wwPDB X-ray Structure Validation Summary Report (i)

Mar 8, 2026 – 05:42 AM UTC

PDB ID	:	7Y4I / pdb_00007y4i
Title	:	Crystal structure of SPINDLY in complex with GDP
Authors	:	Xu, S.T.; Wan, L.H.
Deposited on	:	2022-06-14
Resolution	:	2.85 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:

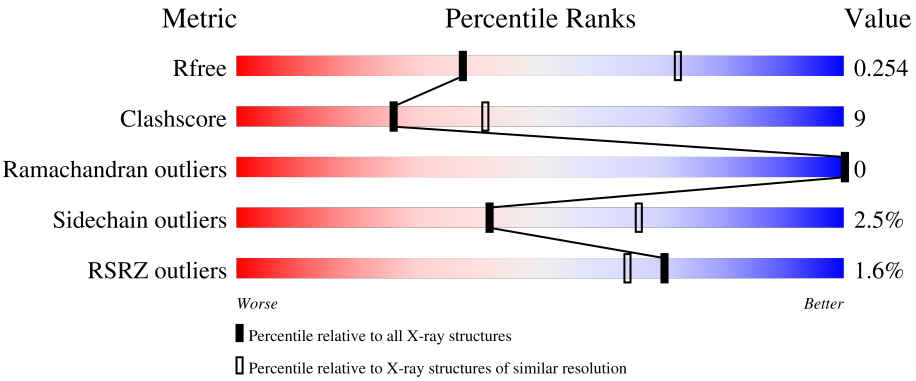
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal 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 (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

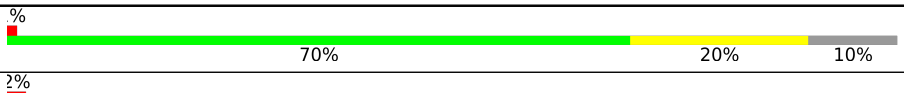
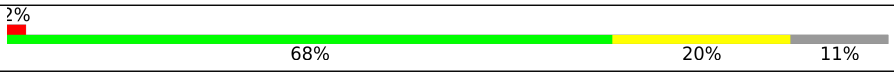
The reported resolution of this entry is 2.85 Å.

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	1407 (2.88-2.84)
Clashscore	190562	1446 (2.88-2.84)
Ramachandran outliers	187476	1406 (2.88-2.84)
Sidechain outliers	187428	1407 (2.88-2.84)
RSRZ outliers	180081	1408 (2.88-2.84)

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	914	
1	B	914	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 12812 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.

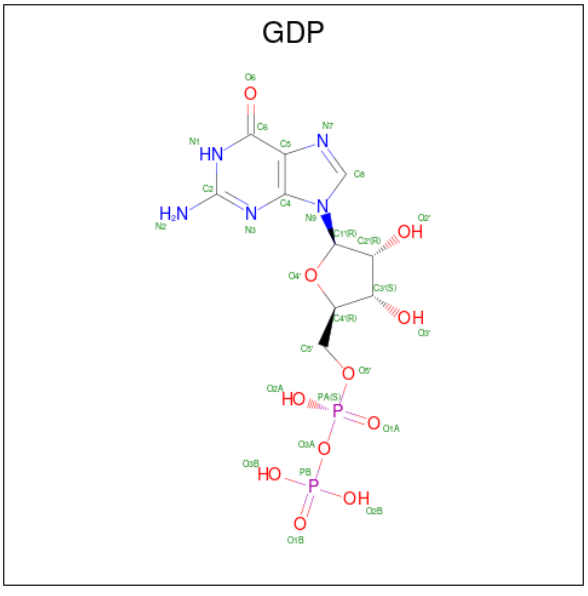
- Molecule 1 is a protein called Probable UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase SPINDLY.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	822	Total	C	N	O	S	0	0	0
			6419	4083	1076	1211	49			
1	B	815	Total	C	N	O	S	0	0	0
			6360	4048	1064	1199	49			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	645	SER	CYS	engineered mutation	UNP Q96301
B	645	SER	CYS	engineered mutation	UNP Q96301

- Molecule 2 is GUANOSINE-5'-DIPHOSPHATE (CCD ID: GDP) (formula: $C_{10}H_{15}N_5O_{11}P_2$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			28	10	5	11	2		

- Molecule 3 is water.

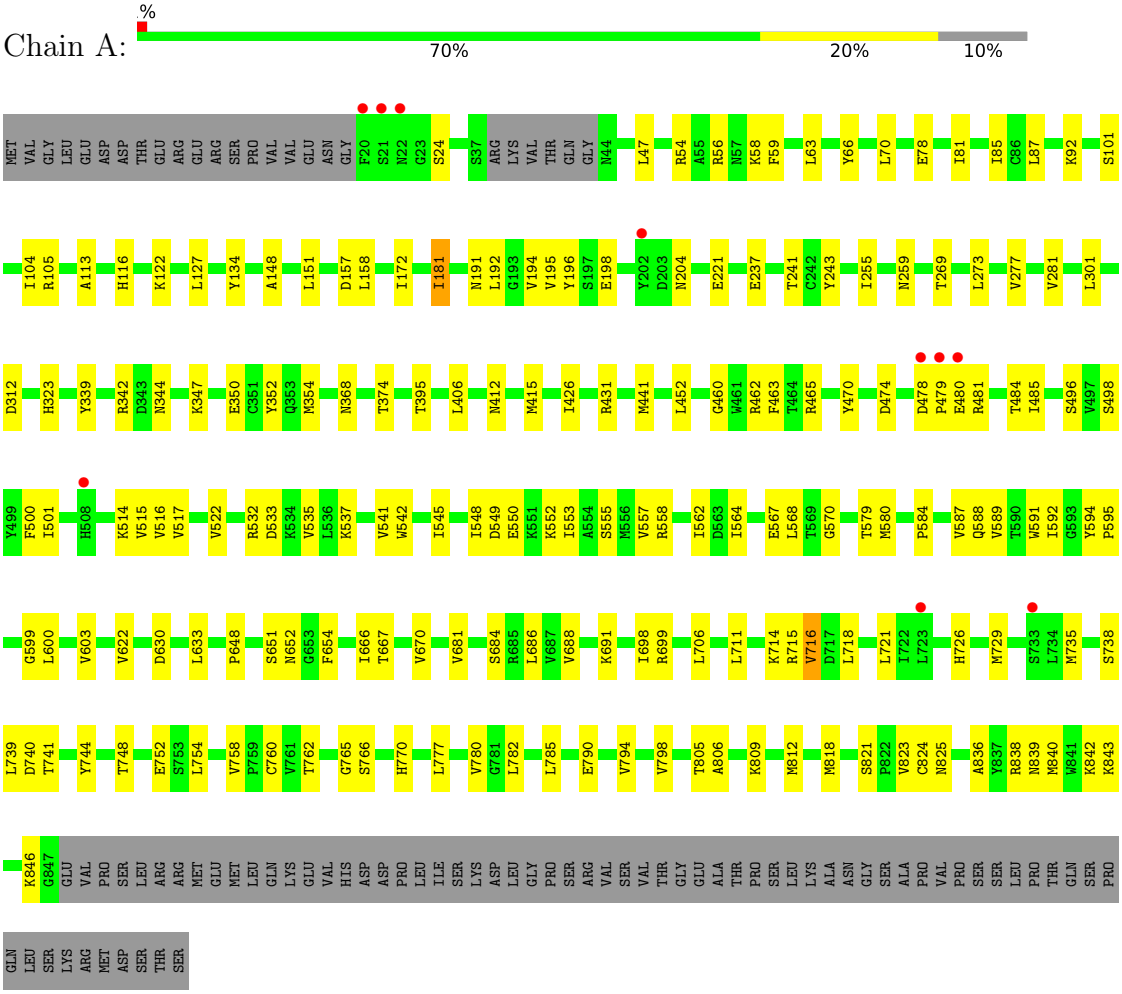
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	3	Total	O	0	0
			3	3		
3	B	2	Total	O	0	0
			2	2		

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.

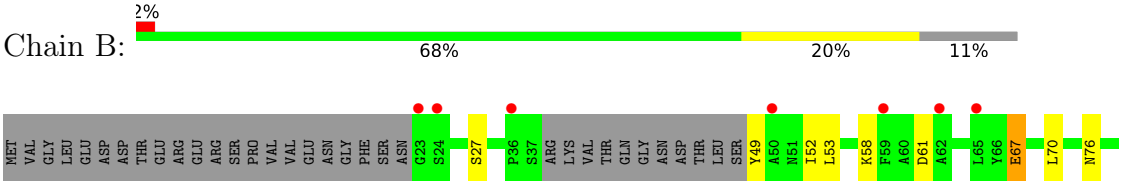
• Molecule 1: Probable UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase SPINDLY

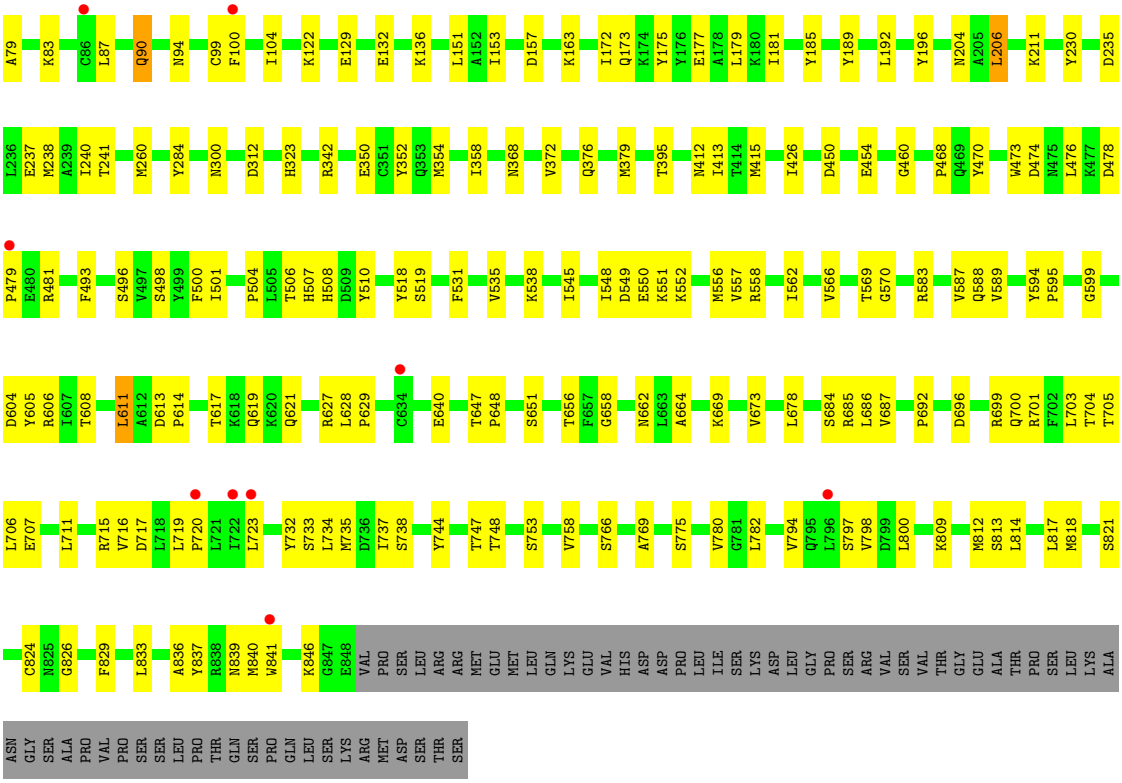
Chain A:



• Molecule 1: Probable UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase SPINDLY

Chain B:




A79 K83 C86 L87 Q90 N94 C99 F100 I104 K122 E129 E132 K136 L151 A152 I153 D157 K163 I172 Q173 K174 Y175 Y176 E177 A178 L179 K180 I181 Y185 Y189 L192 Y196 N204 A205 L206 K211 Y230 D235
L236 E237 M238 A239 I240 T241 M260 Y284 N300 D312 H323 R342 E350 C351 Y352 Q353 M354 I358 N368 V372 Q376 M379 T395 N412 I413 T414 M415 I426 D450 E454 G460 P468 Q469 Y470 W473 D474 N475 L476 K477 D478
P479 E480 R481 F493 S496 V497 S498 Y499 F500 I501 P504 L505 T506 H507 H508 D509 Y510 Y518 S519 F531 V535 K538 I545 I548 D549 E550 K551 K552 M556 V557 R558 I562 V566 T569 G570 R583 V587 Q588 V589 Y594 P595 G599
D604 Y605 R606 I607 T608 L611 A612 D613 P614 T617 K618 Q619 K620 Q621 R627 L628 P629 C634 E640 T647 P648 S651 T656 F657 G658 N662 L663 A664 K669 V673 L678 S684 R685 L686 V687 P692 D696 R699 Q700 R701 F702 L703 T704 T705
L706 E707 L711 R715 V716 D717 L718 L719 P720 L721 I722 L723 Y732 S733 L734 M735 D736 I737 S738 Y744 T747 T748 S753 V758 S766 A769 S775 V780 G781 L782 V794 Q795 L796 S797 V798 D799 L800 K809 M812 S813 L814 L817 M818 S821
C824 N825 G826 F829 L833 A836 Y837 R838 N839 M840 W841 K846 G847 E848 VAL PRO SER LEU ARG ARG MET GLU MET LEU GLN LYS GLU VAL HIS ASP ASP PRO LEU ILE SER LYS ASP LEU GLY PRO SER ARG VAL SER VAL THR GLY GLU ALA THR PRO SER LEU LYS ALA
ASN GLY SER ALA PRO VAL PRO SER SER LEU PRO THR GLN SER PRO GLN LEU SER LYS ARG MET ASP SER THR SER

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	127.73Å 130.68Å 141.37Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.34 – 2.85 35.34 – 2.85	Depositor EDS
% Data completeness (in resolution range)	99.6 (35.34-2.85) 99.6 (35.34-2.85)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.43 (at 2.85Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487, PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.213 , 0.270 (Not available) , 0.254	Depositor DCC
R_{free} test set	2930 reflections (4.46%)	wwPDB-VP
Wilson B-factor (Å^2)	65.6	Xtriage
Anisotropy	0.302	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 78.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	12812	wwPDB-VP
Average B, all atoms (Å^2)	86.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 40.21 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.8287e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GDP

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	0.15	0/6562	0.33	0/8905
1	B	0.14	0/6502	0.33	0/8824
All	All	0.15	0/13064	0.33	0/17729

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

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	6419	0	6336	120	0
1	B	6360	0	6278	130	0
2	A	28	0	12	0	0
3	A	3	0	0	0	0
3	B	2	0	0	0	0
All	All	12812	0	12626	241	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 9.

The worst 5 of 241 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:549:ASP:H	1:B:552:LYS:HE3	1.43	0.83
1:A:545:ILE:HA	1:A:548:ILE:HD12	1.64	0.78
1:B:678:LEU:HD11	1:B:686:LEU:HB2	1.69	0.75
1:A:479:PRO:O	1:A:838:ARG:NH1	2.20	0.74
1:A:780:VAL:HG12	1:A:782:LEU:HD23	1.70	0.74

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	818/914 (90%)	813 (99%)	5 (1%)	0	100	100
1	B	811/914 (89%)	807 (100%)	4 (0%)	0	100	100
All	All	1629/1828 (89%)	1620 (99%)	9 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	691/774 (89%)	675 (98%)	16 (2%)	44	68
1	B	683/774 (88%)	664 (97%)	19 (3%)	38	62
All	All	1374/1548 (89%)	1339 (98%)	35 (2%)	42	66

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	647	THR
1	B	651	SER
1	B	747	THR
1	A	711	LEU
1	A	698	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	279	GLN
1	B	588	GLN
1	B	839	ASN
1	B	652	ASN
1	B	467	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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

1 ligand is 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	$\#\|Z\|>2$
2	GDP	A	1901	-	29,30,30	1.17	3 (10%)	45,47,47	1.77	7 (15%)

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	GDP	A	1901	-	-	3/16/32/32	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1901	GDP	C5-C4	3.15	1.47	1.38
2	A	1901	GDP	C6-N1	-2.37	1.34	1.38
2	A	1901	GDP	C5-N7	-2.04	1.35	1.39

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1901	GDP	C5-C4-N3	-5.96	118.90	128.39
2	A	1901	GDP	C2-N3-C4	4.95	120.82	112.30
2	A	1901	GDP	N9-C4-N3	4.49	134.94	125.95
2	A	1901	GDP	C6-C5-N7	3.24	136.18	130.29
2	A	1901	GDP	C4-C5-N7	-2.46	106.78	110.67

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1901	GDP	PA-O3A-PB-O2B
2	A	1901	GDP	C3'-C4'-C5'-O5'
2	A	1901	GDP	PA-O3A-PB-O1B

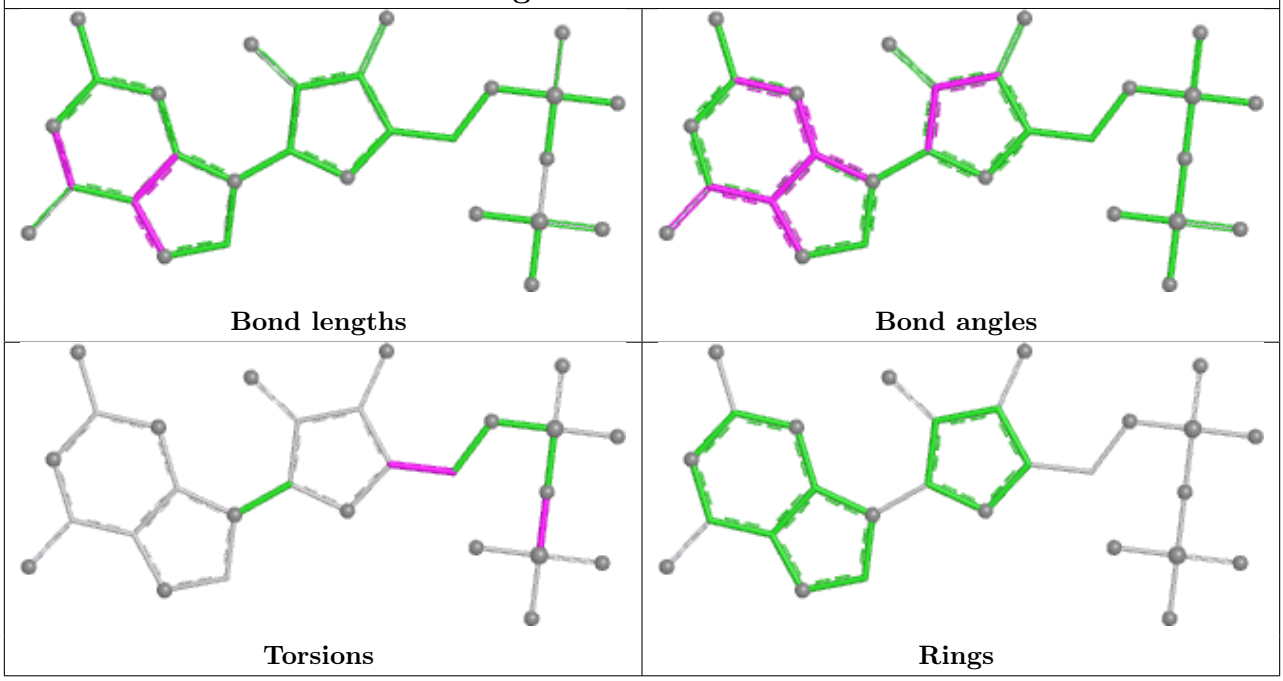
There are no ring outliers.

No monomer is involved in short contacts.

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 GDP A 1901



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

6.1 Protein, DNA and RNA chains

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>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	822/914 (89%)	-0.02	10 (1%)	76	71	40, 84, 133, 168	0
1	B	815/914 (89%)	-0.11	16 (1%)	65	57	37, 83, 134, 178	0
All	All	1637/1828 (89%)	-0.06	26 (1%)	70	64	37, 84, 134, 178	0

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	21	SER	4.9
1	A	478	ASP	4.4
1	A	20	PHE	3.5
1	A	480	GLU	3.3
1	A	508	HIS	2.9

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no oligosaccharides in this entry.

6.4 Ligands

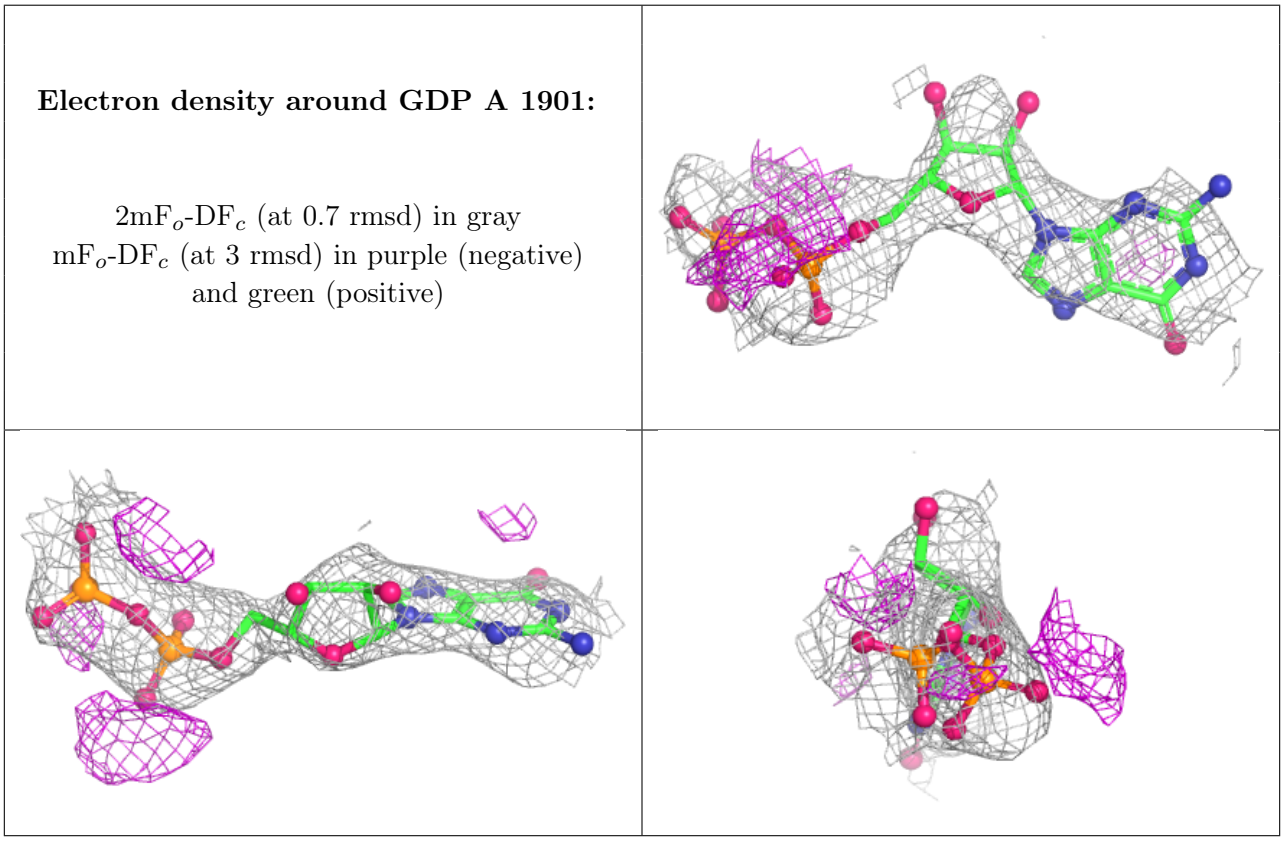
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	GDP	A	1901	28/28	0.71	0.13	161,168,183,190	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 GDP A 1901:

$2mF_o$-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.