



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 11, 2016 – 02:56 PM EDT

PDB ID : 5HPT
Title : System-wide modulation of HECT E3 ligases with selective ubiquitin variant probes: WWP1, Ubv P2.3 and UBCH7
Authors : Wu, K.-P.; Schulman, B.A.
Deposited on : 2016-01-20
Resolution : 2.84 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : unknown
Xtriage (Phenix) : 1.9-1692
EDS : rb-20027107
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) : rb-20027107

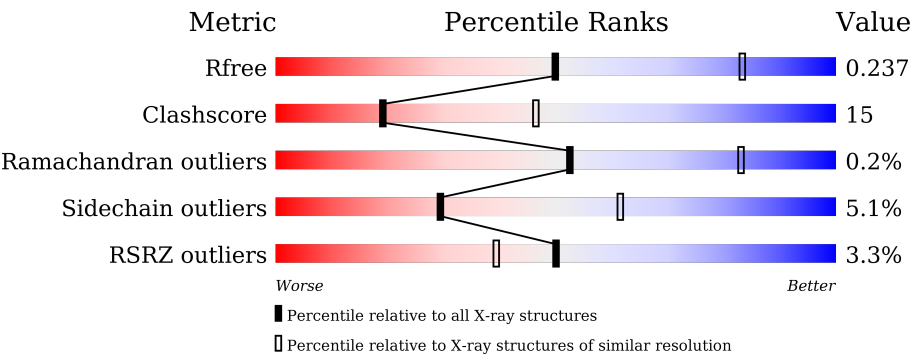
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.84 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	3170 (2.88-2.80)
Clashscore	102246	3658 (2.88-2.80)
Ramachandran outliers	100387	3591 (2.88-2.80)
Sidechain outliers	100360	3594 (2.88-2.80)
RSRZ outliers	91569	3184 (2.88-2.80)

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 $\leq 5\%$. 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	383	<div><div>2%</div><div></div><div>73%</div><div>20%</div><div>.</div><div>.</div></div>
1	D	383	<div><div>4%</div><div></div><div>66%</div><div>27%</div><div>.</div><div>.</div></div>
1	G	383	<div><div>2%</div><div></div><div>68%</div><div>25%</div><div>.</div><div>.</div></div>
2	B	83	<div><div></div><div></div><div>67%</div><div>23%</div><div>.</div><div>8%</div></div>
2	E	83	<div><div>%</div><div></div><div>55%</div><div>29%</div><div>7%</div><div>8%</div></div>
2	H	83	<div><div>6%</div><div></div><div>65%</div><div>20%</div><div>5%</div><div>10%</div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	C	161	<div><div></div><div>3%</div><div>57%</div><div>34%</div><div>6%</div></div>
3	F	161	<div><div></div><div>8%</div><div>55%</div><div>35%</div><div>7%</div></div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 13529 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 NEDD4-like E3 ubiquitin-protein ligase WWP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	373	Total	C	N	O	S	0	0	0
			3111	2013	520	558	20			
1	D	371	Total	C	N	O	S	0	0	0
			3082	1994	516	552	20			
1	G	367	Total	C	N	O	S	0	0	0
			3058	1980	509	549	20			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	535	GLY	-	expression tag	UNP Q9H0M0
A	536	SER	-	expression tag	UNP Q9H0M0
D	535	GLY	-	expression tag	UNP Q9H0M0
D	536	SER	-	expression tag	UNP Q9H0M0
G	535	GLY	-	expression tag	UNP Q9H0M0
G	536	SER	-	expression tag	UNP Q9H0M0

- Molecule 2 is a protein called Ubiquitin variant P2.3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	76	Total	C	N	O	Se	0	0	0
			607	388	101	116	2			
2	E	76	Total	C	N	O	Se	0	0	0
			607	388	101	116	2			
2	H	75	Total	C	N	O	Se	0	0	0
			598	383	99	114	2			

- Molecule 3 is a protein called Ubiquitin-conjugating enzyme E2 L3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	151	Total	C	N	O	S	0	0	0
			1234	790	214	225	5			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	F	150	Total	C	N	O	S	0	0	0
			1227	785	213	224	5			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	155	GLY	-	expression tag	UNP P68036
C	156	GLY	-	expression tag	UNP P68036
C	157	HIS	-	expression tag	UNP P68036
C	158	HIS	-	expression tag	UNP P68036
C	159	HIS	-	expression tag	UNP P68036
C	160	HIS	-	expression tag	UNP P68036
C	161	HIS	-	expression tag	UNP P68036
C	162	HIS	-	expression tag	UNP P68036
F	155	GLY	-	expression tag	UNP P68036
F	156	GLY	-	expression tag	UNP P68036
F	157	HIS	-	expression tag	UNP P68036
F	158	HIS	-	expression tag	UNP P68036
F	159	HIS	-	expression tag	UNP P68036
F	160	HIS	-	expression tag	UNP P68036
F	161	HIS	-	expression tag	UNP P68036
F	162	HIS	-	expression tag	UNP P68036

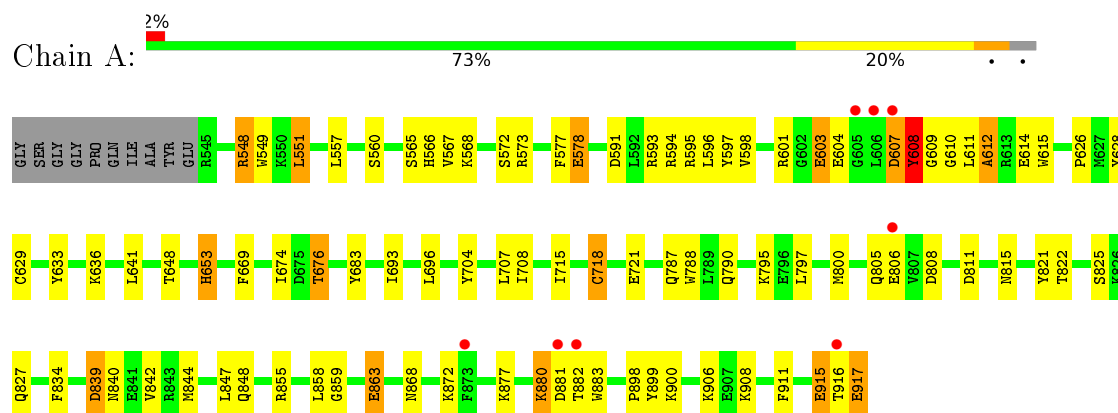
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	3	Total	O	0	0
			3	3		
4	C	1	Total	O	0	0
			1	1		
4	D	1	Total	O	0	0
			1	1		

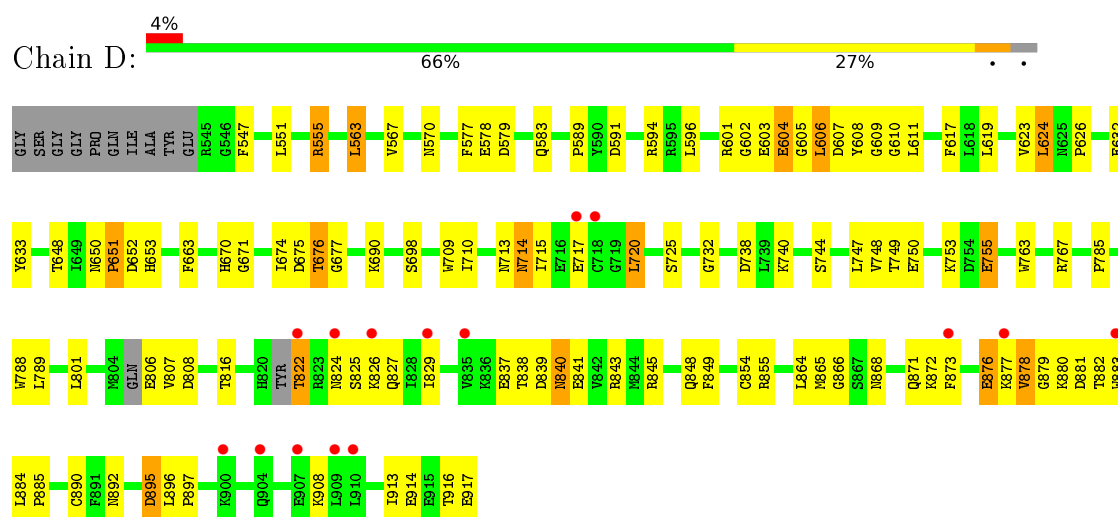
3 Residue-property plots

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.

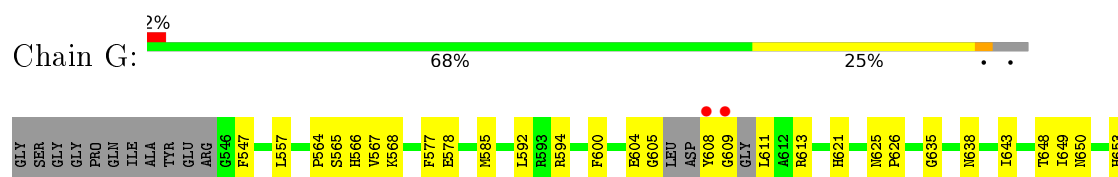
- Molecule 1: NEDD4-like E3 ubiquitin-protein ligase WWP1

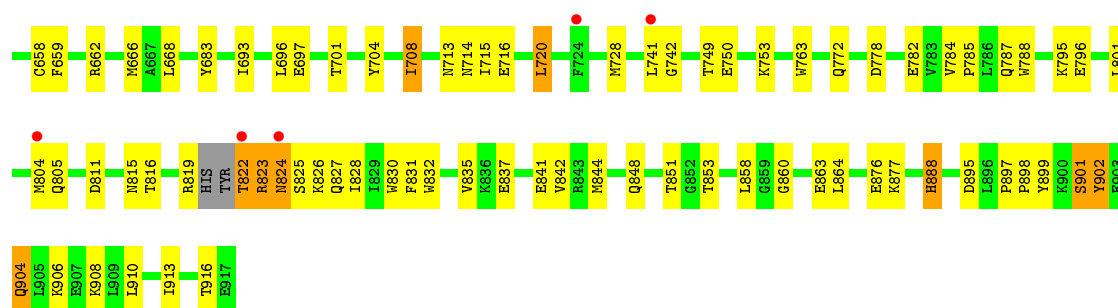


- Molecule 1: NEDD4-like E3 ubiquitin-protein ligase WWP1



- Molecule 1: NEDD4-like E3 ubiquitin-protein ligase WWP1





- Molecule 2: Ubiquitin variant P2.3



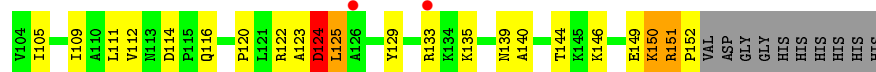
- Molecule 2: Ubiquitin variant P2.3



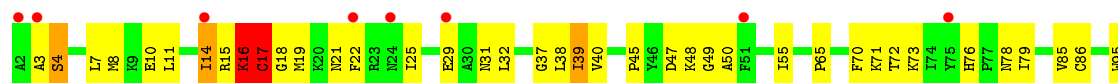
- Molecule 2: Ubiquitin variant P2.3

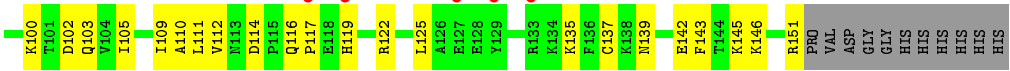


- Molecule 3: Ubiquitin-conjugating enzyme E2 L3



- Molecule 3: Ubiquitin-conjugating enzyme E2 L3





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	114.00Å 118.90Å 158.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	95.15 – 2.84 95.15 – 2.84	Depositor EDS
% Data completeness (in resolution range)	99.5 (95.15-2.84) 91.3 (95.15-2.84)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.73 (at 2.86Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.218 , 0.237 0.218 , 0.237	Depositor DCC
R_{free} test set	3814 reflections (3.90%)	DCC
Wilson B-factor (Å ²)	67.3	Xtriage
Anisotropy	0.343	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 55.1	EDS
Estimated twinning fraction	0.023 for k,h,l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 51451 reflections	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	13529	wwPDB-VP
Average B, all atoms (Å ²)	86.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.63% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.62	0/3193	0.79	5/4309 (0.1%)
1	D	0.61	2/3161 (0.1%)	0.76	2/4264 (0.0%)
1	G	0.53	0/3135	0.72	2/4226 (0.0%)
2	B	0.55	0/615	0.72	0/828
2	E	0.51	0/615	0.85	4/828 (0.5%)
2	H	0.53	0/606	0.74	1/816 (0.1%)
3	C	0.52	0/1265	0.76	2/1709 (0.1%)
3	F	0.44	0/1257	0.88	5/1697 (0.3%)
All	All	0.57	2/13847 (0.0%)	0.77	21/18677 (0.1%)

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	2
1	D	0	2
1	G	0	1
2	H	0	1
3	C	0	1
3	F	0	2
All	All	0	9

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	717	GLU	CB-CG	6.11	1.63	1.52
1	D	717	GLU	CG-CD	5.42	1.60	1.51

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	17	CYS	CA-C-N	-14.21	87.79	116.20
3	F	17	CYS	O-C-N	13.18	145.61	123.20
3	F	17	CYS	C-N-CA	-12.85	95.32	122.30
2	E	64	GLY	N-CA-C	-9.02	90.56	113.10
3	C	124	ASP	N-CA-C	-8.73	87.43	111.00

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	603	GLU	Peptide
1	A	612	ALA	Peptide
3	C	151	ARG	Peptide
1	D	714	ASN	Peptide
1	D	876	GLU	Peptide

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	3111	0	3044	77	0
1	D	3082	0	3008	106	1
1	G	3058	0	2994	93	2
2	B	607	0	633	15	0
2	E	607	0	633	29	0
2	H	598	0	625	16	0
3	C	1234	0	1245	45	1
3	F	1227	0	1236	44	0
4	A	3	0	0	1	0
4	C	1	0	0	0	0
4	D	1	0	0	1	0
All	All	13529	0	13418	409	2

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 15.

The worst 5 of 409 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:D:878:VAL:HB	1:D:879:GLY:CA	1.62	1.29
1:G:822:THR:O	1:G:825:SER:HB3	1.18	1.27
1:A:881:ASP:OD2	1:A:900:LYS:HE2	1.31	1.22
3:C:13:GLU:O	3:C:17:CYS:SG	2.02	1.17
1:D:806:GLU:HG3	1:D:807:VAL:HB	1.17	1.16

All (2) 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)
3:C:146:LYS:O	1:G:568:LYS:NZ[4_445]	1.99	0.21
1:D:608:TYR:OH	1:G:908:LYS:NZ[4_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	371/383 (97%)	360 (97%)	11 (3%)	0	100	100
1	D	365/383 (95%)	343 (94%)	21 (6%)	1 (0%)	46	77
1	G	359/383 (94%)	335 (93%)	24 (7%)	0	100	100
2	B	74/83 (89%)	73 (99%)	1 (1%)	0	100	100
2	E	74/83 (89%)	71 (96%)	3 (4%)	0	100	100
2	H	73/83 (88%)	70 (96%)	3 (4%)	0	100	100
3	C	149/161 (92%)	138 (93%)	9 (6%)	2 (1%)	15	42
3	F	148/161 (92%)	140 (95%)	8 (5%)	0	100	100
All	All	1613/1720 (94%)	1530 (95%)	80 (5%)	3 (0%)	52	83

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	120	PRO
1	D	878	VAL
3	C	125	LEU

5.3.2 Protein sidechains ⓘ

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	338/344 (98%)	320 (95%)	18 (5%)	28	60
1	D	334/344 (97%)	310 (93%)	24 (7%)	18	43
1	G	333/344 (97%)	320 (96%)	13 (4%)	39	72
2	B	69/71 (97%)	66 (96%)	3 (4%)	35	69
2	E	69/71 (97%)	66 (96%)	3 (4%)	35	69
2	H	68/71 (96%)	65 (96%)	3 (4%)	35	68
3	C	135/143 (94%)	129 (96%)	6 (4%)	35	68
3	F	134/143 (94%)	129 (96%)	5 (4%)	41	74
All	All	1480/1531 (97%)	1405 (95%)	75 (5%)	29	61

5 of 75 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	624	LEU
1	D	744	SER
1	G	888	HIS
1	D	651	PRO
1	D	698	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	871	GLN
2	H	25	ASN
1	G	827	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	827	GLN
1	G	888	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 ⓘ

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, 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	373/383 (97%)	0.18	8 (2%) 67 57	42, 64, 104, 134	0
1	D	371/383 (96%)	0.39	15 (4%) 42 31	39, 69, 162, 199	0
1	G	367/383 (95%)	0.28	7 (1%) 70 61	49, 83, 124, 151	0
2	B	74/83 (89%)	0.09	0 100 100	54, 78, 94, 104	0
2	E	74/83 (89%)	0.33	1 (1%) 78 71	64, 103, 130, 138	0
2	H	73/83 (87%)	0.48	5 (6%) 20 12	76, 95, 119, 133	0
3	C	151/161 (93%)	0.40	5 (3%) 50 39	56, 92, 144, 159	0
3	F	150/161 (93%)	0.73	13 (8%) 13 6	68, 121, 153, 167	0
All	All	1633/1720 (94%)	0.34	54 (3%) 50 39	39, 82, 141, 199	0

The worst 5 of 54 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	F	2	ALA	7.6
1	G	609	GLY	6.1
2	H	75	GLY	5.8
1	A	882	THR	5.1
1	G	822	THR	5.1

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 [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.