

# wwPDB X-ray Structure Validation Summary Report (i)

#### Feb 18, 2024 – 11:59 AM EST

PDB ID : 4E7Z

Title: Myosin VI (MD) pre-powerstroke state, P21 crystal form

Authors: Isabet, T.; Sweeney, H.L.; Houdusse, A.

Deposited on : 2012-03-19

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 (i) 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 (1)) were used in the production of this report:

MolProbity: 4.02b-467

Mogul : 1.8.5 (274361), CSD as541be (2020)

Xtriage (Phenix) : 1.13

EDS: 2.36

buster-report : 1.1.7 (2018)

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)

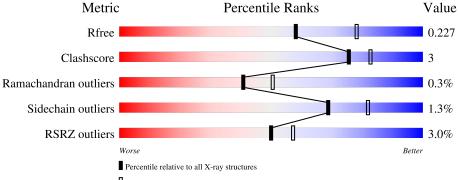
Validation Pipeline (wwPDB-VP) : 2.36

## 1 Overall quality at a glance (i)

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.



Percentile relative to X-ray structures of similar resolution

Metric	Whole archive	Similar resolution
Metric	$(\# \mathrm{Entries})$	$(\#  ext{Entries},  ext{ resolution range}( ext{Å}))$
$R_{free}$	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 <=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	798	87%	7%	5%
1	В	798	85%	8%	7%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:



Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	GOL	A	804	_	-	X	-



# 2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 12444 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 Myosin-VI.

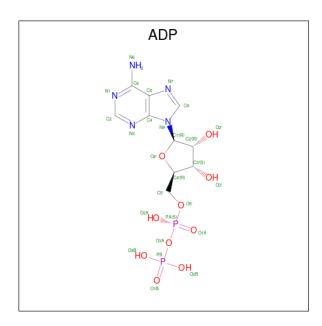
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	A	759	Total 6037	C 3845	N 1033	O 1128	S 31	0	3	0
1	В	745	Total 5849	C 3728	N 1001	O 1093	S 27	0	1	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-8	MET	-	expression tag	UNP F1RQI7
A	-7	ASP	-	expression tag	UNP F1RQI7
A	-6	TYR	-	expression tag	UNP F1RQI7
A	-5	LYS	-	expression tag	UNP F1RQI7
A	-4	ASP	-	expression tag	UNP F1RQI7
A	-3	ASP	-	expression tag	UNP F1RQI7
A	-2	ASP	-	expression tag	UNP F1RQI7
A	-1	ASP	-	expression tag	UNP F1RQI7
A	0	LYS	-	expression tag	UNP F1RQI7
A	1	GLY	-	expression tag	UNP F1RQI7
В	-8	MET	-	expression tag	UNP F1RQI7
В	-7	ASP	-	expression tag	UNP F1RQI7
В	-6	TYR	-	expression tag	UNP F1RQI7
В	-5	LYS	-	expression tag	UNP F1RQI7
В	-4	ASP	-	expression tag	UNP F1RQI7
В	-3	ASP	-	expression tag	UNP F1RQI7
В	-2	ASP	-	expression tag	UNP F1RQI7
В	-1	ASP	-	expression tag	UNP F1RQI7
В	0	LYS	-	expression tag	UNP F1RQI7
В	1	GLY	_	expression tag	UNP F1RQI7

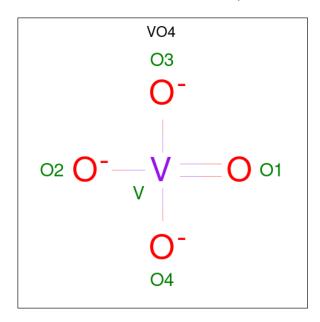
• Molecule 2 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula:  $C_{10}H_{15}N_5O_{10}P_2$ ).





Mol	Chain	Residues		Ato	oms			ZeroOcc	AltConf
9	Λ	1	Total	С	N	О	Р	0	0
2	A	1	27	10	5	10	2	U	0
9	D	1	Total	С	N	О	Р	0	0
2	2   B	1	27	10	5	10	2	U	0

 $\bullet$  Molecule 3 is VANADATE ION (three-letter code: VO4) (formula:  $\mathrm{O_4V}).$ 



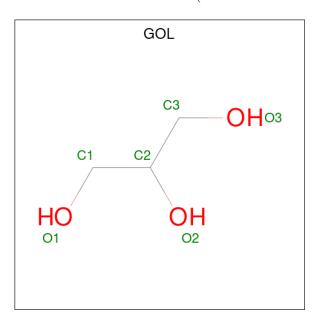
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O V 5 4 1	0	0
3	В	1	Total O V 5 4 1	0	0



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

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

• Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	В	1	Total C O 6 3 3	0	0
5	В	1	Total C O 6 3 3	0	0

• Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	255	Total O 255 255	0	0

Continued on next page...



Continued from previous page...

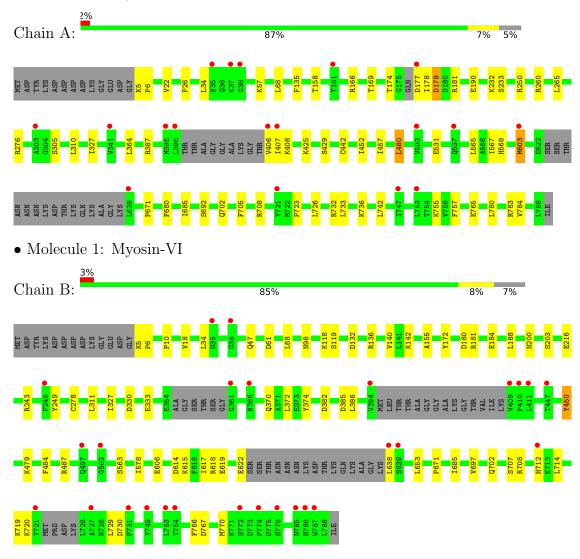
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	В	201	Total O 201 201	0	0



# 3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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: Myosin-VI





# 4 Data and refinement statistics (i)

Property	Value	Source	
Space group	P 1 21 1	Depositor	
Cell constants	98.15Å 93.32Å 101.88Å	Domositon	
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $90.60^{\circ}$ $90.00^{\circ}$	Depositor	
Resolution (Å)	29.65 - 2.30	Depositor	
rtesolution (A)	29.65 - 2.30	EDS	
% Data completeness	98.4 (29.65-2.30)	Depositor	
(in resolution range)	98.4 (29.65-2.30)	EDS	
$R_{merge}$	(Not available)	Depositor	
$R_{sym}$	(Not available)	Depositor	
$< I/\sigma(I) > 1$	3.43  (at  2.31Å)	Xtriage	
Refinement program	PHENIX (phenix.refine: 1.6.4_486)	Depositor	
$R, R_{free}$	0.176 , $0.231$	Depositor	
It, Itfree	0.173 , $0.227$	DCC	
$R_{free}$ test set	4029 reflections (5.01%)	wwPDB-VP	
Wilson B-factor (Å <sup>2</sup> )	40.1	Xtriage	
Anisotropy	0.341	Xtriage	
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.32, 41.6	EDS	
L-test for twinning <sup>2</sup>	$< L > = 0.49, < L^2> = 0.32$	Xtriage	
	0.001 for l,k,-h		
Estimated twinning fraction	0.022  for h,-k,-l	Xtriage	
	0.015  for  l,-k,h		
$F_o, F_c$ correlation	0.96	EDS	
Total number of atoms	12444	wwPDB-VP	
Average B, all atoms $(\mathring{A}^2)$	43.0	wwPDB-VP	

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 3.17% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;sup>2</sup>Theoretical values of <|L|>,  $<L^2>$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

## 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: ADP, VO4, GOL, MG

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).

Mal	Chain	Во	nd lengths	Bond angles		
Mol Chain		RMSZ	# Z  > 5	RMSZ	# Z >5	
1	A	0.49	0/6164	0.59	0/8323	
1	В	0.47	1/5967~(0.0%)	0.58	0/8066	
All	All	0.48	$1/12131 \ (0.0\%)$	0.59	0/16389	

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\operatorname{Observed}(\text{\AA})$	$Ideal(\AA)$
1	В	278	CYS	CB-SG	-5.04	1.73	1.81

There are no bond angle outliers.

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	A	6037	0	5873	35	0
1	В	5849	0	5619	39	0
2	A	27	0	12	0	0
2	В	27	0	12	0	0
3	A	5	0	0	0	0
3	В	5	0	0	0	0
4	A	1	0	0	0	0
4	В	1	0	0	0	0

Continued on next page...



I 'omtamalod	trom	mmonia	maaa
Continued	11 0116	DICUIUUS	Daue
	.,	10	1

Mol	Chain	Non-H	H(model)	$\mathbf{H}(\mathbf{added})$	Clashes	Symm-Clashes
5	A	24	0	32	6	0
5	В	12	0	16	3	0
6	A	255	0	0	5	0
6	В	201	0	0	6	0
All	All	12444	0	11564	74	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 3.

The worst 5 of 74 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$\begin{array}{c} {\rm Interatomic} \\ {\rm distance} ({\rm \AA}) \end{array}$	$\begin{array}{c} \text{Clash} \\ \text{overlap } (\text{\AA}) \end{array}$
1:A:692:SER:HB3	5:A:805:GOL:H12	1.57	0.87
1:B:618:ARG:O	6:B:1060:HOH:O	2.05	0.72
1:B:68:LEU:HD13	1:B:702:GLN:HG3	1.74	0.69
1:B:618:ARG:NH2	5:B:805:GOL:O2	2.27	0.67
1:B:142:LYS:NZ	1:B:450:TYR:OH	2.28	0.64

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	Perce	entiles	
1	A	754/798 (94%)	733 (97%)	19 (2%)	2 (0%)	41	50
1	В	736/798 (92%)	706 (96%)	27 (4%)	3 (0%)	34	42
All	All	1490/1596 (93%)	1439 (97%)	46 (3%)	5 (0%)	41	50

All (5) Ramachandran outliers are listed below:



Mol	Chain	Res	Type
1	A	408	LYS
1	A	26	PRO
1	В	720	LYS
1	В	712	HIS
1	В	140	VAL

#### 5.3.2 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	651/705~(92%)	639 (98%)	12 (2%)	59 75		
1	В	617/705 (88%)	612 (99%)	5 (1%)	81 91		
All	All	1268/1410 (90%)	1251 (99%)	17 (1%)	69 82		

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	В	638	LEU
1	В	708	ARG
1	A	603	MET
1	A	708	ARG
1	A	733	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

$\mathbf{Mol}$	Chain	$\operatorname{Res}$	Type
1	В	20	ASN
1	В	290	GLN

#### 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)

Of 12 ligands modelled in this entry, 2 are monoatomic - leaving 10 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	Trino	Chain	Res	Link	В	ond leng	$_{ m gths}$	В	ond ang	les
MIOI	Type	Chain	nes	Lilik	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	VO4	A	802	4,2	1,4,4	5.45	1 (100%)	-		
2	ADP	A	801	3,4	24,29,29	1.03	1 (4%)	29,45,45	1.28	3 (10%)
5	GOL	A	807	-	5,5,5	0.36	0	5,5,5	0.66	0
5	GOL	В	805	-	5,5,5	0.38	0	5,5,5	0.32	0
2	ADP	В	801	3,4	24,29,29	1.01	0	29,45,45	1.34	5 (17%)
5	GOL	A	805	-	5,5,5	0.41	0	5,5,5	0.77	0
5	GOL	В	804	-	5,5,5	0.32	0	5,5,5	0.35	0
3	VO4	В	802	4,2	1,4,4	5.42	1 (100%)	-		
5	GOL	A	804	-	5,5,5	0.51	0	5,5,5	0.61	0
5	GOL	A	806	-	5,5,5	0.30	0	5,5,5	0.46	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	ADP	A	801	3,4	-	3/12/32/32	0/3/3/3
5	GOL	A	807	-	-	4/4/4/4	-
5	GOL	В	805	-	-	4/4/4/4	-

Continued on next page...



Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	ADP	В	801	3,4	-	3/12/32/32	0/3/3/3
5	GOL	A	805	-	-	4/4/4/4	-
5	GOL	В	804	-	-	2/4/4/4	-
5	GOL	A	804	-	-	3/4/4/4	-
5	GOL	A	806	-	-	4/4/4/4	-

#### All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\operatorname{Observed}(\text{\AA})$	Ideal(A)
3	A	802	VO4	O1-V	5.45	1.94	1.63
3	В	802	VO4	O1-V	5.42	1.94	1.63
2	A	801	ADP	C5-C4	2.13	1.46	1.40

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

Mol	Chain	Res	Type	Atoms	Z	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
2	В	801	ADP	N3-C2-N1	-3.45	123.29	128.68
2	A	801	ADP	N3-C2-N1	-3.15	123.75	128.68
2	A	801	ADP	C4-C5-N7	-2.93	106.34	109.40
2	В	801	ADP	O2B-PB-O1B	2.56	120.70	110.68
2	В	801	ADP	O3A-PB-O1B	-2.47	97.49	111.19

There are no chirality outliers.

5 of 27 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	805	GOL	C1-C2-C3-O3
5	A	806	GOL	O1-C1-C2-C3
5	A	806	GOL	C1-C2-C3-O3
5	A	807	GOL	O1-C1-C2-C3
5	В	804	GOL	O1-C1-C2-C3

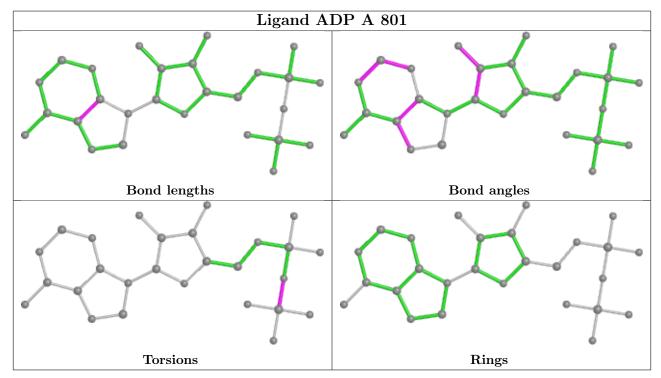
There are no ring outliers.

4 monomers are involved in 9 short contacts:

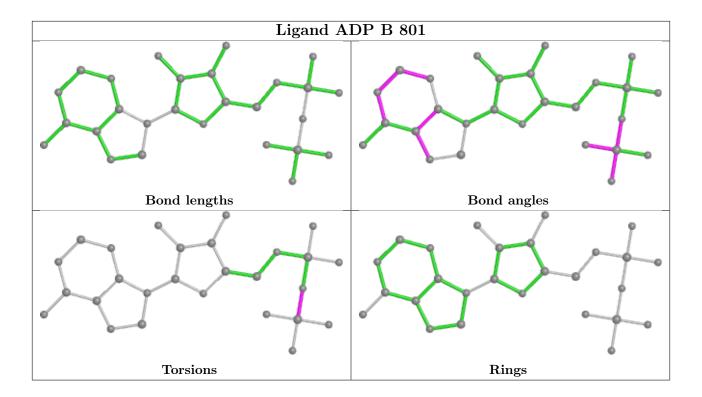
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	В	805	GOL	3	0
5	A	805	GOL	1	0
5	A	804	GOL	4	0
5	A	806	GOL	1	0



The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.







## 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,  $95^{th}$  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 $>$	$\#\mathrm{RSRZ}{>}2$	$OWAB(A^2)$	Q<0.9
1	A	759/798~(95%)	-0.26	18 (2%) 59 66	23, 39, 61, 90	0
1	В	745/798 (93%)	-0.07	27 (3%) 42 49	24, 43, 80, 114	0
All	All	1504/1596~(94%)	-0.17	45 (2%) 50 57	23, 41, 71, 114	0

The worst 5 of 45 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	638	LEU	5.5
1	В	638	LEU	5.3
1	A	35	ASN	5.2
1	В	786	HIS	4.5
1	A	406	VAL	4.1

## 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.

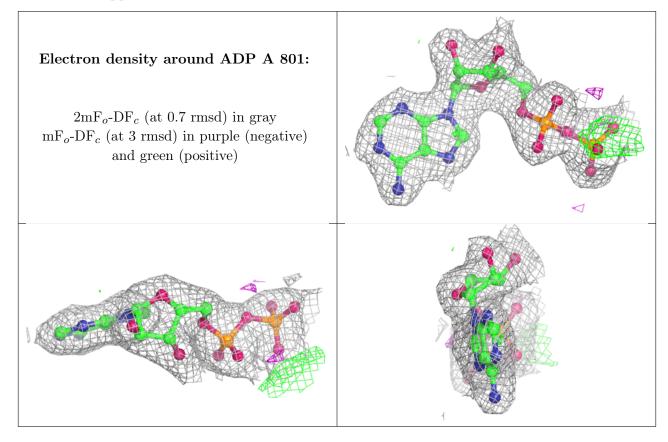
### 6.4 Ligands (i)

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,  $95^{th}$  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.

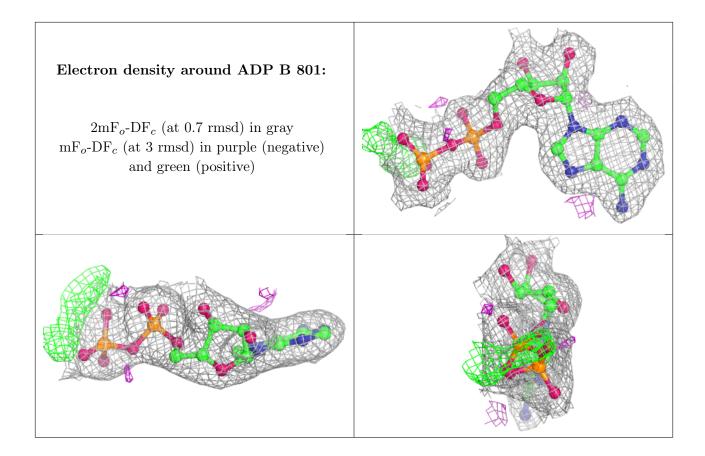


Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\operatorname{B-factors}(\mathring{\mathbf{A}}^2)$	Q<0.9
5	GOL	В	805	6/6	0.84	0.36	42,53,55,59	0
5	GOL	A	807	6/6	0.90	0.17	32,48,54,61	0
5	GOL	A	804	6/6	0.91	0.17	42,42,43,48	0
5	GOL	A	805	6/6	0.91	0.13	36,40,42,53	0
5	GOL	A	806	6/6	0.93	0.22	41,49,55,57	0
4	MG	В	803	1/1	0.95	0.33	32,32,32,32	0
5	GOL	В	804	6/6	0.97	0.20	49,51,52,55	0
4	MG	A	803	1/1	0.97	0.38	32,32,32,32	0
2	ADP	A	801	27/27	0.99	0.14	20,24,28,30	0
2	ADP	В	801	27/27	0.99	0.12	22,25,32,34	0
3	VO4	A	802	5/5	0.99	0.16	22,23,26,27	0
3	VO4	В	802	5/5	0.99	0.15	22,23,28,30	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.







# 6.5 Other polymers (i)

There are no such residues in this entry.

