DockAFM: benchmarking protein structures by docking under AFM topographs

Rui C. Chaves* and Jean-Luc Pellequer*

CEA, iBEB, Service de Biochimie et Toxicologie Nucléaire, F-30207 Bagnols sur Cèze, France

Associate Editor: Anna Tramontano

ABSTRACT

Summary: Proteins can adopt a variety of conformations. We present a simple server for scoring the agreement between 3D atomic structures and experimental envelopes obtained by atomic force microscopy. Three different structures of immunoglobulins (IgG) or blood coagulation factor V activated were tested and their agreement with several topographical surfaces was computed. This approach can be used to test structural variability within a family of proteins.

Availability and implementation: DockAFM is available at http://biodev.cea.fr/dockafm.

Contact: chaves.rui.c@gmail.com or jlpellequer@cea.fr

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 BACKGROUND

A new family of microscopy, atomic force microscopy (AFM), enables the probing of the surface of a single molecule in physiological conditions using a sharp nanoscopic tip (Binnig *et al.*, 1986). Essentially, an AFM height image of a single molecule represents the surface topography of that molecule. AFM imaging yields 2D images with an exceptional signal/noise ratio that allows the observation of single molecules at a lateral resolution of nanometers and a vertical resolution of angstrom (Ido *et al.*, 2013; Schabert *et al.*, 1995). AFM has intrinsic advantages when applied to large proteins, which are difficult to study by traditional techniques (Schroder *et al.*, 2010; Trinh *et al.*, 2012). AFM can handle imaging in liquid compared with the frozen or crystallized state in electronic microscopy (EM) and X-ray diffraction, respectively. In addition, AFM does not rely on symmetry averaging as in traditional EM or on physical averaging as in crystallized sample (Fechner *et al.*, 2009).

AFM imaging allows the observation of the conformational dynamics of single molecules such as membrane proteins (Colom *et al.*, 2012; Scheuring *et al.*, 1999), nucleic acids (Ido *et al.*, 2013; Witz *et al.*, 2011) and on antibodies (Chaves *et al.*, 2013). Such dynamics is also recently observed with high-speed AFM (Ando *et al.*, 2007; Casuso *et al.*, 2012). Thus, it is possible to distinguish putative conformational states of single proteins for instance. However, despite the exceptional signal/noise ratio of high-resolution AFM images, it is not straightforward to interpret atomic changes in AFM images (Chen *et al.*, 2013).

*To whom correspondence should be addressed.

The DockAFM tool establishes a link between topographic images from AFM and molecular dynamics of single proteins. DockAFM computes the fit of input conformations of a molecule with the topographic surface of AFM images. Thus, DockAFM can be used to benchmark protein 3D structures or models against an experimental data obtained by atomic force microscopy. DockAFM uses a real-space description of atoms and surfaces and has been developed as the first step to assemble large macromolecules using their individual constituent (Chaves *et al.*, 2013; Trinh *et al.*, 2012) and AFM images.

2 IMPLEMENTATION

DockAFM uses DOT 2.0 (Roberts *et al.*, 2013) that has demonstrated useful applications on studying protein–protein interactions (Roberts and Pique, 1999) and protein-DNA (Fan and Roberts, 2006). DOT 2.0 allows different grid spacing in the three dimensions. It is well adapted to AFM images using a grid spacing in XY plane of 10 Å and an improved grid spacing in Z-direction of 1 Å.

A typical DockAFM run takes 5 min on the server (64 matrix side). Optionally, DockAFM can erode the AFM image to correct for tip shape artifact but the user must enter appropriate tip shape values (usually taken from the manufacturer). The AFM image is then transformed into favorable and forbidden layers. The docking is performed at a constant null electrostatic potential so that only DOT van der Waals interactions is counted for atoms residing inside the favorable region. Thus, the top ranking solution has the minimum energy that corresponds to the maximum number of atoms located within the favorable layer beneath the AFM surface.

Upon completion of the calculation, the server generates and returns the lowest energy conformation in a Protein Data Bank (PDB) file and a ranked list of 10 000 docking possibilities that can be received either by email or by downloading the results in a zip file. A good docking is obtained when the largest number of atoms is fitted within the favorable layer and when most of top docking orientations are clustered in the topographic surface.

3 RESULTS

Two examples are presented in Figure 1. In the first case, AFM images of single molecules of antibodies were obtained. The question asked to DockAFM server was to show which known 3D structures of complete antibodies (PDB codes for T-shape: 1IGT, Y-shape: 1IGY or intermediate: 1HZH) fit better the AFM images. In the second case, AFM image of human blood coagulation factor V activated (FVa) was obtained.
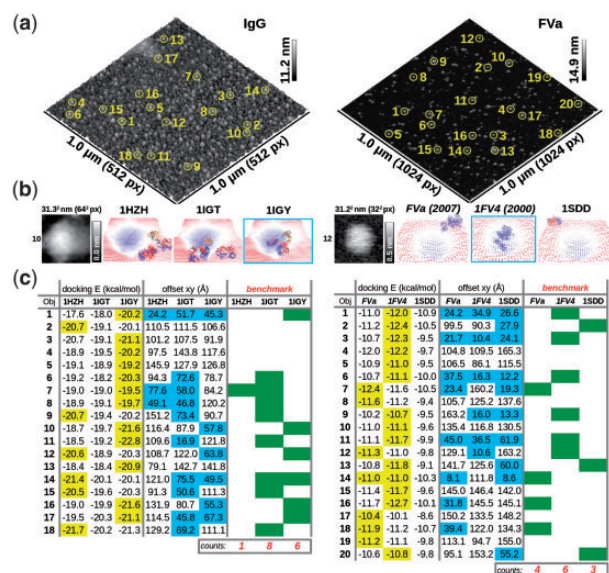
Fig. 1. DockAFM test cases: immunoglobulin (IgG) and activated blood coagulation factor V (FVa): (**a**) AFM images of IgG (left) and FVa (right) show each individual topographic surface used for running DockAFM (**b**) AFM topographic surfaces used for docking protein structures are shown in gray scale as well as three different docked structures are shown for both IgG and FVa systems (bottom side view) (**c**) For each system, the docking score as well as the shift of the docked structure from the center of the topographic surface are shown. The most favorable structure is chosen as that having the smallest shift from the center. Counting most favorable structures is used for the benchmarking purpose of DockAFM

Because no complete structure of FVa exists, the question asked to DockAFM is to show which of the three available model of FVa fits better the AFM image. The three models used for FVa are a comparative model, 1FV4 (Pellequer *et al.*, 2000), a partial crystal structure, 1SDD (Adams *et al.*, 2004) and a more recent completed FVa model, which uses the C domains orientations according to the partial crystal structure, FVa (Gale *et al.*, 2007).

Both AFM images (Fig. 1a) were obtained on a multimode V microscope using the PeakForce tapping mode in air (Bruker, AXS). The IgG image was scaled four times by linear interpolation to a resolution of 4.9 Å/px using Gwyddion, whereas the FVa image was used at the raw resolution of 9.75 Å/px. For the docking, a cubic matrix of 64 nodes with a step of 4.9 Å and a favorable layer of 12 Å was used for IgG, whereas a cubic matrix of 32 nodes with a step of 9.75 Å and a favorable layer of 20 Å was used for FVa. Among the multiple individual molecules on the AFM images, several were cropped and selected. Each tested structure was docked individually in each crop image. Results for a single crop in both molecular systems are shown in Figure 1b, full information is available in Supplementary Material (Supplementary Figs S1 and S2).

The output of DockAFM is indicated in Figure 1c. For each crop and for each structure, docking energy and the offset relative to the center is given. Best docking solutions are highlighted in yellow (first column). Docking solutions having acceptable offsets are highlighted in cyan (second column), i.e. those with a maximum deviation between the center of the docked molecule and that of the topographic image was <25% of the image lateral

size. Finally, the most probable fitted structures are highlighted in green (third column) and correspond to those having the lowest energy with well centering into the topographic surface. In the two cases presented here, the structure 1IGT is the most probable conformation found in the IgG image, whereas the model 1FV4 fits better the FVa AFM image.

4 CONCLUSIONS

The DockAFM server is a tool for docking 3D structures under AFM surfaces. It can be used to benchmark alternative 3D structures or models against experimental data.

REFERENCES

Adams,T.E. *et al.* (2004) The crystal structure of activated protein C-inactivated bovine factor Va: implications for cofactor function. *Proc. Natl Acad. Sci. USA*, **101**, 8918–8923.

Ando,T. *et al.* (2007) High-speed atomic force microscopy for observing dynamic biomolecular processes. *J. Mol. Recognit.*, **20**, 448–458.

Binnig,G. *et al.* (1986) Atomic force microscope. *Phys. Rev. Lett.*, **56**, 930–933.

Casuso,I. *et al.* (2012) Characterization of the motion of membrane proteins using high-speed atomic force microscopy. *Nat. Nanotechnol.*, **7**, 525–529.

Chaves,R.C. *et al.* (2013) Conformational dynamics of individual antibodies using computational docking and AFM. *J. Mol. Recognit.*, **26**, 596–604.

Chen,S.W. *et al.* (2013) Nanoscale structural features determined by AFM for single virus particles. *Nanoscale*, doi: 10.1039/c3nr02706f.

Colom,A. *et al.* (2012) High-speed atomic force microscopy: cooperative adhesion and dynamic equilibrium of junctional microdomain membrane proteins. *J. Mol. Biol.*, **423**, 249–256.

Fan,L. and Roberts,V.A. (2006) Complex of linker histone H5 with the nucleosome and its implications for chromatin packing. *Proc. Natl Acad. Sci. USA*, **103**, 8384–8389.

Fechner,P. *et al.* (2009) Structural information, resolution, and noise in high-resolution atomic force microscopy topographs. *Biophys. J.*, **96**, 3822–3831.

Gale,A.J. *et al.* (2007) Characterization of a factor Xa binding site on factor Va near the Arg506 APC cleavage site. *J. Biol. Chem.*, **282**, 21848–21855.

Ido,S. *et al.* (2013) Beyond the helix pitch: direct visualization of native DNA in aqueous solution. *ACS Nano*, **7**, 1817–1822.

Pellequer,J.L. *et al.* (2000) Three-dimensional model of the coagulation factor Va bound to activated protein C. *Thromb. Haemost.*, **84**, 849–857.

Roberts,V.A. and Pique,M.E. (1999) Definition of the interaction domain for cytochrome c on cytochrome c oxidase. III. Prediction of the docked complex by a complete, systematic search. *J. Biol. Chem.*, **274**, 38051–38060.

Roberts,V.A. *et al.* (2013) DOT2: macromolecular docking with improved biophysical models. *J. Comput. Chem.*, **34**, 1743–1758.

Schabert,F.A. *et al.* (1995) Native *Escherichia coli* OmpF porin surfaces probed by atomic force microscopy. *Science*, **268**, 92–94.

Scheuring,S. *et al.* (1999) High resolution AFM topographs of the *Escherichia coli* water channel aquaporin Z. *EMBO J.*, **18**, 4981–4987.

Schroder,G.F. *et al.* (2010) Super-resolution biomolecular crystallography with low-resolution data. *Nature*, **464**, 1218–1222.

Trinh,M.H. *et al.* (2012) Computational reconstruction of multidomain proteins using atomic force microscopy data. *Structure*, **20**, 113–120.

Witz,G. *et al.* (2011) Conformation of ring polymers in 2D constrained environments. *Phys. Rev. Lett.*, **106**, 248301.