Structural Studies of Signalling Proteins

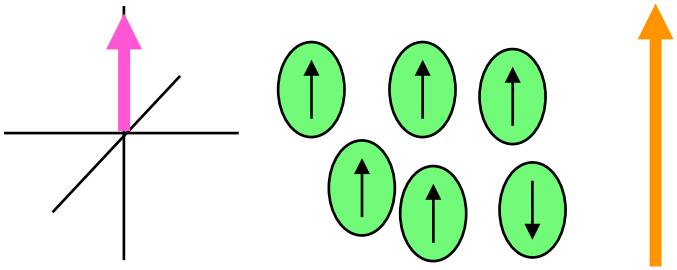
Using CamGRID to Calculate Protein Structures from NMR Data

Helen Mott

Department of Biochemistry
University of Cambridge

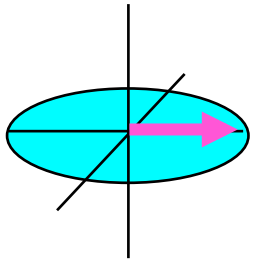
Nuclear Magnetic Resonance

bulk magnetization

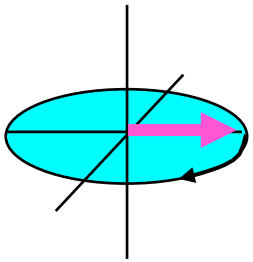


magnetic field

Nuclei with spin (e.g. ^{1}H) align (mostly) with the applied magnetic field



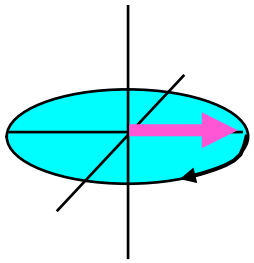
Application of a "pulse" tips the bulk magnetization by 90˚



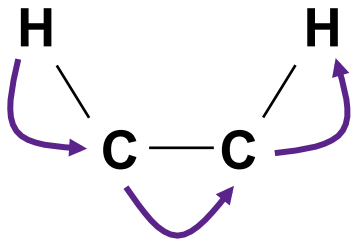
Magnetization vectors then rotate in this plane at a frequency that depends on the chemical environment of each nucleus

Each nucleus in a protein is in a different environment so a frequency can be assigned to each ^{1}H.

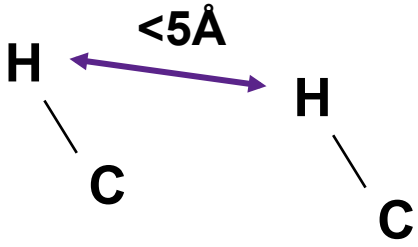
Nuclear Magnetic Resonance



Nuclei can interact with each other:



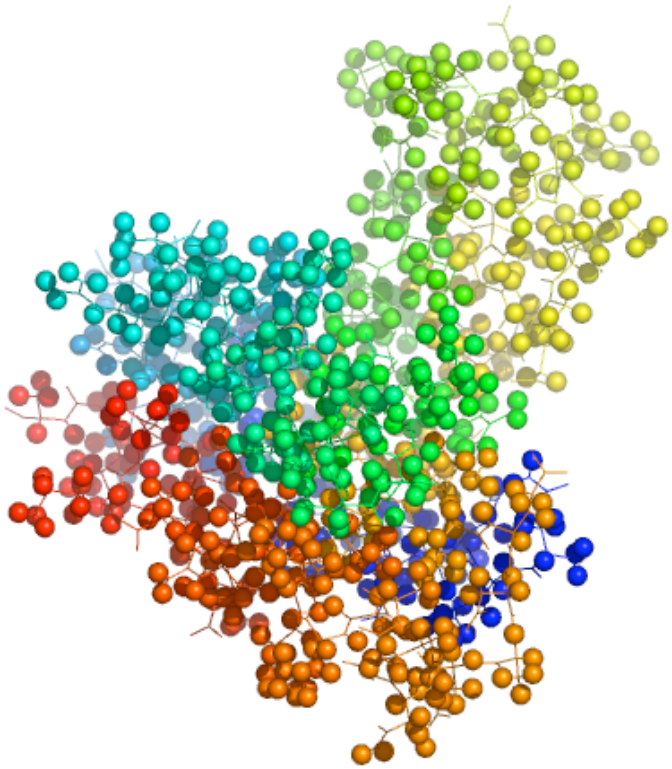
Through bonds (scalar coupling)

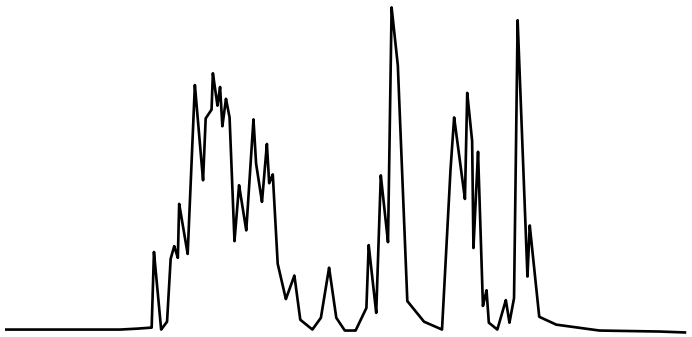


Through space (nuclear Overhauser effect - NOE)
- distance restraints

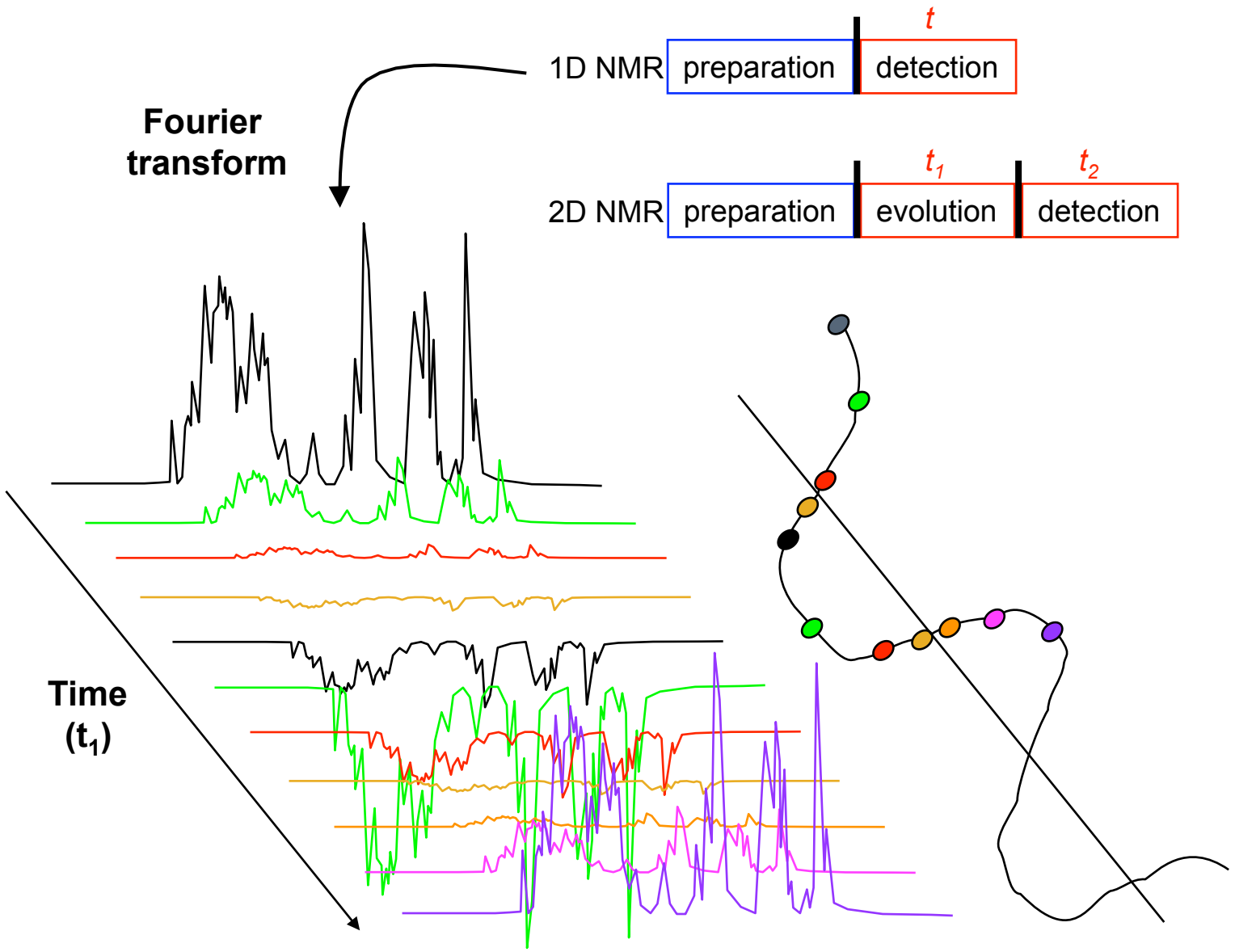
Network of distance restraints (NOEs) leads to structures

Even Small Proteins Contain too Many Hydrogens

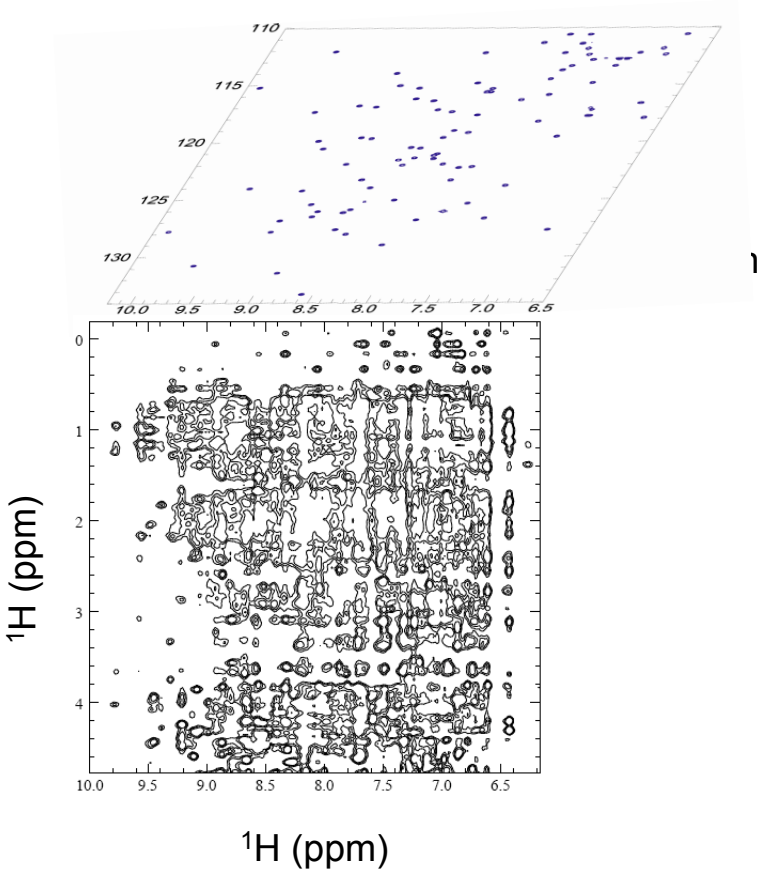


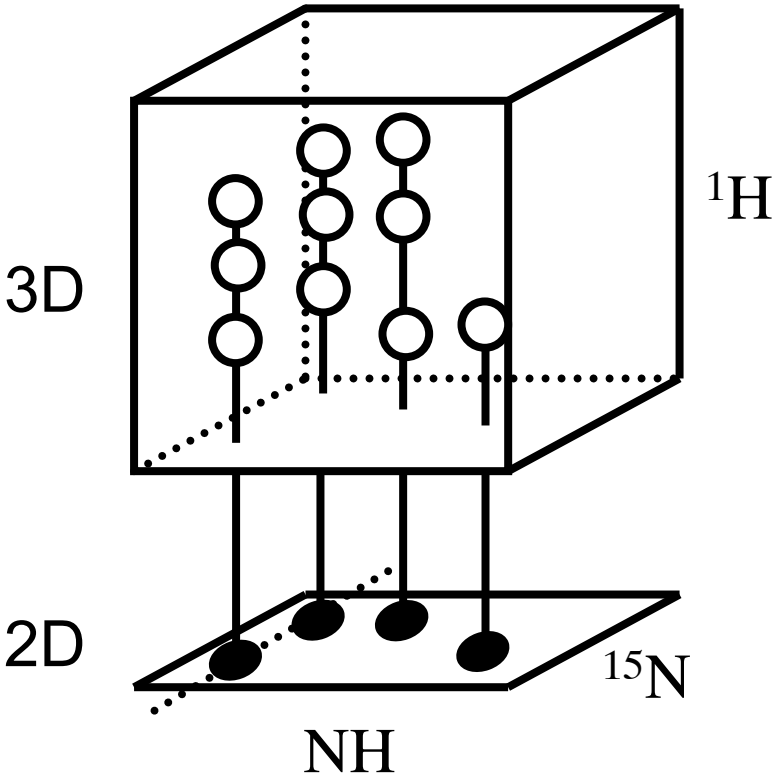


Two-dimensional NMR

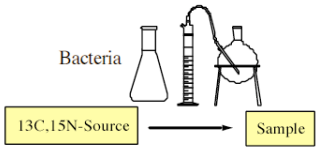


Three-dimensional NMR





Structure Determination by NMR



1.) sample preparation

sample

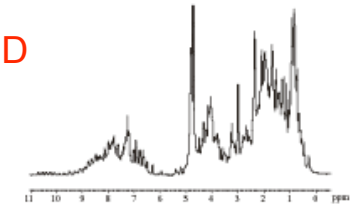
NMR spectrometer
(magnet + console)

pulse sequence

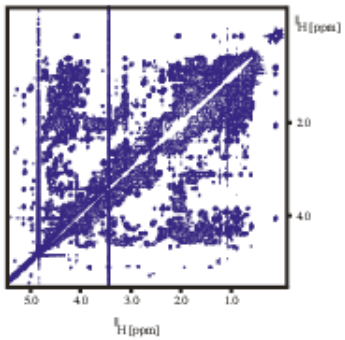
2.) data collection

3.) data processing

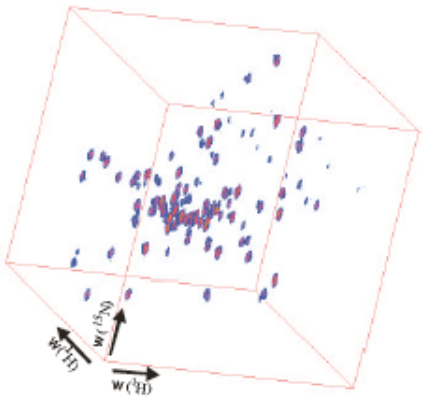
1D



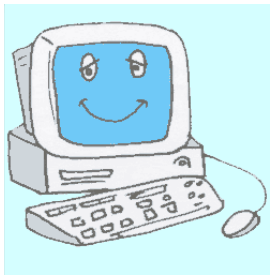
2D



3D

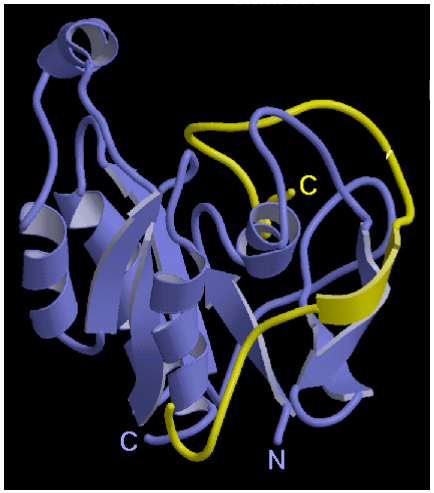


4.) assignment
& analysis



computer

5.) structure
calculation



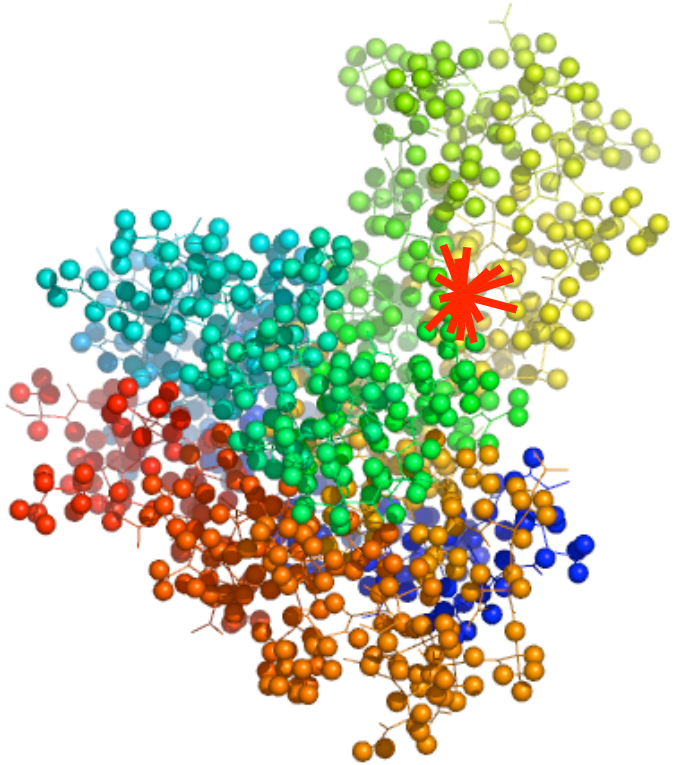
3D structure

Structure Calculation – from NOE to Structural Ensemble

Peak intensities are measured and are calibrated against known distances to derive proton/proton distance constraints (NOE is proportional to $1/ r^6$).

Upper distance limit for NOEs is about 5Å

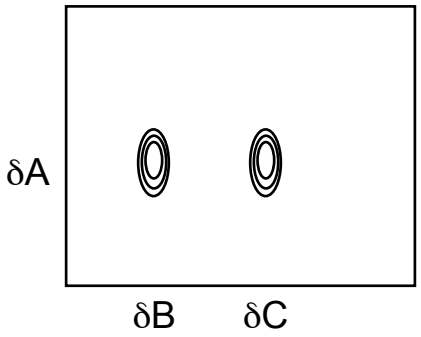
Different or random structure starting points are used to obtain ensemble of calculated structures which are consistent with the experimental data



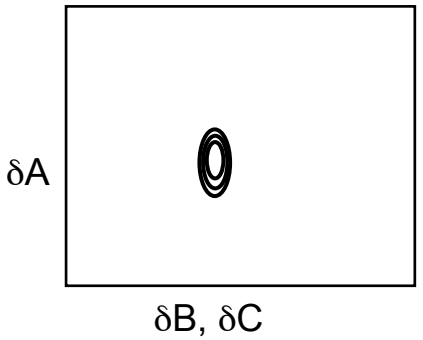
Even a small protein contains several hundred hydrogen nuclei

Ambiguity in the NMR Data

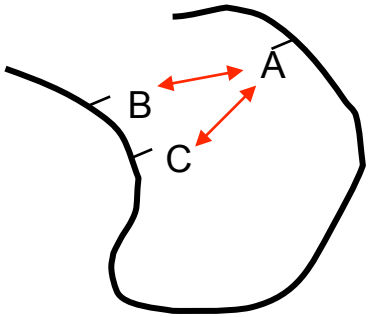
Overlap in through-space spectra

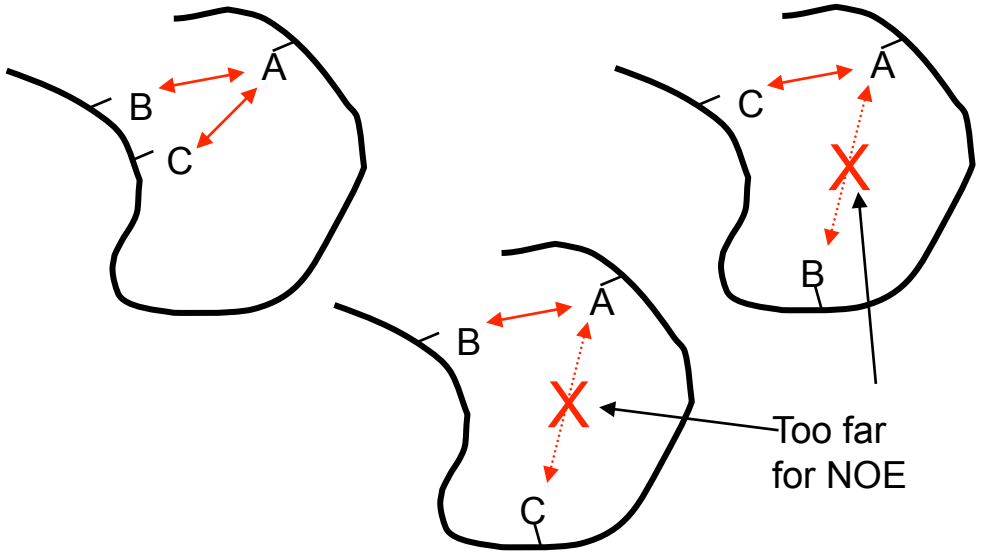


NOEs A-B and A-C can be assigned if the positions of B and C peaks are distinct.



If the position of B and C peaks are the same these possibilities cannot be distinguished.





Ambiguity in the NMR Data

Distance restraints are treated as ambiguous i.e. each is a sum of contributions:

$$\overline{D} \equiv \left(\sum_{a=1}^{N\delta} d_a^{-6} \right)^{-1/6}$$

Where $\overline{D}$ is the effective distance restraint and the individual contributions are d_a.

The structures are calculated using these restraints and the contribution of each possibility is then ranked. Possibilities that contribute little to the peak intensity are discarded.

The structures are then calculated again with the new set of restraints and the analysis is repeated.

The cutoff for the contributions is more stringent with each iteration, thus the ambiguity of the restraints is decreased.

Calculation of three-dimensional structures

Search conformational space for low energy:

Molecular dynamics simulated annealing from random structures using torsion angle dynamics.

Only angles around bonds are allowed to move during dynamics (computationally more efficient)

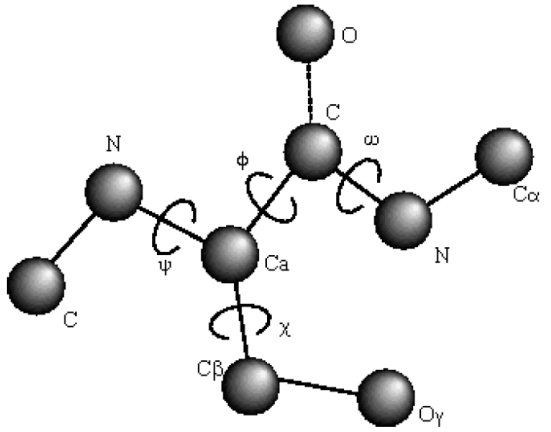
High temperature torsion angle dynamics, followed by slow cooling with Cartesian dynamics (i.e. all atoms are now allowed to move)

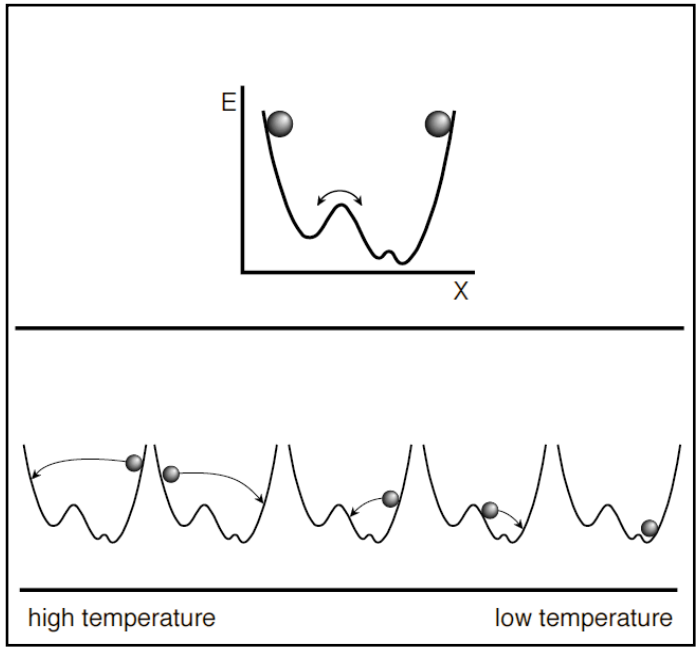
Local energy barriers are overcome by the high temperatures.

Structure calculation is performed using CNS
http://cns-online.org/v1.2/

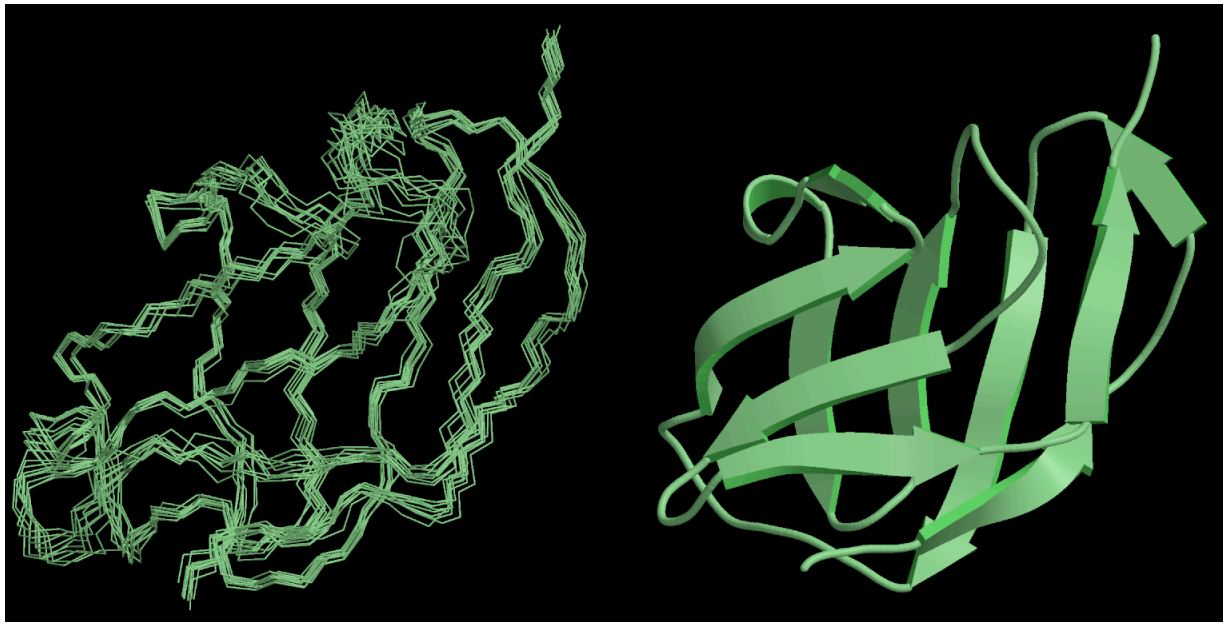
Interfaced with ARIA, which handles all the data using Python
http://aria.pasteur.fr/

Use CamGRID for structure calculations - 9 iterations of 20 structures each takes about 24 hours for a 300 residue protein

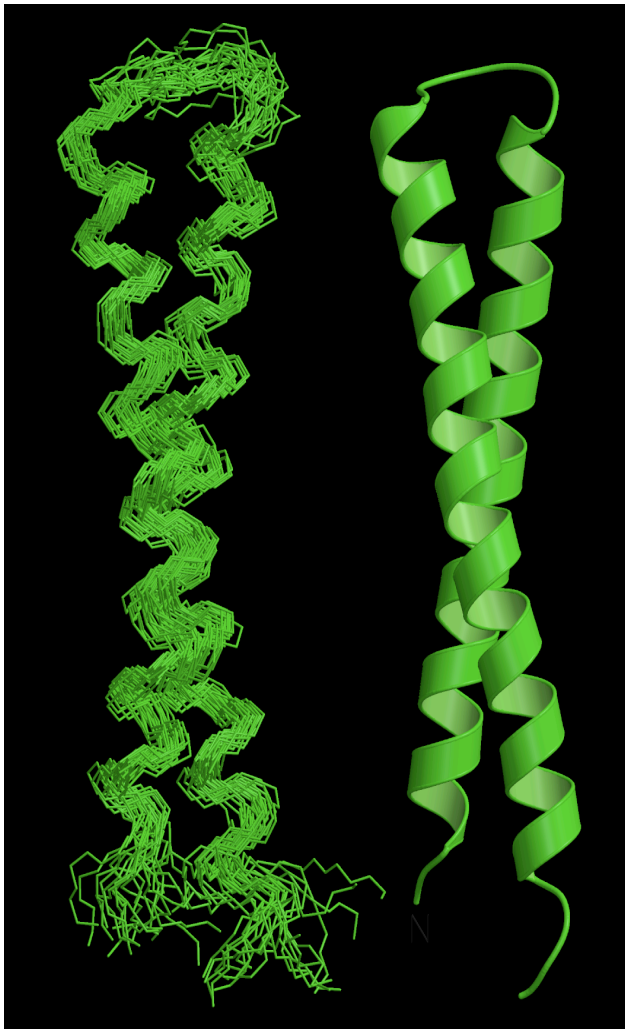




NMR Structures are Ensembles Consistent with the Data

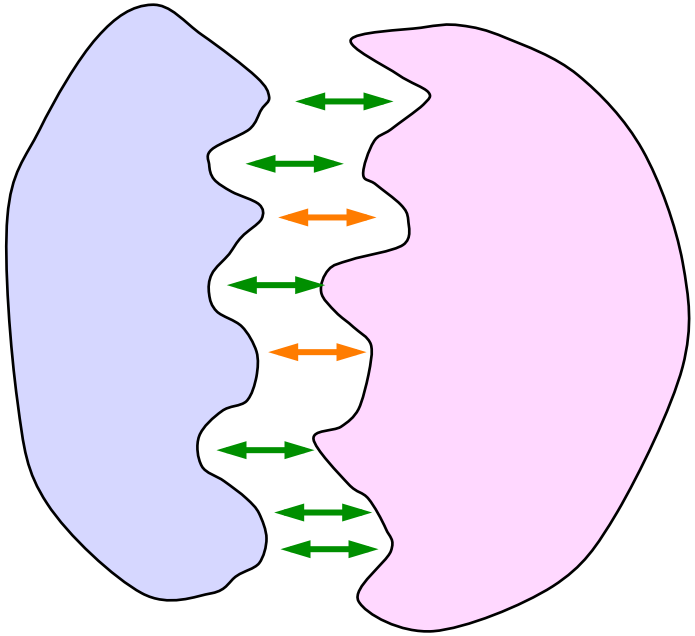


Sec5 - all β-sheet



HR1b - all α-helix

Protein-protein Interfaces

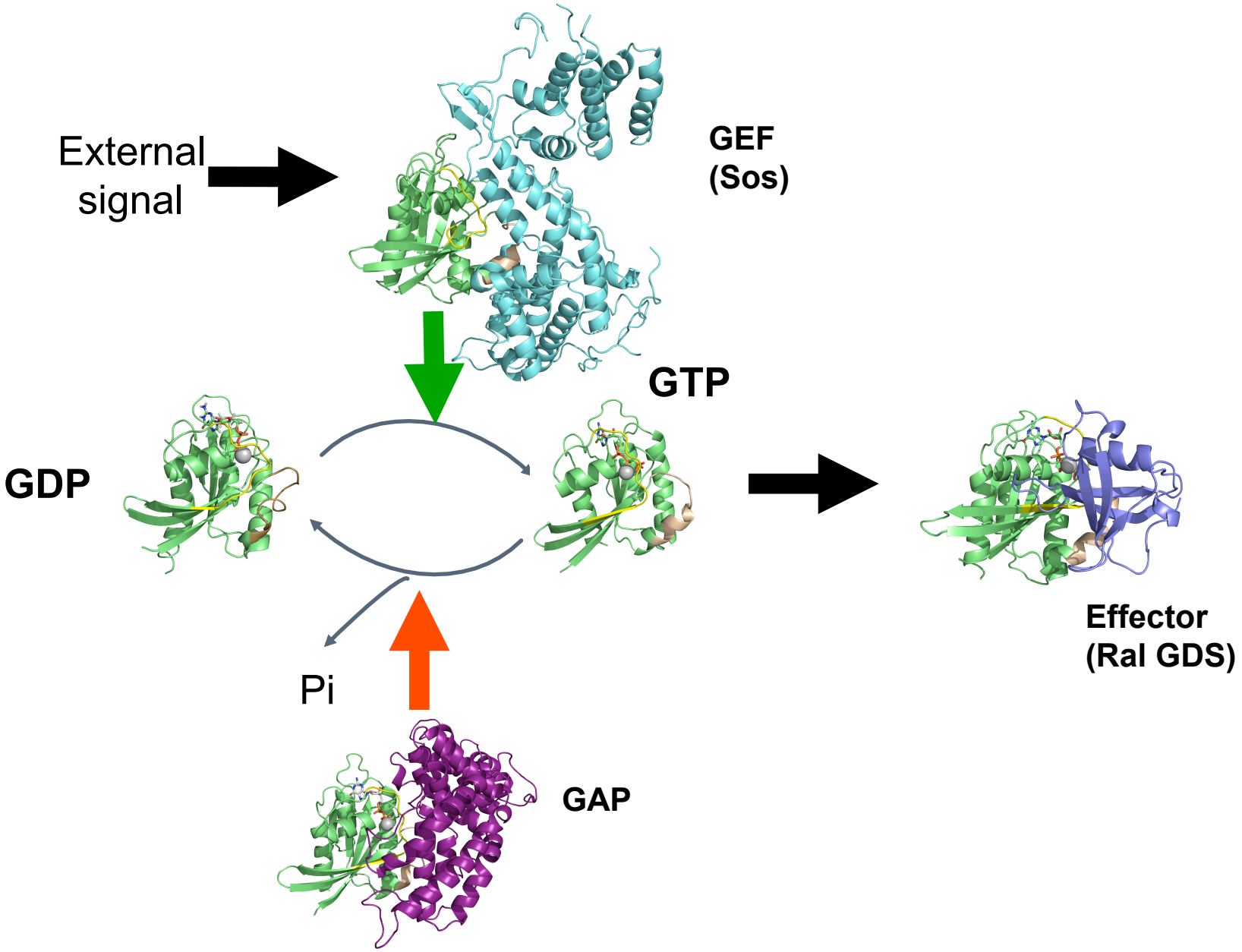


Prediction of protein structures is possible if a homologue is known but interfaces are harder to predict

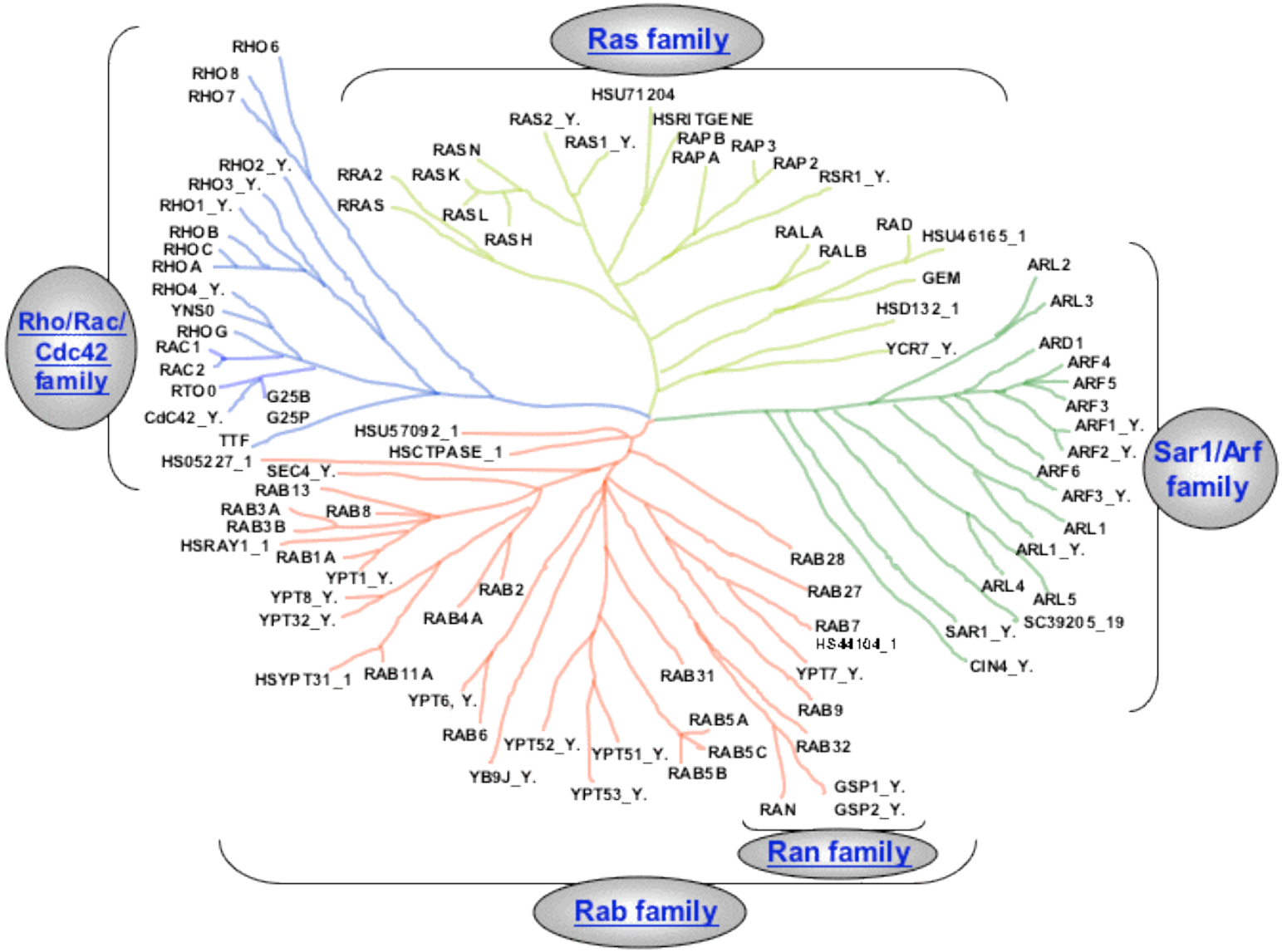
Proteins interact through large, flat surfaces using multiple contacts

Traditionally considered a difficult target for drug design but "hot spots" may define important interactions

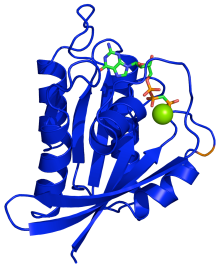
Small G Proteins are Molecular Switches



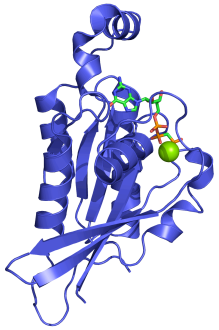
The Ras Superfamily Includes Five Groups of Proteins



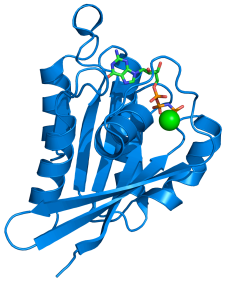
The Ras Superfamily, their Effectors and Effects



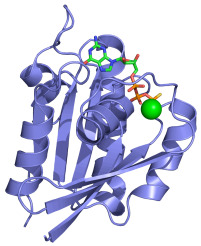
Ras



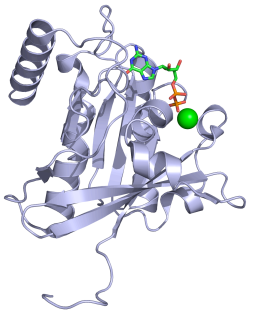
Rho



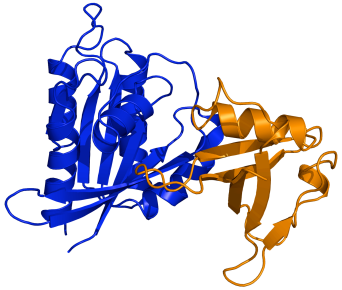
Rab



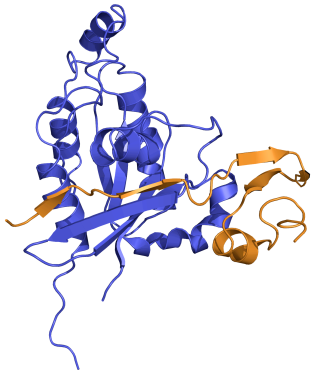
Arf



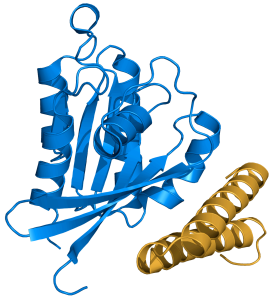
Ran



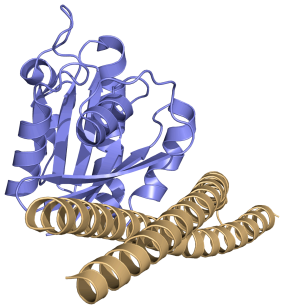
Cell growth and differentiation

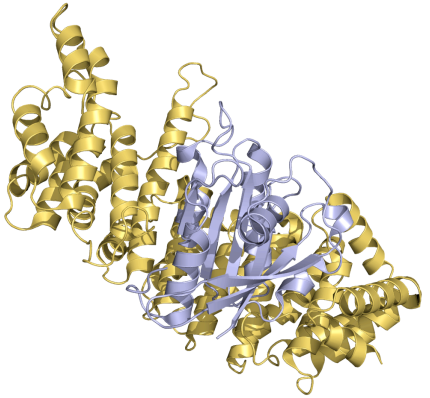


Cytoskeletal organization



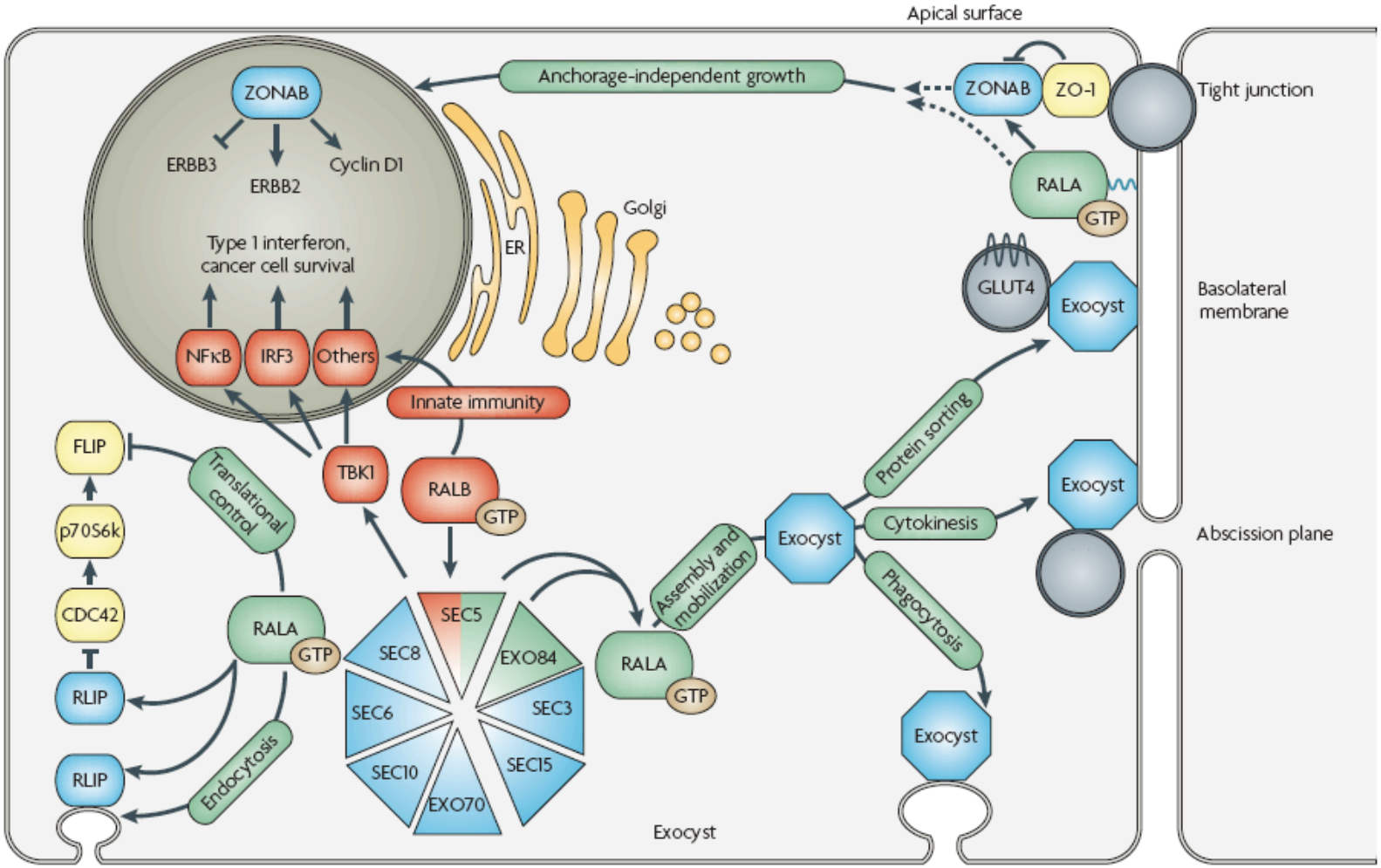
Intracellular vesicle trafficking



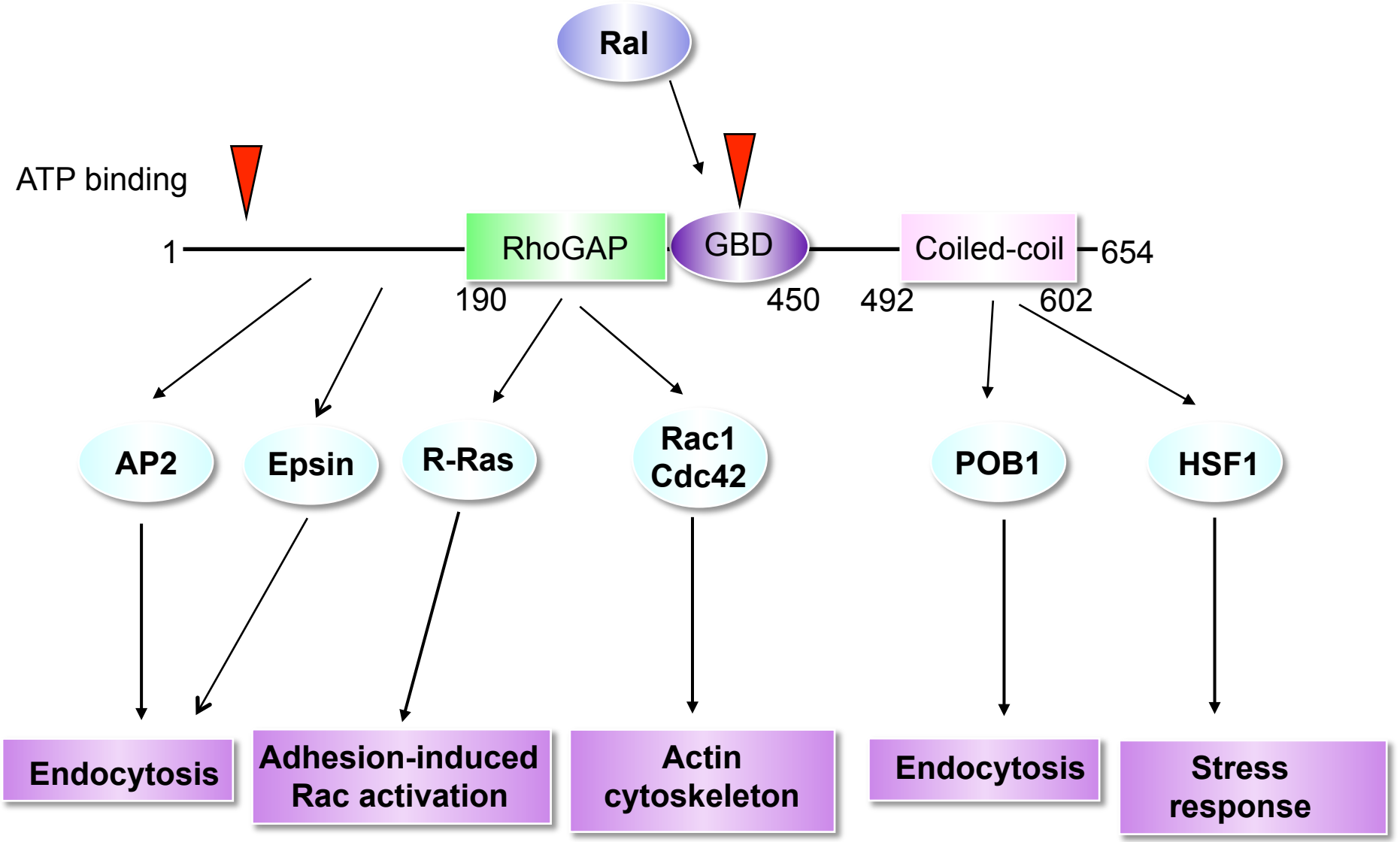


Nucleo-cytoplasmic transport

Ral is a Ras Family Member Involved in Multiple Cellular Processes

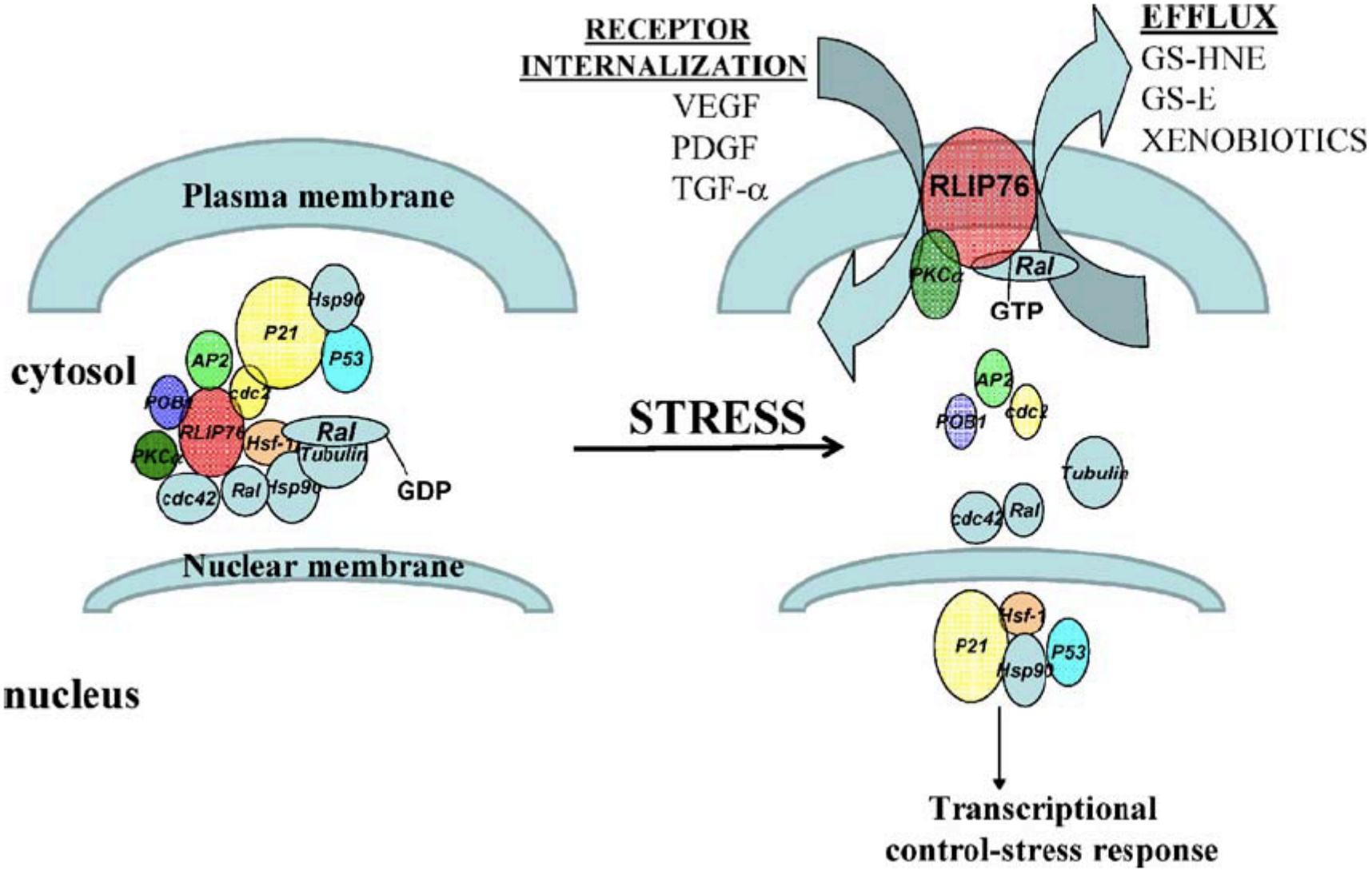


RLIP76 is a Multidomain Ral Effector

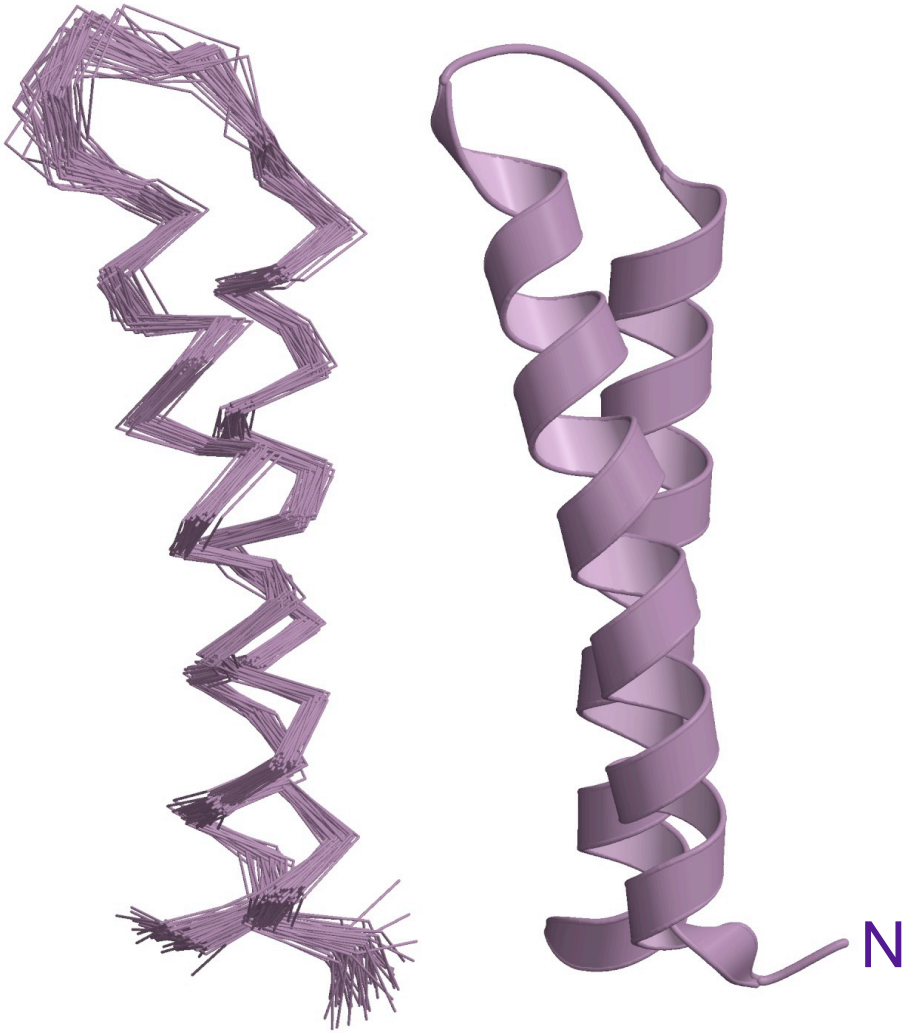


RLIP76 is a Transporter for Toxins and Metabolites in Response to Stress

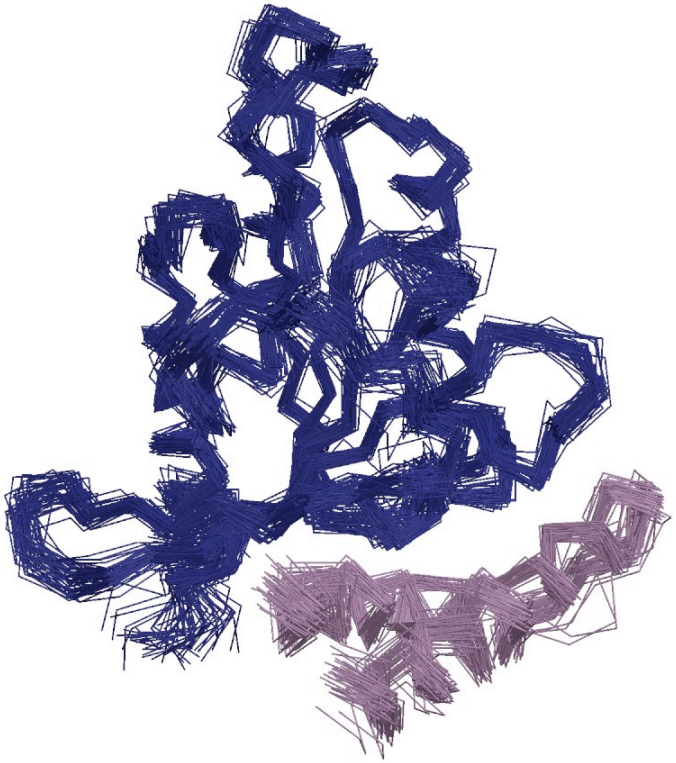
R. Vatsyayan et al./Biochemical Pharmacology 79 (2010) 1699–1705

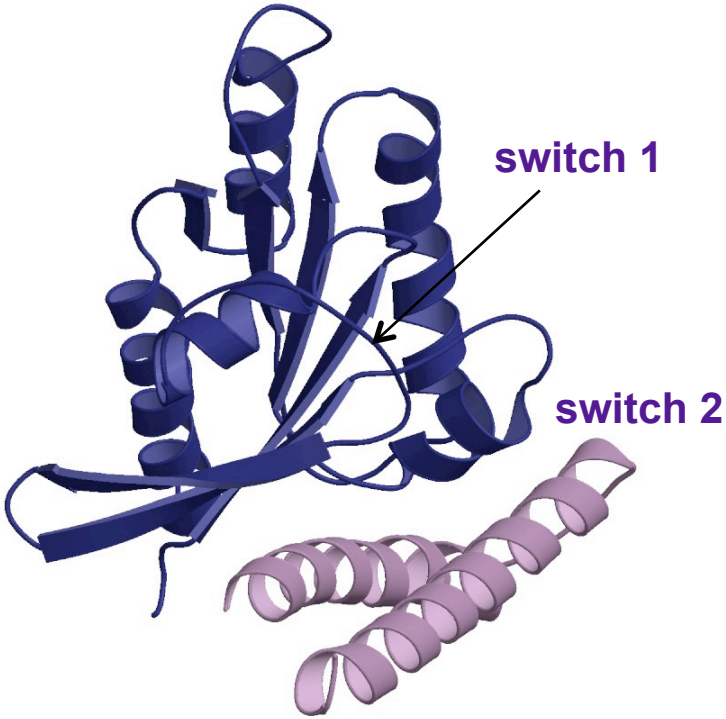


Ral Binding Domain of RLIP76



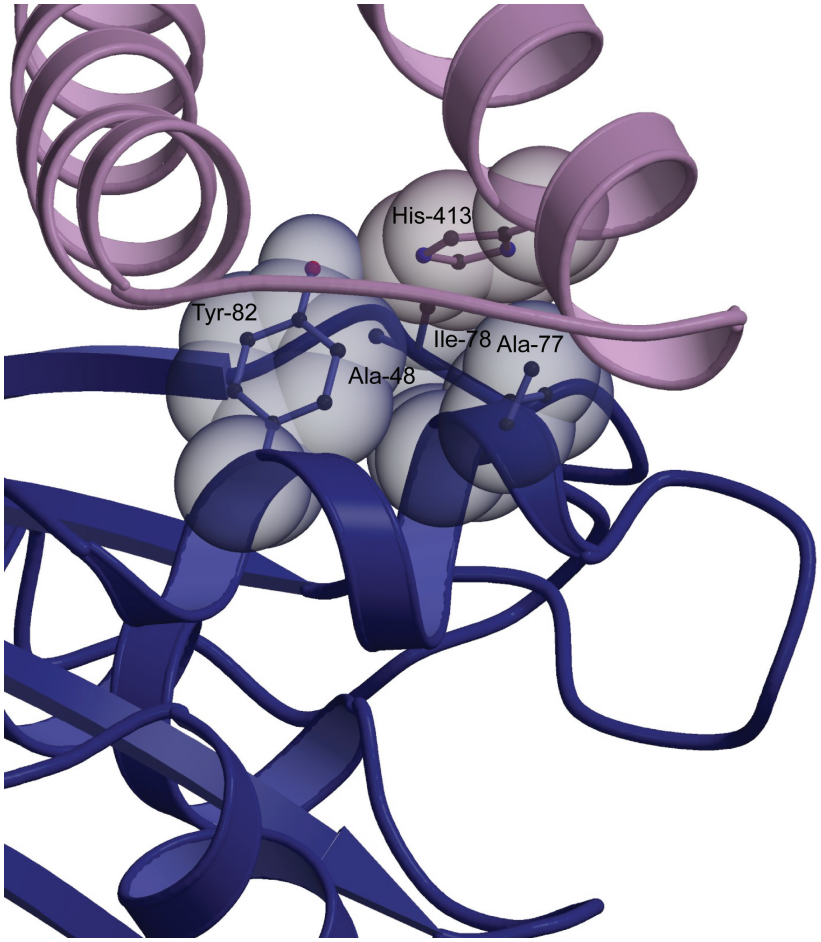
Structure of the RalB-RLIP76 GBD Complex

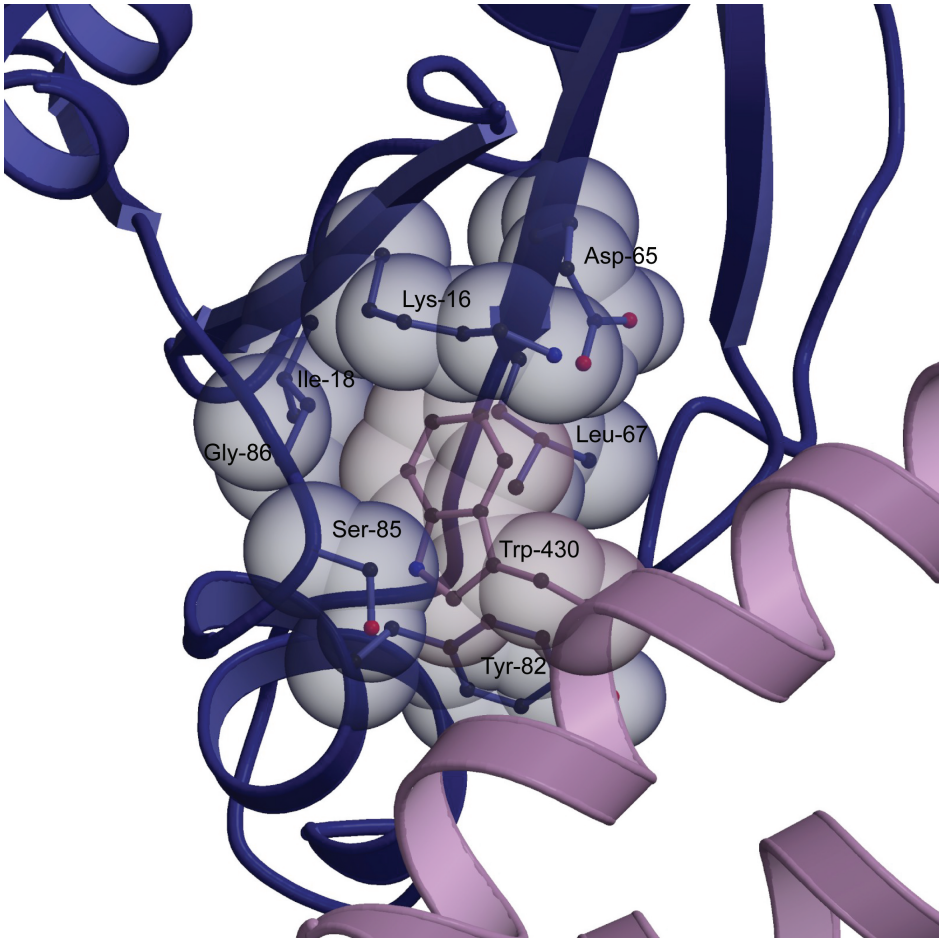




Fenwick et al (2010) Structure Vol 18 985

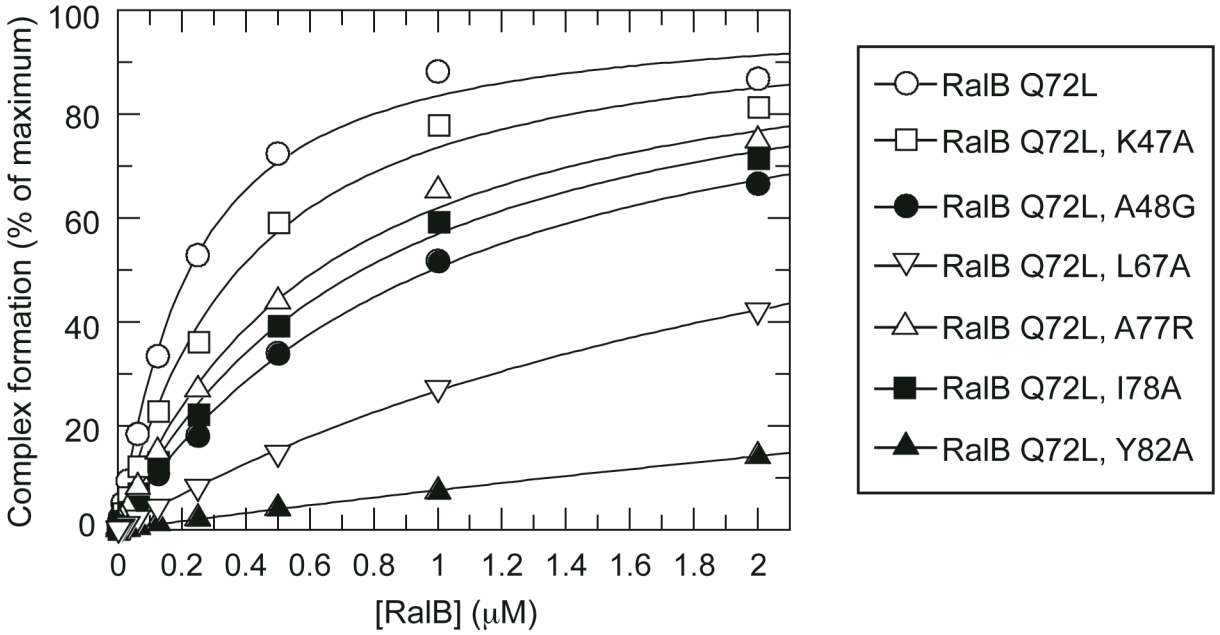
Conserved Residues in RLIP76 are in the Interface





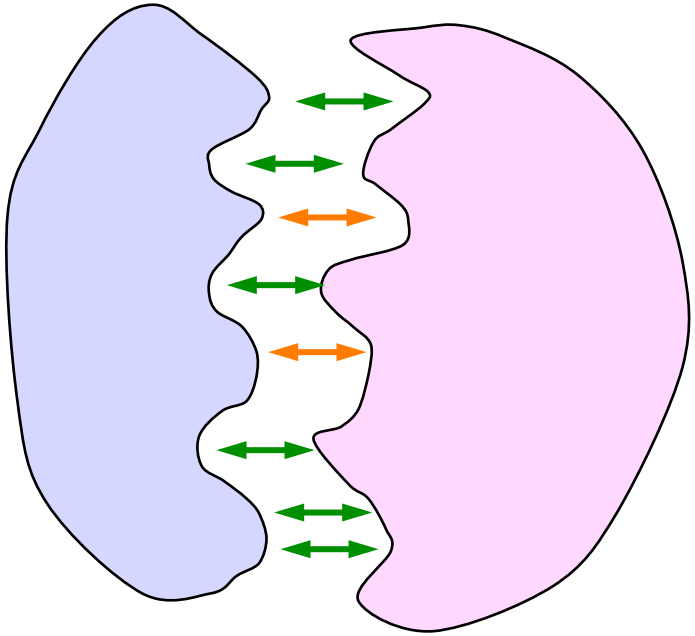
Fenwick et al (2010) Structure Vol 18 985

Mutation of RalB Residues in the Interface Disrupts Binding



Fenwick et al (2010) Structure Vol 18 985

Protein-protein Interfaces



Prediction of protein structures is possible if a homologue is known but interfaces are harder to predict

Proteins interact through large, flat surfaces using multiple contacts

Traditionally considered a difficult target for drug design but "hot spots" may define important interactions

Acknowledgements

R. Bryn Fenwick

Jacques Camonis (Institut Curie)
Peter Lowe (Biomolecular Interactions Consultancy)

NMR Facility
Daniel Nietlispach

BBSRC bioscience for the future

wellcome trust

MRC Medical Research Council

Isaac Newton Trust

Department of Biochemistry
University of Cambridge