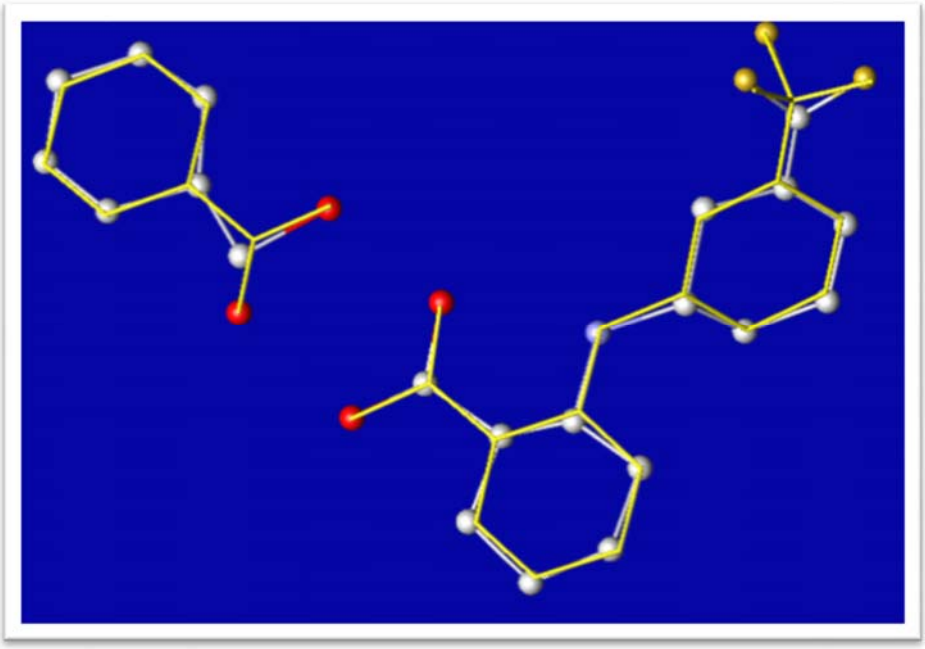
Rigaku
Leading With Innovation



Appropriate Restraint Settings for Crystal Structure Refinement

This is the digest version of the oral presentation given at PPXRD-15, in Hyderabad, India on 20th, August, 2017

By Akito Sasaki and Hisashi Konaka
XRD Application & Software Development Dept.,
X-ray Instrument Division, Rigaku Corporation

Normally, the first choice for crystal structure analysis is the single crystal method. But, sometimes good single crystals cannot be prepared in terms of size and quality, or single crystal state cannot be maintained due to phase transitions, etc. The single crystal method cannot be applied in those cases, and you have to try structure analysis from powder diffraction data to obtain crystal structure information.

There are several advantages to crystal structure analysis from powder, but there are several difficulties at the same time. It is difficult to decompose diffraction peaks in powder diffraction data because there are more peak overlaps in higher diffraction angles. This is often the case even at medium angles, due to low crystal quality, small crystallite sizes, low symmetry of the crystal system, etc., especially for organic compounds.

This makes it difficult to obtain precise diffraction peak positions and precise diffraction peak intensities.

In order to determine precise atomic coordinates, it is necessary to obtain precise diffraction intensities in the high angle region. However, as described above, precise peak positions and intensities cannot be obtained from powder diffraction data in high angle regions, which is a fatal problem for crystal structure analysis.

Figure 1 shows the differences between the crystal structure of a cocrystal of benzoic acid and flufenamic acid obtained from powder and the corresponding structure analysis result obtained from a single crystal.

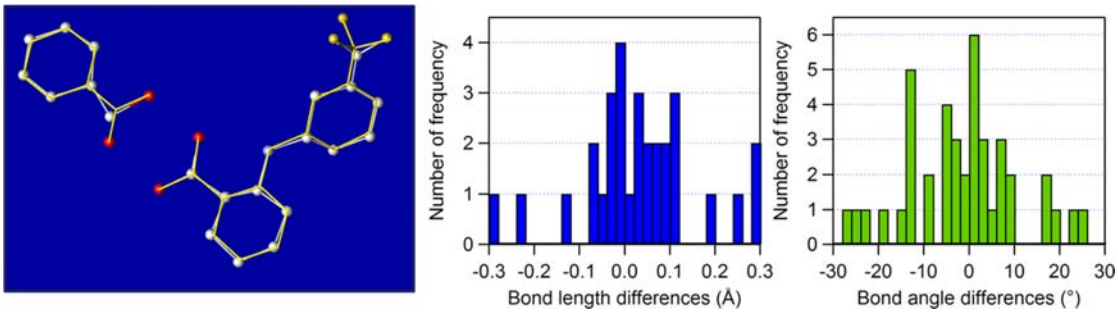


Figure 1. The ball-and-stick molecules are obtained from powder, and the yellow skeletons are from single crystal. The histograms show bond length differences and bond angle differences. You can see 0.2 or 0.3-Å differences in bond lengths, and 20 or 30-degree differences in bond angles.

Of course, very good results are sometimes obtained if the target compound has a very simple structure. But normally it is difficult to determine good atomic coordinates from powder diffraction data without any assumptions/restrictions. To apply assumptions/restrictions, we would like to introduce a new method related to restraint settings. So far, a lot of discussions have been made on restraint settings to obtain good crystal structures from powder diffraction data. The new method may not be the best, but there are several advantages which will be described later.

“Restraint” is a kind of moderate condition. Structure parameters can be refined under restraints as keeping molecular structures. The overall cost function used in least-squares refinement is expressed by the formula,

$$R = R_{wp} + R_{res}^{d} + R_{res}^{a}$$

where $\boldsymbol{R_{wp}}$ is the residual sum of squares to represent the degree of coincidence between measured and calculated data. $\boldsymbol{R_{res}^{d}}$, $\boldsymbol{R_{res}^{a}}$ are the cost functions to represent the degree of appropriateness of bond lengths and angles. And $\boldsymbol{R_{res}}$ values are expressed as follows.

$$R_{res}^{d} = \sum_{i}^{M}\left(\frac{d_{0i} - d_i}{\sigma_i^{d}}\right)^2 \qquad R_{res}^{a} = \sum_{j}^{N}\left(\frac{a_{0j} - a_j}{\sigma_j^{a}}\right)^2$$

Each R_{res} is the sum of squares of the differences between each bond length($\boldsymbol{d_i}$)/bond angle($\boldsymbol{a_j}$), and the target value($\boldsymbol{d_{0i}}$, $\boldsymbol{a_{0j}}$) divided by each sigma($\boldsymbol{\sigma_i^{d}}$, $\boldsymbol{\sigma_j^{a}}$), standard uncertainty.

But there are no guidelines how to set $\boldsymbol{d_{0i}}$, a_{0j}, sigma values ($\boldsymbol{\sigma_i^{d}}$, $\boldsymbol{\sigma_j^{a}}$).

So the purpose of this study is to determine how to set the most appropriate restraint parameters ($\boldsymbol{d_{0i}}$, $\boldsymbol{a_{0j}}$, $\boldsymbol{\sigma_i^{d}}$, $\boldsymbol{\sigma_j^{a}}$) in order to achieve the best and unique results even without knowledge of the molecular structure(s).

First, it is considered to be better to bring the restraint parameter values from a database rather than using general theoretical bond lengths. CSD/Mogul provides statistical values such as average, median, and dispersion of the target bond length, bond angle, and torsion angle based on many similar compounds or partial structures in CSD.

Second, we introduced $\boldsymbol{s_{res}}$, which determines the weight of the restraint terms. The restraint stiffness can be changed with this parameter.

$$R = R_{wp} + s_{res}\left(R_{res}^{d} + R_{res}^{a}\right)$$

If the difference between a bond length/angle and the average value is normalized by each $\boldsymbol{\sigma}$ properly, the normalized standard uncertainty of a certain number of differences in bond lengths/angles should become statistically 1. There is a certain number of bond lengths and angles in a structure. By adjusting $\boldsymbol{s_{res}}$ such that the normalized standard uncertainty ($\boldsymbol{\sigma_{Norm}}$) becomes 1, an appropriate result is considered to be obtained.

$$\sigma_{Norm} = \sqrt{\frac{1}{M+N}\left[\sum_{i}^{M}\left(\frac{d_{0i} - d_i}{\sigma_i^{d}}\right)^2 + \sum_{j}^{N}\left(\frac{a_{0j} - a_j}{\sigma_j^{a}}\right)^2\right]} = 1$$

To verify the practical effectiveness of this method, some experiments were done according to the following procedure. First, the average values and standard uncertainties for each bond length and angle of the compound to be analyzed were obtained from CSD/Mogul. Second, after the structure parameters were converged to some degree, the normalized standard uncertainty ($\boldsymbol{\sigma_{Norm}}$) was re-estimated. Then, the structure refinement and the estimation of the normalized standard uncertainty ($\boldsymbol{\sigma_{Norm}}$) were iterated. Third, when the

normalized standard uncertainty (σ_{Norm}) became 1±0.01, the optimization iteration was finished.

Table 1. Structure analysis results of the 13 compounds using restraints. ***M*** and ***N*** are the numbers of the bond lengths and angles in the asymmetric unit, and ***S*** is the goodness of fit.

#	Compound	Formula in asym. unit	S.G.	V (Å^3)	M	N	s_{res}	R_{wp}(%)	S
1	$[Ag_2(cis\text{-dbe})(C_2F_5COO)_2]_n$	$C_{12}H_9AgF_5NO_2S$	*C*2/*c*	3017.73(8)	20	30	8.0	3.18	1.2747
2	$[Ni(deen)_2](CF_3SO_3)_2$	$C_7H_{16}F_3N_2Ni_{0.5}O_3S$	*P*-1	623.92(5)	14	19	1.4	4.46	1.0980
3	$[Ni(deen)_2(H_2O)_2]Cl_2$	$C_6H_{18}ClN_2Ni_{0.5}O$	$P2_1/n$	969.13(4)	7	7	0.01	4.29	1.1228
4	Flufenamic acid benzoic acid	$C_{21}H_{16}F_3NO_4$	*P*-1	946.67(3)	30	41	71.3	2.72	1.3258
5	Flufenamic acid nicotinamide	$C_{20}H_{16}F_3N_3O_3$	$P2_1/c$	1863.70(3)	33	47	6.5	2.74	1.1959
6	$(Caffeine)_2$(oxalic acid)	$C_9H_{11}N_4O_4$	$P2_1/c$	1062.19(5)	18	26	3.1	3.98	1.2598
7	Lactose monohydrate	$C_{12}H_{24}O_{12}$	$P2_1$	778.382(17)	24	35	6.0	3.18	1.2714
8	Sucrose	$C_{12}H_{22}O_{11}$	$P2_1$	716.797(5)	24	36	1.5	2.90	1.2646
9	Tolazamide	$C_{14}H_{21}N_3O_3S$	*P*-1	784.55(3)	22	30	4.3	3.20	1.1321
10	Chlomipramine hydrochloride	$C_{19}H_{24}Cl_2N_2$	$P2_1/c$	1873.40(5)	24	33	17.0	3.59	1.5296
11	Carbamazepine	$C_{15}H_{12}N_2O$	*R*-3	5690.7(2)	20	28	29.6	5.88	1.1776
12	Phenol red	$C_{19}H_{14}O_5S$	*Pbca*	3297.05(14)	27	39	8.5	1.29	1.2262
13	Indomethacin	$C_{19}H_{16}ClN_1O_4$	*P*-1	870.05(3)	27	39	3.0	3.71	1.1752

Figure 2 shows the result of tolazamide. According to the histograms of bond length and angle differences between single crystal analysis result and powder analysis result, there are a few big differences in bond lengths and angles in terms of absolute values. However, after the differences are normalized by each standard uncertainty, all the normalized length difference and angle difference are within ±2.5 σ from each target value.

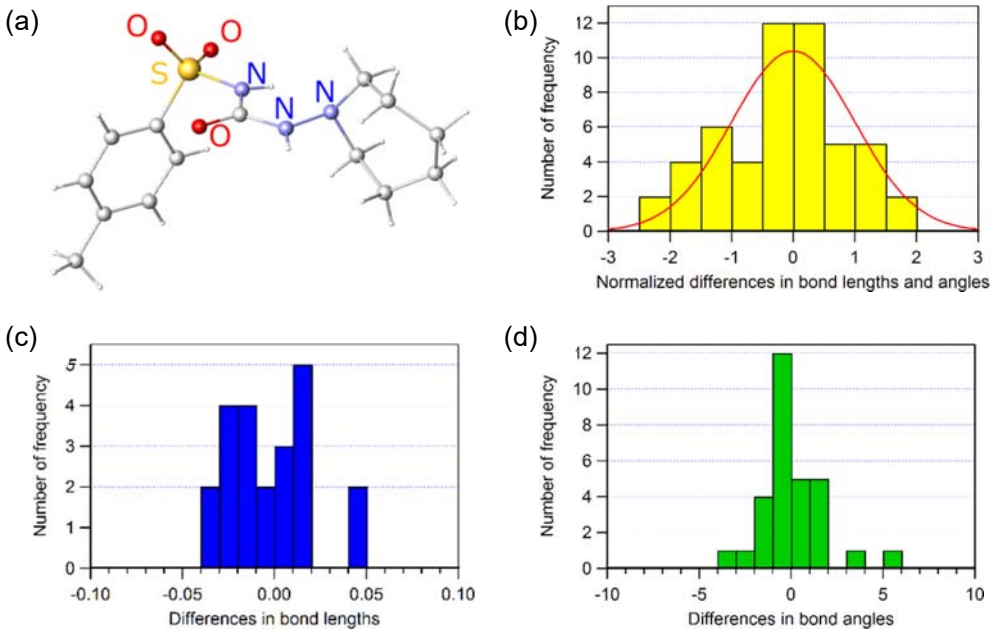


Figure 2. (a) Molecular structure of tolazamide, (b)histogram of the normalized differences in bond lengths and angles, (c) histogram of the differences in bond length, (d) histogram of the differences in bond angles.

In conclusion, chemically-proper results can be obtained by this method so that the final simulated data can reproduce the measured data. Moreover, this method can be applied to your structure analysis without a lot of knowledge about structure chemistry by making use of the database and tool, CSD/Mogul. In other words, the result does not depend on the knowledge or experience of the analyst. Last but not least, this method is suitable for automation. We hope to provide automated software to perform crystal structure analysis from powder diffraction data in near future.