

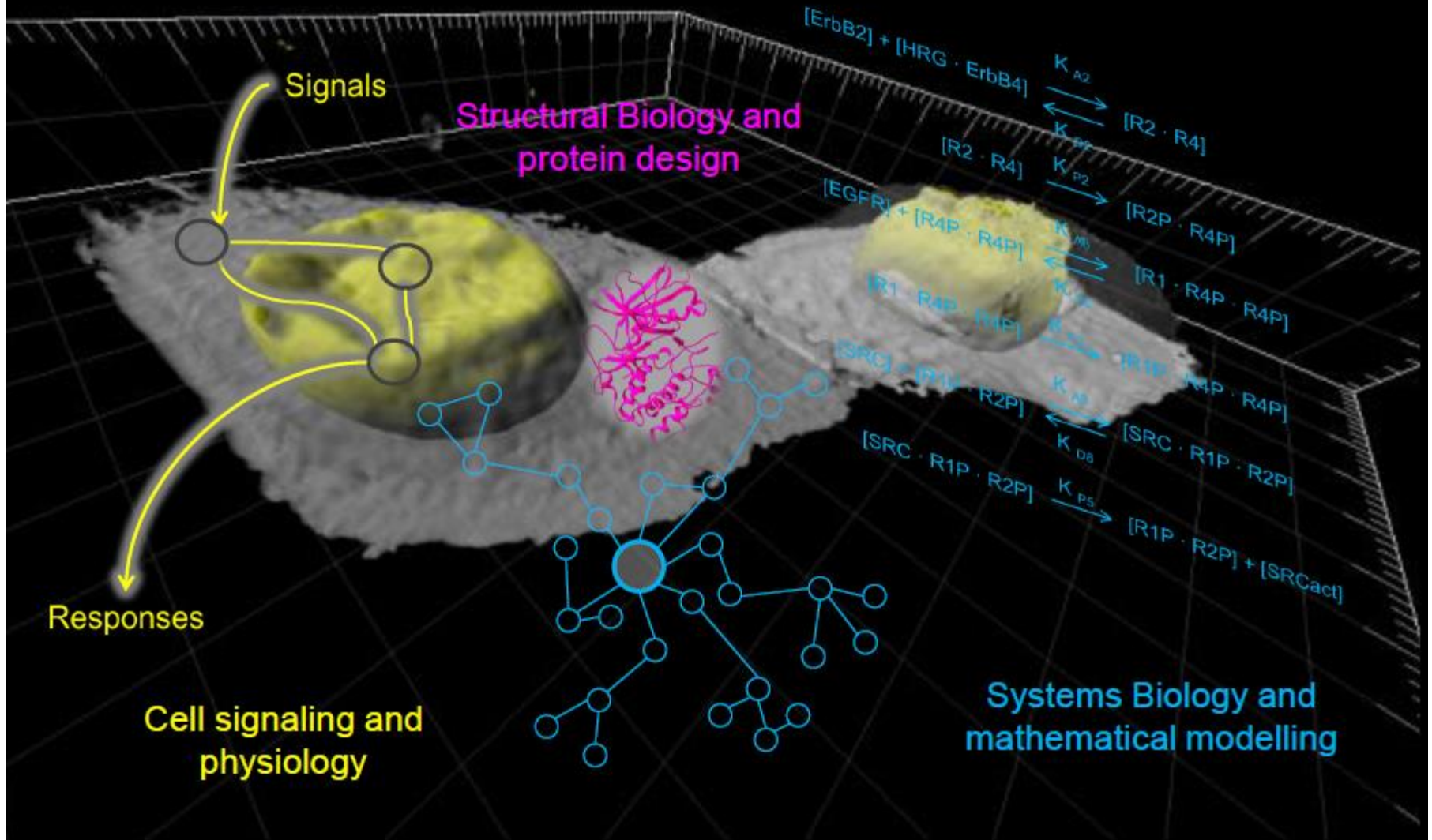
Integrative and quantitative analysis of disease mutations in protein interaction networks and implications for personalized medicine

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

CRG Barcelona: <http://crg.eu>



Exploring the molecular and quantitative mechanisms that underlie cell signaling and contribute to human disease

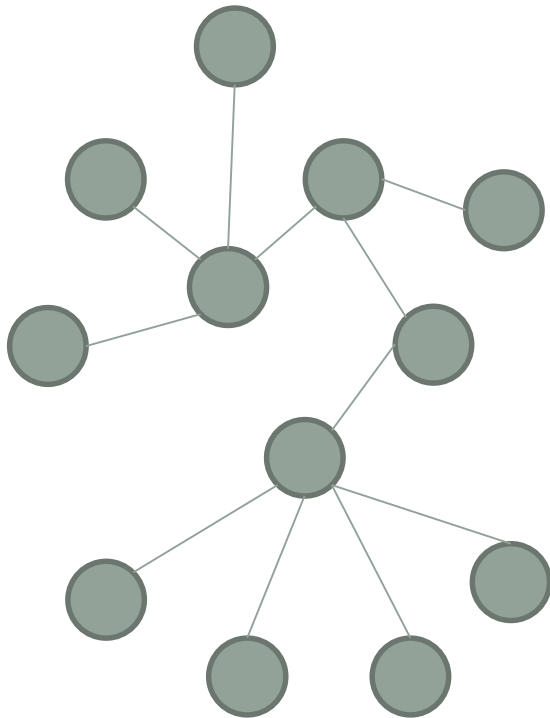


Outline

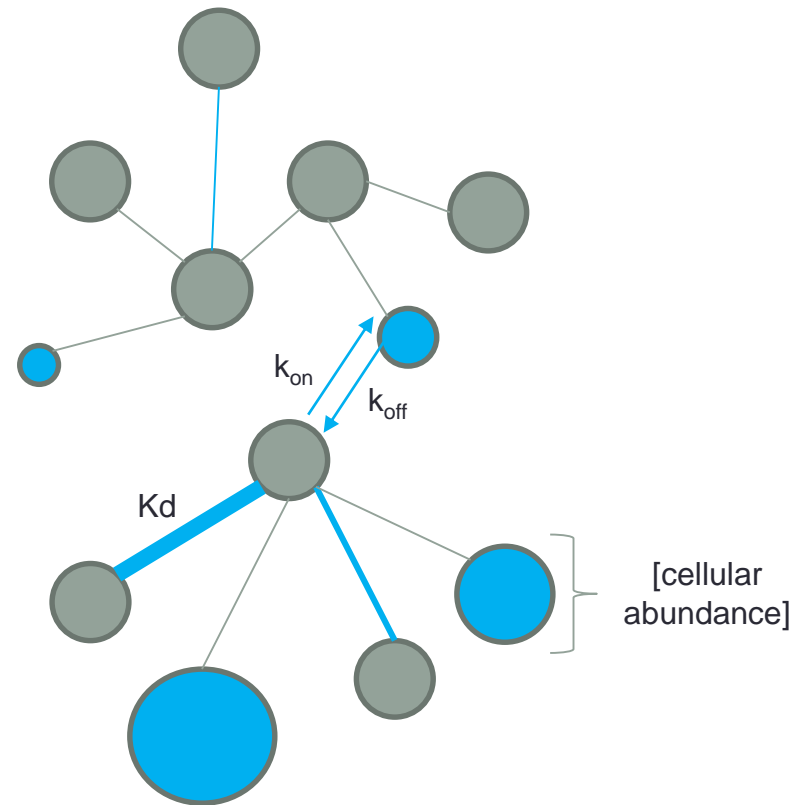
- I.** The effect of affinities, kinetic constants and network topology in PPI networks
- II.** The effect of protein abundance perturbations and interaction competition in PPI networks
- III.** Methods to quantify protein abundances, affinities, and kinetic constants
- IV.** Disease mutations and their principle effect on PPI networks
- V.** Examples for quantitative effects in disease networks
 1. RASopathy vs cancer mutations: a matter of quantity
 2. Rhodopsin stability and disease onset
 3. BRAF mutation frequency: prediction of oncogenic drivers
- VI.** Summary tools & websites
- VII.** Wrap up/ discussion/ conclusions

Quantitative information in protein-protein interaction (PPI) networks

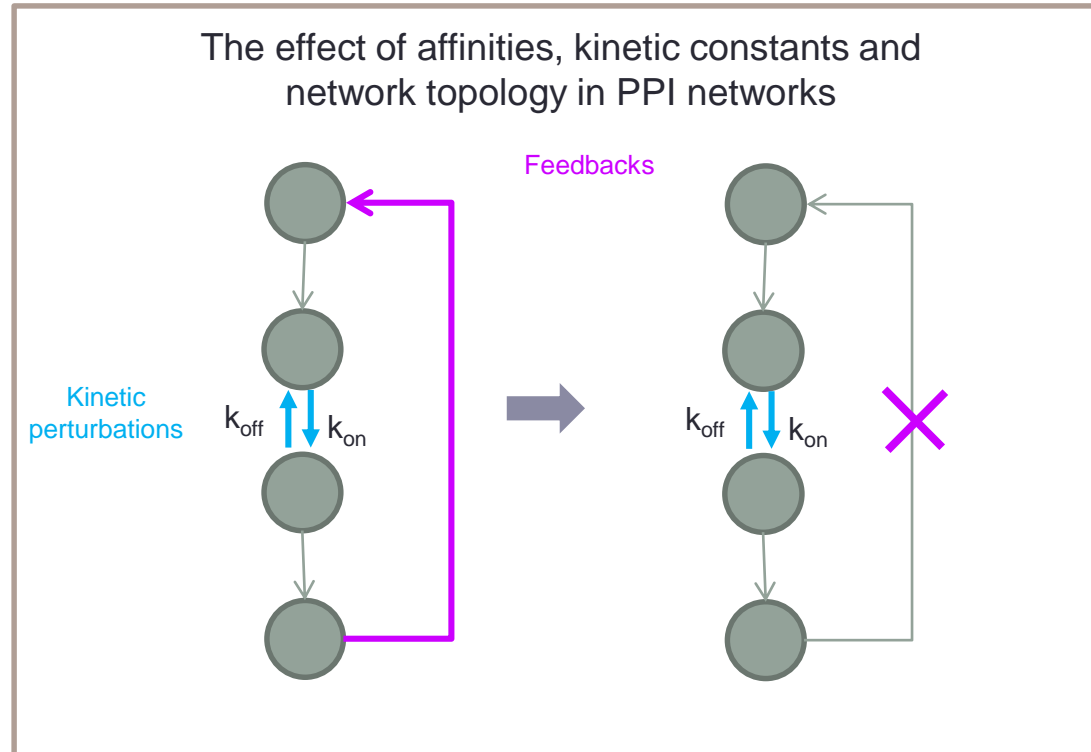
Qualitative PPI networks



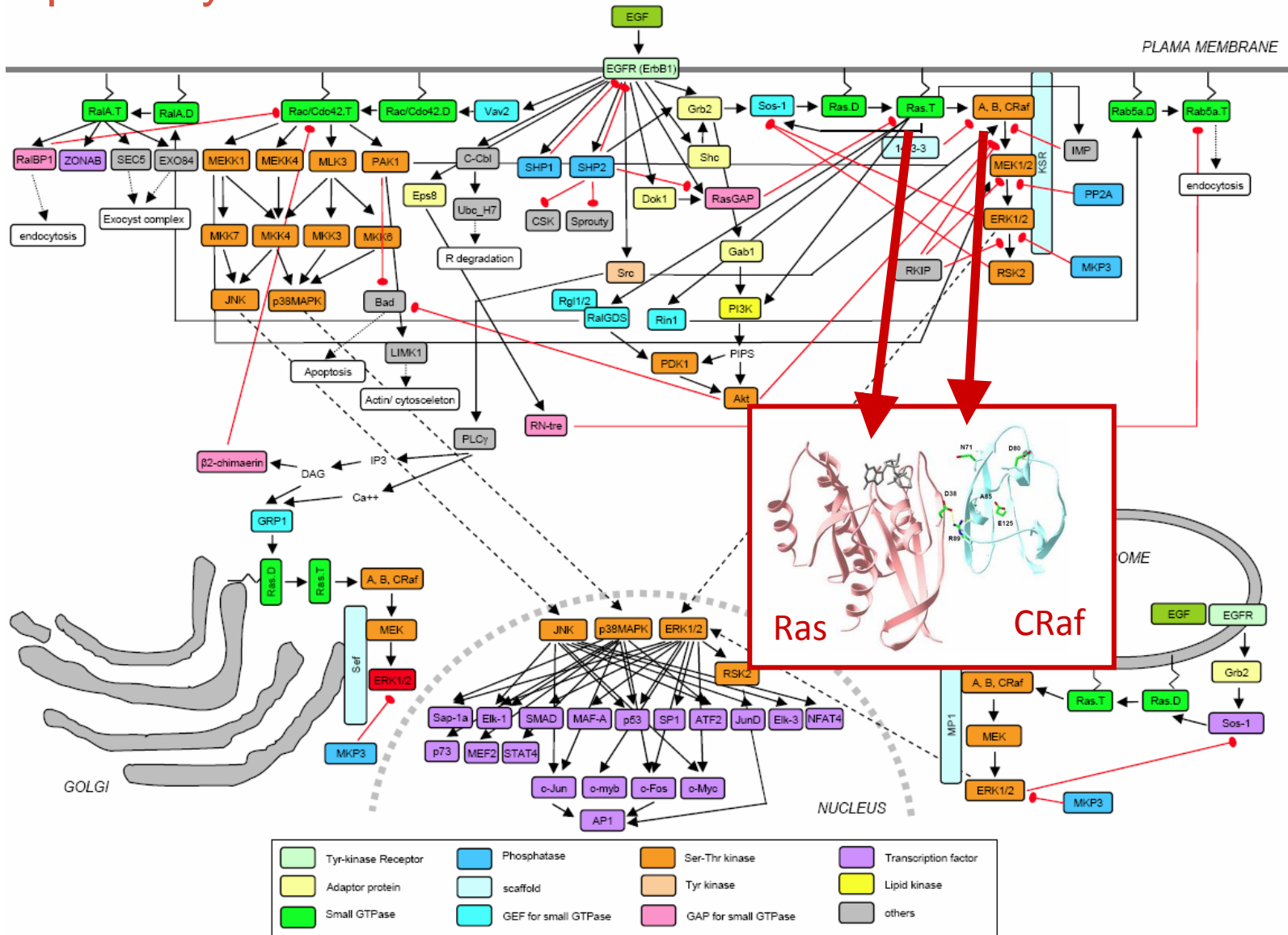
Quantitative PPI networks



Considering protein abundances and affinities/
kinetic constants



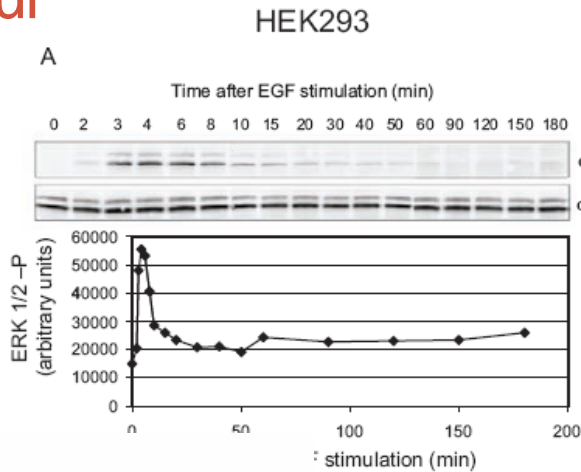
Epidermal growth factor (EGF) activates the RAS-RAF-MEK-ERK pathway



Different network 'wiring' /feedbacks causes the different behaviour

HEK293 cells

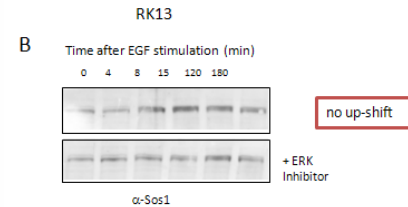
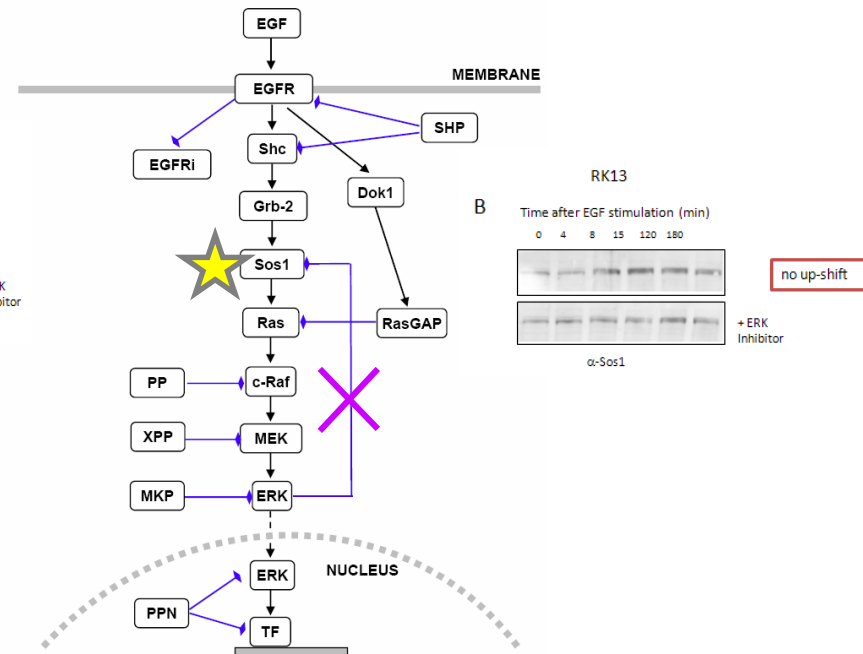
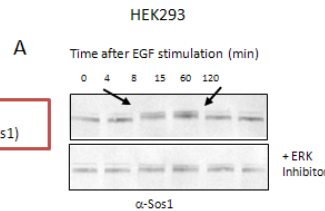
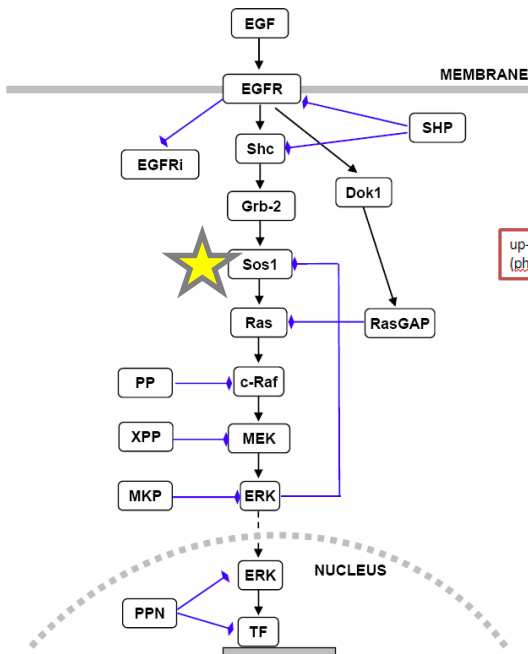
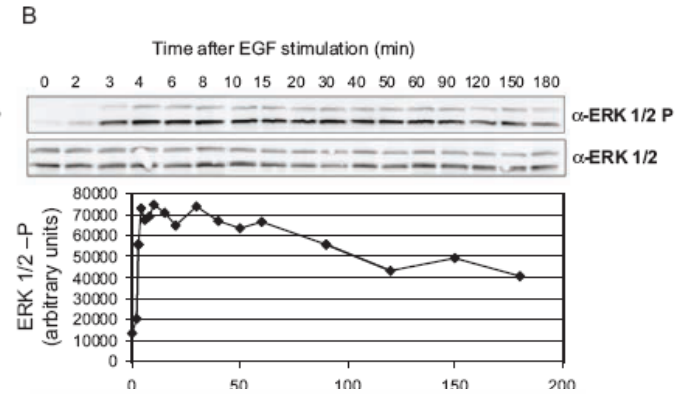
Transient response



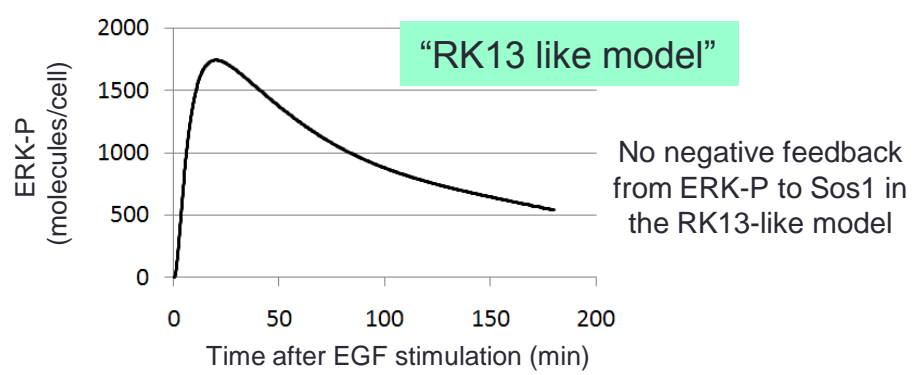
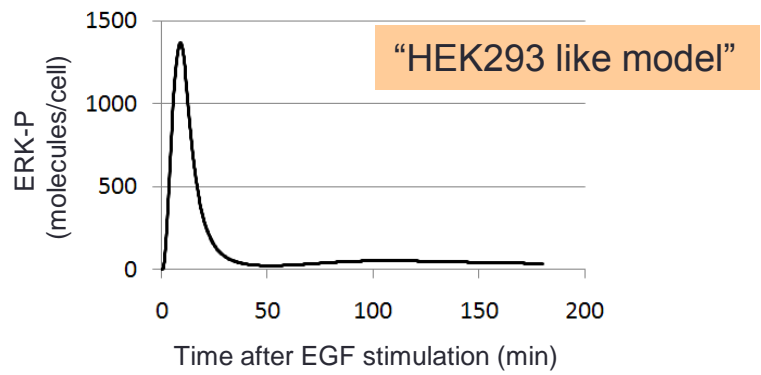
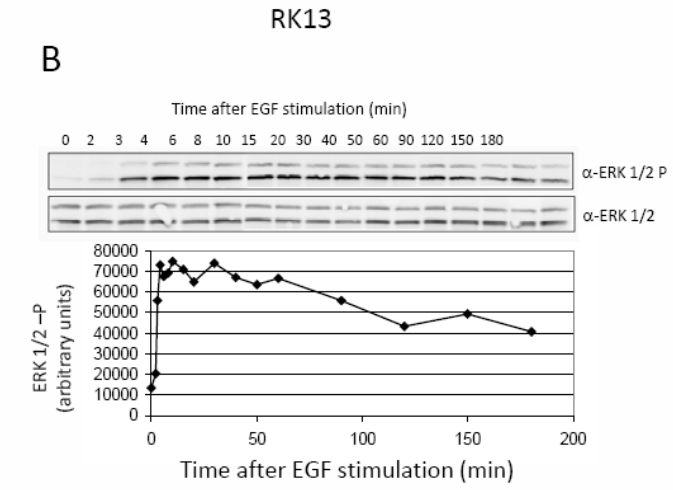
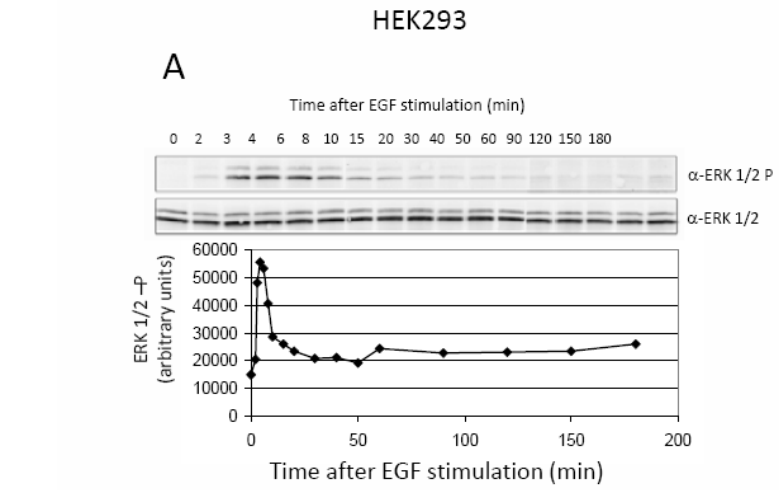
RK13

RK13 cells

Sustained response

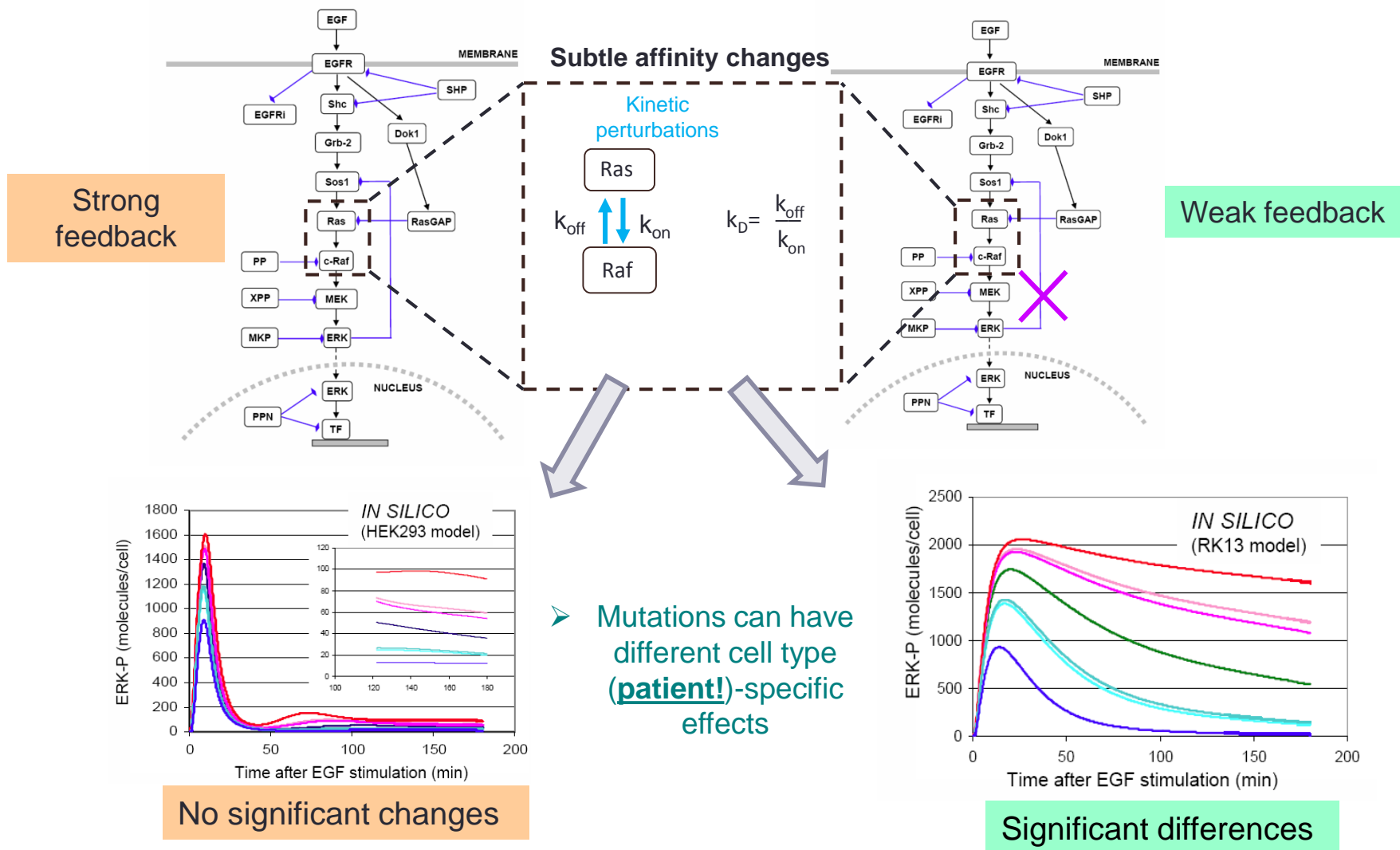


A simple computer model of ERK activation in HEK293 and RK13 cells

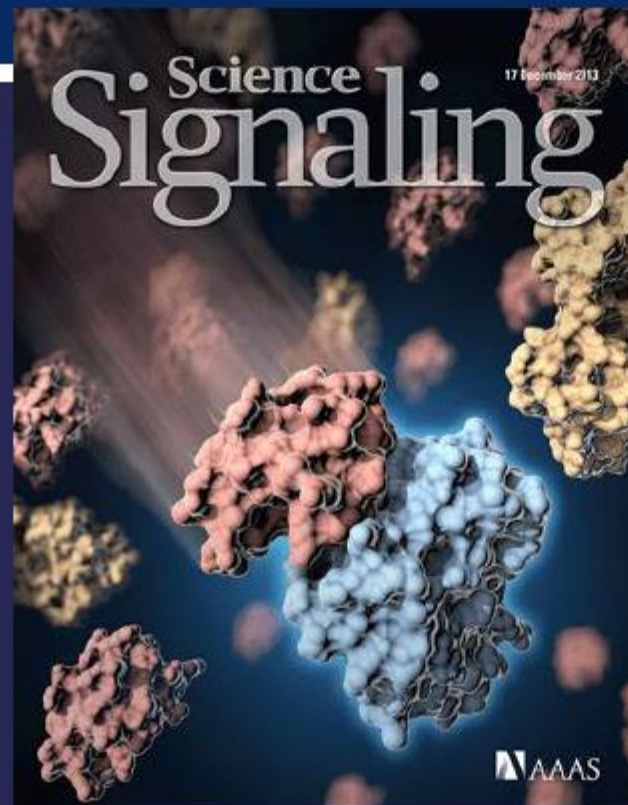


➤ Good agreement of experiment and model predictions

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



Kiel et al, 2013



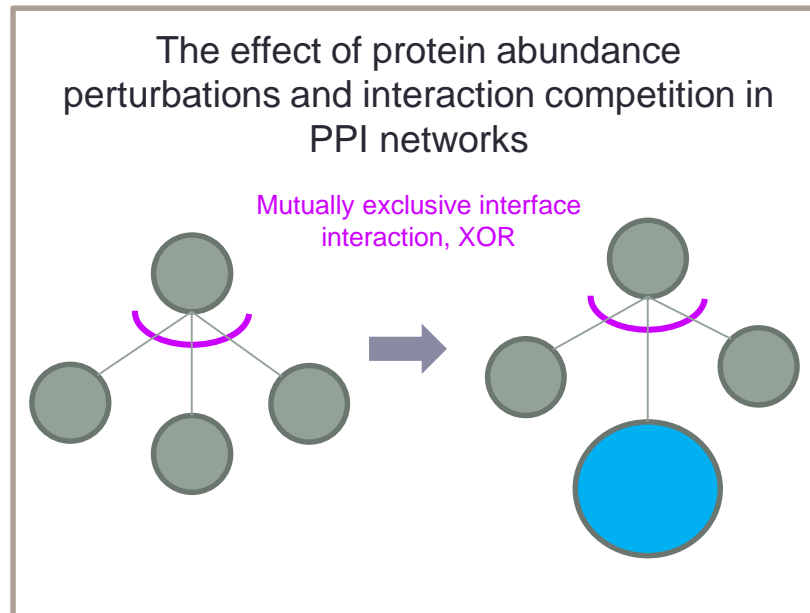
Interaction competition

A means to redirect signaling flows?

CELL BIOLOGY

2013: Signaling Breakthroughs of the Year

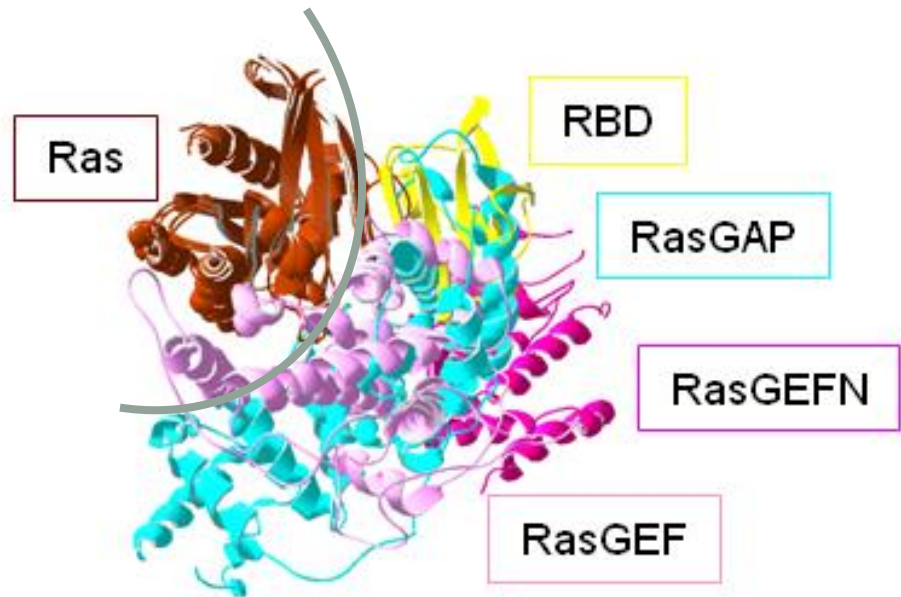
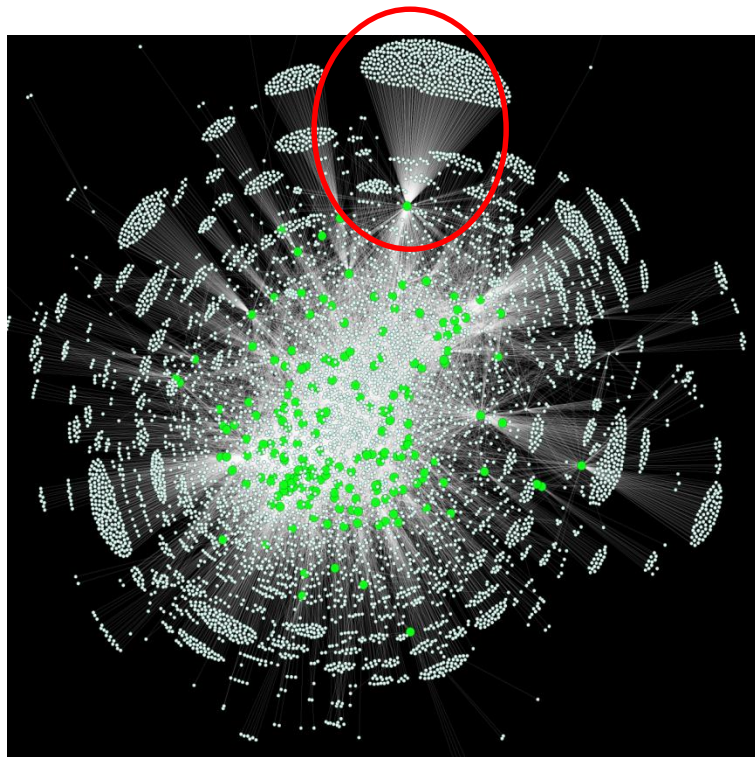
Another mechanism by which hubs decode different inputs is exemplified in a study by Kiel *et al.* (7). The authors showed that competition among proteins binding to a hub can govern how cells produce different responses to the activation of the same receptor. For proteins with a similar affinity for a common hub that bind in a mutually exclusive fashion, the relative abundance of these



How could interaction competition and protein concentration affect downstream signaling?

Signaling complexes: > 300 partners for one protein??

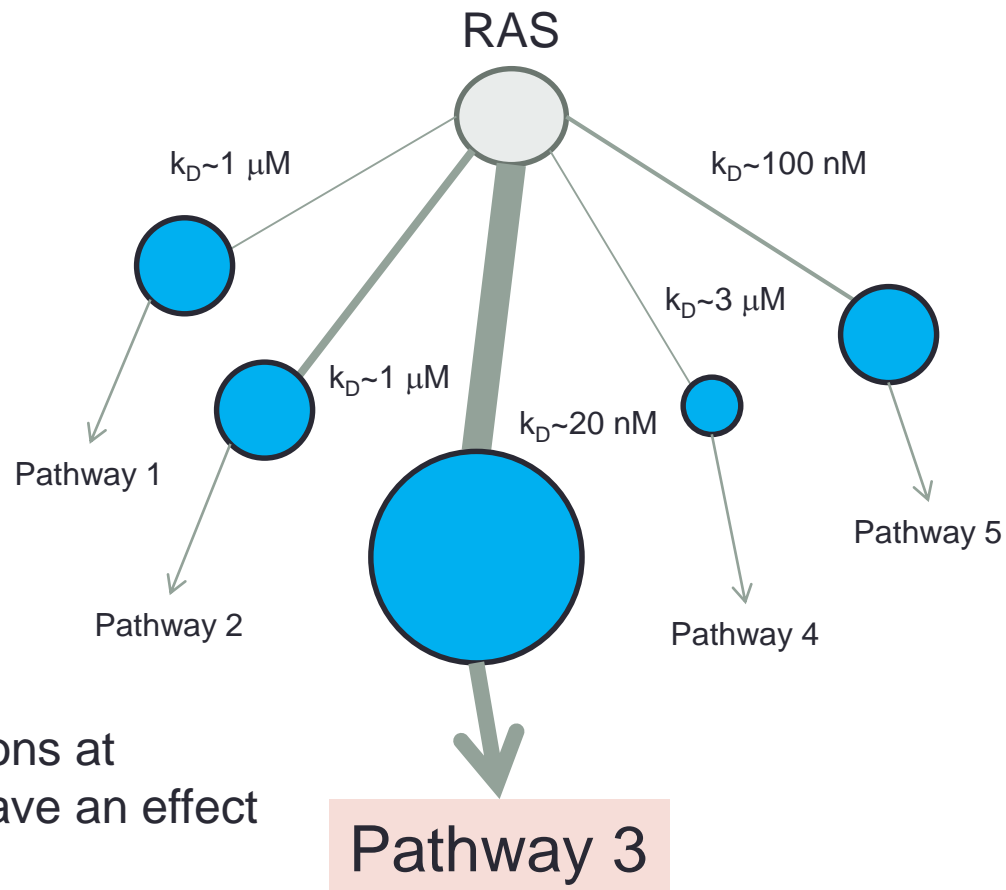
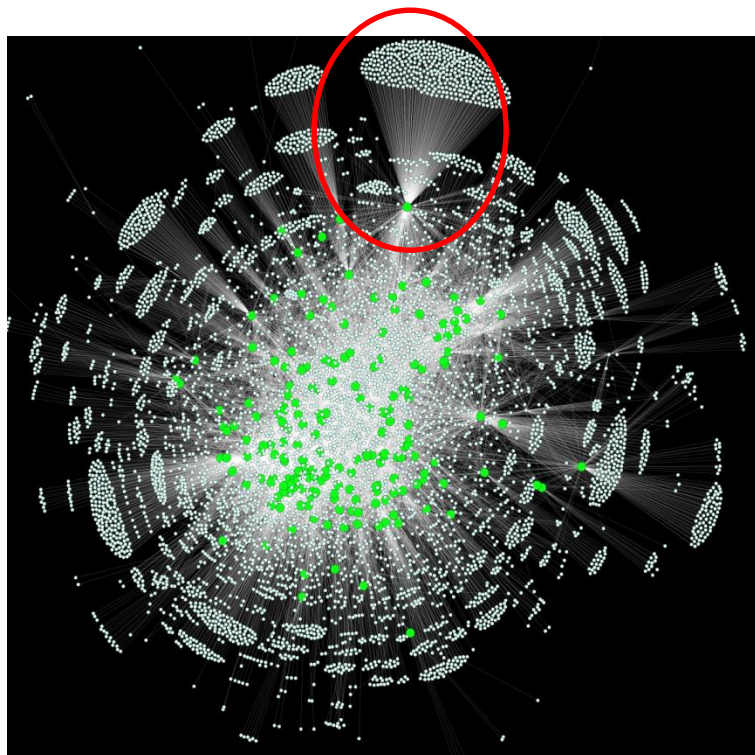
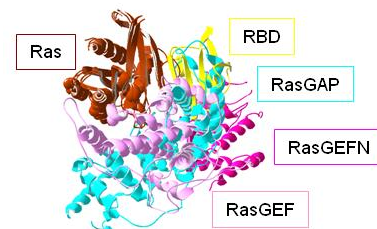
Some proteins will use similar binding surfaces for interaction with other molecules: 'mutually exclusive interactions' / 'XOR'



How could interaction competition and protein concentration affect downstream signaling?

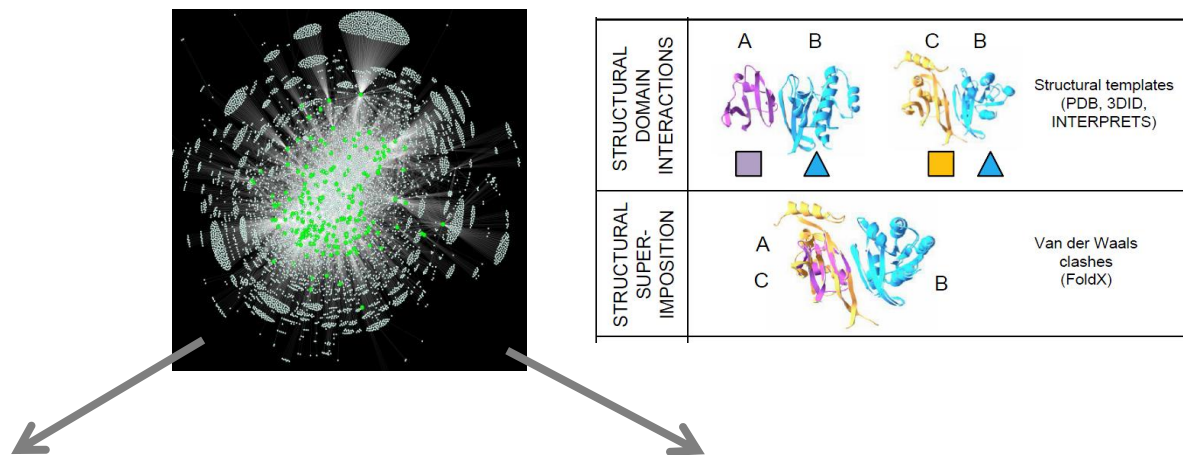
Signaling complexes: > 300 partners for one protein??

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

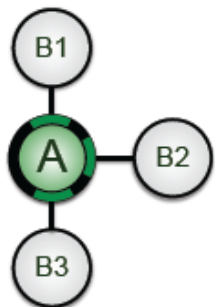


Changes in concentration (ie mutations at promoters, enhancers etc..) could have an effect in signalling

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



Compatible ('AND')

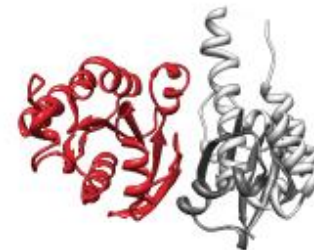
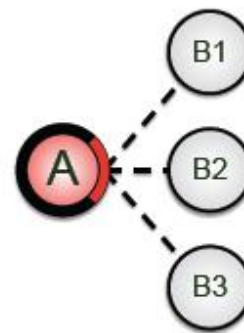


SAPIN (structural analysis of protein interaction networks)

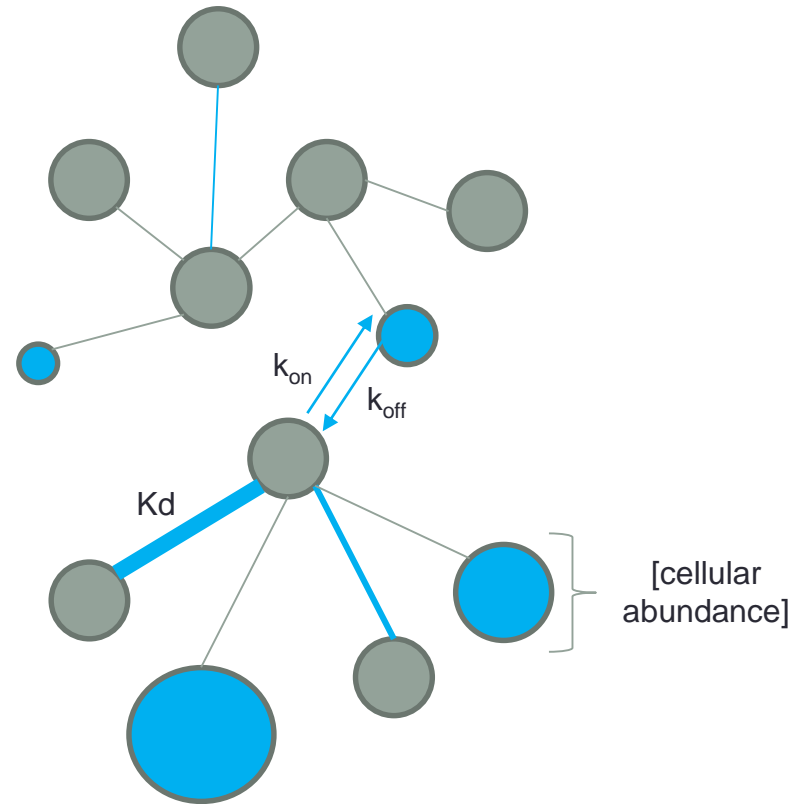
webservice

<http://sapin.crg.es/>

Exclusive ('XOR')



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



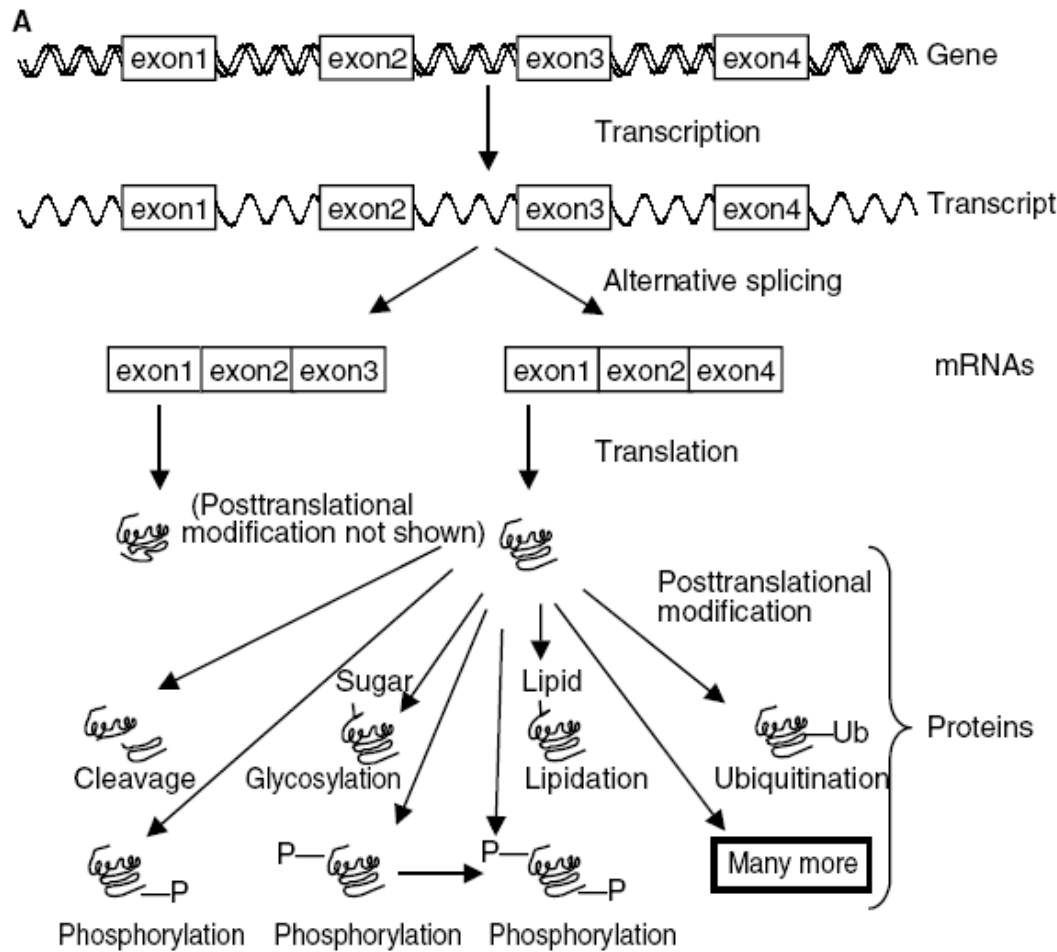
Why proteomics in times of deep RNA sequencing?

- ❑ mRNA does not translate 1:1 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

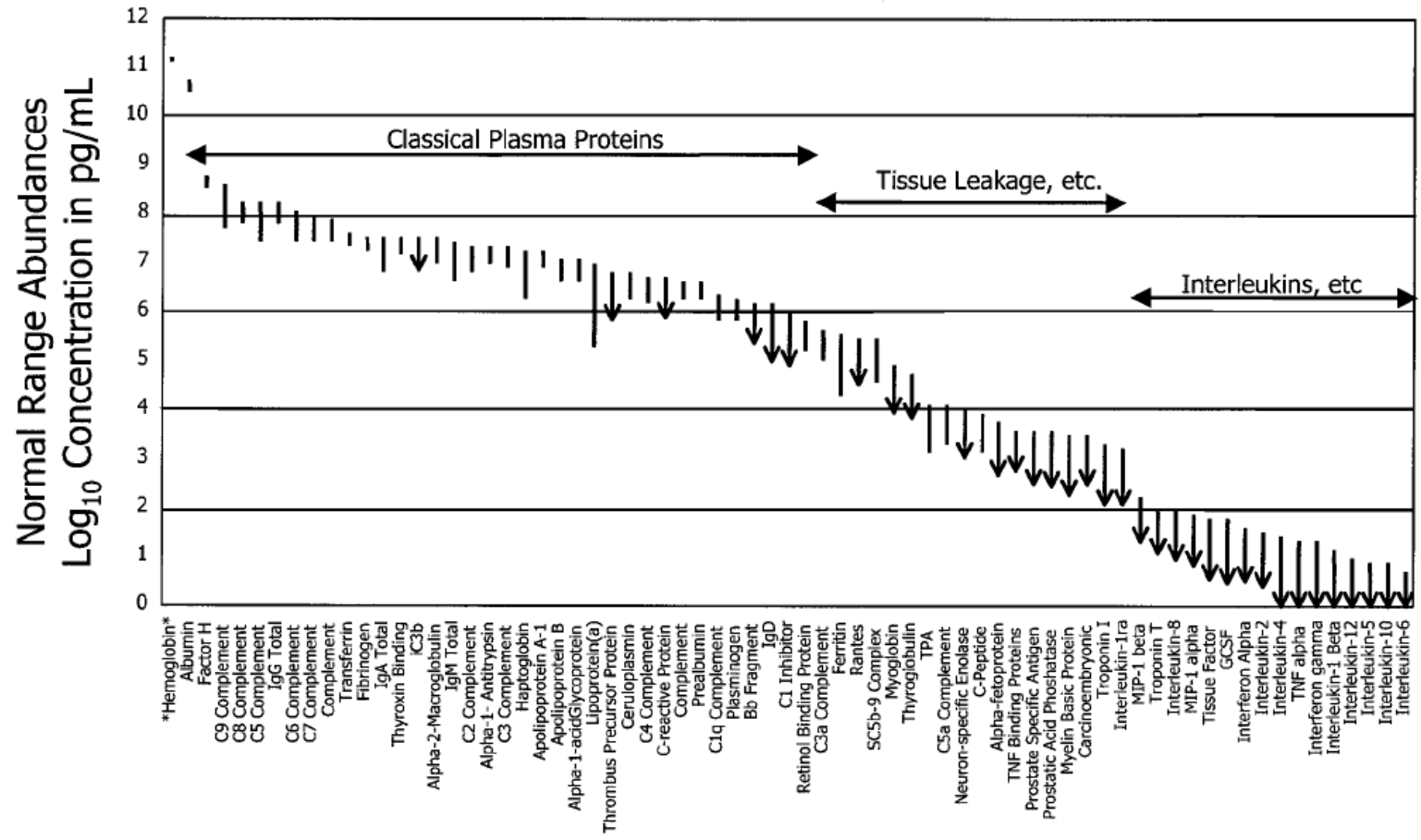


30,000 coding genes per cell

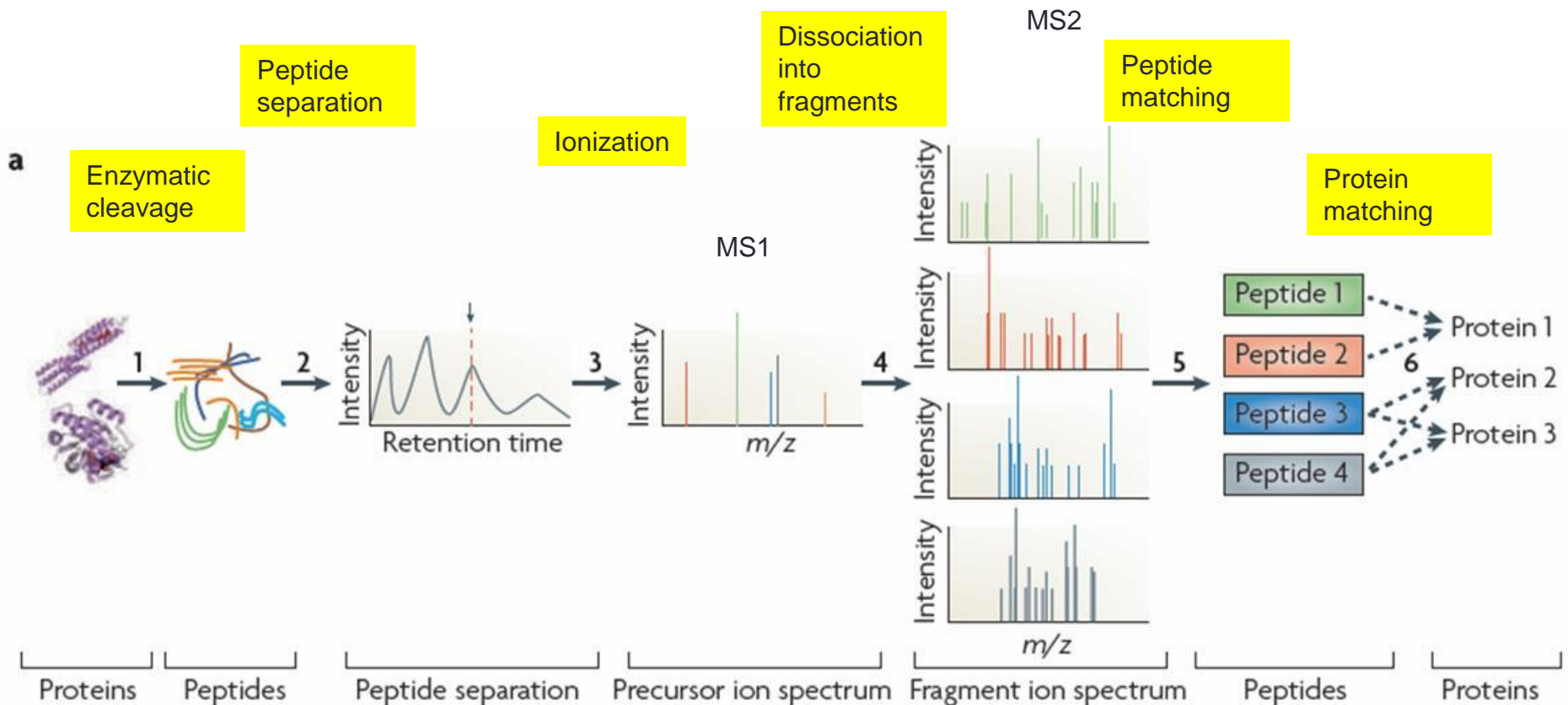
Alt.splicing: 2-3 x 30,000
= 90,000 proteins

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

High dynamic range of the proteome



Protein identification by mass spectrometry



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

Human deep proteome mapping

Molecular Systems Biology 7; Article number 549; doi:10.1038/msb.2011.82
Citation: *Molecular Systems Biology* 7: 549
© 2011 EMBO and Macmillan Publishers Limited All rights reserved 1744-4292/11
www.molecularsystemsbiology.com



REPORT

The quantitative proteome of a human cell line

Martin Beck^{1,9}, Alexander Schmidt^{2,9}, Johan Malmstroem^{3,4}, Manfred Claassen⁵, Alessandro Ori¹, Anna Szymborska¹, Franz Herzog⁶, Oliver Rinner⁴, Jan Ellenberg¹ 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

Molecular Systems Biology 7; Article number 548; doi:10.1038/msb.2011.81
Citation: *Molecular Systems Biology* 7: 548
© 2011 EMBO and Macmillan Publishers Limited All rights reserved 1744-4292/11
www.molecularsystemsbiology.com



REPORT

Deep proteome and transcriptome mapping of a human cancer cell line

Nagarjuna Nagaraj¹, Jacek R Wisniewski¹, Tamar Geiger¹, Juergen Cox¹, Martin Kircher², Janet Kelso², Svante Pääbo² 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

- R. Aebersold lab

~10,000 proteins quantified

Beck et al, MSB, 2011

- M Mann lab

10,255 proteins quantified

Nagaraj et al, MSB, 2011

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

ARTICLE

2014 Pandey lab

doi:10.1038/nature13302

A draft map of the human proteome

Min-Sik Kim^{1,2}, Sneha M. Pinto³, Derese Getnet^{1,4}, Raja Sekhar Nirujogi¹, Srikanth S. Manda¹, Raghothama Chaerkady^{1,2}, Anil K. Madugundu¹, Dhanashree S. Kelkar¹, Ruth Isserlin⁵, Shobhit Jain⁵, Joji K. Thomas¹, Babyakshmi Muthusamy¹, Pamela Leal-Rojas^{1,6}, Praveen Kumar¹, Nandini A. Sahasrabudde¹, Lavanya Balakrishnan¹, Jayshree Advani¹, Bijesh George¹, Santosh Renuse¹, Lakshmi Dhevi N. Selvan¹, Arun H. Patil¹, Vishalakshi Nanjappa¹, Aneesha Radhakrishnan¹, Samarjeet Prasad¹, Tejaswini Subbannayya¹, Rajesh Raju¹, Manish Kumar¹, Sreelakshmi K. Sreenivasamurthy¹, Arivusudar Marimuthu¹, Gajanan J. Sathe¹, Sandip Chavan¹, Keshava K. Datta¹, Yashwanth Subbannayya¹, Apeksha Sahu¹, Soujanya D. Yelamanchi¹, Savita Jayaram¹, Pavithra Rajagopalan¹, Jyoti Sharma¹, Krishna R. Murthy¹, Nazia Syed¹, Renu Goel¹, Aafaque A. Khan¹, Sartaj Ahmad¹, Gourav Dey¹, Keshav Mudgal¹, Aditi Chatterjee¹, Tai-Chung Huang¹, Jun Zhong¹, Xinyan Wu^{1,2}, Patrick G. Shaw¹, Donald Freed¹, Muhammad S. Zahar¹, Kanchan K. Mukherjee¹, Subramanian Shankar¹, Anita Mahadevan^{10,11}, Henry Lam¹², Christopher J. Mitchell¹, Susarla Krishna Shankar^{10,11}, Parthasarathy Satishchandra¹³, John T. Schroeder¹⁴, Ravi Sirdeshmukh¹, Anirban Maitra^{15,16}, Steven D. Leach^{1,17}, Charles G. Drake^{16,18}, Marc K. Halushka¹⁵, T. S. Keshava Prasad¹, Ralph H. Hruban^{15,16}, Candace L. Kerr¹⁹, Gary D. Bader⁵, Christine A. Iacobuzio-Donahue^{15,16,17}, Harsha Gowda¹ & Akhilesh Pandey^{1,2,3,4,15,16,20}

ARTICLE

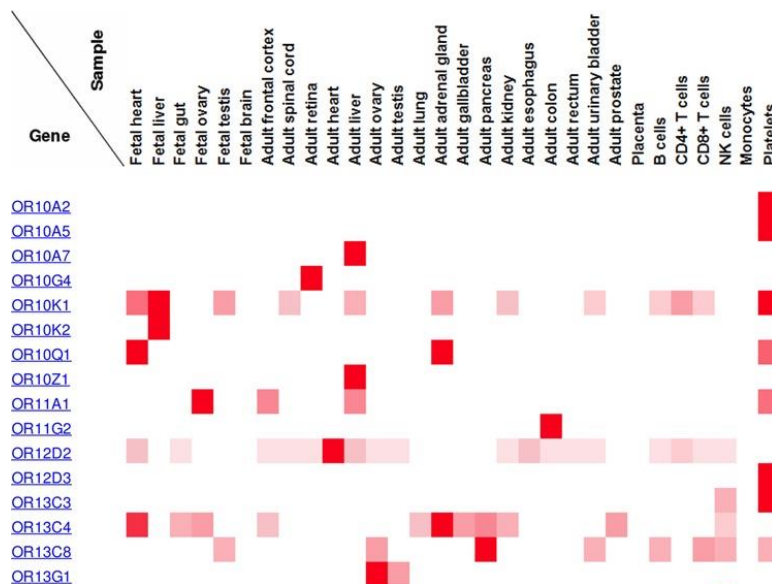
2014 Kuster lab

doi:10.1038/nature13319

Mass-spectrometry-based draft of the human proteome

Mathias Wilhelm^{1,2*}, Judith Schlegli^{2*}, Hannes Hahne^{1*}, Amin Moghaddas Gholami^{1*}, Marcus Lieberenz², Mikhail M. Savitski¹, Emanuel Ziegler², Lars Butzmann², Siegfried Gessulat², Harald Marx¹, Toby Mathieson¹, Simone Lemeer¹, Karsten Schnatbaum¹, Ulf Reimer¹, Holger Wenschuh¹, Martin Mollenhauer¹, Julia Slotta-Huspenina¹, Joos-Hendrik Boese², Marcus Bantscheff¹, Anja Gerstmaier², Franz Faerber² & Bernhard Kuster^{1,6}

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



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

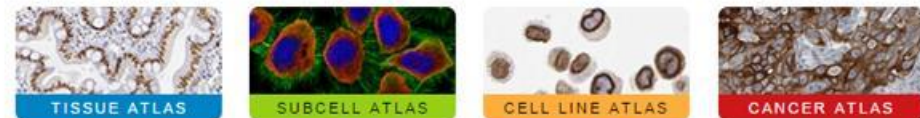
THE HUMAN PROTEIN ATLAS

ABOUT & HELP

A Tissue-Based Map of the Human Proteome

Here, we summarize our current knowledge regarding the human proteome mainly achieved through antibody-based methods combined with transcriptomics analysis across all major tissues and organs of the human body. A large number of lists can be accessed with direct links to gene-specific images of the corresponding proteins in the different tissues and organs.

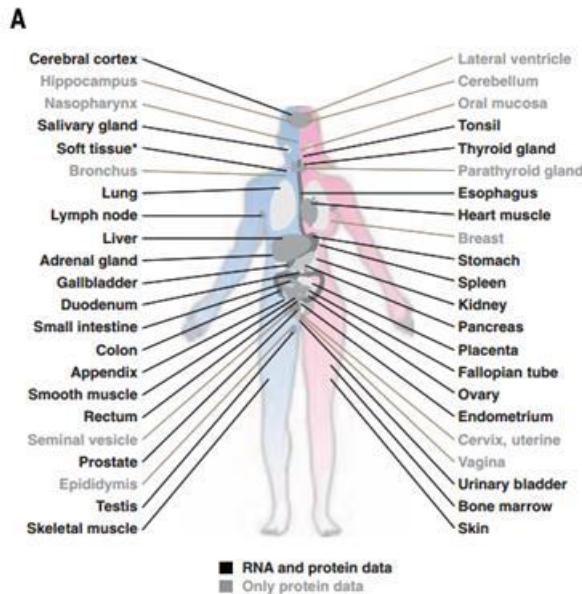
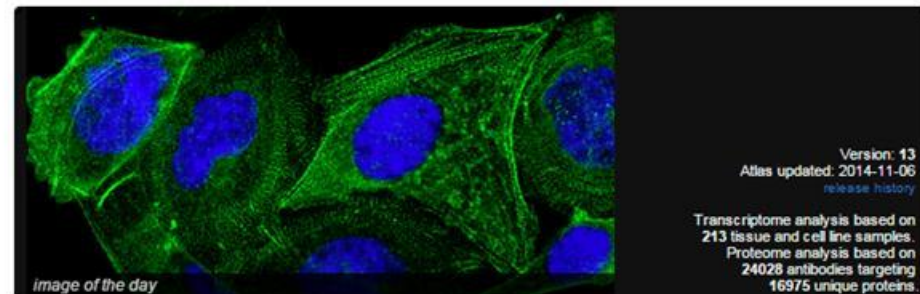
[Read more](#)



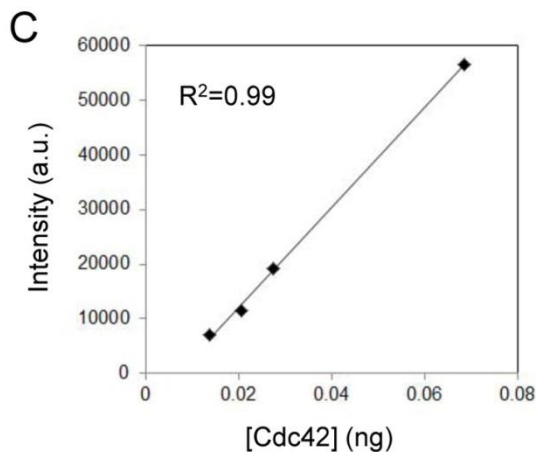
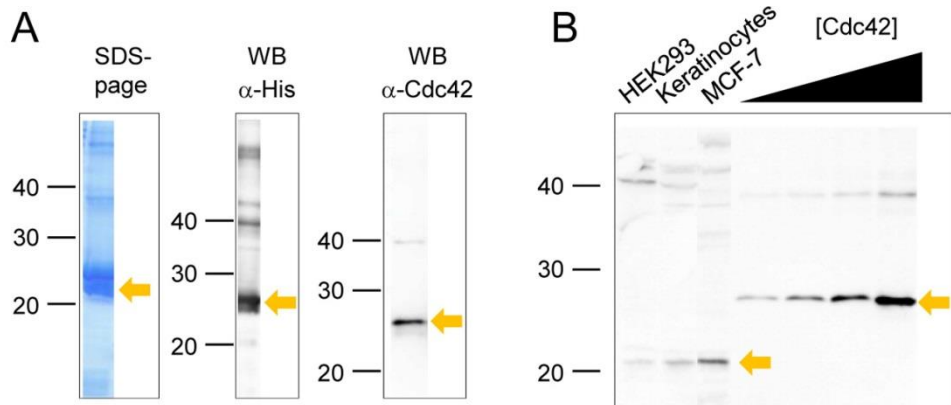
SEARCH ? »

e.g. insulin, PGR, CD36

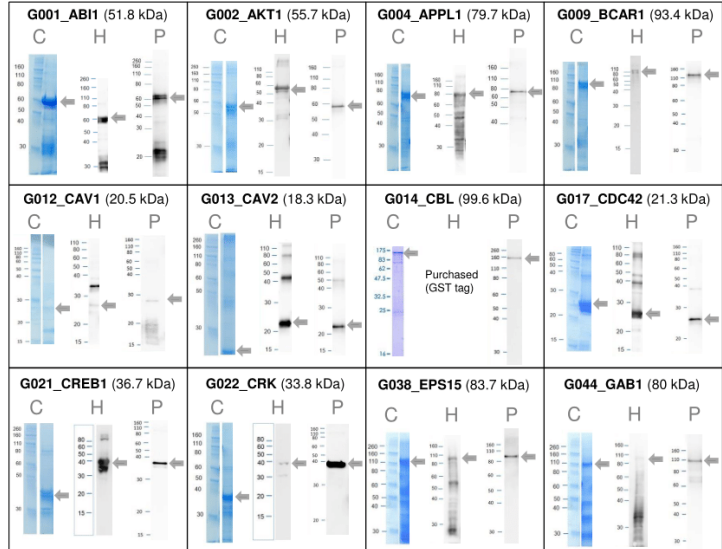
Search Fields »



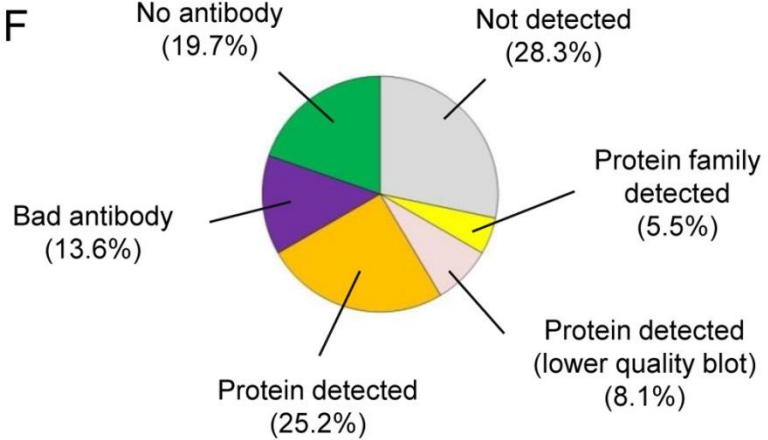
Quantitative Western blotting



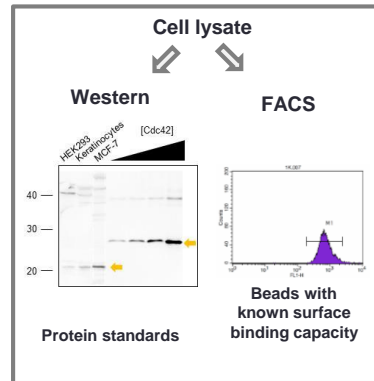
Protein standards: expression, purification and quantification



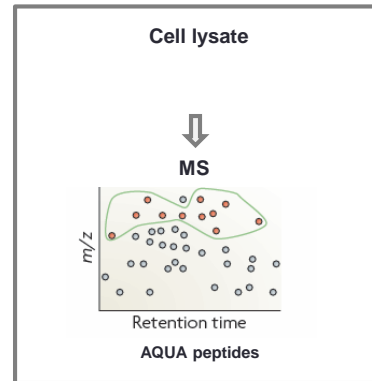
Summary statistic for quantitative Western blotting of 198 ErbB-related proteins



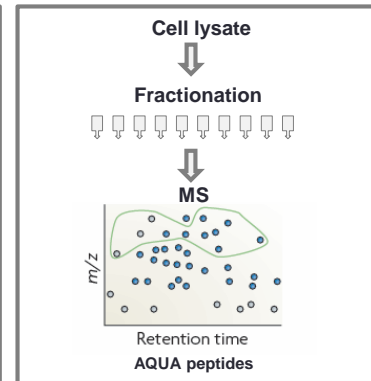
Combining different quantitative approaches to quantify 198 proteins in the ErbB signaling pathway



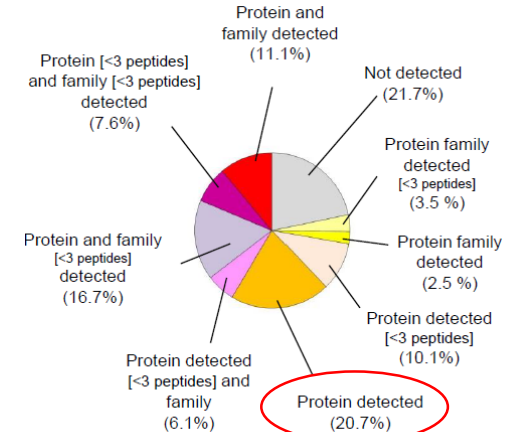
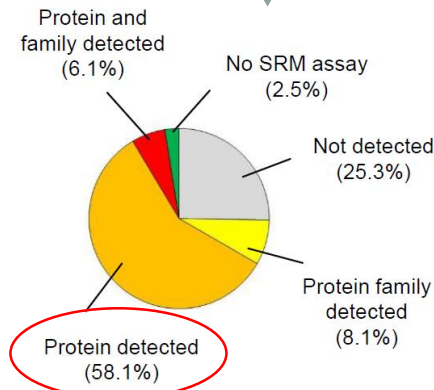
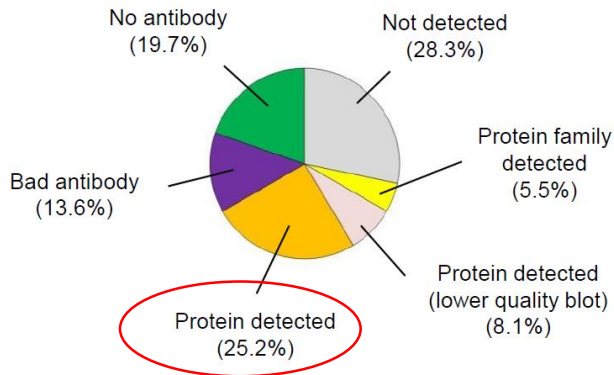
Quantitative Western blotting and quantitative FACS



Targeted mass spectrometry (MS)



Fractionation + shot-gun mass spectrometry (MS)

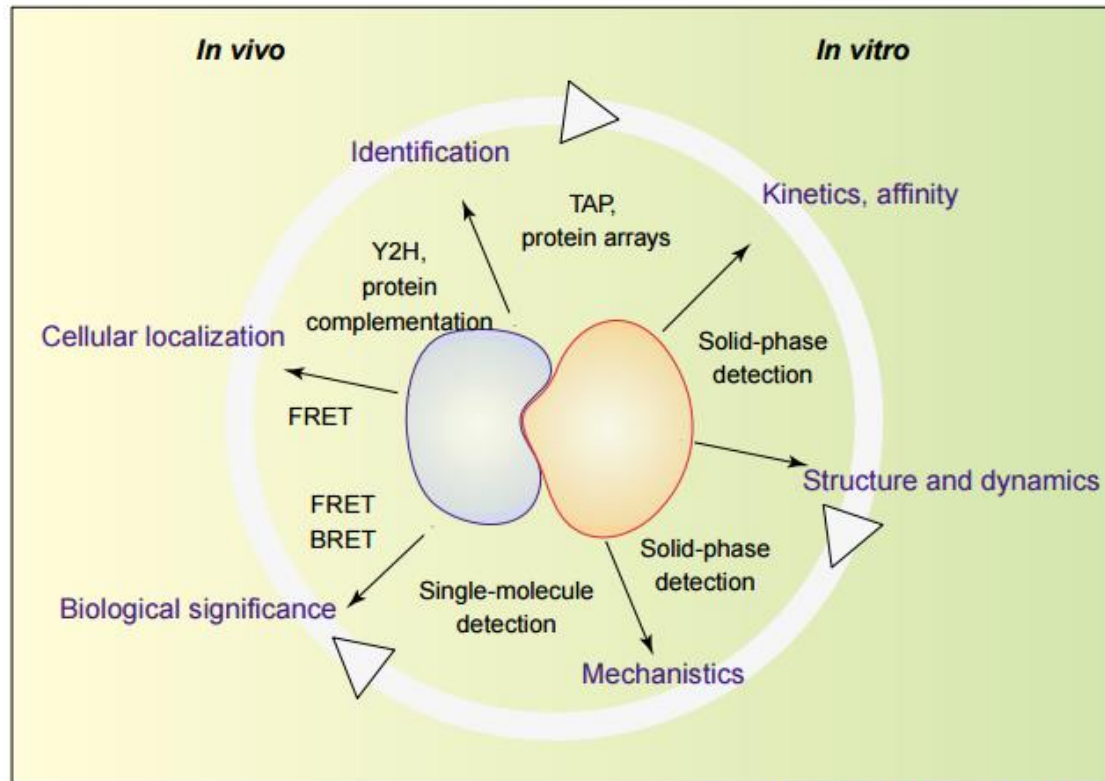


- SRM has a higher sensitivity compared to quantitative western blotting (but some proteins are only detected by Western blotting)
- Problem with isoforms and protein families: as a consequence of frequent gene duplication events in mammals, often similar proteins (e.g. AKT1 and AKT2) cannot be distinguished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

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

The challenge:

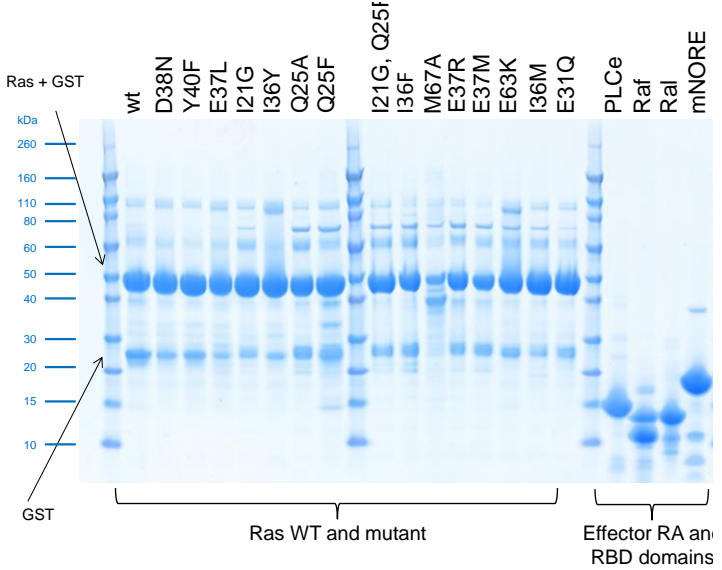
- most *in vivo* techniques are high-throughput, but do not provide affinities (only qualitative binding detection)
- *in vitro* techniques can provide affinities and kinetic constants, but are not high-throughput methods



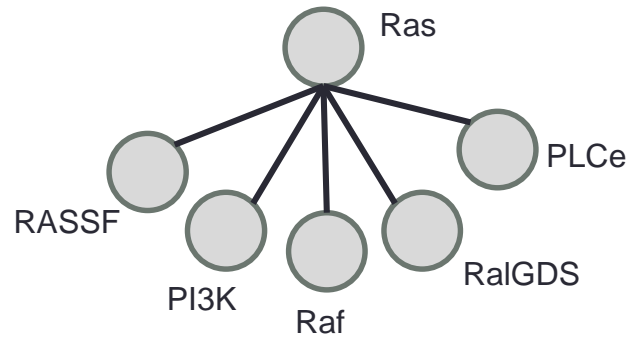
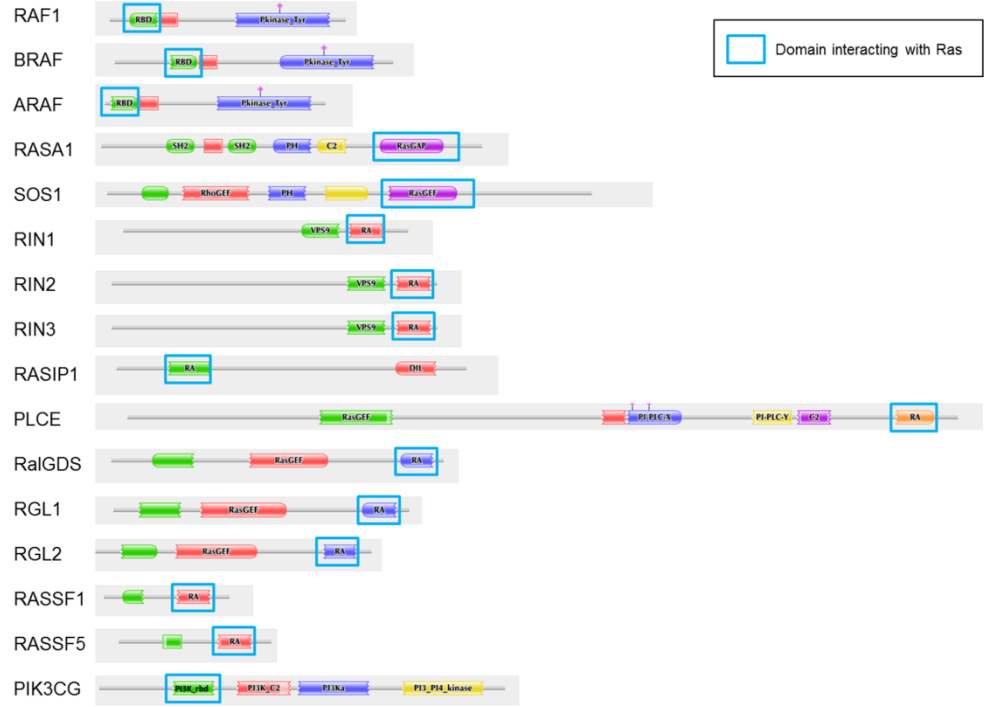
Current Opinion in Structural Biology

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

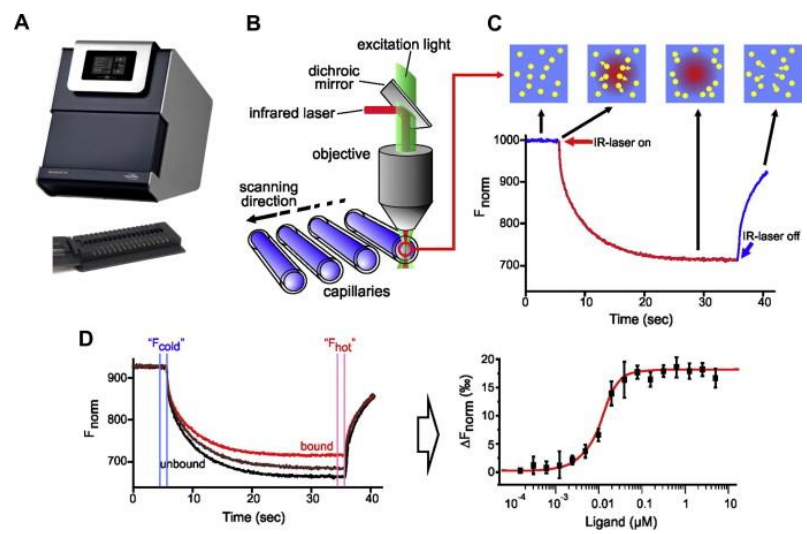


Large proteins are often not soluble: expression and purification of protein domains

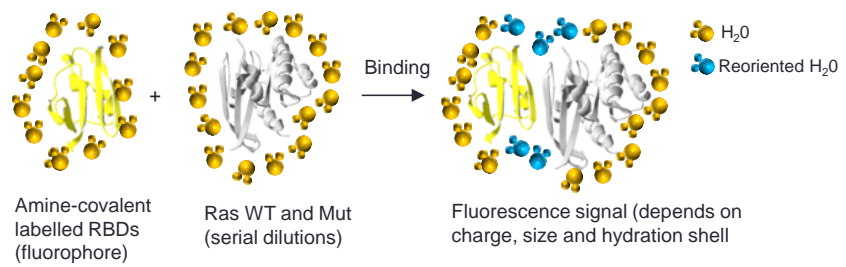


Two main methods to measure affinities and kinetic constants

Microscale thermophoresis



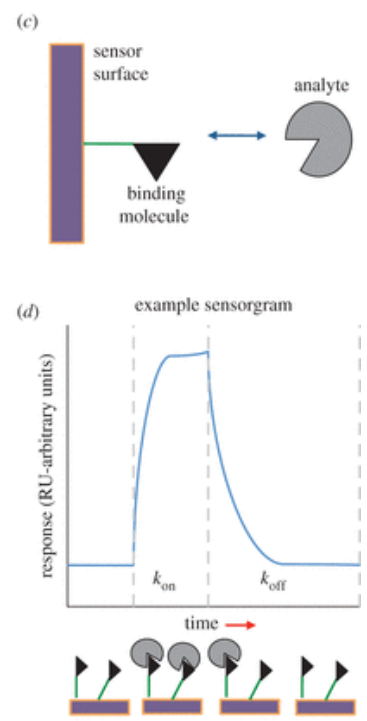
Jerabek-Willemsen et al, J Mol Struct, 2014



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

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

Surface plasmon resonance



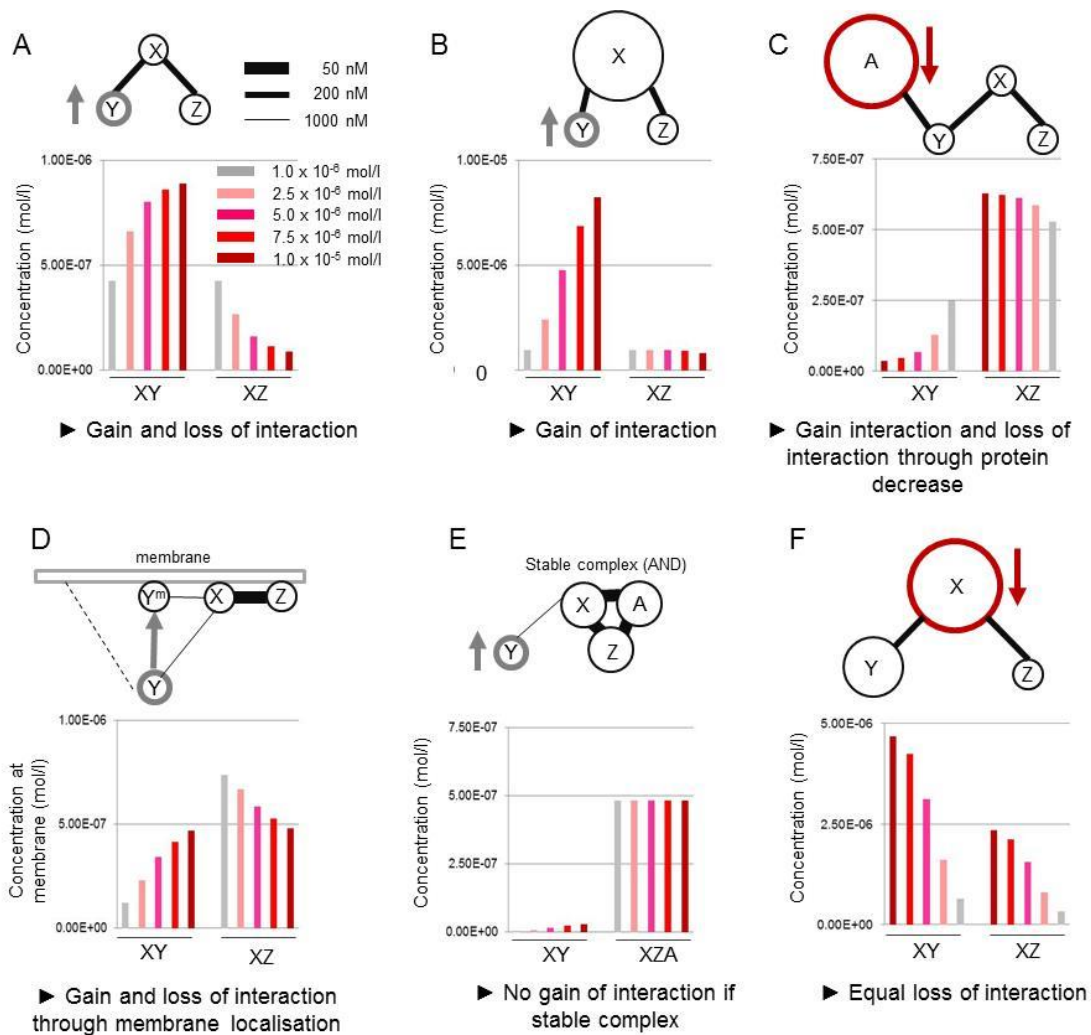
Optical method to measure the refractive index near a sensor surface

Kastritis et al, 2012

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

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

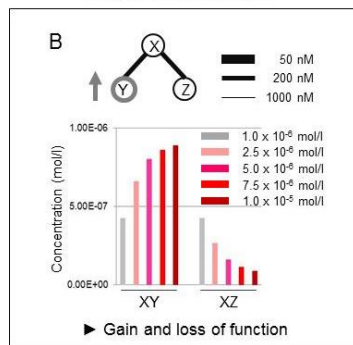
The effect of abundance variation at XOR network motifs



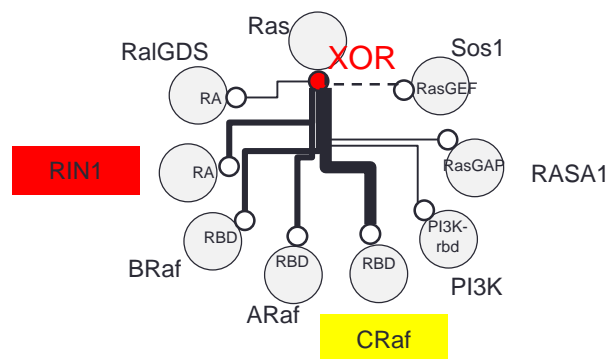
➤ The output/ function depends on both, network structure and abundance: we need to know the network very well to understand

Competition at the Ras XOR node

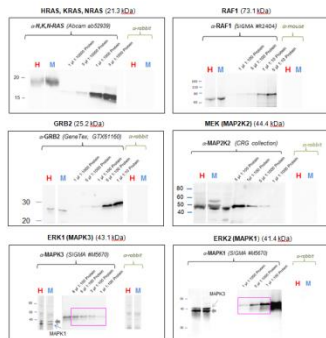
Network motif



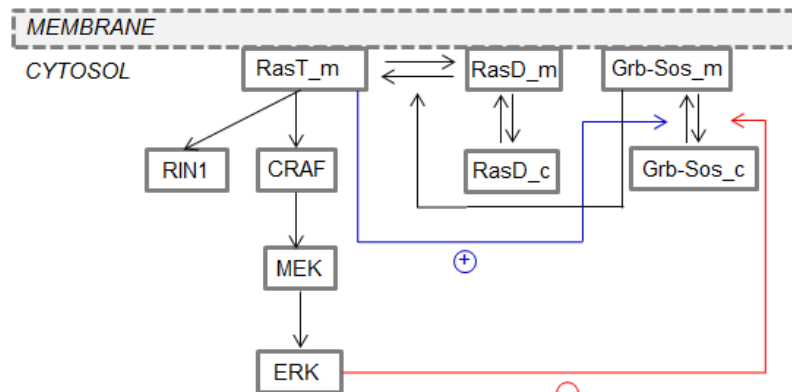
The Ras XOR node



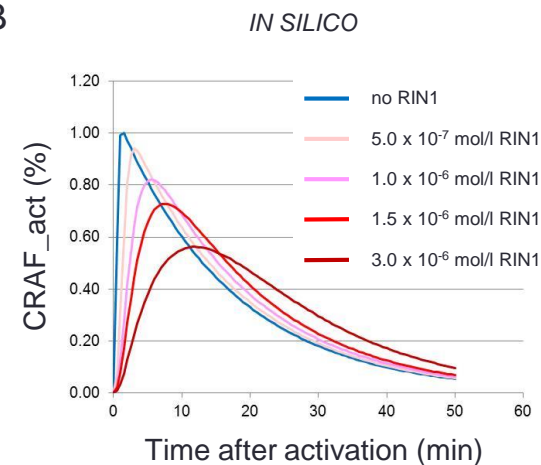
Experimental abundances



A

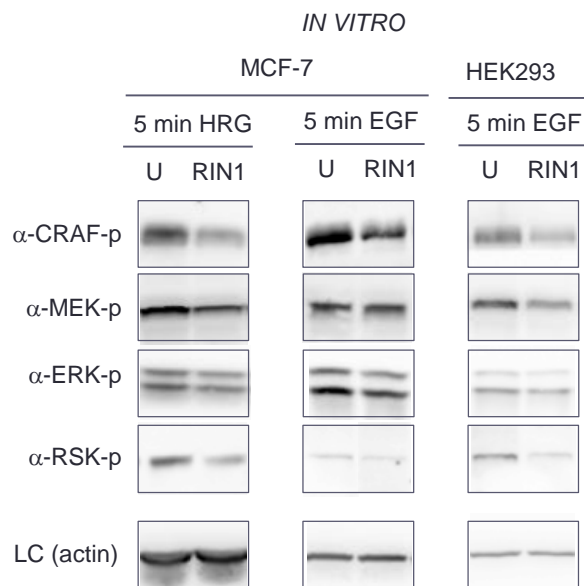


B

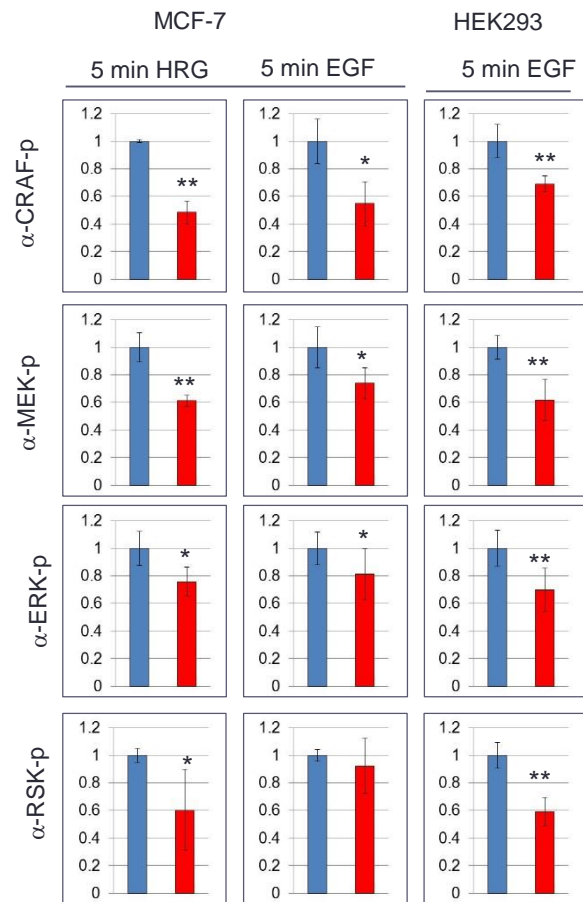


Mathematical network modeling: increasing RIN1 to 10-fold higher of CRAF expression should decrease CRAF activation

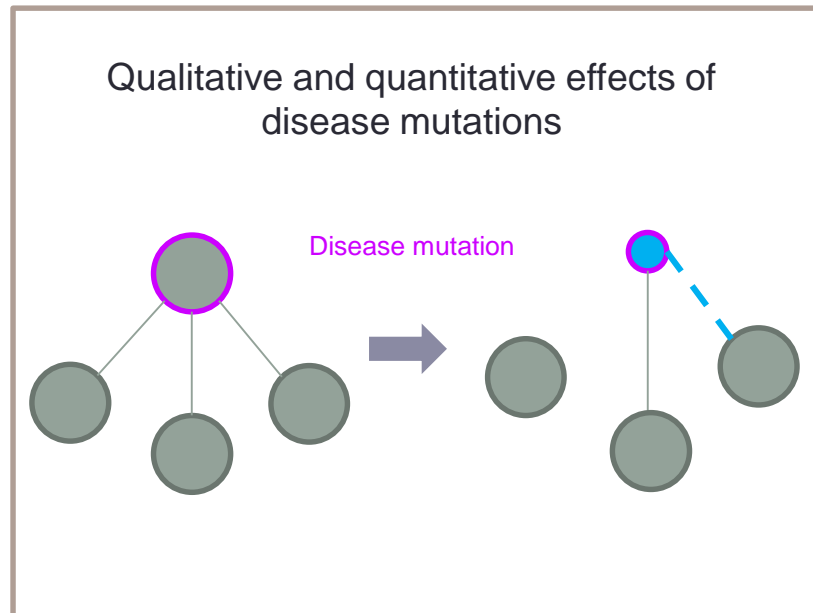
Experimental testing of competition at the Ras node



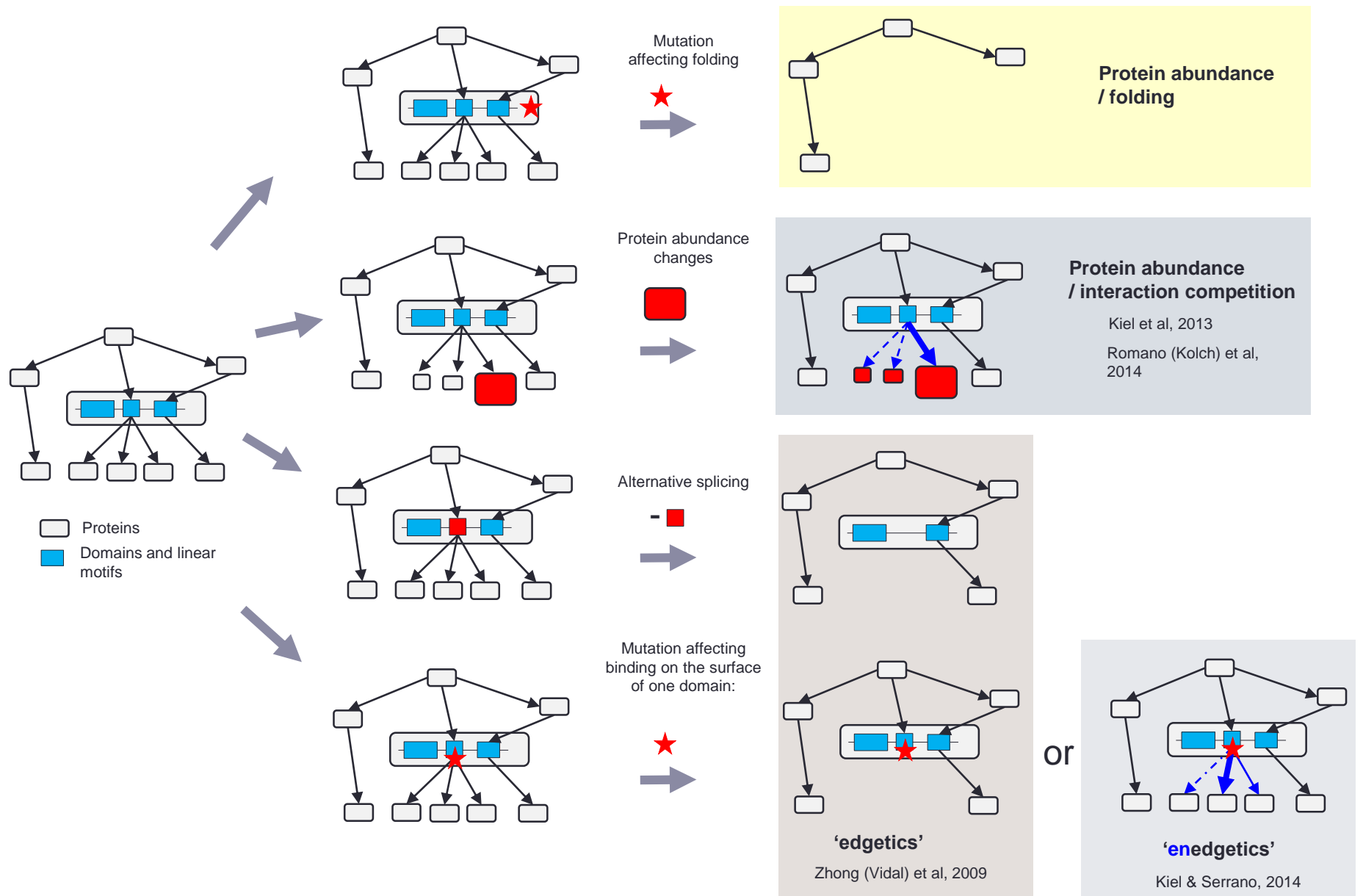
Expression of RIN1 in MCF-7 and HEK293 cells decreases CRAF, MEK, and ERK activation



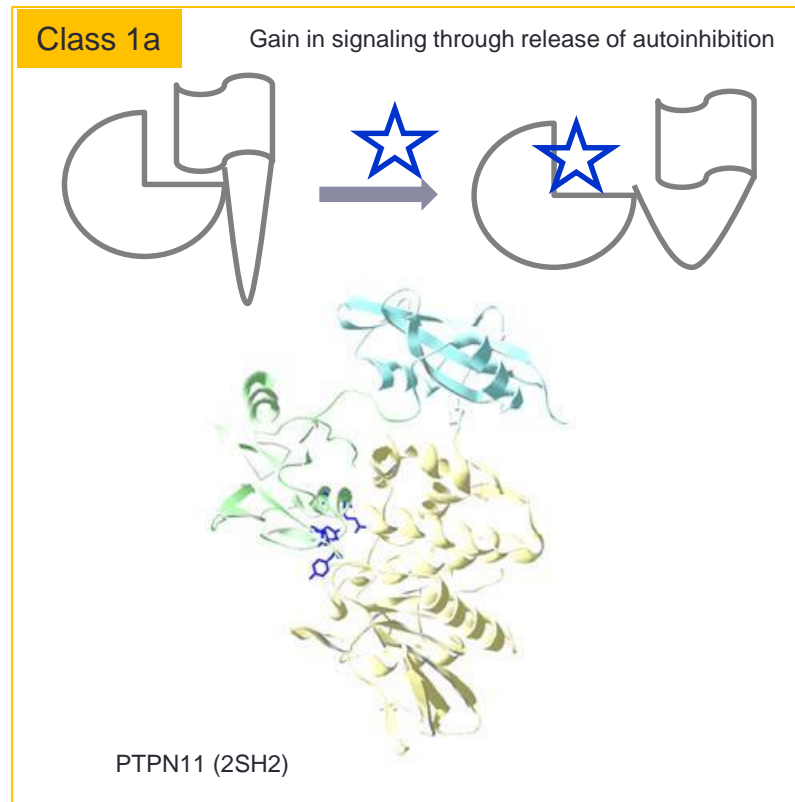
- Alterations in the abundance of one of two hub-binding partners affected downstream signaling



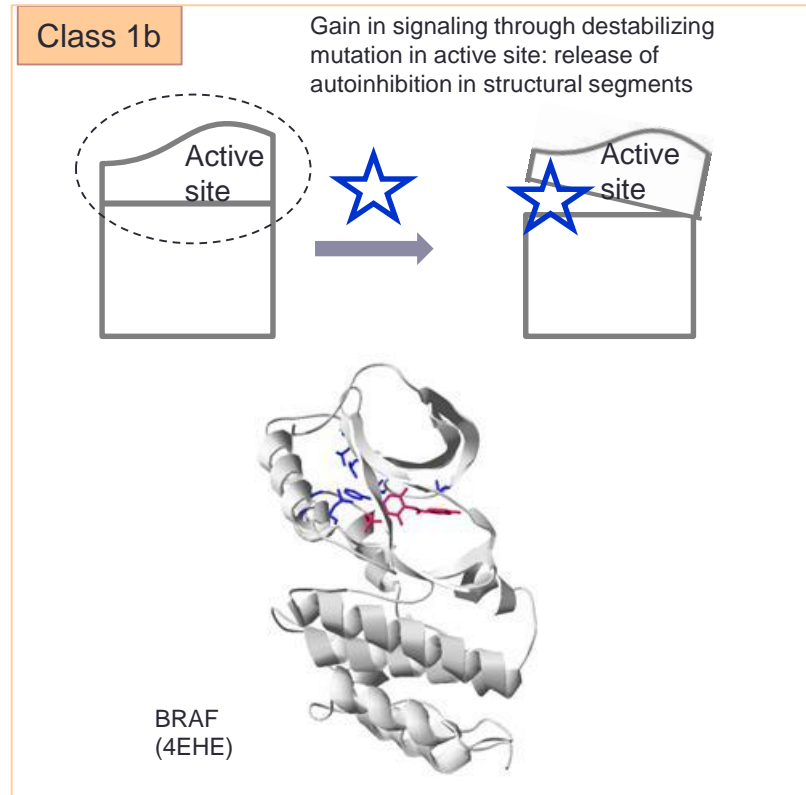
General concepts of interaction ('edge') rewiring



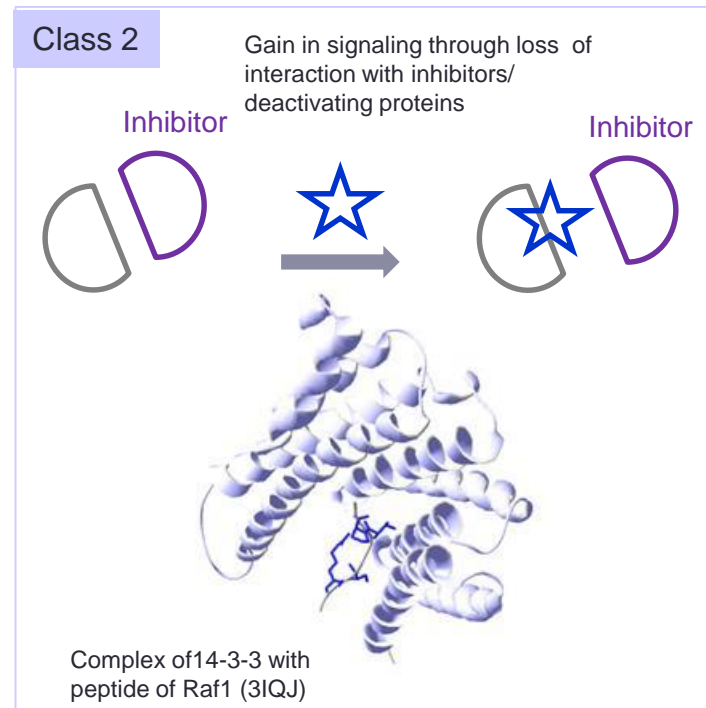
Examples how missense mutations can affect the network: a 3D structural perspective



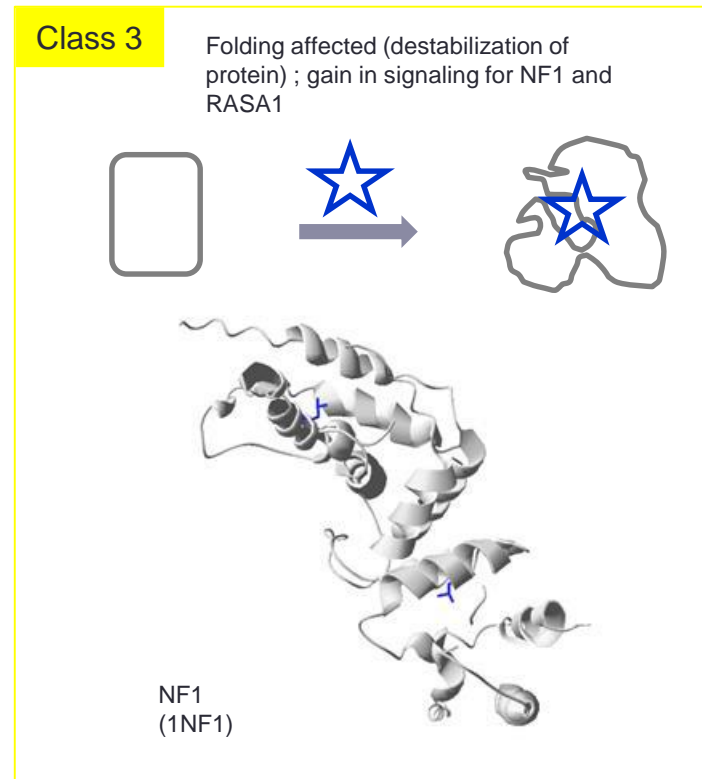
Examples how missense mutations can affect the network: a 3D structural perspective



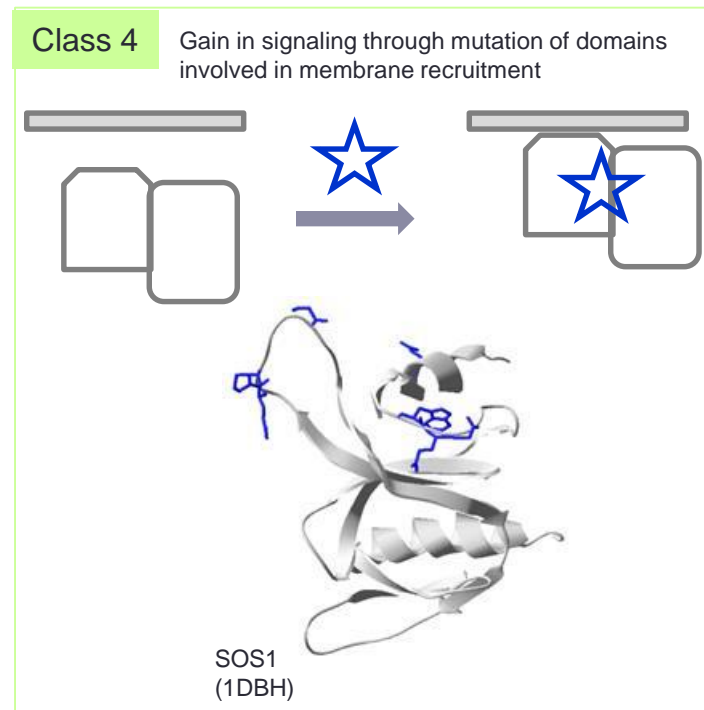
Examples how missense mutations can affect the network: a 3D structural perspective



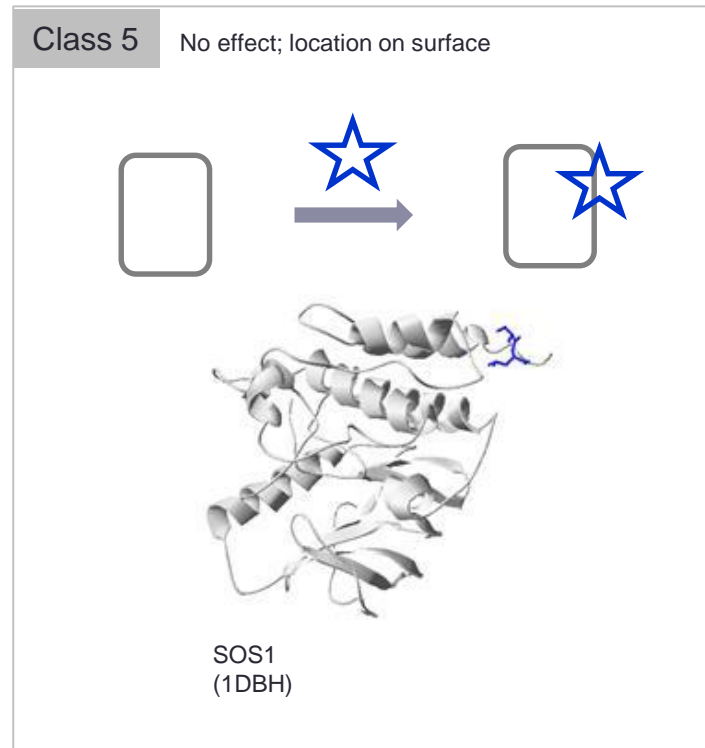
Examples how missense mutations can affect the network: a 3D structural perspective



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



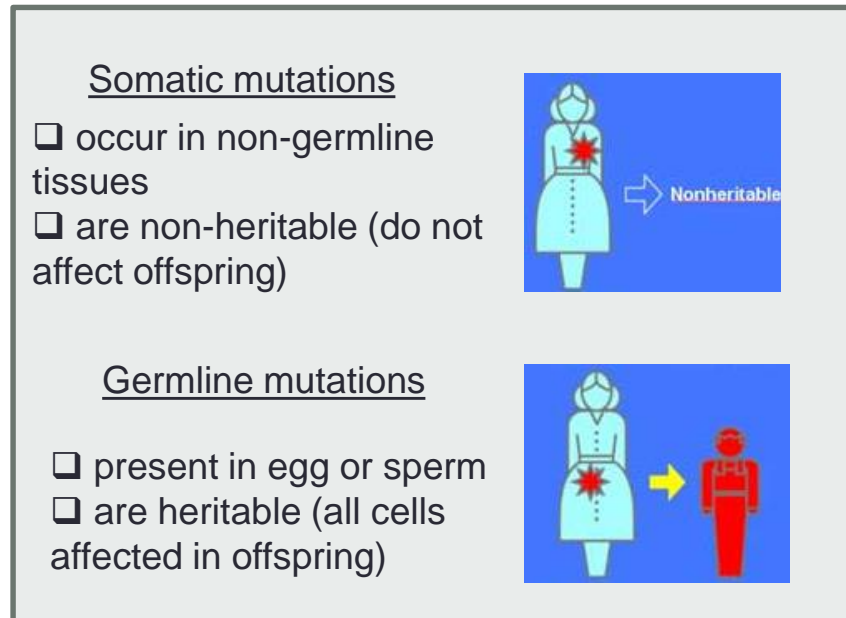
Examples how missense mutations can affect the network: a 3D structural perspective



Example 1: RASopathy and cancer disease mutations

RASopathies: Developmental syndromes of Ras/ MAPK pathway dysregulation

- RASopathies are a group of developmental disorders characterized by postnatal reduced growth facial dysmorphism, cardiac defects, mental retardation, skin defects, musculo-skeletal defects, short stature, cryptorchidism
- RASopathies are caused by germline mutations in genes that encode protein components of the Ras/ 12 proteins involved (HRAS, NF1, MAP2K1, MAP2K2, RASA1, SPRED1, SOS1, PTPN11, RAF1, KRAS, NRAS, BRAF)
- majority of mutations result in increased signal transduction down the Ras/MAPK pathway, but usually to a smaller extent than somatic mutations associated with cancer



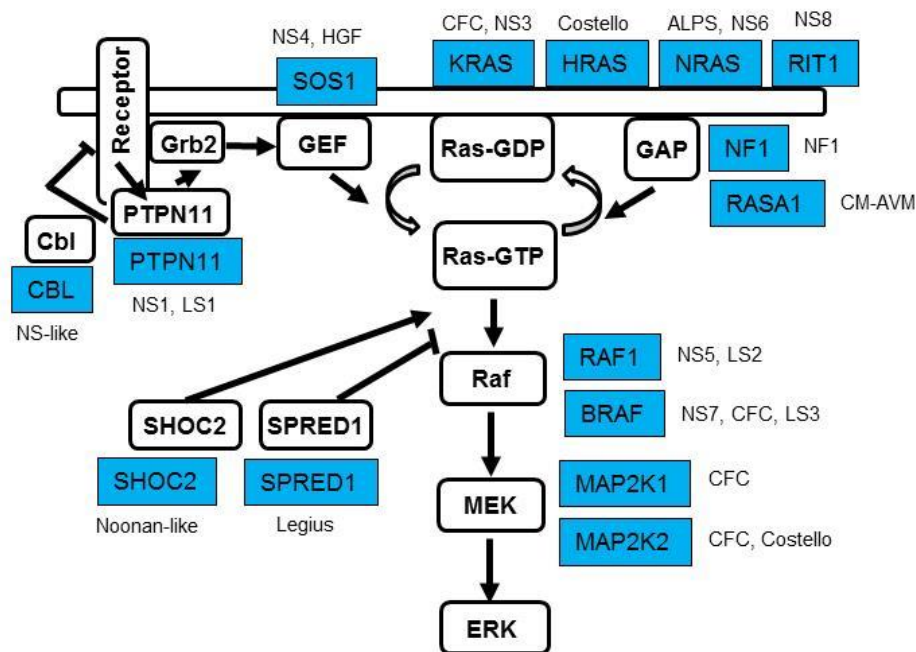
Christina
Kiel



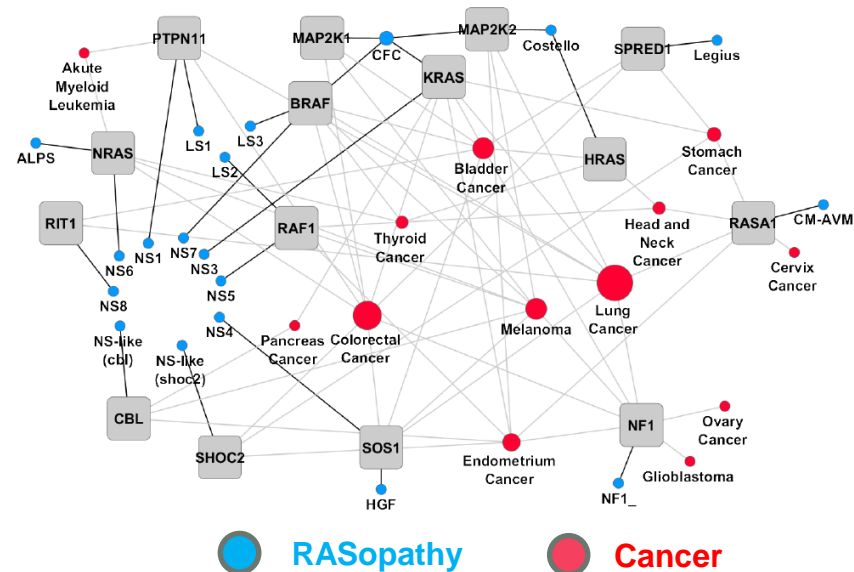
Hannah
Benisty

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

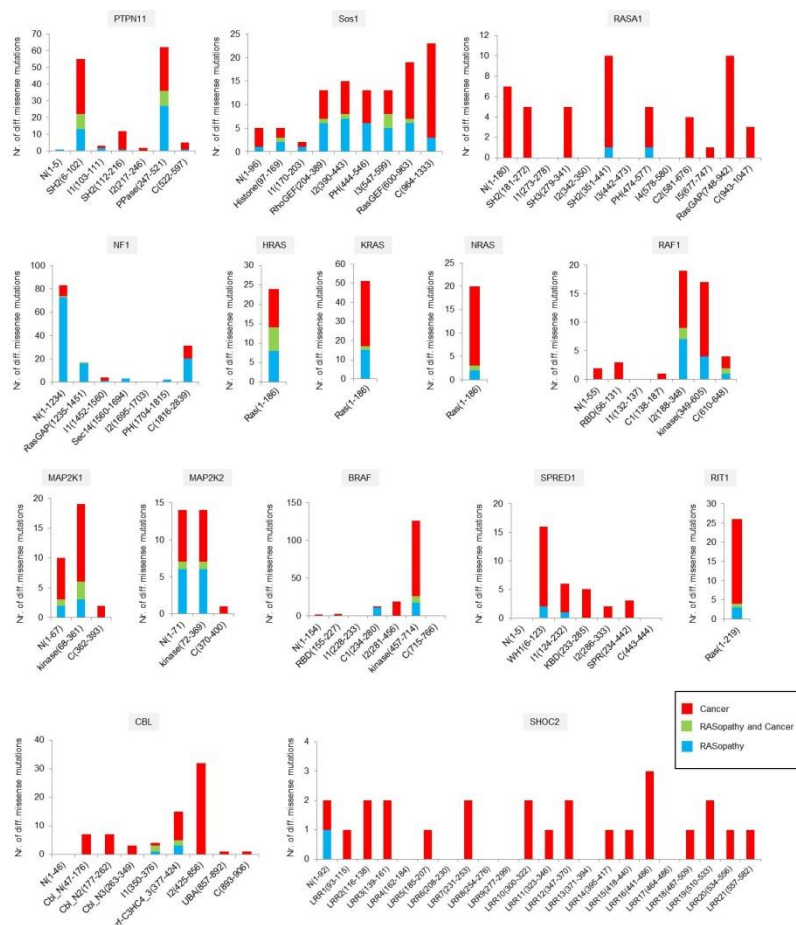


- Proteins in Ras/MAPK syndromes ('RASopathies') are also found in cancer

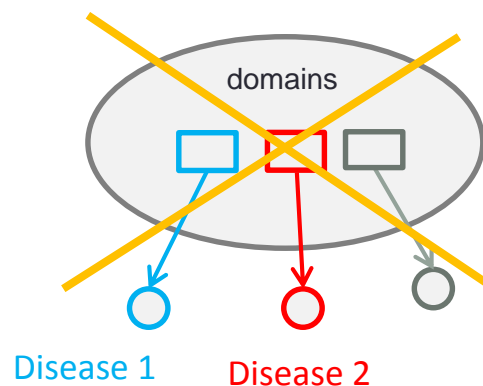
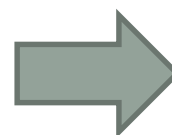


Location of mutations in different domains does not explain the difference between RASopathy and cancer mutations

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



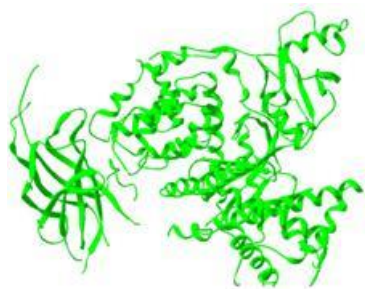
Disease 1 Disease 2

FoldX-based energy calculations of proteins

3D Structural information

A force field for energy calculations and protein design

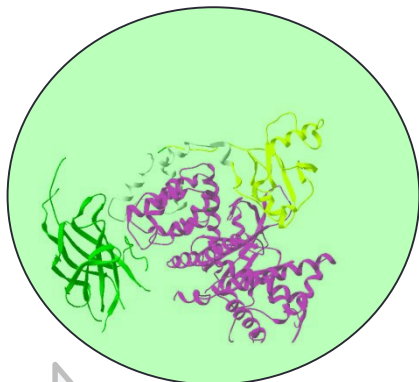
Schymkowitz et al, *Nucleic Acids Res*, 2005



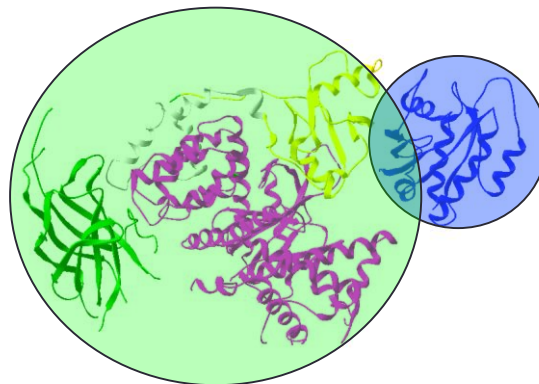
$$+ \text{FoldX} = \Delta G$$

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

✓ Total free energy

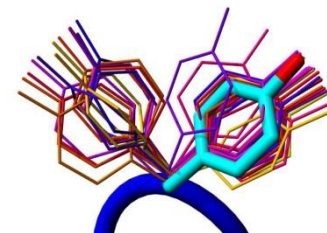


✓ Interaction energy



✓ Mutagenesis

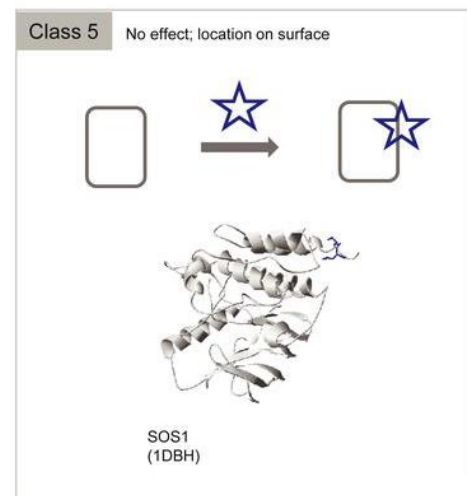
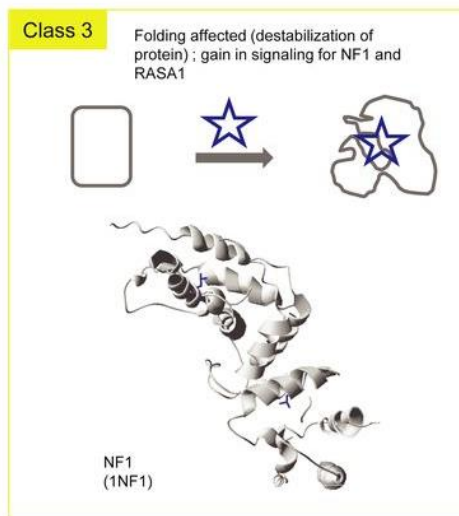
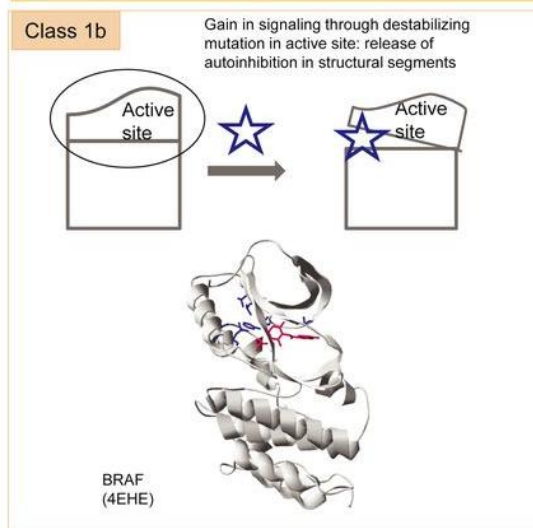
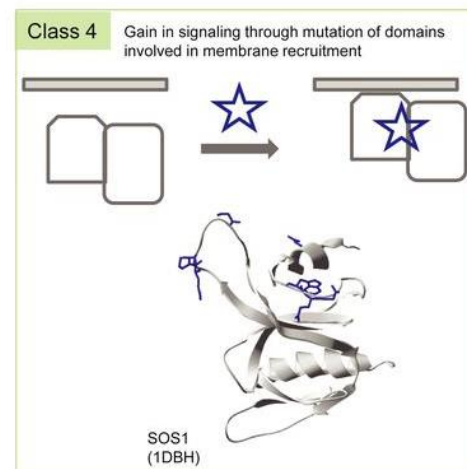
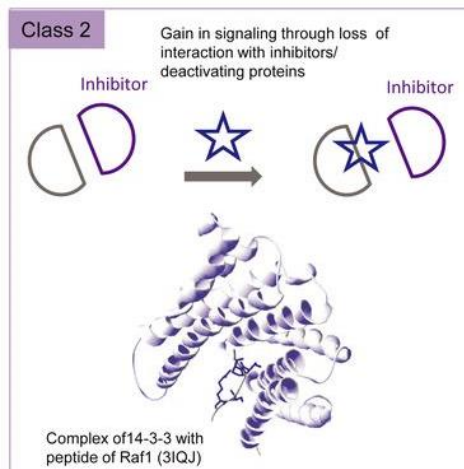
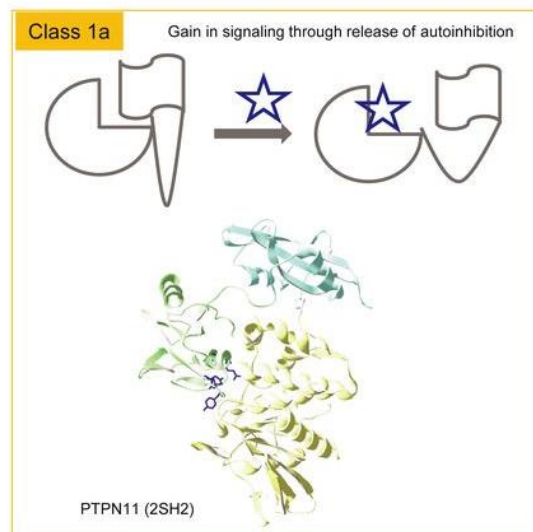
A rotamer library to replace the 20 amino acids



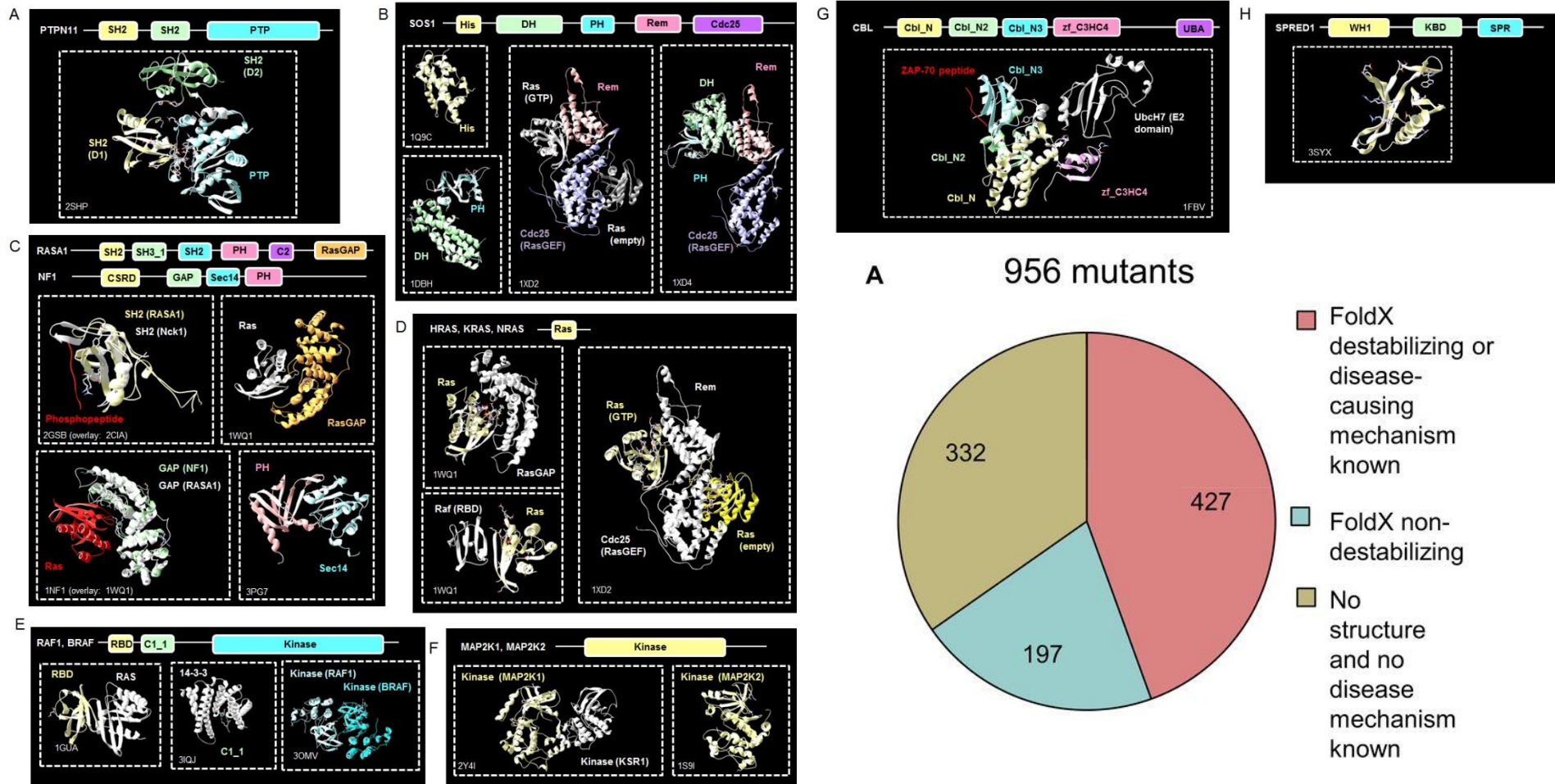
Protein design

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

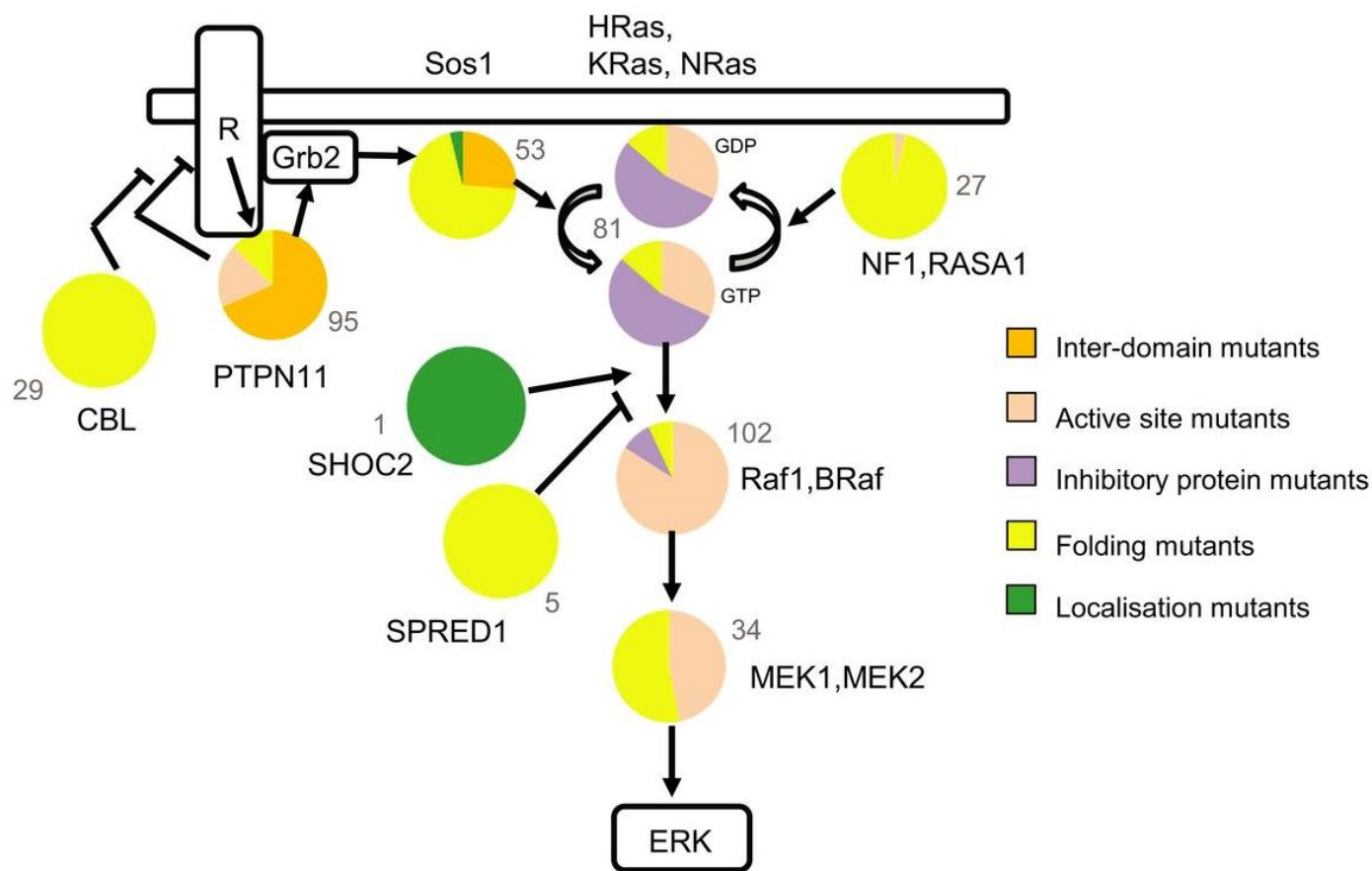
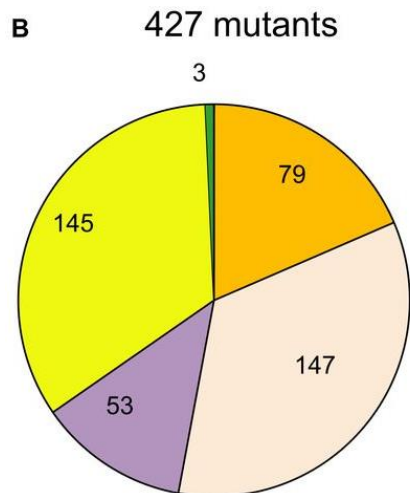
B



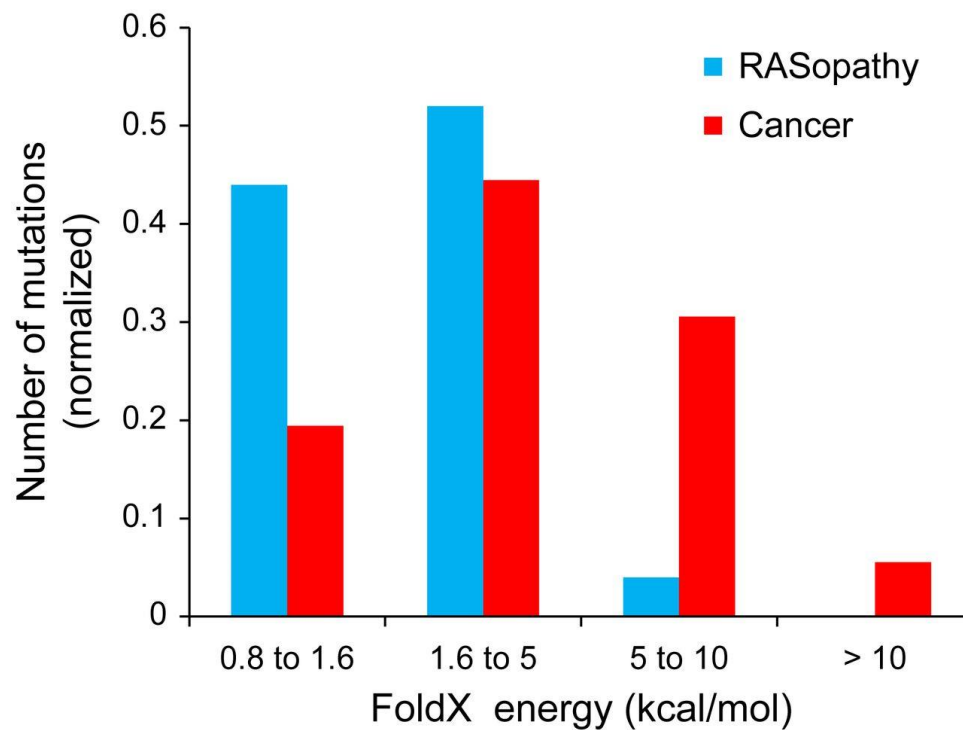
Analysis of 956 missense mutations in RASopathies and cancer: high structural coverage



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

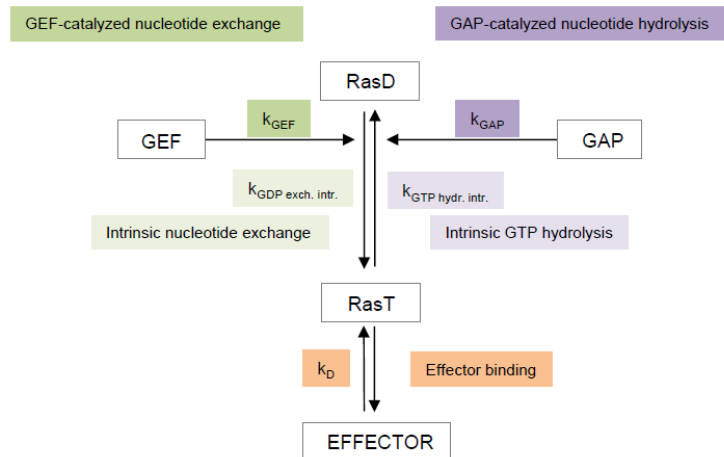


Cancer mutations tend to have higher destabilization values (on average)



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

A



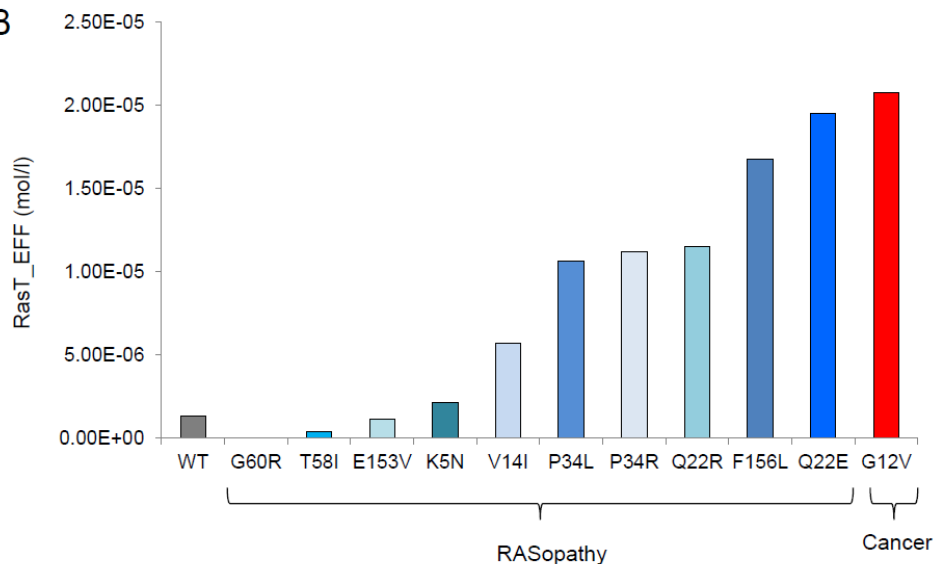
Simulation of Ras activation

‘Enegetics’: quantitative edge effects

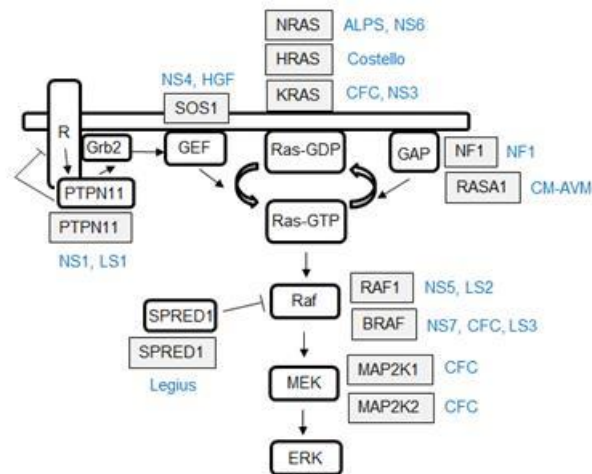
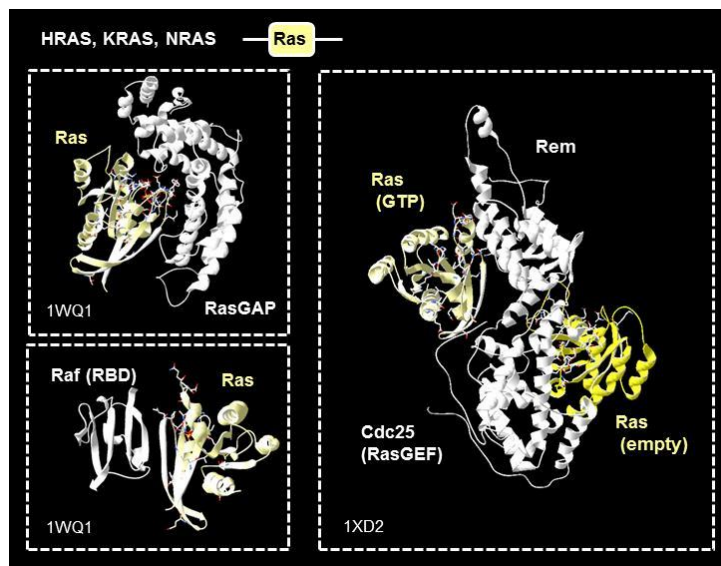
‘Edgetics’ + energies = ‘enegetics’

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

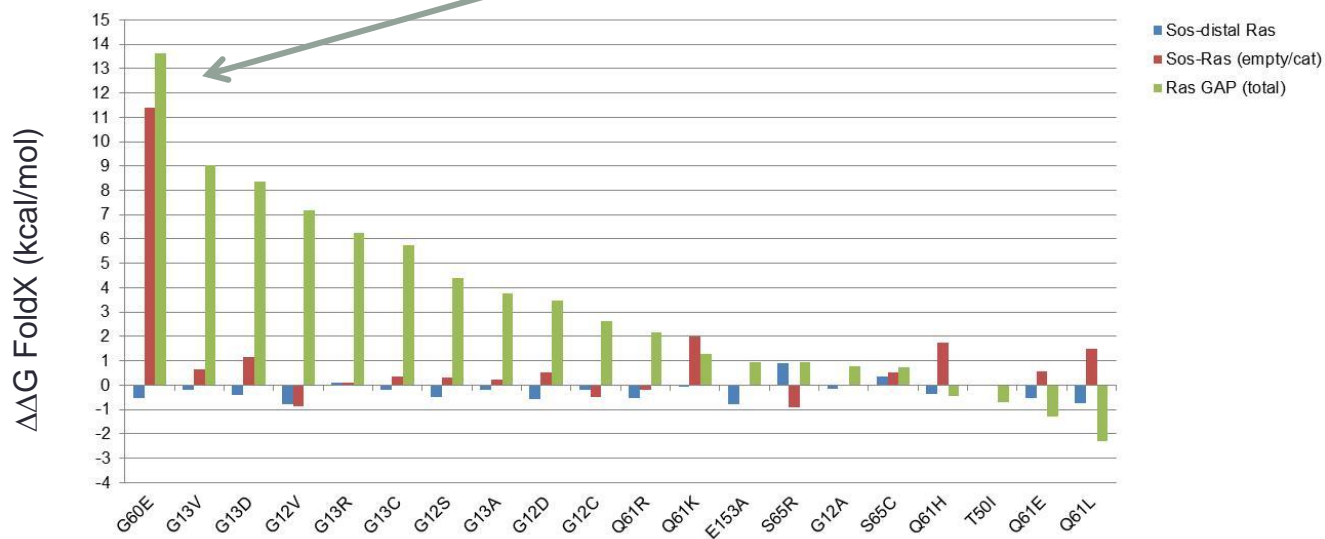
B



Compensatory effects of mutations on different interaction partners

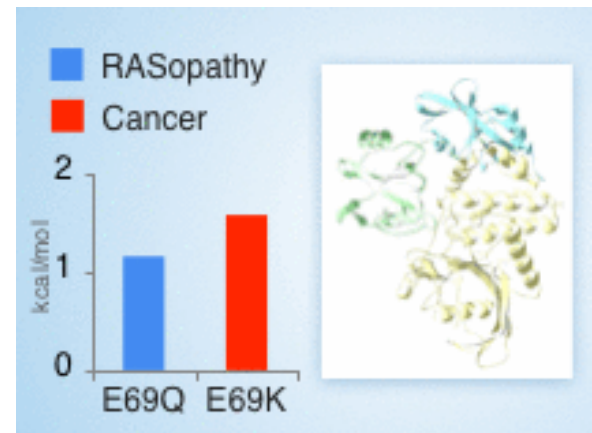


NRAS G60E



Conclusions example 1: RASopathy vs cancer

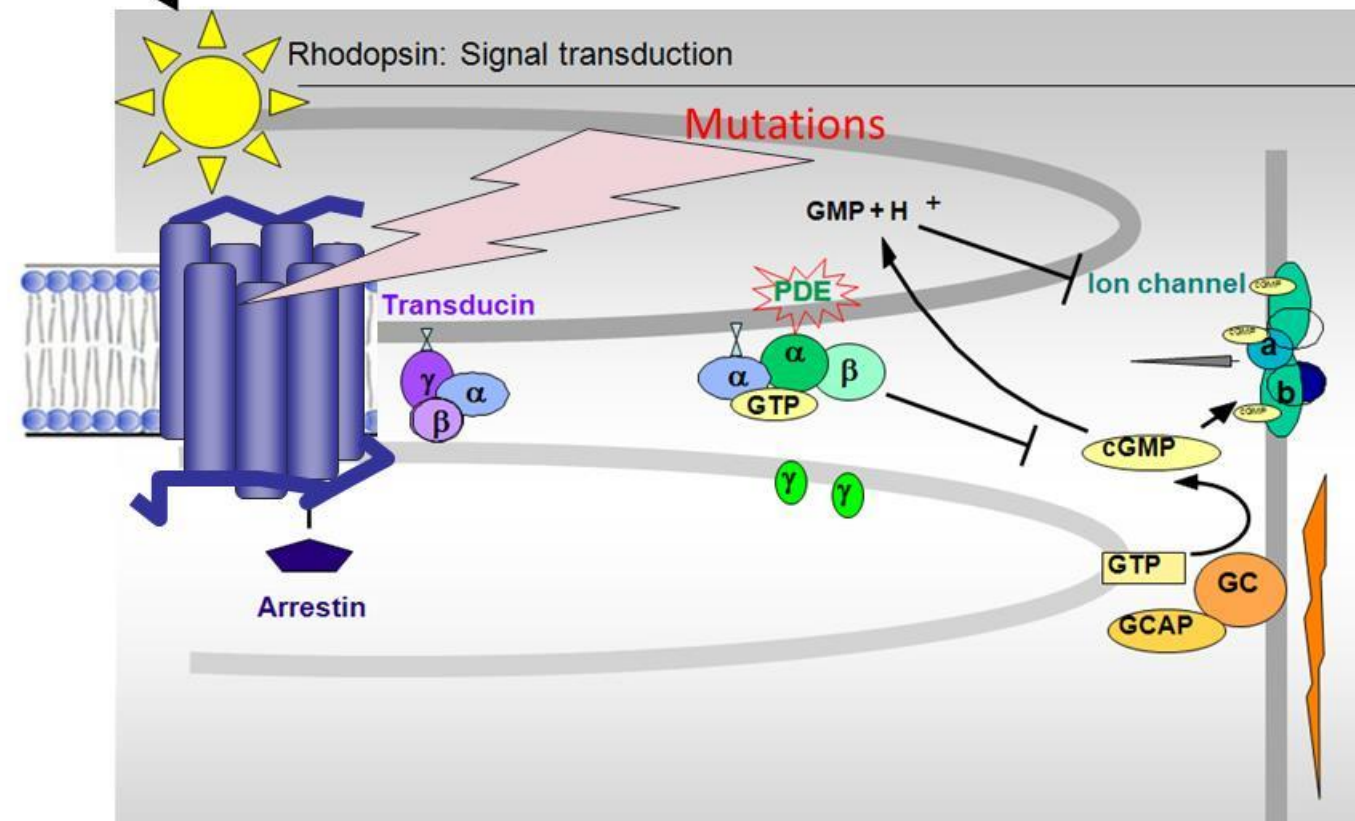
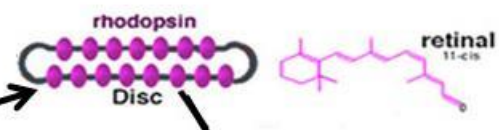
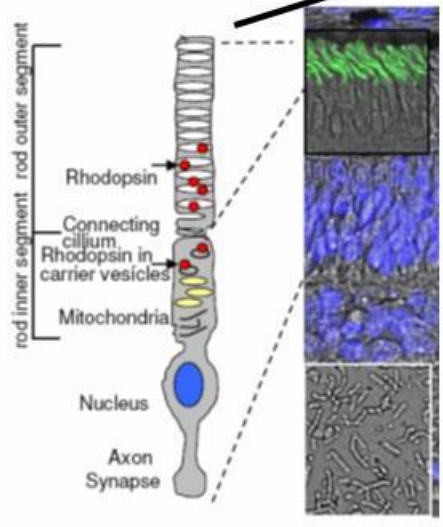
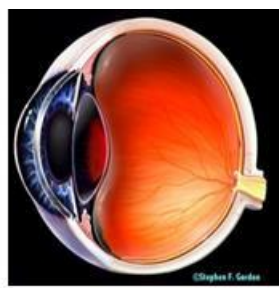
- A systematic analysis of 956 RASopathy and cancer mutations based on structures and energy predictions is presented.
 - Even for the same gene, different disease-causing mechanisms exist depending on the type of mutation.
 - Energy changes are higher for cancer compared to RASopathy mutations.
 - In 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: Rhodopsin disease mutations

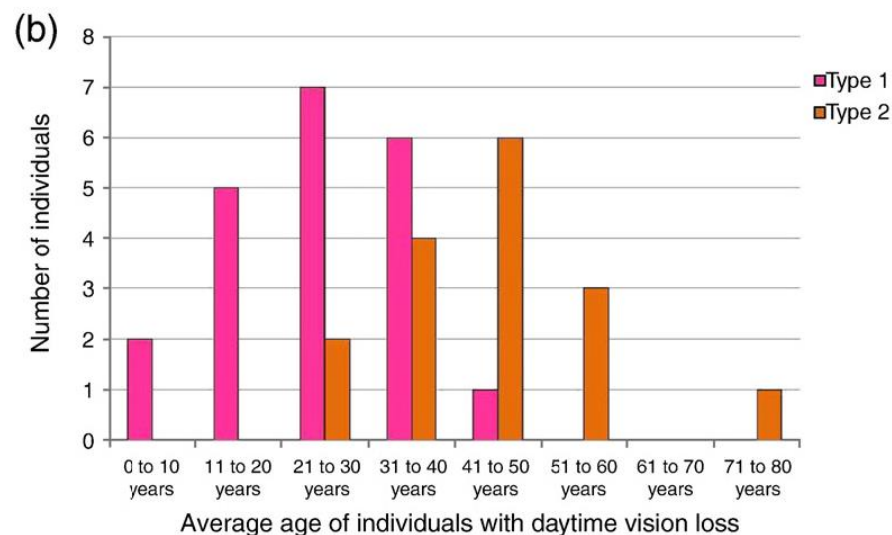
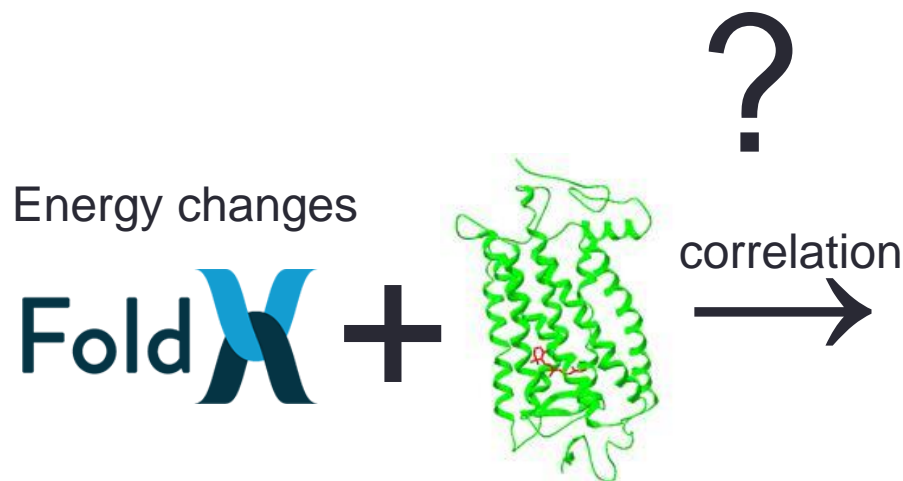
Rhodopsin: involved in light perception in rod outer segment

Understanding disease mutations in rhodopsin, a common cause of retinitis pigmentosa (RP)



Analysis of 103 mutations in rhodopsin linked to RP

Is there a correlation between **energy changes** of rhodopsin missense-mutations and their potential affect on **clinical severity** of Retinitis Pigmentosa (RP)?



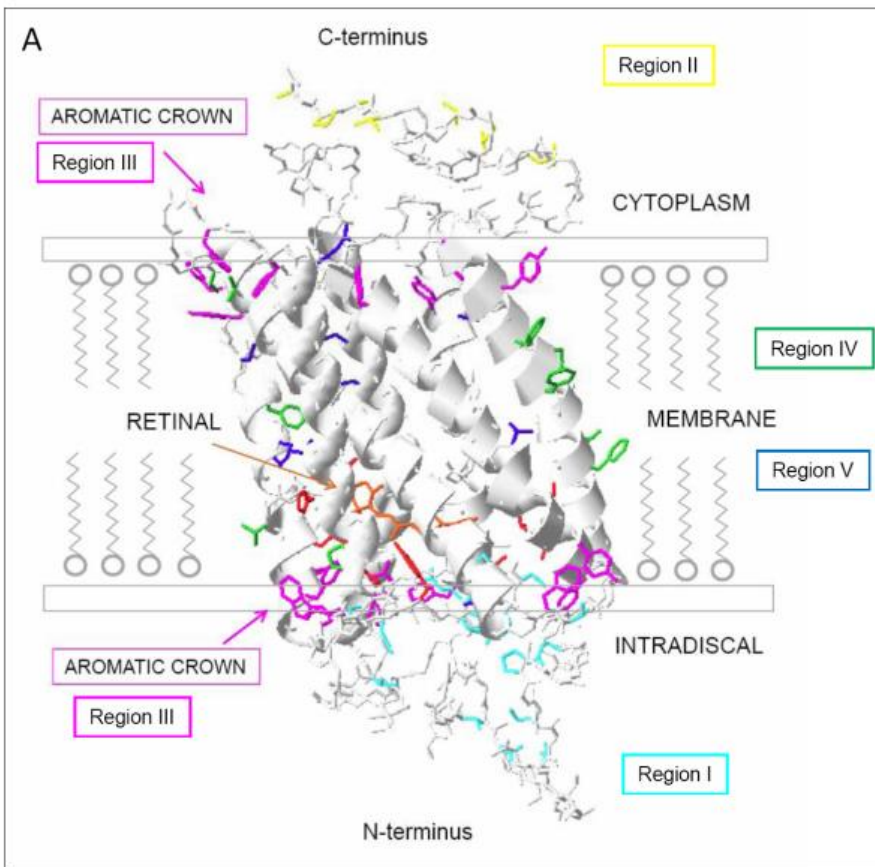
Several consideration for studying the effect of missense mutations in rhodopsin

1) Rhodopsin is a membrane protein: can we use FoldX, a design algorithm developed for *soluble proteins*, for predicting the effect of mutants for a membrane protein?

Region I mutants (intradiscal):
✓ **YES**, not in membrane

Region II mutants (cytoplasm):
✓ **YES**, not in membrane

Region IV mutants (residues pointing outside and facing the lipid bilayer):
NO, a mutation from hydrophobic to polar residue could be predicted favorable by FoldX, but would prevent proper integration of rhodopsin into the membrane.



For analyzing Region IV mutants (residues pointing outside and facing the lipid bilayer): use a different algorithm

Linking amino acid sequence to membrane insertion efficiency

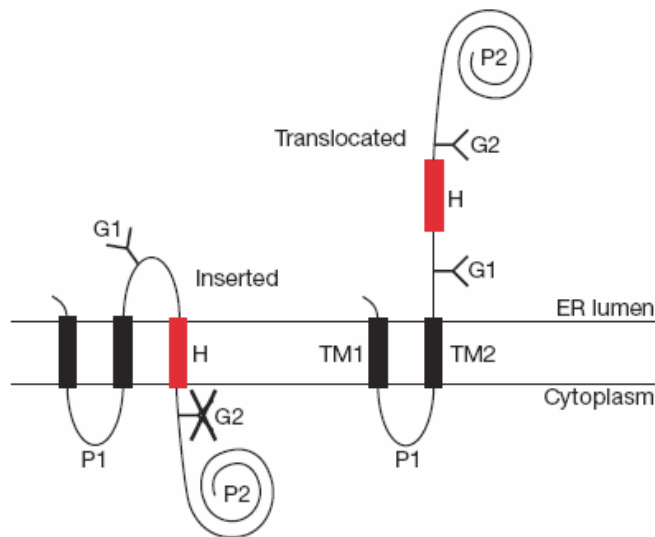


Figure 1 | The Lep model protein. *Escherichia coli* leader peptidase (Lep) has two TM helices (TM1 and TM2) and a large luminal domain (P2). It inserts into rough microsomes in an $N_{lum}-C_{lum}$ orientation. H-segments (red) are engineered into the P2 domain with two flanking Asn-X-Thr glycosylation acceptor sites (G1, G2). Constructs for which the H-segment is integrated into the endoplasmic reticulum membrane as a TM helix are glycosylated only on the G1 site (left), whereas those for which the H-segment is translocated across the membrane are glycosylated on both the G1 and G2 sites (right).

This algorithm is based on experimental results, in which systematically designed 19-residue long amino acid sequences have been expressed and tested in-vitro for TM insertion.

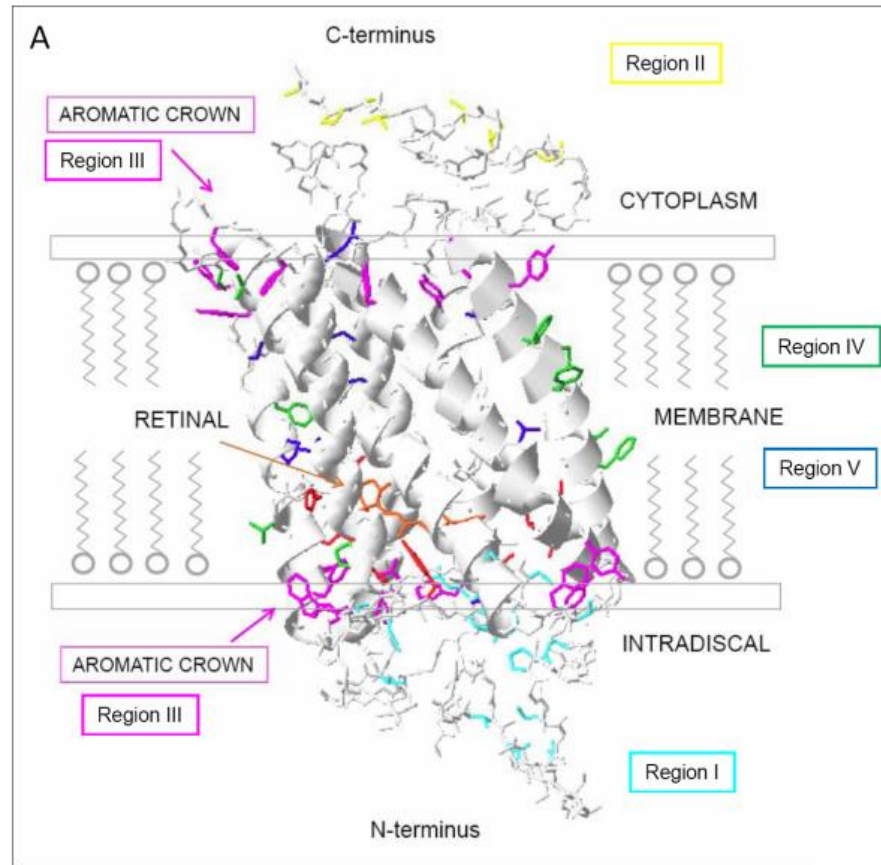
Home [Delta G prediction](#) [Full protein scan](#) [Instructions](#) [Publications](#) [Contact](#)

Delta G prediction server v1.0

Given the amino acid sequence of a putative transmembrane (TM) helix, the server gives a prediction of the corresponding apparent free energy difference, ΔG_{app} , for insertion of this sequence into the Endoplasmic Reticulum (ER) membrane by means of the Sec61 translocon. The server runs in two different "modes", for two different types of queries:

- [Delta G prediction](#) Predict ΔG_{app} for membrane insertion of a potential TM helix.
- [Full protein scan](#) Scan a protein sequence for putative TM helices.

Several considerations for studying the effect of missense mutations in rhodopsin



1) Rhodopsin is a membrane protein: can we use FoldX, a design algorithm developed for *soluble proteins*, for predicting the effect of mutants for a membrane protein?

Region I mutants (intradiscal):

✓ **YES**, not in membrane

Region II mutants (cytoplasm):

✓ **YES**, not in membrane

Region IV mutants (residues pointing outside and facing the lipid bilayer):

NO, a mutation from hydrophobic to polar residue could be predicted favorable by FoldX, but would prevent proper integration of rhodopsin into the membrane.

Region V mutants (residues facing inside the helices):

NO, FoldX desolvation effect is possibly not appropriate since the reference state in soluble proteins is water and in membranes, lipids.

BUT: VanderWaal's clashes of course will be the same for a soluble or membrane protein. To avoid issues related to the proper calibration of the desolvation effect for buried residues in membrane proteins for residues in Region V we determined both the *overall change in energy* and the *Vander Waals' clashes*.

Several consideration for studying the effect of missense mutations in rhodopsin

2) Retinal-free Rhodopsin is unstable: If an amino acid residue contributes to binding a mutation might not necessarily lead to destabilization (energies of retinal not calibrated) → We need to identify all residues in the retinal binding area, and treat the results of mutations involving these residues, separately.

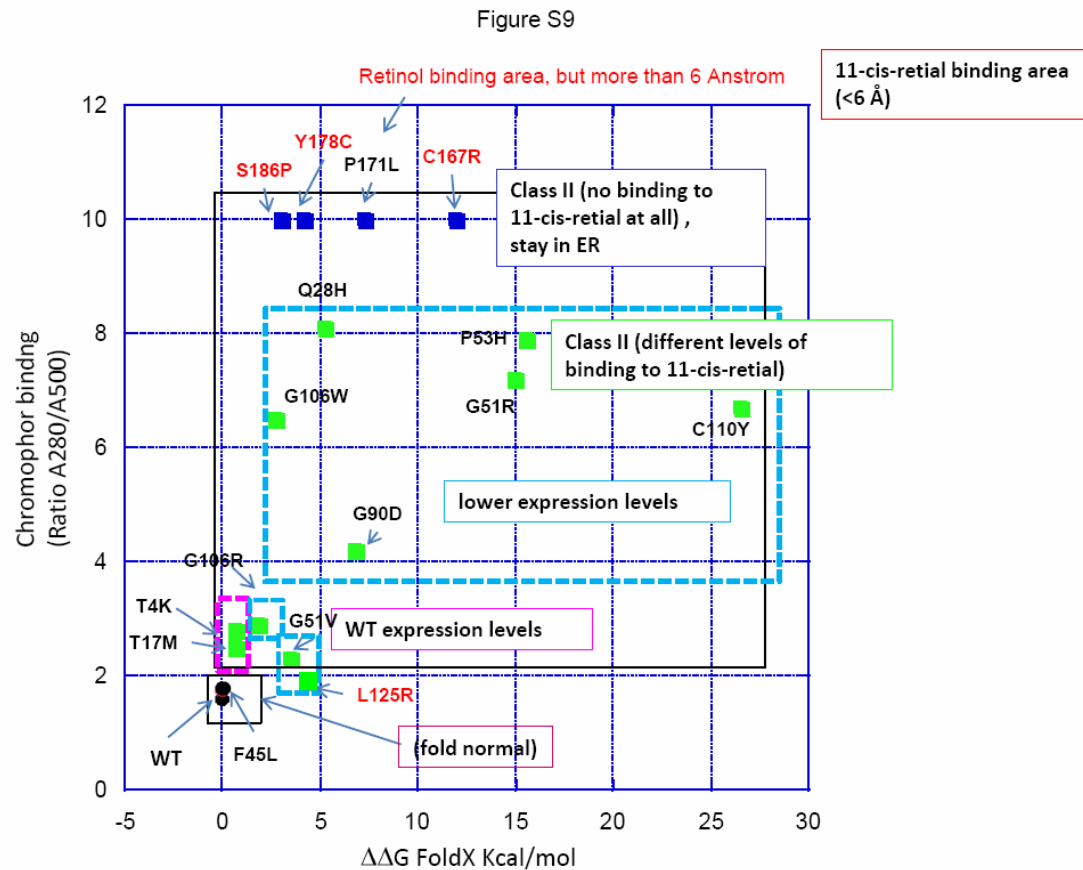
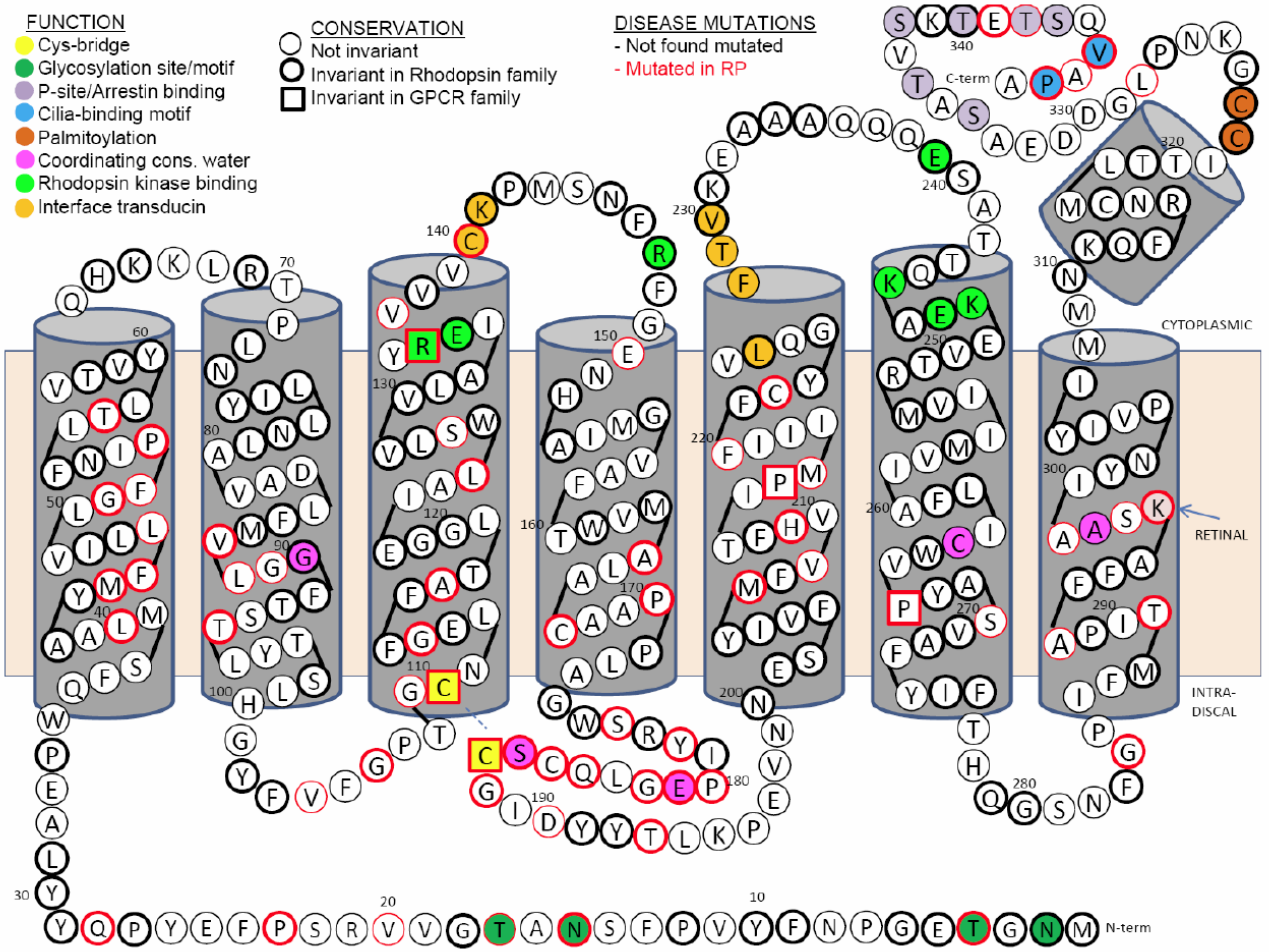


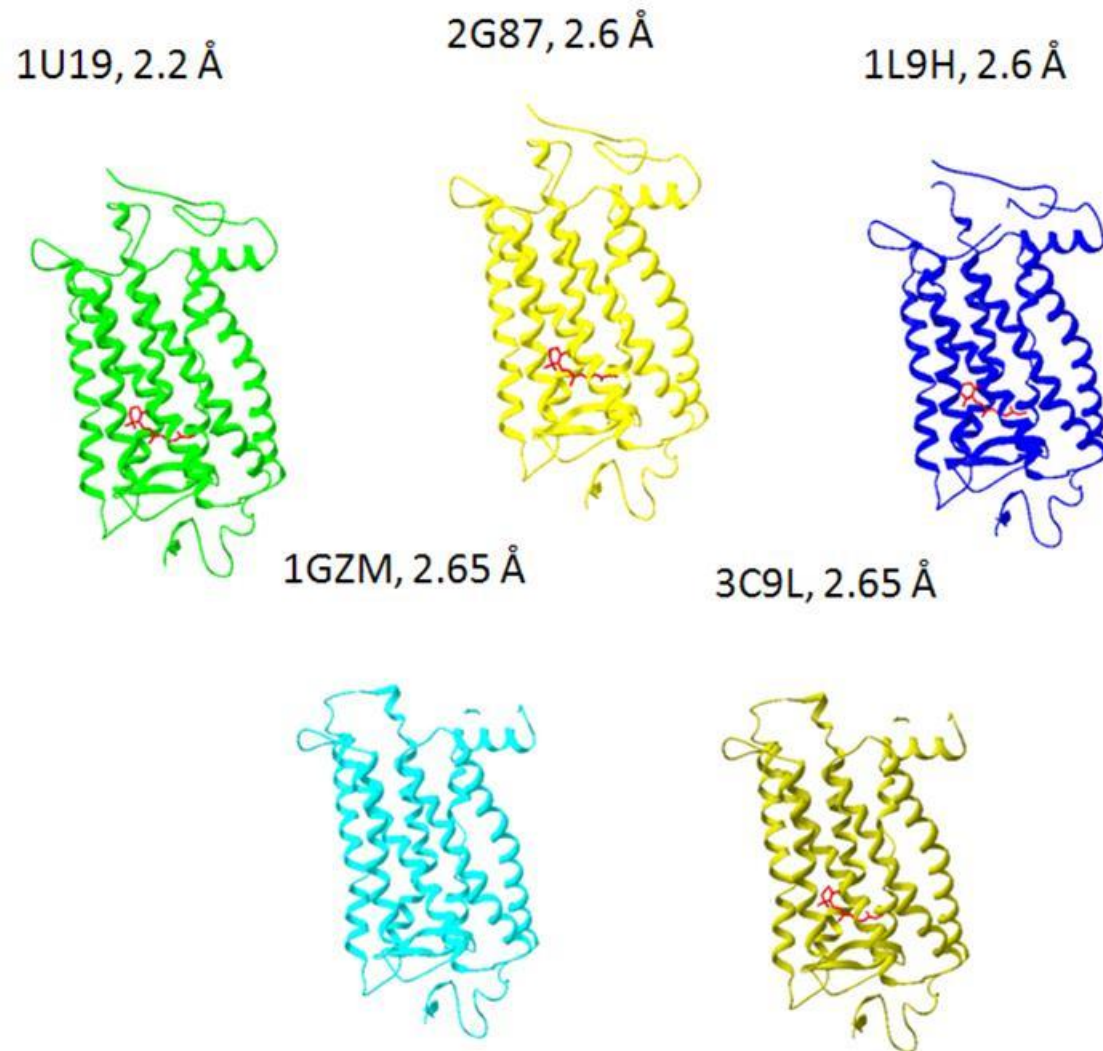
Figure S9 Correlation of FoldX $\Delta\Delta G$ values with in vitro expression levels and chromophore binding properties. In vitro data for protein expression levels and chromophore binding capabilities were taken from Kaushal and Khorana [19] and Briscoe et al. [49]. Involvement of mutants in different Classes (according to the Mendes [20] classification) are indicated. For mutants that show no binding of 11-cis-retinal at all (no peak at 500 nm), we assumed that it corresponds to a ratio A280/A500 of 10. (The reason to assume this was that when A500 is very low, the ratio is round 8 to 9).

Several consideration for studying the effect of missense mutations in rhodopsin

3) Rhodopsin is involved in other functions (e.g. binding to partner proteins): A mutation might cause disease but not be predicted destabilizing with FoldX → We need to know as much as possible about rhodopsin function.

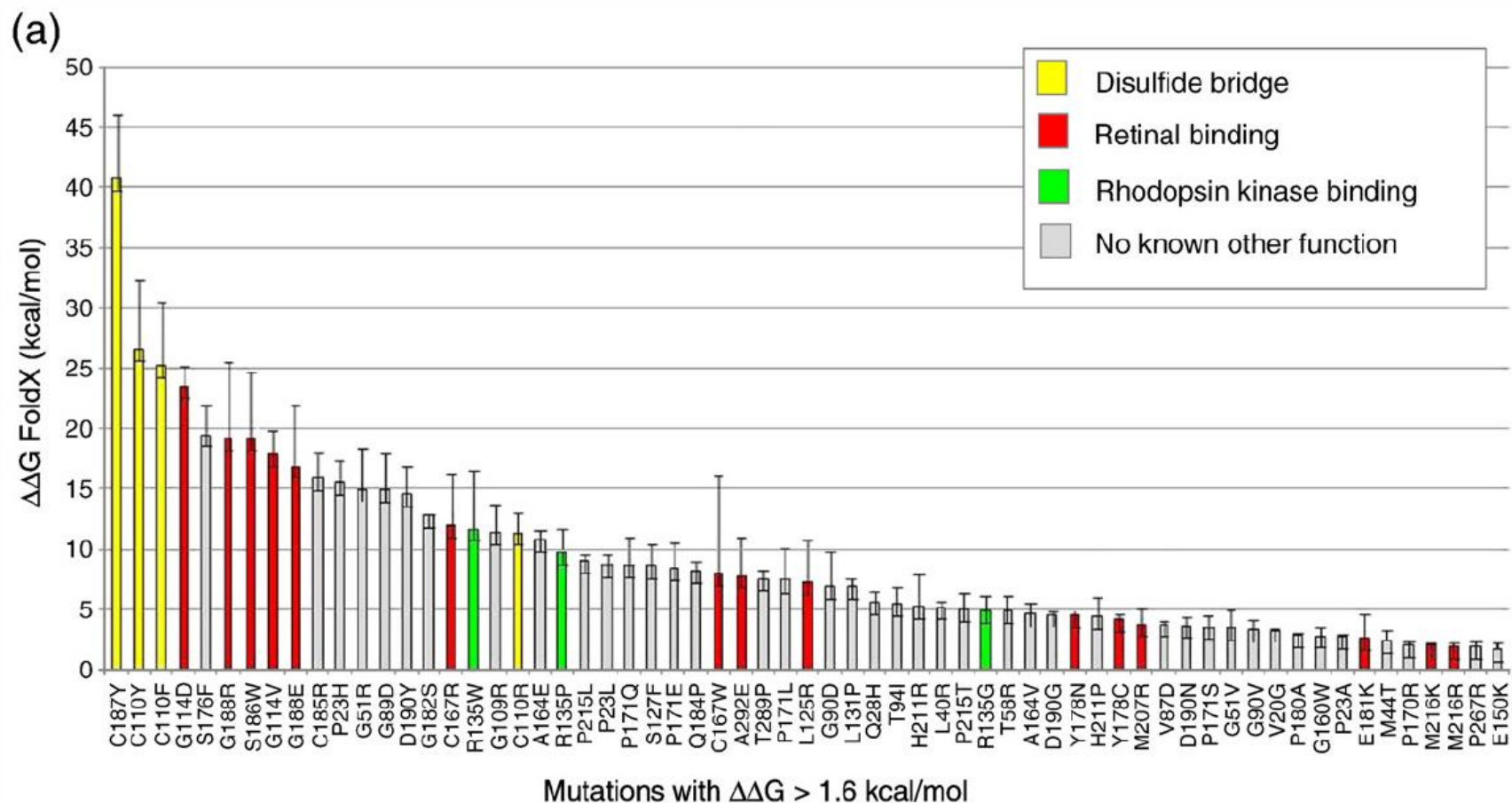


Five structures of bovine rhodopsin were selected ($<2.6 \text{ \AA}$) for mutagenesis and protein stability analysis using FoldX



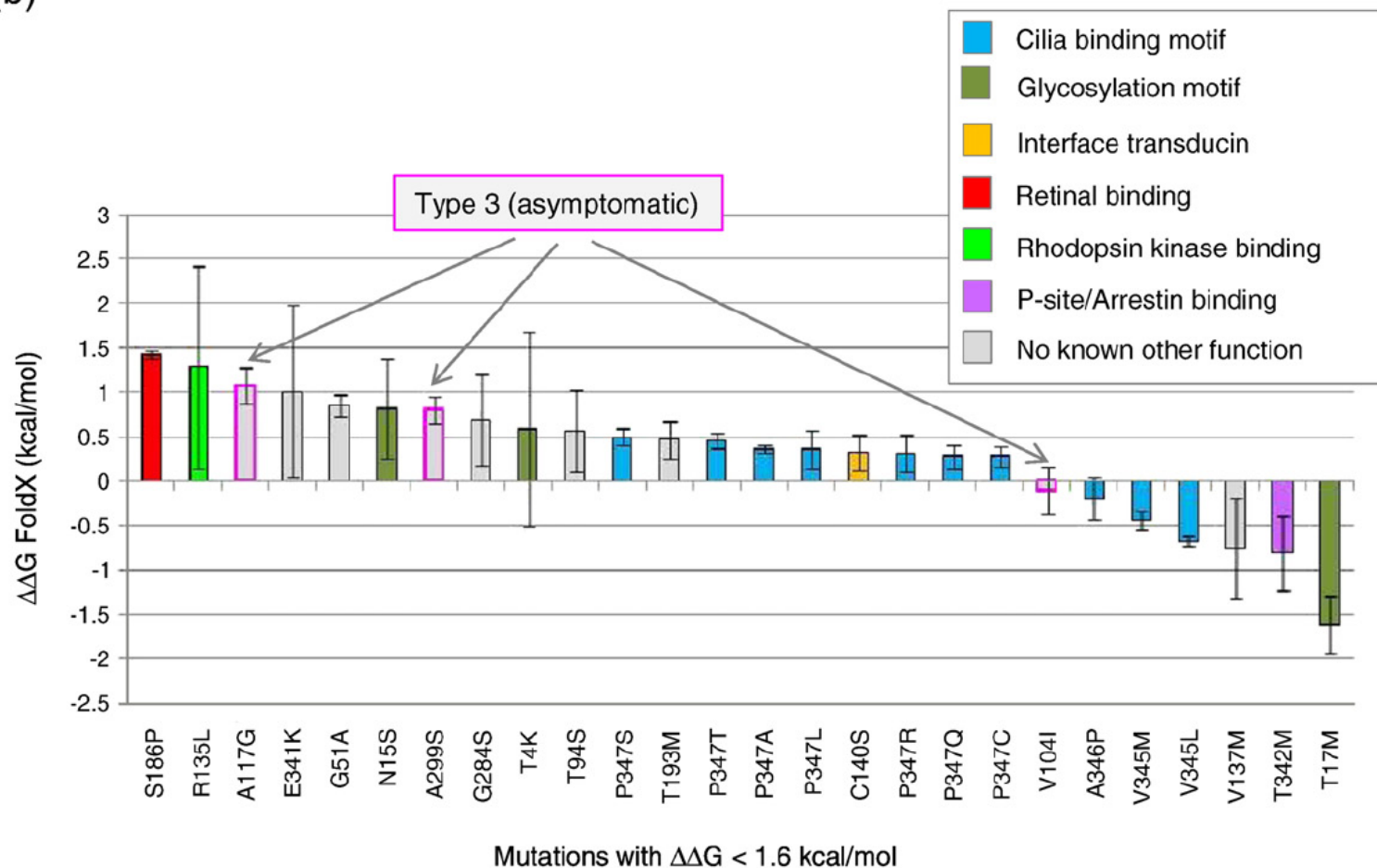
FoldX energy results and involvement in other function

Mutants that are destabilizing ($\Delta\Delta G > 1.6$ kcal/mol)



FoldX energy results and involvement in other function

(b)



- Mutants that are not destabilizing, are usually involved in other functions, which can explain their disease-causing effect.

FoldX calculations and comparing with phenotypic data

V20G	P23A	P23L	P23H	Q28H	L40R	M44T	G51R	G51V	T58R
V87D	G89D	G90D	G90V	T94I	G106R	G106W	G109R	C110Y	C110F
C110R	G114D	G114V	L125R	S127F	L131P	R135G	R135W	R135P	E150K
A164V	A164E	C167R	C167W	P170R	P171E	P171S	P171L	P171Q	S176F
Y178N	Y178C	P180A	E181K	G182S	Q184P	C185R	S186W	C187Y	G188R
G188E	D190N	D190G	D190Y	M207R	H211P	H211R	P215T	P215L	M216R
M216K	T289P	A292E							

(bovine: Leu)

Residues cannot be analysed with FoldX:
12 (Region IV, Van Heijne)

F45L	L46R	F52Y
P53R	F56Y	V209M
F220C	F220L	P267L
P267R	S270R	S297R

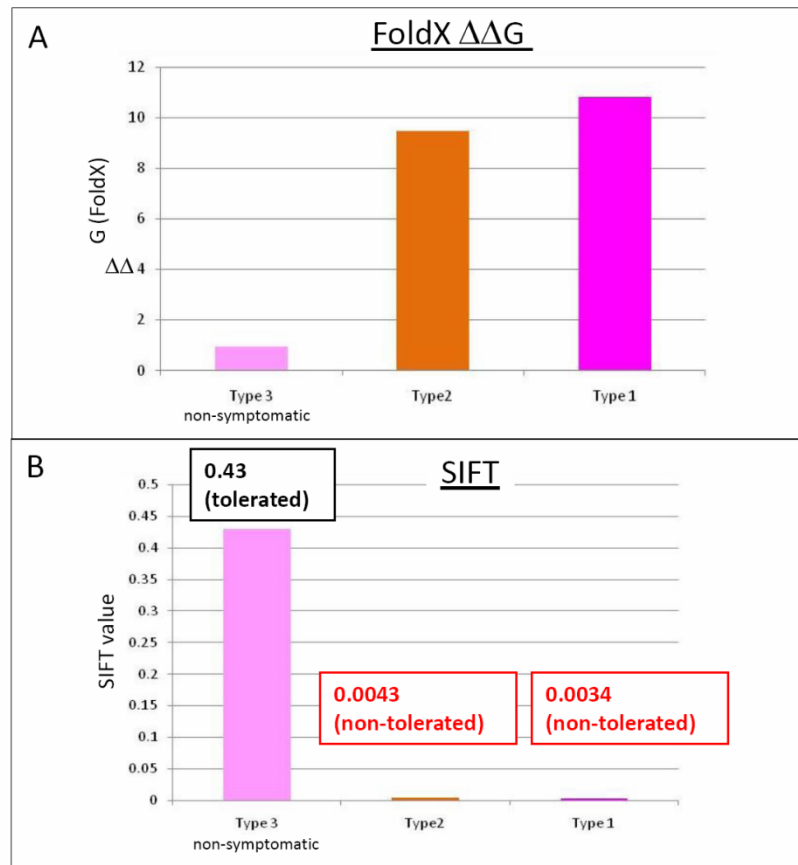
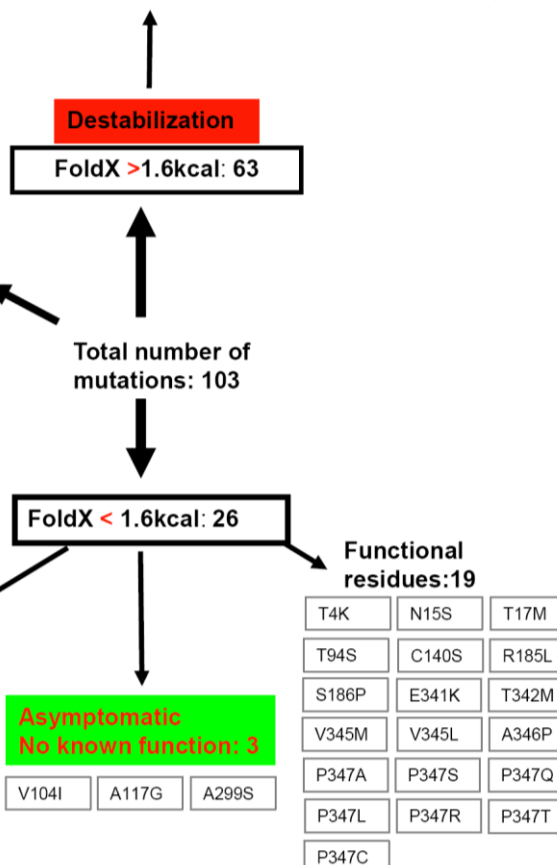
(bovine: Gly) (bovine: Thr)

Bad structure: 2

C222R	L328P
-------	-------

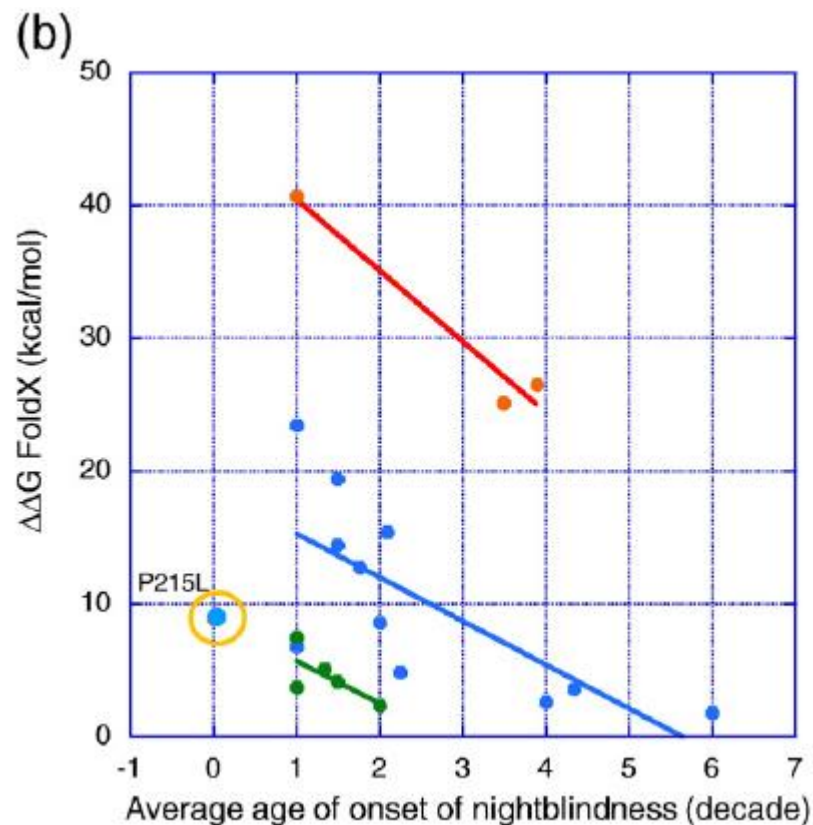
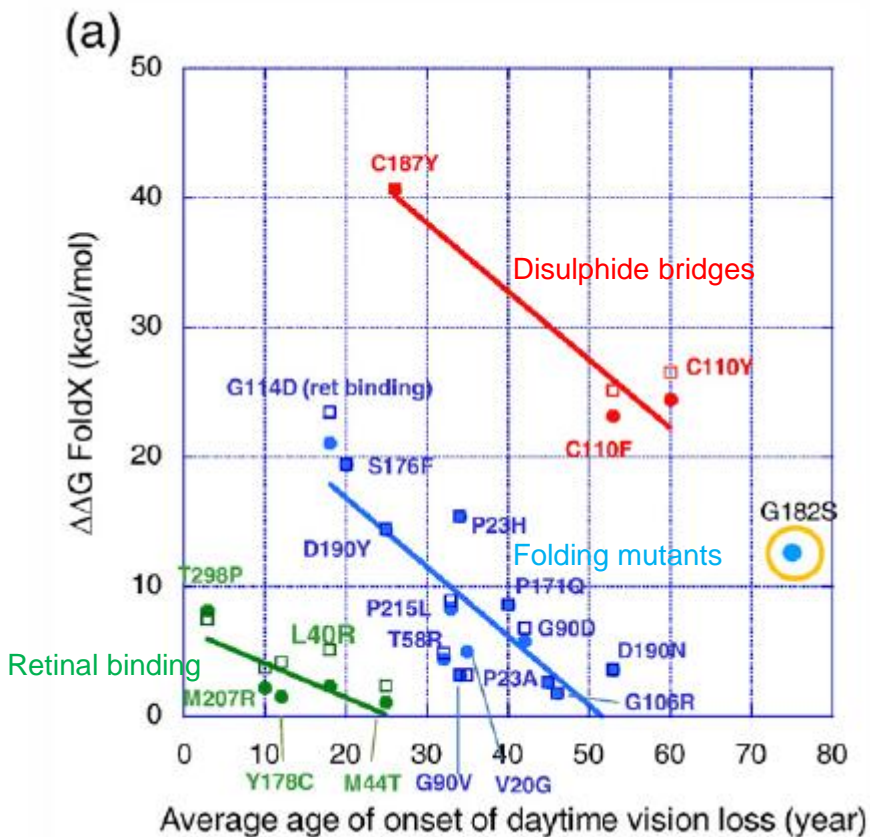
Unexplained/
misdiagnosed: 4

G51A	G284S
V137M	T193M



Correlation of daytime vision loss and night blindness with FoldX energy calculations

Different therapies should be used for the three different types of mutations



Conclusions example 2: Rhodopsin mutations

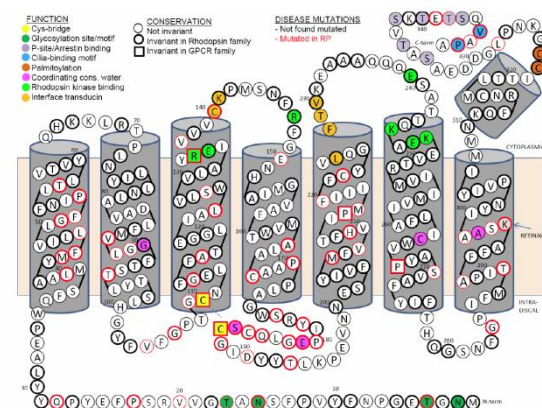
- The majority of the mutants is located within the hydrophobic core of the corresponding proteins and are therefore likely to cause misfolding.

Quantitative predictive assessment for the severity and onset of the disease:

- For folding mutations where sub-typing was available we found a significant correlation between FoldX energy changes and both the average onset age of night-blindness, daytime vision loss and visual acuity.

Most important conclusion:

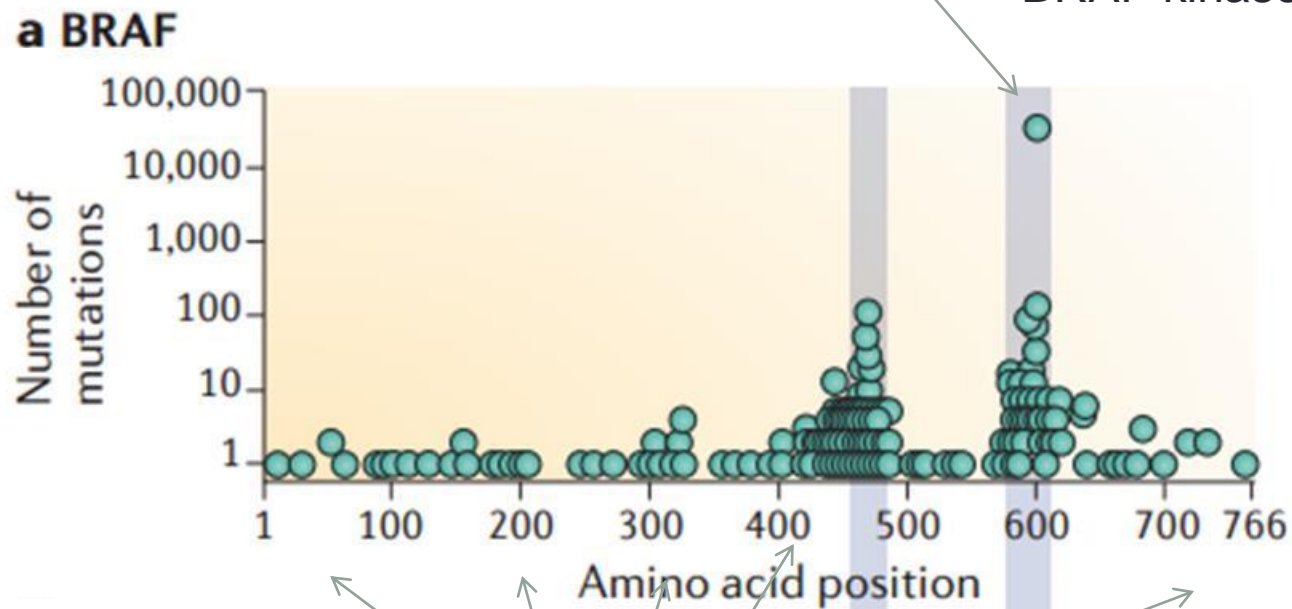
- a high level of functional understanding was necessary for our analysis and the observed energy-phenotype correlation.



Example 3: 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



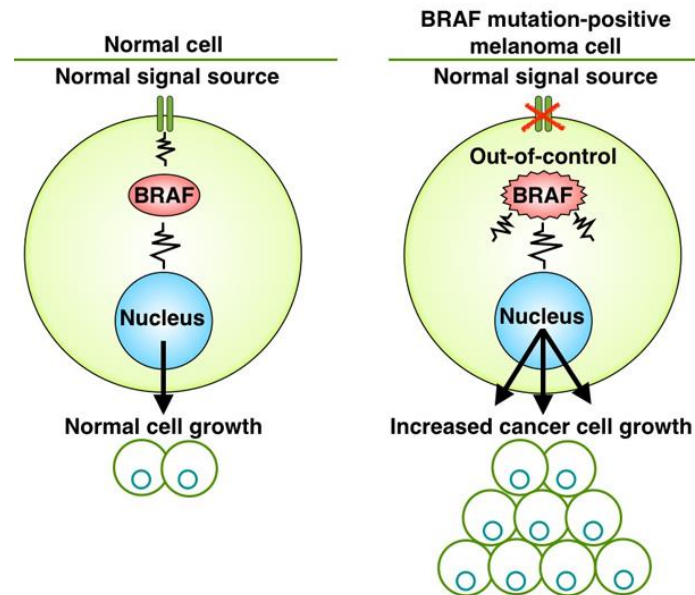
Shall we only treat patients which harbour V600E mutations or also patients with non-V600E mutations?

Catalytic activity of kinases is usually tightly controlled

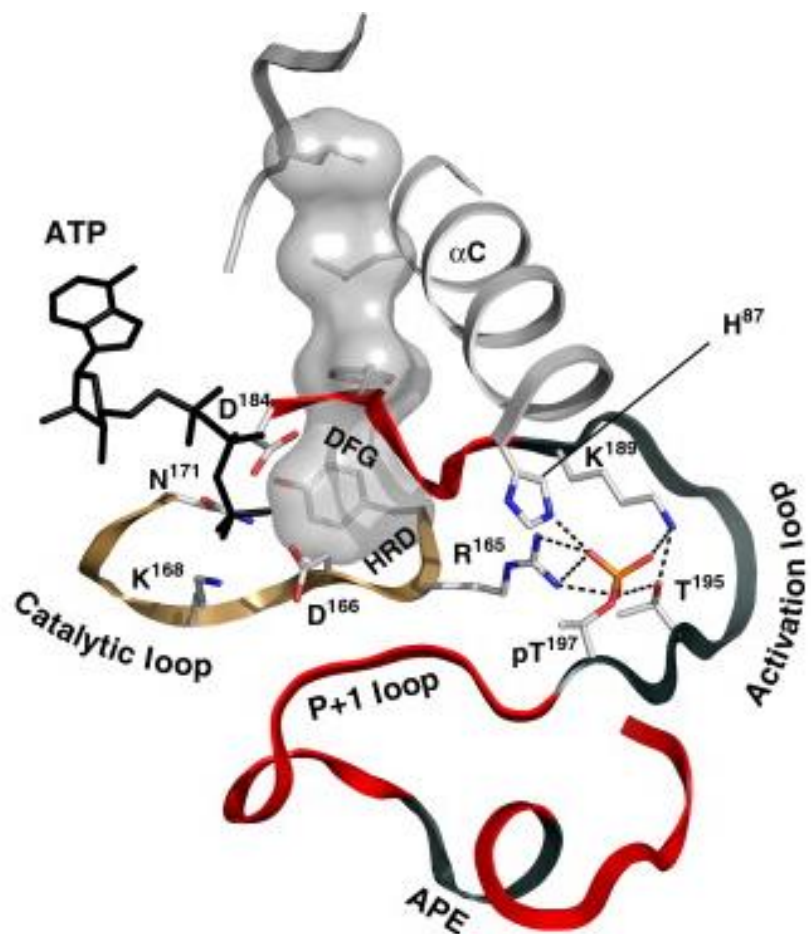
Mechanisms for kinase activation are:

- phosphorylation
- additional domains or subunits of the kinase
- scaffolding proteins
- kinase dimerization

Mutations in kinases (e.g. BRAF) can cause constitutive kinase activation and over activation of downstream signaling, which can cause cancer



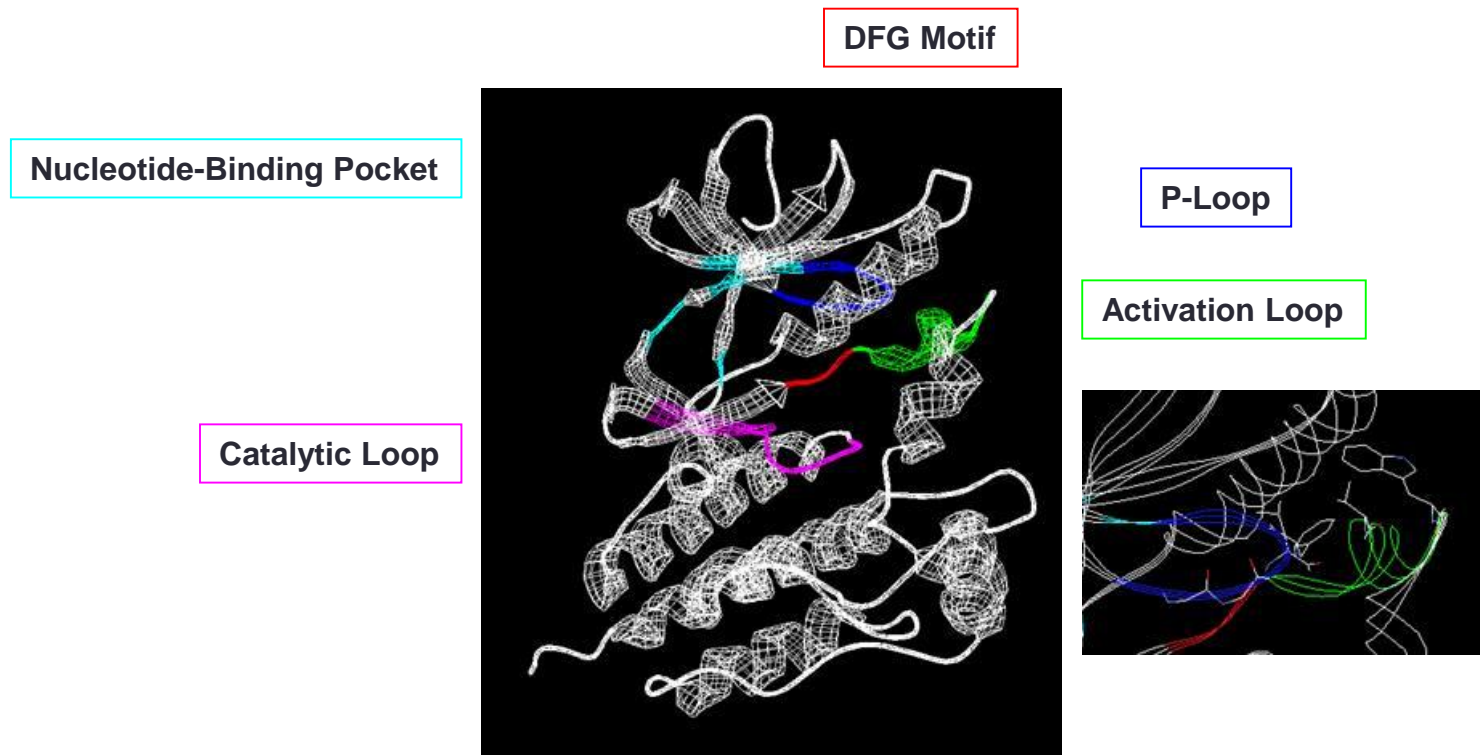
Kinases are activated through mutations in the activation loop (activation segment)



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

BRAF kinase activation through oncogenic mutations (e.g. V600E)

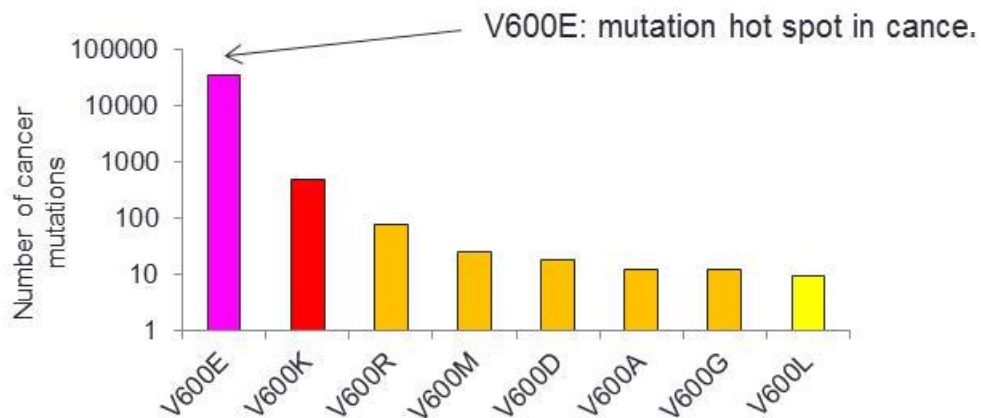
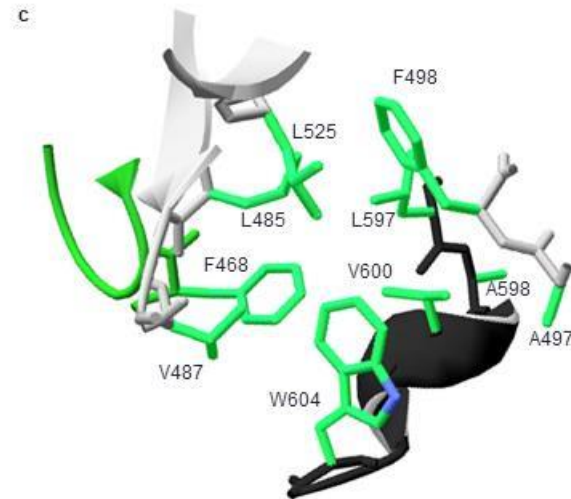
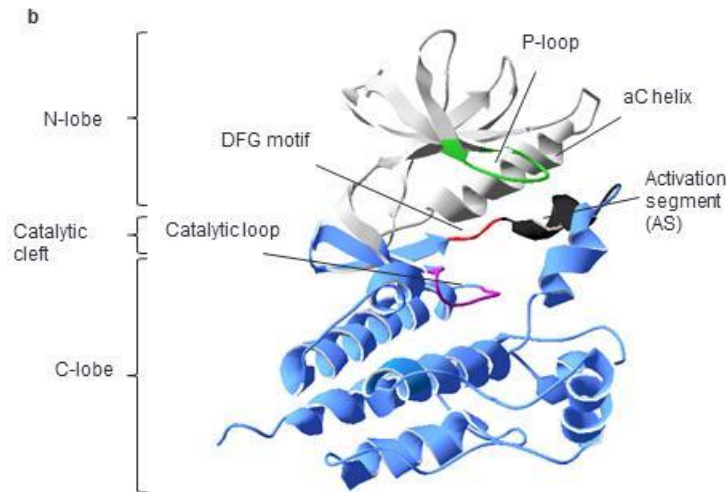
V600E mimics the negative charge of the neighbouring phosphorylated Thr599-P



Activation loop residues: form **strong hydrophobic interactions with the P-loop in the inactive conformation of the kinase, locking the kinase in its inactive state** until the activation loop is phosphorylated, destabilizing these interactions with the presence of negative charge. This triggers the shift to the active state of the kinase. Specifically, L597 and **V600** of the **activation loop** interact with G466, F468, and V471 of the **P-loop** to keep the kinase domain inactive until it is phosphorylated

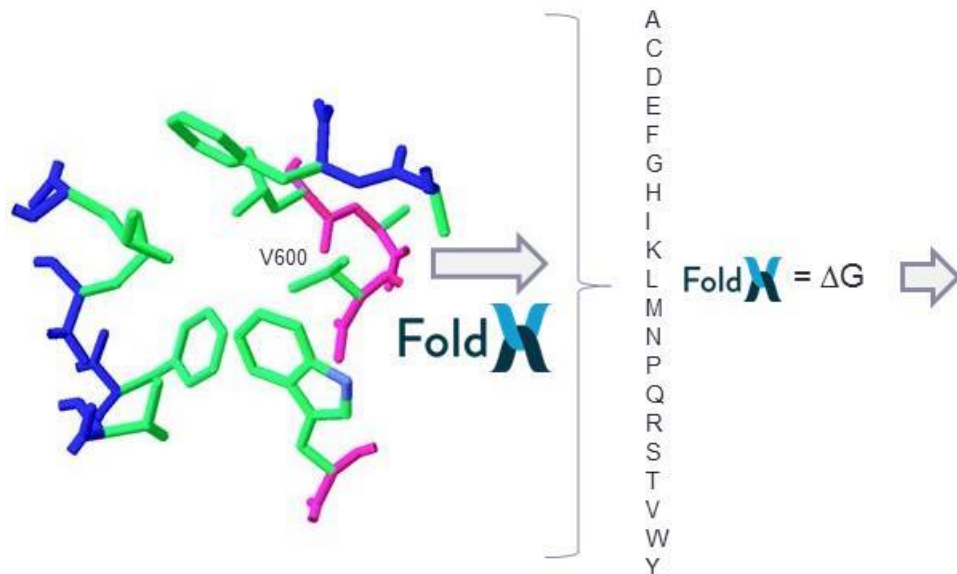
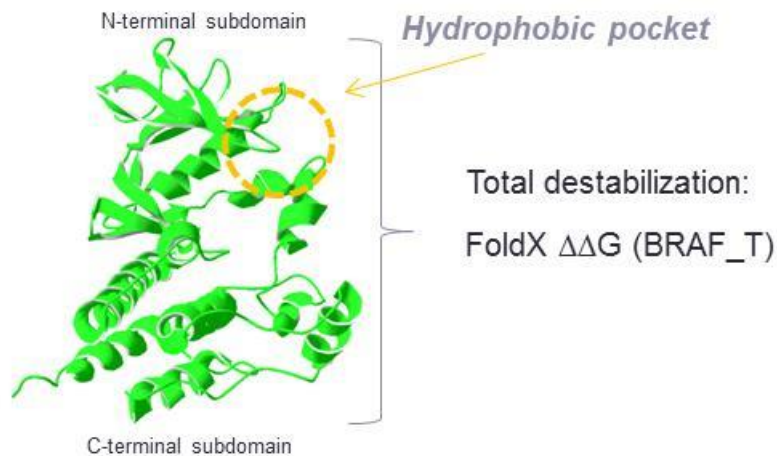
Focus on the position Val600 in the kinase BRAF

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

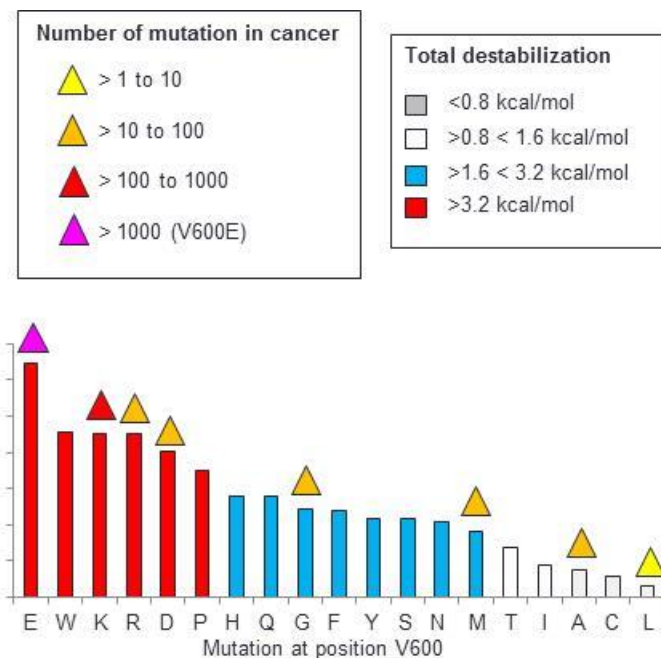


Differences in mutation frequencies: a [quantitative effect?](#)

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



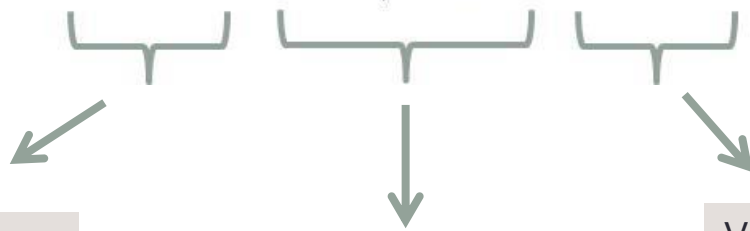
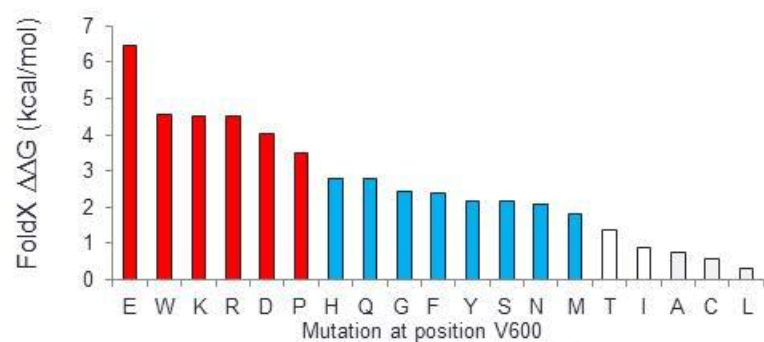
Destabilization of inactive state



No destabilization of active state (data not shown)

Distinguishing driver from passenger mutations

FoldX



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

Fitness??

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

V600G behaves more like a RASopathy mutation

Google search for “V600G BRAF CFC syndrome”: V600G found as a RASopathy mutation 😊

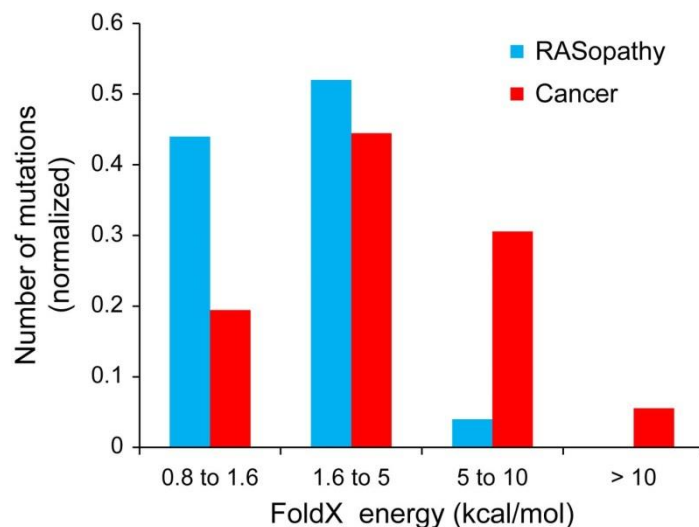
Germline mutation in *BRAF* codon 600 is compatible with human development: *de novo* p.V600G mutation identified in a patient with CFC syndrome

[Champion, KJ¹](#); [Bunag, C²](#); [Estep, AL²](#); [Jones, JR¹](#); [Bolt, CH¹](#); [Rogers, RC¹](#); [Rauen, KA³](#); [Everman, DB¹](#)

Clinical Genetics, Volume 79, issue 5 (May 2011), p. 468-474.

ISSN: 0009-9163 DOI: 10.1111/j.1399-0004.2010.01495.x

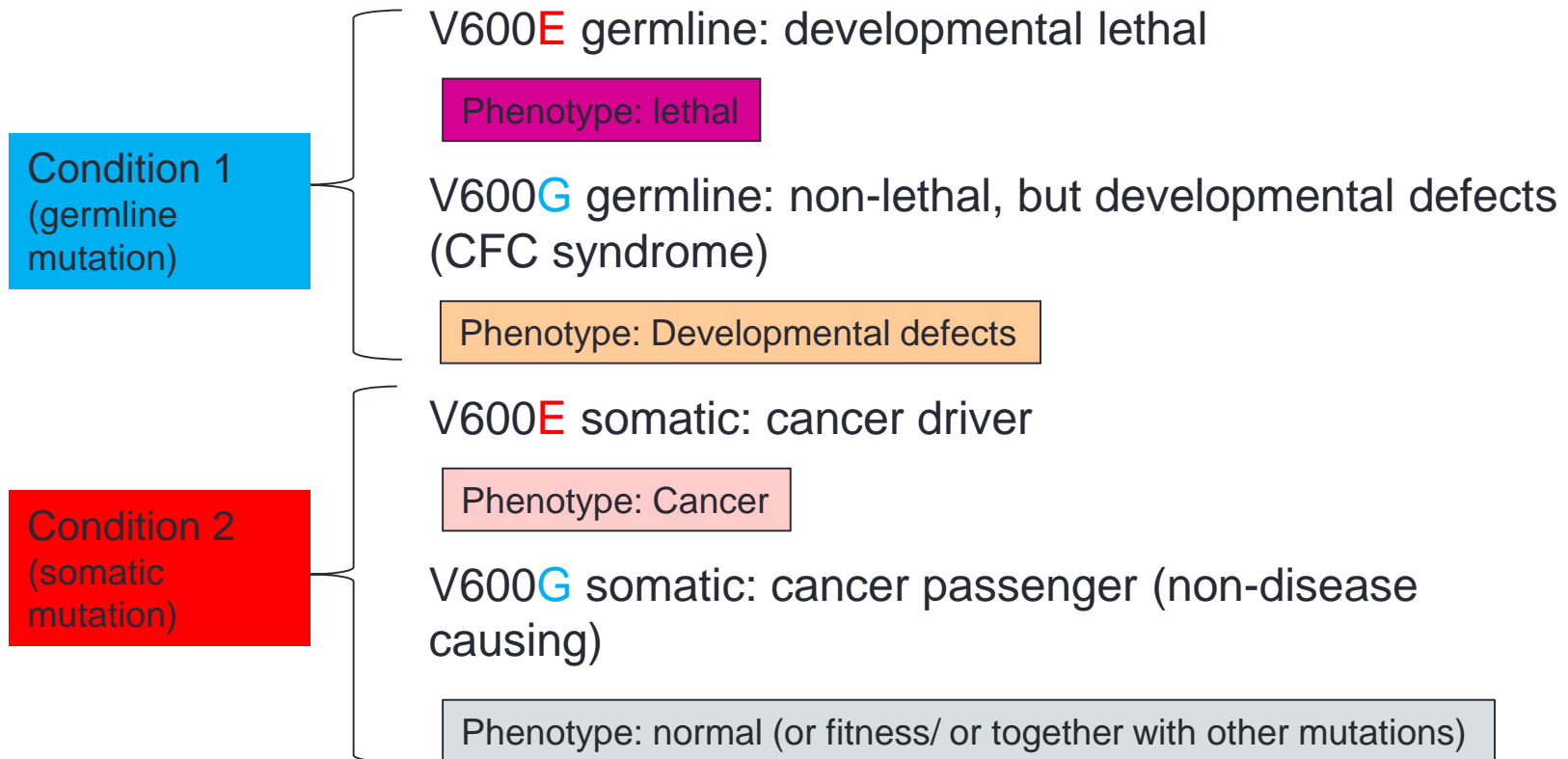
Blackwell Publishing Ltd



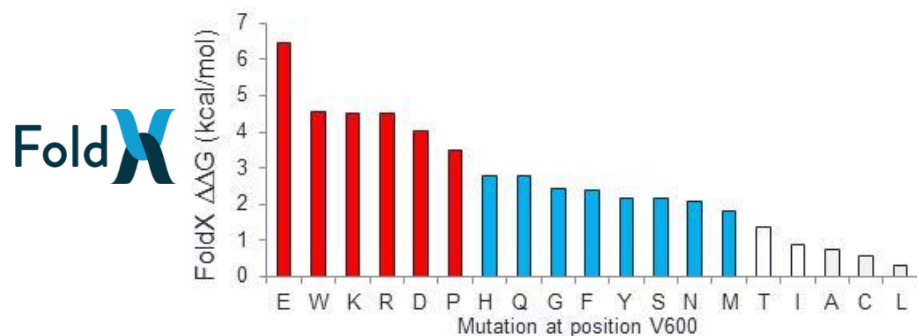
“enedgetics”
Cancer mutations tend to have higher destabilization values (on average)

Kiel & Serrano, 2014

Different energy thresholds for germline and somatic mutations? 'Condition-dependent phenotypes'



Why different cancer frequencies for V600E, V600D and V600K?

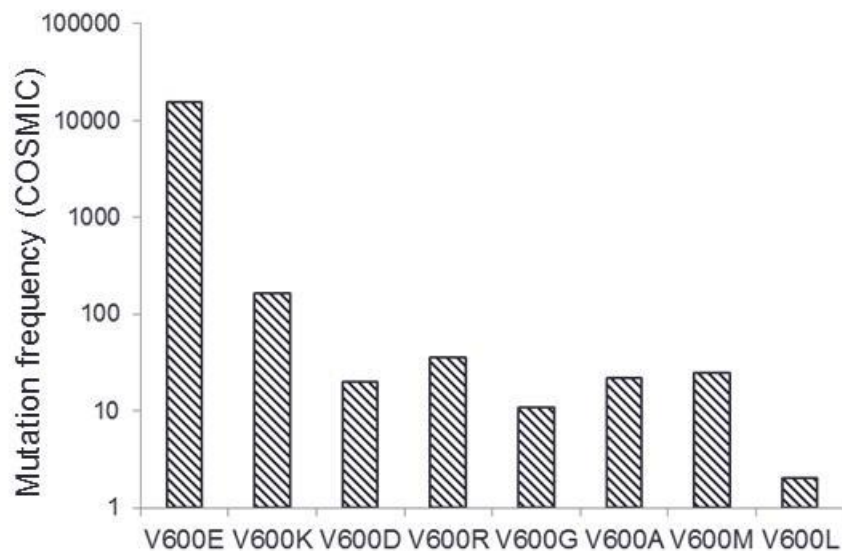


V600K, D, and R have very similar destabilizing energies

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

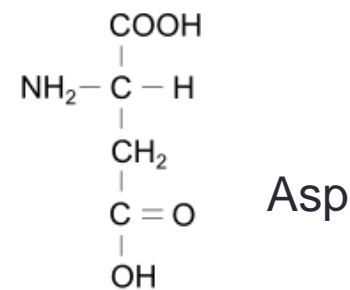
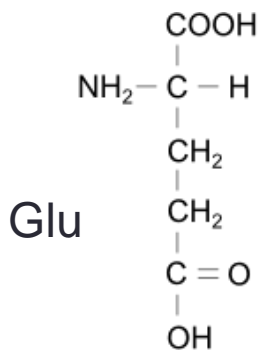
Why different cancer frequencies for V600E, V600D and V600K?



V600E: 15474 frequency

V600D: 20 frequency

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



Why different cancer frequencies for V600E, V600D and V600K?

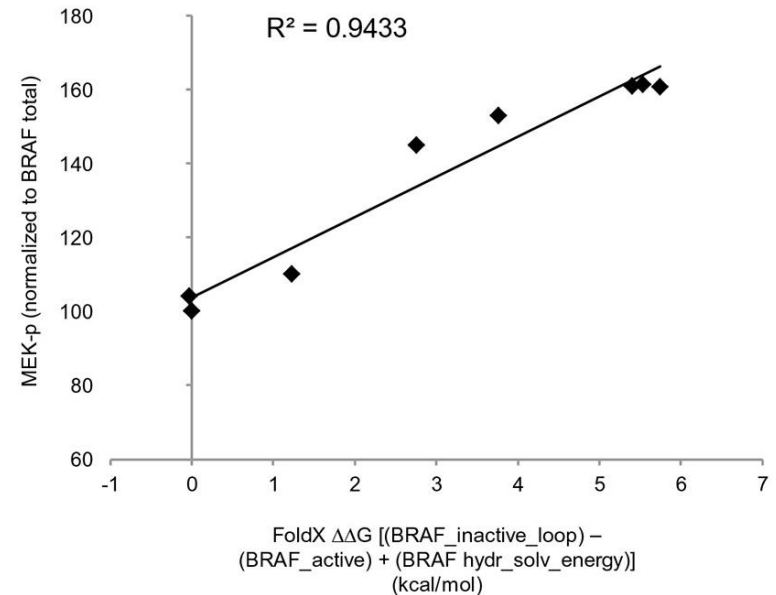
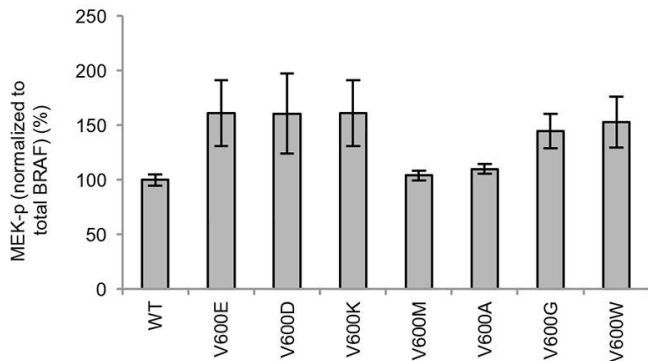
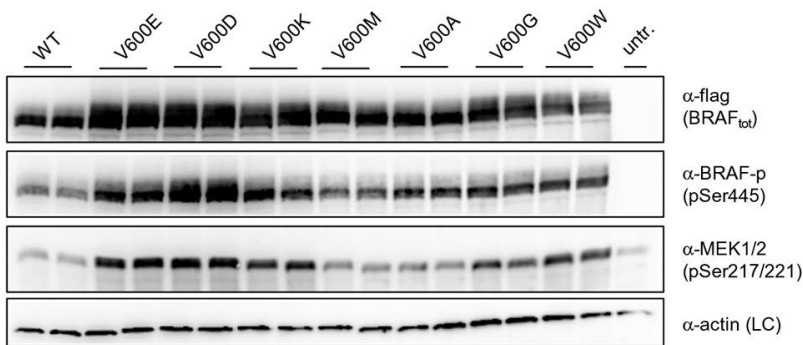
		Second Letter					
		U	C	A	G		
1st letter	U	UUU Phe UUC UUA Leu UUG	UCU UCC Ser UCA UCG	UAU Tyr UAC UAA Stop UAG Stop	UGU Cys UGC UGA Stop UGG Trp	U C A G	
	C	CUU CUC Leu CUA CUG	CCU CCC Pro CCA CCG	CAU His CAC CAA Gln CAG	CGU CGC Arg CGA CGG	U C A G	
	A	AUU AUC Ile AUA AUG Met	ACU ACC Thr ACA ACG	AAU 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

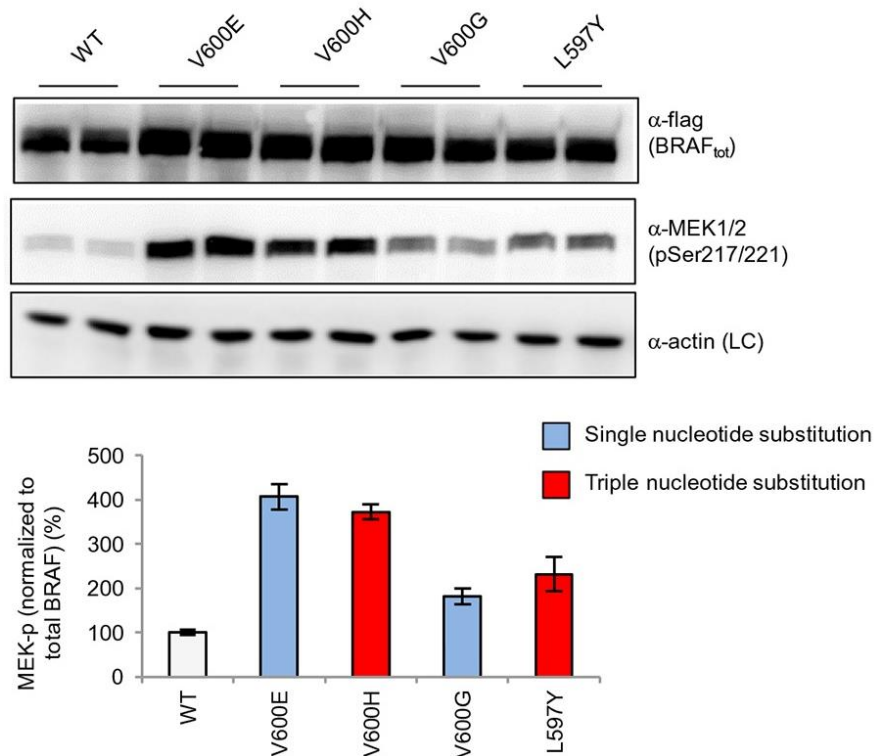
V600K: AAG
 V600R: AGG
 V600D: GAC/T

- 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)



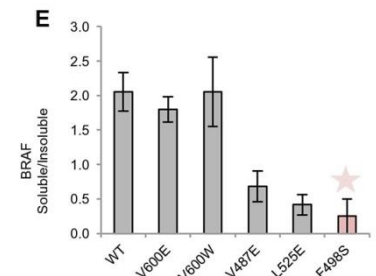
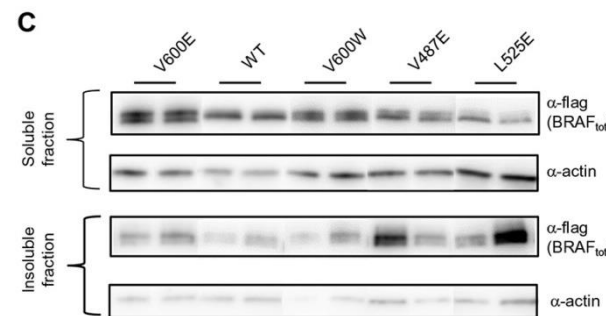
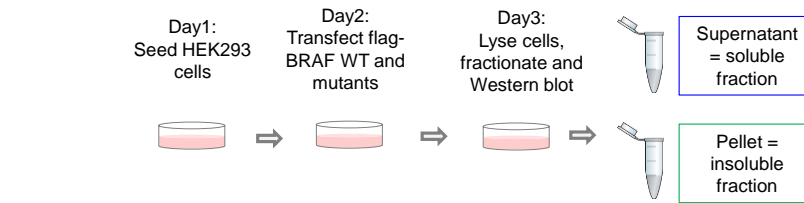
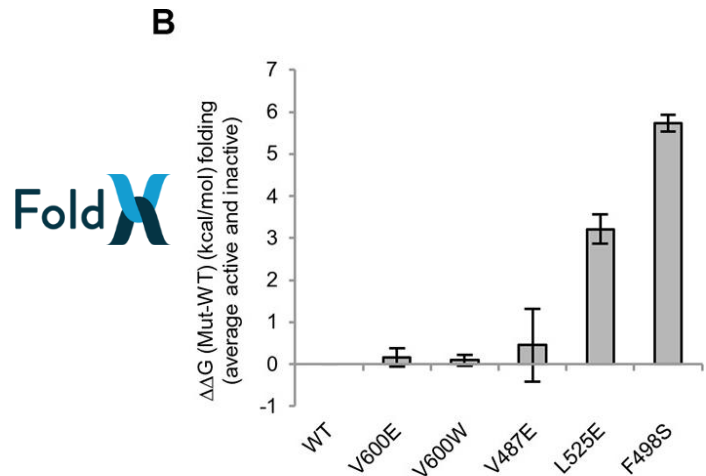
Experimentally validate the effect of BRAF mutations by monitoring downstream MEK activation (HEK293 cells)



- **V600H** (requires 3 nucleotide substitutions) is as active as V600E, but NOT found in cancer. Similarly L597Y is not found in rasopathy patients.

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)

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
- Structural analysis of mutations could suggest for different therapies for mutations happening at different regions of the protein
- In GWAS analysis the number of base changes required for a mutation should be considered in the analysis. Two mutations with the same frequency, one could be neutral and the other deleterious if the first one requires one base change and the second one, two.



Acknowledgements

Collaborations cell signaling team:

Walter Kolch, Boris Kholodenko, Katja Rybakova, Lan Nguyen (SBI Dublin)

Edda Klipp, Friedemann Uchner (Humboldt-Universität zu Berlin)

David Lynn, Kenneth Bryan (EMBL Australia)

Owen Sansom, Fatih Ceteci (Beatson Cancer Research, Glasgow)

Jörg Stelling (ETH Zürich)

Marius Ueffing, Karsten Boldt, (University of Tübingen)

Francesc Posas, Klement Stojanovski, Gerhard Seisenbacher (UPF Barcelona)

Ola Söderberg (Uppsala University)

Philippe Bastiaens, Aneta Koseska (MPI Dortmund)

Imre Berger (EMBL Grenoble)

Serrano lab

Luis Serrano

Cell signaling team

Violeta Beltran

Hannah Benisty

Tony Ferrar

Jae-Seong Yang

Martin Schaefer

Javier Delgado

Kiana Toufighi

Besray Unal

CRG core facilities

Proteomics Core Facility

Advanced Light Microscopy Unit

Protein Screening

FACS Core Facility

Bioinformatics Core Facility



Acknowledgements

Systems Biology Group

Christina Kiel
Hanna Benisty
Violeta Beltran
Martin Schiefer



erc

2009-2014



erc

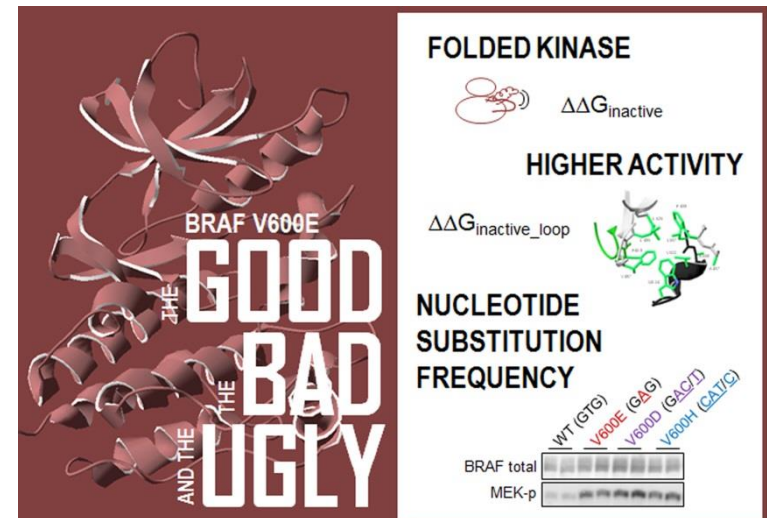
Conclusions example 3: Why BRAF V600E?

- 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.



Protein abundances

← → ↻ pax-db.org

 PaxDb: Protein Abundance Database Downloads Help Archives ▾ About

All Organisms ✕

protein id/name

Search

Browse species

All ▾ all organisms (56)



X. tropicalis



G. gallus



P. troglodytes



H. sapiens



C. fa

Data Overview

species name

Species ↕	Predicted proteome size ↕	Datasets ▲	Proteins Covered ↕
Homo sapiens	20457	170	98%
Mus musculus	22668	75	90%
Arabidopsis thaliana	27416	48	76%
Danio rerio	26163	20	59%
Escherichia coli str. K-12 substr. MG1655	4146	18	98%
Saccharomyces cerevisiae	6692	17	96%
Caenorhabditis elegans	20517	10	60%
Drosophila melanogaster	13937	10	95%
Schizosaccharomyces pombe	5144	8	90%
Zea mays	92413	7	18%

Affinities and kinetic constants



myBDB logout

Search and Browse

Target

Sequence

Name &

Ki IC50 Kd EC50

Rate constants

ΔG° ΔH° $-\Delta S^\circ$

pH (Enzymatic Assay)

pH (ITC)

Substrate or Competitor

Compound Mol. Wt.

Chemical Structure

Pathways

Source Organism

Number of Compounds

Monomer List in csv

Het List in SDF

Compound

FDA Drugs

Important Compounds

Chemical Structure

Name

SMILES

Number of Data / Targets

Special tools

3D Structure Series

Find My Compound's Targets

Find Compounds for My Targets

Do Virtual Screening

SCOP

Citation

Author

Journal/Citation

Institution

PubMed

PubChem BioAssay

US Patent

The Binding Database

Home Info Download About us Email us Contribute data Web Services

BindingDB is a public, web-accessible database of measured binding affinities, focusing chiefly on the interactions of protein considered to be drug-targets with small, drug-like molecules. BindingDB contains 1,207,821 binding data, for 6,265 protein targets and 529,618 small molecules.

There are 2907 protein-ligand crystal structures with BindingDB affinity measurements for proteins with 100% sequence identity, and 7392 crystal structures allowing proteins to 85% sequence identity.

Simple Search

Article Titles, Authors, Assays, Compound Names, Target Names

Use ? for single-letter wild-card or * for general wild-card. For example, "adeny?" or "adeny?". Query cannot start with wild card.

Advanced Search

Combine multiple search criteria, such as chemical structures, target names, and numerical affinities; restrict searches by data source, such as BindingDB, ChEMBL, PubChem, and Patents.

Messages

BindingDB's Advanced Search now allows you to download your search results in Excel format. (March 2016)

We are delighted to announce that Elsevier's Science Direct journals now include links from articles to BindingDB datasets, where available! For an example, go to this article, and see the "Data for this Article, BindingDB" link on the right. (December 2015)

Journal Curation by BindingDB

BindingDB continually curates a set of journals not covered by other public databases. As of January 2016, the status of our current curation effort is as follows:

- [ACS Chemical Biology](#) 2006-2015 (vol 1-10)
- [ACS BioChemistry](#) 1962-1970 (vol 1-9), 1991-2015 (vol 30-54)
- [Bioorganic Chemistry](#) 1971-2015 (vol 1-62)
- [BMC Chemical Biology](#) 2001-2012 (vol 1-12)
- [ChemBioChem](#) 2000-2015 (vol 1-16)
- [Chemical Biology & Drug Design](#) 2006-2015 (vol 67-86)
- [Chemistry & Biology](#) 1994-2014 (vol 1-20)
- [Journal of Biological Chemistry](#) 1988-2013 (vol 264-288)
- [Journal of Chemical Biology](#) 2008-2013 (vol 1-6)
- [Journal of Enzyme Inhibition and Medicinal Chemistry](#) 1997-2009 (vol 11-24)
- [Nature Chemical Biology](#) 2005-2014 (vol 1-10)
- [Medicinal Chemistry Research](#) 2004-2013 (vol 13-22)

myBDB

Username Password
 Username is your registered email in BindingDB.

Video Tutorials



BindingDB News

September 2015. We to post the latest BindingDB user survey or about October 5. ¹ would greatly appreciate your feedback and suggestions! Note, however, that you are always free to click the survey to the registration page.

September 2015. All compounds in BindingDB have now been assigned a BindingDB Molecule ID such as BDBM5018 (The numeric compound ID also known as internal Monomer ID.)

September 2015. BindingDB should give you faster performance if we have upgraded the server. Please let us know immediately if you are having any problems.

July 2015. Please try our new tool to map from one or more proteins of interest to known potential ligands. Find Compounds for My Targets

April 2015. BindingDB has improved security. We use SSL to transmit passwords securely, forgotten passwords are now handled with a link.

March 2015. The BindingDB results table now provide links for protein targets to articles in Antibodypedia, and ligands to UniChem.

February 2015. Full has been replaced by Simple Search, with greatly improved display.

General 'numbers' in biology

B10NUMB3R5 THE DATABASE OF USEFUL BIOLOGICAL NUMBERS

[Home \ Search](#)
[Browse](#)
[Resources](#)
[Cell Biology by the Numbers](#)
[About Us](#)
[Login \ Submit](#)

Key Numbers for Cell Biologists

Cell size

1. Bacteria (*E. coli*): $\approx 0.7\text{-}1.4\ \mu\text{m}$ diameter, $\approx 2\text{-}4\ \mu\text{m}$ length, $\approx 0.5\text{-}5\ \mu\text{m}^3$ in volume; $10^8\text{-}10^9$ cells/ml for culture with $\text{OD}_{600} \approx 1$
2. Yeast (*S. cerevisiae*): $\approx 3\text{-}6\ \mu\text{m}$ diameter, $\approx 20\text{-}160\ \mu\text{m}^3$ in volume
3. Mammalian cell volume: $100\text{-}10000\ \mu\text{m}^3$; HeLa: $500\text{-}5000\ \mu\text{m}^3$ (adherent on slide $\approx 15\text{-}30\ \mu\text{m}$ diameter)

Length Scales Inside Cells

4. Nucleus volume $\approx 10\%$ of cell volume
5. Cell membrane thickness $\approx 4\text{-}10\ \text{nm}$
6. "Average" protein diameter $\approx 3\text{-}6\ \text{nm}$
7. Base pair: $2\ \text{nm}$ (D) x $0.34\ \text{nm}$ (H)
8. Water molecule diameter $\approx 0.3\ \text{nm}$

Division, Replication, Transcription, Translation & Degradation Rates

- at 37°C with a temperature dependence Q_{10} of $\approx 2\text{-}3$
9. Cell cycle time (exponential growth in rich media): *E. coli* $\approx 20\text{-}40\ \text{min}$; yeast $70\text{-}140\ \text{min}$; human cell line (HeLa): $15\text{-}30\ \text{hours}$
 10. Rate of replication by DNA polymerase *E. coli* $\approx 200\text{-}1000$ bases/s; human ≈ 40 bases/s. Transcription by RNA polymerase $10\text{-}100$ bases/s
 11. Translation rate by ribosome $10\text{-}20$ aa/s
 12. Degradation rates (proliferating cells): mRNA half life $<$ cell cycle time; protein half life \approx cell cycle time

Concentration

13. Concentration of $1\ \text{nM}$ in: *E. coli* is ≈ 1 molecule/cell; HeLa $\approx 1,000$ molecules/cell
14. Characteristic concentration for a signaling protein $\approx 10\ \text{nM}\text{-}1\ \mu\text{M}$
15. Water content: $\approx 70\%$ by mass; General elemental composition (dry weight) of *E. coli*: $\approx \text{C}_4\text{H}_7\text{O}_2\text{N}$; Yeast $\approx \text{C}_6\text{H}_{10}\text{O}_2\text{N}$
16. Composition of *E. coli* (dry weight): $\approx 55\%$ protein, 20% RNA, 10% lipids, 15% others
17. Protein conc. $\approx 100\ \text{mg/ml} = 3\ \text{mM}$. $10^8\text{-}10^7$ per *E. coli* (depending on growth rate); Total metabolites (MW $< 1\text{kD}$) $\approx 300\ \text{mM}$

Energetics

18. Membrane potential $\approx 70\text{-}200\ \text{mV}$ $\rightarrow 2\text{-}6\ k_B T$ per electron ($k_B T$ = thermal energy)
19. Free energy (ΔG) of ATP hydrolysis under physiological conditions $\approx 40\text{-}60\ \text{kJ/mole}$ $\rightarrow \approx 20 k_B T$ /molecule ATP; ATP molecules required to make an *E. coli* cell $\approx 10\text{-}50 \times 10^9$
20. ΔG° resulting in order of magnitude ratio between products and reactants concentrations: $\approx 6\ \text{kJ/mol} = 60\ \text{meV} \approx 2\ k_B T$

Useful biological numbers extracted from the literature. Numbers and ranges should only serve as "rule of thumb" values. References are in the online annotated version at the BioNumbers website. Consult website and original references to learn about the details of the system under study including growth conditions, method of measurement, etc.

Diffusion and Catalysis Rate

21. Diffusion coefficient for an "average" protein: in cytoplasm $D \approx 5\text{-}15\ \mu\text{m}^2/\text{s}$ $\rightarrow \approx 10$ millisecond to traverse an *E. coli* $\rightarrow \approx 10$ s to traverse a mammalian (HeLa) cell; small metabolite in water $D \approx 500\ \mu\text{m}^2/\text{s}$
22. Diffusion limited on-rate for characteristic protein $\approx 10^8\text{-}10^9\ \text{s}^{-1}\text{M}^{-1}$ \rightarrow for a protein substrate of concentration $\approx 1\ \mu\text{M}$ the diffusion limited on-rate is $\approx 100\text{-}1000\ \text{s}^{-1}$ thus limiting the catalytic rate k_{cat}

Genome sizes & Error Rates

23. Genome size: *E. coli* $\approx 5\ \text{Mbp}$; *S. cerevisiae* (yeast) $\approx 12\ \text{Mbp}$; *C. elegans* (nematode) $\approx 100\ \text{Mbp}$; *D. melanogaster* (fruit fly) $\approx 120\ \text{Mbp}$; *A. thaliana* (arabidopsis) $\approx 120\ \text{Mbp}$; *M. musculus* (mouse) $\approx 2.5\ \text{Gbp}$; *H. sapiens* (human) $\approx 2.9\ \text{Gbp}$; *T. aestivum* (wheat) $\approx 16\ \text{Gbp}$
24. Number of protein-coding genes: *E. coli* $\approx 4,000$; *S. cerevisiae* $\approx 6,000$; *C. elegans*, *A. thaliana*, *M. musculus*, *H. sapiens* $\approx 20,000$
25. Mutation rate in DNA replication $\approx 10^{-8}\text{-}10^{-10}$ per bp
26. Misincorporation rate: transcription $\approx 10^{-4}$ per nucleotide; translation $\approx 10^{-2}\text{-}10^{-1}$ per amino-acid

Click on a number to see full description and reference
www.BioNumbers.org

Protein structures

RCSB PDB Deposit Search Visualize Analyze Download Learn More MyPDB Login

RCSB PDB An Information Portal to 118087 Biological Macromolecular Structures

Search by PDB ID, author, macromolecule, sequence, or ligand **Go**

Advanced Search | Browse by Annotations

PDB-101 PDB EMDatabank NUCLEIC ACID DATABASE StructuralBiology Knowledgebase

f t y a g

Welcome

Deposit

Search

Visualize

Analyze

Download

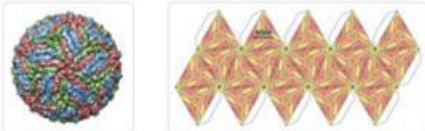
Learn

A Structural View of Biology

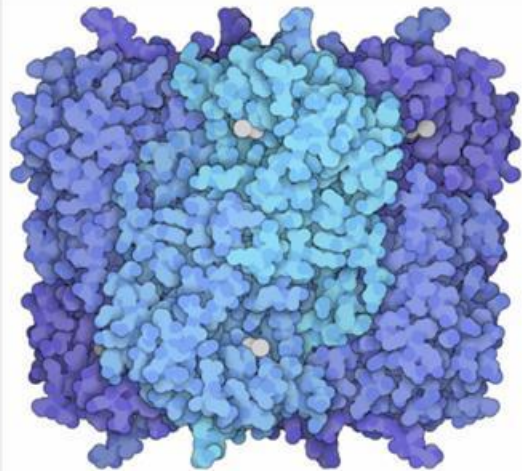
This resource is powered by the Protein Data Bank archive-information about the 3D shapes of proteins, nucleic acids, and complex assemblies that helps students and researchers understand all aspects of biomedicine and agriculture, from protein synthesis to health and disease.

The RCSB PDB builds upon the data by creating tools and resources for research and education in molecular biology, structural biology, computational biology, and beyond.

Zika Virus Structure




April Molecule of the Month



Lead Poisoning

3D structures of protein interactions



Interactome3D

Home Download References Statistics News Tutorials Help About

Interactome3D is a web service for the structural annotation of protein-protein interaction networks. Submit your interactions and the server will find all the available structural data for both the single interactors and the interactions themselves. Additionally you can also visualize and download structural information for interactions involving a set of proteins or interactomes for one of the precalculated organisms.

If you have any doubts read our section of [Frequently Asked Questions](#).

The current version of Interactome3D is **2015_12** [Release notes](#)

Submit your interactions

Enter a name for your dataset: [Example](#)

Enter a list of interactions (max. 10000). Every interaction has to be entered in a separate line, as a pair of space-separated Uniprot ACs (*): [?](#)

For example...
 A0A5B9 A0A5B9
 A0A5B9 P01848
 A0AQH0 O61443

...or upload your interactions from a file: [?](#)

Email (*):

Tutorial
 Learn how to use Interactome3D

Query interactions with proteins

Enter a list of Uniprot ACs (*) or gene names: [?](#)

For example...
 A0A5B9
 P01848
 A0AQH0
 O61443


Only show the proteins in the list [?](#)

Browse for organism

Select one of the pre-calculated organisms:

- > Arabidopsis thaliana
- > Bacillus subtilis
- > Bos taurus
- > Caenorhabditis elegans
- > Campylobacter jejuni
- > Mus musculus
- > Mycobacterium tuberculosis
- > Mycoplasma pneumoniae
- > Plasmodium falciparum
- > Rattus norvegicus
- > Saccharomyces cerevisiae
- > Schizosaccharomyces pombe

3D structures of protein interactions/ mapping of disease mutations




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 your 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...
ETF A, ETF B, 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

Protein design



[PRODUCTS ▾](#)
[LICENSING AND SERVICES ▾](#)
[SUPPORT ▾](#)
[DOWNLOAD ▾](#)
[ABOUT ▾](#)

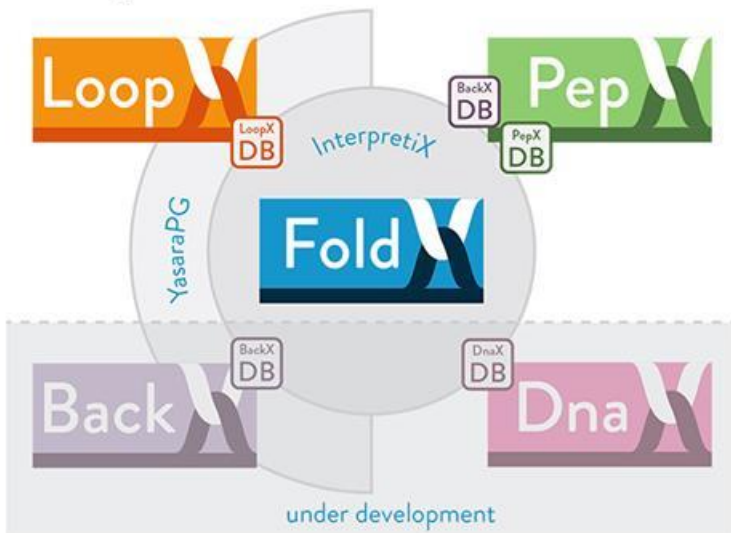


[LOGIN](#) [REGISTER](#)

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):0935-9. 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.

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

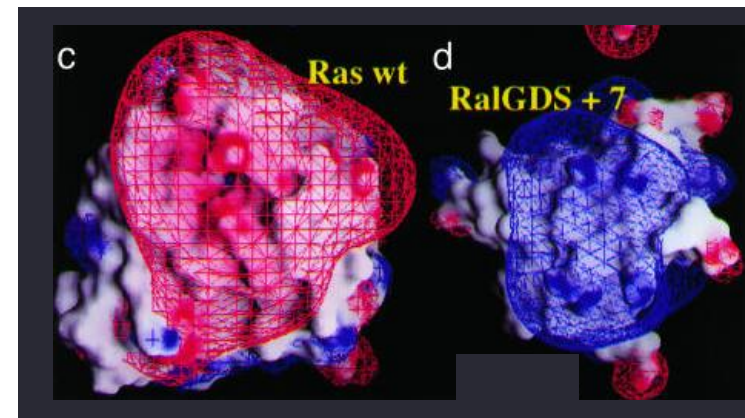
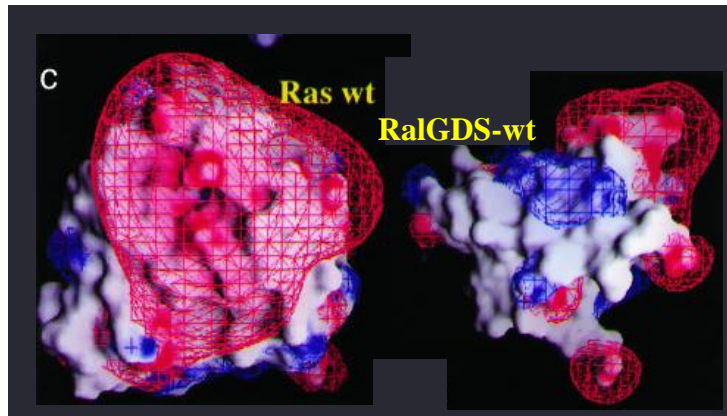
Experimental design of mutants that introduce kinetic perturbations

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

k_{off} ← Dissociation rate constant
 k_{on} ← Association rate constant

E.g.:

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



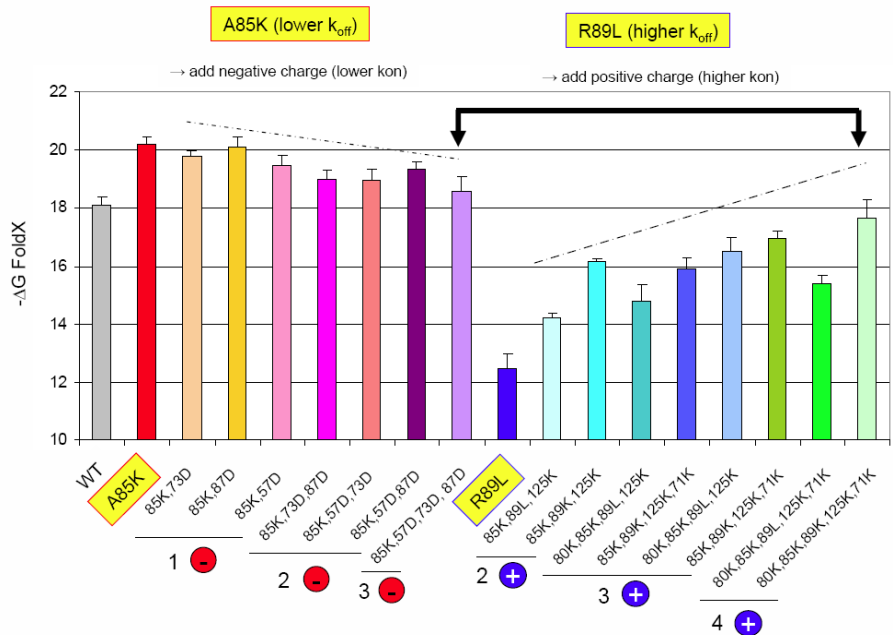
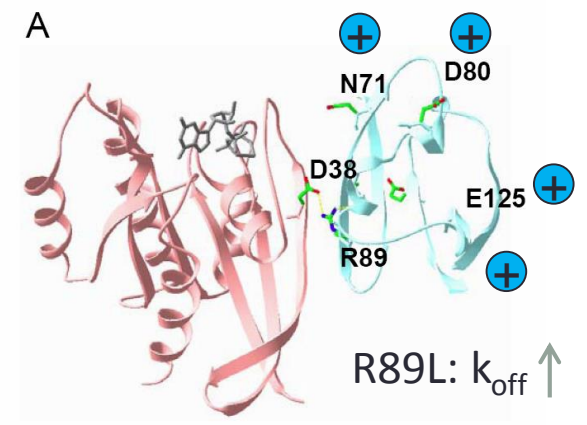
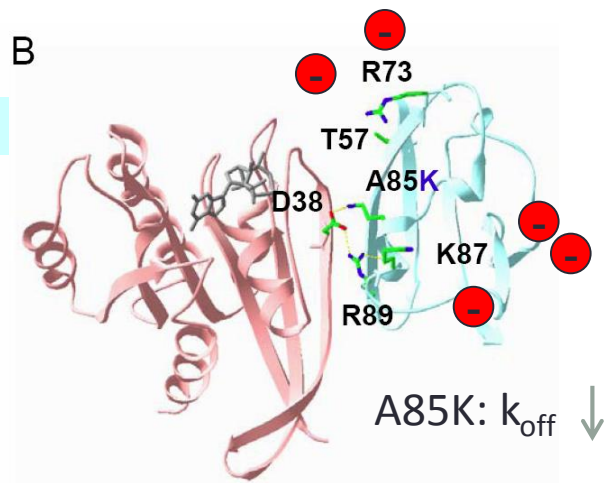
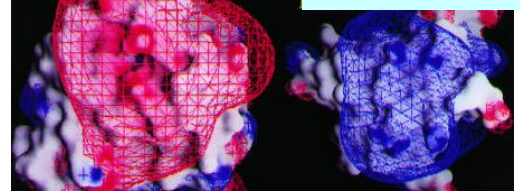
Kiel et al., PNAS, 2004

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

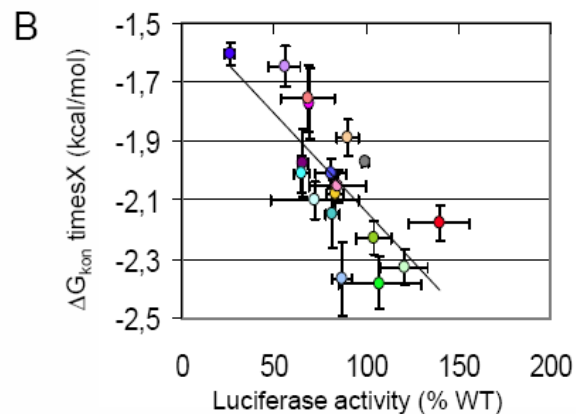
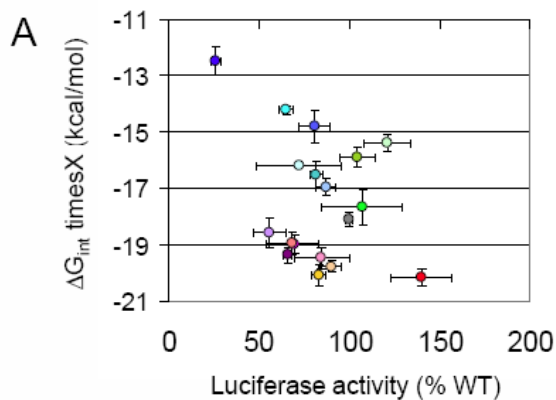
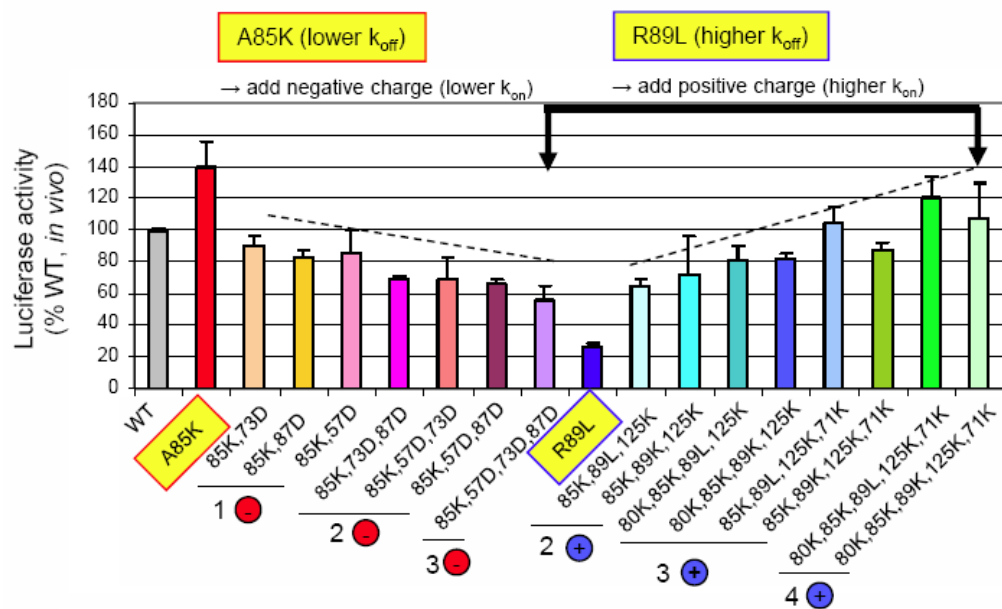
Summary of the protein mutant design

Ras surface negative

Raf surface positive

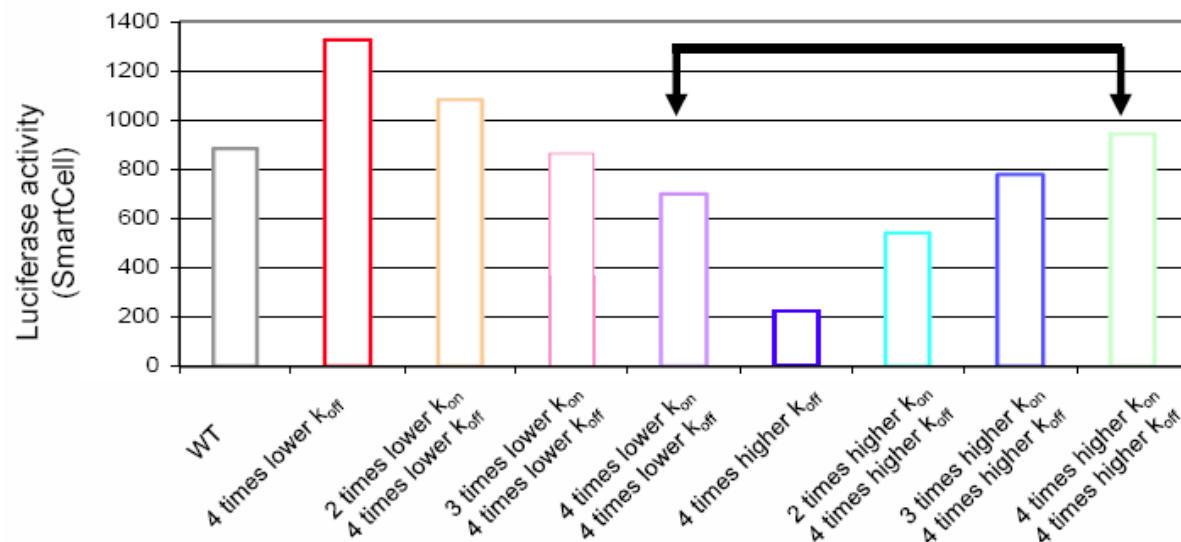


Analysis of all mutants in RK13 cells (luciferase activity assay)



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

Results from 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