wwPDB X-ray Structure Validation Summary Report (i)

Feb 17, 2024 – 04:14 PM EST

PDB ID	:	3TJZ
Title	:	Crystal Structure of Arf1 Bound to the gamma/zeta-COP Core Complex
Authors	:	Goldberg, J.; Yu, X.; Breitman, M.
Deposited on	:	2011-08-25
Resolution	:	2.90 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 (i)) were used in the production of this report:

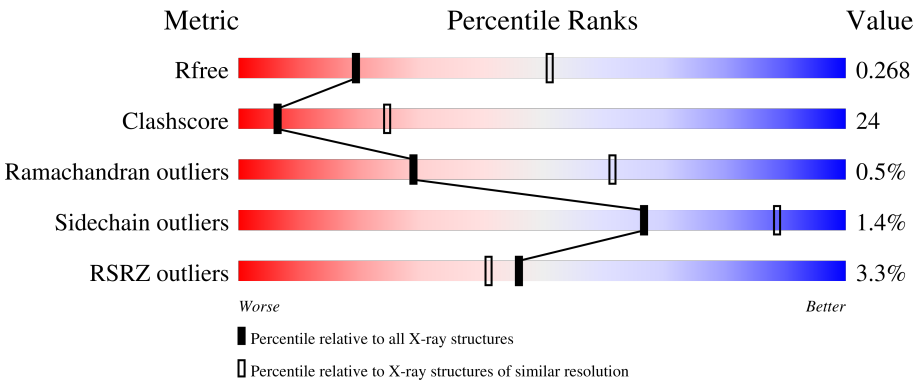
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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	164	76% 21% •
1	D	164	% 66% 30% • •
2	B	355	3% 48% 23% • 26%
2	E	355	3% 43% 30% • 27%
3	C	153	65% 25% • 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	153	10% 41% 50% 8%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 8970 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 ADP-ribosylation factor 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	159	1273	805	224	238	6	0	0	0
1	D	159	1273	805	224	238	6	0	0	0

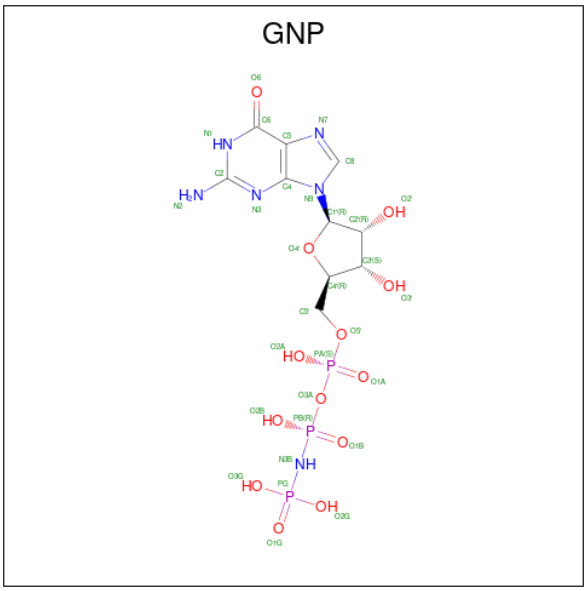
- Molecule 2 is a protein called Coatomer subunit gamma.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	264	2069	1305	366	379	19	0	0	0
2	E	260	2033	1283	359	372	19	0	0	0

- Molecule 3 is a protein called Coatomer subunit zeta-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	140	1128	723	179	221	5	0	0	0
3	F	140	1128	723	179	221	5	0	0	0

- Molecule 4 is PHOSPHOAMINOPHOSPHONIC ACID-GUANYLATE ESTER (three-letter code: GNP) (formula: $C_{10}H_{17}N_6O_{13}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
4	A	1	32	10	6	13	3	0	0
			Total	C	N	O	P		
4	D	1	32	10	6	13	3	0	0

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

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

3 Residue-property plots (i)

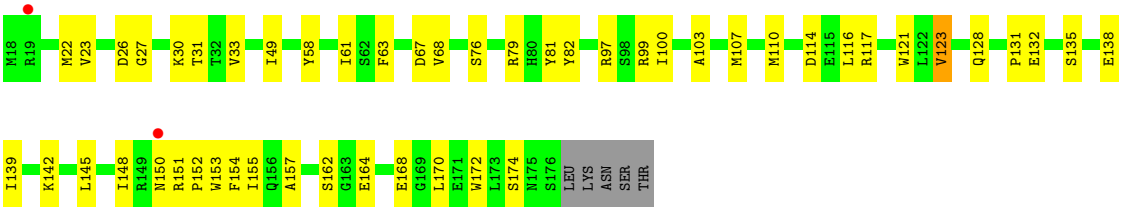
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: ADP-ribosylation factor 1

Chain A:

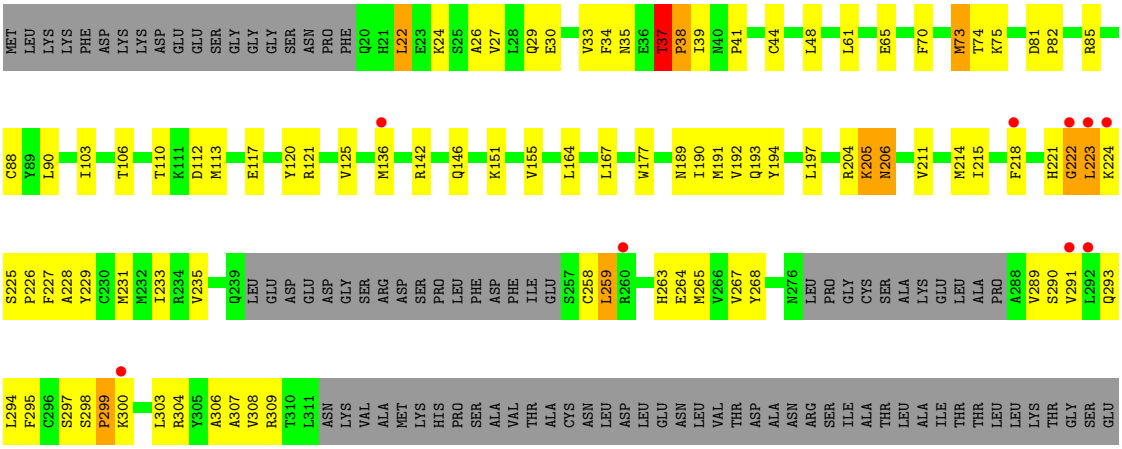
• Molecule 1: ADP-ribosylation factor 1

Chain D:



• Molecule 2: Coatomer subunit gamma

Chain B:



GLY

- Molecule 2: Coatomer subunit gamma

Chain E:

MET LEU LYS LYS PHE ASP LYS LYS ASP GLU GLU SER GLY GLY GLY SER ASN PRO PHE GLN HIS LEU GLU K24 S25 A26 E30 F34 N35 E36 T37 P38 I39 N40 P41 I47 I51 L52 Y53 L54 I55 N56 Q57 F70 F71 A72 M73 T74 K75 L76 F77 Q78 D81 P82

T83 L84 R85 R86 M87 C88 Y89 L90 T91 I92 K93 E94 I104 V105 T106 D112 M113 E117 Y120 R121 V125 S134 L137 V149 D150 K151 V161 L164 V173 V174 K175 V178 N179 E180 N189 I190 M191 V192 Q193 L197 G198 L199 L200 Y201 H202 V203

R204 K205 N206 D207 R208 V211 I215 S216 K217 F218 T219 R220 H221 G222 L223 K224 S225 P226 F227 A228 Y229 C230 M231 M232 V235 A236 Q239 LEU GLU ASP GLU ASP GLY SER ARG ASP SER PRO LEU PHE ASP PHE ILE GLU S257 C258 L259 R260 N261 K262 H263 E264 M265 V266 V267 Y268

E269 A270 A271 I274 V275 N276 LEU PRO GLY CYS SER ALA LYS GLU LEU ALA PRO A288 V289 S290 V291 L292 Q293 L294 F295 C296 S297 S298 P299 K300 A301 A302 L303 R304 Y305 R309 T310 L311 ASN LYS VAL ALA MET LYS HIS PRO SER ALA VAL THR ALA CYS ASN LEU ASP LEU GLU ASN

LEU VAL THR ASP ALA ASN ARG SER ILE ALA THR LEU ALA ILE THR THR LEU LEU LYS THR GLY SER GLU GLY

- Molecule 3: Coatomer subunit zeta-1

Chain C:

MET GLU ALA LEU ILE LEU E7 L10 K14 N21 D32 D33 T34 Y35 P36 F44 I48 F49 D56 S57 E58 I59 L62 E63 V68 Y69 K70 S71 S72 I73 D74 S82 E85 N86 E87 L88 M91 T92 D99 Q103 R106 K107 K111 R112

A113 E116 L123 A124 I128 D138 P139 Q140 Q141 V142 V143 H144 R145 V146 ALA LEU ARG GLY GLU ASP VAL

- Molecule 3: Coatomer subunit zeta-1

Chain F:

MET GLU ALA LEU ILE LEU E7 P8 S9 L10 V13 K14 A15 I16 L17 I18 L19 D20 N21 D22 G23 D24 R25 L26 F27 A28 K29 Y30 Y31 Y35 P36 S37 E40 Q41 K42 A43 F44 E45 K46 N47 I48 F49 N50 K51 T52 T55 D56 S57 E58 I59 A60 L61 L62 E63 G64 L65

T66 V67 V68 Y69 K70 S71 S72 I73 D74 L75 Y76 F77 Y78 V79 I80 G81 S82 S83 Y84 E85 N86 E87 L88 M89 L90 M91 T92 V93 L94 F98 L101 M104 L105 V109 E110 K111 R112 A113 L114 L115 L121 F122 L123 A124 V125 D126 E127 I128 V129 L135 E136 S137 D138

Q141 V146 ALA LEU ARG GLY GLU ASP VAL

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, α, β, γ	163.58Å 163.58Å 145.19Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.00 – 2.90 48.73 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.9 (45.00-2.90) 96.5 (48.73-2.90)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.64 (at 2.91Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.1_743)	Depositor
R, R_{free}	0.206 , 0.260 0.224 , 0.268	Depositor DCC
R_{free} test set	2000 reflections (4.53%)	wwPDB-VP
Wilson B-factor (Å^2)	67.6	Xtriage
Anisotropy	0.059	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 57.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	8970	wwPDB-VP
Average B, all atoms (Å^2)	72.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.60% 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: GNP, 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).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\|>5$	RMSZ	$\#\|Z\|>5$
1	A	0.51	0/1297	0.70	0/1755
1	D	0.46	0/1297	0.66	0/1755
2	B	0.47	1/2100 (0.0%)	0.68	4/2831 (0.1%)
2	E	0.42	0/2063	0.63	0/2781
3	C	0.50	0/1146	0.66	0/1549
3	F	0.37	0/1146	0.57	0/1549
All	All	0.46	1/9049 (0.0%)	0.65	4/12220 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	38	PRO	N-CD	-5.97	1.39	1.47

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	206	ASN	CB-CA-C	-9.38	91.64	110.40
2	B	37	THR	N-CA-C	5.53	125.93	111.00
2	B	37	THR	CB-CA-C	-5.11	97.81	111.60
2	B	223	LEU	N-CA-C	5.03	124.58	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	1273	0	1263	20	0
1	D	1273	0	1263	47	0
2	B	2069	0	2127	131	0
2	E	2033	0	2095	128	0
3	C	1128	0	1127	47	0
3	F	1128	0	1127	75	0
4	A	32	0	13	1	0
4	D	32	0	13	1	0
5	A	1	0	0	0	0
5	D	1	0	0	0	0
All	All	8970	0	9028	431	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 24.

The worst 5 of 431 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)
2:B:37:THR:CG2	2:B:38:PRO:HD3	1.35	1.57
2:B:37:THR:HG22	2:B:38:PRO:CD	1.46	1.45
1:D:150:ASN:OD1	1:D:151:ARG:HG2	1.25	1.33
2:B:299:PRO:O	2:B:300:LYS:HG3	1.23	1.29
1:D:150:ASN:OD1	1:D:151:ARG:CG	1.88	1.22

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	157/164 (96%)	153 (98%)	4 (2%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	157/164 (96%)	152 (97%)	5 (3%)	0	100	100
2	B	258/355 (73%)	243 (94%)	12 (5%)	3 (1%)	13	40
2	E	254/355 (72%)	237 (93%)	16 (6%)	1 (0%)	34	66
3	C	138/153 (90%)	132 (96%)	5 (4%)	1 (1%)	22	54
3	F	138/153 (90%)	127 (92%)	11 (8%)	0	100	100
All	All	1102/1344 (82%)	1044 (95%)	53 (5%)	5 (0%)	29	61

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	222	GLY
2	E	291	VAL
2	B	307	ALA
2	B	299	PRO
3	C	73	ILE

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	136/141 (96%)	135 (99%)	1 (1%)	84	95
1	D	136/141 (96%)	134 (98%)	2 (2%)	65	87
2	B	231/307 (75%)	226 (98%)	5 (2%)	52	81
2	E	227/307 (74%)	224 (99%)	3 (1%)	69	90
3	C	128/138 (93%)	126 (98%)	2 (2%)	62	86
3	F	128/138 (93%)	127 (99%)	1 (1%)	81	94
All	All	986/1172 (84%)	972 (99%)	14 (1%)	67	89

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	68	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	76	SER
3	F	71	SER
2	E	229	TYR
2	E	265	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	130	GLN
2	B	193	GLN
2	B	293	GLN
2	E	78	GLN
2	E	193	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 4 ligands modelled in this entry, 2 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	GNP	A	201	5	29,34,34	1.55	6 (20%)	33,54,54	2.55	10 (30%)
4	GNP	D	201	5	29,34,34	1.40	5 (17%)	33,54,54	2.51	9 (27%)

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	GNP	A	201	5	-	8/14/38/38	0/3/3/3
4	GNP	D	201	5	-	5/14/38/38	0/3/3/3

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	201	GNP	PB-O3A	-3.75	1.54	1.59
4	A	201	GNP	C6-N1	3.45	1.39	1.33
4	D	201	GNP	C6-N1	3.43	1.39	1.33
4	D	201	GNP	PB-O3A	-2.97	1.55	1.59
4	A	201	GNP	PG-O1G	2.95	1.50	1.46

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	201	GNP	C5-C6-N1	-8.71	111.52	123.43
4	D	201	GNP	C5-C6-N1	-8.68	111.56	123.43
4	D	201	GNP	C2-N1-C6	6.04	125.53	115.93
4	A	201	GNP	C2-N1-C6	6.01	125.48	115.93
4	D	201	GNP	O3G-PG-O1G	-5.02	100.83	113.45

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	201	GNP	PG-N3B-PB-O1B
4	A	201	GNP	PA-O3A-PB-O1B
4	A	201	GNP	PA-O3A-PB-O2B
4	A	201	GNP	C5'-O5'-PA-O3A
4	A	201	GNP	C5'-O5'-PA-O1A

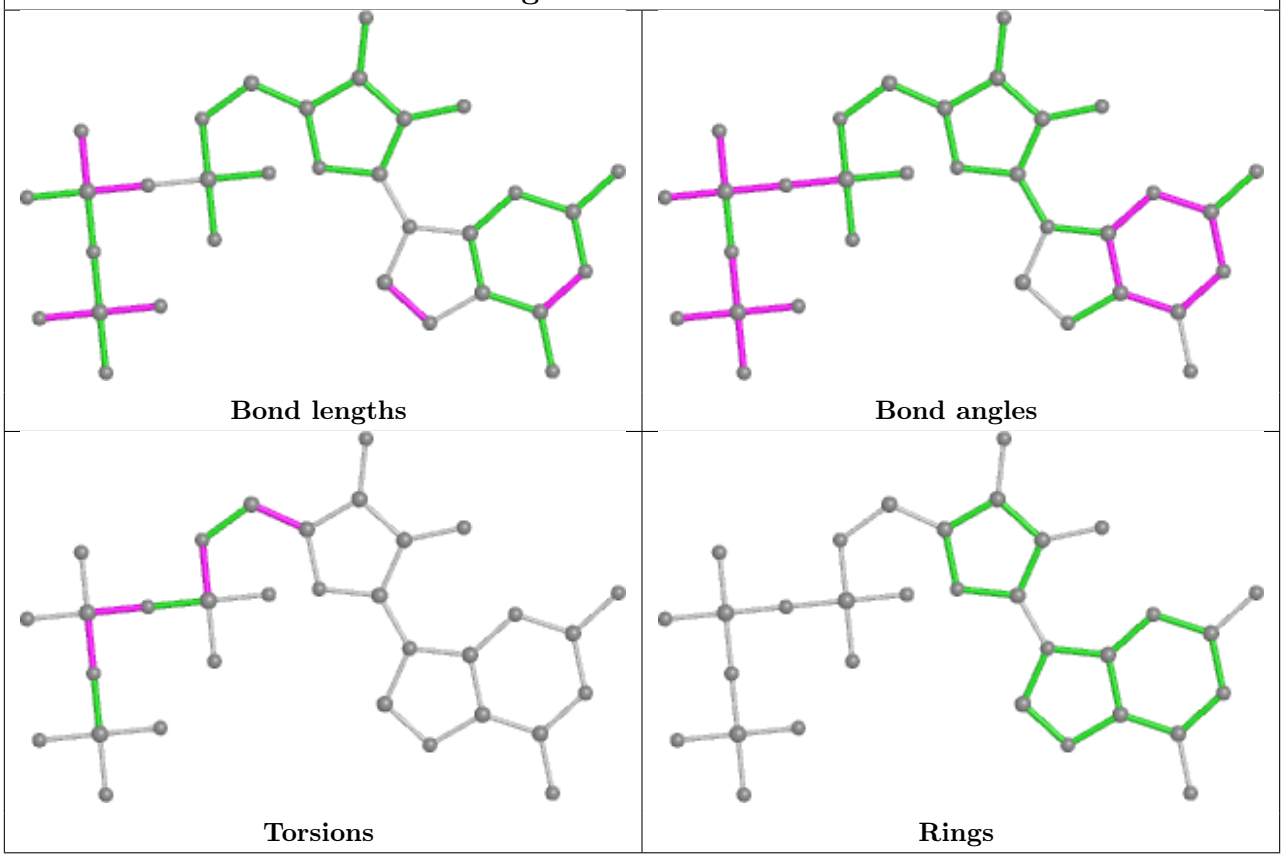
There are no ring outliers.

2 monomers are involved in 2 short contacts:

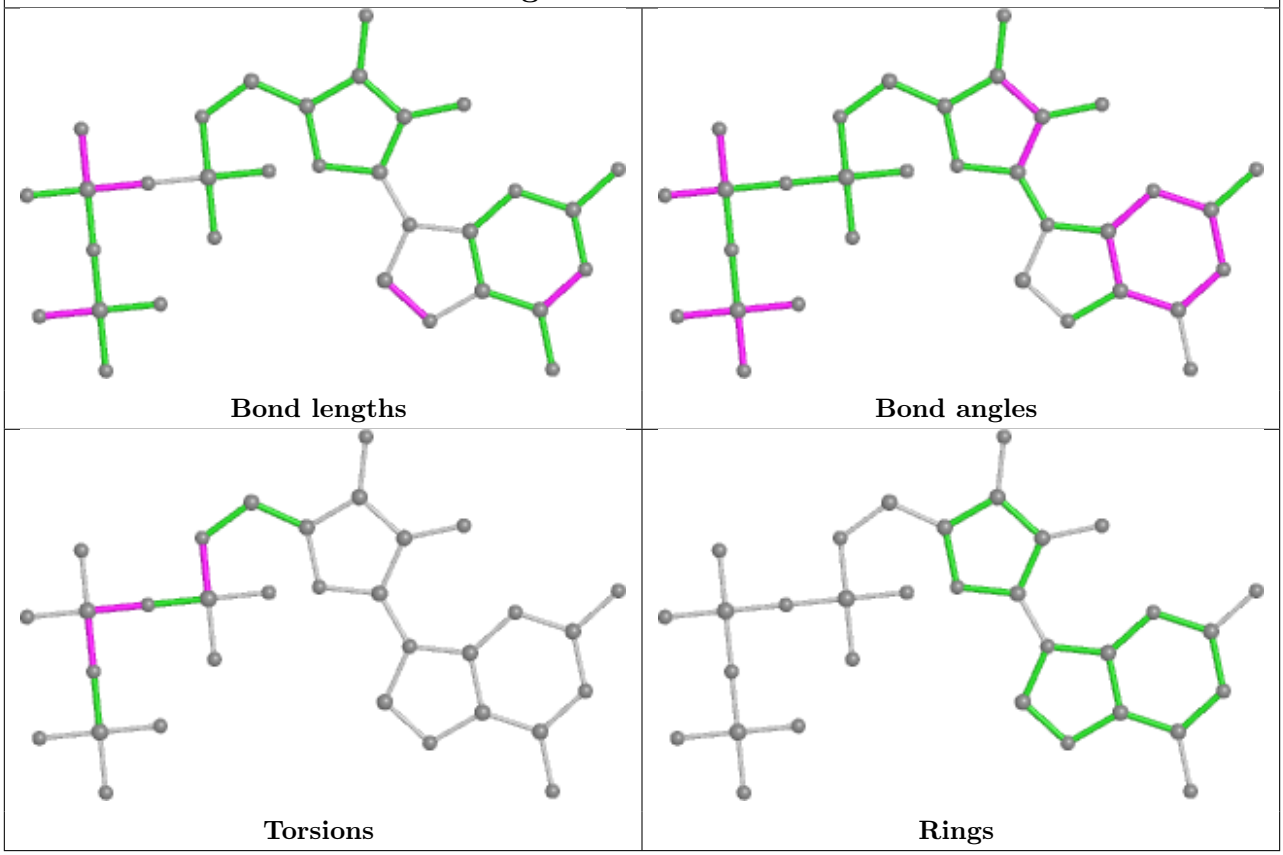
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	201	GNP	1	0
4	D	201	GNP	1	0

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 GNP A 201



Ligand GNP D 201



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	159/164 (96%)	-0.26	0 100 100	35, 51, 75, 96	0
1	D	159/164 (96%)	0.03	2 (1%) 77 77	37, 66, 95, 123	0
2	B	264/355 (74%)	0.06	9 (3%) 45 40	41, 67, 106, 137	0
2	E	260/355 (73%)	0.10	11 (4%) 36 32	45, 76, 120, 136	0
3	C	140/153 (91%)	-0.19	0 100 100	36, 56, 87, 110	0
3	F	140/153 (91%)	0.71	15 (10%) 6 4	44, 98, 136, 154	0
All	All	1122/1344 (83%)	0.07	37 (3%) 46 41	35, 67, 116, 154	0

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	260	ARG	5.9
3	F	50	ASN	5.4
2	B	300	LYS	4.7
3	F	57	SER	3.9
2	B	223	LEU	3.6

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)

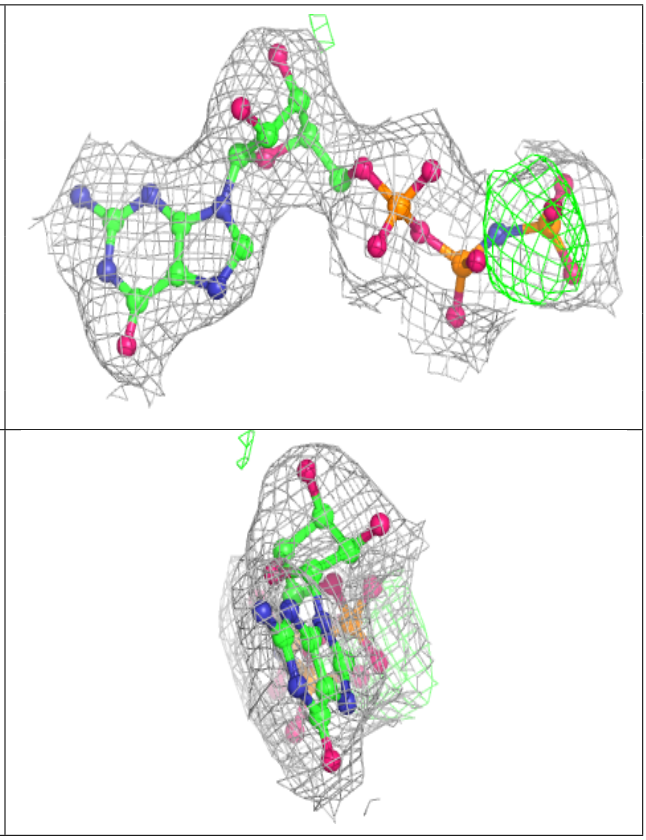
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	GNP	A	201	32/32	0.97	0.18	30,49,68,70	0
4	GNP	D	201	32/32	0.97	0.17	32,53,63,71	0
5	MG	D	202	1/1	0.97	0.30	51,51,51,51	0
5	MG	A	202	1/1	0.98	0.41	47,47,47,47	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.

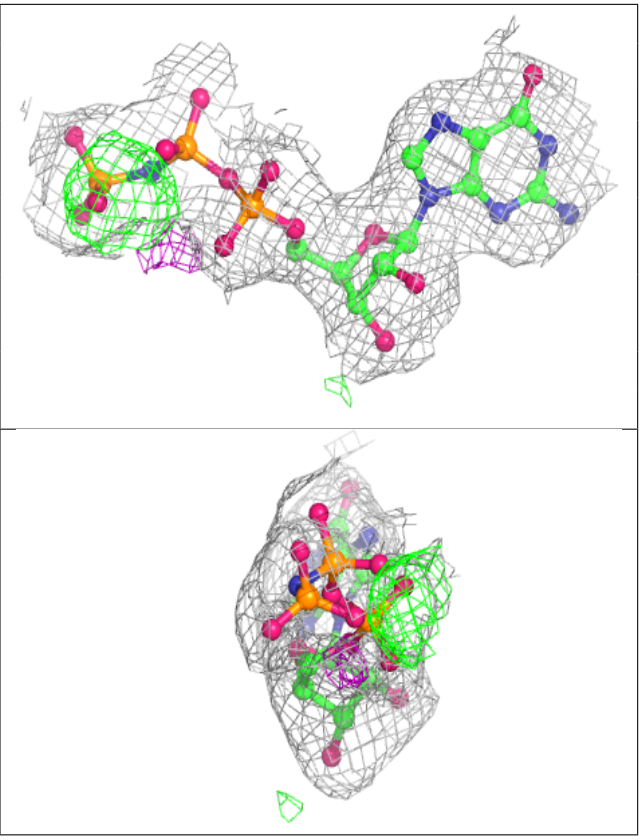
Electron density around GNP A 201:

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 GNP D 201:

$2mF_o$-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.