



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 08:30 AM UTC

PDB ID : 2D40 / pdb\_00002d40  
Title : Crystal Structure of Z3393 from Escherichia coli O157:H7  
Authors : Adams, M.A.; Jia, Z.; Montreal-Kingston Bacterial Structural Genomics Initiative (BSGI)  
Deposited on : 2005-10-05  
Resolution : 2.41 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

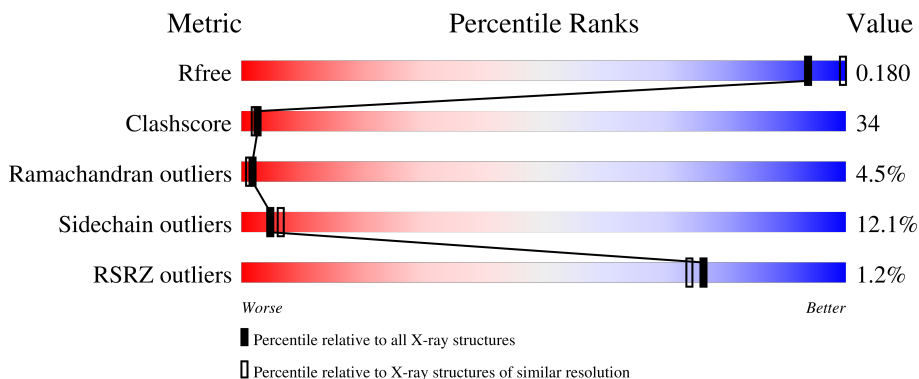
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.41 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))
$R_{free}$	180053	6062 (2.44-2.40)
Clashscore	190562	6562 (2.44-2.40)
Ramachandran outliers	187476	6481 (2.44-2.40)
Sidechain outliers	187428	6482 (2.44-2.40)
RSRZ outliers	180081	6066 (2.44-2.40)

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	354	<div> <div>20%</div> <div>39%</div> <div>17%</div> <div>6%</div> <div>18%</div> </div>
1	B	354	<div> <div>20%</div> <div>36%</div> <div>16%</div> <div>8%</div> <div>20%</div> </div>
1	C	354	<div> <div>23%</div> <div>32%</div> <div>23%</div> <div>•</div> <div>18%</div> </div>
1	D	354	<div> <div>22%</div> <div>33%</div> <div>19%</div> <div>6%</div> <div>19%</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 9579 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 putative gentisate 1,2-dioxygenase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	289	Total	C	N	O	S	Se	0	0	0
			2280	1447	399	426	2	6			
1	B	284	Total	C	N	O	S	Se	0	0	0
			2235	1418	392	417	2	6			
1	C	289	Total	C	N	O	S	Se	0	0	0
			2281	1447	399	427	2	6			
1	D	287	Total	C	N	O	S	Se	0	0	0
			2263	1437	396	422	2	6			

There are 76 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	MSE	-	expression tag	UNP Q8X655
A	-10	GLY	-	expression tag	UNP Q8X655
A	-9	SER	-	expression tag	UNP Q8X655
A	-8	SER	-	expression tag	UNP Q8X655
A	-7	HIS	-	expression tag	UNP Q8X655
A	-6	HIS	-	expression tag	UNP Q8X655
A	-5	HIS	-	expression tag	UNP Q8X655
A	-4	HIS	-	expression tag	UNP Q8X655
A	-3	HIS	-	expression tag	UNP Q8X655
A	-2	HIS	-	expression tag	UNP Q8X655
A	-1	GLY	-	expression tag	UNP Q8X655
A	0	SER	-	expression tag	UNP Q8X655
A	1	MSE	MET	modified residue	UNP Q8X655
A	96	MSE	MET	modified residue	UNP Q8X655
A	131	MSE	MET	modified residue	UNP Q8X655
A	199	MSE	MET	modified residue	UNP Q8X655
A	241	MSE	MET	modified residue	UNP Q8X655
A	253	MSE	MET	modified residue	UNP Q8X655
A	256	MSE	MET	modified residue	UNP Q8X655
B	-11	MSE	-	expression tag	UNP Q8X655
B	-10	GLY	-	expression tag	UNP Q8X655

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-9	SER	-	expression tag	UNP Q8X655
B	-8	SER	-	expression tag	UNP Q8X655
B	-7	HIS	-	expression tag	UNP Q8X655
B	-6	HIS	-	expression tag	UNP Q8X655
B	-5	HIS	-	expression tag	UNP Q8X655
B	-4	HIS	-	expression tag	UNP Q8X655
B	-3	HIS	-	expression tag	UNP Q8X655
B	-2	HIS	-	expression tag	UNP Q8X655
B	-1	GLY	-	expression tag	UNP Q8X655
B	0	SER	-	expression tag	UNP Q8X655
B	1	MSE	MET	modified residue	UNP Q8X655
B	96	MSE	MET	modified residue	UNP Q8X655
B	131	MSE	MET	modified residue	UNP Q8X655
B	199	MSE	MET	modified residue	UNP Q8X655
B	241	MSE	MET	modified residue	UNP Q8X655
B	253	MSE	MET	modified residue	UNP Q8X655
B	256	MSE	MET	modified residue	UNP Q8X655
C	-11	MSE	-	expression tag	UNP Q8X655
C	-10	GLY	-	expression tag	UNP Q8X655
C	-9	SER	-	expression tag	UNP Q8X655
C	-8	SER	-	expression tag	UNP Q8X655
C	-7	HIS	-	expression tag	UNP Q8X655
C	-6	HIS	-	expression tag	UNP Q8X655
C	-5	HIS	-	expression tag	UNP Q8X655
C	-4	HIS	-	expression tag	UNP Q8X655
C	-3	HIS	-	expression tag	UNP Q8X655
C	-2	HIS	-	expression tag	UNP Q8X655
C	-1	GLY	-	expression tag	UNP Q8X655
C	0	SER	-	expression tag	UNP Q8X655
C	1	MSE	MET	modified residue	UNP Q8X655
C	96	MSE	MET	modified residue	UNP Q8X655
C	131	MSE	MET	modified residue	UNP Q8X655
C	199	MSE	MET	modified residue	UNP Q8X655
C	241	MSE	MET	modified residue	UNP Q8X655
C	253	MSE	MET	modified residue	UNP Q8X655
C	256	MSE	MET	modified residue	UNP Q8X655
D	-11	MSE	-	expression tag	UNP Q8X655
D	-10	GLY	-	expression tag	UNP Q8X655
D	-9	SER	-	expression tag	UNP Q8X655
D	-8	SER	-	expression tag	UNP Q8X655
D	-7	HIS	-	expression tag	UNP Q8X655
D	-6	HIS	-	expression tag	UNP Q8X655

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	-5	HIS	-	expression tag	UNP Q8X655
D	-4	HIS	-	expression tag	UNP Q8X655
D	-3	HIS	-	expression tag	UNP Q8X655
D	-2	HIS	-	expression tag	UNP Q8X655
D	-1	GLY	-	expression tag	UNP Q8X655
D	0	SER	-	expression tag	UNP Q8X655
D	1	MSE	MET	modified residue	UNP Q8X655
D	96	MSE	MET	modified residue	UNP Q8X655
D	131	MSE	MET	modified residue	UNP Q8X655
D	199	MSE	MET	modified residue	UNP Q8X655
D	241	MSE	MET	modified residue	UNP Q8X655
D	253	MSE	MET	modified residue	UNP Q8X655
D	256	MSE	MET	modified residue	UNP Q8X655

- Molecule 2 is FE (III) ION (CCD ID: FE) (formula: Fe).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Fe 1 1	0	0
2	B	1	Total Fe 1 1	0	0
2	C	1	Total Fe 1 1	0	0
2	D	1	Total Fe 1 1	0	0

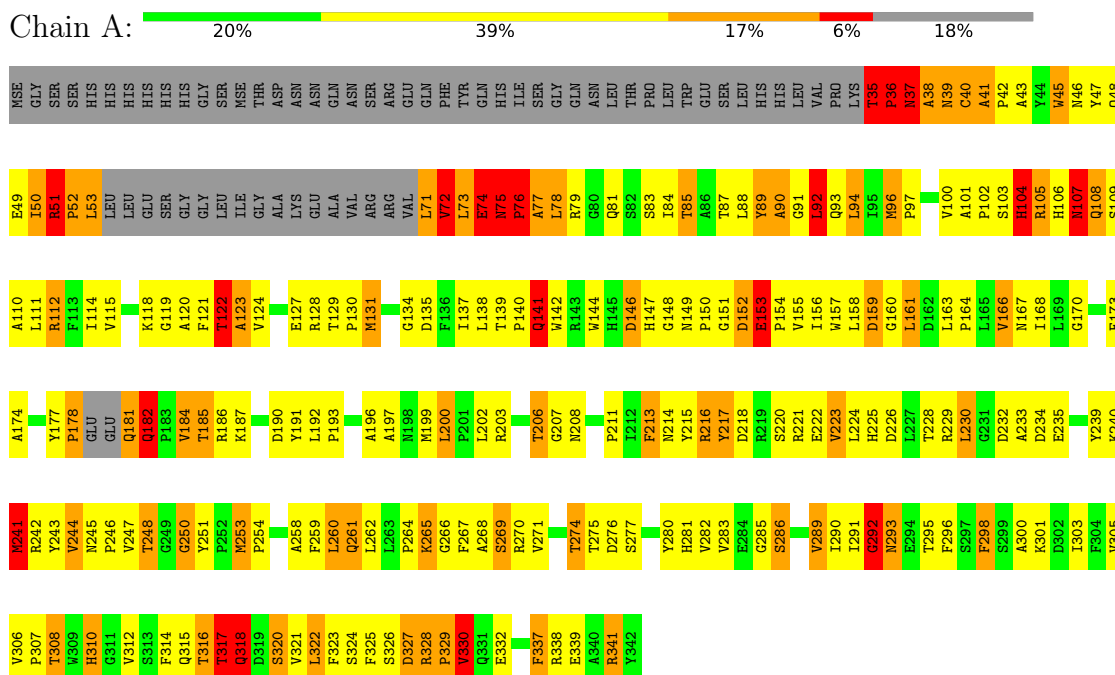
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	110	Total O 110 110	0	0
3	B	129	Total O 129 129	0	0
3	C	130	Total O 130 130	0	0
3	D	147	Total O 147 147	0	0

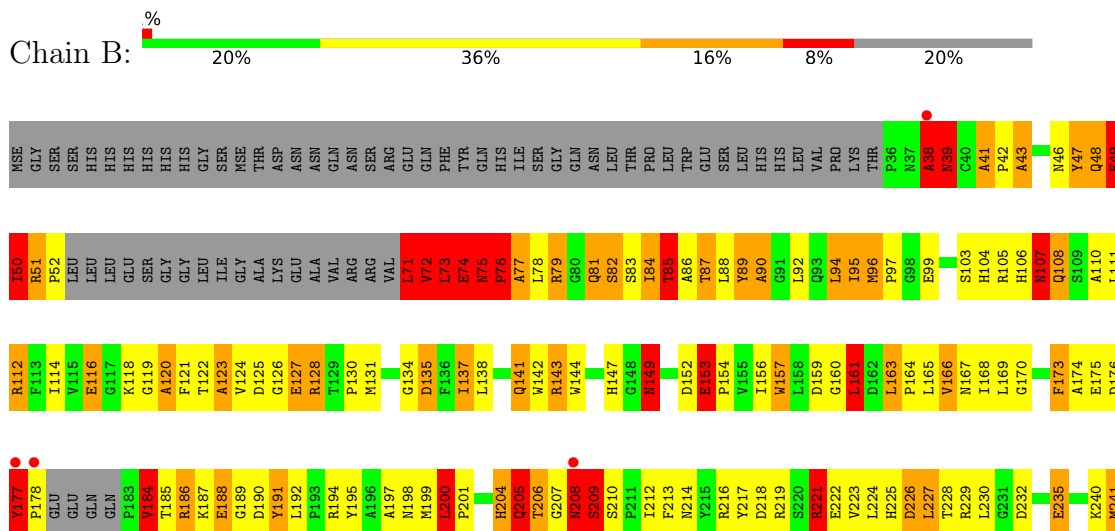
### 3 Residue-property plots

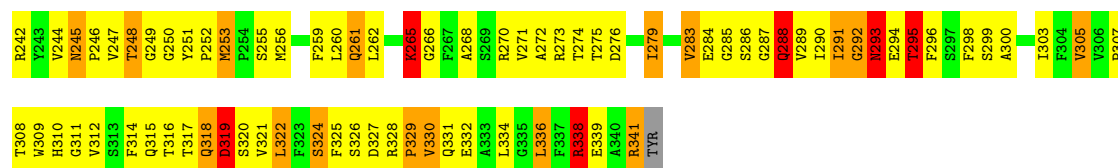
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: putative gentisate 1,2-dioxygenase

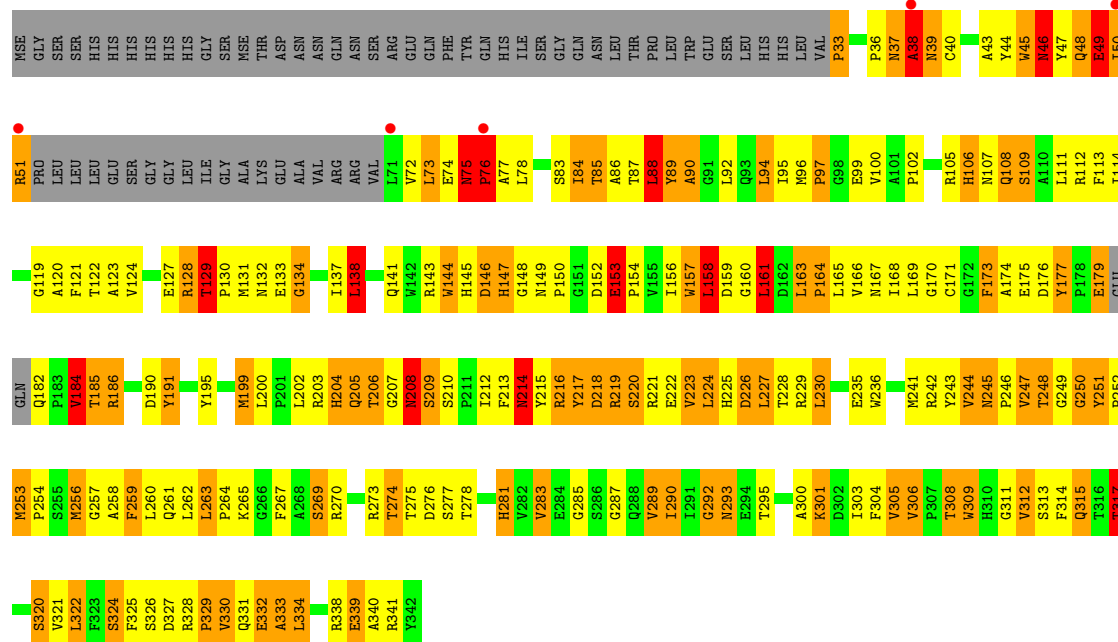
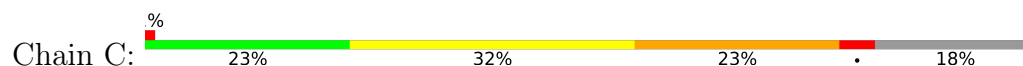


- Molecule 1: putative gentisate 1,2-dioxygenase

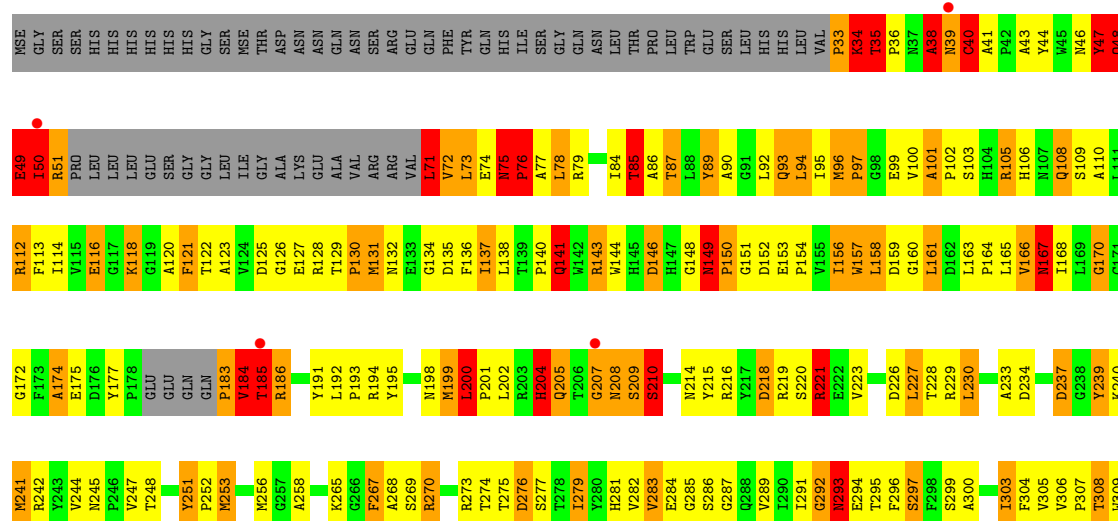
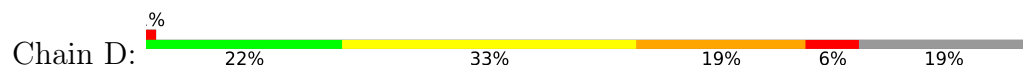


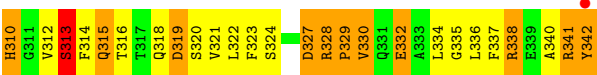


• Molecule 1: putative gentisate 1,2-dioxygenase



• Molecule 1: putative gentisate 1,2-dioxygenase







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	54.13Å 76.08Å 85.46Å 114.08° 94.93° 108.11°	Depositor
Resolution (Å)	34.50 – 2.41 34.50 – 2.41	Depositor EDS
% Data completeness (in resolution range)	100.0 (34.50-2.41) 93.6 (34.50-2.41)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	6.71 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, $R_{free}$	0.173 , 0.249 0.188 , 0.180	Depositor DCC
$R_{free}$ test set	2054 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	27.0	Xtriage
Anisotropy	0.459	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 48.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	9579	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.47% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: FE

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	2.92	201/2340 (8.6%)	1.88	36/3182 (1.1%)
1	B	2.80	176/2294 (7.7%)	1.88	42/3118 (1.3%)
1	C	2.84	189/2341 (8.1%)	1.95	69/3181 (2.2%)
1	D	2.93	202/2323 (8.7%)	1.93	68/3156 (2.2%)
All	All	2.87	768/9298 (8.3%)	1.91	215/12637 (1.7%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	2
1	D	0	2
All	All	0	5

The worst 5 of 768 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	50	ILE	CA-CB	21.08	1.83	1.54
1	A	77	ALA	CA-CB	-19.15	1.21	1.53
1	D	96	MSE	SE-CE	16.07	2.43	1.95
1	A	40	CYS	C-O	15.22	1.43	1.24
1	D	50	ILE	N-CA	15.21	1.65	1.46

The worst 5 of 215 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	40	CYS	N-CA-C	-13.13	82.83	110.80

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	76	PRO	CA-C-N	10.04	137.00	120.23
1	C	76	PRO	C-N-CA	10.04	137.00	120.23
1	B	81	GLN	N-CA-C	-9.75	101.38	113.18
1	C	128	ARG	NE-CZ-NH2	9.67	127.91	119.20

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	37	ASN	Peptide
1	B	39	ASN	Peptide
1	B	76	PRO	Peptide
1	D	149	ASN	Mainchain
1	D	47	TYR	Peptide

## 5.2 Too-close contacts

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	2280	0	2134	139	0
1	B	2235	0	2093	185	0
1	C	2281	0	2135	145	0
1	D	2263	0	2122	150	2
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	110	0	0	9	0
3	B	129	0	0	26	0
3	C	130	0	0	22	0
3	D	147	0	0	25	2
All	All	9579	0	8484	593	2

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 34.

The worst 5 of 593 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:D:76:PRO:CB	1:D:76:PRO:CA	1.75	1.61
1:B:227:LEU:CG	1:B:227:LEU:CD2	1.74	1.59
1:C:50:ILE:CB	1:C:50:ILE:CG2	1.76	1.59
1:D:39:ASN:CG	1:D:39:ASN:CB	1.74	1.56
1:D:50:ILE:CA	1:D:50:ILE:CB	1.83	1.56

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:35:THR:N	3:D:1122:HOH:O[1_455]	1.91	0.29
1:D:33:PRO:CB	3:D:1136:HOH:O[1_455]	1.98	0.22

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	283/354 (80%)	257 (91%)	13 (5%)	13 (5%)	2	1
1	B	278/354 (78%)	248 (89%)	15 (5%)	15 (5%)	1	1
1	C	283/354 (80%)	251 (89%)	25 (9%)	7 (2%)	4	4
1	D	281/354 (79%)	249 (89%)	16 (6%)	16 (6%)	1	0
All	All	1125/1416 (79%)	1005 (89%)	69 (6%)	51 (4%)	2	1

5 of 51 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	38	ALA
1	A	50	ILE
1	A	77	ALA
1	A	141	GLN
1	A	182	GLN

### 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	232/293 (79%)	207 (89%)	25 (11%)	6	9
1	B	227/293 (78%)	198 (87%)	29 (13%)	4	5
1	C	232/293 (79%)	204 (88%)	28 (12%)	5	6
1	D	230/293 (78%)	201 (87%)	29 (13%)	4	6
All	All	921/1172 (79%)	810 (88%)	111 (12%)	5	6

5 of 111 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	73	LEU
1	D	322	LEU
1	C	205	GLN
1	D	313	SER
1	D	167	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 44 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	205	GLN
1	D	75	ASN
1	C	208	ASN
1	C	315	GLN
1	D	208	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 oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	283/354 (79%)	-0.42	0 <b>100</b> <b>100</b>	13, 26, 47, 55	0
1	B	278/354 (78%)	-0.33	4 (1%) 73 70	15, 27, 51, 64	0
1	C	283/354 (79%)	-0.28	5 (1%) 67 64	17, 28, 50, 66	0
1	D	281/354 (79%)	-0.35	5 (1%) 67 64	14, 26, 49, 65	0
All	All	1125/1416 (79%)	-0.34	14 (1%) 76 74	13, 27, 49, 66	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	50	ILE	4.5
1	C	50	ILE	3.9
1	B	208	ASN	3.2
1	D	342	TYR	3.2
1	D	207	GLY	2.8

### 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 oligosaccharides 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( $\text{\AA}^2$ )	Q<0.9
2	FE	B	1002	1/1	0.99	0.08	26,26,26,26	0
2	FE	D	1004	1/1	0.99	0.05	31,31,31,31	0
2	FE	C	1003	1/1	1.00	0.05	31,31,31,31	0
2	FE	A	1001	1/1	1.00	0.02	27,27,27,27	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.