Review:

"My biggest concern with the paper is that I think the authors are pushing the resolution limit. They state that most crystals don't diffract well, but they collected some data to 3.0Å resolution for the X mutant. However, the over R-merge is extremely high at 13.8%. Plus, the R-merge in the highest resolution shell is over 100%. In fact the authors don't even give the exact value because they state that this parameter is meaningless, and state they cut the resolution by <I>/sigma<I> values. I disagree with this statement. R-merge over 100% is purely noise.

Additionally, I personally have never seen an R-merge over 100% with an <I>/sigma<I> of 2 or more in the outer resolution shell. I personally would be embarrassed to publish such bad data! What is the overall R-merge if you use a reasonable R-merge cutoff in the highest resolution shell like 50%?"

We respectfully disagree with the reviewer's opinion about the quality of our diffraction data and our choice of statistical indicators to define the diffraction limit for the crystals described in the manuscript.

The reviewer implies that:

(1) the crystals did not diffract to such a high resolution as presented;

(2) if they would diffract to that resolution, the R-merge statistics would be different;

(3) the choice of <I>/σ<I> as the statistic defining the resolution limit is questionable.

The reviewer also asks:

(4) "What is the overall R-merge if you use a reasonable R-merge cutoff in the highest resolution shell like 50\%?"

(1) <u>Crystals of x, y, z diffracted to resolutions of 3.0, 2.5, and 2.26 Å, respectively.</u>

We attached figures with representative diffraction patterns for each crystal. Reflections with intensities above 3σ of image noise are labeled with yellow (partial reflections) and green (full reflections) circles [Fig. 1].

A. B. C.

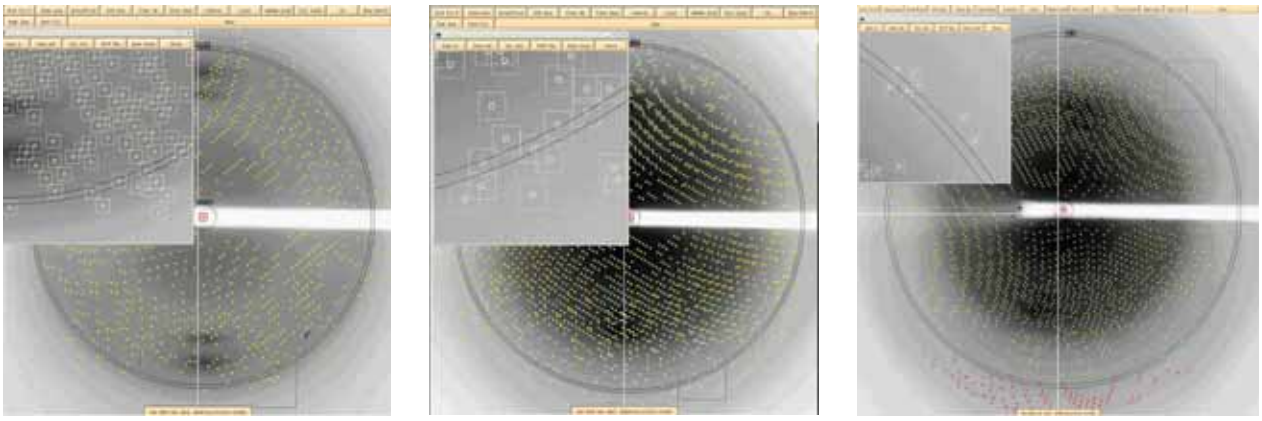


Figure 1. Diffraction images for crystals of V57D (1), V57D (2), and V57P mutants. The last resolution shell is represented by two black circles. Diffraction spots with intensities equal to or above 3σ of image noise are shown. The insert in each image represents a magnified fragment of the particular image, showing that in the last resolution shell the meaningful and informative observations are present.

The most important part of the validation of our decision to include those resolution shells comes from R- and R-free statistics calculated in resolution shells (Table 1, Figure 2). If the reviewer is right, the R-free values in resolution shells with high R-merge values should correspond to R-free values for "pure noise," i.e. R-free values around 59%. This is not the case. For both solved structures, R- and R-free values are consistent with

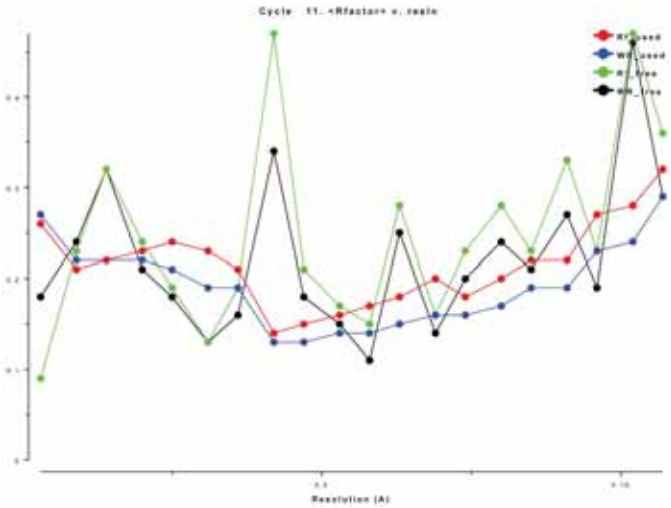
the selected resolution limits. Their values increase with increasing resolution, which is normal behavior[1], but never reach values close to "pure noise."

Table 1. R- and R_{free}- factors statistics for the refined structures of mutants X and Y. The occasional sudden increase of weighted R-free factor in some resolution shells results from a very small number of reflections used for R-free calculations in the resolution shells (we listed them as well). Nonetheless, the R-free for X is 29% in the last resolution shell and for Y it is 24%, clearly indicating that the data in those resolution shells are informative.

resolution range start	No. of reflections used for R-free calculations	weighted R-factor	weighted R-free
18.26	2	0.27	0.18
10.54	7	0.22	0.24
8.45	5	0.22	0.32
7.07	5	0.22	0.21
6.32	4	0.21	0.18
5.68	11	0.19	0.13
5.27	10	0.19	0.16
4.88	14	0.13	0.34
4.61	13	0.13	0.18
4.34	10	0.14	0.15
4.15	11	0.14	0.11
3.98	8	0.15	0.25
3.81	13	0.16	0.14
3.68	16	0.16	0.20
3.54	12	0.17	0.24
3.43	18	0.19	0.21
3.31	16	0.19	0.27
3.23	13	0.23	0.19
3.13	17	0.24	0.46
3.06	22	0.29	0.29

resolution range start	No. of reflections used for R-free calculations	weighted R-factor	weighted R-free
12.91	9	0.27	0.42
7.91	21	0.26	0.34
6.20	13	0.28	0.36
5.27	19	0.24	0.21
4.71	24	0.20	0.21
4.26	17	0.18	0.24
3.95	19	0.18	0.28
3.68	20	0.19	0.19
3.45	29	0.19	0.20
3.28	24	0.20	0.18
3.12	29	0.21	0.22
2.97	34	0.22	0.32
2.85	32	0.20	0.23
2.75	38	0.23	0.24
2.65	31	0.23	0.25
2.56	29	0.24	0.24
2.49	30	0.24	0.27
2.42	33	0.24	0.31
2.35	43	0.25	0.30
2.29	43	0.26	0.24

A.



B.

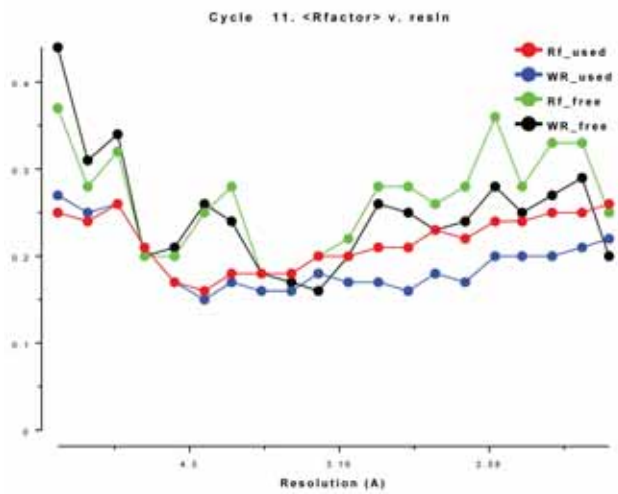


Figure 2. R- and R_{free}-factors behavior across the resolution shells for A) X and B) Y. Weighted R-free is shown in green. In both cases, the R-free values for the last resolution shell are below 40%, indicating that those shells contain informative reflections.

[1] Weaker observations have a relatively higher level of experimental errors, and structural models tend to have higher levels of uncertainty associated with high resolution details. Both those factors contribute to the increase of R-factor and R-free with resolution.

(2) <u>There are several reasons why R-merge values are high in our case.</u>

R-merge is an unweighted statistic that describes how well the unmerged, equivalent observations agree with one another. Hence, the significance of R-merge values depends strongly on multiplicity of observations. Our crystals belong to high symmetry space groups (cubic and tetragonal) and we collected data sets with high multiplicity of observation (MOI = ~20 for cubic space group, MOI = ~10 for tetragonal space group). For this reason, our R-merge values are higher than they would be for instance for the space group $P2_12_12_1$, the most popular space group in macromolecular crystallography. However, high multiplicity is only one of the contributors to R-merge values. R-merge is also affected by systematic errors resulting from instrumental factors, from uncorrected absorption, and most importantly, from decay induced by X-rays during experiments. Indicators (Diederichs and Karplus 1997; Weiss 2001) which are corrected for by multiplicity of observations do not solve the problem because they do not account for all other contributing factors.
The R-merge values in low-resolution shells for all crystals described in the manuscript are in the same range of 4%, which are standard values for experiments performed at a third-generation synchrotron source with fast detectors and relatively short exposures. Those values tell us that we did not have unusually high levels of errors arising from instrumental problems and uncorrected absorption.

The upper limit of the recommended dose is around 10 to 20 MGy, at which the overall intensity of the diffraction image decays 50%. In our experiments scaling B-factors (Borek, Ginell et al. 2007; Borek, Cymborowski et al. 2010) increased to 11.3 $Å^2$, 7.5 $Å^2$, 5.9 $Å^2$ for =X, =Y, =Z respectively. These numbers are equivalent to 1.1 MGy, 0.8 MGy, and 0.6 MGy dose, which indicates that we selected exposures well below the maximal advised limit (Owen, Rudino-Pinera et al. 2006). However, those exposures were high enough to induce significant decay of diffraction intensities (Kmetko, Husseini et al. 2006). The decay of diffracted intensities induced by X-ray radiation is the biggest contributor to high R-merge values at high resolution in our case. Correcting for intensity decay during the experiment requires scaling up the decayed observations. Due to the high multiplicity of observations, such decayed reflections contribute little to the output structure factors as they are down-weighted by their scaled-up uncertainty. Thus, even though R-merge is increased, it does not have a negative impact on the resulting map, but as discussed later, including those observations has significant, positive impact on the downstream calculations.
Eliminating more exposed parts of the dataset has a very big impact on R-merge, but results in a simultaneous decrease of data quality due to the reduction of multiplicity of observations. Using this approach, one can easily reduce R-merge to the values stated by the reviewer, with reduction of data quality that would still allow one to solve the structure.

(3) **<u>The <I>/σ<I> as an optimal choice was discussed before in many publications (Dauter 1999; Evans 1999).</u>**

In this case, <I>/σ<I> choice is validated by goodness of fit (normalized χ^2) being closed to one for the selected error model. The error model used was typical for such experiment (~3 to 4 % of systematic errors). There is no indication that our choice of error model was unjustified.

The reviewer pointed out that he/she never observed <I>/σ<I>=2 with corresponding R-merge above 50% before. There could be many reasons for this: using the minimum required redundancy, *e.g.* 4 in orthorhombic symmetry; using crystals that were larger than our, which therefore did not require synchrotron exposure to extract high resolution information, so they did not undergo decay; their choice of software – some old versions of software artificially reduce R-merge by eliminating negative observations before averaging; and

many others. We routinely observe <I>/σ<I>=2 with corresponding R-merge above 100%, and have even been criticized for this criterion being too stringent for the resolution limit choice.

Weak reflections contribute little information to unsharpened maps, so they do not have a significant impact on map calculations. However, they have a big impact on refinement (Schwarzenbach, Abrahams et al. 1989), and experts in the area of macromolecule refinement (Garib Murshudov) suggest even using <I>/σ<I>=1 for reflections properly weighted by statistical errors. One of the examples where the <I>/σ<I>=~1 criterion was used is the structure of the fatty acid synthetase (Maier, Leibundgut et al. 2008), and R-merge values in the last resolution shells for those data were above 100%.

(4) "What is the overall R-merge if you use a reasonable R-merge cutoff in the highest resolution shell like 50\%?"

The overall R-merge would naturally be much lower after excluding shells contributing high values of R-merge. However, the reviewer's request would require cutting the data at <I>/σ<I> between 7 to 9, which means removing very informative observations. It would not only artificially decrease resolution, but also negatively affect all downstream procedures that are sensitive to exclusion of observations (Schwarzenbach, Abrahams et al. 1989).

Overall, R-merge is not an appropriate statistic to judge the data quality. In the low resolution shells, R-merge provides valuable information about potential problems with instruments or with uncorrected absorption. At high resolution, the uncertainty of the measured intensities is dominated by intensity decay induced by X-ray exposure; thus R-merge does not provide any information about the sources of problems other than decay. The lack of agreement between symmetrically equivalent reflections at high resolution is not a problem by itself as long as intensities are properly weighted by their uncertainties. The validation of how well the resolution limit was chosen is agreement between the data and the final model, which is presented in Table 1 and 2, and as discussed above in our case, it clearly indicates that we did not overestimate the resolution limit.

Borek, D., M. Cymborowski, et al. (2010). "Diffraction data analysis in the presence of radiation damage." Acta Crystallogr D Biol Crystallogr **66**(Pt 4): 426-436.

Borek, D., S. L. Ginell, et al. (2007). "The many faces of radiation-induced changes." J Synchrotron Radiat **14**(Pt 1): 24-33.

Dauter, Z. (1999). "Data-collection strategies." Acta Crystallographica Section D-Biological Crystallography **55**: 1703-1717.

Diederichs, K. and P. A. Karplus (1997). "Improved R-factors for diffraction data analysis in macromolecular crystallography." Nature Structural Biology **4**(4): 269-275.

Evans, P. R. (1999). "Some notes on choices in data collection." Acta Crystallographica Section D-Biological Crystallography **55**: 1771-1772.

Kmetko, J., N. S. Husseini, et al. (2006). "Quantifying X-ray radiation damage in protein crystals at cryogenic temperatures." Acta Cryst. D **62**(9): 1030-1038.

Maier, T., M. Leibundgut, et al. (2008). "The crystal structure of a mammalian fatty acid synthase." Science **321**(5894): 1315-1322.

Owen, R. L., E. Rudino-Pinera, et al. (2006). "Experimental determination of the radiation dose limit for cryocooled protein crystals." Proceedings of the National Academy of Sciences of the United States of America **103**(13): 4912-4917.

Schwarzenbach, D., S. C. Abrahams, et al. (1989). "Statistical Descriptors in Crystallography Report of the International Union of Crystallography Subcommittee on Statistical Descriptors." Acta Crystallographica Section A **45**: 63-75.

Weiss, M. S. (2001). "Global indicators of X-ray data quality." J of Applied Crystallography **34**: 130-135.