wwPDB X-ray Structure Validation Summary Report (i)

Sep 17, 2023 – 04:19 AM EDT

PDB ID	:	4WHV
Title	:	E3 ubiquitin-protein ligase RNF8 in complex with Ubiquitin-conjugating enzyme E2 N and Polyubiquitin-B
Authors	:	Hodge, C.D.; Edwards, R.A.; Glover, J.N.M.
Deposited on	:	2014-09-23
Resolution	:	8.30 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

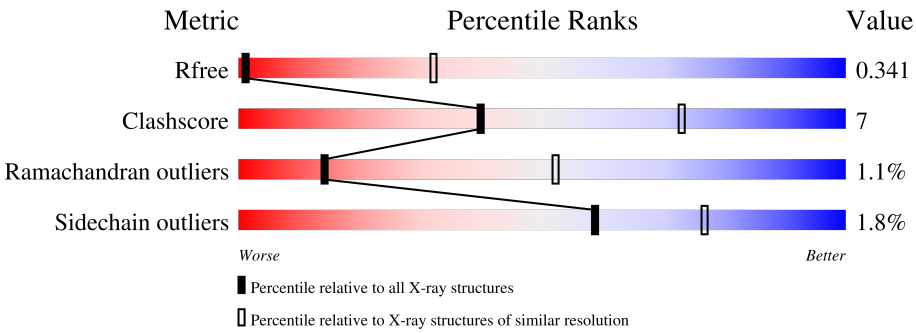
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 8.30 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	1005 (11.50-3.90)
Clashscore	141614	1070 (11.50-3.90)
Ramachandran outliers	138981	1003 (11.50-3.90)
Sidechain outliers	138945	1003 (11.50-3.86)

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 of 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%

Mol	Chain	Length	Quality of chain
1	B	160	76% 13% • 8%
1	E	160	74% 15% • 8%
1	H	160	77% 12% • 8%
1	K	160	79% 10% • 8%
2	C	149	79% 11% •• 9%
2	D	149	81% 12% • 7%
2	I	149	56% 10% •• 32%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	J	149	58% 9% • 32%
3	A	83	81% 5% 14%
3	F	83	76% 10% 14%
3	G	83	83% • 14%
3	L	83	75% 11% 14%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 9620 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-conjugating enzyme E2 N.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	147	1123	723	189	209	2	0	0	0
1	E	147	1123	723	189	209	2	0	0	0
1	H	147	1123	723	189	209	2	0	0	0
1	K	147	1123	723	189	209	2	0	0	0

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-7	GLY	-	expression tag	UNP P61088
B	-6	PRO	-	expression tag	UNP P61088
B	-5	LEU	-	expression tag	UNP P61088
B	-4	GLY	-	expression tag	UNP P61088
B	-3	SER	-	expression tag	UNP P61088
B	-2	PRO	-	expression tag	UNP P61088
B	-1	GLU	-	expression tag	UNP P61088
B	0	PHE	-	expression tag	UNP P61088
B	87	LYS	CYS	engineered mutation	UNP P61088
E	-7	GLY	-	expression tag	UNP P61088
E	-6	PRO	-	expression tag	UNP P61088
E	-5	LEU	-	expression tag	UNP P61088
E	-4	GLY	-	expression tag	UNP P61088
E	-3	SER	-	expression tag	UNP P61088
E	-2	PRO	-	expression tag	UNP P61088
E	-1	GLU	-	expression tag	UNP P61088
E	0	PHE	-	expression tag	UNP P61088
E	87	LYS	CYS	engineered mutation	UNP P61088
H	-7	GLY	-	expression tag	UNP P61088
H	-6	PRO	-	expression tag	UNP P61088
H	-5	LEU	-	expression tag	UNP P61088

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	-4	GLY	-	expression tag	UNP P61088
H	-3	SER	-	expression tag	UNP P61088
H	-2	PRO	-	expression tag	UNP P61088
H	-1	GLU	-	expression tag	UNP P61088
H	0	PHE	-	expression tag	UNP P61088
H	87	LYS	CYS	engineered mutation	UNP P61088
K	-7	GLY	-	expression tag	UNP P61088
K	-6	PRO	-	expression tag	UNP P61088
K	-5	LEU	-	expression tag	UNP P61088
K	-4	GLY	-	expression tag	UNP P61088
K	-3	SER	-	expression tag	UNP P61088
K	-2	PRO	-	expression tag	UNP P61088
K	-1	GLU	-	expression tag	UNP P61088
K	0	PHE	-	expression tag	UNP P61088
K	87	LYS	CYS	engineered mutation	UNP P61088

- Molecule 2 is a protein called E3 ubiquitin-protein ligase RNF8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	136	916	563	166	177	10	0	0	0
			Total	C	N	O	S			
2	D	139	930	571	169	180	10	0	0	0
			Total	C	N	O	S			
2	I	101	741	458	131	142	10	0	0	0
			Total	C	N	O	S			
2	J	101	741	458	131	142	10	0	0	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	337	GLY	-	expression tag	UNP O76064
C	338	PRO	-	expression tag	UNP O76064
C	339	LEU	-	expression tag	UNP O76064
C	340	GLY	-	expression tag	UNP O76064
C	341	SER	-	expression tag	UNP O76064
C	342	PRO	-	expression tag	UNP O76064
C	343	GLU	-	expression tag	UNP O76064
C	344	PHE	-	expression tag	UNP O76064
D	337	GLY	-	expression tag	UNP O76064
D	338	PRO	-	expression tag	UNP O76064
D	339	LEU	-	expression tag	UNP O76064

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	340	GLY	-	expression tag	UNP O76064
D	341	SER	-	expression tag	UNP O76064
D	342	PRO	-	expression tag	UNP O76064
D	343	GLU	-	expression tag	UNP O76064
D	344	PHE	-	expression tag	UNP O76064
I	337	GLY	-	expression tag	UNP O76064
I	338	PRO	-	expression tag	UNP O76064
I	339	LEU	-	expression tag	UNP O76064
I	340	GLY	-	expression tag	UNP O76064
I	341	SER	-	expression tag	UNP O76064
I	342	PRO	-	expression tag	UNP O76064
I	343	GLU	-	expression tag	UNP O76064
I	344	PHE	-	expression tag	UNP O76064
J	337	GLY	-	expression tag	UNP O76064
J	338	PRO	-	expression tag	UNP O76064
J	339	LEU	-	expression tag	UNP O76064
J	340	GLY	-	expression tag	UNP O76064
J	341	SER	-	expression tag	UNP O76064
J	342	PRO	-	expression tag	UNP O76064
J	343	GLU	-	expression tag	UNP O76064
J	344	PHE	-	expression tag	UNP O76064

- Molecule 3 is a protein called Polyubiquitin-B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	71	448	292	74	81	1	0	0	0
			Total	C	N	O	S			
3	F	71	448	292	74	81	1	0	0	0
			Total	C	N	O	S			
3	G	71	448	292	74	81	1	0	0	0
			Total	C	N	O	S			
3	L	71	448	292	74	81	1	0	0	0

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	GLY	-	expression tag	UNP P0CG47
A	-5	PRO	-	expression tag	UNP P0CG47
A	-4	GLY	-	expression tag	UNP P0CG47
A	-3	TYR	-	expression tag	UNP P0CG47
A	-2	GLN	-	expression tag	UNP P0CG47

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	ASP	-	expression tag	UNP P0CG47
A	0	PRO	-	expression tag	UNP P0CG47
F	-6	GLY	-	expression tag	UNP P0CG47
F	-5	PRO	-	expression tag	UNP P0CG47
F	-4	GLY	-	expression tag	UNP P0CG47
F	-3	TYR	-	expression tag	UNP P0CG47
F	-2	GLN	-	expression tag	UNP P0CG47
F	-1	ASP	-	expression tag	UNP P0CG47
F	0	PRO	-	expression tag	UNP P0CG47
G	-6	GLY	-	expression tag	UNP P0CG47
G	-5	PRO	-	expression tag	UNP P0CG47
G	-4	GLY	-	expression tag	UNP P0CG47
G	-3	TYR	-	expression tag	UNP P0CG47
G	-2	GLN	-	expression tag	UNP P0CG47
G	-1	ASP	-	expression tag	UNP P0CG47
G	0	PRO	-	expression tag	UNP P0CG47
L	-6	GLY	-	expression tag	UNP P0CG47
L	-5	PRO	-	expression tag	UNP P0CG47
L	-4	GLY	-	expression tag	UNP P0CG47
L	-3	TYR	-	expression tag	UNP P0CG47
L	-2	GLN	-	expression tag	UNP P0CG47
L	-1	ASP	-	expression tag	UNP P0CG47
L	0	PRO	-	expression tag	UNP P0CG47

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
4	C	2	2	2	0	0
4	D	2	2	2	0	0
4	I	2	2	2	0	0
4	J	2	2	2	0	0

3 Residue-property plots (i)

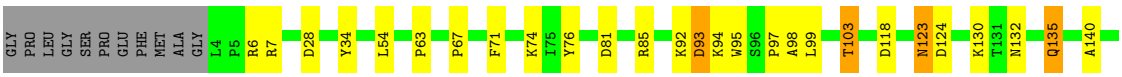
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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. 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-conjugating enzyme E2 N

Chain B:

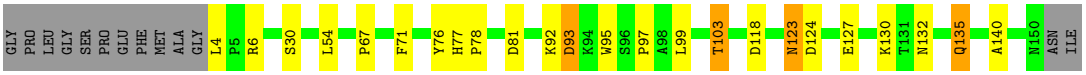
• Molecule 1: Ubiquitin-conjugating enzyme E2 N

Chain E:



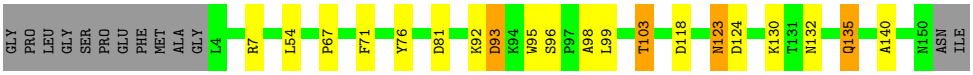
• Molecule 1: Ubiquitin-conjugating enzyme E2 N

Chain H:



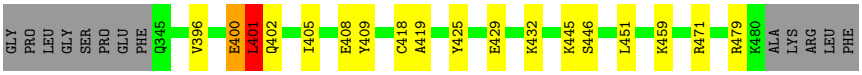
• Molecule 1: Ubiquitin-conjugating enzyme E2 N

Chain K:



• Molecule 2: E3 ubiquitin-protein ligase RNF8

Chain C:



• Molecule 2: E3 ubiquitin-protein ligase RNF8

Chain D:

• Molecule 2: E3 ubiquitin-protein ligase RNF8

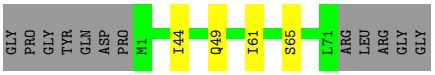
Chain I:

• Molecule 2: E3 ubiquitin-protein ligase RNF8

Chain J:

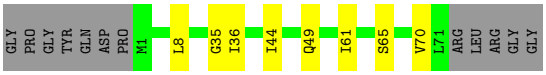
• Molecule 3: Polyubiquitin-B

Chain A:



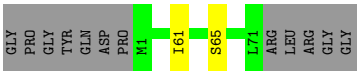
• Molecule 3: Polyubiquitin-B

Chain F:



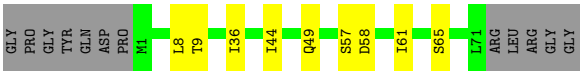
• Molecule 3: Polyubiquitin-B

Chain G:



• Molecule 3: Polyubiquitin-B

Chain L:


GLY
PRO
GLY
TYR
GLN
ASP
PRO
M1
L8
T9
I36
I44
Q49
S57
D58
I61
S65
L71
ARG
LEU
ARG
GLY
GLY

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 64 2 2	Depositor
Cell constants a, b, c, α, β, γ	341.38Å 341.38Å 113.45Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.27 – 8.30 49.27 – 8.21	Depositor EDS
% Data completeness (in resolution range)	99.4 (49.27-8.30) 99.8 (49.27-8.21)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	8.65 (at 8.33Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.329 , 0.337 0.339 , 0.341	Depositor DCC
R_{free} test set	189 reflections (4.62%)	wwPDB-VP
Wilson B-factor (Å^2)	468.4	Xtriage
Anisotropy	0.337	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 822.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.67	EDS
Total number of atoms	9620	wwPDB-VP
Average B, all atoms (Å^2)	522.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.44% 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)

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	B	0.27	0/1151	0.48	0/1575
1	E	0.27	0/1151	0.48	0/1575
1	H	0.27	0/1151	0.48	0/1575
1	K	0.27	0/1151	0.48	0/1575
2	C	0.48	0/922	0.68	0/1255
2	D	0.48	0/936	0.65	0/1274
2	I	0.49	0/747	0.63	0/1010
2	J	0.49	0/747	0.60	0/1010
3	A	0.37	0/452	0.55	0/624
3	F	0.37	0/452	0.55	0/624
3	G	0.37	0/452	0.55	0/624
3	L	0.37	0/452	0.55	0/624
All	All	0.38	0/9764	0.56	0/13345

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 (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	B	1123	0	1090	12	3

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	1123	0	1090	32	8
1	H	1123	0	1090	13	3
1	K	1123	0	1090	18	0
2	C	916	0	747	19	0
2	D	930	0	751	39	0
2	I	741	0	668	18	0
2	J	741	0	668	16	0
3	A	448	0	394	2	0
3	F	448	0	394	10	0
3	G	448	0	394	1	0
3	L	448	0	394	10	6
4	C	2	0	0	0	0
4	D	2	0	0	0	0
4	I	2	0	0	0	0
4	J	2	0	0	0	0
All	All	9620	0	8770	134	19

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 7.

The worst 5 of 134 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:K:98:ALA:O	3:L:9:THR:HA	1.22	1.27
2:D:438:PRO:O	1:E:98:ALA:HB3	1.49	1.13
2:D:408:GLU:OE2	1:E:6:ARG:CB	2.01	1.08
1:K:98:ALA:O	3:L:9:THR:CA	2.03	1.05
2:C:446:SER:HA	2:I:445:LYS:O	1.58	1.04

The worst 5 of 19 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:H:4:LEU:N	1:H:30:SER:O[7_555]	1.16	1.04
3:L:58:ASP:CA	3:L:58:ASP:CA[7_554]	1.24	0.96
1:E:34:TYR:CD2	1:E:74:LYS:NZ[12_554]	1.31	0.89
1:H:4:LEU:N	1:H:30:SER:C[7_555]	1.54	0.66
3:L:57:SER:O	3:L:57:SER:O[7_554]	1.55	0.65

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	B	145/160 (91%)	136 (94%)	7 (5%)	2 (1%)	11	46
1	E	145/160 (91%)	136 (94%)	7 (5%)	2 (1%)	11	46
1	H	145/160 (91%)	136 (94%)	7 (5%)	2 (1%)	11	46
1	K	145/160 (91%)	136 (94%)	7 (5%)	2 (1%)	11	46
2	C	134/149 (90%)	128 (96%)	4 (3%)	2 (2%)	10	46
2	D	137/149 (92%)	131 (96%)	5 (4%)	1 (1%)	22	63
2	I	99/149 (66%)	93 (94%)	4 (4%)	2 (2%)	7	38
2	J	99/149 (66%)	94 (95%)	4 (4%)	1 (1%)	15	55
3	A	69/83 (83%)	68 (99%)	1 (1%)	0	100	100
3	F	69/83 (83%)	68 (99%)	1 (1%)	0	100	100
3	G	69/83 (83%)	68 (99%)	1 (1%)	0	100	100
3	L	69/83 (83%)	68 (99%)	1 (1%)	0	100	100
All	All	1325/1568 (84%)	1262 (95%)	49 (4%)	14 (1%)	14	52

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	123	ASN
1	B	124	ASP
2	C	401	LEU
1	E	123	ASN
1	E	124	ASP

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	B	113/135 (84%)	110 (97%)	3 (3%)	44	65
1	E	113/135 (84%)	110 (97%)	3 (3%)	44	65
1	H	113/135 (84%)	110 (97%)	3 (3%)	44	65
1	K	113/135 (84%)	110 (97%)	3 (3%)	44	65
2	C	74/140 (53%)	72 (97%)	2 (3%)	44	65
2	D	74/140 (53%)	74 (100%)	0	100	100
2	I	74/140 (53%)	72 (97%)	2 (3%)	44	65
2	J	74/140 (53%)	74 (100%)	0	100	100
3	A	33/73 (45%)	33 (100%)	0	100	100
3	F	33/73 (45%)	33 (100%)	0	100	100
3	G	33/73 (45%)	33 (100%)	0	100	100
3	L	33/73 (45%)	33 (100%)	0	100	100
All	All	880/1392 (63%)	864 (98%)	16 (2%)	59	77

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	K	103	THR
1	K	93	ASP
1	H	93	ASP
2	I	429	GLU
1	E	135	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 8 ligands modelled in this entry, 8 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 (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)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.