Small angle Xray scattering: a way to study the conformation, assemblies and interactions of biological macromolecules in solution

- SAXS for determining structures of complexes
- SAXS for determining conformational changes
  - Kinetics of a conformational transition
  - Order-disorder transition
    - SAXS for highly flexible proteins

Recent SAXS reviews

Small-angle scattering for structural biology—Expanding the frontier while avoiding the pitfalls
Jacques D.A., Trewhella J.
Protein Sci. (2010) 19: 642-57.

Structural characterization of proteins and complexes using small-angle X-ray solution scattering
Mertens H.D.T. and Svergun D. I.
J. Struct. Biol. (2010) 172: 128-41.

Small and Wide Angle X-ray Scattering from Biological Macromolecules and their Complexes in Solution
Doniach S., Lipfert J.
Comprehensive Biophysics – Vol. I (2012) Elsevier, p. 376-97

Super-resolution in Solution X-Ray Scattering and Its Applications to Structural Systems Biology
Rambo R.P. and Tainer J.A.
Annu. Rev. Biophys. (2013) 42: 415-41.

Websites

Svergun's team: http://www.embl-hamburg.de/biosaxs/software.html
J. Tainer and coworkers: http://www.bioisis.net

Historique

Pionnier: **A. Guinier (1911-2000)**

Observation de diffusion aux petits angles dans des alliages au cours du vieillissement (**~1937**)

→ Zones de Guinier-Preston=zones de concentration de l'un des types d'atomes dans un alliage (AlCu)

→**Définition du " rayon de giration " des agrégats de toute nature**

« *Simple experiments convinced me that small-angle scattering is indeed characteristic for the division of matter into submicroscopic particles (<1000Å) and does not depend on the atomic structure. Thus it is observed for very small crystals in colloïdal metals as well as for amorphous particles (silica gel) and colloïdal solutions (albumen).* » (A. Guinier, Personal Reminiscences, 1962)

→ Ouvrage de référence: Small Angle Scattering of X-rays, A. Guinier, G. Fournet 1955

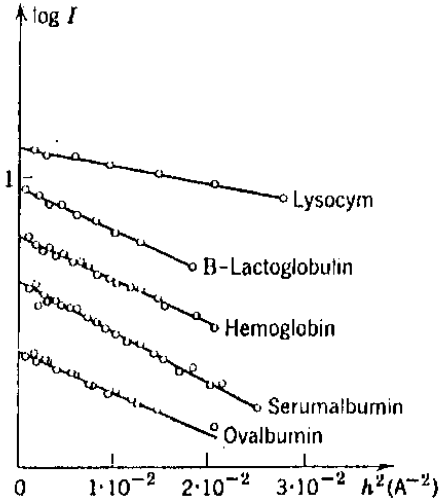
THE JOURNAL OF CHEMICAL PHYSICS VOLUME 18, NUMBER 9 SEPTEMBER, 1950

An X-Ray Investigation of the Shapes and Hydrations of Several Protein Molecules in Solution*

H N RITLAND,† PAUL KAESBERG, AND W. W. BEEMAN
Department of Physics, University of Wisconsin, Madison, Wisconsin

(Received May 10, 1950)

The x-ray scattering curves of five proteins in solution have been measured at small angles The radii of gyration of the protein molecules are determined from the scattering curves These data together with the known molecular weights, densities, and frictional ratios are used to estimate the axial ratios and hydrations of the molecules. Some possibilities and limitations of the method are pointed out.



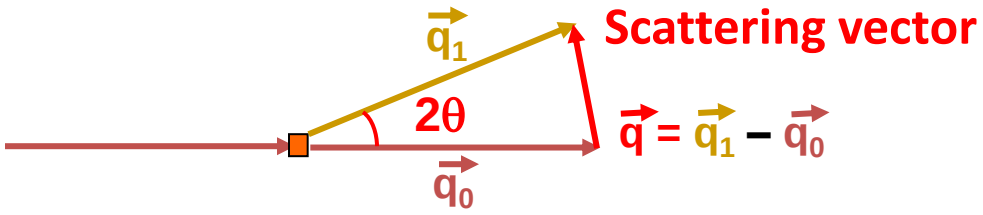
Bond en avant > 1990

Amélioration de la qualité des données (synchrotron)

Développement de nouveaux programmes d'analyse (Svergun et al.)

Solution X-ray scattering

Diagram of the experimental set-up

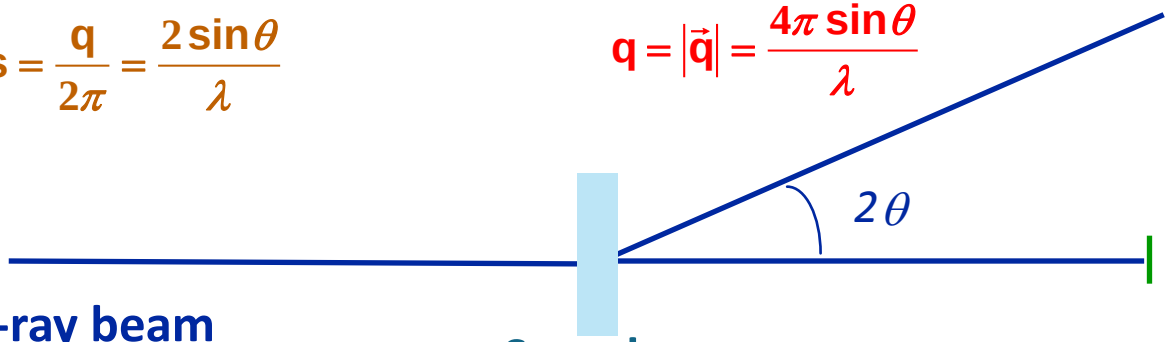


Scattering vector

$\vec{q} = \vec{q}_1 - \vec{q}_0$

$$s = \frac{q}{2\pi} = \frac{2\sin\theta}{\lambda}$$

$$q = |\vec{q}| = \frac{4\pi\sin\theta}{\lambda}$$

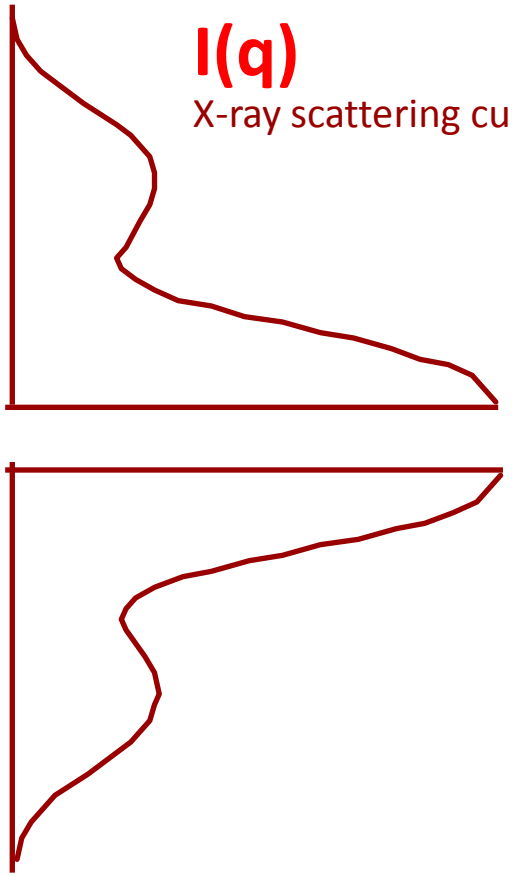


X-ray beam

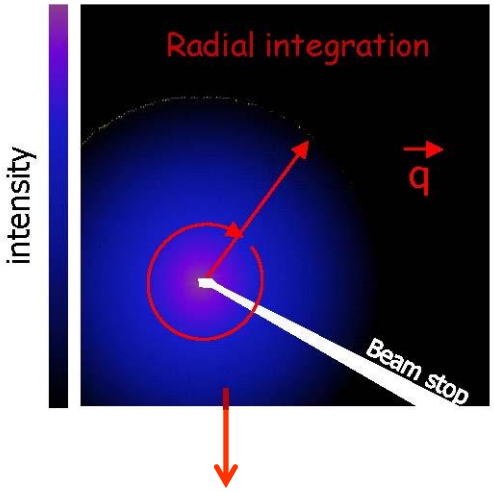
Sample

10µl – 50µl

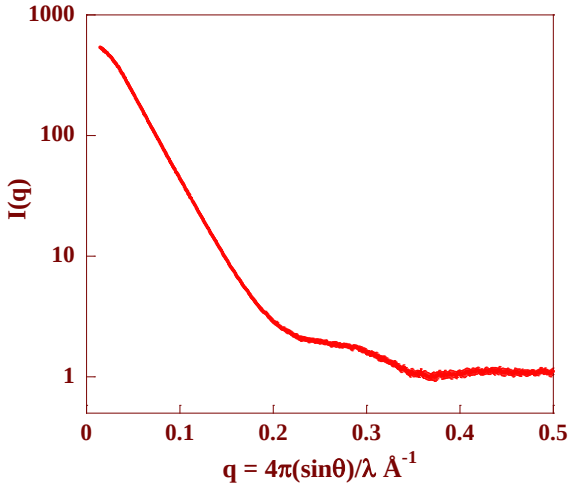
0.1mg/ml –(>)10mg/ml



A 1 mg/ml solution of a globular protein 15kDa molecular mass will scatter in the order of 1 photon in 10^6 incident photons.



SAXS pattern



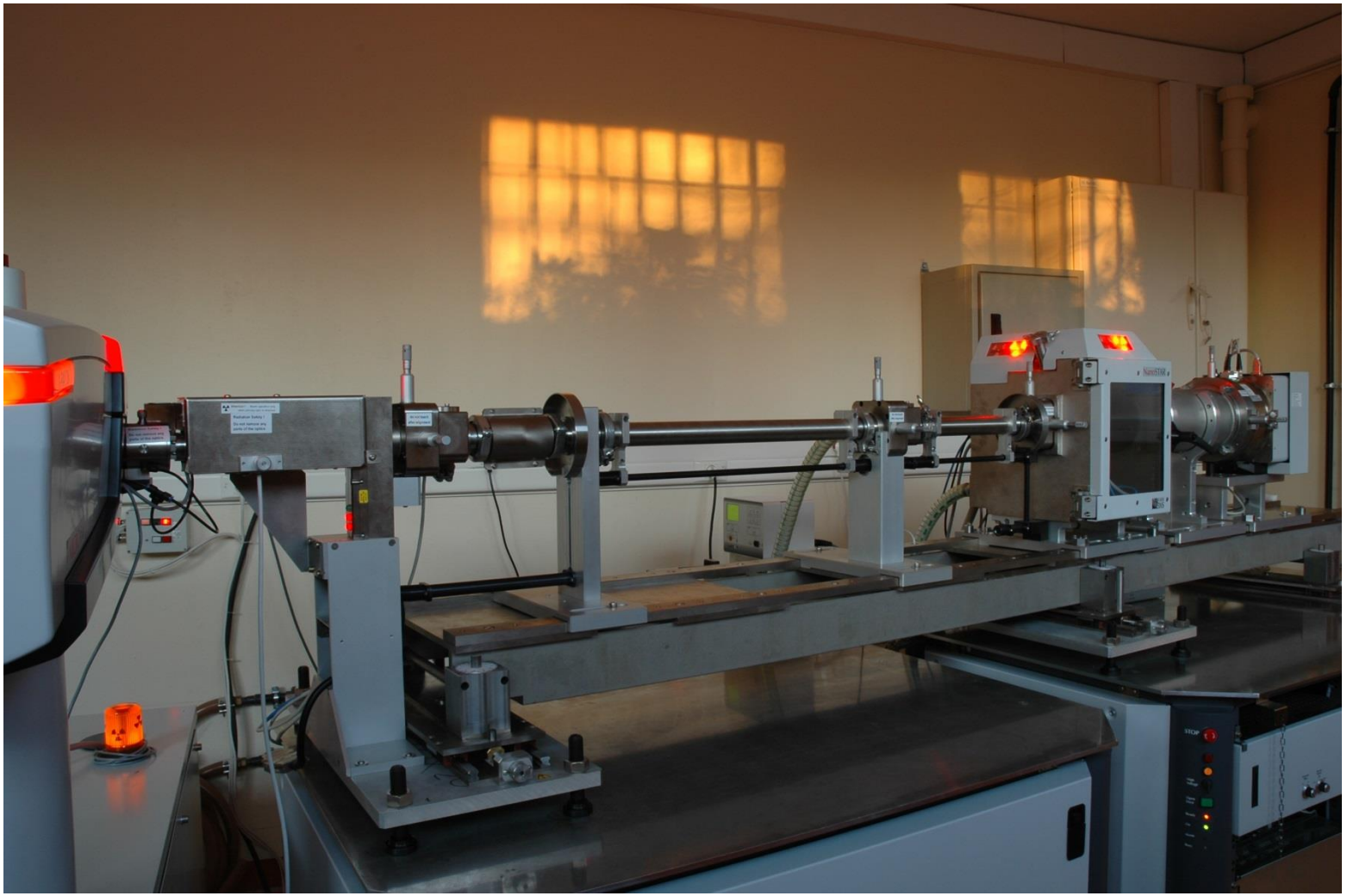
- **No crystal required, no special sample preparation.**
- **No mass limitation : from a few kDa proteins to 100 MDa viruses.**
- **Can be used under quasi-physiological conditions.**
- **Makes possible the static or kinetic study of conformational changes.**

BUT

✓ **Low resolution method best used in combination with other structural (cristallography, NMR, EM), dynamic (NMR, fluorescence), biochemical (e.g. cross-link + MS) and/or computational (data-driven docking) approaches.**

Lignes à haut flux de laboratoire

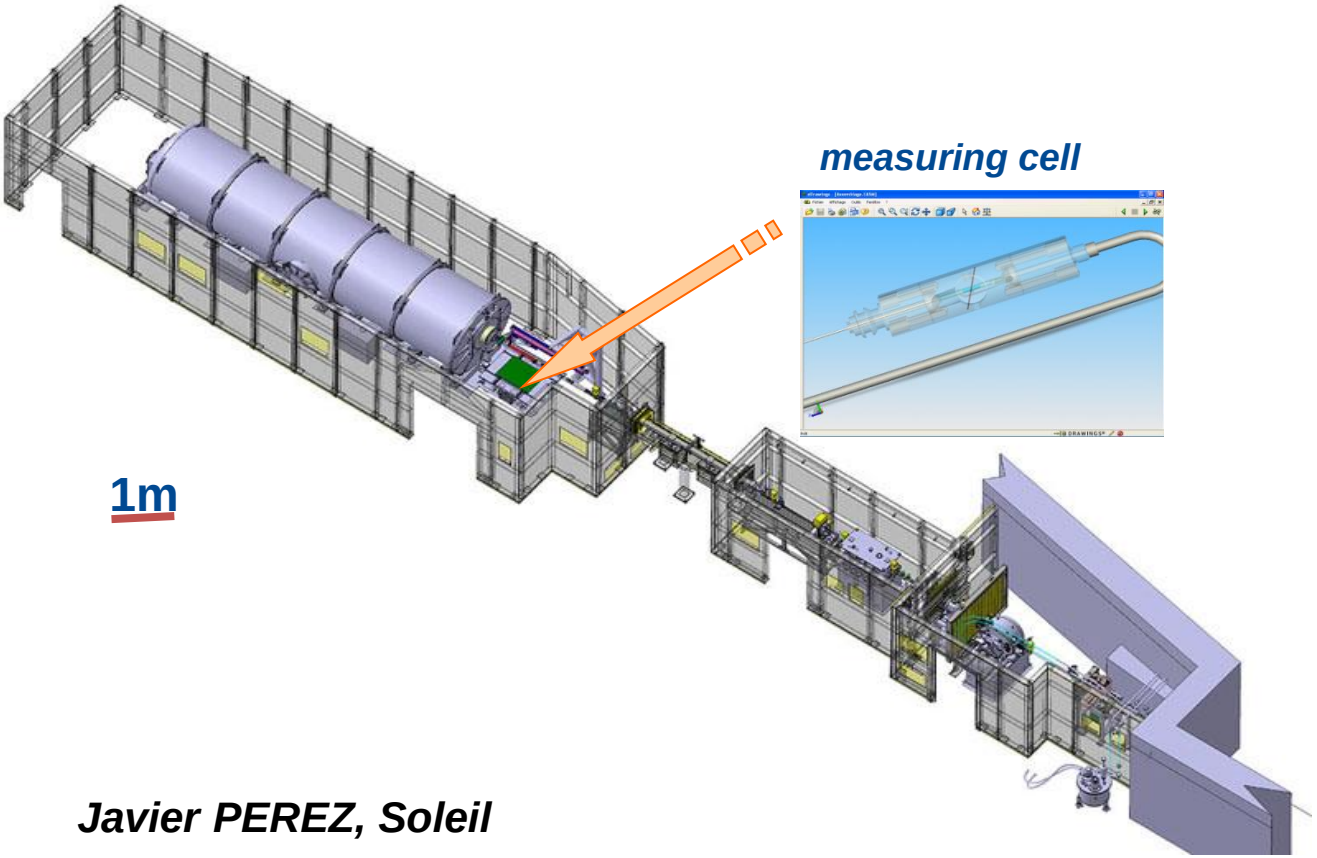
- Nanostar I2BC (Orsay), Aarhus
- BioSAXS2000 IGBMC (Strasbourg)



Flux x10000 !!!!!

Synchrotron beamlines

P12/Petra (Hamburg)
SWING/SOLEIL (Saclay)
BM-29/ESRF (Grenoble)
I22/Diamond
ID-18 BioSAXS/APS
BL4.2/SSR
SR13 ID01/Australian Synchrotron
…

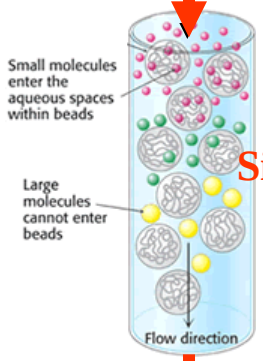


Javier PEREZ, Soleil

SE-HPLC / Solution Sampler

Flow rate 200 µl/min

- Monodispersity is essential for SAXS measurements
- Aggregation should be eliminated
- Oligomeric conformations can be distinguished
- Equilibrium states can be transiently separated
- No time lost in collecting solution from HPLC

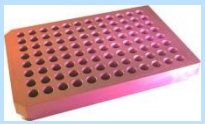


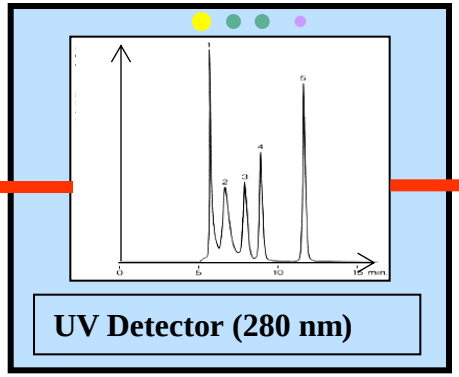
Size Exclusion

Incident X-ray

Pump

Injection-mixing





SAXS

Cell

SLS

refractometry

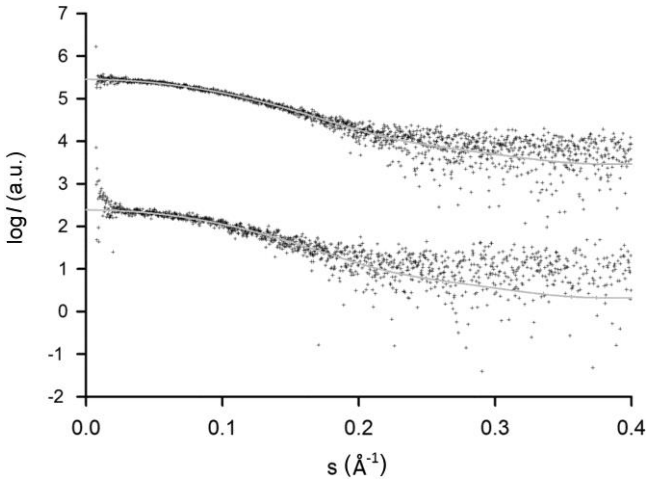
viscosity

courtesy of J. Pérez & G. David
(Swing, SOLEIL)

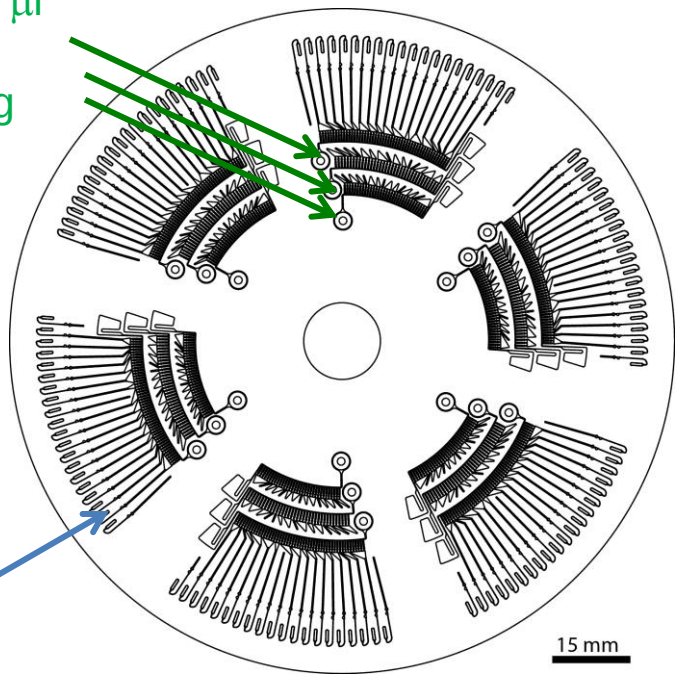
Microfluidics

Rapid high-throughput analysis

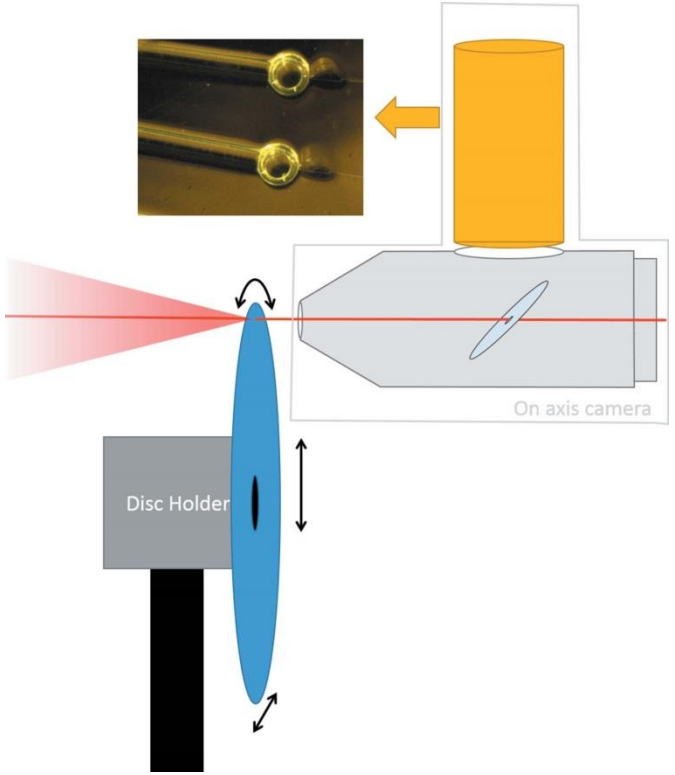
50 ms



3 reservoirs 2-3 μl sample, mixing buffer, screening buffer



120 aliquots of 40nl each



P12/PETRA (Hamburg)

C. Blanchet … D.Svergun, J. Appl. Cryst. 2015

★ **Particle in solution** => the particle adopts all orientations / X-ray beam.

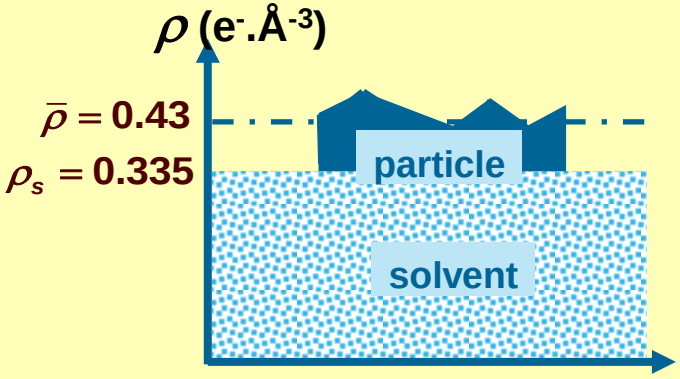
=> only the *spherical average* of the scattered intensity is experimentally accessible.

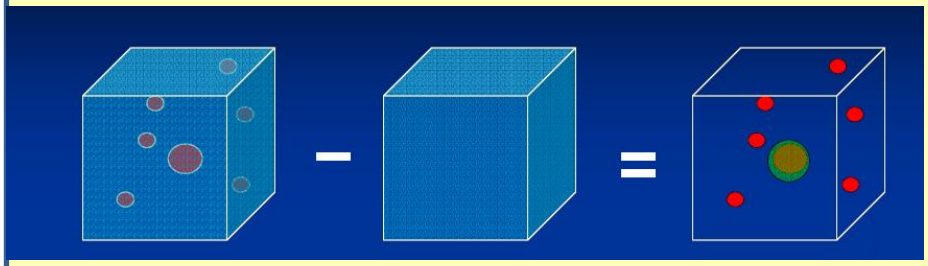
$$I(q) = \langle I(\vec{q}) \rangle_{\Omega}$$

$$I(q) = \sum_{i=1}^{N} \sum_{j=1}^{N} f_i f_j \frac{\sin(qr_{ij})}{qr_{ij}}$$

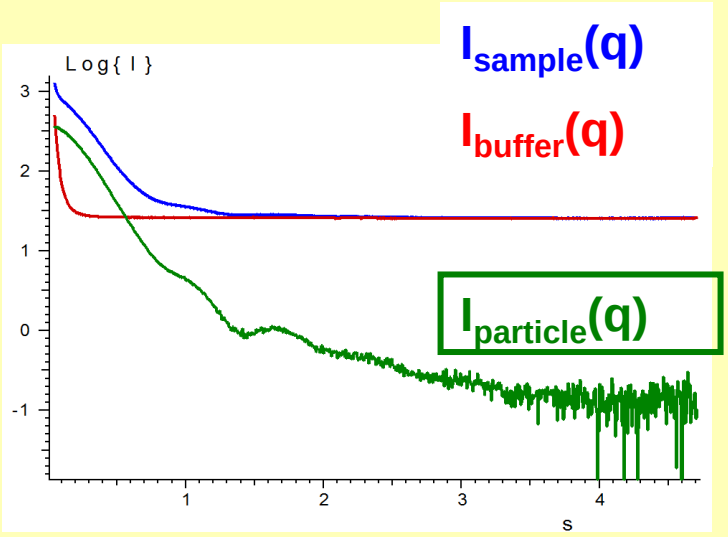
Debye, 1915

★ In solution, what matters is the ***contrast*** **of electron density between the particle and the solvent** $\Delta\rho(\underline{r}) = \rho(\underline{r}) - \rho_s$, small for biological samples.





$I_{sample}(q)$ $I_{buffer}(q)$ $I_{particle}(q)$



Ideal and monodisperse solution

$I_{ideal}(q)=Ni_1(q)$

N particles

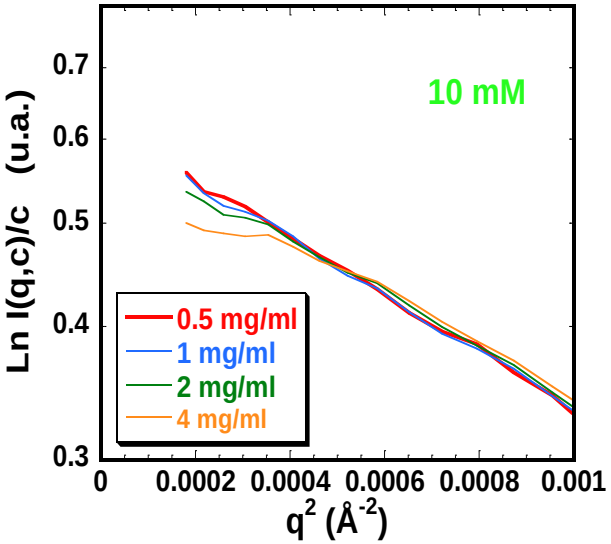
- **No interactions between particles**
- **Identical particles**

$$I_{ideal}(q)=N\left|\int_{V_p}(\rho(\vec{r})-\rho_s)\exp(-i\vec{q}.\vec{r})\,d\vec{r}\right|^2$$

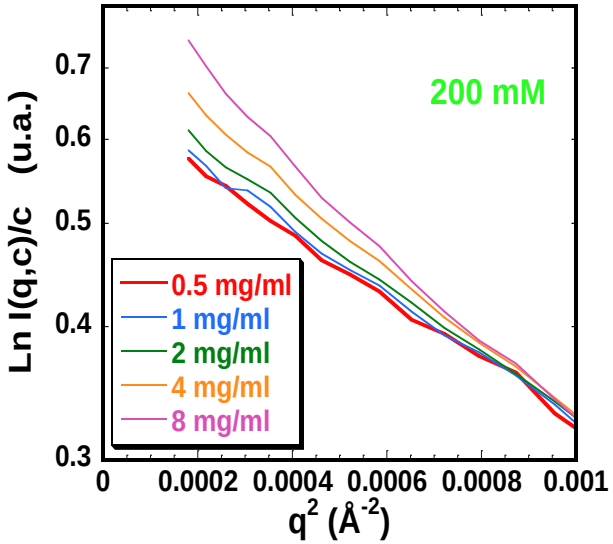
Interactions between particles

$I(q,c)=I(q)_{ideal}\,S(q,c)$

Repulsive



Attractive



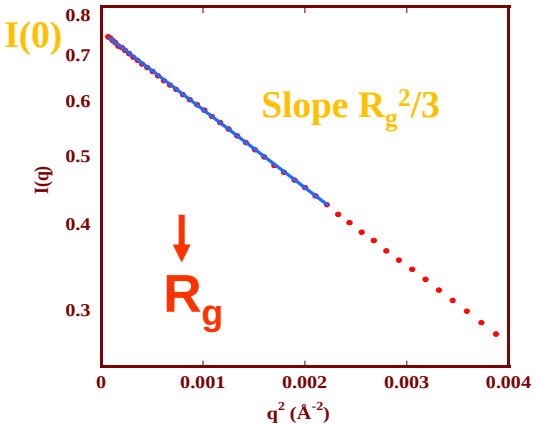
data quality assessment

1st stage **Model-free analysis**

Guinier analysis

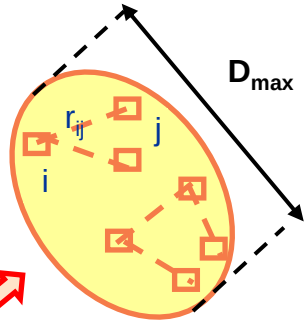
$$ln[I(q)] \cong ln[I(0)] - \frac{R_g^2}{3} q^2$$

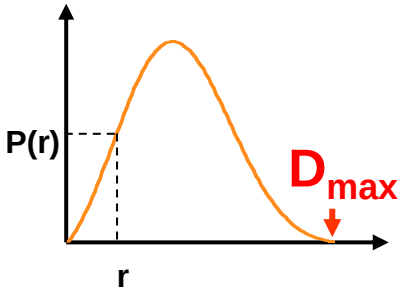
$$\frac{I(0)}{c} \propto M$$



Distance distribution function

$$p(r) = \frac{r^2}{2\pi^2} \int_0^\infty q^2 \, I(q) \frac{\sin(qr)}{qr} dq$$





I(q)

Volumes

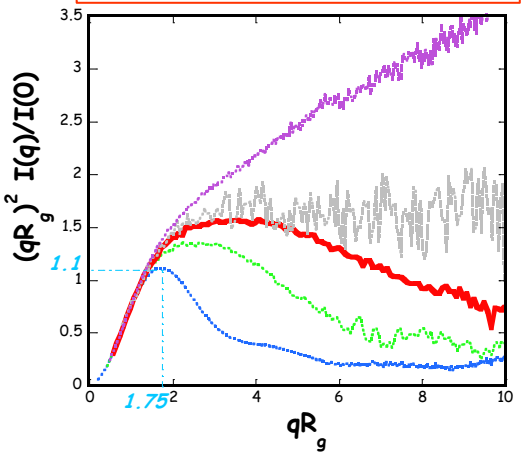
Porod volume *not valid for unfolded proteins*

$$V = 2\pi^2 \cdot \frac{I(0)}{Q} \qquad Q = \int_0^\infty q^2 \cdot I(q)\, dq \rightarrow M$$

Volume-of-correlation *Rambo R.*

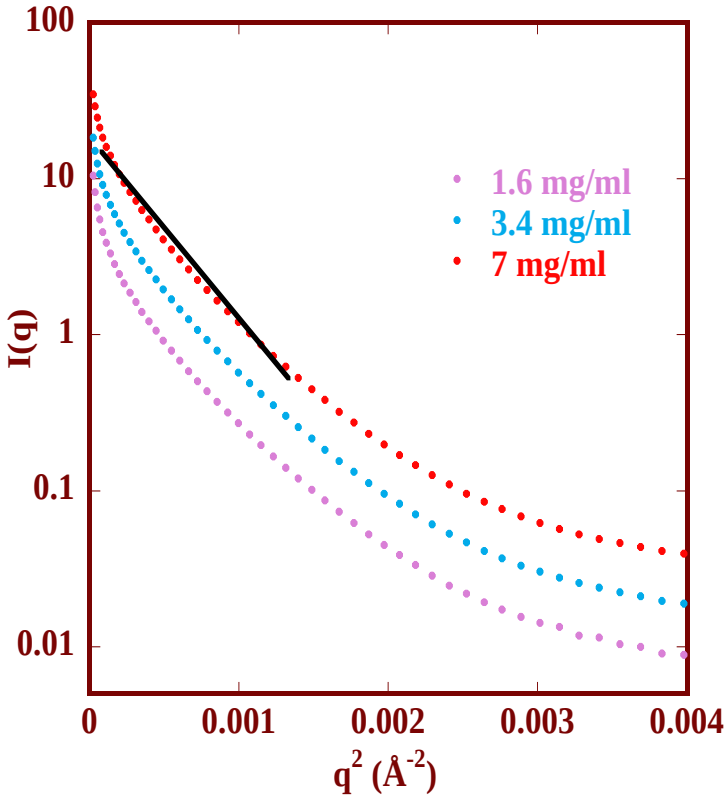
$$V_c = \frac{I(0)}{\int q \cdot I(q)\, dq} \rightarrow M$$

Dimensionless Kratky representation



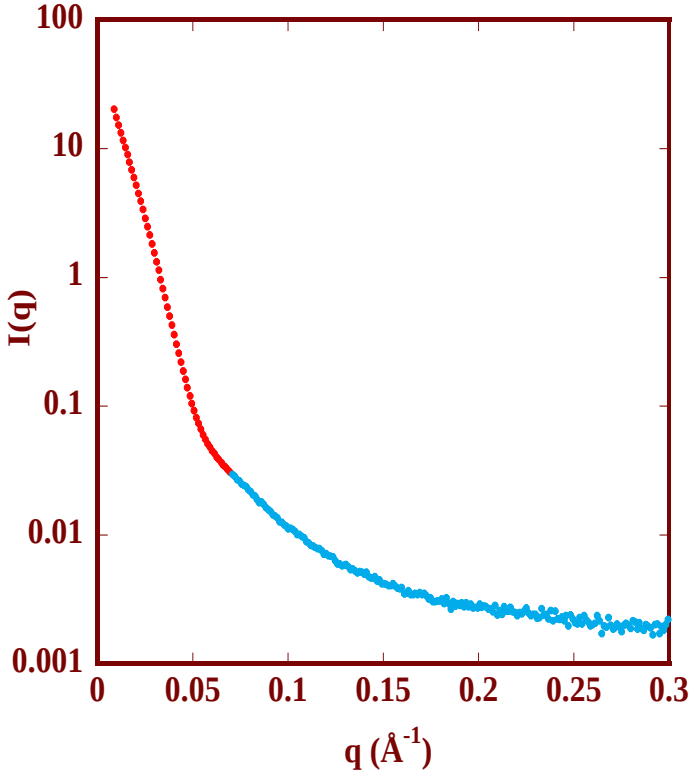
Data quality assessment

Irreversible aggregation



I(0): > 150 fold the expected value for the given MM

Useless data: the whole curve is affected



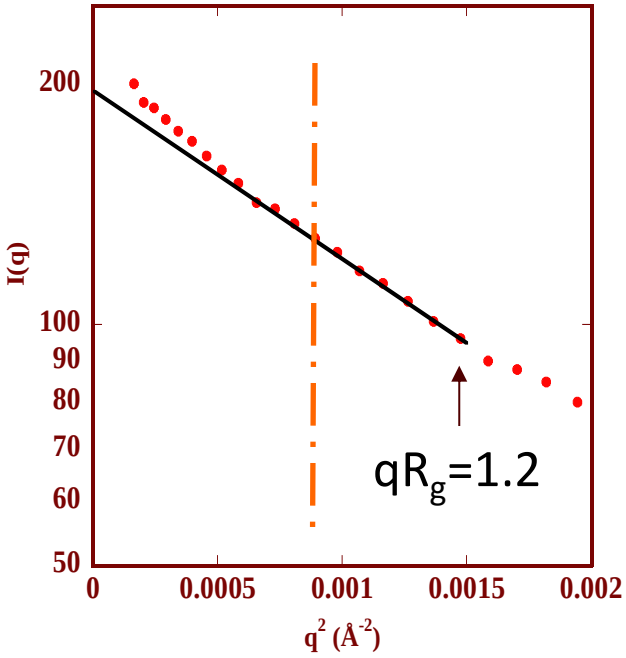
Swing – Domaine 1-242 de RRP44 – 07/08

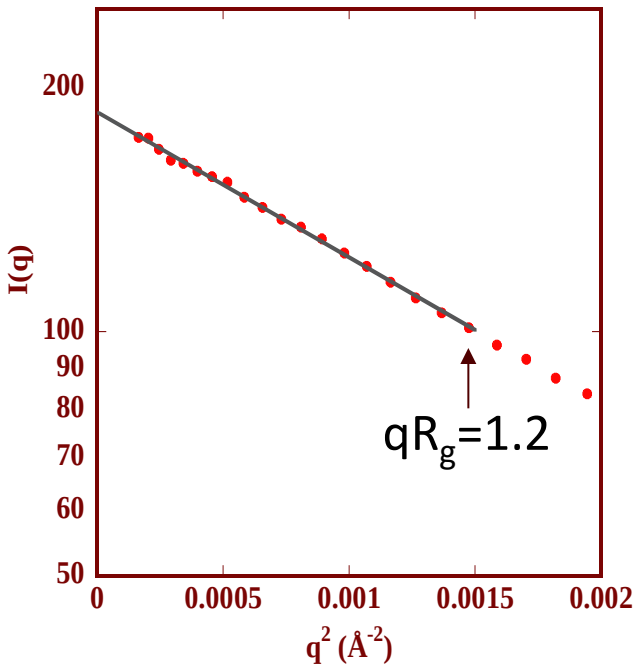
Data quality assessment

weak aggregation → possible improvement

centrifugation, buffer change

Nanostar –PR65 protein



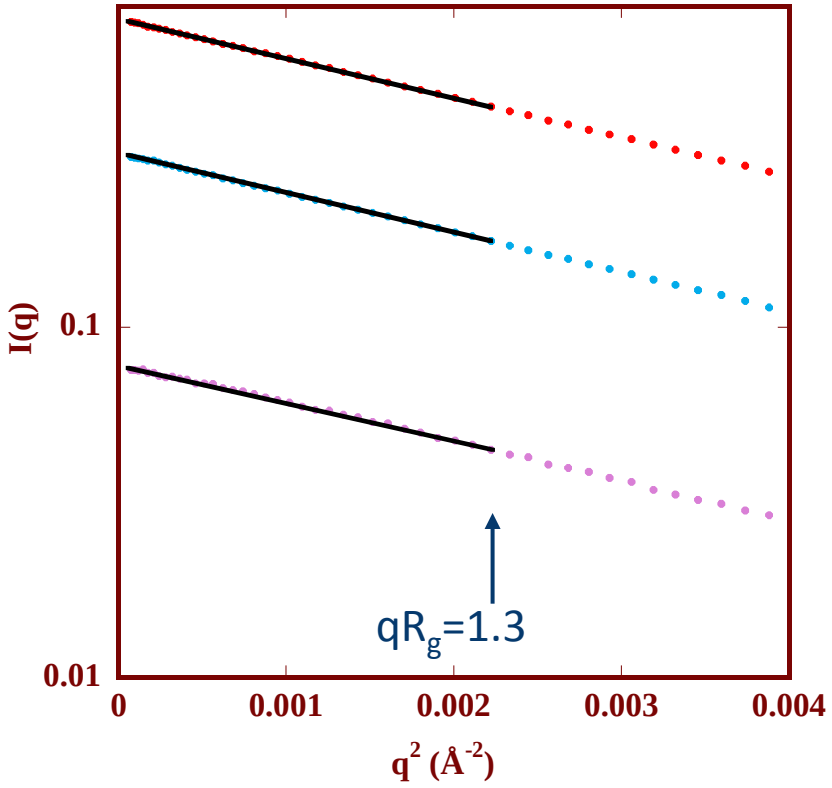


R_g ~ 38 Å – too high!!

R_g ~ 36 Å

Data quality assessment

Guinier plot



same R_g at all three concentrations

⇨ No aggregation, no interactions.

Swing – Polymérase – 07/08

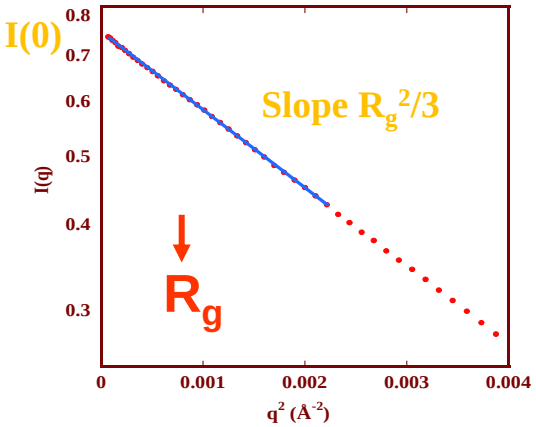
data quality assessment

1st stage **Model-free analysis**

Guinier analysis

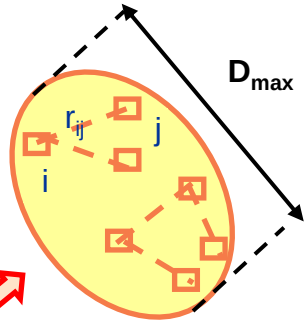
$$ln[I(q)] \cong ln[I(0)] - \frac{R_g^2}{3} q^2$$

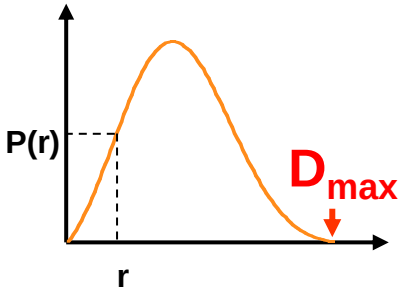
$$\frac{I(0)}{c} \propto M$$



Distance distribution function

$$p(r) = \frac{r^2}{2\pi^2} \int_0^\infty q^2 \, I(q) \frac{\sin(qr)}{qr} dq$$





I(q)

Volumes

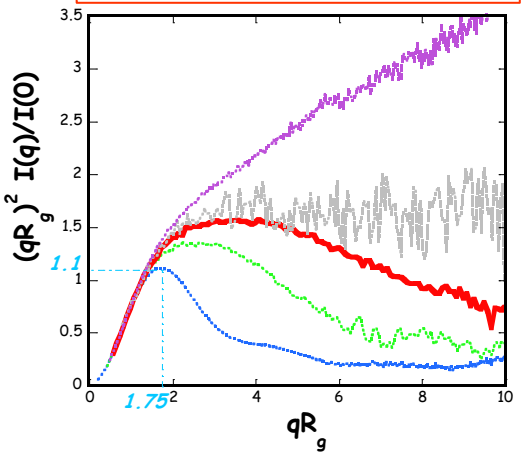
Porod volume *not valid for unfolded proteins*

$$V = 2\pi^2 \cdot \frac{I(0)}{Q} \qquad Q = \int_0^\infty q^2 \cdot I(q)\, dq \rightarrow M$$

Volume-of-correlation *Rambo R.*

$$V_c = \frac{I(0)}{\int q \cdot I(q)\, dq} \rightarrow M$$

Dimensionless Kratky representation



Guinier analysis

$$R_g^2 = \frac{\int_{V_p} r^2(\rho(\vec{r}) - \rho_s)d\vec{r}}{\int_{V_p} (\rho(\vec{r}) - \rho_s)d\vec{r}}$$

R_g^2 is the mean square distance to the center of mass weighted by the contrast of electron density.

If $\Delta\rho(r) \approx$ constant then R_g is a geometrical quantity.

Compact globular protein with n residues:

$R_g \approx 3(n)^{1/3}$

$$I(0) = V_{sol}\frac{cM}{N_a}\left(\frac{m_p N_a}{M} - \rho_s \bar{v}_p\right)^2$$

m_p: electron number

Intensity on an absolute scale (cm^{-1}) using water as primary reference

If the concentration
the specific volume are known

→ **Estimate of M (molar mass)**

$\bar{v}_p$ depends on the protein
0.70-0.76 cm^3/g
Calculated using SEDNTERP
http://sednterp.unh.edu/

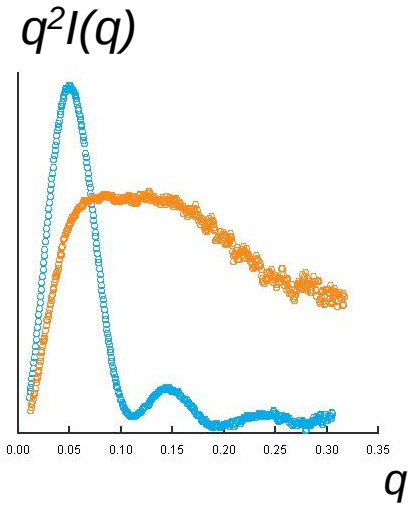
Oligomerization state

Volumes

Porod volume *not valid for unfolded proteins*

$$V = 2\pi^2 \cdot \frac{I(0)}{Q} \qquad Q = \int_0^\infty q^2 \cdot I(q)\, dq$$

Q: Porod invariant
depends only on contrast



? M ↔ V

~~Svergun: M=V/1.6~~ M: kDa V: nm³

$v_p V/1.6 < M < V/1.6$

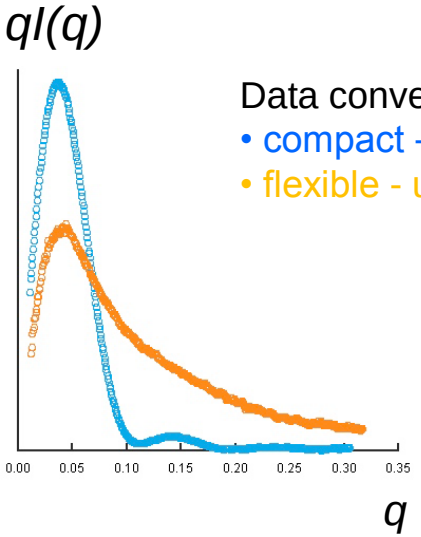
Program SAXSMow
http://www.if.sc.usp.br/~saxs/

Volume-of-correlation *Rambo R.*

$$V_c = \frac{I(0)}{\int q \cdot I(q)\, dq}$$

M ↔ V

Program SCÅTTER
http://www.bioisis.net/

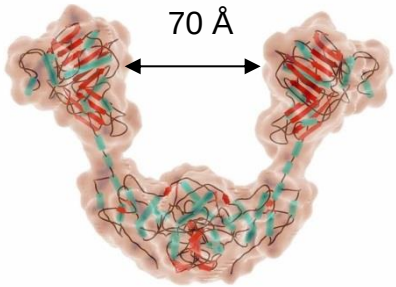


Data converges for both
- compact - folded
- flexible - unfolded

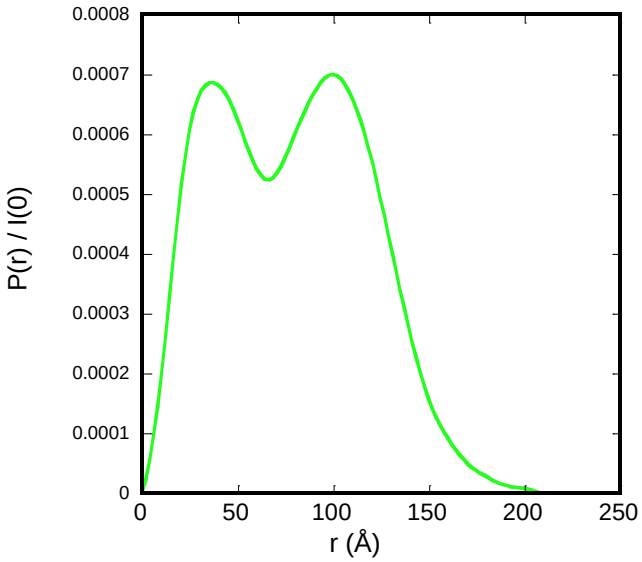
Distance distribution function

P(r) depends on the shape of the particle

Dimère



Topoisomerase VI

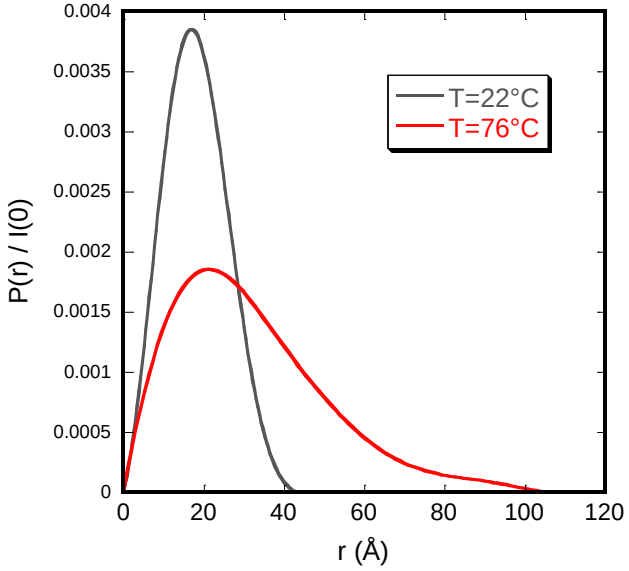


Unstructured protein

Neocarsinostatin

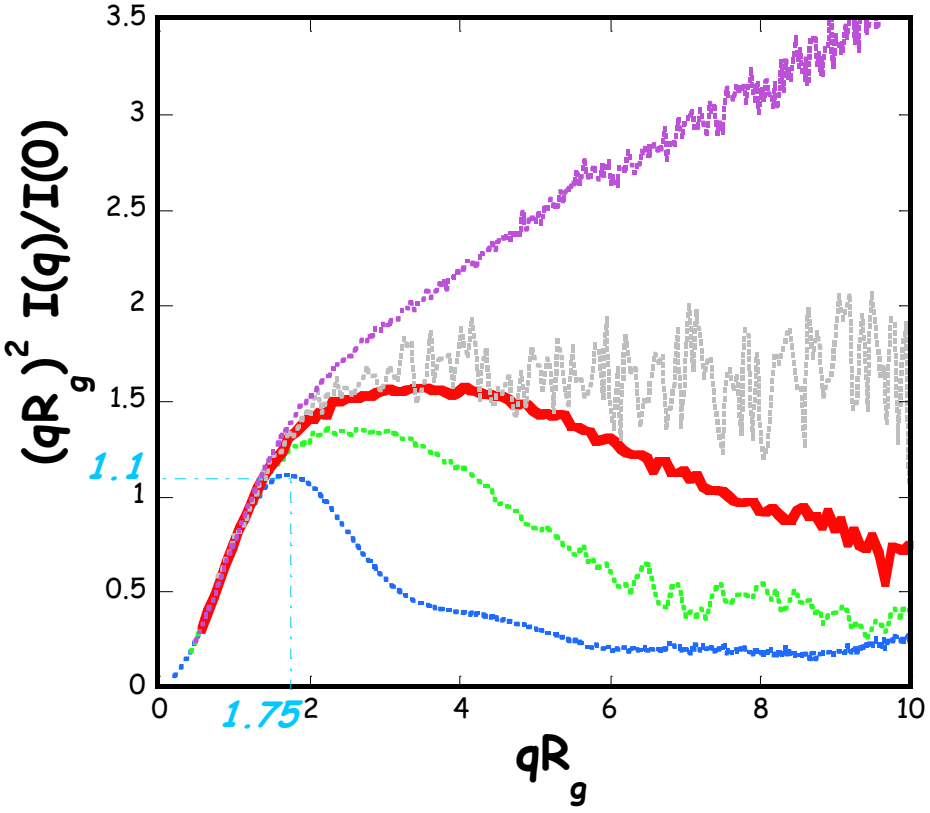
T=22°C

T=76°C

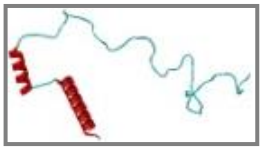


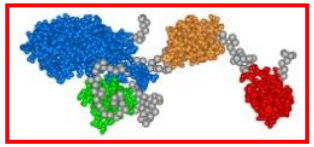
Dimensionless Kratky representation

Fully unstructured proline-rich protein (salivary protein IB5)
Boze et al, Biophys.J 99(2010)656

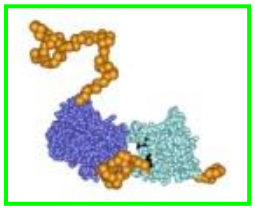


Unstructured protein with some residual secondary structure elements (C-XPC)
Miron et al, Biochemistry 47(2008)1403





« beads on a string »



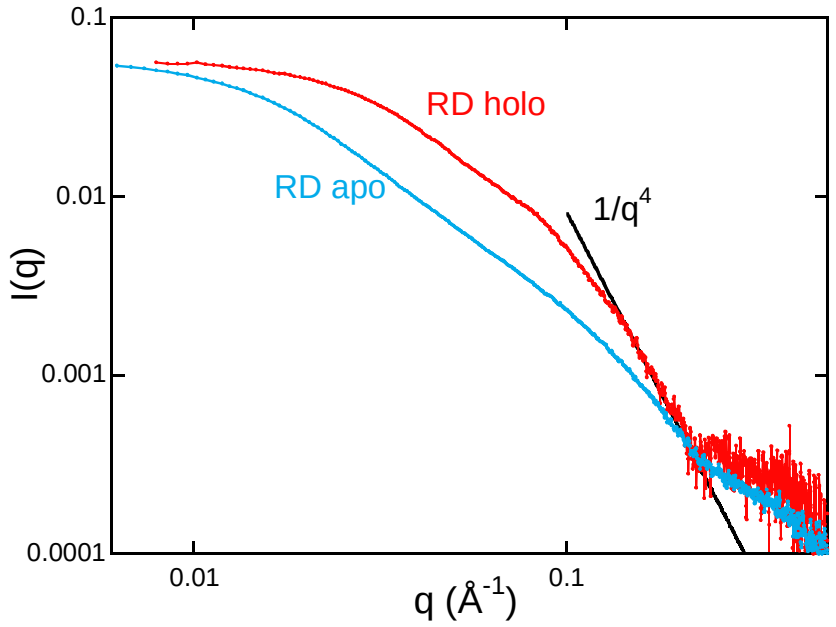
Modular protein p47 with interactions between both domains
Durand et al, Biochemistry 45(2006)7185

For structured globular proteins all curves are nearly superimposed in the range 0≤qRg≤3

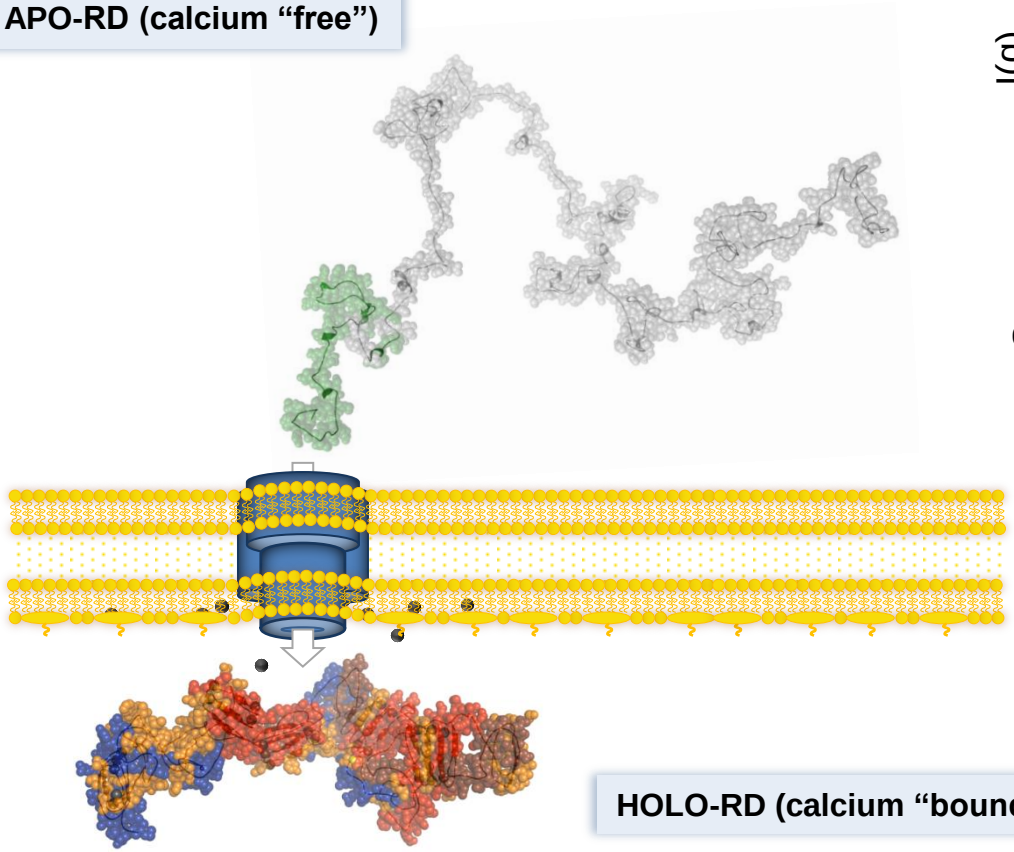
Polymerase PolX$_{Dr}$

Using Porod law for unstructured proteins

SAXS: useful to probe the transition
flexible « spaghetti »
→structured domains



APO-RD (calcium "free")



HOLO-RD (calcium "bound")

Porod's law $1/q^4$
→ well-defined surface

Chenal et al. in preparation

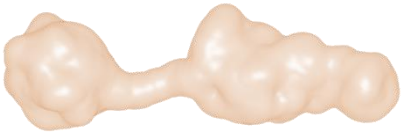
2d stage

Modelling methods

The solution is not unique

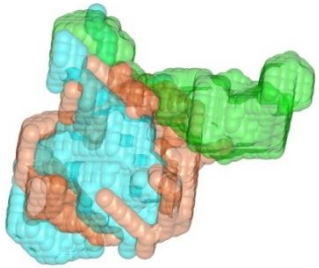
ab initio modelling

DAMMIF, DAMMIN, GASBOR



Vernhes E.,… Boulanger P., in preparation

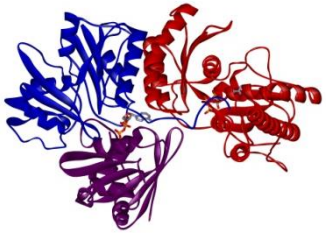
MONSA



Pilotto S.
…Mattevi A.
PNAS 2015

Rigid body analysis : quaternary structure of complexes

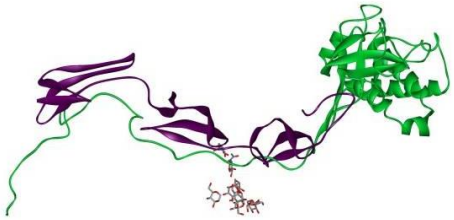
SASREF



Zhang W.
…van Tilbeurgh H
NAR 2015

Rigid body analysis coupled with addition of missing fragments

BUNCH, CORAL



Tiouajni M
…van Tilbeurgh H
FEBS J. 2014

Ensemble methods

EOM
Svergun

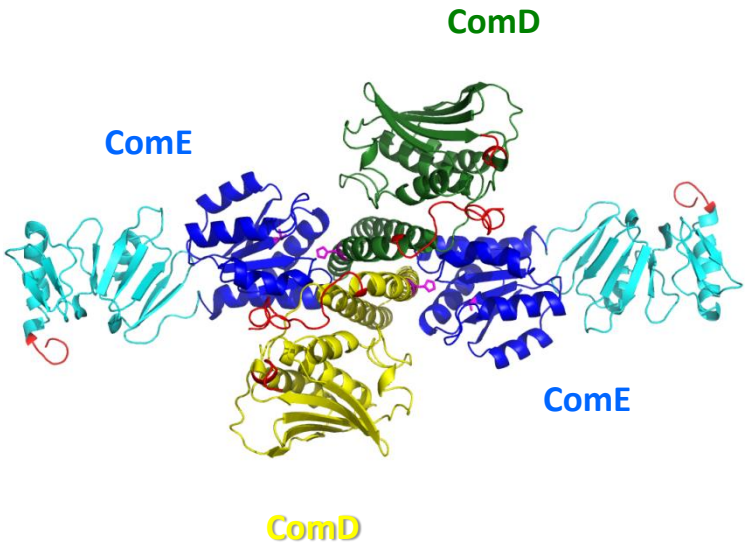
MES
Pelikan et al. (Tainer)

BSS-SAXS
Makowski, Roux

EROS
Hummer

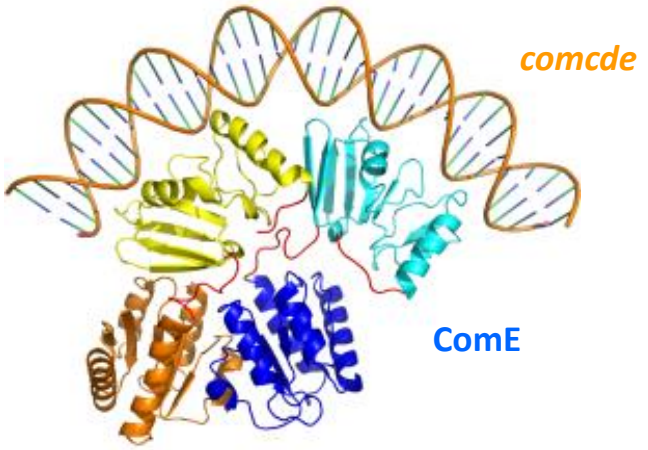
❑ SAXS for determining structures of complexes

S. Quevillon-Cheruel
M. Boudes
D. Sanchez

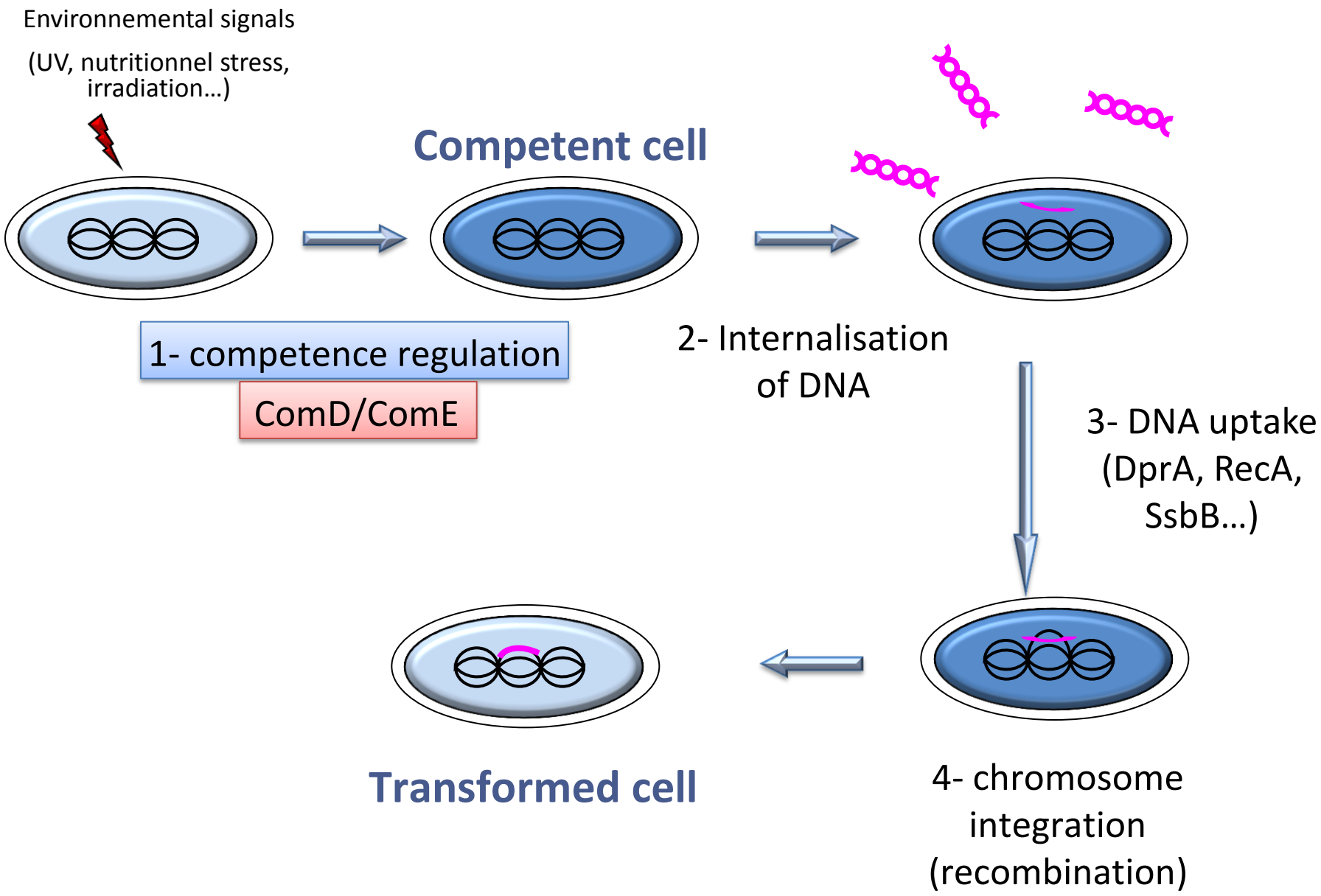


M. Boudes…S. Quevillon-Cheruel, NAR (2014)
D. Sanchez…S. Quevillon-Cheruel, FEBS J. (2015)

No crystals!

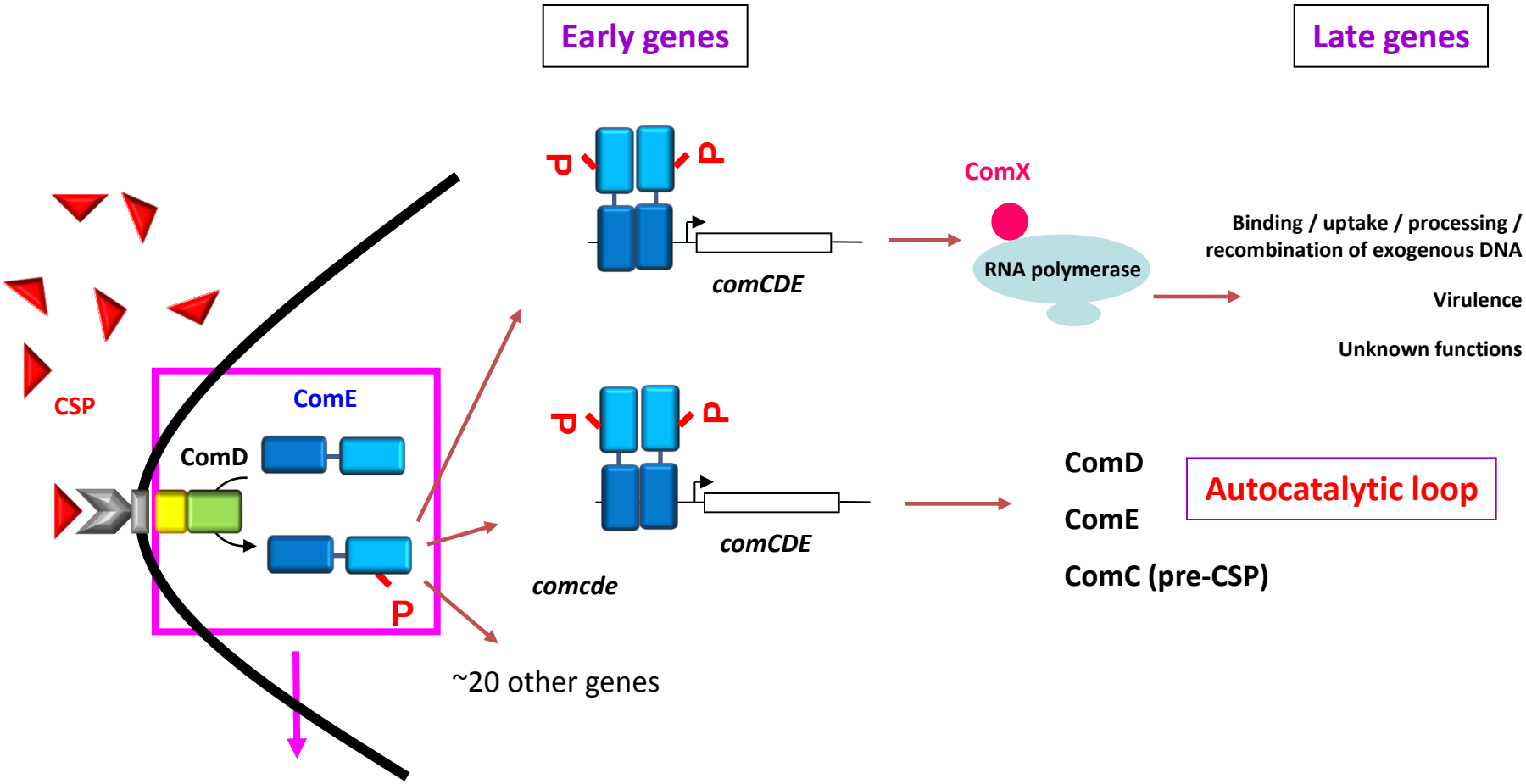


The natural genetic transformation



Mortier-Barrière I et al., *Cell*, **2007** 130:824-36

Induction of the competence



Two-component system

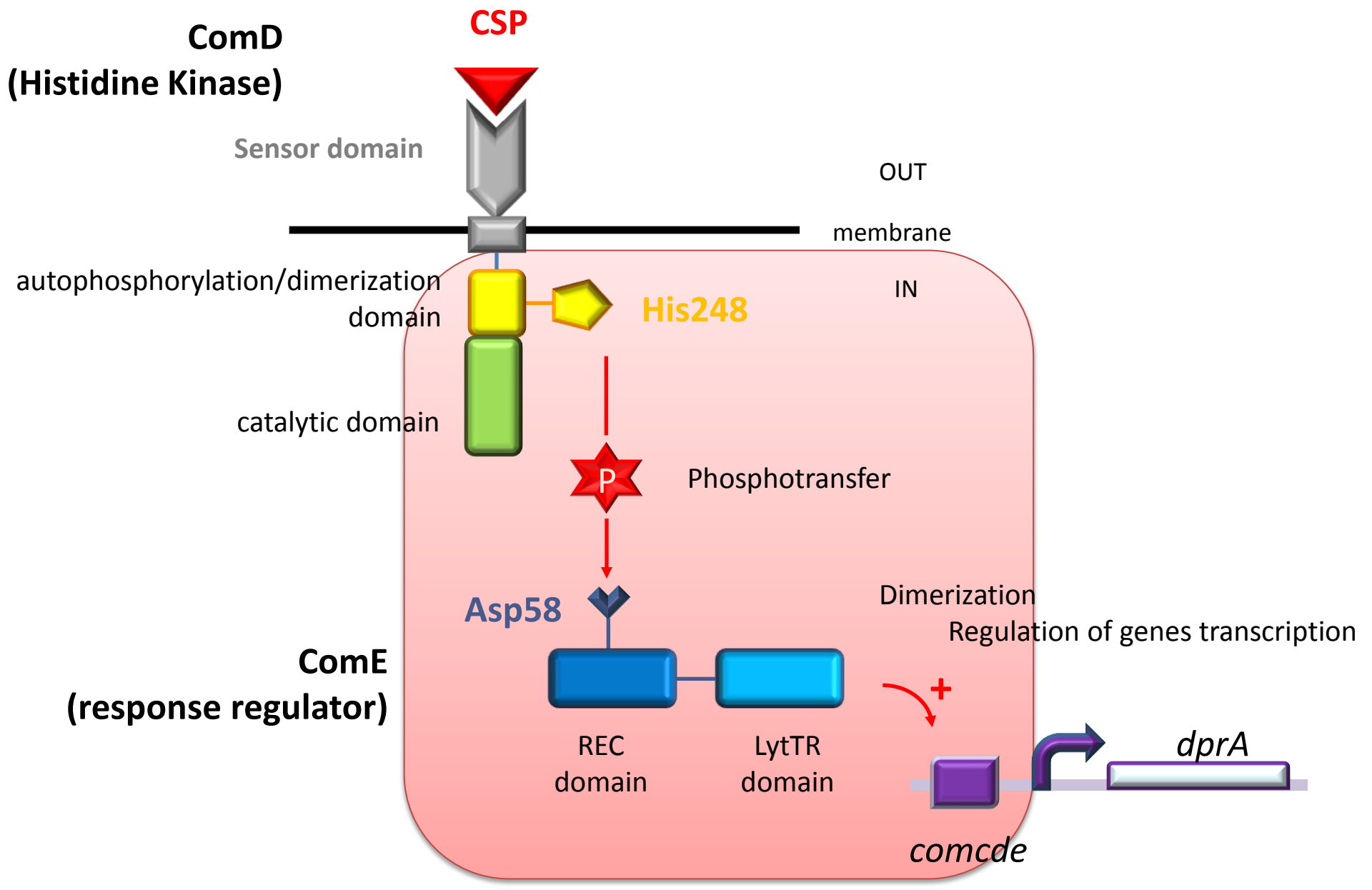
= prevalent strategy for coupling environmental stimuli to adaptive responses in bacteria

Histidine kinase (ComD) + response regulator (ComE)

Courtesy of S. Quevillon-Cheruel

Gao R et al. *Annu Rev Microbiol* **2009**, 63:133-54

The ComD/ComE Two-component system

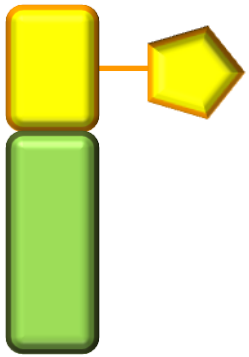


Proteins, mutants and DNA…

ComD

Autophosphorylation
dimerization
domain

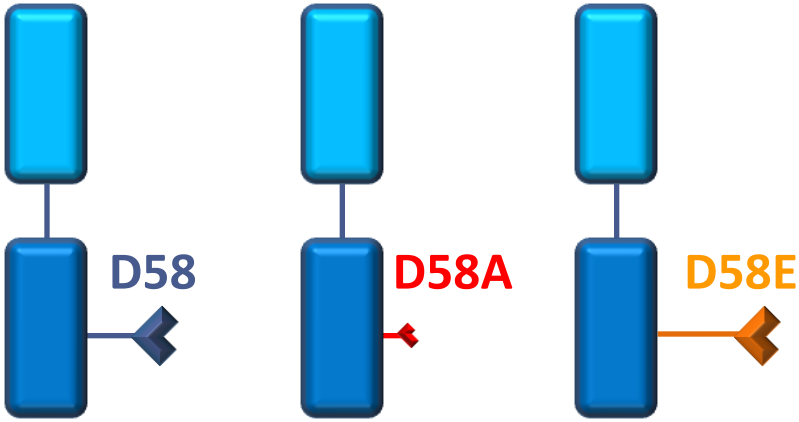
catalytic domain



ComE

LytTR
domain

REC
domain

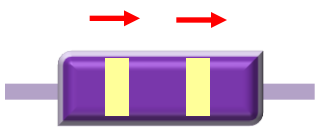


WT **inactive form** **active form**

Phosphomimetic
***In vivo*: constitutively competent**

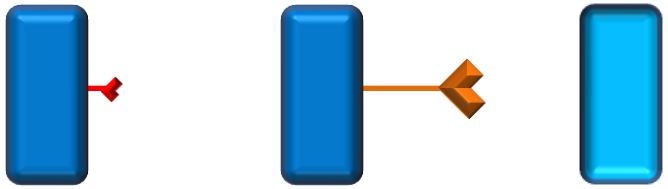
Martin et al. *Mol Microbiol*. 2013 Jan;87(2):394-411.

comcde

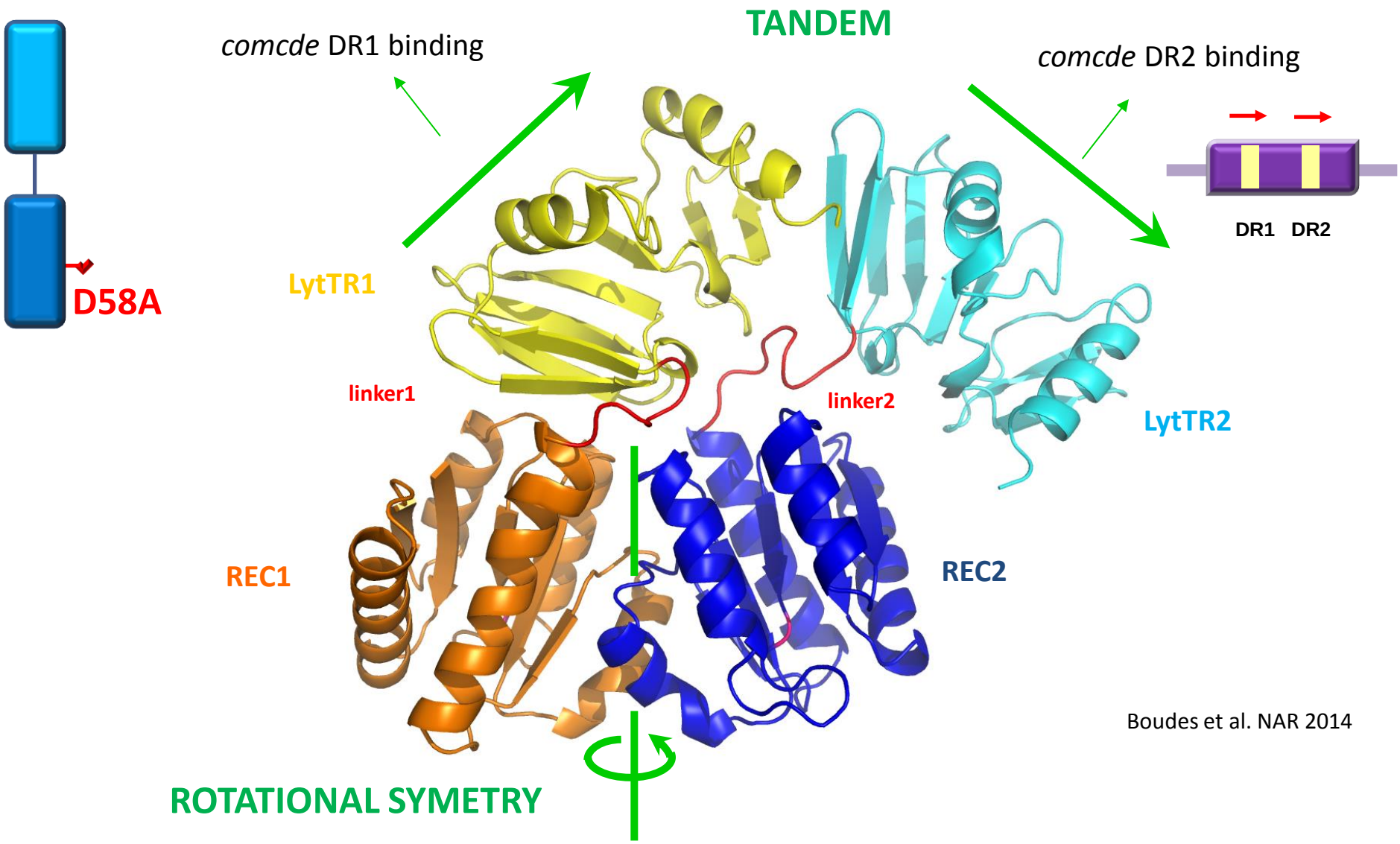


dsDNA

Isolated domains



The crystallographic dimer is asymmetric

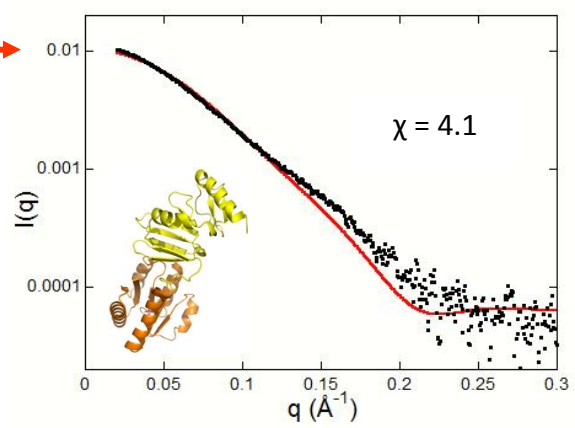


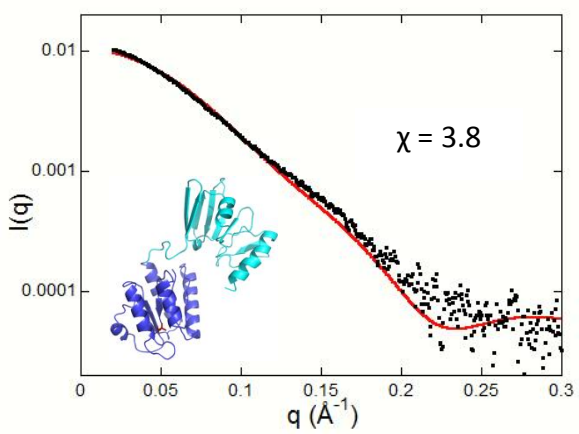
Boudes et al. NAR 2014

What is the oligomerisation state of ComE in solution? ➡ **SAXS**

ComE^{D58A}

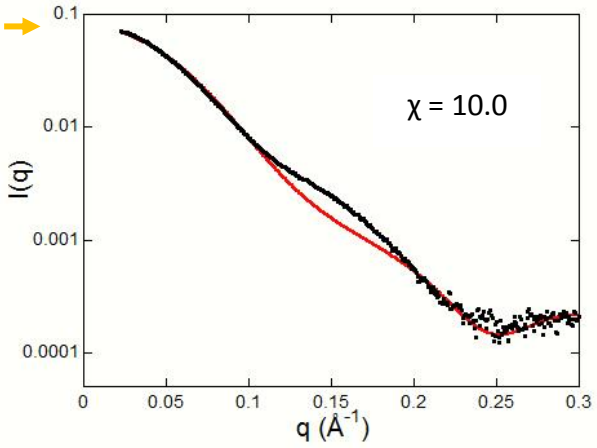
M monomer →

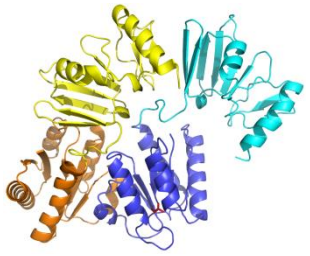




ComE^{D58E}

M dimer →



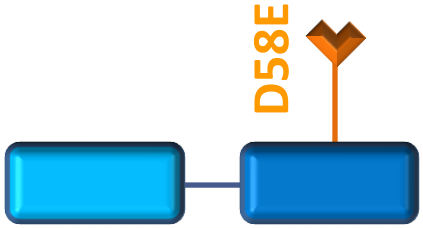


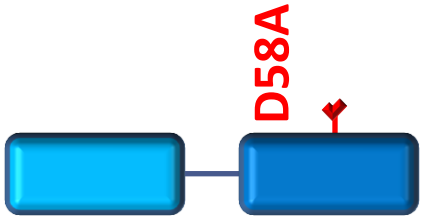
• experimental data

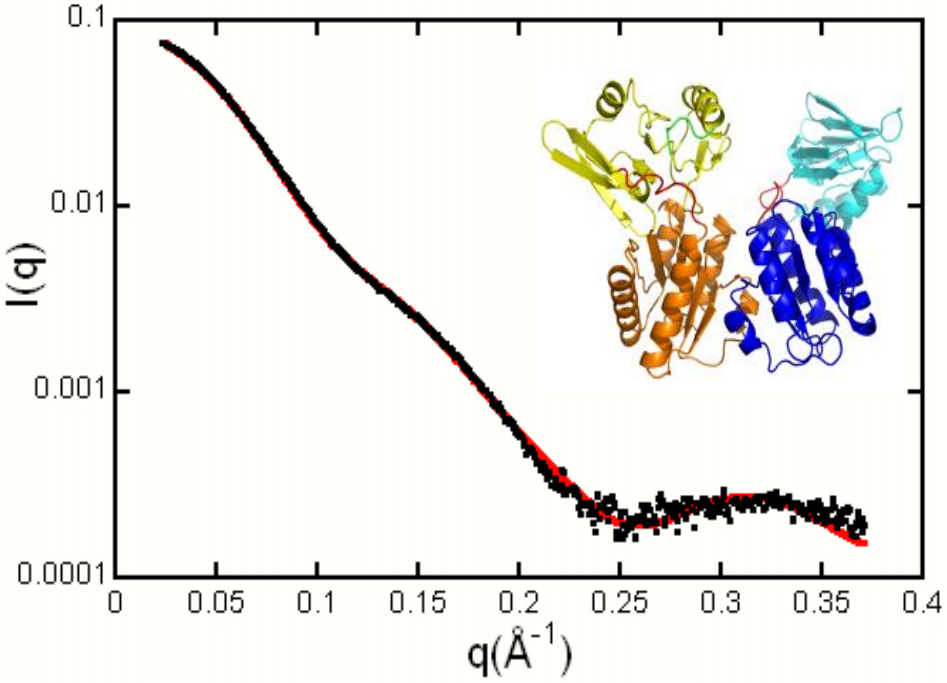
Calculated curves using **CRYSOL**

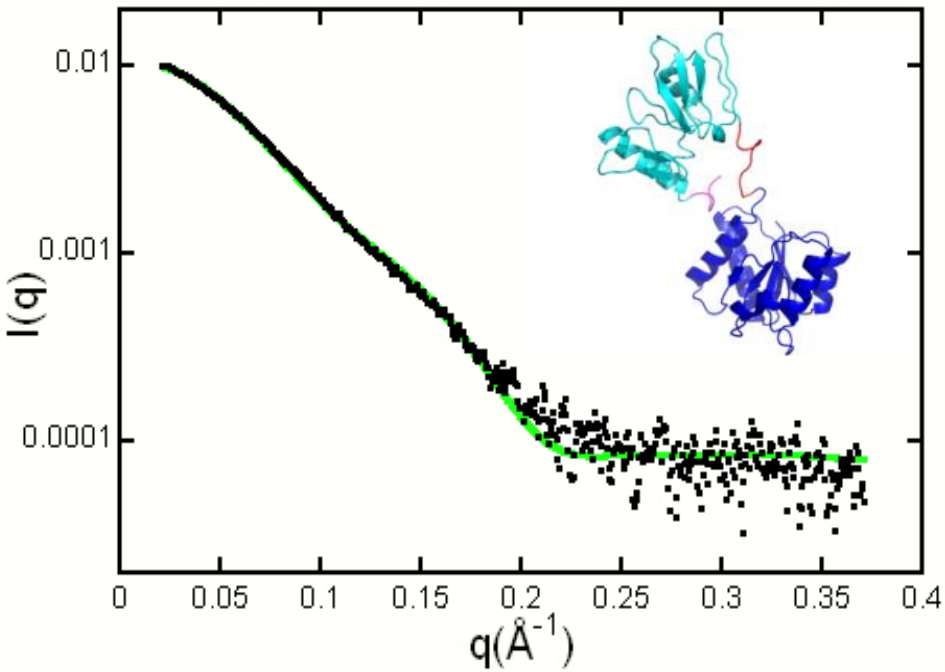
conformations in solution ≠ conformations in crystal

Rigid-body modelling: BUNCH
Ensemble Methods: EOM









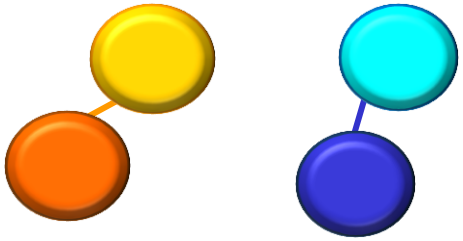
Dimer via REC

Monomer

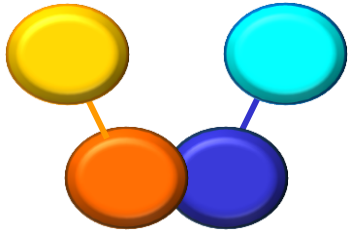
Flexibility between the REC and the LytTR domains in both cases

To Summarize…

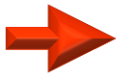
ComE-D58A (as well as the WT) is monomeric in solution

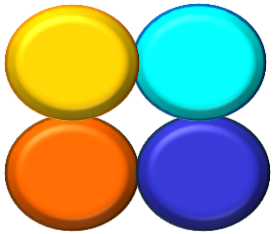


ComE-D58E is a dimer – but not the crystallographic one



Did the crystallographic dimer mimick the interaction with *comcde*?

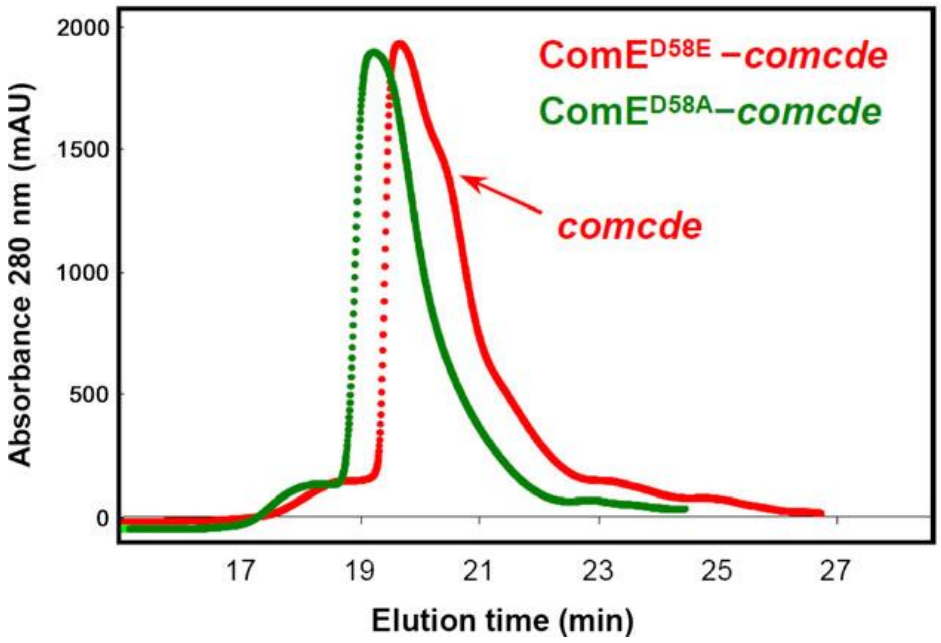




SAXS on ComE-D58A and D58E /*comcde* complexes

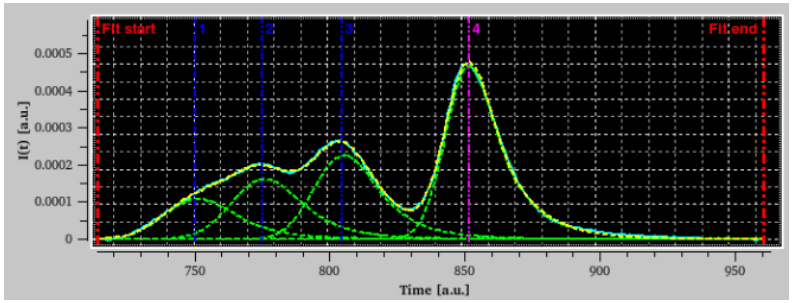
SAXS study of complexes ComE-*comcde*:

requires the use of on-line HPLC



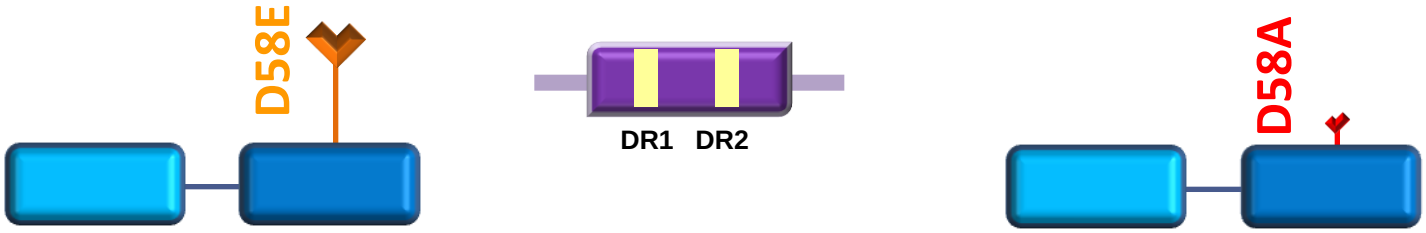
When peaks are not very well resolved : profile deconvolution using gaussians.

HPLC-SAXS : a module within the US-SOMO software.

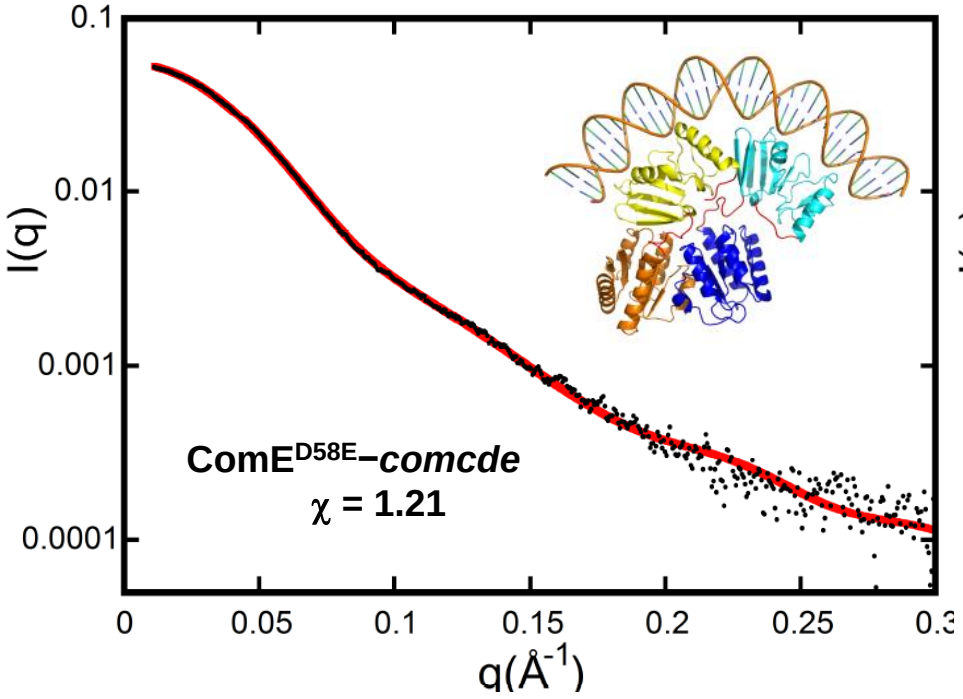


E. Brookes, M. Rocco, J. Pérez, P. Vachette, J. Appl. Cryst. (2013)

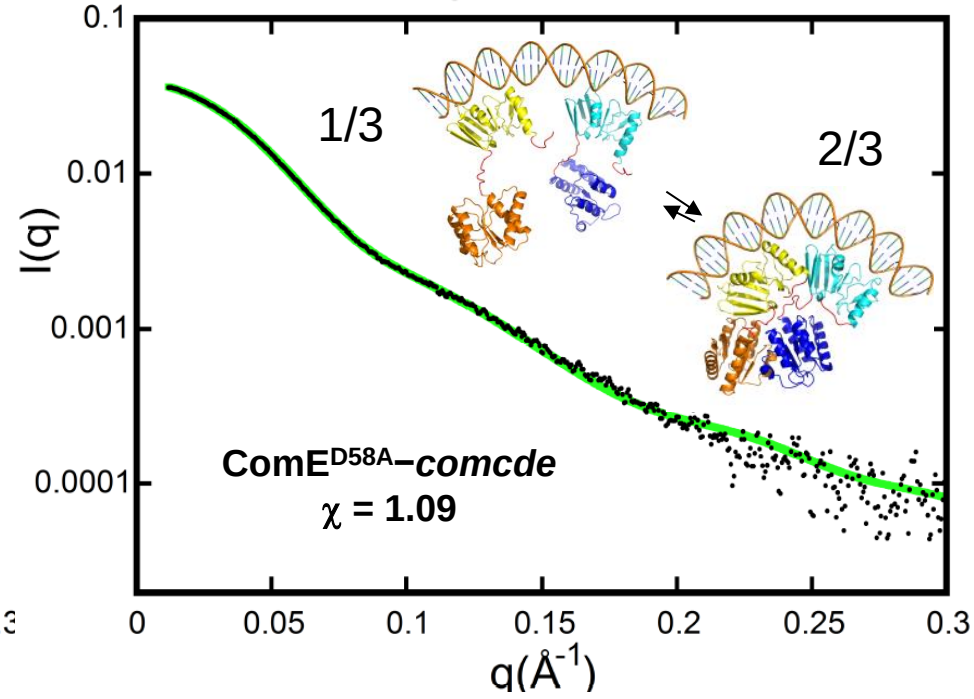
ComE-D58E and D58A binding on the *comcde* promoter



CRYSOL



EOM, OLIGOMER

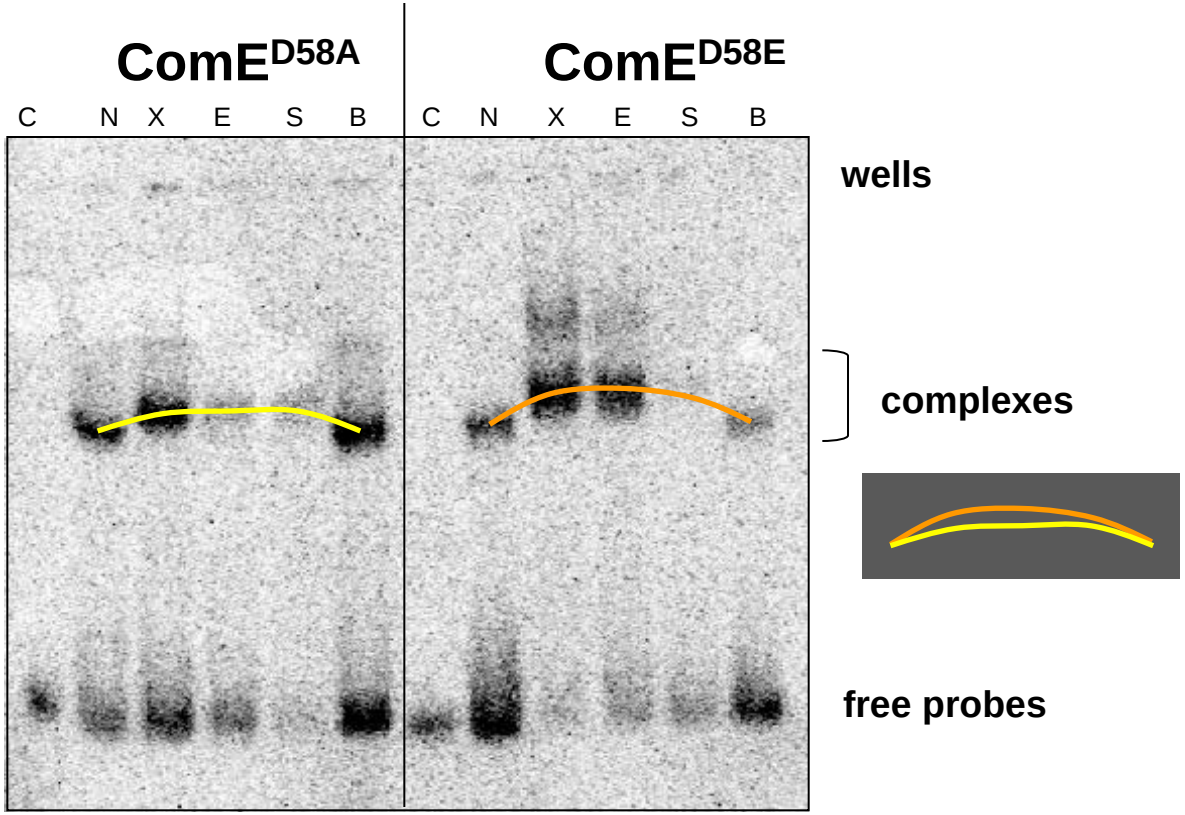


Two monomers of ComE bind a *comcde* promoter

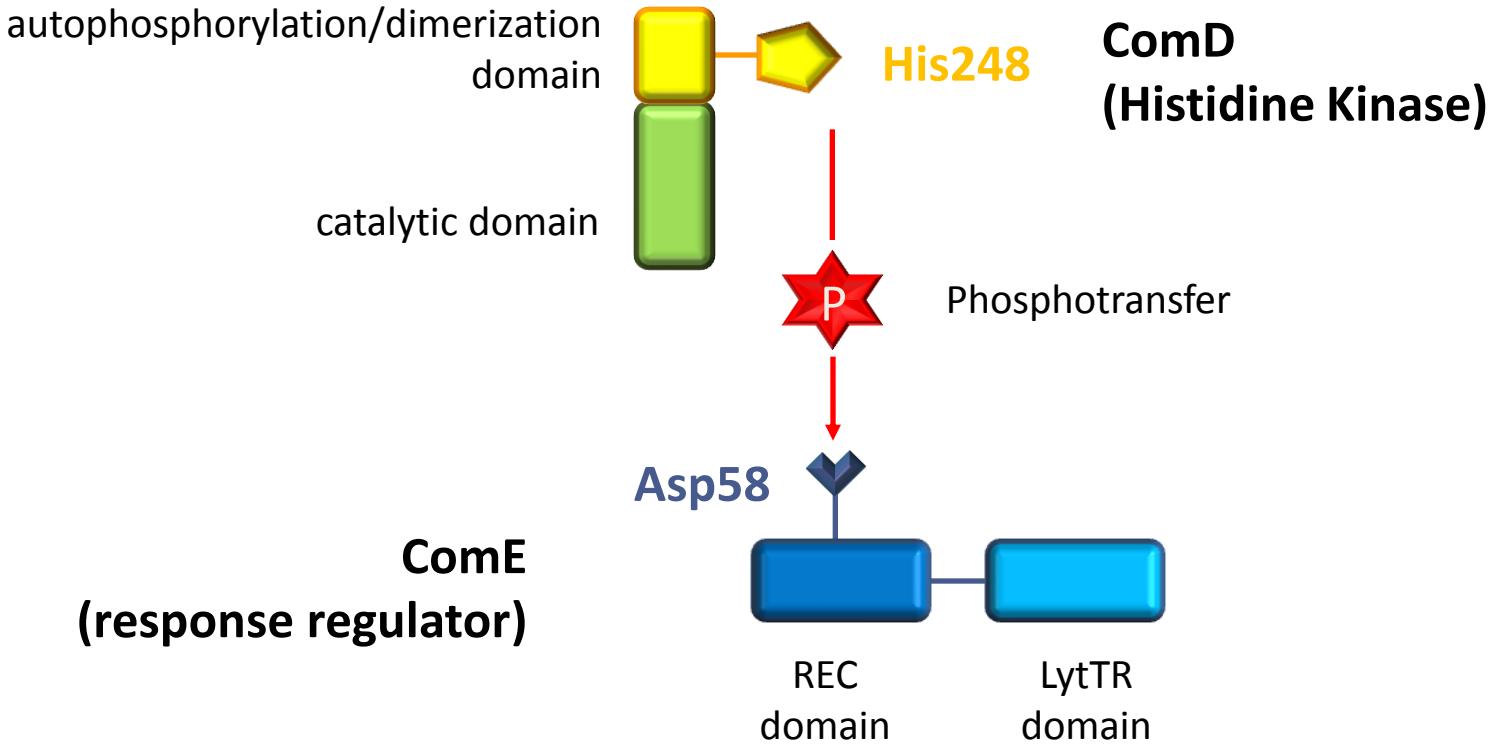
but with various bending

Biochemical validation of the various degrees of DNA bending

Kim et al. *Gene* (1989) 85, 15-23



What about the ComD-ComE interaction? The phosphotransfer?

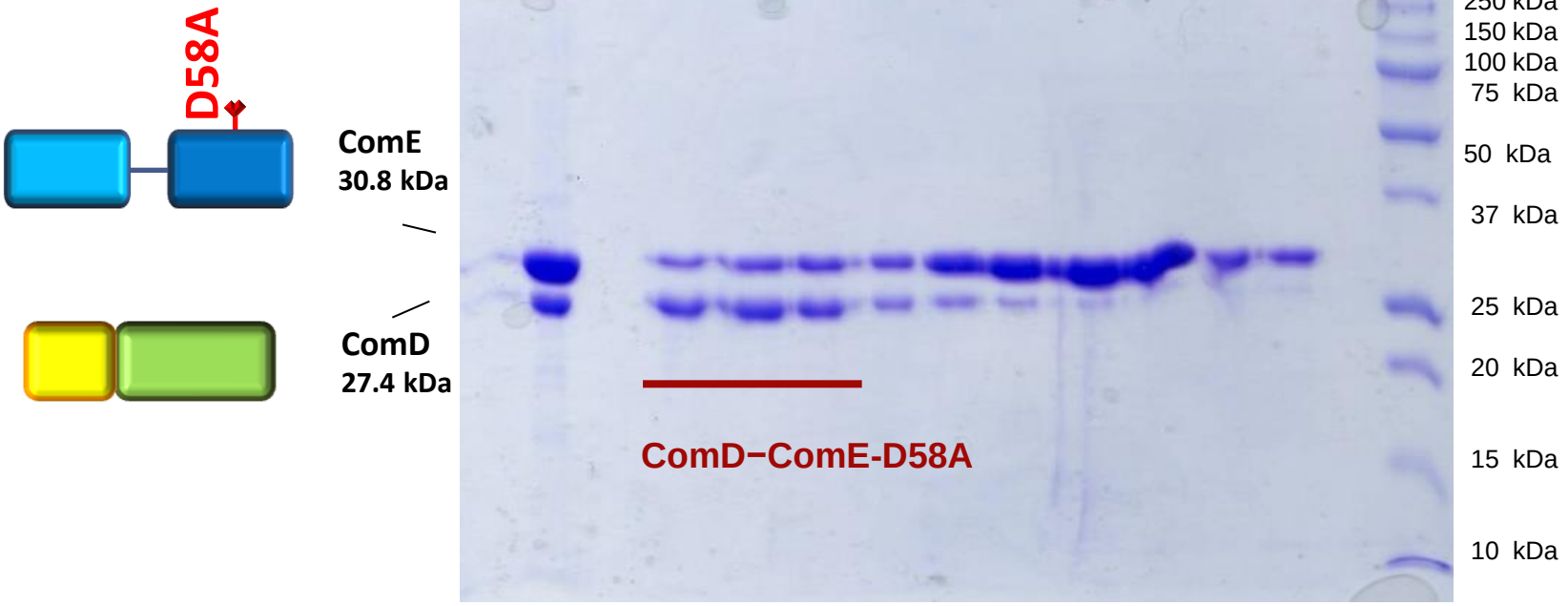


Reconstitution of the ComD-ComE complex

ComD: Insoluble/instable when expressed alone

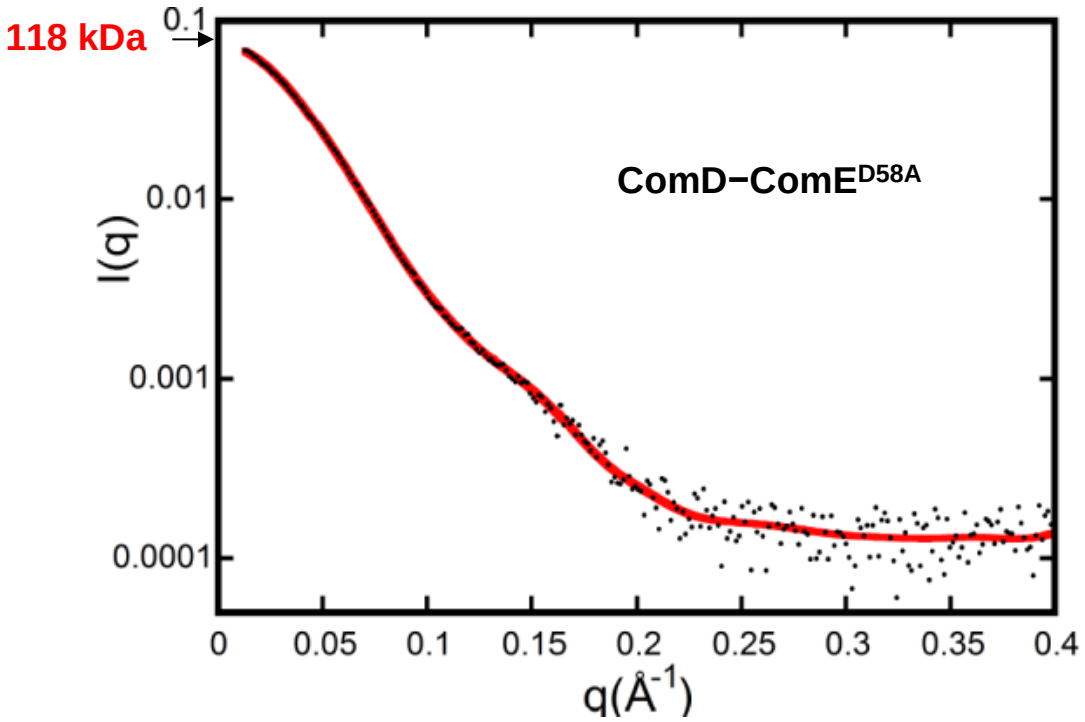
Stabilisation via its interaction with ComE :

- stable with ComE-D58A, transient with ComE-WT and non observed with D58E



No crystals → **SAXS**

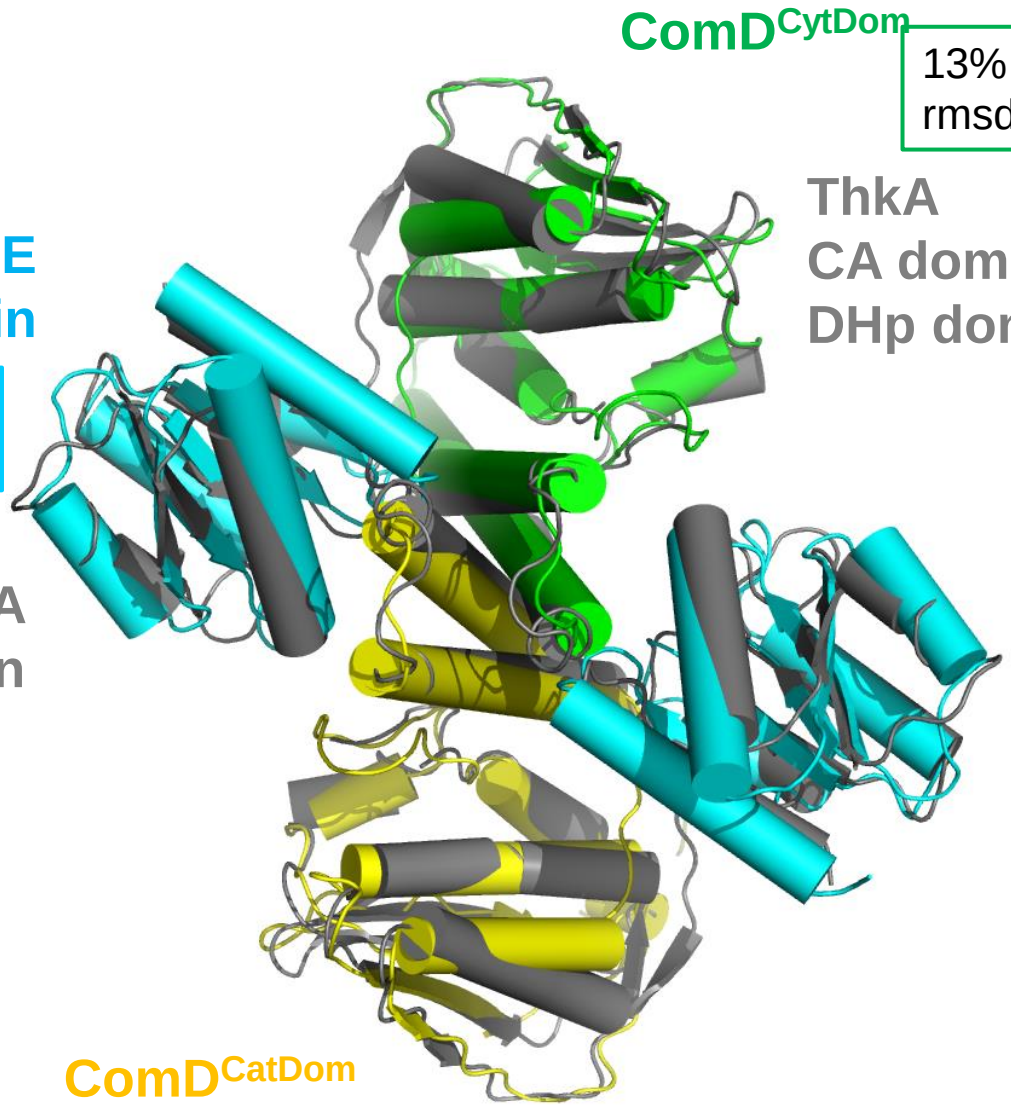
ComD-ComE → heterotetramer



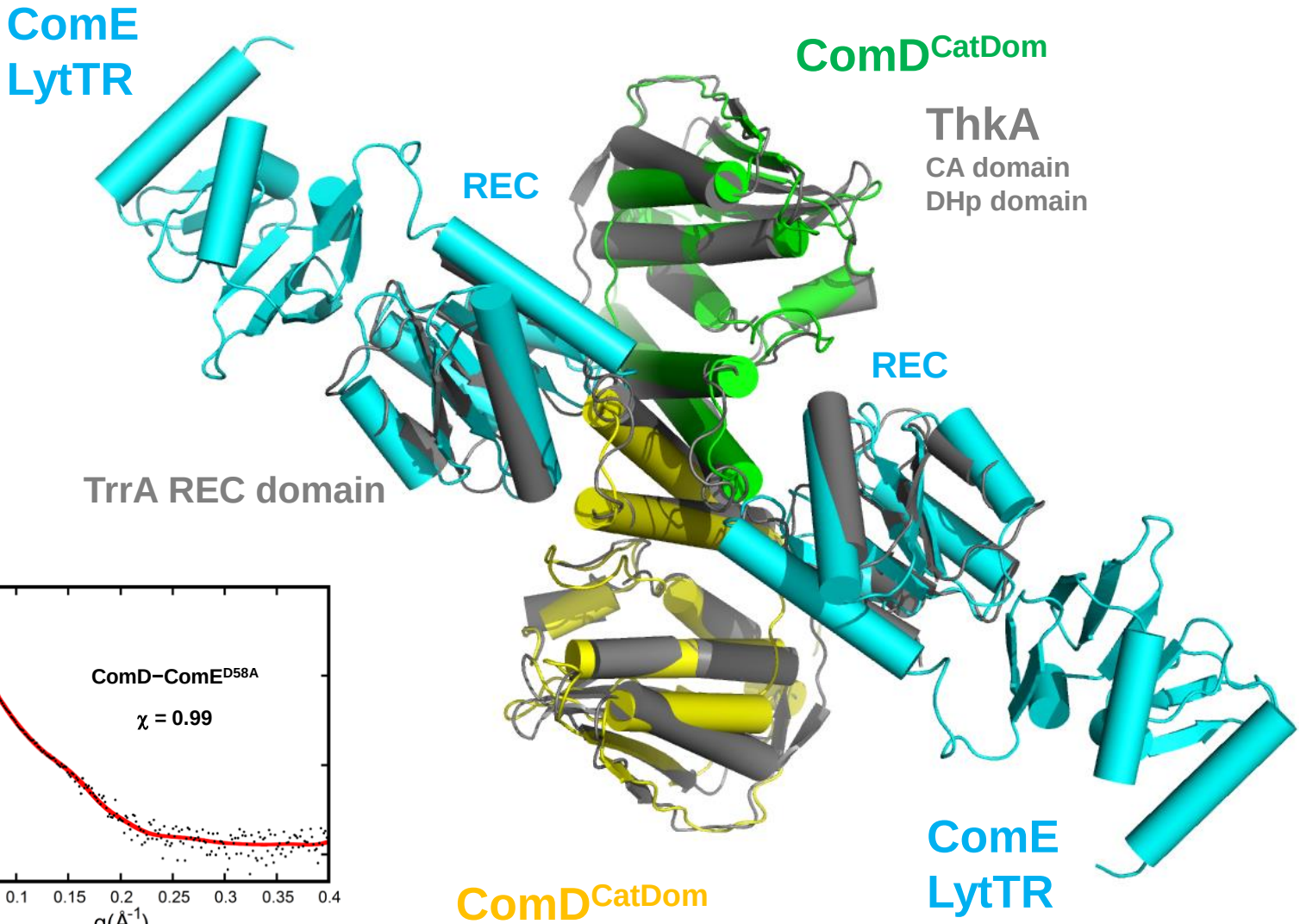
2 ComD + 2 ComE = 114 kDa

26kDa **31kDa**

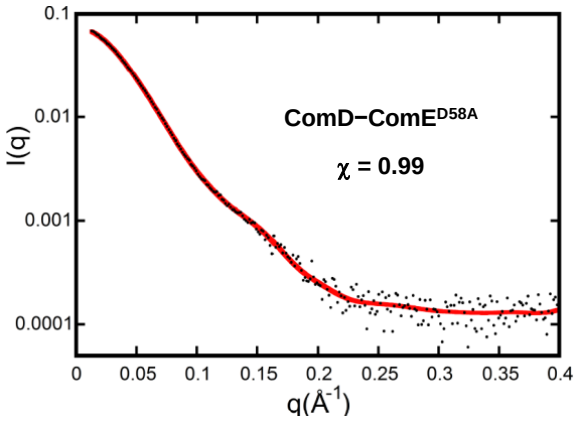
Modelisation of the ComD-ComE complex using ThkA-TrrA (3A0R)



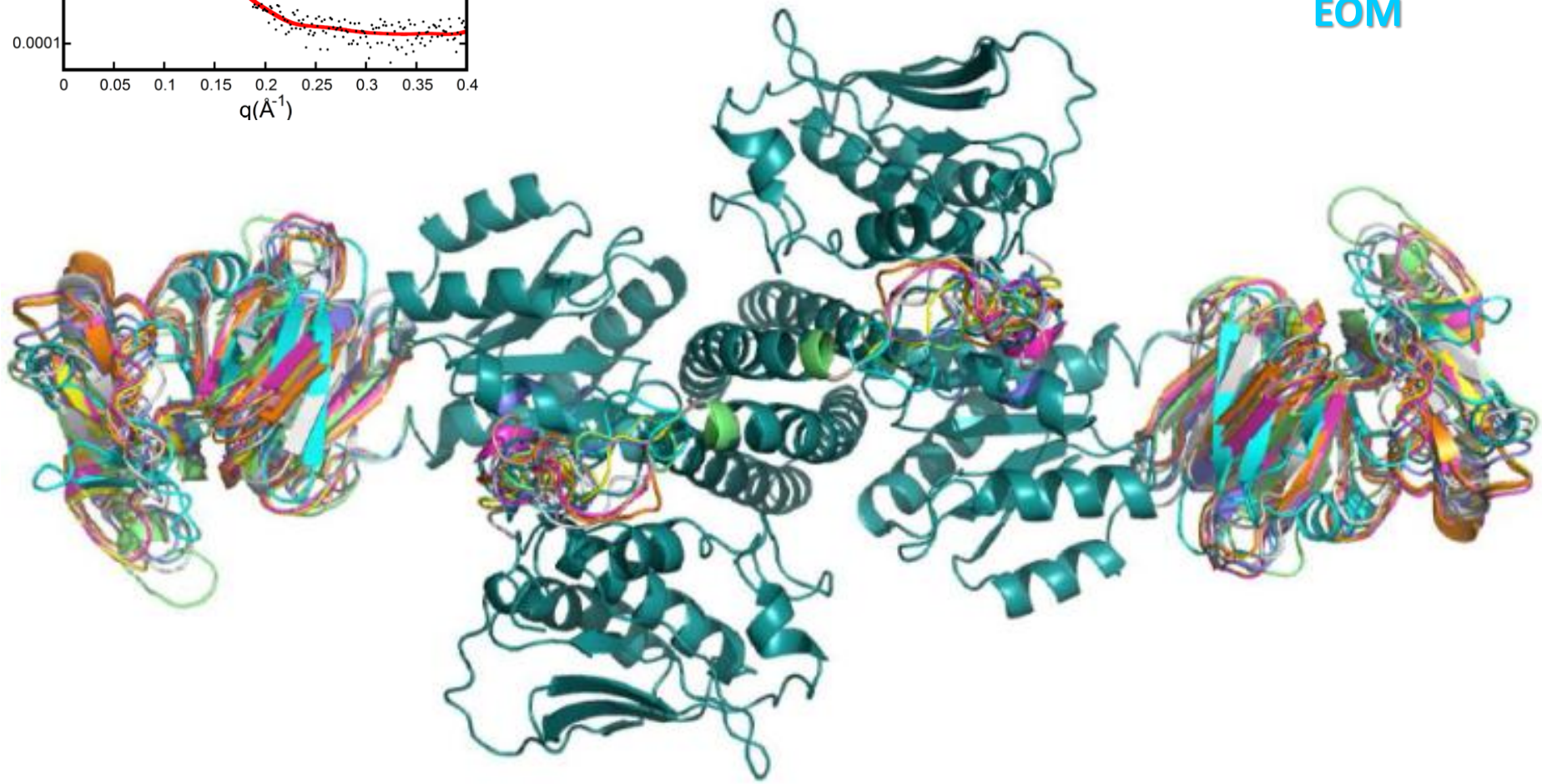
Addition of the LytTR domains



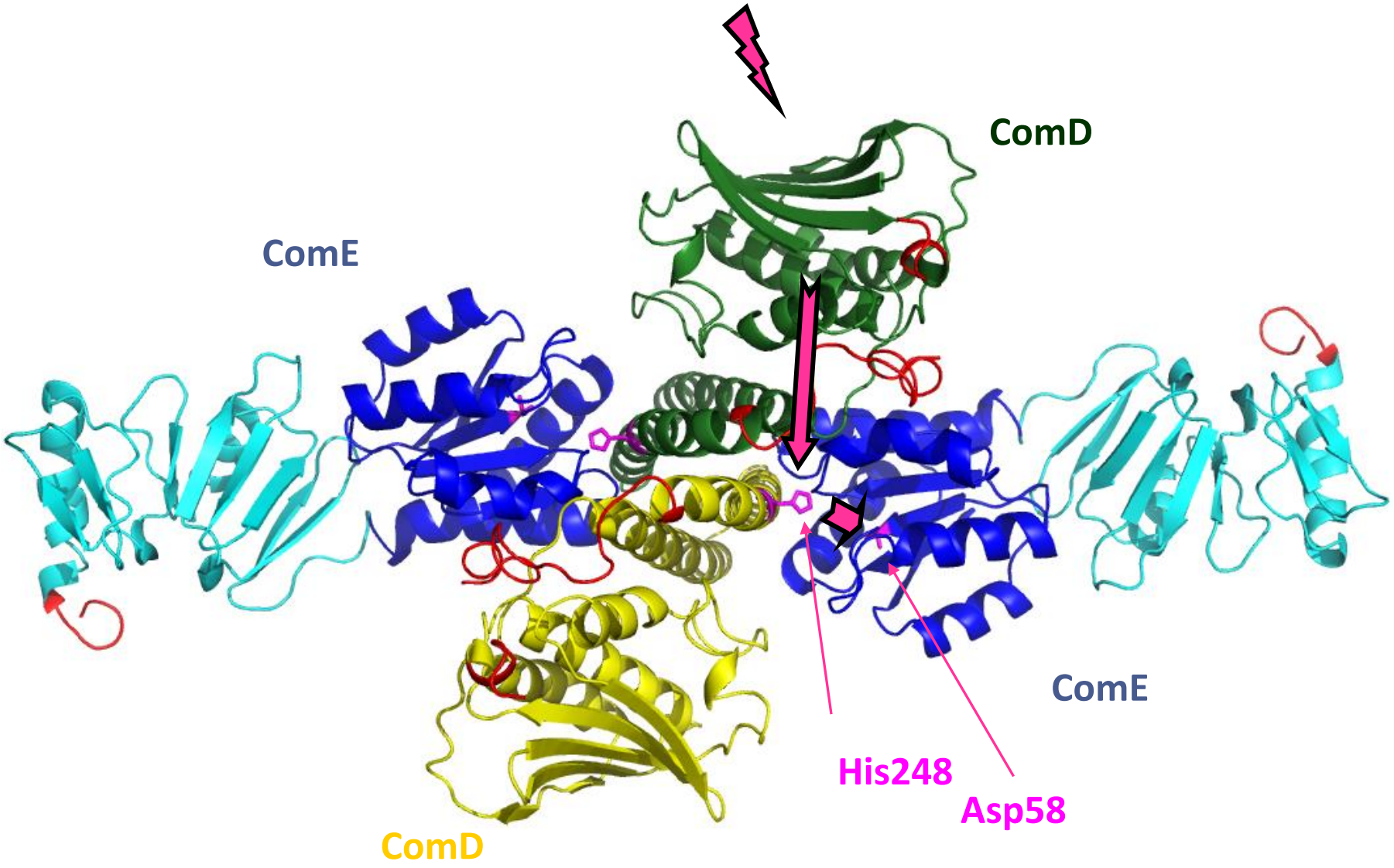
Final models: flexible regions
LytTR and HisTags



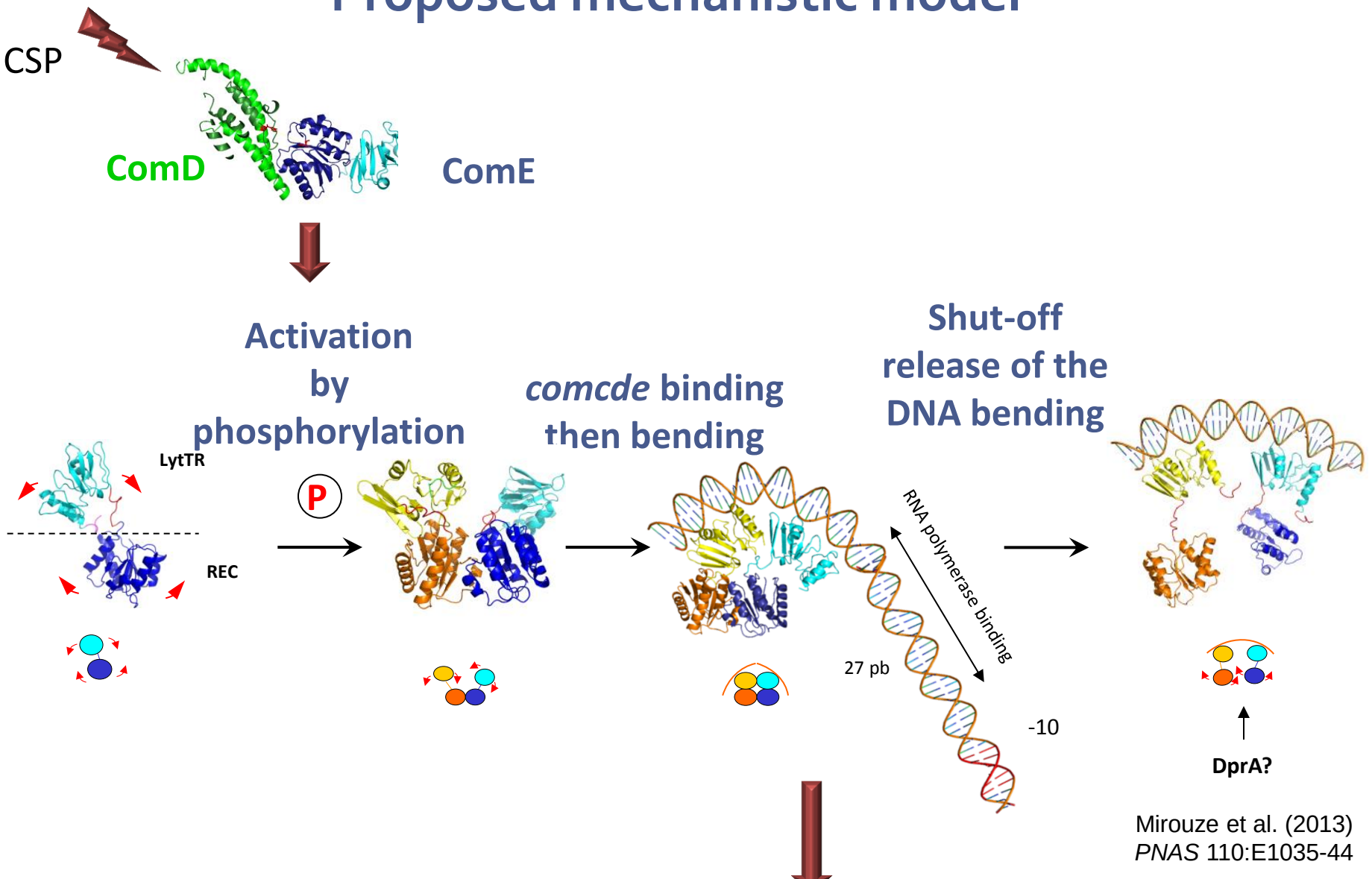
BUNCH
CORAL
EOM



Trans-phosphorylation activation model



Proposed mechanistic model



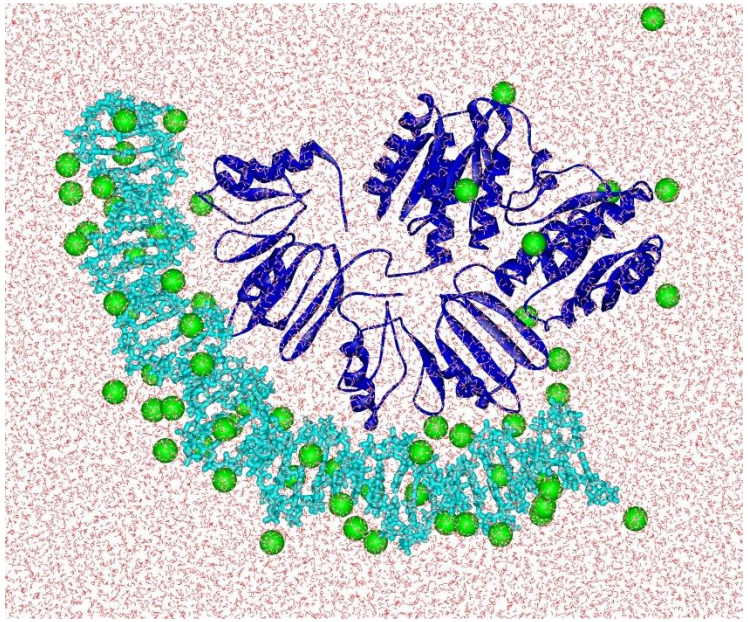
Biophysical Journal Volume 107 July 2014 435–447

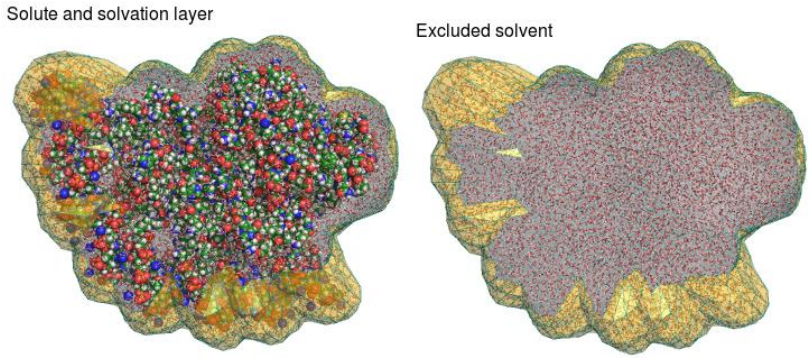
Article

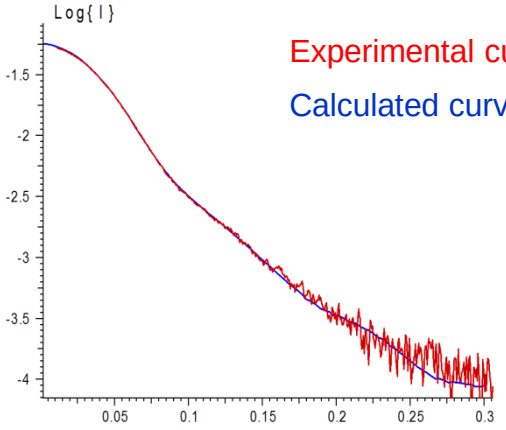
Validating Solution Ensembles from Molecular Dynamics Simulation by Wide-Angle X-ray Scattering Data

Po-chia Chen[1] and Jochen S. Hub[1,*]
[1]Institute for Microbiology and Genetics, Georg-August-University Göttingen, Göttingen, Germany

Calculation of SAXS profiles from explicit-solvent molecular dynamics simulations → No free parameters→ minimizing the risk of overfitting







Experimental curve

Calculated curve using WAXSIS

No adjustable parameters!!

SWAXS-driven MD: directing the simulations into conformations satisfying the experimental data, BJ 108 (2015)2573

FAAM – B3S – I2BC

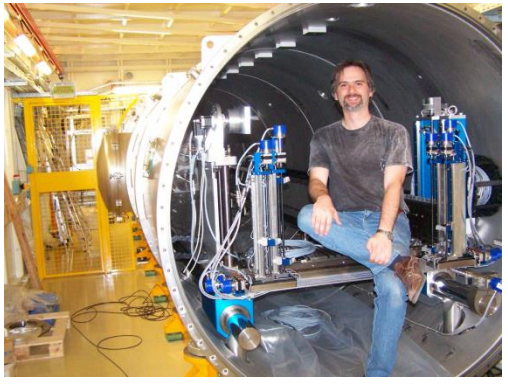
H. Van Tilbeurgh
S. Quevillon-Cheruel
B. Collinet
K. Blondeau
S. Nessler
D. Liger
I. Gallay
N. Lazar

D. Sanchez
M. Ma
A. Talagas
J. Vercruyssen
S. Ben Rejeb

Nanostar *P. Vachette – C. Mérigoux*

SAXS - WAXS - GISAXS
SWING

J. Pérez, P. Roblin, A. Thureau



ESRF
BM29

P. Pernot, A. Round