



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 07:30 PM GMT

PDB ID : 1FW6  
Title : CRYSTAL STRUCTURE OF A TAQ MUTS-DNA-ADP TERNARY COMPLEX  
Authors : Junop, M.S.; Obmolova, G.; Rausch, K.; Hsieh, P.; Yang, W.  
Deposited on : 2000-09-21  
Resolution : 2.70 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](mailto: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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

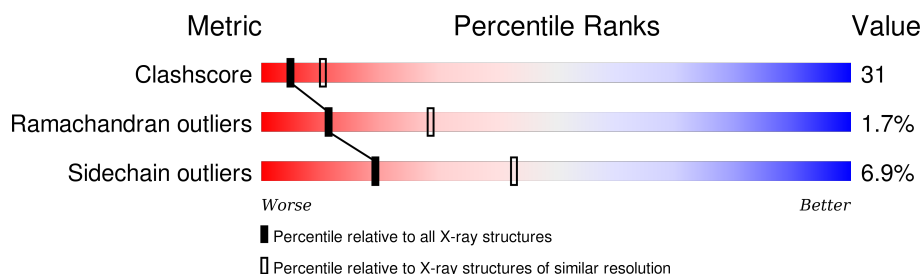
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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))
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)

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  $\geq 3$ , 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	C	23	
2	D	22	
3	A	768	
3	B	768	

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 13165 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 DNA chain called 5'-D(\*GP\*CP\*GP\*AP\*CP\*GP\*CP\*TP\*AP\*GP\*CP\*GP\*TP\*GP\*CP\*GP\*GP\*CP\*TP\*CP\*GP\*TP\*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	23	Total	C	N	O	P	0	0	0
			469	222	87	138	22			

- Molecule 2 is a DNA chain called 5'-D(\*GP\*GP\*AP\*CP\*GP\*AP\*GP\*CP\*CP\*GP\*CP\*C\*P\*GP\*CP\*TP\*AP\*GP\*CP\*GP\*TP\*CP\*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	22	Total	C	N	O	P	0	0	0
			450	212	88	129	21			

- Molecule 3 is a protein called DNA MISMATCH REPAIR PROTEIN MUTS.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	A	759	Total	C	N	O	S	Se	0	0	0
			6006	3825	1068	1099	1	13			
3	B	749	Total	C	N	O	S	Se	0	0	0
			5936	3784	1055	1083	1	13			

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q56215
A	4	MSE	MET	MODIFIED RESIDUE	UNP Q56215
A	70	MSE	MET	MODIFIED RESIDUE	UNP Q56215
A	88	MSE	MET	MODIFIED RESIDUE	UNP Q56215
A	201	MSE	MET	MODIFIED RESIDUE	UNP Q56215
A	250	MSE	MET	MODIFIED RESIDUE	UNP Q56215
A	481	MSE	MET	MODIFIED RESIDUE	UNP Q56215
A	574	MSE	MET	MODIFIED RESIDUE	UNP Q56215
A	586	MSE	MET	MODIFIED RESIDUE	UNP Q56215
A	640	MSE	MET	MODIFIED RESIDUE	UNP Q56215

*Continued on next page...*

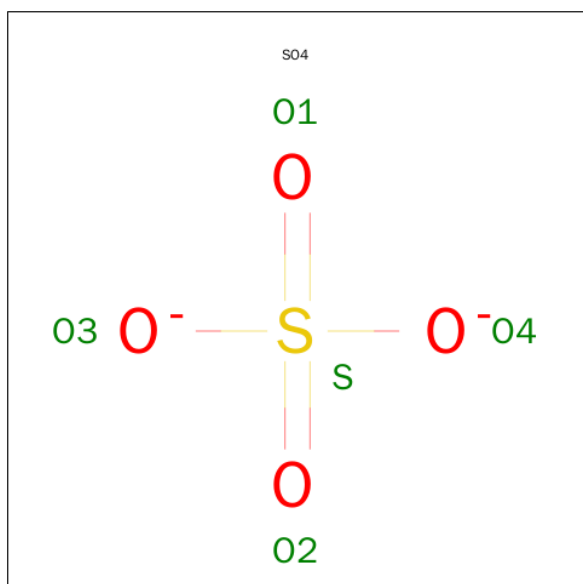
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	643	MSE	MET	MODIFIED RESIDUE	UNP Q56215
A	744	MSE	MET	MODIFIED RESIDUE	UNP Q56215
A	762	MSE	MET	MODIFIED RESIDUE	UNP Q56215
B	1001	MSE	MET	MODIFIED RESIDUE	UNP Q56215
B	1004	MSE	MET	MODIFIED RESIDUE	UNP Q56215
B	1070	MSE	MET	MODIFIED RESIDUE	UNP Q56215
B	1088	MSE	MET	MODIFIED RESIDUE	UNP Q56215
B	1201	MSE	MET	MODIFIED RESIDUE	UNP Q56215
B	1250	MSE	MET	MODIFIED RESIDUE	UNP Q56215
B	1481	MSE	MET	MODIFIED RESIDUE	UNP Q56215
B	1574	MSE	MET	MODIFIED RESIDUE	UNP Q56215
B	1586	MSE	MET	MODIFIED RESIDUE	UNP Q56215
B	1640	MSE	MET	MODIFIED RESIDUE	UNP Q56215
B	1643	MSE	MET	MODIFIED RESIDUE	UNP Q56215
B	1744	MSE	MET	MODIFIED RESIDUE	UNP Q56215
B	1762	MSE	MET	MODIFIED RESIDUE	UNP Q56215

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

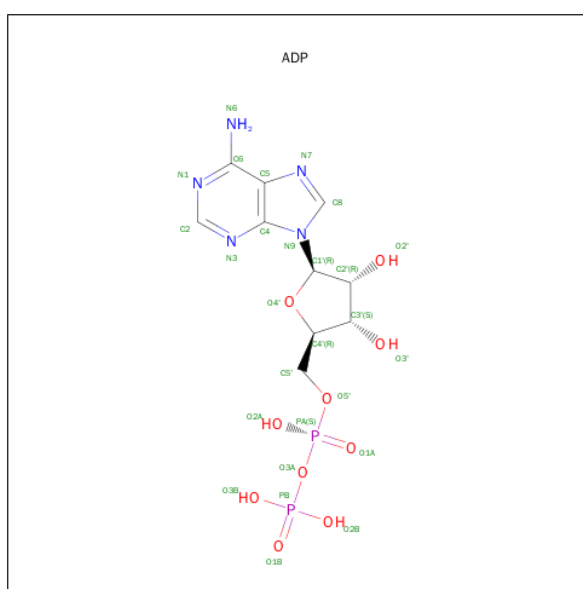
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Mg 1 1	0	0
4	A	1	Total Mg 1 1	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total O S 5 4 1	0	0
5	B	1	Total O S 5 4 1	0	0
5	B	1	Total O S 5 4 1	0	0
5	A	1	Total O S 5 4 1	0	0

- Molecule 6 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula:  $C_{10}H_{15}N_5O_{10}P_2$ ).



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

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	105	Total O 105 105	0	0
7	B	115	Total O 115 115	0	0
7	C	4	Total O 4 4	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	D	4	Total	O	0	0
			4	4		

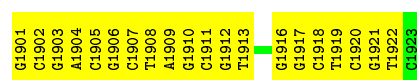
### 3 Residue-property plots

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.

Note EDS was not executed.

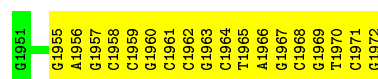
- Molecule 1: 5'-D(\*GP\*CP\*GP\*AP\*CP\*GP\*CP\*TP\*AP\*GP\*CP\*GP\*TP\*GP\*CP\*GP\*GP\*CP\*TP\*CP\*GP\*TP\*C)-3'

Chain C: 



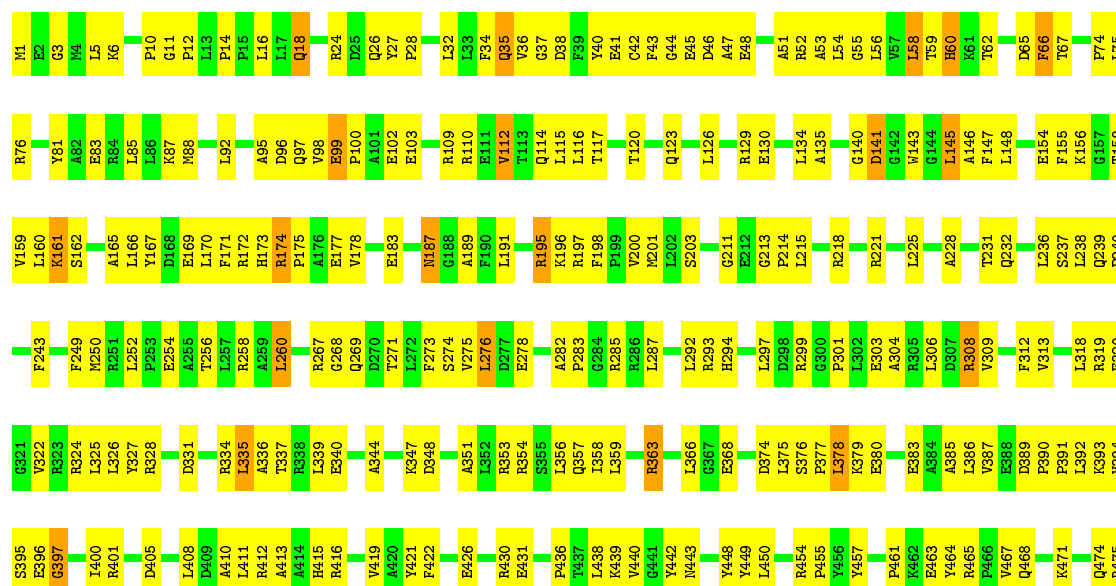
- Molecule 2: 5'-D(\*GP\*GP\*AP\*CP\*GP\*AP\*GP\*CP\*CP\*GP\*CP\*CP\*GP\*CP\*TP\*AP\*GP\*CP\*GP\*TP\*CP\*G)-3'

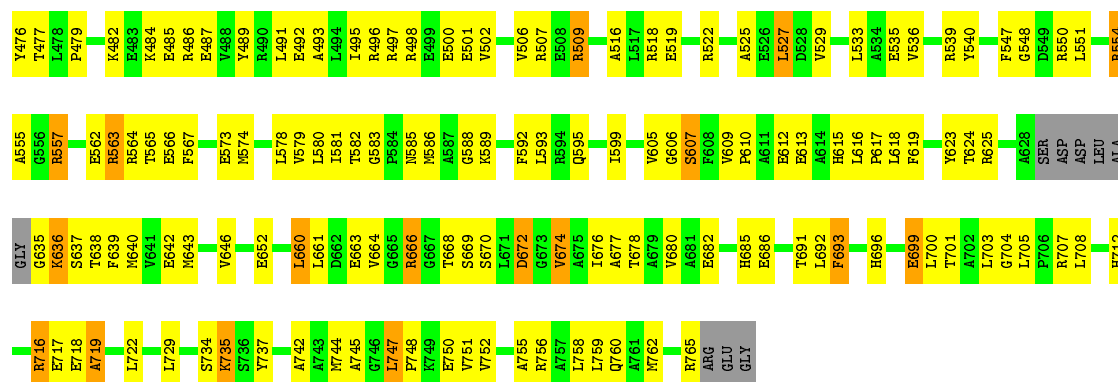
Chain D: 



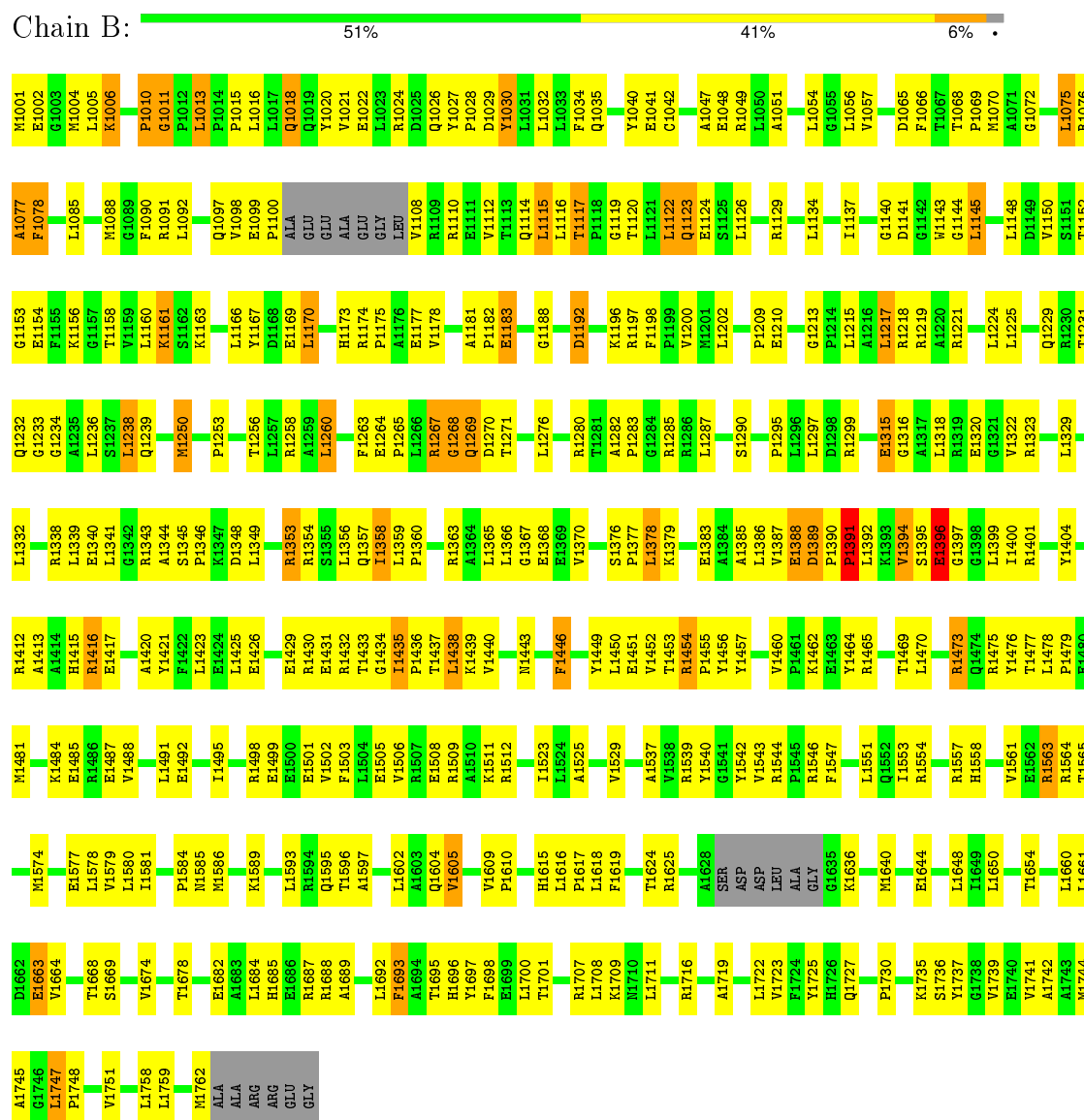
- Molecule 3: DNA MISMATCH REPAIR PROTEIN MUTS

Chain A: 





### • Molecule 3: DNA MISMATCH REPAIR PROTEIN MUTS





## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	103.72Å 113.50Å 160.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	23.17 – 2.70	Depositor
% Data completeness (in resolution range)	90.8 (23.17-2.70)	Depositor
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS 0.9	Depositor
R, $R_{free}$	0.218 , 0.280	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	13165	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

## 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: MG, SO4, 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	C	0.32	0/525	0.66	0/809
2	D	0.29	0/505	0.66	0/778
3	A	0.39	0/6111	0.66	1/8249 (0.0%)
3	B	0.39	0/6040	0.68	2/8152 (0.0%)
All	All	0.38	0/13181	0.67	3/17988 (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	B	1268	GLY	N-CA-C	-5.30	99.86	113.10
3	A	134	LEU	N-CA-C	-5.13	97.14	111.00
3	B	1329	LEU	CA-CB-CG	5.09	127.02	115.30

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	C	469	0	259	35	0
2	D	450	0	246	38	0
3	A	6006	0	6093	362	0
3	B	5936	0	6023	381	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	10	0	0	0	0
5	B	10	0	0	0	0
6	A	27	0	12	2	0
6	B	27	0	12	3	0
7	A	105	0	0	13	0
7	B	115	0	0	20	0
7	C	4	0	0	6	0
7	D	4	0	0	0	0
All	All	13165	0	12645	798	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 31.

The worst 5 of 798 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)
3:B:1663:GLU:HG3	7:B:150:HOH:O	1.40	1.18
3:B:1597:ALA:HB2	3:B:1660:LEU:HD11	1.34	1.10
3:A:267:ARG:HB2	3:A:269:GLN:HE21	1.12	1.07
1:C:1916:DG:H2''	1:C:1917:DG:H5'	1.35	1.07
3:B:1723:VAL:HG13	7:B:149:HOH:O	1.54	1.06

There are no symmetry-related clashes.

## 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
3	A	755/768 (98%)	688 (91%)	58 (8%)	9 (1%)	16 39
3	B	743/768 (97%)	661 (89%)	65 (9%)	17 (2%)	8 20

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1498/1536 (98%)	1349 (90%)	123 (8%)	26 (2%)	11	29

5 of 26 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	B	1010	PRO
3	B	1011	GLY
3	B	1233	GLY
3	B	1269	GLN
3	B	1391	PRO

### 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	
3	A	613/606 (101%)	572 (93%)	41 (7%)	20	44
3	B	608/606 (100%)	565 (93%)	43 (7%)	18	41
All	All	1221/1212 (101%)	1137 (93%)	84 (7%)	19	43

5 of 84 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	699	GLU
3	B	1075	LEU
3	B	1563	ARG
3	A	735	LYS
3	B	1006	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
3	A	585	ASN
3	B	1097	GLN
3	B	1604	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	A	685	HIS
3	B	1114	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules 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, 2 are monoatomic - leaving 6 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	$\# Z  > 2$
5	SO4	A	1853	-	4,4,4	0.30	0	6,6,6	0.09	0
5	SO4	A	852	-	4,4,4	0.18	0	6,6,6	0.27	0
6	ADP	A	999	4	22,29,29	1.13	3 (13%)	27,45,45	2.11	8 (29%)
5	SO4	B	1852	-	4,4,4	0.26	0	6,6,6	0.30	0
6	ADP	B	1999	4	22,29,29	1.07	2 (9%)	27,45,45	1.95	8 (29%)
5	SO4	B	853	-	4,4,4	0.23	0	6,6,6	0.15	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
5	SO4	A	1853	-	-	0/0/0/0	0/0/0/0
5	SO4	A	852	-	-	0/0/0/0	0/0/0/0
6	ADP	A	999	4	-	0/12/32/32	0/3/3/3
5	SO4	B	1852	-	-	0/0/0/0	0/0/0/0
6	ADP	B	1999	4	-	0/12/32/32	0/3/3/3
5	SO4	B	853	-	-	0/0/0/0	0/0/0/0

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	999	ADP	PB-O2B	-2.24	1.46	1.54
6	B	1999	ADP	PB-O2B	-2.11	1.47	1.54
6	A	999	ADP	C8-N7	-2.07	1.30	1.34
6	B	1999	ADP	C8-N7	-2.02	1.30	1.34
6	A	999	ADP	C2-N3	2.12	1.35	1.32

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	B	1999	ADP	N3-C2-N1	-5.26	124.87	128.89
6	A	999	ADP	C5'-C4'-C3'	-4.57	97.07	115.21
6	A	999	ADP	N3-C2-N1	-4.42	125.51	128.89
6	B	1999	ADP	O3A-PA-O5'	-3.92	92.53	102.94
6	B	1999	ADP	O4'-C4'-C5'	-2.51	100.34	109.32

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	999	ADP	2	0
6	B	1999	ADP	3	0

## 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section will therefore be empty.