

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

(19) World Intellectual Property  
Organization

International Bureau



(10) International Publication Number

WO 2022/221956 A1

(43) International Publication Date  
27 October 2022 (27.10.2022)

## (51) International Patent Classification:

*GI6B 15/00* (2019.01)      *GI6C 20/00* (2019.01)  
*GI6B 5/00* (2019.01)      *G01N 23/00* (2006.01)  
*GI6B 15/20* (2019.01)

## (21) International Application Number:

PCT/CA2022/050614

## (22) International Filing Date:

21 April 2022 (21.04.2022)

## (25) Filing Language:

English

## (26) Publication Language:

English

## (30) Priority Data:

63/201,260      21 April 2021 (21.04.2021)      US

(71) **Applicant:** STRUCTURA BIOTECHNOLOGY INC.  
[CA/CA]; 8 Crescent Town Road, Unit 412, Toronto, Ontario M4C 5L3 (CA).

(72) **Inventors:** PUNJANI, Ali; 412-8 Crescent Town Road, Toronto, Ontario M4C5L3 (CA). FLEET, David; 35 Playter Boulevard, Toronto, Ontario M4K2W1 (CA).

(74) **Agent:** BHOLE IP LAW; 130 Queens Quay East, Suite 1214, Toronto, Ontario M5A 0P6 (CA).

(81) **Designated States** (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM,

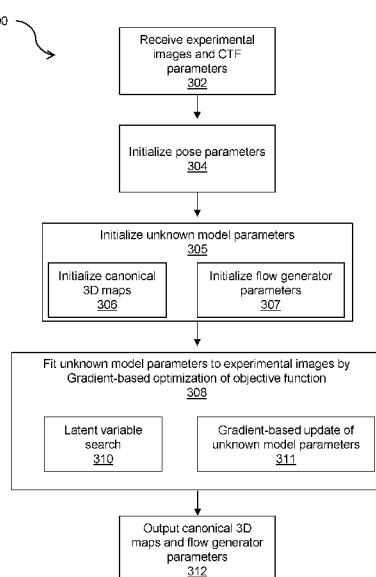
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, 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).

**Published:**

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

## (54) Title: METHODS AND SYSTEMS FOR RECONSTRUCTION OF THREE-DIMENSIONAL STRUCTURE AND THREE-DIMENSIONAL MOTION OF A PROTEIN MOLECULE



(57) **Abstract:** Provided are systems and methods for determining 3D structure and 3D motion of a protein molecule from 2D or 3D particle observation images. The method including: initializing pose parameters and unknown model parameters; the parameters of the one or more flow generators; image formation including: generating one or more 3D deformation fields by inputting the latent coordinate vector into the one or more flow generators; performing a convection and projection operation; and performing CTF corruption; fitting the unknown model parameters to the experimental images by gradient-based optimization of an objective function; latent variable search for a given experimental image including: performing the image formation one or more times to generate simulated images; selecting one or more latent coordinate vectors based on similarity; updating the at least one of the unknown model parameters including: generating simulated images; evaluating the objective function; computing the gradient of the objective function.

FIG. 3

1           METHODS AND SYSTEMS FOR RECONSTRUCTION OF THREE-DIMENSIONAL  
2           STRUCTURE AND THREE-DIMENSIONAL MOTION OF A PROTEIN MOLECULE

3    TECHNICAL FIELD

4 [0001] The following relates generally to electron Cryo-microscopy; and more specifically, to  
5 systems and methods for reconstruction of three-dimensional structure and three-dimensional  
6 motion of a protein molecule.

7    BACKGROUND

8 [0002] Proteins form the molecular machinery of the cell. They are inherently dynamic, often  
9 exhibiting a continuous landscape of energetically favorable conformations, with motion tightly  
10 coupled to function. Methods that uncover protein motion and the conformational landscape have  
11 the potential to illuminate fundamental questions in structural biology, and to enhance our ability  
12 to design therapeutic molecules that elicit specific functional changes in a target protein.  
13 Revealing protein dynamics is a frontier of structural biology, and in the absence of experimental  
14 techniques, molecular dynamics simulation and other approximate predictions have been widely  
15 adopted.

16 [0003] Single particle cryo-EM collects thousands of static two-dimensional (2D) protein  
17 particle images that, in aggregate, span the target protein's three-dimensional (3D)  
18 conformational space. Cryo-EM therefore is potentially able to experimentally uncover both the  
19 atomic-resolution structure and motion of biologically functional moving parts. Nevertheless,  
20 methods for resolving continuous motion and structure from static 2D images have remained  
21 elusive.

22    SUMMARY

23 [0004] In an aspect, there is provided a method for determining the three-dimensional (3D)  
24 structure and 3D motion of a molecule from two-dimensional (2D) or 3D experimental images, the  
25 3D structure in the form of one or more canonical 3D maps, and the 3D motion in the form of  
26 parameters of one or more flow generators, the method comprising: receiving the experimental  
27 images; receiving contrast transfer function (CTF) parameters for each experimental image;  
28 initializing pose parameters for each experimental image; initializing unknown model parameters  
29 of image formation, the unknown model parameters comprising: one or more canonical 3D maps;  
30 and the parameters of the one or more flow generators, each flow generator comprising a  
31 parameterized generator function taking as input a latent coordinate vector and outputting a 3D

1 deformation field; wherein the image formation taking as input at least a latent coordinate vector  
2 and outputting a simulated image, the image formation comprising: generating one or more 3D  
3 deformation fields by inputting the latent coordinate vector into the one or more flow generators;  
4 performing a convection and projection operation by convecting one or more of the canonical 3D  
5 maps by the one or more 3D deformation fields, and projecting using the pose parameters for the  
6 given experimental image; and performing CTF corruption of the projected result using the CTF  
7 parameters of the given experimental image to generate the simulated image; fitting the unknown  
8 model parameters to the experimental images by performing one or more iterations of gradient-  
9 based optimization of an objective function, the objective function taking as input at least  
10 simulated images and experimental images, wherein in at least one of the iterations, performing  
11 latent variable search for at least one experimental image, latent variable search for a given  
12 experimental image comprising: performing the image formation one or more times to generate  
13 simulated images from one or more latent coordinate vectors; and selecting one or more latent  
14 coordinate vectors based on the similarity between simulated images and the experimental  
15 image; and wherein in at least one of the iterations, updating at least one of the unknown model  
16 parameters using at least one of the experimental images, updating the at least one of the  
17 unknown model parameters comprising: generating simulated images by performing the image  
18 formation using the one or more selected latent coordinate vectors associated with the  
19 experimental images; evaluating the objective function using at least the simulated images and  
20 the experimental images; computing the gradient of the objective function with respect to the at  
21 least one unknown model parameter to be updated; and updating the unknown model parameter  
22 using the gradient; and outputting the one or more canonical 3D maps, the parameters of the one  
23 or more flow generators, or both.

24 [0005] In a particular case, the pose parameters for each experimental image are initialized  
25 by receiving them as input.

26 [0006] In another case, the experimental images and the simulated images are 2D.

27 [0007] In yet another case, the canonical 3D map is represented as a 3D voxel array of real-  
28 space density values.

29 [0008] In yet another case, the convection and projection operation comprises interpolating  
30 the canonical 3D map to form the convected and projected image, the interpolation operation  
31 defined by the deformation field and the pose parameters.

32 [0009] In yet another case, the flow generator is a feed-forward neural network.

- 1 [0010] In yet another case, a 3D deformation field is represented as deformed positions of  
2 vertices of a mesh of volumetric elements, the deformation vector field values defined by  
3 interpolation within the volume of each mesh element.
- 4 [0011] In yet another case, the objective function comprises at least the negative log-  
5 likelihood of a simulated image given an experimental image using a noise model.
- 6 [0012] In yet another case, the noise model is a Gaussian noise model
- 7 [0013] In yet another case, the latent variable search comprises coordinate descent search  
8 over the latent coordinate vector space.
- 9 [0014] In yet another case, the latent variable search comprises gradient descent search over  
10 the latent coordinate vector space.
- 11 [0015] In yet another case, the latent variable search comprises selecting latent coordinate  
12 vectors that are equal to the latent coordinate vectors determined to optimize the similarity plus a  
13 component of random noise.
- 14 [0016] In yet another case, each iteration of the gradient-based optimization comprises  
15 selecting a random subset of the experimental images.
- 16 [0017] In yet another case, updating at least one of the unknown model parameters using the  
17 gradient is performed using the Adam update rule.
- 18 [0018] In yet another case, updating at least one of the unknown model parameters using the  
19 gradient is performed using the Stochastic Gradient Descent update rule.
- 20 [0019] In yet another case, updating at least one of the unknown model parameters using the  
21 gradient is performed using the Stochastic Gradient Descent with momentum update rule.
- 22 [0020] In yet another case, updating at least one of the unknown model parameters using the  
23 gradient is performed using the L-BFGS update rule.
- 24 [0021] In yet another case, updating the canonical 3D map using the gradient is performed  
25 using the L-BFGS update rule.
- 26 [0022] In yet another case, the objective function comprises regularizing terms.
- 27 [0023] In yet another case, regularizing terms comprise terms that penalize non-rigid  
28 deformation in a 3D deformation field.
- 29 [0024] In yet another case, a 3D deformation field is represented as deformed positions of  
30 vertices of a mesh of volumetric elements, the deformation vector field values defined by

1 interpolation within the volume of each mesh element, and regularizing terms comprise terms that  
2 penalize non-rigid deformation within each mesh element.

3 [0025] In yet another case, the optimization of the unknown model parameters is performed  
4 using spatial frequency marching

5 [0026] In yet another case, the optimization of the unknown model parameters is first  
6 performed using a lower spatial frequency limit and subsequently performed using a higher spatial  
7 frequency limit.

8 [0027] These and other aspects are contemplated and described herein. It will be appreciated  
9 that the foregoing summary sets out representative aspects of systems and methods to assist  
10 skilled readers in understanding the following detailed description.

## 11 DESCRIPTION OF THE DRAWINGS

12 [0028] The features of the invention will become more apparent in the following detailed  
13 description in which reference is made to the appended drawings wherein:

14 [0029] FIG. 1 shows a system for reconstruction of three-dimensional structure and three-  
15 dimensional motion of a protein molecule, according to an embodiment;

16 [0030] FIG. 2 shows a computing environment of the system of FIG. 1;

17 [0031] FIG. 3 shows a method for reconstruction of three-dimensional structure and three-  
18 dimensional motion of a protein molecule, according to an embodiment;

19 [0032] FIG. 4 illustrates an example of a model to determine flexible 3D structure of a protein  
20 as deformations of a single canonical 3D density map;

21 [0033] FIG. 5 illustrates experimental results of the system of FIG. 1 with a  $K = 5$ -  
22 dimensional latent space on 102,500 particles of an snRNP Spliceosome complex

23 [0034] FIG. 6 illustrates experimental results of the system of FIG. 1 on 102,500 particles of  
24 an snRNP Spliceosome complex;

25 [0035] FIG. 7 illustrates experimental results of the system of FIG. 1 with a  $K = 5$ -  
26 dimensional latent space on 1002,500 particles of an snRNP Spliceosome complex;

27 [0036] FIG. 8 illustrates experimental results of the system of FIG. 1 with a  $K = 2$ -  
28 dimensional latent space on 200,000 particles of a TRPV1 ion channel protein;

- 1 [0037] FIG. 9 illustrates experimental results of the system of FIG. 1 with a  $K = 2$ -  
2 dimensional latent space on 200,000 particles of a TRPV1 ion channel protein;
- 3 [0038] FIG. 10A shows experimental results of a FSC of an entire density of the ion channel  
4 of the system of FIG. 1;
- 5 [0039] FIG. 10B shows experimental results of a FSC using a mask around one of the flexible  
6 peripheral domains in the soluble region of the ion channel; and
- 7 [0040] FIG. 10C shows experimental results of a mask in accordance with FIGS. 10A and  
8 10B.

9 DETAILED DESCRIPTION

10 [0041] For simplicity and clarity of illustration, where considered appropriate, reference  
11 numerals may be repeated among the Figures to indicate corresponding or analogous elements.  
12 In addition, numerous specific details are set forth in order to provide a thorough understanding  
13 of the embodiments described herein. However, it will be understood by those of ordinary skill in  
14 the art that the embodiments described herein may be practiced without these specific details. In  
15 other instances, well-known methods, procedures and components have not been described in  
16 detail so as not to obscure the embodiments described herein. Also, the description is not to be  
17 considered as limiting the scope of the embodiments described herein.

18 [0042] Various terms used throughout the present description may be read and understood  
19 as follows, unless the context indicates otherwise: “or” as used throughout is inclusive, as though  
20 written “and/or”; singular articles and pronouns as used throughout include their plural forms, and  
21 vice versa; similarly, gendered pronouns include their counterpart pronouns so that pronouns  
22 should not be understood as limiting anything described herein to use, implementation,  
23 performance, etc. by a single gender; “exemplary” should be understood as “illustrative” or  
24 “exemplifying” and not necessarily as “preferred” over other embodiments. Further definitions for  
25 terms may be set out herein; these may apply to prior and subsequent instances of those terms,  
26 as will be understood from a reading of the present description.

27 [0043] Any module, unit, component, server, computer, terminal, engine or device  
28 exemplified herein that executes instructions may include or otherwise have access to computer  
29 readable media such as storage media, computer storage media, or data storage devices  
30 (removable and/or non-removable) such as, for example, magnetic disks, optical disks, or tape.  
31 Computer storage media may include volatile and non-volatile, removable and non-removable  
32 media implemented in any method or technology for storage of information, such as computer

1 readable instructions, data structures, program modules, or other data. Examples of computer  
2 storage media include RAM, ROM, EEPROM, flash memory or other memory technology, CD-  
3 ROM, digital versatile disks (DVD) or other optical storage, magnetic cassettes, magnetic tape,  
4 magnetic disk storage or other magnetic storage devices, or any other medium which can be used  
5 to store the desired information and which can be accessed by an application, module, or both.  
6 Any such computer storage media may be part of the device or accessible or connectable thereto.  
7 Further, unless the context clearly indicates otherwise, any processor or controller set out herein  
8 may be implemented as a singular processor or as a plurality of processors. The plurality of  
9 processors may be arrayed or distributed, and any processing function referred to herein may be  
10 carried out by one or by a plurality of processors, even though a single processor may be  
11 exemplified. Any method, application or module herein described may be implemented using  
12 computer readable/executable instructions that may be stored or otherwise held by such  
13 computer readable media and executed by the one or more processors.

14 [0044] Established high-resolution cryo-EM refinement methods often assume rigidity of the  
15 target molecule, and result in blurred, low resolution density for flexible regions. Methods with  
16 spatially adaptive regularization mitigate the adverse effects of local motion on rigid refinement,  
17 but do not estimate the motion per se. Local refinement and multi-body methods use masks to  
18 focus on a sub-region of a protein, but only provide improved resolution on rigid parts of relatively  
19 large molecular weight. Subspace methods approximate a particle's space of conformation as a  
20 linear combination of basis density maps, but without an underlying concept of motion.

21 [0045] The development of a method that, in the presence of continuous flexibility, can  
22 uncover protein motion and thereby improve the resolution of fine structural details is a substantial  
23 technical problem. Effective solutions face several key challenges. First, there are a large number  
24 of unknowns that must be jointly estimated from the data, including the 3D structure of the density  
25 map, a representation of the space of conformational changes, and the position of each particle  
26 image on that conformational landscape. Second, the protein motion and the conformational  
27 landscape are generally non-linear. Third, in order to resolve 3D map details beyond what  
28 conventional (static) reconstructions can provide, structural information must be aggregated from  
29 many different conformations. Finally, despite the high levels of noise, and computational difficulty  
30 of the underlying optimization problem, the unknowns must be estimated with enough precision  
31 to enable recovery of high-resolution details.

32 [0046] The present embodiments advantageously provide an approach for 3D Flexible  
33 Refinement (referred to informally as '3DFlex'), a deep neural network model of continuously

1 flexible protein molecules. 3DFlex is a motion-based heterogeneity model that directly exploits  
2 the notion that conformational variability of a protein is a result of physical processes which tend  
3 to conserve mass and preserve the local geometry of the underlying molecular structure. This is in  
4 contrast to density-based techniques that model conformational variation as a manifold of 3D  
5 density maps without a physically plausible motion model that explicitly captures the transport of  
6 density from one location to another. Some approaches have also sought to model motion in  
7 limited forms. The formulation of 3DFlex is based on a generative architecture that captures  
8 conformational variability in terms of a single high-resolution canonical 3D density map of the  
9 molecule, and a parameterized latent space of deformation fields encoding non-rigid motion. The  
10 deformation fields "bend" the canonical density via convection, yielding all conformations captured  
11 by the model. In 3DFlex canonical density, the deformation field generator, and the latent  
12 coordinates of each particle image are jointly learned from the image data using a specialized  
13 training algorithm, without any prior knowledge about the flexibility of the molecule.

14 [0047] Results obtained with experimental cryo-EM data show that 3DFlex effectively  
15 addresses the challenges of flexible refinement. We show that the model can jointly learn the  
16 structure of the flexible molecule, the underlying non-linear, non-rigid motion that unites its  
17 conformational landscape, and the positions of each single particle image on that landscape.  
18 Given a dataset of tri-snRNP spliceosome particles, 3DFlex learns non-rigid motions that range  
19 from a  $\alpha$ -helices and  $\beta$ -sheets moving fractions of an Angstrom to the large motions of sub-units  
20 bending across a span of 20+ Å. In doing so, it aggregates structural information from all  
21 conformations into a single, optimized density map where high-resolution features of flexible  
22 domains are well-resolved. In fact, 3DFlex can model continuous motion with enough precision  
23 to improve the resolution of small flexible parts that are otherwise poorly resolved in both  
24 conventional and local focused refinements; which the experiments demonstrate on a dataset of  
25 TRPV1 ion-channel particles; where 3DFlex improves FSC resolution and map quality of  
26 peripheral  $\alpha$ -helices in the flexible soluble domains.

27 [0048] The present embodiments provide an approach to determine both the structure and  
28 motion of flexible particles, and effectively use the motion to enhance map detail. This  
29 determination allows for the study of biological mechanisms and function involving motion, at the  
30 frontier of both cryo-EM and structural biology.

31 [0049] The present embodiments can have several possible variations, as understood by a  
32 person skilled in the art and should not be limited to variations specifically described herein. The  
33 present embodiments provide a substantial improvement in the field of structural biology of protein

1 molecules, which generally requires simultaneous consideration of hundreds of thousands of non-  
2 trivial numerical representations of 3D protein structures with high precision.

3 [0050] The present embodiments provide a generative neural network method that  
4 determines the structure and motion of flexible protein molecules from input cryo-EM images of  
5 individual molecules, referred to herein as experimental image or a particle image. The  
6 experimental images can be two dimensional (2D) or three dimensional (3D) images resulting  
7 from, in some instances, single-particle cryo-electron microscopy, and in some other instances  
8 cryo-electron tomography; though in general they can result from any type of particle imaging  
9 approach. Generally, multiple conformations of a dynamic protein are related to each other  
10 through deformation of a single 3D structure. Specifically, the maps for flexible molecule are  
11 represented in terms of i) a canonical 3D map, ii) latent coordinate vectors that specify positions  
12 over the protein's conformational landscape, and iii) a flow generator that converts a latent  
13 coordinate vector into a deformation field; which converts the canonical map into the  
14 corresponding protein conformation. The unknown model parameters include the canonical 3D  
15 map, the parameters of the flow generator, and the latent coordinate vector for each particle  
16 image. These parameters can be jointly learned from the input images by performing gradient-  
17 based optimization.

18 [0051] Referring now to FIG. 1, a system 100 for reconstruction of three-dimensional structure  
19 and three-dimensional motion of a protein molecule, in accordance with an embodiment, is  
20 shown. The system 100 can be executed on a suitable computing device; for example, a desktop  
21 computer, a laptop computer, a server, or the like.

22 [0052] FIG. 1 shows various physical and logical components of an embodiment of the  
23 system 100. As shown, the system 100 has a number of physical and logical components,  
24 including a central processing unit ("CPU") 102, random access memory ("RAM") 104, an input  
25 interface 106, an output interface 108, a network interface 110, non-volatile storage 112, and a  
26 local bus 114 enabling CPU 102 to communicate with the other components. CPU 102 executes  
27 various modules 120, as described below in greater detail. RAM 104 provides relatively  
28 responsive volatile storage to CPU 102. The input interface 106 enables an administrator or user  
29 to provide input via an input device, for example a keyboard and mouse. The output interface 108  
30 outputs information to output devices, such as a display and/or speakers. The network interface 110  
31 permits communication with other systems, such as other computing devices and servers  
32 remotely located from the system 100, such as for a typical cloud-based access model. Non-  
33 volatile storage 112 stores computer-executable instructions for implementing the modules, as

1 well as any data used by these services. Additional stored data can be stored in a database 116.  
2 During operation of the system 100, the modules and the related data may be retrieved from the  
3 non-volatile storage 112 and placed in RAM 104 to facilitate execution.

4 [0053] In an embodiment, as described in more detail in the following, the system 100  
5 includes various modules 120; including an inputs module 122, a reconstruction module 124, and  
6 an output module 126. In further cases, the functions of some or all of the various modules 120  
7 can be combined or performed on other modules, or can be performed on dedicated pieces of  
8 hardware. In some cases, some or all of the various modules 120 can be executed on a server-  
9 side device 32 or a client-side device 26 (shown in FIG. 2), and be in communication with the  
10 other modules. An imaging system 130 may further be linked to the system 100 to obtain cryo-  
11 EM images. The imaging system 130 generally comprises one or more electron microscopes, or  
12 other suitable devices.

13 [0054] In some cases, as shown in a diagram of a computing environment 10 in FIG. 2, the  
14 system 100 can communicate with, and retrieve data, from other computing devices; for example,  
15 from the server 32 to the client computing device 26. The system 100 can communicate with  
16 these devices over a data communication network; for example, the Internet 24.

17 [0055] Turning to FIG. 4, shown is a flowchart for a method 200 to simulate the formation of  
18 cryo-EM images containing three-dimensional structure and three-dimensional motion of a protein  
19 molecule, in accordance with an embodiment. The method 200 can be referred to as ‘image  
20 formation’.

21 [0056] FIG. 4 illustrates an example of image formation in the 3DFlex model that models the  
22 flexible 3D structure of a protein as deformations of a canonical 3D density map  $V$ . Under the  
23 model, a single particle image is associated with a low-dimensional latent coordinate  $z$  that  
24 encodes the particular conformation for the particle in the image. A neural flow generator network  
25  $f_\theta$  converts the latent coordinate into the flow field  $u$  and a convection operator then deforms  
26 the canonical density to generate a convected map  $W$ . It can then be projected along the particle  
27 viewing direction, and contrast transfer function (CTF)-corrupted to generate a simulated image.  
28 The simulated image can be compared against the experimental image.

29 [0057] At block 210, the inputs module 122 receives  $K$ -dimensional latent coordinates  $z_i$  of  
30 a particle under investigation in a received cryo-EM image. The cry-EM image can be received  
31 from the database 116, the network interface 110, or the imaging system 130 via the input  
32 interface 106.

1 [0058] At block 211, the inputs module 122 feeds the latent coordinates  $z_i$  to a flow generator  
 2  $f_\theta(z_i)$  of the reconstruction module 124. This provides a 3D deformation field, denoted  $u(x)$ ,  
 3 where  $x$  is a 3D position and  $\theta$  denotes the parameters of the generator.

4 [0059] At block 213, the deformation vector field and a canonical 3D density  $V$  are input by  
 5 the reconstruction module 124 to a convection operator, denoted  $D(u_i, V)$ , which outputs a  
 6 convected density, denoted  $W_i$ .

7 [0060] At block 215 and block 217, the reconstruction module 124 then determines a 2D or  
 8 3D simulated particle image  $I_i$  as a contrast transfer function (CTF)-corrupted projection of  $W_i$ ,  
 9 plus additive noise  $\eta$ ; i.e.:

$$\begin{aligned} 10 \quad I_i &= C_i P(\phi_i) W_i + \eta \\ 11 \quad &= C_i P(\phi_i) D(f_\theta(z_i), V) + \eta \end{aligned} \quad (1)$$

12 Here,  $C_i$  denotes the CTF operator and  $P(\phi_i)$  is the projection operator for the pose  $\phi_i$ , which  
 13 specifies the transformation between the microscope coordinate frame and the coordinate frame  
 14 of the canonical map. The output of this block 217 is a simulated image, to be compared, at block  
 15 218, with an experimental image.

16 [0061] Turning to FIG. 3, shown is a flowchart for a method 300 to reconstruct the three-  
 17 dimensional structure and three-dimensional motion of a protein molecule from input experimental  
 18 images. At block 302, the input module 122 receives experimental images and CTF parameters.  
 19 Experimental images and CTF parameters can be received from an imaging system 130 via the  
 20 input interface 106, the database 116, and/or the network interface 110. At block 304, the  
 21 reconstruction module initializes pose parameters for each experimental image. At block 305, the  
 22 reconstruction module initializes unknown model parameters; including, at block 306, initializing  
 23 canonical 3D maps and, at block 307, initializing flow generator parameters. At block 308, the  
 24 reconstruction module 124 fits the unknown model parameters to experimental images by  
 25 optimizing an objective function. As described herein, this fitting at block 308 includes, over one  
 26 or more iterations, at block 310, a latent variable search and, at block 311, a Gradient-based  
 27 update of unknown model parameters. In an embodiment, the optimization optimizes the flow  
 28 generator parameters  $\theta$ , the canonical 3D map  $V$ , and the latent coordinates  $z_i$ , in order to  
 29 maximize the likelihood of the experimental data under the probabilistic model of image formation  
 30 (Equation (1)). This is equivalent to minimizing the negative log-likelihood:

$$31 \quad E_{\text{data}}(V, \theta, z_{1:M}) = \frac{1}{2} \sum_{i=1}^M \| I_i - C_i P(\phi_i) D(f_\theta(z_i), V) \|^2 \quad (2)$$

1 where  $M$  is the number of particle images. For notational simplicity, it can be assumed that it is  
2 additive white noise; however the formulation can be extended to handle colored noise as  
3 understood by a person of skill in the art. In some cases, it can be assumed that poses  $\phi_i$  and  
4 CTF estimates are known from input, for example from a standard cryo-EM refinement algorithm,  
5 though these both could also be re-optimized. Note that the particular form of the noise and  
6 corresponding error function Equation (2) can be changed as suitable because any suitable noise  
7 model can be used, or equivalently, any suitable error function defining the data term  $E_{\text{data}}$  can  
8 be used.

9 [0062] At block 312, the output module 126 outputs the canonical 3D map and the flow  
10 generator parameters to the output interface 108, the database 116, or the network interface 110.  
11 In some cases, it also outputs the latent coordinate vectors for each experimental image. The 3D  
12 canonical map contains the reconstructed 3D structure of the protein molecule, and the flow  
13 generator parameters contain the reconstructed 3D motion of the protein.

14 [0063] Within the above formulation, there are several important design choices that define  
15 the architecture of the model. Solving structure and motion from noisy cryo-EM data is a  
16 challenging problem. As such, discussion of the design choices below provides insight into the  
17 working model, reflecting the present inventors' extensive experimental iterations to arrive at the  
18 present invention.

19 [0064] In a particular case, the reconstruction module 124 can use a fully-connected deep  
20 neural network with ReLU activations for the flow generator. The input  $z$  is the low-dimensional  
21 latent coordinate vector for each image, and the output is a 3D flow field  $u(x)$ . The number of  
22 hidden units per layer and the number of layers are adjustable hyperparameters. The final layer  
23 is a linear (without biases or nonlinear activation).

24 [0065] The neural flow generator gives the system 100 the capacity to learn complex,  
25 nonlinear deformation fields from data, and the inherent inductive bias of the architecture helps  
26 avoid over-fitting. Nevertheless, the data is noisy and the number of network parameters is large.  
27 Explicit regularization therefore plays a role in reducing the risk of over-fitting; as described herein.  
28 Note that the method described herein is general with respect to the form of the flow generator  
29 function. A neural network of any suitable architecture, or any other suitable function approximator  
30 with adjustable parameters  $\theta$ , can be used. In particular, neural implicit functions, where the flow  
31 generator function does not output an entire flow field but rather only the flow at a specified input  
32 coordinate  $u(x) = f_\theta(z, x)$ , are another suitable family of flow generators.

1 [0066] The latent space represents a conformational landscape, as different latent positions  
2 correspond to different deformations of the canonical map. The system 100 determines a latent  
3 state (or embedding) for each input image. In probabilistic terms, given an image  $I$ , the goal is to  
4 infer the posterior distribution  $p(z | I)$ , such that the high probability latent states are those for  
5 which the flow generator and canonical map explain the image well (i.e., minimizing Equation (2)).

6 [0067] Determining the exact posterior distribution is generally intractable for problems such  
7 as those solved by the system 100, so instead the reconstruction module 124 uses approximate  
8 inferences. One approach, that can be used in variatioinal auto-encoders (VAE), is so-called  
9 *amortized variational inference*, in which a feed-forward neural network (the encoder) is used to  
10 approximate the posterior for any given image; e.g., from the input image it computes the mean  
11 and covariance over latent coordinates. This approach has been used by deep-learning based  
12 heterogeneity methods. In the present context, the amortized variational inference can be trained  
13 with the flow generator and the canonical map to maximize the likelihood of the particle images.  
14 This approach is typically fast and stable to train. As with VAEs, it incorporates a prior over latent  
15 positions that helps to regularize the structure of the latent space to be smooth, mitigating the  
16 risks of over-fitting.

17 [0068] The challenge with amortized inference is that it can be extremely difficult for one  
18 neural network to do a good job approximating the posterior. In the context of cryo-EM, the  
19 encoder network has to effectively invert the decoder network, requiring information about the  
20 protein, its motion, and the image formation process. For example, when the flow generator shifts  
21 a particular subunit up or down, the encoder must simultaneously learn the same thing in order  
22 to determine the best latent states for a given image. This is difficult given image noise, and the  
23 lack of explicit access to the canonical density and the 3D deformation field. In general, amortized  
24 inference with an encoder has not been found to be sufficiently precise to resolve high-resolution  
25 structure and motion.

26 [0069] An alternative, potentially more accurate approach is to perform inference of latent  
27 coordinate vectors by applying variational optimization individually for each input image. This  
28 entails  $M$  optimization problems, in each case finding the optimal approximate posterior over  
29 latent coordinate vectors for each experimental image. The task of each optimization problem is  
30 to find latent coordinate vectors that, through the image formation process, produce simulated  
31 images that are similar to the experimental image (i.e. minimizing Equation (2)). Herein, each of  
32 these optimization problems is referred to as latent variable search. Latent variable search takes  
33 as input the experimental image and uses the image formation process to infer one or more

1 inferred estimates of the latent coordinate vector associated with this experimental image.  
2 Although computationally more expensive than amortized inference with an encoder network,  
3 inference by latent variable search is more precise and learning naturally incorporates information  
4 across images and viewing directions, capturing structure and motion with sufficient fidelity to  
5 resolve flexible protein regions to higher resolution than other approaches.

6 [0070] The use of latent variable search is characterized by the use of the image formation  
7 process to measure the similarity of simulated images against the experimental image, and to use  
8 this similarity measurement to find latent coordinate vectors that optimize the similarity. This  
9 contrasts with methods that use an encoder (such as the VAE) where inference does not involve  
10 performing image formation and is instead done in a separate predictive step. Multiple different  
11 types of latent variable search can be used in embodiments of the present method, an illustrative  
12 selection of which are described herein. The similarity measure of simulated and experimental  
13 images used in latent variable search can also vary. In some embodiments, the similarity measure  
14 can be chosen to be the same as the negative log-likelihood in Equation (2).

15 [0071] In some embodiments, latent variable search can use an auto-decoder model, in which  
16 it directly optimizes a point estimate of the latent coordinate vector for a given image, taking  
17 advantage of the structure of the generative model. This entails performing latent variable search  
18 by individually optimizing the latent coordinate vector for each experimental image so as to explain  
19 the experimental image well (i.e. minimizing Equation (2)). It can also be viewed as variational  
20 inference with a Gaussian variational family in the limit as the variance tends to 0.

21 [0072] Because the system 100 uses an end-to-end differentiable generative model, it can  
22 compute gradients of the data likelihood with respect to the latent coordinate vectors for each  
23 image, and then use gradient-based optimization in order to perform latent variable search. When  
24 the dimensionality  $K$  of the latent space is small enough, it is also possible to use coordinate-  
25 descent in order to perform latent variable search. The latter approach was found to be simpler  
26 and equally effective in the experiments.

27 [0073] One benefit of explicitly modeling the posterior distribution  $p(z | I)$ , rather than a point  
28 estimate for  $z_i$  given  $I_i$ , is that the learned latent representations are often more meaningful. As  
29 with VAE training, the likelihood of the observed images should be conditioned on samples from  
30 the posterior. Uncertainty in  $p(z | I)$  means that samples of  $z$  in the vicinity of the mean should  
31 yield accurate reconstructions of the data. This tends to regularize the model, encouraging  
32 smoothness of the latent representation.

1 [0074] While such regularization does not occur naturally with a point estimate of  $z_i$ , there is  
2 a heuristic that produces a similar positive effect, by directly adding noise to the point estimate  
3 during latent variable search. This can be likened to variational inference with a Gaussian  
4 variational family with a fixed covariance, and can be used to regularize deterministic auto-  
5 encoders. The reconstruction module 124 lets the variance of the injected noise be determined  
6 by the distance of a given latent point to other nearby embedded points. In addition to noise  
7 injection, the reconstruction module 124 can use a Gaussian prior on latent coordinates with unit  
8 variance to help control the spread of the latent embedding for a given dataset, and to center it at  
9 the origin in the latent space.

10 [0075] Note that for the method described herein, noise injection and the particular prior over  
11 latent coordinates are not strictly necessary. The method can use any suitable technique for  
12 smoothing the latent space and accounting for uncertainty in latent inference. For example, it is  
13 suitable to compute multiple samples from the posterior and use an importance sampling or  
14 variational method to draw samples from the resulting approximation to the posterior in order to  
15 perform latent variable search.

16 [0076] Algorithms for single-particle reconstruction commonly represent 2D and 3D maps and  
17 2D and 3D images in the Fourier domain. This reduces the computational cost of the CTF-  
18 corruption operation and image projection (via Fourier-slice theorem). It also allows maximum-  
19 likelihood 3D reconstruction with known poses in closed-form, as is familiar to those skilled in the  
20 art. On the other hand, the convection of density between conformations is more naturally  
21 formulated as a real-space operation. Features and structures in the canonical density map  $V$   
22 need to be shifted, rotated, and potentially deformed to produce densities consistent with the  
23 observed particles.

24 [0077] The reconstruction module 124 can represent the canonical map  $V$  in real-space, as  
25 a voxel array of size  $N^3$ . Convection and projection can be performed in real-space, and in  
26 practice are combined into a single operator that does not store  $W_i$  explicitly. Once the projected  
27 image of the convected map is formed, it is transformed to Fourier-space, CTF-corrupted, and  
28 transformed back to real-space to be used with the observed image for likelihood computation.  
29 Interestingly, the reconstruction module 124 can also find that 3D reconstruction of the high  
30 resolution canonical map is also possible in real-space using suitable optimization techniques.  
31 Computationally, real-space convection and projection are far more expensive than Fourier-space  
32 slicing, and the FFT for CTF modulation must be applied for every image in the forward pass, and

1 also in the backwards pass for computing gradients. Nevertheless real-space reconstruction is  
2 effective, as demonstrated herein.

3 [0078] Note that any suitable representation of the canonical map can be used. The present  
4 disclosure uses a real-space representation but it is equally appropriate to use any other  
5 representation. For example, neural implicit functions that model the density as a function of  
6 spatial position  $V(x) = g_\psi(x)$  where  $g$  is a neural implicit function and  $\psi$  are its parameters.  
7 Any other representation can be used, for example a wavelet basis, or a learned or optimized  
8 basis.

9 [0079] Convection of density is used to model the physical nature of protein motion, thereby  
10 allowing high-resolution structural detail from experimental data to backpropagate through the  
11 model. There are several ways to construct a convection operator. One way is to express the flow  
12 field as a mapping from convected coordinates (i.e., of voxels in  $W_i$ ) to canonical coordinates.  
13 Convection then requires interpolating the canonical density  $V$  at positions specified by the flow  
14 field. However, in order to maintain conservation of mass the interpolated density must be  
15 modulated by the determinant of the Jacobian of the mapping, which can be challenging to  
16 compute and differentiate.

17 [0080] Instead, the flow,  $u_i(x)$ , represents a forward mapping from canonical coordinates in  
18  $V$  to the deformed coordinates in  $W_i$ . This naturally conserves density, as every voxel in  $V$  has  
19 a destination in  $W_i$  where its contribution is accumulated through an interpolant function. The  
20 convected density at  $x$  can be written as:

21 
$$W_i(x) = \sum_y k(x - u_i(y))V(y) \quad (3)$$

22 where  $u_i = f_\theta(z_i)$ , and  $k(x)$  is an interpolation kernel with finite support. In this case, divergence  
23 and convergence of the flow field must be treated carefully to avoid undesirable artifacts such as  
24 holes, Moiré patterns, and discontinuities. It was determined to be advantageous to use high-  
25 order (e.g., tricubic) interpolation and strong regularization (described herein) to ensure accurate  
26 interpolation and artefact-free gradients.

27 [0081] Note that any suitable convection operator, that can be differentiated, can be used.

28 [0082] As capacity is added to the model, the propensity for over-fitting becomes problematic  
29 without well designed regularization. Over-fitting can result in the formation of localized, high-  
30 density points ("blips") in the canonical map, along with flow fields that translate these aberrations  
31 by large distances to explain noise in the experimental images. This can be especially pronounced

1 with smaller proteins, higher levels of image noise, and membrane proteins containing disordered  
 2 micelle or nanodisc regions (i.e., structured noise). Over-fitting also occurs when the  
 3 regularization is not strong enough to force the model to separate structure from motion. For  
 4 example, rather than improve the canonical density with structure common to all conformations,  
 5 the model sometimes learned to deform a low-resolution canonical density to create high-  
 6 resolution structure (with highly variable local deformations).

7 [0083] To address such issues, the reconstruction module 124 exploits prior knowledge of  
 8 smoothness and local rigidity in the deformation field. In particular, it is unlikely that natural  
 9 deformations would involve large discontinuities in regions of high density; e.g., an  $\alpha$ -helix should  
 10 not be sheared into disjoint pieces. It is also unlikely that formations will be highly non-rigid at fine  
 11 scales in regions of high density; at the extreme, bond lengths should not stretch or compress  
 12 substantially. While simple regularizers can be used, like limiting the frequency content of the flow  
 13 field, or penalizing its curvature, these are difficult to tune and do not prevent over-fitting reliably.

14 [0084] The reconstruction module 124 can instead model flow generation using finite-element  
 15 methods. A tetrahedral mesh covering regions of high density is generated in the canonical frame,  
 16 based on a preliminary consensus refinement. The deformation field is parameterized by a 3D  
 17 flow vector at each vertex of the tetrahedral mesh. The deformation field is then interpolated using  
 18 linear FEM shape functions within each mesh element. Smoothness is enforced implicitly through  
 19 interpolation as a function of the size of the mesh elements, which is an adjustable parameter,  
 20 and the fact that adjacent elements share vertices.

21 [0085] Local rigidity of the flow is also encouraged in each mesh element. In more detail, the  
 22 deformation field within the  $j$ th tetrahedral element for image  $i$ , denoted  $u_{ij}(x)$  can be written  
 23 as a linear mapping:

$$24 \quad u_{ij}(x) = A_{ij}x + b_{ij} \quad (4)$$

25 where matrix  $A$  and vector  $b$  are uniquely determined from 3D flow vectors at the element  
 26 vertices. Local non-rigidity is quantified in terms of the distance between  $A$  and the nearest  
 27 orthogonal matrix (in a MSE sense). In particular, the reconstruction module 124 measures the  
 28 squared deviation of the singular values of  $A$  from unity. Letting  $s_{ij}^\ell$  be the  $\ell$ th singular value of  
 29  $A_{ij}$ , we express the local rigidity regularization loss as:

$$30 \quad E_{\text{rigid}} = \sum_i \sum_j w_j \sum_{\ell=1}^3 (s_{ij}^\ell - 1)^2 \quad (5)$$

1 where  $w_j$  are weights defining the strength of the prior within each mesh element, based on the  
2 density present within the  $j$ th mesh element. The densest elements have weight 1. Empty  
3 elements have weight 0. This ensures that deformation fields are permitted to compress and  
4 expand empty space around the protein.

5 [0086] Note that for the method described herein, regularization using a tetrahedral mesh is  
6 not the only suitable regularization. Other regularizers can be used, such as those that limit the  
7 Fourier-space content of deformation fields, penalize curvature, divergence, and/or curl of the  
8 deformation fields. It is also suitable to use any other regularizer, for example a learned regularizer  
9 that computes the penalty as the output of a particular function on deformation fields. For example,  
10 a suitable regularizer may use an atomic or pseudoatom model of the target protein to define  
11 regions that should be continuous and rigid and interfaces where shear, expansion, and  
12 contraction can be allowed. The regularizer may be in the form of a penalty or direct constraints  
13 on the flow generator network. The regularizers may operate on any or all of the variables,  
14 including the canonical density, the flow generator, and the latent variables.

15 [0087] Image formation 200 is end-to-end differentiable, so in reconstruction method 300, at  
16 block 308, gradient-based optimization can be used to optimize the unknown model parameters  
17 of the flow generator the canonical map that best explains the experimental images. Any type of  
18 gradient-based optimization that uses the gradient to compute and apply an update to the  
19 parameters can be used. It may be advantageous to use either Adam or Stochastic Gradient  
20 Descent (SGD) with Nesterov acceleration, with minibatches of size at least 500 because of the  
21 high levels of image noise. In a given iteration of gradient-based optimization, inference of the  
22 latent coordinate vectors (i.e. latent variable search) for each experimental image in a minibatch  
23 can be performed prior to computing gradients with respect to the canonical density and flow  
24 parameters, or values of inferred latent coordinate vectors from previous iterations or initialization  
25 can be used prior to computing the gradients, or only latent variable search can be performed and  
26 gradient-based updating can be skipped in some iterations. The loss function being optimized in  
27 gradient-based optimization can be a weighted sum of the data log likelihood (Equation (2)) and  
28 the non-rigidity penalty (Equation (5)):

$$29 \quad L = E_{\text{data}} + \lambda_{\text{rigid}} E_{\text{rigid}} \quad (6)$$

30 [0088] During optimization, the reconstruction module 124 can use spatial frequency  
31 matching, learning the model in a coarse-to-fine manner. The canonical density  $V$  is constrained  
32 to be low-pass, with an upper frequency band-limit which increases over iterations. The frequency

1 and learning rate schedule, and  $\lambda_{\text{rigid}}$ , must be tuned for each dataset in our current  
2 implementation. Optimization is done with a box size,  $N = N_L$ , that is typically smaller than the  
3 raw size of the particle images, i.e.,  $N_H$ . As such, optimization only uses information below the  
4 Nyquist frequency for the smaller box size, and therefore is limited in spatial frequency to a lower  
5 limit. Optimization can also be run at the full box size, therefore being limited to a higher spatial  
6 frequency limit. During optimization, it is also possible to apply a real-space mask to the flow  
7 generator output to ensure the deformation is zero except in a region of interest; e.g., to exclude  
8 motion of a micelle or nanodisc.

9 [0089] To initialize training, the canonical density  $V$  can be set to be a low-pass filtered  
10 version of a consensus refinement given the same particles or can be initialized from another  
11 suitable source. The parameters of flow generator can be randomly initialized. The latent  
12 coordinates for the particles are either initialized to zero or to the output of another heterogeneity  
13 embedding method. In some experiments, especially on smaller, low SNR particles, it was  
14 determined that initializing with latent coordinates from 3D Variability Analysis (3DVA) in  
15 cryoSPARC improves results.

16 [0090] It was determined that simultaneously training of the canonical density  $V$  and flow  
17 generator parameters  $\theta$  leads to over-fitting after thousands of gradient iterations, despite strong  
18 regularization. In these cases,  $V$  and  $\theta$  can be initially trained with latent coordinates  $z_i$  fixed  
19 to their initial values from 3DVA for 5 epochs. Then, the latent coordinates are unlocked and latent  
20 inference is performed for each minibatch while updating  $\theta$ , but with  $V$  fixed, for 10 epochs.  
21 Then, alternating updating  $V$  with  $\theta$  fixed for 10 epochs, and repeated until convergence. In  
22 general, any schedule of optimizing some variables while holding others fixed is suitable.

23 [0091] With the ability of the system 100 to capture detailed motion and precise latent  
24 coordinates of particle images, it becomes possible in principle to recover high resolution detail  
25 of the flexible parts of protein molecules that move, that would otherwise be blurred in standard  
26 reconstruction results. In some cases, "focal" reconstructions can be used, where a particular  
27 focal point in 3D space is chosen, and for each particle image, the flow  $u_i = f_\theta(z_i)$  is  
28 approximated by a local rigid transform around the focal point. The local rigid transform, given by  
29 a rotation and shift, is composed with the overall pose of the particle  $\phi_i$  to arrive and a new  
30 "focal" pose for each image. The particle images are then subject to standard Fourier-space  
31 reconstruction from these updated poses, and the result is a 3D density map where the flexible  
32 motion around the focal point has been corrected, and local density may be improved. Focal  
33 reconstructions can be repeated at many focal points covering the protein molecule, and the

1 resulting maps stitched together. However, this approach is generally tedious and encounters  
2 issues when there is any curvature or non-rigidity in the flow fields.

3 [0092] As discussed herein, one can reasonably assume that the deformation flow fields will  
4 be smoother than the high resolution canonical map. Accordingly, the entire model can be  
5 optimized at a small box size,  $N = N_L$ . Once optimization is complete, the flow generator  
6 parameters  $\theta$  and the latent coordinates  $z_i$  can be ‘frozen’, and then transferred to a new model  
7 at full resolution, with  $N = N_H$ . The particles can be partitioned using the same split that was used  
8 in the consensus refinement (from which the poses  $\phi_i$  are obtained). For each half-set, the  
9 canonical density  $V$  is initialized to zero, and re-optimized at full box size  $N_H$ . This yields two  
10 half-maps that can be compared via Fourier-Shell Correlation (FSC). In this way, any FSC  
11 correlation beyond the Nyquist limit at box size  $N_L$  represents true signal recovered in common  
12 in both half-maps that was never seen during training. As a consequence it cannot be spurious  
13 or a result of over-fitting.

14 [0093] To this end, the system 100 optimizes  $V$  at high resolution under the full model using  
15 the same optimization procedures described herein. In some cases, using minibatch SGD was  
16 found to not yield high quality results. One can speculate that noise in the minibatch gradient  
17 estimator is problematic for this task. Nevertheless, the present inventors were able to solve the  
18 problem using full-batch gradient-based optimization using the Limited-memory-Broyden–  
19 Fletcher–Goldfarb–Shanno (L-BFGS) gradient-based optimization method. This is generally more  
20 computationally expensive than direct Fourier-space reconstruction, and requires many iterative  
21 passes over the whole dataset. But it is notable that the present inventors were able to obtain  
22 high quality results in this way. In particular, this approach solves high resolution detail in all  
23 flexible parts of the protein molecule simultaneously, without making assumptions of local rigidity  
24 or smoothness, validated with FSC.

25 [0094] The present inventors conducted example experiments where the application of the  
26 system 100 to two experimental cryo-EM datasets was examined. The experiments demonstrated  
27 the ability of the method to resolve multiple dimensions of non-linear non-rigid protein motion with  
28 sufficient fidelity to improve reconstruction resolution of flexible parts. The first dataset contains  
29 snRNP spliceosome particle images, and the second contains TRPV1 ion-channel particle  
30 images.

31 [0095] For each dataset, a rigid consensus refinement was computed using all particle  
32 images. Non-uniform refinement is used to improve image alignments and overall resolution. The  
33 resulting particle images and poses  $\phi_i$  are fixed and used in training the model of the present

1 embodiments. No prior information is provided about the type or form of heterogeneity in each  
2 dataset. The system 100 was run with a real-space mask that excludes solvent in the canonical  
3 density  $V$ , and for membrane proteins a separate mask is used to enforce zero deformation of  
4 detergent micelle or lipid nanodisc. For the TRPV1 channel, 3D Variability Analysis (3DVA) was  
5 first run on the particle images to generate initializations for latent coordinates  $z_i$ , again without  
6 any prior information about heterogeneity. Each experiment was run on a single NVIDIA™ V100  
7 GPU.

8 [0096] To display continuous motion results as static figures, multiple points  $z_{\text{display}}$  in the  
9 latent space of the model are selected and the corresponding convected densities as  $W_{\text{display}} =$   
10  $D(f_\theta(z_{\text{display}}), V)$  are generated. These densities can be rendered overlayed in multiple colors  
11 and with reference position guide markers to aid the visibility of motion.

12 [0097] The U4/U6.U5 tri-snRNP complex represents a large part of the spliceosome and is a  
13 molecular machine that has several moving parts, linkages, and flexible domains. The example  
14 experiments processed a dataset of 138,899 snRNP particles (EMPIAR-10073). The raw particles  
15 had a box size of 380 pixels and a pixel size of 1.4Å and were first processed through  
16 heterogeneous refinement in *cryoSPARC* software to separate broken particles that are missing  
17 the "head" region. This yielded 102,500 final particles that are downsampled to a box size of 180  
18 pixels (pixel size 2.95Å) and used to train the model. The original raw particles are then input to  
19 the system described herein and used to reconstruction the canonical density (in two separate  
20 half-maps) to high resolution, once the system 100 has learned the motion.

21 [0098] The system 100 was trained starting from random initialization for latent coordinates  
22 and the flow generator network, with a  $K = 5$  dimensional latent space. The flow generator was  
23 a 6-layer MLP with 64 units in each hidden layer and output layer, using ReLU activations. A  
24 tetrahedral mesh with 1601 vertices and 5859 cells is created, covering the density, and the flow  
25 generator outputs the deformation field at these vertices.

26 [0099] The system 100 recovered five dimensions of latent motion (illustrated in FIG. 5), with  
27 each dimension explaining a different type of bending or twisting in the molecule. There are two  
28 large moving parts, the head region and foot region. Both are attached to the more rigid central  
29 body region. In the learned deformation fields from 3DFlex, the foot region largely moves as a  
30 rigid sub-part, with a hinge-like linkage to the body. The head region has significant internal  
31 flexibility and a large range of motion, and all five latent directions encode some motion in the

1 head region. Notably, the system 100 recovers the motion of all regions from a random  
2 initialization without any labels defining parts or linkages.

3 [0100] FIG. 5 illustrates results of the system 100 with a  $K = 5$ -dimensional latent space on  
4 102,500 particles of an snRNP Spliceosome complex, demonstrating the capacity for the system  
5 100 to resolve multiple modes of non-rigid deformation simultaneously. Shown are scatter plots  
6 showing the final distribution of particle latent coordinates across the dataset and convected  
7 densities from the system 100 at minus one and plus one standard deviations in the latent space,  
8 along each of the five latent dimensions. Each dimension resolves a different type of motion within  
9 the same model.

10 [0101] Along with the motion, the system 100 was able to recover high-resolution detail in the  
11 canonical density map (illustrated in FIG. 6). Individual  $\alpha$ -helices can be seen translating several  
12 Angstroms while retaining side-chain features. Likewise, a  $\beta$ -sheet in the flexible head region is  
13 sufficiently resolved to separate  $\beta$ -strands, despite the presence of non-rigid motion. The system  
14 100 only had access to experimental data at a maximum Nyquist resolution of 5.9 $\text{\AA}$  during training,  
15 and so these structural features represent additional signal that is resolved from the data due to  
16 precise modelling of the motion.

17 [0102] FIG. 6 illustrates results of the system 100 on 102,500 particles of an snRNP  
18 Spliceosome complex, demonstrating the capacity for the system 100 to resolve detailed non-  
19 rigid motion and high-resolution structure simultaneously. Shown are a series of convected  
20 densities from the model, at latent coordinates along the first latent dimension, and with focus on  
21 key structural details. The top row shows an  $\alpha$ -helix in the head region of the protein that  
22 translated by several Angstroms and the bottom row shows a  $\beta$ -sheet in the head region that  
23 translates and deforms.

24 [0103] As expected, in regions of substantial motion and flexibility, improvements between a  
25 static conventional refinement and the system are dramatic (illustrated in FIG. 7). For example,  
26 local resolution in the center of the head region is improved from 5.7 $\text{\AA}$  to 3.8 $\text{\AA}$ . For a complex as  
27 large as the snRNP, it is possible to create manual masks around regions that are expected to be  
28 rigid and then to perform local or multi-body refinement. These techniques can improve resolution  
29 and map quality in some domains, such as the foot region, that remain rigid despite motion relative  
30 to the remainder of the molecule. In contrast, the system 100 does not require any manual  
31 masking or prior knowledge about the motion of the molecule. It can detect and then correct for  
32 non-rigid flexibility across the entire molecule at once.

1 [0104] FIG. 7 illustrates results of the system 100 with a  $K = 5$ -dimensional latent space on  
2 1002,500 particles of an snRNP Spliceosome complex. On the left is a density map from  
3 conventional refinement. On the right is a canonical density map from the system 100. The two  
4 maps are filtered by local resolution to aid in visualizing weak density in low resolution areas in  
5 the conventional refinement.

6 [0105] The TRPV1 ion channel is a 380kDa tetrameric membrane protein that acts as a heat-  
7 and capsaicin-activated sensory receptor. A dataset of 200,000 particle images of TRPV1 in  
8 nanodisc (EMPIAR-10059) with a box size of 224 pixels and pixel size of 1.21Å was processed.  
9 These particles were downsampled to 128 pixels (pixel size 2.15Å) and used to train the system  
10 100. The original raw particles were then used to reconstruction the canonical density (in two  
11 separate half-maps) to high resolution, once the system 100 had learned the motion.

12 [0106] The experiments first ran 3D Variability Analysis with  $K = 2$  components. This  
13 provides initialization for  $K = 2$  latent coordinates  $z_i$ . The flow generator, a 3-layer MLP with 32  
14 units in each hidden and output layer, was randomly initialized. The system 100 was trained for 5  
15 epochs over the data with the latent coordinates fixed. The latents were then unlocked and were  
16 optimized at each iteration. Training proceeded for 10 epochs at a time, alternating between fixing  
17 the canonical density and fixing the flow generator, for a total of 50 epochs.

18 [0107] The final result is a model that has captured  $K = 2$  types of flexible coordinated  
19 motion amongst the four peripheral soluble domains of the ion channel (illustrated in FIG. 8).  
20 Along the first latent dimension, each pair of opposing subunits bends towards each other while  
21 the other pair bends apart. The second motion involves all four subunits twisting concentrically  
22 around the channel's pore axis. In both cases, the peripheral-most helices move by approximately  
23 6Å. Both motions are non-rigid and involve flexure of substantial regions of the protein density.

24 [0108] FIG. 8 illustrates results of the system 100 with a  $K = 2$ -dimensional latent space on  
25 200,000 particles of a TRPV1 ion channel protein, demonstrating the capacity to resolve detailed  
26 motion of smaller, membrane proteins. Shown are scatter plots showing the final distribution of  
27 particle latent coordinates across the dataset. Also shown are canonical density that is solved by  
28 the system 100. The micelle is not excluded in the density but is masked to have zero deformation,  
29 so that the system 100 focuses on motion of the protein density. Also shown are convected  
30 densities from the system 100 at minus one and plus one standard deviations in the latent space,  
31 along each of the two latent dimensions. The first dimension resolves a motion where opposite  
32 soluble domains move together or apart. The second dimension resolves a motion where all four  
33 soluble domains twist around the axis of the central pore.

1 [0109] In a conventional refinement of the TRPV1 channel structure, these motions are  
2 detrimental to reconstruction quality and resolution (illustrated in FIG. 9). Several  $\alpha$ -helices in the  
3 soluble region are so poorly resolved that helical pitch is barely visible. Local resolution reaches  
4 2.8 $\text{\AA}$  in the rigid core of the channel, but only 4 $\text{\AA}$  in the periphery. The system 100, on the other  
5 hand, estimates and accounts for the motion of these domains, and substantially improves  
6 resolution and map quality. The system 100 only has access to experimental data up to a  
7 maximum Nyquist resolution of 4.3 $\text{\AA}$  during training, but FSC and local resolution measurements  
8 using the two separate half-set reconstruction show that it recovers consistent structural  
9 information beyond this resolution. Local resolutions in peripheral helices improve to 3.2 $\text{\AA}$   
10 revealing helical pitch and side chain details.

11 [0110] FIG. 9 illustrates results of the system 100 with a  $K = 2$ -dimensional latent space on  
12 200,000 particles of a TRPV1 ion channel protein. On the left is shown a density map from  
13 conventional refinement with local resolution. On the right side is shown canonical density map  
14 from the system 100, using the same local resolution scale. The two maps are identically filtered  
15 and sharpened and displayed at the same threshold level, so that visual comparison of map  
16 quality is possible. The result of the system 100 shows clear improvement in map quality and local  
17 resolution in peripheral flexible domains of the protein. Also shown are detailed views showing  
18 improvement in helical density in the flexible soluble domains.

19 [0111] The two separate half-set reconstructions from the system 100 allow for the ability to  
20 use established validation procedures to measure the improvement from modelling motion. FIG.  
21 10A shows that the global FSC curve of the entire density improves slightly with the system 100  
22 compared to conventional refinement. This indicates that in the highest resolution rigid core region  
23 of the molecule, the system 100 has not lost any structural information. To investigate the effect  
24 in the peripheral domains, the experiments constructed a soft-edged mask around one of the  
25 flexible domains (shown in FIG. 10C). Computing FSC curves within this mask (illustrated in FIG.  
26 10B) shows that the system 100 improves the average resolution from 3.4 $\text{\AA}$  to 3.2 $\text{\AA}$  as well as  
27 the SNR at lower resolutions. This means that the system 100 has resolved more structural  
28 information than conventional refinement for this flexible protein, and validates that the motion  
29 learned by the system 100 is a better model of the particle than a null hypothesis of no motion.

30 [0112] FIGS. 10A to 10C show validation and resolution estimation results of the system 100  
31 with a  $K = 2$ -dimensional latent space on 200,000 particles of a TRPV1 ion channel protein. FSC  
32 curves were measured between half-sets of particles that are used to compute two half-map  
33 reconstructions. The system 100 only had access to experimental data up to a Nyquist resolution

1 limit of 4.3Å during training. Therefore, any correlation beyond this resolution in half-set  
2 reconstructions indicates resolved signal rather than spurious correlation or over-fitting. FIG. 10A  
3 shows a FSC of the entire density of the ion channel, showing that the system 100 resolves high-  
4 resolution details in the rigid core of the protein. FIG. 10B shows FSC using a mask around one  
5 of the flexible peripheral domains in the soluble region of the ion channel. The mask is depicted  
6 in FIG. 10C and is soft-edged. In this region, the system 100 provides a substantially improved  
7 FSC curve and resolution estimate of 3.2Å; versus 3.4Å for conventional refinement. Notably, a  
8 local focused refinement using the same mask is unable to improve resolution beyond the  
9 conventional refinement due to the small size and non-rigidity of the flexible region.

10 [0113] The system 100 has improved the reconstruction of this flexible protein by explicitly  
11 modelling non-rigid deformation. As a baseline, the experiments also performed a local focused  
12 refinement using the same mask (FIG. 10C) to isolate one soluble domain. Local refinement is  
13 unable to improve the density or resolution of the domain and beyond the conventional  
14 refinement. This is expected, as each soluble domain is less than 50kDa in size and deforms  
15 flexibly. It is believed that this comparison illustrates an additional advantage of the system 100.  
16 Unlike local and multi-body refinement methods that assume rigidity and attempt to fit separate  
17 pose parameters for each masked region, the system 100 can exploit correlations between  
18 different moving parts that may move together, making it possible to infer the position of all parts,  
19 even though individually each is too small to align reliably. In the case of TRPV1, the four soluble  
20 domains deform in different directions by different amounts, but the system 100 infers their  
21 positions in a given image jointly.

22 [0114] Advantageously, the system 100 can also be used to model heterogeneity. Discrete  
23 heterogeneity is a long-studied form of heterogeneity that can be easily modelled and computed  
24 from cryo-EM image data. Discrete models approximate continuous conformational landscapes  
25 with a finite number of points as a crude approximation. As such, they cannot aggregate structural  
26 information across conformations, and have no notion of flexibility or protein motion. Some  
27 recently proposed deep learning methods for heterogeneity attempt to approximate continuous  
28 conformational landscapes with finer and more elaborate discrete clustering methods, still without  
29 a notion of motion or unifying structural details across conformational space. They require  
30 massive datasets due to the partitioning of conformational space into discrete reconstructions.

31 [0115] Local and multibody refinements are another long-studied method to attempt to deal  
32 with continuous heterogeneity. Local refinements use masks to cut a 3D protein density into  
33 finitely many (usually a small number less than 5) rigid sub-parts that can rotate and/or shift

1 relative to one-another. This approximation to a continuous deformable object can allow for  
2 improvement in map quality and resolution, but is only applicable when there are actually rigid  
3 substructures, each with enough density to be individually alignable. These methods also require  
4 substantial manual interaction to define masks.

5 [0116] Several other methods have been proposed to model continuous conformational  
6 heterogeneity using density-based models. The simplest of these, methods based on eigen-  
7 analysis, model conformational landscapes as linear subspaces in 3D density. More advanced  
8 techniques use non-linear manifold embedding or deep generative models to construct a non-  
9 linear manifold in the space of 3D density. These methods do succeed in modelling continuous  
10 heterogeneity in the sense that the models are actually continuous, but the models do not have a  
11 notion of protein motion. Instead, density-based models are limited to adding and subtracting  
12 density from different areas of a 3D structure, and do not aggregate structural information across  
13 the landscape or enable improved reconstructions of heterogeneous proteins.

14 [0117] Other methods attempt to model continuous conformational change with an underlying  
15 notion of protein motion. Hypermolecules define continuous conformational change by  
16 representing protein density using a higher-dimensional space that could, in principle, capture  
17 deformation and structure together. This technique is quite different from the architecture of the  
18 present embodiments as it does not explicitly separate canonical density and deformation, and it  
19 has generally yet to produce notable results on experimental data. Other methods fit a Gaussian  
20 Mixture Model (GMM) as a representation of canonical density, and uses a deep generative model  
21 to adjust the parameters of the GMM to model continuous conformational change of protein  
22 density. While GMM parameters can be adjusted to convect density in modeling motion, the GMM  
23 parameters can equally be used to add or subtract density like a density-based model, and no  
24 regularization is used to enforce that the model actually learns motion. Furthermore, the deep  
25 generative model is a (variational) auto-encoder meaning that a separate encoder network is  
26 necessary, unlike the present embodiments where latent variable search is used instead. The  
27 encoder network in the GMM model also does not perform inference by operating directly on  
28 experimental data, instead it takes as input a gradient of the data likelihood with respect to initial  
29 GMM parameters. This means that the encoder is limited to encoding motion that is small relative  
30 to the size scale of Gaussians used in the mixture, limiting motion the model can learn to locally  
31 linear small motion of high resolution objects, or locally linear large motion of low resolution  
32 objects. In contrast, the present embodiments can capture large motion even of high-resolution

1 objects. Finally, due to these limitations and computational limitations, the GMM model does not  
2 improve the resolution or map quality of the canonical density.

3 [0118] The present embodiments can also be used in other applications, extensions and  
4 variations. For example, in addition to a deformation flow generator, the model can include an  
5 additional module that generates a "presence/absence" field that captures the  
6 association/dissociation of subunits in a complex. This module would take in the latent coordinate  
7  $z_i$  and output one or more masks indicating which parts of the canonical density are present in  
8 the  $i$ 'th particle image. The mask would then be used to multiply the canonical density before  
9 deforming through the convection operator. In this way, the model can handle both flexible  
10 conformational variability and compositional variability. For instance, this could be helpful when  
11 working with a dataset containing a protein and a ligand, or multiple subunits in a complex.

12 [0119] The present embodiments can also be used on multiple different datasets of protein  
13 particle images jointly. For example, data can be collected with particles in the presence and  
14 absence of a ligand, or at different temperatures, or with different sample characteristics of any  
15 kind, or at different times after a reaction has taken place. The data can be combined into one  
16 particle stack that is used by the system 100, and the distribution of particles in the latent space  
17 can be analyzed with respect to the originating sample conditions. This can be used for example  
18 to understand the changes in conformational dynamics after ligand binding or after an enzyme  
19 has begun acting on it's substrate.

20 [0120] In some cases, the model can be initialized with latent coordinates or flow generator  
21 parameters, or a canonical density, from any source; for example, other algorithms that uncover  
22 the conformational landscape.

23 [0121] The latent coordinates output from the system 100 can be used in conjunction with the  
24 flow generator to compute "free energy" or related molecular energies to determine the likelihood  
25 of a given conformational state.

26 [0122] In some cases, the model can fit initially to image data from one source, for example  
27 a low energy, low cost microscope or a microscope set with imaging conditions that maximize  
28 low/mid-resolution contrast at the expense of high-resolution contrast. Once the model is trained  
29 on this data, it can then be fine-tuned on a different source of image data, for example high-energy  
30 high-cost microscope data, or data from a microscope set to maximize high-resolution information  
31 at the expense of low/mid-resolution information, or from data collected with minimal radiation  
32 damage, or from data of a different target sample (eg. with or without a ligand).

1 [0123] In some cases, the model can be extended to support particles with various  
2 geometries, sub-units, symmetries, and connected parts by expanding the generative architecture  
3 to have multiple canonical maps and multiple flow generators. In some uses, each canonical map  
4 can represent the density of the canonical structure in one spatial region, and each flow generator  
5 can generate the deformation for that region. In other uses, each canonical density can represent  
6 one discrete object, and the model can have an additional discrete latent variable capturing which  
7 object is present in the particle image. Furthermore, regularizers and constraints can be applied  
8 straightforwardly to, for example, enforce that the output of the flow generator in a certain region  
9 is zero, or that two regions of the protein object have the same structure (symmetry) or that one  
10 region moves rigidly but is attached to another region that deforms non-rigidly. In general, any  
11 combination of at least one canonical map, flow generator, and any other components can be  
12 used to compose the image formation process within the method described herein.

13 [0124] Although the invention has been described with reference to certain specific  
14 embodiments, various modifications thereof will be apparent to those skilled in the art without  
15 departing from the spirit and scope of the invention as outlined in the claims appended hereto.  
16 The entire disclosures of all references recited above are incorporated herein by reference.

## CLAIMS

1. A method for determining the three-dimensional (3D) structure and 3D motion of a molecule from two-dimensional (2D) or 3D experimental images, the 3D structure in the form of one or more canonical 3D maps, and the 3D motion in the form of parameters of one or more flow generators, the method comprising:
  - receiving the experimental images;
  - receiving contrast transfer function (CTF) parameters for each experimental image;
  - initializing pose parameters for each experimental image;
  - initializing unknown model parameters of image formation, the unknown model parameters comprising:
    - one or more canonical 3D maps; and
    - the parameters of the one or more flow generators, each flow generator comprising a parameterized generator function taking as input a latent coordinate vector and outputting a 3D deformation field;
  - wherein the image formation taking as input at least a latent coordinate vector and outputting a simulated image, the image formation comprising:
    - generating one or more 3D deformation fields by inputting the latent coordinate vector into the one or more flow generators;
    - performing a convection and projection operation by convecting one or more of the canonical 3D maps by the one or more 3D deformation fields, and projecting using the pose parameters for the given experimental image; and
    - performing CTF corruption of the projected result using the CTF parameters of the given experimental image to generate the simulated image;
  - fitting the unknown model parameters to the experimental images by performing one or more iterations of gradient-based optimization of an objective function, the objective function taking as input at least simulated images and experimental images,

wherein in at least one of the iterations, performing latent variable search for at least one experimental image, latent variable search for a given experimental image comprising:

performing the image formation one or more times to generate simulated images from one or more latent coordinate vectors; and

selecting one or more latent coordinate vectors based on the similarity between simulated images and the experimental image;

and wherein in at least one of the iterations, updating at least one of the unknown model parameters using at least one of the experimental images, updating the at least one of the unknown model parameters comprising:

generating simulated images by performing the image formation using the one or more selected latent coordinate vectors associated with the experimental images;

evaluating the objective function using at least the simulated images and the experimental images;

computing the gradient of the objective function with respect to the at least one unknown model parameter to be updated; and

updating the unknown model parameter using the gradient; and

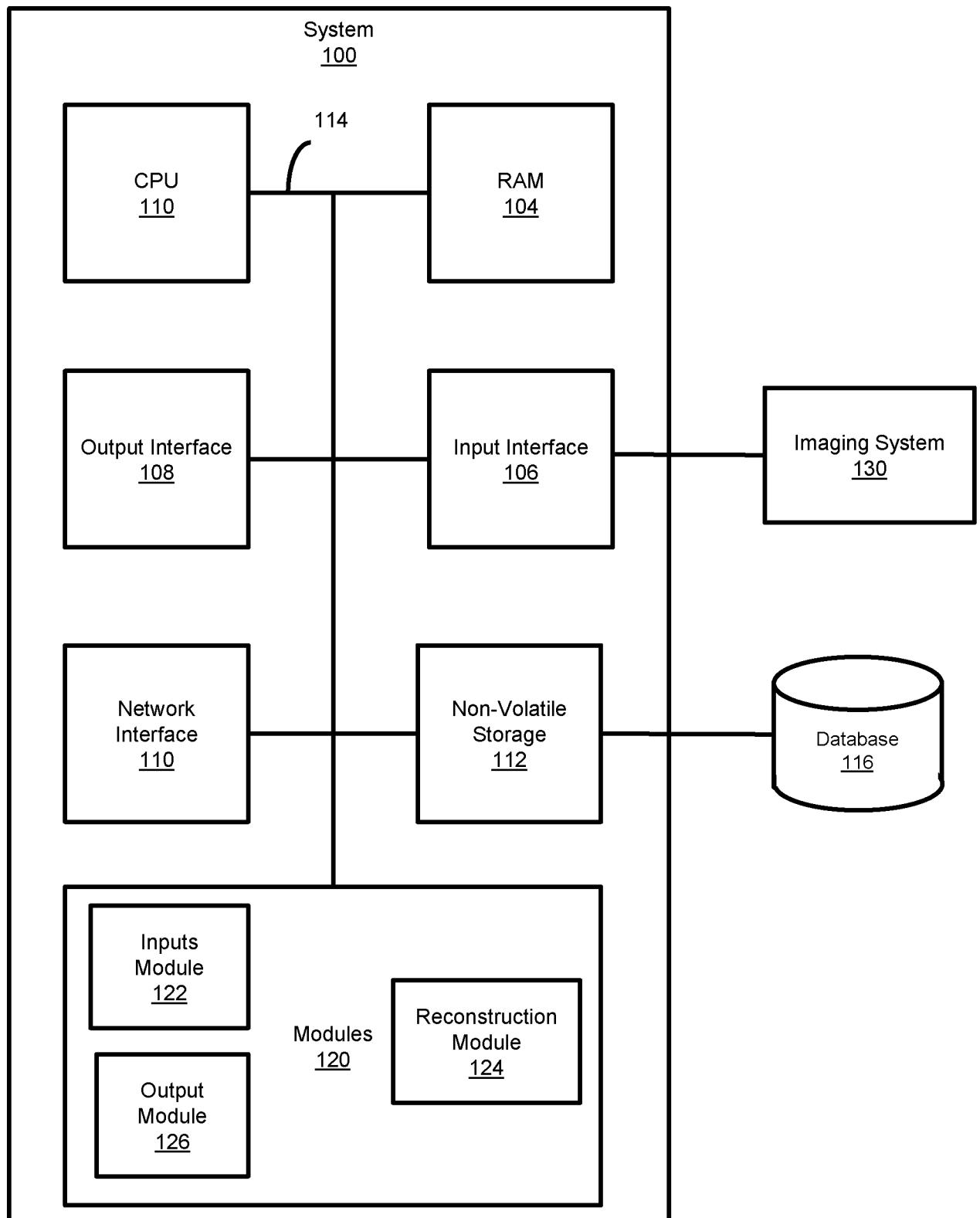
outputting the one or more canonical 3D maps, the parameters of the one or more flow generators, or both.

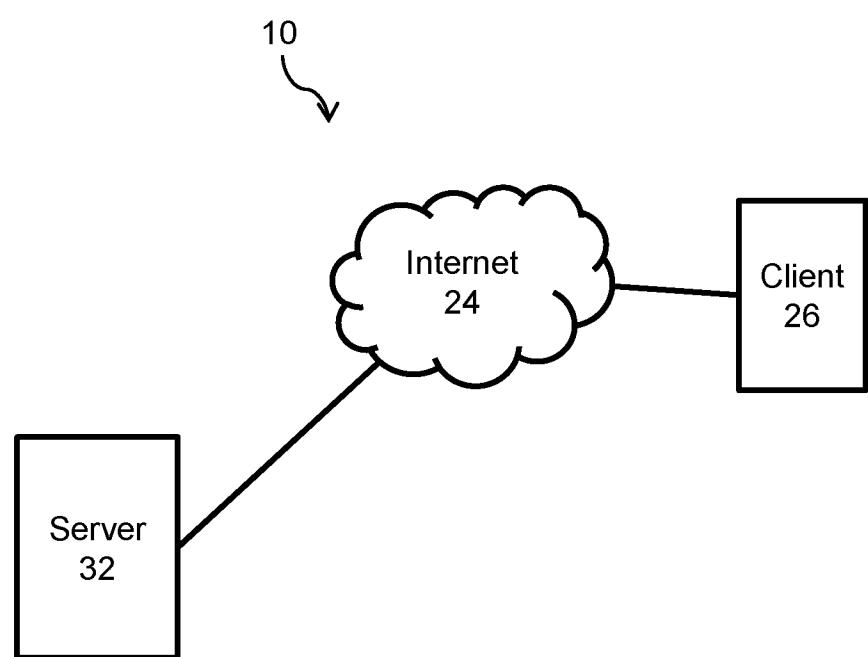
2. The method of claim 1, wherein the pose parameters for each experimental image are initialized by receiving them as input.
3. The method of claim 1, wherein the experimental images and the simulated images are 2D.
4. The method of claim 1, wherein the canonical 3D map is represented as a 3D voxel array of real-space density values.
5. The method of claim 1, wherein the convection and projection operation comprises interpolating the canonical 3D map to form the convected and projected image, the interpolation operation defined by the deformation field and the pose parameters.
6. The method of claim 1, wherein the flow generator is a feed-forward neural network.

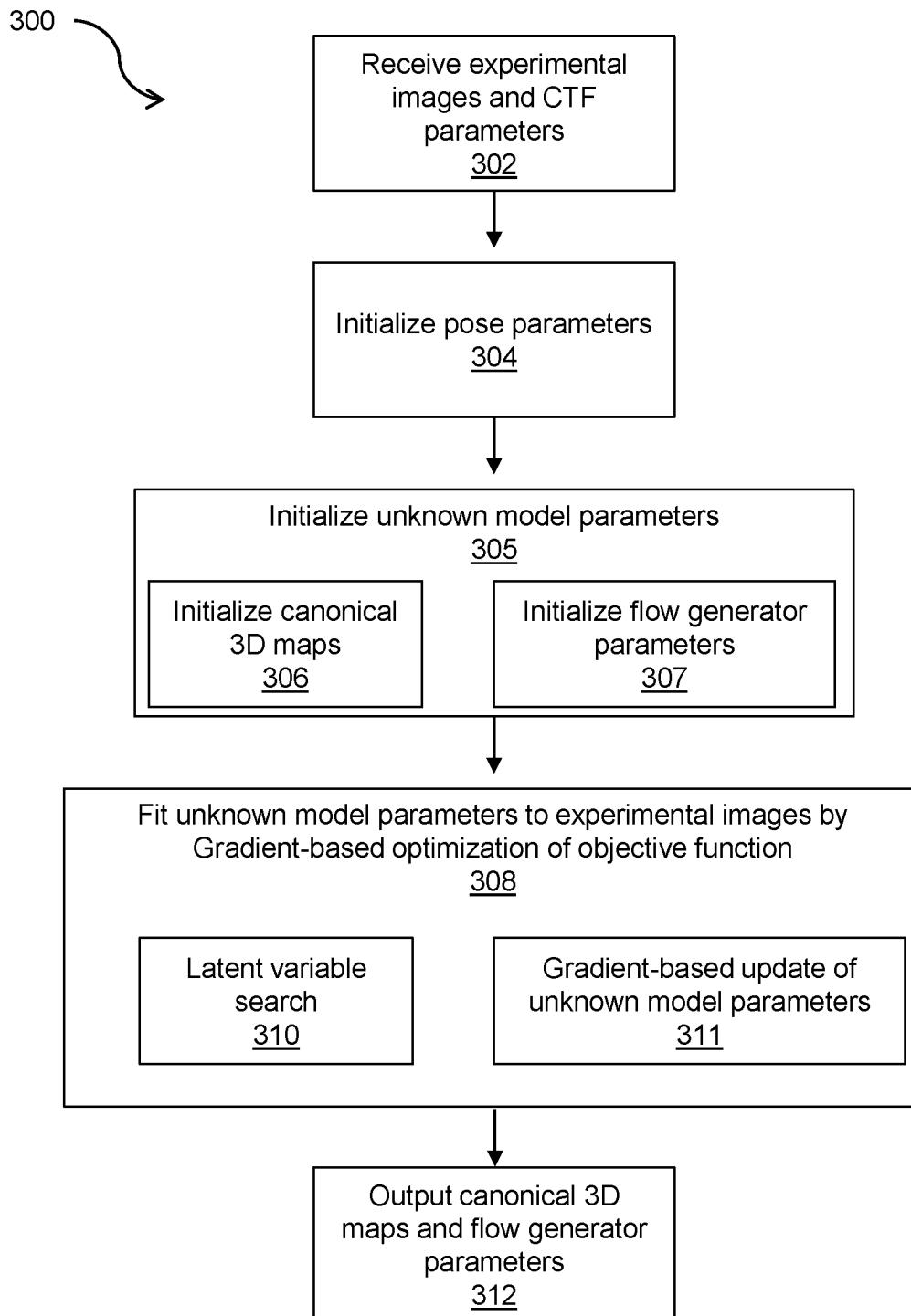
7. The method of claim 1, wherein a 3D deformation field is represented as deformed positions of vertices of a mesh of volumetric elements, the deformation vector field values defined by interpolation within the volume of each mesh element.
8. The method of claim 1, wherein the objective function comprises at least the negative log-likelihood of a simulated image given an experimental image using a noise model.
9. The method of claim 8, wherein the noise model is a Gaussian noise model
10. The method of claim 1, wherein the latent variable search comprises coordinate descent search over the latent coordinate vector space.
11. The method of claim 1, wherein the latent variable search comprises gradient descent search over the latent coordinate vector space.
12. The method of claim 1, wherein the latent variable search comprises selecting latent coordinate vectors that are equal to the latent coordinate vectors determined to optimize the similarity plus a component of random noise.
13. The method of claim 1, wherein each iteration of the gradient-based optimization comprises selecting a random subset of the experimental images.
14. The method of claim 1, wherein updating at least one of the unknown model parameters using the gradient is performed using the Adam update rule.
15. The method of claim 1, wherein updating at least one of the unknown model parameters using the gradient is performed using the Stochastic Gradient Descent update rule.
16. The method of claim 1, wherein updating at least one of the unknown model parameters using the gradient is performed using the Stochastic Gradient Descent with momentum update rule.
17. The method of claim 1, wherein updating at least one of the unknown model parameters using the gradient is performed using the L-BFGS update rule.
18. The method of claim 1, wherein updating the canonical 3D map using the gradient is performed using the L-BFGS update rule.
19. The method of claim 1, wherein the objective function comprises regularizing terms.
20. The method of claim 19, wherein regularizing terms comprise terms that penalize non-rigid deformation in a 3D deformation field.
21. The method of claim 20, wherein a 3D deformation field is represented as deformed

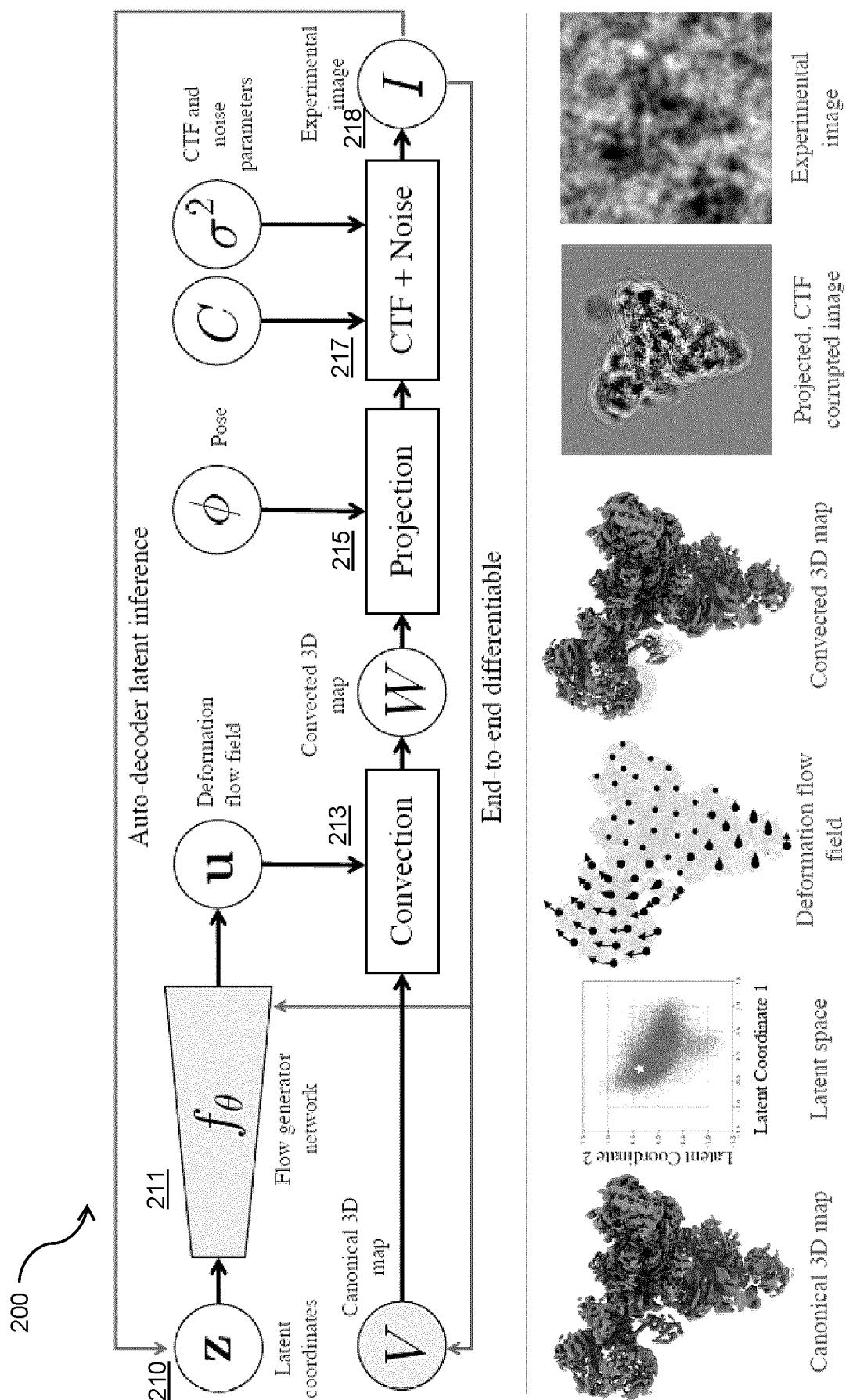
positions of vertices of a mesh of volumetric elements, the deformation vector field values defined by interpolation within the volume of each mesh element, and regularizing terms comprise terms that penalize non-rigid deformation within each mesh element.

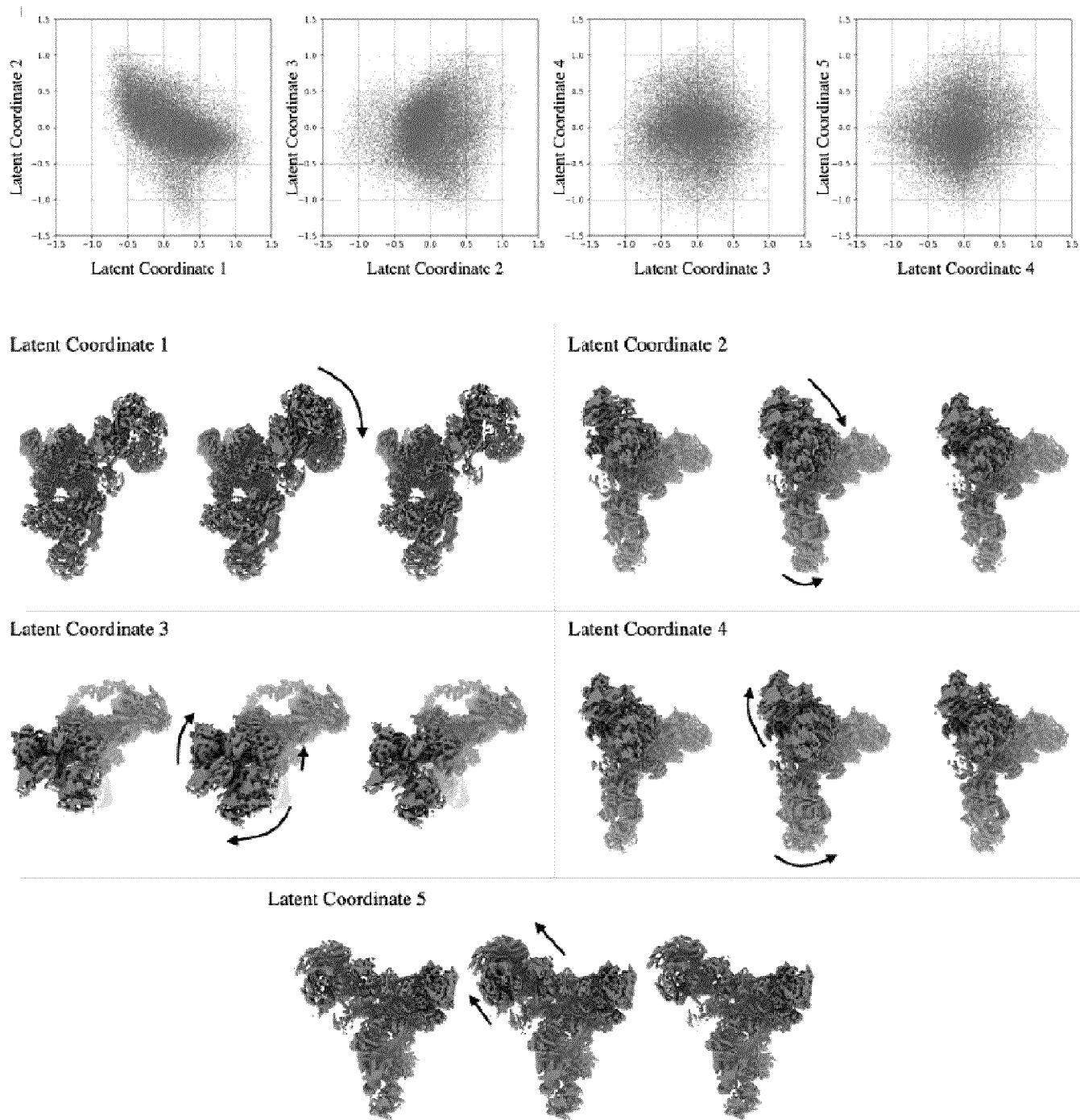
22. The method of claim 1, wherein the optimization of the unknown model parameters is performed using spatial frequency marching
23. The method of claim 1, wherein the optimization of the unknown model parameters is first performed using a lower spatial frequency limit and subsequently performed using a higher spatial frequency limit.

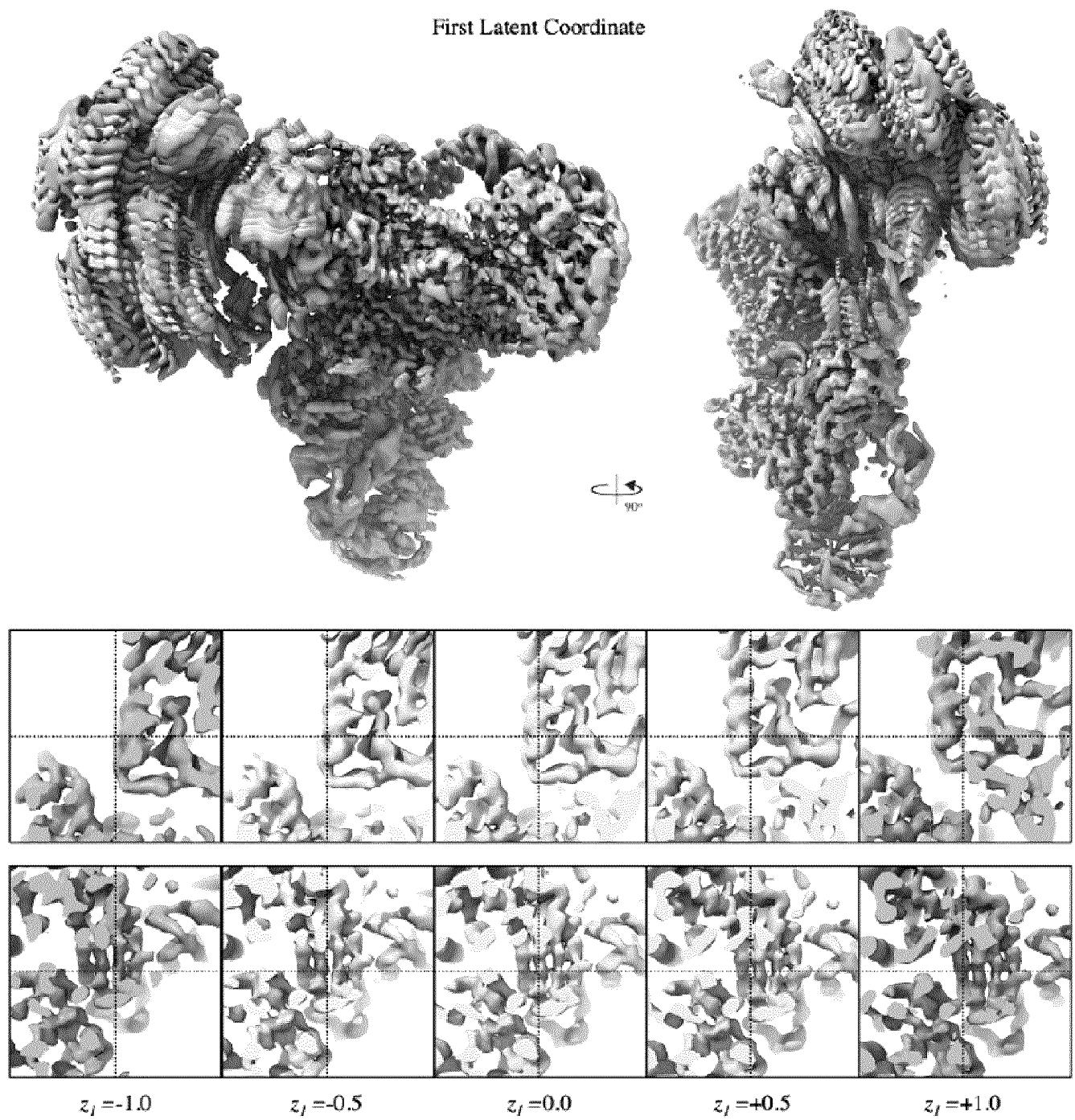
**FIG. 1**

**FIG. 2**

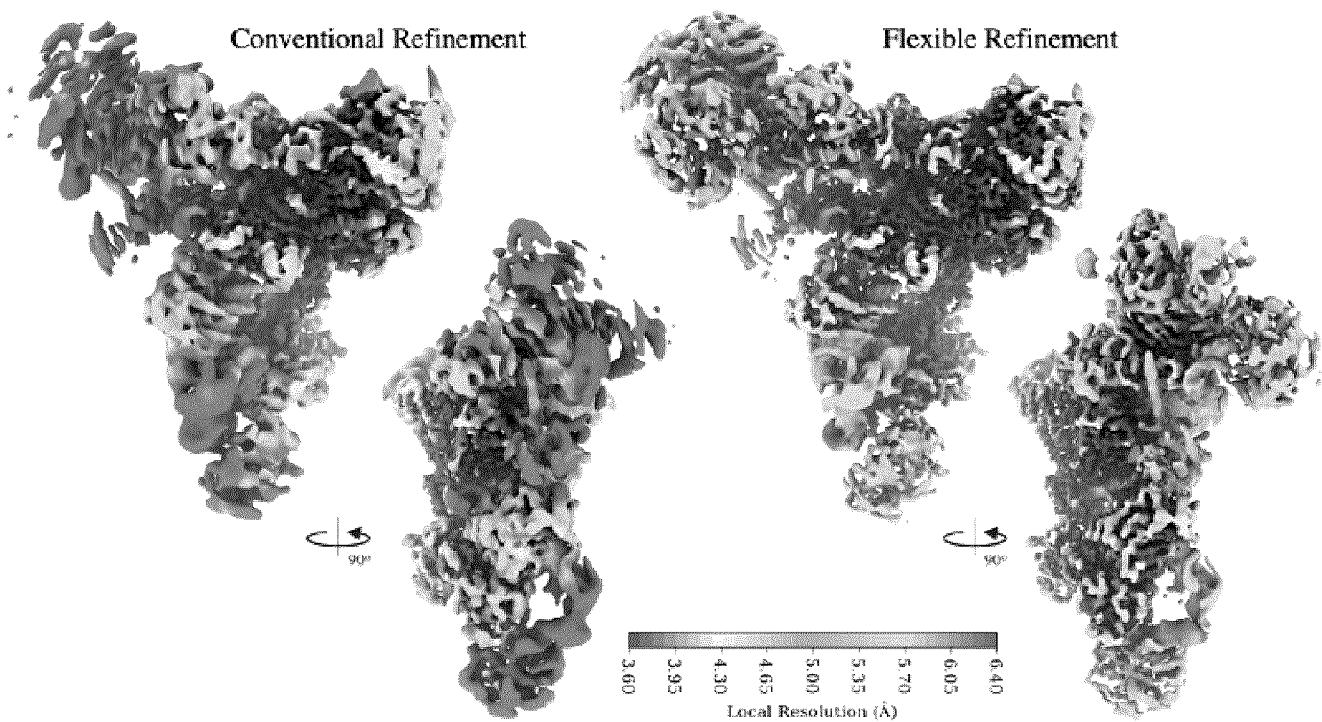
**FIG. 3**

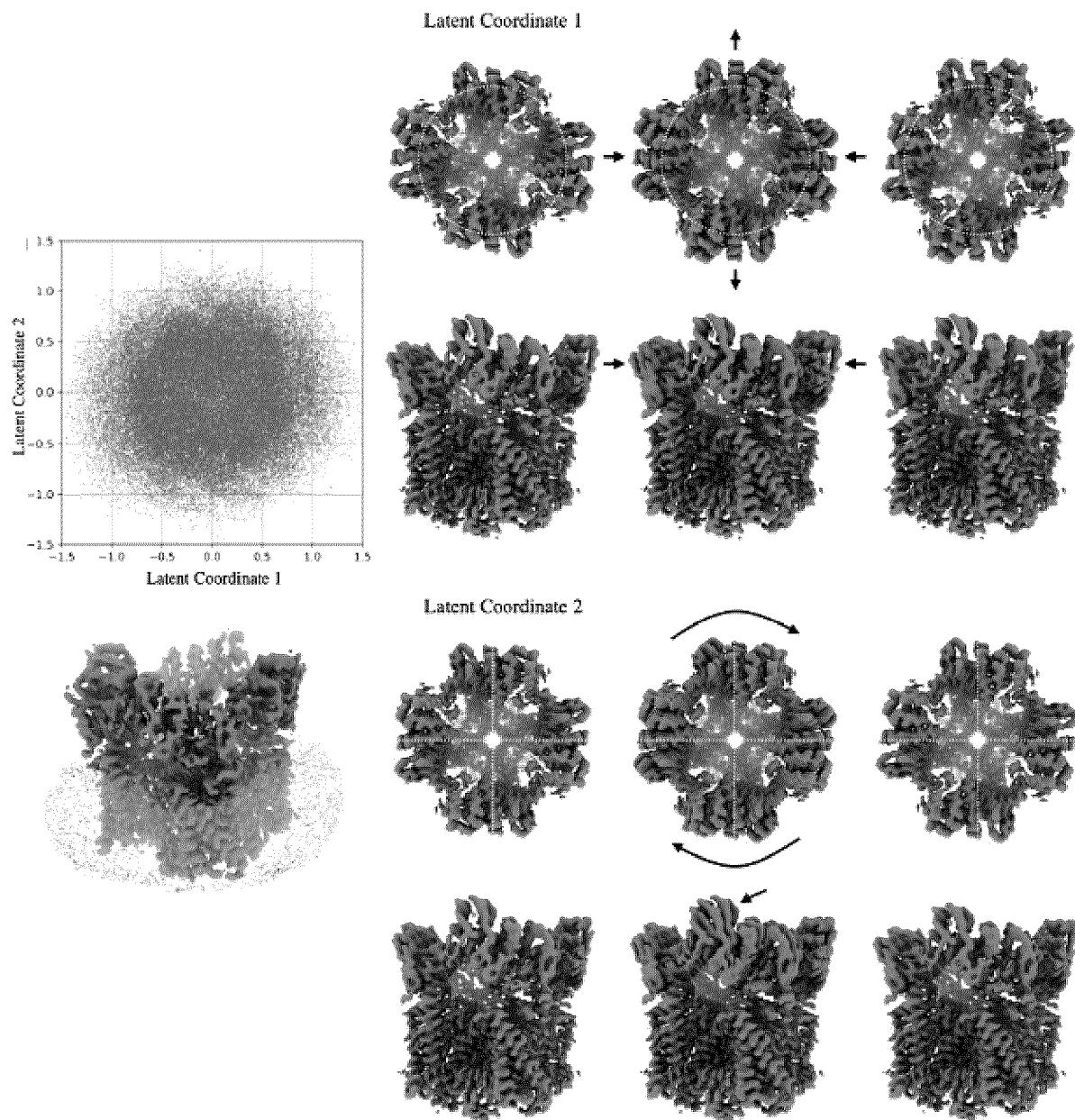


**FIG. 5**

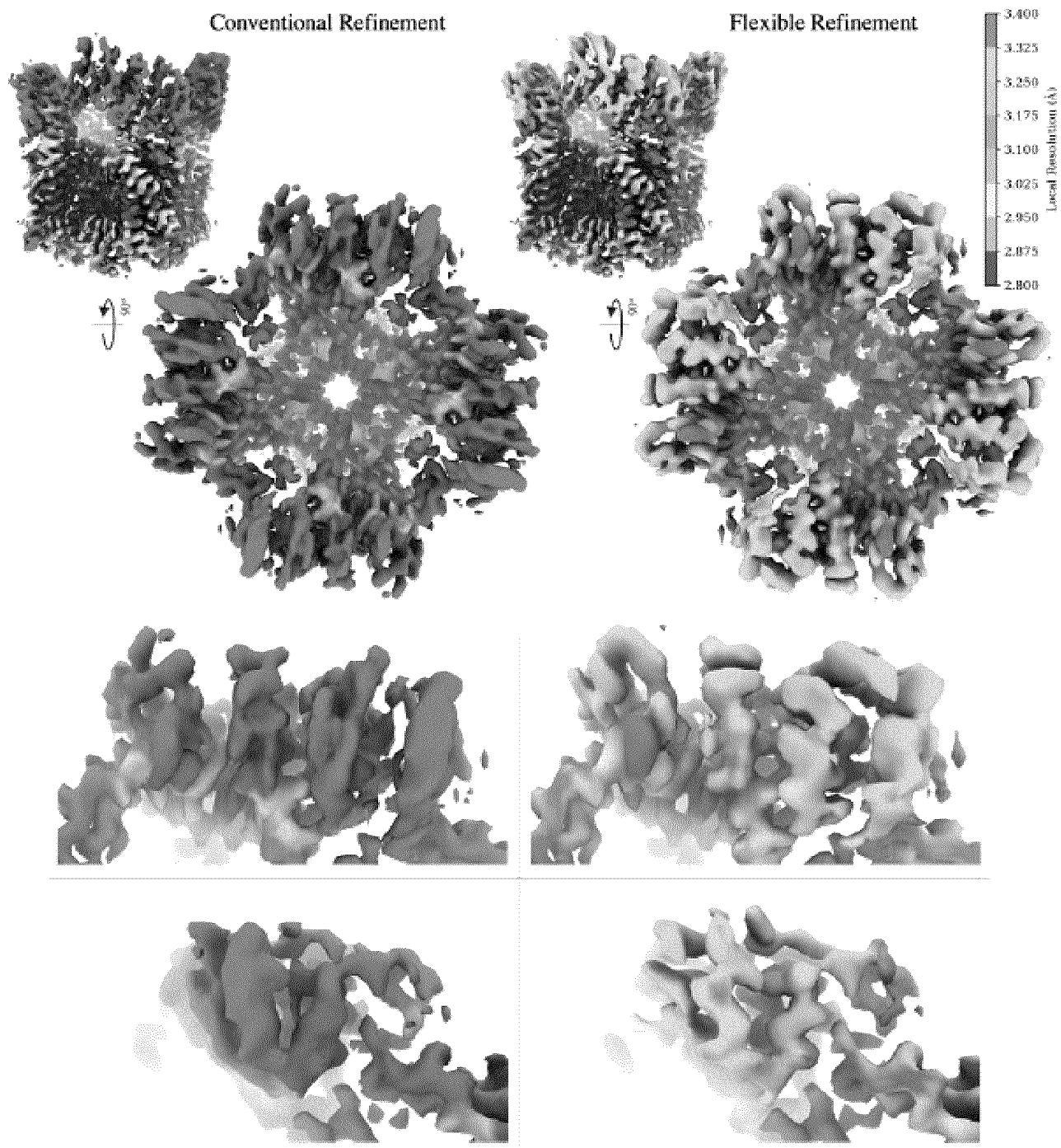


**FIG. 6**

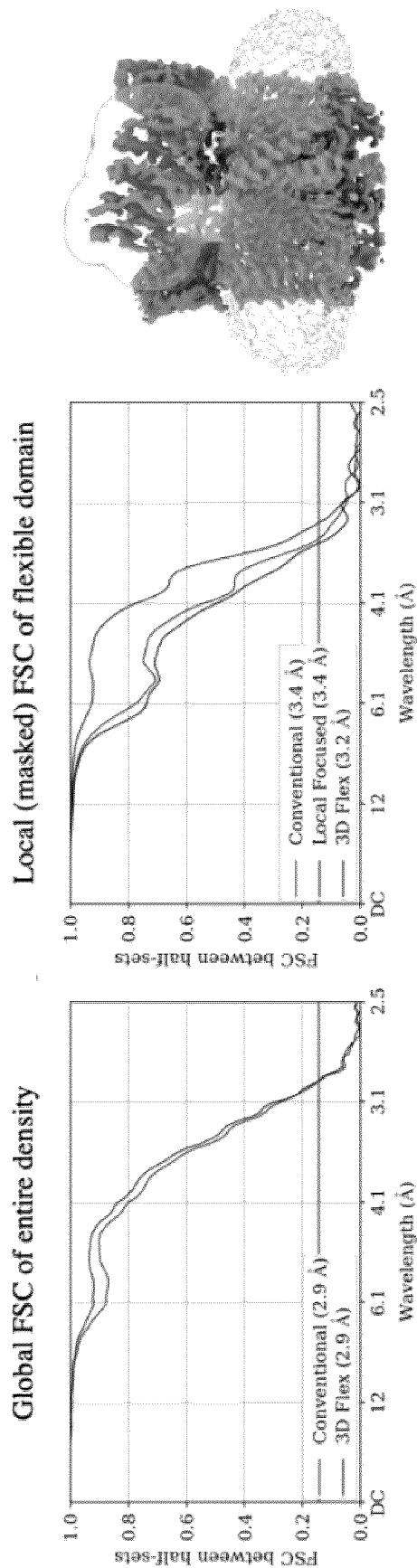
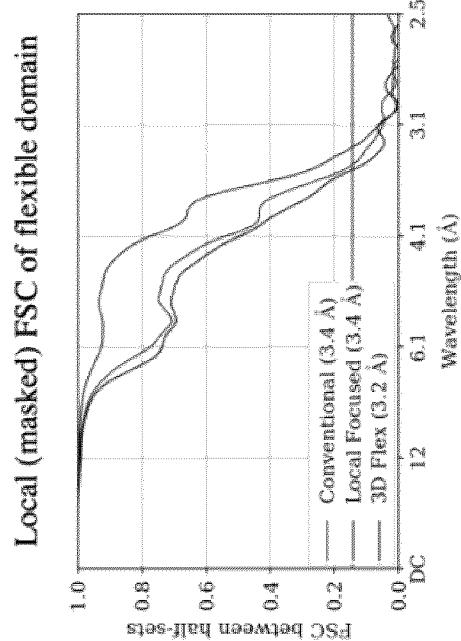
**FIG. 7**

**FIG. 8**

9/10

**FIG. 9**

10/10

**FIG. 10A****FIG. 10B****FIG. 10C**

**INTERNATIONAL SEARCH REPORT**

International application No.  
**PCT/CA2022/050614**

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC: **G16B 15/00** (2019.01), **G16B 5/00** (2019.01), **G16B 15/20** (2019.01), **G16C 20/00** (2019.01), **G01N 23/00** (2006.01)

CPC: G16B 5/00 (2022.02), G16B 15/00 (2022.02), G16B 15/20 (2020.01), G16C 20/00 (2022.02), G01N 23/00 (2020.01),  
G01N 2223/418 (2020.01)

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): G16B 15/00 (2019.01), G16B 5/00 (2019.01), G16B 15/20 (2019.01), G16C 20/00 (2019.01), G01N 23/00 (2006.01); (CPC): G16B 5/00 (2022.01), G16B 15/00 (2022.01), G16B 15/20 (2020.01), G16C 20/00 (2022.01), G01N 23/00 (2020.01), G01N 2223/418 (2020.01)

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

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

Databases: Questel Orbit, Google Patents, Scopus, PubMed, IEEE Xplore. Keywords: 3d structure, motion, canonical 3D map, flow generator (neural network), contrast transfer function (CTF), unknown parameter, latent coordinate vector, 3D deformation field, gradient optimization, objective function (negative log likelihood), latent variable search.

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category* | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No. |
|-----------|---|-----------------------|
| P,X       | P,X PUNJANI, A. et al., 3D Flexible Refinement: Structure and Motion of Flexible Proteins from Cryo-EM, <i>bioRxiv</i> , 22 April 2021 (22-04-2021)<br>* see whole document * | 1-23                  |
| A         | CA3140807A1 (PUNJANI, A.), 10 December 2020 (10-12-2020)<br>* see whole document *  | 1-23                  |
| A         | US20200066371A1 (BRUBAKER, M. A. et al.), 27 February 2020 (27-02-2020)<br>* see whole document *   | 1-23                  |
| A         | CA3078256A1 (PUNJANI, A. et al.), 11 April 2019 (11-04-2019)<br>* see whole document *  | 1-23                  |

Further documents are listed in the continuation of Box C.

See patent family annex.

|  |   |
|--|---|
| *<br>“A”<br>document defining the general state of the art which is not considered to be of particular relevance   | “T”<br>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  |
| “D”<br>document cited by the applicant in the international application  | “X”<br>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   |
| “E”<br>earlier application or patent but published on or after the international filing date   | “Y”<br>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 |
| “L”<br>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) | “&”<br>document member of the same patent family  |
| “O”<br>document referring to an oral disclosure, use, exhibition or other means  |   |
| “P”<br>document published prior to the international filing date but later than the priority date claimed  |   |

Date of the actual completion of the international search  
06 July 2022 (06-07-2022)

Date of mailing of the international search report  
11 July 2022 (11-07-2022)

Name and mailing address of the ISA/CA  
Canadian Intellectual Property Office  
Place du Portage I, C114 - 1st Floor, Box PCT  
50 Victoria Street  
Gatineau, Quebec K1A 0C9  
Facsimile No.: 819-953-2476

Authorized officer  
Abdallah Hamed (819) 639-5290

## INTERNATIONAL SEARCH REPORT

International application No.  
**PCT/CA2022/050614**

| C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT |  |                       |
|---|--|-----------------------|
| Category*   | Citation of document, with indication, where appropriate, of the relevant passages   | Relevant to claim No. |
| A   | WO2019018040A9 (ZHANG, T. et al.), 14 March 2019 (14-03-2019)<br>* see whole document *  | 1-23                  |
| A   | WO2020058174A1 (SENIOR, W. A. et al.), 26 March 2020 (26-03-2020)<br>* see whole document *  | 1-23                  |
| A   | PUNJANI, A. et al., 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM, <i>Journal of Structural Biology</i> , vol. 213, issue 2, 11 February 2021 (11-02-2021)<br>* see whole document * | 1-23                  |
| A   | BARNETT, A. et al., Rapid Solution of the Cryo-EM Reconstruction Problem by Frequency Marching, <i>SIAM Journal on Imaging Sciences</i> , vol. 10, issue 3, pages 1170-1195, 01 August 2017 (01-08-2017)<br>* see whole document *                       | 1-23                  |
| A   | SINGER, A. et al., Computational Methods for Single-Particle Electron Cryomicroscopy, <i>Annual Review of Biomedical Data Science</i> , vol. 3, issue 1, pages 163-190 , 20 July 2020 (20-07-2020)<br>* see whole document *                             | 1-23                  |
| A   | ZHONG, E. D. et al., CryoDRGN: Reconstruction of heterogeneous cryo-EM structures using neural networks, <i>Nature Methods</i> , vol. 18, issue 2, pages 176-185, 4 February 2021 (04-02-2021)<br>* see whole document *                                 | 1-23                  |
| A   | LYUMKIS, D., Challenges and opportunities in cryo-EM single-particle analysis, <i>The Journal of biological chemistry</i> , vol. 294, issue 13, pages 5181-5197, 25 February 2019 (25-02-2019)<br>* see whole document *                                 | 1-23                  |

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.  
**PCT/CA2022/050614**

| Patent Document Cited in Search Report | Publication Date              | Patent Family Member(s)   | Publication Date  |
|--|-------------------------------|---|---|
| CA3140807A1                            | 10 December 2020 (10-12-2020) | EP3980763A1<br>WO2020243839A1   | 13 April 2022 (13-04-2022)<br>10 December 2020 (10-12-2020)   |
| US2020066371A1                         | 27 February 2020 (27-02-2020) | US2017103161A1<br>US10282513B2  | 13 April 2017 (13-04-2017)<br>07 May 2019 (07-05-2019)  |
| CA3078256A1                            | 11 April 2019 (11-04-2019)    | EP3692360A1<br>EP3692360A4<br>US2020333270A1<br>WO2019068201A1  | 12 August 2020 (12-08-2020)<br>21 July 2021 (21-07-2021)<br>22 October 2020 (22-10-2020)<br>11 April 2019 (11-04-2019)  |
| WO2019018040A1                         | 24 January 2019 (24-01-2019)  | WO2019018040A9<br>US2021142498A1<br>US11227403B2  | 14 March 2019 (14-03-2019)<br>13 May 2021 (13-05-2021)<br>18 January 2022 (18-01-2022)  |
| WO2020058174A1                         | 26 March 2020 (26-03-2020)    | CA3110200A1<br>CA3110242A1<br>CA3110395A1<br>CN112585684A<br>CN112585685A<br>CN112585686A<br>EP3821433A1<br>EP3821434A1<br>EP3821435A1<br>JP2022501694A<br>JP2022501695A<br>JP2022501696A<br>US2021304847A1<br>US2021313008A1<br>US2021407625A1<br>WO2020058176A1<br>WO2020058177A1 | 26 March 2020 (26-03-2020)<br>26 March 2020 (26-03-2020)<br>26 March 2020 (26-03-2020)<br>30 March 2021 (30-03-2021)<br>30 March 2021 (30-03-2021)<br>30 March 2021 (30-03-2021)<br>19 May 2021 (19-05-2021)<br>19 May 2021 (19-05-2021)<br>19 May 2021 (19-05-2021)<br>06 January 2022 (06-01-2022)<br>06 January 2022 (06-01-2022)<br>06 January 2022 (06-01-2022)<br>30 September 2021 (30-09-2021)<br>07 October 2021 (07-10-2021)<br>30 December 2021 (30-12-2021)<br>26 March 2020 (26-03-2020)<br>26 March 2020 (26-03-2020) |