wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 02:34 PM GMT

PDB ID	:	3ZUS
Title	:	Crystal structure of an engineered botulinum neurotoxin type A- SNARE23 derivative, LC-A-SNAP23-Hn-A
Authors	:	Masuyer, G.; Stancombe, P.; Chaddock, J.A.; Acharya, K.R.
Deposited on	:	2011-07-19
Resolution	:	2.95 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

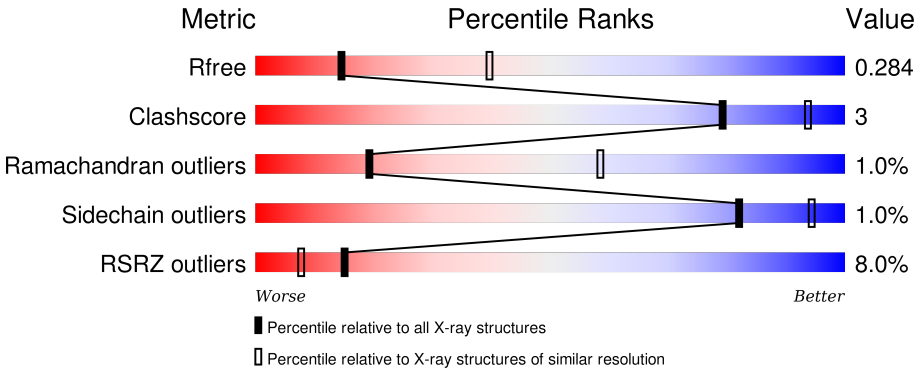
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.95 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	2184 (3.00-2.92)
Clashscore	102246	2552 (3.00-2.92)
Ramachandran outliers	100387	2468 (3.00-2.92)
Sidechain outliers	100360	2471 (3.00-2.92)
RSRZ outliers	91569	2201 (3.00-2.92)

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 <=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	927	4% 84% 7% 8%
1	B	927	6% 84% 7% • 8%
1	C	927	11% 83% 8% 8%
1	D	927	8% 84% 7% • 8%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 27661 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 A, SYNAPTOSOMAL-ASSOCIATED PROTEIN 23.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	851	Total	C	N	O	S	0	0	0
			6889	4427	1112	1329	21			
1	B	851	Total	C	N	O	S	0	1	0
			6898	4432	1113	1332	21			
1	C	852	Total	C	N	O	S	0	0	0
			6898	4432	1113	1332	21			
1	D	853	Total	C	N	O	S	0	0	0
			6902	4436	1114	1331	21			

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	ALA	-	EXPRESSION TAG	UNP P10845
A	1	MET	-	EXPRESSION TAG	UNP P10845
A	2	GLU	-	EXPRESSION TAG	UNP P10845
A	27	ALA	VAL	VARIANT	UNP P10845
A	466	ARG	PRO	SEE REMARK 999	UNP O00161
A	916	LEU	-	EXPRESSION TAG	UNP P10845
A	917	GLU	-	EXPRESSION TAG	UNP P10845
A	918	ALA	-	EXPRESSION TAG	UNP P10845
A	919	HIS	-	EXPRESSION TAG	UNP P10845
A	920	HIS	-	EXPRESSION TAG	UNP P10845
A	921	HIS	-	EXPRESSION TAG	UNP P10845
A	922	HIS	-	EXPRESSION TAG	UNP P10845
A	923	HIS	-	EXPRESSION TAG	UNP P10845
A	924	HIS	-	EXPRESSION TAG	UNP P10845
A	925	HIS	-	EXPRESSION TAG	UNP P10845
A	926	HIS	-	EXPRESSION TAG	UNP P10845
A	927	HIS	-	EXPRESSION TAG	UNP P10845
A	928	HIS	-	EXPRESSION TAG	UNP P10845
B	0	ALA	-	EXPRESSION TAG	UNP P10845
B	1	MET	-	EXPRESSION TAG	UNP P10845

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	2	GLU	-	EXPRESSION TAG	UNP P10845
B	27	ALA	VAL	VARIANT	UNP P10845
B	466	ARG	PRO	SEE REMARK 999	UNP O00161
B	916	LEU	-	EXPRESSION TAG	UNP P10845
B	917	GLU	-	EXPRESSION TAG	UNP P10845
B	918	ALA	-	EXPRESSION TAG	UNP P10845
B	919	HIS	-	EXPRESSION TAG	UNP P10845
B	920	HIS	-	EXPRESSION TAG	UNP P10845
B	921	HIS	-	EXPRESSION TAG	UNP P10845
B	922	HIS	-	EXPRESSION TAG	UNP P10845
B	923	HIS	-	EXPRESSION TAG	UNP P10845
B	924	HIS	-	EXPRESSION TAG	UNP P10845
B	925	HIS	-	EXPRESSION TAG	UNP P10845
B	926	HIS	-	EXPRESSION TAG	UNP P10845
B	927	HIS	-	EXPRESSION TAG	UNP P10845
B	928	HIS	-	EXPRESSION TAG	UNP P10845
C	0	ALA	-	EXPRESSION TAG	UNP P10845
C	1	MET	-	EXPRESSION TAG	UNP P10845
C	2	GLU	-	EXPRESSION TAG	UNP P10845
C	27	ALA	VAL	VARIANT	UNP P10845
C	466	ARG	PRO	SEE REMARK 999	UNP O00161
C	916	LEU	-	EXPRESSION TAG	UNP P10845
C	917	GLU	-	EXPRESSION TAG	UNP P10845
C	918	ALA	-	EXPRESSION TAG	UNP P10845
C	919	HIS	-	EXPRESSION TAG	UNP P10845
C	920	HIS	-	EXPRESSION TAG	UNP P10845
C	921	HIS	-	EXPRESSION TAG	UNP P10845
C	922	HIS	-	EXPRESSION TAG	UNP P10845
C	923	HIS	-	EXPRESSION TAG	UNP P10845
C	924	HIS	-	EXPRESSION TAG	UNP P10845
C	925	HIS	-	EXPRESSION TAG	UNP P10845
C	926	HIS	-	EXPRESSION TAG	UNP P10845
C	927	HIS	-	EXPRESSION TAG	UNP P10845
C	928	HIS	-	EXPRESSION TAG	UNP P10845
D	0	ALA	-	EXPRESSION TAG	UNP P10845
D	1	MET	-	EXPRESSION TAG	UNP P10845
D	2	GLU	-	EXPRESSION TAG	UNP P10845
D	27	ALA	VAL	VARIANT	UNP P10845
D	466	ARG	PRO	SEE REMARK 999	UNP O00161
D	916	LEU	-	EXPRESSION TAG	UNP P10845
D	917	GLU	-	EXPRESSION TAG	UNP P10845
D	918	ALA	-	EXPRESSION TAG	UNP P10845

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	919	HIS	-	EXPRESSION TAG	UNP P10845
D	920	HIS	-	EXPRESSION TAG	UNP P10845
D	921	HIS	-	EXPRESSION TAG	UNP P10845
D	922	HIS	-	EXPRESSION TAG	UNP P10845
D	923	HIS	-	EXPRESSION TAG	UNP P10845
D	924	HIS	-	EXPRESSION TAG	UNP P10845
D	925	HIS	-	EXPRESSION TAG	UNP P10845
D	926	HIS	-	EXPRESSION TAG	UNP P10845
D	927	HIS	-	EXPRESSION TAG	UNP P10845
D	928	HIS	-	EXPRESSION TAG	UNP P10845

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

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

- Molecule 3 is water.

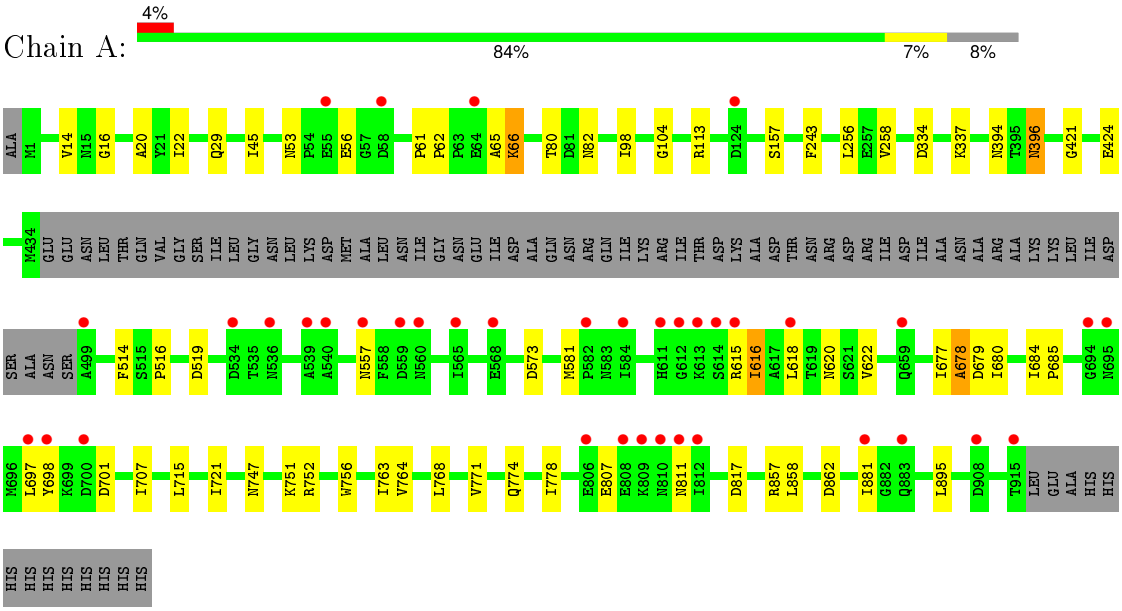
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	27	Total O 27 27	0	0
3	B	21	Total O 21 21	0	0
3	C	9	Total O 9 9	0	0
3	D	13	Total O 13 13	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

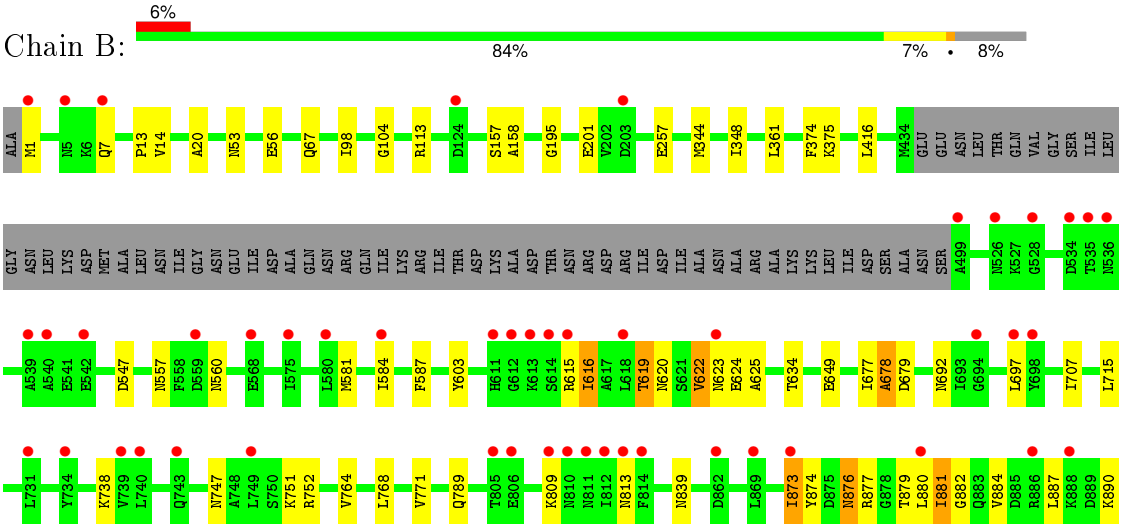
• Molecule 1: BOTULINUM NEUROTOXIN TYPE A, SYNAPTOSOMAL-ASSOCIATED PROTEIN 23

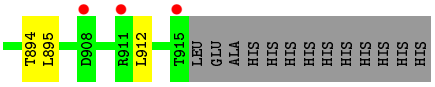
Chain A:



• Molecule 1: BOTULINUM NEUROTOXIN TYPE A, SYNAPTOSOMAL-ASSOCIATED PROTEIN 23

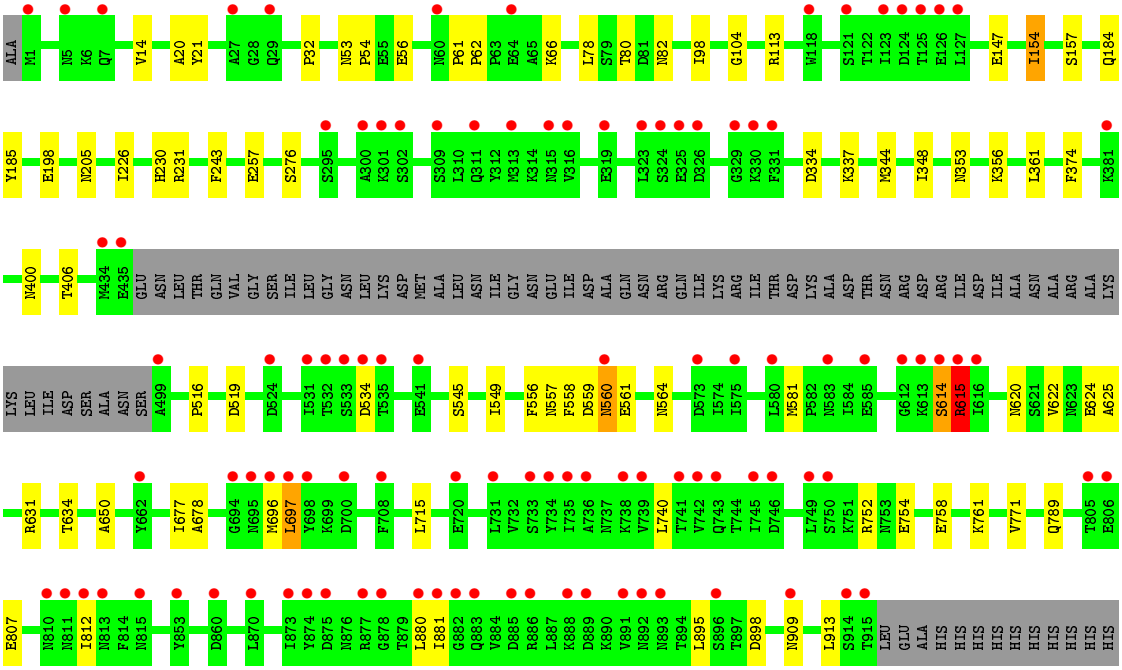
Chain B:





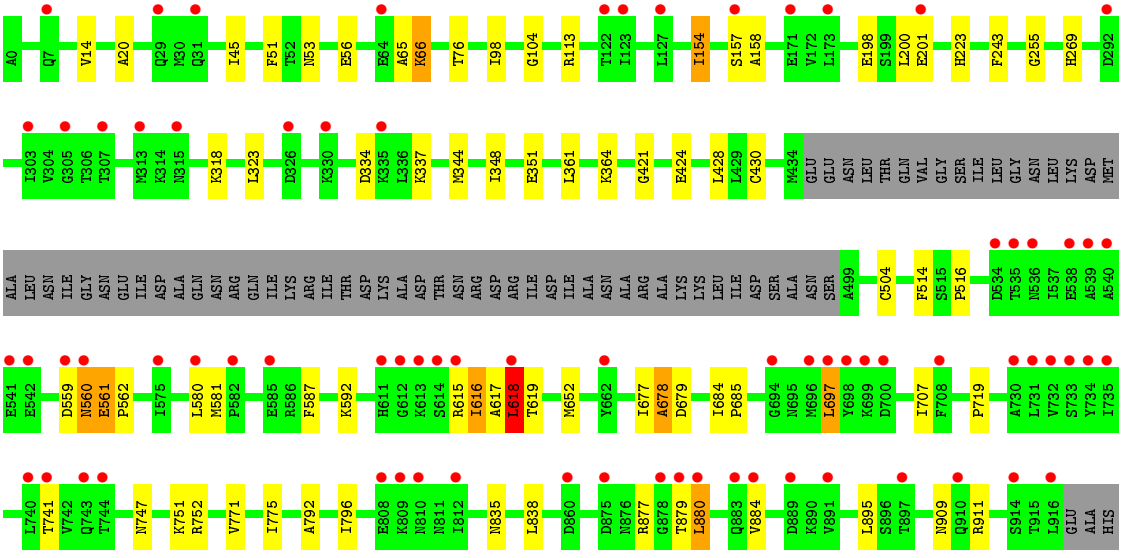
• Molecule 1: BOTULINUM NEUROTOXIN TYPE A, SYNAPTOSOMAL-ASSOCIATED PROTEIN 23

Chain C:



• Molecule 1: BOTULINUM NEUROTOXIN TYPE A, SYNAPTOSOMAL-ASSOCIATED PROTEIN 23

Chain D:



HIS
HIS
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	89.18Å 204.97Å 130.88Å 90.00° 91.91° 90.00°	Depositor
Resolution (Å)	130.81 – 2.95 46.31 – 2.95	Depositor EDS
% Data completeness (in resolution range)	99.3 (130.81-2.95) 99.4 (46.31-2.95)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.62 (at 2.96Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.248 , 0.293 0.243 , 0.284	Depositor DCC
R_{free} test set	4885 reflections (5.26%)	DCC
Wilson B-factor (Å^2)	51.7	Xtriage
Anisotropy	0.336	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 37.5	EDS
Estimated twinning fraction	0.036 for h,-k,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Outliers	0 of 97778 reflections	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	27661	wwPDB-VP
Average B, all atoms (Å^2)	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.90% 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, 0.375 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: ZN

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	0.32	0/7033	0.45	0/9526
1	B	0.33	0/7042	0.46	0/9538
1	C	0.32	0/7042	0.45	0/9538
1	D	0.33	0/7046	0.45	0/9544
All	All	0.33	0/28163	0.45	0/38146

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	6889	0	6799	36	0
1	B	6898	0	6804	59	0
1	C	6898	0	6805	45	0
1	D	6902	0	6817	51	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	A	27	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	21	0	0	0	0
3	C	9	0	0	0	0
3	D	13	0	0	0	0
All	All	27661	0	27225	189	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 3.

The worst 5 of 189 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:874:TYR:CD2	1:B:877:ARG:NH2	1.94	1.33
1:B:874:TYR:CE2	1:B:877:ARG:NH2	1.99	1.29
1:B:880:LEU:O	1:B:881:ILE:HG22	1.55	1.07
1:D:255:GLY:HA3	1:D:587:PHE:CD1	1.91	1.05
1:C:560:ASN:O	1:C:561:GLU:HG3	1.61	1.00

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	847/927 (91%)	812 (96%)	27 (3%)	8 (1%)	21	61
1	B	848/927 (92%)	812 (96%)	27 (3%)	9 (1%)	17	56
1	C	848/927 (92%)	812 (96%)	28 (3%)	8 (1%)	21	61
1	D	849/927 (92%)	804 (95%)	35 (4%)	10 (1%)	16	53
All	All	3392/3708 (92%)	3240 (96%)	117 (3%)	35 (1%)	19	58

5 of 35 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	616	ILE
1	B	616	ILE
1	C	615	ARG
1	D	618	LEU
1	D	619	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	767/831 (92%)	761 (99%)	6 (1%)	86 95
1	B	768/831 (92%)	762 (99%)	6 (1%)	86 95
1	C	768/831 (92%)	759 (99%)	9 (1%)	78 93
1	D	768/831 (92%)	757 (99%)	11 (1%)	74 92
All	All	3071/3324 (92%)	3039 (99%)	32 (1%)	82 94

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	519	ASP
1	C	697	LEU
1	D	741	THR
1	C	615	ARG
1	C	740	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	743	GLN
1	C	205	ASN
1	D	743	GLN
1	C	29	GLN
1	C	503	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 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

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

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^{th} 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	851/927 (91%)	0.24	38 (4%) 37 21	20, 44, 72, 85	0
1	B	851/927 (91%)	0.29	52 (6%) 25 13	23, 47, 93, 118	0
1	C	852/927 (91%)	0.63	106 (12%) 5 2	25, 59, 117, 140	0
1	D	853/927 (92%)	0.42	75 (8%) 12 6	23, 52, 97, 118	0
All	All	3407/3708 (91%)	0.39	271 (7%) 15 8	20, 50, 100, 140	0

The worst 5 of 271 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	885	ASP	8.2
1	C	736	ALA	8.2
1	A	614	SER	6.9
1	C	499	ALA	6.8
1	B	811	ASN	6.6

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

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^{th} 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(Å^2)	Q<0.9
2	ZN	D	1917	1/1	0.95	0.16	-	83,83,83,83	0
2	ZN	C	1916	1/1	0.97	0.14	-	45,45,45,45	0
2	ZN	A	1916	1/1	0.99	0.13	-	38,38,38,38	0
2	ZN	B	1916	1/1	0.94	0.17	-	74,74,74,74	0

6.5 Other polymers (i)

There are no such residues in this entry.