wwPDB X-ray Structure Validation Summary Report (i)

Oct 22, 2023 – 09:07 AM EDT

PDB ID	:	2ZR2
Title	:	Crystal structure of seryl-tRNA synthetase from Pyrococcus horikoshii complexed with ATP
Authors	:	Itoh, Y.; Sekine, S.; Yokoyama, S.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on	:	2008-08-22
Resolution	:	2.80 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:

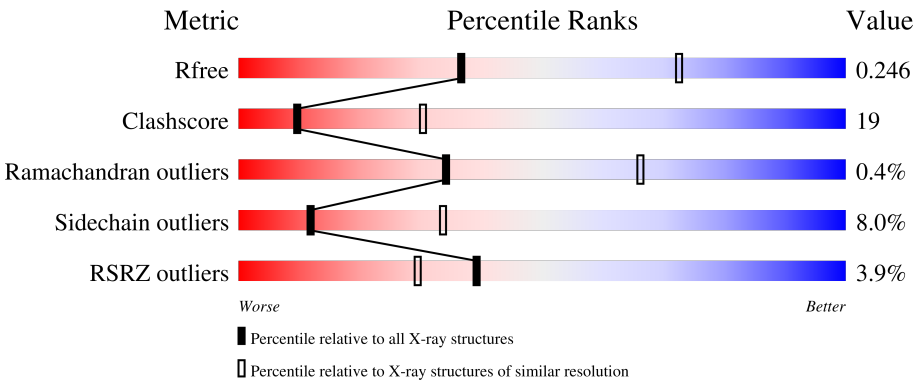
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 2.80 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}	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	455	4% 58% 37%
1	B	455	4% 60% 34%

2 Entry composition (i)

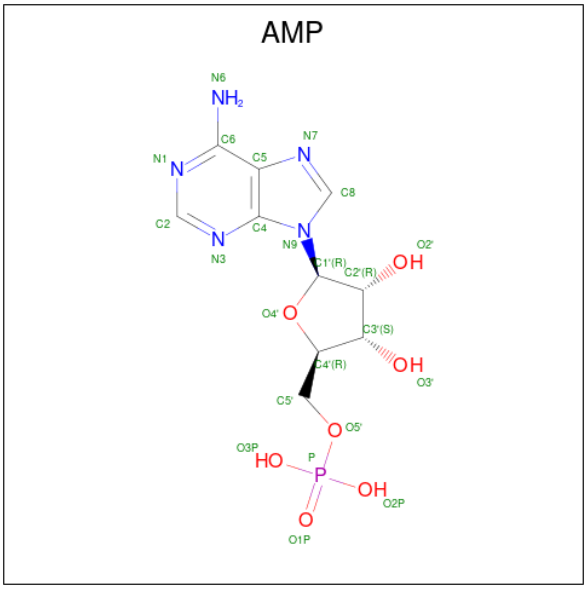
There are 4 unique types of molecules in this entry. The entry contains 7551 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Seryl-tRNA synthetase.

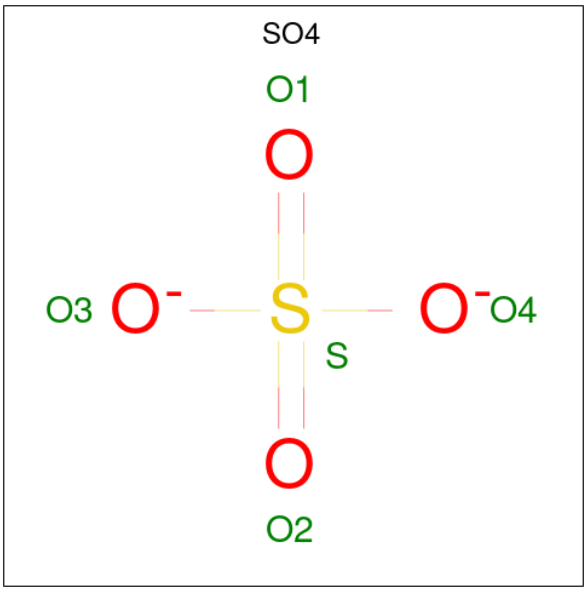
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	447	3699	2376	645	670	8	0	0	0
1	B	447	3699	2376	645	670	8	0	0	0

- Molecule 2 is ADENOSINE MONOPHOSPHATE (three-letter code: AMP) (formula: $C_{10}H_{14}N_5O_7P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
2	A	1	23	10	5	7	1	0	0
2	B	1	23	10	5	7	1	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

- Molecule 4 is water.

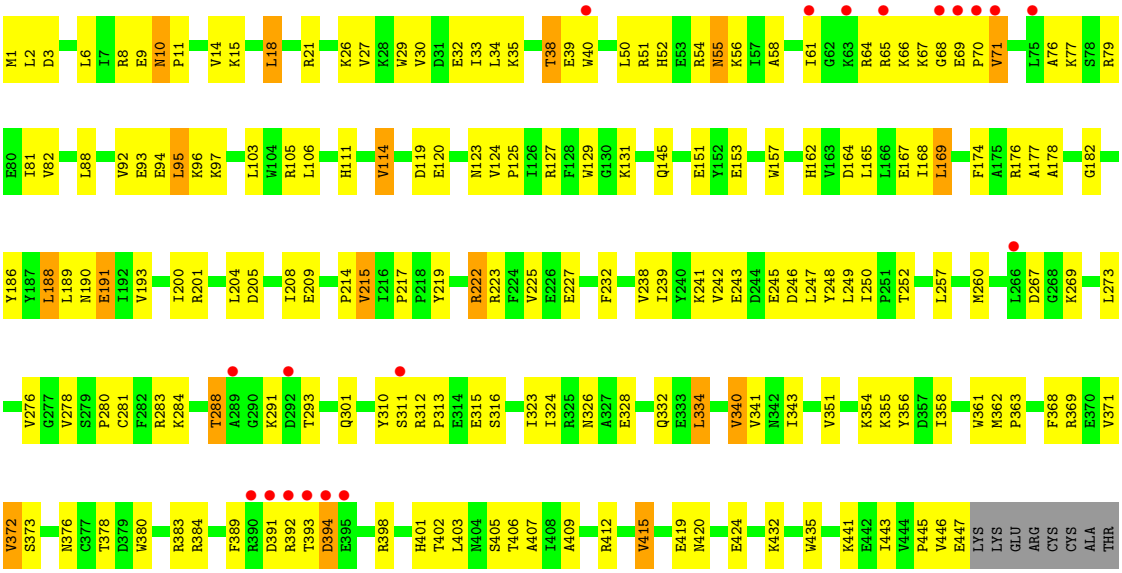
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	35	Total O 35 35	0	0
4	B	37	Total O 37 37	0	0

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.

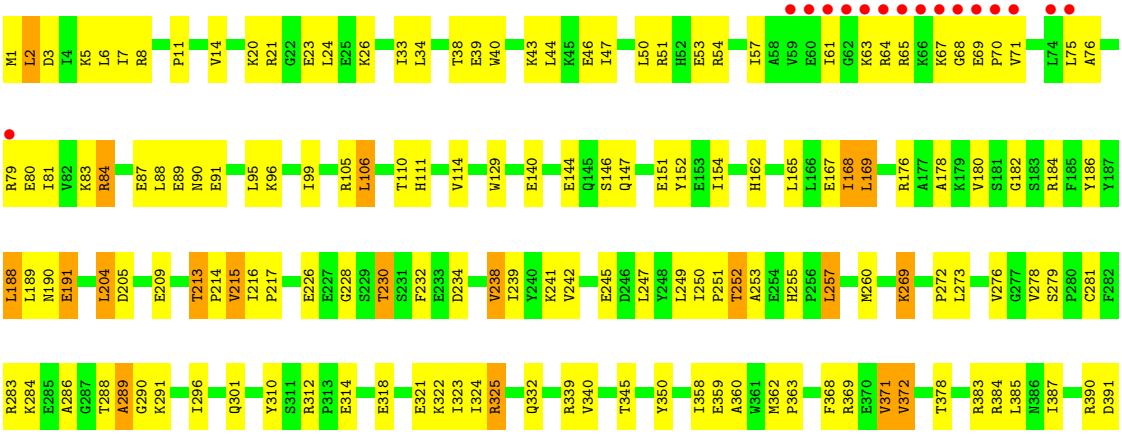
- Molecule 1: Seryl-tRNA synthetase

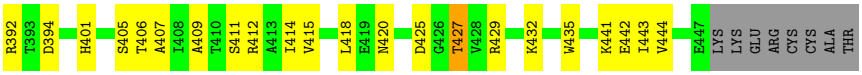
Chain A:



- Molecule 1: Seryl-tRNA synthetase

Chain B:




R392
T393
D394
H401
S405
T406
A407
I408
A409
T410
S411
R412
A413
I414
V415
L418
E419
N420
D425
G426
T427
V428
R429
K432
W435
K441
E442
I443
V444
E447
LYS
LYS
GLU
ARG
CYS
CYS
ALA
THR

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	96.66Å 120.50Å 126.78Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.52 – 2.80 48.52 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.52-2.80) 99.5 (48.52-2.80)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	2.90 (at 2.81Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.192 , 0.251 0.189 , 0.246	Depositor DCC
R_{free} test set	1873 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	67.7	Xtriage
Anisotropy	0.385	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 62.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.008 for -h,l,k	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	7551	wwPDB-VP
Average B, all atoms (Å^2)	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.42% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]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.

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: SO4, AMP

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	0.46	0/3783	0.71	2/5109 (0.0%)
1	B	0.49	0/3783	0.73	2/5109 (0.0%)
All	All	0.47	0/7566	0.72	4/10218 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	407	ALA	N-CA-C	-5.52	96.11	111.00
1	A	409	ALA	N-CA-C	-5.50	96.16	111.00
1	B	409	ALA	N-CA-C	-5.49	96.19	111.00
1	B	407	ALA	N-CA-C	-5.38	96.48	111.00

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	3699	0	3737	153	0
1	B	3699	0	3737	148	0
2	A	23	0	12	1	0
2	B	23	0	12	3	0
3	A	20	0	0	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	15	0	0	0	0
4	A	35	0	0	2	0
4	B	37	0	0	4	0
All	All	7551	0	7498	290	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 19.

The worst 5 of 290 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:40:TRP:HE1	1:B:96:LYS:HE3	1.27	0.96
1:B:358:ILE:HB	1:B:372:VAL:HG13	1.50	0.92
1:A:64:ARG:HH21	1:A:70:PRO:HD2	1.30	0.92
1:B:273:LEU:HB2	1:B:310:TYR:HB2	1.53	0.91
1:A:168:ILE:HD12	1:A:445:PRO:HB3	1.57	0.85

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	445/455 (98%)	414 (93%)	28 (6%)	3 (1%)	22 53
1	B	445/455 (98%)	416 (94%)	28 (6%)	1 (0%)	47 78
All	All	890/910 (98%)	830 (93%)	56 (6%)	4 (0%)	34 66

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	10	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	393	THR
1	B	289	ALA
1	A	71	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	398/405 (98%)	368 (92%)	30 (8%)	13	37
1	B	398/405 (98%)	364 (92%)	34 (8%)	10	31
All	All	796/810 (98%)	732 (92%)	64 (8%)	12	34

5 of 64 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	272	PRO
1	B	340	VAL
1	A	334	LEU
1	A	293	THR
1	B	371	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	55	ASN
1	A	147	GLN
1	B	90	ASN
1	B	342	ASN

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)

9 ligands are modelled in this entry.

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	SO4	A	2007	-	4,4,4	0.30	0	6,6,6	0.22	0
3	SO4	A	2003	-	4,4,4	0.27	0	6,6,6	0.14	0
2	AMP	B	1002	-	22,25,25	1.35	1 (4%)	25,38,38	1.46	2 (8%)
3	SO4	A	2005	-	4,4,4	0.32	0	6,6,6	0.17	0
3	SO4	B	2004	-	4,4,4	0.25	0	6,6,6	0.10	0
3	SO4	B	2006	-	4,4,4	0.31	0	6,6,6	0.12	0
2	AMP	A	1001	-	22,25,25	1.52	2 (9%)	25,38,38	2.08	6 (24%)
3	SO4	B	2002	-	4,4,4	0.28	0	6,6,6	0.17	0
3	SO4	A	2001	-	4,4,4	0.28	0	6,6,6	0.07	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	AMP	B	1002	-	-	0/6/26/26	0/3/3/3
2	AMP	A	1001	-	-	1/6/26/26	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1001	AMP	O4'-C1'	5.93	1.49	1.41
2	B	1002	AMP	O4'-C1'	5.09	1.48	1.41
2	A	1001	AMP	P-O3P	2.10	1.62	1.54

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1001	AMP	N3-C2-N1	-5.25	120.47	128.68
2	B	1002	AMP	N3-C2-N1	-5.12	120.67	128.68
2	A	1001	AMP	O3P-P-O5'	-4.92	93.64	106.73
2	A	1001	AMP	O2P-P-O5'	-4.74	94.12	106.73
2	B	1002	AMP	O5'-P-O1P	3.48	116.22	106.47

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1001	AMP	C5'-O5'-P-O3P

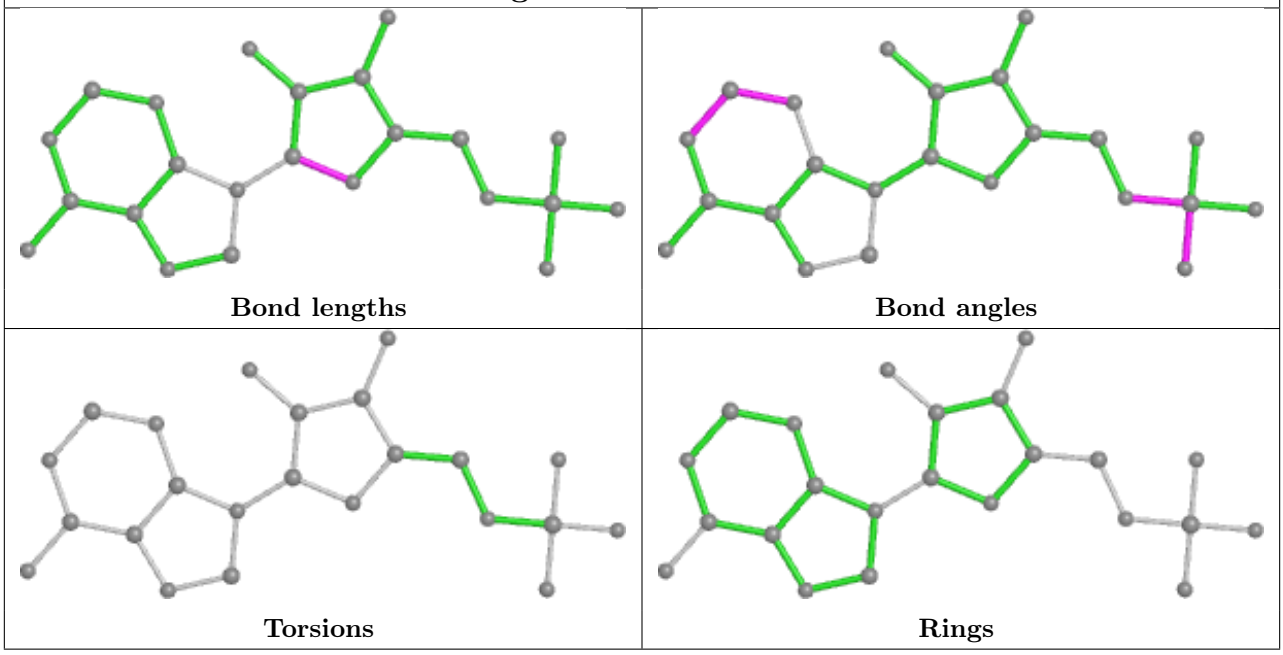
There are no ring outliers.

4 monomers are involved in 6 short contacts:

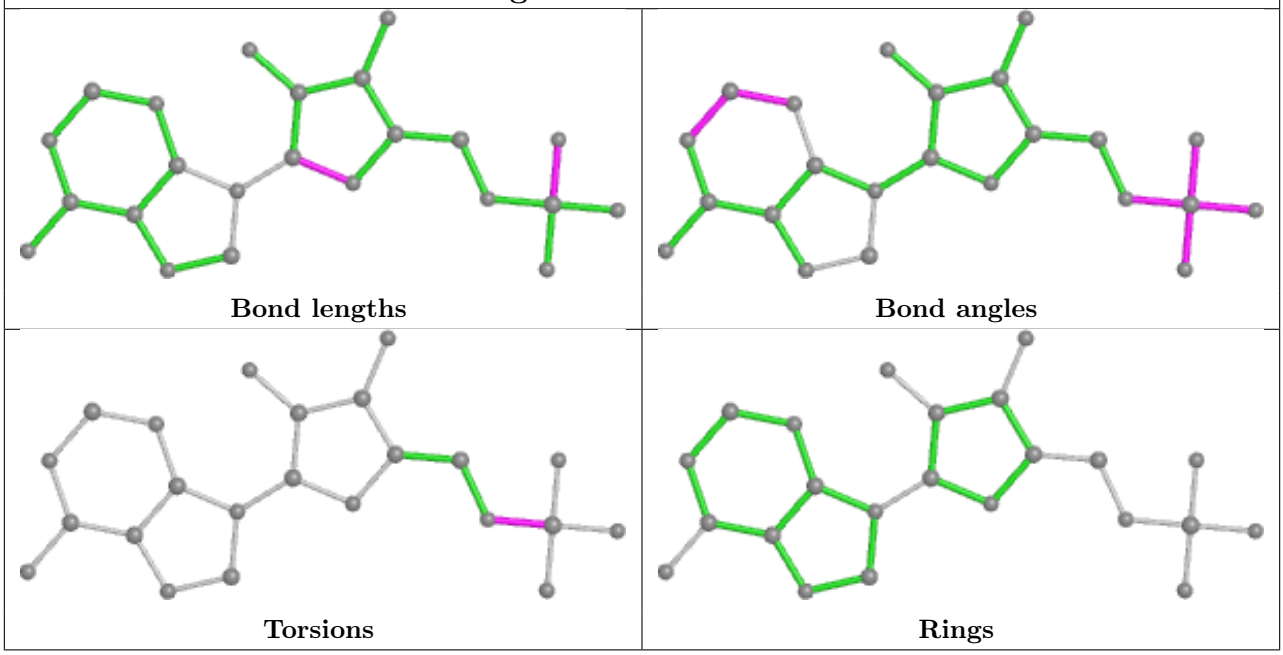
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	2007	SO4	1	0
2	B	1002	AMP	3	0
3	A	2005	SO4	1	0
2	A	1001	AMP	1	0

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.

Ligand AMP B 1002



Ligand AMP A 1001



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, 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	447/455 (98%)	0.06	19 (4%)	35	25	36, 61, 103, 159	0
1	B	447/455 (98%)	-0.05	16 (3%)	42	32	34, 57, 111, 165	0
All	All	894/910 (98%)	0.00	35 (3%)	39	29	34, 59, 108, 165	0

The worst 5 of 35 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	70	PRO	5.5
1	A	71	VAL	4.9
1	B	68	GLY	4.8
1	B	67	LYS	4.1
1	B	61	ILE	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)

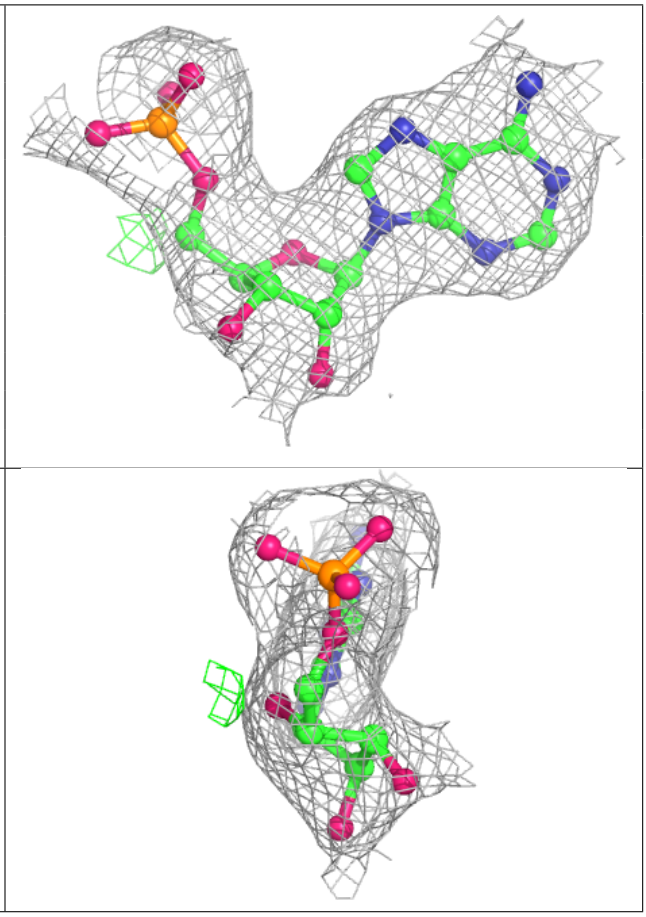
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	SO4	B	2006	5/5	0.84	0.24	127,129,131,133	0
3	SO4	A	2003	5/5	0.91	0.17	113,117,119,123	0
3	SO4	B	2004	5/5	0.94	0.15	122,126,128,129	0
3	SO4	A	2005	5/5	0.94	0.26	100,100,108,111	0
2	AMP	B	1002	23/23	0.96	0.18	45,55,60,64	0
3	SO4	A	2001	5/5	0.96	0.22	89,91,95,99	0
3	SO4	B	2002	5/5	0.97	0.12	90,91,98,100	0
2	AMP	A	1001	23/23	0.97	0.23	44,56,64,73	0
3	SO4	A	2007	5/5	0.97	0.22	86,90,94,98	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.

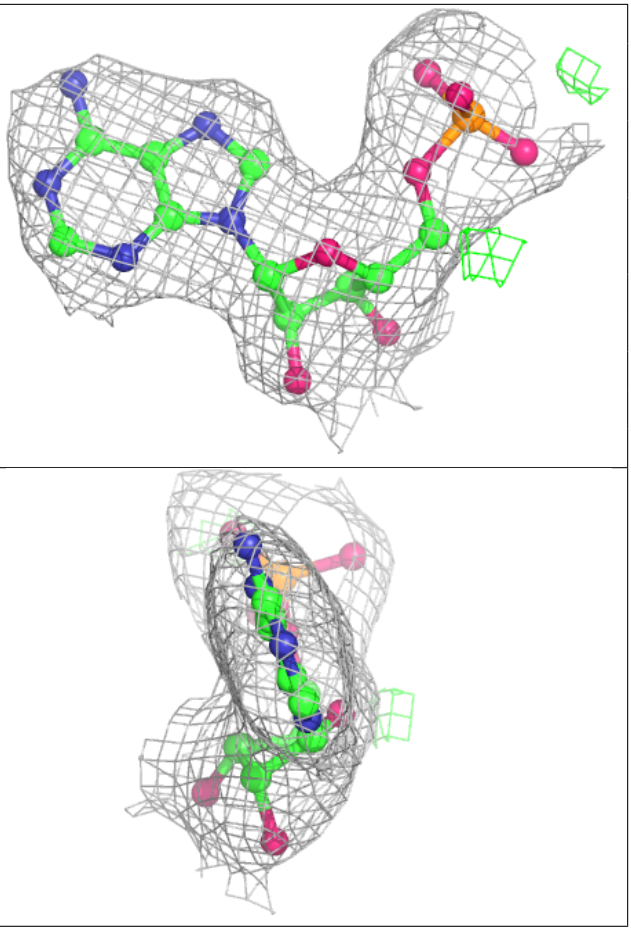
Electron density around AMP B 1002:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative) and green (positive)



Electron density around AMP A 1001:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



6.5 Other polymers (i)

There are no such residues in this entry.