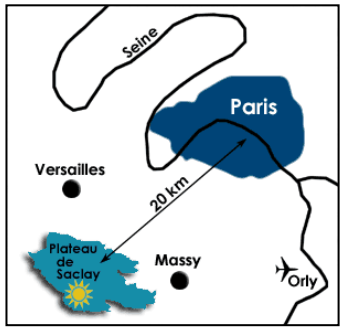
Memprot & Dadimodo: programs for modeling the detergent belt in solubilized membrane protein complexes & re-orienting domains of multi-domain proteins

Javier Pérez

Synchrotron SOLEIL



Outline

- ✓ Short presentation of current available techniques at beamline SWING
- ✓ Memprot program: why and how

  AQP0 as the « Guinea pig » protein
- ✓ Example of a recent application using Memprot:

  MhsT protein in 4 different detergents
- ✓ Dadimodo program: why and how
- ✓ Example of application coupling Memprot & Dadimodo:

  HasA-HasR protein with two different constructs
- ✓ Short conclusion: foreseen new functionalities for Memprot and Dadimodo

Outline

- ✓ Short presentation of current available techniques at beamline SWING
- ✓ Memprot program: why and how

  AQP0 as the « Guinea pig » protein
- ✓ Example of a recent application using Memprot:

  MhsT protein in 4 different detergents
- ✓ Dadimodo program: why and how
- ✓ Example of application coupling Memprot & Dadimodo:

  HasA-HasR protein with two different constructs
- ✓ Short conclusion: foreseen new functionalities for Memprot and Dadimodo

Current techniques at beamline SWING

https://www.synchrotron-soleil.fr/en/beamlines/swing

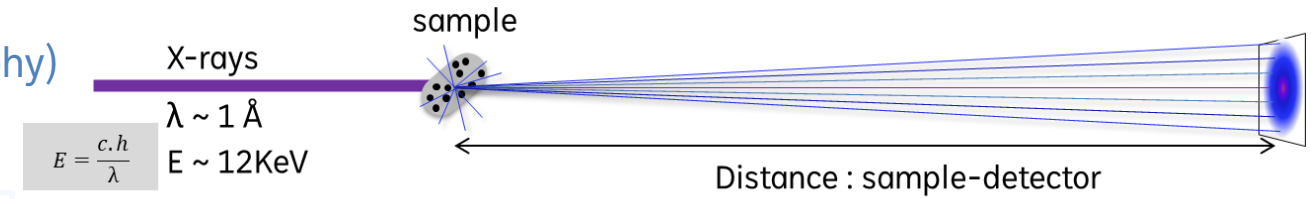
Since 2008

- **S**mall and **W**ide angle X-ray scatter**ING**

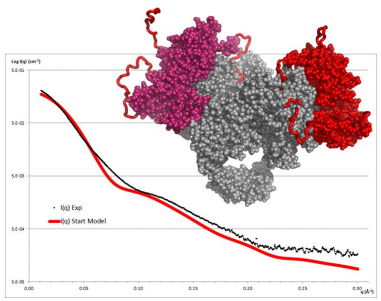
From macromolecule to material.

- Hard X-rays Energy (U20 undulator): 5 keV to 16 keV (wavelength: 2.5 Å to 0.75 Å)
- Usual beam size (FWHM): 25-400 (H) x 25-100 (V) μm²
- Structural information from non-crystalline samples (scale: nm to μm)

- BioSAXS
- « Classical » SWAXS
- μSAXS mapping
- PXCT (ptycho-tomography)

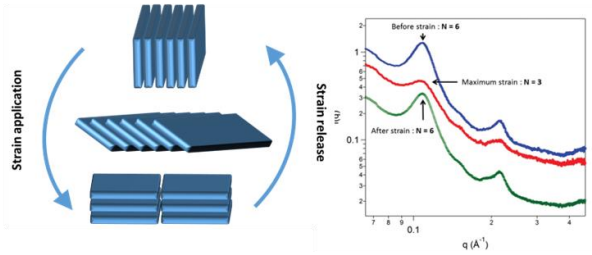


Protein conformations



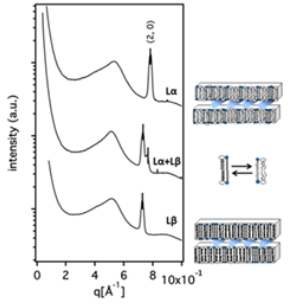
Roudenko O., Thureau A. & Pérez J., March 2018
Petrella, S., et al. (2019). Structure, 27, 579–589.

Deformation under strain

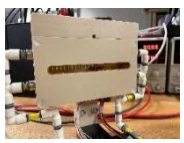
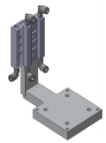


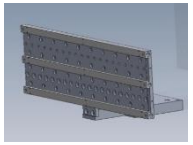
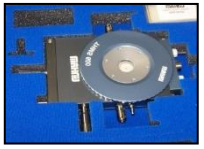
Beaudoin, E., *et al.* (2017), *Nanoscale*, 9(44): 17371-17377

Phase transition of lipids/ liquid crystals



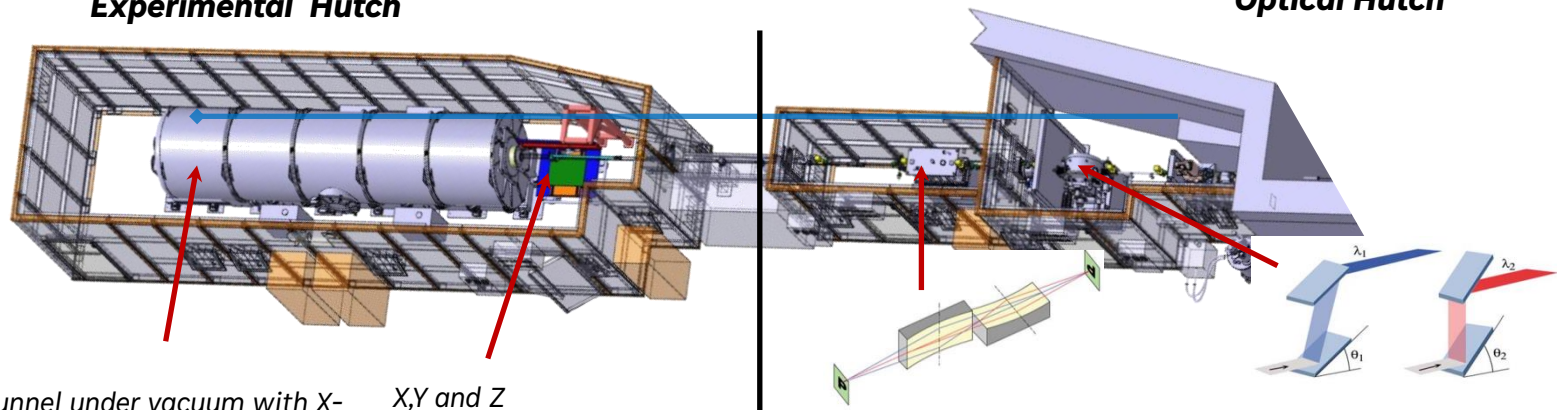
Bizien, T., *et al.* (2015), *Langmuir*, 31(44): 12148–12154.

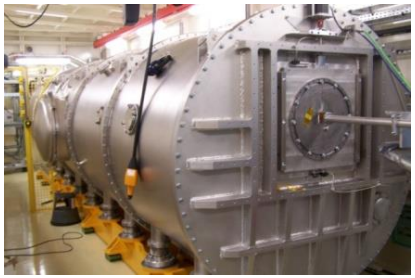



SWING Beamline

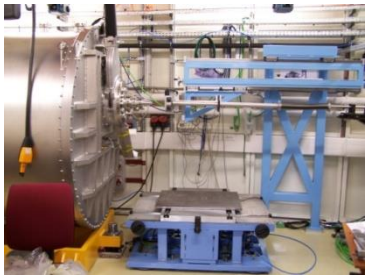
Experimental Hutch | **Optical Hutch**



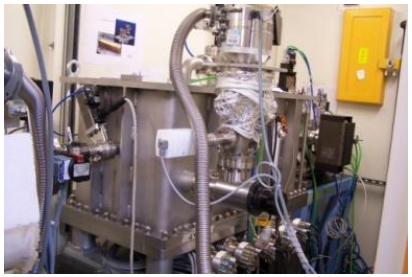
Tunnel under vacuum with X-rays detector



X,Y and Z motorized table



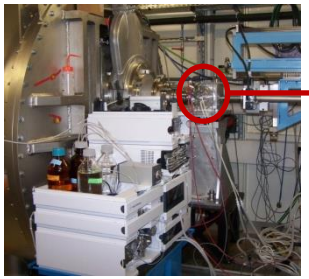
Double cristals mirror



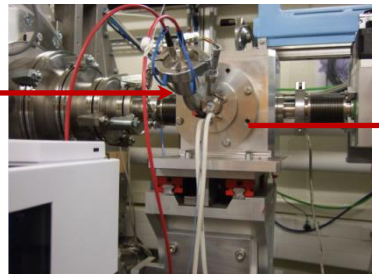
Monochromator

Sample environment dedicated to the biology :

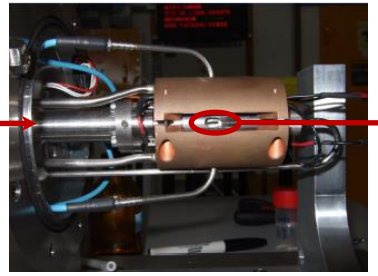
SEC-HPLC device



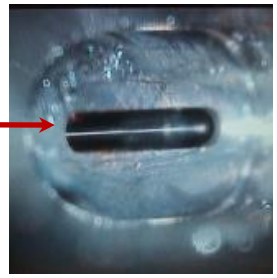
SEC-SAXS



SAXS cell vacuum chamber



Details of the SAXS cell

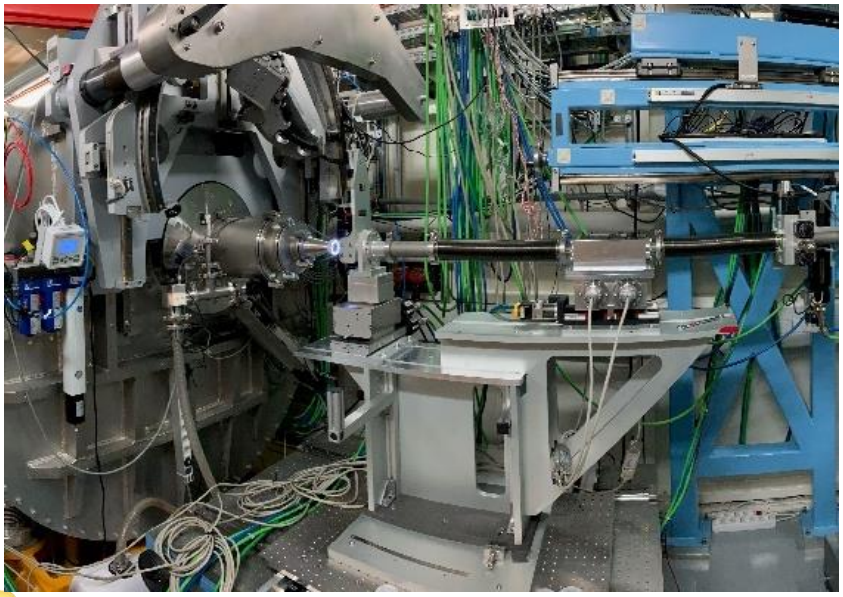


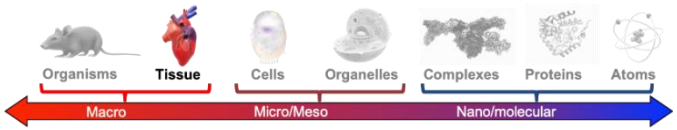
Quartz capillary

Imaging activities

Microfocus set-up

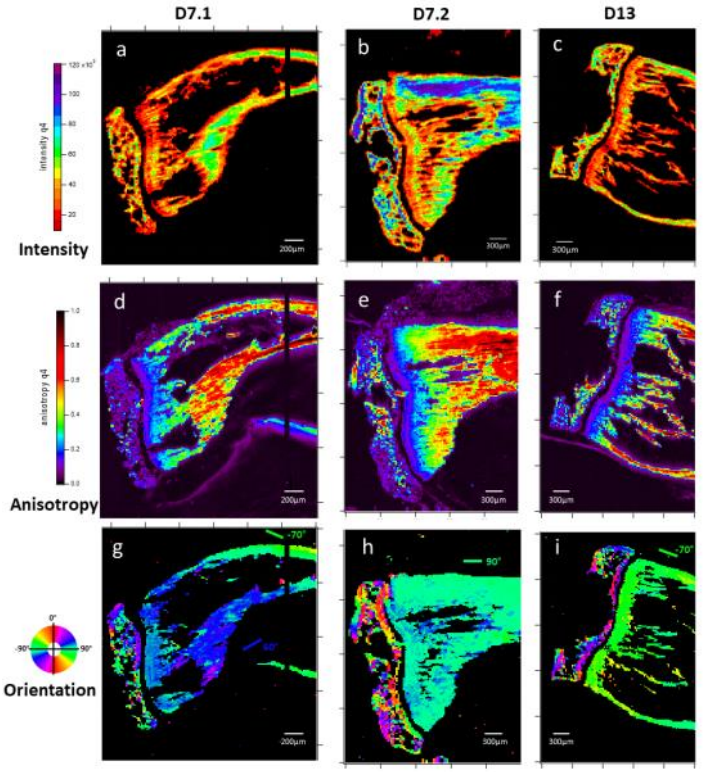
Scanning **SAXS** cartography
2D scattering/diffraction maps (spatial resol 25 µm)
Sample size: a few mm





Local ultrastructure of Bone sarcoma

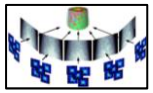
Slice ~100 µm thick



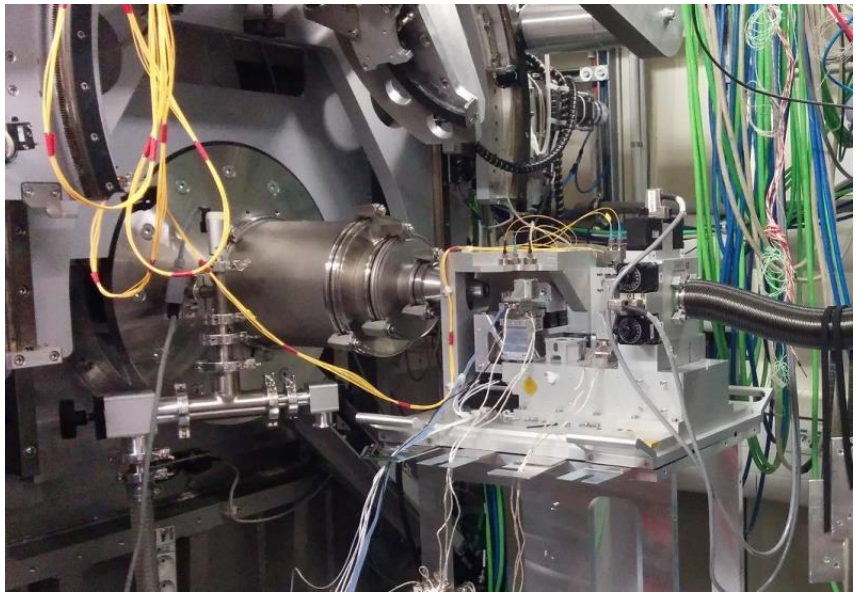
A. Bardouil, F. Artzner *et al.*, submitted.

Ptycho-tomography (PXCT)

Nanoprobe set-up

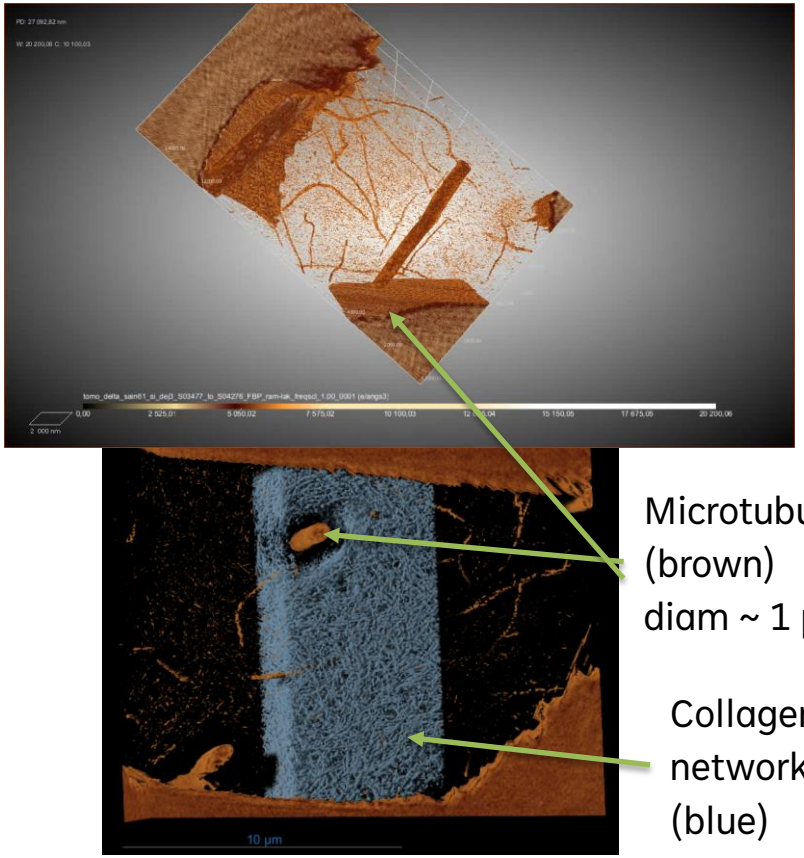


Ptychography **X**-rays **C**omputed **T**omography
Based on diffraction from a **coherent X-rays** beam
3D density maps at high spatial resolution (< 50 nm)
Sample size : 10 – 30 µm



Dentin high resolution density map

Dentinogenesis Imperfecta study (cyl. Ø 15 µm)

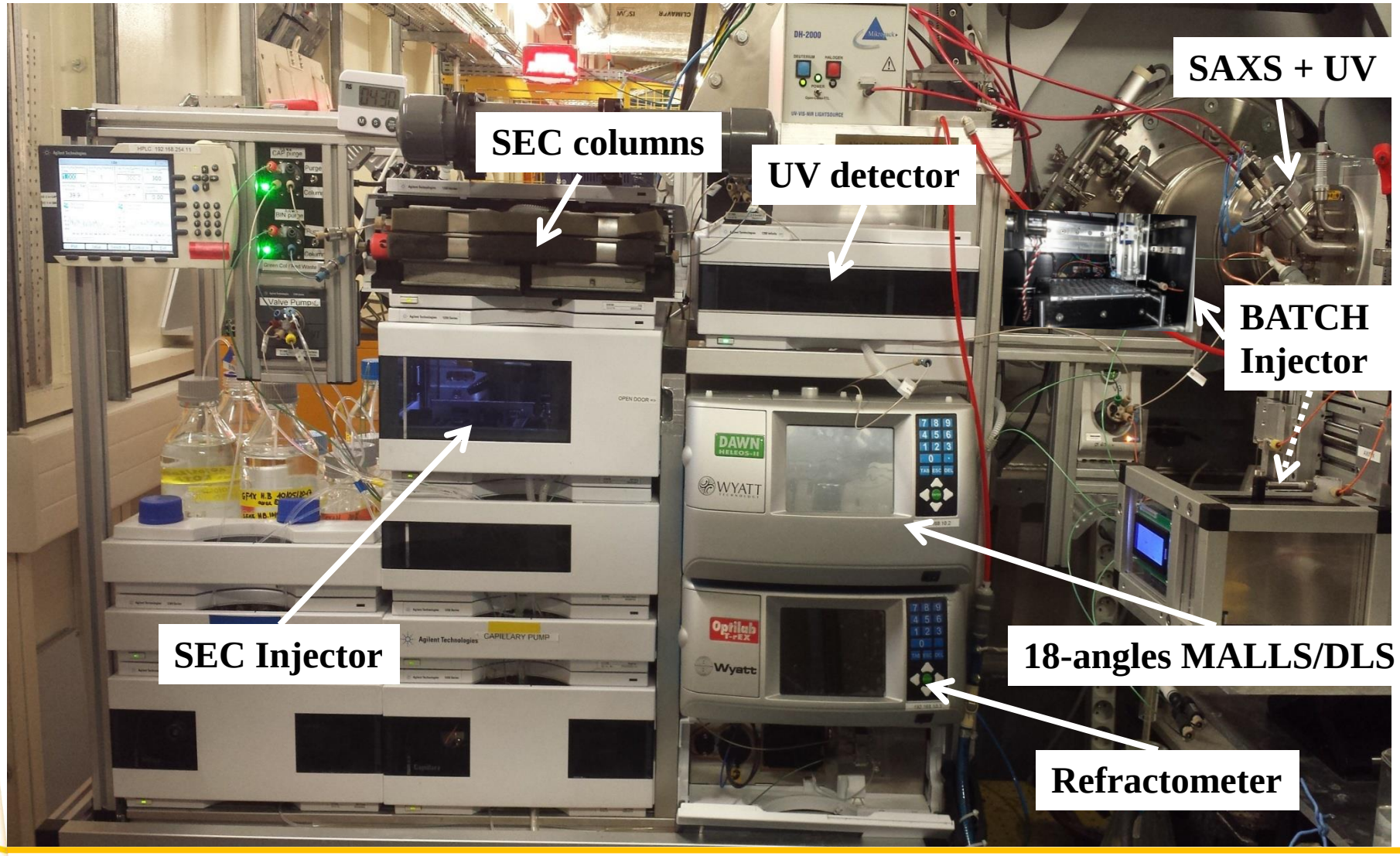


Microtubule (brown) diam ~ 1 µm

Collagen network (blue)

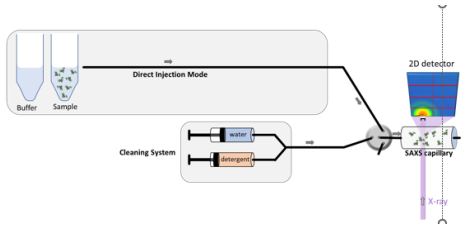
Elsa Vennat
Margot Riou
Muriel De la Dure-Molla
Alexandre Jeetoo
Thomas Reiss
Alexis Nicolay

HPLC – MALLS/QELS – SAXS – RI online



SWING: a robotic arm for BioSAXS

- High-throughput injecting robot
  - Duty cycle : 3'30''
    - Injection / measurement / cleaning / drying
  - Injection volumes 10 to 50 µL
  - No dilution effect : [0.2 – 2 mg/mL]



~ 170 µL of tubing from injection to SAXS capillary

- Wetting effect induce loss of sample volume
- Difficulties for injecting viscous samples
- Cleaning and drying take time
- Cross contamination might occur

Ideally : Directly inject the sample into the capillary

0 µL of tubing from injection to SAXS capillary

- No sample loss
- Faster Cleaning and Drying process
- Viscous samples are allowed
- No cross contamination

ISAC Support Group @ SOLEIL (Laura Muñoz)

anr DeepSAXS

+ Automatic switch with SEC-SAXS mode

Outline

- ✓ Short presentation of current available techniques at beamline SWING
- ✓ Memprot program: why and how

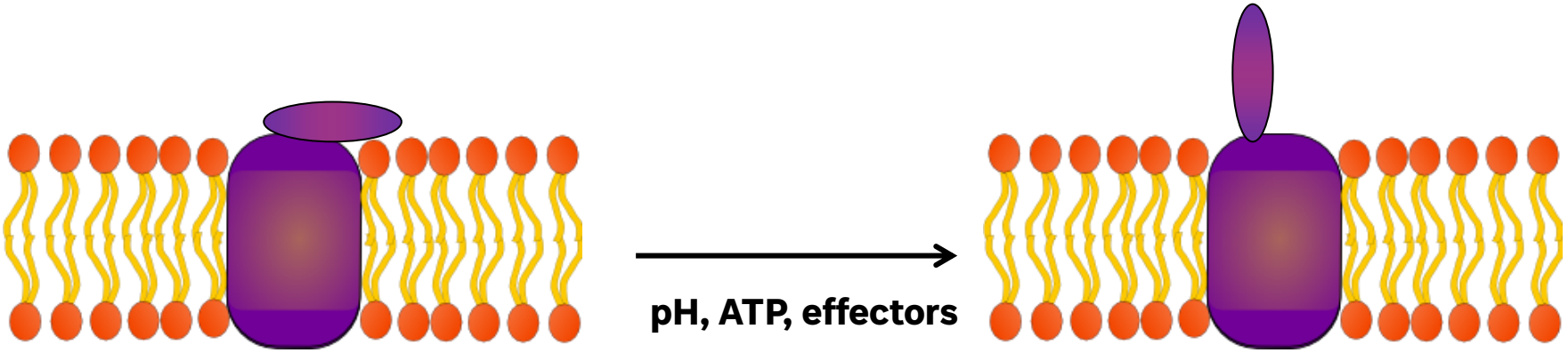
  AQP0 as the « Guinea pig » protein
- ✓ Example of a recent application using Memprot:

  MHST protein in 4 different detergents
- ✓ Dadimodo program: why and how
- ✓ Example of application coupling Memprot & Dadimodo:

  HasA-HasR protein with two different constructs
- ✓ Short conclusion: foreseen new functionalities for Memprot and Dadimodo

Motivation: apply SAXS to Membrane proteins

- SAXS is good at monitoring conformation changes
- Membrane proteins undergo conformational changes

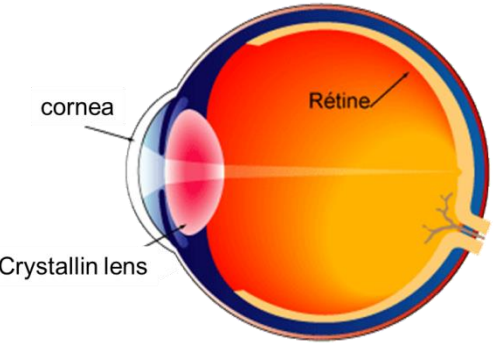


One conformation A **One conformation B**

- How can we use SAXS to monitor membrane proteins conformation changes ?
- How can we use SAXS with a membrane protein of known structure ?

Our starting model protein: Aquaporin-0

Crystalline lens (eye)

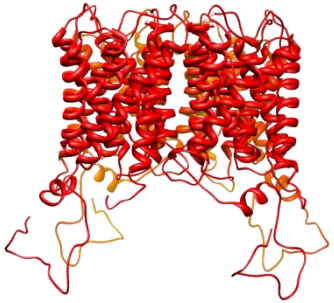


AQP0 (ex-MIP)
60 % of the membrane protein content

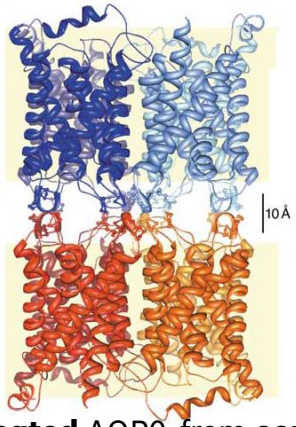
Natively tetramer

Water transport across cell membranes

✓ Two types of known existing states
✓ 3D already obtained



Full AQP0, from **cortex**
→ **Tetramer**



Truncated AQP0, from **core**
→ **Octamer**

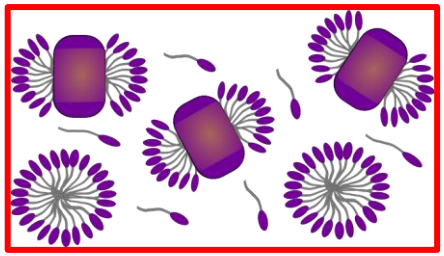
Gonen et al., Nature 2004

Purification of Full AQP0

- From bovine eye to lens membrane

- From lens membrane to AQP0 in solution
- Detergent:
  Dodecyl-β-D-maltopyranoside (DDM)

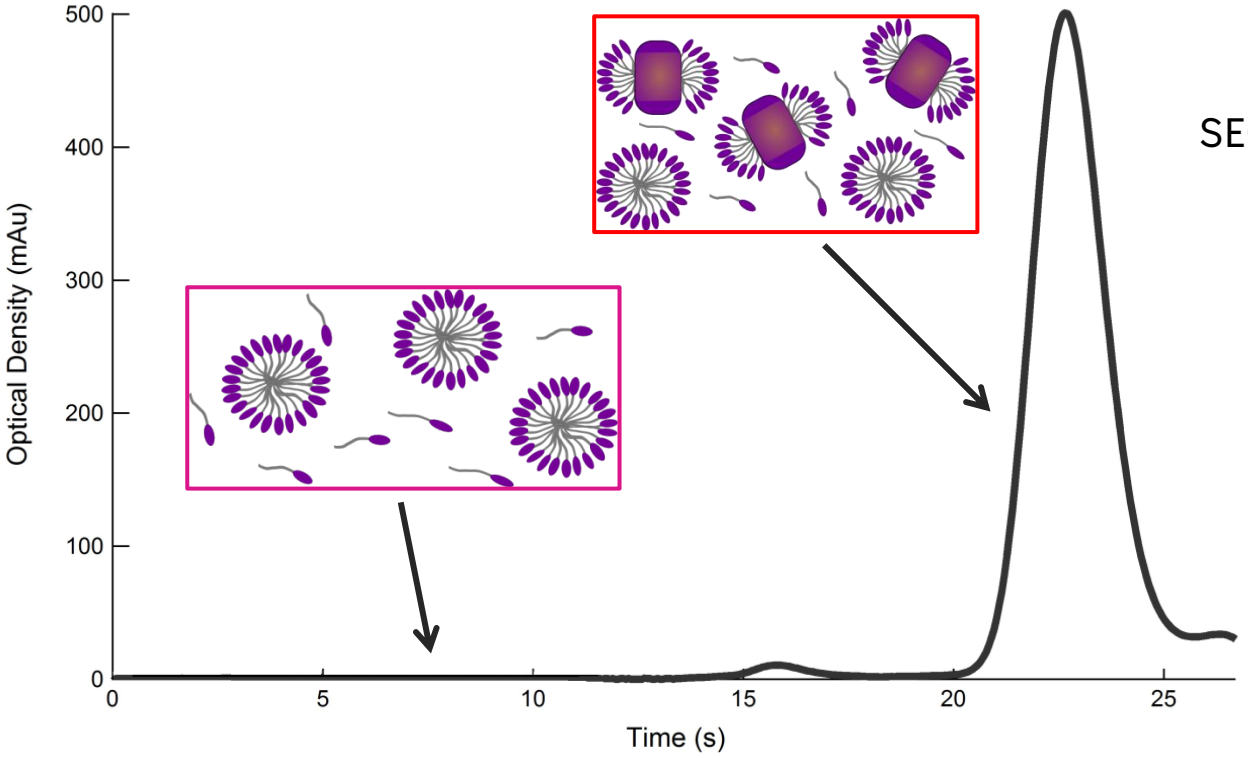
✓ Concentration reached **4 mg/ml** (2ml)



2 problems for SAXS:
- **Mixture**
- **Detergent belt**

SEC elution using a 0.2% DDM buffer

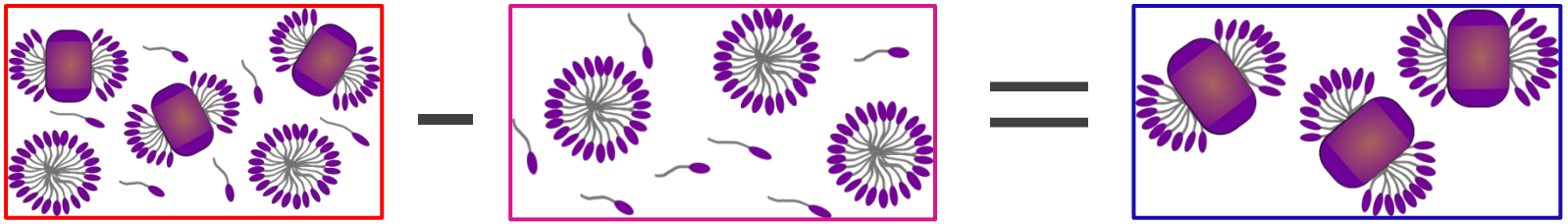
- **Mixture problem solved with the HPLC**



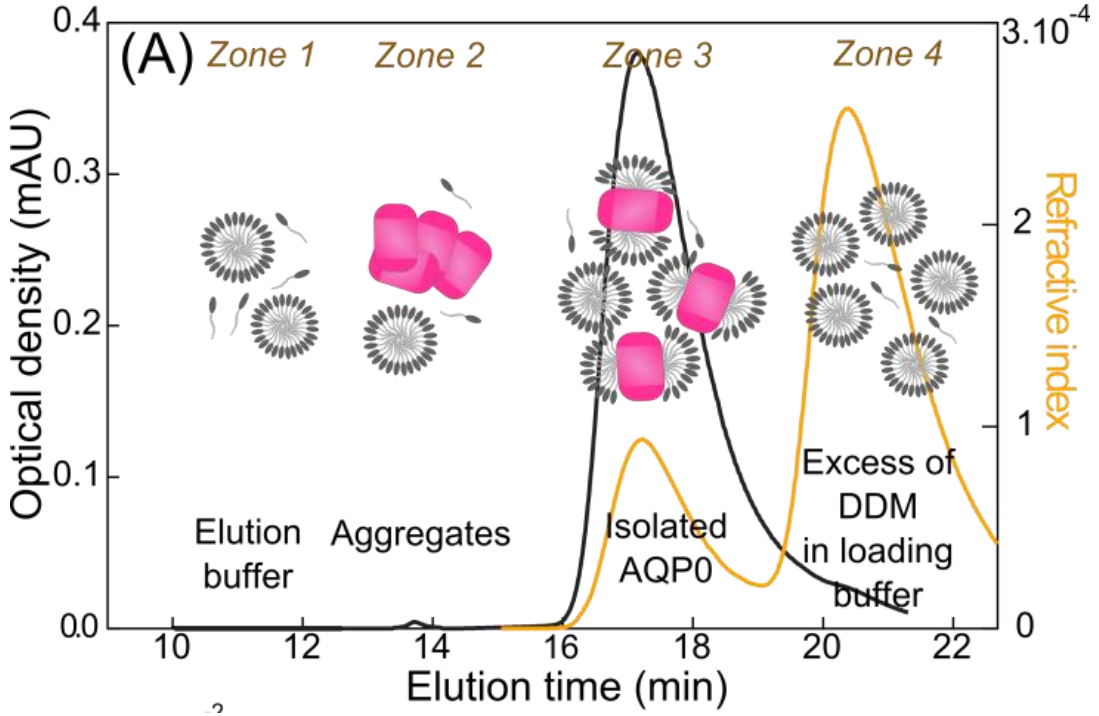
SEC-HPLC / SAXS combination

✓ Prevents the aggregates in the sample

✓ Subtraction of free micelles of detergent



SEC elution : O.D. (280nm) + R.I.

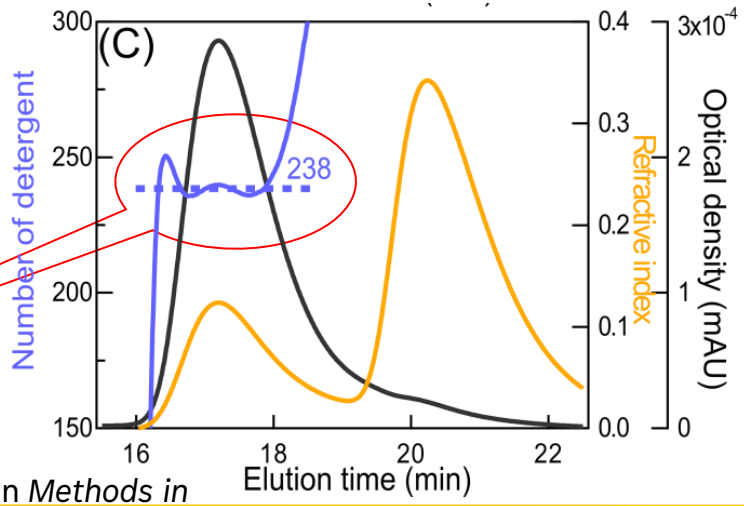


From Refractive Index & UV abs (280nm)

$$\varphi = \frac{OD \times (dn/dc)_{Det}}{\varepsilon_{AQP0}} \times \left[RI - \frac{OD \times \left[(dn/dc)_{Prot} - (dn/dc)_{Det} \right]}{\varepsilon_{AQP0}} \right]^{-1}$$

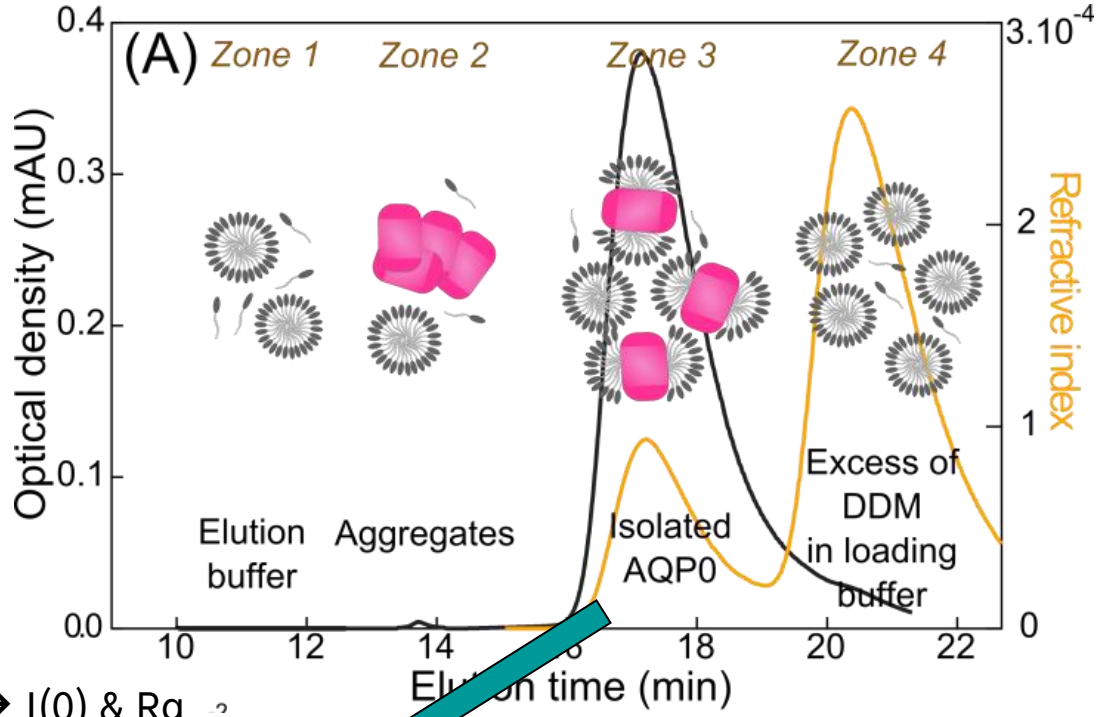
$$N_{Det} = \frac{1-\varphi}{\varphi} \times \frac{M_{Prot}}{M_{Det}}$$

N_{Det} = 238 ± 15 molecules per protein

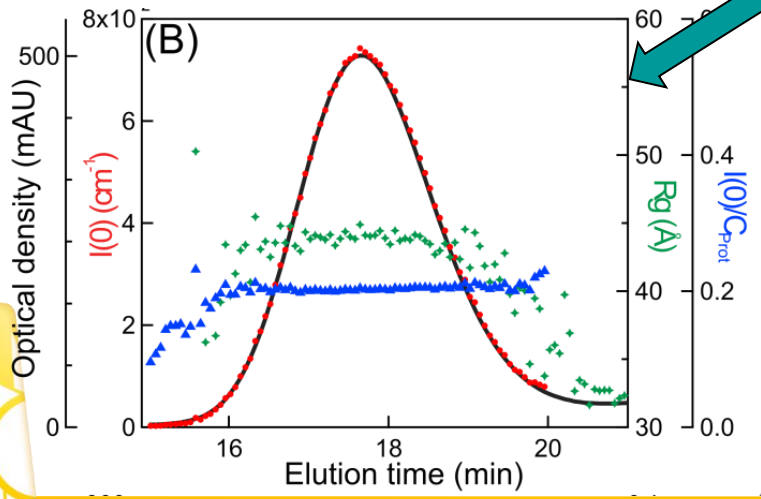


Baranowski, M., Pérez, J. in *Methods in Molecular Biology., 2168: 177-197. (2021)*

SEC elution : O.D. (280nm) + SAXS



+ SAXS → I(0) & Rg



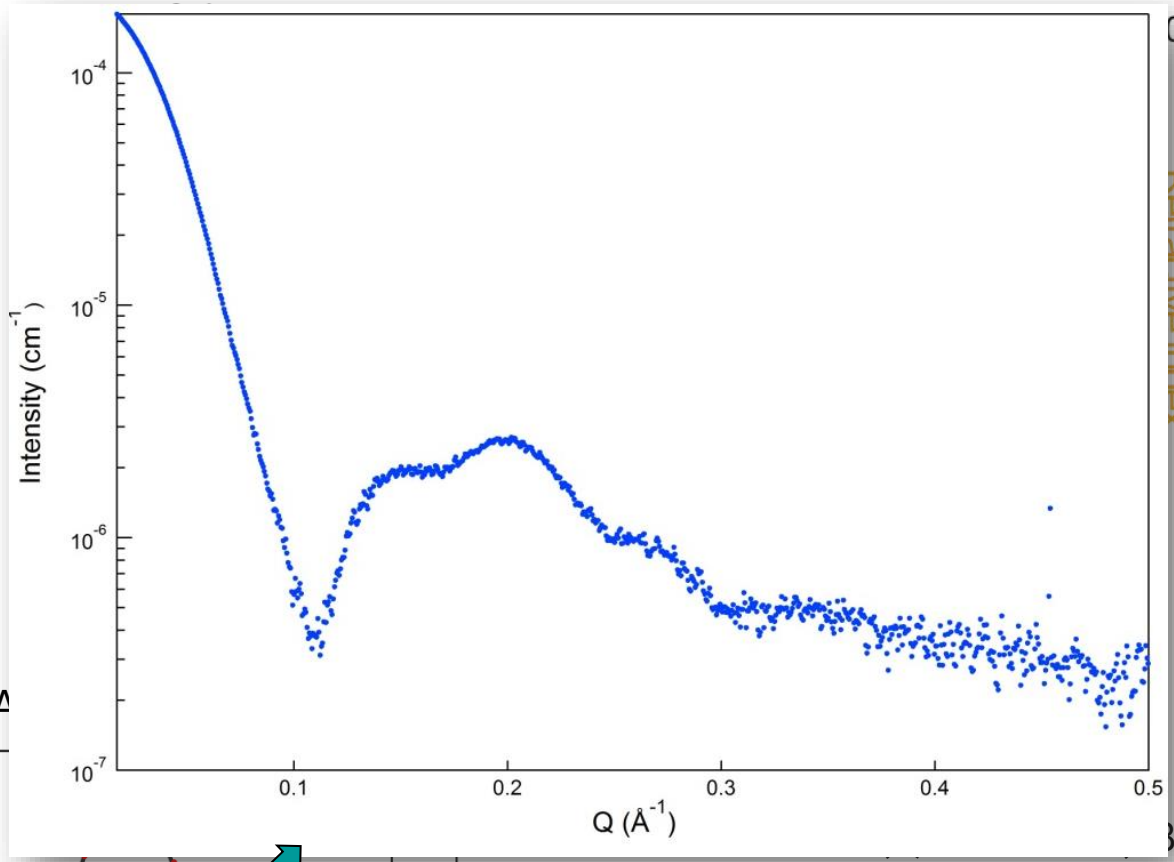
$$\frac{I(0)}{C} = M \cdot \frac{{r_e}^2}{N_A}\left(\frac{n_e N_A}{M} - \rho_0 \bar{v}\right)^2$$

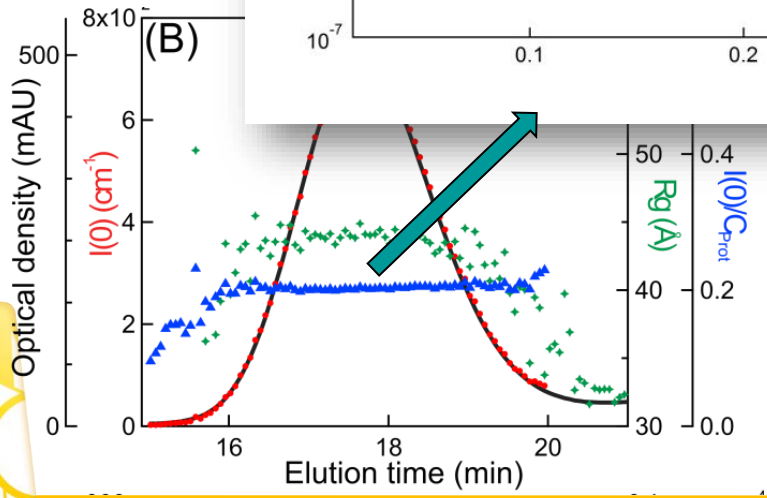
⇒ **M** is constant

⇒ No depletion of detergent

⇒ Monodisperse solution

SEC elution : O.D. (280nm) + SAXS

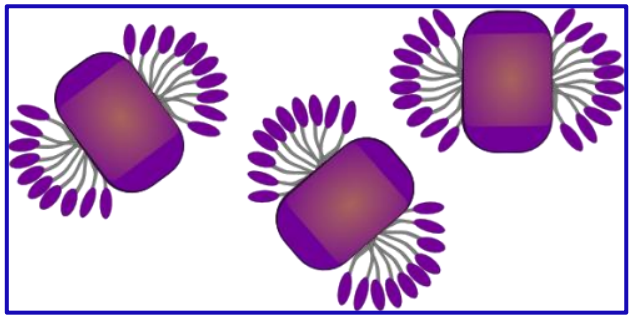




$$N_{Det} = \frac{\sqrt{\frac{I(0) M_{prot} N_A}{C_{prot} f^2}} - (n_{prot} N_A - \rho_S \bar{\nu}_{prot} M_{prot})}{(n_{det} N_A - \rho_S \bar{\nu}_{det} M_{det})}$$

N_{Det} = 225 ± 25 molecules per protein

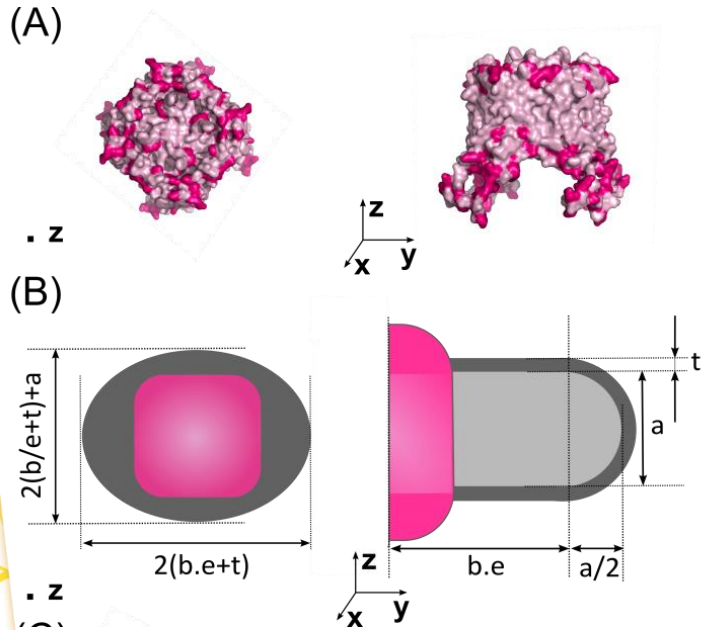
How to find a model which fits the curve ?



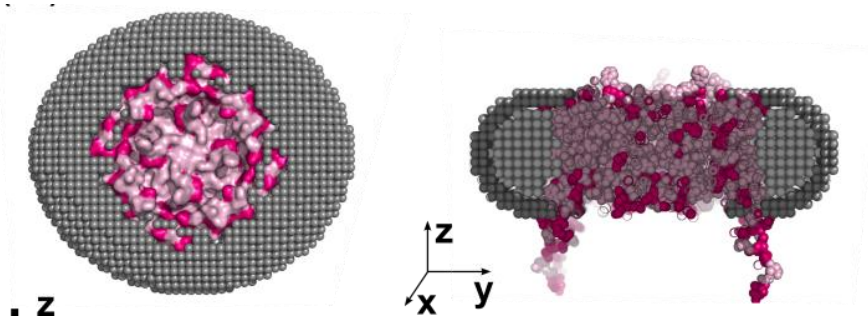
Several electronic densities : protein/detergent

→ ~~Simple Ab initio method~~

A parametrized torus with two electronic densities



The torus volume is filled with beads.
The SAXS curve is calculated with CRYSOL



Beads « atoms » and grid parameters chosen for Crysol input :

ρ_{tails} = 0.282 Å → Leu/CD2

ρ_{heads} = 0.520 Å → Lys/NZ

Lipfert et al. (2007), Phys.Chem.B, 111, 12427–12438

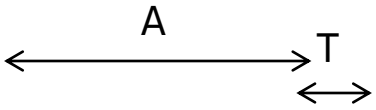
Validating the beads modeling

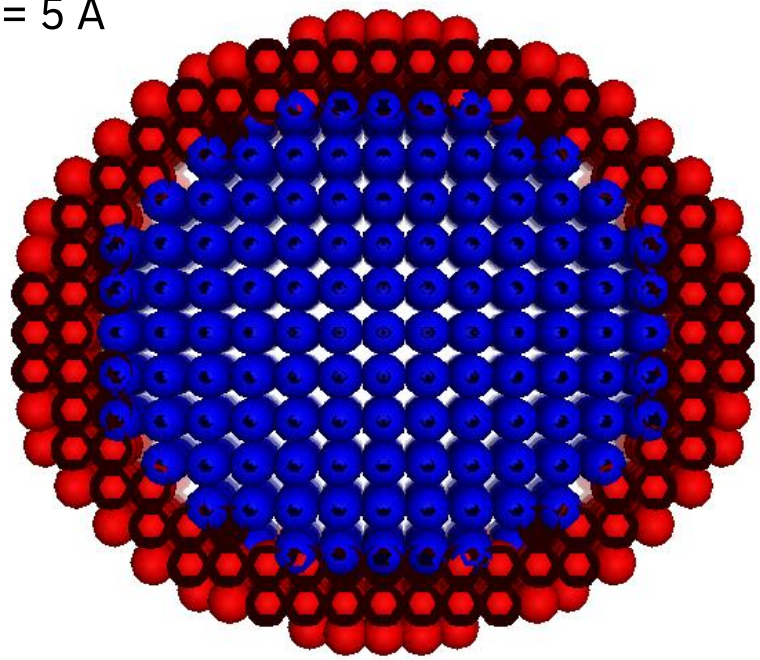
Core-shell ellipsoid

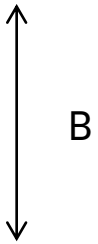
A = 22 Å

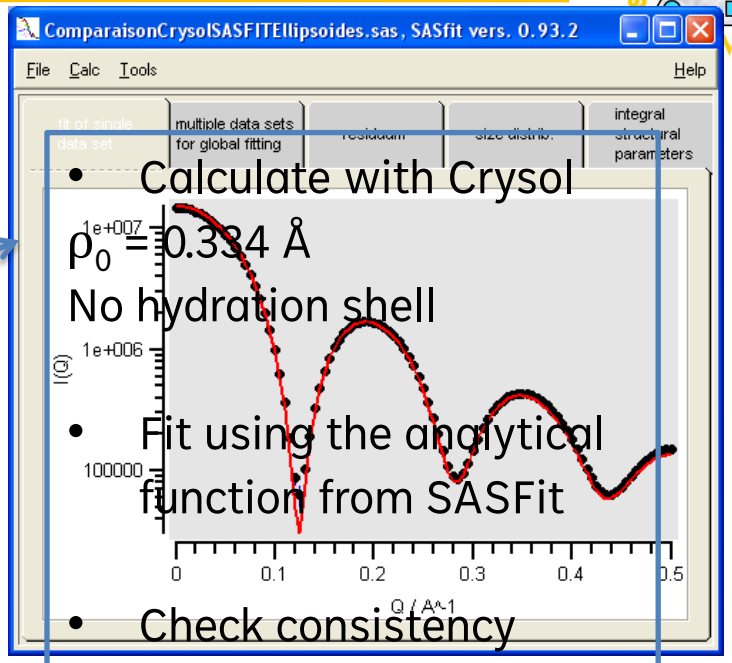
B = 18.2 Å

T = 5 Å







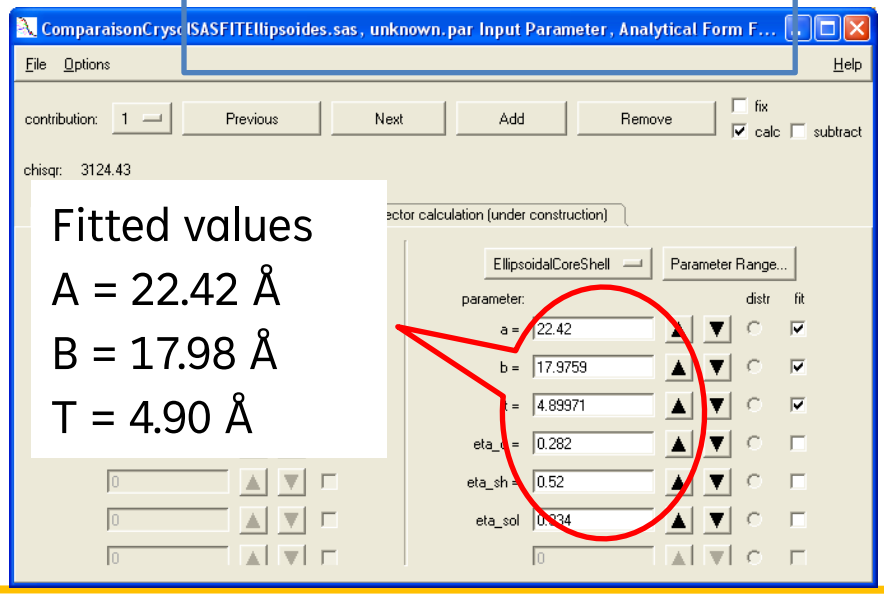


- Calculate with Crysol

ρ_0 = 0.334 Å

No hydration shell

- Fit using the analytical function from SASFit
- Check consistency



Fitted values

A = 22.42 Å

B = 17.98 Å

T = 4.90 Å

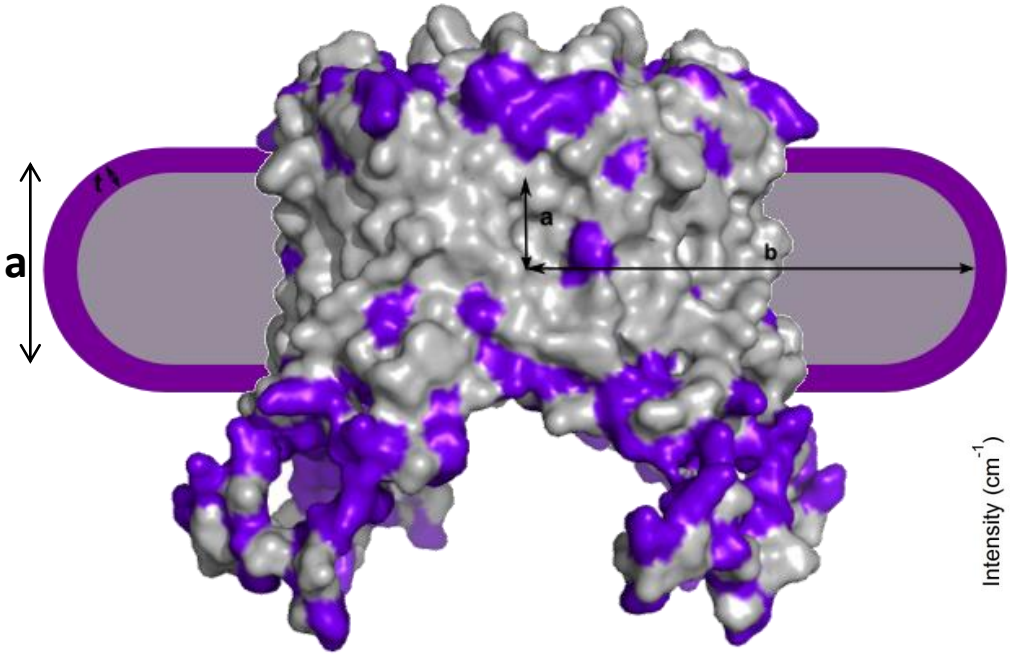
Beads « atoms » and grid parameters chosen for Crysol input :

ρ_{in} = 0.282 $e^-/Å^3$

ρ_{out} = 0.520 $e^-/Å^3$

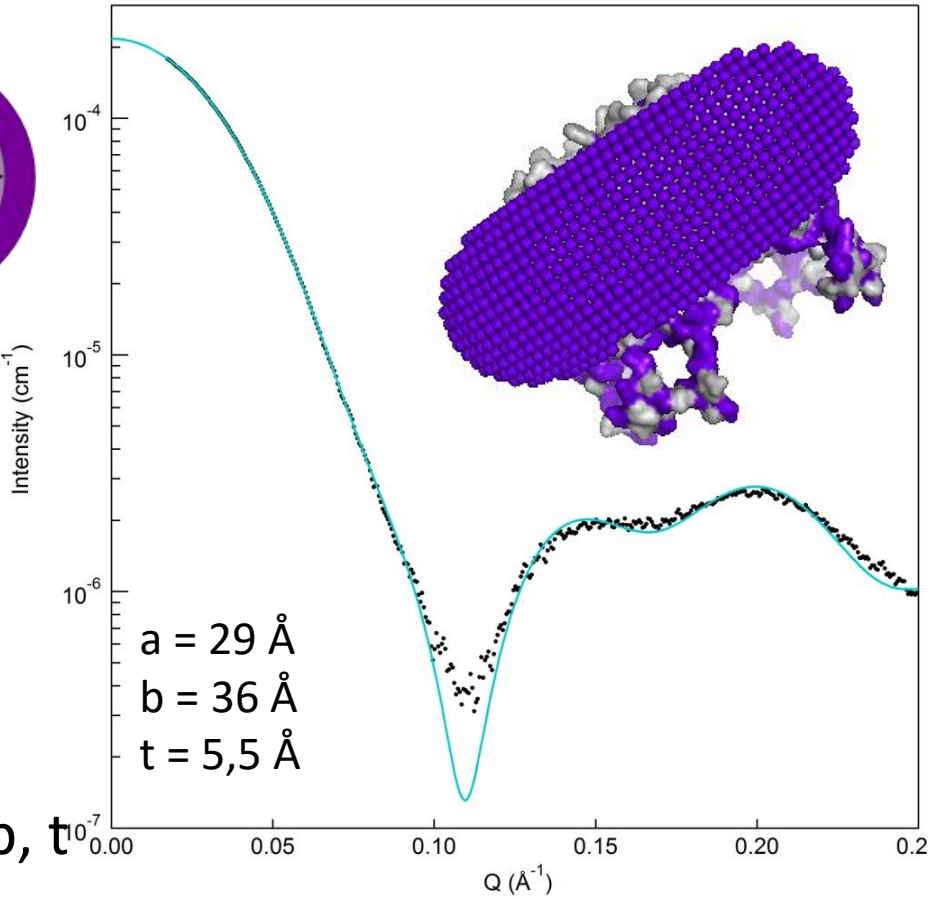
Validated

Circular detergent corona



Modelization of a circular torus of detergent

Three free geometric parameters, a, b, t

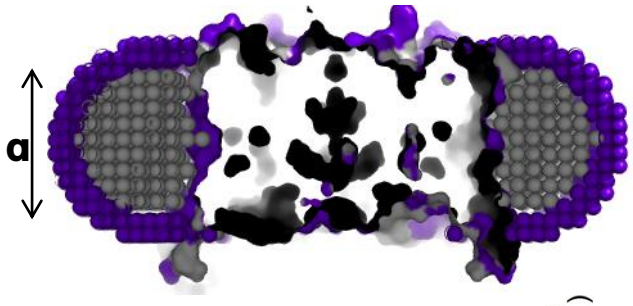


Elliptical detergent corona

Berthaud et al. (2012), *JACS*, 134 (24), 10080-10088

Addition of the ellipticity parameter : e

a = 30 Å
b = 35 Å
t = 5.5 Å
e = 1.12

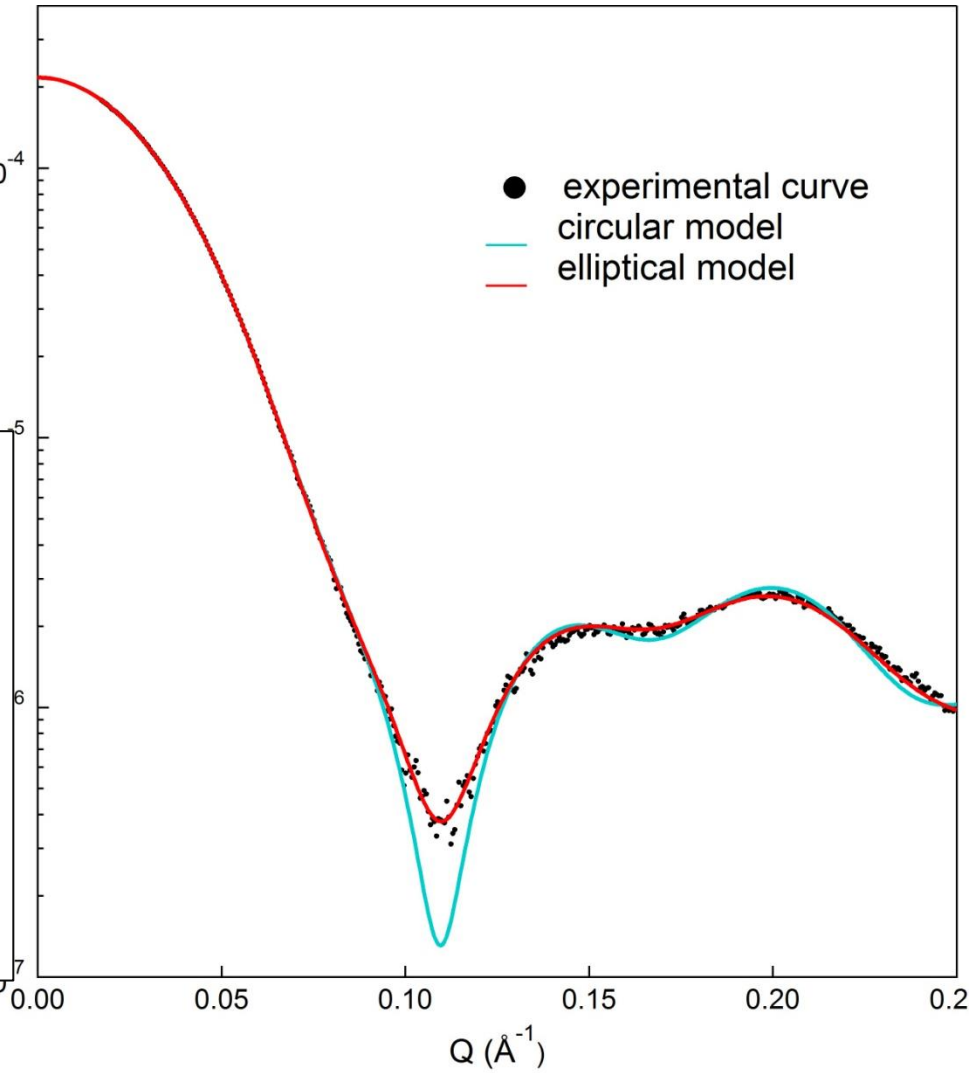


Relevancy of the model

Number of detergent molecules from the coarse grain model by volume's calculations:

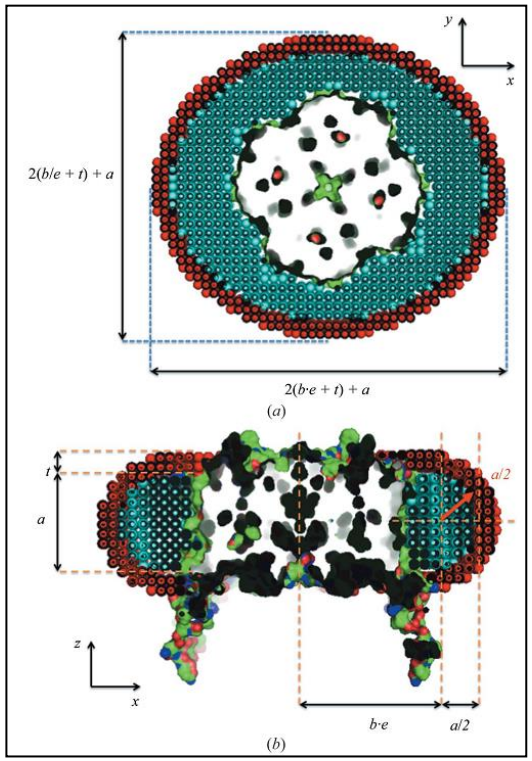
N = 270 ± 30 detergent molecules

⇒ Good agreement with previous values !



Memprot : a program to generate/optimize the corona geometrical model

Pérez J. & Koutsioubas, A. (2015), *Acta Cryst., D71, 86-93*



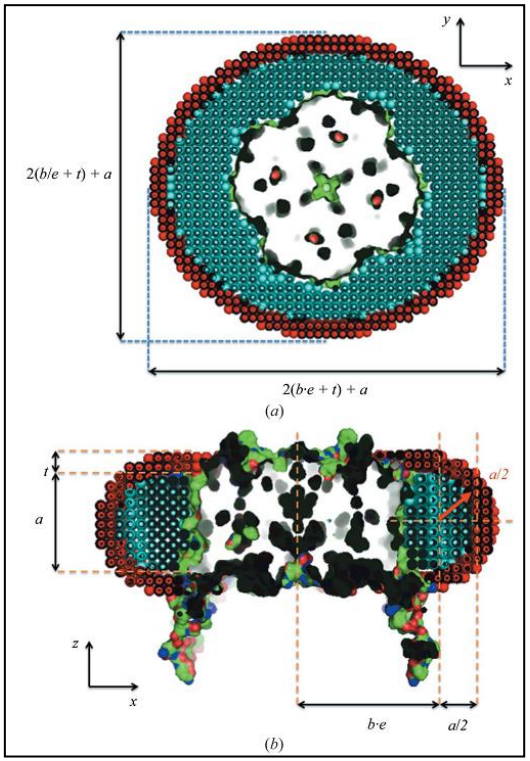

```
minimum_chi ← infinite
 for each a in the range [a_min,a_max], do
  for each b in the range [b_min,b_max], do
   for each t in the range [t_min,t_max], do
    for each e in the range [e_min,e_max], do
     for each phi in the range [phi_min,phi_max], do
      generate corona_model(a,b,t,e,phi)
      calculate chi (corona, protein pdb, experimental data) calling CRYSOL
      if minimum_chi > chi, then
       minimum_chi ← chi
 return chi
```

Algorithm of the *Memprot* program. The program essentially creates PDB files with the models made of the full-atom protein structure and the parameterized coarse-grained detergent corona, and *CRYSOL* is called to calculate the SAXS curves. An overall sorting on the χ value is performed to keep the best model.

$$\chi^2 = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{I_{exp}(q_i) - I_{calc}(q_i)}{\sigma(q_i)}\right]^2$$

Memprot : a program to generate/optimize the corona geometrical model

Pérez J. & Koutsioubas, A. (2015), *Acta Cryst., D71, 86-93*



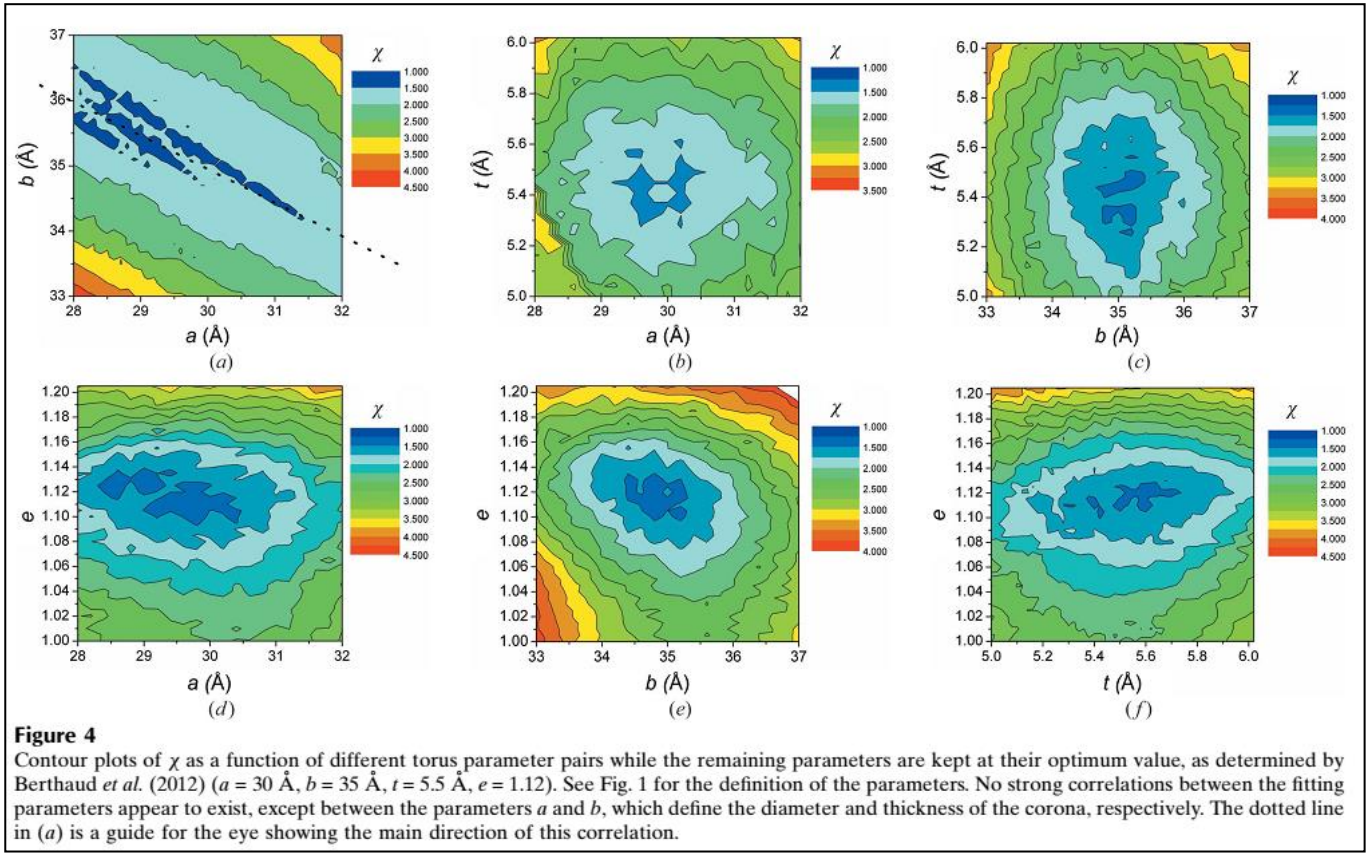
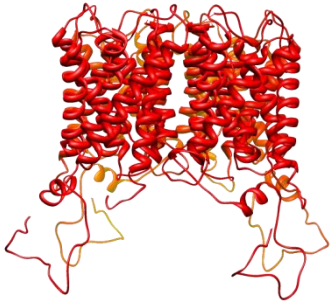


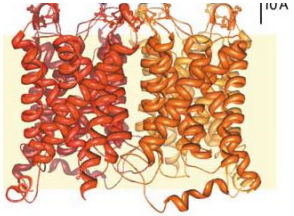
Figure 4
Contour plots of χ as a function of different torus parameter pairs while the remaining parameters are kept at their optimum value, as determined by Berthaud *et al.* (2012) (a = 30 Å, b = 35 Å, t = 5.5 Å, e = 1.12). See Fig. 1 for the definition of the parameters. No strong correlations between the fitting parameters appear to exist, except between the parameters a and b, which define the diameter and thickness of the corona, respectively. The dotted line in (*a*) is a guide for the eye showing the main direction of this correlation.

$$\chi^2 = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{I_{exp}(q_i) - I_{calc}(q_i)}{\sigma(q_i)}\right]^2$$

How far can we discriminate different protein constructs ?



Full Aqp-0
(2b6p)



Truncated Aqp-0
(2b6o)

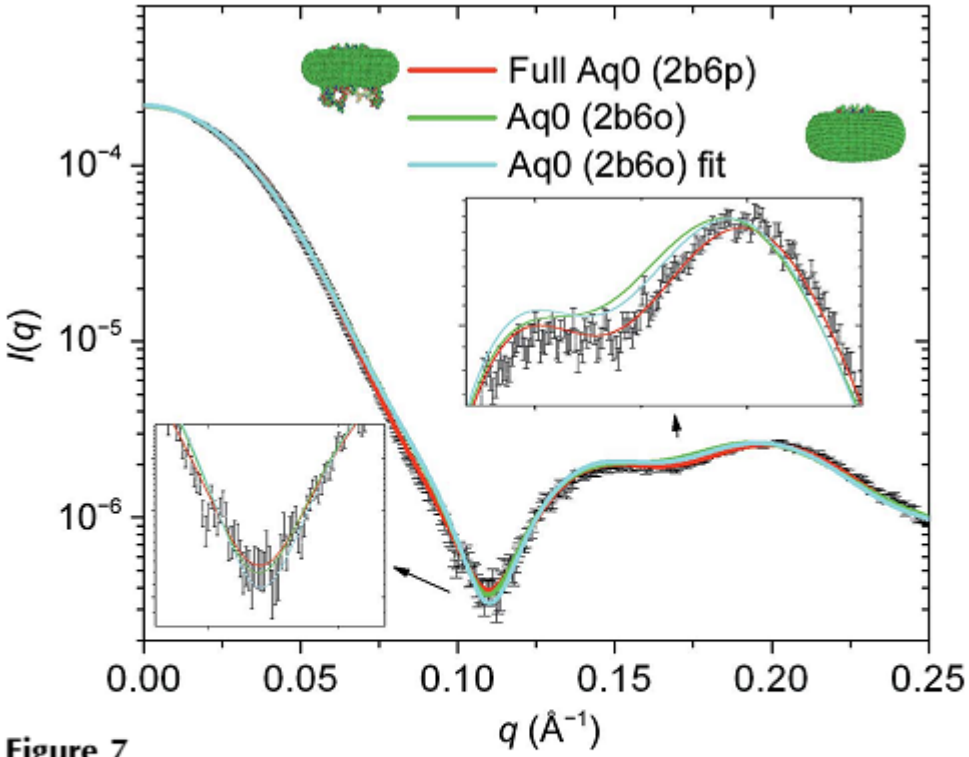


Figure 7
Scattering curves corresponding to corona parameters a = 29.6 Å, b = 35.4 Å, t = 5.6 Å, e = 1.12, e'_{heads} = 0.512 e Å^{-3}, e'_{tails} = 0.270 e Å^{-3}) for the full (2b6p) and truncated (2b6o) structures of aquaporin-0. The respective χ values are 1.31 and 3.79. The curve corresponding to an artificial optimized corona using the truncated form of aquaporin-0 is also plotted. The associated χ value is 3.47, which is still much higher than that for the complex based on the actual 2b6p structure.

Outline

- ✓ Short presentation of current available techniques at beamline SWING
- ✓ Memprot program: why and how

  AQP0 as the « Guinea pig » protein
- ✓ Example of a recent application using Memprot:

  MHST protein in 4 different detergents
- ✓ Dadimodo program: why and how
- ✓ Example of application coupling Memprot & Dadimodo:

  HasA-HasR protein with two different constructs
- ✓ Short conclusion: foreseen new functionalities for Memprot and Dadimodo

MhsT in different detergents

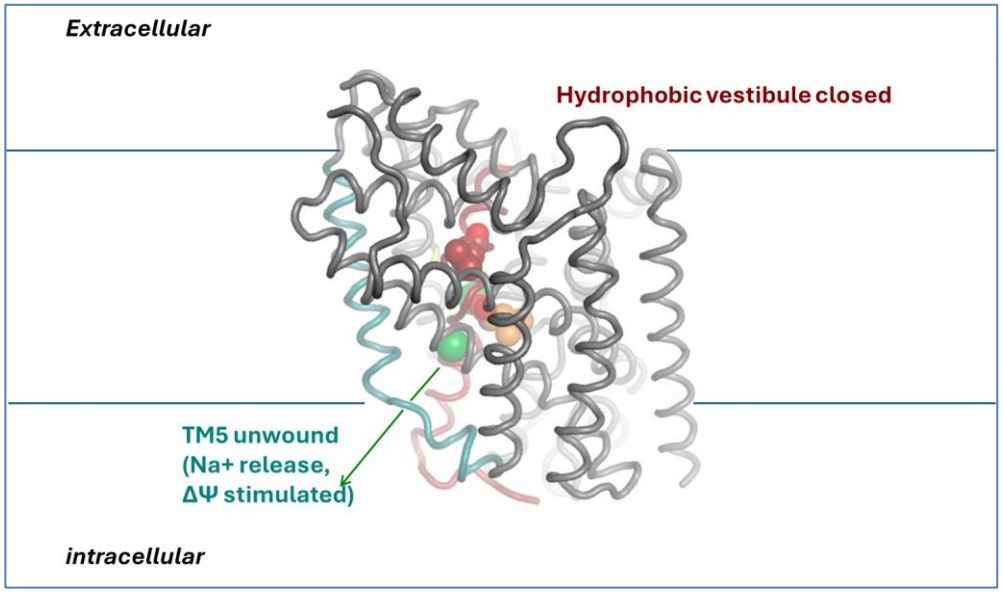
De Pol et al., submitted

B. halodurans multi-hydrophic amino acid transporter MhsT – bacterial SLC6 transporter

The **SLC6 family** of secondary active transporters:

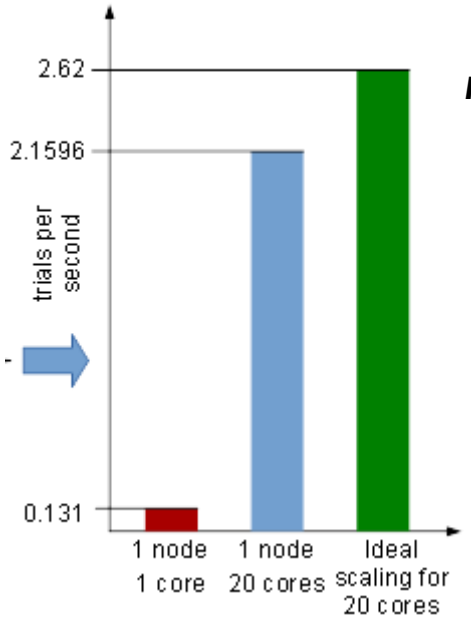
- **integral membrane** solute carrier proteins
- Na+-dependent **translocation of small amino acid** or amino acid-like substrates
- includes the serotonin, dopamine, norepinephrine, GABA, taurine, creatine
- associated with a number of human diseases and disorders making this family a **critical target for therapeutic development**
- several members **directly involved in the action of drugs of abuse such as cocaine, amphetamines, and ecstasy**.

(coll: Poul Nissen, Aarhus University, Denmark)

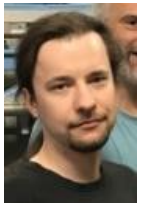


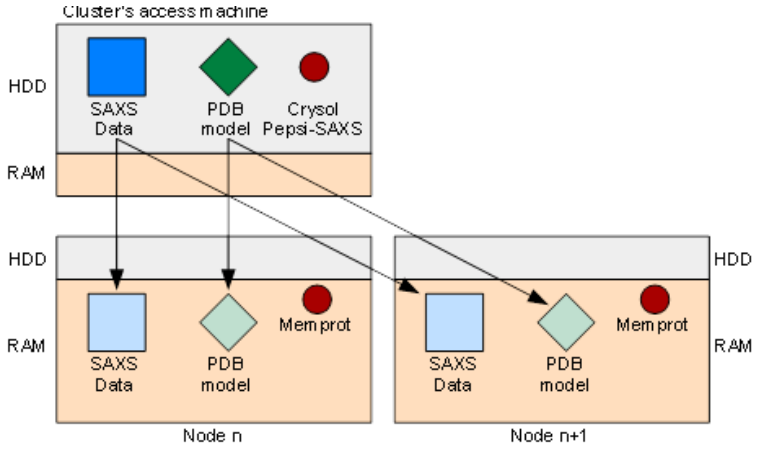
Parallelization of Memprot code

- Typical Memprot runs range from thousands to hundreds of thousands of trials – the speed of calculations and scalability is an important issue
- We have implemented MPI-based, data-driven parallelization in Memprot to benefit from HPC clusters (here SOLEIL HPC)
- 449 residue protein MHST (PDB id 4us3) was used as a test case (sample provided by collaborators and measured at SWING)
- To prevent saturation of the cluster's network, Memprot stores all i/o files (experimental data, protein's PDB model, intermediate files) locally in the node's RAM, utilising /dev/shm partition



M. Baranowski (2016-2018)

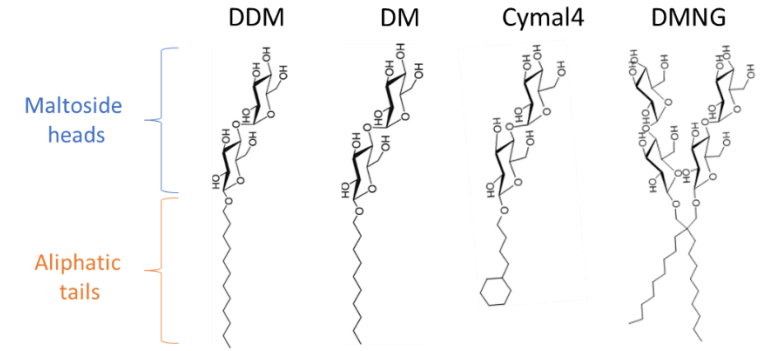


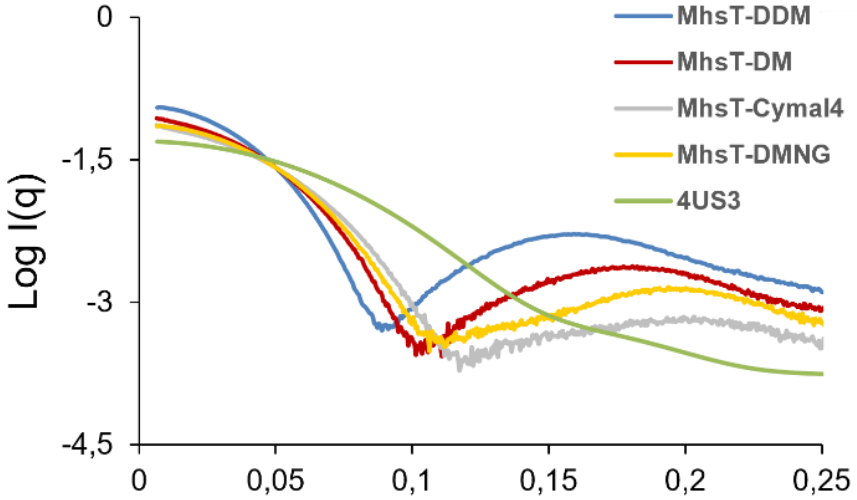


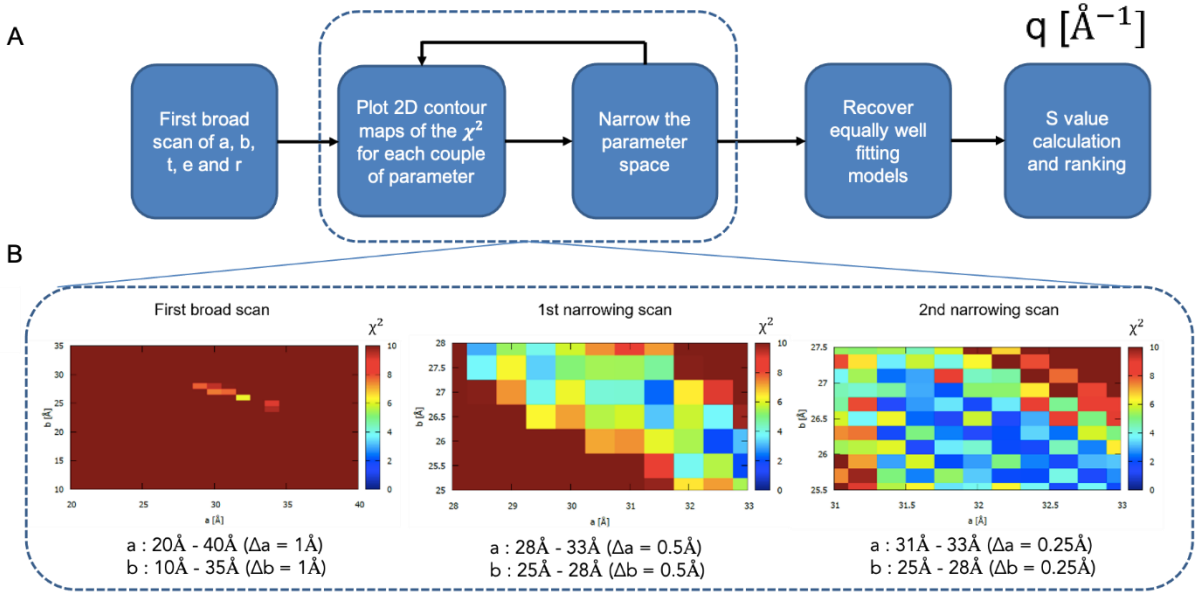
MhsT in different detergents

(coll: Poul Nissen, Aarhus University, Denmark)

De Pol et al., submitted







9 × 60,000 fits

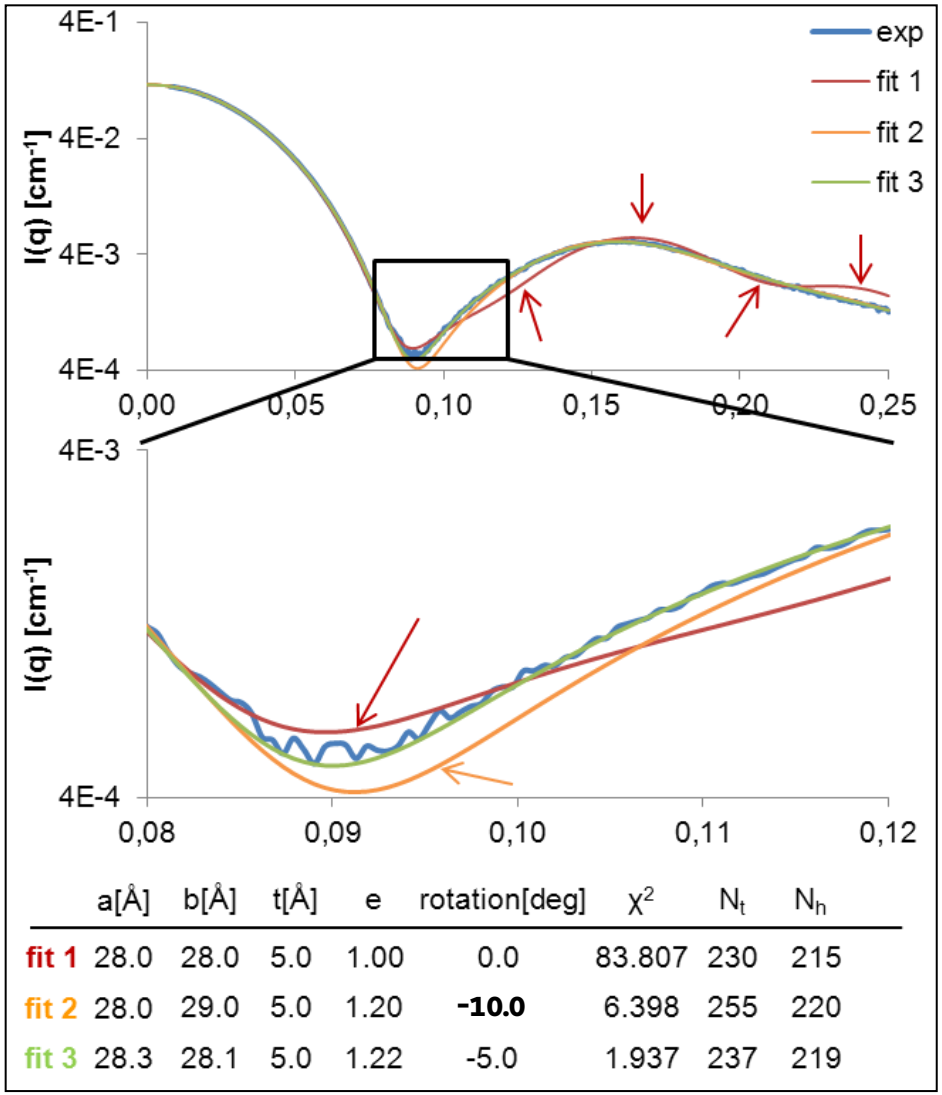
9 × 90,000 fits

9 × 200,000 fits

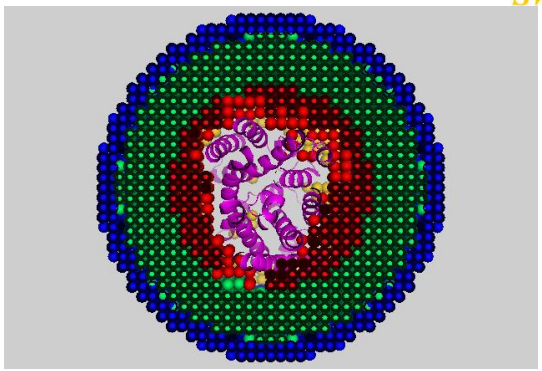
64 cores x 2.65 GHz
→ 1 min for 1000 models

Example of fitting improvement steps on MhsT protein

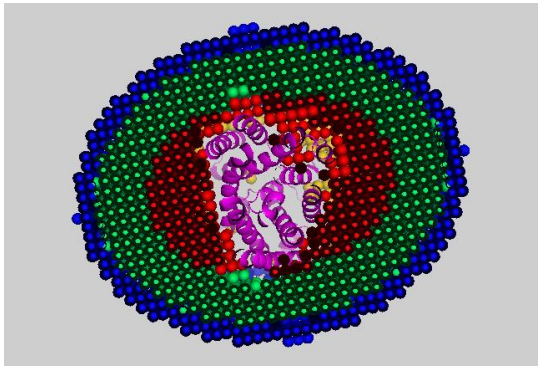
(coll: Poul Nissen, Aarhus University, Denmark)



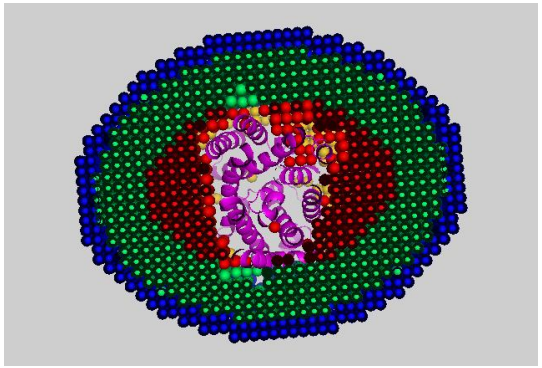
	a[Å]	b[Å]	t[Å]	e	rotation[deg]	χ^2	N_t	N_h
fit 1	28.0	28.0	5.0	1.00	0.0	83.807	230	215
fit 2	28.0	29.0	5.0	1.20	**-10.0**	6.398	255	220
fit 3	28.3	28.1	5.0	1.22	-5.0	1.937	237	219



Fit 1



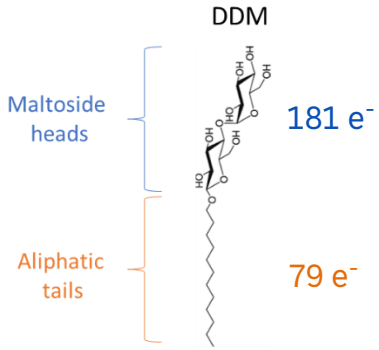
Fit 2

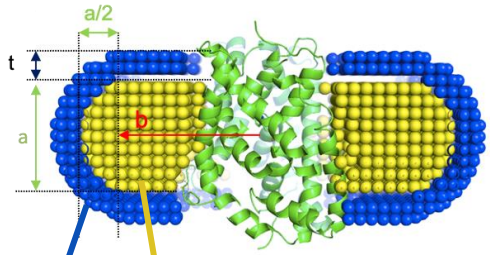


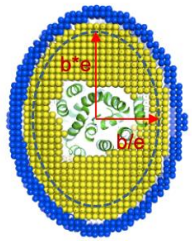
Fit 3

Keep only physically meaningful models

De Pol et al., submitted







From total number of yellow beads

→ Ne^-_{inner} = total # electrons of the hydrophobic area

→ $N_{tails} = Ne^-_{inner} / 79$

From total number of blue beads

→ Ne^-_{outer} = total # electrons of the hydophilic area

→ $N_{heads} = Ne^-_{outer} / 181$

In principle, N_{tails} should be equal to N_{heads} !

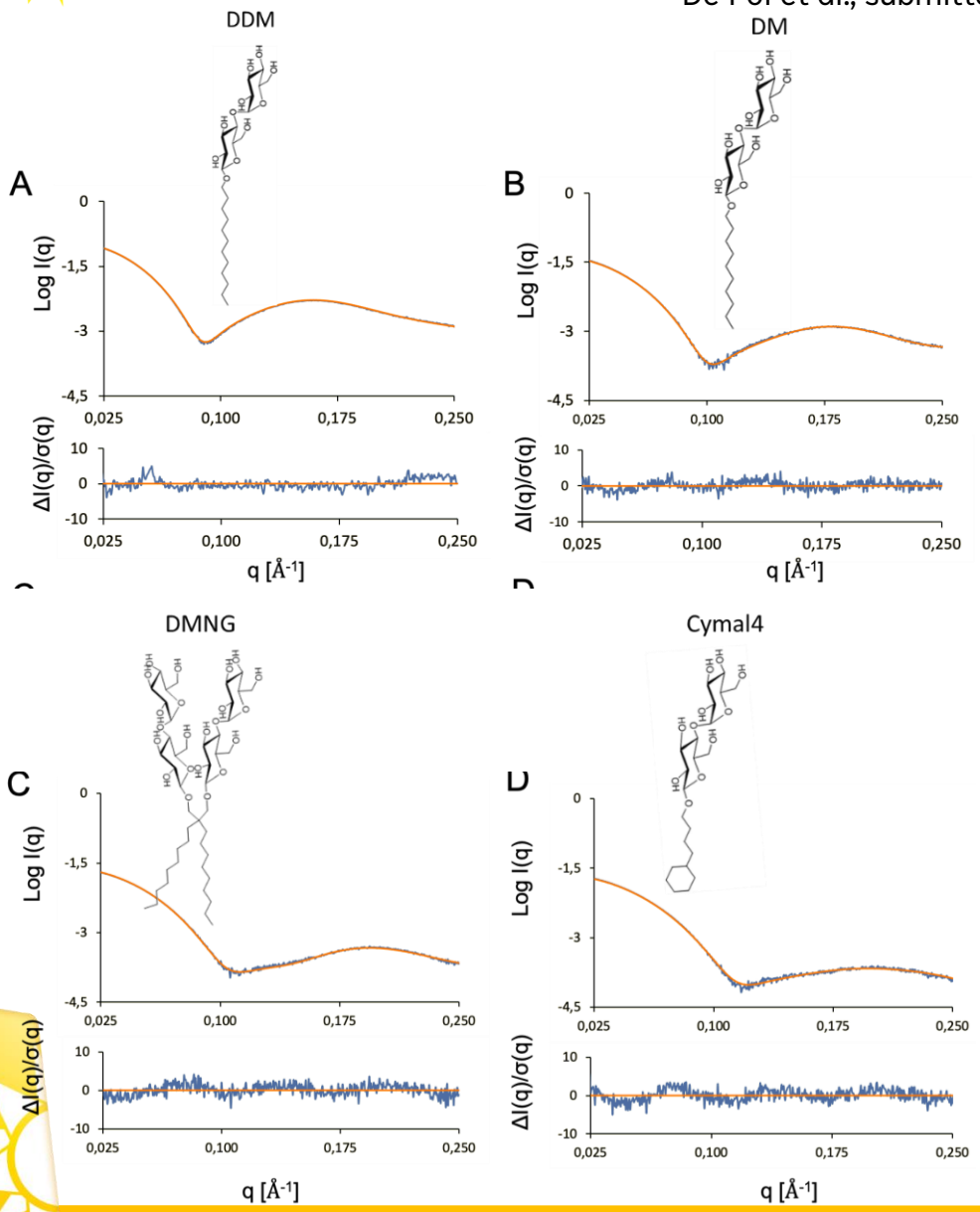
In practice, not always exactly the case.

A scoring function was designed to account for this constraint:

$$\mathbf{S} = (\mathbf{1} + |\mathbf{1} - \boldsymbol{\alpha}| + |\mathbf{1} - \boldsymbol{ToH}|) * \boldsymbol{\chi}^{\mathbf{2}}$$, where $ToH = N_{tails} / N_{heads}$

MhsT in different detergents

De Pol et al., submitted



Rank	a [Å]	b [Å]	t [Å]	e	r [°]	ρ_T [e⁻ / Å³]	ρ_H [e⁻ / Å³]	χ^2	α	#heads	#tails	ToH	S value
1	30.3	26.2	5.7	1.25	0	0.541	0.281	2.046	1.04	210	210	1.000	2.128
2	30.3	26.2	5.6	1.25	180	0.541	0.281	2.052	1.04	210	210	1.000	2.134
3	30.4	26.1	5.8	1.25	0	0.536	0.278	2.069	1.03	211	210	1.005	2.141
4	30.4	26.1	5.8	1.25	180	0.536	0.278	2.081	1.03	211	210	1.005	2.154
5	30.1	26.3	5.9	1.24	0	0.541	0.281	2.106	1.04	211	210	1.005	2.201
6	30.1	26.3	5.7	1.24	180	0.541	0.281	2.119	1.04	211	210	1.005	2.214
7	30.3	26.2	5.7	1.24	0	0.536	0.278	2.164	1.03	212	210	1.010	2.251
8	30.0	26.3	5.8	1.25	0	0.536	0.278	2.160	1.03	211	208	1.014	2.255
9	30.1	26.3	5.9	1.25	0	0.541	0.281	2.154	1.04	211	209	1.010	2.262
10	30.7	25.9	6.0	1.26	0	0.536	0.278	2.197	1.03	210	210	1.000	2.263

Rank	a [Å]	b [Å]	t [Å]	e	r [°]	ρ_T [e⁻ / Å³]	ρ_H [e⁻ / Å³]	χ^2	α	#heads	#tails	ToH	S value
1	33.5	21.0	6.3	1.28	120	0.515	0.288	1.700	1.03	189	182	1.038	1.816
2	32.0	20.5	5.8	1.28	120	0.530	0.275	1.759	1.02	176	169	1.041	1.866
3	28.5	22.5	6.0	1.22	30	0.530	0.275	1.821	1.02	169	172	0.983	1.888
4	31.5	21.5	6.1	1.28	30	0.536	0.288	1.818	1.03	191	187	1.021	1.911
5	32.5	20.0	5.8	1.28	120	0.530	0.275	1.772	1.02	173	162	1.068	1.928
6	31.5	21.5	6.6	1.31	120	0.505	0.273	1.900	1.01	196	199	0.985	1.948
7	35.0	21.1	6.1	1.37	120	0.536	0.288	1.900	1.03	188	188	1.000	1.957
8	32.5	20.5	6.6	1.31	120	0.510	0.275	1.839	1.02	187	199	0.940	1.986
9	31.5	21.0	5.9	1.25	30	0.541	0.291	1.859	1.04	178	173	1.029	1.987
10	32.0	20.5	6.6	1.31	120	0.510	0.275	1.862	1.02	178	191	0.932	2.026

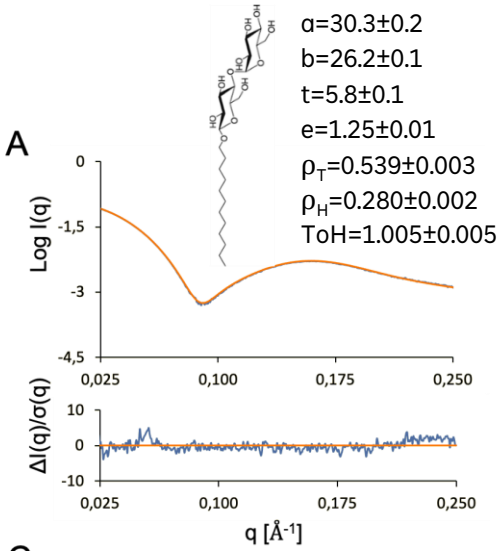
Rank	a [Å]	b [Å]	t [Å]	e	r [°]	ρ_T [e⁻ / Å³]	ρ_H [e⁻ / Å³]	χ^2	α	#heads	#tails	ToH	S value
1	31	19	5.6	1.35	90	0.524	0.281	2.748	0.97	149	151	0.987	2.866
2	30	20	5.6	1.28	90	0.510	0.284	2.829	0.98	142	142	1.000	2.886
3	31	19	5.9	1.38	90	0.508	0.263	2.618	0.94	152	159	0.956	2.890
4	32	19	5.8	1.41	90	0.513	0.276	2.739	0.95	166	163	1.018	2.925
5	31	19	5.8	1.38	90	0.518	0.278	2.762	0.96	150	153	0.98	2.928
6	31	19	5.9	1.38	90	0.502	0.251	2.677	0.93	156	160	0.975	2.931
7	31	19	5.8	1.35	90	0.508	0.263	2.686	0.94	152	157	0.968	2.933
8	30	20	5.6	1.35	90	0.518	0.278	2.816	0.96	156	157	0.994	2.946
9	31	20	5.7	1.28	60	0.510	0.284	2.893	0.98	142	142	1.000	2.951
10	31	19	5.6	1.38	90	0.518	0.278	2.835	0.96	149	148	1.007	2.968

Rank	a [Å]	b [Å]	t [Å]	e	r [°]	ρ_T [e⁻ / Å³]	ρ_H [e⁻ / Å³]	χ^2	α	#heads	#tails	ToH	S value
1	32.2	19.5	6.2	1.28	90	0.536	0.299	2.388	1.03	84	86	0.977	2.515
2	32.4	19.6	6.1	1.28	90	0.541	0.302	2.465	1.04	84	84	1.000	2.564
3	32.0	19.8	6.3	1.28	90	0.541	0.302	2.369	1.04	84	88	0.955	2.570
4	31.9	19.1	6.4	1.28	90	0.536	0.299	2.403	1.03	84	90	0.933	2.636
5	31.3	20.5	6.1	1.25	90	0.541	0.302	2.534	1.04	81	83	0.976	2.696
6	31.2	19.0	6.9	1.30	90	0.520	0.270	2.331	1.00	82	98	0.837	2.711
7	31.7	19.0	6.9	1.30	90	0.520	0.270	2.331	1.00	82	98	0.837	2.711
8	32.0	20.5	6.0	1.25	90	0.541	0.302	2.643	1.04	81	81	1.000	2.749
9	32.0	19.1	6.0	1.28	90	0.536	0.299	2.656	1.03	84	83	1.012	2.768
10	31.5	22.0	6.3	1.28	90	0.536	0.299	2.636	1.03	87	89	0.978	2.773

MhsT in different detergents

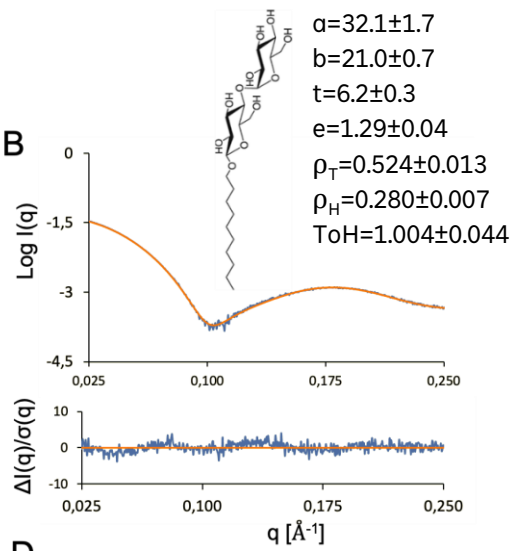
De Pol et al., submitted

DDM



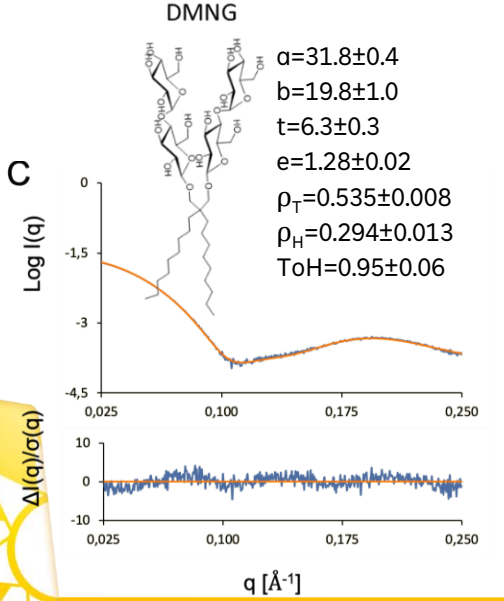
α=30.3±0.2
b=26.2±0.1
t=5.8±0.1
e=1.25±0.01
ρ_T=0.539±0.003
ρ_H=0.280±0.002
ToH=1.005±0.005

DM



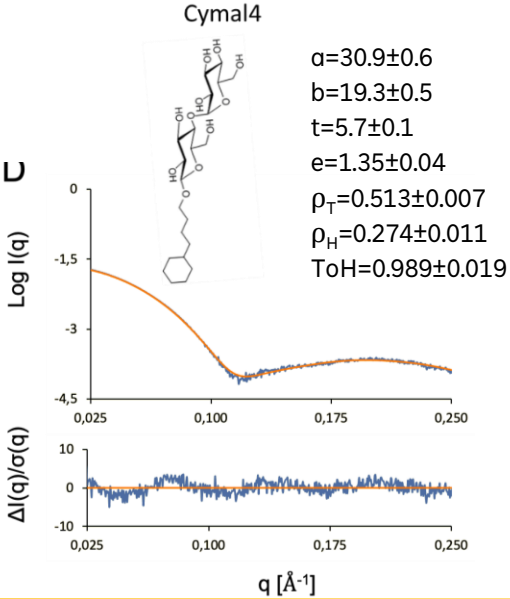
α=32.1±1.7
b=21.0±0.7
t=6.2±0.3
e=1.29±0.04
ρ_T=0.524±0.013
ρ_H=0.280±0.007
ToH=1.004±0.044

DMNG



α=31.8±0.4
b=19.8±1.0
t=6.3±0.3
e=1.28±0.02
ρ_T=0.535±0.008
ρ_H=0.294±0.013
ToH=0.95±0.06

Cymal4



α=30.9±0.6
b=19.3±0.5
t=5.7±0.1
e=1.35±0.04
ρ_T=0.513±0.007
ρ_H=0.274±0.011
ToH=0.989±0.019

Rank	a [Å]	b [Å]	t [Å]	e	r [°]	ρ_T [e⁻/Å³]	ρ_H [e/Å³]	χ^2	α	#heads	#tails	ToH	S value
1	30.3	26.2	5.7	1.25	0	0.541	0.281	2.046	1.04	210	210	1.000	2.128
2	30.3	26.2	5.6	1.25	180	0.541	0.281	2.052	1.04	210	210	1.000	2.134
3	30.4	26.1	5.8	1.25	0	0.536	0.278	2.069	1.03	211	210	1.005	2.141
4	30.4	26.1	5.8	1.25	180	0.536	0.278	2.081	1.03	211	210	1.005	2.154
5	30.1	26.3	5.9	1.24	0	0.541	0.281	2.106	1.04	211	210	1.005	2.201
6	30.1	26.3	5.7	1.24	180	0.541	0.281	2.119	1.04	211	210	1.005	2.214
7	30.3	26.2	5.7	1.24	0	0.536	0.278	2.164	1.03	212	210	1.010	2.251
8	30.0	26.3	5.8	1.25	0	0.536	0.278	2.160	1.03	211	208	1.014	2.255
9	30.1	26.3	5.9	1.25	0	0.541	0.281	2.154	1.04	211	209	1.010	2.262
10	30.7	25.9	6.0	1.26	0	0.536	0.278	2.197	1.03	210	210	1.000	2.263

Rank	a [Å]	b [Å]	t [Å]	e	r [°]	ρ_T [e⁻/Å³]	ρ_H [e/Å³]	χ^2	α	#heads	#tails	ToH	S value
1	33.5	21.0	6.3	1.28	120	0.515	0.288	1.700	1.03	189	182	1.038	1.816
2	32.0	20.5	5.8	1.28	120	0.530	0.275	1.759	1.02	176	169	1.041	1.866
3	28.5	22.5	6.0	1.22	30	0.530	0.275	1.821	1.02	169	172	0.983	1.888
4	31.5	21.5	6.1	1.28	30	0.536	0.288	1.818	1.03	191	187	1.021	1.911
5	32.5	20.0	5.8	1.28	120	0.530	0.275	1.772	1.02	173	162	1.068	1.928
6	31.5	21.5	6.6	1.31	120	0.505	0.273	1.900	1.01	196	199	0.985	1.948
7	35.0	21.1	6.1	1.37	120	0.536	0.288	1.900	1.03	188	188	1.000	1.957
8	32.5	20.5	6.6	1.31	120	0.510	0.275	1.839	1.02	187	199	0.940	1.986
9	31.5	21.0	5.9	1.25	30	0.541	0.291	1.859	1.04	178	173	1.029	1.987
10	32.0	20.5	6.6	1.31	120	0.510	0.275	1.862	1.02	178	191	0.932	2.026

Rank	a [Å]	b [Å]	t [Å]	e	r [°]	ρ_T [e⁻/Å³]	ρ_H [e/Å³]	χ^2	α	#heads	#tails	ToH	S value
1	31	19	5.6	1.35	90	0.524	0.281	2.748	0.97	149	151	0.987	2.866
2	30	20	5.6	1.28	90	0.510	0.284	2.829	0.98	142	142	1.000	2.886
3	31	19	5.9	1.38	90	0.508	0.263	2.618	0.94	152	159	0.956	2.890
4	32	19	5.8	1.41	90	0.513	0.276	2.739	0.95	166	163	1.018	2.925
5	31	19	5.8	1.38	90	0.518	0.278	2.762	0.96	150	153	0.98	2.928
6	31	19	5.9	1.38	90	0.502	0.251	2.677	0.93	156	160	0.975	2.931
7	31	19	5.8	1.35	90	0.508	0.263	2.686	0.94	152	157	0.968	2.933
8	30	20	5.6	1.35	90	0.518	0.278	2.816	0.96	156	157	0.994	2.946
9	31	20	5.7	1.28	60	0.510	0.284	2.893	0.98	142	142	1.000	2.951
10	31	19	5.6	1.38	90	0.518	0.278	2.835	0.96	149	148	1.007	2.968

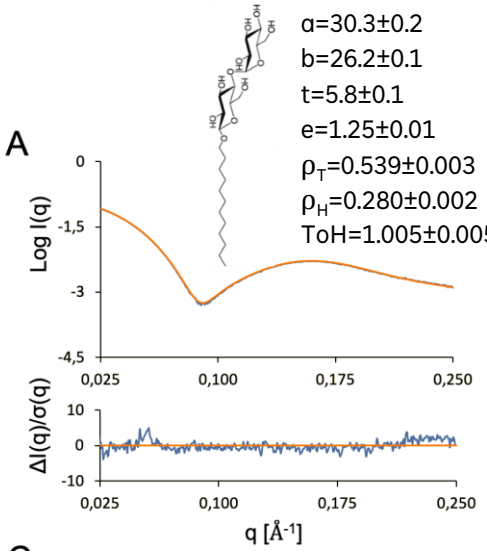
Rank	a [Å]	b [Å]	t [Å]	e	r [°]	ρ_T [e⁻/Å³]	ρ_H [e/Å³]	χ^2	α	#heads	#tails	ToH	S value
1	32.2	19.5	6.2	1.28	90	0.536	0.299	2.388	1.03	84	86	0.977	2.515
2	32.4	19.6	6.1	1.28	90	0.541	0.302	2.465	1.04	84	84	1.000	2.564
3	32.0	19.8	6.3	1.28	90	0.541	0.302	2.369	1.04	84	88	0.955	2.570
4	31.9	19.1	6.4	1.28	90	0.536	0.299	2.403	1.03	84	90	0.933	2.636
5	31.3	20.5	6.1	1.25	90	0.541	0.302	2.534	1.04	81	83	0.976	2.696
6	31.2	19.0	6.9	1.30	90	0.520	0.270	2.331	1.00	82	98	0.837	2.711
7	31.7	19.0	6.9	1.30	90	0.520	0.270	2.331	1.00	82	98	0.837	2.711
8	32.0	20.5	6.0	1.25	90	0.541	0.302	2.643	1.04	81	81	1.000	2.749
9	32.0	19.1	6.0	1.28	90	0.536	0.299	2.656	1.03	84	83	1.012	2.768
10	31.5	22.0	6.3	1.28	90	0.536	0.299	2.636	1.03	87	89	0.978	2.773

MhsT in different detergents

De Pol et al., submitted

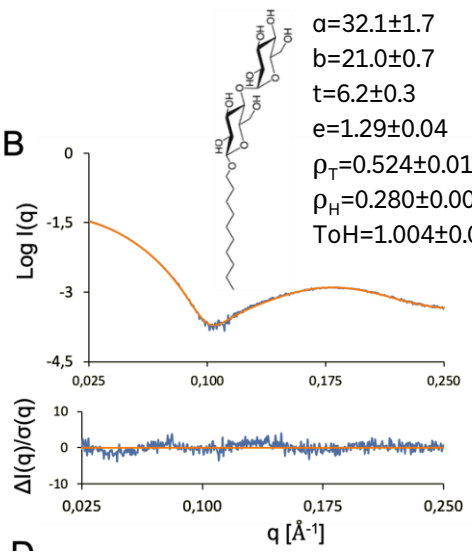
DDM

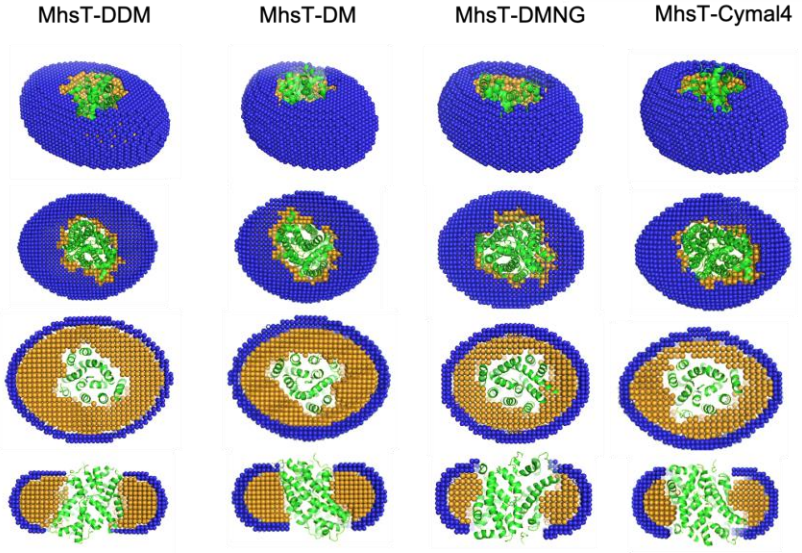
α=30.3±0.2
b=26.2±0.1
t=5.8±0.1
e=1.25±0.01
ρ_T=0.539±0.003
ρ_H=0.280±0.002
ToH=1.005±0.005



DM

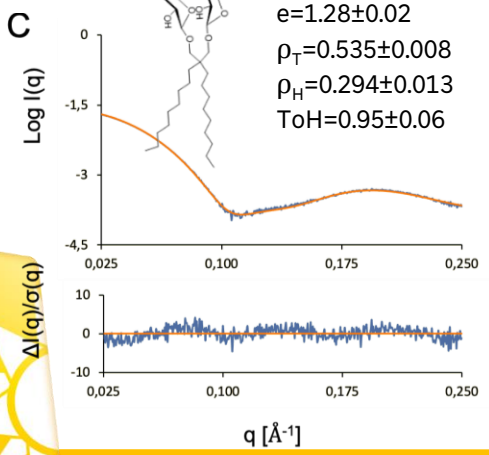
α=32.1±1.7
b=21.0±0.7
t=6.2±0.3
e=1.29±0.04
ρ_T=0.524±0.013
ρ_H=0.280±0.007
ToH=1.004±0.044





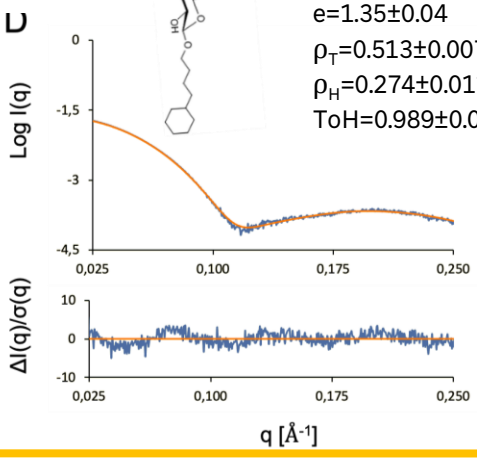
DMNG

α=31.8±0.4
b=19.8±1.0
t=6.3±0.3
e=1.28±0.02
ρ_T=0.535±0.008
ρ_H=0.294±0.013
ToH=0.95±0.06

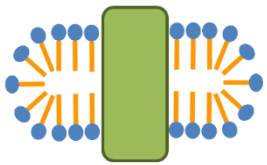


Cymal4

α=30.9±0.6
b=19.3±0.5
t=5.7±0.1
e=1.35±0.04
ρ_T=0.513±0.007
ρ_H=0.274±0.011
ToH=0.989±0.019



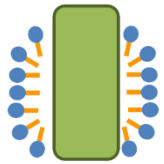
Case of MhsT-DDM, MhsT-DM



Can be well modelled by a flat hemi-torus

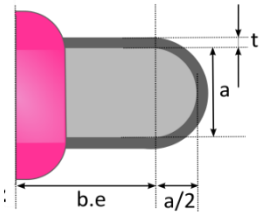
VS

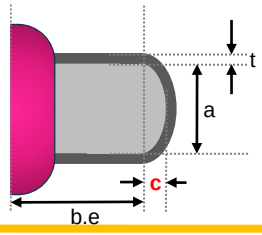
Case of MhsT-Cymal-4 and MhsT-DMNG



Can be only approximated by a flat hemi-torus

A “bicelle-like” model should be implemented in the future





Outline

- ✓ Short presentation of current available techniques at beamline SWING
- ✓ Memprot program: why and how

  AQP0 as the « Guinea pig » protein
- ✓ Example of a recent application using Memprot:

  MHST protein in 4 different detergents
- ✓ Dadimodo program: why and how
- ✓ Example of application coupling Memprot & Dadimodo:

  HasA-HasR protein with two different constructs
- ✓ Short conclusion: foreseen new functionalities for Memprot and Dadimodo

DADIMODO : Data-Driven Modules Docking

Collab : Christina Sizun & François Bontems (ICSN, Gif sur Yvette))

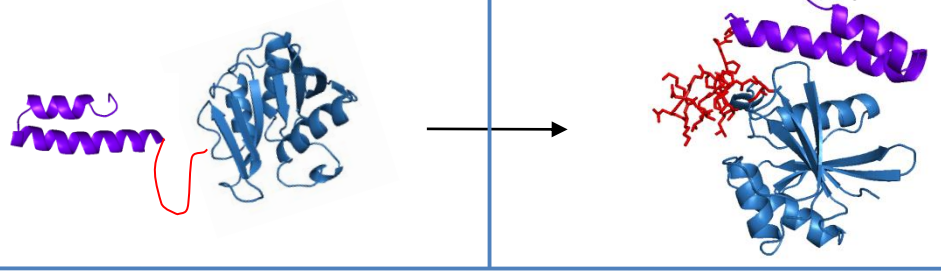
F. Mareuil, et al. (2007) *Eur Biophys J.*

Evrard et al. (2011), *J. Appl. Cryst.*

Modelling approach : complete atomic model

Full structure initiated with :

- Crystal or NMR domain structures
- Homology models



Prior knowledge:

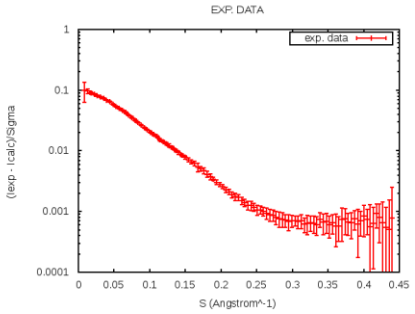
- Sequence
- Sub-parts moved as rigid-bodies (user-defined)
- A correct stereochemistry is maintained at all steps by minimizing energy

Experimental data:

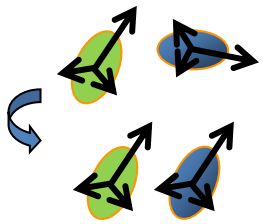
- SAXS
- NMR

  RDC

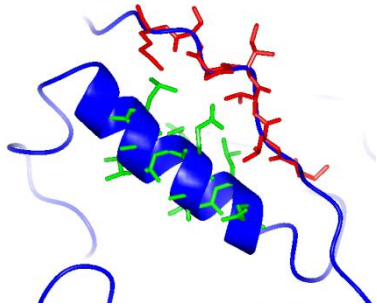
  ADR (chem. shift map.)



SAXS score



RDC score



ADR score

Optimisation of the **all-atom** structure *via* a genetic algorithm

Dadimodo Implementation

- Initial (slow) version : Evrard et al. (2011), *J. Appl. Cryst.*, 44:1264-1271.

- Current (faster) version : O. Roudenko , A. Thureau, J. Pérez (2019), *GECCO '19, ACM*, NY, USA, 401-402.

  - Parallel implementation of the genetic algorithm
    - 7300 Atoms → 7 hours on a 20 processor node (200 generations)

  - User-friendly input
    - Tools for completion of pdb input files (if needed)
    - User-defined topology : Pdb file + rigid bodies definitions

  - Web server since end 2018
    - Accessible to external users (after login in Soleil DB)
    - Five independent runs launched in parallel

Before launching Dadimodo: PDB generation step

A. Sali & T.L. Blundell. Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234, 779-815, 1993.

Shell script launching « Modeller »

A script is available on Swing Website
https://www.synchrotron-soleil.fr/en/beamlines/swing

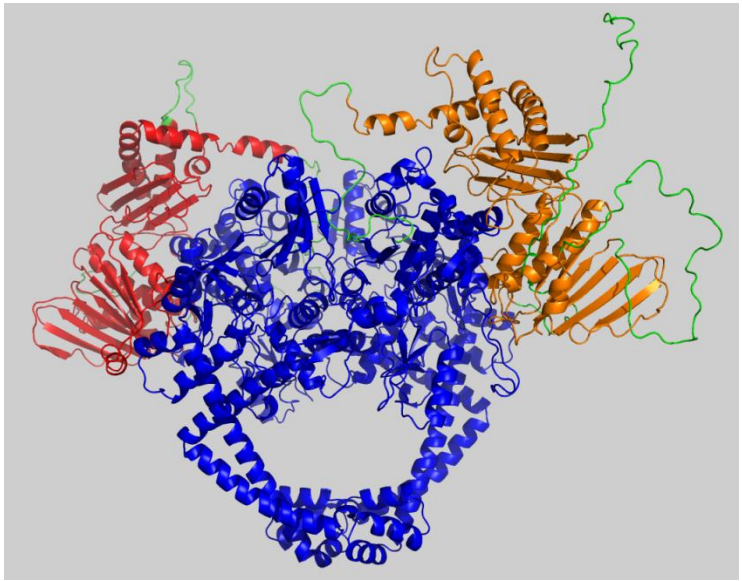
Original PDB files
missing atoms & residues

FASTA sequence of the entire protein

Generates

- missing atoms coordinates
- missing residues in linkers and tails
- aleatory orientations for flexibly connected domains

Complete PDB file
Interpretable by MMTK, the MD python library used in Dadimodo

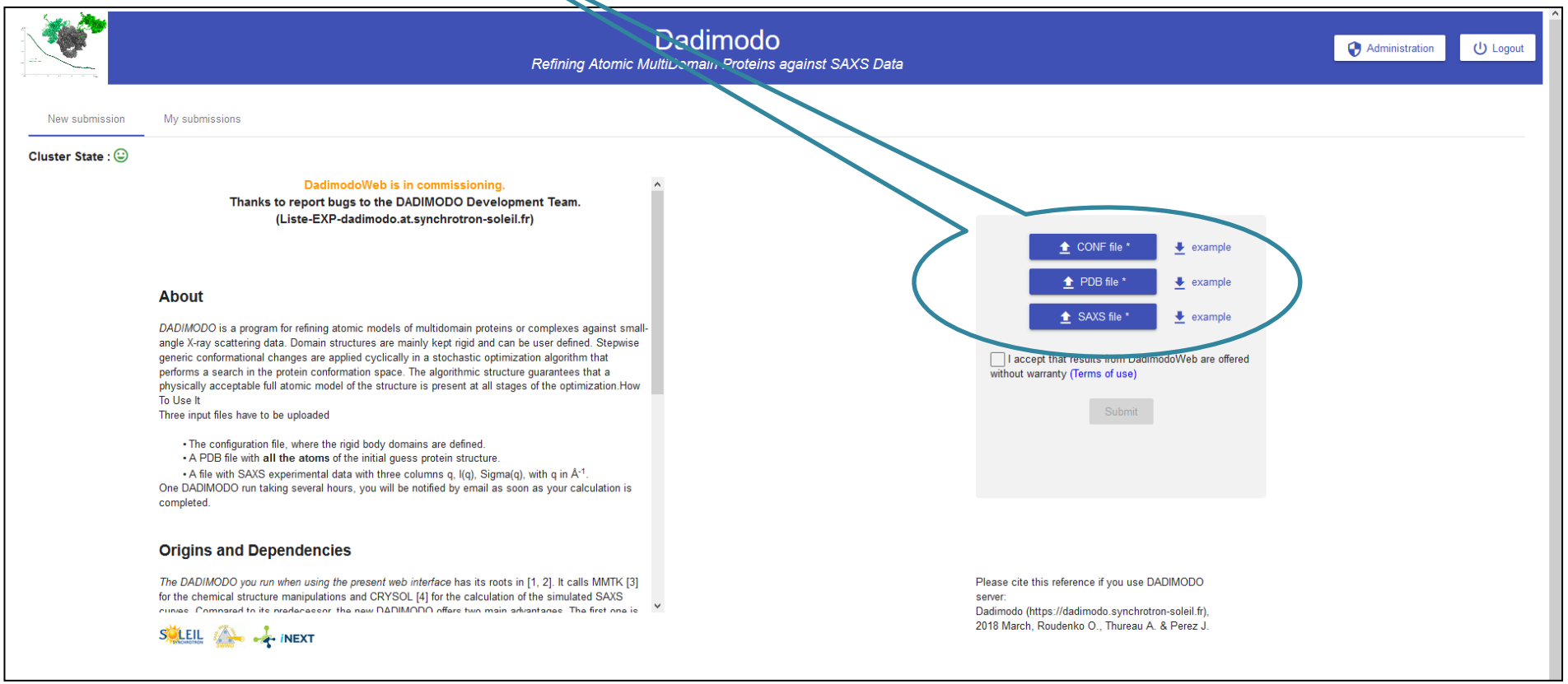


Mycobacterium tuberculosis DNA Gyrase (PDB 6GAV)

Dadimodo Web Server

https://dadimodo.synchrotron-soleil.fr

3 input files needed to launch Dadimodo on the Web Server

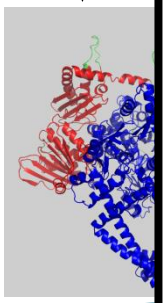


Dadimodo Web Server

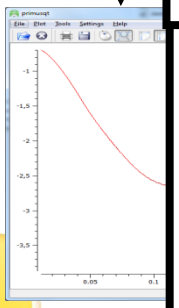
https://dadimodo.synchrotron-soleil.fr

3 input files needed to launch Dadimodo on the Web Server

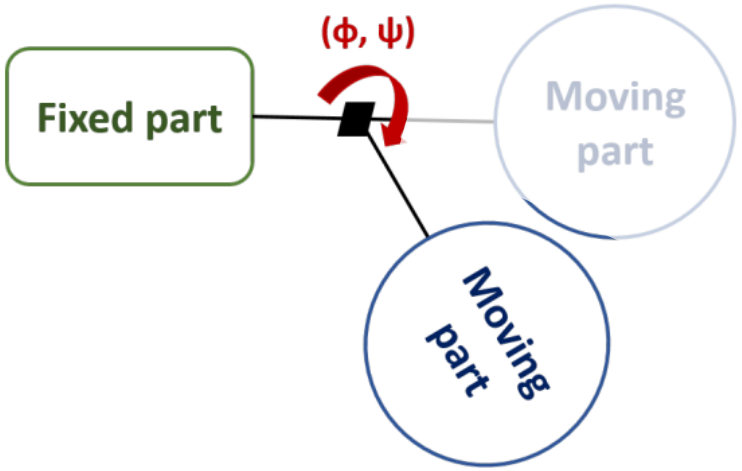
Complet



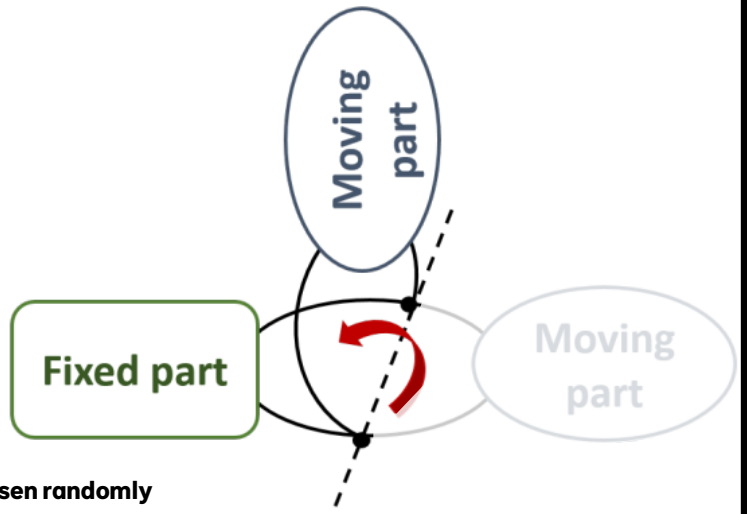
SAX



Linker (ϕ, ψ) mutation



Double linker (axis) mutation



Mutated residue is chosen randomly

B: 83 – 124

loop

B: 427 – 454

linker

A: 422 - 452

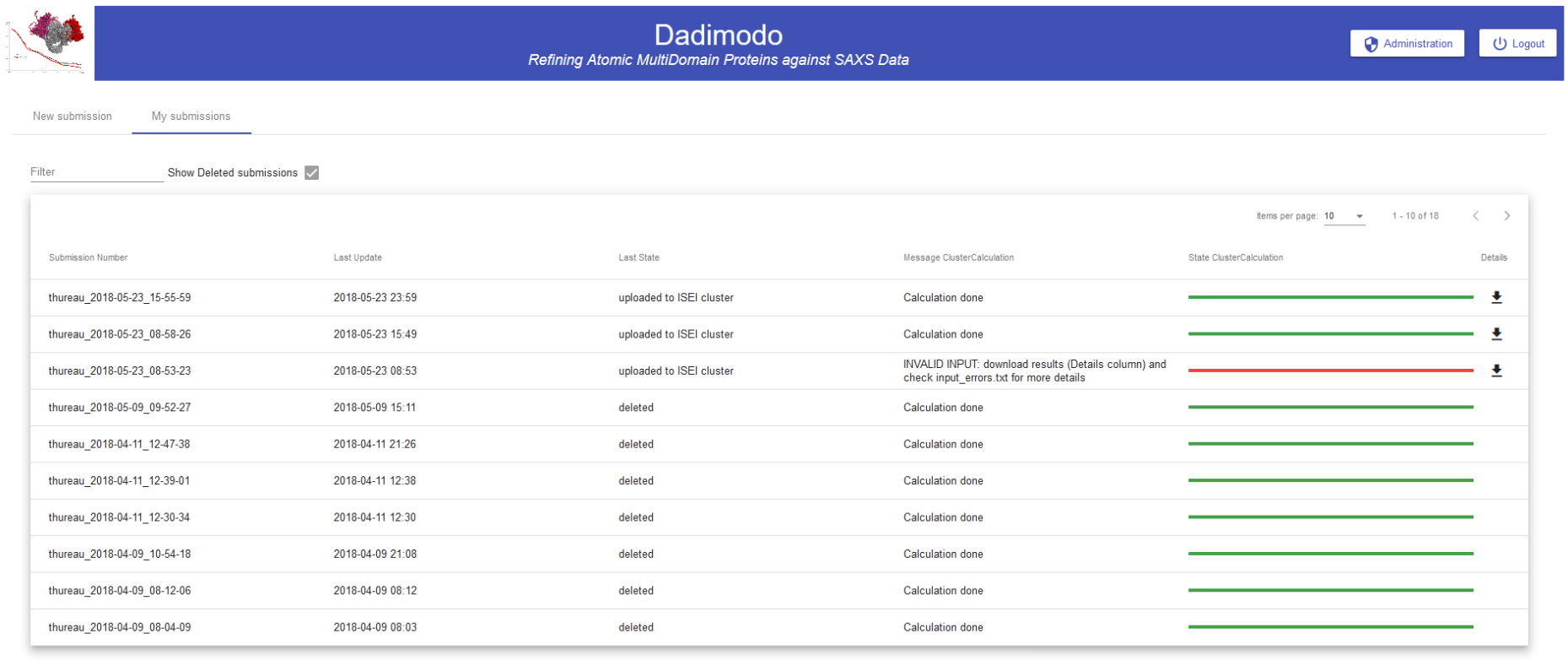
A: 453 – 1179
&
B: 455 – 1179

Dadimodo Web Server

https://dadimodo.synchrotron-soleil.fr

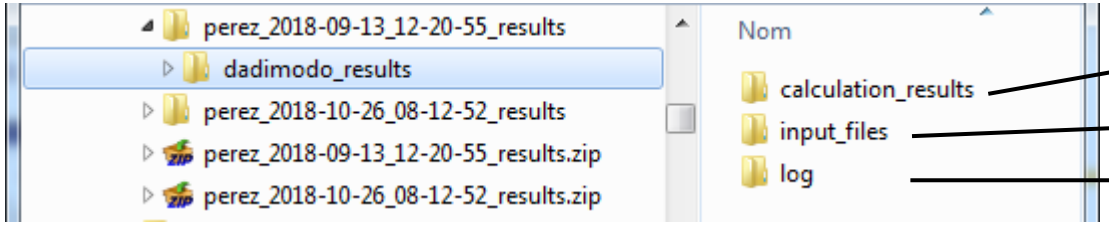
« My submissions » tab:

- ***Status of current submission and history of past jobs***
- ***Results download (zip file)***



Dadimodo results files

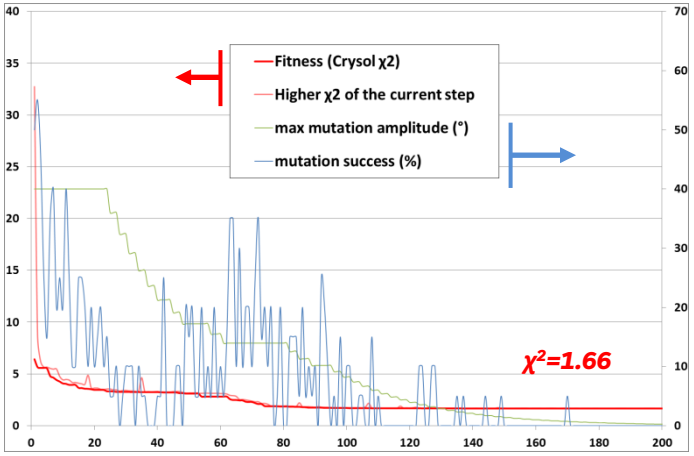
https://dadimodo.synchrotron-soleil.fr

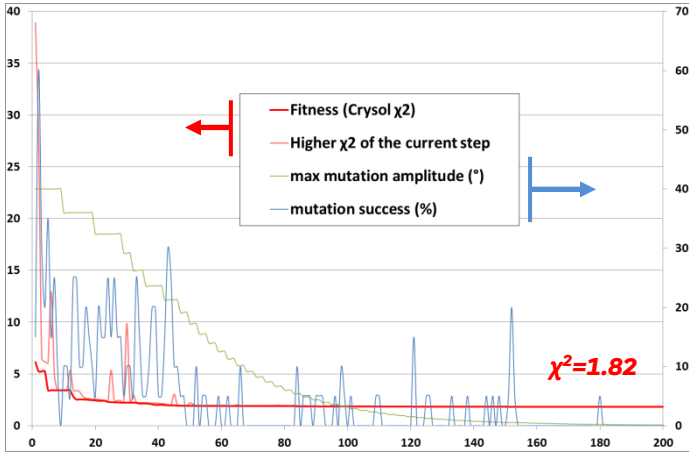


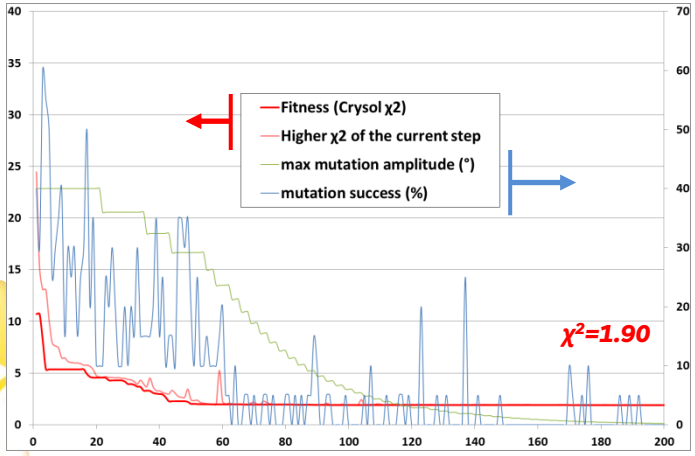
The final pdb files from the several runs

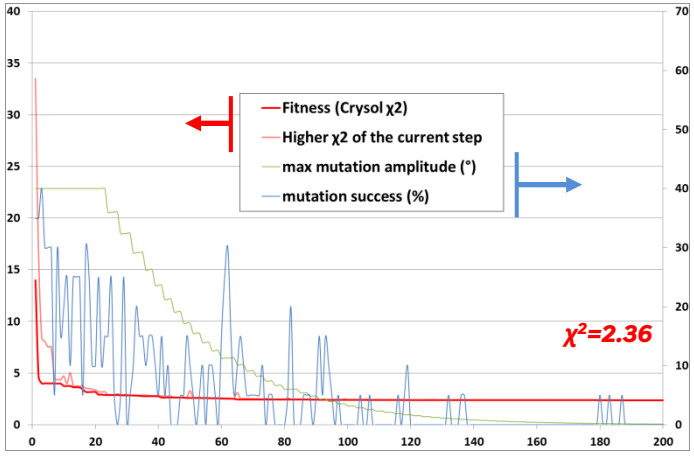
The 3 input files

The exhaustive log files + the statistics files









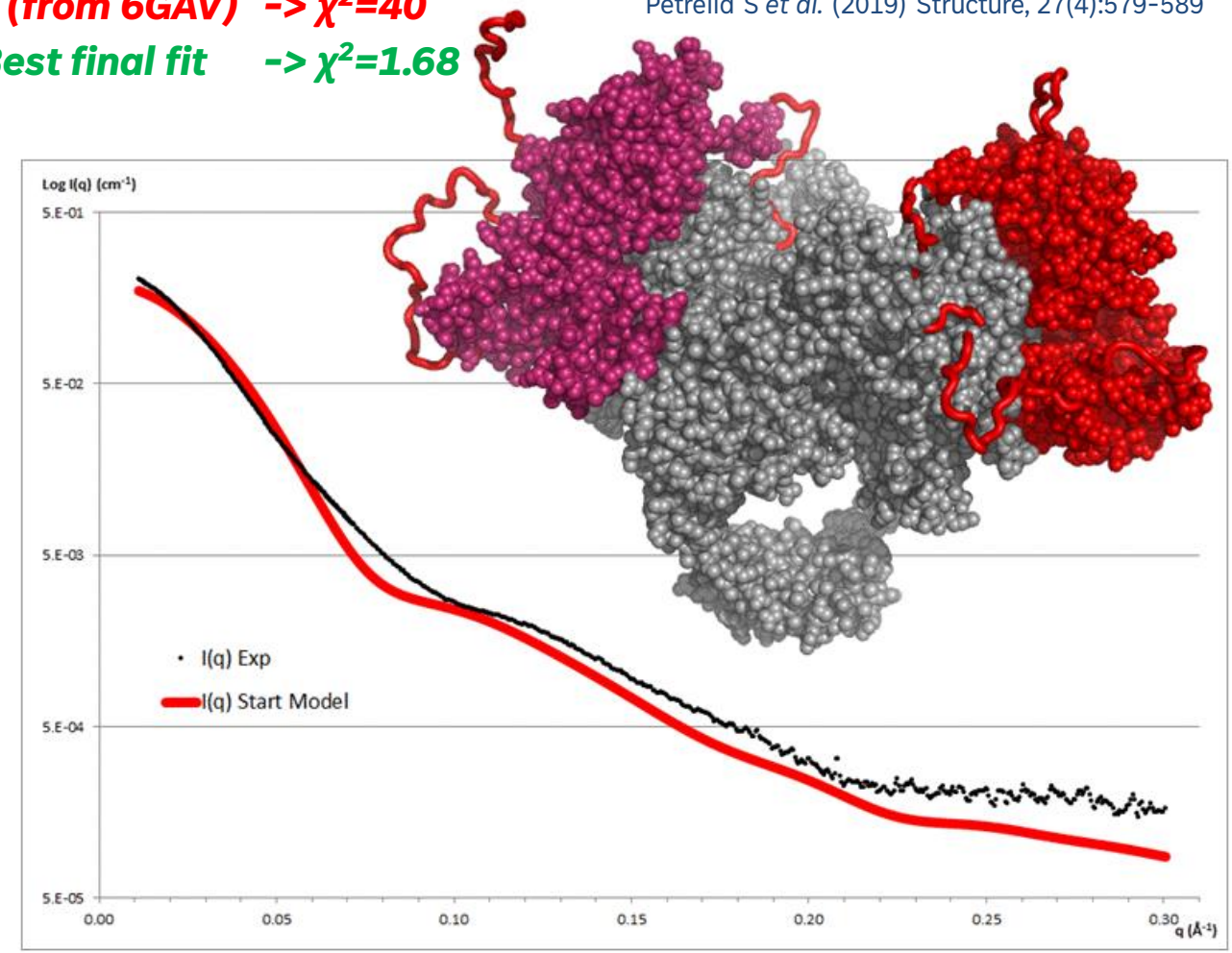
Example: DNA Gyrase

Mycobacterium tuberculosis DNA Gyrase

Petrella S *et al.* (2019) Structure, 27(4):579-589

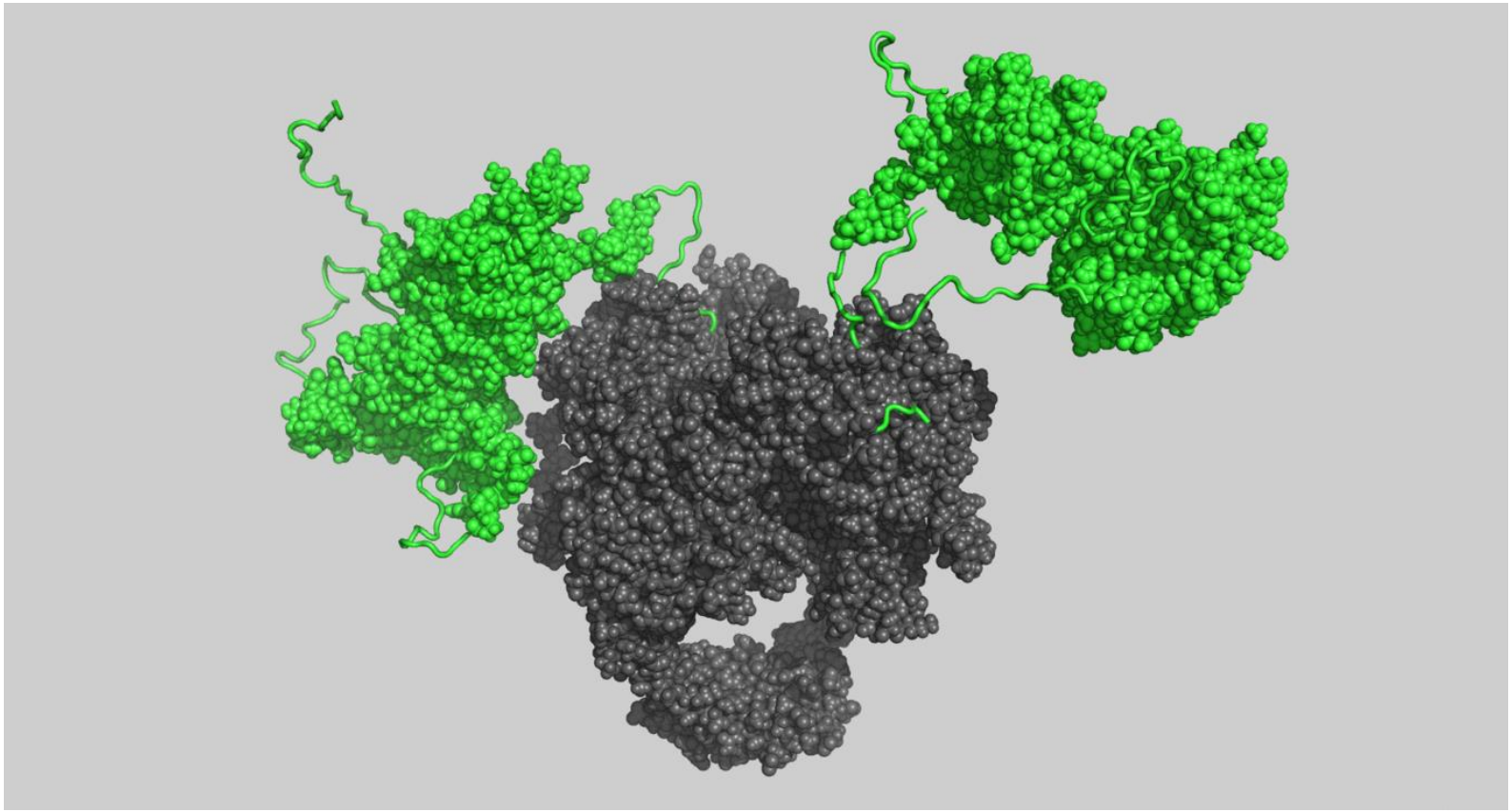
Start model (from 6GAV) -> χ^2=40

Best final fit -> χ^2=1.68



Example: DNA Gyrase

5 best final fits : $1.68 < \chi^2 < 1.76$



Mycobacterium tuberculosis DNA Gyrase

Outline

- ✓ Short presentation of current available techniques at beamline SWING
- ✓ Memprot program: why and how

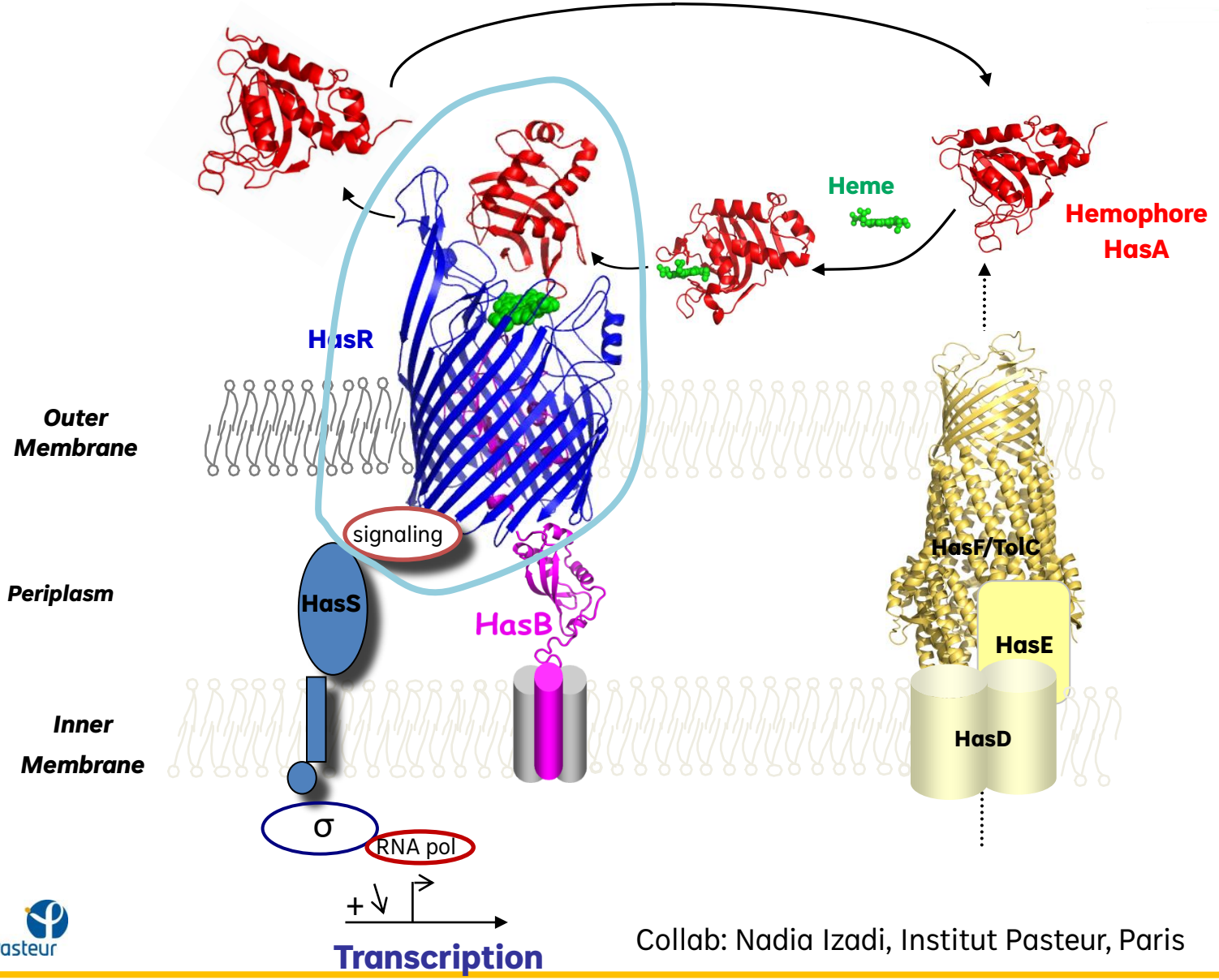
  AQP0 as the « Guinea pig » protein
- ✓ Example of a recent application using Memprot:

  MHST protein in 4 different detergents
- ✓ Dadimodo program: why and how
- ✓ Example of application coupling Memprot & Dadimodo:

  HasA-HasR protein with two different constructs
- ✓ Short conclusion: foreseen new functionalities for Memprot and Dadimodo

Has: Heme acquisition system

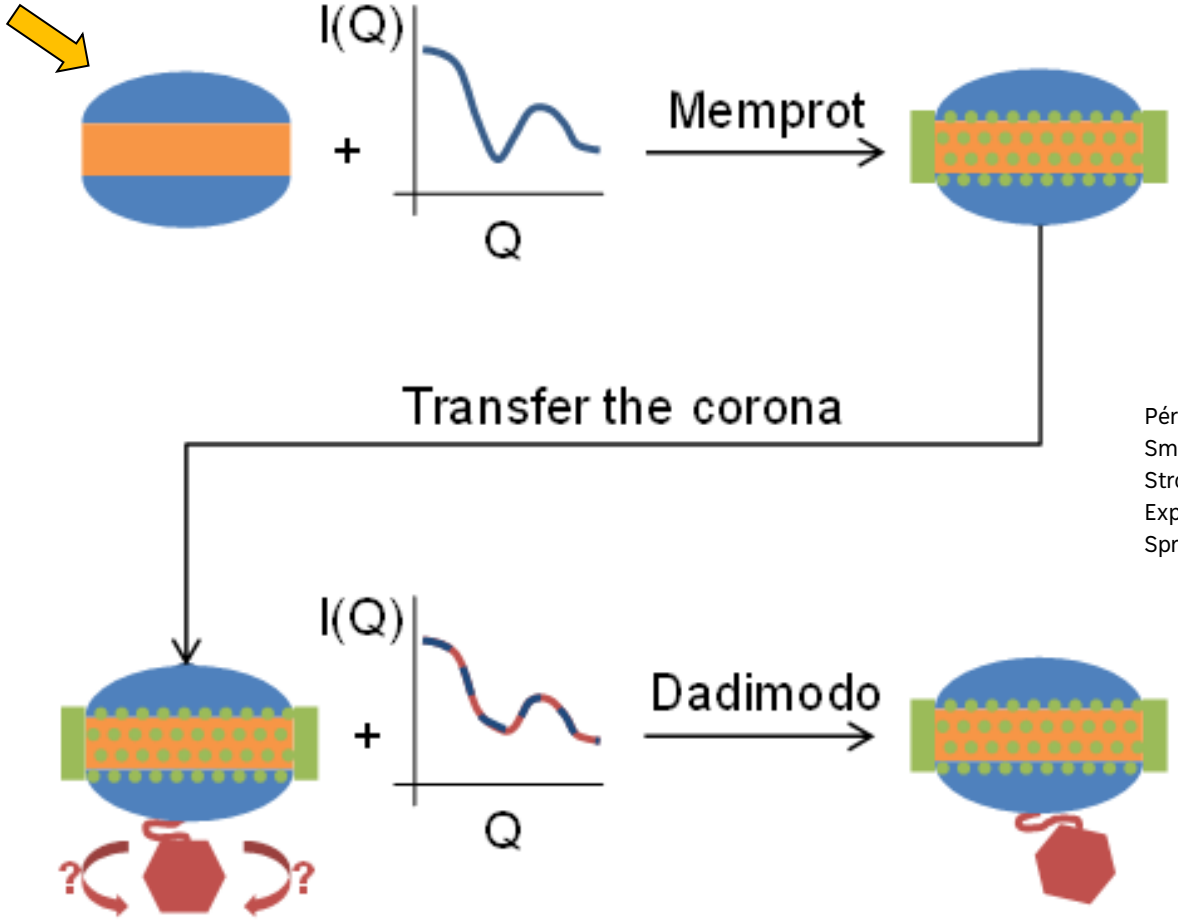
Gram- Bacteria, e.g. Serracia marcescens, opportunist pathogen



Collab: Nadia Izadi, Institut Pasteur, Paris

Our approach: a two-step strategy

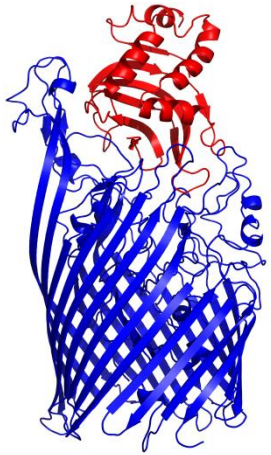
A construct of known structure is needed



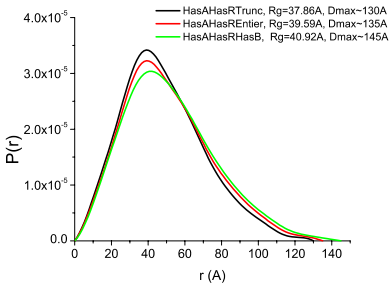
Pérez J., Vachette P. (2017) In: Biological Small Angle Scattering: Techniques, Strategies and Tips. Advances in Experimental Medicine and Biology, vol 1009. Springer, Singapore

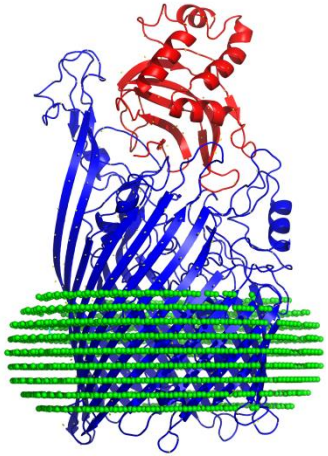
HasA-HasR: Modeling strategy

1/

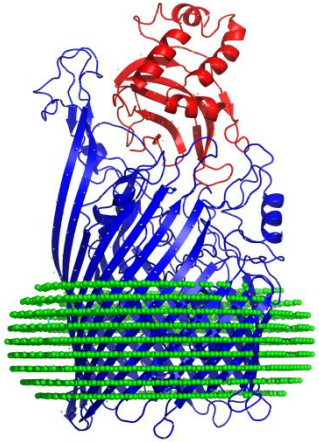


SAXS measurements
(distance constraints)

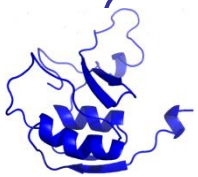


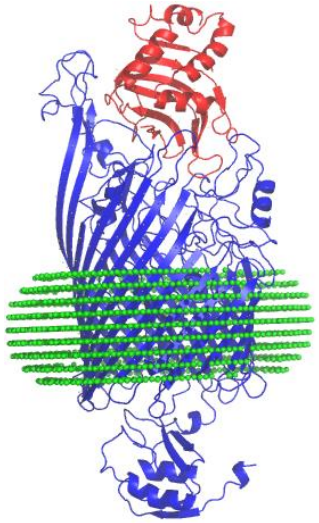


2/



NMR structure of HasR signaling domain
+ SAXS data

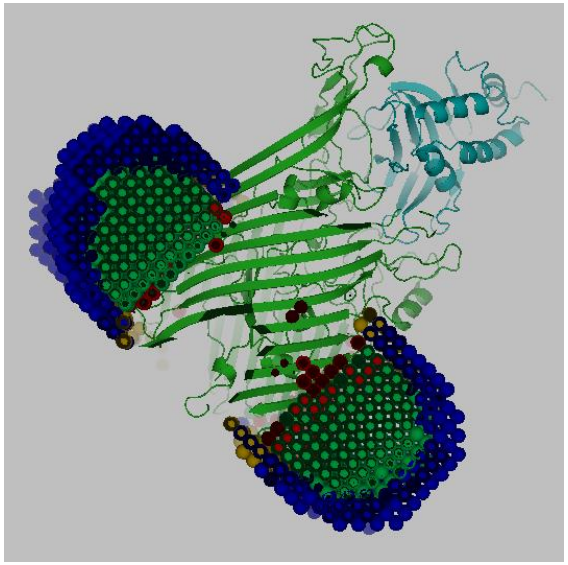


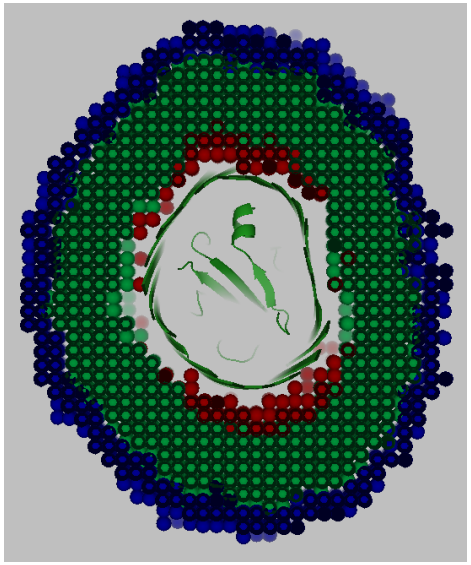


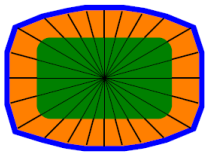
Entire HasR complex including the signaling domain

Corona geometrically adapted to protein shape

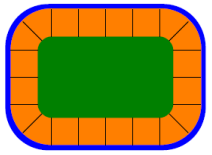
Wojtowicz *et al.*, Biochem. J. (2016) **473**, 2239–2248





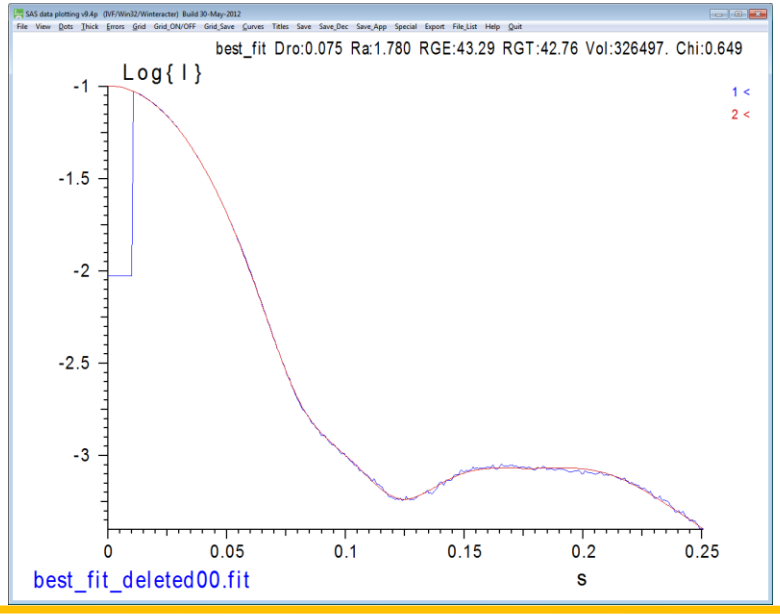


ASA1: Constant thickness on the **line between a given pseudoatom** of the corona **and the center** of the corona.



ASA2: Thickness defined as **the shortest distance to the protein's surface.**

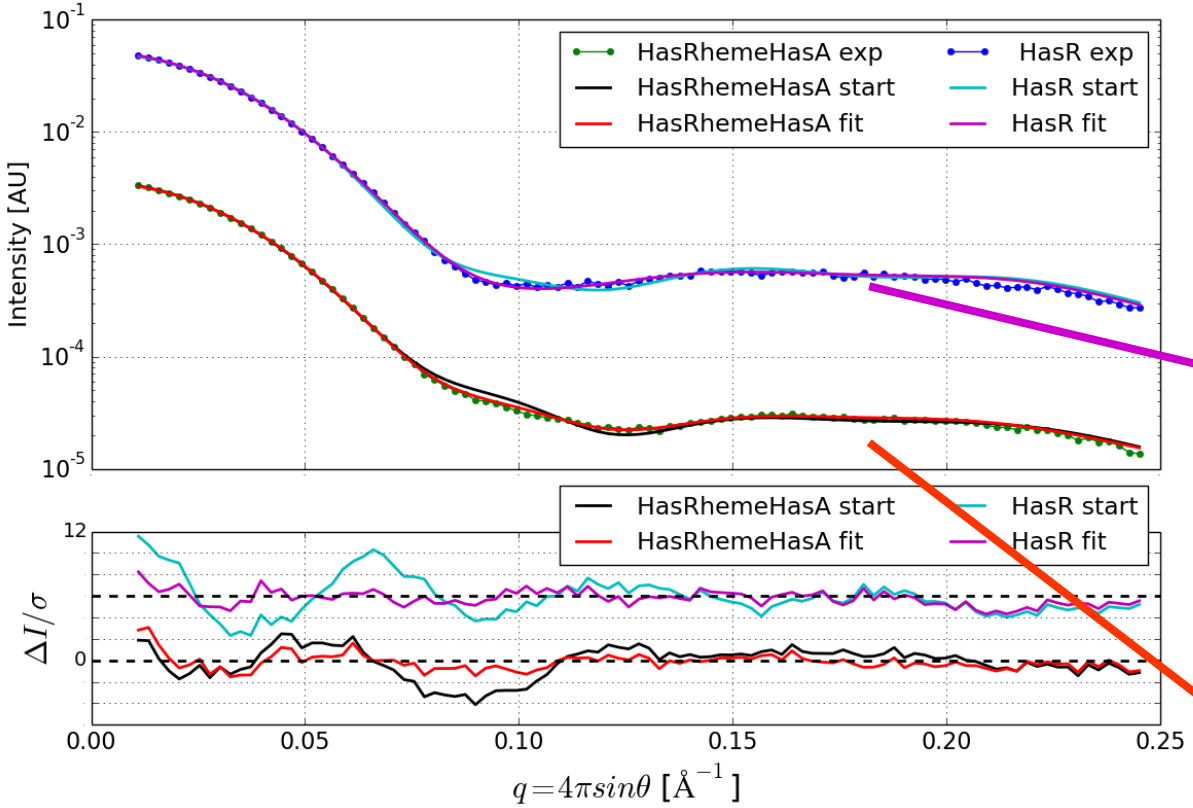
a = 33.500
b = 2.600
t = 5.400
e = 1.110
chi^2 = 2.005
electron density of hydrophobic part= 0.272
electron density of hydrophilic part= 0.506
Number of detergents (tails calc) = 285
Number of detergents (heads calc) = 240

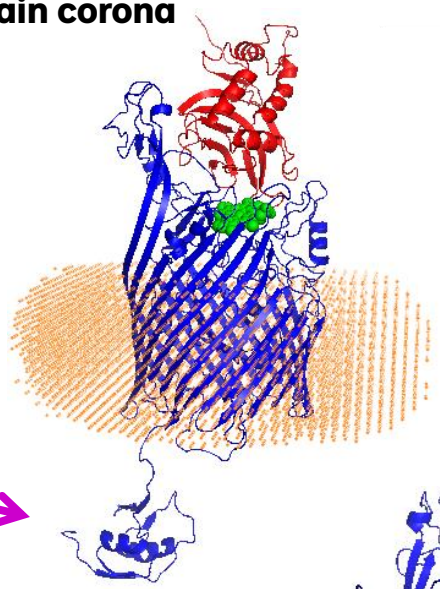


HasR-HasA, now including the signal domain

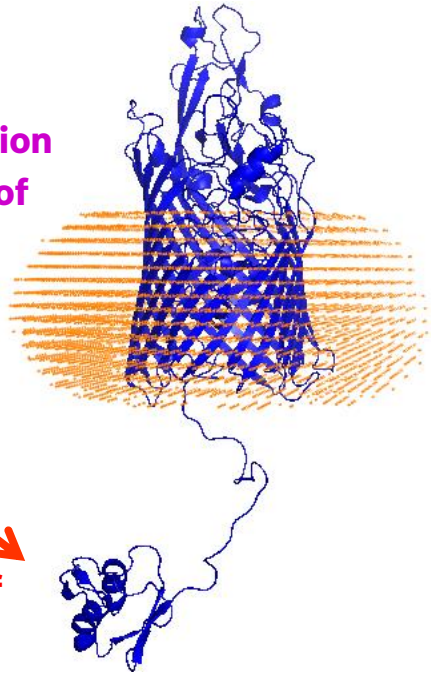
Dadimodo → specially adapted for coarse grain corona

Wojtowicz *et al.*, Biochem. J. (2016) **473**, 2239–2248





Average position in presence of HasA

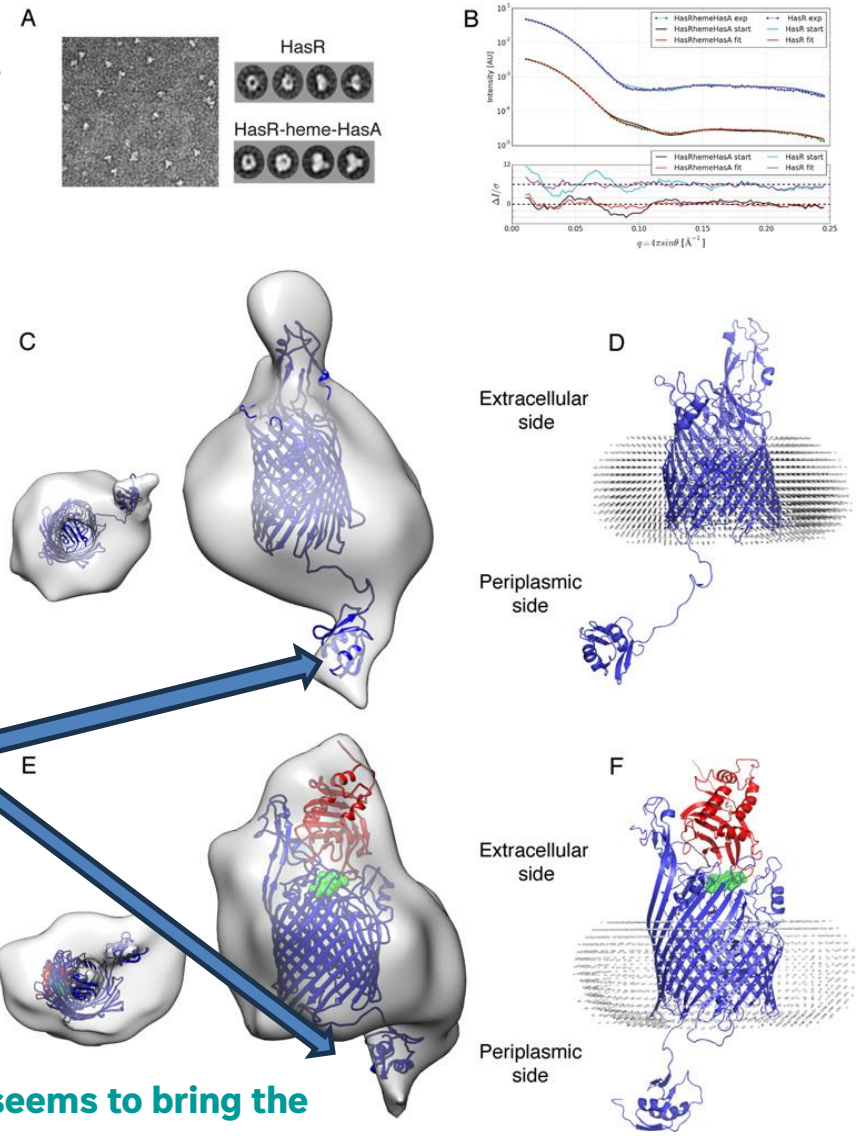


Average position in the absence of HasA

The interaction of HasA with HasR seems to bring the signaling domain closer to the membrane

HasR-HasA, now including the signal domain

Wojtowicz *et al.*, Biochem. J. (2016) **473**, 2239–2248



Dadimodo models are in agreement with further EM analysis !

The interaction of HasA with HasR seems to bring the signaling domain closer to the membrane

Outline

- ✓ Short presentation of current available techniques at beamline SWING
- ✓ Memprot program: why and how

  AQP0 as the « Guinea pig » protein
- ✓ Example of a recent application using Memprot:

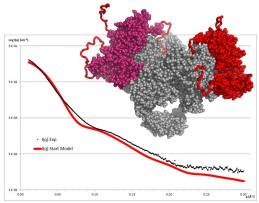
  MHST protein in 4 different detergents
- ✓ Dadimodo program: why and how
- ✓ Example of application coupling Memprot & Dadimodo:

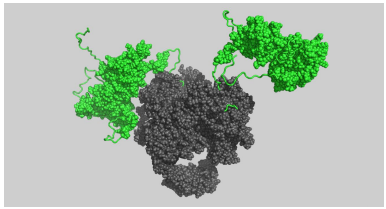
  HasA-HasR protein with two different constructs
- ✓ Short conclusion: foreseen new functionalities for Memprot and Dadimodo

Roadmap

Dadimodo

- Better stop criteria (different from number of generations)
- More friendly output for each run (plot figures,...)
- Summary file for all runs (classification of individual results)
- ADR constraints available on WebServer version (currently only available on local version)

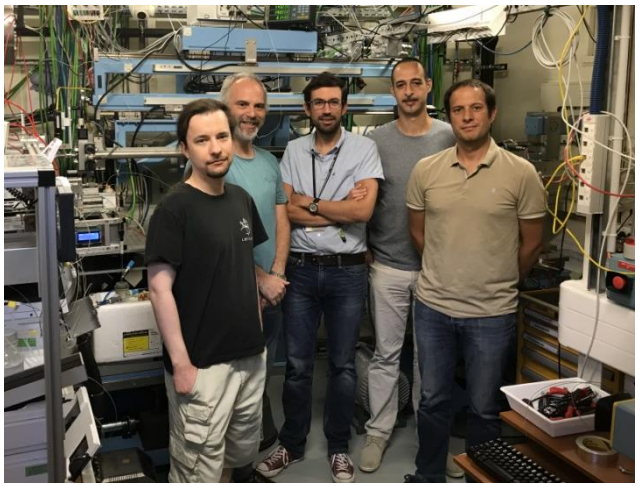




Memprot

- Commissioning of other geometries (bicelles & nanodiscs)
- Web server for direct access by users (currently only staff can use the HPC)

Collaborations

- AQP-0
  - **Alice Berthaud**, Institut Curie
  - Stéphanie Mangenot, Institut Curie
  - **Alexandros Koutsioumpas**, Swing + Julich ForschungsZentrum
- MHST
  - Poul Nissen team, Aarhus University
  - **Maciej Baranowski**, Swing
- HasA-HasR
  - Nadia Izadi
  - **Alexandros Koutsioumpas**, Julich ForschungsZentrum
- DNA Gyrase
  - Stéphanie Petrella, Unité de Microbiologie Structurale
- Memprot
  - **Maciej Baranowski**
  - **Alexandros Koutsioumpas**
- Dadimodo on the Web
  - **Olga Roudenko**, SOLEIL
  - Aurélien Thureau, Swing
  - Alejandro Diaz, SOLEIL

Beamline SWING
- Maciej Baranowski
- Javier Pérez
- Thomas Bizien
- Youssef Liatimi
- Aurélien Thureau
- Arnaud Gibert