



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:39 AM GMT

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](mailto: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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

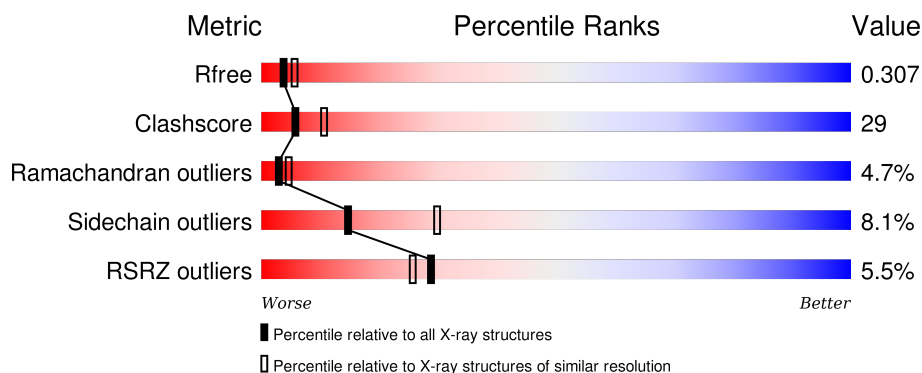
# 1 Overall quality at a glance ⓘ

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))
$R_{free}$	91344	3152 (2.70-2.62)
Clashscore	102246	3524 (2.70-2.62)
Ramachandran outliers	100387	3469 (2.70-2.62)
Sidechain outliers	100360	3469 (2.70-2.62)
RSRZ outliers	91569	3161 (2.70-2.62)

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  $\geq 3$ , 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	1252	<div> <div>4%</div> <div>53%</div> <div>39%</div> <div>7%</div> </div>
1	B	1252	<div> <div>7%</div> <div>47%</div> <div>45%</div> <div>6%</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 [i](#)

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:  $C_2H_3O_2$ ).

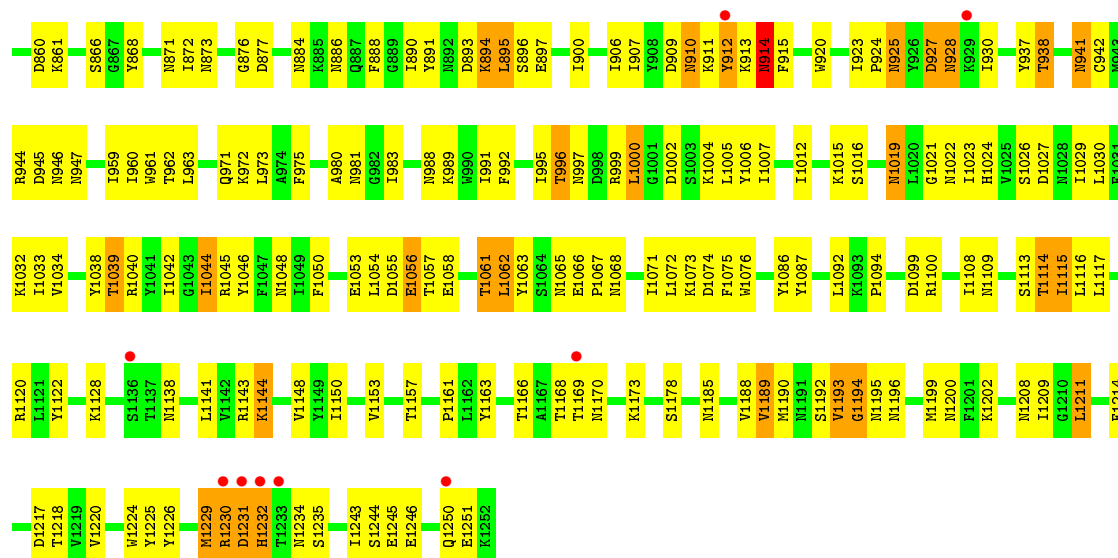


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 4 2 2	0	0
4	B	1	Total C O 4 2 2	0	0

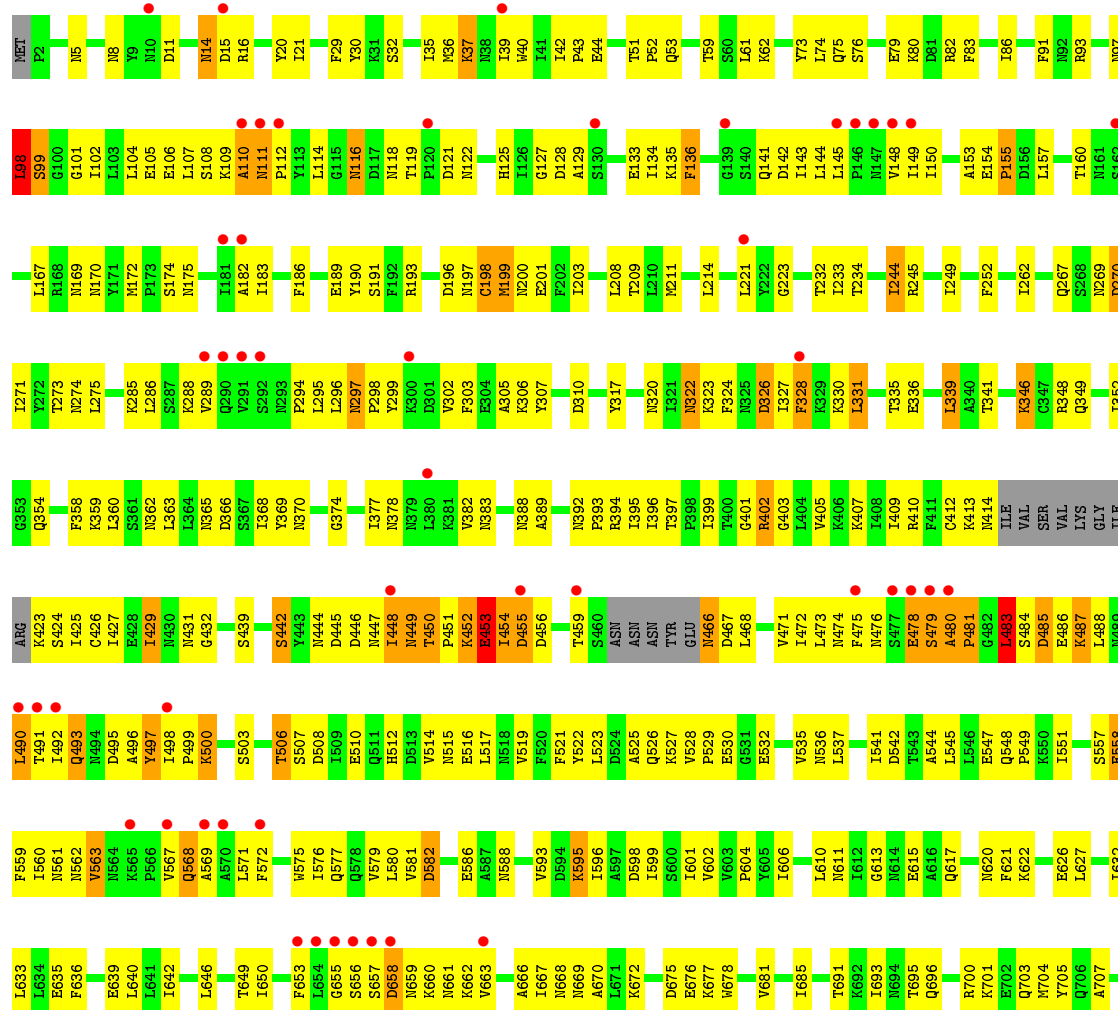
- Molecule 5 is water.

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





• Molecule 1: Botulinum neurotoxin type E



E1246	T1166	N1090	I1012	P929	R856	N710
W1249	A1167	V1091	D1013	I930	Y857	Q711
Q1250	T1168		K1014	V931	L782	V712
E1251	T1169		K1015		Y785	N713
K1252	N1094	P1094	S1016	I939	D860	A714
	N1095	N1095		I940	D864	I715
	N1096	N1096	L1020	R941		K716
	F1097	F1097	G1021	C942	Y868	T717
	I1098	I1098	M1022	M943		I718
	D1099	D1099	I1023	R944	I872	I719
	R1100	R1100	D945	D945	N873	E720
	R1101	R1101	D1027	N946	I874	V723
	K1102	K1102	N1028	N947	N875	N724
	D1103	D1103	I1029	S948	G876	N725
	S1104	S1104	L1030	G949	D877	Y726
			F1031	W950	V878	L727
	S1107	S1107	K1032	L954	Y879	G801
	I1108	I1108	I1033	N955		L728
			V1034		R884	E729
	I1111	I1111	N1035	E988	K885	E730
	R1112	R1112			N886	K731
	T1113	T1113	Y1038	W961	Q887	E732
	T1114	T1114	T1039	T962	F888	E733
	I1115	I1115	R1040	I963	G889	L734
	L1116	L1116	Y1041	Q964	I890	T735
	L1117	L1117	I1042		N812	N736
			G1043	D965	S813	K737
			I1044	N966	D893	Y738
	L1121	L1121	R1045	A967	M814	D739
				G968	T818	I740
	K1126	K1126		I969	L895	K741
			M1048	N970	S897	Q742
	Q1130	Q1130	I1049	N971	F825	I743
	R1131	R1131	F1050	K972	S828	E744
	N1132	N1132	E1053	L973	S829	N745
	N1133	N1133	L1054		Y830	E746
	S1135	S1135	D1055	N979	T831	L747
	S1136	S1136	E1056	A980	D832	N748
	T1137	T1137		N981	D833	Q749
			I1059	G982	K834	K750
				I983	I835	V751
	D1139	D1139	S1064		L836	S752
	N1140	N1140	N1065	Y986	I837	I753
	L1141	L1141	E1066		K911	A754
	V1142	V1142	F1067	I991	S838	N755
	R1143	R1143	M1068		T839	
	K1144	K1144	T1069	T994	N841	I758
			N1070	I995	K842	D759
	Y1148	Y1148	I1071	T996	F843	R760
	Y1149	Y1149	L1072			F761
	I1150	I1150	K1073	R999	R846	L762
	M1151	M1151	D1074	L1000	I847	
			F1075	G1001	S848	S768
	F1152	F1152	M1076	D1002	R921	Y769
				S1003	I923	L770
	T1157	T1157	G1077	K1004	P924	N771
	H1158	H1158	M1078	L1005	N925	K772
				Y1006	D926	
			Y1082	I1007	N927	V777
	P1161	P1161	Y1087		N854	K778
	L1162	L1162			N855	



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	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
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.97 (at 2.48Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.253 , 0.309 0.252 , 0.307	Depositor DCC
$R_{free}$ test set	4751 reflections (5.08%)	DCC
Wilson B-factor (Å <sup>2</sup> )	43.4	Xtriage
Anisotropy	0.475	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 56.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	1 of 104429 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	20217	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	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.*

<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, 0.375 respectively for untwinned datasets, and 0.333, 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: 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%)	3	4
1	B	1232/1252 (98%)	1003 (81%)	171 (14%)	58 (5%)	3	4
All	All	2474/2504 (99%)	2059 (83%)	299 (12%)	116 (5%)	3	4

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%)	14	27
1	B	1139/1153 (99%)	1051 (92%)	88 (8%)	16	33
All	All	2285/2306 (99%)	2101 (92%)	184 (8%)	15	30

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	A	1303	-	1,3,3	2.02	1 (100%)	0,3,3	0.00	-
4	ACT	B	1303	-	1,3,3	2.77	1 (100%)	0,3,3	0.00	-

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
4	ACT	A	1303	-	-	0/0/0/0	0/0/0/0
4	ACT	B	1303	-	-	0/0/0/0	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1303	ACT	CH3-C	2.02	1.51	1.48
4	B	1303	ACT	CH3-C	2.77	1.52	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, 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	1246/1252 (99%)	-0.03	51 (4%) 41 39	12, 45, 86, 109	0
1	B	1238/1252 (98%)	0.27	85 (6%) 20 17	24, 62, 96, 109	0
All	All	2484/2504 (99%)	0.12	136 (5%) 29 26	12, 53, 93, 109	0

The worst 5 of 136 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	655	GLY	8.9
1	A	477	SER	8.8
1	B	477	SER	8.1
1	A	655	GLY	7.6
1	A	657	SER	6.8

### 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 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, 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	NA	A	1301	1/1	0.96	0.08	-1.51	62,62,62,62	0
3	NA	B	1302	1/1	0.97	0.08	-2.45	55,55,55,55	0
2	ZN	A	1300	1/1	0.99	0.10	-2.97	36,36,36,36	0
3	NA	A	1302	1/1	0.98	0.07	-3.91	38,38,38,38	0
4	ACT	B	1303	4/4	0.80	0.20	-	65,65,66,66	0
2	ZN	B	1300	1/1	0.99	0.04	-	47,47,47,47	0
4	ACT	A	1303	4/4	0.75	0.29	-	70,70,71,71	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.