**Characterizing functionally important conformational dynamics of a radiation-sensitive enzyme using time-resolved, substrate-mixing crystallography and continuous scattering at room temperature.**

Henry van den Bedem (SLAC; spokesperson), Mark Wilson (UNL), Mike Wall (LANL), Nicholas Sauter (LBNL)

## 1. Experimental Team

| Name | Institution | contact | role | onsite |
|---|---|---|---|---|
| Henry van den Bedem | SLAC/BSD | vdbedem@slac.stanford.edu | PI. experimental design, sample prep, data processing, analysis & collection | Y |
| Mark Wilson | UNL | mwilson13@unl.edu | PI. experimental design, sample prep, data processing, analysis & collection | Y |
| Mike Wall | LANL | mewall@lanl.gov | PI. experimental design, sample prep, data processing, analysis & collection | Y |
| Michael Soltis | SLAC/SSRL | soltis@slac.stanford.edu | PI. experimental design, sample prep, data processing, analysis & collection | Y |
| Aina Cohen | SLAC/SSRL | aina@slac.stanford.edu | PI. experimental design, sample prep, sample delivery | Y |
| Nick Sauter | LBNL | mksauter@lbl.gov | PI. data processing | Y |
| Ray Sierra | SLAC/LCLS | rsierra@slac.stanford.edu | experimental design, sample prep, sample delivery, data collection | Y |
| Aaron Brewster | LBNL | ASBrewster@lbl.gov | data processing and analysis | Y |
| Chuck Yoon | SLAC/LCLS | yoon82@stanford.edu | data processing and analysis | Y |

## 2. Scientific Case

*XFELs hold unique promise for the characterization of macromolecular structure far-from-equilibrium*

Protein motion is critical for the remarkable functional diversity of macromolecules. However, a detailed understanding of how protein conformational dynamics couples to function has been limited by experimental barriers that prevent the direct observation of protein motion under physiologically relevant conditions. In addition to three dimensional structural information, X-ray crystallography provides detailed, but often disregarded, information about the spatial properties of conformational heterogeneity for every residue in the protein. Considerable effort spanning several decades has been expended to extract and analyze information about protein mobility from X-ray crystallographic data, including joint work by the collaborators on this proposal (Wall *et al.*, 1997; Wilson & Brunger, 2000; van den Bedem *et al.*, 2013; Keedy *et al.*, 2014; Lin *et al.*, 2016; van Benschoten, A.H. *et al.*, 2016). Critically lacking in these analyses is characterization of the time domain, which is an inherent limitation of conventional monochromatic X-ray crystallographic experiments (van den Bedem & Fraser, 2015). To observe synchronized changes in the molecules in a microcrystal, the system must be perturbed and brought into an non-equilbirium state, from which its relaxation can be observed (Schmidt & Saldin, 2014). Importantly, the introduction of ligands or substrates perturb equilibrium. In combination with the unique characteristics of X-ray free electron lasers (XFELs), this perturbation enables the mapping of protein conformational changes to a temporal coordinate at room temperature. Therefore, XFELs can interrogate the time dimension of functionally important protein dynamical processes, opening the field of diffraction-based studies of proteins under non-equilibrium conditions. Unlike single-crystal Laue time-resolved crystallography, XFEL experiments are not constrained by the challenging requirement for synchronization over the full volume of a macroscopic crystal or the use of photo-activated processes.

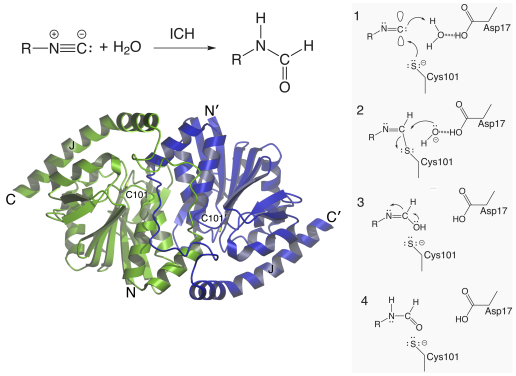


**Fig 1 ICH reaction, structure, and mechanism.** Top left shows the reaction catalyzed by ICH. Bottom left is the dimeric crystal structure, with the catalytic C101. Proposed mechanism on the right.

*Conformational dynamics of isocyanide hydratase (ICH) are important for enzyme turnover*

We propose to use XFEL radiation to observe the role of correlated motion in the full catalytic cycle of isocyanide hydratase (ICH, Lakshminarasimhan *et al.*, 2010). ICH hydrates diverse isocyanides to yield N-formamide and is only one of two enzymes characterized to date that degrades organic isocyanides. ICH was isolated from *P.*

*fluorescens*, which is an abundant and critical bacterial component of the healthy soil microbiome. Pseudomonads protect plant roots by cyanogenesis, and play a key role in creating disease-suppressive 'super-soils', increasing crop resilience against climate change-induced fungi and diseases. Therefore, understanding the role of ICH in improving the fitness of Pseudomonads in competitive soil microbiotic niches would advance the understanding of molecular determinants of soil microbiome health and species composition, with important implications for predicting or (bio-)engineering responses to more extreme environments owing to climate change, or for extra-planetary soils.

ICH is a 230-residue homodimeric enzyme (Fig. 1) that we have previously characterized using both cryogenic (100K) and ambient (277K) conventional synchrotron radiation (SR) data to 1.0 Å resolution (Lakshminarasimhan *et al.*, 2010 and unpublished). Our prior work identified a collective ~2 Å shift in an α-helix proximal to the active site which is prominent in the 277K SR data and present but less evident in the 100K dataset (Fig. 2). This helical shift alters the microenvironment of the catalytic nucleophile, with important consequences for catalysis. This concerted displacement is facilitated by the strained backbone torsion angles of Ile152 and a highly conserved G150-G151 diglycine motif. A liquid-like motions (LLM) model of continuous data we obtained from the 277K diffraction images yields a strong correlation of 0.64 to 1.4 Å resolution (Fig 3). The correlation is exquisitely sensitive to the details of the atomic displacements, indicating that continuous scattering can be used to evaluate alternative models of ICH conformational variation.

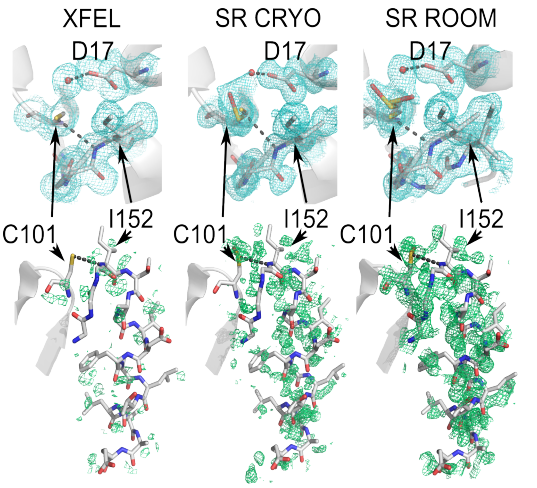


**Fig 2 Radiation dose-response perturbation of C101 charge state shifts conformational ensemble.** Top: 2Fo-Fc density at 0.3σ shows increasingly populated alternate active site substates around I152. Bottom, Fo-Fc difference maps reveal that functional motions propagate over long, 25Å distances.

We have conducted extensive pre-steady state kinetic characterization of ICH with the chromogenic substrate para-nitrophenyl isocyanide (p-NPIC) and found that the enzyme displays burst kinetics with ~60% of the protein being apparently inactive. Importantly, the pre-steady state rate constant is ~18 $s^{-1}$ (Fig. 4A, B), meaning that early steps in catalysis occur on a timescale that requires XFEL approaches to observe. A G150T mutation locks the helix into the shifted conformation (verified by X-ray crystallography), increasing the active fraction ~2-fold, but substantially decreases the pre-steady state rate constant, demonstrating that the dynamic displacement of this helix affects ICH catalytic activity (Fig. 4).

We collected a complete XFEL data set to 1.6 Å resolution at the LCLS on 50 μm apo ICH crystals at room temperature using the microfluidic electrokinetic sample holder (MESH) injector under helium (Sierra et al., 2015). Surprisingly, despite the excellent quality of the maps, we found no evidence of the helical shift. The principal difference between the 277K SR and XFEL data is the oxidation state of the active site nucleophile Cys101 (Fig 2). A cysteine-sulfenate (Cys-SOH) is apparent in the 277K SR data, which closely resembles the proposed covalent thioimidate intermediate that forms at Cys101 in the catalytic cycle of ICH. Because the XFEL data suffered no radiation-induced oxidation of Cys101, this modification and the associated helical displacement were not observed. Radiation dose-induced oxidation of the Cys-SOH thiolate neutralizes the charge of the Sγ, thereby making it a poorer hydrogen bond acceptor. The weakened hydrogen bond between the amide H of Ile152 and Cys101 Sγ enables the helical shift, relieving backbone strain around Ile152 (Fig 2). Long (> 1μs) molecular dynamics simulations of ICH crystals further confirmed this hypothesis (Fig 4). The formation of an obligatory enzymatic intermediate thioimidate at Cys101 will likely have the same effect, triggering a similar conformational change required for efficient ICH catalysis. Therefore, we propose that the collective motions observed in 277K SR data are a snapshot of the conformational ensemble of intermediate states populated during enzyme turnover. However, several important questions remain unanswered: 1) What is the role of these motions, which are propagated over 25Å from the

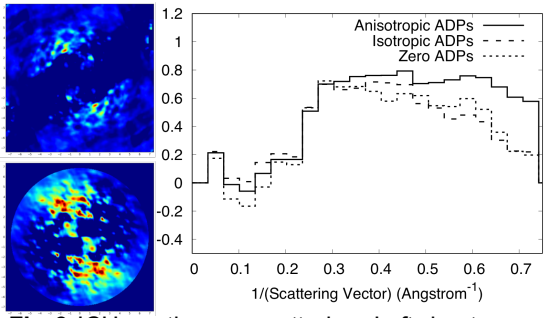


**Fig 3** ICH continuous scattering. Left: heatmap of 3D experimental continuous intensity (top) and refined LLM model (bottom) Right: Fourier shell correlation between ICH continuous intensity and an LLM model.

active site? Are they essential for the chemical step of catalysis, or do they facilitate product release? 2) What is the role of other dynamic responses throughout the protein, and what are their timescales and mechanisms? Analysis of our apo XFEL data indicates sharply reduced continuous scattering. A mix-and-inject experiment would trigger internal motions, enabling a direct comparison of the effect of internal motions on continuous intensities, and help validate atomic displacement correlations thoughout the ensemble. Our proposed MFX experiment can map the full ensemble of transient conformational substates to reaction intermediates. Therefore, ICH provides a powerful model system for interrogating the connection between multi-lengthscale dynamics and catalysis in enzymes with high resolution Bragg data and continuous scattering. This would allow a complete structural characterization of an enzyme during catalysis that could be mapped to a previously measured pre-steady state kinetic profile.

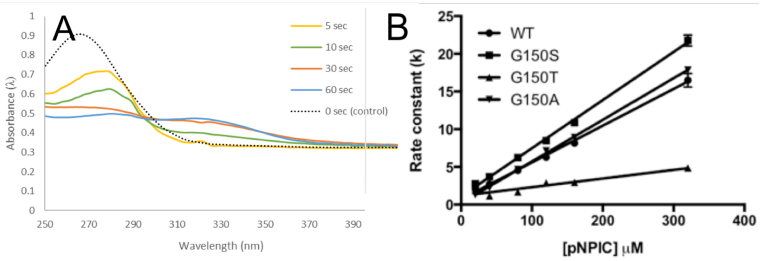


**Fig 4**: A) Absorbance spectra of catalytically active ICH crystals in 1mM *p-napthyl isocyanide* at time points 0-60 secs. 270 nm indicates substrate concentration, 315 nm indicates product B) Pre-steady state rate constants versus concentration for wild-type ICH and mutants G150S, G150T, and G150A in solution.

**XFEL data can map ICH enzyme conformational dynamics to the reaction coordinate.** ICH is highly amenable to characterization by XFEL radiation; our high-resolution and complete data set (1.6Å) from apo enzyme was collected at XPP in 45 minutes from crushed, 20-50μm crystals. Moreover, the subtrate p-NPIC is a small (148 g/mol), distinctively-shaped molecule whose thioimidate intermediate and formamide products are both clearly structurally distinct from the substrate. ICH is ideally suited for a time-resolved substrate-mixing crystallography experiment. ICH acting on p-NPIC has a low $K_m$ of 7.2 +/- 1.0 μM and a relatively low steady-state turnover number of 0.3 $s^{-1}$ at room temperature in solution. The pre-steady state rate constant is ~18-20 $s^{-1}$, during which period the proposed intermediate is likely to form. We have verified catalysis in the crystal (Fig. 4B), and preliminary observations suggest the catalytic rate in the crystal slows by approximately an order of magnitude (turnover number ~0.03 $s^{-1}$, approximately two turnovers a minute). We are currently measuring both steady state and pre-steady state rate constants for pulverized and capillary-mounted crystals that will help us precisely tailor the experiment to the time regime of interest. We further calculated mixing and diffusion time-scales to be on ~10-20 ms time-scale for the small solution volume containing ICH micro-crystals. Therefore, characterizing the key steps in ICH catalysis using p-NPIC as a substrate with SFX is feasible.

## 3. Experimental Procedure

Our time-resolved study will collect a series of ICH data sets, in which the XFEL pulses probe the crystals after the introduction of substrate at varying time intervals. We will use the microfluidic electrokinetic sample holder (coMESH) injector to deliver crystals to the XFEL beam under atmospheric pressure, and at room temperature. Our crystals will be prepared at high concentration (~$10^8$ $ml^{-1}$), and will be diluted if required to reduce multiple hits. We propose to collect four data sets: A 'dark' data set of native (apo) ICH, and a t=0 data set corresponding to the total mixing and diffusion time. Additionally, we will collect data sets at t=10 ms and t=100 ms. We expect that the 10 ms time point will capture the catalytic intermediate, and will allow us to determine how enzyme motion is coupled to its formation and eventual turnover. The 100 ms time point will represent the final product-bound complex and will show the enzyme is a post-catalysis state but prior to product release. We will cool the samples to slow the reaction if required to access these time-points.

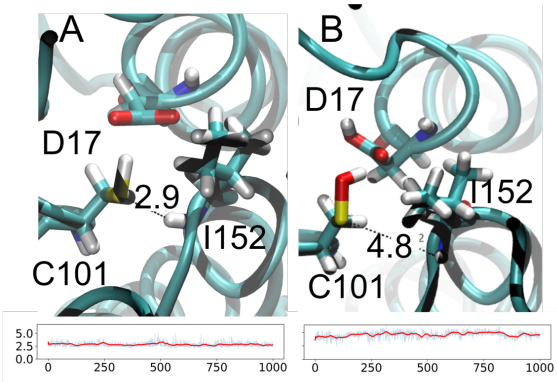


**Fig 5 Simulations of ICH** with C101 (A) and C101_SOH (B) indicate that a C101-SOH weakened hydrogen bond is disrupted early in the trajectory (B lower panel), in contrast to C101 (A lower panel).

*Crystals*

The experiment requires about 1 ml of a crystal slurry sample. For our preliminary collection of ICH XFEL data, we crushed large ICH crystals by vortexing with a 0.5 mm steel ball, resulting in ~50 μm fragments that delivered excellent data. ICH is easily expressed and purified in conditions that have been well-established by the Wilson group, and have already been produced in mL quantities. We have grown large numbers of these crystals at SLAC/SSRL and

crush them to monodisperse sizes suitable for XFEL data collection. We have also grown microcrystals (~50 µm) directly from extensively pulverized crystals as seed stock.

*Data Processing and Analysis*

We successfully processed the XFEL diffraction data from our previous experiment with cctbx.xfel (Sauter et al., 2013). Postrefinement was carried out with PRIME (Uervirojnangkoorn et al., 2015), resulting in an initial model with $R_{free}$/R of 0.27/0.21, and 95% completeness in the high-resolution shell. We will follow the same strategy for diffraction data at each time slice of our time-resolved experiment. The resulting reduced data will be used for standard structure determination with PHENIX (Adams et al., 2010).

In addition to conventional modeling of harmonic motions in diffraction data with atomic displacement parameters, we will used advanced, automated methods to analyze larger, collective motion from electron density maps (van den Bedem et al., 2009; Burnley et al., 2012). qFit multi-conformer models (van den Bedem et al., 2009) represent each residue by up to four conformers that are selected by exploring a vast number of collective interpretations of low-level electron density. qFit models can uncover conformational ensembles from X-ray data (Fraser et al., 2011), and is sensitive to altered occupancies as equilibrium of the conformational ensemble in the crystals shifts (Tenboer et al., 2014; Keedy et al. 2015). qFit models have led to insight into collective motions and molecular mechanisms in DHFR (van den Bedem et al., 2013; Keedy et al. 2014), HIV protease (Keedy, Fraser, van den Bedem, 2015) and cypA (Keedy et al., 2015). We have already successfully created multi-conformer models of ICH using our room temperature XFEL data set (unpublished, Fig. 6). Two of us (van den Bedem and Wilson) recently jointly published papers on conformational heterogeneity analyzed with qFit in $Ca^{2+}$-Calmodulin (Lin et al., 2016) and DHFR (Keedy et al., 2014). We have added qFit to our 'online' computational workflow for rapidly creating multi-conformer models including intermediates, which will augment the extensive computational and data processing skills of our team. We expect to observe substantial changes in the electron density, which will be interpreted in terms of occupancy shifts for the multi-conformer models. Difference density maps and multi-conformer models allow us to immediately adjust our experimental setup to access intermediates.

Continuous scattering. We will integrate and analyze continuous scattering data with our LUNUS software (Wall, 2009) (https://github.com/mewall/lunus). LUNUS uniquely measures and characterizes the three-dimensional reciprocal-space distribution of continuous X-ray reflections for macromolecular crystallography. LUNUS already has been adapted to process data indexed using cctbx.xfel methods (van Benschoten *et al*, 2016). We will also use our newly developed, custom algorithm to identify continuous intensity on XFEL images (unpublished). Our high quality SR continuous data and LLM model (Fig. 3) suggest that our experiments will be especially well suited for validating the application of LUNUS to SFX. Algorithm development in this direction is of high priority; MEW's work is partially supported by an ECP ExaFEL award to LCLS. The SFX experiments will enable us to overcome the radiation damage limitations of prior single-crystal synchrotron datasets. LUNUS has previously characterized motions in $Ca^{2+}$-calmodulin, staphylococcal nuclease (Wall *et al*, 1997); and cyclophilin A and trypsin (van Benschoten *et al*, 2016). To validate qFit models, we will make use of an extension of the LLM model for ensemble models (unpublished). We will map our crystal MD simulations (Fig 5) to the continuous intensities. Our strategy will result in a time-resolved series of qFit models that reflect the time-evolution of conformational heterogeneity in ICH following introduction of substrate from joint Bragg and continuous data (Fig 6).

*Finally, we wish to address two concerns raised by a previous PRP.* First, the reviewers pointed out that our proposal lacked evidence that ICH crystals are catalytically active. We have confirmed activity in micro-crystals with a UV-vis spectrophotometric assay, which monitored the evolution of the product, para-nitrophenyl formamide (p-NPF) continuously at 315 nm (Fig 4A). Therefore, ICH turns over in the crystal, addressing a key concern about feasibility. Secondly, addressing the importance of ICH, understanding of molecular determinants of soil microbiome health and species composition will enable us to engineer enzymes for extreme environments owing to climate change, or extra-planetary soils. Third, it is increasingly recognized that post-translational modifications by cysteine oxidation play a key role in controling cell-signaling and enzyme catalysis critical to human health, for example in the ICH homologue human

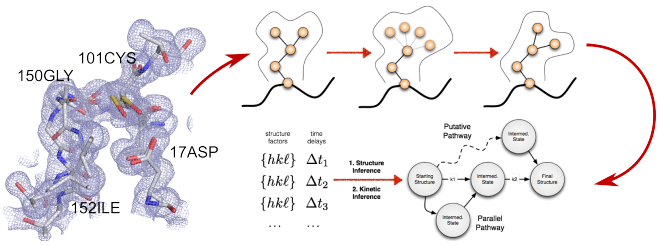


**Fig 6.** qFit multi-conformer models (left) are a snapshot of the conformational ensemble along the reaction coordinate. By tracking shifting occupancies as the reaction cycle proceeds (top right), we can identify collective motion. This results in a model of structural conformational substates which is kinetically consistent with the data (bottom right)

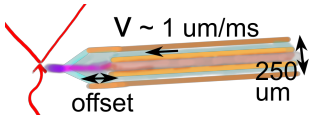

**Fig 7** coMESH nozzle with capillary offsets to control time points at interaction region.

DJ-1 and in protein tyrosine phosphatases. ICH is a unique and generalizable model system to study the dynamic response of proteins to oxidative post-translational modification—a largely unexplored but important topic. Finally, the reviewers asked if continuous scattering from the 100K data or the native XFEL data was distinct from that at 277K. We developed a new algorithm to extract continuous intensities from the XFEL data. Preliminary results indicate intensities were sharply reduced. However, we were unable to confirm that reduced intensities were related to the lack of internal motion, due to the absence of a comparison XFEL data set with triggered motions. Our current proposal will overcome that limitation.

**4. Technical Feasibility**

*Equipment*

We will use MFX standard configuration #2. The coMESH will be positioned above the XFEL beam, such that the electrospun liquid microjet flows vertically downward, perpendicular to the X-ray beam. The interaction region will be contained in helium atmosphere to reduce air scatter, improving S/N for continuous scatter.

*Delivery of para-nitrophenyl isocyanide (p-NPIC) substrate and mixing.*

p-NPIC will be mixed with protein crystals by adding buffer and substrate to the outer concentric coMESH flow. We can precisely control substrate and enzyme concentrations. By varying the capillary offsets and flow rates, we can control mixing rates. By varying the distance in between the concentric capillary exits, we can access different time points. The difference in conductivity of the two fluids creates a gradient in applied potential and drives mixing to occur on the order of microseconds. At flow rates on the order of μL/min in the outer capillary's 250 μm inner diameter, flow velocities of 1 μm/ms are achieved. Offsetting the capillary exits by 0, 10 and 100 μm, easily achieved under a microscope, we are able to get time delays of 0, 10 and 100 ms, respectively (Fig 7). Electrokinetics induces ion and surrounding fluid velocities on the order of ms at μm length scales without increasing the sample consumption, in contrast to hydrodynamic mixing time scales which are diffusion limited at these low Reynolds numbers.

We will use the Rayonix MX170-HS detector, because the large dynamic range is ideal for collecting data from well-diffracting protein crystals. The energy spectrometer will be used to record the intensity and energy fluctuations for every LCLS pulse (Zhu et al., 2012) and used to aid in data processing and scaling. Data collection will be carried out using the MFX control and DAQ systems. The standard MFX optics and monochromator system will be employed.

*Parameters*

Following the data collection parameters used for our previous ICH data collection, we propose to use the highest fundamental X-ray energy achievable using the MFX monochromator and optics system, maximum pulse energy, native energy bandwidth, and 80fs pulse width. The X-ray beam will be focused to 10μm or less. We will use a seeded FEL beam, operating at 10Hz, in order to match the readout rate of the Rayonix detector. MFX also offers the opportunity to collect during multiplexing beam time which increases the opportunity of beam time to be available for our experiments.

**5. Shift Requirements**

We request a total of two 12-hour shifts, assuming the standard configuration is in place and tested. We expect that we can collect two datasets per 12-hour shift, such that two shifts will allow us to collect the apo enzyme dataset as well as datasets corresponding to our timepoints with introduced substrate.

**References**

M.A. Wilson & A.T. Brunger. (2000) The 1.0 Å crystal structure of $Ca^{2+}$-bound calmodulin: an analysis of disorder and implications for functionally relevant plasticity. J Mol Biol 301:1237-56

J. Lin, H. van den Bedem, A.T. Brunger and M.A. Wilson (2016) Atomic resolution experimental phase information reveals extensive disorder and bound 2-methyl-2,4-pentanediol in $Ca^{2+}$-calmodulin. Act Cryst D72:83-92

D.A. Keedy et al (2014) Crystal Cryocooling Distorts Conformational Heterogeneity in a Model Michaelis Complex of DHFR. Structure 22:899-910.

H. van den Bedem et al (2013) Automated identification of functional dynamic contact networks from X-ray crystallography. Nature Meth 10:896-902.

H. van den Bedem & J.S.Fraser (2015) Integrative, Dynamic Structural Biology at Atomic Resolution--It's About Time. Nature Meth 12:307-318.

Wall, M., Ealick, S. & Gruner, S. (1997) Three-dimensional continuous X-ray scattering from crystals of Staphylococcal nuclease. Proc Natl Acad Sci USA 94, 6180–6184.
Schmidt & Saldin (2014) Enzyme transient state kinetics in crystal and solution from the perspective of a time resolved crystallographer. Structural Dynamics 1, 024701.
Lakshminarasimhan et al., (2010) Evolution of new enzymatic function by structural modulation of cysteine reactivity in pseudomonas fluorescens Isocyanide Hydratase. J Biol. Chem. 285:29651–29661.
Sierra et al (2015) Concentric-flow electrokinetic injector enables serial crystallography of ribosome and photosystem II. doi:10.1038/nmeth.3667 (Adv. Online publ.)
H. van den Bedem et al (2009) Modeling discrete heterogeneity in X-ray diffraction data by fitting multi-conformers Acta Cryst D65:1107-1117
Burnley et al (2012) Modelling dynamics in protein crystal structures by ensemble refinement. eLife 1:1-29
Fraser et al (2011) Accessing protein conformational ensembles using room-temperature X-ray crystallography. Proc Nat. Acad. Sci. USA 108:16247-16252
Ten Boer et al (2014) Time Resolved Serial Crystallography Captures High Resolution Intermediates of Photoactive Yellow Protein. Science 346:1242-1246.
D.A. Keedy, J.S. Fraser & H. van den Bedem (2015) Exposing Hidden Alternative Backbone Conformations in X-ray Crystallography Using qFit. PLoS Comput. Biol. 11:e1004507
Wall, M. E. in Micro and Nano Technologies in Bioanalysis 544, 269–279 (2009).
Van Benschoten, A.H. *et al.* (2016) Measuring and modeling continuous scattering in protein X-ray crystallography. Proc Natl Acad Sci USA. 113:4069-74

# MFX Parameter Table (Run 17 Standard Configuration #2)

If multiple samples are planned, please add rows to the table and list all the samples proposed.

| Parameter Table for the MFX Standard Configuration #2 | | |
|---|---|---|
| **Sample & sample delivery** | **Sample name, description** | Isocyanide Hydratase, 230-residue homodimeric enzyme |
| | **Sample State (crystalline, solution, etc.)** | crystalline |
| | **Sample Temperature (C)** | 277K |
| | **Sample Delivery Method (GDVN, LCP, MESH, etc.)** | coMESH (mix-and-inject) |
| | **High-resolution jet viewing needed?** | yes |
| **Scattering geometry** | **Desired Measurement (SAXS/WAXS/SFX)** | SFX |
| | **Post-sample attenuator needed? Yes/No. Specify** | No |
| | **Desired Highest Resolution (Å)** | 1.5A |
| | **Front Detector Distance (mm)** | |
| **X-ray parameters** | **X-ray Energy (keV)** | max available |
| **Optical beam parameters** | **Pump laser needed?** | No |
| | **Wavelength (nm)** | |
| | **Pulse duration range?** | |
| | **Maximum pulse energy (µJ)** | |
| | **Focal size (FWHM) (µm)** | |
| | **Minimum fluence on target ($mJ/cm^2$)** | |
| | **Polarization requirements?** | |
| **X-ray Beam Time** | **Number of shifts [1 shift = 12 hrs]** | 2 |
| **Any additional comments** | | |