



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 07:47 PM GMT

PDB ID : 4PWX  
Title : Crystal structure of an ATP-bound Get3-Get4-Get5 complex from *S.cerevisiae*  
Authors : Gristick, H.B.; Clemons Jr., W.M.  
Deposited on : 2014-03-21  
Resolution : 5.40 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) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

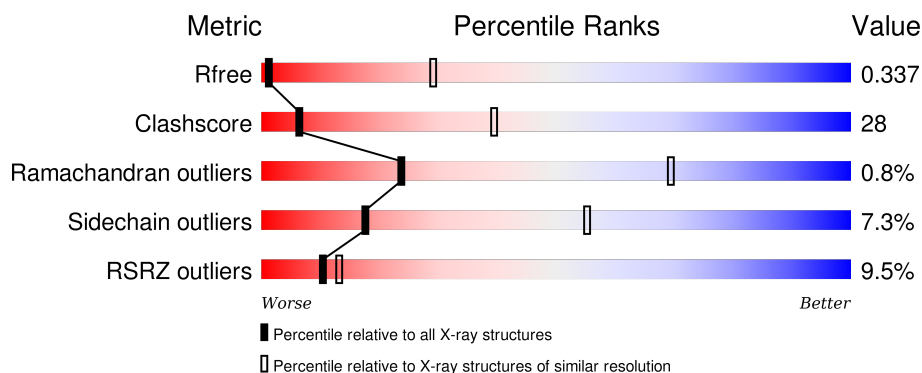
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 5.40 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))
$R_{free}$	91344	1013 (7.10-3.62)
Clashscore	102246	1020 (7.10-3.70)
Ramachandran outliers	100387	1013 (7.10-3.64)
Sidechain outliers	100360	1011 (7.10-3.62)
RSRZ outliers	91569	1012 (7.10-3.62)

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.

Mol	Chain	Length	Quality of chain
1	A	356	<div> <div>7%</div> <div>42%</div> <div>37%</div> <div>•</div> <div>17%</div> </div>
1	B	356	<div> <div>10%</div> <div>42%</div> <div>38%</div> <div>•</div> <div>16%</div> </div>
2	C	288	<div> <div>6%</div> <div>52%</div> <div>41%</div> <div>• •</div> </div>
2	E	288	<div> <div>6%</div> <div>51%</div> <div>43%</div> <div>• •</div> </div>
3	D	54	<div> <div>11%</div> <div>61%</div> <div>31%</div> <div>• •</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	54	

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
4	ATP	A	401	-	-	-	X
4	ATP	B	401	-	-	-	X
5	MG	A	402	-	-	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 10159 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 ATPase GET3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	294	Total	C	N	O	S	0	0	0
			2319	1475	382	446	16			
1	B	299	Total	C	N	O	S	0	0	0
			2358	1497	388	456	17			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	SER	-	EXPRESSION TAG	UNP Q12154
A	0	VAL	-	EXPRESSION TAG	UNP Q12154
A	1	ASP	-	EXPRESSION TAG	UNP Q12154
A	57	VAL	ASP	ENGINEERED MUTATION	UNP Q12154
B	-1	SER	-	EXPRESSION TAG	UNP Q12154
B	0	VAL	-	EXPRESSION TAG	UNP Q12154
B	1	ASP	-	EXPRESSION TAG	UNP Q12154
B	57	VAL	ASP	ENGINEERED MUTATION	UNP Q12154

- Molecule 2 is a protein called Golgi to ER traffic protein 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	277	Total	C	N	O	S	0	0	0
			2301	1498	362	436	5			
2	C	276	Total	C	N	O	S	0	0	0
			2296	1495	361	435	5			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	9	MET	-	EXPRESSION TAG	UNP Q12125
E	10	GLY	-	EXPRESSION TAG	UNP Q12125
E	258	ALA	LYS	ENGINEERED MUTATION	UNP Q12125
E	260	ALA	LYS	ENGINEERED MUTATION	UNP Q12125

*Continued on next page...*

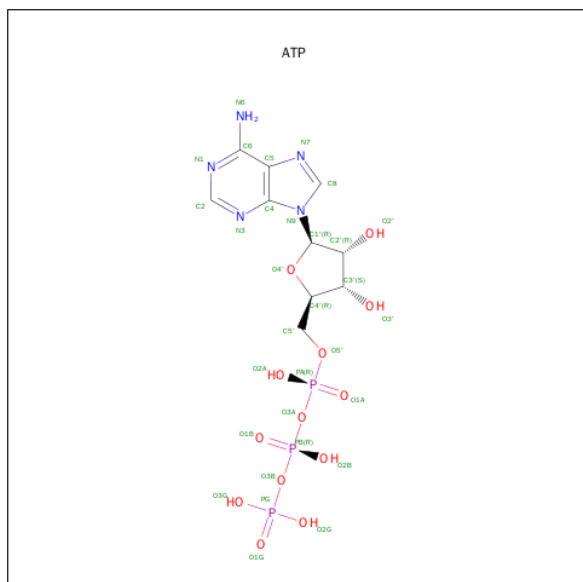
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	291	GLU	-	EXPRESSION TAG	UNP Q12125
E	292	ASN	-	EXPRESSION TAG	UNP Q12125
E	293	LEU	-	EXPRESSION TAG	UNP Q12125
E	294	TYR	-	EXPRESSION TAG	UNP Q12125
E	295	PHE	-	EXPRESSION TAG	UNP Q12125
E	296	GLN	-	EXPRESSION TAG	UNP Q12125
C	9	MET	-	EXPRESSION TAG	UNP Q12125
C	10	GLY	-	EXPRESSION TAG	UNP Q12125
C	258	ALA	LYS	ENGINEERED MUTATION	UNP Q12125
C	260	ALA	LYS	ENGINEERED MUTATION	UNP Q12125
C	291	GLU	-	EXPRESSION TAG	UNP Q12125
C	292	ASN	-	EXPRESSION TAG	UNP Q12125
C	293	LEU	-	EXPRESSION TAG	UNP Q12125
C	294	TYR	-	EXPRESSION TAG	UNP Q12125
C	295	PHE	-	EXPRESSION TAG	UNP Q12125
C	296	GLN	-	EXPRESSION TAG	UNP Q12125

- Molecule 3 is a protein called Ubiquitin-like protein MDY2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	F	52	Total	C	N	O	0	0	0
			410	270	64	76			
3	D	52	Total	C	N	O	0	0	0
			410	270	64	76			

- Molecule 4 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula:  $C_{10}H_{16}N_5O_{13}P_3$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
4	B	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

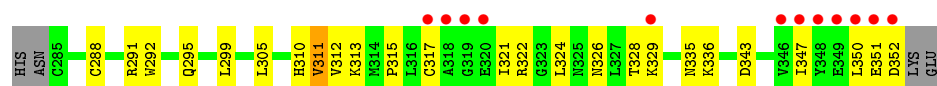
- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

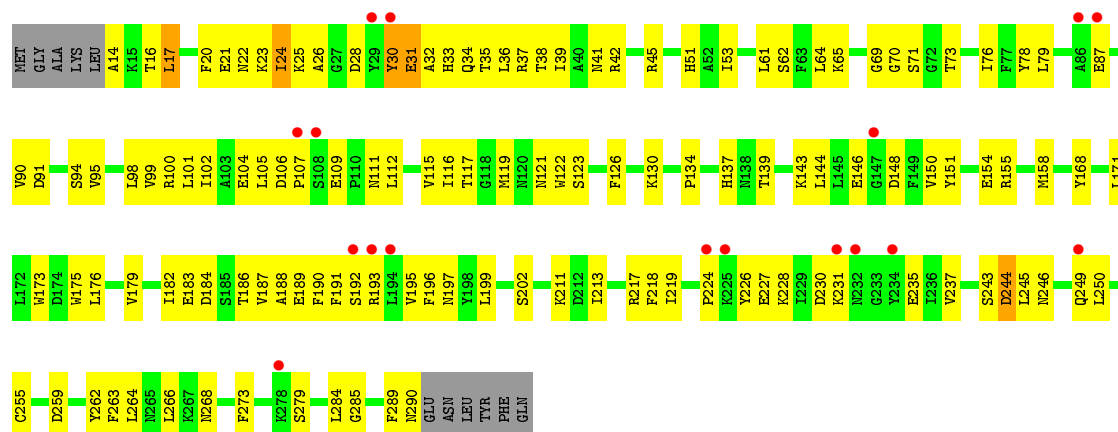
- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Zn	0	0
			1	1		

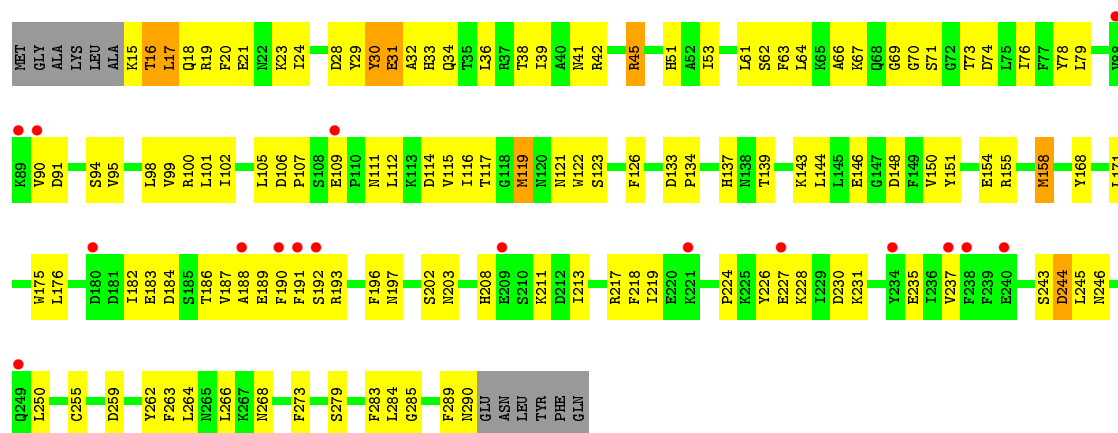




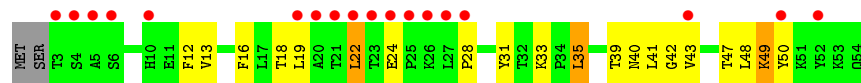
- Molecule 2: Golgi to ER traffic protein 4



- Molecule 2: Golgi to ER traffic protein 4



- Molecule 3: Ubiquitin-like protein MDY2



- Molecule 3: Ubiquitin-like protein MDY2







## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	166.34Å 134.50Å 84.06Å 90.00° 113.37° 90.00°	Depositor
Resolution (Å)	30.00 – 5.40 29.98 – 5.40	Depositor EDS
% Data completeness (in resolution range)	(Not available) (30.00-5.40) 94.1 (29.98-5.40)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.68 (at 5.34Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.270 , 0.328 0.283 , 0.337	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	355.6	Xtriage
Anisotropy	0.264	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , 287.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	0 of 5529 reflections	Xtriage
$F_o, F_c$ correlation	0.86	EDS
Total number of atoms	10159	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	347.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.34% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

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

## 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: ZN, MG, ATP

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	0.25	0/2357	0.45	0/3177
1	B	0.25	0/2396	0.45	0/3230
2	C	0.27	0/2353	0.42	0/3179
2	E	0.27	0/2358	0.42	0/3186
3	D	0.26	0/421	0.47	0/572
3	F	0.26	0/421	0.47	0/572
All	All	0.26	0/10306	0.44	0/13916

There are no bond length outliers.

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	2319	0	2324	164	0
1	B	2358	0	2356	164	0
2	C	2296	0	2218	121	0
2	E	2301	0	2223	130	0
3	D	410	0	433	22	0
3	F	410	0	433	24	2
4	A	31	0	12	5	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	31	0	12	3	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
6	A	1	0	0	0	0
All	All	10159	0	10011	557	2

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 28.

The worst 5 of 557 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:305:LEU:HG	2:C:42:ARG:HH22	1.17	1.06
1:A:28:GLY:HA3	1:B:28:GLY:HA3	1.44	0.98
1:A:305:LEU:HG	2:C:42:ARG:NH2	1.81	0.94
1:A:12:LEU:HD21	1:A:237:THR:HG21	1.48	0.93
2:E:158:MET:SD	2:E:197:ASN:HB3	2.11	0.91

All (2) 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	Interatomic distance (Å)	Clash overlap (Å)
3:F:22:LEU:CD2	3:F:22:LEU:CD2[2_855]	2.10	0.10
3:F:22:LEU:CD1	3:F:22:LEU:CD1[2_855]	2.14	0.06

## 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
1	A	286/356 (80%)	250 (87%)	32 (11%)	4 (1%)	14 58
1	B	291/356 (82%)	252 (87%)	36 (12%)	3 (1%)	19 64

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	C	274/288 (95%)	257 (94%)	16 (6%)	1 (0%)	39	80
2	E	275/288 (96%)	257 (94%)	16 (6%)	2 (1%)	26	71
3	D	50/54 (93%)	46 (92%)	4 (8%)	0	100	100
3	F	50/54 (93%)	45 (90%)	5 (10%)	0	100	100
All	All	1226/1396 (88%)	1107 (90%)	109 (9%)	10 (1%)	24	69

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	171	GLY
1	B	171	GLY
1	A	189	LYS
1	B	277	ALA
2	E	70	GLY

### 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	262/311 (84%)	237 (90%)	25 (10%)	11	41
1	B	267/311 (86%)	242 (91%)	25 (9%)	11	42
2	C	249/258 (96%)	235 (94%)	14 (6%)	26	63
2	E	249/258 (96%)	238 (96%)	11 (4%)	35	69
3	D	48/50 (96%)	44 (92%)	4 (8%)	14	49
3	F	48/50 (96%)	45 (94%)	3 (6%)	22	59
All	All	1123/1238 (91%)	1041 (93%)	82 (7%)	17	55

5 of 82 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	188	GLU
1	B	263	ASP

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
2	C	158	MET
1	B	189	LYS
1	B	240	CYS

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	172	HIS
1	B	301	GLN
2	C	121	ASN
1	B	222	ASN
1	B	310	HIS

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 3 are monoatomic - leaving 2 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	ATP	A	401	5	24,33,33	1.37	2 (8%)	31,52,52	3.63	12 (38%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	ATP	B	401	5	24,33,33	1.37	1 (4%)	31,52,52	3.60	10 (32%)

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	ATP	A	401	5	-	0/18/38/38	0/3/3/3
4	ATP	B	401	5	-	0/18/38/38	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	401	ATP	C8-N7	-2.02	1.30	1.34
4	B	401	ATP	O4'-C1'	4.34	1.46	1.41
4	A	401	ATP	O4'-C1'	4.45	1.46	1.41

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	401	ATP	N3-C2-N1	-9.87	121.33	128.89
4	A	401	ATP	N3-C2-N1	-9.54	121.59	128.89
4	B	401	ATP	PA-O3A-PB	-9.37	106.40	132.73
4	A	401	ATP	PA-O3A-PB	-8.96	107.56	132.73
4	A	401	ATP	O3A-PA-O5'	-7.59	82.80	102.94

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	401	ATP	5	0
4	B	401	ATP	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](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	294/356 (82%)	0.44	24 (8%) 14 16	193, 339, 374, 396	0
1	B	299/356 (83%)	0.62	37 (12%) 5 9	193, 339, 378, 396	0
2	C	276/288 (95%)	0.25	17 (6%) 24 23	280, 349, 386, 429	0
2	E	277/288 (96%)	0.34	17 (6%) 25 23	280, 349, 387, 429	0
3	D	52/54 (96%)	0.48	6 (11%) 6 10	333, 345, 405, 419	0
3	F	52/54 (96%)	1.22	18 (34%) 0 3	333, 345, 406, 419	0
All	All	1250/1396 (89%)	0.46	119 (9%) 10 13	193, 344, 382, 429	0

The worst 5 of 119 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	98	ASN	6.9
1	B	352	ASP	5.4
1	B	348	TYR	4.8
1	B	278	GLU	4.8
1	B	157	GLY	4.8

### 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 carbohydrates 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. 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
4	ATP	A	401	31/31	0.70	0.59	0.40	309,312,314,315	0
4	ATP	B	401	31/31	0.78	0.44	0.01	316,321,324,324	0
5	MG	A	402	1/1	0.76	0.41	-1.07	327,327,327,327	0
5	MG	B	402	1/1	0.96	0.29	-3.02	327,327,327,327	0
6	ZN	A	403	1/1	0.85	0.08	-	327,327,327,327	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.