Journal of Structural Biology 209 (2020) 107447

Contents lists available at ScienceDirect

Journal of Structural Biology

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

Re-examining the spectra of macromolecules. Current practice of spectral quasi B-factor flattening

J.L. Vilas[a], J. Vargas[b], M. Martinez[a], E. Ramirez-Aportela[a], R. Melero[a], A. Jimenez-Moreno[a], E. Garduño[c], P. Conesa[a], R. Marabini[d], D. Maluenda[a], J.M. Carazo[a], C.O.S. Sorzano[a,e,*]

[a] Biocomputing Unit, Centro Nacional de Biotecnologia (CNB-CSIC), Darwin, 3, Campus Universidad Autonoma, 28049 Cantoblanco, Madrid, Spain
[b] Dept. Anatomy and Cell Biology, McGill Univ., Montreal, Canada
[c] Department of Computer Science, Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México, 04510 Mexico City, Mexico
[d] Escuela Politecnica Superior, Universidad Autónoma de Madrid, 28049 Cantoblanco, Madrid, Spain
[e] Univ. San Pablo – CEU, Campus Urb. Monteprincipe, 28668 Boadilla del Monte, Madrid, Spain

ARTICLE INFO

Keywords:
Electron microscopy
Single Particle Analysis
Sharpening
B-factor correction
Structure factor

2010 MSC::
00–01
99–00

ABSTRACT

The analysis of structure factors in 3D cryo-EM Coulomb potential maps and their "enhancement" at the end of the reconstruction process is a well-established practice, normally referred to as sharpening. The aim is to increase contrast and, in this way, to help tracing the atomic model. The most common way to accomplish this enhancement is by means of the so-called B-factor correction, which applies a global filter to boost high frequencies with some dampening considerations related to noise amplification. The results are maps with a better visual aspect and a quasiflat spectrum at medium and high frequencies. This practice is so widespread that most map depositions in the Electron Microscopy Data Base (EMDB) only contain sharpened maps. Here, the use in cryoEM of global B-factor corrections is theoretically and experimentally analyzed. Results clearly illustrate that protein spectra present a falloff. Thus, spectral quasi-flattening may produce protein spectra with distortions when compared with experimental ones, this fact, combined with the practice of reporting only sharpened maps, generates a sub-optimal situation in terms of data preservation, reuse and reproducibility. Now that the field is more advanced, we put forward two suggestions: (1) to use methods which keep more faithfully the original experimental signal properties of macromolecules when "enhancing" the map, and (2) to further stress the need to deposit the original experimental maps without any postprocessing or sharpening, not only the enhanced maps. In the absence of access to these original maps data is lost, preventing their future analysis with new methods.

The use of Single Particle Analysis (SPA) techniques aims at elucidating the 3D structure of macromolecular complexes with the ultimate goal of helping to infer its biological function (Henderson, 2015; Frank, 2018). Macromolecules scatter the incoming electrons and produce images from which the three-dimensional structure is estimated. Once the Coulomb potential map has been reconstructed, it has become customary to correct/enhance the scattering amplitudes in Fourier space using sharpening techniques, the most widely used being the so-called B-factor correction (Rosenthal and Henderson, 2003; Fernandez et al., 2008). This enhancement normally consists in applying a filter that boosts high frequency components. The filter is an exponentially growing filter weighted by the signal to noise ratio whose parameter is estimated using the Guinier plot (logarithm of the squared amplitudes radially averaged of the Fourier transform of the macromolecule *vs* frequency squared). The result in most implementations is a quasi-flat spectrum over a large frequency range. The boost of medium and high frequency components increases the contrast of many structural features of the map and helps to model the atomic structure. In this sense, B-factor correction is a tremendously effective method to increase the interpretability of the reconstructed map. Other sharpening methods based on B-factor also exist, these perform a boosting of medium and high frequencies exhibiting very good results but with certain constraints, for instance the Phenix-AutoSharpen maximizes the map connectivity and minimizes the map isosurface (Terwilliger et al., 2018).

Note that while spectral flattening may be a very useful tool for map analysis, especially oriented to model refinement, it produces maps that

* Corresponding author at: Biocomputing Unit, Centro Nacional de Biotecnologia (CNB-CSIC), Darwin, 3, Campus Universidad Autonoma, 28049 Cantoblanco, Madrid, Spain.
E-mail addresses: carazo@cnb.csic.es, coss@cnb.csic.es (C.O.S. Sorzano).

https://doi.org/10.1016/j.jsb.2020.107447
Received 16 May 2019; Received in revised form 29 December 2019; Accepted 1 January 2020
Available online 03 January 2020
1047-8477/

do not conform with our *a priori* knowledge of the spectrum of a macromolecule, since a spectral decay is expected, larger for electrons than for photons, as we will demonstrate in this work. If it is correctly used as an "inspection tool", the latter inconsistency would not be a problem. However, the reality is that the majority of depositions at the Electron Microscopy Data Base (EMDB) (Lawson et al., 2016) since 2013 only contain these spectral flattened maps, which in practice prevents the access to the experimentally derived information. It is an example of how a "visualization tool" has substituted the experimental data. It is important to understand how we have reached the current situation of the widespread deposition of maps with quasi flats spectra, and we approached this issue by analyzing the fundamental basis underlying B-factor correction. The arguments vary from: 1) a stated theoretical result based on Guinier approximation and Wilson statistics, 2) compensation of excessive amplitude dampening caused by the 3D reconstruction process in general (a combination of inaccuracies in the estimation of the contrast transfer function, CTF of the microscope, inaccuracies in the angular assignment, the high noise level of the original images or numerical approximations of some of the algorithms involved, among others), or 3) the fact that higher frequency content helps to better trace the atomic model and map visualization. With regards to this last point, macromolecule visualization is indeed considerably improved (although noise is also boosted), this is probably the most important reason why it is used, producing good results, as an aid for model tracing.

In this letter, the spectral behavior of macromolecules was studied. The starting point will be the atomic models deposited in the Protein Data Bank (PDB) (Berman et al., 2000), that hypothetically define the ideal protein structure; from them, cryo-EM maps will be calculated using the electron atomic scattering factors and their spectral properties will be analyzed (a number of limitations apply, that will be considered along the manuscript). Additionally, and just for illustrative reasons, we performed some experiments modifying the protein structure of a protein and checking the effects within its spectrum, such as: random displacement of atoms and substitution of carbon by iron atoms, among others.

Based on the studies referred to above, we analyzed the limits of our very common quasi-B factor practice, since scattering theory predicts a certain decay of the protein structure factors (power spectra). In fact, structure factor decay with resolution is an intrinsic property and it is well known in the literature (Shmueli and Weiss, 1995; Feigin and Svergun, 1987; Drenth, 2006). In this way, the starting point will be to rigorously revisit the fundamentals of the scattering theory, in particular, the Guinier approximation (at very low frequencies) and the Wilson statistics (at higher frequencies) (see Appendix), to understand the fundamentals of B-factor correction. This analysis will not show anything new from the point of view of electron scattering theory, but it will provide the context to understand how sharpening is normally performed in cryoEM. Essentially, 1) Guinier approximation is only valid at very low frequencies, and 2) Wilson statistics supports that proteins spectra present a falloff, in contrast, with the widespread idea that it justifies the spectra flattening.

The B-factor quasi-flattening correction has produced very good results aiding in the understanding and interpretation of biological macromolecules, its working mechanism is simple: a boosting of high frequencies weighted by a merit function to avoid noise amplification in some way. Nevertheless, there are certain issues which must be considered: first, once a structure has been reconstructed there is no well accepted criteria about which is the best B-factor to be applied, although some papers have already made some suggestions (Terwilliger et al., 2018; Ramírez-Aportela et al., in press); second, the B-factor correction is a global transformation while the signal to noise ratio is spatially dependent (we remind the reader of the clear success of local resolution analysis in cryoEM) and therefore, a global B-factor choice might enhance some parts of the protein more than required, or produce the reverse. Despite these two issues, the B-factor correction only affects the amplitudes of the Fourier transform of the protein, so the global structure of the macromolecule would not be affected. Taking into account all the above considerations, the B-factor quasi-flattening correction is clearly not the best solution, although it provides good visual results. In this context, we think that its application might be revisited with the aim of developing new algorithms that not only provide nice visual results, but that also produce data that can be more faithfully used in subsequent steps of the processing. Indeed, a clear current trend in the cryoEM field is advancing toward local analysis, and in this context the use of new tools like LocScale (Jakobi et al., 2017) (that searches for the similarity between the structure factor of a model and the structure factor of the map by local scaling of the amplitudes) or LocalDeBlur (Ramírez-Aportela et al., in press) (that applies a local deblurring based on local resolution values of the reconstructed map) seem to be good candidates as alternative methods that overcome this problem. In this regard, we would like to encourage the cryo-EM community to deposit in EMDB the sharpened map, the original reconstruction, as well as other relevant data (Henderson et al., 2012), such as: half-maps, used masks, or FSC curves, among others. Note that undoing a sharpening transformation might be practically impossible and, consequently, the original data are no longer accessible. In addition, to have the original reconstruction opens new horizons to the community, as it allows for the possibility of applying future sharpening algorithms that could enhance the deposited data and, consequently, the biological understanding of the complex.

1. Dependence of the structure factors on macromolecular features

The goal of this section is to analyze the behavior of structure factors at different frequency ranges in a practical manner, and we will approach this issue using simulations. The atomic structure of a macromolecule defines its structure factors (its frequency behavior); a simple way to proceed is by converting the atomic model into a Coulomb potential map (Sorzano et al., 2015) and then determining its spectrum, that will be given by the structure factors. Thus, the logarithm of the structure factor, $Ln|F(q)|^2$, in terms of the frequency square, q^2, was represented for several proteins (see Fig. 1 black line), this is the named Guinier plot. Note that, for simplicity reasons, it is convenient to work with the normalized structure factor $(Ln(|F(q)|^2/|F(0)|^2) - q^2$. In all cases, the profile decays with frequency, in particular, there is a steep decay at very low frequencies that allows distinguishing two regimes: the low and the medium–high frequencies.

The folding of proteins is related to medium-resolution details of the macromolecule. In contrast, Guinier frequencies range describe only the shape of the macromolecule (as we will show in the following section). Thus, we wanted to study how resilient the spectral decay previously indicated was under harsh modifications of the structure, i.e. if the secondary structure was modified or even destroyed, how did it affect the structure factors? In Fig. 1 we present, in continuous black, the spectra of beta-galactosidase (PDB-3j7h), TRPV1 (PDB-3j9j), a triple mutant of the NHAA dimer (PDB-4atv) and the Yeast 20S proteasome in complex with Ac-PAE-ep (PDB-4y6v) up to high resolution, showing a clear spectral decay, far from being flat. In fact, macromolecular spectra behave, in a general way, similar to the ones of typical natural scenes, with a decay in the order of $1/q^2$, where q is the frequency, (van der Schaaf and van Hateren, 1996). We then proceeded to check if this decay was resilient to relatively harsh (and completely artificial) modifications on the macromolecule itself. In order to do so, we artificially changed a number of key characteristics of the maps themselves to verify the effect of these modifications upon the spectra. Thus, the following experiments were carried out:

1. *Experiment 1. Relationship to the atom description*: the atomic description used in the experiment above was performed with very

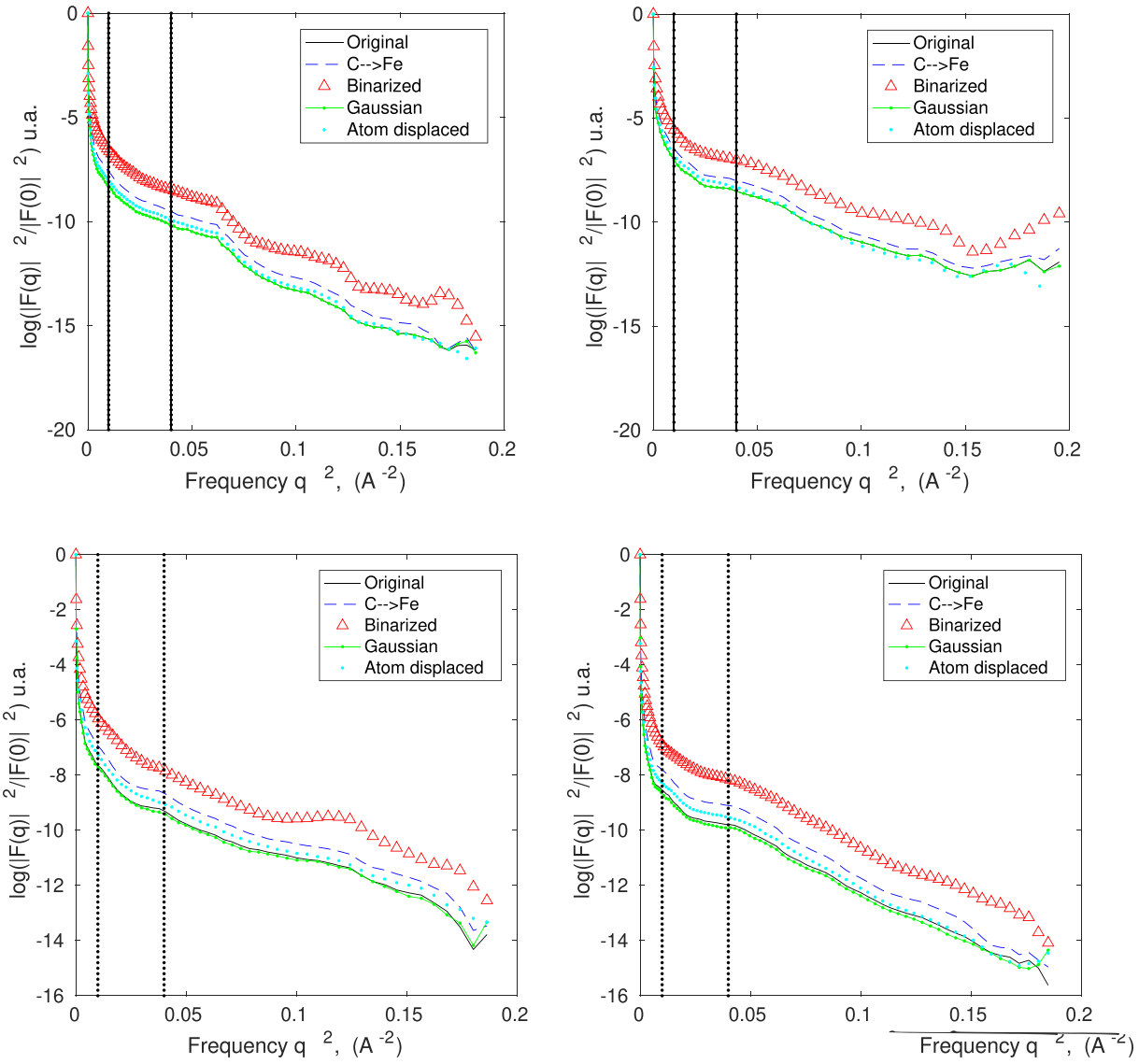

Fig. 1. Normalized structure factors up to high resolution for (upper-left) PDBPDB3j7j7h, (upper-right) 3j9j, (lower-left) PDBPDB4atatv and (lower-right) PDB4 4y6v. a) (Continuous-black) original converted Coulomb potential map, b) (Blue-dashed) Substituted C atoms by Fe atoms. c) (Red-triangles) Binarized Coulomb potential map, d) (Green-continuous-dotted) Map converted from PDB substituting atoms by Gaussian functions, e) (Cyan-dotted) Random displacement of atoms. The vertical lines determine resolutions of 10 and 5 Å.

accurate atomic scattering factors (Sorzano et al., 2015). We used a far simpler atomic description consisting in a Gaussian function whose variance is proportional to the atomic number. In Fig. 1, Green-continuous-dotted line shows that this modification had almost no impact on the structure factor profile.

2. *Experiment 2. Relationship to the atom nature*: we changed all carbon atoms in the atomic model to iron atoms. In Fig. 1, Blue dashed line, can be observed that the structure factor profile was simply shifted up (iron scatters more than carbon), but the overall shape did not significantly change.
3. *Experiment 3. Relationship to the overall macromolecule shape*: we binarized the Coulomb potential map and computed its structure factor profile. This was a major change of the macromolecule representation which kept its overall shape but completely destroyed its internal information. However, as can be seen in Fig. 1, Red triangles line, aside from a global shift of the profile and small differences at high resolution, the structure factor profile did not significantly change.
4. *Experiment 4. Relationship to the relative atomic positions*. The atoms were randomly displaced from their original position with a maximum shift of 6 Å (the random shifts were uniformly distributed between -6 and 6 Å in all directions, note that the radius of an α-helix is 6 Å, and the sideways distance between α carbons in a β sheet is approximately 5 Å). Again, the overall shape of the macromolecule was kept, but the transformation destroyed the information about the secondary structure information. In Fig. 1, Cyan dots, the shape of the structure factor displayed a similar curve to that of the previous experiment, but with a slight shift. The decay was preserved however, the features of the secondary structure were altered completely.

The conclusion from these experiments was clear: structure factors coming from cryoEM maps never present a quasi-flat profile, in spite of some “brutal” modifications introduced in the macromolecular structure. The radial structure factor profile of a macromolecule is an intrinsic feature of its overall shape and one that does not greatly depend upon the type of atoms used for its representation, their shape or their exact position in space (within limits). In all experiments the structure factor presented a non-flat profile exhibiting a decay with the frequency. This fact occurs at all frequency ranges.

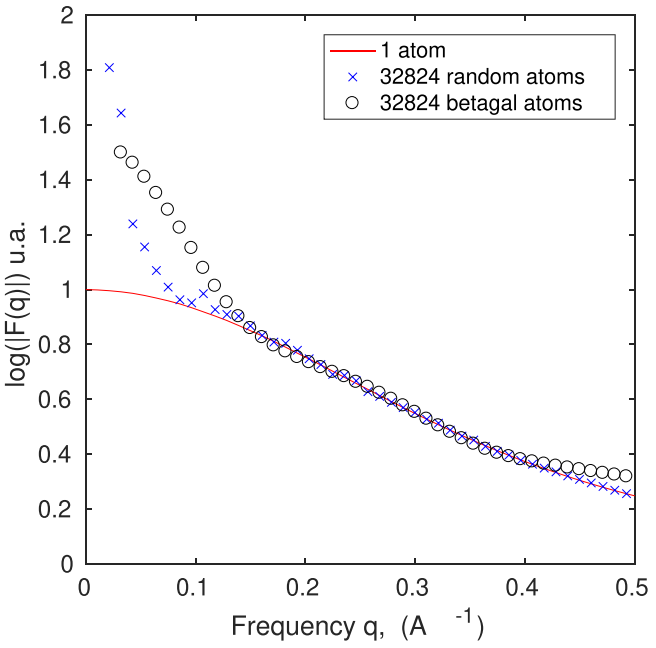

Fig. 2. Structure factor curves in terms of frequency for: (Red-curve) a single carbon atom, (Blue circles) a configuration of 32824 carbon atoms randomly distributed inside a box of 200 × 200 × 200 Å considering the atomic form factor for electrons, (Black crosses) the structure factor of the *beta*-galactosidase.

Finally, it must be commented that the secondary structure will present certain characteristic frequencies as an α-helix pitch, and interatomic distances. The structure factor curve reflects these structural features as peaks at specific frequencies. Note that they are identified as deviations from an ideal falloff (Morris et al., 2004).

2. On the Guinier law

After the first, practical, global analysis presented in the previous section, we now turn to an analytic understanding of both the low and the medium–high frequency regimes; we used scattering theory (Feigin and Svergun, 1987; Guinier and Fournet, 1955; Shmueli and Weiss, 1995; Drenth, 2006), see Appendix. The structure factor is a general expression that depends on the structure geometry and on the nature of the atoms. However, under the small angle scattering approximation, an explicit expression can be obtained, the so-called Guinier law or Guinier approximation (Guinier and Fournet, 1955), as stated more than half a century ago. For completeness, the proof can be found in Section below *On the Guinier law* of Appendix. We were interested in obtaining the quantitative limits of the Guinier law. There exist some validity criteria (Feigin and Svergun, 1987), but to our knowledge, this is the first time that those limits have been analytically determined. The most important reason for performing this analysis is that the B-factor correction is often considered to be connected to Guinier modelling, however there is no connection whatsoever. Indeed, Guinier law applies only at very low frequencies, while spectral quasi-flattening is an operation performed at the medium–high resolution range. In general, Guinier approximation is valid for resolutions lower than the gyration radius of a macromolecule. Furthermore, it is not about flattening the spectrum, but instead it states that at very low frequencies (in the order of the radius of gyration of the macromolecule) the structure factors decay as an exponential function.

3. On Wilson statistics

Wilson statistics is also usually employed to support B-factor correction. Consequently, we analyzed the fundamentals behind the quasi-flattening spectra based on Wilson statistics. Again, the starting point of our analysis was the scattering theory (Feigin and Svergun, 1987; Guinier and Fournet, 1955; Shmueli and Weiss, 1995; Drenth, 2006) (see Appendix) and the general expression for the structure factor, see Eq. (A.17) in the Appendix. The basic hypothesis underlying Wilson statistics is that atoms are randomly distributed on the "unit cell" of the protein (Wilson, 1949). Following this principle, and as presented in Section *On the Wilson statistics* of Appendix, the expected value for the structure factors under this hypothetical distribution results in the sum of the form factor of all atoms, that is, the protein atomistic nature. However, this result must be further analyzed. First, macromolecules largely violate the hypothesis that atoms are randomly distributed on the "unit cell". Second, the expected value as sum of atomic form factor fails at low frequencies (Guinier approximation) and when the atoms are very close to each other. Third, the expected value of the structure factor is not a constant, the form factor of atoms presents a falloff with the frequency that implies a decay of the structure factor. To analyze macromolecular spectral features and Wilson statistics, the following experiment was carried out. A model was created consisting of 32,824 atoms randomly distributed inside a box with dimensions 200 × 200 × 200 Å. The number of atoms was chosen in order to establish a comparison with the β - galactosidase that has the same number of atoms. For the sake of simplicity, all atoms were considered to be the same as carbon atoms. Then, the radial average of the structure factor for such configuration was calculated. To do that, the form factor of atoms were simulated as sum of Gaussians according to (Sorzano et al., 2015; Prince, 2006), and neglecting thermal vibrations. Since all atoms are equal, the expected value of the structure factor must be proportional to the form factor of a single atom (see Eq. (B.8)), this is exactly what Fig. 2 shows. Note that the structure factor of the randomly placed atoms (Blue-circles) presents a decay with the frequency confirming our assessment of non-flat spectra and following the curve of a single atom (red curve) at medium and high frequencies but exhibiting differences at low frequencies. In addition, we also wanted to compare this curve with the one obtained for the β - galactosidase, obtaining a structure factor curve close to the one obtained for the randomly placed atoms except at low frequencies. Interestingly, the results from this experiment would suggest that although some basic assumptions required by Wilson statistics are largely violated in macromolecules, the effect of this violation is still only significant at low frequencies, a regime for which Guinier law provides already a very good model.

Finally, we wanted to check if the flat structure factor of protein could be obtained when considering atomic form factor of atoms for X-rays. Hence, we decided to estimate again the atomic form factor of a single carbon atom, and compare it with the one derived from the scattering form factor for electrons. This comparison can be observed in Fig. 3. The atomic from factors for X-rays were taken from (Prince, 2006). Note that the structure factor plot for X-rays is always higher than the one obtained for electrons. However, both show a similar behavior with a non-negligible decay with the frequency. It should be noted, however, that the electron form factors values used in these experiments, come from the classical Crystallography Tables (Prince, 2006). These do not capture changes due to interatomic interactions and consequently, changes to these graphs should be expected in the future, but currently is all that can be done. In fact, some works have been published in this regard (Yonekura et al., 2017).

However, the results with non-flat profiles pose the question as to why the quasi-flattening spectra based on B-factor enhances the visualization so much? Intuitively, this question can be answered from the uncertainty principle that establishes that when an object is well localized in space then, it is not well localized in frequency terms (the opposite implication is also true). Thus, what flattening is doing is delocalizing the protein in frequency, and as a consequence, its shape becomes sharper in the real space.

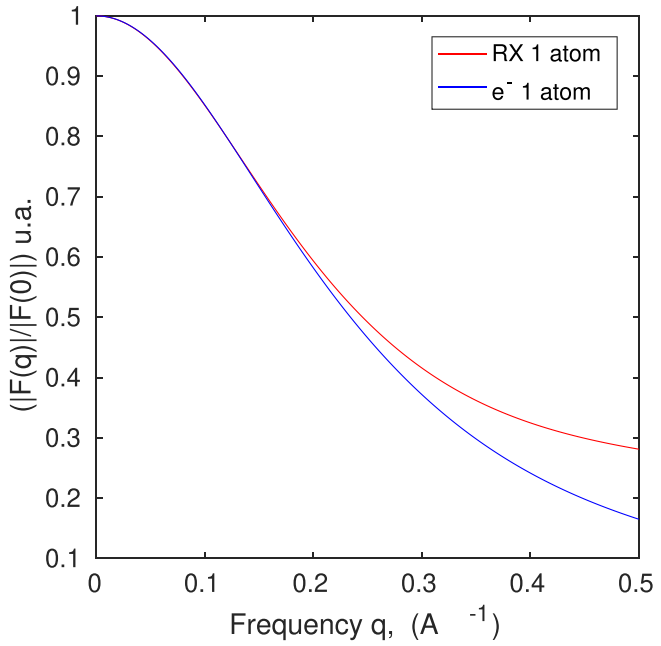

Fig. 3. Atomic form factor curves in terms of frequency for a single carbon atom looking at for X-ray (Red curve) and electron scattering (Blue curve).

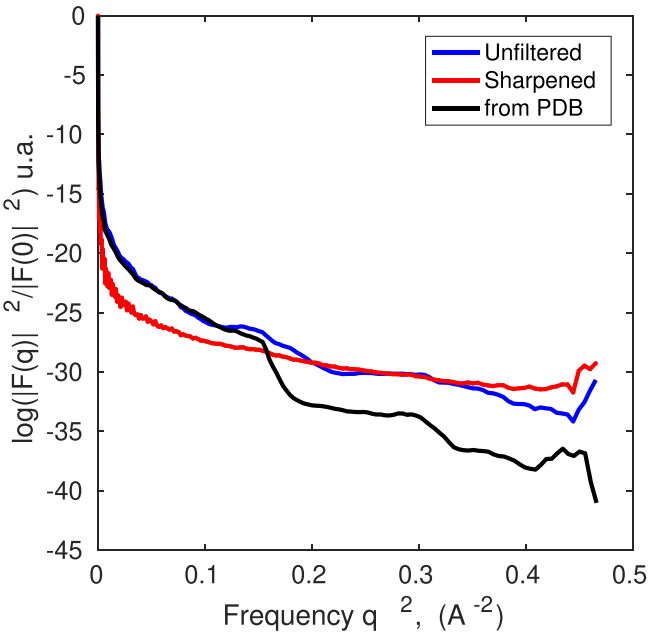

Fig. 4. Normalized structure factor from (Blue) unfiltered map, (Red) B-factor corrected map and (Continuous-black) atomic model (PDB entry 3j7h) converted into Coulomb potential map.

A more complex but accurate explanation comes again from the scattering theory, and it is explained in depth in Appendix B. This explanation starts considering ideal atoms that lack any structure and are, therefore, just points in the space (dimensionless points). Note that in the absence of structure, the form factor of atoms is constant, this is because if the atoms represent a single point all scattered electrons have the same origin (the point) and there is no phase difference among them. This fact implies that the expected value of the structure factor in Wilson statistics should result in a constant. As a consequence, for the B-factor correction that makes flat spectra, what it is trying to achieve is to make comparable the measured structure factors with the structure factors of point atoms. Due to the duality between positions and frequencies, the larger variance in frequency and the lesser uncertainty in position, this transformation makes the map sharper. Unfortunately, atoms do present structure and thus this model is not representative.

4. B-factor correction to compensate overdampened spectra

Another argument for performing B-factor quasi-flattening correction could be to compensate for over-dampened map spectra associated to the accumulation of errors in image processing steps. Indeed, it might be the case that the reconstruction process has over-dampened the spectrum of the macromolecular map, and that we need to boost high frequencies to recover the original spectral fall-off. However, it has not been proven that reconstruction algorithms necessarily over-dampen high frequency components or, expressed in a more general manner, that they have not already incorporated features to compensate for this sort of over-dampening. In Fig. 4 we show the structure factor profile for β-galactosidase obtained in the Map Challenge (Heynmann, 2018) together with the structure factor profile of the map calculated from the corresponding PDB entry 3j7h. As shown in Fig. 4, the structure factor of the PDB falls faster (implying a larger B-factor) than the processed reconstruction. Although this behavior is not necessarily the case for all reconstructions, in general we have not appreciated any obvious spectral over-dampening in experimental cryo-EM maps with respect to their corresponding models, implying that the algorithms used in the reconstruction process have carefully avoided spectral attenuation and that, therefore, the use of B-factor quasi flattening is not justified on the basis of a posterior compensation of potential defects of the reconstruction algorithms.

5. Conclusions

Based on these considerations, and witnessing how quickly cryo-EM is advancing to produce quantitative estimations of the Coulomb potential of a macromolecule, we consider it is the appropriate time to raise the important issue that spectral quasi-flattening, currently the most widely used post-processing enhancement operation applied to a cryoEM map, should be used with great care, although it certainly produces visually appealing maps, these maps have the wrong spectral characteristics. It might be argued that spectral quasi-flattening is only an “aid”, a tool, for tracing; however, the hard fact is that the majority of all cryoEM entries deposited in EMDB since 2013 only contain these flattened maps, not the experimentally derived one. It is the case of how an otherwise useful inspection tool has led to the substitution of the experimental data themselves, completely removing the capacity to access any more detailed quantitative cryo-EM information from a large set of EMDB entries.

Obviously, there are several ways to improve current practices. The most obvious and immediate change would be the deposition of both, the unsharpened and sharpened maps, as well as any other experimental data, chiefly the usual “two half maps” that are part of virtually any Single Particle Analysis workflow. Currently this practice is recommended in EMDB, but given the evidence of possible errors, as presented in this paper, it provides additional grounds to propose making this practice compulsory, rather than merely recommended.

Other suggestions go in the direction of developing more careful sharpening techniques that do not enforce a quasi spectral flattening, but that tend to provide results that follows more closely the spectra of experimentally solved macromolecules. It is worth noting that some recent proposals of new post-processing (sharpening) methods have more or less explicitly started to deviate from global B-factor quasi flattening, going from localized B-factor flattening (Jakobi et al., 2017) to local map re-scaling (Hryc et al., 2017) or the recent one based on local resolution information LocalDeBlur (Ramírez-Aportela et al., in press).

In short, in this communication we have presented a framework in

which the limitations of B-factor quasi-flattening have been systematically explored, clarifying a number of misconceptions in the field and trying to contribute to the establishment of better deposition schema and to the development of a new generation of sharpening algorithms.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The author would like to express their gratitude to Dr. J.J. Fernandez and Morag Gunson for their interest, following and advice on this work. The authors would like to acknowledge economical support from The Spanish Ministry of Economy and Competitiveness through Grants BIO2016-76400-R(AEI/FEDER, UE), the "Comunidad Autónoma de Madrid" through Grant: S2017/BMD-3817. The authors acknowledge the support and the use of resources of Instruct-ERIC.

Appendix A. On the Guinier approximation

A.1. Scattering theory

Consider a scatterer element located in position $\mathbf{r}_k$ and an incident electron beam in the direction defined by the unitary vector $\mathbf{s}_0$. Electron interaction with the charge density of the scatterer element will cause the scattering of the incident electron beam. The direction of the scattered beam will be determined by the unitary vector $\mathbf{s}$, which forms an angle 2θ with respect to the direction of the incident beam, $\mathbf{s}_0$. A scheme of the system is shown in Fig. A.5. The goal will be to calculate the intensity of the output beam.

If the macromolecule is composed by N atoms, then, the problem is equivalent to obtaining the output intensity of a set of N small scatterers. Scattering is the consequence of the electron-matter interaction. Assuming a thin sample and taking into account that energies in TEM are high enough ($100keV - 1MeV$), as it is the most common case in TEM for the imaging of biological samples, there are two possible scenarios: the incident electron can interact once with the matter, or it can be transmitted through the sample without any interaction. In both cases the energy of the transmitted beam is conserved, being in the first one an elastic interaction. The reason for the scattering will be the electrostatic interactions being macromolecule described by its Coulomb potential $\rho(\mathbf{r})$.

In a far point, P, from the macromolecule, a set of N elastic scattered plane waves will interfere. This interference will depend on the optical path difference. Without loss of generality, the reference system can be fixed in the mass center of the macromolecule. Thus, the phase difference in a far point P between a reference ray through the mass center and a scatterer element at the location $\mathbf{r}_k$ will be proportional to $\mathbf{s}\cdot\mathbf{r}_k - \mathbf{s}_0\cdot\mathbf{r}_k = (\mathbf{s} - \mathbf{s}_0)\cdot\mathbf{r}_k$, see Fig. A.5 (here we assume that the observer is far from the scatterer and that the phase difference is independent of the observer position, Fraunhofer regime). For the sake of clarity the notation, we define the scattering vector $\mathbf{q}$ as:

$$\mathbf{q} = \frac{2\pi}{\lambda}(\mathbf{s} - \mathbf{s}_0). \tag{A.1}$$

and its absolute module will be:

$$q = \frac{4\pi}{\lambda}\sin\theta, \tag{A.2}$$

being λ the wavelength of the electron beam. Note that q has frequency units ($1/\text{Å}$).

The incident wave at the position $\mathbf{r}_{k,\lambda}$, will be given by the wave function of the incident beam with amplitude A and frequency, $2\pi/\lambda$, as $\Psi_{in} = Ae^{i\frac{2\pi}{\lambda}\mathbf{s}_0\cdot\mathbf{r}_k}$. Defining an arbitrary point, P, by the position vector $\mathbf{l}$, measured from $\mathbf{r}_k$ along the scattering direction $\mathbf{s}$; the path difference allows us to establish a linear relation between the input and scattered wave functions, $\Psi_k(\mathbf{q})$, as follows:

$$\begin{aligned}\Psi_k(\mathbf{s}) &= \Psi_{in} f_{k,\lambda} e^{-i\frac{2\pi}{\lambda}\mathbf{s}\cdot(\mathbf{r}_k+\mathbf{l})} \\ &= Ae^{i\frac{2\pi}{\lambda}\mathbf{s}_0\cdot\mathbf{r}_k} f_{k,\lambda} e^{-i\frac{2\pi}{\lambda}\mathbf{s}\cdot(\mathbf{r}_k+\mathbf{l})}\end{aligned} \tag{A.3}$$

where $f_{k,\lambda}$ is the electron atomic scattering factor at wavelength λ. Equivalently, we could have used the $\mathbf{q}$ variable instead of $\mathbf{s}$:

$$\Psi_k(\mathbf{q}) = Af_k(q)e^{-i\mathbf{q}\cdot\mathbf{r}_k}e^{-i\frac{2\pi}{\lambda}\mathbf{s}\mathbf{l}}, \tag{A.4}$$

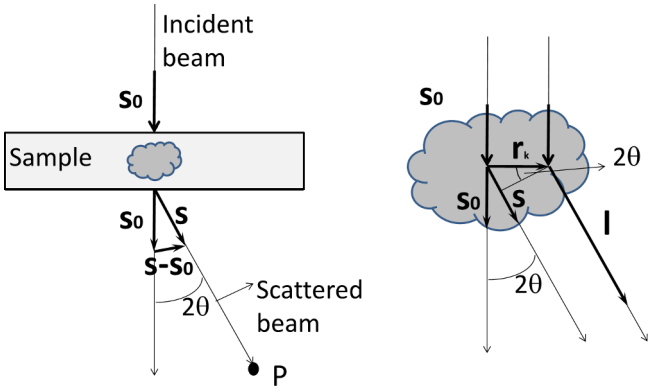

Fig. A.5. Scheme of the scattering system. An incident beam lights a scatterer element and the beam is scattered forming an angle 2θ respect to the original incident direction.

Note that there is a direct relationship between $f_k(q)$ and $f_{k,\lambda}$ since the only change is the conversion from wavelength to frequency variables.

The output beam is determined by the contribution of many scattered waves, therefore, by applying the superposition principle, the scattered wave at the point P can be written by means of:

$$\Psi(\mathbf{q}) = \sum_{k=1}^{N} \Psi_k(\mathbf{q}) = Ae^{-i\frac{2\pi}{\lambda}\mathbf{sl}} \sum_{k=1}^{N} f_k(q)e^{-i\mathbf{q}\cdot\mathbf{r}_k}, \tag{A.5}$$

where N represents the number of scatterers, i.e. atoms in the macromolecule. Let us now relate this equation to the Coulomb potential of the macromolecule, $\rho(\mathbf{r})$. By definition, this Coulomb potential, from the electron scattering point of view can be seen as the superposition of the electron scattering form factors:

$$\rho(\mathbf{r}) = \sum_{k=1}^{N} b_k(\mathbf{r}-\mathbf{r}_k) = \sum_{k=1}^{N} b_k(\mathbf{r}) \bigstar \delta(\mathbf{r}-\mathbf{r}_k), \tag{A.6}$$

where $b_k(\mathbf{r})$ is the function resulting from the inverse Fourier transform of the $f_k(\mathbf{q})$ coefficients:

$$b_k(\mathbf{r}) = \int f_k(\mathbf{q})e^{i\mathbf{q}\cdot\mathbf{r}}d\mathbf{q} \tag{A.7}$$

If we now do the Fourier transform of this equation, we obtain:

$$F(\mathbf{q}) = \int \rho(\mathbf{r})e^{-i\mathbf{q}\mathbf{r}}d\mathbf{r} = \sum_{k=1}^{N} f_k(\mathbf{q})e^{-i\mathbf{q}\cdot\mathbf{r}_k} \tag{A.8}$$

By comparing Eq. (A.5) with Eq. (A.8) we clearly see that:

$$\Psi(\mathbf{q}) = Ae^{-i\frac{2\pi}{\lambda}\mathbf{sl}}F(\mathbf{q}) \tag{A.9}$$

Nevertheless, detectors record intensity instead of wave functions:

$$I(\mathbf{q}) = |\Psi(\mathbf{q})|^2 = A^2|F(\mathbf{q})|^2 = A^2 \iint \rho(\mathbf{r}')\rho(\mathbf{r})e^{-i\mathbf{q}\cdot\mathbf{r}'}e^{i\mathbf{q}\cdot\mathbf{r}}d\mathbf{r}d\mathbf{r}'. \tag{A.10}$$

The value $F(\mathbf{q})$ is called the structure factor and it is analyzed in the next section.

A.2. Structure factor

The structure factor given by Eq. (A.10) can rarely ever be integrated in terms of elemental functions, and usually requires numerical integration. The Coulomb potential, $\rho(\mathbf{r})$, determines the shape of the macromolecule. Therefore, the structure factor of only a few geometries can be calculated analytically. The goal of this section is to provide a general expression for integrating the structure factor under specific conditions.

The first step is to simplify Eq. (A.10) by means of the variable change $\hat{\mathbf{r}} = \mathbf{r}' - \mathbf{r}$,:

$$|F(\mathbf{q})|^2 = \iint \rho(\mathbf{r})\rho(\hat{\mathbf{r}}+\mathbf{r})e^{-i\mathbf{q}\cdot\hat{\mathbf{r}}}d\mathbf{r}d\hat{\mathbf{r}}. \tag{A.11}$$

Note that the expression of $|F(\mathbf{q})|^2$ must be an even function. Indeed, $|F(-\mathbf{q})|^2 = |F^*(\mathbf{q})|^2 = |F(\mathbf{q})|^2$.

Let us define the autocorrelation function as:

$$\gamma(\hat{\mathbf{r}}) = \int \rho(\mathbf{r})\rho(\hat{\mathbf{r}}+\mathbf{r})d\mathbf{r}. \tag{A.12}$$

Note that if the scatterer element is effectively limited to a maximum radius R, then the autocorrelation function is limited to $2R$. With this function, the integral Eq. (A.11) gets even simpler:

$$|F(\mathbf{q})|^2 = \int \gamma(\hat{\mathbf{r}})e^{-i\mathbf{q}\cdot\hat{\mathbf{r}}}d\hat{\mathbf{r}}. \tag{A.13}$$

That is, the structure factor is the Fourier transform of the autocorrelation function.

Let us consider the structure factor integral in spherical coordinates with the angle between $\hat{\mathbf{r}}$ and $\mathbf{q}$ being β :

$$|F(\mathbf{q})|^2 = \iiint \gamma(\hat{r}\sin\beta\cos\phi,\ \hat{r}\sin\beta\sin\phi,\ \hat{r}\cos\beta)e^{-iq\hat{r}\cos\beta}\hat{r}^2\sin\beta d\hat{r}d\beta d\phi. \tag{A.14}$$

For simplicity, let us assume that autocorrelation function has radial symmetry, then the structure factor also has radial symmetry (another way of reasoning and reaching the same point would be by calculating the radial average along a given frequency):

$$\begin{aligned} |F(q)|^2 &= \iiint \gamma(\hat{r})e^{-iq\hat{r}\cos\beta}\hat{r}^2\sin\beta d\hat{r}d\beta d\phi \\ &= 4\pi\int \hat{r}^2\gamma(\hat{r})\mathrm{sinc}(q\hat{r})d\hat{r} \end{aligned} \tag{A.15}$$

Note that:

$$|F(0)|^2 = 4\pi\int \hat{r}^2\gamma(\hat{r})d\hat{r} \tag{A.16}$$

and we can write the structure factor as a function of its value at zero frequency:

$$|F(q)|^2 = |F(0)|^2\frac{\int \hat{r}^2\gamma(\hat{r})\mathrm{sinc}(q\hat{r})d\hat{r}}{\int \hat{r}^2\gamma(\hat{r})d\hat{r}}, \tag{A.17}$$

where $\mathrm{sinc}(q\hat{r})$ function is defined as $\sin(q\hat{r})/q\hat{r}$.

A.3. Validity of the Guinier law

Virtually all works about sharpening start talking about the Guinier law. The essentials are simple, in the sense that Guinier (Guinier and Fournet, 1955), based on certain approximations, could predict the shape of the spectrum of a macromolecule up to a given (relatively low) resolution. However, the rest of the spectrum has nothing to do with Guinier explanation. So, in this work we wanted to analyze the extend of Guinier approximations so that we could trace a clearer boundary between what is Guinier-based and what is not.

In this section we discuss about the Guinier law and its applicability to sharpening by B-factor correction. We show that the Guinier law is derived after two critical approximations that restrict its validity to a very small frequency region. In particular, the two approximations are: 1) the Taylor expansion of the *sinc* function, and 2) the Guinier approximation of a second order polynomial by an exponential.

A.3.1. Approximation 1: Taylor approximation of the sinc.

Guinier's approximation (Guinier and Fournet, 1955) considers small angle scattering, justifying a Taylor approximation up to second order of the sinc function to integrate Eq. (A.17) resulting in

$$|F(q)|^2 \approx |F(0)|^2\left(1-\frac{R_g^2}{3}q^2\right), \tag{A.18}$$

where $R_g^2=\frac{1}{2}\frac{\int \hat{r}^4\gamma(\hat{r})d\hat{r}}{\int \hat{r}^2\gamma(\hat{r})d\hat{r}}$, is the so called gyration radius. Note that this Taylor expansion is of 3rd order because the *sinc* is an even function and it lacks odd power terms in its Taylor expansion. To check the validity of this approximation, we have calculated the relative error between the *sinc* function and its polynomial approximation:

$$\epsilon=\left|\frac{\mathrm{sinc}(q\hat{r})-(1-\frac{1}{3!}(q\hat{r})^2)}{\mathrm{sinc}(q\hat{r})}\right| \tag{A.19}$$

The error committed in making this approximation is always less than 10% for the range $q\hat{r}\leqslant 5/3$. Since this approximation is performed within an integral (see Eq. A.17), $\hat{r}$ takes the value R (particle radius) in the worse case, implying that $qR\leqslant 5/3$. This is only achieved for resolutions, $1/q$, larger than $3R/5$. That is, for a macromolecule of radius $R=100$ Å, the frequencies for which the *sinc* approximation is valid (the relative error is smaller than 10%) are lower than 60 Å.

A.3.2. Approximation 2: Approximation of a second order polynomial by an exponential

The second Guinier step is to approximate the polynomial in Eq. A.18 by an exponential function

$$|F(q)|^2 \approx |F(0)|^2 e^{-\frac{R_g^2}{3}q^2} \tag{A.20}$$

The product $B=\frac{1}{3}R_g^2$ is called the B-factor. This is the parameter being estimated in electron microscopy and being corrected for. As we did in the previous section, we analyze the relative error of this approximation

$$\epsilon=\left|\frac{\left(1-\frac{1}{3}R_g^2q^2\right)-e^{-\frac{1}{3}R_g^2q^2}}{1-\frac{1}{3}R_g^2q^2}\right| \tag{A.21}$$

This equation can be easily solved with the help of the W-lambert function (Corless et al., 1996). In our case

$$\frac{1}{3}R_g^2q^2=1+W\left(-\frac{1}{(1+\epsilon)e}\right) \tag{A.22}$$

For $\epsilon=10\%$, we have $\frac{1}{3}R_g^2q^2=0.3755$, the error is smaller than 10% for those frequencies such that

$$qR_g\leqslant 1.06 \tag{A.23}$$

This is only achieved for resolutions, $1/q$, larger than $R_g/1.06$. That is, for a macromolecule of radius of gyration $R_g=100$ Å, the frequencies for which the polynomial by exponential approximation is valid (the relative error is smaller than 10%) are those lower than 94 Å, or 100 Å for the sake of roundness. Eq. (A.23) establishes a criterion by which the Guinier approximation is valid. It is noteworthy that the gyration radius is related to the macromolecular radius. As a consequence, the Guinier approximation is only valid at very low frequencies. Even if higher errors are considered, the limits of Guinier approximation only remain valid at low frequencies; in fact, if an error higher than 20% is accepted for a macromolecule of $R_g=100$ Å, then the Guinier range will be valid for frequencies lower than 77 Å only.

A.4. Experimental verification of the Guinier law

In this section, we verify that the statements presented above are experimentally observed. With this aim, the atomic model of a structure of β-galactosidase (Bartesaghi et al., 2015) (PDB-3j7h), TRPV1 (Liao et al., 2013) (PDB-3j9j), a triple mutant of the NHAA dimer (Lee et al., 2014) (PDB-4atv) and the Yeast 20S proteasome in complex with Ac-PAE-ep (Huber et al., 2015) (PDB-4y6v) were used. The PDB provides the atomic structure of the macromolecule. As a consequence, the corresponding structure factor can be derived from it without any ambiguity. We converted the atomic models into Coulomb potential maps at a sampling rate of 1Å/pixel using `xmipp_volume_from_PDB` (Sorzano et al., 2015) in Scipion v.1.1 (de la Rosa-Trevin et al., 2016). The Guinier plot maps the structure factor in terms of the frequency square, $\log(|F(q)|^2)$ vs. q^2. For simplicity, it is convenient to work with the normalized structure factor $\log(|F(q)|^2/|F(0)|^2)$.

As can be seen in Fig. A.6, the spectrum follows a straight line up to about 100 Å (marked with a vertical line in Fig. A.6); this is exactly what would be expected in a proper use of scattering theory. However, beyond this point, the theory is not applicable and any additional processing steps

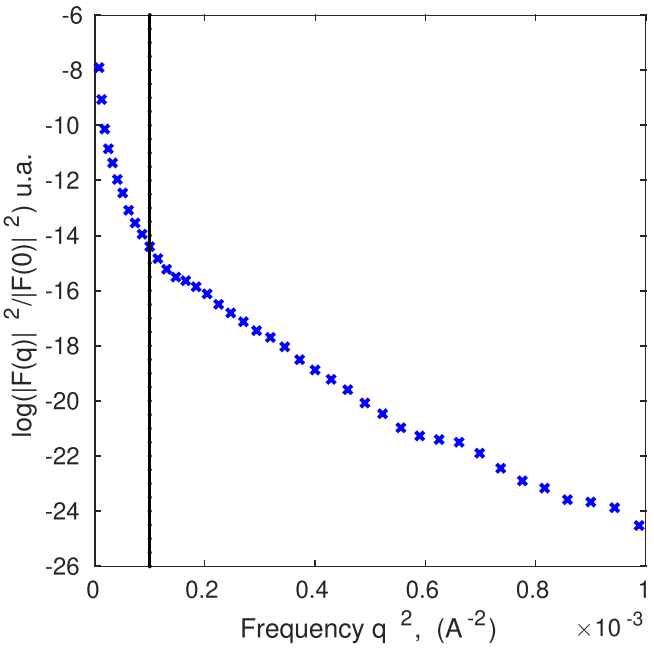

Fig. A.6. Guinier plot for the atomic model PDB 3j7h converted into Coulomb potential map, on the low frequency range up to 30 Å. Note that the Guinier approximation is shown as a straight line for resolutions lower than 100Å.

are proposed must be based on other principles. In particular, it is clear that past 100 Å the spectrum profile is not flat, so the theory cannot be used to justify a quasi B-factor flattening of the spectrum, in agreement with the results obtained by the scattering theory.

Appendix B. On Wilson statistics

This appendix is divided in two parts: The form factor of atoms and Wilson statistics respectively. The first part covers part of the scattering theory that was not explained in Appendix A, the scattering of electrons by single atoms, and how to model this interaction. It will introduce the basic concepts that are required to fully understand the second part, in which the expression of the structure factor at medium and high frequencies are obtained and the Wilson statistics are discussed.

B.1. Form factor of atoms

Let us consider an individual atom that is lighted by an electron beam along the direction $\mathbf{s}_0$. The electromagnetic interaction between the radiation and the electrostatic potential of the atom results in the electron scattering. To understand this scattering is necessary to model the form factor of the atom. First, atoms presents a given structure that can be considered as rotational symmetric (despite the orbital shapes make false this assumption it is a good approximation, orbital s approximation). Second, it is also known that atoms position are modulated by a distribution as a consequence of thermal vibrations. Hence, under all these hypothesis the form factor of the atom is modelled as (Rupp, 2010)

$$f_{at}(q) = f_{st}(q) \cdot f_{th}(q), \tag{B.1}$$

where the terms f_{st} and f_{th} represent the form factor for the structural component of the atom and thermal vibrations respectively (normally modelled as a Gaussian with a temperature factor). The structural component, is usually modelled as sum of Gaussians (Peng et al., 1996; Sorzano et al., 2015)

$$f_{st}(q) = \sum_{i=1}^{5} a_i e^{-b_i q^2}, \tag{B.2}$$

with a_i and b_i constants that depends on the atom nature. Note that if the atom is idealized as a point, lacking of structure, then, whatever two scattered directions will present the same phase difference. In other words, the phase difference is a consequence of the structure, see Fig. A.5, because the phase difference appears when there is a difference of optical path length between two rays and in this case all rays have the same optical path.

B.2. Wilson statistics

It has been proved that Guinier law can only be applied at very low resolutions. It is yet to be known what happens in the medium and high resolution range. The starting point is the most general expression of the structure factor (Eq. A.8), thus, assuming that the protein is defined by a finite number of atoms, N,

$$F(\mathbf{q}) = \sum_{k=1}^{N} f_k e^{-i\mathbf{q}\cdot\mathbf{r}_k}, \tag{B.3}$$

where the dependence of atom form factor f_k on the scattering vector $\mathbf{q}$ was omitted for simplicity, and it is considered as product of the position,

structure and thermal factors, Eq. (B.1). The exact expression of the square of the structure factor will be calculated as

$$|F(\mathbf{q})|^2 = \left[\sum_{k=1}^{N} f_k(q)\cos(\mathbf{q}\cdot\mathbf{r}_k)\right]^2 + \left[\sum_{k=1}^{N} f_k(q)\sin(\mathbf{q}\cdot\mathbf{r}_k)\right]^2. \tag{B.4}$$

which can be expanded as

$$|F(\mathbf{q})|^2 = \sum_{k=1}^{N} f_k^2(q) + \sum_{\substack{k,j \\ k\neq j}}^{N} f_k(q)f_j(q)[\cos(\mathbf{q}\cdot\mathbf{r}_k)\cos(\mathbf{q}\cdot\mathbf{r}_j) + \sin(\mathbf{q}\cdot\mathbf{r}_k)\sin(\mathbf{q}\cdot\mathbf{r}_j)], \tag{B.5}$$

and rewritten

$$|F(\mathbf{q})|^2 = \sum_{k=1}^{N} f_k^2(q) + \sum_{\substack{k,j \\ k\neq j}}^{N} f_k(q)f_j(q)\cos(\mathbf{q}\cdot(\mathbf{r}_k - \mathbf{r}_j)). \tag{B.6}$$

Up to this point, the structure factor expression is absolutely general. What Wilson statistics assumes is that atoms are randomly distributed inside the protein, it means, that the position distributions $\mathbf{r}_k$ is uniform (this hypothesis will be discussed latter). Thus, the expected value of the structure factor in is calculated

$$\mathbb{E}[|F((\mathbf{q})|^2] = \mathbb{E}\left[\sum_{k=1}^{N} f_k^2(q)\right] + \mathbb{E}\left[\sum_{\substack{k,j \\ k\neq j}}^{N} f_k(q)f_j(q)\cos(\mathbf{q}\cdot(\mathbf{r}_k - \mathbf{r}_j))\right], \tag{B.7}$$

and therefore

$$\mathbb{E}[|F((\mathbf{q})|^2] = \sum_{k=1}^{N} |f_k(q)|^2. \tag{B.8}$$

However, Wilson assumption must be discussed. Note, that the interference terms (cosine terms) are neglected as a consequence of the oscillating character of the cosine function. However, there exist two scenarios in which it cannot be neglected: first, at very low frequency, the cosines poorly oscillate, this behaviour supports that Wilson statistics only apply at high frequency and that low ones are governed by the Guinier approximation; second, when the atoms are very close to each other, the vector $\mathbf{r}_k - \mathbf{r}_j$ will be very small casting again slow cosine oscillations (this is exactly the case of biological macromolecules); third, it is not necessary that atoms are randomly distributed, in fact, reality is different and proteins exhibit atoms at specific positions. Fortunately, the same result can be obtained if the radial average of the structure factor is calculated, thus, the interference term is neglected.

$$|F(q)|^2 = \sum_{k=1}^{N} |f_k(q)|^2. \tag{B.9}$$

Finally, it must be highlighted that the expected value or radial average of the structure factor still presents a dependence on the frequency, q, in the form factor of the atoms, Eq. (B.9). This means, that the expected value is not constant and will present a falloff with the frequency. This decay can be even more pronounced if thermal vibrations are considered. Note that, the B-factor correction quasi-flattening is justified in many cases considering that this expected value is constant. However, as our arguments have just shown, this idea should be rejected. This is an important issue because it still uncertain as to whether any reason in Physics exist to cast flat spectra. The answer is affirmative, but in the unrealistic scenario that considers atoms randomly distributed, occupying the whole space available, and without structure (dimensionless/point atoms and in absence of thermal vibrations). Under these conditions the form factor of atoms are just simple constants that depends on the kind of atom, and therefore Eq. (B.9) or Eq. (B.8) turns into

$$F(q) = \sum_{k=1}^{N} |f_k|^2 = constant. \tag{B.10}$$

The result of the flattening transformations is a sharpened protein that enhances in visualization and helps the model tracing. However, we insist in that this scenario is unreallistic. In addition, it also must be noted that in Eq. (B.9) the distribution of structure factor was obtained and its mean value was obtained without knowledge about the structure, taking only into account the nature of the atoms (scattering factors f_k) that constitute them. The expected value is, therefore, the best approximation for the structure factor in absence of knowledge about the macromolecular shape, and for that reason the B-factor correction becomes flat, or quasi-flat spectra enhances the macromolecular visualization. But the corrected structures do not represent a macromolecular object, because the structure factor must decay as Wilson statistics show. Hence, sharpening methods that incorporate knowledge *a priori* about the structure will be good alternative candidates to the current practice based on B-factor quasi flattening.

References

Bartesaghi, A., Merk, A., Banerjee, S., Matthies, D., Wu, X., Milne, J., Subramaniam, S., 2015. '2.2 Å resolution cryo-EM structure of β-galactosidase in complex with a cell-permeant inhibitor. Science 348, 1147–1151.

Berman, H., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T., Weissig, H., Shindyalov, I., Bourne, P., 2000. The protein data bank. Nucl. Acids Res. 28 (1), 235–242. URL:https://www.rcsb.org/.

Corless, R., Gonnet, G., Hare, D., Jeffrey, D., Knuth, D., 1996. On the lambert w function. Adv. Comput. Math 5, 329–359.

de la Rosa-Trevin, J., 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, J., Marabini, R., Sorzano, C., Carazo, J.,

2016. Scipion: a software framework toward integration, reproducibility, and validation in 3d electron microscopy. J. Struct. Biol. 195, 93–99.
Drenth, J., 2006. Principles of Protein X-Ray Crystallography. Springer.
Feigin, L., Svergun, D., 1987. Structure Analysis by Small-Angle X-Ray and Neutron Scattering. Springer.
Fernandez, J.J., Luque, D., Caston, J.R., Carrascosa, J.L., 2008. Sharpening high resolution information in single particle electron cryomicroscopy. J. Struct. Biol. 164, 170–175.
Frank, J., 2018. Ionic scattering factors of atoms that compose biological molecules. IUCrJ 5, 348–353.
Guinier, A., Fournet, G., 1955. Small-Angle Scattering of X-Rays. John Wiley and Sons Inc.
Henderson, R., 2015. Overview and future of single particle electron cryomicroscopy. Arch Biochem. Biophys. 581, 19–24.
Henderson, R., Sali, A., Baker, M., Carragher, B., Devkota, B., Downing, K., Egelman, E., Feng, Z., Frank, J., Grigorieff, N., Jiang, W., Ludtke, S., Medalia, O., Penczek, P., Rosenthal, P., Rossmann, M., Schmid, M., Schröder, G., Steven, A., Stokes, D., Westbrook, J., Wriggers, W., Yang, H., Young, J., Berman, H., Chiu, W., Kleywegt, G., Lawson, C., 2012. Outcome of the first electron microscopy validation task force meeting. Structure 20 (2), 205–214.
Heynmann, B., 2018. Map challenge assessment: fair comparison of single particle cryoem reconstructions. J. Struct. Biol. 204, 360–367.
Hryc, C.F., Chen, D.H., Afonine, P.V., Jakana, J., Wang, Z., Haase-Pettingell, C., Jiang, W., Adams, P.D., King, J.A., Schmid, M., Chiu, W., 2017. Accurate model annotation of a near-atomic resolution cryo-EM map. Proc. Natl. Acad. Sci. U.S.A. 114, 3103–3108.
Huber, E., de Bruin, G., Heinemeyer, W., Soriano, G.P., Overkleeft, H., Groll, M., 2015. Systematic analyses of substrate preferences of 20S proteasomes using peptidic epoxyketone inhibitors. J. Am. Chem. Soc. 137, 7835–7842.
Jakobi, A.J., Wilmanns, M., Sachse, C., 2017. Model-based local density sharpening of cryo-EM maps. ELife 6.
Lawson, C., Patwardhan, A., Baker, M., Hryc, C., Garcia, E., Hudson, B., Lagerstedt, I., Ludtke, S., Pintilie, G., Sala, R., Westbrook, J., Berman, H., Kleywegt, G., Chiu, W., 2016. Emdatabank unified data resource for 3dem. Nucl. Acids Res. 44 (D1), D396–D403. URL:http://www-ebi.emdatabank.org/.
Lee, C., Yashiro, S., Dotson, D., Uzdavinys, P., Iwata, S., Sansom, M., von Ballmoos, C., Beckstein, O., Drew, D., Cameron, A., 2014. Systematic analyses of substrate preferences of 20s proteasomes using peptidic epoxyketone inhibitors. J. Gen. Physiol 1444, 529–544.
Liao, M., Cao, E., Julius, D., Cheng, Y., 2013. Structure of the TRPV1 ion channel determined by electron cryo-microscopy. Nature 504, 107–112.
Morris, R., Blanc, E., Bricogne, G., 2004. On the interpretation and use of $<|e|^2>$ (d^*) profiles. Acta Crystallogr. Section D 227–240.
Peng, L., Gen, R., Dudarev, S., Whelan, M., 1996. Robust parameterization of elastic and absorptive electron atomic scattering factors. Acta Crystallogr. A 257–276.
Prince, E., 2006. International Tables for Crystallography Volume C: Mathematical, physical and chemical tables. Wiley.
Ramírez-Aportela, E., Vilas, J.L., Melero, R., Conesa, P., Martínez, M., Maluenda, D., Mota, J., Jiménez, A., Vargas, J., Marabini, R., Carazo, J.M., Sorzano, C.O.S., in press. Automatic local resolution-based sharpening of cryo-EM maps, Bioinformatics.
Rosenthal, R., Henderson, H., 2003. Determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. J. Mol. Biol. 333, 721–745.
Rupp, B., 2010. Biomolecular Crystallography. Garland Science.
Shmueli, U., Weiss, G., 1995. Introduction to Crystallographic Statistics. International Union of Crystallography.
Sorzano, C., Vargas, J., Oton, J., Abrishami, V., de la Rosa-Trevin, J., del Riego, S., Fernandez-Alderete, A., Martinez-Rey, C., Marabini, R., Carazo, J., 2015. Fast and accurate conversion of atomic models into electron density maps. AIMS Biophys. 2, 8–20.
Terwilliger, T., Sobolev, O., Anfonine, P., Adams, P.D., 2018. Automated map sharpening by maximization of detail and connectivity. Acta Crstallogr. Section D 74, 545–559.
van der Schaaf, A., van Hateren, J., 1996. Modelling the power spectra of natural images: Statistics and information. Vision Res. 36, 2759–2770.
Wilson, A., 1949. Determination of absolute from relative x-ray intensity data. Nature 150 152-152.
Yonekura, K., Matsuoka, R., Yamashita, Y., Yamane, T., Ikeguchi, M., Kidera, A., Maki-Yonekura, S., 2017. Advances in the field of single-particle cryo-electron microscopy over the last decade. Nat. Protocols 12, 209–212.