Local resolution estimates of cryoEM reconstructions

JL Vilas[1,5], JB Heymann[2,5], HD Tagare[1,5], E Ramirez-Aportela[3,5], JM Carazo[3,5] and COS Sorzano[3,4,5]

The field of cryoEM has quickly advanced in last years with the new biochemical, technological, methodological and computational developments. It has allowed significant progresses in Structural Biology, typically reaching quasi-atomic resolutions in the reconstructed maps. However, this rapid advance has also generated new questions relevant to resolution estimates. The global resolution metrics and their criteria have been deeply discussed in the last decade, but despite that, it remains as an important issue in the field. Recently, the introduction of local resolution measurements has changed how cryoEM reconstructions are interpreted, providing information about the existence of heterogeneity, flexibility, and angular assignment errors, and using it as a tool to aid in modeling. In this review we revisit the concept of local resolution and the different algorithms in the current state of the art. However, the concept of local resolution is not uniquely defined, and each implementation measures different features. This may lead to inappropriate interpretation of local resolution maps. Hence, a set of good practices is provided in this review to avoid misleading and over-interpretation of the reconstructions.

Addresses
[1] Department of Biomedical Engineering, Yale University, New Haven, United States

[2] Laboratory of Structural Biology Research, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD, United States

[3] Biocomputing Unit, Centro Nacional de Biotecnologia (CNB-CSIC), Campus Universidad Autonoma, 28049 Cantoblanco, Madrid, Spain

[4] Univ. San Pablo—CEU, Campus Urb. Monteprincipe, 28668 Boadilla del Monte, Madrid, Spain

Corresponding author: Sorzano, COS (coss@cnb.csic.es)
[5] All authors have equally contributed to this article.

Structural Biology has as main objective the elucidation of the three-dimensional structure of biological macromolecules to predict and understand their action mechanisms. One of the important techniques is cryo-electron microscopy (cryoEM), which makes use of a transmission electron microscope to acquire images of a specimen under study. Initially, cryoEM produced low resolution structures that revealed only the overall shapes of the complexes, and so the term *blobology* was coined. However, in the last decade rapid advances allowed the determination of highly detailed structures (even atomic structures), bringing about the so-called resolution revolution [1]. Nevertheless, the field always had the need to quantify the quality of the reconstructed maps in terms of the spatial reliability of the measured features, that is, resolution. Over time, many metrics have been proposed for the estimation of the map quality, the main ones differential phase residual [2], spectral signal to Noise ratio [3] or the Fourier shell correlation (FSC) [4], the latter being the current standard in cryoEM (for an in-depth analysis of their theoretical properties, see [5•]). Many factors such as structural variation and preferred orientation result in varying local details in reconstructions. The question is, therefore, to properly represent these variations and how to interpret them. Thus, local resolution offers a spatially limited quality metric to capture structural variability, particularly important in building models into the maps. The first method developed to assess local resolution, named *blocres*, uses a local form of the FSC between two half maps [6]. A second algorithm, *ResMap*, followed a year later and has become popular because it calculates local resolution of a single reconstruction, as well as between two half maps [7]. Since then, new approaches include *MonoRes*, which uses a monogenic signal for accessing the local properties of the map [8], and *DeepRes* based on deep learning from filtered atomic models at different frequencies [9•]. Local resolution has also been applied in electron tomography by adapting the *MonoRes* algorithm, [10•]. Recently, assessment of deficiencies in angular assignment by directional resolution estimates [11•,12•] was extended to local-directional resolution anisotropy (*Monodir*, [13••]). An important consequence of this latter work is the introduction into the field of the concept that resolution is simultaneously local and directional. The number of different ways in which local resolution can be estimated provides a rich basis for understanding structural variability within cryoEM reconstructions.

This review addresses the latest advances on local resolution estimation, as well as the good practices about how it should be performed and interpreted. Despite the extensive use of local resolution analyses in the last

few years, there still remain questions of proper application and avoiding over-interpretation. We present the main local resolution programs available and comment on their strengths and weaknesses, as well as best practices for each.

Current local resolution methods and its limitations

Currently, there exist several local resolution algorithms that we summarize here. We illustrate the use of the different methods with an example. We also provide tables of the strengths and weakness of each method in the supplementary material.

blocres

The Bsoft package [14] offers the program blocres to produce a local resolution map that can be used to filter a map with the program blocfilt to the appropriate local detail [6]. The algorithm in blocres extracts smaller sub-maps ('boxes') from halfmaps and calculates the FSC between them. The edges of each box can be softened with a windowing function (the Hanning window is the default) to suppress high frequency correlations due to the edges. The box size is important to obtain sufficient counting statistics and avoid over-estimating the local resolution. An empirical rule was developed that sets a good box size from 5 to 7 times the global resolution.

The program calculates a resolution estimate at every voxel. This is an expensive operation that is accelerated by parallel processing, calculating on a coarser grid and interpolating, and excluding unnecessary calculations in the background and symmetry-related parts using a mask.

The output is a resolution map with the estimates in angstrom at the FSC curve cutoff specified. The voxels that are not computed or interpolated are set to a background value that the user specifies, usually below the lowest resolution expected (such as 20 or 50 Å). The map can be used in UCSF Chimera [15] to color the corresponding reconstruction.

The program blocres can also estimate the resolution in radial shells, useful when examining reconstructions of icosahedral viral capsids or similar structures. The algorithm extracts shells with the user-specified width from each halfmap and compare them by FSC. The shell edges should be smoothed to suppress high frequency components.

ResMap

ResMap algorithm attempts to determine the local resolution of density maps by locally fitting a sine-like function [7]. Local resolution is then defined as the wavelength of the fitted sine function. ResMap starts by applying a Gaussian window function centered in the pixel of interest, the standard deviation of this Gaussian establishes the locality of the measurement. Then, a steerable basis functions composed by 17 functions is defined. Any rotation can be expressed as a linear combination of a finite set of these functions. In the *ResMap*, the steerable basis functions, named H2 [16], are the product of a Gaussian function with the second-order Hermite polynomials or their corresponding Hilbert Transforms. The set H2 approximates the sine and cosine functions up to a second order accuracy. These basis functions are used to fit the local densities and determining the wavelength (local resolution) that minimizes the residual between the linear combination and the local density. To determine if the linear combination provides a signal that fits the local density above the noise level a hypothesis test is carried out, establishing a comparison with the background noise. The measurement of this noise level can be carried out by the difference of two half maps or defining a mask that defines a border between signal and noise. The use of two halves is preferable to avoid possible mismatches in the noise statistics (inside and outside the mask), that is the reason why the last version of *ResMap* only accepts this input.

MonoRes

This algorithm [17•] estimates the local resolution by establishing a comparison (hypothesis test) at different frequencies between the energy of local signals and the energy of noise. The frequency for which the local signal at a particular voxel cannot be distinguished from the corresponding voxel in a noise distribution defines the local resolution. The local energy is calculated using the monogenic signal [8,18] of the given map, where the amplitude term defines the strength of the signal. Thus, *MonoRes* starts by high pass filtering the original map from low to high frequency, and estimating at each frequency its corresponding monogenic amplitude. Then, a hypothesis test is performed to determine if the strength of the signal can be significantly detected above the level of noise. This estimation of the local resolution can be carried out using a single map or two half maps. In both cases a mask is required to determine the distribution of noise. In the case of a single map, *MonoRes* assumes that the protein is inside the mask, with only noise outside. In contrast, if two half maps are considered, the difference of the maps provides the noise distribution within the mask. The second method is preferable and recommended. Note that it avoids possible bias in the noise estimation when a map is masked or denoised.

The algorithm in *MonoRes* has also been extended in two new forms: *MonoDir* and *MonoTomo*. *MonoDir* [13••] estimates the local-directional resolution of the reconstructed map and from it, determining the local resolution anisotropy, existence of preferred directions, and possible alignment errors. The directionality is introduced by analyzing cones in Fourier space (directional filter). Once

the map is directionally filtered, the local resolution of that map can be calculated to obtain a local-directional resolution map. Other modifications are also introduced, to avoid the rippling produced by the directional filter and ensuring a correct noise estimation (for details see [13••]). The strengths and weakness are summarized in Table 3. *MonoTomo* estimates the local resolution in tomograms. In electron tomography noise is spatially variant and a mask with only noise cannot be defined as in *MonoRes*. By considering two half tomograms reconstructed under the same alignment parameters both problems are solved. These half tomograms can be obtained by splitting the set of images in two subsets with the odd and even images, or alternatively the set of frames for each tilt angle in two subsets (from a single movie two movie with the half of frames are obtained). This latter method is preferable as the angular sampling is not modified.

DeepRes

This method makes use of deep learning to estimate local resolution of maps [9•]. To do that, a convolutional neural network (CNN), followed by a dense layer, is defined. The training uses atomic models converted into maps. Thus, by low pass filtering the map at different resolutions and considering a small sliding window, the network is trained on the filtering resolution, setting the corresponding weights of the CNN. As a consequence, a model for data is defined and local resolution can be estimated with no more information than the density map and a mask to specify the voxels to be analyzed. Conceptually, the algorithm learns what the shapes of the densities and their textures are at different frequencies. Importantly, it produces different results for original unfiltered maps compared to sharpened maps. Comparing the histograms of these resolution distributions can thus be used to detect over/under-sharpening.

Good practices and thoughts on local resolution

The concept of resolution as a reflection of detail is fraught with different definitions, assumptions on what is actually measured, and different choices made in calculating a quantitative measure. Even the widely accepted and commonly used FSC remains a point of contention in discussions on how detailed a cryoEM map is. This becomes even more complicated when we want to devise a measure of local resolution. In this review the four different algorithms provide considerably different impressions of local detail. Three (*blocres*, *ResMap* and *MonoRes*) are based on the relationship between the signal and noise, but with various assumptions and interpretations. In contrast, *DeepRes* is based on learning typical protein morphology and textures. The distinctions between the results from these methods are often at the level of single voxels, while there is more agreement in larger local regions. Hence, local resolution measure inherently embodies a trade-off between locality and resolution: the smaller an area, the higher the uncertainty in the local resolution value.

From a structural biology point of view, variation in local resolution represents either anisotropy of information (as arising from preferred orientations), or a variation in structural flexibility. These can be distinguished by employing the directional resolution method of *MonoDir*. Resolution (global or local) is a necessary but not a sufficient criterion for being able to interpret structure. It augments visual inspection, providing a statistical basis for either accepting a particular density as a representative structural element. Empirical observation has shown that alpha helices become distinguishable around 10 Å, their helical twists become apparent at 6 Å, beta strands and large side chains appear at 4 Å, and *de novo* building can be done at 3 Å. Nevertheless, and because of the uncertain principle it is not possible localize a given

Figure 1
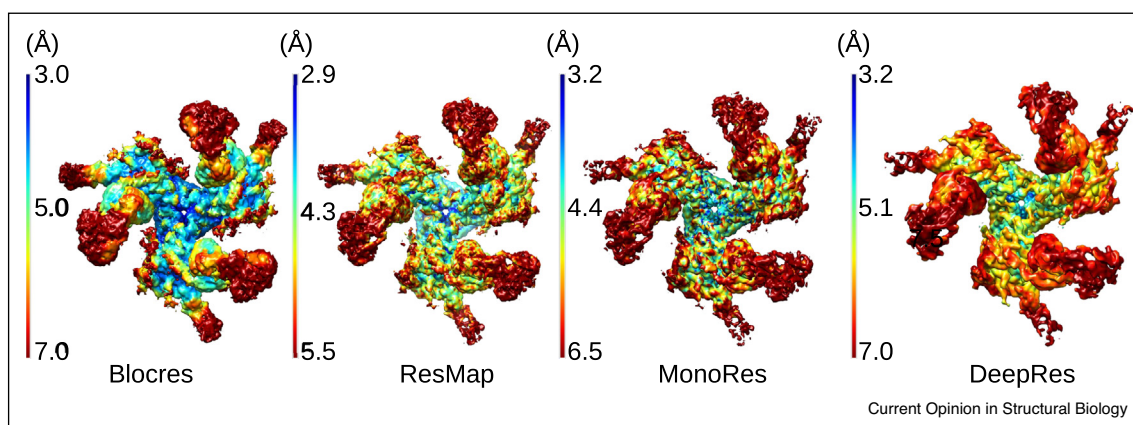
Local resolution maps of the HIV env trimer in complex with CD4 and an antibody [19]. In all cases the core has the highest resolution (blue regions), while the bond proteins at the periphery show the lowest resolution (red regions).

signal in space and frequency at the same time. It makes that the 'true frequency' of this signal is, is not clear — the 'true frequency' could be located anywhere within the frequency uncertainly window.

Performing a local resolution analysis requires an understanding of the algorithms and how to properly apply them. Each method defines local resolution differently, with its corresponding limitations about locality, mask dependency, sharpening, symmetries, resolution range, significance, among many others. For algorithms based on the SNR, it is better to estimate the resolution using two half maps. In cases where a single map is used, a noise model is required, typically calculated from a background area outside the particle density. This may be complicated if the reconstruction underwent a sharpening, denoising or masking procedure. It is therefore important to analyze the maps as 'raw' reconstructions before any further manipulation.

An example

We used a fragment of the HIV virus (EMDB-8713) [19] to estimate the local resolution with all the algorithms described here. This map represents a core of high structural rigidity and a flexible periphery. The results (Figure 1) for the various methods are quantitatively different, but qualitatively similar with resolution ranges around the reported global resolution of 3.7 Å [19]. It is therefore an over-interpretation to infer meaning in individual voxel values. Thus local resolution is both relative in a quantitative sense and spatially diffuse, subject to a form of the uncertainty principle. Proper interpretation is therefore based on an understanding of the balance between the detail in the local resolution map and the fact that it inherently has to be calculated from an region larger than one or a few voxels.

Conclusions

Local resolution analysis enhances our understanding of the variations in detail in cryoEM maps. It is now commonly performed in the final stage of single particle analysis to characterize the reconstruction and to inform the user on potential issues that need addressing. However, if the analysis is not properly done or the results are misinterpreted, it stifles progress. Here we reviewed the main algorithms for estimating local resolution, emphasizing that: Firstly, the definition of local resolution is not unique, it depends on the applied algorithm; secondly, locality and resolution relevance suffers from the uncertainty principle: we called this problem, 'How local is local?'; thirdly, resolution estimation is necessary to judge what we are looking at, but is not sufficient for interpretation; fourthly, apart from DeepRes, local resolution should always be estimated using the raw reconstructions; and finally, if possible, with two half maps. We hope these guidelines help the community in the estimation and interpretation of cryoEM maps. Finally, we look forward

to further development of the idea of local resolution in new implementations that look at it in different ways. All methods discussed here can be used in their corresponding packages or via Scipion [20] as well as the web tools of Scipion [21].

Conflict of interest statement
Nothing declared.

Appendix A. Supplementary data
Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.sbi.2020.06.005.

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of special interest

1. Kuhlbrandt W: **The resolution revolution**. *Science* 2014, **343**:1443-1444.

2. Frank J, Verschoor A, Boublik M: **Computer averaging of electron micrographs of 40s ribosomal subunits**. *Science* 1981, **214**:1353-1355.

3. Unser M, Trus B, Steven A: **A new resolution criterion based on spectral signal-to-noise ratio**. *Ultramicroscopy* 1987, **23**:39-52.

4. Harauz G, van Heel M: **Exact filters for general geometry three dimensional reconstruction**. *Optik* 1986, **73**:146-156.

5. Sorzano COS, Vargas J, Oton J, Abrishami V, de la Rosa-
  • Trevin JM, Gomez-Blanco J, Vilas JL, Marabini R, Carazo JM: **A review of resolution measures and related aspects in 3d electron microscopy**. *Prog Biophys Mol Biol* 2017, **124**:1-30
  A review of global resolution measurements. This article explains all mathematical details behind the resolution. A rigorous analysis.

6. Cardone G, Heymann JB, Steven AC: **One number does not fit all: mapping local variations in resolution in cryo-EM reconstructions**. *J Struct Biol* 2013, **184**:226-236.

7. Kucukelbir A, Sigworth FJ, Tagare HD: **Quantifying the local resolution of cryo-EM density maps**. *Nat Methods* 2014, **11**:63-65.

8. Felsberg M, Sommer G: **The monogenic signal**. *IEEE Trans Signal Process* 2001, **49**:3136-3144.

9. Ramirez-Aportela E, Mota J, Conesa P, Carazo J, Sorzano C:
  • **DeepRes: a new deep-learning- and aspect-based local resolution method for electron-microscopy maps**. *IUCrJ* 2019, **6**
  An algorithm for estimating the local resolution of reconstructed maps based on deep learning.

10. Vilas JL, Oton J, Messaoudi C, Melero R, Conesa P, Ramirez-
  • Aportela E, Mota J, Martinez M, Jimenez A, Marabini R, Carazo JM, Vargas J, Sorzano COS: **Measurement of local resolution in electron tomography**. *J Struct Biol X* 2020, **4**:100016.

11. Tan YZ, Baldwin PR, Davis JH, Williamson JR, Potter CS,
  • Carragher B, Lyumkis D: **Addressing preferred specimen

orientation in single-particle cryo-EM through tilting. *Nat Methods* 2017, **14**:793-796

This manuscript introduces the global resolution anisotropy concept join to the existence of preferred directions.

12. Naydenova K, Russo CJ: **Measuring the effects of particle**
  • **orientation to improve the efficiency of electron cryomicroscopy**. *Nat Commun* 2017, **8**:629

This manuscript introduces the global resolution anisotropy concept.

13. Vilas JL, Tagare HD, Vargas J, Sorzano C, Carazo JM: **Measuring**
  •• **local-directional resolution and local anisotropy in cryo-EM maps**. *Nat Commun* 2020, **11**:55

This references introduces the concept of local directional resolution, local resolution anisotropy, their measurement and its applications.

14. Heymann JB: **Guidelines for using bsoft for high resolution reconstruction and validation of biomolecular structures from electron micrographs**. *Protein Sci* 2018, **27**:159-171 http://dx.doi.org/10.1002/pro.3293.

15. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE: **Ucsf chimera — a visualization system for exploratory research and analysis**. *J Comput Chem* 2004, **25**:1605-1612 http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15264254.

16. Freeman WT, Adelson EH: **The design and use of steerable filters**. *IEEE Trans Pattern Anal Mach Intell* 1991, **13**:891-906.

17. Vilas JL, Gomez-Blanco J, Conesa P, Melero R, de la Rosa
  • Trevin JM, Oton J, Cuenca J, Marabini R, Carazo JM, Vargas J,
  Sorzano COS: **Monores: automatic and accurate estimation of local resolution for electron microscopy maps**. *Structure* 2018, **26**:337-344

The MonoRes method for measuring the local resolution based on monogenic signals is explained. This reference also explains the algorithmic core of MonoTomo and MonoDir, algorithms for estimations of local resolution in tomography and local-directional resolution.

18. Unser M, Van De Ville D: **Wavelet steerability and the higher-order Riesz transform**. *IEEE Trans Image Process* 2010, **19**:636-652.

19. Ozorowski G, Pallesen J, de Val N, Lyumkis D, Cottrell CA, Torres JL, Copps J, Stanfield R, Cupo A, Pugach P, Moore JP, Wilson IA, Ward AB: **Open and closed structures reveal allostery and pliability in the hiv-1 envelope spike**. *Nature* 2019, **547**:360-363.

20. de la Rosa-Trevin JM, Quintana A, del Cano L, Zaldivar-Peraza A, Foche I, Gutierrez J, Gomez-Blanco J, Burguet-Castells J, Cuenca J, Abrishami V, Vargas J, Oton J, Sharov G, Navas J, Conesa P, Vilas JL, Marabini R, Sorzano COS, Carazo JM: **Scipion: a software framework toward integration, reproducibility, and validation in 3d electron microscopy**. *J Struct Biol* 2016, **195**:93-99.

21. Conesa P, Gutierrez J, Quintana A, de la Rosa-Trevin JM, Zaldivar-Peraza A, Cuenca J, Kazemi M, Vargas J, del Cano L, Segura J, Sorzano COS, Carazo JM: **Scipion web tools: easy to use cryo-EM image processing over the web**. *Protein Sci* 2018, **27**:269-275.

4. Supplementary Material

Method	Weaknesses
Blocres	1. The analysis cannot be done on a single map, but requires halfmaps.
	2. The mask must be provided by the user, with the usual care not to introduce high frequency correlations.
	3. The brute force nature of the calculation means that it may be slow, alleviated by running it on a machine with multiple cores.
ResMap	1. Noise is assumed to be Gaussian and white, and therefore a pre-whitening step with user intervention, is required to get a flat spectrum. This has special relevance when the map was sharpened by a B-factor correction, in those cases the pre-whitening must be carried out with special care.
	2. Harsh low pass filtered maps break the condition of white spectra and therefore are not suitable for ResMap.
	3. ResMap requires noise to estimate properly the local resolution, as a consequence, masked maps or denoised reconstructions are not suitable for the algorithm. However, the use of two half maps can overcome this drawback.
MonoRes	1. For masked or denoised reconstructions, it is not possible to estimate the local resolution using a single map, and the use of half maps is required to overcome this problem.
	2. A mask must be provided, despite *MonoRes* is able to create it, it is highly recommendable to provide one.
	3. A resolution range for which the local resolution values are distributed must be provided as an input.
Deepres	1. It lacks of an analytical or mathematical model of local resolution, the model is learnt.
	2. The local resolution presents a dependency with sharpening transformations like B-factor corrections.

Table 2: Weaknesses of the different local resolution algorithms of SPA.

Method	Strengths/Weaknesses
MonoDir Local-directional resolution in SPA	**Strengths** 1. Local resolution anisotropy and its dispersion is calculated. 2. The existence of preferred direction is detected only with the reconstructed map and without access to the set particles that generated such map. 3. It allows to identify the existence of angular alignment error by measuring the slope of the radial average curve of the radial and tangential local-directional resolution maps. **Weaknesses** 1. A mask is required to estimate the local-directional resolution. 2. The box size in which the macromolecule is embedded must be large enough. There must be more than 10 pixel between the macromolecule and the faces of the box to be able of estimating the noise model properly for each direction.
MonoTomo Local resolution in electron tomograms	**Strengths** 1. It is the first local resolution algorithm in electron tomography. 2. Local resolution maps are useful as a segmentation tool and open new possibilities in the field as: weighted subtomograms averaging, local filtering and sharpening, among other. 3. The algorithm is mask free, avoiding the problem of *MonoRes* 4. It is invariant under global isotropic transformations. **Weaknesses** 1. A range in which the local resolution values are must be provided. 2. The algorithm requires two half tomograms with the same alignment parameters. This fact complicates the standard workflow carried out in electron tomography. 3. Splitting the set of images in odd and even, implies that each half tomogram will be reconstructed considering twice the angular sampling of the tilt series. This is avoided by splitting the frames.

Table 3: Strengths and weaknesses for the local direction resolution algorithm of SPA, *MonoDir*, and the local resolution method of electron tomography, MonoTomo.