Journal of Structural Biology 194 (2016) 325–336

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Hybrid approach for structural modeling of biological systems from X-ray free electron laser diffraction patterns

Atsushi Tokuhisa [a], Slavica Jonic [b], Florence Tama [a,c], Osamu Miyashita [a,*]

[a] RIKEN Advanced Institute for Computational Science, 6-7-1, Minatojima-minami-machi, Chuo-ku, Kobe, Hyogo 650-0047, Japan
[b] IMPMC, Sorbonne Universités – CNRS UMR 7590, UPMC Univ Paris 6, MNHN, IRD UMR 206, Paris 75005, France
[c] Department of Physics, Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8602, Japan

ARTICLE INFO

Article history:
Received 31 December 2015
Received in revised form 9 March 2016
Accepted 9 March 2016
Available online 10 March 2016

Keywords:
XFEL
Coherent diffraction
Diffraction pattern
Protein structure
Computational modeling
Experimental conditions
Conformational transitions

ABSTRACT

We present a new hybrid approach for structural modeling using X-ray free electron laser (XFEL) diffraction patterns from non-crystalline biological samples. Reconstruction of a 3D structure requires a large number of diffraction patterns; however, in the current XFEL experiments with biological systems, the analysis often relies on a small number of 2D diffraction patterns. In this study, we explore the strategies to identify plausible 3D structural models by combining the 2D analysis of such diffraction patterns with computational modeling (normal mode analysis or molecular dynamics simulations). As the first step toward such hybrid modeling, we established a protocol to assess the agreement between the model structure and the target XFEL diffraction pattern and showed that XFEL data can be used to study the conformational transitions of biological molecules. We tested the proposed algorithms using data of three biomolecular complexes of different sizes (elongation factor 2, CCM virus, and ribosome) and examined the experimental conditions that are required to perform such studies, in particular the XFEL beam intensity requirements. The results indicate that the current beam intensity is close to a strength that enables us to study conformational transitions of macromolecules, such as ribosomes. The proposed algorithm can be combined with molecular mechanics approaches, such as molecular dynamics simulations and normal mode analysis, to generate a large number of candidate structures to perform hybrid structural modeling.

1. Introduction

Three-dimensional (3D) structures of biological macromolecules at atomic resolution provide valuable information for understanding their activities and mechanisms, potentially leading to pharmacological applications. One of the most successful methods to determine the 3D structure at atomic level of resolution is X-ray crystallography. However, X-ray crystallography cannot be used for systems that cannot be crystallized. In order to obtain structural information from non-crystalline samples, various alternative methods have been utilized, such as electron microscopy (EM), small-angle X-ray scattering (SAXS) and fluorescence resonance energy transfer (FRET) (van den Bedem and Fraser, 2015). These methods have been used to elucidate macromolecular structure and dynamics that could not be deciphered by X-ray crystallography, although not at an atomic level of detail.

Recently, X-ray free electron laser (XFEL) (Emma et al., 2010; Ishikawa et al., 2012) emerged and started to be used as fourth-generation X-ray sources that provide a femto-second X-ray pulse with the peak brilliance of at least one billion times higher than the one employed at synchrotrons (Waldrop, 2014). These new X-ray sources offer possibilities of providing new biological structural information. The observation techniques using XFEL is based on the concept of 'probe before destruction' scheme (Gaffney and Chapman, 2007; Neutze et al., 2000; Solem, 1986).

The current experimental applications of XFEL on biological systems may be classified into two categories: observation of the systems in crystalline and non-crystalline states. In micro- and nano-crystallography, due to the high intensity of XFEL beam, small crystals that are inadequate for conventional X-ray crystallography can be used for structure determination. In addition, XFEL creates very short pulses, and thus the radiation damage could be avoided. Utilizing these properties, high-resolution 3D structures of novel proteins have been determined (Barends et al., 2014; Boutet et al., 2012; Chapman et al., 2011; Liu et al., 2013). In the single particle coherent diffraction imaging (CDI), randomly oriented non-crystalline samples are injected into or placed in front of the X-ray laser beam to obtain two-dimensional (2D) coherent diffraction patterns. These diffraction patterns are approximately the squares of the amplitudes of Fourier transforms of 2D real-space projections of the 3D object, but do not contain phase information.

* Corresponding author.
E-mail address: osamu.miyashita@riken.jp (O. Miyashita).

http://dx.doi.org/10.1016/j.jsb.2016.03.009
1047-8477/

Although the 2D phase information cannot be measured, it could be determined computationally from an oversampled 2D diffraction pattern with a sufficient scattering intensity. The most widely used algorithm for the phase retrieval is hybrid input output (HIO) algorithm (Fienup, 1982; Gerchberg and Saxton, 1972; Miao et al., 1998; Rodriguez et al., 2013; Sayre, 1952). This single-shot "lensless" 2D imaging has been used to elucidate internal structure of large biological objects, such as cell and virus (Gallagher-Jones et al., 2014; Hantke et al., 2014; Kimura et al., 2014; Seibert et al., 2011; van der Schot et al., 2015).

CDI can be extended to 3D structure determinations. A single diffraction pattern is a part of the Ewald sphere that provides fragmentary information of the 3D structure factor of the macromolecular complex. The 3D structure determination requires many 2D diffraction patterns in different orientations as well as the estimation of these orientations from the diffraction patterns, since, generally, the molecular orientations cannot be determined experimentally (Kassemeyer et al., 2012). Many theoretical studies (Hosseinizadeh et al., 2014; Huldt et al., 2003; Loh and Elser, 2009; Tegze and Bortel, 2013, 2012) show the possibility of the 3D structure determination at atomic resolution using noisy data resulting from single particle diffraction patterns. Alternatively, the use of diffraction data from ice plates that embed a large number of particles has also been proposed (Kodama and Nakasako, 2011; Oroguchi and Nakasako, 2013). Several 3D structure reconstructions were performed using experimental data of non-biological samples with strong scattering power resulting in diffraction patterns with a higher signal-to-noise ratios (Chapman et al., 2006; Kassemeyer et al., 2013; Loh et al., 2010). In addition, the 3D structure of a gold nano-crystal was reconstructed at ~5.5 nm structural resolution from one single-shot coherent diffraction pattern by taking advantage of the symmetry of the nano-crystal and the curvature of the Ewald sphere (Xu et al., 2014).

In contrast, the number of CDI experiments with biological samples aiming at 3D structure determination is still small. The difficulty of the CDI experiments with biological macromolecular complexes is their poor scattering power implying that their diffraction intensity is weak, even with intense XFEL beams. Thus, in a large scattering region signal to noise ratio becomes very low, which makes the structural reconstruction processes difficult. Current CDI applications target large biological samples at the micrometer scale and the reconstruction resolutions are still low. Recently, 3D structure determination of the Giant Mimivirus was successfully performed at ~125 nm structural resolution (Ekeberg et al., 2015). In this pioneering work with biological samples, it has been proposed to analyze only the diffraction patterns with a high signal-to-noise ratio in order to reduce the data analysis complexity.

The quantum noise has an influence on the convergence of the solution in the phase retrieval process. When HIO algorithms are applied to noisy diffraction patterns, it may result in, non-deterministic, various reconstructed real images (Sekiguchi et al., 2016). Noise may be reduced by averaging patterns with similar orientations; however, the average diffraction patterns result in lowering the reconstruction resolution. To overcome the problem, several phase retrieval algorithms robust to noise (Ikeda and Kono, 2012; Martin et al., 2012; Rodriguez et al., 2013) as well as statistical evaluation of image set from multiple reconstruction trials have been presented (Sekiguchi et al., 2016). An approach to enhance signals from biological samples by interference with a strong signal from gold colloid particles randomly dispersed around the biological sample was proposed (Takayama et al., 2015). However, this approach may be limited to 2D structural determination because it is necessary to prepare multiple samples with the same colloid particle positions for 3D structural determination.

The weakness of biomolecular diffraction intensity could be compensated using a very large number of diffraction data (Huldt et al., 2003), however, which is currently unattainable. Using an X-ray pulse focused to a spot of ~5 μm in diameter and ~115 nm size particles, it was recently shown that the overall hit rate can achieve 79% while the single hit rate can be only 33% (Hantke et al., 2014). To obtain high-resolution experimental data, the X-ray beam should be focused to a spot as small as 50 nm in diameter (Mimura et al., 2014). However, this extreme X-ray beam focusing unfortunately decreases the hit rate. Recent development of new concentric-flow electrokinetic injector can improve the hit rate (Sierra et al., 2016), however, yet to be applied to single particle analyses. In addition, inherent conformational diversity of biological samples, particularly for larger systems, hinders the high-resolution 3D structural reconstruction process.

Thus, XFEL data are currently limited in resolution and typically to 2D instead of 3D, and therefore, there is not sufficient information to determine the high-resolution structure solely from the diffraction patterns. In this regard, the current XFEL data are similar to other "low-resolution" data such as cryo-EM and SAXS. For these data, hybrid approaches, which combine the computational molecular modeling with experimental data, play an increasingly important role in elucidating structural information at atomic level resolution. In particular, they have been used to reveal the details of important conformational dynamics of several biological molecules (Alber et al., 2008; Fritz et al., 2013; Gorba et al., 2008; Jin et al., 2014; Russel et al., 2012; Tama et al., 2004; van den Bedem and Fraser, 2015).

In this paper, we propose a new hybrid approach that can be applied to XFEL diffraction patterns from non-crystalline biological samples. Using computer modeling, the method generates a large number of model structures and identifies the model structure that is in the best agreement with the experimental data. Such identification of "plausible" structures overcomes the problem of 3D reconstruction from current small-size diffraction data sets.

We also address the question of the XFEL data quality that is required to distinguish between different conformations of biological molecules. Diffraction patterns from biological molecules are under strong influence of noise and the orientation of the target molecule is unknown. We examined how the experimental conditions, in particular beam intensity, affect our ability to differentiate between the different conformations from the data. In this study, the proposed methodology was tested using three biomolecular complexes of different sizes, elongation factor 2 (EF2), CCM virus (CCMV) and ribosome. We selected two conformations of each system and examined the XFEL experimental conditions that are required to distinguish between the conformations using the diffraction patterns. For EF2, we also created intermediate structures between two known conformations to further examine the limit of structural differences that can be differentiated with XFEL data. Our results show that the proposed methodology can be used to identify the candidate structural model. In addition, the required XFEL beam intensity to study each system is discussed in detail. The proposed matching scheme could be combined with molecular mechanics simulation techniques, such as molecular dynamics simulations and normal mode analysis, to construct candidate structures and achieve model reconstructions from XFEL diffraction patterns.

2. Methods

2.1. Simulated diffraction patterns and similarity measurement

One of the goals of this study is to examine the necessary conditions in which the XFEL diffraction patterns can be used to study conformational dynamics of biological molecules. We performed tests using three protein molecules/complexes with multiple known conformations, and examined the conditions in which the diffraction patterns can be used to differentiate between these