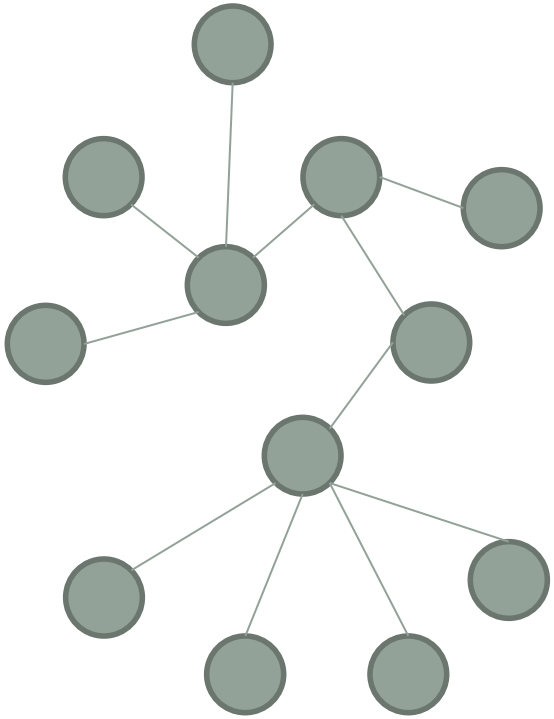
Integrative and quantitative analysis of disease mutations in protein interaction networks and implications for personalized medicine

Christina Kiel, Staff scientist
Department of Systems Biology, Luis Serrano group
CRG Barcelona

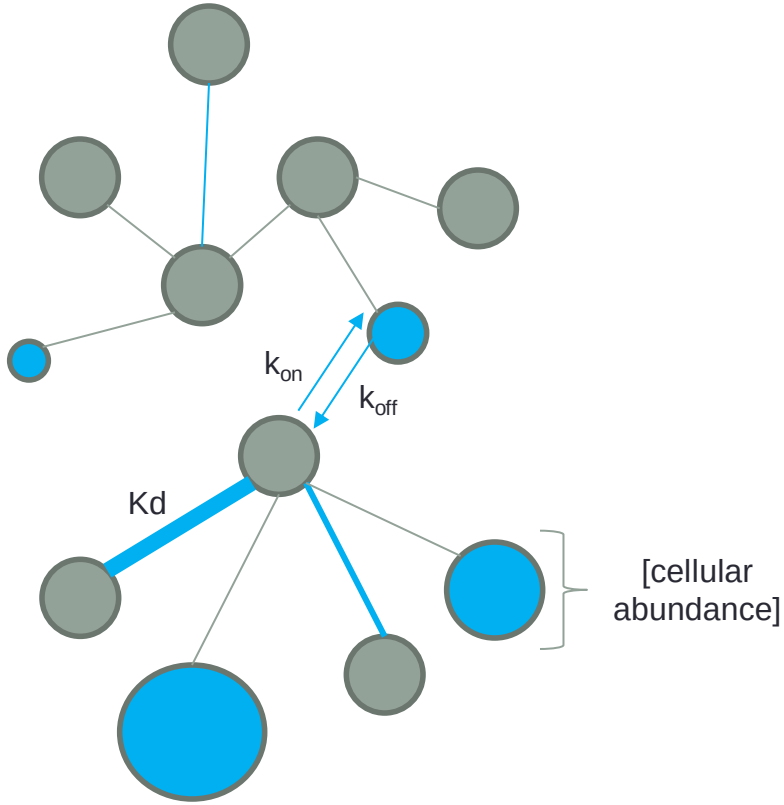
CRG Barcelona: http://crg.eu

Quantitative information in protein-protein interaction (PPI) networks

Qualitative PPI networks



Quantitative PPI networks



Considering protein abundances and affinities/ kinetic constants

Outline

I. The effect of affinities, kinetic constants and network topology in PPI networks

II. The effect of protein abundance perturbations and interaction competition in PPI networks

III. Methods to quantify protein abundances, affinities, and kinetic constants

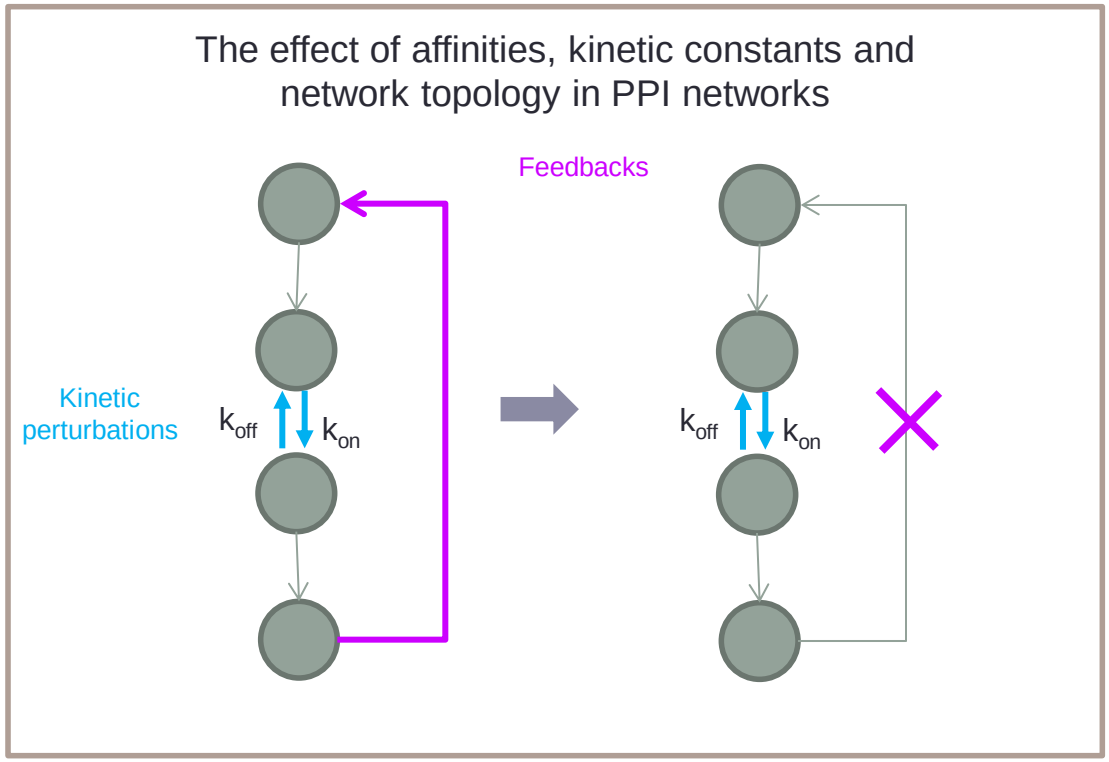
IV. Disease mutations and their principle effect on PPI networks

V. Examples for quantitative effects in disease networks

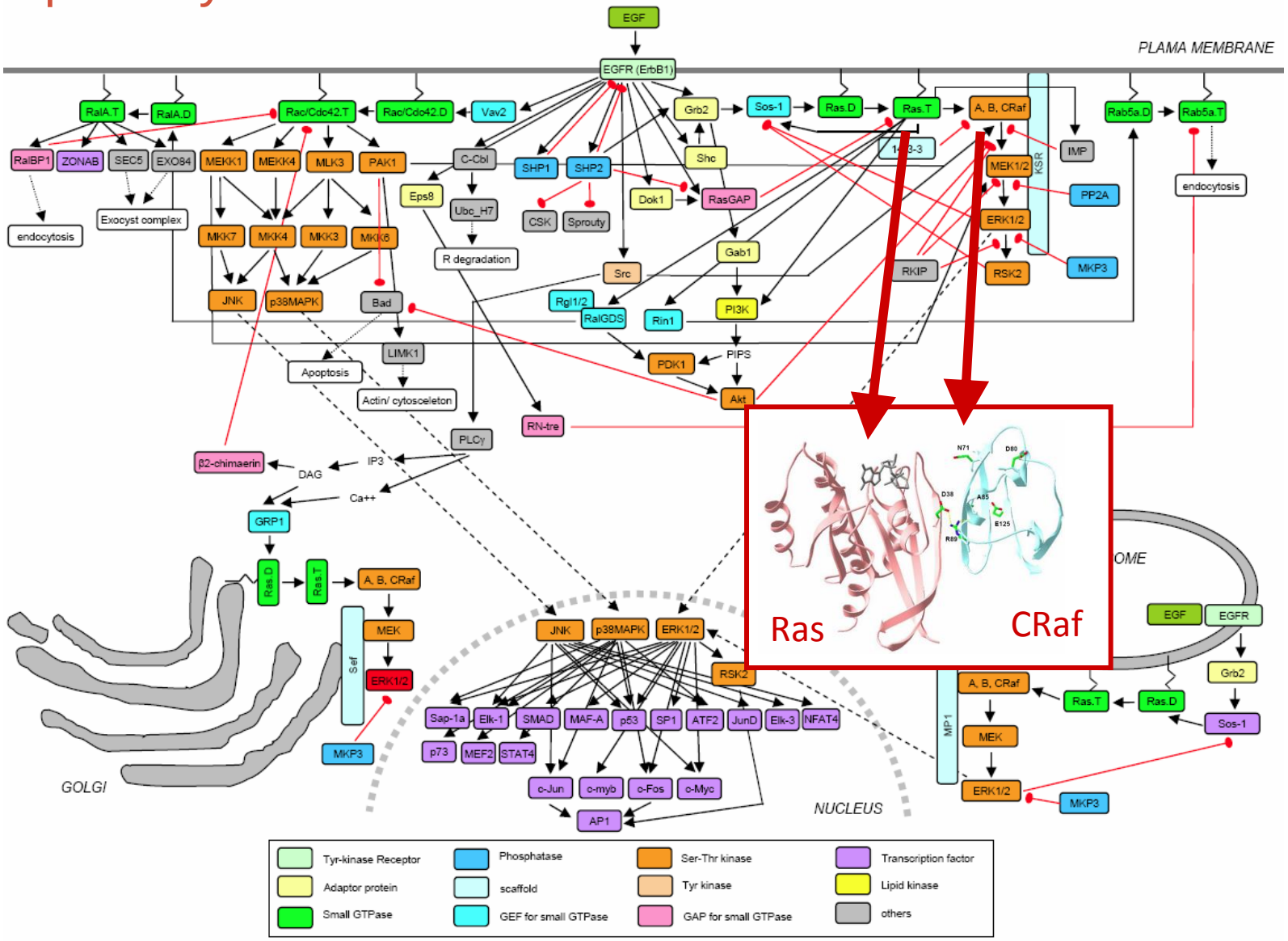
1. RASopathy vs cancer mutations: a matter of quantity
2. BRAF mutation frequency: prediction of oncogenic drivers

VI. Summary tools & websites

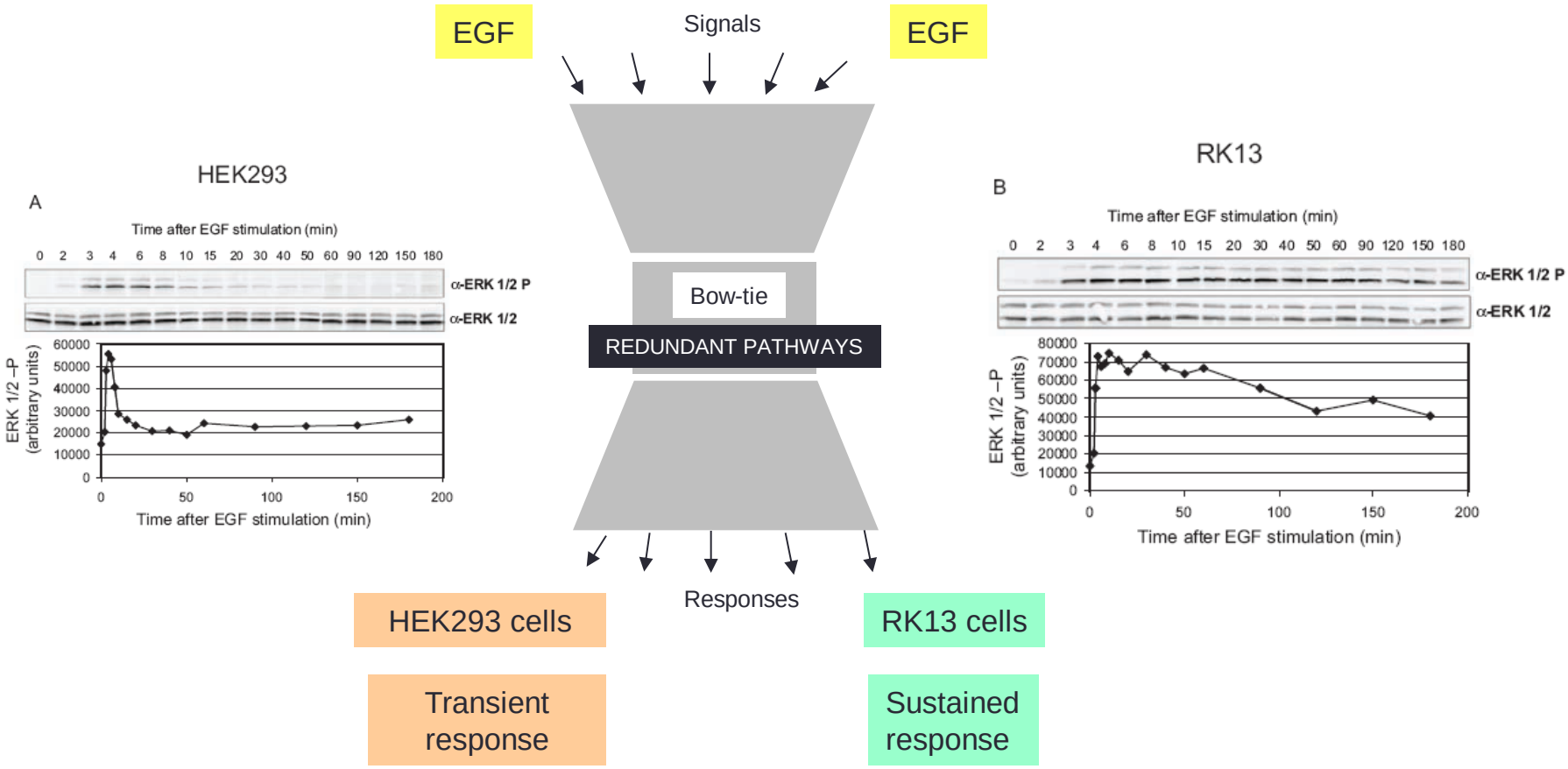
VII. Wrap up/ discussion/ conclusions


The effect of affinities, kinetic constants and network topology in PPI networks
Feedbacks
Kinetic perturbations
k_{off}
k_{on}
k_{off}
k_{on}

Epidermal growth factor (EGF) activates the RAS-RAF-MEK-ERK pathway



Cell type-specific ERK activation in HEK293 and RK13 cells



Different signaling response (ERK-p) with similar ligand (EGF)

Kiel & Serrano, Sci Signal, 2009

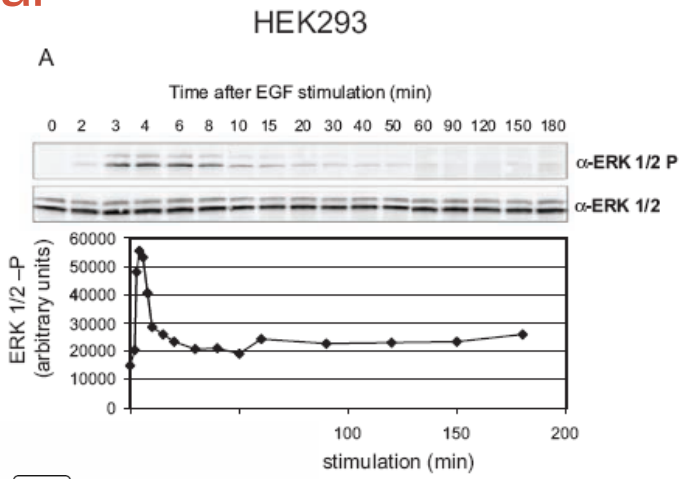
Different network 'wiring' /feedbacks causes the different behaviour

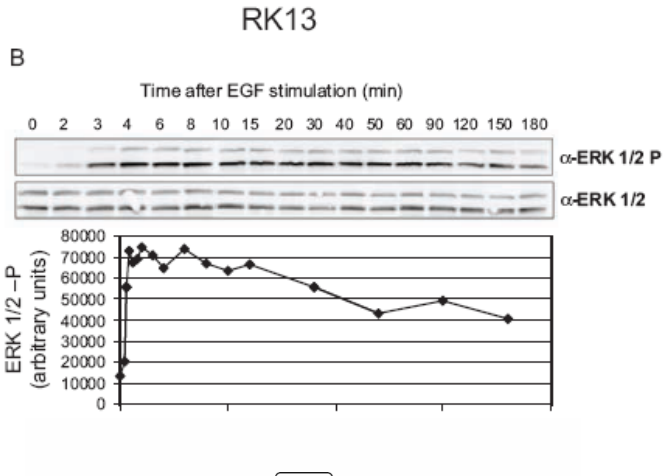
HEK293 cells

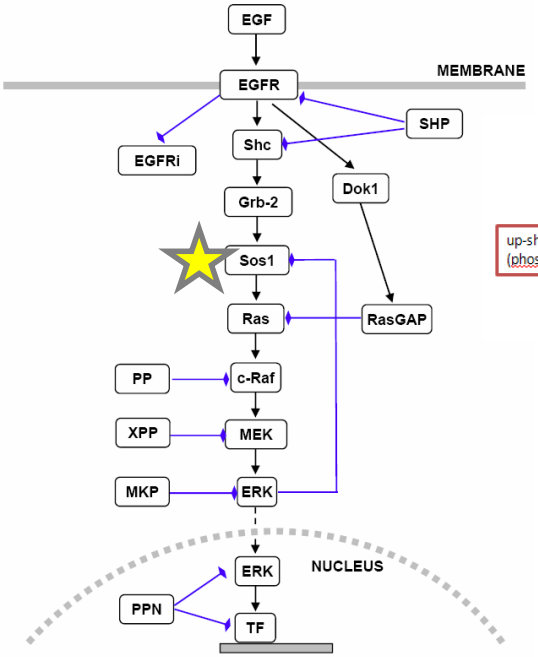
Transient response

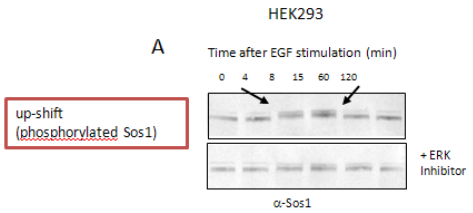
RK13 cells

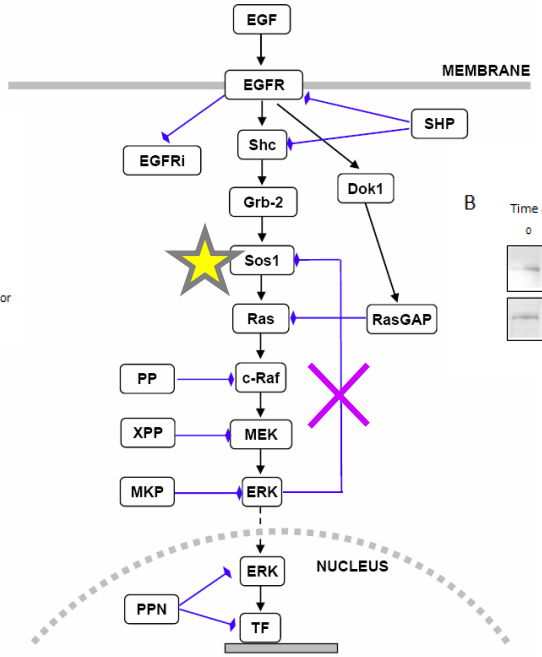
Sustained response

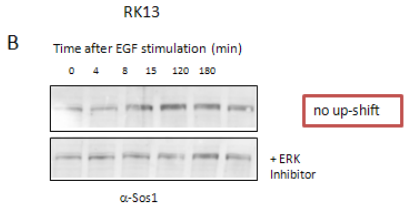






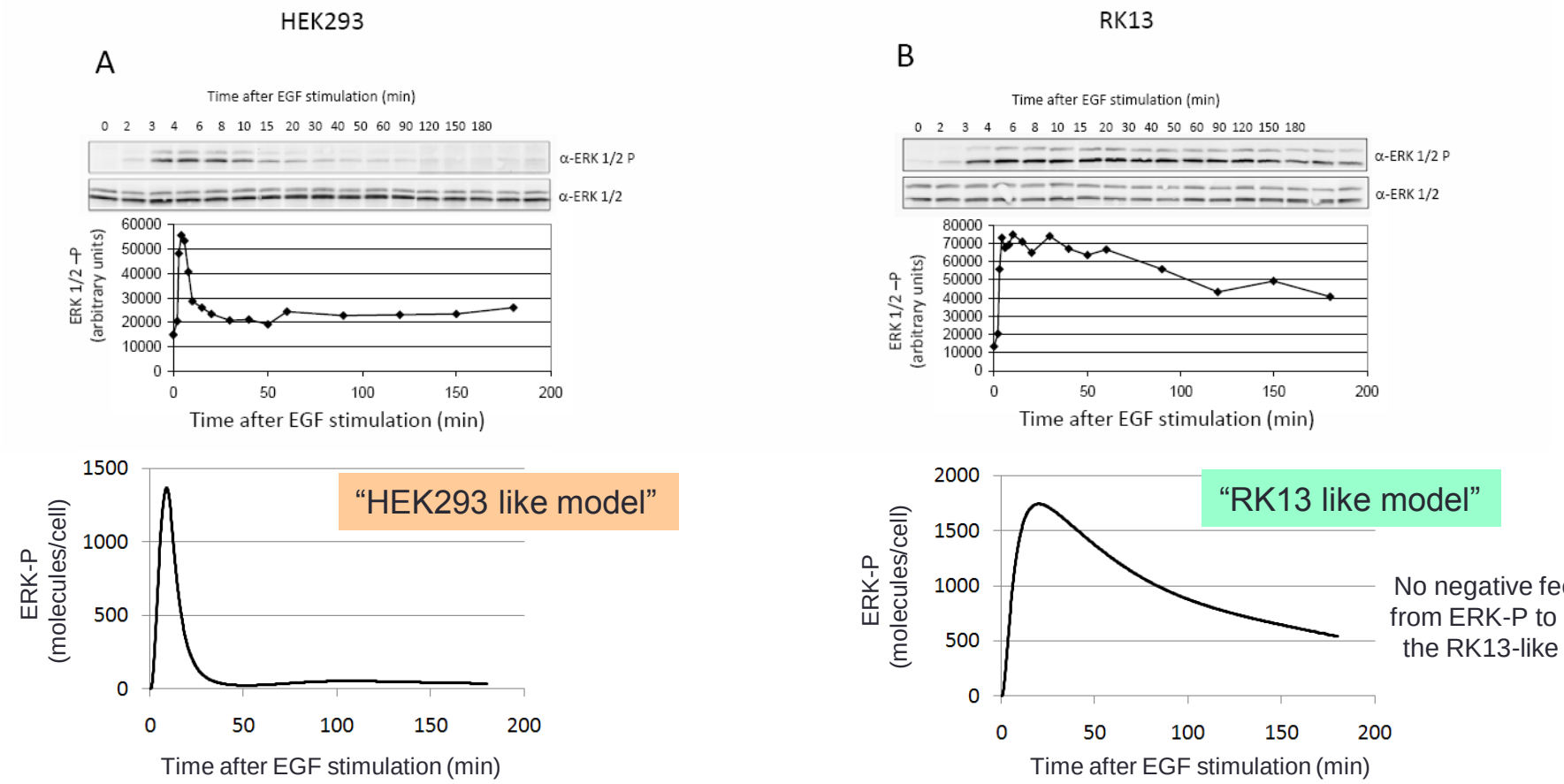






Kiel & Serrano, Sci Signal, 2009

A simple computer model of ERK activation in HEK293 and RK13 cells



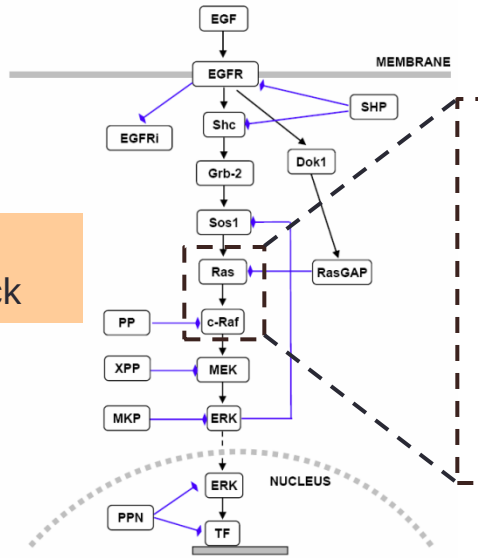
No negative feedback from ERK-P to Sos1 in the RK13-like model

- Good agreement of experiment and model predictions

Kiel & Serrano, Sci Signal, 2009

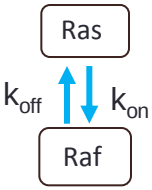
Model predictions: different cell type-specific wiring results in different responses to affinity perturbations

Strong feedback

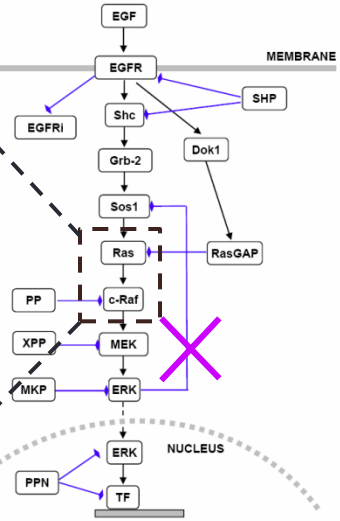


Subtle affinity changes

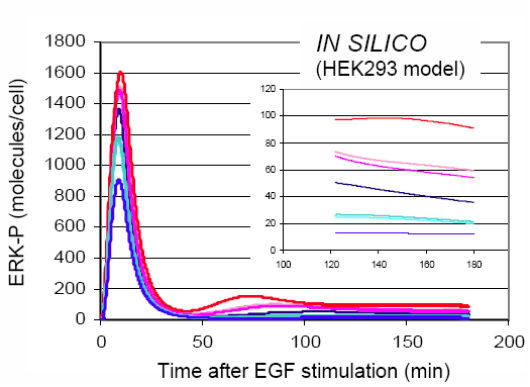
Kinetic perturbations



$k_D = \frac{k_{off}}{k_{on}}$

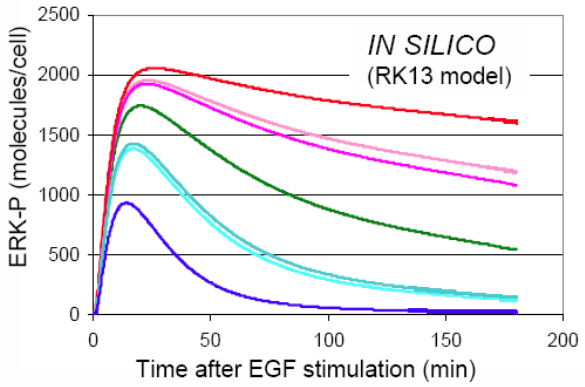


Weak feedback



No significant changes

- Mutations can have different cell type (**patient!**)-specific effects



Significant differences

Kiel & Serrano, Sci Signal, 2009

Experimental validation of the role of kinetic parameters in MCF7 cells (weak feedback)

Experimental design of mutants that introduce kinetic perturbations

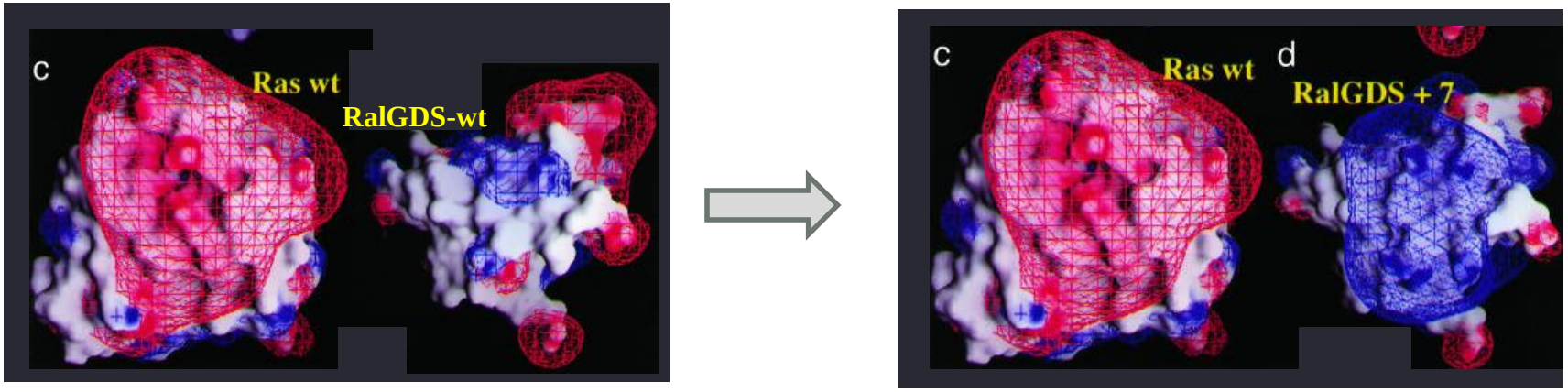
Affinity (Dissociation constant) → $K_d = \frac{k_{off}}{k_{on}}$

k_{off}: Dissociation rate constant

k_{on}: Association rate constant

E.g.:

↑ Increase k_{on}: improve electrostatic surface complementarity; *'electrostatic steering'*



Kiel et al., PNAS, 2004

↑ Increase k_{off}: mutate hot-spot residues in the interface

FoldX-based energy calculations of proteins

3D Structural information

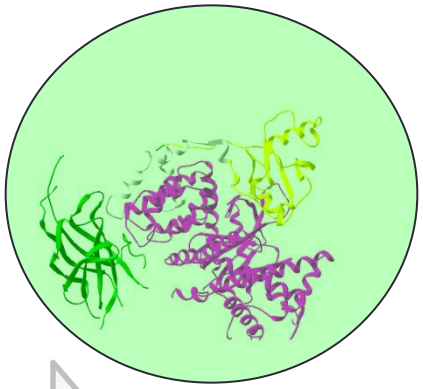
A force field for energy calculations and protein design

Schymkowitz et et al, *Nucleic Acids Res*, 2005

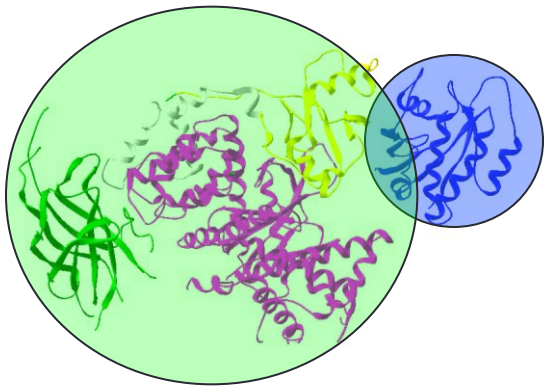
+ FoldX = ΔG

Relation to affinity: $\Delta G = RT \ln K_d$

✓ Total free energy

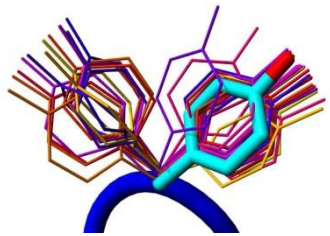


✓ Interaction energy



✓ Mutagenesis

A rotamer library to replace the 20 amino acids

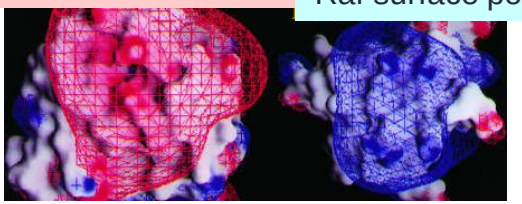


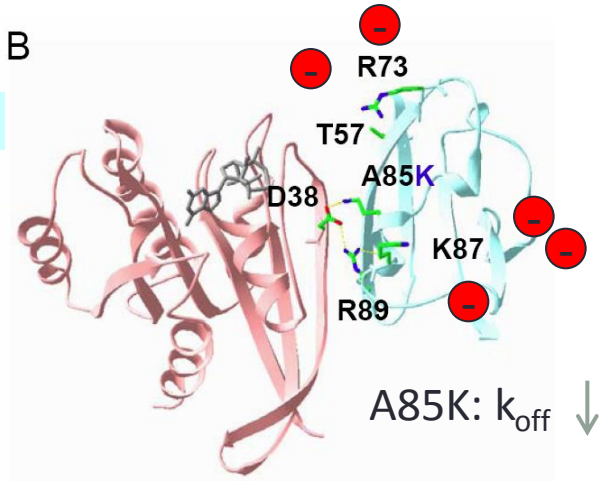
Protein design

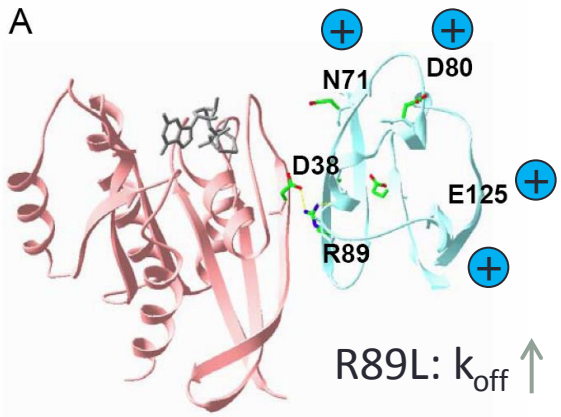
Summary of the protein mutant design

Ras surface negative

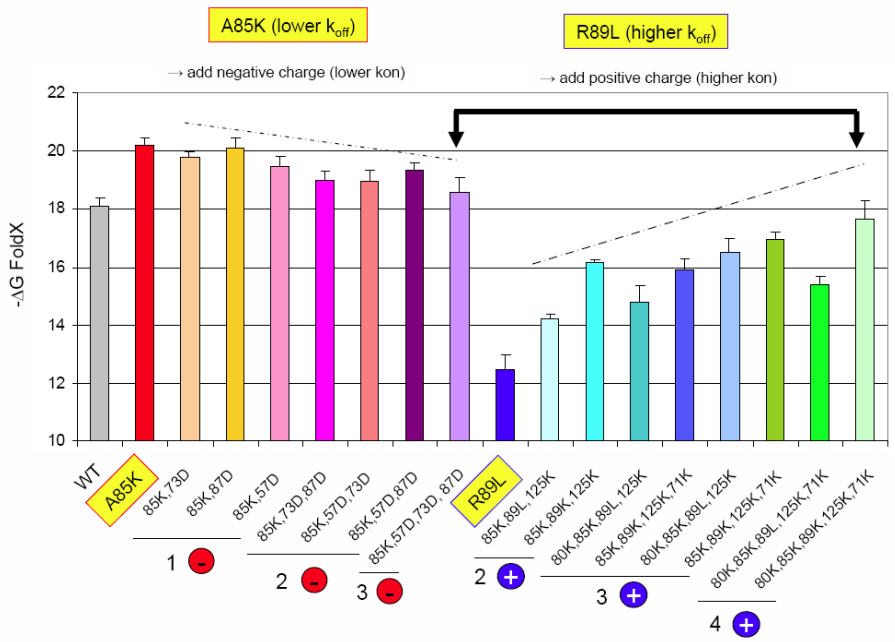
Raf surface positive





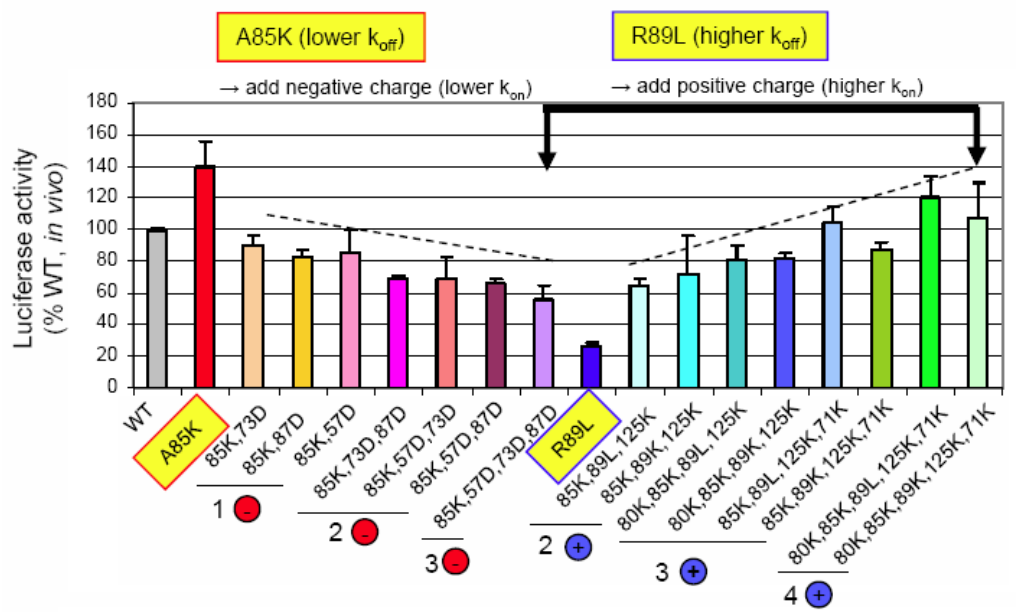


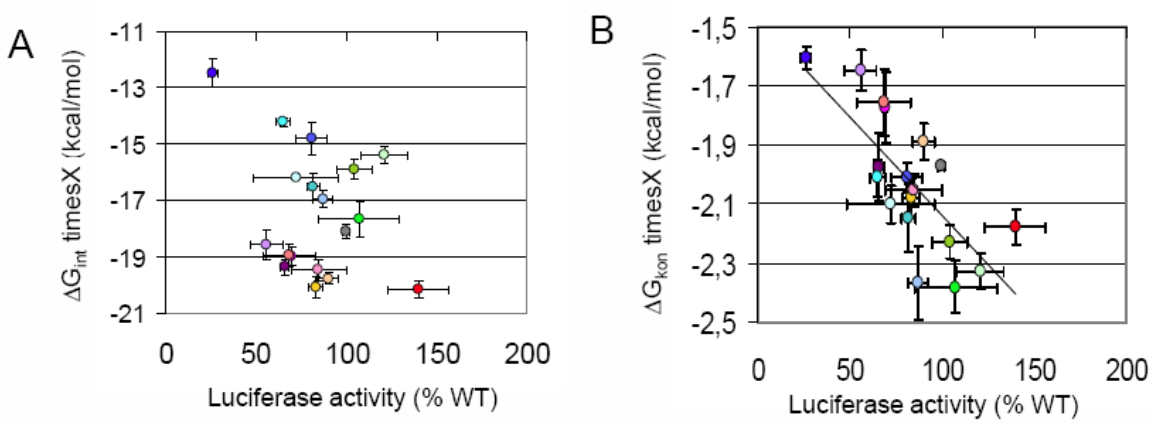
FoldX



Kiel & Serrano, Sci Signal, 2009

Analysis of all mutants in RK13 cells (luciferase activity assay)

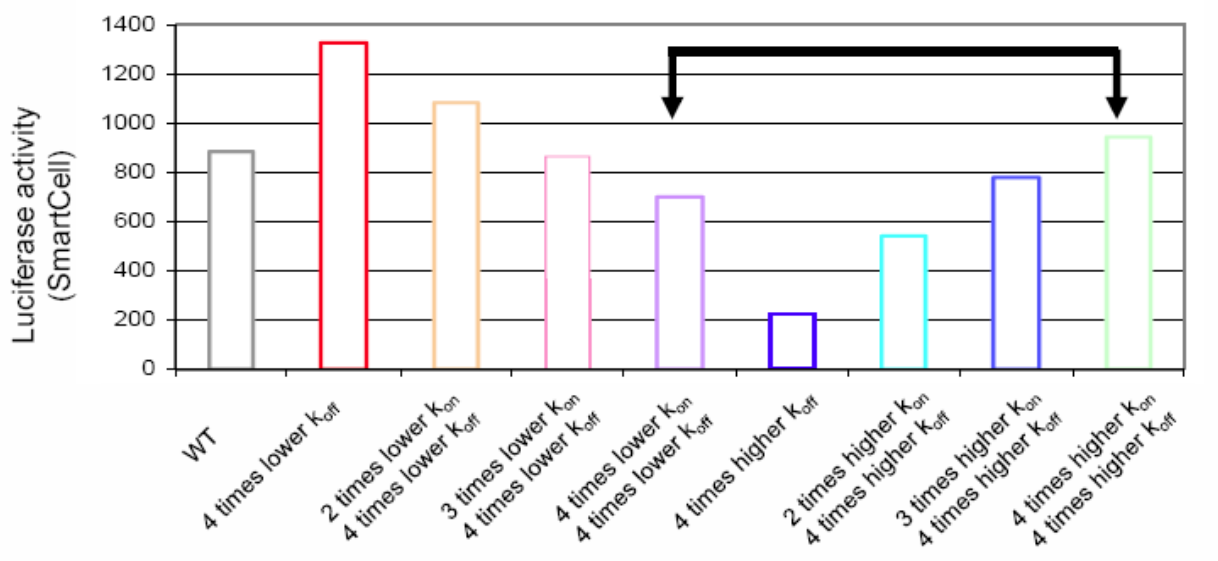




Correlation between predicted changes in k_{on} is very high, while correlation with affinity (ΔG) is poorer

Kiel & Serrano, Sci Signal, 2009

Results from the network model for designed mutants

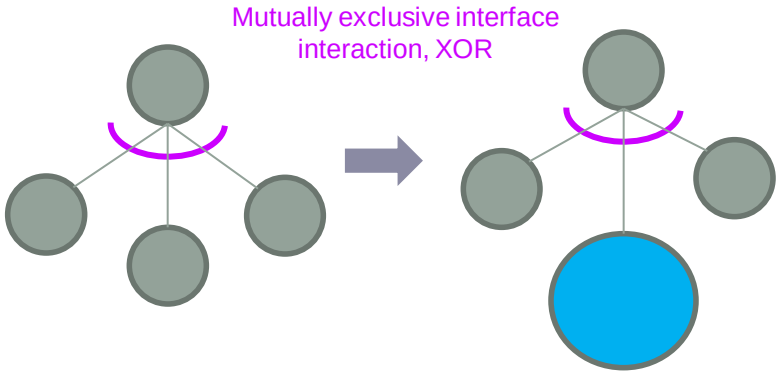


Confirms experimental findings:
Mutant with 4 time lower k_{on} and 4 times lower k_{off} (same K_D) has less predicted luciferase activity (and opposite for mutant with 4 times higher k_{on}/k_{off})

- Experiments and simulations suggest that association rate constants of Ras-Raf complex formation are important for signaling

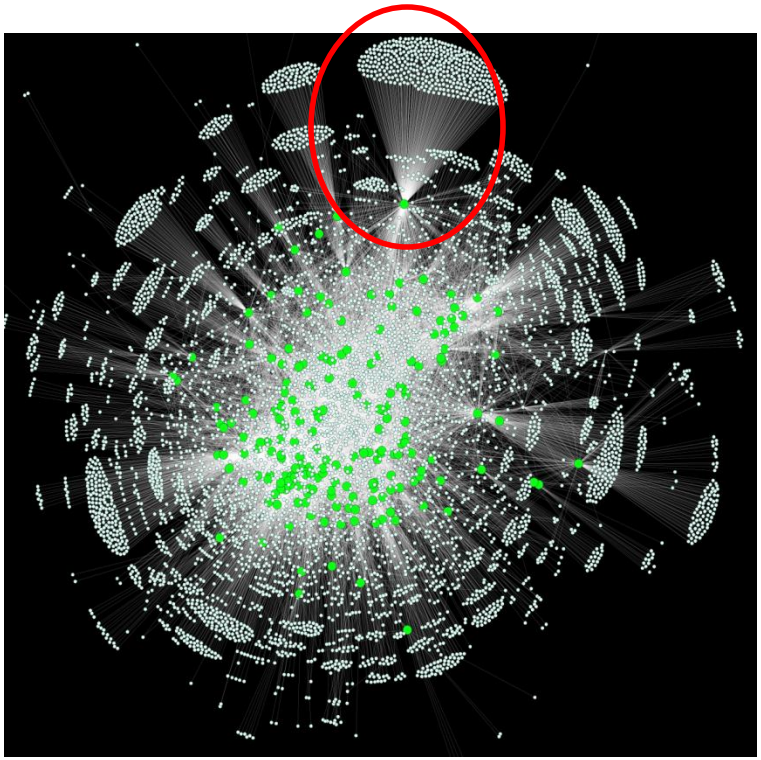
Kiel & Serrano, Sci Signal, 2009

The effect of protein abundance perturbations and interaction competition in PPI networks

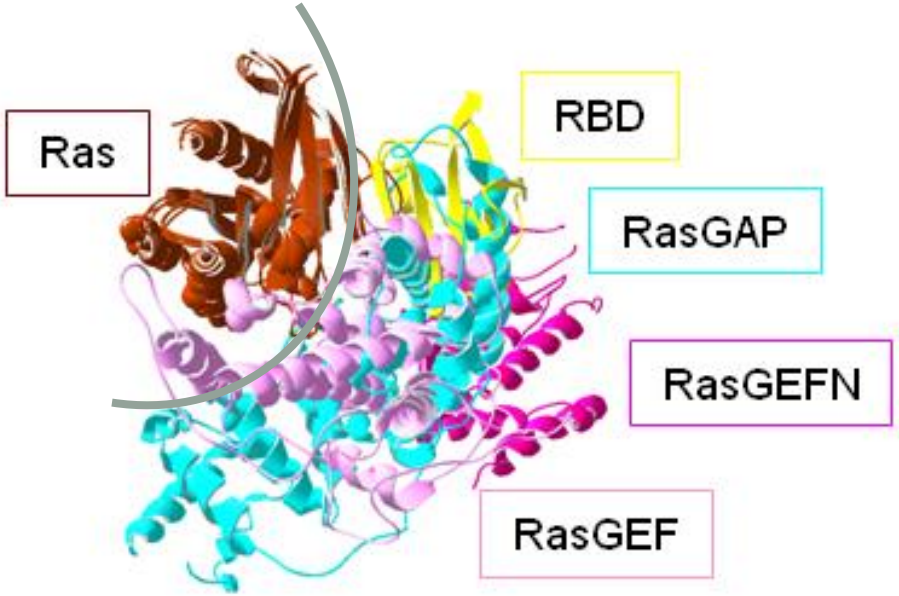


How could interaction competition and protein concentration affect downstream signaling?

Signaling complexes: > 300 partners for one protein??

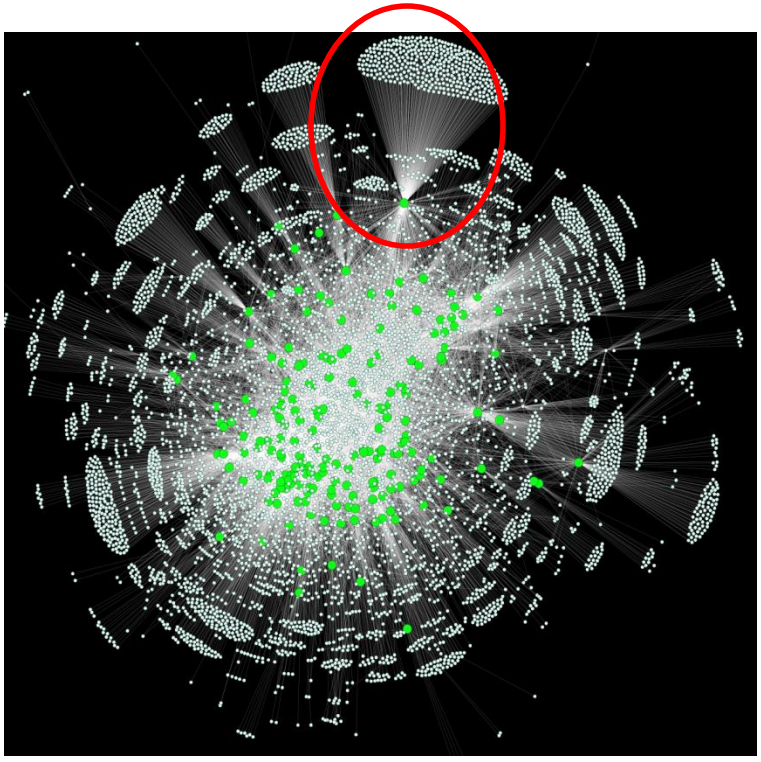


Some proteins will use similar binding surfaces for interaction with other molecules: 'mutually exclusive interactions'/ 'XOR'

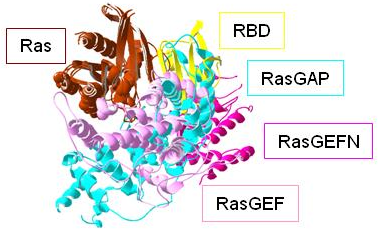


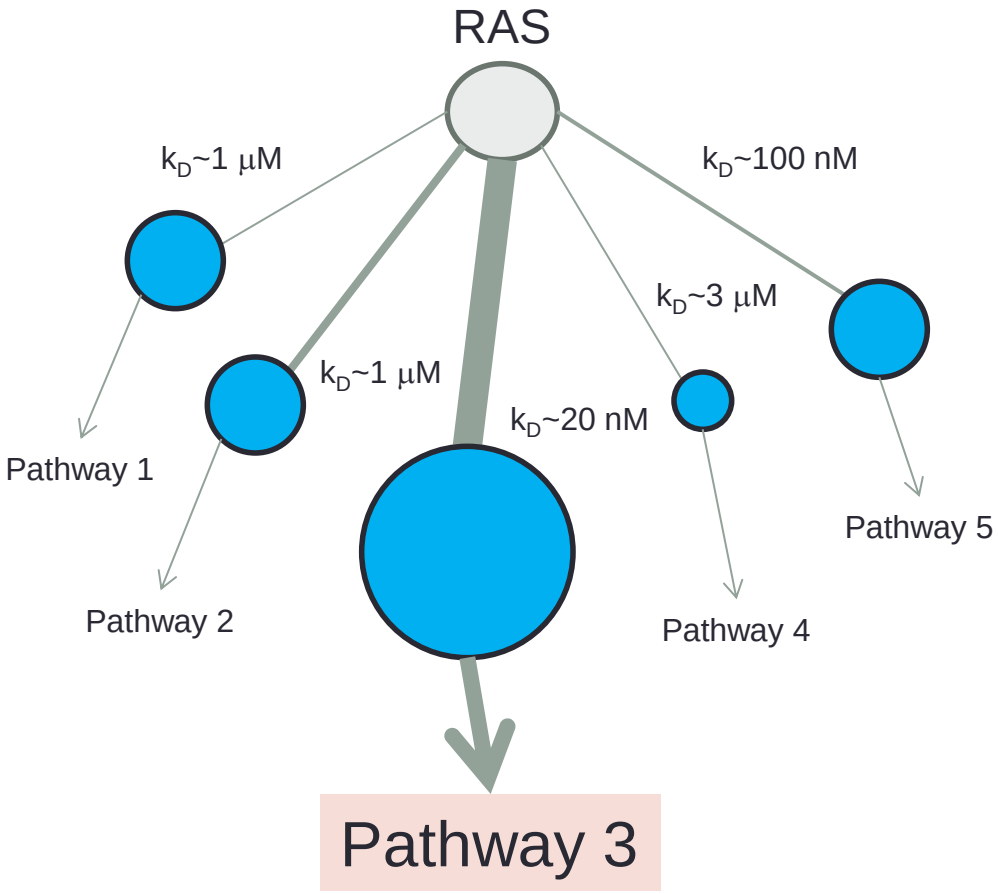
How could interaction competition and protein concentration affect downstream signaling?

Signaling complexes: > 300 partners for one protein??



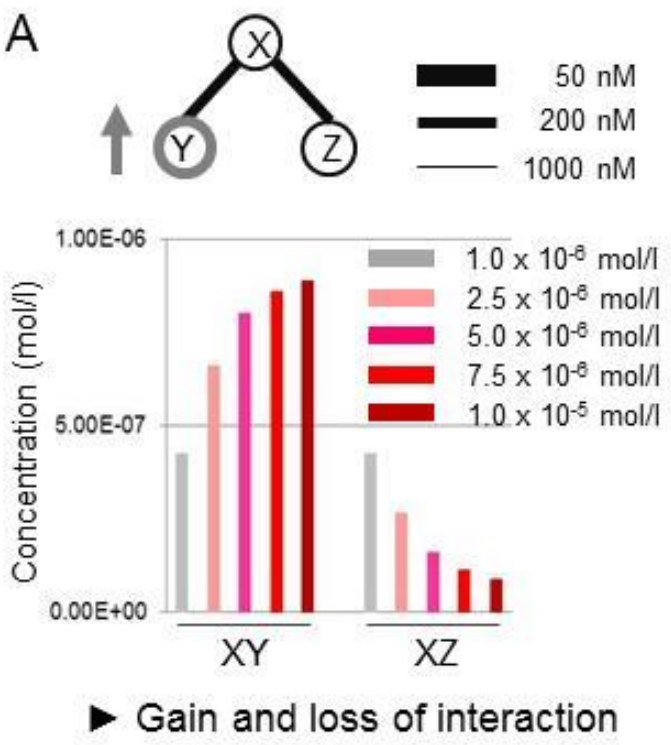
In a simple world: concentration and k_D will determine the signaling output





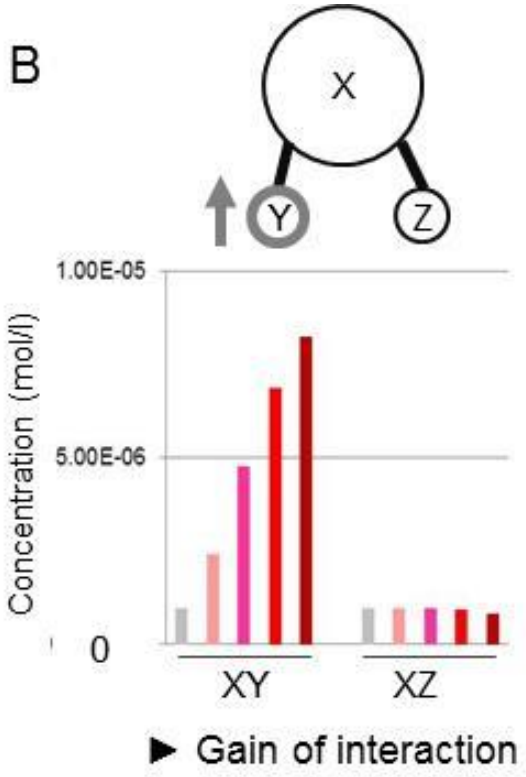
The effect of abundance variation at XOR network motifs

Competition



► Gain and loss of interaction

No competition



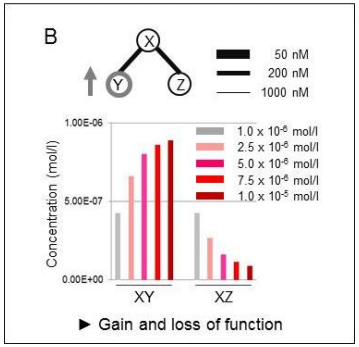
► Gain of interaction

➢ The output/ function depends on both, network structure and abundance

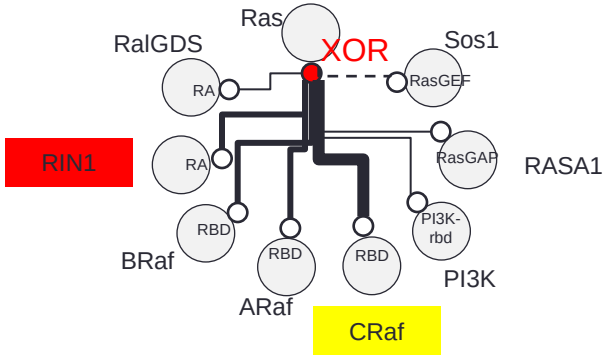
Kiel et al, Sci Signal, 2013

Competition at the Ras XOR node

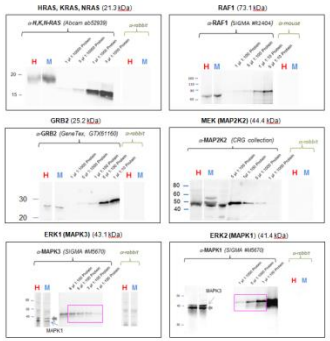
Network motif



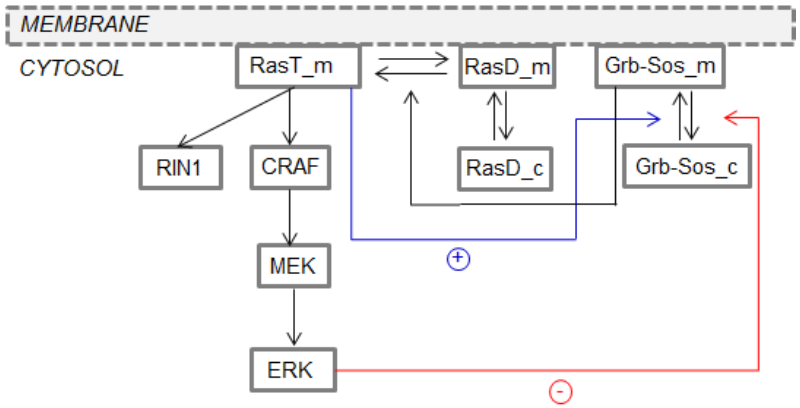
The Ras XOR node



Experimental abundances

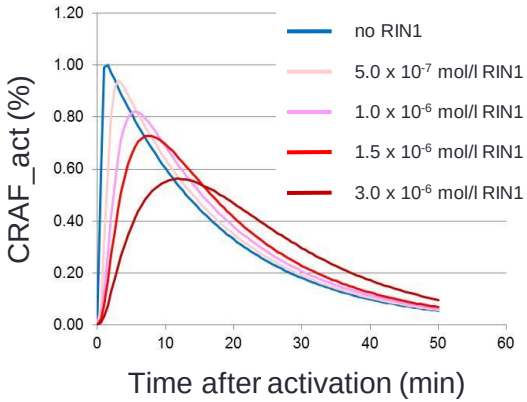


A



B

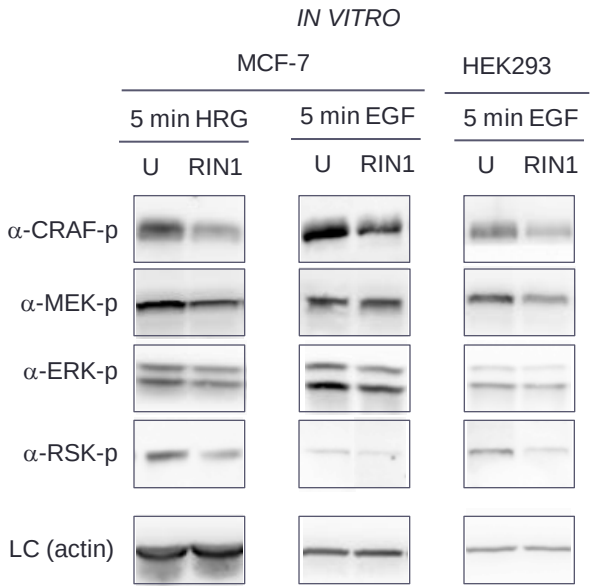
IN SILICO

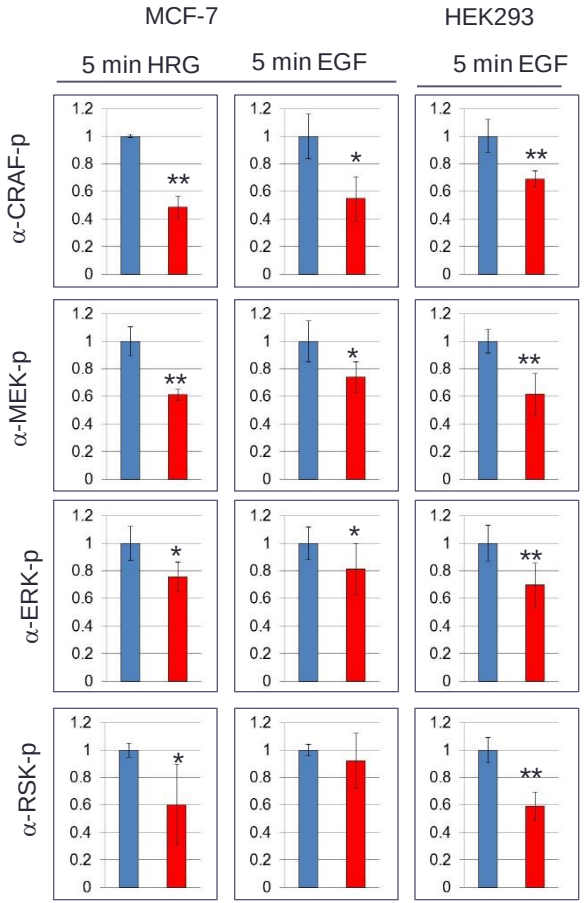


Mathematical network modeling: increasing RIN1 to 10-fold higher of CRAF expression should decrease CRAF activation

Kiel et al, Sci Signal, 2013

Experimental testing of competition at the Ras node



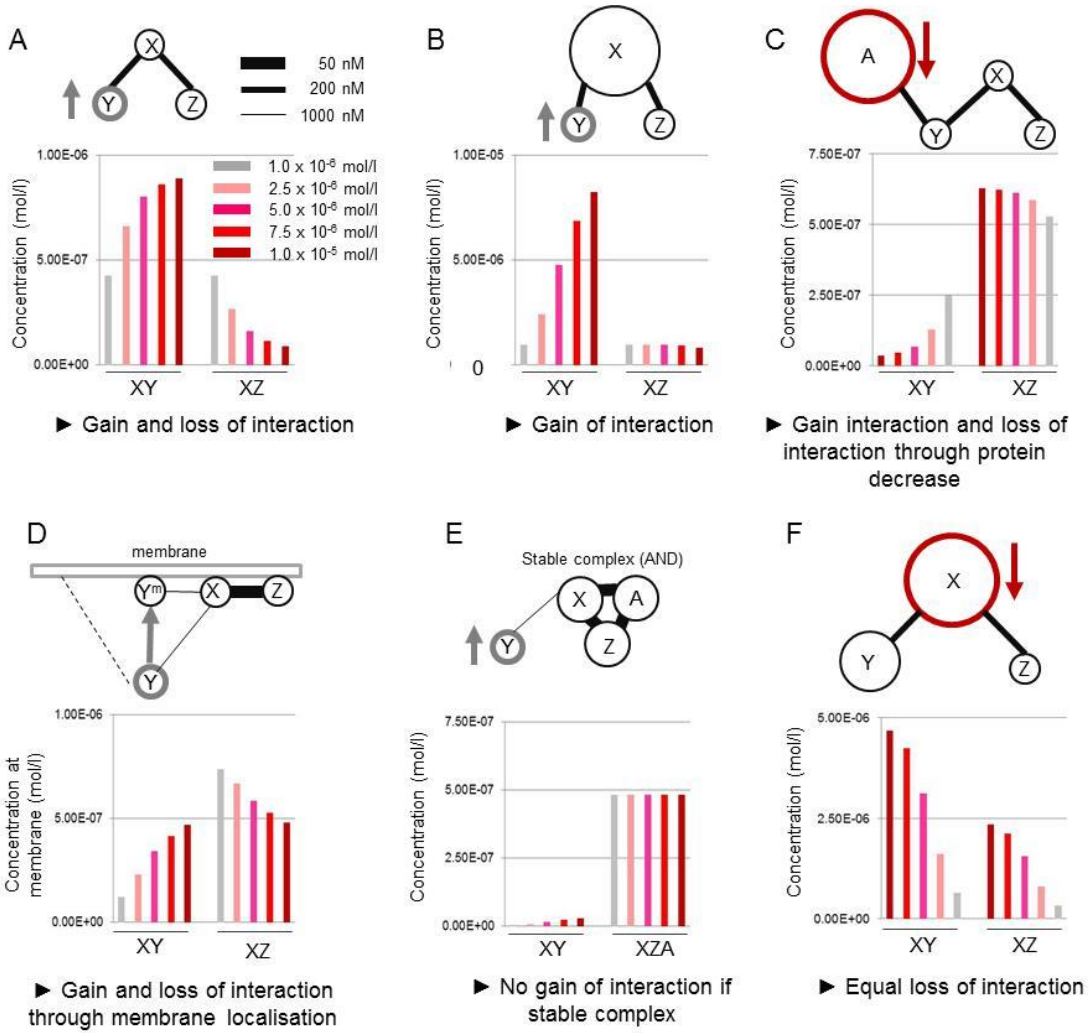


Expression of RIN1 in MCF-7 and HEK293 cells decreases CRAF, MEK, and ERK activation

- Alterations in the abundance of one of two hub-binding partners affected downstream signaling

Kiel et al, Sci Signal, 2013

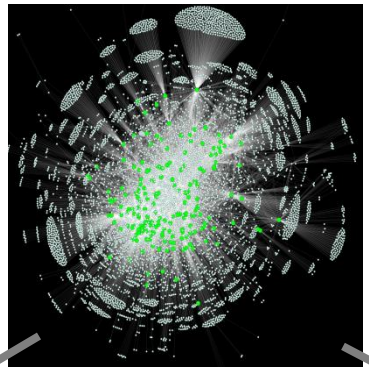
The effect of abundance variation at XOR network motifs

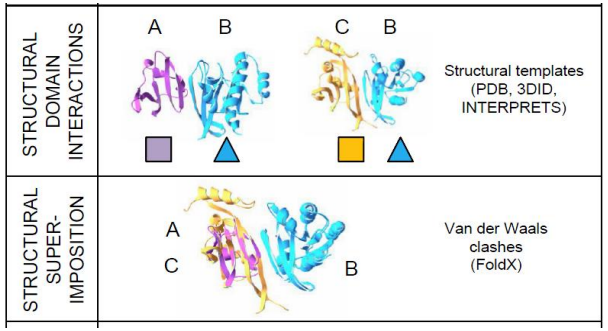


- The output/ function depends on both, network structure and abundance: we need to know the network very well to understand

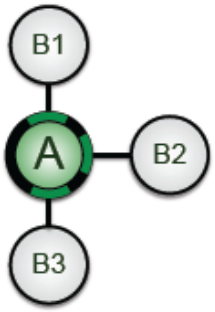
Kiel et al, Sci Signal, 2013

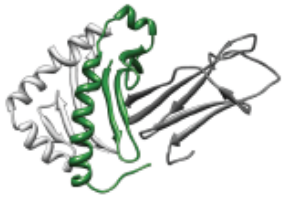
A bioinformatics tool to distinguish mutually exclusive from compatible interactions in large-scale PPI





Compatible ('AND')

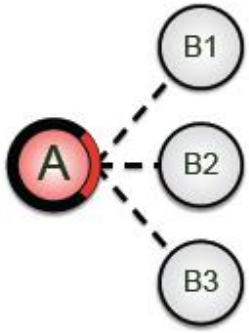


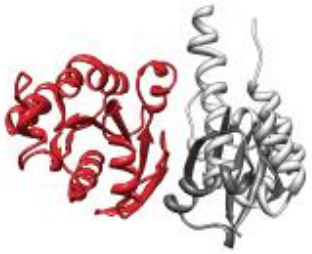


SAPIN (structural analysis of protein interaction networks) webserver

http://sapin.crg.es/

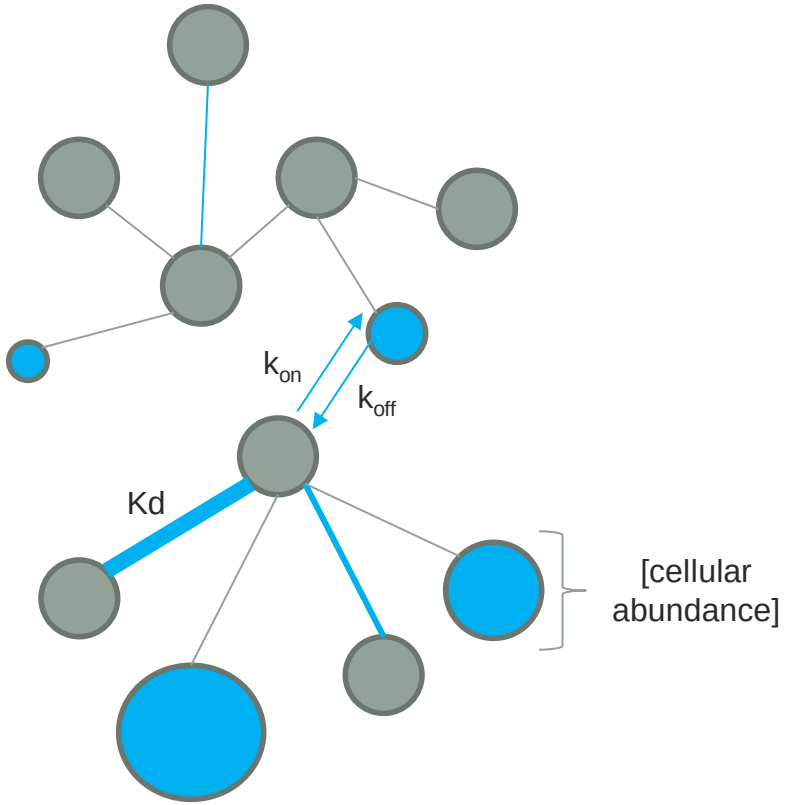
Exclusive ('XOR')





Yang et al, Bioinformatics, 2012

Experimental methods to quantify protein abundances, affinities, and kinetic constants



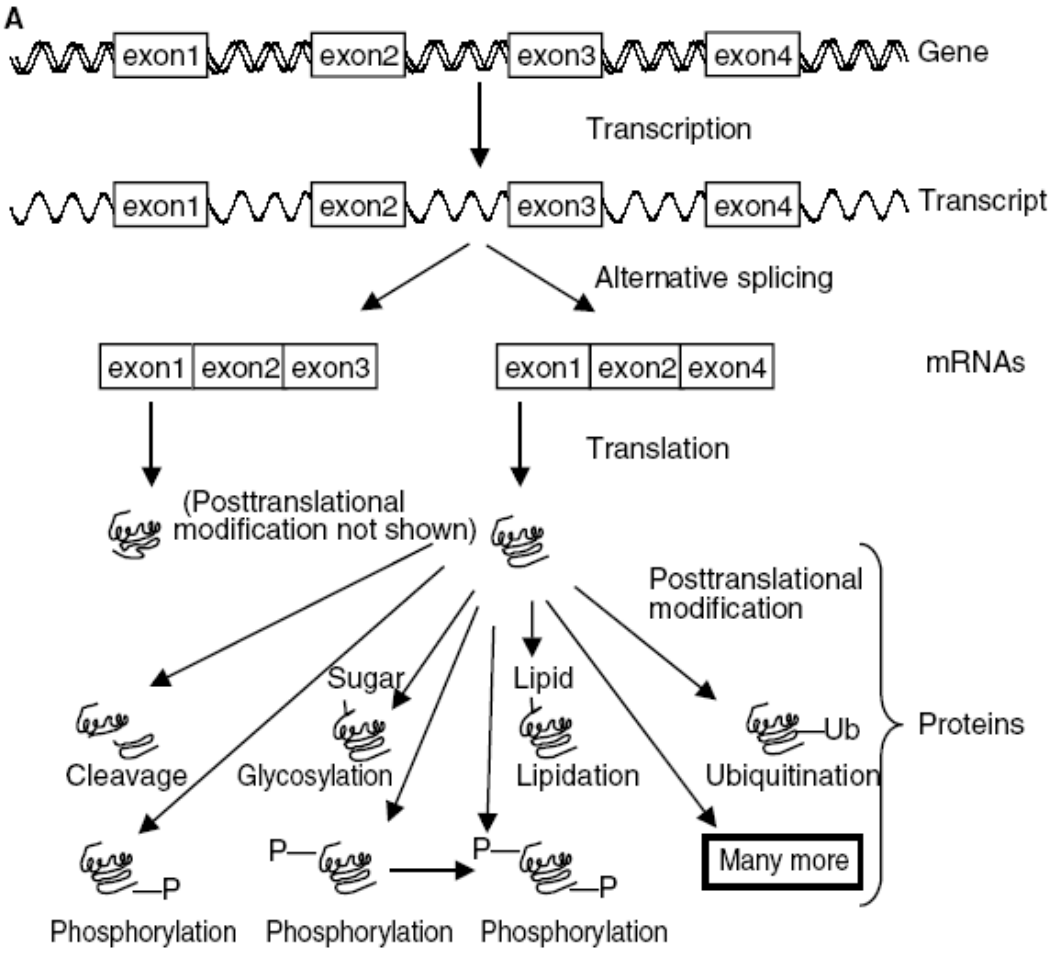
Why proteomics in times of deep RNA sequencing?

- mRNA does not translate1:1 into protein; keywords:
  (i) translation efficiency,
  (ii) mRNA stability,
  (iii) protein stability,
- Posttranslational modification (PTMs) of proteins, e.g. phosphorylation

Two main aims: IDENTIFICATION and QUANTIFICATION

Two main techniques: MASS SPECTROMETRY and ANTIBODY-BASED

High complexity of the proteome



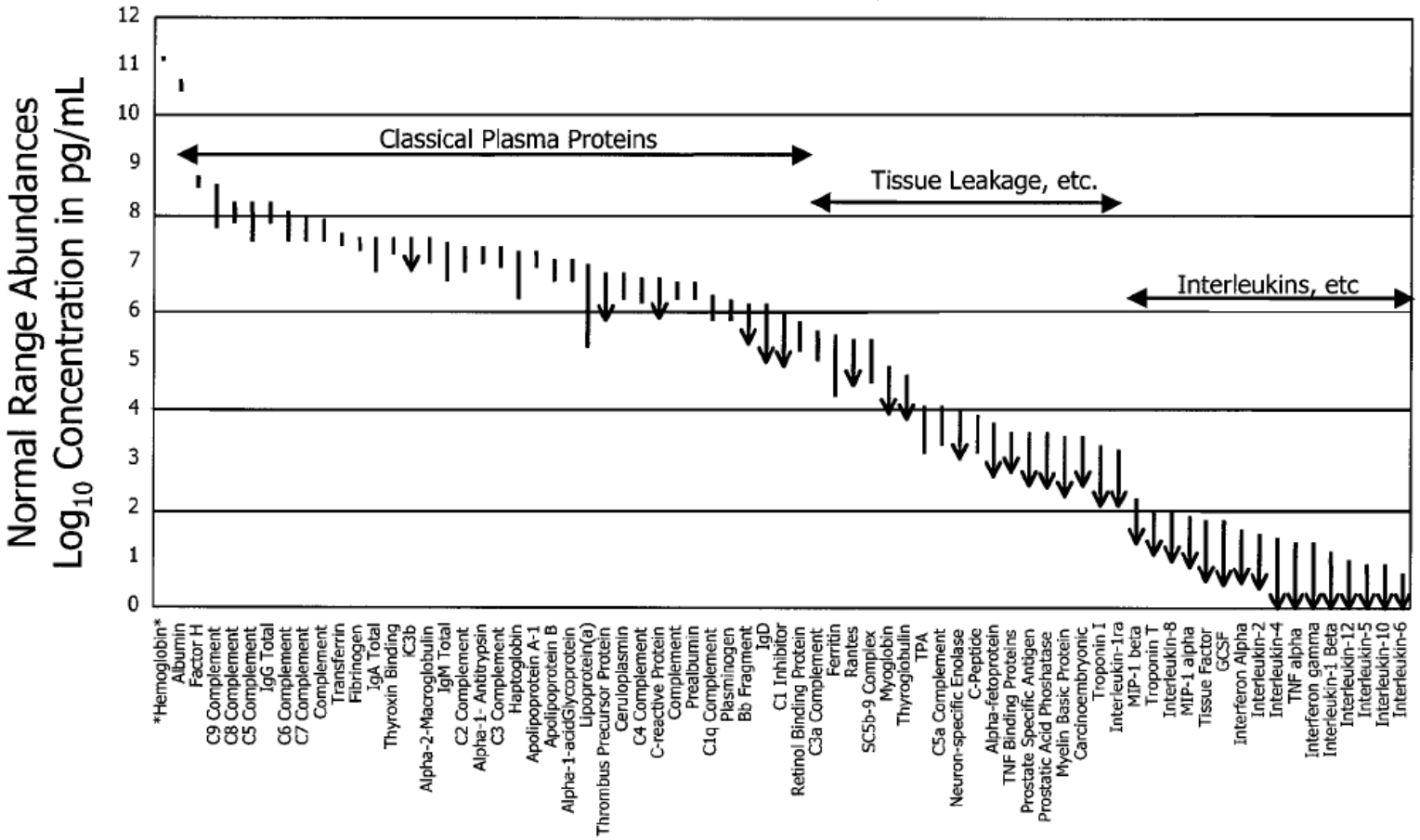
30,000 coding genes per cell

Alt.splicing: 2-3 x 30,000
= 90,000 proteins

Post-translational modifications
> 10 x 90,000
= 900,000 proteins

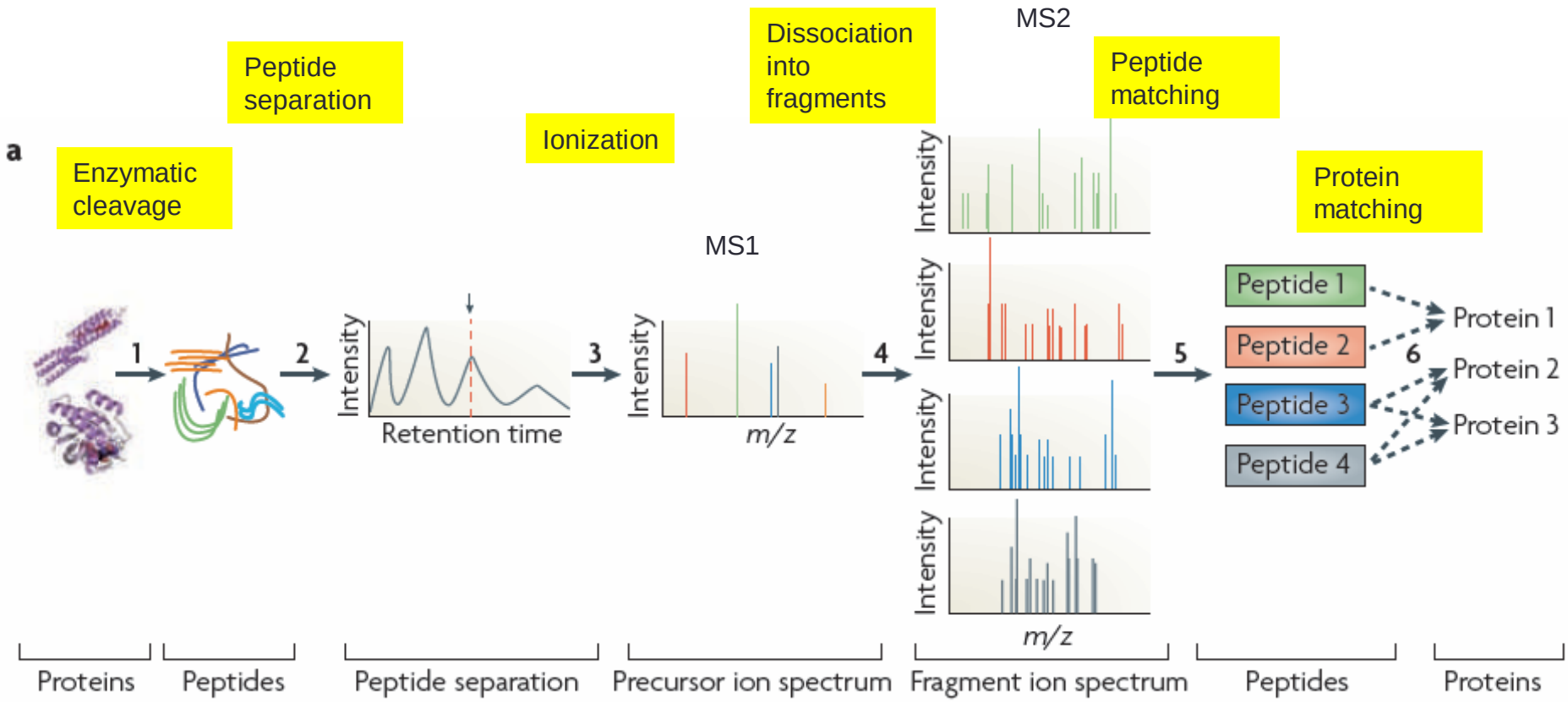
Peng and Gygi, JMS, 2001

High dynamic range of the proteome



Anderson and Anderson, MCP, 2002

Protein identification by mass spectrometry

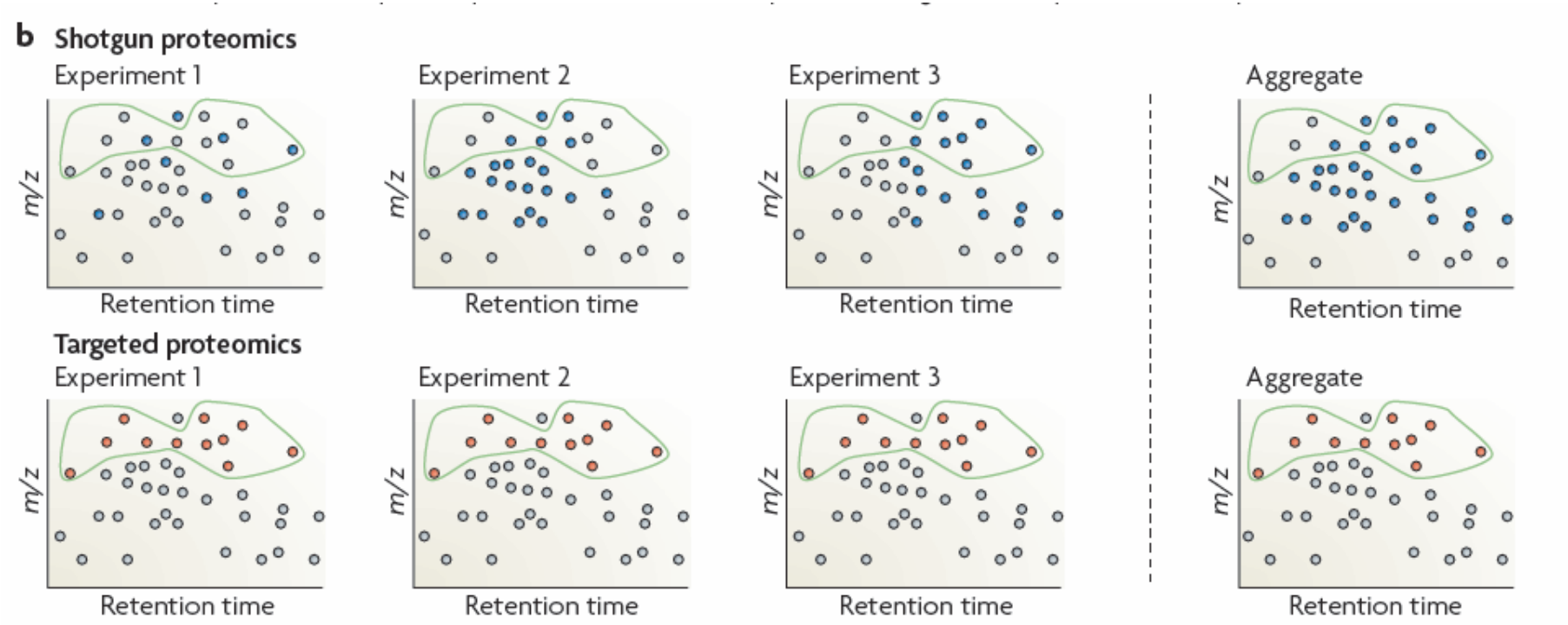


- ❑ Address problem of cellular complexity by fractionation, e.g. liquid chtromatography
- ❑ Address problem of cellular dynamic range by better and better (and better…) mass spectrometers…

Ahrens et al, 2010

‘Shotgun’ compared to ‘targeted’ approach

Targeted proteomics is the method of choice for studying (a limited number of) signaling proteins



Ahrens et al, 2010

Human deep proteome mapping

Molecular Systems Biology 7; Article number 549; doi:10.1038/msb.2011.82
Citation: *Molecular Systems Biology* 7: 549

www.molecularsystemsbiology.com

molecular systems biology

REPORT

The quantitative proteome of a human cell line

Martin Beck[1,9], Alexander Schmidt[2,9], Johan Malmstroem[3,4], Manfred Claassen[5], Alessandro Ori[1], Anna Szymborska[1], Franz Herzog[6], Oliver Rinner[4], Jan Ellenberg[1] and Ruedi Aebersold[6,7,8,*]

[1] European Molecular Biology Laboratory, Heidelberg, Germany, [2] Biozentrum, University of Basel, Basel, Switzerland, [3] Department of Immunotechnology, BMC, Lund, Sweden, [4] Biognosys AG, Schlieren, Switzerland, [5] Department of Computer Science, ETH Zurich, Zurich, Switzerland, [6] Department of Biology, Institute of Molecular Systems Biology, ETH Zurich, Zurich, Switzerland; [7] Competence Center for Systems Physiology and Metabolic Diseases, Zurich, Switzerland and [8] Department of Science, University of Zurich, Zurich, Switzerland
[9] These authors contributed equally to this work
* Corresponding author. Department of Biology, Institute of Molecular Systems Biology, ETH Zurich, Wolfgang-Pauli-Strasse 16, Zurich 8093, Switzerland. Tel.: +41 44 633 1071; Fax: +41 44 633 1051; E-mail: aebersold@imsb.biol.ethz.ch

Received 15.8.11; accepted 29.9.11

- R. Aebersold lab

~10,000 proteins quantified

Beck et al, MSB, 2011

Molecular Systems Biology 7; Article number 548; doi:10.1038/msb.2011.81
Citation: *Molecular Systems Biology* 7: 548

www.molecularsystemsbiology.com

molecular systems biology

REPORT

Deep proteome and transcriptome mapping of a human cancer cell line

Nagarjuna Nagaraj[1], Jacek R Wisniewski[1], Tamar Geiger[1], Juergen Cox[1], Martin Kircher[2], Janet Kelso[2], Svante Pääbo[2] and Matthias Mann[1,*]

[1] Department for Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Martinsried, Germany and [2] Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany
* Corresponding author. Department for Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Am Klopferspitz 18, D-82152 Martinsried, Germany. Tel.: +49 89 8578 2557; Fax: +49 89 8578 2219; E-mail: mmann@biochem.mpg.de

Received 15.7.11; accepted 29.10.11

- M Mann lab

10,255 proteins quantified

Nagaraj et al, MSB, 2011

Human deep proteome mapping: where are we now? Complete?

2014 Pandey lab

ARTICLE

doi:10.1038/nature13302

A draft map of the human proteome

Min-Sik Kim[1,2], Sneha M. Pinto[3], Derese Getnet[1,4], Raja Sekhar Nirujogi[3], Srikanth S. Manda[3], Raghothama Chaerkady[1,2], Anil K. Madugundu[3], Dhanashree S. Kelkar[3], Ruth Isserlin[5], Shobhit Jain[5], Joji K. Thomas[3], Babylakshmi Muthusamy[3], Pamela Leal-Rojas[1,6], Praveen Kumar[3], Nandini A. Sahasrabuddhe[3], Lavanya Balakrishnan[3], Jayshree Advani[3], Bijesh George[3], Santosh Renuse[3], Lakshmi Dhevi N. Selvan[3], Arun H. Patil[3], Vishalakshi Nanjappa[3], Aneesha Radhakrishnan[3], Samarjeet Prasad[1], Tejaswini Subbannayya[3], Rajesh Raju[3], Manish Kumar[3], Sreelakshmi K. Sreenivasamurthy[3], Arivusudar Marimuthu[3], Gajanan J. Sathe[3], Sandip Chavan[3], Keshava K. Datta[3], Yashwanth Subbannayya[3], Apeksha Sahu[3], Soujanya D. Yelamanchi[3], Savita Jayaram[3], Pavithra Rajagopalan[3], Jyoti Sharma[3], Krishna R. Murthy[3], Nazia Syed[3], Renu Goel[3], Aafaque A. Khan[3], Sartaj Ahmad[3], Gourav Dey[3], Keshav Mudgal[7], Aditi Chatterjee[3], Tai-Chung Huang[1], Jun Zhong[1], Xinyan Wu[1,2], Patrick G. Shaw[1], Donald Freed[1], Muhammad S. Zahari[2], Kanchan K. Mukherjee[8], Subramanian Shankar[9], Anita Mahadevan[10,11], Henry Lam[12], Christopher J. Mitchell[1], Susarla Krishna Shankar[10,11], Parthasarathy Satishchandra[13], John T. Schroeder[14], Ravi Sirdeshmukh[3], Anirban Maitra[15,16], Steven D. Leach[1,17], Charles G. Drake[16,18], Marc K. Halushka[15], T. S. Keshava Prasad[3], Ralph H. Hruban[15,16], Candace L. Kerr[19]†, Gary D. Bader[5], Christine A. Iacobuzio-Donahue[15,16,17], Harsha Gowda[3] & Akhilesh Pandey[1,2,3,4,15,16,20]

2014 Kuster lab

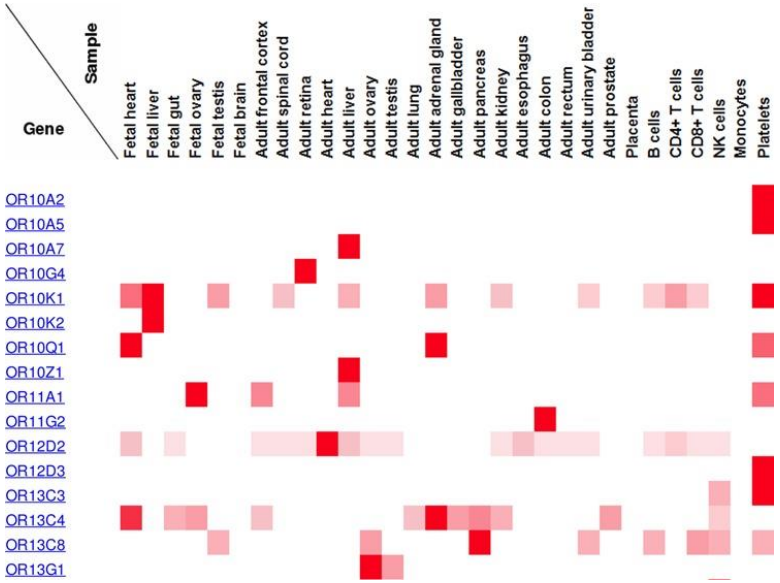
ARTICLE

doi:10.1038/nature13319

Mass-spectrometry-based draft of the human proteome

Mathias Wilhelm[1,2]*, Judith Schlegl[2]*, Hannes Hahne[1]*, Amin Moghaddas Gholami[1]*, Marcus Lieberenz[2], Mikhail M. Savitski[3], Emanuel Ziegler[2], Lars Butzmann[2], Siegfried Gessulat[2], Harald Marx[1], Toby Mathieson[3], Simone Lemeer[1], Karsten Schnatbaum[4], Ulf Reimer[4], Holger Wenschuh[4], Martin Mollenhauer[5], Julia Slotta-Huspenina[5], Joos-Hendrik Boese[2], Marcus Bantscheff[3], Anja Gerstmair[2], Franz Faerber[2] & Bernhard Kuster[1,6]

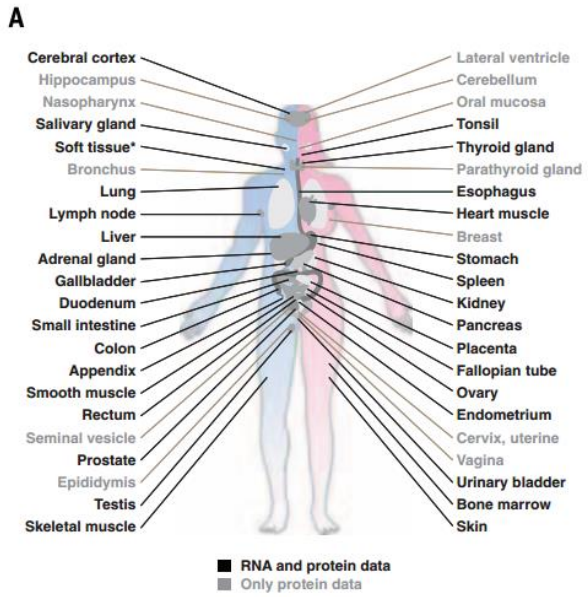
Many proteins are identified with peptides belonging to more than one protein (e.g. isoforms)

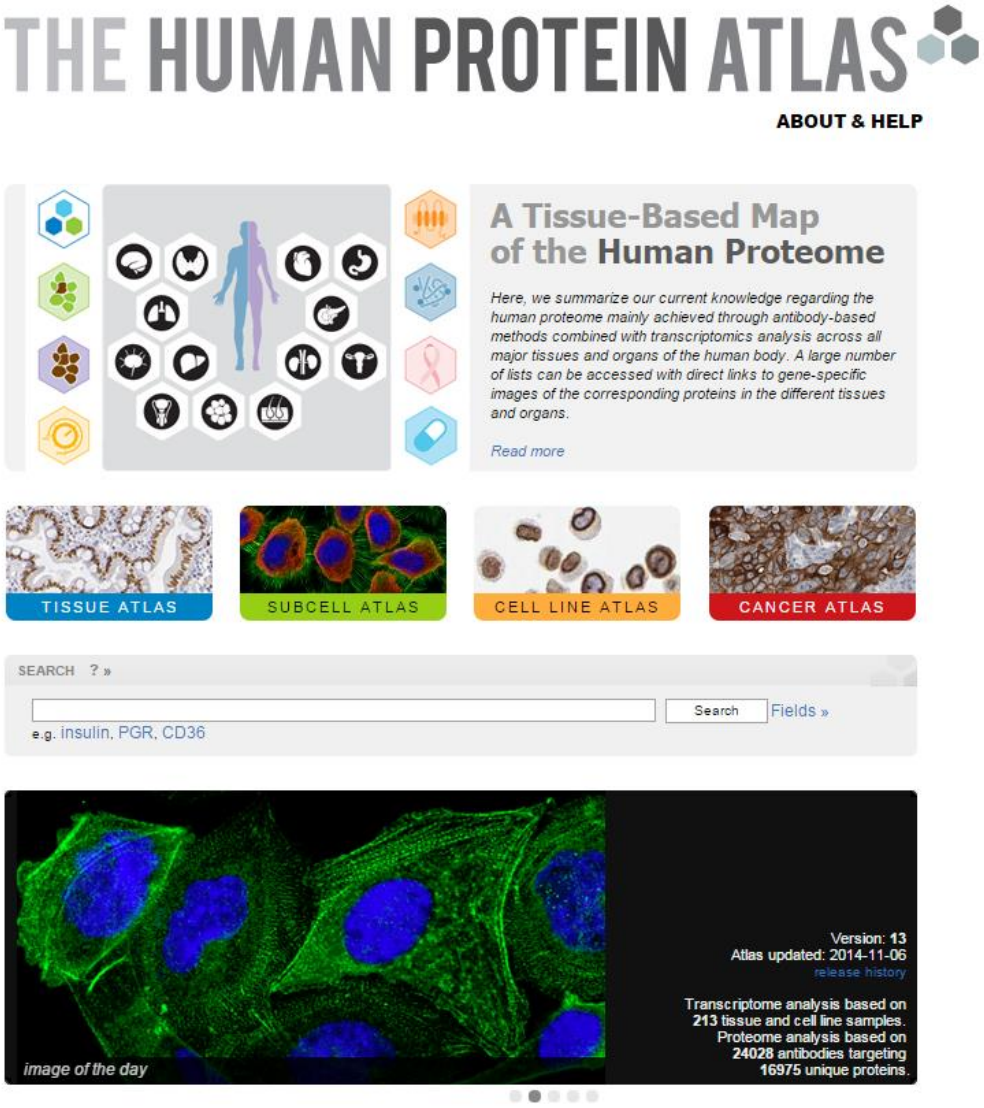


Ezkurdia et al, J Proteome Res, 2014

Antibody-based proteomics: only semi-quantitative abundances

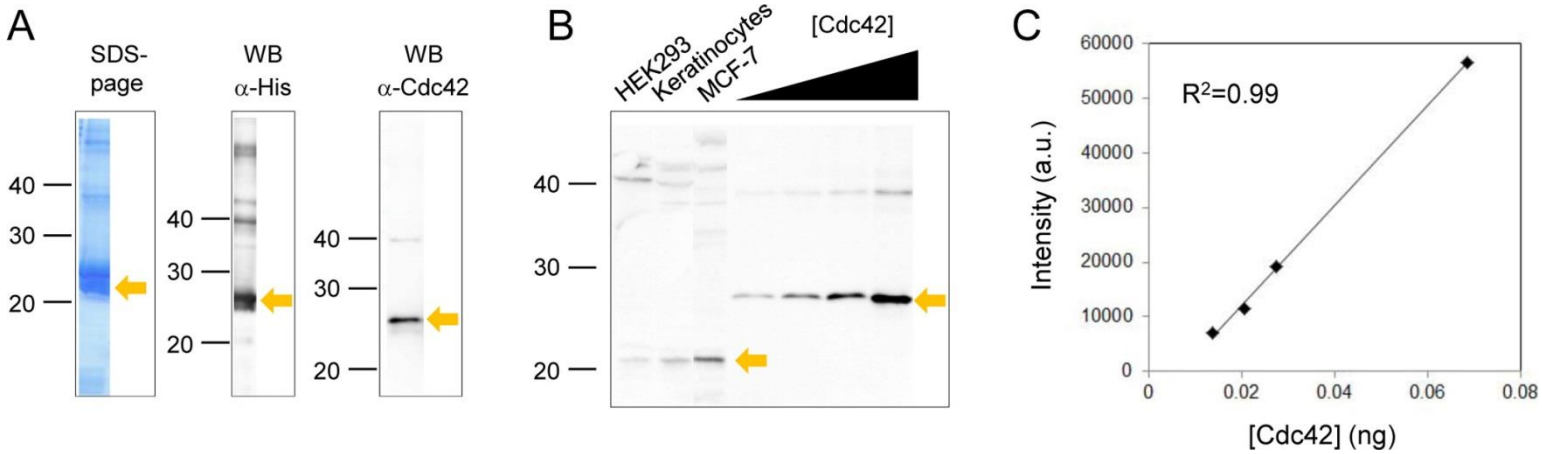
- Tissue-based map of the human proteome
- 44 major tissues and organs in the human body
- 24,028 antibodies corresponding to 16,975 protein-encoding genes



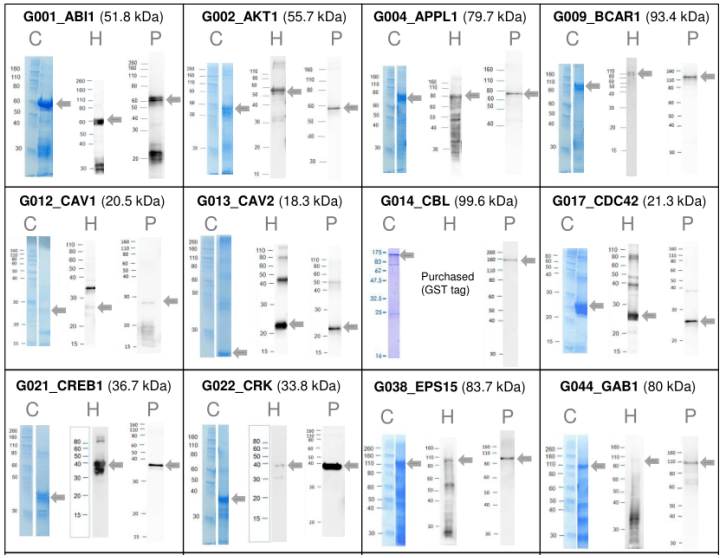


Uhlen et al, Science, 2015

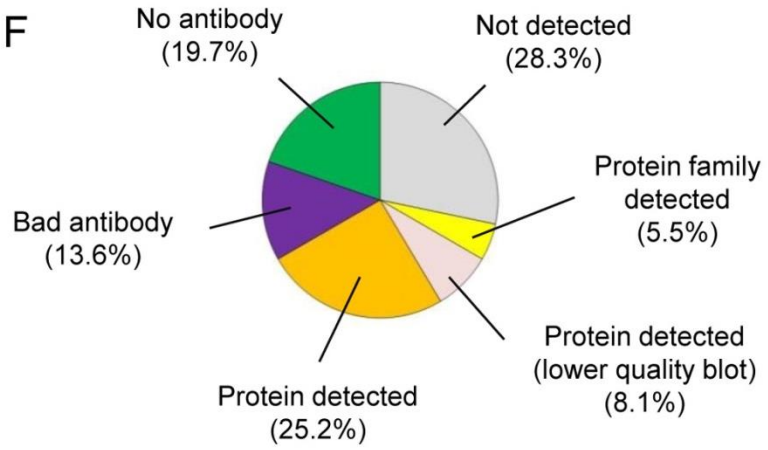
Quantitative Western blotting



Protein standards: expression, purification and quantification

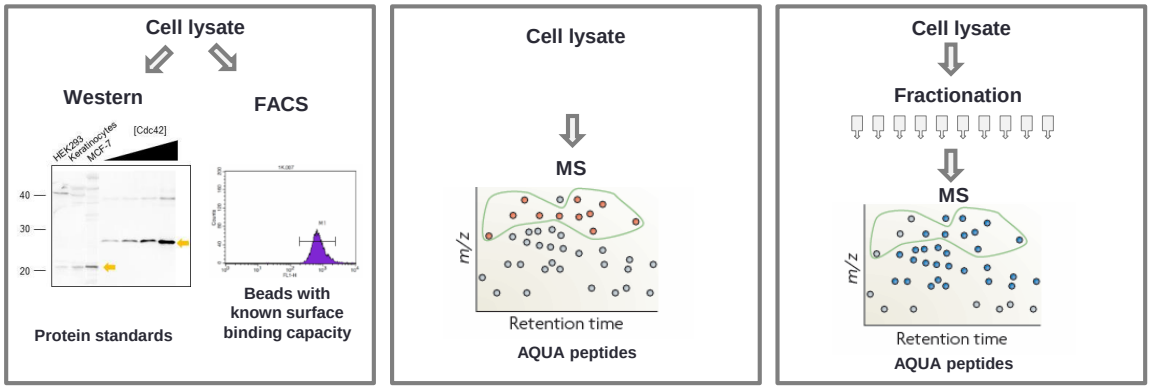


Summary statistic for quantitative Western blotting of 198 ErbB-related proteins



Kiel et al, J Prot Res, 2014

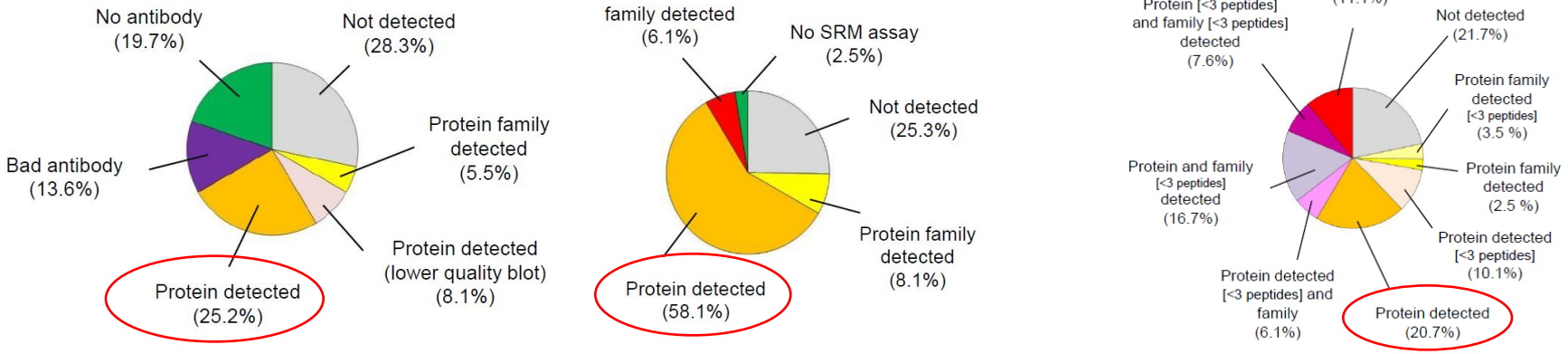
Combining different quantitative approaches to quantify 198 proteins in the ErbB signaling pathway



Quantitative Western blotting and quantitative FACS

Targeted mass spectrometry (MS)

Fractionation + shot-gun mass spectrometry (MS)



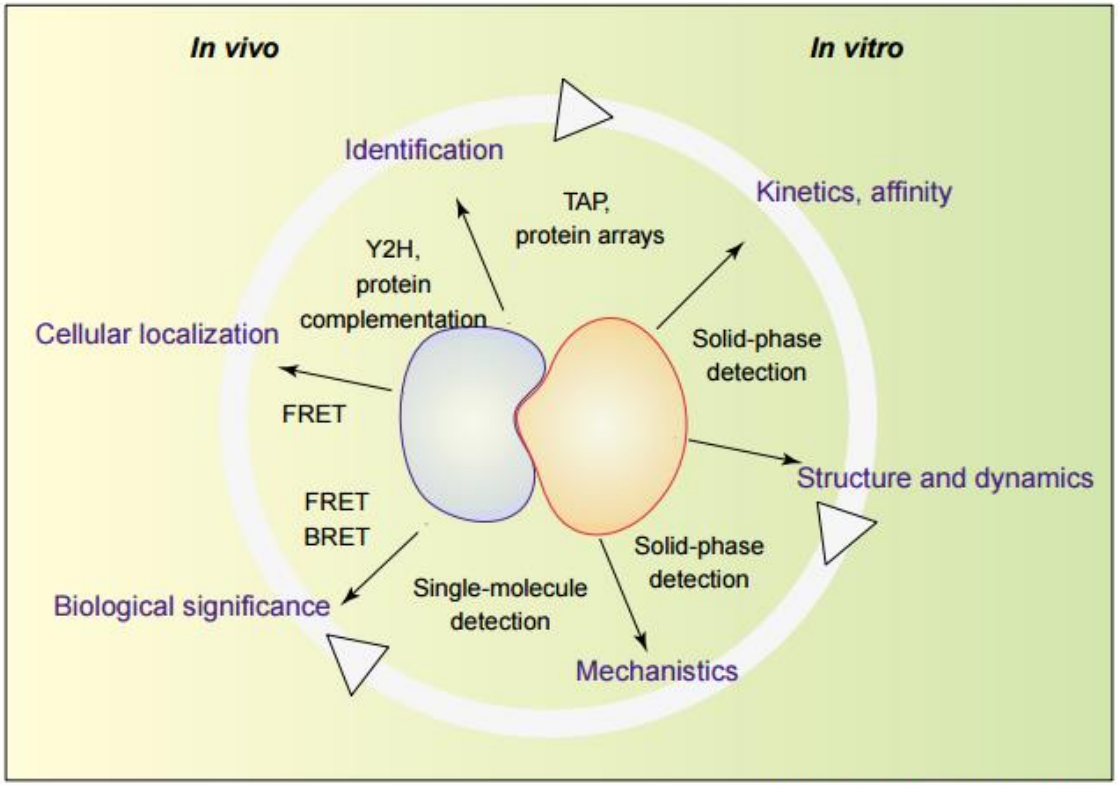
- SRM has a higher sensitivity compared to quantitative western blotting (but some proteins are only detected by Western blotting)
- Problem with isoforms and protein families: as a consequence of frequent gene duplication events in mammals, often similar proteins (e.g. AKT1 and AKT2) cannot be distinguished using the peptides detected by MS. > they can only be assigned to a protein group/ family

Kiel et al, J Prot Res, 2014

Measuring protein interactions *in vivo* and *in vitro*

The challenge:

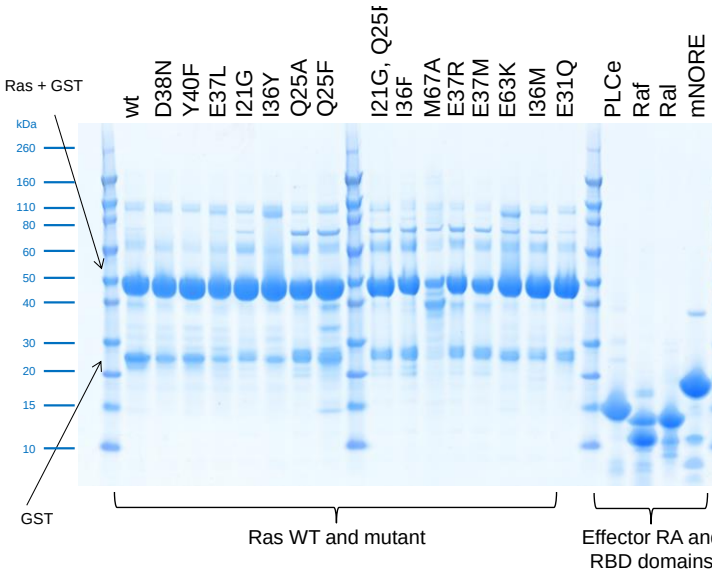
- most *in vivo* techniques are high-throughput, but do not provide affinities (only qualitative binding detection)
- *in vitro* techniques can provide affinities and kinetic constants, but are not high-throughput methods

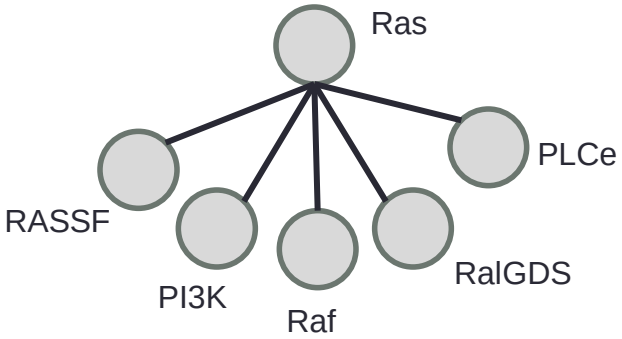


Piehler, Curr Opin Struct Biol, 2005

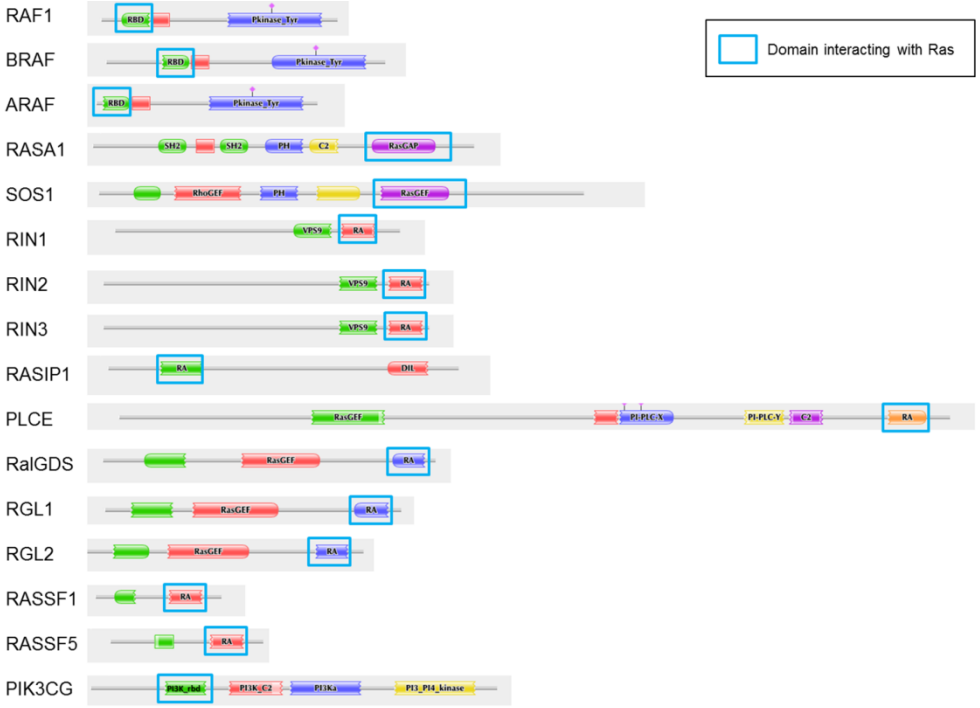
Measuring protein affinities in vitro requires the expression and purification of proteins (e.g. using bacteria)

Example: Bacterial expressed and purified Ras protein mutants and interactors



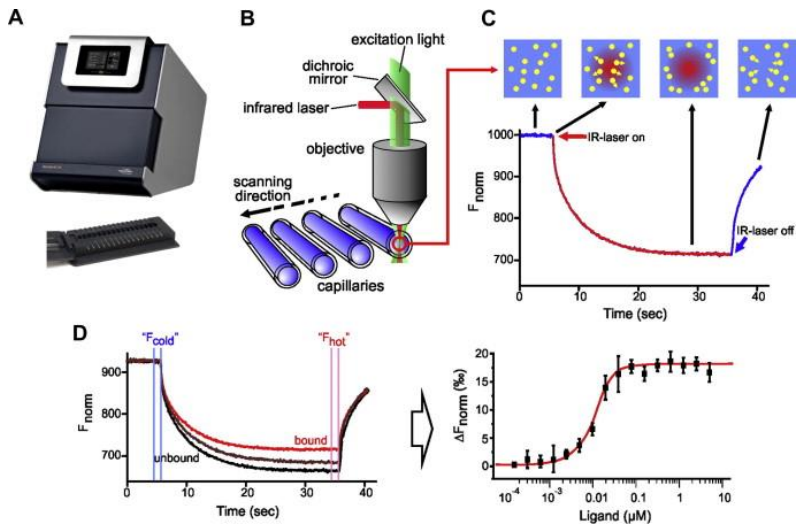


Large proteins are often not soluble: expression and purification of protein domains

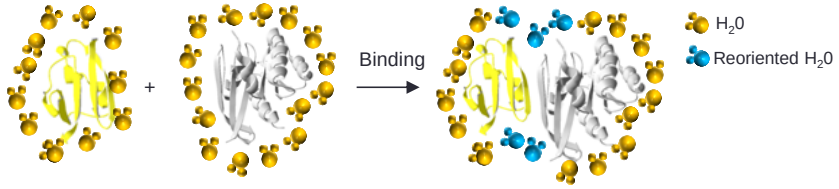


Two main methods to measure affinities and kinetic constants

Microscale thermophoresis



Jerabek-Willemsen et al, J Mol Struct, 2014



Amine-covalent labelled RBDs (fluorophore)

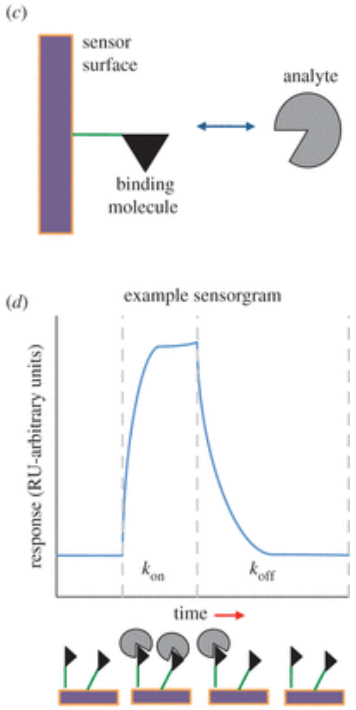
Ras WT and Mut (serial dilutions)

Fluorescence signal (depends on charge, size and hydration shell

- Provides only the affinity in equilibrium (K_d value), but not kinetic constants

$$K_d = \frac{[A] \times [B]}{[AB]}$$

Surface plasmon resonance

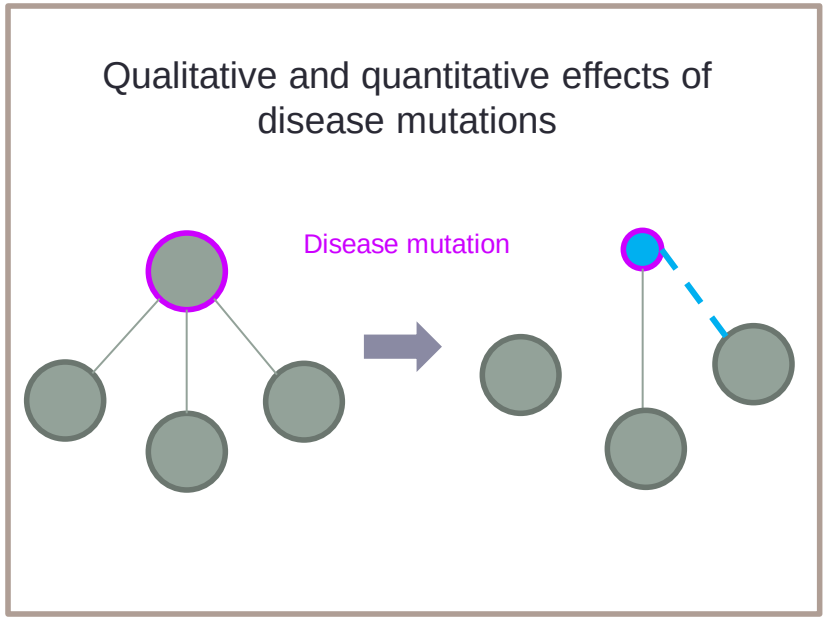


Kastritis et al, 2012

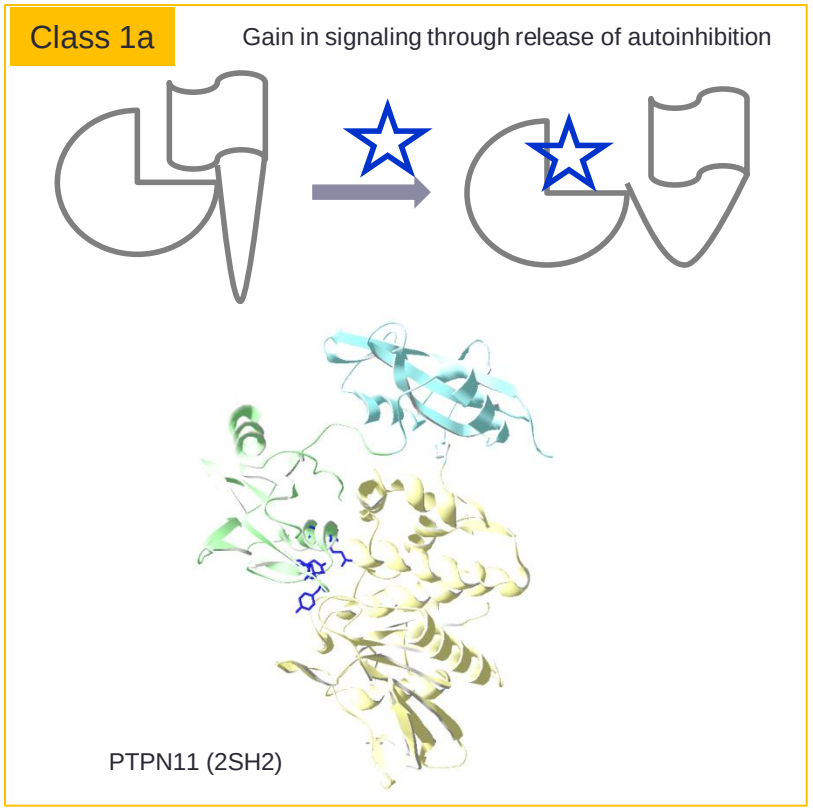
Optical method to measure the refractive index near a sensor surface

- Provides kinetic constants (k_{on} and k_{off})

$$K_d = \frac{k_{off}}{k_{on}}$$

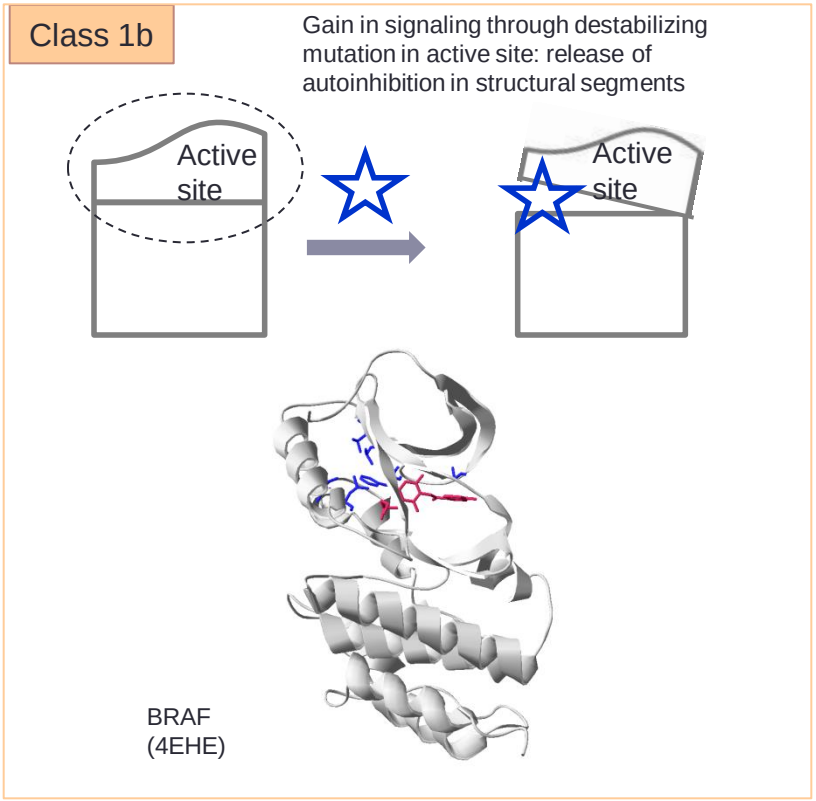

Qualitative and quantitative effects of disease mutations
Disease mutation

Examples how missense mutations can affect the network: a 3D structural perspective



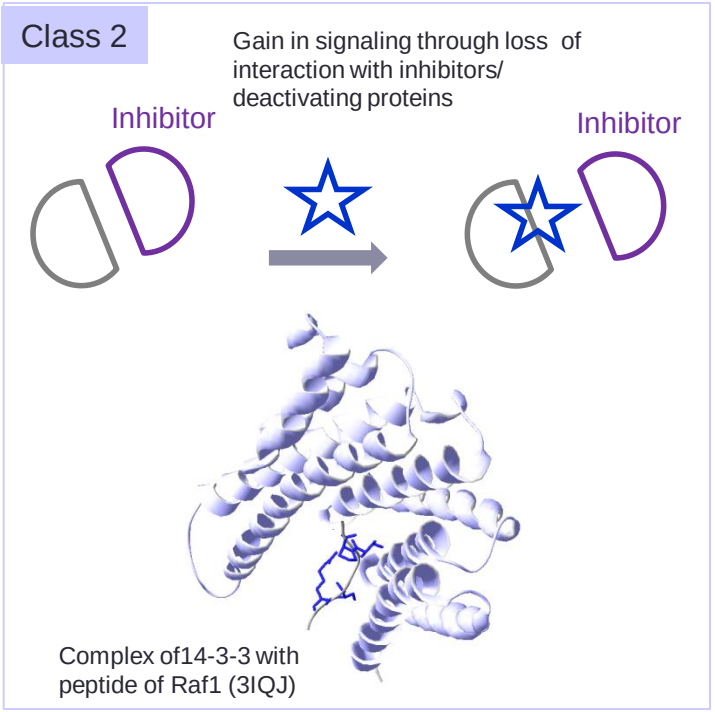
Kiel & Serrano, Mol Sys Biol, 2014

Examples how missense mutations can affect the network: a 3D structural perspective



Kiel & Serrano, Mol Sys Biol, 2014

Examples how missense mutations can affect the network: a 3D structural perspective

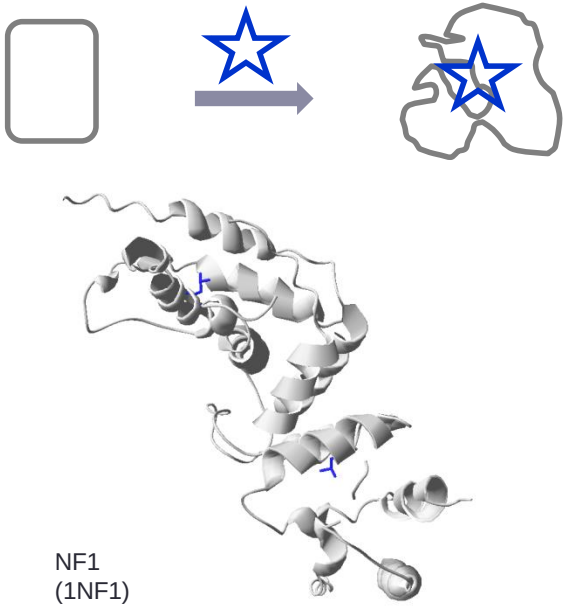


Kiel & Serrano, Mol Sys Biol, 2014

Examples how missense mutations can affect the network: a 3D structural perspective

Class 3

Folding affected (destabilization of protein) ; gain in signaling for NF1 and RASA1



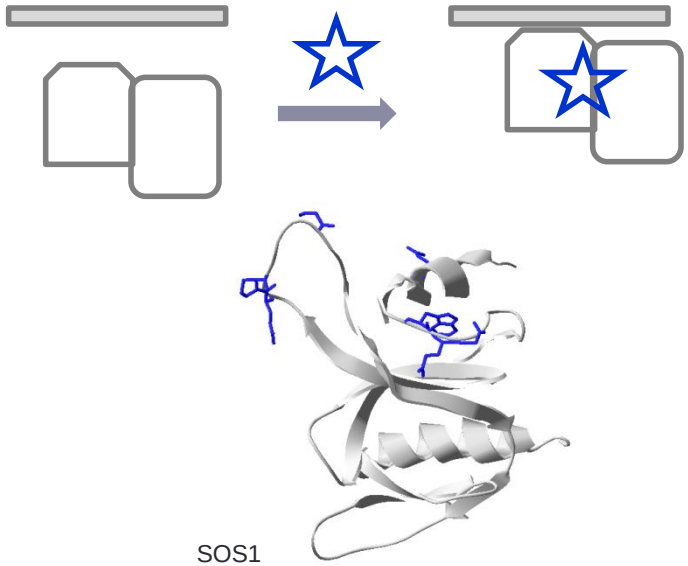
NF1
(1NF1)

Kiel & Serrano, Mol Sys Biol, 2014

Examples how missense mutations can affect the network: a 3D structural perspective

Class 4

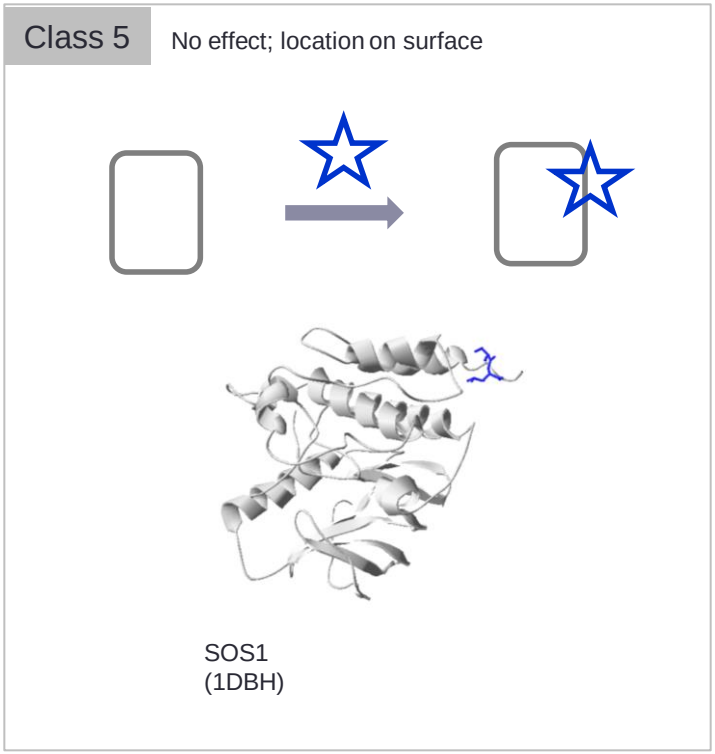
Gain in signaling through mutation of domains involved in membrane recruitment



SOS1
(1DBH)

Kiel & Serrano, Mol Sys Biol, 2014

Examples how missense mutations can affect the network: a 3D structural perspective

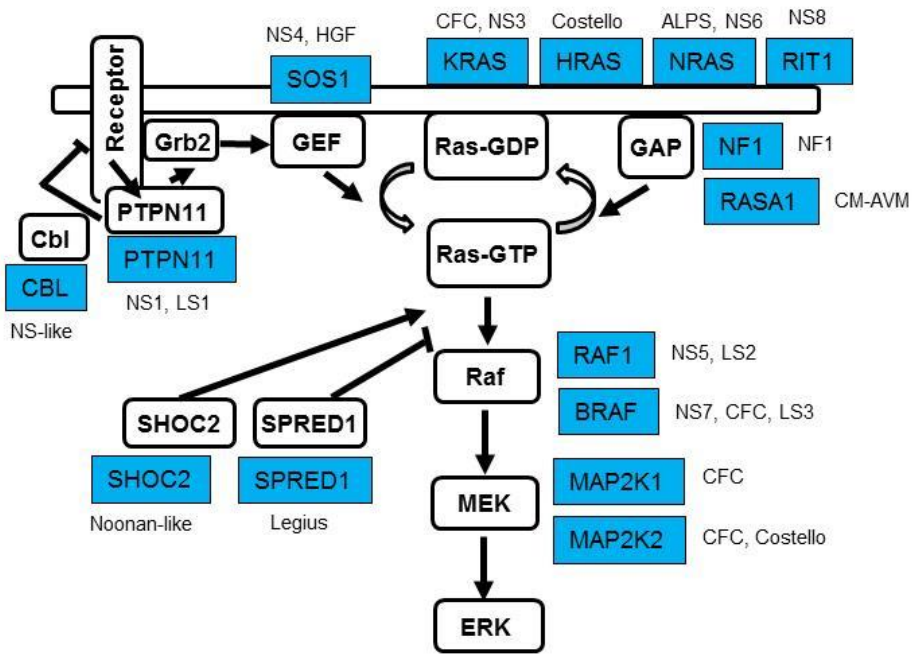


Kiel & Serrano, Mol Sys Biol, 2014

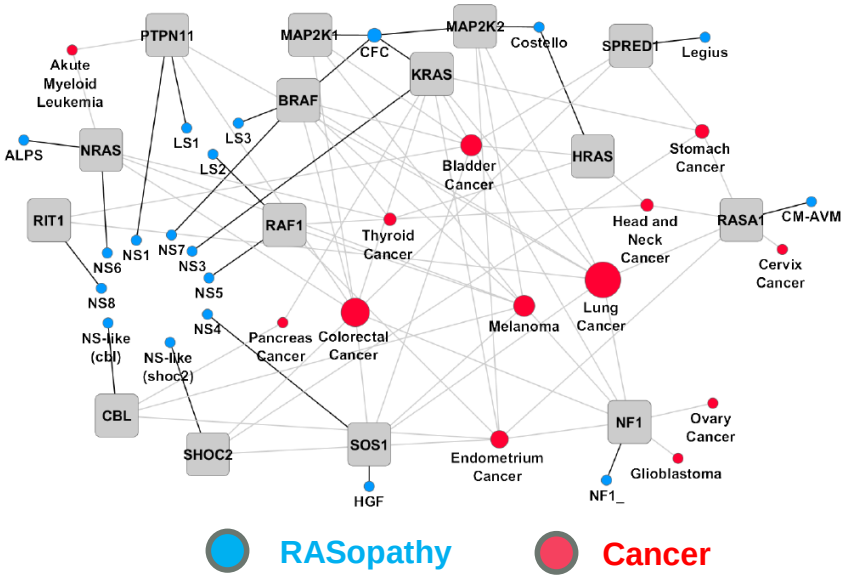
Example 1: RASopathy and cancer disease mutations

What are the differences in mutations of the same protein causing different disease (e.g. RASopathies or cancer)?

- Ras/MAPK syndromes ('RASopathies') are a class of developmental disorders caused by germline mutations



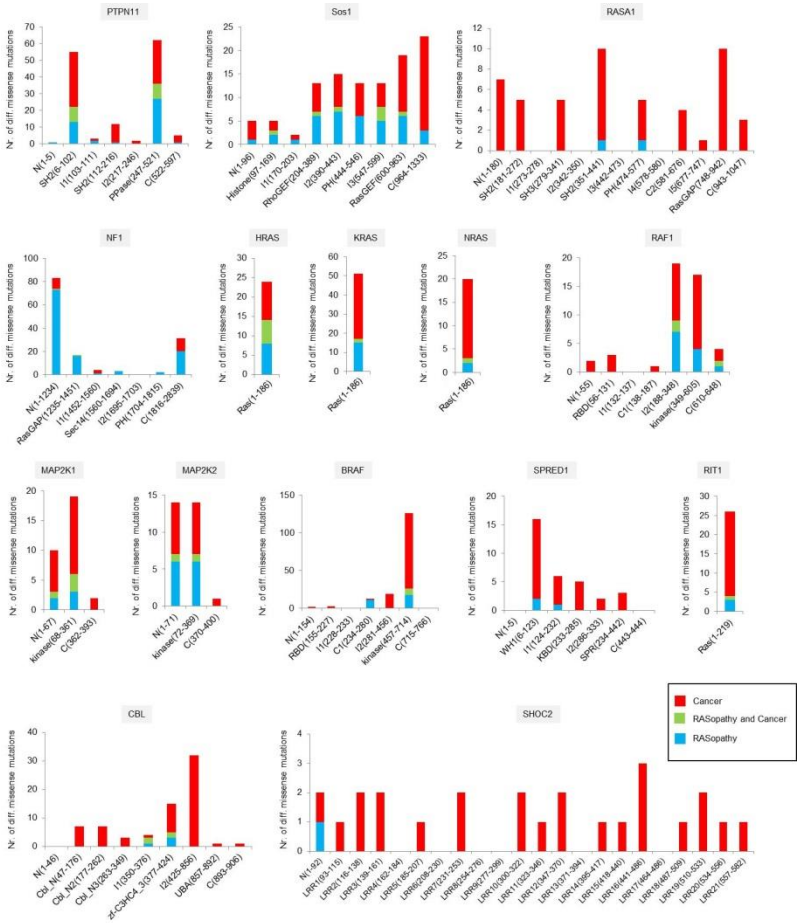
- Proteins in Ras/MAPK syndromes ('RASopathies') are also found in cancer



Kiel & Serrano, Mol Sys Biol, 2014

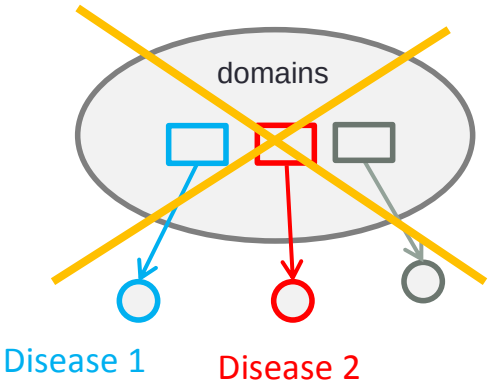
Location of mutations in different domains does not explain the difference between RASopathy and cancer mutations

Distribution of somatic and germline mutations in 98 different structural domains and inter-structural regions



'Edgetics' does not explain it

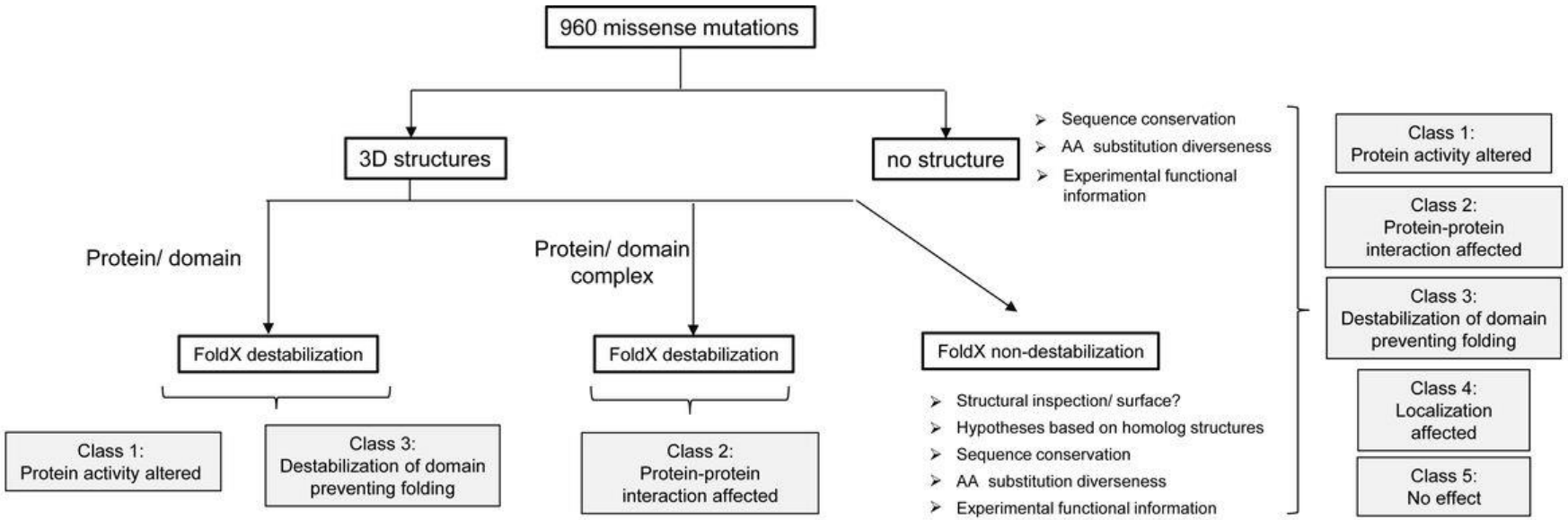
Domain localization of mutation does not explain why a particular mutation will cause RASopathy or cancer



Kiel & Serrano, Mol Sys Biol, 2014

Analysis of 956 missense mutations in RASopathies and cancer based on structural information and FoldX energies

Pipeline:



Kiel & Serrano, Mol Sys Biol, 2014

FoldX-based energy calculations of proteins

3D Structural information

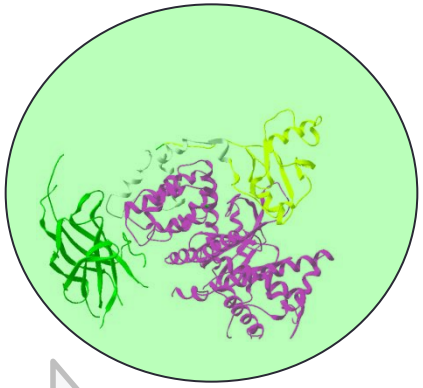
A force field for energy calculations and protein design

Schymkowitz et et al, *Nucleic Acids Res*, 2005

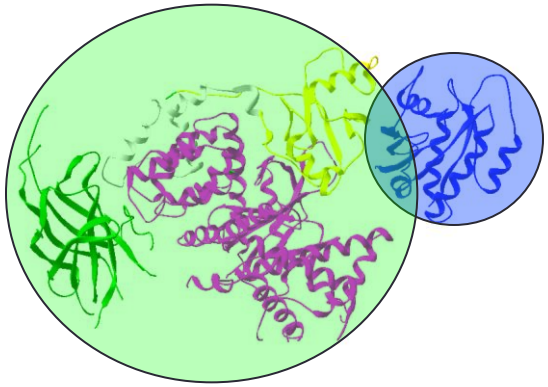
+ FoldX = ΔG

Relation to affinity: $\Delta G = RT \ln K_d$

✓ Total free energy

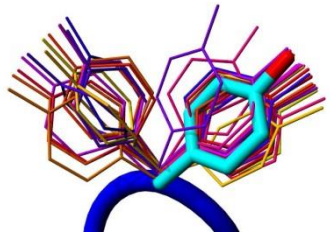


✓ Interaction energy



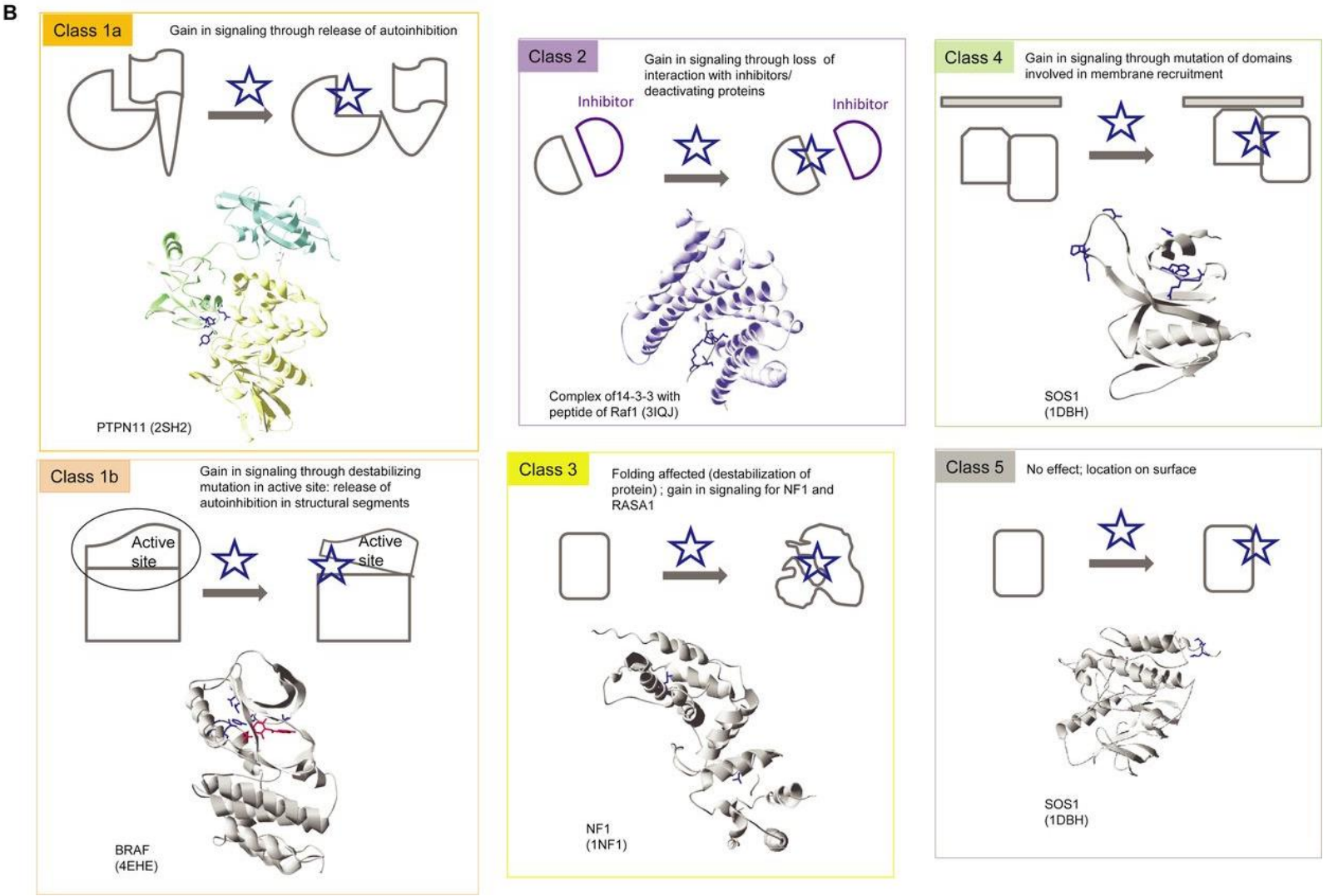
✓ Mutagenesis

A rotamer library to replace the 20 amino acids



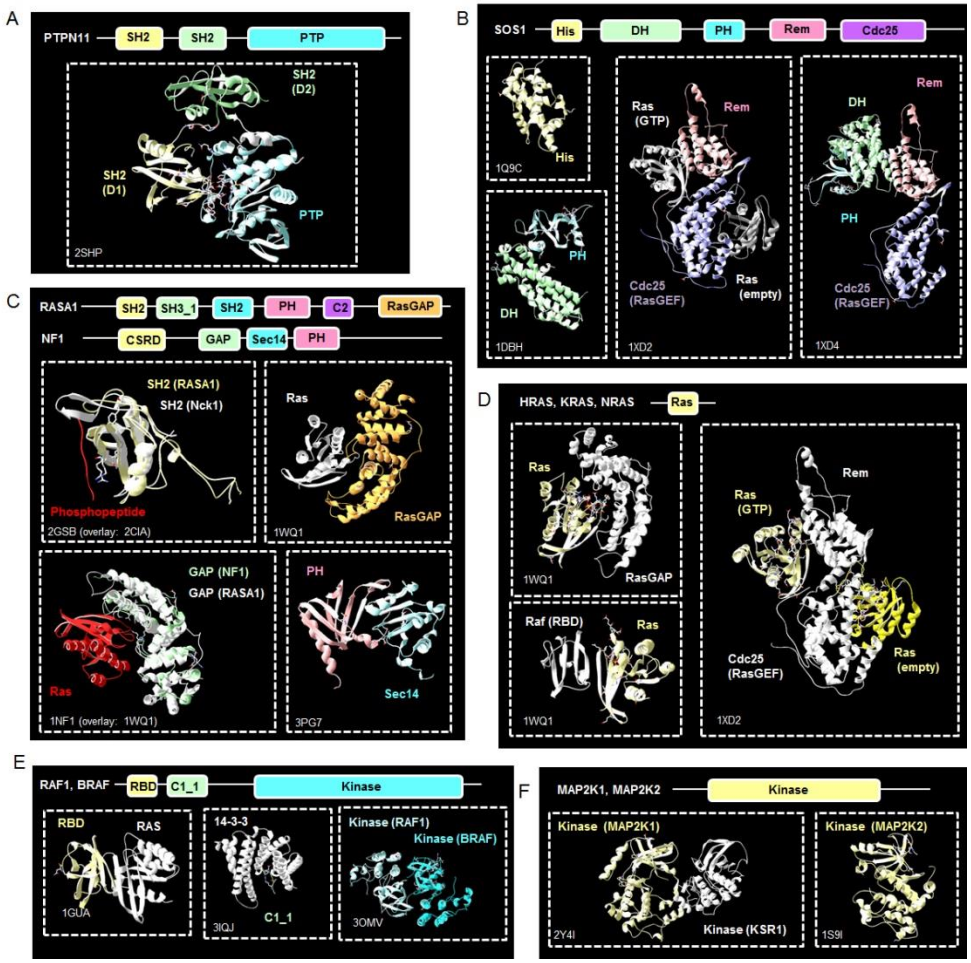
Protein design

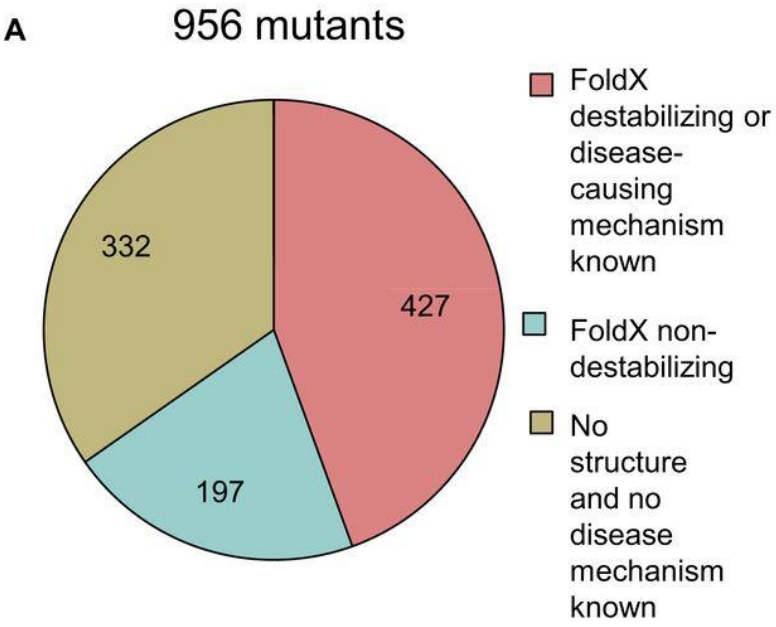
Analysis of 956 missense mutations in RASopathies and cancer based on structural information and FoldX energies



Kiel & Serrano, Mol Sys Biol, 2014

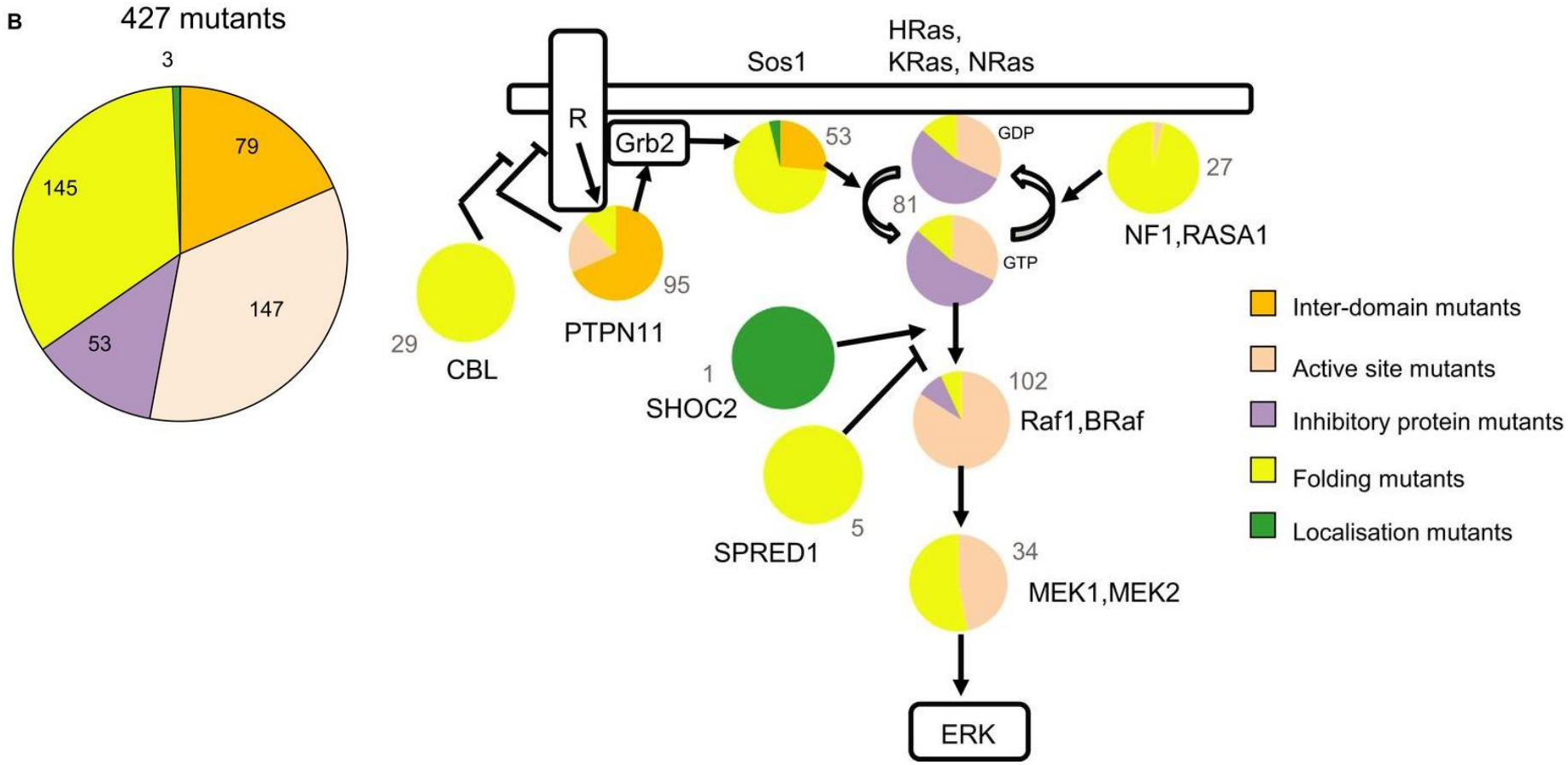
Analysis of 956 missense mutations in RASopathies and cancer: high structural coverage





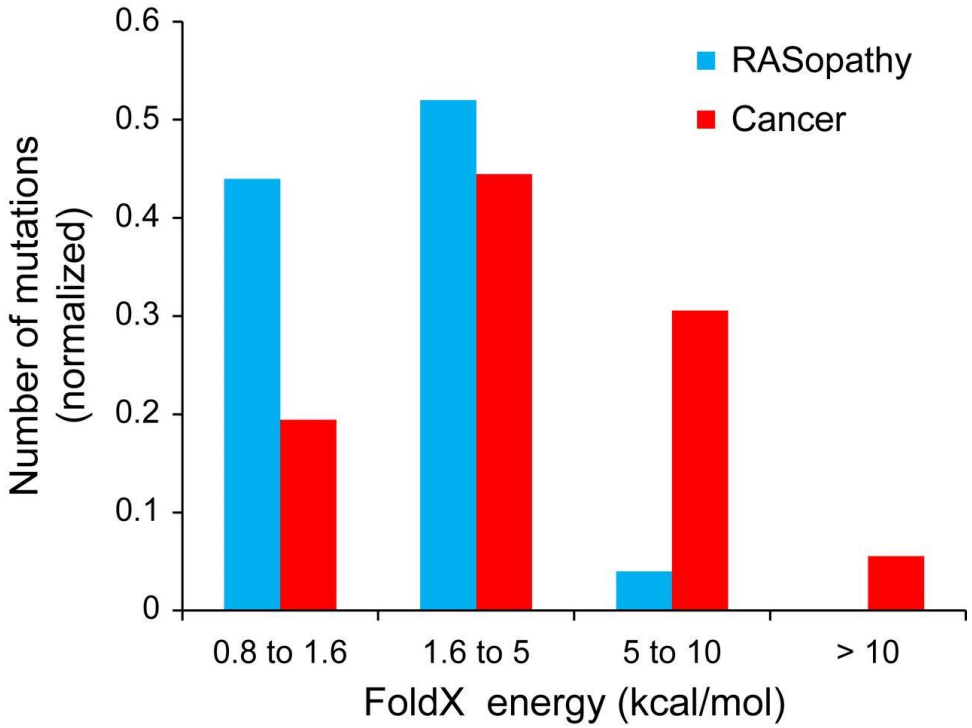
Kiel & Serrano, Mol Sys Biol, 2014

Multiple effects of a mutation even for the same protein/ protein class



Kiel & Serrano, Mol Sys Biol, 2014

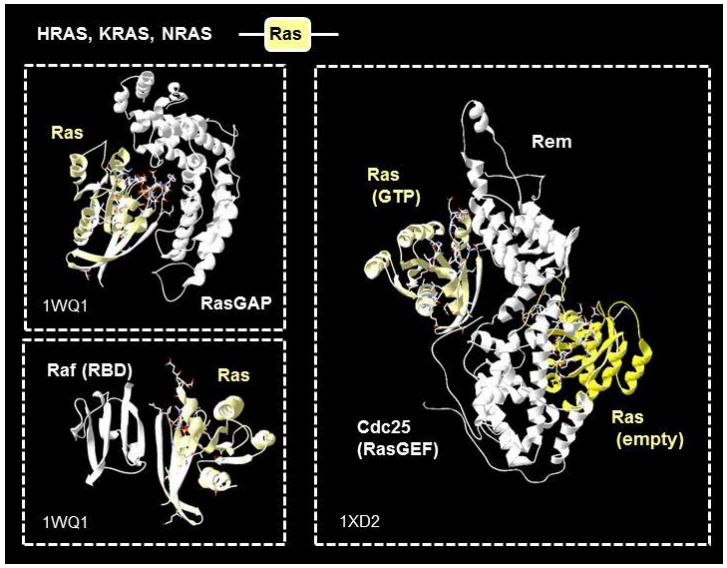
Cancer mutations tend to have higher destabilization values (on average)

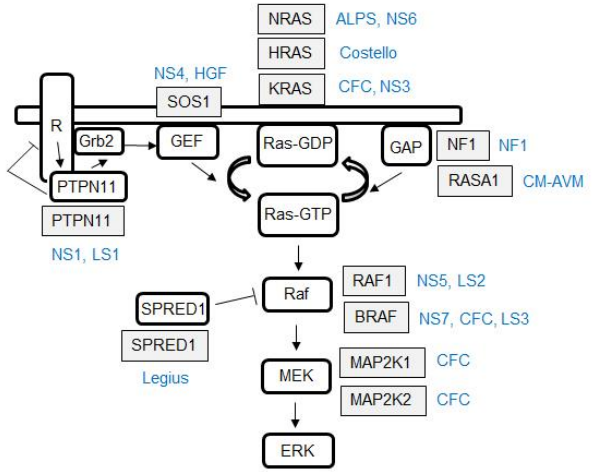


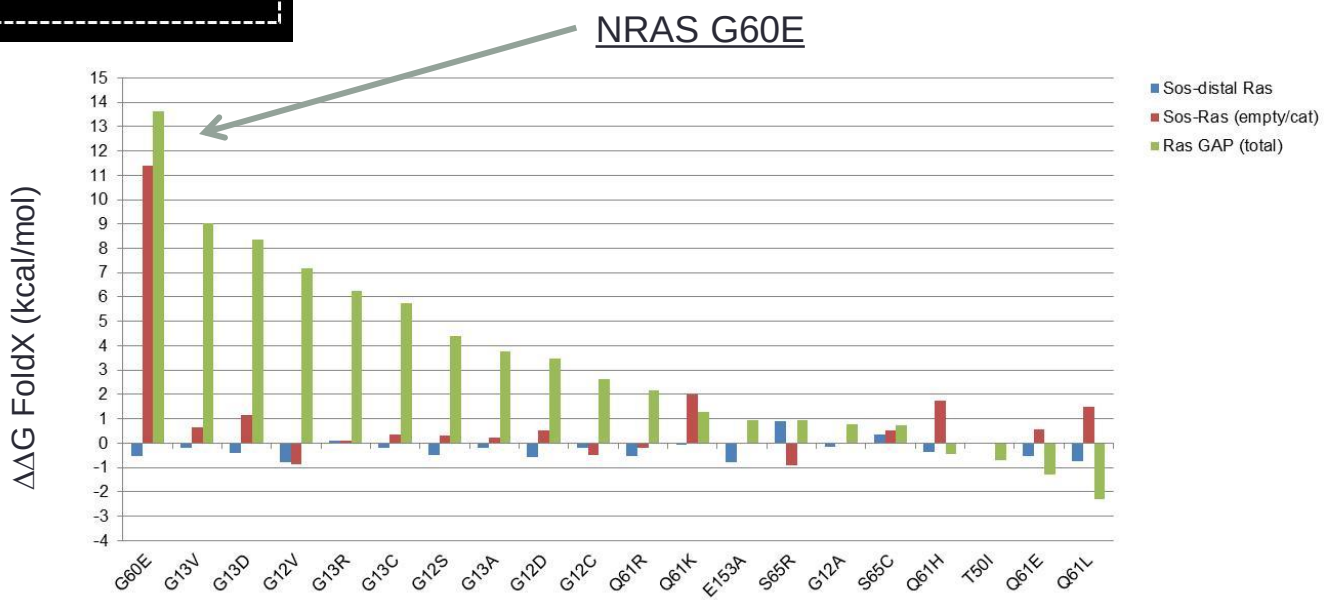
FoldX

Kiel & Serrano, Mol Sys Biol, 2014

Compensatory effects of mutations on different interaction partners



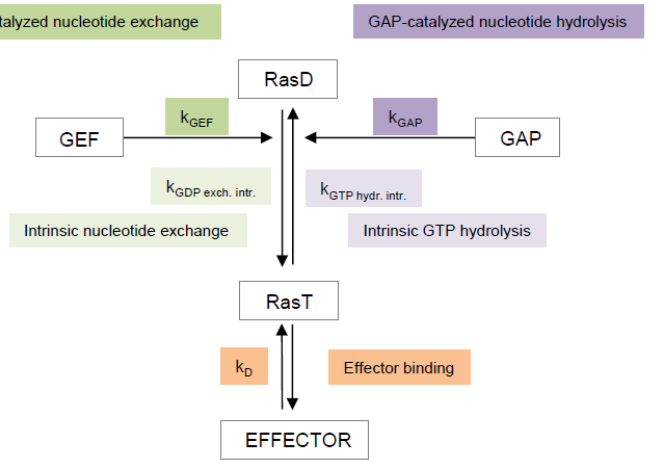




Kiel & Serrano, Mol Sys Biol, 2014

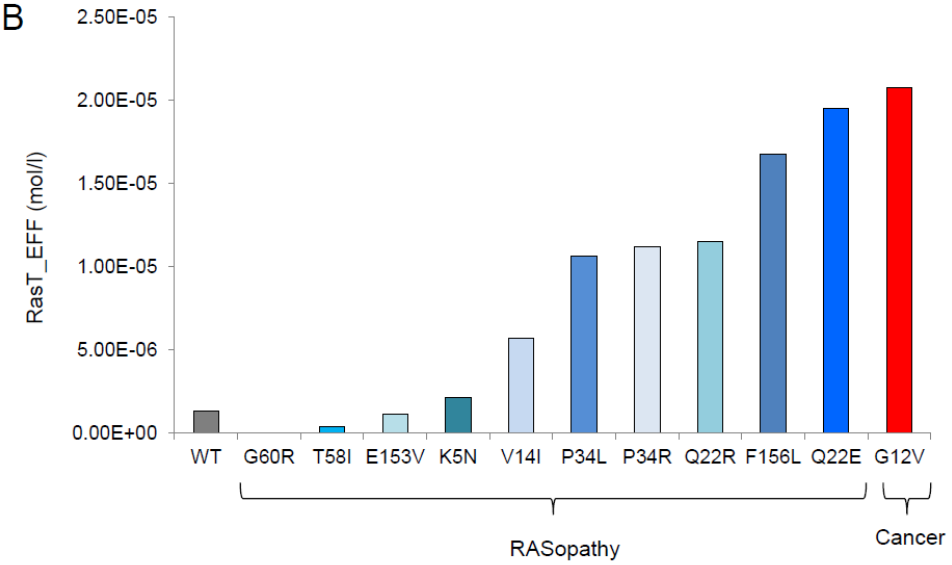
Quantitative effects on protein stability, or activity could explain in some cases the different phenotype: cancer or RASopathy

A



Simulation of Ras activation

B



‘Enedgetics’: quantitative edge effects

‘Edgetics’ + energies = ‘enedgetics’

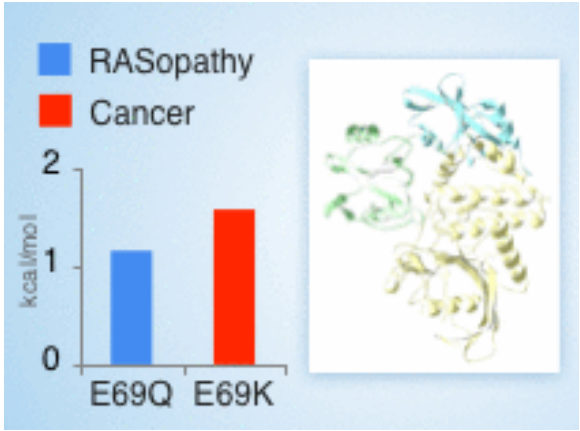
Quantitative effects on protein stability, activity, or folding explains in some cases the different phenotype

Kiel & Serrano, Mol Sys Biol, 2014

Conclusions example 1: RASopathy vs cancer

- A systematic analysis of 956 RASopathy and cancer mutations based on structures and energy predictions is presented.
- Even for the same gene, different disease-causing mechanisms exist depending on the type of mutation.
- Energy changes are higher for cancer compared to RASopathy mutations.
- In some cases, RASopathy mutations show compensatory changes that, as predicted by network modelling, result only in minor pathway deregulation.

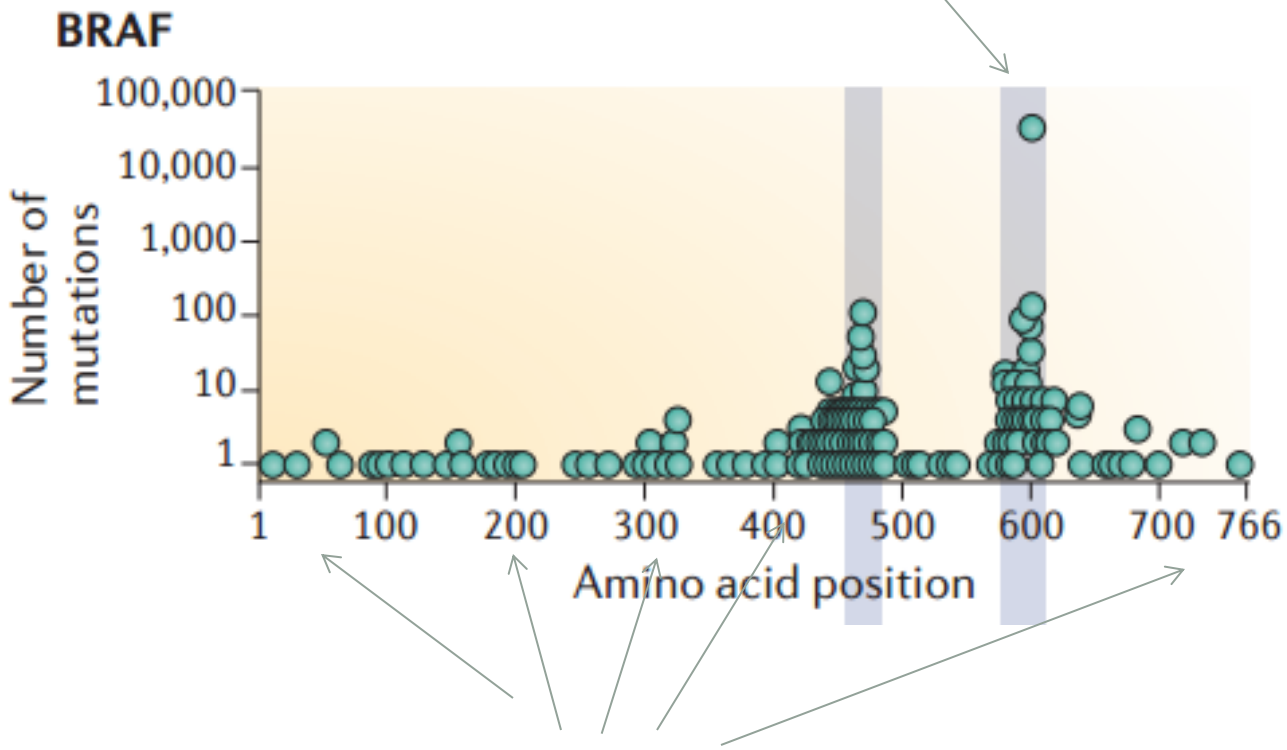
➢ Combined network-based and structural analyses show that quantitative changes rather than all-or-none rewiring underlie the difference between RASopathy and Cancer mutations.



Example 2: BRAF mutations in cancer. Why V600E?

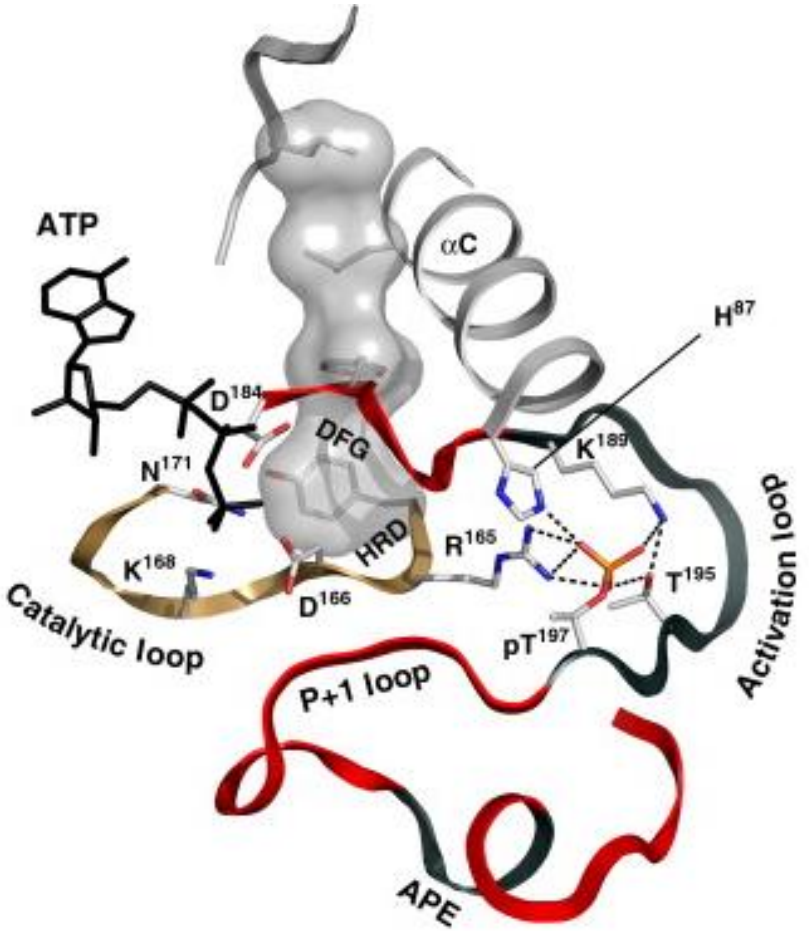
The most common BRAF mutation is V600E and induces constitutive kinase activation

Patients are treated with a BRAF kinase inhibitor



Shall we only treat patients which harbour V600E mutations or also patients with non-V600E mutations?

Kinases are activated through mutations in the activation loop (activation segment)

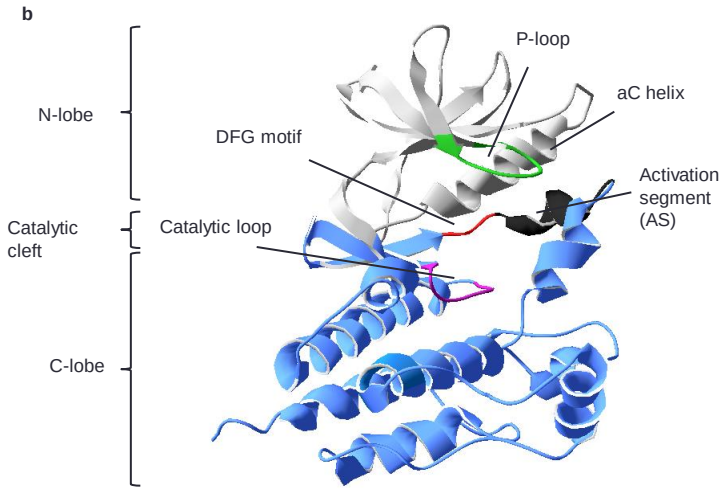


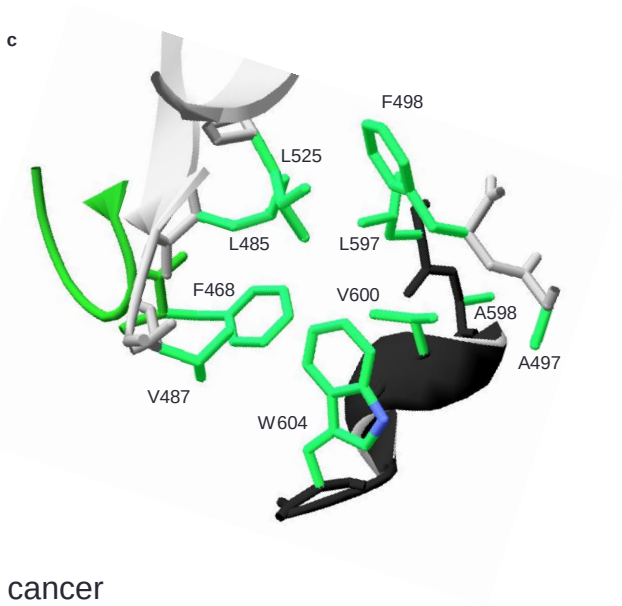
- phosphorylation in the activation segment causes structural **rearrangements of the activation segment and the aC helix**. This reorients the DFG loop resulting in activation of the kinase

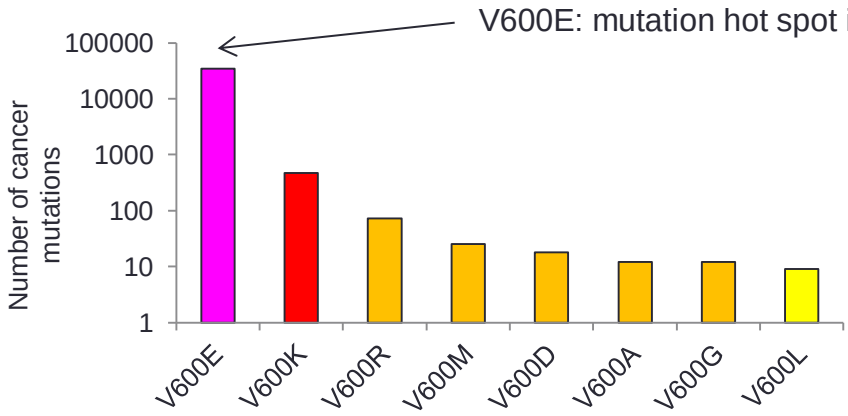
Taylor & Kornev, TIBS, 2011

Focus on the position Val600 in the kinase BRAF

V600 is buried in a hydrophobic pocket formed by the activation segment (AS) and the aC helix







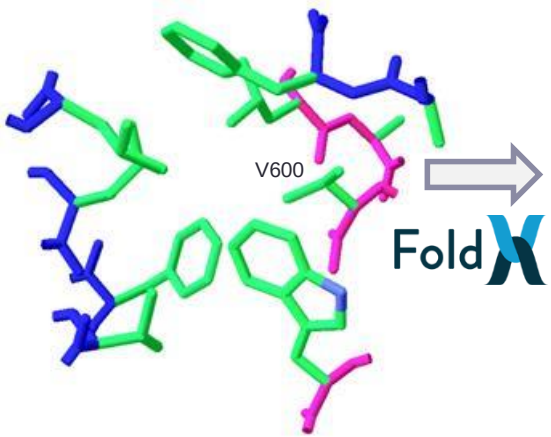
Differences in mutation frequencies: a quantitative effect?

Kiel et al, Elife, 2016

The V600E mutation causes a high destabilization of the inactive state (aC helix/AS hydrophobic pocket)

Total destabilization:

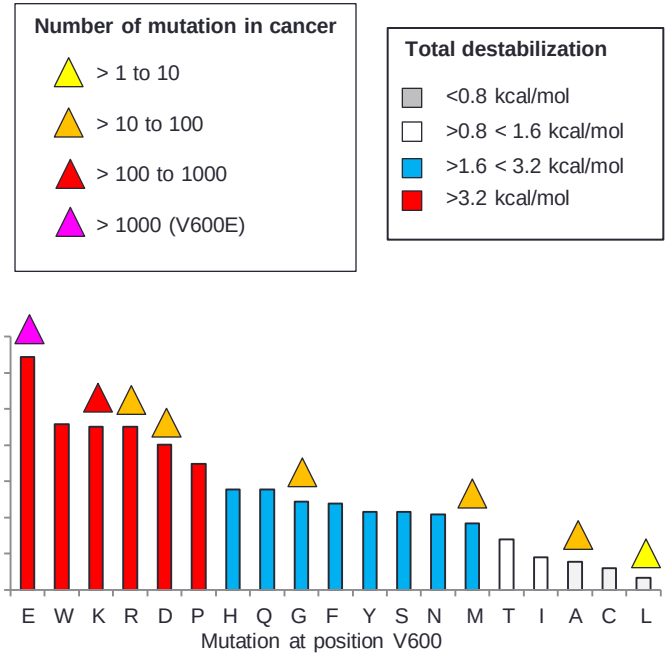
FoldX ΔΔG (BRAF_T)



A C D E F G H I K L M N P Q R S T V W Y

FoldX = ΔG

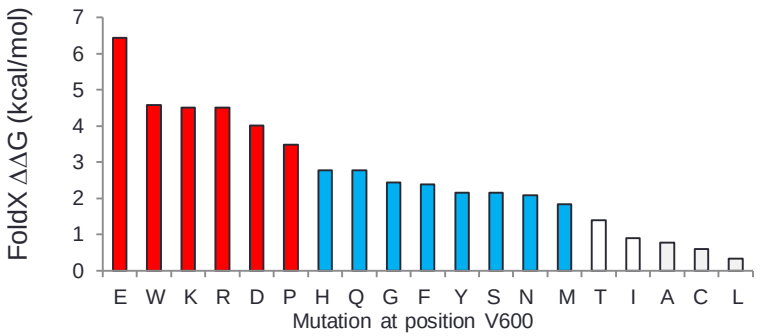
Destabilization of inactive state



No destabilization of active state (data not shown)

Kiel et al, Elife, 2016

Distinguishing driver from passenger mutations



V600K, D, and R have very similar destabilizing energies > cancer driver

Fitness??

V600A, M, and L are not very destabilizing > cancer passenger

Kiel et al, Elife, 2016

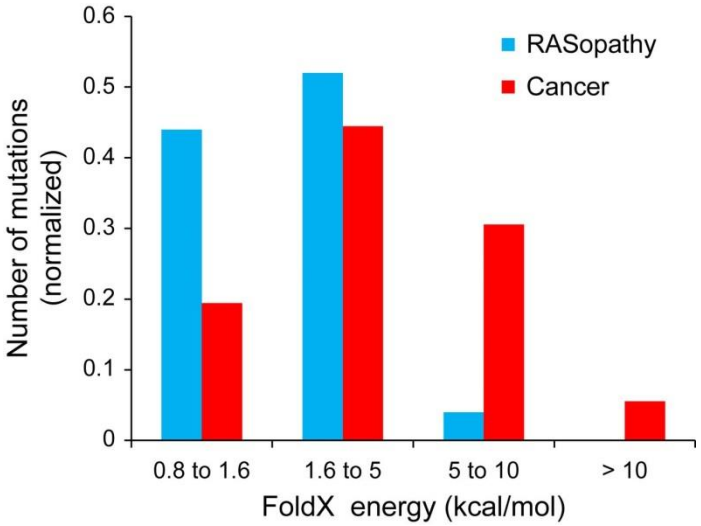
V600G behaves more like a RASopathy mutation

Google search for "V600G BRAF CFC syndrome": V600G found as a RASopathy mutation☺

Germline mutation in *BRAF* codon 600 is compatible with human development: *de novo* p.V600G mutation identified in a patient with CFC syndrome

Champion, KJ[1]; Bunag, C[2]; Estep, AL[2]; Jones, JR[1]; Bolt, CH[1]; Rogers, RC[1]; Rauen, KA[3]; Everman, DB[1]

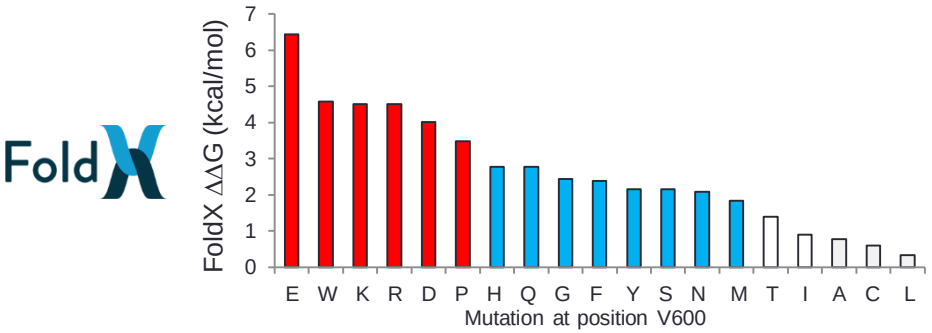
Clinical Genetics, Volume 79, issue 5 (May 2011), p. 468-474.
ISSN: 0009-9163 DOI: 10.1111/j.1399-0004.2010.01495.x
Blackwell Publishing Ltd



"enedgetics"
Cancer mutations tend to have higher destabilization values (on average)

Kiel & Serrano, 2014

Why different cancer frequencies for V600E, V600D and V600K?

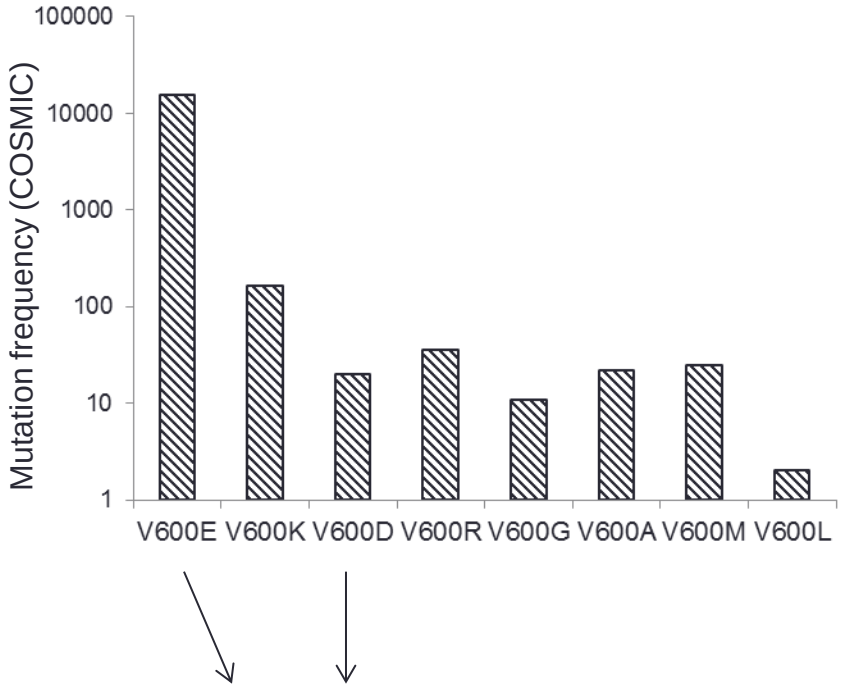


V600K, D, and R have very similar destabilizing energies

Why is V600E the by far most frequent mutation?

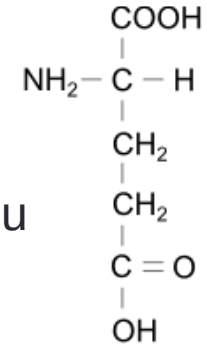
aa	frequency
Glu	15474
Lys	164
Arg	36
Met	25
Ala	22
Asp	20
Gly	11
Leu	2

Why different cancer frequencies for V600E, V600D and V600K?

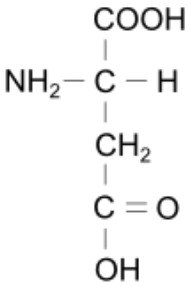


V600E: 15474 frequency
V600D: 20 frequency

Glu



Distinguishing cancer driver from passenger mutations:
Is V600E a driver mutation and V600D a passenger mutation?
On the molecular level: Glu and Asp have similar biochemical properties



Asp

Why different cancer frequencies for V600E, V600D and V600K?

1st letter	Second Letter: U	Second Letter: C	Second Letter: A	Second Letter: G	3rd letter
U	UUU Phe UUC UUA Leu UUG	UCU UCC Ser UCA UCG	UAU Tyr UAC UAA Stop UAG Stop	UGU Cys UGC UGA Stop UGG Trp	U C A G
C	CUU CUC Leu CUA CUG	CCU CCC Pro CCA CCG	CAU His CAC CAA Gln CAG	CGU CGC Arg CGA CGG	U C A G
A	AUU AUC Ile AUA AUG Met	ACU ACC Thr ACA ACG	AAU Asn AAC AAA Lys AAG	AGU Ser AGC AGA Arg AGG	U C A G
G	GUU GUC Val GUA GUG	GCU GCC Ala GCA GCG	GAU Asp GAC GAA Glu GAG	GGU GGC Gly GGA GGG	U C A G

V600E: GAG

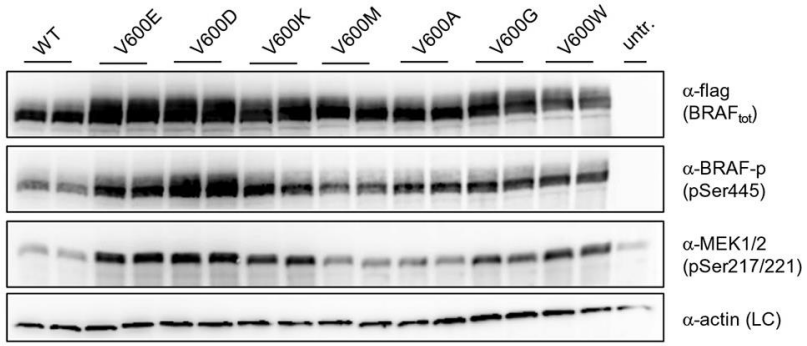
V600K: AAG

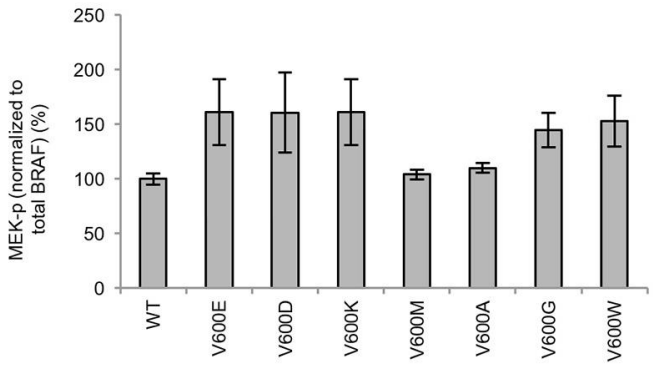
V600R: AGG

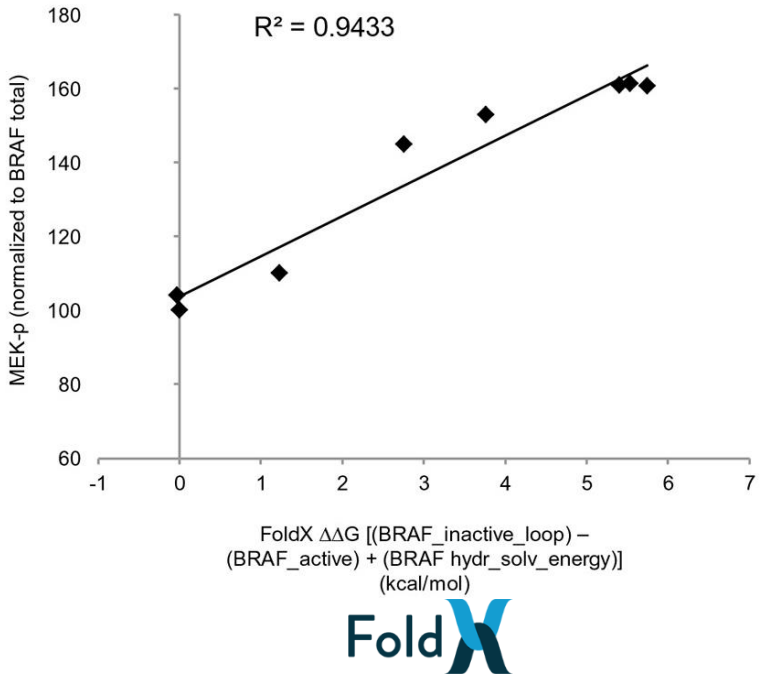
V600D: GAC/T

- The higher mutation frequency of V600E compared to V600D can be explained based on the number of nucleotide substitutions needed: V600D requires 2 nucleotide substitutions

Experimentally validate the effect of BRAF mutations by monitoring downstream MEK activation (HEK293 cells)

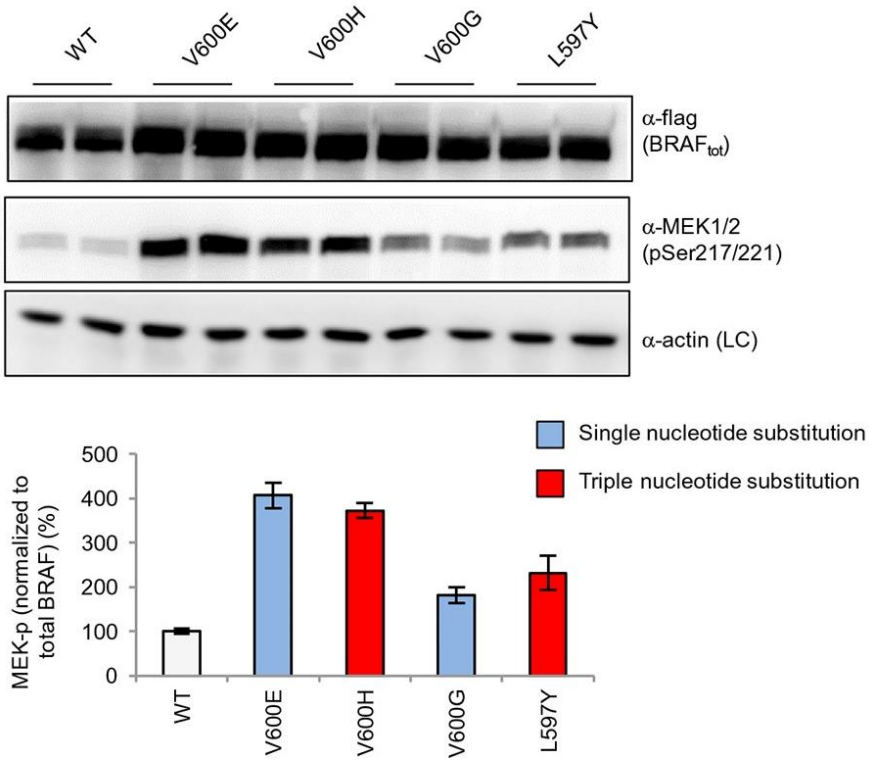






Kiel et al, Elife, 2016

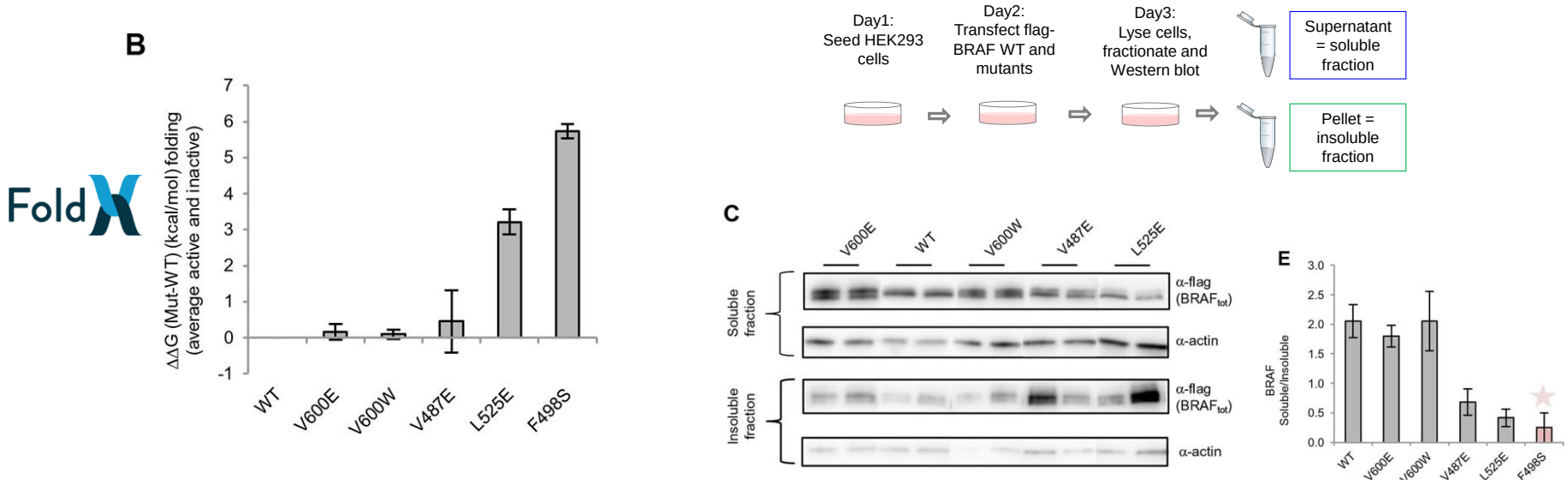
Experimentally validate the effect of BRAF mutations by monitoring downstream MEK activation (HEK293 cells)



- V600E (requires 3 nucleotide substitutions) is as active as V600E, but NOT found in cancer

Why are no mutations at other positions in the hydrophobic pocket - in a different position to Val600 - found frequently mutated in cancer?

FoldX prediction: other mutations in the hydrophobic pocket destabilize the pocket and may thereby release the AS, would also affect the folding of the inactive and/or active kinase



- Experimentally: lower BRAF expression levels (and MEK phosphorylation)

Kiel et al, Elife, 2016

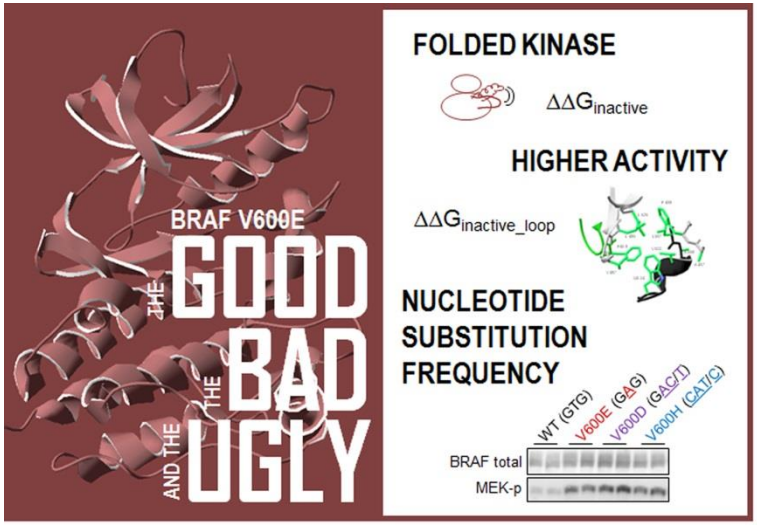
Conclusions example 3: Why BRAF V600E?

- BRAF mutation frequencies depend on the equilibrium between the destabilization of the hydrophobic pocket, the overall folding energy, the activation of the kinase and the number of bases required to change the corresponding amino acid.

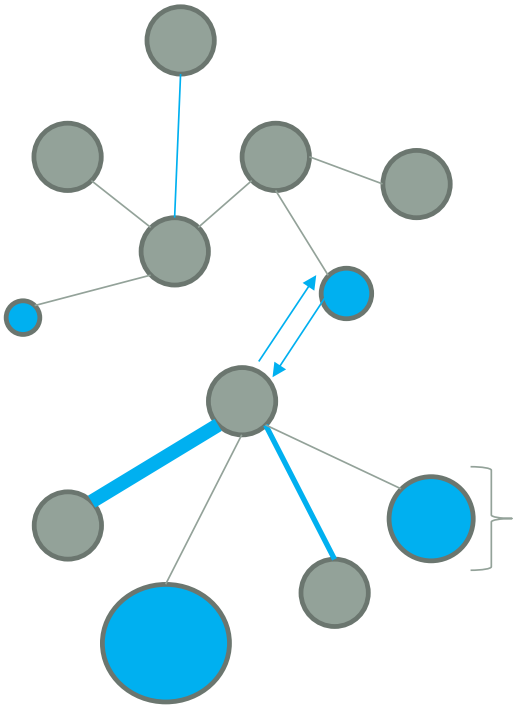
Why BRAF V600E?

- V600E is the only single nucleotide substitution (Asp, Lys, and Arg, require two bases substitutions) that opens the AS through destabilization of autoinhibitory interactions, without significantly impairing the folding of the inactive or active kinase domain.

➢ The results underscore the importance of considering changes at both the DNA and protein level when attempting to understand why certain cancer-causing mutations are more common than others.



Quantitative PPI networks



FoldX

RCSB PDB PROTEIN DATA BANK

BindingDB

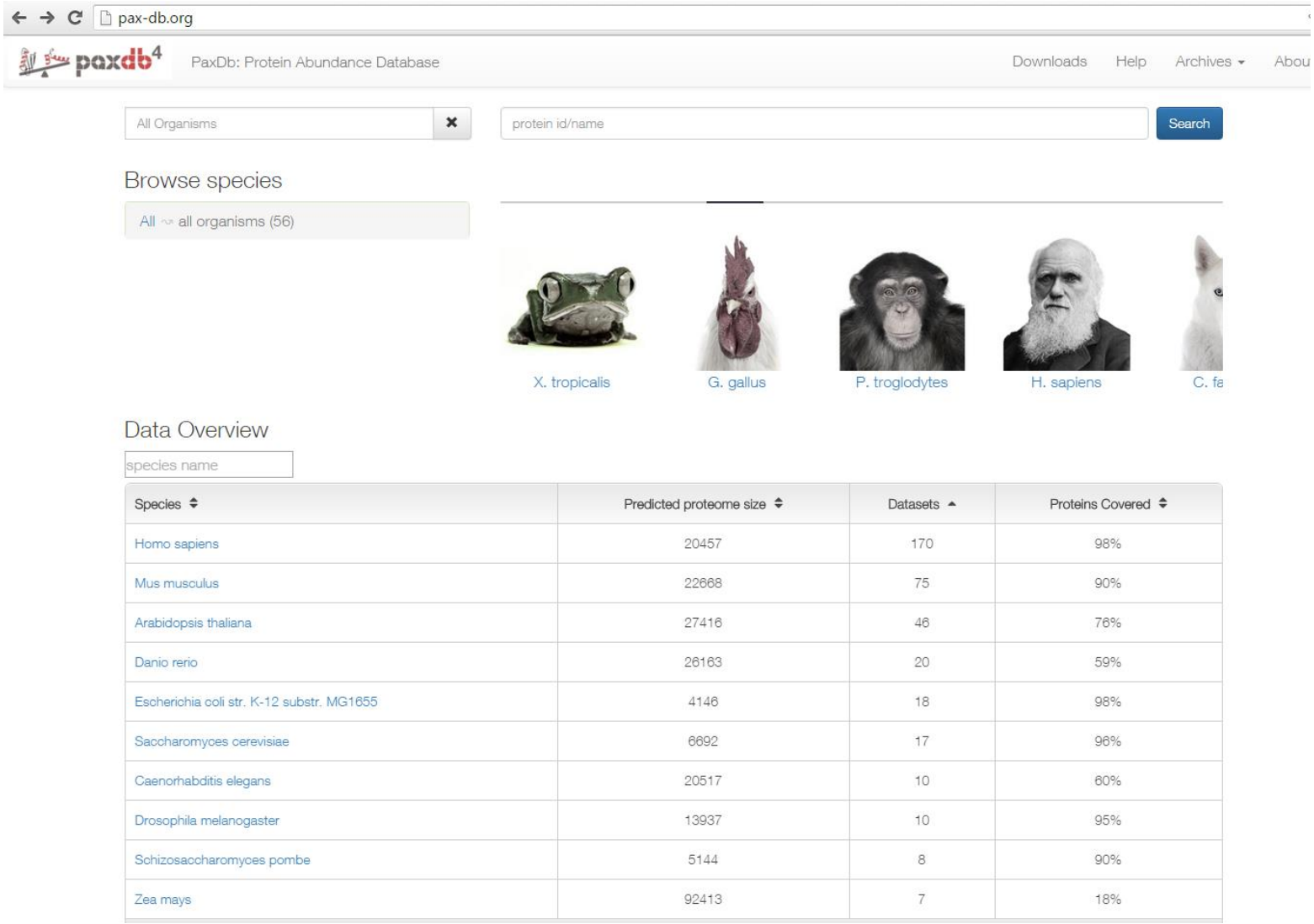
Interactome3D

dsysmap

paxdb4

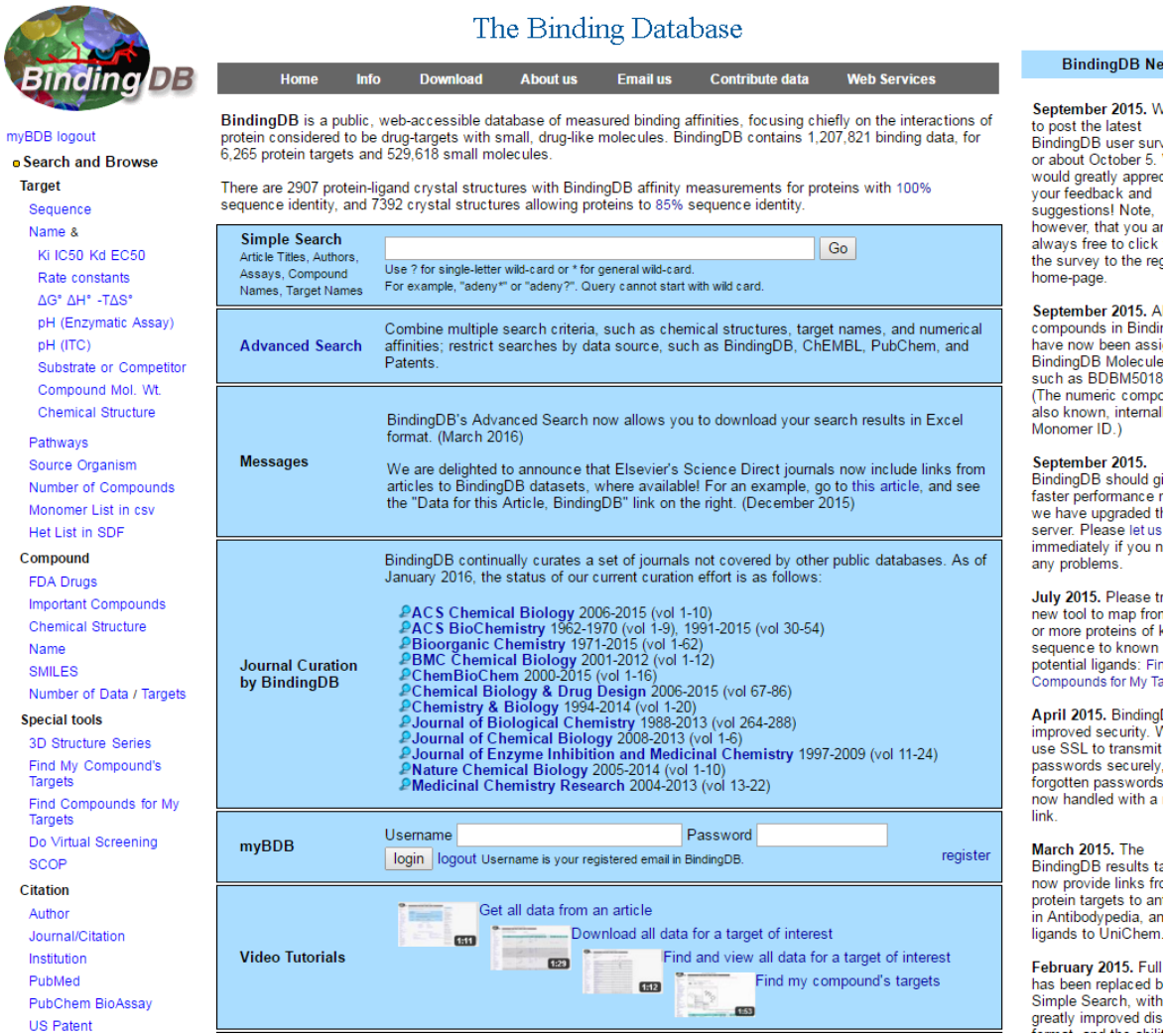
B10NUMB3R5 THE DATABASE OF USEFUL BIOLOGICAL NUMBERS

Protein abundances



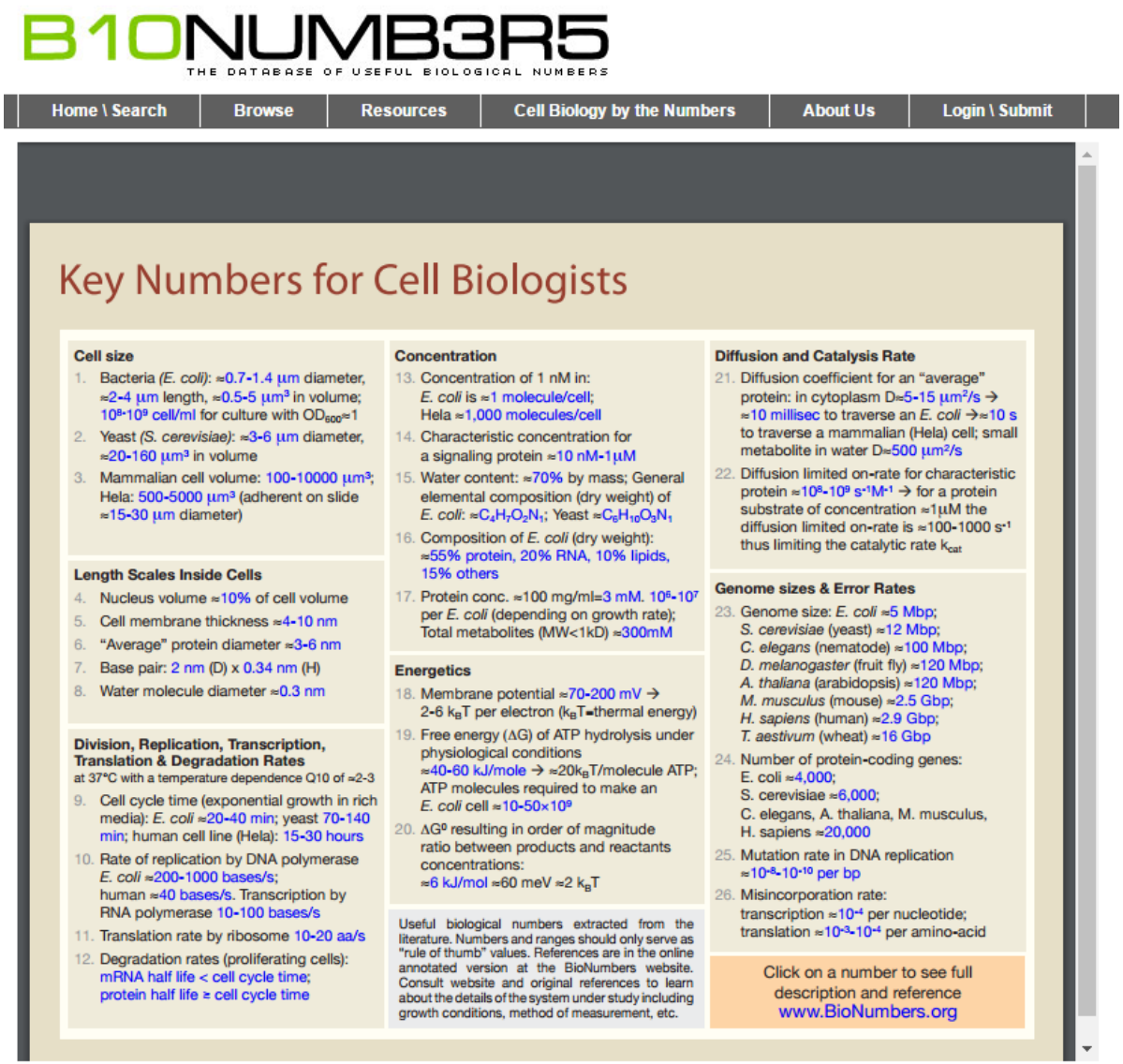
http://pax-db.org/

Affinities and kinetic constants



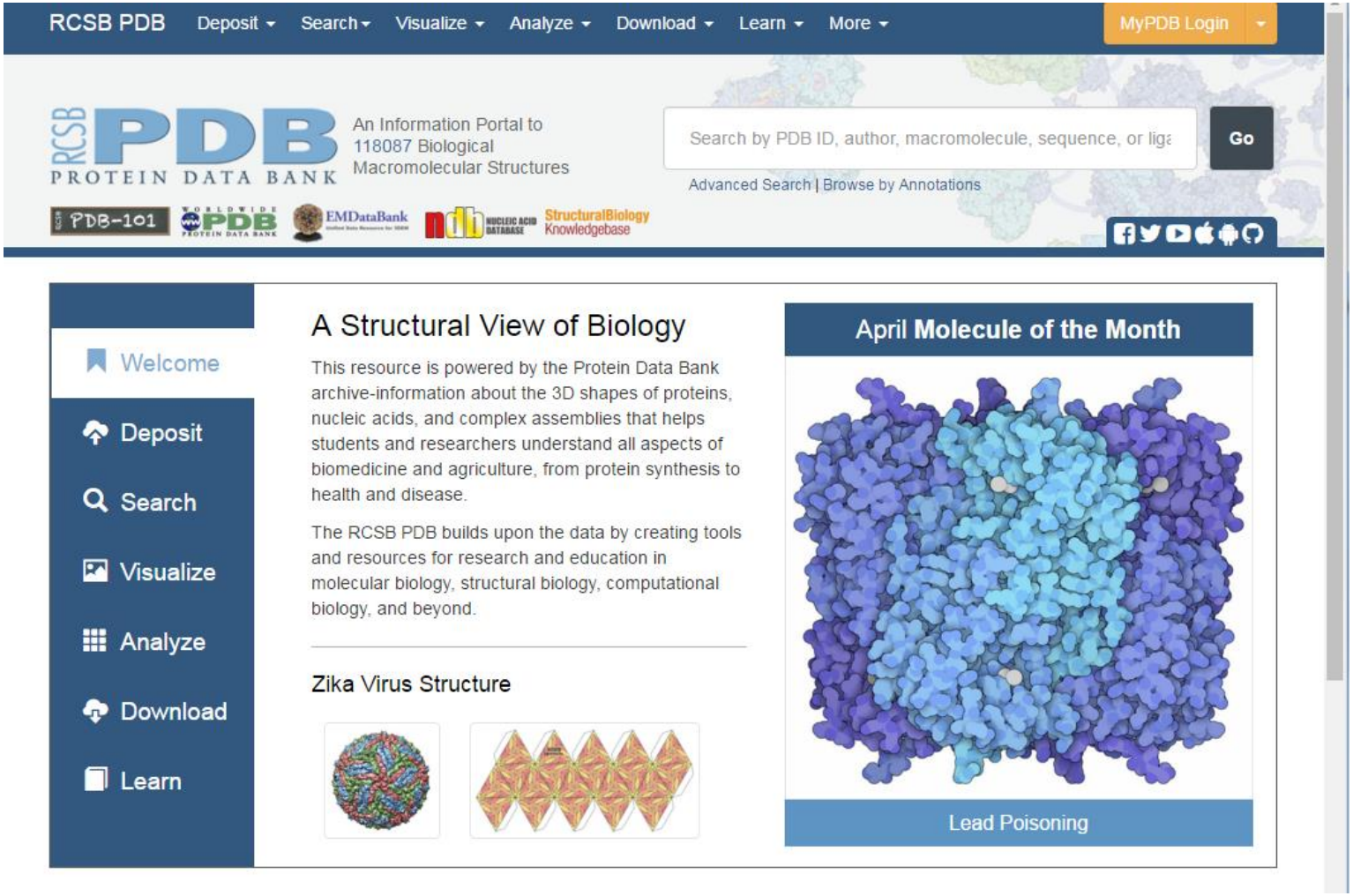
https://www.bindingdb.org/bind/index.jsp

General 'numbers' in biology



http://bionumbers.hms.harvard.edu/

Protein structures



http://www.rcsb.org/pdb/home/home.do

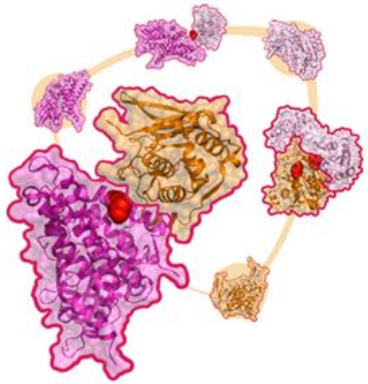
3D structures of protein interactions

http://interactome3d.irbbarcelona.org/

3D structures of protein interactions/ mapping of disease mutations

dsysmap

Home Browse Download Stats Tutorial Help About



dSysMap (**Mapping of Human disease-related mutations at the systemic level**) displays Human disease-related mutations on the structural interactome. Mapping of mutations on protein structures and on interaction interfaces allows you to visualize the region of the interactome that they affect and helps in rationalizing their mechanism of action.

The current version of **dSysMap** is 2015_05

Is this you first time with dSysMap? **Take a 5 minutes Tutorial!**

Tutorial :: Learn how to use dSysMap

Browse diseases

Select a disease from the following list.
Example: Loeys-Dietz syndrome

Type here the name of a disease or browse the list...

- Bacterial infection or mycosis
- Blood disease
- Cancer
- Cardiovascular disease
- Congenital abnormality
- Connective tissue disease
- Digestive system disease
- Ear-nose-throat disease
- Endocrine system disease
- Eye disease
- Fetal disease
- Genetic disease
- Immune system disease
- Infant-newborn disease

Query with a list of proteins

Enter a list of proteins (Uniprot AC or gene name) Example

For example...
ETFA, ETFB, ACADM, ACADS, ACADVL, SOCS3, IRF7, GPHN, RPSA

Submit

Submit your mutations

Enter a list of mutations (which format?) Example

For example...
APC: p.Ala1582Lys, p.Thr506Trp
AXIN1: p.Phe119Ala, p.Gln190Arg
DLG1

Submit

http://dsysmap.irbbarcelona.org/

Protein design

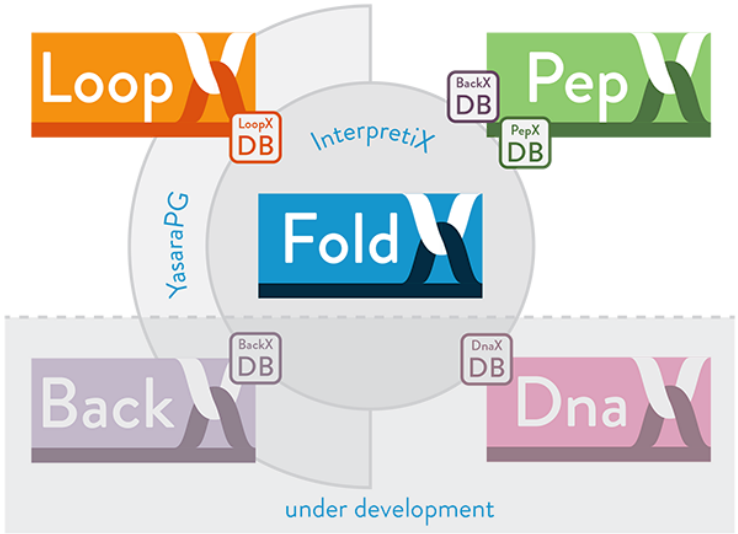
FoldX PRODUCTS LICENSING AND SERVICES SUPPORT DOWNLOAD ABOUT CRG Centre for Genomic Regulation

LOGIN REGISTER

Products / Products

PRODUCTS

Technology Overview



The FoldX Suite

The FoldX Suite builds on the strong fundament of advanced protein design features already implemented in the oldest FoldX versions and integrates new capabilities: loop reconstruction (LoopX) and peptide docking (PepX). The Suite also features an improved usability thanks to a new boost Command Line Interface.

FOLDX PAPERS

McKeone R, Wikstrom M, Kiel C, Rakoczy EP. **"Assessing the correlation between mutant rhodopsin stability and the severity of retinitis pigmentosa."** *Mol. Vis.*. 2014;20:183-99.
2014

Kiel C, Serrano L. **"Structure-energy-based predictions and network modelling of RASopathy and cancer missense mutations."** *Mol. Syst. Biol.*. 2014;10:727.
2014

De Baets G, Van Durme J, Reumers J, et al. **"SNPeffect 4.0: on-line prediction of molecular and structural effects of protein-coding variants."** *Nucleic Acids Res.*. 2012;40(Database issue):D935-9.
2012

Simões-Correia J, Figueiredo J, Lopes R, et al. **"E-cadherin destabilization accounts for the pathogenicity of missense mutations in hereditary diffuse gastric cancer."** *PLoS ONE.*
2012;7(3):e33783.
2012

Kimberley FC, van der Sloot AM, Guadagnoli M, et al. **"The design and characterization of receptor-selective APRIL variants."** *J. Biol. Chem.*.
2012;287(44):37434-46.
2012

Van Durme J, Delgado J, Stricher F, Serrano L, Schymkowitz J, Rousseau F. **"A graphical interface for the FoldX forcefield."** *Bioinformatics.*
2011;27(12):1711-2.

http://foldxsuite.crg.eu/products#foldx

Conclusions/ Wrap up

- Quantitative information is important to consider in PPI networks; however, it is often difficult to address these quantities experimentally.
- Protein quantification is not a solved problem; especially in mammalian cells, because of the problem of shared peptides for isoforms and splice variants
- It is impossible to measure binding affinities and kinetic constants in a high-throughput manner (protein expression and purification needed)
- The effect of mutations can be assessed in a quantitative manner using protein design tools, provided 3D structural information is available