



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 20, 2026 – 07:55 PM UTC

PDB ID : 2AJ4 / pdb\_00002aj4  
Title : Crystal structure of *Saccharomyces cerevisiae* Galactokinase in complex with galactose and Mg:AMPPNP  
Authors : Thoden, J.B.; Sellick, C.A.; Timson, D.J.; Reece, R.J.; Holden, H.M.  
Deposited on : 2005-08-01  
Resolution : 2.40 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

<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 : 4-5-2 with Phenix2.0  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtrriage (Phenix) : 2.0  
EDS : 3.0  
Buster-report : wwPDB partial adaption of 1.1.7 (2018)  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

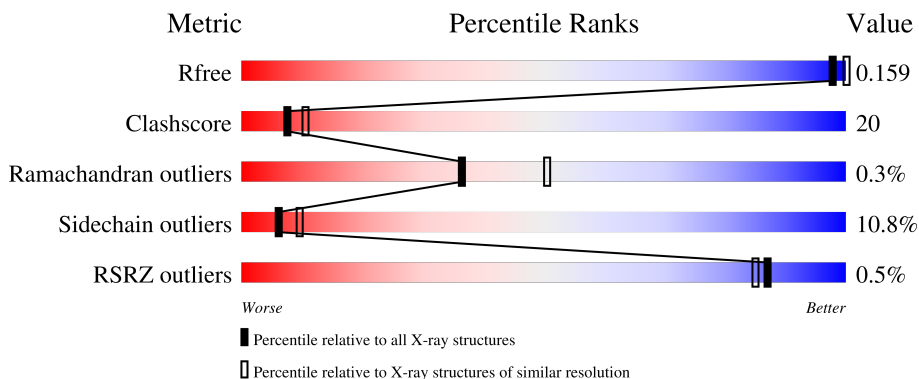
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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))
$R_{free}$	180053	4912 (2.40-2.40)
Clashscore	190562	5391 (2.40-2.40)
Ramachandran outliers	187476	5320 (2.40-2.40)
Sidechain outliers	187428	5321 (2.40-2.40)
RSRZ outliers	180081	4916 (2.40-2.40)

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  $\geq 3$ , 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	548	 53% 32% 6% • 7%
1	B	548	 47% 36% 8% • 8%

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 8129 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 Galactokinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	509	3938	2517	656	742	23	0	0	0
1	B	505	3910	2498	652	737	23	0	0	0

There are 42 discrepancies between the modelled and reference sequences:

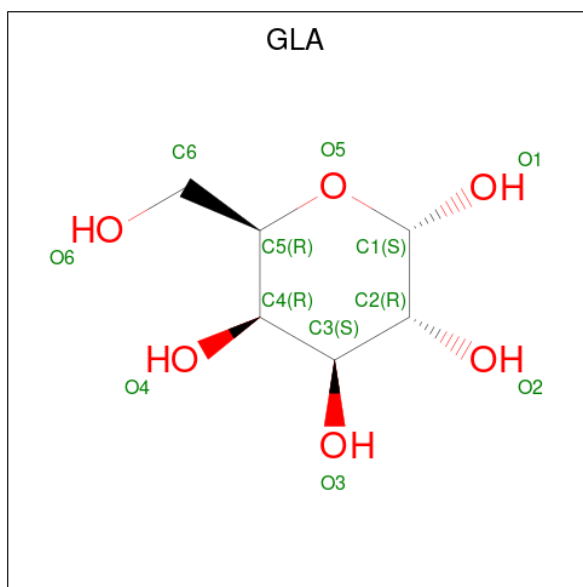
Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	cloning artifact	UNP P04385
A	-18	GLY	-	cloning artifact	UNP P04385
A	-17	SER	-	cloning artifact	UNP P04385
A	-16	SER	-	cloning artifact	UNP P04385
A	-15	HIS	-	expression tag	UNP P04385
A	-14	HIS	-	expression tag	UNP P04385
A	-13	HIS	-	expression tag	UNP P04385
A	-12	HIS	-	expression tag	UNP P04385
A	-11	HIS	-	expression tag	UNP P04385
A	-10	HIS	-	expression tag	UNP P04385
A	-9	SER	-	cloning artifact	UNP P04385
A	-8	SER	-	cloning artifact	UNP P04385
A	-7	GLU	-	cloning artifact	UNP P04385
A	-6	ASN	-	cloning artifact	UNP P04385
A	-5	LEU	-	cloning artifact	UNP P04385
A	-4	TYR	-	cloning artifact	UNP P04385
A	-3	PHE	-	cloning artifact	UNP P04385
A	-2	GLN	-	cloning artifact	UNP P04385
A	-1	GLY	-	cloning artifact	UNP P04385
A	0	HIS	-	cloning artifact	UNP P04385
A	1	MET	-	cloning artifact	UNP P04385
B	-19	MET	-	cloning artifact	UNP P04385
B	-18	GLY	-	cloning artifact	UNP P04385
B	-17	SER	-	cloning artifact	UNP P04385
B	-16	SER	-	cloning artifact	UNP P04385

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-15	HIS	-	expression tag	UNP P04385
B	-14	HIS	-	expression tag	UNP P04385
B	-13	HIS	-	expression tag	UNP P04385
B	-12	HIS	-	expression tag	UNP P04385
B	-11	HIS	-	expression tag	UNP P04385
B	-10	HIS	-	expression tag	UNP P04385
B	-9	SER	-	cloning artifact	UNP P04385
B	-8	SER	-	cloning artifact	UNP P04385
B	-7	GLU	-	cloning artifact	UNP P04385
B	-6	ASN	-	cloning artifact	UNP P04385
B	-5	LEU	-	cloning artifact	UNP P04385
B	-4	TYR	-	cloning artifact	UNP P04385
B	-3	PHE	-	cloning artifact	UNP P04385
B	-2	GLN	-	cloning artifact	UNP P04385
B	-1	GLY	-	cloning artifact	UNP P04385
B	0	HIS	-	cloning artifact	UNP P04385
B	1	MET	-	cloning artifact	UNP P04385

- Molecule 2 is alpha-D-galactopyranose (CCD ID: GLA) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			12	6	6		
2	B	1	Total	C	O	0	0
			12	6	6		

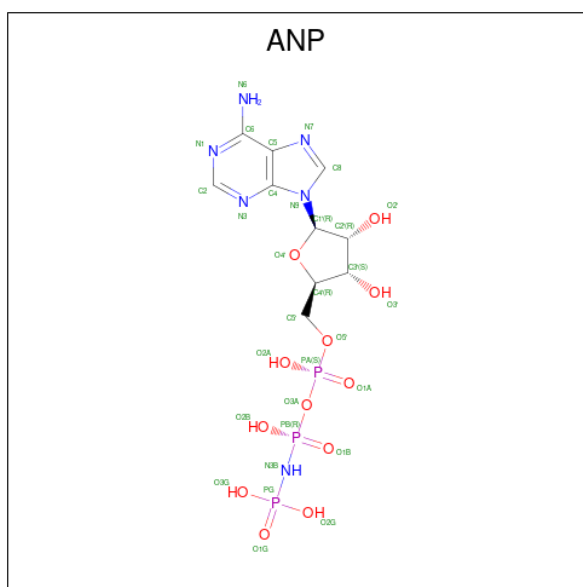
- Molecule 3 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

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

- Molecule 4 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

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

- Molecule 5 is PHOSPHOAMINOPHOSPHONIC ACID-ADENYLATE ESTER (CCD ID: ANP) (formula: C<sub>10</sub>H<sub>17</sub>N<sub>6</sub>O<sub>12</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C N O P 31 10 6 12 3	0	0
5	B	1	Total C N O P 31 10 6 12 3	0	0

- Molecule 6 is water.

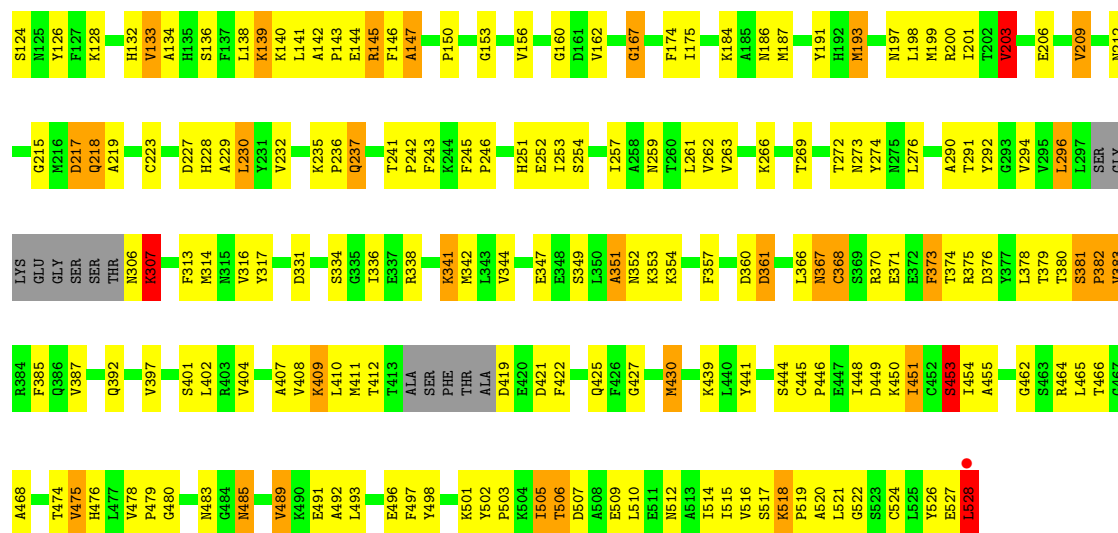
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	106	Total O 106 106	0	0

Continued on next page...

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
6	B	85	Total	O	0	0
			85	85		





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	63.30Å 85.00Å 112.30Å 90.00° 96.70° 90.00°	Depositor
Resolution (Å)	30.00 – 2.40 30.00 – 2.40	Depositor EDS
% Data completeness (in resolution range)	95.5 (30.00-2.40) 95.5 (30.00-2.40)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.75 (at 2.39Å)	Xtrriage
Refinement program	TNT	Depositor
R, $R_{free}$	0.172 , 0.205 0.164 , 0.159	Depositor DCC
$R_{free}$ test set	4490 reflections (10.12%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.1	Xtrriage
Anisotropy	0.356	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 111.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	8129	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.44% 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, 0.333 respectively for untwinned datasets, and 0.375, 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: MG, CL, GLA, ANP

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	1.15	1/4022 (0.0%)	1.69	57/5451 (1.0%)
1	B	1.17	3/3993 (0.1%)	1.68	57/5411 (1.1%)
All	All	1.16	4/8015 (0.0%)	1.69	114/10862 (1.0%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	528	LEU	C-OXT	15.41	1.54	1.23
1	A	528	LEU	C-OXT	9.12	1.41	1.23
1	B	528	LEU	C-O	-8.43	1.06	1.23
1	B	528	LEU	CA-C	6.92	1.67	1.52

The worst 5 of 114 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	528	LEU	CB-CA-C	-11.10	89.02	110.10
1	B	217	ASP	CA-CB-CG	-9.65	102.95	112.60
1	B	87	LYS	N-CA-C	-9.27	100.74	111.03
1	A	188	GLY	CA-C-N	9.17	129.23	119.05
1	A	188	GLY	C-N-CA	9.17	129.23	119.05

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	3938	0	3924	140	0
1	B	3910	0	3898	179	0
2	A	12	0	12	1	0
2	B	12	0	12	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	31	0	13	2	0
5	B	31	0	13	3	0
6	A	106	0	0	2	0
6	B	85	0	0	1	0
All	All	8129	0	7872	320	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 20.

The worst 5 of 320 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:B:203:VAL:HG22	1:B:215:GLY:HA3	1.42	0.98
1:B:218:GLN:HE21	1:B:218:GLN:H	1.11	0.95
1:A:218:GLN:NE2	1:A:218:GLN:H	1.67	0.93
1:B:218:GLN:H	1:B:218:GLN:NE2	1.66	0.92
1:B:505:ILE:HD13	1:B:510:LEU:HD12	1.53	0.89

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	
1	A	501/548 (91%)	483 (96%)	17 (3%)	1 (0%)	43	58
1	B	497/548 (91%)	460 (93%)	35 (7%)	2 (0%)	30	43
All	All	998/1096 (91%)	943 (94%)	52 (5%)	3 (0%)	36	50

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	74	ASP
1	B	23	ARG
1	B	453	SER

### 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	433/467 (93%)	392 (90%)	41 (10%)	8	13
1	B	431/467 (92%)	379 (88%)	52 (12%)	5	7
All	All	864/934 (92%)	771 (89%)	93 (11%)	6	9

5 of 93 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	133	VAL
1	B	347	GLU
1	B	139	LYS
1	B	209	VAL
1	B	383	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	364	GLN
1	B	495	ASN
1	B	88	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	102	GLN
1	B	155	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 oligosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

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
5	ANP	A	532	3	33,33,33	1.42	6 (18%)	45,52,52	1.31	4 (8%)
2	GLA	A	529	-	12,12,12	0.70	0	17,17,17	1.25	1 (5%)
5	ANP	B	532	3	33,33,33	1.52	6 (18%)	45,52,52	1.37	8 (17%)
2	GLA	B	529	-	12,12,12	0.61	0	17,17,17	0.91	1 (5%)

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	ANP	A	532	3	-	4/18/38/38	0/3/3/3
2	GLA	A	529	-	-	1/2/22/22	0/1/1/1
5	ANP	B	532	3	-	3/18/38/38	0/3/3/3
2	GLA	B	529	-	-	1/2/22/22	0/1/1/1

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	532	ANP	PB-O1B	4.10	1.52	1.46
5	A	532	ANP	C6-N6	-3.83	1.24	1.34
5	B	532	ANP	C6-N6	-3.76	1.24	1.34
5	A	532	ANP	PB-O1B	3.67	1.51	1.46
5	A	532	ANP	PG-O1G	3.29	1.51	1.46

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	532	ANP	C2-N1-C6	3.77	124.92	118.73
5	A	532	ANP	O1B-PB-N3B	-3.67	106.37	111.77
5	A	532	ANP	C2-N1-C6	3.63	124.70	118.73
5	B	532	ANP	N3-C2-N1	-3.46	123.35	128.58
5	A	532	ANP	N3-C2-N1	-3.31	123.57	128.58

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	532	ANP	PB-N3B-PG-O1G
5	A	532	ANP	PG-N3B-PB-O1B
5	B	532	ANP	PB-N3B-PG-O1G
5	B	532	ANP	PG-N3B-PB-O1B
2	A	529	GLA	O5-C5-C6-O6

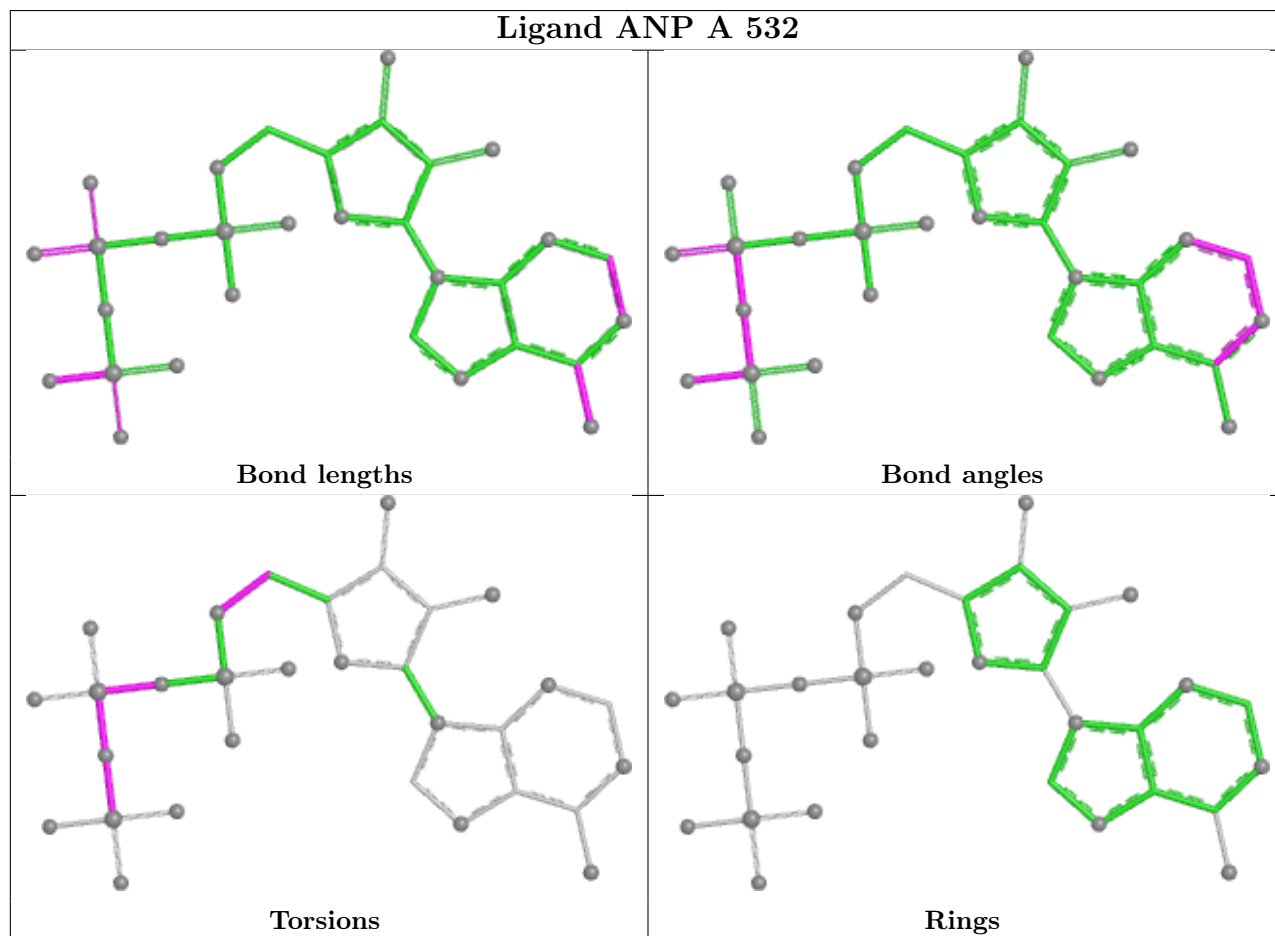
There are no ring outliers.

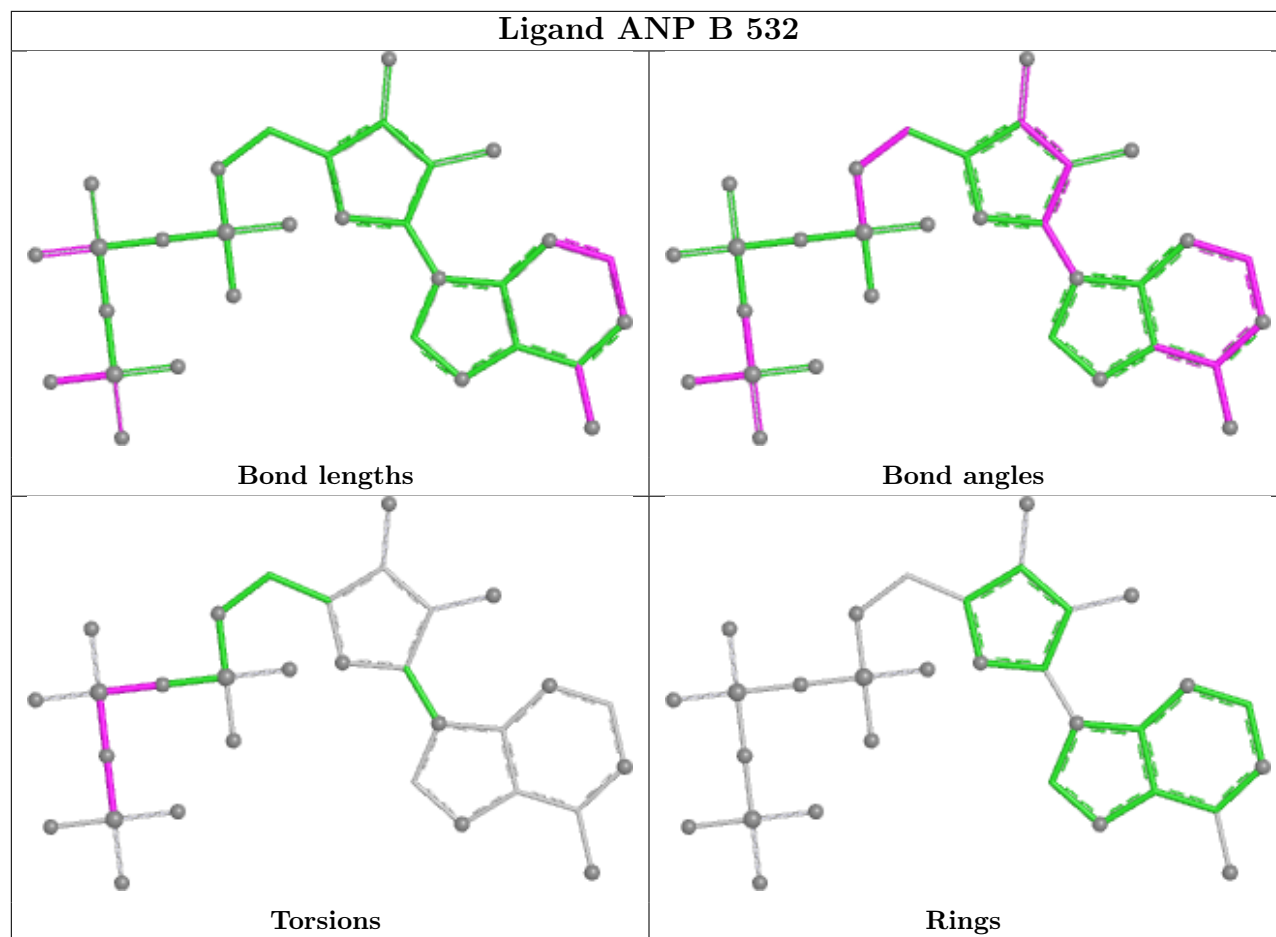
3 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	532	ANP	2	0
2	A	529	GLA	1	0
5	B	532	ANP	3	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 than 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<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	509/548 (92%)	-0.55	3 (0%) 85 83	12, 28, 67, 96	0
1	B	505/548 (92%)	-0.18	2 (0%) 88 86	13, 39, 85, 98	0
All	All	1014/1096 (92%)	-0.37	5 (0%) 87 85	12, 33, 80, 98	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	416	PHE	4.4
1	B	528	LEU	3.2
1	A	120	VAL	2.9
1	B	14	PHE	2.4
1	A	413	THR	2.0

### 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](#)

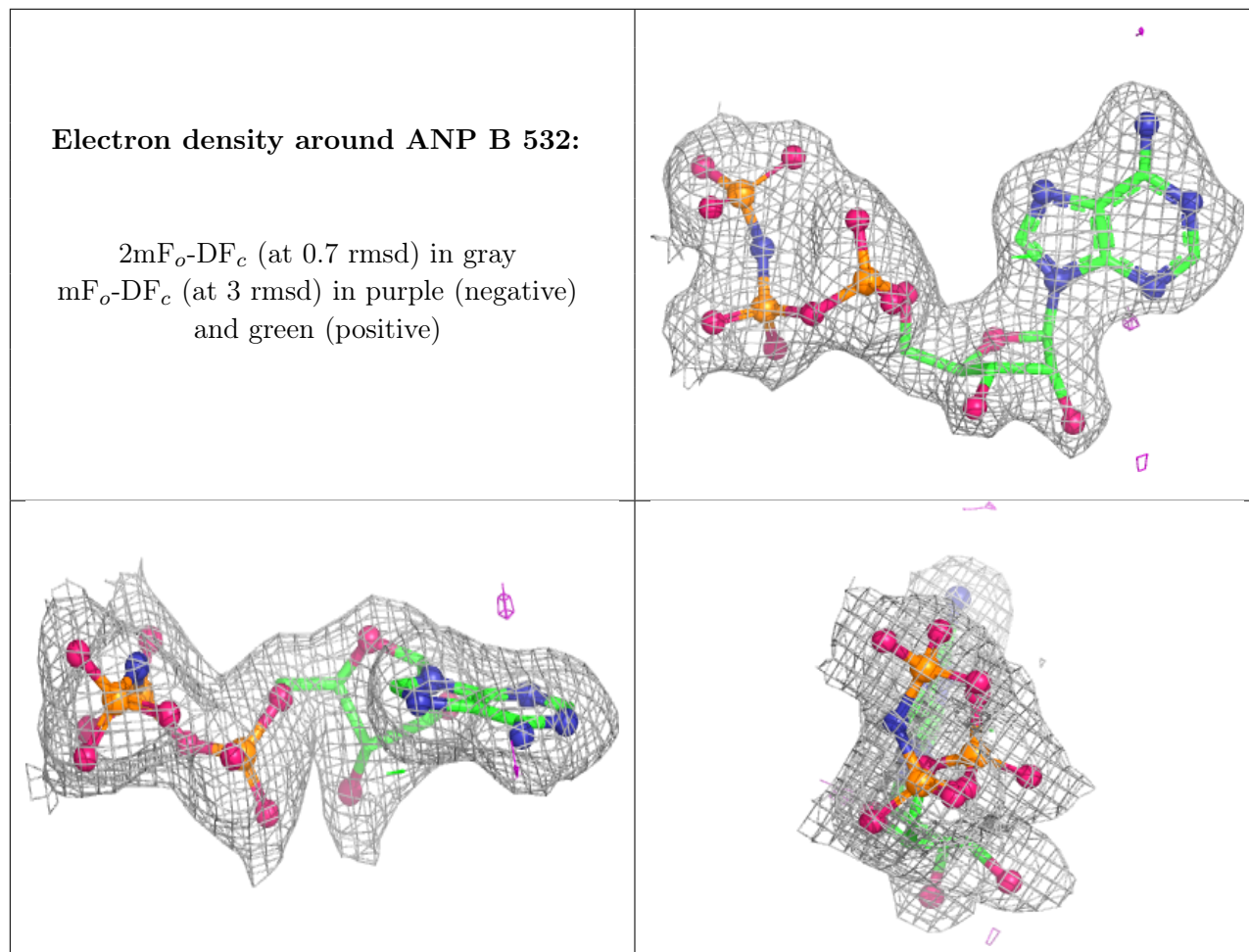
There are no oligosaccharides in this entry.

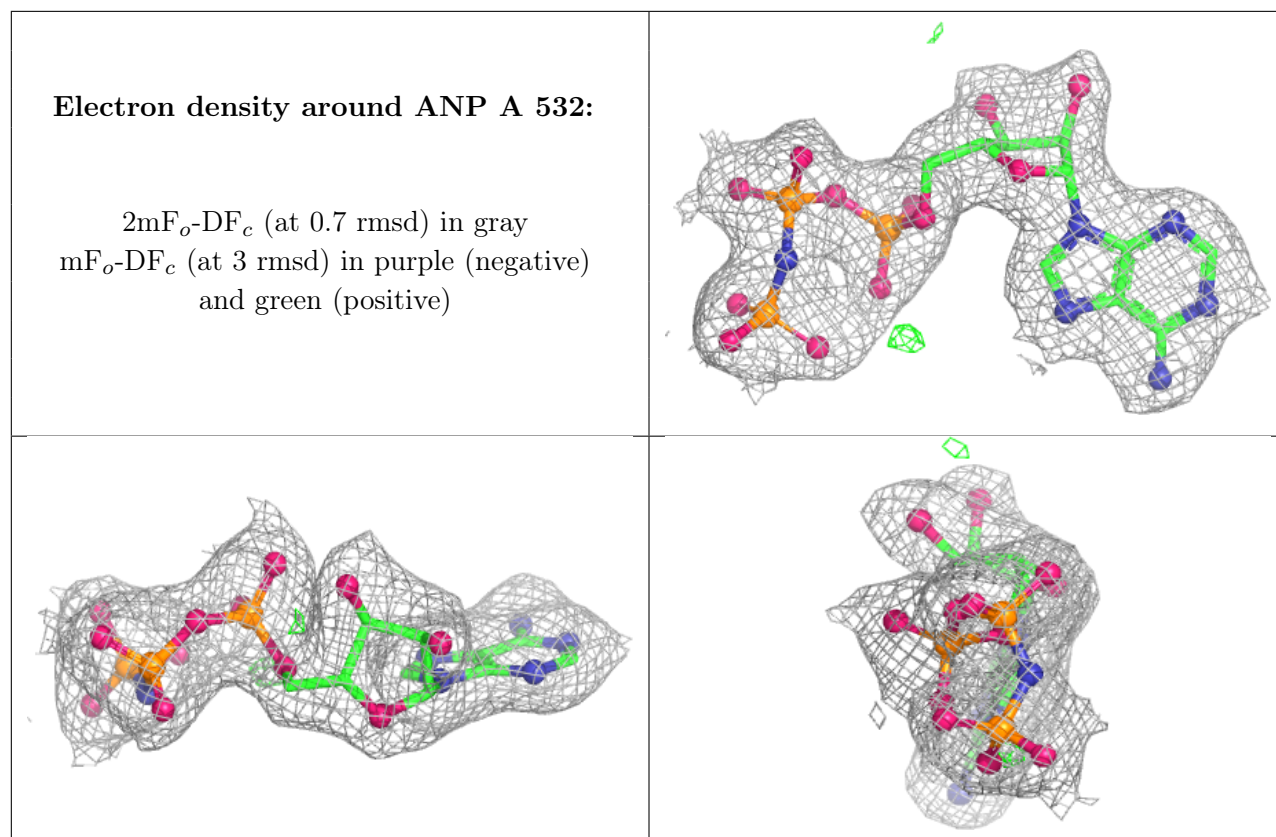
### 6.4 Ligands [i](#)

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<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	B-factors( $\text{\AA}^2$ )	Q<0.9
3	MG	B	530	1/1	0.95	0.10	44,44,44,44	0
4	CL	B	531	1/1	0.96	0.07	48,48,48,48	0
3	MG	A	530	1/1	0.97	0.07	34,34,34,34	0
2	GLA	B	529	12/12	0.98	0.05	11,17,30,33	0
4	CL	A	531	1/1	0.98	0.03	33,33,33,33	0
2	GLA	A	529	12/12	0.98	0.05	5,14,22,27	0
5	ANP	B	532	31/31	0.98	0.05	9,28,51,95	0
5	ANP	A	532	31/31	0.99	0.05	10,22,36,51	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.