



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 29, 2024 – 12:07 am GMT

PDB ID : 8A1Q  
Title : HIV-1 Integrase Catalytic Core Domain and C-Terminal Domain in Complex with Allosteric Integrase Inhibitor STP0404 (Pirmitegravir)  
Authors : Singer, M.R.; Pye, V.E.; Cook, N.J.; Cherepanov, P.  
Deposited on : 2022-06-01  
Resolution : 2.06 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 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 : 4.02b-467  
Mogul : 1.8.4, 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

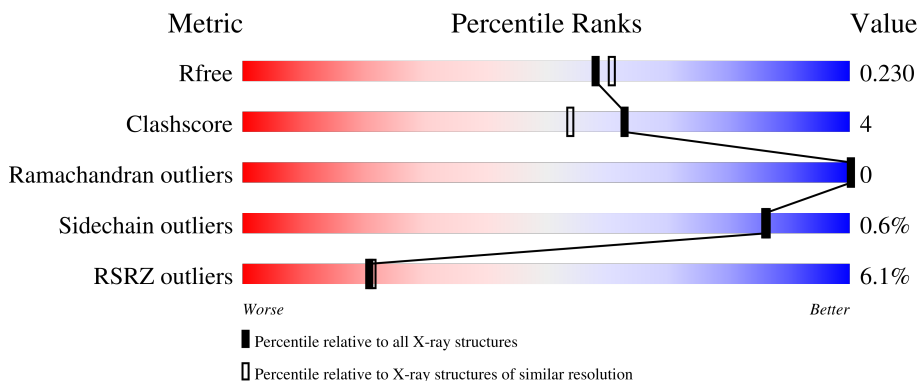
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.06 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	2684 (2.08-2.04)
Clashscore	141614	2801 (2.08-2.04)
Ramachandran outliers	138981	2768 (2.08-2.04)
Sidechain outliers	138945	2768 (2.08-2.04)
RSRZ outliers	127900	2646 (2.08-2.04)

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  $\geq 3$ , 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	233	
1	B	233	
1	C	233	
1	D	233	

## 2 Entry composition i

There are 5 unique types of molecules in this entry. The entry contains 3498 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 Integrase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	57	Total 470	C 297	N 91	O 81	S 1	0	1	0
1	B	151	Total 1176	C 752	N 205	O 215	S 4	0	1	0
1	C	58	Total 469	C 295	N 88	O 85	S 1	0	0	0
1	D	147	Total 1141	C 728	N 200	O 209	S 4	0	1	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	219	SER	-	expression tag	UNP P12497
A	243	GLU	TRP	engineered mutation	UNP P12497
A	424	LYS	PHE	engineered mutation	UNP P12497
B	-20	SER	-	expression tag	UNP P12497
B	4	GLU	TRP	engineered mutation	UNP P12497
B	185	LYS	PHE	engineered mutation	UNP P12497
C	219	SER	-	expression tag	UNP P12497
C	243	GLU	TRP	engineered mutation	UNP P12497
C	424	LYS	PHE	engineered mutation	UNP P12497
D	-20	SER	-	expression tag	UNP P12497
D	4	GLU	TRP	engineered mutation	UNP P12497
D	185	LYS	PHE	engineered mutation	UNP P12497

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	C	1	Total C O 4 2 2	0	0
2	C	1	Total C O 4 2 2	0	0
2	D	1	Total C O 4 2 2	0	0
2	D	1	Total C O 4 2 2	0	0
2	D	1	Total C O 4 2 2	0	0

*Continued on next page...*



*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
5	B	42	Total 42	O 42	0	0
5	C	6	Total 6	O 6	0	0
5	D	54	Total 54	O 54	0	0



ALA  
VAL  
PHE  
ILE  
HIS  
ASN  
LYS  
LYS  
ARG  
LYS  
GLY  
GLY  
ILE  
GLY  
TYR  
SER  
ALA  
GLY  
GLU  
ARG  
ILE  
VAL  
ASP  
ASP  
ILE  
ILE  
ALA  
THR  
ASP  
ILE  
GLN  
THR  
LYS  
GLU

● Molecule 1: Integrase

Chain D: 61% 37%

SER  
ILE  
GLN  
ASN  
PHE  
SER  
VAL  
TYR  
TYR  
ARG  
ASP  
SER  
ARG  
ASP  
PRO  
VAL  
TRP  
CYS  
LYS  
GLY  
PRO  
ALA  
LYS  
LEU  
LEU  
GLU  
LYS  
GLY  
GLU  
GLY  
ALA  
VAL  
VAL  
ILE  
GLN  
ASP  
ASN  
SER  
ASP  
ILE  
LYS  
VAL  
VAL  
PRO  
ARG  
ARG  
LYS  
ALA  
LYS  
ILE  
ILE  
ARG  
ASP  
TYR  
GLY  
LYS  
GLN  
MET  
ALA  
GLY  
ASP

ASP  
CYS  
VAL  
ALA  
ARG  
GLN  
ASP  
GLU  
ASP  
MET  
HIS  
GLY  
GLN  
VAL  
ASP  
CYS  
S57  
V77  
I84  
T93  
W131  
G140  
ILE  
PRO  
TYR  
ASN  
PRO  
GLN  
SER  
GLN  
G149  
W150  
I151  
K173  
R187  
K211  
GLU



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	62.26Å 69.26Å 64.11Å 90.00° 102.38° 90.00°	Depositor
Resolution (Å)	62.61 – 2.06 62.62 – 2.06	Depositor EDS
% Data completeness (in resolution range)	99.7 (62.61-2.06) 99.8 (62.62-2.06)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.04 (at 2.07Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, $R_{free}$	0.219 , 0.235 0.218 , 0.230	Depositor DCC
$R_{free}$ test set	1634 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	57.4	Xtrriage
Anisotropy	0.257	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 66.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.095 for l,-k,h	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	3498	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	79.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.86% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: WBV, EDO, 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	0.24	0/481	0.54	0/643
1	B	0.24	0/1200	0.43	0/1623
1	C	0.24	0/477	0.51	0/639
1	D	0.23	0/1163	0.43	0/1571
All	All	0.24	0/3321	0.46	0/4476

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 [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	470	0	490	5	0
1	B	1176	0	1186	12	0
1	C	469	0	477	10	0
1	D	1141	0	1153	6	0
2	A	8	0	12	1	0
2	B	28	0	42	1	0
2	C	8	0	12	0	0
2	D	16	0	24	2	0
3	B	35	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	D	35	0	0	0	0
4	B	1	0	0	0	0
4	D	1	0	0	0	0
5	A	8	0	0	0	0
5	B	42	0	0	0	0
5	C	6	0	0	0	0
5	D	54	0	0	0	0
All	All	3498	0	3396	29	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 4.

All (29) 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:C:244:LYS:HE2	1:C:267:ILE:HD13	1.71	0.72
1:B:95:GLN:HG2	2:B:302:EDO:H11	1.76	0.66
1:B:185:LYS:HD3	1:C:276:ALA:HB2	1.83	0.61
1:B:77:VAL:HG22	1:B:84:ILE:HG22	1.86	0.57
1:D:173:LYS:HZ2	2:D:1303:EDO:H12	1.68	0.57
1:D:77:VAL:HG22	1:D:84:ILE:HG22	1.88	0.56
1:C:227:TYR:HB2	1:C:260:VAL:HG11	1.89	0.53
1:C:244:LYS:HG2	1:C:249:VAL:HG12	1.91	0.52
1:C:224:ARG:HG2	1:D:131:TRP:CE2	2.45	0.52
1:C:241:LEU:HA	1:C:251:ILE:HG22	1.94	0.50
1:B:211:LYS:HE2	1:D:187:ARG:HH22	1.78	0.49
1:B:152:GLU:HG2	1:B:156:LYS:HE3	1.94	0.49
1:C:268:ILE:HG22	1:D:131:TRP:HD1	1.79	0.47
1:A:266:LYS:HB3	2:A:502:EDO:H22	1.98	0.45
1:C:270:ASP:HB2	1:C:273:LYS:HB3	1.97	0.45
1:B:103:LYS:O	1:B:107:ARG:HG2	2.16	0.45
1:C:225:VAL:HG22	1:C:267:ILE:HG12	1.99	0.44
1:D:93:THR:HA	2:D:1305:EDO:H22	1.99	0.44
1:A:226:TYR:HD1	1:A:268:ILE:HD13	1.83	0.43
1:B:66:THR:HG23	1:B:73:ILE:HB	2.01	0.42
1:B:57:SER:HB3	1:B:58:PRO:HD3	2.03	0.41
1:B:167:ASP:OD1	1:B:167:ASP:N	2.54	0.41
1:B:62:GLN:HG2	1:B:114:HIS:HB2	2.03	0.41
1:A:270:ASP:HB2	1:A:273:LYS:HE2	2.02	0.41
1:B:101:LEU:HD23	1:B:101:LEU:HA	1.88	0.41
1:C:278:ASP:OD1	1:C:279:ASP:N	2.53	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:68:LEU:HD11	1:B:159:LYS:HD2	2.03	0.40
1:A:224[B]:ARG:HD2	1:A:238:PRO:HB2	2.02	0.40
1:A:240:LYS:HA	1:A:240:LYS:HD3	1.94	0.40

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	56/233 (24%)	52 (93%)	4 (7%)	0	100	100
1	B	148/233 (64%)	146 (99%)	2 (1%)	0	100	100
1	C	56/233 (24%)	53 (95%)	3 (5%)	0	100	100
1	D	144/233 (62%)	139 (96%)	5 (4%)	0	100	100
All	All	404/932 (43%)	390 (96%)	14 (4%)	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	49/193 (25%)	49 (100%)	0	100	100
1	B	123/193 (64%)	121 (98%)	2 (2%)	62	59

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	49/193 (25%)	49 (100%)	0	100	100
1	D	119/193 (62%)	119 (100%)	0	100	100
All	All	340/772 (44%)	338 (99%)	2 (1%)	86	86

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	144	ASN
1	B	211	LYS

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

Mol	Chain	Res	Type
1	C	222	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](#)

Of 19 ligands modelled in this entry, 2 are monoatomic - leaving 17 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
2	EDO	B	302	-	3,3,3	0.46	0	2,2,2	0.32	0
2	EDO	D	1305	-	3,3,3	0.46	0	2,2,2	0.35	0
2	EDO	D	1302	-	3,3,3	0.46	0	2,2,2	0.37	0
2	EDO	C	501	-	3,3,3	0.44	0	2,2,2	0.36	0
2	EDO	B	305	-	3,3,3	0.46	0	2,2,2	0.34	0
2	EDO	B	306	-	3,3,3	0.47	0	2,2,2	0.30	0
3	WBV	D	1301	-	34,38,38	1.75	5 (14%)	37,58,58	1.84	7 (18%)
2	EDO	B	304	-	3,3,3	0.46	0	2,2,2	0.31	0
2	EDO	B	301	-	3,3,3	0.46	0	2,2,2	0.33	0
2	EDO	B	307	-	3,3,3	0.45	0	2,2,2	0.32	0
2	EDO	C	502	-	3,3,3	0.46	0	2,2,2	0.36	0
3	WBV	B	308	-	34,38,38	1.74	7 (20%)	37,58,58	1.89	6 (16%)
2	EDO	A	502	-	3,3,3	0.46	0	2,2,2	0.34	0
2	EDO	A	501	-	3,3,3	0.46	0	2,2,2	0.33	0
2	EDO	B	303	-	3,3,3	0.47	0	2,2,2	0.34	0
2	EDO	D	1304	-	3,3,3	0.45	0	2,2,2	0.36	0
2	EDO	D	1303	-	3,3,3	0.46	0	2,2,2	0.34	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	EDO	B	302	-	-	0/1/1/1	-
2	EDO	D	1305	-	-	0/1/1/1	-
2	EDO	D	1302	-	-	0/1/1/1	-
2	EDO	C	501	-	-	0/1/1/1	-
2	EDO	B	305	-	-	0/1/1/1	-
2	EDO	B	306	-	-	0/1/1/1	-
3	WBV	D	1301	-	-	2/18/21/21	0/4/4/4
2	EDO	B	304	-	-	0/1/1/1	-
2	EDO	B	301	-	-	0/1/1/1	-
2	EDO	B	307	-	-	0/1/1/1	-
2	EDO	C	502	-	-	0/1/1/1	-
3	WBV	B	308	-	-	3/18/21/21	0/4/4/4
2	EDO	A	502	-	-	0/1/1/1	-
2	EDO	A	501	-	-	0/1/1/1	-
2	EDO	B	303	-	-	0/1/1/1	-
2	EDO	D	1304	-	-	0/1/1/1	-
2	EDO	D	1303	-	-	0/1/1/1	-

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	1301	WBV	CBB-CBA	5.40	1.56	1.49
3	B	308	WBV	CBB-CBA	5.22	1.56	1.49
3	D	1301	WBV	CAP-NBG	-4.21	1.31	1.35
3	B	308	WBV	CAP-NBG	-4.15	1.31	1.35
3	B	308	WBV	CAZ-NBH	-3.47	1.32	1.39
3	D	1301	WBV	CAZ-NBH	-3.44	1.32	1.39
3	B	308	WBV	CBB-CBC	2.25	1.41	1.38
3	D	1301	WBV	CBC-CBF	2.22	1.56	1.52
3	B	308	WBV	CBC-CBF	2.22	1.56	1.52
3	D	1301	WBV	CBB-CBC	2.14	1.41	1.38
3	B	308	WBV	CAV-CLA	2.05	1.79	1.74
3	B	308	WBV	CAA-CAW	2.04	1.53	1.50

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	308	WBV	CAO-NAR-NBG	5.74	110.21	104.23
3	D	1301	WBV	CAO-NAR-NBG	5.72	110.19	104.23
3	B	308	WBV	CAX-CAQ-NBH	-5.52	104.05	112.63
3	D	1301	WBV	CAX-CAQ-NBH	-4.90	105.01	112.63
3	D	1301	WBV	CAD-NBG-NAR	3.55	124.68	120.50
3	B	308	WBV	CAD-NBG-NAR	3.42	124.52	120.50
3	B	308	WBV	CAY-CAZ-NBH	3.15	108.66	106.66
3	D	1301	WBV	CAY-CAZ-NBH	3.15	108.66	106.66
3	B	308	WBV	CAW-NAS-CBD	3.11	121.99	117.75
3	D	1301	WBV	CAW-NAS-CBD	3.02	121.87	117.75
3	B	308	WBV	CAC-CAZ-NBH	2.46	125.79	122.69
3	D	1301	WBV	CBA-CBB-CBC	-2.38	119.25	121.55
3	D	1301	WBV	CAC-CAZ-NBH	2.27	125.55	122.69

There are no chirality outliers.

All (5) torsion outliers are listed below:

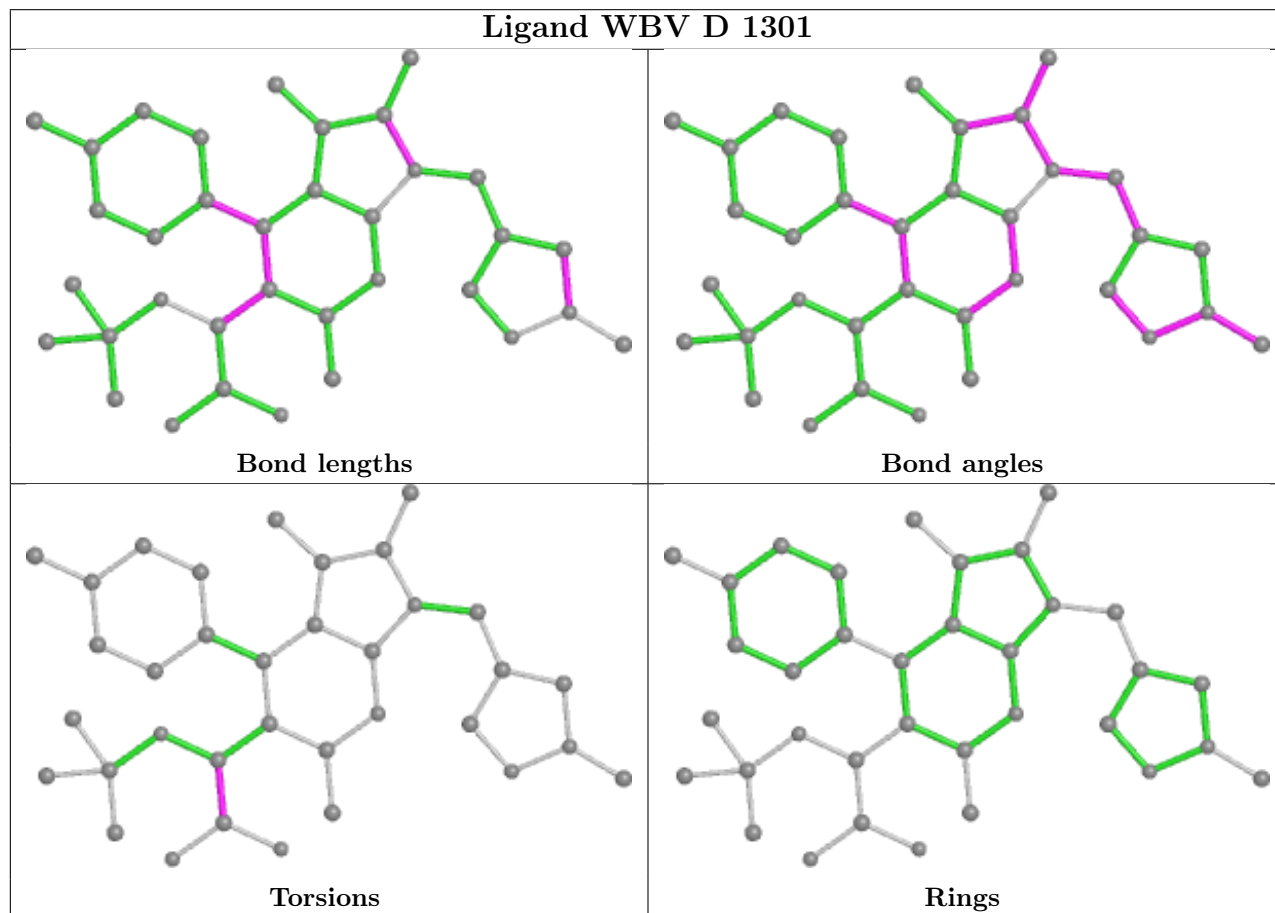
Mol	Chain	Res	Type	Atoms
3	B	308	WBV	OAH-CAU-CBF-CBC
3	B	308	WBV	OAI-CAU-CBF-CBC
3	D	1301	WBV	OAH-CAU-CBF-CBC
3	D	1301	WBV	OAI-CAU-CBF-CBC
3	B	308	WBV	OAI-CAU-CBF-OAT

There are no ring outliers.

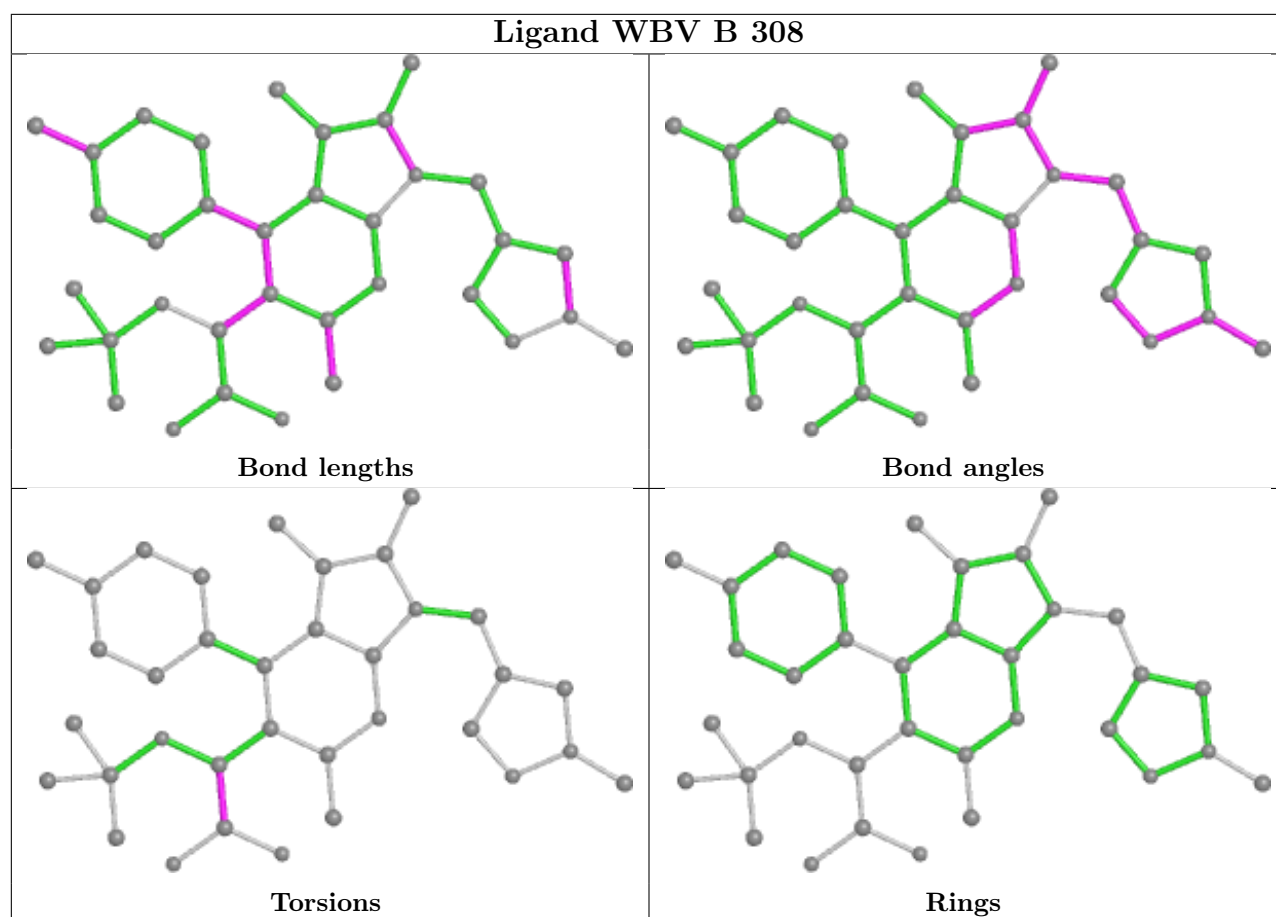
4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	302	EDO	1	0
2	D	1305	EDO	1	0
2	A	502	EDO	1	0
2	D	1303	EDO	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

### 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<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	57/233 (24%)	0.74	6 (10%) <b>6</b> <b>6</b>	77, 98, 123, 139	0
1	B	151/233 (64%)	0.45	2 (1%) <b>77</b> <b>78</b>	48, 66, 119, 134	0
1	C	58/233 (24%)	1.42	16 (27%) <b>0</b> <b>0</b>	77, 108, 151, 164	0
1	D	147/233 (63%)	0.44	1 (0%) <b>87</b> <b>88</b>	46, 65, 95, 117	0
All	All	413/932 (44%)	0.62	25 (6%) <b>21</b> <b>21</b>	46, 75, 123, 164	0

All (25) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	262	ARG	5.7
1	C	244	LYS	5.4
1	C	260	VAL	5.4
1	C	249	VAL	4.9
1	C	250	VAL	4.8
1	A	259	VAL	3.5
1	A	257	ILE	3.5
1	C	242	LEU	3.2
1	C	238	PRO	3.2
1	C	278	ASP	3.1
1	C	259	VAL	3.0
1	A	251	ILE	2.9
1	D	151	ILE	2.9
1	C	258	LYS	2.8
1	A	250	VAL	2.7
1	A	246	GLU	2.6
1	C	243	GLU	2.5
1	C	246	GLU	2.4
1	A	267	ILE	2.3
1	C	251	ILE	2.3
1	B	58	PRO	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	143	TYR	2.2
1	C	225	VAL	2.2
1	C	248	ALA	2.0
1	C	240	LYS	2.0

## 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](#)

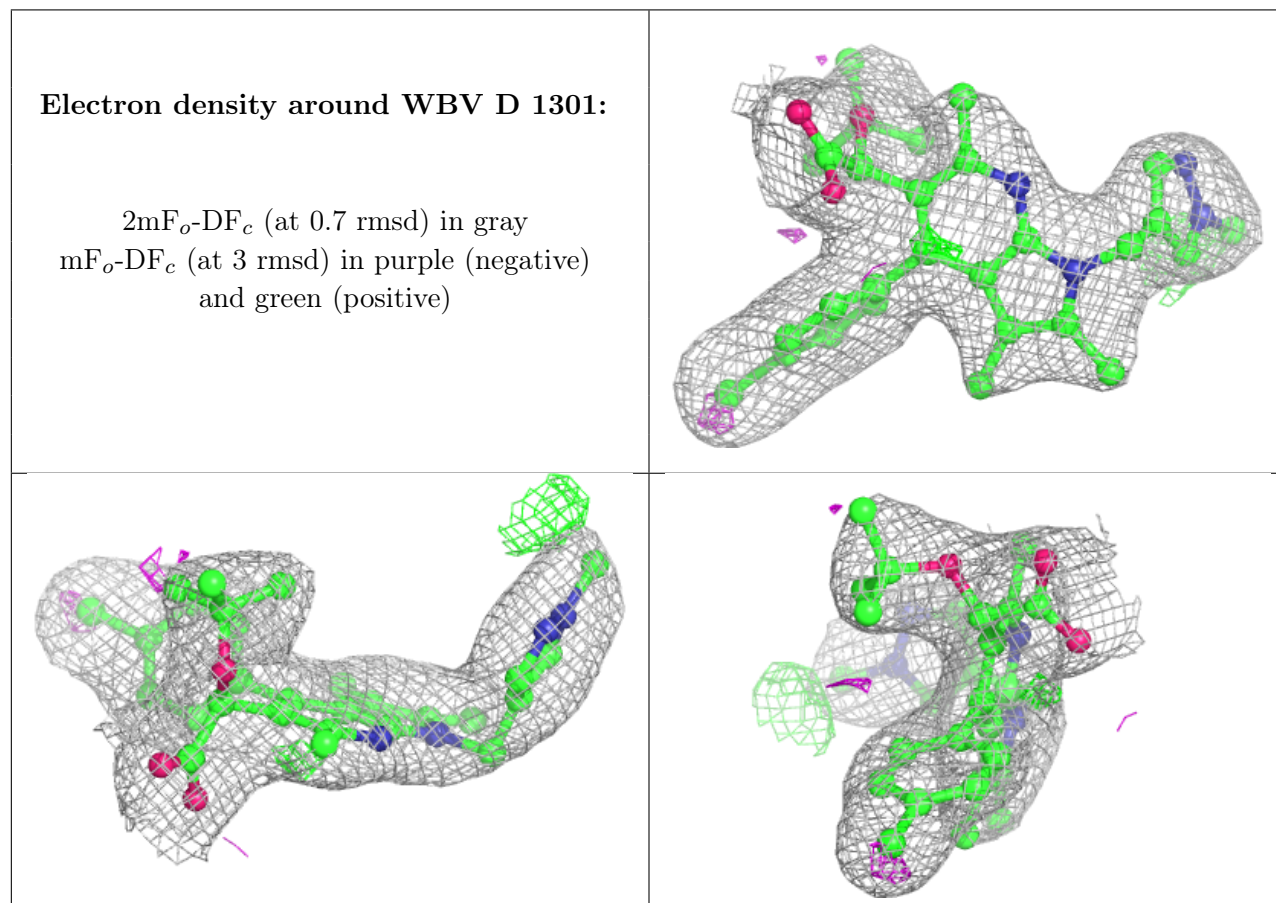
There are no monosaccharides in this entry.

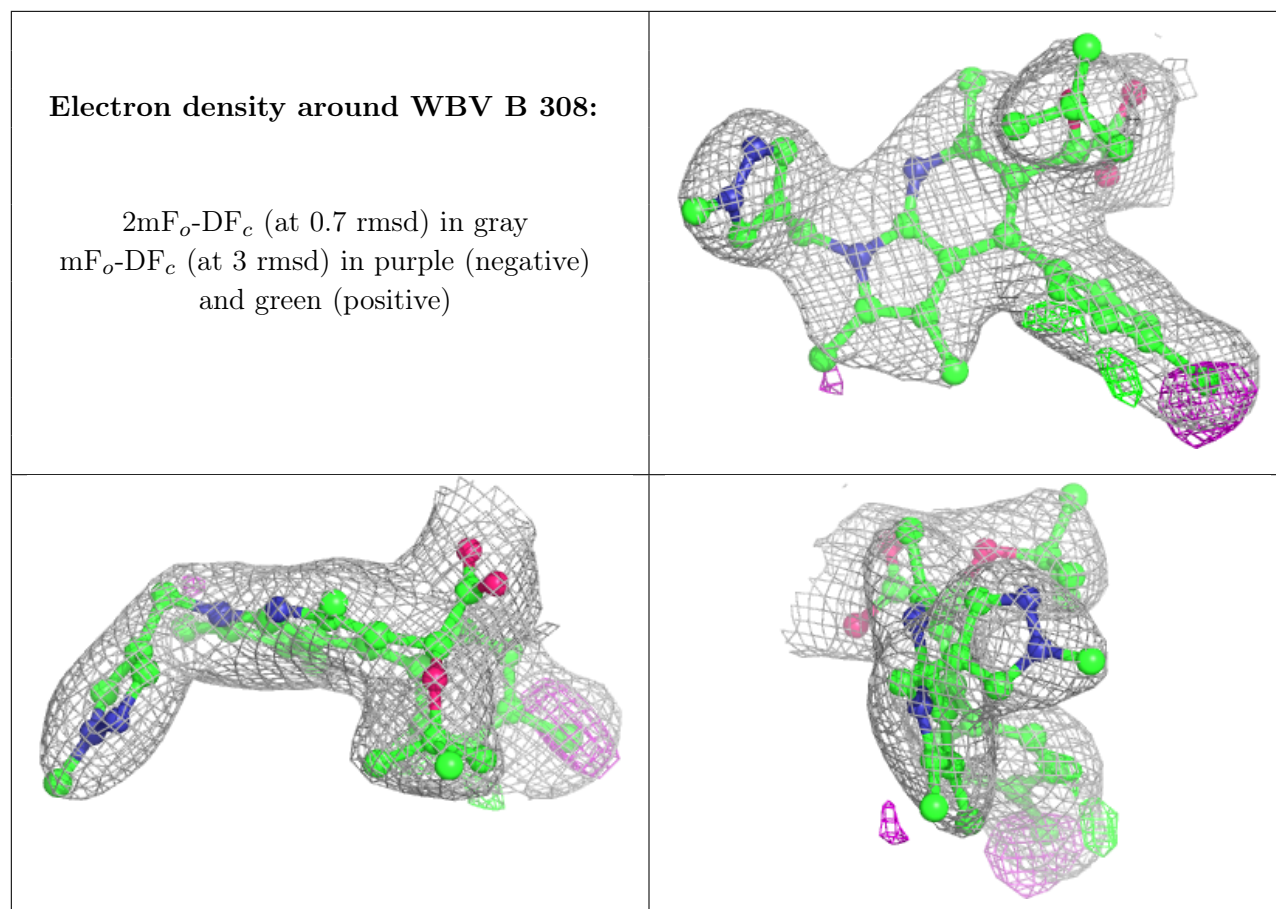
## 6.4 Ligands [i](#)

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, 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
4	MG	D	1306	1/1	0.59	0.16	69,69,69,69	0
2	EDO	B	301	4/4	0.68	0.27	79,80,82,83	0
2	EDO	B	304	4/4	0.71	0.22	65,69,72,74	0
2	EDO	D	1305	4/4	0.75	0.26	75,76,80,82	0
2	EDO	A	501	4/4	0.76	0.18	110,111,114,115	0
2	EDO	D	1303	4/4	0.78	0.16	69,72,73,75	0
2	EDO	C	502	4/4	0.79	0.20	75,80,83,85	0
2	EDO	A	502	4/4	0.80	0.30	66,72,73,78	0
4	MG	B	309	1/1	0.84	0.09	57,57,57,57	0
2	EDO	D	1304	4/4	0.84	0.23	72,76,79,80	0
2	EDO	B	302	4/4	0.89	0.17	70,73,79,86	0
2	EDO	B	305	4/4	0.89	0.14	68,69,74,79	0
2	EDO	D	1302	4/4	0.91	0.16	65,68,73,74	0
2	EDO	B	307	4/4	0.91	0.30	63,63,67,71	0
2	EDO	B	303	4/4	0.91	0.08	64,65,70,75	0
2	EDO	C	501	4/4	0.94	0.33	80,81,82,83	0
3	WBV	D	1301	35/35	0.95	0.15	54,61,72,73	0
2	EDO	B	306	4/4	0.95	0.26	70,77,78,83	0
3	WBV	B	308	35/35	0.95	0.16	57,65,72,75	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.