



EMBO

EMBO PPI Bioinformatics Course

Rome, 5-11-18

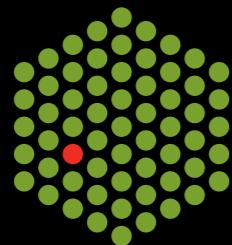


*Sequence/Structure/Function/Networks
All Contribute To The
Understanding of Cell Regulation
And How It Goes wrong in Disease*

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



EMBL



Bioinformatics is now often a part of interdisciplinary biomedical studies

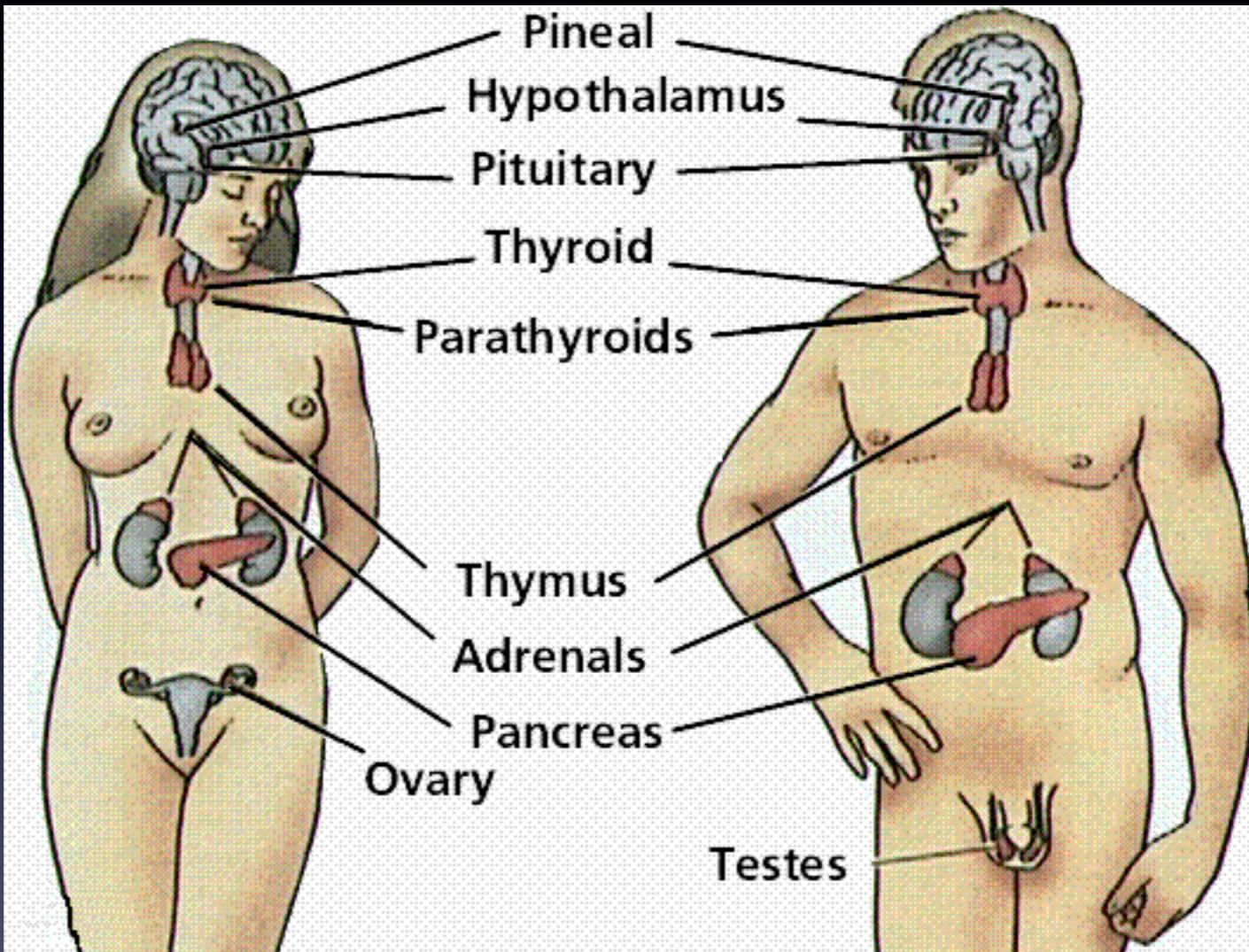
This talk visits some relevant topics to try and set background context about why PPI bioinformatics is useful.

In particular, aspects of cancer, viral and bacterial diseases

Interdisciplinary Signalling Research
requires addressing communication
at multiple levels
between cells and within cells

Let's drill down from the top level

Organismal Level Signalling



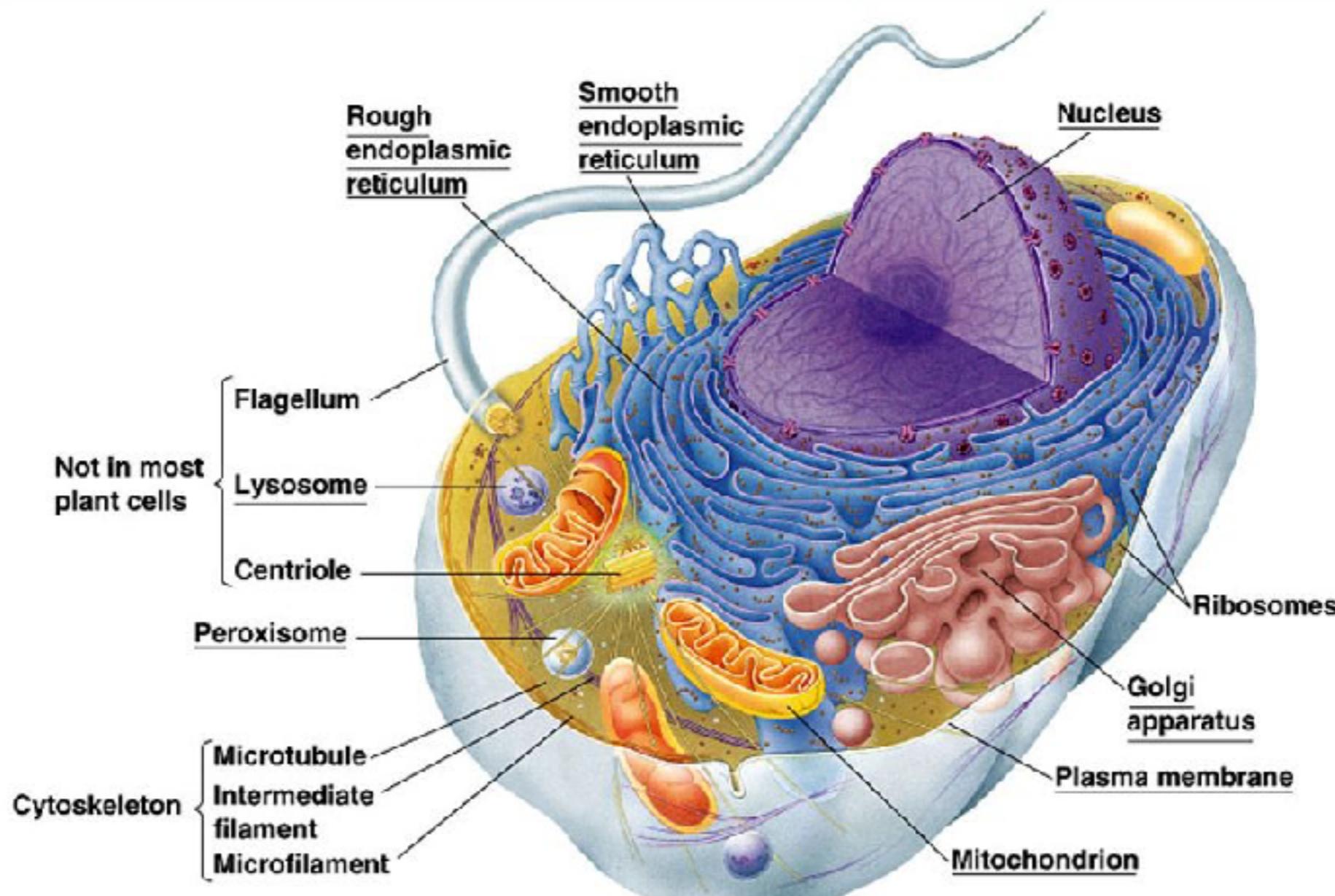
Insulin, endocrine system etc. could be researched long before in-cell signalling

Circadium Rhythm has organismal and cellular level components. So why would anyone study circadium rhythm in Hela cells?

The screenshot shows a search results page from PubMed. At the top, there is a search bar with the query "hela circadium rhythm". Below the search bar are three buttons: "Create RSS", "Create alert", and "Advanced". Underneath the search bar, there are dropdown menus for "Format: Summary" and "Sort by: Most Recent". The "Per page:" dropdown is set to 20. To the right of these dropdowns is a "Send to" button. Below the search controls, the text "Search results" is displayed in bold. Underneath "Search results", it says "Items: 1 to 20 of 40". At the bottom of the page, there are navigation links: "<< First", "< Prev", "Page 1 of 2", "Next >", and "Last >".

In-cell signalling is primarily studied in cancer cells adapted to tissue culture: (1) Biased (2) Broken

A Generic Animal Cell



Copyright © 2003 Pearson Education, Inc., publishing as Benjamin Cummings.

Copyright © 2004 Pearson Education, Inc. publishing as Benjamin Cummings

Molecules have to get into cell compartments but also be able to get around them

Cell Regulation: Cooperative and Spatially arranged

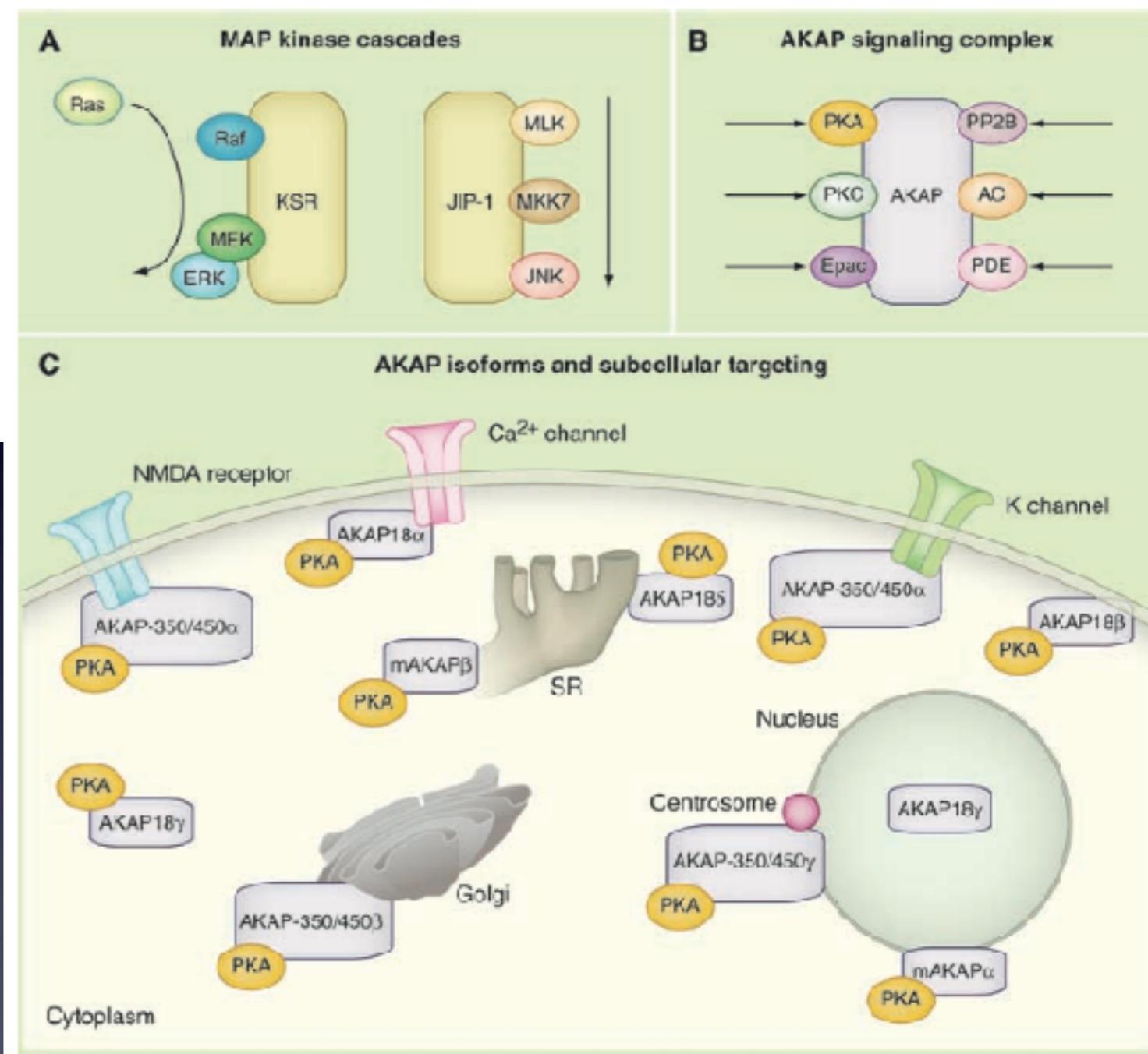
Spatial Cell Biology

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



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.

Spatial Exclusivity of Mitotic Kinases

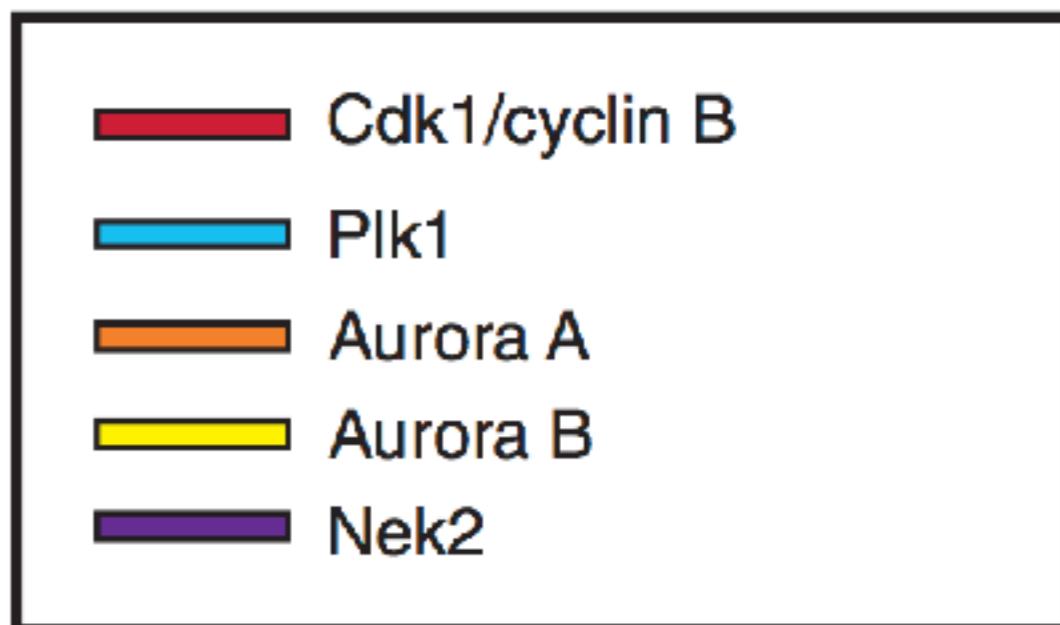
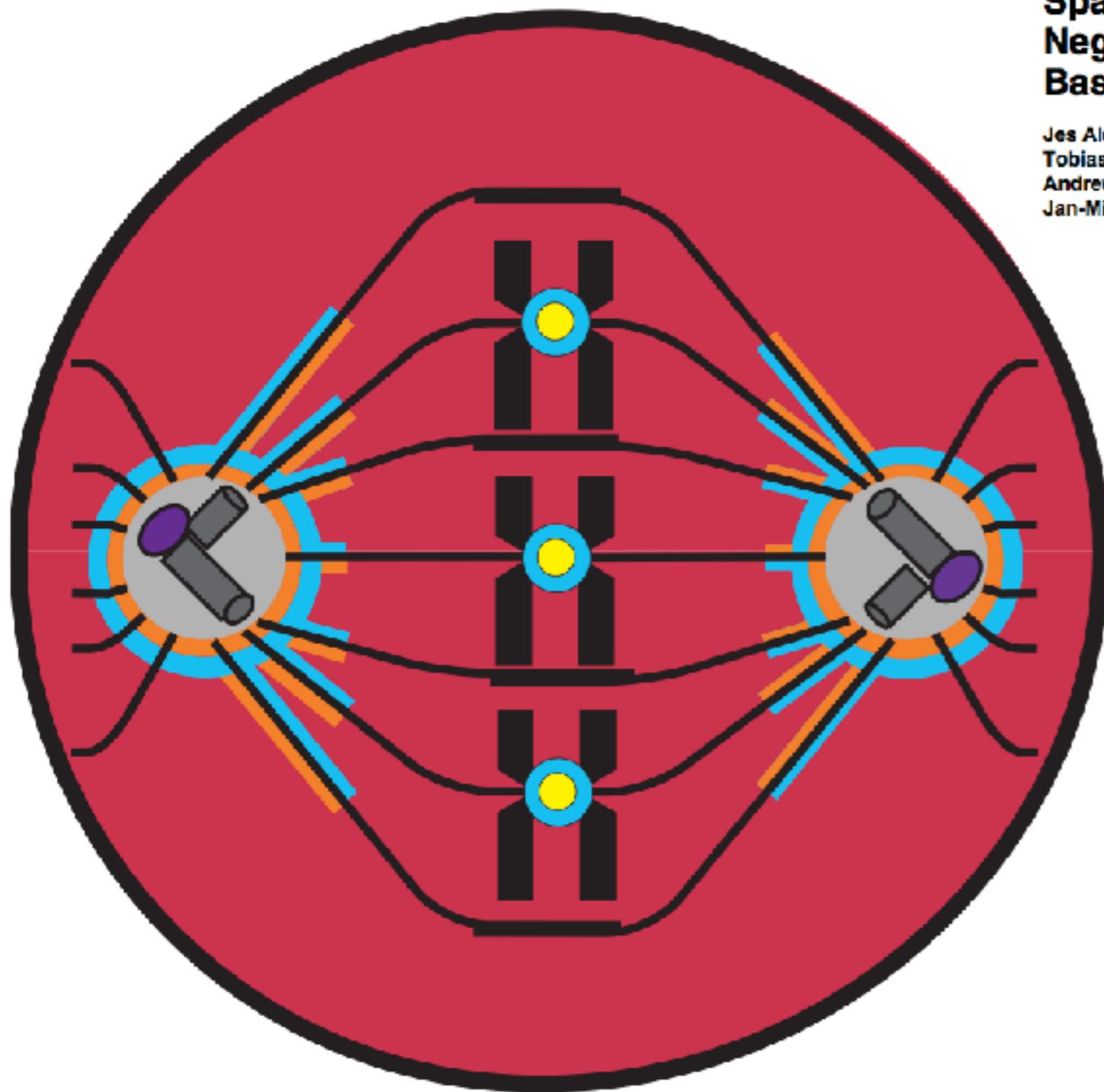
RESEARCH ARTICLE

MITOTIC KINASES

Sci. Sig., 6-2011

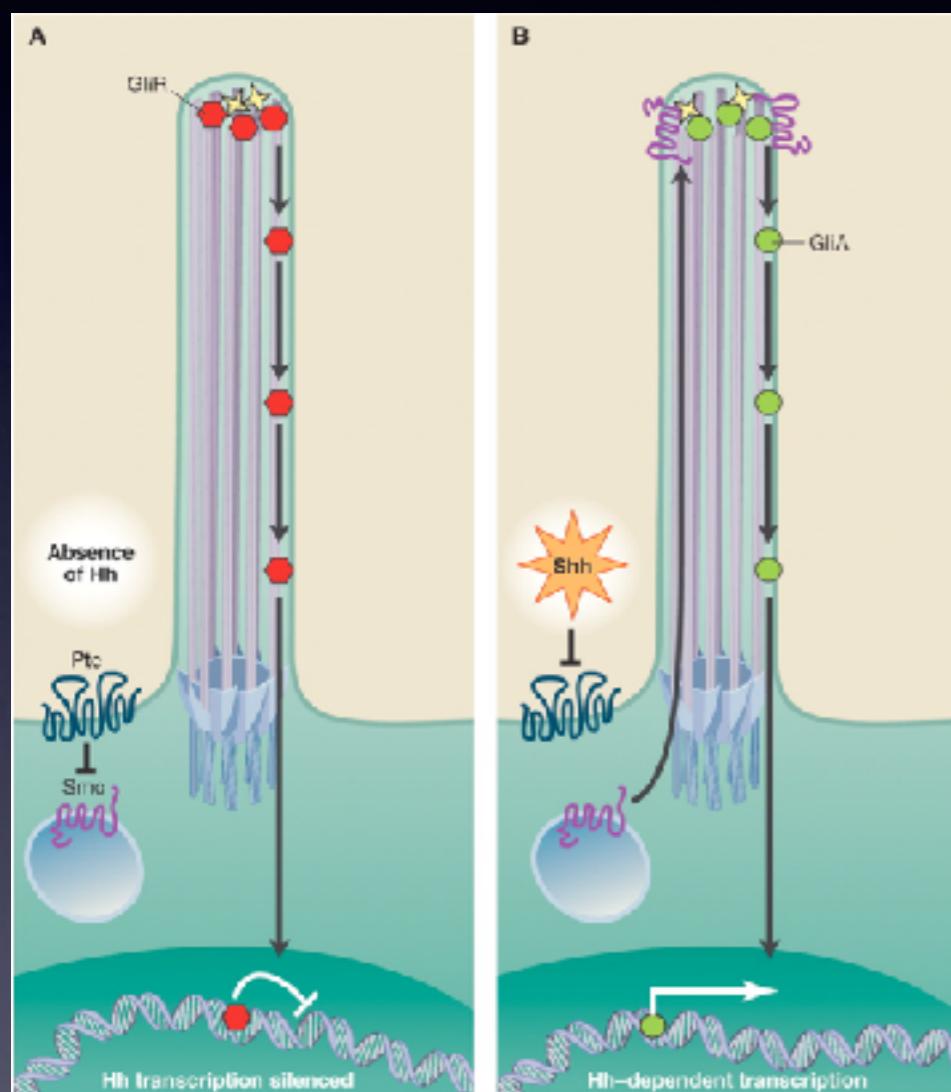
Spatial Exclusivity Combined with Positive and Negative Selection of Phosphorylation Motifs Is the Basis for Context-Dependent Mitotic Signaling

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‡}



Primary cilia represent the extreme of space-time regulation in signaling

Hedgehog signalling in Cilia utilise Gli/Glis transcription factors trafficked to nucleus.



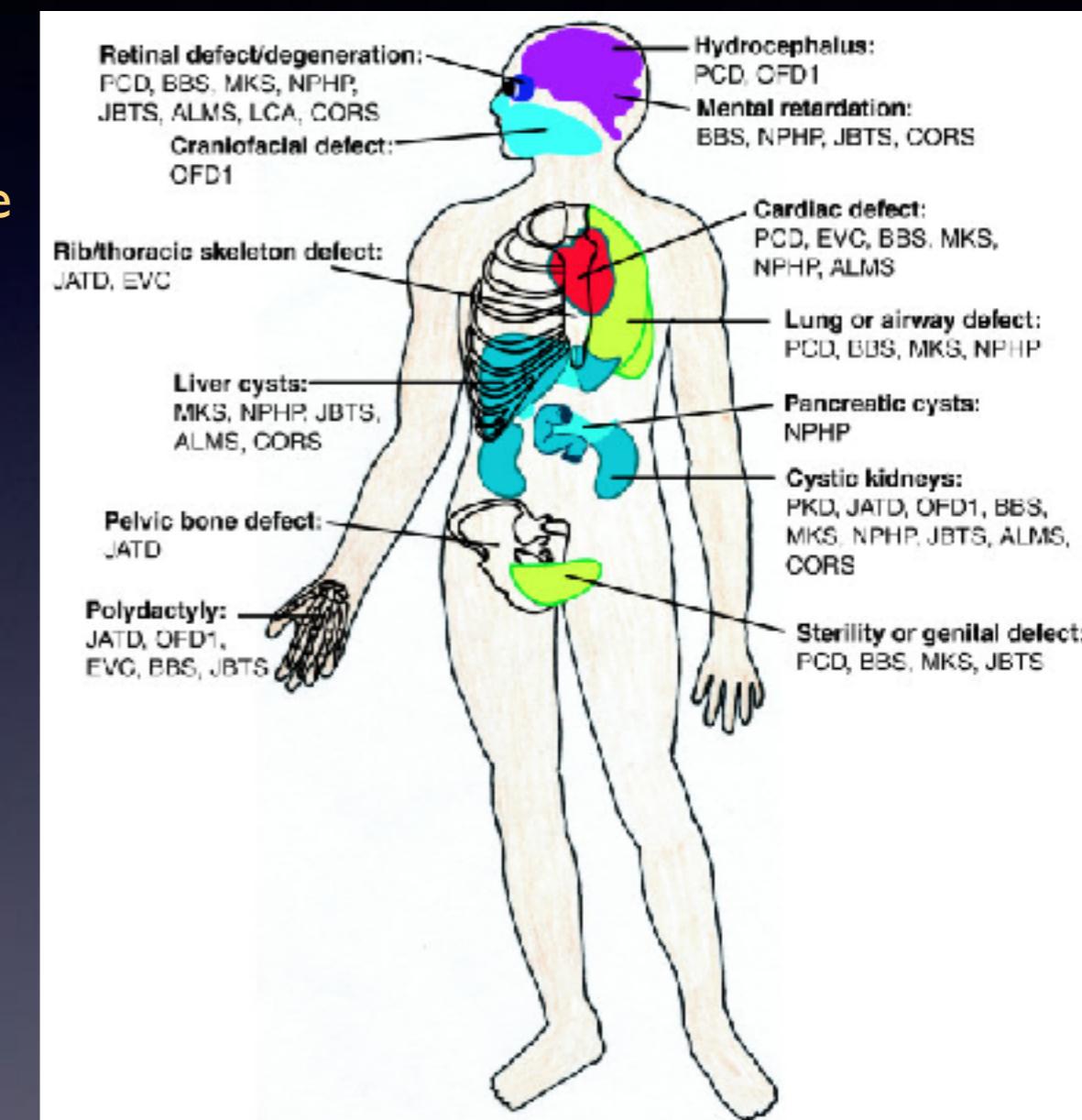
Singla and Reiter (2006) Science, 313, 629

How many
Angels can dance
upon a pinhead?



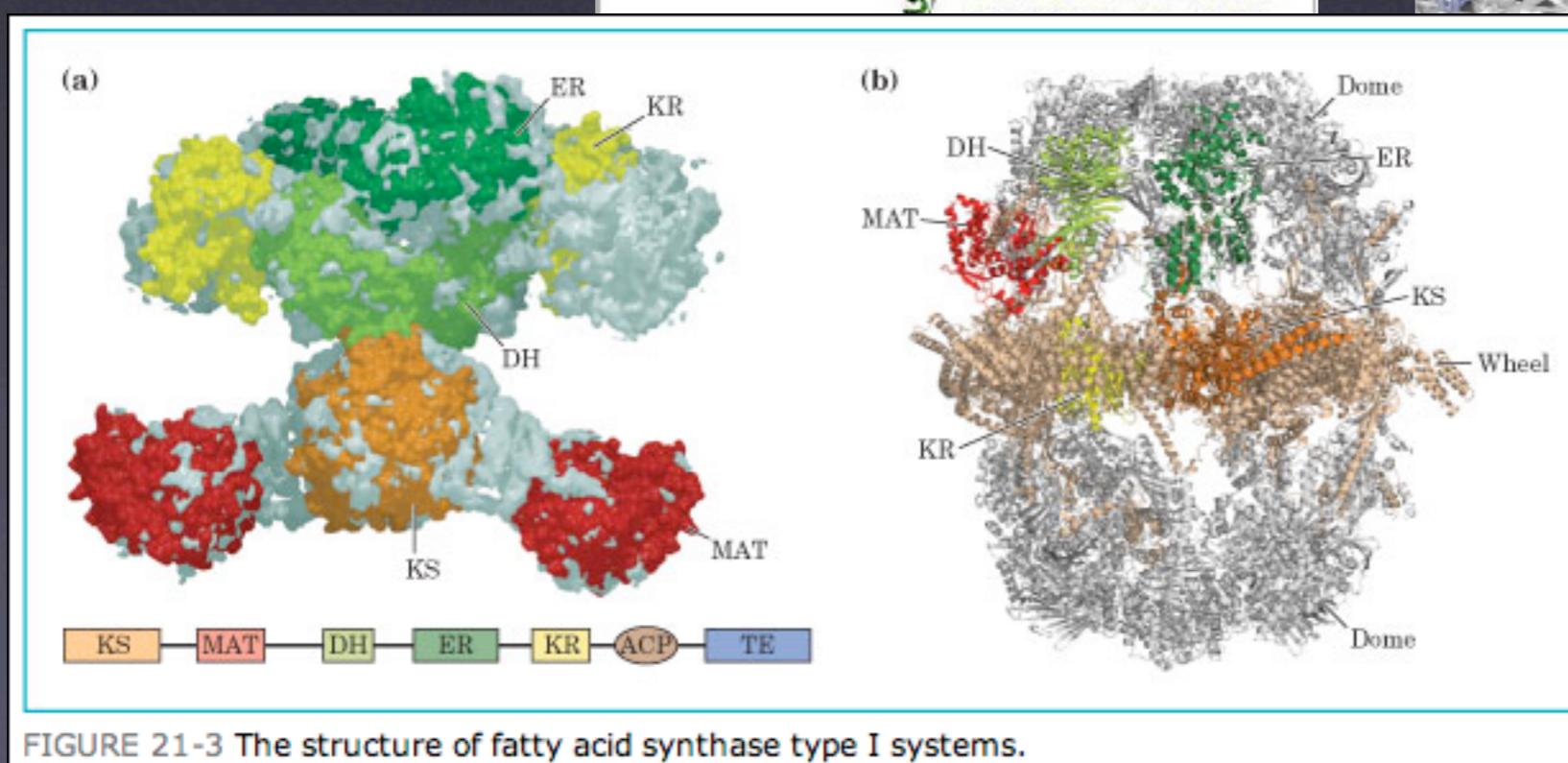
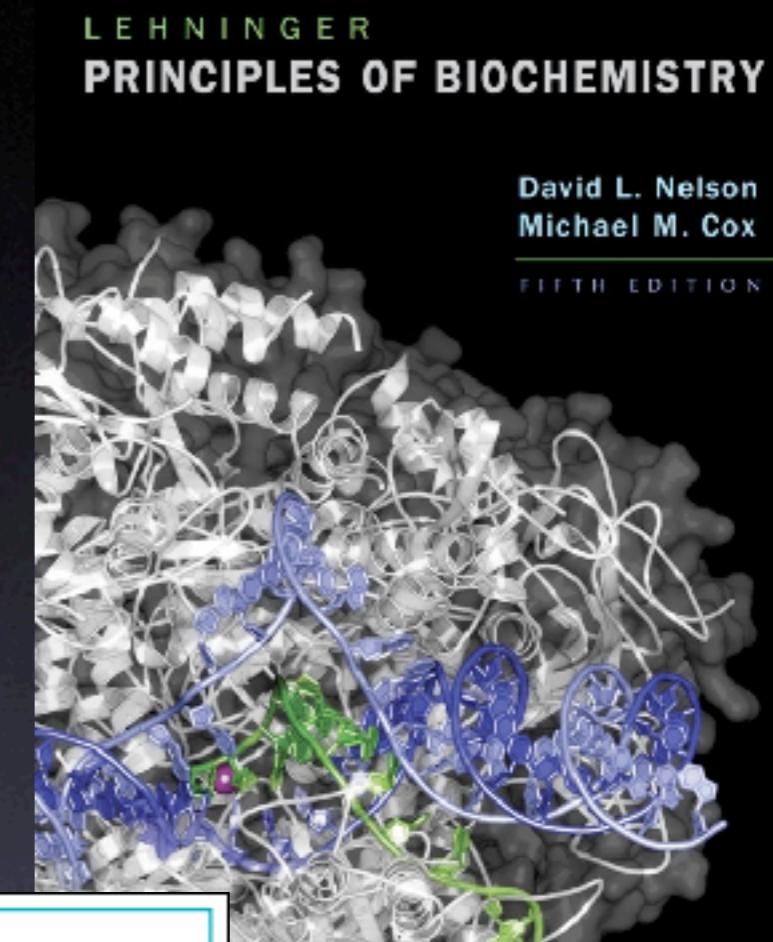
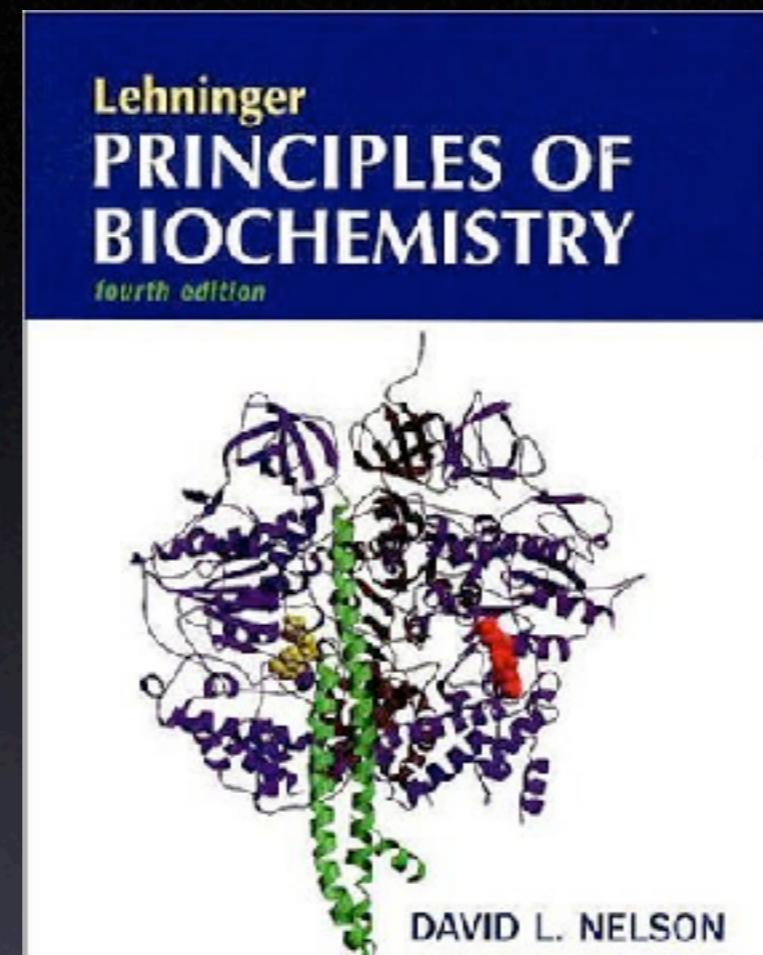
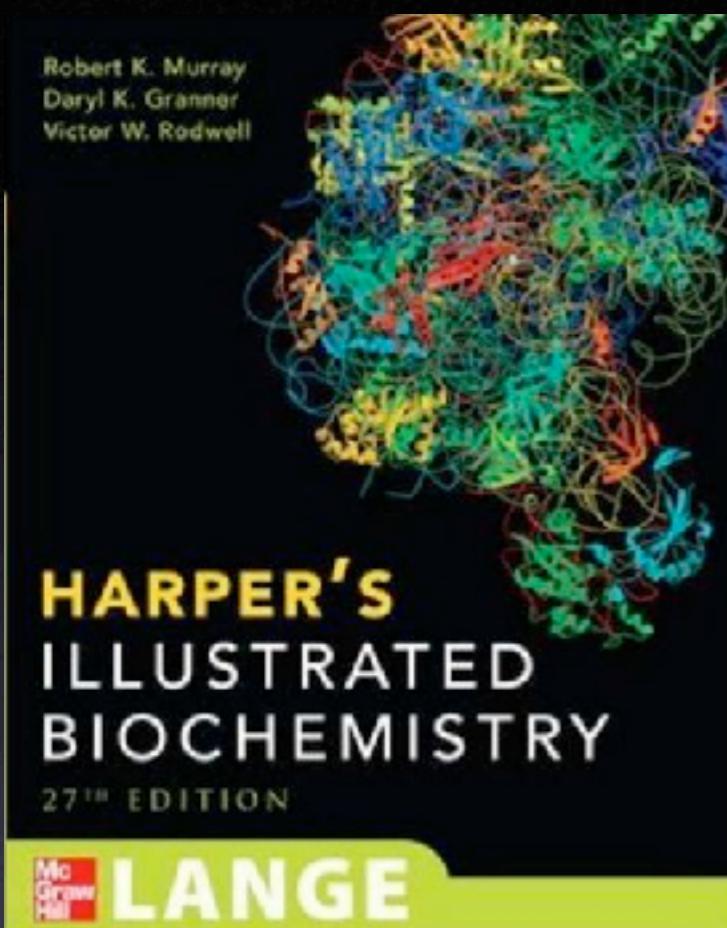
How many TFs
can fit on the tip
of a cilium?

Ciliopathies affect nearly every tissue in the body



Lee and Gleeson Genome Medicine 2011 3:59

Biochemistry Text Book Figures: Engines of the Cell



Complexes
Complexes
Complexes

U4/U6.U5 tri-snRNP by cryo-EM

A Cryo-EM complex that is:

Large

Flexible

Dynamic

Has big holes

Very hard to crystallise

There are plenty of
complexes that are:

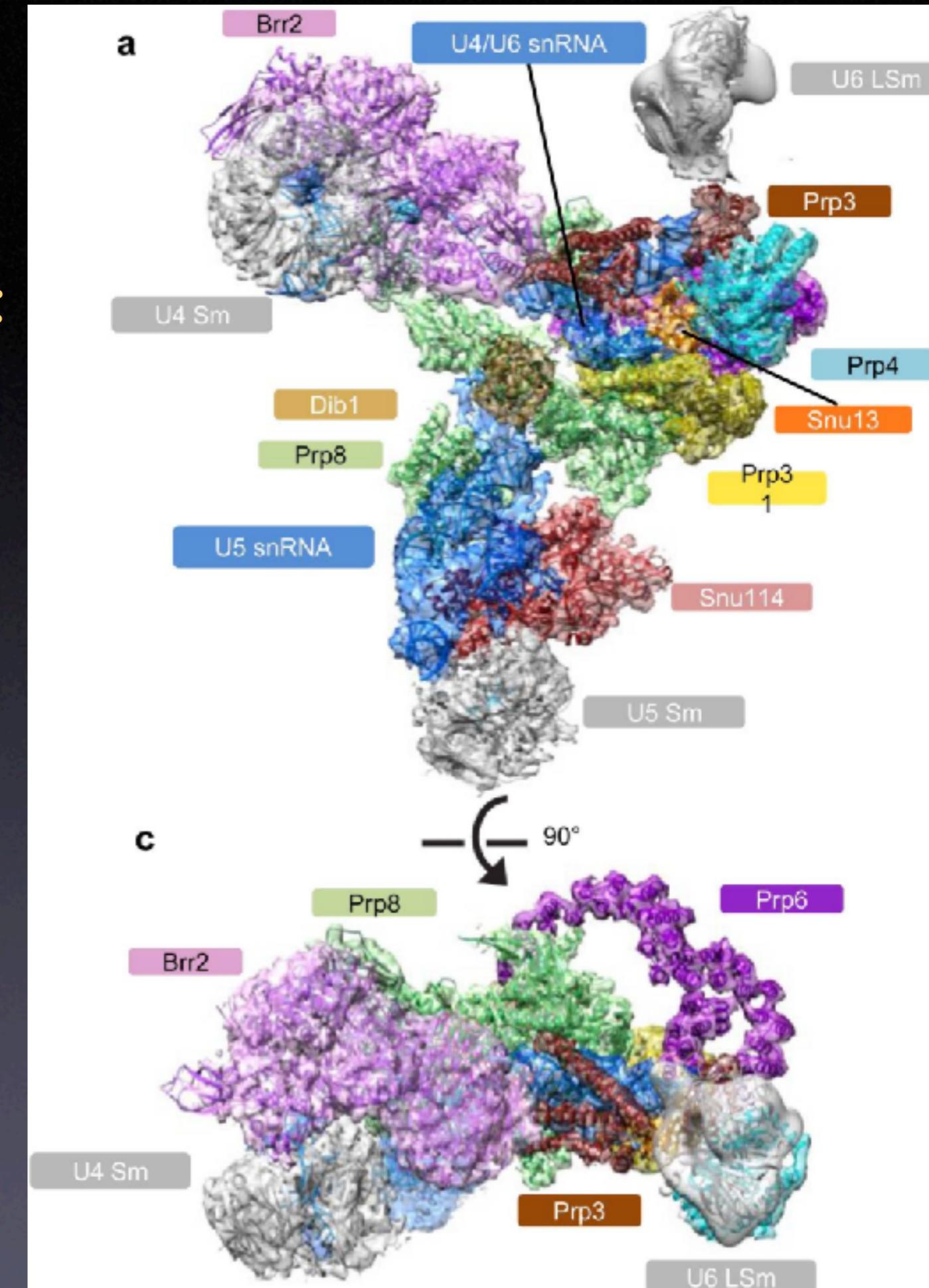
Much larger

More flexible

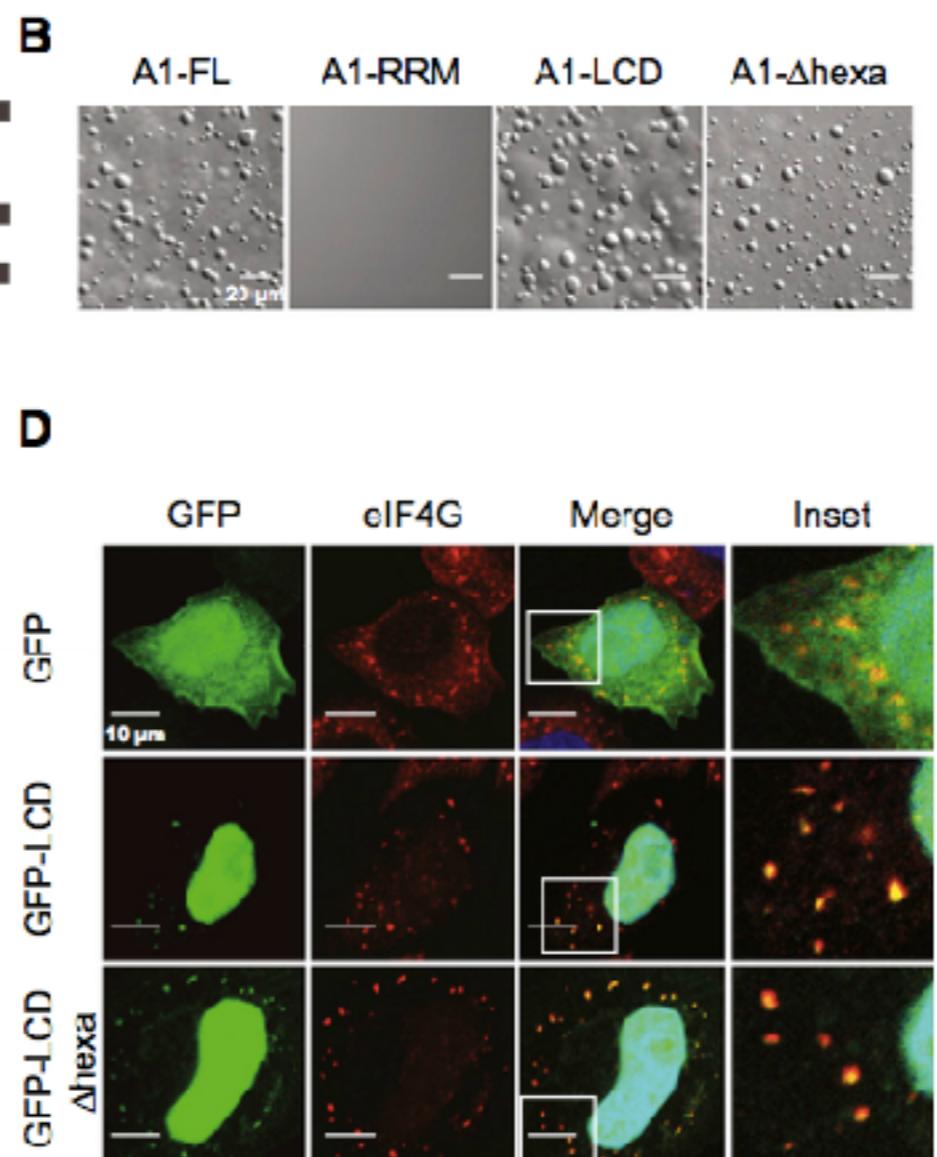
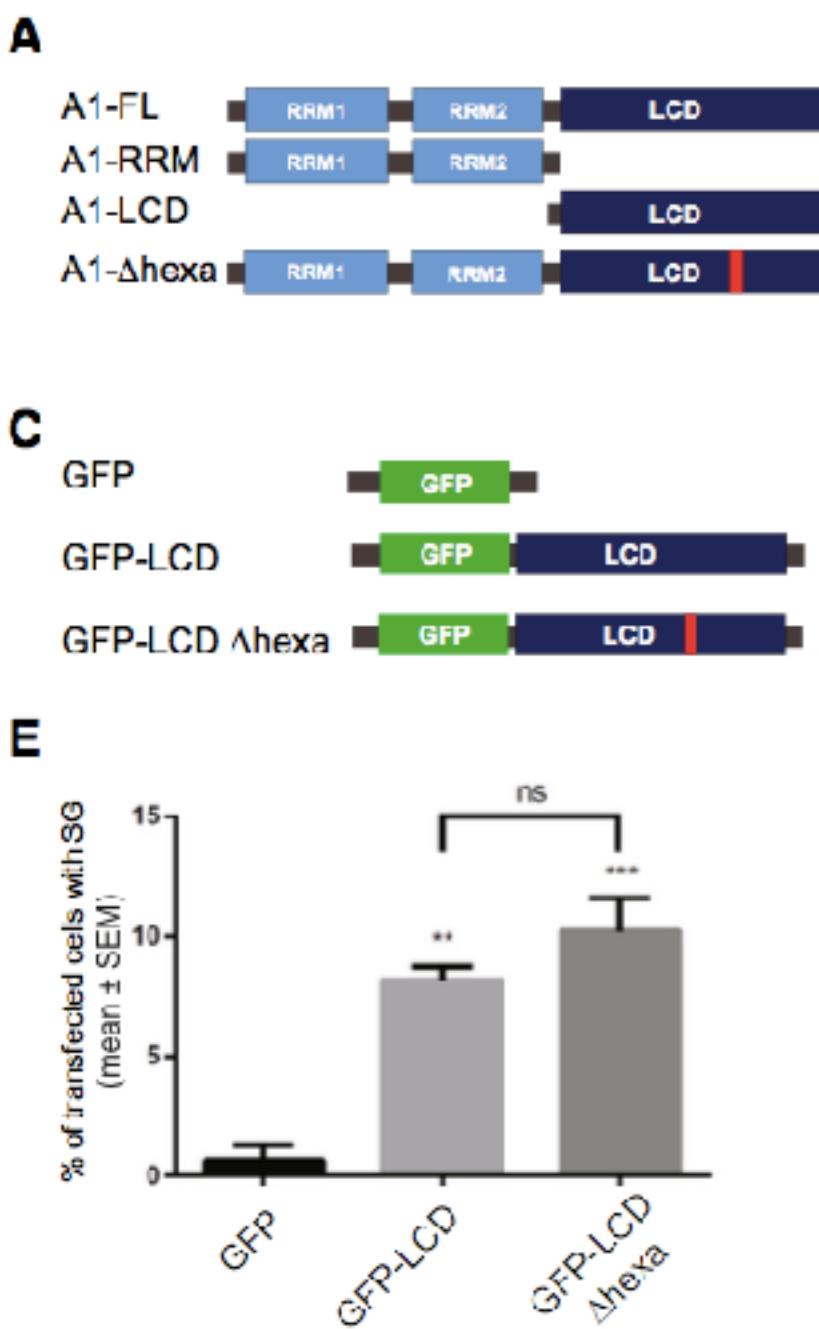
More dynamic

Have even bigger holes

Will never be crystallised



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- Δ hexa).

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

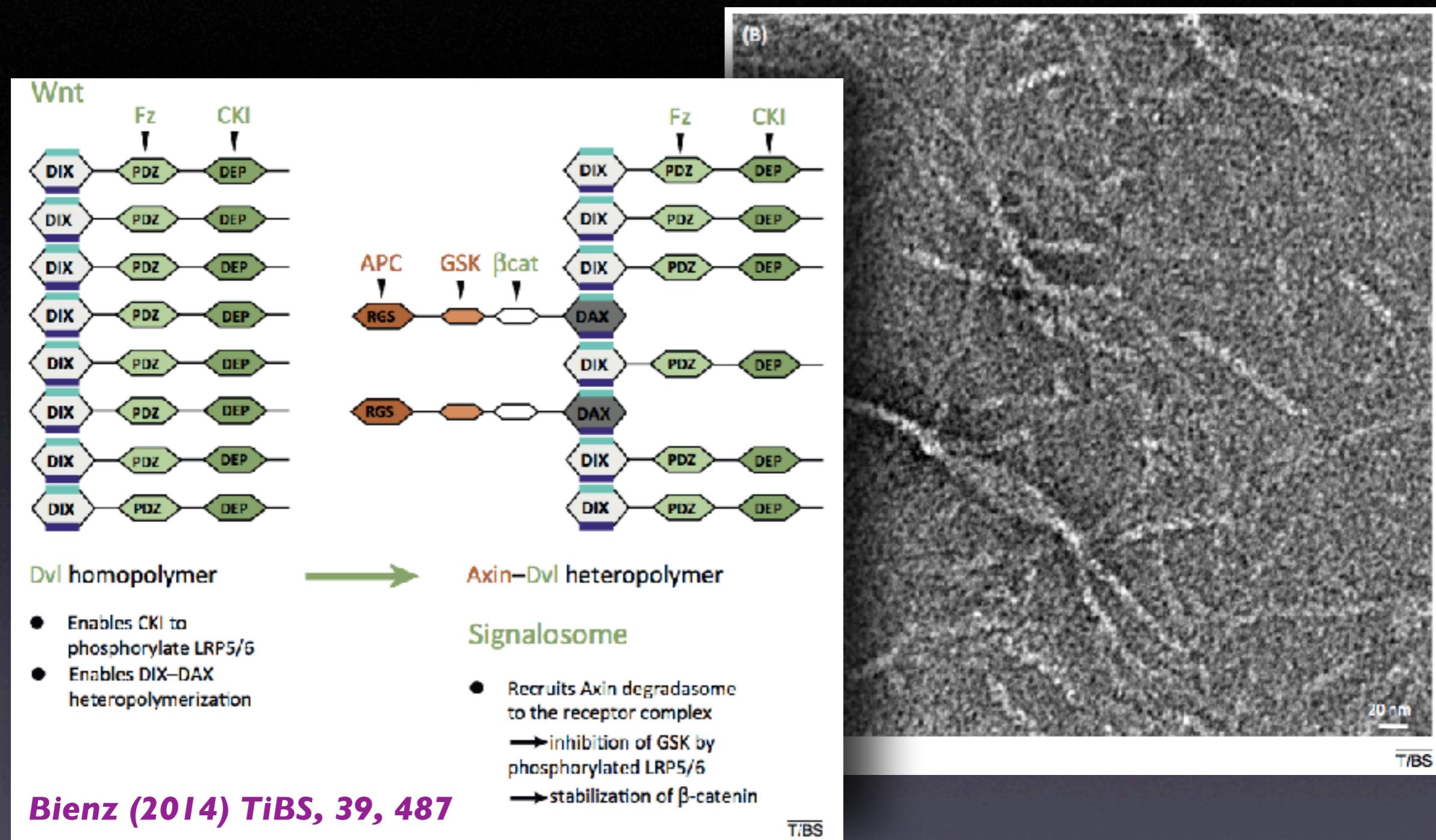
(C) Schematic of the constructs transiently expressed in HeLa cells.

(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] \times 100) was plotted as mean \pm 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)

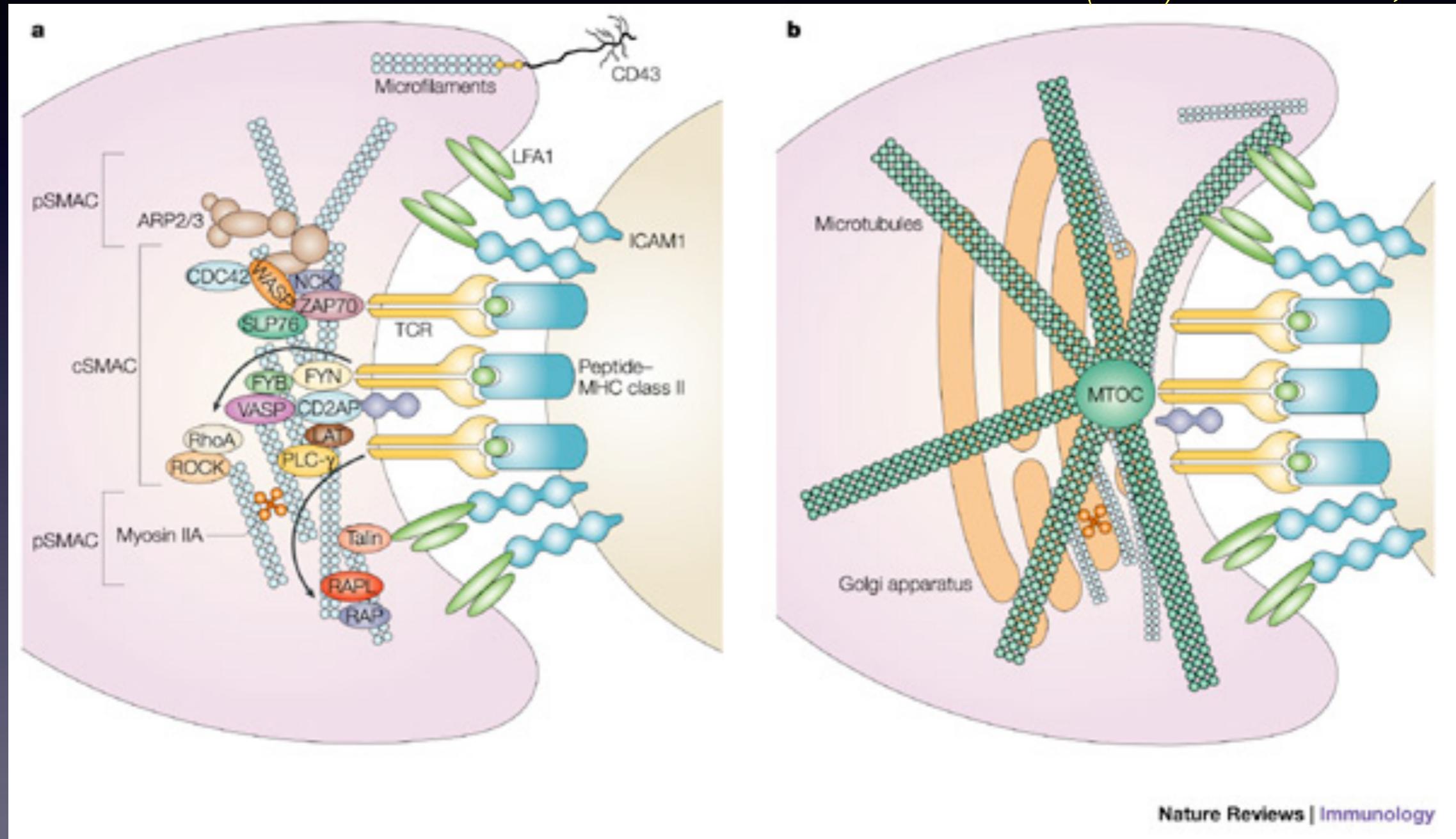
Some types of complex tolerate stoichiometry violations



Wnt-signalling Polymeric helical signalosomes have different proportions of component proteins: They also act as scaffolds to assemble other regulatory proteins/complexes

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

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

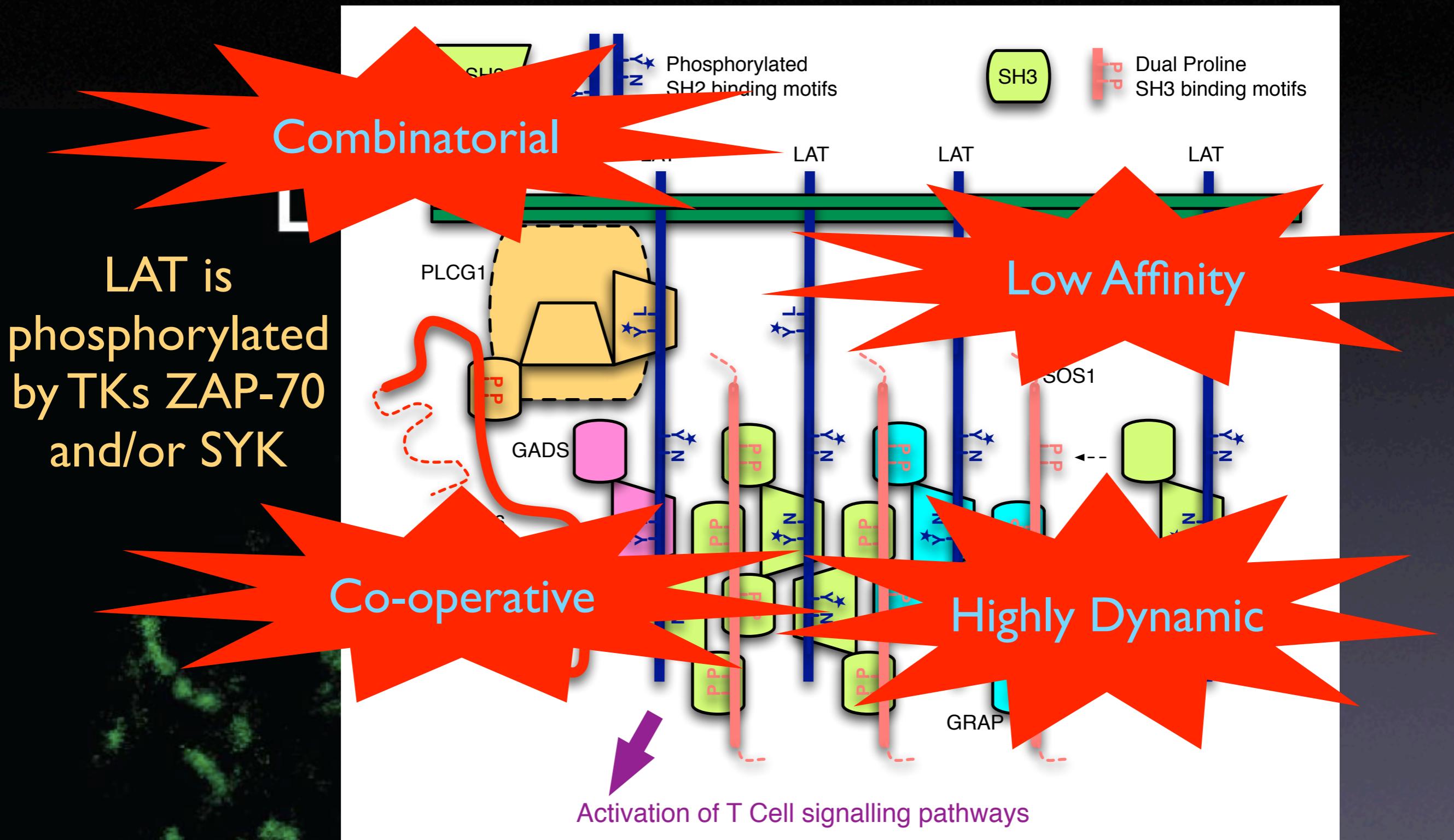


Nature Reviews | Immunology

Effectively the whole cell becomes a giant macromolecular complex

Initiation and propagation of T Cell signalling

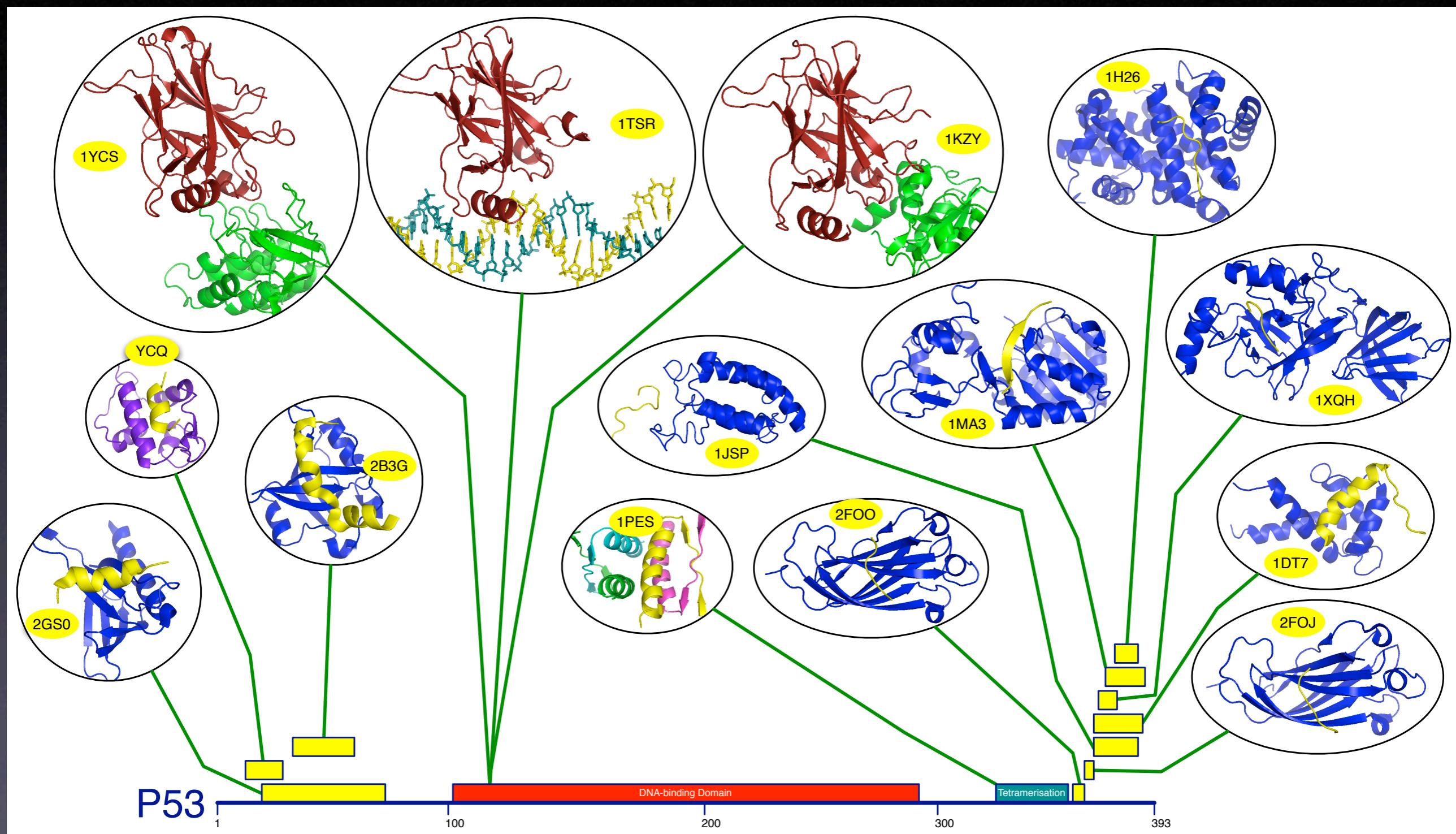
Multivalent assembly of the LAT signalling complex by short linear motifs



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

Solved structures of p53 fragments

p53 exhibits switch-like properties



p53 is a famous example of a protein with a modular structural architecture

Thousands of cell regulatory proteins are composed of three main protein module classes:

- Natively folded domains
- Intrinsically disordered domains (IDDs)
- Short Linear Motifs (SLiMs)

You will learn more about these modules during the course

Degrons are a type of SLiM

SCIENCE SIGNALING | REVIEW

CANCER

Degrons in cancer

Bálint Mészáros,^{1*} Manjeet Kumar,^{2*} Toby J. Gibson,^{2†} Bo

Degrons are the elements that are used by E3 ubiquitin ligases to target short linear motifs embedded within the sequences of modular proteins; they are important for many different cellular processes, such as programmed cell death and cellular hypoxia. Degrons enable the elimination of proteins that are no longer required, preventing their possible dysfunction. Although the human genome encodes ~600 E3 ubiquitin ligases, only a fraction of these enzymes have well-defined target degrons. Thus, for most cellular proteins, the destruction mechanisms are poorly understood. This is important for many diseases, especially for cancer, a disease that involves the enhanced expression of oncogenes and the persistence of encoded oncoproteins coupled with reduced abundance of tumor suppressors. Loss-of-function mutations occur in the degrons of several oncoproteins, such as the transcription factors MYC and NRF2, and in various mitogenic receptors, such as NOTCH1 and several receptor tyrosine kinases. Mutations eliminating the function of the β -catenin degron are found in many cancers and are considered one of the most abundant mutations driving carcinogenesis. In this Review, we describe the current knowledge of degrons in cancer and suggest that increased research on the “dark degrome” (unknown degron-E3 relationships) would enhance progress in cancer research.



Amy Wheeler

@DJAhmei

I misread this paper's title as "Dragons in cancer"

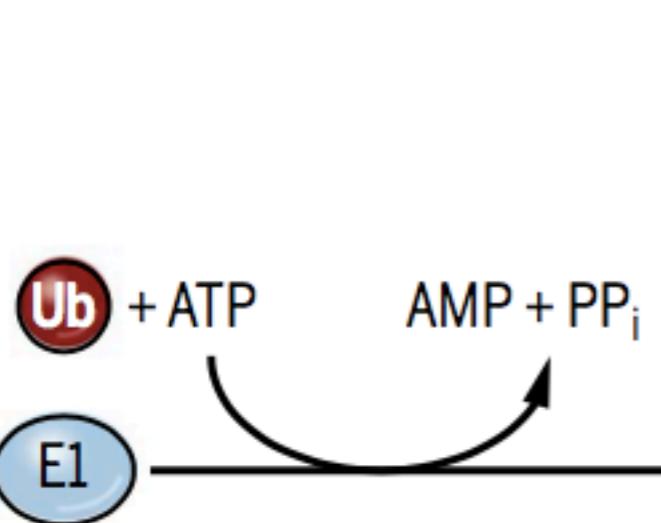
<https://t.co/cmdQIFEUhH>

14 Mar 2017

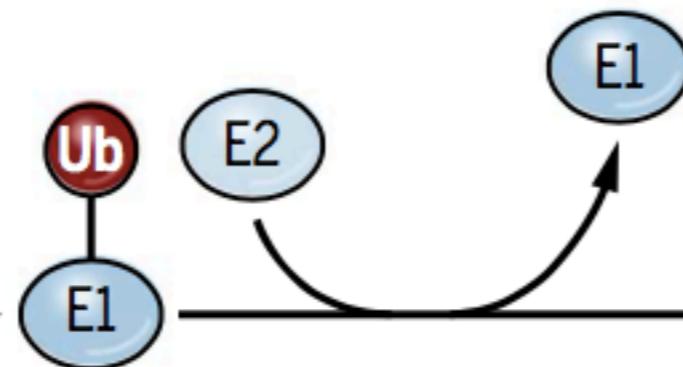


Proteins are recycled by the ubiquitin-proteasome system (UPS)

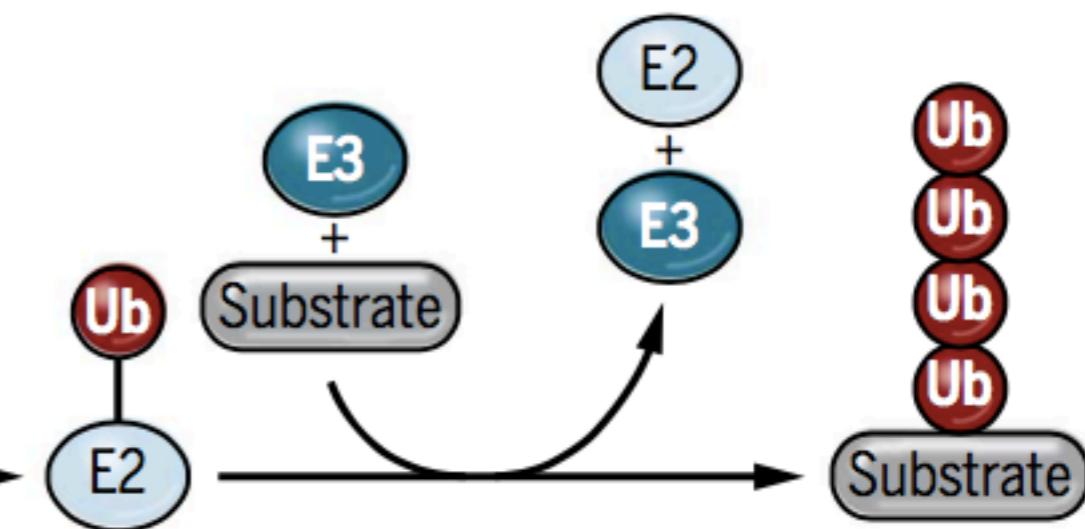
Ubiquitin activation



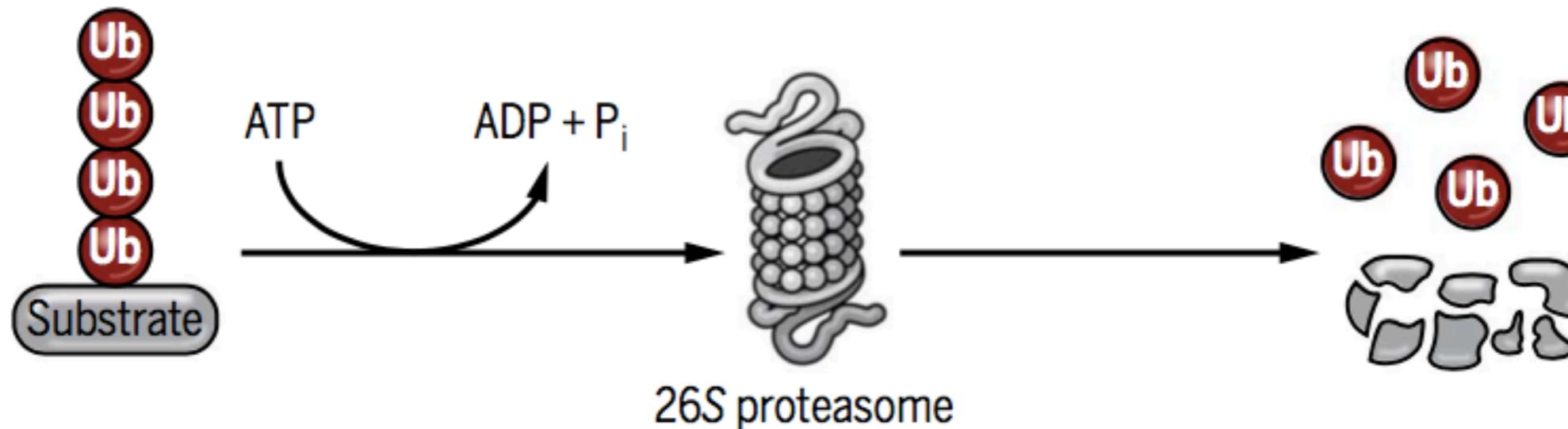
Ubiquitin transfer to E2



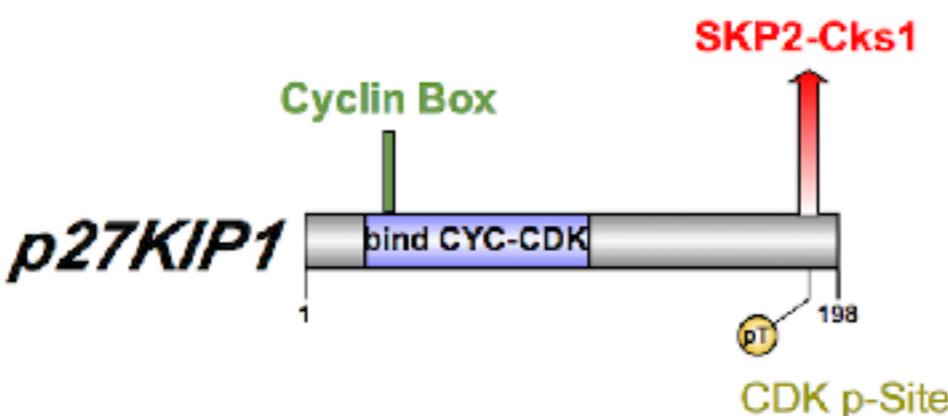
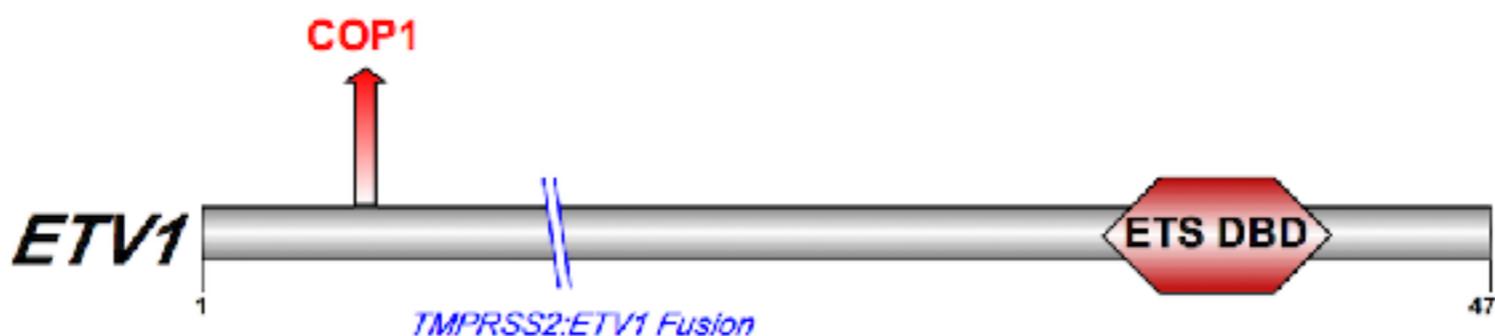
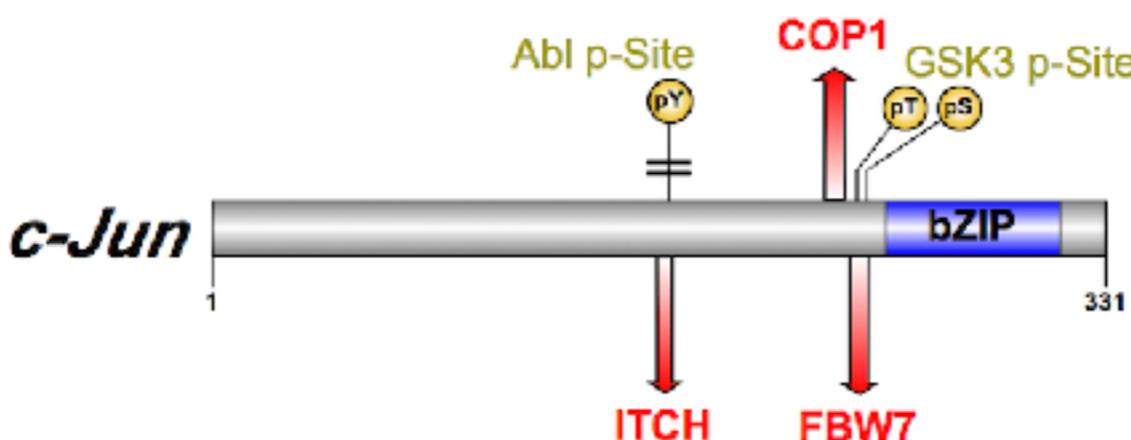
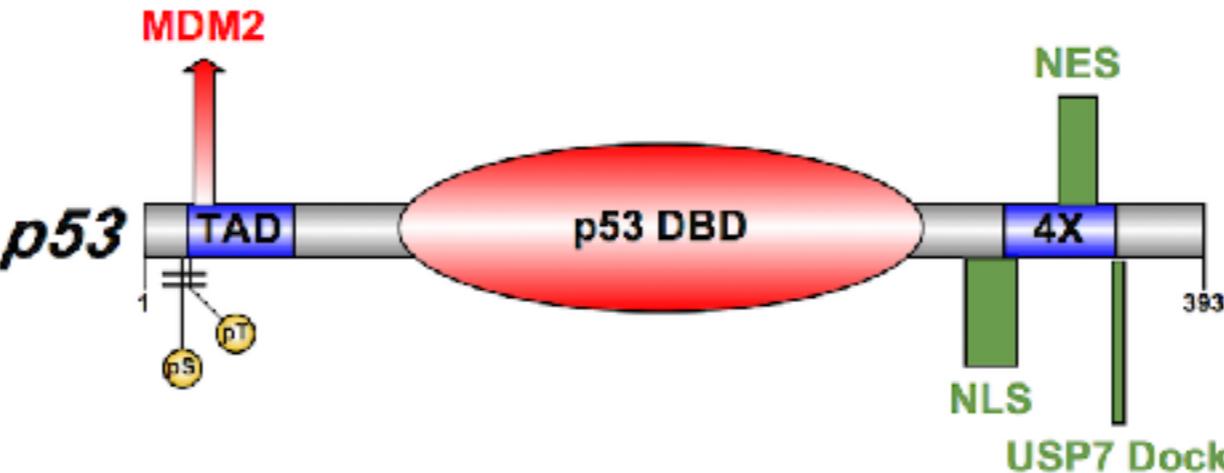
Ubiquitin attachment to substrate



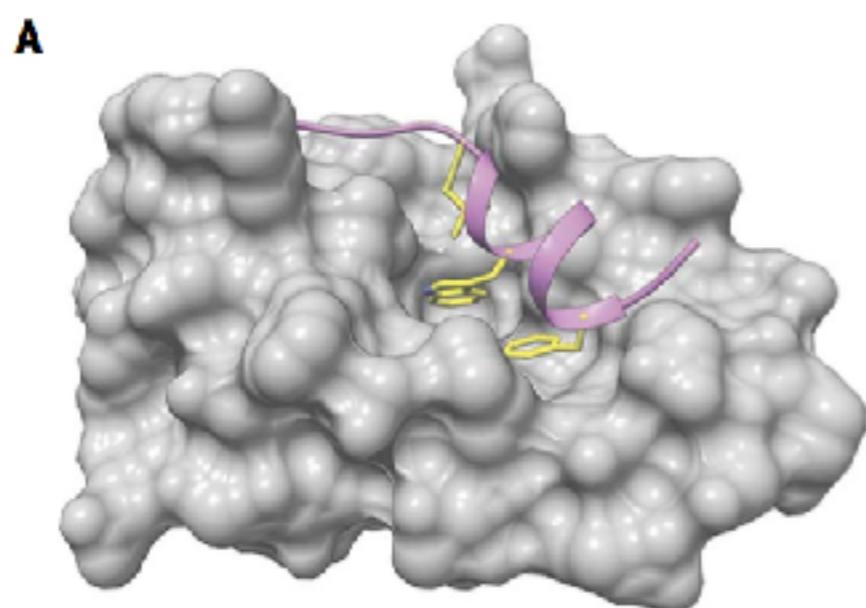
Ub-tagged substrate degradation



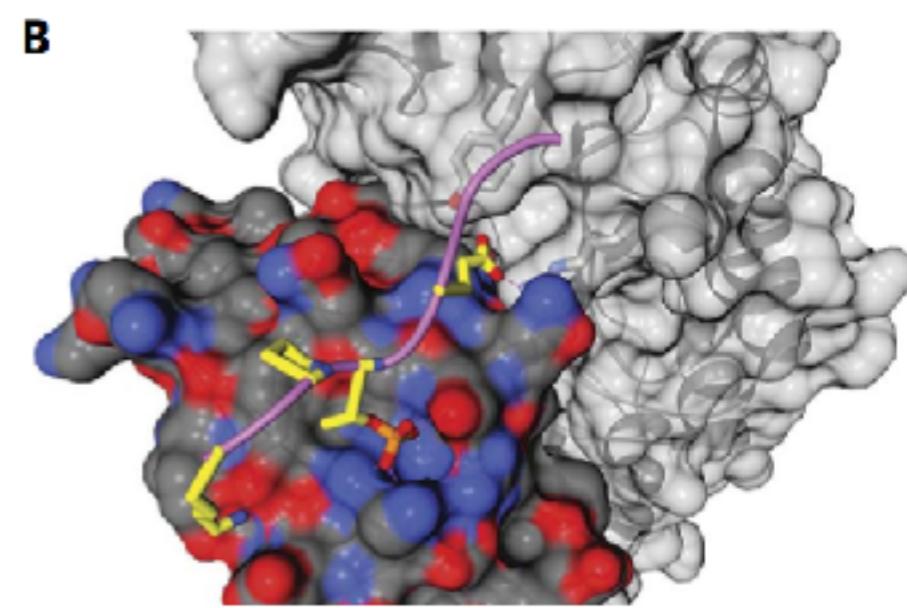
E3 ligases recognise “degrons” on the target proteins



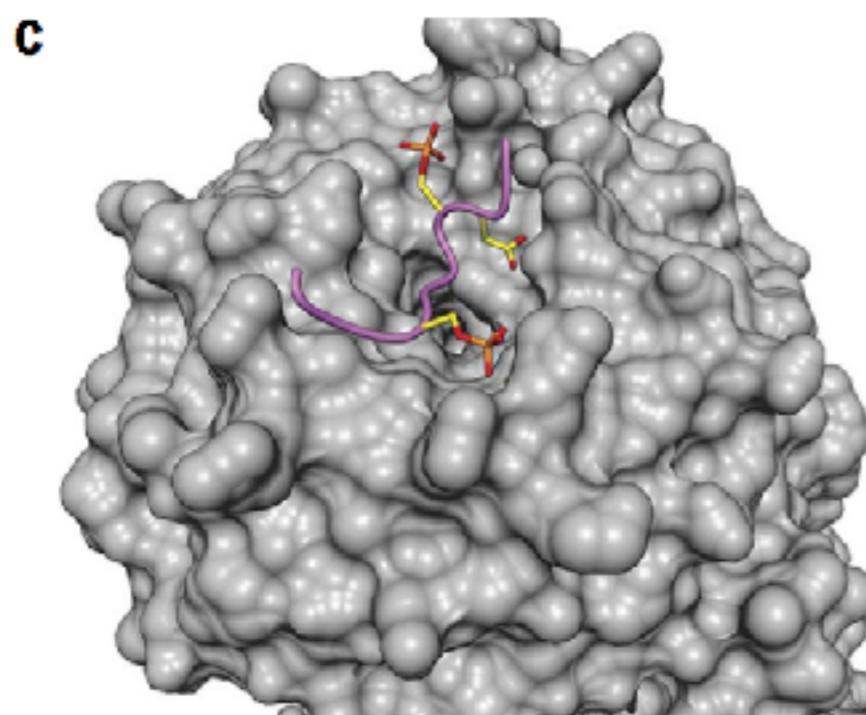
A. p53 - MDM2



B. p27 - SKP2-CKS1



C. Beta-Catenin - TrCP1



D. TRB1 - COPI

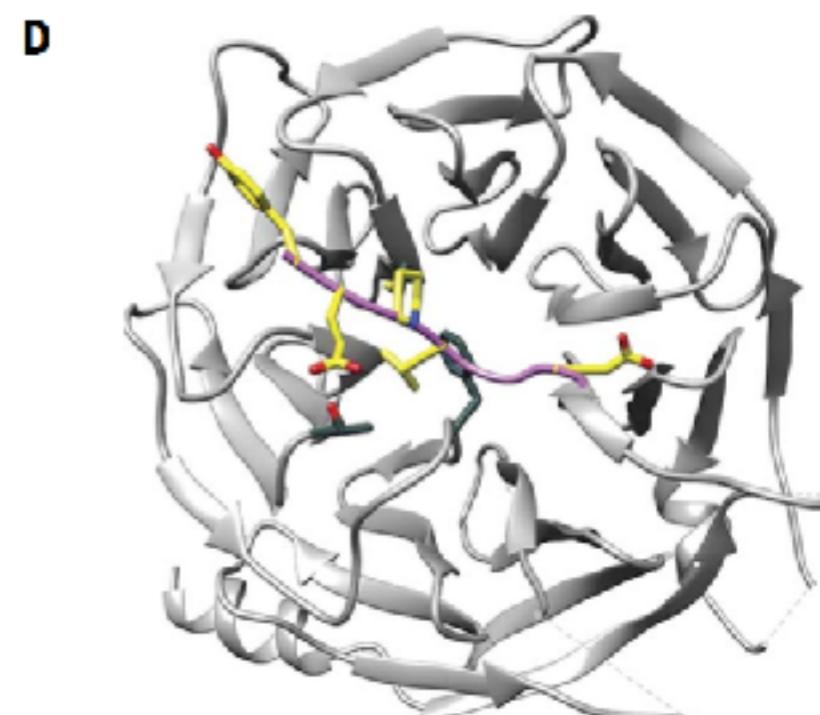
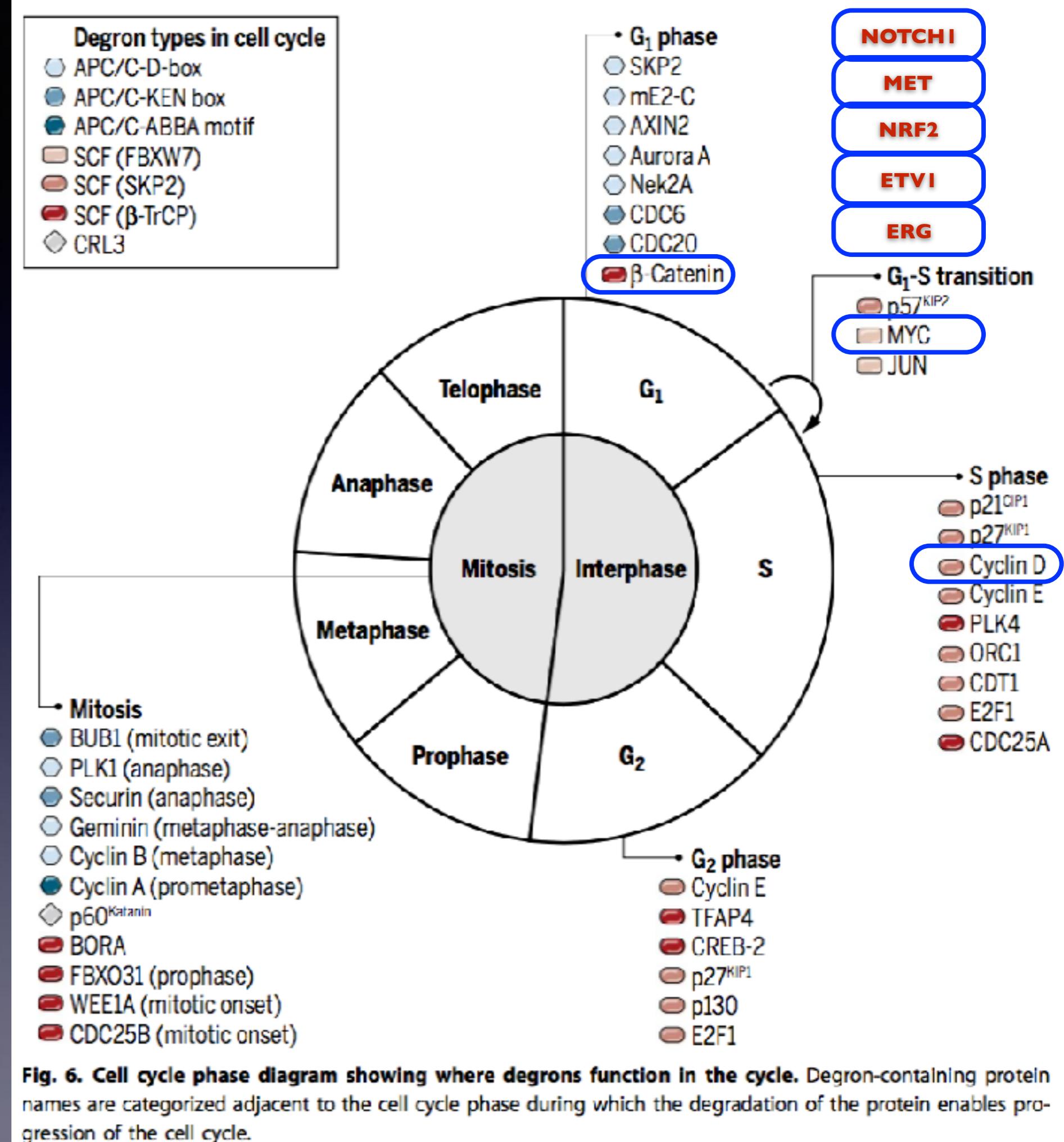


Fig. 3. E3 ligase-degron complexes. (A) p53-MDM2, the p53 degron peptide, enters a deep hydrophobic pocket on MDM2. The three key hydrophobic residues from p53 peptide are shown [Protein Data Bank (PDB): 1YCR]. (B) Structure of SKP1-SKP2-CKS1 in complex with the p27^{KIP1} phosphodegron peptide. SKP2 is rendered in gray surface, whereas CKS1 is represented in dark gray surface, with red (oxygen) and blue (nitrogen) polar functional groups. The small peptide from the p27^{KIP1} is rendered in ribbon representation with purple color: Its phosphothreonine interacts with positively charged CKS1 surface (blue region), which provides the phosphospecificity. SKP1 is not shown for clarity (PDB: 2AST). (C) Doubly phosphorylated β-catenin degron motif (see Table 1) in complex with β-TrCP1 and SKP1. Both of these molecules are shown in gray surface representation, whereas the β-catenin fragment is rendered as purple ribbon (PDB: 1P22). (D) COP1 E3 ligase with TRB1 degron. TRB1 binds on the conserved surface of COP1 in an extended manner. COP1 with WD40 repeats (β-propellers) is rendered as ribbon, whereas TRB1 peptide is rendered in stick representation and colored purple. (PDB: 5IGQ). Interchain H bonds are represented in magenta. Motif-defining positions have been rendered as yellow sticks for all the cases.

Many degrons get mutated in cancer

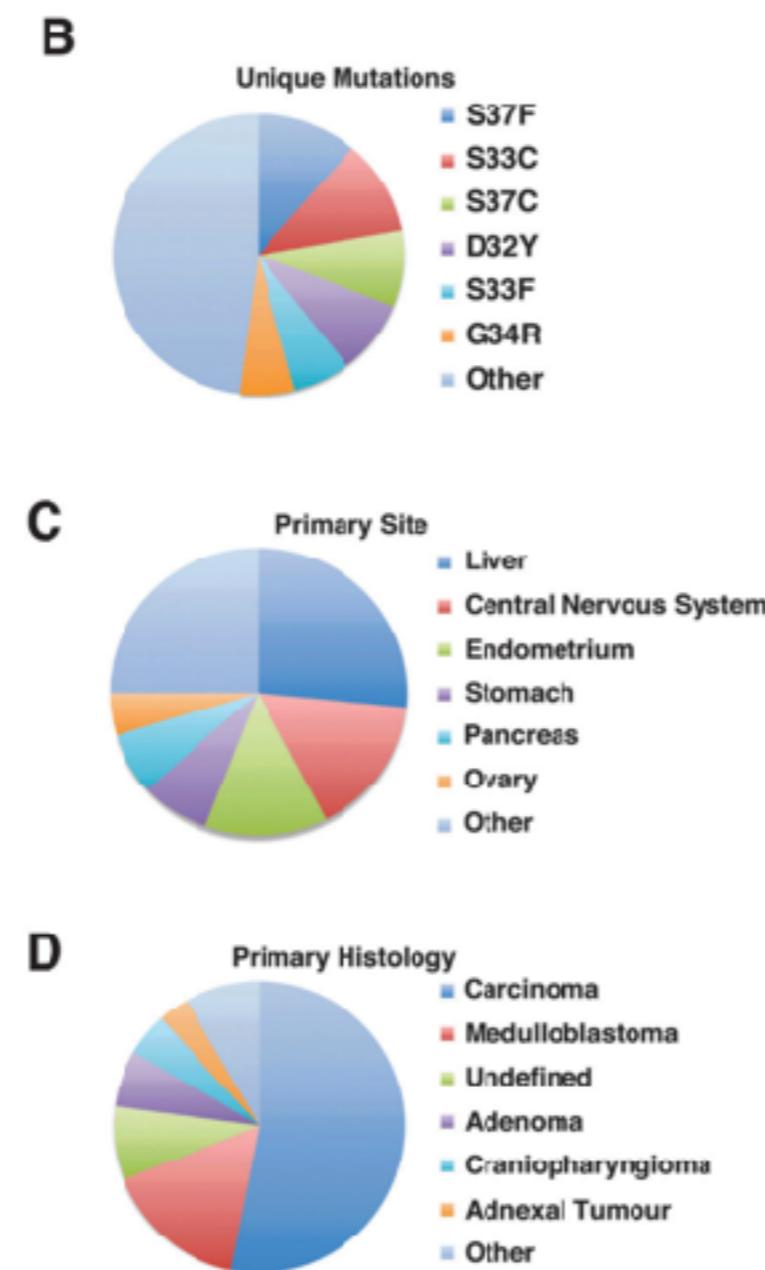
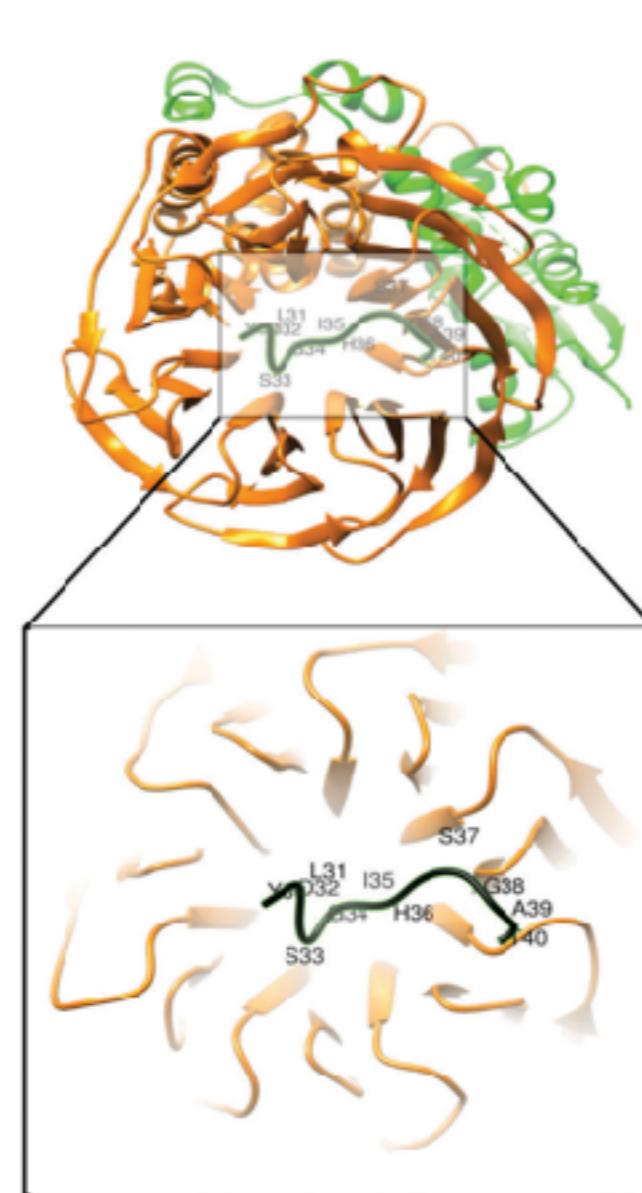
-
But only in
tumour
promoters,
never in tumour
suppressors

-
Degrongs will
need to be
scored for
targeted
therapies



Proteome-wide analysis of human disease mutations in short linear motifs: neglected players in cancer?†

Bora Uyar,*^a Robert J. Weatheritt,^{bc} Holger Dinkel,^a Norman E. Davey^{ad} and Toby J. Gibson*^a

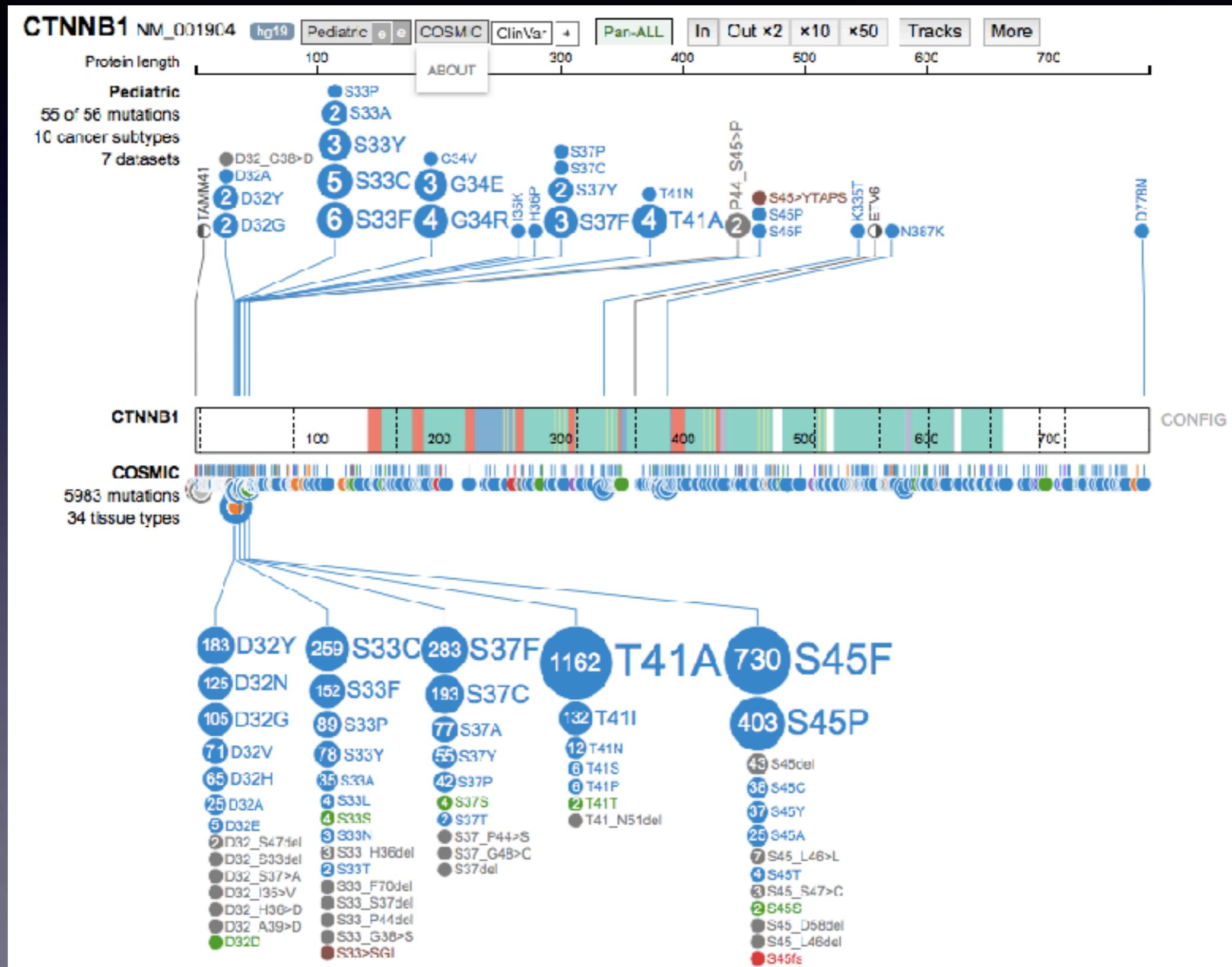


Mol. BioSyst., 2014, 10, 2626–2642

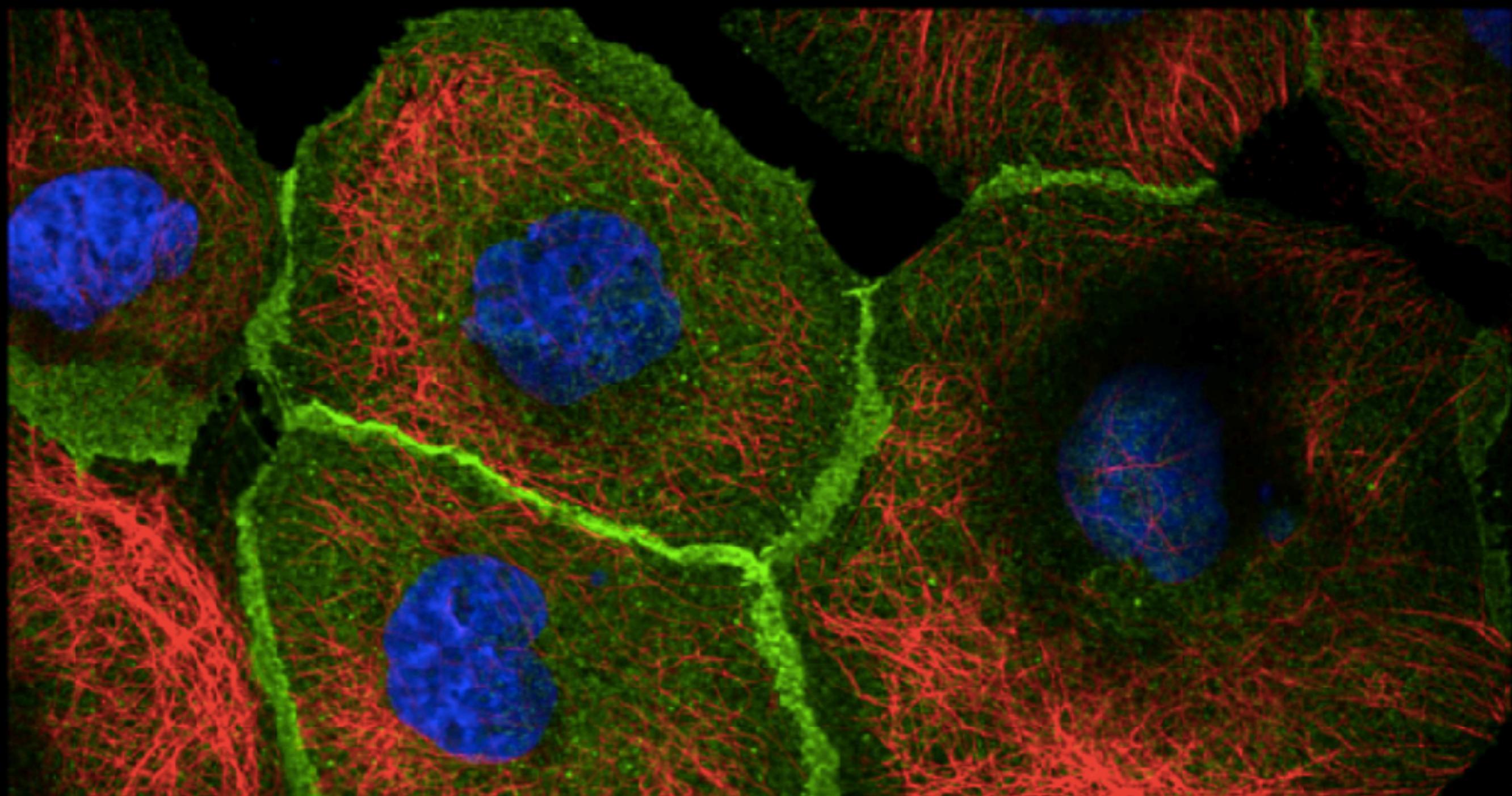
The Beta-Catenin
DSGxxS degron is
the most mutated
SLiM in cancer

According to the available disease-related missense mutation datasets, the most recurrently mutated experimentally validated SLiM is the conserved proteasomal degradation motif (“degron”) in the highly disordered N-terminal region of β-catenin ([Fig. 3A](#)). This motif (DEG_SCF_TRCP1_1, ₃₂D_PSGIH_PS₃₇)

Cancer mutations in Beta-Catenin target its degron (displayed by PeCan (St Judes))



Wnt signalling Beta-Catenin transcription factor imaged at the Human Protein Atlas



Degrons are at the **nexus** between
mutation and gene expression changes
in cancer

Their roles in cancer are a manifestation of
the importance of correct gene dosage in
cell regulation

Viral hijack of cell regulation:

Fragility in a complex system

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, 99, 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

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)

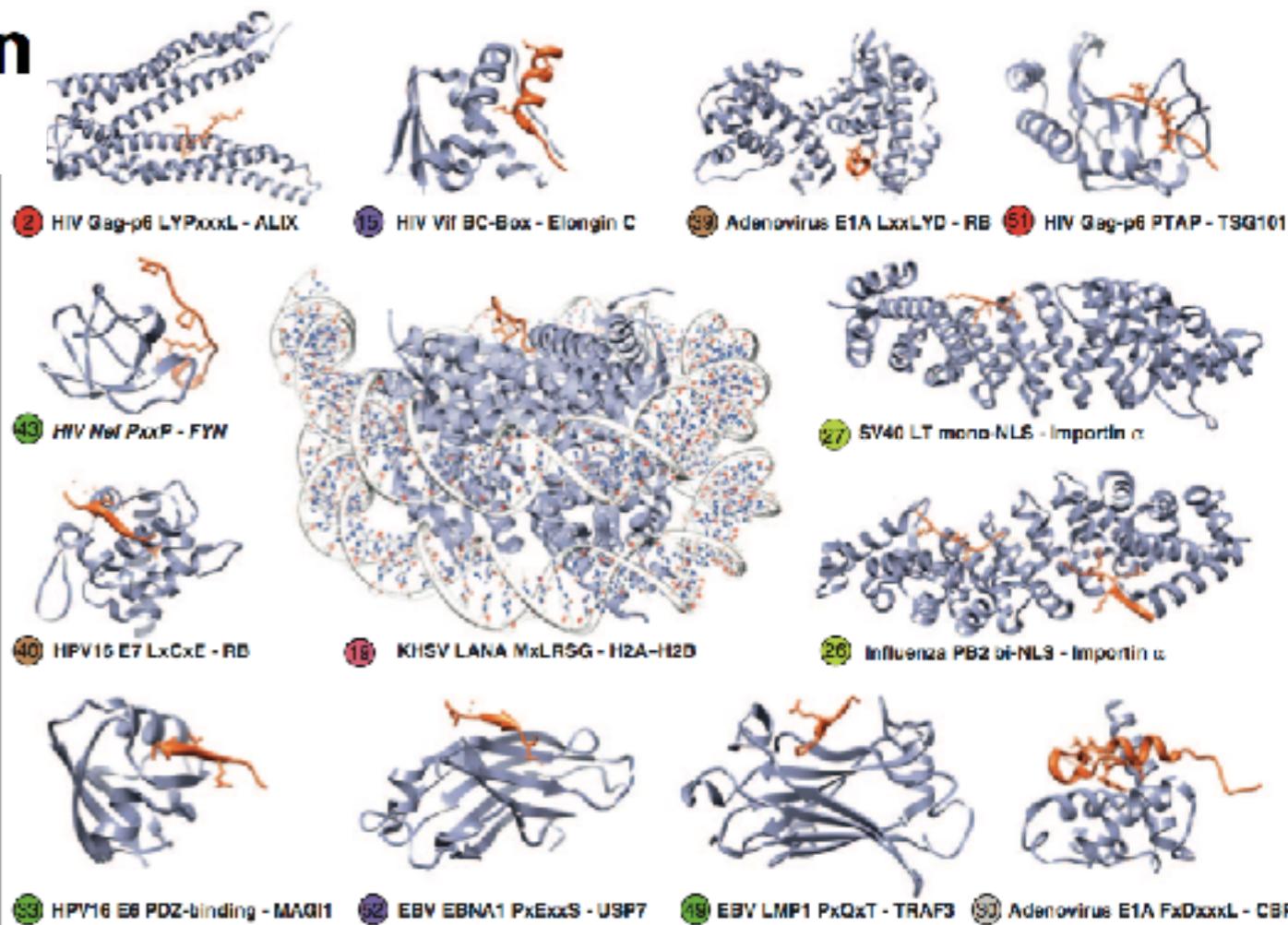
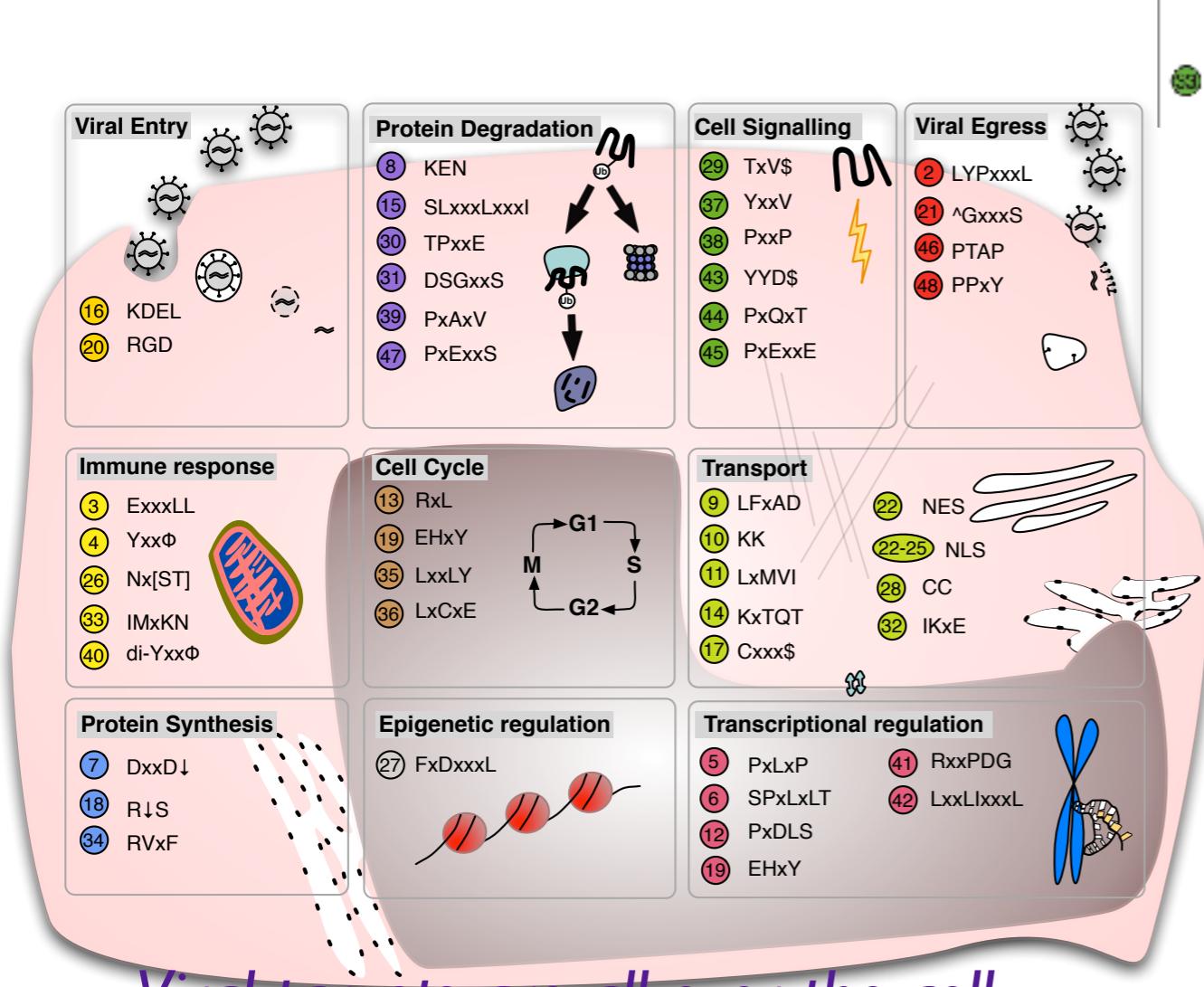
that is, unaffected by random component failures but vulnerable to targeted attacks on its key components.

How viruses hijack cell regulation

Norman E. Davey¹, Gilles Travé² and Toby J. Gibson¹

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

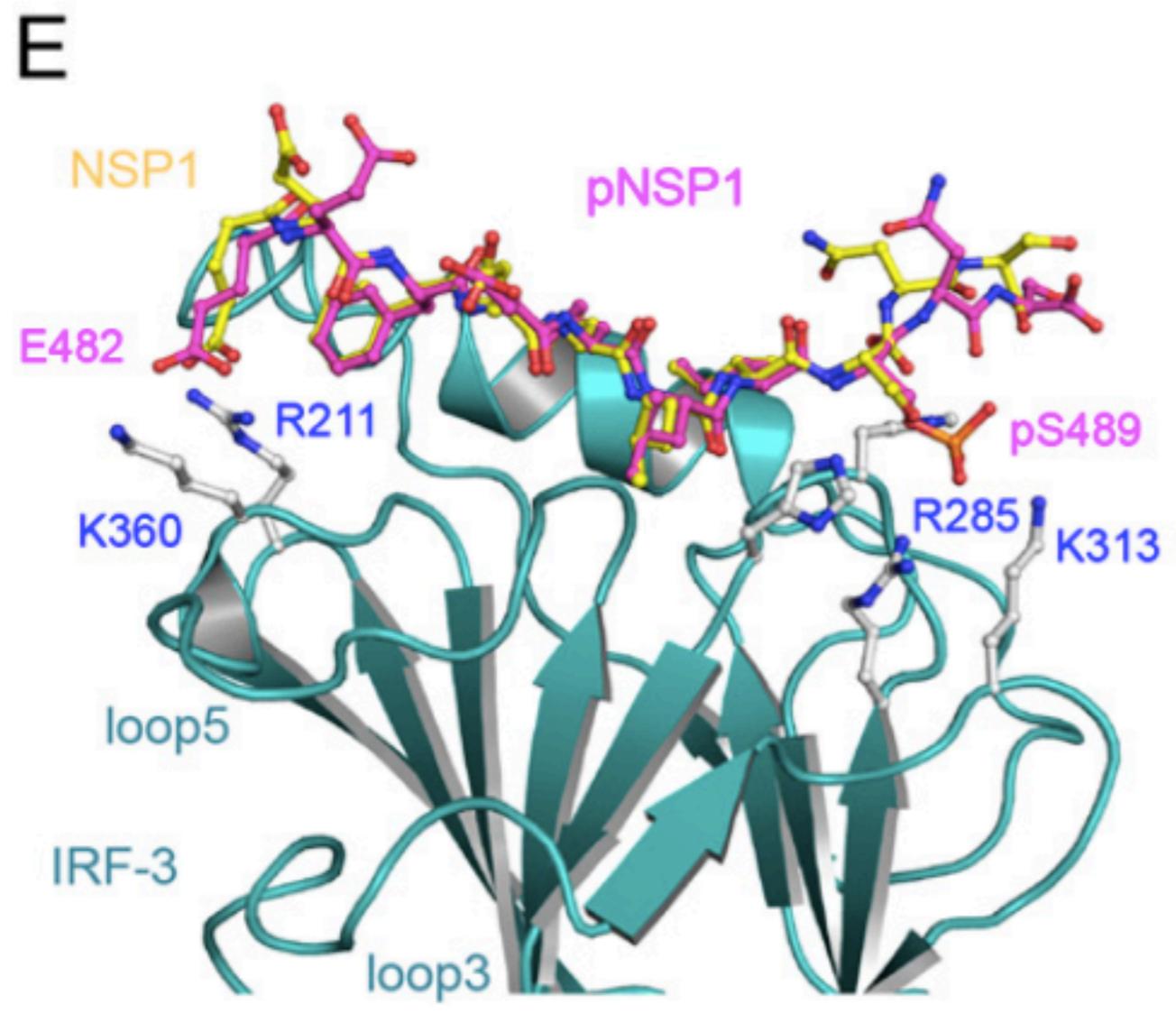


- Why is there “always”
a cellular protein motif
interaction for a virus
to subvert?
- What does this tell
us about the nature of
the cell?

Structural basis for concerted recruitment and activation of IRF-3 by innate immune adaptor proteins

Baoyu Zhao^{a,1}, Chang Shu^{a,1,2}, Xinsheng Gao^b, Banumathi Sankaran^c, Fenglei Du^a, Catherine L. Shelton^{d,e}, Andrew B. Herr^{d,e}, Jun-Yuan Ji^b, and Pingwei Li^{a,2}

approved April 27, 2016



Significance

Type I IFNs are key cytokines involved in antiviral immunity. A number of innate sensing pathways regulate the induction of type I IFNs. These pathways converge at the activation of the transcription factor IRF-3 (IFN regulatory factor 3). Three different adaptors mediate the recruitment of IRF-3 using a conserved structural motif. In this study, we determined the molecular mechanisms by which these adaptors recruit IRF-3 upon phosphorylation, the mechanism of IRF-3 activation, and how rotavirus subverts these signaling mechanisms to evade innate immune surveillance. These results provide critical insights into the molecular basis of innate immunity against microbial and viral infections.

The pLxIS motif of rotavirus NSP1 targets IRF-3 for degradation

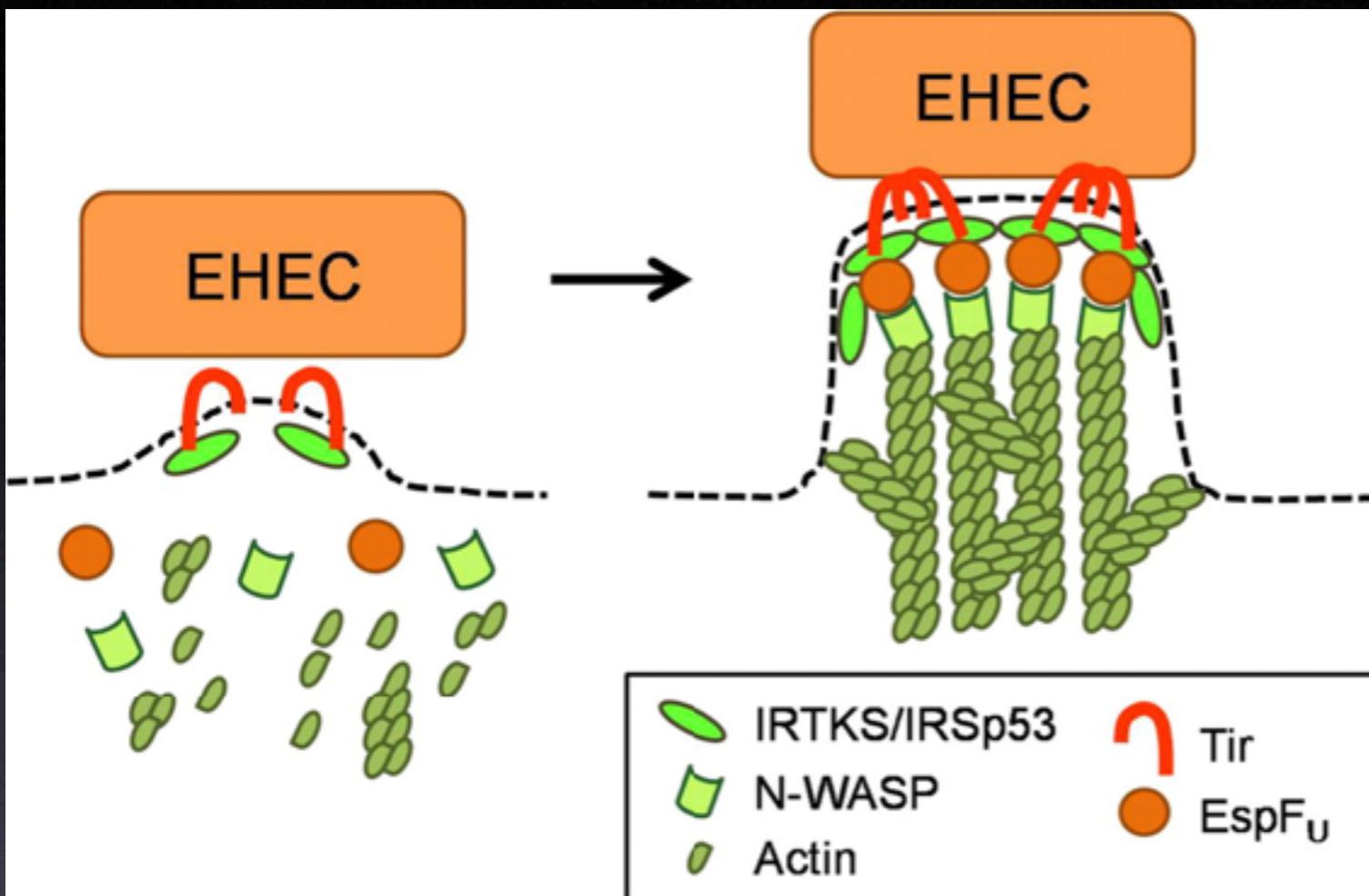
See *ELM Entry LIG_IRF3_LxIS_1*

Living pathogens use three main approaches to perturb host cells

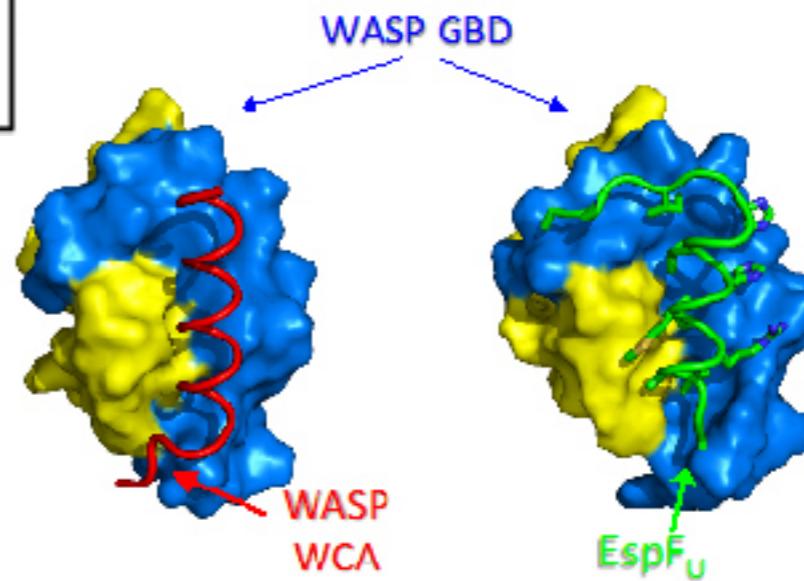
- Toxins
 - **Staphylococcus:** Pore-forming toxin
 - *The cell will die*
 - Enzymatic irreversible inactivation
 - **Yersinia:** Acetylates MapK activation loop Ser and Thr
 - *The cell will eventually die*
 - Linear motif hijack
 - **Toxoplasma:** MapK docking motifs activate MapK
 - *Cell can be re-tuned and kept alive indefinitely*

Pathogenic Pedestal Formation

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

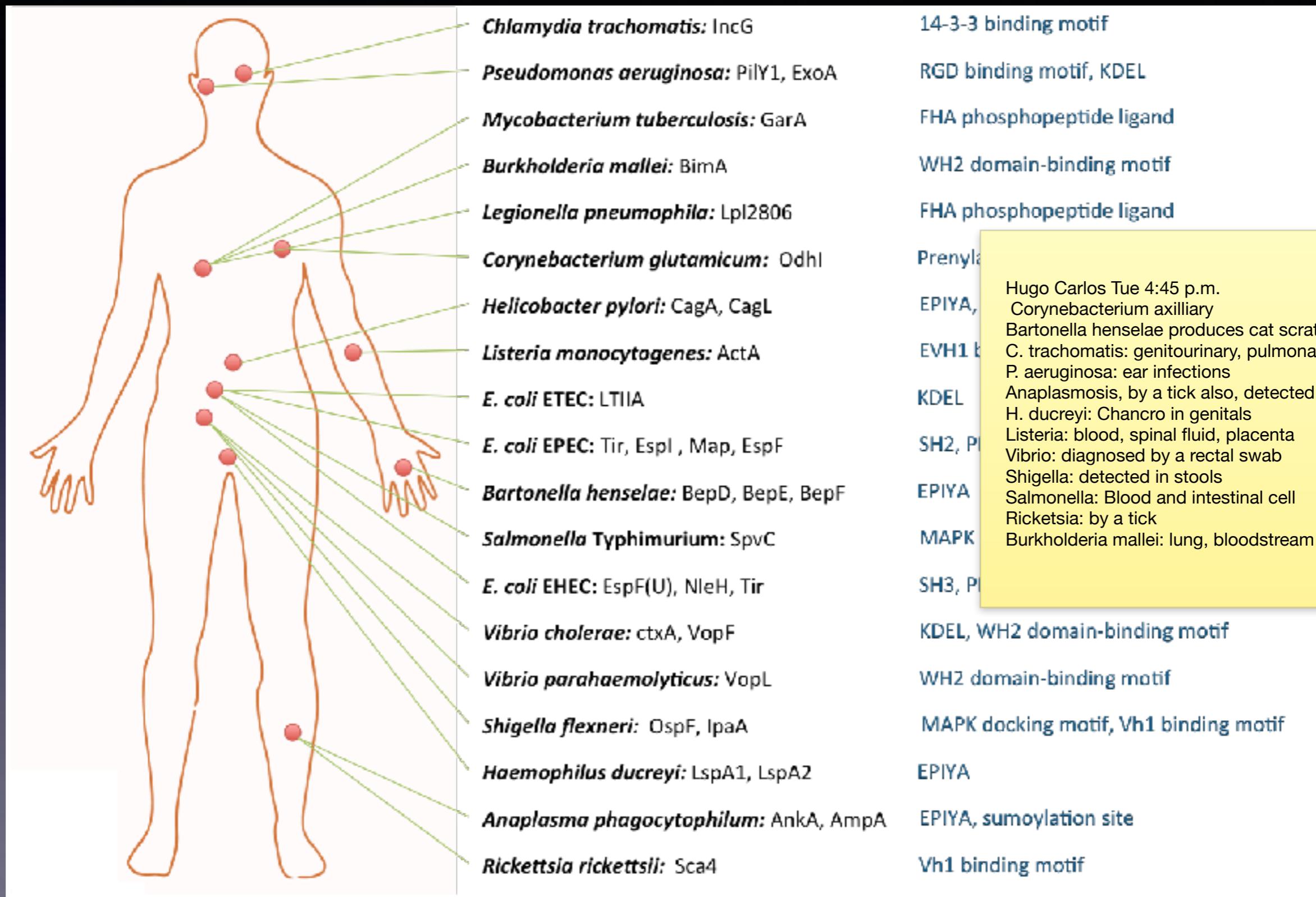


Yi PNAS, 106, 6431 (2009)

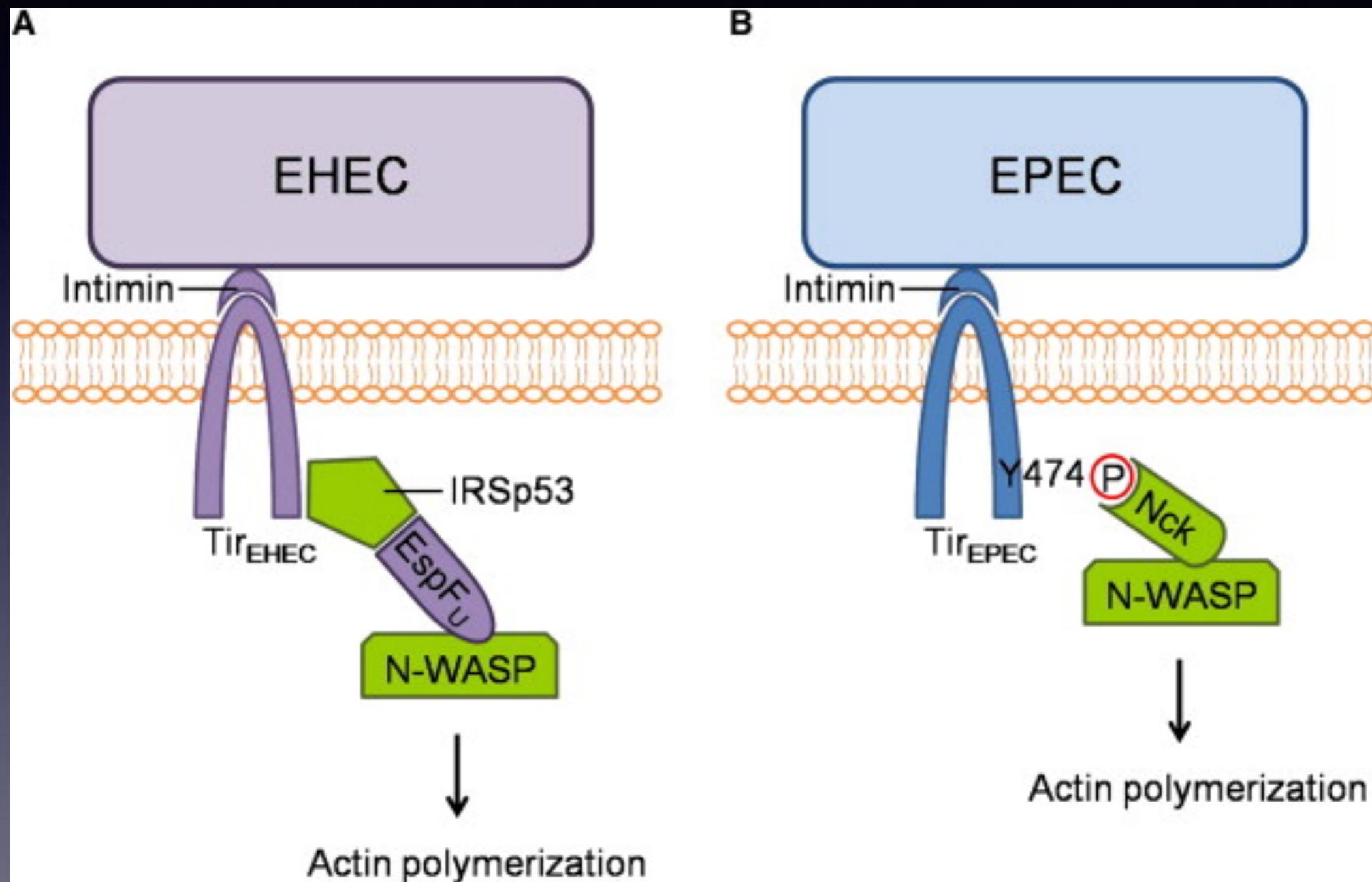


Cheng, Nature, 454, 1009 (2008)

We are collecting reported examples of SLiMs for benchmarking e.g. Human Bacterial Pathogens Using Motif Mimics

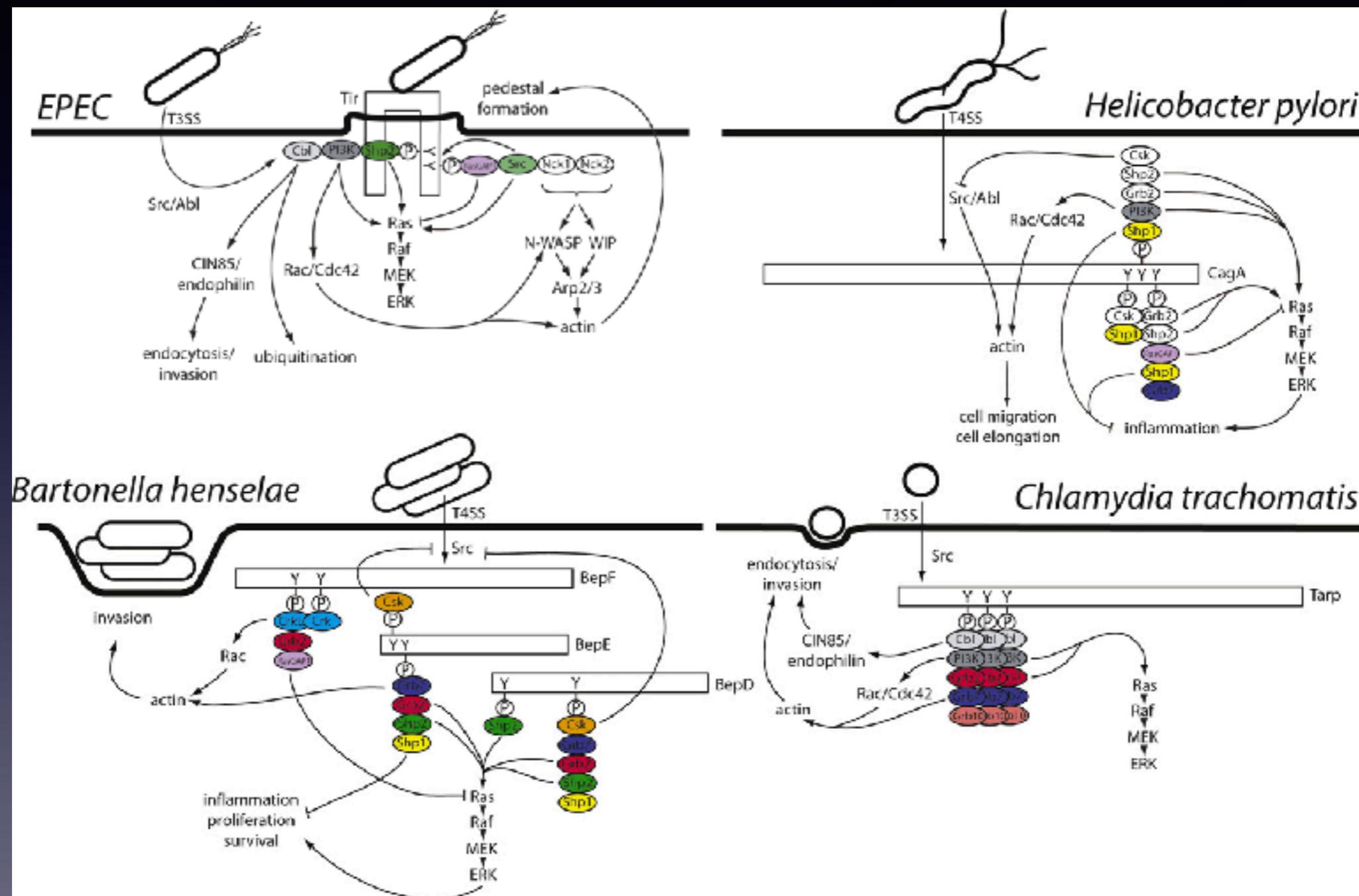


Pathogenic *E. coli* EHEC and EPEC use Abl and Src kinases to polymerise actin and make pedestals



Yi et al. Cell Host Micro. (2009)

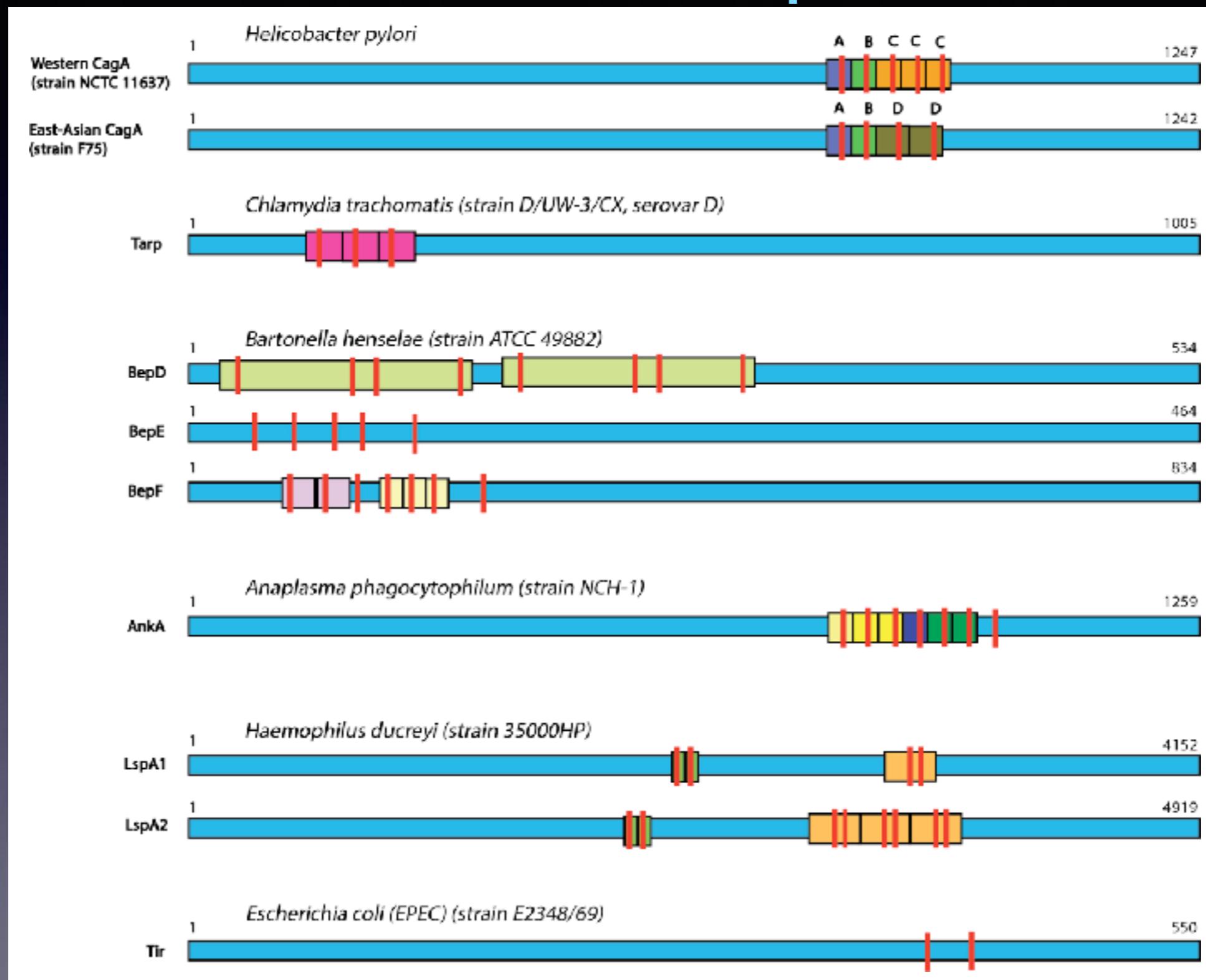
Cellular proteins that interact with tyrosine-phosphorylated bacterial proteins



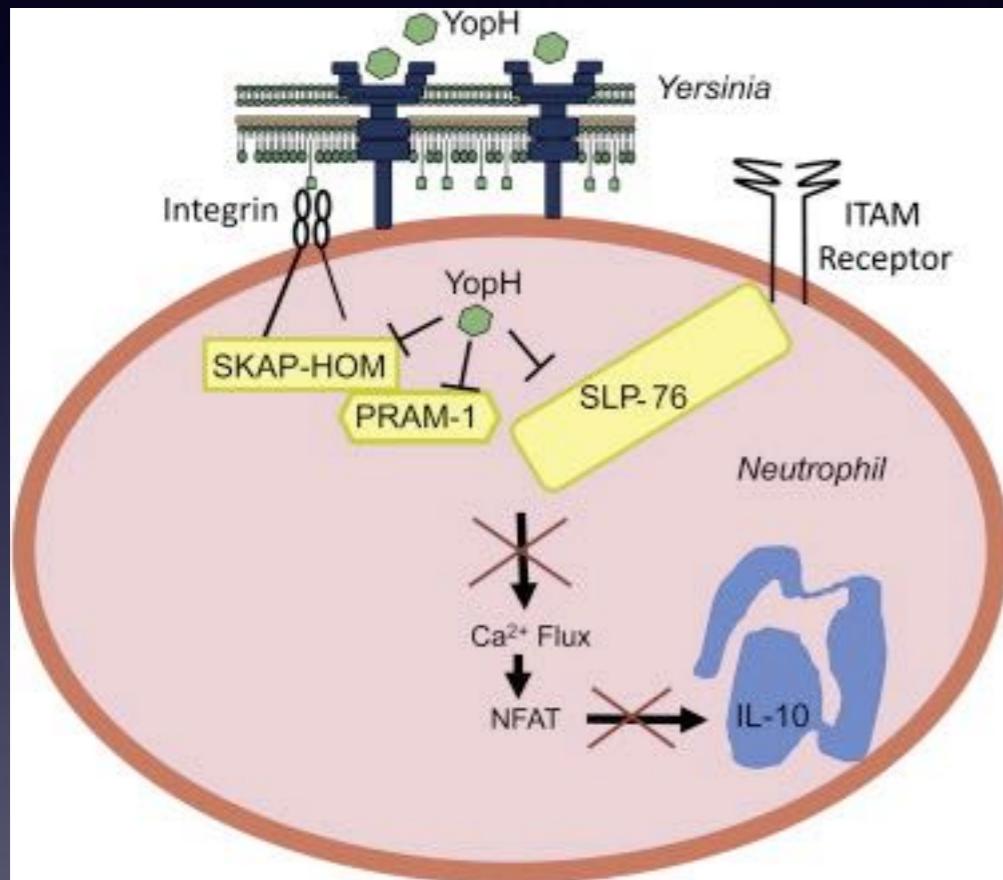
SILAC experiment recovered interactions with tyrosine phosphorylated bacterial proteins

Selbach, M., et al. 2009

EPIYA-like motifs in *H. pylori* CagA and other bacterial effector proteins



Yersinia pestis introduces a tyrosine phosphatase to impair neutrophil responses



YopH selectively dephosphorylates host proteins

Calcium flux is blocked and cytokine production altered

Pseudomonas aeruginosa internalization is Abl kinase dependent

OPEN ACCESS Freely available online

PLOS PATHOGENS

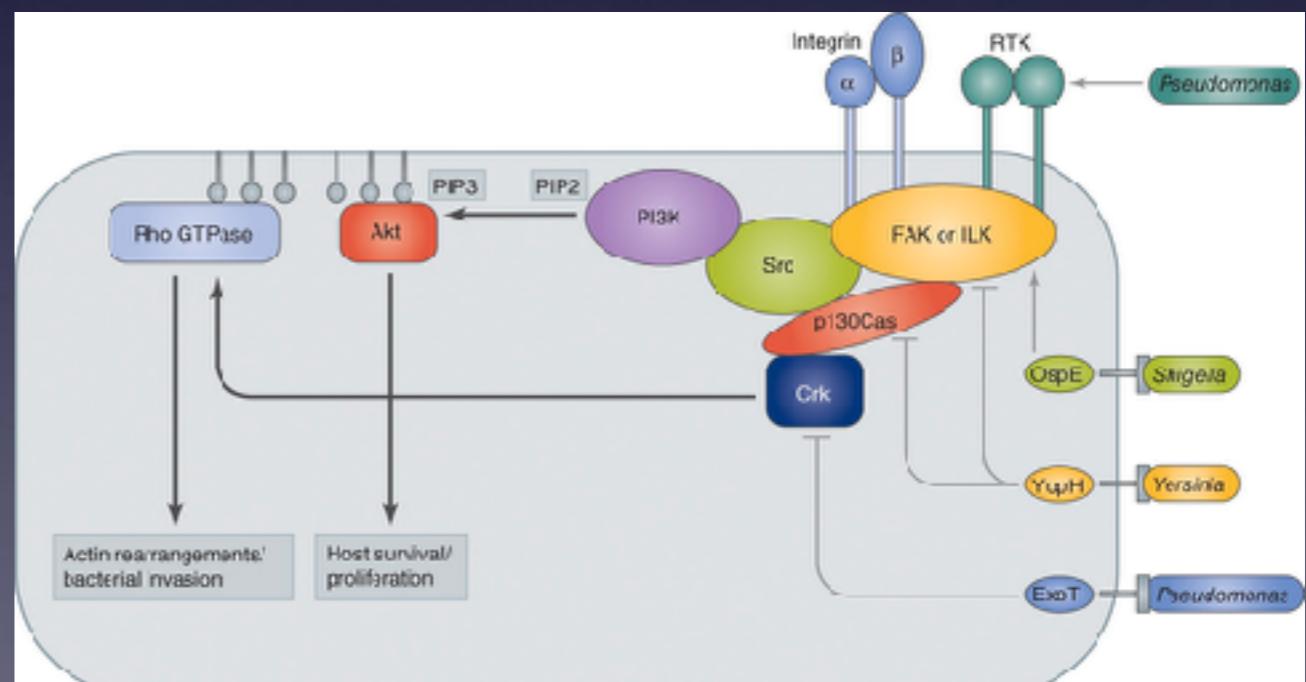
RNAi Screen Reveals an Abl Kinase-Dependent Host Cell Pathway Involved in *Pseudomonas aeruginosa* Internalization

Julia F. Pielage^{1,2,3}, Kimberly R. Powell⁴, Daniel Kalman⁴, Joanne N. Engel^{1,2,3*}

~80 genes were inactivated using RNAi

Abl is a component of the host signaling pathway that leads to internalisation of *P. aeruginosa*

ExoT abrogates *P. aeruginosa*-induced Abl-dependent Crk phosphorylation



Pielage, J. F., et al., 2008
Krachler A. M., et al., 2011

Can we use Tyrosine Kinase inhibitors to treat bacterial infections?

Pathogenic perturbation of signalling networks might enable a general therapeutic strategy:

Can inhibitors of signalling set the cell into a condition that is hostile to the pathogen?

Would pathogens find it difficult to evolve resistance to such therapies?

In this course, we want to help you to use bioinformatics tools to augment your research and help answer questions like this one

