High-intensity phasing with x-ray free-electron lasers

Sang-Kil Son,[1,*] Lorenzo Galli,[1,2] Henry N. Chapman,[1,2] and Robin Santra[1,2]
[1] Center for Free-Electron Laser Science (CFEL), DESY, Germany
[2] Department of Physics, University of Hamburg, Germany

Abstract

X-ray free-electron lasers (XFELs) show promise for revealing molecular structure using **serial femtosecond crystallography (SFX)**, but the associated phase problem remains largely unsolved. Many of the ab initio methods that are used for phasing diffraction data collected with synchrotron radiation employ anomalous scattering from heavy atoms, for example, **multiwavelength anomalous diffraction (MAD)**. Because of the extremely high intensity of XFELs, samples experience severe and unavoidable electronic radiation damage, especially to heavy atoms. The scattering factors of heavy atoms are dramatically changed due to ionization during an intense x-ray pulse, which hinders direct implementation of those phasing techniques with XFELs. A generalized version of MAD at high x-ray intensity has been proposed previously, suggesting that element-specific and fluence-dependent electronic damage could be used to determine phases. Here, we show two recent results towards a new **high-intensity phasing (HIP)**. We demonstrate that simulated SFX data of Cathepsin B can be phased by the different ionization degree of S atoms between two datasets at low and high x-ray fluences, similar to the technique of radiation induced phasing. We present an experimental evidence of the different ionization degree of Gd atoms between two datasets obtained from a Gd derivative of lysozyme microcrystals, which is used to identify the positions of the Gd atoms. New opportunities and challenges of high-intensity phasing methods with XFELs will be discussed.

Introduction

Femtosecond x-ray nanocrystallography

> One bottleneck of x-ray crystallography is the need for large-scale high-quality crystals, which are very difficult to be grown or are simply not available in many cases of interest.

> The unprecedented high x-ray fluence from XFELs provides a sufficiently large number of photons to enable structure determination from diffraction measurements of streams of single molecules and nanocrystals.

> Due to an extremely high fluence that is ~100 times larger than the conventional damage limit, samples are subject to severe damage.

> The ultrashort x-ray pulses generated by XFELs enable us to carry out “diffraction-before-destruction” within femtosecond timescales to suppress nuclear motion.

> Electronic radiation damage is unavoidable, which is characterized by multiphoton multiple ionization via a sequence of one-photon inner-shell ionizations and relaxations.

> Another bottleneck of x-ray crystallography, the phase problem, remains largely unsolved for femtosecond x-ray nanocrystallography.

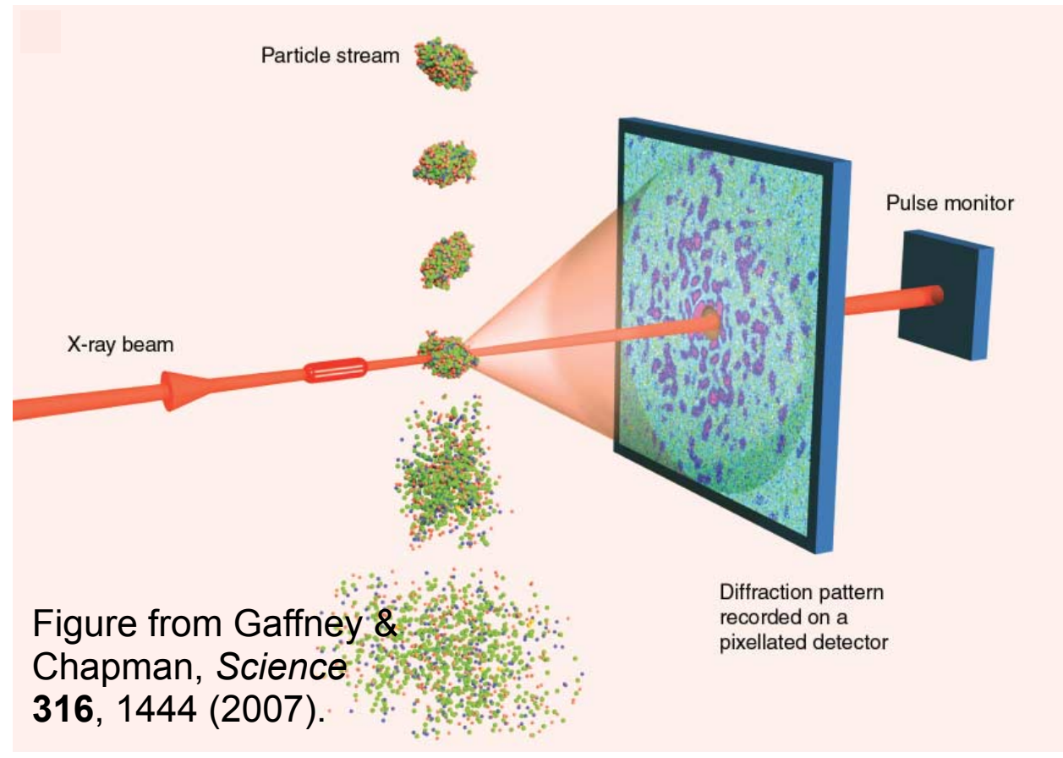


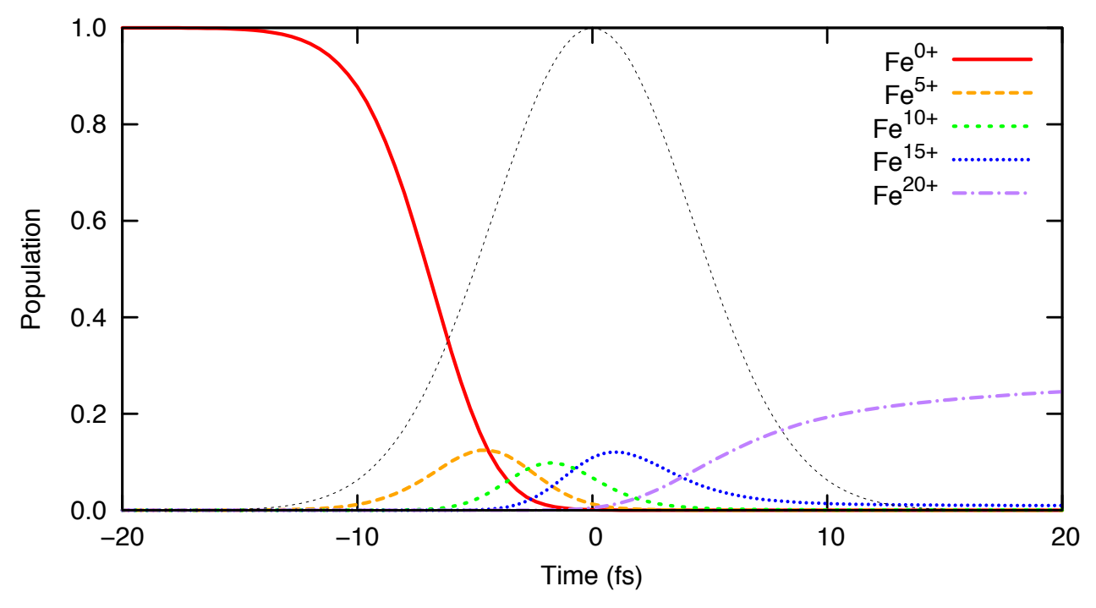
Figure from Gaffney & Chapman, *Science* **316**, 1444 (2007).

Electronic damage to heavy atoms

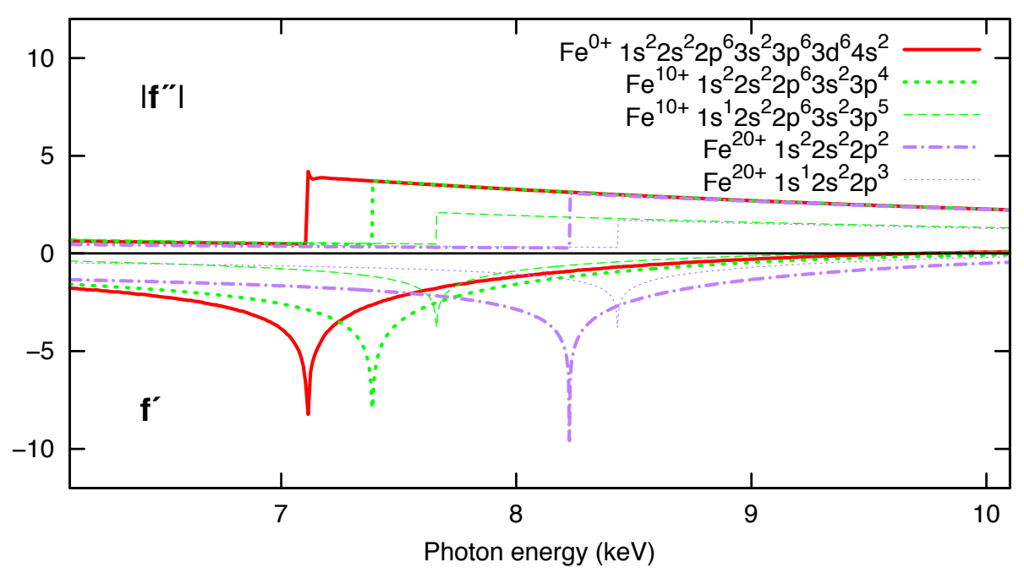
The ground-state configuration for neutral Fe,

$$\text{Fe: } 1s^2 2s^2 2p^6 3s^2 3p^6 3d^6 4s^2$$

For electronic damage dynamics of a Fe atom, 27,783 coupled rate equations are solved.



Population dynamics of Fe charge states during an XFEL pulse of 8 keV, 5×10^{12} photons/μm^2, 10 fs FWHM



Dispersion corrections of atomic form factors of Fe and its ions

High-intensity phasing (HIP)

> Multiwavelength anomalous diffraction (MAD): well-established phasing method with synchrotron radiation, employing the dispersion correction of x-ray scattering from heavy atoms

> Karle–Hendrickson equation: the key formula in MAD

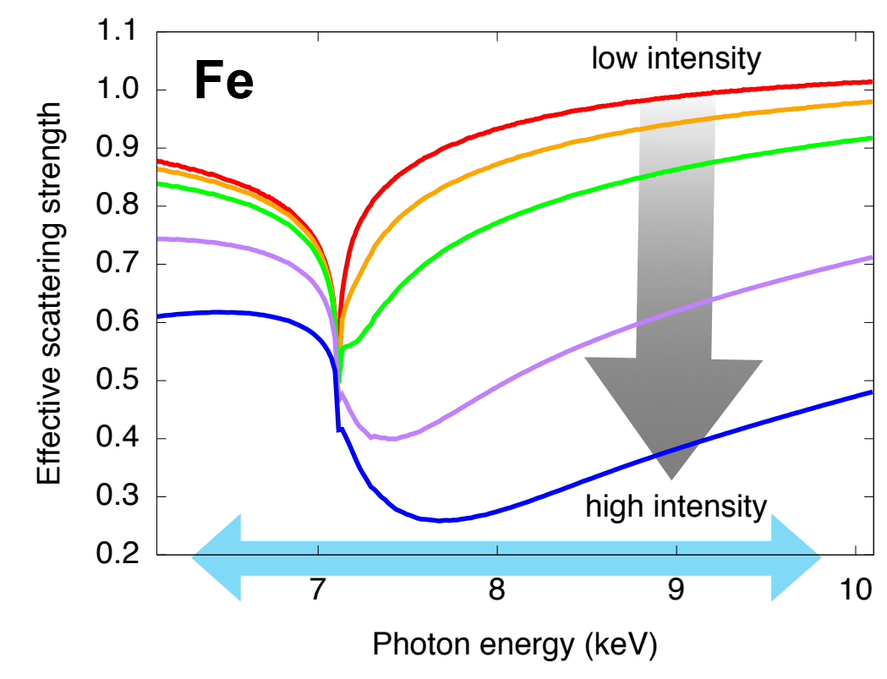
> Generalized Karle–Hendrickson equation: the key formula in MAD at high x-ray intensity, including ionization dynamics

$$\begin{aligned}\frac{dI(\mathbf{Q},\mathcal{F},\omega)}{d\Omega} = \mathcal{F}C(\Omega)\Big[&\left|F_P^0(\mathbf{Q})\right|^2 + \left|F_H^0(\mathbf{Q})\right|^2 \tilde{a}(\mathbf{Q},\mathcal{F},\omega)\\ &+ \left|F_P^0(\mathbf{Q})\right|\left|F_H^0(\mathbf{Q})\right| b(\mathbf{Q},\mathcal{F},\omega)\cos\Delta\phi^0(\mathbf{Q})\\ &+ \left|F_P^0(\mathbf{Q})\right|\left|F_H^0(\mathbf{Q})\right| c(\mathbf{Q},\mathcal{F},\omega)\sin\Delta\phi^0(\mathbf{Q})\\ &+ N_H\left|f_H^0(\mathbf{Q})\right|^2\{a(\mathbf{Q},\mathcal{F},\omega)-\tilde{a}(\mathbf{Q},\mathcal{F},\omega)\}\Big]\end{aligned}$$

> **3 unknown variables**: solvable with 3 measurements

> **MAD coefficients** are known: expressed by time evolution of form factors; to be measured or calculated

$$\tilde{a}(\mathbf{Q},\mathcal{F},\omega) = \frac{1}{\{f_H^0(\mathbf{Q})\}^2}\int_{-\infty}^{\infty} dt\, g(t)\left|\sum_{I_H} P_{I_H}(\mathcal{F},\omega,t) f_{I_H}(\mathbf{Q},\omega)\right|^2$$



Dependence of MAD coefficients on photon energy (ω) and fluence (*F*)

> Bleaching effect by high x-ray intensity: new path to phasing ➔ **HIP**

Native cathepsin B (heavy atom: S)

– simulated datasets @ 6 keV
– LCLS datasets @ 6 keV

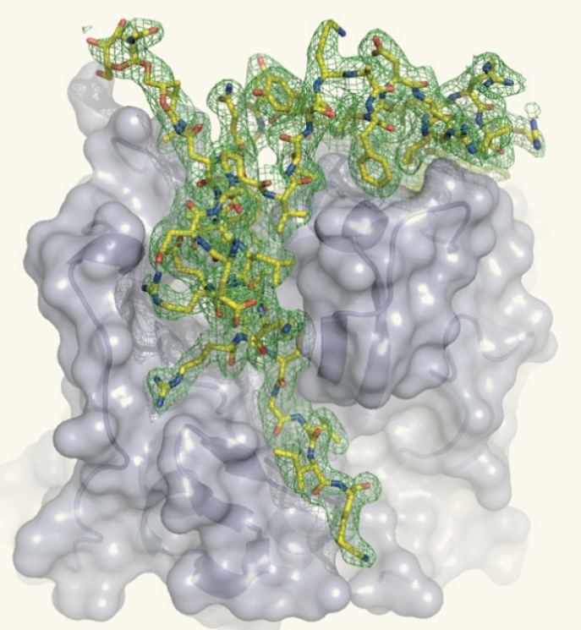


Figure from *Nature* **505**, 620 (2014).

Towards HI-RIP (high-intensity radiation-damage-induced phasing)

> Selective ionization of heavy atoms by using high x-ray intensity

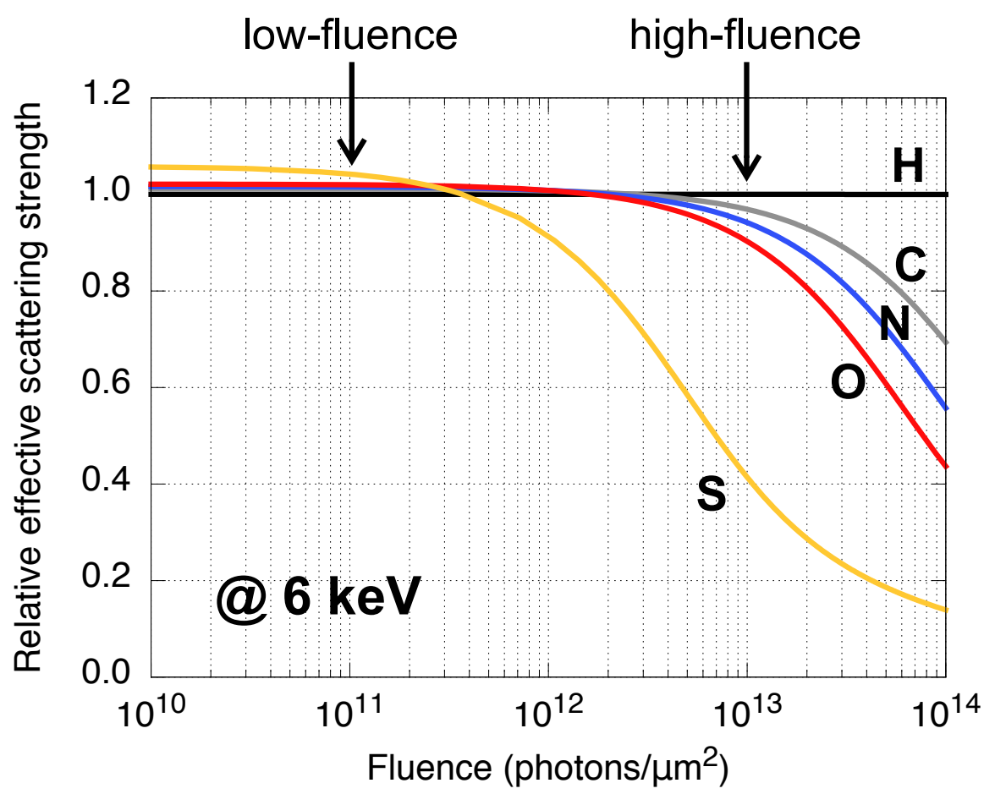
> Targeting S atoms in native proteins

> Like RIP (radiation-damage-induced phasing), but no chemical rearrangement during the x-ray pulses

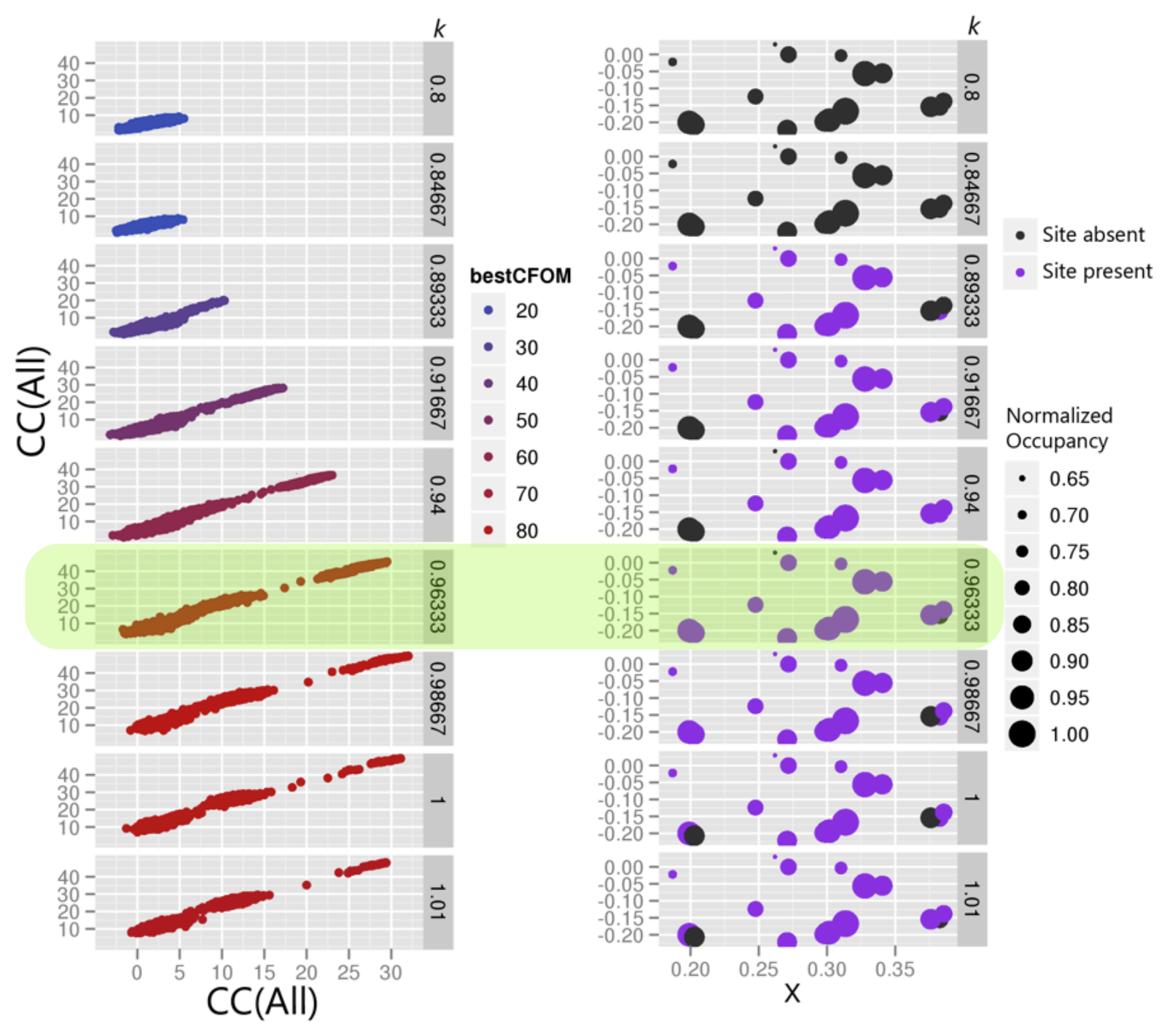
> Like SIR (single isomorphic replacement), but no atomic replacement in sample preparation

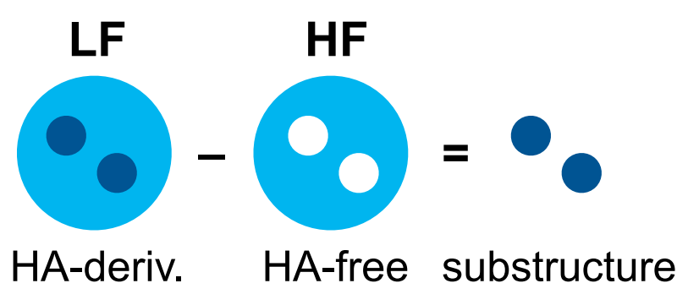
> Simulation of scattering patterns: stochastic ionization to every single atom in the sample

> Substructure determination and phasing with simulated LF/HF datasets using standard software adopted for RIP



Substructure determination



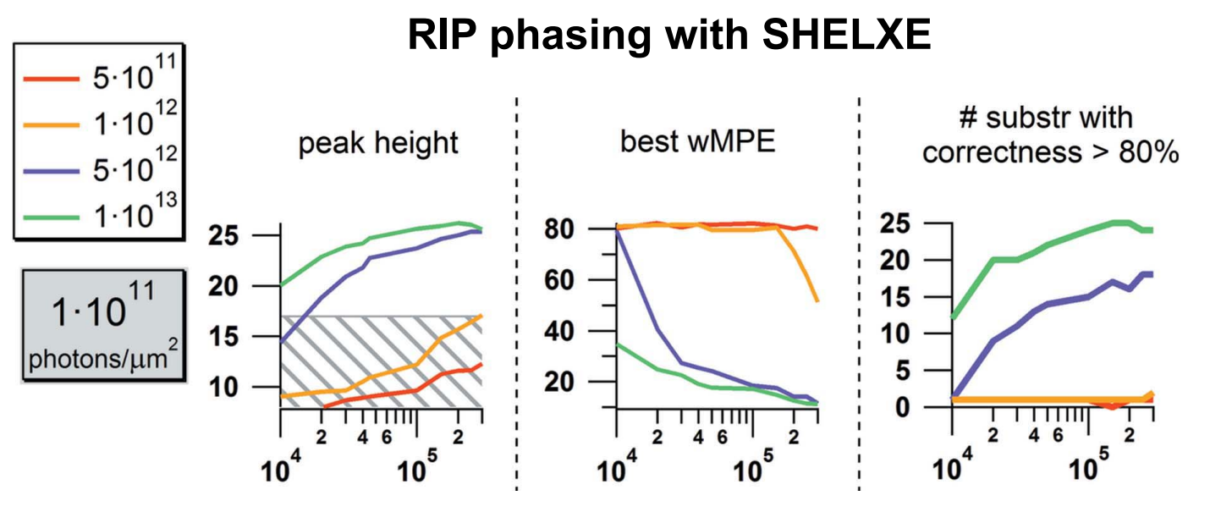


(Left) Correlation coefficient ratios with SHELXD as a function of scaling factor *k*, (Right) Correctness of sub-structures as a function of *k*, purple: correctly identified, black: not identified

Galli *et al.*, *J. Synch. Radiat.* (in press, 2015).

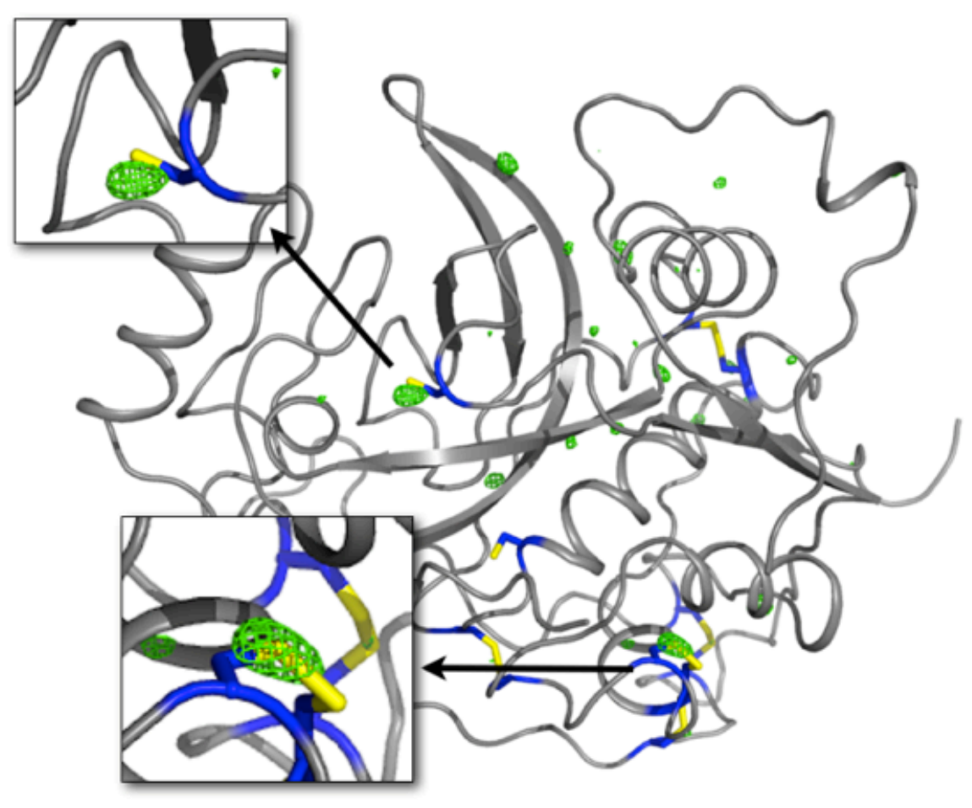
Phase determination

RIP phasing with SHELXE



Least # of patterns needed per dataset

HF (photons/μm^2)	LF (photons/μm^2): 10^{11}	$5\cdot10^{11}$	10^{12}	$5\cdot10^{12}$
$5\cdot10^{11}$	XXX			
10^{12}	XXX	XXX		
$5\cdot10^{12}$	20,000	30,000	40,000	
10^{13}	10,000	20,000	20,000	80,000



Substructure determination with experimental data

CXI beamline, 6 keV, 40 fs pulse duration, 3 mJ pulse energy, 0.1 μm focus

HF: est. 4×10^{12} ph/μm^2 ➔ ~37,000 patterns collected

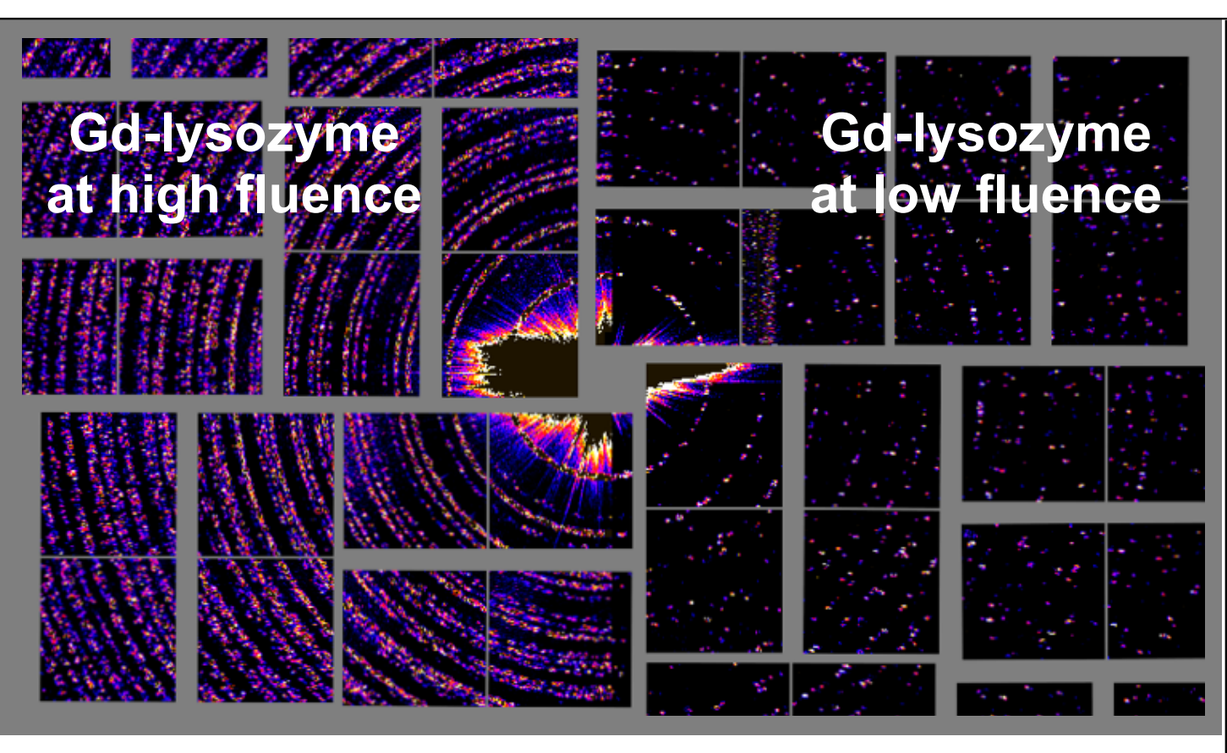
LF: 5~15% ➔ ~32,000 patterns collected

Phased difference (Fo–Fo) Fourier map superposed to the CatB model. Contoured at 4.5σ. S atoms are represented by yellow sticks.

Galli *et al.*, (in preparation).

Derivatized lysozyme (heavy: Gd)

– LCLS datasets @ 8.5 keV



Towards HIP (high-intensity phasing) with experimental data

> Low-fluence: Gd-derivatized lysozyme ➔ 391,214 patterns collected

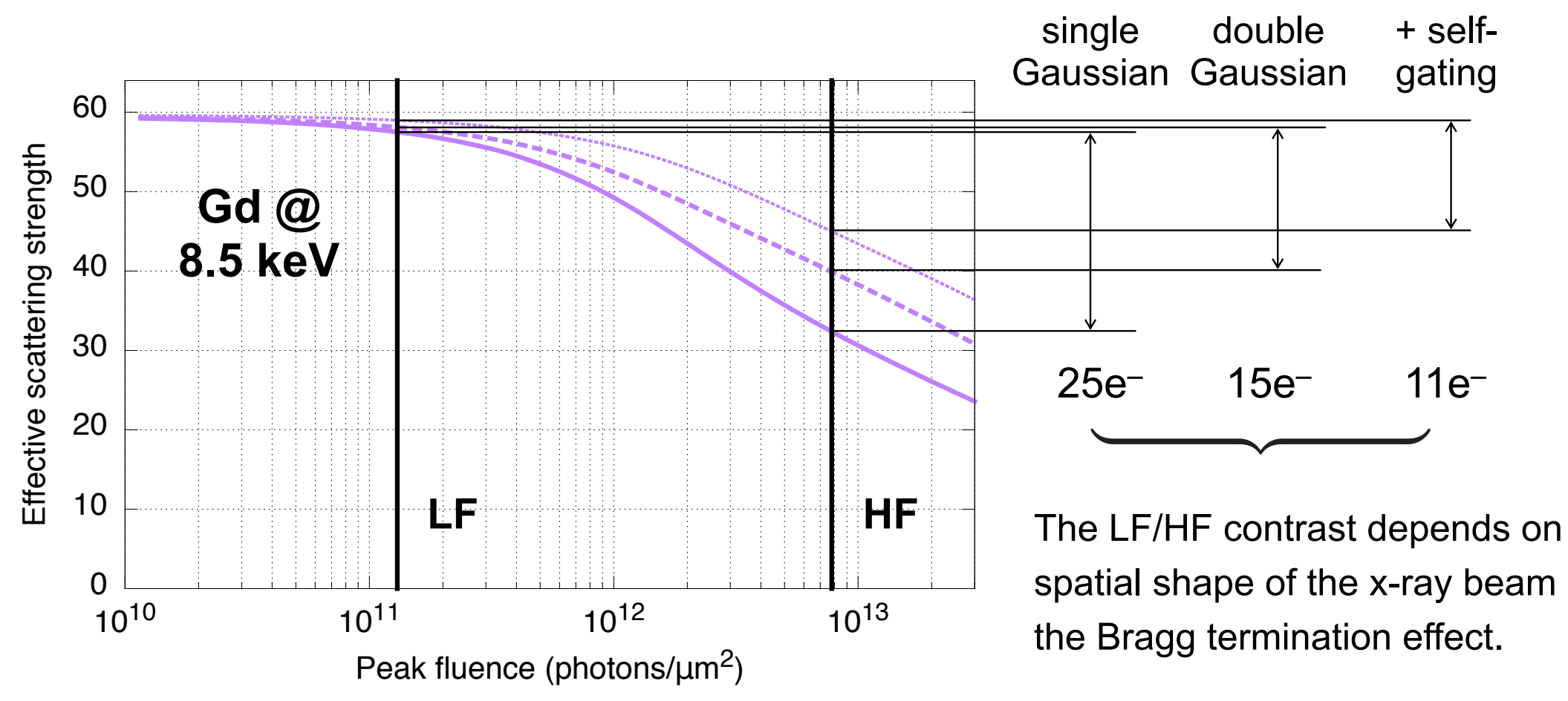
> High-fluence: native (Gd-free) lysozyme ➔ 591,966 patterns collected

> Demonstrate an experimental evidence of a scattering strength contrast between LF/HF datasets

Theoretical estimation

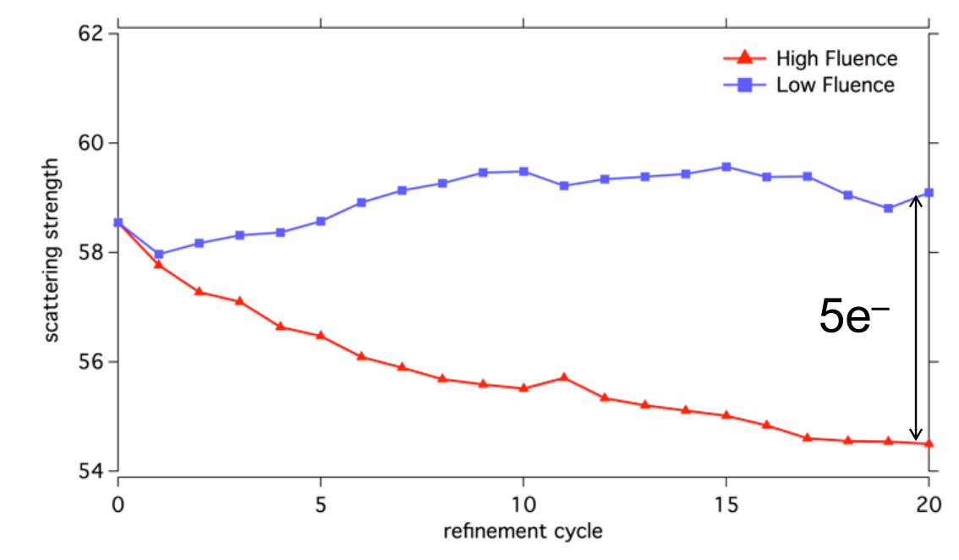
Scattering strength after volume averaging over spatial and temporal profile of the x-ray beam:

$$f_{\text{eff}} = \sqrt{\frac{\int d^3x \int dt \mathcal{F}(\mathbf{x})g(t)|\tilde{f}(\mathbf{Q},\mathcal{F},\omega,t)|^2}{\int d^3x \int dt \mathcal{F}(\mathbf{x})g(t)}}$$

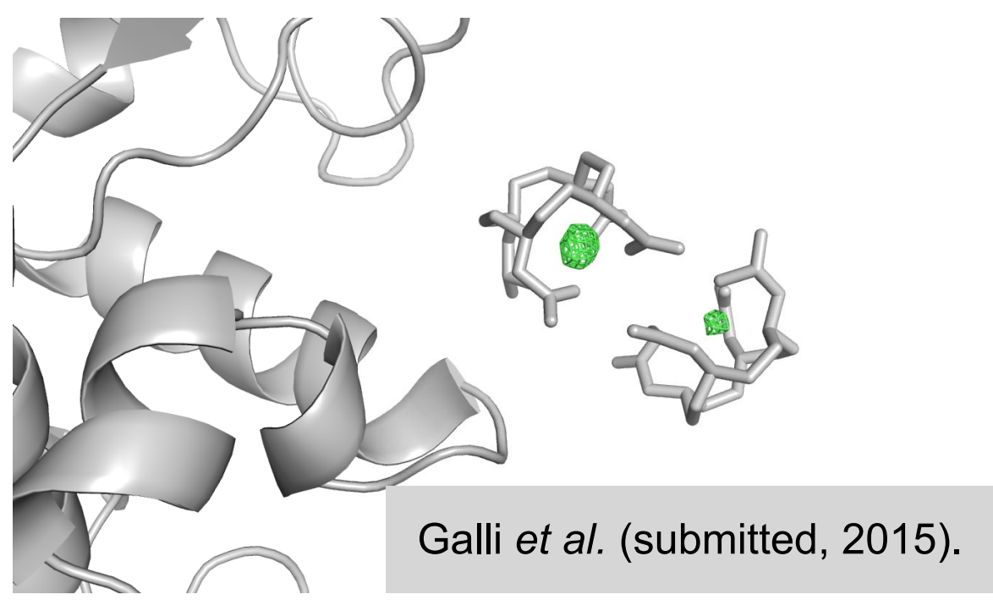


The LF/HF contrast depends on the spatial shape of the x-ray beam and on the Bragg termination effect.

Experimental analysis



Anomalous refinement of f’ and f”



Galli *et al.* (submitted, 2015).

Phased difference (Fo–Fc) Fourier map superposed to the Gd-Lys model. Data to 1.9 Å, contoured at 4.5σ. ➔ 8.8~12e$^-$

Issues on discrepancy between theory and experiment

> ionization dynamics based on an atomic model
  - relativistic treatment for heavy atoms
  - molecular environment
  - local plasma environment / collisional ionization

> unknown x-ray beam profile

> self-gating of the Bragg peaks

> ionization-induced fluctuation at high x-ray intensity

> crystal size

> scaling procedure

Theoretical estimation: 11~25e$^-$

Experimental analysis: 5~12e$^-$

$$I \propto \left|F_P^0 + \bar{f}\sum_{j=1}^{N_{\text{Gd}}} e^{i\mathbf{Q}\cdot\mathbf{R}_j}\right|^2 + N_{\text{Gd}}\bar{V}_1 + \left|\sum_{j=1}^{N_{\text{Gd}}} e^{i\mathbf{Q}\cdot\mathbf{R}_j}\right|^2 V_2$$

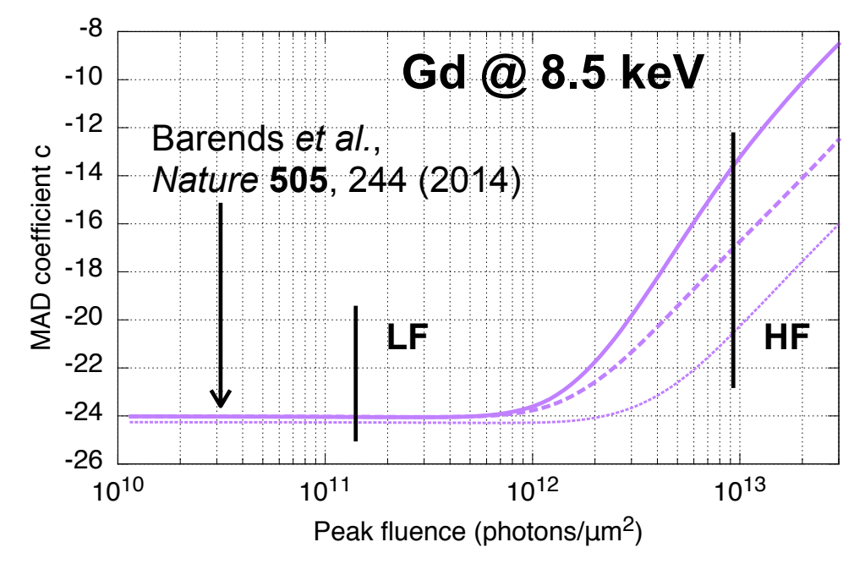
conventional crystallographic software

full expression from generalized KH equation

Towards HI-SAD (high-intensity single-wavelength anomalous diffraction)

SAD phasing with Phenix

	LF	HF
R(solve)	0.338	0.355
R(free)	0.264	0.457



XATOM toolkit

Theory

We implement an integrated toolkit, XATOM, to treat x-ray-induced processes based on nonrelativistic quantum electrodynamics and perturbation theory within the Hartree–Fock–Slater model.

Physical processes

> Photoionization

> Auger (Coster–Kronig) decay

> Fluorescence

> Shake-off

> Elastic x-ray scattering; inelastic x-ray scattering

> Resonant elastic x-ray scattering (dispersion correction)

Using the plasma extension, these processes can be treated with screening effect in a plasma environment.

Damage dynamics

To simulate electronic damage dynamics in intense x-ray pulses, we use the rate equation approach with photoionization cross sections, Auger rates, and fluorescence rates, for all possible *n*-hole electronic configurations for all possible +*n* charge states.

Applications

> Ionization, relaxation, and scattering dynamics at high intensity

> Charge distribution analysis of noble gases in XFELs

> Photoelectron / Auger / fluorescence spectra

> Multiwavelength anomalous diffraction at high intensity

> Ionization potential depression for ions embedded in plasmas

Conclusions

> Electronic radiation damage: unavoidable at high x-ray intensity

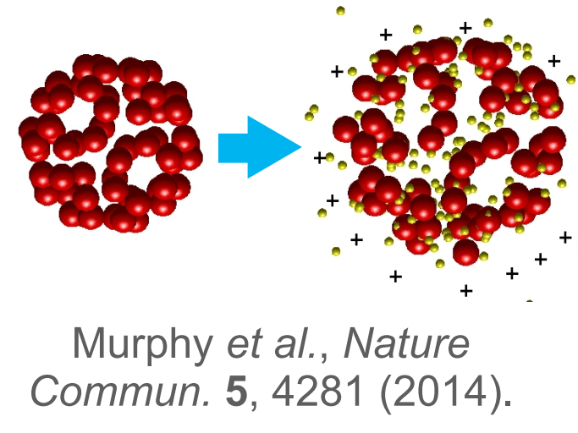
> Turning electronic radiation damage into an advantage for phasing

> Generalized Karle–Hendrickson equation in extreme conditions of ionizing radiations: to be used not only in phasing but also in refinement

> High-Intensity Phasing (HIP): new opportunities for solving the phase problem in nanocrystallography with XFELs

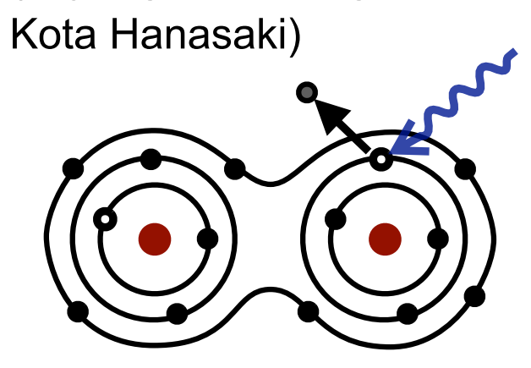
> Outlook: new software developments for better ionization model

XMDYN: molecular dynamics for all particles (Zoltan Jurek)



Murphy *et al.*, *Nature Commun.* **5**, 4281 (2014).

XMOLECULE: detailed description on molecules (Yajiang Hao, Ludger Inhester, Kota Hanasaki)



Hao *et al.*, (in preparation).

References

> HI-MAD and generalized Karle-Hendrickson equation:
Son, Chapman & Santra, *Phys. Rev. Lett.* **107**, 218102 (2011);
ibid., *J. Phys. B* **46**, 164015 (2013).

> S-CatB simulation: Galli *et al.*, *J. Synch. Radiat.* (in press, 2015).

> S-CatB experiment: Galli *et al.*, (in preparation).

> Gd-Lys experiment: Galli *et al.*, (submitted).

> XATOM: Son, Young & Santra, *Phys. Rev. A* **83**, 033402 (2011).

* sangkil.son@cfel.de

Center for Free-Electron Laser Science
CFEL is a scientific cooperation of the three organizations:
DESY – Max Planck Society – University of Hamburg