

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property

Organization

International Bureau

(43) International Publication Date

25 November 2021 (25.11.2021)



(10) International Publication Number

WO 2021/237116 A1

(51) International Patent Classification:

G01N 23/225 (2018.01) H01J 37/285 (2006.01)

G01N 23/00 (2006.01) G01N 1/42 (2006.01)

B01L 99/00 (2010.01) G02B 21/00 (2006.01)

H01J 37/26 (2006.01) G01N 21/77 (2006.01)

Declarations under Rule 4.17:

— of inventorship (Rule 4.17(iv))

Published:

— with international search report (Art. 21(3))

(21) International Application Number:

PCT/US2021/033701

(22) International Filing Date:

21 May 2021 (21.05.2021)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/028,759 22 May 2020 (22.05.2020) US

(71) Applicant: THE REGENTS OF THE UNIVERSITY OF

CALIFORNIA [US/US]; 1111 Franklin Street, Twelfth Floor, Oakland, CA 94607-5200 (US).

(72) Inventors: GONEN, Tamir; 10889 Wilshire Blvd.,

Suite 920, Los Angeles, CA 90095-7191 (US). MAR-

TYNOWYCZ, Michael; 10889 Wilshire Blvd., Suite 920,

Los Angeles, CA 90095-7191 (US).

(74) Agent: HALSTEAD, David, P. et al.; Foley Hoag LLP,

155 Seaport Boulevard, Boston, MA 02210-2600 (US).

(81) Designated States (unless otherwise indicated, for every

kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every

kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(54) Title: METHODS FOR ANALYZING INTERMOLECULAR INTERACTIONS IN MICROCRYSTALS

(57) Abstract: Methods of introducing a small molecule into a crystal of a macromolecule, of obtaining a microcrystal having a macromolecule and a small molecule from a crystal of the macromolecule, of determining a structural model for a complex having a macromolecule and a small molecule, of identifying a small molecule that complexes with a macromolecule, and of screening a library of small molecules for their binding to a macromolecule are disclosed.



WO 2021/237116 A1

METHODS FOR ANALYZING INTERMOLECULAR INTERACTIONS IN MICROCRYSTALS

RELATED APPLICATION

This application claims a right of priority to and the benefit of the filing date of U.S. Provisional Application No. 63/028,759, filed on May 22, 2020, which is hereby incorporated by reference in its entirety.

STATEMENT OF RIGHTS

This invention was made with Government support under Grant Number GM1365081, awarded by the National Institutes of Health. The Government has certain rights in the invention.

BACKGROUND

Atomic resolution structures are critical to understanding how proteins and small molecules function. Knowing how the atoms in a molecule are arranged allow for rational design of ligands such as drugs to bind and inhibit unwanted function or otherwise alter the structure or function of the protein. Investigation of protein structures with these small molecule ligands bound is an important part of drug discovery.

Macromolecular crystallography has been adapted to investigate these ligand-protein interactions by incorporating the ligand into the protein crystal lattice by either soaking ligands into crystals or co-crystallizing the protein with the ligand. However, successfully coercing ligands into protein crystals is challenging. Co-crystallization of the protein with the ligand can change the crystallization condition or result in the ligand forming close contacts between protein molecules in the lattice instead of effectively diffusing into the binding site. Therefore, soaking already formed crystals with ligands can be preferable, but still has its drawbacks.

Crystals soaked with ligands will often crack or dissolve as the additional ligand diffuses into the protein lattice. Cracked or dissolved crystals can eliminate the order that allows the crystals to diffract. It is also unknown how deeply or efficiently the ligand will soak into the protein crystals typically used in X-ray diffraction experiments. Ligand will often only partially occupy the protein crystal leading to weak density in its binding location. Another challenge in ligand soaking experiments for macromolecular crystals arises in modeling the weak density that the ligands occupy. Improved methods for studying and analyzing these interactions are needed.

SUMMARY OF THE INVENTION

In some aspects, methods of introducing a small molecule into a crystal of a macromolecule include applying a first sample having the crystal of the macromolecule onto an electron microscopy (EM) grid; and applying a second sample having the small molecule onto the EM grid to introduce the small molecule into the crystal.

In certain aspects, methods of obtaining a microcrystal having a macromolecule and a small molecule from a crystal of the macromolecule include applying a first sample having the crystal of the macromolecule onto an electron microscopy (EM) grid; applying a second sample having the small molecule onto the EM grid to obtain a crystal having the macromolecule and the small molecule; plunging the EM grid into a cryogenic liquid; placing the EM grid into a focused ion beam scanning electron microscope (FIB/SEM); and milling the crystal having the macromolecule and the small molecule in the FIB/SEM using a focused ion beam to obtain a microcrystal having the macromolecule and the small molecule.

In some aspects, methods of determining a structural model for a complex having a macromolecule and a small molecule include, in addition to features of the aspects for obtaining a microcrystal having a macromolecule and a small molecule from a crystal of the macromolecule, placing the microcrystal having the macromolecule and the small molecule into a transmission electron microscope (TEM); collecting a plurality of electron diffraction patterns from the microcrystal; and determining a structural model based on the plurality of electron diffraction patterns.

In certain aspects, methods of identifying a small molecule that complexes with a macromolecule include determining a structural model for a complex, and characterizing the small molecule as one that complexes with the macromolecule based on the structural model. Such aspects can be used in the context of screening a library of small molecules for their binding to a macromolecule.

Further embodiments and details for each of these aspects is presented throughout the disclosure.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1. On-grid soaking of ligands into protein microcrystals. (A) Schematic cartoon of the on-grid soaking procedure. From left to right: protein solution is applied to the carbon side of a glow-discharged holey carbon film, the grid is blotted from the back in high humidity for ~20s, the ligand solution is applied to the same grid and allowed to incubate for 20s, the grid is blotted from the back again for ~20s. The grid is then vitrified by plunging

into liquid ethane and stored in liquid nitrogen. (B) SEM image of a grid after on-grid soaking of proteinase K crystals with I3C solution with hundreds of crystals to be potentially investigated. (C) High resolution SEM image showing crystal density and distribution. Scale bar 1mm in (B) and 50 μ m in (C).

Figure 2. Cryo-FIB milling of a proteinase K microcrystal soaked with I3C. (A) SEM and (B) FIB images of a 10 μ m protein crystal before cryo-FIB milling. The crystal after milling the back crystal to a thickness of 200nm imaged using the (C) SEM and (D) FIB. (E) The resulting diffraction from this lamella at 0° for 1s. Inset showing visible high resolution spots. Scale bars 10 μ m in all images. Red arrow indicates crystal selected for investigation.

Figure 3. I3C ligands soaked into the protein lattice. (Top Left) Cartoon representation of the determined protein (beige) and the identified ligand binding sites as sticks and spheres colored by element. (Right and bottom) corresponding numbered sites with their densities before modeling (left side), and after the ligand was modeled and refined (right). Blue meshes are 2F_o-F_c maps contoured at the 1.0 σ level, and green meshes are F_o-F_c maps contoured at the 3 σ levels. The iodine atoms form an equilateral triangle with side lengths ~6Å.

Figure 4. Comparison of ligand placement between a microcrystal determined by MicroED and a large crystal determined using X-ray diffraction. (Top left) MicroED (beige) and X-ray determined (light blue) structures overlaid with their bound ligands. The overall Ca RMSD between the models is 0.18 Å. (Right and bottom) For each numbered ligand, the corresponding 2F_o-F_c and F_o-F_c densities using only the protein phases are shown with the MicroED on the left and X-ray on the right for comparison. Ligand 3 in the X-ray structure was not observed. All 2F_o-F_c maps are contoured at the 1s level, and F_o-F_c maps are contoured at the 3s level. Maps are presented with a 1.75 Å carve for clarity. Density maps for each are calculated at 1.78 Å resolution.

DETAILED DESCRIPTION OF THE INVENTION

A method for soaking ligands into macromolecule (e.g., protein) crystals (e.g., microcrystals) on EM (e.g., TEM) grids is presented. One or more crystals on the grid are soaked simultaneously using standard cryo-EM vitrification equipment. An embodiment of the method is demonstrated, through the Examples, using proteinase K microcrystals soaked with the 5-amino-2,4,6-triiodoisophthalic acid (I3C) magic triangle. A soaked microcrystal is milled to a thickness of 200 nm using a focused ion-beam, and microcrystal electron diffraction (MicroED) data are collected. A high-resolution structure of the protein with four

ligands at high occupancy is determined. Results suggest that soaking ligands into microcrystals may result in a more efficient uptake than in larger crystals that are typically used in drug discovery pipelines.

The discoveries underlying the disclosed methods can be used to introduce a small molecule into a crystal of a macromolecule, to obtain a microcrystal having a macromolecule and a small molecule from a crystal of the macromolecule, to determine a structural model for a complex having a macromolecule and a small molecule, to identify a small molecule that complexes with a macromolecule, and to screen a library of small molecules for their binding to a macromolecule.

Definitions

As used in the description, the words “a” and “an” can mean one or more than one. As used in the claims in conjunction with the word “comprising,” the words “a” and “an” can mean one or more than one. As used in the description, “another” can mean at least a second or more.

A “structural model” of a molecule (or compound) is a model of its molecular geometry, and it describes the three-dimensional arrangement of its atoms. For example, in a Cartesian coordinate system for a three-dimensional space, the set of (x, y, z) coordinates for each atom of the molecule would describe the three-dimensional arrangement of the atoms. The structural model of a molecule can describe the three-dimensional arrangement of all of its atoms, or can describe the three-dimensional arrangement of fewer than all of its atoms (e.g., if some atoms, such as hydrogen atoms, have not been resolved due to experimental limitations). In particular, the structural model of a molecule describes the identity (e.g., via an atomic number) and three-dimensional arrangement (e.g., with respect to a chosen reference frame, for example in a Cartesian coordinate system having an atom of the molecule at its center) of at least 60% (e.g., 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100%) of the non-hydrogen atoms of the molecule. The structural model of a molecule can describe the connectivity of the molecule’s atoms, and it can identify the molecule; it need not describe a solution conformation of the molecule. It also need not identify the stereochemistry of the molecule (although in some of the disclosed embodiments, it does identify stereochemical information). In addition, the bond lengths and angles need not correspond to those of an average ensemble solution conformation.

The “unit-cell parameters” include the lengths of the three edges of the unit cell. Depending on the type of unit cell, fewer than three values may suffice to describe the unit-

cell parameters. For example, for a cubic unit cell, each of the edges would have the same length; thus, a single value for an edge can be sufficient as length information. The unit-cell parameters also include the three angles of the unit cell. As with edges, fewer than three values may suffice to describe the angles for some of the unit-cell types.

The “space group” describes the overall symmetry of the microcrystal, and it can be one of the 230 possible space groups. Exemplary space groups include P_1 , $P2_12_12_1$, and C_2 . For a biological molecule (e.g., peptide), 65 of the 230 space groups can be relevant, since the absence of mirror planes, glide planes, centers of symmetry, or rotation inversion axes in such molecules make the remaining space groups inapplicable. *See, e.g., Cantor & Schimmel, Biophysical Chemistry, Part II: Techniques for the Study of Biological Structure and Function*, p. 736 (W. H. Freeman and Company 1980).

“Collecting” an electron-diffraction pattern includes obtaining (e.g., via recording) a description of the electron-diffraction pattern in any of the available detection means, such as on a film, an image plate, via an indirect charge-coupled device (CCD) detector, via an indirect complementary metal–oxide–semiconductor (CMOS) detector, via direct monolithic active pixel sensors (MAPS) (e.g., based on CMOS technology or pnCCDs), or via a hybrid pixel detector (e.g., HPAD). As further detailed herein, certain embodiments focus on the use of direct electron detectors.

A “cryogenic fluid” is a liquefied gas that can be kept in liquid state at temperatures below -150°C . An exemplary cryogenic fluid is liquid nitrogen.

The term “phase” is used herein in the sense of the phase of the diffracted electrons described as matter behaving as waves having an amplitude as a function of time and space that depends on phase, wavelength, and maximal amplitude.

A “polymorphic form” of a molecule results from an arrangement or conformation of the constituents of its microcrystal that might differ from another microcrystal of the same molecule.

Samples

Crystals of macromolecules can be obtained via standard methods. For example, they can be grown through controlled precipitation from an aqueous solution (e.g., by using ammonium sulfate, by using polyethylene glycol).

In some embodiments, a full electron diffraction pattern set can be collected from a single microcrystal, allowing for the determination of a three-dimensional structural model for the molecule from that single microcrystal without relying on any other microcrystal. The

microcrystal can have two of its dimensions (e.g., edges) longer than a micrometer (e.g., 1.1, 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10 micrometers), while one of its dimensions can be shorter than two micrometers (e.g., 20, 40, 60, 80, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000, 1500 nanometers). The methods disclosed herein can be used without confirming the presence of the microcrystals beforehand (e.g., via regular human vision or via common light microscopes, which may fail to detect the microcrystals); the samples can be loaded into an electron microscope in order to locate the microcrystals once the sample grid is already in the electron microscope.

The sample can be applied to an electron microscopy grid. Before applying the sample, the electron microscopy grid can be layered with a thin layer of amorphous carbon by evaporating carbon graphite on it (other materials for support can be used such as formvar, silicone monoxide, lacey carbon, and graphene). The electron microscopy grid can be further subjected to glow discharging to improve the way the microcrystals in the sample distribute on the grid. Once the sample is applied to the grid, the grid may further be surrounded by a cryogenic fluid to maintain the microcrystals at a cryogenic temperature for further data collection.

Systems for Determining Structural Models

In some aspects, the methods disclosed herein use electron microscopes (e.g., transmission electron microscopes). Standard electron microscopes can be used, without modification, for carrying out the methods disclosed herein. Some of the sources for obtaining electron microscopy instruments include Thermo Fisher Scientific (168 Third Avenue, Waltham, MA USA 02451); Hitachi High Technologies America, Inc. (10 North Martingale Road, Suite 500, Schaumburg, IL 60173-2295); and JEOL Ltd. (3-1-2 Musashino, Akishima, Tokyo 196-8558, Japan). An electron microscope can be operated at a suitable voltage, for example between 40 kV and 400 kV. Suitability of a voltage can be determined by considering the de Broglie wavelength of the electrons in light of the desired resolution as well as by considering the permissible range of electron doses in light of the tolerable radiation damage. For example, the de Broglie relationship provides the wavelength of an electron as $\lambda = h/(m_e v)$, in which λ is the wavelength of the electron, h is the Planck's constant, m_e is the mass of the electron, and v is the velocity of the electron. Given that an electron accelerated by a voltage difference of Φ in volts will have a kinetic energy of $e\Phi = (1/2)m_e v^2$, solving the two equations provides the wavelength in Angstroms of the electron in terms of the electron microscope voltage in volts as $\lambda = 12.3/\sqrt{\Phi}$. See, e.g., Cantor &

Schimmel, *Biophysical Chemistry, Part II: Techniques for the Study of Biological Structure and Function*, p. 820 (W. H. Freeman and Company 1980). Using this equation, one can determine that the wavelength of the electron would approximate 0.06 Å at a voltage of 40kV, 0.04 Å at a voltage of 100kV, 0.03 Å at a voltage of 200kV, and 0.02 Å at a voltage of 300 kV. Each of these wavelengths is sufficient to resolve the individual atoms of a molecule, and each is shorter than the typical wavelengths of X-rays used in X-ray crystallography.

Some of the particular types of electron microscopes that can be used include Focused Ion Beam (FIB) systems, DualBeam Focused Ion Beam-Scanning Electron Microscope systems (FIB-SEMs), and Transmission Electron Microscopes (TEMs).

Various detectors can be used to record electron diffraction patterns. For example, films, image plates, indirect charge-coupled device detectors, indirect complementary metal-oxide-semiconductor detectors, direct monolithic active pixel sensors, or hybrid pixel detectors can be used. These may be obtained from sources similar to those for electron microscopes, as well as from other sources such as Gatan, Inc. (5794 W. Las Positas Blvd., Pleasanton, CA 94588) (e.g., for CCD or CMOS cameras, as well as for specimen holders) and Electron Microscopy Sciences (P.O. Box 550, 1560 Industry Road, Hatfield, PA 19440) (e.g., for Kodak films). These detectors can be used without modification or without creating any specialized arrangements that differ from those used for traditional transmission electron microscopy (e.g., used in diffraction mode). Although hybrid pixel detectors are more advanced than the other types of detectors in some ways, use of them is not necessary for the methods disclosed herein: any of the commonly available detectors can be used to practice the disclosed methods. For methods that require obtaining structural information quickly however (e.g., within an hour), traditional Kodak films are impractical; thus, in such cases, digital detectors (e.g., indirect CCD, indirect CMOS, direct CMOS or other MAPS, or HPAD) are preferred.

In some embodiments, the used detectors are direct electron detectors or electron counting devices. For example, direct detection devices (DDD) from Gatan Inc. (5794 W. Las Positas Blvd., Pleasanton, CA 94588), FEI (5350 NE Dawson Creek Drive, Hillsboro, Oregon 97124 USA; operating as a subsidiary of Thermo Fisher Scientific), or Direct Electron, LP (13240 Evening Creek Drive, Suite 311, San Diego, CA 92128 USA) can be used. Some of the particular devices include Gatan K2 and K3, FEI/ThermoFisher Falcon 3EC (and newer Thermo Fisher Falcon 4), and Direct Electron DE-20. *See generally* G. McMullan, A. R. Faruqi, and R. Henderson, Direct Electron Detectors, *Methods in Enzymology* 579: 1-17.

Electron microscopy grids can be obtained from Ted Pella, Inc. (P.O. Box 492477, Redding, CA 96049-2477); Millipore Sigma (3050 Spruce St., St. Louis, MO 63103); as well as from Electron Microscopy Sciences (P.O. Box 550, 1560 Industry Road, Hatfield, PA 19440).

Methods of Processing Electron-diffraction Patterns

In some aspects, the methods used to determine structural models of molecules are the standard ones used in X-ray crystallography. For example, using the collection of programs within the CCP4 suite (available from the World Wide Web at ccp4.ac.uk/), one can index and integrate the reflections (e.g., using iMosflm, xia2/DIALS); reduce the data to a higher quality subset (e.g., using AIMLESS); estimate phases from molecular replacement (e.g., using PHASER, MOLREP); estimate phases using direct methods (e.g., SHELXD); estimate phases using multiple isomorphous replacement (e.g., peakmax, rsps, vecref, vectors); estimate phases using multiwavelength anomalous dispersion (e.g., crossec, rantan, revise); and refine structural models (e.g., using REFMAC5), in addition to performing numerous other steps for determining structural models. The theories underlying these implementations are available through numerous publications (e.g., B. Rupp, *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology*, (Garland Science, ed. 1, 2009) for biomolecules; G. Taylor, Introduction to Phasing, *Acta Crystallographica Section D* 66: 325–38 (2010); K. Cowtan, Phase Problem in X-ray Crystallography, and Its Solution, *Encyclopedia of Life Sciences* 1–5 (2001); and I. Uson & G. M. Sheldrick, Advances in Direct Methods for Protein Crystallography, *Current Opinion in Structural Biology* 9: 643–48 (1999) for phasing). For indexing and integration, one can also use XDS, and for direct methods of estimating phases, one can also use SHELXT or SnB. Alternatively, various independently developed scripts can be used for data processing (e.g., those available at cryoem.ucla.edu/pages/MicroED).

Methods of Forming Microcrystals

According to some aspects, methods are disclosed to introduce a small molecule into a crystal of a macromolecule, and to obtain a microcrystal having a macromolecule and a small molecule from a crystal of the macromolecule.

Some of the disclosed methods include applying a first sample having the crystal of the macromolecule onto an electron microscopy (EM) grid; and applying a second sample having the small molecule onto the EM grid to introduce the small molecule into the crystal,

e.g., by diffusion. In some embodiments, the grid can contain more than one crystal, which would enable introducing the small molecule into more than one crystal via one process. This contrasts with the traditional method of soaking a crystal, one at a time, into a solution of the small molecule. The application steps can be performed on the carbon-coated side of the EM grid, and blotting, for example with a filter paper, can be performed on the other side (e.g., the copper side). Any of the steps/methods can be under cryogenic conditions (e.g., with the sample/grid in liquid nitrogen).

The steps, when performed, can be performed for various durations. For example, each of the blotting after application of the first sample, the incubation of the second sample once it is applied, and the blotting after the application of the second sample can be for 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, or 90 seconds.

Some of the disclosed methods include placing the EM grid into a focused ion beam scanning electron microscope (FIB/SEM); and milling the crystal having the macromolecule and the small molecule in the FIB/SEM using a focused ion beam to obtain a microcrystal having the macromolecule and the small molecule. The crystal can be milled to have a dimension that is less than a certain thickness, for example less than 100, 150, 200, 250, 300, 350, or 400 nanometers. Before the milling, some methods include depositing platinum on the crystal having the macromolecule and the small molecule by sputter coating the grid.

Methods of Processing Microcrystals

In some aspects, the disclosure relates to methods of determining a structural model for a complex having a macromolecule and a small molecule, of identifying a small molecule that complexes with a macromolecule, and of screening a library of small molecules for their binding to a macromolecule.

Some of these methods rely on forming a microcrystal as described in the previous section, for example via milling in a FIB/SEM. Such a microcrystal can be placed into a transmission electron microscope (TEM), and electron diffraction patterns from the microcrystal can be collected. Thereafter, a structural model can be determined using suitable methods, for example such as those detailed in the Methods of Processing Electron-diffraction Patterns section.

The collection of electron diffraction patterns can include continuously rotating the EM grid in the TEM. In various embodiments, the microcrystal is rotated around the axis by at least 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, or 140 degrees, or by any value/range between these values. The diffraction patterns, in some embodiments, are collected not from the whole microcrystal, but from a portion of it (e.g., the portion that has been milled to have a suitable thickness). The diffraction patterns can be collected using an electron beam in a direction orthogonal to the milling direction.

The disclosed methods for determining a structural model, in conjunction with the disclosed methods for introducing a small molecule into a crystal of a macromolecule and the disclosed methods for milling such a crystal, can be used to identify a small molecule that binds to the macromolecule and to identify how such a binding occurs (e.g., the stoichiometry of binding, the sites of binding). Such an identification can be in the context of a high-throughput drug discovery using MicroED, for example for drug molecules that affect the function of the macromolecule.

EXAMPLES

Example 1: Efficient, high-throughput ligand incorporation into protein microcrystals by on-grid soaking

A method for soaking ligands into protein microcrystals on TEM grids is presented. Every crystal on the grid is soaked simultaneously using only standard cryoEM vitrification equipment. The method is demonstrated using proteinase K microcrystals soaked with the 5-amino-2,4,6-triodoisophthalic acid (I3C) magic triangle. A soaked microcrystal is milled to a thickness of 200nm using a focused ion-beam, and microcrystal electron diffraction (MicroED) data are collected. A high-resolution structure of the protein with four ligands at high occupancy is determined. Results suggest that soaking ligands into microcrystals may result in a more efficient uptake than in larger crystals that are typically used in drug discovery pipelines.

Introduction

Atomic resolution structures are critical to understanding how proteins and small molecules function. Crystal structures of small molecule pharmaceuticals are also important to determine how drugs are formulated in their delivery (Datta & Grant, 2004). Knowing how the atoms in a molecule are arranged allow for rational design of ligands such as drugs to bind and inhibit unwanted function (He *et al.*, 2010; Daga *et al.*, 2010). Investigation of

protein structures with these small molecule ligands bound is an important part of drug discovery (Kuhn *et al.*, 2002; Tickle *et al.*, 2004; Blundell, 2017). Macromolecular crystallography has been adapted to investigate these ligand-protein interactions by incorporating the ligand into the protein crystal lattice by either soaking ligands into crystals or co-crystallizing the protein with the ligand (Beck *et al.*, 2009). Successfully coercing ligands into protein crystals is challenging (McNae *et al.*, 2005). Co-crystallization of the protein with the ligand can change the crystallization condition or result in the ligand forming close contacts between protein molecules in the lattice instead of effectively diffusing into the binding site (Reynolds, 2014). Soaking already formed crystals with ligands is the standard approach (Beck *et al.*, 2009; Lebioda & Zhang, 1992; López-Jaramillo *et al.*, 2002). Crystals soaked with ligands will often crack or dissolve as the additional ligand diffuses into the protein lattice. Cracked or dissolved crystals can eliminate the order that allows the crystals to diffract. It is also unknown how deep or efficiently the ligand will soak into the large protein crystals typically used in X-ray diffraction experiments. Ligand will often only partially occupy the protein crystal leading to weak density in its binding location. A large challenge in ligand soaking experiments for large macromolecular crystals arises in modeling the weak density that the ligands occupy (Pozharski *et al.*, 2013; Liebschner *et al.*, 2017; Pearce *et al.*, 2017).

Microcrystal electron diffraction (MicroED) is an electron cryo-microscopy (cryoEM) method for determining atomic resolution structures from protein microcrystals in a transmission electron microscope (Shi *et al.*, 2013; Nannenga, Shi, Leslie *et al.*, 2014; Nannenga, Shi, Hattne *et al.*, 2014). MicroED has been used to solve novel protein and small molecule pharmaceutical crystal structures (Jones *et al.*, 2018; Sawaya *et al.*, 2016; Rodriguez *et al.*, 2015). Previous ligand-protein interactions determined using MicroED relied on co-crystallization of the protein and ligand prior to applying the crystals to the grid (Purdy *et al.*, 2018; Seidler *et al.*, 2018; Clabbers *et al.*, 2020). In this way, novel co-crystals of the HIV-GAG bevirimat complex were determined (Purdy *et al.*, 2018). The first-generation HIV-I maturation inhibitor bevirimat was found to occupy the six-fold axis of the protein hexamer along the c axis of the unit cell. The density suggested that the drug bound non-specifically in this position to any of the six protein interfaces, and served as a non-specific inhibitor of the protein capsid, ultimately preventing maturation of the viral capsid. A recent investigation of protein microcrystals using MicroED was also able to correctly locate the inhibitor acetazolamide (AZM) in the active site of human carbonic anhydrase isoform II (HCA II) that it had been co-crystallized with (Clabbers *et al.*, 2020). Other drug discovery

efforts utilizing MicroED did not utilized ligand-protein interactions. For example, MicroED was used to determine structures of fibrillar segments of human tau protein. These structures were used to template the beta sheet interface and ultimately design a peptide inhibitor to fibril formation (Seidler *et al.*, 2018). No crystals of this fibril inhibitor complex could be determined since fibrils were no longer formed when incubated together. Time resolved experiments that expose the protein to stimuli just prior to vitrification for cryoEM imaging were similarly pioneered by Unwin (Unwin, 1995).

We demonstrate here a simple, high-throughput method for soaking ligands into protein microcrystals on a TEM grid. The method involves first applying crystals to a grid similarly to any other cryoEM experiment, back blotting, adding the ligand solution, and back blotting again prior to plunging into liquid ethane. This method is demonstrated by on-grid soaking the I3C magic triangle ligand into microcrystals of proteinase K. On-grid soaking is coupled with cryo-FIB milling which removes any excess material and increases the quality of data from the selected crystal. CryoFIB milling also allows investigation of ligand soaked protein crystals of any size, that greatly expands the scope of the method. Our data shows that the I3C ligand soaks into the crystals with higher efficiency when compared to using much larger macromolecular crystals under similar conditions. These results have far-reaching implications in future investigations utilizing MicroED for drug discovery. The approach may also serve as a platform for soaking experiments of interest in other cryoEM modalities, such as single particle or cryotomography.

Results

Soaking the ligand into the crystals

Proteinase K crystals were grown in batch to an average size of 1-20um. Crystal slurry was applied by a pipette to the carbon side of the grid, and the grid was blotted from the back inside a temperature and humidity controlled blotting chamber. In this way, the liquid flows through the holes of the holey carbon film and the crystals remain behind. The grid essentially acts as a filter for the crystals. To the grid of blotted crystals, I3C magic triangle was added to the grid and allowed to incubate. This step is mimetic of soaking a looped crystal into a ligand solution. The grid was blotted from the back again, and immediately plunged into liquid ethane (Figure 1). The grids were then transferred to either a Thermo-Fisher Aquilos dual beam FIB/SEM or a Thermo-Fisher Talos Arctica TEM for inspection (Figure 1).

Identification and selection of ligand soaked crystals

Crystals on the TEM grid were identified in the SEM after platinum coating. The crystal density was high with crystals being present on both the thin carbon film and the film above the copped grid bars (Figure 1). A crystal was identified in both SEM and FIB imaging (Figure 2) as a cluster of two crystals directly next to one another in the approximate center of a grid square, with the front and back crystals measuring approximately 5 μ m and 10 μ m across, respectively. The view of the back crystal was occluded in the FIB image at 18° by another crystal near the grid bar.

Cryo-FIB milling

The occluding crystals in front of and abutting the crystal of interest were milled through in order to access the crystal furthest behind the rest (Figure 2). Rough milling was done in cleaning cross sections from both the top and bottom of the desired lamella location. A final thickness of approximately 200nm was achieved after two polishing steps using a very low gallium beam current.

MicroED data collection

The grid was transferred to a Talos Arctica TEM for MicroED data collection. Crystal lamellae were identified by taking a low-magnification montage. The target lamella appeared as a long streak of white with a semitransparent lamella suspended in the gap. The crystal was screened for diffraction by collecting a single diffraction image at 0° for 1s using a 70 μ m aperture in order to isolate the signal from a small area on the desired crystal. Diffraction spots to >2Å were observed (Figure 2). A continuous rotation MicroED data set was collected between -30 to +30°.

MicroED data analysis

MicroED data were indexed, integrated, and scaled in DIALS as described (Hattne *et al.*, 2015; Parkhurst *et al.*, 2016; Clabbers *et al.*, 2018; Winter *et al.*, 2018; Waterman *et al.*, 2016). A resolution cutoff was applied after integration at 1.78Å, where the CC_{1/2} fell to a value of 0.33. This dataset from a single crystal was found to have an overall completeness of 94.5%, a I/ σ I of 7.8, and R_{pim} of 14%.

Structure determination and modeling

Molecular replacement was performed using the proteinase K model with PDBID 6CL7 to determine the structure (McCoy *et al.*, 2007). This model of proteinase K contains no solvent or ions in the model, making it an ideal for molecular replacement when looking

for ligands. A single, unambiguous solution in $P 4_3 2_1 2$ was found. The initial solution from molecular replacement was inspected for differences in the density corresponding to ligands. We found >25 peaks in the $F_o - F_c$ map with values $>6\sigma$. From these peaks, four I3C ligands could be immediately identified in both the $2F_o - F_c$ and $F_o - F_c$ maps (Figure 3). Prior to placing and refining the molecules, the rings of the I3C molecules were apparent - in some cases even without lowering the contour levels. The quality of these maps made for simple, unambiguous identification of the ligand molecules without any special treatment (such as Polder maps) of the density.

Structure refinement

After placing the four I3C ligands, the structure was refined and solvent molecules were added automatically. Another round of refinement was conducted that allowed the occupancy of the I3C molecules to vary. The four I3C ligands were found to have occupancies between 58 and 75%, with average B-factors of $20 \text{ \AA}^2$. These B-factors are approximately equal to that of the modeled solvent atoms, and slightly higher than the average B-factor of the protein, or $13 \text{ \AA}^2$. The entire model was refined using individual B-factors, rather than group B-factors for the ligands, attesting to the quality of the data. Detailed data collection and refinement statistics are presented in *Table 1*.

Discussion

It has been suggested that the soaking kinetics of ligands into protein crystals should be more efficient for smaller crystals. Our results corroborate this assumption. Soaking I3C into proteinase K crystals $\sim 200 \mu\text{m}$ across under the same conditions that were then investigated using a rotating anode home source resulted in a similarly high resolution of $1.76 \text{ \AA}$ (Beck *et al.*, 2010). This X-ray investigation located 3, I3C molecules with occupancies between 12 - 75% in the proteinase K asymmetric unit. These molecules could be unambiguously identified by the triangular arrangement of iodine atoms. Our experiment showed an unambiguous additional I3C ligand off of Arg290. This site is similar to a site occupied in a brominated analogue to the I3C magic triangle in similarly sized crystals that were soaked in an even higher concentration in this X-ray investigation. The occupancy of the four I3C ligands found in our investigation was higher than that in the X-ray investigation that used large crystals, 58 - 75% vs. 12-75%. The ligands in our experiment also had similarly low B-factors for of only $20 \text{ \AA}^2$. The density of all four of the I3C just after molecular replacement were clear without any modelling, whereas the densities in the X-ray structure were difficult to interpret without refinement (Figure 4). Our data suggest that

ligands such as I3C may soak into much smaller crystals more efficiently without sacrificing map quality or resolution.

Investigating ligand-protein interactions using on-grid soaking has several advantages to typical macromolecular X-ray crystallography ligand soaking pipelines. By initially placing an entire crystallization drop on the grid, many crystals are added to the grid at once, and then all soaked simultaneously. Traditionally, crystals had to be individually looped, soaked into a high concentration ligand solution, back soaked into a cryo-protective solution, and then plunged into liquid nitrogen. This approach requires a high degree of dexterity and has multiple points of failure from operator error. Our method soaks all the crystals at once, requires no manual manipulation of the individual crystals, and does not require a cryo-protective back soak, since properly vitrified crystals show no ice rings (Figure 2). We demonstrate our method on a crystal only 10 μ m across at its widest. Crystals of this size also pose a considerable challenge to accurately fish out with a nylon loop and optimally shoot at synchrotron end stations without microfocus capabilities.

The additional step of a dual beam FIB/SEM to this method adds significant utility. Milling protein crystals has recently seen a high amount of interest for preparing crystals that are too small for synchrotron X-ray experiments, but still too large for MicroED. By milling the ligand soaked crystals we are no longer limited to crystals that happen to be in thin ice and happen to be smaller than 1 μ m thick, though this method is still tractable when a FIB/SEM is unavailable by means of crystal fragmentation (Martynowycz *et al.*, 2017; de la Cruz *et al.*, 2017). Using the FIB/SEM, even crystals large enough for synchrotron or home source X-ray experiments become amenable to our method, while avoiding the laborious looping and plunging steps. Milling crystals also reduces the background contributed from the vitreous ice layer and carbon film surrounding crystals on TEM grids. Soaking microcrystals on TEM grids may even prove to be a preferable emerging method for synchrotron X-ray investigations. Recent results have shown that vitrified protein microcrystals on TEM grids can be located ahead of time using an SEM, and then subsequently used for X-ray diffraction experiments. We also envision applications to single particle cryoEM and cryotomography workflows where active compounds are added on-grid and allowed to interact just before vitrification, allowing for time resolved investigations of ligand activity on varying scales.

The I3C magic triangle was chosen to demonstrate on-grid crystal soaking because it has a combination of amine and carboxyl groups. These exposed groups make the triangle highly hydrophilic, and more likely to incorporate into the crystal. Other ligands such as

small molecule drugs may be more hydrophobic and may not soak into the crystals as well. The I3C ligand was designed for phasing in-house macromolecular protein crystal structures using X-ray diffraction, since the anomalous signal from iodine at the wavelength of a standard Cu $K\alpha$ of 1.54 Å is particularly high (Beck *et al.*, 2008, 2009, 2010). There is no similarly strong anomalous signal in MicroED experiments, and the zero degree scattering amplitude from iodine is only marginally stronger than that of carbon (11 vs 2.5) compared to X-ray diffraction experiments (53 vs 6)(Colliex *et al.*, 2006). This compounded with the lack of anomalous signal suggests that the I3C magic triangle may not be particularly useful for phasing MicroED data. Though recent investigations have demonstrated phasing MicroED data using isomorphous differences is plausible, an experimentally phased protein structure has yet to be demonstrated (Martynowycz *et al.*, 2020). It is likely that incorporations of heavy atoms using the on-grid soaking method will prove invaluable to this effort.

Quantitative experiments investigating crystal size, soaking efficiency, and effects on protein-ligand interacts is an avenue for exciting future investigations. On a single grid, crystals can be found in a variety of sizes (Figure 1). Since all the crystals on the grid have been subjected to the same soaking condition without differences between operators or soaking techniques, a systematic investigation of different crystal sizes and ligand occupancies would be of great potential interest. Furthermore, cryo-FIB milling of ligand soaked crystals in specific regions of individual crystals provides a unique opportunity for investigation that may not be possible using traditional means. In this way, it may be possible to determine whether the occupancy of ligands in very large crystals depends on the where crystals are milled – near the edge or middle.

Conclusions

We present a method for high throughput, on-grid soaking of ligands into protein microcrystals that is fast, simple, and effective. On-grid soaking of protein microcrystals is demonstrated and shown to effectively incorporate ligands. The results suggest that soaked microcrystals of the same protein-ligand complex have higher occupancy and number of ligands bound than using a much larger crystal required for X-ray scattering experiments. This method can be used on current cryoEM hardware for sample preparation and data collection workflows. Adding an additional milling step in a FIB/SEM expands the scope of the method to macromolecular protein crystals of any size, potentially even synchrotron X-ray investigations. Our approach has implications for drug discovery, the investigation of

protein-ligand interactions, and investigating the properties of how ligands interact with and flow through crystals.

Example 2: Materials and Methods for Example 1

Materials

Proteinase K (*E. Album*) was purchased from Sigma and used without further purification. 5-amino-2,4,6-triiodoisophthalic acid (I3C) magic triangles were purchased from Hampton and prepared as described (Beck *et al.*, 2010).

Crystallization

Proteinase K crystals were grown in batch by dissolving 5mg of lyophilized proteinase K powder into 1 mL of 1.25M ammonium sulfate at 4°C. Crystals between 1-50µm were formed within one day, and typically crystals were visible under the light microscope within minutes. The crystal soaking solution was 1.25M ammonium sulfate 0.1M I3C.

Grid preparation

Quantifoil R2/2 Cu200 grids were glow discharged for 30s immediately prior to use. Grids were loaded on to a Leica GP2 cryo-plunger inside of a cold room held at 4°C and 35% relative humidity. The blotting chamber was set to 4°C and 90% humidity. Filter paper was added prior to use, and the system was allowed to equilibrate for 15 mins prior to use. Grids were loaded into the plunger and 3µL of proteinase K slurry were applied to the carbon side (front) of the grid and allowed to incubate with the protein drop for 30s. The grid was then gently blotted from the back (copper side) for 20s. After blotting, 3µL of the crystal soaking solution (1.25M ammonium sulfate, 0.1M I3C) was added to the same grid and allowed to incubate again for 20s. The grids were then blotted again from the back for 20s and immediately plunged into liquid ethane and transferred to liquid nitrogen for storage.

Cryo-FIB milling of the protein crystals

Stored grids were clipped with the carbon facing side up and transferred into a cryogenically cooled Thermo-Fisher Aquilos dual beam FIB-SEM for milling. Grid clips were marked with a dot on the top with a sharpie to indicate the milling direction. A thin layer of platinum (~10nm) was deposited on the grids by sputter coating prior to inspecting the grids using the SEM (Martynowycz *et al.*, 2019b). An additional layer (~300nm) of carbon-rich platinum was added on top of this using the gas injection system. We applied the GIS platinum layer in the mapping position to prevent shadowing of the platinum layer and

opened the GIS valve at 12mm rather than 7mm in order to have a slower, more controlled deposition. Crystals were identified in the SEM and FIB, brought to eucentric height, and milled using cleaning cross sections in steps into lamellae until a thickness of ~200nm as described (Martynowycz *et al.*, 2019a; Duyvesteyn *et al.*, 2018). The final polishing step used a gallium beam current of 10pA to remove the last few nm of crystal from the lamellae (Martynowycz *et al.*, 2019b).

MicroED data collection

MicroED data collection was performed very similarly to previous experiments (Martynowycz *et al.*, 2019a,b; Hattne *et al.*, 2015, 2018; Jason de la Cruz *et al.*, 2019; Hattne *et al.*, 2019). Grids with milled crystals were transferred to a Thermo-Fisher Talos Arctica transmission electron microscope after rotating the grids by 90°. Rotation by 90° assures the tilt axis during MicroED data collection is perpendicular to the milling direction, allowing for a greater rotation range of the crystal lamellae. The TEM was operated at liquid nitrogen temperature at an accelerating voltage of 200kV. Data were collected between the real space wedge between -30 and +30° at a rate of 0.25 °/s. Frames were read out every 1s and binned by 2. The total exposure to the crystal lamellae was 2.4 e⁻Å⁻². The diffraction distance was set to 1900mm which corresponds to a crystal to detector distance of 1853mm after taking post-column magnification into account. Camera length was calibrated using diffraction from a molybdenum film prior to loading the protein grid and to check for any distortions in the diffraction (Clabbers *et al.*, 2018, 2017).

MicroED data processing

Data were converted from MRC to SMV format as described, and an ADSC offset of 512 was applied to compensate for negative bias and pixel truncation (Hattne *et al.*, 2016). The data were indexed, integrated, and scaled in DIALS using the general linear background model using a frame pedestal of 512 and a detector gain of 14 (Clabbers *et al.*, 2018; Parkhurst *et al.*, 2016). A resolution cutoff of 1.78Å was selected by the CC_{1/2} = 0.33 criterion (Evans & Murshudov, 2013; Karplus & Diederichs, 2012). The structure was determined by molecular replacement in PHASER using the search model PDB 6CL7 using electron scattering factors (Hattne *et al.*, 2018; McCoy *et al.*, 2007). The TFZ and LLG of the solution were 69.1 and 6856, respectively. Four I3C molecules were visible in the density prior to any refinement or modeling. These are apparent as the I – I distance is known to be ~6Å (Beck *et al.*, 2008). The four I3C molecules were placed manually in COOT using the get monomer command using the three letter code I3C (Emsley & Cowtan, 2004). The

structure was refined in Phenix.refine after creating restraints in Elbow as described (Moriarty *et al.*, 2009; Afonine *et al.*, 2012). The final model was refined using individual isotropic B-factors for the protein, solvent molecules, and I3C molecules. Occupancy was allowed to vary for the I3C molecules.

Figures and tables

Figures were arranged in Microsoft Powerpoint. Images were adjusted and cropped in FIJI (Schindelin *et al.*, 2012). Tables were arranged in Microsoft Excel. Protein models and meshes were generated using PyMol (Schrödinger LLC, 2014).

Table(s)

Table 1: Refinement statistics for a proteinase K microcrystal soaked with I3C.

Proteinase K soaked on-grid with I3C	
Wavelength (Å)	0.0251
Resolution range (Å)	43.32 - 1.78 (1.844 - 1.78)
Space group (#)	P 4 ₃ 2 ₁ 2
Unit cell (a = b, c, $\alpha = \beta = \gamma$)	67.55, 102.77, 90
Total reflections	109326 (10458)
Multiplicity	4.9 (4.8)
Completeness (%)	94.52 (94.56)
Mean I/sigma(I)	7.83 (2.19)
Wilson B-factor	14.37
R _{merge}	0.2622 (0.9953)
R _{pim}	0.1304 (0.5076)
CC _{1/2}	0.971 (0.365)
R _{work}	0.1694
R _{free}	0.2094
Number of non-hydrogen atoms	2339
macromolecules	2029
ligands	64
solvent	246
RMS (bonds)	0.005
RMS (angles)	0.74
Ramachandran favored (%)	97.83
Ramachandran allowed (%)	2.17
Ramachandran outliers (%)	0
Rotamer outliers (%)	0
Clashscore	5.21
Average B-factor	13.51
macromolecules	12.59
ligands	19.44
solvent	19.61

References

- Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68**, 352–367.
- Beck, T., Da Cunha, C. E. & Sheldrick, G. M. (2009). *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **65**, 1068–1070.
- Beck, T., Gruene, T. & Sheldrick, G. M. (2010). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 374–380.
- Beck, T., Krasauskas, A., Gruene, T. & Sheldrick, G. M. (2008). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **64**, 1179–1182.
- Blundell, T. L. (2017). *IUCrJ.* **4**, 308–321.
- Clabbers, M. T. B., Fisher, S. Z., Coincon, M., Zou, X. & Xu, H. (2020). *BioRxiv Mol. Biol.*
- Clabbers, M. T. B., van Genderen, E., Wan, W., Wiegers, E. L., Gruene, T. & Abrahams, J. P. (2017). *Acta Crystallogr. Sect. D Struct. Biol.* **73**, 738–748.
- Clabbers, M. T. B., Gruene, T., Parkhurst, J. M., Abrahams, J. P. & Waterman, D. G. (2018). *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 506–518.
- Colliex, C., Cowley, J. M., Dudarev, S. L., Fink, M., Gjønnnes, J., Hilderbrandt, R., Howie, A., Lynch, D. F., Peng, L. M., Ren, G., Ross, a W., Smith, V. H., Spence, J. C. H., Steeds, J. W., Wang, J., Whelan, M. J. & Zvyagin, B. B. (2006). *International Tables for Crystallography*, Vol. C, pp. 259–429. Chester, England: International Union of Crystallography.
- Daga, P., Patel, R. & Doerksen, R. (2010). *Curr. Top. Med. Chem.* **10**, 84–94.
- Datta, S. & Grant, D. J. W. (2004). *Nat. Rev. Drug Discov.* **3**, 42–57.
- Duyvesteyn, H. M. E., Kotecha, A., Ginn, H. M., Hecksel, C. W., Beale, E. V., de Haas, F., Evans, G., Zhang, P., Chiu, W. & Stuart, D. I. (2018). *Proc. Natl. Acad. Sci.* **115**, 9569–9573.
- Emsley, P. & Cowtan, K. (2004). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2126–2132.
- Evans, P. R. & Murshudov, G. N. (2013). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 1204–1214.
- Hattne, J., Martynowycz, M. W., Penczek, P. A. & Gonen, T. (2019). *IUCrJ.* **6**, 921–926.

- Hattne, J., Reyes, F. E., Nannenga, B. L., Shi, D., De La Cruz, M. J., Leslie, A. G. W. & Gonen, T. (2015). *Acta Crystallogr. Sect. A Found. Adv.* **71**, 353–360.
- Hattne, J., Shi, D., Glynn, C., Zee, C.-T. C.-T., Gallagher-Jones, M., Martynowycz, M. W. M. W., Rodriguez, J. A. J. A. & Gonen, T. (2018). *Structure*. **26**, 759-766.e4.
- Hattne, J., Shi, D., De La Cruz, M. J., Reyes, F. E. & Gonen, T. (2016). *J. Appl. Crystallogr.* **49**, 1029–1034.
- He, Z., Zhang, J., Shi, X.-H., Hu, L.-L., Kong, X., Cai, Y.-D. & Chou, K.-C. (2010). *PLoS One*. **5**, e9603.
- Jason de la Cruz, M., Martynowycz, M. W., Hattne, J. & Gonen, T. (2019). *Ultramicroscopy*. **201**, 1–4.
- Jones, C. G., Martynowycz, M. W., Hattne, J., Fulton, T. J., Stoltz, B. M., Rodriguez, J. A., Nelson, H. M. & Gonen, T. (2018). *ACS Cent. Sci.* **4**, 1587–1592.
- Karplus, P. A. & Diederichs, K. (2012). *Science (80-.)*. **336**, 1030–1033.
- Kuhn, P., Wilson, K., Patch, M. G. & Stevens, R. C. (2002). *Curr. Opin. Chem. Biol.* **6**, 704–710.
- de la Cruz, M. J., Hattne, J., Shi, D., Seidler, P., Rodriguez, J., Reyes, F. E., Sawaya, M. R., Cascio, D., Weiss, S. C., Kim, S. K., Hinck, C. S., Hinck, A. P., Calero, G., Eisenberg, D. & Gonen, T. (2017). *Nat. Methods*. **14**, 399–402.
- Lebioda, L. & Zhang, E. (1992). *J. Appl. Crystallogr.* **25**, 323–324.
- Liebschner, D., Afonine, P. V., Moriarty, N. W., Poon, B. K., Sobolev, O. V., Terwilliger, T. C. & Adams, P. D. (2017). *Acta Crystallogr. Sect. D Struct. Biol.* **73**, 148–157.
- López-Jaramillo, F. J., Moraleda, A. B., González-Ramírez, L. A., Carazo, A. & García-Ruiz, J. M. (2002). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **58**, 209–214.
- Martynowycz, M., Glynn, C., Miao, J., de la Cruz, M. J., Hattne, J., Shi, D., Cascio, D., Rodriguez, J. & Gonen, T. (2017). *BioRxiv*.
- Martynowycz, M. W., Hattne, J. & Gonen, T. (2020). *Structure*. **28**, 458-464.e2.
- Martynowycz, M. W., Zhao, W., Hattne, J., Jensen, G. J. & Gonen, T. (2019a). *Structure*. **27**, 545-548.e2.
- Martynowycz, M. W., Zhao, W., Hattne, J., Jensen, G. J. & Gonen, T. (2019b). *Structure*. **27**, 1594-1600.e2.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Crystallogr.* **40**, 658–674.

- McNae, I. W., Kan, D., Kontopidis, G., Patterson, A., Taylor, P., Worrall, L. & Walkinshaw, M. D. (2005). *Crystallogr. Rev.* **11**, 61–71.
- Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. (2009). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **65**, 1074–1080.
- Nannenga, B. L., Shi, D., Hattne, J., Reyes, F. E. & Gonen, T. (2014). *Elife*. **3**, e03600.
- Nannenga, B. L., Shi, D., Leslie, A. G. W. & Gonen, T. (2014). *Nat. Methods*. **11**, 927–930.
- Parkhurst, J. M., Winter, G., Waterman, D. G., Fuentes-Montero, L., Gildea, R. J., Murshudov, G. N. & Evans, G. (2016). *J. Appl. Crystallogr.* **49**, 1912–1921.
- Pearce, N. M., Krojer, T. & von Delft, F. (2017). *Acta Crystallogr. Sect. D Struct. Biol.* **73**, 256–266.
- Pozharski, E., Weichenberger, C. X. & Rupp, B. (2013). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 150–167.
- Purdy, M. D., Shi, D., Chrastowicz, J., Hattne, J., Gonen, T. & Yeager, M. (2018). *Proc. Natl. Acad. Sci.* **115**, 13258–13263.
- Reynolds, C. H. (2014). *ACS Med. Chem. Lett.* **5**, 727–729.
- Rodriguez, J. A., Ivanova, M. I., Sawaya, M. R., Cascio, D., Reyes, F. E., Shi, D., Sangwan, S., Guenther, E. L., Johnson, L. M., Zhang, M., Jiang, L., Arbing, M. A., Nannenga, B. L., Hattne, J., Whitelegge, J., Brewster, A. S., Messerschmidt, M., Boutet, S., Sauter, N. K., Gonen, T. & Eisenberg, D. S. (2015). *Nature*. **525**, 486–490.
- Sawaya, M. R., Rodriguez, J., Cascio, D., Collazo, M. J., Shi, D., Reyes, F. E., Hattne, J., Gonen, T. & Eisenberg, D. S. (2016). *Proc. Natl. Acad. Sci.* **113**, 11232–11236.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P. & Cardona, A. (2012). *Nat. Methods*. **9**, 676–682.
- Schrödinger LLC (2014).
- Seidler, P. M., Boyer, D. R., Rodriguez, J. A., Sawaya, M. R., Cascio, D., Murray, K., Gonen, T. & Eisenberg, D. S. (2018). *Nat. Chem.* **10**, 170–176.
- Shi, D., Nannenga, B. L., Iadanza, M. G. & Gonen, T. (2013). *Elife*. **2013**, e01345.
- Tickle, I., Sharff, A., Vinkov, M., Yon, J. & Jhoti, H. (2004). *Chem. Soc. Rev.* **33**, 558.
- Unwin, N. (1995). *Nature*. **373**, 37–43.

Waterman, D. G., Winter, G., Gildea, R. J., Parkhurst, J. M., Brewster, A. S., Sauter, N. K. & Evans, G. (2016). *Acta Crystallogr. Sect. D, Struct. Biol.* **72**, 558–575.

Winter, G., Waterman, D. G., Parkhurst, J. M., Brewster, A. S., Gildea, R. J., Gerstel, M., Fuentes-Montero, L., Vollmar, M., Michels-Clark, T., Young, I. D., Sauter, N. K. & Evans, G. (2018). *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 85–97.

INCORPORATION BY REFERENCE

Each publication and patent mentioned herein is hereby incorporated by reference in its entirety. In case of conflict, the present specification, including any definitions herein, will control.

EQUIVALENTS

While specific embodiments of the subject invention have been discussed, the above specification is illustrative and not restrictive. Many variations of the invention will become apparent to those skilled in the art upon review of the preceding description and the following claims. The full scope of the invention should be determined by reference to the claims, along with their full scope of equivalents, and by reference to the rest of the specification, along with such variations.

CLAIMS

What is claimed is:

1. A method of introducing a small molecule into a crystal of a macromolecule, comprising
applying a first sample comprising the crystal of the macromolecule onto an electron microscopy (EM) grid; and
contacting the crystal with a second sample comprising the small molecule to introduce the small molecule into the crystal.
2. The method of claim 1, wherein the EM grid has a carbon side and a copper side, and wherein both said applying the first sample and said contacting with the second sample are on the carbon side.
3. The method of claim 2, further comprising blotting the EM grid from the copper side after applying said first sample but before contacting with said second sample.
4. The method of claim 2 or 3, further comprising blotting the EM grid from the copper side after contacting with said second sample.
5. The method of any one of claims 1 to 4, further comprising plunging the EM grid into a cryogenic liquid.
6. The method of claim 5, wherein the cryogenic liquid is liquid ethane.
7. The method of any one of claims 1 to 6, comprising simultaneously introducing a plurality of said small molecules into a plurality of crystals of said macromolecule, e.g., on one EM grid, such that different crystals are contacted with the same small molecule simultaneously, or, e.g., in a spaced array, such that different crystals are contacted with a different small molecule.
8. The method of any one of claims 1 to 6, comprising introducing a plurality of said small molecule into a plurality of crystals of said macromolecule at a stoichiometric ratio between said small molecule and said macromolecule.
9. The method of any one of claims 1 to 8, wherein the macromolecule is a protein.

10. A method of obtaining a microcrystal comprising a macromolecule and a small molecule from a crystal of the macromolecule, comprising
 - applying a first sample comprising the crystal of the macromolecule onto an electron microscopy (EM) grid;
 - contacting the crystal with a second sample comprising the small molecule to obtain a crystal comprising the macromolecule and the small molecule;
 - plunging the EM grid into a cryogenic liquid;
 - placing the EM grid into a focused ion beam scanning electron microscope (FIB/SEM); and
 - milling the crystal comprising the macromolecule and the small molecule in the FIB/SEM using a focused ion beam to obtain a microcrystal comprising the macromolecule and the small molecule.
11. The method of claim 10, wherein the EM grid has a carbon side and a copper side, and wherein both said applying the first sample and said contacting with the second sample are on the carbon side.
12. The method of claim 11, further comprising blotting the EM grid from the copper side after applying said first sample but before contacting with said second sample.
13. The method of claim 11 or 12, further comprising blotting the EM grid from the copper side after a contacting with said second sample.
14. The method of any one of claims 10 to 13, wherein the cryogenic liquid is liquid ethane.
15. The method of any one of claims 10 to 14, wherein the microcrystal obtained using said milling has at least one dimension that is at most 100, 150, 200, 250, 300, 350, or 400 nanometers.
16. The method of any one of claims 10 to 15, wherein the macromolecule is a protein.
17. The method of any one of claims 10 to 16, wherein said milling is under cryogenic conditions.

18. The method of any one of claims 10 to 17, further comprising depositing platinum on the crystal comprising the macromolecule and the small molecule by sputter coating the grid before said milling.
19. A method of determining a structural model for a complex comprising a macromolecule and a small molecule, comprising
- applying a first sample comprising a crystal of the macromolecule onto an electron microscopy (EM) grid;
 - contacting the crystal with a second sample comprising the small molecule to obtain a crystal comprising the macromolecule and the small molecule;
 - plunging the EM grid into a cryogenic liquid;
 - placing the EM grid into a focused ion beam scanning electron microscope (FIB/SEM);
 - milling the crystal comprising the macromolecule and the small molecule in the FIB/SEM using a focused ion beam to obtain a microcrystal comprising the macromolecule and the small molecule;
 - placing the microcrystal comprising the macromolecule and the small molecule into a transmission electron microscope (TEM);
 - collecting a plurality of electron diffraction patterns from the microcrystal; and
 - determining a structural model based on the plurality of electron diffraction patterns.
20. The method of claim 19, wherein said collecting a plurality of electron diffraction patterns comprises continuously rotating the EM grid in the TEM.
21. The method of claim 19 or 20, wherein said plurality of electron diffraction patterns are collected from a selected portion of the microcrystal.
22. The method of any one of claims 19 to 21, wherein the macromolecule is a protein.
23. The method of any one of claims 19 to 22, wherein said milling is under cryogenic conditions.

24. The method of any one of claims 19 to 23, further comprising depositing platinum on the crystal comprising the macromolecule and the small molecule by sputter coating the grid before said milling.
25. A method of identifying a small molecule that complexes with a macromolecule, comprising
- applying a first sample comprising a plurality of crystals of the macromolecule onto an electron microscopy (EM) grid;
 - contacting the crystals with a second sample comprising a plurality of small molecules on the EM grid to obtain a crystal comprising the macromolecule and the small molecule;
 - plunging the EM grid into a cryogenic liquid;
 - placing the EM grid into a focused ion beam scanning electron microscope (FIB/SEM);
 - milling the crystal comprising the macromolecule and the small molecule in the FIB/SEM using a focused ion beam to obtain a microcrystal comprising the macromolecule and the small molecule;
 - placing the microcrystal comprising the macromolecule and the small molecule into a transmission electron microscope (TEM);
 - collecting a plurality of electron diffraction patterns from the microcrystal;
 - determining a structural model for a complex of the macromolecule and the small molecule based on the plurality of electron diffraction patterns; and
 - characterizing interactions between the small molecule and the macromolecule based on the structural model, e.g., characterizing the small molecule as a small molecule that complexes with the macromolecule.
26. The method of claim 25, wherein said identifying is based on the occupancy values for the small molecule in the structural model.
27. The method of claim 25 or 26, wherein the macromolecule is a protein.

28. A method of screening a library of small molecules for their binding to a macromolecule, comprising identifying a small molecule that complexes with the macromolecule using the method of claim 25.

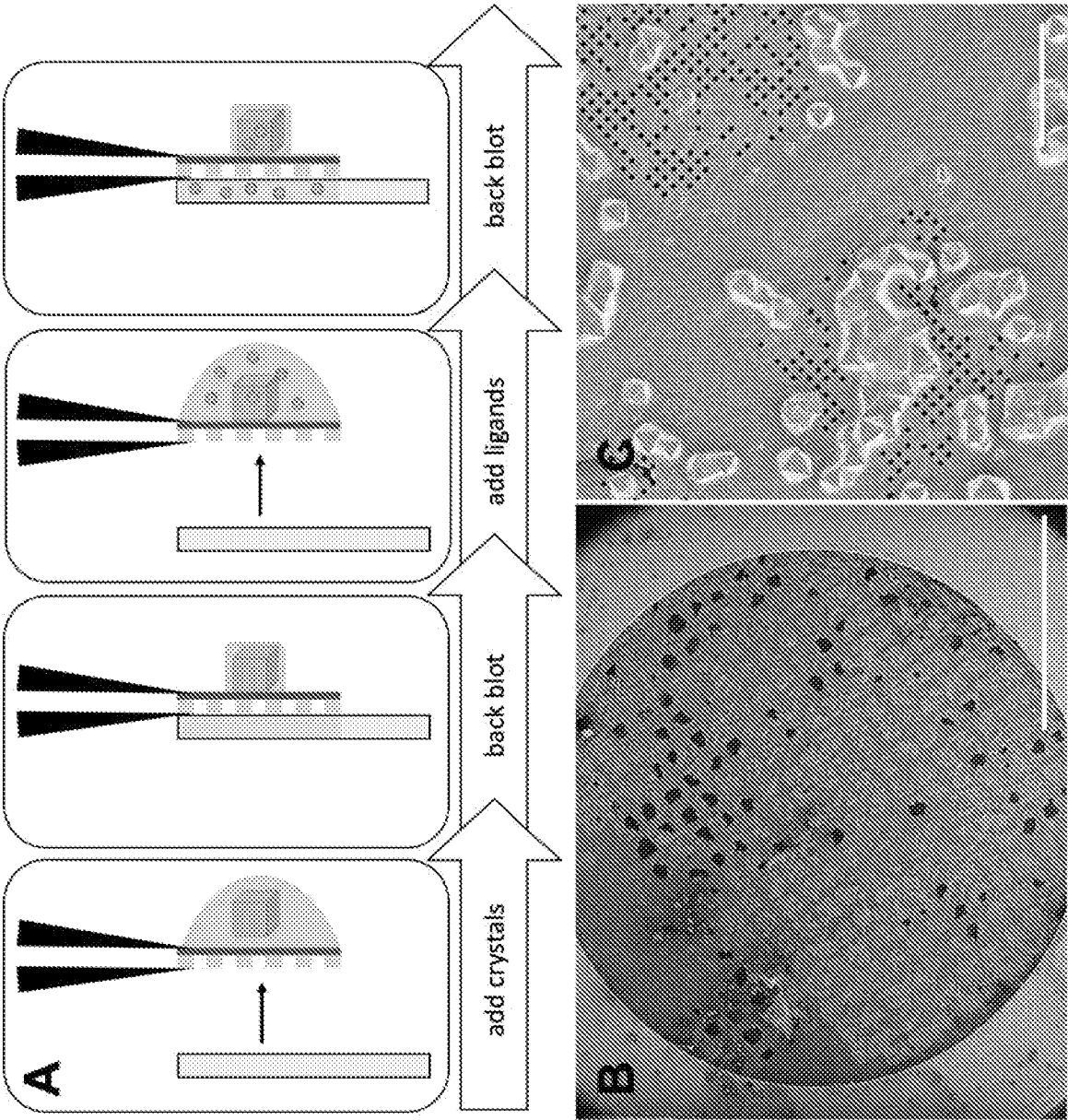


Figure 1

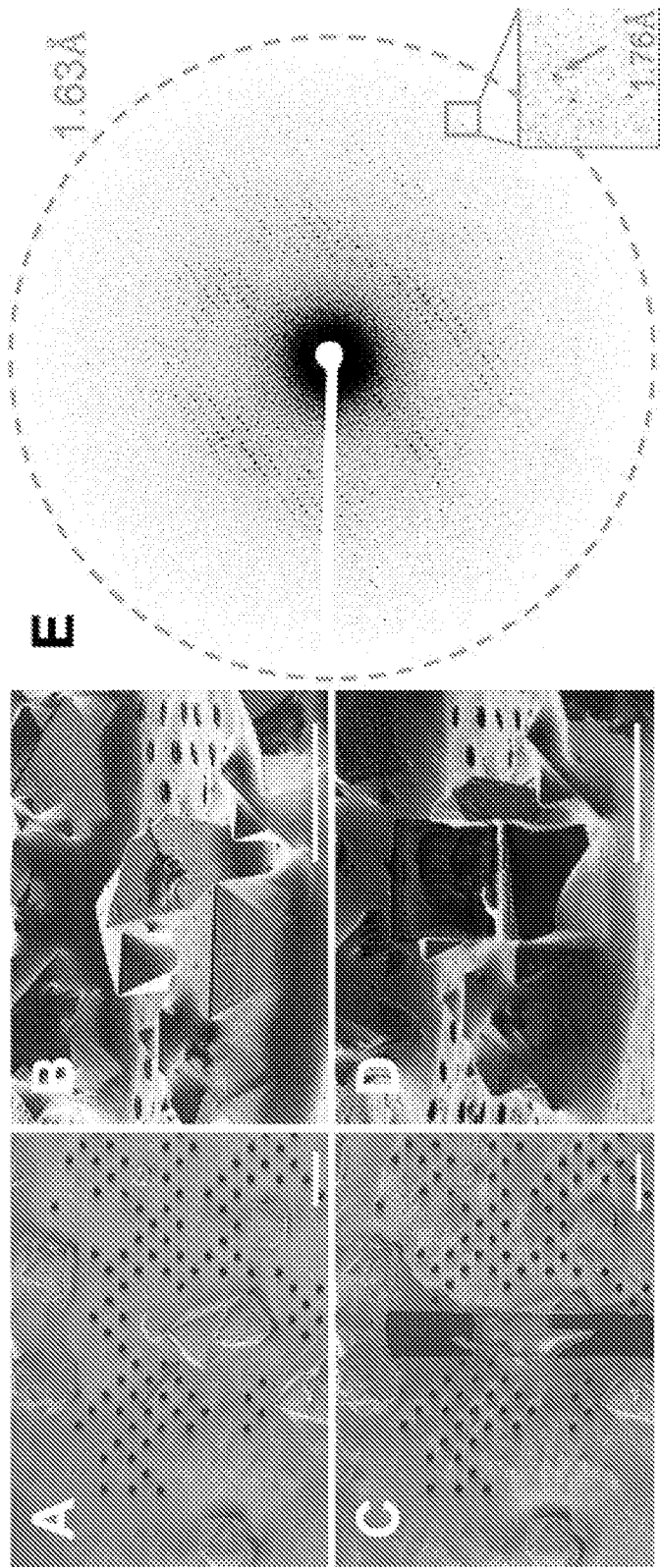


Figure 2

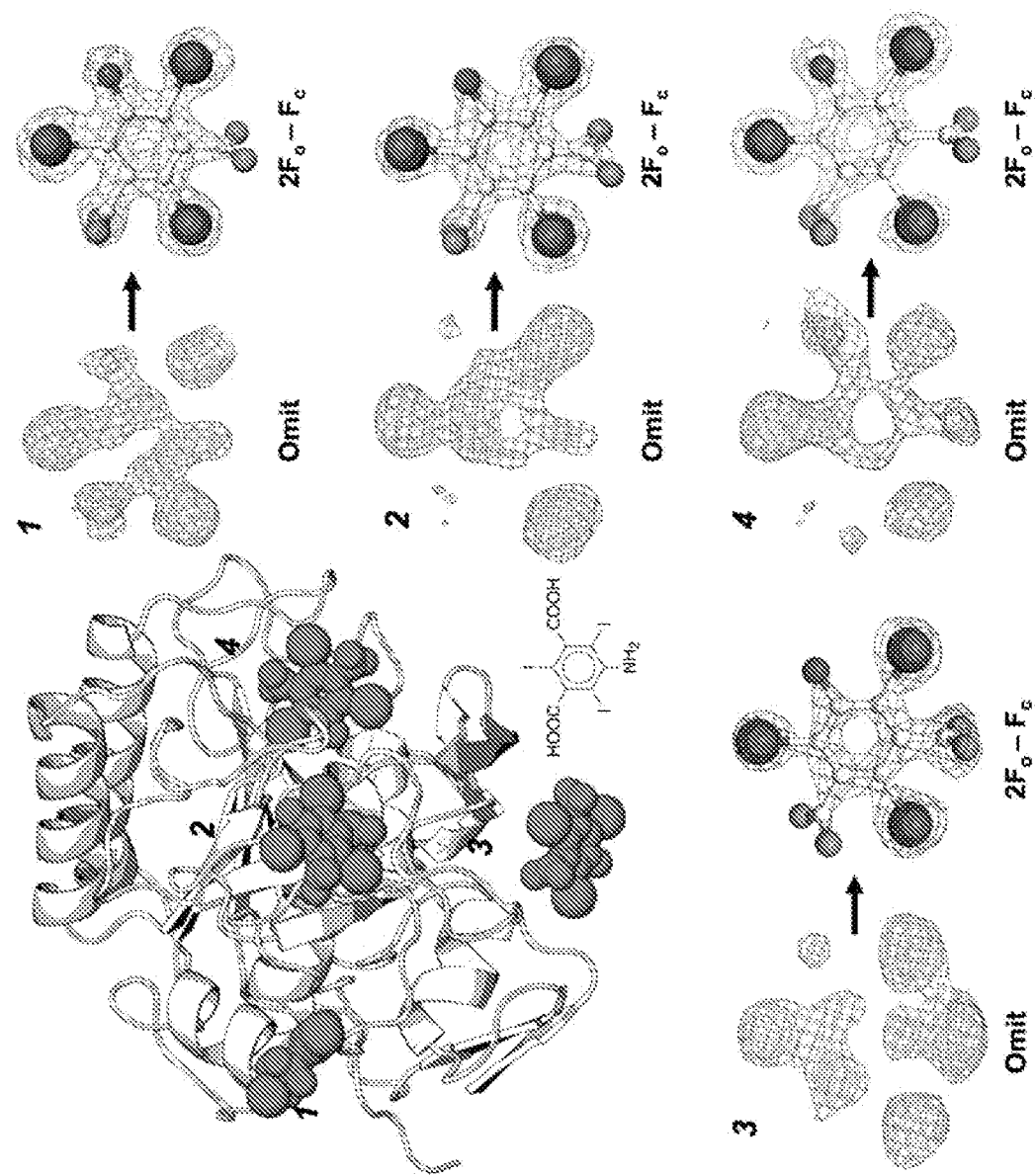


Figure 3

Figure 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2021/033701

A. CLASSIFICATION OF SUBJECT MATTER

See extra sheet.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC (20210101) G01N 23/225, G01N 23/00, B01L 99/00, H01J 37/26, H01J 37/285, G01N 1/42, G02B 21/00, G01N 21/77

CPC (20180201) G01N 23/225, G01N 23/00, B01L 99/00, H01J 37/26, H01J 37/285, G01N 1/42, G02B 21/002, G02B 21/00, G01N 2021/7783

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Databases consulted: Google Patents, Google Scholar, Derwent Innovation, Orbit, Similari (AI-based)

Search terms used: protein, small molecule, grid

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2016209907 A1 UNIV CALIFORNIA [US] 29 Dec 2016 (2016/12/29) [0009], [0023], [0025], [0031], [0034], [0041], [0085], Figure 1A	1-9
Y	[0009], [0023], [0025], [0031], [0034], [0041], [0085], Figure 1A	10-28
Y	M.W. Martynowycz et al., 'Collection of Continuous Rotation MicroED Data from Ion Beam-Milled Crystals of Any Size', Structure, Volume 27, Issue 3, 5 March 2019, Pages 545-548.e2, Available online 17 January 2019, https://doi.org/10.1016/j.str.2018.12.003 M.W. Martynowycz et al. 17 Jan 2019 (2019/01/17) p. 545 (right column), p. 546 (left column)	10-18
Y	M.W. Martynowycz et al., 'Qualitative analyses of polishing and pre-coating FIB milled crystals for MicroED', Structure 27, 1594–1600, October 1, 2019, Available online 15 August 2019, https://doi.org/10.1016/j.str.2019.07.004 M.W. Martynowycz et al. 15 Aug 2019 (2019/08/15) Summary, p. 6, 7	10-18
Y	p. 2, 6-8, Figure 2	19-28



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“D” document cited by the applicant in the international application

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

08 Sep 2021

Date of mailing of the international search report

09 Sep 2021

Name and mailing address of the ISA:

Israel Patent Office

Technology Park, Bldg.5, Malcha, Jerusalem, 9695101, Israel

Email address: pctoffice@justice.gov.il

Authorized officer

KORBAKOV Nina

Telephone No. 972-73-3927110

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2021/033701

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	C.G. Jones et al., 'The CryoEM Method MicroED as a Powerful Tool for Small Molecule Structure Determination', ACS Cent. Sci., 4, 1587?1592, 2 November 2018, DOI: 10.1021/acscentsci.8b00760 C.G. Jones et al. 02 Nov 2018 (2018/11/02) whole document	1-28
A	S.A. Arnold et al., 'Blotting-free and lossless cryo-electron microscopy grid preparation from nanoliter-sized protein samples and single-cell extracts', Journal of Structural Biology 197 (2017) 220–226, Available online 15 November 2016, http://dx.doi.org/10.1016/j.jsb.2016.11.002 S.A. Arnold et al. 15 Nov 2016 (2016/11/15) whole document	1-28

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/US2021/033701

Patent document cited search report	Publication date	Patent family member(s)	Publication Date
WO 2016209907 A1	29 Dec 2016	WO 2016209907 A1	29 Dec 2016
		US 2018108510 A1	19 Apr 2018
		US 10559449 B2	11 Feb 2020

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2021/033701

A. CLASSIFICATION OF SUBJECT MATTER:

IPC (20210101) G01N 23/225, G01N 23/00, B01L 99/00, H01J 37/26, H01J 37/285, G01N 1/42, G02B 21/00, G01N 21/77

CPC (20180201) G01N 23/225, G01N 23/00, B01L 99/00, H01J 37/26, H01J 37/285, G01N 1/42, G02B 21/002, G02B 21/00, G01N 2021/7783