Conformational dynamics of individual antibodies using computational docking and AFM

Rui C Chaves, Jean-marie Teulon, Michael Odorico, Pierre Parot, Shu-wen W Chen, Jean-luc Pellequer

Conformational dynamics of individual antibodies using computational docking and AFM

Rui C. Chaves[a,*], Jean-Marie Teulon[a], Michael Odorico[a], Pierre Parot[a], Shu-wen W. Chen[b], Jean-Luc Pellequer[a,*,+]

[a]CEA, iBEB, Service de Biochimie et Toxicologie Nucléaire, F-30207 Bagnols sur Cèze.

[b]13 Avenue de la Mayre, F-30200 Bagnols sur Cèze, France.

[*]**For correspondence**: chaves.rui.c@gmail.com (RCC); jlpellequer@cea.fr (JLP)

[+]**Address**:

Jean-Luc Pellequer

Institut de Biologie Structurale

71, avenue des Martyrs

CS 10090

38044 Grenoble Cedex 9

Tel : +33 (0)457 42 8756

Fax : +33 (0)476 50 1890

ABSTRACT

Molecular recognition between a receptor and a ligand requires a certain level of flexibility in macromolecules. In this study, we aimed at analyzing the conformational variability of receptors portrayed by monoclonal antibodies which have been individually imaged using Atomic Force Microscopy (AFM). Individual antibodies were chemically coupled to activated mica surface and they have been imaged using AFM in ambient conditions. The resulting topographical surface of antibodies was used to assemble the three subunits constituting antibodies: 2 Fab and 1 Fc using a surface-constrained computational docking approach. Reconstructed structures based on 10 individual topographical surfaces of antibodies are presented and for which separation and relative orientation of the subunits were measured. When compared with three X-ray structures of antibodies present in the PDB database, results indicate that several arrangements of reconstructed subunits are comparable with those of known structures. Nevertheless, no reconstructed structure superimposes adequately to any particular X-ray structure consequence of the antibody flexibility. We conclude that high-resolution AFM imaging with appropriate computational reconstruction tools are adapted to study the conformational dynamics of large individual macromolecules deposited on mica.

Keywords:

Protein structure reconstruction

Structure dynamics

Immunoglobulin G (IgG)

Atomic Force Microscopy (AFM)

Docking

INTRODUCTION

The high specificity of antibody has promoted this protein family to the indispensable role in therapeutics and biotechnology which has been demanding continuous efforts in understanding its structure and dynamics {Tong, 2013 #3371; Correia, 2013 #3336}. High flexibility of antibodies allows the reduction of steric hindrance during recognition {Pellequer, 1993 #1715; Thouvenin, 1997 #976}, and the binding to antigens with different shape and size {Dangl, 1988 #3352; Roux, 1997 #3343; Saphire, 2002 #3342}. In addition, increasing binding affinity has been observed when the antibody linker provided more flexibility in symmetroadhesins (synthetic antibody mimetic) {Capon, 2011 #3387}. Eventually, extreme potency against HIV has been attributed to the structural flexibility of polyvalent antibody constructs such as dodecameric CD4 replacing Fab regions {Bennett, 2007 #3388}. Consequently, an ensemble of antibody conformations is more representative than one static state for antibody molecules toward the understanding of biological activities {James, 2003 #2383; James, 2003 #3373}. This was confirmed from analysis on the Immunoglobulin G (IgG) complete structures determined by X-ray {Burton, 1990 #1429; Harris, 1995 #3344; Harris, 1998 #65; Kuznetsov, 2000 #3346; Larson, 1991 #3347; Saphire, 2001 #3379; Saphire, 2003 #3349; Saphire, 2002 #3342; Stura, 1994 #3350}. X-ray determined structures can be considered as snapshots of a wide range of varying conformations. In computational fields, using a static structure as the basis of molecular docking led to poor successes {Ghemtio, 2013 #3381} whereas taking into account structural variability, a greater success in molecular modeling was observed {Pellequer, 2006 #2483}.

The flexibility of IgG molecules has been studied by NMR {Kim, 1994 #3358}, SAXS {Gregory, 1987 #3359}, and EM {Roux, 1989 #3383; Roux, 1994 #3357; Roux, 1997 #3343; Roux, 1998 #3355; Roux, 1999 #3356; Sandin, 2004 #3448; Correia, 2013 #3336}. The dynamics of antibodies has been investigated by individual-particle electron tomography (IPET) and the optimized negative staining (OpNS) protocol {Tong, 2013 #3371; Zhang, 2012 #3378; Zhang, 2011 #3377; Zhang, 2010 #3376}. However, the current techniques in structural biology are hampered by molecular global flexibility, which is characterized by domain-to-domain movements and intrinsic protein disordering

{Dunker, 2001 #3380}. Therefore, new approaches need to be developed for studying dramatic changes in protein conformations.

Compared to cryo-electron microscopy (Cryo-EM) and small angle X-ray scattering (SAXS), Atomic Force Microscopy (AFM) has intrinsic advantages when applied to large proteins which are difficult to study (large, flexible, partially disordered, hard to purify). AFM can handle a sample concentration down to 1 µg/ml, 3 orders of magnitude lower compared to the demands for X-ray diffraction or NMR. Possible imaging can be performed on a liquid system instead of frozen or crystallized state required for TEM and X-ray diffraction. In addition, AFM does not rely on symmetry averaging as in traditional EM or on physical averaging as in crystallized sample. Finally, the information extracted from AFM images occurs in real space rather than in the Fourier space. AFM topography is obtained with a raster scanning of a tip crossing the image plane and monitoring the vertical position (Z) of tip. The AFM microcantilever, which mounts the probing tip, can passively sense the localized forces between the tip and the specimen surface {Binnig, 1986 #2920}. AFM provides high signal-to-noise ratio (SNR) allowing observations of single molecules at a lateral resolution of about 1 nanometer (nm) and a vertical resolution near 1 Å on proteins {Schabert, 1994 #2922} while recent developments suggest lateral resolution below 1 nm {Ido, 2013 #3447}.

A large range of antibody conformations has been observed with high-resolution AFM images when deposited on a variety of substrates {Yang, 1996 #2976; San Paulo, 2000 #3131; Thomson, 2005 #3247; Martinez, 2008 #3253; Czajkowsky, 2009 #3204}. These antibody images allow the analysis of conformational variability in antibody structures. To study quantitatively such structural variability, it would be useful to reach atomic-level description of antibodies based on AFM images. The relation between high-resolution AFM topography and molecular shape of 3D structure has already been addressed {Asakawa, 2011 #3207; Buzhynskyy, 2009 #3203; Davies, 2005 #2926; Scheuring, 2007 #2679; Scheuring, 2005 #3133; Trinh, 2012 #3182, Chen and Pellequer, 2013?#3385}. A recent computational protocol for reconstructing multidomain proteins has been developed using experimental evidence, i.e., AFM topography {Trinh, 2012

#3182}. In this work, we applied this method to study conformational variability of bio-molecules, as exemplified by imaged antibodies. A conceptual scheme of the semi-empirical strategy is explained in **Figure 1**. In brief, the strategy includes four steps: the first one requires the preparation of structural units of the target macromolecule, here it is two Fabs and one Fc; the second step requires a topographic surface of the complete target molecule, here it is obtained from AFM; the third step allows the docking of structural units beneath the topographic surface; the fourth step consists in assembling from docked positions one Fc domain to two Fab domains.

MATERIALS AND METHODS

Sample preparation and AFM instrumentation

A freshly cleaved mica disk (Muscovite, Bruker) was activated with a solution of 3-Mercaptopropyl triethoxysilane (MPTES) with 1 % (v/v) ethanol/chloroform and incubated overnight under the condition of 60 ºC with a nitrogen saturated atmosphere. The substrate plate (mica disk) was rinsed with 100 % ethanol for 3 min, and then washing drops were progressively diluted with pure water. Later, the substrate plate was completely rinsed in water. The cleaned surface of substrate was modified by reducing MPTES on the surface with 1 mM of Dithiothreitol (DTT) in water for 30 min, and then washed with HEPES buffer (a mixture of HEPES at 10 mM, KCl at 150 mM and 1 mM of EDTA at pH = 7.0). Lastly, the substrate was chemically activated with sulfosuccinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate (Sulfo-SMCC) at 0.5 mM in HEPES buffer during 30 min.

Mouse monoclonal antibody (10 nM in HEPES buffer) was deposited and incubated on the activated substrate plate for 1 hr at room temperature. The supernatant on the substrate plate was removed and the substrate was gently rinsed three times in HEPES buffer, then in water. Prior to AFM imaging, the whole sample was dried using a Laboport vacuum pump (KNF Neuberger, Trenton, NJ). In order to minimize the applied pressures on molecular surfaces, the PeakForce Tapping mode of MultiMode 8 AFM

(Bruker AXS, Santa Barbara) was used to perform the imaging. PeakForce tapping operates in a non-resonant mode by performing rapid force-distance curves at a single pixel. The feed-back control is performed at the force level by interpreting force curves and identifying the peak force value. The ramp size of the force-distance curve was set to 150 nm and the ramping frequency was 2 kHz. The engaging setpoint was set at 0.15 V and the setpoint was continuously monitored and updated using the ScanAsyst interface. A SNL tip with a spring constant of 0.38 N/m (Bruker AFM probes) and a nominal tip radius of 2 nm was attached to the silicon nitride lever for probing the sample surface in air. The raw AFM image was acquired with 512 x 512 pixels in a physical dimension of 1.0 x 1.0 μm^2, see **Figure 2**.

Protocol of structure reconstruction

Reduction of stripe noises was performed on the raw image using *DeStripe* {Chen, 2011 #3015}. A total of 10 image objects, with an averaged height of 5.1 ± 0.8 nm (maximum height 8 nm), was cropped from the parent image, thus 10 child images were generated. These objects were selected based on their lateral size (≤20 nm), with maximal height, and no particular pixel artifacts. They are labeled A-J and highlighted with a yellow box in **Figure 2**. A molecular topography at a fine scale is prerequisite to docking experiments. Therefore, each child image was rebuilt with smaller pixel width, 4.9 Å, which is four times the original probed resolution. The intensity values at newly made pixels were calculated using the Bspline algorithm in Gwyddion {Gwyddion, 2013 #3294} as elaborated in **Figure 3**. Each child image was processed with the morphological erosion {Villarrubia, 1997 #2928} to remove the size effect of tip. The model tip used in erosion was formed of a side wall angle 15°, while the apex diameter (ϕ) was selected as the maximum size that do not jeopardize the maximum height in the eroded molecular topograph {Trinh, 2011 #3014}. Consequently, the respective ϕ values for A-J child images were 30, 46, 44, 30, 38, 34, 38, 46, 30, 42 Å. Regarding the molecular subunits to be assembled, the atomic coordinates of Fab fragments were obtained from PDB code 1AY1, while that of Fc fragment were taken from 1H3T. The

coordinates of missing atoms in the Fc fragment of 1H3T were modeled using the symmetry-related coordinates {Gale, 2007 #2603}.

The docking space was provided with a cubic grid of 256x256x256 grid points whose respective grid spacing is 4.9x4.9x1.0 Å. The image plane was placed in the middle of the grid so that its center locates at the grid center and its base plane parallels the X-Y plane of the grid. Two zones, namely forbidden (B) and favorable (L), sandwiched the AFM topography. Zone B corresponds to the space right above the topography, while zone L is the region beneath with a thickness of 12 Å. The Fab and Fc fragments of antibody were independently docked against the AFM topography. The docking was performed using the program DOT 2.0 {Mandell, 2001 #9; Ten Eyck, 1995 #252}. Each docking displacement includes translation of an antibody fragment from one visiting point to the next and rotation of the fragment with an angular step 6° (54000 different rotations). DOT provides a systematic FFT-based search in translation and rotation for optimum docking solution. Each run of DOT yielded more than $9x10^{11}$ docking results for each antibody fragment. Only top 10^5 of potential solutions were retained for later reconstruction, yet it is still too large for the full combinatorial analysis. On a high-resolution cubic matrix of 256 nodes, typical running time on 12 Xeon E5-2620 CPUs at 2.0 GHz is 150 min.

The combination of docked solutions uses the in-house software, *combine* {Trinh, 2012 #3182}. To reduce the computational time we imposed a range for the distance between subunits. For that, we measured the distance between the C_α coordinates of the residues Pro227 on each Fab and Ser252 on the Fc domain of the known three X-ray structures, PDB codes: 1IGT, 1IGY and 1HZH. The distances obtained were (38, 52), (35, 40), and (35, 45) Å, respectively. Knowing this structural information, we imposed a laxer condition during the assembling of the antibody structure, i.e., 28.0~62.4 Å, while discarding the potential solutions whose Fc-Fab separations were unfitted. To further reduce the computational tasks, we used Adepth {Chen, 2013? #3268} to acquire the skin layer of atoms of each subunit within 3 Å to its surface and then these extruded structures were used in *combine*. At the coarse level, *combine* grouped the 10^5 potential docked

solutions for each antibody subunit into 4000 clusters represented by one reference solution and removed all others which were within a range of translational and rotational variations to the reference. Totally, there were 4000^3 combinations for two Fabs and one Fc fragment. The range values were set to as (5 Å, 6º), (5 Å, 12º), (10 Å, 12º), (20 Å, 12º), or (20 Å, 24º). For the final assembly process a refined (fine) combination further screened out the coarse results for those translational and rotational values differing in the range of 20 Å and 25º (**Figure 4**). A typical run on 12 CPUs is about 10h but varies greatly depending on the selected set-up.

3D structures comparison by RMSD

To compare 3D structures, see **Figure 5** and **Figure 6**, the first method used was based in the root mean square deviation (RMSD) given by:

$$RMSD_{ij} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(d_{n,ij} - \overline{d_{ij}}\right)^2}\ ,\ \overline{d_{ij}} = \sum_{n=1}^{N}\frac{d_{n,ij}}{N}\ ,$$

where, the Euclidean distance between N, Cα, C coordinates of the positions *i* and *j* in the structure *n* is denoted as $d_{n,ij}$. The pseudo-symmetry of Fab domains was taken into account by producing four reference models for each of the three known X-rays structures using a global superimposition with sup3d {Chen, 2004 #1442}. To simplify the superimposition of reconstructed antibodies with those of X-ray crystal structures of complete antibodies, Fab and Fc domains used for the reconstruction were superimposed first on their relative counterpart in the X-ray structures 1IGT, 1IGY, IHZH using conserved disulfide bridges. Then, the computation of RMSDs between reconstructed antibodies and X-ray structures were done with an all-atom calculation since now all the structures have the same number of atoms. The same process was applied to compare the combine ranking (**Figure 5**). Thus, the facing up/down in one or both Fab was included, and the minimum RMSD values among the four models are reported here.

Geometry of the assembled structures

Descriptive vectors were determined for representing dispositions of individual antibody fragments using pairs of conserved disulfide bridges, see **Figure 7**. Each disulfide bridge resides respectively in one of the four Ig domains composing the fragment and the orientation of the vector was defined using the C_α atoms in the residues (R) of the conserved disulfide bridges (SS): SS1 = R261-R321, SS2 = R367-R425, SS3 = R321-R261, SS4 = R425-R367, SS5 = R23-R88, SS6 = R134-R194, SS7 = R92-R22, SS8 = R191-R136, SS9 = R23-R88, SS10 = R134-R194, SS11 = R92-R22, SS12 = R191-R136. We denoted this geometric vector by $\mathbf{Fab_1}$ (red), $\mathbf{Fab_2}$ (gray) and $\mathbf{F_c}$ (blue) for two Fabs and Fc, respectively. In addition, two geometric vectors were defined, $\mathbf{L_1}$ (olive) and $\mathbf{L_2}$ (green), to describe the linker connecting Ser239 of Fc to Pro208 of the two Fab heavy chains, and their length was measured between the C_α atoms of connected residues.

RESULTS AND DISCUSSIONS

IgG 3D structure assembly

The reconstruction protocol was performed on 10 child images from AFM topography, see **Figure 2**. Top ranked solutions are highlighted in green frames of **Figure 3**. The observed separations between the Fc and Fab1/Fab2 domains are in average $|\mathbf{L_1}| = 56.6 \pm 7.6$ Å, and $|\mathbf{L_2}| = 51.3 \pm 13.3$ Å, where the deviation term was expressed as the maximum deviation from the average. The angle between two linkers, $\angle\mathbf{L_1L_2}$, is 64.6 ± 43.6^o.

Because of the combinatorial explosion when assembling a three-domain structure, coarse combination runs were set initially with only 4000 docking orientations for each domain (a total of 4000^3 combinations). On each coarse run a different set of docking solutions were considered, filtered with a different translation and rotation interval, i.e. only the solutions with a larger translation or rotation change than the threshold are considered (**Figure 4**). Threshold used to filter out docking orientations are {(5 Å, 6º), (5 Å, 12º), (10 Å, 12º), (20 Å, 12º), (20 Å, 24º)}. To obtain the best ranked assembly from

all coarse combinations, a finer combination was performed by considering only nearby docking solutions of those selected in the coarse best-ranked assembly but within a translation and rotation window of 20 Å and 25 degrees (**Figure 4c**). The ranking of macromolecular assemblies is performed using a quality-of-fit parameter ($S_{EFactor}$) that determines the agreement between the experimental AFM topographic surface and the surface of the assembled antibody {Trinh, 2012 #3182}; thus, the ideal score for a reconstruction is 0.

To evaluate the convergence of the reconstruction protocol, RMSD were calculated within the reconstruction ranking (**Figure 5**). For both coarse and fine reconstructions (**Figure 5a** and **Figure 5b**, respectively), RMSDs of the ranking reconstructions converges uniformly. The observed convergence should not be confused with reconstruction accuracy but rather to illustrate whether top ranking solutions are unique. Each point on the fitting curves shown in **Figure 5a** and **Figure 5b** are obtained using an unconstrained logistic function: $y = A_2 + (A_1-A_2)/(1 + (x/x_0)^p)$ where A_1, A_2, x_0 and p are fitted parameters. An example of how a fit is obtained is indicated in **Figure 5c**. For the coarse and fine plot fits, quality of fit (R^2) was 0.027 and 0.021, whereas the asymptote was 60.8 ± 2.4 Å and 21.7 ± 9.0 Å, respectively. Although individual RMSD values vary significantly due to the dispersion of possible computed orientations (**Figure 5c**), the logistic regression fit clearly show identical trend for all experimental crop data. When the reference molecule was taken randomly (instead of the top 1) results were completely different and fitting lines were mostly linear (no convergence, data not shown). Thus, the computational reconstruction protocol produces assemblies in which geometry converges toward a small number of solutions that are ranked in the top 10.

IgG Dynamics

Flexibility of antibodies is well known to virologists who study binding of antibodies to large viruses. In case of symmetric antigens, such as viruses, antibodies usually have to behave like contortionists to be able to bind to antigens with both Fabs. Only by looking at individual molecules, it is possible to observe the wide range of structural variability

among antibodies as shown with cryo-EM {Sandin, 2004 #3448; Bongini, 2004 #3488}, AFM {Kienberger, 2004 #3404} or SAXS data {Lilyestrom, 2012 #3449}.

First, it is interesting to look at the variability among the reconstructions made for all the 10 child images. Using their Fc domain, each top reconstruction structure is superimposed onto that of the top 1 obtained from crop A (used as a reference only for this purpose) (**Figure 6**). The distribution of orientations observed for the Fab domains illustrate the flexibility of this molecule. This variation can be expressed by the RMSD as shown in **Figure 6b**. Both visually (**Figure 6a**) and numerically (**Figure 6b**) it shows that conformation of antibodies deposited on MPTES-grafted mica is highly variable.

Second, to test how variable are the top antibody reconstructions in the 10 child images, their conformation were compared with that of known complete X-ray structure of antibodies (**Figure 6c**). Results show very large RMSD values (> 60 Å) whatever is the reference X-ray structure indicating a clear difference between whole antibody X-ray structures and each reconstructed antibody where all the reconstructions resemble more to that of 1IGT than any other X-ray structures of antibodies.

Third, antibody domains flexibility can be described numerically by defining vectors along the axes through each domain along the length, and vectors between the linking points between the Fab and Fc domains (see **Figure 7**). The hinge/linker region was characterized by $|\mathbf{L_1}|$ and $|\mathbf{L_2}|$ (olive and green). Qualitatively, the superimposition of the antibody structures, using Fc as reference, illustrates once more that the Fabs and Fc can adopt a large variety of dispositions relative to each other. **Figure 8** presents a 3D vectorial representation of the top assemblies for the 10 child images. The flexibility in the linker/hinge region is clearly visible. In addition, flexibility in Fab domains allows them to move and rotate relative to each other and relative to the Fc. Quantitatively, vector lengths, vector-vector angles, as well as plane-plane orientations indicates that the geometry of top 10 reconstructed antibodies resembles at least one of the three X-ray crystal structures of whole antibody (**Table 1**). It can be noted that the range of Fab-Fc angles (28°-97°) obtained in our reconstruction is well included in the two limiting values (15°-127°) obtained by cryo-EM analysis {Bongini, 2004 #3488}. This result illustrates the plausibility of reconstructions by judging the quantitative variation in the geometry of reconstructed antibodies which adopt values that are commonly seen in known antibody

structures. It is important to recall that no other constraints than the AFM topography and a range of $|\mathbf{L_1}|$ and $|\mathbf{L_2}|$ were set in the reconstruction. We may conclude that the conformation of antibodies deposited on mica adopt a large variety of geometry that are compatible with known values observed from whole antibody structures. Although it is difficult to adequately compare projection results of EM with real-space 3D reconstruction of whole antibody molecules with our AFM-based protocol, the overall observed flexibility in our study is in good agreement with recently published data from EM {Bongini, 2004 #3488; Correia, 2013 #3336}.

CONCLUSIONS

AFM-Assembly is a complete computational protocol that uses the AFM imaging capabilities for the reconstruction of protein structures. AFM high signal/noise ratio provides imaging at the single molecule level capturing an ensemble of macromolecular conformations. Despite the lack of atomic-resolution in AFM imaging on isolated macromolecules, the AFM-Assembly process makes it possible to interpret the dynamics of flexible molecules such as antibodies.

Acknowledgements

We are grateful to Dr. L. Bellanger (DSV/iBEB/SBTN/LICB) for generous gift of monoclonal antibodies. This work was supported initially by French ANR [ANR-07-PCVI-0002-01] and partly from [ANR-P2N-2010-NANO-003] and the Commissariat à l'énergie atomique et aux énergies alternatives (CEA)

Additional information

Competing interests

The authors declare that no competing interests exist.

Author contributions

RCC and JLP, Conception and design, Analysis and interpretation of data, Drafting or revising the article; JMT, Acquisition of data; SWC Drafting or revising the article; MO and PP support on the acquisition of data.

REFERENCES

Asakawa H, Ikegami K, Setou M, Watanabe N, Tsukada M, Fukuma T. 2011. Submolecular-Scale Imaging of alpha-Helices and C-Terminal Domains of Tubulins by Frequency Modulation Atomic Force Microscopy in Liquid. *Biophys. J.* **101**: 1270-1276. Doi:10.1016/j.bpj.2011.07.020

Bennett A, Liu J, Van Ryk D, Bliss D, Arthos J, Henderson RM, Subramaniam S. 2007. Cryoelectron tomographic analysis of an HIV-neutralizing protein and its complex with native viral gp120. *J. Biol. Chem.* **282**: 27754-27759. Doi:10.1074/jbc.M702025200

Binnig G, Quate CF, Gerber C. 1986. Atomic force microscope. *Phys. Rev. Lett.* **56**: 930-933. Doi:10.1103/PhysRevLett.56.930

Burton DR. 1990. Antibody: the flexible adaptor molecule. *Trends Biochem. Sci.* **15**: 64-69.

Buzhynskyy N, Golczak M, Lai-Kee-Him J, Lambert O, Tessier B, Gounou C, Berat R, Simon A, Granier T, Chevalier JM and others. 2009. Annexin-A6 presents two modes of association with phospholipid membranes. A combined QCM-D, AFM and cryo-TEM study. *J. Struct. Biol.* **168**: 107-116. Doi:10.1016/j.jsb.2009.03.007

Capon DJ, Kaneko N, Yoshimori T, Shimada T, Wurm FM, Hwang PK, Tong X, Adams SA, Simmons G, Sato TA and others. 2011. Flexible antibodies with nonprotein hinges. *Proc. Jpn Acad. Ser. B* **87**: 603-616. Doi:JST.JSTAGE/pjab/87.603

Chen S-wW, Pellequer JL. 2004. Identification of functionally important residues in proteins using comparative models. *Curr. Med. Chem.* **11**: 595-605. Doi:10.2174/0929867043455891

Chen S-wW, Pellequer JL. 2013. Adepth: new representation and its implications for atomic depths of macromolecules. *Nucl. Acids Res.* **41**: W412-W416. Doi:10.1093/nar/gkt299

Chen S-wW, Pellequer JL. 2011. DeStripe: frequency-based algorithm for removing stripe noises from AFM images. *BMC Struct. Biol.* **11**: 7. Doi:10.1186/1472-6807-11-7

Correia I, Sung J, Burton R, Jakob CG, Carragher B, Ghayur T, Radziejewski C. 2013. The structure of dual-variable-domain immunoglobulin molecules alone and bound to antigen. *mAbs* **5**: 364-372. Doi:10.4161/mabs.24258

Czajkowsky DM, Shao Z. 2009. The human IgM pentamer is a mushroom-shaped molecule with a flexural bias. *Proc. Natl. Acad. Sci. USA* **106**: 14960-14965. Doi:10.1073/pnas.0903805106

Dangl JL, Wensel TG, Morrison SL, Stryer L, Herzenberg LA, Oi VT. 1988. Segmental flexibility and complement fixation of genetically engineered chimeric human, rabbit and mouse antibodies. *EMBO J.* **7**: 1989-1994.

Davies E, Teng KS, Conlan RS, Wilks SP. 2005. Ultra-high resolution imaging of DNA and nucleosomes using non-contact atomic force microscopy. *FEBS Lett.* **579**: 1702-1706. Doi:10.1016/j.febslet.2005.02.028

Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, Oldfield CJ, Campen AM, Ratliff CM, Hipps KW and others. 2001. Intrinsically disordered protein. *J. Mol. Graph. Model.* **19**: 26-59. Doi:10.1016/S1093-3263(00)00138-8

Gale AJ, Yegneswaran S, Xu X, Pellequer J-L, Griffin JH. 2007. Characterization of a factor Xa binding site on factor Va near the Arg506 APC cleavage site. *J. Biol. Chem.* **282**: 21848–21855. Doi:10.1074/jbc.M702192200

Ghemtio L, Muzet N. 2013? Retrospective molecular docking study of WY-25105 ligand to beta-secretase and bias of the three-dimensional structure flexibility. *J. Mol. Model.* **in press**. Doi:10.1007/s00894-013-1821-0

Gregory L, Davis KG, Sheth B, Boyd J, Jefferis R, Nave C, Burton DR. 1987. The solution conformations of the subclasses of human IgG deduced from sedimentation and small angle X-ray scattering studies. *Mol. Immunol.* **24**: 821-829.

Gwyddion. 2013. http://gwyddion.net/.

Harris LJ, Skaletsky E, McPherson A. 1995. Crystallization of intact monoclonal antibodies. *Proteins* **23**: 285-289. Doi:10.1002/prot.340230218

Harris LJ, Skaletsky E, McPherson A. 1998. Crystallographic structure of an intact IgG1 monoclonal antibody. *J. Mol. Biol.* **275**: 861-872. Doi:10.1006/jmbi.1997.1508

Humphrey W, Dalke A, Schulten K. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**: 33-38. Doi:0263785596000185 [pii]

Ido S, Kimura K, Oyabu N, Kobayashi K, Tsukada M, Matsushige K, Yamada H. 2013. Beyond the helix pitch: direct visualization of native DNA in aqueous solution. *ACS Nano* **7**: 1817-1822. Doi:10.1021/nn400071n

James LC, Roversi P, Tawfik DS. 2003. Antibody multispecificity mediated by conformational diversity. *Science* **299**: 1362-1367. Doi:10.1126/science.1079731

James LC, Tawfik DS. 2003. Conformational diversity and protein evolution--a 60-year-old hypothesis revisited. *Trends Biochem. Sci.* **28**: 361-368. Doi:10.1016/S0968-0004(03)00135-X

Kienberger F, Mueller H, Pastushenko V, Hinterdorfer P. 2004. Following single antibody binding to purple membranes in real time. *EMBO Rep.* **5**: 579-583. Doi:10.1038/sj.embor.7400149

Kim H, Matsunaga C, Yoshino A, Kato K, Arata Y. 1994. Dynamical structure of the hinge region of immunoglobulin G as studied by 13C nuclear magnetic resonance spectroscopy. *J. Mol. Biol.* **236**: 300-309. Doi:10.1006/jmbi.1994.1136

Kuznetsov YG, Day J, Newman R, McPherson A. 2000. Chimeric human-simian anti-CD4 antibodies form crystalline high symmetry particles. *J. Struct. Biol.* **131**: 108-115. Doi:10.1006/jsbi.2000.4282

Larson S, Day J, Greenwood A, Skaletsky E, McPherson A. 1991. Characterization of crystals of an intact monoclonal antibody for canine lymphoma. *J. Mol. Biol.* **222**: 17-19. Doi:10.1016/0022-2836(91)90731-K

Lilyestrom WG, Shire SJ, Scherer TM. 2012. Influence of the cosolute environment on IgG solution structure analyzed by small-angle X-ray scattering. *J. Phys. Chem. B* **116**: 9611-9618. Doi:10.1021/jp303839t

Mandell JG, Roberts VA, Pique ME, Kotlovyi V, Mitchell JC, Nelson E, Tsigelny I, Ten Eyck LF. 2001. Protein docking using continuum electrostatics and geometric fit. *Protein Eng.* **14**: 105-113. Doi:10.1093/protein/14.2.105

Martinez NF, Lozano JR, Herruzo ET, Garcia F, Richter C, Sulzbach T, Garcia R. 2008. Bimodal atomic force microscopy imaging of isolated antibodies in air and liquids. *Nanotechnology* **19**: 384011. Doi:10.1088/0957-4484/19/38/384011

Pellequer J-L, Van Regenmortel MHV. 1993. Affinity of monoclonal antibodies to large multivalent antigens: influence of steric hindrance on antibody affinity constants calculated from Scatchard plots. *Mol. Immunol.* **30**: 955-958. Doi:10.1016/0161-5890(93)90022-4

Pellequer J-L, Chen S-wW. 2006. Multi-template approach to modeling engineered disulfide bonds. *Proteins* **65**: 192-202. Doi:10.1002/prot.21059

Roux KH, Strelets L, Michaelsen TE. 1997. Flexibility of human IgG subclasses. *J. Immunol.* **159**: 3372-3382.

Roux KH. 1999. Immunoglobulin structure and function as revealed by electron microscopy. *Int. Arch. Allergy Immunol.* **120**: 85-99. Doi:10.1159/000024226

Roux KH, Strelets L, Brekke OH, Sandlie I, Michaelsen TE. 1998. Comparisons of the ability of human IgG3 hinge mutants, IgM, IgE, and IgA2, to form small immune complexes: a role for flexibility and geometry. *J. Immunol.* **161**: 4083-4090.

Roux KH, Greenspan NS. 1994. Monitoring the formation of soluble immune complexes composed of idiotype and anti-idiotype antibodies by electron microscopy. *Mol. Immunol.* **31**: 599-606. Doi:10.1016/0161-5890(94)90167-8

Roux KH. 1989. Immunoelectron microscopy of idiotype-anti-idiotype complexes. *Methods Enzymol.* **178**: 130-144. Doi:10.1016/0076-6879(89)78010-1

San Paulo A, Garcia R. 2000. High-resolution imaging of antibodies by tapping-mode atomic force microscopy: attractive and repulsive tip-sample interaction regimes. *Biophys. J.* **78**: 1599-1605. Doi:10.1016/S0006-3495(00)76712-9

Sandin S, Ofverstedt LG, Wikstrom AC, Wrange O, Skoglund U. 2004. Structure and flexibility of individual immunoglobulin G molecules in solution. *Structure* **12**: 409-415. Doi:10.1016/j.str.2004.02.011

Saphire EO, Stanfield RL, Crispin MD, Morris G, Zwick MB, Pantophlet RA, Parren PW, Rudd PM, Dwek RA, Burton DR and others. 2003. Crystal structure of an intact human IgG: antibody asymmetry, flexibility, and a guide for HIV-1 vaccine design. *Adv. Exp. Med. Biol.* **535**: 55-66.

Saphire EO, Parren PW, Pantophlet R, Zwick MB, Morris GM, Rudd PM, Dwek RA, Stanfield RL, Burton DR, Wilson IA. 2001. Crystal structure of a neutralizing human IGG against HIV-1: a template for vaccine design. *Science* **293**: 1155-1159. Doi:10.1126/science.1061692

Saphire EO, Stanfield RL, Crispin MD, Parren PW, Rudd PM, Dwek RA, Burton DR, Wilson IA. 2002. Contrasting IgG structures reveal extreme asymmetry and flexibility. *J. Mol. Biol.* **319**: 9-18. Doi:10.1016/S0022-2836(02)00244-9

Schabert FA, Engel A. 1994. Reproducible acquisition of Escherichia coli porin surface topographs by atomic force microscopy. *Biophys. J.* **67**: 2394-2403. Doi:10.1016/S0006-3495(94)80726-X

Scheuring S, Busselez J, Levy D. 2005. Structure of the dimeric PufX-containing core complex of Rhodobacter blasticus by in situ atomic force microscopy. *J. Biol. Chem.* **280**: 1426-1431. Doi:10.1074/jbc.M411334200

Scheuring S, Boudier T, Sturgis JN. 2007. From high-resolution AFM topographs to atomic models of supramolecular assemblies. *J. Struct. Biol.* **159**: 268-276. Doi:10.1016/j.jsb.2007.01.021

Stura EA, Satterthwait AC, Calvo JC, Stefanko RS, Langeveld JP, Kaslow DC. 1994. Crystallization of an intact monoclonal antibody (4B7) against Plasmodium

falciparum malaria with peptides from the Pfs25 protein antigen. *Acta Crystallogr. D* **50**: 556-562. Doi:10.1107/S0907444994001782

Ten Eyck LF, Mandell J, Roberts VA, Pique ME. http://www.supercomp.org/sc95/proceedings/636_LTEN/SC95.HTM. In: Hayes A, Simmons M, editors; 1995; San Diego. IEEE Computer Society Press, Los Alamitos, CA.

Thomson NH. 2005. The substructure of immunoglobulin G resolved to 25 kDa using amplitude modulation AFM in air. *Ultramicroscopy* **105**: 103-110. Doi:10.1016/j.ultramic.2005.06.024

Thouvenin E, Laurent S, Madelaine M-F, Rasschaert D, Vautherot J-F, Hewat EA. 1997. Bivalent binding of a neutralising antibody to a calicivirus involves the torsional flexibility of the antibody hinge. *J. Mol. Biol.* **270**: 238-246. Doi:10.1006/jmbi.1997.1095

Tong H, Zhang L, Kaspar A, Rames MJ, Huang L, Woodnutt G, Ren G. 2013. Peptide-conjugation induced conformational changes in human IgG1 observed by optimized negative-staining and individual-particle electron tomography. *Sci. Rep.* **3**: 1089. Doi:10.1038/srep01089

Trinh M-H, Odorico M, Pique ME, Teulon J-M, Roberts VA, Ten Eyck LF, Getzoff ED, Parot P, Chen S-wW, Pellequer J-L. 2012. Computational reconstruction of multidomain proteins using atomic force microscopy data. *Structure* **20**: 113-120. Doi:10.1016/j.str.2011.10.023

Trinh M-H, Odorico M, Bellanger L, Jacquemond M, Parot P, Pellequer J-L. 2011. Tobacco mosaic virus as an AFM tip calibrator. *J. Mol. Recognit.* **24**: 503-510. Doi:10.1002/jmr.1119

Villarrubia JS. 1997. Algorithms for scanned probe microscope image simulation, surface reconstruction, and tip estimation. *J. Res. Natl. Inst. Stand. Technol.* **102**: 425-454.

Yang J, Mou J, Yuan JY, Shao Z. 1996. The effect of deformation on the lateral resolution of atomic force microscopy. *J. Microsc.* **182**: 106-113. Doi:10.1046/j.1365-2818.1996.140422.x

Zhang L, Song J, Newhouse Y, Zhang S, Weisgraber KH, Ren G. 2010. An optimized negative-staining protocol of electron microscopy for apoE4 POPC lipoprotein. *J. Lipid Res.* **51**: 1228-1236. Doi:10.1194/jlr.D002493

Zhang L, Ren G. 2012. IPET and FETR: experimental approach for studying molecular structure dynamics by cryo-electron tomography of a single-molecule structure. *PLoS One* **7**: e30249. Doi:10.1371/journal.pone.0030249

Zhang L, Song J, Cavigiolio G, Ishida BY, Zhang S, Kane JP, Weisgraber KH, Oda MN, Rye KA, Pownall HJ and others. 2011. Morphology and structure of lipoproteins revealed by an optimized negative-staining protocol of electron microscopy. *J. Lipid Res.* **52**: 175-184. Doi:10.1194/jlr.D010959

Table 1 - Dynamics of the relative angles and separation of the three domains in each individual antibody.

	Linker					*Domains*					
	separation (Å)		axis (°)			axis (°)			planes (°)		
Obj	|L1|	|L2|	L1^L2	L1^Fc	L2^Fc	Fab1^Fab2	Fab1^Fc	Fab2^Fc	Fab1^Fab2	Fab1^Fc	Fab2^Fc
A	58.5	49.1	63.1	118.7	144.1	23.8	29.5	50.7	31.5	55.1	46.9
B	60.5	51.4	87.3	104.7	144.5	31.6	28.1	37.5	43.3	83.3	41.7
C	59.7	51.0	74.8	121.1	140.1	22.9	40.9	31.0	58.0	83.0	28.5
D	56.4	50.9	103.2	101.1	138.8	49.8	34.2	32.9	32.1	68.1	37.1
E	54.8	58.6	68.8	145.9	130.4	25.7	53.7	44.4	32.8	46.5	13.8
F	59.4	59.3	20.7	145.4	150.8	97.2	45.2	90.0	60.1	40.3	85.2
G	58.4	47.4	44.1	158.9	157.0	11.5	43.7	53.3	54.6	47.8	86.6
H	56.6	61.6	77.5	97.8	128.2	31.1	70.0	57.3	64.0	45.6	24.9
I	49.3	45.9	71.5	113.9	154.1	90.2	37.8	97.3	67.9	35.0	61.9
J	54.5	38.4	35.7	117.4	121.5	40.6	51.7	48.3	43.3	48.5	54.3
mean	**57**	**51**	**65**	**122**	**141**	**42**	**43**	**54**	**49**	**55**	**48**
range	49:60	38:62	21:103	98:159	121:157	12:97	28:70	31:97	31:68	35:83	14:87
1HZH[a,b]	34.8	45.7	41.8	135.4	107.0	143.4	70.2	73.3	57.7	79.6	89.7
1IGT[a,b]	51.6	35.5	50.6	113.4	146.0	169.5	112.6	64.9	35.0	73.1	88.1
1IGY[a,b]	36.6	34.0	74.6	72.5	116.7	112.8	99.7	56.6	39.2	56.4	85.5

[a]Superimpositions of Fab and Fc on the X-ray structures (see methods).

[b]Highlighted boxes indicate values from X-ray structures that are closest to mean values from reconstructed antibodies.

Figure Legends

Figure 1 – *AFM-Assembly* concept: the 3D structure of a large protein can be reconstructed (4) by combining computed orientations obtained from the docking stage (3) with the constituent units of the target protein (1) under experimental constraints that is high-resolution AFM topograph (2).

Figure 2 – AFM topographic image of multiple isolated antibodies deposited on chemically-modified mica substrate. Antibodies are chemically grafted on the substrate using a short linker. AFM imaging was performed in air with SNL tips using the PeakForce tapping mode of a MultiMode 8 AFM. The 10 selected isolated antibodies are indicated by yellow boxes (child images A-J).

Figure 3 – Antibody structure reconstruction obtained by AFM-Assembly for 10 child images (A-J) highlighted in Figure 2. Interpolated crops of these objects are shown in yellow frames. Height profiles are given according to xy (black) and yx (red) directions; location of profile on a crop image is shown for crop A (xy profile in shown in white). 3D antibody reconstruction is shown in green frames. On the top of each frame a backside view of the reconstructed antibody in the AFM topographic surface is shown. On the bottom is shown a cartoon representation of each reconstructed antibody where Fab1 is colored in orange and gray while Fab2 is colored in yellow and olive; Fc fragments are colored in red and blue. The Fab1-Fab2 angle values as well as the linker lengths are indicated. Images are drawn using VMD {Humphrey, 1996 #2696}.

Figure 4 – Multiple reconstruction of antibody using the child image A. a) Effect of thresholds used for selecting docking solutions: the top row uses a linker distance ranging from 35 to 52Å whereas the range in the bottom row is from 28 to 62.4Å. Each column represents a different set of translation (in Å) and rotation (in °) used to select docking solutions. Below each 3D representation of the reconstruction is shown the docking rank obtained from the docking with DOT. b) Top five antibody reconstructions from coarse combination using parameters found in the bottom right frames of a). In this

reconstruction only the top 1, i.e. the lowest score, is in visual agreement with an antibody shape. c) Refinement of the reconstruction using a finer set of docking solutions that are selected around those of top score reconstruction of b).

Figure 5 – Representation of the reconstruction protocol convergence in the 10 different child images (A-J) using the $S_{Efactor}$. Logistic regression fit between RMSD and ranking solutions obtained from a) coarse and b) fine combination. RMSD is computed between each reconstruction after superimposition onto the top 1 combination (step size = 100). Typical raw data in shown in c) where the logistic regression fit is displayed on red. No constraints have been used in the fit. These results indicate that the top 50 reconstructions are most similar with each other than those compared with low ranking solutions. It also indicates that low ranking solutions are always dissimilar from the top 1.

Figure 6 – Structural variability in antibody reconstruction. a) Superimposition of the top 1 antibody reconstruction for each of the 10 child images using the Fc fragment as reference; the color code is identical to that of fig. 3. b) Representation of the RMSD computed between all atoms of each reconstruction in a) taking the object A as an arbitrarily reference. c) RMSD computed between each top 1 reconstruction for the 10 child images and the three known complete X-ray structures of antibodies available in the PDB. [a]Superimpositions of Fab and Fc on the X-ray structures (see methods).

Figure 7 – Vectorial representation of the subunits in antibody structures. L1 and L2 are the linker vectors between Fc and Fabs subunits. Orientation vectors, thick arrows, in Fab and Fc subunits are defined using conserved disulfide bridges, labeled SS**X**. Fab and Fc structures are colored as indicated in Figure 3.

Figure 8 – 3D vectorial representation of the linking vectors (L1 and L2) and the domains vectors (Fc, Fab1 and Fab2) for the 10 child images (A-J).

Fig. 1

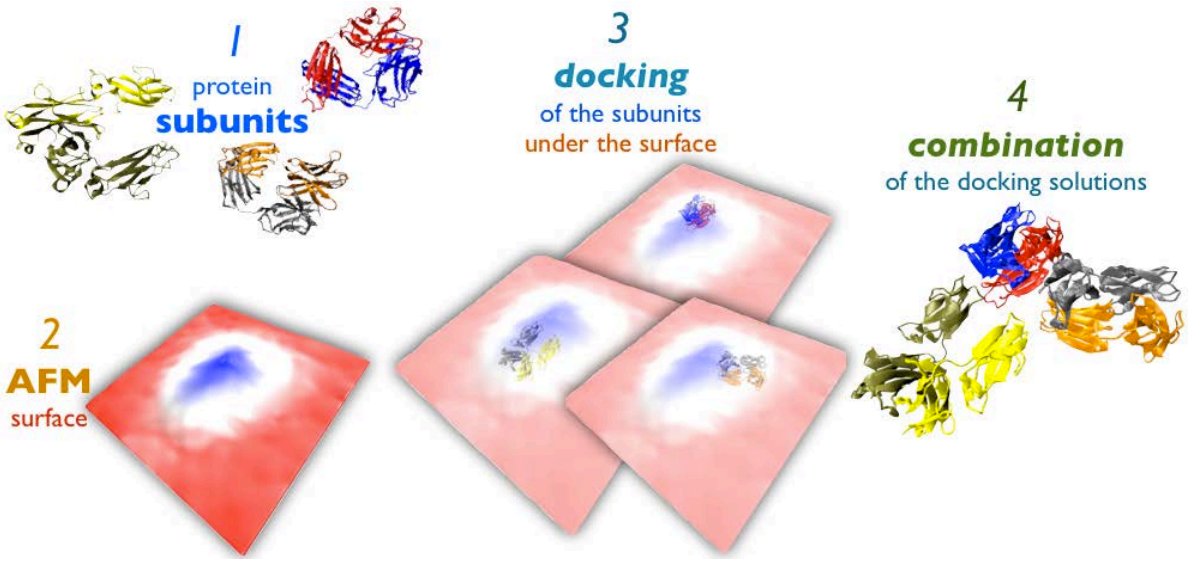
1
protein
subunits
2
AFM
surface
3
docking
of the subunits
under the surface
4
combination
of the docking solutions

Fig. 2

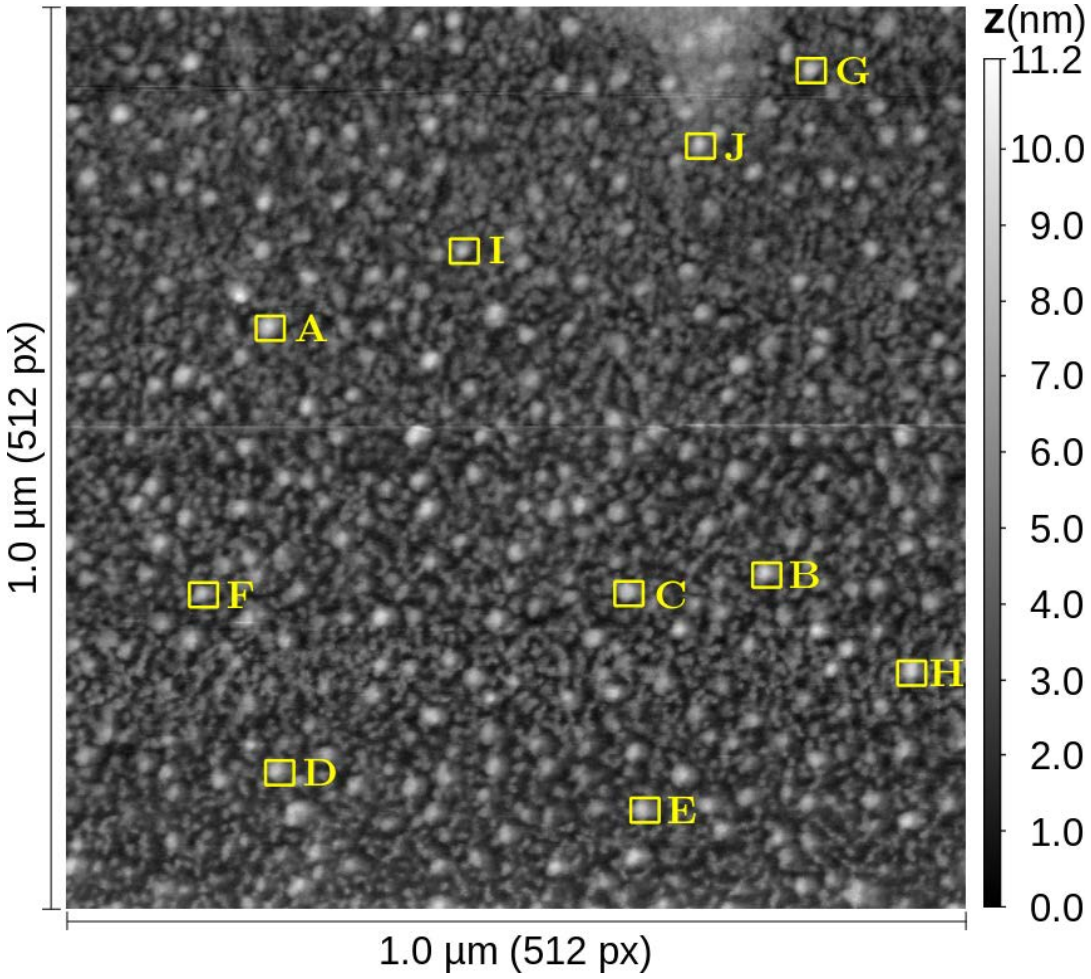
z(nm)
11.2
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
G
J
I
A
F
C
B
H
D
E
1.0 µm (512 px)
1.0 µm (512 px)

Fig. 3

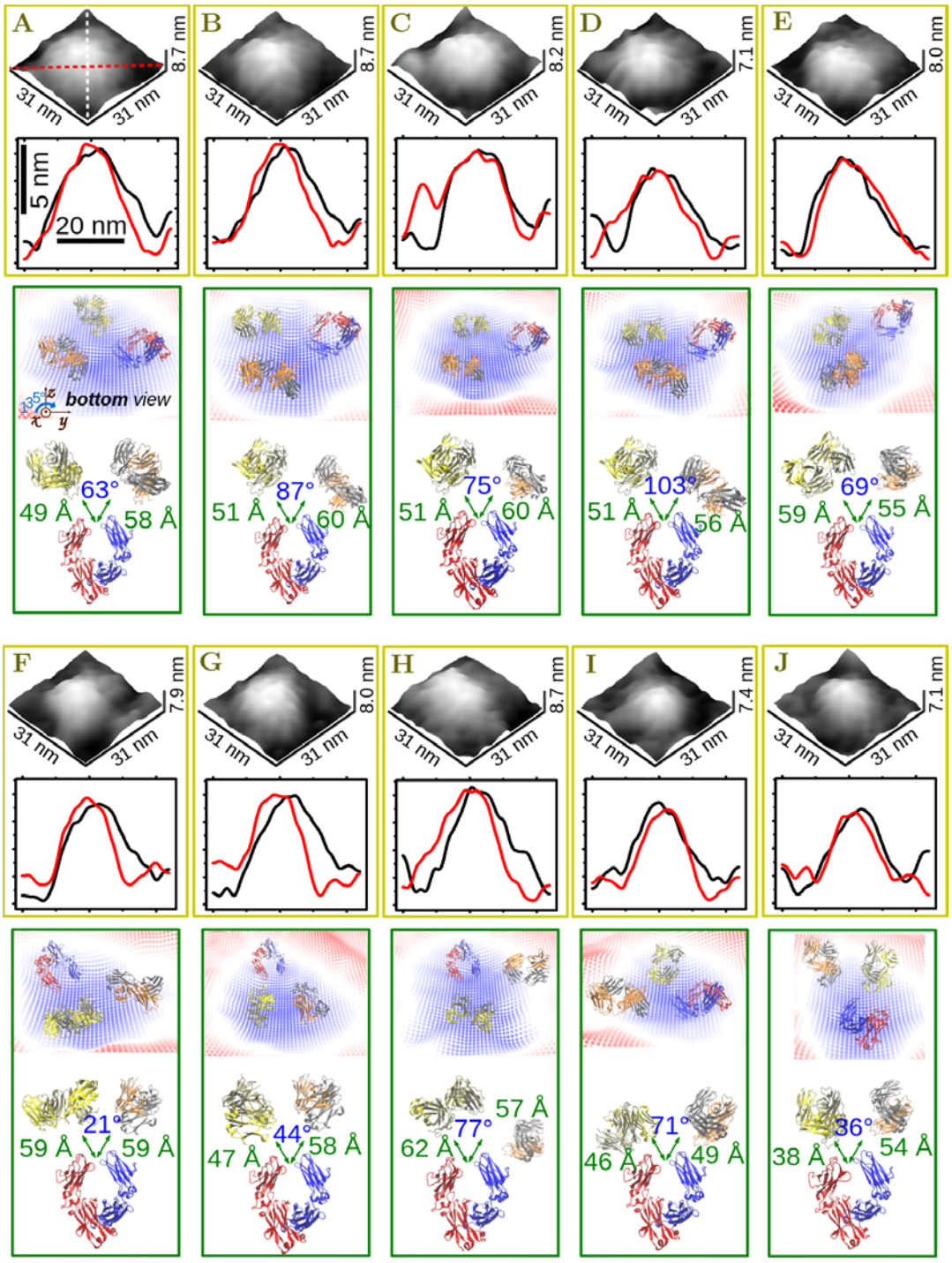
A
B
C
D
E
8.7 nm
8.7 nm
8.2 nm
7.1 nm
8.0 nm
31 nm
31 nm
5 nm
20 nm
bottom view
63°
49 Å
58 Å
87°
51 Å
60 Å
75°
51 Å
60 Å
103°
51 Å
56 Å
69°
59 Å
55 Å
F
G
H
I
J
7.9 nm
8.0 nm
8.7 nm
7.4 nm
7.1 nm
21°
59 Å
59 Å
44°
47 Å
58 Å
77°
62 Å
57 Å
71°
46 Å
49 Å
36°
38 Å
54 Å

Fig. 4

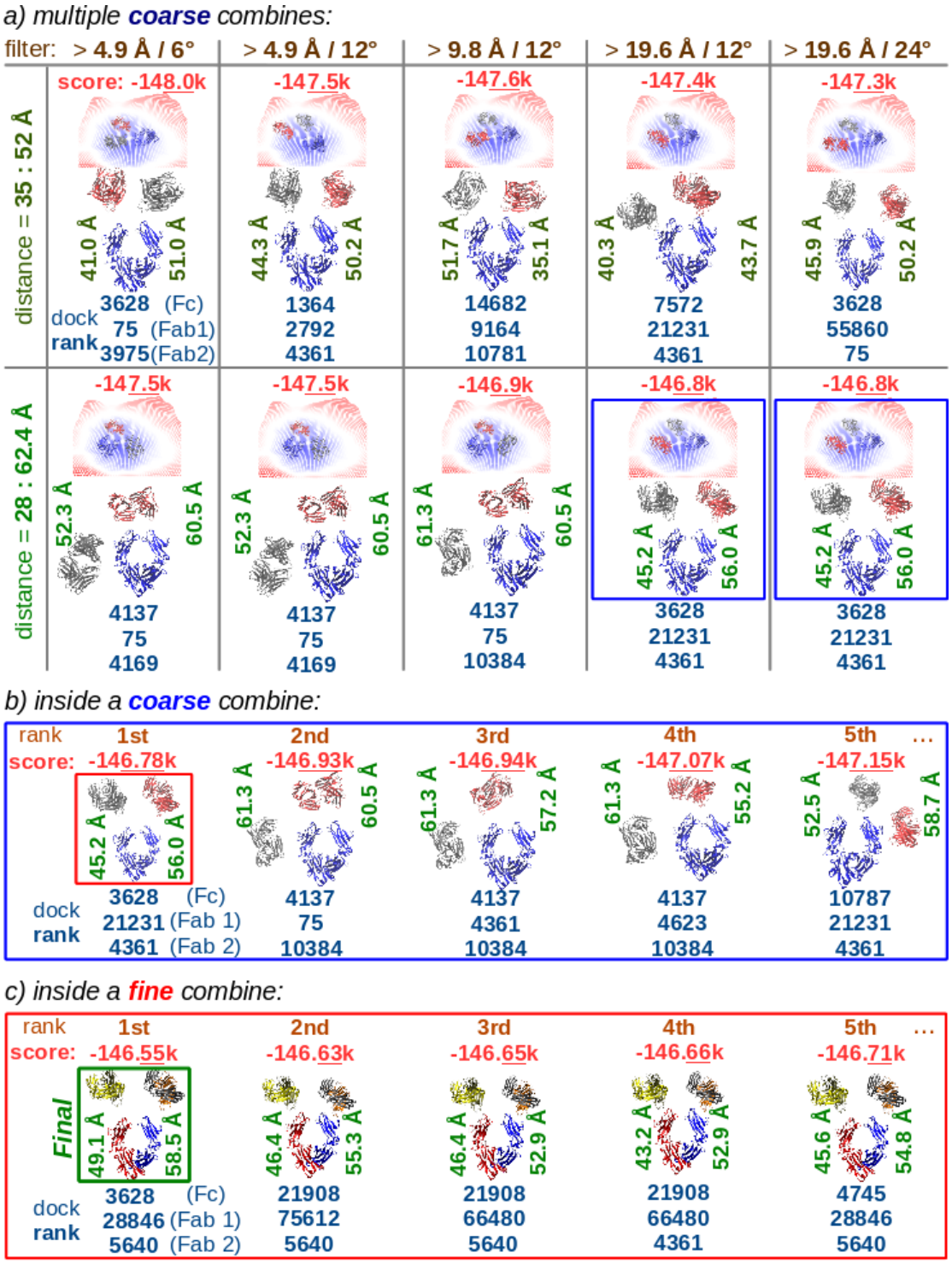
a) multiple coarse combines:
filter: > 4.9 Å / 6°
> 4.9 Å / 12°
> 9.8 Å / 12°
> 19.6 Å / 12°
> 19.6 Å / 24°
distance = 35 : 52 Å
distance = 28 : 62.4 Å
score: -148.0k
-147.5k
-147.6k
-147.4k
-147.3k
-147.5k
-147.5k
-146.9k
-146.8k
-146.8k
41.0 Å
51.0 Å
44.3 Å
50.2 Å
51.7 Å
35.1 Å
40.3 Å
43.7 Å
45.9 Å
50.2 Å
52.3 Å
60.5 Å
52.3 Å
60.5 Å
61.3 Å
60.5 Å
45.2 Å
56.0 Å
45.2 Å
56.0 Å
dock rank
3628 (Fc)
75 (Fab1)
3975 (Fab2)
1364
2792
4361
14682
9164
10781
7572
21231
4361
3628
55860
75
4137
75
4169
4137
75
4169
4137
75
10384
3628
21231
4361
3628
21231
4361
b) inside a coarse combine:
rank
1st
2nd
3rd
4th
5th
...
score:
-146.78k
-146.93k
-146.94k
-147.07k
-147.15k
61.3 Å
60.5 Å
61.3 Å
57.2 Å
61.3 Å
55.2 Å
52.5 Å
58.7 Å
45.2 Å
56.0 Å
dock rank
3628 (Fc)
21231 (Fab 1)
4361 (Fab 2)
4137
75
10384
4137
4361
10384
4137
4623
10384
10787
21231
4361
c) inside a fine combine:
rank
1st
2nd
3rd
4th
5th
...
score:
-146.55k
-146.63k
-146.65k
-146.66k
-146.71k
Final
49.1 Å
58.5 Å
46.4 Å
55.3 Å
46.4 Å
52.9 Å
43.2 Å
52.9 Å
45.6 Å
54.8 Å
dock rank
3628 (Fc)
28846 (Fab 1)
5640 (Fab 2)
21908
75612
5640
21908
66480
5640
21908
66480
4361
4745
28846
5640

Fig. 5

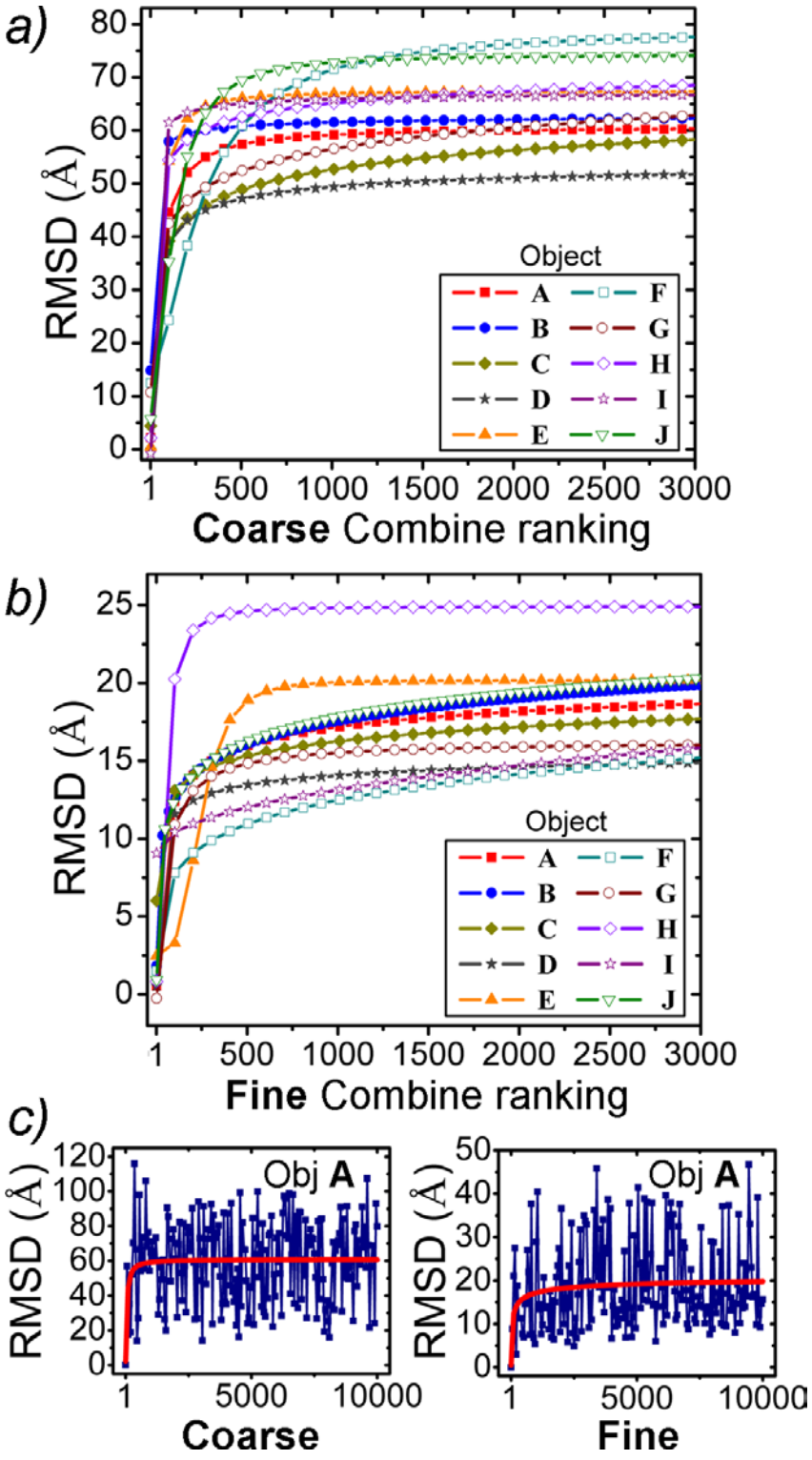
a)
RMSD (Å)
Coarse Combine ranking
Object
A
B
C
D
E
F
G
H
I
J
b)
RMSD (Å)
Fine Combine ranking
c)
Obj A
RMSD (Å)
Coarse
Fine

Fig. 6

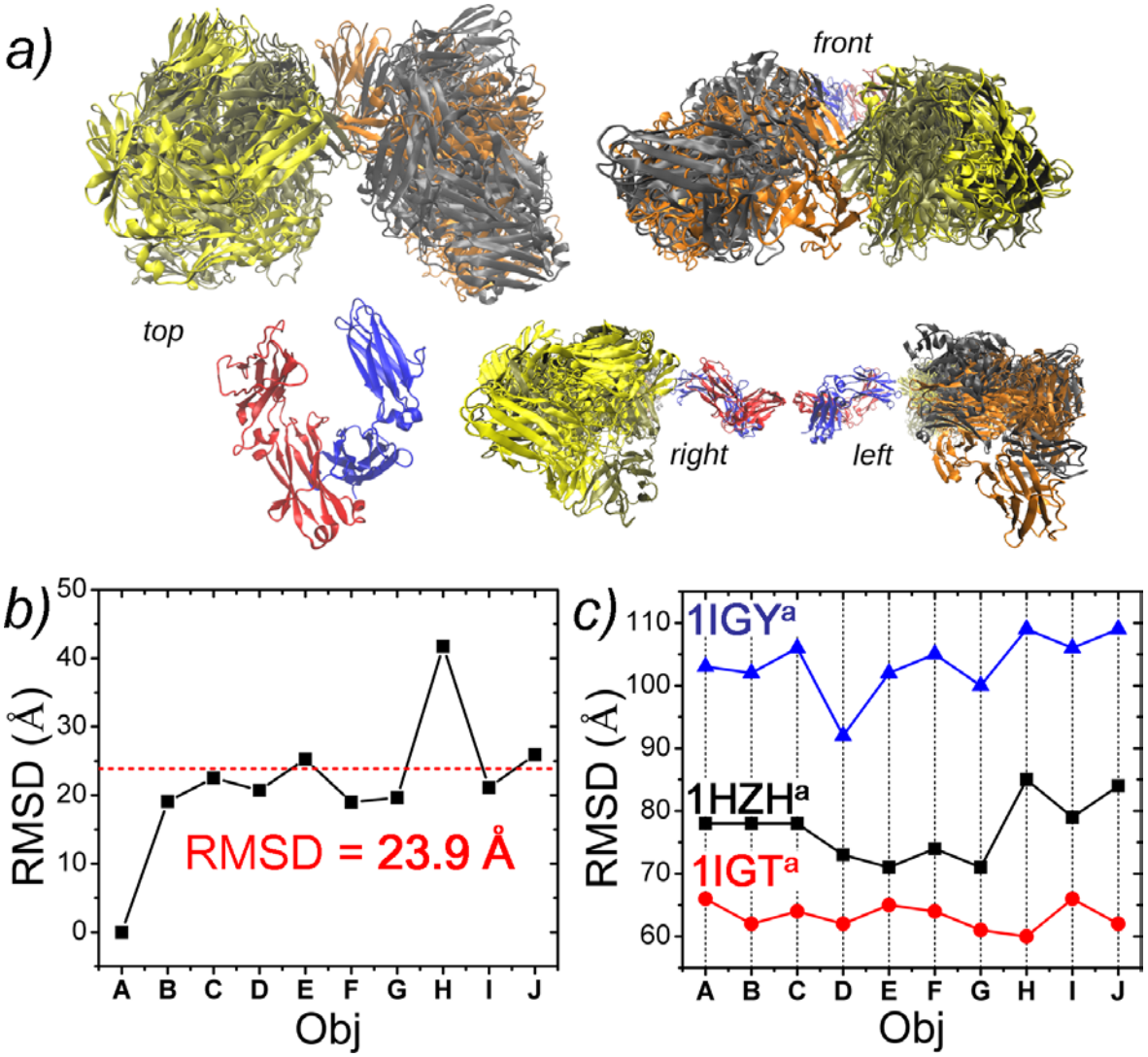
a)
front
top
right
left
b)
RMSD (Å)
RMSD = 23.9 Å
Obj
A B C D E F G H I J
c)
1IGY^a
1HZH^a
1IGT^a
RMSD (Å)
Obj

Fig. 7

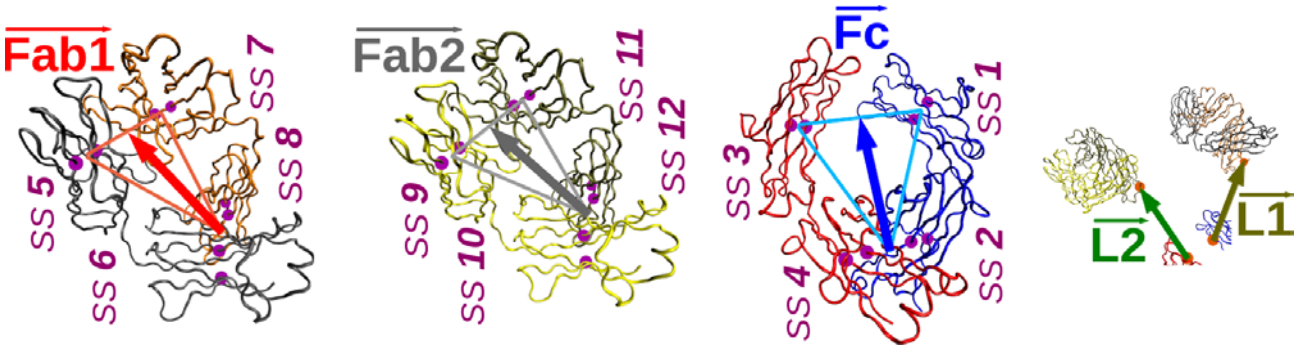
Fab1
SS 5
SS 6
SS 7
SS 8
Fab2
SS 9
SS 10
SS 11
SS 12
Fc
SS 3
SS 4
SS 1
SS 2
L2
L1

Fig. 8