W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Oct 5, 2016 – 05:36 PM EDT

|		
PDB ID	:	5JTV
Title	:	USP7CD-UBL45 in complex with Ubiquitin
Authors	:	Murray, J.M.; Rouge, L.
Deposited on	:	2016-05-09
Resolution	:	3.31 Å(reported)

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 3.31 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.
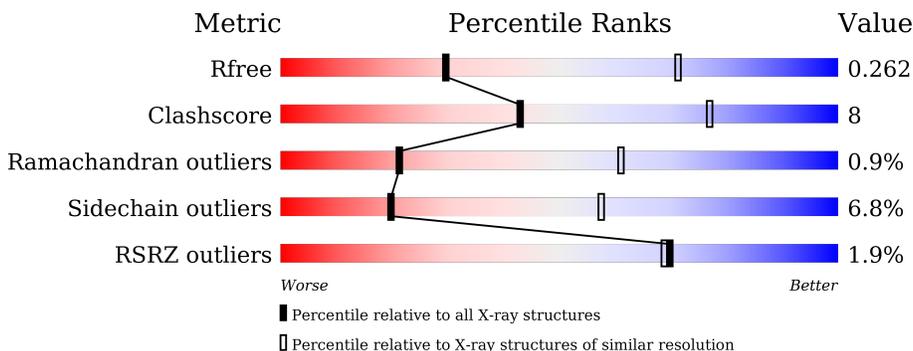
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1198 (3.40-3.24)
Clashscore	102246	1280 (3.40-3.24)
Ramachandran outliers	100387	1260 (3.40-3.24)
Sidechain outliers	100360	1259 (3.40-3.24)
RSRZ outliers	91569	1203 (3.40-3.24)

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	594	2% 70% 23% · 5%
1	C	594	% 74% 19% · 5%
1	E	594	2% 72% 20% ·· 5%
1	G	594	2% 72% 20% · 5%
2	B	76	84% 16%
2	D	76	84% 16%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	76	4% ████████ 89% ██ 11%
2	H	76	████████ 68% ███ 30% ·
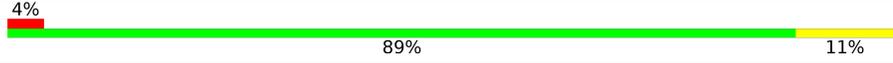
2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 20520 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 Ubiquitin carboxyl-terminal hydrolase 7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	564	Total 4562	C 2899	N 784	O 854	S 25	2	0	0
1	C	564	Total 4533	C 2878	N 784	O 846	S 25	26	0	0
1	E	565	Total 4530	C 2881	N 778	O 846	S 25	11	0	0
1	G	563	Total 4507	C 2861	N 776	O 845	S 25	0	0	0

There are 100 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	192	MET	-	initiating methionine	UNP Q93009
A	193	HIS	-	expression tag	UNP Q93009
A	194	HIS	-	expression tag	UNP Q93009
A	195	HIS	-	expression tag	UNP Q93009
A	196	HIS	-	expression tag	UNP Q93009
A	197	HIS	-	expression tag	UNP Q93009
A	198	HIS	-	expression tag	UNP Q93009
A	199	GLY	-	expression tag	UNP Q93009
A	200	GLU	-	expression tag	UNP Q93009
A	201	ASN	-	expression tag	UNP Q93009
A	202	LEU	-	expression tag	UNP Q93009
A	203	TYR	-	expression tag	UNP Q93009
A	204	PHE	-	expression tag	UNP Q93009
A	205	GLN	-	expression tag	UNP Q93009
A	206	GLY	-	expression tag	UNP Q93009
A	872	GLY	-	linker	UNP Q93009
A	873	GLY	-	linker	UNP Q93009
A	874	SER	-	linker	UNP Q93009
A	875	GLY	-	linker	UNP Q93009
A	876	GLY	-	linker	UNP Q93009
A	877	SER	-	linker	UNP Q93009

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	878	GLY	-	linker	UNP Q93009
A	879	GLY	-	linker	UNP Q93009
A	880	SER	-	linker	UNP Q93009
A	881	GLY	-	linker	UNP Q93009
C	192	MET	-	initiating methionine	UNP Q93009
C	193	HIS	-	expression tag	UNP Q93009
C	194	HIS	-	expression tag	UNP Q93009
C	195	HIS	-	expression tag	UNP Q93009
C	196	HIS	-	expression tag	UNP Q93009
C	197	HIS	-	expression tag	UNP Q93009
C	198	HIS	-	expression tag	UNP Q93009
C	199	GLY	-	expression tag	UNP Q93009
C	200	GLU	-	expression tag	UNP Q93009
C	201	ASN	-	expression tag	UNP Q93009
C	202	LEU	-	expression tag	UNP Q93009
C	203	TYR	-	expression tag	UNP Q93009
C	204	PHE	-	expression tag	UNP Q93009
C	205	GLN	-	expression tag	UNP Q93009
C	206	GLY	-	expression tag	UNP Q93009
C	872	GLY	-	linker	UNP Q93009
C	873	GLY	-	linker	UNP Q93009
C	874	SER	-	linker	UNP Q93009
C	875	GLY	-	linker	UNP Q93009
C	876	GLY	-	linker	UNP Q93009
C	877	SER	-	linker	UNP Q93009
C	878	GLY	-	linker	UNP Q93009
C	879	GLY	-	linker	UNP Q93009
C	880	SER	-	linker	UNP Q93009
C	881	GLY	-	linker	UNP Q93009
E	192	MET	-	initiating methionine	UNP Q93009
E	193	HIS	-	expression tag	UNP Q93009
E	194	HIS	-	expression tag	UNP Q93009
E	195	HIS	-	expression tag	UNP Q93009
E	196	HIS	-	expression tag	UNP Q93009
E	197	HIS	-	expression tag	UNP Q93009
E	198	HIS	-	expression tag	UNP Q93009
E	199	GLY	-	expression tag	UNP Q93009
E	200	GLU	-	expression tag	UNP Q93009
E	201	ASN	-	expression tag	UNP Q93009
E	202	LEU	-	expression tag	UNP Q93009
E	203	TYR	-	expression tag	UNP Q93009
E	204	PHE	-	expression tag	UNP Q93009

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	205	GLN	-	expression tag	UNP Q93009
E	206	GLY	-	expression tag	UNP Q93009
E	872	GLY	-	linker	UNP Q93009
E	873	GLY	-	linker	UNP Q93009
E	874	SER	-	linker	UNP Q93009
E	875	GLY	-	linker	UNP Q93009
E	876	GLY	-	linker	UNP Q93009
E	877	SER	-	linker	UNP Q93009
E	878	GLY	-	linker	UNP Q93009
E	879	GLY	-	linker	UNP Q93009
E	880	SER	-	linker	UNP Q93009
E	881	GLY	-	linker	UNP Q93009
G	192	MET	-	initiating methionine	UNP Q93009
G	193	HIS	-	expression tag	UNP Q93009
G	194	HIS	-	expression tag	UNP Q93009
G	195	HIS	-	expression tag	UNP Q93009
G	196	HIS	-	expression tag	UNP Q93009
G	197	HIS	-	expression tag	UNP Q93009
G	198	HIS	-	expression tag	UNP Q93009
G	199	GLY	-	expression tag	UNP Q93009
G	200	GLU	-	expression tag	UNP Q93009
G	201	ASN	-	expression tag	UNP Q93009
G	202	LEU	-	expression tag	UNP Q93009
G	203	TYR	-	expression tag	UNP Q93009
G	204	PHE	-	expression tag	UNP Q93009
G	205	GLN	-	expression tag	UNP Q93009
G	206	GLY	-	expression tag	UNP Q93009
G	872	GLY	-	linker	UNP Q93009
G	873	GLY	-	linker	UNP Q93009
G	874	SER	-	linker	UNP Q93009
G	875	GLY	-	linker	UNP Q93009
G	876	GLY	-	linker	UNP Q93009
G	877	SER	-	linker	UNP Q93009
G	878	GLY	-	linker	UNP Q93009
G	879	GLY	-	linker	UNP Q93009
G	880	SER	-	linker	UNP Q93009
G	881	GLY	-	linker	UNP Q93009

- Molecule 2 is a protein called Polyubiquitin-B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	76	597	375	104	117	1	0	0	0

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	76	Total	C	N	O	S	0	0	0
			597	375	104	117	1			
2	F	76	Total	C	N	O	S	0	0	0
			597	375	104	117	1			
2	H	76	Total	C	N	O	S	0	0	0
			597	375	104	117	1			

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 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: Ubiquitin carboxyl-terminal hydrolase 7
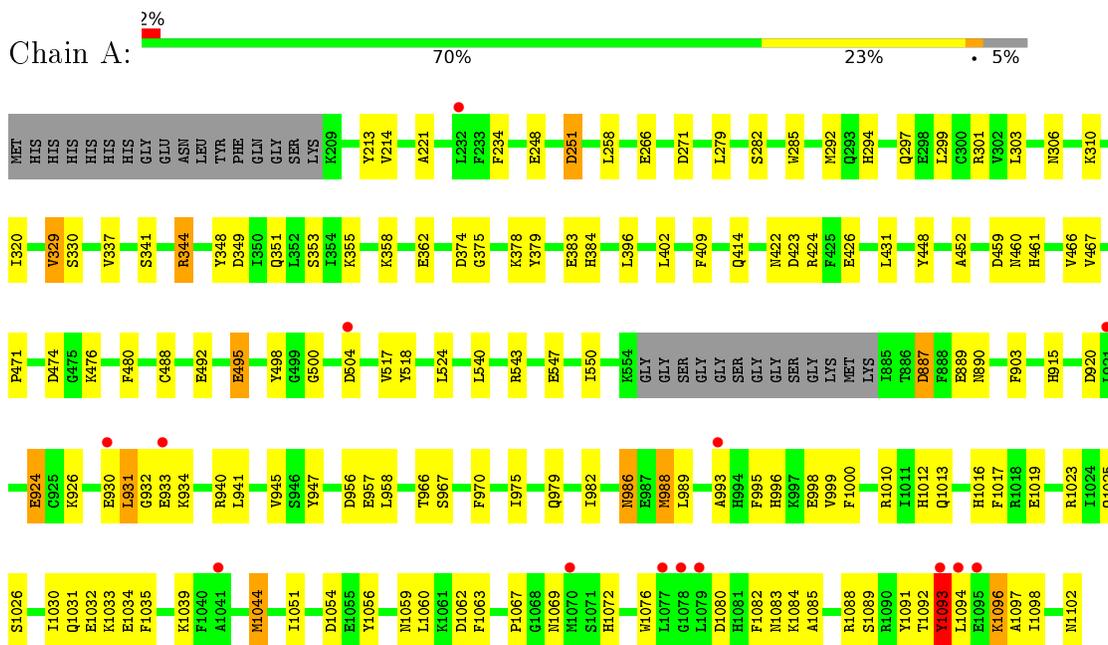
• Molecule 1: Ubiquitin carboxyl-terminal hydrolase 7

• Molecule 1: Ubiquitin carboxyl-terminal hydrolase 7
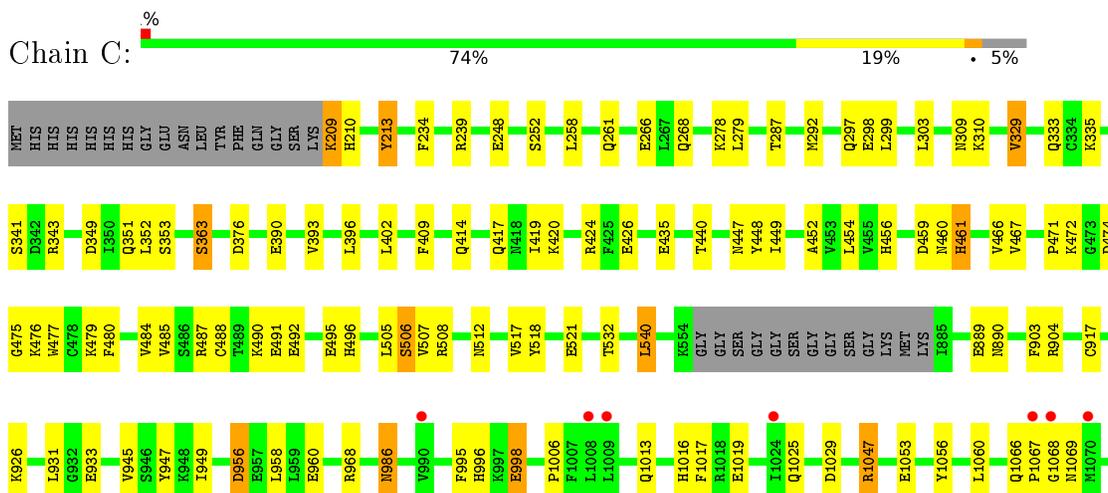
Chain E:
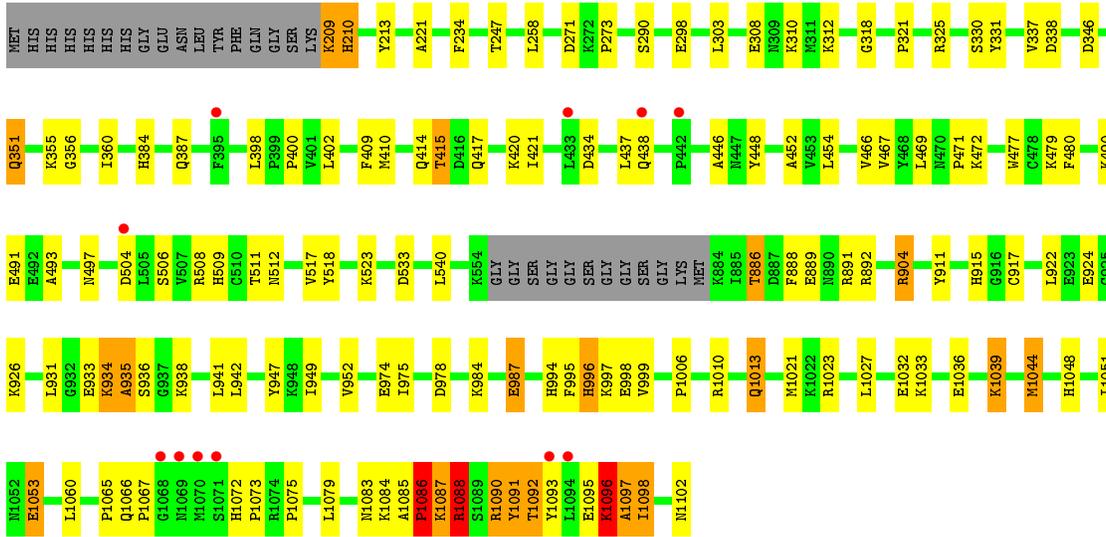
• Molecule 1: Ubiquitin carboxyl-terminal hydrolase 7
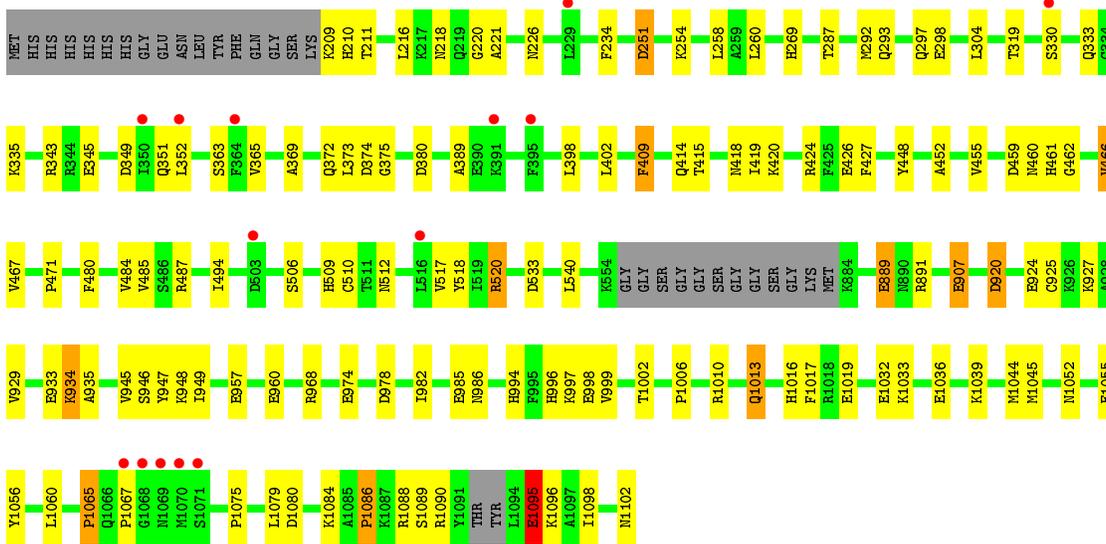
Chain G:

• Molecule 2: Polyubiquitin-B
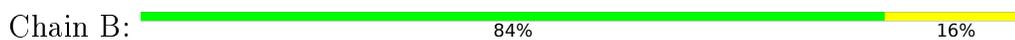
Chain B:

• Molecule 2: Polyubiquitin-B

Chain D:

84%   16%

M1 E18 P19 E24 N25 K29 Q49 G53 R54 T55 D58 K63 E64 R74 G75 G76

• Molecule 2: Polyubiquitin-B

Chain F:

4%

89%   11%

M1 L8 L15 I30 Q41 F45 L50 L67 H68 L69 R74 G75 G76

• Molecule 2: Polyubiquitin-B

Chain H:

68%   30%

M1 Q2 T14 L15 E18 P19 I23 E24 N25 V26 K33 Q40 Q41 R42 G53 R54 T55 L56 S57 D58 K63 E64 S65 H68 L69 V70 L71 R72 L73 R74 G75 G76

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	113.61Å 115.52Å 257.82Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.95 – 3.31 46.95 – 3.31	Depositor EDS
% Data completeness (in resolution range)	99.4 (46.95-3.31) 92.6 (46.95-3.31)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.64 (at 3.32Å)	Xtriage
Refinement program	PHENIX (dev_2210: ???)	Depositor
R, R_{free}	0.239 , 0.269 0.235 , 0.262	Depositor DCC
R_{free} test set	2395 reflections (5.05%)	DCC
Wilson B-factor (Å^2)	95.4	Xtriage
Anisotropy	0.341	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 28.3	EDS
L-test for twinning[2]	$< |L| > = 0.47, < L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.036 for k,h,-l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	20520	wwPDB-VP
Average B, all atoms (Å^2)	114.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.76% 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 (i)

5.1 Standard geometry (i)

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	0.34	3/4663 (0.1%)	0.53	1/6290 (0.0%)
1	C	0.26	0/4634	0.49	1/6254 (0.0%)
1	E	0.28	0/4630	0.54	4/6249 (0.1%)
1	G	0.29	0/4605	0.54	4/6215 (0.1%)
2	B	0.32	0/603	0.68	1/812 (0.1%)
2	D	0.27	0/603	0.52	0/812
2	F	0.28	0/603	0.60	0/812
2	H	0.30	0/603	0.67	1/812 (0.1%)
All	All	0.29	3/20944 (0.0%)	0.54	12/28256 (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	A	0	1
1	E	0	2
1	G	0	2
All	All	0	5

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	498	TYR	CD2-CE2	-7.40	1.28	1.39
1	A	498	TYR	CD1-CE1	-6.54	1.29	1.39
1	A	498	TYR	CE2-CZ	-5.01	1.32	1.38

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	G	506	SER	CB-CA-C	13.36	135.49	110.10
1	E	1088	ARG	N-CA-C	8.57	134.13	111.00

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	1096	LYS	CD-CE-NZ	-8.38	92.44	111.70
2	B	71	LEU	CB-CA-C	-7.57	95.81	110.20
1	G	996	HIS	N-CA-C	-7.00	92.11	111.00

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	504	ASP	Peptide
1	E	1092	THR	Peptide
1	E	504	ASP	Peptide
1	G	1090	ARG	Peptide
1	G	1095	GLU	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	4562	0	4423	91	1
1	C	4533	0	4370	71	0
1	E	4530	0	4366	86	0
1	G	4507	0	4335	88	1
2	B	597	0	618	12	0
2	D	597	0	618	8	0
2	F	597	0	618	12	0
2	H	597	0	618	21	0
All	All	20520	0	19966	330	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 8.

The worst 5 of 330 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)
2:B:74:ARG:NH1	1:E:1096:LYS:HB2	1.64	1.12

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1096:LYS:HB3	2:F:74:ARG:NH1	1.69	1.08
1:C:292:MET:HE1	1:G:1095:GLU:O	1.59	1.03
1:E:934:LYS:HZ3	1:E:934:LYS:H	1.03	1.01
1:C:292:MET:CE	1:G:1095:GLU:O	2.09	1.00

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:383:GLU:OE2	1:G:287:THR:OG1[4_545]	2.08	0.12

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	560/594 (94%)	523 (93%)	33 (6%)	4 (1%)	26	66
1	C	560/594 (94%)	522 (93%)	31 (6%)	7 (1%)	15	53
1	E	561/594 (94%)	516 (92%)	37 (7%)	8 (1%)	14	50
1	G	557/594 (94%)	517 (93%)	37 (7%)	3 (0%)	34	72
2	B	74/76 (97%)	74 (100%)	0	0	100	100
2	D	74/76 (97%)	74 (100%)	0	0	100	100
2	F	74/76 (97%)	74 (100%)	0	0	100	100
2	H	74/76 (97%)	74 (100%)	0	0	100	100
All	All	2534/2680 (95%)	2374 (94%)	138 (5%)	22 (1%)	21	61

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1093	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1097	ALA
1	E	1087	LYS
1	C	475	GLY
1	C	1096	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	494/535 (92%)	453 (92%)	41 (8%)	14	47
1	C	487/535 (91%)	447 (92%)	40 (8%)	14	48
1	E	484/535 (90%)	448 (93%)	36 (7%)	17	52
1	G	483/535 (90%)	453 (94%)	30 (6%)	23	60
2	B	67/68 (98%)	65 (97%)	2 (3%)	48	79
2	D	67/68 (98%)	67 (100%)	0	100	100
2	F	67/68 (98%)	67 (100%)	0	100	100
2	H	67/68 (98%)	66 (98%)	1 (2%)	72	88
All	All	2216/2412 (92%)	2066 (93%)	150 (7%)	20	57

5 of 150 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	454	LEU
1	E	209	LYS
1	G	934	LYS
1	C	484	VAL
1	C	945	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	387	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	915	HIS
1	G	915	HIS
1	E	230	GLN
1	G	1012	HIS

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)

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(Å²)	Q<0.9
1	A	564/594 (94%)	-0.05	14 (2%)	61	60	91, 114, 145, 166	1 (0%)
1	C	562/594 (94%)	-0.12	7 (1%)	81	81	88, 109, 137, 161	1 (0%)
1	E	565/594 (95%)	-0.13	11 (1%)	70	69	89, 110, 144, 168	2 (0%)
1	G	563/594 (94%)	-0.10	14 (2%)	61	60	88, 114, 153, 181	0
2	B	76/76 (100%)	-0.07	0	100	100	94, 105, 115, 123	0
2	D	76/76 (100%)	-0.09	0	100	100	89, 104, 113, 118	0
2	F	76/76 (100%)	-0.01	3 (3%)	43	41	103, 124, 139, 155	0
2	H	76/76 (100%)	-0.12	0	100	100	105, 135, 152, 162	0
All	All	2558/2680 (95%)	-0.10	49 (1%)	70	69	88, 112, 145, 181	4 (0%)

The worst 5 of 49 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1094	LEU	11.2
1	E	504	ASP	6.6
1	E	1094	LEU	5.2
1	G	1070	MET	5.0
1	A	504	ASP	4.8

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.