wwPDB X-ray Structure Validation Summary Report (i)

Oct 5, 2023 – 03:58 PM EDT

PDB ID	:	6VMF
Title	:	Crystal structure of the Y766F mutant of GoxA soaked with glycine
Authors	:	Yukl, E.T.
Deposited on	:	2020-01-27
Resolution	:	2.24 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 (i) 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 (i)) were used in the production of this report:

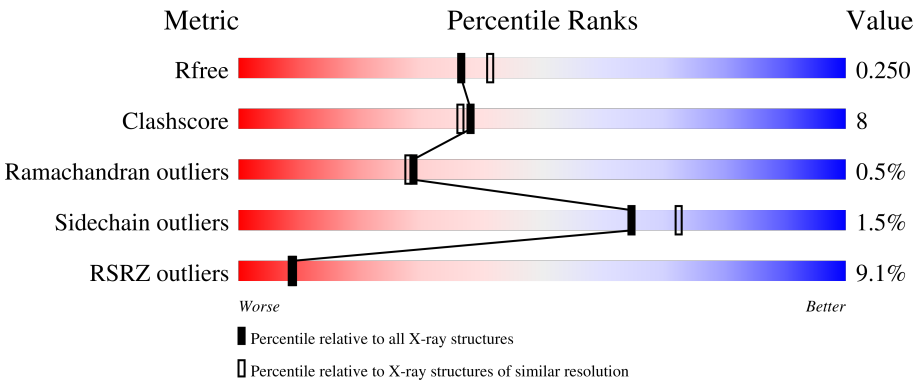
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry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

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.24 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	2391 (2.26-2.22)
Clashscore	141614	2539 (2.26-2.22)
Ramachandran outliers	138981	2489 (2.26-2.22)
Sidechain outliers	138945	2490 (2.26-2.22)
RSRZ outliers	127900	2353 (2.26-2.22)

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 >=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 <=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	6% 78% 14% • 8%
1	B	816	12% 76% 18% • 5%
1	C	816	8% 79% 14% • 6%
1	D	816	8% 76% 19% • 5%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 25629 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 Glycine oxidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	754	6060	3832	1030	1178	20	7	14	0
1	B	773	6202	3920	1054	1207	21	0	11	0
1	C	770	6148	3889	1047	1192	20	0	7	0
1	D	778	6194	3916	1054	1204	20	0	4	0

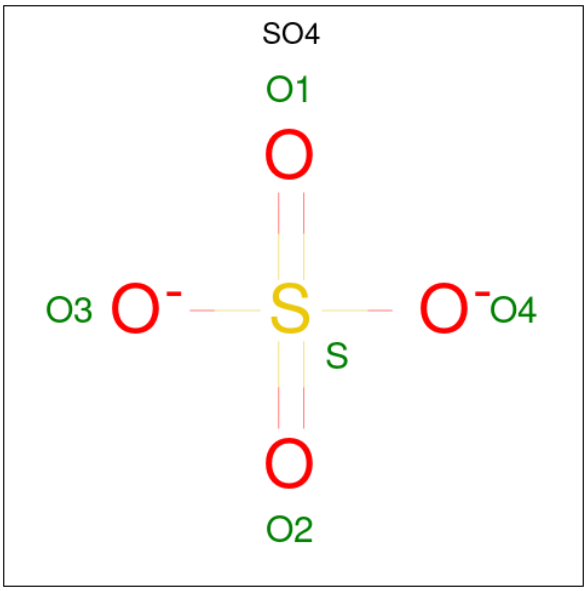
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	766	PHE	TYR	engineered mutation	UNP A0A161XU12
B	766	PHE	TYR	engineered mutation	UNP A0A161XU12
C	766	PHE	TYR	engineered mutation	UNP A0A161XU12
D	766	PHE	TYR	engineered mutation	UNP A0A161XU12

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

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

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	C	1	Total O S 5 4 1	0	0
3	C	1	Total O S 5 4 1	0	0
3	D	1	Total O S 5 4 1	0	0

- Molecule 4 is water.

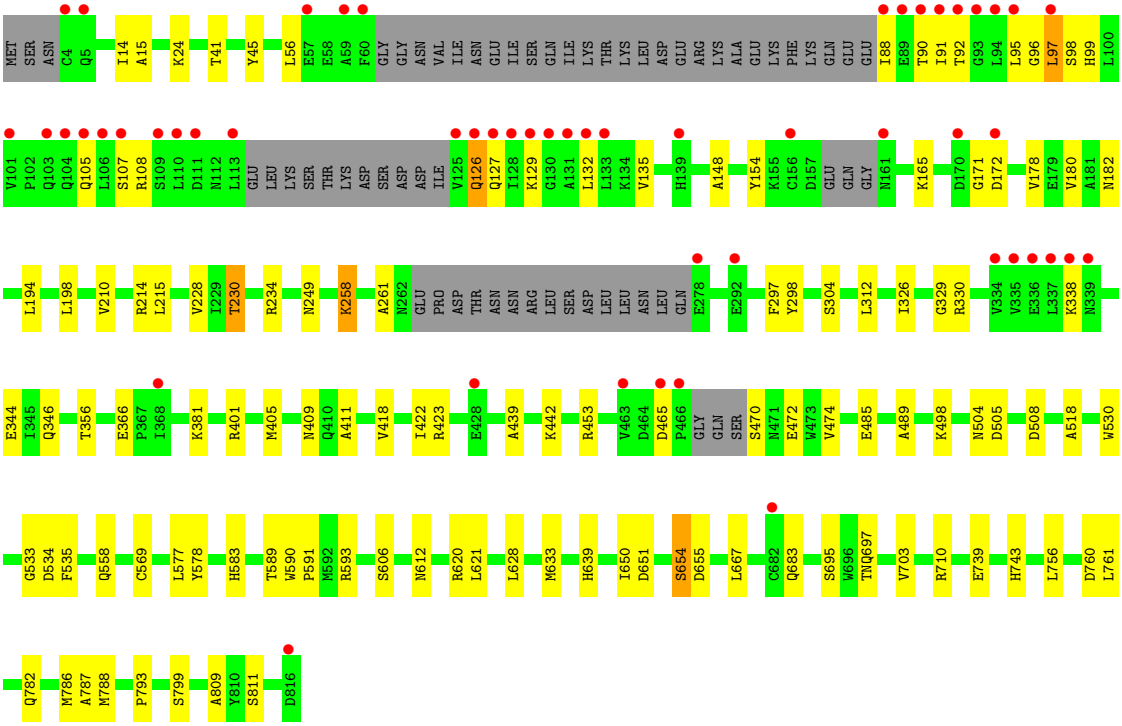
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	293	Total O 293 293	0	0
4	B	208	Total O 208 208	0	0
4	C	258	Total O 258 258	0	0
4	D	247	Total O 247 247	0	0

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.

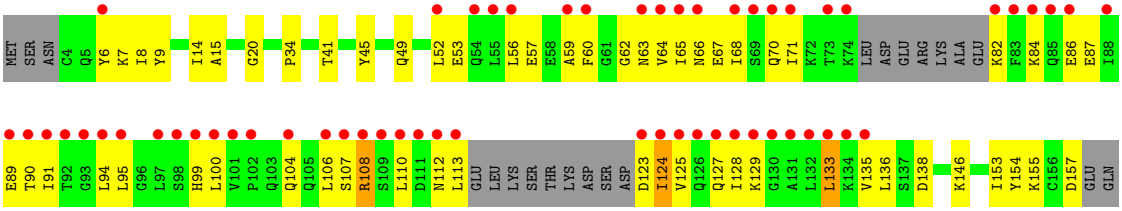
- Molecule 1: Glycine oxidase

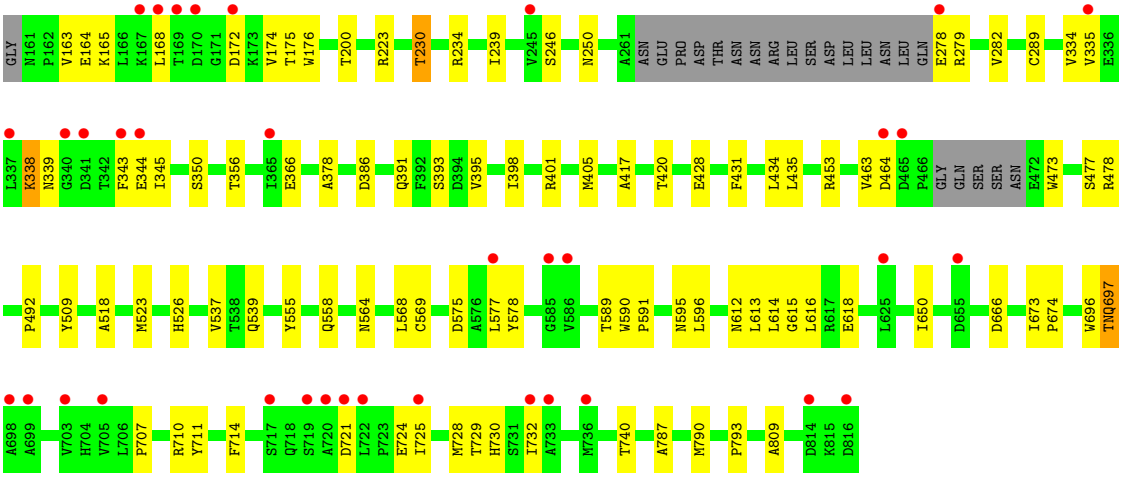
Chain A:



- Molecule 1: Glycine oxidase

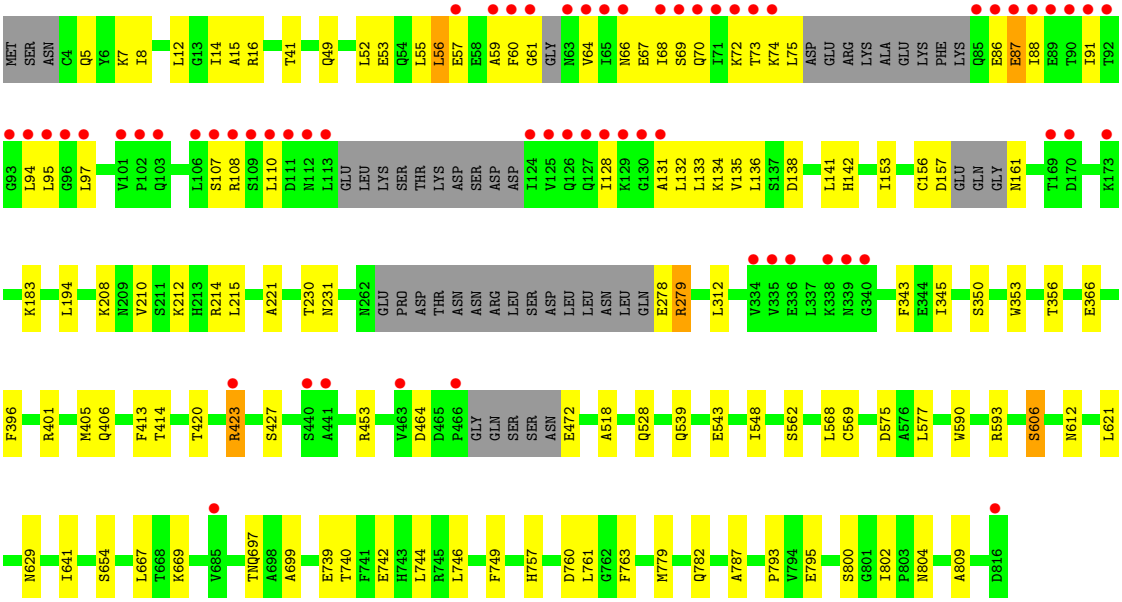
Chain B:





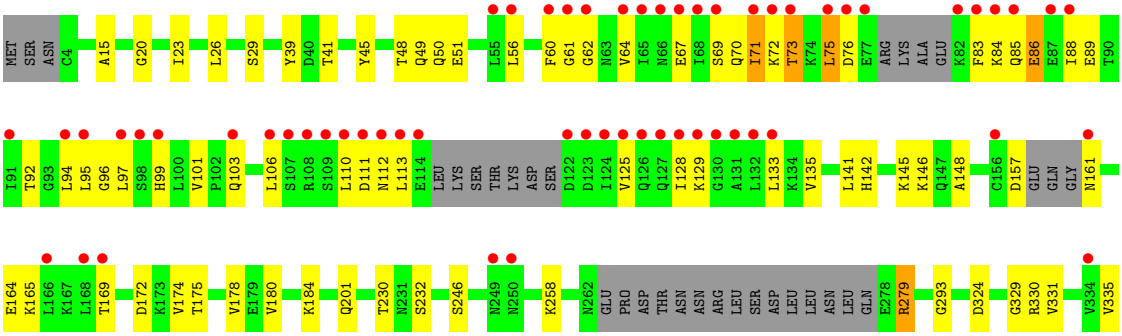
- Molecule 1: Glycine oxidase

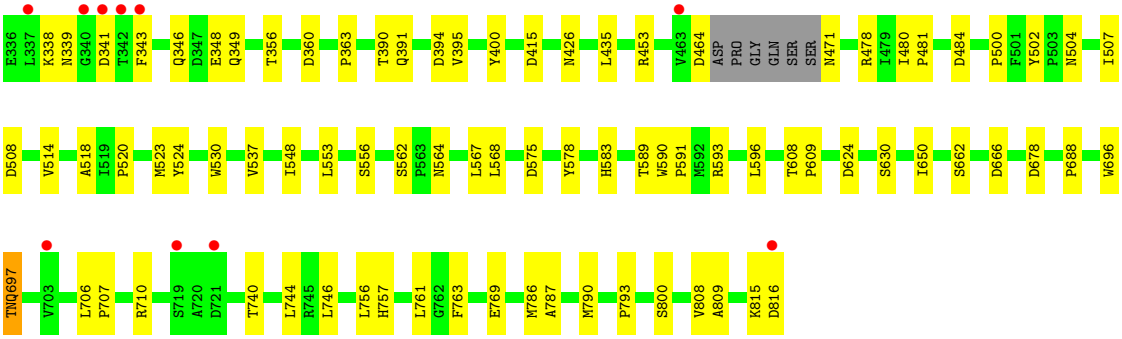
Chain C:



- Molecule 1: Glycine oxidase

Chain D:




E336 L337 K338 N339 G340 D341 T342 F343 Q346 D347 E348 Q349 T356 D360 P363 T390 Q391 D394 V395 Y400 D415 N426 L435 R453 V463 D464 ASP PRO GLY GLN SER SER N471 R478 I479 I480 P481 D484 P500 F501 Y502 P503 N504 I507
D508 V514 A518 I519 P520 M523 Y524 W530 V537 I548 L553 S556 S562 P563 N564 L567 L568 D575 Y578 H583 T589 W590 P591 M592 R593 L596 T608 P609 D624 S630 I650 S662 D666 D678 P688 W696
TNQ697 V703 L706 P707 R710 S719 A720 D721 T740 L744 R745 L746 L756 H757 L761 G762 F763 E769 M786 A787 M790 P793 S800 V808 A809 K815 D816

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	107.86Å 92.98Å 187.05Å 90.00° 95.06° 90.00°	Depositor
Resolution (Å)	48.42 – 2.24 48.42 – 2.24	Depositor EDS
% Data completeness (in resolution range)	98.0 (48.42-2.24) 98.0 (48.42-2.24)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.10 (at 2.24Å)	Xtriage
Refinement program	PHENIX 1.14_3260	Depositor
R, R_{free}	0.202 , 0.250 0.202 , 0.250	Depositor DCC
R_{free} test set	2000 reflections (1.15%)	wwPDB-VP
Wilson B-factor (Å^2)	33.7	Xtriage
Anisotropy	0.757	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 48.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	25629	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.65% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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: SO4, TNQ, 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	0.29	0/6185	0.47	0/8425
1	B	0.29	0/6328	0.49	0/8609
1	C	0.29	0/6272	0.49	2/8536 (0.0%)
1	D	0.27	0/6319	0.48	1/8598 (0.0%)
All	All	0.28	0/25104	0.48	3/34168 (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	D	0	2
All	All	0	3

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	75	LEU	CA-CB-CG	7.45	132.44	115.30
1	C	56	LEU	CA-CB-CG	5.99	129.08	115.30
1	C	87	GLU	CA-CB-CG	5.49	125.48	113.40

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	123	ASP	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	D	71	ILE	Peptide
1	D	73	THR	Peptide

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	6060	0	5805	76	0
1	B	6202	0	5948	119	0
1	C	6148	0	5909	81	0
1	D	6194	0	5949	98	0
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	C	10	0	0	0	0
3	D	5	0	0	0	0
4	A	293	0	0	8	0
4	B	208	0	0	10	0
4	C	258	0	0	5	0
4	D	247	0	0	11	0
All	All	25629	0	23611	367	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 8.

The worst 5 of 367 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:B:104:GLN:HA	1:B:107:SER:HB3	1.43	1.01
1:D:394:ASP:OD2	4:D:1001:HOH:O	1.92	0.87
1:A:95:LEU:HD13	1:A:97:LEU:HG	1.54	0.87
1:B:568:LEU:HD22	1:B:616:LEU:HD11	1.58	0.84
1:C:91:ILE:HA	1:C:94:LEU:HD12	1.61	0.83

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	755/816 (92%)	724 (96%)	26 (3%)	5 (1%)	22	20
1	B	770/816 (94%)	733 (95%)	34 (4%)	3 (0%)	34	35
1	C	762/816 (93%)	733 (96%)	26 (3%)	3 (0%)	34	35
1	D	769/816 (94%)	736 (96%)	30 (4%)	3 (0%)	34	35
All	All	3056/3264 (94%)	2926 (96%)	116 (4%)	14 (0%)	29	28

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	97	LEU
1	A	126	GLN
1	B	124	ILE
1	D	96	GLY
1	A	230	THR

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	668/711 (94%)	658 (98%)	10 (2%)	65	72
1	B	682/711 (96%)	673 (99%)	9 (1%)	69	76
1	C	676/711 (95%)	661 (98%)	15 (2%)	52	59
1	D	680/711 (96%)	673 (99%)	7 (1%)	76	82
All	All	2706/2844 (95%)	2665 (98%)	41 (2%)	65	72

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	575	ASP
1	D	99	HIS
1	C	606[A]	SER
1	C	654	SER
1	D	575	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	734	GLN
1	C	70	GLN
1	D	471	ASN
1	D	112	ASN
1	D	426	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)

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	TNQ	B	697	1	18,21,22	3.14	3 (16%)	16,29,31	2.71	4 (25%)
1	TNQ	A	697	1	18,21,22	3.15	3 (16%)	16,29,31	3.29	3 (18%)
1	TNQ	D	697	1	18,21,22	3.18	3 (16%)	16,29,31	2.85	3 (18%)
1	TNQ	C	697	1	18,21,22	3.15	3 (16%)	16,29,31	3.24	3 (18%)

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	TNQ	B	697	1	-	4/9/11/13	0/2/2/2
1	TNQ	A	697	1	-	4/9/11/13	0/2/2/2
1	TNQ	D	697	1	-	4/9/11/13	0/2/2/2
1	TNQ	C	697	1	-	4/9/11/13	0/2/2/2

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	697	TNQ	C2-N1	-11.96	1.26	1.45
1	C	697	TNQ	C2-N1	-11.86	1.26	1.45
1	A	697	TNQ	C2-N1	-11.81	1.26	1.45
1	B	697	TNQ	C2-N1	-11.79	1.26	1.45
1	D	697	TNQ	CZ2-CE2	3.87	1.49	1.42

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	697	TNQ	C1-C2-N1	8.78	127.07	110.96
1	A	697	TNQ	C2-N1-CH2	-8.54	111.74	123.98
1	C	697	TNQ	C1-C2-N1	8.46	126.48	110.96
1	C	697	TNQ	C2-N1-CH2	-8.42	111.91	123.98
1	D	697	TNQ	C1-C2-N1	7.29	124.34	110.96

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	697	TNQ	CZ2-CH2-N1-C2
1	B	697	TNQ	CZ2-CH2-N1-C2
1	B	697	TNQ	O3-C1-C2-N1
1	C	697	TNQ	CZ2-CH2-N1-C2
1	C	697	TNQ	O2-C1-C2-N1

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	697	TNQ	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	D	697	TNQ	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 7 ligands modelled in this entry, 4 are monoatomic - leaving 3 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	SO4	C	903	-	4,4,4	0.15	0	6,6,6	0.04	0
3	SO4	C	902	-	4,4,4	0.14	0	6,6,6	0.07	0
3	SO4	D	902	-	4,4,4	0.13	0	6,6,6	0.05	0

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	753/816 (92%)	0.26	53 (7%)	16	15	20, 37, 89, 143	0
1	B	772/816 (94%)	0.64	94 (12%)	4	3	26, 54, 123, 169	0
1	C	769/816 (94%)	0.38	63 (8%)	11	11	22, 43, 112, 160	0
1	D	777/816 (95%)	0.42	69 (8%)	9	9	21, 44, 112, 164	0
All	All	3071/3264 (94%)	0.43	279 (9%)	9	9	20, 44, 112, 169	0

The worst 5 of 279 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	61	GLY	8.7
1	C	110	LEU	8.4
1	B	128	ILE	8.4
1	C	65	ILE	8.2
1	A	106	LEU	8.2

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	TNQ	A	697	20/21	0.93	0.18	21,39,48,50	0
1	TNQ	C	697	20/21	0.95	0.15	23,35,48,60	0
1	TNQ	B	697	20/21	0.96	0.16	21,39,57,60	0
1	TNQ	D	697	20/21	0.96	0.17	18,37,48,48	0

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, 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
2	MG	B	901	1/1	0.80	0.15	49,49,49,49	0
2	MG	C	901	1/1	0.85	0.19	40,40,40,40	0
3	SO4	C	903	5/5	0.93	0.25	97,101,106,110	0
3	SO4	C	902	5/5	0.95	0.20	93,95,96,97	0
3	SO4	D	902	5/5	0.95	0.21	74,76,82,97	0
2	MG	D	901	1/1	0.97	0.10	45,45,45,45	0
2	MG	A	901	1/1	0.98	0.18	38,38,38,38	0

6.5 Other polymers (i)

There are no such residues in this entry.