



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 23, 2017 – 10:52 PM EDT

PDB ID : 4UBF
Title : HsMCAK motor domain complex
Authors : Welburn, J.P.I.; Talapatra, S.K.
Deposited on : unknown
Resolution : 3.00 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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20029824
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
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-20029824

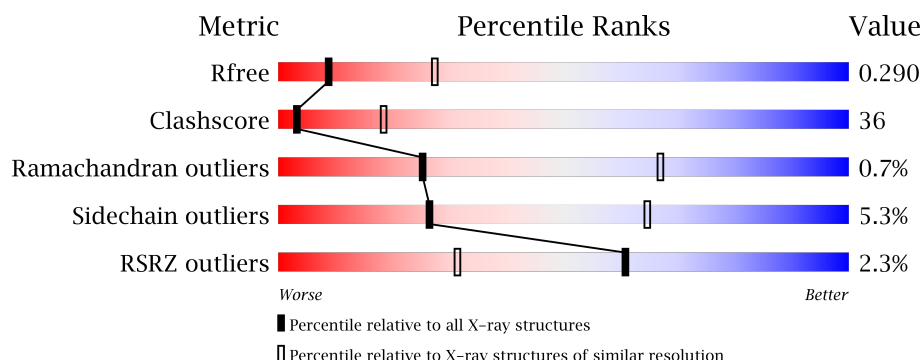
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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}	100719	1692 (3.00-3.00)
Clashscore	112137	2037 (3.00-3.00)
Ramachandran outliers	110173	1973 (3.00-3.00)
Sidechain outliers	110143	1976 (3.00-3.00)
RSRZ outliers	101464	1716 (3.00-3.00)

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	387	<div> <div>2%</div> <div> <div></div> <div>46%</div> <div>31%</div> <div>•</div> <div>20%</div> </div> </div>
1	B	387	<div> <div>2%</div> <div> <div></div> <div>43%</div> <div>36%</div> <div>••</div> <div>18%</div> </div> </div>
1	C	387	<div> <div>2%</div> <div> <div></div> <div>50%</div> <div>30%</div> <div>•</div> <div>17%</div> </div> </div>
1	D	387	<div> <div>3%</div> <div> <div></div> <div>38%</div> <div>40%</div> <div>5%</div> <div>•</div> <div>17%</div> </div> </div>
2	P	12	<div> <div></div> <div> <div>17%</div> <div>33%</div> <div>8%</div> <div>42%</div> </div> </div>

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 9707 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 Kinesin-like protein KIF2C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	309	Total	C	N	O	S	0	3	0
			2322	1490	392	419	21			
1	B	319	Total	C	N	O	S	0	1	0
			2404	1537	412	434	21			
1	C	322	Total	C	N	O	S	0	1	0
			2412	1546	401	444	21			
1	D	322	Total	C	N	O	S	0	0	0
			2337	1501	389	428	19			

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	207	MET	-	expression tag	UNP Q99661
A	208	GLY	-	expression tag	UNP Q99661
A	209	SER	-	expression tag	UNP Q99661
A	210	SER	-	expression tag	UNP Q99661
A	211	HIS	-	expression tag	UNP Q99661
A	212	HIS	-	expression tag	UNP Q99661
A	213	HIS	-	expression tag	UNP Q99661
A	214	HIS	-	expression tag	UNP Q99661
A	215	HIS	-	expression tag	UNP Q99661
A	216	HIS	-	expression tag	UNP Q99661
A	217	SER	-	expression tag	UNP Q99661
A	218	SER	-	expression tag	UNP Q99661
A	219	GLY	-	expression tag	UNP Q99661
A	220	LEU	-	expression tag	UNP Q99661
A	221	VAL	-	expression tag	UNP Q99661
A	222	PRO	-	expression tag	UNP Q99661
A	223	ARG	-	expression tag	UNP Q99661
A	224	GLY	-	expression tag	UNP Q99661
B	207	MET	-	expression tag	UNP Q99661
B	208	GLY	-	expression tag	UNP Q99661
B	209	SER	-	expression tag	UNP Q99661

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	210	SER	-	expression tag	UNP Q99661
B	211	HIS	-	expression tag	UNP Q99661
B	212	HIS	-	expression tag	UNP Q99661
B	213	HIS	-	expression tag	UNP Q99661
B	214	HIS	-	expression tag	UNP Q99661
B	215	HIS	-	expression tag	UNP Q99661
B	216	HIS	-	expression tag	UNP Q99661
B	217	SER	-	expression tag	UNP Q99661
B	218	SER	-	expression tag	UNP Q99661
B	219	GLY	-	expression tag	UNP Q99661
B	220	LEU	-	expression tag	UNP Q99661
B	221	VAL	-	expression tag	UNP Q99661
B	222	PRO	-	expression tag	UNP Q99661
B	223	ARG	-	expression tag	UNP Q99661
B	224	GLY	-	expression tag	UNP Q99661
C	207	MET	-	expression tag	UNP Q99661
C	208	GLY	-	expression tag	UNP Q99661
C	209	SER	-	expression tag	UNP Q99661
C	210	SER	-	expression tag	UNP Q99661
C	211	HIS	-	expression tag	UNP Q99661
C	212	HIS	-	expression tag	UNP Q99661
C	213	HIS	-	expression tag	UNP Q99661
C	214	HIS	-	expression tag	UNP Q99661
C	215	HIS	-	expression tag	UNP Q99661
C	216	HIS	-	expression tag	UNP Q99661
C	217	SER	-	expression tag	UNP Q99661
C	218	SER	-	expression tag	UNP Q99661
C	219	GLY	-	expression tag	UNP Q99661
C	220	LEU	-	expression tag	UNP Q99661
C	221	VAL	-	expression tag	UNP Q99661
C	222	PRO	-	expression tag	UNP Q99661
C	223	ARG	-	expression tag	UNP Q99661
C	224	GLY	-	expression tag	UNP Q99661
D	207	MET	-	expression tag	UNP Q99661
D	208	GLY	-	expression tag	UNP Q99661
D	209	SER	-	expression tag	UNP Q99661
D	210	SER	-	expression tag	UNP Q99661
D	211	HIS	-	expression tag	UNP Q99661
D	212	HIS	-	expression tag	UNP Q99661
D	213	HIS	-	expression tag	UNP Q99661
D	214	HIS	-	expression tag	UNP Q99661
D	215	HIS	-	expression tag	UNP Q99661

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	216	HIS	-	expression tag	UNP Q99661
D	217	SER	-	expression tag	UNP Q99661
D	218	SER	-	expression tag	UNP Q99661
D	219	GLY	-	expression tag	UNP Q99661
D	220	LEU	-	expression tag	UNP Q99661
D	221	VAL	-	expression tag	UNP Q99661
D	222	PRO	-	expression tag	UNP Q99661
D	223	ARG	-	expression tag	UNP Q99661
D	224	GLY	-	expression tag	UNP Q99661

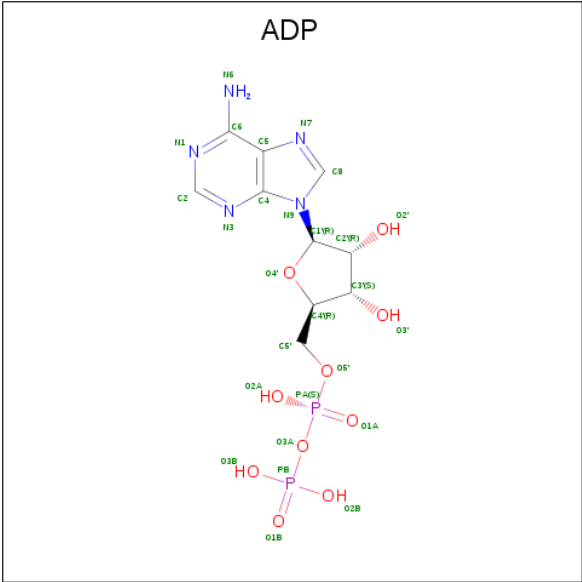
- Molecule 2 is a protein called Kinesin-like protein KIF2C.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	P	7	Total	C	N	O	0	0	0
			48	27	8	13			

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	2	Total	Mg	0	0
			2	2		
3	A	2	Total	Mg	0	0
			2	2		
3	D	1	Total	Mg	0	0
			1	1		
3	C	2	Total	Mg	0	0
			2	2		

- Molecule 4 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula: C₁₀H₁₅N₅O₁₀P₂).

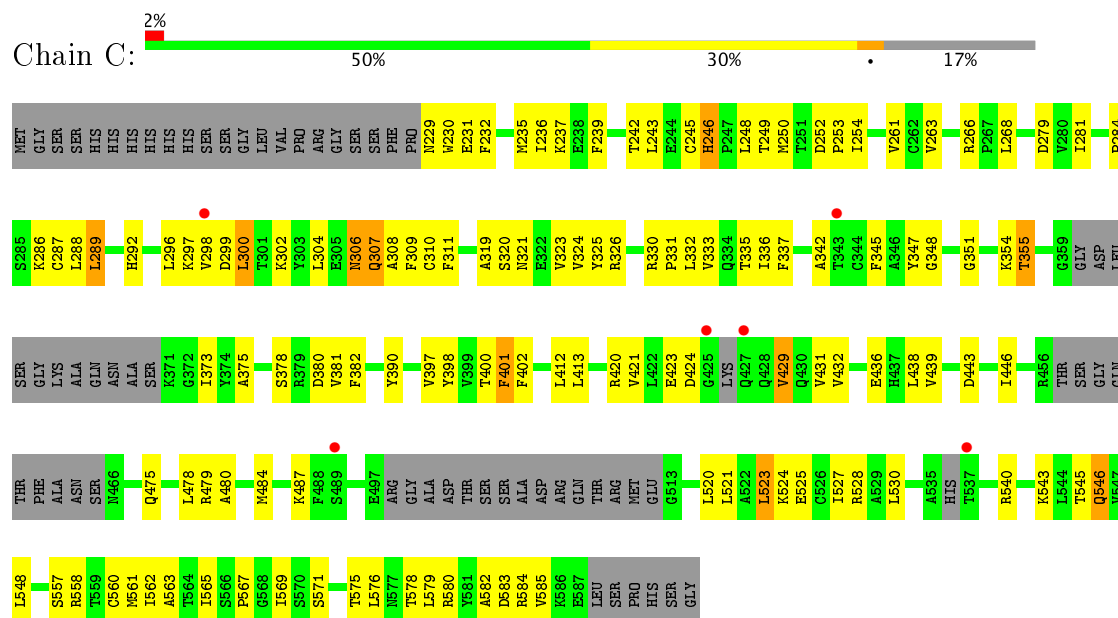


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	0
			27	10	5	10	2		
4	B	1	Total	C	N	O	P	0	0
			27	10	5	10	2		
4	C	1	Total	C	N	O	P	0	0
			27	10	5	10	2		
4	D	1	Total	C	N	O	P	0	0
			27	10	5	10	2		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	14	Total	O	0	0
			14	14		
5	B	20	Total	O	0	0
			20	20		
5	C	17	Total	O	0	0
			17	17		
5	D	16	Total	O	0	0
			16	16		
5	P	2	Total	O	0	0
			2	2		

- Molecule 1: Kinesin-like protein KIF2C



GLOBAL-STATISTICS INFOmissingINFO

4 Model quality [i](#)

4.1 Standard geometry [i](#)

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

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	0.39	0/2368	0.64	5/3198 (0.2%)
1	B	0.47	1/2443 (0.0%)	0.87	9/3295 (0.3%)
1	C	0.32	1/2451 (0.0%)	0.63	7/3307 (0.2%)
1	D	0.37	2/2372 (0.1%)	0.78	13/3208 (0.4%)
2	P	0.35	0/47	1.39	0/62
All	All	0.39	4/9681 (0.0%)	0.74	34/13070 (0.3%)

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	B	0	2
1	C	0	1
1	D	0	2
All	All	0	5

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	229	ASN	C-N	-14.82	0.99	1.34
1	D	247	PRO	N-CD	5.13	1.55	1.47
1	D	538	PRO	N-CD	5.10	1.54	1.47
1	C	253	PRO	N-CD	5.01	1.54	1.47

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	229	ASN	O-C-N	-19.66	91.24	122.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	229	ASN	C-N-CA	18.62	168.25	121.70
1	B	229	ASN	CA-C-N	13.64	147.21	117.20
1	B	230	TRP	N-CA-C	9.43	136.45	111.00
1	B	230	TRP	N-CA-CB	-9.42	93.64	110.60

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	229	ASN	Mainchain,Peptide
1	C	560[A]	CYS	Mainchain
1	D	330	ARG	Peptide
1	D	510	ARG	Peptide

4.2 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	2322	0	2270	115	0
1	B	2404	0	2346	220	0
1	C	2412	0	2331	127	0
1	D	2337	0	2214	226	0
2	P	48	0	33	5	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
3	C	2	0	0	0	0
3	D	1	0	0	0	0
4	A	27	0	12	5	0
4	B	27	0	12	1	0
4	C	27	0	12	3	0
4	D	27	0	12	7	0
5	A	14	0	0	1	0
5	B	20	0	0	3	0
5	C	17	0	0	5	0
5	D	16	0	0	7	0
5	P	2	0	0	1	0
All	All	9707	0	9242	684	0

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 36.

The worst 5 of 684 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:B:413:LEU:HD12	1:B:437:HIS:CD2	1.14	1.61
1:B:413:LEU:CD1	1:B:437:HIS:HD2	1.02	1.59
1:D:395:LEU:HD12	1:D:480:ALA:CA	1.25	1.55
1:D:395:LEU:CD1	1:D:480:ALA:HA	1.36	1.51
1:D:384:LEU:HD21	1:D:390:TYR:CE2	1.48	1.47

There are no symmetry-related clashes.

4.3 Torsion angles [i](#)

4.3.1 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	300/387 (78%)	261 (87%)	36 (12%)	3 (1%)	18	59
1	B	306/387 (79%)	265 (87%)	40 (13%)	1 (0%)	44	81
1	C	309/387 (80%)	278 (90%)	30 (10%)	1 (0%)	44	81
1	D	308/387 (80%)	266 (86%)	38 (12%)	4 (1%)	14	51
2	P	5/12 (42%)	4 (80%)	1 (20%)	0	100	100
All	All	1228/1560 (79%)	1074 (88%)	145 (12%)	9 (1%)	25	67

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	230	TRP
1	A	353	GLY
1	A	455[A]	CYS
1	A	455[B]	CYS
1	D	330	ARG

4.3.2 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	238/332 (72%)	228 (96%)	10 (4%)	34	73
1	B	244/332 (74%)	231 (95%)	13 (5%)	26	65
1	C	244/332 (74%)	233 (96%)	11 (4%)	32	71
1	D	225/332 (68%)	211 (94%)	14 (6%)	21	58
2	P	4/11 (36%)	2 (50%)	2 (50%)	0	0
All	All	955/1339 (71%)	905 (95%)	50 (5%)	26	65

5 of 50 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	414	ASN
1	C	307	GLN
1	D	537	THR
1	C	246	HIS
1	C	300	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	306	ASN
1	D	546	GLN
1	D	356	HIS
1	A	556	ASN
1	C	475	GLN

4.3.3 RNA [i](#)

There are no RNA molecules in this entry.

4.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

4.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

4.6 Ligand geometry ⓘ

Of 11 ligands modelled in this entry, 7 are monoatomic - leaving 4 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$
4	ADP	A	602	3	25,29,29	0.99	1 (4%)	24,45,45	1.56	2 (8%)
4	ADP	B	601	3	25,29,29	1.03	1 (4%)	24,45,45	1.61	2 (8%)
4	ADP	C	602	-	25,29,29	1.01	1 (4%)	24,45,45	1.61	2 (8%)
4	ADP	D	602	3	25,29,29	1.03	1 (4%)	24,45,45	1.61	2 (8%)

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
4	ADP	A	602	3	-	0/12/32/32	0/3/3/3
4	ADP	B	601	3	-	0/12/32/32	0/3/3/3
4	ADP	C	602	-	-	0/12/32/32	0/3/3/3
4	ADP	D	602	3	-	0/12/32/32	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	602	ADP	C5-C4	3.17	1.47	1.40
4	D	602	ADP	C5-C4	3.18	1.47	1.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	601	ADP	C5-C4	3.20	1.47	1.40
4	A	602	ADP	C5-C4	3.21	1.47	1.40

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	602	ADP	N3-C2-N1	-5.85	123.76	128.86
4	D	602	ADP	N3-C2-N1	-5.82	123.79	128.86
4	B	601	ADP	N3-C2-N1	-5.80	123.81	128.86
4	A	602	ADP	N3-C2-N1	-5.42	124.14	128.86
4	C	602	ADP	C4-C5-N7	-3.07	106.45	109.41

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	602	ADP	5	0
4	B	601	ADP	1	0
4	C	602	ADP	3	0
4	D	602	ADP	7	0

4.7 Other polymers [i](#)

There are no such residues in this entry.

4.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	229:ASN	C	230:TRP	N	1.00

5 Fit of model and data [i](#)

5.1 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(Å ²)	Q<0.9
1	A	309/387 (79%)	-0.19	6 (1%) 67 37	28, 65, 100, 114	0
1	B	319/387 (82%)	-0.07	6 (1%) 67 37	29, 64, 93, 113	0
1	C	322/387 (83%)	-0.19	6 (1%) 67 37	25, 74, 118, 149	0
1	D	322/387 (83%)	0.09	11 (3%) 46 20	30, 77, 117, 129	0
2	P	7/12 (58%)	-0.43	0 100 100	43, 54, 66, 67	0
All	All	1279/1560 (81%)	-0.09	29 (2%) 61 31	25, 70, 111, 149	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	537	THR	4.1
1	D	344	CYS	4.0
1	A	533[A]	ASN	3.3
1	C	489	SER	3.2
1	D	489	SER	3.1

5.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.4 Ligands [i](#)

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	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
3	MG	C	601	1/1	0.88	0.20	1.89	53,53,53,53	0
4	ADP	B	601	27/27	0.93	0.20	0.30	38,50,57,59	0
4	ADP	C	602	27/27	0.94	0.15	-0.09	48,66,75,81	0
4	ADP	A	602	27/27	0.90	0.16	-0.74	54,79,94,105	0
4	ADP	D	602	27/27	0.93	0.16	-1.05	49,74,85,92	0
3	MG	D	601	1/1	0.86	0.13	-1.97	49,49,49,49	0
3	MG	A	603	1/1	0.98	0.08	-3.51	42,42,42,42	0
3	MG	C	603	1/1	0.65	0.27	-	78,78,78,78	0
3	MG	B	602	1/1	0.96	0.30	-	41,41,41,41	0
3	MG	B	603	1/1	0.99	0.11	-	30,30,30,30	0
3	MG	A	601	1/1	0.93	0.07	-	54,54,54,54	0

5.5 Other polymers ⓘ

There are no such residues in this entry.