Journal of Structural Biology xxx (2013) xxx–xxx

Contents lists available at ScienceDirect

Journal of Structural Biology

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

iMODFIT: Efficient and robust flexible fitting based on vibrational analysis in internal coordinates

José Ramón Lopéz-Blanco, Pablo Chacón *

Department of Biological Physical Chemistry, Rocasolano Physical Chemistry Institute, CSIC, Serrano 119, Madrid 28006, Spain

ARTICLE INFO

Article history:
Received 14 May 2013
Received in revised form 20 August 2013
Accepted 22 August 2013
Available online xxxx

Keywords:
Hybrid methods
Molecular modeling
Flexible fitting
Normal mode analysis
Internal coordinates
Coarse-grained models

ABSTRACT

Here, we employed the collective motions extracted from Normal Mode Analysis (NMA) in internal coordinates (torsional space) for the flexible fitting of atomic-resolution structures into electron microscopy (EM) density maps. The proposed methodology was validated using a benchmark of simulated cases, highlighting its robustness over the full range of EM resolutions and even over coarse-grained representations. A systematic comparison with other methods further showcased the advantages of this proposed methodology, especially at medium to lower resolutions. Using this method, computational costs and potential overfitting problems are naturally reduced by constraining the search in low-frequency NMA space, where covalent geometry is implicitly maintained. This method also effectively captures the macromolecular changes of a representative set of experimental test cases. We believe that this novel approach will extend the currently available EM hybrid methods to the atomic-level interpretation of large conformational changes and their functional implications.

1. Introduction

The merging of information from electron microscopy (EM) and X-ray crystallography constitutes a fruitful approach to decipher the functional mechanisms of essential macromolecular complexes. Computational fitting techniques enable the interpretation of EM data that are captured in different functional states in terms of available atomic structures. Early methods attempted to fit the optimal position and orientation of a rigid atomic structure into a given EM density map (Fabiola and Chapman, 2005; Wriggers and Chacón, 2001). However, available atomic structures frequently exhibit different conformations than those captured using EM. Thus, understanding the observed conformational changes often requires a flexible fit. Various approaches have been developed to address the high dimensionality of this problem. Notable examples have simplified the fitting search using a linear combination of low-frequency normal modes (Hinsen et al., 2005; Siebert and Navaza, 2009; Suhre et al., 2006; Tama et al., 2004a), including the toxic anthrax complex (Tama et al., 2006), the Ca-ATPase pump (Hinsen et al., 2005) and the *Escherichia coli* protein-conducting channel bound to a translating ribosome (Mitra et al., 2005). Reduced vector-based representations (Rusu et al., 2008; Wriggers et al., 2004) have also produced efficient fitting protocols; the functional motions of RNA polymerase II and its GreB factor were characterized using this procedure (Darst et al., 2002; Opalka et al., 2003). Several molecular dynamics (MD)-based methods (Chen et al., 2003; Orzechowski and Tama, 2008; Trabuco et al., 2008; Trabuco et al., 2009) incorporate an extra force term to adapt the atomic structure into a targeted high-resolution EM map. Biased MD techniques have also been applied to fit impressive high-resolution maps of the ribosome in different conformational states (Ratje et al., 2010; Villa et al., 2009; Whitford et al., 2011). Furthermore, the computational cost of MD methods has been alleviated by employing coarse-grained (CG) approximations (Grubisic et al., 2010; Whitford et al., 2010; Whitford et al., 2011). Other methods include simplified force fields and elastic network models (Schröder et al., 2007; Tan et al., 2008; Zheng, 2011). Finally, both homology modeling (Topf and Sali, 2005; Velazquez-Muriel and Carazo, 2007; Zhu et al., 2010) and rigidity-based constraints (Jolley et al., 2008) have also proven to be useful strategies for flexible fitting. Finally, other authors proposed a consensus fit obtained with several flexible fitting approaches (Ahmed and Tama, 2013; Ahmed et al., 2012).

Although many successful applications exist for flexible fitting, the major challenge is to combine efficiency and accuracy while limiting overfitting problems. Such problems arise when the information encoded in the low-resolution map is insufficient to guide the process correctly. In this situation, atomic structures can be distorted by local or unrealistic motions. Rather than introducing arbitrary constraints or regularization steps, we instead present an efficient alternative based on Normal Mode Analysis (NMA) in internal coordinates (IC). NMA in IC was established in computational biology in the early 1980s. The seminal works of Go, Levitt

* Corresponding author.
E-mail address: pablo@chaconlab.org (P. Chacón).

1047-8477/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jsb.2013.08.010

and others (Brooks and Karplus, 1985; Go et al., 1983; Levitt et al., 1985; Noguti and Go, 1983; Tirion, 1996) set the foundation for decomposing molecular motions into series of deformation modes using dihedral angles as variables. NMA is based on a harmonic approximation of the dynamics of a system around the equilibrium conformation. The potential directions of the motion encoded in each deformation mode are obtained by diagonalizing the Hessian (the second derivative of the potential energy) and kinetic energy matrices. High-frequency modes represent localized displacements, whereas low-energy modes correspond to collective conformational changes. A great deal of evidence indicates that conformational changes can be described by collections of low-frequency modes. Moreover, it has been demonstrated that low-frequency modes computed in IC provide a reasonable and inexpensive direct view of the relevant conformational space, even at different CG levels (Bray et al., 2011; Lopez-Blanco et al., 2011). In the current study, we used the probable directions encoded in this essential space to flex the atomic structure while maximizing the density overlap with the target experimental map. NMA in IC naturally reduces the conformational search space to physically realistic collective motions and implicitly maintains the covalent structures, thus preventing distortions.

In the following discussion, we first present the methodological details of our automatic flexible fitting tool, termed iMODFIT. We validated this novel approach using representative simulated test cases within a wide range of resolutions. Next, a systematic comparison against other methods that revealed the advantages of iMODFIT, particularly at lower resolutions, was performed. Finally, we illustrate the excellent comparative performance of this method in the flexible fitting of a set of representative experimental cases.

2. Methods

First, we will look briefly at the employed NMA methodology, and then we will discuss the actual fitting approach guided by low-frequency modes in more detail.

2.1. NMA computation in internal coordinates

NMA merged with CG models is a powerful and popular alternative that can be used to simulate the collective motions of macromolecular complexes at extended timescales (Bahar et al., 2010; Cui and Bahar, 2007; Skjaerven et al., 2009; Tama and Brooks, 2006). Vibrational analysis evaluates the relevant collective motions based on harmonic approximation around a local minimum and allows for motion decomposition into a series of deformation modes. Computing the modes using the canonical backbone dihedral angles as variables has two main advantages over standard Cartesian approaches; first, it implicitly preserves the model geometry by minimizing potential distortions, and second, it substantially reduces the number of variables, thereby improving efficiency (Bray et al., 2011; Kovacs et al., 2005; Lopez-Blanco et al., 2011; Lu et al., 2006; Mendez and Bastolla, 2010). To compute low-frequency modes in IC, we employed the NMA engine of iMod (Lopez-Blanco et al., 2011). Briefly, the macromolecule is simplified as a set of pseudo-atoms connected by harmonic springs, and the modes are computed by solving the following generalized eigenvalue problem:

$$\mathbf{HU} = \lambda_k \mathbf{TU} \quad \text{where} \quad U = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N), \tag{1}$$

where $\mathbf{u}_k$ is the k^{th} deformation vector with its associated λ_k eigenvalue. In other words, the potential directions of the motion encrypted in the modes are obtained by diagonalizing the second derivative matrices of the potential (**H**) and the kinetic (**T**) energies. The potential energy was defined as:

$$V = \sum_{i<j} F_{ij}\left(r_{ij} - r_{ij}^0\right)^2 + s\sum_{\alpha}\left(\theta_\alpha - \theta_\alpha^0\right)^2 \quad \text{where}$$
$$F_{ij} = 1/\left(1 + \left(r_{ij}^0/3.8\right)^6\right) \tag{2}$$

The first term in this equation corresponds to Hooke's potential, where the superindex 0 indicates the initial equilibrium conformation, r_{ij} is the distance between atoms i and j, and F_{ij} is the spring stiffness matrix, which is represented by a sigmoid function. The second term prevents irrational low frequencies generated from floppy tips by adding an extra-torsional stiffness to each dihedral angle, θ_α (Lu et al., 2006).

Another advantage derived from using iMod framework is versatility. iMODFIT can handle multiple chains of proteins, nucleic acids or even small rigid ligands. Moreover, different graining levels can be chosen to represent protein structures. It is possible to select between all-heavy atoms (HA), a five pseudo-atoms (C5) representation (NH, Cα, CO, Cβ and virtual mass located at the mass center of the remaining side chain atoms) or a single Cα atom per amino acid (CA) representation. In addition, greater computational savings can be obtained by randomly selecting different subsets of dihedral angles for the NMA. As we will demonstrate, such CG strategies are very useful for large systems.

2.2. NMA-guided cross-correlation fitting

The proposed flexible fitting method, which is schematized in Fig. 1, searches the conformational space spanned by the lowest frequency torsional modes for the best cross-correlation fit of an atomic model into a given target density map. The initial pose of this atomic model comes from a preliminary coarse registration. Correlation-based rigid body docking tools such as COLORES (Chacón and Wriggers, 2002) or ADP_EM (Garzon et al., 2007) are used to approximately localize the model into the target map.

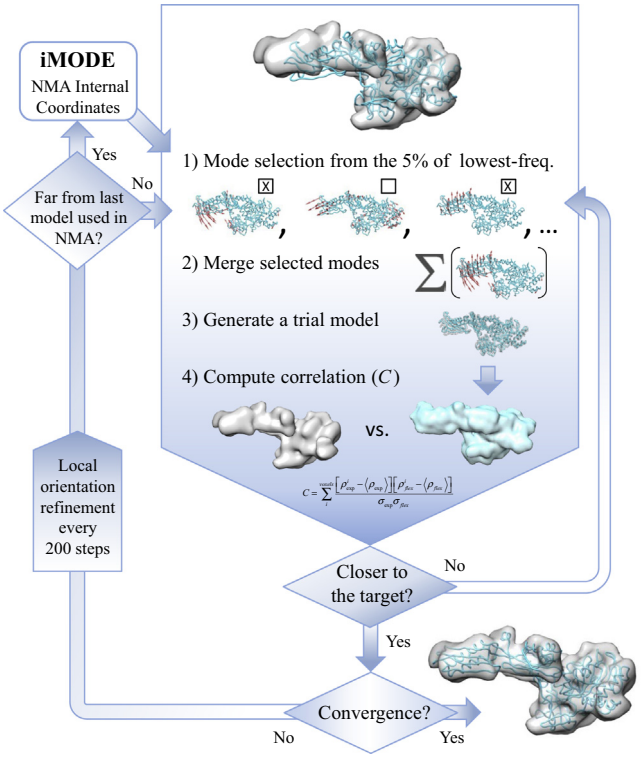

Fig. 1. iMODFIT flexible fitting flowchart.

The actual fitting procedure starts by calculating the low-frequency modes in IC from this initial pose using iMod (Lopez-Blanco et al., 2011). Once the lowest 5% of the modes are calculated, the 10% of them are randomly selected and merged into a single deformation vector that will be used to generate trial structures. The amplitudes of the selected modes are scaled following the collective variable method (Yamashita et al., 2001) to prevent artificially low acceptance rates for the motions with the highest frequency. The selection probability is inversely proportional to the mode's frequency. This stochastic selection permits a fair exploration of the low-frequency torsional modes and naturally avoids the definition of an *ad hoc* number of modes. In our fitting tests with resolutions of 5–25 Å, we empirically found good performance using ~5% of the low-frequency modes. Larger percentages can also be employed, but we obtained slightly distorted models for values greater than 10%, suggesting that higher frequencies contributed to increase over-fitting. In contrast, percentages below 2% were insufficient to satisfactorily describe all of the motions encoded in our test. In certain cases, the conformational space spanned by these few modes was too small, and the flexible fitting got stuck in a local minimum.

Applying the merged deformation vector generates a new trial model that is low-pass-filtered to produce a simulated density map. The fitting score is defined as the normalized cross-correlation between the EM experimental map, ρ_{exp}, and this simulated map, ρ_{trial}:

$$C = \sum_{i}^{voxels} \frac{[\rho^i_{exp} - \langle\rho_{exp}\rangle][\rho^i_{trial} - \langle\rho_{trial}\rangle]}{\sigma_{exp}\sigma_{trial}} \quad (3)$$

The flexed conformation is accepted only if the cross-correlation improves; otherwise, new trial deformations are generated and tested. This procedure is repeated iteratively until convergence except when the accepted structure deviates by more than 0.1 Å RMSD from the previous one used for NMA. In this case, the NMA is repeated with the last accepted structure. Additionally, a local rigid-body optimization is performed every 200 iterations to realign the flexed structure using four-steps of a simple correlation-based parabolic interpolation (Press et al., 2007). This local optimization procedure only compensates for small changes in the center of mass and orientation when the structure is flexed. The magnitude of the displacement produced by this optimization is typically below the grid sampling of the map.

2.3. Benchmarks

To validate the fitting performance, we employed a set of 23 different pairs of open and closed protein structures from the molecular motions database (Flores et al., 2006). This representative set of large conformations was previously described (Lopez-Blanco et al., 2011) and corresponds to large motions (from 2 to 18 Å RMSD) that can eventually be captured by standard EM procedures. For each test pair, the open structure was flexibly fitted inside a simulated EM map generated from the closed conformation and *vice versa*. The simulated maps were generated by low-pass filtering of the atomic structures at different resolutions, as described elsewhere (Garzon et al., 2007). The chosen resolutions (5, 8, 10, 15, 20 and 25 Å) covered the typical experimental range of EM measurements. The average displacement between all conformation pairs was 7.6 ± 4.4 Å. In all cases, the sampling was 2 Å/voxel except for the 5 Å maps, where 1 Å/voxel was used instead. The complete list of protein cases is detailed in Table S1.

To verify the robustness of the method with experimental maps we employed the flexible fitting test cases from the cryo-EM modeling challenge (Ludtke et al., 2012). Because this test set is biased towards both high-resolution maps and small-amplitude motions, we enriched it by adding several medium- to low-resolution cases available at the EM data bank (EMDB) with large-amplitude motions. See Table S3 for the complete list of the 28 experimental test cases.

Table 1
iMODFIT fitting results obtained with simulated cases using different coarse-grained representations at different resolutions.

	Resolution (Å)					
CG[a]	5	8	10	15	20	25
HA	1.1 ± 0.8	1.1 ± 0.6	1.1 ± 0.6	1.2 ± 0.7	1.2 ± 0.7	1.3 ± 0.8
C5	1.2 ± 0.9	1.1 ± 0.7	1.1 ± 0.6	1.2 ± 0.7	1.2 ± 0.7	1.3 ± 0.8
Cα	1.6 ± 1.6	1.4 ± 0.9	1.3 ± 0.7	1.4 ± 0.7	1.5 ± 0.7	1.6 ± 0.8
HA50	1.2 ± 0.9	1.1 ± 0.7	1.1 ± 0.7	1.2 ± 0.7	1.2 ± 0.7	1.3 ± 0.8
HA90	1.6 ± 1.5	1.4 ± 1.2	1.4 ± 1.1	1.5 ± 1.0	1.6 ± 1.0	1.7 ± 1.1

[a] Coarse-grained (CG) protein models: Cα: one C_α atom per residue; C5: 3 atoms for backbone and 2 for the side chain; HA: considering all heavy atoms; HA50: HA with 50% of the dihedral angles randomly fixed; HA90: HA with 90% fixed.

2.4. Technical details

The method was implemented in C/C++ and is freely available at http://chaconlab.org/iMODFIT with complete documentation, including tutorials.

3. Results

3.1. Validation with simulated maps

The performance of this novel method was validated by flexibly fitting a set of test cases comprising a wide variety of collective conformational changes (see Section 2). In each test, the atomic structure of a given conformer must be flexed and fitted into a simulated EM map generated from other conformer coordinates. In all cases, the fitting was performed in the two senses of the conformational change at experimental-like resolutions from 5 to 25 Å. Under these controlled conditions, the RMSD, which compares the final fitted model with the original structure, defines the fitting accuracy. As shown in Table 1, this RMSD average was remarkably small, with values close to 1 Å when iMODFIT was used with the all-heavy atoms representation. The fitting accuracy was slightly reduced as the resolution decreased, although even at 25 Å, the deviation was only ~1.5 Å. In other words, the method recovered the original structure using only the lowest-frequency conformational space. The RMSDs at 5 Å were slightly higher than those at 8–10 Å because two of the 46 cases were trapped in a local minimum. By removing these two transition cases (1wdn to 1ggg and 1t5s to 1sn4), the deviations dropped below 1 Å. Table 2 includes the structural quality validation of the final fitted models as analyzed by the Molprobity web server (Chen et al., 2010). As expected, because of the implicit maintenance of the covalent

Table 2
Structural quality assessment of the results obtained with simulated cases.

	Resolution (Å)											
	5		8		10		15		20		25	
CG	C^a	R^b	C	R	C	R	C	R	C	R	C	R
HA	35	0.8	35	0.9	35	0.9	36	0.8	35	0.8	35	0.8
C5	36	0.7	38	0.7	38	0.7	36	0.7	39	0.7	39	0.7
Cα	50	0.6	64	0.7	70	0.7	67	0.7	65	0.7	61	0.7
HA50	35	0.6	35	0.6	35	0.6	35	0.6	35	0.6	34	0.6
HA90	49	0.6	50	0.6	49	0.6	50	0.6	49	0.6	46	0.6

[a-b] Molprobity scores (Chen et al., 2010): *C*, number of clashes per 1000 atoms, and *R*, percentage of Ramachandran backbone φ and ψ angles outside the allowed regions.

structure, the flexed conformations preserved the initial crystallographic geometry. Independent of the map resolution, there were no additional incorrect bond lengths or angles, and only the average clashes and percentages of dihedral angles outside Ramachandran regions increased slightly. The clashes were augmented from 21 to 35, and the percentage of not allowed dihedral angles increased from 0.6 to 0.9. Moreover, these few extra clashes typically involved collisions of side-chain hydrogen atoms. It is worth noting that hydrogen atoms were ignored during the flexible fitting and were only added afterwards to perform the Molprobity check.

The fitting results were similar for the three different available CG representations (see Section 2). We found no difference between the results obtained with the HA or C5 models (Table 1), and even the simplest Cα representation maintained satisfactory fitting results. Additionally, the random removal of dihedral variables emerged as an efficient CG strategy for flexible fitting. For example, by randomly removing 50% of the dihedral variables, we obtained nearly the same fitting results but halved the computational demands. Furthermore, it should be noted that, in this CG procedure, none of the regions were permanently rigid because the dihedrals to be fixed were randomly selected for every NMA calculation (typically 50–100 times per run). Freezing larger percentages reduced the accuracy and increased the number of collisions. For example, when 90% was frozen, the RMSD was 1.6 ± 1.1 Å with an average of 50 collisions, which suggests some structural degradation. Below ∼50%, the random dihedral removal strategy alone or combined with a simplified atomic representation (e.g., C5 or Cα) led to large computational savings without compromising fitting accuracy. These savings largely enhanced the computational efficiency, thereby extending the applicability to larger macromolecular systems.

3.2. Comparison with other flexible fitting methods

Although other methods exist (see Section 1), we performed a systematic comparison with two representative and fully available flexible fitting tools, NMFF and YUPSCX. The former is also an NMA-based approach, but it employs Cartesian coordinates (Tama et al., 2004a; Tama et al., 2004b). In this method, the atoms in one or more residues are grouped to form CG rigid-body blocks. In the flexible fitting process, NMFF uses gradient following and Newton–Raphson minimizations in normal modes space to optimize the overall cross-correlation between the atomic structure and the EM map. YUPSCX is a molecular mechanics-based approach that uses a simplified potential and a simulated annealing protocol as the search engine. YUPSCX was selected over other MD approaches, such as MDFF (Trabuco et al., 2008; Trabuco et al., 2009), FLEX-EM (Topf et al., 2008), or MDfit (Whitford et al., 2011), because of its simplicity and considerably lower computational cost. In any case, as others (Vashisth et al., 2012) have demonstrated, MDFF is limited at low resolutions ($\geqslant$15 Å), even if larger structural restraints and lower steering forces are used to prevent overfitting.

Table 3 shows the deviations of the fitted models obtained with NMFF and YUPSCX using the same test cases. It is important to mention that we substantially improved the performance of NMFF by adapting several default parameters to prevent slow convergence problems. We increased the selected modes from 26 to 36 (including null modes) and the convergence rate from 1 to 0.001. YUPSCX was employed using the default parameters. The comparative results are shown for the 5, 10, 15, and 20 Å results. The RMSD values for all of the methods were reasonably low considering the map resolutions. At higher resolutions, the difference between the methods was not significant, and only deviations of a few tenths were observed. As the resolution decreased, accuracy was gradually lost, and the lowest RMSD values corresponded to our approach. For example, at a resolution of 15 Å, the average RMSD values for iMODFIT, NMFF and YUPSCX were 1.1, 1.2 and 1.7 Å, respectively. The deviations were also higher at lower resolutions; for example, at 20 Å, we obtained fit model deviations of 1.2, 1.4 and 2.0 Å. Remarkably, at a resolution of 5 Å, failure in a local minimum was more frequent, and all of the methods presented higher sigma values. At this very high resolution for EM, the fitness landscape is likely to be more rugged, frustrating the fitting search at least in a few cases. This suggests the need for a different strategy or parameter set to handle more detailed and more rugged density maps. For example, an obvious alternative would be lowering the map resolution by a few angstroms.

Although the final fitted models were convergent, we detected clear differences in their geometric quality. In the case of NMFF, even with simulated maps, we found many incorrect bond lengths and angles, up to 17% and 19%, respectively. These systematic deviations are likely effects of overfitting because the atomic structure displacement along a combination of normal modes in Cartesian coordinates can accumulate important geometrical distortions. For YUPCSX, the bond lengths remained correct, but the percentage of unacceptable bond angles deteriorated from an initial value of 0.9% to ∼7–9%. Small local deformations, which eventually accumulate during the fitting process, could be the reason for such bond angle distortions. In contrast, our methodology maintained the initial crystallographic quality with the introduction of only a few additional clashes.

Table 3
Comparative results of different flexible fitting methods with simulated cases.*

	Resolution (Å)	RMSD (Å)	Clashes (‰)	Rama (%)	Lengths (%)	Angles (%)
iMODFIT	5	1.1 ± 0.8	35 ± 17	0.8 ± 0.8	0.2 ± 0.6	2.6 ± 4.7
	10	1.1 ± 0.6	35 ± 17	0.9 ± 0.8	0.2 ± 0.6	2.6 ± 4.7
	15	1.1 ± 0.7	36 ± 17	0.8 ± 0.8	0.2 ± 0.6	2.6 ± 4.7
	20	1.2 ± 0.7	35 ± 17	0.8 ± 0.7	0.2 ± 0.6	2.6 ± 4.7
NMFF	5	1.3 ± 1.4	28 ± 12	1.2 ± 0.9	15.7 ± 9.5	17.0 ± 9.0
	10	1.1 ± 0.5	28 ± 12	1.2 ± 0.9	16.1 ± 10.5	17.6 ± 10.7
	15	1.2 ± 0.6	28 ± 12	1.2 ± 0.8	16.1 ± 12.2	17.9 ± 11.7
	20	1.4 ± 0.6	29 ± 12	1.3 ± 1.0	16.9 ± 12.2	18.8 ± 12.0
YUPSCX	5	1.1 ± 1.1	22 ± 7	0.7 ± 0.6	0.6 ± 0.9	6.6 ± 5.7
	10	1.3 ± 0.6	41 ± 6	0.6 ± 0.6	0.6 ± 0.8	8.0 ± 6.3
	15	1.7 ± 0.8	46 ± 7	0.7 ± 0.6	0.6 ± 0.7	8.6 ± 6.4
	20	2.0 ± 0.7	49 ± 7	0.7 ± 0.7	0.7 ± 0.7	8.7 ± 6.3
INITIAL		7.6 ± 4.3	21 ± 11	0.6 ± 0.6	0.2 ± 0.6	2.6 ± 4.7

* Clashes and Rama are defined in Table 2. Notice that only iMODFIT maintained the initial percentages of incorrect bond lengths and angles.

Please cite this article in press as: Lopéz-Blanco, J.R., Chacón, P. iMODFIT: Efficient and robust flexible fitting based on vibrational analysis in internal coordinates. J. Struct. Biol. (2013), http://dx.doi.org/10.1016/j.jsb.2013.08.010

3.3. Timing

Comparatively, YUPSCX and iMODFIT are fast approaches, whereas NMFF is quite slow. For example, iMODFIT and YUPSCX were able to process the whole-protein benchmark using an Intel Quad Q6600 Linux box in 6.5 and 4.7 h, respectively. However, for NMFF, the same run required more than 28 h to complete. In contrast, YUPSCX scaled with molecular size more efficiently than our approach did; for example, YUPSCX only took 15 min to flexibly fit the largest example of the benchmark (1su4, 994 residues), whereas iMODFIT, using the costly HA atomic representation, required approximately one hour. Nevertheless, by randomly selecting 50% of the canonical dihedral angles, the whole benchmark could be processed with iMODFIT in less than 4 h without compromising either the fitting accuracy or structural quality.

3.4. Experimental cases

We next confirmed the effectiveness of our approach using a representative set of experimental test cases. The fitting procedure was initiated by docking the corresponding available atomic structure into the experimental EM map using a rigid-body approximation (Garzon et al., 2007). In contrast to the simulated cases, there were no atomic gold standards for comparison, the exact solution is unknown and the determination of fitting accuracy is not possible. In addition to determining the cross-correlation coefficient, we checked the fitting quality by visual inspection and by direct comparison with previous fitting results from other sources.

First we describe the performance of iMODFIT with the 2010 cryo-EM modeling challenge test cases. This benchmark is mainly composed of high-resolution density maps including GroEL (4 Å), the GroEL–GroES complex (7.7 Å), Mm-cpn in its closed state (4.3 Å), aquaporin-0 (2.5 Å), the VP6 component of rotavirus (3.8 Å) and a bacterial ribosome (6.4 Å). All of them were obtained from cryo-EM excepting the aquaporin-0 map that was resolved by electron crystallography. In all of these cases, the original rigid body fittings were successfully refined by iMODFIT, yielding higher cross-correlation values while preserving the initial crystallographic structural quality. These results illustrate the ability of our approach to capture small backbone changes (<1.5 Å Cα RMSD) when the quality and resolution of the map are sufficiently high. The challenge benchmark includes a lidless Mm-cpn in its open state (8 Å), elongation factor eEF2 (8.9 Å) and GroEL–GroES (23.5 Å) density maps, where we observed larger conformational changes. The Cα RMSD between the initial and final models in these cases was ~3–4 Å. In the Mm-cpn and eEF2 cases, the RMSD difference is attributable to the actual conformational change, whereas in the case of the GroEL–GroES complex, it was also the result of the lack of resolution. As described before for simulated cases, the stereochemistry of the initial and fitted structures remained almost the same without requiring any regularization step. Moreover, the deviations between the symmetry-related subunits of GroEL and Mm-cpn of the fitted structures were only 0.6 Å, though no symmetry restraints were used. All of the results obtained using iMODFIT fully agreed with previous results (Chan et al., 2012; Wang and Schroder, 2012).

We extended the resolution test range of the challenge benchmark with more difficult flexible fitting cases that undergo larger conformational changes. Most of these cases were obtained from the electron microscopy data bank and have been used by other

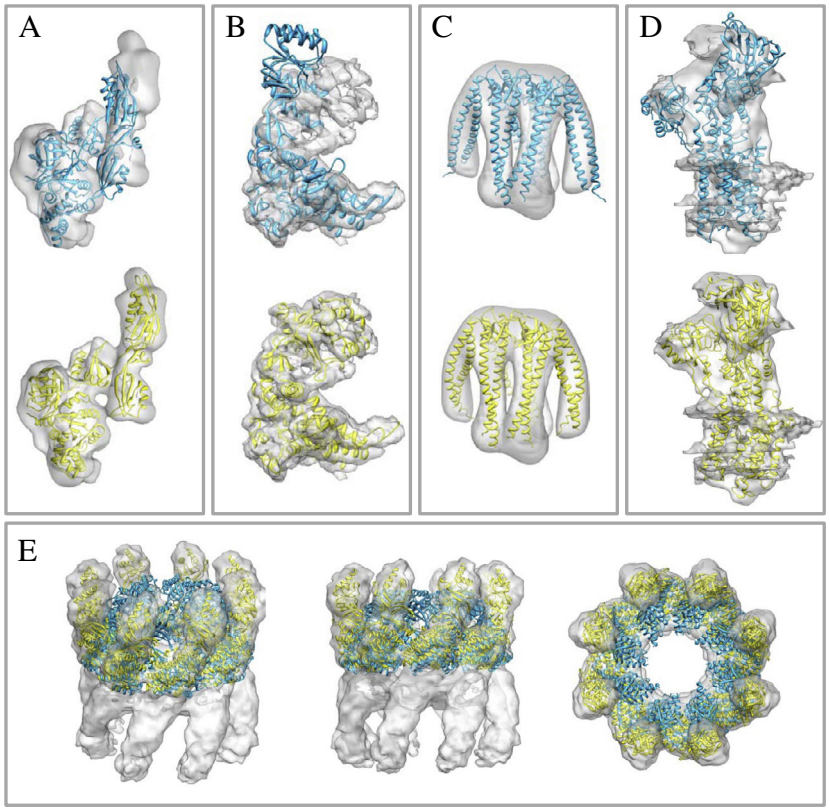

Fig. 2. Flexible fitting of representative experimental maps. The initial pose was obtained by rigid body fitting between the initial atomic structures (cyan) and their corresponding experimental EM maps (transparent) using ADP_EM (Garzon et al., 2007). In panels A–E, the final fitted model is represented with a yellow ribbon. The examples shown include the following: (A) elongation factor G (PDB 1FNM, EMD-1364), (B) the GroEL monomer (PDB 1SX4, EMD-1181), (C) prefoldin (PDB 1FXK), (D) calcium ATPase (PDB 1SU4) and (E) the thermosome (PDB 1A6D, EMD-1396). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Please cite this article in press as: Lopéz-Blanco, J.R., Chacón, P. iMODFIT: Efficient and robust flexible fitting based on vibrational analysis in internal coordinates. J. Struct. Biol. (2013), http://dx.doi.org/10.1016/j.jsb.2013.08.010

researchers as flexible fitting targets. All test cases were successfully fitted using iMODFIT, and only two cases required the modification of default parameters. In all cases, the cross-correlation coefficients were significantly improved (Tables S3 and S4), and the visual inspection revealed an excellent overlap of the flexed structures with the experimental densities (representative examples are shown in Fig. 2). A well-known fitting case involving a large conformational change was revealed by the cryo-EM reconstruction of elongation factor G (EFG) bound to a ribosome (Valle et al., 2003). We observed a huge displacement in one of the terminal domains with respect to a large rotation of the intermediate domains (see Fig. 2A and the corresponding animation in the supplementary material). The same motion was previously characterized by other researchers using different fitting techniques (Ahmed et al., 2012; Kovacs et al., 2008; Orzechowski and Tama, 2008; Velazquez-Muriel and Carazo, 2007). We also studied the GroEL open-to-close transition between the nucleotide-bound (R″ state) and the high-affinity (T state) conformations. We employed a 8.7 Å resolution map of the $GroEL_{14}$–ADP_7–$GroES_7$ complex (Ranson et al., 2006) (EMD-1181), which includes two types of GroEL heptameric rings stacked back-to-back. One of the rings is formed by ADP-bound monomers (R″ state), whereas the other is formed by apo-GroEL monomers (T state). The equivalent structure at an atomic resolution is available in the PDB (ID 1SX4). First, we performed a flexible fitting with a single monomer atomic structure in the R″ conformation into a segmented region of the experimental map that roughly included a monomer in the T state. The fitting revealed a large clockwise rotation of the apical domain. In particular, an outside-in motion of the K and L helices eventually led to a reduction of the chaperonin chamber (Fig. 2B). The apical domain twisting motions were further analyzed based on an animation of the actual fitting trajectory, which is included as supplemental material. This conformational change has been demonstrated both experimentally (Clare et al., 2012; Ranson et al., 2006; Roseman et al., 1996; Xu et al., 1997) and computationally (Hyeon et al., 2006; Ma et al., 2000). The comparison between the fitted and the corresponding crystallographic T state structures yielded an RMSD of approximately 3 Å. This deviation was reasonable, considering the resolution of the map and the large conformational change (>13 Å). We next tried the flexible fitting in the opposite direction, i.e., from the T-state atomic structure to the R″ segmented map, but the results were trapped in a local minima with the default parameters. However, well-fitted models (<3 Å from atomic R″ conformation) were obtained by reducing the map resolution to 15 Å and disabling the local orientation refinement for the first iterations. As a result of these changes, we were able to smooth the fitness landscape, thereby facilitating the convergence to the correct fit. The fitting of the prefoldin "jellyfish" structure is shown in Fig. 2C. The characteristic coiled-coil tentacles were better adjusted inside the EM density. This flexibility is related to the ability of this chaperone to bind different substrates (Martin-Benito et al., 2007). We also characterized both the open-to-close and close-to-open transitions of the two Mm-Cpn high-resolution maps from the Cryo-EM modeling challenge by flexible fitting. These large motions were not tested in the challenge. The whole lidless atomic structures were successfully fitted in the corresponding maps. Remarkably, the differences between identical subunits were very small, considering that symmetry restraints

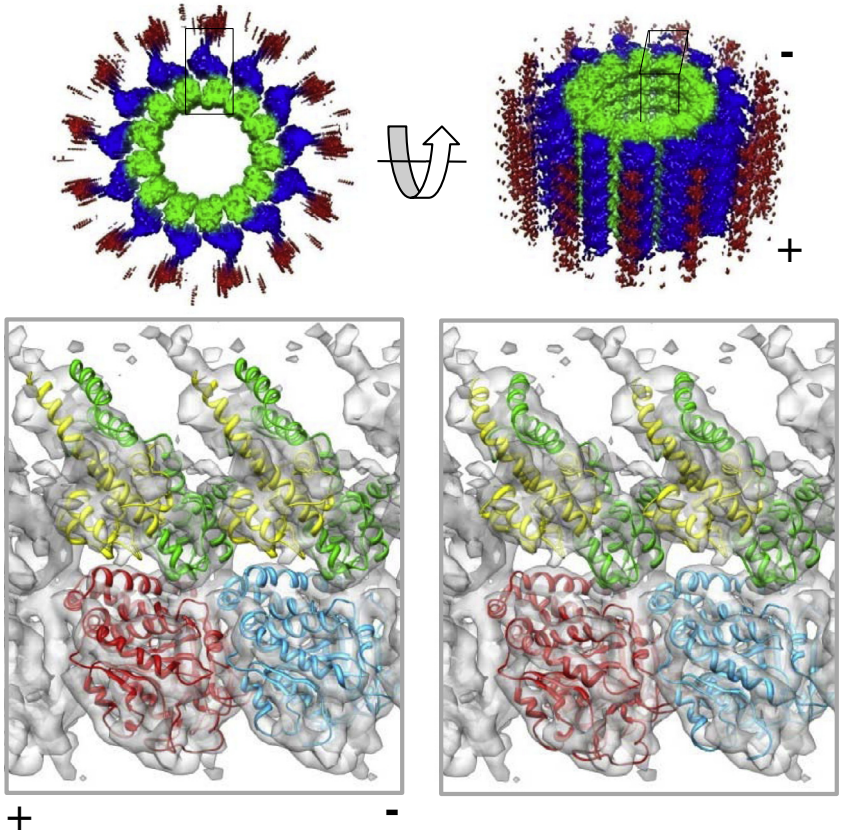

Fig. 3. The top images show a cryo-EM map of a microtubule (green) decorated with the Ndc80/Nuf2 bonsai complex (blue) (PDB 2VE7 and 1JFF, EMD-5223). Poorly defined densities (not considered) are highlighted in red. The lower panels show the initial rigid body fit (left) and the iMODFIT flexible fit (right) of a single tubulin dimer and two Ndc80/Nuf2 complexes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Please cite this article in press as: Lopéz-Blanco, J.R., Chacón, P. iMODFIT: Efficient and robust flexible fitting based on vibrational analysis in internal coordinates. J. Struct. Biol. (2013), http://dx.doi.org/10.1016/j.jsb.2013.08.010

Table 4
Averages obtained using iMODFIT, NMFF and YUPSCX with experimental maps.

	CC[a]	Clashes (‰)	Rama (%)	Lengths (%)	Angles (%)
Initial	0.50	32 ± 27	0.7 ± 0.8	0.2 ± 0.2	0.4 ± 0.4
iMODFIT	0.84	49 ± 32	1.1 ± 1.0	0.3 ± 0.3	0.5 ± 0.4
NMFF	0.82	40 ± 29	1.2 ± 1.0	10.9 ± 12.7	9.6 ± 11.9
YUPSCX	0.86	33 ± 14	0.9 ± 1.0	0.7 ± 0.8	3.8 ± 4.0

[a] Normalized cross-correlation between experimental target maps and the final fitted structures (Eq. (1)). See the footnote of Table 2 for abbreviations. The cases where either NMFF or YUPCSX failed to yield results were removed from the averages (see Table S4).

were not used. The difference between the symmetric subunits was 0.22 and 0.45 Å RMSD for the 4.3 Å (closed) and 8 Å (open) resolution maps, respectively. Fig. 2E shows another impressive fitting of an octameric ring of a thermosome inside a 10 Å cryo-EM map (see also animation in supplementary data). Again, this fitting was in full agreement with the previous results (Clare et al., 2008).

As another example of effective fitting, the results of Hinsen and collaborators (Hinsen et al., 2005) with Ca-ATPase were reproduced. Our flexible fit also revealed large cytoplasmic domain motions and small helix arrangements in the transmembrane region (Fig. 2D). In another pump with poorer resolution, the copper pump (CopAdC), many of the transmembrane helices were compressed and collapsed inside the density map using the default fitting parameters. However, this distortion was easily corrected using 2% of the modes (~60 modes) rather than the default 5% (supplementary Figure S1) and by fixing sheets and helices. This result is a good example of how simple constraints can prevent overfitting problems caused by missing density in EM maps. Our experimental tests also included a cryo-EM map of the Ndc80/Nuf2 kinetochore microtubule complex at a resolution of 8 Å (Alushin et al., 2010) (see Fig. 3, top panel). In this case, the original rigid body fittings of the Ncd80/Nuf2 and tubulin dimer structures were visually improved, as shown in the bottom panel. These results illustrate the ability of iMODFIT to perform the fitting without requiring detailed map segmentation. Here, a generous box selection around the target densities was sufficient.

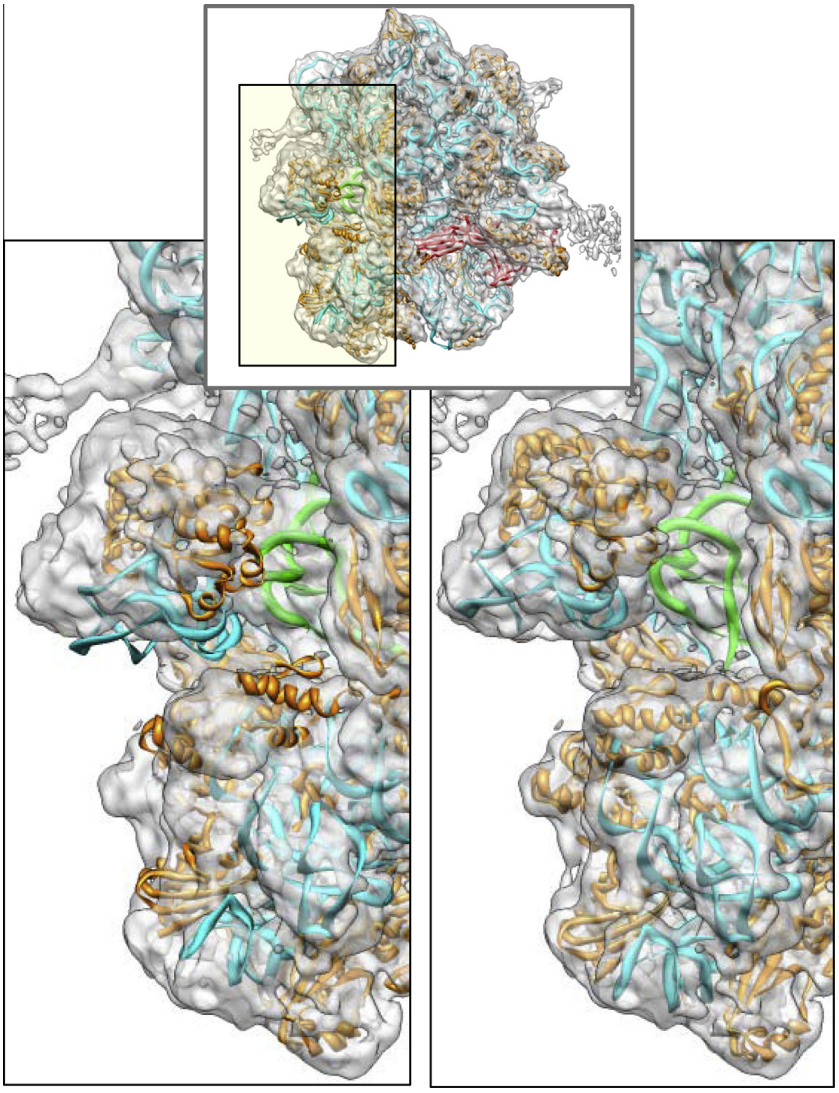

Fig. 4. Ribosome flexible fitting. On the left, the initial rigid-body fitting of the TIpre atomic structure is shown inside an EM map of the TIpost conformation (EMD-1799) with a resolution of 7.6 Å. On the right, the final fitted results obtained with iMODFIT are shown. The bottom images represent the corresponding zoomed regions of the 30S head and L1, where major differences were observed. The color codes are orange for protein, blue for rRNA, green for tRNA and red for EFG. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Please cite this article in press as: Lopéz-Blanco, J.R., Chacón, P. iMODFIT: Efficient and robust flexible fitting based on vibrational analysis in internal coordinates. J. Struct. Biol. (2013), http://dx.doi.org/10.1016/j.jsb.2013.08.010

Similar to the simulated cases, we repeated the whole set of experimental fittings using NMFF and YUPSCX as representative methods. These results are summarized in Table 4. In principle, we observed an overall convergence to the same final fitted model, and the final correlation values were quite similar. YUPSCX presented slightly higher correlation values than did iMODFIT. Similar to the simulated map cases, NMFF suffered from large distortions in both bond lengths and bond angles (>10%). For YUPSCX, the deviation was ~4% for the angles, whereas iMODFIT preserved the initial structural quality. Only a few extra clashes, typically involving side-chain hydrogen atoms, and a few outliers from the allowed regions of the Ramachandran plot were detected with our tool (Table 4). In general, our approach was more robust than YUPSCX and NMFF, which did not even complete the fitting in a few cases.

A detailed pairwise comparison between the tested fitting methods revealed that the convergence was greater, with RMSD differences of approximately 1–2 Å, at higher resolutions ($\leqslant$10 Å) (see Table S5). This discrepancy increased as the resolution decreased or as the motion amplitude increased. Nevertheless, all fitted models remained close to 3 Å. On average, the NMFF models were the most divergent because they contained more distortions. In the challenging case of GroEL, YUPSCX did not fall within the local minima between the R″ and T state transition; however, in the opposite direction, small distortions became apparent. In fact, this method seemed to be more sensitive to lower resolution and the presence of extra densities. For example, in the case of the high-resolution eEF2, some of the loops were displaced towards nearby ribosome density.

Finally, to illustrate iMODFIT's capability to fit huge macromolecular structures composed of proteins, nucleic acids and small ligands, we conducted the challenging flexible fitting of the whole ribosome (~2.5 MDa) in two conformations. As shown in Fig. 4, we reproduced the fitting of the TIpre state atomic coordinates into a TIpost conformation cryo-EM map at a resolution of 7.6 Å (EMD-1799). This fitting was previously performed by Whitford et al. using MDfit, a state-of-the-art MD-based flexible fitting method (Whitford et al., 2011). First, the TIpre atomic model was fitted using a rigid-body approximation into the TIpost EM map. This initial fitting revealed clear differences in the 30S L1 region and in the tRNA (left panel on Fig. 4), although such differences disappeared after flexible fitting with iMODFIT. The final fitted structure agreed very well with the experimental map density (right panel on Fig. 4). The observed fitting trajectory revealed a collective motion involving a swivel motion of the head and a large displacement of the L1 region, both related to tRNA translocation (see the corresponding fitting trajectory in the supplemental material). These structural arrangements of ribosomal translocation intermediates have previously been well characterized and modeled (Agirrezabala et al., 2012; Ratje et al., 2010). It is also worth noting that our fitted model, obtained with a single PC box, differed only by 1 Å from the corresponding MDfit model obtained at supercomputing facilities. Additionally, similar conclusions were drawn by performing the fitting in the opposite direction using the TIpre map (data not shown).

4. Discussion and conclusions

Normal modes in IC provide an excellent mechanism to introduce flexibility in the fitting of atomic structures into electron density maps. Using the low-frequency essential NMA space, iMODFIT successfully fitted a representative set of 46 simulated and 28 experimental test cases. These results revealed the efficiency, accuracy and robustness of this method over the full range of EM experimental resolutions. Working in IC reduced the nonphysical geometric distortions of previous NMA approaches in Cartesian coordinates, and constraining the conformational space to low-frequency modes prevented unrealistic or local motions beyond the map resolution. These two key features of iMODFIT generally led to better fits than did other methods. This was especially evident at medium to low resolutions (<10 Å), where our approximation obtained better fits while maintaining the crystallographic quality of the initial models. It is worth noting that currently more than 85% of the maps deposited in EMDB have resolutions lower than 10 Å. One major concern of flexible fitting is overfitting. Poor resolution, map defects (e.g., scale inaccuracies, missing densities, resolution anisotropy) or partial correspondence between the map and the atomic structure can lead to distorted models. Flexible fitting with a reduced number of modes in internal coordinates provides a natural way to alleviate such overfitting distortions. However, low-frequency modes encode global rearrangements rather than local changes (side chain motions, loop rearrangements or small backbone displacements). Thus, at very high EM resolutions (5–8 Å), MD-based atomistic approaches could be more appropriate. In turn, these atomistic methods become limited at more challenging lower resolutions or larger motions (Vashisth et al., 2012) where iMODFIT is more robust. As already suggested by Trabuco et al. (2008), a hybrid approach using MD-based simulations to refine NMA fitted models may represent an interesting and efficient alternative, especially when large conformational changes occur.

One of the major advantages of our method is that it is extremely easy to use because it does not require any *ad hoc* inputs or elaborated preprocessing steps. It is highly customizable, and the user can control all of the fitting parameters, select CG representation or fix parts of the molecule. Moreover, the computational cost is quite low, and only a few minutes are needed to perform the fitting from raw data on a standard PC. The efficiency gain obtained by merging internal coordinates and CG models extended the applicability of iMODFIT to include large macromolecules. This feature is particularly useful when working with the megadalton targets regularly characterized by cryo-EM. Thus, we believe our novel approach represents a key contribution to the automatic interpretation of the collective conformational changes of large complexes and their functional implications at the atomic level. As a proof of iMODFIT's usefulness, different groups have already obtained impressive flexible fittings with ATP synthase (Lau and Rubinstein, 2012), yeast vacuolar ATPase (Oot et al., 2012) and coxsackievirus (Seitsonen et al., 2012). Future research should expand the method's versatility to handle different fitting scores and search algorithms. We are actively working to extend it to even larger systems (López-Blanco et al., 2013).

Acknowledgments

This study was supported by BFU2009-09552, CAM S2010/BMD-2353 and by the Human Frontier Science Program – RGP0039/2008.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jsb.2013.08.010.

References

Agirrezabala, X., Liao, H.Y., Schreiner, E., Fu, J., Ortiz-Meoz, R.F., Schulten, K., Green, R., Frank, J., 2012. Structural characterization of mRNA–tRNA translocation intermediates. Proc. Natl. Acad. Sci. USA 109, 6094–6099.

Ahmed, A., Tama, F., 2013. Consensus among multiple approaches as a reliability measure for flexible fitting into cryo-EM data. J. Struct. Biol. 182, 67–77.

Ahmed, A., Whitford, P.C., Sanbonmatsu, K.Y., Tama, F., 2012. Consensus among flexible fitting approaches improves the interpretation of cryo-EM data. J. Struct. Biol. 177, 561–570.

Alushin, G.M., Ramey, V.H., Pasqualato, S., Ball, D.A., Grigorieff, N., Musacchio, A., Nogales, E., 2010. The Ndc80 kinetochore complex forms oligomeric arrays along microtubules. Nature 467, 805–810.

Bahar, I., Lezon, T.R., Bakan, A., Shrivastava, I.H., 2010. Normal Mode Analysis of biomolecular structures: functional mechanisms of membrane proteins. Chem. Rev. 110, 1463–1497.

Bray, J.K., Weiss, D.R., Levitt, M., 2011. Optimized torsion-angle normal modes reproduce conformational changes more accurately than cartesian modes. Biophys. J. 101, 2966–2969.

Brooks, B., Karplus, M., 1985. Normal modes for specific motions of macromolecules: application to the hinge-bending mode of lysozyme. Proc. Natl. Acad. Sci. USA 82, 4995–4999.

Chacón, P., Wriggers, W., 2002. Multi-resolution contour-based fitting of macromolecular structures. J. Mol. Biol. 317, 375–384.

Chan, K.Y., Trabuco, L.G., Schreiner, E., Schulten, K., 2012. Cryo-electron microscopy modeling by the molecular dynamics flexible fitting method. Biopolymers 97, 678–686.

Chen, J.Z., Fürst, J., Chapman, M.S., Grigorieff, N., 2003. Low-resolution structure refinement in electron microscopy. J. Struct. Biol. 144, 144–151.

Chen, V.B., Arendall 3rd, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C., 2010. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D 66, 12–21.

Clare, D.K., Stagg, S., Quispe, J., Farr, G.W., Horwich, A.L., Saibil, H.R., 2008. Multiple states of a nucleotide-bound group 2 chaperonin. Structure 16, 528–534.

Clare, D.K., Vasishtan, D., Stagg, S., Quispe, J., Farr, G.W., Topf, M., Horwich, A.L., Saibil, H.R., 2012. ATP-triggered conformational changes delineate substrate-binding and -folding mechanics of the GroEL chaperonin. Cell 149, 113–123.

Cui, Q., Bahar, I., 2007. Normal Mode Analysis Theoretical and Applications to Biological and Chemical Systems. Chapman and Hall, CRC.

Darst, S.A., Opalka, N., Chacon, P., Polyakov, A., Richter, C., Zhang, G., Wriggers, W., 2002. Conformational flexibility of bacterial RNA polymerase. Proc. Natl. Acad. Sci. USA 99, 4296–4301.

Fabiola, F., Chapman, M.S., 2005. Fitting of high-resolution structures into electron microscopy reconstruction images. Structure 13, 389–400.

Flores, S., Echols, N., Milburn, D., Hespenheide, B., Keating, K., Lu, J., Wells, S., Yu, E.Z., Thorpe, M., Gerstein, M., 2006. The database of macromolecular Motions: new features added at the decade mark. Nucleic Acids Res. 34, D296–301.

Garzón, J.I., Kovacs, J., Abagyan, R., Chacón, P., 2007. ADP_EM: fast exhaustive multi-resolution docking for high-throughput coverage. Bioinformatics 23, 427–433.

Go, N., Noguti, T., Nishikawa, T., 1983. Dynamics of a small globular protein in terms of low-frequency vibrational modes. Proc. Natl. Acad. Sci. USA 80, 3696–3700.

Grubisic, I., Shokhirev, M.N., Orzechowski, M., Miyashita, O., Tama, F., 2010. Biased coarse-grained molecular dynamics simulation approach for flexible fitting of X-ray structure into cryo electron microscopy maps. J. Struct. Biol. 169, 95–105.

Hinsen, K., Reuter, N., Navaza, J., Stokes, D.L., Lacapere, J.J., 2005. Normal mode-based fitting of atomic structure into electron density maps: application to sarcoplasmic reticulum Ca-ATPase. Biophys. J. 88, 818–827.

Hyeon, C., Lorimer, G.H., Thirumalai, D., 2006. Dynamics of allosteric transitions in GroEL. Proc. Natl. Acad. Sci. USA 103, 18939–18944.

Jolley, C.C., Wells, S., Fromme, P., Thorpe, M.F., 2008. Fitting low-resolution cryo-EM maps of proteins using constrained geometric simulations. Biophys. J. 94, 1613–1621.

Kovacs, J.A., Cavasotto, C.N., Abagyan, R., 2005. Conformational sampling of protein flexibility in generalized coordinates: application to ligand docking. J. Comput. Theor. Nanosci. 2, 354–361.

Kovacs, J.A., Yeager, M., Abagyan, R., 2008. Damped-dynamics flexible fitting. Biophys. J. 95, 3192–3207.

Lau, W.C., Rubinstein, J.L., 2012. Subnanometre-resolution structure of the intact *Thermus thermophilus* H+-driven ATP synthase. Nature 481, 214–218.

Levitt, M., Sander, C., Stern, P.S., 1985. Protein normal-mode dynamics: trypsin inhibitor, crambin, ribonuclease and lysozyme. J. Mol. Biol. 181, 423–447.

López-Blanco, J.R., Garzón, J.I., Chacón, P., 2011. IMod: multipurpose Normal Mode Analysis in internal coordinates. Bioinformatics 27, 2843–2850.

López-Blanco, J.R., Reyes, R., Aliaga, J.I., Badia, R.M., Chacón, P., Quintana-Ortí, E.S., 2013. Exploring large macromolecular functional motions on clusters of multicore processors. J. Comput. Phys. 246, 275–288.

Lu, M., Poon, B., Ma, J., 2006. A new method for coarse-grained elastic normal-mode analysis. J. Chem. Theory Comput. 2, 464–471.

Ludtke, S.J., Lawson, C.L., Kleywegt, G.J., Berman, H., Chiu, W., 2012. The 2010 cryo-EM modeling challenge. Biopolymers 97, 651–654.

Ma, J., Sigler, P.B., Xu, Z., Karplus, M., 2000. A dynamic model for the allosteric mechanism of GroEL. J. Mol. Biol. 302, 303–313.

Martin-Benito, J., Gomez-Reino, J., Stirling, P.C., Lundin, V.F., Gomez-Puertas, P., Boskovic, J., Chacón, P., Fernandez, J.J., Berenguer, J., Leroux, M.R., Valpuesta, J.M., 2007. Divergent substrate-binding mechanisms reveal an evolutionary specialization of eukaryotic prefoldin compared to its archaeal counterpart. Structure 15, 101–110.

Mendez, R., Bastolla, U., 2010. Torsional network model: normal modes in torsion angle space better correlate with conformation changes in proteins. Phys. Rev. Lett. 104, 228103–228107.

Mitra, K., Schaffitzel, C., Shaikh, T., Tama, F., Jenni, S., Brooks Iii, C.L., Ban, N., Frank, J., 2005. Structure of the *E. coli* protein-conducting channel bound to a translating ribosome. Nature 438, 318–324.

Noguti, T., Go, N., 1983. Dynamics of native globular proteins in terms of dihedral angles. J. Phys. Soc. Jpn. 52, 3283–3288.

Oot, R.A., Huang, L.S., Berry, E.A., Wilkens, S., 2012. Crystal structure of the yeast vacuolar ATPase heterotrimeric EGC(head) peripheral stalk complex. Structure 20, 1881–1892.

Opalka, N., Chlenov, M., Chacón, P., Rice, W.J., Wriggers, W., Darst, S.A., 2003. Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase. Cell 114, 335–345.

Orzechowski, M., Tama, F., 2008. Flexible fitting of high-resolution X-ray structures into cryoelectron microscopy maps using biased molecular dynamics simulations. Biophys. J. 95, 5692–5705.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 2007. Numerical recipes, . third ed. The Art of Scientific Computing third ed. Cambridge University Press.

Ranson, N.A., Clare, D.K., Farr, G.W., Houldershaw, D., Horwich, A.L., Saibil, H.R., 2006. Allosteric signaling of ATP hydrolysis in GroEL–GroES complexes. Nat. Struct. Mol. Biol. 13, 147–152.

Ratje, A.H., Loerke, J., Mikolajka, A., Brünner, M., Hildebrand, P.W., Starosta, A.L., Dönhöfer, A., Connell, S.R., Fucini, P., Mielke, T., Whitford, P.C., Onuchic, J.N., Yu, Y., Sanbonmatsu, K.Y., Hartmann, R.K., Penczek, P.A., Wilson, D.N., Spahn, C.M.T., 2010. Head swivel on the ribosome facilitates translocation by means of intra-subunit tRNA hybrid sites. Nature 468, 713–716.

Roseman, A.M., Chen, S., White, H., Braig, K., Saibil, H.R., 1996. The chaperonin ATPase cycle: mechanism of allosteric switching and movements of substrate-binding domains in GroEL. Cell 87, 241–251.

Rusu, M., Birmanns, S., Wriggers, W., 2008. Biomolecular pleiomorphism probed by spatial interpolation of coarse models. Bioinformatics 24, 2460–2466.

Schröder, G.F., Brunger, A.T., Levitt, M., 2007. Combining efficient conformational sampling with a deformable elastic network model facilitates structure refinement at low resolution. Structure 15, 1630–1641.

Seitsonen, J.J., Shakeel, S., Susi, P., Pandurangan, A.P., Sinkovits, R.S., Hyvonen, H., Laurinmaki, P., Yla-Pelto, J., Topf, M., Hyypia, T., Butcher, S.J., 2012. Structural analysis of coxsackievirus A7 reveals conformational changes associated with uncoating. J. Virol. 86, 7207–7215.

Siebert, X., Navaza, J., 2009. UROX 2.0: an interactive tool for fitting atomic models into electron-microscopy reconstructions. Acta Crystallogr. D 65, 651–658.

Skjaerven, L., Hollup, S.M., Reuter, N., 2009. Normal Mode Analysis for proteins. J. Mol. Struct. (Theochem) 898, 42–48.

Suhre, K., Navaza, J., Sanejouand, Y.H., 2006. NORMA: a tool for flexible fitting of high-resolution protein structures into low-resolution electron-microscopy-derived density maps. Acta Crystallogr. D 62, 1098–1100.

Tama, F., Brooks, C.L., 2006. Symmetry, form, and shape: guiding principles for robustness in macromolecular machines. Annu. Rev. Biophys. Biomol. Struct. 35, 115–133.

Tama, F., Miyashita, O., Brooks 3rd, C.L., 2004a. Flexible multi-scale fitting of atomic structures into low-resolution electron density maps with elastic network Normal Mode Analysis. J. Mol. Biol. 337, 985–999.

Tama, F., Miyashita, O., Brooks 3rd, C.L., 2004b. Normal mode based flexible fitting of high-resolution structure into low-resolution experimental data from cryo-EM. J. Struct. Biol. 147, 315–326.

Tama, F., Ren, G., Brooks 3rd, C.L., Mitra, A.K., 2006. Model of the toxic complex of anthrax: responsive conformational changes in both the lethal factor and the protective antigen heptamer. Protein Sci. 15, 2190–2200.

Tan, R.K.Z., Devkota, B., Harvey, S.C., 2008. YUP.SCX: coaxing atomic models into medium resolution electron density maps. J. Struct. Biol. 163, 163–174.

Tirion, M.M., 1996. Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. Phys. Rev. Lett. 77, 1905–1908.

Topf, M., Sali, A., 2005. Combining electron microscopy and comparative protein structure modeling. Curr. Opin. Struct. Biol. 15, 578–585.

Topf, M., Lasker, K., Webb, B., Wolfson, H., Chiu, W., Sali, A., 2008. Protein structure fitting and refinement guided by cryo-EM density. Structure 16, 295–307.

Trabuco, L.G., Villa, E., Mitra, K., Frank, J., Schulten, K., 2008. Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. Structure 16, 673–683.

Trabuco, L.G., Villa, E., Schreiner, E., Harrison, C.B., Schulten, K., 2009. Molecular dynamics flexible fitting: a practical guide to combine cryo-electron microscopy and X-ray crystallography. Methods 49, 174–180.

Valle, M., Zavialov, A., Sengupta, J., Rawat, U., Ehrenberg, M., Frank, J., 2003. Locking and unlocking of ribosomal motions. Cell 114, 123–134.

Vashisth, H., Skiniotis, G., Brooks 3rd, C.L., 2012. Using enhanced sampling and structural restraints to refine atomic structures into low-resolution electron microscopy maps. Structure 20, 1453–1462.

Velazquez-Muriel, J.A., Carazo, J.M., 2007. Flexible fitting in 3D-EM with incomplete data on superfamily variability. J. Struct. Biol. 158, 165–181.

Villa, E., Sengupta, J., Trabuco, L.G., LeBarron, J., Baxter, W.T., Shaikh, T.R., Grassucci, R.A., Nissen, P., Ehrenberg, M., Schulten, K., Frank, J., 2009. Ribosome-induced changes in elongation factor Tu conformation control GTP hydrolysis. Proc. Natl. Acad. Sci. USA 106, 1063–1068.

Wang, Z., Schroder, G.F., 2012. Real-space refinement with DireX: from global fitting to side-chain improvements. Biopolymers 97, 687–697.

Whitford, P.C., Geggier, P., Altman, R.B., Blanchard, S.C., Onuchic, J.N., Sanbonmatsu, K.Y., 2010. Accommodation of aminoacyl-tRNA into the ribosome involves reversible excursions along multiple pathways. RNA 16, 1196–1204.

Please cite this article in press as: Lopéz-Blanco, J.R., Chacón, P. iMODFIT: Efficient and robust flexible fitting based on vibrational analysis in internal coordinates. J. Struct. Biol. (2013), http://dx.doi.org/10.1016/j.jsb.2013.08.010

Whitford, P.C., Ahmed, A., Yu, Y., Hennelly, S.P., Tama, F., Spahn, C.M.T., Onuchic, J.N., Sanbonmatsu, K.Y., 2011. Excited states of ribosome translocation revealed through integrative molecular modeling. Proc. Natl. Acad. Sci. USA 108, 18943–18948.

Wriggers, W., Chacón, P., 2001. Modeling tricks and fitting techniques for multiresolution structures. Structure 9, 779–788.

Wriggers, W., Chacón, P., Kovacs, J.A., Tama, F., Birmanns, S., 2004. Topology representing neural networks reconcile biomolecular shape, structure, and dynamics. Neurocomputing 56, 365–379.

Xu, Z., Horwich, A.L., Sigler, P.B., 1997. The crystal structure of the asymmetric GroEL–GroES–(ADP)7 chaperonin complex. Nature 388, 741–750.

Yamashita, H., Endo, S., Wako, H., Kidera, A., 2001. Sampling efficiency of molecular dynamics and Monte Carlo method in protein simulation. Chem. Phys. Lett. 342, 382–386.

Zheng, W., 2011. Accurate flexible fitting of high-resolution protein structures into cryo-electron microscopy maps using coarse-grained pseudo-energy minimization. Biophys. J. 100, 478–488.

Zhu, J., Cheng, L., Fang, Q., Zhou, Z.H., Honig, B., 2010. Building and refining protein models within cryo-electron microscopy density maps based on homology modeling and multiscale structure refinement. J. Mol. Biol. 397, 835–851.

Please cite this article in press as: Lopéz-Blanco, J.R., Chacón, P. iMODFIT: Efficient and robust flexible fitting based on vibrational analysis in internal coordinates. J. Struct. Biol. (2013), http://dx.doi.org/10.1016/j.jsb.2013.08.010