



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 21, 2017 – 09:25 AM EDT

PDB ID : 5XN7
Title : Crystal structure of the effector domain RID of *Vibrio vulnificus* MARTX toxin
Authors : Yin, L.; Zhu, Y.
Deposited on : unknown
Resolution : 2.70 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

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20029824
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20029824

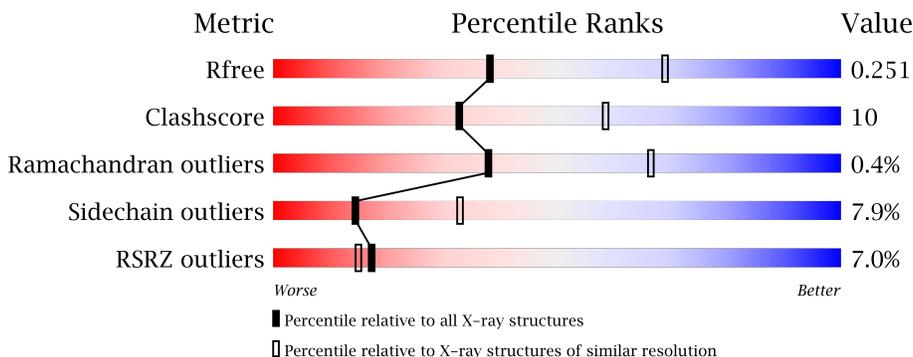
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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}	100719	2259 (2.70-2.70)
Clashscore	112137	2590 (2.70-2.70)
Ramachandran outliers	110173	2550 (2.70-2.70)
Sidechain outliers	110143	2550 (2.70-2.70)
RSRZ outliers	101464	2275 (2.70-2.70)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	616	 7% 66% 20% 12%
1	B	616	 5% 69% 19% 11%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	DIO	A	2902	-	-	-	X
6	CL	A	2906	-	-	X	-
6	CL	A	2907	-	-	-	X

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 8384 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.

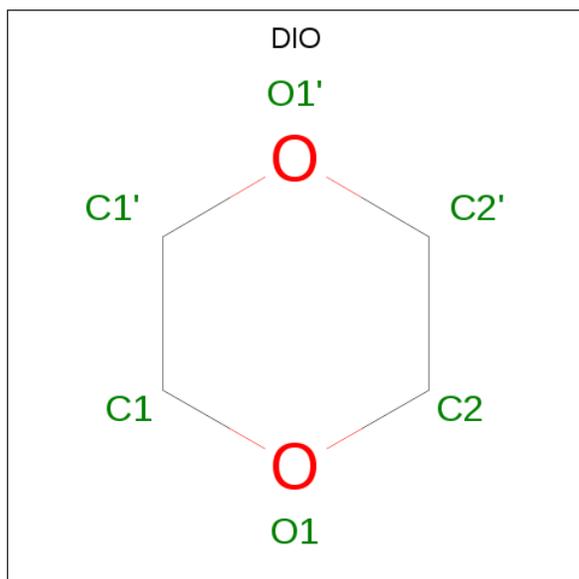
- Molecule 1 is a protein called Putative RTX-toxin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	Se			
1	A	545	Total 4105	C 2576	N 726	O 794	Se 9	0	0	0
1	B	551	Total 4143	C 2597	N 731	O 807	Se 8	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

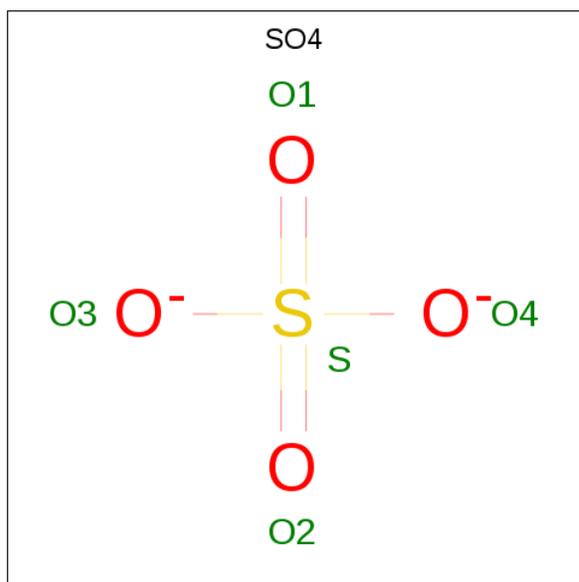
Chain	Residue	Modelled	Actual	Comment	Reference
A	2361	ARG	GLN	conflict	UNP F1CLG1
A	2541	ASP	GLY	conflict	UNP F1CLG1
A	2835	ALA	CYS	engineered mutation	UNP F1CLG1
B	2361	ARG	GLN	conflict	UNP F1CLG1
B	2541	ASP	GLY	conflict	UNP F1CLG1
B	2835	ALA	CYS	engineered mutation	UNP F1CLG1

- Molecule 2 is 1,4-DIETHYLENE DIOXIDE (three-letter code: DIO) (formula: C₄H₈O₂).



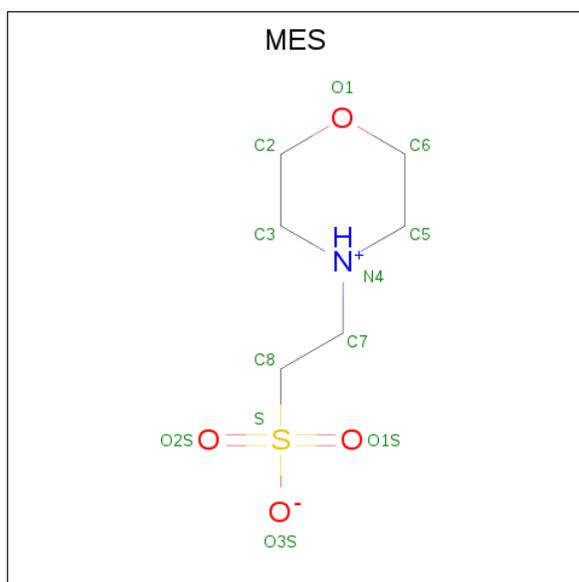
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			6	4	2		
2	A	1	Total	C	O	0	0
			6	4	2		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



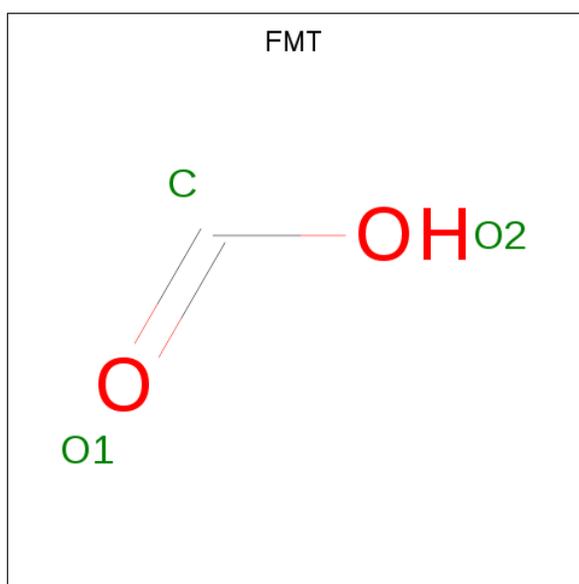
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: C₆H₁₃NO₄S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	N	O	S			
4	A	1	Total	12	6	1	4	1	0	0
4	B	1	Total	12	6	1	4	1	0	0

- Molecule 5 is FORMIC ACID (three-letter code: FMT) (formula: CH₂O₂).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	
			Total	C	O			
5	A	1	Total	3	1	2	0	0

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	1	Total Cl 1 1	0	0
6	A	2	Total Cl 2 2	0	0

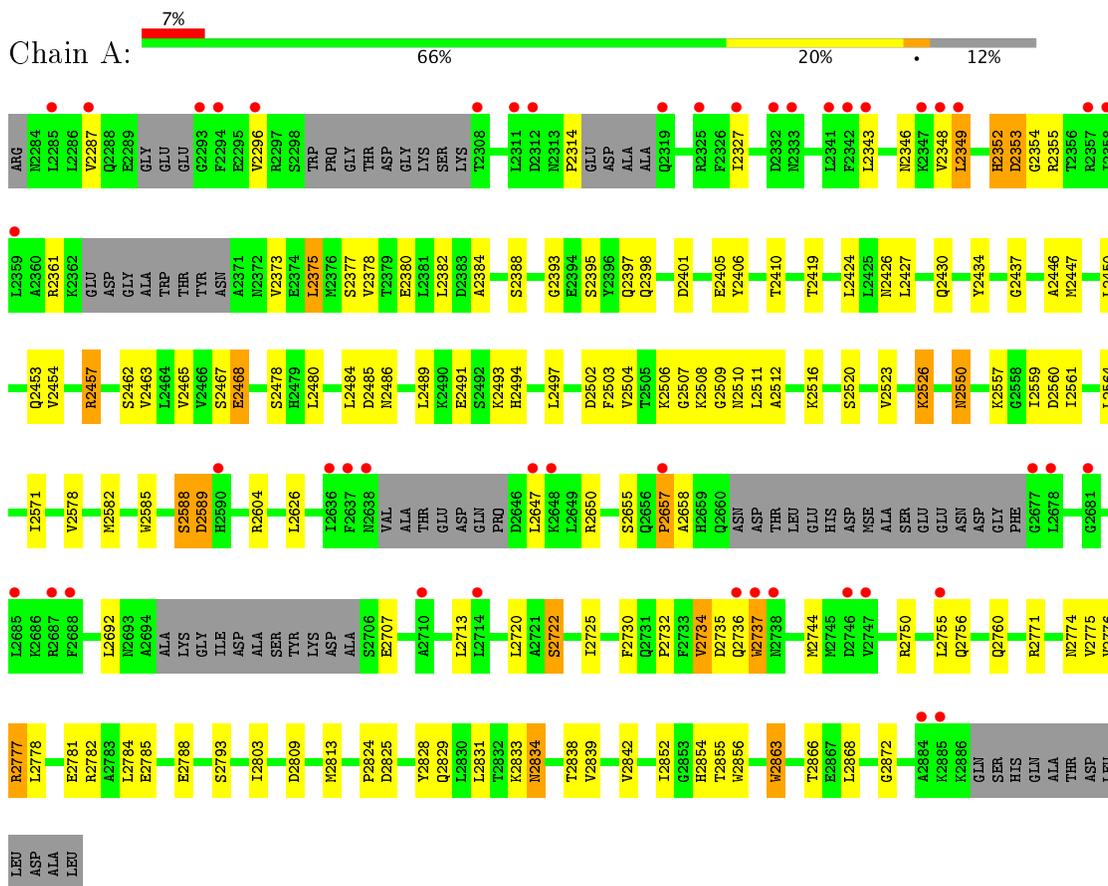
- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	60	Total O 60 60	0	0
7	B	29	Total O 29 29	0	0

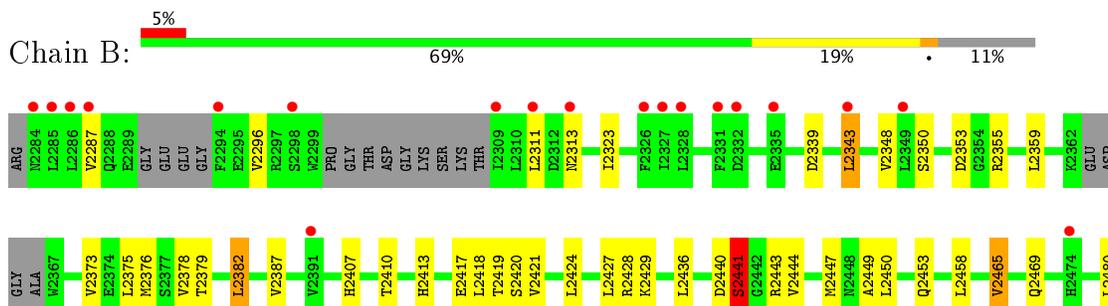
3 Residue-property plots

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: Putative RTX-toxin



- Molecule 1: Putative RTX-toxin



Y2481	Y2482	L2495	Y2496	L2497	V2504	T2505	K2506	Q2507	H2510	LEU	ALA	ASN	ILE	ASP	LYS	LEU	GLY	S2520	L2524	V2527	A2530	I2543	L2552	K2553	K2554	L2564	I2571	M2575	S2576	P2577	A2580	W2585	H2595	L2598	Q2599	I2600	L2607	E2608	G2609	Q2610	N2619				
Y2620	V2621	S2622	W2623	W2624	W2632	I2636	A2640	D2643	Q2644	P2645	D2646	L2649	R2650	A2658	HIS	GLN	ASN	ASP	THR	LEU	GLU	HIS	ASP	MSE	ALA	SER	GLU	GLU	ASN	ASP	GLY	PHE	GLY	L2678	G2681	E2682	V2689	N2693	K2696	GLY	ILE	D2699	V2712	M2716	M2719
S2722	H2728	Q2731	P2732	F2733	N2734	D2735	Q2736	W2737	W2738	Q2739	T2740	D2743	W2744	W2745	A2748	W2749	R2750	F2751	D2764	V2768	V2775	L2784	E2788	R2802	V2810	A2811	A2812	W2813	Q2814	W2817	D2823	F2824	D2825	A2835	T2838	V2839	A2840	I2841	A2848	E2854	T2855	W2856			
R2867	F2868	V2862	W2863	E2867	L2868	L2875	E2881	ILE	ALA	ALA	LYS	LYS	GLN	SER	HIS	GLN	ALA	THR	ASP	LEU	LEU	ASP	ALA	LEU																					

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	124.74Å 185.19Å 81.36Å 90.00° 95.60° 90.00°	Depositor
Resolution (Å)	50.00 – 2.70 46.30 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.5 (50.00-2.70) 99.5 (46.30-2.70)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.82 (at 2.69Å)	Xtrriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.203 , 0.244 0.207 , 0.251	Depositor DCC
R_{free} test set	2554 reflections (5.37%)	DCC
Wilson B-factor (Å ²)	55.5	Xtrriage
Anisotropy	0.803	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 62.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	8384	wwPDB-VP
Average B, all atoms (Å ²)	78.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.38% 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: FMT, DIO, SO4, MES, CL

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	0.64	3/4161 (0.1%)	0.72	2/5619 (0.0%)
1	B	0.59	4/4205 (0.1%)	0.70	8/5689 (0.1%)
All	All	0.61	7/8366 (0.1%)	0.71	10/11308 (0.1%)

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

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	2856	TRP	CD2-CE2	5.97	1.48	1.41
1	B	2856	TRP	CD2-CE2	5.91	1.48	1.41
1	B	2817	TRP	CD2-CE2	5.58	1.48	1.41
1	B	2863	TRP	CD2-CE2	5.29	1.47	1.41
1	A	2863	TRP	CD2-CE2	5.28	1.47	1.41

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	2743	ASP	N-CA-C	8.41	133.70	111.00
1	A	2834	ASN	CB-CA-C	-7.95	94.50	110.40
1	B	2744	MSE	CB-CA-C	6.32	123.04	110.40
1	A	2314	PRO	N-CA-CB	6.11	110.63	103.30
1	B	2745	MSE	N-CA-C	6.00	127.21	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	2588	SER	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	4105	0	3918	80	0
1	B	4143	0	3896	81	0
2	A	12	0	16	0	0
3	A	5	0	0	0	0
4	A	12	0	13	0	0
4	B	12	0	13	0	0
5	A	3	0	1	0	0
6	A	2	0	0	2	0
6	B	1	0	0	0	0
7	A	60	0	0	1	0
7	B	29	0	0	1	0
All	All	8384	0	7857	160	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 10.

The worst 5 of 160 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:2736:GLN:HE22	1:B:2750:ARG:HH22	0.95	0.93
1:B:2736:GLN:HE22	1:B:2750:ARG:NH2	1.70	0.89
1:A:2777:ARG:O	1:A:2781:GLU:HG3	1.85	0.75
1:A:2737:TRP:NE1	1:A:2744:MSE:HG2	2.04	0.72
1:A:2377:SER:OG	1:A:2380:GLU:HG3	1.89	0.72

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	529/616 (86%)	486 (92%)	41 (8%)	2 (0%)	38	66
1	B	537/616 (87%)	498 (93%)	37 (7%)	2 (0%)	38	66
All	All	1066/1232 (86%)	984 (92%)	78 (7%)	4 (0%)	38	66

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2589	ASP
1	A	2657	PRO
1	B	2441	SER
1	B	2737	TRP

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	412/496 (83%)	381 (92%)	31 (8%)	16	36
1	B	412/496 (83%)	378 (92%)	34 (8%)	13	30
All	All	824/992 (83%)	759 (92%)	65 (8%)	14	33

5 of 65 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2855	THR
1	B	2379	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	2813	MSE
1	A	2866	THR
1	B	2353	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	2319	GLN
1	B	2834	ASN
1	B	2756	GLN
1	A	2716	ASN
1	B	2736	GLN

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 3 are monoatomic - leaving 6 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	$\# Z > 2$
2	DIO	A	2901	-	6,6,6	0.61	0	6,6,6	0.92	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	DIO	A	2902	-	6,6,6	0.55	0	6,6,6	0.73	0
3	SO4	A	2903	-	4,4,4	0.46	0	6,6,6	0.14	0
4	MES	A	2904	-	12,12,12	2.03	1 (8%)	14,16,16	1.92	4 (28%)
5	FMT	A	2905	-	0,2,2	0.00	-	0,1,1	0.00	-
4	MES	B	2901	-	12,12,12	2.17	1 (8%)	14,16,16	2.58	5 (35%)

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
2	DIO	A	2901	-	-	0/0/6/6	0/1/1/1
2	DIO	A	2902	-	-	0/0/6/6	0/1/1/1
3	SO4	A	2903	-	-	0/0/0/0	0/0/0/0
4	MES	A	2904	-	-	0/6/14/14	0/1/1/1
5	FMT	A	2905	-	-	0/0/0/0	0/0/0/0
4	MES	B	2901	-	-	0/6/14/14	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	2901	MES	C8-S	-6.99	1.67	1.77
4	A	2904	MES	C8-S	-6.62	1.67	1.77

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	2901	MES	O2S-S-C8	-2.67	104.50	106.79
4	B	2901	MES	O3S-S-O1S	-2.59	105.43	111.37
4	B	2901	MES	O1-C2-C3	-2.07	107.19	111.83
4	A	2904	MES	C2-C3-N4	2.05	112.98	110.11
4	A	2904	MES	O3S-S-C8	2.69	109.36	106.06

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	536/616 (87%)	0.41	45 (8%) 12 9	32, 69, 134, 188	0
1	B	542/616 (87%)	0.33	30 (5%) 26 24	38, 74, 127, 162	0
All	All	1078/1232 (87%)	0.37	75 (6%) 17 15	32, 71, 132, 188	0

The worst 5 of 75 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	2687	ARG	5.4
1	B	2285	LEU	5.4
1	B	2286	LEU	5.3
1	A	2343	LEU	5.3
1	A	2678	LEU	5.2

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 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
2	DIO	A	2902	6/6	0.98	0.24	2.65	63,70,71,75	0
6	CL	A	2907	1/1	0.95	0.23	2.10	63,63,63,63	0
3	SO4	A	2903	5/5	0.93	0.22	1.42	96,100,109,110	0
4	MES	B	2901	12/12	0.94	0.20	-0.25	58,63,72,73	0
6	CL	B	2902	1/1	0.99	0.18	-0.66	45,45,45,45	0
5	FMT	A	2905	3/3	0.95	0.15	-0.94	54,54,60,62	0
6	CL	A	2906	1/1	0.95	0.13	-	92,92,92,92	0
2	DIO	A	2901	6/6	0.85	0.19	-	104,108,115,117	0
4	MES	A	2904	12/12	0.91	0.27	-	106,126,136,138	0

6.5 Other polymers [i](#)

There are no such residues in this entry.