



wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 04:39 pm BST

PDB ID : 6UFQ
Title : Crystal structure of D678N GoxA bound to glycine
Authors : Yukl, E.T.
Deposited on : 2019-09-24
Resolution : 2.51 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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

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 : 2.11
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) : 2.11

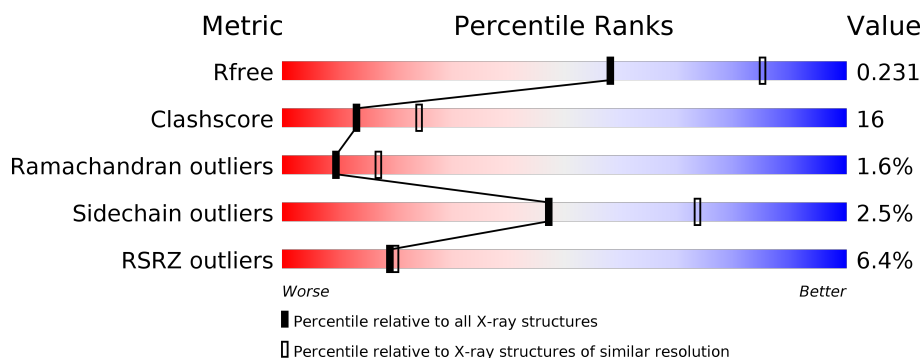
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.51 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 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	816	<div> <div>5%</div> <div> <div></div> <div>70%</div> <div>25%</div> <div>• •</div> </div> </div>
1	B	816	<div> <div>6%</div> <div> <div></div> <div>69%</div> <div>25%</div> <div>• •</div> </div> </div>
1	C	816	<div> <div>4%</div> <div> <div></div> <div>67%</div> <div>24%</div> <div>• 7%</div> </div> </div>
1	D	816	<div> <div>9%</div> <div> <div></div> <div>70%</div> <div>25%</div> <div>• •</div> </div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 25587 atoms, of which 20 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 Glycine Oxidase GoxA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	790	Total	C	N	O	S	0	0	0
			6247	3948	1064	1215	20			
1	B	790	Total	C	N	O	S	0	0	0
			6246	3946	1063	1217	20			
1	C	760	Total	C	N	O	S	0	0	0
			6002	3797	1021	1164	20			
1	D	787	Total	C	N	O	S	0	0	0
			6221	3931	1060	1210	20			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	678	ASN	ASP	engineered mutation	UNP A0A161XU12
B	678	ASN	ASP	engineered mutation	UNP A0A161XU12
C	678	ASN	ASP	engineered mutation	UNP A0A161XU12
D	678	ASN	ASP	engineered mutation	UNP A0A161XU12

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

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

- Molecule 3 is GLYCINE (three-letter code: GLY) (formula: C₂H₅NO₂).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	H	N	O	0	0
			10	2	5	1	2		
3	B	1	Total	C	H	N	O	0	0
			10	2	5	1	2		
3	C	1	Total	C	H	N	O	0	0
			10	2	5	1	2		
3	D	1	Total	C	H	N	O	0	0
			10	2	5	1	2		

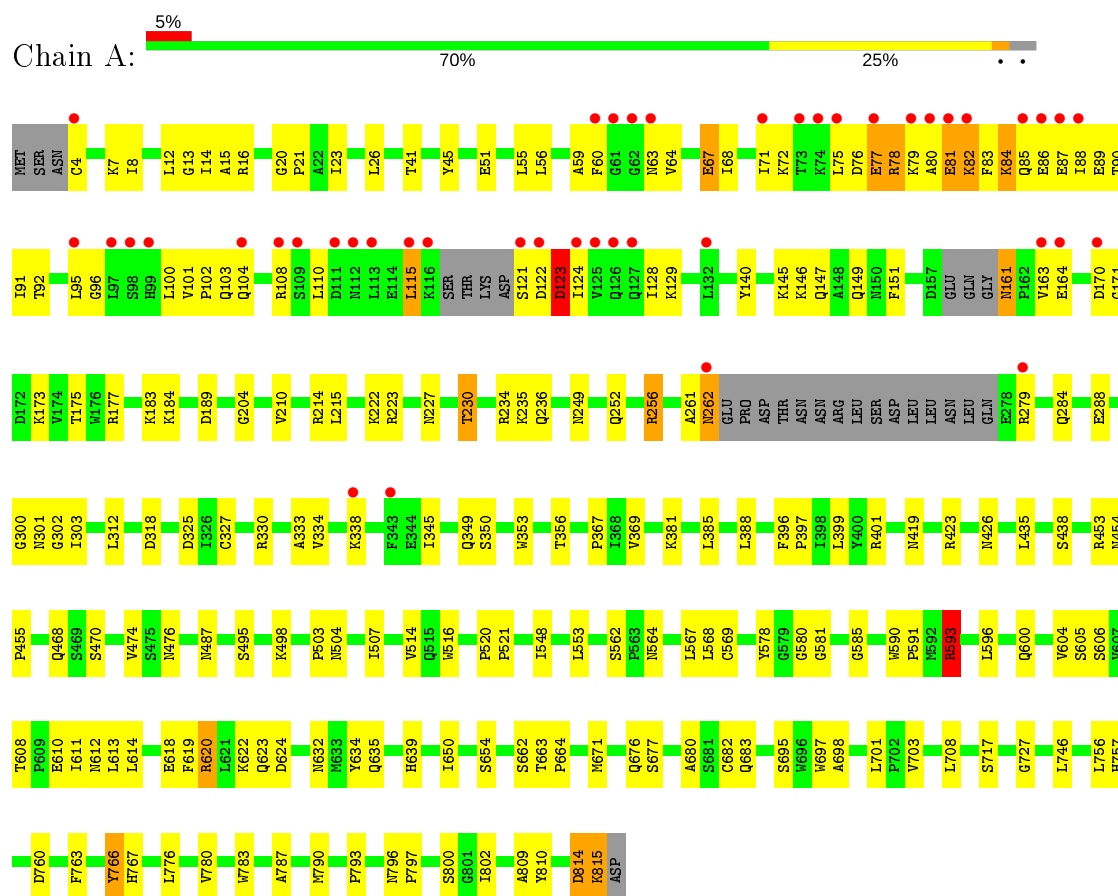
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	240	Total	O	0	0
			240	240		
4	B	215	Total	O	0	0
			215	215		
4	C	195	Total	O	0	0
			195	195		
4	D	177	Total	O	0	0
			177	177		

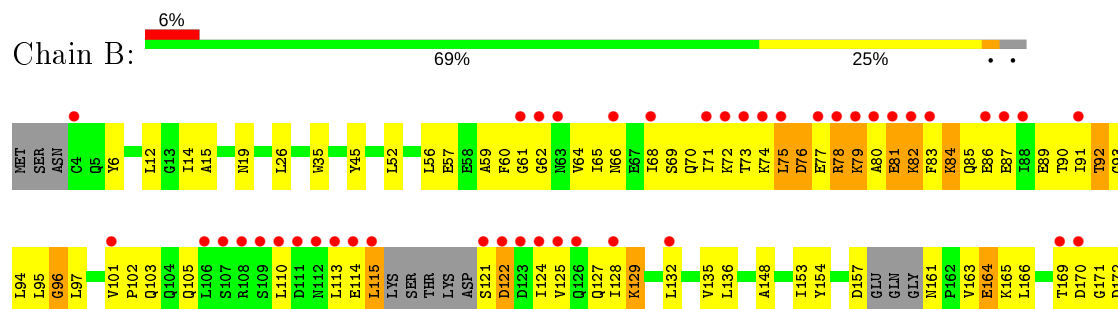
3 Residue-property plots

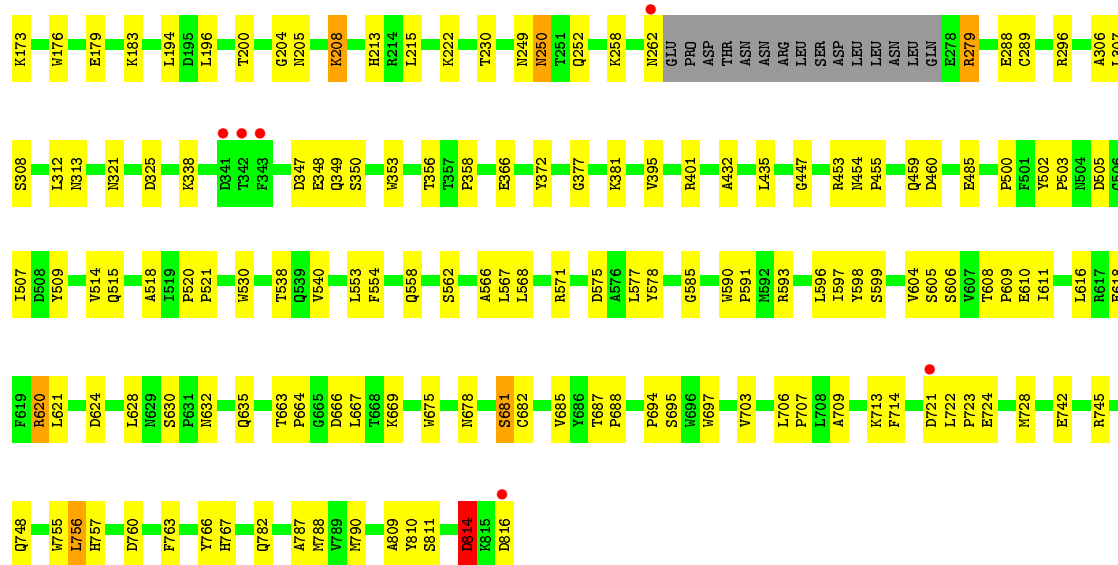
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Glycine Oxidase GoxA

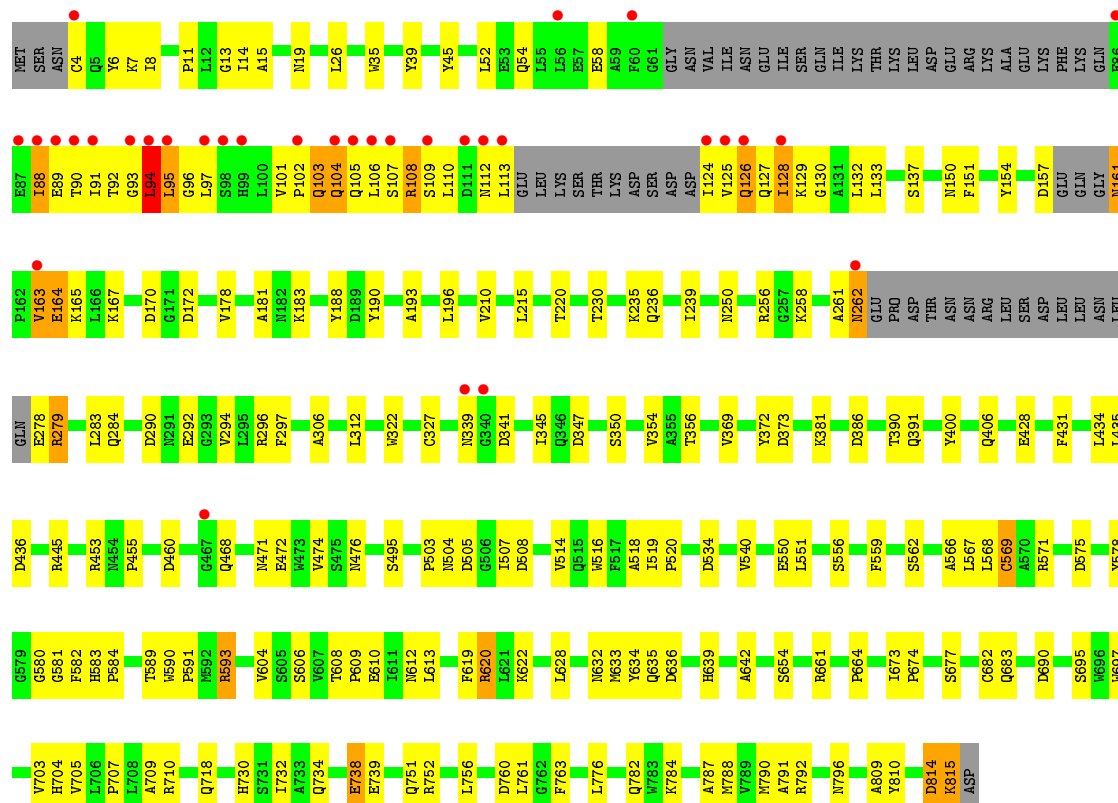


• Molecule 1: Glycine Oxidase GoxA



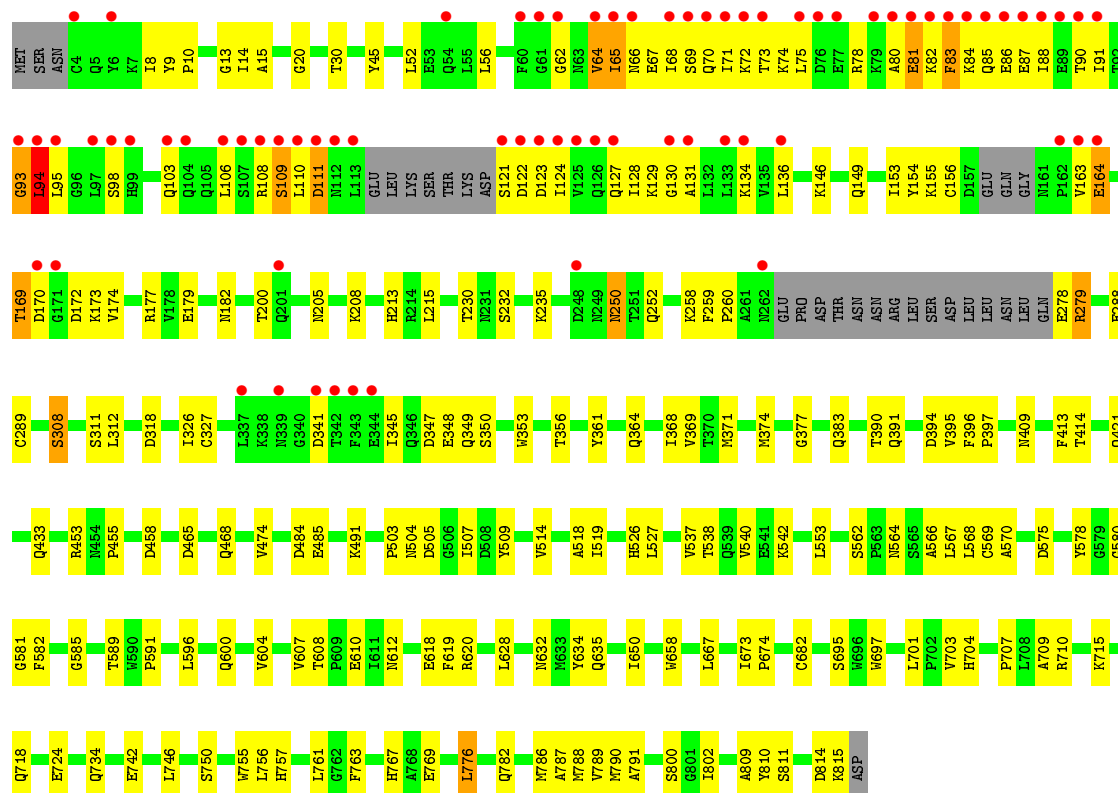


• Molecule 1: Glycine Oxidase GoxA



• Molecule 1: Glycine Oxidase GoxA





4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	109.14Å 91.57Å 178.41Å 90.00° 91.53° 90.00°	Depositor
Resolution (Å)	46.87 – 2.51 46.87 – 2.51	Depositor EDS
% Data completeness (in resolution range)	99.6 (46.87-2.51) 99.6 (46.87-2.51)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.26 (at 2.51Å)	Xtriage
Refinement program	PHENIX 1.14 _3260	Depositor
R, R_{free}	0.176 , 0.231 0.176 , 0.231	Depositor DCC
R_{free} test set	6022 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å ²)	31.9	Xtriage
Anisotropy	0.656	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 41.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	0.045 for h,-k,-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	25587	wwPDB-VP
Average B, all atoms (Å ²)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.56% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: TRQ, MG

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	0.51	3/6379 (0.0%)	0.63	0/8680
1	B	0.49	2/6378 (0.0%)	0.62	3/8680 (0.0%)
1	C	0.50	5/6132 (0.1%)	0.66	4/8351 (0.0%)
1	D	0.46	0/6353	0.62	1/8646 (0.0%)
All	All	0.49	10/25242 (0.0%)	0.63	8/34357 (0.0%)

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	B	0	1
1	C	0	2
1	D	0	1
All	All	0	4

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	327	CYS	CB-SG	7.13	1.94	1.82
1	A	327	CYS	CB-SG	6.87	1.94	1.82
1	B	755	TRP	NE1-CE2	-6.12	1.29	1.37
1	A	766	TYR	CE1-CZ	-6.11	1.30	1.38
1	C	569	CYS	CB-SG	-6.00	1.72	1.82

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	167	LYS	CD-CE-NZ	-13.56	80.51	111.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	167	LYS	CB-CG-CD	-9.30	87.41	111.60
1	B	75	LEU	CB-CG-CD1	-7.01	99.09	111.00
1	C	94	LEU	CA-CB-CG	6.85	131.06	115.30
1	D	776	LEU	CA-CB-CG	6.16	129.46	115.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	814	ASP	Peptide
1	C	93	GLY	Peptide
1	C	94	LEU	Peptide
1	D	94	LEU	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	6247	0	6021	192	0
1	B	6246	0	6012	209	0
1	C	6002	0	5767	200	1
1	D	6221	0	5991	207	1
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	5	5	2	3	0
3	B	5	5	2	3	0
3	C	5	5	2	2	0
3	D	5	5	2	0	0
4	A	240	0	0	13	0
4	B	215	0	0	11	0
4	C	195	0	0	19	0
4	D	177	0	0	18	0
All	All	25567	20	23799	790	1

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 16.

The worst 5 of 790 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:C:91:ILE:C	1:C:94:LEU:HD13	1.22	1.50
1:C:91:ILE:O	1:C:94:LEU:HD13	1.36	1.24
1:C:91:ILE:C	1:C:94:LEU:CD1	2.06	1.23
1:A:75:LEU:HD11	1:A:80:ALA:HB3	1.23	1.16
1:D:80:ALA:CA	1:D:81:GLU:HB2	1.76	1.15

All (1) 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:C:534:ASP:OD1	1:D:383:GLN:NE2[1_545]	2.09	0.11

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	781/816 (96%)	727 (93%)	44 (6%)	10 (1%)	12	21
1	B	781/816 (96%)	731 (94%)	37 (5%)	13 (2%)	9	16
1	C	749/816 (92%)	700 (94%)	37 (5%)	12 (2%)	9	17
1	D	778/816 (95%)	720 (92%)	45 (6%)	13 (2%)	9	16
All	All	3089/3264 (95%)	2878 (93%)	163 (5%)	48 (2%)	9	17

5 of 48 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	77	GLU
1	A	78	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	81	GLU
1	A	122	ASP
1	B	84	LYS

5.3.2 Protein sidechains ⓘ

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	686/711 (96%)	667 (97%)	19 (3%)	43	70
1	B	686/711 (96%)	672 (98%)	14 (2%)	55	79
1	C	658/711 (92%)	638 (97%)	20 (3%)	41	68
1	D	683/711 (96%)	669 (98%)	14 (2%)	55	79
All	All	2713/2844 (95%)	2646 (98%)	67 (2%)	47	73

5 of 67 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	756	LEU
1	C	137	SER
1	D	311	SER
1	B	814	ASP
1	C	95	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	539	GLN
1	B	678	ASN
1	D	433	GLN
1	B	459	GLN
1	C	730	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

4 non-standard protein/DNA/RNA residues are modelled in this entry.

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$
1	TRQ	C	697	1	13,17,18	1.69	2 (15%)	14,24,26	2.02	4 (28%)
1	TRQ	B	697	1	13,17,18	1.43	1 (7%)	14,24,26	2.32	5 (35%)
1	TRQ	D	697	1	13,17,18	1.26	1 (7%)	14,24,26	2.26	4 (28%)
1	TRQ	A	697	1	13,17,18	1.85	3 (23%)	14,24,26	1.72	3 (21%)

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
1	TRQ	C	697	1	-	0/4/19/21	0/2/2/2
1	TRQ	B	697	1	-	0/4/19/21	0/2/2/2
1	TRQ	D	697	1	-	0/4/19/21	0/2/2/2
1	TRQ	A	697	1	-	1/4/19/21	0/2/2/2

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	697	TRQ	CE2-CZ2	-4.60	1.44	1.50
1	C	697	TRQ	CE2-CZ2	-4.04	1.44	1.50
1	B	697	TRQ	CE2-CZ2	-3.36	1.45	1.50
1	C	697	TRQ	CB-CG	-3.17	1.47	1.51
1	A	697	TRQ	CB-CG	-2.94	1.47	1.51

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	B	697	TRQ	CZ2-CE2-NE1	5.58	128.85	119.94
1	D	697	TRQ	CZ2-CE2-NE1	5.25	128.32	119.94
1	A	697	TRQ	CZ2-CE2-NE1	4.18	126.61	119.94
1	C	697	TRQ	CZ2-CE2-NE1	3.84	126.07	119.94
1	D	697	TRQ	O7-CZ2-CE2	3.79	125.85	121.84

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	697	TRQ	CA-CB-CG-CD1

There are no ring outliers.

4 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	C	697	TRQ	1	0
1	B	697	TRQ	4	0
1	D	697	TRQ	2	0
1	A	697	TRQ	3	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 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, 95th 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 ²)	Q<0.9
1	A	789/816 (96%)	-0.01	44 (5%) 24 25	19, 31, 80, 116	0
1	B	789/816 (96%)	0.06	49 (6%) 20 21	23, 34, 92, 127	0
1	C	759/816 (93%)	-0.08	33 (4%) 35 38	21, 32, 78, 119	0
1	D	786/816 (96%)	0.23	73 (9%) 8 8	23, 37, 99, 131	0
All	All	3123/3264 (95%)	0.05	199 (6%) 19 20	19, 34, 88, 131	0

The worst 5 of 199 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	61	GLY	10.9
1	C	88	ILE	9.1
1	D	83	PHE	8.1
1	A	80	ALA	7.3
1	A	122	ASP	7.3

6.2 Non-standard residues in protein, DNA, RNA chains [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, 95th 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 ²)	Q<0.9
1	TRQ	C	697	16/17	0.96	0.15	21,27,33,33	0
1	TRQ	B	697	16/17	0.96	0.14	25,28,32,33	0
1	TRQ	A	697	16/17	0.96	0.15	22,27,31,33	0
1	TRQ	D	697	16/17	0.97	0.14	23,29,34,35	0

6.3 Carbohydrates [i](#)

There are no carbohydrates 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, 95th 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(\AA^2)	Q<0.9
2	MG	A	901	1/1	0.89	0.14	30,30,30,30	0
3	GLY	C	902	5/5	0.91	0.19	31,40,45,45	0
3	GLY	A	902	5/5	0.93	0.29	39,47,59,59	0
3	GLY	D	902	5/5	0.94	0.21	35,44,53,53	0
2	MG	B	901	1/1	0.95	0.15	32,32,32,32	0
2	MG	D	901	1/1	0.95	0.12	34,34,34,34	0
3	GLY	B	902	5/5	0.97	0.14	26,32,43,43	0
2	MG	C	901	1/1	0.98	0.13	32,32,32,32	0

6.5 Other polymers [i](#)

There are no such residues in this entry.