



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 30, 2024 – 11:59 pm BST

PDB ID : 4UMP
Title : Structure of MELK in complex with inhibitors
Authors : Johnson, C.N.; Berdini, V.; Beke, L.; Bonnet, P.; Brehmer, D.; Coyle, J.E.; Day, P.J.; Frederickson, M.; Freyne, E.J.E.; Gilissen, R.A.H.J.; Hamlett, C.C.F.; Howard, S.; Meerpoel, L.; McMenamin, R.; Patel, S.; Rees, D.C.; Sharff, A.; Sommen, F.; Wu, T.; Linders, J.T.M.
Deposited on : 2014-05-20
Resolution : 2.30 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
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)	:	2.36.2

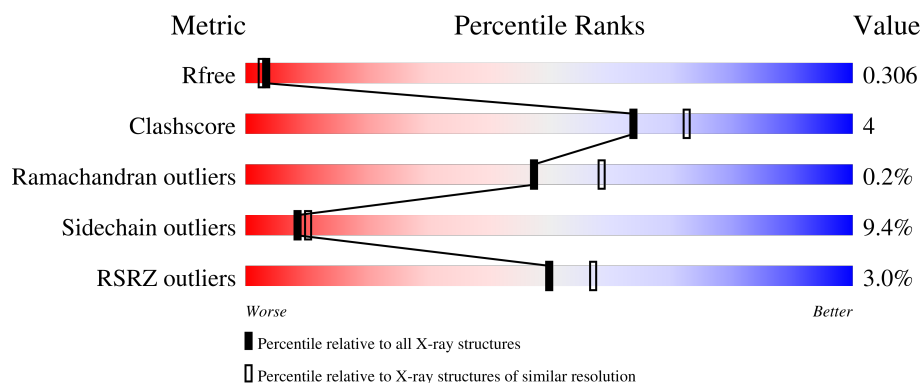
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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 $\leq 5\%$. 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	356	<div> <div>0%</div> <div>75%</div> <div>11%</div> <div>•</div> <div>12%</div> </div>
1	B	356	<div> <div>2%</div> <div>74%</div> <div>15%</div> <div>•</div> <div>9%</div> </div>
1	C	356	<div> <div>3%</div> <div>73%</div> <div>14%</div> <div>•</div> <div>13%</div> </div>
1	D	356	<div> <div>5%</div> <div>73%</div> <div>15%</div> <div>•</div> <div>10%</div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 10944 atoms, of which 36 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 MATERNAL EMBRYONIC LEUCINE ZIPPER KINASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	313	Total	C	N	O	S	0	1	0
			2545	1642	430	456	17			
1	B	324	Total	C	N	O	S	0	0	0
			2614	1680	444	472	18			
1	C	311	Total	C	N	O	S	0	1	0
			2536	1636	427	456	17			
1	D	321	Total	C	N	O	S	0	1	0
			2593	1672	439	467	15			

There are 108 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	expression tag	UNP Q14680
A	-18	GLY	-	expression tag	UNP Q14680
A	-17	SER	-	expression tag	UNP Q14680
A	-16	SER	-	expression tag	UNP Q14680
A	-15	HIS	-	expression tag	UNP Q14680
A	-14	HIS	-	expression tag	UNP Q14680
A	-13	HIS	-	expression tag	UNP Q14680
A	-12	HIS	-	expression tag	UNP Q14680
A	-11	HIS	-	expression tag	UNP Q14680
A	-10	HIS	-	expression tag	UNP Q14680
A	-9	SER	-	expression tag	UNP Q14680
A	-8	SER	-	expression tag	UNP Q14680
A	-7	GLY	-	expression tag	UNP Q14680
A	-6	LEU	-	expression tag	UNP Q14680
A	-5	VAL	-	expression tag	UNP Q14680
A	-4	PRO	-	expression tag	UNP Q14680
A	-3	ARG	-	expression tag	UNP Q14680
A	-2	GLY	-	expression tag	UNP Q14680
A	-1	SER	-	expression tag	UNP Q14680
A	0	HIS	-	expression tag	UNP Q14680
A	167	ALA	THR	engineered mutation	UNP Q14680

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	171	ALA	SER	engineered mutation	UNP Q14680
A	213	THR	ASN	engineered mutation	UNP Q14680
A	214	ALA	VAL	engineered mutation	UNP Q14680
A	215	ALA	MET	engineered mutation	UNP Q14680
A	218	VAL	TYR	engineered mutation	UNP Q14680
A	219	ALA	LYS	engineered mutation	UNP Q14680
B	-19	MET	-	expression tag	UNP Q14680
B	-18	GLY	-	expression tag	UNP Q14680
B	-17	SER	-	expression tag	UNP Q14680
B	-16	SER	-	expression tag	UNP Q14680
B	-15	HIS	-	expression tag	UNP Q14680
B	-14	HIS	-	expression tag	UNP Q14680
B	-13	HIS	-	expression tag	UNP Q14680
B	-12	HIS	-	expression tag	UNP Q14680
B	-11	HIS	-	expression tag	UNP Q14680
B	-10	HIS	-	expression tag	UNP Q14680
B	-9	SER	-	expression tag	UNP Q14680
B	-8	SER	-	expression tag	UNP Q14680
B	-7	GLY	-	expression tag	UNP Q14680
B	-6	LEU	-	expression tag	UNP Q14680
B	-5	VAL	-	expression tag	UNP Q14680
B	-4	PRO	-	expression tag	UNP Q14680
B	-3	ARG	-	expression tag	UNP Q14680
B	-2	GLY	-	expression tag	UNP Q14680
B	-1	SER	-	expression tag	UNP Q14680
B	0	HIS	-	expression tag	UNP Q14680
B	167	ALA	THR	engineered mutation	UNP Q14680
B	171	ALA	SER	engineered mutation	UNP Q14680
B	213	THR	ASN	engineered mutation	UNP Q14680
B	214	ALA	VAL	engineered mutation	UNP Q14680
B	215	ALA	MET	engineered mutation	UNP Q14680
B	218	VAL	TYR	engineered mutation	UNP Q14680
B	219	ALA	LYS	engineered mutation	UNP Q14680
C	-19	MET	-	expression tag	UNP Q14680
C	-18	GLY	-	expression tag	UNP Q14680
C	-17	SER	-	expression tag	UNP Q14680
C	-16	SER	-	expression tag	UNP Q14680
C	-15	HIS	-	expression tag	UNP Q14680
C	-14	HIS	-	expression tag	UNP Q14680
C	-13	HIS	-	expression tag	UNP Q14680
C	-12	HIS	-	expression tag	UNP Q14680
C	-11	HIS	-	expression tag	UNP Q14680

Continued on next page...

Continued from previous page...

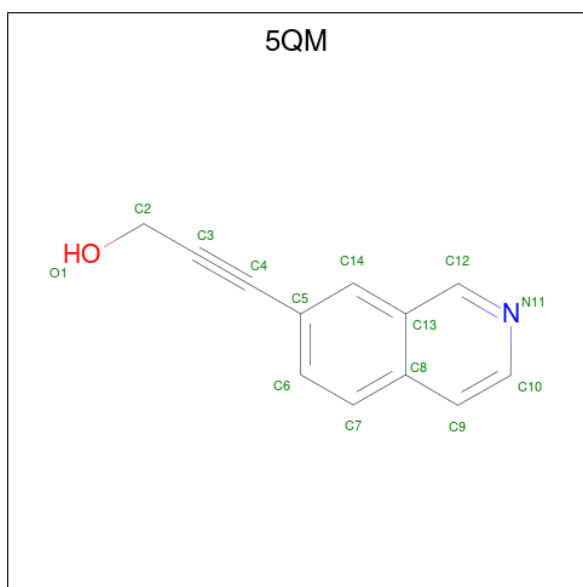
Chain	Residue	Modelled	Actual	Comment	Reference
C	-10	HIS	-	expression tag	UNP Q14680
C	-9	SER	-	expression tag	UNP Q14680
C	-8	SER	-	expression tag	UNP Q14680
C	-7	GLY	-	expression tag	UNP Q14680
C	-6	LEU	-	expression tag	UNP Q14680
C	-5	VAL	-	expression tag	UNP Q14680
C	-4	PRO	-	expression tag	UNP Q14680
C	-3	ARG	-	expression tag	UNP Q14680
C	-2	GLY	-	expression tag	UNP Q14680
C	-1	SER	-	expression tag	UNP Q14680
C	0	HIS	-	expression tag	UNP Q14680
C	167	ALA	THR	engineered mutation	UNP Q14680
C	171	ALA	SER	engineered mutation	UNP Q14680
C	213	THR	ASN	engineered mutation	UNP Q14680
C	214	ALA	VAL	engineered mutation	UNP Q14680
C	215	ALA	MET	engineered mutation	UNP Q14680
C	218	VAL	TYR	engineered mutation	UNP Q14680
C	219	ALA	LYS	engineered mutation	UNP Q14680
D	-19	MET	-	expression tag	UNP Q14680
D	-18	GLY	-	expression tag	UNP Q14680
D	-17	SER	-	expression tag	UNP Q14680
D	-16	SER	-	expression tag	UNP Q14680
D	-15	HIS	-	expression tag	UNP Q14680
D	-14	HIS	-	expression tag	UNP Q14680
D	-13	HIS	-	expression tag	UNP Q14680
D	-12	HIS	-	expression tag	UNP Q14680
D	-11	HIS	-	expression tag	UNP Q14680
D	-10	HIS	-	expression tag	UNP Q14680
D	-9	SER	-	expression tag	UNP Q14680
D	-8	SER	-	expression tag	UNP Q14680
D	-7	GLY	-	expression tag	UNP Q14680
D	-6	LEU	-	expression tag	UNP Q14680
D	-5	VAL	-	expression tag	UNP Q14680
D	-4	PRO	-	expression tag	UNP Q14680
D	-3	ARG	-	expression tag	UNP Q14680
D	-2	GLY	-	expression tag	UNP Q14680
D	-1	SER	-	expression tag	UNP Q14680
D	0	HIS	-	expression tag	UNP Q14680
D	167	ALA	THR	engineered mutation	UNP Q14680
D	171	ALA	SER	engineered mutation	UNP Q14680
D	213	THR	ASN	engineered mutation	UNP Q14680
D	214	ALA	VAL	engineered mutation	UNP Q14680

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	215	ALA	MET	engineered mutation	UNP Q14680
D	218	VAL	TYR	engineered mutation	UNP Q14680
D	219	ALA	LYS	engineered mutation	UNP Q14680

- Molecule 2 is 3-(isoquinolin-7-yl)prop-2-yn-1-ol (three-letter code: 5QM) (formula: C₁₂H₉NO).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	H	N	O	0	0
			23	12	9	1	1		
2	B	1	Total	C	H	N	O	0	0
			23	12	9	1	1		
2	C	1	Total	C	H	N	O	0	0
			23	12	9	1	1		
2	D	1	Total	C	H	N	O	0	0
			23	12	9	1	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	143	Total	O	0	0
			143	143		
3	B	162	Total	O	0	0
			162	162		
3	C	130	Total	O	0	0
			130	130		

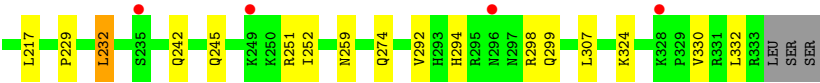
Continued on next page...

Continued from previous page...

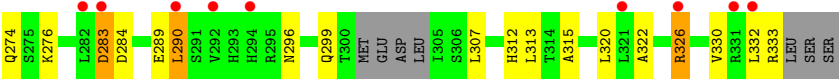
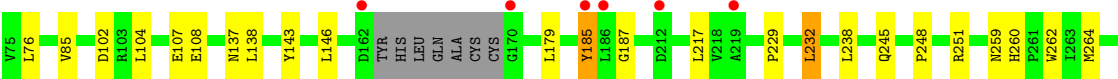
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	D	129	Total 129	O 129	0	0

● Molecule 1: MATERNAL EMBRYONIC LEUCINE ZIPPER KINASE





● Molecule 1: MATERNAL EMBRYONIC LEUCINE ZIPPER KINASE



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	65.82Å 74.97Å 77.09Å 86.47° 70.06° 89.94°	Depositor
Resolution (Å)	50.38 – 2.30 50.38 – 2.30	Depositor EDS
% Data completeness (in resolution range)	95.7 (50.38-2.30) 95.8 (50.38-2.30)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.24 (at 2.29Å)	Xtriage
Refinement program	BUSTER 2.11.5	Depositor
R, R_{free}	0.235 , 0.291 0.243 , 0.306	Depositor DCC
R_{free} test set	2989 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å ²)	50.3	Xtriage
Anisotropy	0.241	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 57.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	10944	wwPDB-VP
Average B, all atoms (Å ²)	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.26% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 5QM

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	0.58	0/2604	0.74	1/3519 (0.0%)
1	B	0.57	0/2671	0.73	0/3609
1	C	0.55	0/2594	0.71	0/3505
1	D	0.58	0/2653	0.73	0/3587
All	All	0.57	0/10522	0.73	1/14220 (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	C	0	1
1	D	0	1
All	All	0	2

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	267	TYR	CB-CA-C	-5.52	99.37	110.40

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	18	GLY	Peptide
1	D	283	ASP	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	2545	0	2568	18	0
1	B	2614	0	2639	23	0
1	C	2536	0	2555	20	0
1	D	2593	0	2623	27	0
2	A	14	9	9	0	0
2	B	14	9	9	0	0
2	C	14	9	9	0	0
2	D	14	9	9	0	0
3	A	143	0	0	2	0
3	B	162	0	0	0	0
3	C	130	0	0	0	0
3	D	129	0	0	1	0
All	All	10908	36	10421	87	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 4.

The worst 5 of 87 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:A:310:TYR:H	1:A:314:THR:HG21	1.47	0.79
1:C:109:THR:HG22	1:C:202:LEU:HB3	1.67	0.75
1:C:138:LEU:HB3	1:C:146:LEU:HD22	1.73	0.69
1:A:104:LEU:HB2	1:A:109:THR:HG22	1.74	0.69
1:A:138:LEU:HB3	1:A:146:LEU:HD22	1.73	0.69

There are no symmetry-related clashes.

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	306/356 (86%)	290 (95%)	16 (5%)	0	100	100
1	B	318/356 (89%)	301 (95%)	17 (5%)	0	100	100
1	C	302/356 (85%)	286 (95%)	16 (5%)	0	100	100
1	D	316/356 (89%)	292 (92%)	22 (7%)	2 (1%)	25	31
All	All	1242/1424 (87%)	1169 (94%)	71 (6%)	2 (0%)	47	58

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	48	GLY
1	D	3	ASP

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	277/314 (88%)	248 (90%)	29 (10%)	7	8
1	B	285/314 (91%)	255 (90%)	30 (10%)	7	8
1	C	277/314 (88%)	260 (94%)	17 (6%)	18	25
1	D	282/314 (90%)	252 (89%)	30 (11%)	6	7
All	All	1121/1256 (89%)	1015 (90%)	106 (10%)	8	10

5 of 106 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	320	LEU
1	C	292	VAL
1	D	296	ASN
1	C	15	GLU
1	C	154	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	144	HIS
1	C	274	GLN
1	D	260	HIS
1	C	265	GLN
1	C	277	ASN

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](#)

4 ligands are modelled in this entry.

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	$\# Z > 2$
2	5QM	A	1334	-	15,15,15	0.65	0	18,19,19	0.47	0
2	5QM	B	1334	-	15,15,15	0.66	0	18,19,19	0.48	0
2	5QM	D	1334	-	15,15,15	0.68	0	18,19,19	0.50	0
2	5QM	C	1334	-	15,15,15	0.56	0	18,19,19	0.49	0

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
2	5QM	A	1334	-	-	0/2/4/4	0/2/2/2
2	5QM	B	1334	-	-	0/2/4/4	0/2/2/2
2	5QM	D	1334	-	-	0/2/4/4	0/2/2/2
2	5QM	C	1334	-	-	0/2/4/4	0/2/2/2

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](#)

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	313/356 (87%)	-0.07	2 (0%) 89 92	35, 56, 86, 112	0
1	B	324/356 (91%)	0.03	7 (2%) 62 69	37, 57, 91, 121	0
1	C	311/356 (87%)	0.05	10 (3%) 47 54	38, 59, 89, 134	0
1	D	321/356 (90%)	0.17	19 (5%) 22 28	37, 60, 106, 138	0
All	All	1269/1424 (89%)	0.04	38 (2%) 50 57	35, 58, 93, 138	0

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	186	LEU	7.1
1	D	47	LEU	6.7
1	D	331	ARG	3.9
1	A	214	ALA	3.9
1	C	182	GLY	3.7

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 monosaccharides in this entry.

6.4 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. 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	B-factors(\AA^2)	Q<0.9
2	5QM	C	1334	14/14	0.93	0.15	38,47,53,53	0
2	5QM	D	1334	14/14	0.94	0.10	36,46,55,57	0
2	5QM	B	1334	14/14	0.95	0.13	38,45,50,50	23
2	5QM	A	1334	14/14	0.96	0.09	35,43,49,55	0

6.5 Other polymers [i](#)

There are no such residues in this entry.