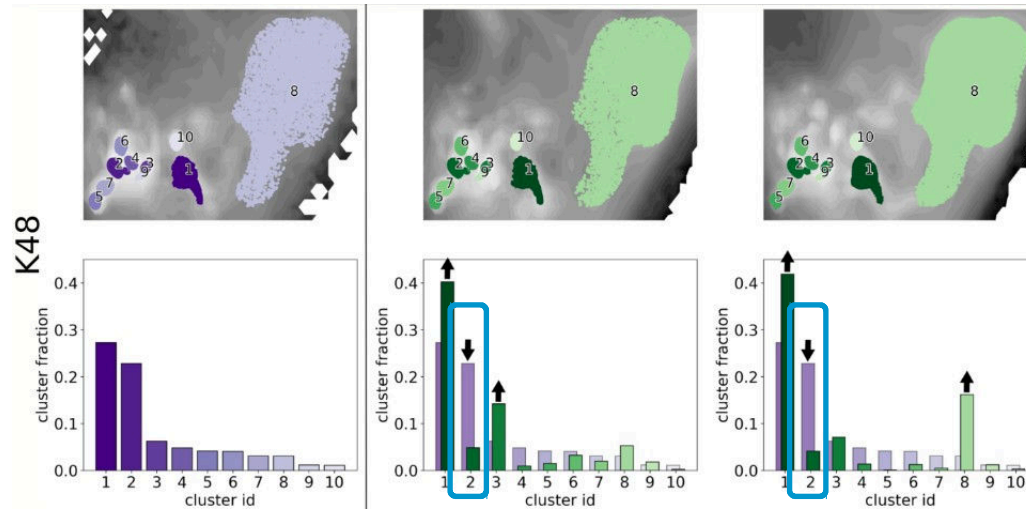
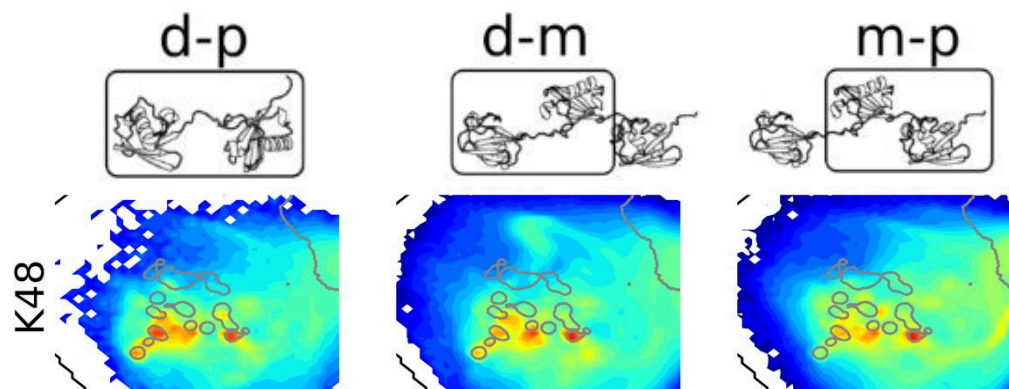
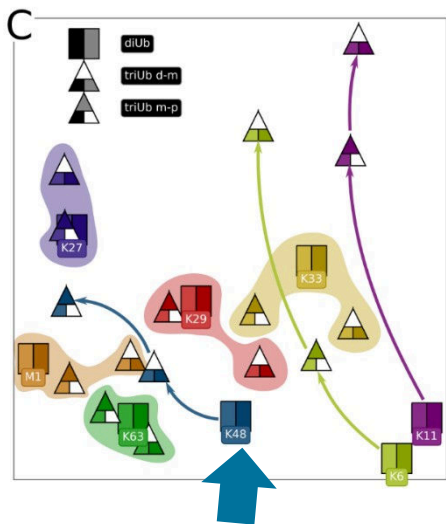


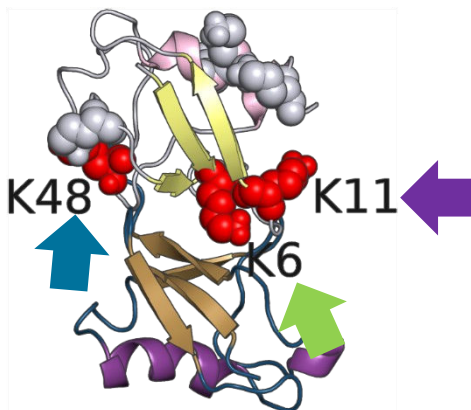
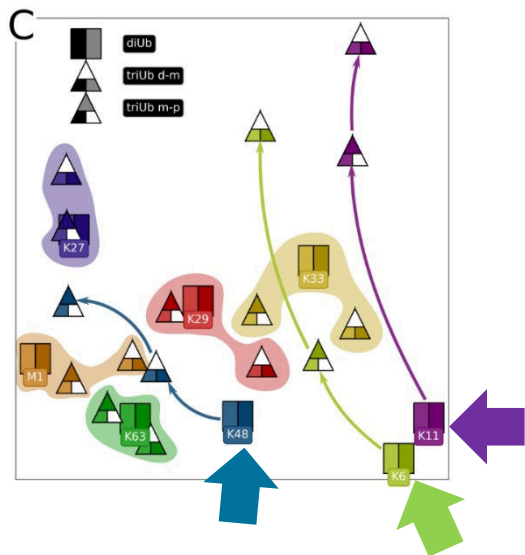
HDBSCAN



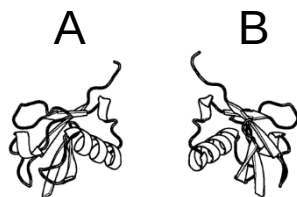








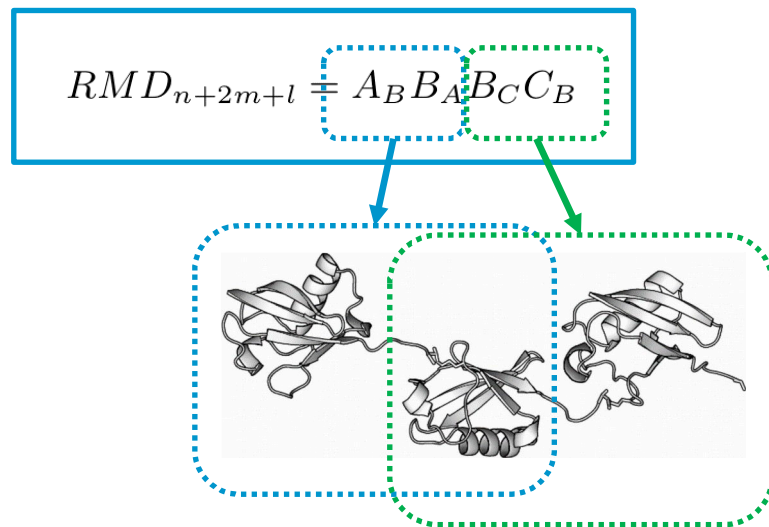
Extension to triUb



$$D_{A,B} = \begin{matrix} & x_1 & \dots & x_m \\ y_1 & \left(\begin{matrix} d_{a_1,b_1} & \dots & d_{a_1,b_m} \\ \vdots & & \vdots \\ d_{a_n,b_1} & \dots & d_{a_n,b_m} \end{matrix} \right) \\ \vdots & & & \\ y_n & & & \end{matrix}$$

$$A_B = [\min(y_1), \dots, \min(y_n)]$$

$$B_A = [\min(x_1), \dots, \min(x_m)]$$

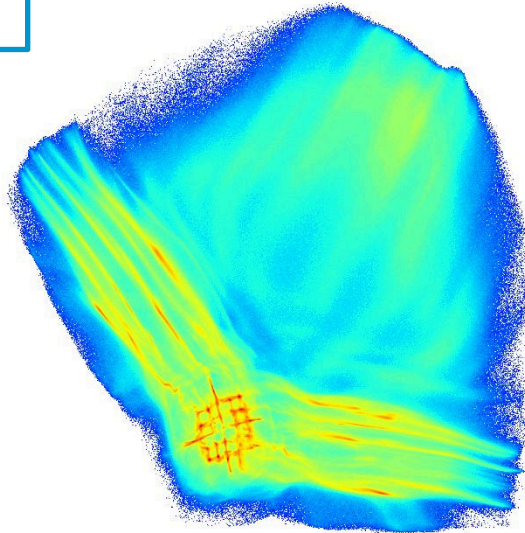
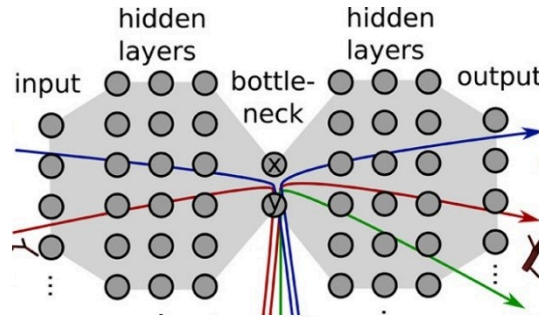


3xUb/triUb: **288** dimensions

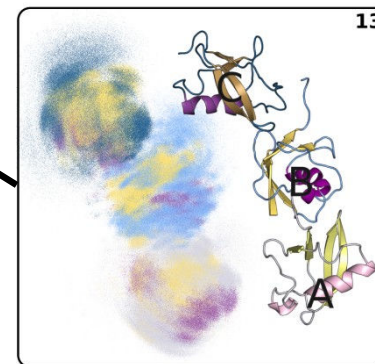
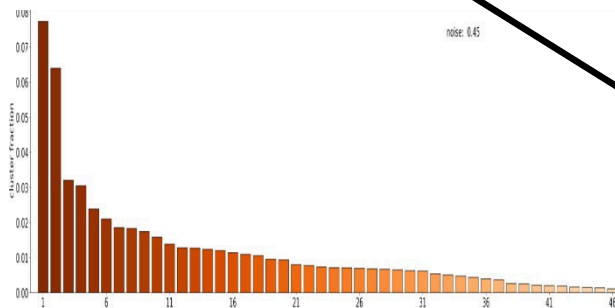
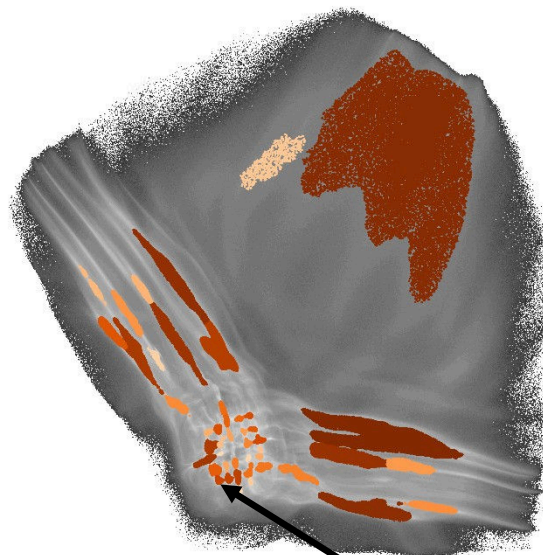
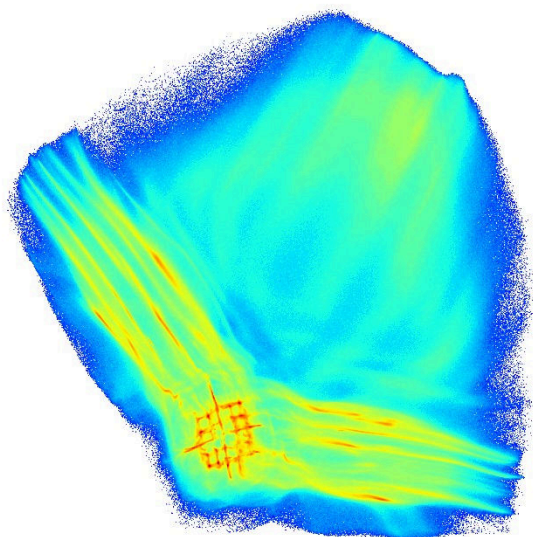
training set



RMD 288D
5 M samples

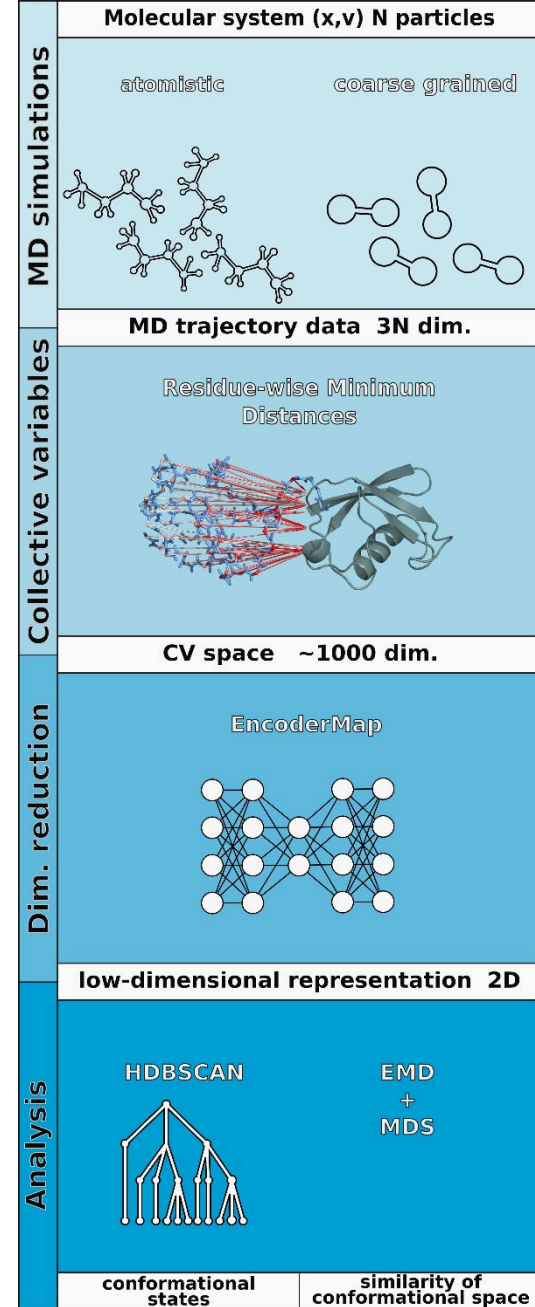


216 M samples
in total



Conclusions

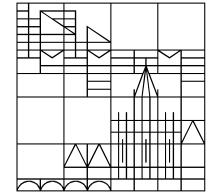
- conformational sampling on CG level
 - transferable and scalable CVs
 - effective dimensionality reduction
 - quantitative metric for conformational maps
 - robust clustering
-
- characterization of Ub-Ub interaction
 - conformation function correlation
 - linkage specific conformational characterization
 - longer chains behave different (K48, K6, K11)



Acknowledgements



Universität
Konstanz



AG Peter

Christine Peter
Christoph Globisch
Oleksandra Kukharenko
Tobias Lehmke
Michael King
Jakob Steuer
Kevin Savade
Alex Flachmüller
Simon Hunkler

Students

Ronja Schneck
Tatiana Meshcheryakova
Melissa Wörner
Julius Seumer
Leon Franke
Valeria Dilger

Collaborators

Tobias Schneider
Michael Kovermann
Martin Scheffner

DFG Deutsche
Forschungsgemeinschaft

