EMBO PPI Training Course
Budapest, 30-5-2016 – 4-6-2016

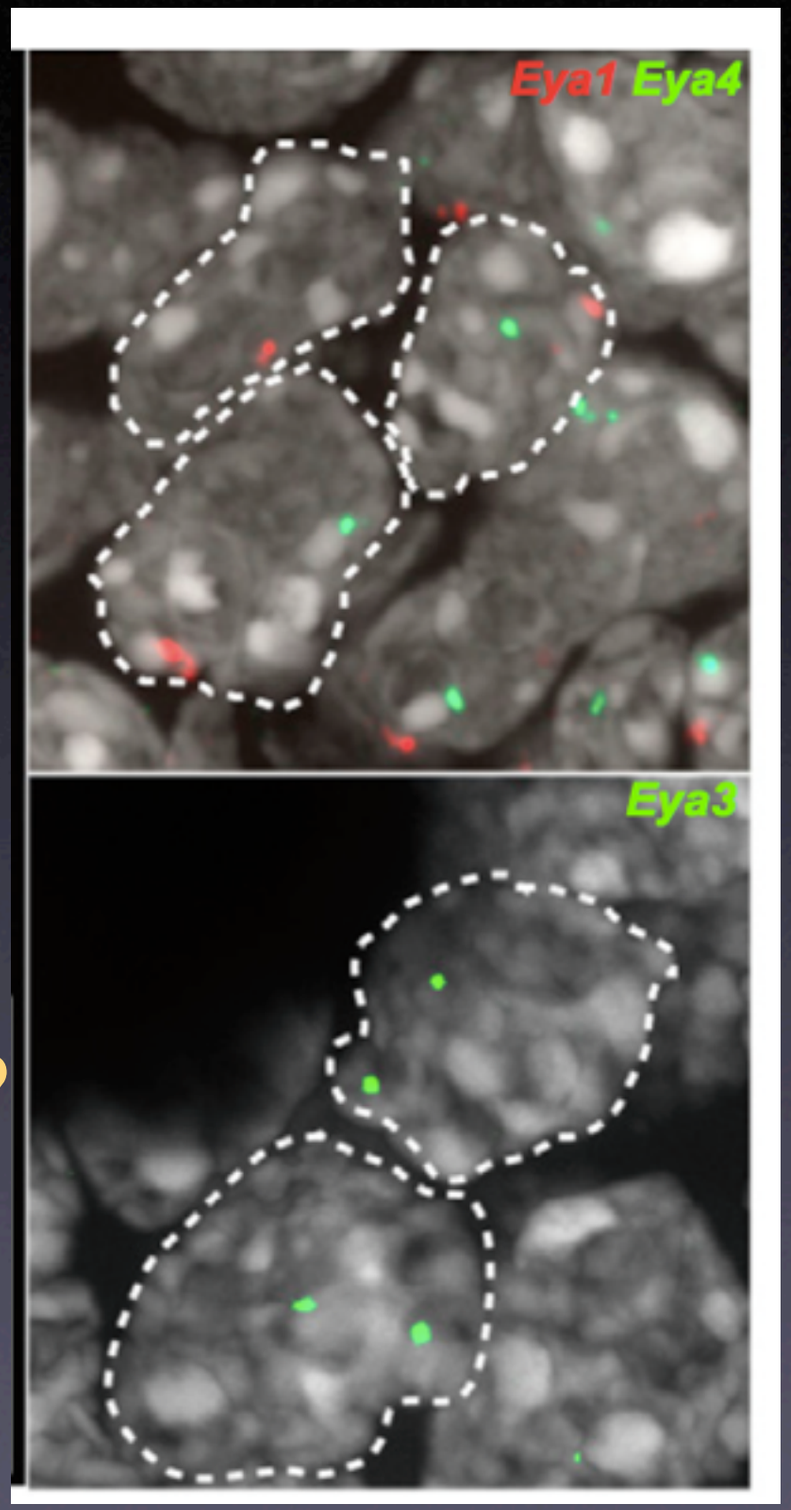
Modular Protein Architecture and the Construction of Cell Regulatory Systems

Toby J. Gibson
Structural & Computational Biology Unit
EMBL, Heidelberg

Using RNA fish, Eya4 shows random monoallelic expression (RME) in eye development

Many, many developmentally important genes show RME

Question: *Why should genes be expressed from just one of the two alleles during development?*

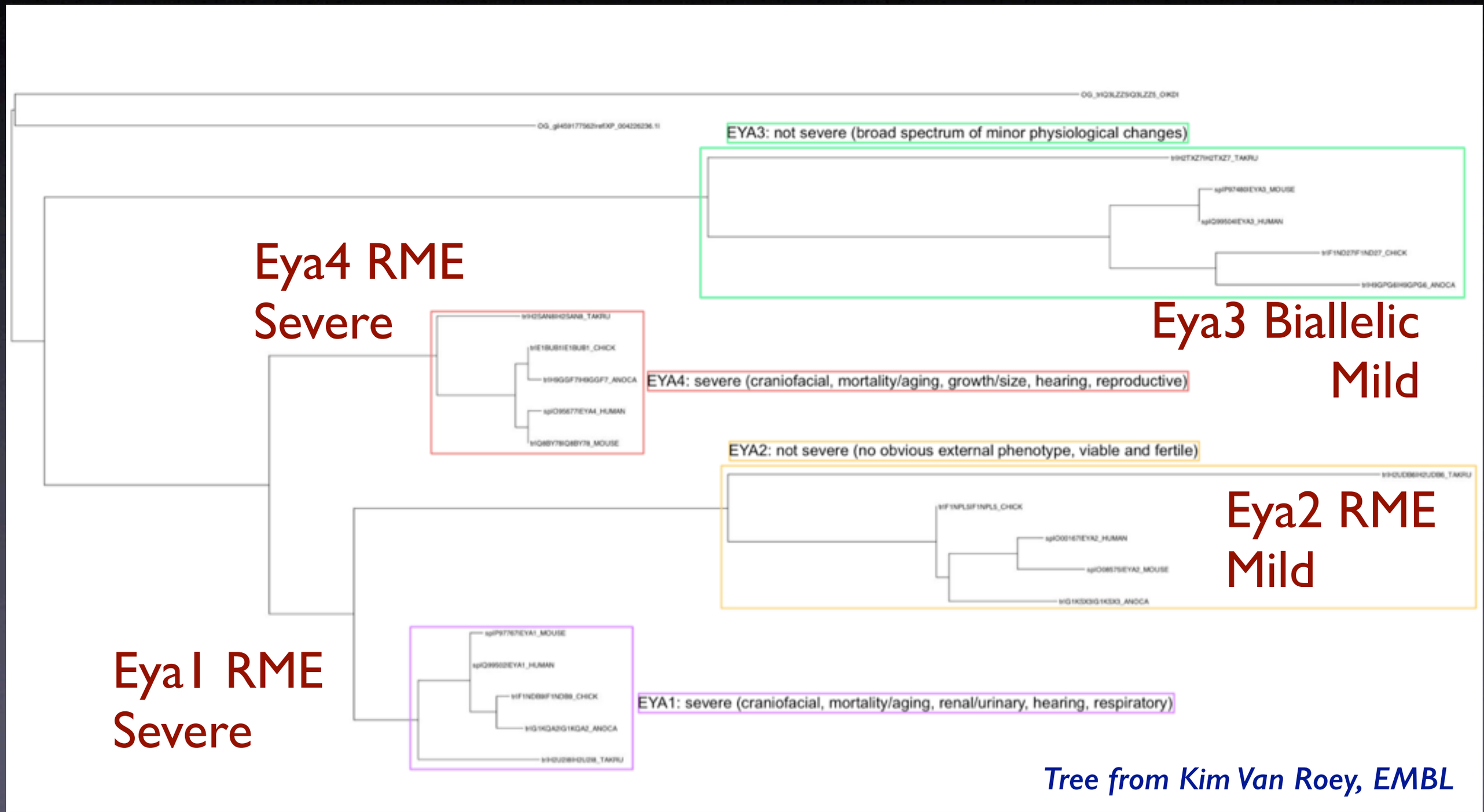


Eya1, Eya4
Monoallelic

Eya3
Biallelic

Gendrel et al. (2014) Dev. Cell 28, 366

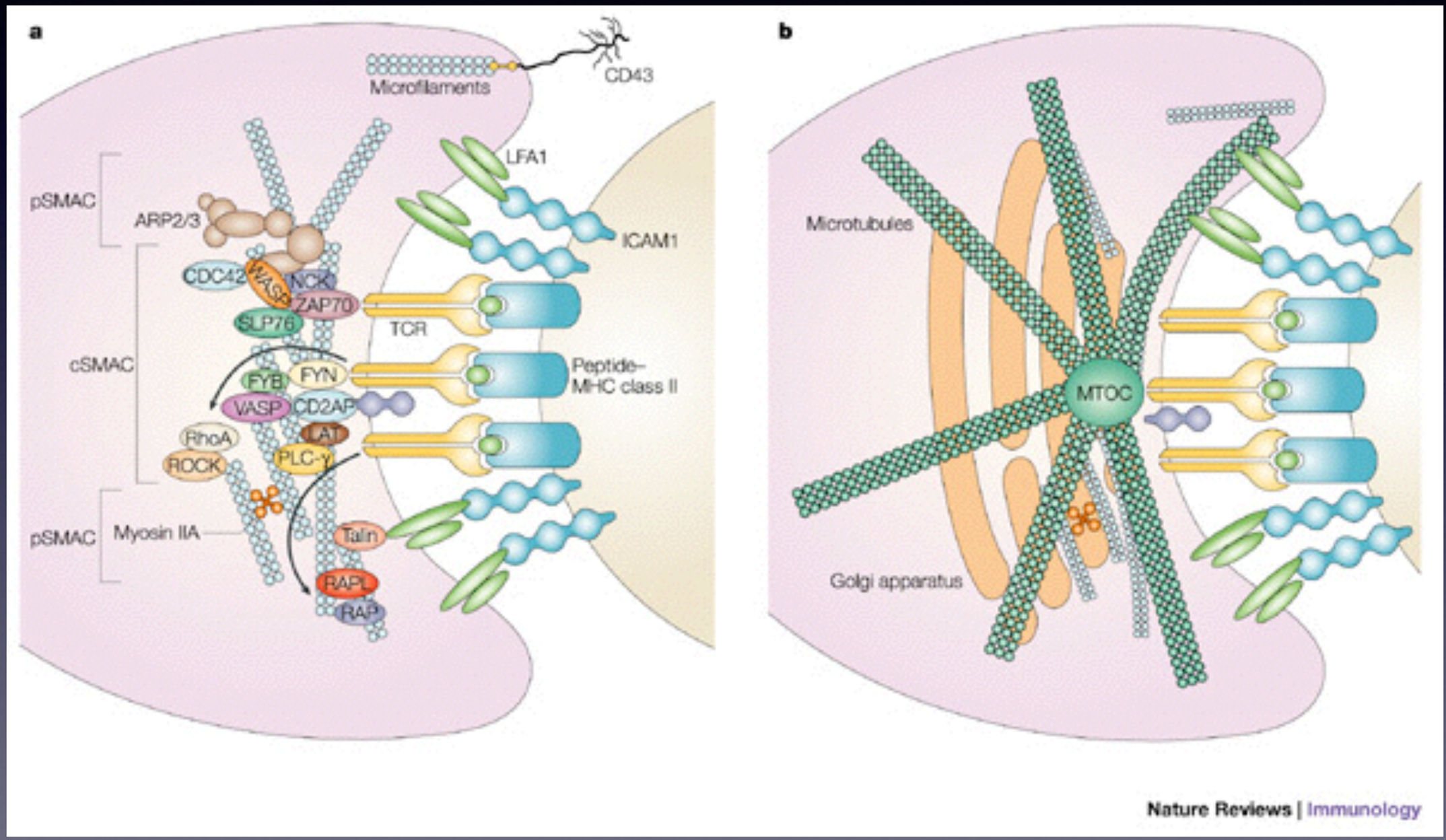
Eya paralogues are evolving at different rates. Gene knockouts have different severities. Eya1 and Eya4 heterozygotes have strong phenotypes.



Tree from Kim Van Roey, EMBL

When a signal is received by a membrane receptor, what happens next?

The Immunological Synapse - A platform for multisignal input and output in T Cell activation

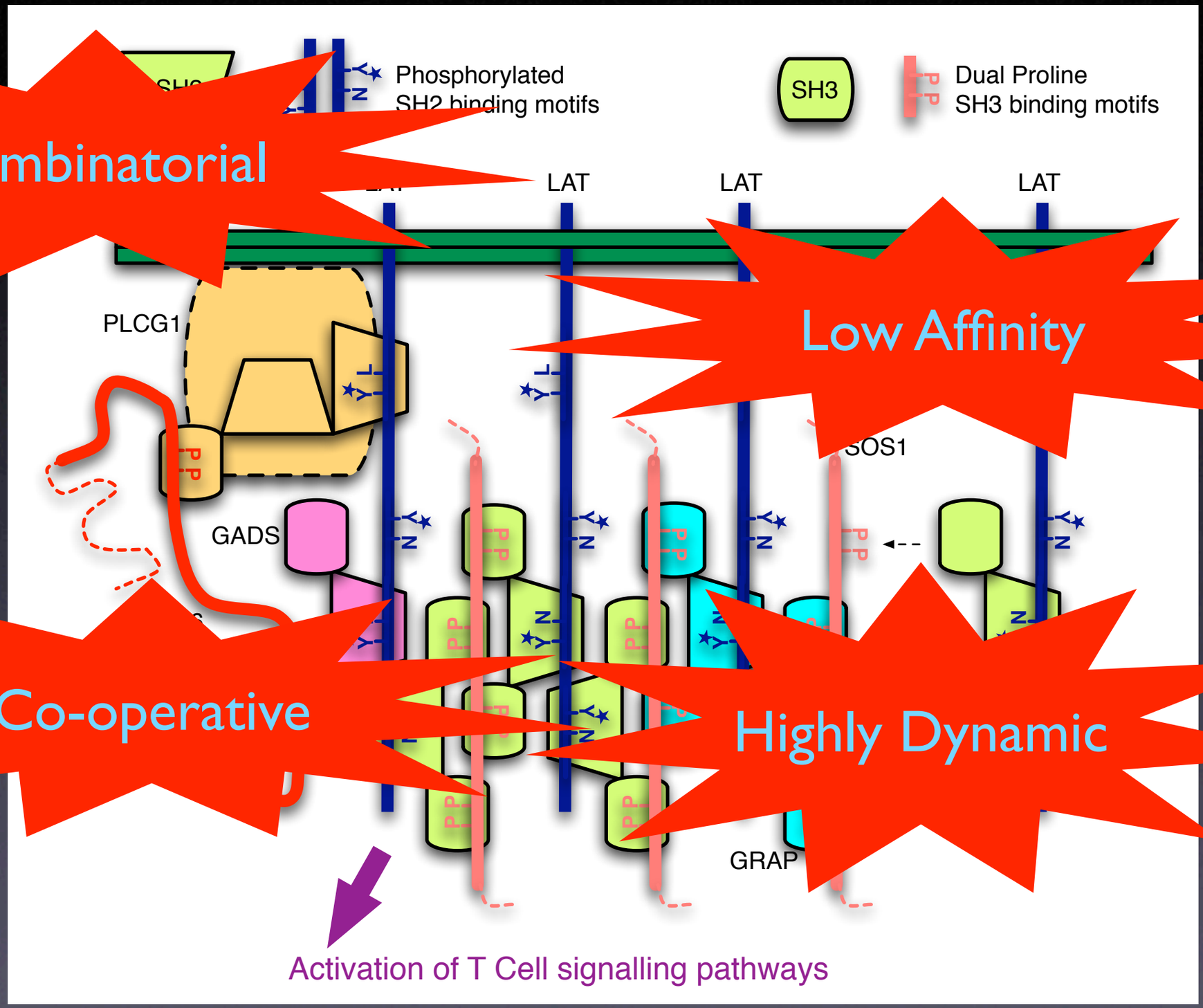


Vicente-Manzanares et al. (2004) Nat. Rev. Imm. 4, 110

Propagation of T Cell signalling

Multivalent assembly of the LAT signalling complex by short linear motifs

LAT is phosphorylated by TKs ZAP-70 and/or SYK



Combinatorial

Low Affinity

Co-operative

Highly Dynamic

After Houtman et al. (2006) NS&MB, 13, 798

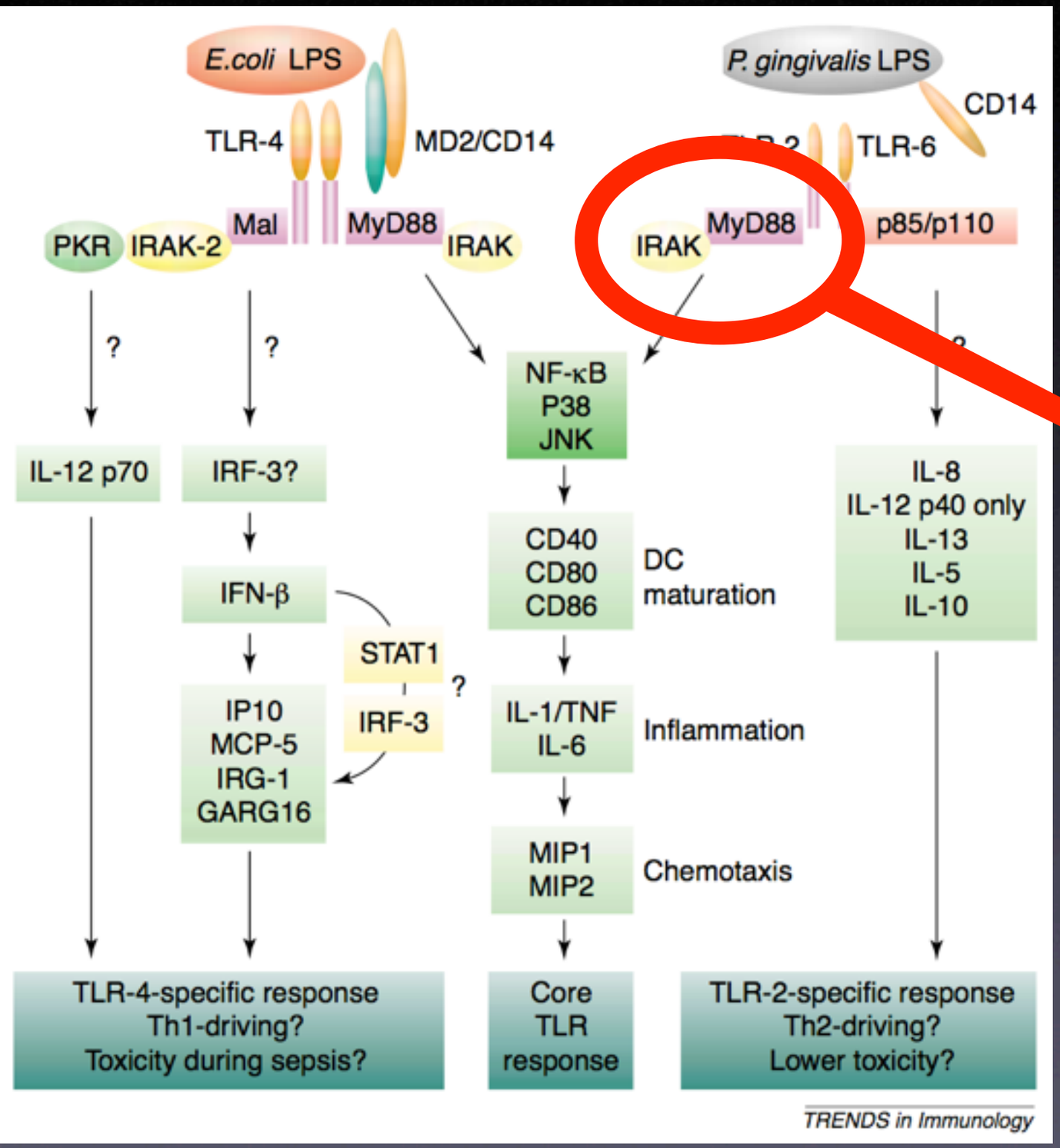
The LAT interaction fur-ball retrieved from the STRING server

Is this a good representation of the molecular details?

STRING is at http://string.embl.de/

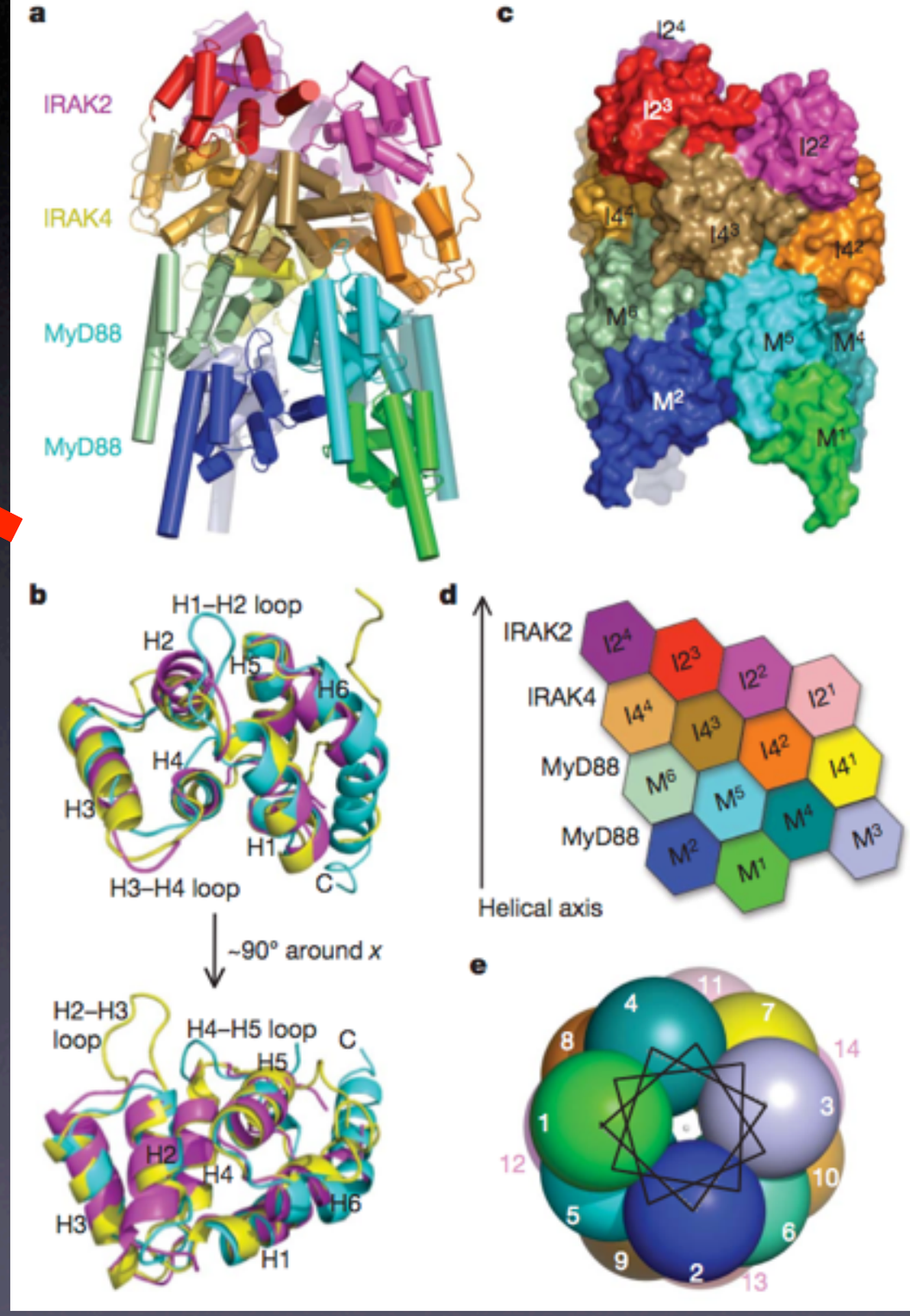
Innate Immunity

Toll-like receptor signalling



O'Neill (2002) Trends Imm. 23, 296

Assembly of the myddosome using death domains from MyD88, IRAK4, IRAK2



Lin et al. (2010) Nature 465, 885

When a signal is received by a membrane receptor, what happens next?

Answer

Typically, a discrete signalling platform is assembled to integrate other cell state signals so that an informed decision leads to the correct outcome

You are an engineer:

If system reliability is critical, would you design a simple system or a complex one?

Robustness of biological systems

Complexity and robustness

J. M. Carlson*† and John Doyle‡

*Department of Physics, University of California, Santa Barbara, CA 93106; and †Control and Dynamical Systems, California Institute of Technology, Pasadena, CA 91125

Carlson and Doyle (2002) PNAS, 66, 2538

...By robustness, we mean the maintenance of some desired system characteristics despite fluctuations in the behavior of its component parts or its environment....

BIOLOGICAL ROBUSTNESS

Hiroaki Kitano

Abstract | Robustness is a ubiquitously observed property of biological systems. It is considered to be a fundamental feature of complex evolvable systems. It is attained by several underlying principles that are universal to both biological organisms and sophisticated engineering systems.

Kitano (2004) Nat Rev Genet, 5, 826

CH Waddington (1905-1975)

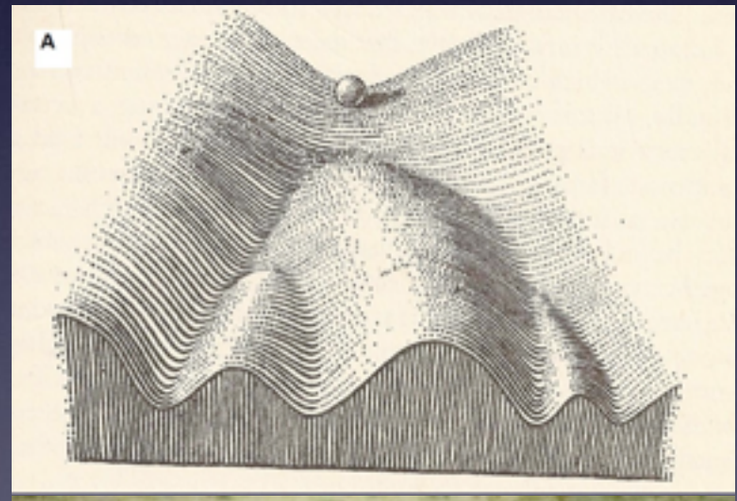
- A unifier of development and genetics

- A forefather of systems biology

- System robustness and weak phenotypes

Some of Waddington's concepts:

- Epigenetic Landscape
  - *Developmental cell fates and increasing irreversibility*
- Canalisation
  - *Robustness in developmental processes*
- COWDUNG
  - *Conventional Wisdom of the DUmiNant Group*



Increases in system complexity due to selection for robustness introduce a new issue: **system fragility**

A good example is the Internet which is:

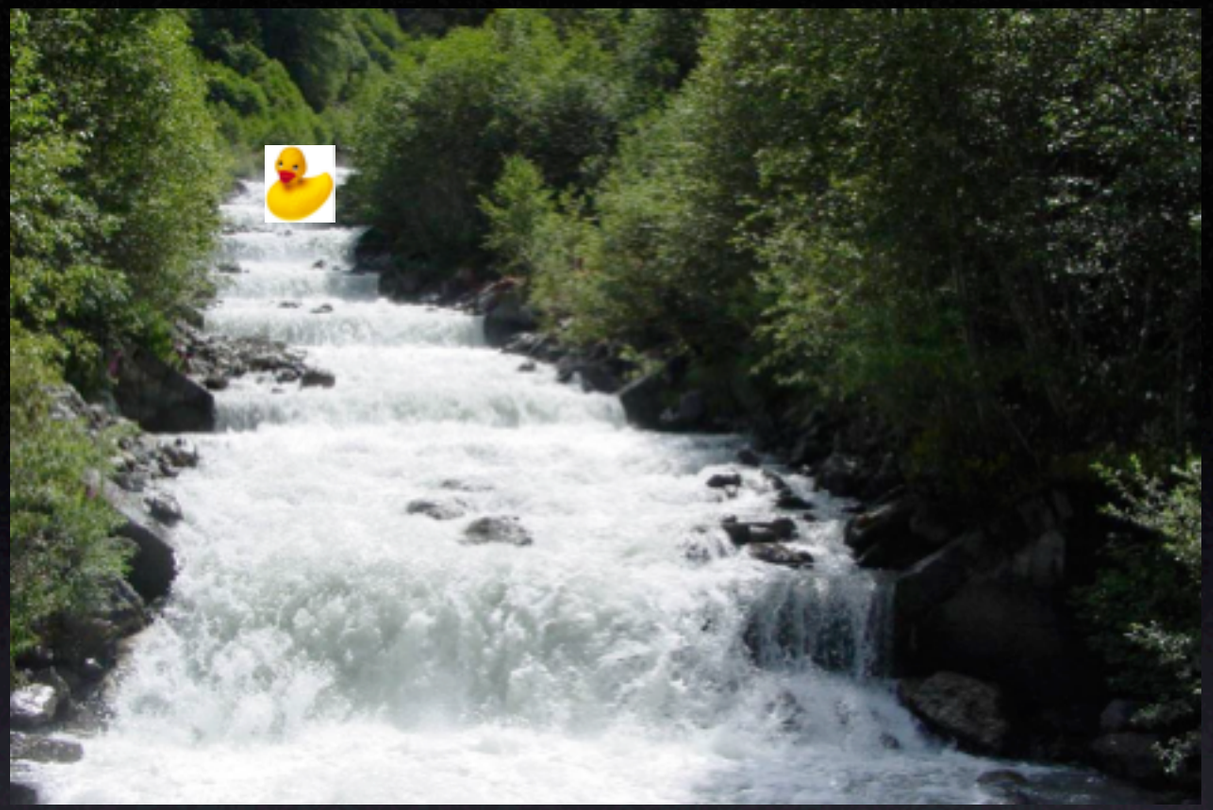
"robust yet fragile" (RYF)

that is, unaffected by random component failures but vulnerable to targeted attacks on its key components.

Cascades

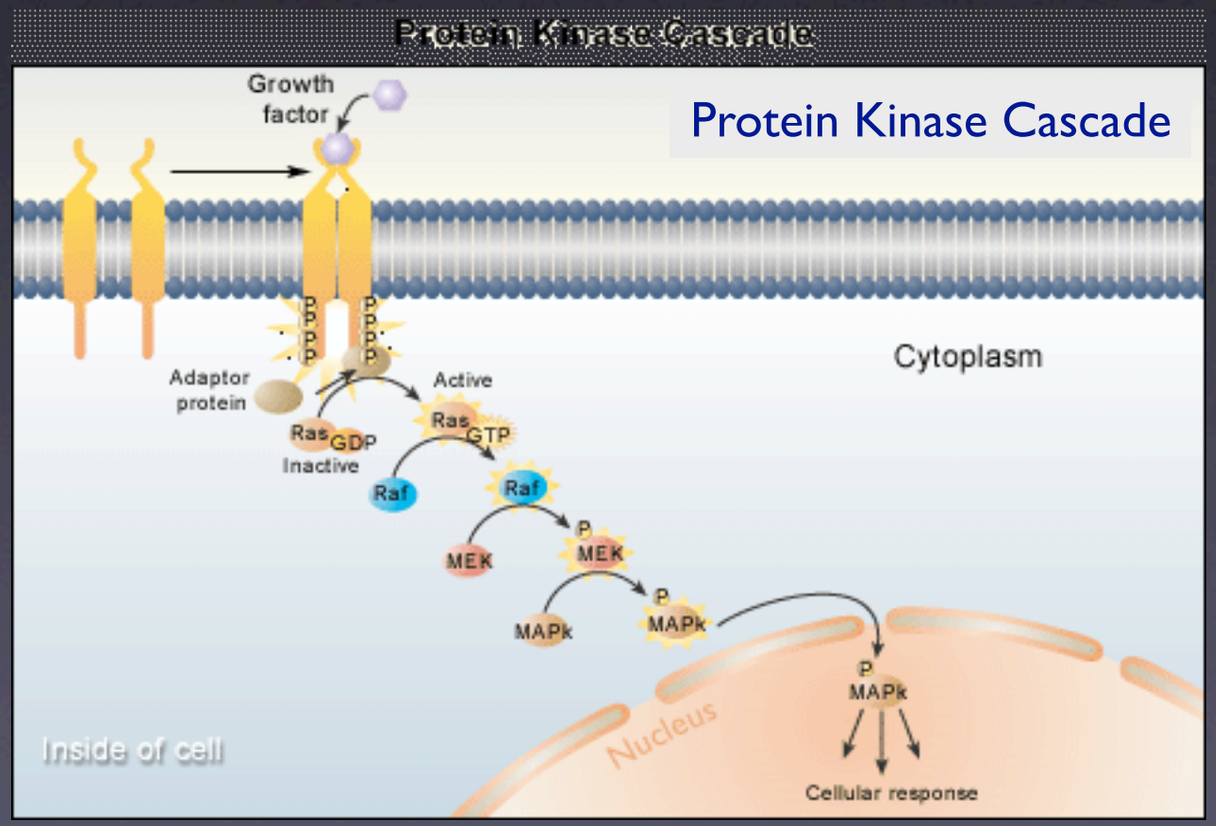
Properties

Linearity
Uneven
Accellerating
Unregulated
Uncertain end point?



Cascade in South Tyrol, source K. Amon and G. Zsoldos

Cascading mechanisms
are neither accurate
nor precise



source http://www.biology.arizona.edu/cell_BIO/

The first report of a protein kinase cascade

Cell

Result list |

Volume 25, Issue 1, July 1981, Pages 9-21

doi:10.1016/0092-8674(81)90227-0 Cite or Link Using D

Article

A mouse homolog to t[illegible]viar[illegible]rcoma virus *src* protein is a mem[illegible]of a p[illegible]ein kinase cascade

Mark Spector, Robe[illegible]B.[illegible]ins, Volker M. Vogt and Efraim Racker

Section of Bioc[illegible]str[illegible]Mole[illegible]ar and Cell Biology Wing Hall Cornell University, Ithaca, [illegible]York 14853, USA

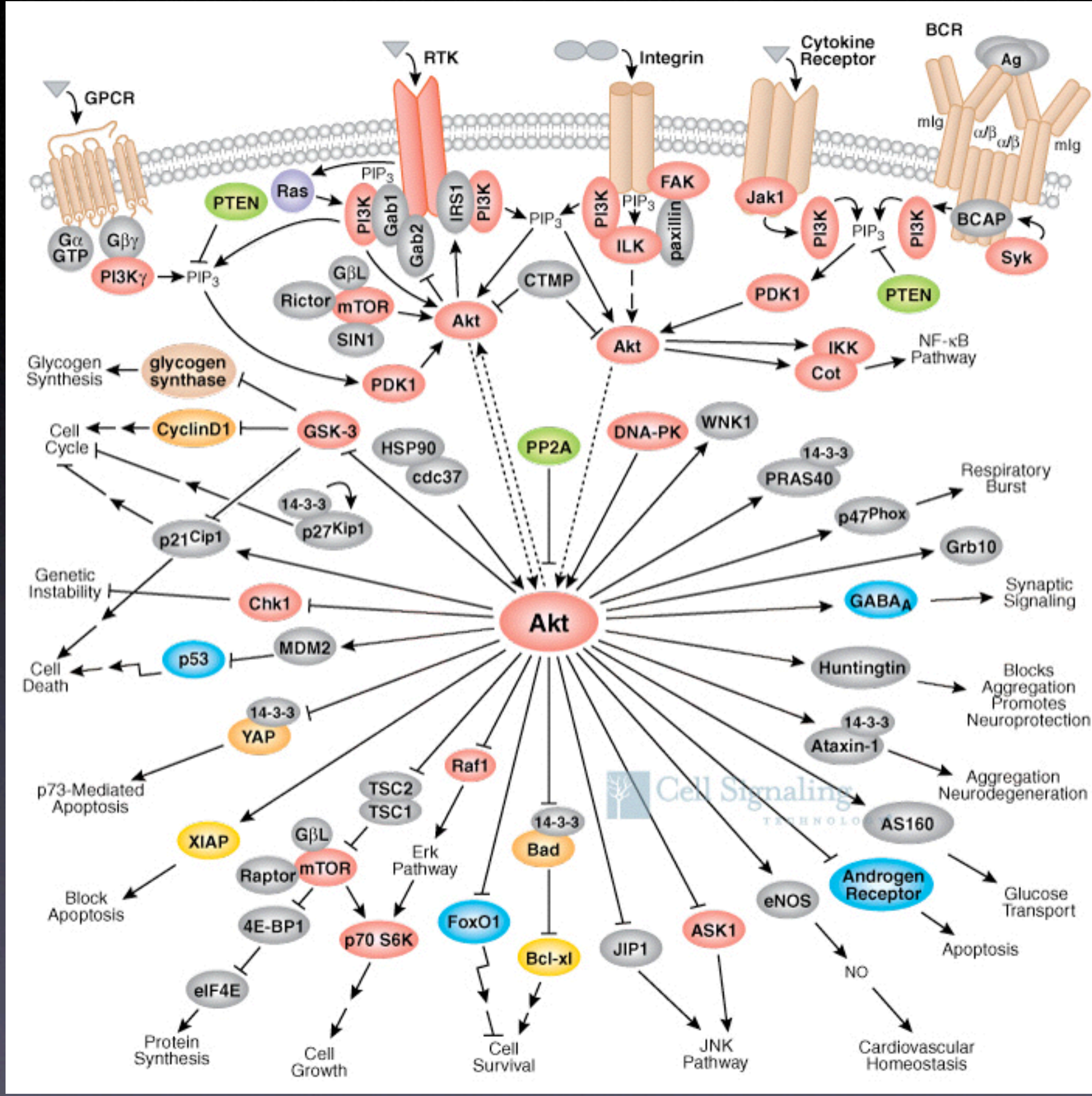
Received 2 March 19[illegible]Revised 21 April 1981. Available online 28 April 2004.

Abstract

Recent work has identified a cascade of membrane-bound protein kinases in Ehrlich ascites tumor cells. These enzymes, designated PK_L, PK_S and PK_M, are present in both Ehrlich tumor and mouse

AKT / PKB Kinase Cascade

Cascade or Network?



http://www.cellsignal.com/reference/pathway/Akt_PKB.jsp

Most Tyrosine Kinases have very limited sequence specificity

- *in vivo* TK substrate detection remains difficult
- *in vivo* substrates ≠ good in *vitro* peptides
- Cannot define a simple sequence pattern at phosphosite
- Problem: how do they avoid each other's substrates?

ELSEVIER

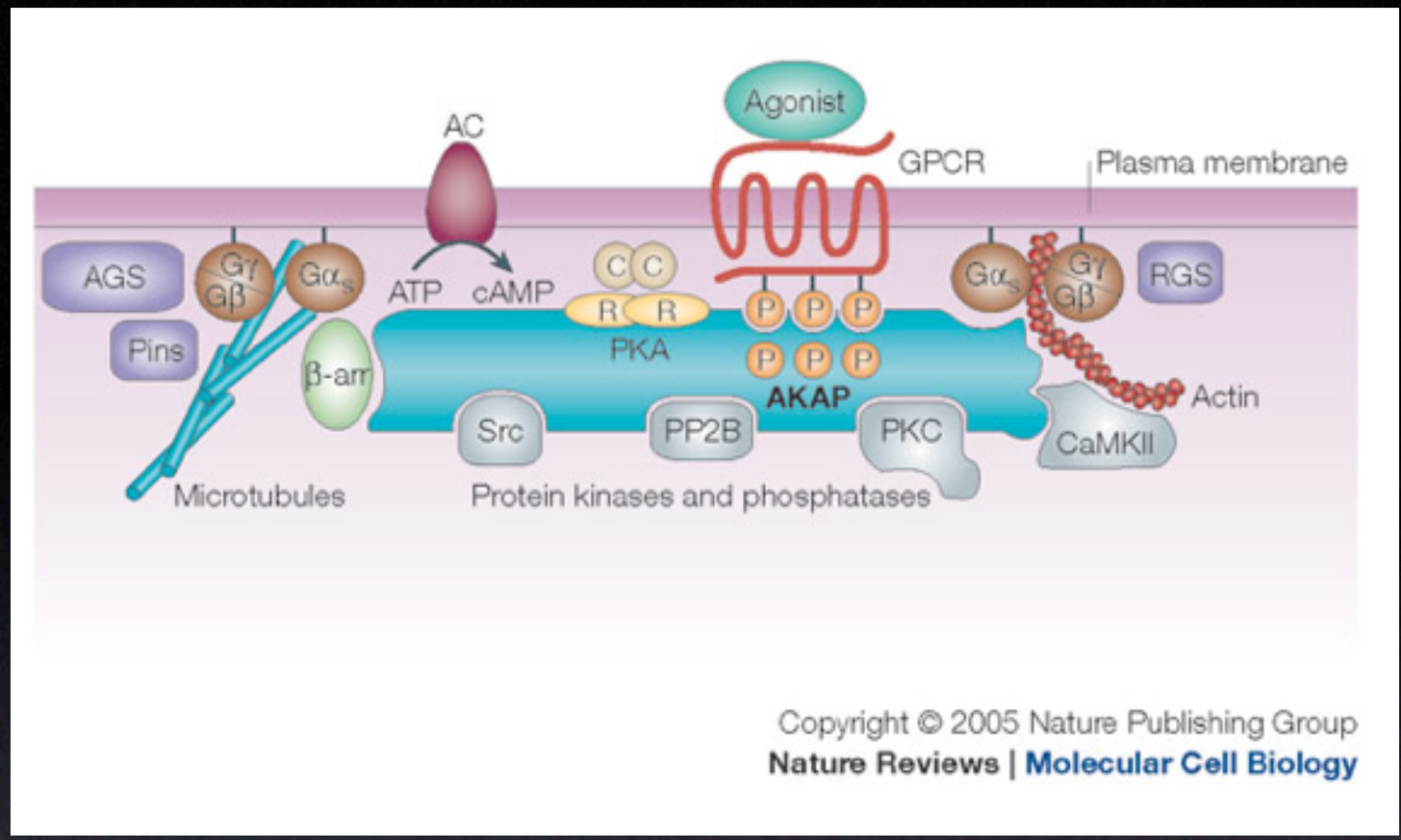
Protein tyrosine kinase–substrate interactions

Ron Bose[1,2,*], Marc A Holbert[1,*], Kerry A Pickin[1,*] and Philip A Cole[1,2]

Curr. Op. Str. Bio. (2006) 16, 668

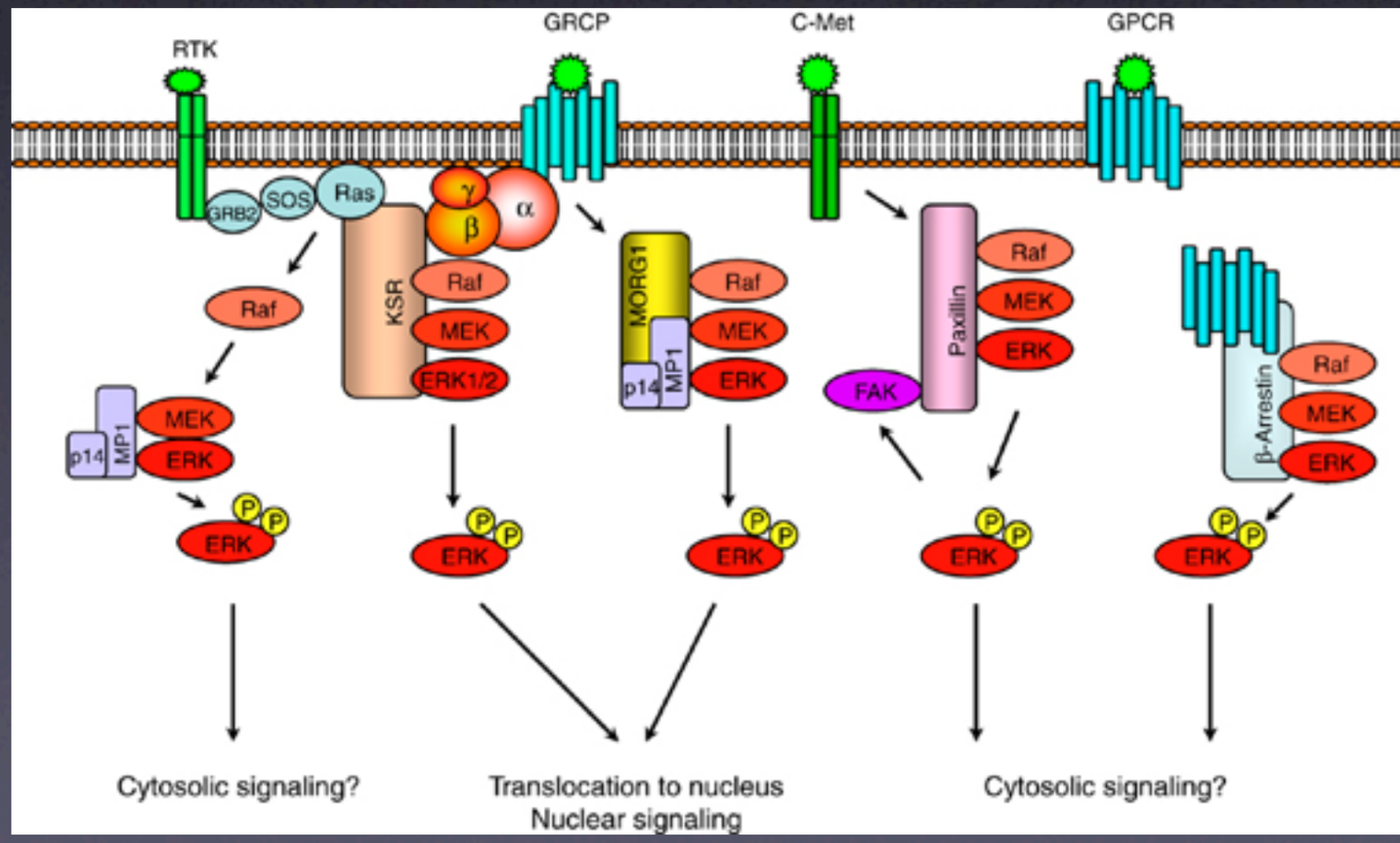
Solution to kinase substrate specificity problem: Scaffolding

PKA/Src/PKC scaffold



Malbon (2005) NRMCB, 6, 689

Map kinase scaffolds



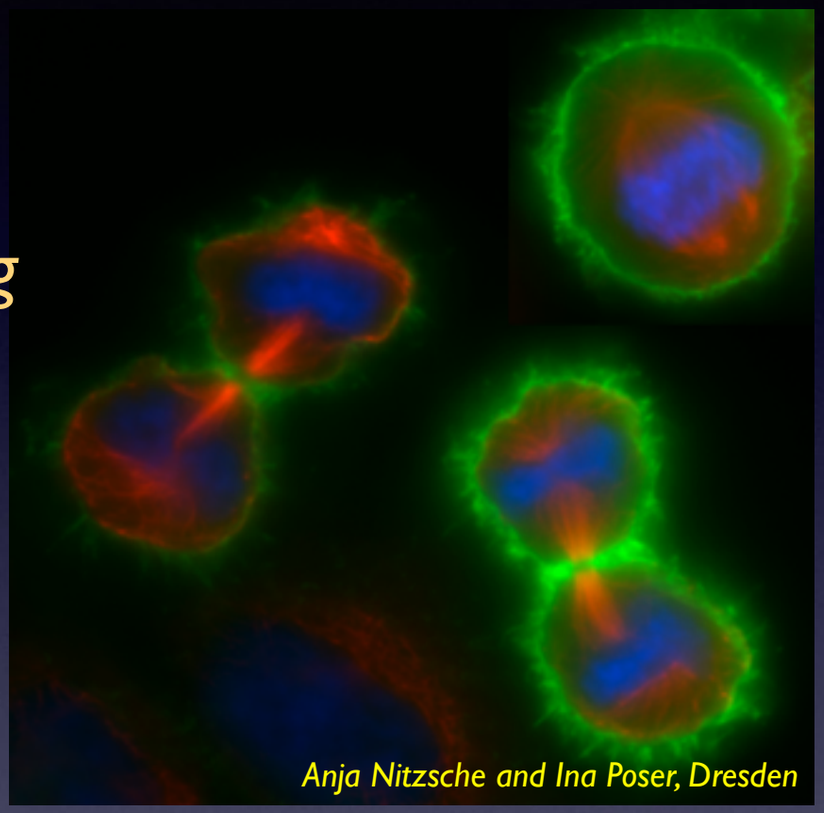
Dhanasekaran (2007) Oncogene, 26, 3185

Lots of different AKAPs scaffold the PKA kinase

Different complexes in different locations

AKAP5

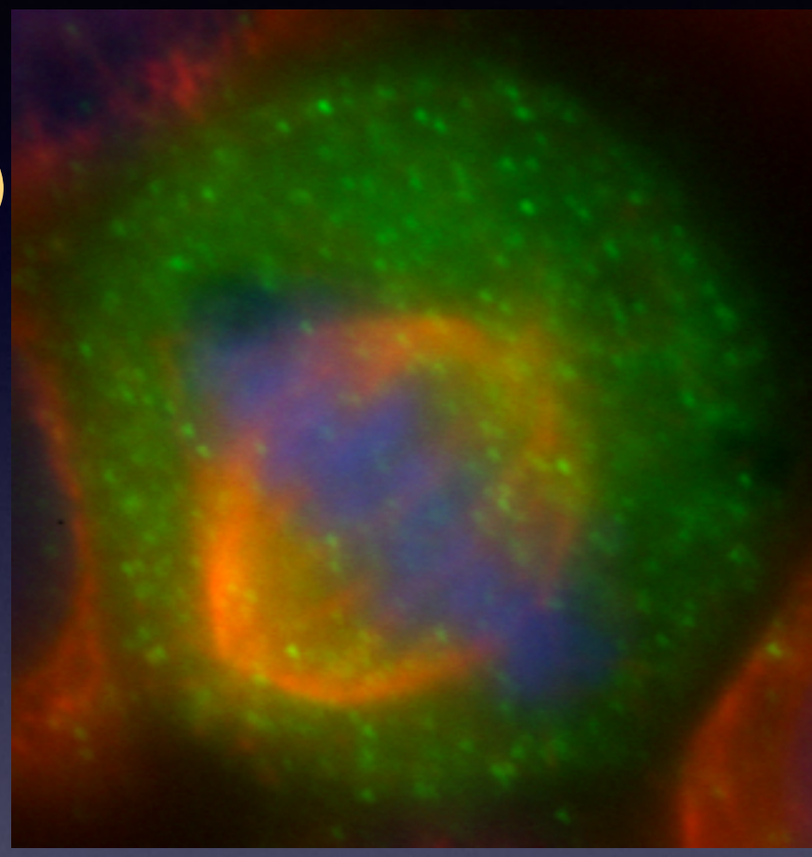
A-kinase (PKA)anchoring protein 5



Anja Nitzsche and Ina Poser, Dresden

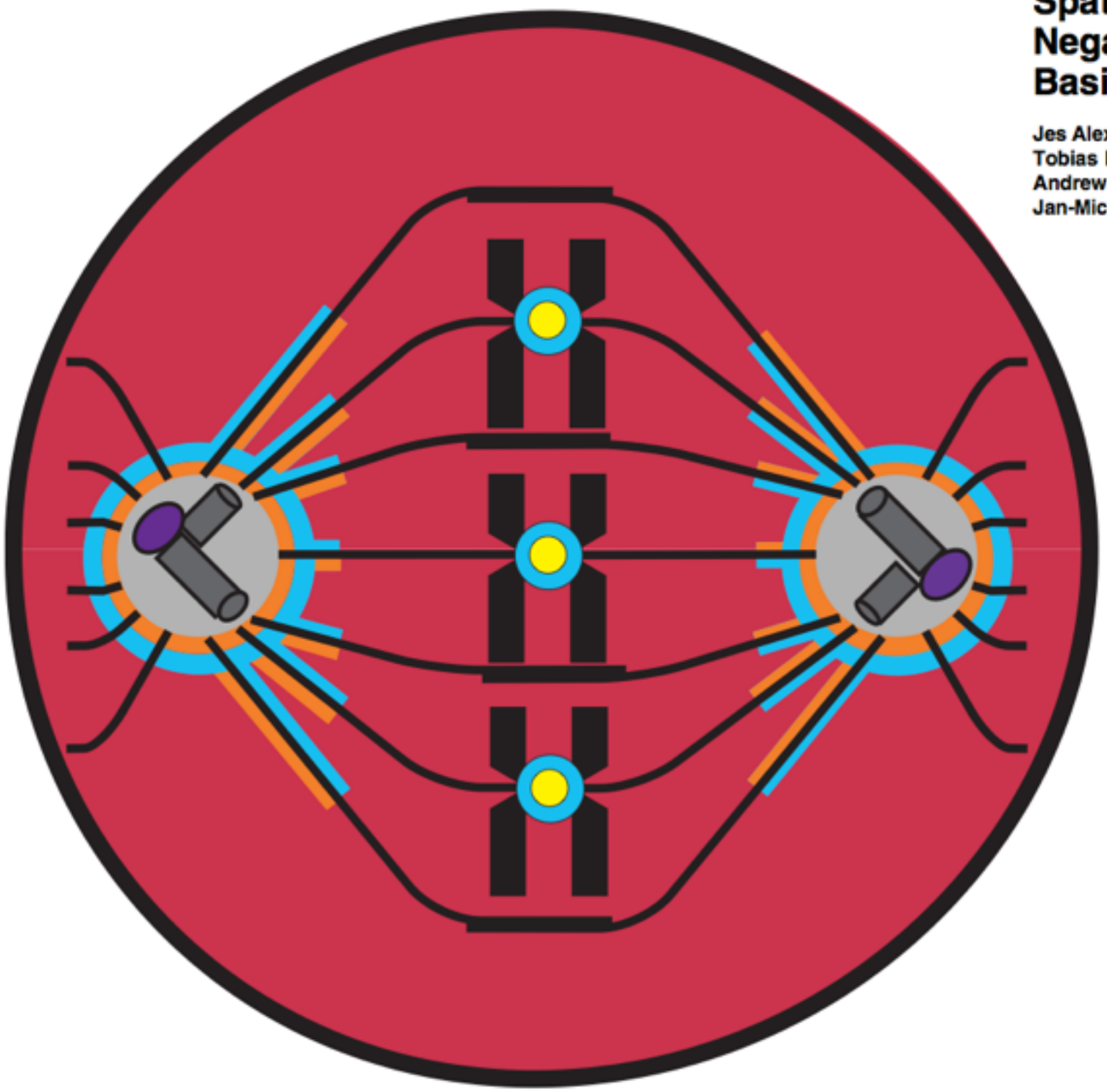
AKAP12

A-kinase (PKA) anchoring protein 12



[Tagged Cell lines made in Dresden as part of the ***Mitocheck*** and ***DiGtoP*** projects]

Spatial Exclusivity of Mitotic Kinases



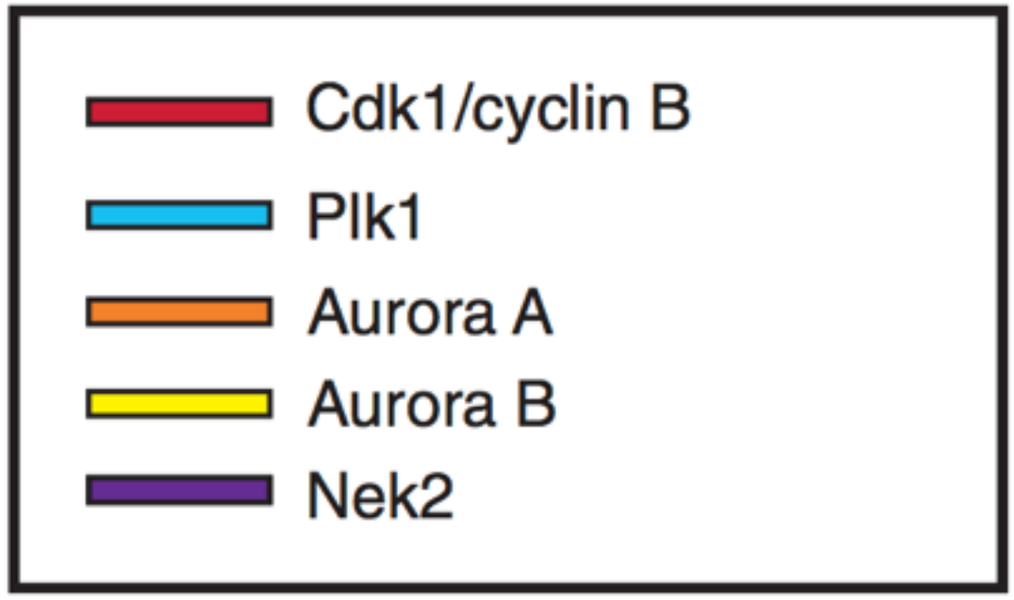
RESEARCH ARTICLE

MITOTIC KINASES

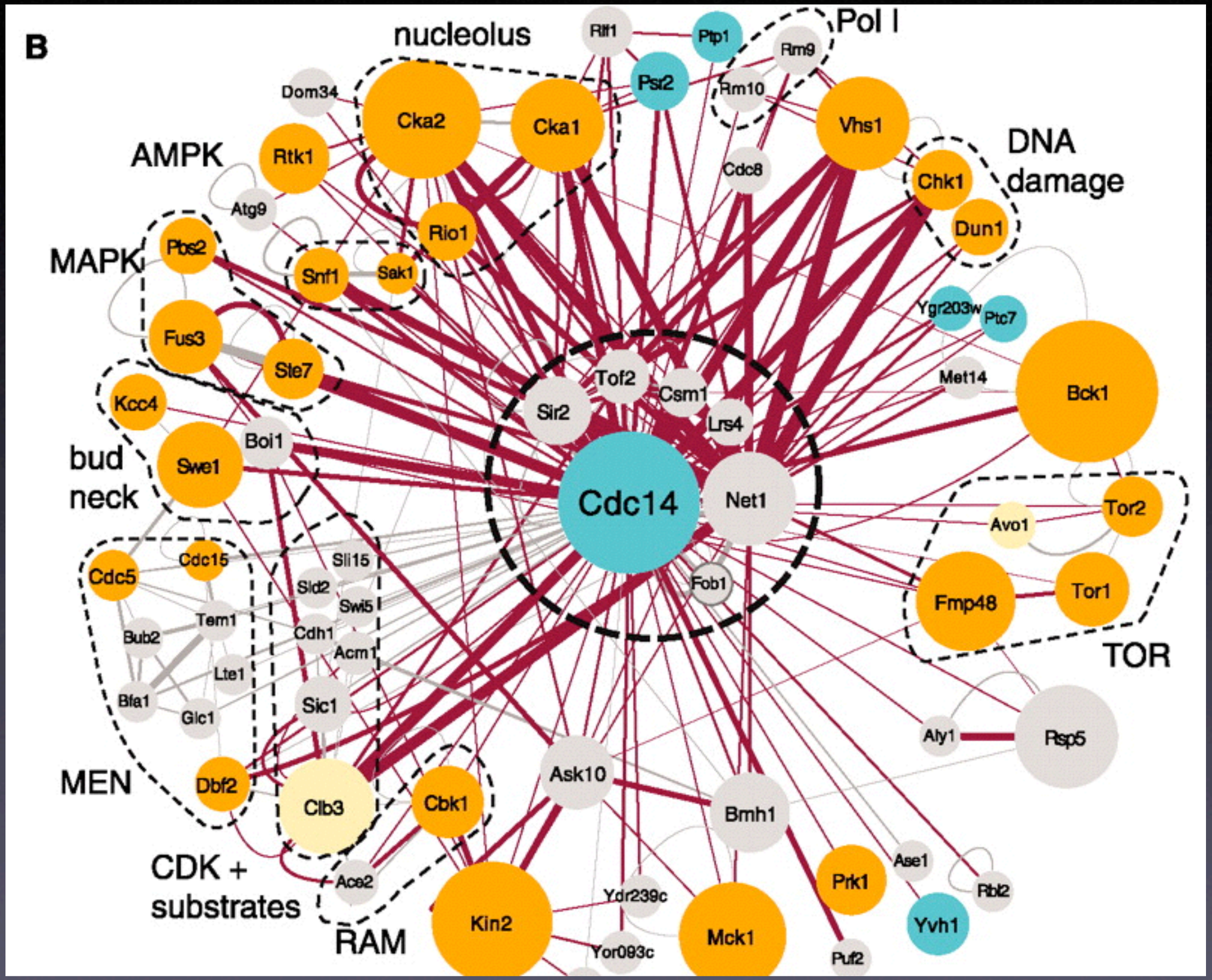
Sci. Sig., 6-2011

Spatial Exclusivity Combined with Positive and Negative Selection of Phosphorylation Motifs Is the Basis for Context-Dependent Mitotic Signaling

Jes Alexander,[1*] Daniel Lim,[1] Brian A. Joughin,[1] Björn Hegemann,[2†] James R. A. Hutchins,[2] Tobias Ehrenberger,[1] Frank Ivins,[3] Fabio Sessa,[4] Otto Hudecz,[2] Erich A. Nigg,[5] Andrew M. Fry,[6] Andrea Musacchio,[4] P. Todd Stukenberg,[7] Karl Mechtler,[2] Jan-Michael Peters,[2] Stephen J. Smerdon,[3] Michael B. Yaffe[1,8‡]



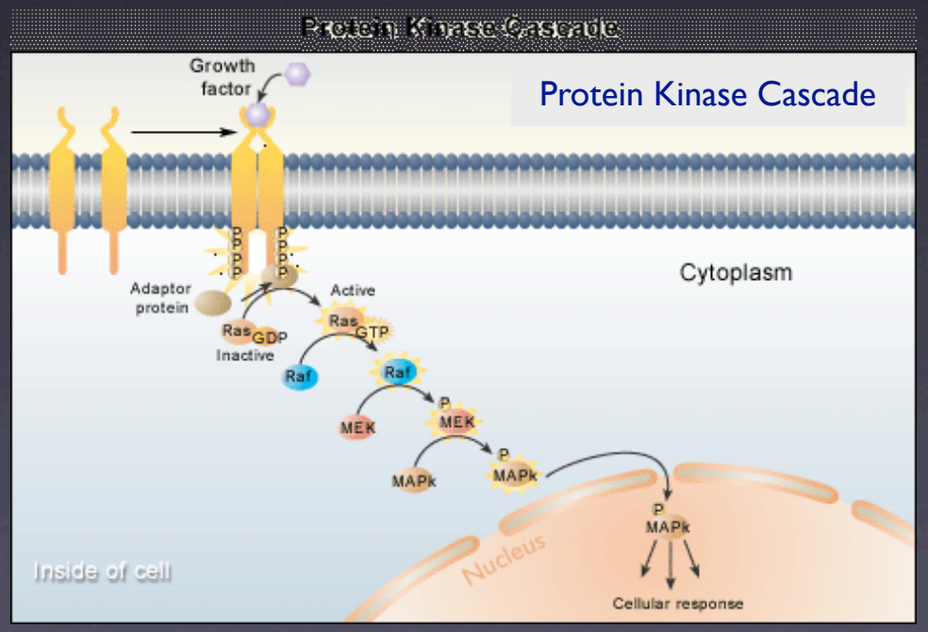
Yeast Cdc14 phosphatase interaction network



Breitkreutz et al. (2010) Science, 328, 10443

Kinases are networked, scaffolded and have limited or nonexistent substrate specificity

- Kinases do not find their substrates by simple free diffusion
  - Widely used Reaction-Diffusion equations are insufficient for modelling kinase signalling
- *"Kinase Cascade" is one of the worst analogies in Biology and its meme needs to become extinct*

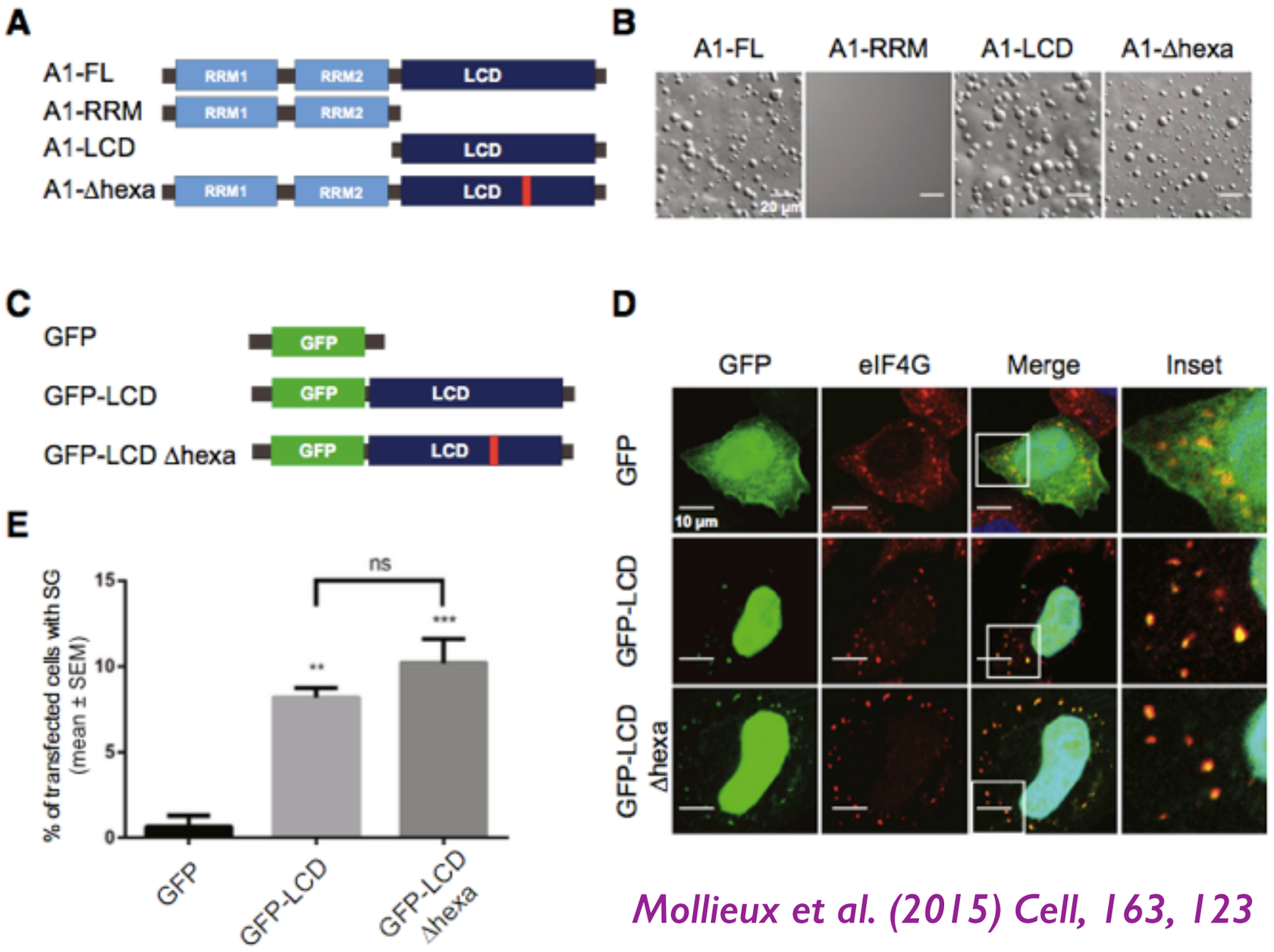


source http://www.biology.arizona.edu/cell_BIO/

Instead of measuring concentration,
[the cell] counts molecules

Sydney Brenner, 2007

Some types of complex tolerate stoichiometry violations



Mollieux et al. (2015) Cell, 163, 123

Figure 2. Liquid-Liquid Phase Separation by hnRNPA1 Is Mediated by the C-Terminal Low Complexity Sequence Domain and Is Distinct from Fibrillization

(A) Schematic of the structure of hnRNPA1 full length (A1-FL), the N terminus comprising the two folded RNA recognition motifs (A1-RRM), the low complexity sequence domain (A1-LCD), and the mutant with a deletion of residues 259–264 (Kim et al., 2013) (A1-Δhexa).

(B) DIC images of A1-FL, A1-RRM, A1-LCD, and A1-Δhexa at 140 μM protein, 150 mg/ml Ficoll in 50 mM HEPES, 300 mM NaCl, and 5 mM DTT.

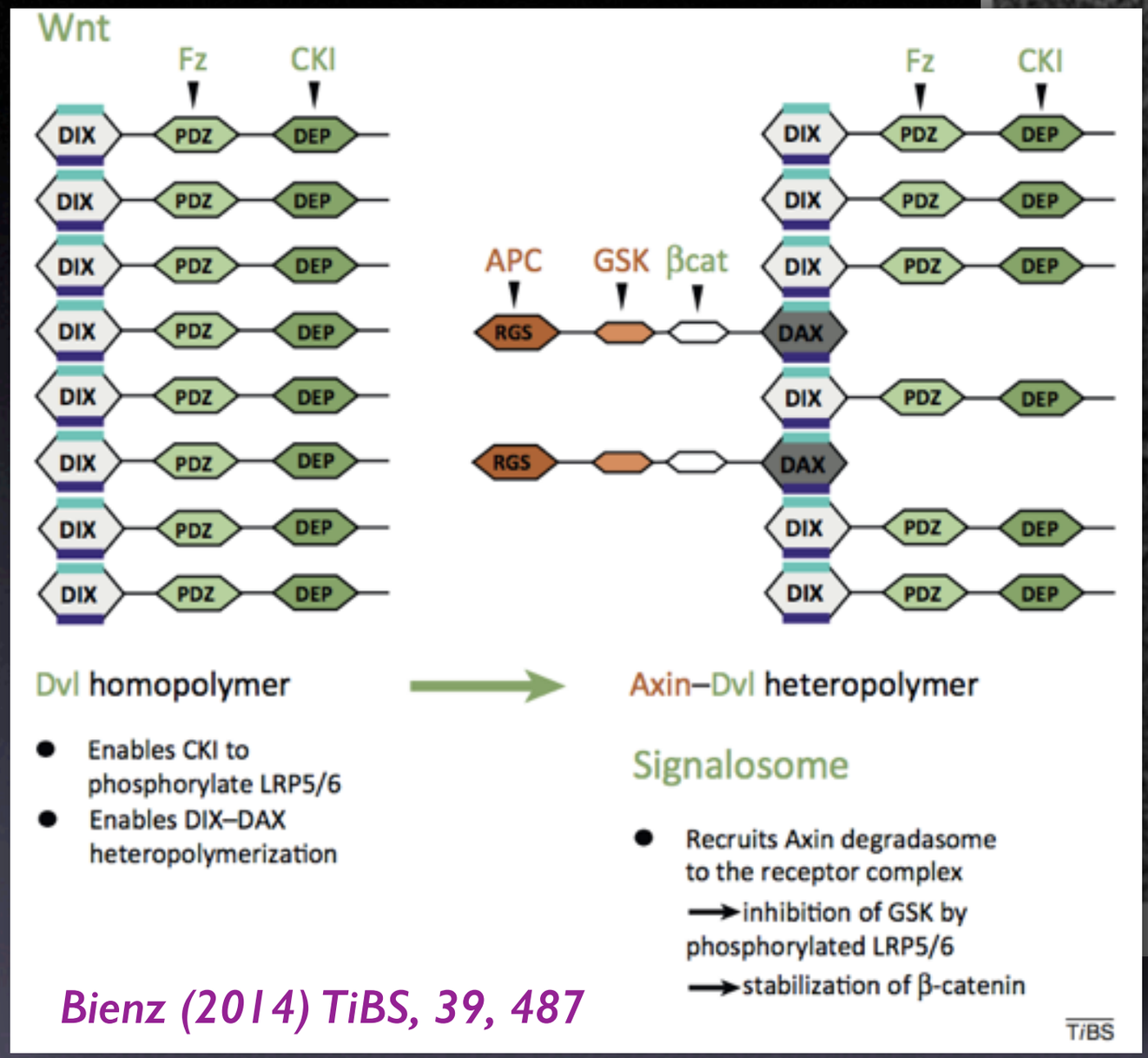
(C) Schematic of the constructs transiently expressed in HeLa cells.

(D) Representative confocal microscopy images of HeLa cells transfected with constructs presented in (C), treated with 0.5 mM sodium arsenite for 15 min, and immunostained with anti-eIF4G (red) and DAPI (blue).

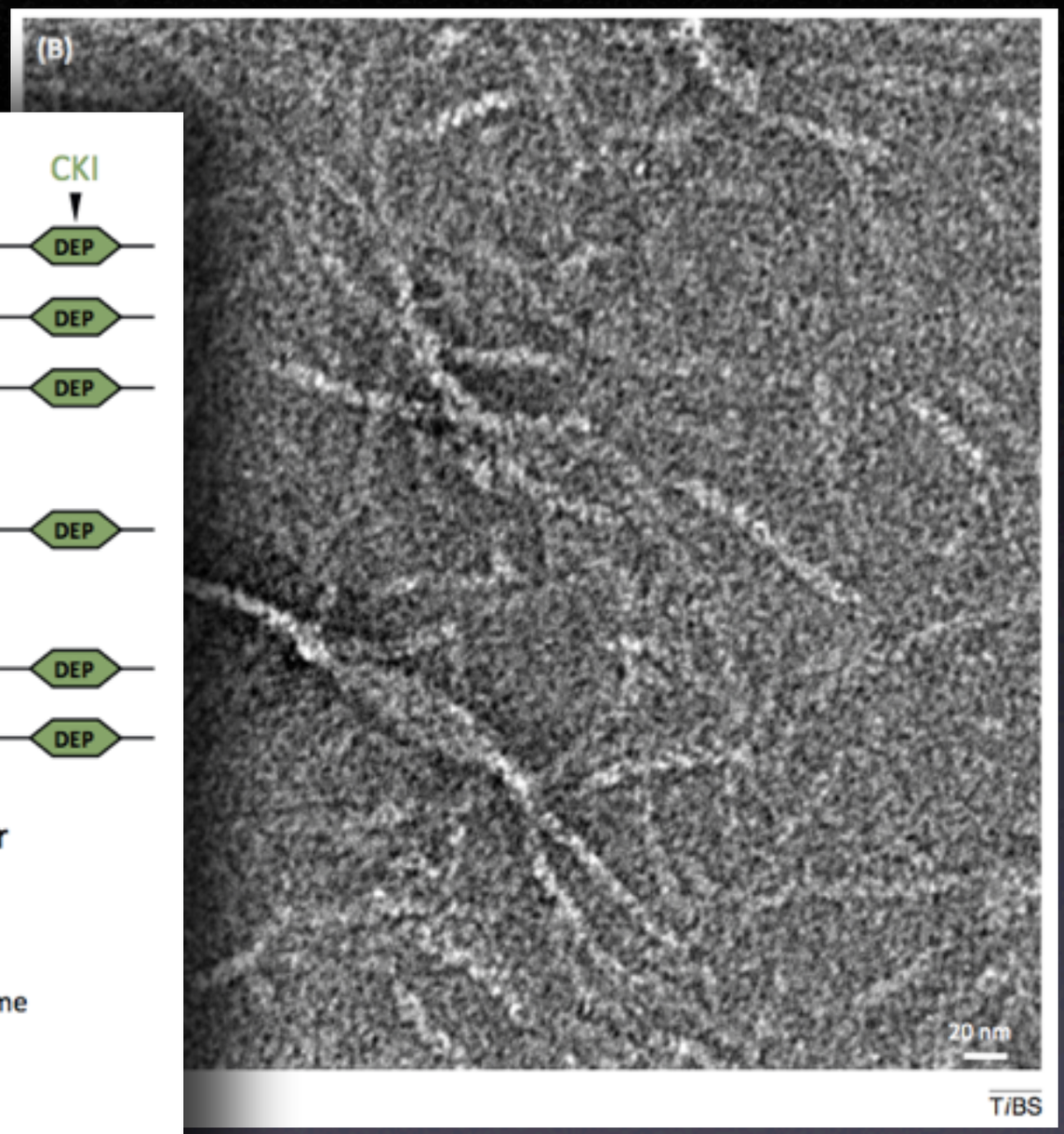
(E) Quantification for data in (D). The percentage of transfected cells displaying GFP signal in SGs ([number of cells with GFP-positive SGs/number of GFP-expressing cells] × 100) was plotted as mean ± SEM; n = 100 cells; **p < 0.005, ***p < 0.001 by one-way ANOVA, Tukey's post hoc test.

Liquid phase separation "complexes" scale to any size (e.g. stress granules, nucleoli)

Some types of complex tolerate stoichiometry violations



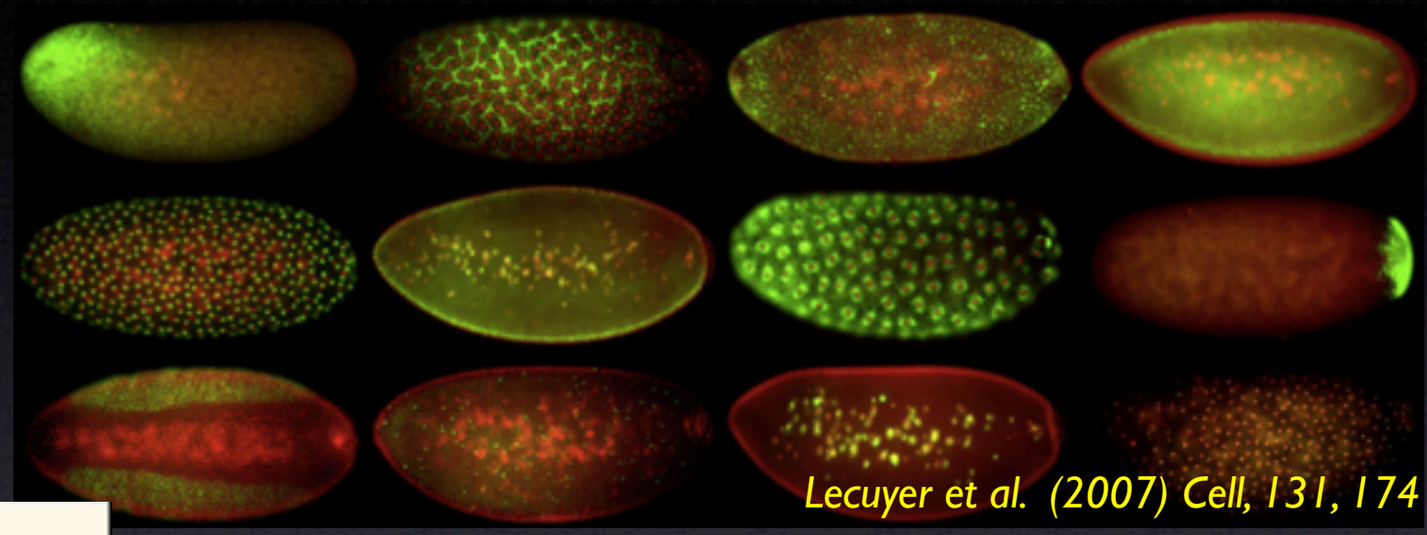
Bienz (2014) TiBS, 39, 487



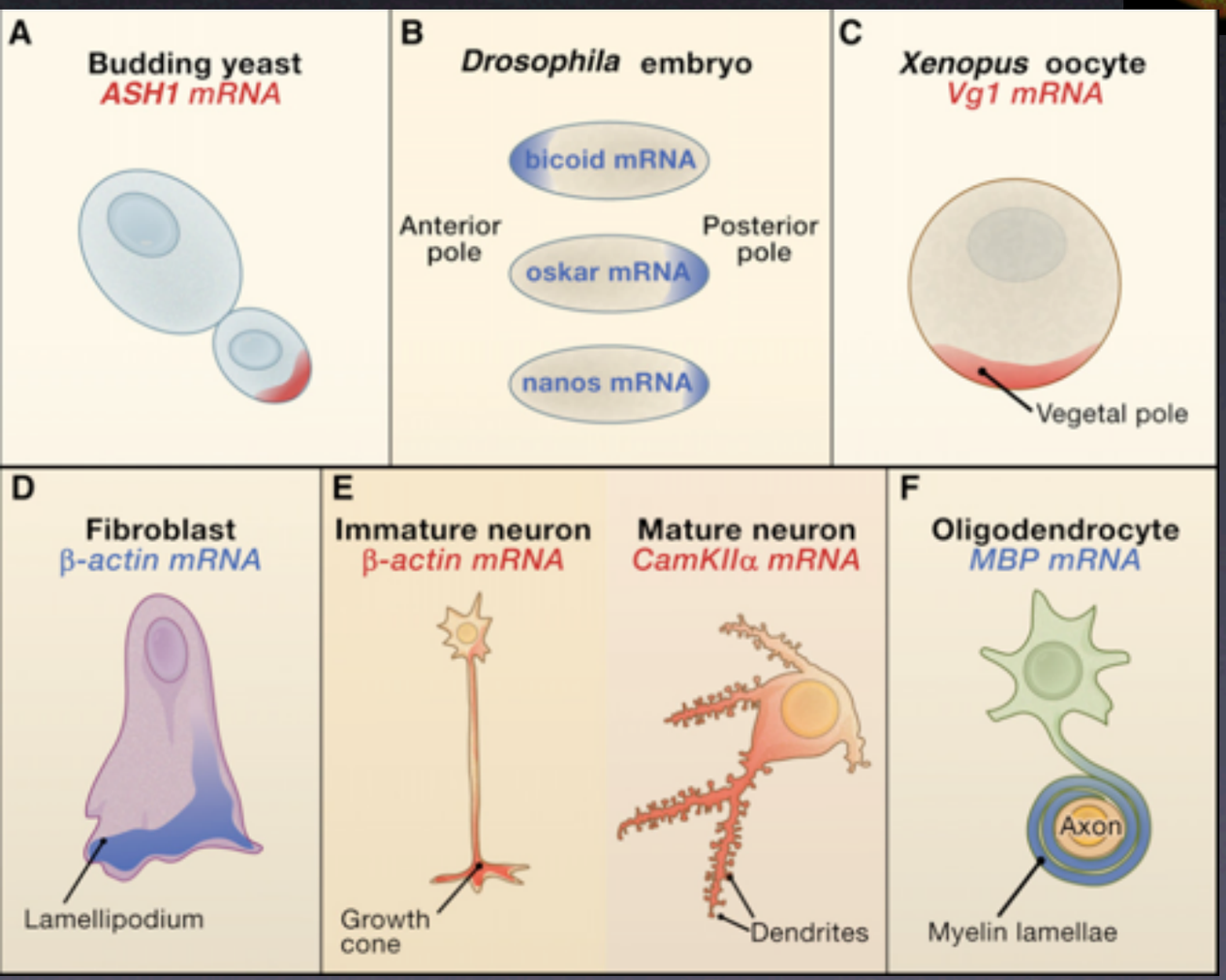
Polymeric helical signalosomes have different proportions of component proteins: They act as scaffolds to assemble variable numbers of other regulatory proteins/complexes

Proteins are often made exactly where they are needed in the cell

70% of mRNAs have striking subcellular localisations in *Drosophila* embryos



Lecuyer et al. (2007) Cell, 131, 174



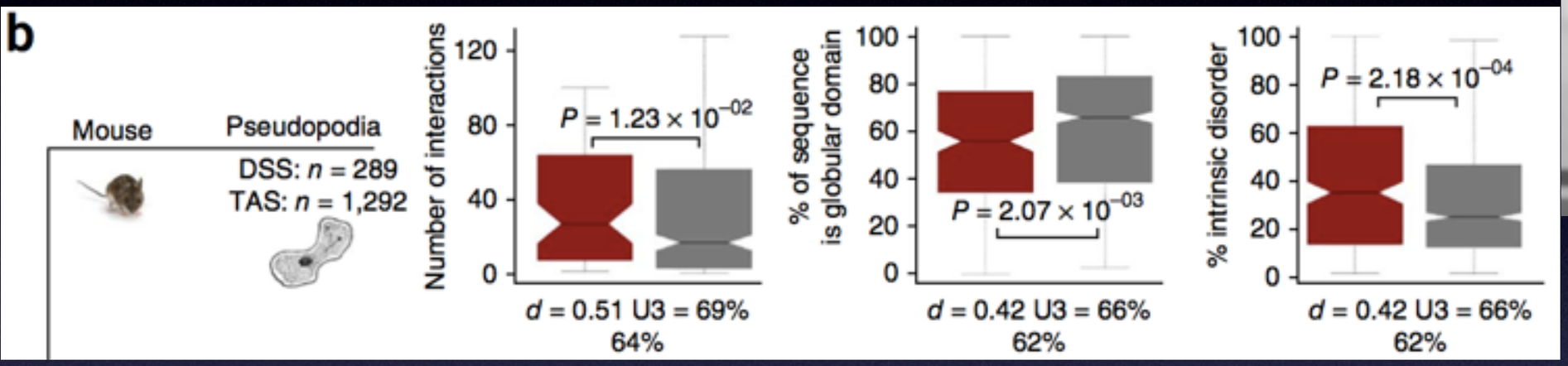
Martin and Ephrussi (2009) Cell, 136, 719

Some examples of localised mRNAs involved in spatially regulated translation

Asymmetric mRNA localization contributes to fidelity and sensitivity of spatially localized systems

Robert J Weatheritt[1,3], Toby J Gibson[2] & M Madan Babu[1]

nature structural & molecular biology

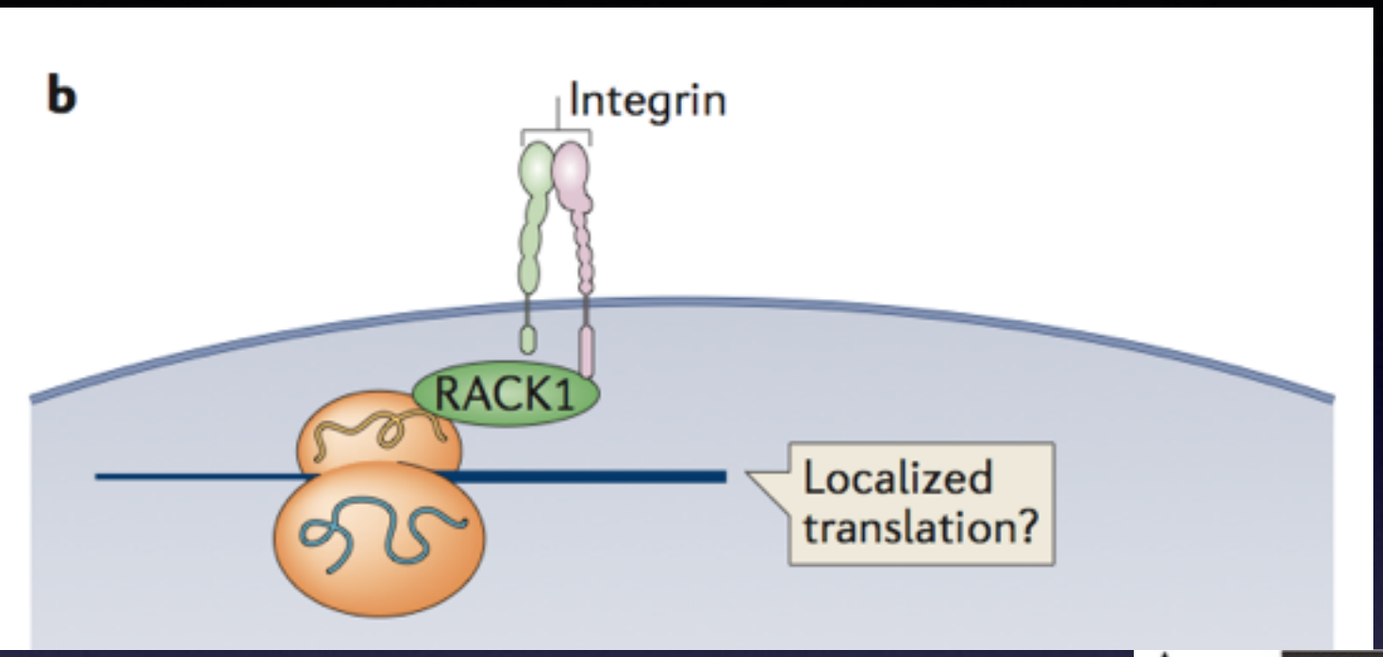


Robert Weatheritt, PhD, now in Toronto with Ben Blencowe

mRNAs in pseudopodia encode proteins enriched for intrinsic disordered regions

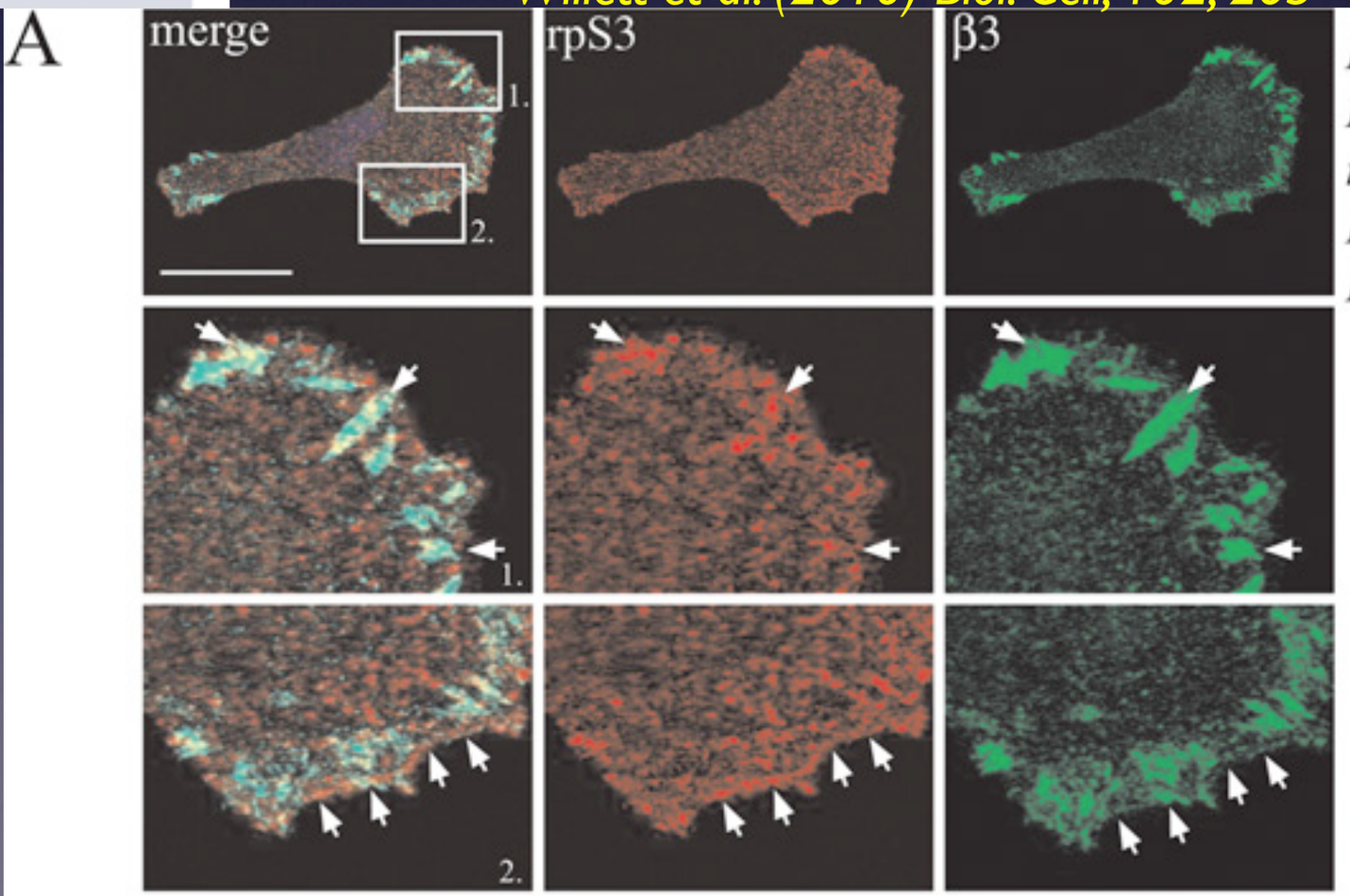
Proteins synthesised on-site often provide essential components required to activate the signalling machinery. They also tend to encode proteins that have the capacity to nucleate and form reversible, non-membranous assemblies

Ribosomal subunits colocalise with beta3 integrin at adhesion foci at the leading edge of migrating fibroblasts



Xue and Barna (2012) Nat Rev MCB, 13, 355

Willett et al. (2010) Biol. Cell, 102, 265



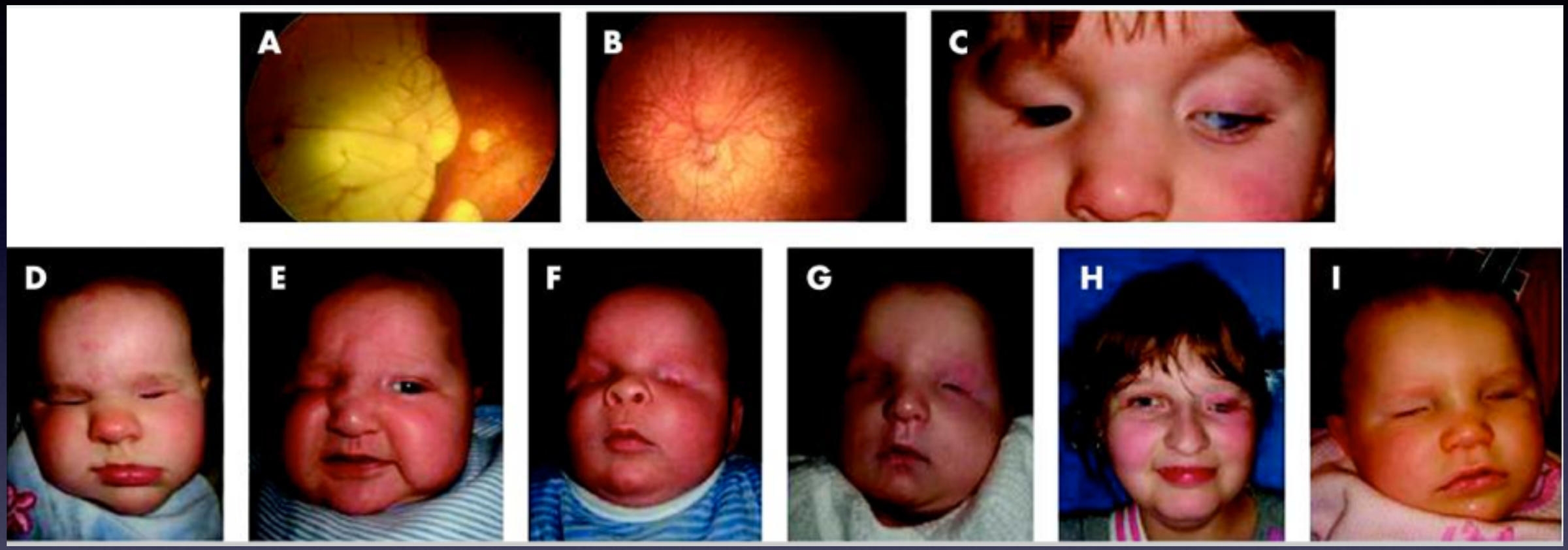
40S subunits are enriched at FACs

Spatial regulation of translation - implications

- Making proteins in the wrong place is often a bad thing
- Cells have been under continual selection pressure to develop systems for precise mRNA targeting

How many proteins can be allowed to freely diffuse in the cell?

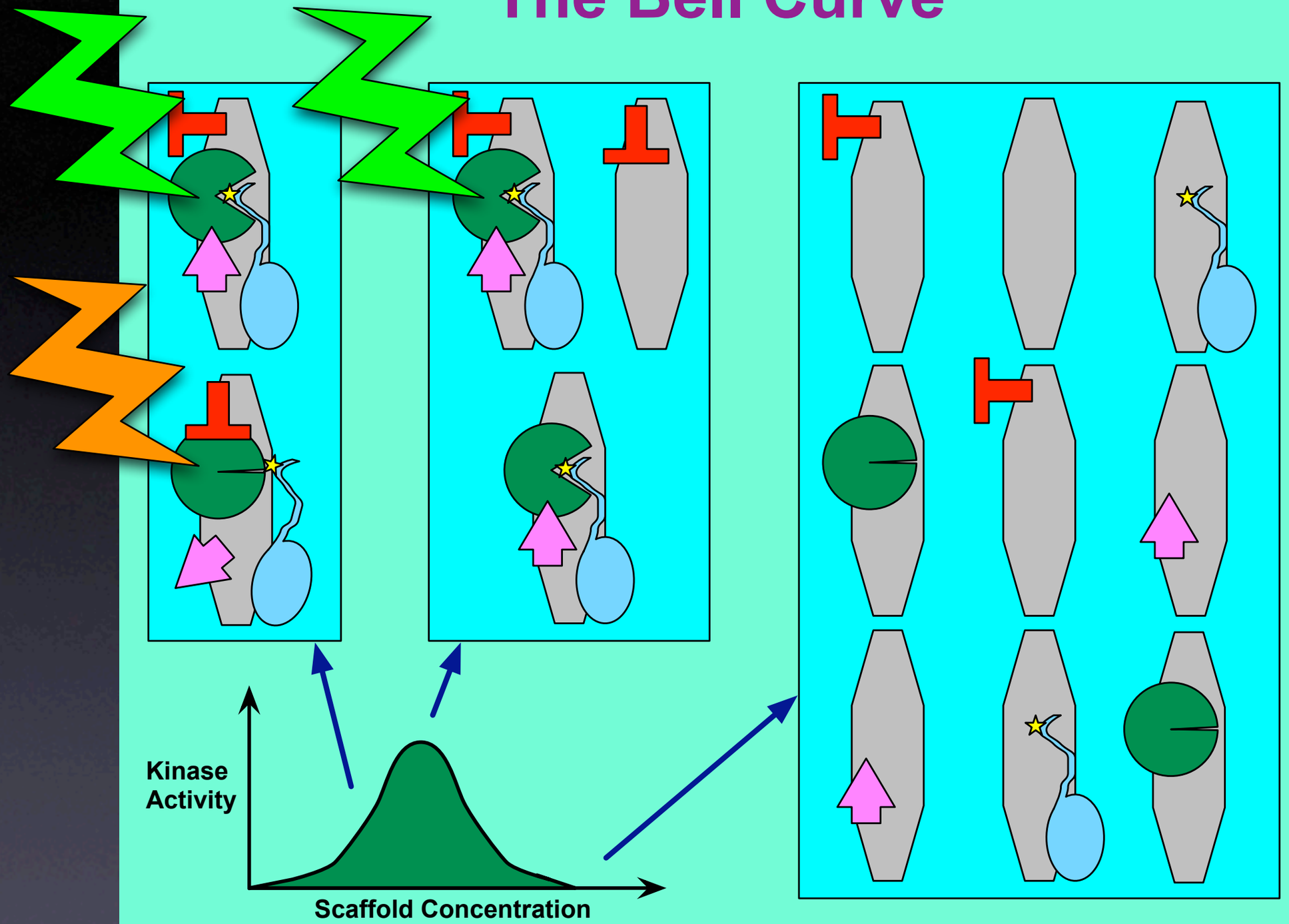
Sox2, Oct4 and Nanog are key stem cell genes



Sox2 haploinsufficiency leads to aniridia

Phenotypes can often give a misleading view of protein function. They highlight the strongest point of failure.

The Bell Curve



effect of KSR varied dramatically with the level of KSR protein expressed. In *Xenopus* oocytes, KSR functioned as a positive regulator of Ras signaling when expressed at low levels, whereas at high levels of expression, KSR blocked Ras-dependent signal transduction. Likewise, overexpression of *Drosophila* KSR blocked R7 photore-

Cacace et al. (1999) MCB, 19, 229

Many components of regulatory complexes exhibit balanced gene dosage

It is not just scaffolds: Foxc1 and Pax6 are, like Sox2, TFs that cannot tolerate dosage alteration in any direction during eye development

The transience of transient overexpression

Toby J Gibson, Markus Seiler & Reiner A Veitia

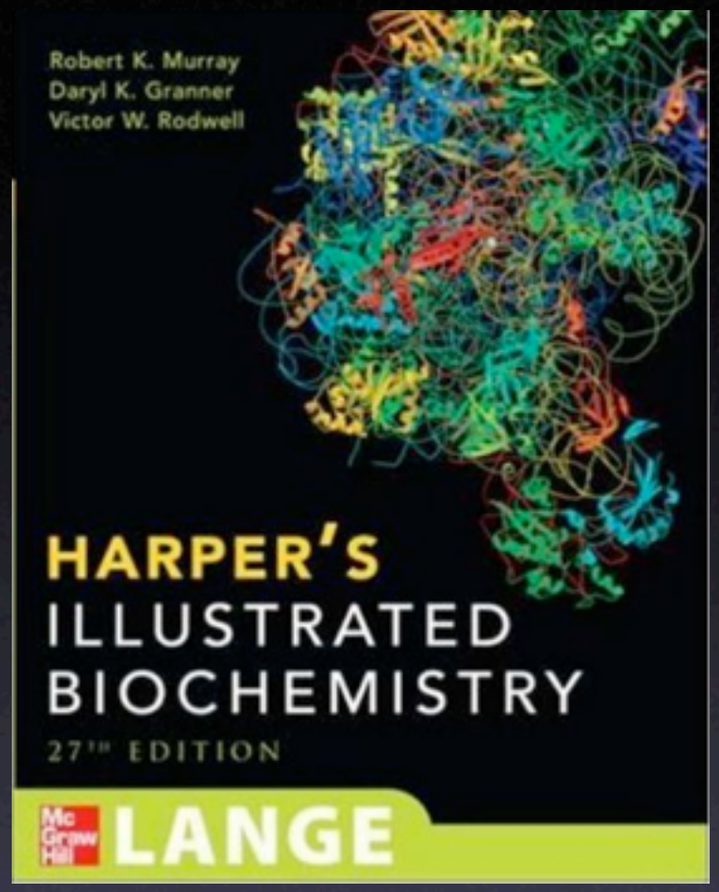
Much of what is known about mammalian cell regulation has been achieved with the aid of transiently transfected cells. However, overexpression can violate balanced gene dosage, affecting protein folding, complex assembly and downstream regulation. To avoid these problems, genome engineering technologies now enable the generation of stable cell lines expressing modified proteins at (almost) native levels.

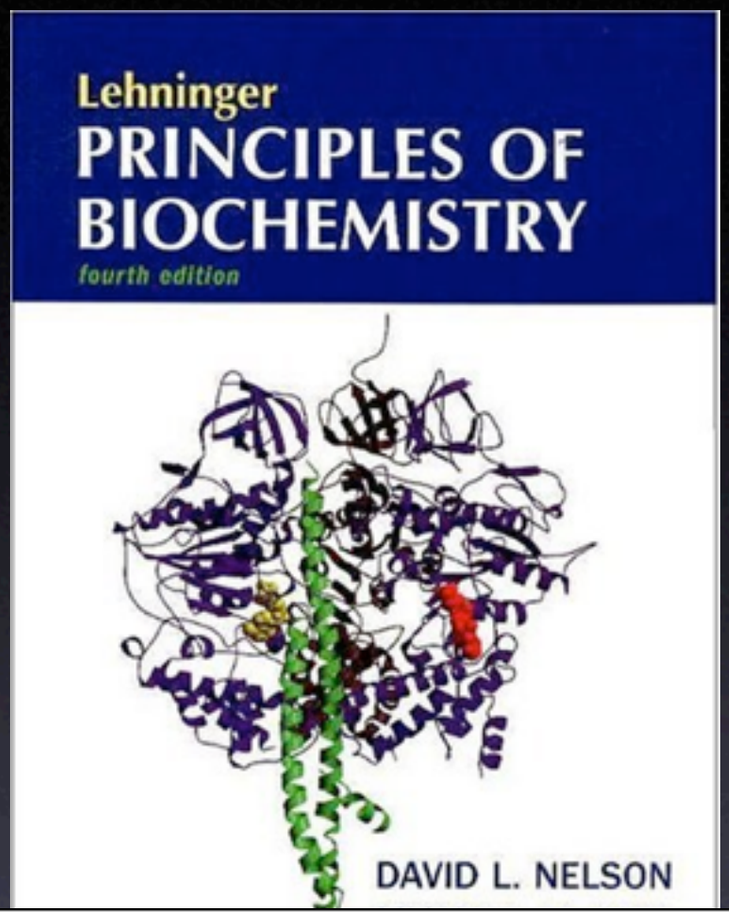
Nature Methods (2013) NCB 10, 715

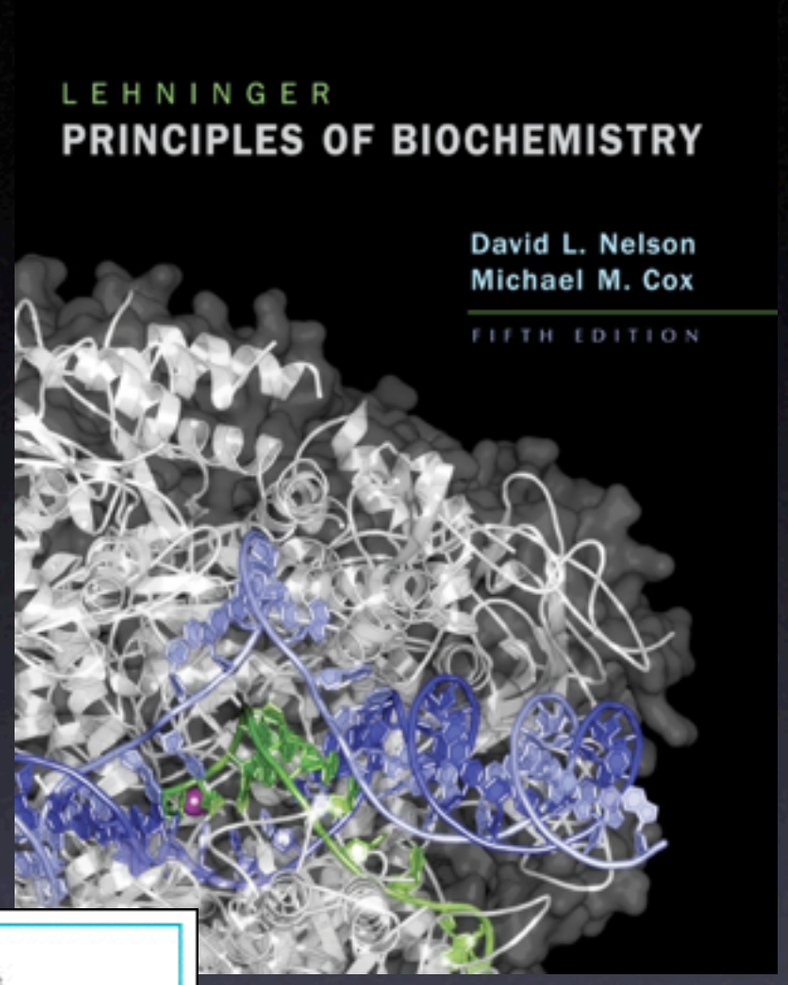
Table 2. Contrasting issues with transient overexpression experiments relative to native expression

Features of Cell Regulation / Effect on Experiment	**Over Expression**	**Native Expression**
Low molecule number (*e.g.* <1000 per cell)	X	√
Spatially arranged protein	X	√
Coupled mRNA transport / Spatial translation	Overload system	√
Mutants that are (unknowingly) unfolded	Amyloid/aggregation	?
Balanced gene dosage of regulators	X	√
Kinases and their substrates are scaffolded	X	√
Laser bleaching to study diffusion (or other motion) of a signalling protein	Meaningless	√
Protein complex by Co-IP	???	√
Proteomics	X	√
Reproducibility	??	√
Synchronised cell population	X	√
Differentiate from stem cell	X	√

Biochemistry Text Books







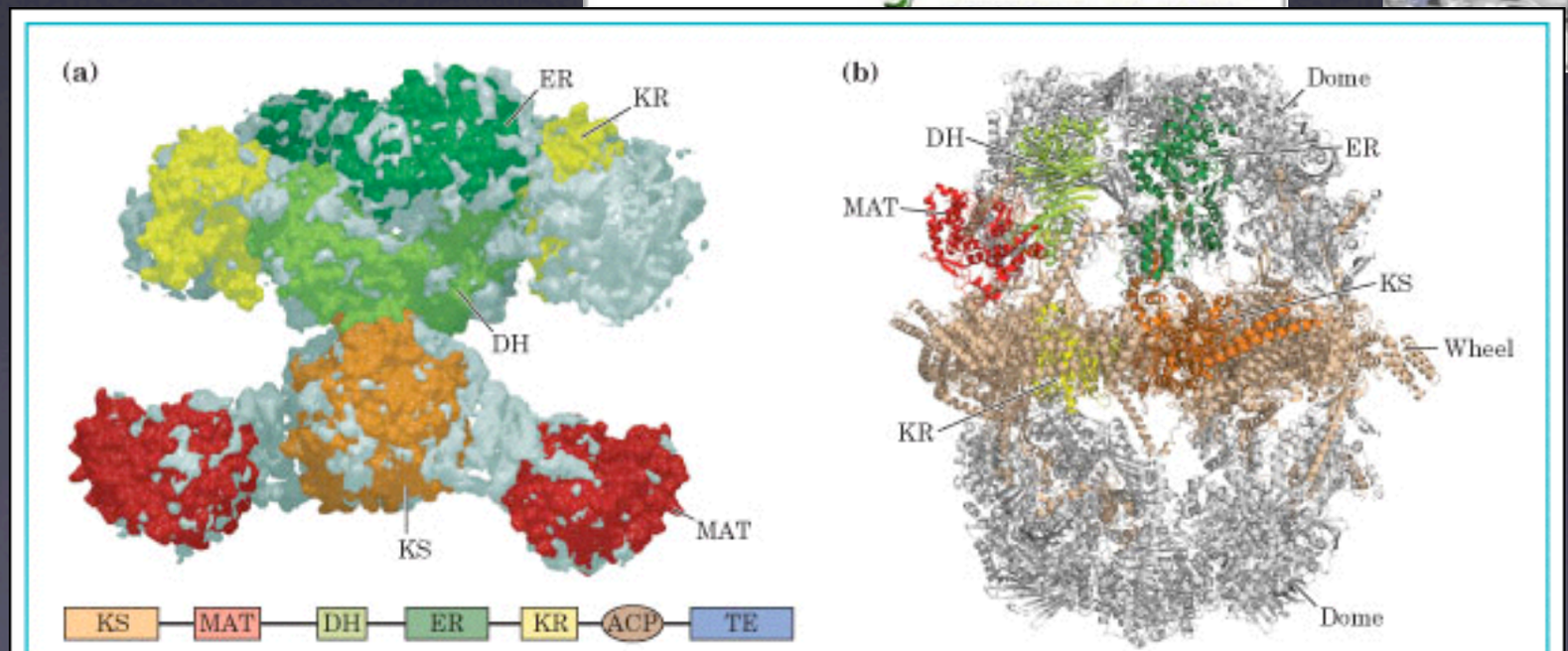


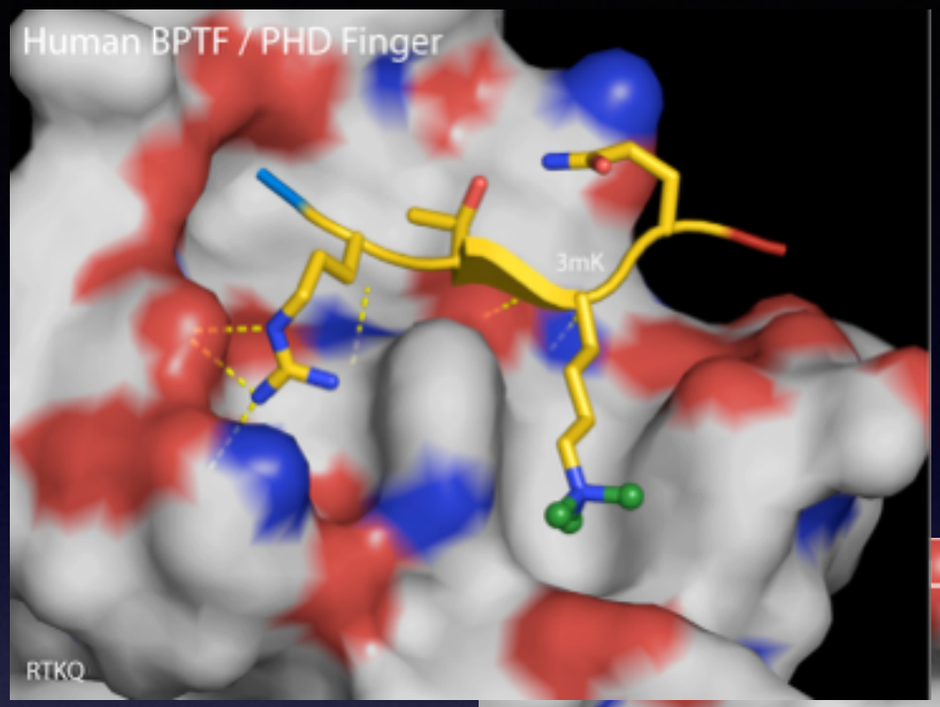
FIGURE 21-3 The structure of fatty acid synthase type I systems.

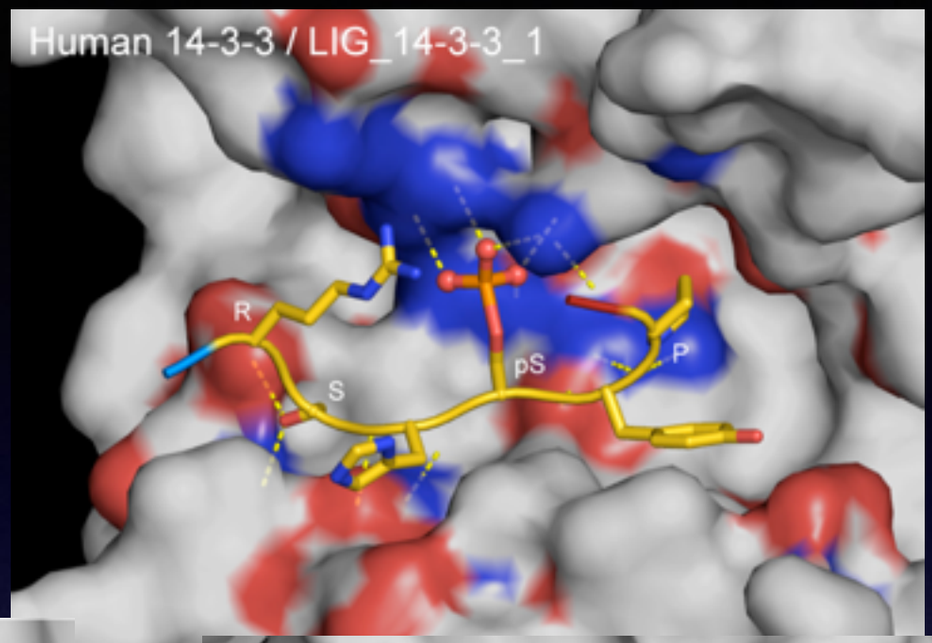
Complexes
Complexes
Complexes

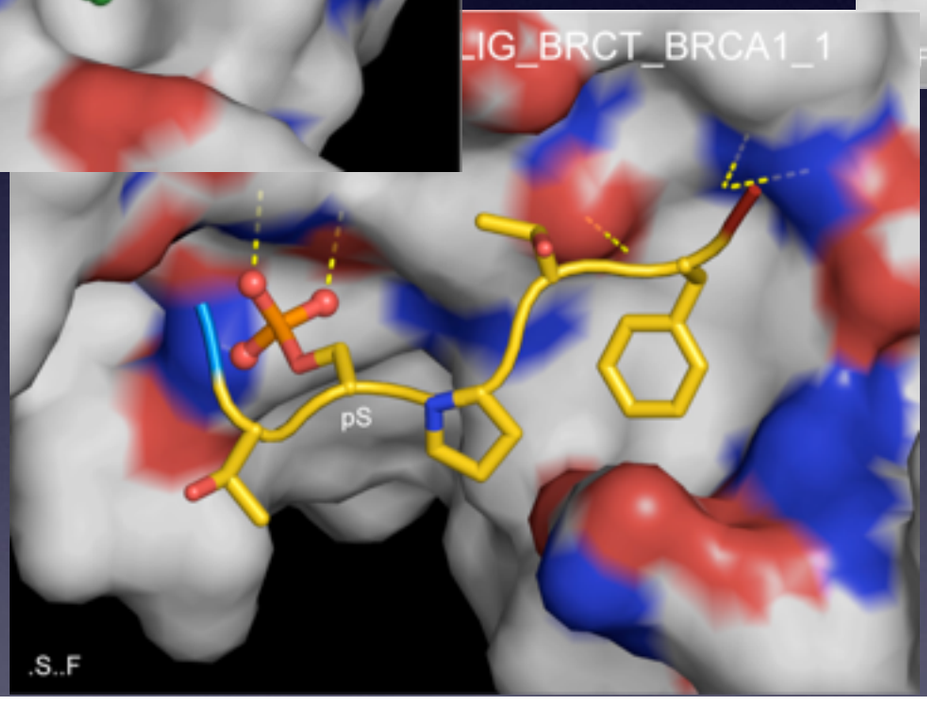
Truth and clarity are complementary

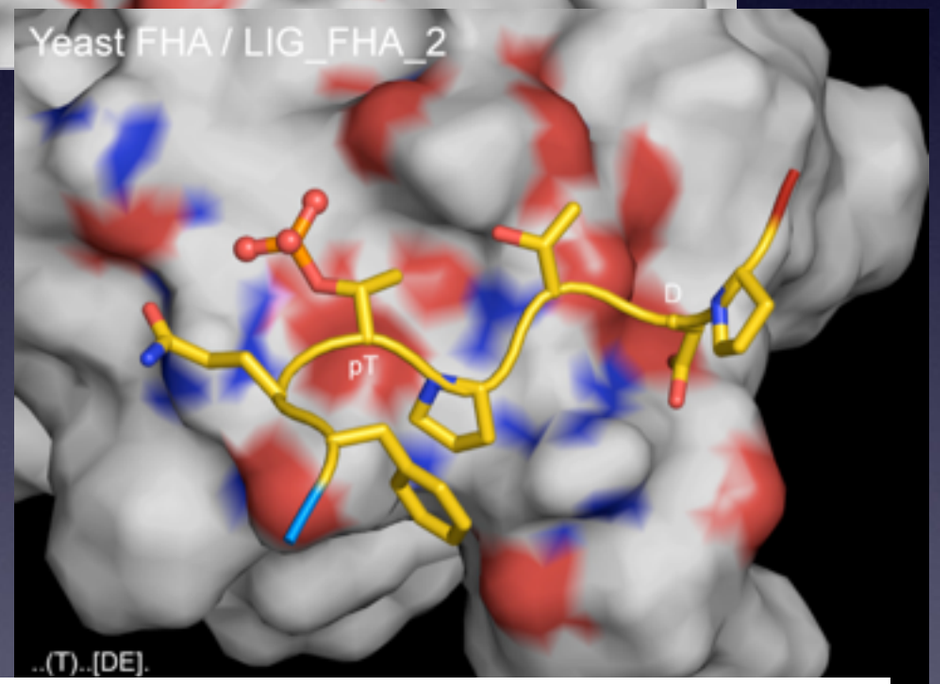
Niels Bohr

Biochemistry books are not so good on regulatory interactions









Understanding eukaryotic linear motifs and their role in cell signaling and regulation

Francesca Diella[1], Niall Haslam[1], Claudia Chica[1], Aidan Budd[1], Sushama Michael[1], Nigel P. Brown[2], Gilles Trave[3] Toby J. Gibson[1]

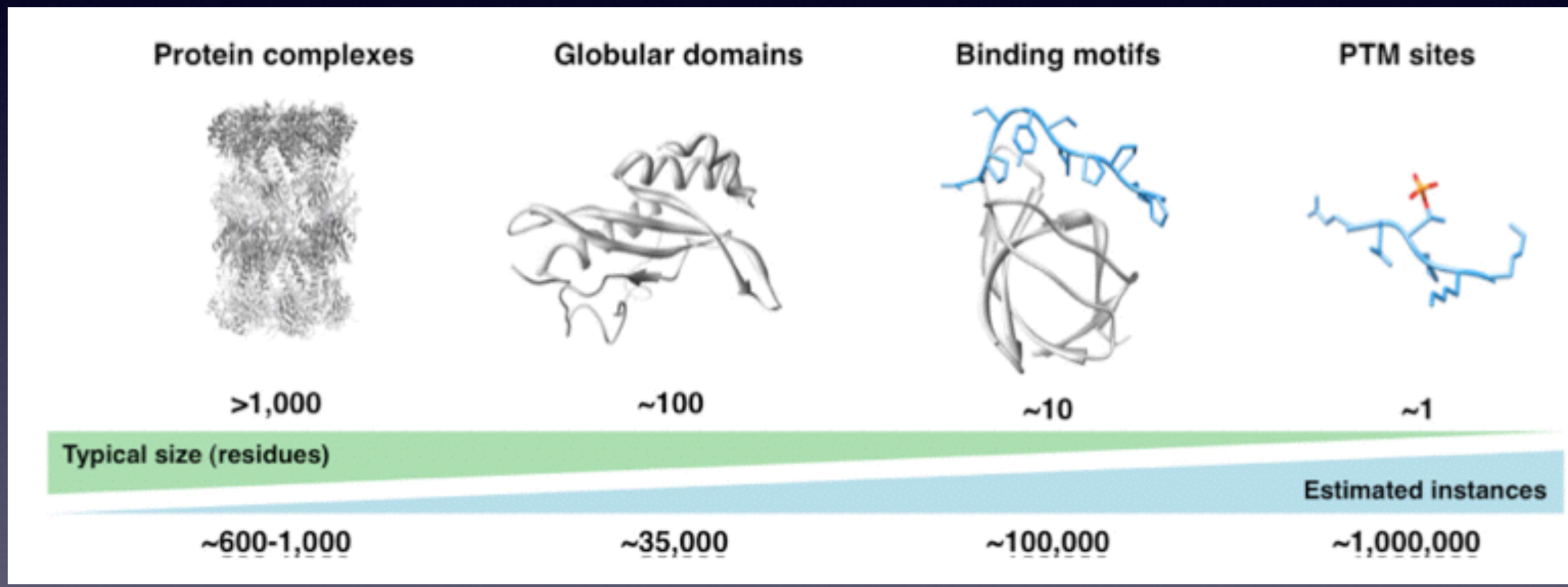
[1]Structural and Computational Biology Unit, European Molecular Biology Laboratory, 69117 Heidelberg, Germany, [2]BIOQUANT, Ruprecht-Karls-Universität Heidelberg, Im Neuenheimer Feld 267, 69120 Heidelberg, Germany, [3]ESBS, 1, Bld Sébastien Brandt, BP10413, 67412-ILLKIRCH, France 3

24 page open access review in Frontiers in Biosciences

Molecular Cell
Review
Molecular Cell 55, July 17, 2014

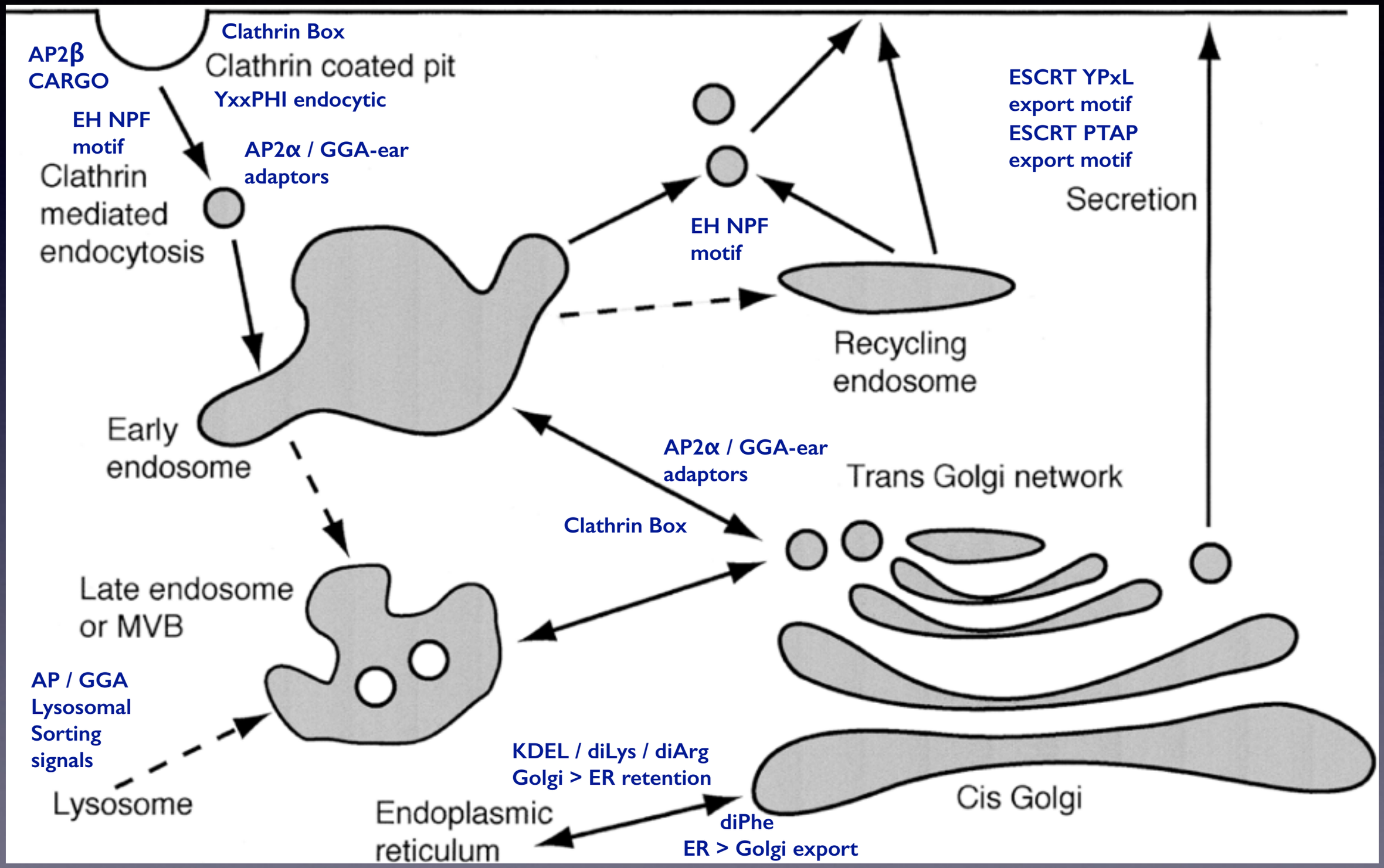
A Million Peptide Motifs for the Molecular Biologist

Peter Tompa,[1,2,*] Norman E. Davey,[3] Toby J. Gibson,[4] and M. Madan Babu[5,*]



Vesicle trafficking in the cell

The cell has to control the movement of subcellular organelles.
Complex and dynamic systems require extensive regulation.

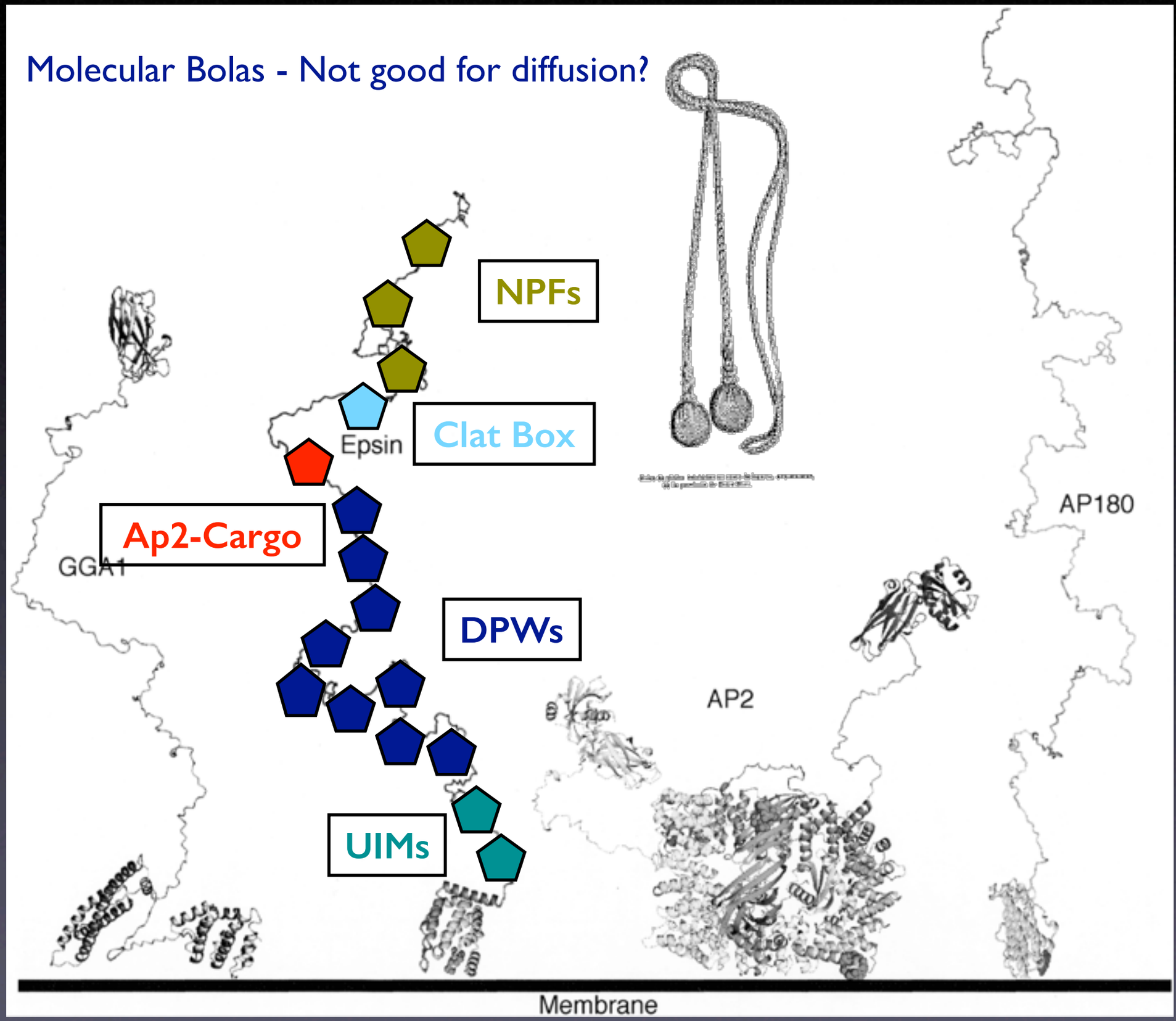


Owen DJ (2004) Biochem. Soc.Trans. 32, 1-14

Modular regulatory proteins involved in endocytosis

Most Endocytosis proteins have a mixture of **globular domains** and **natively disordered** regions. The disordered regions are proving to be rich in **Linear Motifs.**

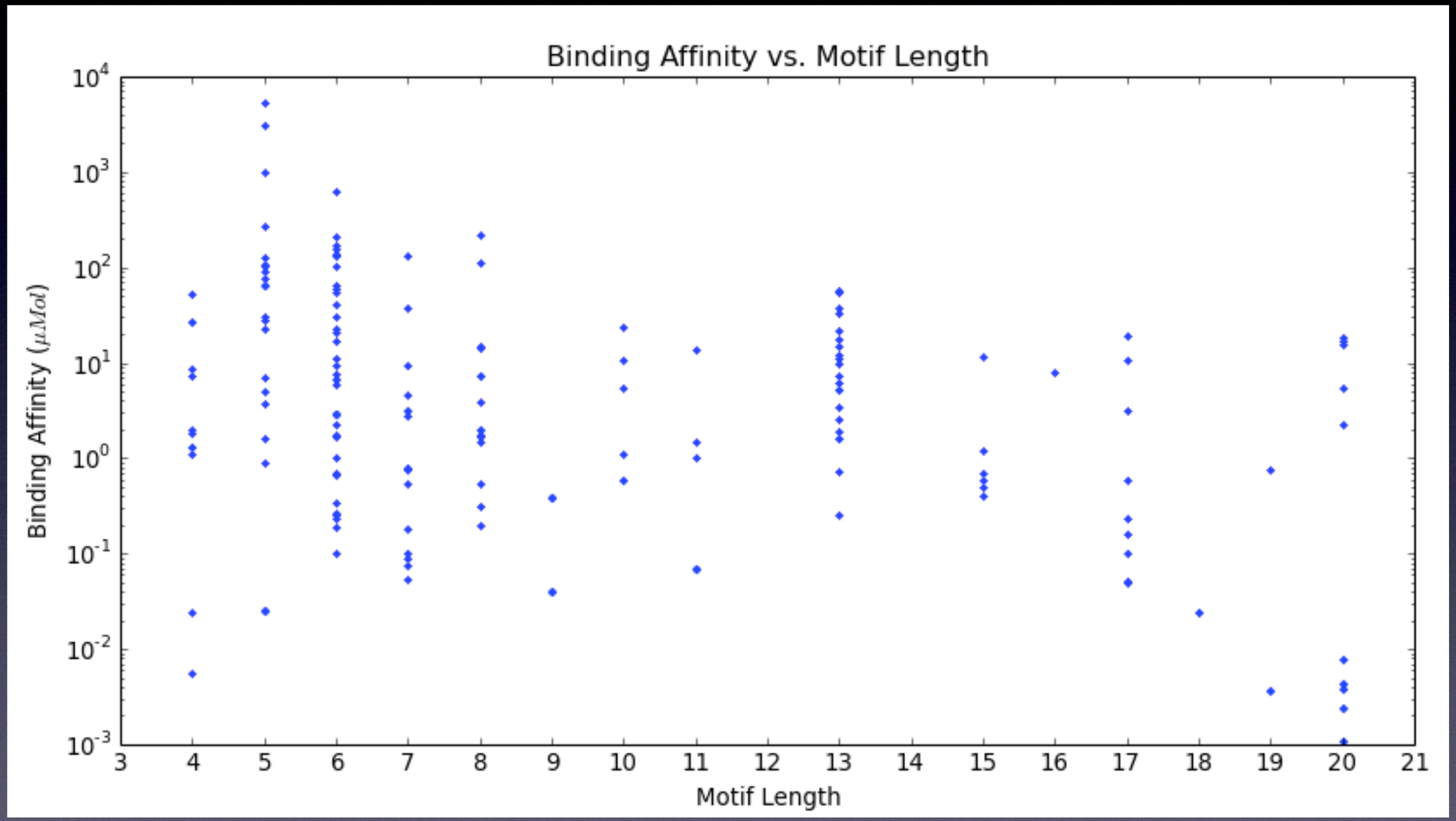
Here the disordered regions are shown to scale with respect to the globular domains



Owen DJ (2004) Biochem. Soc. Trans. 32, 1-14.

Binding Affinity vs. Motif Length

http://elm.eu.org



K_d (μM)

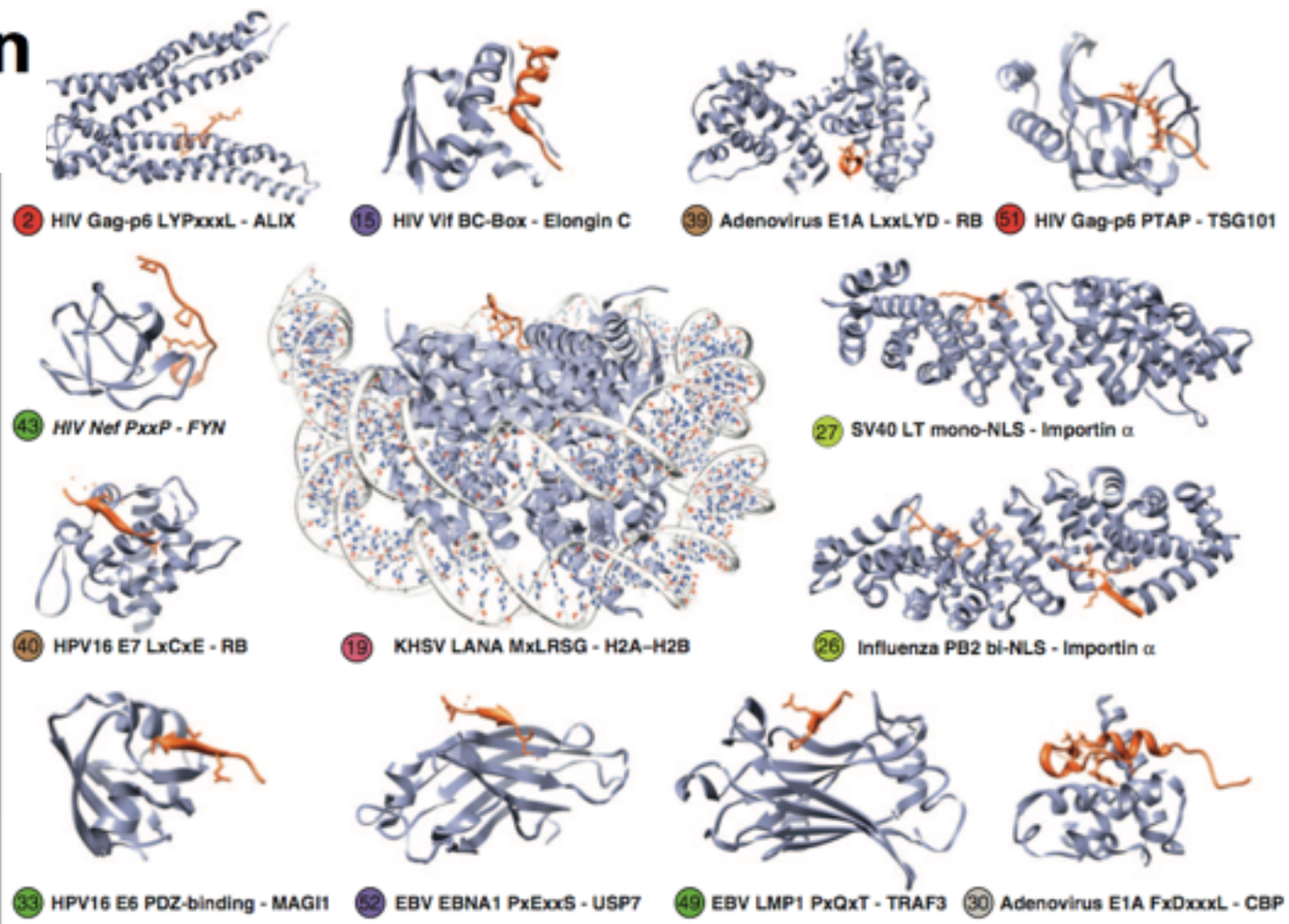
Motif length (aa residues)

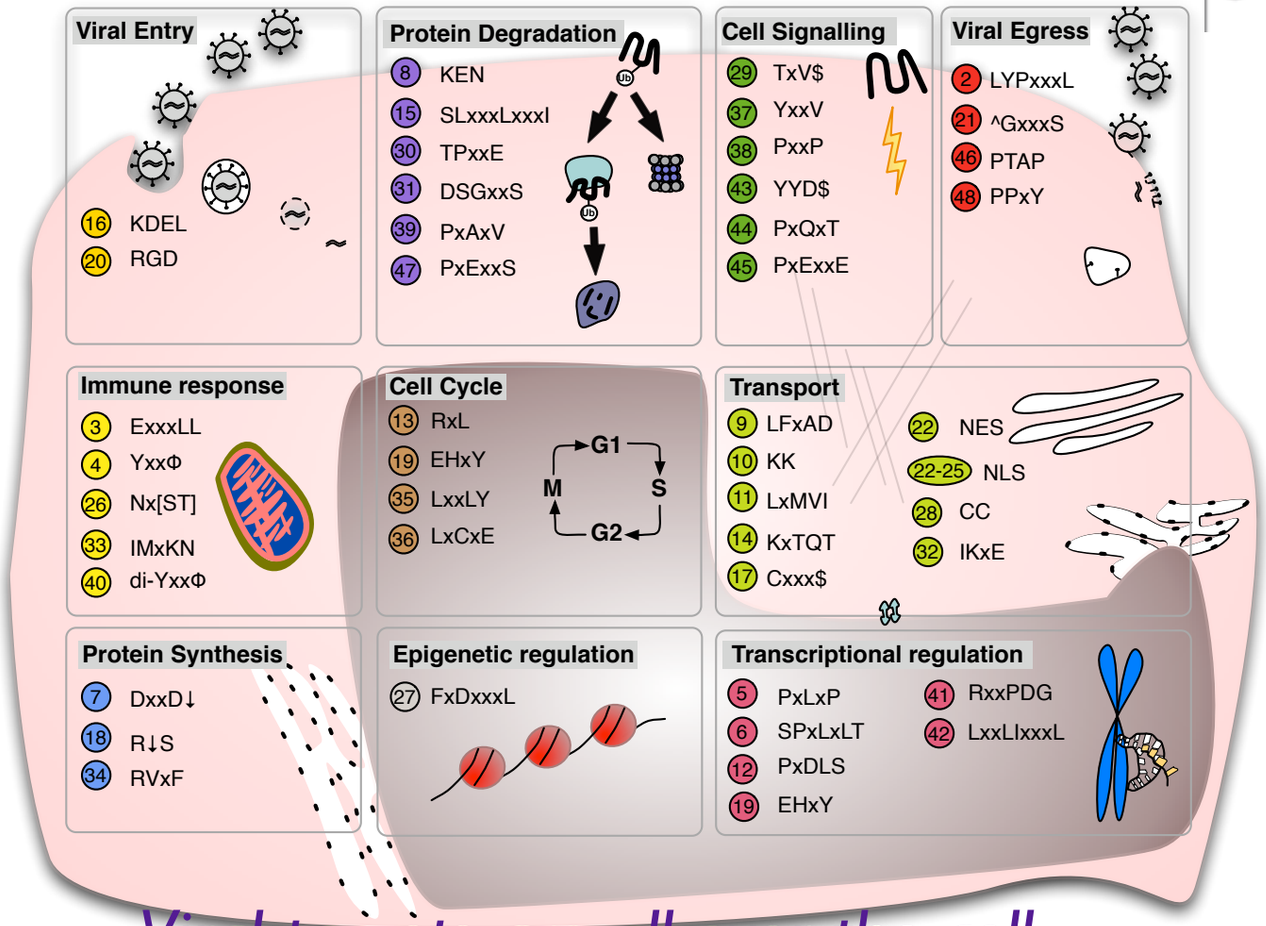
How viruses hijack cell regulation

Norman E. Davey[1], Gilles Travé[2] and Toby J. Gibson[1]

TiBS (2011) 36, 159

More than a third of the motif classes annotated in our ELM Resource (http:elm.eu.org) are already known to be used by viruses

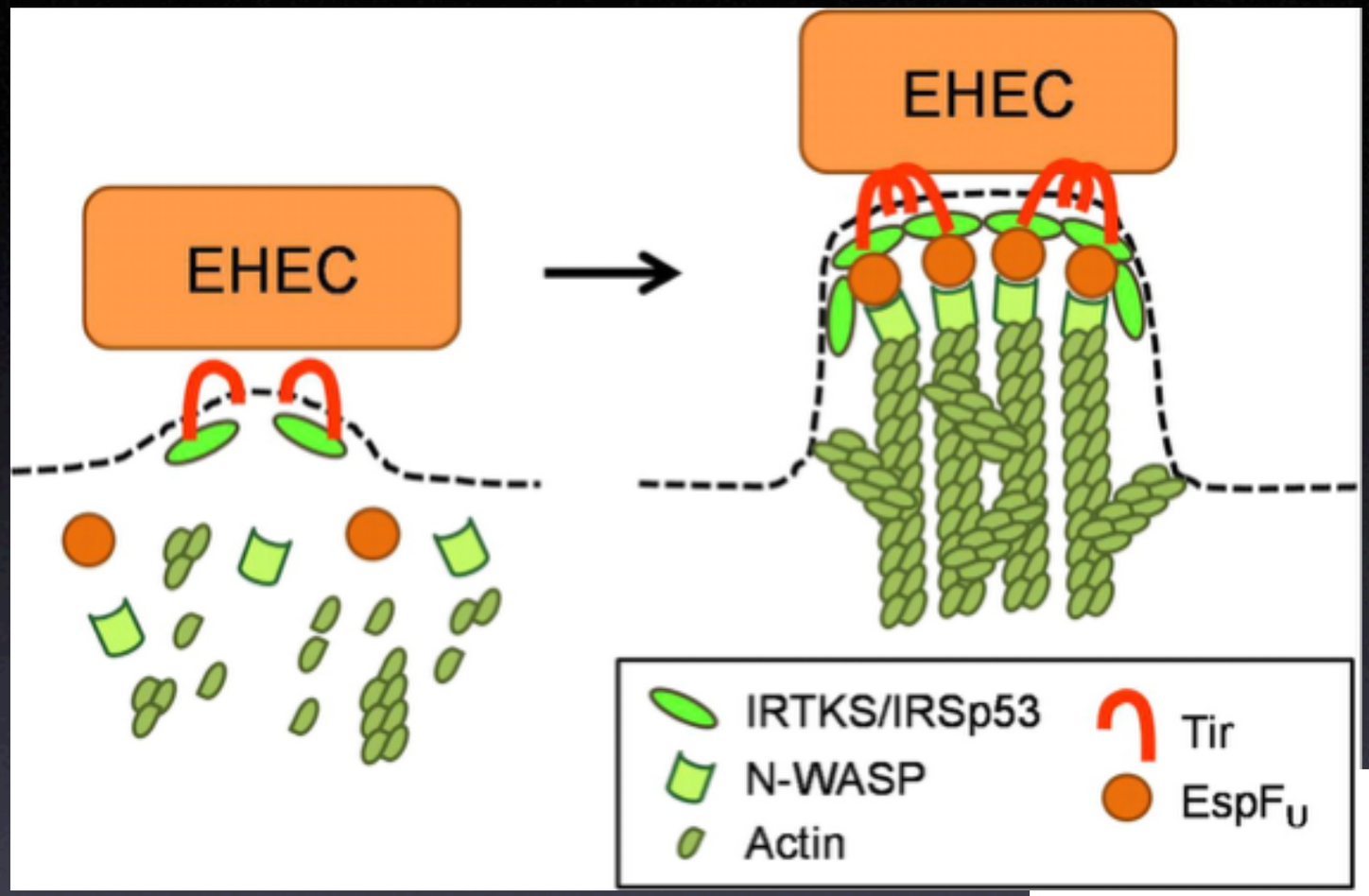




Viral targets are all over the cell

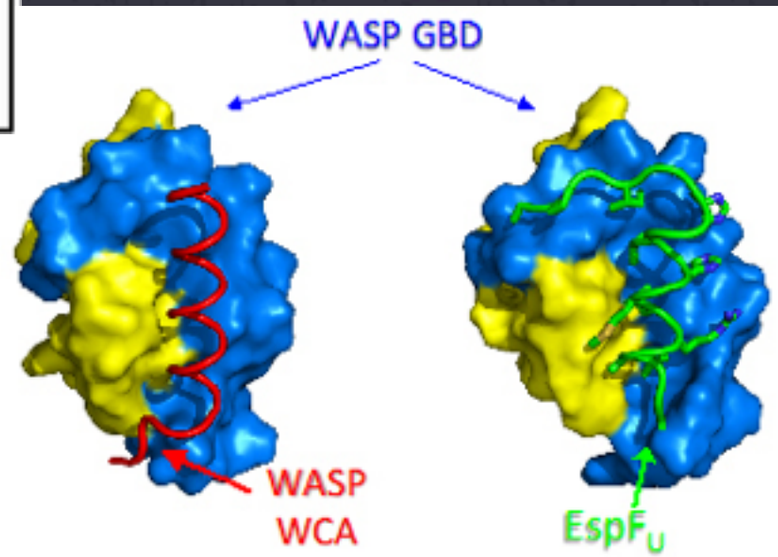
- Why is there "always" a cellular protein motif interaction for a virus to subvert?
- What does this tell us about the nature of the cell?

Pathogenic Pedestal Formation



Yi PNAS, 106, 6431 (2009)

A linear motif in *E. coli* EHEC EspFu binds N-WASP leading to Actin polymerisation



Cheng, Nature, 454, 1009 (2008)

Cell Regulation: Cooperative and Spatially arranged

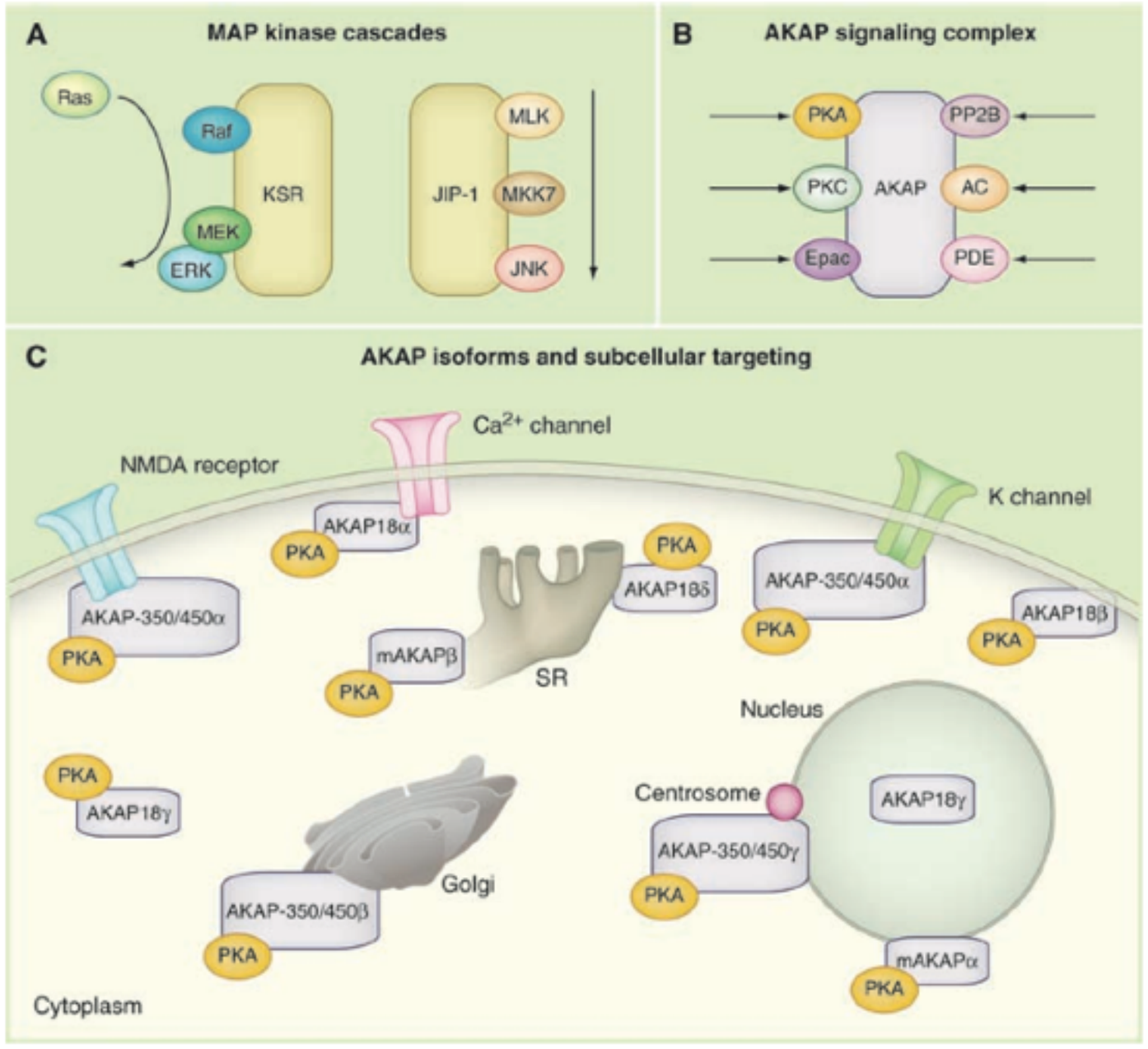
Spatial Cell Biology

REVIEW

Cell Signaling in Space and Time: Where Proteins Come Together and When They're Apart

John D. Scott[1]* and Tony Pawson[2,3]*

Science, 326, 2009



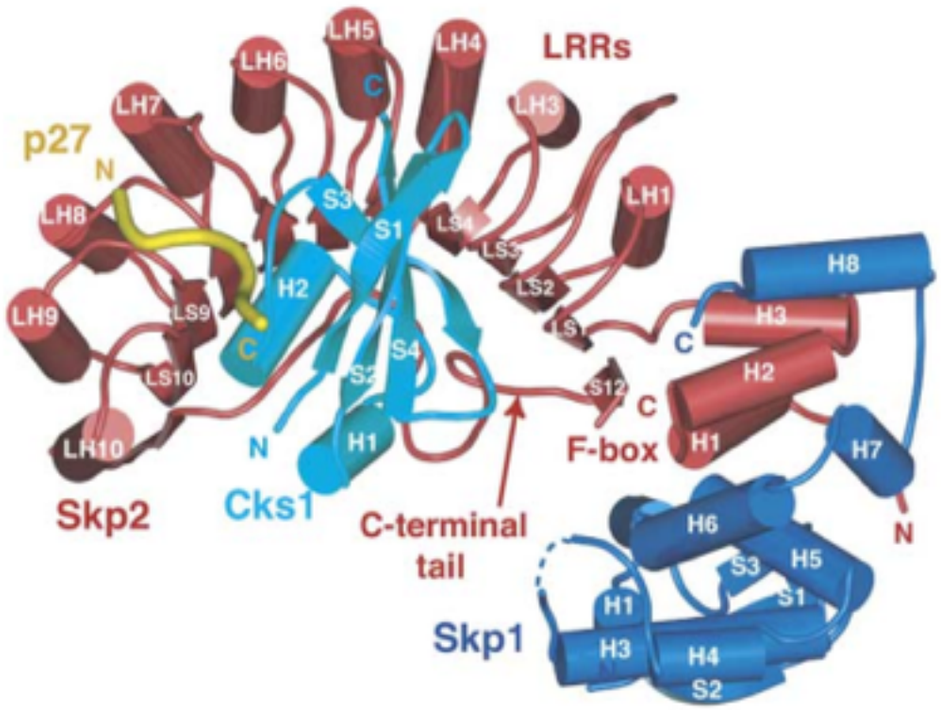
Tony Pawson, Cell (2004)

While there is still much debate about these ideas, the spatial segregation of signaling pathways is likely to be an important topic for the future.

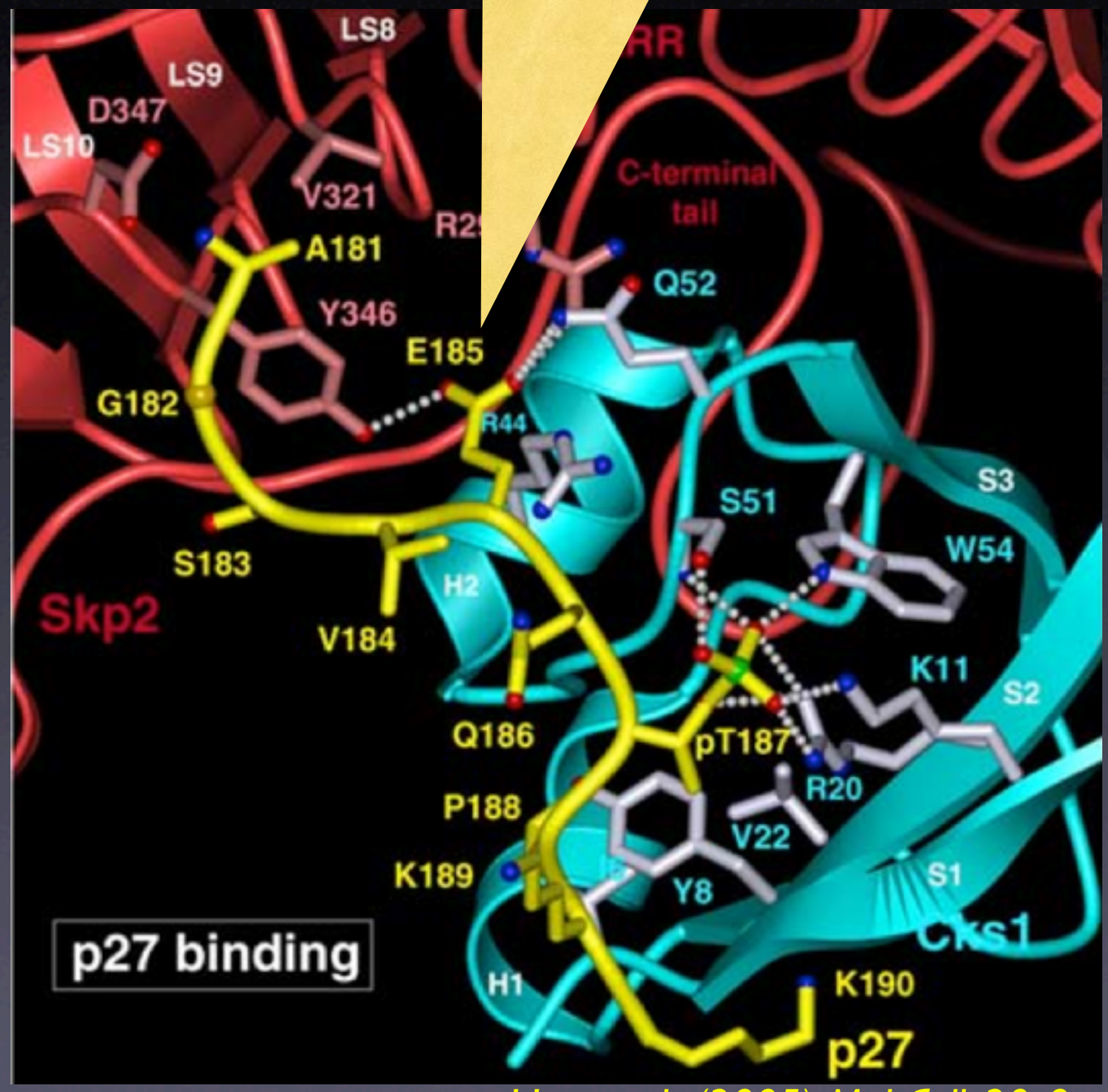
Cooperativity by preassembly: P27kip1 phosphorylated motif bound by a complex of Skp1-Skp2-Cks1

C

```
                                  181                 190
p27 human      175 SDGSPNAGSVEQTPKKPGLRRRQT
p27 pig        175 SDGSPNSASVEQTPKKPGLRRRQT
p27 mouse      175 SDGSPNAGTVEQTPKKPGLRR-QT
p27 duck       175 SEDSPSASSVEQTPKKSSPRRHQT
p27 chicken    175 SEDSPSASSVEQTPKKSSPRRHQT
p27 hamster    175 SDGSLNAGSVEQTPKKPGLRRHQT
p27 xenopus    192 TKGVHLLCPLEQTPRKK-IR
```



Glu185 is bound co-operatively by Skp2 and Cks1

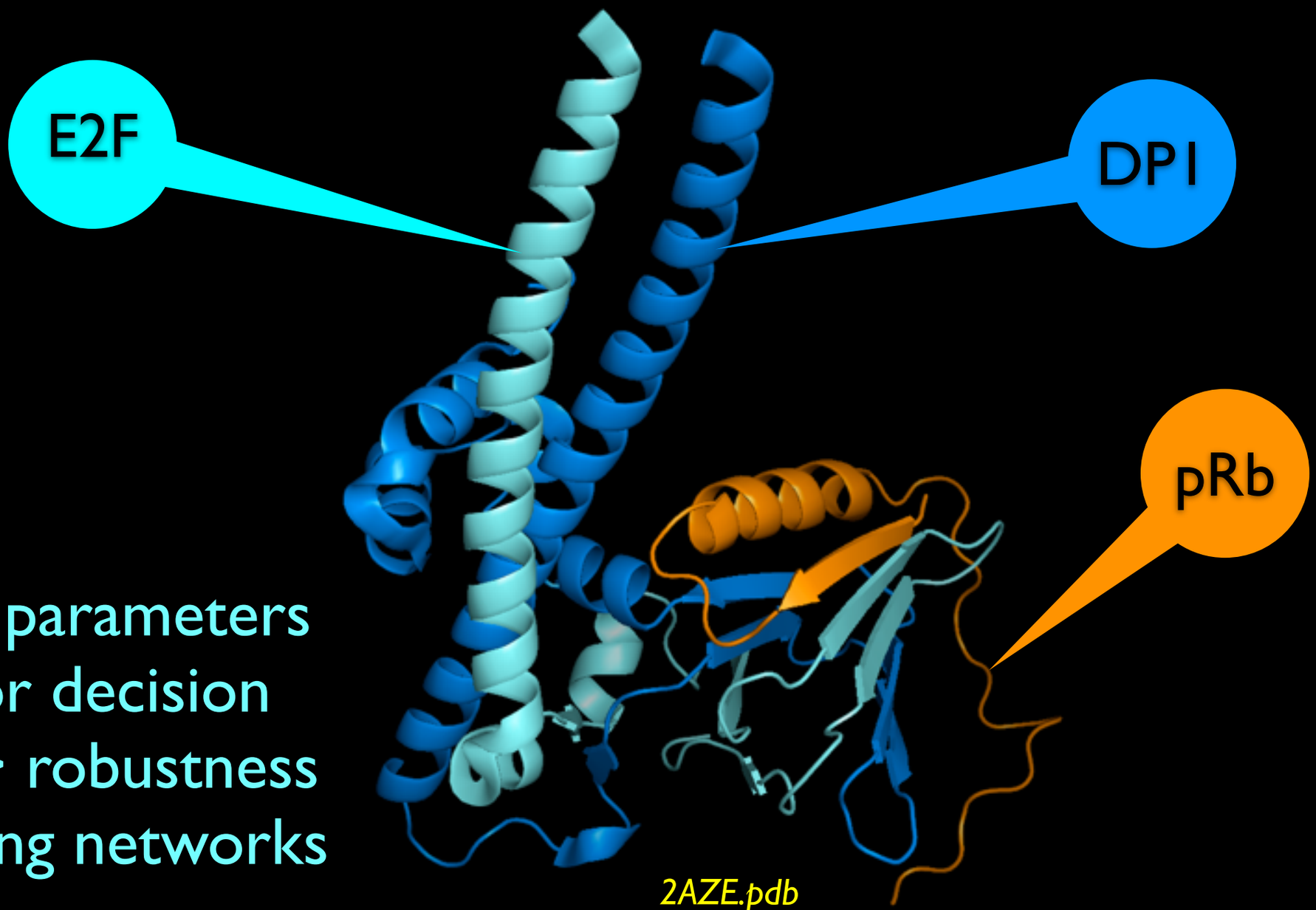


Hao et al., (2005) Mol Cell. 20, 9

Cooperativity of IDRs - Intrinsically Disordered Regions

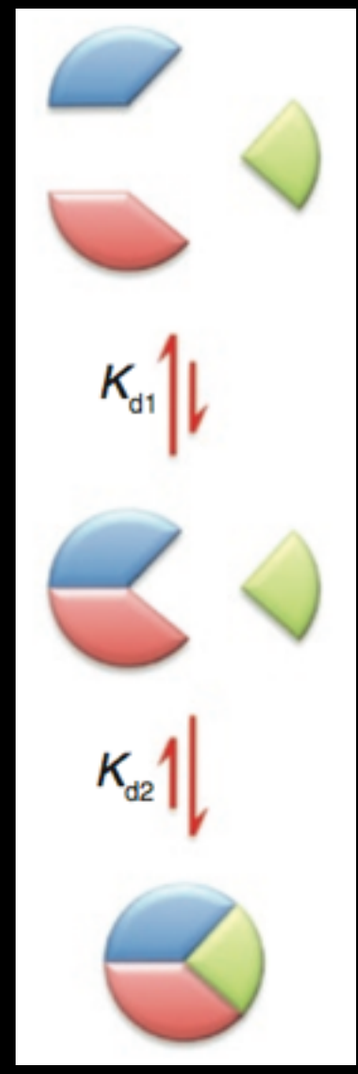
Regulation by cooperative assembly of E2F1, DP1 and Rb

Mutual induced fit assembly of a repressive heterotrimer from three natively disordered protein segments



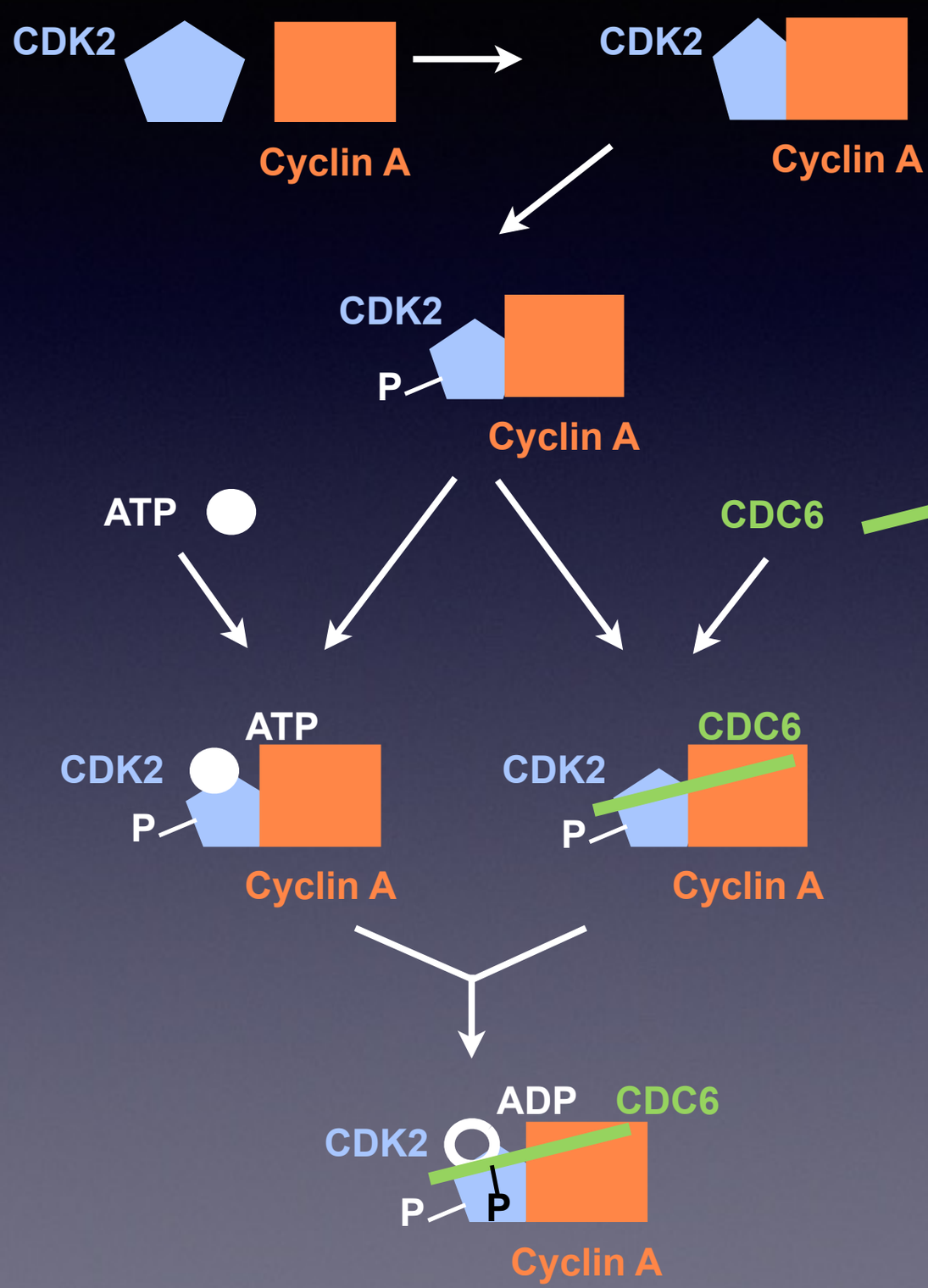
2AZE.pdb
Rubin et al. (2005) Cell 123, 1093

Multiple parameters used for decision making > robustness in signalling networks



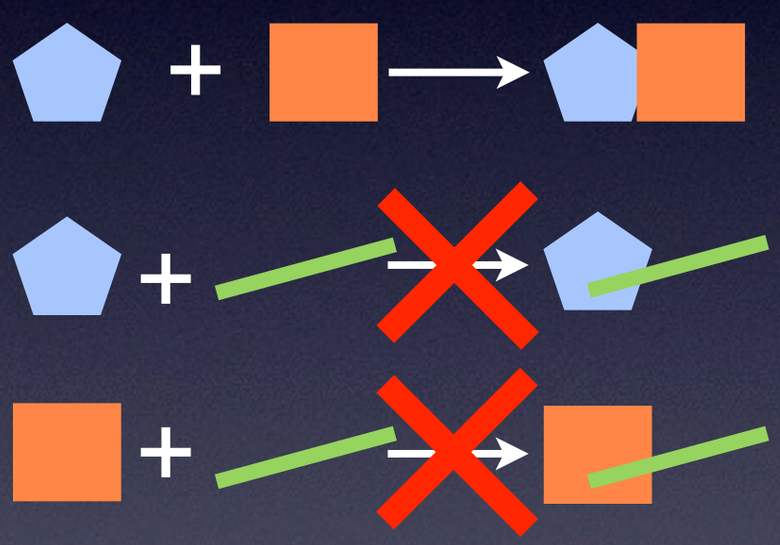
Whitty (2008)
NCB, 4, 435

Phosphorylation of CDC6 by Cdk2-CyclinA



How Bioinformatics interaction standards work: Capturing Phosphorylation of CDC6 by Cdk2-CyclinA

Current representation
Binary Interactions

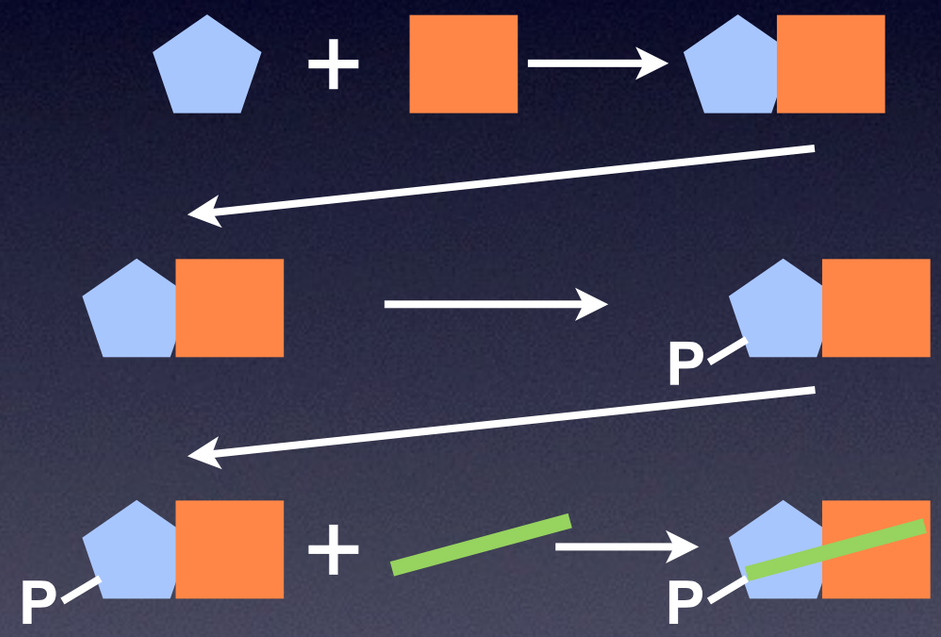


Binary

Distinct

Independent

Desired representation
Cooperative Interactions



Multivalent

Allosteric

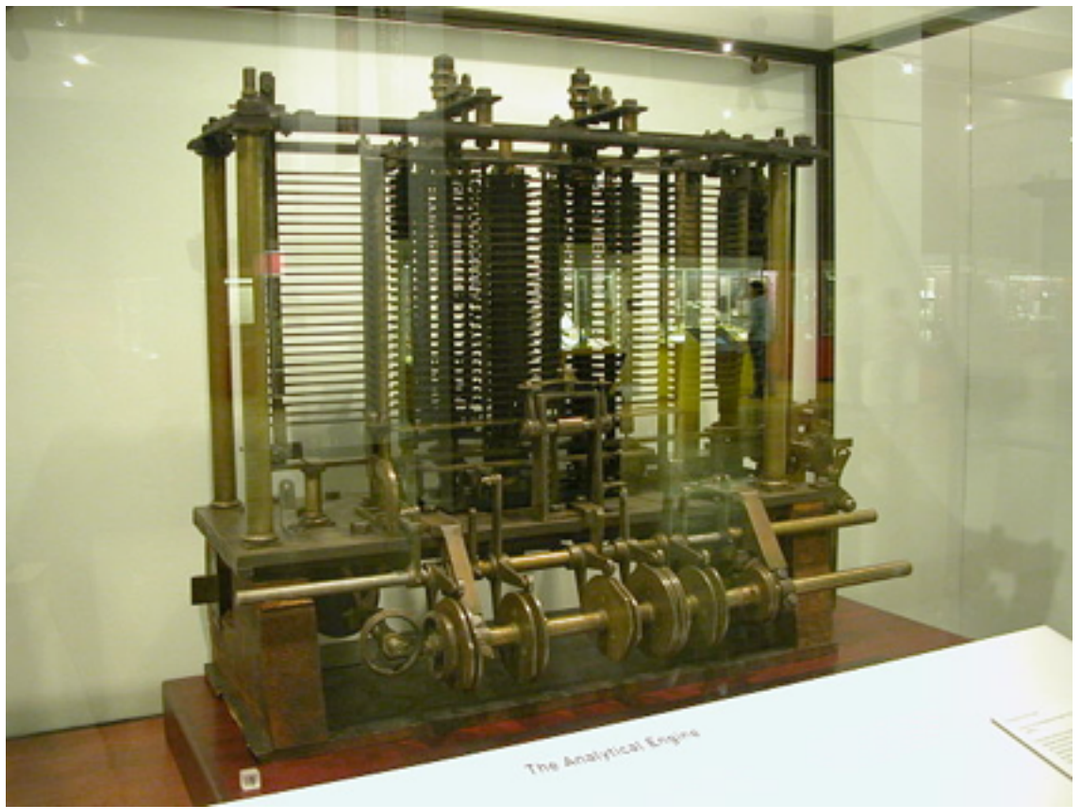
Interdependency of binding events

Allostery

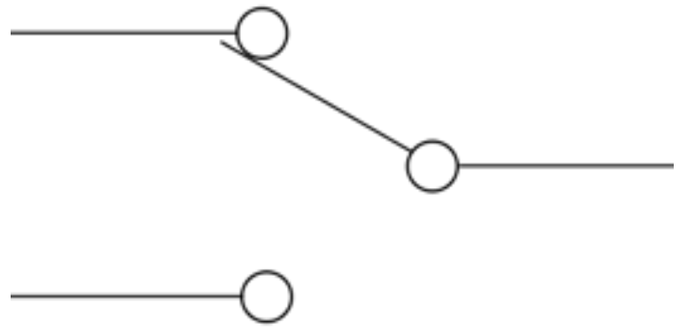
“The second secret of life”

Jacques Monod

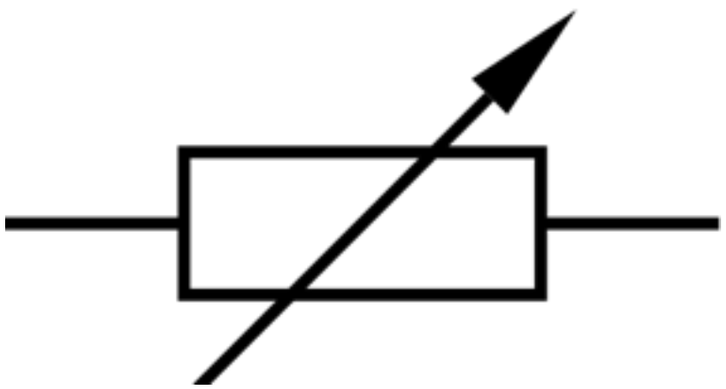
Logic processing is always done by machines with switches



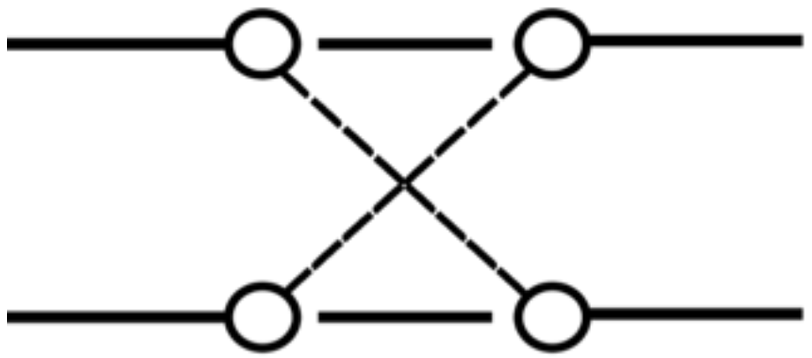
Babbage analytical engine



3-way switch



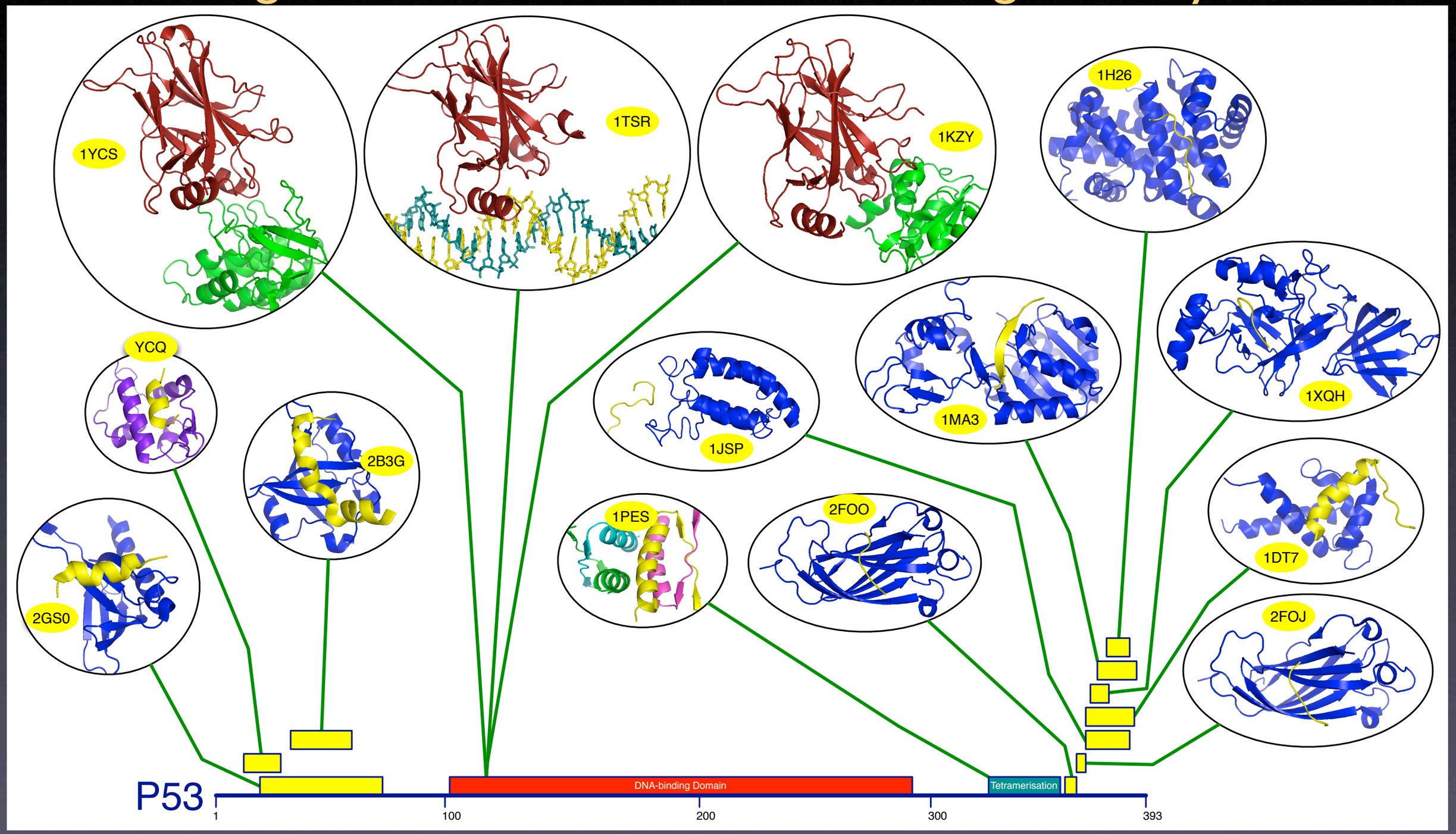
Rheostat



4-way switch

Molecular switching with P53

IUP makes more interactions than Globdom / Mutually exclusive binding / Alternative conformations / Regulated by PTM



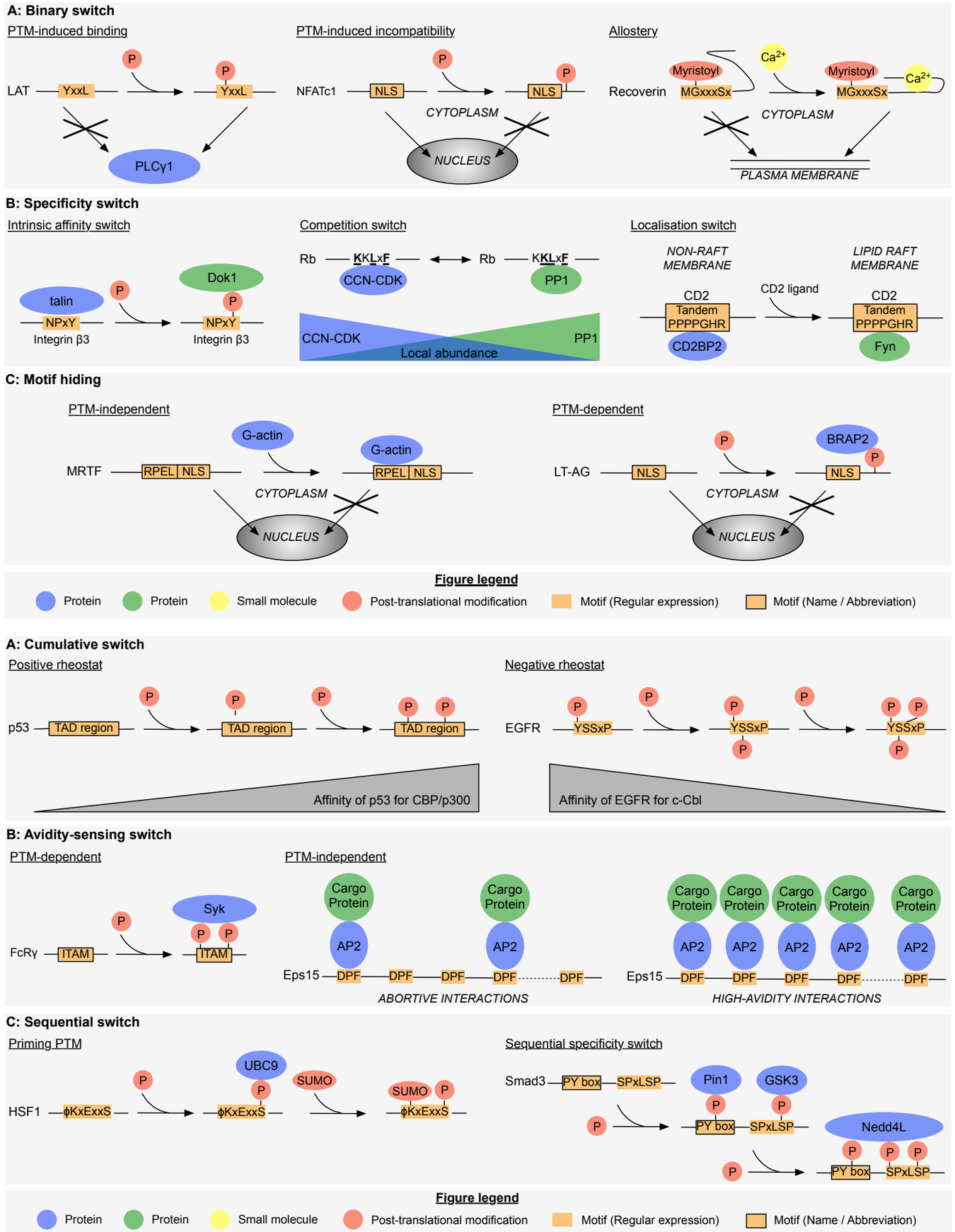
Available online at www.sciencedirect.com
SciVerse ScienceDirect
Current Opinion in Structural Biology

Switches.elm.eu.org

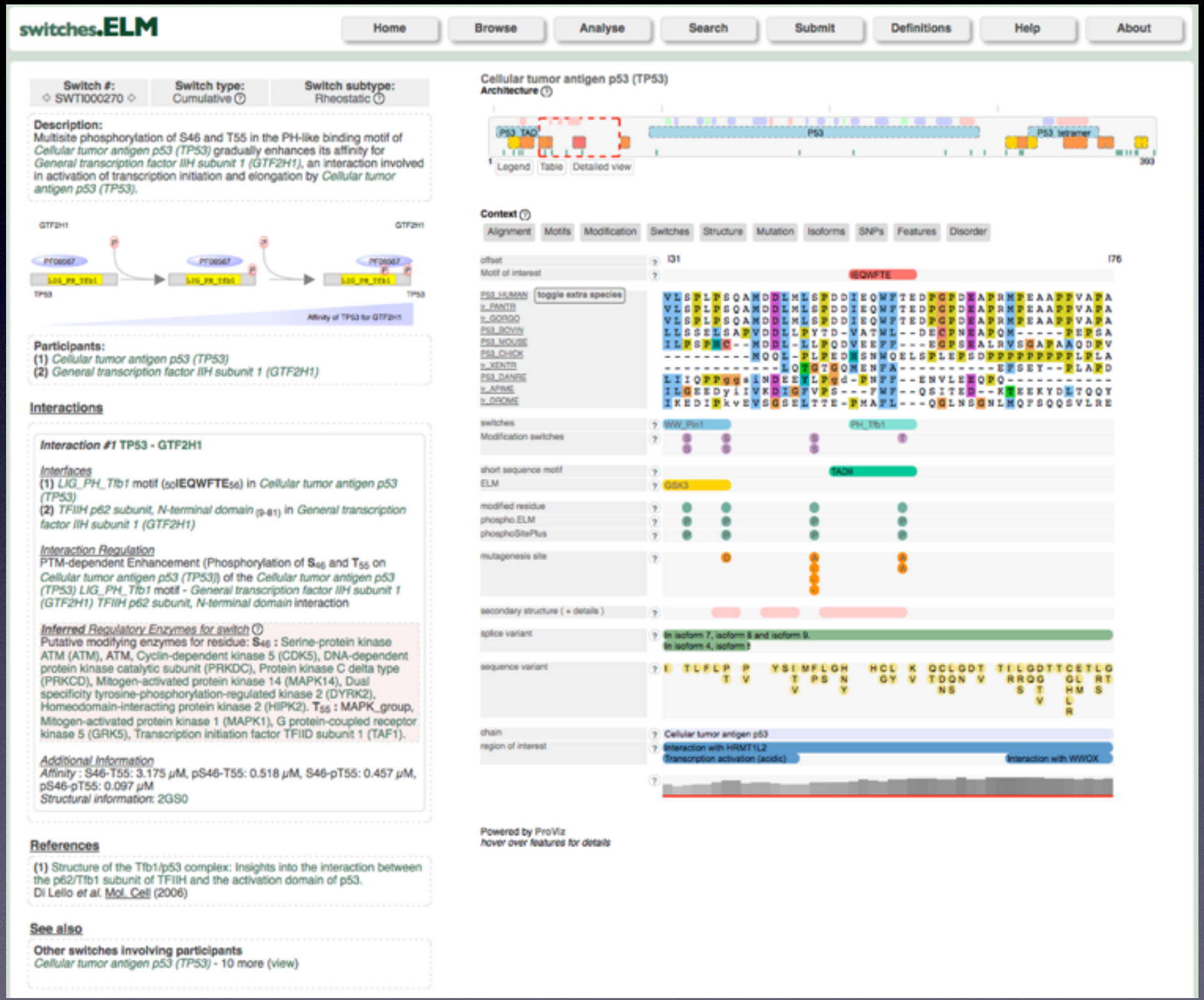
Motif switches: decision-making in cell regulation
Kim Van Roey[1], Toby J Gibson[1] and Norman E Davey[1,2]

Six classes of molecular switch involving IDP

- Binary Switch
  - Simple On-Off
- Specificity Switch
  - Multiple On states
- Motif-Hiding Switch
  - Conditional motif accessibility
- Cumulative Switch
  - Graduated rheostat-like behaviour
- Avidity sensing
  - Sharp, cooperative affinity shift
- Sequential Switch
  - Strict logical dependence of execution



switches.ELM p53 rheostatic switch example



Cell Regulatory Decisions

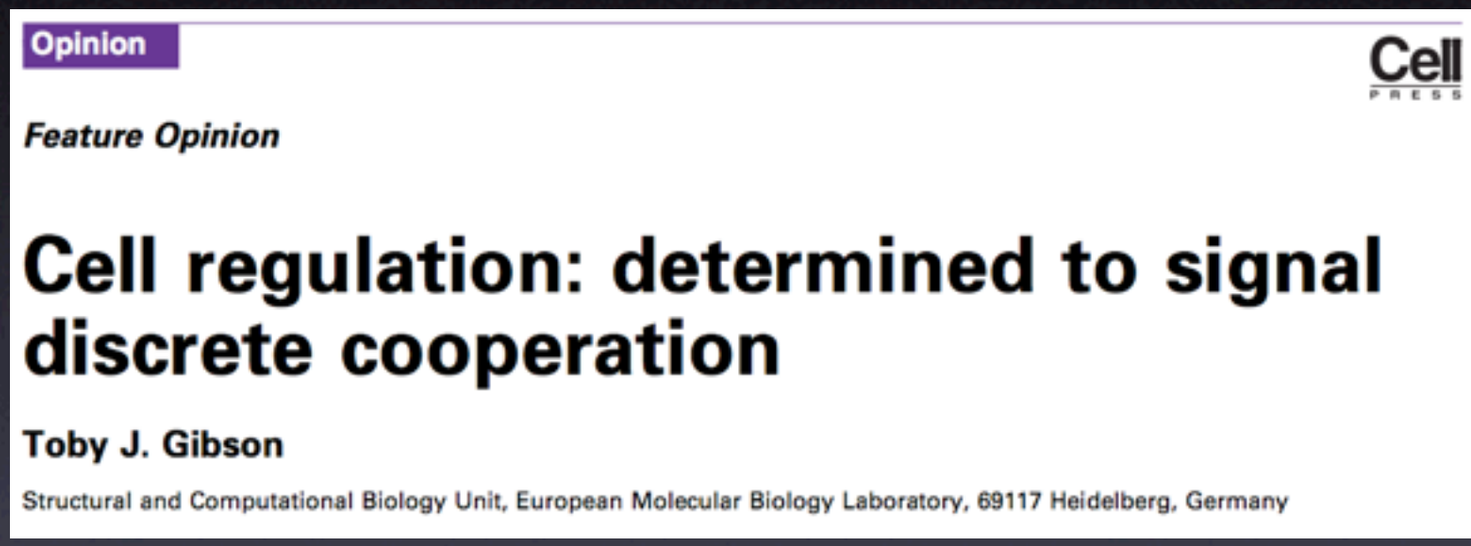
- Are made in large complexes
  - by in-complex molecular switching
    - including addition and subtraction of proteins to complexes
  - using switches assembled from low affinity interacting components
- Allostery is a major switching mechanism
- Pre-assembly is a major switching mechanism
  - and variations on pre-assembly switches include rheostats, avidity sensors, motif-hiding switches, sequential switches....

Everything should be made as simple as possible, but not simpler

Albert Einstein

Cell regulation is networked and redundant being effected by discrete, precise and cooperative molecular switches in large regulatory protein complexes

- No cellular dictator
- No master regulator
- No first among equals
- No top-down system of governance



Opinion

Cell PRESS

Feature Opinion

Cell regulation: determined to signal discrete cooperation

Toby J. Gibson

Structural and Computational Biology Unit, European Molecular Biology Laboratory, 69117 Heidelberg, Germany

TIBS 10/09

The "politics" of the Cell is Anarcho-Syndicalist

Homage to Catalonia

Some Cooperative Interactors from the past and present

Current Group Members

Brigitte Altenberg (V)
Aidan Budd
Francesca Diella (V)
Holger Dinkel
Manjeet Kumar
Ben Lang
Vlada Milchevskaya
Hugo Samanas
Grischa Tödt

EU Grant Coordinators

SysCilia FP7 6/2010 - 15
Ronald Roepman (Nijmegen)

SyBoSS FP7 6/2010 - 15
Francis Stewart (Dresden)

ELM Resource Collaborators

ELM Founder Groups
Toby Gibson (EMBL)
Rein Aasland (Bergen)
Bill Hunter (Dundee)
Manuela Helmer-Citterich (Rome)
Leszek Rychlewski (Poznan)
Bernhard Kuster (Cellzome)

Phospho.ELM
Nikolaj Blom (Lyngby)
Martin Miller (Lyngby)
Thomas Sicheritz-Pontén (Lyngby)
Scott Cameron (Dundee)
Bernhard Kuster (Cellzome)
Lars Jensen (Bork)
Allegra Via (Rome)

Cooperative Standards
Henning Hermjakob (EBI)
Sandra Orchard (EBI)
Chris Workman (Copenhagen)
Olga Rigina (Copenhagen)
Fred de Masi (Copenhagen)

Clustal W/X 2.0
Mark Larkin (Dublin)
Des Higgins (Dublin)
Chenna Ramu (Berlin)
Nigel Brown (Heidelberg)
Rodrigo Lopez (EBI)
Julie Thompson (Strasbourg)

Ataxin-1 Molecular Switch
Annalisa Pastore (Mill Hill)
Cesira de Chiara (Mill Hill)

Transient overexpression
Reiner Veitia (Paris)

Million Motifs
Madan Babu (Cambridge)
Peter Tompa (Brussels)

DiGtoP BMBF/2008-2013
Wolfgang Wurst (München)
Francis Stewart (Dresden)
Matthias Mann (München)
Tony Hyman (Dresden)
Oliver Brüstle (Bonn)
.....