FOUNDATIONS
ADVANCES

Volume 75 (2019)

Supporting information for article:

Multipole Electron Densities and Atomic Displacement Parameters in Urea from Accurate Powder X-Ray Diffraction

Bjarke Svane, Kasper Tolborg, Lasse Rabøl Jørgensen, Martin Roelsgaard, Mads Ry Vogel Jørgensen and Bo Brummerstedt Iversen

S1. Image Plate Integration

Incorrect integration that does not follow the diffraction rings will result in asymmetric peaks in the resulting one-dimensional diffractogram. As the curvature of the rings depend on diffraction angle, the asymmetry will likewise vary with 2θ. Proper intensity partitioning of overlapping peaks thus requires that the integration path follows the rings. Assuming that the detector is perfectly circular with a perfectly centered sample, it can be derived analogously to (Straasø *et al.*, 2013) that the coordinates of a point P on the surface of a cylinder is:

$$P = (x, y, z) = \frac{\left(Rsin2\theta sin\phi, \sqrt{R^2 - (Rsin2\theta sin\phi)^2 - (Rsin2\theta cos\phi)^2}, Rsin2\theta cos\phi\right)}{\sqrt{1 - sin^2 2\theta sin^2\phi}}$$

Where R is the sample to detector distance, 2θ is twice the Bragg angle and ϕ is the azimuthal angle. An integration routine based on this analytical expression has so far been used for data collected using AVID. Detailed examination of the integration path reveals that some deviation from the true Debye-Scherrer footprint is present at the edge of the IP (see Figure S1). The degree of deviation is not constant between measurements or even across a single IP in a single measurement.

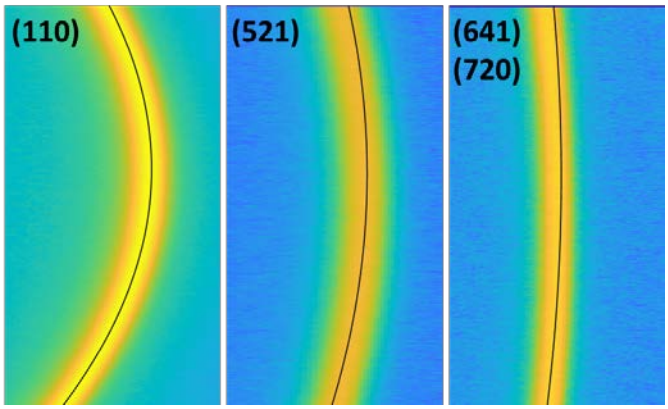

Figure S1 Diffraction rings of LaB6 reference overlaid with Debye-Scherrer cone integration path (black line). (hkl) in upper left corner.

One possibility is that the arc description as a perfect cylinder is not appropriate. Independent recalibration of the AVID setup revealed that after extended use of AVID the arc deviated from a perfect circle. The topmost part has sunk slightly, with a kink around $2\theta = 30°$. The deformation is small enough that it is reasonable to approximate the arc as a perfect circle, but with the center in $(y, z) = (-0.01mm, -0.21mm)$, where $(0,0)$ is the ideal sample position. The deviation in peak position resulting from this deformation is 8-10 pixels, depending on angle. The angular dependence

is smooth, however, so this deformation does not explain the integration path deviation, which is much larger on a small angular interval.

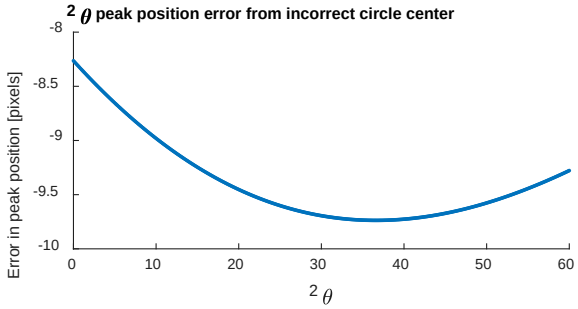

Figure S2 Calculation of error in peak position prior to AVID recalibration. Measurement done in February 2018 at JJ X-Ray, Lyngby.

The integration path error arises from the fact that no fitting of detector alignment is implemented. The conventional method of solving this problem would be to work out an analytical expression, fit the pitch, roll, and yaw angles along with movement of IPs on the arc perpendicular to the beam both along and across the arc. As the IPs are not placed identically each measurement, this position determination must be repeated for every measurement. While this is possible on the IPs with distinct peaks, it is very unstable for the IPs higher in 2θ and it does not provide the required accuracy. The actual IP positions during each measurement must be determined, prohibiting the use of a standard, as the IPs are not placed identically in each measurement.

To overcome the issue of IP position fitting on IPs with only low intensity peaks, a new routine for the entire integration routine has been implemented. It is based on simulating the diffraction pattern of the sample on the entire arc using the real, non-idealized arc geometry. The IP position and rotation are determined based on minimization of mean square error between the observed and simulated diffractogram. Initial coarse alignment is based on down-sampled data (¼ resolution) and only fits IP position. The algorithm then fits both position and rotation on heavily down-sampled data (1/16 resolution) and proceeds pyramidically on ¼ resolution data and finally on the full resolution data. The pyramidical registration process speeds up the image alignment significantly. The transformation registration is done automatically and takes ~10 minutes per IP on a standard laptop computer. The algorithm is robust enough to determine IP position and rotation reproducibly and accurately. The last statement is based on visual inspection of the integration path compared to the measured 2D diffraction rings. It is a requirement for the use of this method that the space group and unit cell lengths are known. However, it is unlikely that an ED determination will be attempted on a crystal without a known structure, making this a minor issue. Furthermore, unit cell parameters can already

be determined to good accuracy even with some asymmetry present in the peaks from the analytical implementation of the integration routine.

Having determined the IP position and rotation based on the simulated data, it is straightforward to determine the corresponding 2θ-angles for each pixel individually. A vectorized integration routine with correct intensity division of pixels between 2θ-bins has been implemented. This division is necessary, as the exact diffraction angle of a pixel most often lies between two 2θ-bins. The new integration routine reduces the asymmetry sufficiently that no asymmetry term is necessary in the Rietveld refinement.

The estimated uncertainty on measured data is important to the quality of the final model. Significant effort has been spent in making sure that the error estimates from the measurement are sound and propagate correctly into the extracted structure factor lists. The uncertainty in a diffractogram point has traditionally been based on Poisson statistics of the photon counts. This is inaccurate, as a significant contribution stems from the limited number of diffraction crystallites (Alexander & Klug, 1950). The structure factor uncertainties in the generated structure factor lists are approximated as the estimated standard deviations. When extracting structure factors from Jana, the estimated standard deviations are either determined based on propagation of diffractogram uncertainties or from the pattern fit directly (Petrícek *et al.*, 2014). The pattern fit is dependent on the uncertainties in the powder pattern, making the uncertainty estimate in the diffractograms important.

Here standard deviations are determined from sample statistics analogously to Wahlberg et al. (2016). This is based on the assumption that each pixel is an independent measurement. The standard deviation can then be calculated for each diffractogram point as:

$$\sigma(\langle I_{2\theta}\rangle) = \sqrt{\frac{\sum_{pixel=1}^{n}\left(I_{pixel} - \langle I_{2\theta}\rangle\right)^2}{(n-1)n}}$$

Here I_{pixel} is the intensity in each individual pixel contributing to a given $\langle I_{2\theta}\rangle$, which is the average intensity in a given 2θ bin. $n$is the total number of pixels contributing to the given average. The routine has been vectorized and accounts correctly for sharing of pixel intensity between adjacent 2θ bins. A sample statistics approach will by definition contain all contributions to the total uncertainty, provided the assumption of independence is sound. While this approximation is good for Hybrid Pixel Array Detectors, the point spread function of CCD and image plate systems means that those intensities will be underestimated. However, this method of assigning uncertainties will be appropriate for all powder diffraction data collected on 2D-detectors, with the caveat of uncorrected point spread functions leading to underestimated uncertainties.

To summarize, we have reduced the observed peak asymmetry by fitting the IP roll angle and correcting for the deviation from a perfect cylinder of the arc. The implemented integration routine gives less asymmetric and narrower peaks with a lower standard deviation than the previous implementations. The efforts in optimizing the integration and extraction accuracy have been necessary to obtain data suitable for PXRD ED modelling. Lacking these corrections severely limits the detail to which the ED can later be modelled.

S2. Rietveld refinement parameters.

A pseudo-Voigt profile description with GU, GX, GW, LY-coefficients, sample strain in the [400], [202] and [220]-directions (Stephens, 1999) and a corresponding Lorentzian mixing parameter (zeta) was used to extract the observed structure factors from the powder data. The integration routine obviated the need for a shift. The background was described using linear interpolation between 35 points and inspected manually to ensure that it did not unduly describe data features. C, O and N coordinates and anisotropic ADPs were refined within symmetry constraints. H atom positions were fixed to the neutron reference (Swaminathan *et al.*, 1984), with the ADPs of H being scaled to neutron using a 2-parameter linear correction factor calculated with the program UijXN based on C, O and N vibrational parameters (Blessing, 1995). The correction factor had a value of 0.835 in the plane of the molecule (*a*,*b*) and 0.956 out of the plane (*c*), corresponding very well to what was expected from linear interpolation between 123 K and 60 K neutron data. An anomalous dispersion correction was implemented based on tabulated values from the program WinGX. Data and model are plotted in Figure S3 for spherical form factors. The quality is similar for MM and MM-HB form factor decriptions.

The profile parameters used in structure factor extraction are given in Table S1. There are only minor differences in the profile description between the different scattering factor descriptions. This gives confidence that the profile description does not absorb the errors from inaccurate form factor description.

Table S1 Rietveld pattern profile fit parameters on AVID long λ (P08) data. ADPs are given as 10^{-4}Å^2.

	IAM	**MM**	**MM-HB**
R/wR(obs)	5.68/4.12	4.84/3.34	4.89/3.43
$R_P/w\,R_p$	3.23/2.54	3.07/2.41	3.06/2.41
GOF	12.10	11.47	11.50
Scale	0.6634(4)	0.6333(4)	0.6345(4)
a [Å]	5.578071(9)	5.578088(9)	5.578094(9)
c [Å]	4.687410(6)	4.687418(6)	4.687413(6)
GU	-6.7(1)	-6.3(2)	-6.3(1)
GW	0.28(5)	0.269(2)	0.267(2)
LX	-0.153(3)	-0.100(3)	-0.096(3)
LY	5.02(5)	4.30(5)	4.25(5)
Zeta	0.448(2)	0.454(2)	0.454(2)
S400	1.657(6)	1.672(6)	1.676(6)
S220	-0.5454(9)	-0.5430(8)	-0.5435(8)
S202	0.018(1)	0.019(1)	0.019(1)

***z*(C)**	0.3286(1)	0.3292(1)	0.3290(1)
***z*(O)**	0.59690(7)	0.59580(6)	0.59547(6)
***x*(N)**	0.14357(5)	0.14390(5)	0.14366(5)
***z*(N)**	0.17823(7)	0.17787(6)	0.17857(6)
U11(C)	128(2)	118(2)	114(2)
U33(C)	93(2)	45(2)	50(2)
U12(C)	13(3)	-8(3)	-6(3)
U11(O)	167(1)	131(1)	133(1)
U33(O)	84(2)	72(2)	72(2)
U12(O)	27(2)	17(2)	16(2)
U11(N)	253(1)	240(1)	237(1)
U33(N)	108(2)	116(2)	114(2)
U12(N)	-125(2)	-124(2)	-122(2)
U13(N)	9(1)	5(1)	4(1)

IAM model uses a spherical form factor. MM and MM-HB corresponds to the simple and more complex multipolar models, respectively, as detailed in the main text.

The fit quality descriptors show that while an IAM description gives a noticeably worse fit in the Rietveld refinement, the difference in fit quality between the simple and complex aspherical scattering factor model is small. Consequently, it is expected that the difference in extracted structure factors is also small between the later two. Multiple methods of determining the optimal peak profile has been attempted, e.g. Le Bail fitting (A. Le Bail , H. Duroy, 1988), multiple asymmetry descriptors, split pseudo-Voigt profile, and fitting the profile to low order non-overlapping reflections only. Whole-pattern Rietveld modelling proved the best model. This is likely because the number of peaks with no significant peak overlap is too low to accurately determine the included profile parameters. In particular, strain broadening determination relies on having many peaks present in the data to adequately resolve strain in different directions.

ADPs and fractional coordinates of non-hydrogen atoms are also fitted in the Rietveld refinement. The values are close to the results of both SCXRD (Birkedal *et al*., 2004) and neutron (Swaminathan et al., 1984) measurements. In general, the IAM model gives larger ADPs than the other two models, showing that ADPs can absorb the errors of a poor ED description. Fractional coordinates show no clear trends between models. The agreement of atomic positions with neutron data is comparable to the agreement between SCXRD and neutron data.

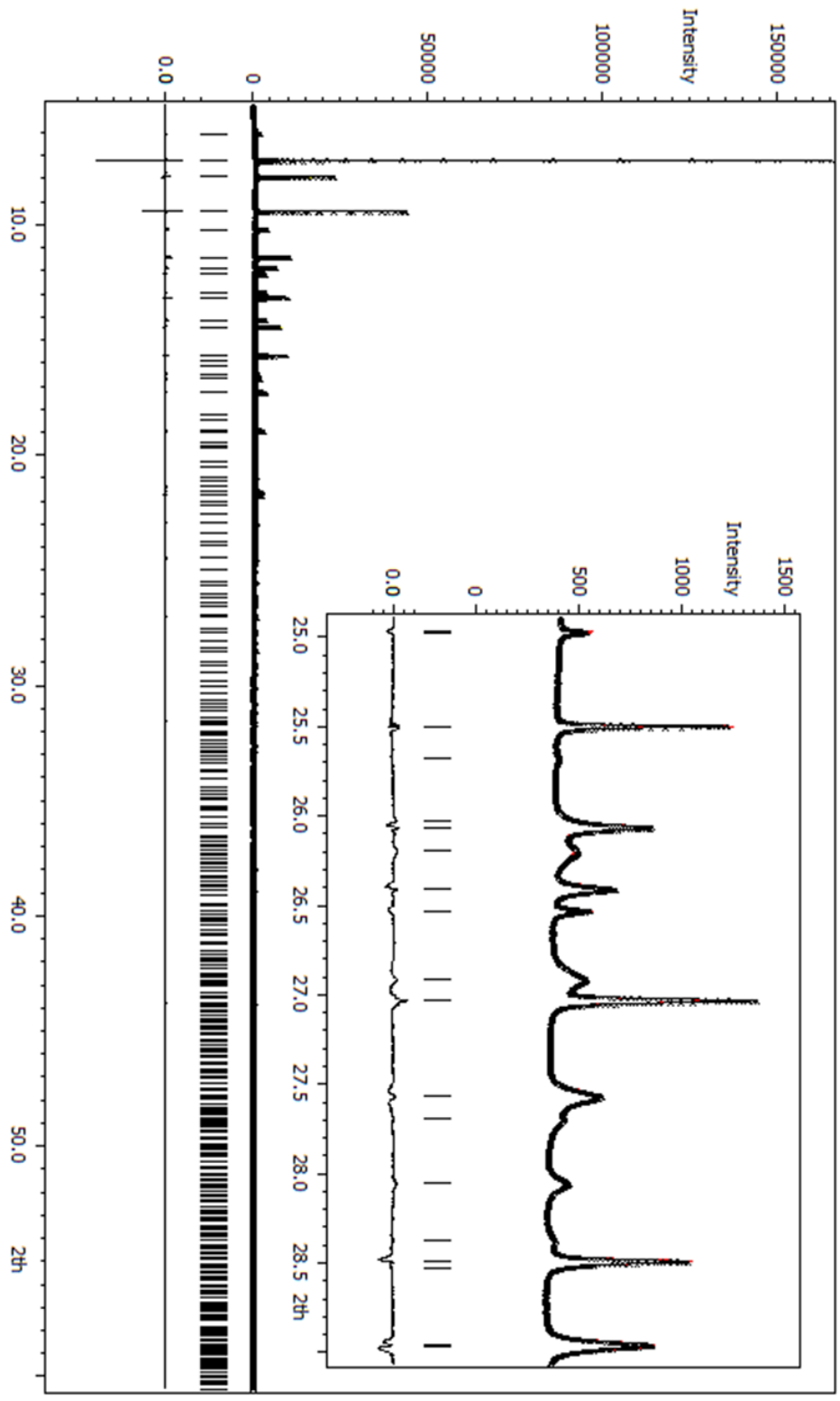
Intensity
150000
100000
50000
0
0.0
10.0
20.0
30.0
40.0
50.0
2th
Intensity
1500
1000
500
0
0.0
25.0
25.5
26.0
26.5
27.0
27.5
28.0
28.5
2th

Figure S3 PXRD data and Rietveld fit using the spherical form factor model. Peak positions are marked with vertical bars. Black line is observed data, red line is calculated. The difference line is shown below the plot. The quality of fits are very similar for the other form factor models.

S3. Atomic Displacement Parameters for multipolar refinement of AVID data

Table S2 Refined parameters for multipolar models against different structure factor lists. First line is Avid MM, second line is AVID IAM-Ex, third line is AVID short lambda, 4th line is approximate 100 K neutron data based on linear interpolation between 60 K and 123 K neutron data.

	$\mathbf{U^{11}}\,[10^{-4}\text{Å}^2]$	$\mathbf{U^{33}}\,[10^{-4}\text{Å}^2]$	$\mathbf{U^{12}}\,[10^{-4}\text{Å}^2]$	$\mathbf{U^{13}}\,[10^{-4}\text{Å}^2]$
C	89(4)	30(6)	-12(6)	-
	88(5)	56(6)	0(6)	-
	78(2)	39(3)	-44(2)	-
	125(5)	58(3)	0(4)	-
O	113(4)	53(6)	25(7)	-
	136(5)	55(7)	29(9)	-
	133(2)	80(4)	13(3)	-
	169(6)	54(4)	11(5)	-
N	226(5)	81(5)	-119(6)	2(3)
	227(5)	66(5)	-129(6)	6(3)
	212(2)	70(3)	-112(3)	-2(1)
	243(4)	86(2)	-125(2)	2(3)

S4. Hansen-Coppens model population parameters for multipolar refinement of AVID and reference data

Table S3 Multipole population parameters refined against different structure factor lists. Naming convention as in Table 2. Local coordinate system definitions are given in the text below.

		AVID MM	**AVID IAM-Ex**	**AVID short λ**	**SC reference**
C	P_{10}	0.15(6)	0.13(6)	0.17(4)	0.042(8)
	P_{20}	0.09(5)	-0.03(6)	0.22(3)	0.168(7)
	P_{22+}	-0.23(6)	-0.15(6)	-0.29(3)	-0.175(9)
	P_{30}	0.37(7)	0.32(7)	0.61(5)	0.312(10)
	P_{32+}	-0.30(10)	-0.22(11)	0.13(7)	0.180(18)
O	P_{10}	-0.08(5)	0.02(5)	-0.10(3)	-0.076(5)
	P_{20}	-0.08(5)	-0.09(5)	-0.08(3)	-0.055(5)
	P_{22+}	0.04(5)	0.04(6)	-0.01(3)	-0.056(7)
	P_{30}	0.05(4)	0.06(4)	0.02(2)	0.036(5)
	P_{32+}	0.31(13)	0.34(13)	0.33(7)	0.050(21)
N	P_{11+}	0.17(5)	0.15(5)	0.10(3)	0.019(6)
	P_{11-}	-0.04(5)	-0.09(5)	-0.08(3)	0.005(6)
	P_{20}	0.00(4)	-0.9(4)	-0.08(3)	-0.040(7)
	P_{22+}	-0.06(4)	-0.2(4)	-0.02(2)	-0.025(5)
	P_{22-}	0.07(3)	0.08(3)	-0.01(2)	-0.033(5)
	P_{31+}	-0.13(6)	-0.16(6)	-0.11(4)	0.079(9)
	P_{31-}	-0.17(8)	-0.15(8)	-0.18(4)	0.005(13)
	P_{33+}	0.40(8)	0.27(8)	0.41(4)	0.220(12)
	P_{33-}	-0.07(3)	-0.06(3)	0.03(2)	-0.012(4)
H1	P_{10}	0.64(9)	0.61(9)	0.50(5)	0.208(11)
	P_{20}	0.40(10)	0.37(9)	0.44(6)	0.022(14)
	P_{22+}	0.20(5)	0.10(5)	-0.04(4)	0.071(8)
	P_{22-}	0.04(6)	0.01(4)	0.02(5)	0.042(16)
H2	P_{10}	0.28(7)	0.22(7)	0.32(5)	0.209(9)
	P_{20}	0.11(5)	0.12(5)	0.22(3)	0.132(7)
	P_{22+}	-0.20(7)	-0.08(7)	-0.03(5)	0.092(11)
	P_{22-}	0.18(8)	0.11(7)	0.01(7)	-0.019(20)

Local coordinate systems are as follows: For carbon, the z-axis points towards the oxygen atom and the y-axis is in the plane of the molecule. The site symmetry is mm2, with the twofold axis parallel to the z-axis and a mirror plane perpendicular to the y- and x-axis.

For oxygen, the z-axis points toward the carbon atom and the y-axis is in the plane of the molecule. The site symmetry is mm2, with the twofold axis parallel to the z-axis and a mirror plane perpendicular to the y- and x-axis.

For nitrogen, the x-axis points towards the carbon atom, with the y-axis perpendicular and in the molecular plane. The site symmetry is m, with the mirror plane perpendicular to the z-axis.

For H1 and H2, the z-axis points toward the nitrogen atom, with the y-axis perpendicular and in the molecular plane. The site symmetry is 2, with the twofold axis parallel to the z-axis.

S5. Residual Density analysis

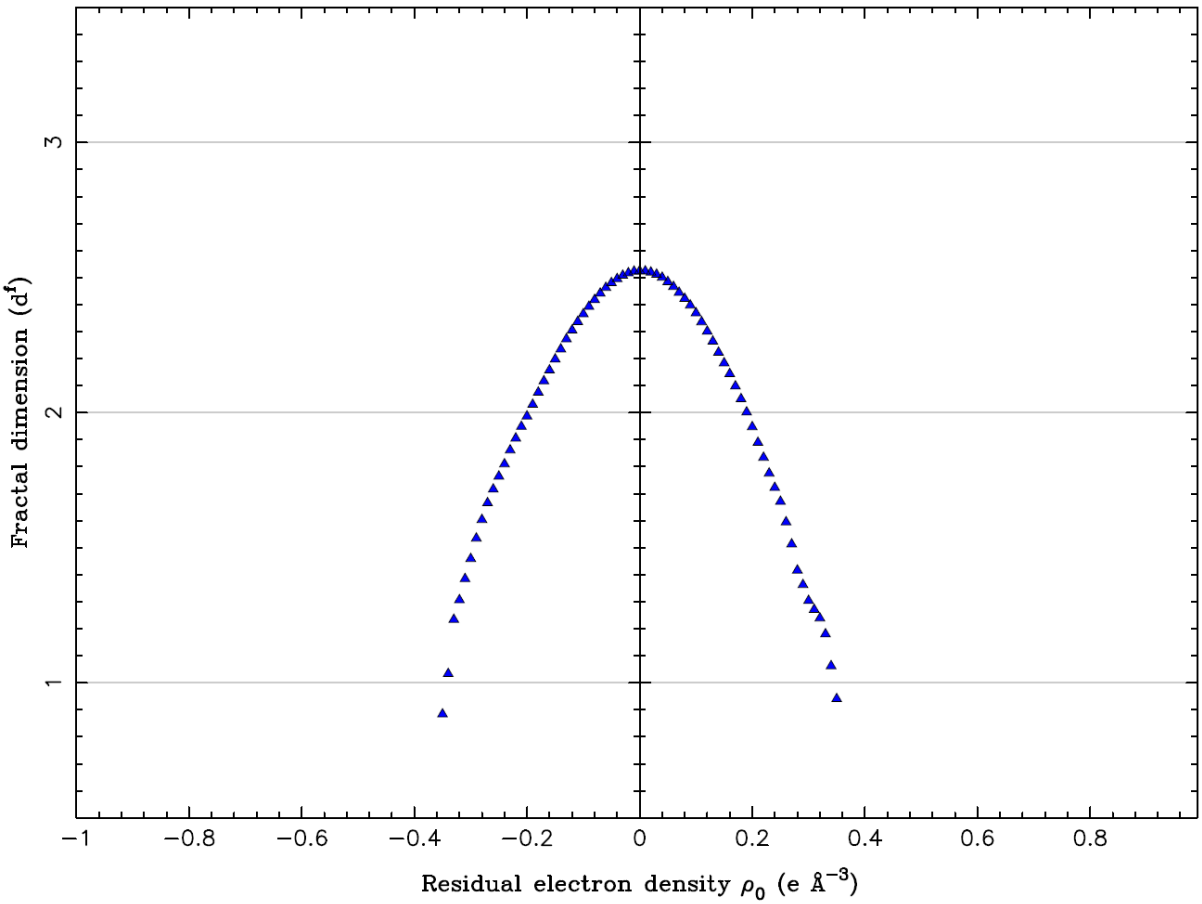

Figure S4 Fractal dimension plot of AVID MM.

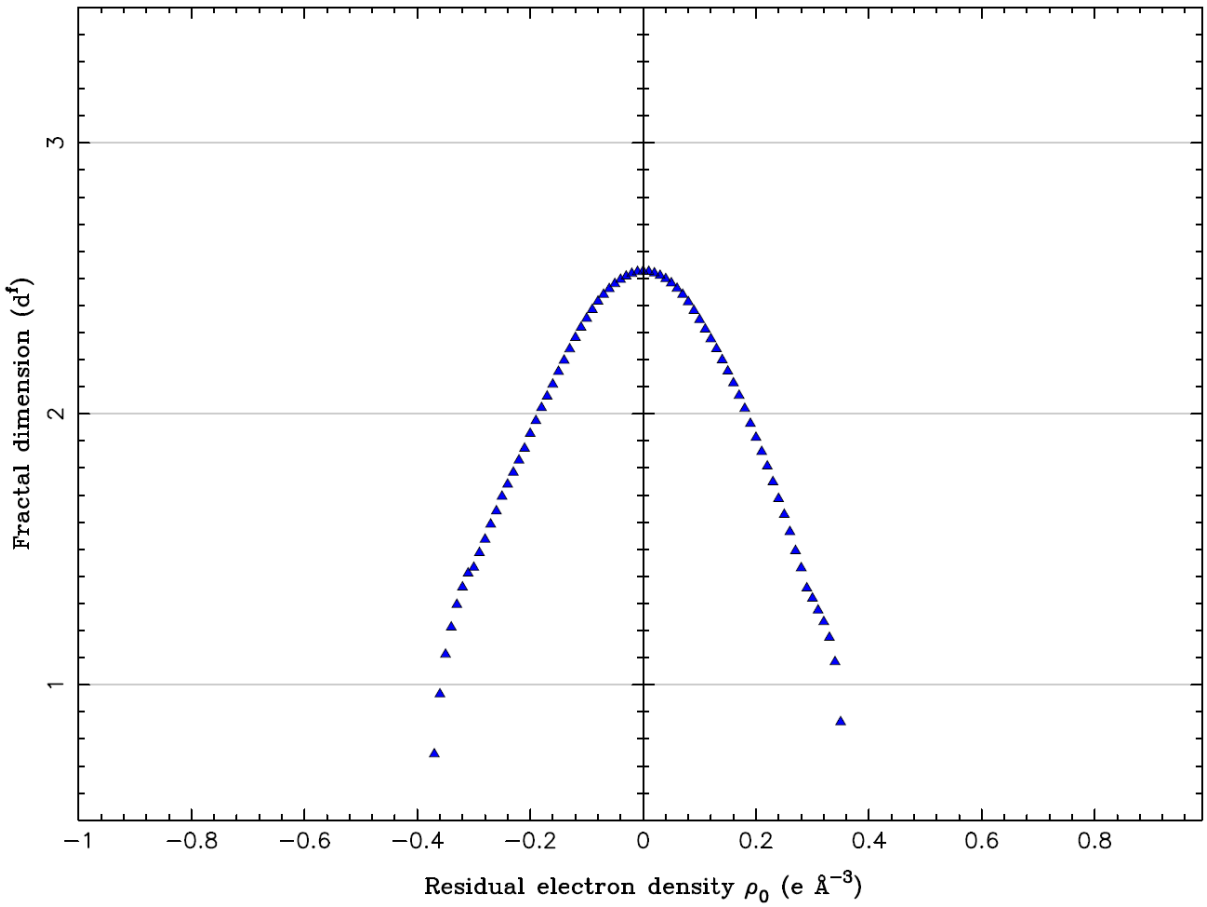

Figure S5 Fractal dimension plot of AVID IAM-Ex