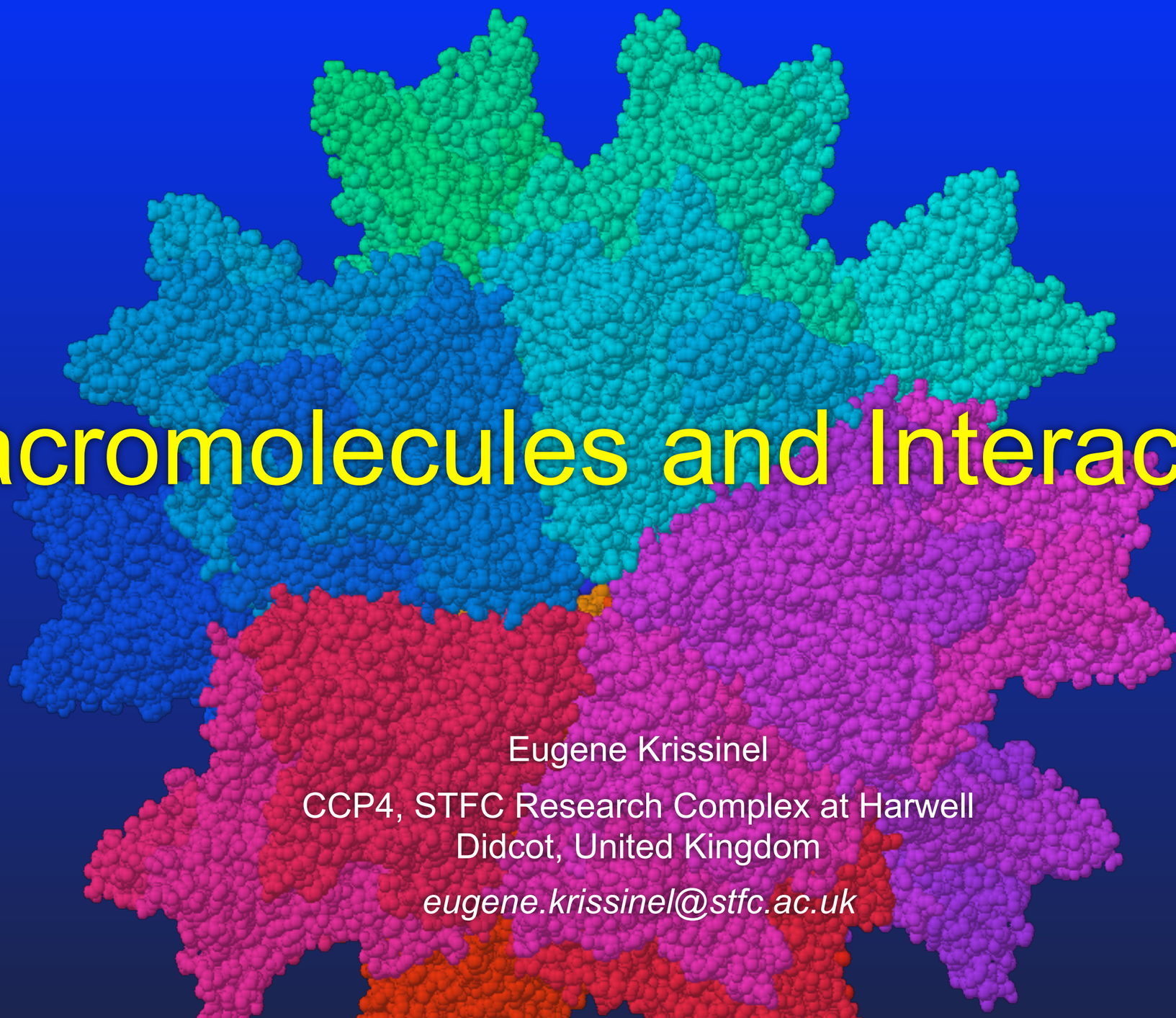


Macromolecules and Interactions

Eugene Krissinel

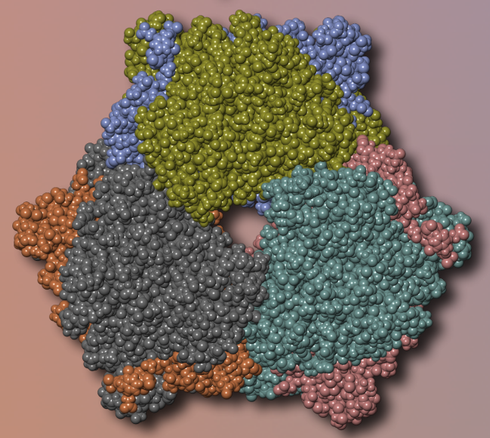
CCP4, STFC Research Complex at Harwell
Didcot, United Kingdom

eugene.krissinel@stfc.ac.uk

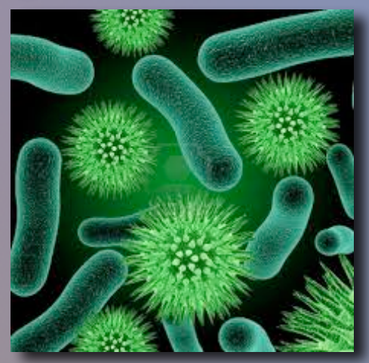
DLS-CCP4 Data Collection and Structure Solution Workshop
Dec 1-8, 2019, Diamond Light Source Ltd., Harwell, UK

MX Schematic Loop

Wild protein



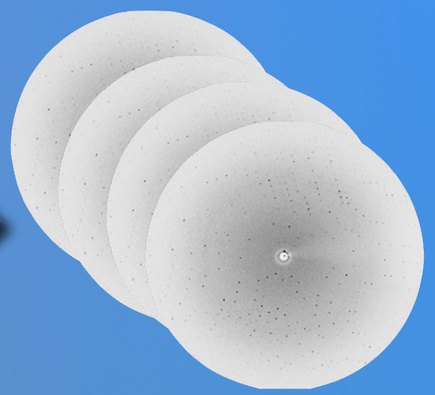
Protein production



Crystallisation



Diffraction

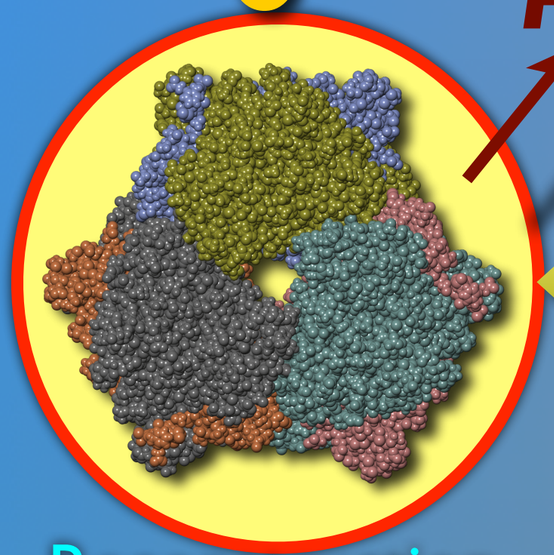


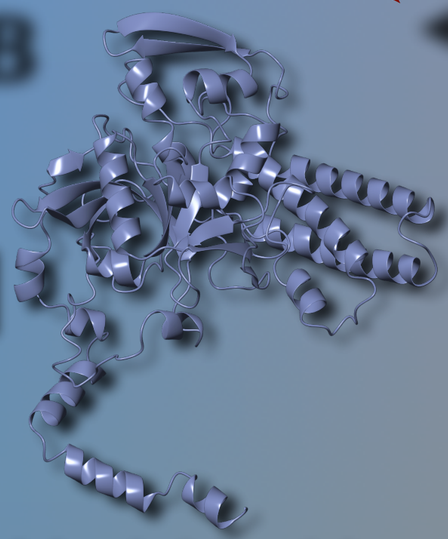


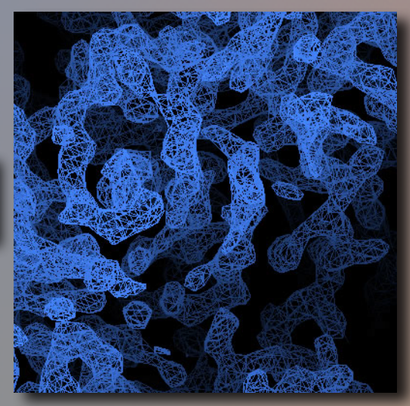
experiment

computations

PDB







H,K,L

Reconstruction and analysis

Model building and Refinement

Phasing

Data processing

What do we solve structures for?

- In general, for learning about interaction between biomolecules:
  - *enzyme functioning*
  - *transport mechanisms*
  - *chemical signalling*
  - .....

- Why structure? There are better methods to study interactions.
  - *hope to understand important interaction in fine details*
  - *hope to learn key features and optimise experimentation*
  - *hope to learn how to predict interactions*
  - *hope to learn how to control interactions (drugs, medicine)*

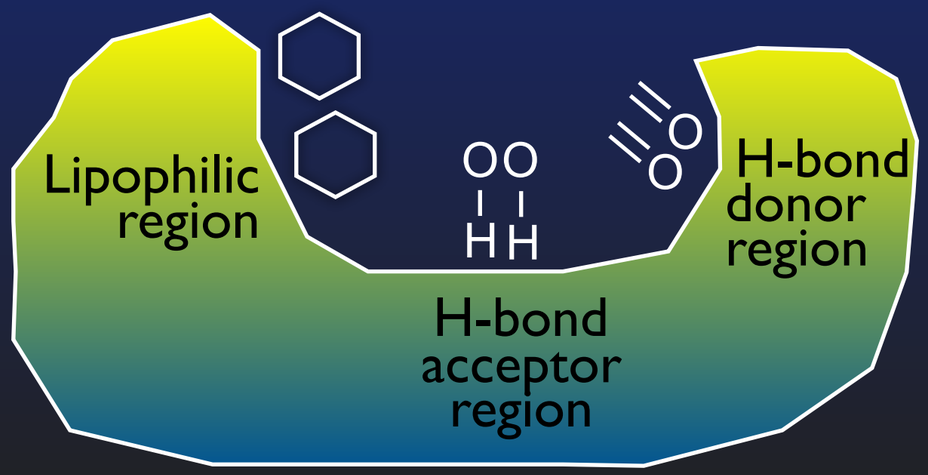


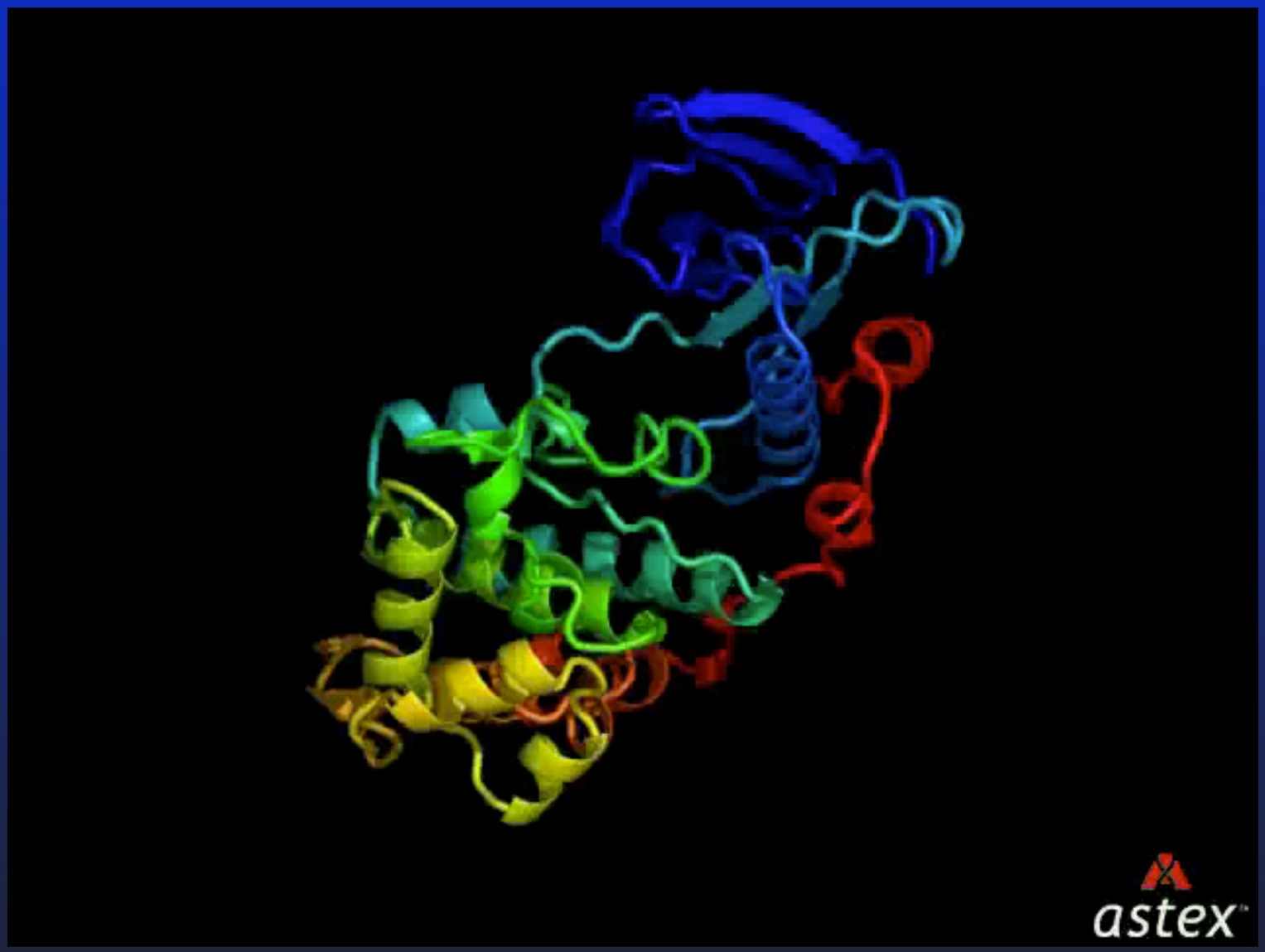
Research Complex at Harwell

Intelligent Drug Design

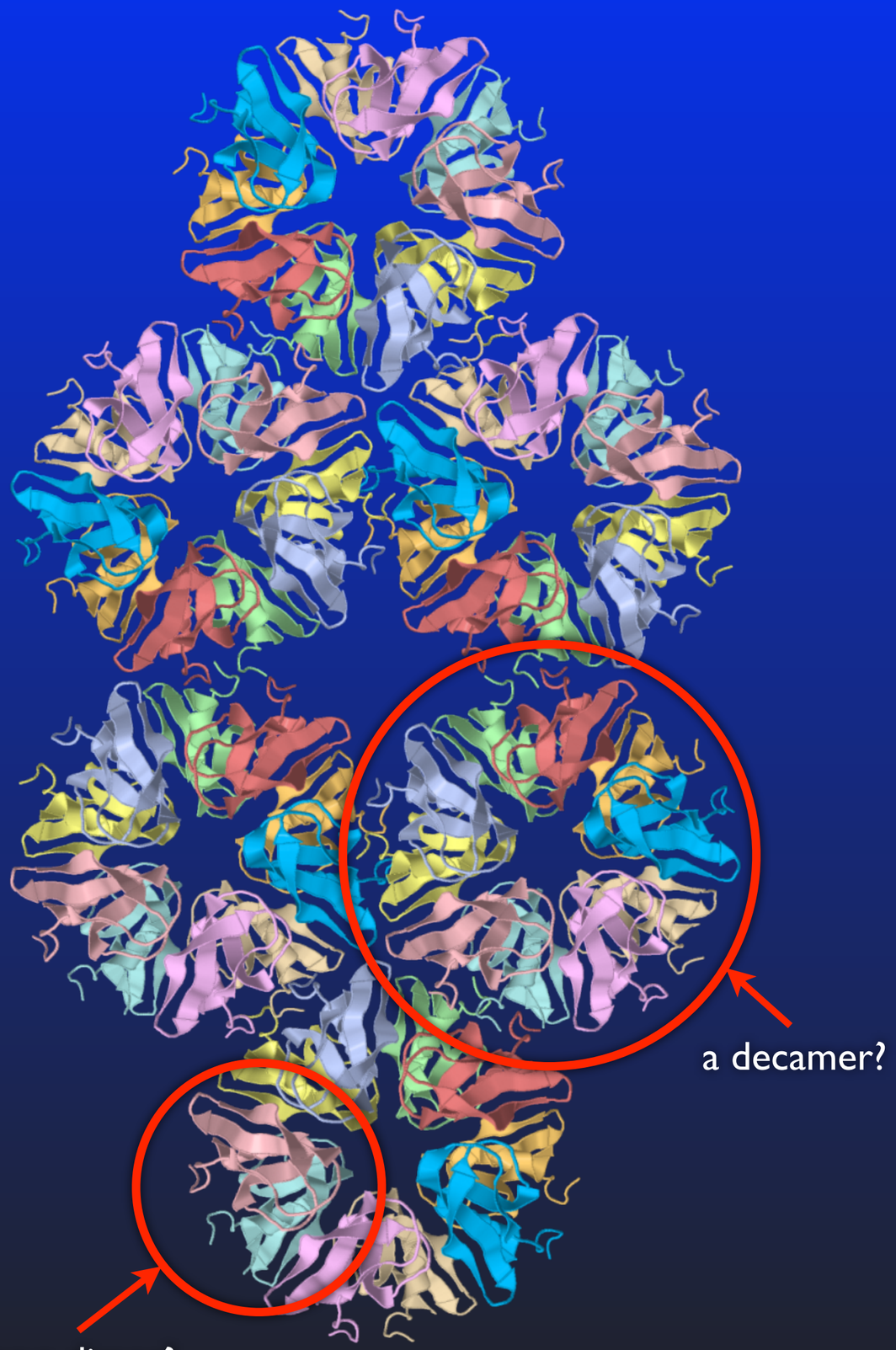
- A generic name for many methods aimed to discover new drugs by means better than at random
  - *Bioinformatics*
  - *Directed Combinatorial Chemistry*
  - *Computer-Assisted Drug Design*
  - *Structure-Based Drug Design*
  - *Fragment-Based Drug Design*

- Basic idea: find a molecule that blocks the "right" protein's active site, or suggest how such a molecule should look like





From web-site of ASTEX Pharmaceutical



★ Macromolecular crystals present us with models of biological structures and interactions between them

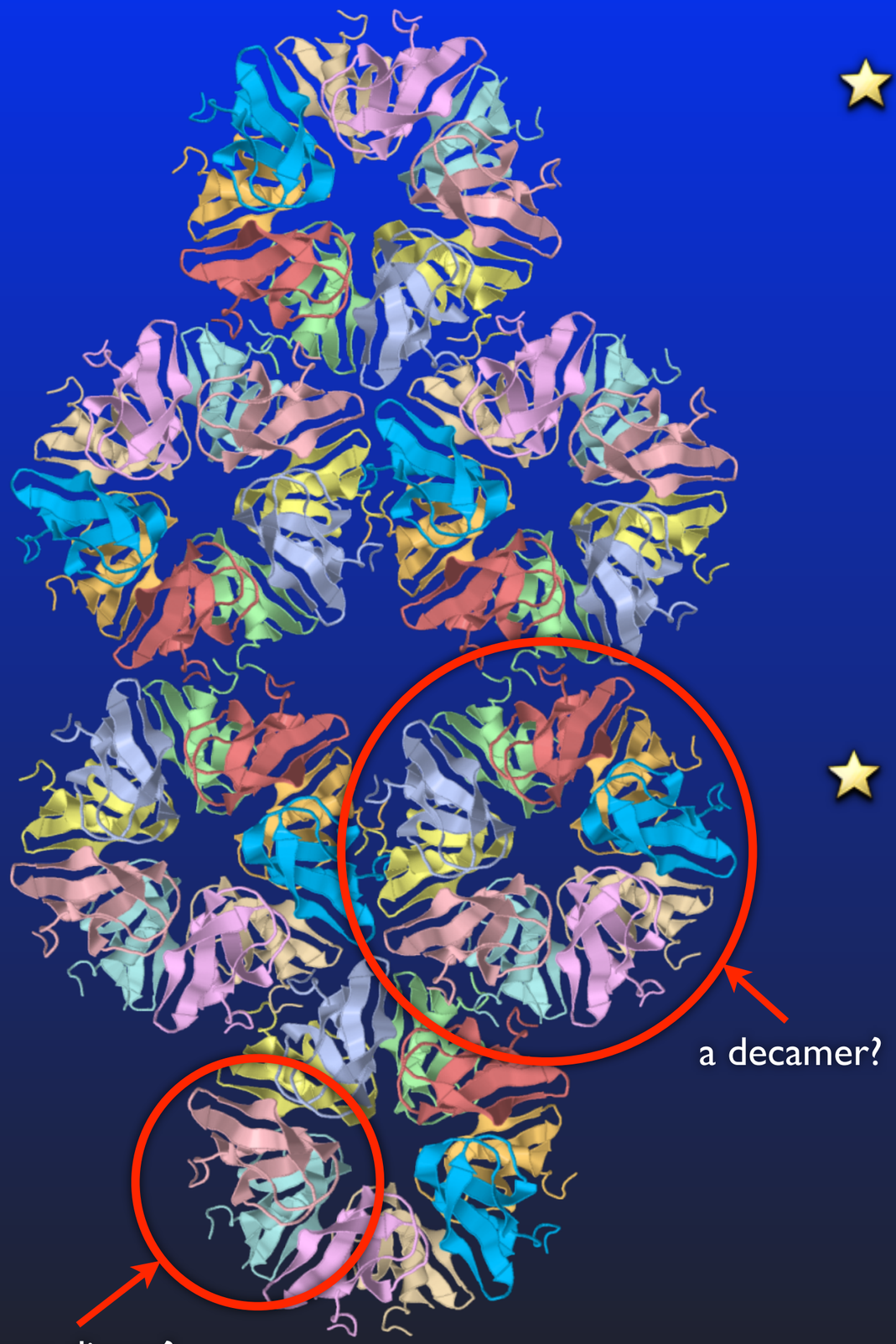
➡ "if you want to know how A interacts with B - crystallise them together!" (crystallographer's sweet dream, *but does this always work?*)

➡ interactions make complexes

➡ complexes make biology

➡ biology tells which drug





★ Crystals present us with both real ("*significant*") and artifactual interactions, which may be difficult to differentiate. Frequently used techniques:

Rules of thumb: e.g. manifestation in different crystal forms

Experimental: complementing studies (MS, EM, NMR, scattering)

Bioinformatical: homology and interface similarity analysis

Computational: energy estimates and modelling

★ PISA software infers significant interactions and macromolecular assemblies from crystal data by evaluating their free Gibbs energy:

$$\Delta G_0 = -\Delta G_{\text{int}} - T\Delta S \quad > \quad 0$$

http://www.ccp4.ac.uk/pisa



Research Complex at Harwell

$$\Delta G_0 = -\Delta G_{int} - T\Delta S > 0$$



The PDB does indeed contain a wealth of experimental data on macromolecular complexes

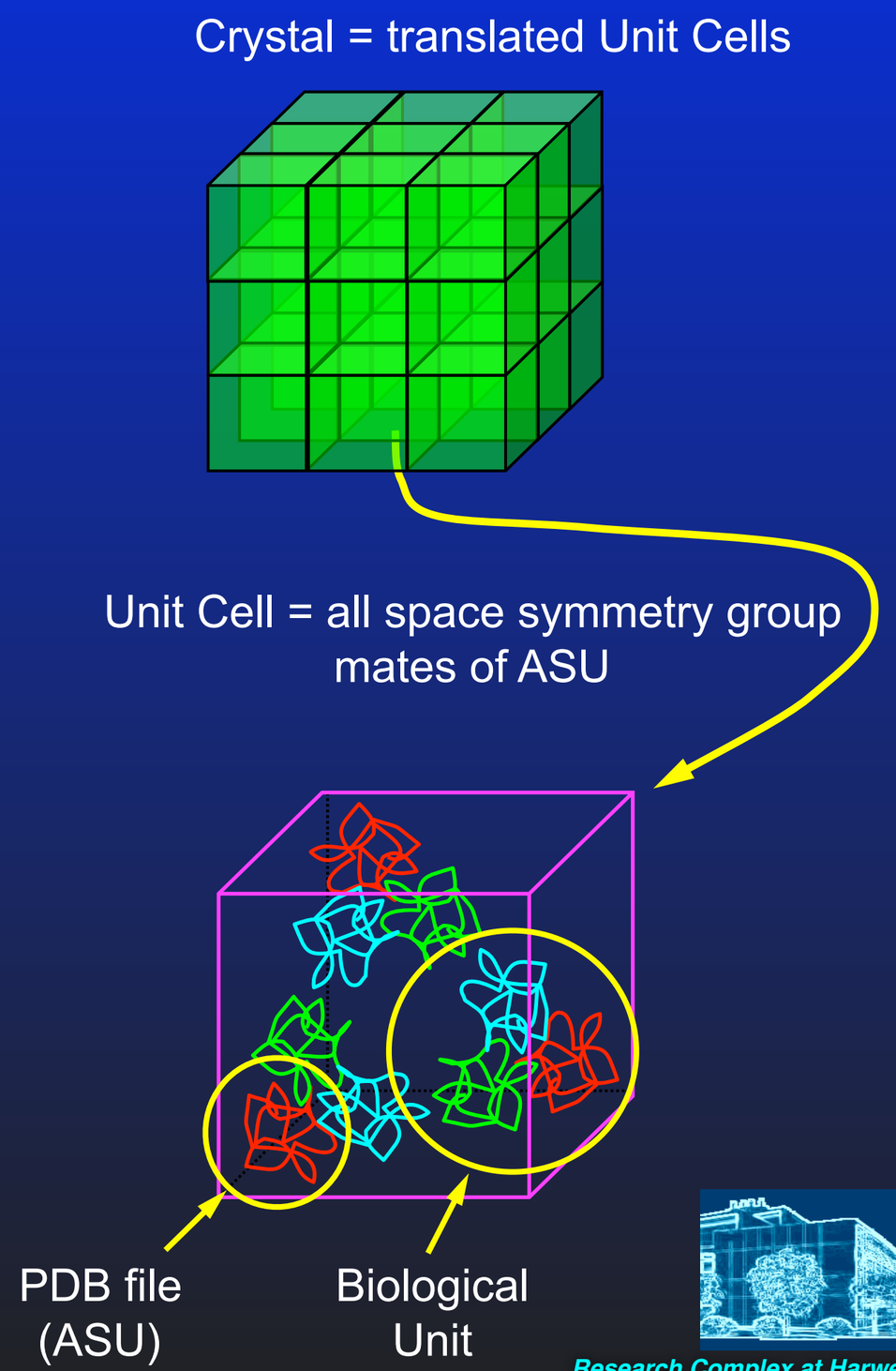
More than 80% of macromolecular structures are solved by means of X-ray diffraction on crystals.

Any crystal represents macromolecular interactions and associations through inter-molecular interfaces

An X-ray diffraction experiment produces atomic coordinates of the Asymmetric Unit (ASU), which is stored as a PDB file.

In general, neither ASU nor Unit Cell has any direct relation to PQS. The PQS may be made of

- a single ASU
- a part of ASU
- several ASU
- several ASU parts



Detection of Biological Units in Crystals: PISA summary

1. Enumerate all possible multimeric assemblies in crystal packing, subject to crystal properties: space symmetry group, geometry and composition of the Asymmetric Unit
    - Achieved with graph-theoretical techniques, by representing crystal as an infinite periodic graph of connected macromolecules
    - Equivalent to splitting the crystal in all possible ways over groups of chemically equivalent interfaces, by considering each group to be engaged or disengaged

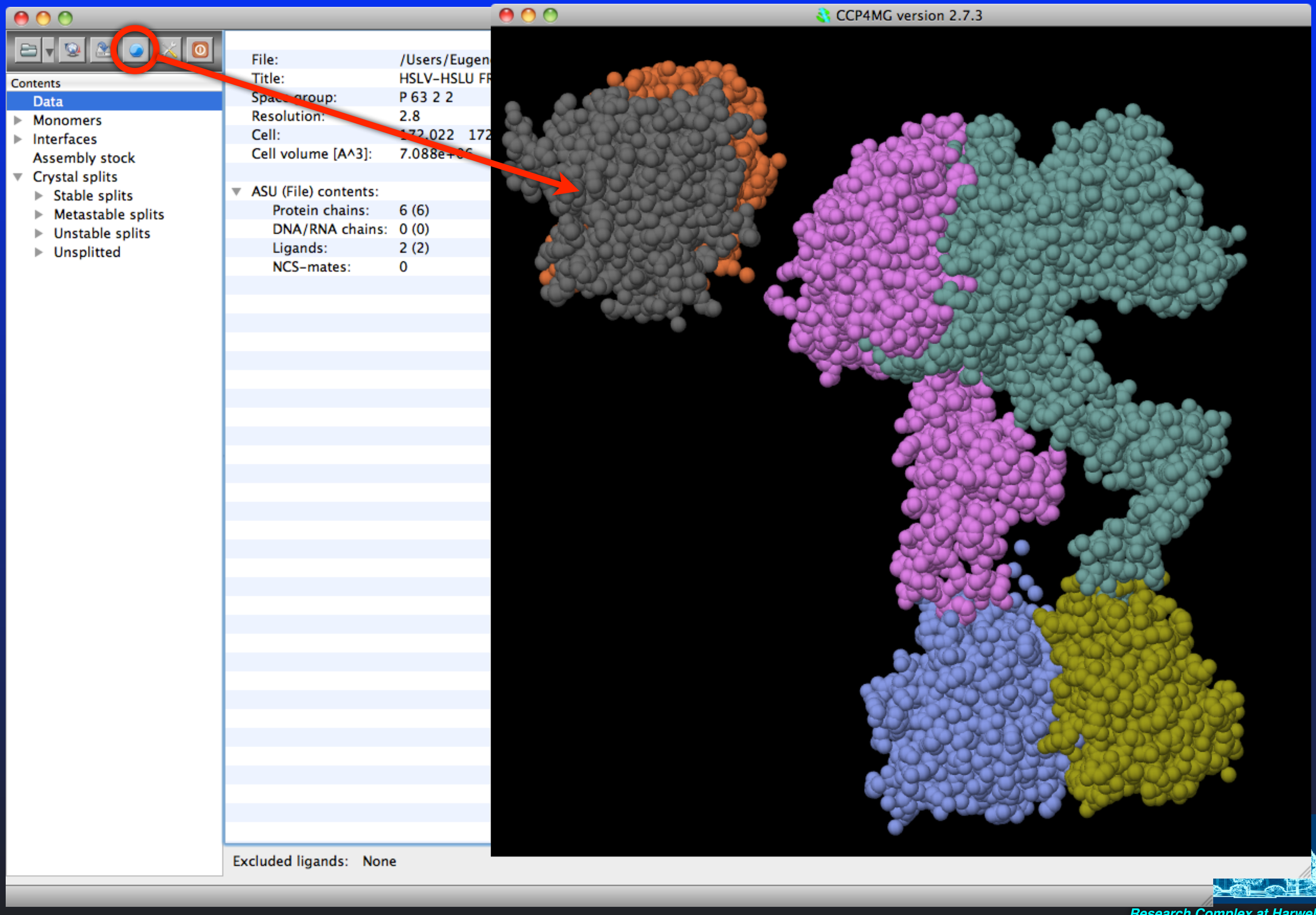
2. Evaluate all candidate assemblies for chemical stability:

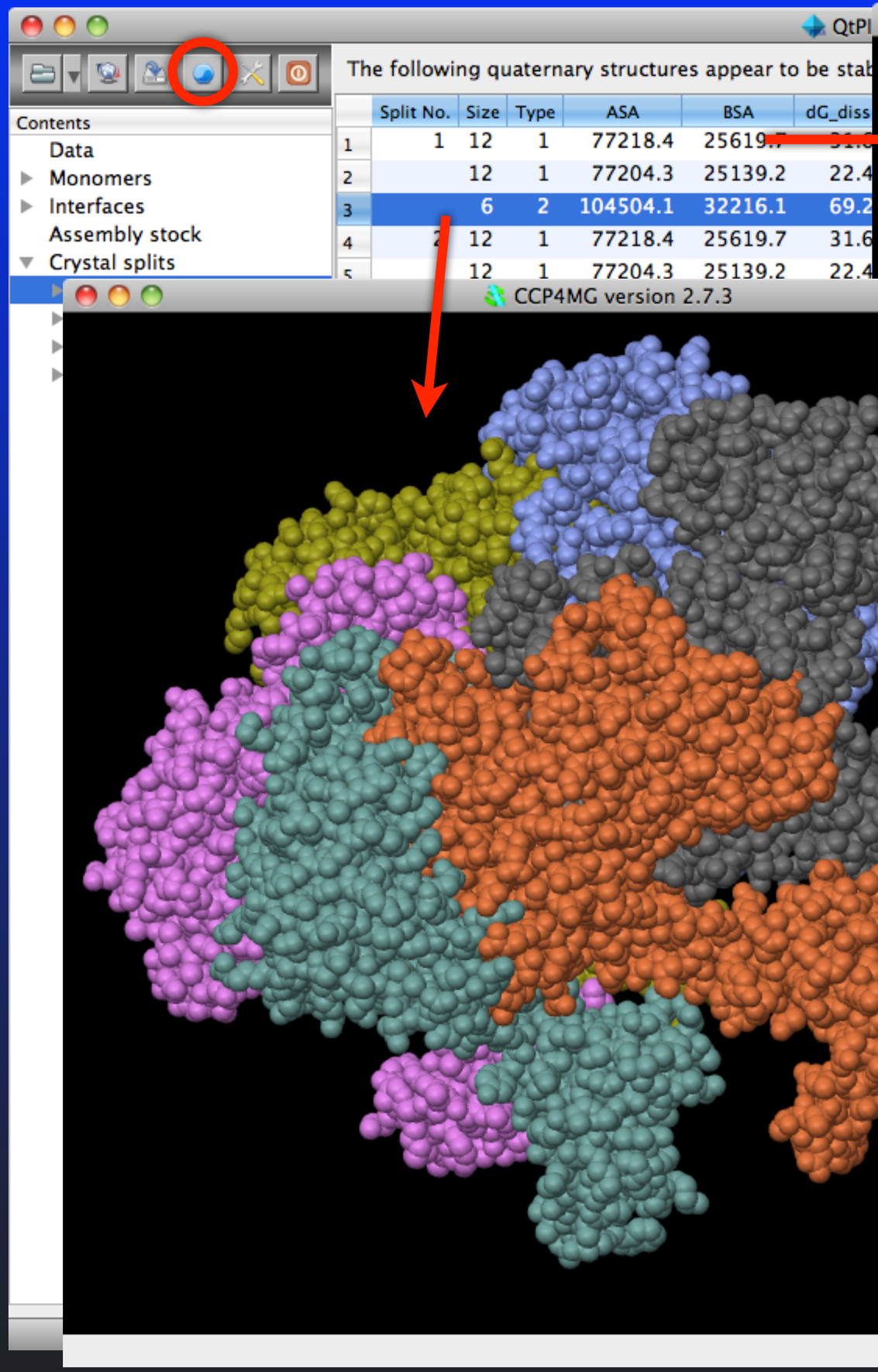
$$\Delta G_0 = -\Delta G_{\text{int}} - T\Delta S \quad > \quad 0$$

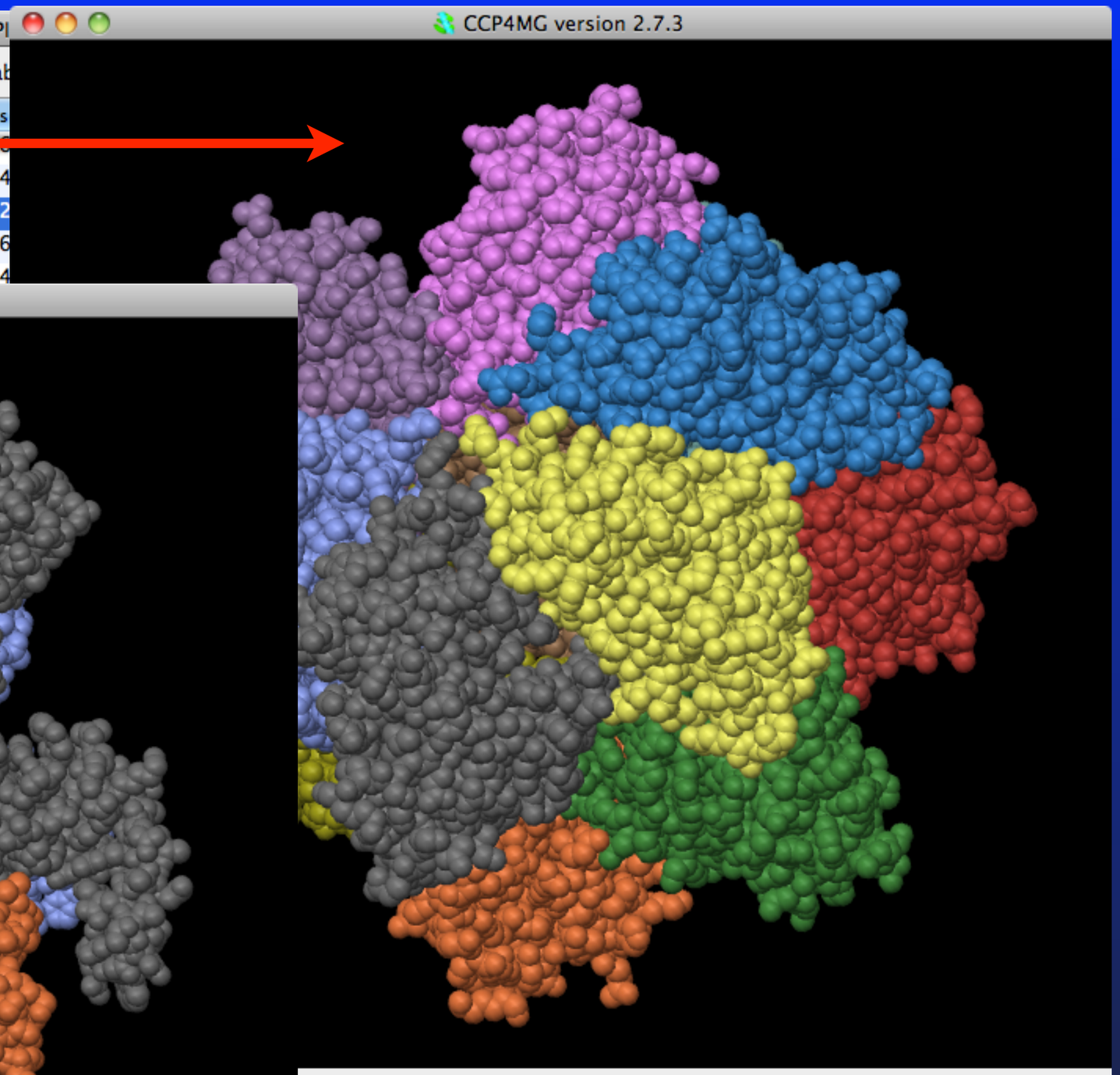
3. Leave only sets of stable assemblies in the list, and range them by chances to be a biological unit:
    - Larger assemblies take preference
    - Single-assembly sets take preference
    - Otherwise, assemblies with higher ΔG_0 take preference



E. Krissinel and K. Henrick (2007) J. Mol. Biol. 372, 774-797


Contents
Data
Monomers
Interfaces
Assembly stock
Crystal splits
Stable splits
Metastable splits
Unstable splits
Unsplitted
File: /Users/Eugen
Title: HSLV-HSLU FR
Space group: P 63 2 2
Resolution: 2.8
Cell: 172.022 172
Cell volume [A^3]: 7.088e+06
ASU (File) contents:
Protein chains: 6 (6)
DNA/RNA chains: 0 (0)
Ligands: 2 (2)
NCS-mates: 0
Excluded ligands: None
CCP4MG version 2.7.3


QtPI
The following quaternary structures appear to be stab
Split No. Size Type ASA BSA dG_diss
1 1 12 1 77218.4 25619.7
2 12 1 77204.3 25139.2 22.4
3 6 2 104504.1 32216.1 69.2
4 12 1 77218.4 25619.7 31.6
5 12 1 77204.3 25139.2 22.4
Contents
Data
Monomers
Interfaces
Assembly stock
Crystal splits
CCP4MG version 2.7.3


CCP4MG version 2.7.3

jsPISA 2.0.3 [PDB 1e94]

www.ccp4.ac.uk/pisa/sessions/LQ-821-SI/index.html

CCP4 on-line

jsPISA 2.0.3 [PDB 1e94]

Print Refresh Help Exit

Input | Monomers | Interfaces | Stock | Crystal Splits | Log file

Interfaces

- List of interfaces
  - 1. F || E
    - Bonds
    - Residues
  - 2. E || F
  - 3. E || E
  - 4. A || B
  - 5. B || A
  - 6. D || C
  - 7. D || C
  - 8. B || A
  - 9. D || C
  - 10. C || C
  - 11. A || A
  - 12. B || B
  - 13. D || D
  - 14. [ANP]E:500 || E
  - 15. [ANP]F:501 || F
  - 16. F || B
  - 17. E || A
  - 18. C || F
  - 19. E || A
  - 20. [ANP]E:500 || F
  - 21. E || [ANP]F:501
  - 22. F || E
  - 23. E || E
  - 24. F || F

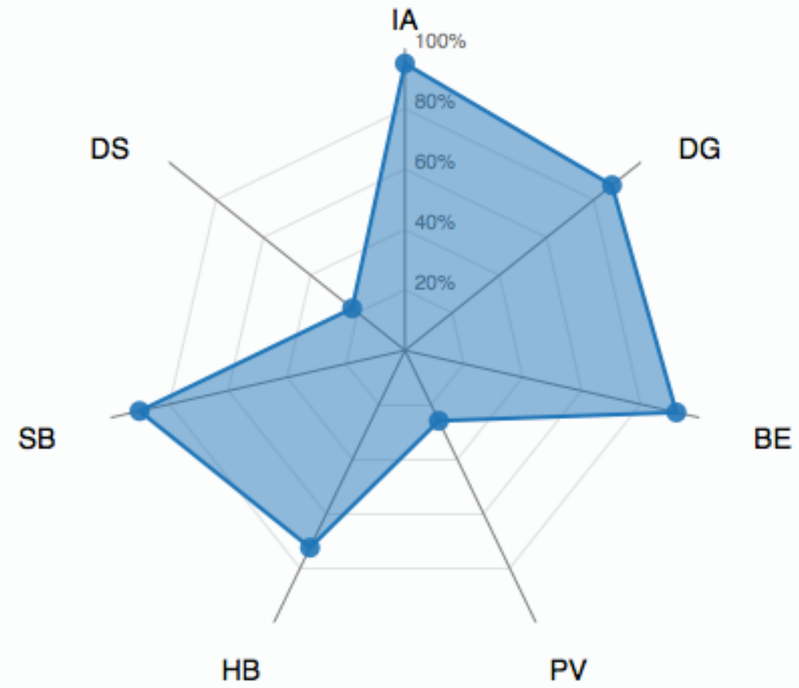
Interface F || E

Summary

	Monomer 1		Monomer 2	
Monomer ID	F		E	
Class	Protein		Protein	
Symmetry operation	X,Y,Z		X,Y,Z	
Symmetry ID	1_555		1_555	
Interface atoms	232	7.3%	221	6.9%
Surface atoms	1925	60.5%	1921	60.4%
Total atoms	3184	100.0%	3183	100.0%
Interface residues	66	16.1%	65	15.9%
Surface residues	380	92.9%	385	94.4%
Total residues	409	100.0%	408	100.0%
BSA, Å²	2146.6	9.6%	2345.5	10.6%
ASA, Å²	22253.4	100.0%	22072.6	100.0%
Solvation energy, kcal/mol	-331.9		-334.3	
SE gain, kcal/mol	-7.5		-5.0	

Download PDB | Download PDBx | View in JSMol

Interaction radar

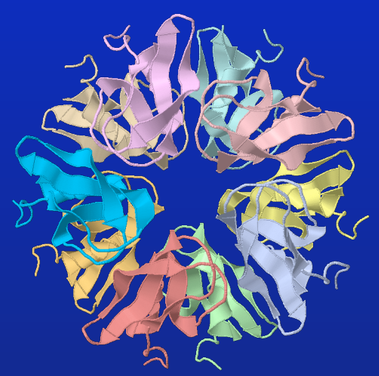


Interface parameters

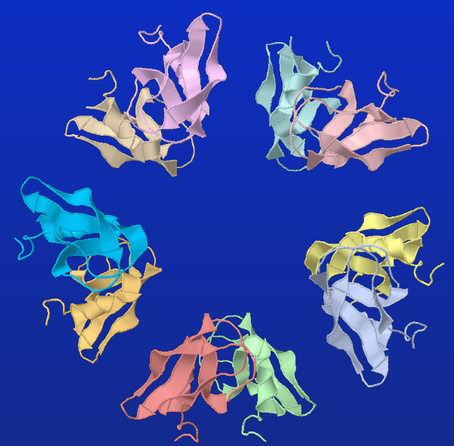
IA	:	Interface area, Å²	2246
DG	:	Solvation Energy, kcal/mol	-12.46
BE	:	Total Binding Energy, kcal/mol	-26.65
PV	:	Hydrophobic P-value	0.3542
HB	:	Number of Hydrogen Bonds	11
SB	:	Number of Salt Bridges	25
DS	:	Number of Disulphide Bonds	0

What is “A Stable Complex?”

★ Chemical systems always move towards equilibrium:



$$K_d = \frac{\prod_i^n [A_i]}{[A]}$$



★ PISA reports the Free Gibbs Energy, $\Delta G_0 = -RT \log K_d$, how to interpret?

✦ *In general, if the equilibrium is shifted to the left ($K_d < 1$), the complex is stable.*

- But does this always mean that stable complex has higher concentration than the dissociates? - no it does not
- And this depends on the concentration anyway? - yes it does
- And it also depends on the dissociation pattern (dissociation into monomers, dimers, trimers etc.)? How to identify the pattern?

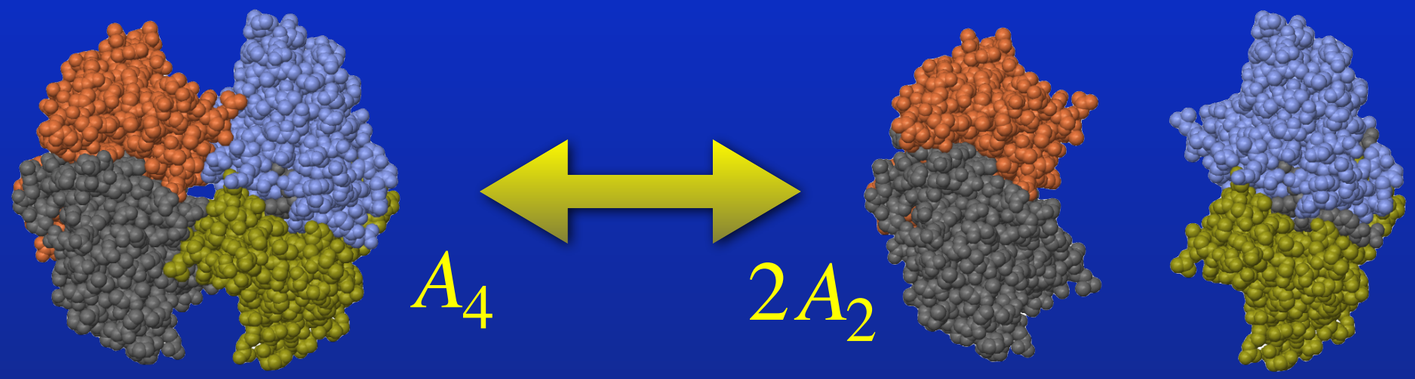
- the pattern is, in essence, the minimum free energy route of dissociation
- is the minimum free energy route always unique?
- does it not depend on concentration (temperature, pH, etc.), too?



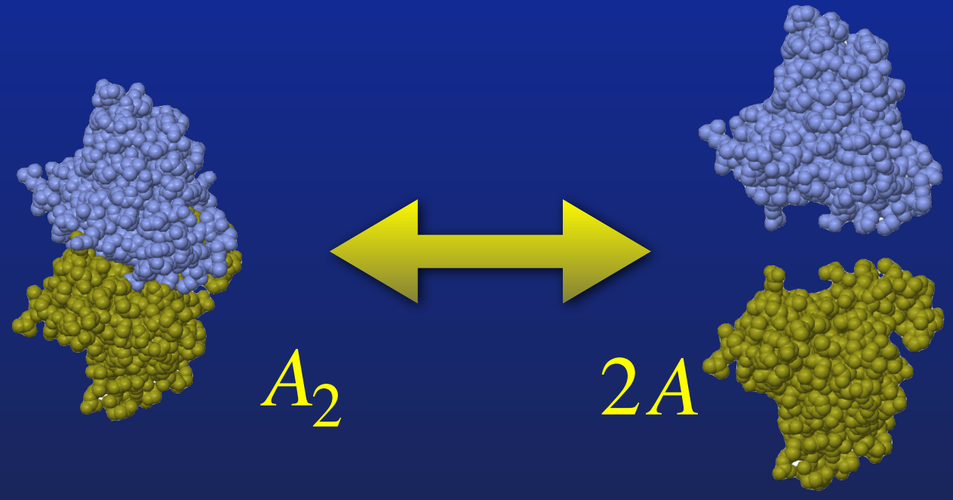
Research Complex at Harwell

Is ΔG_0 Sufficient An Indicator?

★ Consider PDB entry 3LT5:



$\Delta G_0 = 3 \ \ kcal / M$



$\Delta G_0 = 10 \ \ kcal / M$

The tetramer is weaker than the dimer, so one may think that the structure is dimeric

But the tetramer is equilibrated with the dimer, so that their concentrations can be comparable

What is the correct answer?



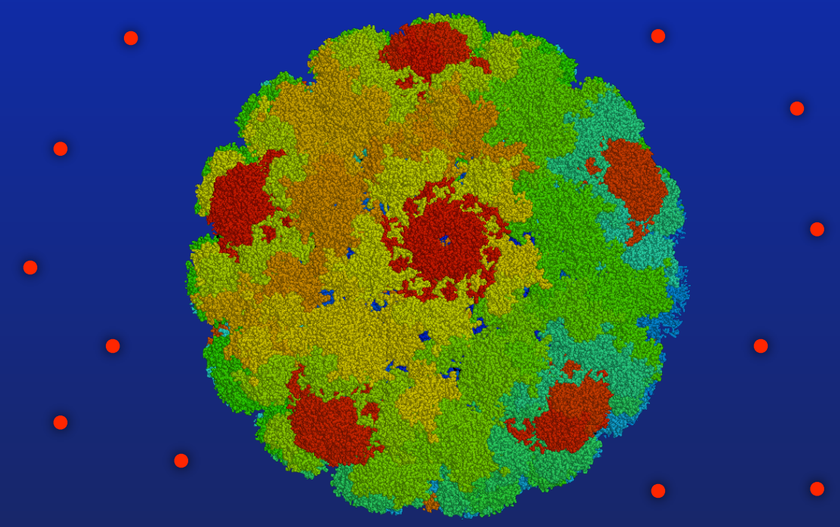
The Stock

- All possible complexes co-exist in dynamic equilibrium and form a "stock"
  - *PISA's Stock is limited to complexes formed by crystal interfaces*
- Their stock concentrations do vary
- Concentrations depend on free energy of dissociation *and stock composition*
- Concentration-based analysis is not very indicative:

for the equilibrium between large complex and its monomeric units on the right,

$$[A_{360}] \ll [A]$$

from which one could conclude that the complex is unstable; but obviously, the protein is highly aggregated



- Aggregated states are better indicated by the aggregation index:

$$A_i = \frac{m_i}{\sum_j m_j}$$

m_i mass of ith species in the Stock

$$0 < A_i < 1$$

fully dissolved *fully aggregated*



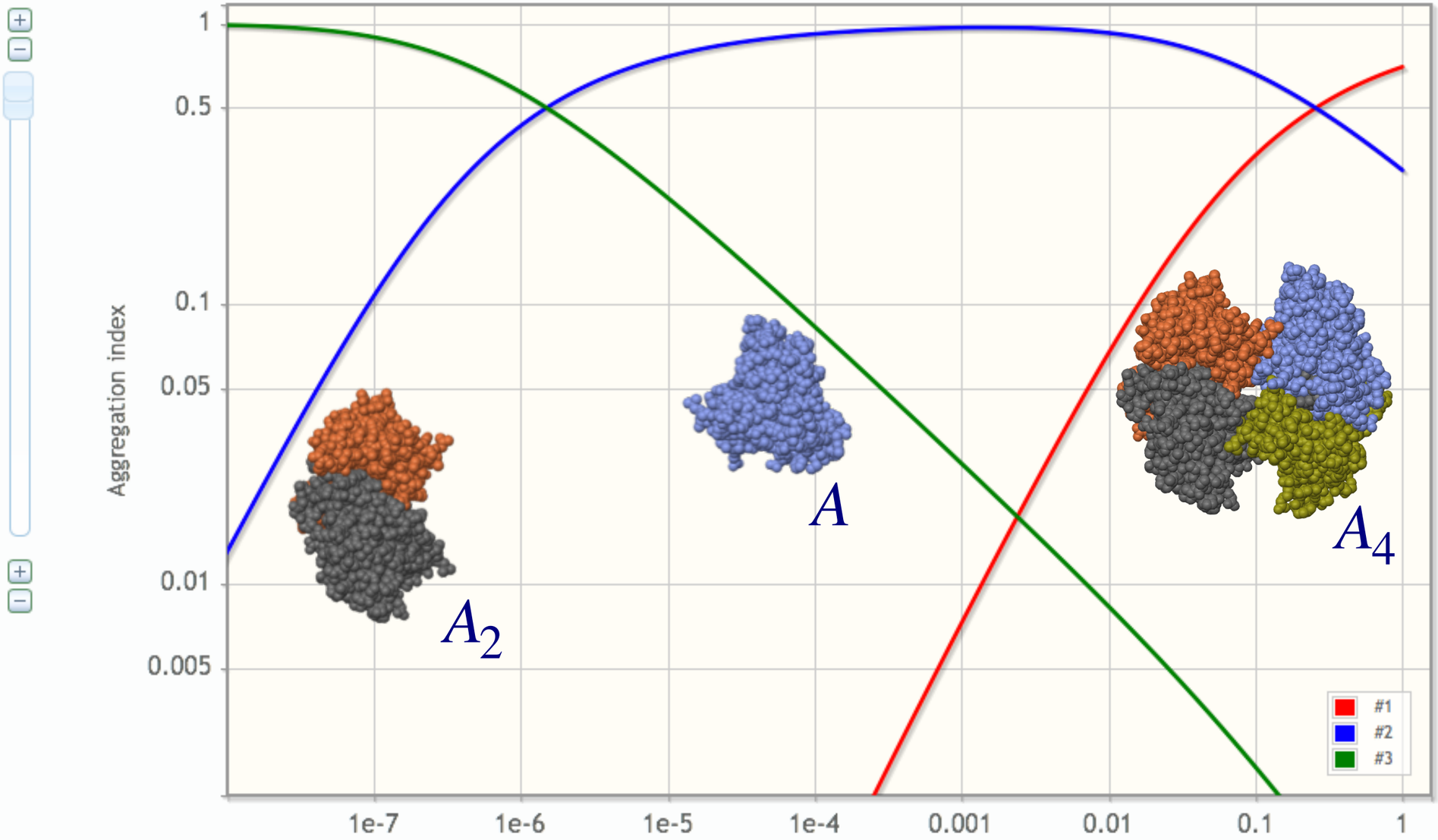
Research Complex at Harwell

CCP4 on-line

jsPISA 2.0.3 [PDB 3lt5]

Print Refresh Help Exit

Input | Monomers | Interfaces | Stock | Crystal Splits | Log file



aggregation index

Crystal dissolution

Stock components

##	Plot	Size	Type	ASA	BSA	dG_diss	dG0	Formula	Composition
1	☑	4	1	32286.1	12422.8	3.3	23.4	A_4a_4	$A_2B_2[FMN]_4$
2	☑	2	2	17385.7	4968.7	10.1	10.1	A_2a_2	$AB[FMN]_2$
3	☑	1	3	10722.4	726.1	0.0	0.0	Aa	A[FMN]

Stock-Based Analysis Makes a Difference

Examlpe of PDB entry **3IMP**

- *"standard" PISA analysis suggests that the structure can be dodecameric*
- *dodecamers are weak, but still marginally stronger than the tetramers*
- *not completely clear which dodecamer to choose*

CCP4 on-line

jsPISA 2.0.3 [PDB 3imp]

Print Refresh Help Exit

Input | Monomers | Interfaces | Stock | Crystal Splits | Log file

Crystal Splits

- ▼ Stable splits
  - 1-1. A(12)
  - 2-2. A(12)
  - 3-3. A(4)
  - 3-4. A(4)
  - 3-5. A(4)
- ▶ Metastable splits
- Unstable splits
- ▶ No-crystal analysis

Crystal splits in Stable Assemblies

The following quaternary structures appear to be stable in solution

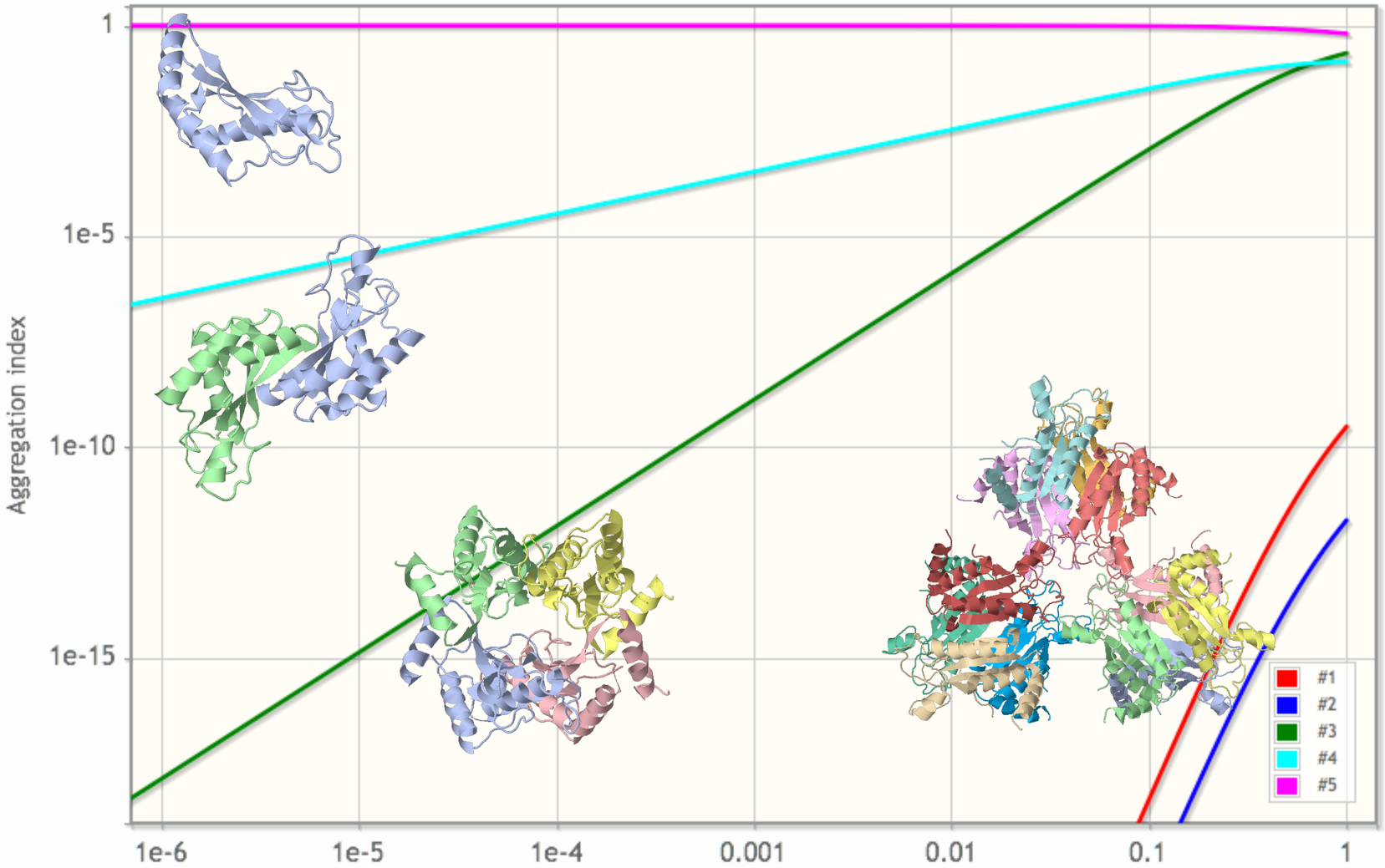
##	Split No.	Size	Type	ASA	BSA	dG_diss	dG0	Formula	Composition
1	1	12	1	69248.0	21061.4	9.9	9.9	A_{12}	ABCDEFGHIJKL
2	2	12	2	69695.5	20613.9	6.9	6.9	A_{12}	ABCDEFGHIJKL
3	3	4	3	24107.1	5972.3	5.9	5.9	A_4	AFGH
4		4	3	24124.9	5941.7	4.5	6.8	A_4	IJKL
5		4	3	24243.1	5920.2	3.9	3.9	A_4	BCDE

CCP4 on-line

jsPISA 2.0.3 [PDB 3imp]

Print Refresh Help Exit

Input | Monomers | Interfaces | Stock | Crystal Splits | Log file



aggregation index

Stock components

##	Plot	Size	Type	ASA	BSA	dG_diss	dG0	Formula	Composition
1	☑	12	1	69248.0	21061.4	9.9	9.9	A_{12}	ABCDEFGHIJKL
2	☑	12	2	69695.5	20613.9	6.9	6.9	A_{12}	ABCDEFGHIJKL
3	☑	4	3	24179.6	5938.6	4.6	5.1	A_4	AFGH
4	☑	2	4	13426.6	1648.3	0.9	0.9	A_2	AH
5	☑	1	5	7525.8	0.0	0.0	0.0	A	A

★ However, in the stock, concentration of dodecamers appears to be negligible comparing to that of lower-multiplicity complexes

- *Reason: given complexes dissociate to the ground state (monomers). Their equilibrium concentrations are:*

$$C_{A_{12}} \sim K_{eq}(C_A)^{12}$$

$$C_{A_4} \sim K_{eq}(C_A)^{4}$$

therefore, at similar equilibrium constants,

$$C_{A_{12}} \ll C_{A_4}$$

★ Stock analysis results:

- *primarily monomeric*
- *co-existence of dimers and tetramers at high concentrations*

How does One Infer on the Dissociation Pattern?

- ★ Calculated by PISA automatically as the most optimal dissociation pathway
- ★ Presented in PISA detailed reports on individual assemblies

CCP4 on-line

jsPISA 2.0.3 [PDB 3imp]

Print Refresh Help Exit

Input | Monomers | Interfaces | Stock | **Crystal Splits** | Log file

Crystal Splits

- ▼ Stable splits
  - 1-1. A(12)
  - 2-2. A(12)
  - 3-3. A(4)
  - 3-4. A(4)
  - 3-5. A(4)
- ▶ Metastable splits
- Unstable splits
- ▶ No-crystal analysis

Split 2, Assembly 2: ABCDEFGHIJKL

Summary

Multimeric state	12	*Surface area, Å²*	69695.5	*Free energy of dissociation, kcal/mol*	6.9
Copies in Unit Cell	2	*Buried area, Å²*	20613.9	*Entropy of dissociation, kcal/mol*	138.0
Symmetry number	2				
Composition	ABCDEFGHIJKL				
Formula	A_{12}				
Dissociation pattern	A + H + F + G + B + E + C + D + I + L + J + K				

Download PDB | Download PDBx | View in JSMol | View dissociated

Engaged Interfaces

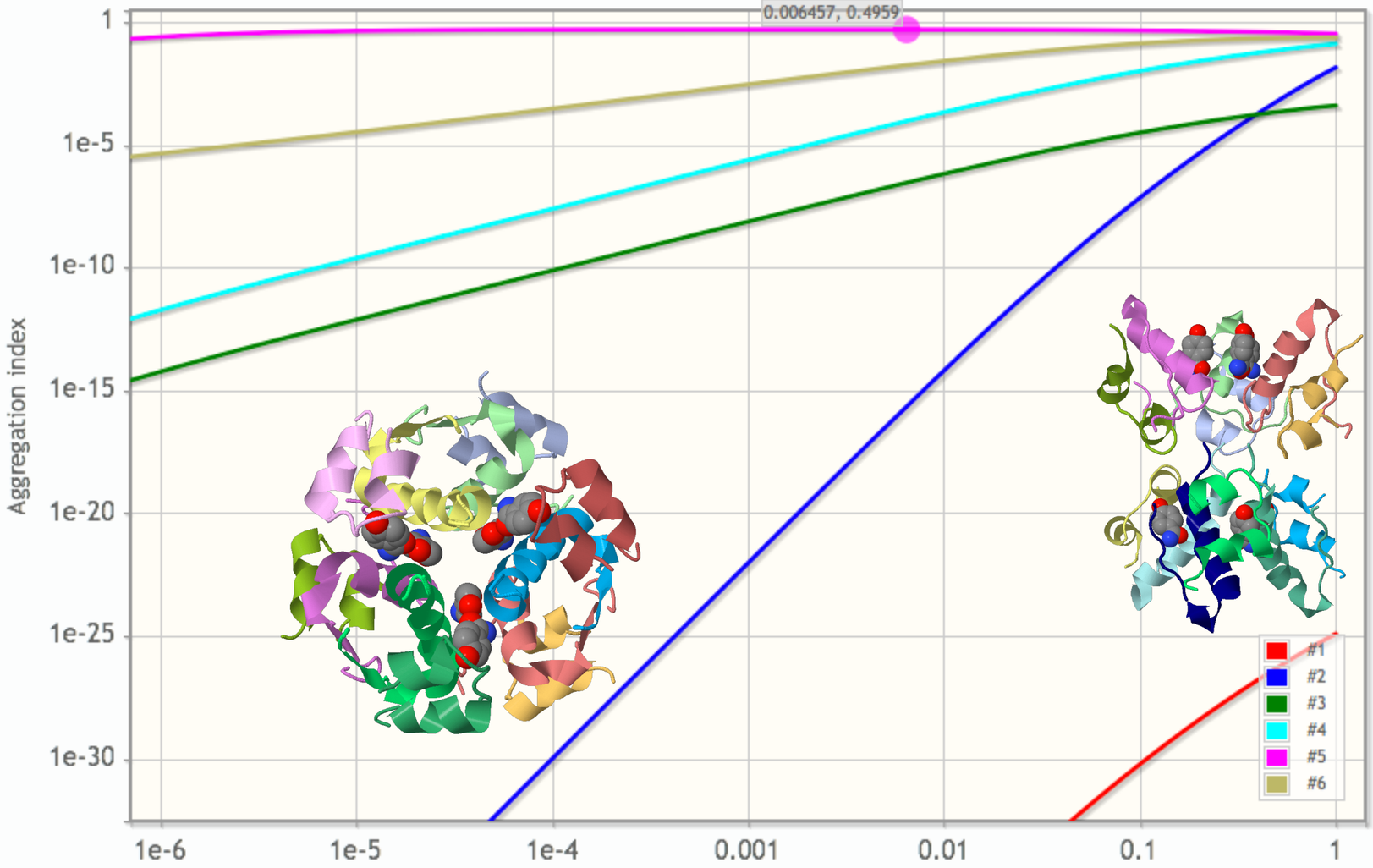
##	Type	Monomer 1	Monomer 2	Occ.	Symmetry operation	Sym. ID	Area	Delta G	P-value	N_{hb}	N_{sb}	N_{ds}
1	1	G	F	1*	X,Y,Z	1_555	865.2	-5.7	0.390	11	3	0
2	1	D	C	1*	X,Y,Z	1_555	819.5	-6.9	0.319	9	3	0
3	1	E	B	1*	X,Y,Z	1_555	816.3	-7.6	0.300	9	3	0
4	1	K	J	1*	X,Y,Z	1_555	812.8	-8.1	0.244	10	2	0

CCP4 on-line

jsPISA 2.0.3 [PDB 1ben]

Print Refresh Help Exit

Input | Monomers | Interfaces | Stock | Crystal Splits | Log file



aggregation index

Stock components

##	Plot	Size	Type	ASA	BSA	dG_diss	dG0	Formula	Composition
1	☑	12	1	18849.1	13007.7	3.8	11.7	$A_6B_3C_3b_6$	$A_3C_3B_3D_3[HBD]_6$
2	☑	12	2	12160.3	19696.5	3.6	43.4	$A_6B_3C_3b_6$	$A_3C_3B_3D_3[HBD]_6$
3	☑	4	3	6672.2	3946.7	7.2	9.8	A_2BCb_2	$ACBD[HBD]_2$
4	☑	4	4	5901.7	4717.3	2.0	13.3	A_2BCb_2	$ACBD[HBD]_2$
5	☑	2	5	3677.3	1669.7	9.6	9.6	ABb	AB[HBD]
6	☑	2	6	3932.6	1339.4	2.6	2.6	ACb	CD[HBD]

An Example of the Opposite

- Dodecamers in insulin **1BEN** appear to have similar dissociation free energies but drastically different stock concentration profiles
  - *Reason: dodecamer #1 dissociates into monomers, while dodecamer #2 dissociates into tetramers*

- Stock analysis results:
  - *primarily dimeric*
  - *co-existence of dimers, tetramers and dodecamer #2 at high concentrations*

Classification of Protein Assemblies

Assembly classification on the benchmark set of 218 protein structures published in

Ponstingl, H., Kabir, T. and Thornton, J. (2003) Automatic inference of protein quaternary structures from crystals. J. Appl. Cryst. 36, 1116-1122.

	1mer	2mer	3mer	4mer	6mer	Other	Sum	Correct
1mer	49	3	0	1	1	1	55	89%
2mer	3	71+11	0	2+1	0	0	76+12	93%
3mer	1	0	22	0	1	0	24	92%
4mer	2	2+1	0	26+6	0	1	31+7	84%
6mer	0	0	0	1	10+2	0	10+3	92%
	196+22 ⇔ 196 homomers and 22 heteromers					Total:	196+22	**90%**

Classification error in ΔG_0 : ± 5 kcal/mol



Research Complex at Harwell

Classification of Protein-DNA Complexes

Assembly classification on the benchmark set of 212 protein-DNA complexes published in

Luscombe, N.M., Austin, S.E., Berman, H.M. and Thornton, J. (2000) An overview of the structures of protein-DNA complexes. Genome Biol. 1, 1-37.

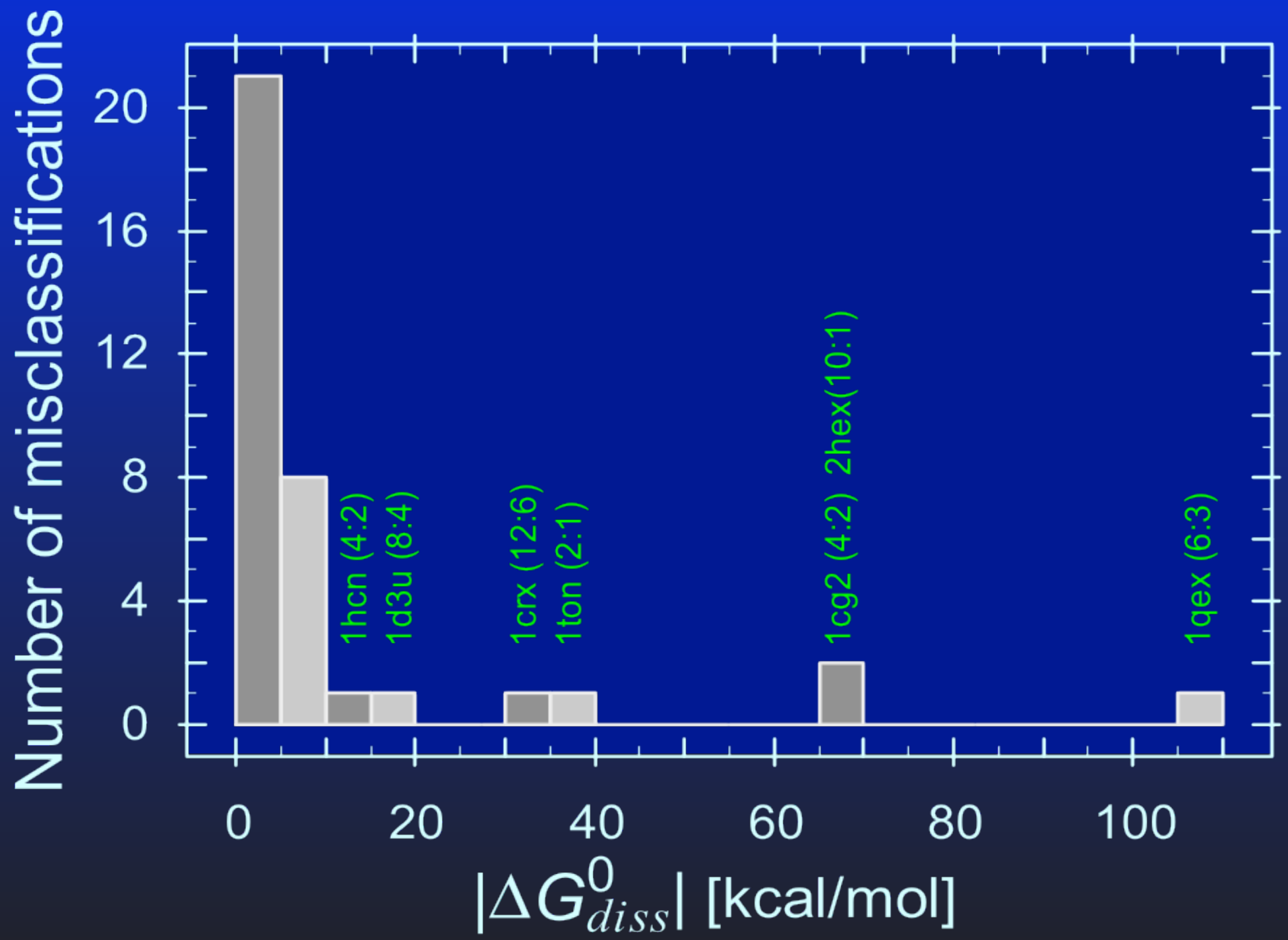
	2mer	3mer	4mer	5mer	6mer	10mer	Other	Sum	Correct
2mer	1	0	0	0	0	0	0	1	100%
3mer	6	96	0	0	1	0	2	105	91%
4mer	0	2	83	0	0	0	0	85	98%
5mer	0	0	2	3	0	0	0	5	60%
6mer	1	0	0	0	13	0	1	15	87%
10mer	0	0	0	0	0	1	0	1	100%
							Total:	212	**93%**

Classification error in ΔG_0 : ± 5 kcal/mol



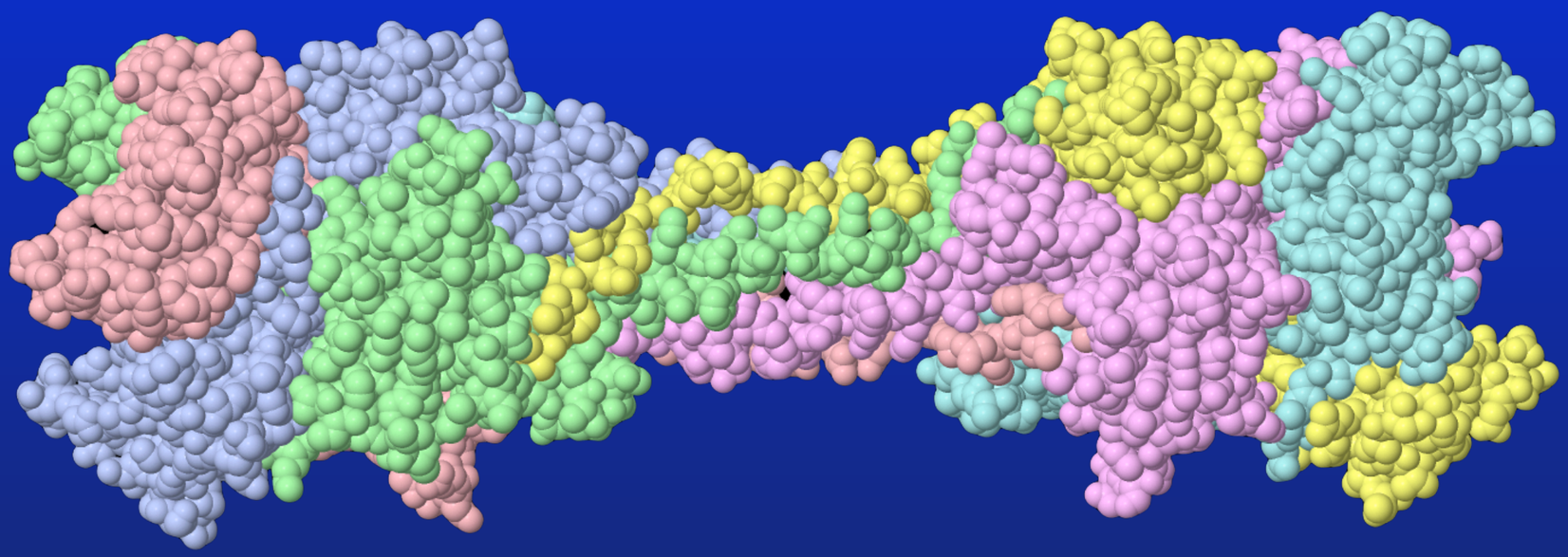
Research Complex at Harwell

Free Energy Distribution of Misclassifications



Example of misclassification: 1QEX

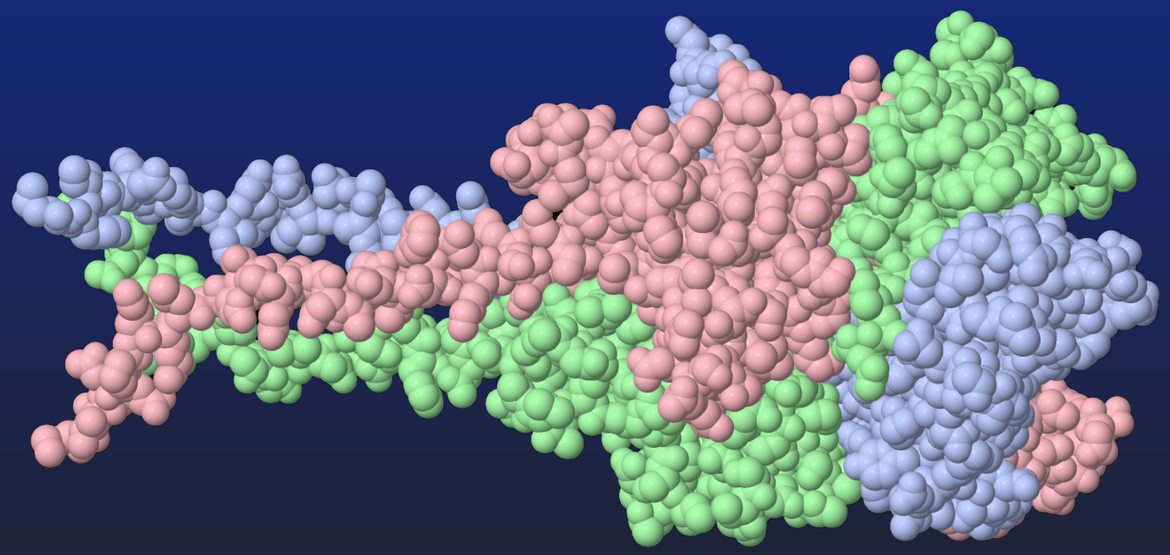
BACTERIOPHAGE T4 GENE PRODUCT 9 (GP9), THE TRIGGER OF TAIL CONTRACTION AND THE LONG TAIL FIBERS CONNECTOR



Predicted: homohexamer

Dissociates into 2 trimers

$\Delta G_0 \simeq 106$ kcal/mol



Biological unit: homotrimer

Dissociates into 3 monomers

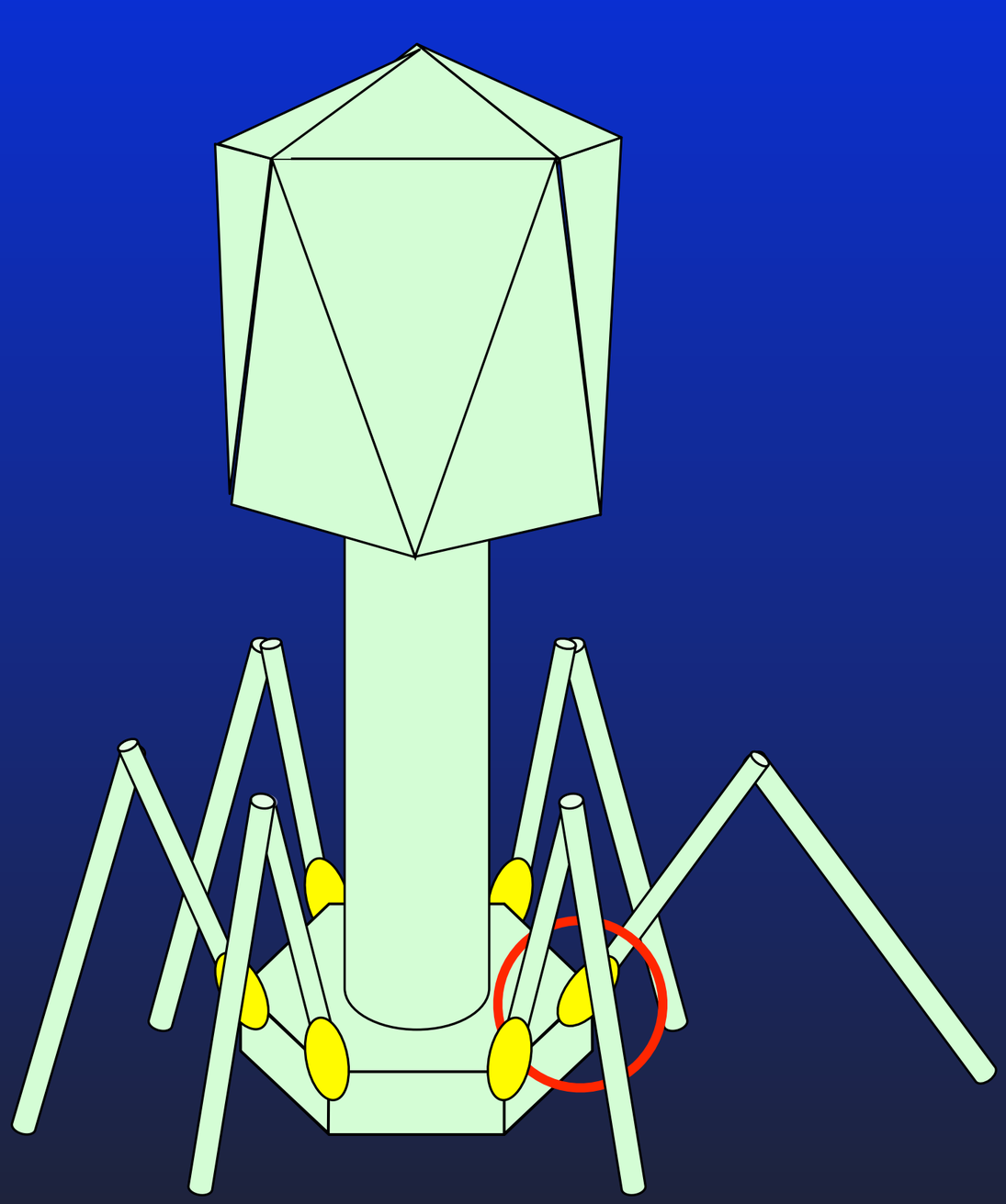
$\Delta G_0 \simeq 90$ kcal/mol

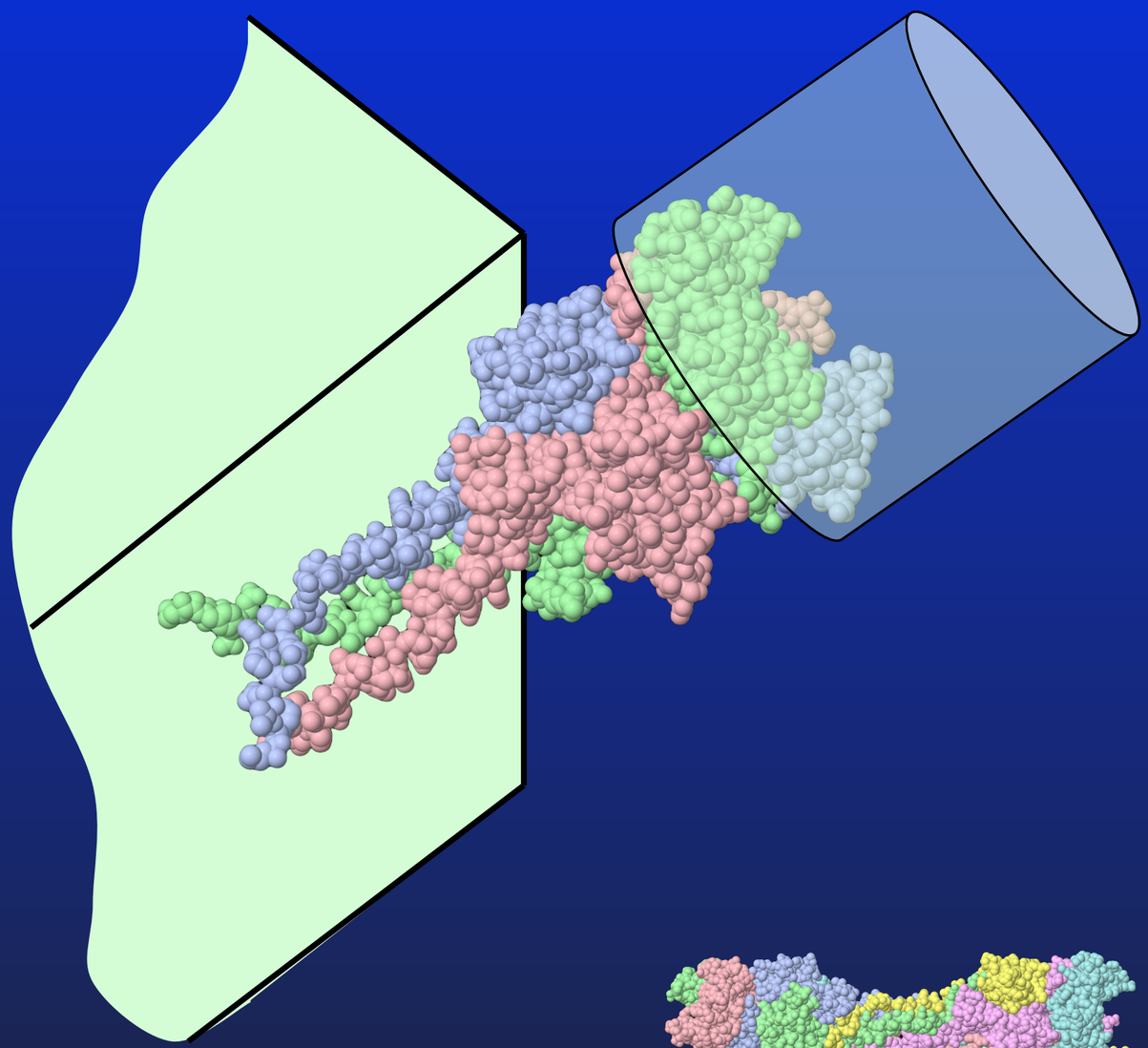


Research Complex at Harwell

Example of misclassification: 1QEX

BACTERIOPHAGE T4 GENE PRODUCT 9 (GP9), THE TRIGGER OF TAIL CONTRACTION AND THE LONG TAIL FIBERS CONNECTOR

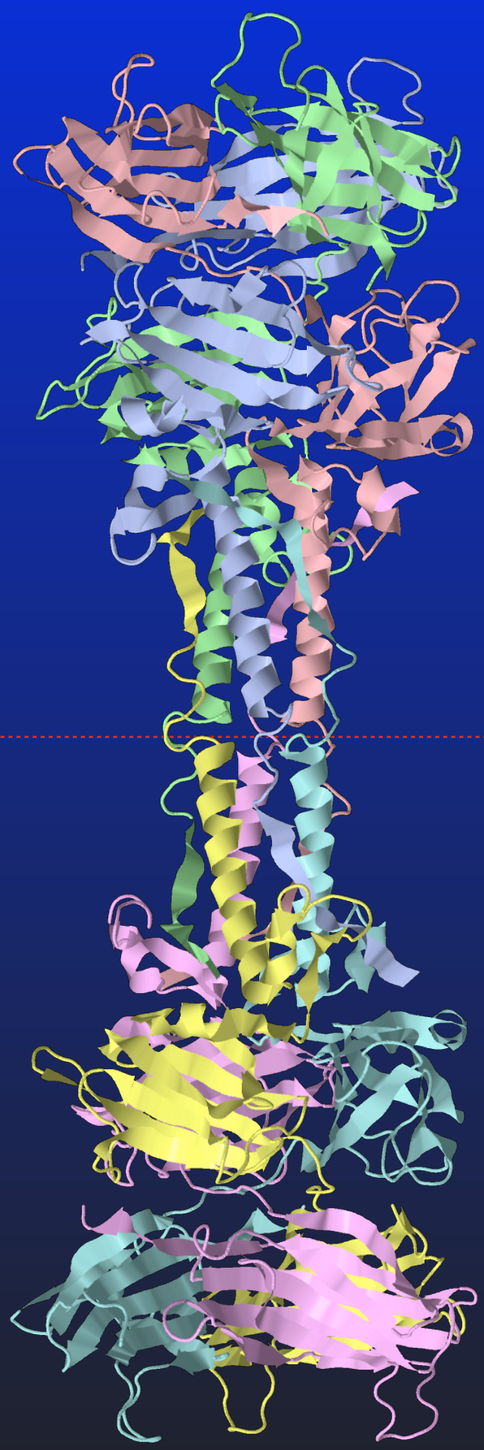




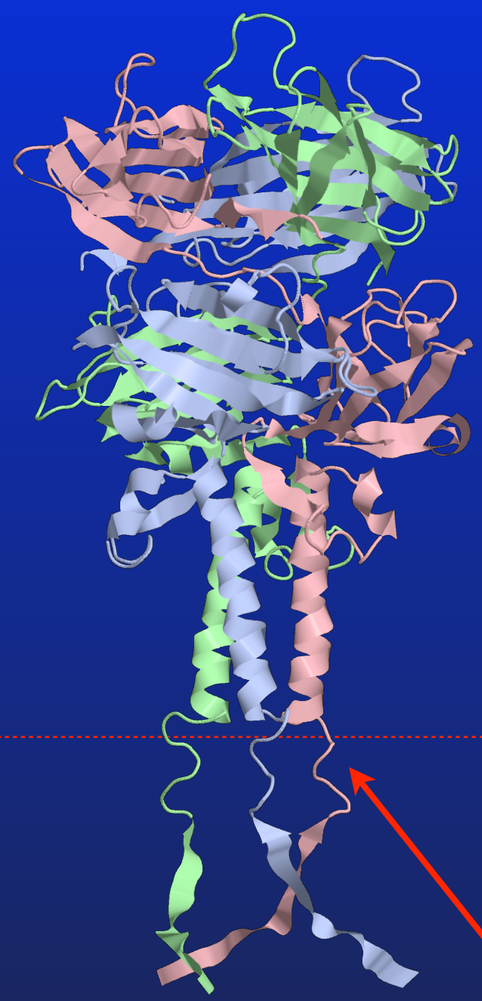
Rossmann M.G., Mesyanzhinov V.V., Arisaka F and Leiman P.G. (2004) *The bacteriophage T4 DNA injection machine*. Curr. Opinion Struct. Biol. 14:171-180.

Example of misclassification: 1QEX

BACTERIOPHAGE T4 GENE PRODUCT 9 (GP9), THE TRIGGER OF TAIL CONTRACTION AND THE LONG TAIL FIBERS CONNECTOR

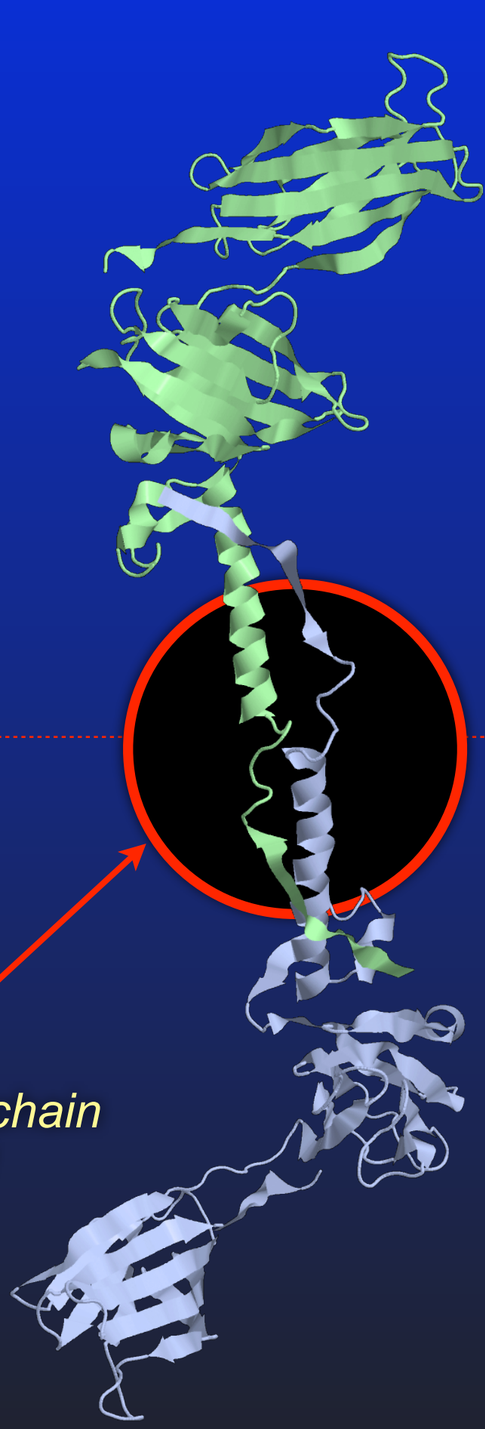


1QEX hexamer

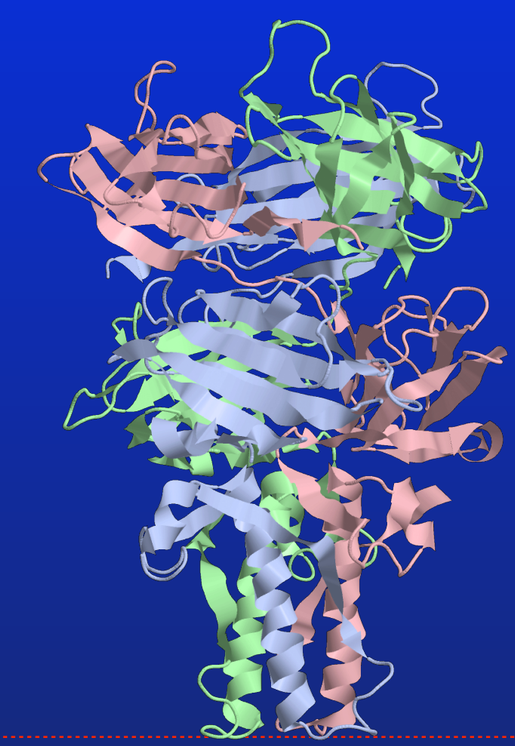


1QEX trimer

Wrong mainchain tracing!



1QEX asu

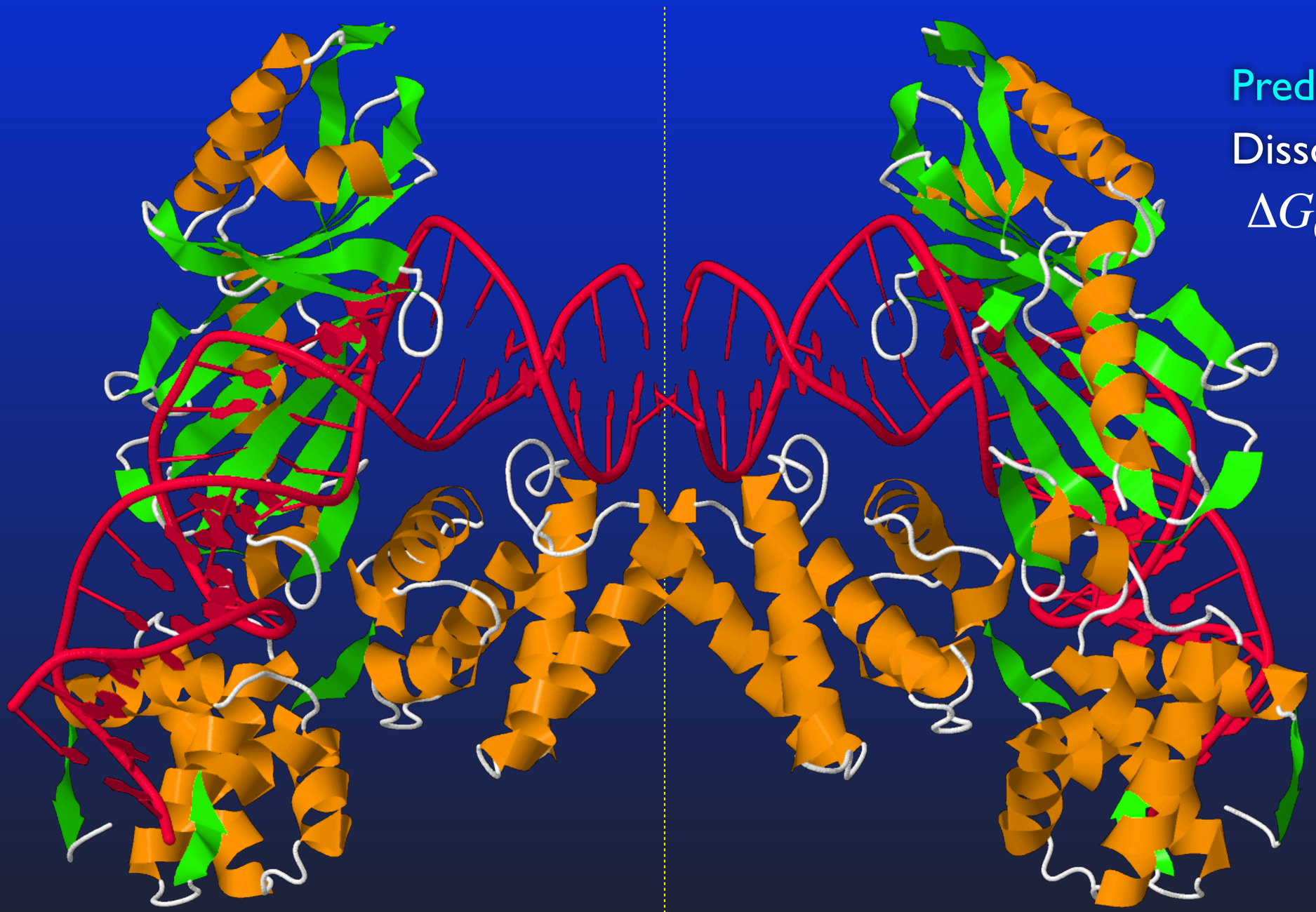


1S2E trimer
Correct mainchain tracing
Identified correctly



Example of misclassification: 1D3U

TATA-BINDING PROTEIN / TRANSCRIPTION FACTOR



Predicted: octamer

Dissociates into 2 tetramers

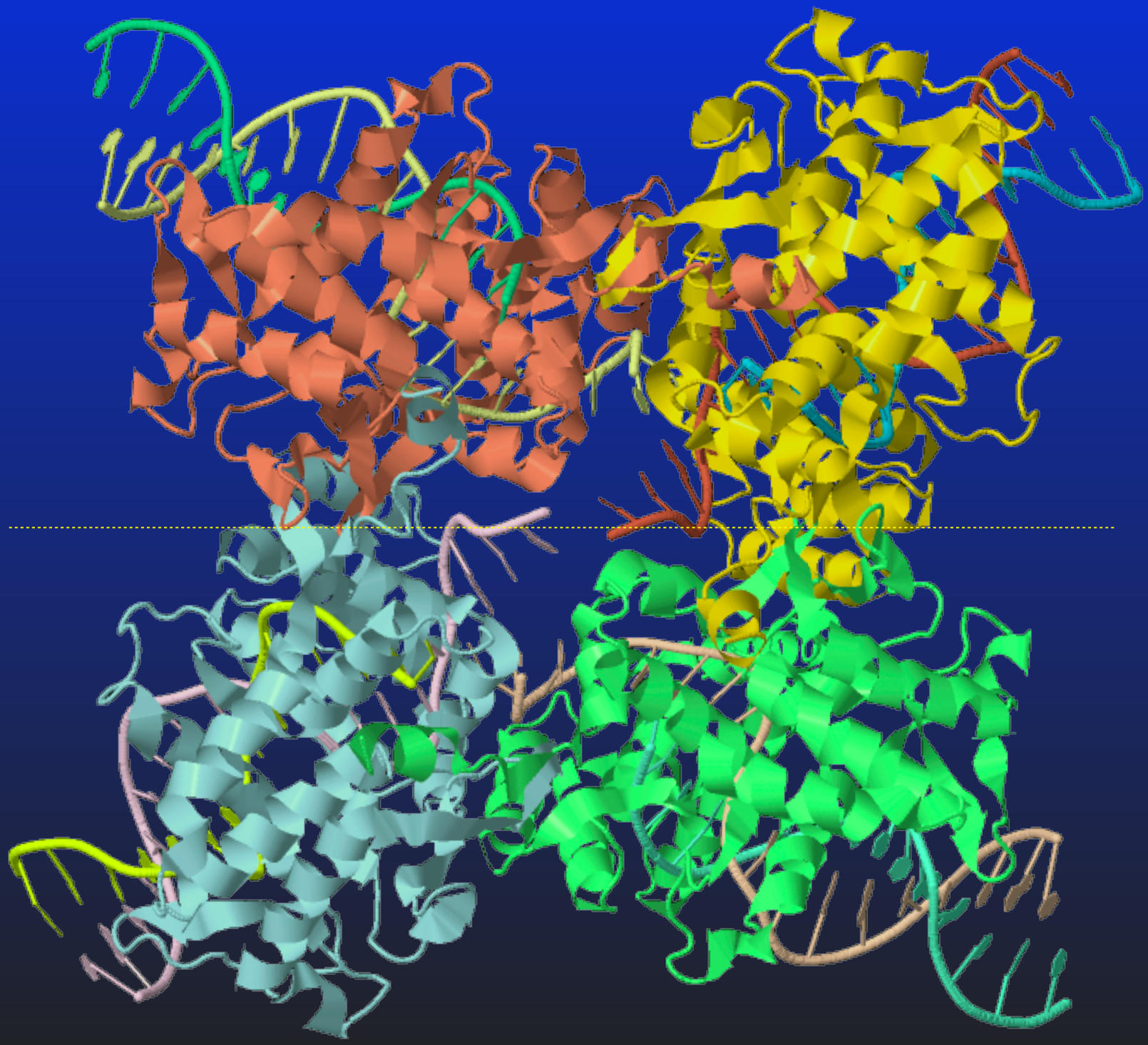
$\Delta G_0 \simeq 20$ kcal/mol

Functional unit:
tetramer



Example of misclassification: 1CRX

CRE RECOMBINASE / DNA COMPLEX REACTION INTERMEDIATE



Predicted: dodecamer

Dissociates into 2 hexamers

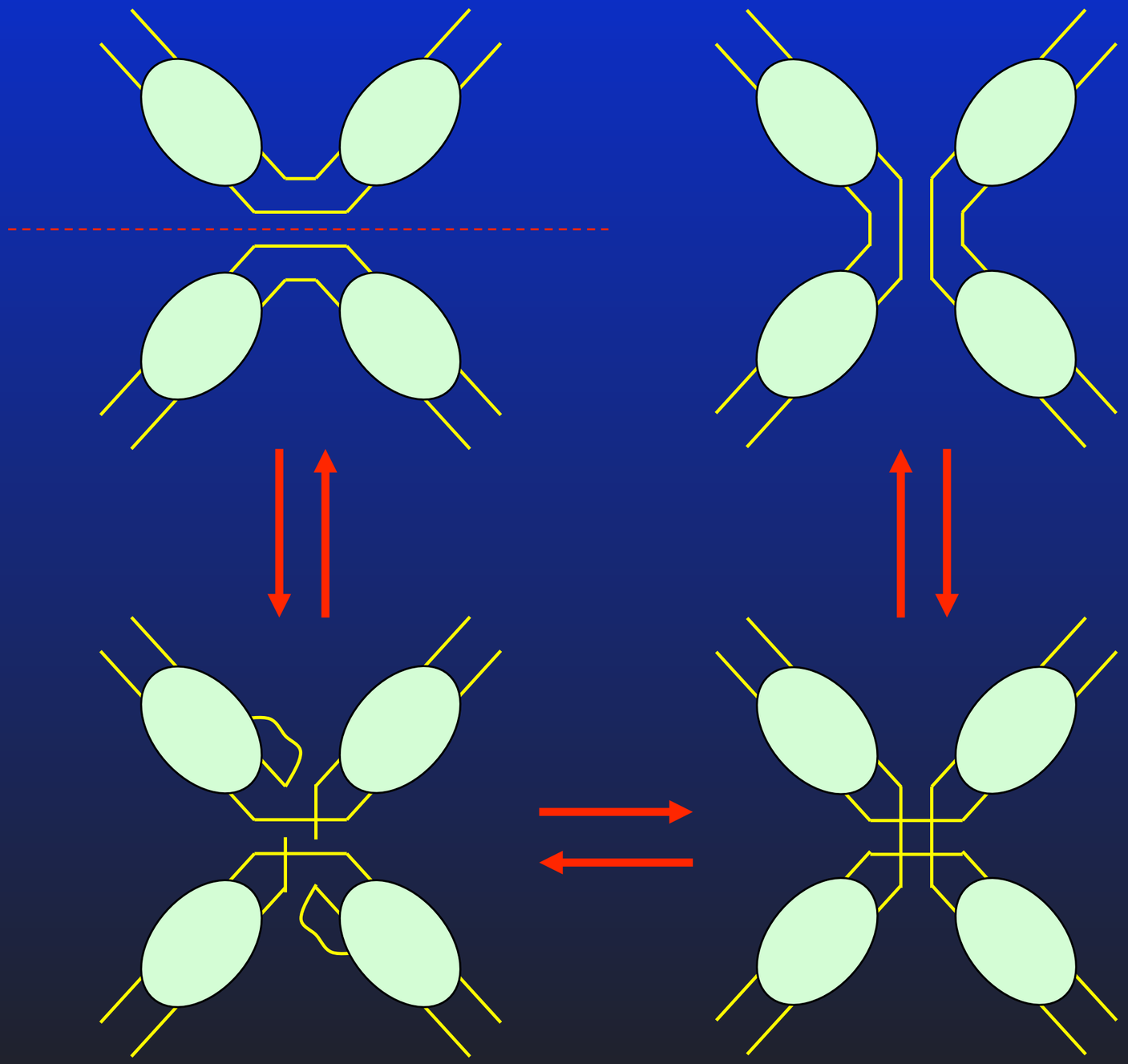
$\Delta G_0 \simeq 28$ kcal/mol

Functional unit: trimer



Example of misclassification: 1CRX

CRE RECOMBINASE / DNA COMPLEX REACTION INTERMEDIATE



Guo F., Gopaul D.N. and van Duyne G.D. (1997)

Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse.

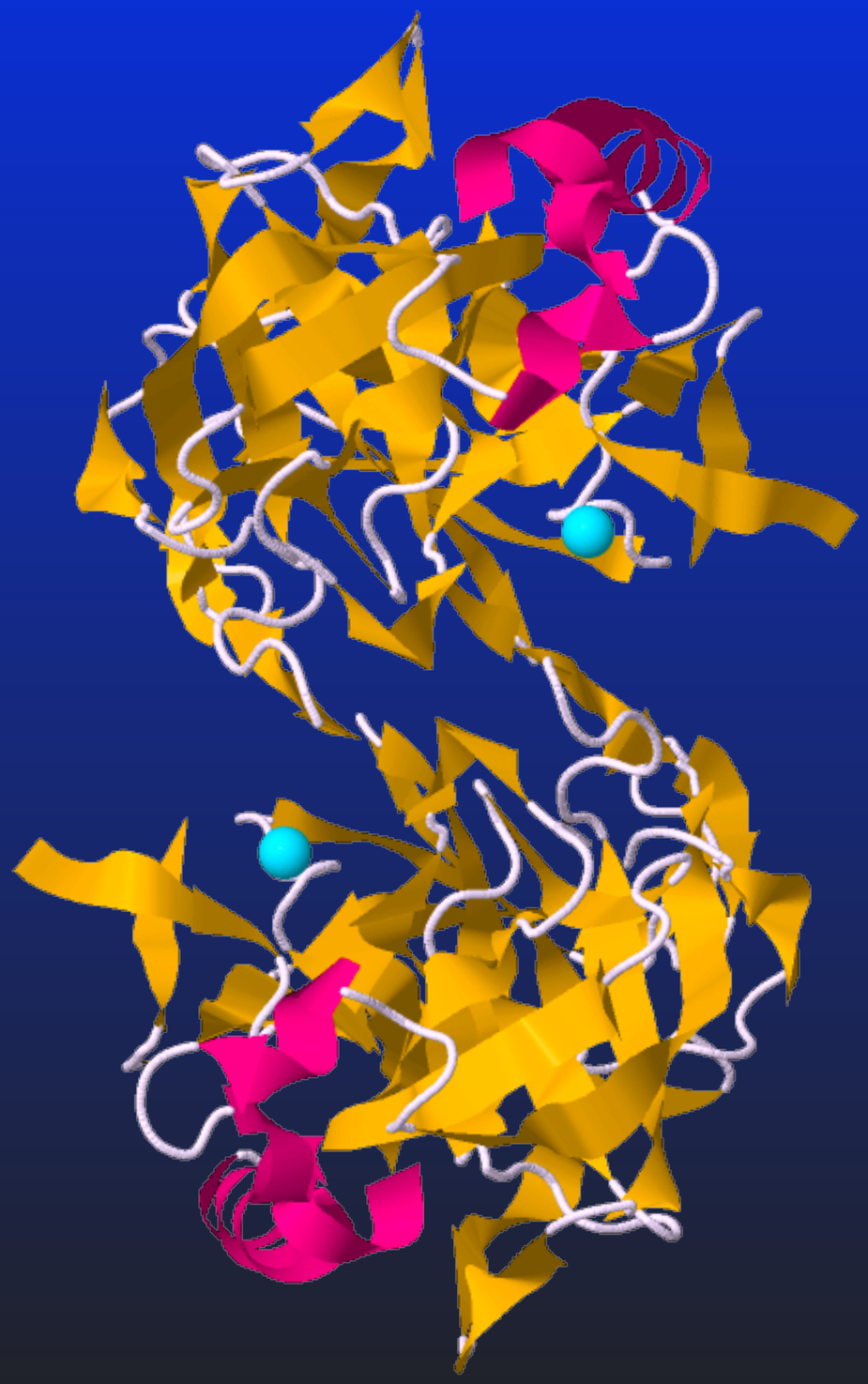
Nature 389:40-46.



Research Complex at Harwell

Example of misclassification: 1TON

TONIN



Predicted: dimer

Dissociates at

$\Delta G_0 \simeq 37$ kcal/mol

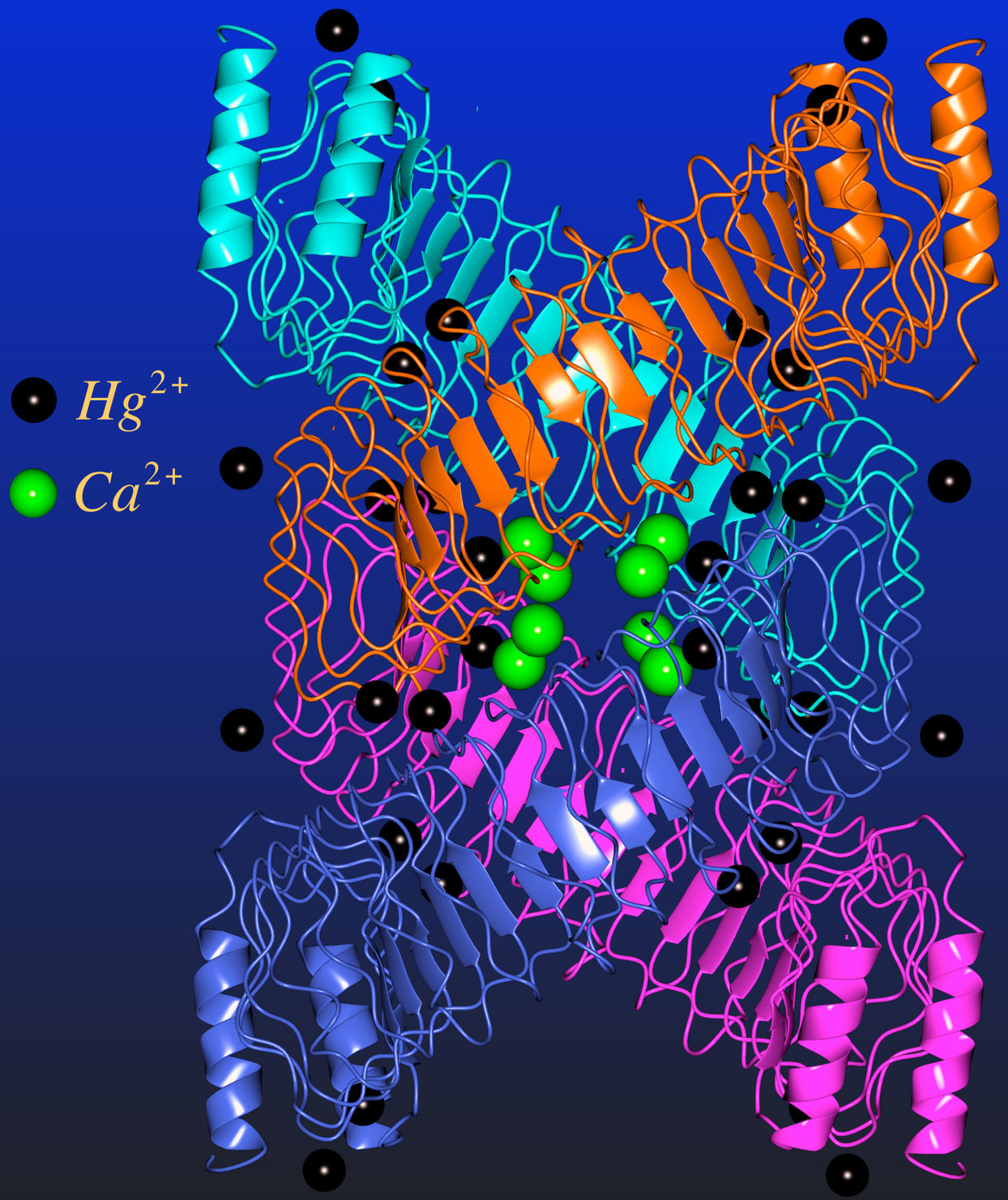
Biological unit: monomer

Apparent dimerization is an artefact due to the presence of Zn^{+2} ions added to the buffer to aid crystallization. Removal of Zn from the file results in $\Delta G_0 \simeq 3$ kcal/mol

Fujinaga M., James M.N.G. (1997) *Rat submaxillary gland serine protease, tonin structure solution and refinement at 1.8 Å resolution.* J.Mol.Biol. 195:373-396.

Example of ion effect: 1G9U vs 1JL5

Y. PESTIS CYTOXIN YopM



Predicted: homotetramer in form of a superhelix featuring a hollow cylinder with an inner diameter of ~35 Å.

	1G9U	1JL5
Space Group	$P4_222$	$I4_122$
ΔG_0, kcal/mol	37	3
Number of ions	40	16

Biological unit: monomer

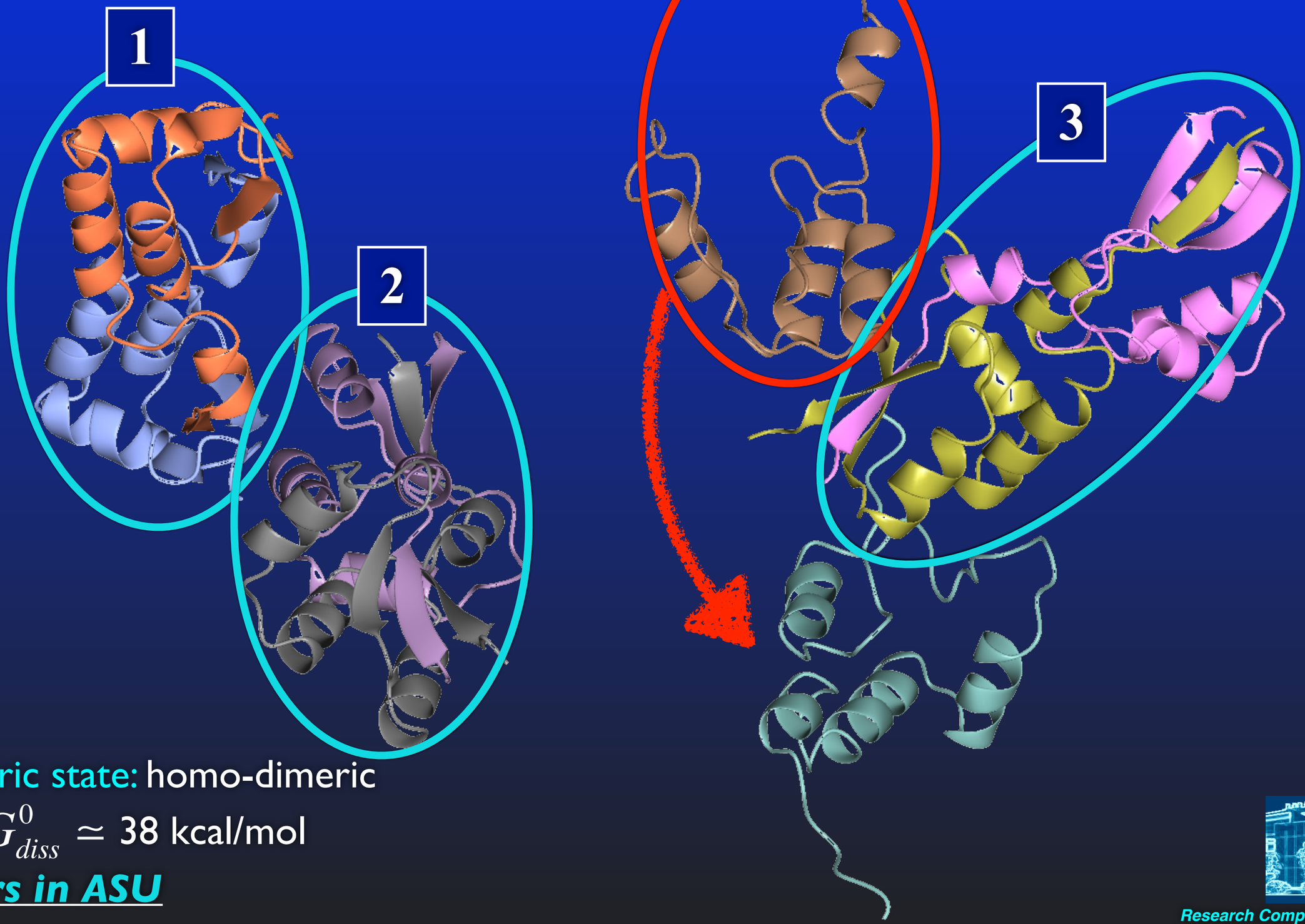
Evdokimov, A. G., Anderson, D. E., Routzahn, K. M. & Waugh, D. S. (2001). J. Mol. Biol. 312, 807–821

Removal of ions makes the structure monomeric in PISA estimates



Research Complex at Harwell

Choice of the Asymmetric Unit: 4BJQ



Oligomeric state: homo-dimeric

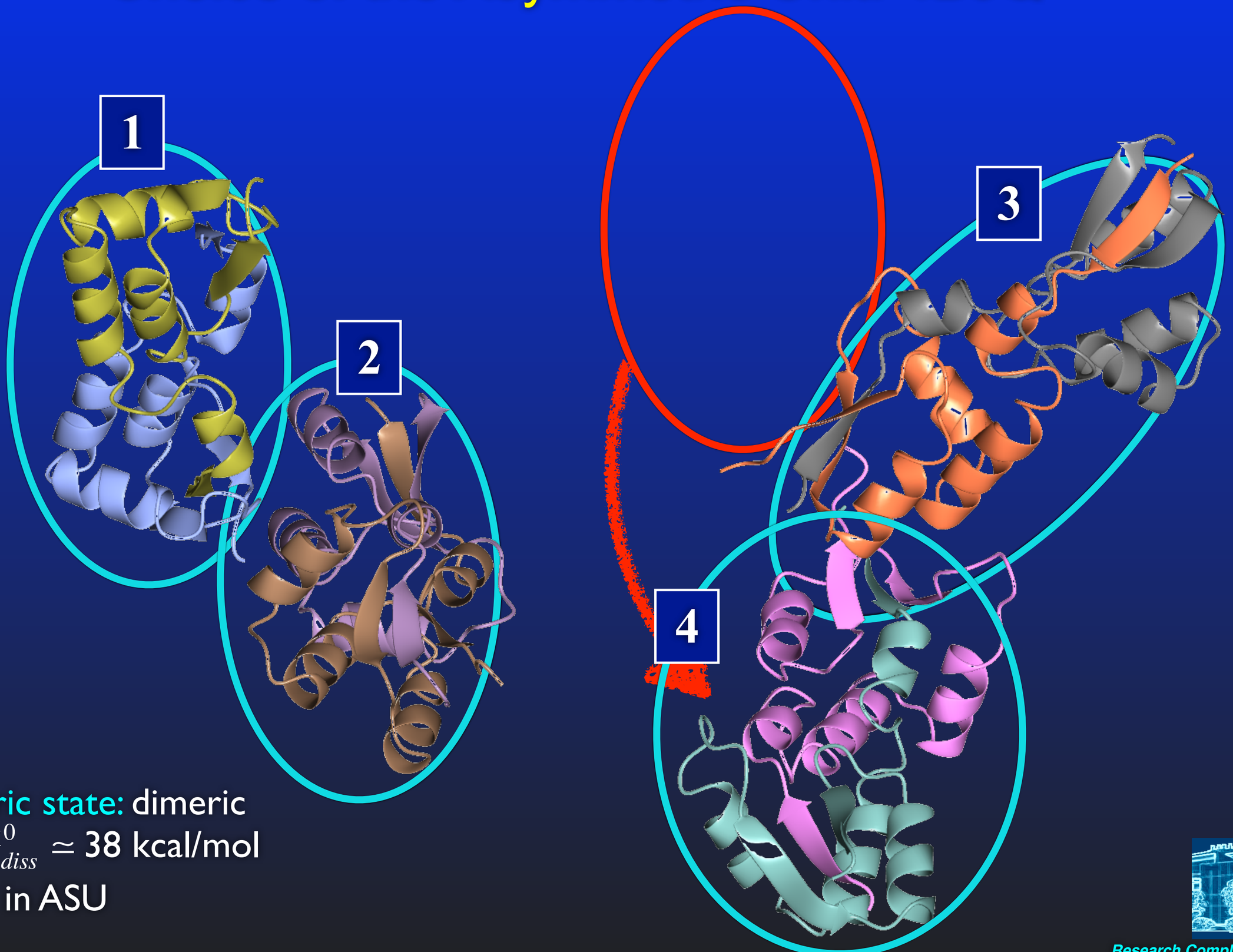
with $\Delta G^{0}_{diss} \simeq$ 38 kcal/mol

4 dimers in ASU



Research Complex at Harwell

Choice of the Asymmetric Unit: 4BJQ



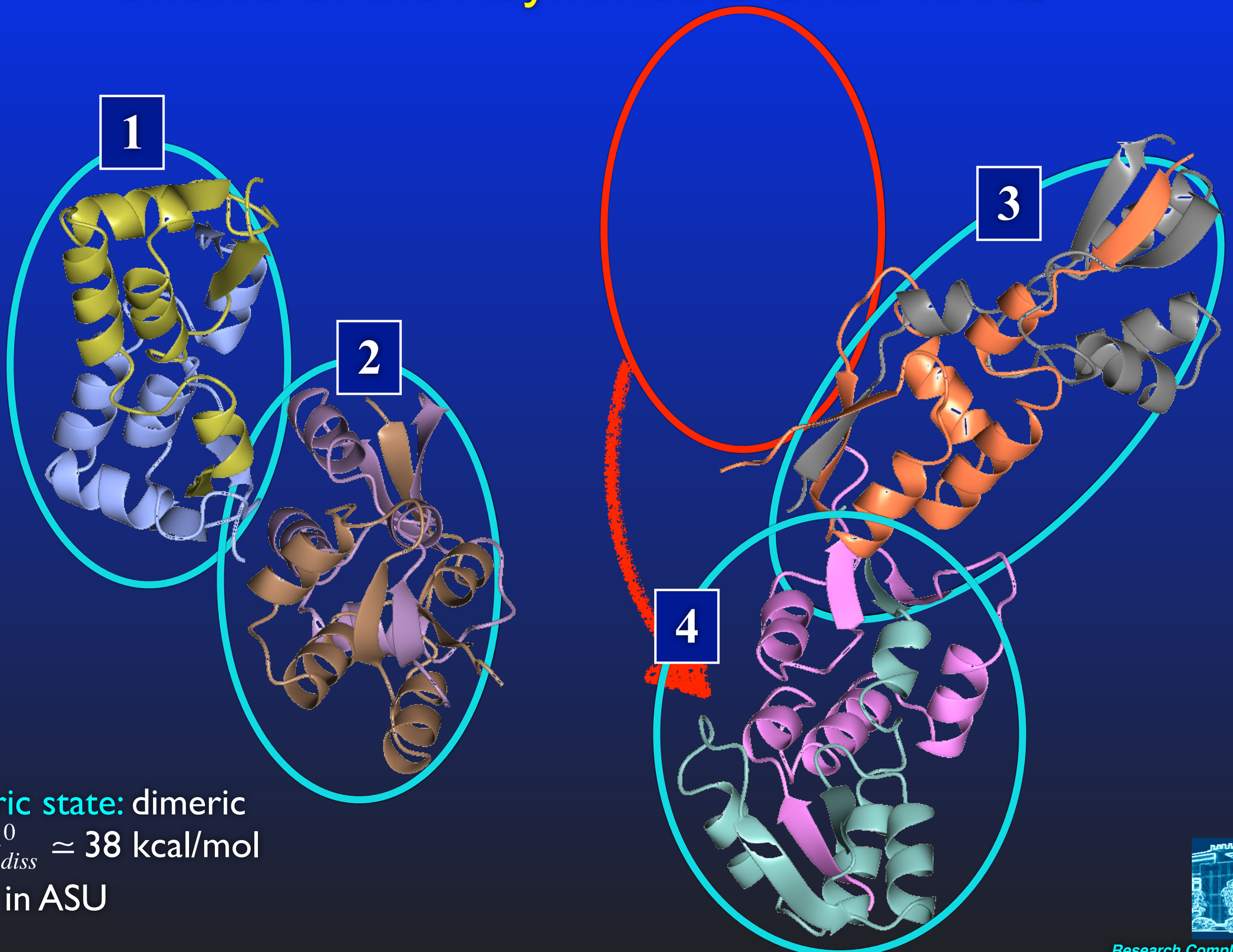
Oligomeric state: dimeric
with $\Delta G^0_{diss} \simeq 38$ kcal/mol

4 dimers in ASU

Choice of the Asymmetric Unit



Choice of the Asymmetric Unit: 4BJQ



Oligomeric state: dimeric
with $\Delta G^0_{diss} \simeq 38$ kcal/mol

4 dimers in ASU

Does it really work?

- PISA appears to work quite well, which seems to be a "problem"
  - 90% success rate achieved on the benchmark set
  - in 2007, wwPDB adopted PISA as a mandatory processing tool for all depositions
  - since that, feedback from wwPDB curators suggests that up to 95% of classifications made by PISA agree with experimental data on oligomeric state, where available, and with intuitive and common-sense considerations where experimental evidence is not given

- Why it might work well? Two reasons:

Energy models and calculations are quite accurate	Obviously wrong
PISA relies on geometry of interactions given by crystal packing. PISA does not dock monomeric units; rather, it uses crystal contacts as "nature's dockings" assuming that they are correct.	Probably correct



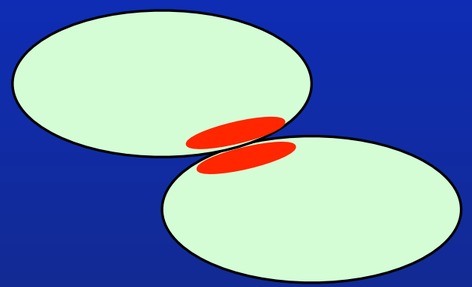
Research Complex at Harwell

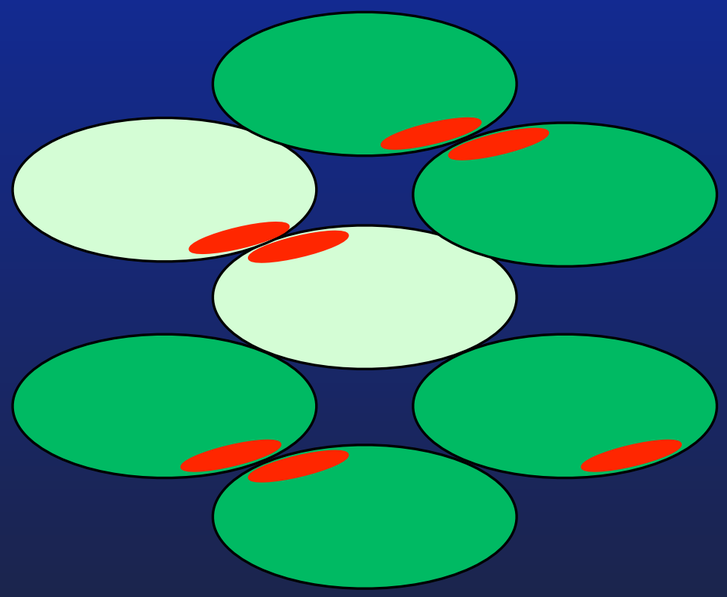
Distortions and Re-assembly

- Crystal optimizes energy globally, therefore it may sacrifice biologically relevant interaction in favour of unspecific crystal contacts

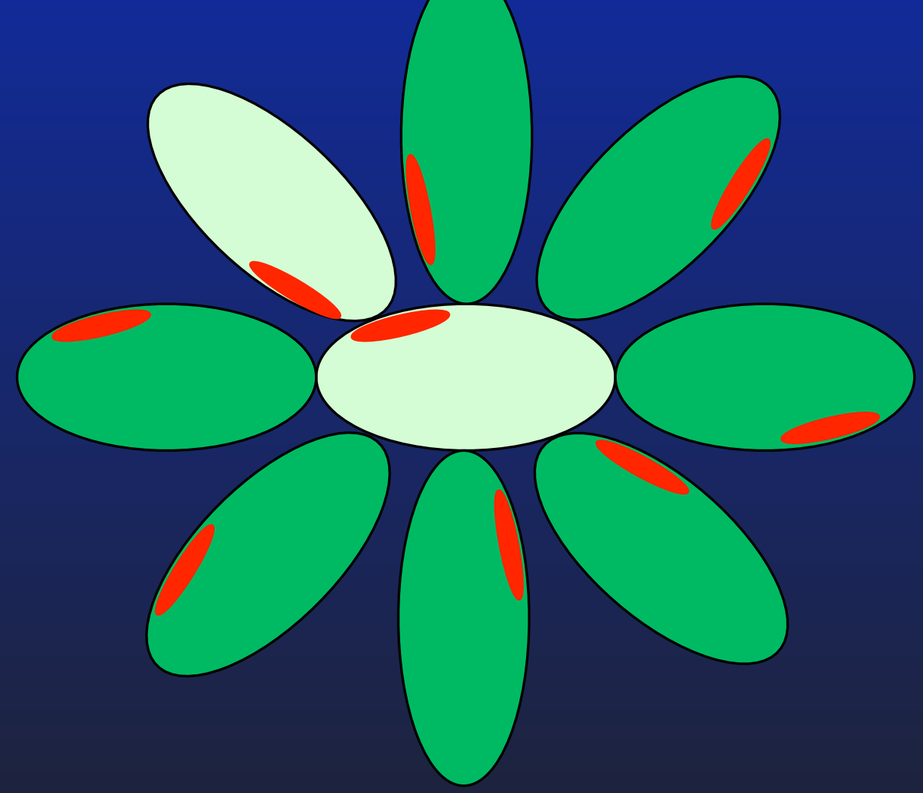
- Distortion

- Re-assembly





Probably, distortions are always there

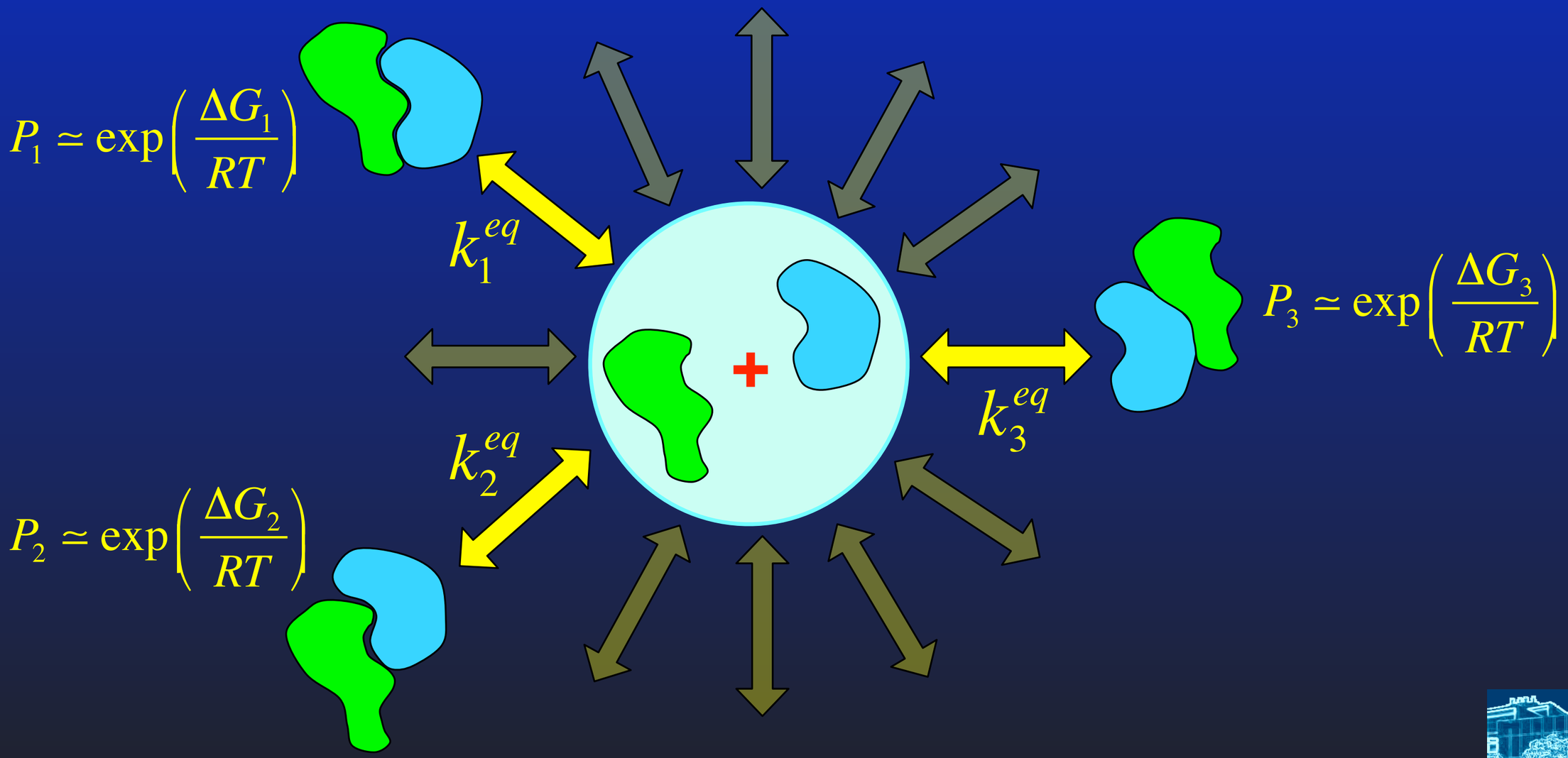


There is a chance for re-assembly if interaction is weak



Alternative assemblies

- All complexes (assemblies) have right to exist in solvent, however with different occurrence probabilities. These probabilities may differ of those in crystal environment, e.g., in case of substantially assisted crystallisation.





Real and superficial crystal contacts

- If a crystal contact remains thermodynamically preferential in solution, the chances are that it represents a biochemically relevant interaction

- Experimental (not crystallographic) data on structure of complexes in solutions is *very* sparse

- One can hope to get some clues using computational docking, assuming that docking approximates in-solvent situation

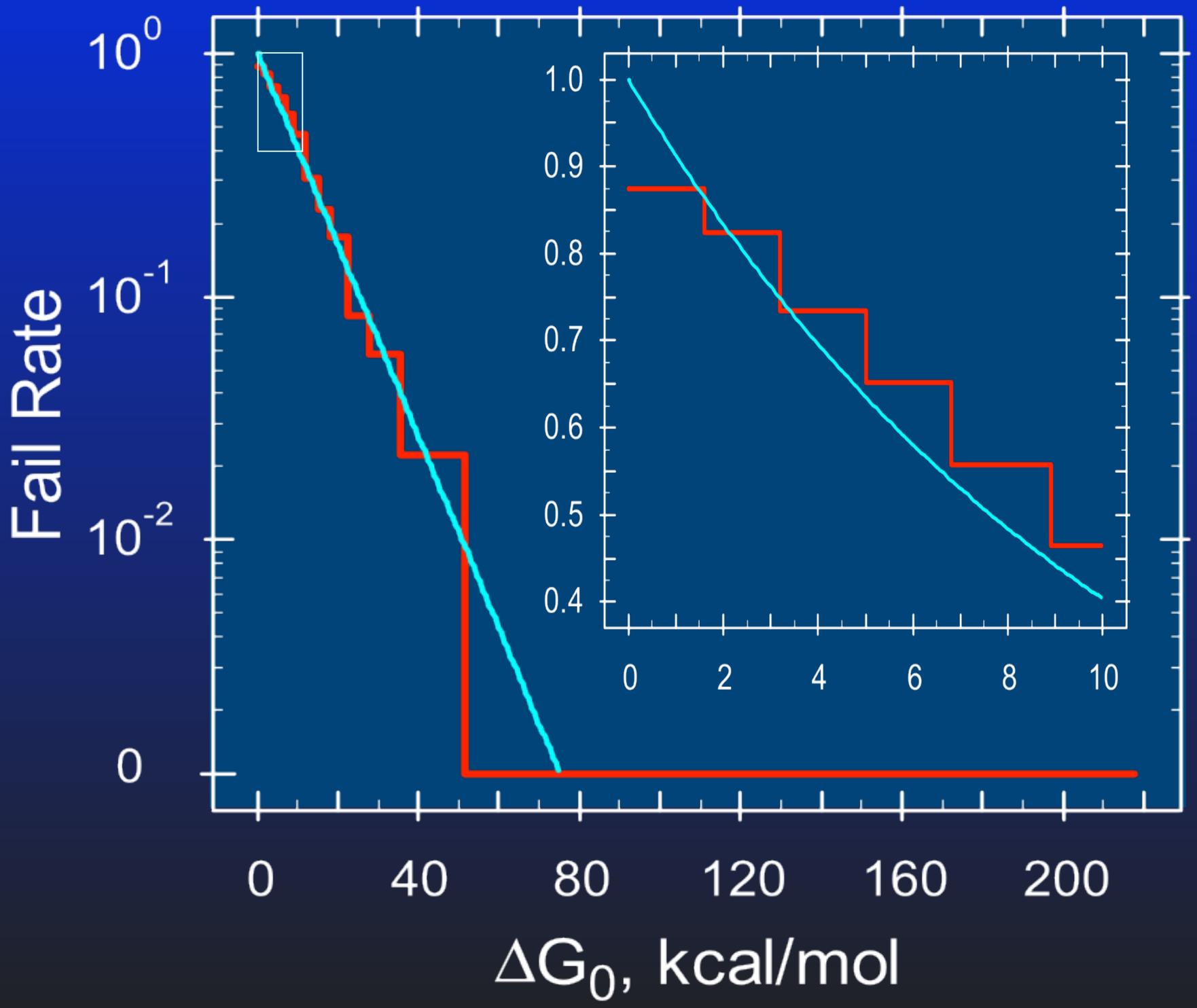
- Being applied to 4065 non-redundant dimers from the PDB, docking fails to arrive at crystal interface in 38% of instances

E. Krissinel (2010) J. Comp. Chem. 31, 133-143



Research Complex at Harwell

Fail rate of docking



The plot shows the probability of docking not to arrive at crystal interface, as a function of interface free energy.

The probabilities are calculated using equipopulated bins.

Overall, 38% of failures.

Good news: at high ΔG^0_{diss} errors disappear

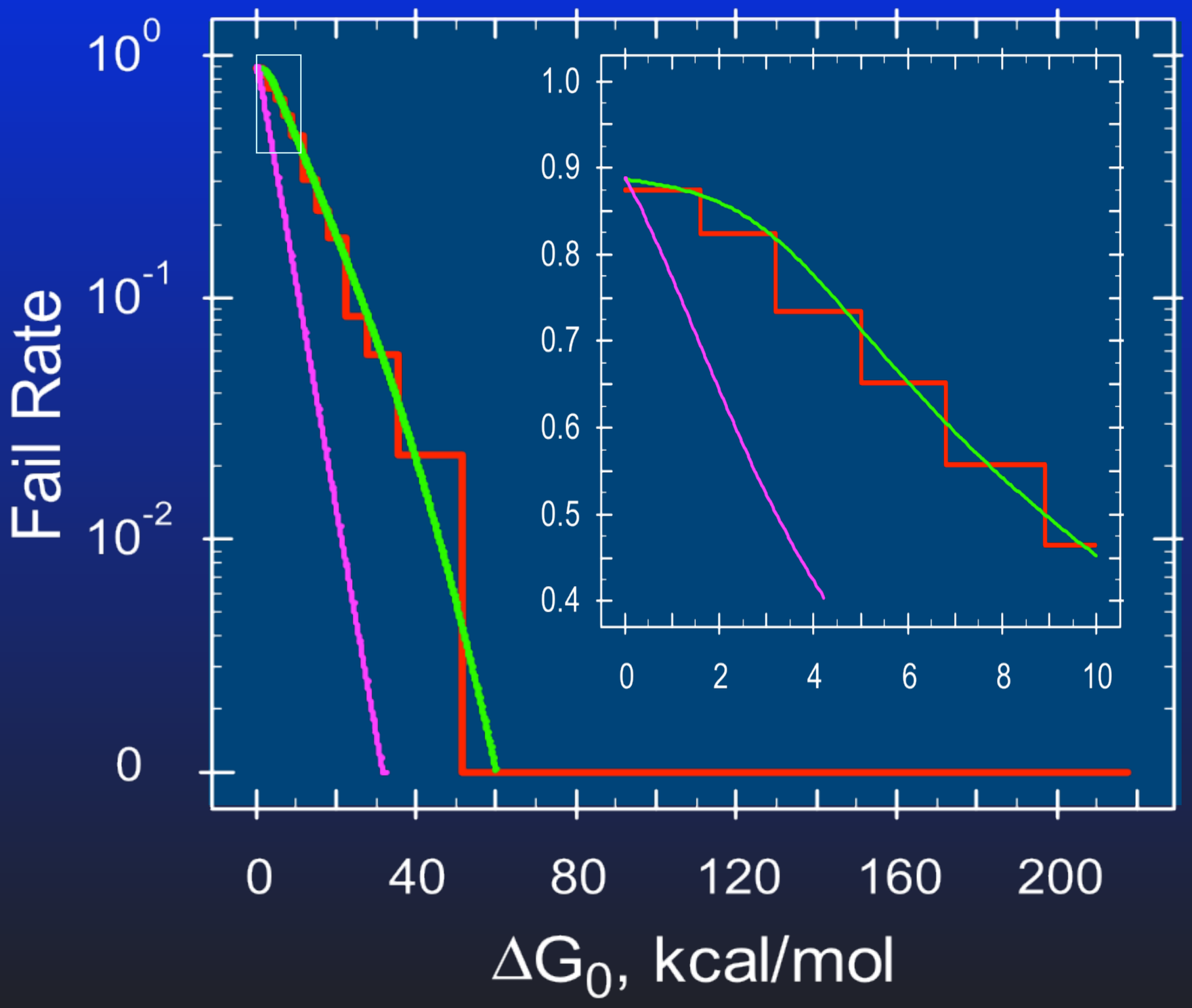
Bad news: biologically interesting interactions are normally weak

E. Krissinel (2010) J. Comp. Chem. 31, 133-143



Research Complex at Harwell

Calculation errors and crystal misrepresentation effects



probability of docking to fail to arrive at crystal contact

best fit (a formula) when both calculation errors and misrepresentation effects are taken into account

pure crystal misrepresentation effect (same formula with "zero" calculation errors substituted)

E. Krissinel (2010)
J. Comp. Chem. **31**, 133-143



Research Complex at Harwell

So what is the practicality of all this?

- PISA *is not* a substitution for experiments on the identification of protein's oligomeric state
  - both the software and (much less likely) experiment may give wrong results
  - in difference of experimental results, calculations do not make a scientific evidence!

- PISA may be used for choosing complex models for molecular replacement
  - already done in BALBES automatic molecular replacement pipeline

- PISA may be used for interpretation of experimental results when evidence is not sufficient for a definite answer
  - which dimer?
  - inconclusive evidence (e.g. oligomeric state highly dependent on concentration/ temperature/ion presence etc.)

- PISA may be used for sanity checks, comparative analysis and flag raising
  - is proposed complex structure compatible with crystal packing?
  - is proposed complex different from close homologs?
  - is there a strong disagreement with biological/biochemical expectations?



Research Complex at Harwell

Acknowledgements

Kim Henrick *European Bioinformatics Institute*	General introduction and PQS expertise
Mark Shenderovich *Structural Bioinformatics Inc.*	Helpful discussion
Hannes Ponstingl *Sanger Centre*	Sharing expertise and benchmark data
Sergei Strelkov *University of Leuven*	“Mystery” of bacteriophage T4
MSD & PDB teams *EBI & Rutgers*	Everyday use of PISA, examples, verification and feedback
CCP4 *Daresbury-York-Oxford*	Encouragement, support and publicity
~10,000 PISA users *Worldwide*	Using PISA and feedback
Biotechnology and Biological Sciences Research Council *(BBSRC) UK*	Research grant No. 721/B19544



Research Complex at Harwell