



# Full wwPDB X-ray Structure Validation Report ⓘ

May 14, 2020 – 10:50 pm BST

PDB ID : 4EEC  
Title : Crystal Structure of the glycopeptide antibiotic sulfotransferase StaL complexed with A3P and desulfo-A47934.  
Authors : Shi, R.; Cygler, M.  
Deposited on : 2012-03-28  
Resolution : 2.70 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 following versions of software and data (see [references ⓘ](#)) 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	:	2.11
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)	:	2.11

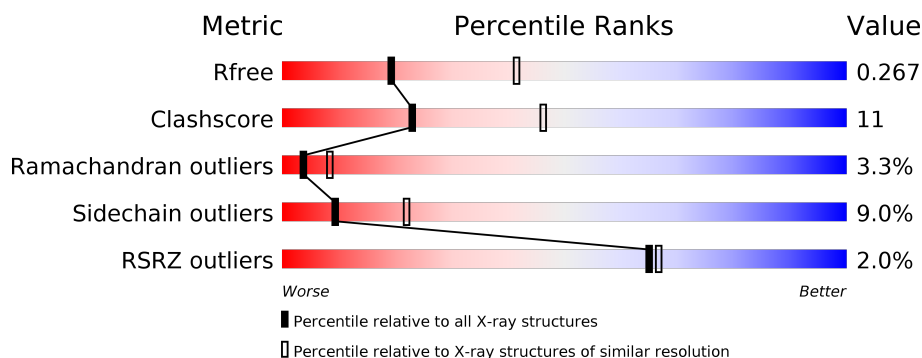
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 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	286	<div> <div>%</div> <div> <div></div> <div>65%</div> <div>18%</div> <div>5%</div> <div>12%</div> </div> </div>
1	B	286	<div> <div>2%</div> <div> <div></div> <div>65%</div> <div>15%</div> <div>•</div> <div>15%</div> </div> </div>
2	C	7	<div> <div> <div></div> <div>14%</div> <div>57%</div> <div>29%</div> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3899 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 StaL.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	253	Total	C	N	O	S	0	0	0
			1937	1228	333	363	13			
1	B	242	Total	C	N	O	S	0	0	0
			1839	1162	315	350	12			

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-15	MET	-	INITIATING METHIONINE	UNP Q8KLM3
A	-14	GLY	-	EXPRESSION TAG	UNP Q8KLM3
A	-13	SER	-	EXPRESSION TAG	UNP Q8KLM3
A	-12	SER	-	EXPRESSION TAG	UNP Q8KLM3
A	-11	HIS	-	EXPRESSION TAG	UNP Q8KLM3
A	-10	HIS	-	EXPRESSION TAG	UNP Q8KLM3
A	-9	HIS	-	EXPRESSION TAG	UNP Q8KLM3
A	-8	HIS	-	EXPRESSION TAG	UNP Q8KLM3
A	-7	HIS	-	EXPRESSION TAG	UNP Q8KLM3
A	-6	HIS	-	EXPRESSION TAG	UNP Q8KLM3
A	-5	SER	-	EXPRESSION TAG	UNP Q8KLM3
A	-4	SER	-	EXPRESSION TAG	UNP Q8KLM3
A	-3	GLY	-	EXPRESSION TAG	UNP Q8KLM3
A	-2	LEU	-	EXPRESSION TAG	UNP Q8KLM3
A	-1	VAL	-	EXPRESSION TAG	UNP Q8KLM3
A	0	PRO	-	EXPRESSION TAG	UNP Q8KLM3
A	1	ARG	-	EXPRESSION TAG	UNP Q8KLM3
A	2	GLY	-	EXPRESSION TAG	UNP Q8KLM3
A	3	SER	-	EXPRESSION TAG	UNP Q8KLM3
B	-15	MET	-	INITIATING METHIONINE	UNP Q8KLM3
B	-14	GLY	-	EXPRESSION TAG	UNP Q8KLM3
B	-13	SER	-	EXPRESSION TAG	UNP Q8KLM3
B	-12	SER	-	EXPRESSION TAG	UNP Q8KLM3
B	-11	HIS	-	EXPRESSION TAG	UNP Q8KLM3
B	-10	HIS	-	EXPRESSION TAG	UNP Q8KLM3

*Continued on next page...*

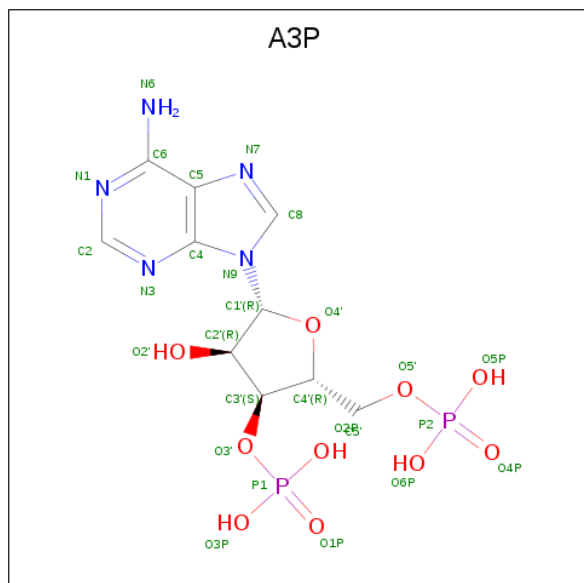
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-9	HIS	-	EXPRESSION TAG	UNP Q8KLM3
B	-8	HIS	-	EXPRESSION TAG	UNP Q8KLM3
B	-7	HIS	-	EXPRESSION TAG	UNP Q8KLM3
B	-6	HIS	-	EXPRESSION TAG	UNP Q8KLM3
B	-5	SER	-	EXPRESSION TAG	UNP Q8KLM3
B	-4	SER	-	EXPRESSION TAG	UNP Q8KLM3
B	-3	GLY	-	EXPRESSION TAG	UNP Q8KLM3
B	-2	LEU	-	EXPRESSION TAG	UNP Q8KLM3
B	-1	VAL	-	EXPRESSION TAG	UNP Q8KLM3
B	0	PRO	-	EXPRESSION TAG	UNP Q8KLM3
B	1	ARG	-	EXPRESSION TAG	UNP Q8KLM3
B	2	GLY	-	EXPRESSION TAG	UNP Q8KLM3
B	3	SER	-	EXPRESSION TAG	UNP Q8KLM3

- Molecule 2 is a protein called desulfo-A47934.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	7	Total	C	Cl	N	O	0	0	0
			86	58	3	7	18			

- Molecule 3 is ADENOSINE-3'-5'-DIPHOSPHATE (three-letter code: A3P) (formula:  $C_{10}H_{15}N_5O_{10}P_2$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			27	10	5	10	2		

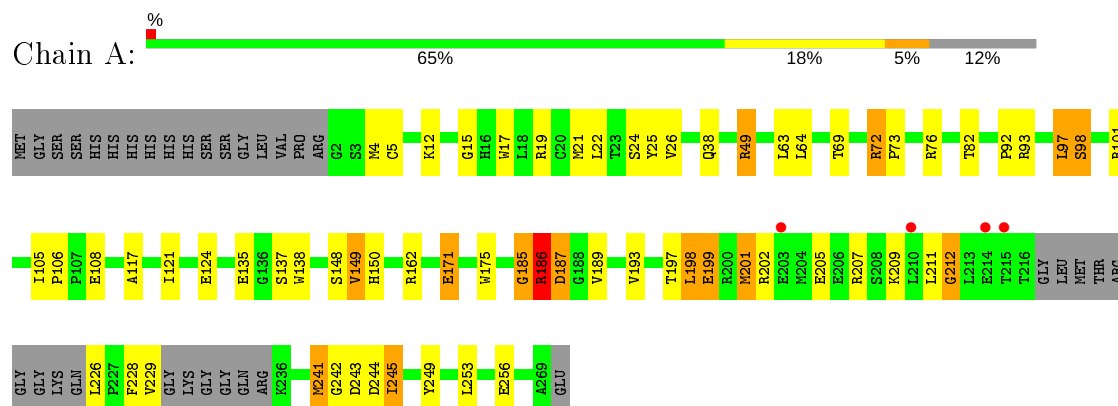
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	3	Total 3	O 3	0	0
4	B	7	Total 7	O 7	0	0

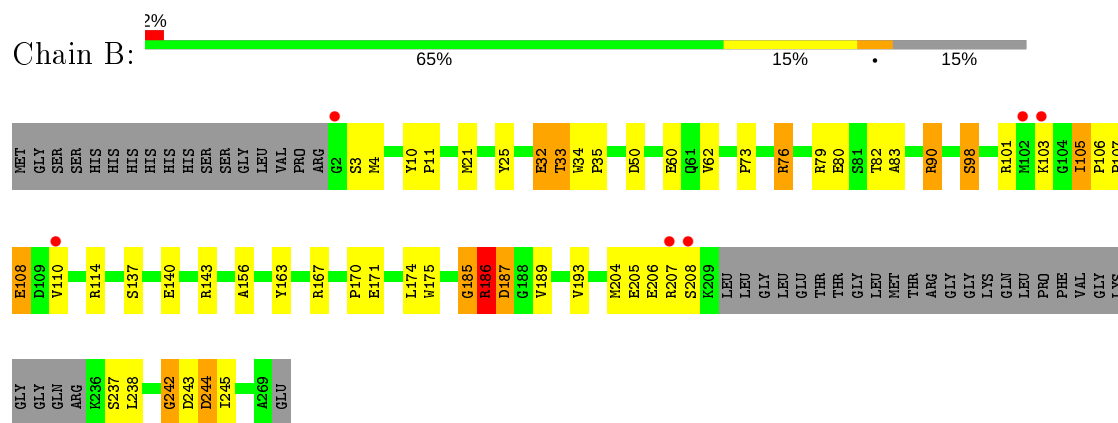
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

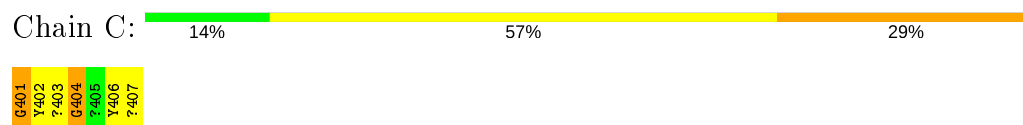
#### • Molecule 1: StaL



#### • Molecule 1: StaL



#### • Molecule 2: desulfo-A47934



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	53.44Å 82.58Å 123.08Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.70 44.87 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.4 (50.00-2.70) 99.4 (44.87-2.70)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.10 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.223 , 0.272 0.215 , 0.267	Depositor DCC
$R_{free}$ test set	778 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	56.5	Xtriage
Anisotropy	0.978	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 31.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	3899	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.24% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: A3P, GHP, 3MY, 0UZ, 3FG, OMY

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	0.80	3/1985 (0.2%)	0.82	0/2701
1	B	0.71	0/1885	0.77	0/2566
All	All	0.76	3/3870 (0.1%)	0.80	0/5267

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
2	C	0	4
All	All	0	5

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	209	LYS	CE-NZ	7.28	1.67	1.49
1	A	256	GLU	CG-CD	5.63	1.60	1.51
1	A	256	GLU	CB-CG	5.30	1.62	1.52

There are no bond angle outliers.

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	241	MET	Peptide
2	C	401	GHP	Mainchain
2	C	404	GHP	Peptide

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Group
2	C	406	OMY	Mainchain,Peptide

## 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	1937	0	1823	47	0
1	B	1839	0	1713	36	0
2	C	86	0	37	5	0
3	A	27	0	11	1	0
4	A	3	0	0	0	0
4	B	7	0	0	1	0
All	All	3899	0	3584	85	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 11.

All (85) 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:242:GLY:HA2	1:A:245:ILE:HG13	1.53	0.90
1:B:33:THR:HG23	1:B:34:TRP:H	1.35	0.88
1:A:73:PRO:HA	1:A:76:ARG:HD3	1.60	0.83
1:B:73:PRO:O	1:B:76:ARG:HB2	1.79	0.83
1:B:33:THR:CG2	1:B:34:TRP:H	1.92	0.83
1:A:205:GLU:OE1	1:A:229:VAL:HG22	1.82	0.80
1:B:185:GLY:O	1:B:186:ARG:HB2	1.85	0.77
1:B:21:MET:CE	1:B:174:LEU:HD22	2.14	0.77
1:B:114:ARG:HD3	1:B:245:ILE:HD11	1.67	0.74
1:B:186:ARG:HH11	1:B:186:ARG:HG3	1.53	0.72
1:A:49:ARG:HG3	2:C:402:3MY:CD2	2.21	0.71
1:B:33:THR:CG2	1:B:34:TRP:N	2.52	0.70
1:A:197:THR:HG23	1:A:198:LEU:HG	1.74	0.69
1:A:4:MET:SD	1:A:64:LEU:HD11	2.32	0.69
1:B:32:GLU:O	1:B:33:THR:HB	1.94	0.67
1:B:21:MET:HE2	1:B:174:LEU:HD22	1.76	0.66

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:197:THR:CG2	1:A:198:LEU:HG	2.26	0.65
1:A:73:PRO:HA	1:A:76:ARG:CD	2.25	0.65
1:B:242:GLY:O	1:B:244:ASP:N	2.30	0.65
1:B:21:MET:HE3	1:B:174:LEU:HD22	1.80	0.63
1:B:34:TRP:HB3	1:B:35:PRO:CD	2.30	0.62
1:A:242:GLY:CA	1:A:245:ILE:HG13	2.28	0.62
1:B:170:PRO:HB2	1:B:193:VAL:HG11	1.81	0.61
1:A:21:MET:CE	1:A:193:VAL:HG22	2.32	0.59
1:A:5:CYS:HB2	1:A:63:LEU:HD23	1.83	0.58
2:C:407:3FG:OXT	2:C:407:3FG:HG2	2.02	0.58
1:A:72:ARG:NH1	1:B:50:ASP:O	2.35	0.58
1:A:171:GLU:OE2	1:A:186:ARG:NH2	2.36	0.57
1:A:69:THR:OG1	1:A:72:ARG:HG2	2.05	0.57
1:B:90:ARG:HB2	1:B:163:TYR:HB3	1.87	0.57
1:A:228:PHE:C	1:A:228:PHE:CD2	2.78	0.56
1:A:197:THR:HG22	1:A:199:GLU:H	1.70	0.56
1:B:175:TRP:CD2	1:B:186:ARG:HG2	2.40	0.56
1:A:4:MET:SD	1:A:64:LEU:CD1	2.94	0.56
1:B:60:GLU:O	1:B:62:VAL:HG23	2.07	0.55
1:B:189:VAL:O	1:B:193:VAL:HG23	2.06	0.55
1:A:98:SER:HA	1:A:101:ARG:NH1	2.22	0.55
1:A:49:ARG:HG3	2:C:402:3MY:CE2	2.38	0.54
1:A:187:ASP:N	1:A:187:ASP:OD1	2.35	0.53
1:A:175:TRP:CE3	1:A:186:ARG:HG3	2.43	0.53
1:A:17:TRP:HA	1:A:201:MET:HE1	1.92	0.52
1:B:79:ARG:HD2	4:B:301:HOH:O	2.09	0.52
1:B:98:SER:HA	1:B:101:ARG:NH1	2.25	0.52
1:A:228:PHE:O	1:A:228:PHE:CD2	2.63	0.51
1:A:124:GLU:O	1:A:137:SER:HB2	2.10	0.51
1:A:25:TYR:OH	1:A:185:GLY:N	2.44	0.50
1:A:73:PRO:O	1:A:76:ARG:HB2	2.11	0.50
1:B:83:ALA:O	1:B:156:ALA:HA	2.10	0.50
1:A:117:ALA:O	1:A:121:ILE:HG13	2.12	0.49
1:B:32:GLU:O	1:B:33:THR:CB	2.60	0.49
1:A:22:LEU:O	1:A:26:VAL:HG23	2.13	0.49
1:B:186:ARG:HH11	1:B:186:ARG:CG	2.24	0.48
1:A:211:LEU:O	1:A:212:GLY:C	2.53	0.47
1:B:187:ASP:OD1	1:B:187:ASP:N	2.45	0.47
1:B:33:THR:HA	1:B:204:MET:HG2	1.96	0.46
1:A:106:PRO:HB2	1:A:108:GLU:OE2	2.15	0.46
1:A:175:TRP:CD2	1:A:186:ARG:HG3	2.49	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:402:3MY:CE1	2:C:404:GHP:C2	2.83	0.46
1:B:10:TYR:CG	1:B:11:PRO:HD2	2.50	0.46
1:A:148:SER:O	1:A:150:HIS:N	2.48	0.46
1:B:4:MET:HB3	1:B:82:THR:HA	1.98	0.45
1:A:197:THR:HG22	1:A:199:GLU:N	2.31	0.45
1:B:25:TYR:OH	1:B:185:GLY:N	2.49	0.45
1:A:241:MET:HG3	1:A:245:ILE:HD12	1.99	0.45
1:B:33:THR:HG22	1:B:34:TRP:N	2.30	0.45
1:A:148:SER:O	1:A:149:VAL:C	2.55	0.45
1:B:34:TRP:HB3	1:B:35:PRO:HD3	1.99	0.45
1:A:249:TYR:HE2	1:A:253:LEU:HD22	1.82	0.44
1:A:93:ARG:O	1:A:97:LEU:HD22	2.17	0.44
2:C:401:GHP:C2	2:C:403:3FG:CG1	2.91	0.44
1:B:33:THR:HG23	1:B:204:MET:O	2.18	0.44
1:B:106:PRO:O	1:B:108:GLU:N	2.51	0.43
1:A:202:ARG:HA	1:A:229:VAL:HG21	2.00	0.43
1:B:137:SER:OG	1:B:140:GLU:HB2	2.18	0.43
1:B:105:ILE:H	1:B:105:ILE:HG13	1.59	0.42
1:A:15:GLY:O	1:A:19:ARG:HG3	2.19	0.42
1:A:72:ARG:HA	1:A:72:ARG:HE	1.83	0.42
1:B:238:LEU:HD23	1:B:238:LEU:HA	1.75	0.42
1:A:186:ARG:HA	1:A:189:VAL:HG23	2.02	0.42
1:A:92:PRO:HG3	1:A:138:TRP:CE2	2.55	0.41
1:A:202:ARG:HG3	1:A:229:VAL:CG2	2.50	0.41
1:A:4:MET:HB3	1:A:82:THR:HA	2.01	0.41
1:A:73:PRO:HA	1:A:76:ARG:NE	2.36	0.41
1:A:12:LYS:HD2	3:A:301:A3P:O6P	2.21	0.40
1:A:38:GLN:OE1	1:A:38:GLN:HA	2.21	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	247/286 (86%)	226 (92%)	15 (6%)	6 (2%)	6	15
1	B	238/286 (83%)	203 (85%)	25 (10%)	10 (4%)	3	5
All	All	485/572 (85%)	429 (88%)	40 (8%)	16 (3%)	4	8

All (16) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	186	ARG
1	B	33	THR
1	B	110	VAL
1	B	243	ASP
1	A	105	ILE
1	A	149	VAL
1	A	185	GLY
1	A	212	GLY
1	B	186	ARG
1	A	243	ASP
1	B	185	GLY
1	B	207	ARG
1	B	103	LYS
1	B	107	PRO
1	B	105	ILE
1	B	242	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	195/234 (83%)	178 (91%)	17 (9%)	10	23
1	B	183/234 (78%)	166 (91%)	17 (9%)	9	21
All	All	378/468 (81%)	344 (91%)	34 (9%)	9	22

All (34) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	24	SER
1	A	49	ARG
1	A	72	ARG
1	A	97	LEU
1	A	98	SER
1	A	135	GLU
1	A	162	ARG
1	A	171	GLU
1	A	186	ARG
1	A	187	ASP
1	A	198	LEU
1	A	199	GLU
1	A	201	MET
1	A	207	ARG
1	A	226	LEU
1	A	244	ASP
1	A	245	ILE
1	B	3	SER
1	B	32	GLU
1	B	76	ARG
1	B	80	GLU
1	B	90	ARG
1	B	98	SER
1	B	108	GLU
1	B	143	ARG
1	B	167	ARG
1	B	171	GLU
1	B	186	ARG
1	B	187	ASP
1	B	205	GLU
1	B	206	GLU
1	B	208	SER
1	B	237	SER
1	B	244	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	150	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

7 non-standard protein/DNA/RNA residues 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
2	GHP	C	404	2	10,11,12	1.35	1 (10%)	11,14,16	0.77	0
2	GHP	C	401	2	10,11,12	1.04	1 (10%)	11,14,16	1.33	2 (18%)
2	3FG	C	407	2	9,13,13	1.76	4 (44%)	13,18,18	1.38	2 (15%)
2	3FG	C	403	2	11,12,13	1.34	1 (9%)	13,16,18	0.67	0
2	3MY	C	402	2	12,13,14	1.37	1 (8%)	14,17,19	1.26	2 (14%)
2	0UZ	C	405	2	11,12,13	1.70	4 (36%)	13,16,18	1.05	0
2	OMY	C	406	2	12,14,15	3.51	2 (16%)	17,19,21	1.88	4 (23%)

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	GHP	C	404	2	-	2/4/6/8	0/1/1/1
2	GHP	C	401	2	-	0/4/6/8	0/1/1/1
2	3FG	C	407	2	-	2/4/8/8	0/1/1/1
2	3FG	C	403	2	-	2/4/6/8	0/1/1/1
2	3MY	C	402	2	-	2/5/6/8	0/1/1/1
2	0UZ	C	405	2	-	0/4/6/8	0/1/1/1
2	OMY	C	406	2	-	1/9/10/12	0/1/1/1

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	406	OMY	CZ-CE1	11.16	1.50	1.39
2	C	403	3FG	CB-CA	-3.51	1.48	1.52
2	C	406	OMY	CE1-CL	3.28	1.81	1.73
2	C	402	3MY	CE2-CL	3.22	1.81	1.73
2	C	404	GHP	C1-CA	-3.01	1.49	1.52
2	C	405	0UZ	C1-CA	-2.97	1.49	1.52
2	C	405	0UZ	C3-CL3	2.74	1.80	1.73
2	C	407	3FG	CG1-CB	2.58	1.43	1.39
2	C	407	3FG	CG2-CB	2.43	1.42	1.39
2	C	407	3FG	CG1-CD1	2.21	1.42	1.39
2	C	401	GHP	C1-CA	-2.19	1.50	1.52
2	C	405	0UZ	CA-C	2.05	1.55	1.51
2	C	407	3FG	CG2-CD2	2.03	1.42	1.39
2	C	405	0UZ	C4-C3	2.02	1.41	1.39

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	406	OMY	CG-CB-CA	-5.28	104.42	111.49
2	C	407	3FG	C-CA-CB	3.98	116.42	111.28
2	C	406	OMY	CD1-CE1-CZ	-3.31	119.00	120.91
2	C	401	GHP	C2-C1-CA	-2.98	115.84	120.65
2	C	406	OMY	CD1-CE1-CL	2.61	122.69	118.49
2	C	402	3MY	CB-CA-C	2.53	116.22	111.47
2	C	402	3MY	CG-CB-CA	-2.50	109.03	114.10
2	C	401	GHP	C6-C1-CA	2.42	124.55	120.65
2	C	407	3FG	CG2-CB-CA	2.29	123.35	119.71
2	C	406	OMY	O-C-CA	-2.19	119.03	124.78

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	407	3FG	C-CA-CB-CG2
2	C	407	3FG	C-CA-CB-CG1
2	C	402	3MY	C-CA-CB-CG
2	C	402	3MY	CA-CB-CG-CD2
2	C	406	OMY	O-C-CA-CB
2	C	404	GHP	C2-C1-CA-C
2	C	403	3FG	C-CA-CB-CG1
2	C	403	3FG	C-CA-CB-CG2
2	C	404	GHP	C6-C1-CA-N

There are no ring outliers.

5 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	404	GHP	1	0
2	C	401	GHP	1	0
2	C	407	3FG	1	0
2	C	403	3FG	1	0
2	C	402	3MY	3	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates 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	$\# Z  > 2$
3	A3P	A	301	-	26,29,29	1.18	3 (11%)	31,45,45	1.54	8 (25%)

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
3	A3P	A	301	-	-	1/11/31/31	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	301	A3P	C5-C4	2.95	1.48	1.40
3	A	301	A3P	O4'-C1'	2.75	1.44	1.41

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	301	A3P	C2-N3	2.19	1.35	1.32

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	301	A3P	O4'-C1'-C2'	-4.09	100.94	106.93
3	A	301	A3P	N3-C2-N1	-3.31	123.51	128.68
3	A	301	A3P	C4-C5-N7	-2.63	106.66	109.40
3	A	301	A3P	O6P-P2-O5P	2.27	116.31	107.64
3	A	301	A3P	O6P-P2-O5'	-2.07	101.24	106.73
3	A	301	A3P	C2-N1-C6	2.01	122.19	118.75
3	A	301	A3P	O3P-P1-O2P	2.01	115.31	107.64
3	A	301	A3P	N6-C6-N1	2.00	122.73	118.57

There are no chirality outliers.

All (1) torsion outliers are listed below:

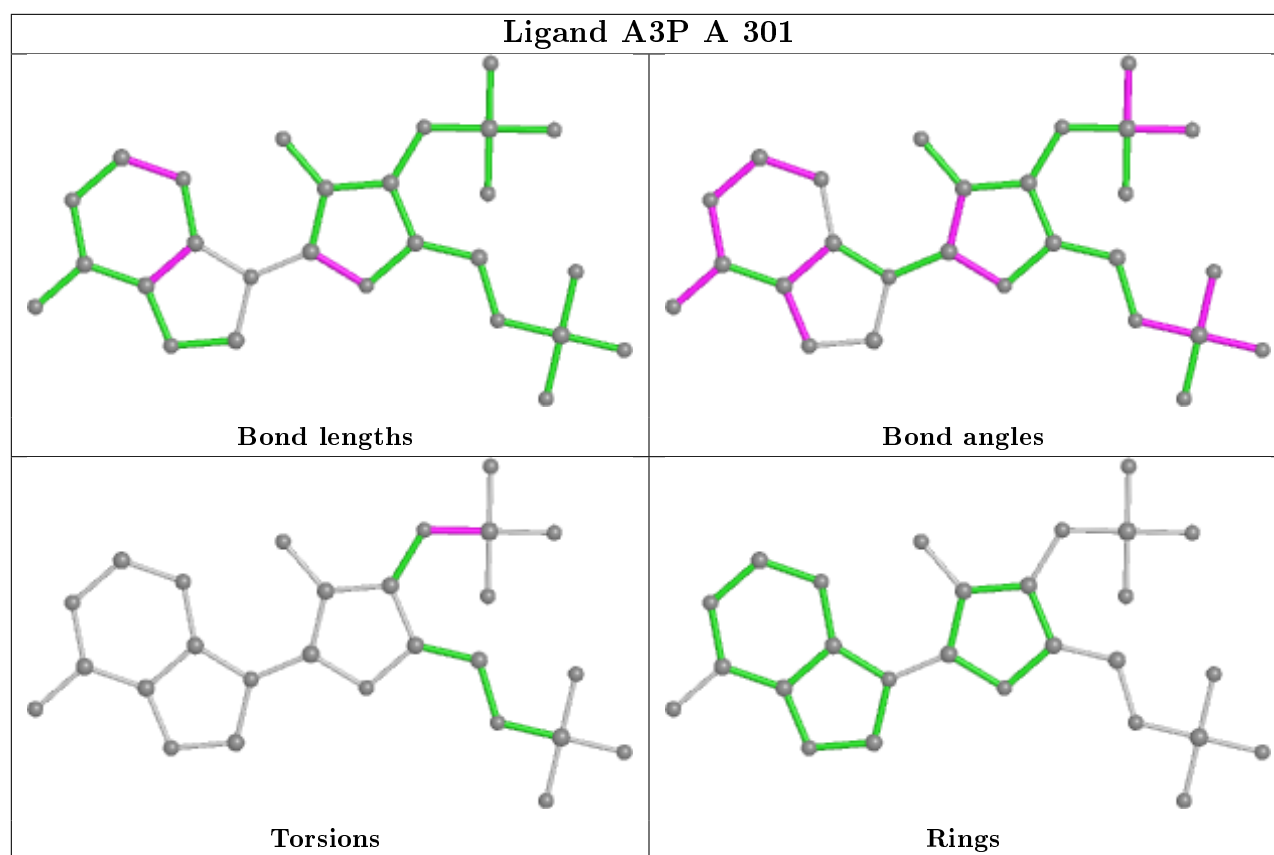
Mol	Chain	Res	Type	Atoms
3	A	301	A3P	C3'-O3'-P1-O1P

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	301	A3P	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 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	253/286 (88%)	-0.14	4 (1%) 72 74	38, 59, 94, 105	0
1	B	242/286 (84%)	-0.05	6 (2%) 57 59	39, 66, 114, 127	0
2	C	0/7	-	-	-	-
All	All	495/579 (85%)	-0.10	10 (2%) 65 67	38, 61, 103, 127	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	215	THR	3.4
1	B	110	VAL	3.2
1	B	2	GLY	3.2
1	A	214	GLU	3.1
1	B	103	LYS	2.8
1	B	208	SER	2.7
1	B	102	MET	2.7
1	B	207	ARG	2.2
1	A	203	GLU	2.0
1	A	210	LEU	2.0

### 6.2 Non-standard residues in protein, DNA, RNA chains [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(Å <sup>2</sup> )	Q<0.9
2	0UZ	C	405	12/13	0.72	0.17	93,97,100,103	0
2	3FG	C	407	13/13	0.73	0.25	104,105,107,108	0
2	3MY	C	402	13/14	0.83	0.26	66,68,72,72	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	OMY	C	406	14/15	0.84	0.28	93,97,101,101	0
2	GHP	C	404	11/12	0.88	0.14	80,82,85,86	0
2	3FG	C	403	12/13	0.91	0.18	70,74,76,80	0
2	GHP	C	401	11/12	0.95	0.22	62,64,65,65	0

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

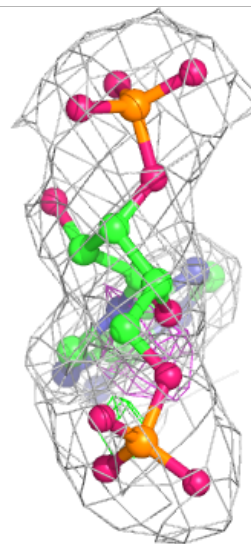
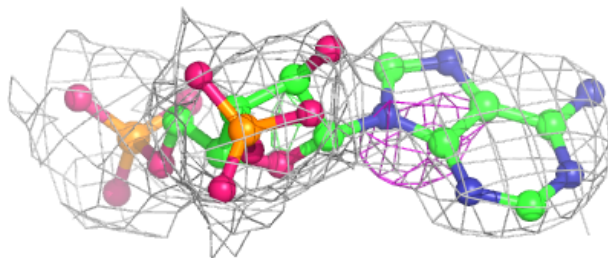
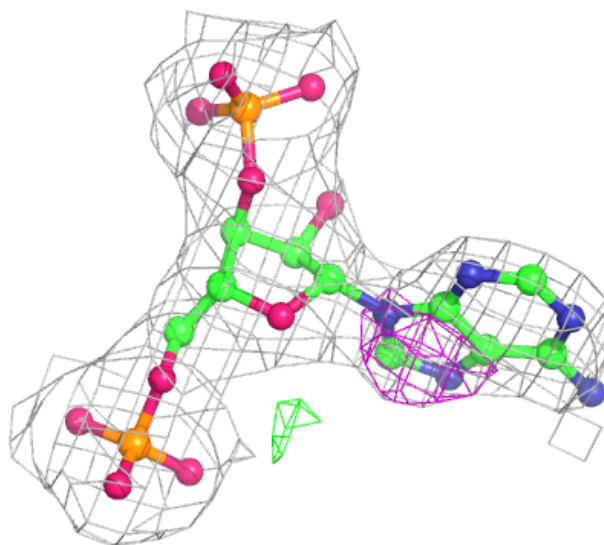
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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	A3P	A	301	27/27	0.94	0.17	52,58,61,62	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 A3P A 301:**

$2mF_o-DF_c$  (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.