Electron density based validation

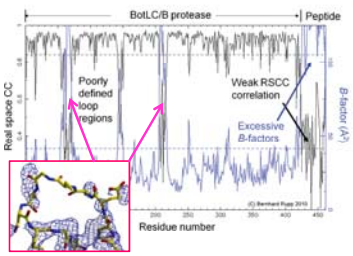


Figure 13-7 Real space correlation coefficient and *B*-factor plot. The superseded PDB entry 1B83 (downloadable from the book website) contains the model coordinates for the BoTLC/B protease–synaptobrevin-II complex[20] refined at 2.0 Å. Shown in black (upper curve) is the residue-by-residue real space correlation coefficient; in blue (lower curve) the *B*-factors are plotted for each residue. The left part of the figure corresponds to the protease, which, with the exception of three loop regions is well defined. The synaptobrevin-II ligand peptide at the right figure edge, however, shows a very worrisome crossover between abysmal real space correlation and excessive *B*-factors. The plot was created by the *ShakeAndWARP* web service.[21] As an exercise, you may convince yourself that the superseding PDB entry 3g94 (in the meantime also retracted) suffers from the same problems.

The real space correlation plot (RSCC) shows in one shot where the uncertain regions of the model are located. Low correlation = poor fit between model and density

RSCC determines seriousness of geometry outliers

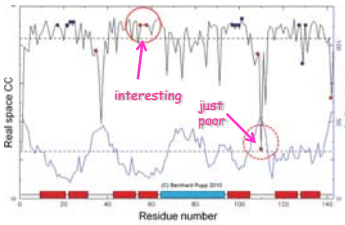


Figure 13-8 Real space correlation coefficients, plasticity, and location of stereochemistry deviations. The real space correlation plot quickly provides an overview of the plasticity of the protein. Note the correspondence of high *B*-factors and density deviations in flexible regions of the molecule. The location of stereochemistry outliers in relation to the regions of intrinsic plasticity or rigidity (as indicated by good RS fit and low *B*-factors) is worth examining. If residues with unusually high backbone torsion angle energy (red dots) are located in regions of high real space density correlation (solid red circle), they deserve attention because they may carry some relevance. If they are located in regions of poor density and high molecular flexibility (dashed red circle), they are probably just reflections of our inability to correctly model the structure because of poor electron density. Crystal contacts are exemplified by blue dots. Regions where a protein forms crystal contacts may not necessarily be in a native conformation, which is relevant if contacts are in the vicinity of binding sites or have other significance associated with them.

Geometry errors in low RSCC region are expected, but outliers in good electron density can be quite interesting and relevant

Long know – persistently neglected

Real-space solution to the problem of full disclosure

SIR — As discussed in Editorials in *Nature* and *Nature Structural Biology*[1,2], authors submitting research papers that describe molecular structures to Nature journals are required upon request to provide structural coordinates for reviewers to assess the quality of the work. But in the competitive field of structural biology, full disclosure of coordinates to anonymous peer review does raise serious concerns.

Nonetheless, the local nature of structure model quality — about which commonly provided global indicators such as R and R_{free} reveal no detail[3] — in fact requires reviewers to be able to access and assess the quality in relevant areas of the structure (in particular around sites with bound ligands), which is not possible without disclosure of coordinates and structure factors.

space correlation plots substantially lowers the risk of over-interpretation of poor local electron density, enhances information available for review, and provides authors with a strong means to demonstrate the quality of their structures without compromising their original data.

Bernhard Rupp
Department of Molecular Biology and Biochemistry, University of California, Irvine, California 92697, USA, and q.e.d. life science discoveries, Livermore, California 94551, USA

1. *Nature* **435**, 1138 (2005).
2. *Nature Struct. Biol.* **10**, 579 (2003).
3. Kleywegt, G. J. & Jones, T. A. *Structure* **10**, 465–472 (2002).
4. Branden, C. I. & Jones, T. A. *Nature* **343**, 687–689 (1990).

Nature* welcomes comments and feedback to this suggestion from researchers in the community, via authors@nature.com, or from readers at our author blog, Nautilus: http://blogs.nature.com/nautilus/2006/12/RSCC_plots.html — Editor, *Nature

Discrete ligand molecules require careful validation

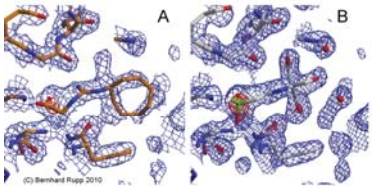
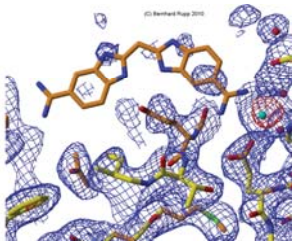


Figure 12-31 The sweetener that was bitter. The panels show 2.1 Å maps contoured at 1σ (blue) and 5σ (red). (A) A super-sweetener presumed to be bound to a taste receptor analog built into electron density map generated from CNS ML $2mF_o - DF_c$ coefficients. Poor density fit and a questionable vdW contact prompted critical re-examination of the model, and the fit and local chemistry proved to be perfect for TES buffer (2-(2-hydroxy-1,1-dihydroxymethyl-ethylamino)-ethanesulfonic acid), a component from the crystallization cocktail (B). The *ShakeAndWARP* map in (B) has slightly less noise and cleaner connectivity and clearly reveals the true nature of the ligand. PDB entry 1yrl.[139]

Binding sites are by nature designed to bind something. They will grab anything in the crystallization cocktail – and high concentration will further increase occupancy. Chemical plausibility and excellent electron density are a must for ligand studies!

Poor density and geometry of ligands

Figure 13-9 Absent ligand density. A bias minimized ML omit $2mF_o - DF_c$ electron density map contoured at 0.8σ is shown for the ligand (a potential inhibitor) in a *Clostridium botulinum* serotype B neurotoxin light chain protease. The mean ligand *B*-factor is ~120 Å² while the mean *B*-factor of the protease is ~36 Å². Despite the low electron density contouring, and as expected from the high *B*-factors, no appreciable density for the ligand is present. Instead, even the catalytic water normally present in the apo-structure next to the Zn^{2+} site is visible in the minimally biased omit density. The two ring systems of the ligand are also in an unusual, perfectly flat conformation (180.0°), indicating lack of refinement. PDB entry 1fgh[87] (withdrawn).



If only weak electron density supports the ligand conformation, then chemical plausibility and correct restraints become even more important!

Common problem: bad ligand restraint files

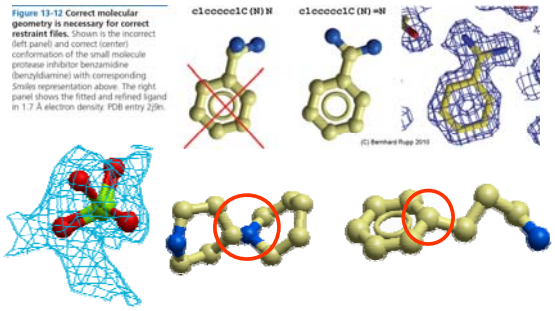



Figure 13-12 Correct molecular geometry is necessary for correct restraint files. Shown is the incorrect (left panel) and correct (center) conformation of the small molecule protease inhibitor benzamidine (benzyldiamine) with corresponding *Smiles* representation above. The right panel shows the fitted and refined ligand in 1.7 Å electron density. PDB entry 2jfm.