



wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 01:58 AM GMT

PDB ID : 2AWN
Title : Crystal structure of the ADP-Mg-bound E. Coli MALK (Crystallized with ATP-Mg)
Authors : Lu, G.; Westbrook, J.M.; Davidson, A.L.; Chen, J.
Deposited on : 2005-09-01
Resolution : 2.30 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>

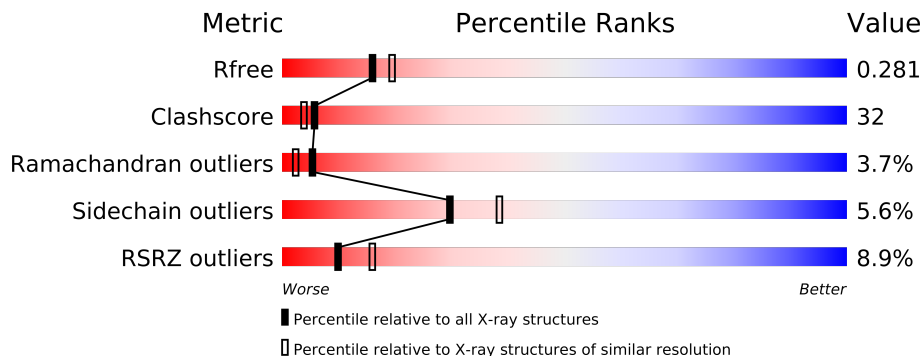
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
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)	:	stable22683

1 Overall quality at a glance

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}	66092	2929 (2.30-2.30)
Clashscore	79885	3679 (2.30-2.30)
Ramachandran outliers	78287	3642 (2.30-2.30)
Sidechain outliers	78261	3641 (2.30-2.30)
RSRZ outliers	66119	2930 (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 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	381	
1	B	381	
1	C	381	
1	D	381	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	MG	B	502	-	X

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 10669 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 Maltose/maltodextrin import ATP-binding protein malK.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	330	Total	C	N	O	S	0	0	0
			2534	1605	445	472	12			
1	B	374	Total	C	N	O	S	0	0	0
			2887	1825	515	534	13			
1	C	344	Total	C	N	O	S	0	0	0
			2596	1642	466	476	12			
1	D	301	Total	C	N	O	S	0	0	0
			2334	1475	412	435	12			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	372	ALA	-	CLONING ARTIFACT	UNP P68187
A	373	SER	-	CLONING ARTIFACT	UNP P68187
A	374	ALA	-	CLONING ARTIFACT	UNP P68187
A	375	SER	-	CLONING ARTIFACT	UNP P68187
A	376	HIS	-	CLONING ARTIFACT	UNP P68187
A	377	HIS	-	CLONING ARTIFACT	UNP P68187
A	378	HIS	-	CLONING ARTIFACT	UNP P68187
A	379	HIS	-	CLONING ARTIFACT	UNP P68187
A	380	HIS	-	CLONING ARTIFACT	UNP P68187
A	381	HIS	-	CLONING ARTIFACT	UNP P68187
B	372	ALA	-	CLONING ARTIFACT	UNP P68187
B	373	SER	-	CLONING ARTIFACT	UNP P68187
B	374	ALA	-	CLONING ARTIFACT	UNP P68187
B	375	SER	-	CLONING ARTIFACT	UNP P68187
B	376	HIS	-	CLONING ARTIFACT	UNP P68187
B	377	HIS	-	CLONING ARTIFACT	UNP P68187
B	378	HIS	-	CLONING ARTIFACT	UNP P68187
B	379	HIS	-	CLONING ARTIFACT	UNP P68187
B	380	HIS	-	CLONING ARTIFACT	UNP P68187
B	381	HIS	-	CLONING ARTIFACT	UNP P68187
C	372	ALA	-	CLONING ARTIFACT	UNP P68187

Continued on next page...

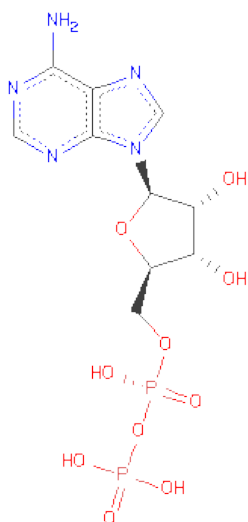
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	373	SER	-	CLONING ARTIFACT	UNP P68187
C	374	ALA	-	CLONING ARTIFACT	UNP P68187
C	375	SER	-	CLONING ARTIFACT	UNP P68187
C	376	HIS	-	CLONING ARTIFACT	UNP P68187
C	377	HIS	-	CLONING ARTIFACT	UNP P68187
C	378	HIS	-	CLONING ARTIFACT	UNP P68187
C	379	HIS	-	CLONING ARTIFACT	UNP P68187
C	380	HIS	-	CLONING ARTIFACT	UNP P68187
C	381	HIS	-	CLONING ARTIFACT	UNP P68187
D	372	ALA	-	CLONING ARTIFACT	UNP P68187
D	373	SER	-	CLONING ARTIFACT	UNP P68187
D	374	ALA	-	CLONING ARTIFACT	UNP P68187
D	375	SER	-	CLONING ARTIFACT	UNP P68187
D	376	HIS	-	CLONING ARTIFACT	UNP P68187
D	377	HIS	-	CLONING ARTIFACT	UNP P68187
D	378	HIS	-	CLONING ARTIFACT	UNP P68187
D	379	HIS	-	CLONING ARTIFACT	UNP P68187
D	380	HIS	-	CLONING ARTIFACT	UNP P68187
D	381	HIS	-	CLONING ARTIFACT	UNP P68187

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 3 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula: C₁₀H₁₅N₅O₁₀P₂).



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

- Molecule 4 is water.

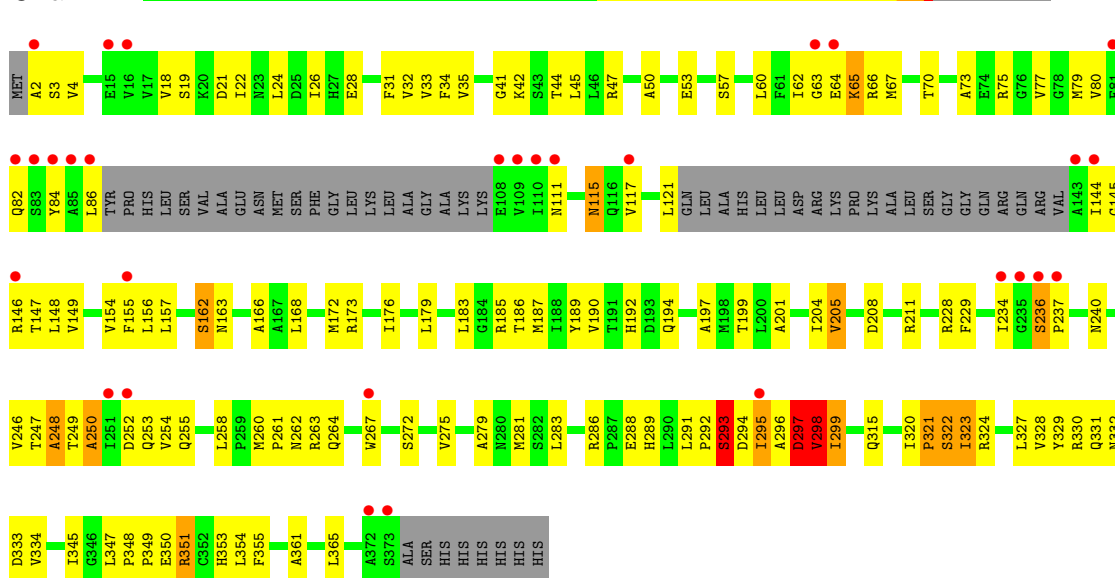
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	44	Total	O	0	0
			44	44		
4	B	81	Total	O	0	0
			81	81		
4	C	36	Total	O	0	0
			36	36		
4	D	45	Total	O	0	0
			45	45		

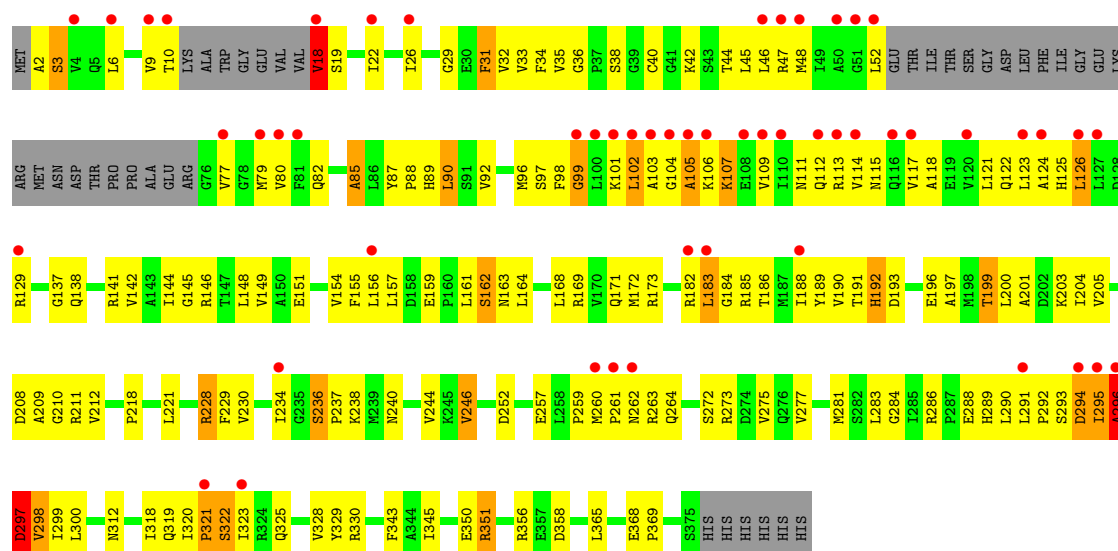
3 Residue-property plots

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: Maltose/maltodextrin import ATP-binding protein malK

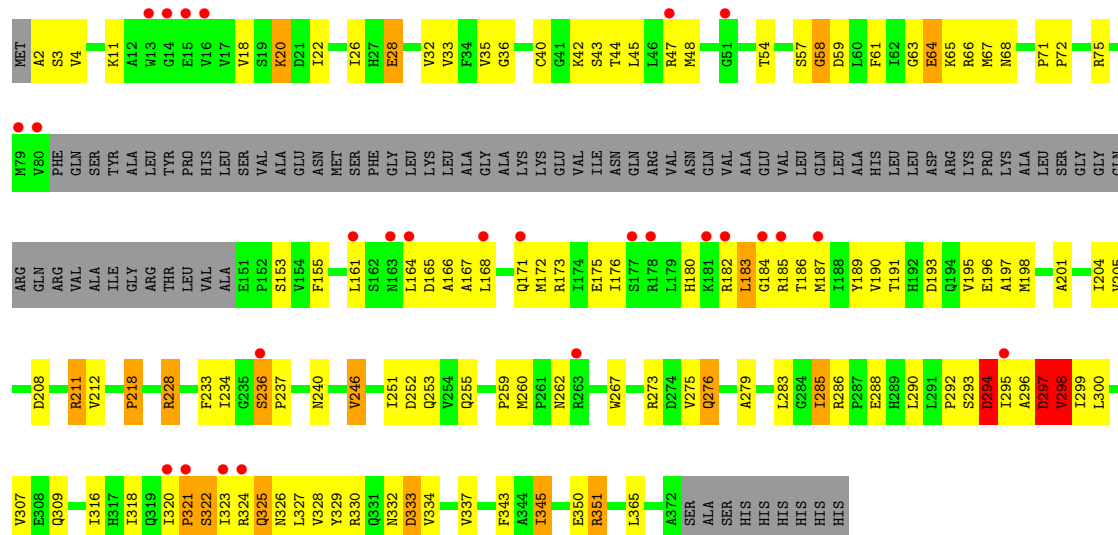
Chain A:





• Molecule 1: Maltose/maltodextrin import ATP-binding protein malK

Chain D:



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	70.31Å 102.67Å 130.78Å 90.00° 90.81° 90.00°	Depositor
Resolution (Å)	37.28 – 2.30 37.28 – 2.28	Depositor EDS
% Data completeness (in resolution range)	85.0 (37.28-2.30) 93.9 (37.28-2.28)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.61 (at 2.29Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.230 , 0.271 0.239 , 0.281	Depositor DCC
R_{free} test set	3843 reflections (5.18%)	DCC
Wilson B-factor (Å ²)	52.4	Xtriage
Anisotropy	0.390	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 50.9	EDS
Estimated twinning fraction	0.023 for h,-k,-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 83374 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	10669	wwPDB-VP
Average B, all atoms (Å ²)	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.24% 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: 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	0.90	3/2576 (0.1%)	0.91	9/3500 (0.3%)
1	B	0.54	2/2937 (0.1%)	0.82	11/3984 (0.3%)
1	C	0.49	2/2637 (0.1%)	0.79	7/3581 (0.2%)
1	D	0.76	3/2375 (0.1%)	0.82	7/3224 (0.2%)
All	All	0.69	10/10525 (0.1%)	0.84	34/14289 (0.2%)

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

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	299	ILE	N-CA	32.04	2.10	1.46
1	D	298	VAL	N-CA	30.24	2.06	1.46
1	A	298	VAL	N-CA	24.84	1.96	1.46
1	C	297	ASP	N-CA	17.16	1.80	1.46
1	B	327	LEU	N-CA	15.96	1.78	1.46

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	298	VAL	N-CA-CB	19.45	154.29	111.50
1	A	298	VAL	N-CA-CB	18.42	152.02	111.50
1	A	298	VAL	C-N-CA	-16.37	80.79	121.70
1	C	297	ASP	N-CA-CB	15.98	139.37	110.60
1	C	297	ASP	C-N-CA	14.99	159.17	121.70

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	293	SER	Peptide
1	B	326	ASN	Peptide
1	C	18	VAL	Mainchain
1	C	296	ALA	Peptide
1	D	297	ASP	Peptide

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	2534	0	2562	150	0
1	B	2887	0	2946	161	0
1	C	2596	0	2624	226	1
1	D	2334	0	2367	132	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	27	0	12	1	0
3	B	27	0	12	3	0
3	C	27	0	12	3	0
3	D	27	0	12	0	0
4	A	44	0	0	1	0
4	B	81	0	0	6	1
4	C	36	0	0	2	0
4	D	45	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	10669	0	10547	663	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 32.

The worst 5 of 663 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:18:VAL:CG2	1:C:44:THR:HG21	1.31	1.58
1:C:296:ALA:CB	1:C:299:ILE:CG1	1.83	1.54
1:B:327:LEU:N	1:B:327:LEU:CA	1.78	1.45
1:C:297:ASP:N	1:C:297:ASP:CA	1.80	1.44
1:B:319:GLN:CA	1:B:326:ASN:HD21	1.29	1.41

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)
1:C:102:LEU:CD2	4:B:765:HOH:O[2.556]	1.97	0.23

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	324/381 (85%)	294 (91%)	20 (6%)	10 (3%)	7	3
1	B	372/381 (98%)	339 (91%)	24 (6%)	9 (2%)	9	6
1	C	338/381 (89%)	290 (86%)	29 (9%)	19 (6%)	3	1
1	D	297/381 (78%)	263 (89%)	23 (8%)	11 (4%)	5	2
All	All	1331/1524 (87%)	1186 (89%)	96 (7%)	49 (4%)	5	2

5 of 49 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	236	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	293	SER
1	A	298	VAL
1	B	236	SER
1	B	323	ILE

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	277/323 (86%)	261 (94%)	16 (6%)	28	36
1	B	315/323 (98%)	299 (95%)	16 (5%)	33	43
1	C	277/323 (86%)	265 (96%)	12 (4%)	40	52
1	D	258/323 (80%)	239 (93%)	19 (7%)	20	24
All	All	1127/1292 (87%)	1064 (94%)	63 (6%)	30	38

5 of 63 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	332	ASN
1	C	126	LEU
1	D	297	ASP
1	B	345	ILE
1	B	351	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 45 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	255	GLN
1	B	353	HIS
1	D	276	GLN
1	B	326	ASN
1	C	112	GLN

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 8 ligands modelled in this entry, 4 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	$\# Z > 2$
3	ADP	A	401	2	29,29,29	1.16	1 (3%)	45,45,45	2.21	10 (22%)
3	ADP	B	402	2	29,29,29	1.26	3 (10%)	45,45,45	2.18	10 (22%)
3	ADP	C	403	2	29,29,29	0.96	1 (3%)	45,45,45	2.15	9 (20%)
3	ADP	D	404	2	29,29,29	1.04	1 (3%)	45,45,45	2.12	10 (22%)

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
3	ADP	A	401	2	-	0/16/32/32	0/1/3/3
3	ADP	B	402	2	-	0/16/32/32	0/1/3/3
3	ADP	C	403	2	-	0/16/32/32	0/1/3/3
3	ADP	D	404	2	-	0/16/32/32	0/1/3/3

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	402	ADP	C2-N1	3.13	1.40	1.33
3	D	404	ADP	C2-N1	3.12	1.40	1.33
3	A	401	ADP	C2-N1	3.11	1.40	1.33
3	C	403	ADP	C2-N1	3.00	1.39	1.33
3	B	402	ADP	PA-O3A	-2.72	1.55	1.59

The worst 5 of 39 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	401	ADP	O5'-PA-O1A	-8.79	74.93	109.37
3	C	403	ADP	O5'-PA-O1A	-8.48	76.14	109.37
3	D	404	ADP	O5'-PA-O1A	-8.23	77.12	109.37
3	B	402	ADP	O5'-PA-O1A	-8.13	77.51	109.37
3	C	403	ADP	N3-C2-N1	-6.61	123.18	128.71

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	330/381 (86%)	0.57	30 (9%) 9 15	39, 63, 107, 146	0
1	B	374/381 (98%)	0.27	11 (2%) 49 59	32, 52, 87, 119	0
1	C	344/381 (90%)	0.84	53 (15%) 3 5	38, 76, 130, 148	0
1	D	301/381 (79%)	0.53	27 (8%) 10 16	34, 64, 102, 115	0
All	All	1349/1524 (88%)	0.55	121 (8%) 10 16	32, 62, 113, 148	0

The worst 5 of 121 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	86	LEU	8.6
1	C	102	LEU	8.2
1	A	84	TYR	7.2
1	A	85	ALA	7.1
1	C	103	ALA	6.8

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
2	MG	B	502	1/1	0.19	39.19	48,48,48,48	0
3	ADP	B	402	27/27	0.17	1.67	40,56,73,81	0
3	ADP	C	403	27/27	0.27	1.03	85,110,116,117	27
3	ADP	A	401	27/27	0.15	0.64	38,74,84,94	0
2	MG	C	503	1/1	0.15	-0.08	109,109,109,109	0
3	ADP	D	404	27/27	0.13	-0.27	54,92,100,106	0
2	MG	D	504	1/1	0.08	-0.84	76,76,76,76	0
2	MG	A	501	1/1	0.05	-5.10	55,55,55,55	0

6.5 Other polymers ⓘ

There are no such residues in this entry.