

## Structural variability and the incoherent addition of scattered intensities in single-particle diffraction

Filipe R. N. C. Maia,<sup>\*</sup> Tomas Ekeberg, Nicușor Tîmneanu, David van der Spoel, and Janos Hajdu<sup>†</sup>  
*Department of Cell and Molecular Biology, Uppsala University, Husargatan 3, P.O. Box 596, Uppsala SE-75124, Sweden*

(Received 1 July 2009; published 18 September 2009)

X-ray lasers may allow structural studies on single particles and biomolecules without crystalline periodicity in the samples. We examine here the effect of sample dynamics as a source of structural heterogeneity on the resolution of the reconstructed image of a small protein molecule. Structures from molecular-dynamics simulations of lysozyme were sampled and aligned. These structures were then used to calculate diffraction patterns corresponding to different dynamic states. The patterns were incoherently summed and the resulting data set was phased using the oversampling method. Reconstructed images of hydrated and dehydrated lysozyme gave resolutions of 3.7 Å and 7.6 Å, respectively. These are significantly worse than the root-mean-square deviation of the hydrated (2.7 Å for all atoms and 1.45 Å for C- $\alpha$  positions) or dehydrated (3.7 Å for all atoms and 2.5 Å for C- $\alpha$  positions) structures. The noise introduced by structural dynamics and incoherent addition of dissimilar structures restricts the maximum resolution to be expected from direct image reconstruction of dynamic systems. A way of potentially reducing this effect is by grouping dynamic structures into distinct structural substates and solving them separately.

DOI: [10.1103/PhysRevE.80.031905](https://doi.org/10.1103/PhysRevE.80.031905)

PACS number(s): 82.53.Ps, 42.30.Wb, 87.64.Bx, 82.37.Rs

### I. INTRODUCTION

The determination of macromolecular structures is of fundamental importance to the progress of biology. Nowadays most of the high-resolution structure determination is done using x-ray crystallography. Application of x-ray crystallography has been hugely successful for resolving protein structures, as seen by the exponential growth of the Protein Data Bank (PDB). Nevertheless, many proteins of biological interest are very difficult or impossible to crystallize. As such, an alternative method for high-resolution structure determination would be extremely valuable.

X-ray lasers can potentially allow high-resolution imaging of single macromolecules without the need for crystals [1,2]. In principle the main factor limiting the resolution achievable by this technique is the radiation damage caused by the intense pulse, which pulverizes the sample [3–6]. The extremely short duration of the pulse enables one to mitigate this problem as it becomes possible to record the diffraction image of the object before it explodes, as has already been demonstrated for micron sized objects [7].

The proposed experiment for imaging individual proteins with a free-electron laser (FEL) can be described in several steps. First a series of proteins are injected in the experimental chamber and exposed to the x-ray beam. The resulting diffraction patterns are then collected, oriented, and averaged in order to create a three-dimensional (3D) diffraction pattern [8–10]. Averaging is necessary as the expected signal from a single shot is very low but sufficiently strong to be used for orientation by recently proposed techniques [9,11]. The resulting 3D diffraction pattern can then be phased by employing the fact that the sample is isolated [12] and to finally obtain the electron density structure.

The resolution limit imposed by radiation damage can only be achieved, for 3D reconstructions, in the case of perfectly reproducible samples. Unfortunately most samples of biological interest (and indeed many other samples) are not reproducible at the atomic level. In particular proteins are inherently flexible and dynamic when viewed at atomic resolution [13]. Even with noiseless diffraction patterns and perfect orientation this atomic variability will introduce noise in the 3D diffraction patterns which will in the end limit the maximum resolution that can be obtained. This is due to the incoherent averaging of diffraction patterns from different structures. In x-ray crystallography the diffractions from the different units cells are combined coherently so the variability between unit cells simply results in a decrease in the signal at high resolution usually modeled using a Debye-Waller factor [14] or through ensemble refinement as recently shown [15].

While the effects of radiation damage, low photon count and orientation, on the achievable resolution have been extensively analyzed, in the context of single-particle imaging [1,4,16–20], there are no publications on the effects of sample variability. Sample variability comes not only from the dynamics of the molecules under study but also from the variability in the hydration layers that will surround many of these molecules. As the wavelength and the pulse duration of the x-ray FELs source becomes shorter, sample variability will be the main factor in determining the achievable resolution; therefore, it is essential to study this effect.

The aim of this study is to quantify the effect of protein motion in the resolution of the reconstructed images. We start by presenting a simple analytical model of the effect of the vibration of two connected atoms on a time integrated diffraction pattern. We then calculate the diffraction pattern resulting from the incoherent sum of diffraction patterns of different conformations of a protein. The resulting pattern is phased and the reconstructed image is then compared to the average protein structure and its resolution is estimated.

<sup>\*</sup>filipe@xray.bmc.uu.se

<sup>†</sup>janos@xray.bmc.uu.se

## II. EFFECT OF ATOM VIBRATION ON AN AVERAGED DIFFRACTION PATTERN

The diffraction pattern of an object is proportional to the square of the absolute value of its wave function at the detector  $I = \psi\psi^*$ . In the far field and ignoring polarization effects  $\psi$  can be approximated by the Fourier transform of the electron density of the object. For a one-dimensional system composed of two vibrating pointlike atoms connected by a bond of length  $l$ ,  $\psi$  as a function of the scattering vector  $q$  is given by

$$\begin{aligned} \psi(q) &= \mathcal{F}[\delta(x + l/2 + \epsilon_1) + \delta(x - l/2 + \epsilon_2)] \\ &= \exp[i\pi q(l + 2\epsilon_1)] + \exp[-i\pi q(l - 2\epsilon_2)], \end{aligned} \quad (1)$$

where  $\delta$  is the Dirac delta function and  $\epsilon_1$  and  $\epsilon_2$  represent small deviations from the average position of each atom.

The instantaneous diffracted intensity is then given by

$$I_{inst}(q) = \psi\psi^* = 2 + 2 \cos[2\pi q(l + \epsilon_1 - \epsilon_2)]. \quad (2)$$

Assuming that the deviations  $\epsilon_1$  and  $\epsilon_2$  are normally distributed throughout multiple exposures with zero mean and identical standard deviation  $\sigma$  as described by

$$G(\epsilon, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\epsilon^2}{2\sigma^2}\right), \quad (3)$$

we get that the diffraction pattern incoherently averaged over those exposures ( $I_{inc}$ ), as proposed in the FEL experiments, is given by

$$\begin{aligned} I_{inc}(q) &= \iint_{-\infty}^{+\infty} \psi\psi^* G(\epsilon_1, \sigma) G(\epsilon_2, \sigma) d\epsilon_1 d\epsilon_2 \\ &= 2 + 2 \cos(2\pi ql) \exp(-4\pi^2 q^2 \sigma^2). \end{aligned} \quad (4)$$

The vibration of the atoms results in a scaling of the interference term in the diffraction pattern by  $\exp(-4\pi^2 q^2 \sigma^2)$  as seen in Fig. 1.

The diffraction intensities of the same system when the amplitudes are added coherently give rise to the following pattern:

$$I_{coh}(q) = [2 + 2 \cos(2\pi ql)] \exp(-4\pi^2 q^2 \sigma^2). \quad (5)$$

In this case the entire diffraction pattern is modulated by  $\exp(-4\pi^2 q^2 \sigma^2)$ .

Following the treatment in [21] one can easily extend the model to three dimensions and  $N$  spherical atoms with atomic form factors  $f_i$  and average positions  $\bar{\mathbf{r}}_i$ , all with the same Gaussian vibration with a standard deviation of  $\sigma$ . The diffraction intensities are composed of self terms

$$I_{self}(\mathbf{q}) = I_{self}(q) = \sum_{i=1}^N |f_i(q)|^2, \quad (6)$$

as well as cross terms

$$I_{cross}(\mathbf{q}) = \sum_{i \neq j} f_i(q) f_j^*(q) \exp[-2\pi i \mathbf{q} \cdot (\bar{\mathbf{r}}_i - \bar{\mathbf{r}}_j)], \quad (7)$$

the latter being the only one that contain information about the relative position of the atoms. The incoherently added intensities are then given by

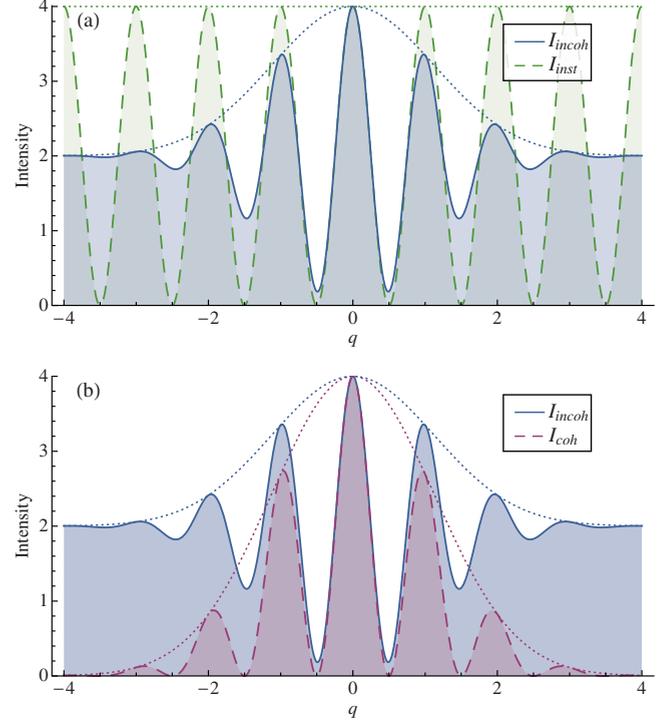


FIG. 1. (Color online) Modulation of the diffraction intensity on a 1D bond of length  $l=1$  due to vibration of pointlike atoms as a function of the scattering vector  $q$ . (a) The dashed line represents the instantaneous diffraction intensity ( $I_{inst}$ ) for  $\epsilon_1 = \epsilon_2 = 0$ . The solid line represents the incoherently added intensities ( $I_{inc}$ ) for an atom vibration  $\sigma=0.1$ . (b) The solid line corresponds to the same as in (a) while the dashed line shows the intensities resulting from the coherently added amplitudes using the same parameters. The envelopes of the intensities are drawn as dotted lines.

$$I_{inc}(\mathbf{q}) = I_{self}(\mathbf{q}) + \exp(-4\pi^2 \sigma^2 q^2) I_{cross}(\mathbf{q}), \quad (8)$$

while the intensities resulting from the coherently added amplitudes are

$$I_{coh}(\mathbf{q}) = \exp(-4\pi^2 \sigma^2 q^2) [I_{self}(\mathbf{q}) + I_{cross}(\mathbf{q})]. \quad (9)$$

For a system with a large number of randomly distributed atoms,  $I_{cross}$  averaged over constant  $q = |\mathbf{q}|$  is given by

$$\overline{I_{cross}(q)} = \sum_{i \neq j} f_i(q) f_j^*(q) \frac{\sin(2\pi q |\bar{\mathbf{r}}_i - \bar{\mathbf{r}}_j|)}{2\pi q |\bar{\mathbf{r}}_i - \bar{\mathbf{r}}_j|}. \quad (10)$$

For such a system the high  $q$  limit of the intensity averaged over a fixed shell is given by

$$\lim_{q \rightarrow \infty} \overline{I_{inc}(q)} = I_{self}(q), \quad (11)$$

$$\lim_{q \rightarrow \infty} \overline{I_{coh}(q)} = \exp(-4\pi^2 \sigma^2 q^2) I_{self}(q). \quad (12)$$

The standard deviation for the intensities at a fixed resolution shell in the limit of high resolution is given by

$$\begin{aligned}
\lim_{q \rightarrow \infty} \text{stdev}(I_{inc}(q)) &= \lim_{q \rightarrow \infty} \text{stdev}(I_{coh}(q)) \\
&= \exp(-4\pi^2\sigma^2q^2) \sqrt{\sum_{i \neq j} |f_i(q)|^2 |f_j(q)|^2}
\end{aligned} \tag{13}$$

because as  $q$  tends to infinity the phase term in  $I_{cross}$  will be randomly distributed.

The ratio between the two is then

$$\frac{\lim_{q \rightarrow \infty} \text{stdev}(I_{inc}(q))}{\lim_{q \rightarrow \infty} I_{inc}(q)} = 0, \tag{14}$$

$$\frac{\lim_{q \rightarrow \infty} \text{stdev}(I_{coh}(q))}{\lim_{q \rightarrow \infty} I_{coh}(q)} \approx 1. \tag{15}$$

This ratio, which can be interpreted as a relative standard deviation, can be used as a measure of the amount of incoherence in a pattern.

These simple systems show that the incoherent sum of diffraction patterns, necessary to obtain a pattern with a sufficient signal to noise ratio, deviates from the coherent sum by modulating only the interference term of the diffraction intensities. This deviation increases for higher scattering vectors, which in turn will affect the resolution at which the image can be reconstructed.

### III. RECONSTRUCTION FROM TIME INTEGRATED DIFFRACTION PATTERNS

#### A. Molecular-dynamics simulations of the proteins in vacuum

For the numerical experiments we used lysozyme as it is a common protein as a model system. Simulations were performed in vacuum using GROMACS [22,23] as to reproduce the electrospray conditions expected in experiments. Two systems were simulated, one in which the protein was totally dehydrated and another one in which the protein was covered in a 3 Å layer of water molecules. The water layer stabilizes the system decreasing the fluctuations of the structure. Three simulations for each system were run and from each simulation ten snapshots were taken, 1 ns apart, for a total of 60 structures. The structures corresponding to each system were aligned by minimizing the root-mean-square distance (RMSD) between all atoms in the protein. The average RMSD after superposition was 3.7 Å for the dehydrated and 2.7 Å for the hydrated systems.

#### B. Diffraction pattern calculation and analysis

The 3D diffraction patterns from each structure were calculated on a  $100 \times 100 \times 100$  cubic grid. The distance between adjacent grid points was  $1.287 \times 10^{-2}$  Å<sup>-1</sup> and the maximum scattering vector on the corner of the grid had a resolution equal to 0.897 Å. The atomic form factors used were calculated using the five term Gaussian approximation

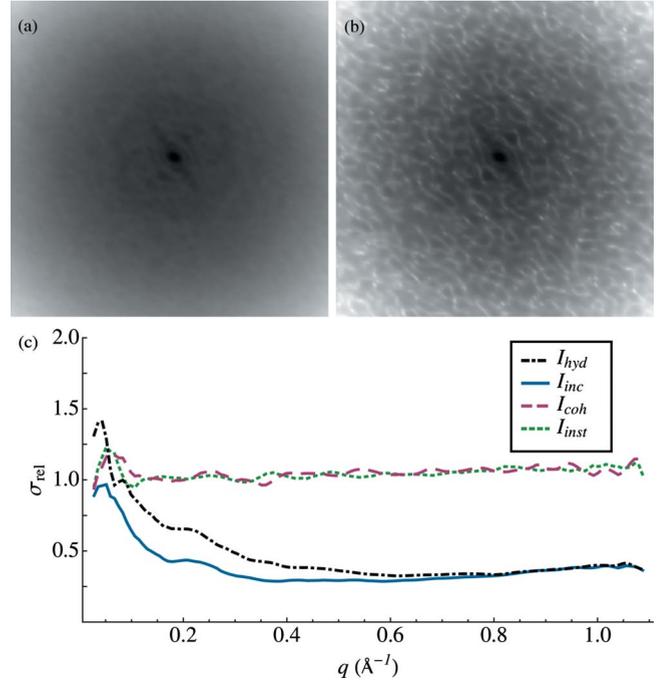


FIG. 2. (Color online) (a) Slice through the origin of the incoherently averaged diffraction intensities of the dehydrated system. (b) Slice through the origin of the diffraction intensities resulting from coherently averaging the amplitudes of the dehydrated system. (c) Relative standard deviation of the intensities ( $\sigma_{rel}$ ) as a function of the magnitude of the scattering vector ( $q$ ), for the intensities resulting from coherently averaged amplitudes ( $I_{coh}$ ), instantaneous intensities ( $I_{inst}$ ), and incoherent averaged intensities ( $I_{inc}$ ) of the dehydrated system, plus for incoherent averaged intensities of the hydrated system ( $I_{hyd}$ ).

to the atomic form factors given in the International Tables for Crystallography [24].

The patterns calculated from each system were averaged together coherently, by averaging the complex amplitudes, and incoherently, by averaging the intensities. The relative standard deviation of the intensities ( $\sigma_{rel}$ ) was then calculated at each resolution shell according to

$$\sigma_{rel}(q) = \frac{1}{\overline{I(q)}} \sqrt{\frac{1}{N_q} \sum_{i=1}^{N_q} [I(q_i) - \overline{I(q)}]^2}. \tag{16}$$

For each resolution shell  $q=|q|$ ,  $N_q$  represents the number of pixels and  $\overline{I(q)}$  the average intensity.

We compared four diffraction patterns. The incoherently ( $I_{inc}$ ) and coherently ( $I_{coh}$ ) averaged patterns from the dehydrated system, a diffraction pattern calculated from the first snapshot from the dehydrated system which we refer to as instantaneous diffraction pattern ( $I_{inst}$ ), and finally the incoherently averaged patterns from the hydrated system ( $I_{hyd}$ ).

As can be seen in Fig. 2 the relative standard deviation drops for larger scattering vectors in the case of the incoherently added diffraction patterns as it was expected. This drop reflects the fact that the interference part of the diffraction pattern is washed away by the movements of the protein leaving only the average atomic structure factor which is

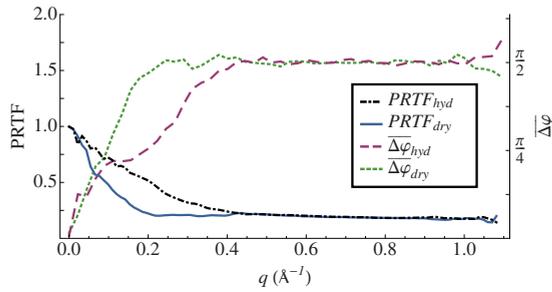


FIG. 3. (Color online) PRTF and mean phase error in radians of the averaged reconstructed image. The PRTF drops below  $1/e$  at  $q=0.132 \text{ \AA}^{-1}$  and  $q=0.271 \text{ \AA}^{-1}$  resulting in a resolution of  $7.6 \text{ \AA}$  and  $3.7 \text{ \AA}$  for the dehydrated and hydrated systems, respectively. The PRTF and the mean phase flatten out at the same resolution supporting the use of the PRTF as a good indicator of the information content of the reconstruction.

constant for a given resolution shell [21]. There is no appreciable drop in either the instantaneous or the coherently added patterns. The drop of  $\sigma_{rel}$  in the dehydrated system is larger due to the greater amount of motion in that system. These results are consistent with the analysis of the vibrating atoms system, where the drop in the variance of the signal is proportional to the atomic motion.

### C. Phasing of the calculated patterns

The incoherently added patterns were phased using the shrinkwrap algorithm [25] combined with Hybrid Input Output [26] using the Hawk image reconstruction package [27]. The reconstructed object was constrained to be real and to have a support which corresponds to 5% of the imaged volume. The volume constraint was based on an estimate of the volume of the protein. The support was updated every 20 iterations and the new support was calculated based on the average of the previous 20 iterations. 24 independent reconstructions were performed for each system and the recovered images were superimposed and averaged to obtain the averaged reconstructed image ( $\rho_{rec}$ ). For comparison we used the average of the electron densities of all the snapshots for each system,  $\rho_{avg}$ , as the average protein structure.

### D. Analysis of the reconstructed image

The phase retrieval transfer function (PRTF) [7] shows  $7.6 \text{ \AA}$  and  $3.7 \text{ \AA}$  resolution for the dehydrated and hydrated systems, respectively. This is significantly worse than the RMSD of the structural ensembles used ( $3.7 \text{ \AA}$  and  $2.7 \text{ \AA}$ ) as can be seen in Fig. 3.

The resolution-dependent mean phase error ( $\overline{\Delta\varphi(q)}$ ) was calculated according to

$$\overline{\Delta\varphi(q)} = \frac{1}{N_q} \sum_{i=1}^{N_q} [\varphi_{avg}(\mathbf{q}_i) - \varphi_{rec}(\mathbf{q}_i)], \quad (17)$$

where  $\varphi_{avg}$  and  $\varphi_{rec}$  represent the phases of  $\rho_{avg}$  and  $\rho_{rec}$ , respectively.

The  $\Delta\varphi$  between the average of the 24 recovered images and  $\rho_{avg}$  is in good agreement with the PRTF, both flattening

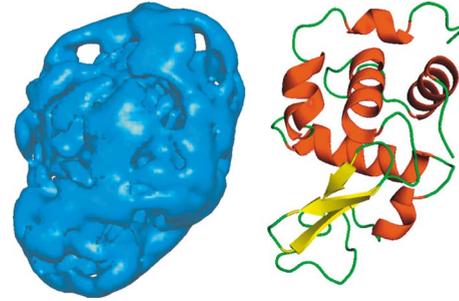


FIG. 4. (Color online) Left: isosurface of the reconstructed electron density map of  $\rho_{rec}$  for hydrated lysozyme assuming perfect image orientation, perfect signal, and no radiation damage to the sample. The map was plotted at  $3\sigma$ . Right: cartoon representation of the secondary structure of 1AKI [28], the PDB entry used in the molecular-dynamics simulations. The isosurface on the left follows the PDB structure, but details are smoothed out as a result of the lower resolution.

out at a resolution of approximately  $5 \text{ \AA}$  and at  $2.5 \text{ \AA}$  for the dehydrated and hydrated systems. These resolutions also correlate well with the resolution at which  $\sigma_{rel}$  stabilizes.

Finally the reconstructed maps match closely  $\rho_{avg}$  at low resolution as seen in Fig. 4, but there are obvious differences at medium and high resolutions as it would be expected from the PRTF. The reconstructed map from the hydrated system is of higher quality compared to the dry system due to the reduced structure fluctuations.

The decreased resolution of the reconstructed image as compared to the RMSD of the used structures can be understood by the fact that there is no real space equivalent to the incoherent averaging of diffraction patterns unlike for coherent averaging. In this sense the incoherent summation introduces artifacts in the images, which increases with increasing resolution due to greater relative structural differences at higher resolution. This noise in turn leads to a degraded reconstructed image which will be blurred by more than simply the motion of the sample.

## IV. SEPARATION OF CONFORMATIONAL SUBSTATES

A full three-dimensional reconstruction requires reproducible samples exposed to the beam one-by-one, and in different orientations. How reproducible is a “reproducible object” and how well can we distinguish between similar and dissimilar structures will affect the resolution in the reconstructed image. Macromolecules fluctuate between distinct conformers [29] and each of these may be present in a number of conformational substates [30].

Possibilities exist to identify and separate distinct structural states from a redundant set of diffraction patterns. We briefly survey two such possibilities here: the use of common arcs of intersection and manifold mapping to separate like from unlike structures.

### A. Application of common arcs of intersection to separate structural classes

Two different projections of the same 3D object have a common line of intersection. In diffraction space, the inter-

section gives an arc, which passes through the origin [8]. The common arc of intersection can be found by an appropriate subtraction of two diffraction patterns provided we have the same object exposed to the beam in the two orientations. If there is no common object then there is no common line of intersection. If the signal is strong enough for the line of intersection to be found, it will then be possible to establish similarities and determine the relative orientation of these images. When identical objects are illuminated, the arc of intersection extends to the highest resolution. In the case of grossly dissimilar objects the arc contracts to a point in the origin. When different conformational states of the same macromolecule are considered, we expect the common arc of intersection to fade at higher resolutions. A pairwise correlation analysis could be used to separate major conformational states. Components can be separated and structural heterogeneity can be reduced in a redundant data set. This procedure can be used to separate different structures from each other in a data set (e.g., from a mixture of different proteins) but requires diffraction patterns with high signal-to-noise ratios at high resolutions. We note that signal-to-noise ratios can be enhanced by invoking classification algorithms and averaging (e.g., [8,31]), in which case the problem is shifted to telling if two weak patterns come from the same structure or from two different structures, as discussed next.

### B. Separation of different structures or structural states by manifold mapping

An innovative approach to the problem of orienting very weak diffraction patterns has recently been published [9]. The key for this method is the recognition that the molecule has only three degrees of rotational freedom (it resides in a 3D space) and so its diffraction patterns must occupy a three-dimensional subset of the set of all possible intensity measurements. The tip of the image vector in the multidimensional intensity space is therefore confined to a 3D manifold. Diffraction patterns from different structures fall onto differ-

ent three-dimensional manifolds and can in principle be distinguished from each other [32]. This is an extremely powerful but computationally expensive method that may work even with extremely weak diffraction patterns.

## V. CONCLUSIONS

This paper evidences the need to achieve a low variation of the samples in single-particle 3D diffraction imaging. While for perfectly reproducible samples the maximum resolution obtainable from diffraction image is only limited by radiation damage, which can be minimized by going to shorter pulse lengths, in the case of proteins their intrinsic dynamics might be the limiting factor. To overcome this limitation, ways to reduce variability in the sample structure have to be explored. For example, the use of a small hydration layer around the proteins would both decrease the amount of radiation damage [20] and the range of motion of the protein samples [23]. Cooling can be used to reduce vibrational amplitudes. Another possibility is to classify the diffraction patterns according to their conformations using the fact that the pattern from each conformation lives on a different manifold [9] or the fact that the extent of the common arc between patterns from different conformations is limited by the similarity between the conformations. The smaller RMSD on each class, as compared to the full ensemble, leads to potentially higher resolution structures for each class.

## ACKNOWLEDGMENTS

We are grateful to A. Patriksson for sharing the simulations and C. Caleman for inspiring discussions. This work was supported by the Swedish Research Council through the Centre of Excellence in FEL-Studies at Uppsala University. Additional support came from the DFG Cluster of Excellence at the Munich Centre for Advanced Photonics and from the Virtual Institute Program of the Helmholtz Society.

- 
- [1] R. Neutze, R. Wouts, D. van der Spoel, E. Weckert, and J. Hajdu, *Nature (London)* **406**, 752 (2000).  
 [2] J. Hajdu, *Curr. Opin. Struct. Biol.* **10**, 569 (2000).  
 [3] P. Persson, S. Lunell, A. Szöke, B. Ziaja, and J. Hajdu, *Protein Sci.* **10**, 2480 (2001).  
 [4] M. Bergh, N. Timneanu, and D. van der Spoel, *Phys. Rev. E* **70**, 051904 (2004).  
 [5] N. Timneanu, C. Caleman, J. Hajdu, and D. van der Spoel, *Chem. Phys.* **299**, 277 (2004).  
 [6] B. Ziaja, R. A. London, and J. Hajdu, *J. Appl. Phys.* **97**, 064905 (2005).  
 [7] H. N. Chapman *et al.*, *Nat. Phys.* **2**, 839 (2006).  
 [8] G. Hultdt, A. Szöke, and J. Hajdu, *J. Struct. Biol.* **144**, 219 (2003).  
 [9] R. Fung, V. Shneerson, D. K. Saldin, and A. Ourmazd, *Nat. Phys.* **5**, 64 (2009).  
 [10] V. L. Shneerson, A. Ourmazd, and D. K. Saldin, *Acta Crystallogr. A* **64**, 303 (2008).  
 [11] Ne-Te Duane Loh and V. Elser, *Phys. Rev. E* **80**, 026705 (2009).  
 [12] J. Miao, P. Charalambous, J. Kirz, and D. Sayre, *Nature (London)* **400**, 342 (1999).  
 [13] H. Frauenfelder, S. G. Sligar, and P. G. Wolynes, *Science* **254**, 1598 (1991).  
 [14] I. Waller, *Z. Phys. A: Hadrons Nucl.* **17**, 398 (1923).  
 [15] E. J. Levin, D. A. Kondrashov, G. E. Wesenberg, and G. N. Phillips, Jr., *Structure* **15**, 1040 (2007).  
 [16] Z. Jurek, G. Faigel, and M. Tegze, *Eur. Phys. J. D* **29**, 127 (2004).  
 [17] S. P. Hau-Riege, R. A. London, and A. Szöke, *Phys. Rev. E* **69**, 051906 (2004).  
 [18] G. Bortel and G. Faigel, *J. Struct. Biol.* **158**, 10 (2007).  
 [19] G. Bortel, G. Faigel, and M. Tegze, *J. Struct. Biol.* **166**, 226 (2009).

- [20] S. P. Hau-Riege, R. A. London, H. N. Chapman, A. Szoke, and N. Timneanu, *Phys. Rev. Lett.* **98**, 198302 (2007).
- [21] M. Bergh, G. Huldt, N. Timneanu, F. R. N. C. Maia, and J. Hajdu, *Q. Rev. Biophys.* **41**, 181 (2008).
- [22] E. Lindahl, B. Hess, and D. van der Spoel, *J. Mol. Model.* **7**, 306 (2001).
- [23] A. Patriksson, E. Marklund, and D. Vanderspoel, *Biochemistry* **46**, 933 (2007).
- [24] *International Tables for Crystallography*, edited by A. J. C. Wilson (Kluwer, Dordrecht, 1995), Vol. C, p. 500.
- [25] S. Marchesini, H. He, H. N. Chapman, S. P. Hau-Riege, A. Noy, M. R. Howells, U. Weierstall, and J. C. H. Spence, *Phys. Rev. B* **68**, 140101(R) (2003).
- [26] J. R. Fienup, *Appl. Opt.* **21**, 2758 (1982).
- [27] F. R. N. C. Maia and T. Ekeberg, Hawk imaging reconstruction package, <http://xray.bmc.uu.se/hawk/> (2009).
- [28] P. J. Artymiuk, C. C. F. Blake, D. W. Rice, and K. S. Wilson, *Acta Crystallogr. B* **38**, 778 (1982).
- [29] K. U. Linderstrøm-Lang and J. A. Shellman, in *The Enzymes*, edited by P. D. Boyer (Academic Press, New York, 1959), Vol. 1, pp. 443–510.
- [30] H. Frauenfelder, F. Parak, and R. D. Young, *Annu. Rev. Biophys. Biophys. Chem.* **17**, 451 (1988).
- [31] M. van Heel *et al.*, *Q. Rev. Biophys.* **33**, 307 (2000).
- [32] A. Ourmazd, Structure by fleeting illumination of faint objects, <http://www.xfel.eu/documents/spb-workshop-2008-ourmazd.pdf> (2008).