



wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2020 – 12:24 am BST

PDB ID : 3FFZ
Title : Domain organization in Clostridium butulinum neurotoxin type E is unique:
Its implication in faster translocation
Authors : Kumaran, D.; Eswaramoorthy, S.; Swaminathan, S.
Deposited on : 2008-12-04
Resolution : 2.65 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

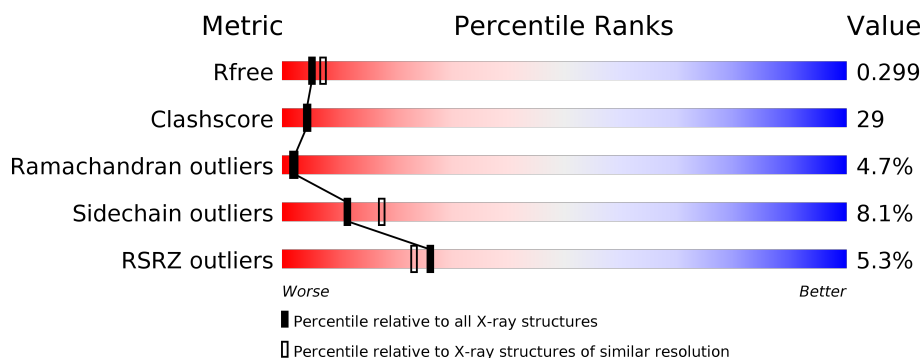
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.65 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	1332 (2.68-2.64)
Clashscore	141614	1374 (2.68-2.64)
Ramachandran outliers	138981	1349 (2.68-2.64)
Sidechain outliers	138945	1349 (2.68-2.64)
RSRZ outliers	127900	1318 (2.68-2.64)

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	1252	<div> <div>4%</div> <div> <div></div> <div>53%</div> <div>39%</div> <div>7%</div> </div> <div>•</div> </div>
1	B	1252	<div> <div>6%</div> <div> <div></div> <div>47%</div> <div>45%</div> <div>6%</div> </div> <div>••</div> </div>

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
3	NA	A	1301	-	-	X	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 20217 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 Botulinum neurotoxin type E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1246	Total	C	N	O	S	0	0	0
			10085	6414	1686	1961	24			
1	B	1238	Total	C	N	O	S	0	0	0
			10025	6375	1674	1952	24			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	177	GLY	ARG	SEE REMARK 999	UNP Q00496
A	340	ALA	ARG	SEE REMARK 999	UNP Q00496
A	963	LEU	PHE	SEE REMARK 999	UNP Q00496
A	964	GLN	GLU	SEE REMARK 999	UNP Q00496
A	967	ALA	ARG	SEE REMARK 999	UNP Q00496
A	1195	ASN	-	INSERTION	UNP Q00496
B	177	GLY	ARG	SEE REMARK 999	UNP Q00496
B	340	ALA	ARG	SEE REMARK 999	UNP Q00496
B	963	LEU	PHE	SEE REMARK 999	UNP Q00496
B	964	GLN	GLU	SEE REMARK 999	UNP Q00496
B	967	ALA	ARG	SEE REMARK 999	UNP Q00496
B	1195	ASN	-	INSERTION	UNP Q00496

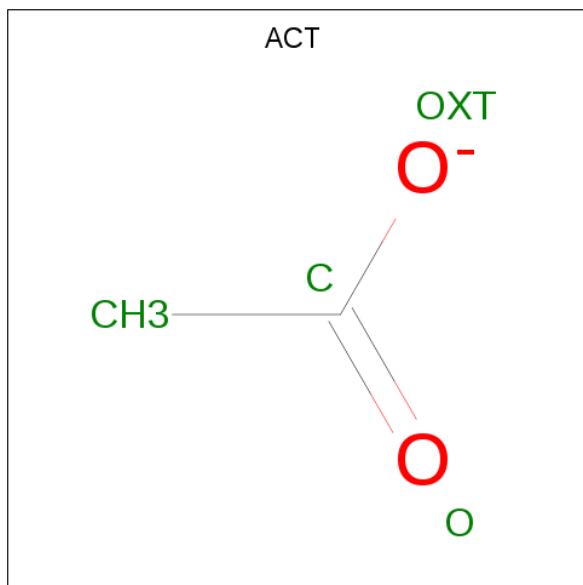
- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Na 1 1	0	0
3	A	2	Total Na 2 2	0	0

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).



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

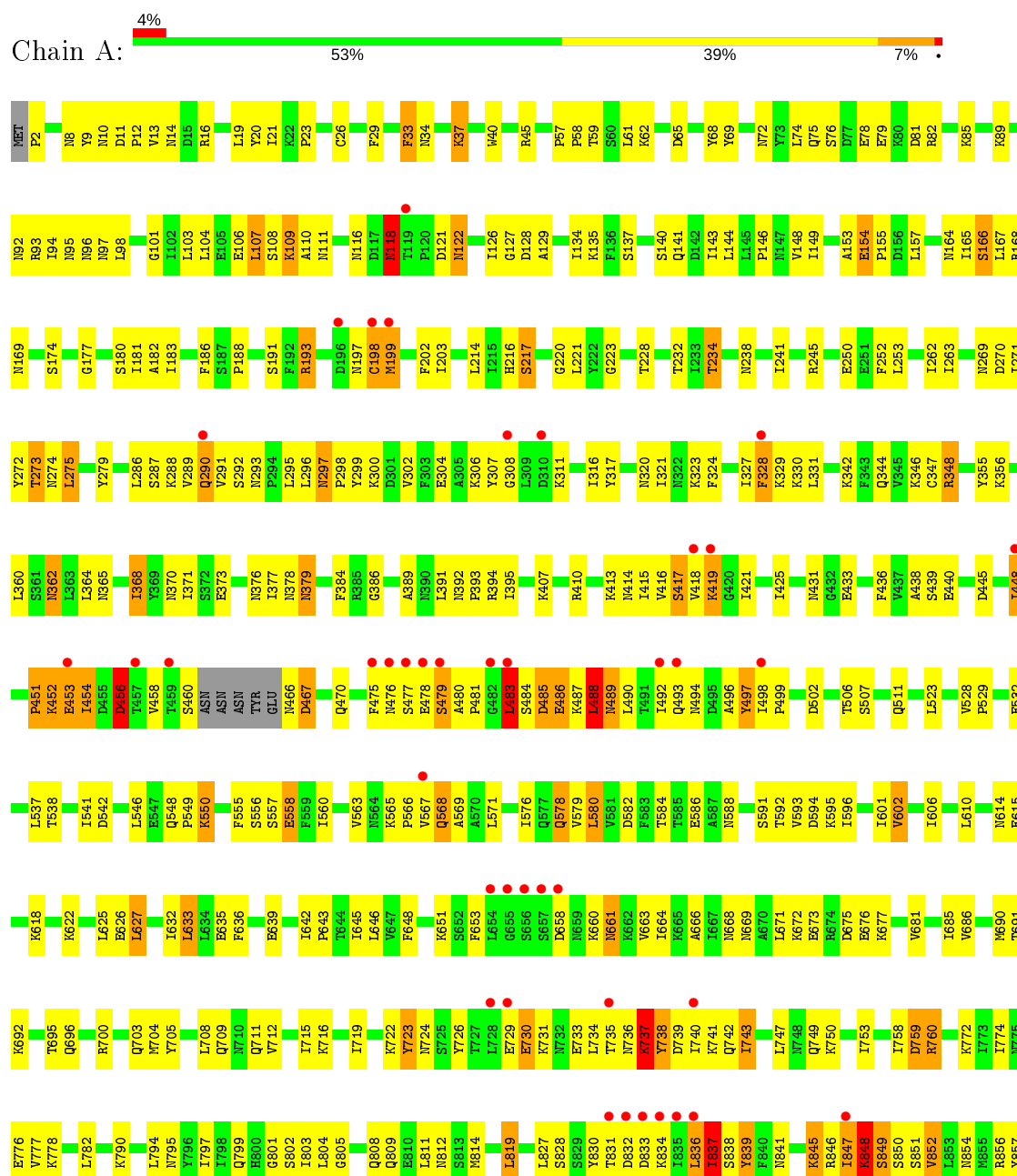
- Molecule 5 is water.

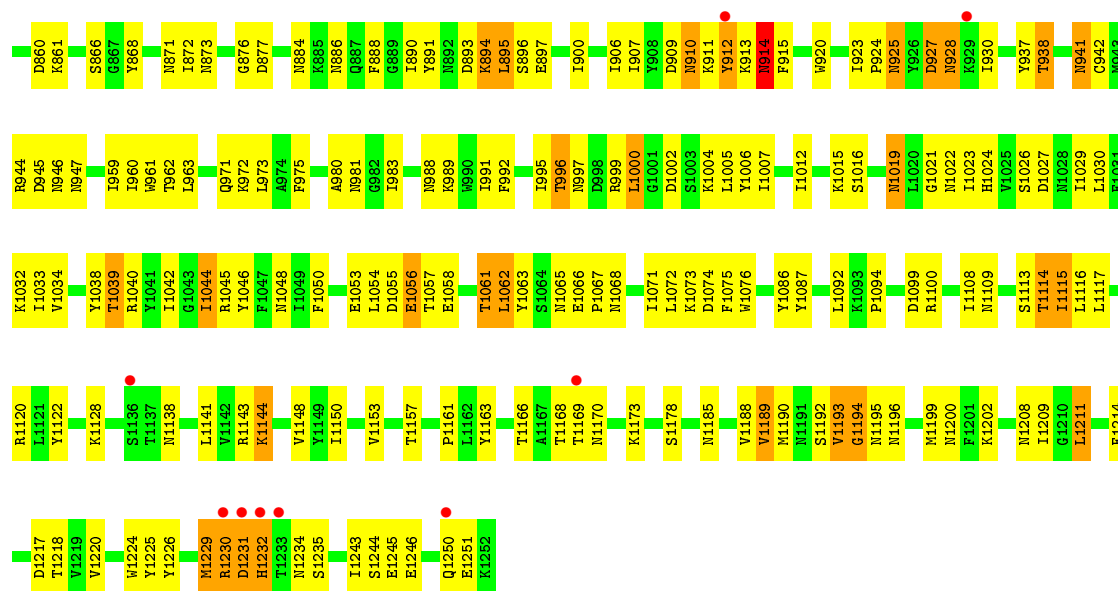
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	62	Total O 62 62	0	0
5	B	32	Total O 32 32	0	0

3 Residue-property plots

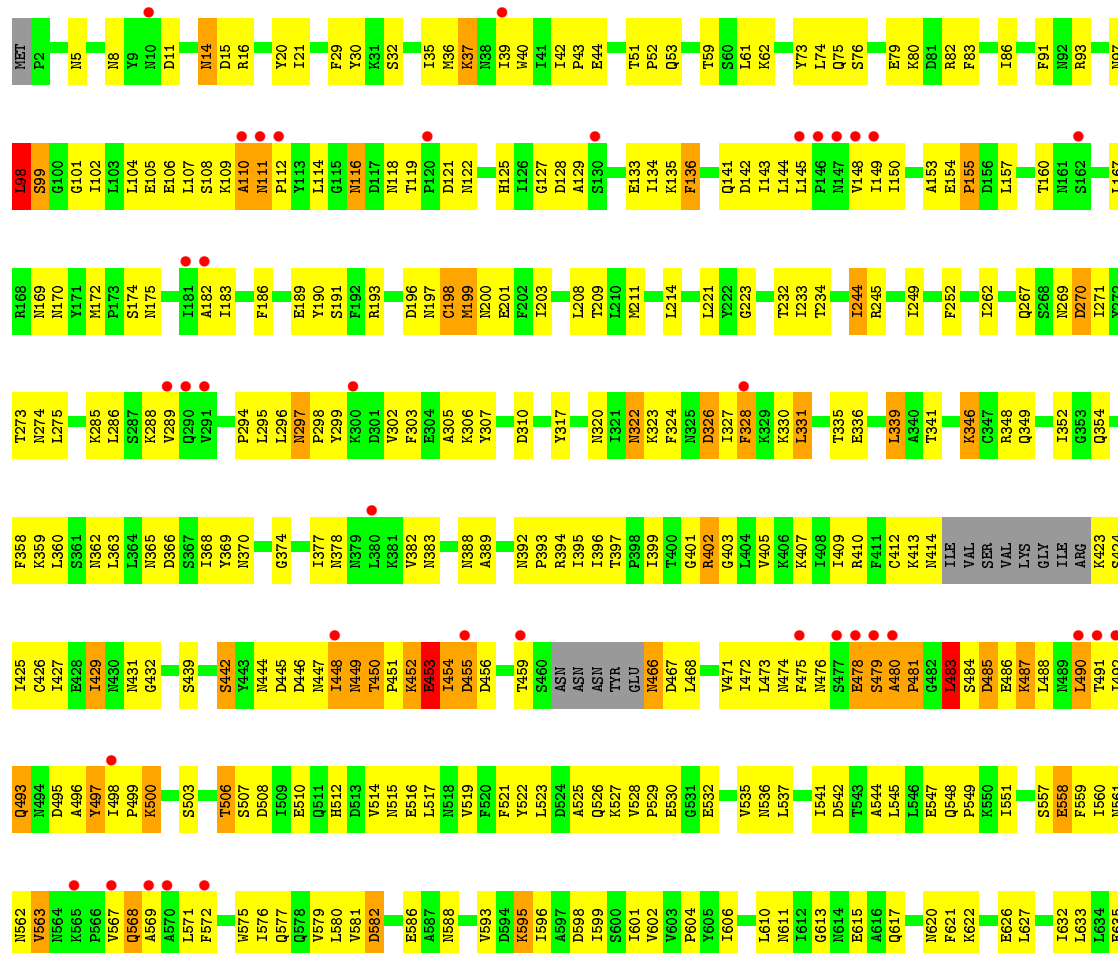
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are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





• Molecule 1: Botulinum neurotoxin type E



Q1250	T1169	P1094	K1015	D860	Y785	N713	F636
E1251	N1170	N1095	S1016	D864	N788	A714	E639
K1252	K1171	N1096	L1020	Y868	Y789	I715	L640
	S1178	F1097	G1021	Y872	L793	I717	L641
	R1183	I1098	N1022	I873	L794	I718	L642
	F1184	D1099	I1023	N874	N795	I719	L646
	N1185	R1100	D1027	I875	Y796	E720	
	Q1186	K1102	N1028	N876	I797	Y723	T649
	V1187	D1103	I1029	G876	I798	N724	I650
	V1188	S1104	L1030	D877	Q799	S725	
	V1189	F1031	F1030	V878	B800	V726	F653
	M1190	K1032	K1032	Y879	G801	T727	L694
	N1191	I1108	I1033	N884	S802	L728	G654
	S1192	F1034	V1034	K885	E729	E729	S656
	V1193	N1035	N1035	N886	L804	E730	S657
	G1194	Y1038	Y1038	Q887	G805	K731	D658
	N1195	T1039	T1039	F888	E806	N732	N659
	M1199	R1040	R1040	Q889	Q809	E733	K660
	N1200	Y1041	Y1041	I890	L734	L734	N661
	F1201	L1116	Q964	Y891	N812	N735	K662
	K1202	L1117	Q965	N892	S813	N736	V663
	N1203	L1121	I966	D893	N814	K737	
	N1204	L1126	A967	K894		Y738	A666
	N1205	K1126	G968	L895	T819	D739	I667
	G1206	Q1130	N969	S896	F825	K741	N668
	N1207	N1049	N970	E897		Q742	N669
	N1208	R1131	Q971	V898	S828	I743	A670
	I1209	V1132	K972	N899	E829	E744	L671
	G1210	N1133	L973	I900	S830	N745	K672
	L1211	N1134	N979	N903	Y830	E746	D675
	K1215	S1135	A980	D904	T831	N747	E676
	A1216	T1137	N981	Y905	D832	N748	K677
		N1138	G982	I906	D833	Q749	W678
		D1139	N983	I907	K834	K750	
	W1224	N1140	Y886	N910	L835	V751	V681
	Y1225	L1141		Y912	L836	S752	T685
	Y1226	V1142		Y912	S838	I753	
	T1227	R1143	Y991	Y912	Y839	A754	T691
	H1228	K1144	T994	N913	F840	N755	K692
	M1229		I995	N914	N841		I693
	R1230	V1146	T996	F915	K842		K694
	D1231	I1149		S916	F843		T695
	H1232	N1151					Q696
	T1233	F1152	R999	F919			
	N1234	V1153	L1000	W920	R846		R700
	S1235		G1001	V921	K848		K701
			D1002	R922	S849		E702
	F1239	T1157	S1003	I923	S850		Q703
	W1240	H1158	K1004	P924	S851		M704
	N1241		L1005	N925	V852		Y705
	F1242	P1161	Y1006	L853	Y926		Q706
	I1243	L1162	I1007	D827	N854		A707
	S1244			N928	N855		
	E1245	T1166	I1012	R929	R856		N710
	E1246	A1167	D1013	I930	Y857		Q711
	W1249	T1168	Q1014				V712

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	81.43Å 172.57Å 137.26Å 90.00° 99.84° 90.00°	Depositor
Resolution (Å)	29.49 – 2.65 33.84 – 2.49	Depositor EDS
% Data completeness (in resolution range)	86.5 (29.49-2.65) 75.9 (33.84-2.49)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.97 (at 2.48Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.253 , 0.309 0.244 , 0.299	Depositor DCC
R_{free} test set	4751 reflections (4.55%)	wwPDB-VP
Wilson B-factor (Å ²)	43.4	Xtriage
Anisotropy	0.475	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 56.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	20217	wwPDB-VP
Average B, all atoms (Å ²)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.24% 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: NA, ZN, ACT

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	0.47	0/10285	0.73	4/13934 (0.0%)
1	B	0.43	1/10224 (0.0%)	0.71	8/13851 (0.1%)
All	All	0.45	1/20509 (0.0%)	0.72	12/27785 (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

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	449	ASN	C-N	6.70	1.49	1.34

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	450	THR	C-N-CD	-14.63	88.41	120.60
1	B	1138	ASN	C-N-CA	6.99	139.17	121.70
1	B	1232	HIS	N-CA-C	6.39	128.25	111.00
1	A	914	ASN	N-CA-C	6.16	127.63	111.00
1	A	483	LEU	N-CA-C	6.13	127.56	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	1138	ASN	Mainchain

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	10085	0	9924	533	0
1	B	10025	0	9849	624	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	2	0	0	2	0
3	B	1	0	0	0	0
4	A	4	0	3	0	0
4	B	4	0	3	0	0
5	A	62	0	0	10	0
5	B	32	0	0	1	0
All	All	20217	0	19779	1158	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 29.

The worst 5 of 1158 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:448:ILE:CG1	1:B:646:LEU:HG	1.80	1.10
1:A:914:ASN:ND2	1:A:914:ASN:H	1.42	1.09
1:A:852:VAL:HG13	1:A:906:ILE:HD12	1.35	1.08
1:A:568:GLN:HG3	1:A:571:LEU:HD22	1.36	1.06
1:B:448:ILE:HG13	1:B:646:LEU:HG	1.34	1.05

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	1242/1252 (99%)	1056 (85%)	128 (10%)	58 (5%)	2	2
1	B	1232/1252 (98%)	1003 (81%)	171 (14%)	58 (5%)	2	2
All	All	2474/2504 (99%)	2059 (83%)	299 (12%)	116 (5%)	2	2

5 of 116 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	153	ALA
1	A	198	CYS
1	A	290	GLN
1	A	418	VAL
1	A	419	LYS

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	1146/1153 (99%)	1050 (92%)	96 (8%)	11	16
1	B	1139/1153 (99%)	1051 (92%)	88 (8%)	13	20
All	All	2285/2306 (99%)	2101 (92%)	184 (8%)	11	17

5 of 184 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1114	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	116	ASN
1	B	1099	ASP
1	A	1128	LYS
1	A	1230	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 124 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1035	ASN
1	B	111	ASN
1	B	1065	ASN
1	A	1048	ASN
1	A	1147	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 7 ligands modelled in this entry, 5 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
4	ACT	B	1303	-	1,3,3	3.07	1 (100%)	0,3,3	0.00	-
4	ACT	A	1303	-	1,3,3	2.26	1 (100%)	0,3,3	0.00	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	1303	ACT	CH3-C	3.07	1.52	1.48
4	A	1303	ACT	CH3-C	2.26	1.51	1.48

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	1246/1252 (99%)	-0.07	50 (4%)	38	34	12, 45, 86, 109	0
1	B	1238/1252 (98%)	0.24	81 (6%)	18	16	24, 62, 96, 109	0
All	All	2484/2504 (99%)	0.09	131 (5%)	26	23	12, 53, 93, 109	0

The worst 5 of 131 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	655	GLY	8.7
1	A	477	SER	8.6
1	B	477	SER	7.9
1	A	655	GLY	7.6
1	A	657	SER	6.6

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. 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
4	ACT	A	1303	4/4	0.75	0.29	70,70,71,71	0
4	ACT	B	1303	4/4	0.81	0.20	65,65,66,66	0
3	NA	A	1301	1/1	0.96	0.08	62,62,62,62	0
3	NA	B	1302	1/1	0.96	0.08	55,55,55,55	0
3	NA	A	1302	1/1	0.98	0.07	38,38,38,38	0
2	ZN	B	1300	1/1	0.99	0.04	47,47,47,47	0
2	ZN	A	1300	1/1	0.99	0.10	36,36,36,36	0

6.5 Other polymers [i](#)

There are no such residues in this entry.