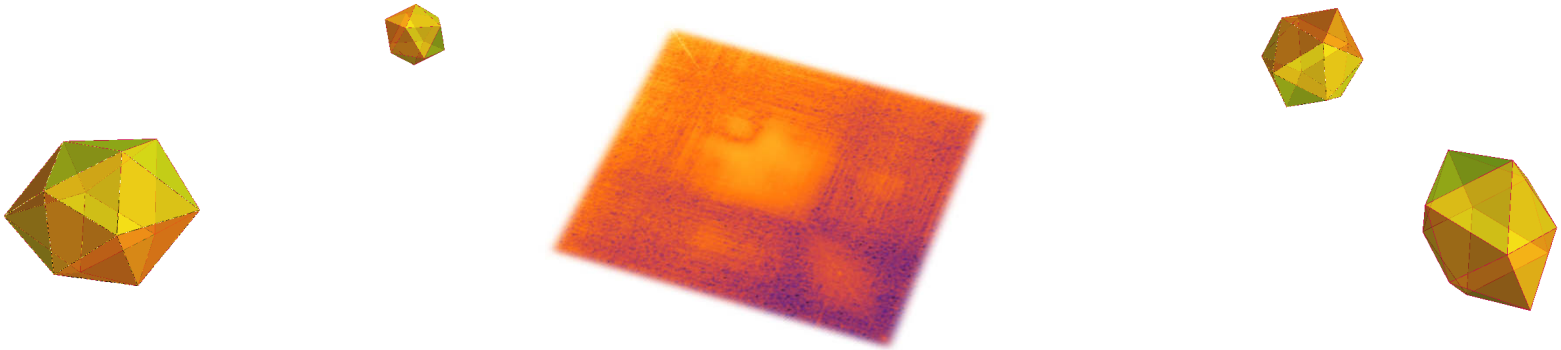
A fluctuation x-ray scattering approach for biological imaging

Ruslan Kurta, European XFEL, Germany

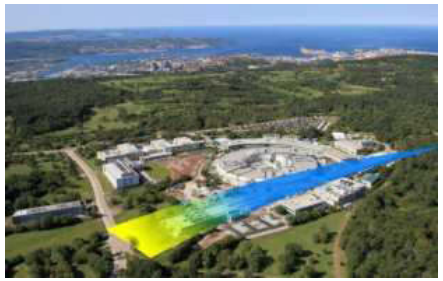


Content

- Introduction: x-ray scattering approaches for biological structure determination
  - Serial femtosecond crystallography (SFX)
  - Single particle coherent diffractive imaging (SPI)
  - Small angle x-ray scattering (SAXS)
- Fluctuation x-ray scattering (FXS)
  - Angular cross-correlation function (CCF): from SAXS to higher-order SAXS
  - Two- and three-point CCFs
- Applications of FXS
  - Single particle structure recovery from 2D disordered ensembles
  - 3D structure of viruses from single-particle imaging experiments
- Summary

X-Ray Free-Electron Lasers (XFELs)

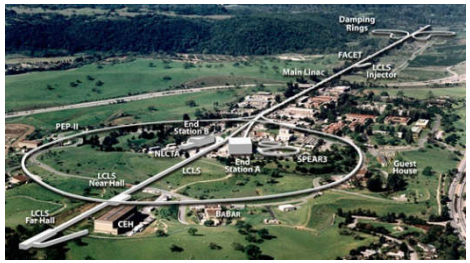
FERMI 2011



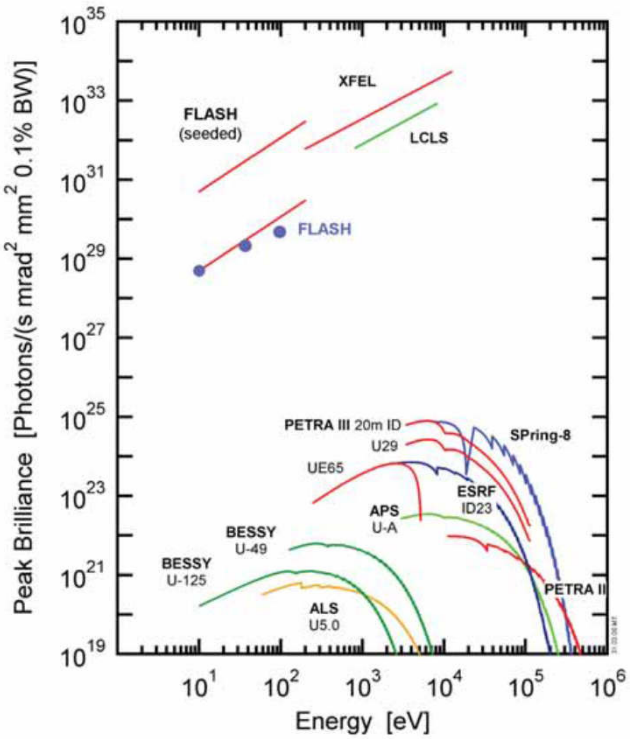
SACLA 2011

PAL XFEL 2016

LCLS 2009



FLASH 2005

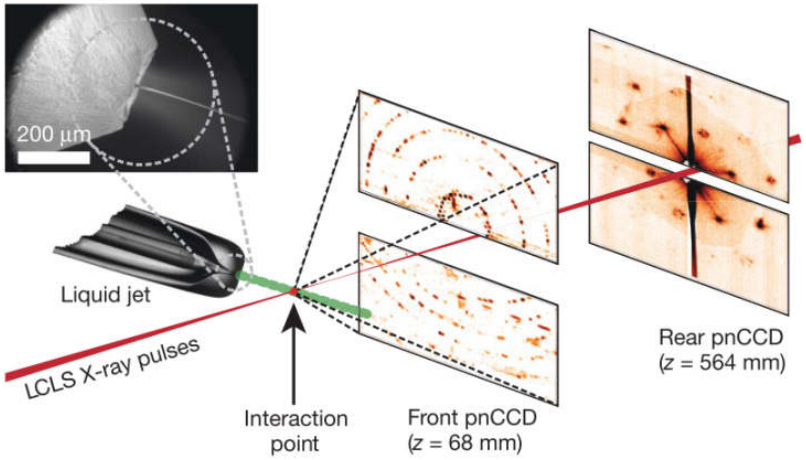


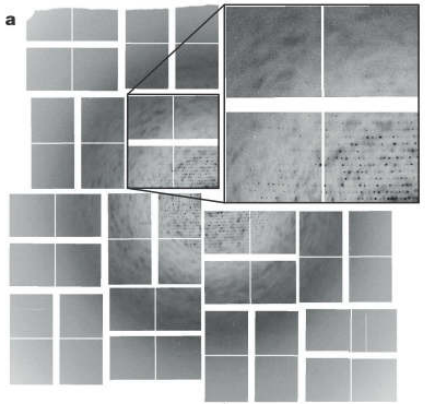
European XFEL (2017)

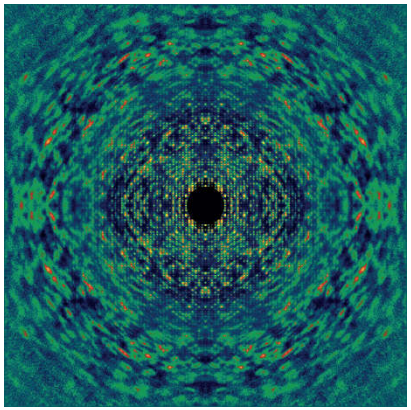
- ✓ **Unprecedented peak brilliance**
- ✓ **Ultra-short pulse duration**
- ✓ **High repetition rate**
- ✓ **High coherence**

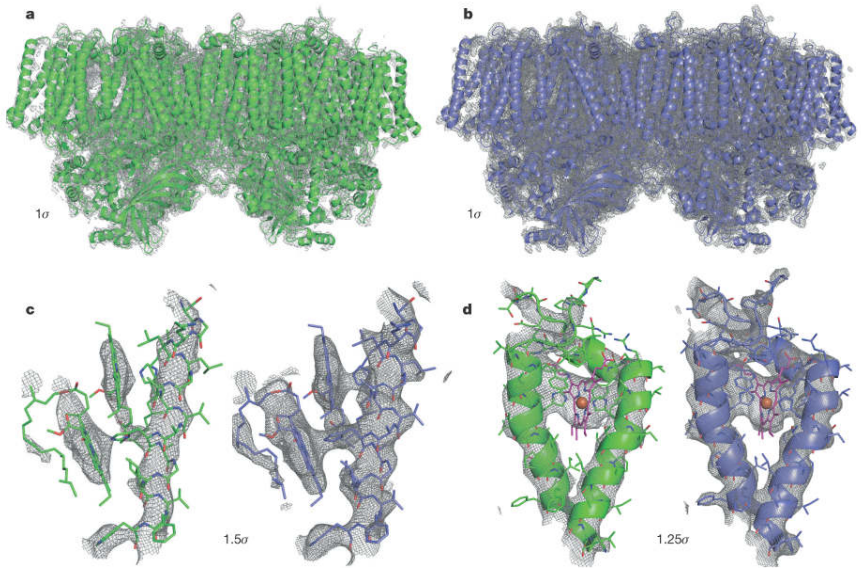
T. Tschentscher *et al., Appl. Sci.* *7*(6), 592 (2017)

Serial Femtosecond Crystallography (SFX)





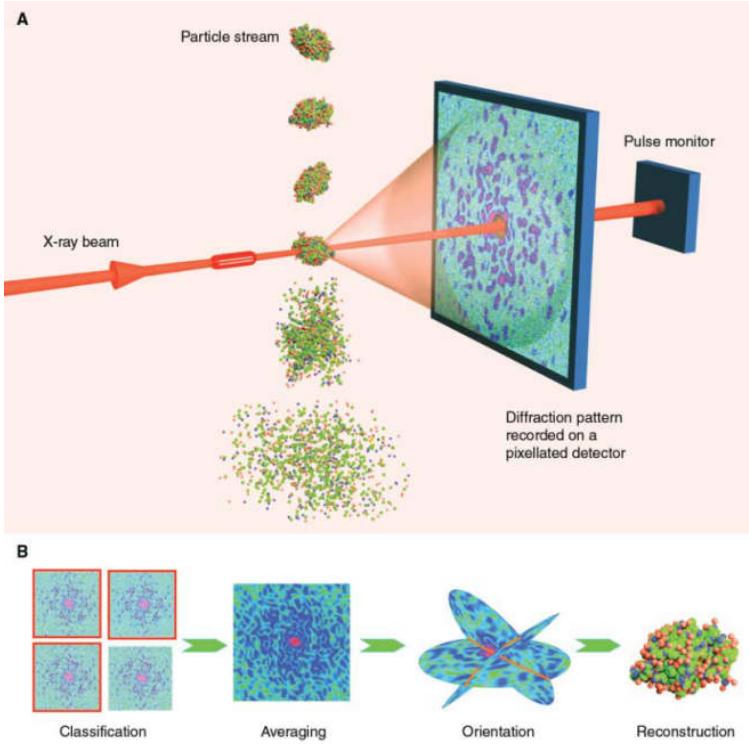




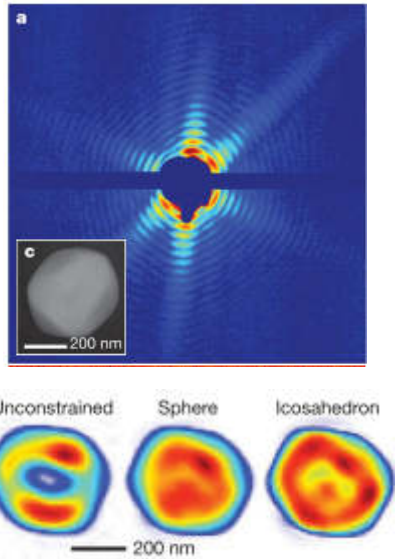
- SFX is an excellent tool for high-resolution structure studies of macromolecules and their complexes.
- Requires crystalline samples.

H. N. Chapman *et al.*, Nature 470, 73 (2011)
K. Ayyer *et al.*, Nature 530, 202 (2016)

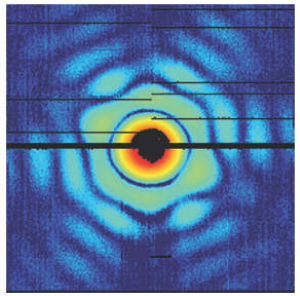
Single Particle Coherent Diffractive Imaging (SPI)

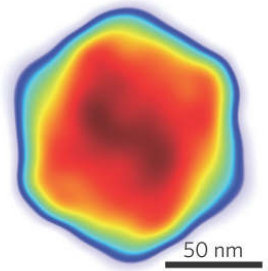


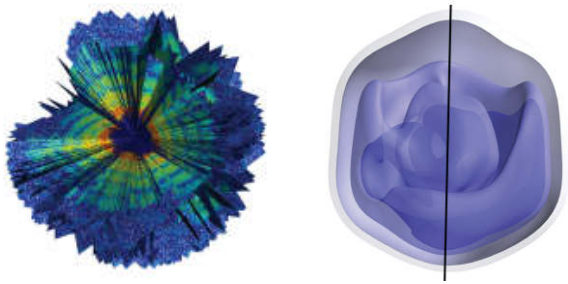
Mimivirus



Coarboxysome (cell organelle)



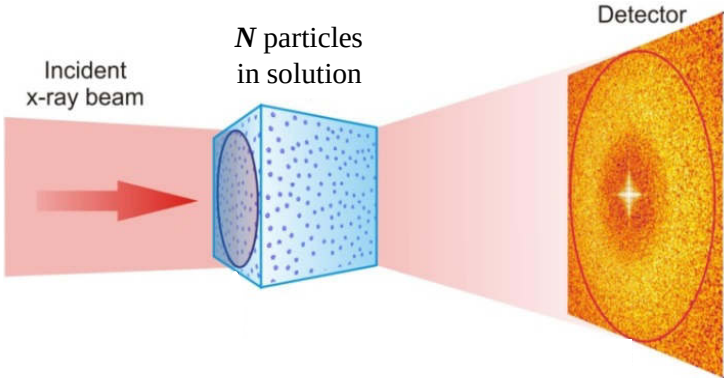


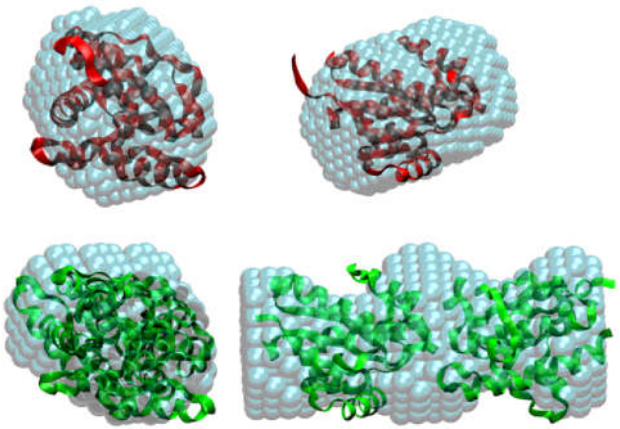


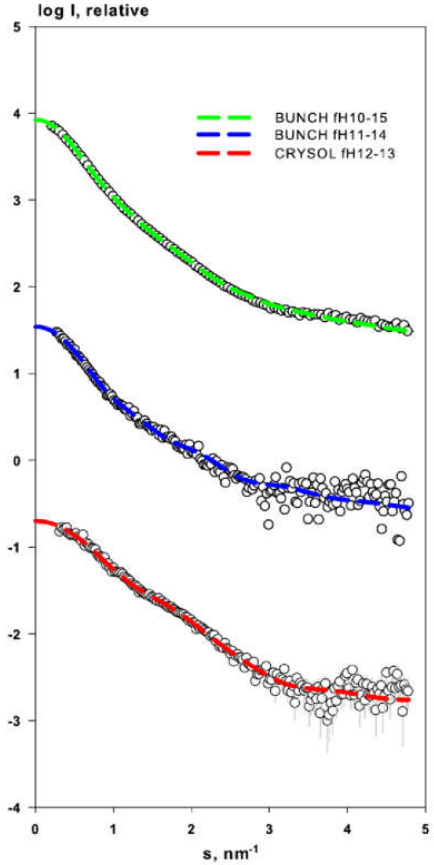
- Requires extremely high photon flux (XFELs)

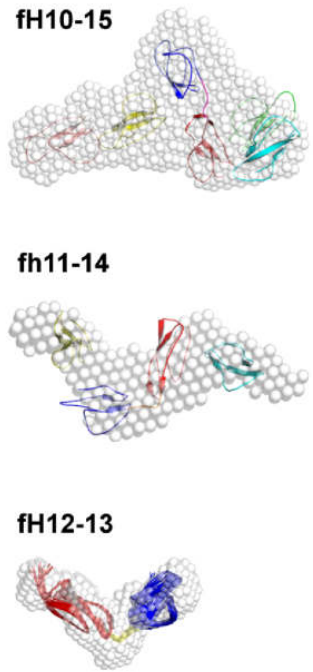
K. J. Gaffney, H. N. Chapman, Science 316, 5830 (2007)
M. M. Seibert *et al.*, Nature 470, 78 (2011)
F. M. Hantke *et al.*, Nat. Phot. 8, 943 (2014)
T. Eckeberg *et al.*, PRL 114, 098102 (2015)

Small-Angle X-ray Scattering







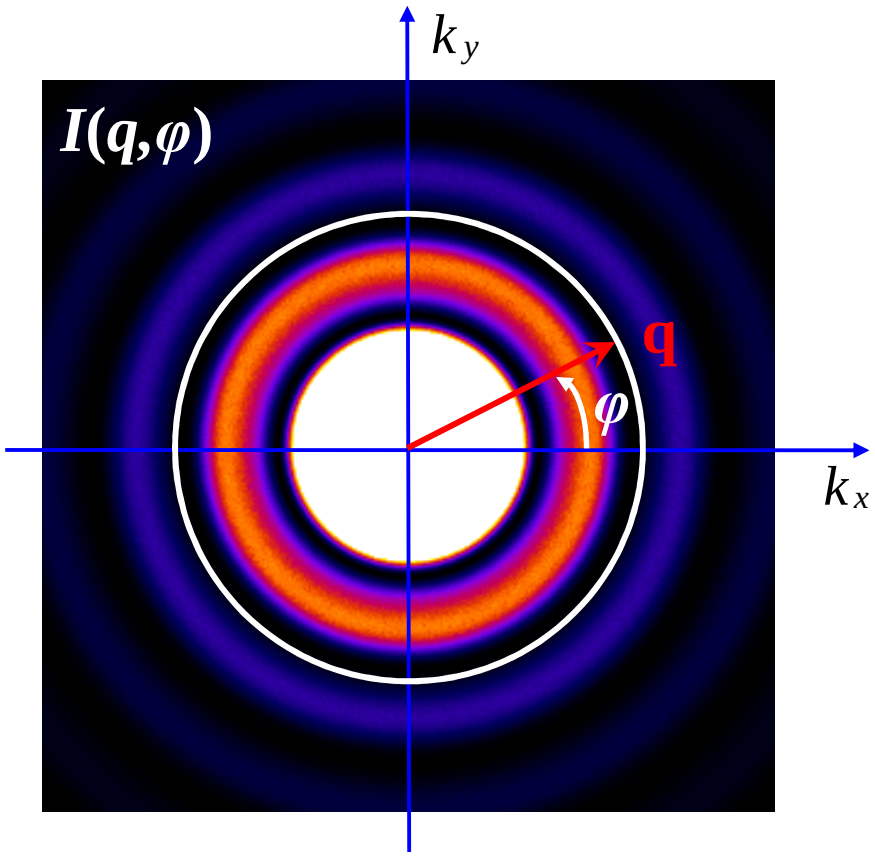


- Particles in native environment
- Low-resolution particle structure

H. D. T. Merens and D. I. Svergun, J. Struct. Biol. 172, 128-141 (2010)

Small-angle x-ray scattering (SAXS):

Dilute system of particles



Scattered intensity from N particles:

$$I(q,\varphi) = \sum_{i=1}^{N} I_i(q,\varphi) \qquad (1)$$

where $I_i(q,\varphi)$ is an individual contribution from the i-th particle.

SAXS intensity:

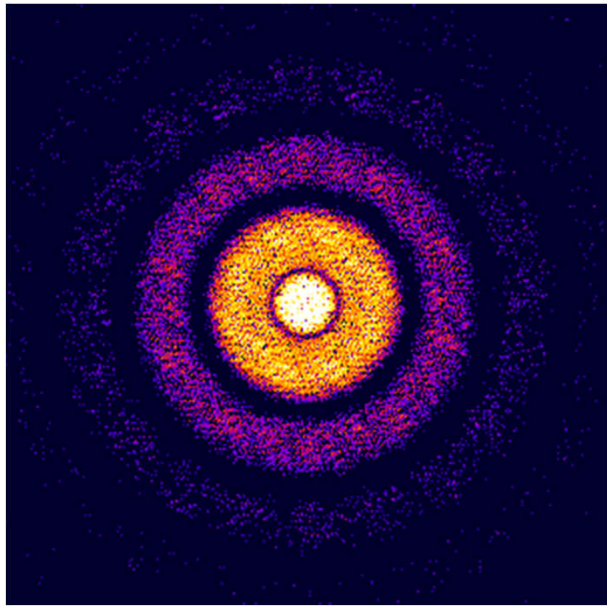
$$I(q) \equiv \langle I(q,\varphi) \rangle_\varphi = \frac{1}{2\pi} \int_0^{2\pi} I(q,\varphi) d\varphi,$$

or

$$I(q) = \left\langle \sum_{i=1}^{N} I_i(q,\varphi) \right\rangle_\varphi \qquad (2)$$

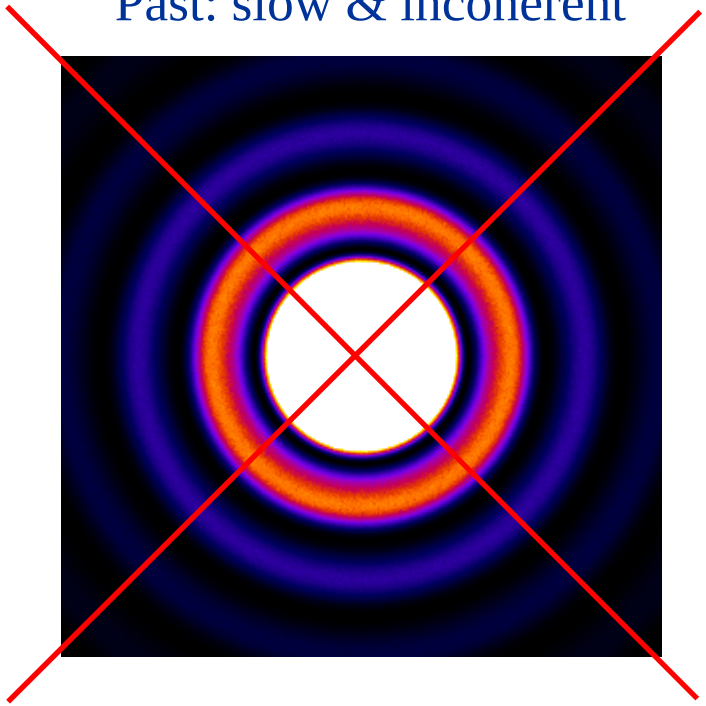
Scattering of coherent ultrashort pulses from a system of *N* particles

Present : fast & coherent

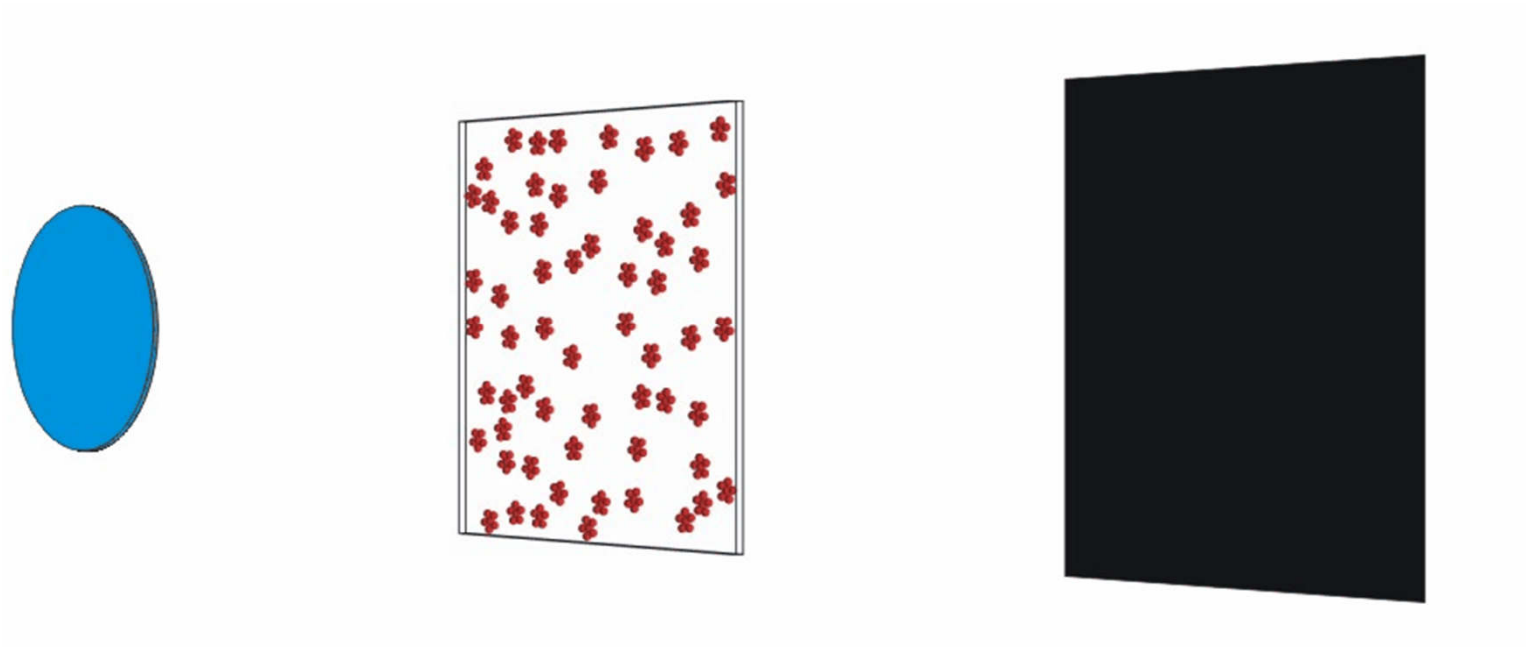


$\sum$

Past: slow & incoherent



Fluctuation X-ray Scattering (FXS)

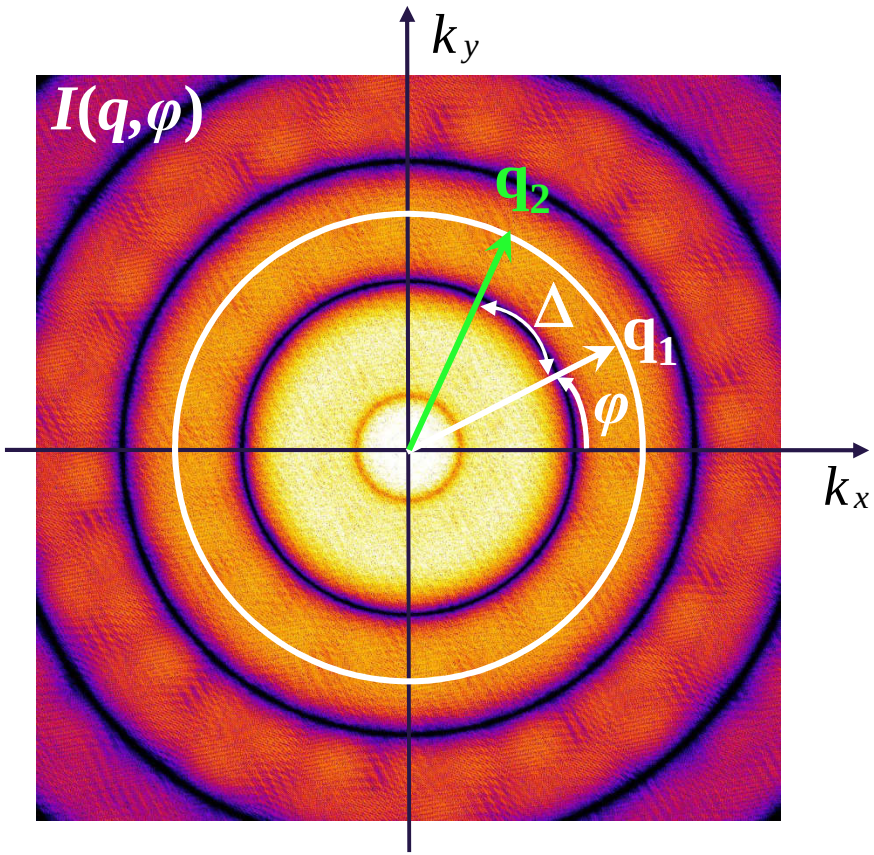


- Avoids intensity smearing due to rotational diffusion by applying ultrashort x-ray pulses (XFELs)
- Advantageous for weakly-scattering particles, since multiple-particle hits can be used along with the single hits

Z. Kam, Macromolecules 10, 927 (1977)
R.P. Kurta, M. Altarelli, I.A. Vartanyants, Adv. Chem. Phys. 161, Ch.1 (2016)

X-ray cross-correlation analysis:

two-point cross-correlation function



$$|\mathbf{q}_1| = |\mathbf{q}_2| = q$$

Two-point cross-correlation function (CCF):

$$C(q,\Delta) = \langle I(q,\varphi) I(q,\varphi+\Delta) \rangle_\varphi \qquad (1)$$

Fourier components:

$$C^n(q) = \frac{1}{2\pi} \int_0^{2\pi} C(q,\Delta) \exp(-in\Delta) d\Delta \qquad (2)$$

$$I^n(q) = \frac{1}{2\pi} \int_0^{2\pi} I(q,\varphi) \exp(-in\varphi) d\varphi \qquad (3)$$

$$\langle \langle I(q,\varphi) \rangle_\varphi \rangle_M = \langle I^0(q) \rangle_M \text{ - conventional SAXS} \qquad (4)$$

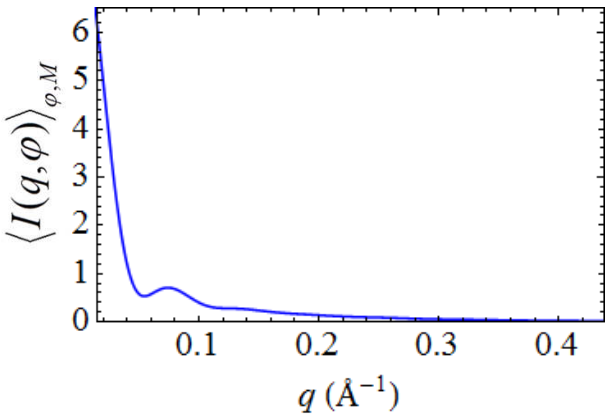
$$\langle C^n(q) \rangle_M = \langle |I^n(q)|^2 \rangle_M \text{ - “higher-order SAXS”} \qquad (5)$$

$\langle \ \rangle_M$ - averaging over diffraction patterns.

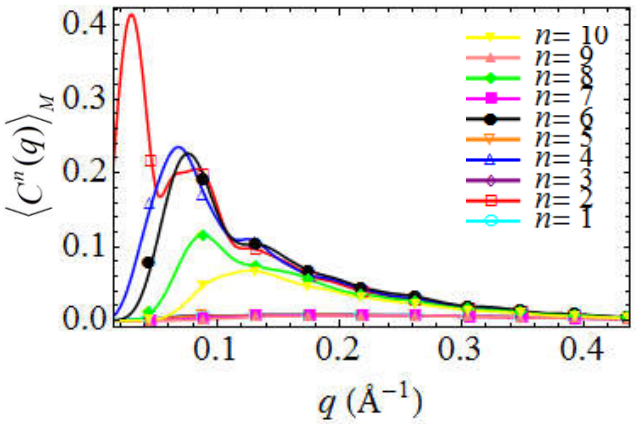
Z. Kam, Macromolecules 10, 927 (1977)
R.P. Kurta, M. Altarelli, I.A. Vartanyants, Adv. Chem. Phys. 161, Ch.1 (2016)

Single-particle structure recovery from solution x-ray scattering

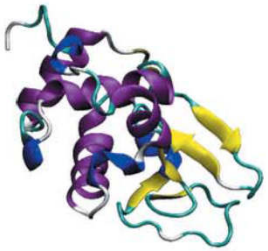
$$\langle I(q,\varphi)\rangle_{\varphi,M} = \langle I^0(q)\rangle_M \qquad \langle C^n(q)\rangle_M \Rightarrow \left\langle \left| I^n(q)\right|^2 \right\rangle_M, n = 1,2,3...$$



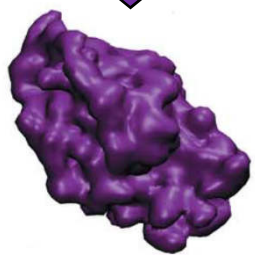
+



=



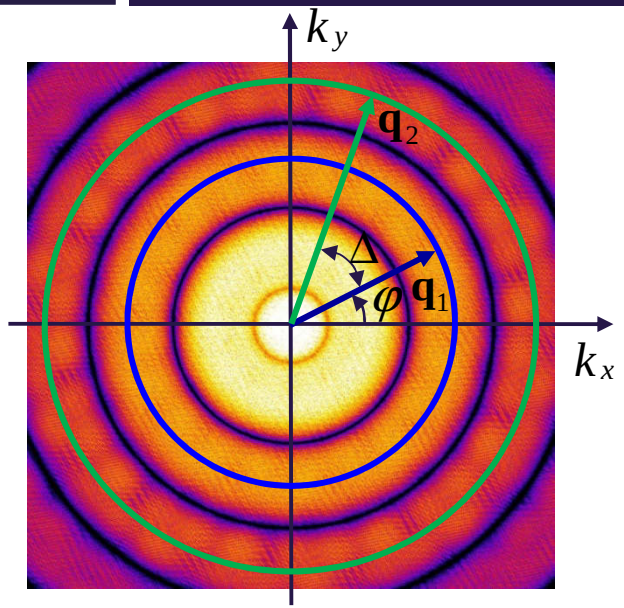
Bead model, *etc*

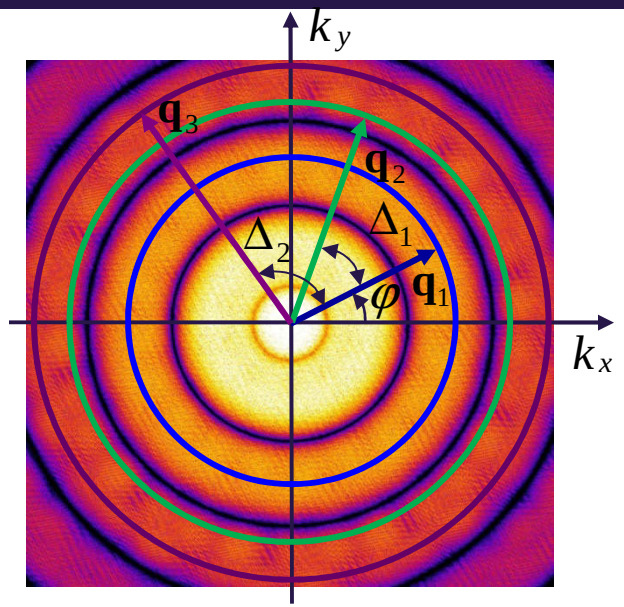


Conventional SAXS analysis

Higher-order SAXS analysis

Two- and three-point CCFs





Two-point CCF:

$$\langle C(q_1,q_2,\Delta)\rangle = \left\langle \langle I(q_1,\varphi)I(q_2,\varphi+\Delta)\rangle_\varphi \right\rangle_M$$

Three-point CCF:

$$\langle C(q_1,q_2,q_3,\Delta_1\Delta_2)\rangle = \left\langle \langle I(q_1,\varphi)I(q_2,\varphi+\Delta_1)I(q_3,\varphi+\Delta_2)\rangle_\varphi \right\rangle_M$$

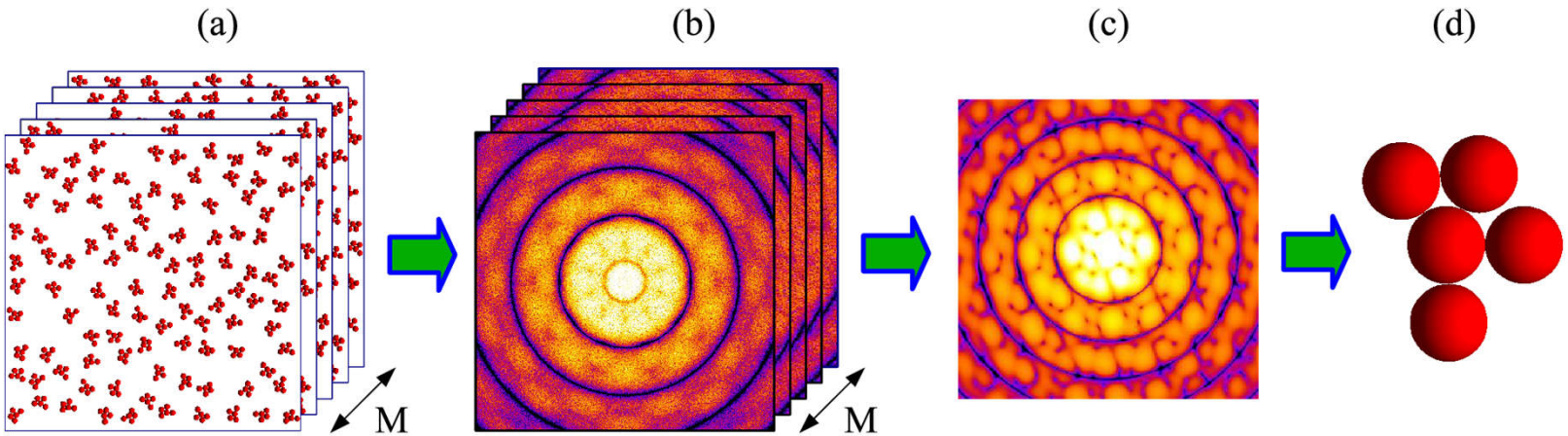
where $\langle f(q,\varphi)\rangle_\varphi = \frac{1}{2\pi}\int_0^{2\pi} f(q,\varphi)d\varphi$ - angular average,

$\langle\ \rangle_M$ - statistical average (over diffraction patterns).

R.P. Kurta, R. Dronyak, M. Altarelli, E. Weckert and I.A. Vartaniants, New J. Phys. 15, 013059 (2013)

Single particle structure recovery from a disordered system of 2D particles

Experiment → XCCA → Phase retrieval → Structure



(a) A large number *M* of realizations of a disordered system composed of *N* reproducible particles;

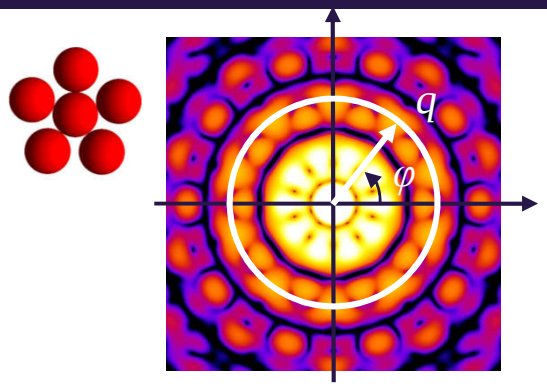
(b) Measured set of diffraction patterns;

(c) Diffraction pattern corresponding to a single particle;

(d) Structure of the single particle.

R.P. Kurta, R. Dronyak, M. Altarelli, E. Weckert and I.A. Vartaniants, New J. Phys. 15, 013059 (2013)
R.P. Kurta, M. Altarelli, I.A. Vartaniants, Adv. Cond. Matt. Phys. 2013, 959835 (2013)

Recovery of the scattered intensity from a single particle by means of XCCA



Fourier series of the scattered intensity from a single particle:

$$I(q,\varphi) = \sum_{n=-\infty}^{\infty} \left|I_q^n\right| \exp(i\phi_q^n)\exp(in\varphi)$$

Fourier components of the CCF:

$$\left\langle C_{q_1,q_2}^n \right\rangle_M = \frac{1}{2\pi}\int_0^{2\pi} \left\langle C(q_1,q_2,\Delta)\right\rangle_M \exp(-in\Delta)d\Delta$$

2-point:

$$\left\langle C_{q_1,q_2}^n \right\rangle_M = I_{q_1}^{n*} \cdot I_{q_2}^n \cdot N$$

Many-particle quantities

single-particle quantities

$$\left|I_q^n\right|$$

$$\arg\left(\left\langle C_{q_1,q_2}^n \right\rangle_M\right) = \phi_{q_2}^n - \phi_{q_1}^n$$

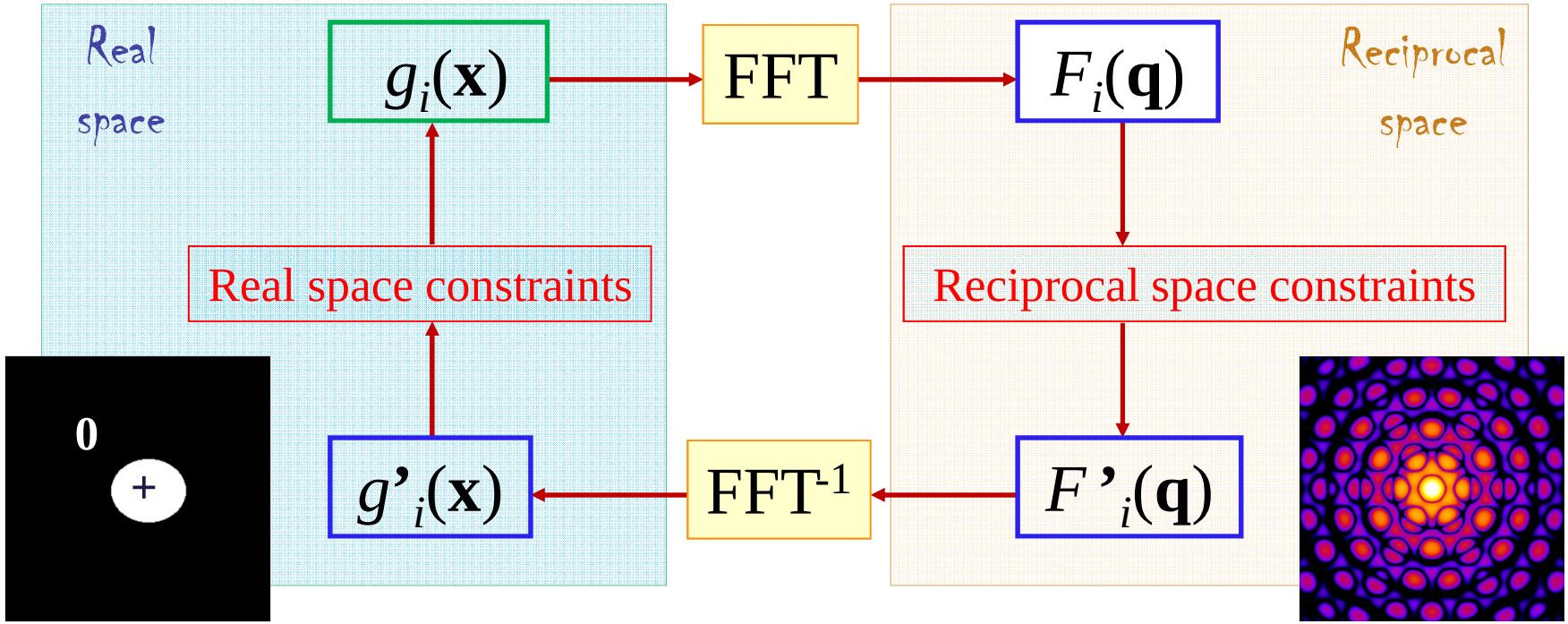
3-point:

$$\left\langle C_{q_1,q_2,q_3}^{n_1,n_2} \right\rangle_M = I_{q_1}^{(n_1+n_2)*} I_{q_2}^{n_1} I_{q_3}^{n_2} \cdot N$$

$$\arg\left(\left\langle C_{q_1,q_2,q_3}^{n_1,n_2} \right\rangle_M\right) = \phi_{q_2}^{n_1} + \phi_{q_3}^{n_2} - \phi_{q_1}^{n_1+n_2}$$

R.P. Kurta *et al.*, New J. Phys. 15, 013059 (2013)

Iterative phase retrieval



Real space constraints:

- Finite support
- Positivity

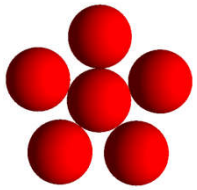
Reciprocal space constraints:

- Intensity constraint (taking into account missing data due to a beamstop, detector gaps, etc)

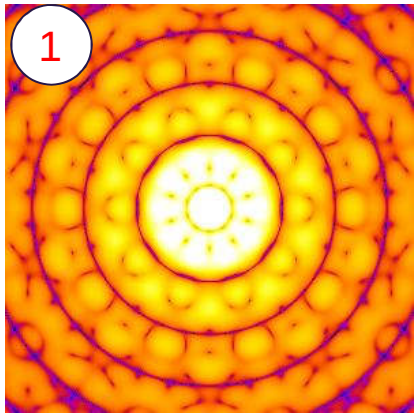
$$|F_i(\mathbf{q})| \rightarrow \sqrt{I_{\exp}(\mathbf{q})}$$

R.W.Gerchberg ,W.O. Saxton, Optik 35, 237 (1972)
J.R. Fienup, Appl. Opt. 21, 2758 (1982)
V. Elser, J. Opt. Soc. Am. A 20, 40 (2003)

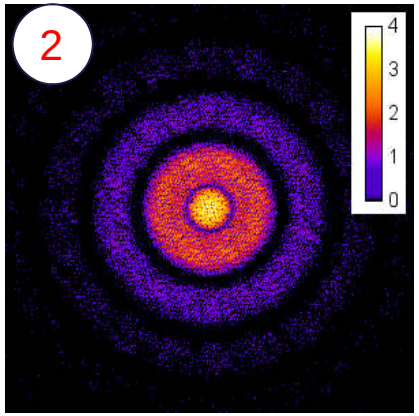
Structure recovery of a single pentagonal cluster



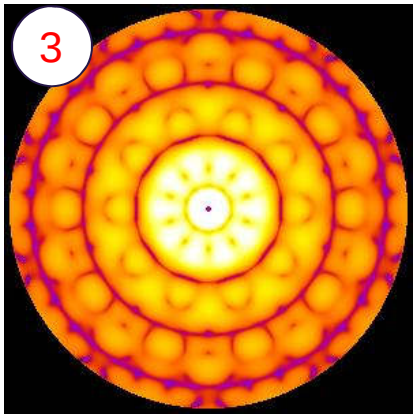
Single pentagonal cluster (N=1)



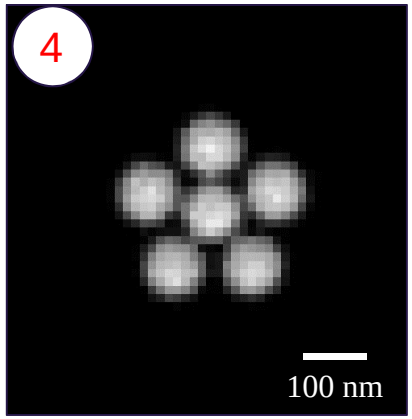
Coherent scattering,
N=10 clusters,
Fluence: $4 \cdot 10^{10}$ ph/µm^2,
Poisson noise,
M=10^5 diffraction patterns.



Diffraction pattern of a single cluster, recovered from M=10^5 patterns of the form (2)

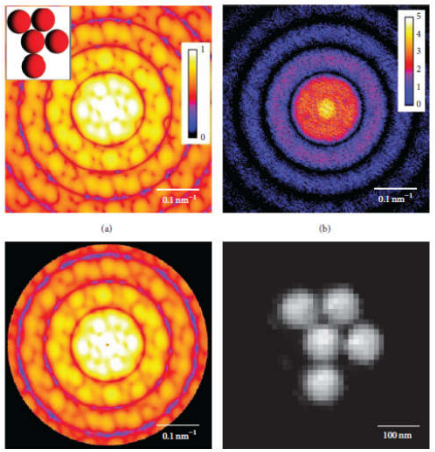


Projected electron density of the cluster, reconstructed using iterative phase retrieval algorithms.

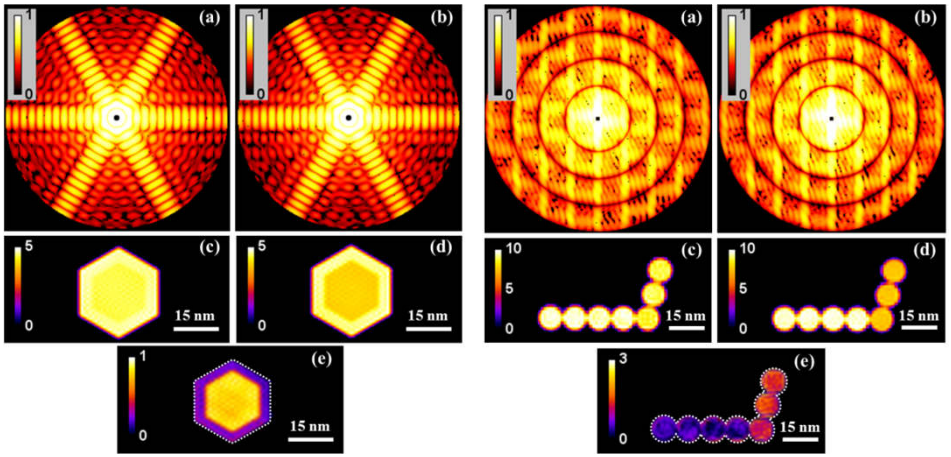


R.P. Kurta *et al.*, New J. Phys. 15, 013059 (2013)

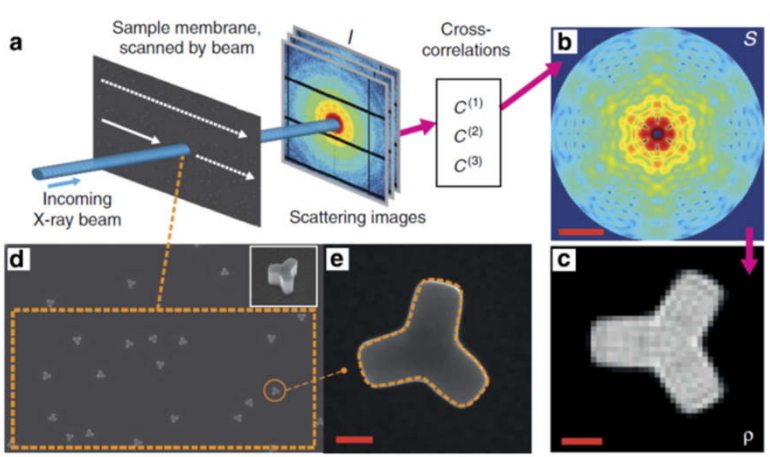
Single-particle structure from disordered 2D ensembles



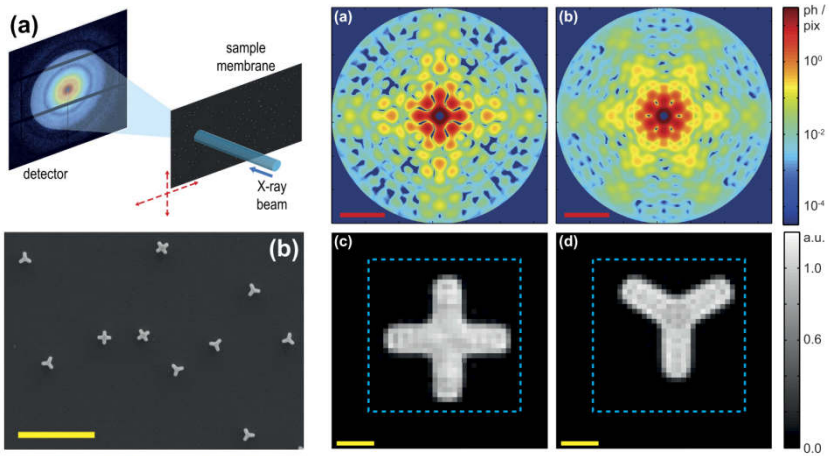
R.P. Kurta et al., New J. Phys. 15, 013059 (2013)
R.P. Kurta et al., Adv. Chem. Phys. 161, Ch. 1 (2016)



R.P. Kurta, J. Phys. B: At. Mol. Opt. Phys. 49, 165001 (2016)

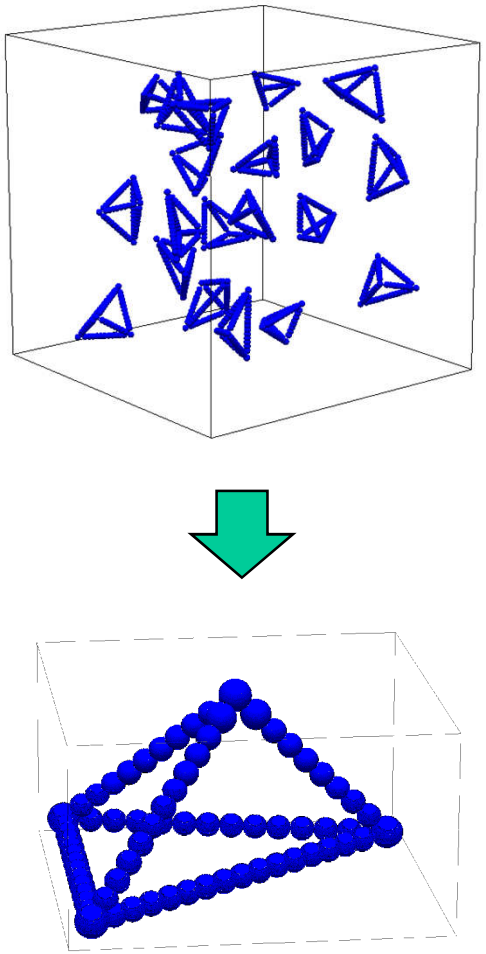


B. Pedrini et al., Nat. Comm. 4, 1647 (2013)



B. Pedrini et al., Sci. Rep. 7, 45618 (2017)

Fluctuation x-ray scattering from 3D systems

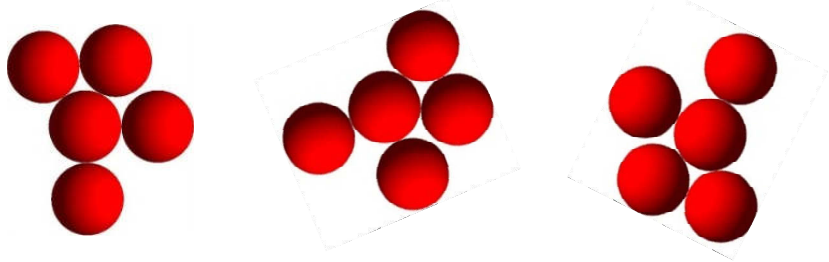


Z. Kam, Macromolecules 10, 927 (1977)

XCCA for a 3D system: the problem of direct phasing

2D systems

ψ_0 ψ_1 ψ_2



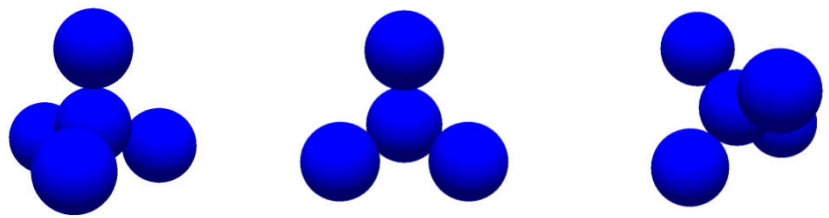
$$\left\langle C^{n}_{q_1,q_2} \right\rangle_M = I^{n*}_{q_1,\psi_0} \cdot I^{n}_{q_2,\psi_0}$$

$$\left\langle C^{n_1,n_2}_{q_1,q_2,q_3} \right\rangle_M = I^{(n_1+n_2)*}_{q_1,\psi_0} I^{n_1}_{q_2,\psi_0} I^{n_2}_{q_3,\psi_0}$$

$\langle \ \rangle_M$ - averaging over diffraction patterns.

3D systems

ψ_0 ψ_1 ψ_2



$$\left\langle C^{n}_{q_1,q_2} \right\rangle_M = \left\langle I^{n*}_{q_1,\psi_i} \cdot I^{n}_{q_2,\psi_i} \right\rangle_\psi$$

$$\left\langle C^{n_1,n_2}_{q_1,q_2,q_3} \right\rangle_M = \left\langle I^{(n_1+n_2)*}_{q_1,\psi_i} I^{n_1}_{q_2,\psi_i} I^{n_2}_{q_3,\psi_i} \right\rangle_\psi$$

$\langle \ \rangle_\psi$ - averaging over different projections.

Analysis of 3D systems: indirect approaches

- ✓ Spherical harmonics

  Z. Kam, Macromolecules 10, 927 (1977)
  Z. Kam, J. Theor. Biol. 82, 1 (1980)
  D. K. Saldin *et al.*, J. Phys.: Cond. Matt. 21, 134014 (2009)
  D. Starodub *et al.*, Nat. Comm. 3, 1276 (2012)

- ✓ Icosahedral harmonics

  D. K. Saldin *et al.*, Opt. Express 19, 17318 (2011)

- ✓ Cylindrical harmonics

  D. K. Saldin *et al.*, Acta Cryst. A 66, 32 (2010)
  G. Chen *et al.*, J. Synch. Rad. 19, 695 (2012)

- ✓ Zernike polynomials

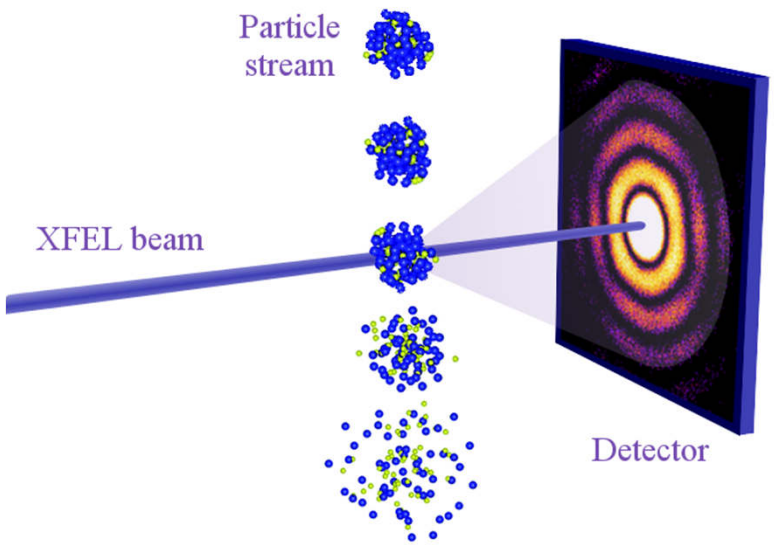
  H. Liu *et al.*, Acta Cryst. A 68, 561 (2012)

  H. Liu *et al.*, Acta Cryst. A 69, 365 (2013)

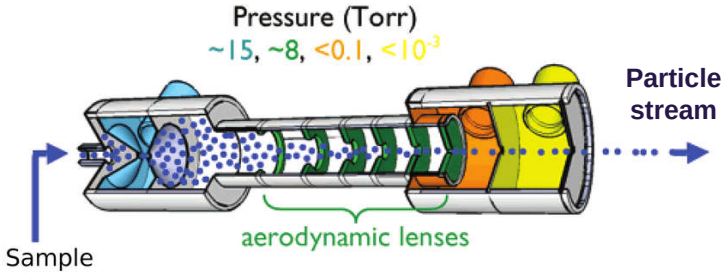
- Indirect methods: fitting or model assumptions are applied
- Better than conventional SAXS: additional information is available from the analysis

Experiment ➡ CCFs ➡ 3D Intensity ➡ Phase retrieval ➡ Structure

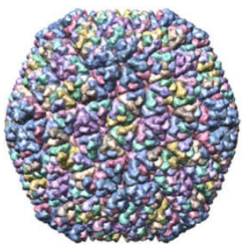
SPI experiments on aerosolized virus particles at the AMO instrument, LCLS



Photon energy: E=1.6keV
Sample-detector distance: 581 mm
Detector: pnCCD
Sample injection: aerodynamic lens stack system with a GDNV



RDV



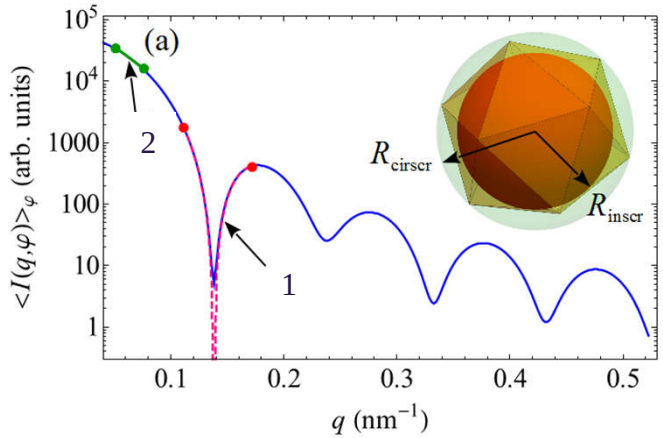
The Rice Dwarf Virus was the first studied plant pathogenic virus; an icosahedral double shelled virus, ranging from 70-75 nm in diameter.

PR772

The Parenteral Drug Association virus filter task force has chosen PR772 as a model bacteriophage to standardize nomenclature for large-pore-size virus-retentive filters (filters designed to retain viruses larger than 50-60 nm in size).

H. K. N. Reddy *et al.*, Scientific Data 4, 170079 (2017)
M. Bogan *et al.*, Nano Lett. 8, 310 (2008)

Size distribution histograms for RDV and PR772



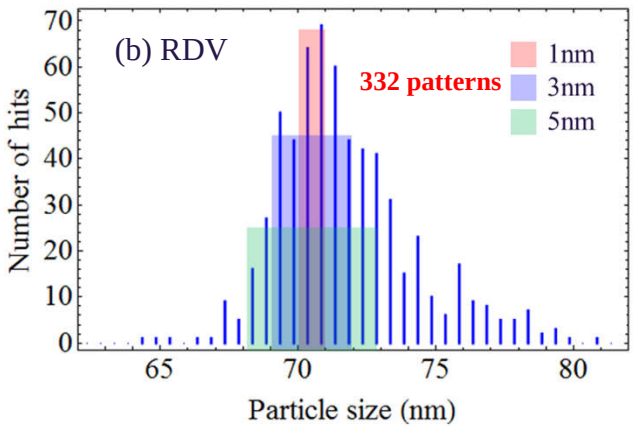
1) Spherical particle form factor fit

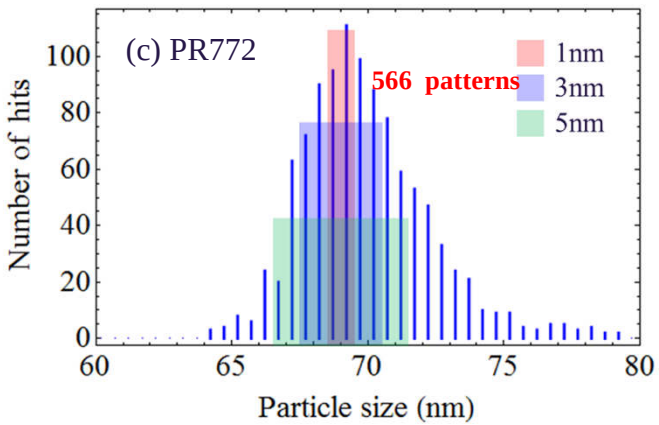
$$\langle I(q)\rangle_\varphi = A\left[\frac{\sin(qR_s)-qR_s\sin(qR_s)}{q^3}\right]^2 \quad (1)$$

2) Guinier-type fit: ($q<1.3/R_g$)

$$\langle I(q)\rangle_\varphi = I(0)\exp[-q^2R_g^2/3] \quad (2)$$

$$y = ax+b \qquad R_g = \sqrt{-3a}$$





Exact definition of particle size:

$$\text{Size} = 2R_{circum} \quad (3)$$

$$R_{circum} = \sqrt{15-6\sqrt{5}}\,R_{inscr} \quad (4)$$

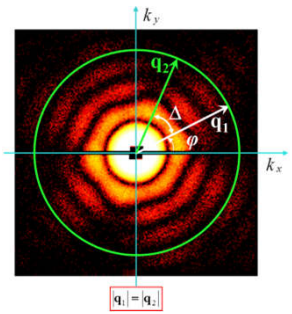
Empirical approximation:

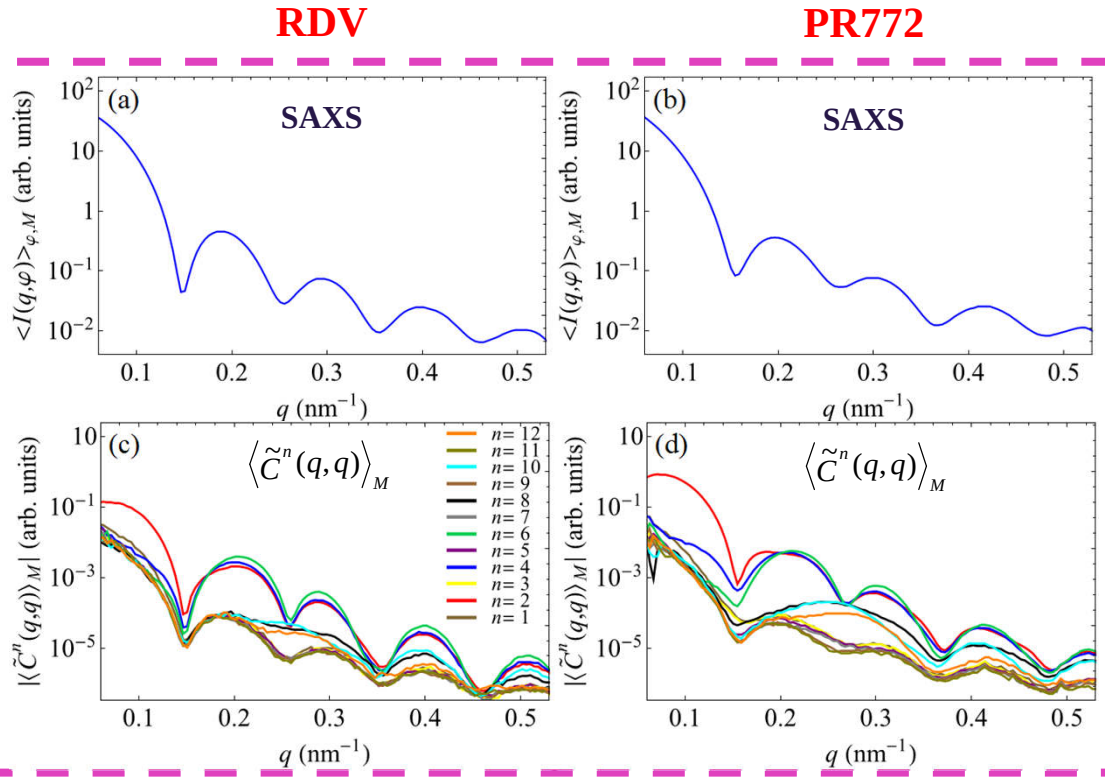
$$R_{inscr} \approx (R_s + R_g)/2 \quad (5)$$

$$\text{Size} \approx 1.26(R_g + R_s) \quad (6)$$

$$(\text{ or Size} \approx 2.36\,R_s\ ) \quad (7)$$

Experimental results: generalized SAXS dataset

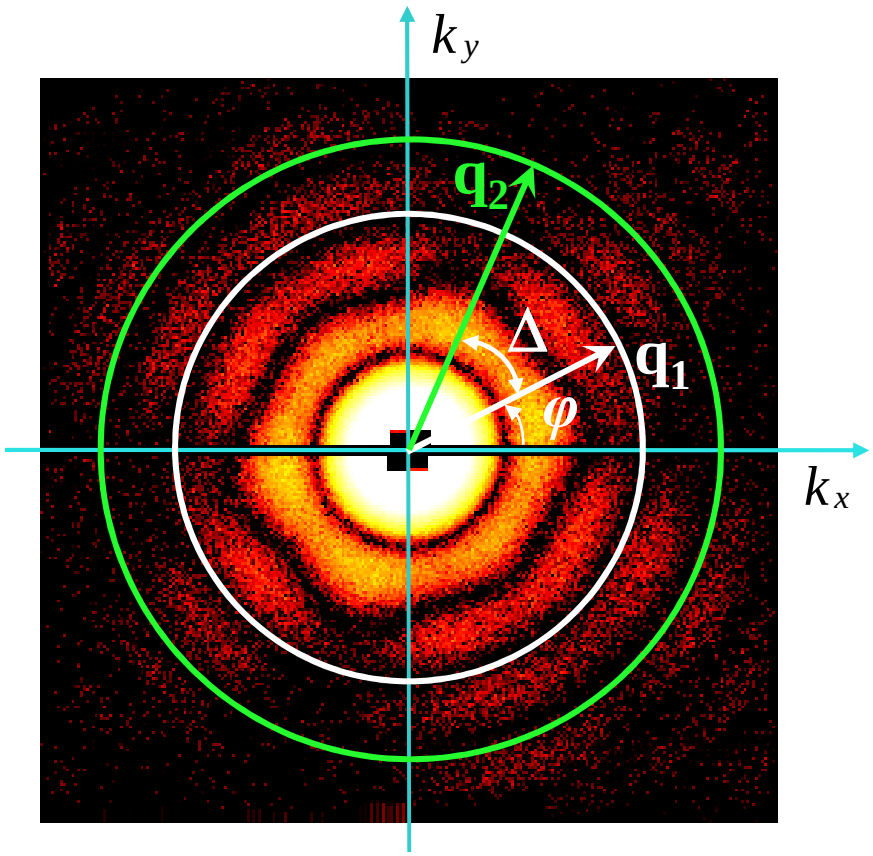




Generalization of SAXS: Higher-order SAXS

R.P. Kurta *et al.*, PRL 119, 158102 (2017)

Advanced correlation analysis



$$|\mathbf{q}_1| \neq |\mathbf{q}_2|$$

CCF defined on the same pattern *i*:

$$\langle C_{ii}(q_1,q_2,\Delta)\rangle_M = \left\langle \langle I_i(q_1,\varphi)\, I_i(q_2,\varphi+\Delta)\rangle_\varphi \right\rangle_M \qquad (1)$$

CCF defined between different patterns *i* and *j*:

$$\langle C_{ij}(q_1,q_2,\Delta)\rangle_M = \left\langle \langle I_i(q_1,\varphi)\, I_j(q_2,\varphi+\Delta)\rangle_\varphi \right\rangle_M \qquad (2)$$

Fourier components of the CCFs:

$$\langle C_{ij}^n(q_1,q_2)\rangle_M = \frac{1}{2\pi}\int_0^{2\pi} \langle C_{ij}(q_1,q_2,\Delta)\rangle_M \exp(-in\Delta)d\Delta \qquad (3)$$

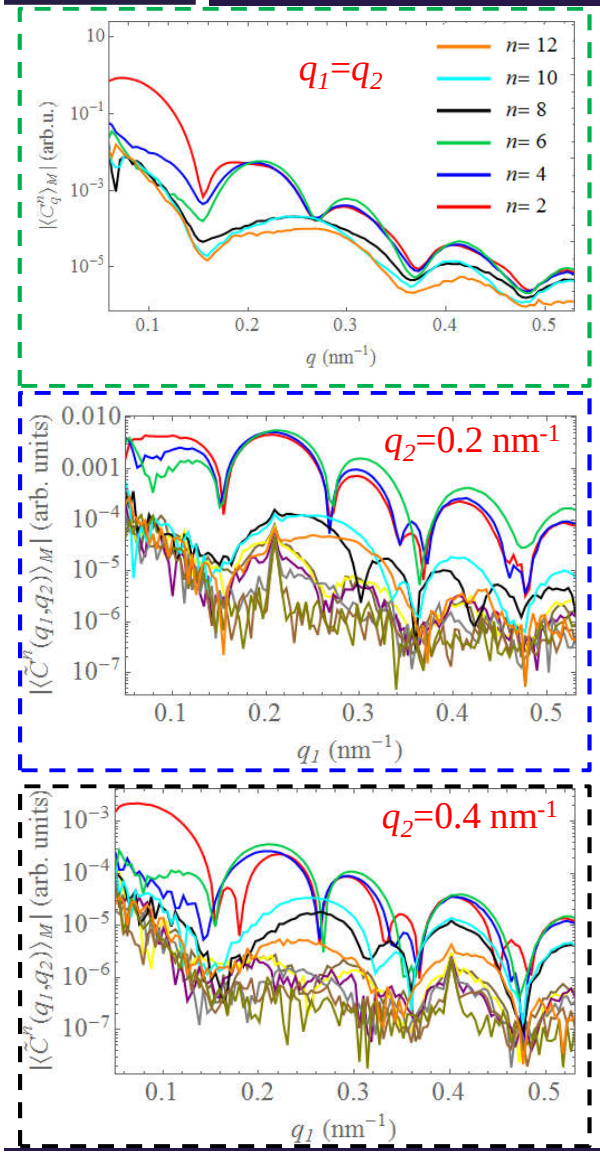
Difference spectrum:

$$\boxed{\langle \tilde{C}^n(q_1,q_2)\rangle_M = \langle C_{ii}^n(q_1,q_2)\rangle_M - \langle C_{ij}^n(q_1,q_2)\rangle_M} \qquad (4)$$

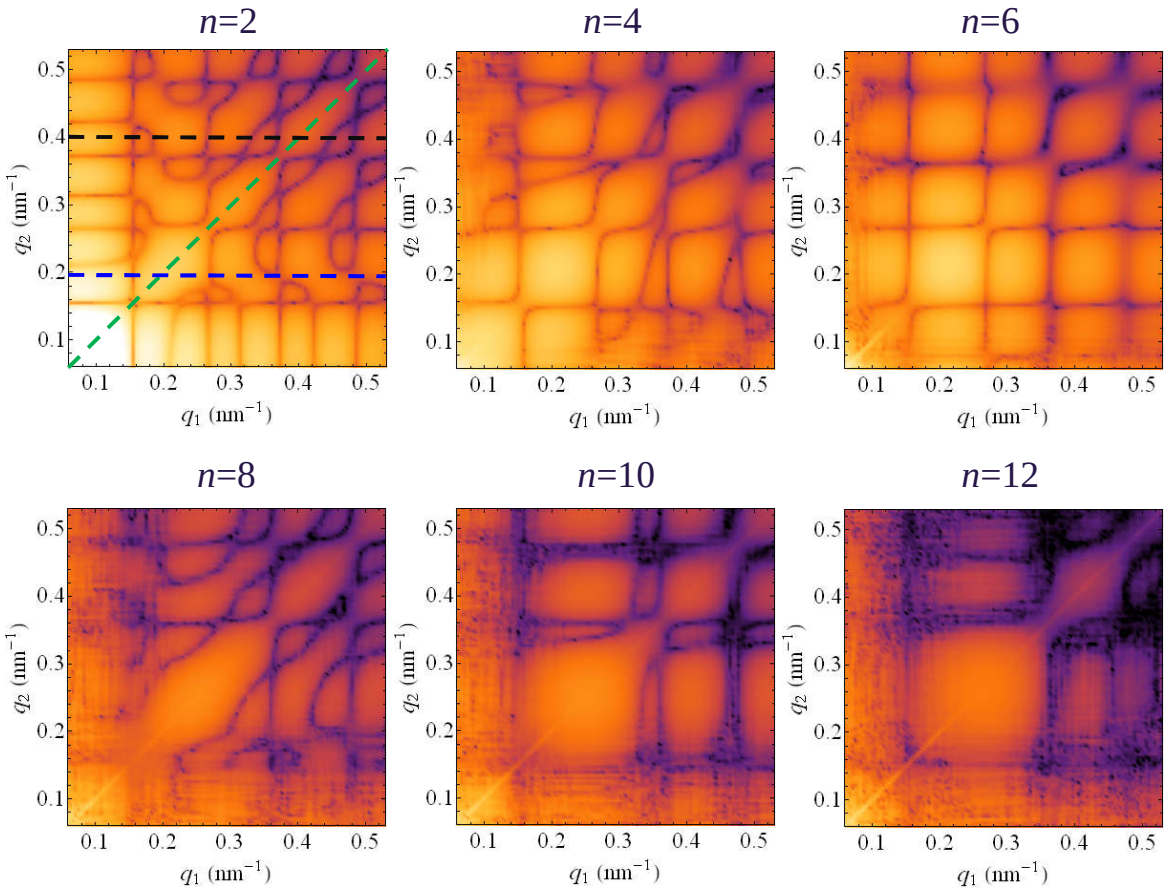
Difference spectrum is used to reduce undesirable effect of various systematic issues and improve the FXS data quality.

R.P. Kurta *et al.*, J. Phys.: Conf. Series 499, 012021 (2014)
G. Chen *et al.*, J. Synch. Rad. 19, 695700 (2012)
R.P. Kurta, M. Altarelli, I.A. Vartanyants, Adv. Chem. Phys. 161, Ch.1 (2016)

Full correlation dataset: 2D correlation maps



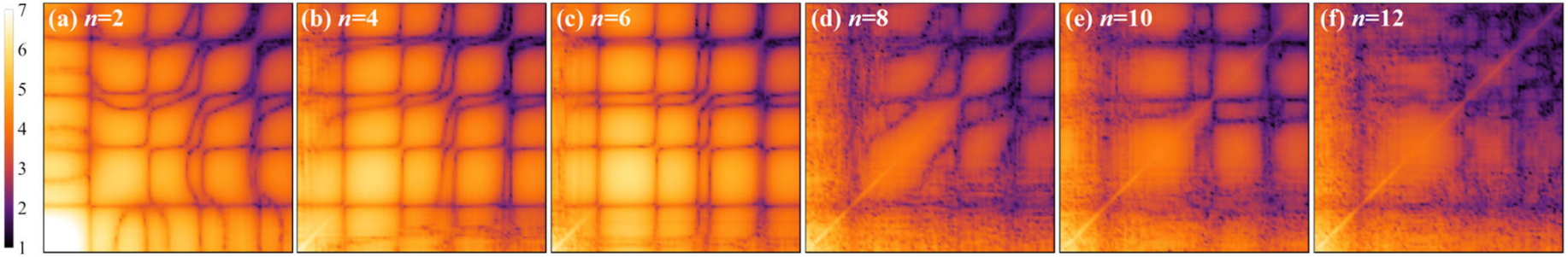
$$\left|\left\langle \tilde{C}^n(q_1,q_2)\right\rangle_M\right| = \left|\left\langle C_{ii}^n(q_1,q_2)\right\rangle_M - \left\langle C_{ij}^n(q_1,q_2)\right\rangle_M\right|$$



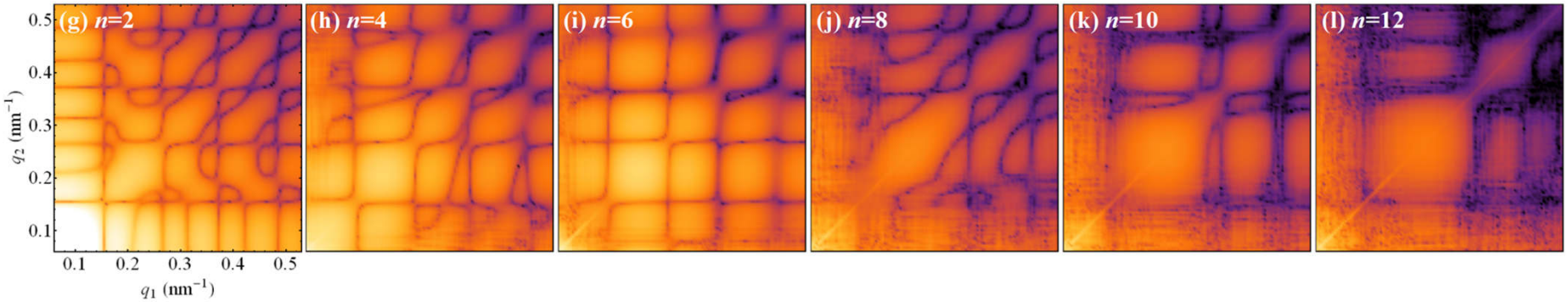
Experimental 2D correlation maps for RDV and PR772

Amplitudes: $\left|\left\langle \tilde{C}^{n}(q_1, q_2)\right\rangle_M\right|$

RDV



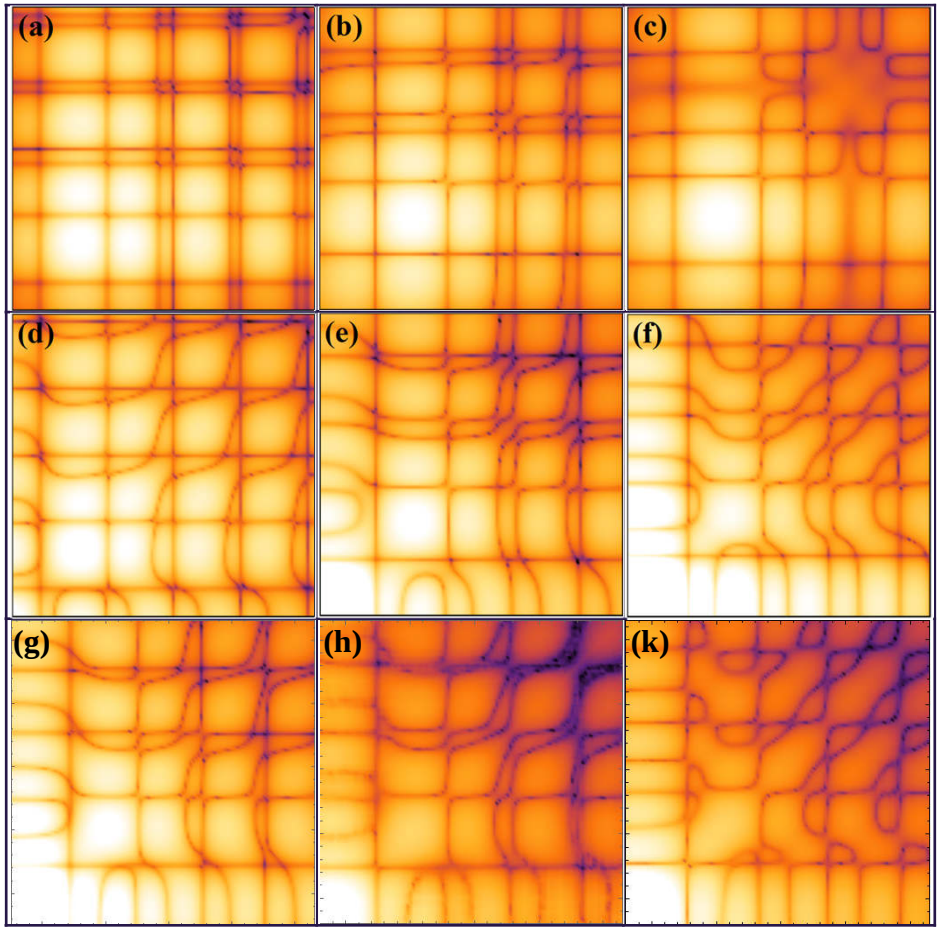
PR772



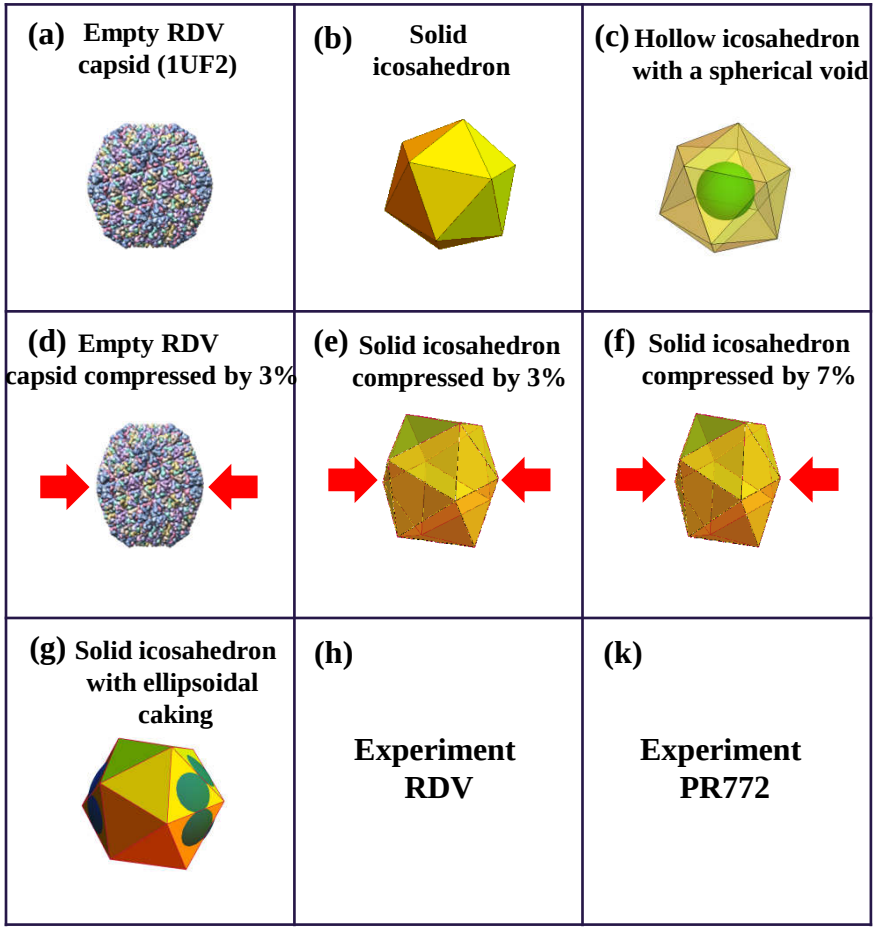
R.P. Kurta *et al.*, PRL 119, 158102 (2017)

Model-based structural analysis

$$\left|\left\langle \tilde{C}^{n=2}(q_1,q_2)\right\rangle_M\right|$$



Model structures

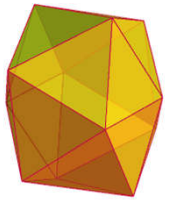


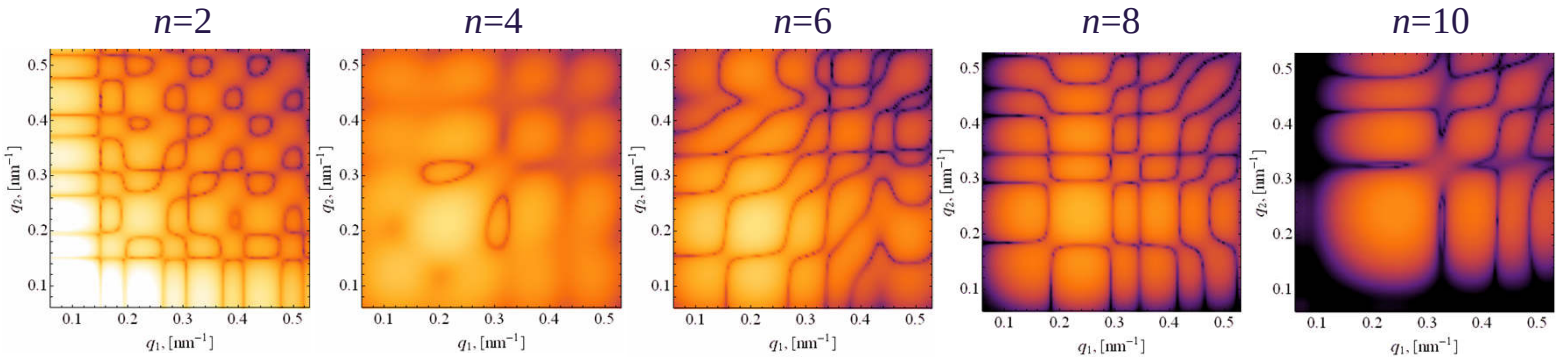
R.P. Kurta *et al.*, PRL 119, 158102 (2017)

Sensitivity of cross-correlations to structure variation

Particle distortion

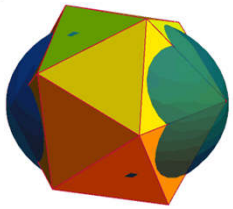
0.750

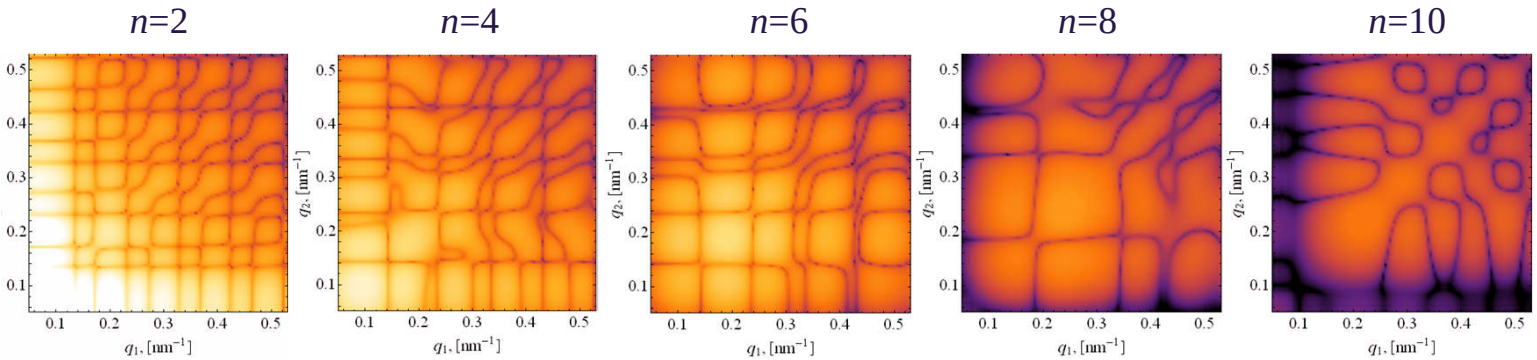




Particle caking

375





Multitiered iterative phasing (MTIP) for structure recovery

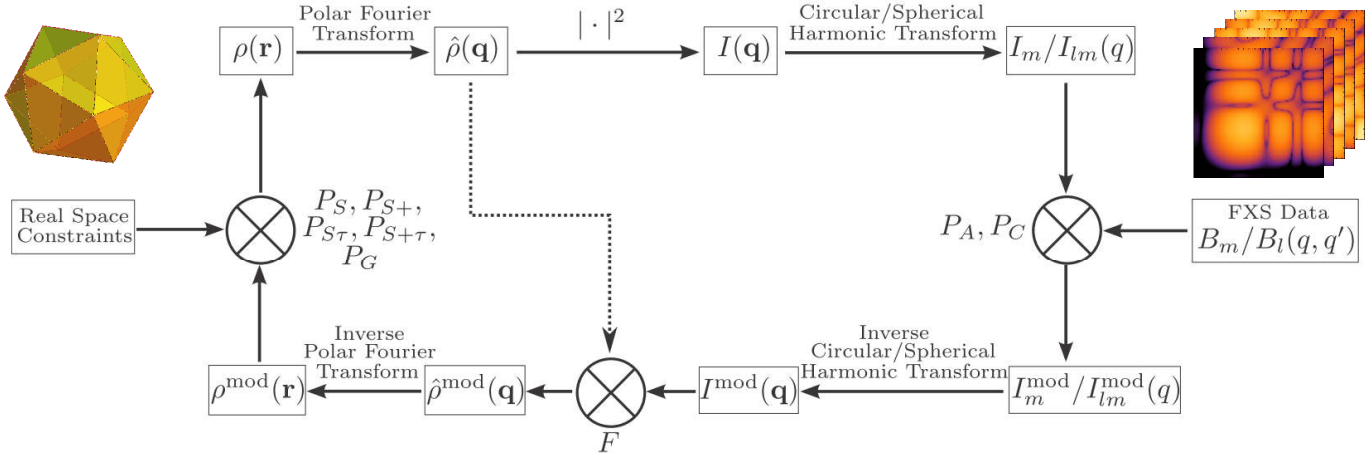
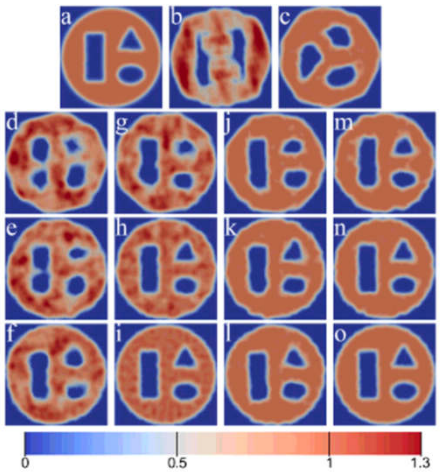
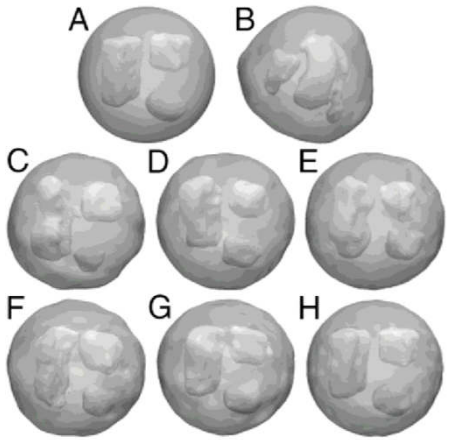
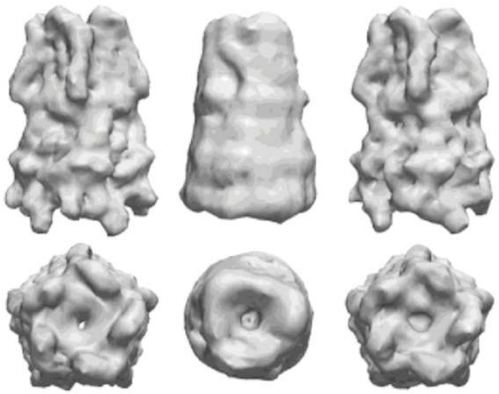


Fig. S7. Flowchart of the multitiered iterative phasing procedure. The mod superscript denotes values that have been modified to agree with the data. One uses circular harmonic transforms, I_m, and B_m for the 2D case and spherical harmonic transforms, I_{lm}, and B_l for the 3D case.

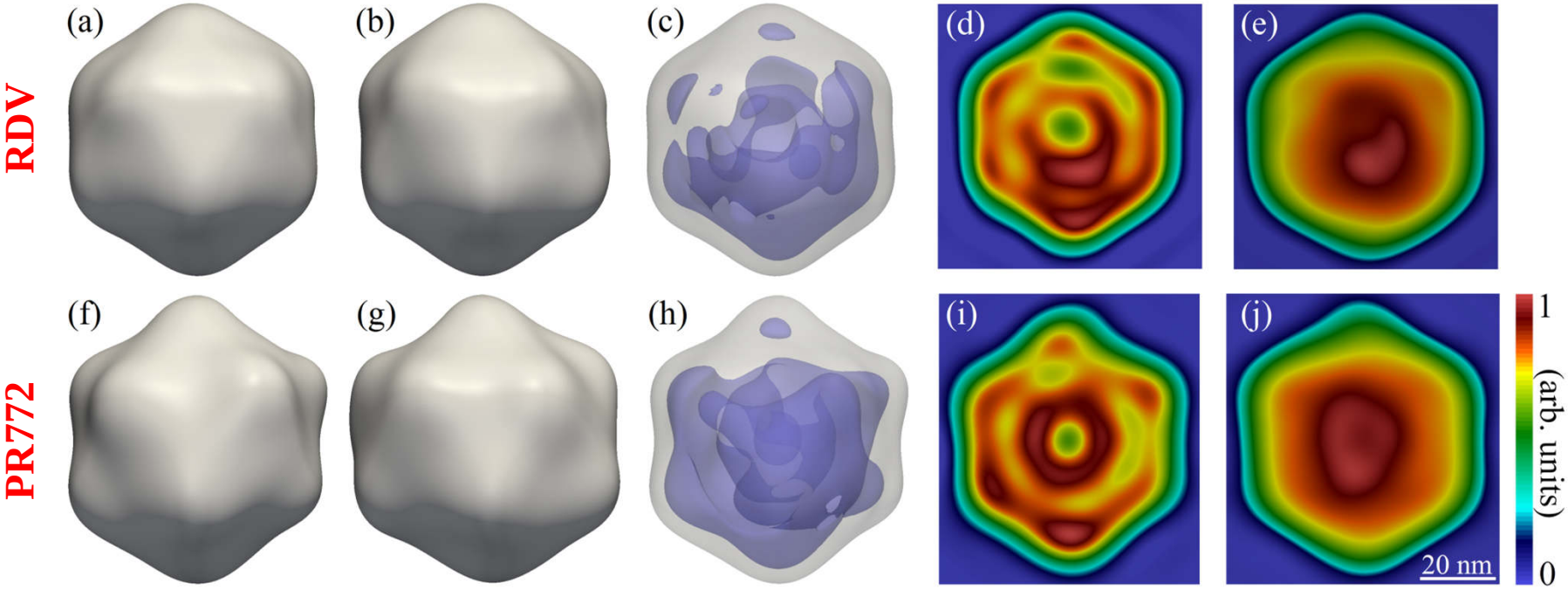






J. J. Donatelli, P. H. Zwart, J. A. Sethian, PNAS 112 (33), 10286 (2015)

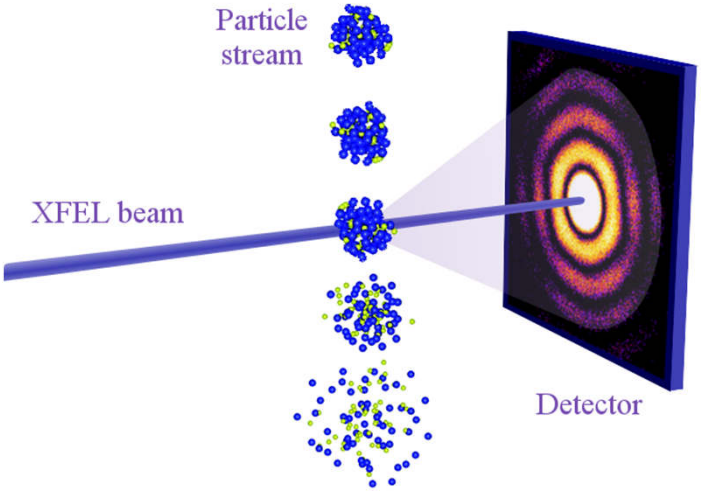
MTIP reconstructions for RDV and PR772



Reconstructed images of RDV and PR772.
Two different views (corresponding to a 72 degree rotation about the top axis) of the reconstructed RDV (a,b) and PR772 (f,g) particles, as well as density plots showing nonuniformities in the internal distribution of material inside RDV (c) and PR772 (h), 2D slices through the center of the reconstructed densities for RDV (d) and PR772 (i), and 2D projections of the reconstructed densities for RDV (e) and PR772 (j).

R.P. Kurta *et al.*, PRL 119, 158102 (2017)

SPI experiments on PR772 virus particles at the AMO instrument, LCLS



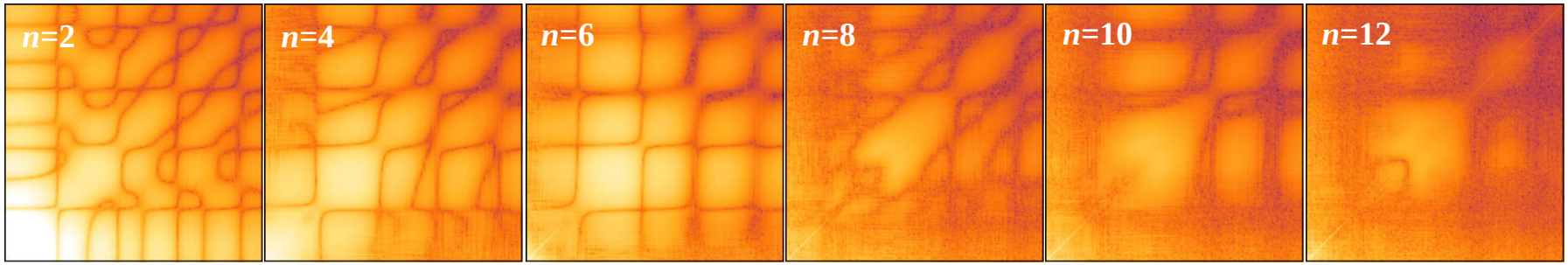
AMO_86615 (August 2015):
Photon energy: E=1.6 keV
Sample-detector distance: 581 mm
Detector: pnCCD
Sample: PR772, RDV
Sample delivery: GDVN

AMO_06516 (April 2016):
Photon energy: E=1.7 keV
Sample-detector distance: 257 mm
Detector: pnCCD
Sample: PR772
Sample delivery: GDVN

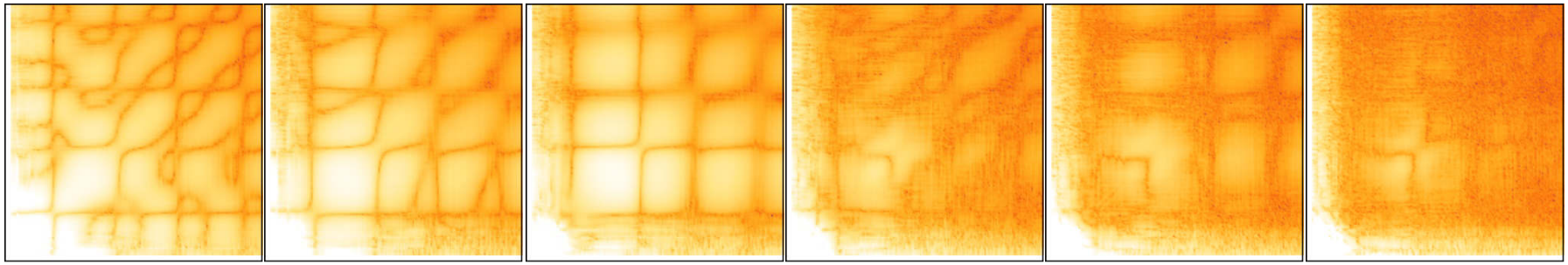
AMO_11416 (August 2016):
Photon energy: E=1.7 keV
Sample-detector distance: 205 mm
Detector: pnCCD
Sample: PR772
Sample delivery: GDVN

Experimental 2D correlation maps for PR772

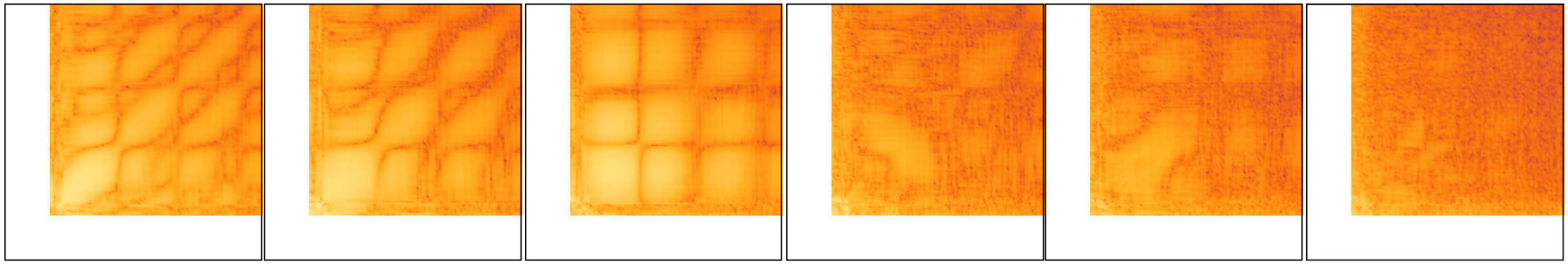
AMO_86615 (PD=3nm, M=2705 patterns)



AMO_06516 (PD=3nm, M=621 patterns)

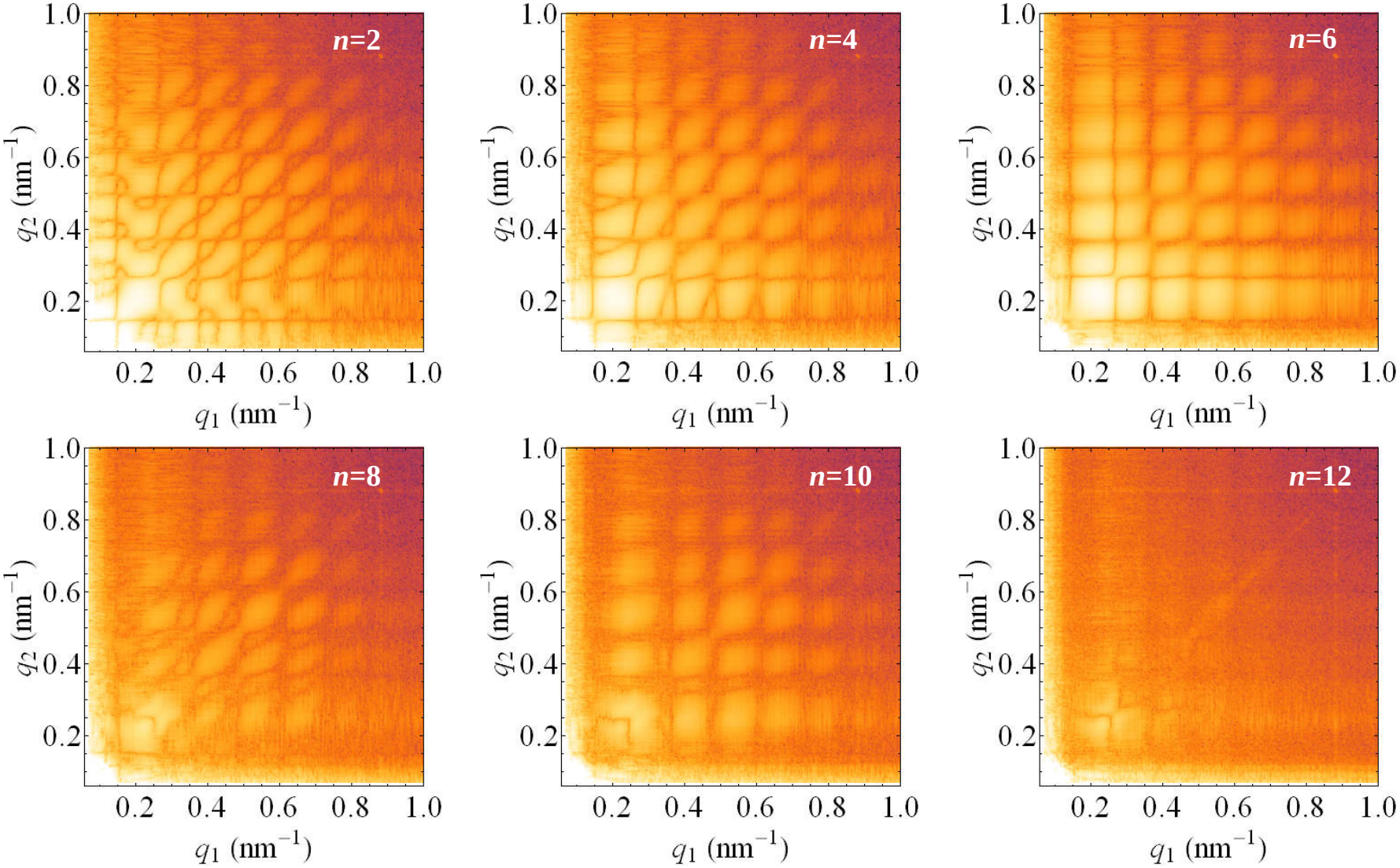


AMO_ 11416 (PD=3nm, M=115 patterns)



Experimental 2D correlation maps:

AMO_06516 data, full range



Summary

- FXS approach offers an alternative way for structural analysis of biological particles with an XFEL;
- FXS is, in the simplest case, a generalization of SAXS;
  In more general case it may give several orders of magnitude increase of information content as compared to SAXS;
- Cross-correlation functions are valuable statistical means, which can be conveniently used for model-based and *ab-initio* structure recovery;
- First application of FXS to biological particles at an XFEL demonstrates substantial potential of the technique for the future studies of structure and dynamics of biomaterials.

Acknowledgements

European XFEL

M. Altarelli
A. Mancuso
A. Madsen
S. Molodtsov

SLAC

A. Aquila
C. H. Yoon
H. Li

LBNL

J. J. Donatelli
P. H. Zwart

DESY

I.A. Vartanyants
I.A. Zaluzhnyy
O.Y. Gorobtsov
A. Shabalin
A. Singer
R. Dronyak
D. Dzhigaev
M. Sprung
E. Weckert

+ >100 more from the SPI initiative

Thank you for your attention