



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 6, 2026 – 02:32 PM UTC

PDB ID : 8PO2 / pdb_00008po2
Title : Discovery and Optimisation of Potent, Efficacious and Selective Inhibitors Targeting EGFR Exon20 Insertion Mutations. Compound 33 bound to EGFRin-sNPG [V948R]
Authors : Hargreaves, D.
Deposited on : 2023-07-03
Resolution : 2.28 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

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

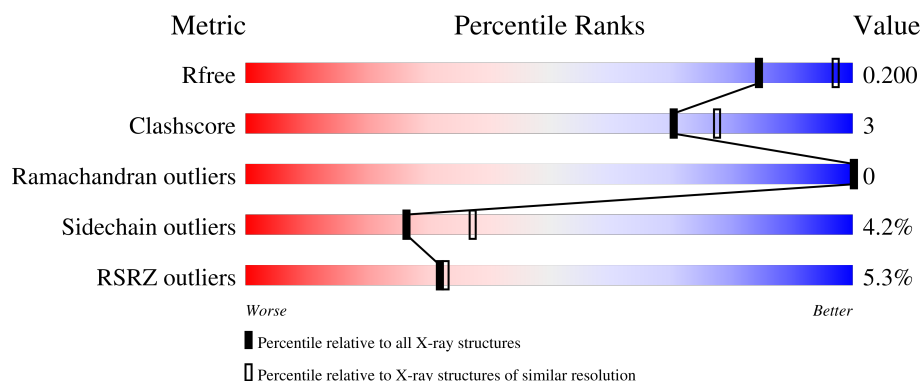
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.28 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	9078 (2.30-2.26)
Clashscore	190562	9802 (2.30-2.26)
Ramachandran outliers	187476	9690 (2.30-2.26)
Sidechain outliers	187428	9691 (2.30-2.26)
RSRZ outliers	180081	9085 (2.30-2.26)

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	332	4% 72% 10% 18%
1	B	332	3% 71% 11% 17%
1	D	332	2% 5% 95%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4810 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.

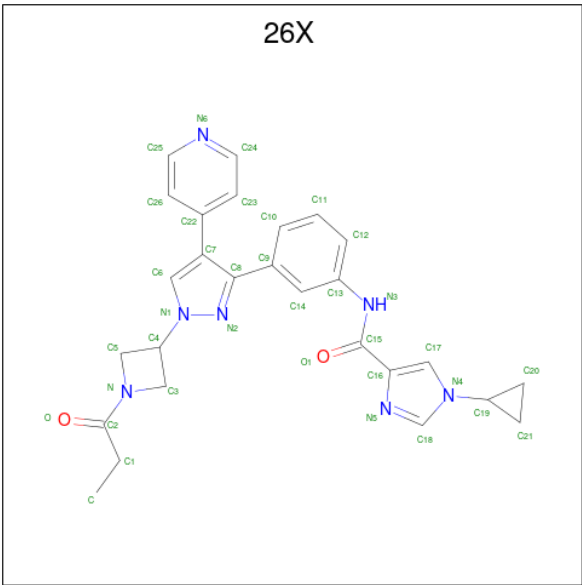
- Molecule 1 is a protein called Epidermal growth factor receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	271	Total	C	N	O	S	0	0	0
			2172	1401	371	385	15			
1	B	274	Total	C	N	O	S	0	0	0
			2193	1416	374	388	15			
1	D	18	Total	C	N	O	S	0	0	0
			145	88	18	37	2			

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	694	GLY	-	expression tag	UNP P00533
A	773	GLY	-	insertion	UNP P00533
A	774	ASN	-	insertion	UNP P00533
A	775	PRO	-	insertion	UNP P00533
A	951	ARG	VAL	engineered mutation	UNP P00533
B	694	GLY	-	expression tag	UNP P00533
B	773	GLY	-	insertion	UNP P00533
B	774	ASN	-	insertion	UNP P00533
B	775	PRO	-	insertion	UNP P00533
B	951	ARG	VAL	engineered mutation	UNP P00533
D	-311	GLY	-	expression tag	UNP P00533
D	-232	GLY	-	insertion	UNP P00533
D	-231	ASN	-	insertion	UNP P00533
D	-230	PRO	-	insertion	UNP P00533
D	-54	ARG	VAL	engineered mutation	UNP P00533

- Molecule 2 is 1-cyclopropyl- {N}-[3-[1-(1-propanoylazetid-3-yl)-4-pyridin-4-yl-pyrazol-3-yl]phenyl]imidazole-4-carboxamide (CCD ID: 26X) (formula: C₂₇H₂₇N₇O₂) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			36	27	7	2		
2	B	1	Total	C	N	O	0	0
			36	27	7	2		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	109	Total	O	0	0
			109	109		
3	B	108	Total	O	0	0
			108	108		
3	D	11	Total	O	0	0
			11	11		

SER	PRO	THR	ASP	SER	ASN	PHE	TYR	ARG	ALA	LEU	WO	D7	V8	V9	B12	E13	Y14	L15	I16	P17	GLN	GLN	GLY																																							
ALA	GLU	GLY	GLY	VAL	PRO	PRO	ILE	CYS	THR	ILE	ASP	VAL	ILE	MET	ARG	LYS	CYS	TRP	MET	ILE	ASP	ALA	ASP	SER	VAL	TRP	SER	TYR	GLY	VAL	THR	THR	VAL	TRP	GLU	LEU	MET	THR	PHE	GLY	SER	LYS	PRO	TYR	ASP	GLY	ILE	PRO	ALA	SER	GLU	ILE	SER	GLY	ILE	SER	GLU	LEU	HIS	LEU	LYS	TYR

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	137.12Å 85.71Å 97.84Å 90.00° 121.76° 90.00°	Depositor
Resolution (Å)	69.06 – 2.28 69.06 – 2.28	Depositor EDS
% Data completeness (in resolution range)	70.2 (69.06-2.28) 70.9 (69.06-2.28)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.74 (at 2.27Å)	Xtriage
Refinement program	BUSTER 2.11.8 (24-FEB-2021)	Depositor
R, R_{free}	0.194 , 0.246 0.191 , 0.200	Depositor DCC
R_{free} test set	1539 reflections (3.45%)	wwPDB-VP
Wilson B-factor (Å ²)	41.2	Xtriage
Anisotropy	0.010	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 38.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	4810	wwPDB-VP
Average B, all atoms (Å ²)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.04% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: 26X

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	0.77	0/2220	1.04	1/3003 (0.0%)
1	B	0.75	1/2241 (0.0%)	1.07	6/3032 (0.2%)
1	D	0.72	0/146	1.09	0/198
All	All	0.76	1/4607 (0.0%)	1.05	7/6233 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	955	MET	SD-CE	-5.35	1.66	1.79

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	770	ASP	CA-CB-CG	7.65	120.25	112.60
1	B	751	THR	N-CA-C	-6.04	104.61	111.14
1	B	740	ILE	N-CA-C	5.99	113.81	107.76
1	B	945	ASP	CA-CB-CG	5.39	117.99	112.60
1	B	754	LYS	N-CA-C	5.04	117.20	110.35
1	B	905	VAL	CA-C-N	5.01	126.96	120.44
1	B	905	VAL	C-N-CA	5.01	126.96	120.44

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	2172	0	2235	15	0
1	B	2193	0	2262	16	0
1	D	145	0	123	2	0
2	A	36	0	0	0	0
2	B	36	0	0	1	0
3	A	109	0	0	0	0
3	B	108	0	0	1	0
3	D	11	0	0	1	0
All	All	4810	0	4620	31	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 3.

All (31) 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:706:ILE:HD11	1:A:760:LEU:HD11	1.76	0.66
1:B:896:HIS:HD2	3:B:1253:HOH:O	1.79	0.65
1:B:798:PHE:HB2	1:B:848:VAL:HB	1.87	0.56
1:B:776:HIS:HE1	3:D:106:HOH:O	1.87	0.56
1:A:853:HIS:CE1	1:D:9:VAL:HG12	2.41	0.56
1:B:775:PRO:O	1:B:853:HIS:HE1	1.89	0.55
1:B:731:TRP:CZ2	1:B:733:PRO:HB3	2.43	0.54
1:B:929:ILE:HG23	1:B:934:GLU:HB2	1.91	0.52
1:A:706:ILE:CD1	1:A:760:LEU:HD11	2.41	0.51
1:B:774:ASN:ND2	1:B:776:HIS:H	2.09	0.49
1:A:882:LYS:HG2	1:A:918:TYR:HD2	1.77	0.49
1:A:707:LEU:HD13	1:A:782:GLY:HA3	1.96	0.47
1:A:774:ASN:HD22	1:A:776:HIS:H	1.62	0.46
1:B:907:VAL:O	1:B:911:MET:HG2	2.17	0.45
1:A:838:HIS:O	1:A:839:ARG:HB2	2.18	0.44
1:B:707:LEU:HD13	1:B:782:GLY:HA3	1.99	0.43
1:A:948:MET:O	1:A:952:LYS:HG3	2.18	0.43
1:B:950:MET:O	1:B:953:CYS:HB2	2.17	0.43
1:B:861:LEU:HD13	2:B:1101:26X:C18	2.48	0.43
1:A:718:LEU:HD21	1:A:728:LYS:HB2	2.02	0.42
1:A:849:LYS:HZ3	1:D:7:ASP:HB3	1.85	0.42
1:B:760:LEU:HG	1:B:785:LEU:HD12	2.00	0.42
1:A:815:GLN:HE21	1:A:819:ASN:ND2	2.19	0.41
1:A:834:ARG:N	1:A:834:ARG:HD2	2.35	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:830:TYR:CE1	1:A:834:ARG:HD3	2.56	0.41
1:B:740:ILE:HA	1:B:741:PRO:HD3	2.00	0.41
1:B:944:ILE:HD12	1:B:944:ILE:HA	1.89	0.41
1:B:707:LEU:HD22	1:B:792:ILE:HD13	2.03	0.40
1:B:718:LEU:HD21	1:B:728:LYS:HB2	2.03	0.40
1:A:707:LEU:HD22	1:A:792:ILE:HD13	2.03	0.40
1:A:785:LEU:HD22	1:A:789:VAL:HG22	2.04	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	267/332 (80%)	260 (97%)	7 (3%)	0	100	100
1	B	270/332 (81%)	261 (97%)	9 (3%)	0	100	100
1	D	16/332 (5%)	14 (88%)	2 (12%)	0	100	100
All	All	553/996 (56%)	535 (97%)	18 (3%)	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	240/290 (83%)	230 (96%)	10 (4%)	26	37
1	B	242/290 (83%)	233 (96%)	9 (4%)	30	43
1	D	17/290 (6%)	15 (88%)	2 (12%)	5	5
All	All	499/870 (57%)	478 (96%)	21 (4%)	26	37

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	720	SER
1	A	730	LEU
1	A	760	LEU
1	A	807	GLU
1	A	861	LEU
1	A	864	LEU
1	A	929	ILE
1	A	951	ARG
1	A	973	LYS
1	A	980	ARG
1	B	705	ARG
1	B	730	LEU
1	B	736	GLU
1	B	769	VAL
1	B	861	LEU
1	B	929	ILE
1	B	951	ARG
1	B	965	ARG
1	B	980	ARG
1	D	0	MET
1	D	9	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (18) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	756	ASN
1	A	776	HIS
1	A	794	GLN
1	A	819	ASN
1	A	829	ASN
1	A	845	ASN
1	A	938	GLN
1	A	979	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	756	ASN
1	B	771	ASN
1	B	774	ASN
1	B	776	HIS
1	B	790	GLN
1	B	815	GLN
1	B	845	ASN
1	B	853	HIS
1	B	938	GLN
1	B	979	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

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	$\# Z > 2$
2	26X	B	1101	1	38,41,41	0.62	2 (5%)	51,59,59	1.36	1 (1%)
2	26X	A	1101	1	38,41,41	0.20	0	51,59,59	1.07	1 (1%)

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	26X	B	1101	1	-	3/26/40/40	0/6/6/6
2	26X	A	1101	1	-	2/26/40/40	0/6/6/6

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1101	26X	C19-N4	-2.91	1.44	1.48
2	B	1101	26X	C4-N1	-2.06	1.43	1.47

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1101	26X	C5-C4-N1	-9.29	111.42	118.18
2	A	1101	26X	C5-C4-N1	-7.11	113.00	118.18

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1101	26X	C3-C4-N1-N2
2	B	1101	26X	C-C1-C2-N
2	B	1101	26X	O-C2-N-C5
2	A	1101	26X	C20-C19-N4-C18
2	B	1101	26X	C-C1-C2-O

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1101	26X	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

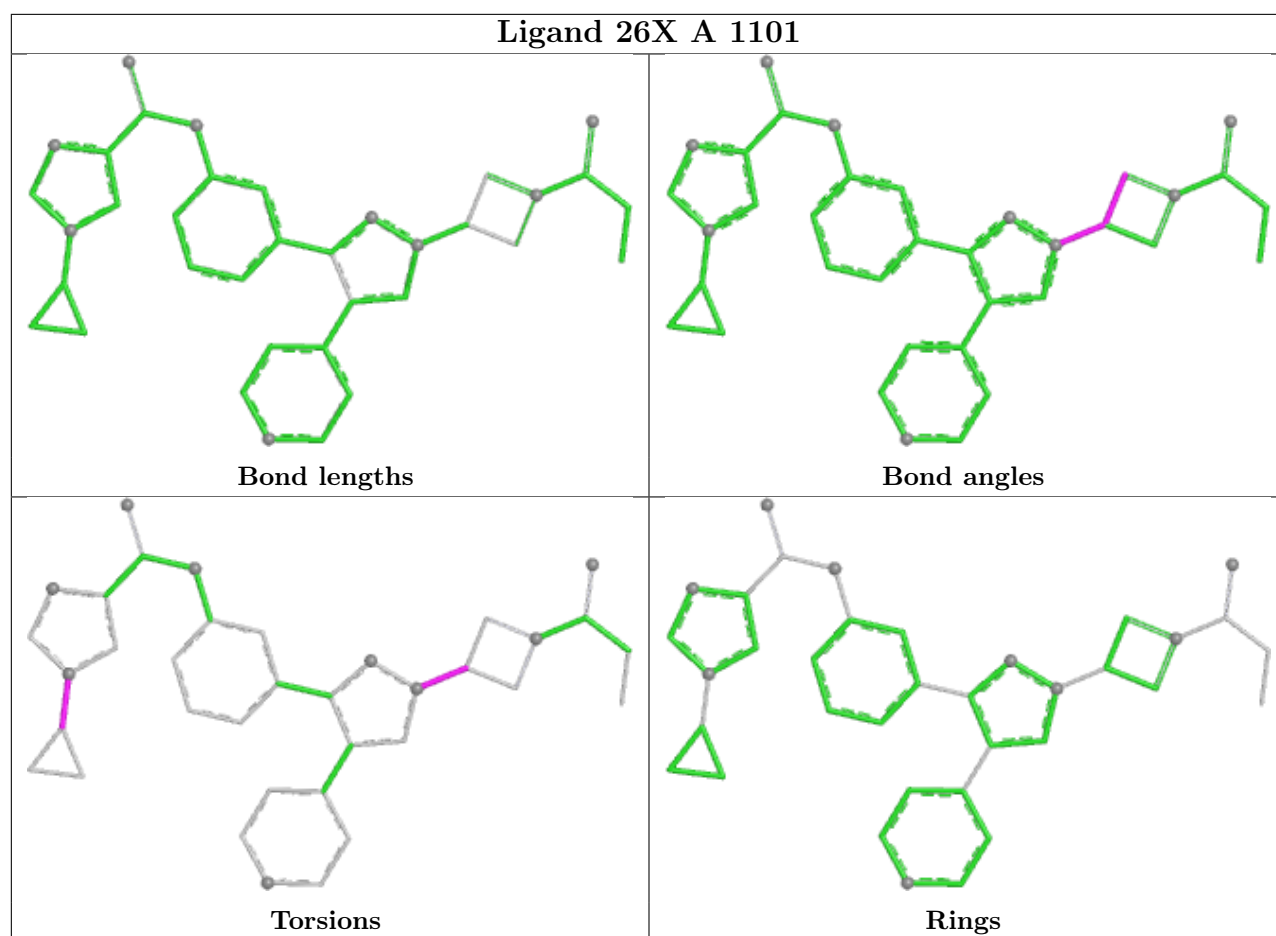
Ligand 26X B 1101

Bond lengths

Bond angles

Torsions

Rings



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, 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	271/332 (81%)	0.32	13 (4%) 35 36	25, 43, 63, 73	0
1	B	274/332 (82%)	0.23	10 (3%) 46 48	25, 42, 65, 82	0
1	D	18/332 (5%)	1.45	7 (38%) 1 0	37, 59, 79, 79	0
All	All	563/996 (56%)	0.31	30 (5%) 32 33	25, 43, 66, 82	0

All (30) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	764	TYR	4.7
1	D	17	PRO	4.1
1	D	14	TYR	3.8
1	B	866	GLY	3.8
1	D	0	MET	3.5
1	A	866	GLY	3.5
1	B	735	GLY	3.5
1	D	12	ASP	3.0
1	D	16	ILE	2.9
1	A	722	ALA	2.8
1	A	816	TYR	2.7
1	B	737	LYS	2.7
1	A	717	VAL	2.6
1	B	736	GLU	2.6
1	B	732	ILE	2.6
1	B	987	ASP	2.5
1	A	794	GLN	2.5
1	A	865	LEU	2.4
1	D	15	LEU	2.3
1	A	879	VAL	2.3
1	A	705	ARG	2.2
1	B	811	ASN	2.2
1	A	779	ARG	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	810	ASP	2.2
1	B	758	GLU	2.2
1	D	13	GLU	2.2
1	A	763	ALA	2.1
1	A	735	GLY	2.0
1	B	925	GLU	2.0
1	B	810	ASP	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 oligosaccharides in this entry.

6.4 Ligands [i](#)

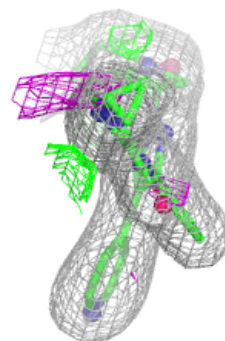
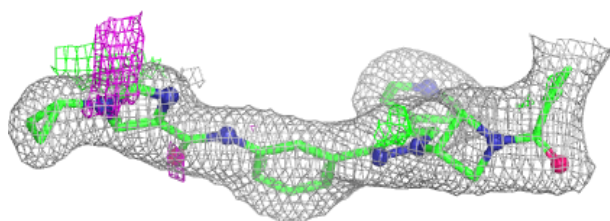
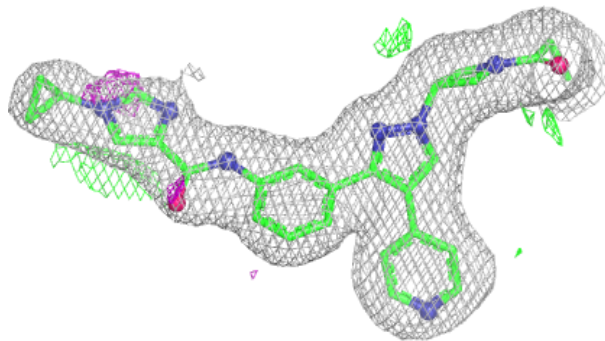
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists 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
2	26X	A	1101	36/36	0.95	0.09	33,36,47,47	0
2	26X	B	1101	36/36	0.95	0.08	36,40,44,44	0

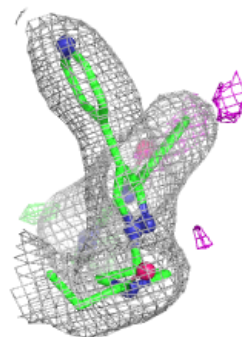
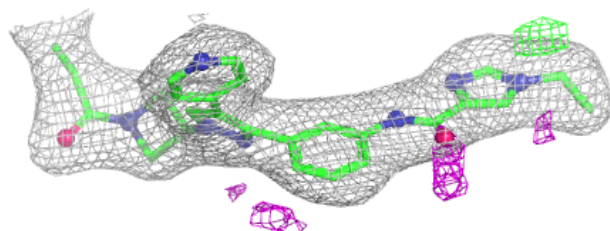
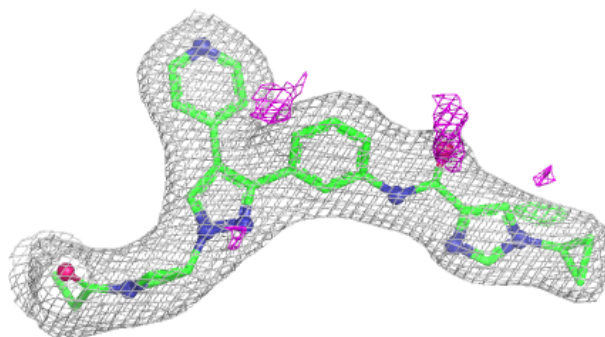
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around 26X A 1101:

$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

**Electron density around 26X B 1101:**

$2mF_o-DF_c$ (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.