IP3R1 – Assessing Map Interpretability at Near Atomic Resolution

Steven J. Ludtke[2], Guizhen Fan[1], Matthew L. Baker[2], Zhao Wang[2], Mariah R. Baker[1], Stephen Murray[2], Pavel A. Synyagovsky[1], Wah Chiu[2], Irina I. Serysheva[1]

[1.] Department of Biochemistry and Molecular Biology, Structural Biology Imaging Center, University of Texas at Houston Medical School, Houston, TX, USA.
[2.] National Center for Macromolecular Imaging, Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX , USA.

Once resolutions beyond ~6 Å have been achieved in single particle reconstruction it becomes possible to begin interpreting maps in terms of the protein sequence. At the low resolution extreme, C-alpha tracing and homology modeling can be applied. As resolution increases towards 4 Å, beta strand separation becomes clear, and larger sidechain densities begin to appear. However, simply citing the resolution of a map, regardless of criteria, does not provide sufficient information to assess the appropriate level of interpretability. It is always possible to filter maps such that, for example, sidechain densities appear, even when such densities are likely due primarily to noise.

If the same data were refined independently in two different software packages or with two somewhat different starting maps, the final maps will never be identical near the resolution limit, since the 0.143 FSC threshold represents a signal to noise ratio of only 0.33. That is, as we approach the limiting resolution of our map, we know that an increasing fraction of the observed features are due to noise, and yet this process is gradual enough that establishing a specific hard limit on interpretation is not feasible.

IP3R1 is a ubiquitous ion channel located in the ER, responsible for the release of Ca^{2+} and involved in a broad array of cellular processes. We have solved the structure of IP3R1 to a gold-standard resolution of 4.7 Å, permitting us to unambiguously trace ~85% of the primary sequence through the map, and through a combination of de-novo modeling, homology modeling, and crystal structures of fragments, constructed an all-atom model. However, this raises the question of whether at this resolution it is appropriate to even consider an all-atom model, which could be easily over-interpreted.

To help assess the level of interpretability of our map and model, we performed a number of assessments and validations. First, we performed completely independent refinements of the particle data in both EMAN2.1 [1] and Relion [2]. The resulting structures were virtually identical, with the EMAN map resolving some peripheral density better than Relion, and Relion showing more sidechain detail in a few helices near the core of the structure. After performing real-space refinement of the all-atom model against the two maps using PHENIX, an RMSD of only 0.8 Å was achieved. While clearly the specific atomic coordinates were heavily influenced by molecular mechanics to achieve this value, it does indicate that no significant disagreements exist between the maps.

Applying Resmap [3], we see that the apparent level of detail is significantly higher in some regions of the map than in others (Fig 1). This 'resolution' varies from 3.6 Å in the transmembrane domain to 6.5 Å in parts of the cytoplasmic domain. This clearly indicates some level of variability is present in the underlying data. This could be due to actual structural flexibility and/or the presence of minor components of the other two highly homologous isoforms of IP3R in a portion of the particles. Clearly, though, the simple statement that we have "4.7 Å resolution" is insufficient as an assessment of the map.

The logical next step is to ask how well the atomistic model represents the map density. After converting the model to a density map we can compute a map vs. model FSC curve. We can also compute an FSC curve between the EMAN2.1 and RELION1.3 maps. As shown in Fig 2, if we interpret these curves with the 0.25 threshold mandated by use of all (rather than half) of the data in the maps, we see agreement to only ~5.9 Å resolution, on average. We can also compute a localized version of the FSC, and again observe that self consistency is much improved in similar areas to those observed by Resmap.

If we take the most conservative position, and considered the map to be only 5.9 Å resolution, it would be ridiculous to claim any sort of significant sidechain interpretation. However, this is an overly pessimistic view. As a simple example, the presence of apparent phenylalanine residues at exactly the location where they would be expected to act as a gate blocking the closed channel would seem to reaffirm that the registration of the primary sequence with our density map is correct, at least in that region. Similar observations can be made elsewhere in the map.

Clearly map interpretability is a complex issue, not easily encompassed by any single number, or even by any single procedure. We will consider these issues, and various strategies for quantifying them, in the context of IP3R1 and other recently solved structures.

References:

[1] G Tang *et al.*, J Struct Biol **157** (2007), p.38-46.
[2] SH Scheres, J Struct Biol **180** (2012), p.519-30.
[3] A Kucukelbir, FJ Sigworth and HD Tagare, Nat Methods **11** (2014), p.63-5.
[4] This work was supported by grants from the National Institutes of Health (R01GM072804, R21AR063255, S10OD016279, P41GM103832, R01GM079429, R01GM080139, and R21GM100229), Department of Defense (R038318-I), the American Heart Association (14RNT1980029), the Muscular Dystrophy Association (295138) and National Science Foundation (DBI-1356306).
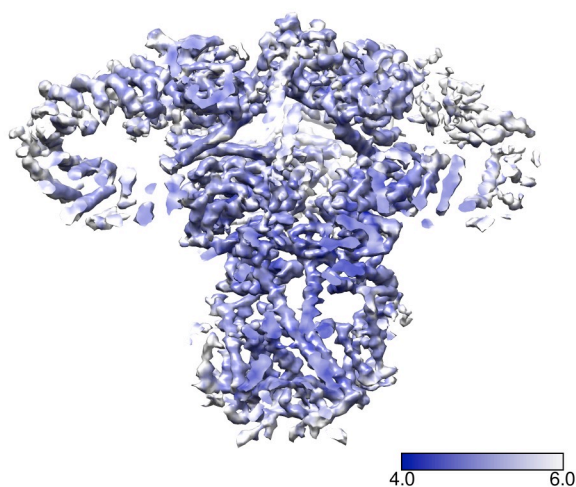
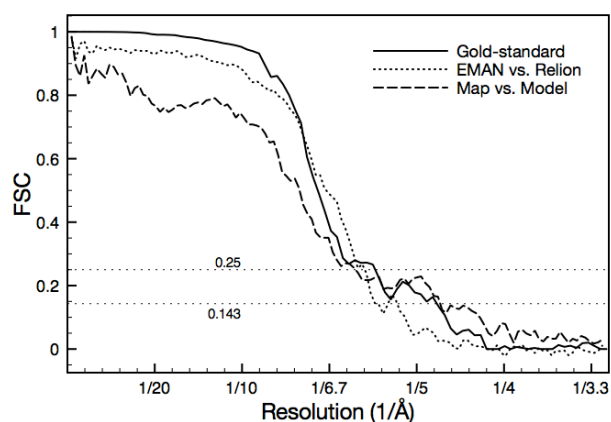**Figure 1.** Final reconstruction sliced in half, colored by Resmap value from 4-6 Å.

Figure 2. Map vs. model (after real-space refinement), EMAN vs. RELION and gold standard FSC