



wwPDB X-ray Structure Validation Summary Report

Jul 31, 2014 – 06:55 PM EDT

PDB ID : 4UQI
Title : AP2 controls clathrin polymerization with a membrane-activated switch
Authors : Kelly, B.T.; Graham, S.C.; Liska, N.; Dannhauser, P.N.; Hoening, S.;
Ungewickell, E.J.; Owen, D.J.
Deposited on : 2014-06-23
Resolution : 2.79 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

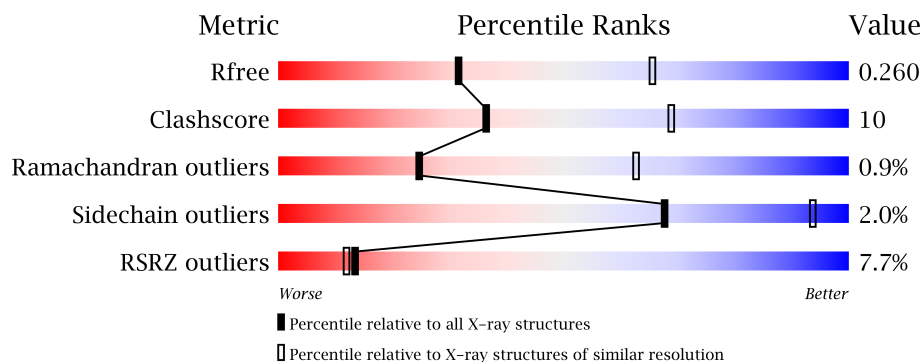
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.16 November 2013
Xtriage (Phenix)	:	dev-1439
EDS	:	stable23489
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable23489

1 Overall quality at a glance

The reported resolution of this entry is 2.79 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}	66092	1799 (2.80-2.80)
Clashscore	79885	2295 (2.80-2.80)
Ramachandran outliers	78287	2252 (2.80-2.80)
Sidechain outliers	78261	2254 (2.80-2.80)
RSRZ outliers	66119	1802 (2.80-2.80)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	628	
2	B	657	
3	M	446	
4	S	142	

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 13889 atoms, of which 0 are hydrogen and 0 are deuterium.

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 AP-2 COMPLEX SUBUNIT ALPHA-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	600	Total	C	N	O	S	0	0	0
			4737	3017	815	884	21			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	8	ASP	GLU	CONFLICT	UNP P18484
A	272	GLU	-	INSERTION	UNP P18484
A	622	GLY	-	EXPRESSION TAG	UNP P18484
A	623	SER	-	EXPRESSION TAG	UNP P18484
A	624	GLY	-	EXPRESSION TAG	UNP P18484
A	625	LEU	-	EXPRESSION TAG	UNP P18484
A	626	VAL	-	EXPRESSION TAG	UNP P18484
A	627	PRO	-	EXPRESSION TAG	UNP P18484
A	628	ARG	-	EXPRESSION TAG	UNP P18484

- Molecule 2 is a protein called AP-2 COMPLEX SUBUNIT BETA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	597	Total	C	N	O	S	0	0	0
			4720	3007	783	905	25			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	652	HIS	-	EXPRESSION TAG	UNP P63010
B	653	HIS	-	EXPRESSION TAG	UNP P63010
B	654	HIS	-	EXPRESSION TAG	UNP P63010
B	655	HIS	-	EXPRESSION TAG	UNP P63010
B	656	HIS	-	EXPRESSION TAG	UNP P63010
B	657	HIS	-	EXPRESSION TAG	UNP P63010

- Molecule 3 is a protein called AP-2 COMPLEX SUBUNIT MU.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	M	396	Total	C	N	O	S	0	0	0
			3192	2053	559	561	19			

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	236A	MET	-	INSERTION	UNP P84092
M	236B	GLU	-	INSERTION	UNP P84092
M	236C	GLN	-	INSERTION	UNP P84092
M	236D	LYS	-	INSERTION	UNP P84092
M	236E	LEU	-	INSERTION	UNP P84092
M	236F	ILE	-	INSERTION	UNP P84092
M	236G	SER	-	INSERTION	UNP P84092
M	236H	GLU	-	INSERTION	UNP P84092
M	236I	GLU	-	INSERTION	UNP P84092
M	236J	ASP	-	INSERTION	UNP P84092
M	236K	LEU	-	INSERTION	UNP P84092

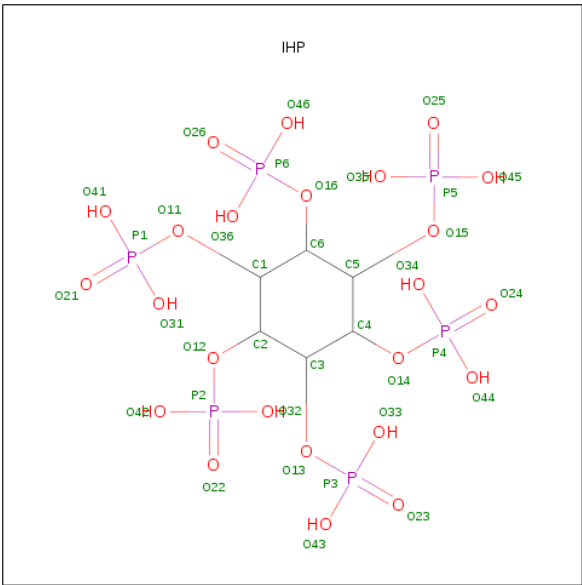
- Molecule 4 is a protein called AP-2 COMPLEX SUBUNIT SIGMA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	S	142	Total	C	N	O	S	0	0	0
			1200	778	200	215	7			

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Cl	0	0
			1	1		

- Molecule 6 is INOSITOL HEXAKISPHOSPHATE (three-letter code: IHP) (formula: C₆H₁₈O₂₄P₆).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	O	P	0	0
			36	6	24	6		

- Molecule 7 is water.

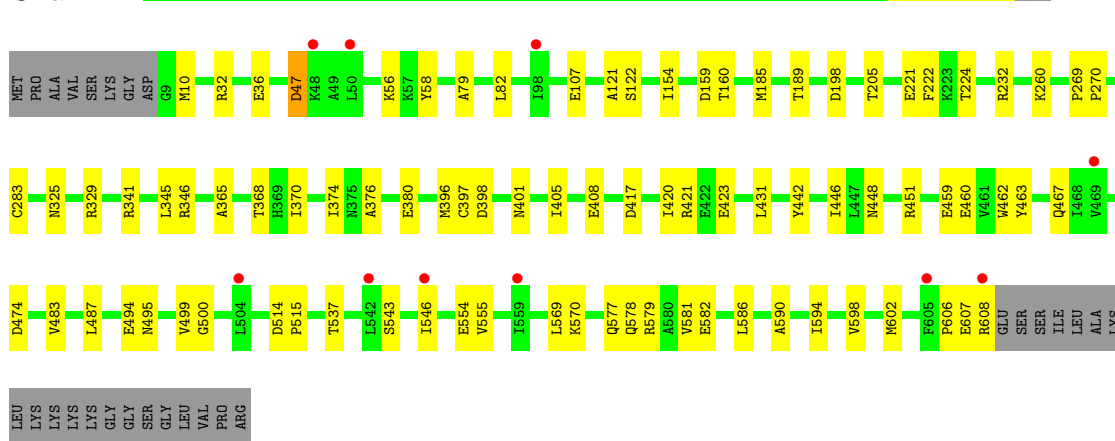
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	1	Total	O	0	0
			1	1		
7	B	2	Total	O	0	0
			2	2		

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.

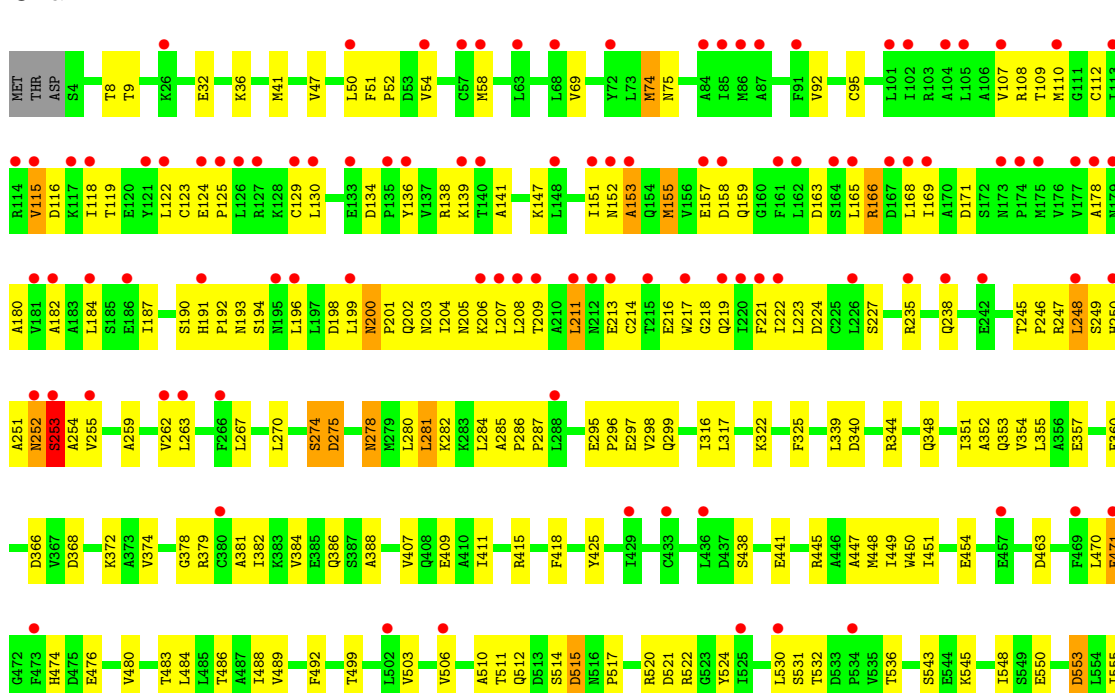
• Molecule 1: AP-2 COMPLEX SUBUNIT ALPHA-2

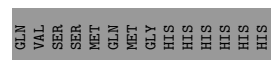
Chain A:



• Molecule 2: AP-2 COMPLEX SUBUNIT BETA

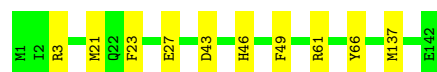
Chain B:





Chain M:

Chain S:



4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	121.30Å 121.30Å 259.37Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	97.37 – 2.79 97.37 – 2.79	Depositor EDS
% Data completeness (in resolution range)	99.9 (97.37-2.79) 99.9 (97.37-2.79)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.73 (at 2.77Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.203 , 0.259 0.206 , 0.260	Depositor DCC
R_{free} test set	2824 reflections (5.34%)	DCC
Wilson B-factor (Å ²)	74.2	Xtriage
Anisotropy	0.245	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 54.7	EDS
Estimated twinning fraction	0.021 for -h,-k,l	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 55666 reflections (0.002%)	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	13889	wwPDB-VP
Average B, all atoms (Å ²)	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.65% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: IHP, CL

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	0.65	0/4822	0.76	0/6541
2	B	0.53	0/4793	0.72	1/6503 (0.0%)
3	M	0.64	0/3255	0.82	2/4382 (0.0%)
4	S	0.78	0/1224	0.80	0/1650
All	All	0.62	0/14094	0.76	3/19076 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	M	162	ARG	NE-CZ-NH1	6.73	123.67	120.30
3	M	323	ARG	NE-CZ-NH1	6.46	123.53	120.30
2	B	366	ASP	CB-CG-OD1	5.36	123.12	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4737	0	4830	64	0
2	B	4720	0	4831	180	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	M	3192	0	3295	55	0
4	S	1200	0	1195	6	0
5	A	1	0	0	0	0
6	A	36	0	6	1	0
7	A	1	0	0	0	0
7	B	2	0	0	0	0
All	All	13889	0	14157	290	1

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 10.

The worst 5 of 290 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:198:ASP:O	1:A:232:ARG:NH2	2.01	0.92
1:A:495:ASN:O	1:A:499:VAL:HG23	1.71	0.88
2:B:204:ILE:HG22	2:B:208:LEU:HD12	1.59	0.85
1:A:365:ALA:O	1:A:368:THR:HB	1.75	0.84
2:B:211:LEU:O	2:B:211:LEU:HG	1.76	0.84

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	Distance(Å)	Clash(Å)
2:B:213:GLU:O	2:B:213:GLU:O[4.555]	2.10	0.10

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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	598/628 (95%)	581 (97%)	17 (3%)	0	100	100
2	B	593/657 (90%)	522 (88%)	57 (10%)	14 (2%)	9	29
3	M	388/446 (87%)	364 (94%)	22 (6%)	2 (0%)	38	76

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	S	140/142 (99%)	135 (96%)	5 (4%)	0	100	100
All	All	1719/1873 (92%)	1602 (93%)	101 (6%)	16 (1%)	25	63

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	200	ASN
2	B	211	LEU
2	B	247	ARG
2	B	252	ASN
2	B	253	SER

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	526/548 (96%)	521 (99%)	5 (1%)	85	98
2	B	535/589 (91%)	516 (96%)	19 (4%)	47	82
3	M	353/398 (89%)	347 (98%)	6 (2%)	73	95
4	S	131/131 (100%)	130 (99%)	1 (1%)	89	98
All	All	1545/1666 (93%)	1514 (98%)	31 (2%)	68	94

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	253	SER
2	B	282	LYS
3	M	337	CYS
2	B	278	ASN
2	B	438	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	250	HIS
2	B	305	ASN
3	M	135	GLN
2	B	203	ASN
3	M	72	ASN

5.3.3 RNA ⓘ

There are no RNA chains 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 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
6	IHP	A	1609	-	36,36,36	2.64	13 (36%)	60,60,60	2.86	22 (36%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	IHP	A	1609	-	-	0/30/54/54	0/1/1/1

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	1609	IHP	P1-O41	5.04	1.72	1.54
6	A	1609	IHP	P1-O31	4.94	1.72	1.54
6	A	1609	IHP	P5-O35	4.77	1.72	1.54
6	A	1609	IHP	P5-O45	4.74	1.71	1.54
6	A	1609	IHP	P4-O44	4.69	1.71	1.54

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	1609	IHP	O11-C1-C6	-8.24	89.08	108.29
6	A	1609	IHP	P5-O15-C5	-7.74	103.90	121.77
6	A	1609	IHP	O16-C6-C5	7.10	124.83	108.29
6	A	1609	IHP	P6-O16-C6	6.95	137.81	121.77
6	A	1609	IHP	C6-C1-C2	6.05	122.96	110.44

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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, 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	600/628 (95%)	0.28	10 (1%) 67 68	42, 72, 97, 130	0
2	B	597/657 (90%)	0.98	108 (18%) 2 2	69, 103, 136, 149	18 (3%)
3	M	396/446 (88%)	0.38	16 (4%) 36 37	48, 78, 111, 152	0
4	S	142/142 (100%)	0.23	0 100 100	45, 55, 74, 96	0
All	All	1735/1873 (92%)	0.54	134 (7%) 13 12	42, 81, 128, 152	18 (1%)

The worst 5 of 134 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	168	LEU	8.2
2	B	169	ILE	7.9
2	B	530	LEU	7.2
2	B	122	LEU	6.2
2	B	174	PRO	6.1

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 ⓘ

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, 95th 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	RSR	LLDF	B-factors(\AA^2)	Q<0.9
6	IHP	A	1609	36/36	0.12	-2.07	60,74,80,82	0
5	CL	A	1608	1/1	0.13	-3.22	60,60,60,60	0

6.5 Other polymers ⓘ

There are no such residues in this entry.