# *Dynamic assembly of proteins: characterization, prediction and design*

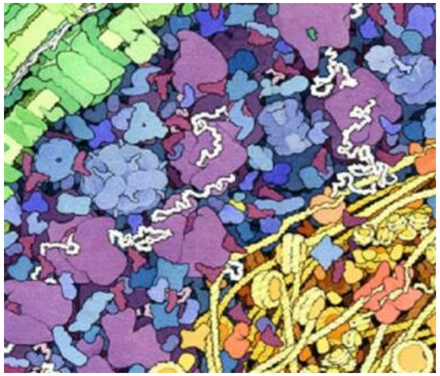
Raphaël GUEROIS

Molecular Assemblies and Signaling
Structural Biology and Radiobiology Lab
iBiTecS – URA CNRS 2096 - CEA Saclay

***Jobim, September 7-9, 2010***

# Protein interaction networks take place in a dense macromolecular environment

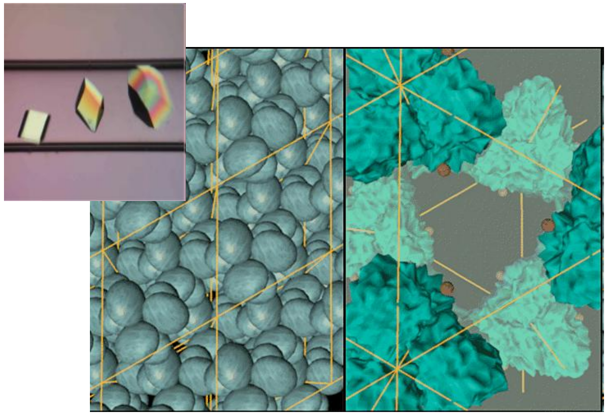
Proteins+RNA *in cellulo*



300-400 mg/ml

*Zimmerman & Trach, J Mol Biol, 1991*

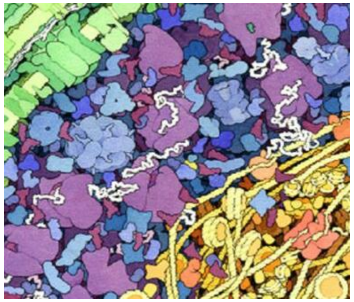
Protein crystals



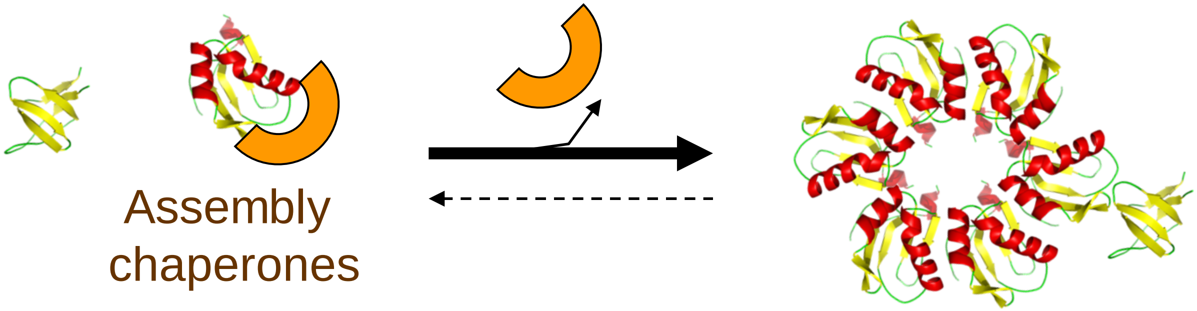
~ 600 mg/ml

- ➔ Challenging environments to establish specific interactions
- ➔ How multi-subunits systems do assemble in this dense environment ?

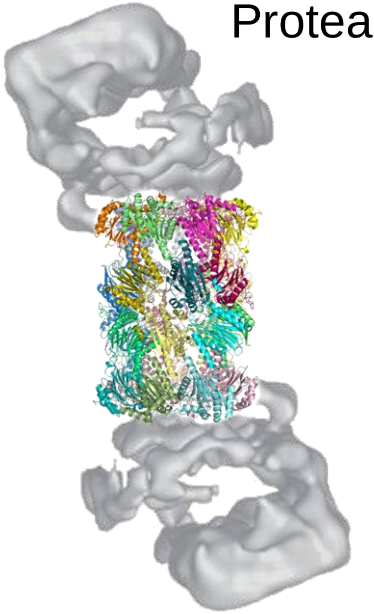
# Proper assembly can be controled by generalist or specialized assembly chaperones



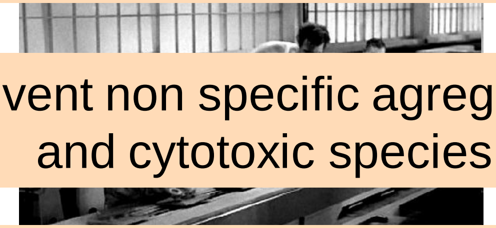
Large molecular assemblies



Assembly chaperones



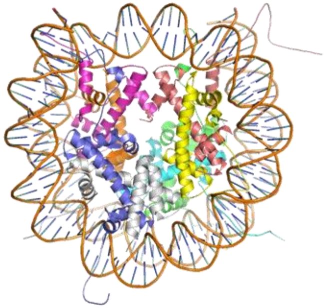
Proteasome



Synchronize assembly steps

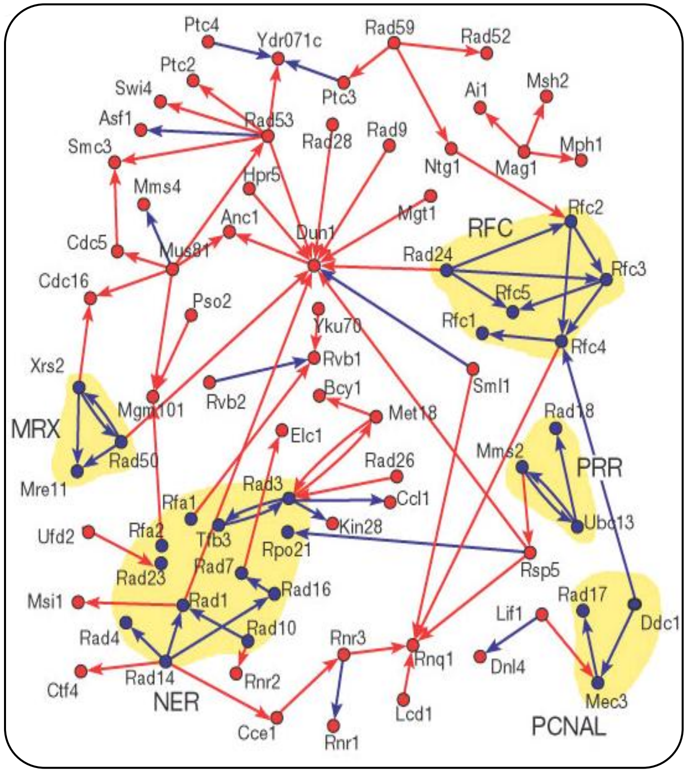
Prevent non specific agregation and cytotoxic species

Control the specificity of subunits assembly



Nucleosome

# Many assembly chaperones recently discovered could already be found in large interaction maps



**Ensemble of physical interactions, both binary and co-membership complexes**

**Dissection of protein-protein interaction networks**

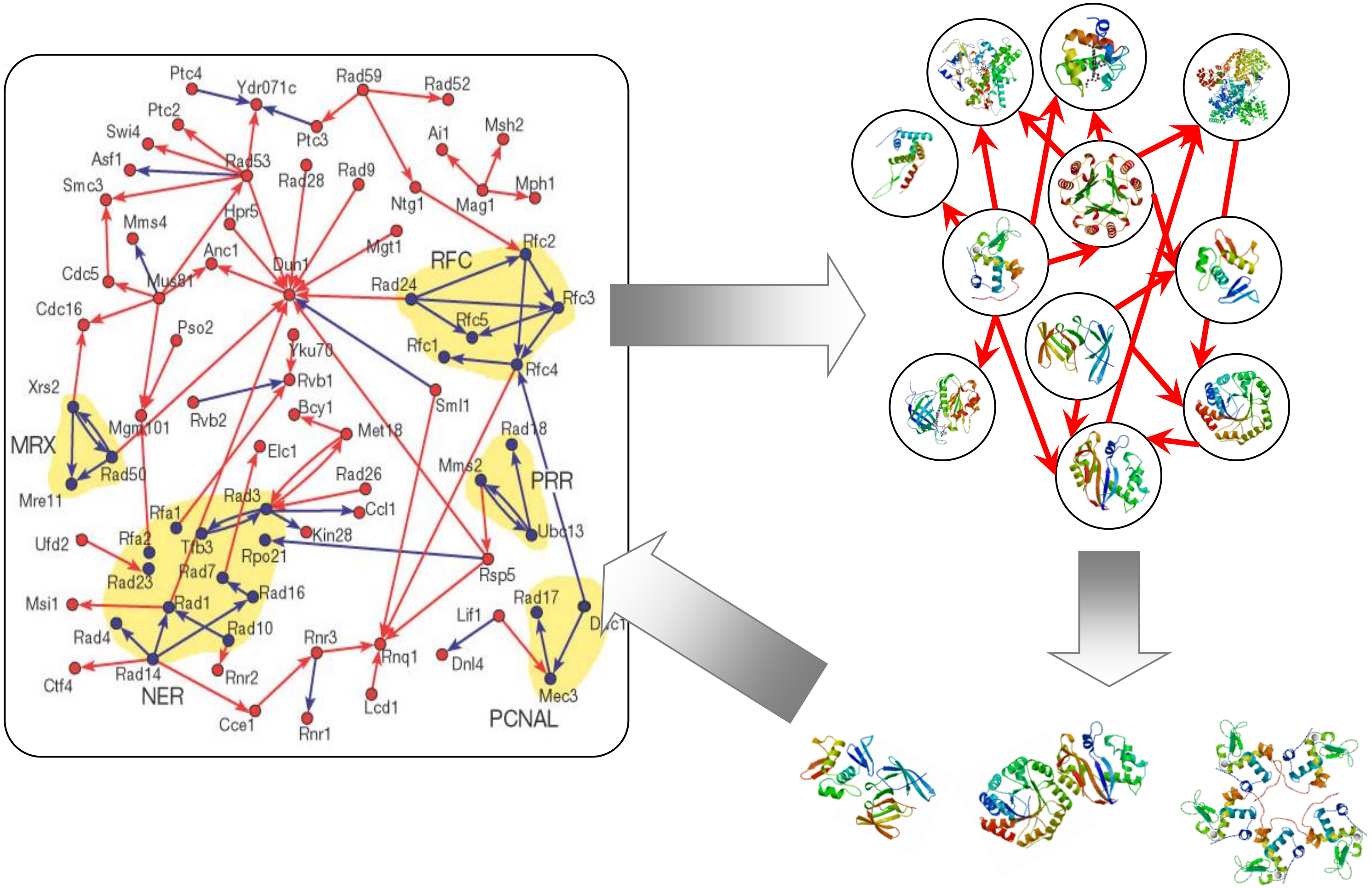
➔ Synergies, competition and crosstalks ?

**Design of compensatory mutants**

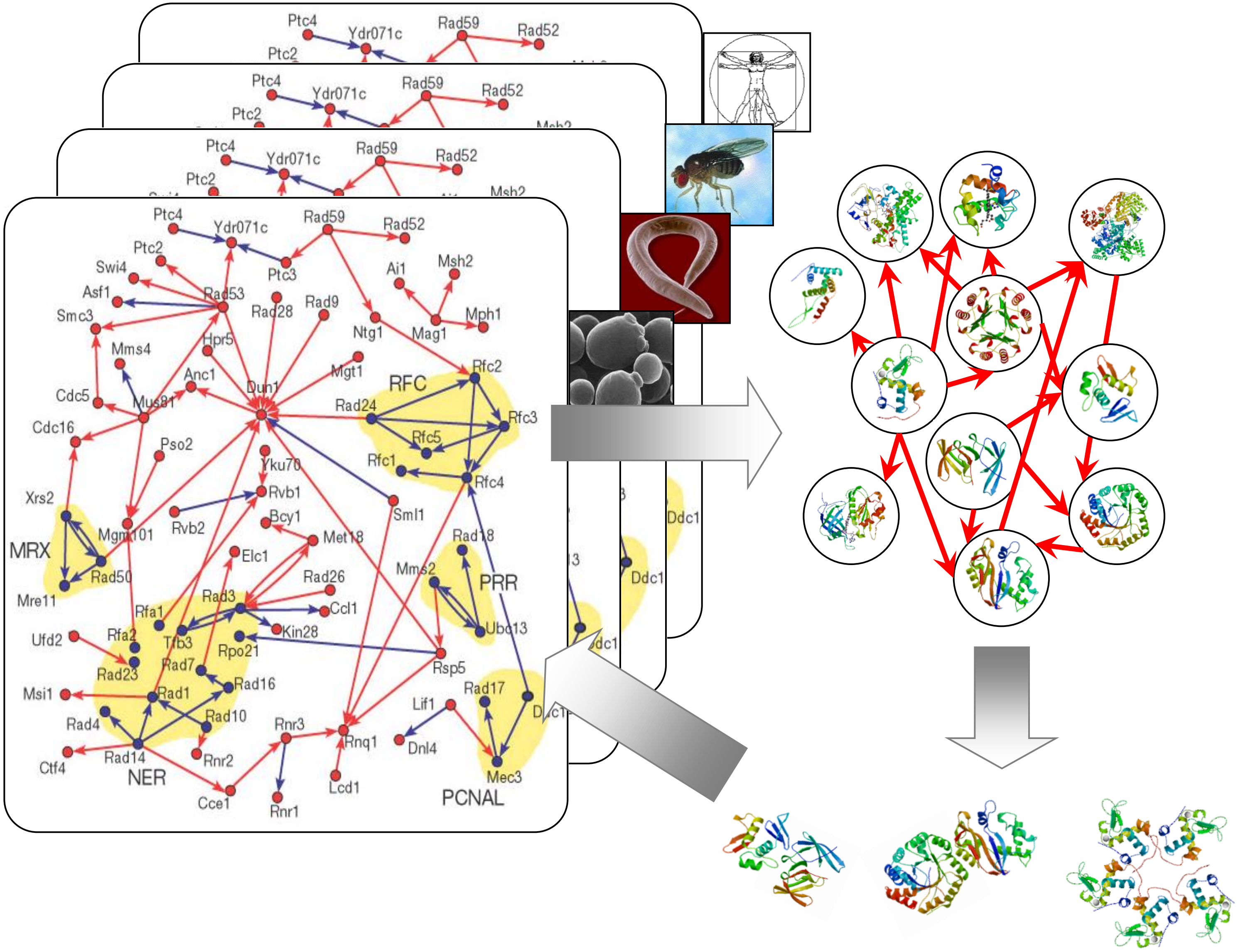
➔ Assess *in vivo* whether a direct interaction is involved in a specific phenotype

How far can structural biology help in exploring these networks ?

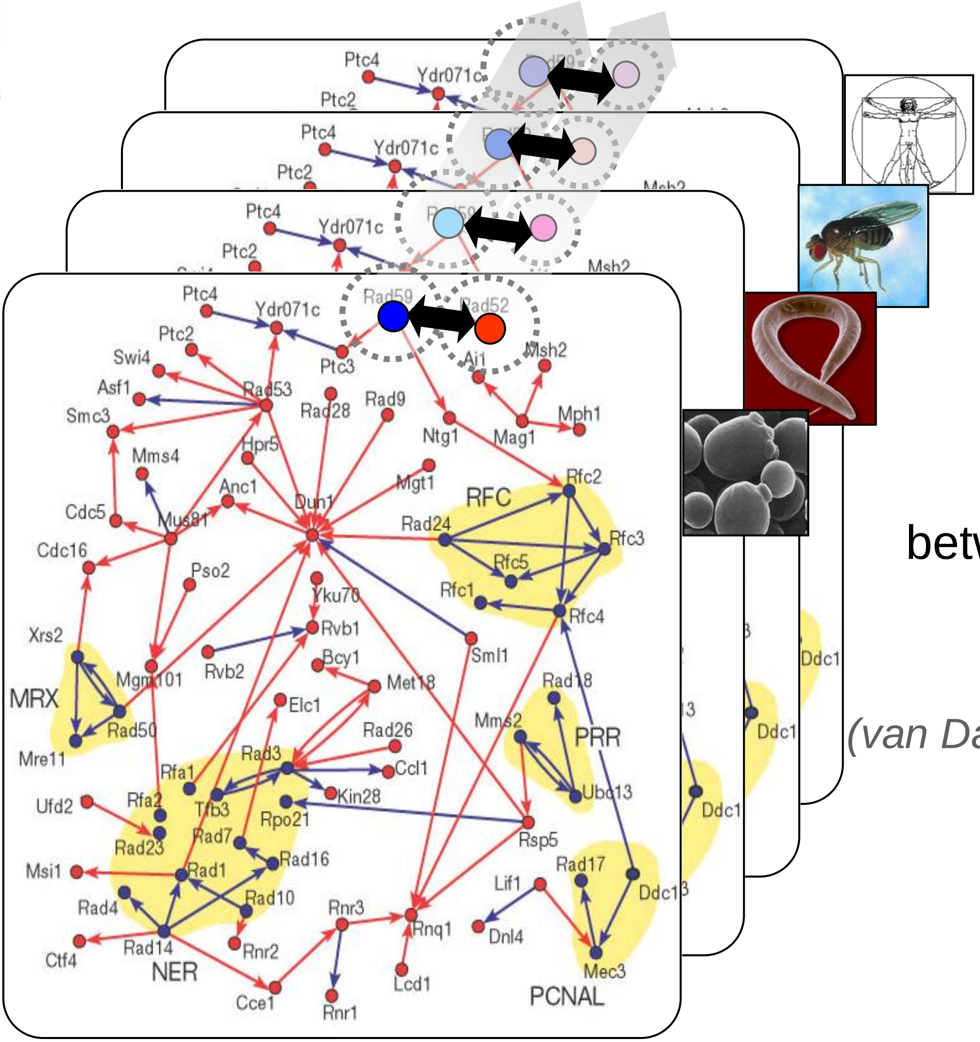
# How far can available interaction data be translated into 3D structures ?



# The evolutionary dimension should provide key information to exploit interaction data under a structural perspective



# Do conserved proteins tend to interact within the same complexes in different species ?

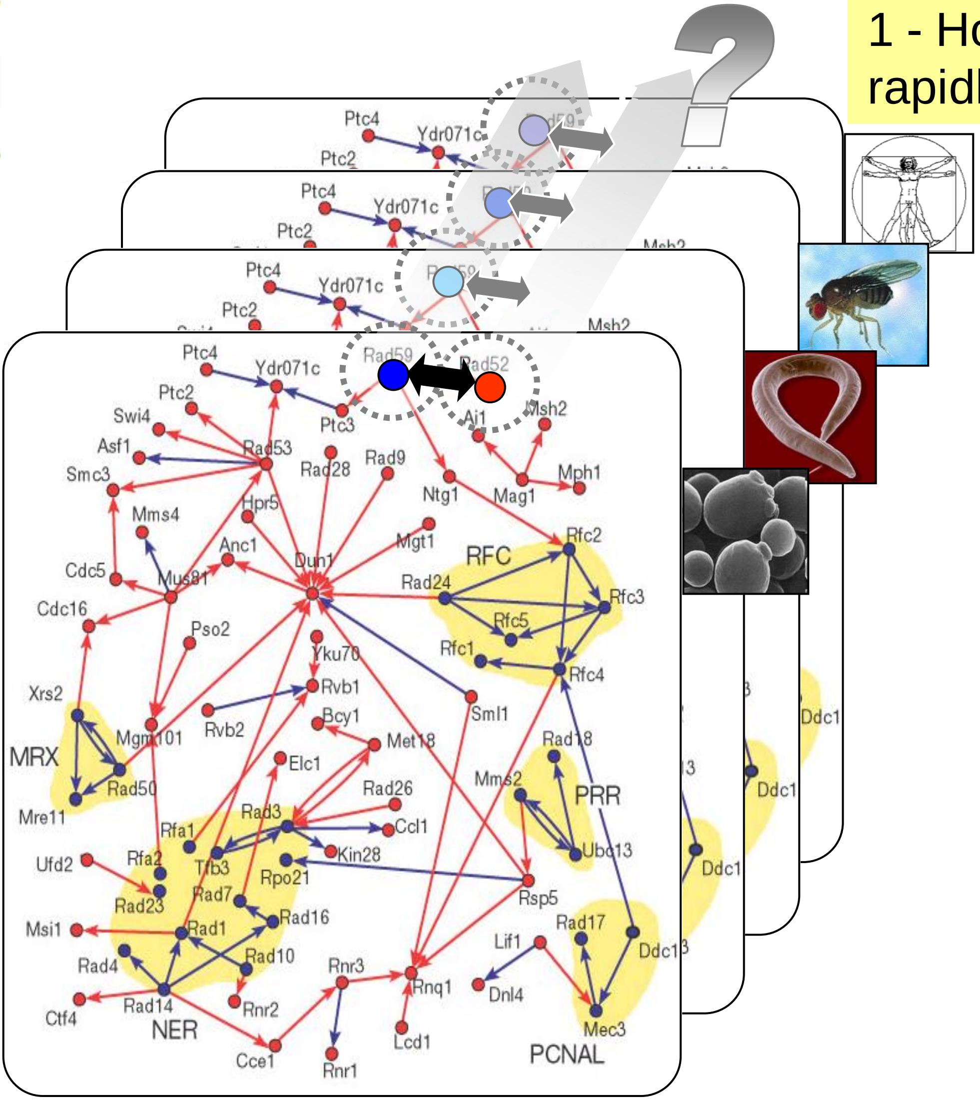


No extensive rewiring  
between yeast and human network  
rather gene loss/gain  
among complex subunits

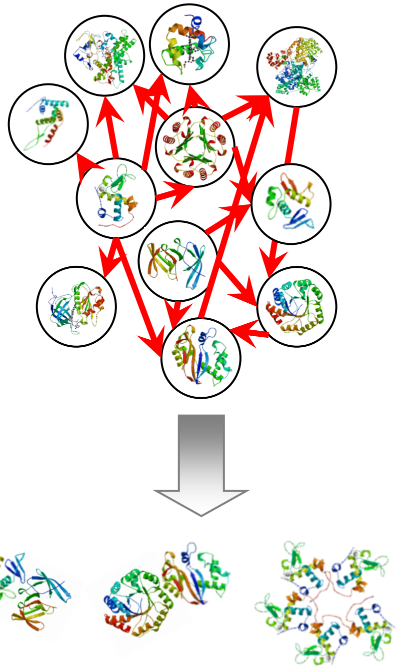
*(van Dam & Snel, PLoS Comput. Biol. (2008))*

# How can evolutionary information help in combining interaction maps with structural data ?

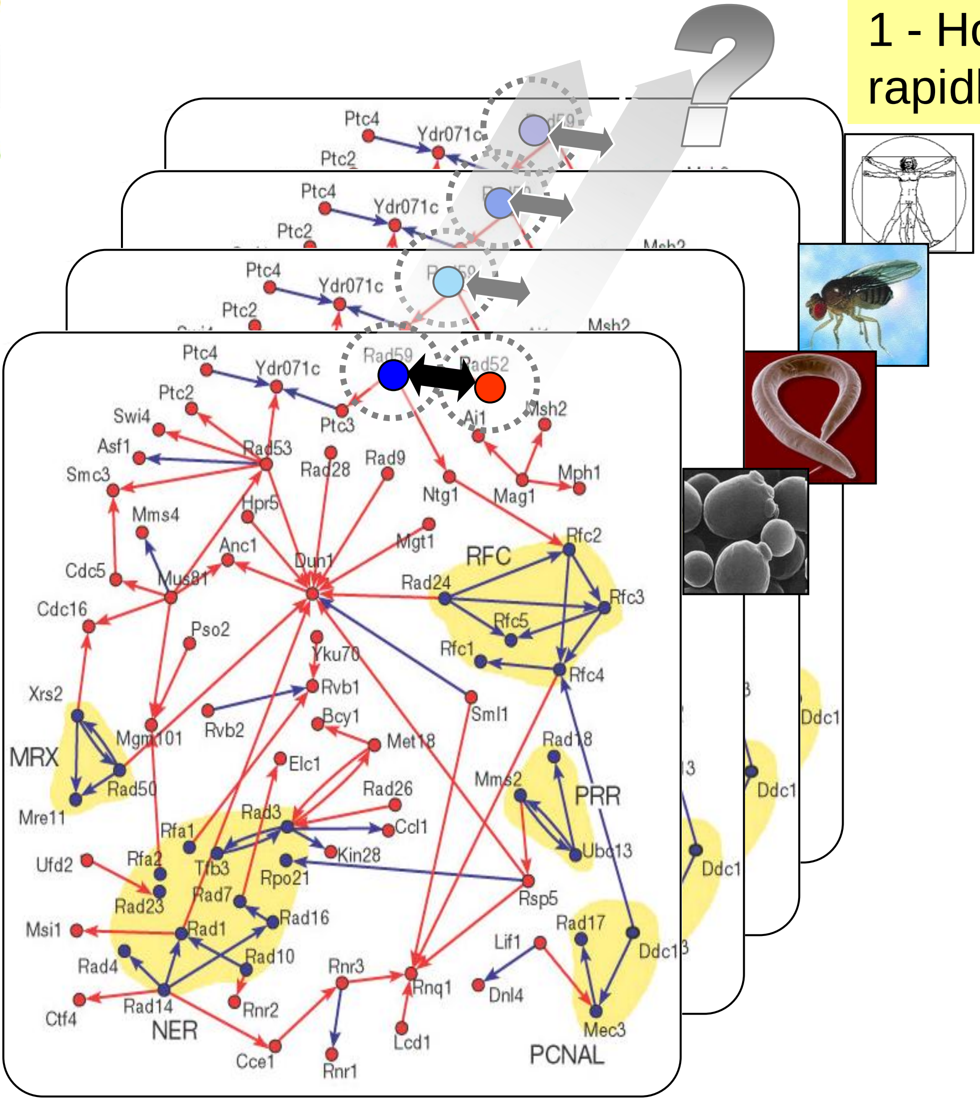
1 - Homology detection for rapidly diverging partners ?



2 – Evolutionary traces at complex interface ?

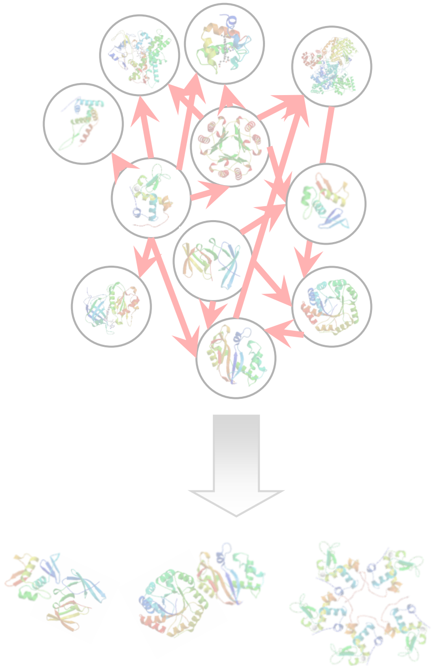


# How can evolutionary information help in combining interaction maps with structural data ?



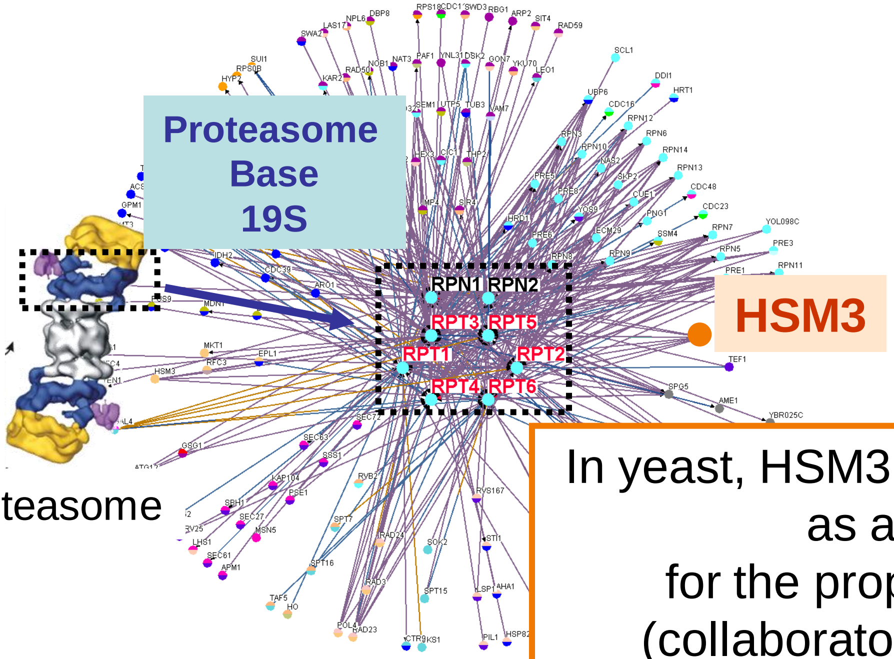
1 - Homology detection for rapidly diverging partners ?

2 – Evolutionary traces at complex interface ?

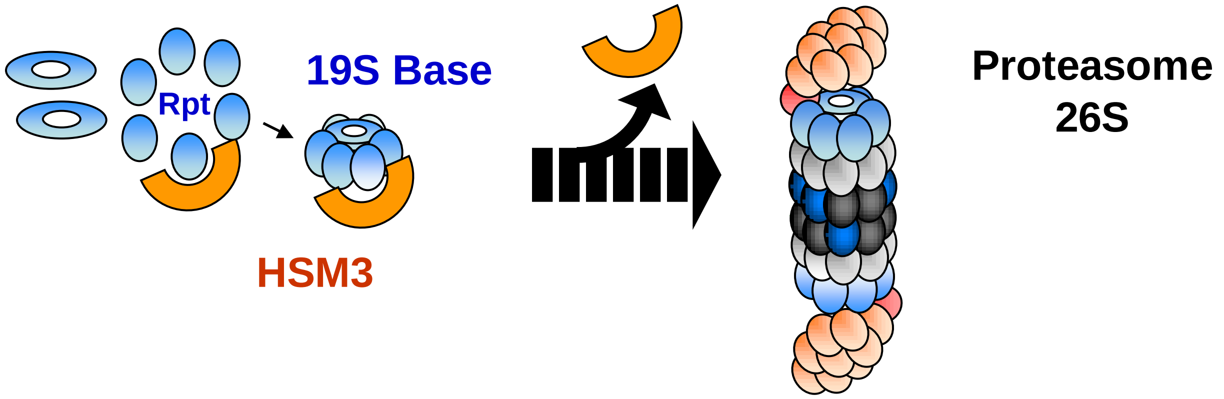


*Jobim, September 7-9, 2010*

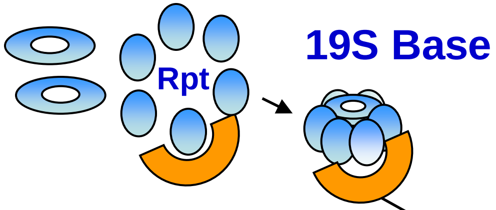
# Proteasome assembly in yeast is controled by assembly chaperones exhibiting high sequence divergence



In yeast, HSM3 was detected by a genetic screen as an important chaperone for the proper assembly of the 19S base (collaborator Anne Peyroche, CEA Saclay)



# Is there a human homolog of HSM3 ?



HSM3 sequence diverged fast making a putative homolog difficult to detect in remotely related species.

HSM3 seq.

*S. cerevisiae*

*K. lactis*

~ **25** % identity

**Closely related fungi**

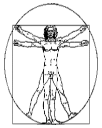
*Ex: Candida Albicans*

?

**Remotely related fungi**

*Ex: Schyzosachoromyces Pombe*

*Ex: Magnaporthe grisae*

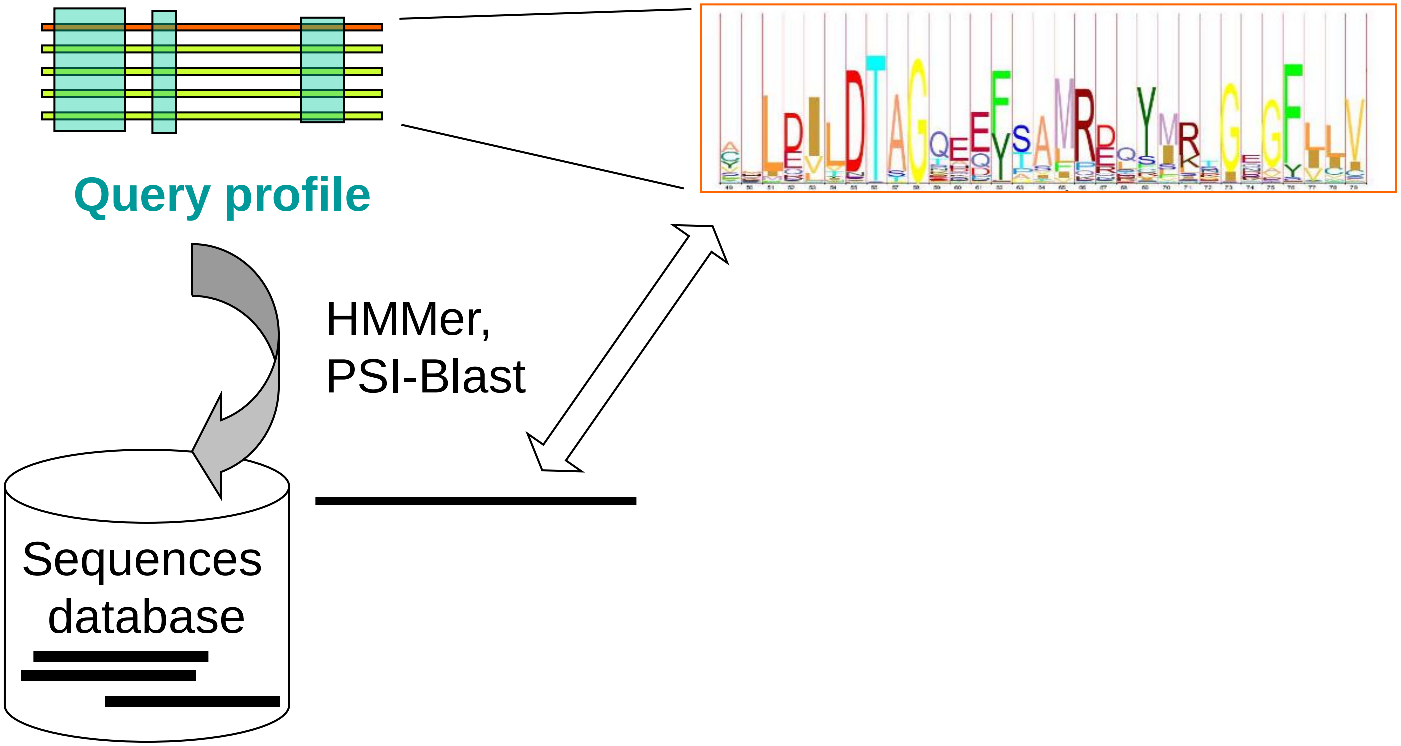


**Euk. sup.**

HEAT Repeats superfamily

➔PSI-Blast rapidly diverges > 1600 sequences hits … too many potential orthologs…

# Methods for searching remote homologs

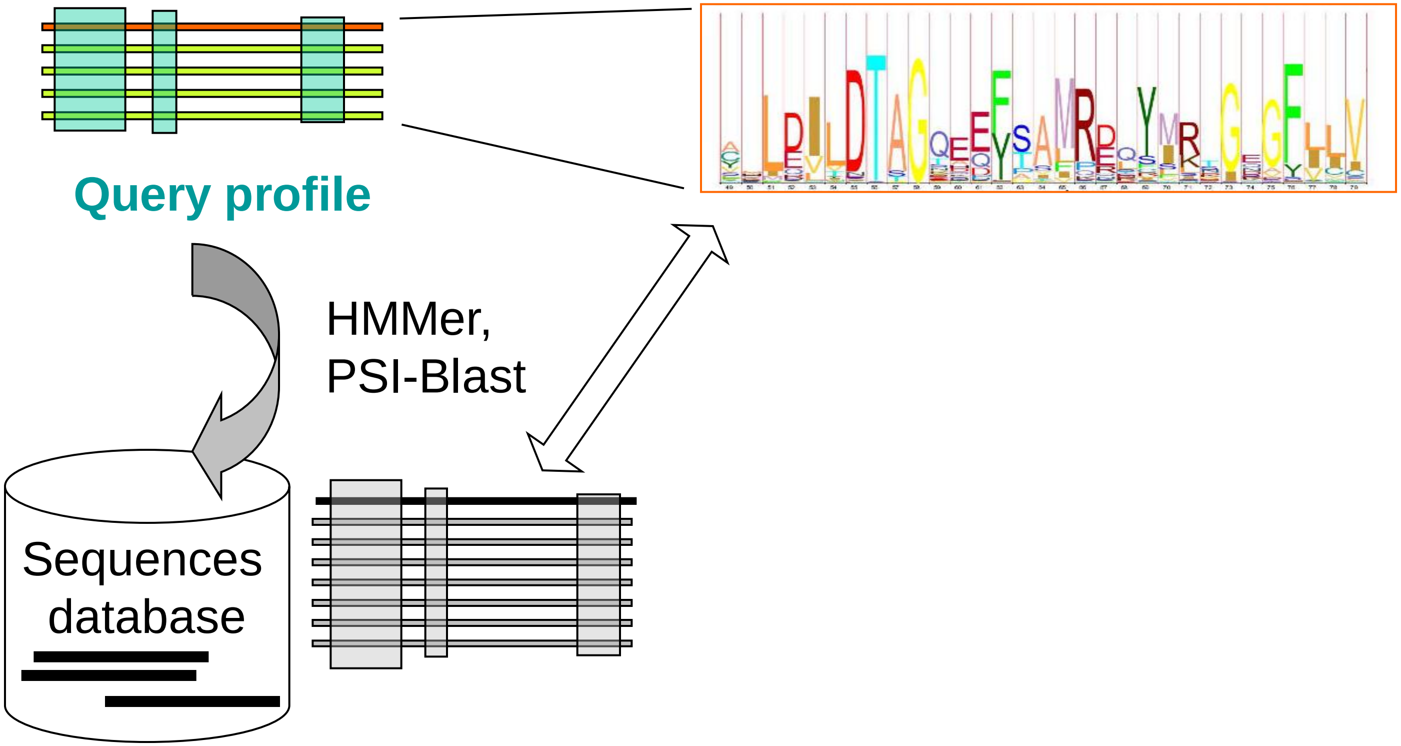


**Query profile**

HMMer,
PSI-Blast

Sequences database

# Methods for searching remote homologs

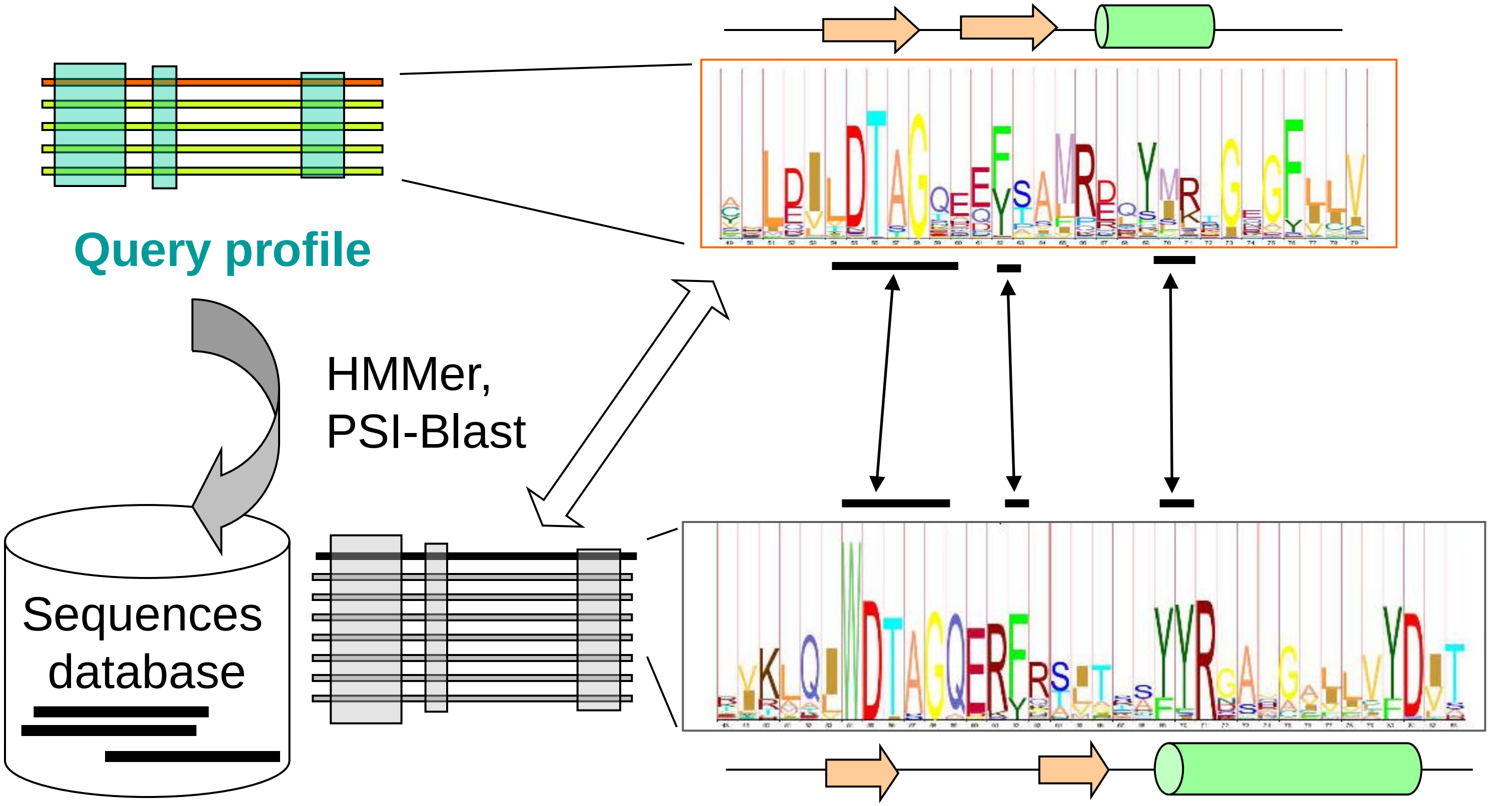


Query profile

HMMer,
PSI-Blast

Sequences
database

# Methods for searching remote homologs



**Query profile**

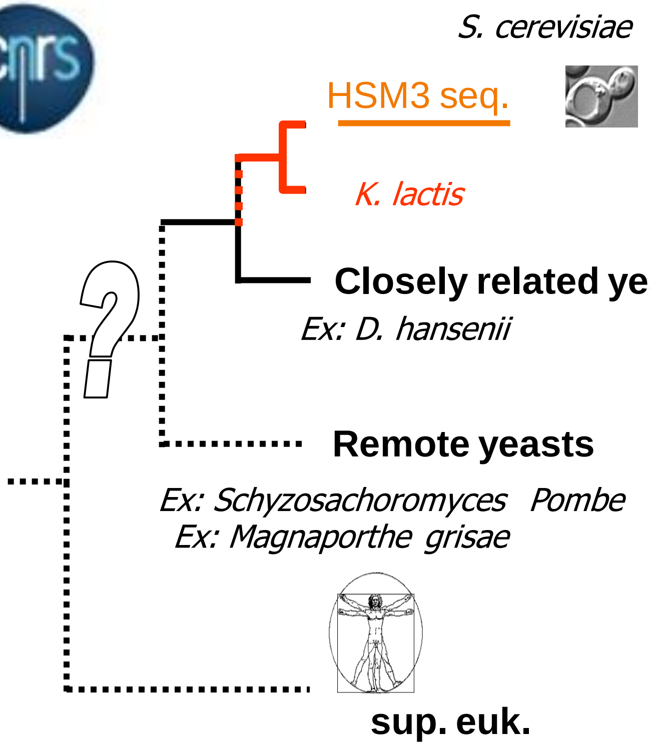
HMMer,
PSI-Blast

Sequences
database

| | |
|---|---|
| FFAS | *Rychlewski et al., Prot Sci. (2000)* |
| PROF_SIM | *Yona & Levitt, JMB (2002)* |
| **COMPASS** | *Sadreyev & Grishin, JMB (2003)* |
| **HHsearch** | *Söding, Bioinformatics (2005)* |
| PRC | *Madera, Bioinformatics (2008)* |
| PROCAIN | *Wang et al, NAR (2009)* |

# How can the correct orthologue be identified without diverging in the HEAT repeat superfamily ?

Only 4 orthologous sequences could be detected before divergence…



*S. cerevisiae*

HSM3 seq.

*K. lactis*

**Closely related yeasts**

*Ex: D. hansenii*

**Remote yeasts**

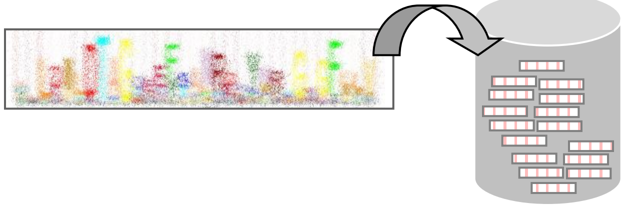
*Ex: Schyzosachoromyces Pombe*
*Ex: Magnaporthe grisae*

**sup. euk.**

- For fully sequenced organism
- Build a tailored profile for each sequence

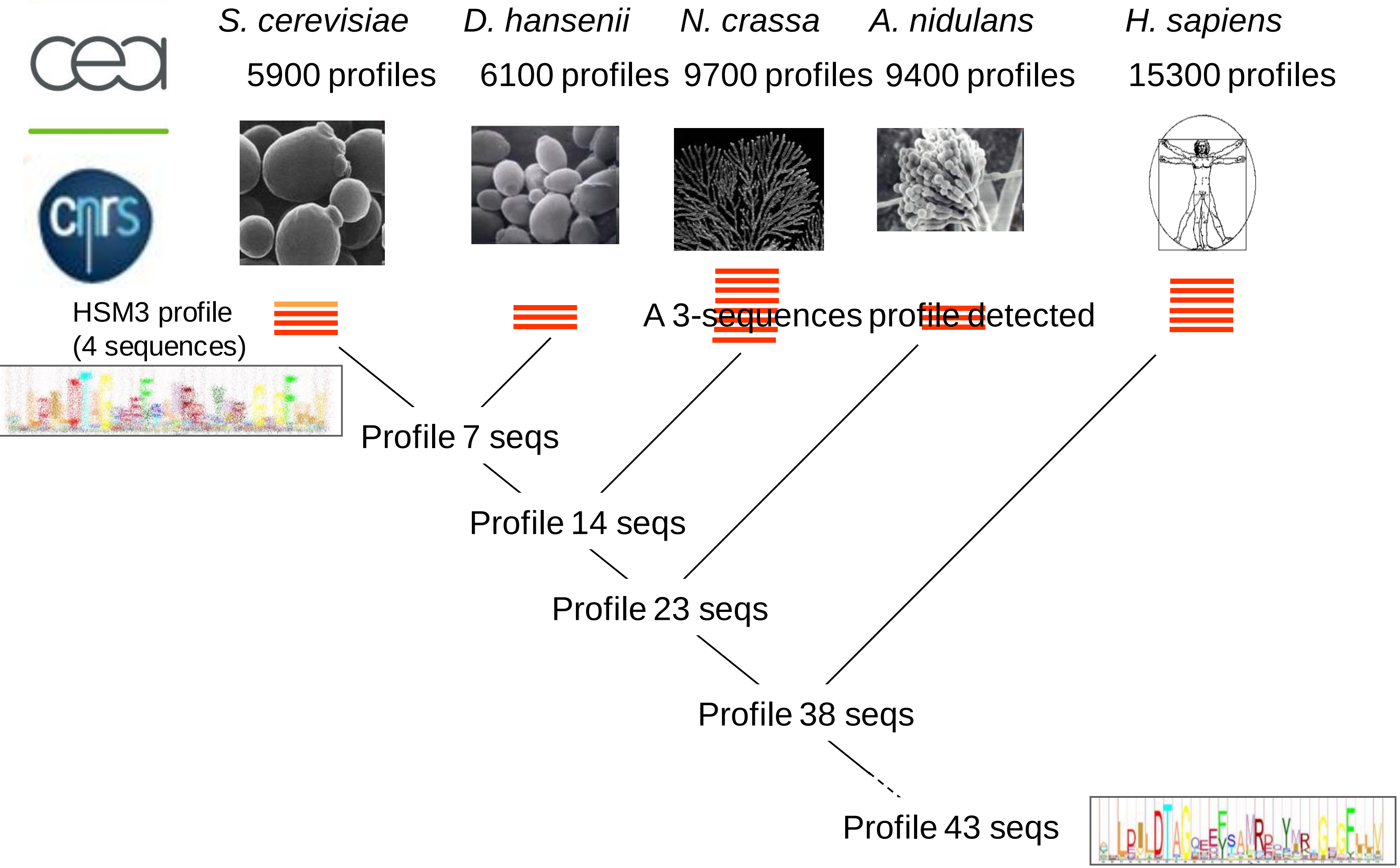
➔ Profile collection/organism

Profile



Profile-profile comparison

# Profile Collection for each model organism
# Iterative Profile/Profile Alignment (IPPA)



*S. cerevisiae* 5900 profiles

*D. hansenii* 6100 profiles

*N. crassa* 9700 profiles

*A. nidulans* 9400 profiles

*H. sapiens* 15300 profiles

HSM3 profile
(4 sequences)

A 3-sequences profile detected

Profile 7 seqs

Profile 14 seqs

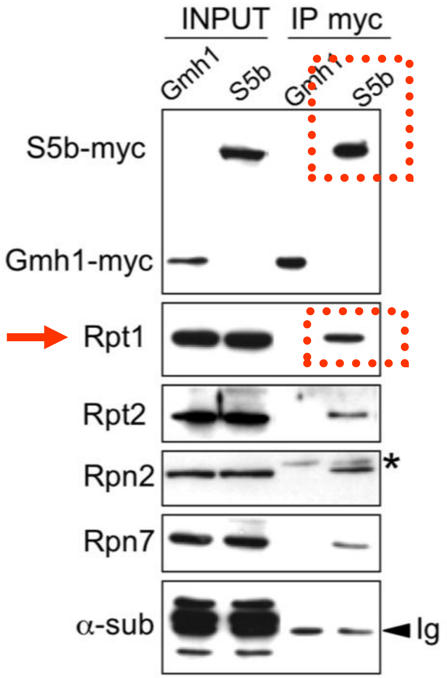
Profile 23 seqs

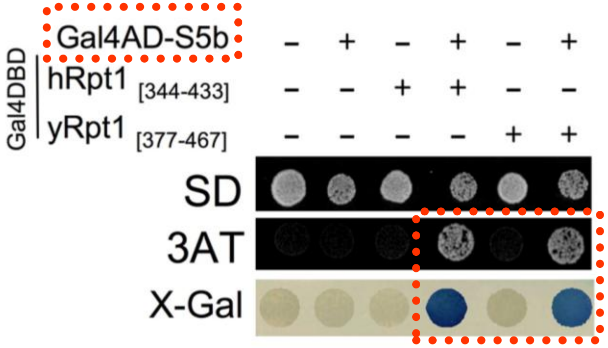
Profile 38 seqs

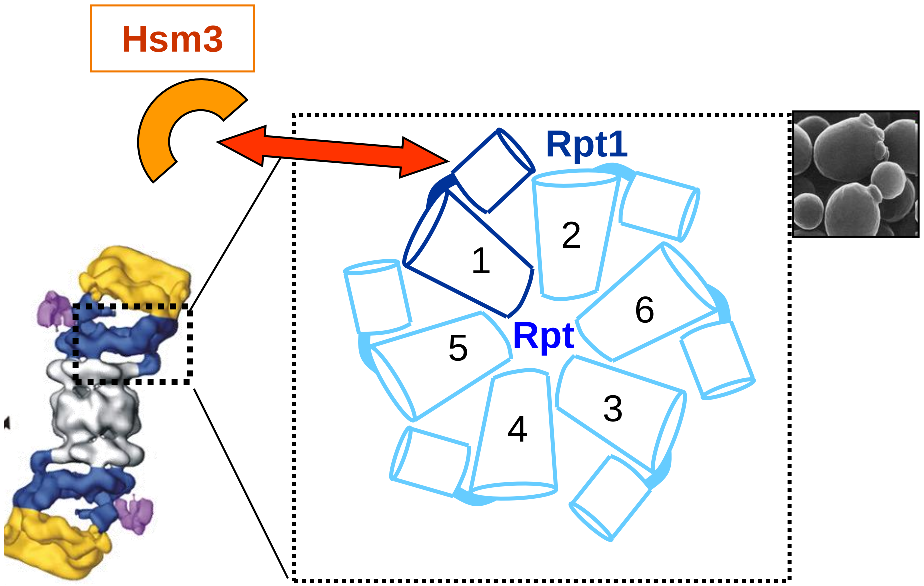
Profile 43 seqs

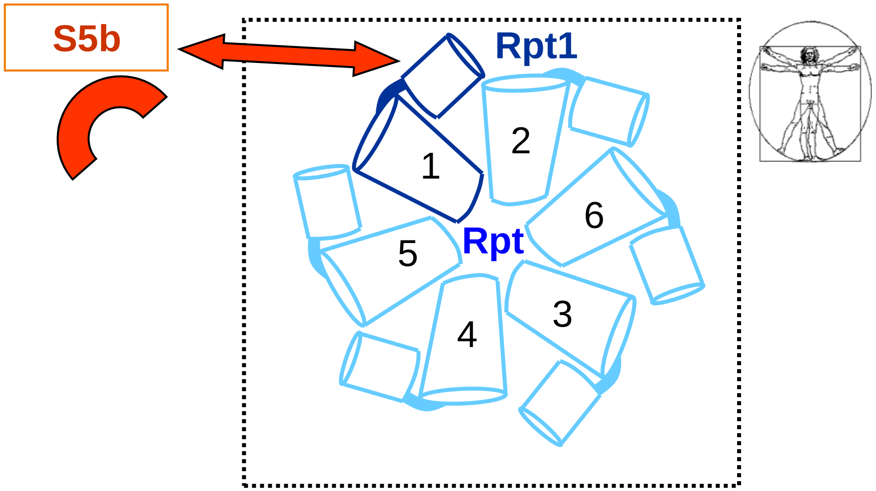
A human orthologue for HSM3 … S5B ? (seq. Id ~ 9%)

# Experimental validation that S5b is a HSM3 human remote ortholog



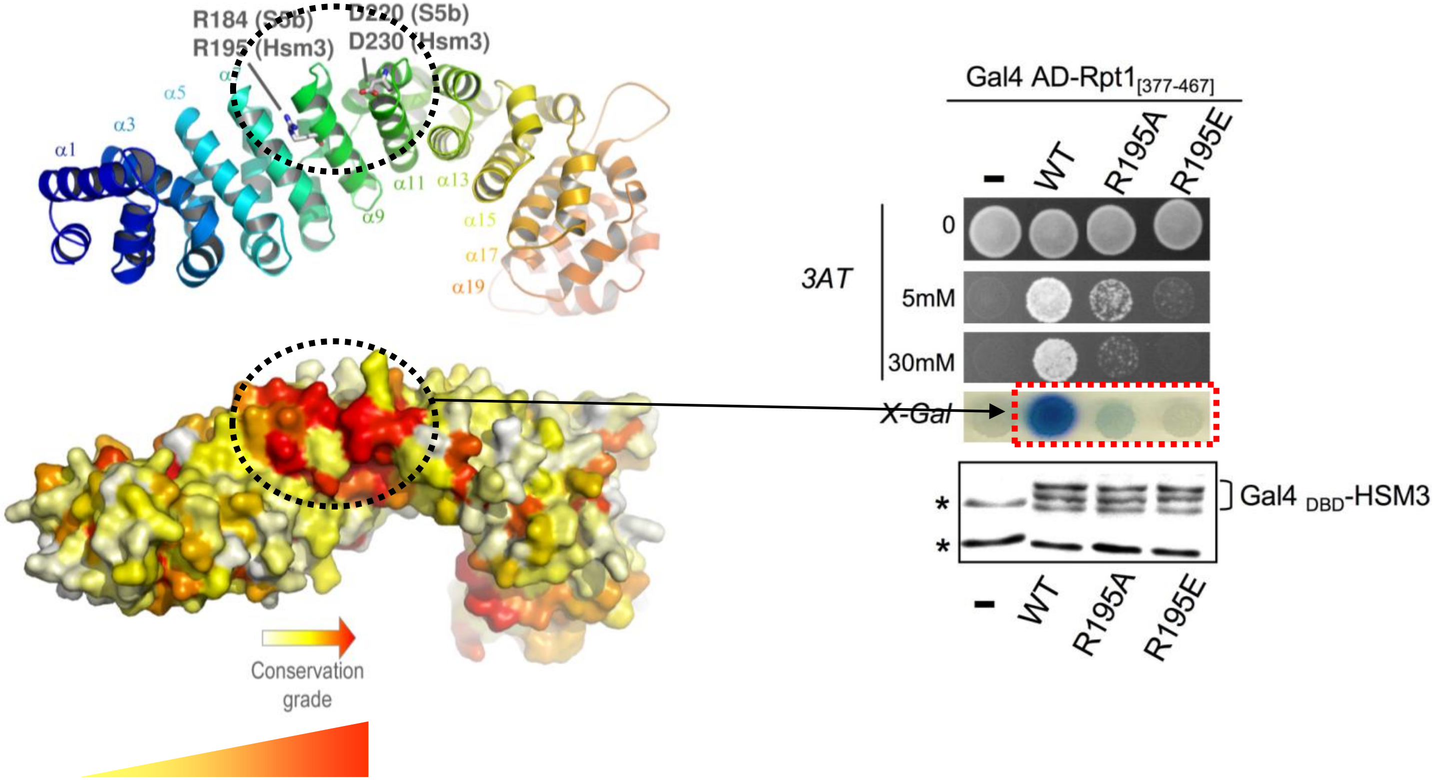






*Le Tallec et al (2009)* ***Mol Cell***

# Design of a interaction mutants between HSM3/S5b and Rpt1

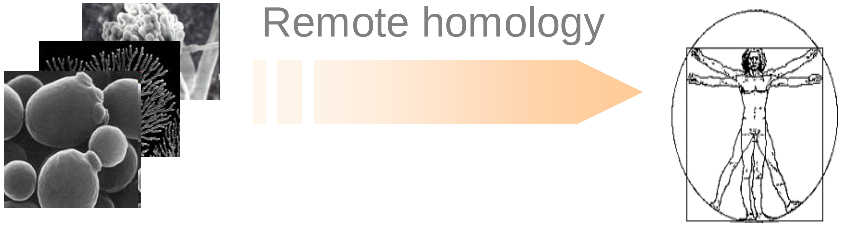


conservation score

*(Rate4Site, Pupko et al Bioinformatics 2002)*

# Remote homologs detection using iterative profile-profile analyses on pre-built profile databases

➔ 3 new proteasome chaperones in human (%id < 10%)



*Coll. A. Peyroche (CEA, Saclay)*

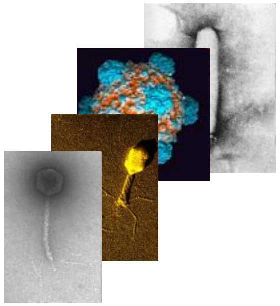
Le Tallec et al (2007) ***Mol Cell***
Le Tallec et al (2009) ***Mol Cell***

➔ DNA repair machineries in *Helicobacter pylori* ?

*Coll. P. Radicella (CEA, Fontenay)*

Marsin et al (2008) ***PLoS Genet***

➔ Recombination machineries in bacteriophages ?

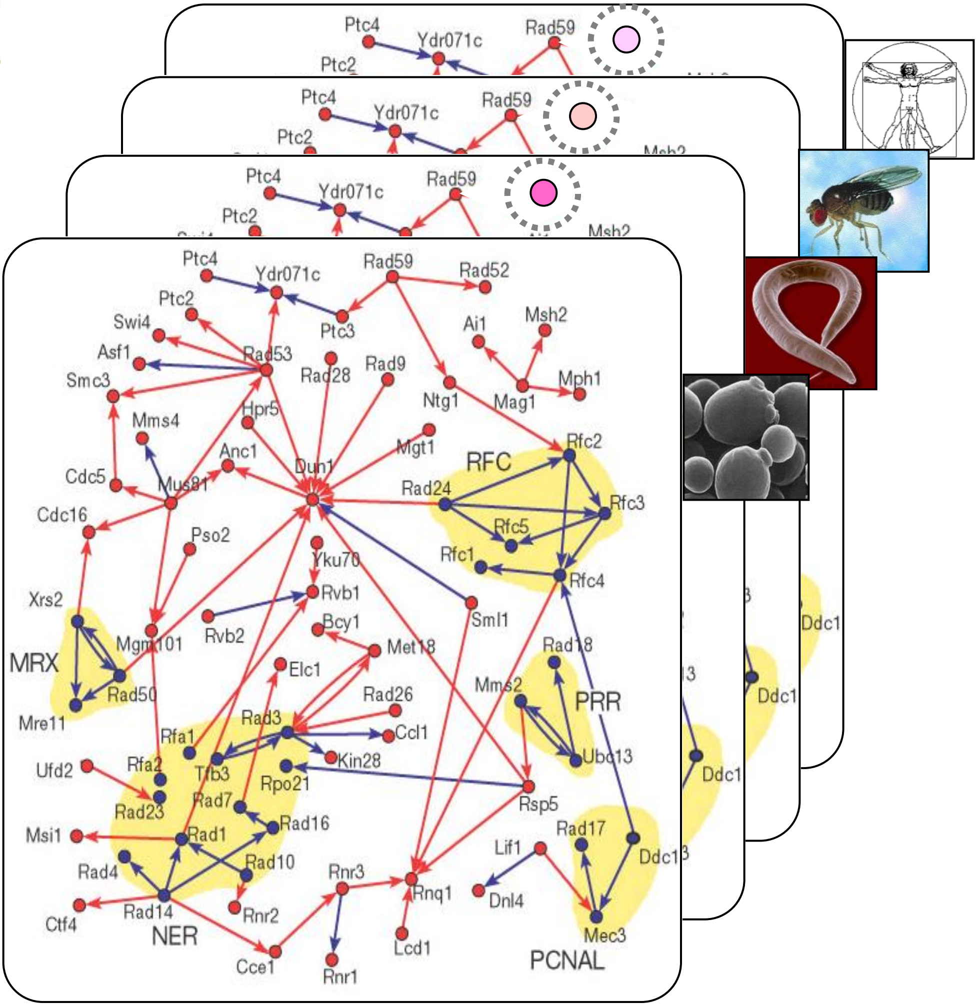


*Coll. M.-A. Petit (INRA, Jouy-enJosas)*

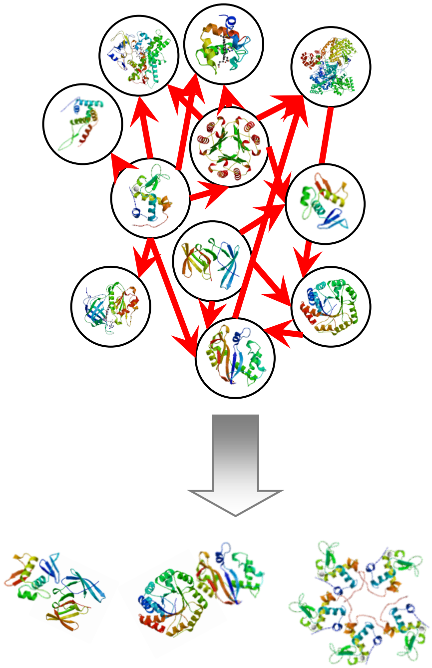
➔ Large scale analysis of 550 genomes taking into account the gene neighborhoods as an additional constraint

Lopes et al (2010) ***NAR***

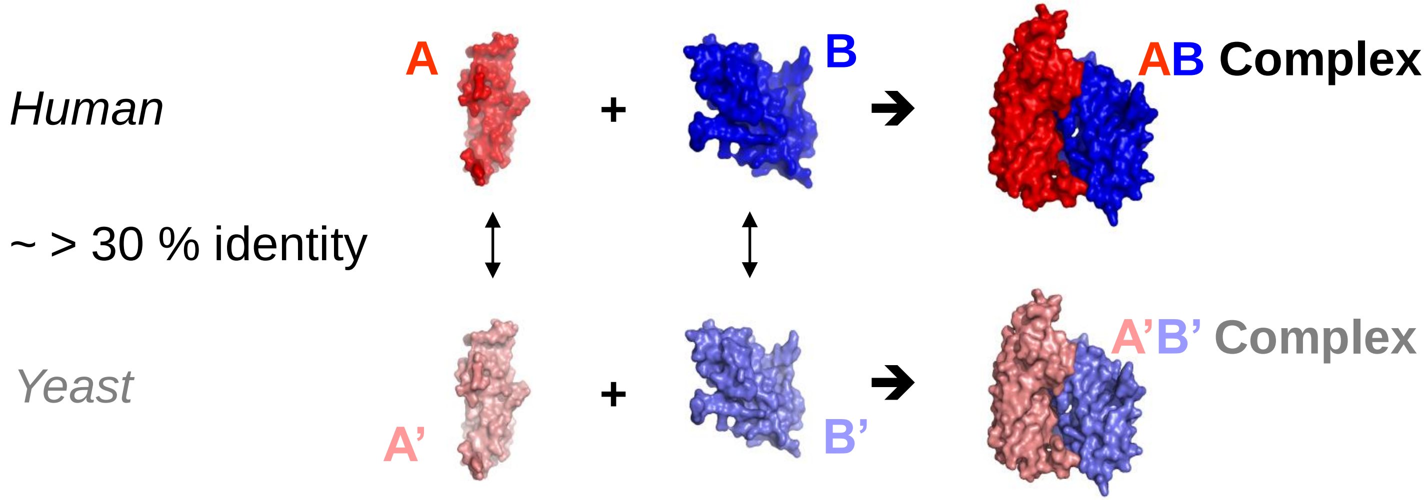
# From the prediction of protein 3D structures to that of protein assemblies…



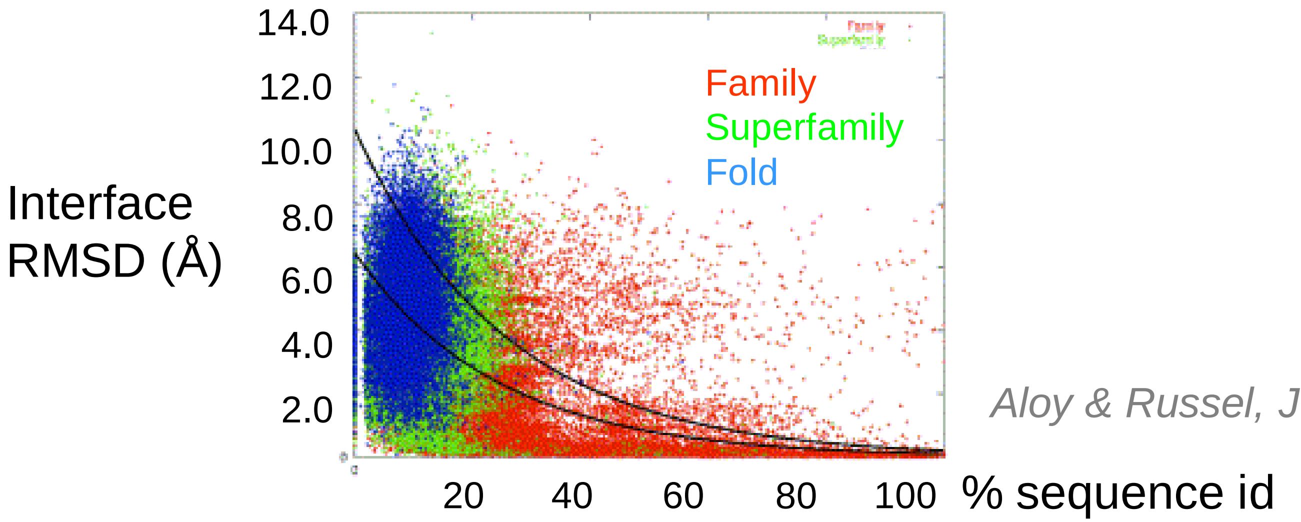
2 – Evolutionary traces at complex interface ?



# Relationships between sequence divergence and conservation of the binding mode

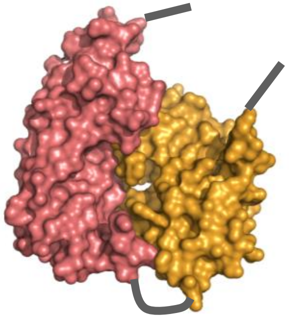


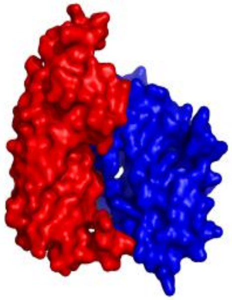
Two homologous complexes (~> 30% identity) generally interact in a similar manner



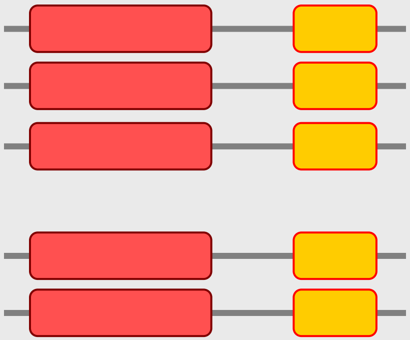
*Aloy & Russel, JMB 2003*

# Database of intra and intermolecular interactions between domains or proteins





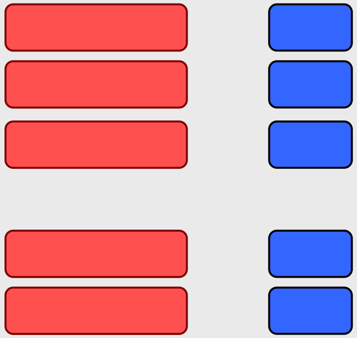
## intra-molecular interactions



**Domain A** **Domain B**

*H. sapiens*
*M. muscu*
*D. rerio*
⋮
*C. albic*
*S. cerev*

## inter-molecular interactions



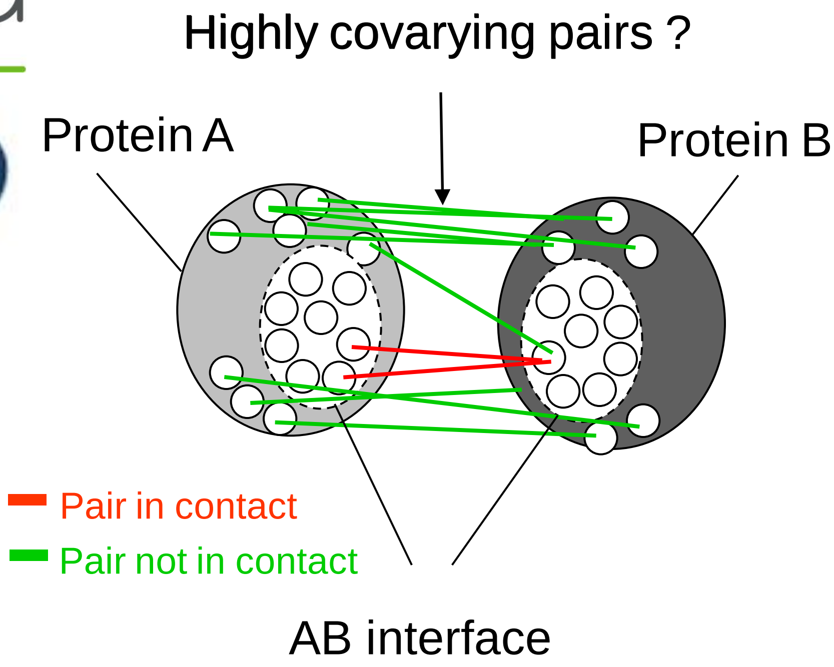
**Prot A** **Prot B**

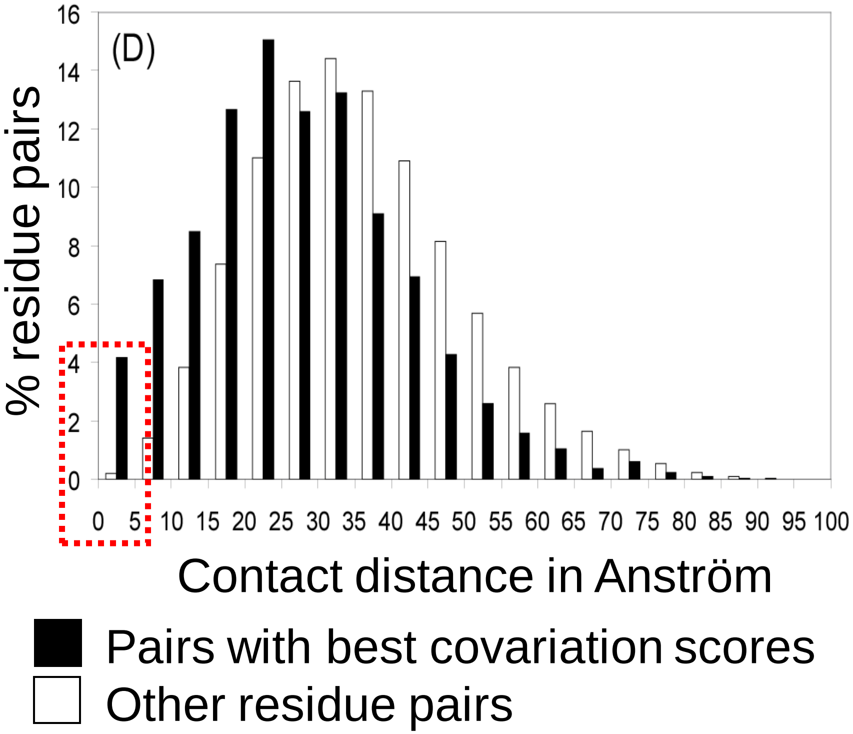
*H. sapiens*
*M. muscu*
*D. rerio*
⋮
*C. albic*
*S. cerev*

84 non redundant interactions extracted from the PSIMAP database *(Kim et al. Bioinformatics 2004)*

132 non redundant inter-molecular interactions compiled from *(Mintseris & Weng Proteins (2003), PNAS (2005))*

# Co-variation analyses at the interface of intra-molecular domain-domain interactions





**➔ Only 2 to 4 % of the residue pairs with highest covariation events are contacting in the complex**

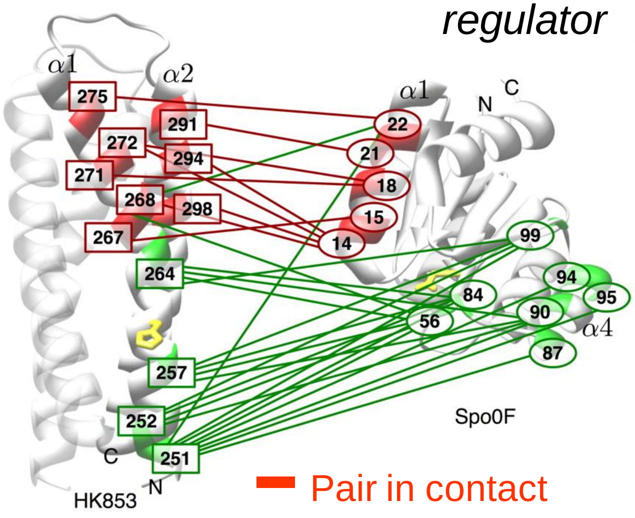
➔ Noise in the covariation signal at the surface ?

➔ Indirect effects between co-varying positions ?

# New methodologies in covariation analyses :
# Disentangling direct from indirect covariation events…

Based on a large set of sequences ~ 8000 sequences

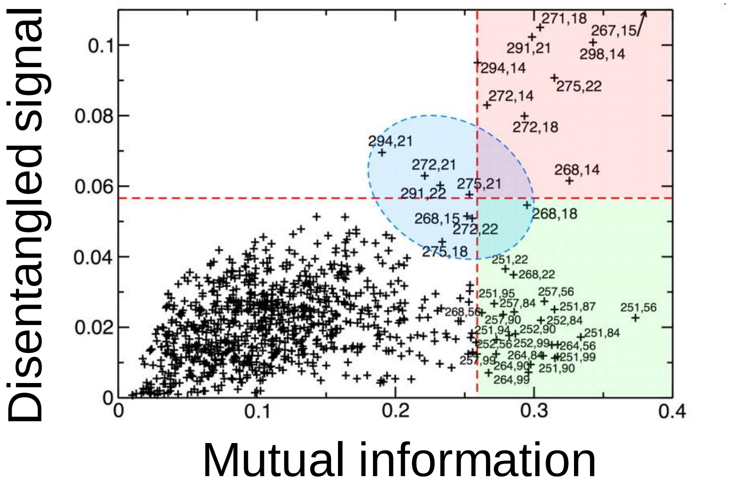
*His-kinase ←→ response regulator*



🟥 Pair in contact

🟩 Pair not in contact

Direct Coupling Analysis (DCA method)



*From Weigt, White … Hwa, PNAS (2009).*

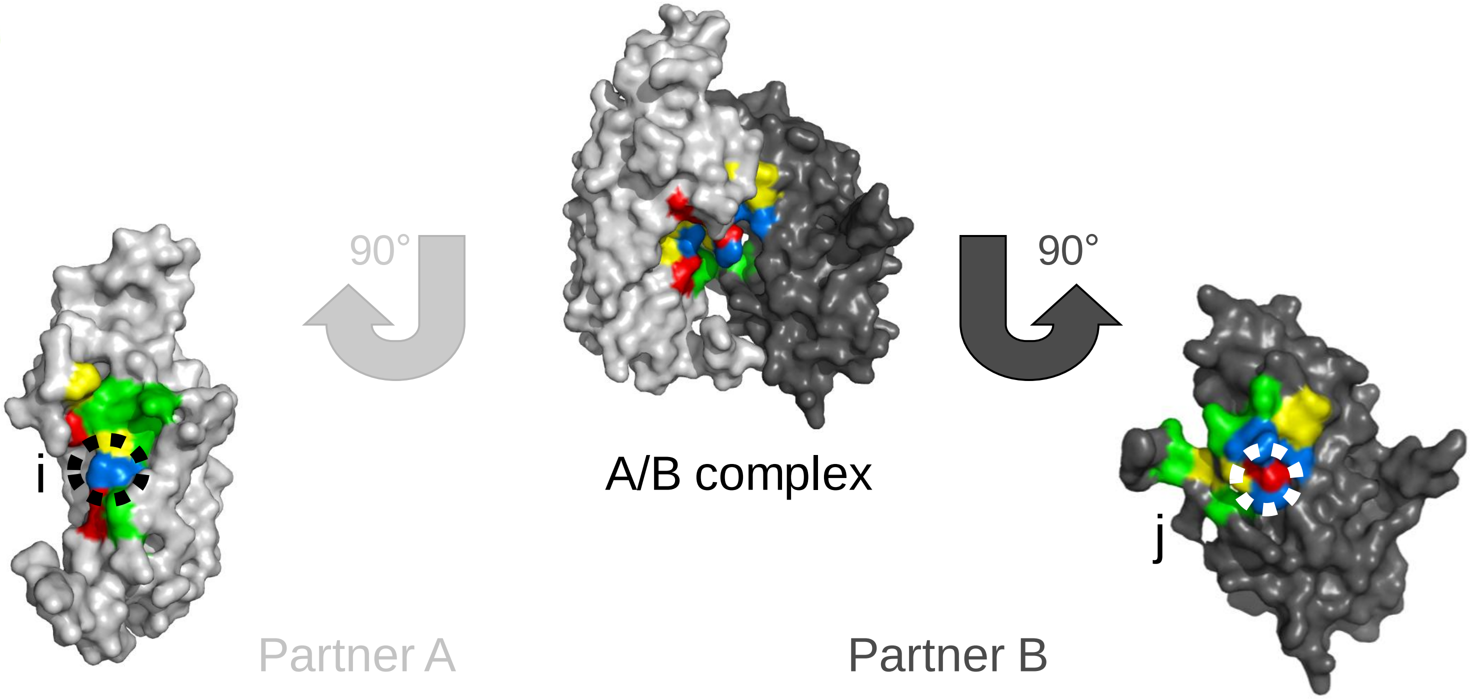
*Weigt, White … Hwa, PNAS (2009).*

➔ ***contact predictions at complex interface***

*Burger & van Nimwegen, PLoS Comput. Biol. (2010)*

➔ ***contact predictions within protein domains***

# How does physico-chemical complementarity evolved at the interface ?



A/B complex

Partner A

Partner B

i j

*Human* ○○○🔵○○○ ○○○🔴○○○

i et j are defined as complementary with respect to their chemical properties

🟢 Hydrophobic
🟡 Polar
🔴 Acidic
🔵 Basic

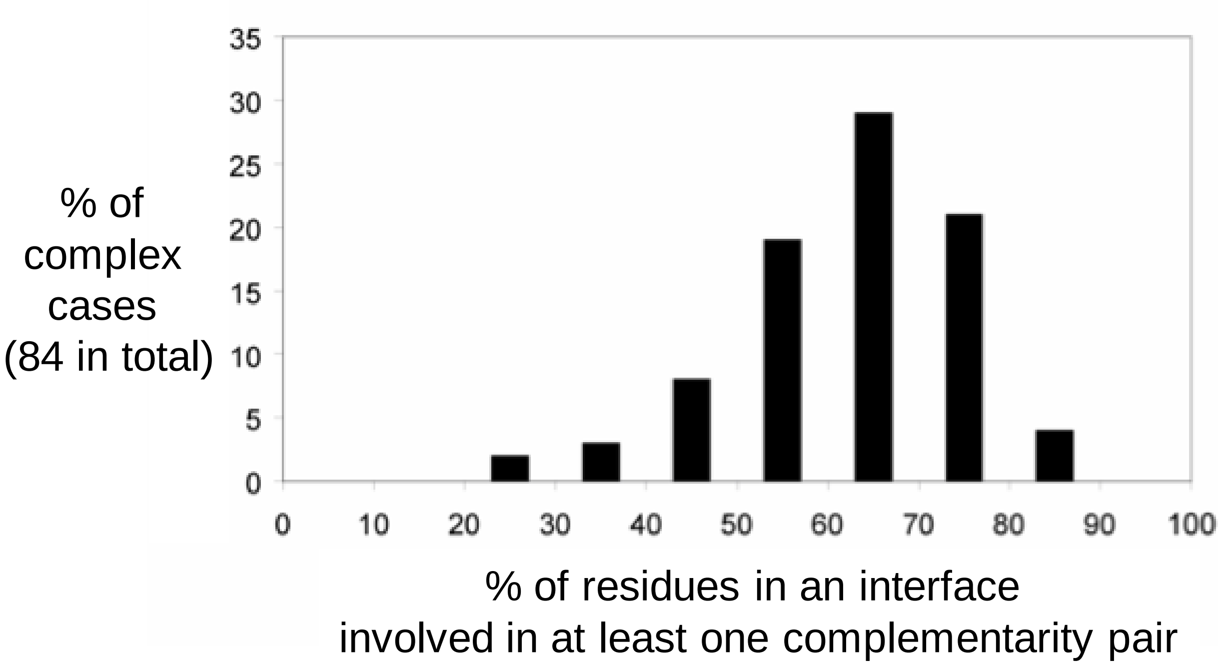
# 3 classes of complementarity for contacting pairs at complex interface

🟢 Hydrophobic
🟢 Hydrophobic

🟡 Polar
🟡 Polar

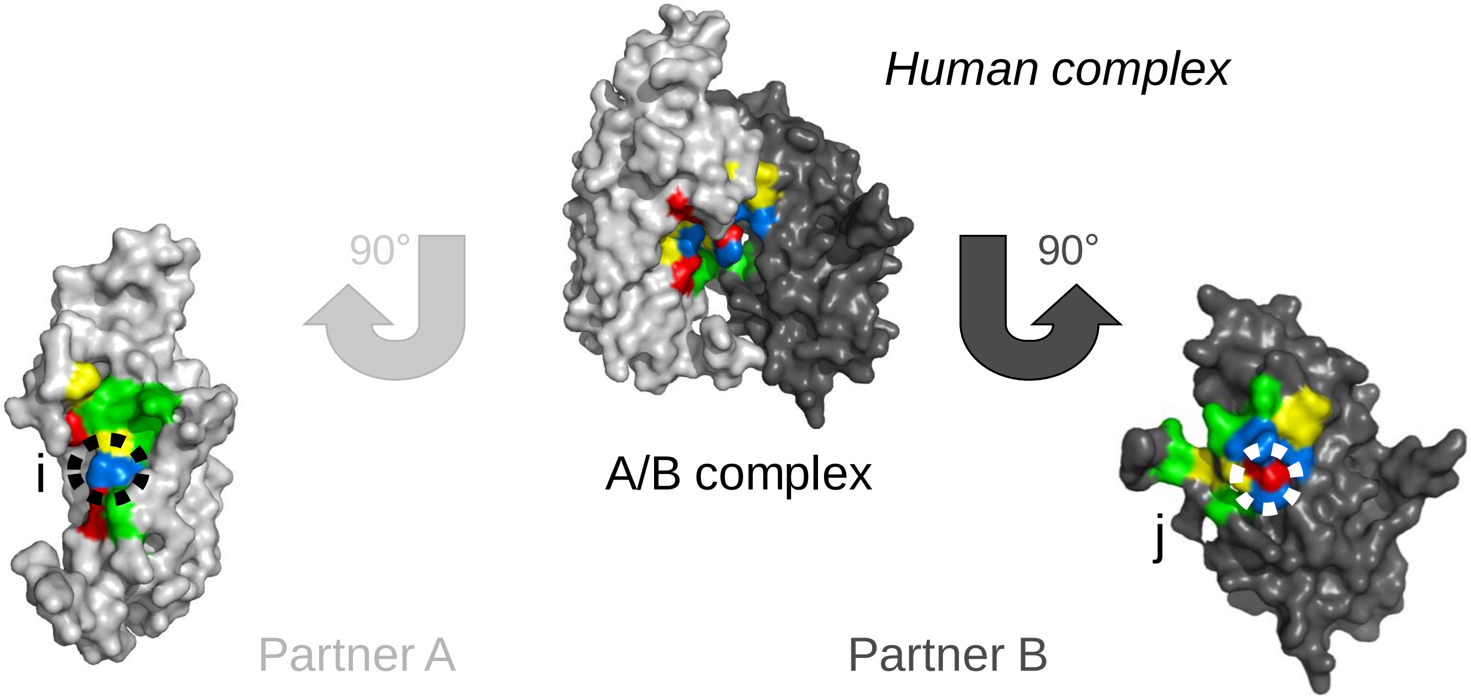
🔴 Acidic
🔵 Basic

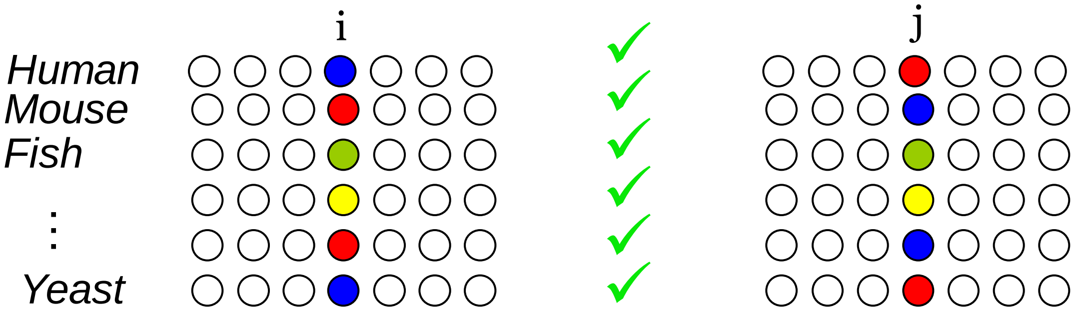
~ 62% of residues of an interface
fulfill one of the 3 complementarity classes in at least one contact



*Database of 84 interacting domains of known 3D structure*
*(extracted from PSIMAP (Kim et al. Proteins 2004))*

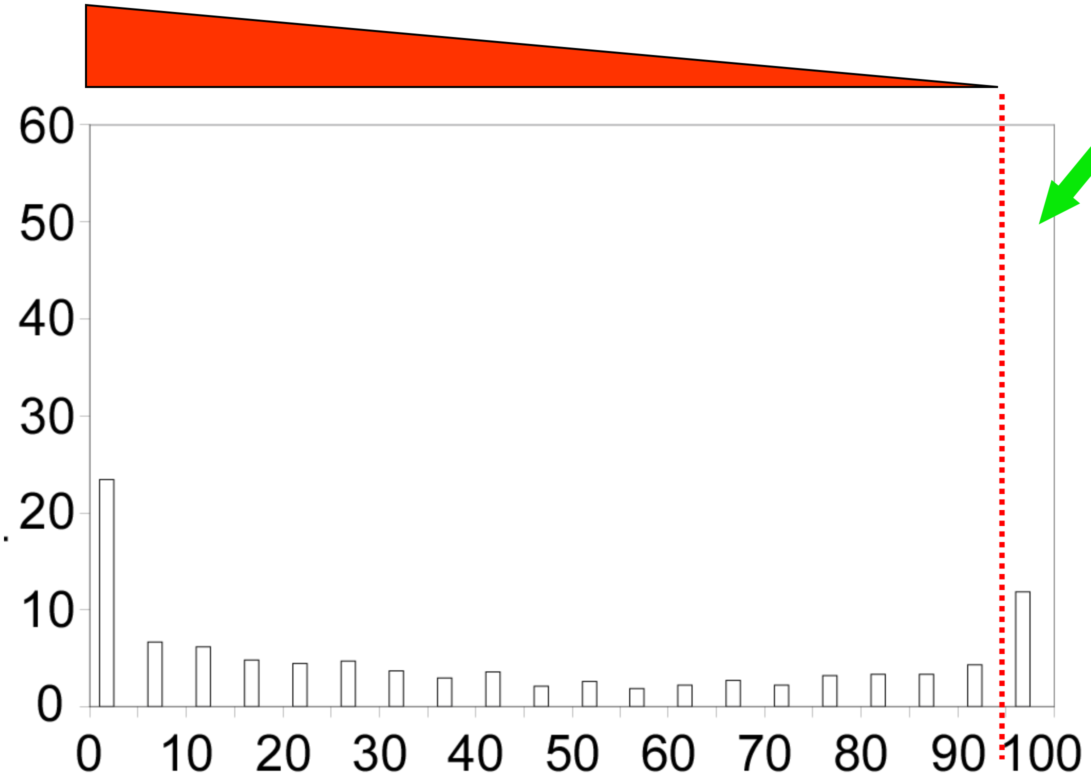
# How is this complementarity maintained through evolution ?
# Are there complementarity breaks ?





i & j are called
**“strictly complementary”**

# How is this complementarity maintained through evolution ?
# Are there complementarity breaks ?



**Complementarity is strictly preserved for only 10 % of the contacting pairs**

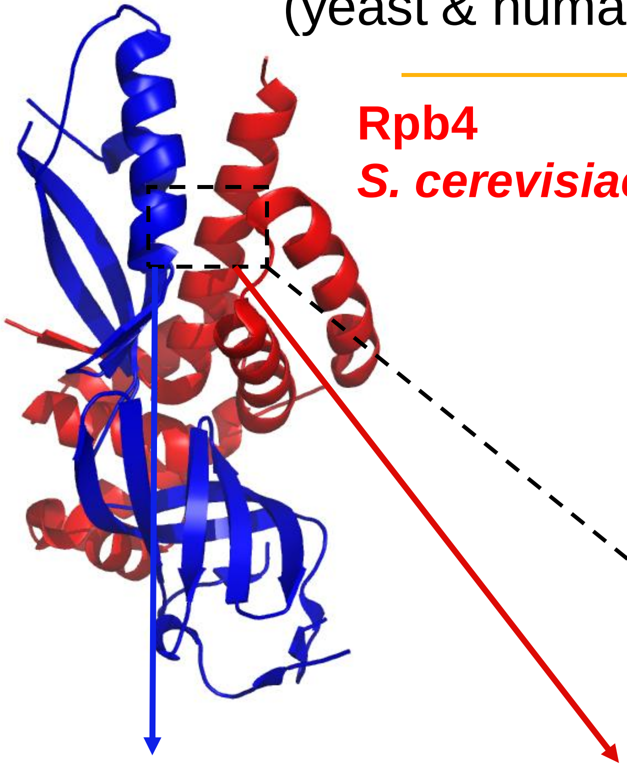
% of contacting pairs

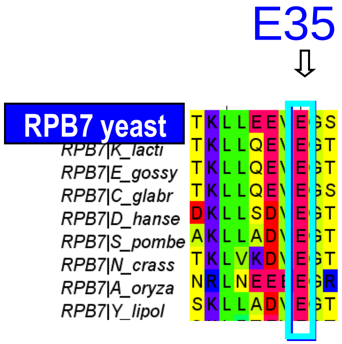
% of sequences in an alignment with preserved complementarity

# Example : Rpb4/Rpb7
# (yeast & human)

**Rpb4**
***S. cerevisiae***

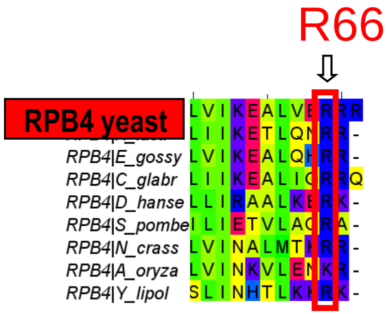
**Rpb7**
***S. cerevisiae***





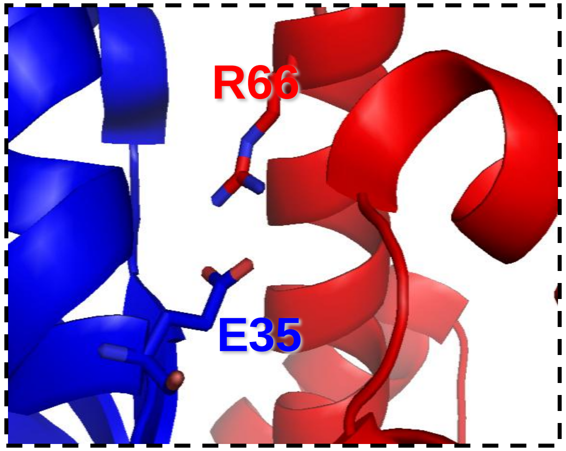
E35

| | |
|---|---|
| **RPB7 yeast** | TKLLEEVEGS |
| *RPB7\|K_lacti* | TKLLQEVEGT |
| *RPB7\|E_gossy* | TKLLQEVEGT |
| *RPB7\|C_glabr* | TKLLQEVEGS |
| *RPB7\|D_hanse* | DKLLSDVEGT |
| *RPB7\|S_pombe* | AKLLADVEGT |
| *RPB7\|N_crass* | TKLVKDVEGT |
| *RPB7\|A_oryza* | NRLNEEEEGR |
| *RPB7\|Y_lipol* | SKLLADVEGT |



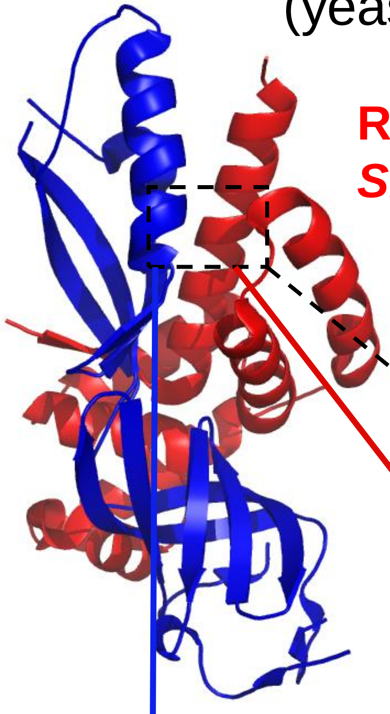
R66

| | |
|---|---|
| **RPB4 yeast** | LVIKEALVERRR |
| | LIIKETLQNRR- |
| *RPB4\|E_gossy* | LVIKEALQHRR- |
| *RPB4\|C_glabr* | LIIKEALIGRRQ |
| *RPB4\|D_hanse* | LLIRAALKERK- |
| *RPB4\|S_pombe* | ILIETVLAQRA- |
| *RPB4\|N_crass* | LVINALMTKRR- |
| *RPB4\|A_oryza* | LVINKVLENKR- |
| *RPB4\|Y_lipol* | SLINHTLKKRK- |



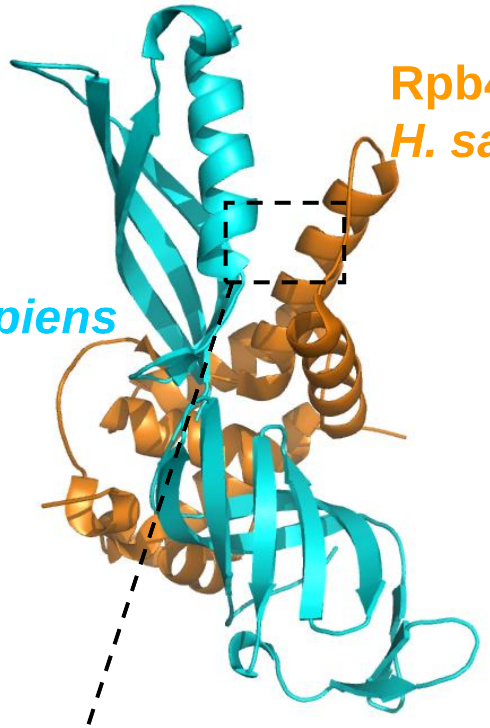
# Example : Rpb4/Rpb7 (yeast & human)

**Rpb7**
***S. cerevisiae***



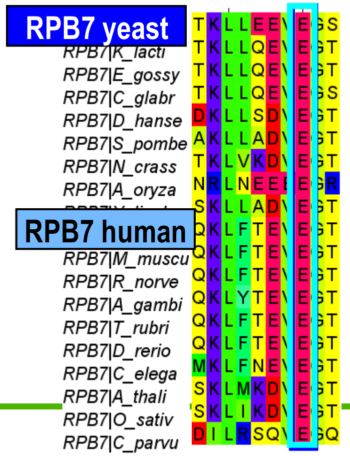
**Rpb4**
***S. cerevisiae***

**Rpb7**
***H. sapiens***



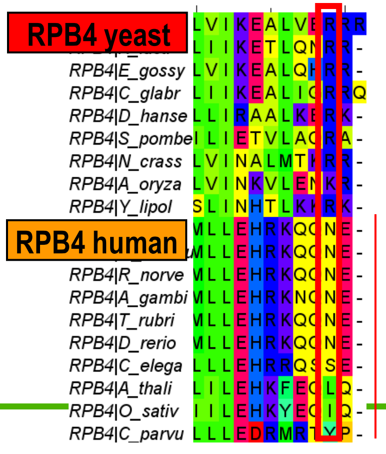
**Rpb4**
***H. sapiens***

E35

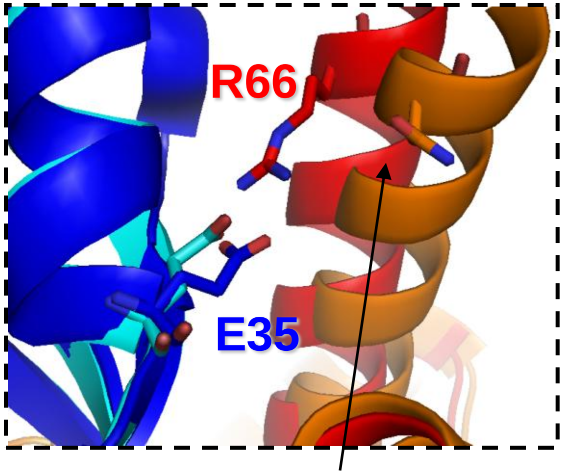


| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| RPB7 yeast | T | K | L | L | E | E | V | E | G | S |
| RPB7\|K_lacti | T | K | L | L | Q | E | V | E | G | T |
| RPB7\|E_gossy | T | K | L | L | Q | E | V | E | G | T |
| RPB7\|C_glabr | T | K | L | L | Q | E | V | E | G | S |
| RPB7\|D_hanse | D | K | L | L | S | D | V | E | G | T |
| RPB7\|S_pombe | A | K | L | L | A | D | V | E | G | T |
| RPB7\|N_crass | T | K | L | V | K | D | V | E | G | T |
| RPB7\|A_oryza | N | R | L | N | E | E | E | E | G | R |
| RPB7\|Y_lipol | S | K | L | L | A | D | V | E | G | T |
| RPB7 human | Q | K | L | F | T | E | V | E | G | T |
| RPB7\|M_muscu | Q | K | L | F | T | E | V | E | G | T |
| RPB7\|R_norve | Q | K | L | F | T | E | V | E | G | T |
| RPB7\|A_gambi | Q | K | L | Y | T | E | V | E | G | T |
| RPB7\|T_rubri | Q | K | L | F | T | E | V | E | G | T |
| RPB7\|D_rerio | Q | K | L | F | T | E | V | E | G | T |
| RPB7\|C_elega | M | K | L | F | N | E | V | E | G | T |
| RPB7\|A_thali | S | K | L | M | K | D | V | E | G | T |
| RPB7\|O_sativ | S | K | L | I | K | D | V | E | G | T |
| RPB7\|C_parvu | D | I | L | R | S | Q | V | E | G | Q |

R66

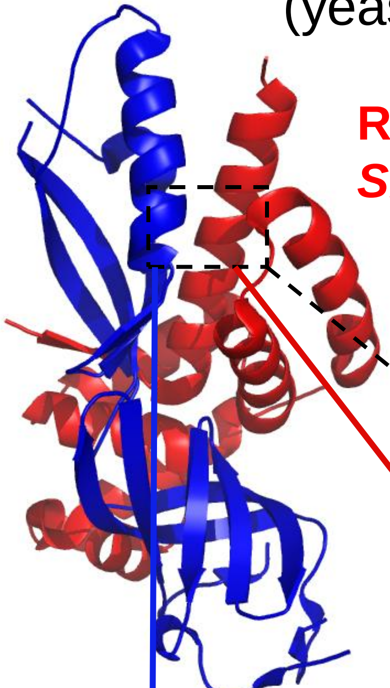


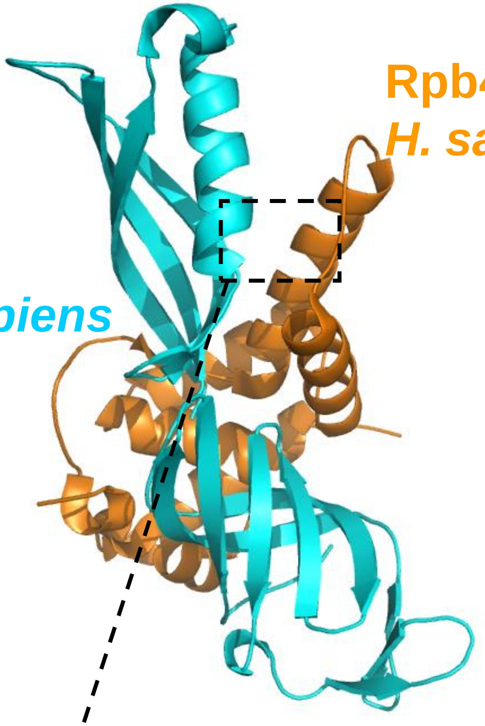
| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| RPB4 yeast | L | V | I | K | E | A | L | V | E | R | R | R |
| RPB4\|K_lacti | L | I | I | K | E | T | L | Q | N | R | R | - |
| RPB4\|E_gossy | L | V | I | K | E | A | L | Q | H | R | R | - |
| RPB4\|C_glabr | L | I | I | K | E | A | L | I | G | R | R | Q |
| RPB4\|D_hanse | L | L | I | R | A | A | L | K | E | R | K | - |
| RPB4\|S_pombe | I | L | I | E | T | V | L | A | G | R | A | - |
| RPB4\|N_crass | L | V | I | N | A | L | M | T | K | R | R | - |
| RPB4\|A_oryza | L | V | I | N | K | V | L | E | N | K | R | - |
| RPB4\|Y_lipol | S | L | I | N | H | T | L | K | K | R | K | - |
| RPB4 human | M | L | L | E | H | R | K | Q | Q | N | E | - |
| RPB4\|M_muscu | M | L | L | E | H | R | K | Q | Q | N | E | - |
| RPB4\|R_norve | M | L | L | E | H | R | K | Q | Q | N | E | - |
| RPB4\|A_gambi | M | L | L | E | H | R | K | N | Q | N | E | - |
| RPB4\|T_rubri | M | L | L | E | H | R | K | Q | Q | N | E | - |
| RPB4\|D_rerio | M | L | L | E | H | R | K | Q | Q | N | E | - |
| RPB4\|C_elega | L | L | L | E | H | R | R | Q | S | S | E | - |
| RPB4\|A_thali | L | I | L | E | H | K | F | E | Q | L | Q | - |
| RPB4\|O_sativ | I | I | L | E | H | K | Y | E | Q | I | Q | - |
| RPB4\|C_parvu | L | L | L | E | D | R | M | R | T | Y | P | - |

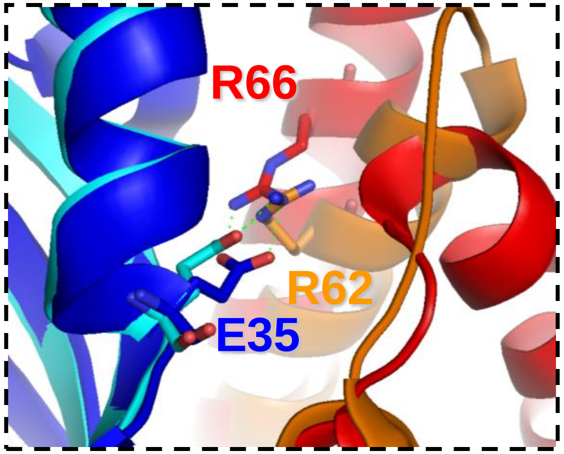


Complementarity break

# Example : Rpb4/Rpb7 (yeast & human)



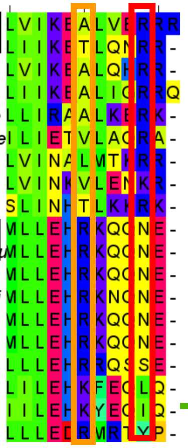




E35 ⇩

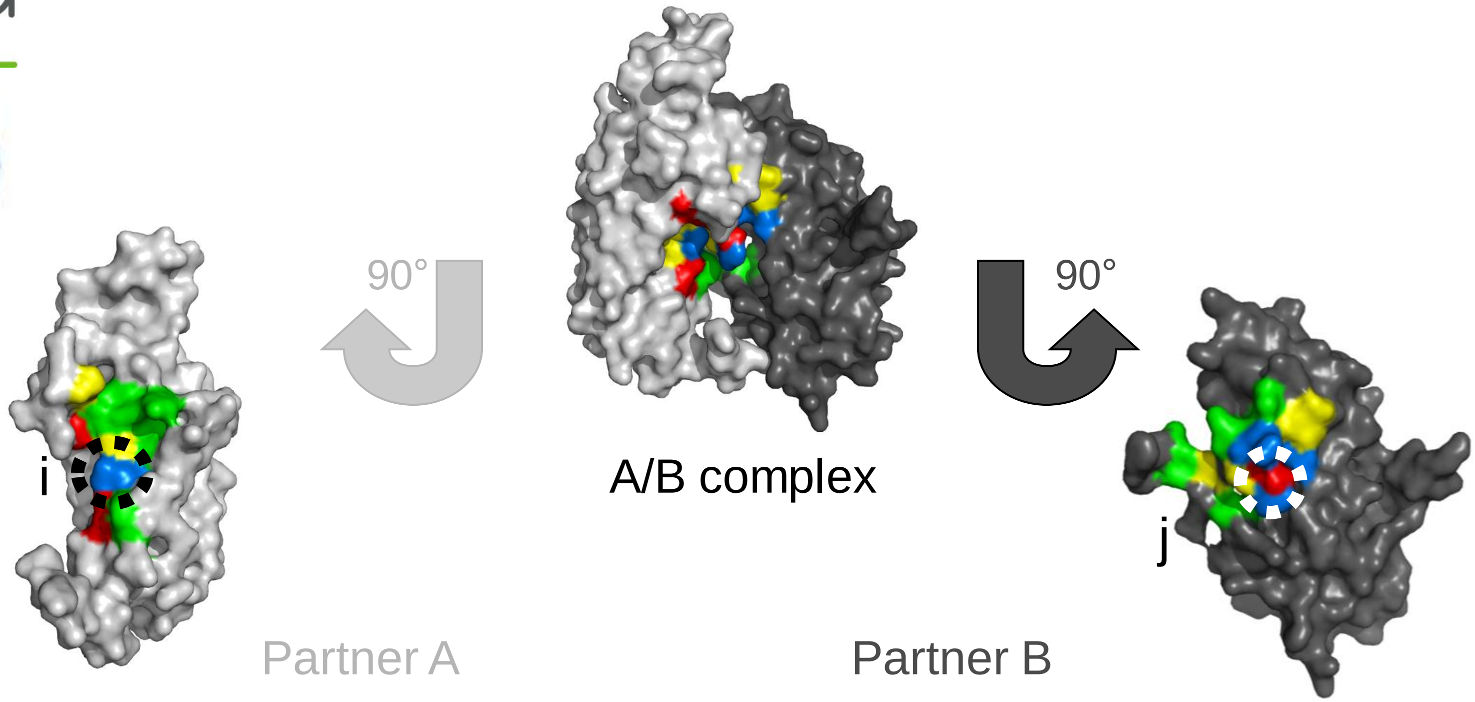
| | |
|---|---|
| **RPB7 yeast** | TKLLEEVEGS |
| RPB7\|K_lacti | TKLLQEVEGT |
| RPB7\|E_gossy | TKLLQEVEGT |
| RPB7\|C_glabr | TKLLQEVEGS |
| RPB7\|D_hanse | DKLLSDVEGT |
| RPB7\|S_pombe | AKLLADVEGT |
| RPB7\|N_crass | TKLVKDVEGT |
| RPB7\|A_oryza | NRLNEEEEGR |
| | SKLLADVEGT |
| **RPB7 human** | QKLFTEVEGT |
| RPB7\|M_muscu | QKLFTEVEGT |
| RPB7\|R_norve | QKLFTEVEGT |
| RPB7\|A_gambi | QKLYTEVEGT |
| RPB7\|T_rubri | QKLFTEVEGT |
| RPB7\|D_rerio | QKLFTEVEGT |
| RPB7\|C_elega | MKLFNEVEGT |
| RPB7\|A_thali | SKLMKDVEGT |
| RPB7\|O_sativ | SKLIKDVEGT |
| RPB7\|C_parvu | DILRSQVEGQ |

R62 R66
⇩ ⇩



| | |
|---|---|
| **RPB4 yeast** | LVIKEALVERRR |
| | LIIKETLQNRR- |
| RPB4\|E_gossy | LVIKEALQHRR- |
| RPB4\|C_glabr | LIIKEALIGRRQ |
| RPB4\|D_hanse | LLIRAALKERK- |
| RPB4\|S_pombe | ILIETVLAGRA- |
| RPB4\|N_crass | LVINALMTKRR- |
| RPB4\|A_oryza | LVINKVLENKR- |
| RPB4\|Y_lipol | SLINHTLKKRK- |
| **RPB4 human** | MLLEHRKQGNE- |
| | MLLEHRKQGNE- |
| RPB4\|R_norve | MLLEHRKQGNE- |
| RPB4\|A_gambi | MLLEHRKNGNE- |
| RPB4\|T_rubri | MLLEHRKQGNE- |
| RPB4\|D_rerio | MLLEHRKQGNE- |
| RPB4\|C_elega | LLLEHRRQSSE- |
| RPB4\|A_thali | LILEHKFEGLQ- |
| RPB4\|O_sativ | IILEHKYEGIQ- |
| RPB4\|C_parvu | LLLEDRMRTYP- |

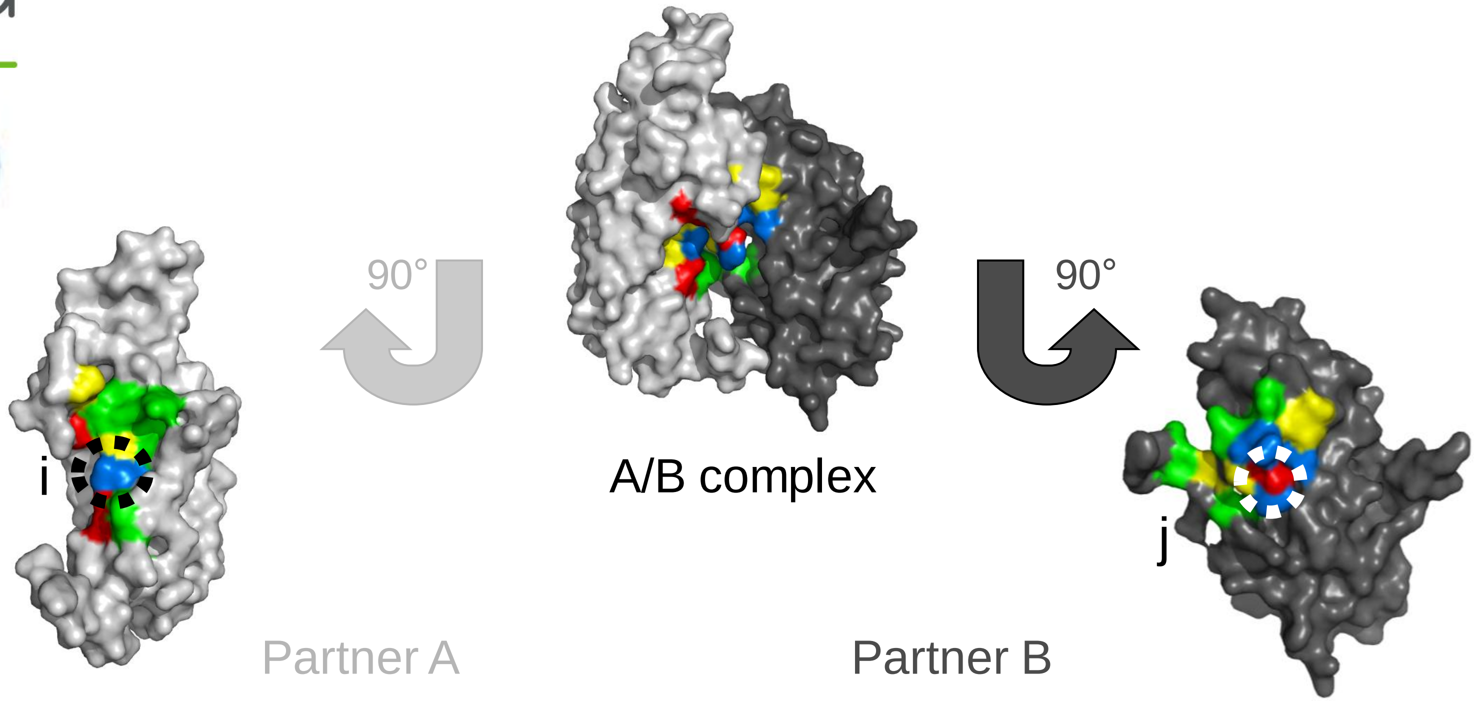
# How is this complementarity maintained through evolution ?
# Are there complementarity breaks ?

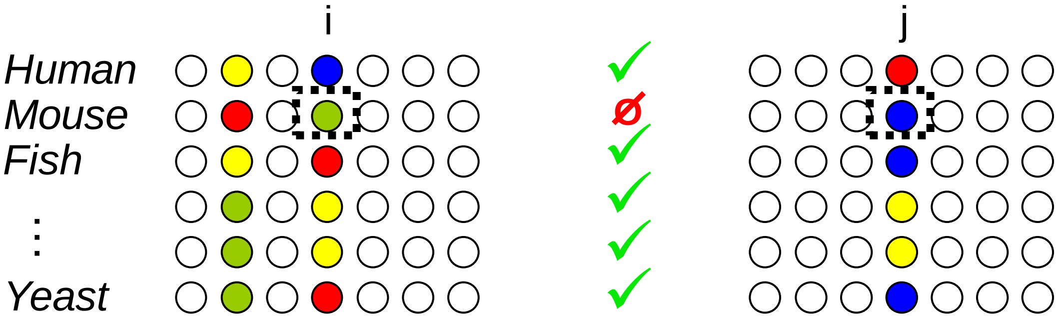


| | | | | i | | | | | | | | j | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Human* | ○ | 🟡 | ○ | 🔵 | ○ | ○ | ○ | ✓ | ○ | ○ | ○ | 🔴 | ○ | ○ | ○ |
| *Mouse* | ○ | 🔴 | ○ | 🟢 | ○ | ○ | ○ | | ○ | ○ | ○ | 🔵 | ○ | ○ | ○ |
| *Fish* | ○ | 🟡 | ○ | 🔴 | ○ | ○ | ○ | | ○ | ○ | ○ | 🔵 | ○ | ○ | ○ |
| ⋮ | ○ | 🟢 | ○ | 🟡 | ○ | ○ | ○ | | ○ | ○ | ○ | 🟡 | ○ | ○ | ○ |
| | ○ | 🟢 | ○ | 🟡 | ○ | ○ | ○ | | ○ | ○ | ○ | 🟡 | ○ | ○ | ○ |
| *Yeast* | ○ | 🟢 | ○ | 🔴 | ○ | ○ | ○ | | ○ | ○ | ○ | 🔵 | ○ | ○ | ○ |

🟢 Hydrophobic
🟡 Polar
🔴 Acidic
🔵 Basic

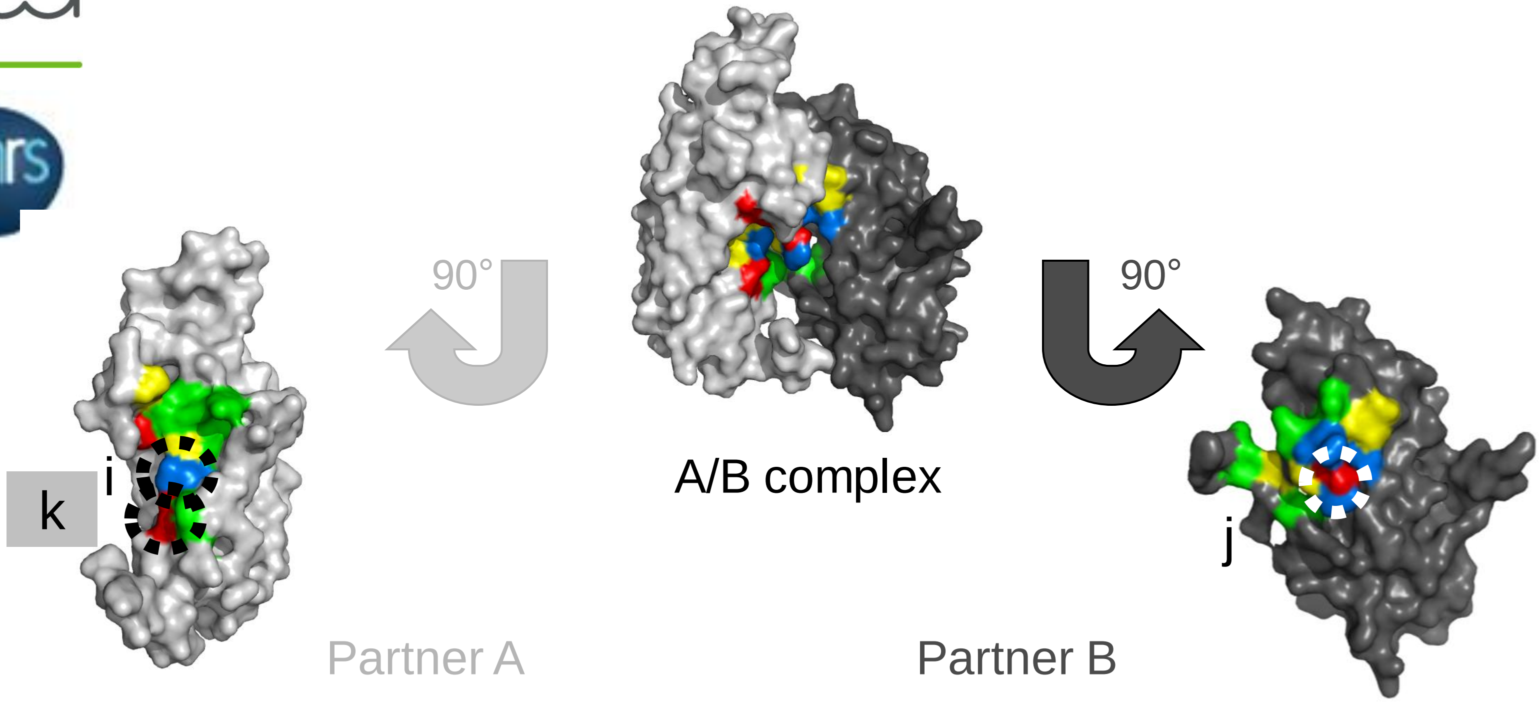
# How is this complementarity maintained through evolution ?
# Are there complementarity breaks ?

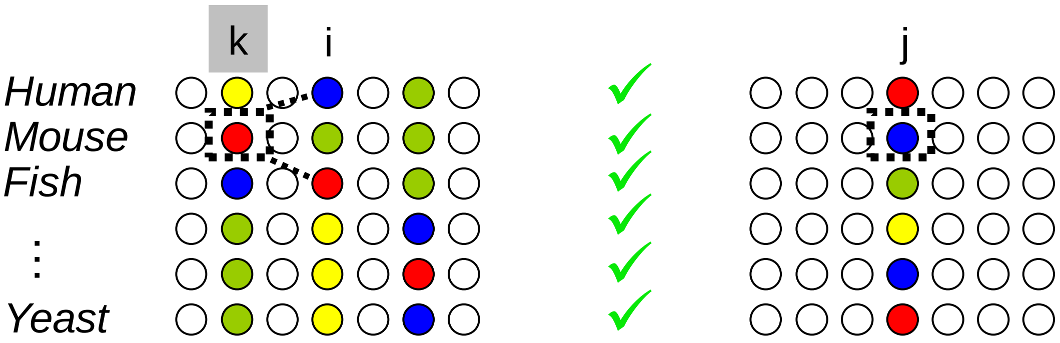




1 Complementarity break

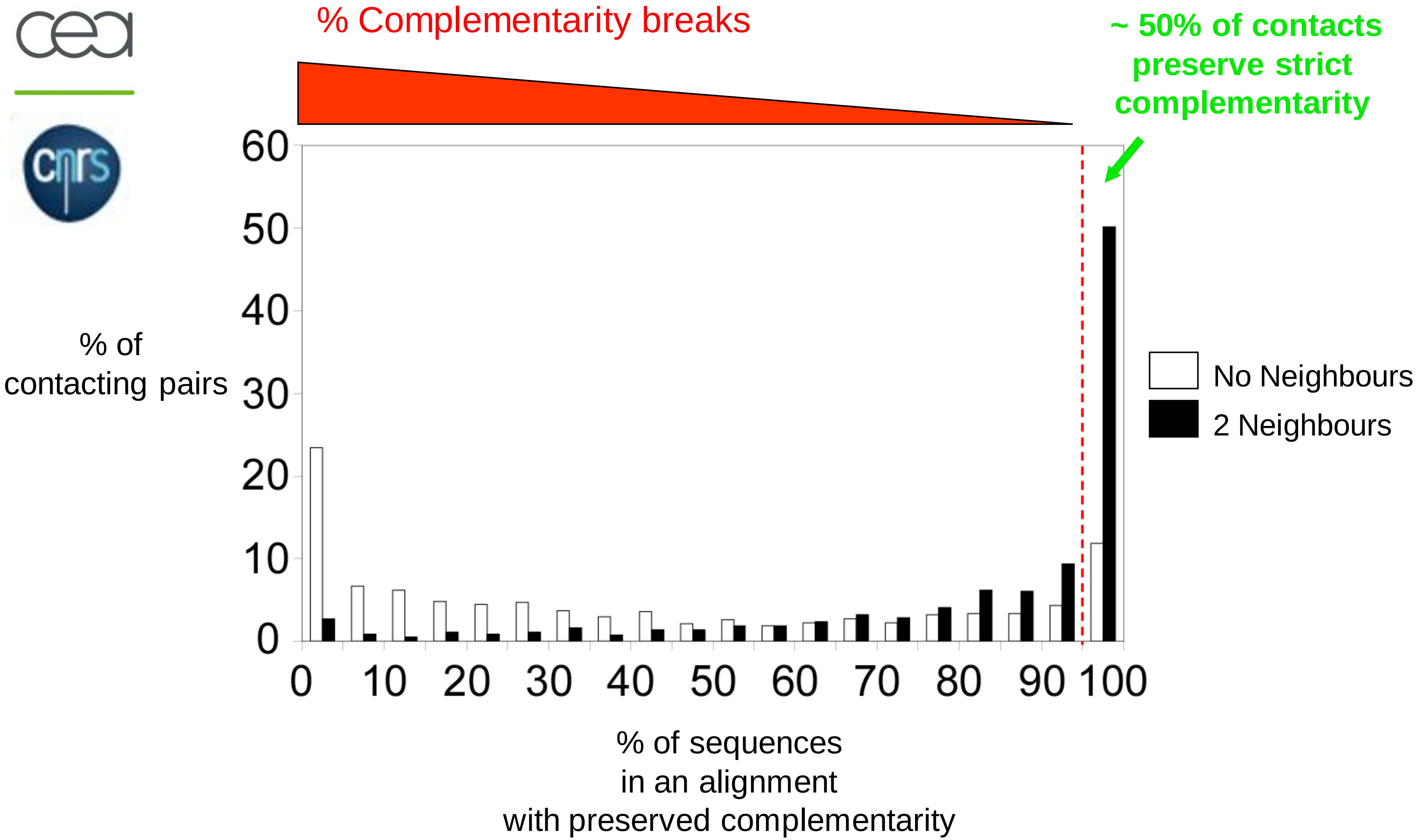
# How is this complementarity maintained through evolution ? Are there complementarity breaks ?



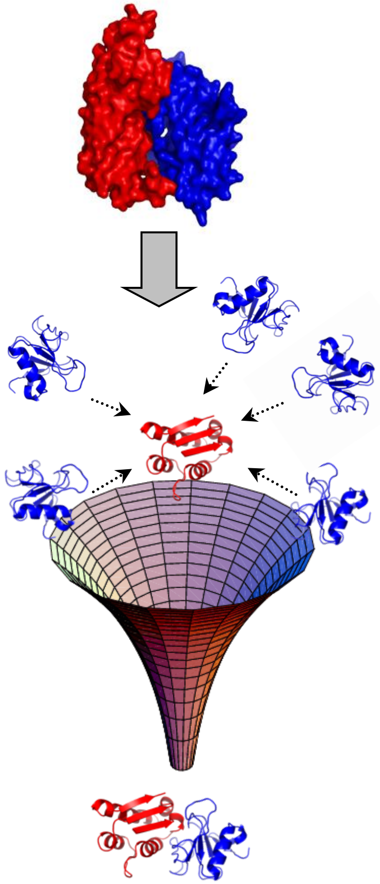


Structural Neighbours may compensate for loss of complementarity

# Impact of considering 2 structural neighbours in the complementarity analysis ?



# Can we discriminate false complexes from correct ones ?



Learning set :
60 complexes

Generate 10000 decoys
(~ 7 are « correct »)

Rigid-body docking with shape complementarity optimization (FTDock)

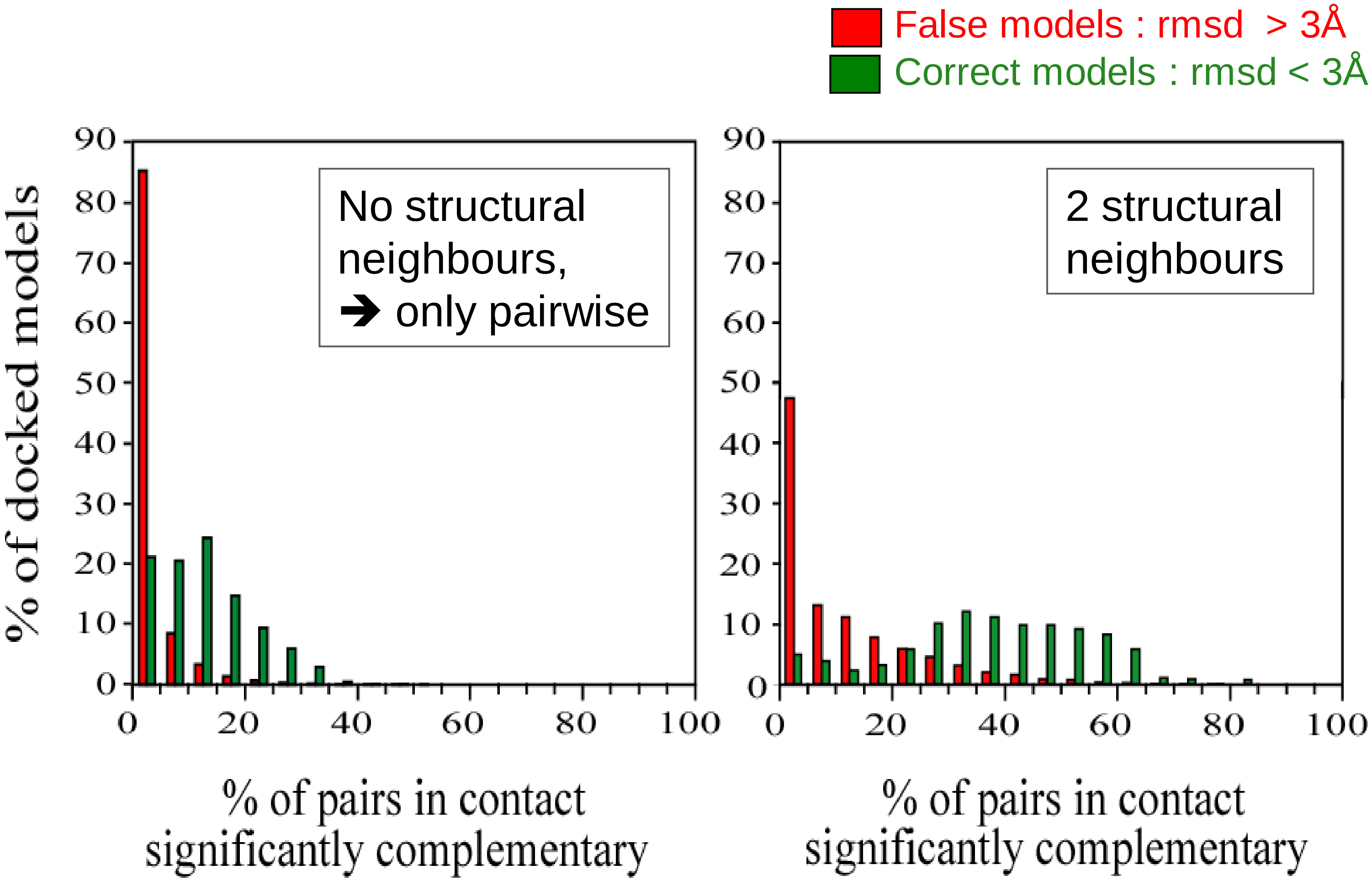
Evolutionary Filter based on complementarity

% of pairs with strict complementarity

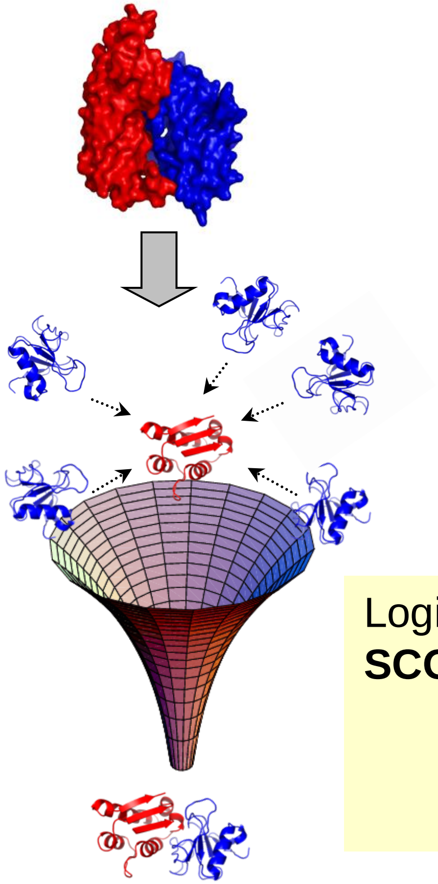
# Structural neighbours are important to account for the conservation of physico-chemical complementarity

- **Number of models generated :** 10 000

  On average 7 “correct models” within the 10 000 false ones



# Can we discriminate false complexes from correct ones ?



Learning set :
60 complexes

Generate 10000 decoys
(~ 7 are « correct »)

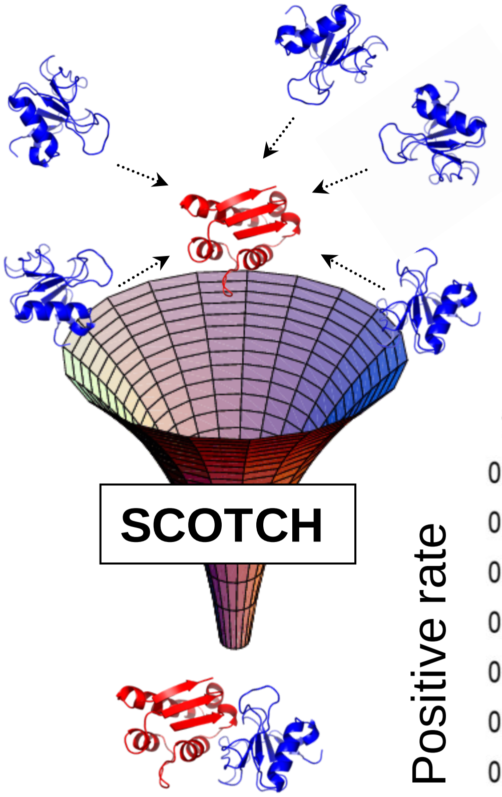
Order the 10000 solutions with respect to the score **SCOTCH**

Logistic-regression to adjust $\alpha$,$\beta$ & $\gamma$

**SCOTCH** $= \alpha\cdot$(%Complementarity)
$+ \beta\cdot$(Npairs)
$+ \gamma\cdot$(Highly_variable_pos)

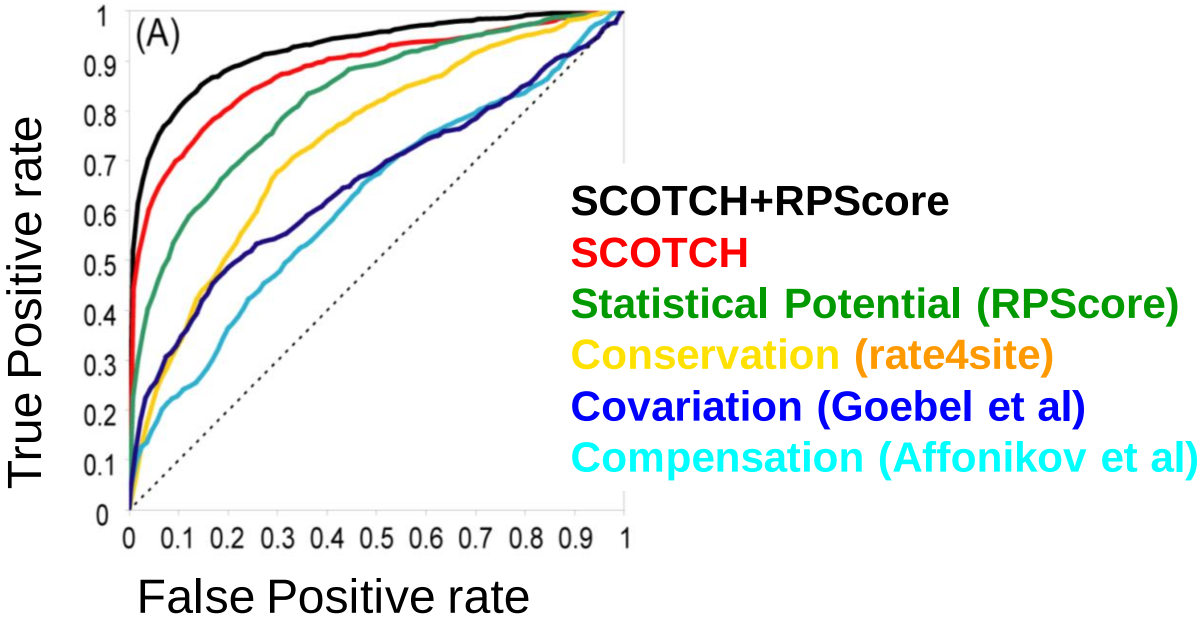
Rank of the « correct » models ?

# SCOTCH performance to filter a set of models at low resolution



**Identification of near native complexes**

- 78 % of 130 intermolecular complexes blind tested contained a near-native model among the 10 best ranked solutions.



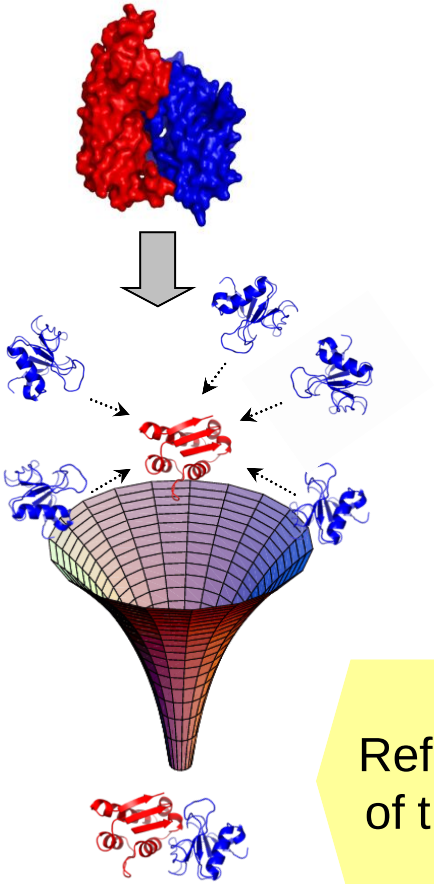
*Madaoui & Guerois, PNAS (2008)*

➔ Evolutionary information can greatly help the identification of near-native assembly modes between proteins

➔ Structural neighbours are key features to account for the structural plasticity of the complex interface throughout evolution

➔ This low-medium resolution step can seed further exploration at high resolution to identify the most likely complex.

# Can we discriminate false complexes from correct ones ?



Generate
$10^3 – 10^4$ decoys

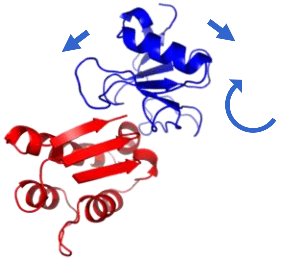
Select
~ 10 models

1 final solution

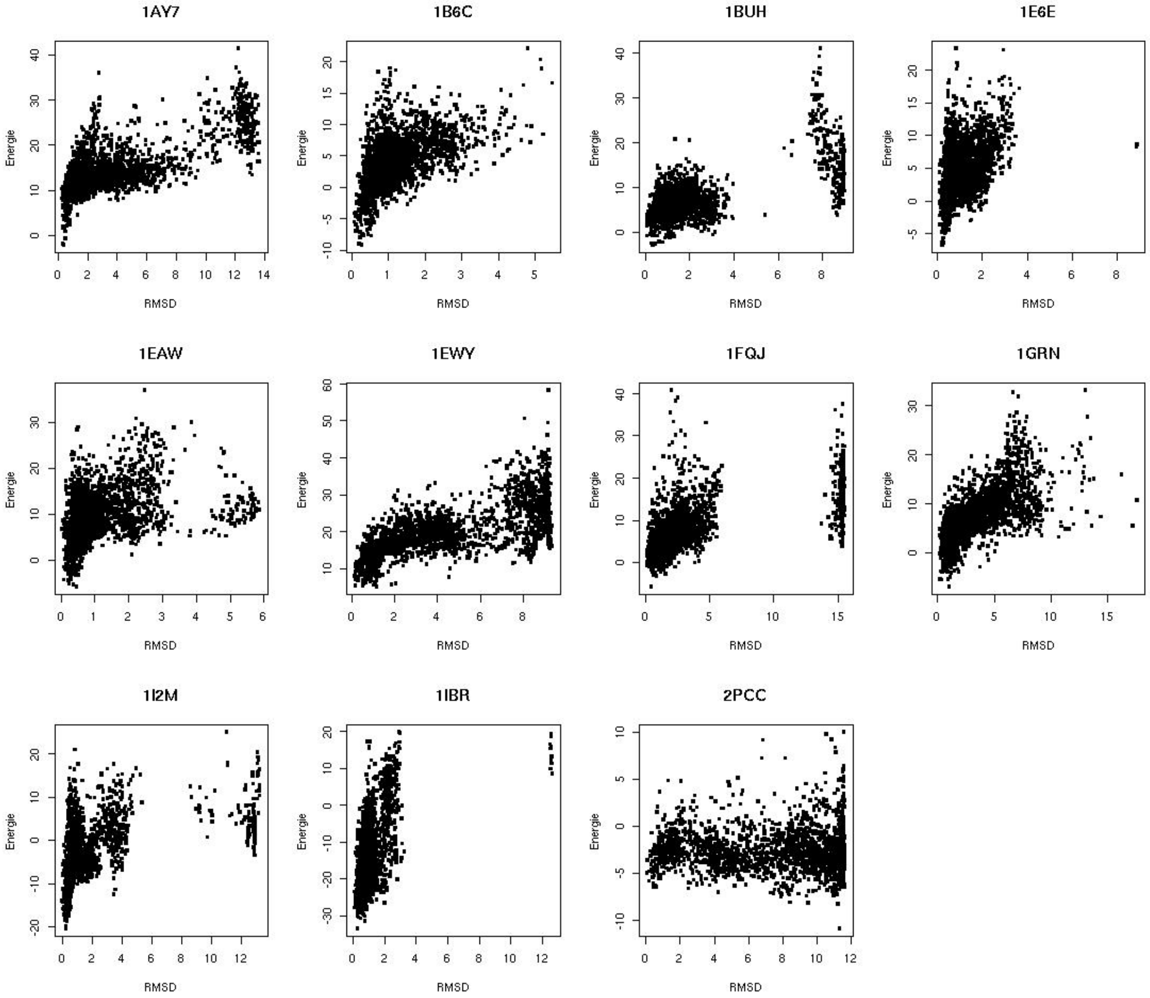
Refinement and selection
of the most likely model ?

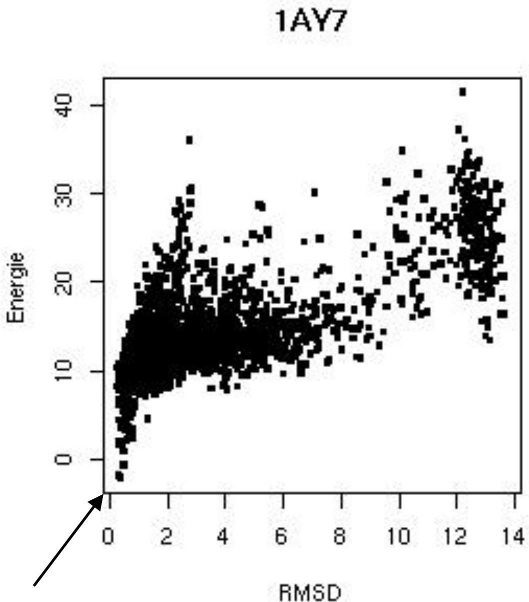
# How to identify relevant models after Coarse-grained filtering ?
# … Looking for the funnel

Small perturbations
Side-chain optimisation
Full-atom description



Analysis on known X-ray structures of complexes





Native
structure

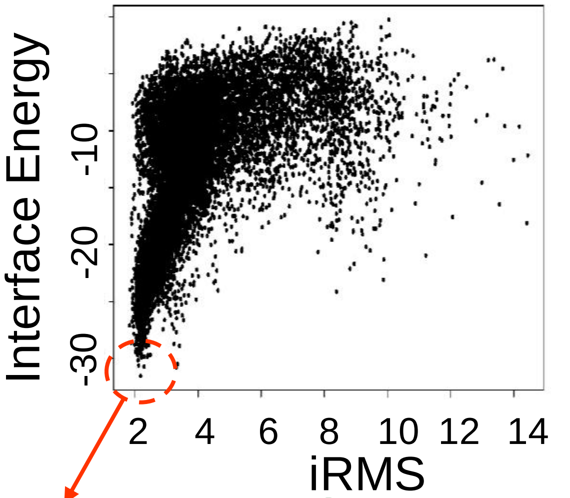
Scored using FoldX *(Guerois et al JMB 2002)* or Rosetta *(Baker’s group)*

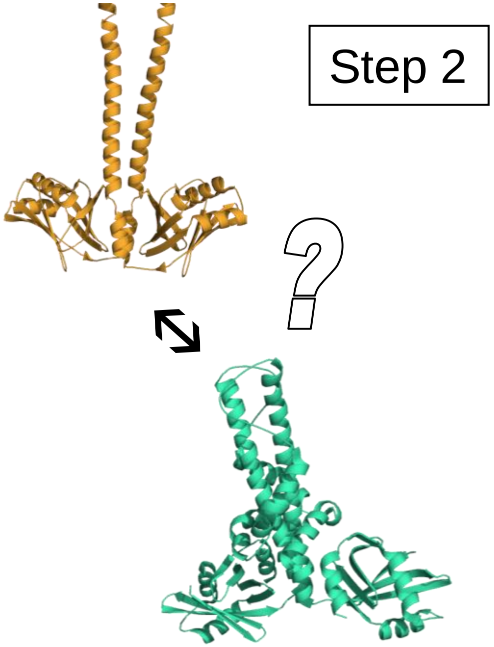
# An Example of Prediction
# Exploiting Evolution and Energy Calculations

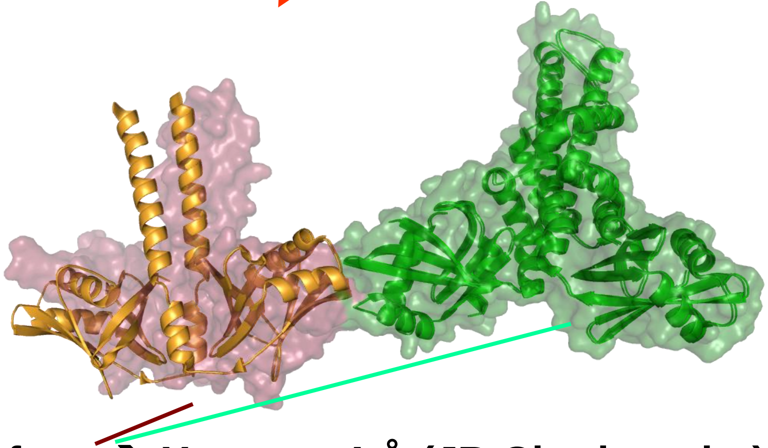
## DNA repair complex (Non-homologous End Joining)

*Coll. JB Charbonnier (LBSR)*

| | |
|---|---|
| Step 1 | Filter solutions using evolutionnary constraints |
| Step 2 | Local perturbations,<br>Optimisations of the interactions<br>… search for funnels |


Rosetta Score (min vs all)

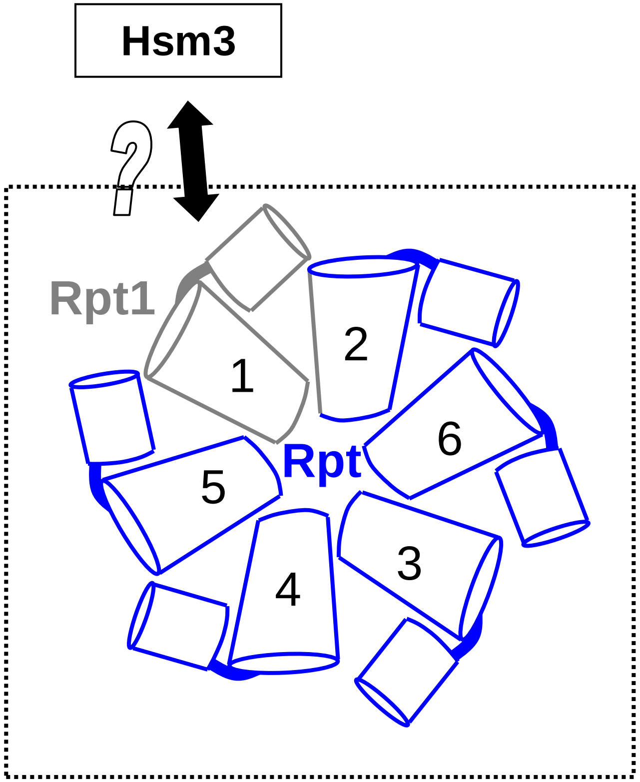


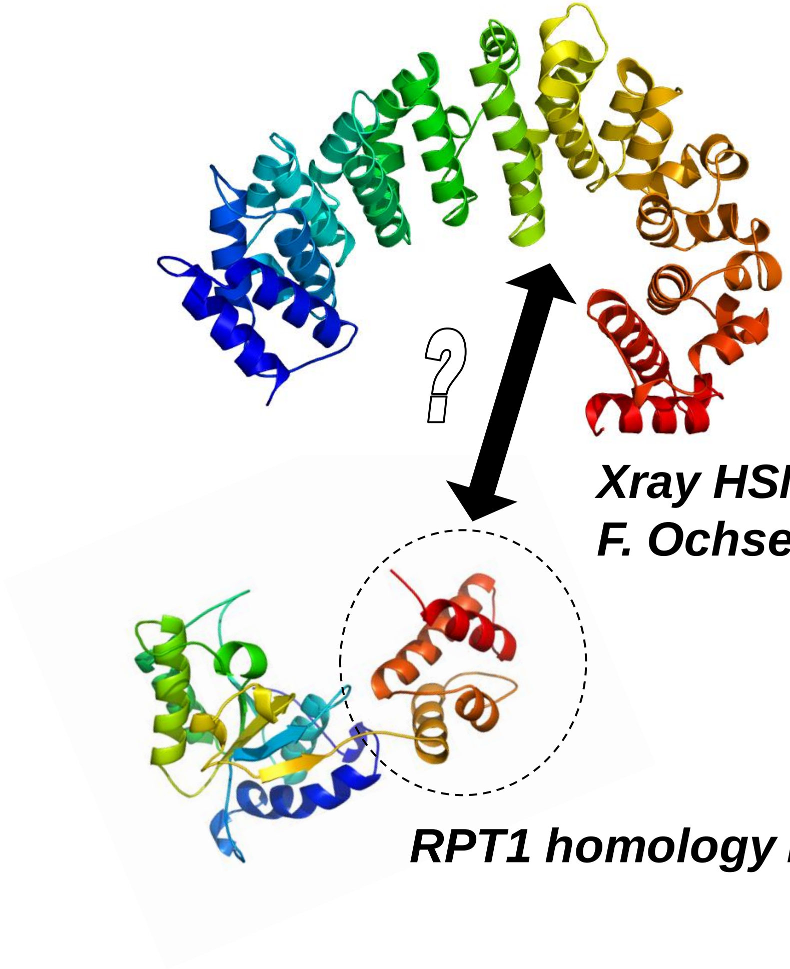


*G. Faure in Malivert et al, JBC (2010)*

**Surface ➔ Xray at ~4 Å (JB Charbonnier)**

# How does Hsm3 specifically interact with only one subunit of the 19S proteasome ?



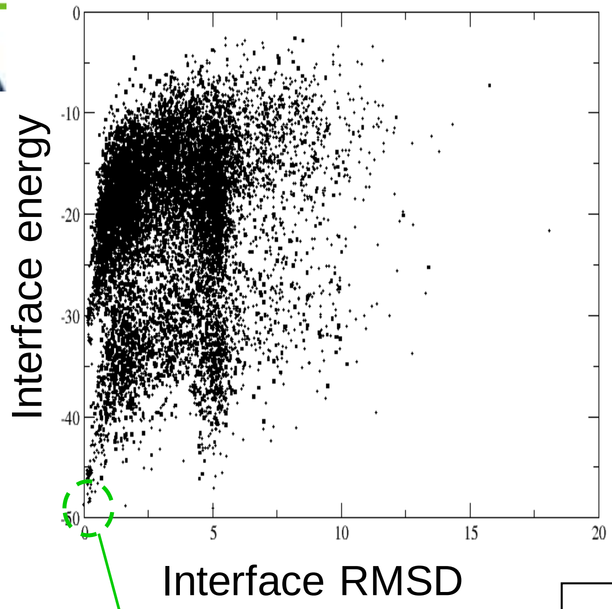


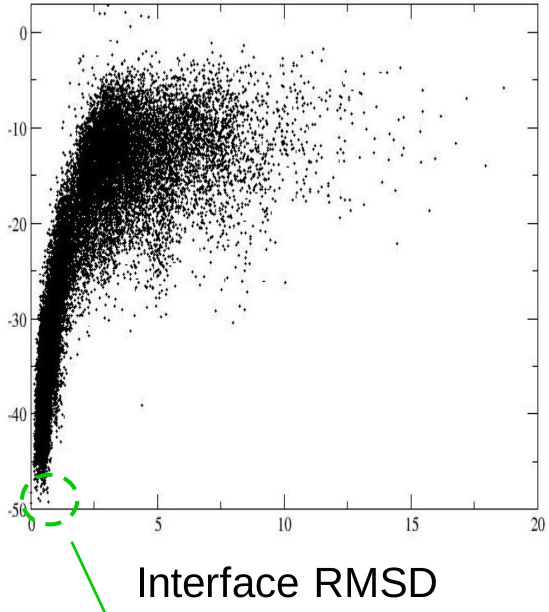
***Xray HSM3***
***F. Ochsenbein***

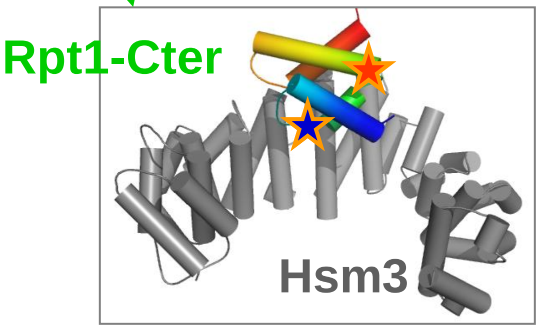
***RPT1 homology model***

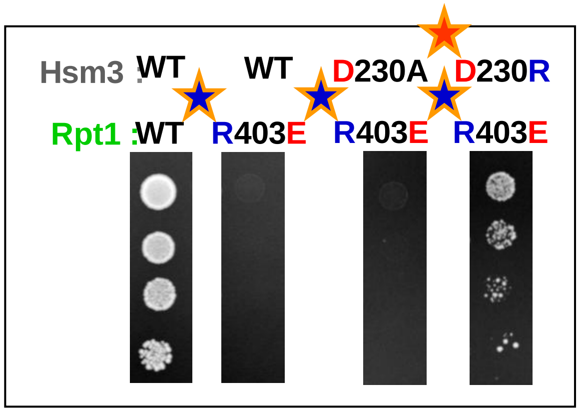
# Two major classes of models give rise to energy funnels … Compensatory mutants help discriminating them

Interface energy (Rosetta score)
vs Interface RMSD

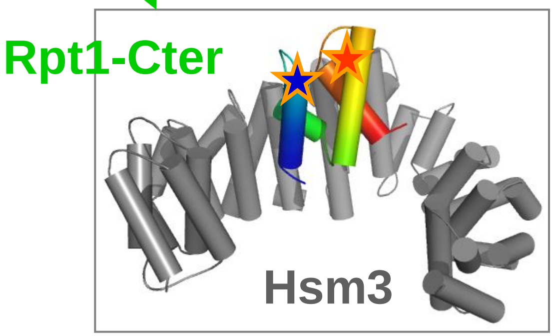




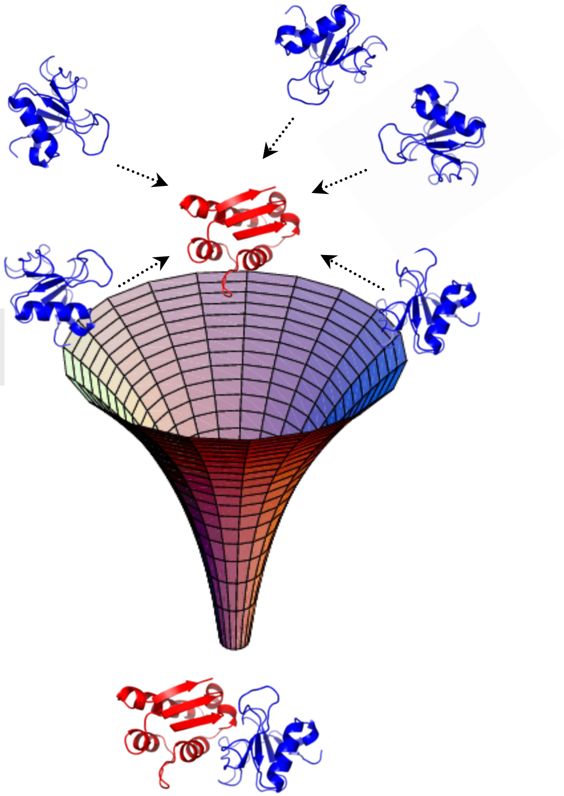




*A. Peyroche's group*



# The coarse-grained selection step further need to be improved

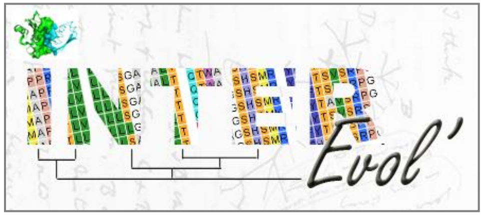


$10^3 – 10^4$ decoys

~ 10 decoys

1 final solution

Structural plasticity ←→ Evolutionary properties

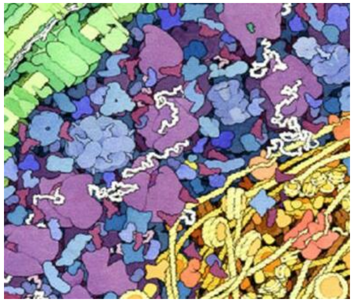


Structural Interologs database

*(Faure G et al, in prep)*

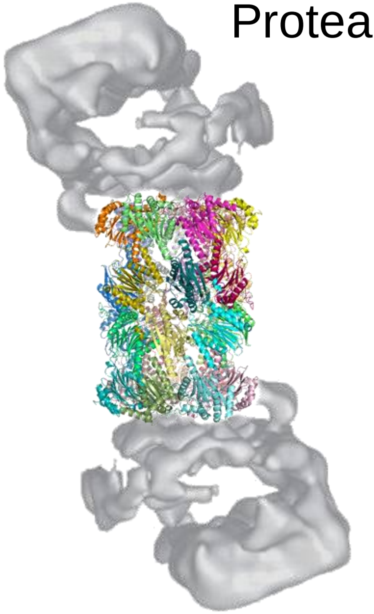
Final discrimination
(sampling and scoring the structure
using evolutionary based information)

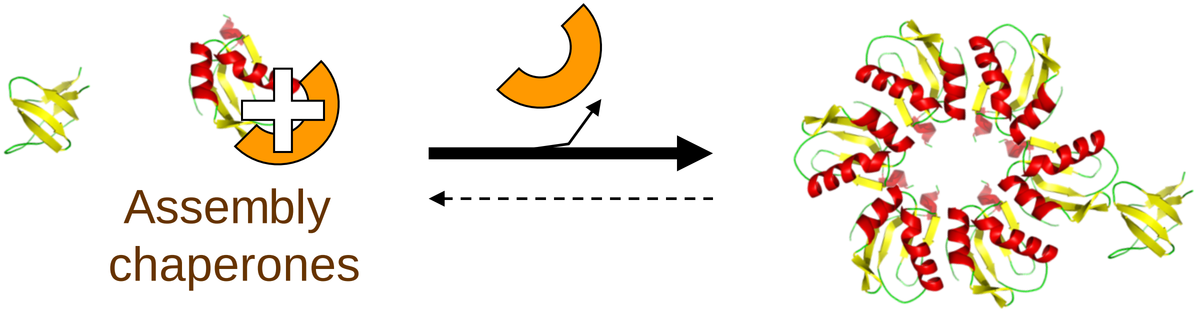
# Effects of perturbing assembly chaperones action ?



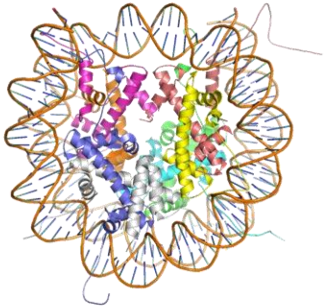
Large molecular assemblies

Proteasome



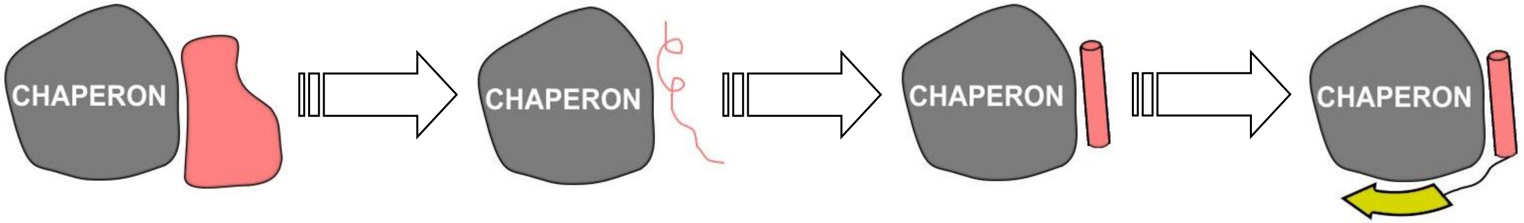


Assembly chaperones



Nucleosome

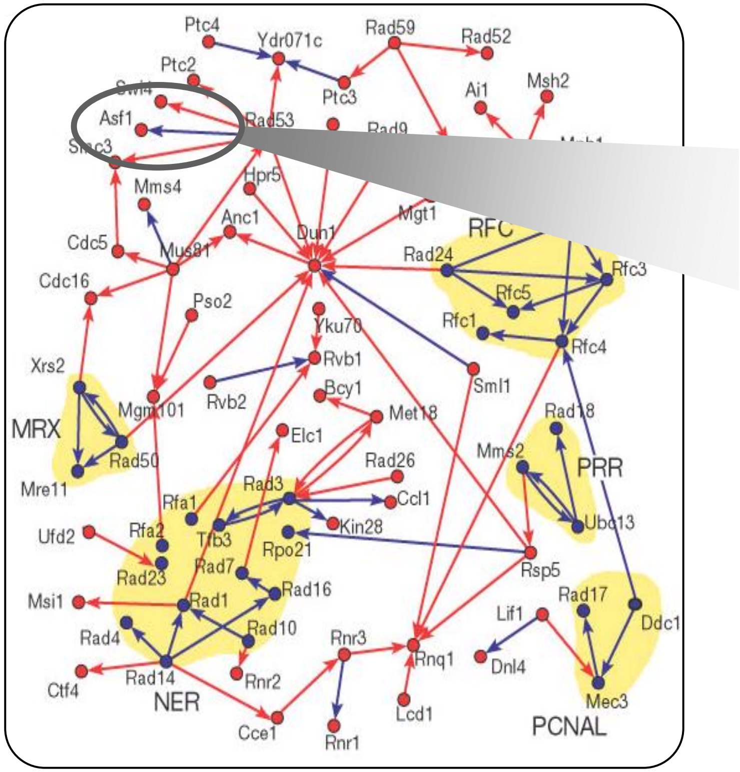
# Exploiting high resolution modelling methodologies for the design of chaperone inhibitors

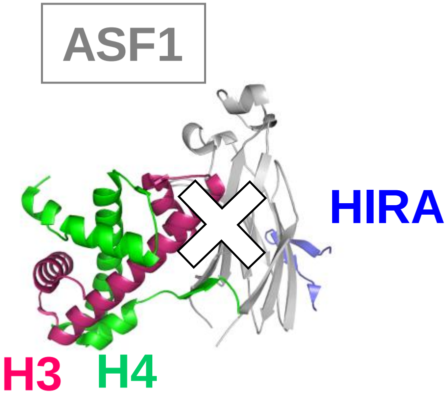


Extraction of
the binding epitope

Epitope stabilization
through design

Epitope tethering



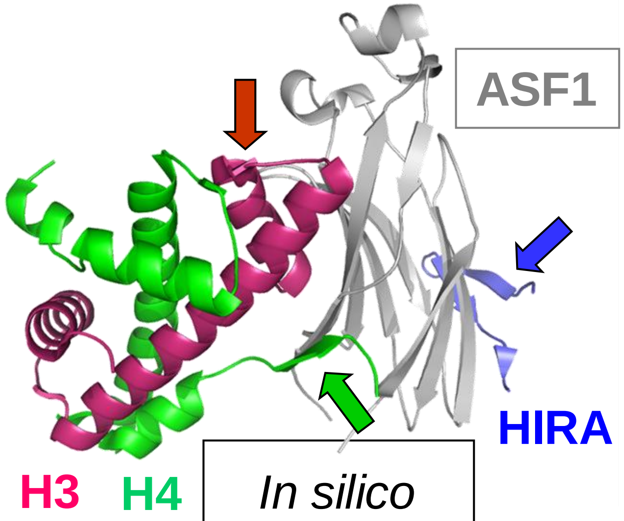


Inhibiting Asf1 : Nucleosome Assembly Chaperone

→ Perturbate nucleosome assembly process

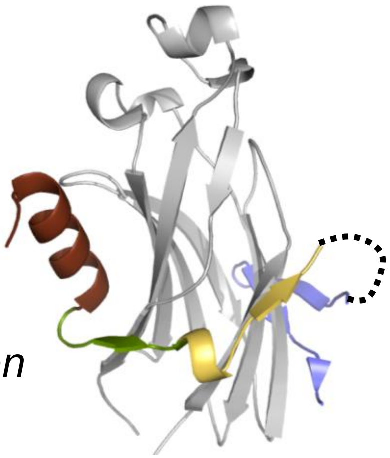
→ Cell-cycle deffects ?

# In silico design of protein-protein interaction inhibitors

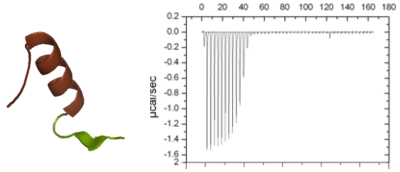


*Nucleosome assembly chaperone*

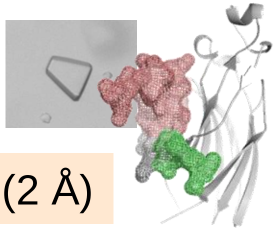
*In silico* designed inhibitor



- *Peptide stabilization*
- *Interface optimization*



10 µM ➔ Kd=50 nM



Xray (2 Å)

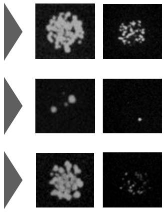
Phenotypes

Stress - +

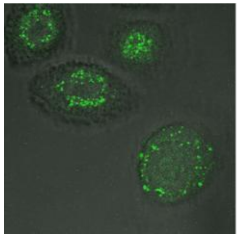
WT

Asf1Δ

WT +inhibiteur



Vectorisation with CPP



***in vitro & in vivo***

***A. Gaubert & F. Ochsenbein***

# Acknowledgements

**Françoise Ochsenbein, PI**
**(NMR, Xray)**

**Guilhem Faure, PhD**
**(Interface co-evolution)**

**Albane Gaubert, PhD**
**(Interaction inhibitors design)**

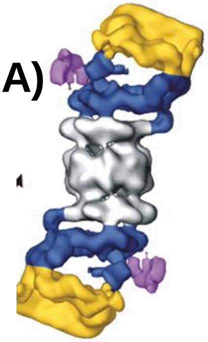
**Anne Lopes, Postdoc**
**(Remote homology detection)**

**Hocine Madaoui,**
**Emmanuelle Becker**

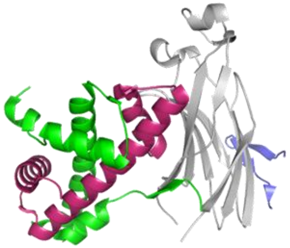
**Tools : [http://biodev.cea.fr/lbsr/](http://biodev.cea.fr/lbsr/)**

**Structural Biology and Radiobiology Lab**

**Anne Peyroche (CEA)**
Benoit le Tallec
Bénedicte Barrault



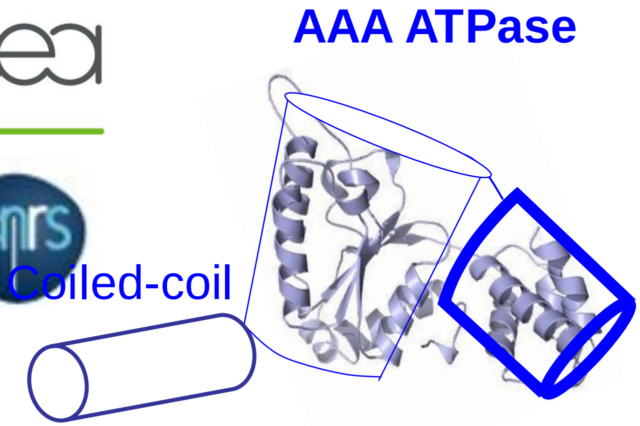
**Vasily Ogryzko**
**(Institut Gustave Roussy)**

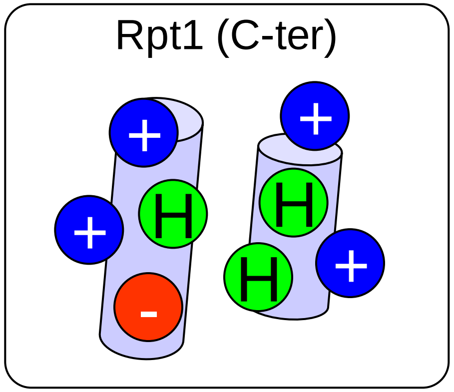


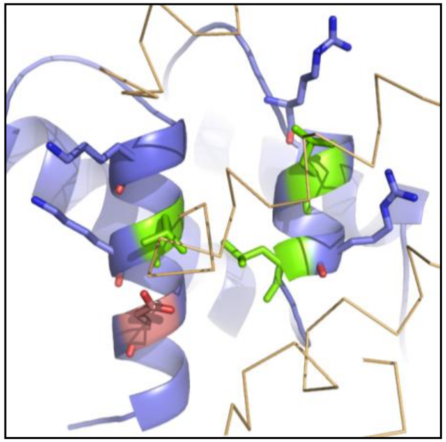
**Marie-Agnès Petit**
**(INRA Jouy-en-Josas)**

**Pablo Radicella**
**(CEA Fontenay)**

# Conservation of the C-terminal motif in other Rpts : Rpt2 -> Rpt6 ?

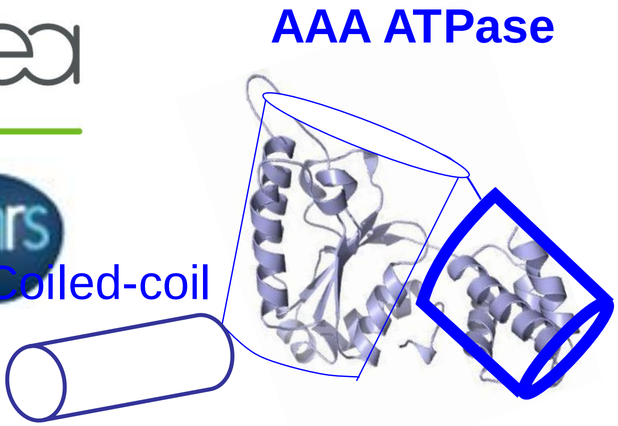


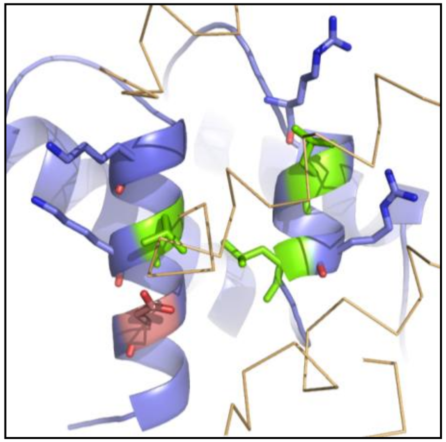


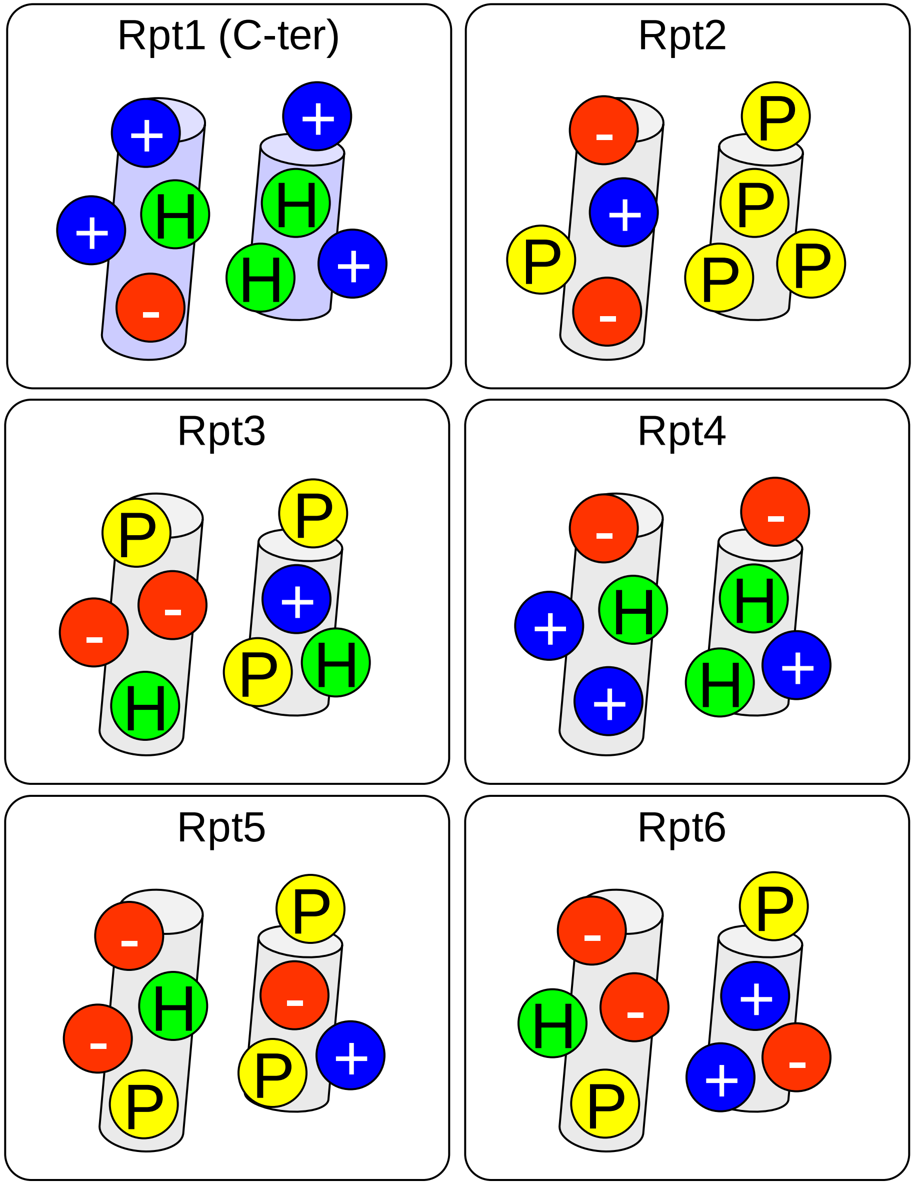


- + basic
- - acidic
- P polaire
- H hydrophobe

# Conservation du motif C-terminal dans les autres Rpts : Rpt2 -> Rpt6 ?



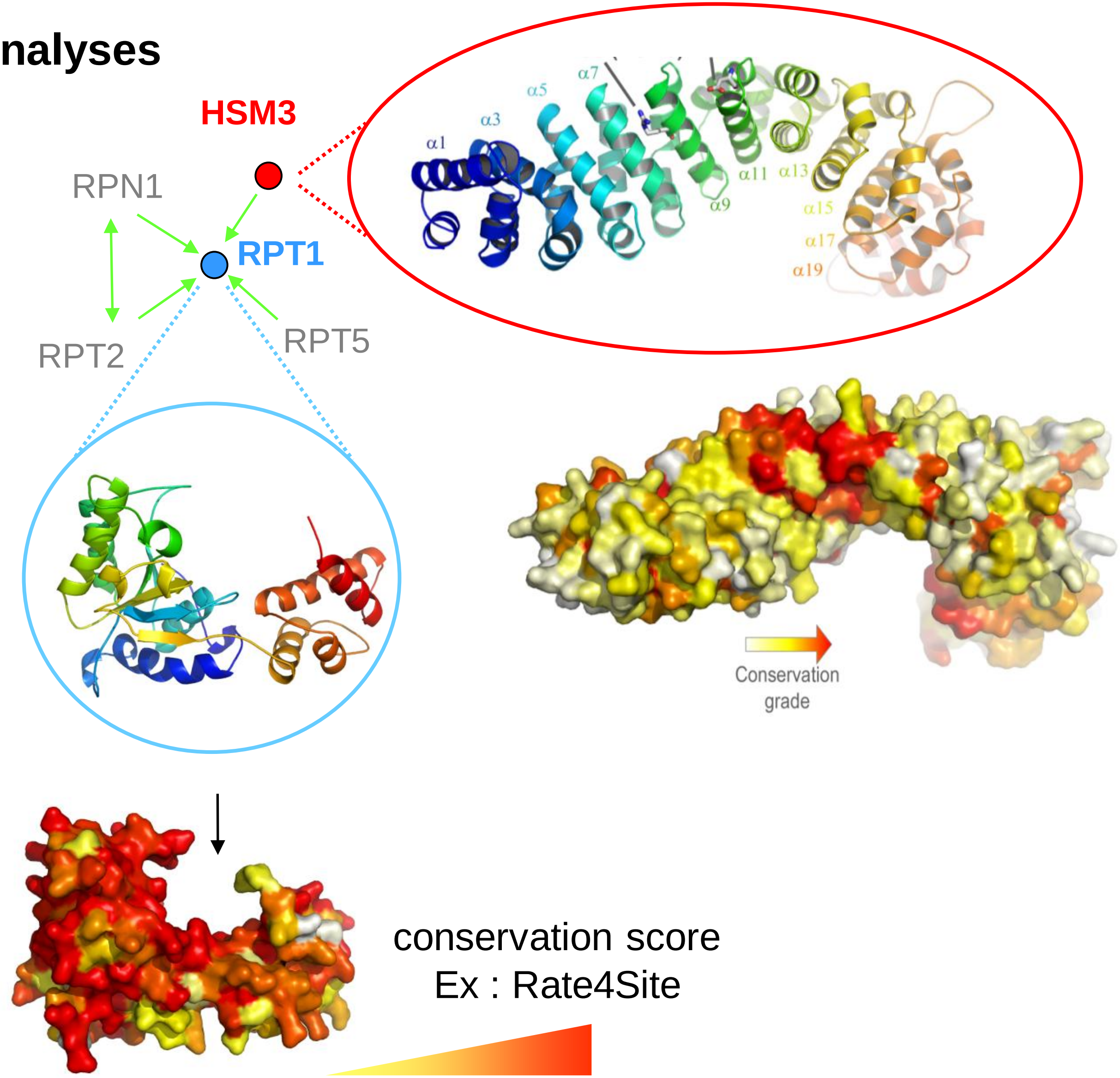




- + basic
- - acidic
- P polaire
- H hydrophobe

# Which evolutionary signals at protein surfaces can be captured to identify the interaction sites ?

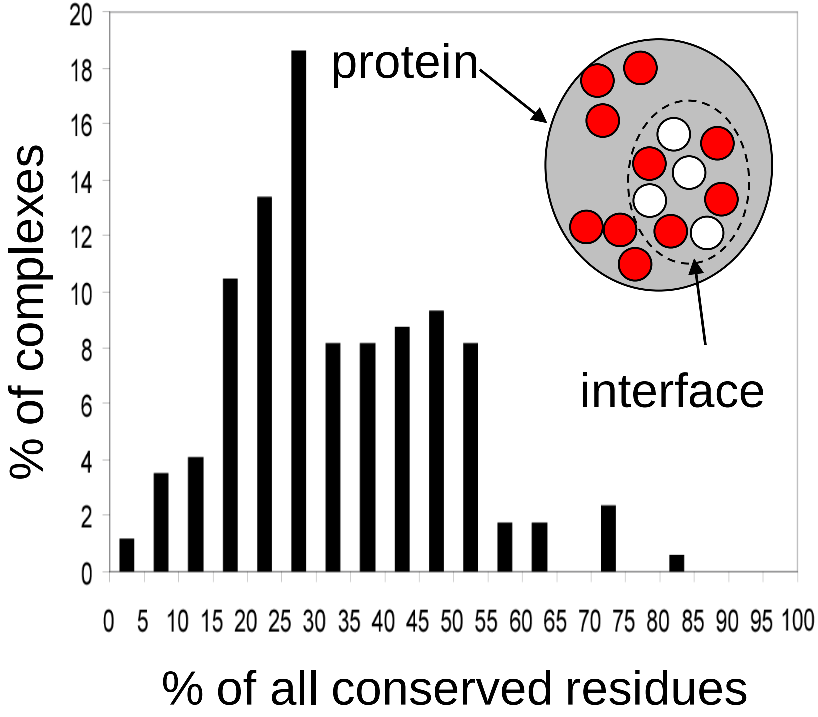
## Conservation analyses



conservation score
Ex : Rate4Site

# Conservation analyses at the interface of intra-molecular domain-domain interactions

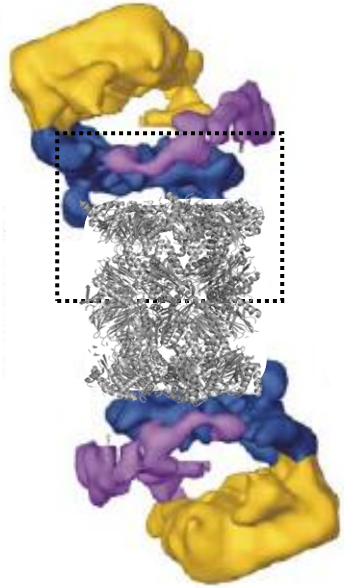
Which ratio of conserved residues
are part of the interface ?

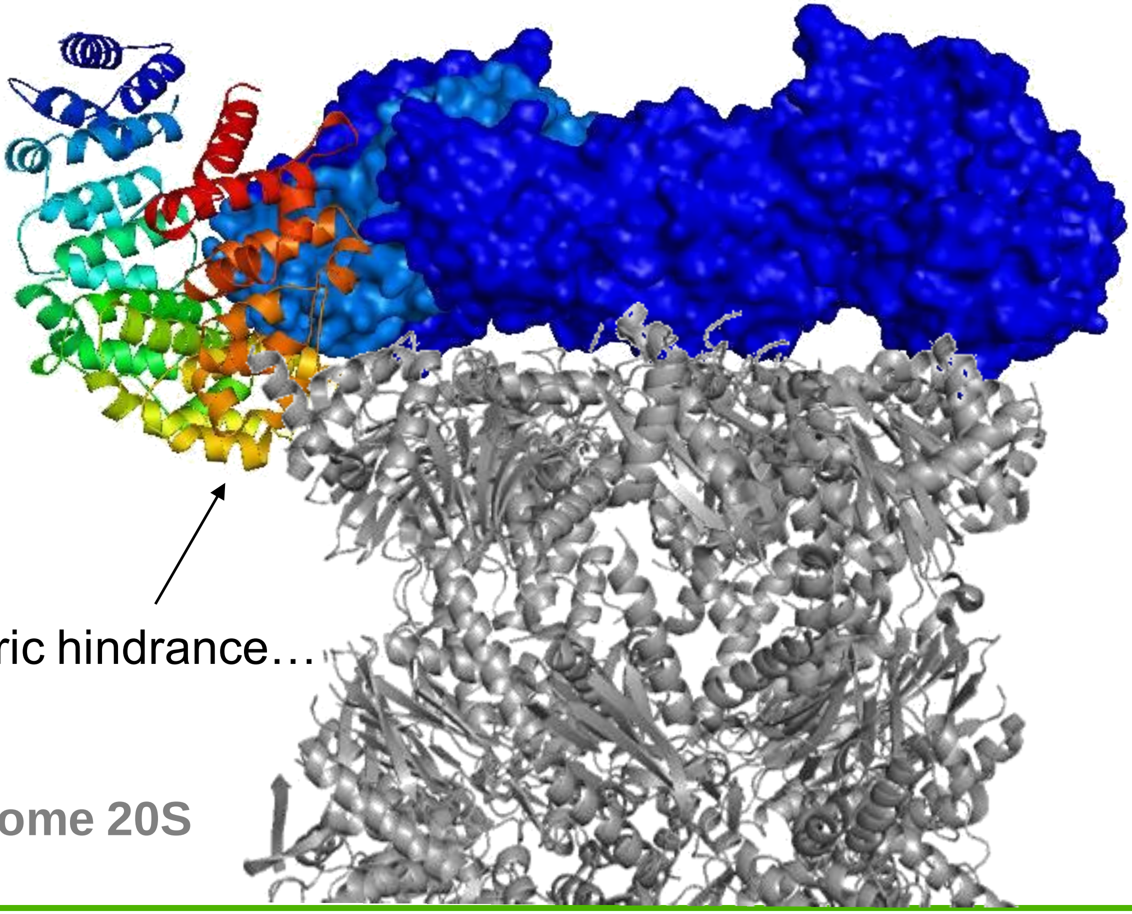


Several approaches combined conservation with other structure and sequence features
to identify potential binding patches ➔ no mutual information

*(ProMate (Neuvirth, JMB, 2004), PINUP (Liang et al, NAR, 2006), SPPIDER (Porollo, Proteins, 2007)*

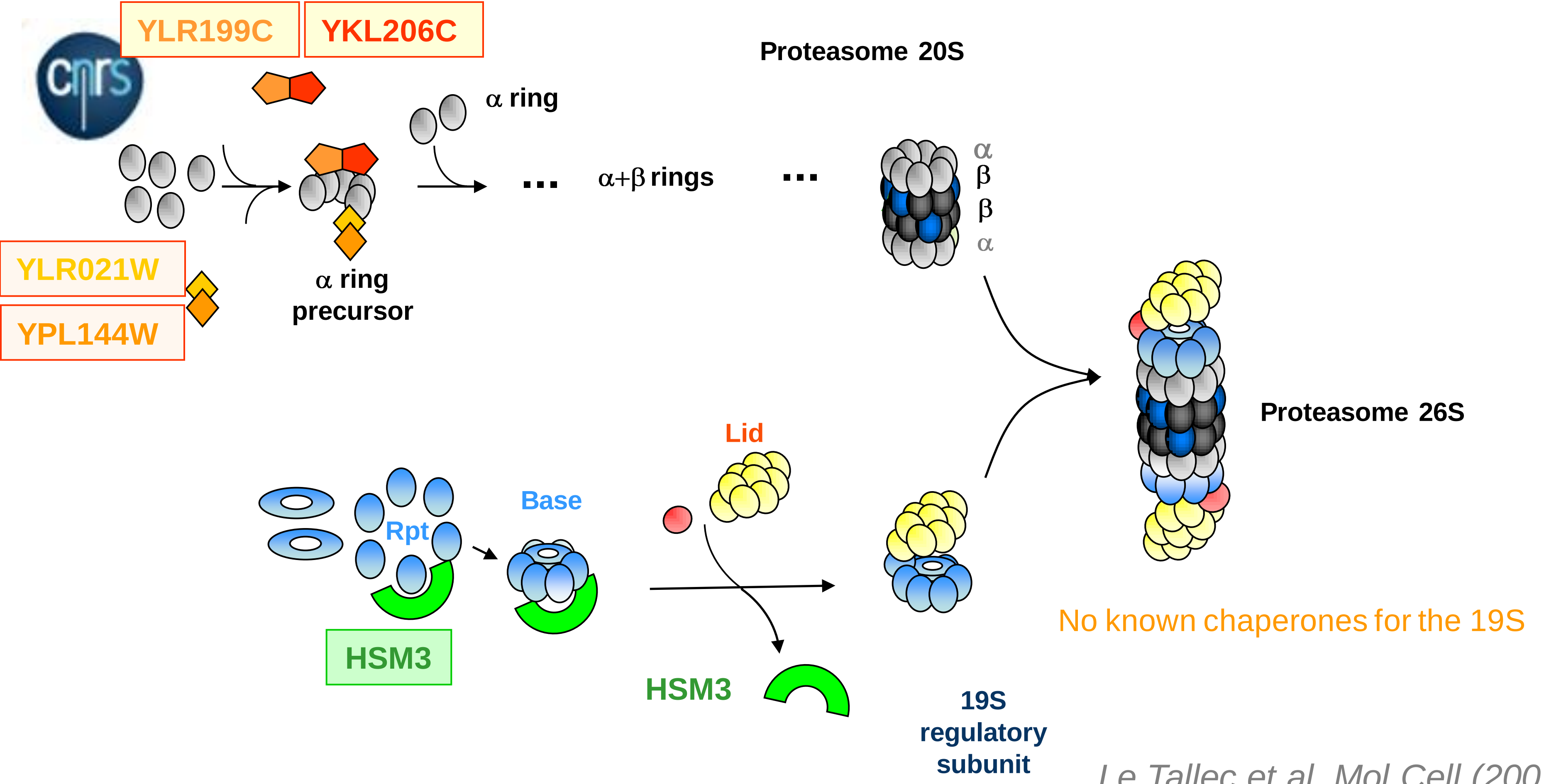
# Speculative model for the release of Hsm3 with the AAA ATPase ring and the 20S





# 5 novel chaperones identified in yeast
# ➔ Required for proper proteasome assembly

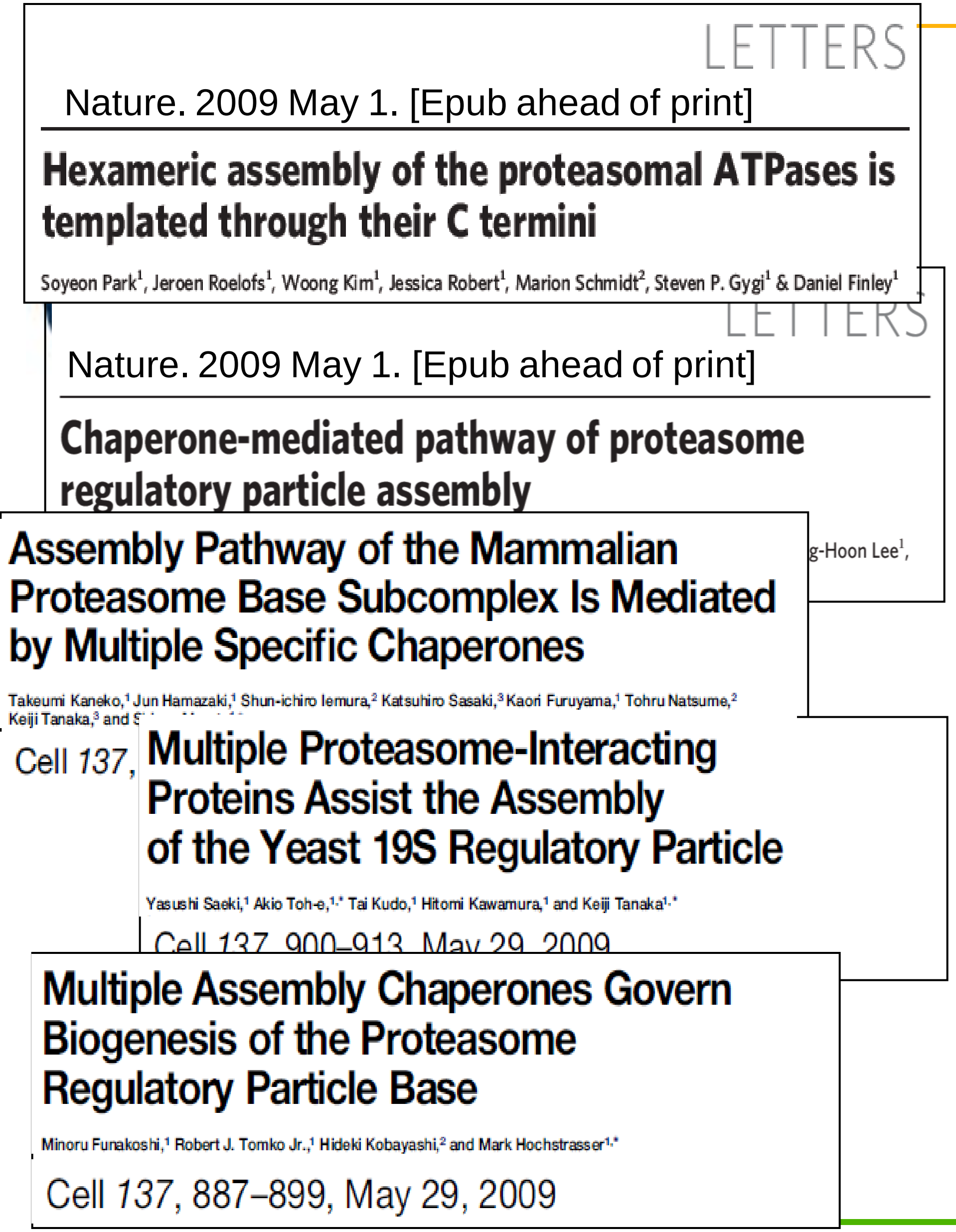
Coll. A. Peyroche’s group (iBiTecS/CEA Saclay)
organism : *S. cerevisiae*



No known chaperones for the 19S

*Le Tallec et al, Mol Cell (2007)*
*Le Tallec et al, Mol Cell (2009)*

# Numerous partners are still to be uncovered …



LETTERS

Nature. 2009 May 1. [Epub ahead of print]

**Hexameric assembly of the proteasomal ATPases is templated through their C termini**

Soyeon Park[1], Jeroen Roelofs[1], Woong Kim[1], Jessica Robert[1], Marion Schmidt[2], Steven P. Gygi[1] & Daniel Finley[1]

---

LETTERS

Nature. 2009 May 1. [Epub ahead of print]

**Chaperone-mediated pathway of proteasome regulatory particle assembly**

g-Hoon Lee[1],

---

**Assembly Pathway of the Mammalian Proteasome Base Subcomplex Is Mediated by Multiple Specific Chaperones**

Takeumi Kaneko,[1] Jun Hamazaki,[1] Shun-ichiro Iemura,[2] Katsuhiro Sasaki,[3] Kaori Furuyama,[1] Tohru Natsume,[2] Keiji Tanaka,[3] and

Cell 137,

---

**Multiple Proteasome-Interacting Proteins Assist the Assembly of the Yeast 19S Regulatory Particle**

Yasushi Saeki,[1] Akio Toh-e,[1,*] Tai Kudo,[1] Hitomi Kawamura,[1] and Keiji Tanaka[1,*]

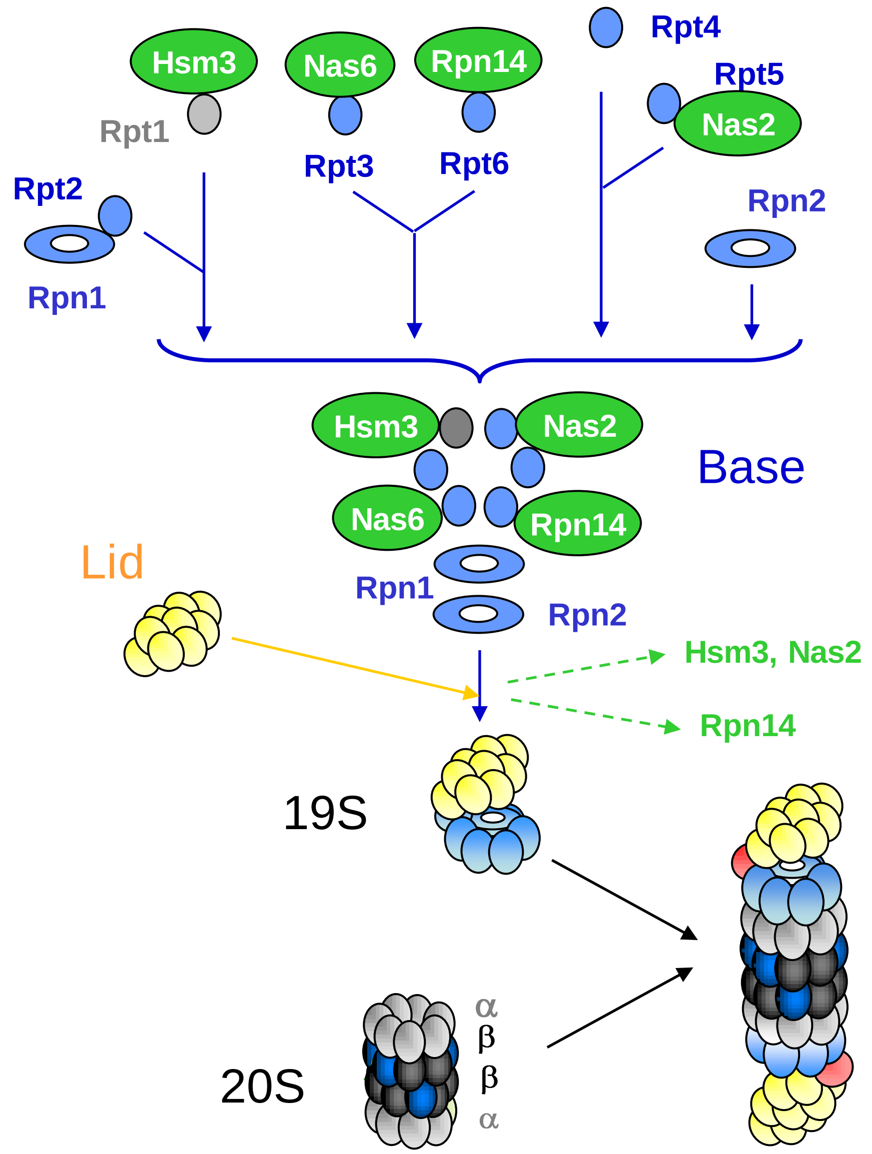
Cell 137, 900–913, May 29, 2009

---

**Multiple Assembly Chaperones Govern Biogenesis of the Proteasome Regulatory Particle Base**

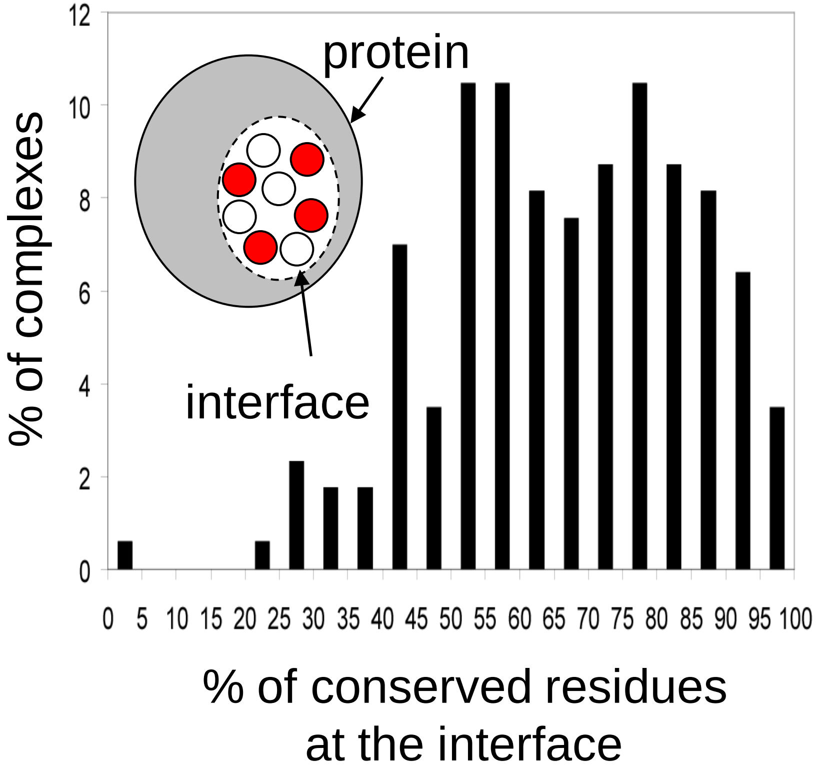
Minoru Funakoshi,[1] Robert J. Tomko Jr.,[1] Hideki Kobayashi,[2] and Mark Hochstrasser[1,*]

Cell 137, 887–899, May 29, 2009

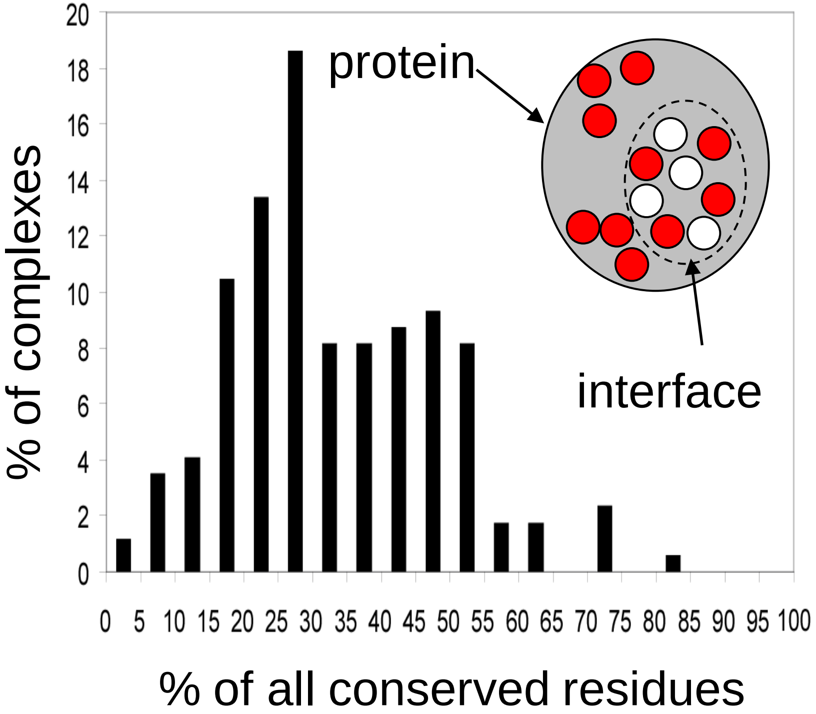


# Conservation analyses at the interface of intra-molecular domain-domain interactions

Which ratio of the interface correspond to conserved residues ?



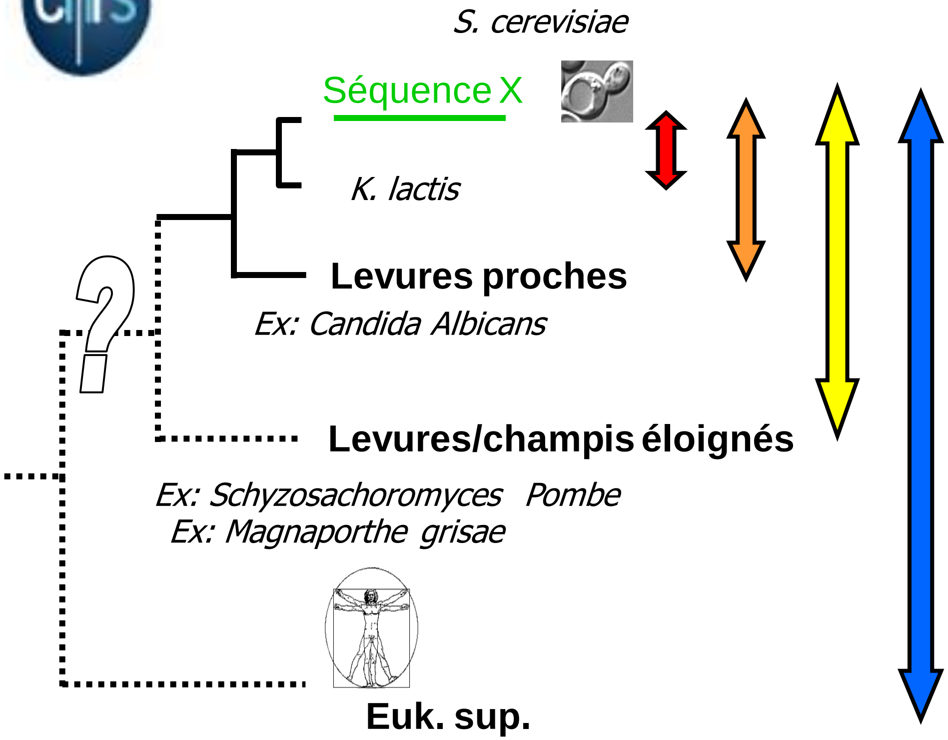
Which ratio of conserved residues are part of the interface ?



Several approaches combined conservation with other structure and sequence features to identify potential binding patches ➔ no mutual information

*(ProMate (Neuvirth, JMB, 2004), PINUP (Liang et al, NAR, 2006), SPPIDER (Porollo, Proteins, 2007)*

# Jusqu'à quel organisme retrouve-t-on des homologues pour les gènes de *S. cerevisiae* ?



*S. cerevisiae*

Séquence X

*K. lactis*

**Levures proches**

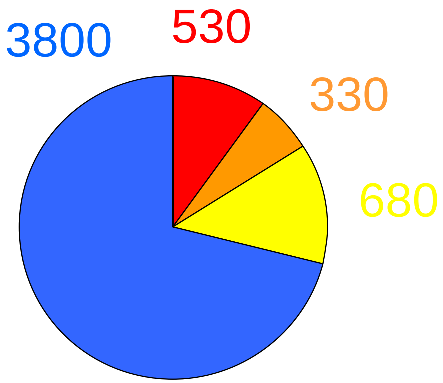
*Ex: Candida Albicans*

**Levures/champis éloignés**

*Ex: Schyzosachoromyces Pombe*
*Ex: Magnaporthe grisae*

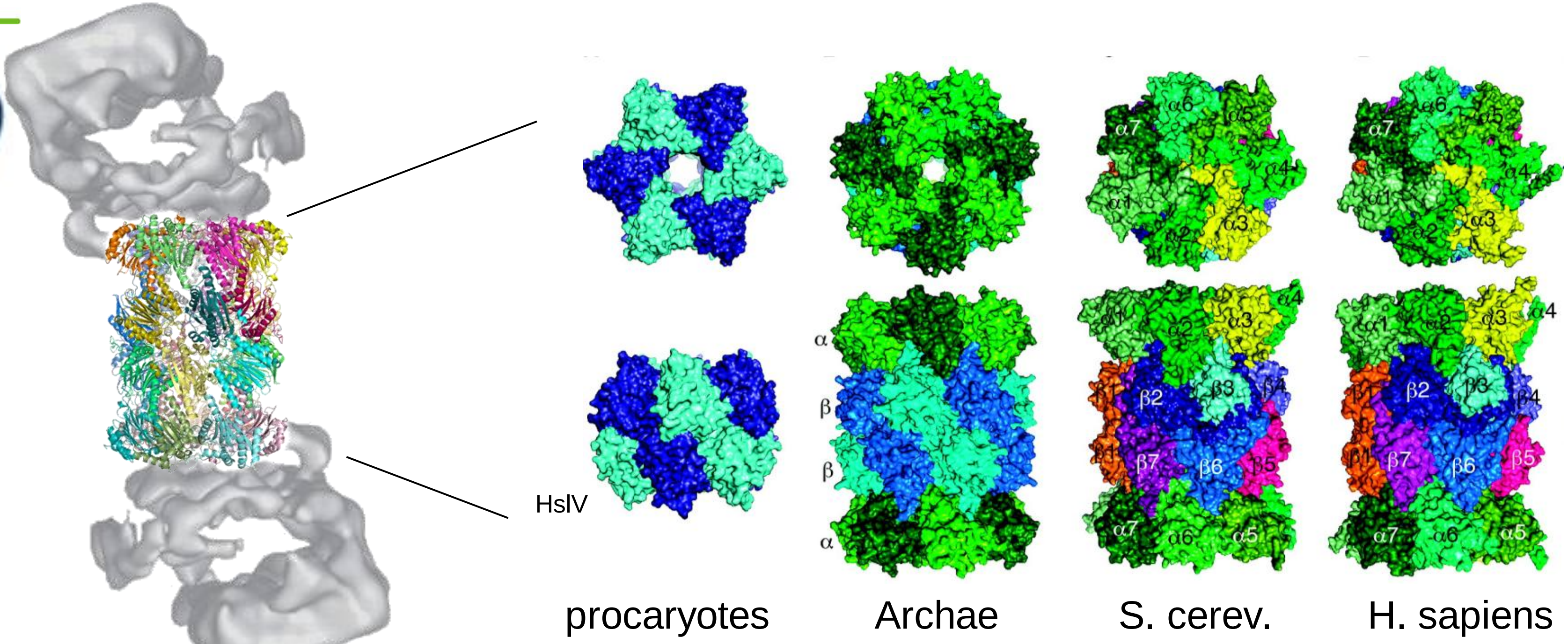
**Euk. sup.**

Répartition des gènes de
*S. cerevisiae* dans les
différentes catégories

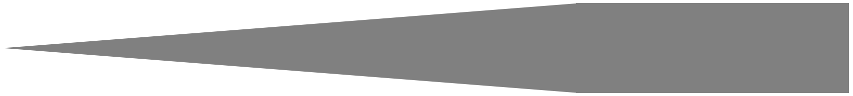


Passage par la prédiction de structure

# Complexité de l’organisation du protéasome en multiples sous-unités



~ 35 sous-unités différentes



Nombre de sous-unités différentes dans le 20S