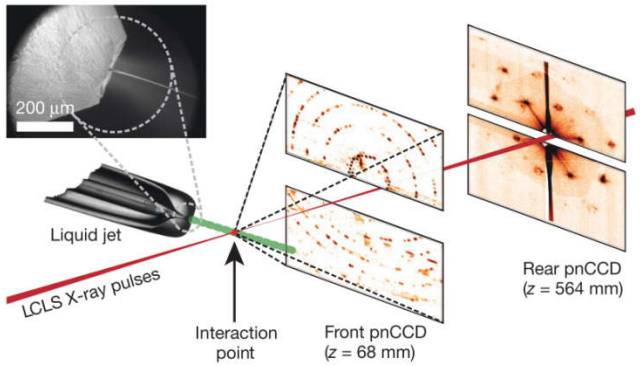
Revealing the nanoscale structure of viruses with XFEL pulses

Ruslan Kurta
Theory Group
Theoretical physicist

Hamburg, 24.01.2018

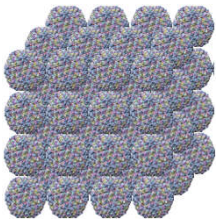
X-ray techniques for biological structure determination at XFELs

Serial Femtosecond Crystallography (SFX)

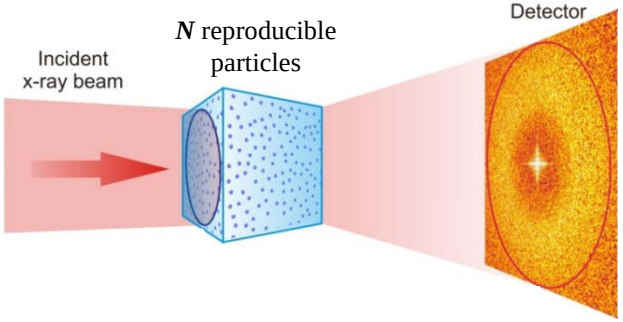


H. N. Chapman *et al*., Nature 470, 73 (2011)

- Requires crystalline samples

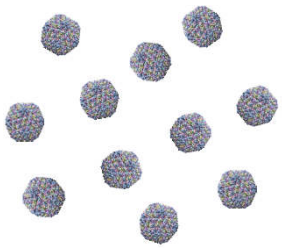


Fluctuation X-ray Scattering (FXS)

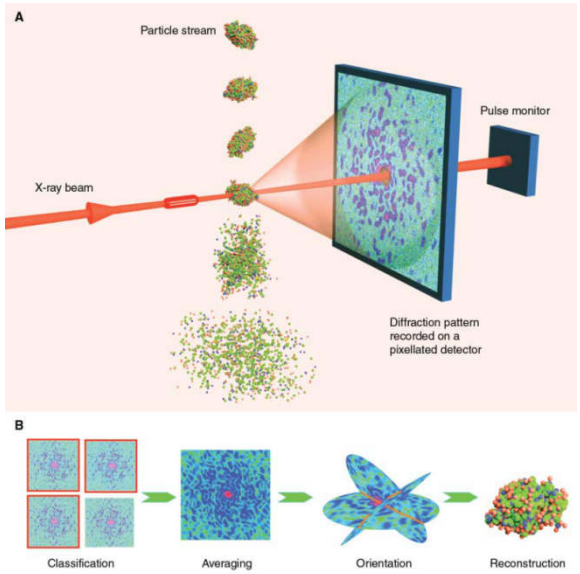


Z. Kam, Macromolecules 10, 927 (1977)

- Can be applied to a disordered system of N particles ($N \geq 1$)
- FXS is a generalization of small-angle x-ray scattering (SAXS)



Single Particle Coherent Diffractive Imaging (SPI)

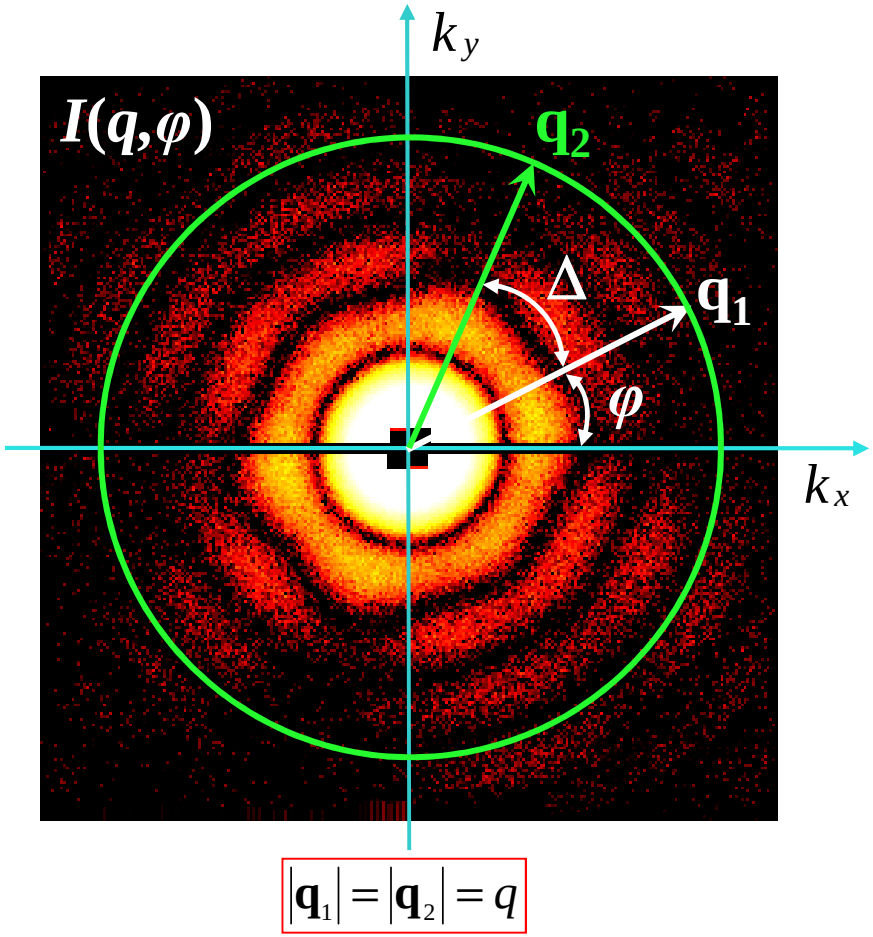


H. N. Chapman *et al*., Nature 470, 73 (2011)

- Requires extremely high photon flux

European XFEL

Fluctuation x-ray scattering: basics



Two-point cross-correlation function (CCF):

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

where $I(q,\varphi)$ - scattered intensity,

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

Fourier components:

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

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

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

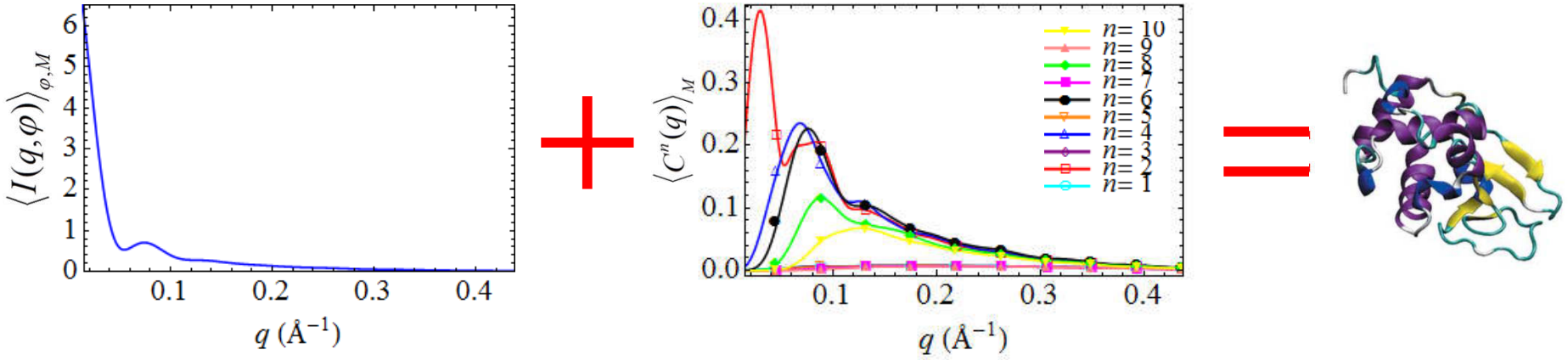
$$\langle C^n(q) \rangle_M = \langle |I^n(q)|^2 \rangle_M \text{ - "higher-order SAXS"} \quad (5)$$

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

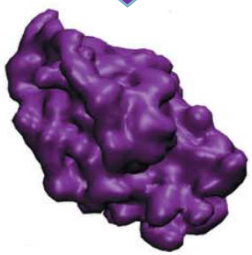
Z. Kam, Macromolecules 10, 927 (1977)
R.P. Kurta, M. Altarelli, I.A. Vartanyants, Adv. Chem. Phys. 161, Ch.1 (2016)

Single-particle structure recovery from fluctuation x-ray scattering

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



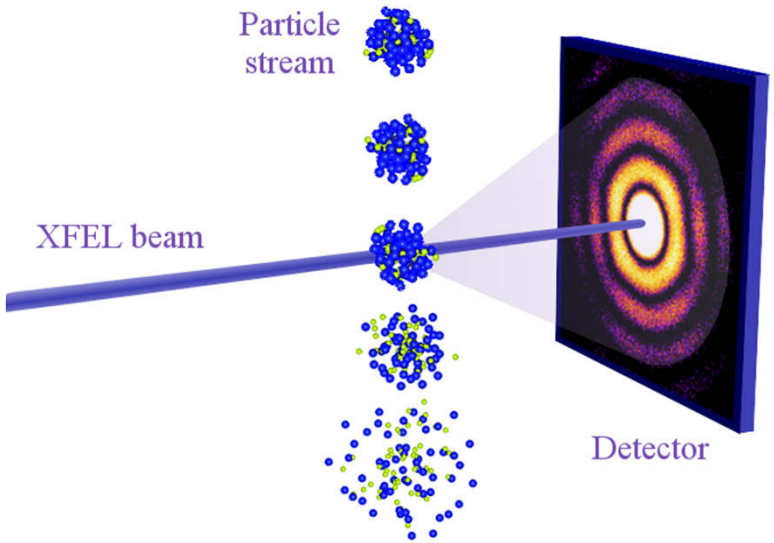
Bead model, *etc*



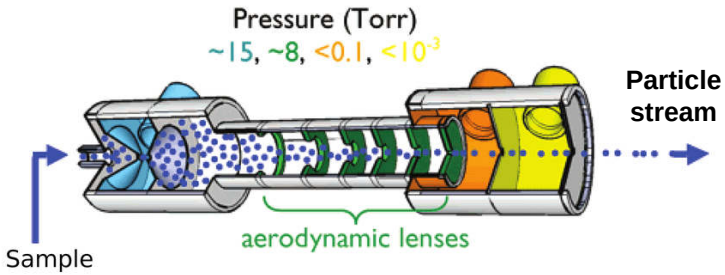
Conventional SAXS analysis

Higher-order SAXS analysis

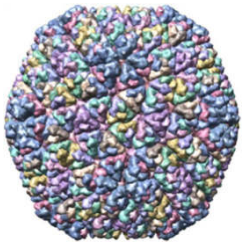
SPI experiments on aerosolized virus particles at the AMO instrument, LCLS



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.

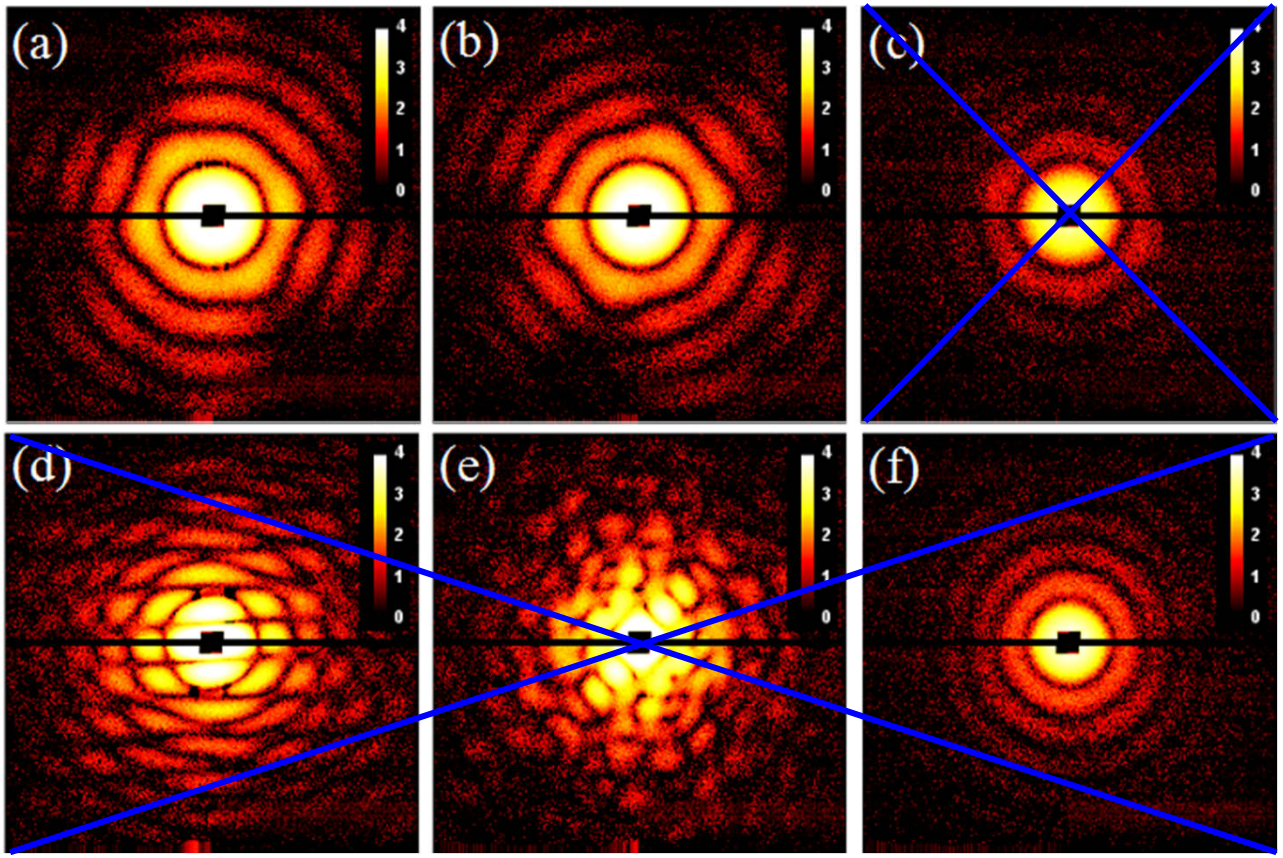
PR772

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)

Diffraction before destruction

- **Initial data classification was done using diffusion map embedding.**
- **Additional filtering was based on the average intensity/pixel threshold & manual rejection of outliers.**

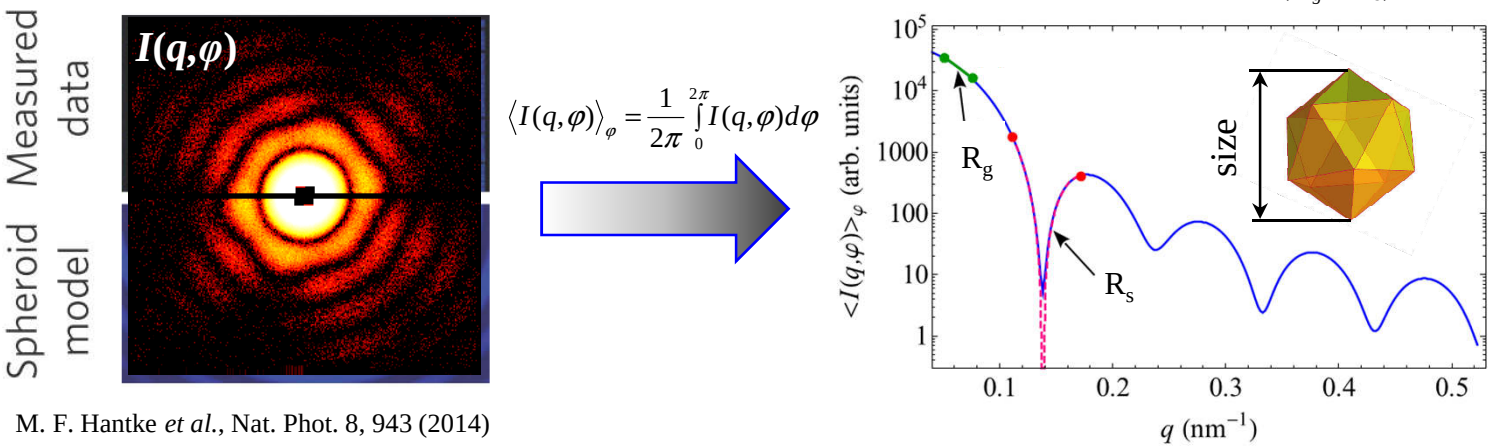


Measured data:
about **3*10^6** patterns were measured in the experiment.

After filtering:
760 patterns for RDV
and
1400 patterns for PR772
corresponding to high-intensity single-particle hits were selected for further analysis.

H. K. N. Reddy *et al.*, Scientific Data 4, 170079 (2017)

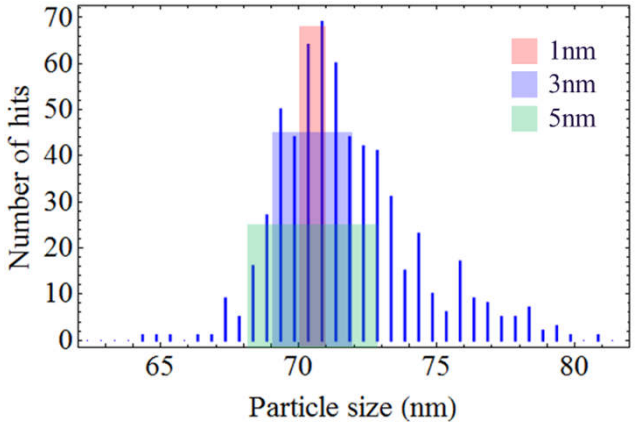
Size distribution histograms for RDV and PR772



M. F. Hantke *et al.*, Nat. Phot. 8, 943 (2014)

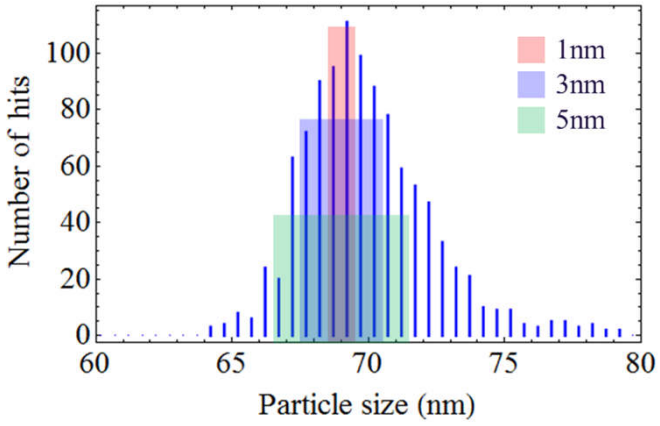
RDV

332 patterns @ PD=3nm



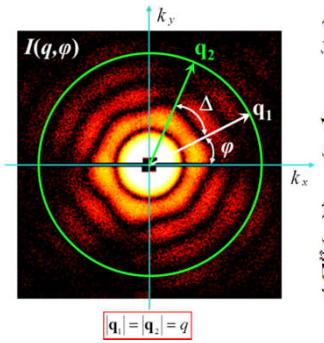
PR772

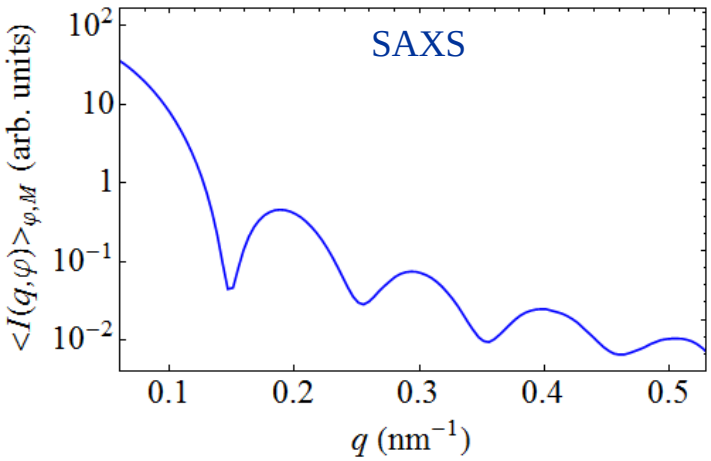
566 patterns @ PD=3nm

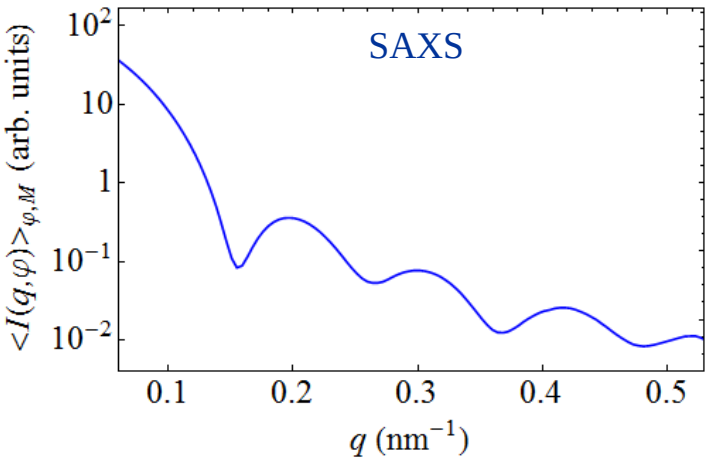


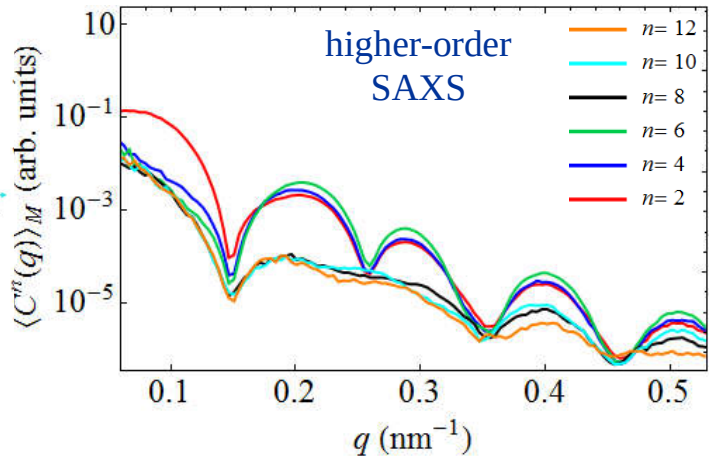
Generalization of SAXS: higher-order SAXS

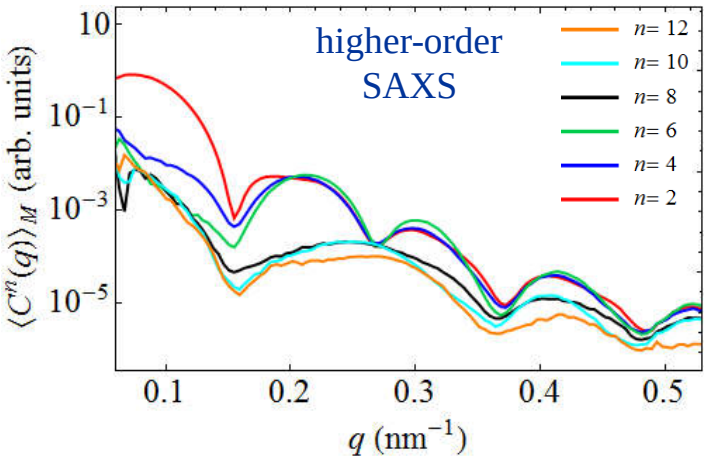
RDV **PR772**





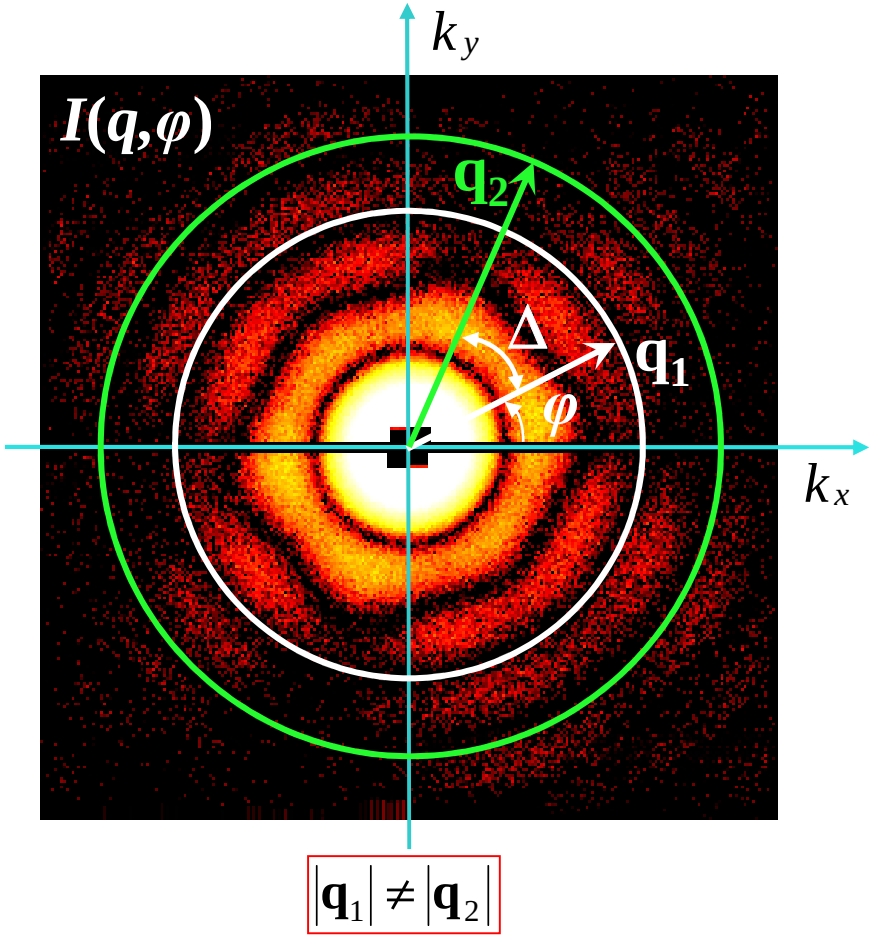






R.P. Kurta *et al.*, PRL 119, 158102 (2017)

Extended set of cross-correlations



Two-point CCF defined at two different resolution rings:

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

Fourier components (FCs) of the CCF:

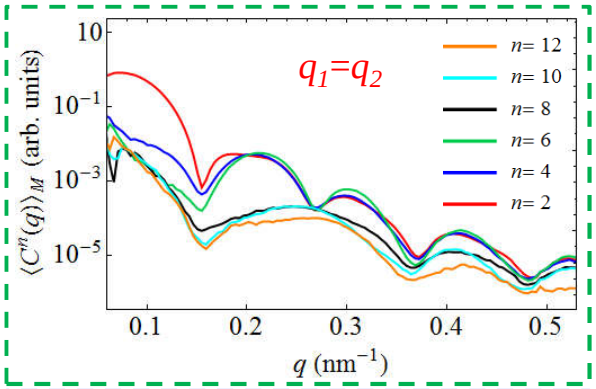
$$C^n(q_1, q_2) = \frac{1}{2\pi} \int_0^{2\pi} C(q_1, q_2, \Delta) \exp(-in\Delta) d\Delta$$

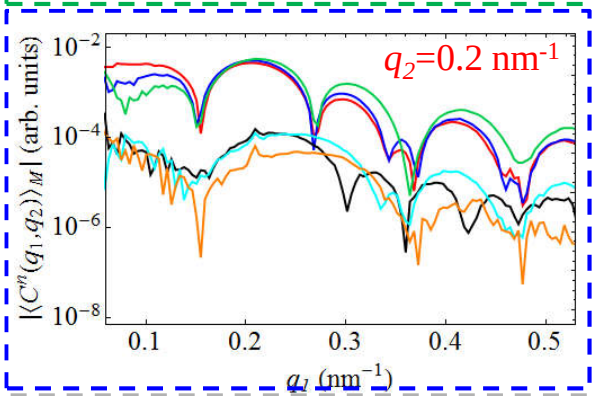
Additional information about particle structure:
$N_q \times n_{tot}$ times higher information content compared to SAXS, where N_q is the momentum transfer sampling, and n_{tot} is the total number of significant FCs.

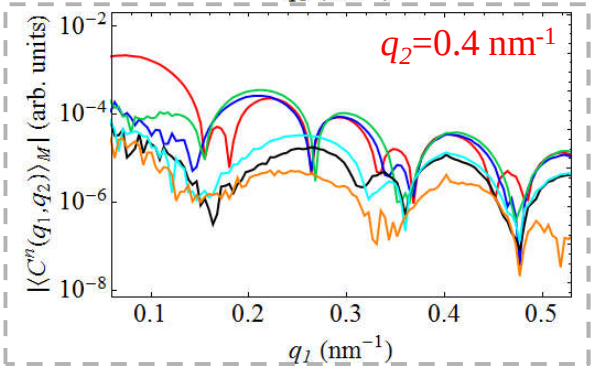
Z. Kam, Macromolecules 10, 927 (1977)
R.P. Kurta, M. Altarelli, I.A. Vartanyants, Adv. Chem. Phys. 161, Ch.1 (2016)

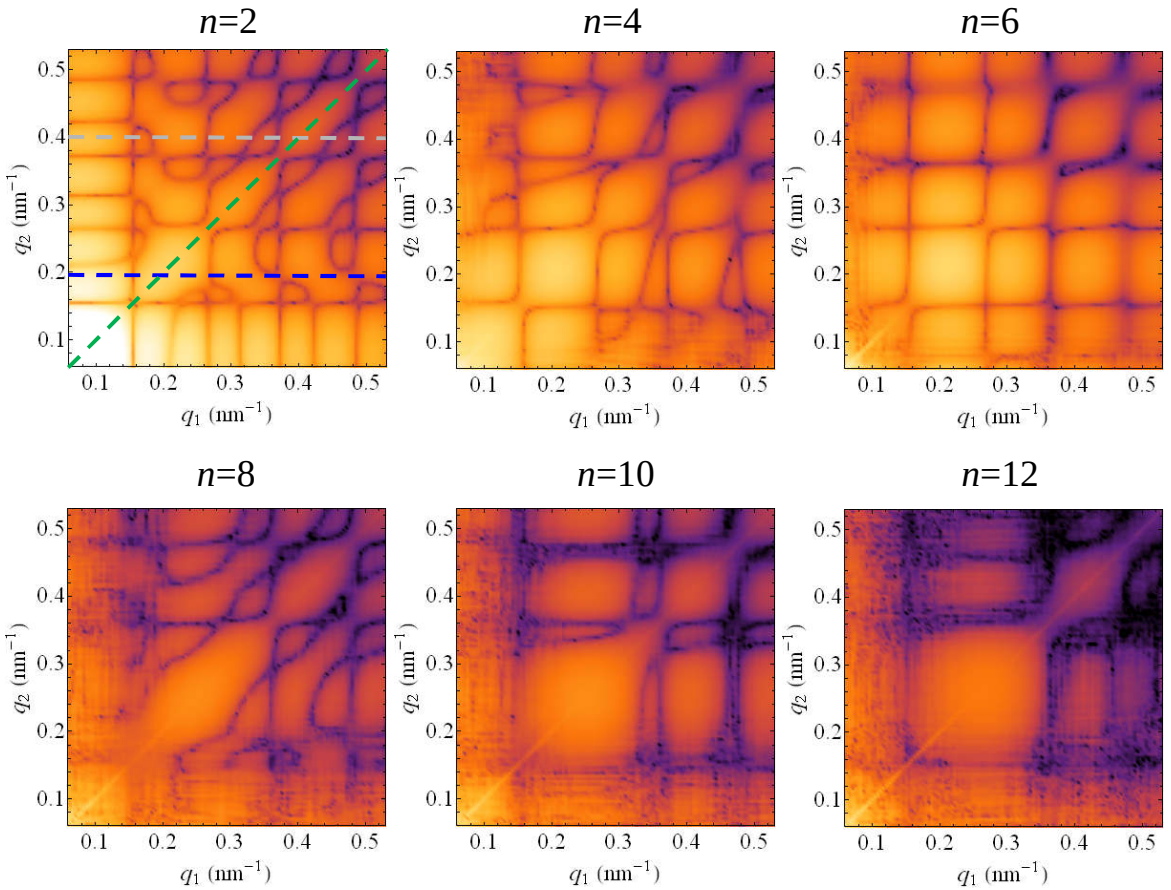
2D correlation maps

$$\left|\left\langle C^n(q_1,q_2)\right\rangle_M\right| = f(q_1,q_2) \text{ for } n = 2,4,6,8,10 \text{ and } 12$$

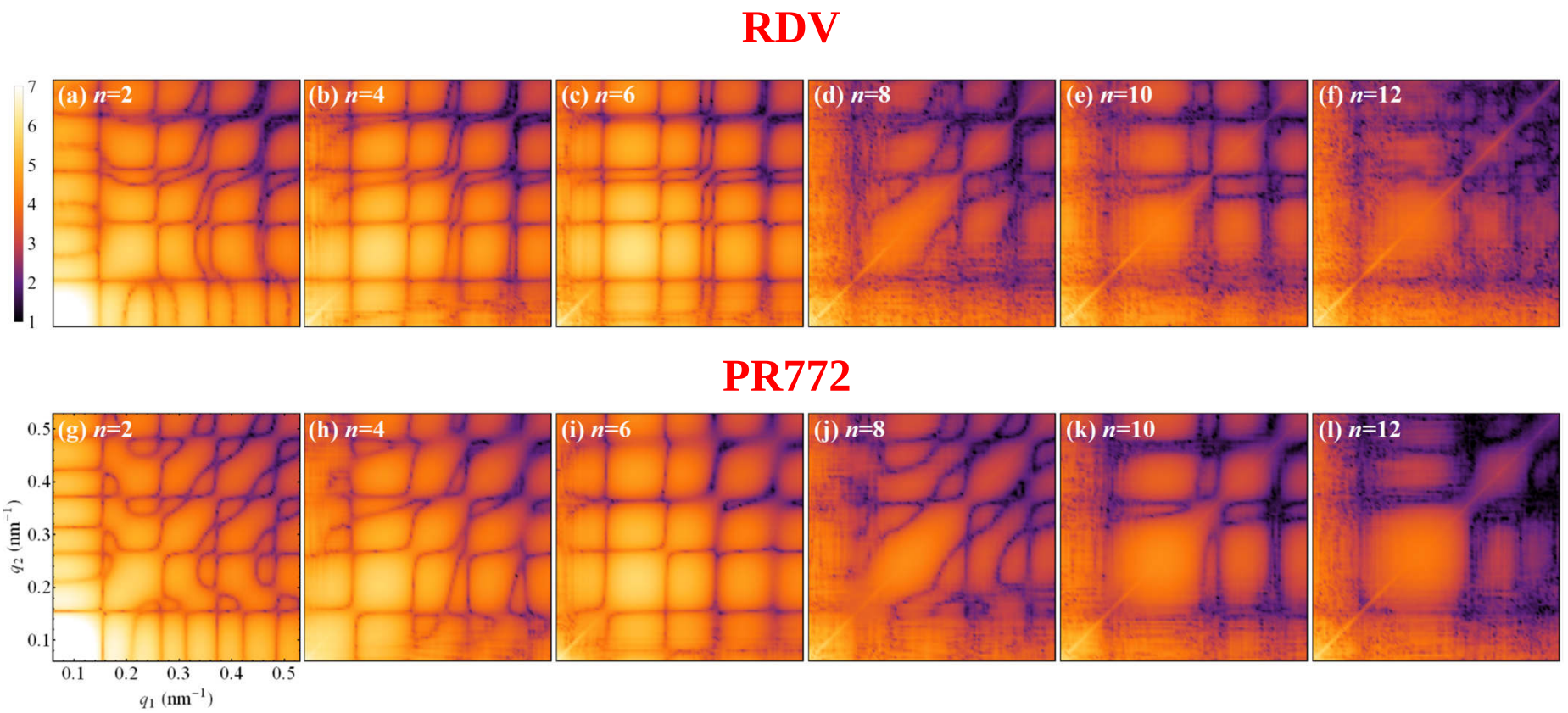








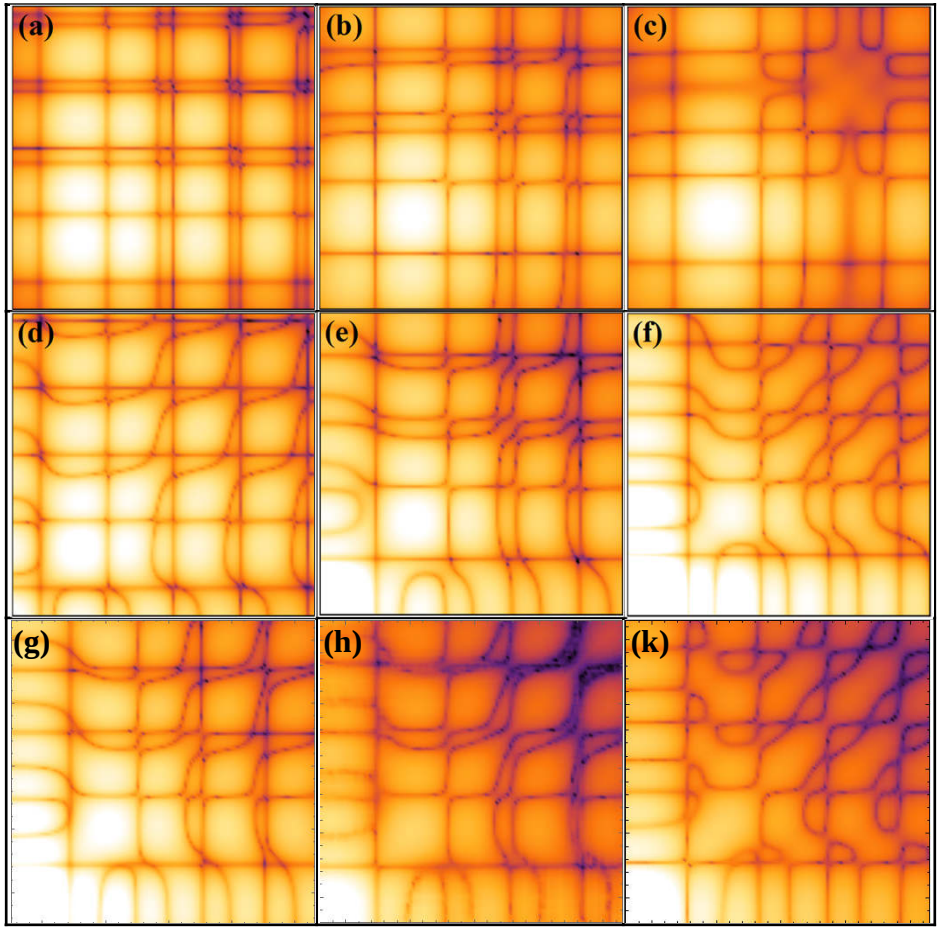
Experimental 2D correlation maps for RDV and PR772



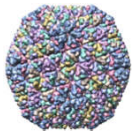
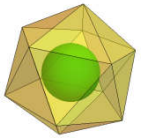
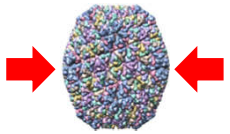
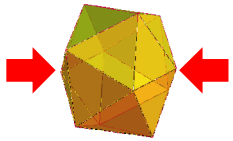
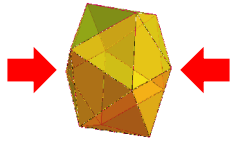
R.P. Kurta *et al.*, PRL 119, 158102 (2017)

Model-based structure analysis

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



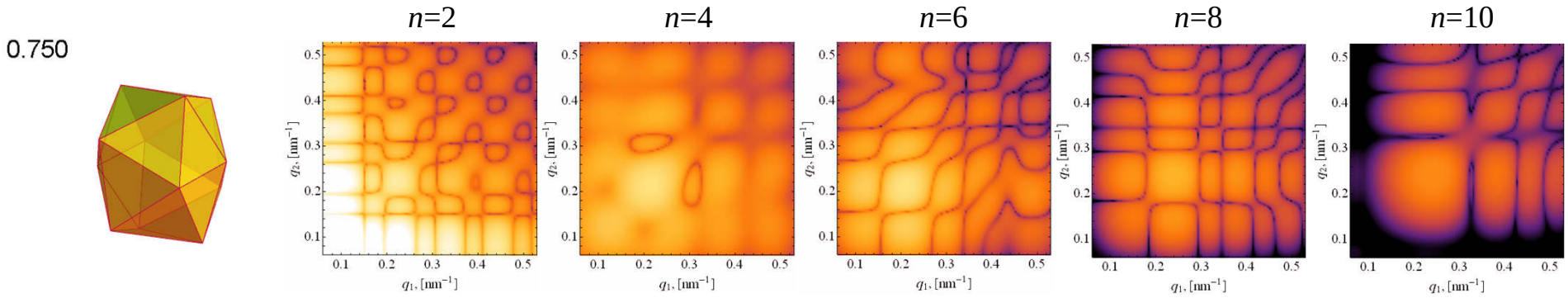
Model structures

(a) Empty RDV capsid (1UF2) 	(b) Solid icosahedron 	(c) Hollow icosahedron with a spherical void 
(d) Empty RDV capsid compressed by 3% 	(e) Solid icosahedron compressed by 3% 	(f) Solid icosahedron compressed by 7% 
(g) Solid icosahedron with ellipsoidal caking 	(h) Experiment RDV	(k) Experiment PR772

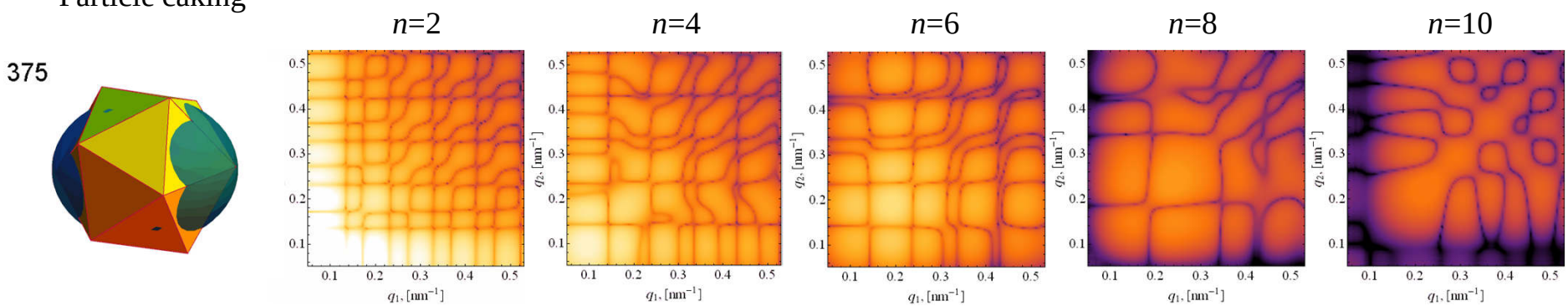
R.P. Kurta *et al.*, PRL 119, 158102 (2017)

Sensitivity of cross-correlations to structure variation

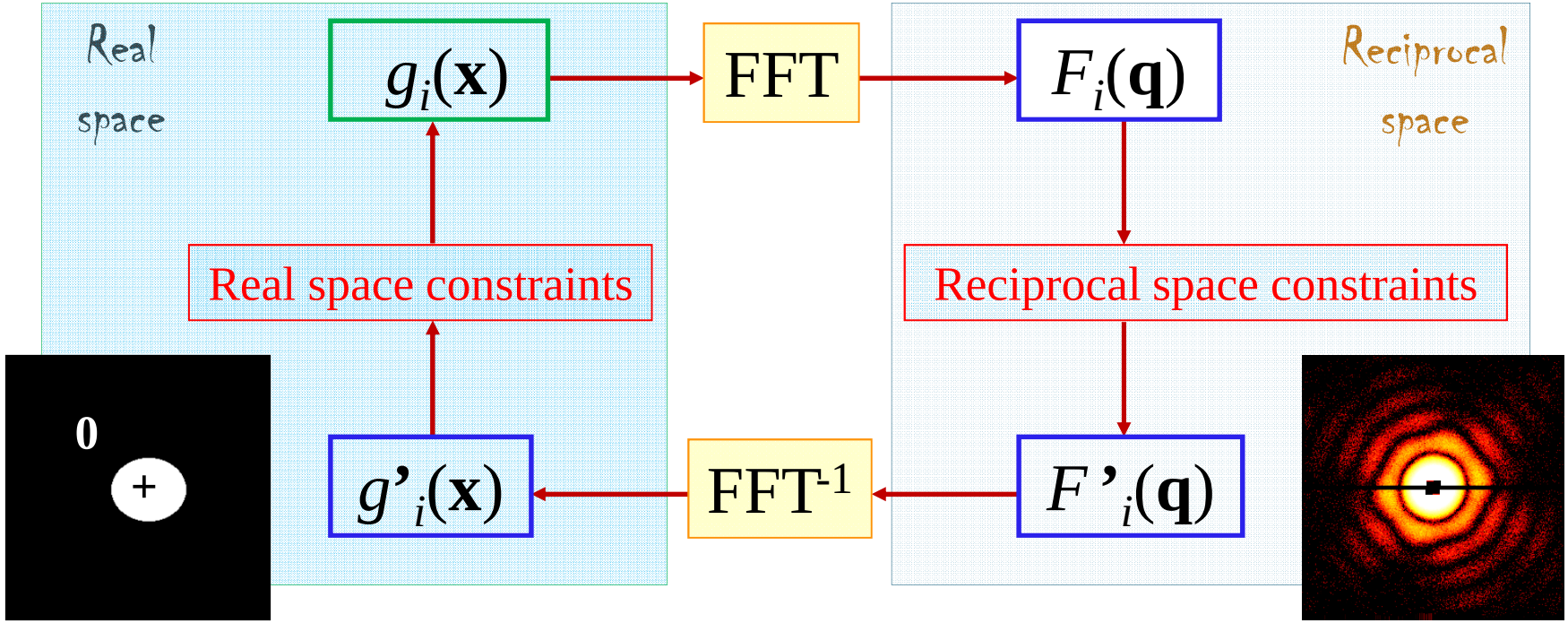
Particle distortion



Particle caking



Conventional iterative phase retrieval:
structure recovery from the measured intensities



Real space constraints:

- Finite support
- Positivity

Reciprocal space constraints:

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

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

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)

Multitiered iterative phasing (MTIP):
structure recovery from the measured correlations

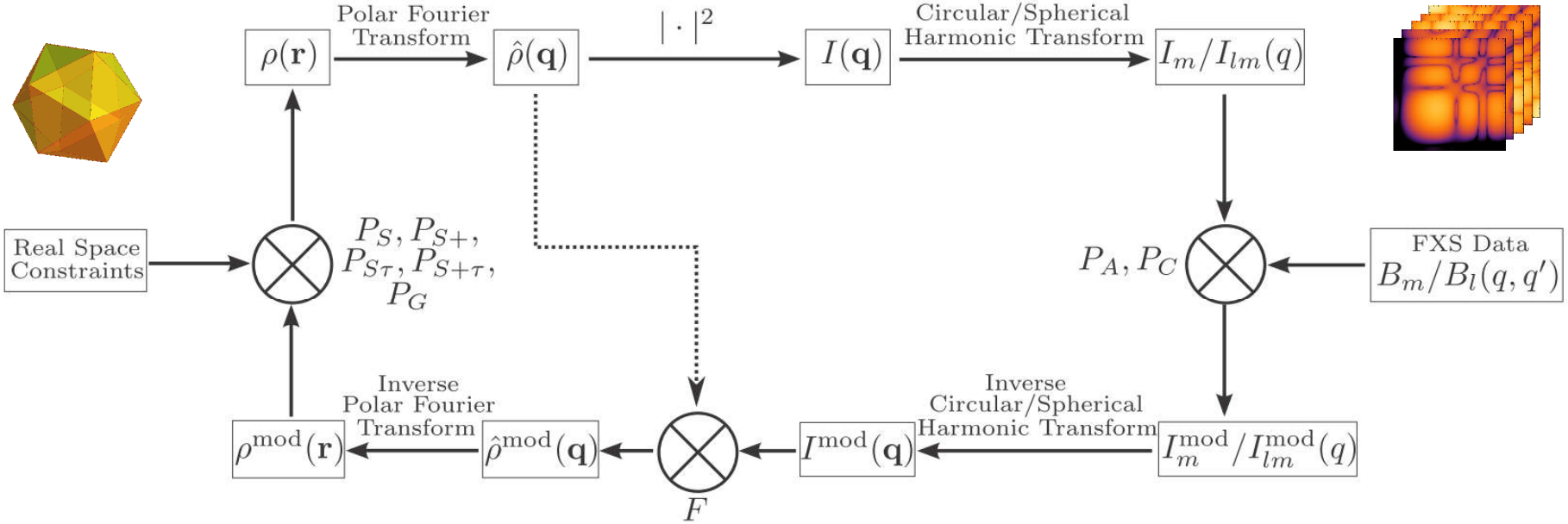
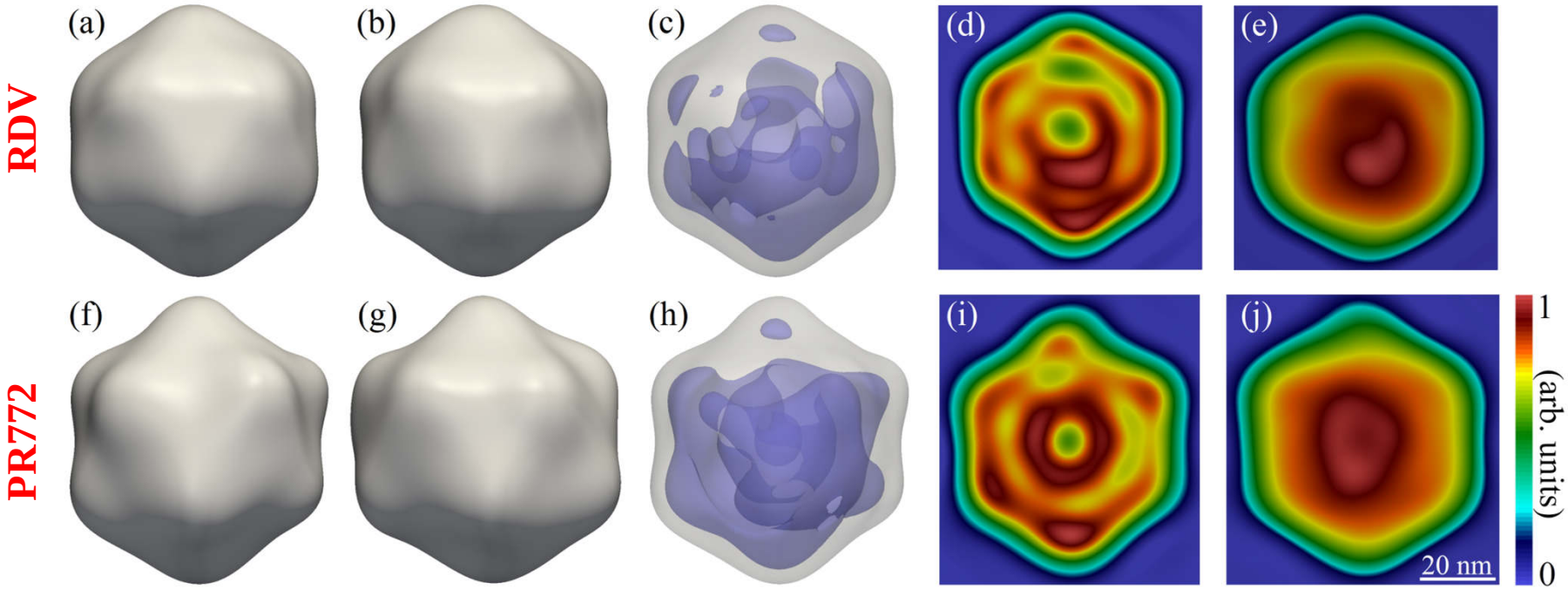


Fig. S7. Flowchart of the multitiered iterative phasing procedure. The mod superscript denotes values that have been modified to agree with the data. One uses circular harmonic transforms, I_m, and B_m for the 2D case and spherical harmonic transforms, I_{lm}, and B_l for the 3D case.

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

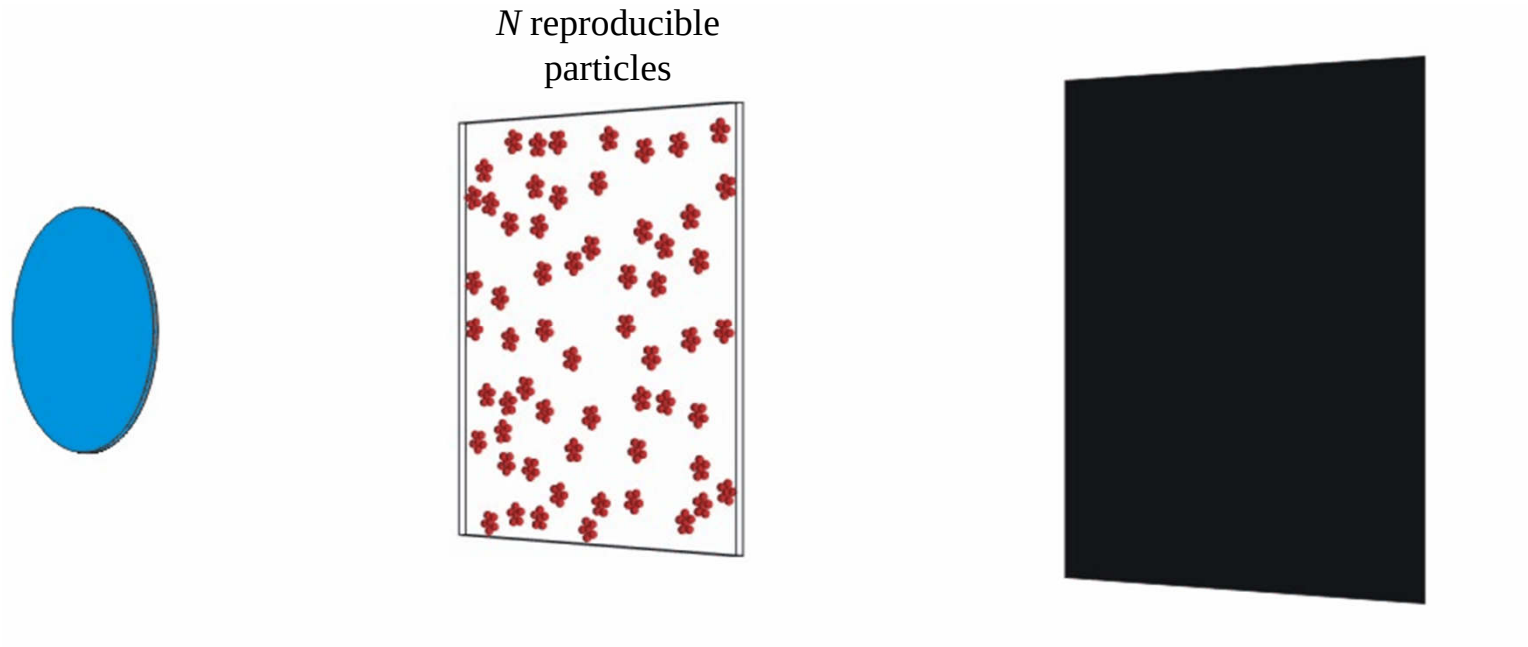
MTIP reconstructions of RDV and PR772



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)

Beyond single particle imaging: multiple-particle FXS



$$\left\langle C_{\text{multiple}}(q,\Delta)\right\rangle_M = N\left\langle C_{\text{single}}(q,\Delta)\right\rangle_M$$

$\left\langle C_{\text{single}}(q,\Delta)\right\rangle_M$ - single-particle CCF (N=1)

$\left\langle C_{\text{multiple}}(q,\Delta)\right\rangle_M$ - multiple particle CCF (N>1)

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

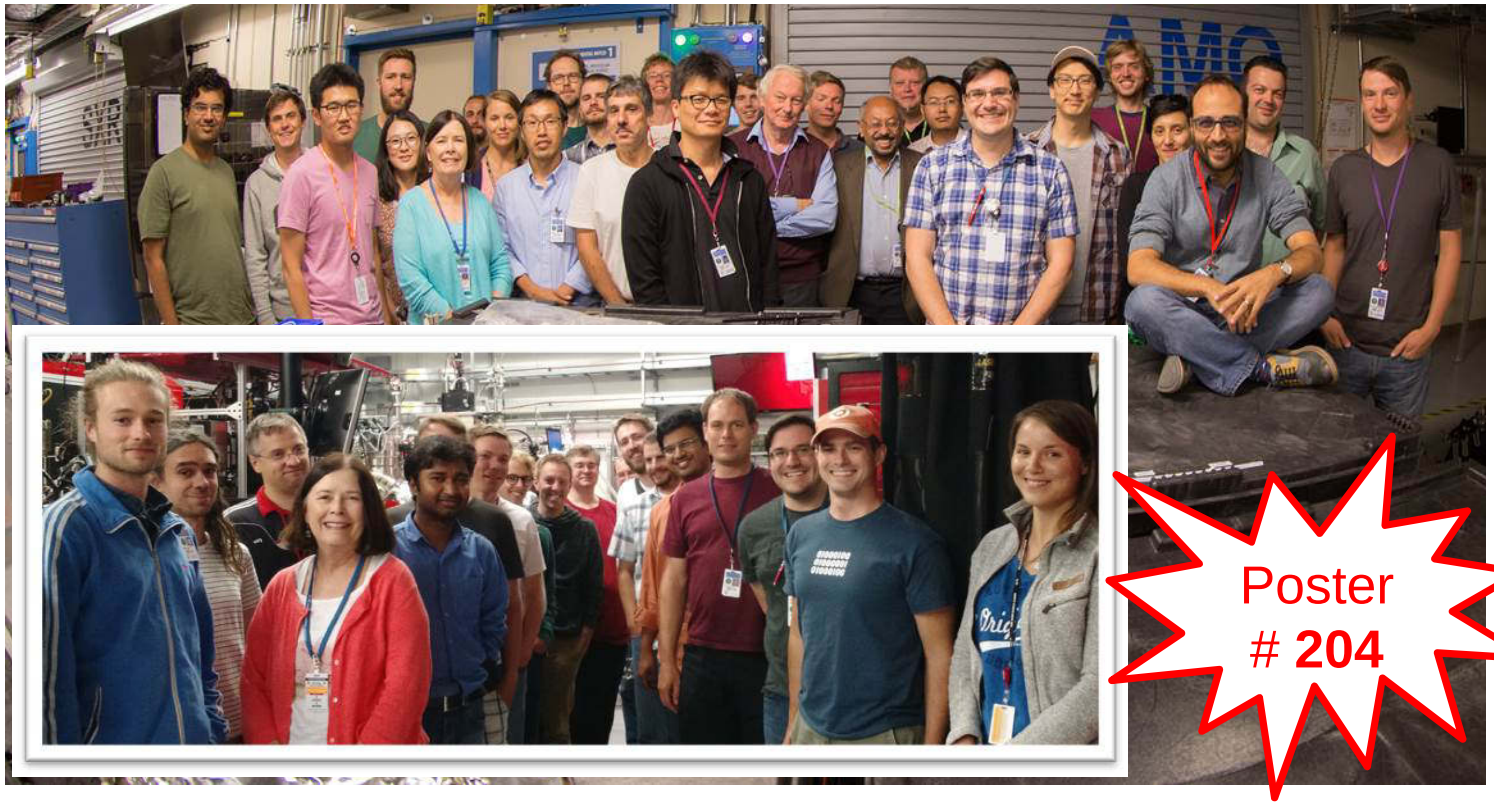
Summary

- 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 LCLS demonstrates substantial potential of the technique for the future studies of structure and dynamics of biomaterials with an XFEL.

Acknowledgements

J. J. Donatelli , P. H. Zwart, C. H. Yoon, A. Aquila, A. P. Mancuso, M. Altarelli, I. A. Vartanyants

SPI Initiative: **ARC (Australia), ASU (US), BNL (US), CBIS (Singapore), CFEL (Germany) , Cornell (US), CSRC (China), DESY (Germany), XFEL.EU (Germany), LBNL (US), LLNL (US), LMB (Sweden), PAL (Korea), PSI (Switzerland), PULSE (US), Rice (US), SLAC (US), Stanford (US), STU (China), TU-Berlin (Germany), UWM (US)**



A. Aquila, et al. The linac coherent light source single particle imaging road map, Struct. Dyn. 2, 041701 (2015).