National
Institute of Chemistry
Slovenia

Designable modularity in synthetic biology

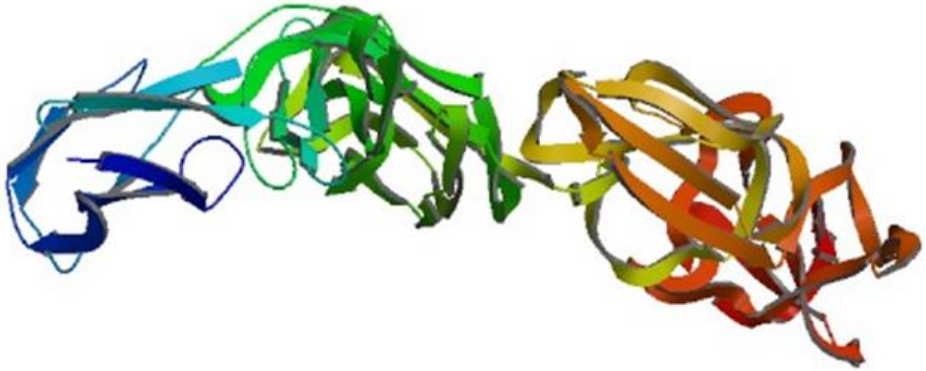
from engineering logic functions into cells to the design of new protein folds

Roman Jerala

Department of biotechnology
National institute of chemistry
Ljubljana, Slovenia

Structural and functional modularity of proteins

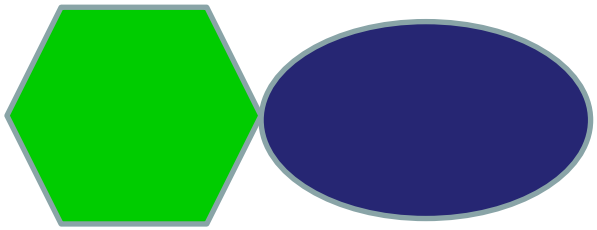
Modularity of proteins



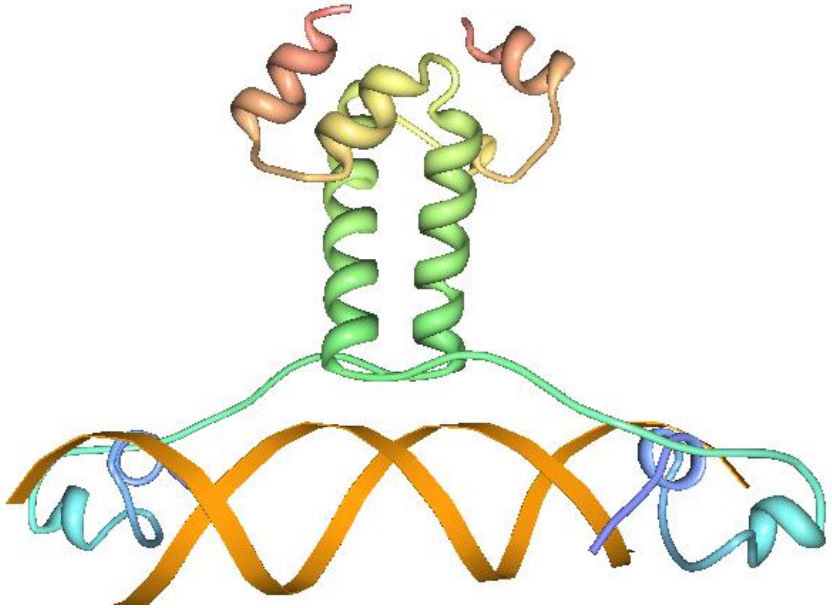
Advantages:

- Lower number of required *de novo* domains
- Combinations of modules increases the set of functionalities – accelerated evolution

Modularity of the transcriptional regulatory elements



DNA binding + effector domain

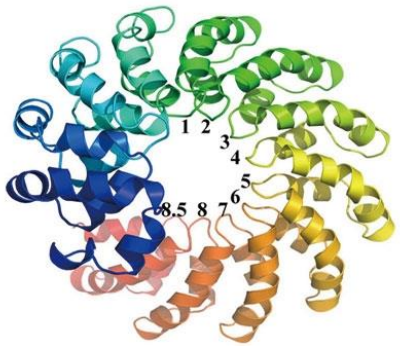


Instructions for the molecular assembly

❖DNA sequence + protein binding domains

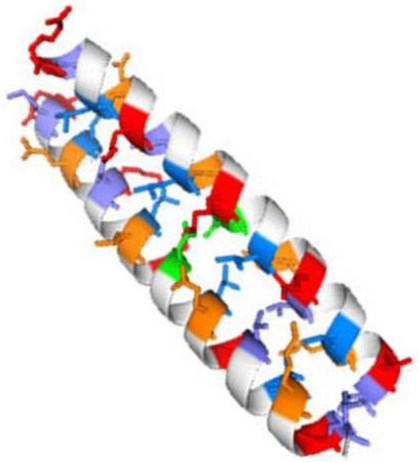
cellular program code with large complexity and can be easily designed (DNA synthesis).

Designable DNA binding domains/complexes
(e.g. zinc fingers, TALE domains, CRISPR/Cas)



❖Protein-protein interactions

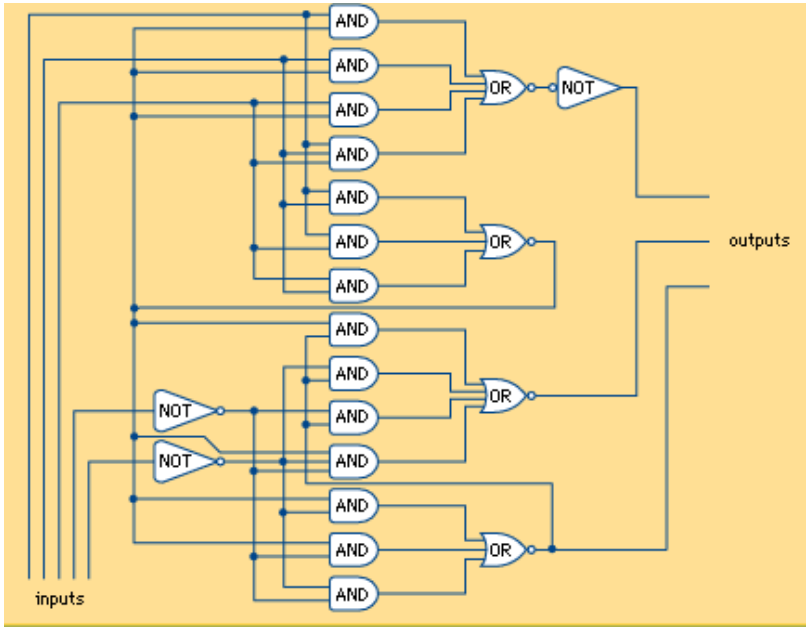
Engineering of polypeptide interactions based on well-understood rules (designed coiled-coil assemblies)



Information transfer in electronic vs. cellular circuits

Electronic circuits

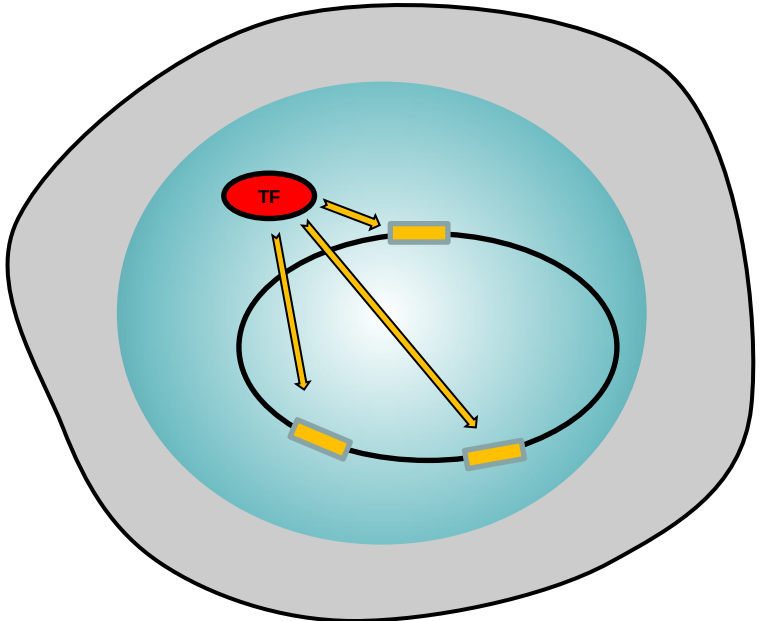
Conductive wires control the flow of information



Cells

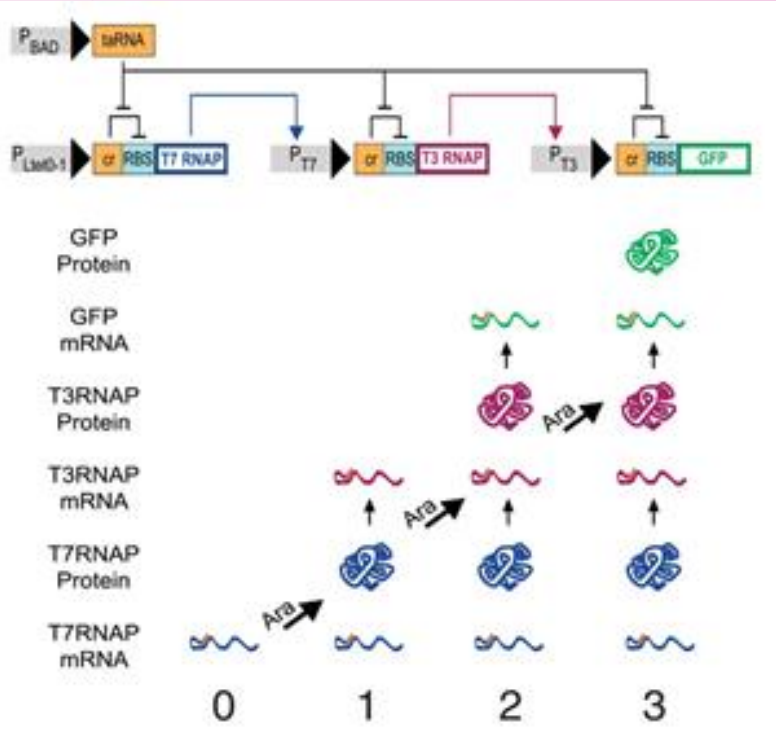
Free diffusion: transcription factors act on all binding sites within each cell

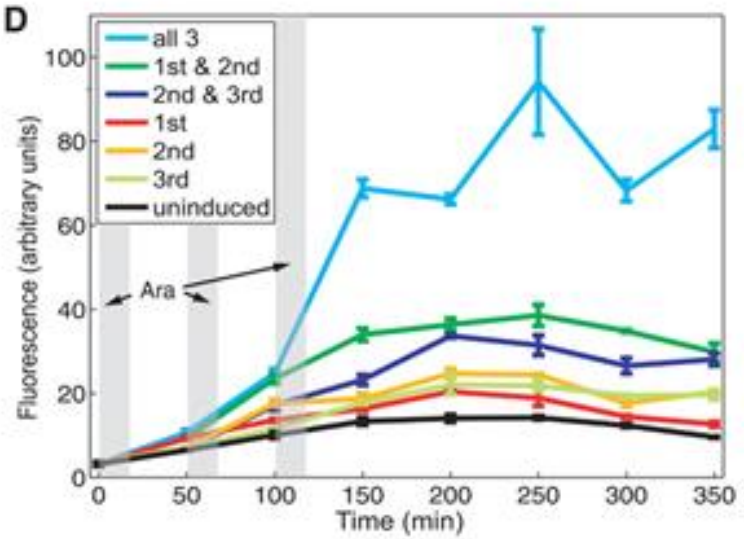
ORTHOGONALITY



Limitations of designed circuits

Cellular circuit that counts (up to 3)

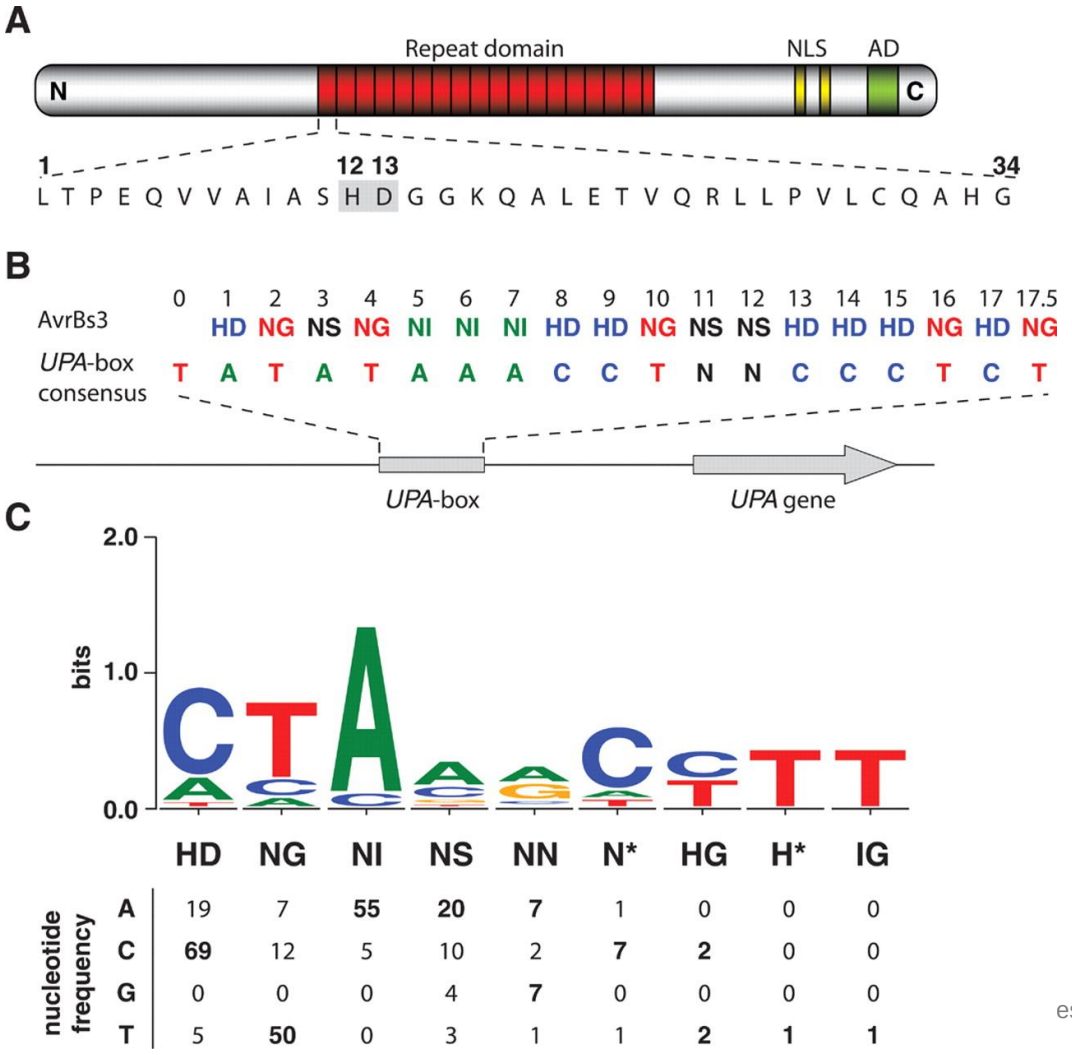




Friedland et al., Science 2009

An ideal toolbox of designed TF

Transcriptional activator-like (TAL) effectors



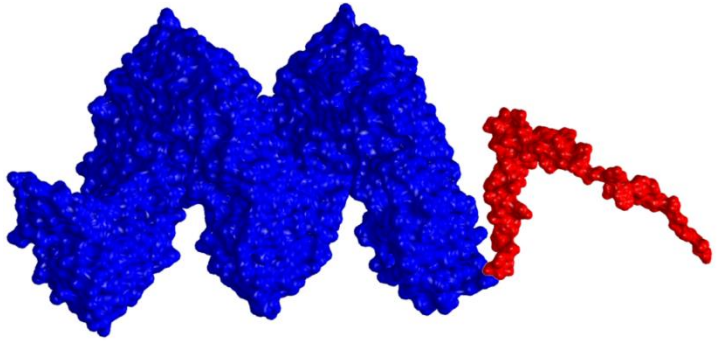
VDR

variable diresidue repeat (34 aa)

DNA recognition code

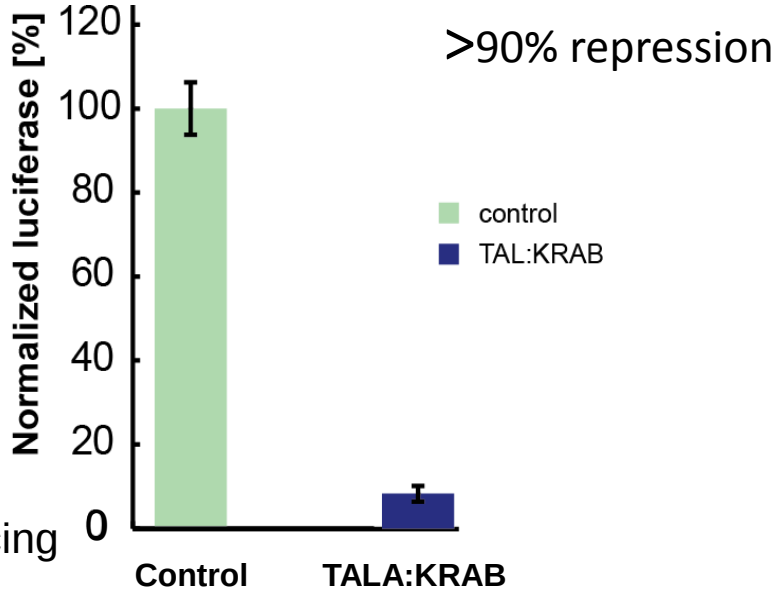
Bloch et al., Science 2009

Designed TAL repressors and activators

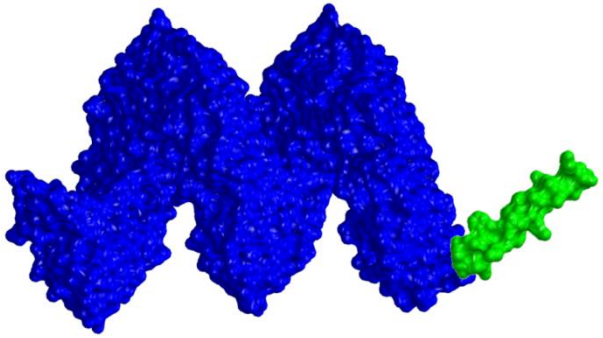


TAL KRAB

KRAB: Krueppel-associated box – chromatin silencing

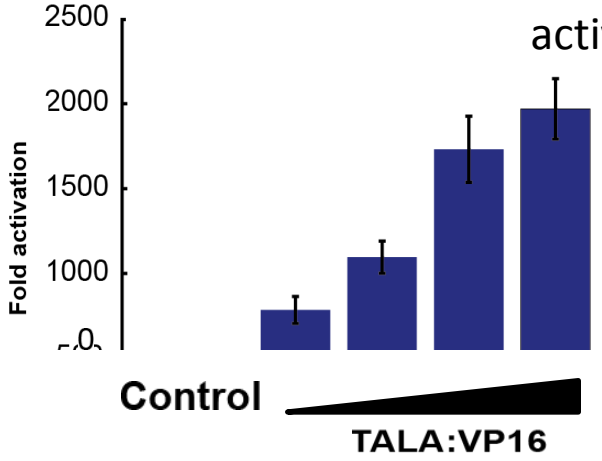


>90% repression



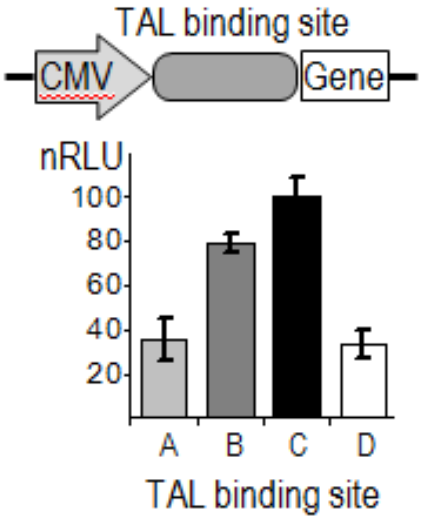
TAL VP16

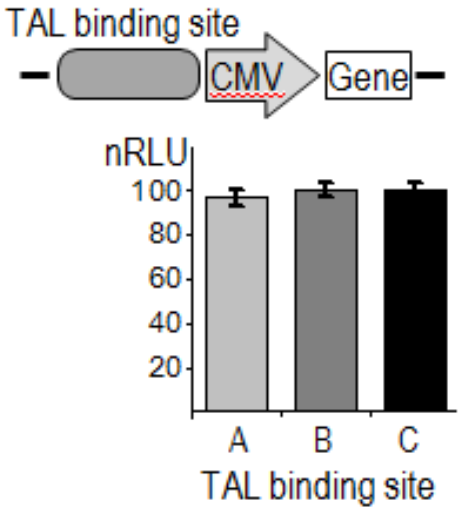
VP16: recruitment of transcriptional machinery

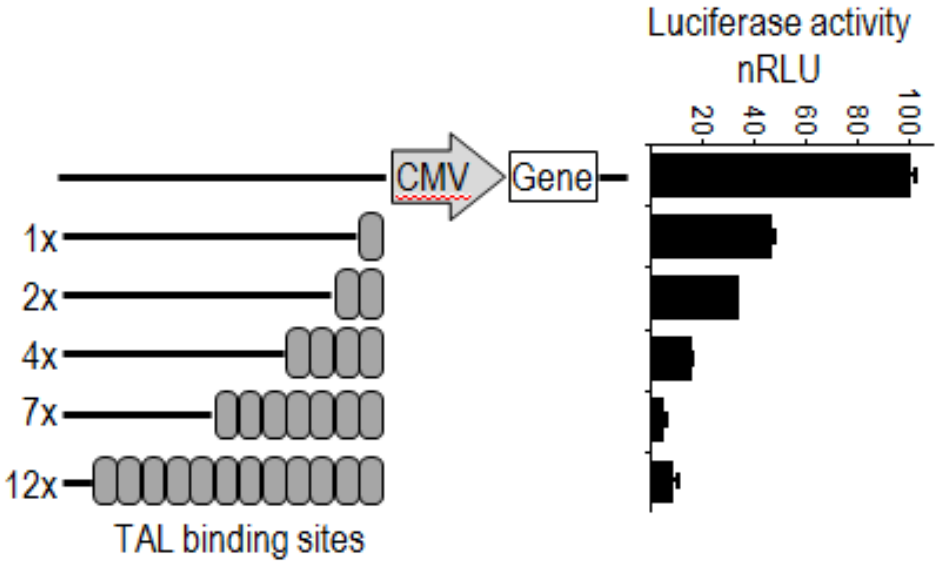


>1500-fold activation

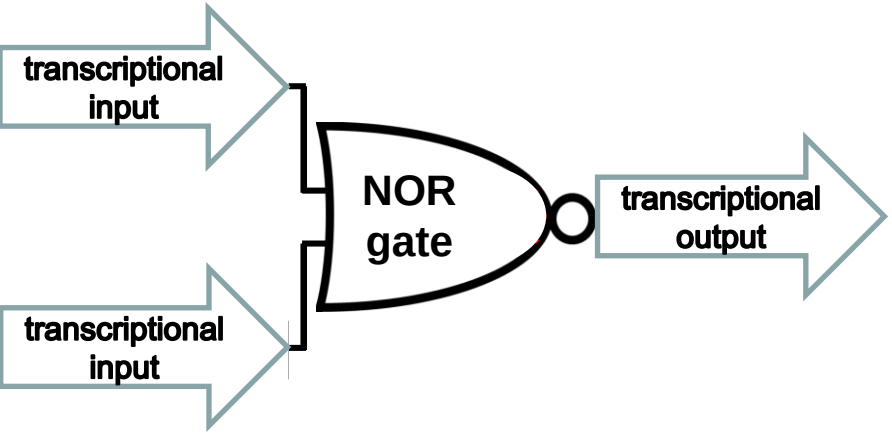
Fine tuning of designed repressors

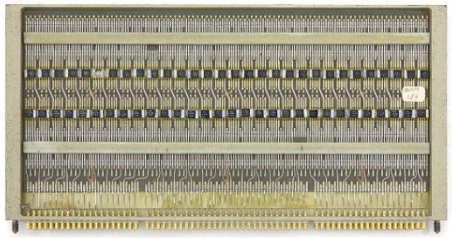




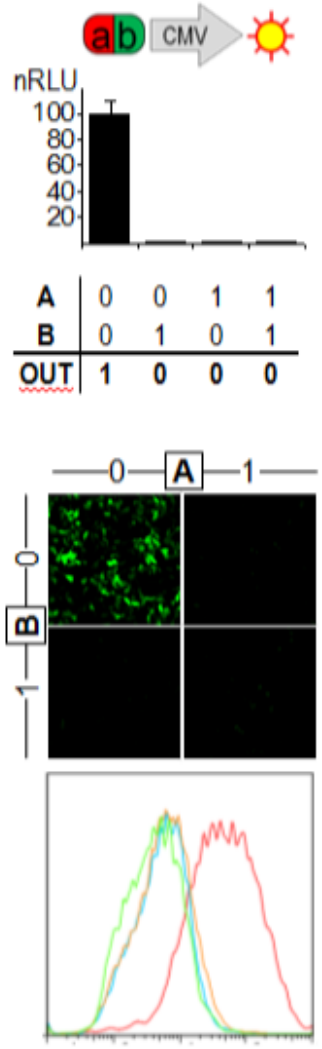


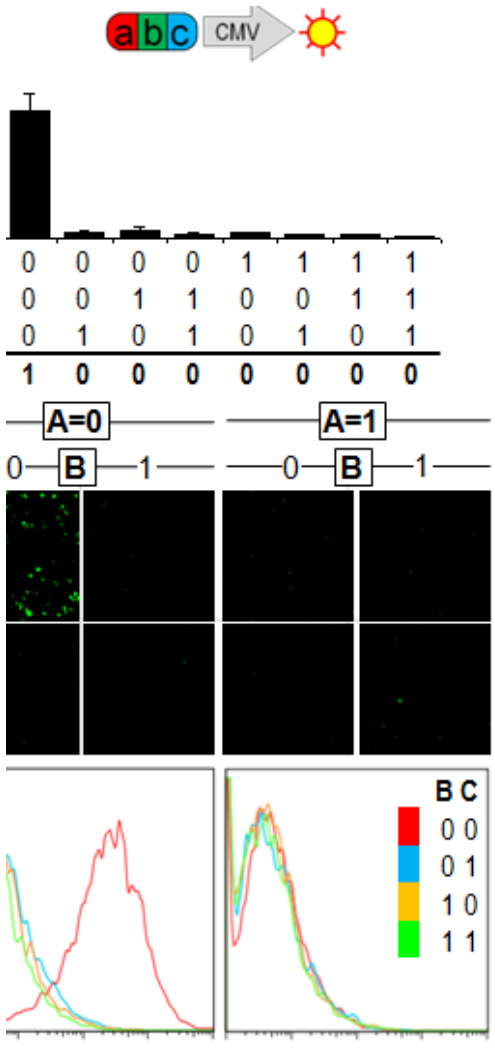
Designed NOR gate



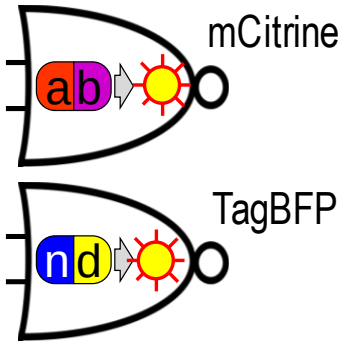


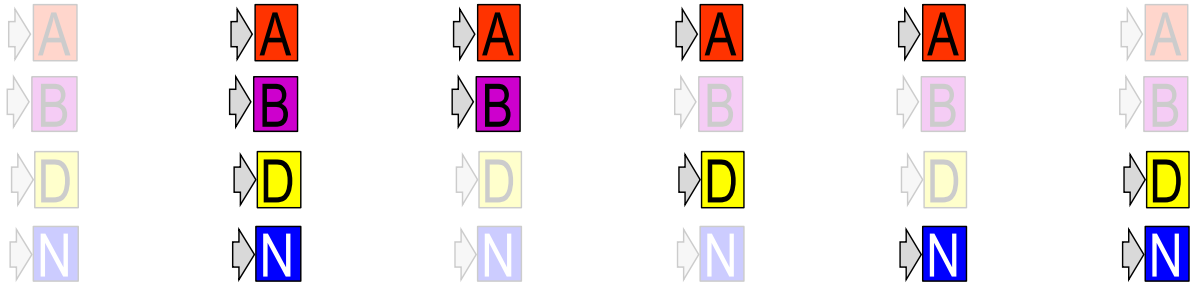
Appolo guidance computer
Micrologic chip –
5600 triple NOR gates

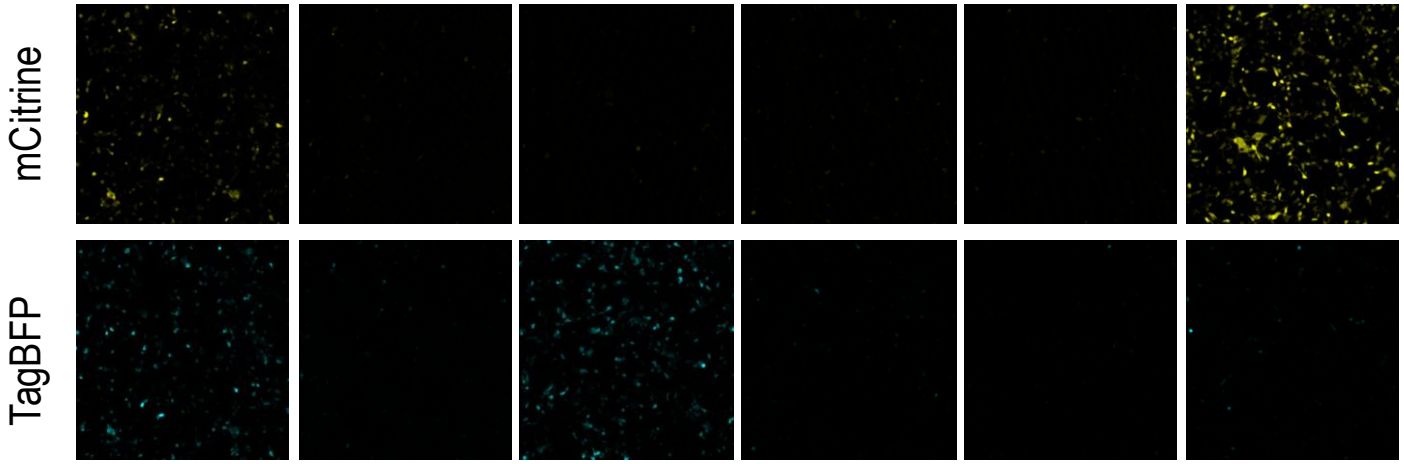




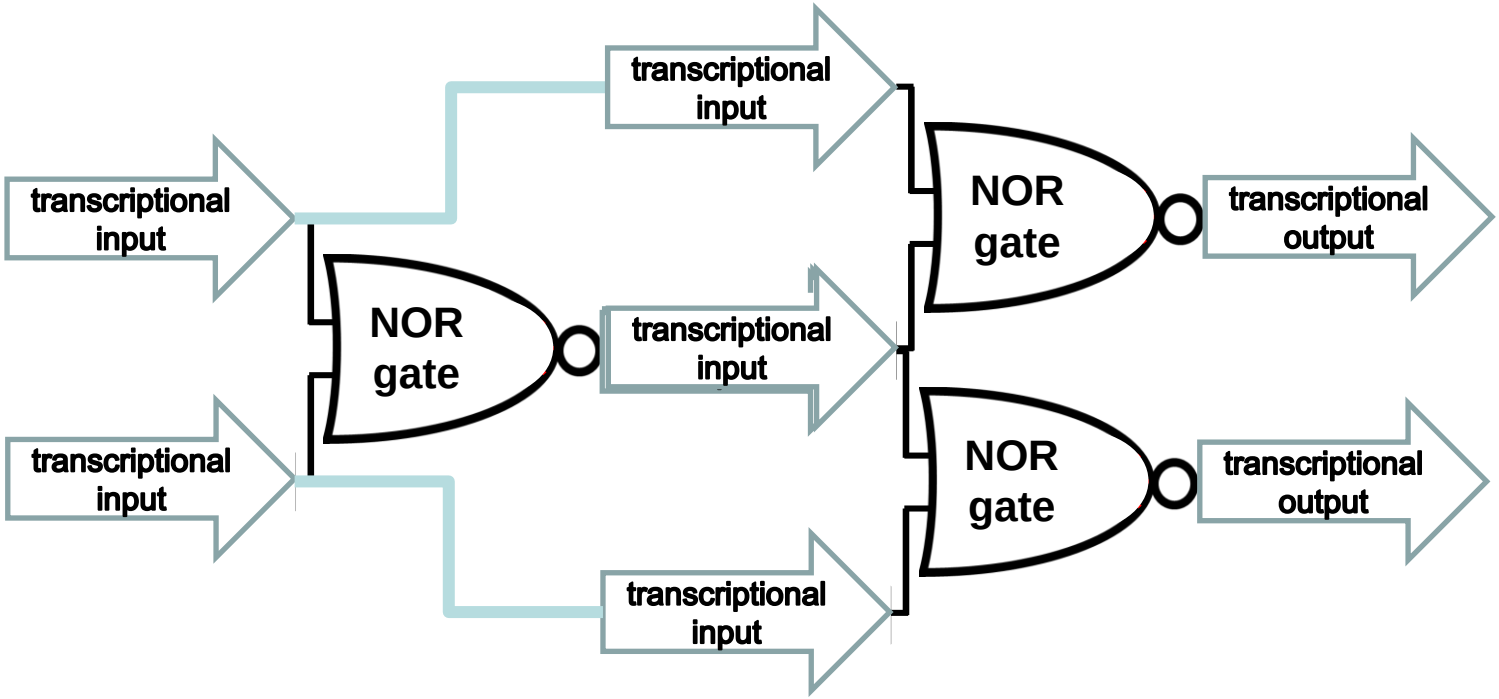
Orthogonality of NOR gate

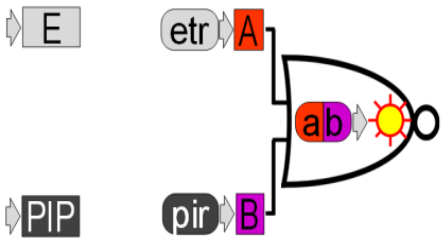


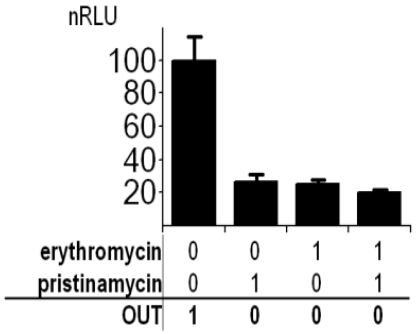




Layered NOR gates for complex functions

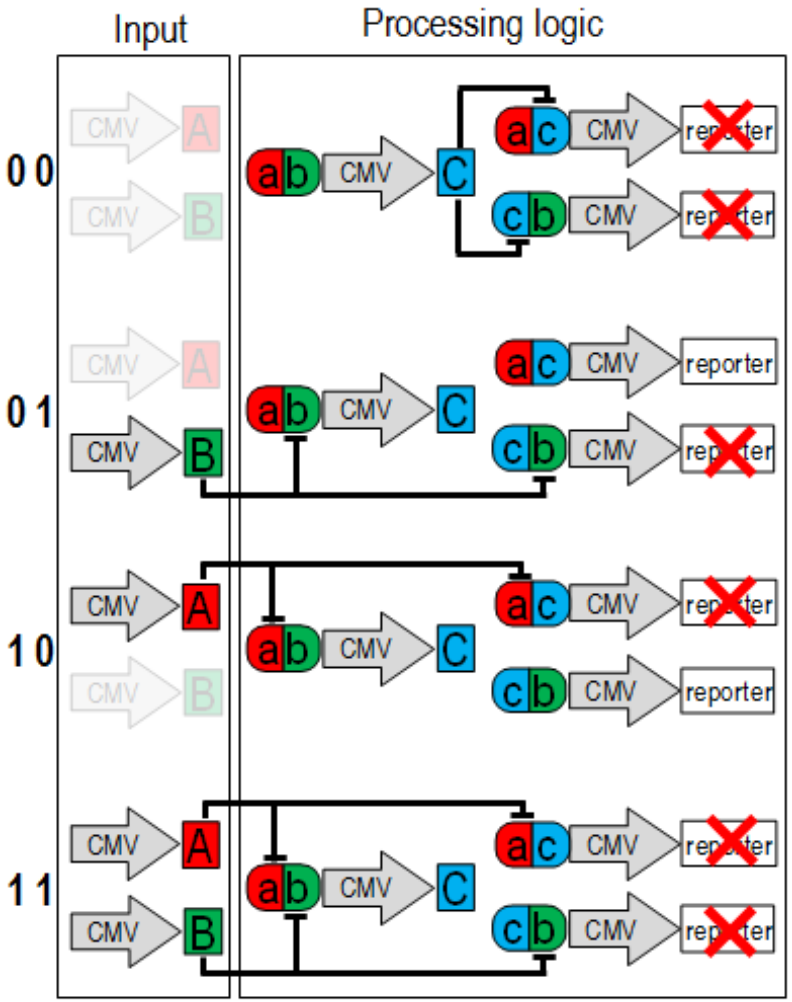


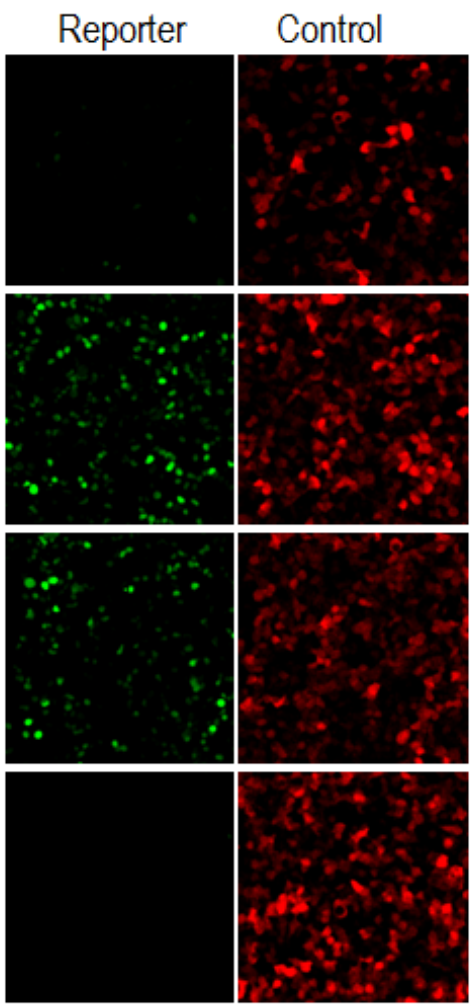


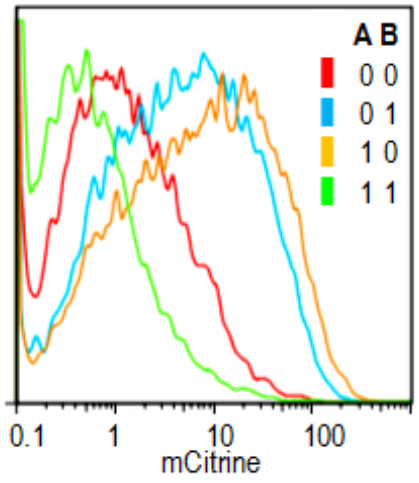


Input based on chemical inducers

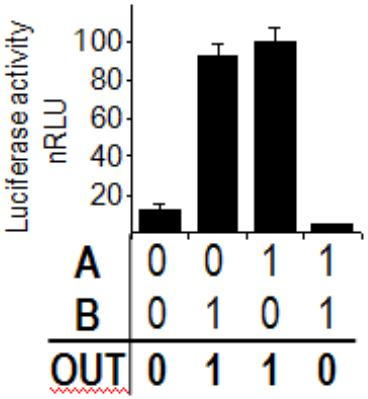
XOR function





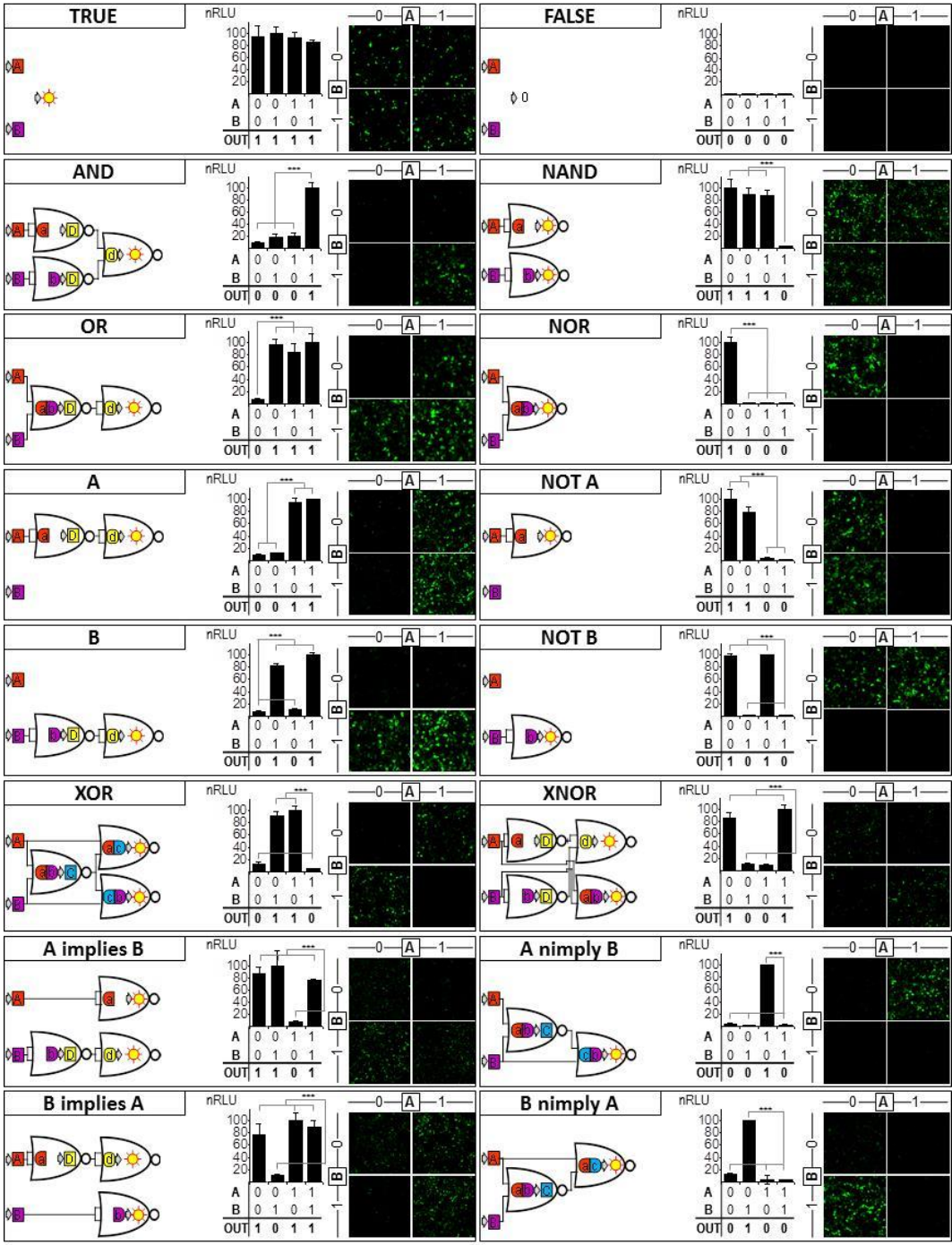


D



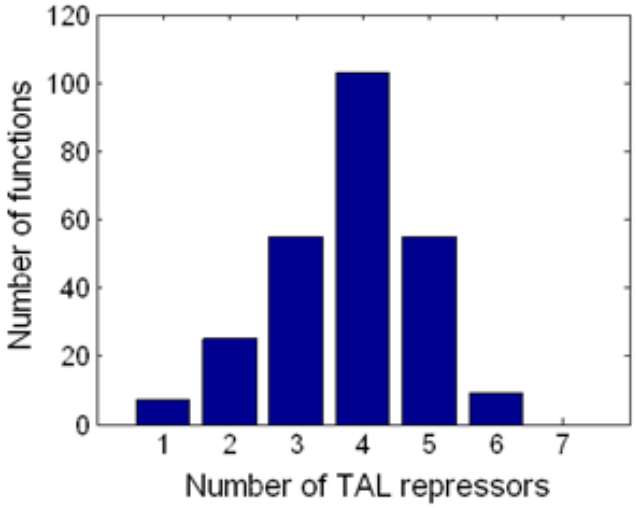
Logic functions

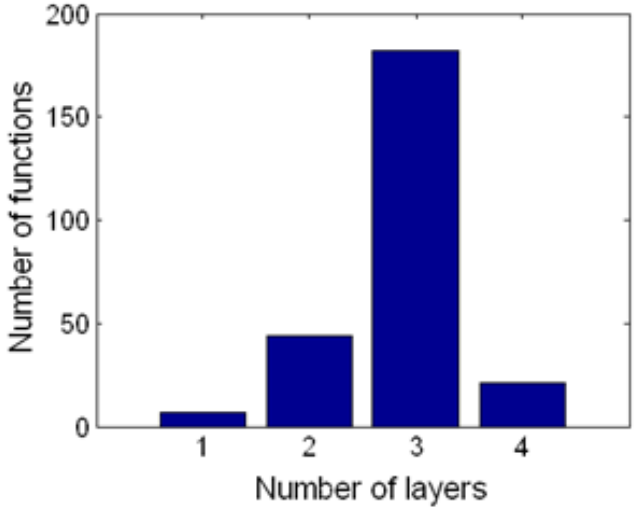
Implementation of all
16 two-input
logic functions based
on genetic NOR gate

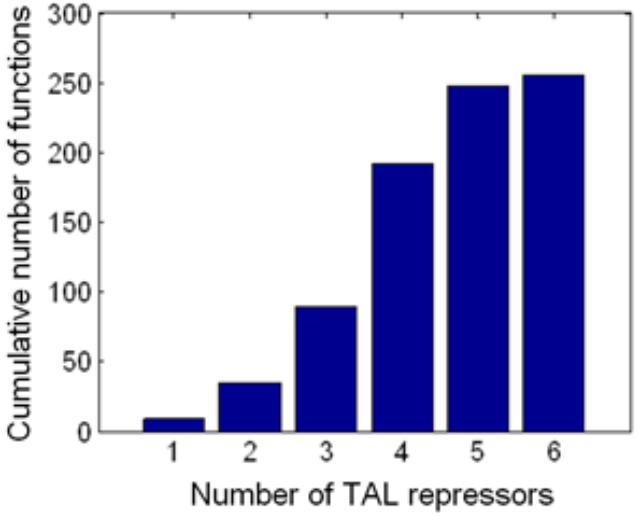


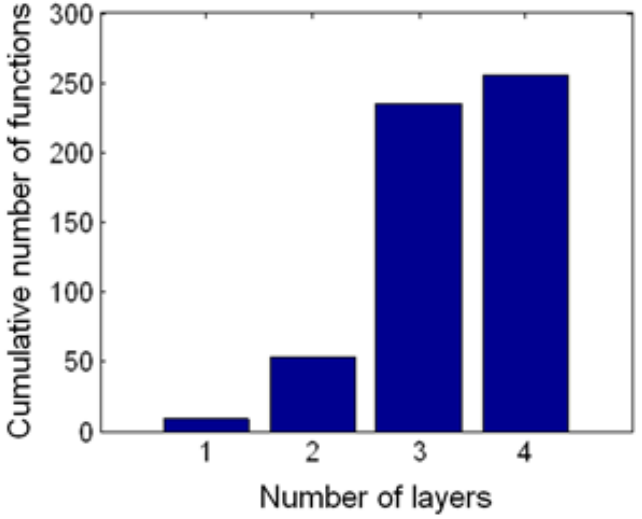
Gaber et al., Nature Chem Biol 2013

Implementation of triple input logic functions with designed repressors

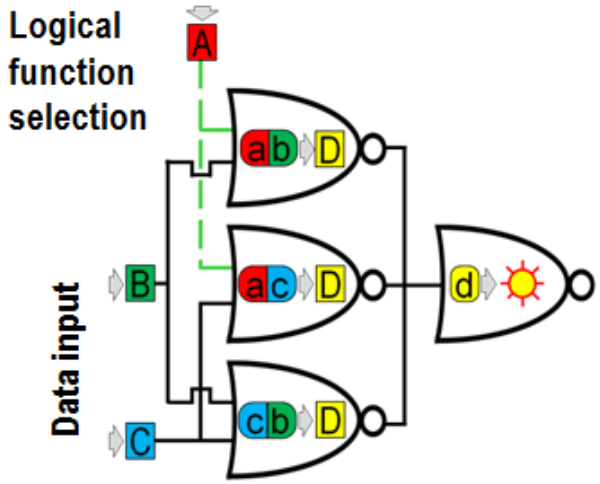


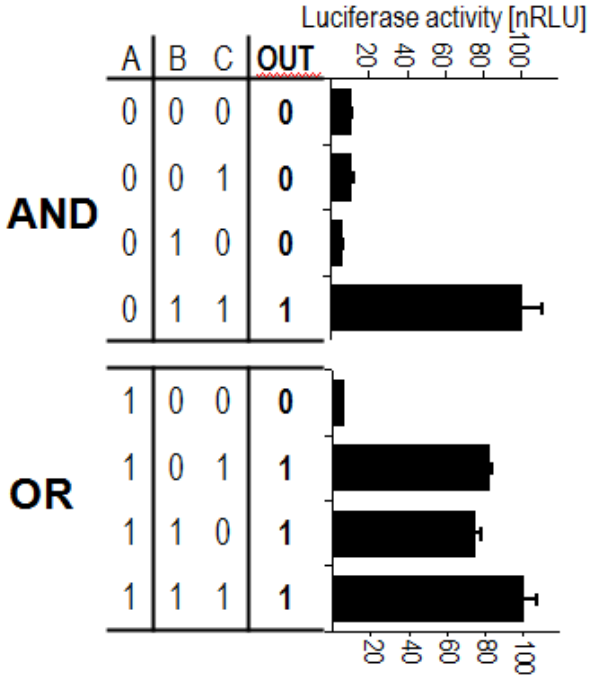


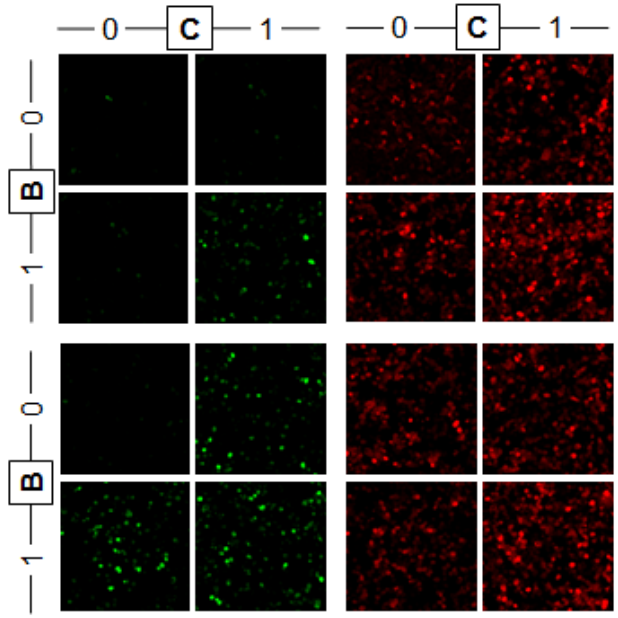




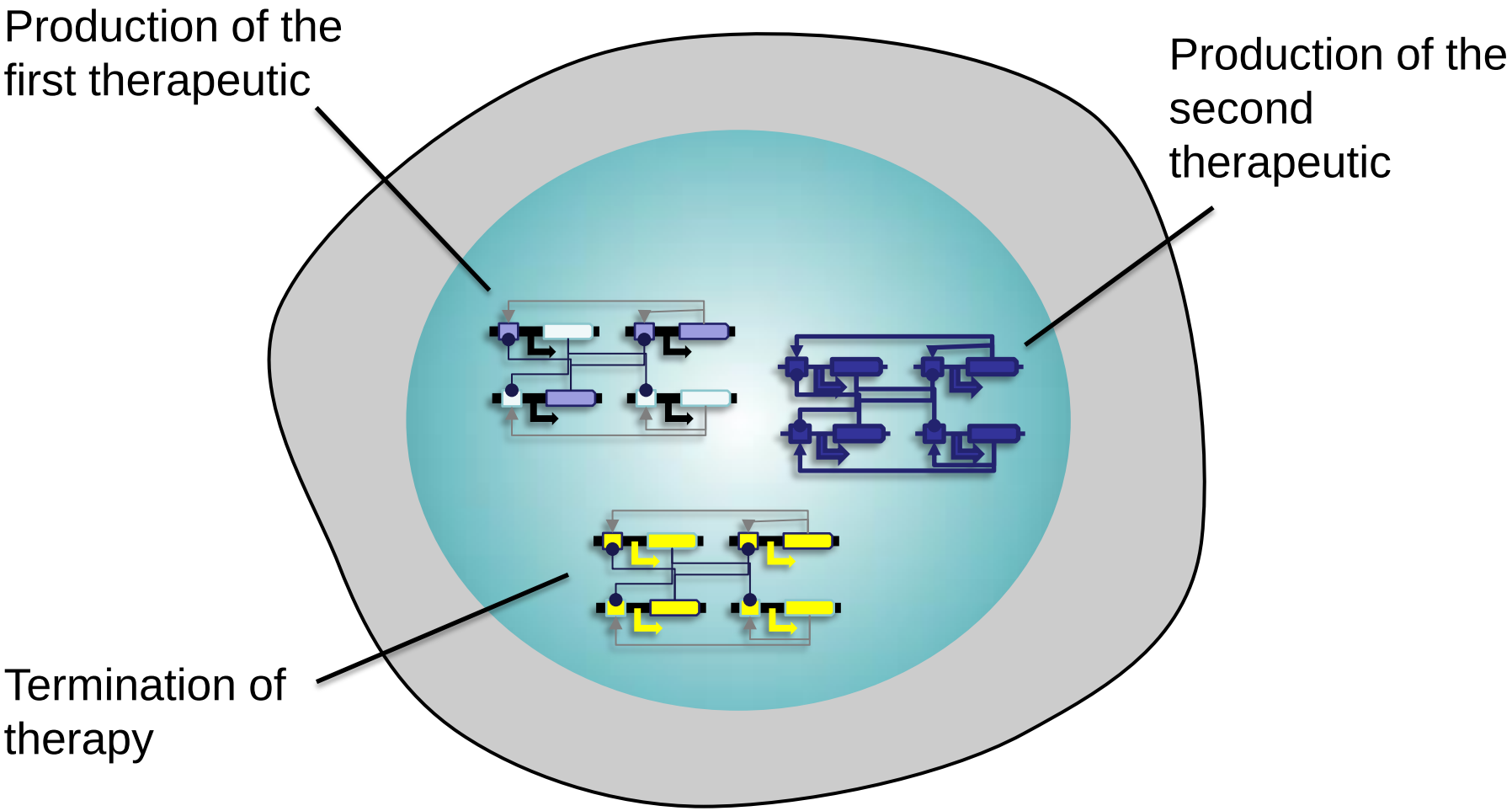
Selection of logic functions





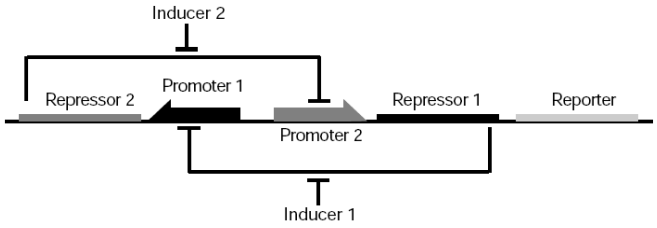


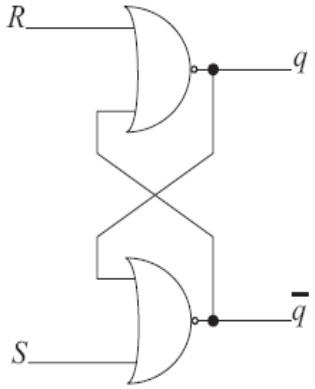
The need for multiple switches within engineered cells



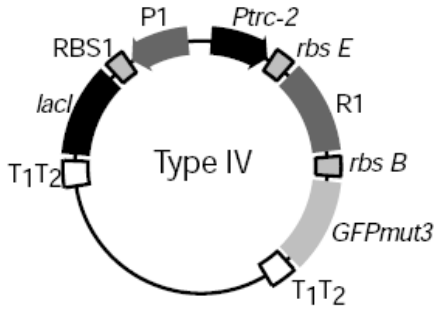
Construction of memory cell from 2 NOR gates

Bistable – toggle switch

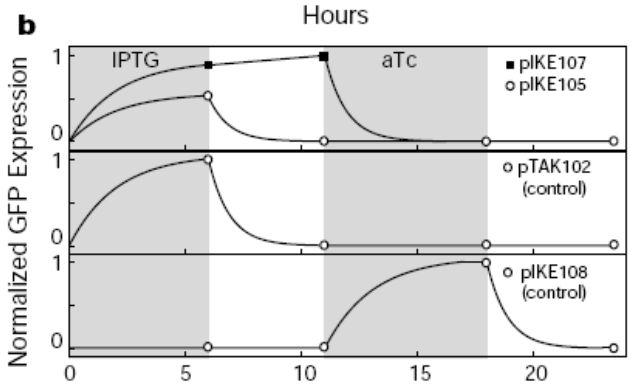




Logical outline of RS memory cell

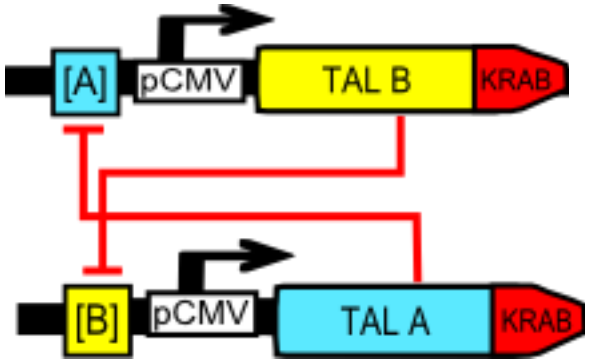


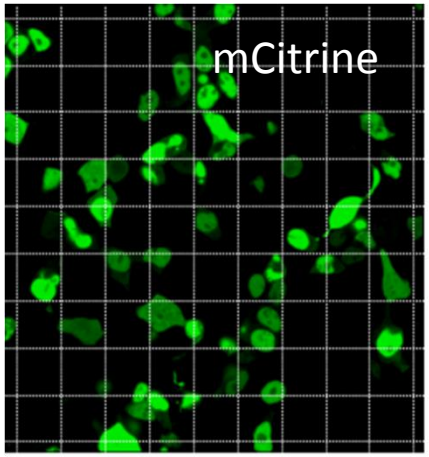
Gardner et al., Nature 2000

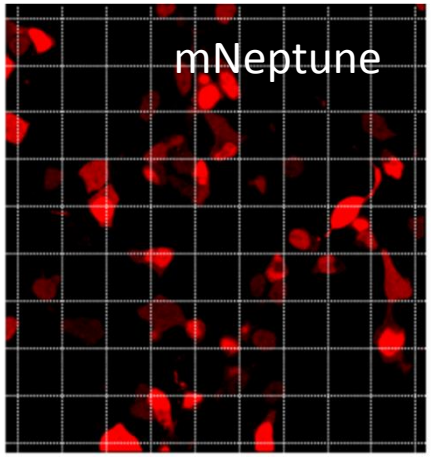


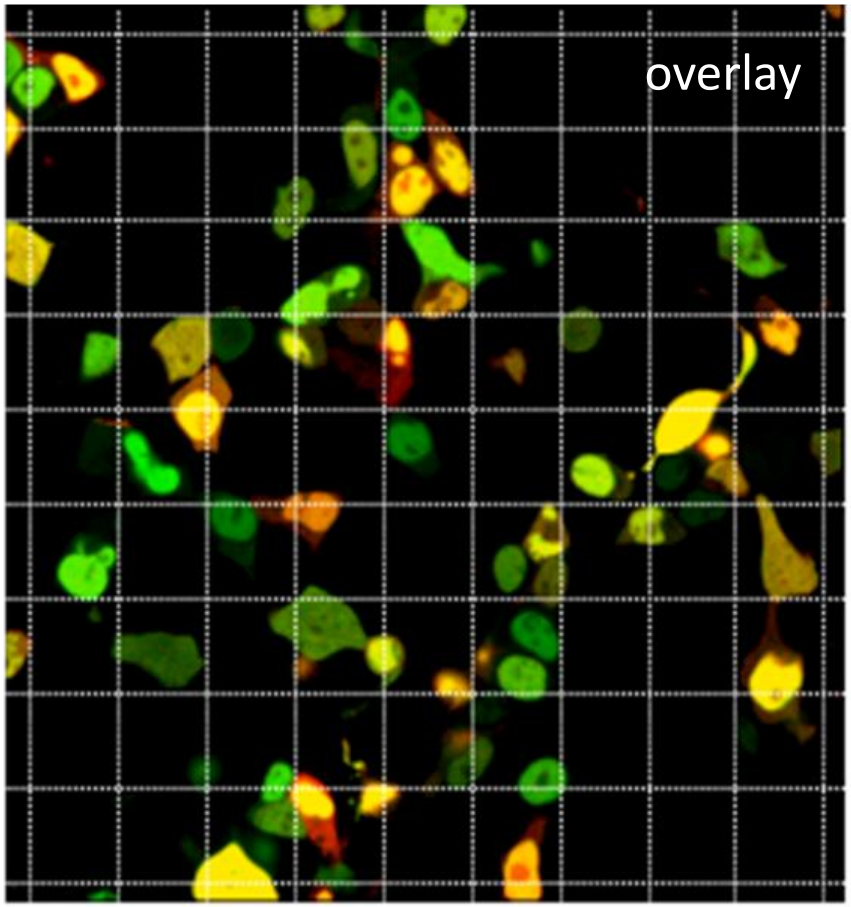
Mutual repressor switch based on designed DNA binding domains

80% of cells express both reporters

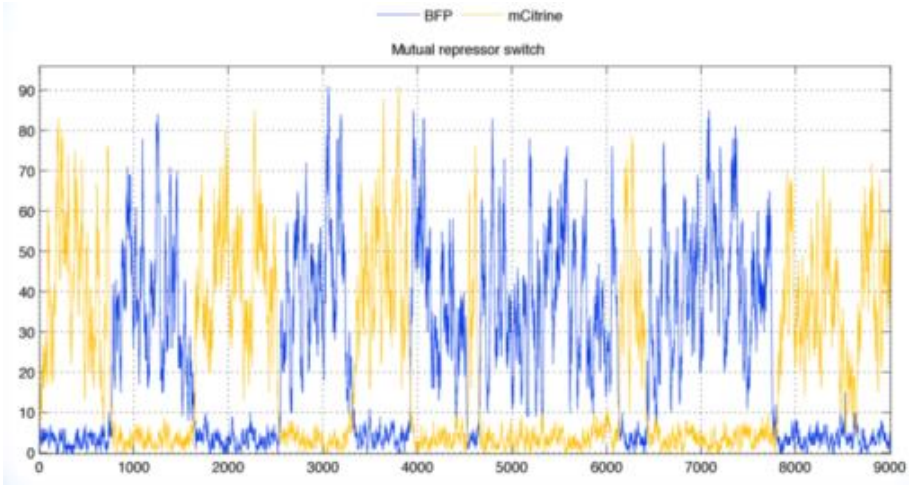




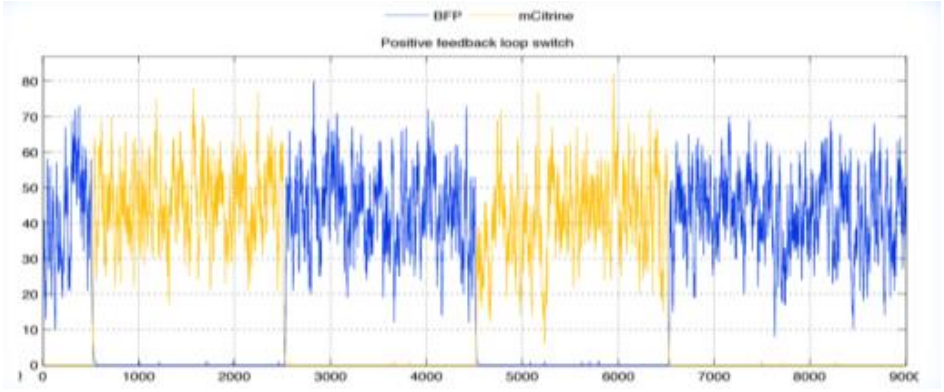




Mutual repressor switch simulation



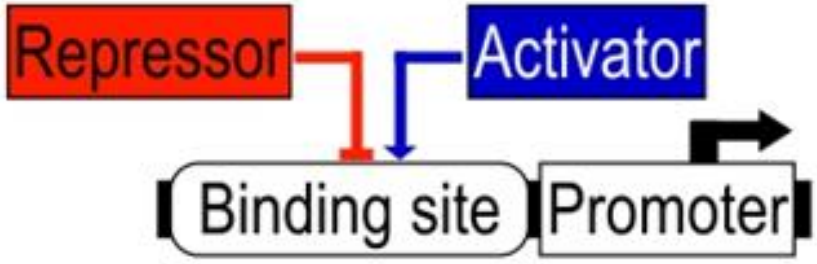
Stochastic switching between the two states: no stable state because of the linear response (monomers)

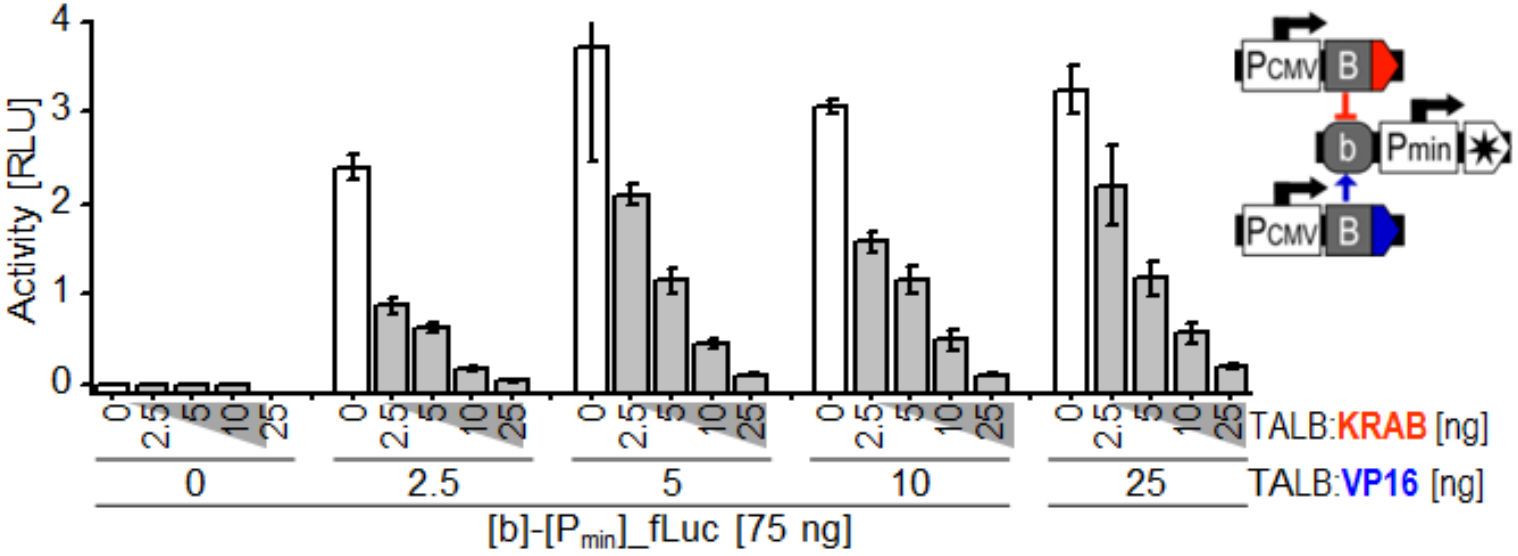


Cooperative behavior of transcription factors introduces nonlinearity and bistability

Introduction of nonlinear response

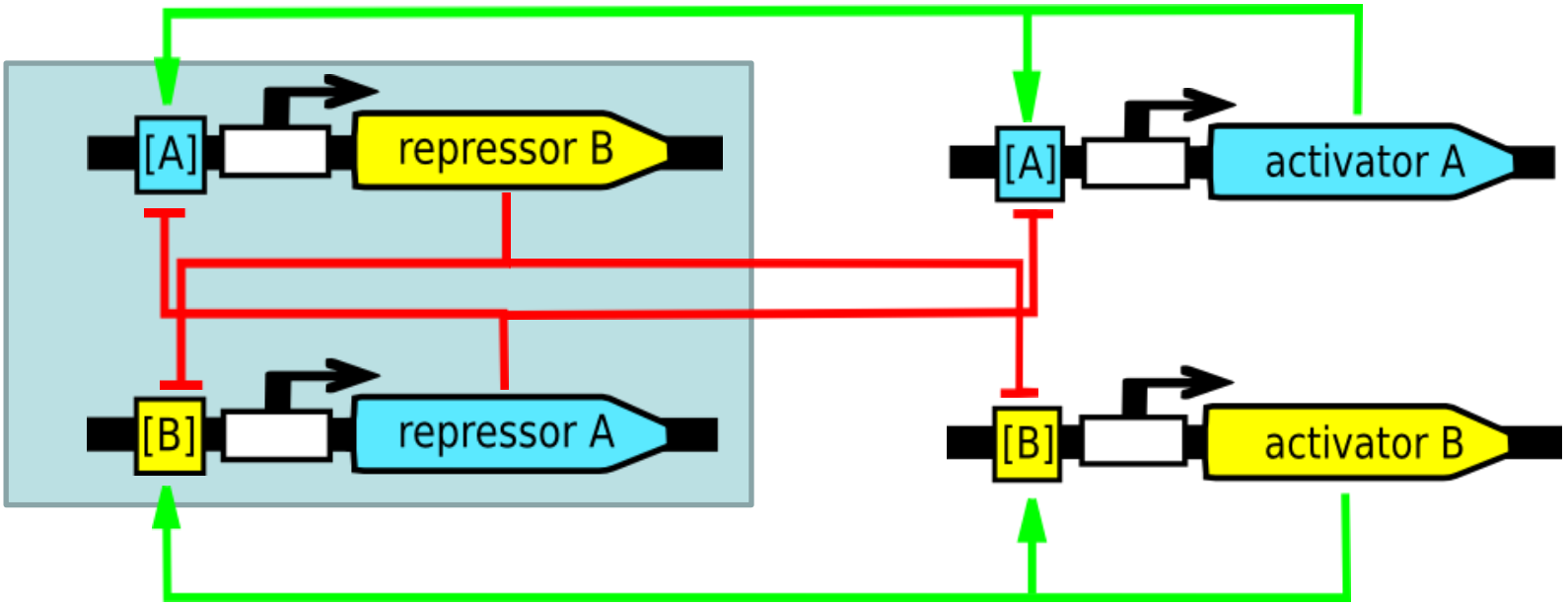
1. **Competition between repressor and activator** for the same binding site





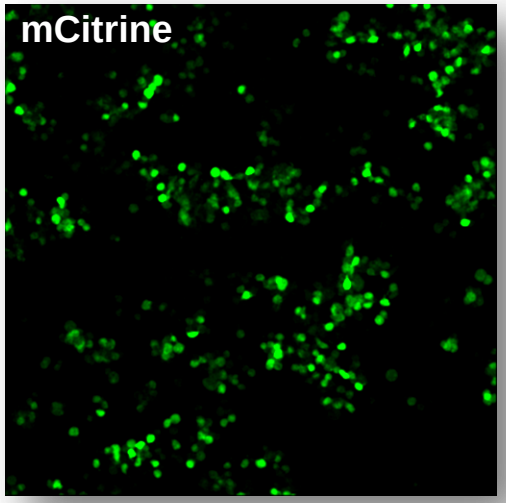
The competitive feedback loop switch

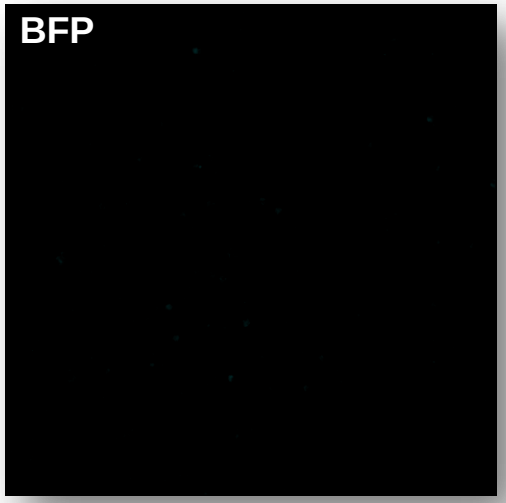
2. Introduction of a **positive feedback** loop

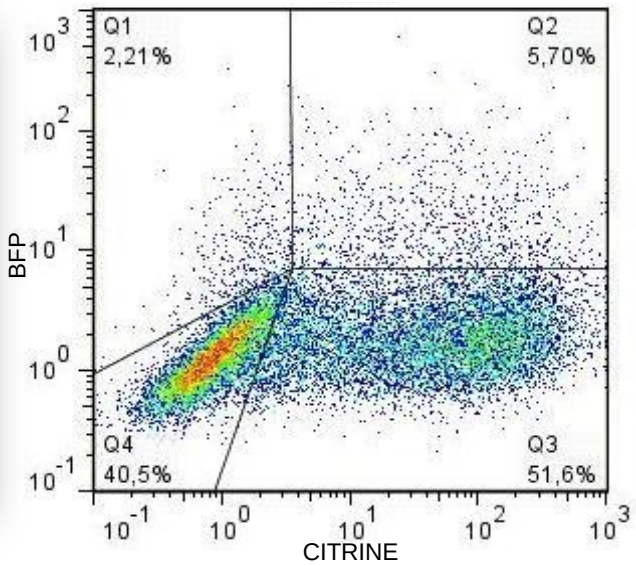


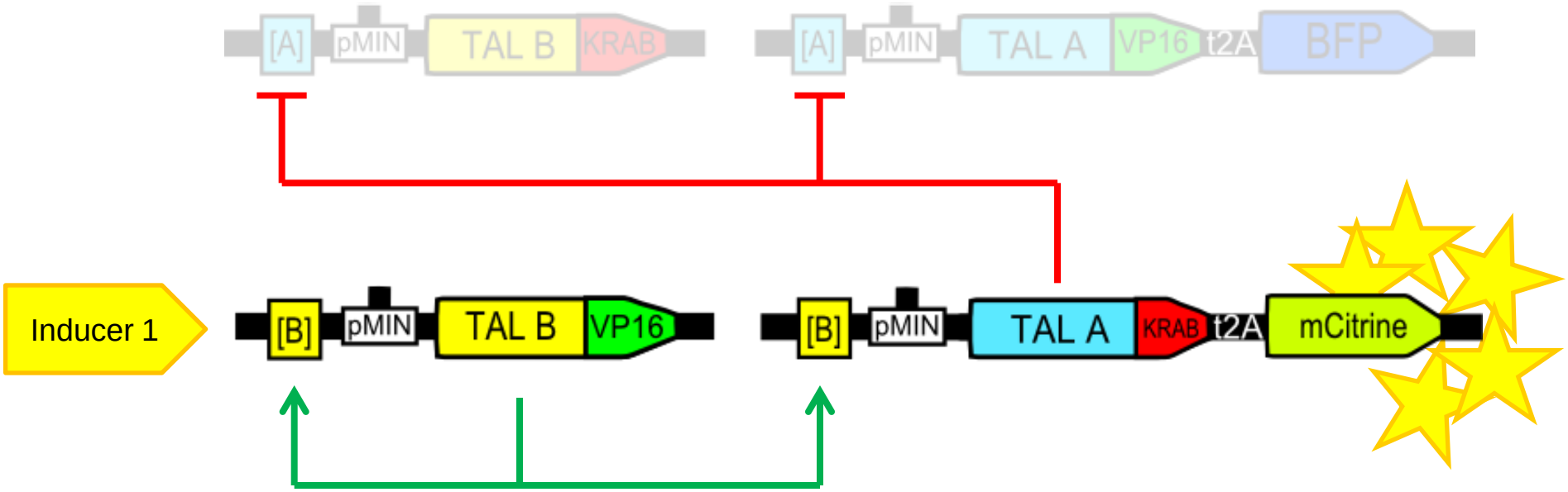
Lebar et al., Nature Communications 2014

+ inducer 1

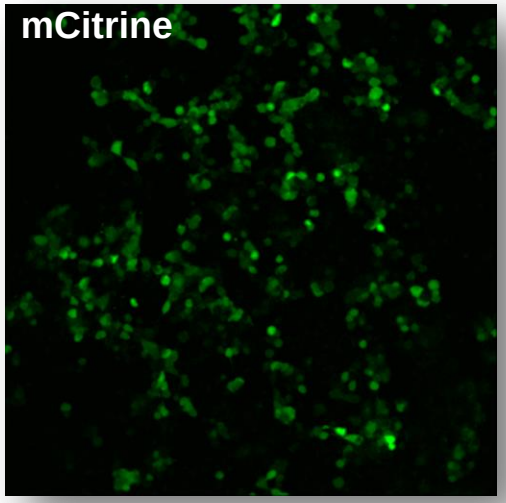

mCitrine

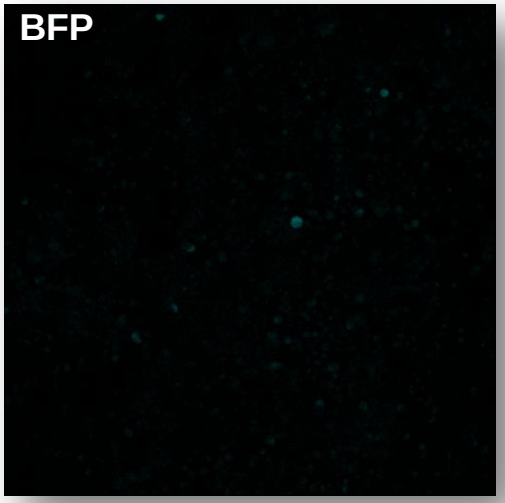

BFP

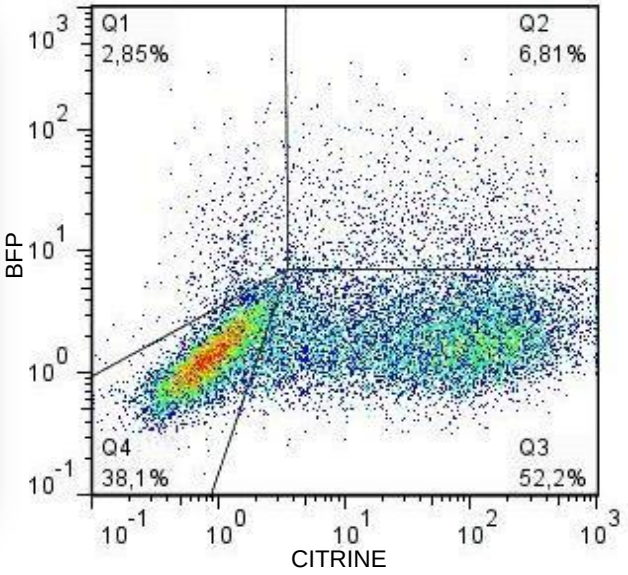

Q1
2,21%
Q2
5,70%
Q4
40,5%
Q3
51,6%
BFP
CITRINE

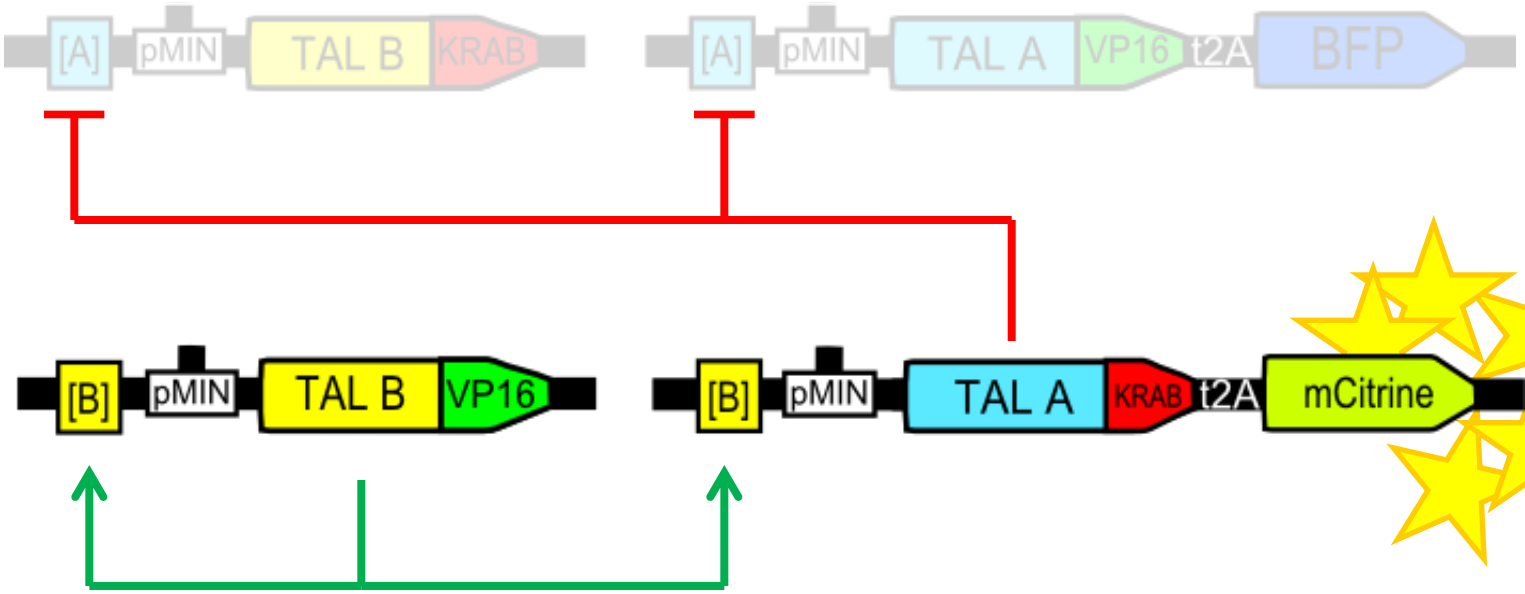

[A]
pMIN
TAL B
KRAB
[A]
pMIN
TAL A
VP16
t2A
BFP
Inducer 1
[B]
pMIN
TAL B
VP16
[B]
pMIN
TAL A
KRAB
t2A
mCitrine

Stable state

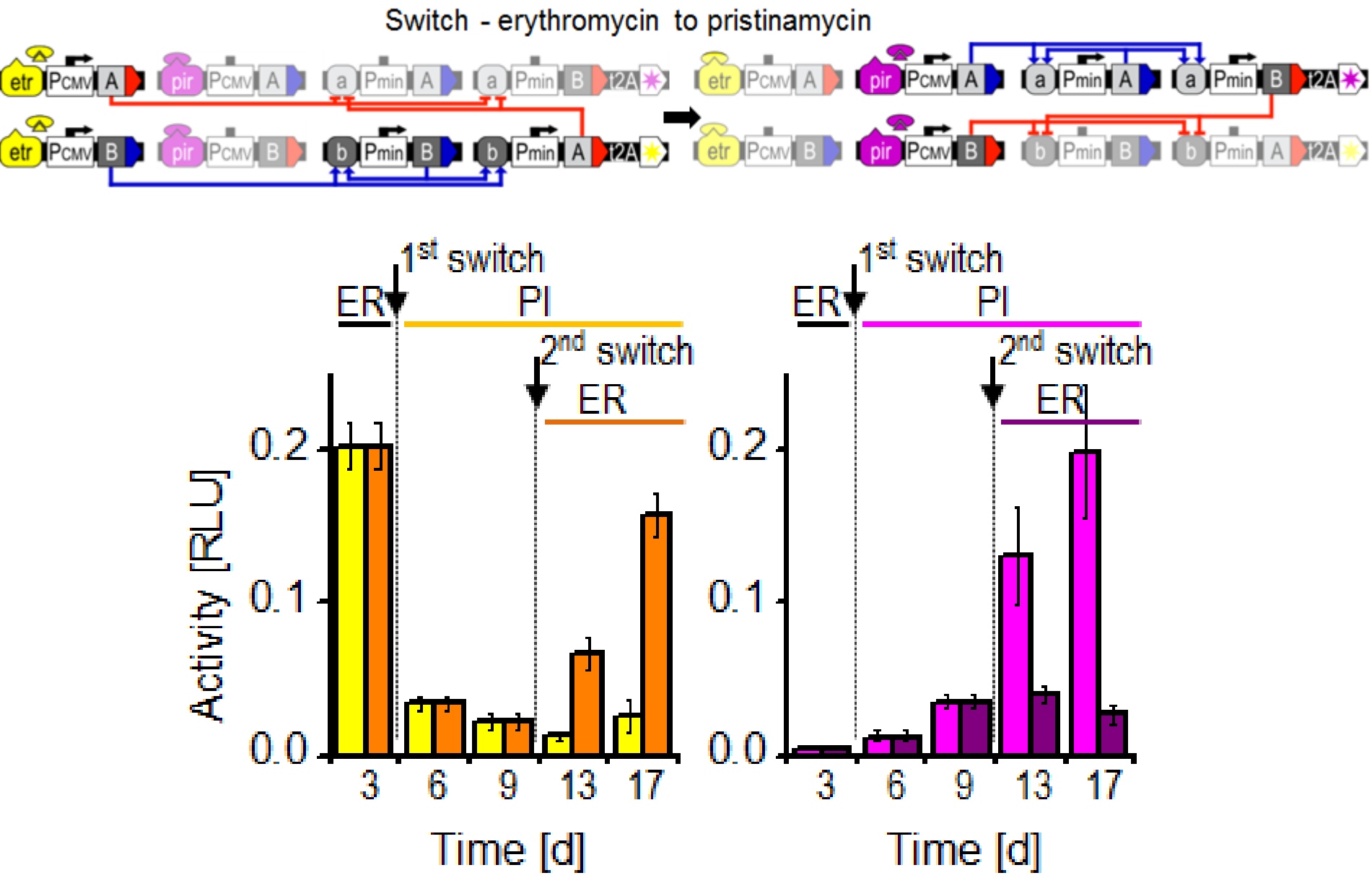








Switching between the 2 states



Designable orthogonal cellular logic

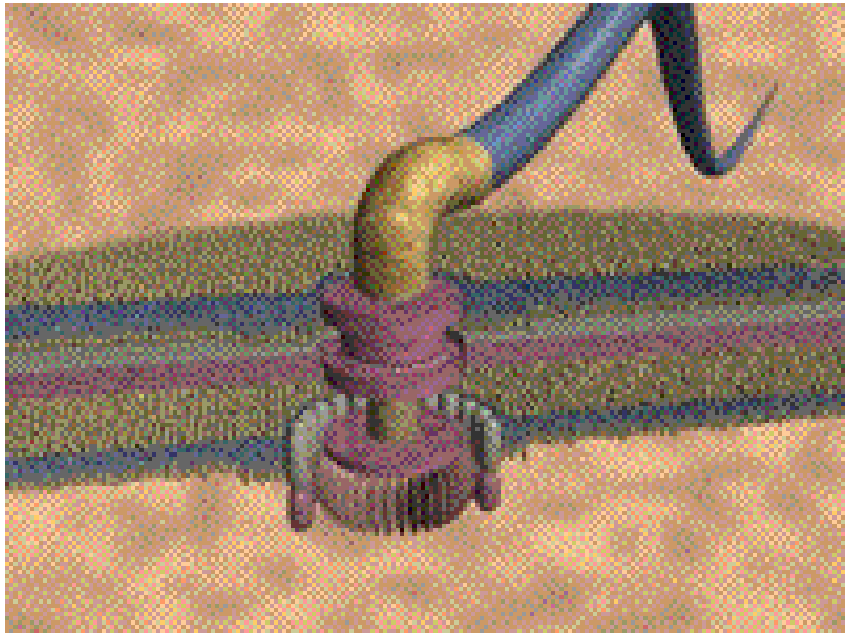
•Designable DNA binding domains can implement orthogonal logic functions in mammalian cells

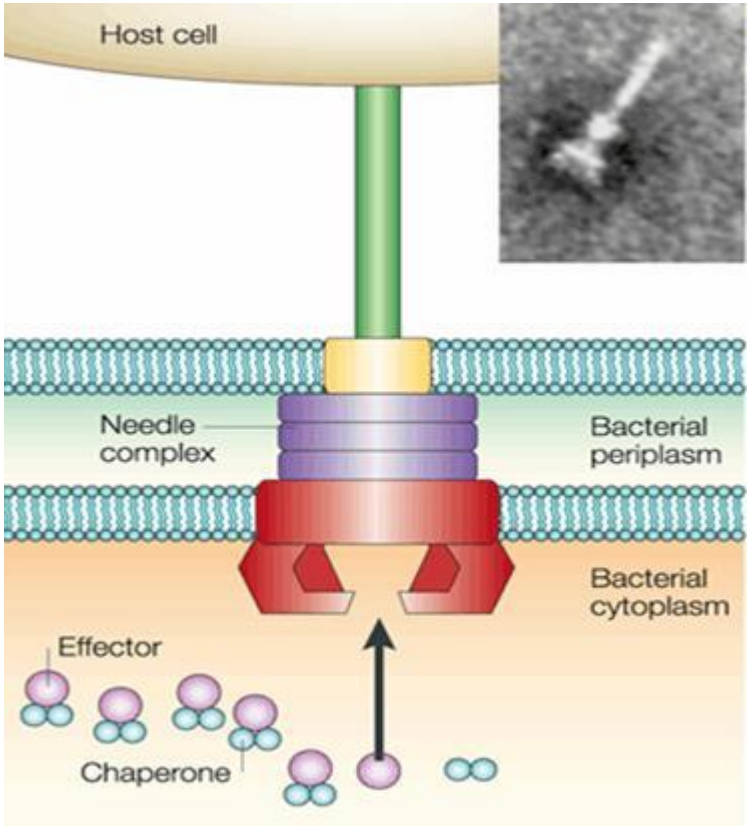
•Layered structure of functionaly complete NOR gate allows construction of complex logic

•Competition for the binding site and positive feedback loop can introduce nonlinarity required for the construction of dynamic logic structures

•Designable DNA bindig domains reprent scalable digital memory elements

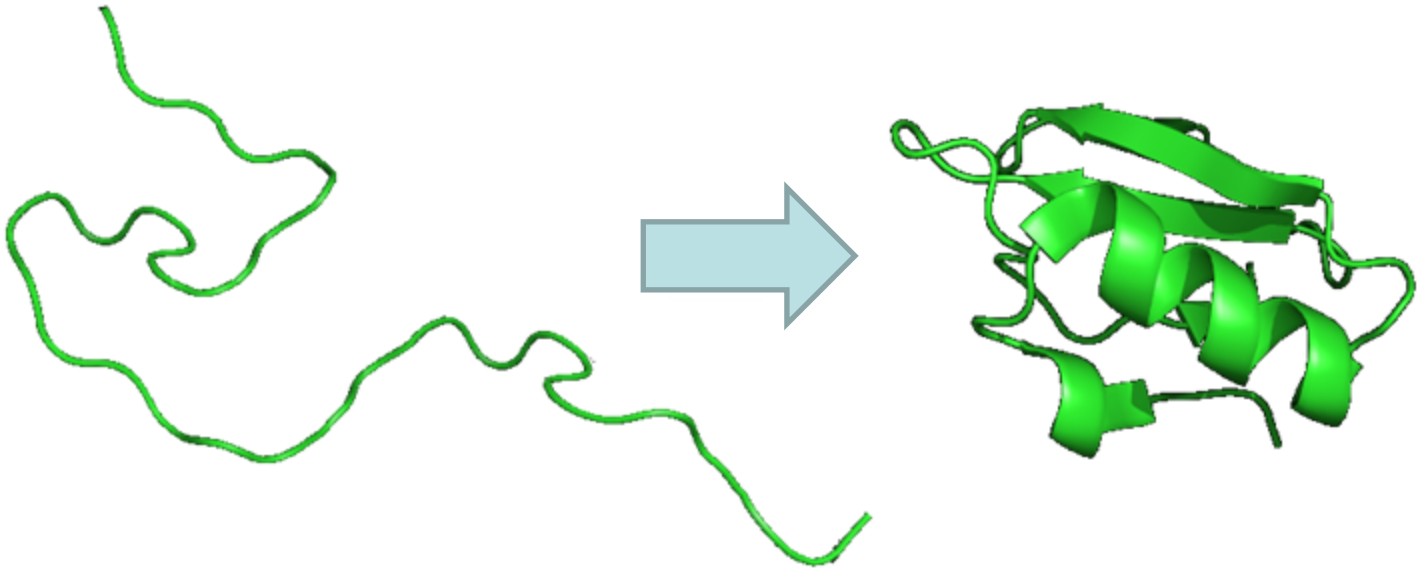
Natural molecular machines



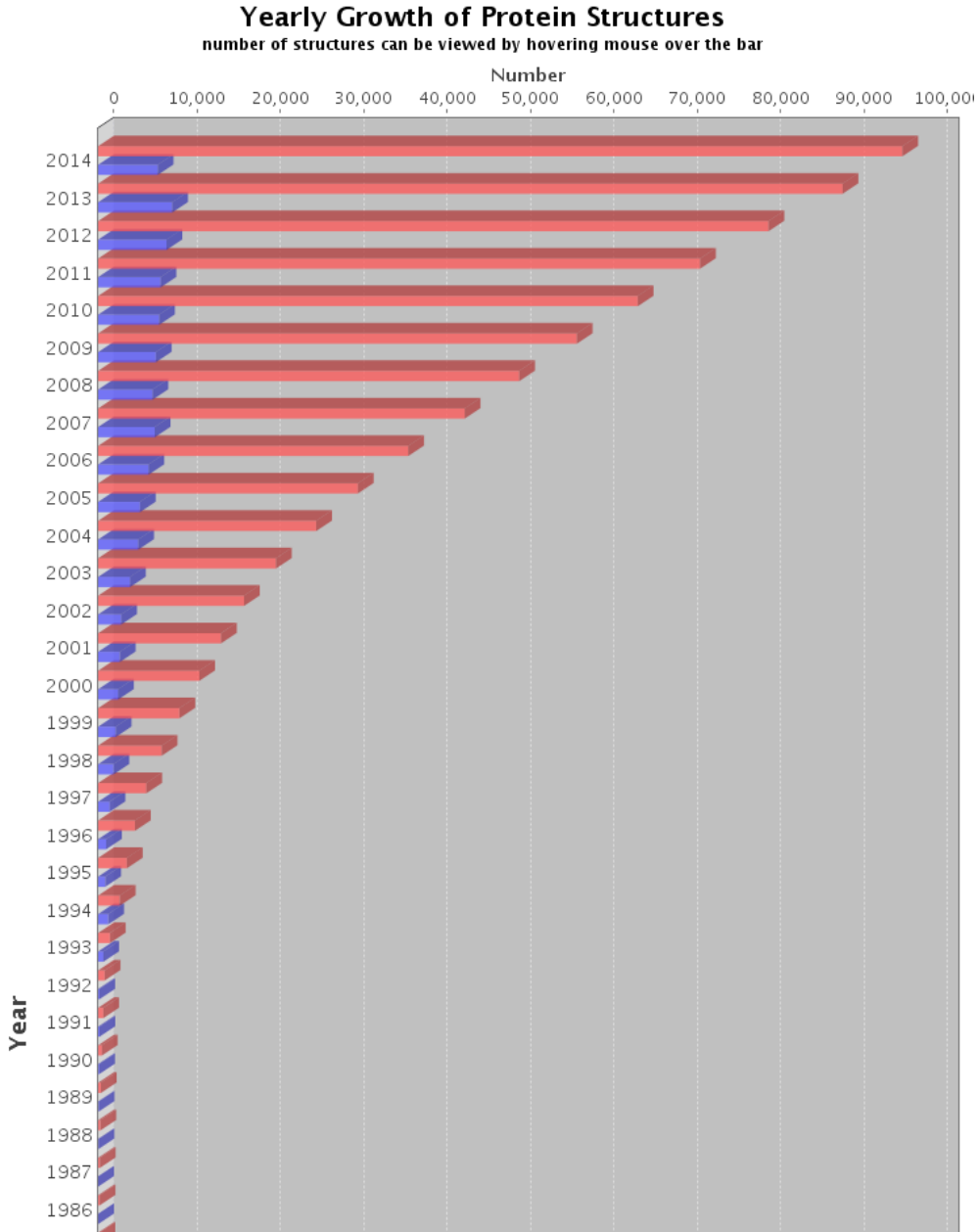


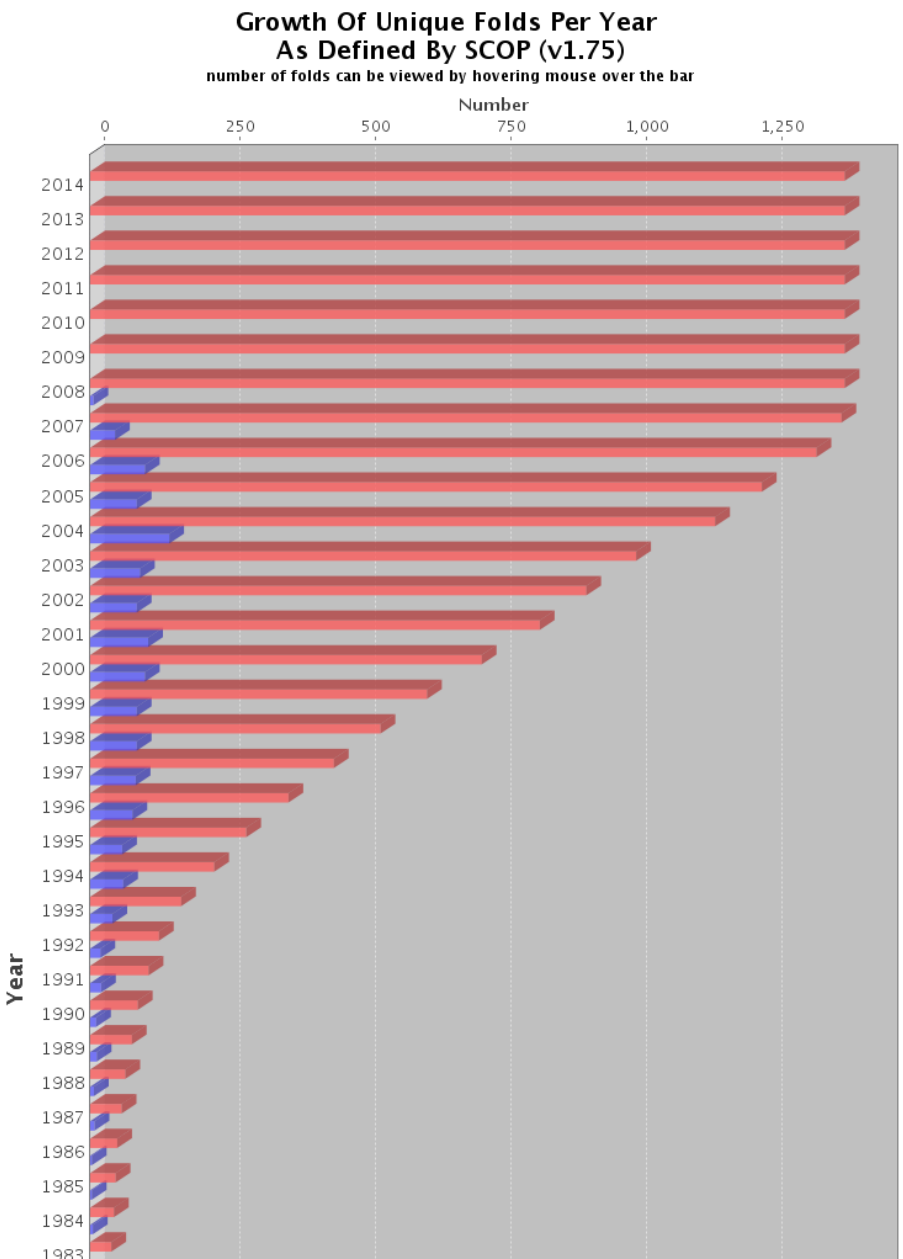
http://www.arn.org/mm/mm.htm

Natural protein origami

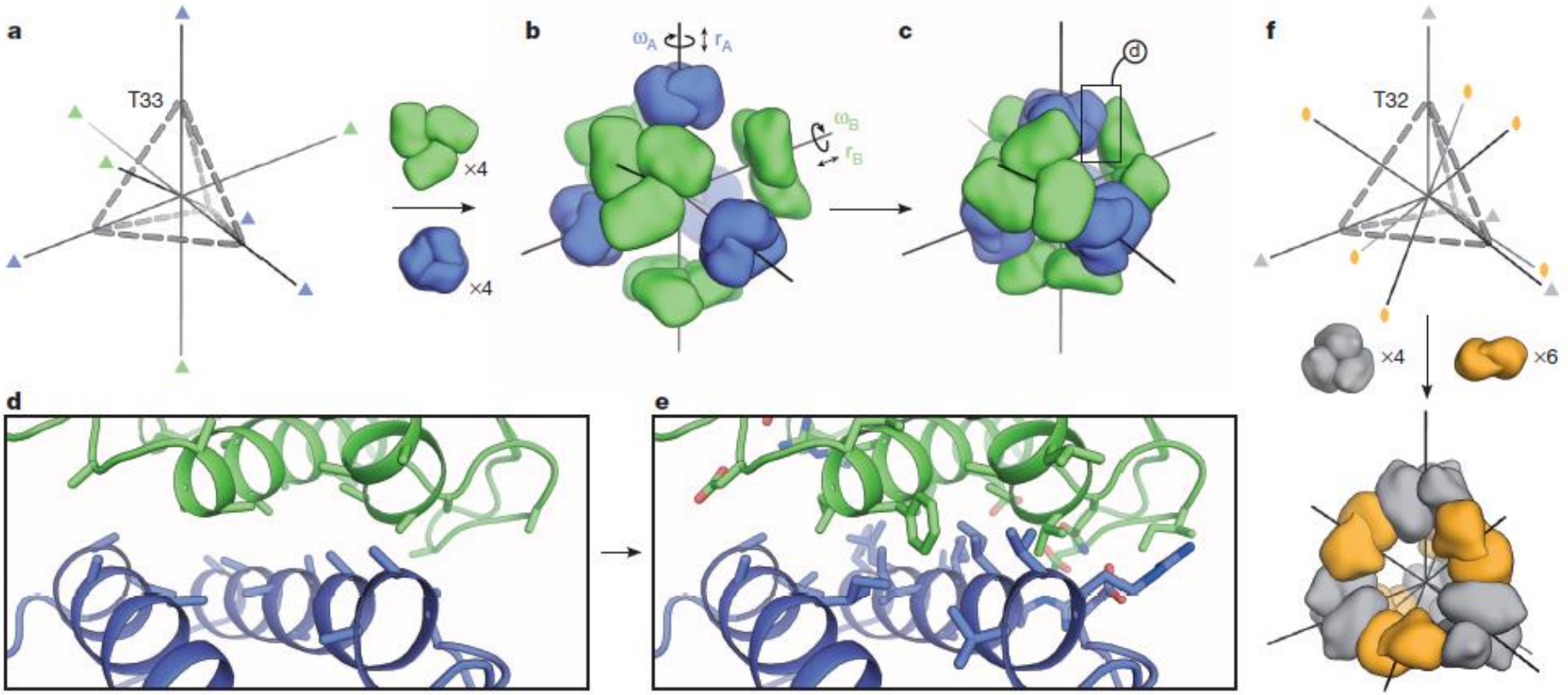


Natural protein folds



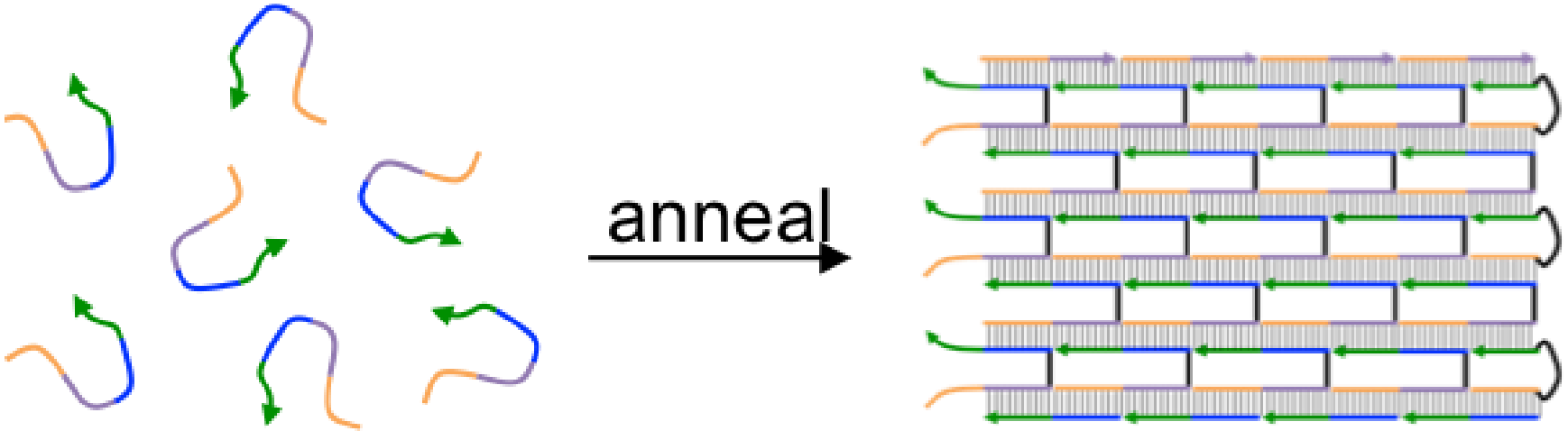


Designed protein domain assemblies

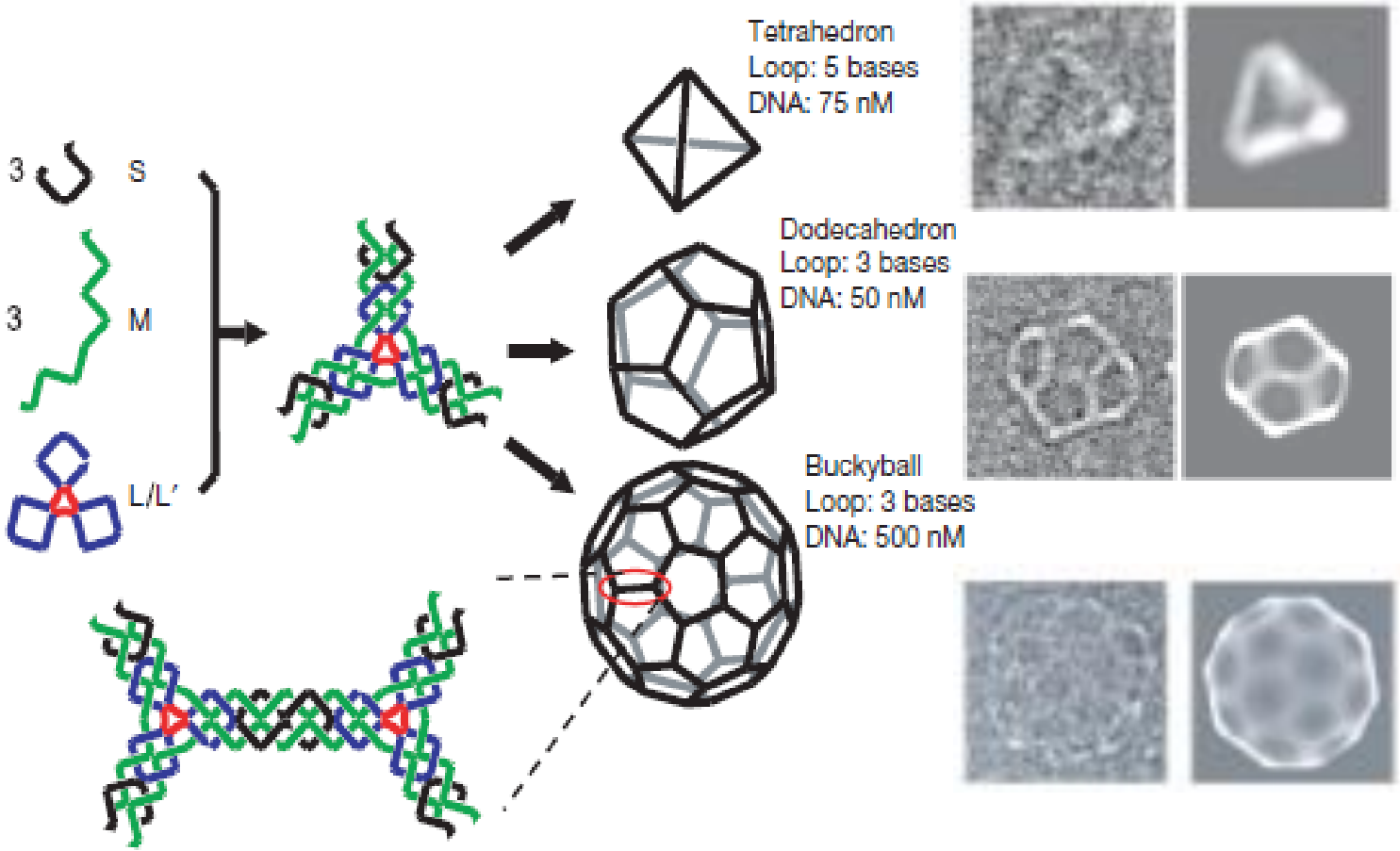


King et al., Science 2014

Long range modular interactions in designed DNA nanostructures

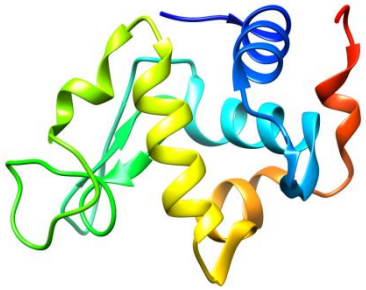
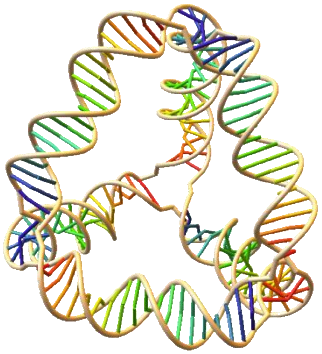
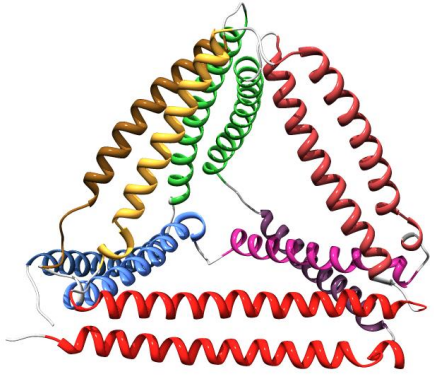


Designed DNA nanostructures

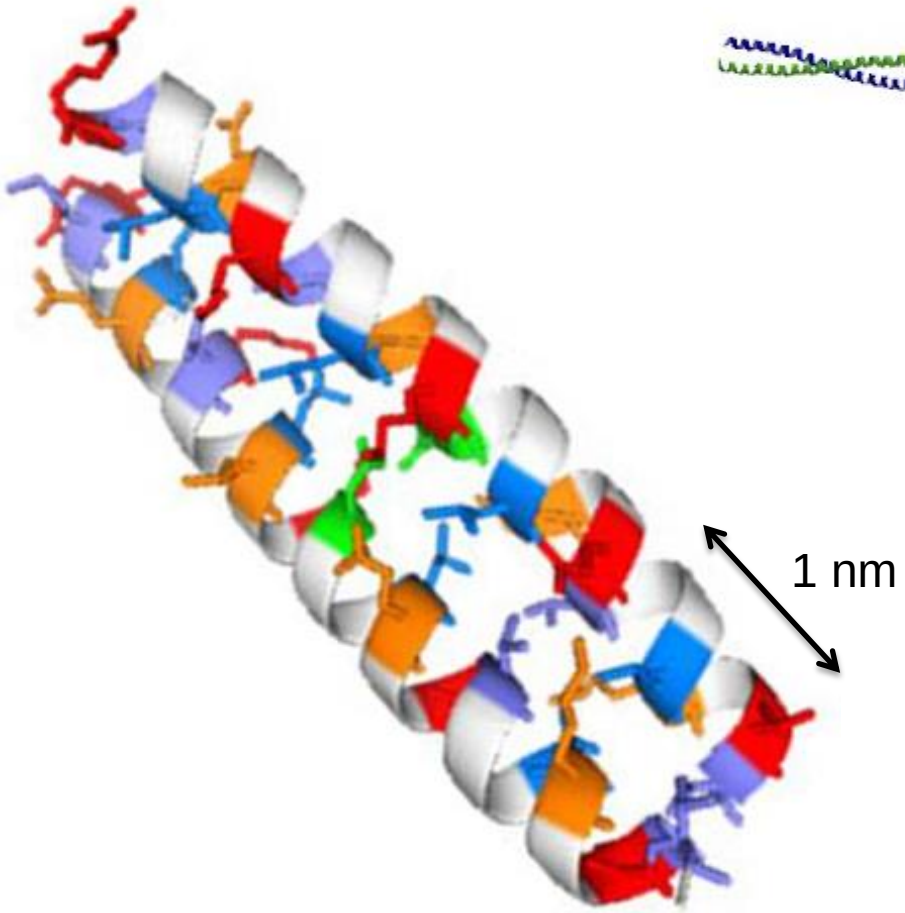


He et al., Nature, 2008

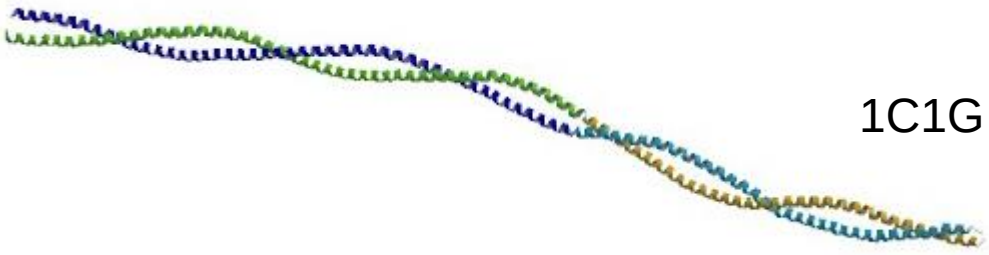
Evolved and designed bionanostructures

	DNA	Protein
Evolved compact fold		
Modular fold		

Coiled-coils as building blocks

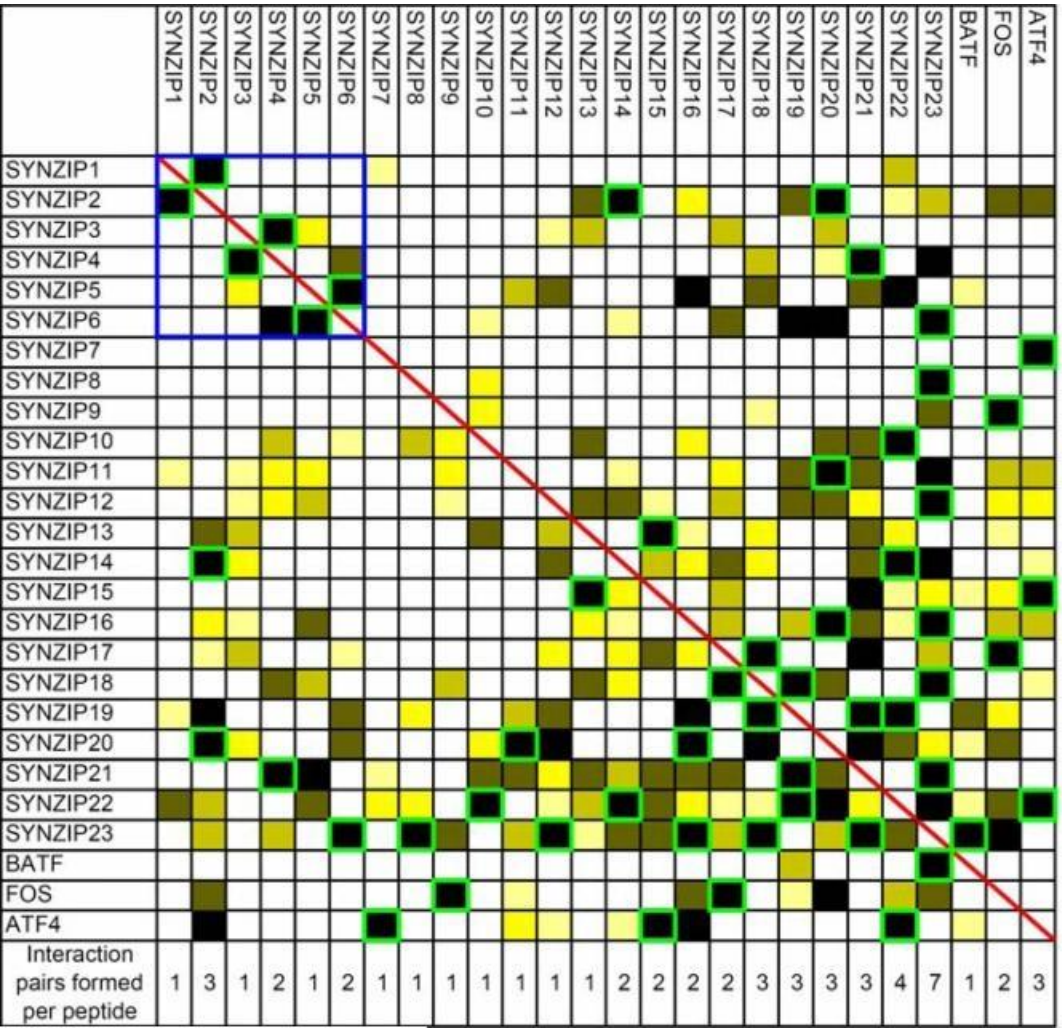


Length: 3 - > 50 nm (hundreds of residues)

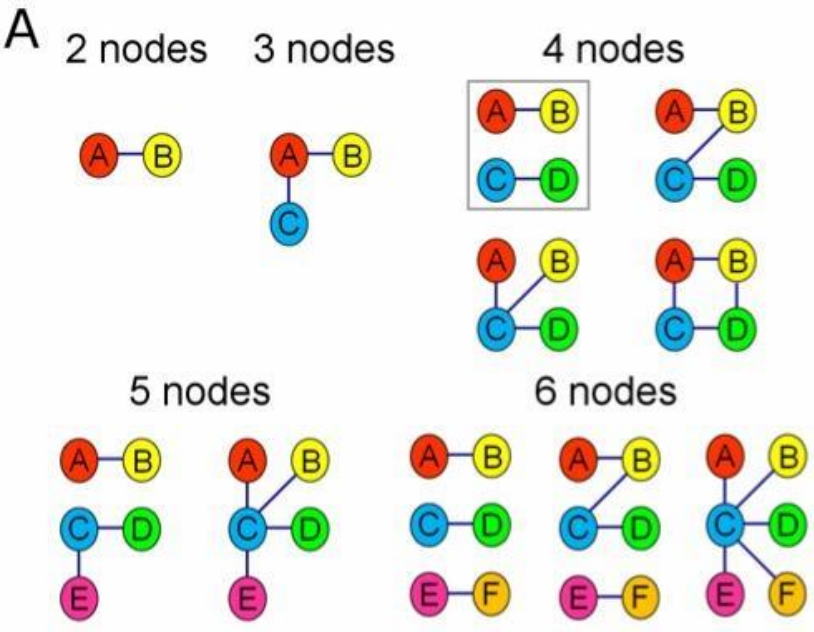


1C1G

Orthogonality of the native coiled-coil dimers



Analysis of pairwise interactions between 55 natural bZIP CC segments

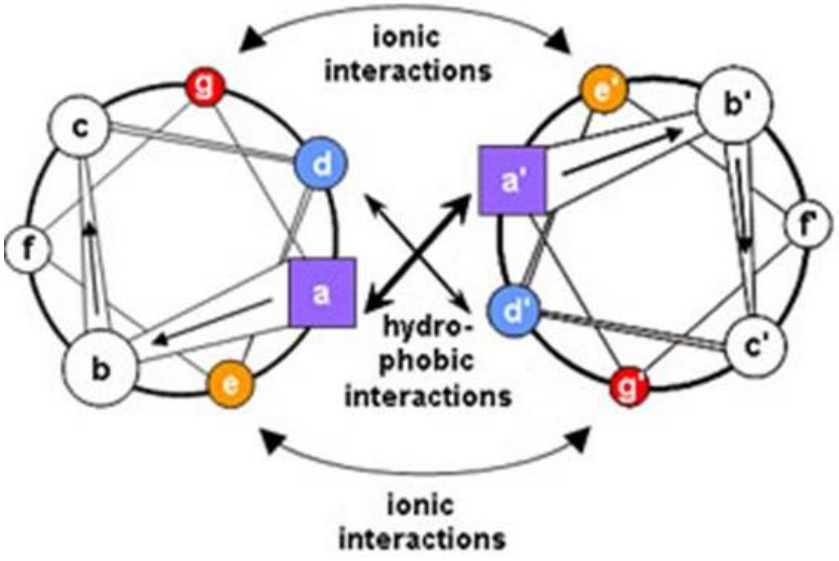


Reinke et al. (Keating lab), JACS 2011

Coiled-coil design rules

➢ We used the principles governing the selectivity and stability of CC segments to **design and experimentally test** a set of peptides

Heptad repeat-specific pattern



Stabilization

- hydrophobic residues at positions ***a*** and ***d***
- opposite charged residues at positions ***e*** and ***g***

Destabilization

- Negative design motif based on burial of polar Asn residues
- maximize the difference between designed (target) and unwanted combinations of residues

- positions ***b, c*** and ***f*** can be chemically modified to introduce desired function into the coiled-coil assembly

Design of orthogonal coiled-coil dimers

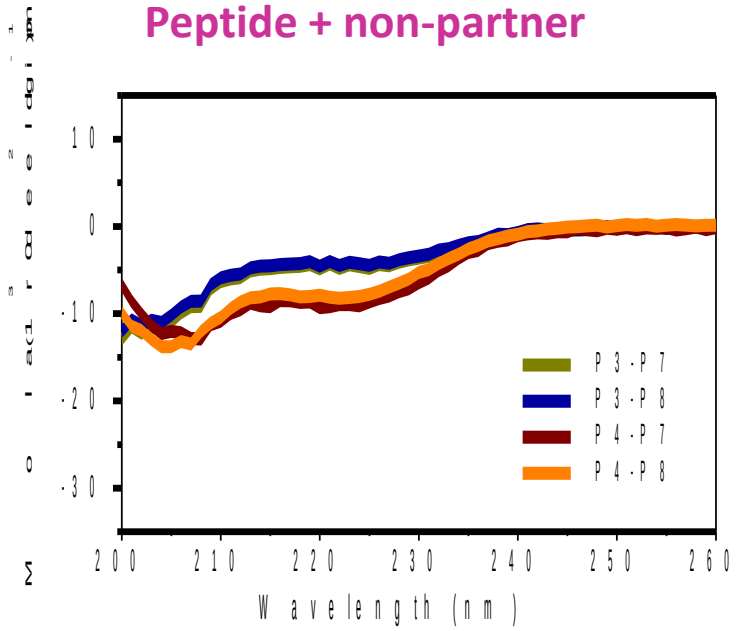
	Sequence[a]					Hydophobic pattern at positions **a**[b]	Electrostatic pattern of heptads[c]
	SPED	gabc**L**ef	gabc**L**ef	gabc**L**ef	gabc**L**e**Y**	**G**	
P1	SPED	**E**IQAL**E**E	**EN**AQL**E**Q	**EN**AAL**E**E	**E**IAQL**E**Y	G	I **N** **N** I	EEEE
P2	SPED	**K**IAQL**K**E	**KN**AAL**K**E	**KN**QQL**K**E	**K**IQAL**K**Y	G	I **N** **N** I	KKKK
P3	SPED	**E**IQQL**E**E	**E**IAQL**E**Q	**KN**AAL**K**E	**KN**QAL**K**Y	G	I I **N** **N**	EEKK
P4	SPED	**K**IAQL**K**Q	**K**IQAL**K**Q	**EN**QQL**E**E	**EN**AAL**E**Y	G	I I **N** **N**	KKEE
P5	SPED	**EN**AAL**E**E	**K**IAQL**K**Q	**KN**AAL**K**E	**E**IQAL**E**Y	G	**N** I **N** I	EKKE
P6	SPED	**KN**AAL**K**E	**E**IQAL**E**E	**EN**QAL**E**E	**K**IAQL**K**Y	G	**N** I **N** I	KEEK
P7	SPED	**E**IQAL**E**E	**KN**AQL**K**Q	**E**IAAL**E**E	**KN**QAL**K**Y	G	I **N** I **N**	EKEK
P8	SPED	**K**IAQL**K**E	**EN**QQL**E**Q	**K**IQAL**K**E	**EN**AAL**E**Y	G	I **N** I **N**	KEKE

Parallel	P1	P2	P3	P4	P5	P6	P7	P8
P1	33	**100**	29	27	31	32	30	29
P2	–	−6	10	7	11	12	11	9
P3	–	–	10	**93**	19	20	19	17
P4	–	–	–	5	17	18	17	15
P5	–	–	–	–	13	**101**	−15	−16
P6	–	–	–	–	–	16	−13	−15
P7	–	–	–	–	–	–	12	**96**
P8	–	–	–	–	–	–	–	9

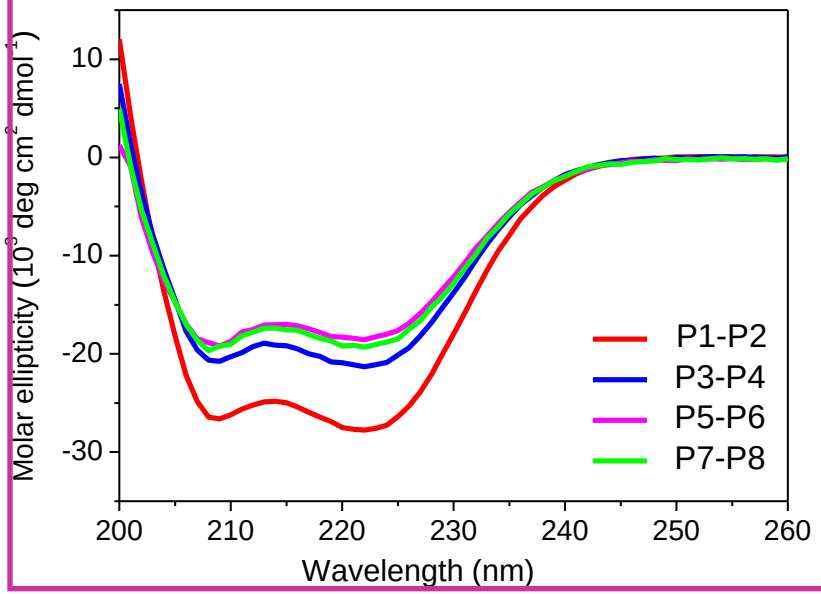
Antiparallel	P1	P2	P3	P4	P5	P6	P7	P8
P1	−62	5	−30	−33	−28	−27	−29	−30
P2		−100	−49	−52	−47	−46	−48	−49
P3			1	−87	−40	−38	−41	−42
P4				−3	−42	−41	−43	−44
P5					−81	7	−39	−40
P6						−78	−37	−38
P7							3	−84
P8								1

Orthogonality of designed coiled-coil peptides

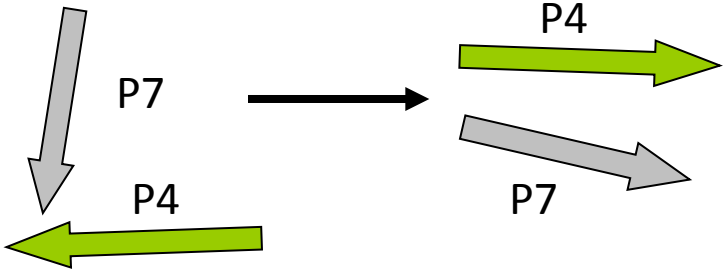
Peptide + non-partner



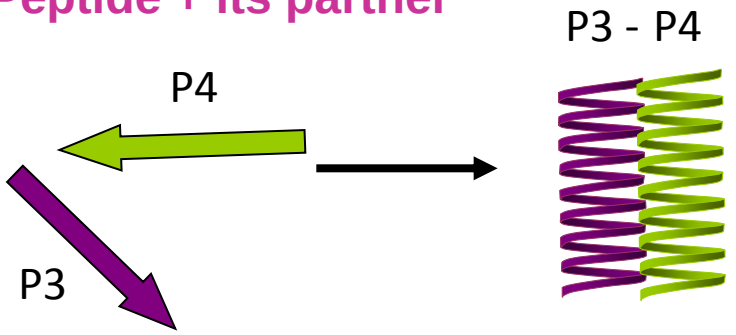
Peptide + designed partner



Peptide + non-partner

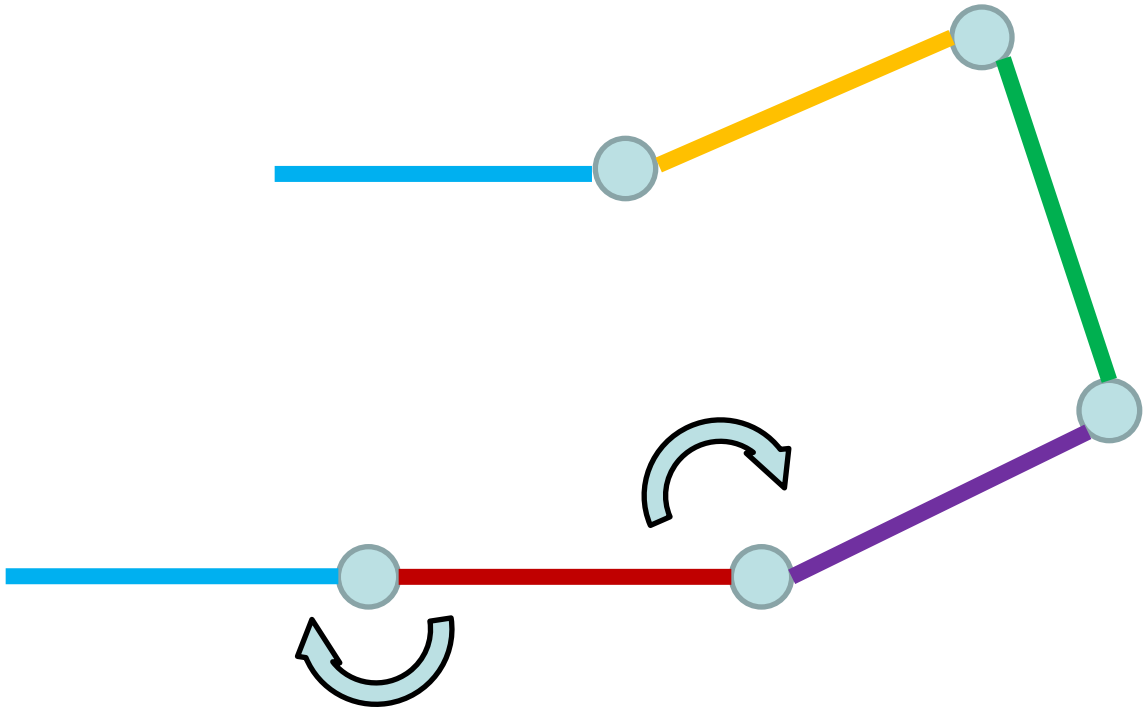


Peptide + its partner

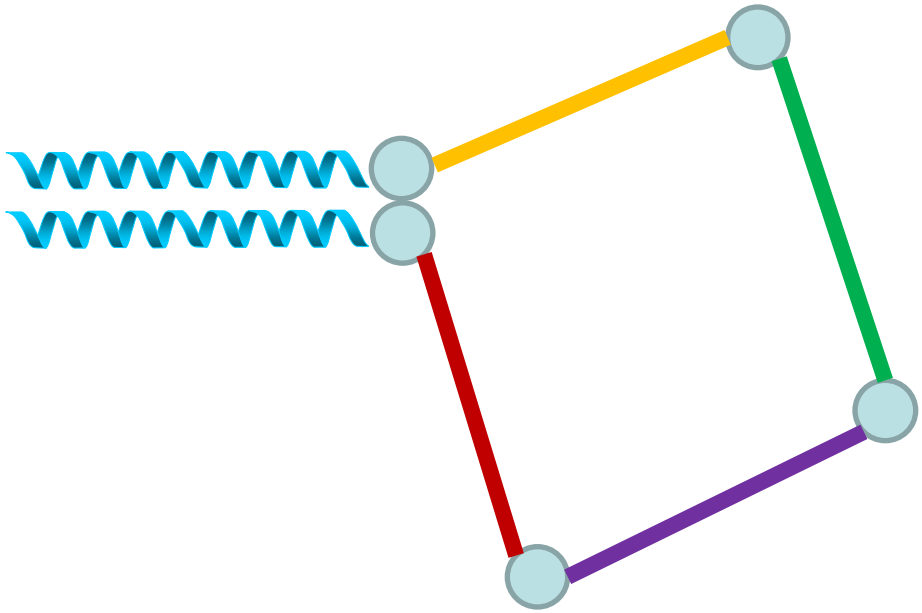


Gradišar and Jerala, J.Pept.Sc., 2011

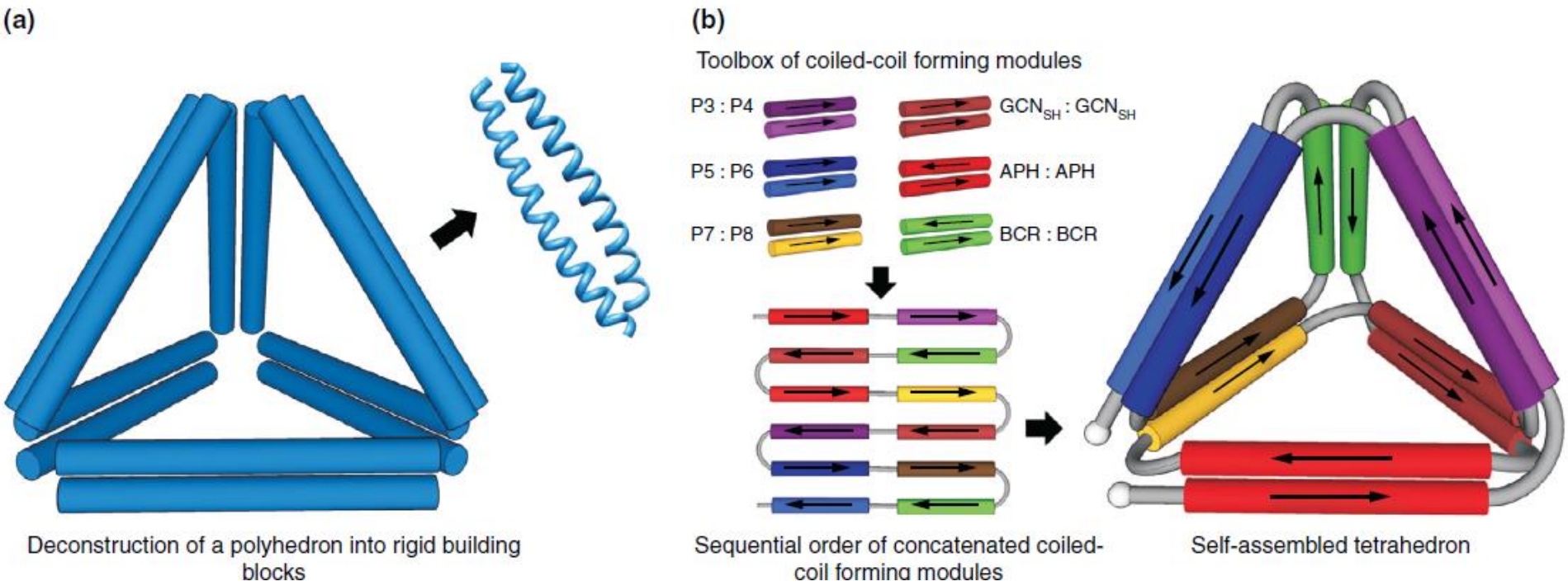
Flexible linker connecting interacting elements



Flexible linker connecting interacting elements

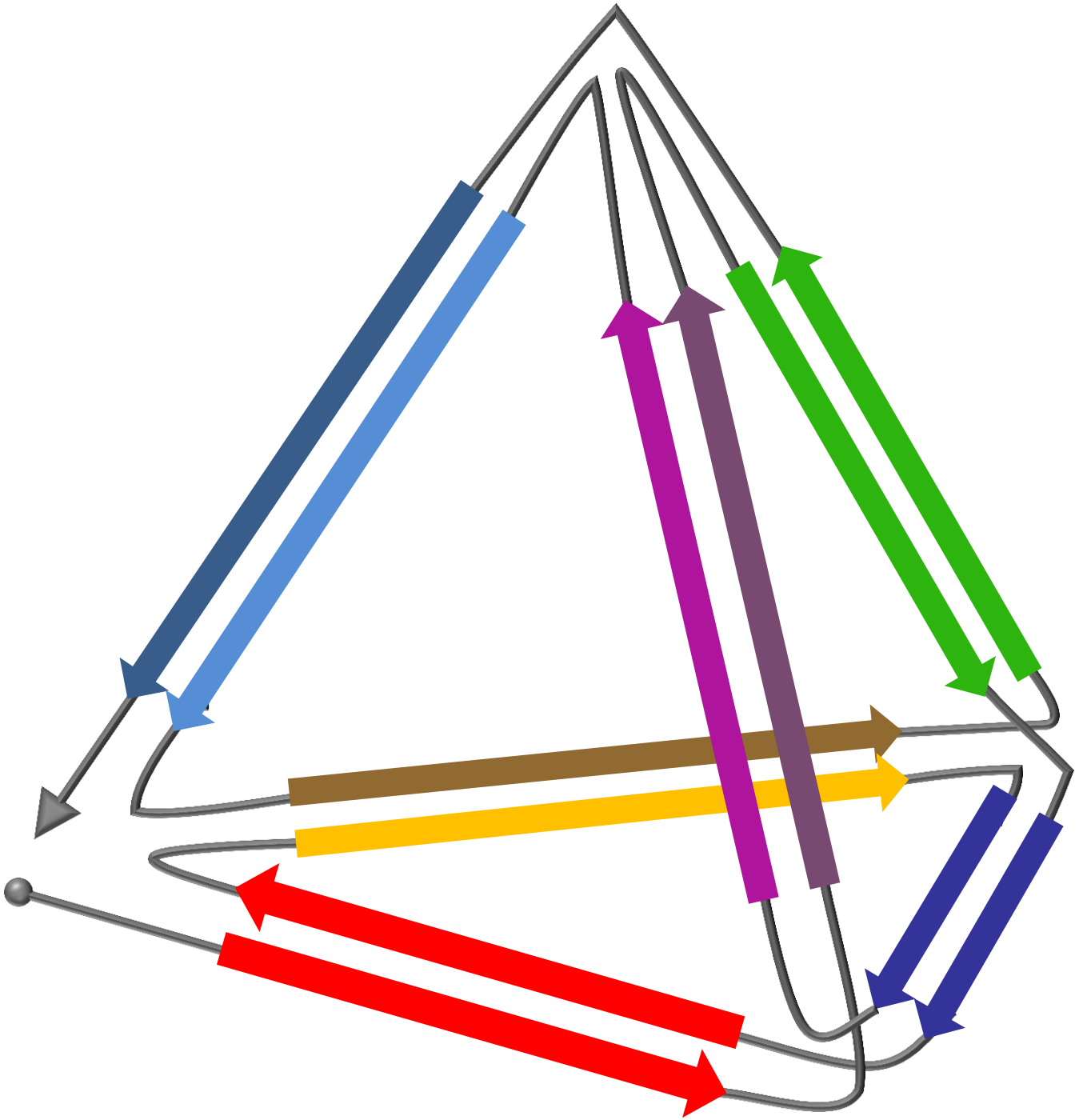


Deconstructing shape into modules

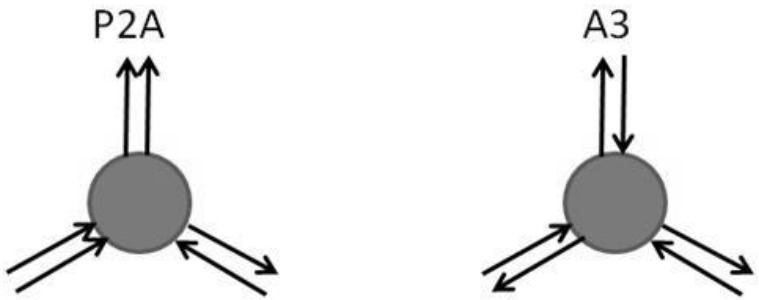


Can the tetrahedral edges be traversed exactly twice, forming coiled-coils at each edge ?

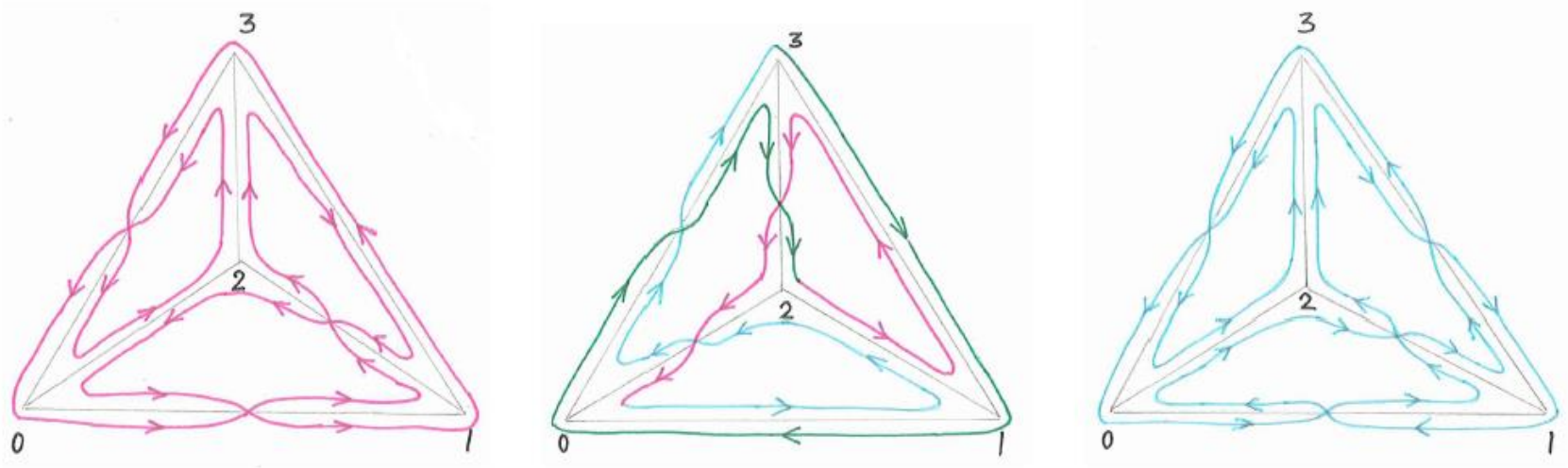
Božič-Abram et al., Cur.Op.Chem.Biol. 2013



Topological solutions for a tetrahedron

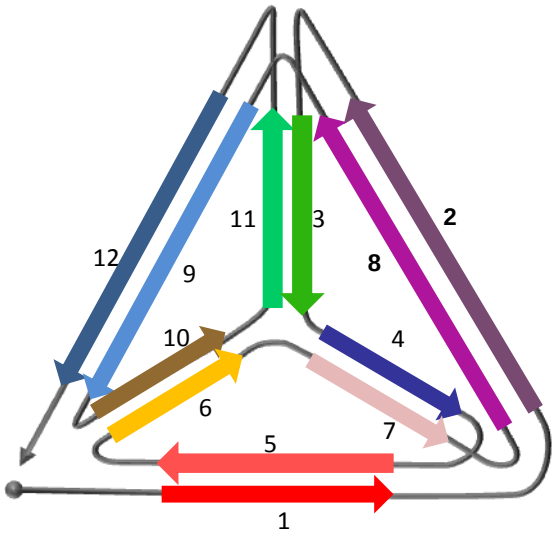


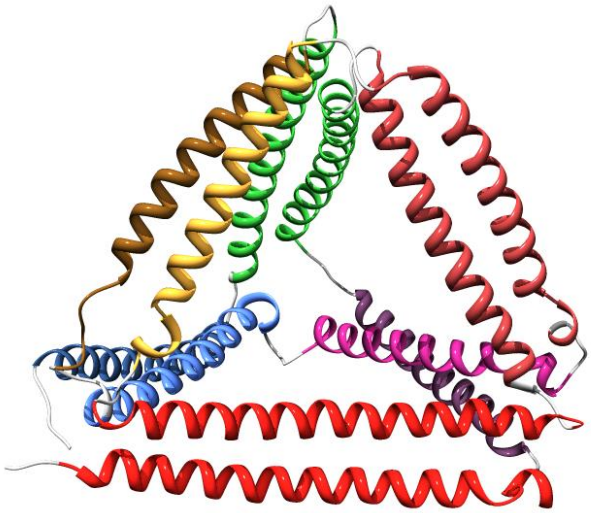
Two possible types of vertices of the degree of 3 (6).

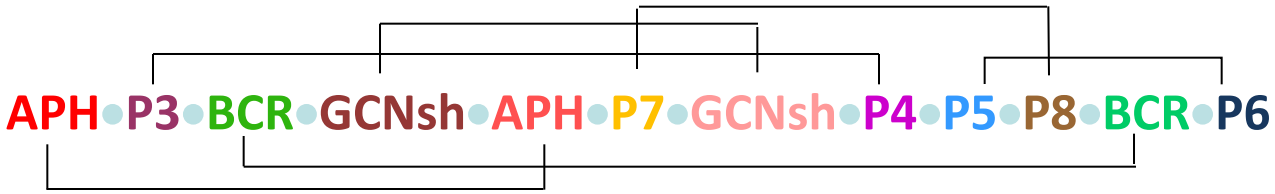


Three possible topologies to construct a tetrahedron but could be realized by 28 different combinations of segments.

Design of a tetrahedron-forming polypeptide



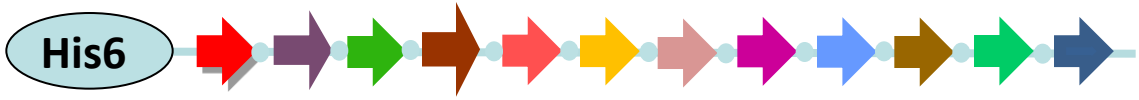




4 parallel dimers

2 antiparallel dimers

TET12

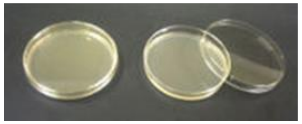


- flexible tetrapeptide linker

SGPG

Polypeptide production, isolation and self-assembly

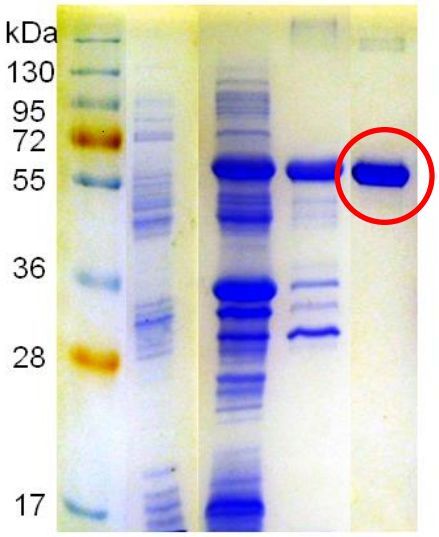
Production in *E.coli*



Protein purification

Affinity chromatography
HPLC-RP

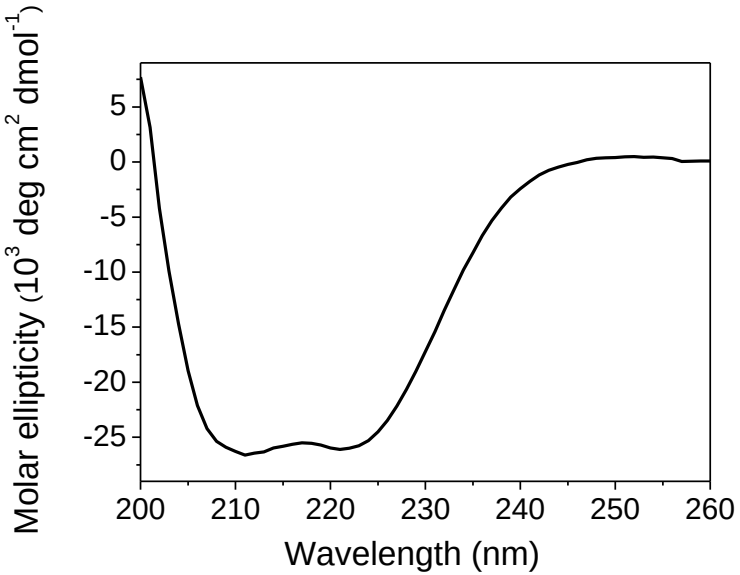
SDS-PAGE



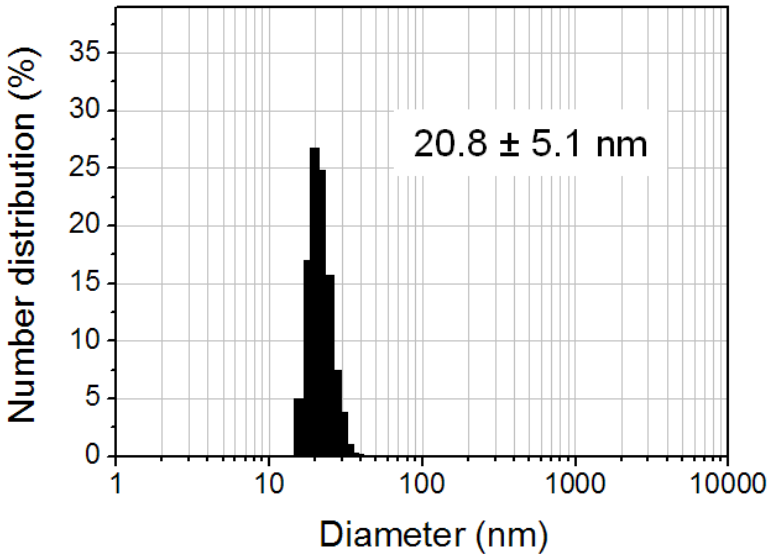
1. Standard
2. Supernatant of cell lysate
3. Inclusion bodies
4. TET12 purified on NiNTA
5. TET12 purified on RP-HPLC

In vitro self-assembly

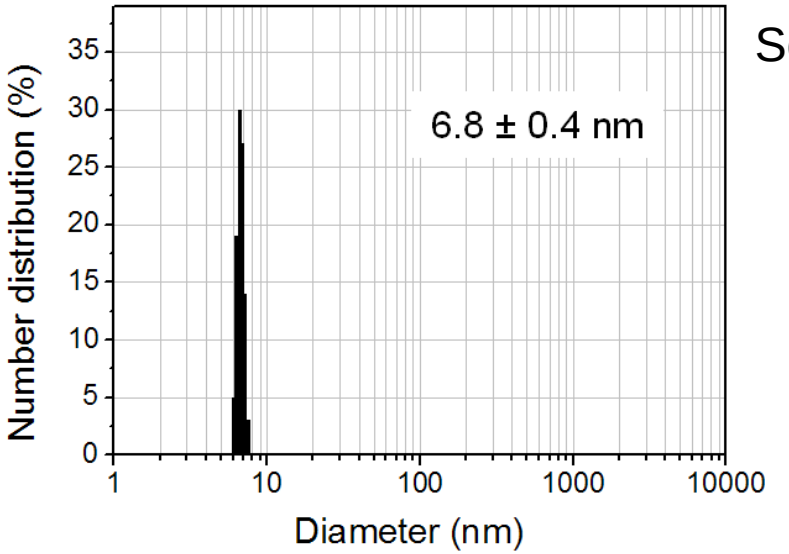
- **Dialysis** at low polypeptide concentration



Characterization of hydrodynamic size by DLS

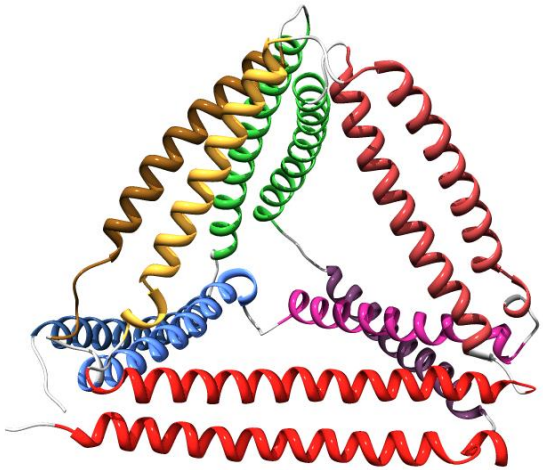


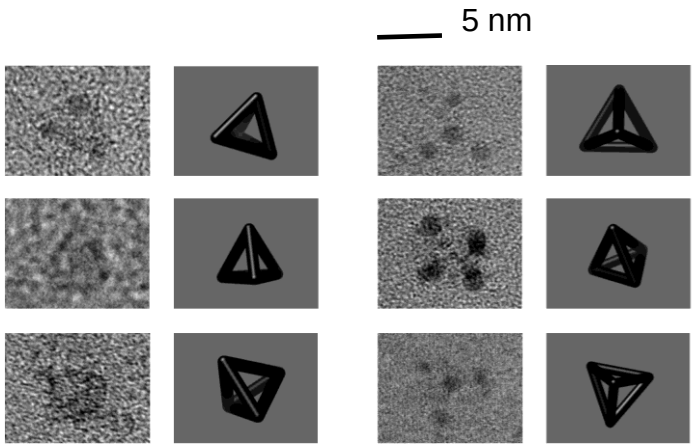
Denatured TET12

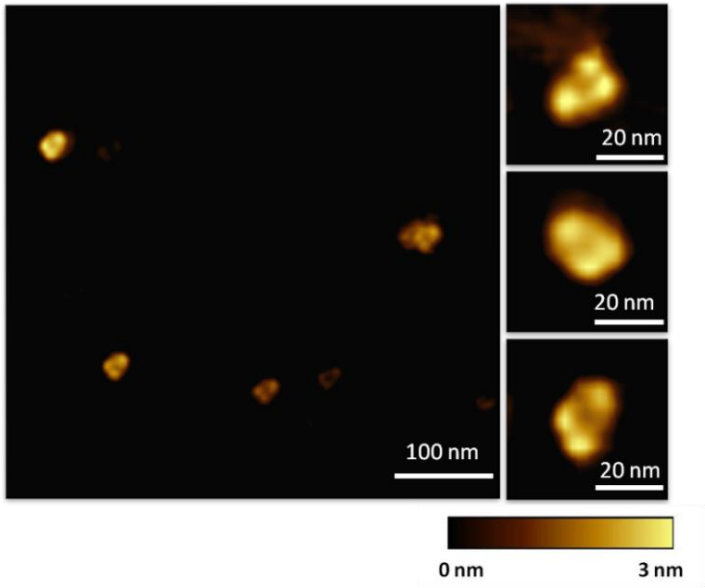


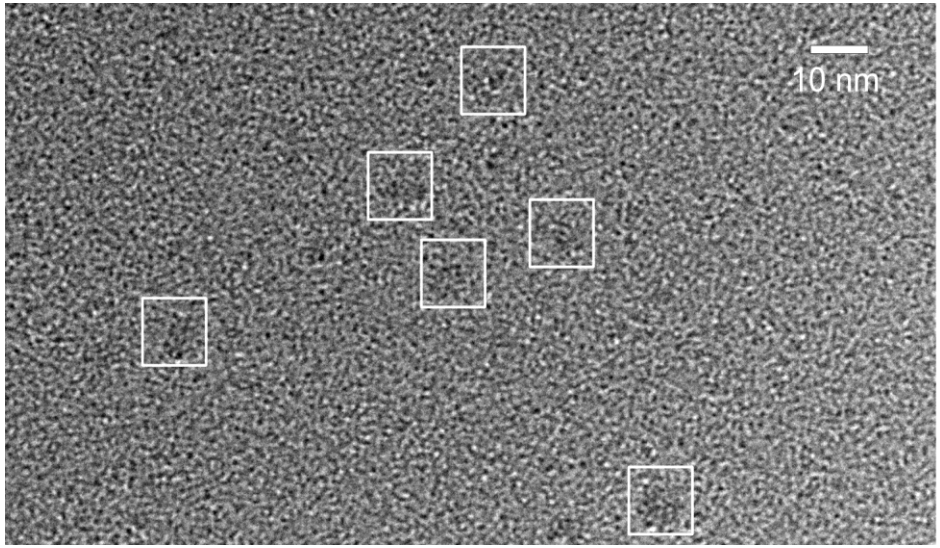
Self-assembled TET12

TEM and AFM imaging



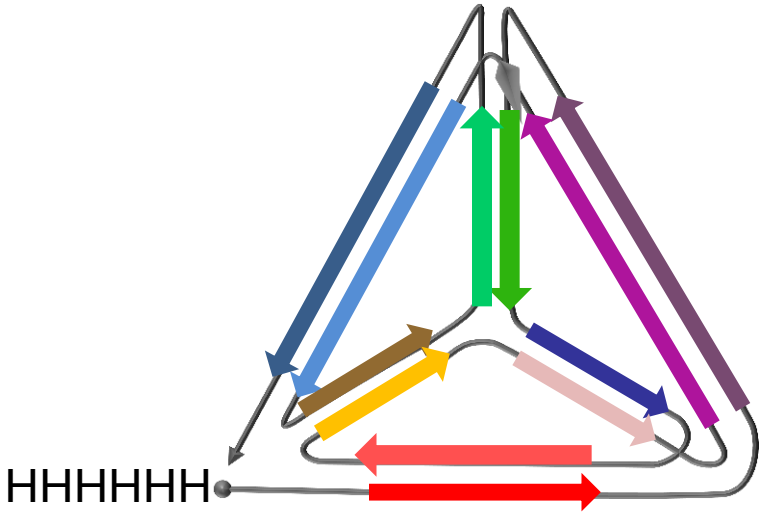


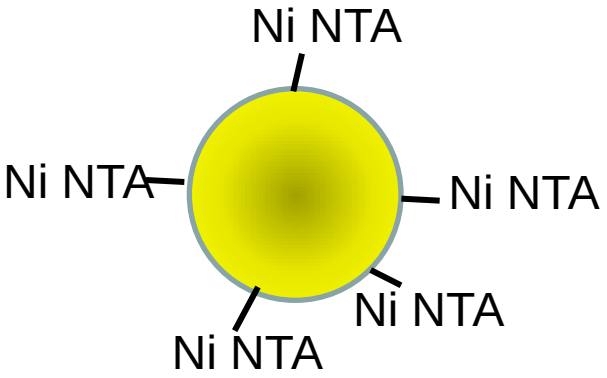


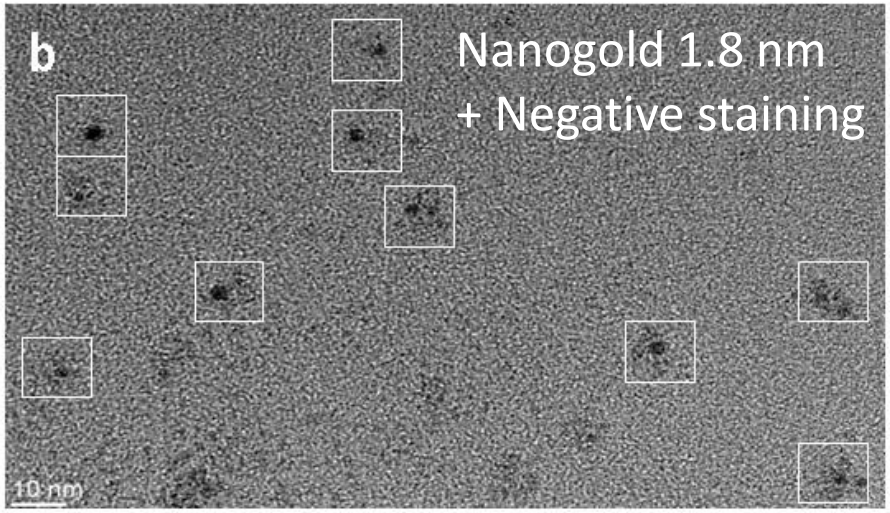


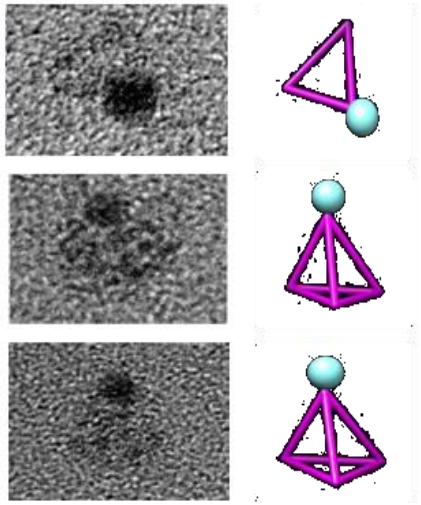
Gradišar et al., Nature Chem. Biol. 2013

Detection of the N-terminal end of TET12

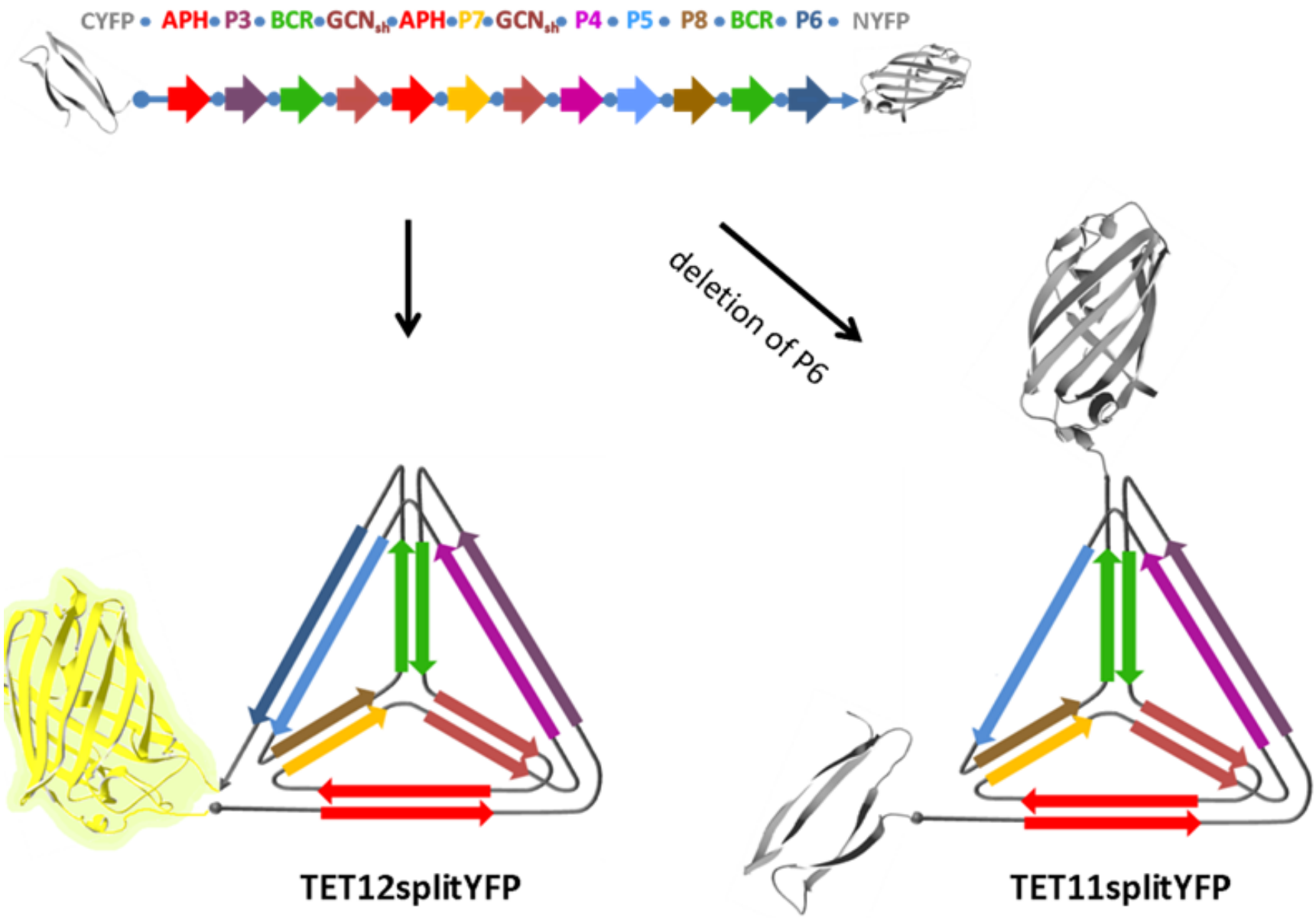




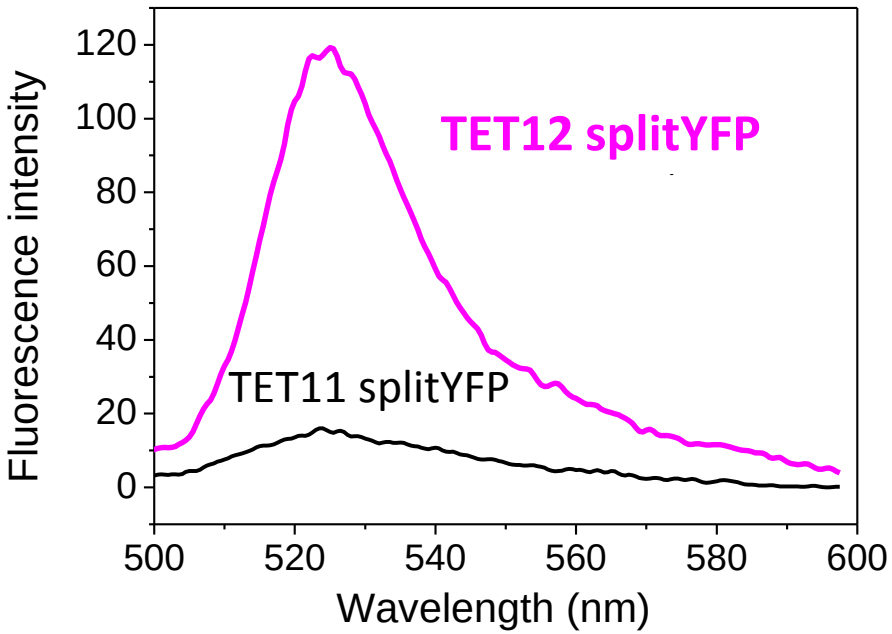




Termini of the tetrahedral path coincide



Coincidence of termini by YFP reconstitution

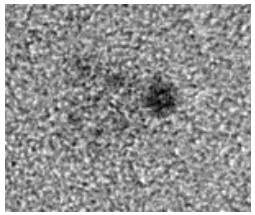


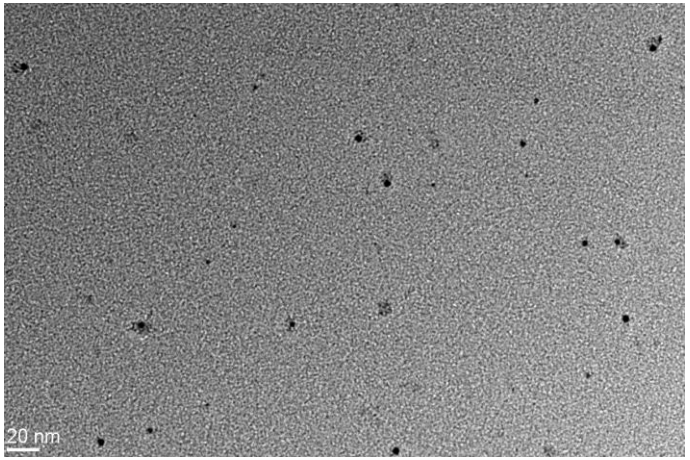
In vitro reconstitution

No fluorescence in producing bacteria

➢Fluorescence is reconstituted only in TET12 splitYFP but not for TET11 splitYFP

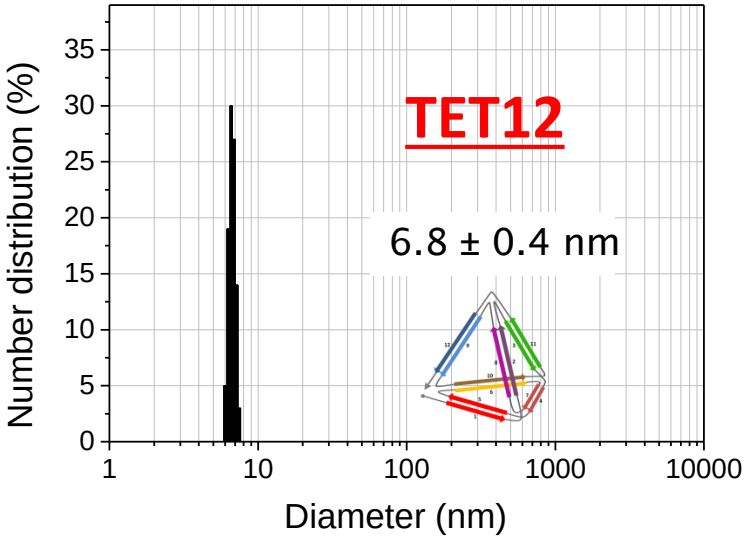
TET12 splitYFP

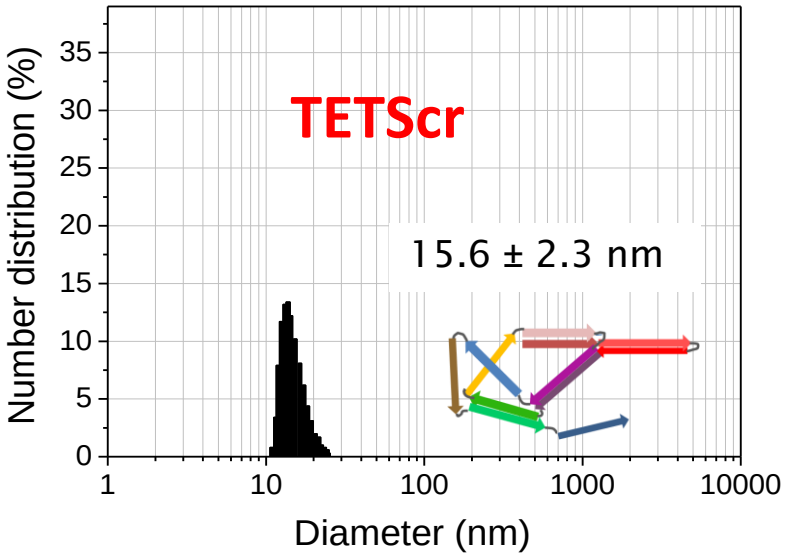


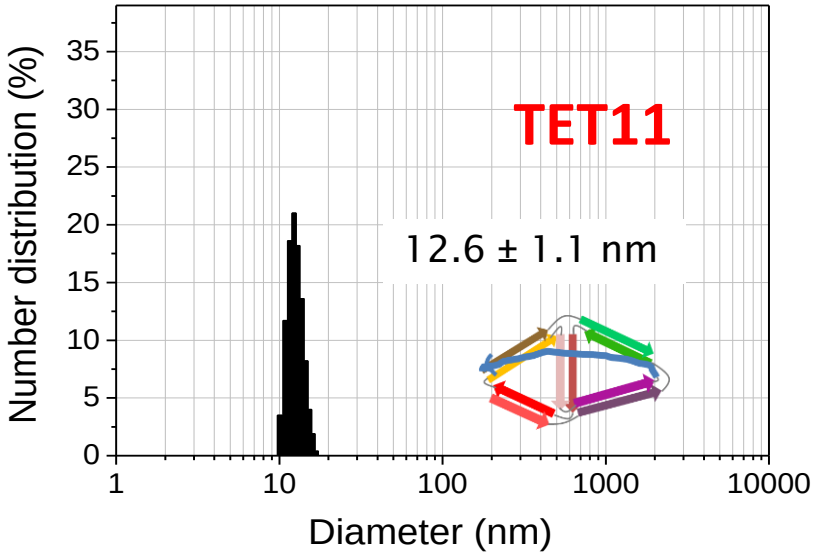


Gradišar et al., Nature Chem. Biol. 2013

Correct order of segments defines the structure

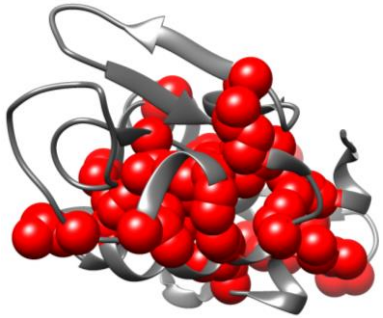






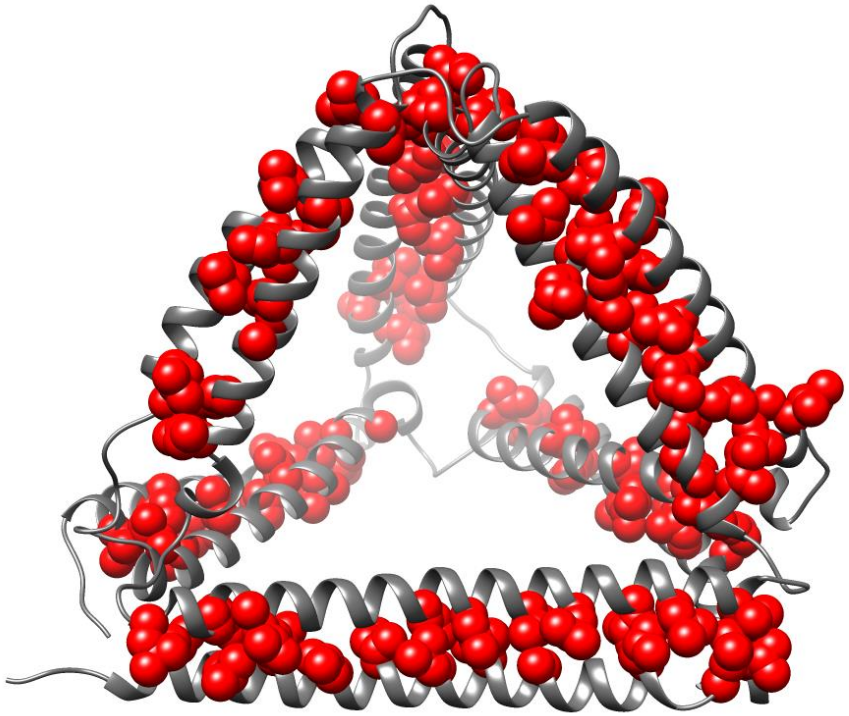
Natural and topological protein fold

NATIVE PROTEIN FOLD



Compact and continuous hydrophobic core joining secondary structure elements

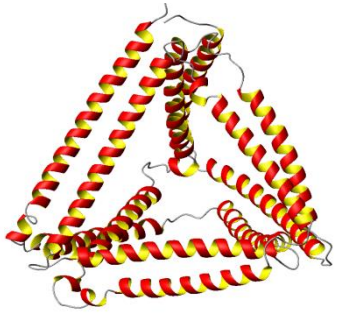
TOPOLOGICAL PROTEIN FOLD

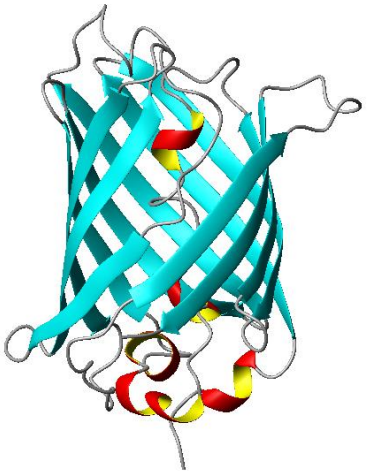


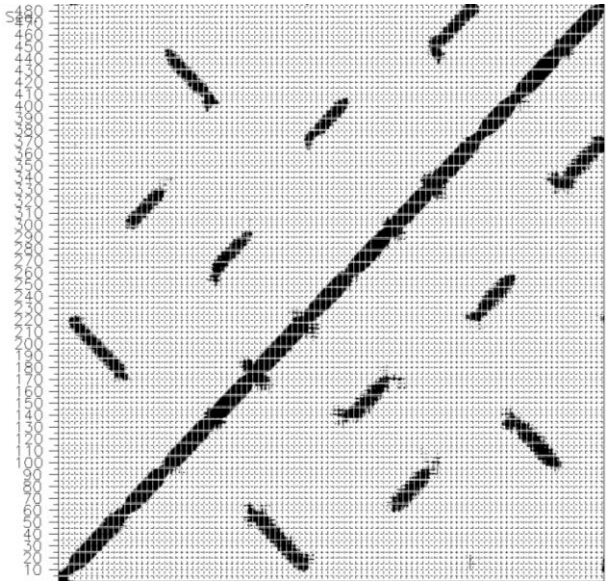
Hydrophobic core limited to within each building block

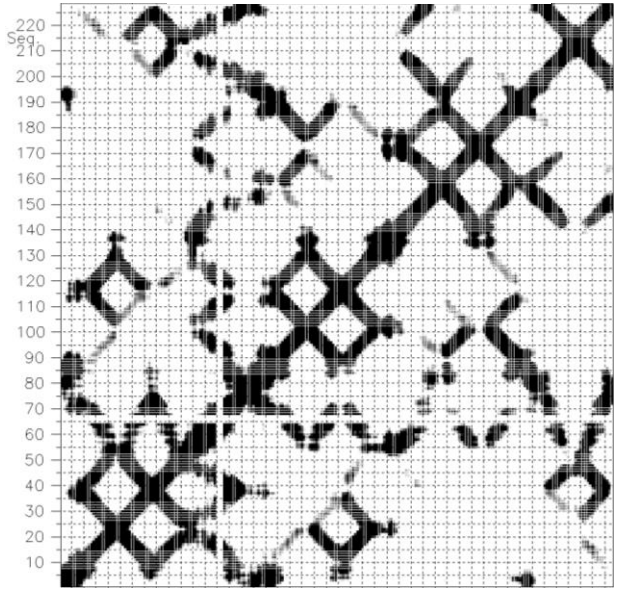
Topology defines the fold !

Fold definition by long-range interactions





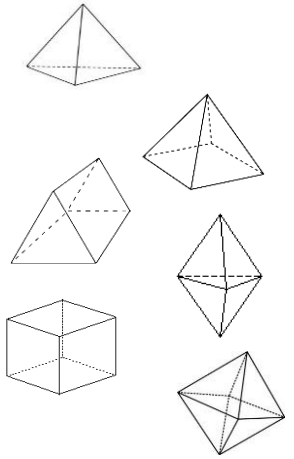




Challenges in the design of modular proteins

- Increasing the complexity of topological folds

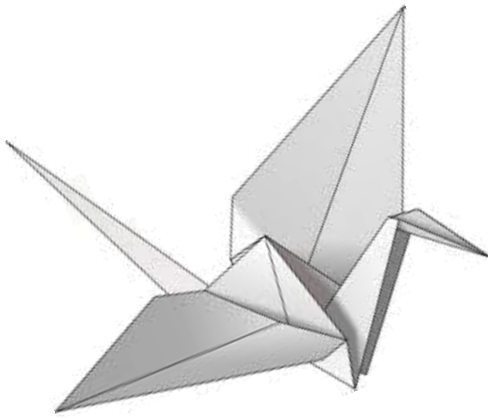
Polyhedron	number of edges	topologies	antiparallel only	parallel only
triangular pyramid (tetrahedron)	6	3	0	0
square pyramid	8	82	5	0
triangular bipyramid	9	470	0	0
triangular prism	9	25	2	0
square prism (cube)	12	40	0	0
square bipyramid (octahedron)	12	22246	0	275



- functional modular protein
- In vivo folding

Summary

- Concatenated coiled-coil-based modules can be used to design new type of a **topological protein fold** based on similar principles as DNA nanostructures
- Orthogonal and topology encoded long-range interactions can define complex nanoscale protein shapes
- Modularity of biopolymers can be used to design folding pathways
- Topological proteins can be designed to fold in vivo



Acknowledgements

Tina Lebar
Rok Gaber
Helena Gradišar
Iva Hafner Bratkovič
Vid Kočar
Sabina Božič
Tibor Doles
Tomaž Pisanski
Nino Bašić
Sandi Klavžar

+members of Slovenian iGEM teams 2009, 2012:
Urban Bezeljak, Boštjan Pirš, Anja Golob, Miha Jerala, Martin Stražar, Uroš Zupančič, Dušan Vučko