



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 01:03 PM UTC

PDB ID : 3DOK / pdb_00003dok
Title : Crystal structure of K103N mutant HIV-1 reverse transcriptase in complex with GW678248.
Authors : Chamberlain, P.P.; Ren, J.; Stammers, D.K.
Deposited on : 2008-07-04
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 ⓘ 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)
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

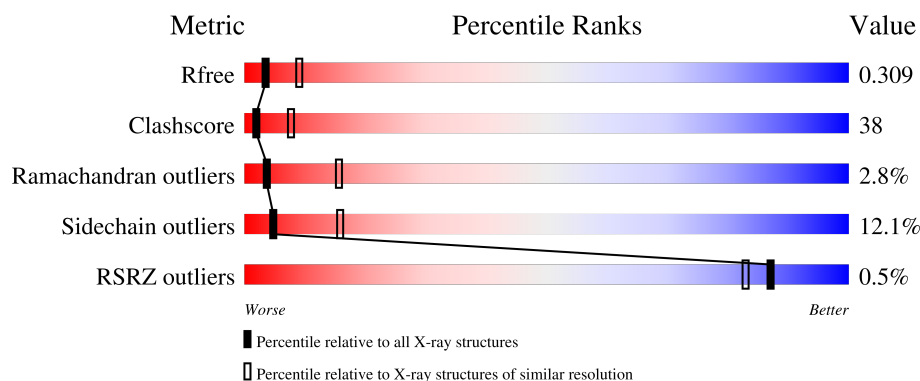
1 Overall quality at a glance

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}	180053	2481 (2.90-2.90)
Clashscore	190562	2690 (2.90-2.90)
Ramachandran outliers	187476	2623 (2.90-2.90)
Sidechain outliers	187428	2625 (2.90-2.90)
RSRZ outliers	180081	2481 (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 $\leq 5\%$. 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	560	
2	B	440	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 7750 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 Reverse transcriptase/ribonuclease H.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	536	Total	C	N	O	S	0	0	0
			4383	2835	729	811	8			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	103	ASN	LYS	engineered mutation	UNP P04585

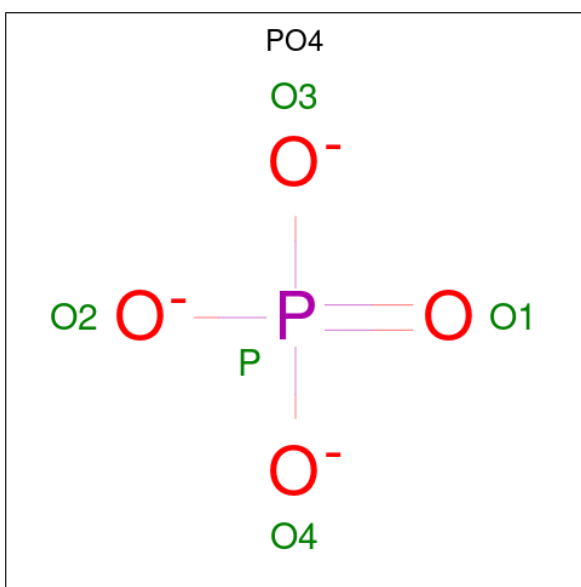
- Molecule 2 is a protein called p51 RT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	402	Total	C	N	O	S	0	0	0
			3320	2158	553	602	7			

There is a discrepancy between the modelled and reference sequences:

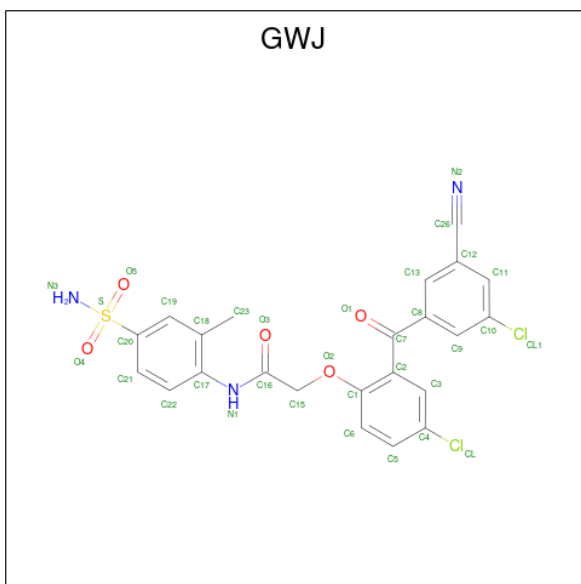
Chain	Residue	Modelled	Actual	Comment	Reference
B	103	ASN	LYS	engineered mutation	UNP P04585

- Molecule 3 is PHOSPHATE ION (CCD ID: PO4) (formula: O₄P).



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

- Molecule 4 is 2-{4-chloro-2-[(3-chloro-5-cyanophenyl)carbonyl]phenoxy}-N-(2-methyl-4-sulfamoylphenyl)acetamide (CCD ID: GWJ) (formula: $C_{23}H_{17}Cl_2N_3O_5S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	Cl	N	O	S	
			34	23	2	3	5	1	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	6	Total 6	O 6	0	0
5	B	2	Total 2	O 2	0	0

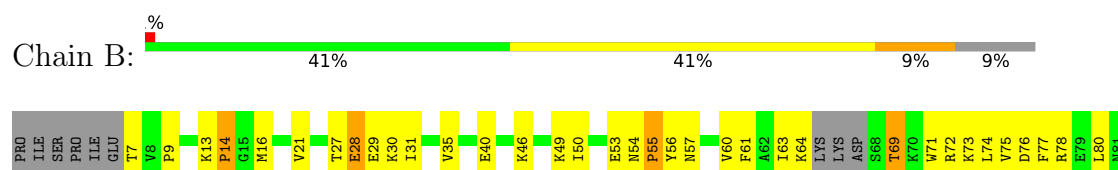
3 Residue-property plots

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: Reverse transcriptase/ribonuclease H



• Molecule 2: p51 RT



G384	K82	Q151	HIS	T296	G384
K385	R83	K154	GLN	E297	K385
T386	T84	GLY	LYS	E298	T386
P387	Q85	I159	PRO	E302	P387
F388	D86	F160	PRO	L303	F388
F389	F87	Q161	PHE	L310	F389
T393	W88	S162	LEU	Q331	T393
W398	GLU	S163	TRP	K331	W398
E399	VAL	M164	GLY	D320	E399
T400	GLN	T165	MET	P321	T400
W401	LEU	K166	TVR	S322	W401
W402	G93	I167	E233	K323	W402
T403	I94	L168	L234	D324	T403
E404	F95	E169	H235	L325	E404
W410	H96	F171	D237	Q332	W410
F416	P97	R172	K238	W337	F416
V417	A98	K173	W239	T338	V417
N418	G99	Q174	Q242	Y339	N418
P421	L100	N175	P243	Q340	P421
K424	K101	F176	I244	I341	K424
L425	V108	D177	V245	F342	L425
W426	D110	I178	D250	Q343	W426
Y427	V111	Y180	V254	F346	Y427
Q428	G112	Y181	Q258	L349	Q428
L429	D113	Q182	K259	G352	L429
E430	A114	D186	L260	K353	E430
K431	Y115	E187	V261	A355	K431
GLU	F116	Y188	W266	K356	GLU
PRO	S117	L193	A267	W357	PRO
I1E	V118	E194	S268	H361	I1E
VAL	P119	I195	Q269	T362	VAL
GLY	L120	G196	I270	D363	GLY
ALA	D121	K199	Y271	D364	ALA
GLU	E122	T200	P272	V365	GLU
THR	E123	K201	Q273	K366	THR
PHE	F124	L202	I274	Q367	PHE
	R125	E203	K275	T368	
	K126	H208	V276	T369	
	Y127	L209	R277	E370	
	T128	L210	Q278	Q373	
	A129	R211	L279	T376	
	F130	G213	K281	T377	
	T131	L214	L282	E378	
	I132	T215	T286	S379	
	F133	T216	K287	T380	
	S134	PRO	L288	V381	
	I135	ASP	L289	I382	
	N136	LYS	T290	W383	
	M137	LYS	L295		
	E138				
	T139				
	R143				
	Y144				
	Q145				
	L149				
	P150				

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	136.94Å 109.77Å 72.48Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.79 – 2.90 29.79 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.5 (29.79-2.90) 99.5 (29.79-2.90)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.46 (at 2.90Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.218 , 0.312 0.215 , 0.309	Depositor DCC
R_{free} test set	1203 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å ²)	80.4	Xtriage
Anisotropy	0.333	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 70.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	7750	wwPDB-VP
Average B, all atoms (Å ²)	75.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.91% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: PO4, CSD, GWJ

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	0.75	3/4491 (0.1%)	1.25	42/6107 (0.7%)
2	B	0.69	0/3412	1.19	29/4635 (0.6%)
All	All	0.72	3/7903 (0.0%)	1.23	71/10742 (0.7%)

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	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	184	MET	SD-CE	5.99	1.94	1.79
1	A	473	THR	N-CA	5.73	1.52	1.46
1	A	108	VAL	CA-CB	5.60	1.60	1.54

The worst 5 of 71 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	139	THR	CA-C-N	11.74	134.52	119.84
1	A	139	THR	C-N-CA	11.74	134.52	119.84
2	B	215	THR	N-CA-C	11.04	134.32	110.80
1	A	118	VAL	CA-C-N	9.95	130.05	119.90
1	A	118	VAL	C-N-CA	9.95	130.05	119.90

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	144	TYR	Sidechain
1	A	188	TYR	Sidechain

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	4383	0	4416	390	0
2	B	3320	0	3346	222	0
3	A	5	0	0	0	0
4	A	34	0	17	5	0
5	A	6	0	0	1	0
5	B	2	0	0	0	0
All	All	7750	0	7779	597	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 38.

The worst 5 of 597 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:139:THR:HB	1:A:140:PRO:HD2	1.28	1.14
1:A:448:ARG:H	1:A:448:ARG:HD3	0.93	1.09
1:A:49:LYS:H	1:A:49:LYS:HD3	0.97	1.05
1:A:49:LYS:H	1:A:49:LYS:CD	1.64	1.03
1:A:448:ARG:HD3	1:A:448:ARG:N	1.75	0.99

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	531/560 (95%)	441 (83%)	72 (14%)	18 (3%)	3	12
2	B	394/440 (90%)	347 (88%)	39 (10%)	8 (2%)	6	22
All	All	925/1000 (92%)	788 (85%)	111 (12%)	26 (3%)	4	15

5 of 26 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2	ILE
1	A	114	ALA
1	A	116	PHE
1	A	140	PRO
1	A	223	LYS

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	480/499 (96%)	409 (85%)	71 (15%)	3	10
2	B	365/400 (91%)	334 (92%)	31 (8%)	10	31
All	All	845/899 (94%)	743 (88%)	102 (12%)	5	16

5 of 102 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	452	LEU
2	B	14	PRO
2	B	386	THR
1	A	459	THR
1	A	497	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	336	GLN
2	B	361	HIS
2	B	428	GLN
1	A	475	GLN
1	A	361	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue 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	$\# Z > 2$
1	CSD	A	280	1	4,7,8	1.22	1 (25%)	1,8,10	0.89	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
1	CSD	A	280	1	-	1/2/6/8	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	280	CSD	OD1-SG	-2.03	1.45	1.47

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	280	CSD	N-CA-CB-SG

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

2 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	PO4	A	1300	-	4,4,4	1.73	2 (50%)	6,6,6	0.46	0
4	GWJ	A	999	-	36,36,36	2.74	9 (25%)	52,52,52	1.58	10 (19%)

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	GWJ	A	999	-	-	4/25/25/25	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	999	GWJ	C5-C6	7.11	1.50	1.38
4	A	999	GWJ	C17-C18	6.41	1.53	1.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	999	GWJ	C2-C1	6.27	1.52	1.40
4	A	999	GWJ	C19-C20	5.71	1.49	1.39
4	A	999	GWJ	C22-C21	5.58	1.47	1.38

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	999	GWJ	O5-S-O4	-4.55	111.80	118.80
4	A	999	GWJ	O5-S-N3	3.93	113.02	107.35
4	A	999	GWJ	C20-S-N3	-3.28	103.84	108.40
4	A	999	GWJ	C22-C17-C18	-2.92	117.37	120.70
4	A	999	GWJ	C15-O2-C1	2.58	122.99	117.60

There are no chirality outliers.

All (4) torsion outliers are listed below:

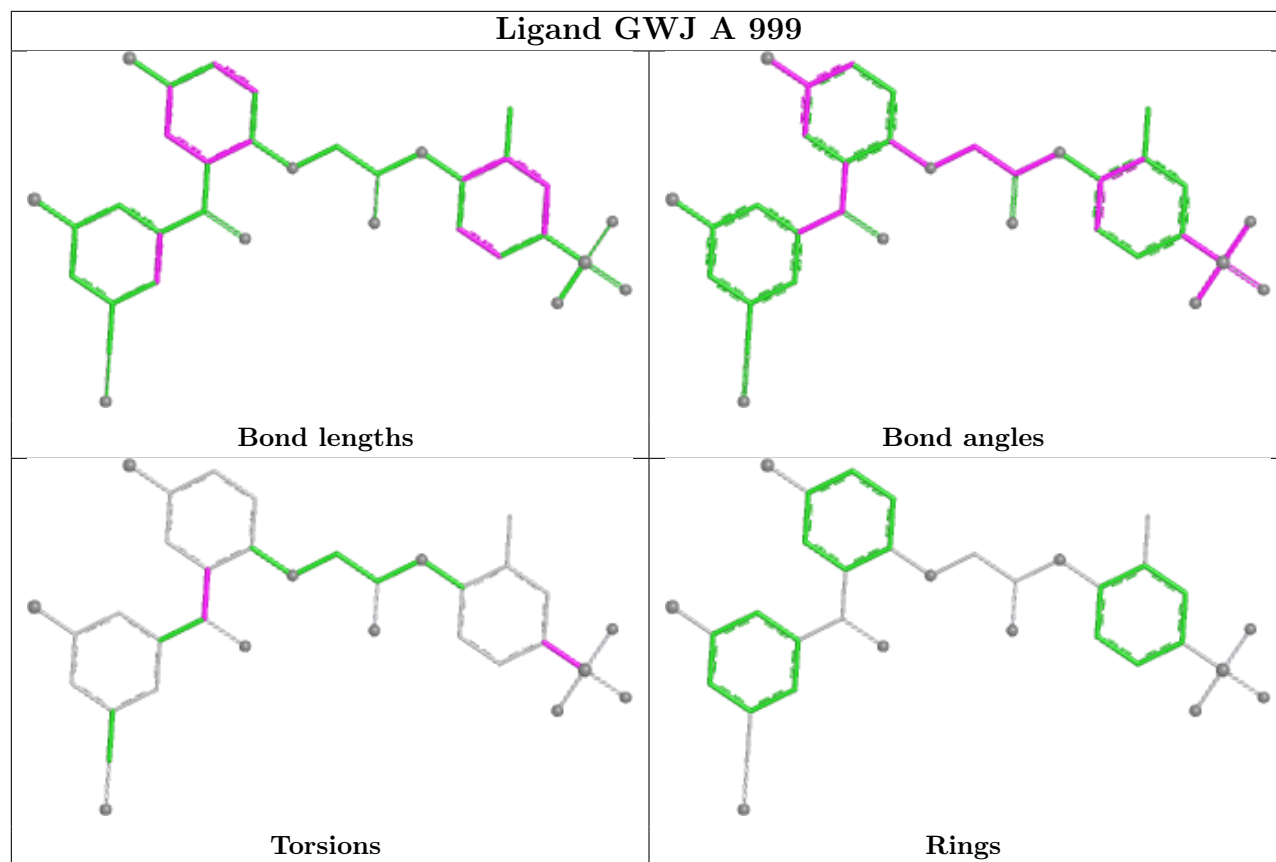
Mol	Chain	Res	Type	Atoms
4	A	999	GWJ	C19-C20-S-O5
4	A	999	GWJ	C21-C20-S-O5
4	A	999	GWJ	C1-C2-C7-O1
4	A	999	GWJ	C1-C2-C7-C8

There are no ring outliers.

1 monomer is involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	999	GWJ	5	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, 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	535/560 (95%)	-0.34	2 (0%) 88 85	37, 72, 116, 174	0
2	B	402/440 (91%)	-0.49	3 (0%) 84 79	32, 69, 120, 148	0
All	All	937/1000 (93%)	-0.40	5 (0%) 87 83	32, 71, 119, 174	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	93	GLY	3.0
1	A	455	ALA	2.9
2	B	215	THR	2.1
1	A	471	ASP	2.0
2	B	426	TRP	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, 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
1	CSD	A	280	8/9	0.95	0.08	58,64,81,94	0

6.3 Carbohydrates [i](#)

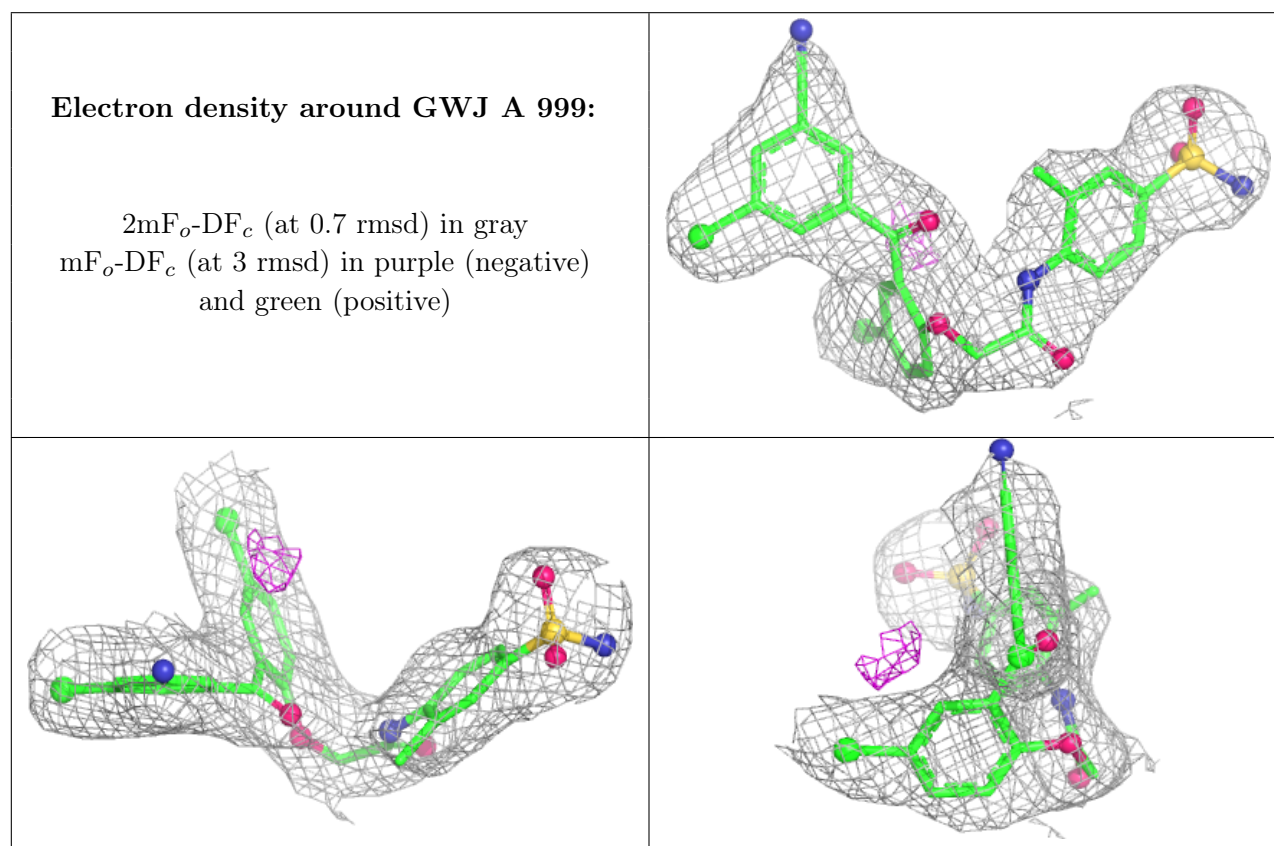
There are no oligosaccharides in this entry.

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, 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($\text{\AA}^2$)	Q<0.9
3	PO4	A	1300	5/5	0.62	0.16	132,137,152,154	0
4	GWJ	A	999	34/34	0.94	0.08	35,59,75,78	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.