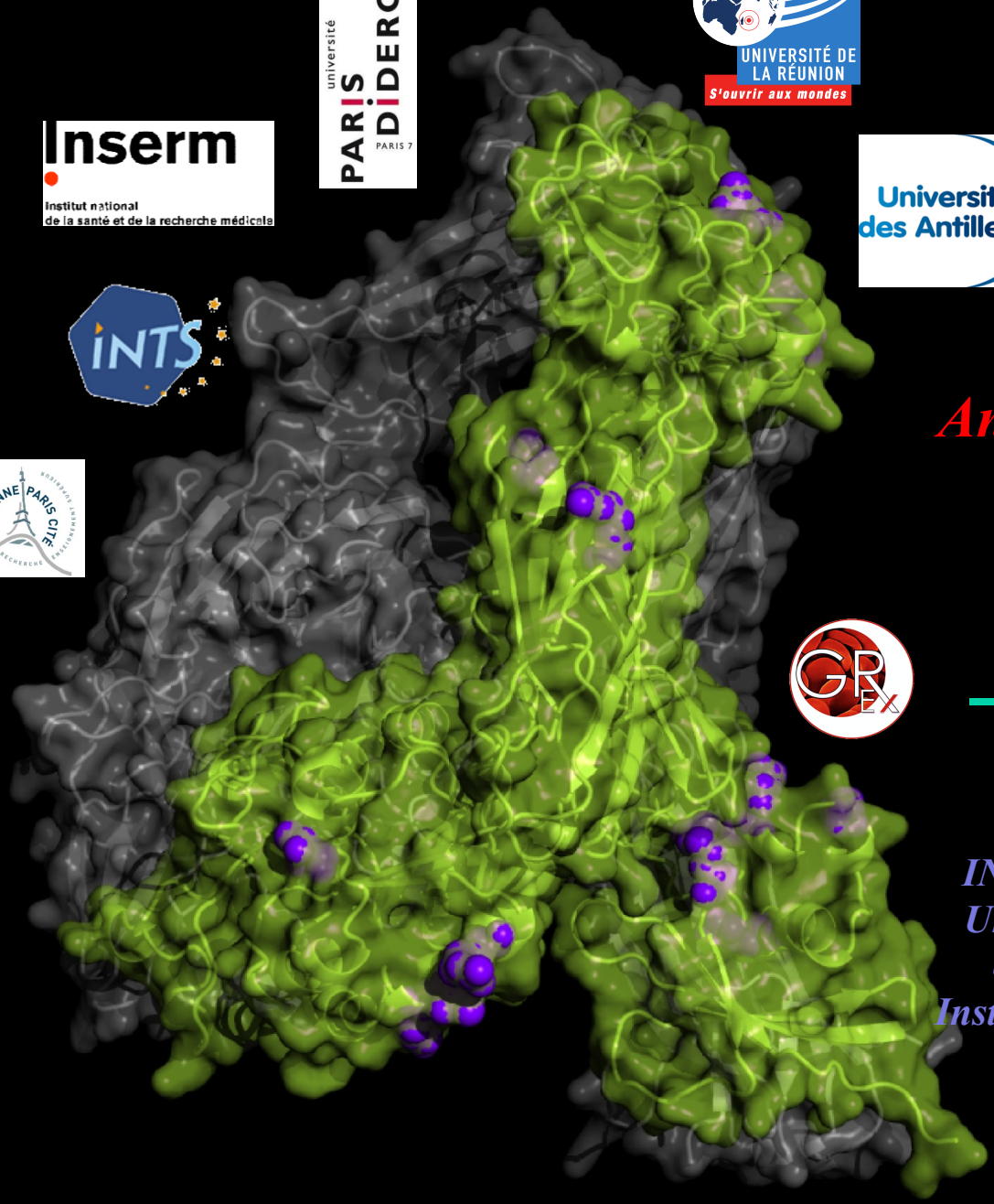
BELBI 2018



BelBi 2018

Analysis of allosteric effect of pathologic variants at the light of local protein conformations

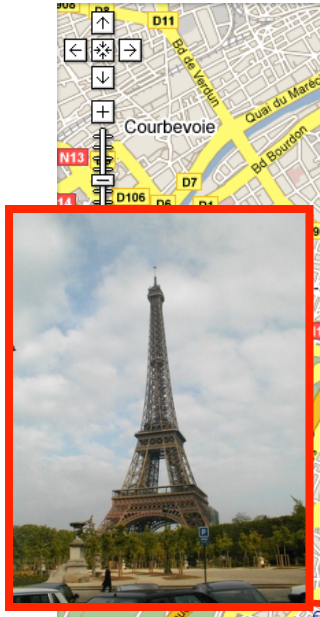
Alexandre G. de Brevern

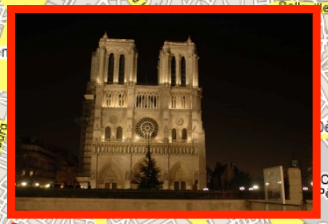
INSERM UMR_S 1134, team 2 – DSIMB, Univ. Paris Diderot, Sorbonne Paris Cité, Univ de la Réunion, Univ des Antilles, Institut National de la Transfusion Sanguine (INTS), GR-Ex

Inserm
Institut national
de la santé et de la recherche médicale

DSIMB: a bioinformatics team

DSIMB
Bioinformatique & Interactions

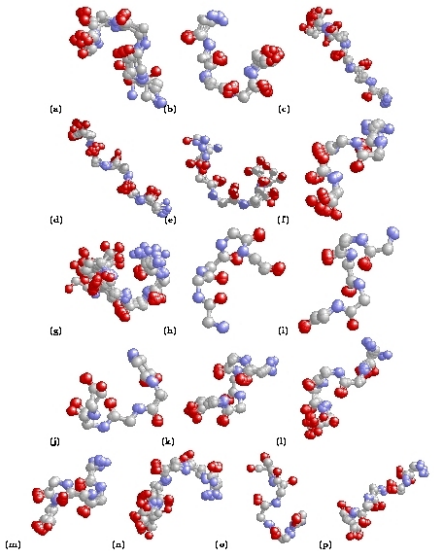




Inserm
Institut national
de la santé et de la recherche médicale

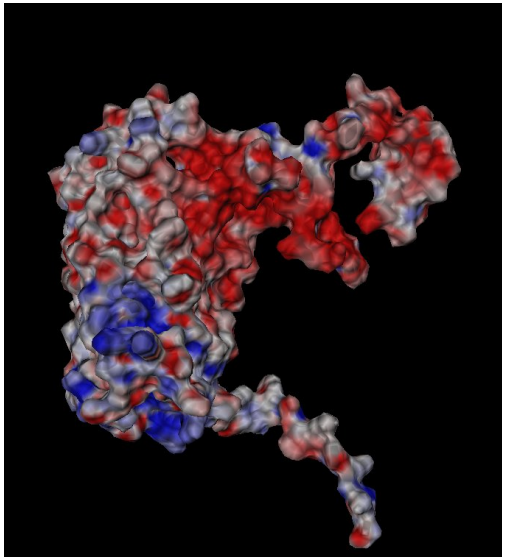
DSIMB: a bioinformatics team

DSIMB
Bioinformatique & Interactions



Our main goal:
Sequence - Structure
Function & Dynamics

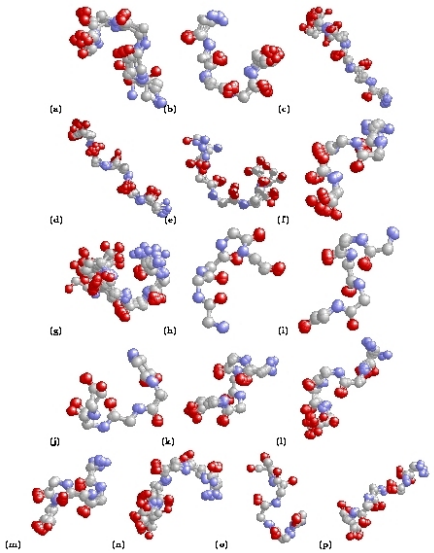
Methodological developments and applications



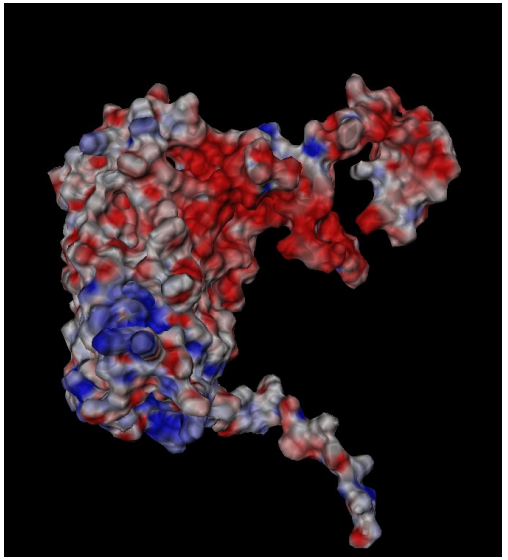
Inserm
Institut national
de la santé et de la recherche médicale

DSIMB: a bioinformatics team

DSIMB
Bioinformatique & Interactions



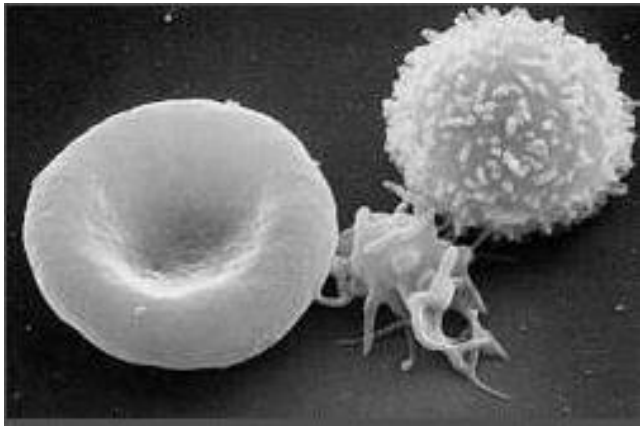
Our main goal:
Sequence - Structure
Function & Dynamics



Methodological developments and applications

Red Blood Cells

and Platelets



Inserm
Institut national
de la santé et de la recherche médicale

A part of the journey

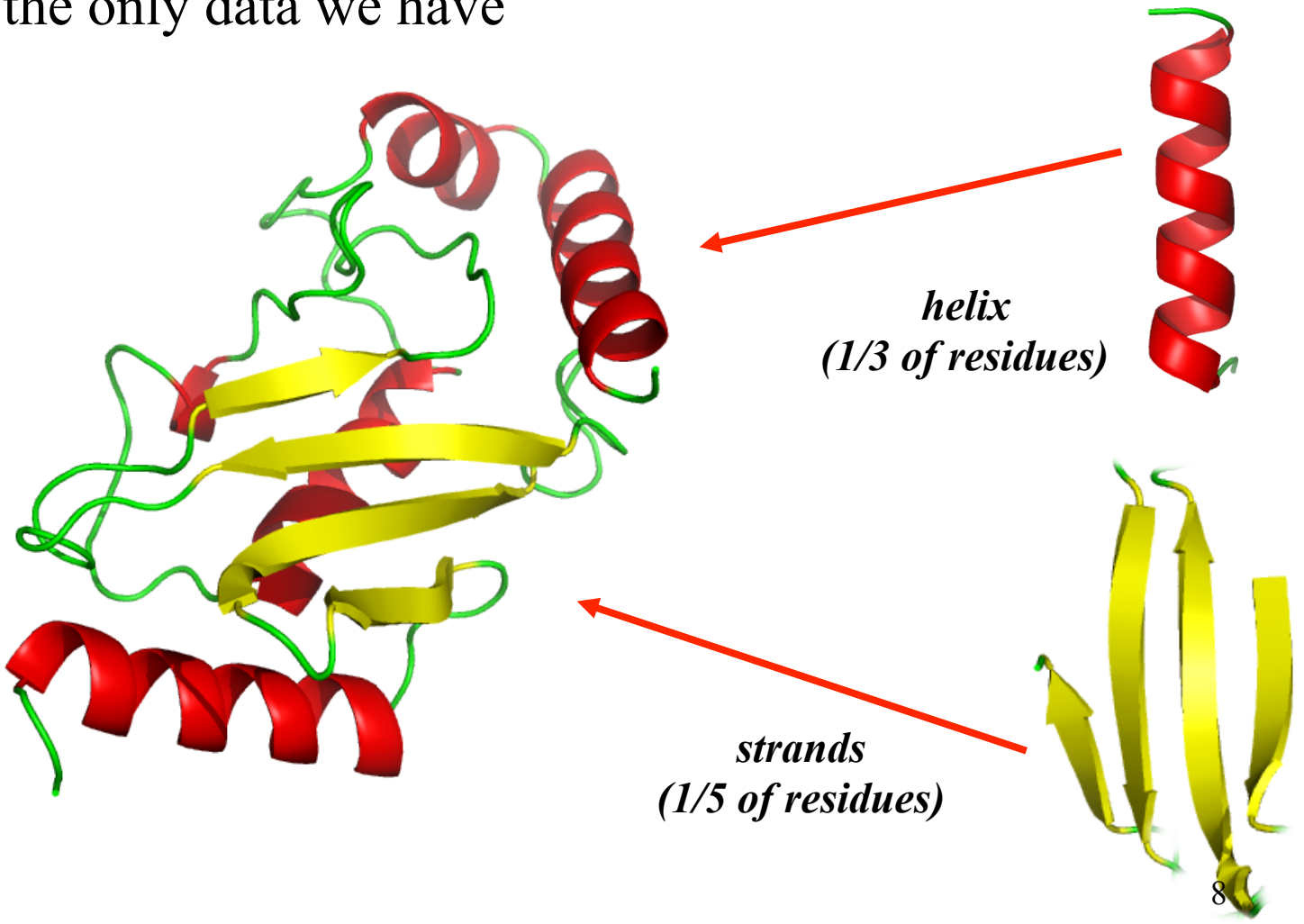
SIMB
Bioinformatique & Interactions

- Why going beyond classical secondary structure ?
- How ?
- How is it useful for analysis of protein dynamics ?
- Can it be used also for protein disorder ?

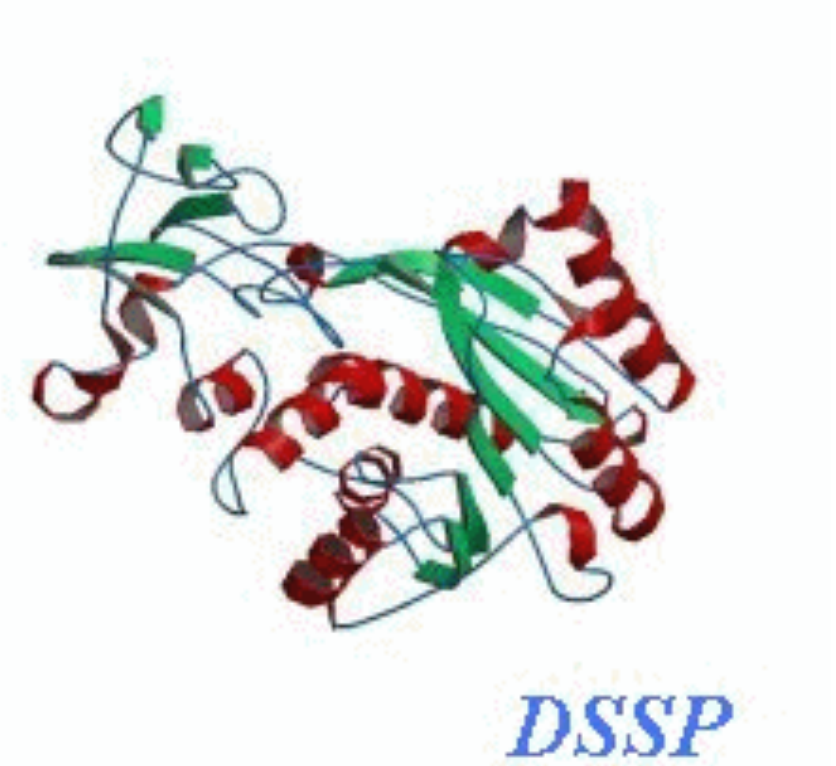
SECONDARY STRUCTURE

The secondary structures

- That's the only data we have



Secondary structures: some questions

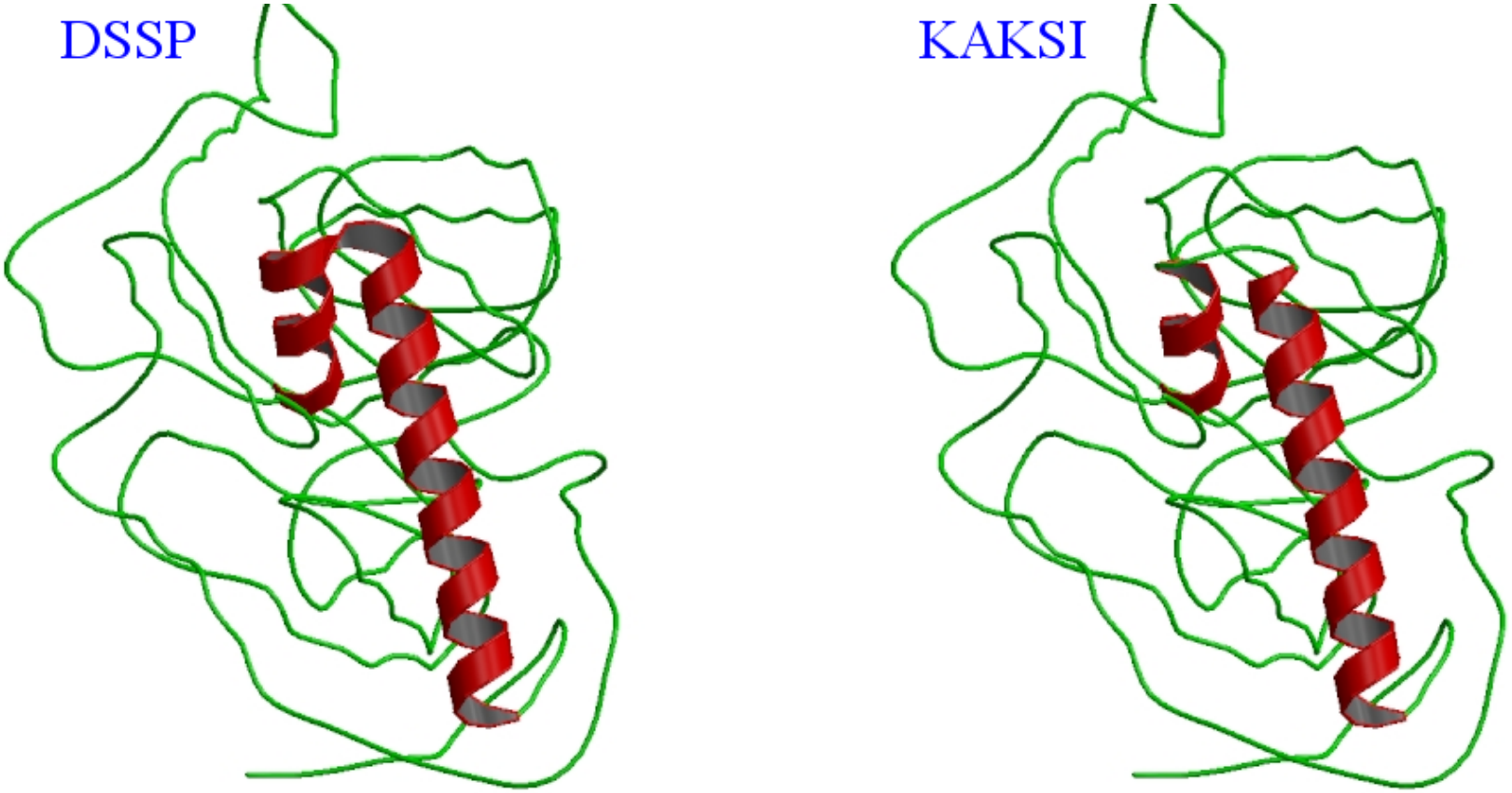


No simple concensus between methods

Assignment example for the *Hhai Méthyltransferase* protein (PDB code :10MH) by DSSP, STRIDE, PSEA, DEFINE, PCURVE, XTLSSTR and SECSTR.

L. Fourrier, C. Benros & A.G. de Brevern (*BMC Bioinfo*, 2004)

Secondary structures: some questions

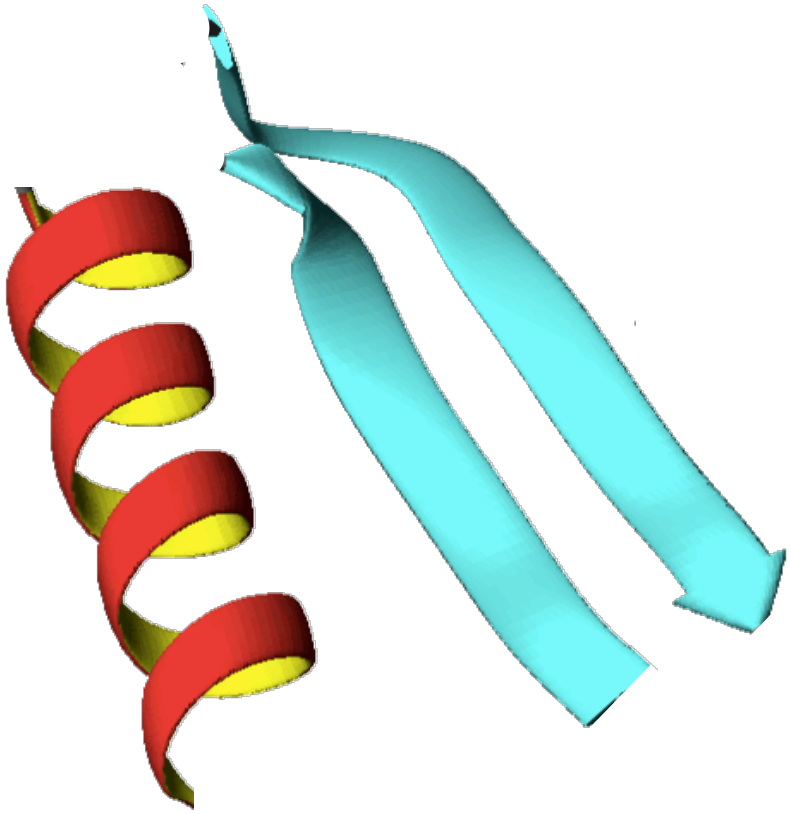


J. Martin, G. Letellier, J.F. Taly, A. Martin, A.G. de Brevern & J.F. Gibrat (*BMC Structural Biology,* 2005)

http://migale.jouy.inra.fr/mig/mig_fr/servlog/kaksi/

Secondary structures: some questions

Moreover .



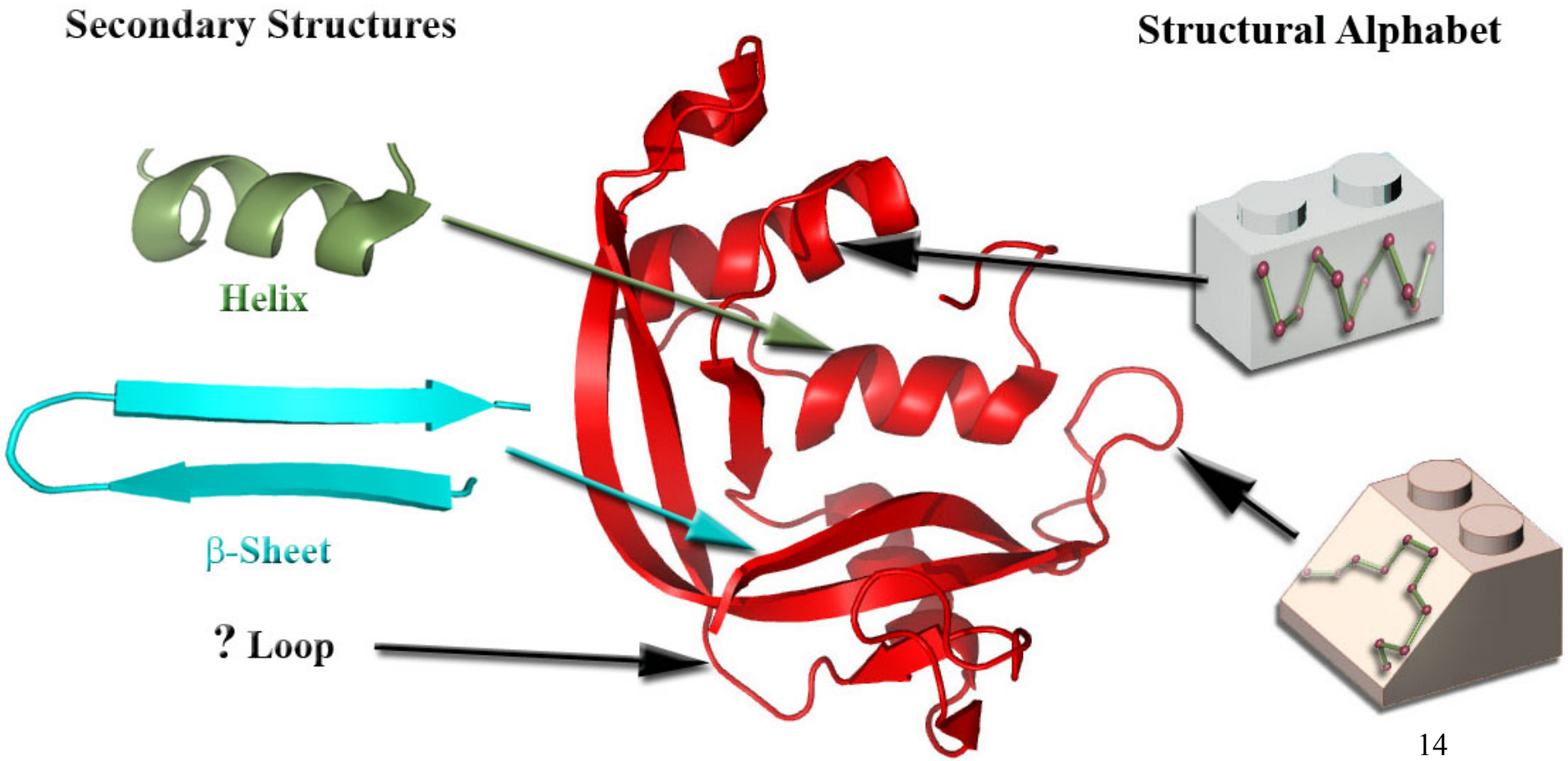
Secondary structures: some questions

- It is possible to look at protein structures in a different ways
- At a local level.

STRUCTURAL ALPHABET

Structural alphabet

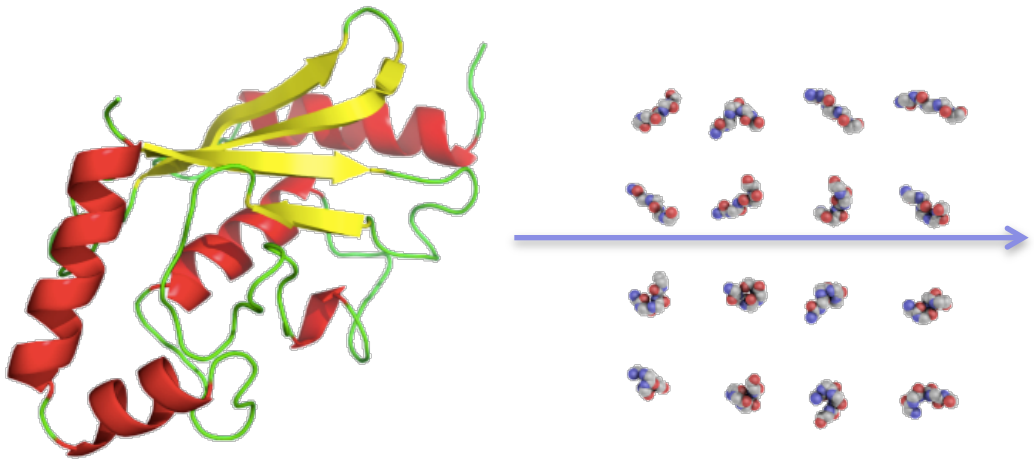
- Another vision



Structural alphabet

- Structural Alphabet (SA):

  Structure translation (3D -> 1D)



```
>2aak
ZZfklmmmmmmmmmmmmcccehi
acdddfkbmkbccddddddehia
kgopacddddddfkbckbccdd
ddfbdcddfklcfkopacdfklm
mmgcfkbcfklmmmmmmmmmmmc
cfklcfblmlmmmmmmmmmlmmm
mmmmmmmmmmZZ
```

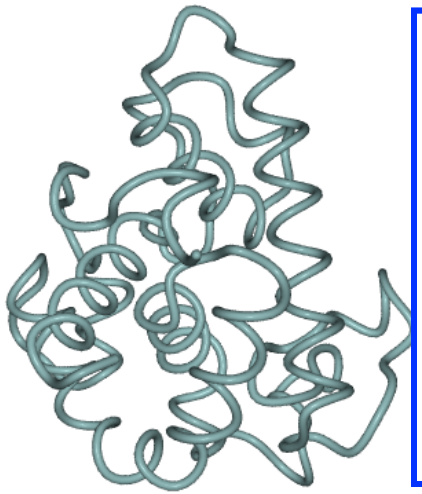
3D protein structure ***1D Sequence***

de Brevern A.G., Etchebest C. & Hazout, S. (*Proteins*, 2000).

Inserm
Institut national
de la santé et de la recherche médicale

Structural alphabet

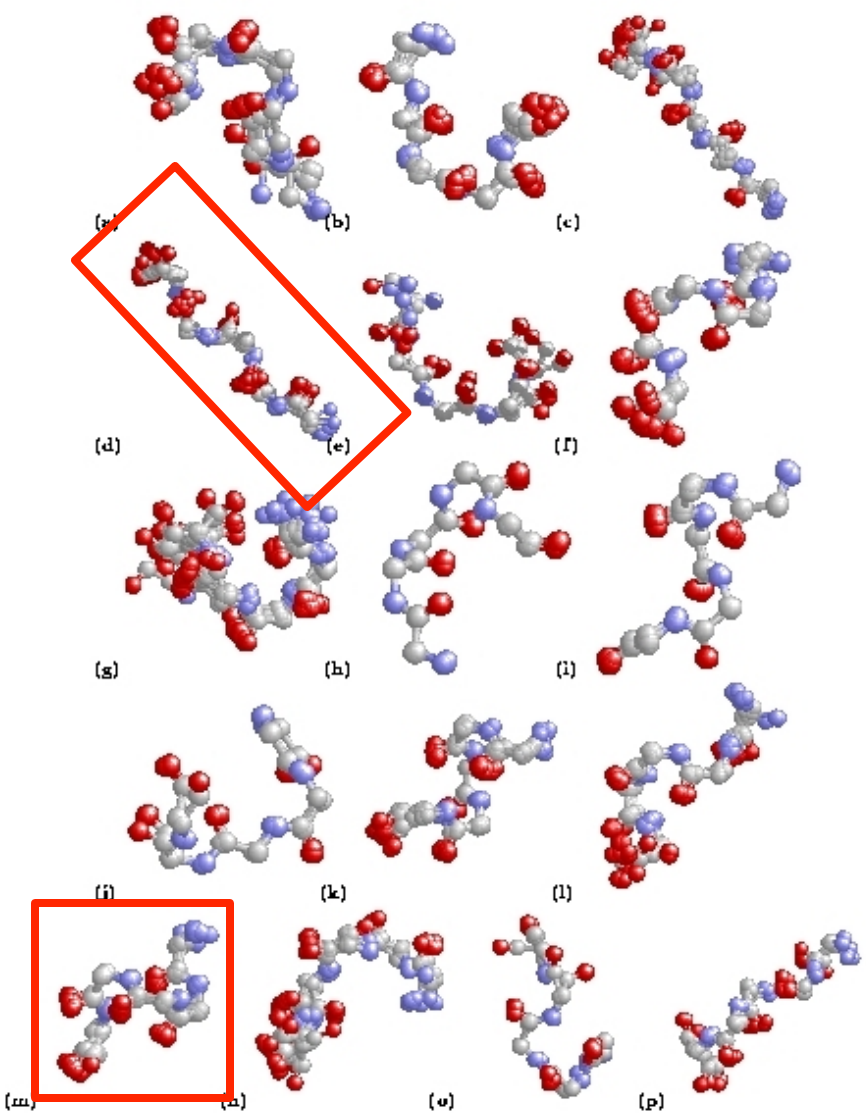
SIMB
Bioinformatique & Interactions



```
>153L [secondary structures]
HHHH                    HHHHHH         HHHHHHH
HHHHHHHHHHHHHHHH          HHHHHHHHH
                                 HHHHHHHHHHHH
HHHHHHHHH      HHHHHHHHHHH     EEEEE
   HHHHHHHHHHHHHH
```

```
>153L [structural alphabet]
ZZmnopfklpccebjafklmmmnopabecjklmmmmmmmmm
mmmmmmmmmmmmmmmmmnopafklmmmmmmmmnooolapgeh
iafkopagcjkopafklmccehjfklmklmmmmmmmmmmm
mmmmmmmmmbcfklmmmmmmmmmnomklmnbfklmmgoia
hilmmmmmmmmmmmmmnoZZ
```

Structural alphabet



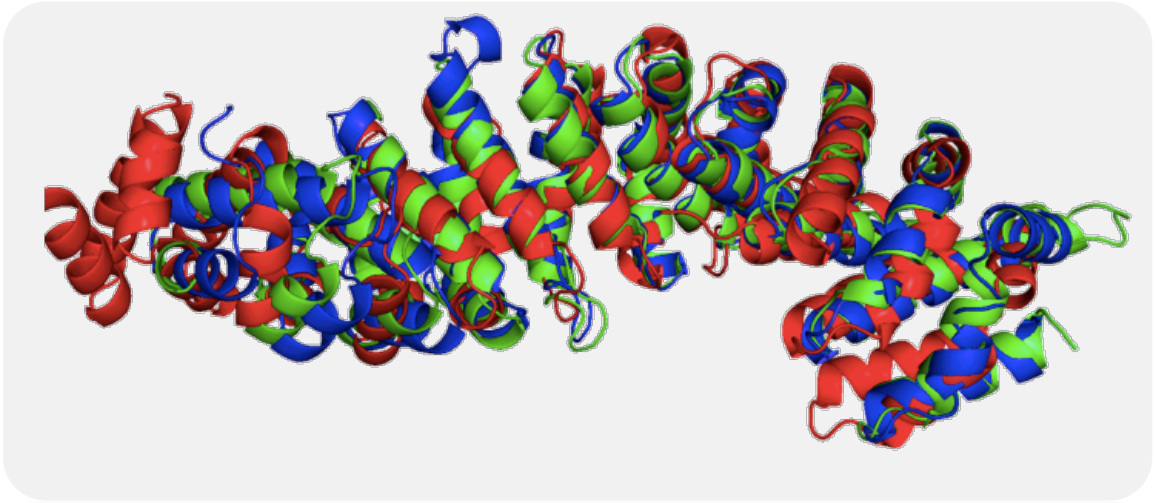
Pr. Serge Hazout

de Brevern A.G., Etchebest C. & Hazout, S. (*Proteins*, 2000).

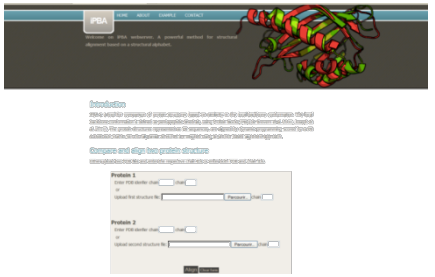
16 Protein Blocks

Structural alphabet

- Superimposition of 3D protein structures



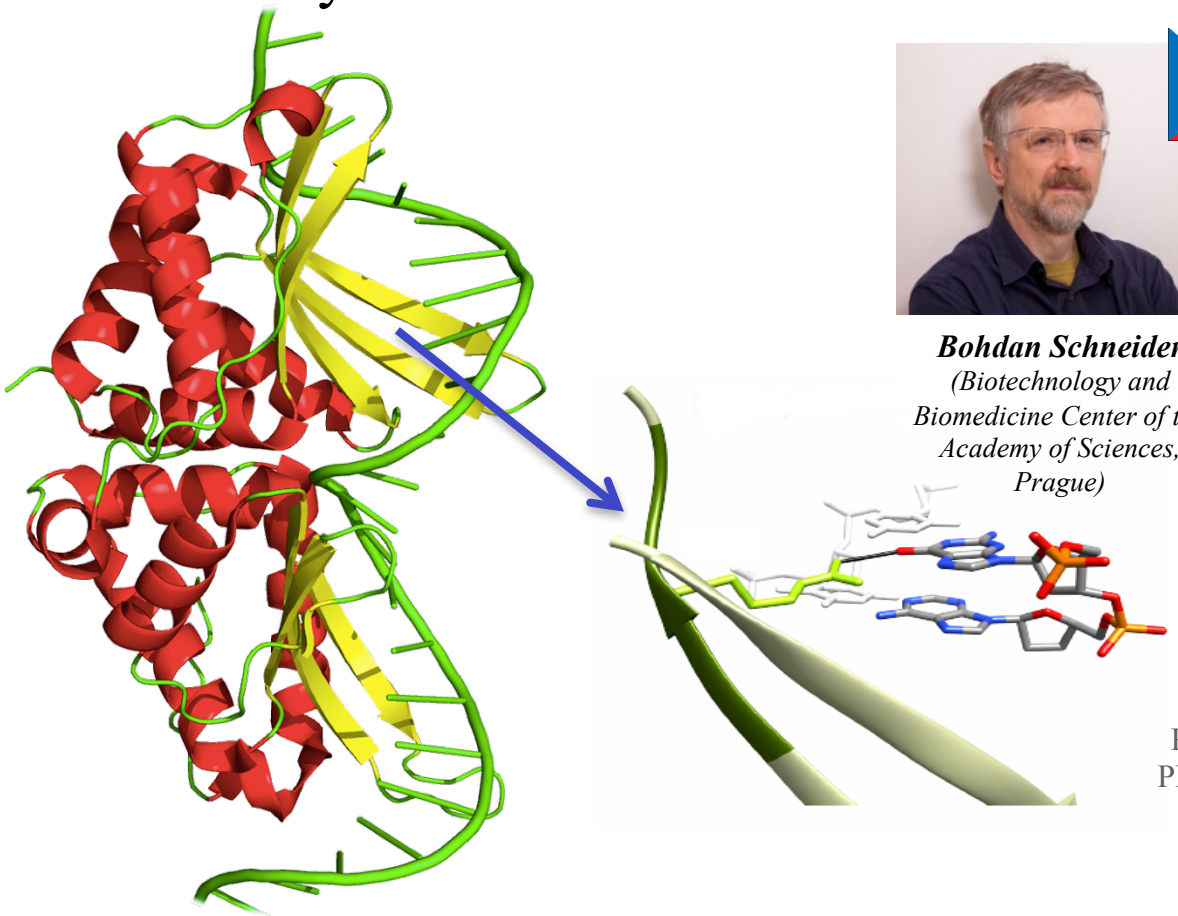
Narayaswami Srinivasan
(IISc Bangalore)



Joseph A.P., Srinivasan N. & de Brevern A.G. (2011) *Biochimie*, **93**(9):1434-45.
Agarwal G., Mahajan S., Srinivasan N. & de Brevern A.G. (2012) *PLoS ONE*, **6**(3):e17826.
Gelly J.-C., Joseph A.P., Srinivasan N. & de Brevern A.G. (2011), *Nucleic Acid Res*, **39**:W18-23.
Joseph A.P., Srinivasan N. & de Brevern A.G. (2012) *Biochimie*, **94**:2025-34.
Léonard S., Joseph A.P., Srinivasan N., Gelly J.-C. & de Brevern A.G. (2014) *J Biol Struct Dyn*, **32**(4):661-8.

Structural alphabet

- Protein interface analyses



Bohdan Schneider
(Biotechnology and Biomedicine Center of the Academy of Sciences, Prague)

PB *d* + Nt 19
PDB code3eh8

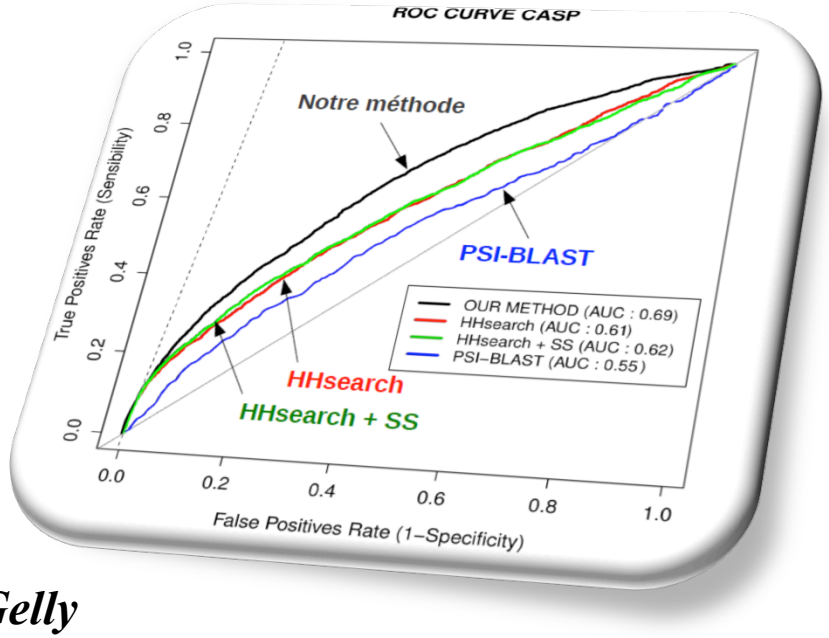
Schneider B., Cerný J., Svozil D., Cech P., Gelly J.-C. & de Brevern A.G. (2014) *Nucl Acid Res*, **42**(5):3381-94.
Schneider B., Gelly J.-C., de Brevern A.G. & Cerný J. (2014) *Acta Crystallographica Section D*, **42**(5):3381-94.

Structural alphabet

```
>153L
TDCYGNVNRIDTTGASCKTAKPEGL
SYCGVSASKKIAERDLQAMDRYKTI
IKKVGEKLCVEPAVIAGIISRESHA
GKVLKNGWGDRGNGFGLMQVDKRSH
KPQGTWNGEVHITQGTTILINFIKT
IQKKFPSWTKDQQLKGGISAYNAGA
GNVRSYARMDIGTTHDDYANDVVAR
AQYYKQHGY
```

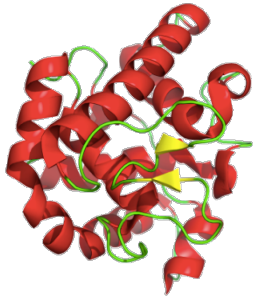
Prediction of PBs

```
>153L
ZZmnopfklpccebjafklmmmnopa
becjklmmmmmmmmmmmmmmmmmmmm
mmmmmnopafklmmmmmmmmnooola
pgehiafkopagcjkopafklmcceh
jfklmklmmmmmmmmmmmmmmmmmmm
mmbcfklmmmmmmmmmmnomklmnbf
klmmgoiahilmmmmmmmmmmmmmmn
oZZ
```



Structural analogy

Structural fold

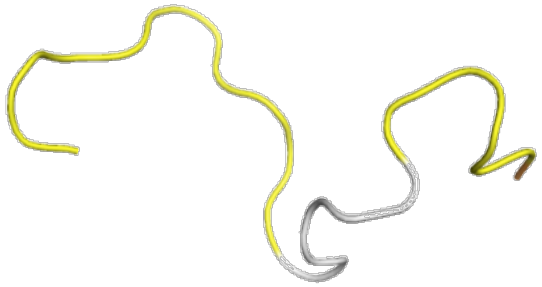


Dr. J.-C. Gelly

(*Bioinformatics*, 2015; *Sci Rep*, 2016)

Structural alphabet

- The most used structural alphabet at this day (Web Of Sciences: 156 citations, Google scholar: 288)
- More than 20 teams worldwide
- Used for local and global analysis, prediction (locale, specific, flexibility), le *protein design*, building of structural models, etc.



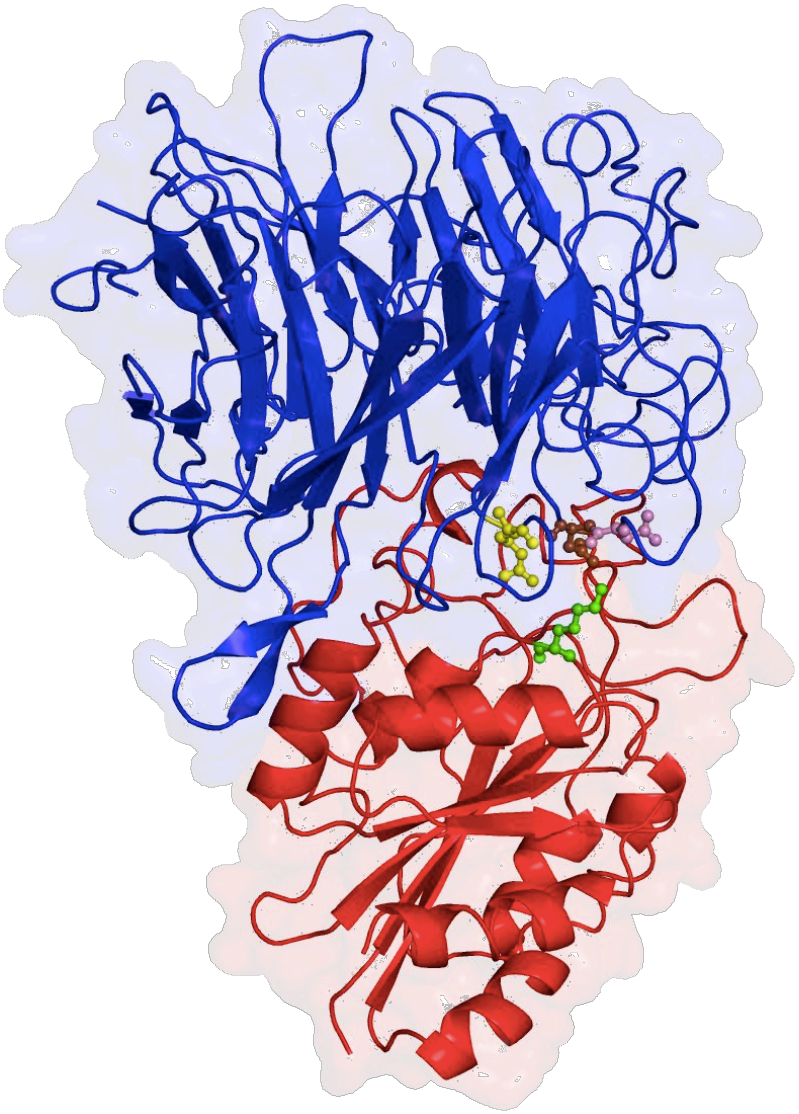
Analysis of molecular dynamics of protein and peptide

de Brevern A.G., Etchebest C. & Hazout S. (2000) *Proteins*, **41**(3), 271-287.

STRUCTURAL ALPHABET & FLEXIBILITY: MD

Integrins

Integrins



Dr. Vincent Jallu
Platelet departement

INTS

Integrins

One complex (Integrin):

- Glanzmann's thrombasthenia is an abnormality of the platelets, an extremely rare coagulopathy.

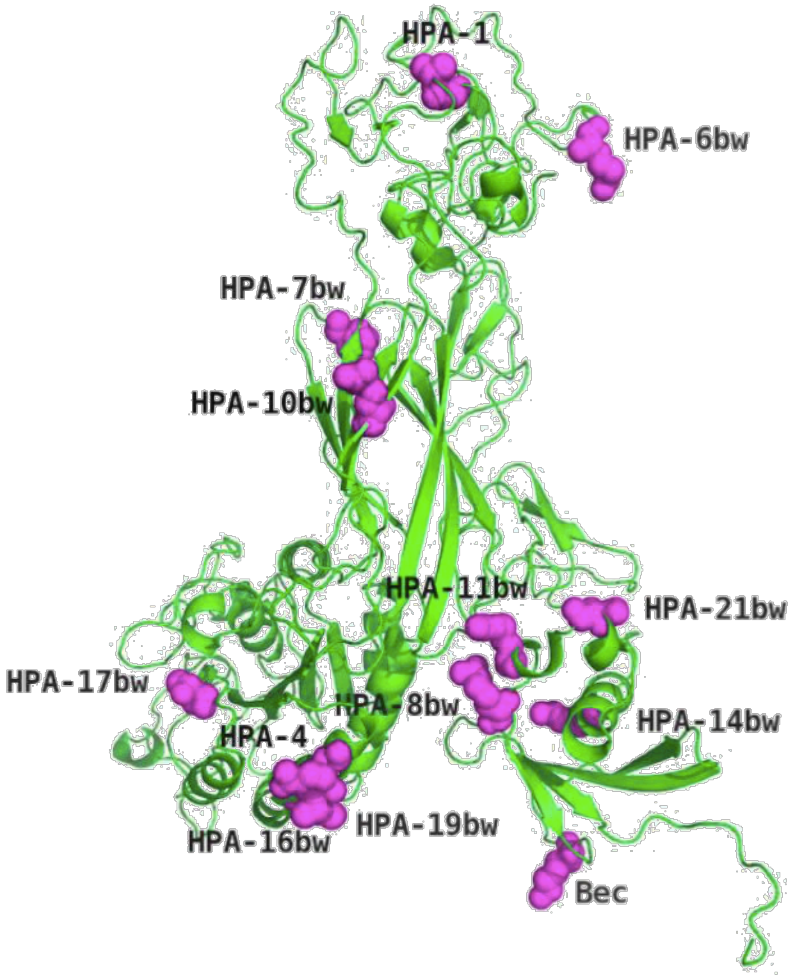
- Fetal / Neonatal alloimmune thrombocytopenia (FNAIT), a severe bleeding syndrome in which fetal / neonatal platelet destruction is mediated by maternal antibodies directed to specific antigens (or alloantigens) inherited from the father

Also the only blood group not on Red Blood Cells.

Inserm
Institut national
de la santé et de la recherche médicale

Integrins

SIMB
Bioinformatique & Interactions



Integrins

- This protein is *very* flexible

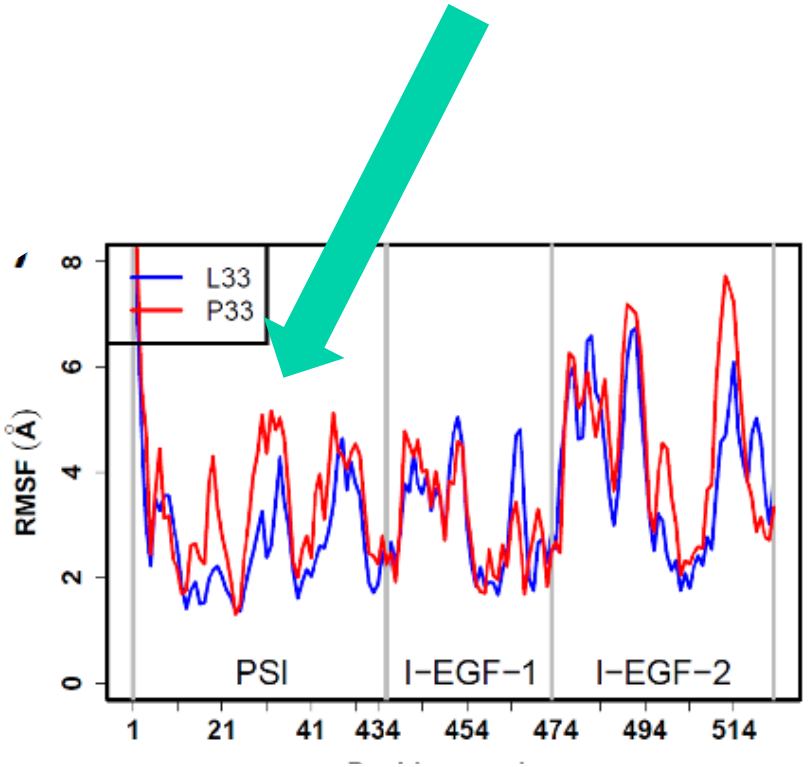
We compute Root Mean Square Fluctuation (RMSF)

$$RMSF_i = \sqrt{\frac{1}{nstep}\sum_{t=1}^{nstep} \|\vec{r}_i(t) - \langle\vec{r}_i\rangle\|^2}$$

Integrins

- This protein is *very* flexible

Residue 33 is associated to high RMSF



Inserm
Institut national
de la santé et de la recherche médicale

Integrins

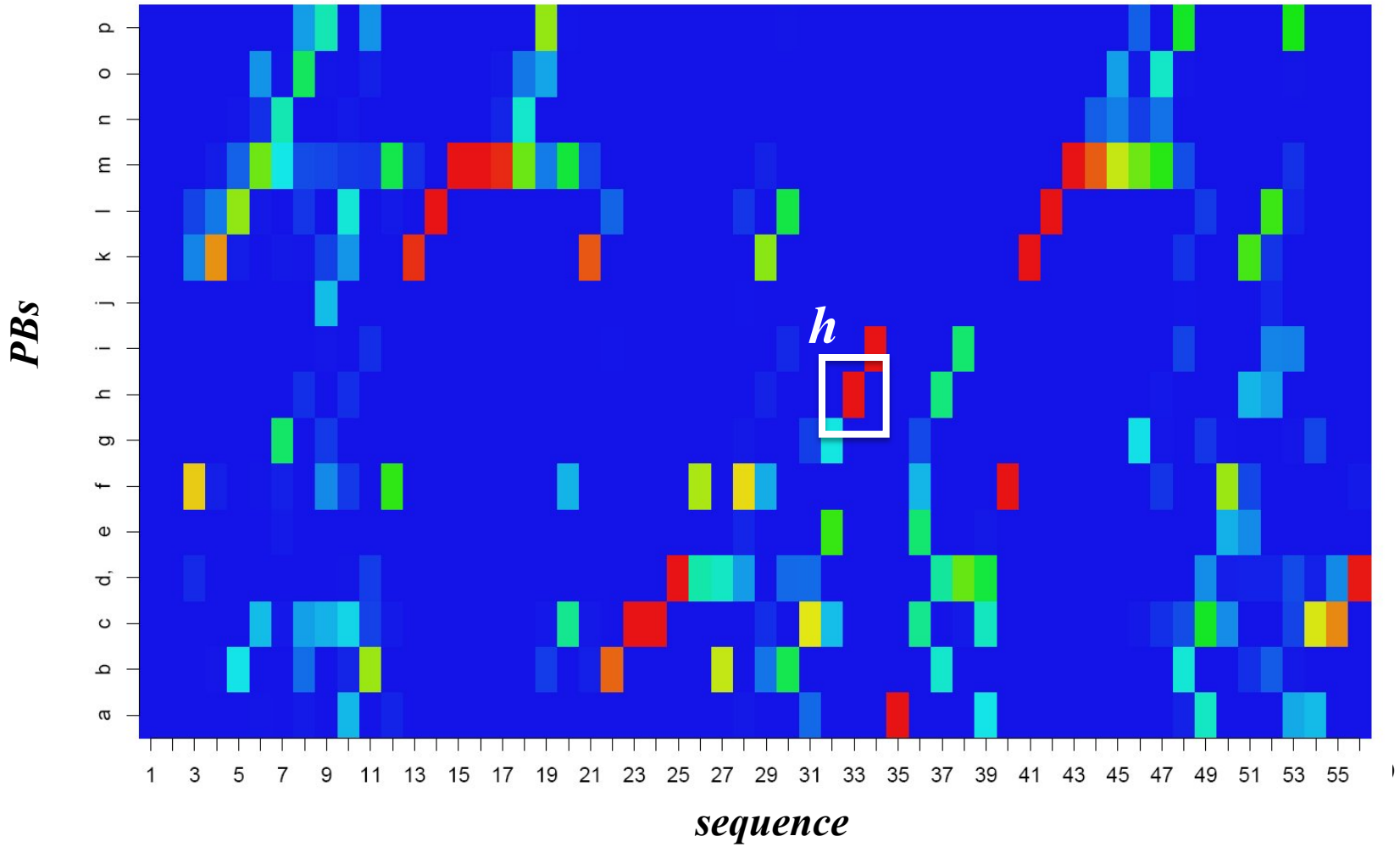
SIMB
Bioinformatique & Interactions

- How to use Protein Blocks here ?

```
ZZfklmnopabfklmmmnpfkbccdfblkbcchiaeh
iafklmmmnopafklpccddddddfbdcdddfblmb
ckbccdfbdcfbdcdddehiacddddddddddehiac
ddddddfbfklmmmmmmmmmgojlmmmmmmmnpcbd
cddddddddddfklpcfbdfklmnomclmmmbdcdd
debiadfbdcbfblklmmmmmmmmccdddfblbdcfk
lmmmmmmmommmnopafkbcddddddfbdcddehkl
mmnopabccehpacbdfkopafblmmpccddfklmmm
mmmmnopacdddbdfklmmmmmmmmmmghiacdddd
fklmmmmmmmmmmmmmmmncccddehiafkbccddd
ddfehjbccehiabdehiaghiacddddddddehjdd
fddddddddehiabdcddddddfbdfklmmpcfbg
cknojhiaccehhiacddehiacjkopafgehjlmng
jiafbfblkbcehiopaccehhiacdfbdchiaaedj
kopaccblbccfblpehhlpaccehhiacddehiacj
kopacfbfklmmgfkopghhlpaccehhiacdfbghi
edjkopacckbccfklmmmmmmmmmnopajklmmmk
lmmmomcdcdddfbdcdebiacdcdddddfkopacdd
dddddfkopacddddfbdcdddZZ
```

Integrins

- Then count for every snapshots, the frequency of PBs

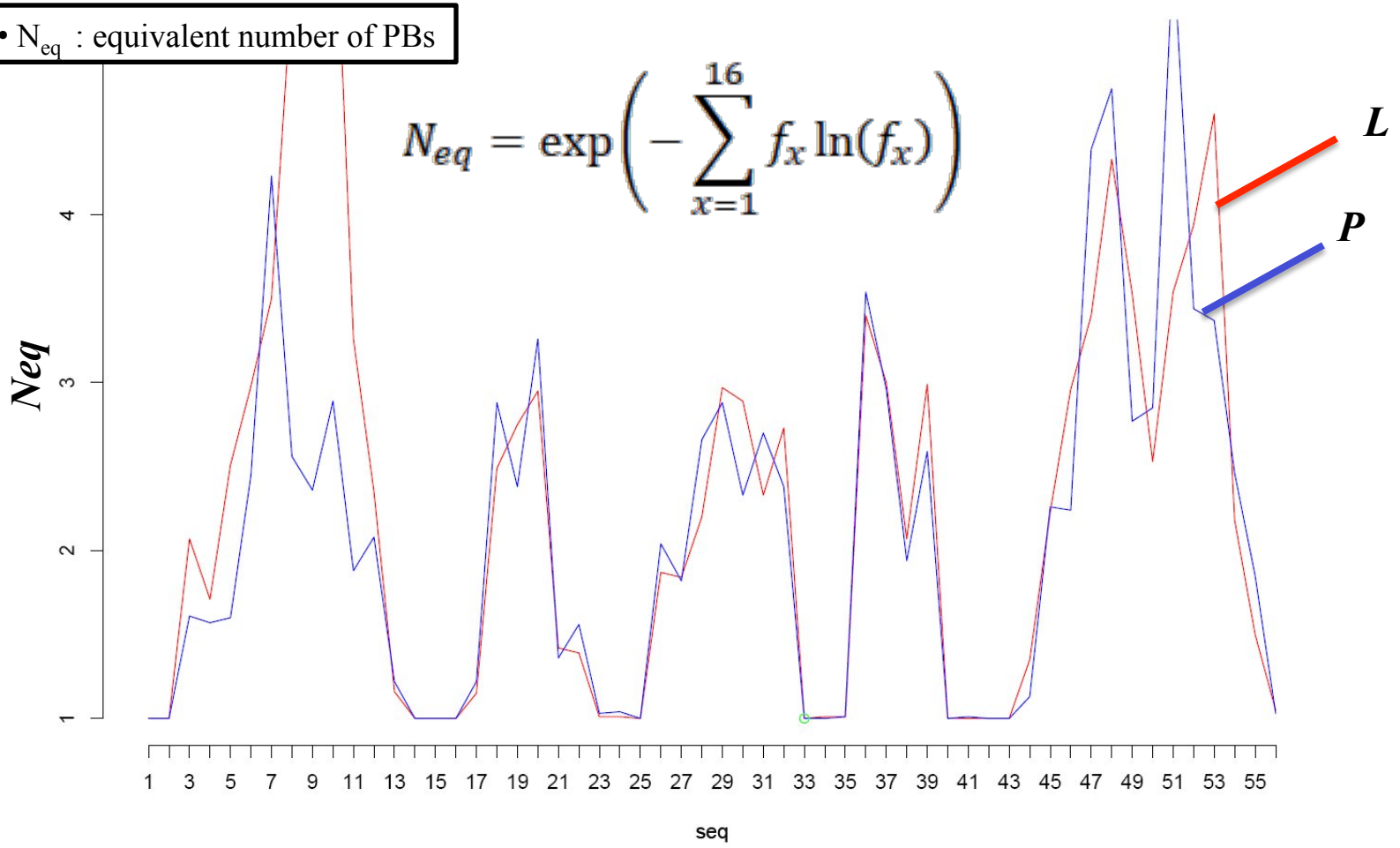


Integrins

- Both sets of simulations are quite similar

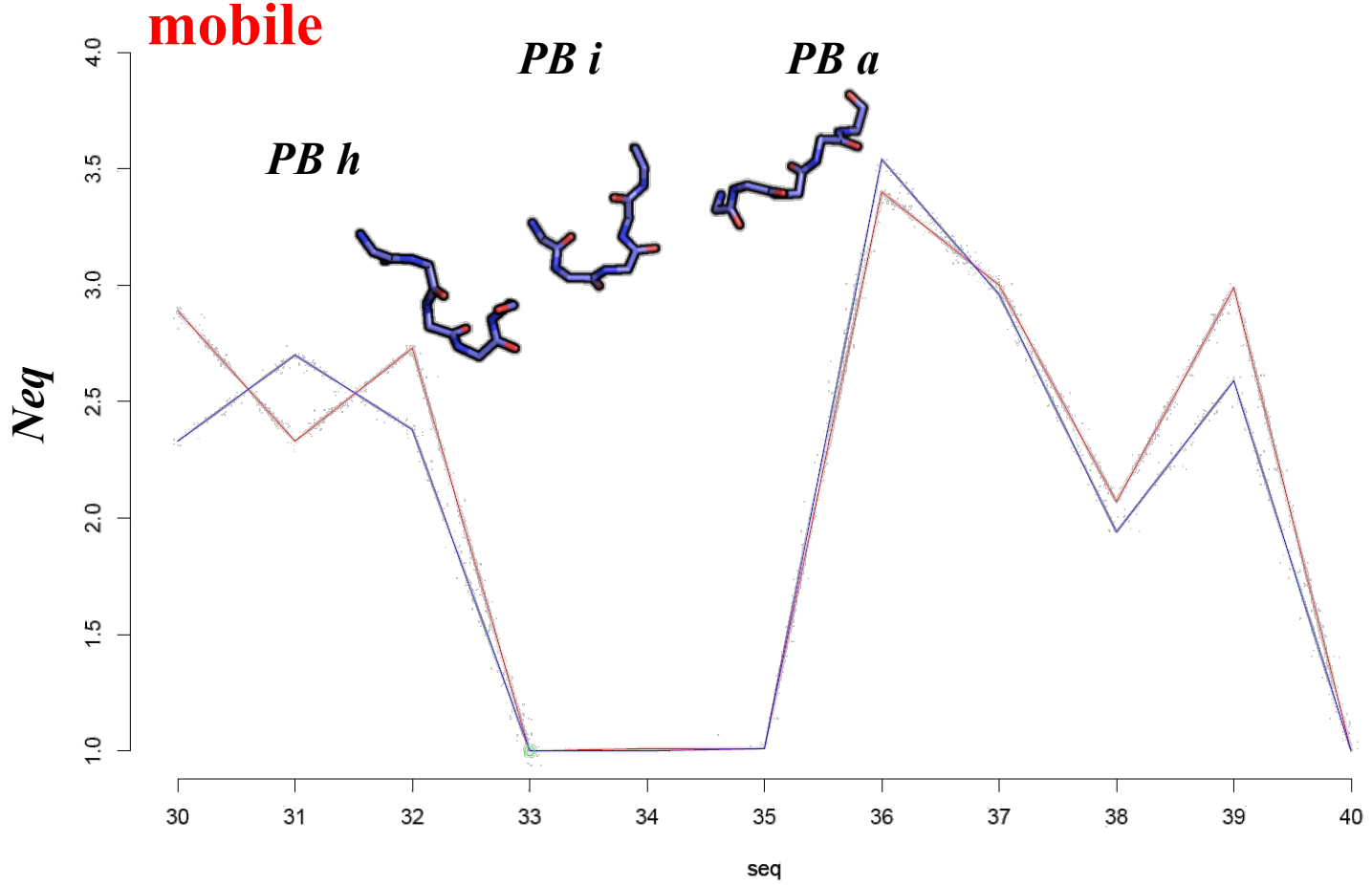
• N_{eq} : equivalent number of PBs

$$N_{eq} = \exp\left(-\sum_{x=1}^{16} f_x \ln(f_x)\right)$$



Integrins

- But position 33 is not ‘flexible’, it stays highly constraint ➔



Integrins

Conclusion

- Not real change of the epitope, always highly accessible
- But not flexible, it is mobile within deformable regions
- Proof of interest of Protein Blocks for the analysis => does not correlate with RMSf

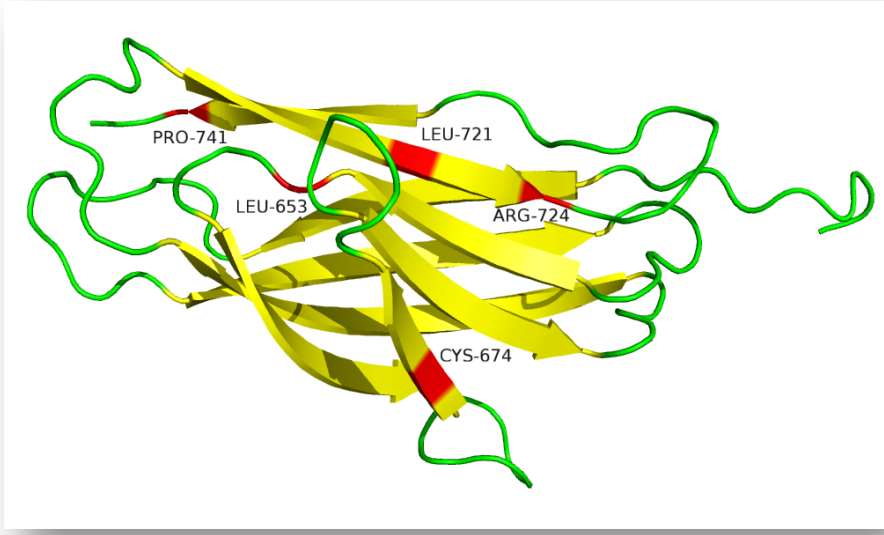
Jallu V. Poulain P., Fuchs P., Kaplan C., de Brevern A.G. (*PlosOne*, 2012)
Jallu V. Poulain P., Fuchs P., Kaplan C., de Brevern A.G. (*Biochimie*, 2014)

Allostery

- A more systematic approach: an example of Calf-1 domain
- Calf-1 domain : 141 residus (603-743) from α chain
- 7 mutations involved in GT:

Matthieu Goguet (master 2 student)

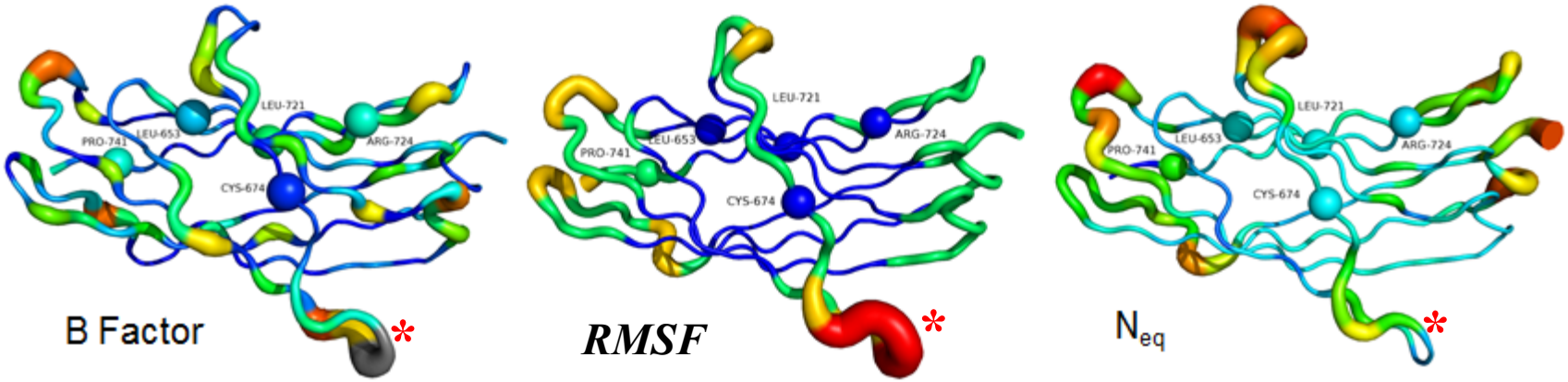
- L653R
- C674R
- L721V / L721R
- R724P / R724Q
- P741R

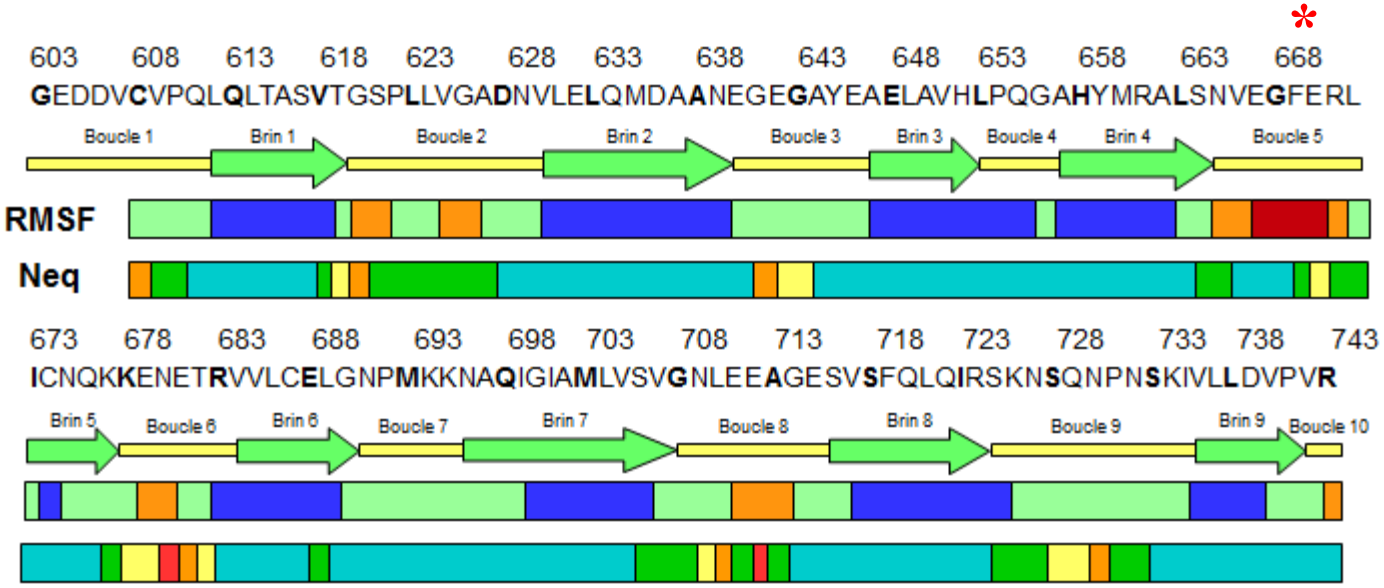


Scientific Reports 2017

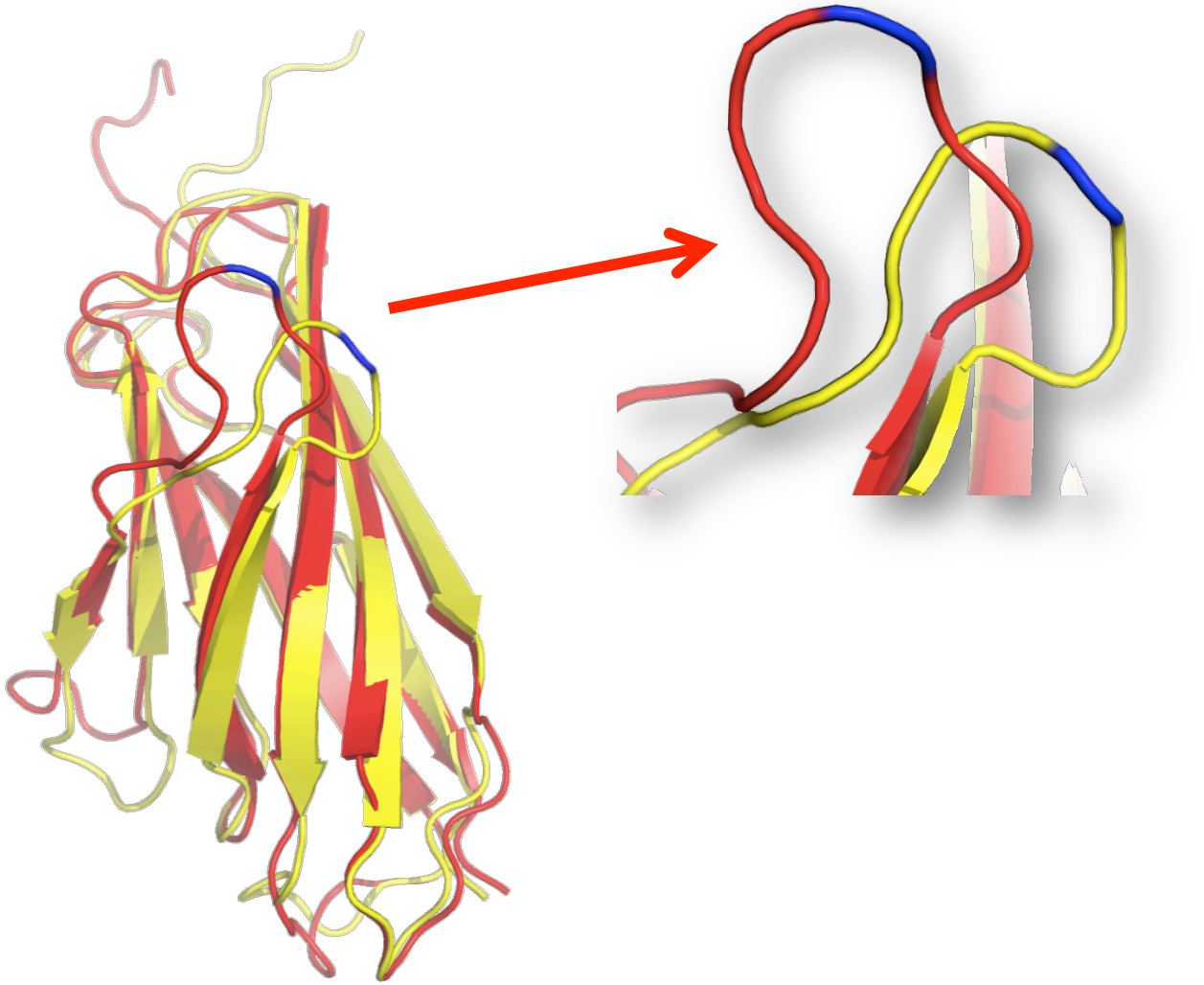
Allostery

- Analysis of WT Calf-1

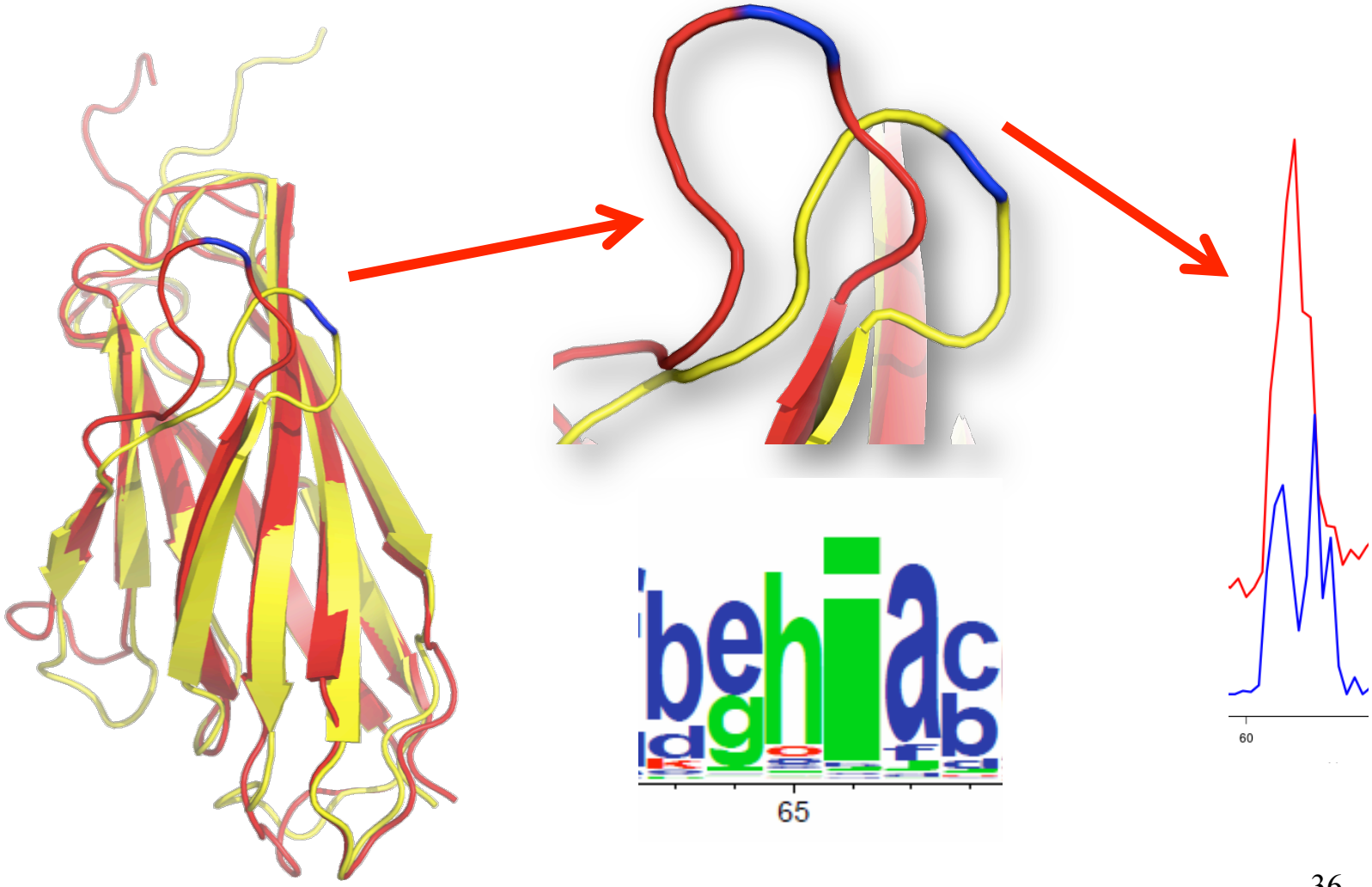




Allostery



Allostery



Inserm
Institut national
de la santé et de la recherche médicale

Allostery

SIMB
Bioinformatique & Interactions

➢ More details in

Craveur *et al*., (*Frontiers in Mol Sci*, 2015)

frontiers in Molecular Biosciences

REVIEW
published: 27 May 2015
doi: 10.3389/fmolb.2015.00020

Protein flexibility in the light of structural alphabets

Pierrick Craveur[1,2,3,4], *Agnel P. Joseph*[5], *Jeremy Esque*[6], *Tarun J. Narwani*[1,2,3,4], *Floriane Noël*[1,2,3,4], *Nicolas Shinada*[1,2,3,4], *Matthieu Goguet*[1,2,3,4], *Sylvain Leonard*[1,2,3,4], *Pierre Poulain*[1,2,3,4,7], *Olivier Bertrand*[1,3,4], *Guilhem Faure*[8], *Joseph Rebehmed*[9], *Amine Ghozlane*[10], *Lakshmipuram S. Swapna*[11,12], *Ramachandra M. Bhaskara*[11,13], *Jonathan Barnoud*[1,2,3,4,14], *Stéphane Téletchéa*[1,2,3,4,15], *Vincent Jallu*[16], *Jiri Cerny*[17], *Bohdan Schneider*[17], *Catherine Etchebest*[1,2,3,4], *Narayanaswamy Srinivasan*[11], *Jean-Christophe Gelly*[1,2,3,4] and *Alexandre G. de Brevern*[1,2,3,4]*

[1] Institut National de la Santé et de la Recherche Médicale U 1134, Paris, France, [2] UMR_S 1134, DSIMB, Université Paris Diderot, Sorbonne Paris Cité, Paris, France, [3] Institut National de la Transfusion Sanguine, DSIMB, Paris, France, [4] UMR_S 1134, DSIMB, Laboratory of Excellence GR-Ex, Paris, France, [5] Rutherford Appleton Laboratory, Science and Technology Facilities Council, Didcot, UK, [6] Institut National de la Santé et de la Recherche Médicale U964,7 UMR Centre National de la Recherche Scientifique 7104, IGBMC, Université de Strasbourg, Illkirch, France, [7] Ets Poulain, Pointe-Noire, Congo, [8] National Library of Medicine, National Center for Biotechnology Information, National Institutes of Health, Bethesda, MD, USA, [9] Centre National de la Recherche Scientifique UMR7590, Sorbonne Universités, Université Pierre et Marie Curie – MNHN – IRD – IUC, Paris, France, [10] Metagenopolis, INRA, Jouy-en-Josas, France, [11] Molecular Biophysics Unit, Indian Institute of Science, Bangalore, Bangalore, India, [12] Hospital for Sick Children, and Departments of Biochemistry and Molecular Genetics, University of Toronto, Toronto, ON, Canada, [13] Department of Theoretical Biophysics, Max Planck Institute of Biophysics, Frankfurt, Germany, [14] Laboratoire de Physique, École Normale Supérieure de Lyon, Université de Lyon, Centre National de la Recherche Scientifique UMR 5672, Lyon, France, [15] Faculté des Sciences et Techniques, Université de Nantes, Unité Fonctionnalité et Ingénierie des Protéines, Centre National de la Recherche Scientifique UMR 6286, Université Nantes, Nantes, France, [16] Platelet Unit, Institut National de la Transfusion Sanguine, Paris, France, [17] Institute of Biotechnology, The Czech Academy of Sciences, Prague, Czech Republic

Protein structures are valuable tools to understand protein function. Nonetheless, proteins are often considered as rigid macromolecules while their structures exhibit specific flexibility, which is essential to complete their functions. Analyses of protein structures and dynamics are often performed with a simplified three-state description, i.e., the classical secondary structures. More precise and complete description of protein backbone conformation can be obtained using libraries of small protein fragments that are able to approximate every part of protein structures. These libraries, called structural alphabets (SAs), have been widely used in structure analysis field, from definition of ligand binding sites to superimposition of protein structures. SAs are also well suited to analyze the dynamics of protein structures. Here, we review innovative approaches that investigate protein flexibility based on SAs description. Coupled to various sources of experimental data (e.g., B-factor) and computational methodology (e.g., Molecular Dynamic simulation), SAs turn out to be powerful tools to analyze protein dynamics, e.g., to examine allosteric mechanisms in large set of structures in complexes, to identify order/disorder transition. SAs were also shown to be quite efficient to predict protein flexibility from amino-acid sequence. Finally, in this review, we exemplify the interest of SAs for studying flexibility with different cases of proteins implicated in pathologies and diseases.

Keywords: protein structures, disorder, secondary structure, structural alphabet, protein folding, allostery, protein complexes, protein–DNA interactions

OPEN ACCESS

Edited by:
Kris Pauwels,
Vrije Universiteit Brussel, Belgium

Reviewed by:
Mark Pfuhl,
King's College London, UK
Elena Papaleo,
University of Copenhagen, Denmark

***Correspondence:**
Alexandre G. de Brevern,
Institut National de la Santé et de la Recherche Médicale U 1134, 6 Rue Alexandre Cabanel, 75015 Paris, France
alexandre.debrevern@univ-paris-diderot.fr

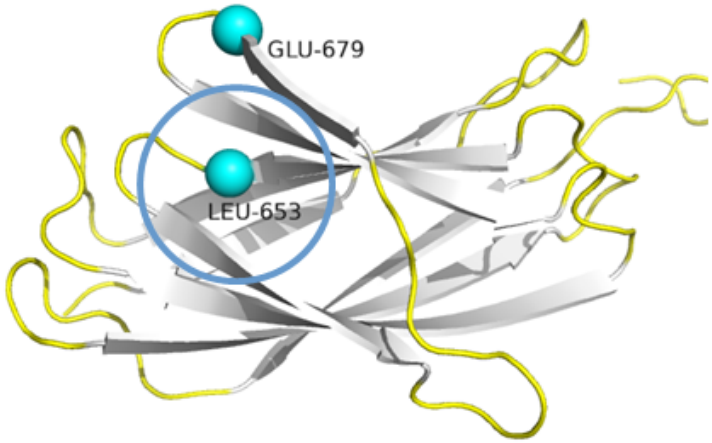
Specialty section:
This article was submitted to Structural Biology, a section of the journal Frontiers in Molecular Biosciences

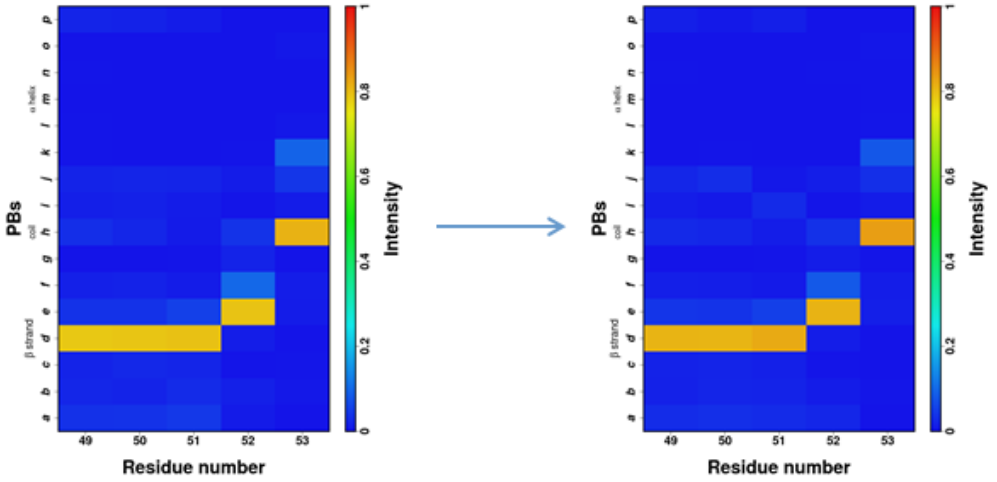
Received: 28 February 2015
Accepted: 30 April 2015
Published: 27 May 2015

Citation:
Craveur P, Joseph AP, Esque J, Narwani TJ, Noël F, Shinada N, Goguet M, Leonard S, Poulain P, Bertrand O, Faure G, Rebehmed J, Ghozlane A, Swapna LS, Bhaskara RM, Barnoud J, Téletchéa S, Jallu V, Cerny J, Schneider B, Etchebest C, Srinivasan N, Gelly J-C and de Brevern AG (2015) Protein flexibility in the light of structural alphabets. Front. Mol. Biosci. 2:20. doi: 10.3389/fmolb.2015.00020

Allostery

Variant L653R : Mutation site



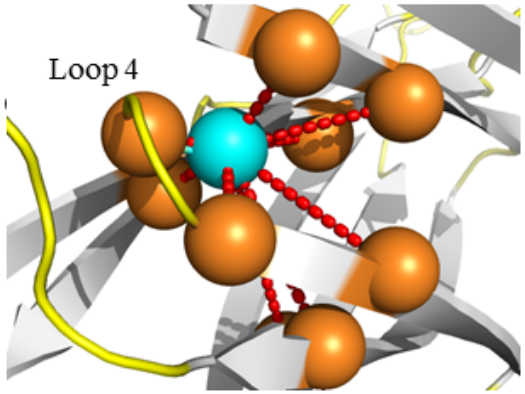


WT : 9 interactions

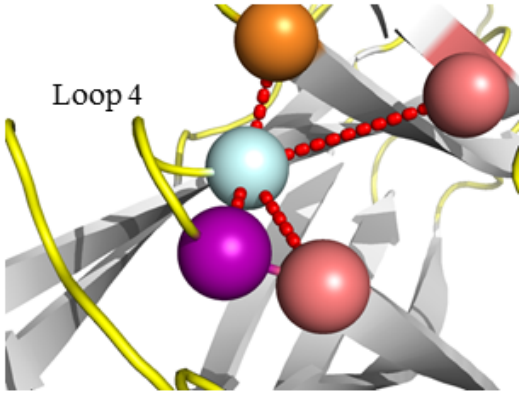
Variant : only 1 conserved but 3 news to

→ No structural changes

WT

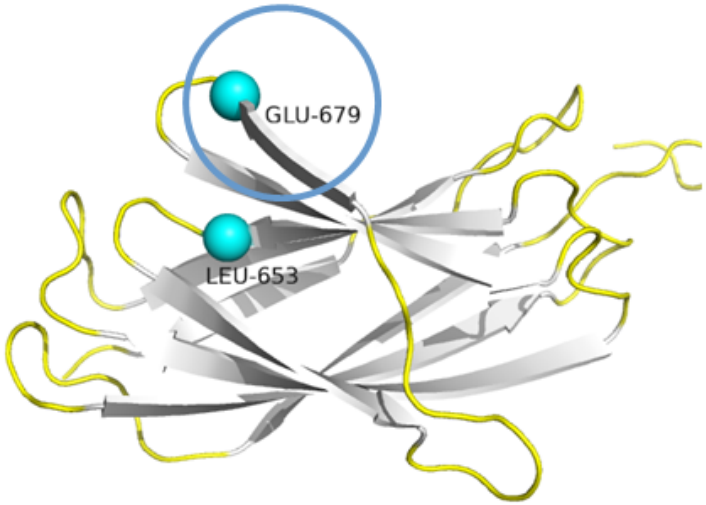


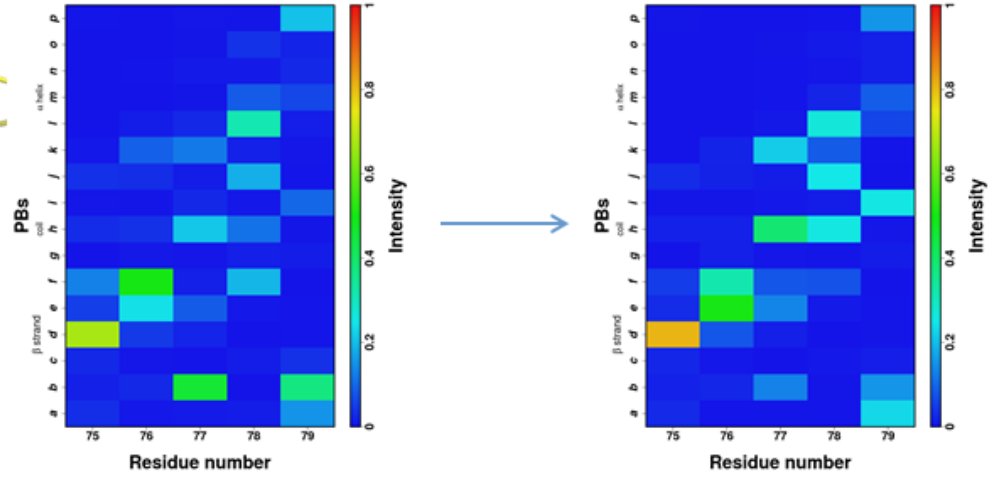
Variant



Allostery

Variant L653R : position E679



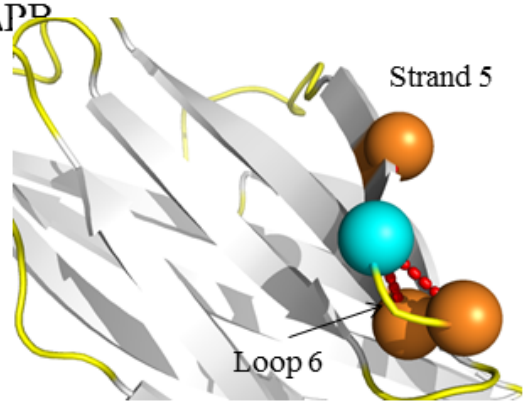


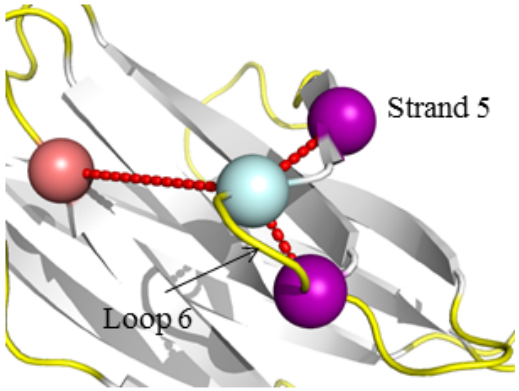
WT Variant

N_{eq} similar in WT and variant but high ΔPP

Interaction reshuffle (no conservation)

→Structural desorganization



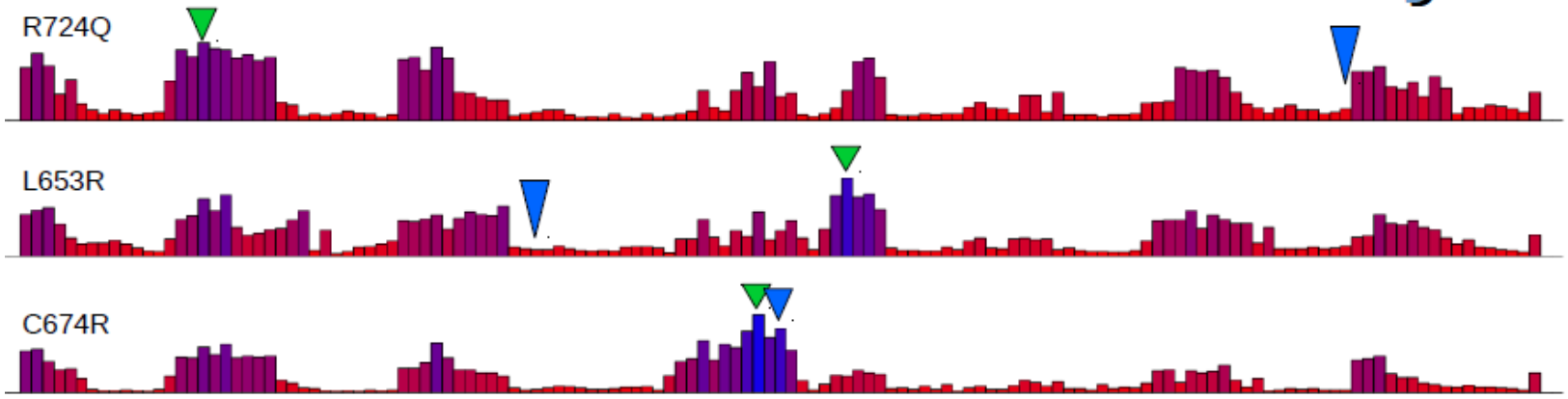


Inserm
Institut national
de la santé et de la recherche médicale

Allostery

SIMB
Bioinformatique & Interactions

Results : Global variant mobility

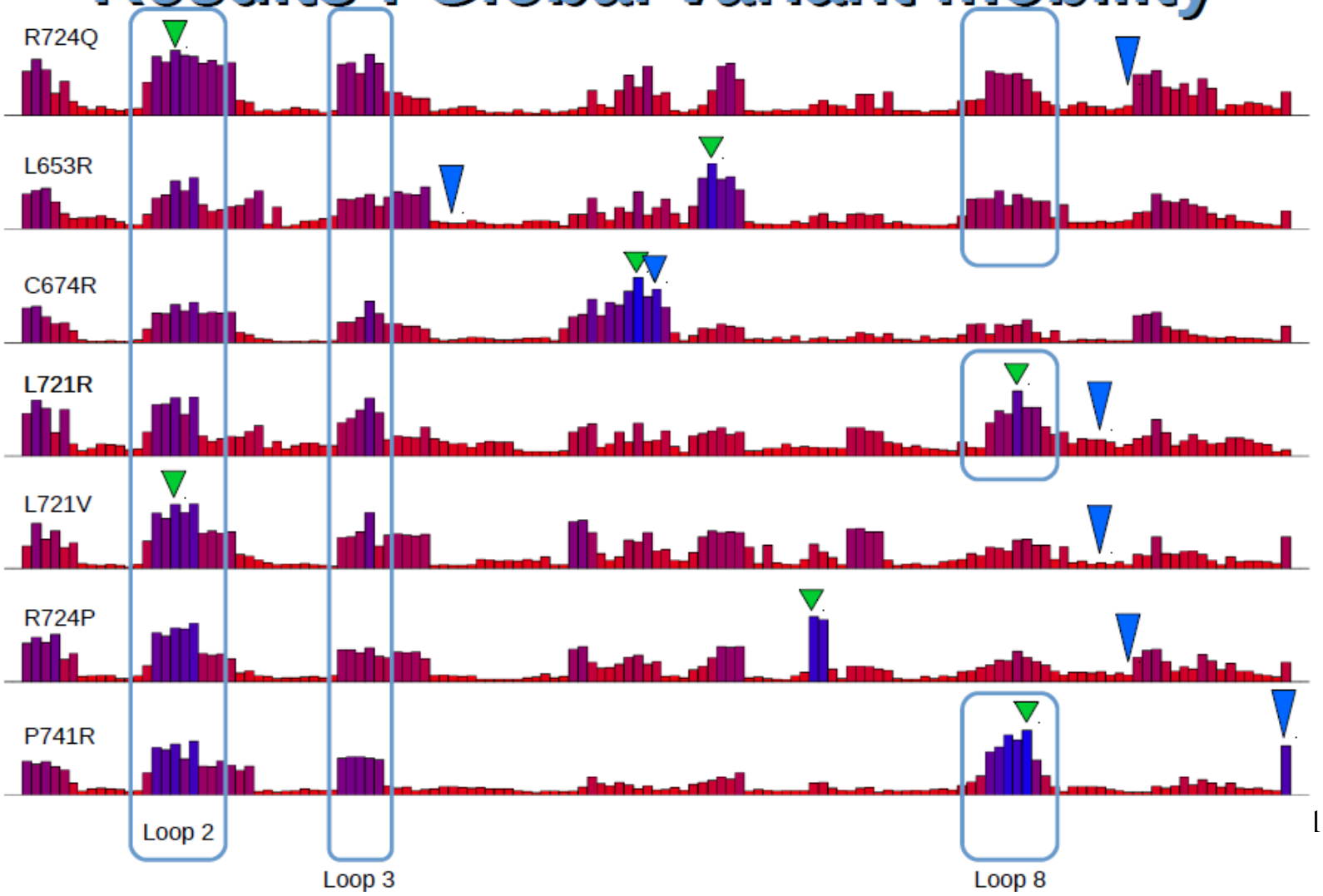


Inserm
Institut national
de la santé et de la recherche médicale

Allostery

SIMB
Bioinformatique & Interactions

Results : Global variant mobility



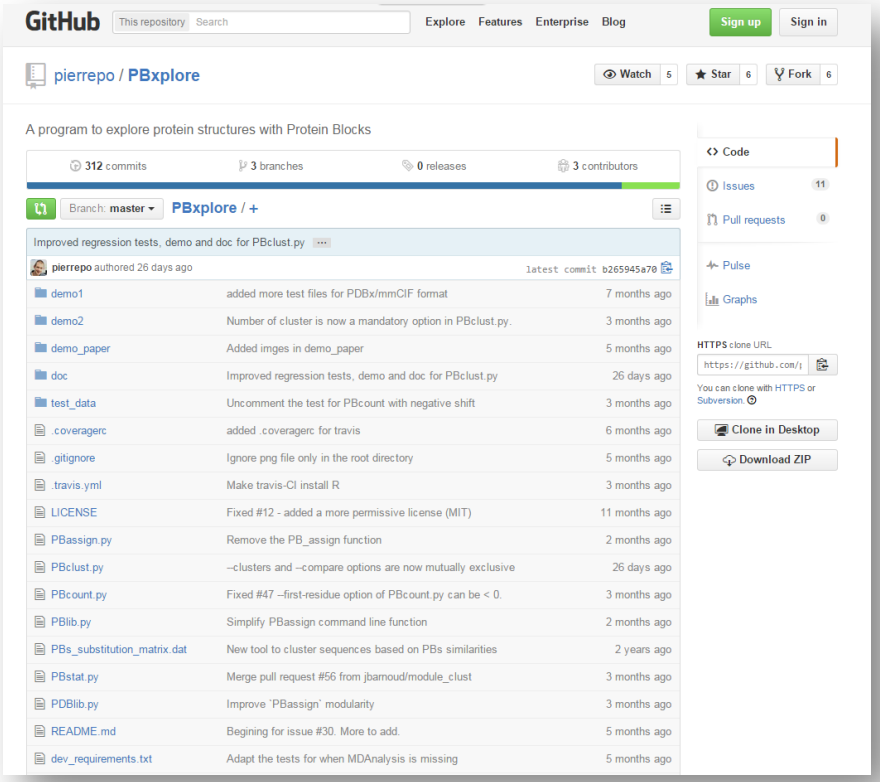
Conclusion (1)

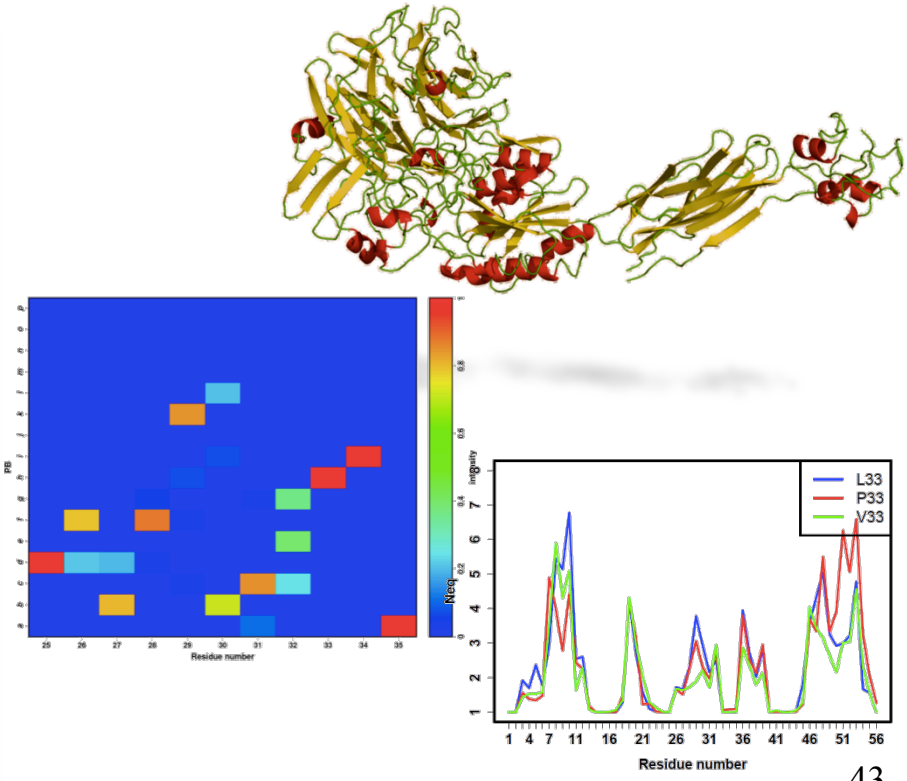
- Reliable, automated and efficient molecular dynamics protocol with interesting use of Protein Blocks
- Results on Calf-1 : no changes on mutation spots but structural local impact observed at another place.
- Does these allosteric effects are common to all of these variants ?
- Future: Study of this and others integrins

Recent results

- Last developments: deposit on ArXiv, now published in PeerJ

https://github.com/pierrepo/PBxplore



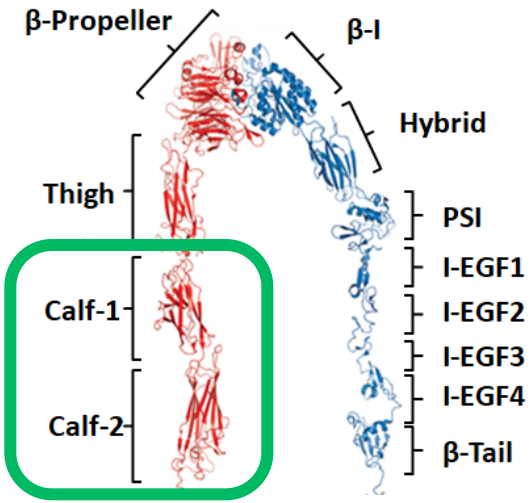


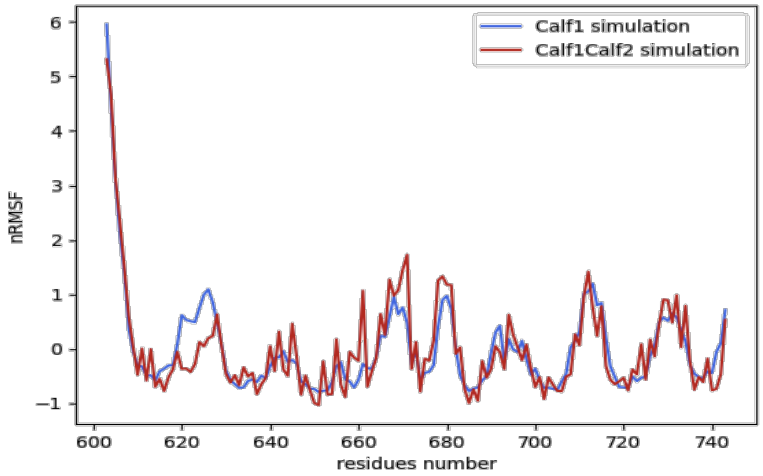
Recent results

Calf-1 + Calf- 2

- No impact of the forcefield used
- No impact on the domains cutting

Soubika Bisoo
(master 2 student)

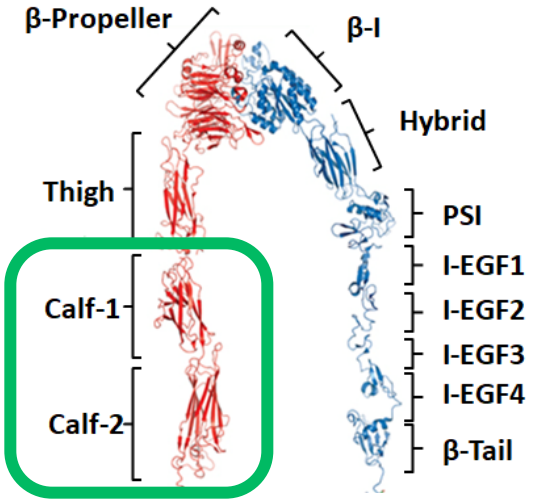


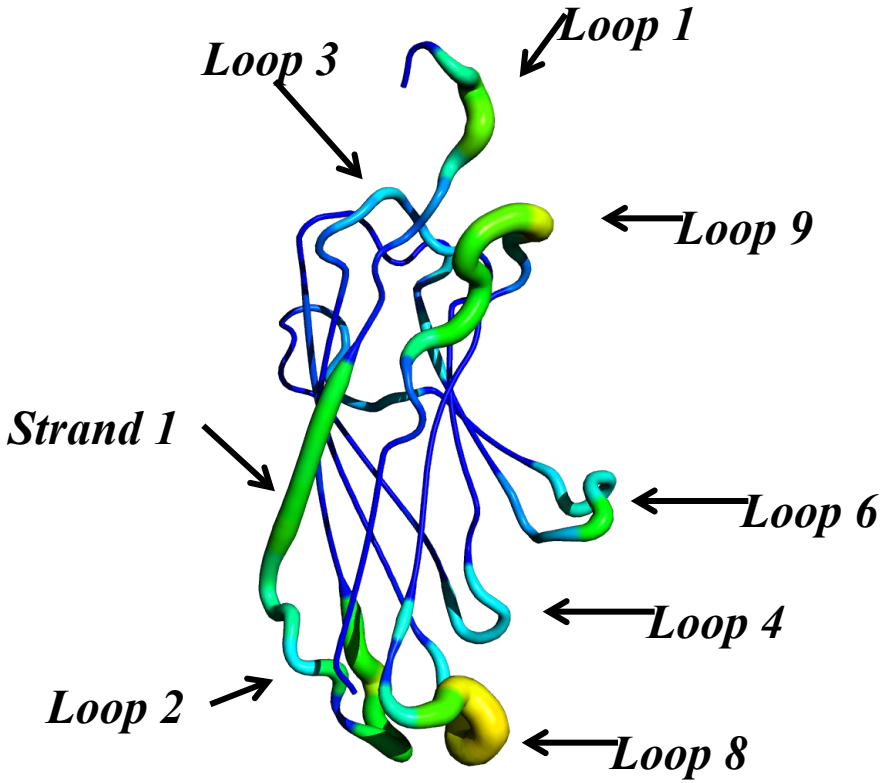


Recent results

Calf-1 + Calf- 2

- BUT allosteric effect of Calf-2 variant on Calf-1 domain !





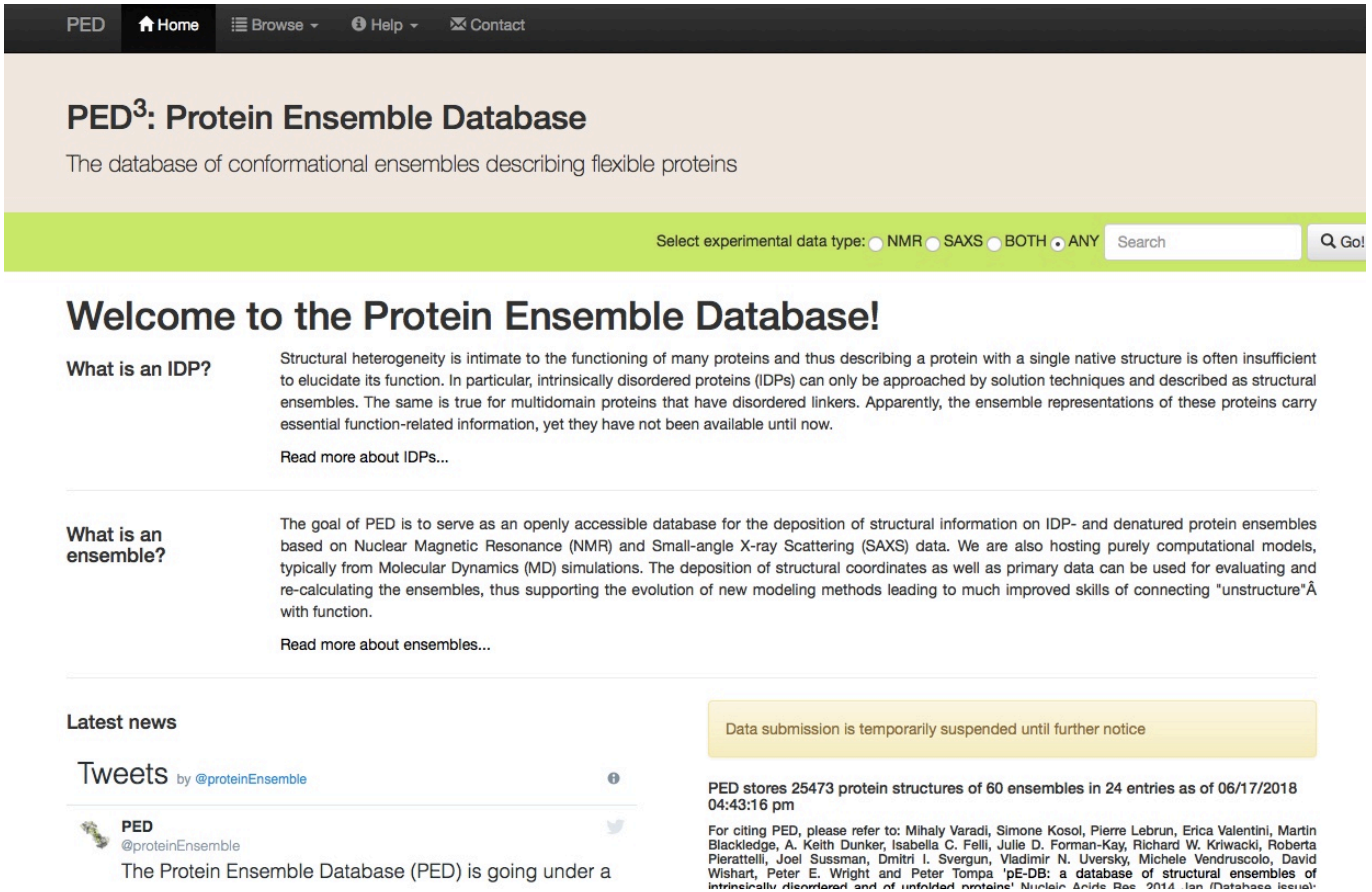
Calf-1: ΔPB

Conclusions (2)

- Secondary structures are useful tools, but they have their own limitations
- Other complementary approaches are interesting, i.e., a simple structural alphabet as the Protein Blocks
- They can help to analyze molecular dynamics, i.e., you can see what you can see without.

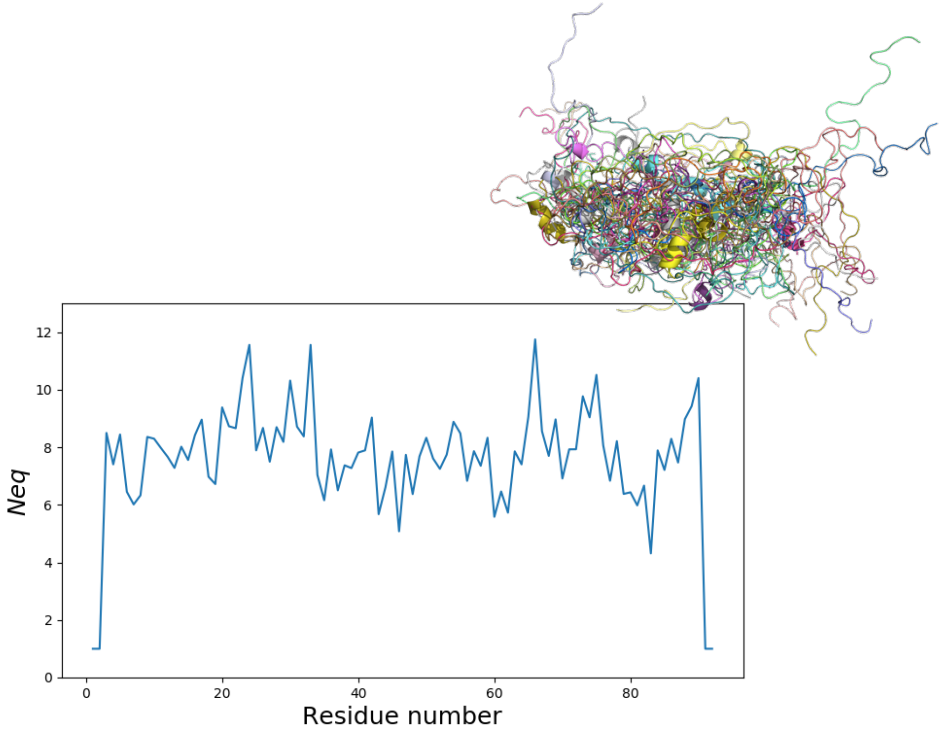
Flexibility, mobility, disorder …

- Thanks to last Belbi 2016, I am now really more interested into this question



Flexibility, mobility, disorder …

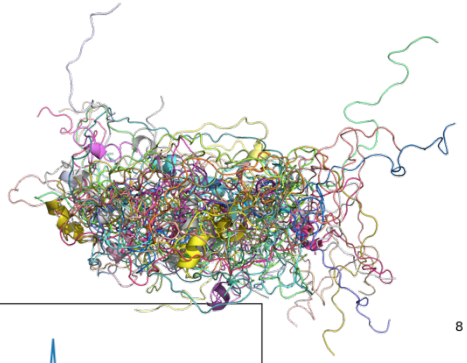
- Thanks to last Belbi 2016, I am now really more interested into this question

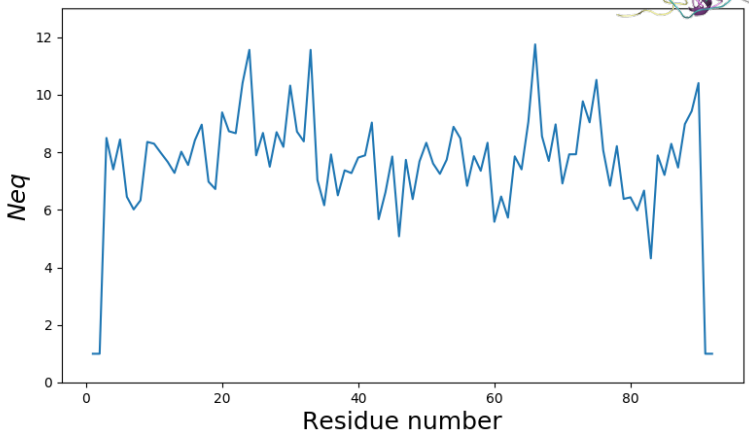


Yes, real disorder everywhere

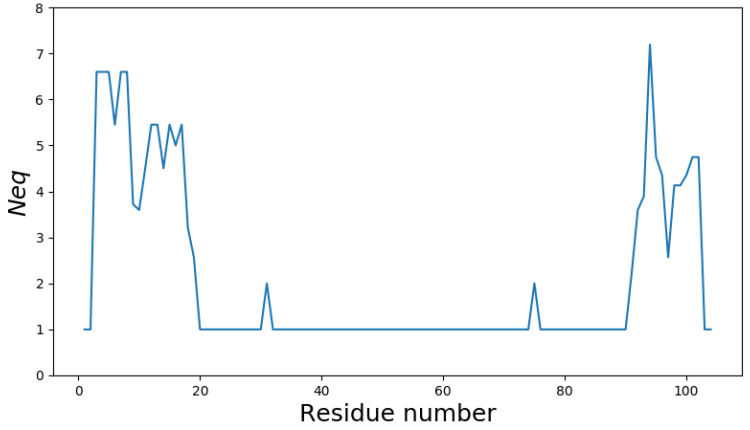
Flexibility, mobility, disorder …

- Thanks to last Belbi 2016, I am now really more interested into this question





Not really



Flexibility, mobility, disorder …

- PBs can be a good tool for that··· needs more works

Rigid ➔ ***flexible*** ➔ ***disorder***

Rigid / mobile ➔ ***flexible/deformable*** ➔ ***disorder***

Acknowledgments

S. Hazout

C. Etchebest

P. Poulain

P. Fuchs

V. Jallu

C. Kaplan

R. Petermann

T.J. Narwani

M. Goguet

S. Sali

S. Kaddhar

S. Bisoo

N. Shinada

N. Srinivasan

B. Schneider

N. Mitic

Inserm
Institut national
de la santé et de la recherche médicale

SIMB
Bioinformatique & Interactions

BELBI 2018

Thank you for your time and attention