



# Full wwPDB X-ray Structure Validation Report

Oct 7, 2023 – 06:16 PM EDT

PDB ID : 6DNP  
Title : Crystal structure of Mycobacterium tuberculosis malate synthase in complex with 2-F-3-Methyl-6-F-phenyldiketoacid  
Authors : Krieger, I.V.; Sacchettini, J.C.; TB Structural Genomics Consortium (TB-SGC); Mycobacterium Tuberculosis Structural Proteomics Project (XMTB)  
Deposited on : 2018-06-07  
Resolution : 1.71 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the  symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references](#) ) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35.1  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.35.1

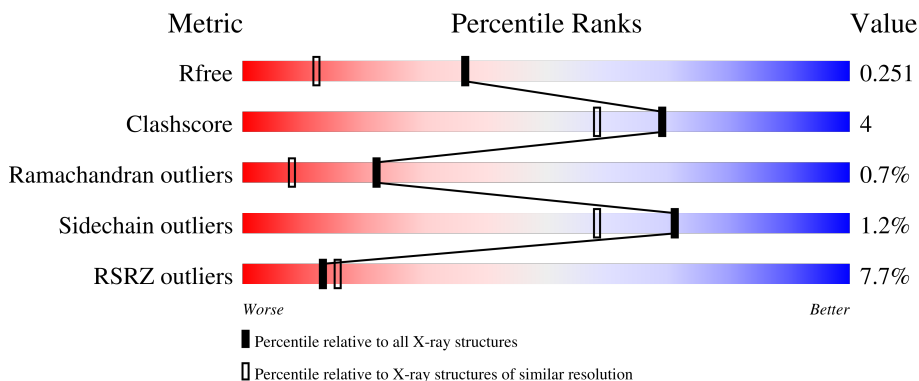
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.71 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	5722 (1.74-1.70)
Clashscore	141614	6152 (1.74-1.70)
Ramachandran outliers	138981	6051 (1.74-1.70)
Sidechain outliers	138945	6051 (1.74-1.70)
RSRZ outliers	127900	5629 (1.74-1.70)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	741	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 5845 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

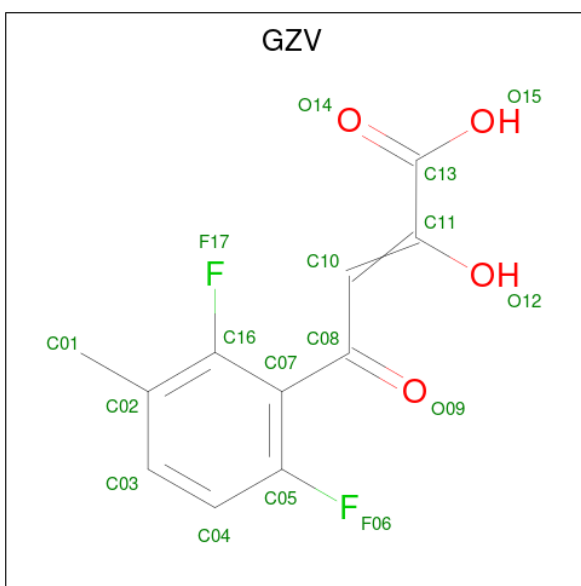
- Molecule 1 is a protein called Malate synthase G.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	712	5484	3448	967	1048	21	0	6	0

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

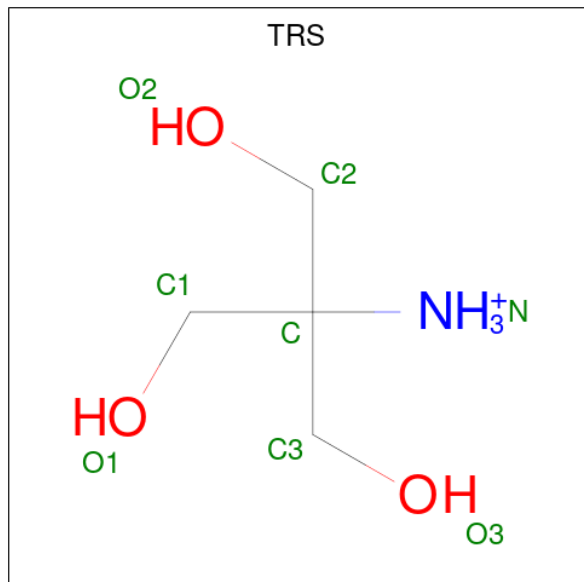
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
2	A	2	2	2	0	0

- Molecule 3 is (2Z)-4-(2,6-difluoro-3-methylphenyl)-2-hydroxy-4-oxobut-2-enoic acid (three-letter code: GZV) (formula: C<sub>11</sub>H<sub>8</sub>F<sub>2</sub>O<sub>4</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	F	O		
3	A	1	17	11	2	4	0	0

- Molecule 4 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula: C<sub>4</sub>H<sub>12</sub>NO<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	A	1	8	4	1	3	0	0

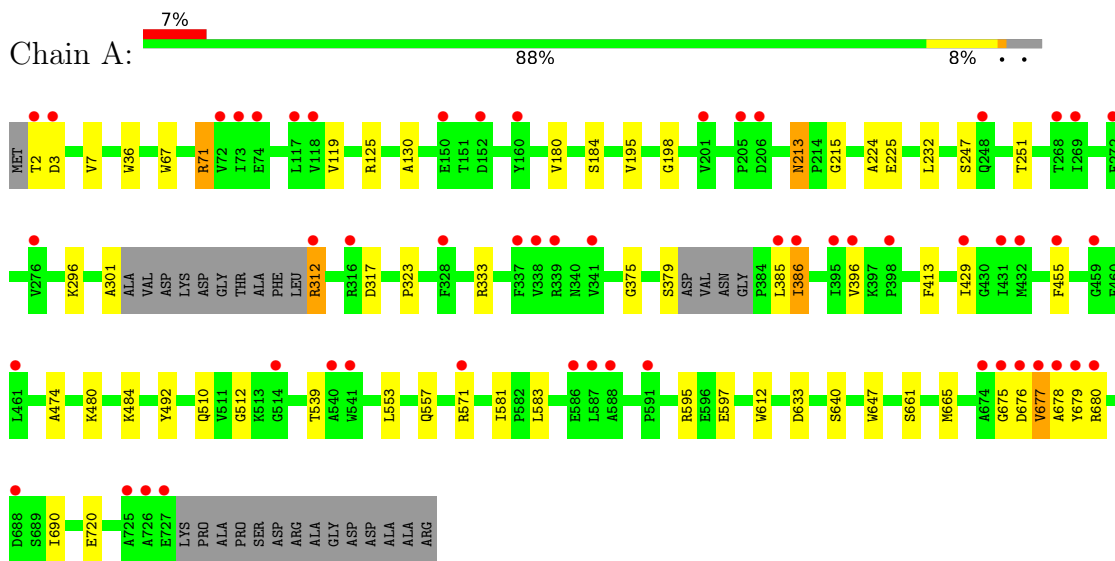
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	334	Total	O	0	0
			334	334		

### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Malate synthase G



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	79.74Å 79.74Å 227.29Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	37.62 – 1.71 37.62 – 1.71	Depositor EDS
% Data completeness (in resolution range)	98.5 (37.62-1.71) 99.4 (37.62-1.71)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.01 (at 1.71Å)	Xtrriage
Refinement program	PHENIX (1.11.1_2575: ???)	Depositor
R, $R_{free}$	0.207 , 0.251 0.206 , 0.251	Depositor DCC
$R_{free}$ test set	3954 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	30.8	Xtrriage
Anisotropy	0.436	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 41.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	5845	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.81% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, GZV, TRS

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.37	0/5605	0.55	1/7620 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	633	ASP	CB-CG-OD1	5.88	123.59	118.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5484	0	5456	40	0
2	A	2	0	0	0	0
3	A	17	0	0	0	0
4	A	8	0	12	1	0
5	A	334	0	0	3	0
All	All	5845	0	5468	41	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 4.

All (41) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:595:ARG:HH21	1:A:595:ARG:HG2	1.56	0.69
1:A:474:ALA:HA	5:A:927:HOH:O	1.96	0.65
4:A:804:TRS:H21	5:A:1004:HOH:O	1.98	0.62
1:A:675:GLY:O	1:A:677:VAL:N	2.32	0.62
1:A:680:ARG:HA	1:A:680:ARG:HE	1.69	0.58
1:A:125:ARG:HG3	5:A:1175:HOH:O	2.03	0.58
1:A:2:THR:HG23	1:A:3:ASP:H	1.70	0.57
1:A:595:ARG:HG2	1:A:595:ARG:NH2	2.20	0.56
1:A:571:ARG:O	1:A:571:ARG:HG3	2.06	0.56
1:A:413:PHE:CE2	1:A:429[B]:ILE:HG12	2.41	0.56
1:A:690:ILE:HD13	1:A:720:GLU:HG2	1.92	0.52
1:A:213:ASN:ND2	1:A:215:GLY:H	2.09	0.49
1:A:317:ASP:OD2	1:A:385:LEU:HD11	2.13	0.47
1:A:455:PHE:HB3	1:A:510:GLN:HB2	1.96	0.47
1:A:333:ARG:NH1	1:A:385:LEU:HB3	2.29	0.47
1:A:413:PHE:HE2	1:A:429[B]:ILE:HG12	1.78	0.47
1:A:180:VAL:HG21	1:A:232:LEU:HD13	1.96	0.46
1:A:296:LYS:O	1:A:379:SER:HB3	2.16	0.46
1:A:680:ARG:HA	1:A:680:ARG:NE	2.31	0.46
1:A:67:TRP:CZ3	1:A:71:ARG:HD3	2.51	0.45
1:A:333:ARG:HB2	1:A:333:ARG:CZ	2.46	0.45
1:A:553:LEU:O	1:A:557:GLN:HG3	2.17	0.44
1:A:247:SER:O	1:A:251:THR:OG1	2.34	0.44
1:A:7:VAL:HG11	1:A:36:TRP:HB3	1.99	0.44
1:A:661:SER:O	1:A:665:MET:HG2	2.18	0.43
1:A:119:VAL:HG11	1:A:130:ALA:HB2	1.99	0.43
1:A:385:LEU:HD13	1:A:385:LEU:HA	1.89	0.42
1:A:581:ILE:HG22	1:A:583:LEU:HG	2.01	0.42
1:A:198:GLY:HA2	1:A:323:PRO:O	2.19	0.42
1:A:396:VAL:HG22	1:A:455:PHE:CZ	2.55	0.42
1:A:677:VAL:O	1:A:679:TYR:N	2.53	0.42
1:A:67:TRP:O	1:A:71:ARG:HD2	2.19	0.42
1:A:195:VAL:HG13	1:A:224:ALA:HB1	2.02	0.41
1:A:125:ARG:HH21	1:A:312:ARG:NE	2.18	0.41
1:A:333:ARG:NH1	1:A:385:LEU:HG	2.35	0.41
1:A:597:GLU:HG2	1:A:647:TRP:CH2	2.56	0.41
1:A:301:ALA:C	1:A:312:ARG:HD2	2.41	0.41
1:A:67:TRP:CE3	1:A:71:ARG:HD3	2.56	0.41
1:A:512:GLY:HA2	1:A:539:THR:O	2.21	0.41
1:A:375:GLY:O	1:A:386:ILE:HB	2.21	0.40

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:480:LYS:NZ	1:A:640:SER:OG	2.54	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	712/741 (96%)	691 (97%)	16 (2%)	5 (1%)	22   8

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	676	ASP
1	A	677	VAL
1	A	678	ALA
1	A	225	GLU
1	A	386	ILE

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	579/594 (98%)	572 (99%)	7 (1%)	71   58

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	71	ARG
1	A	184	SER
1	A	213	ASN
1	A	312	ARG
1	A	484	LYS
1	A	492	TYR
1	A	612	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	213	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
3	GZV	A	803	2	16,17,17	1.11	1 (6%)	22,24,24	2.11	6 (27%)
4	TRS	A	804	-	7,7,7	0.99	0	9,9,9	1.13	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GZV	A	803	2	-	7/12/12/12	0/1/1/1
4	TRS	A	804	-	-	6/9/9/9	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	803	GZV	O12-C11	2.52	1.40	1.34

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	803	GZV	O12-C11-C13	4.72	116.86	114.01
3	A	803	GZV	C05-C07-C16	4.35	119.14	115.84
3	A	803	GZV	O15-C13-C11	3.74	119.53	115.91
3	A	803	GZV	C10-C11-C13	-3.28	118.73	122.69
3	A	803	GZV	C03-C02-C16	2.70	120.11	115.75
3	A	803	GZV	C11-C10-C08	2.50	121.94	120.57

There are no chirality outliers.

All (13) torsion outliers are listed below:

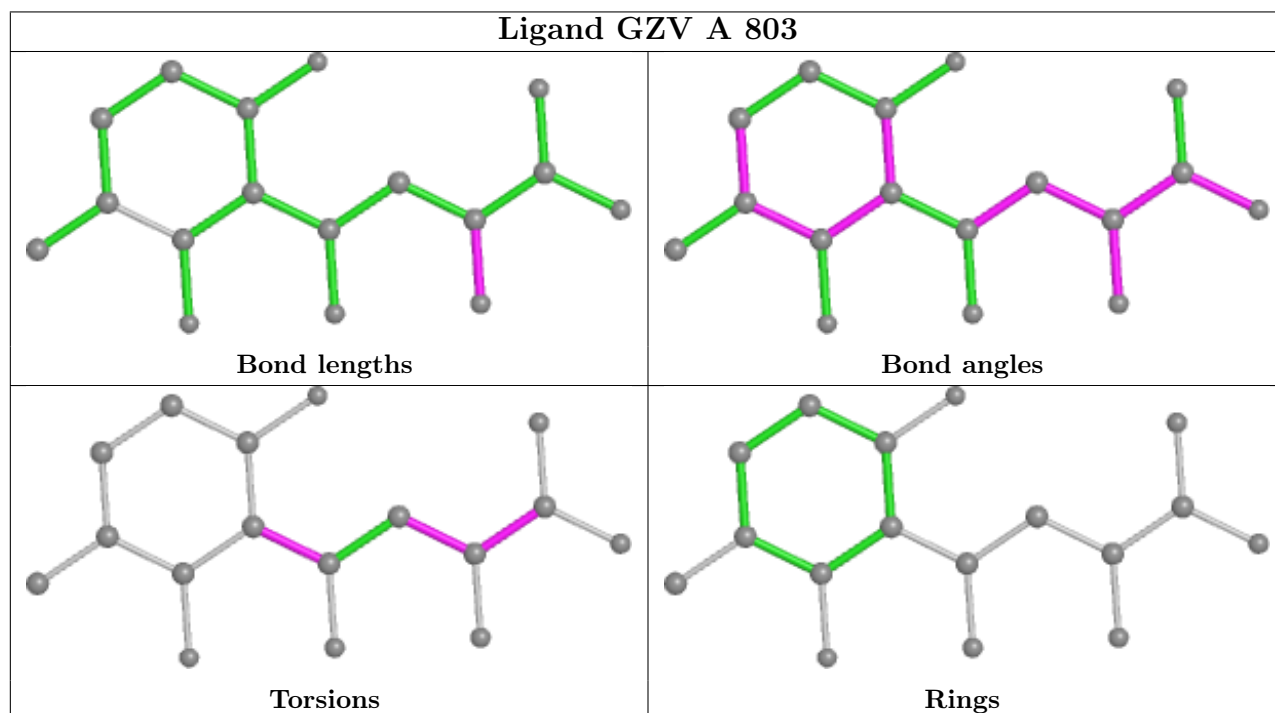
Mol	Chain	Res	Type	Atoms
3	A	803	GZV	C08-C10-C11-C13
3	A	803	GZV	C08-C10-C11-O12
3	A	803	GZV	O12-C11-C13-O15
4	A	804	TRS	C1-C-C2-O2
4	A	804	TRS	C3-C-C2-O2
4	A	804	TRS	C2-C-C3-O3
3	A	803	GZV	C05-C07-C08-O09
4	A	804	TRS	N-C-C2-O2
4	A	804	TRS	N-C-C3-O3
3	A	803	GZV	C10-C11-C13-O14
3	A	803	GZV	C10-C11-C13-O15
4	A	804	TRS	C1-C-C3-O3
3	A	803	GZV	O12-C11-C13-O14

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	804	TRS	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	712/741 (96%)	0.25	55 (7%) <b>13</b> <b>16</b>	25, 36, 61, 92	0

All (55) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	385	LEU	8.8
1	A	74	GLU	7.9
1	A	587	LEU	7.7
1	A	677	VAL	6.0
1	A	152	ASP	6.0
1	A	675	GLY	5.4
1	A	726	ALA	5.0
1	A	73	ILE	4.9
1	A	72	VAL	4.5
1	A	674	ALA	4.5
1	A	118	VAL	4.2
1	A	312	ARG	4.1
1	A	269	ILE	4.1
1	A	588	ALA	4.1
1	A	680	ARG	4.1
1	A	248	GLN	4.1
1	A	676	ASP	4.0
1	A	338	VAL	3.6
1	A	386	ILE	3.5
1	A	2	THR	3.4
1	A	396	VAL	3.4
1	A	398	PRO	3.2
1	A	725	ALA	3.1
1	A	461	LEU	3.1
1	A	541	TRP	3.1
1	A	117	LEU	3.1
1	A	727	GLU	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	571	ARG	3.0
1	A	341	VAL	2.8
1	A	339	ARG	2.7
1	A	337	PHE	2.5
1	A	395[A]	ILE	2.5
1	A	206	ASP	2.5
1	A	432	MET	2.5
1	A	272	PHE	2.5
1	A	678	ALA	2.4
1	A	201	VAL	2.4
1	A	586	GLU	2.4
1	A	455	PHE	2.3
1	A	429[A]	ILE	2.3
1	A	160	TYR	2.3
1	A	205	PRO	2.2
1	A	514	GLY	2.2
1	A	3	ASP	2.2
1	A	679	TYR	2.1
1	A	591	PRO	2.1
1	A	276	VAL	2.1
1	A	540	ALA	2.1
1	A	316	ARG	2.1
1	A	328	PHE	2.1
1	A	268	THR	2.0
1	A	688	ASP	2.0
1	A	150	GLU	2.0
1	A	431	ILE	2.0
1	A	459	GLY	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

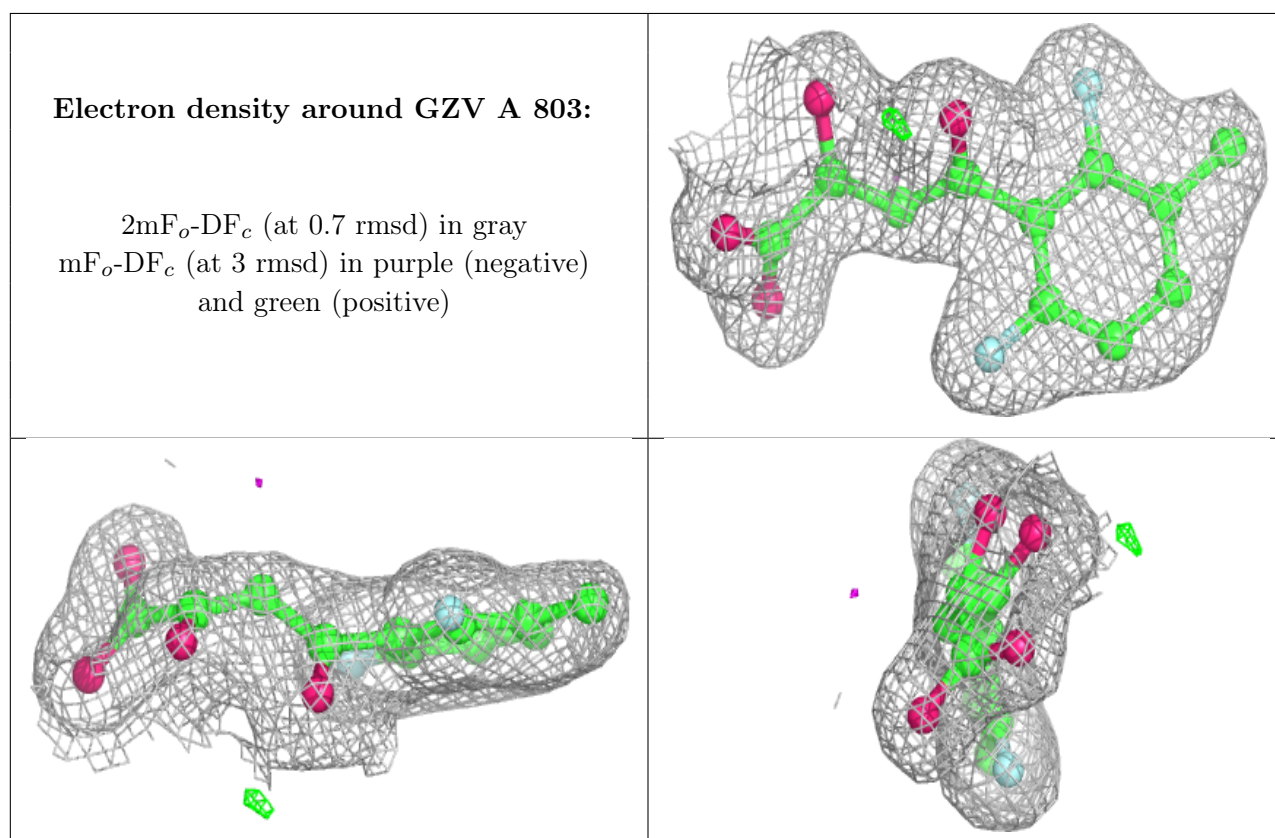
## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,

median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q < 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	TRS	A	804	8/8	0.87	0.25	39,49,55,56	0
2	MG	A	802	1/1	0.90	0.15	41,41,41,41	0
3	GZV	A	803	17/17	0.95	0.19	25,29,32,37	0
2	MG	A	801	1/1	0.97	0.19	26,26,26,26	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.



## 6.5 Other polymers [i](#)

There are no such residues in this entry.