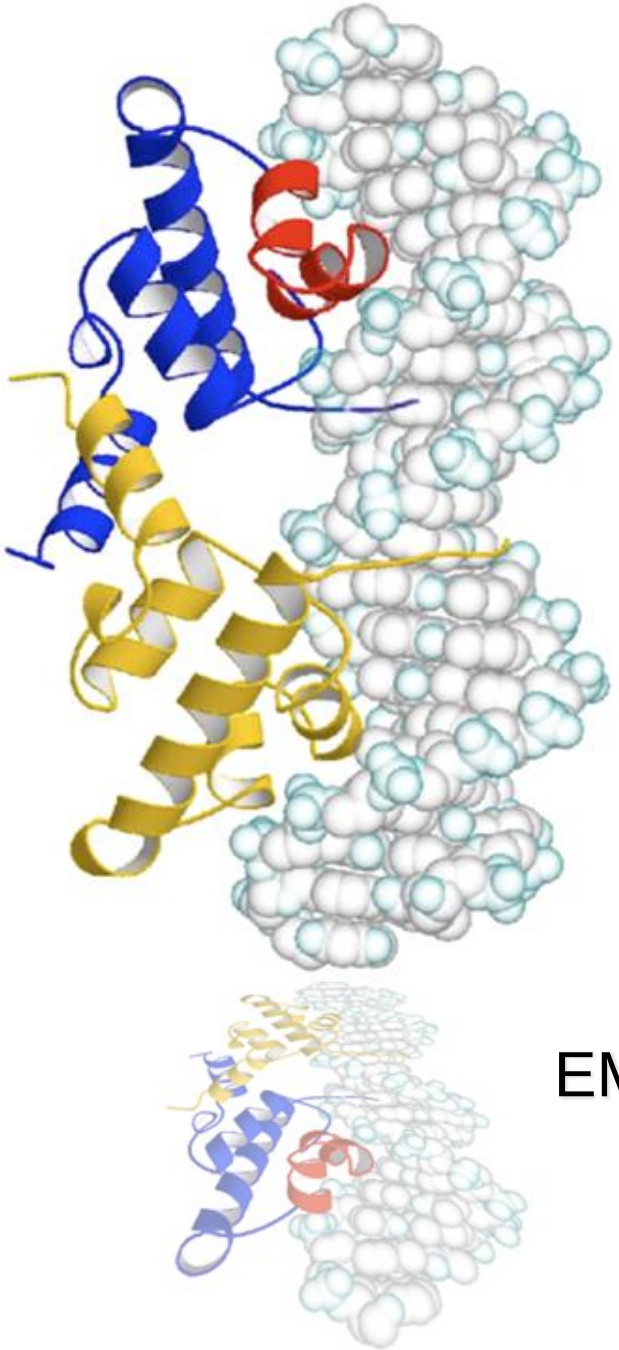


Evolution and assembly of protein complexes

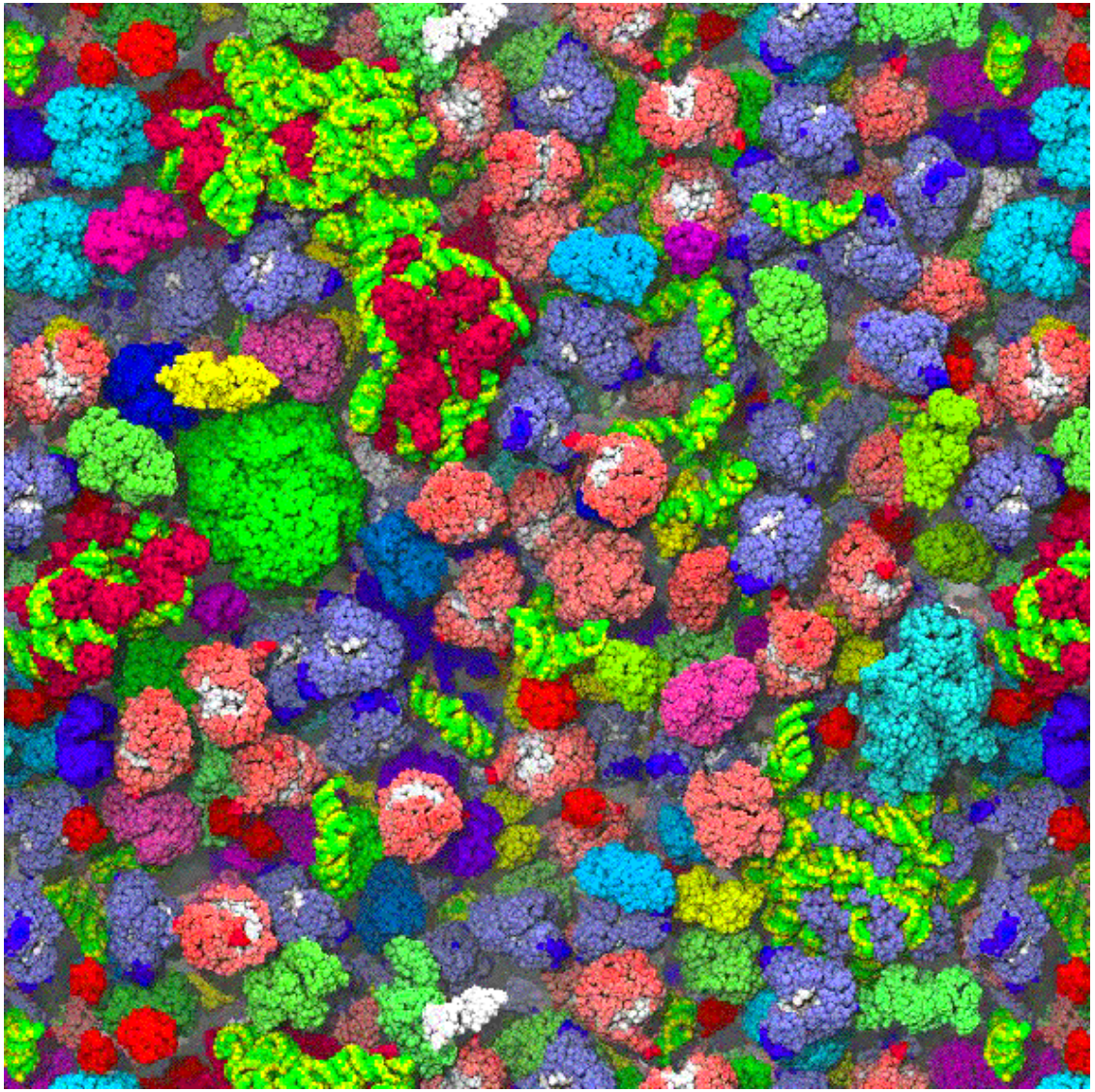
Sarah Teichmann

EMBL-European Bioinformatics Institute &

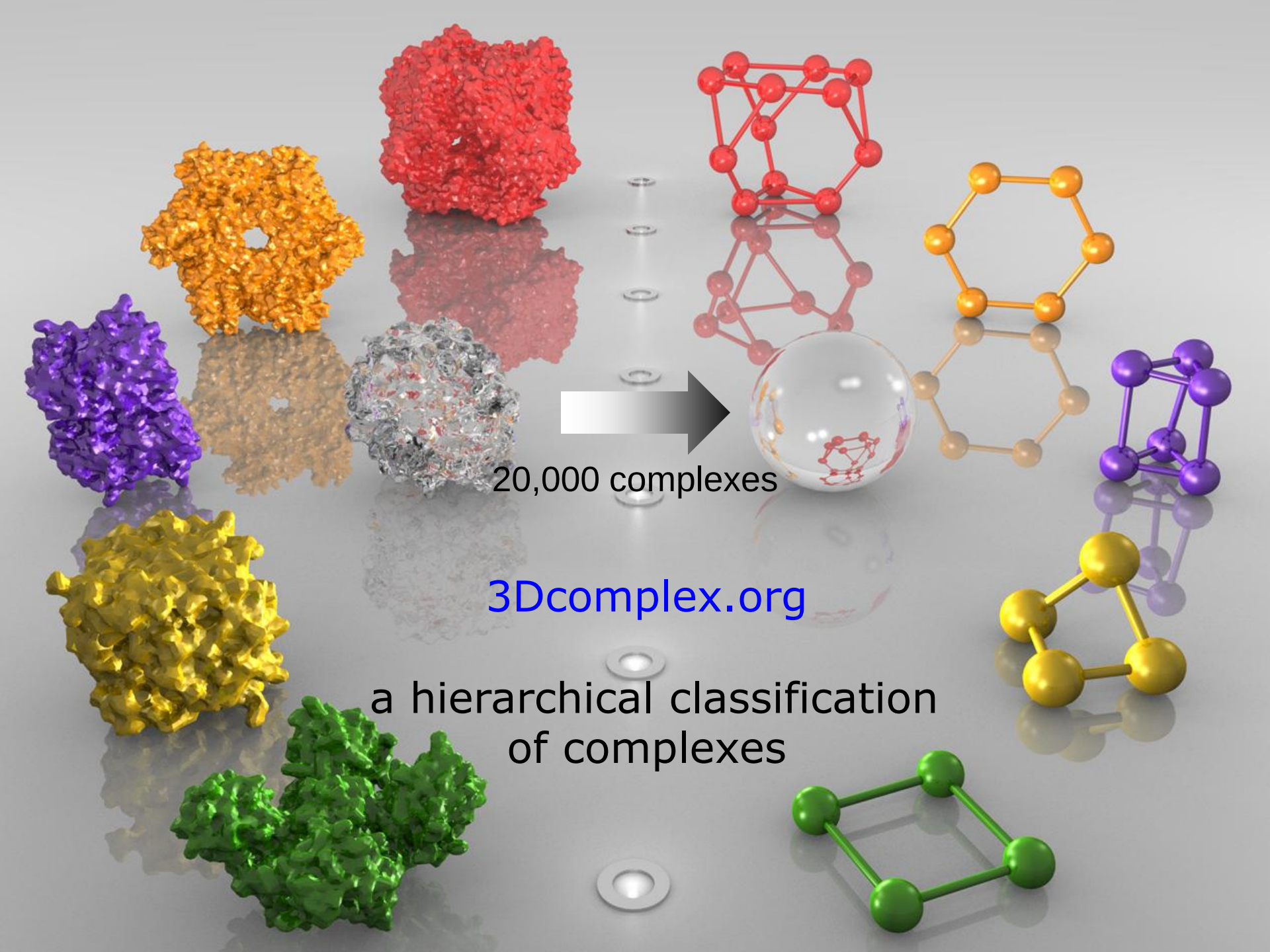
Wellcome Trust Sanger Institute

Cambridge, UK

Protein complexes assemble specifically



Adrian Elcock, U. Iowa


20,000 complexes
3Dcomplex.org
a hierarchical classification
of complexes

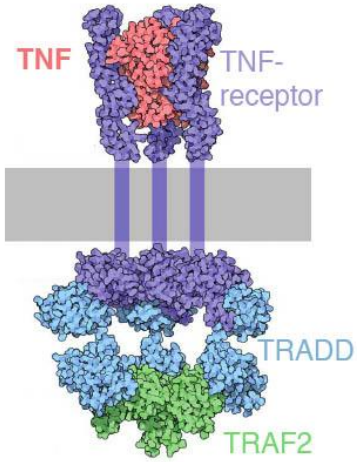
Two types of protein complexes

- Homomers

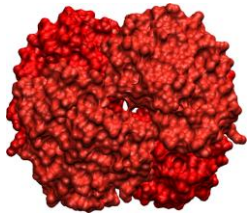
Caspase-9

Apoptosis

- Heteromers



Hemoglobin



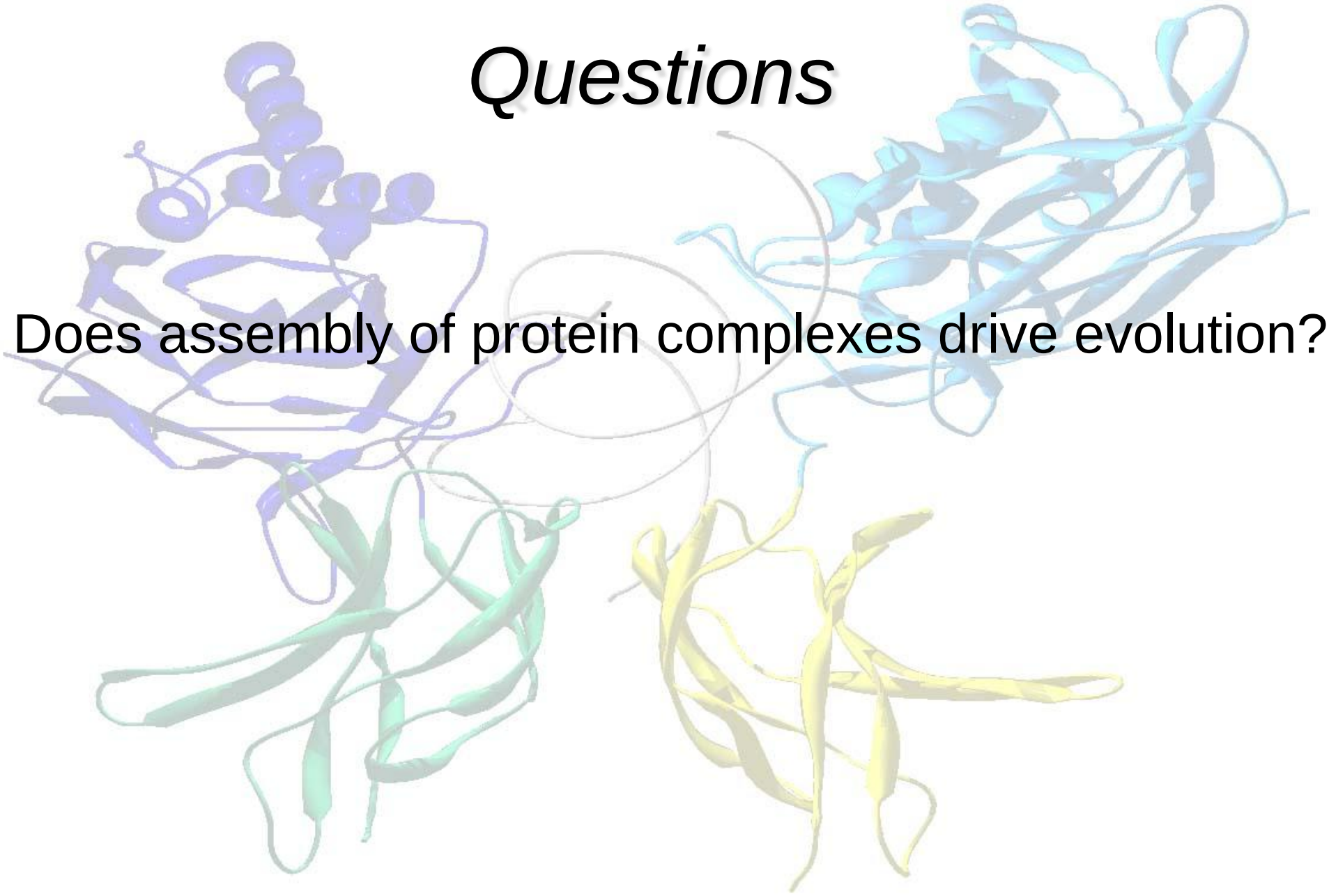
Questions

Does assembly of protein complexes drive evolution?

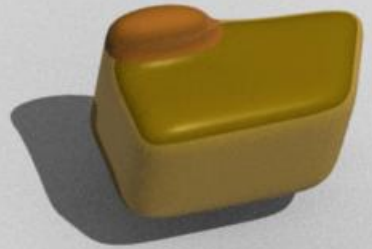
What are mutational mechanisms?

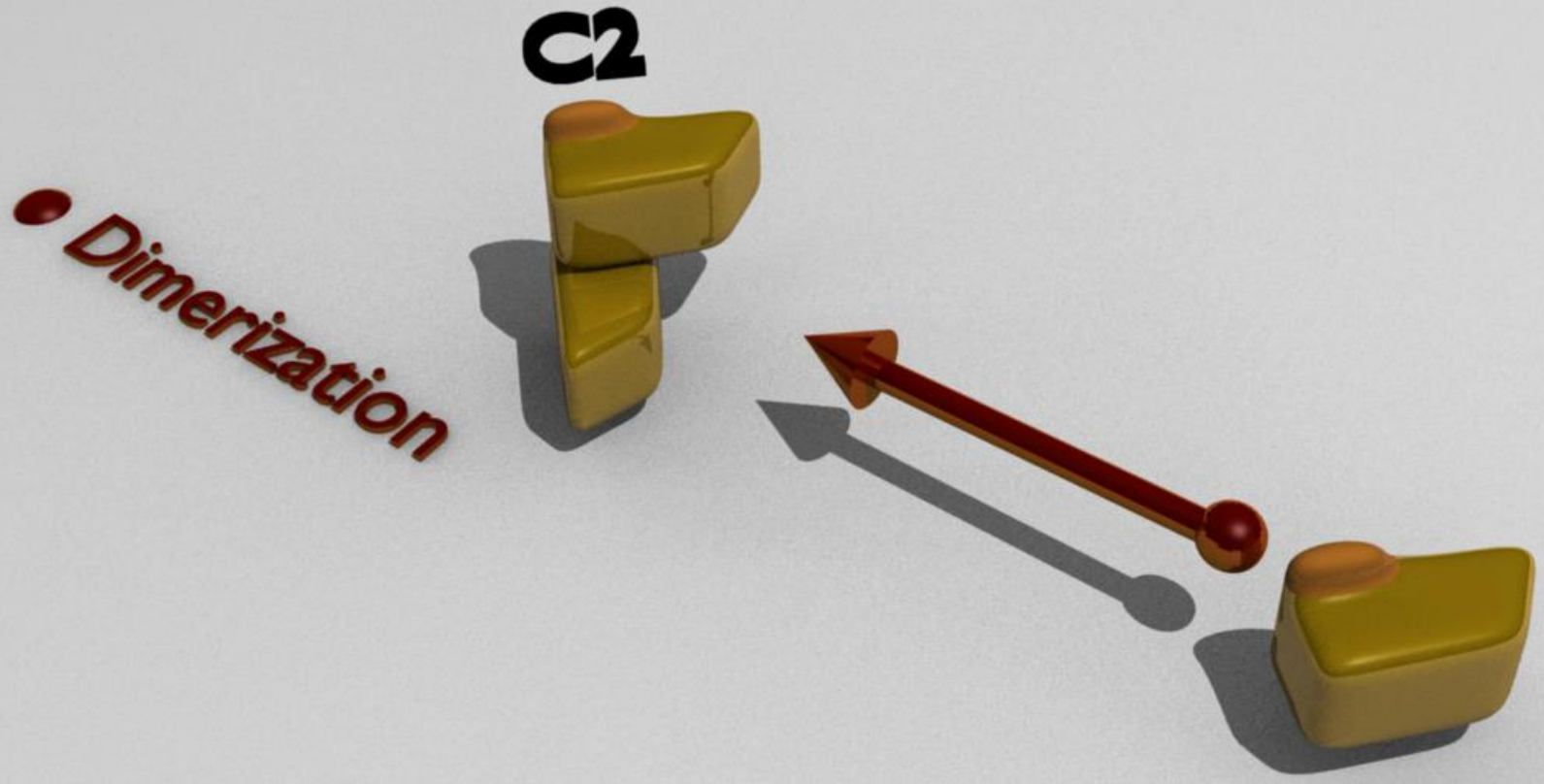
Can principles of assembly to predict topologies?

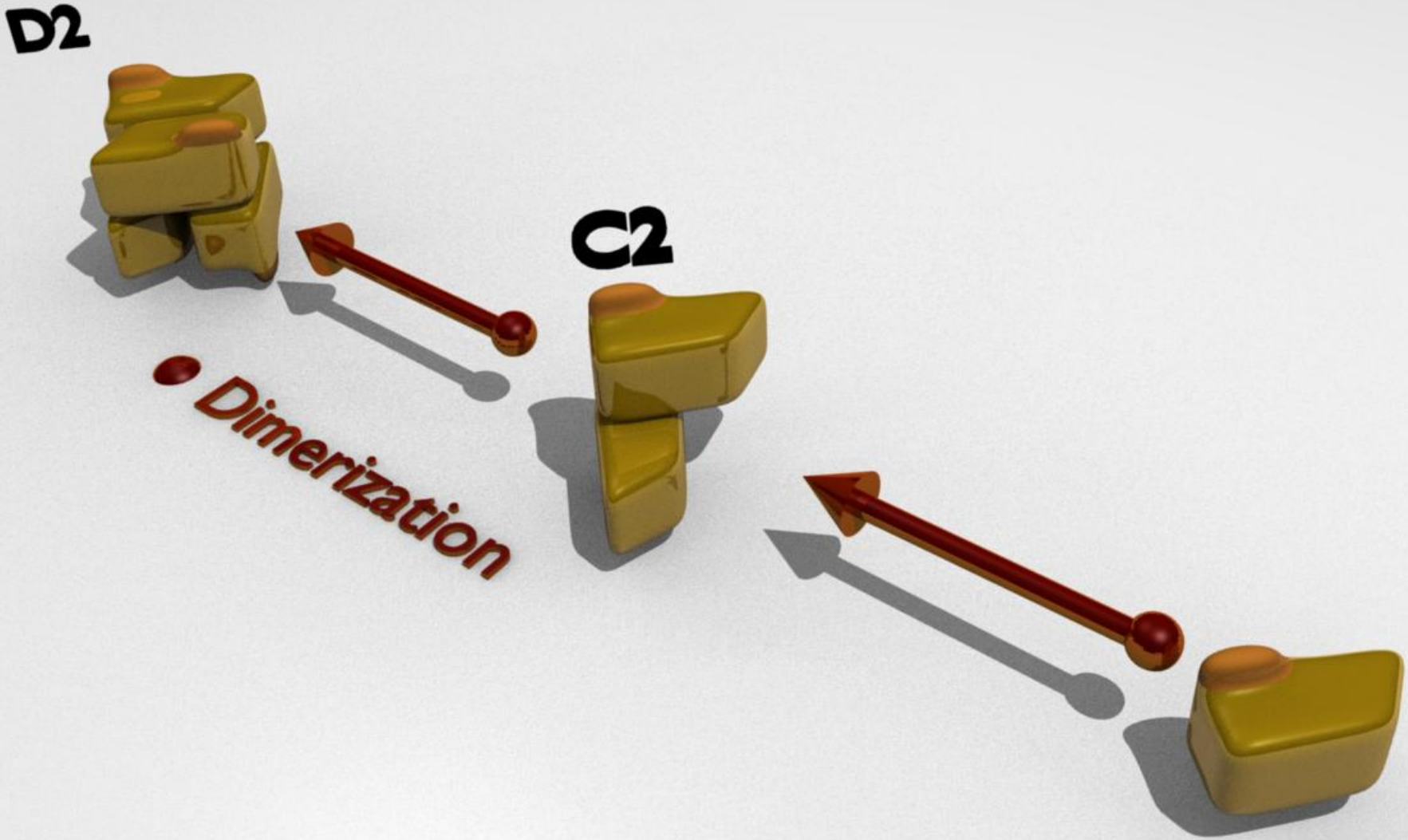
Questions

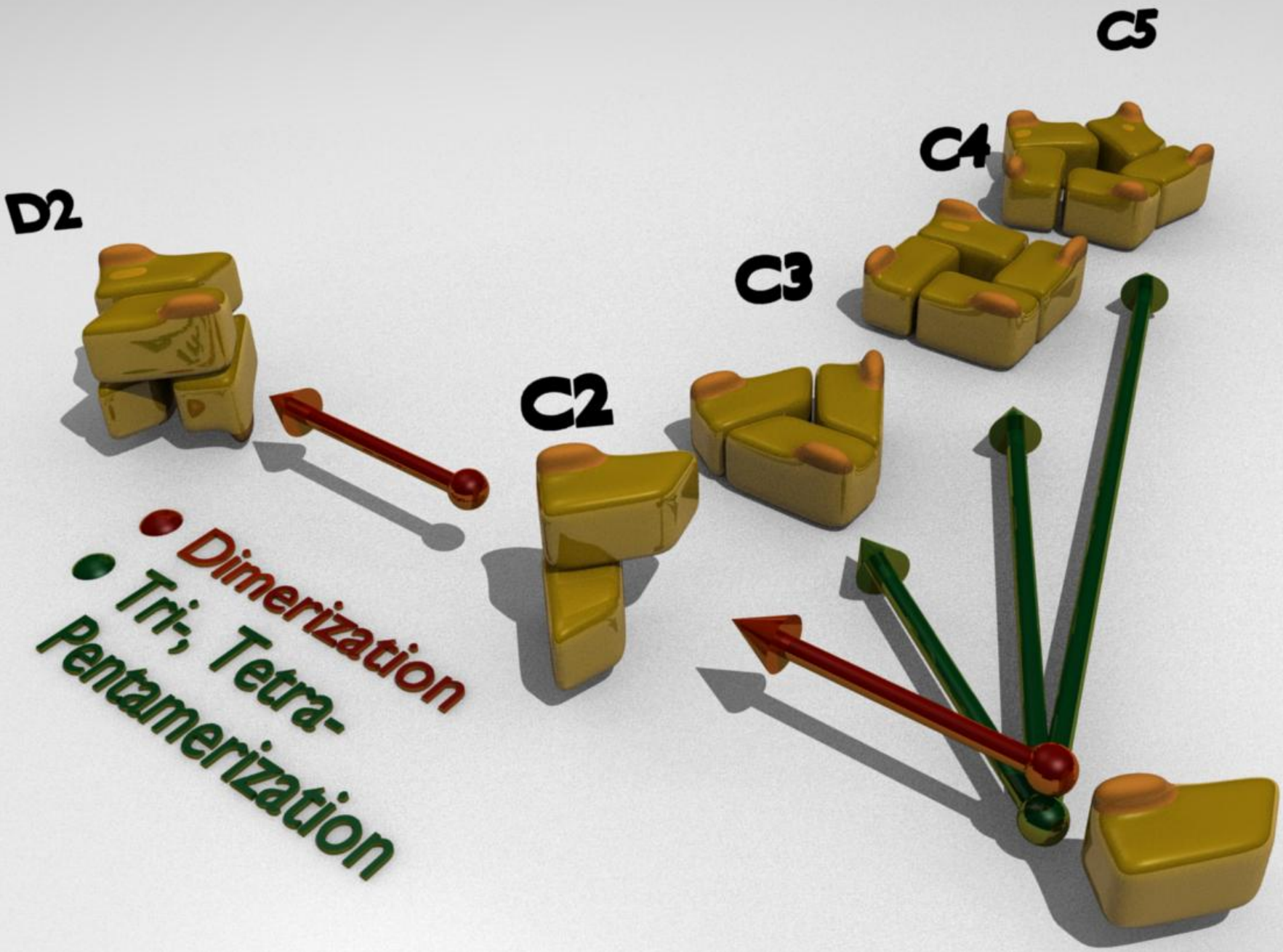


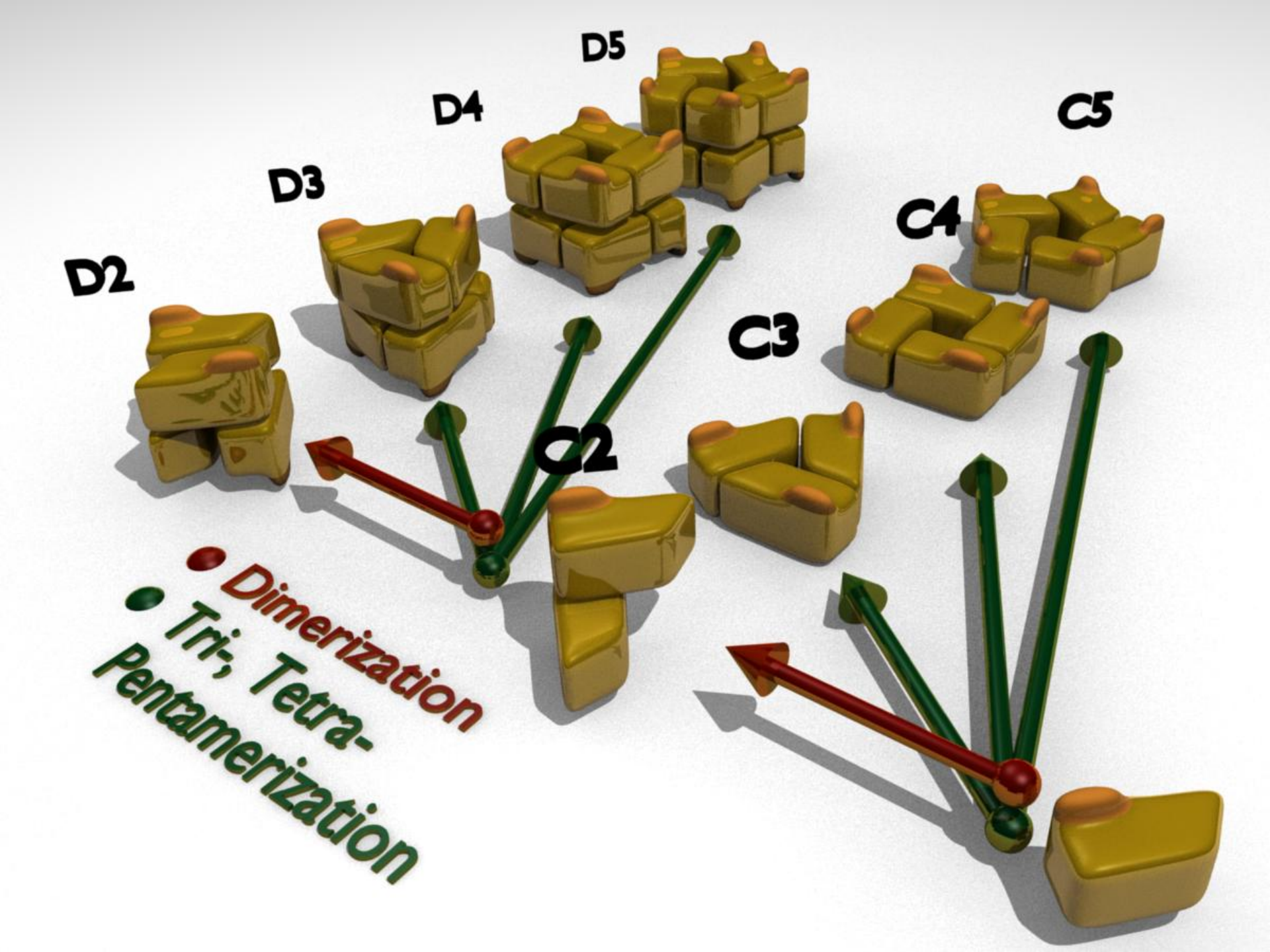
Does assembly of protein complexes drive evolution?

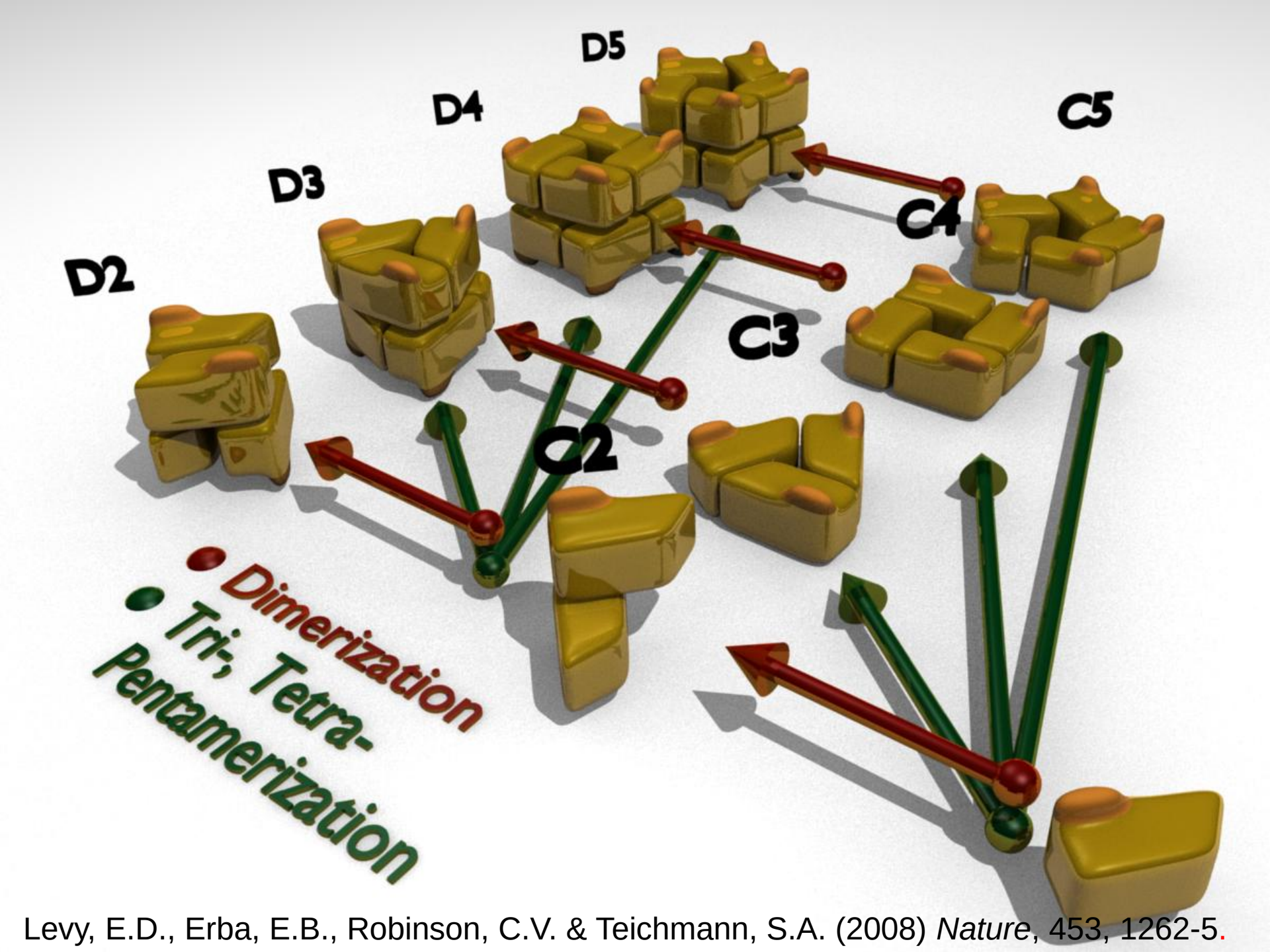



C2
Dimerization


D2
C2
Dimerization


D2
C2
C3
C4
C5
Dimerization
Tri-, Tetra-
Pentamerization


D5
D4
D3
D2
C5
C4
C3
C2
Dimerization
Tri-, Tetra-
Pentamerization



Levy, E.D., Erba, E.B., Robinson, C.V. & Teichmann, S.A. (2008) *Nature*, 453, 1262-5.

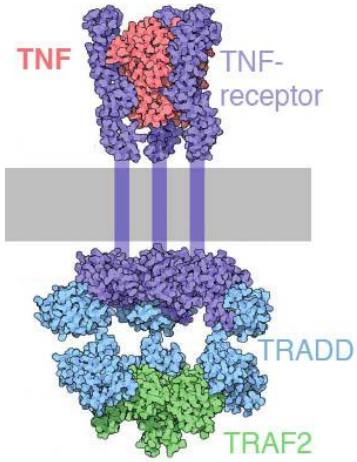
Two types of protein complexes

- Homomers

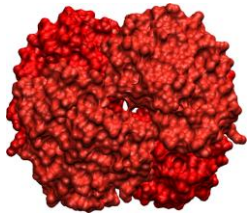
Caspase-9

Apoptosis

- Heteromers

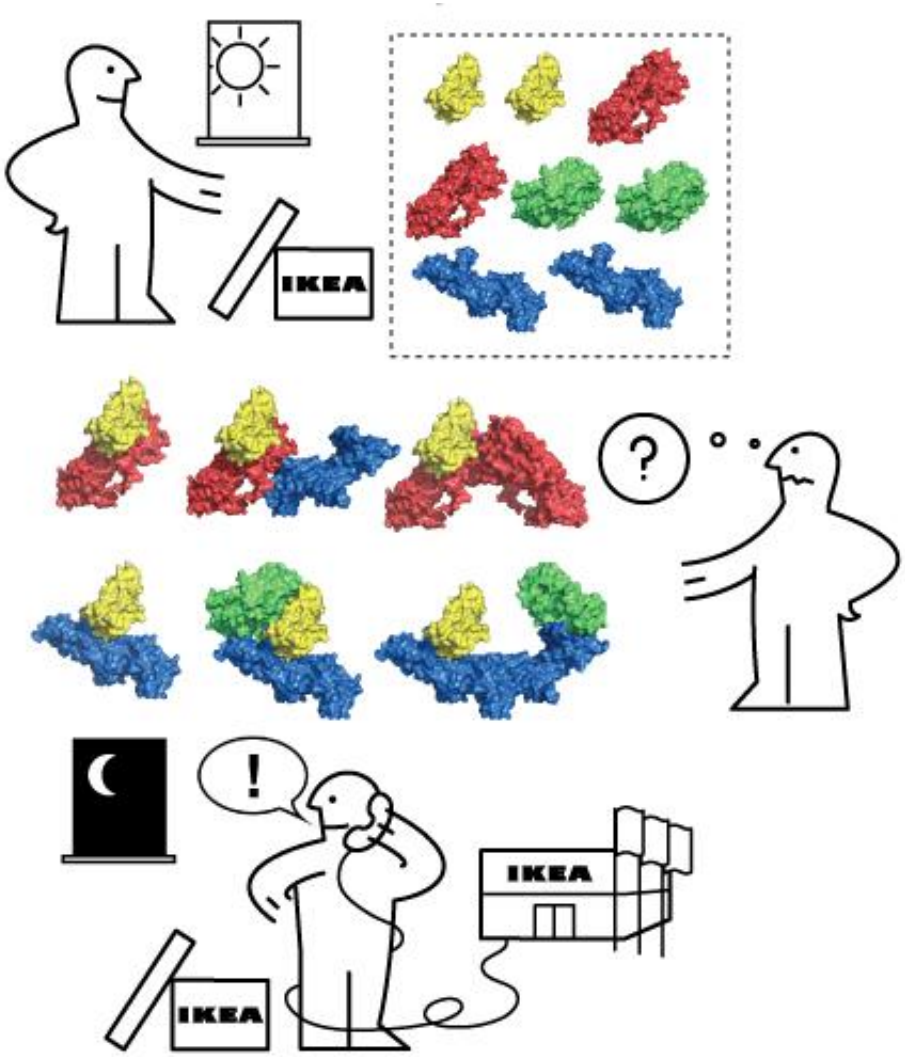


Hemoglobin



The order of assembly is important

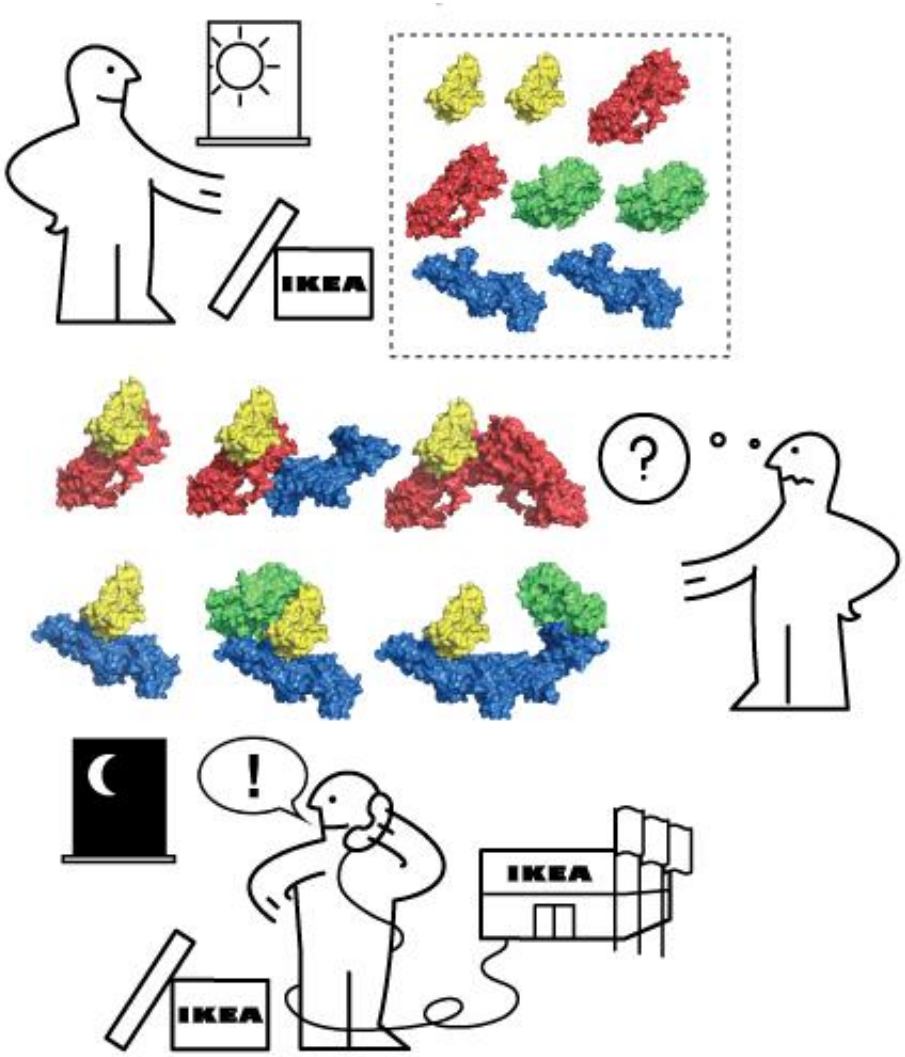
Random assembly



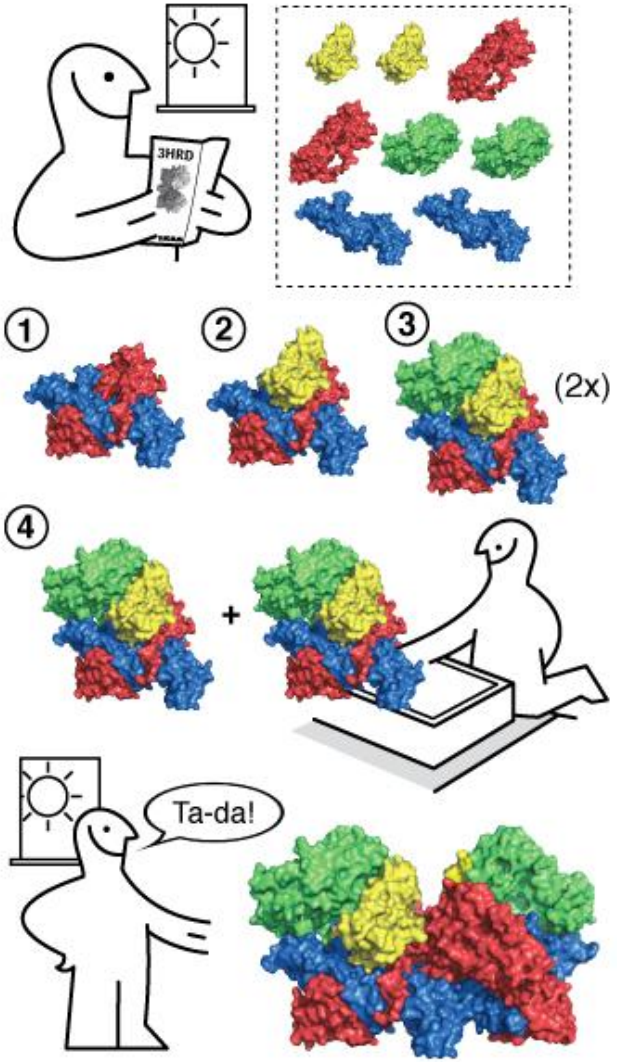
Joe Marsh

The order of assembly is important

Random assembly



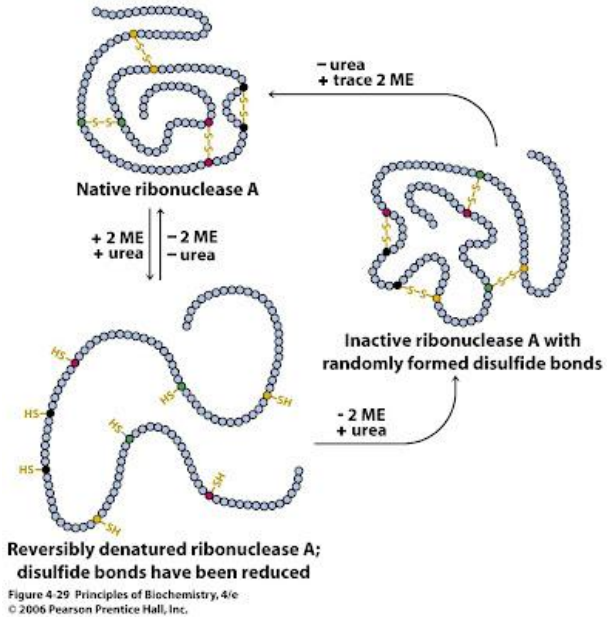
Ordered assembly



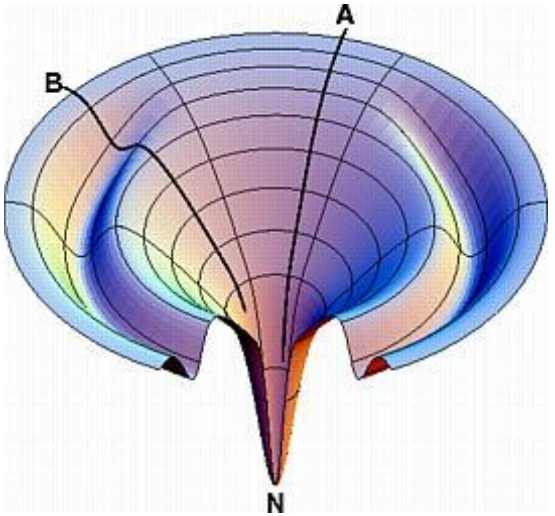
Joe Marsh

Protein Folding: Ordered, Fast, Spontaneous

Anfinsen: bovine pancreatic ribonuclease

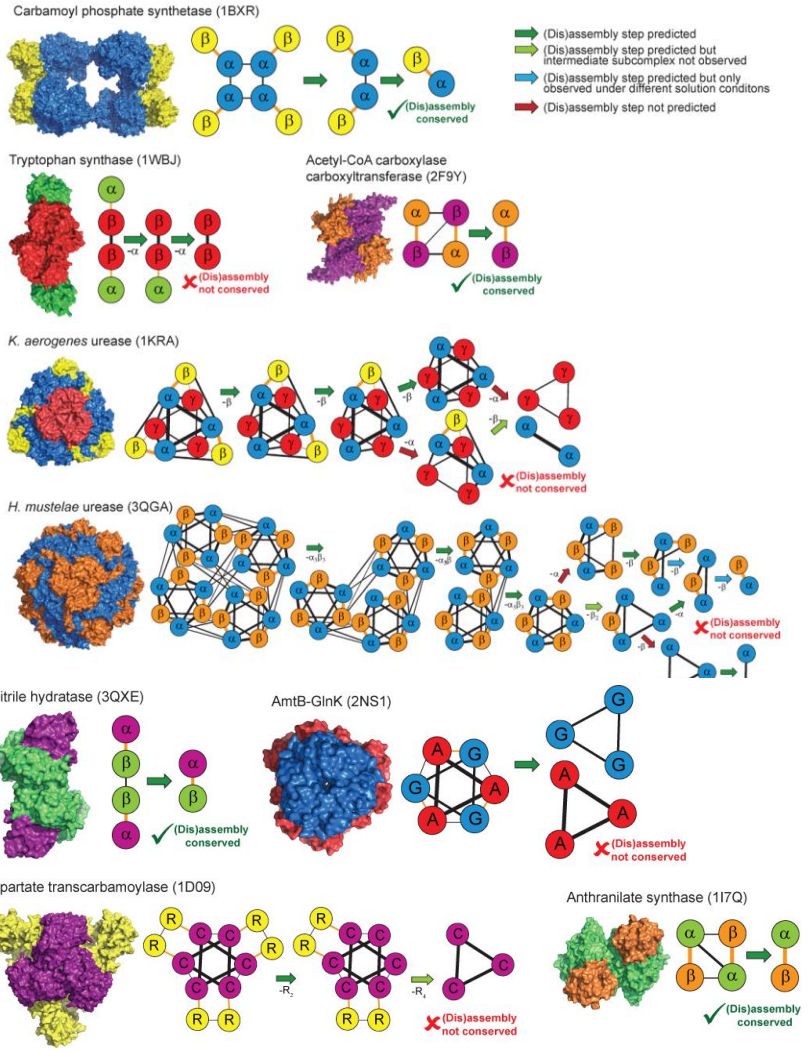


Protein folding landscape vs Levinthal's paradox



Predictable? Conserved?

Ordered (dis)assembly pathways can be predicted from crystal structures



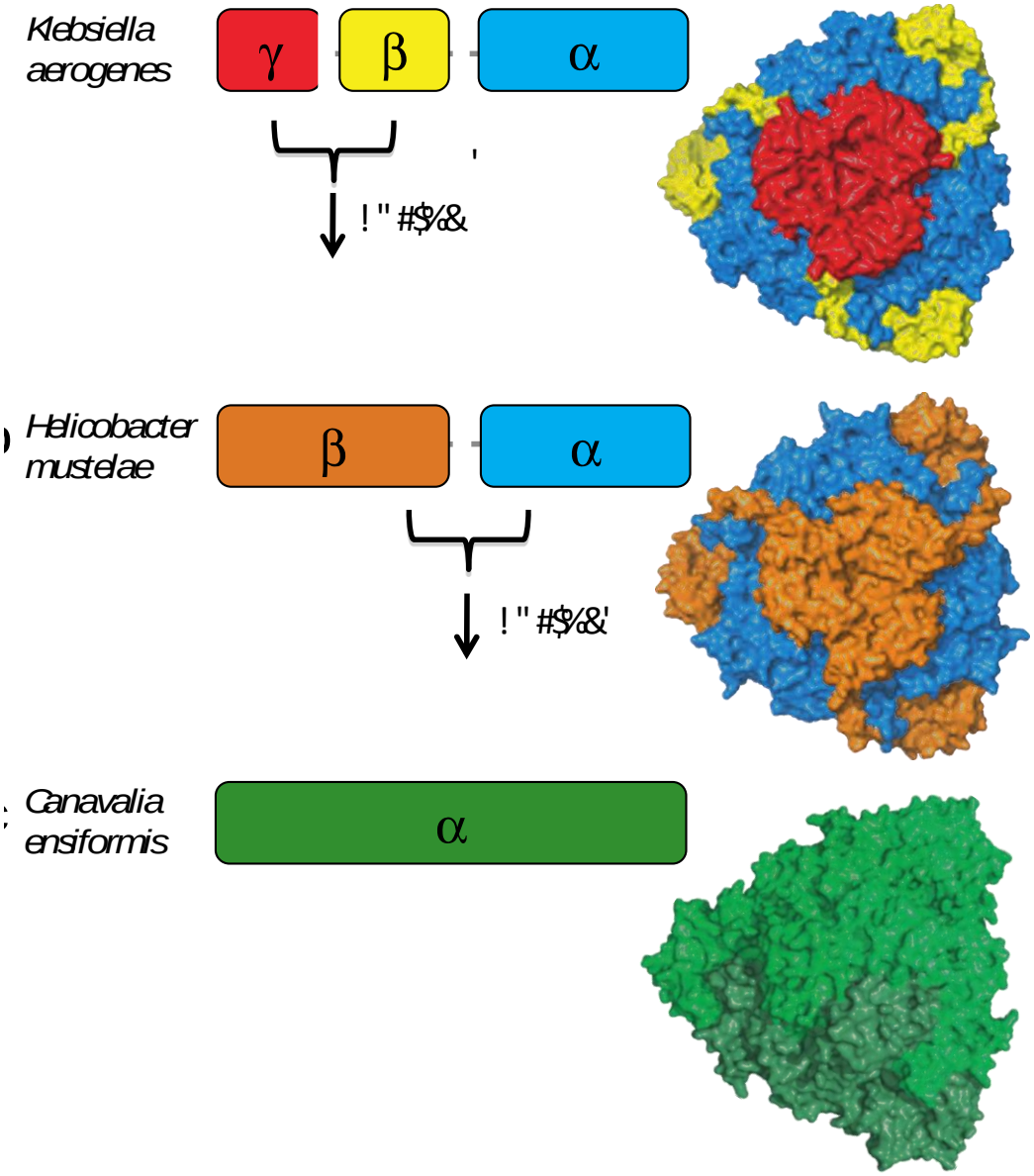
Complex name	PDB ID	Correctly predicted steps
nESI-MS		
Carbamoyl phosphate synthetase	1BXR	2/2
Tryptophan synthase	1WBJ	2/2
Acetyl-CoA carboxylase carboxytransferase	2F9Y	1/1
Klebsiella aerogenes urease	1KRA	4/6
Helicobacter mustelae urease	3QGA	9/11
Literature		
Nitrile hydratase	3QXE	1/1
AmtB-GlnK	2NS1	1/1
Aspartate transcarbamoylase	1D09	2/2
Anthranilate synthase	1I7Q	1/1
	Total	**23/27**

7/9 complexes predicted perfectly

Gene fusion

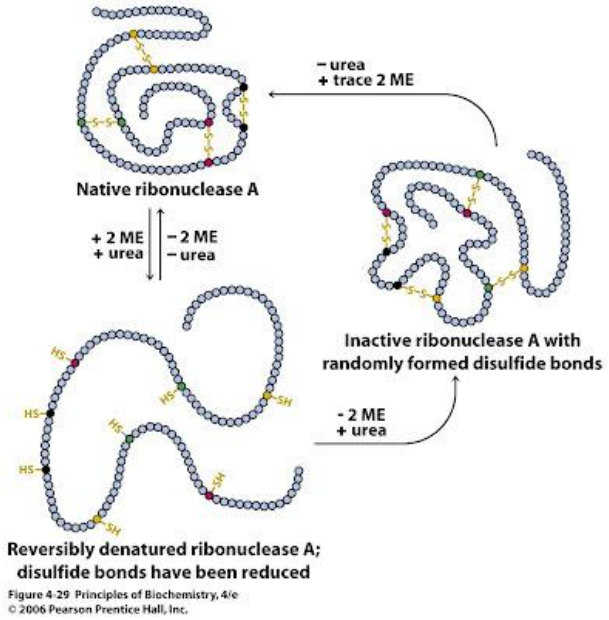
and

fission

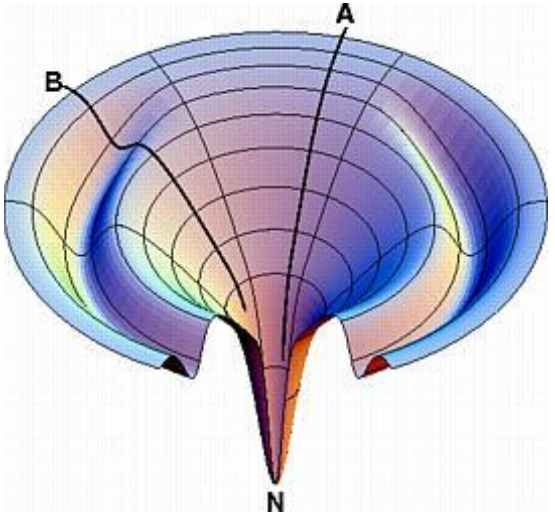


Ordered, Fast, Spontaneous (c.f. Protein Folding)

Anfinsen: bovine pancreatic ribonuclease



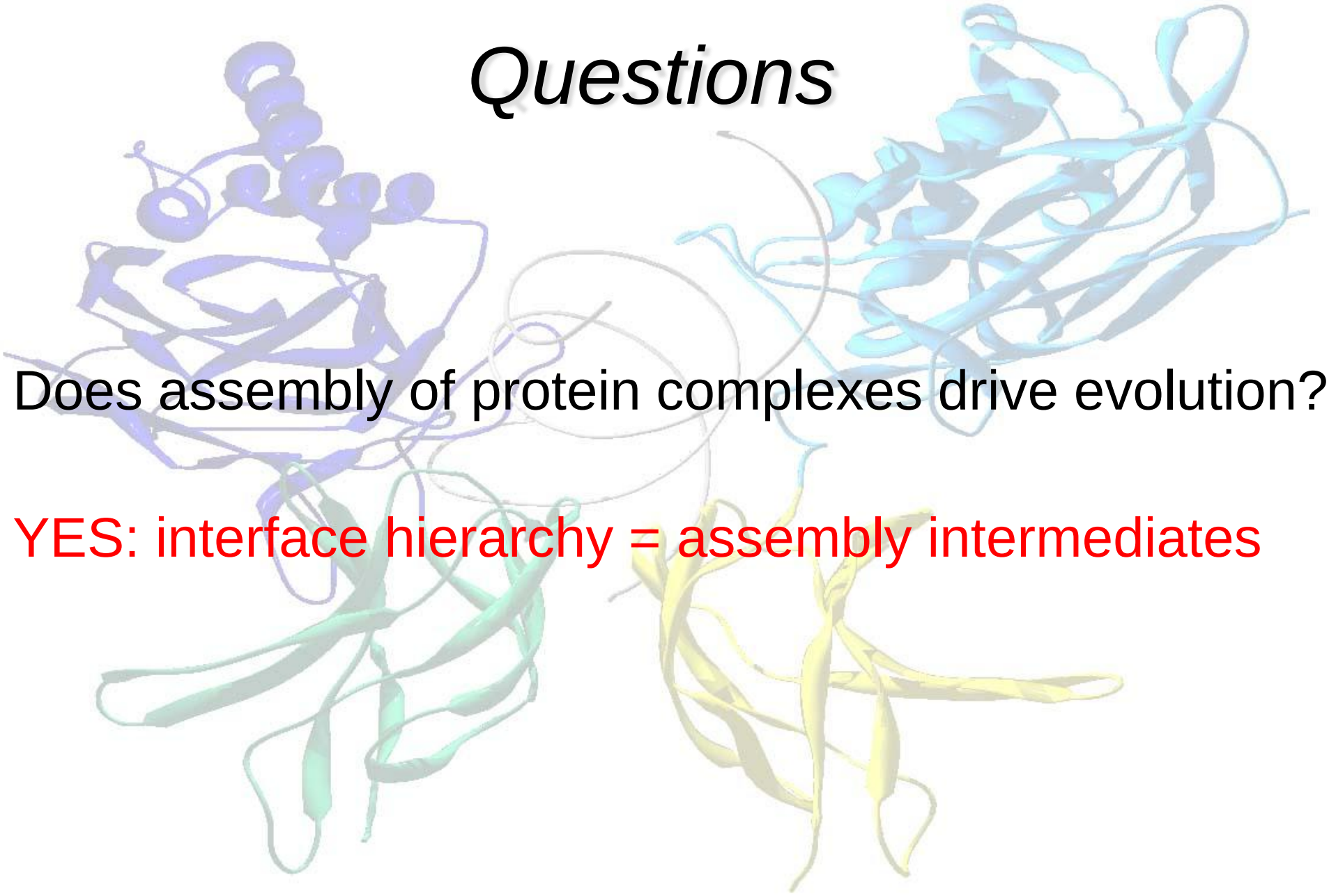
Protein folding landscape vs Levinthal's paradox



Predictable? Conserved? YES!

Marsh, J., Hernandez, H., Hall, Z., Ahnert, S., Perica, T., Robinson, C.V., Teichmann, S.A., *Cell*, 2013.

Questions



Does assembly of protein complexes drive evolution?

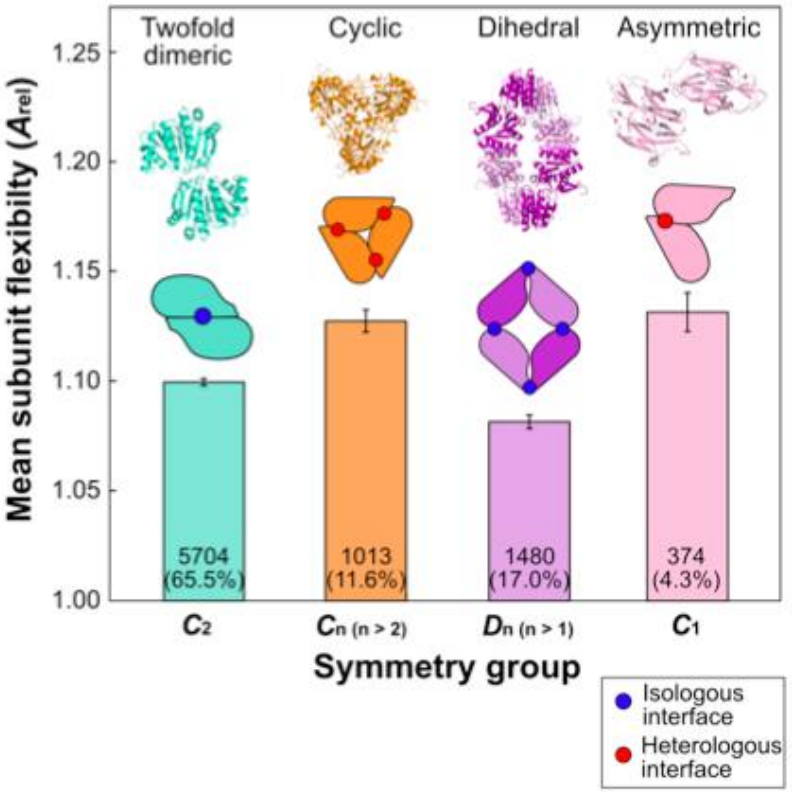
YES: interface hierarchy = assembly intermediates

Questions

Does assembly of protein complexes drive evolution?

YES

And flexibility



Marsh & Teichmann, PLoS Biol, 2014

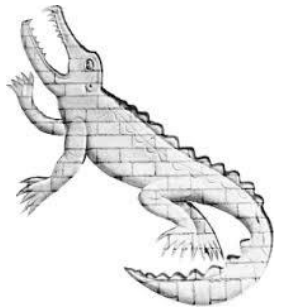
Acknowledgements

Joseph Marsh (now MRC-HGU, Edinburgh)

Emmanuel Levy (now Weizmann Institute)

Carol Robinson & Zoe Hall , Dept Chemistry, Oxford

Sebastian Ahnert, Dept Physics, Cambridge

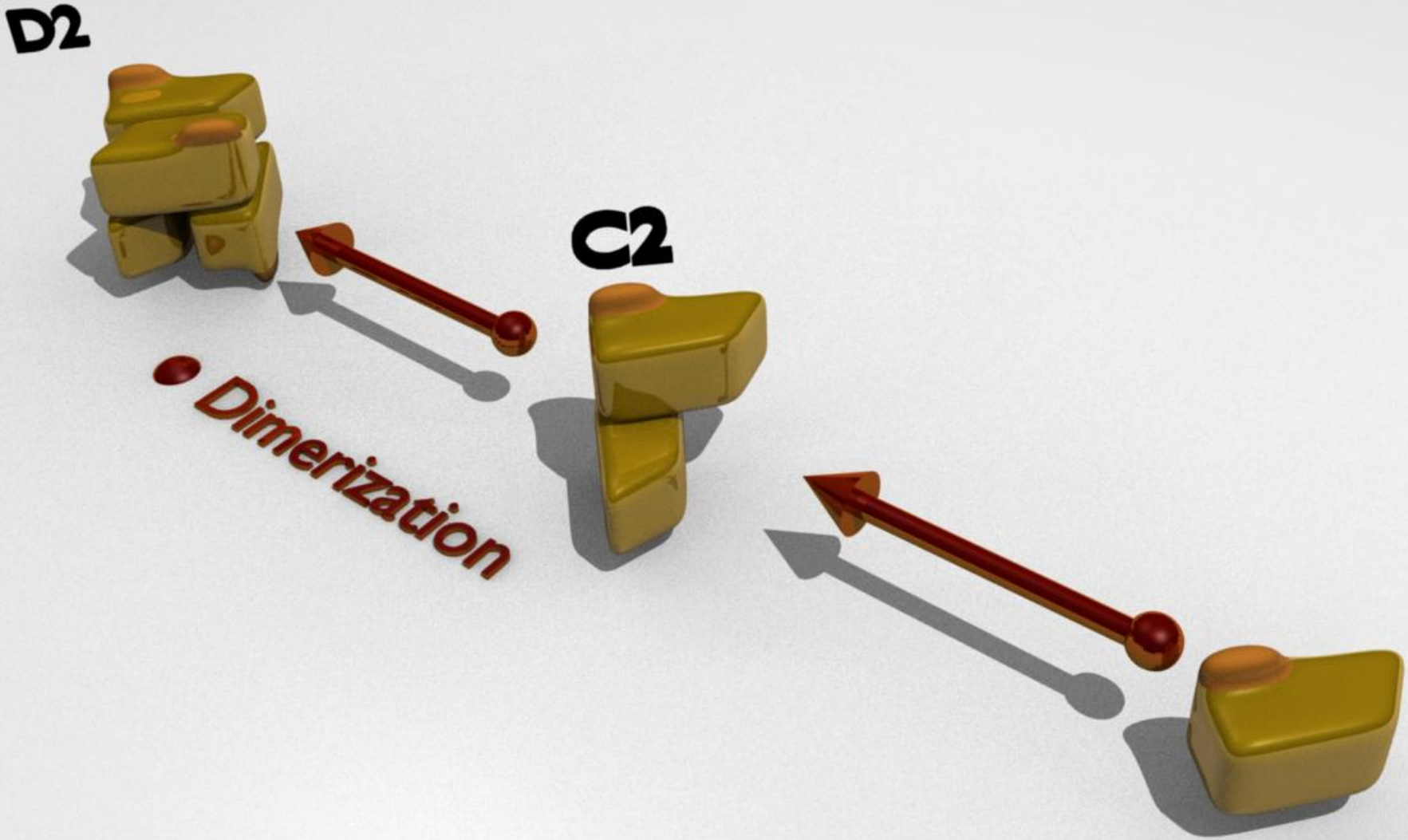


Questions

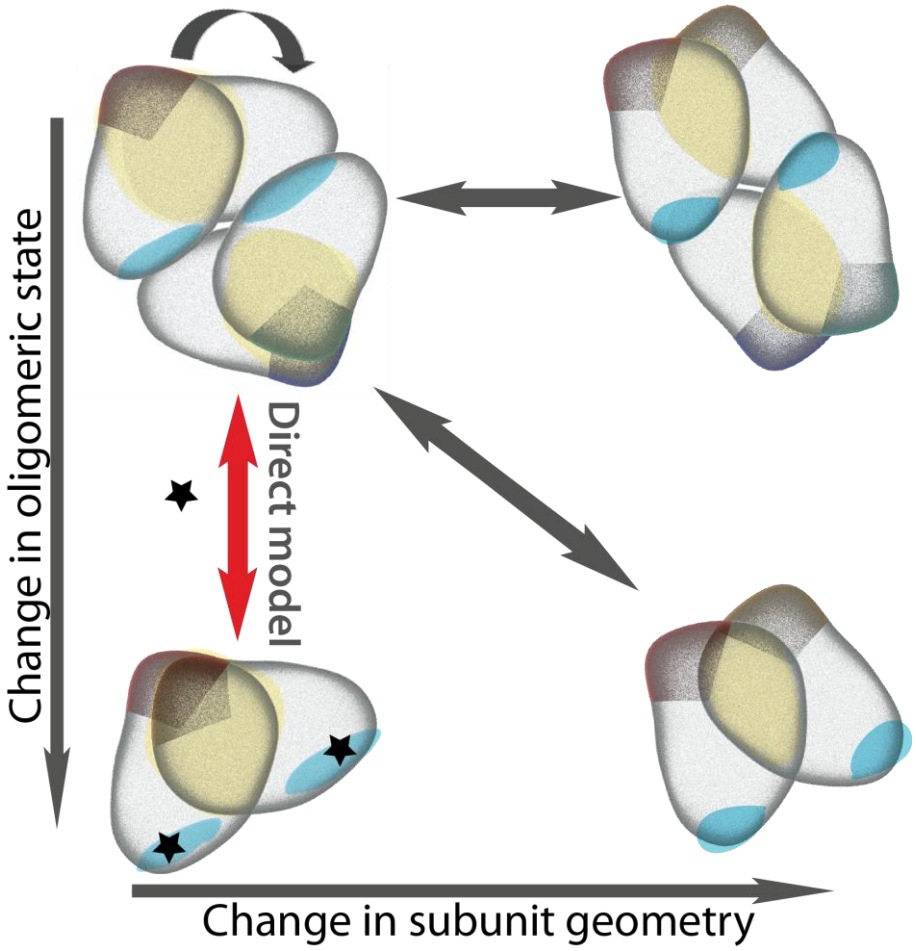
Does assembly of protein complexes drive evolution?

YES – homomers and heteromers

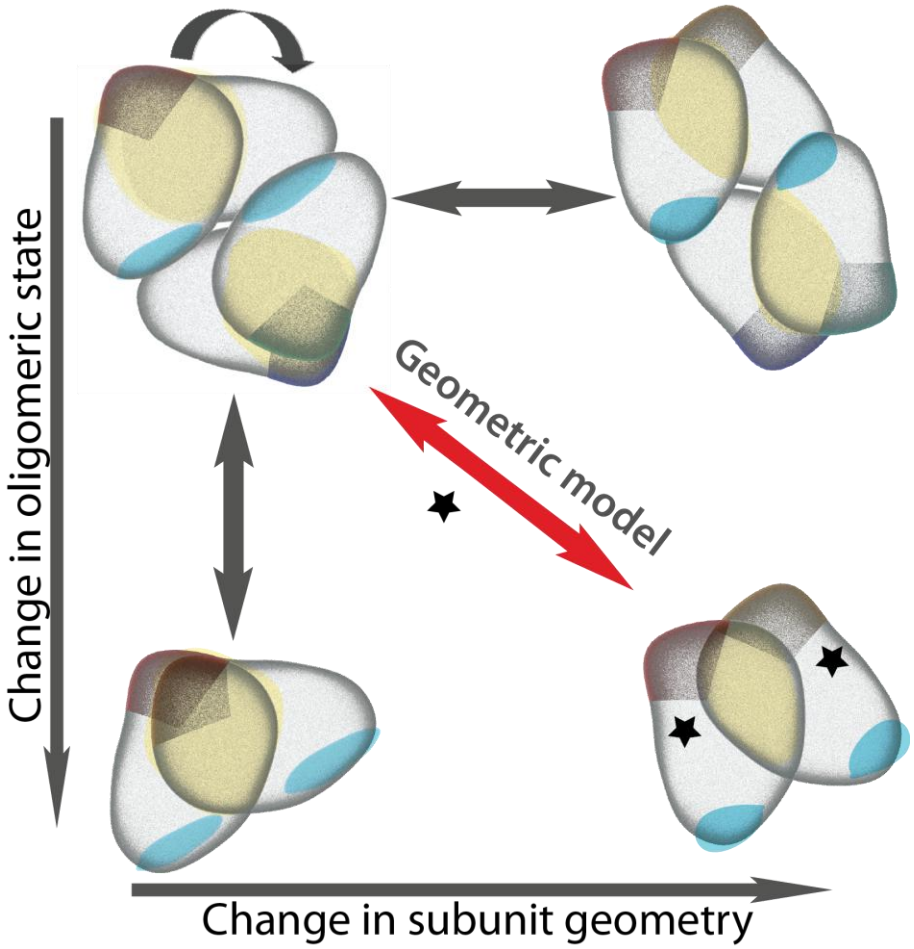
What are mutational mechanisms?


D2
C2
Dimerization

Direct vs indirect effect of mutations



Direct vs indirect effect of mutations

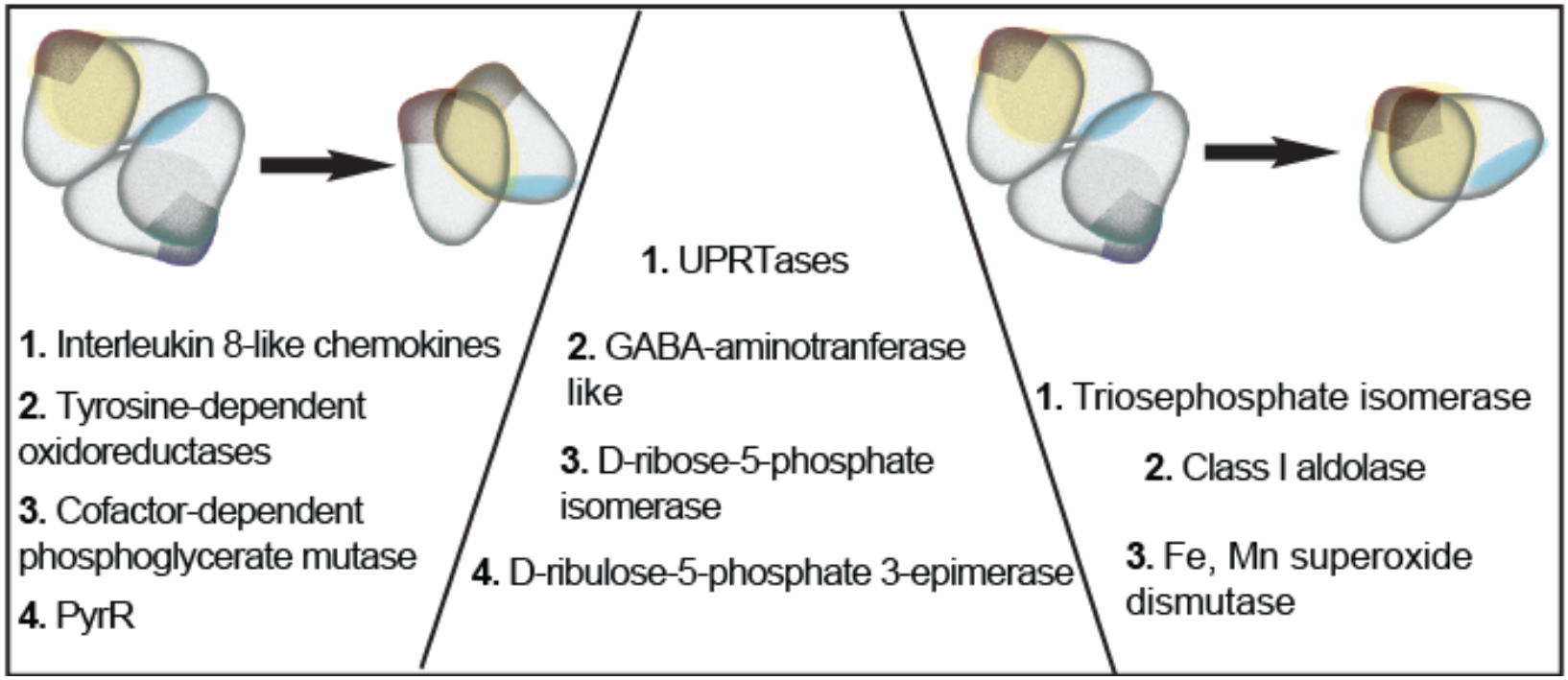


Equally important

4 families
Geometric model

4 families

3 families
Direct model

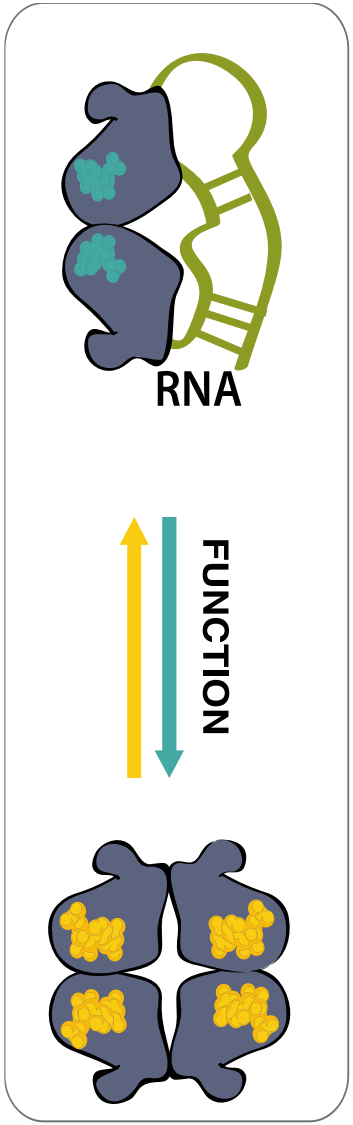


Perica T, Chothia C, and Teichmann SA, PNAS, 2012

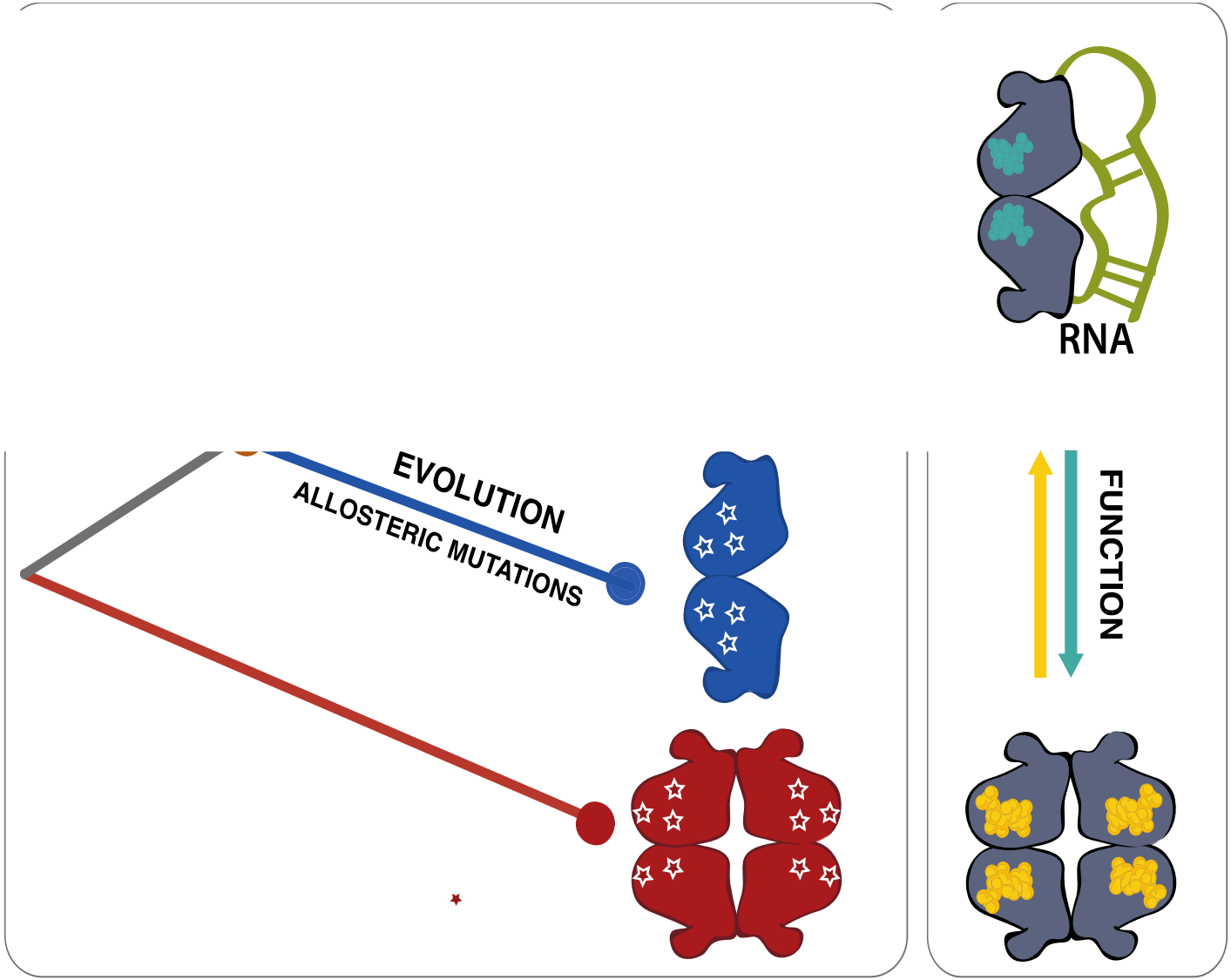
What is the mechanism for quaternary structure evolution?

Indirect mutations?
How do these work?

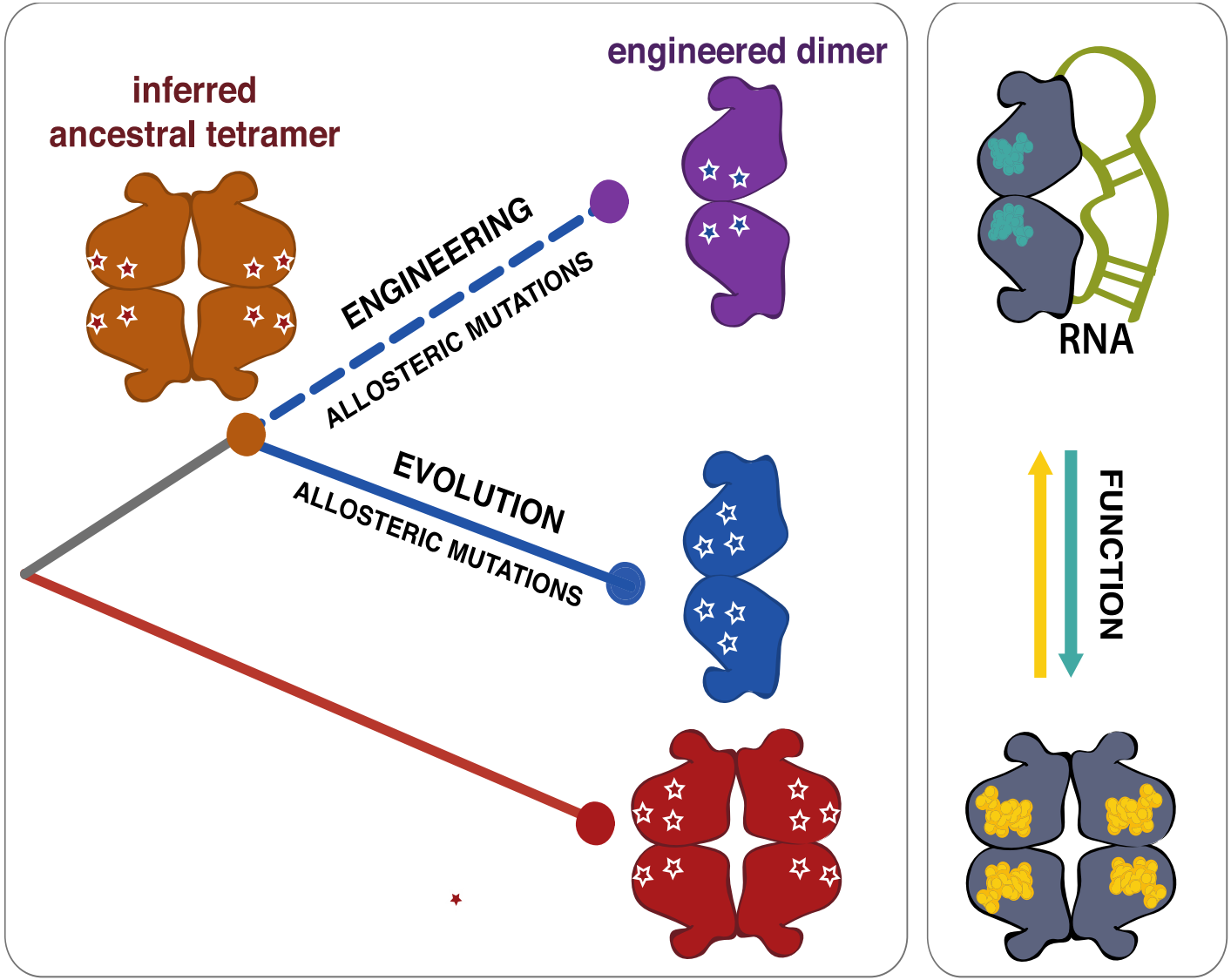
PyrR: a bacterial attenuator family



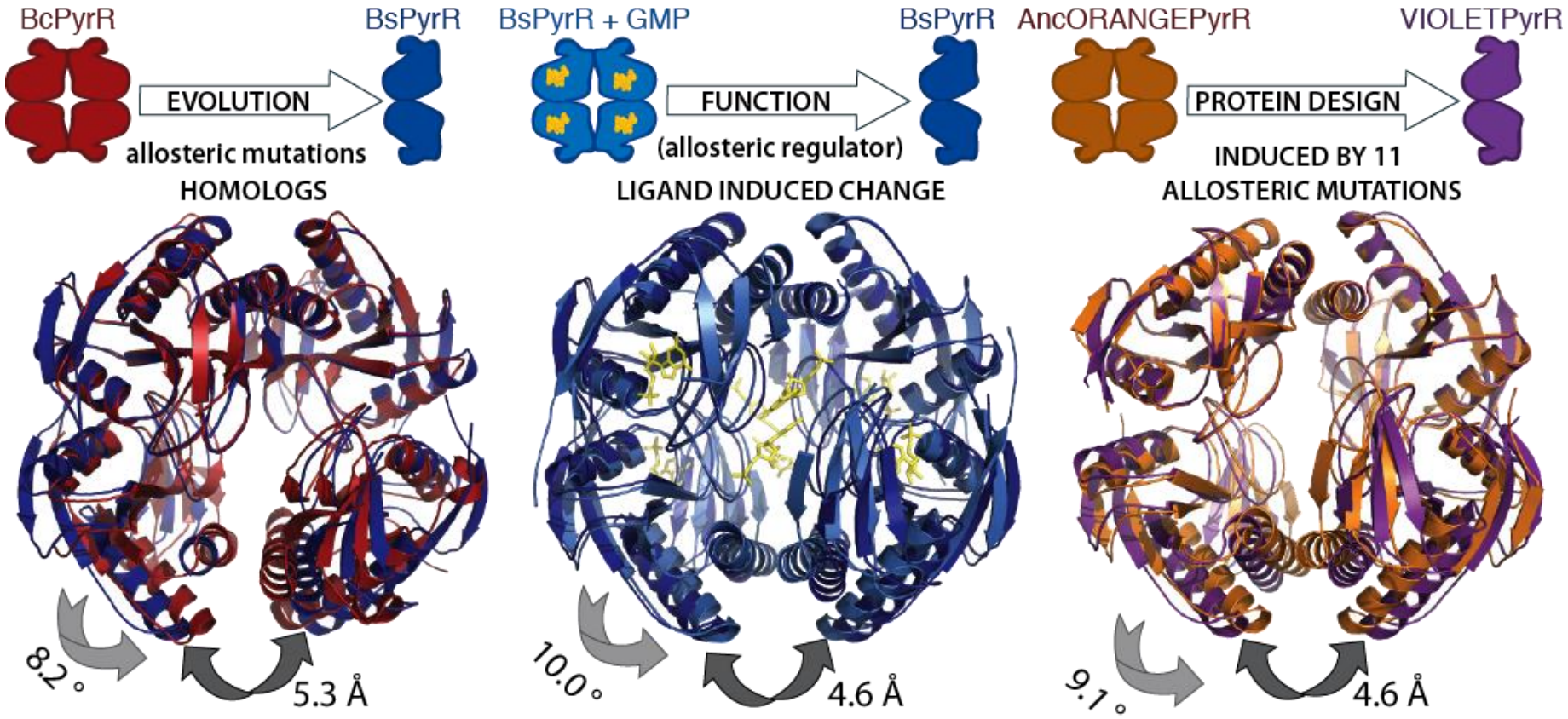
PyrR: a bacterial attenuator family



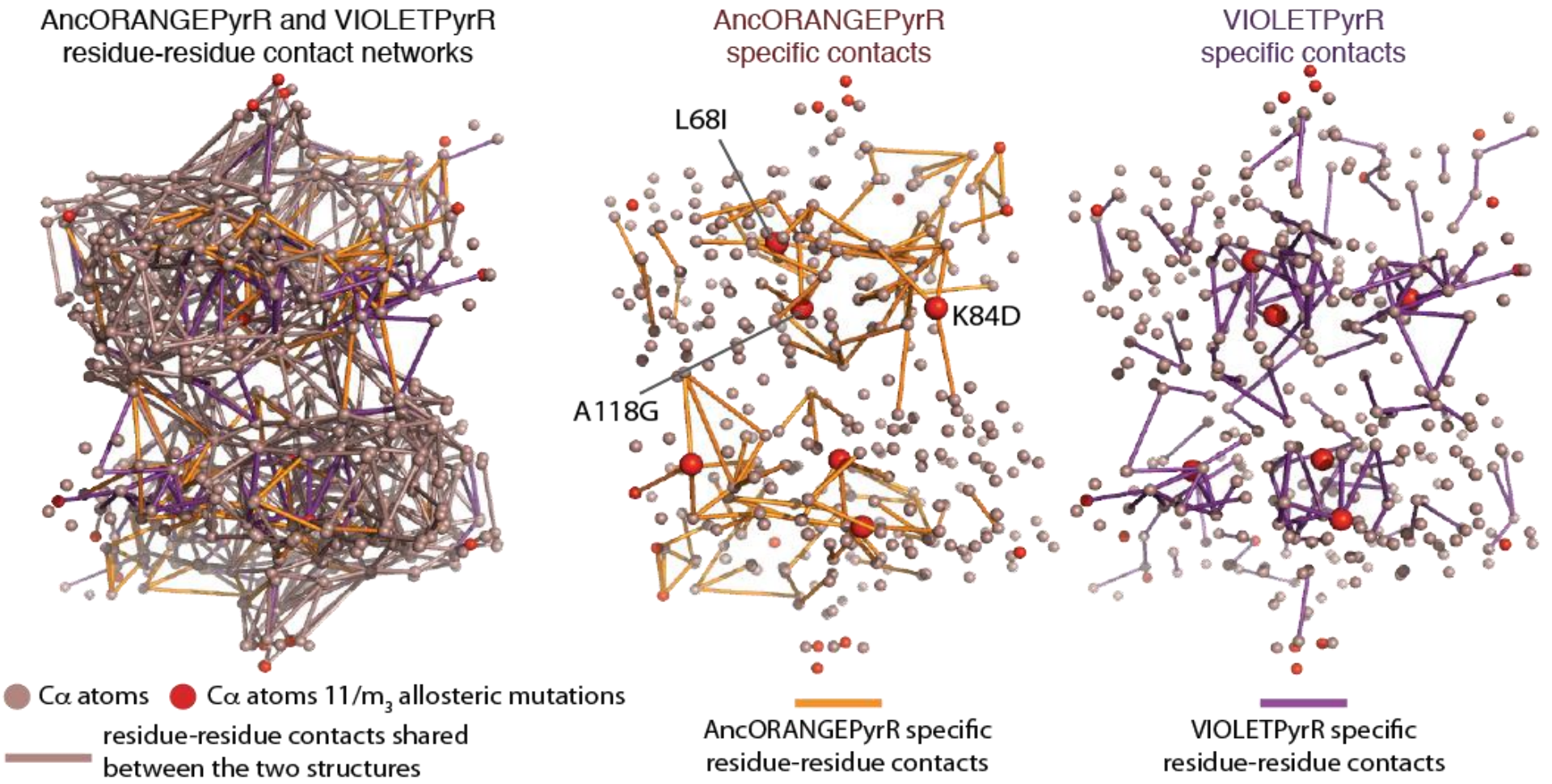
PyrR: a bacterial attenuator family



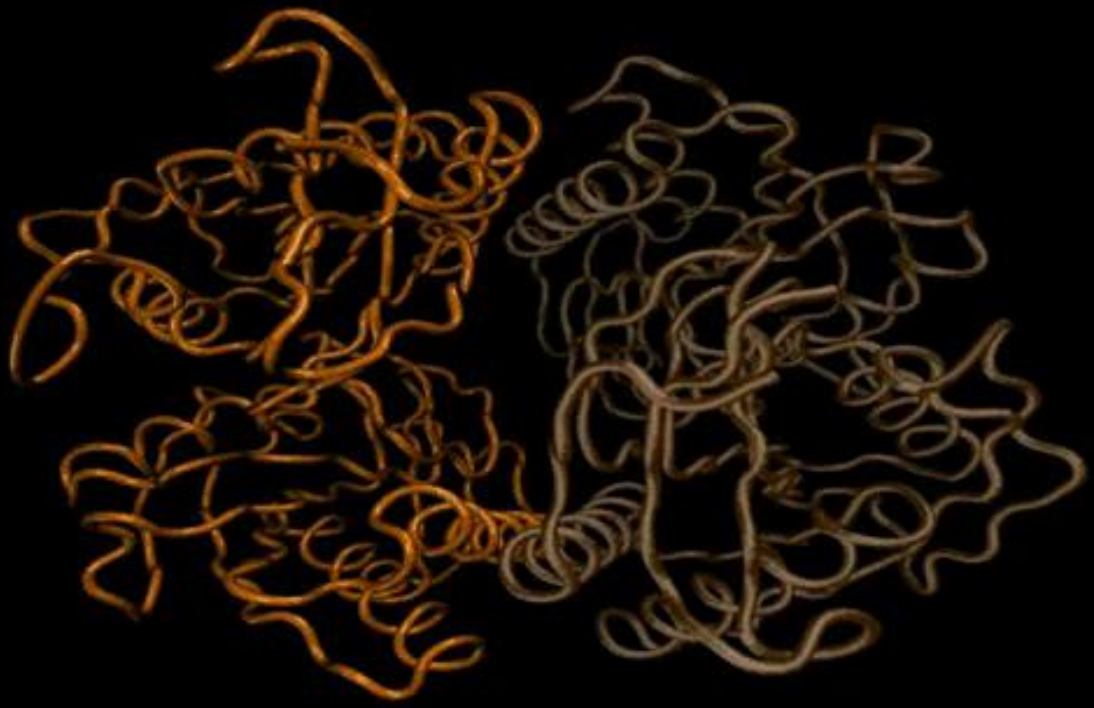
Same geometric changes: allostery and evolution



3 mutations - extensive contact rewiring

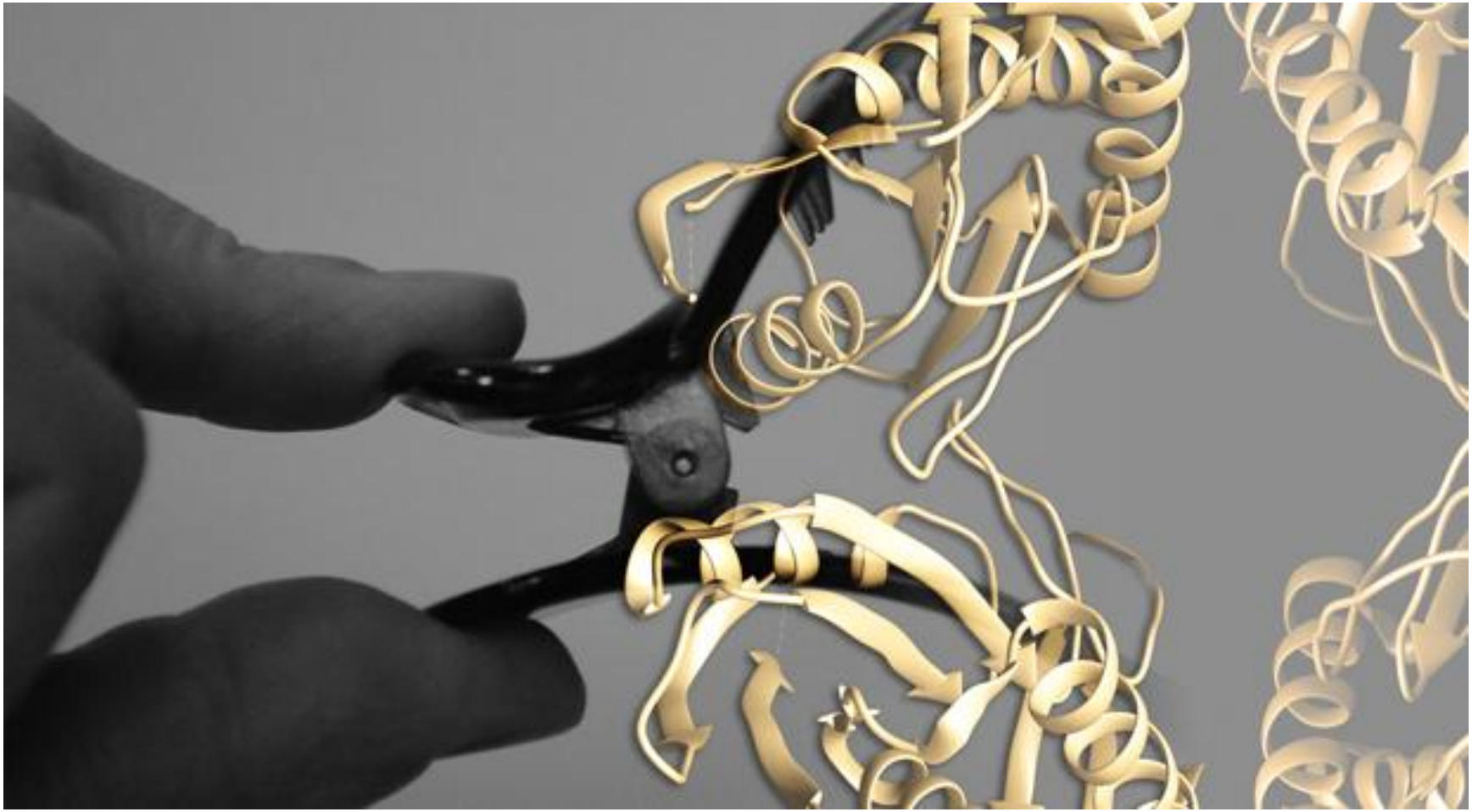


Contact rewiring – dynamics



Same conformational change as allostery

Dimeric unit = Hairclip



Allostery constrains evolutionary path

Perica, T., Kondo, Y., Tiwari, S.P., McLaughlin, S.H., Kemplen, K.R., Zhang, X., Steward, A., Reuter, N., Clarke, J. & Teichmann, S.A. (2014) *Science*, 346, 1254346

Questions

Does assembly of protein complexes drive evolution?

YES – homomers and heteromers

What are mutational mechanisms?

Direct & **indirect** – allosteric mutations

Acknowledgements

Tina Perica

Annette Steward &

Jane Clarke, Chemistry/Cambridge

Stephen McLaughlin, MRC-LMB

Yasushi Kondo, MRC-LMB

Sandhya Tiwari &

Nathalie Reuter, Bergen

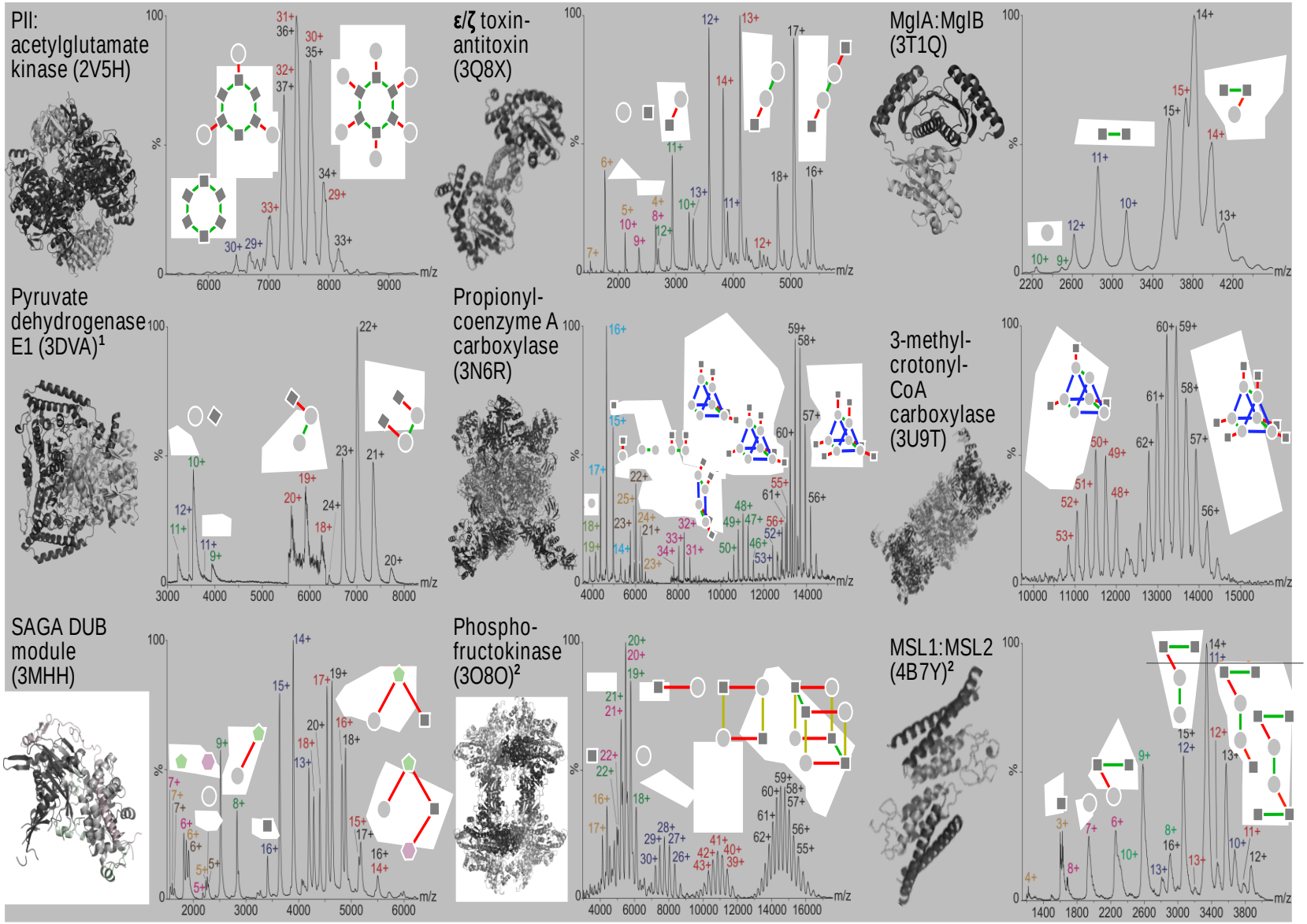
Questions

Does assembly of protein complexes drive evolution?

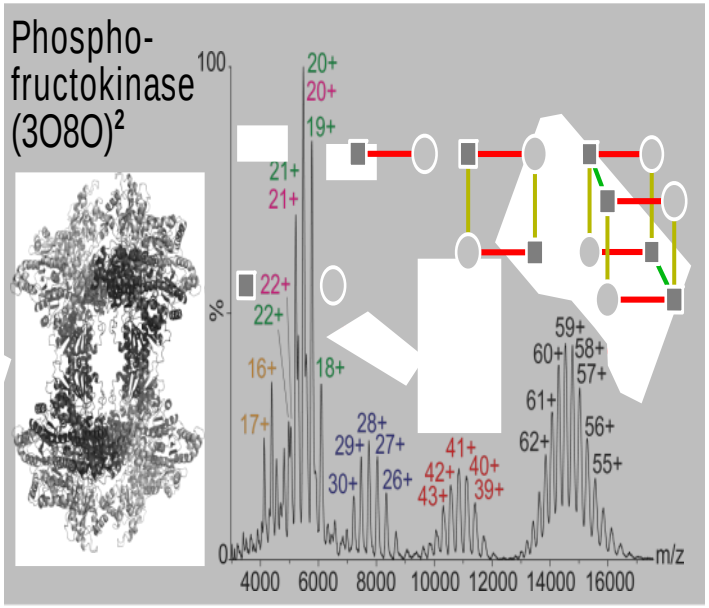
What are mutational mechanisms?

Can principles of assembly to predict topologies?

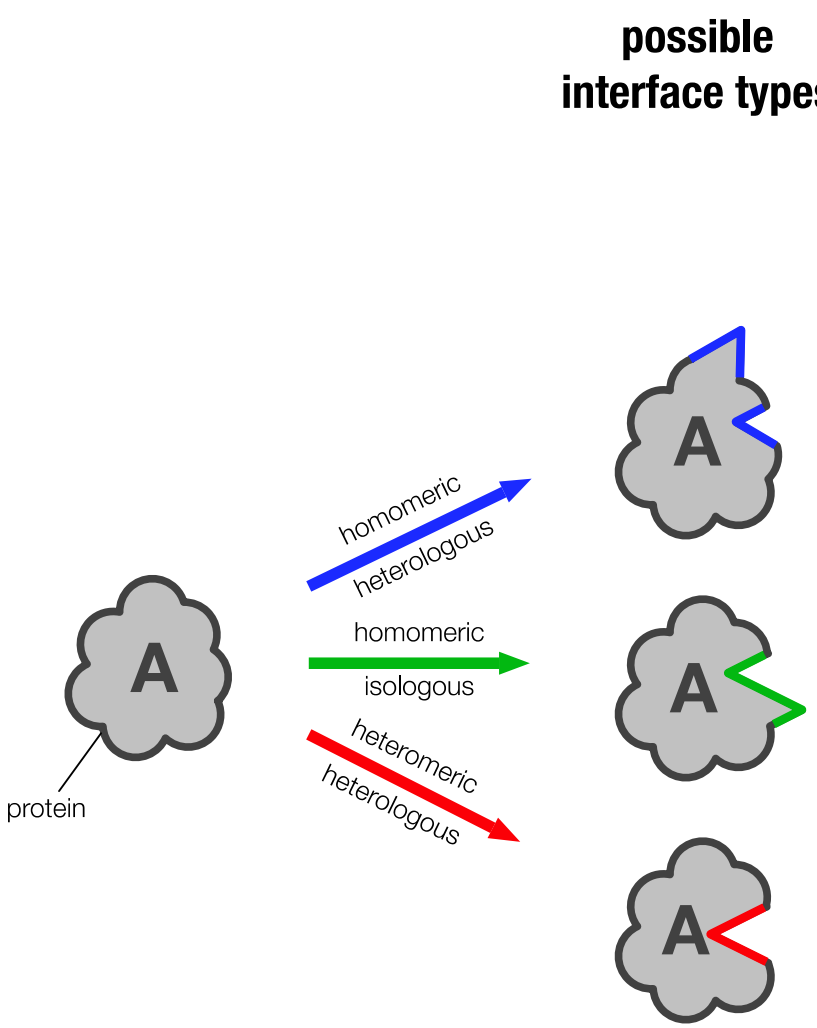
(Dis)assembly intermediates



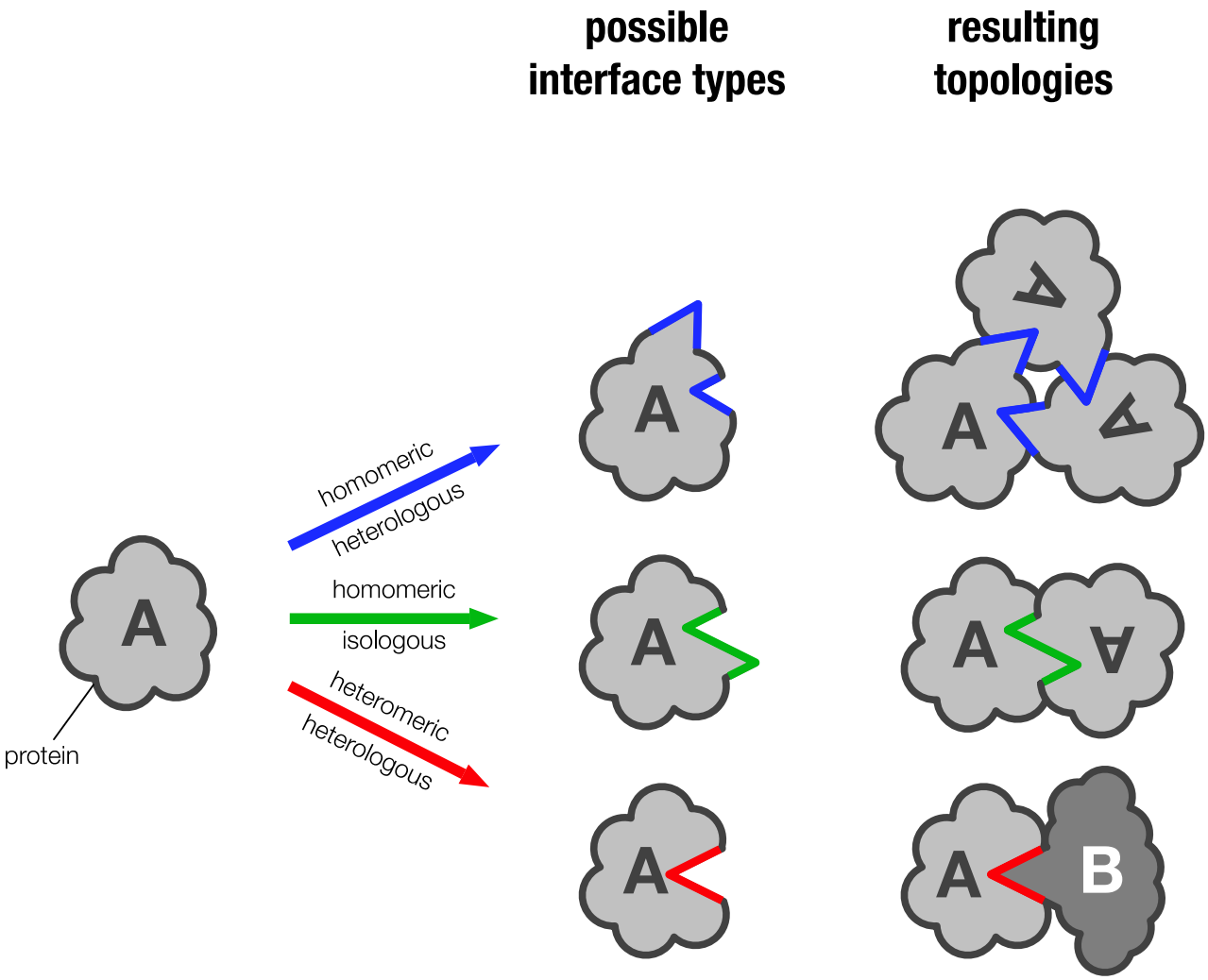
(Dis)assembly intermediates



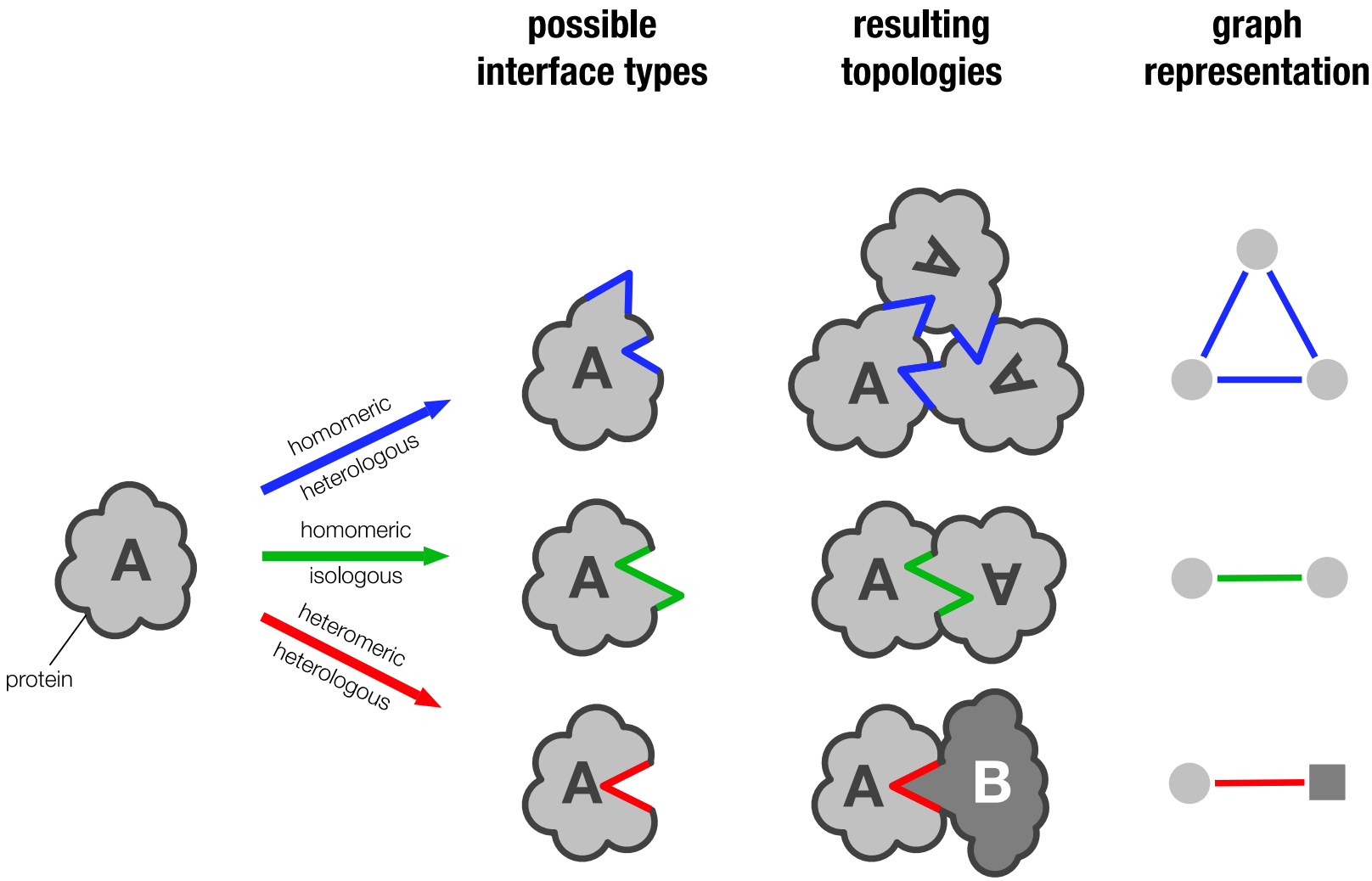
What types of assembly transitions occur?



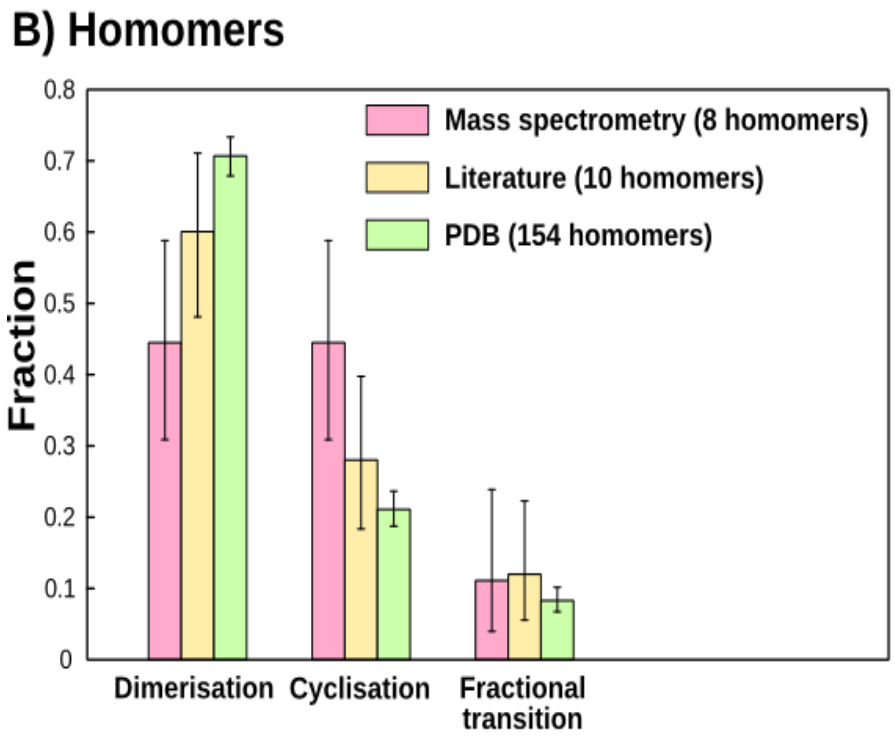
What types of assembly transitions occur?



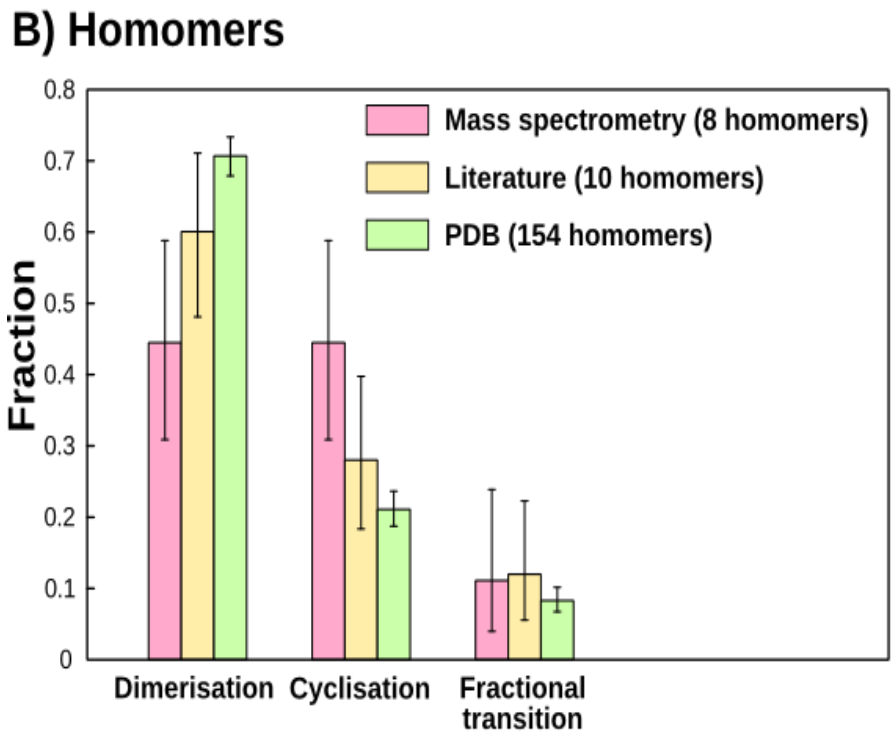
What types of assembly transitions occur?

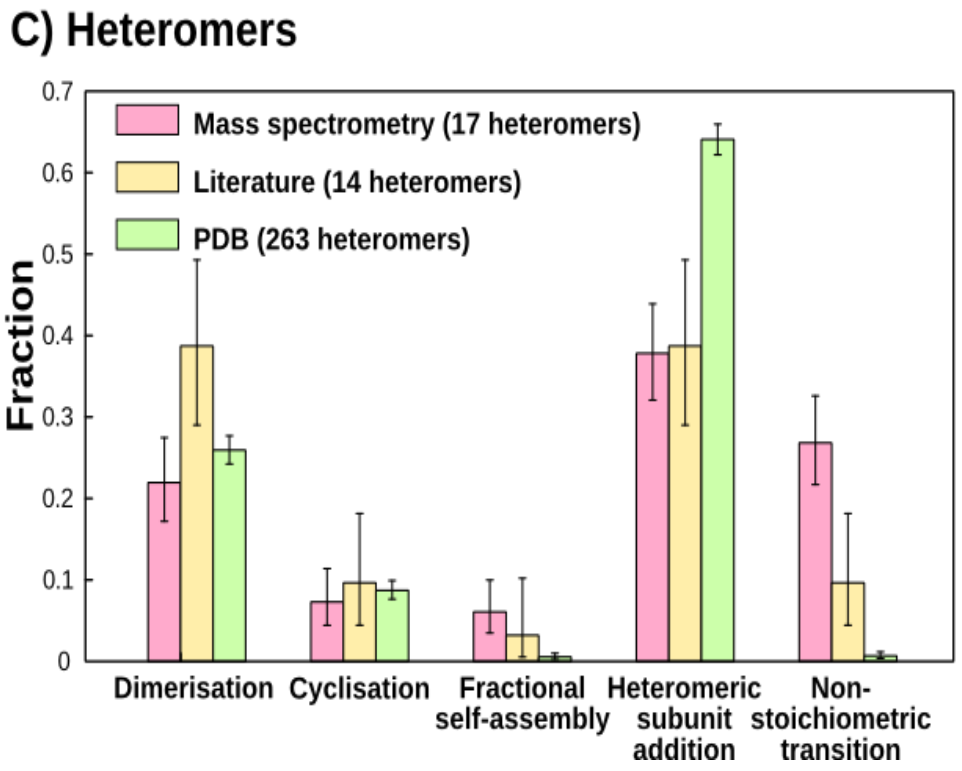


Frequencies of transitions types?

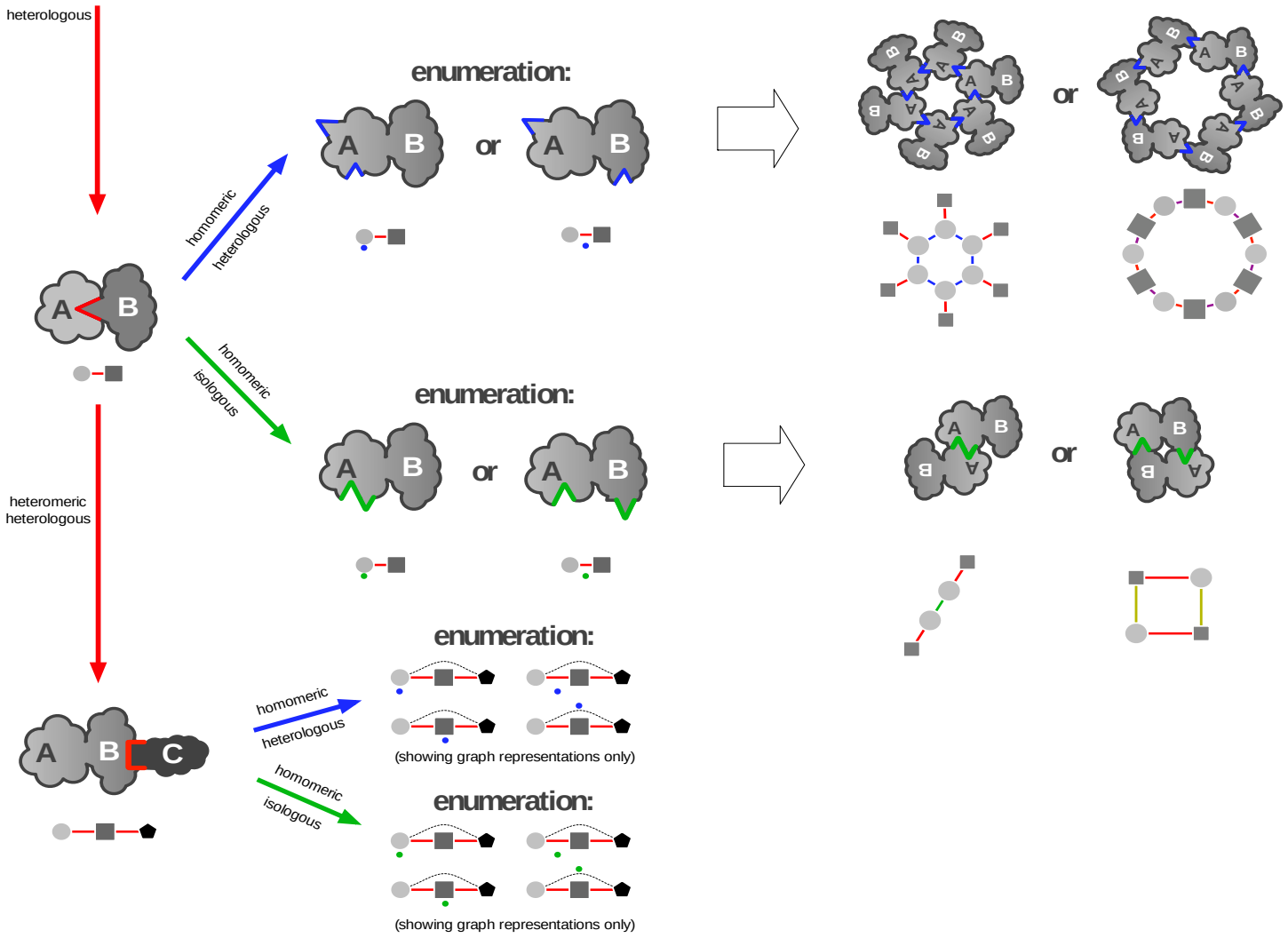


Frequencies of transitions types?

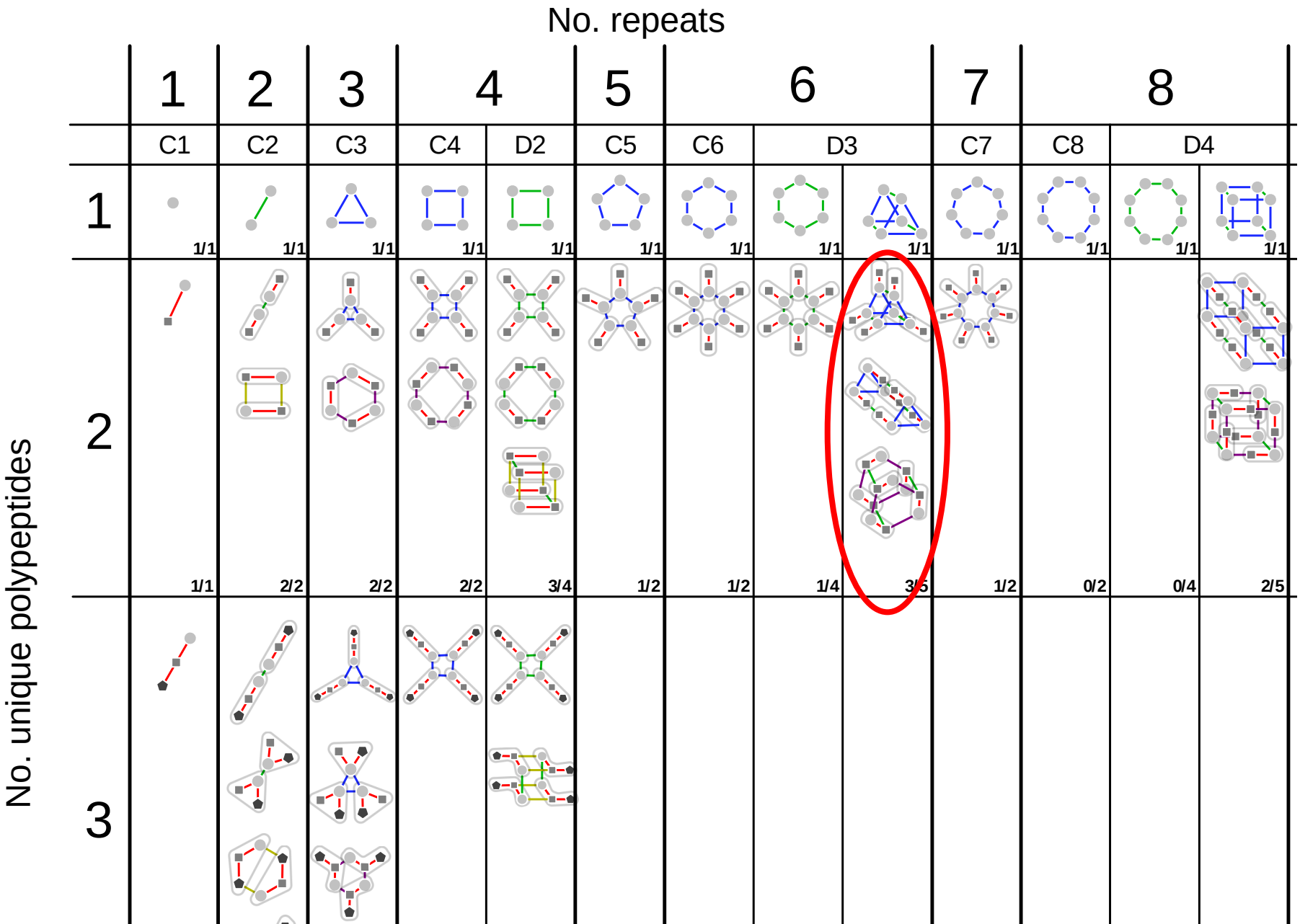




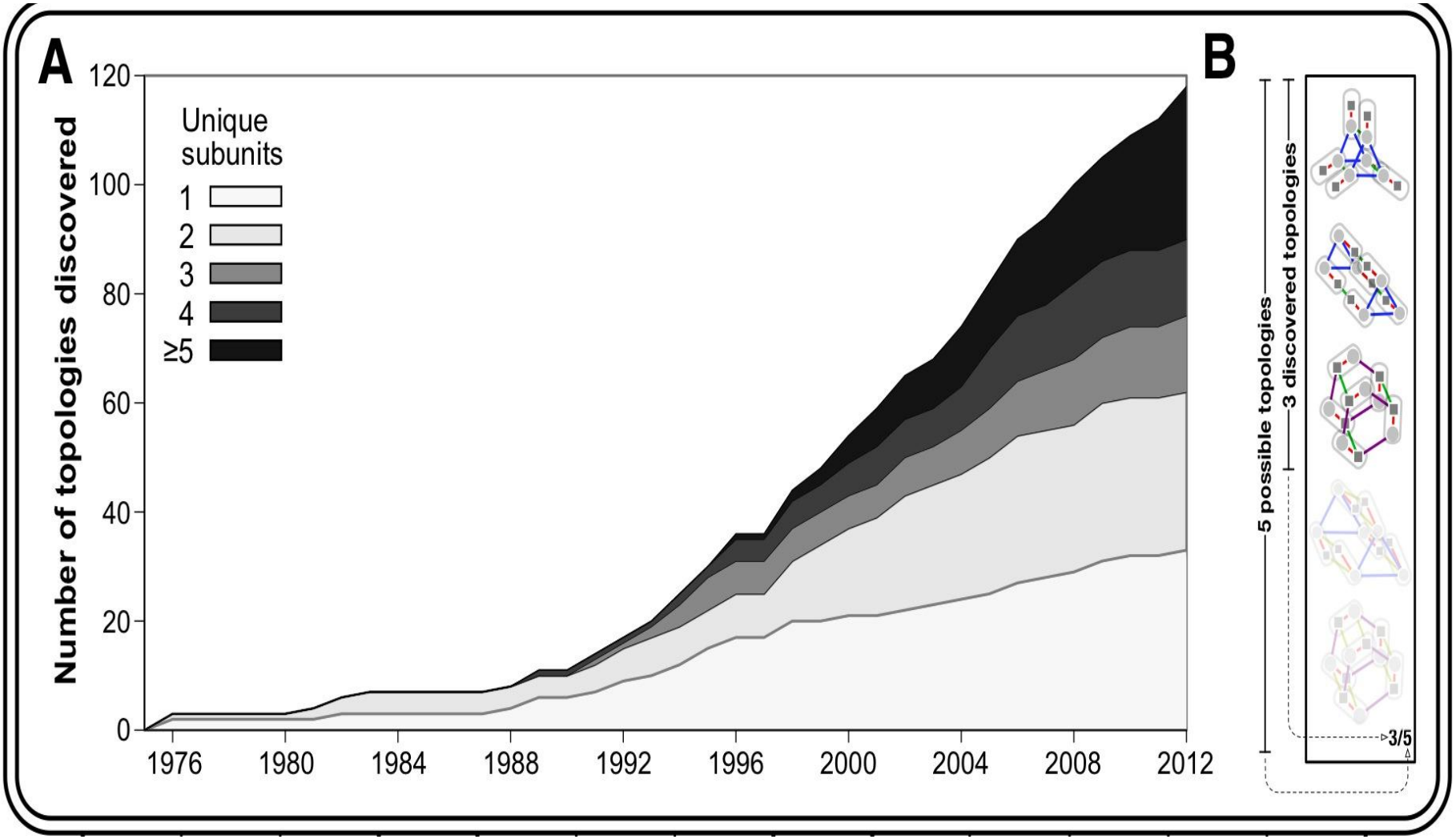
Enumerating all heteromeric topologies



Periodic Table of Protein Complexes

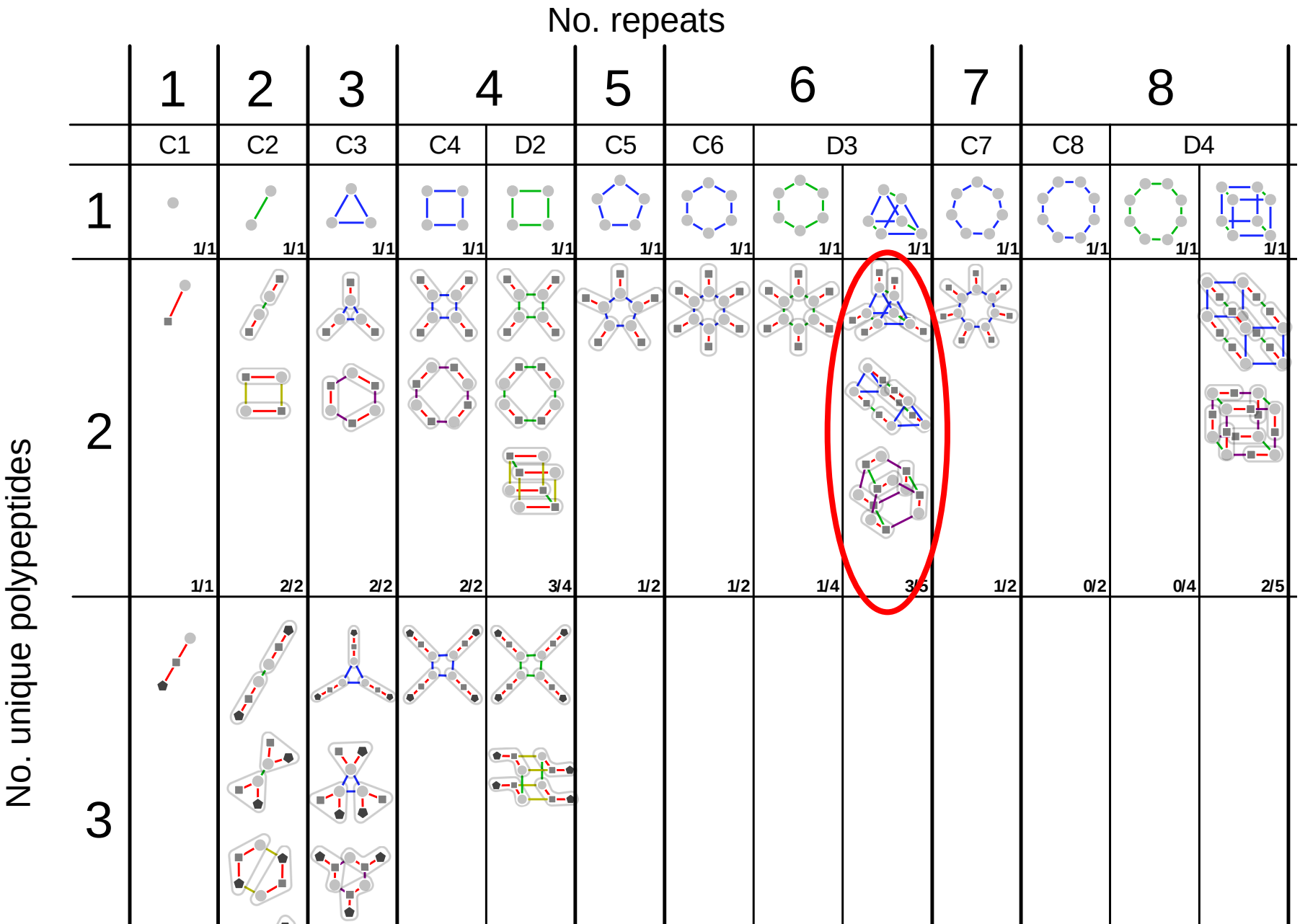


Rate of discovery of new topologies

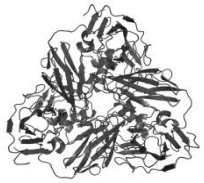
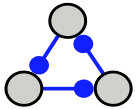
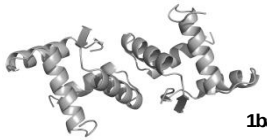
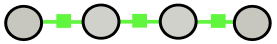
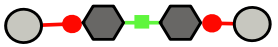
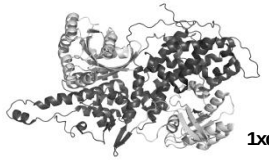
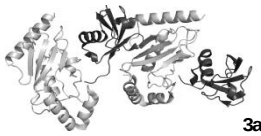


About 4-5 new topologies/year, 120 in total

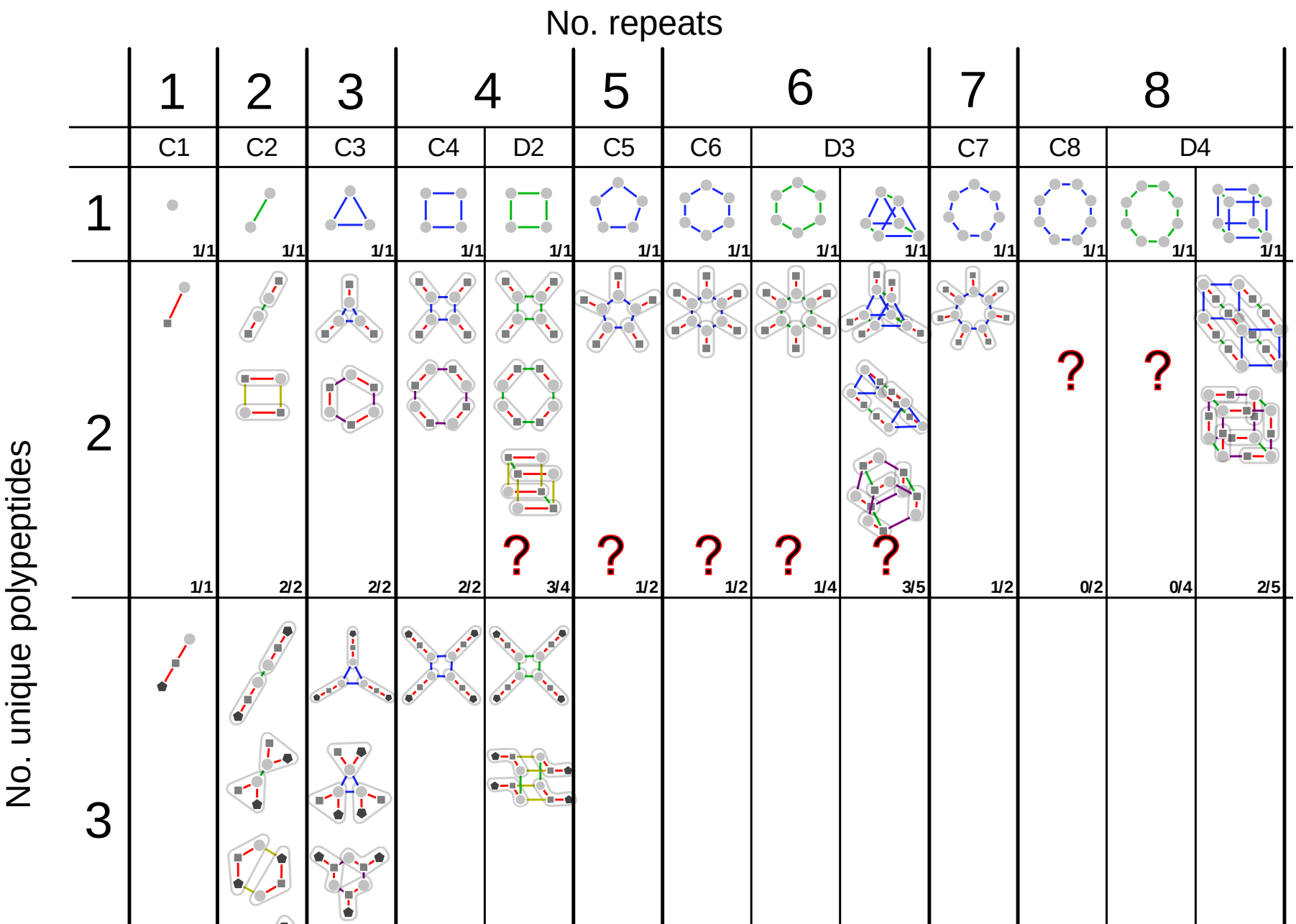
Periodic Table of Protein Complexes



Topologies and Quaternary Structure Errors

	Example	Graph representation	Bijectivity	Occurrence	Error rate
Homomers	 1as6		**bijective homomer** one sequence, one topological environment	**92.5%**	**10.6%**
	 1baz		**non-bijective homomer** one sequence, more than one topological environment	**7.5%**	**60.1%**
Heteromers	 1wbj		**bijective heteromer** multiple sequences, which map bijectively (i.e. one-to-one) to topological environments	**91.7%**	**9.0%**
	 1xd2		**non-bijective heteromer with uneven stoichiometry** multiple sequences, which **do not** map bijectively (i.e. one-to-one) to topological environments, and **do not** all appear an equal number of times.	**6.4%**	**20.4%**
	 3a33		**non-bijective heteromer with even stoichiometry** multiple sequences, which **do not** map bijectively (i.e. one-to-one) to topological environments, but **do** all appear an equal number of times.	**1.9%**	**58.6%**

Periodic Table of Protein Complexes



Graph/motif frequency predictions

- Transcriptional network motifs:

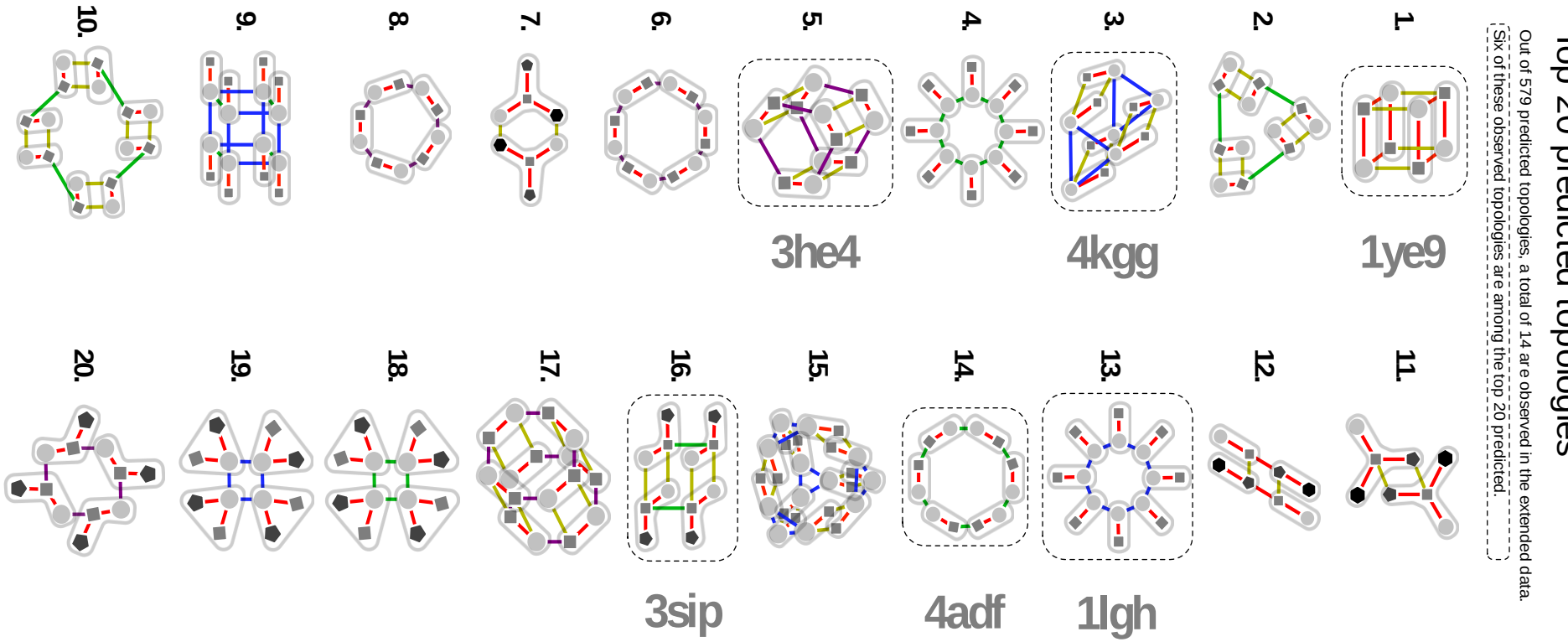
Shen-Orr….Alon U (2004) *Nature Genetics*

- Metabolic network motifs

Ma….Chao TS (2009) *Cell*

- HERE: Protein complex topologies

Predicting frequencies of topologies



6/14 new topologies are in top 20 predictions (out of 567)

Questions

Does assembly of protein complexes drive evolution?

What are mutational mechanisms?

Can principles of assembly to predict topologies?

~120 observed, ca 4 new per year

Predict top new topologies

Identify quaternary structure errors

THANKS TO

- **Joseph Marsh**, HFSP Postdoc @ EBI, now MRC-HGU, Edinburgh

- **Sebastian Ahnert**, Royal Soc URF @ Dept Physics/Cavendish, Cambridge

- **Carol Robinson & Helena Hernandez** @ Dept Chemistry, Oxford