wwPDB X-ray Structure Validation Summary Report (i)

May 25, 2020 – 05:55 pm BST

PDB ID	:	2AQ3
Title	:	Crystal structure of T-cell receptor V beta domain variant complexed with superantigen SEC3
Authors	:	Cho, S.; Swaminathan, C.P.; Yang, J.; Kerzic, M.C.; Guan, R.; Kieke, M.C.; Kranz, D.M.; Mariuzza, R.A.; Sundberg, E.J.
Deposited on	:	2005-08-17
Resolution	:	2.30 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 following versions of software and data (see references (1)) were used in the production of this report:

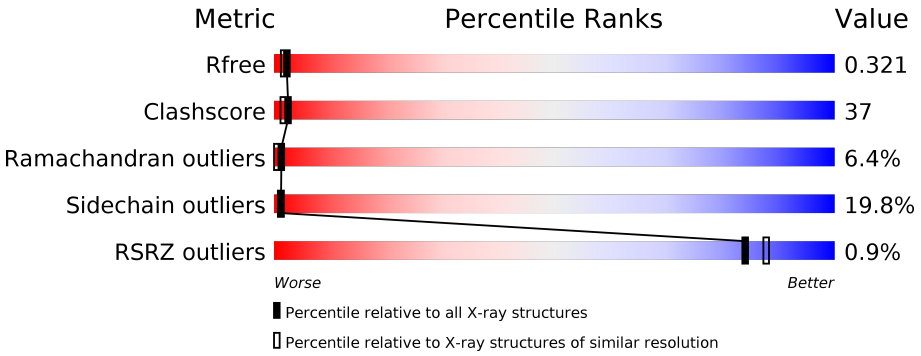
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.30 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.



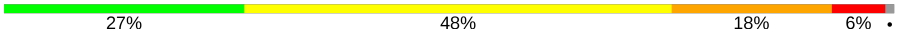
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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 <=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	112	% 25% 39% 29%
1	C	112	% 27% 33% 31% 6%
1	E	112	% 40% 38% 18%
1	G	112	29% 48% 17%
2	B	237	34% 37% 23% 5%
2	D	237	% 30% 48% 19%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	237	
2	H	237	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 11155 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 T-cell receptor beta chain V.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	109	Total	C	N	O	S	0	0	0
			828	513	147	165	3			
1	C	109	Total	C	N	O	S	0	0	0
			828	513	147	165	3			
1	E	109	Total	C	N	O	S	0	0	0
			828	513	147	165	3			
1	G	109	Total	C	N	O	S	0	0	0
			828	513	147	165	3			

- Molecule 2 is a protein called Enterotoxin type C-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	234	Total	C	N	O	S	0	0	0
			1904	1207	312	375	10			
2	D	235	Total	C	N	O	S	0	0	0
			1913	1211	314	378	10			
2	F	234	Total	C	N	O	S	0	0	0
			1906	1205	312	379	10			
2	H	237	Total	C	N	O	S	0	0	0
			1922	1215	315	382	10			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	ASN	DELETION	UNP P0A0L5
B	?	-	VAL	DELETION	UNP P0A0L5
D	?	-	ASN	DELETION	UNP P0A0L5
D	?	-	VAL	DELETION	UNP P0A0L5
F	?	-	ASN	DELETION	UNP P0A0L5
F	?	-	VAL	DELETION	UNP P0A0L5
H	?	-	ASN	DELETION	UNP P0A0L5
H	?	-	VAL	DELETION	UNP P0A0L5

- Molecule 3 is water.

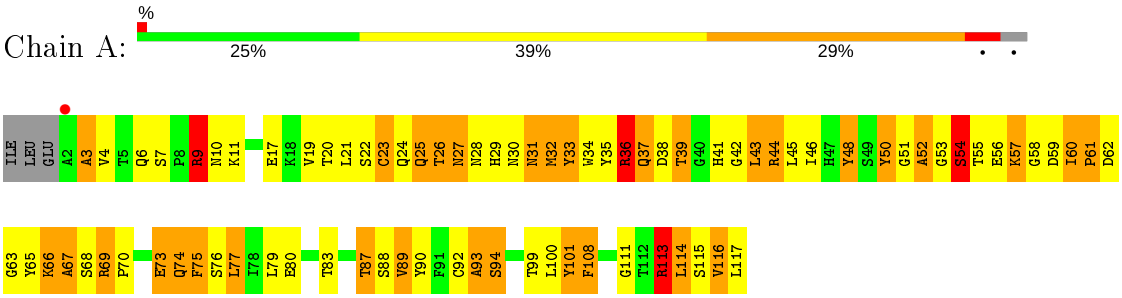
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	21	Total O 21 21	0	0
3	B	48	Total O 48 48	0	0
3	C	15	Total O 15 15	0	0
3	D	31	Total O 31 31	0	0
3	E	11	Total O 11 11	0	0
3	F	31	Total O 31 31	0	0
3	G	15	Total O 15 15	0	0
3	H	26	Total O 26 26	0	0

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 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.

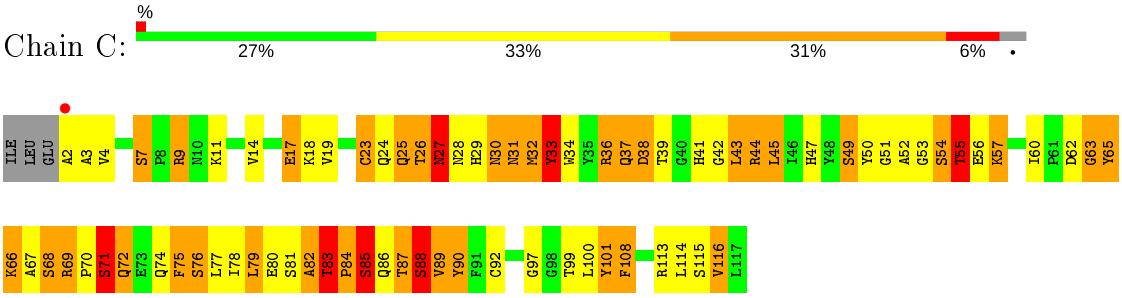
- Molecule 1: T-cell receptor beta chain V

Chain A:



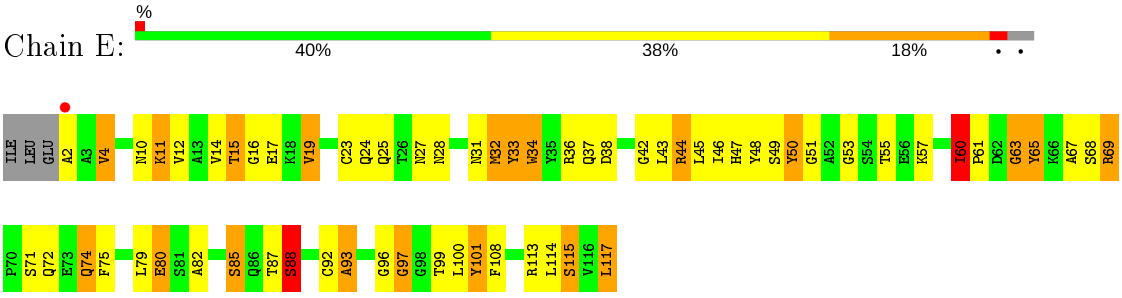
- Molecule 1: T-cell receptor beta chain V

Chain C:



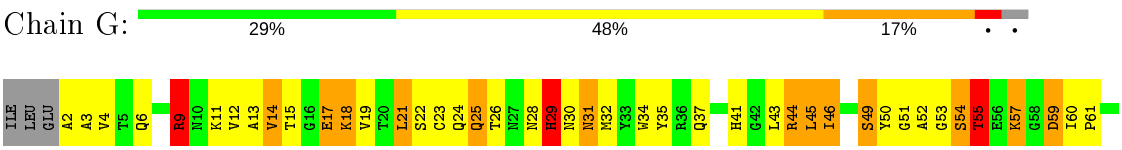
- Molecule 1: T-cell receptor beta chain V

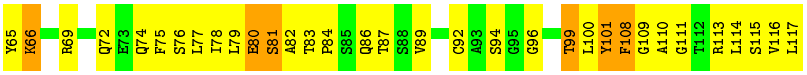
Chain E:



- Molecule 1: T-cell receptor beta chain V

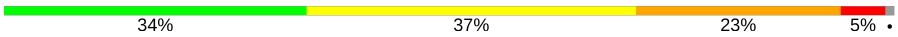
Chain G:

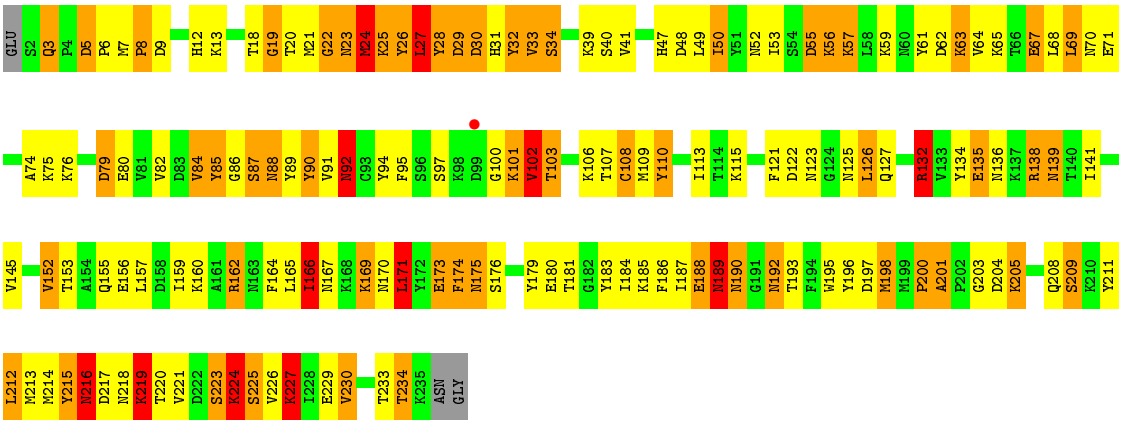




• Molecule 2: Enterotoxin type C-3

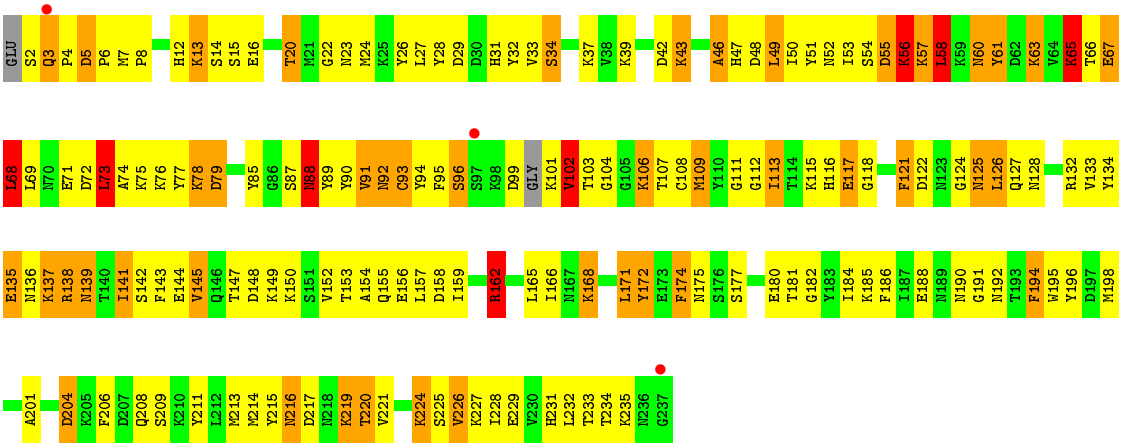
Chain B:





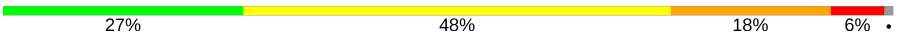
• Molecule 2: Enterotoxin type C-3

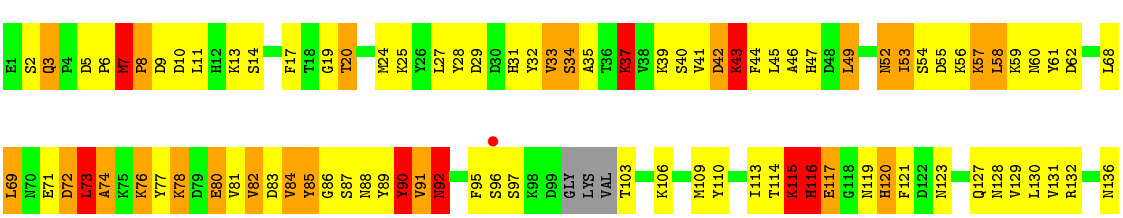
Chain D:

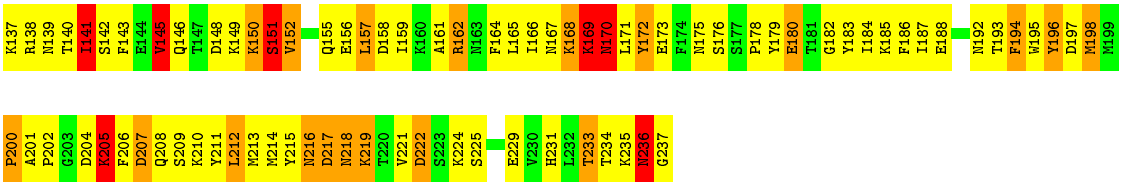


• Molecule 2: Enterotoxin type C-3

Chain F:

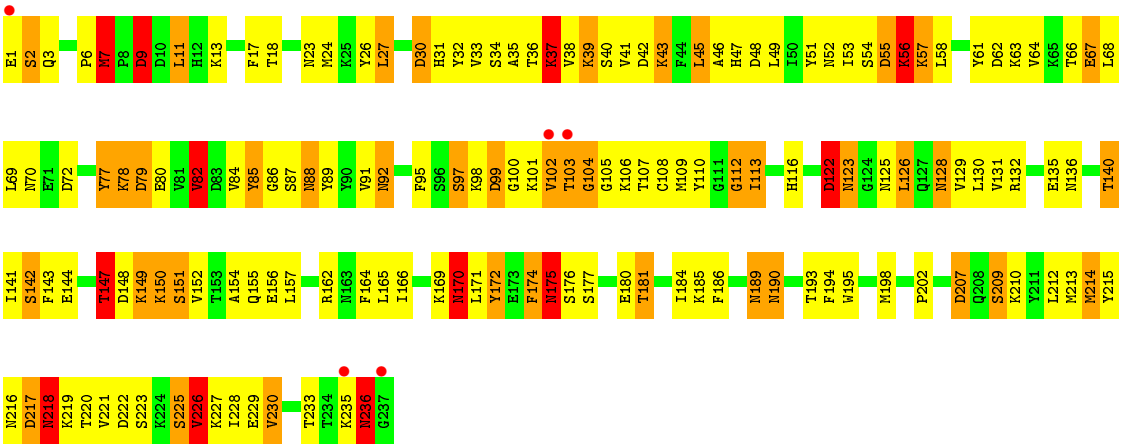






- Molecule 2: Enterotoxin type C-3

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	64.16Å 70.46Å 98.37Å 74.18° 75.76° 88.40°	Depositor
Resolution (Å)	40.00 – 2.30 63.42 – 2.28	Depositor EDS
% Data completeness (in resolution range)	94.2 (40.00-2.30) 92.4 (63.42-2.28)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.62 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.206 , 0.271 0.212 , 0.321	Depositor DCC
R_{free} test set	3402 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	38.3	Xtriage
Anisotropy	0.357	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 67.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	11155	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.60% 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

5.1 Standard geometry

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	2.51	38/845 (4.5%)	2.06	26/1142 (2.3%)
1	C	2.25	31/845 (3.7%)	1.88	19/1142 (1.7%)
1	E	2.20	24/846 (2.8%)	1.74	15/1145 (1.3%)
1	G	2.13	15/846 (1.8%)	1.98	22/1145 (1.9%)
2	B	2.05	63/1945 (3.2%)	1.72	37/2619 (1.4%)
2	D	1.83	26/1953 (1.3%)	1.67	32/2627 (1.2%)
2	F	1.83	31/1946 (1.6%)	1.60	22/2618 (0.8%)
2	H	1.77	30/1962 (1.5%)	1.56	22/2640 (0.8%)
All	All	2.00	258/11188 (2.3%)	1.73	195/15078 (1.3%)

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	6
1	E	0	1
1	G	0	1
2	B	0	5
2	D	0	1
2	F	0	4
2	H	0	1
All	All	0	19

The worst 5 of 258 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	101	TYR	C-N	23.04	1.87	1.34
1	E	101	TYR	C-N	21.36	1.83	1.34
1	C	63	GLY	C-N	20.06	1.80	1.34
1	E	63	GLY	C-N	18.65	1.76	1.34
1	G	101	TYR	C-N	17.43	1.74	1.34

The worst 5 of 195 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	G	101	TYR	O-C-N	-23.91	84.44	122.70
1	A	101	TYR	O-C-N	-22.11	87.32	122.70
1	E	101	TYR	O-C-N	-18.64	92.88	122.70
2	D	162	ARG	NE-CZ-NH1	-16.09	112.25	120.30
2	D	29	ASP	CB-CG-OD1	-13.04	106.57	118.30

There are no chirality outliers.

5 of 19 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	3	ALA	Peptide
1	A	35	TYR	Peptide
1	A	54	SER	Mainchain
1	A	6	GLN	Peptide
1	A	61	PRO	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	828	0	786	79	1
1	C	828	0	786	79	1
1	E	828	0	786	60	0
1	G	828	0	787	66	0
2	B	1904	0	1847	125	0
2	D	1913	0	1852	121	0
2	F	1906	0	1839	135	0
2	H	1922	0	1853	155	0
3	A	21	0	0	1	0
3	B	48	0	0	12	0
3	C	15	0	0	1	0
3	D	31	0	0	3	0
3	E	11	0	0	0	0
3	F	31	0	0	9	0
3	G	15	0	0	1	0
3	H	26	0	0	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	11155	0	10536	803	1

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 37.

The worst 5 of 803 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:A:26:THR:CB	1:A:26:THR:CG2	1.76	1.60
2:B:27:LEU:CG	2:B:27:LEU:CD1	1.86	1.53
1:E:80:GLU:CD	1:E:80:GLU:CG	1.75	1.51
1:G:101:TYR:C	1:G:108:PHE:N	1.74	1.41
1:A:32:MET:SD	1:A:32:MET:CE	2.08	1.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:26:THR:OG1	1:C:26:THR:CG2[1_455]	2.00	0.20

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	105/112 (94%)	91 (87%)	10 (10%)	4 (4%)	3	1
1	C	105/112 (94%)	92 (88%)	9 (9%)	4 (4%)	3	1
1	E	107/112 (96%)	83 (78%)	16 (15%)	8 (8%)	1	0
1	G	107/112 (96%)	88 (82%)	16 (15%)	3 (3%)	5	3
2	B	232/237 (98%)	188 (81%)	26 (11%)	18 (8%)	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
2	D	231/237 (98%)	175 (76%)	34 (15%)	22 (10%)	0 0
2	F	230/237 (97%)	185 (80%)	31 (14%)	14 (6%)	1 0
2	H	233/237 (98%)	167 (72%)	52 (22%)	14 (6%)	1 0
All	All	1350/1396 (97%)	1069 (79%)	194 (14%)	87 (6%)	1 0

5 of 87 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	7	SER
1	A	28	ASN
1	A	39	THR
2	B	80	GLU
2	B	101	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	88/91 (97%)	78 (89%)	10 (11%)	5 6
1	C	88/91 (97%)	60 (68%)	28 (32%)	0 0
1	E	88/91 (97%)	71 (81%)	17 (19%)	1 1
1	G	88/91 (97%)	67 (76%)	21 (24%)	0 0
2	B	215/217 (99%)	183 (85%)	32 (15%)	3 3
2	D	216/217 (100%)	177 (82%)	39 (18%)	1 1
2	F	215/217 (99%)	159 (74%)	56 (26%)	0 0
2	H	216/217 (100%)	179 (83%)	37 (17%)	2 2
All	All	1214/1232 (98%)	974 (80%)	240 (20%)	1 1

5 of 240 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	224	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	F	39	LYS
2	H	140	THR
1	E	37	GLN
1	E	85	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 46 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	88	ASN
2	D	163	ASN
2	H	189	ASN
2	D	92	ASN
2	D	128	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)

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)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	2
1	C	2
1	E	2
2	H	1
1	G	1

The worst 5 of 8 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	H	101:LYS	C	102:VAL	N	7.11
1	C	101:TYR	C	108:PHE	N	2.52
1	A	63:GLY	C	65:TYR	N	2.04
1	A	101:TYR	C	108:PHE	N	1.87
1	E	101:TYR	C	108:PHE	N	1.83

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	109/112 (97%)	-0.31	1 (0%) 84 88	8, 35, 52, 59	0
1	C	109/112 (97%)	-0.28	1 (0%) 84 88	21, 39, 58, 66	0
1	E	109/112 (97%)	-0.20	1 (0%) 84 88	27, 47, 64, 70	0
1	G	109/112 (97%)	-0.21	0 100 100	24, 40, 59, 63	0
2	B	234/237 (98%)	-0.32	1 (0%) 92 95	10, 41, 65, 86	0
2	D	235/237 (99%)	-0.19	3 (1%) 77 81	23, 47, 71, 87	0
2	F	234/237 (98%)	-0.28	1 (0%) 92 95	29, 45, 75, 91	0
2	H	237/237 (100%)	-0.08	5 (2%) 63 70	27, 50, 76, 103	0
All	All	1376/1396 (98%)	-0.23	13 (0%) 84 88	8, 44, 69, 103	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	237	GLY	4.9
2	H	102	VAL	4.0
2	H	1	GLU	3.6
2	D	237	GLY	3.5
2	H	103	THR	3.4

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.