



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 9, 2026 – 03:26 AM UTC

PDB ID : 8HV1 / pdb_00008hv1
Title : Crystal structure of EGFR_DMx in complex with covalently bound fragment 1
Authors : Dokurno, P.
Deposited on : 2022-12-26
Resolution : 2.40 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

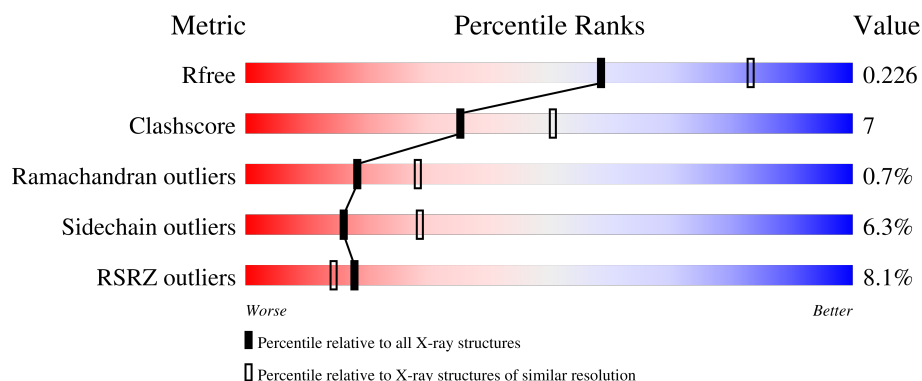
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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	4912 (2.40-2.40)
Clashscore	190562	5391 (2.40-2.40)
Ramachandran outliers	187476	5320 (2.40-2.40)
Sidechain outliers	187428	5321 (2.40-2.40)
RSRZ outliers	180081	4916 (2.40-2.40)

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	331	

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 2518 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.

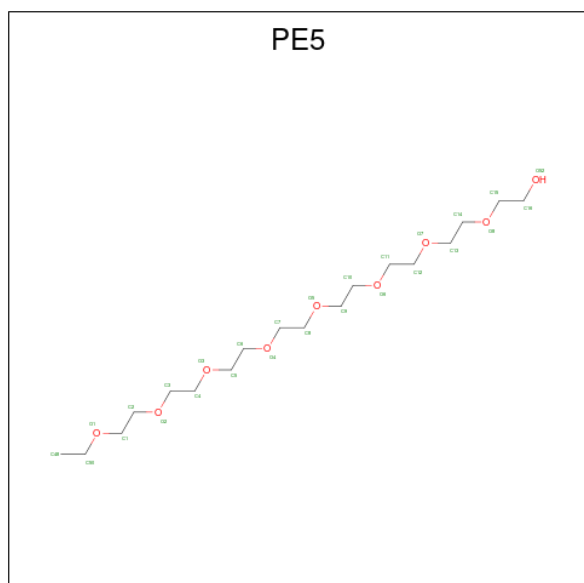
- Molecule 1 is a protein called Epidermal growth factor receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	309	2406	1541	408	440	17	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

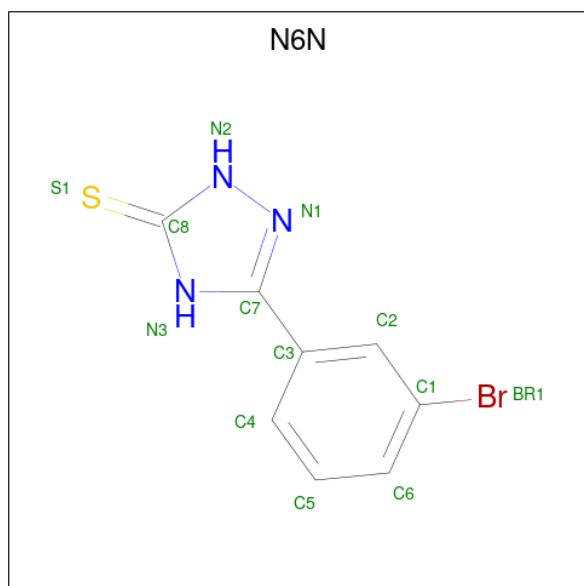
Chain	Residue	Modelled	Actual	Comment	Reference
A	694	GLY	-	expression tag	UNP P00533
A	790	MET	THR	variant	UNP P00533
A	858	ARG	LEU	variant	UNP P00533
A	865	ALA	GLU	engineered mutation	UNP P00533
A	866	ALA	GLU	engineered mutation	UNP P00533
A	867	ALA	LYS	engineered mutation	UNP P00533
A	1023	ASN	-	expression tag	UNP P00533
A	1024	SER	-	expression tag	UNP P00533

- Molecule 2 is 3,6,9,12,15,18,21,24-OCTAOXAHEXACOSAN-1-OL (CCD ID: PE5) (formula: $C_{18}H_{38}O_9$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			26	18	8		

- Molecule 3 is 3-(3-bromophenyl)-1,4-dihydro-1,2,4-triazole-5-thione (CCD ID: N6N) (formula: C₈H₆BrN₃S) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	Br	C	N	S	0	0
			13	1	8	3	1		

- Molecule 4 is SODIUM ION (CCD ID: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Na	0	0
			1	1		

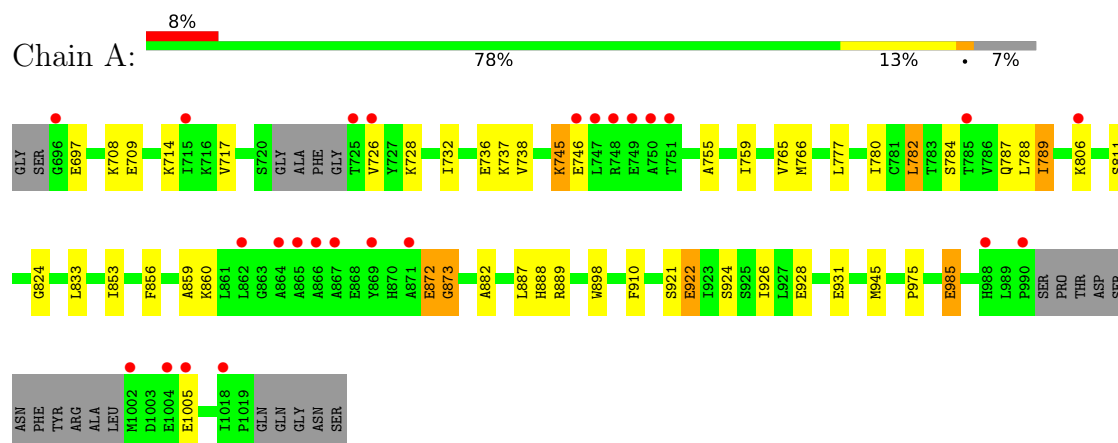
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	72	Total	O	0	0
			72	72		

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: Epidermal growth factor receptor



4 Data and refinement statistics

Property	Value	Source
Space group	I 2 3	Depositor
Cell constants a, b, c, α , β , γ	144.91Å 144.91Å 144.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.00 – 2.40 25.00 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.9 (25.00-2.40) 99.8 (25.00-2.40)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.17 (at 2.41Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.189 , 0.219 0.194 , 0.226	Depositor DCC
R_{free} test set	995 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	59.7	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 49.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.023 for -l,-k,-h	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	2518	wwPDB-VP
Average B, all atoms (Å ²)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.13% 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: N6N, NA, PE5

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	1.04	0/2456	1.38	3/3334 (0.1%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	856	PHE	CA-C-N	6.00	126.75	120.03
1	A	856	PHE	C-N-CA	6.00	126.75	120.03
1	A	873	GLY	CA-C-O	-5.25	118.04	122.29

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	2406	0	2387	35	0
2	A	26	0	35	1	0
3	A	13	0	0	1	0
4	A	1	0	0	0	0
5	A	72	0	0	1	0
All	All	2518	0	2422	36	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 7.

All (36) 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)
2:A:1101:PE5:C16	5:A:1264:HOH:O	1.67	1.33
1:A:806:LYS:HG2	1:A:910:PHE:HB3	1.46	0.97
1:A:872:GLU:HA	1:A:889:ARG:NH2	1.88	0.88
1:A:745:LYS:NZ	3:A:1102:N6N:BR1	2.75	0.74
1:A:872:GLU:HA	1:A:889:ARG:CZ	2.22	0.68
1:A:789:ILE:HD12	1:A:789:ILE:N	2.09	0.67
1:A:746:GLU:HA	1:A:787:GLN:HG2	1.78	0.63
1:A:788:LEU:C	1:A:789:ILE:HD12	2.22	0.63
1:A:806:LYS:HE2	1:A:910:PHE:O	1.97	0.62
1:A:872:GLU:CA	1:A:889:ARG:NH2	2.62	0.61
1:A:746:GLU:CB	1:A:787:GLN:HG2	2.34	0.58
1:A:985:GLU:H	1:A:985:GLU:CD	2.10	0.57
1:A:780:ILE:HG22	1:A:788:LEU:HD23	1.88	0.56
1:A:872:GLU:C	1:A:889:ARG:HH21	2.15	0.55
1:A:746:GLU:CA	1:A:787:GLN:HG2	2.38	0.53
1:A:872:GLU:C	1:A:889:ARG:NH2	2.67	0.52
1:A:806:LYS:HG2	1:A:910:PHE:CB	2.31	0.51
1:A:873:GLY:N	1:A:889:ARG:HH21	2.08	0.51
1:A:924:SER:O	1:A:928:GLU:HG3	2.12	0.50
1:A:926:ILE:HG23	1:A:931:GLU:HB3	1.93	0.49
1:A:789:ILE:N	1:A:789:ILE:CD1	2.74	0.49
1:A:887:LEU:HB3	1:A:888:HIS:CD2	2.48	0.49
1:A:811:SER:OG	1:A:975:PRO:HB2	2.12	0.49
1:A:746:GLU:CB	1:A:787:GLN:HE21	2.28	0.47
1:A:755:ALA:O	1:A:759:ILE:HG12	2.14	0.47
1:A:806:LYS:HE2	1:A:910:PHE:CB	2.45	0.46
1:A:766:MET:HB3	1:A:777:LEU:HB2	1.98	0.45
1:A:945:MET:HE2	1:A:945:MET:HB2	1.84	0.45
1:A:833:LEU:HD21	1:A:859:ALA:HB1	1.98	0.44
1:A:765:VAL:HG13	1:A:833:LEU:HD11	1.99	0.44
1:A:882:ALA:HA	1:A:898:TRP:CD2	2.54	0.43
1:A:782:LEU:HD12	1:A:782:LEU:HA	1.84	0.41
1:A:736:GLU:HB3	1:A:738:VAL:HG22	2.02	0.41
1:A:824:GLY:HA3	1:A:853:ILE:HD12	2.03	0.41
1:A:922:GLU:H	1:A:922:GLU:HG2	1.80	0.40
1:A:860:LYS:HA	1:A:860:LYS:HD3	1.80	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	303/331 (92%)	287 (95%)	14 (5%)	2 (1%)	18	28

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1005	GLU
1	A	714	LYS

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	255/287 (89%)	239 (94%)	16 (6%)	16	29

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	697	GLU
1	A	708	LYS
1	A	709	GLU
1	A	717	VAL
1	A	726	VAL
1	A	728	LYS
1	A	732	ILE
1	A	737	LYS
1	A	745	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	782	LEU
1	A	784	SER
1	A	789	ILE
1	A	872	GLU
1	A	921	SER
1	A	922	GLU
1	A	985	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	773	HIS
1	A	787	GLN
1	A	888	HIS
1	A	976	GLN
1	A	982	GLN

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](#)

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 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	PE5	A	1101	-	25,25,26	0.30	0	24,24,25	0.19	0
3	N6N	A	1102	1	13,14,14	0.61	0	14,19,19	0.69	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	PE5	A	1101	-	-	10/23/23/24	-
3	N6N	A	1102	1	-	0/4/4/4	0/2/2/2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (10) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1101	PE5	C13-C14-O8-C15
2	A	1101	PE5	O7-C13-C14-O8
2	A	1101	PE5	O6-C10-C9-O5
2	A	1101	PE5	O3-C5-C6-O4
2	A	1101	PE5	O1-C1-C2-O2
2	A	1101	PE5	C16-C15-O8-C14
2	A	1101	PE5	C7-C8-O5-C9
2	A	1101	PE5	O4-C7-C8-O5
2	A	1101	PE5	C8-C7-O4-C6
2	A	1101	PE5	C10-C9-O5-C8

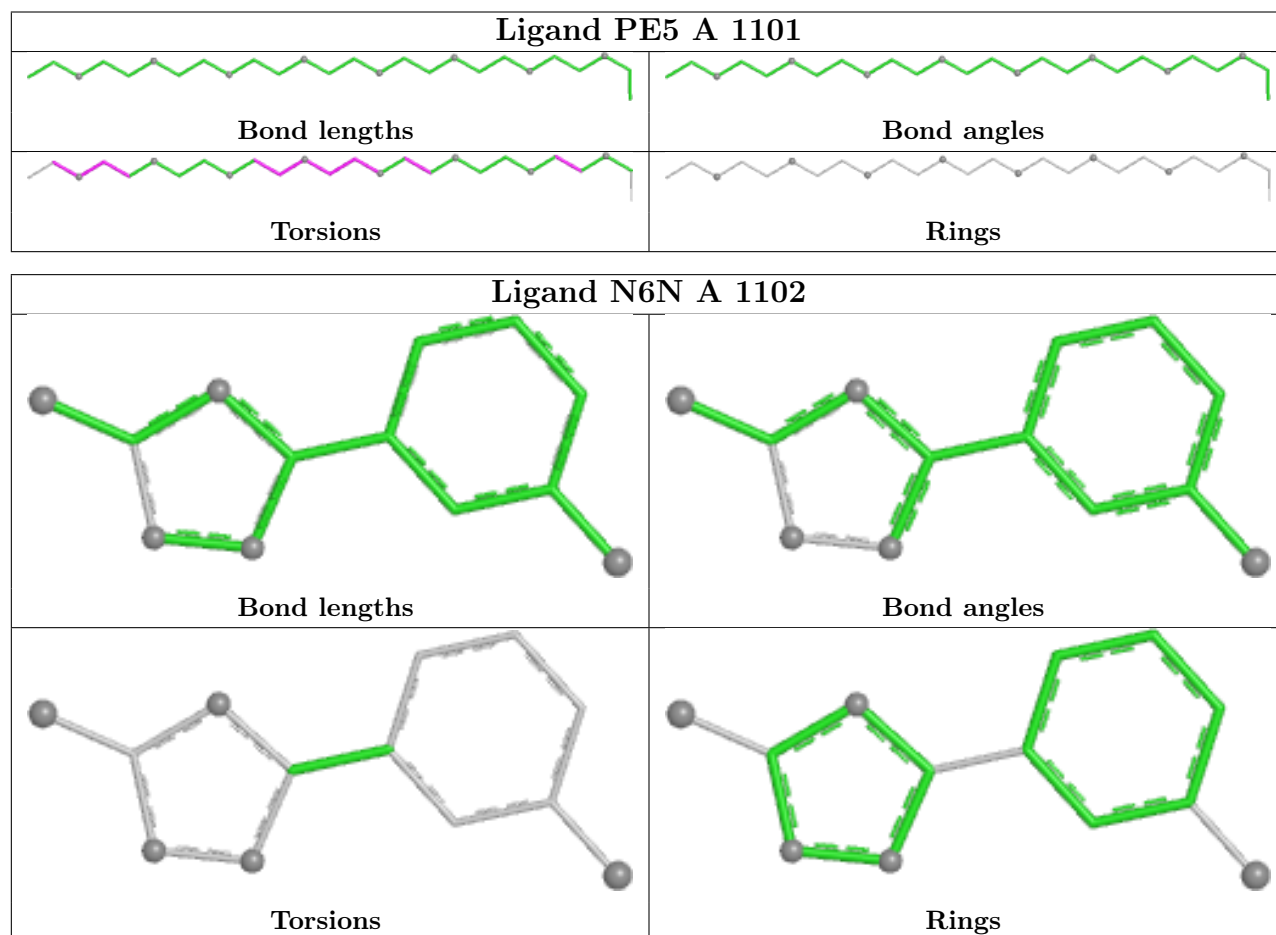
There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1101	PE5	1	0
3	A	1102	N6N	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, 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	309/331 (93%)	0.07	25 (8%) 18 14	40, 58, 126, 164	0

All (25) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	726	VAL	4.7
1	A	749	GLU	4.3
1	A	748	ARG	4.3
1	A	1018	ILE	4.2
1	A	715	ILE	4.1
1	A	696	GLY	3.9
1	A	867	ALA	3.9
1	A	747	LEU	3.8
1	A	750	ALA	3.7
1	A	785	THR	3.4
1	A	746	GLU	3.3
1	A	864	ALA	3.0
1	A	806	LYS	2.9
1	A	751	THR	2.9
1	A	988	HIS	2.7
1	A	1002	MET	2.5
1	A	869	TYR	2.4
1	A	1005	GLU	2.4
1	A	990	PRO	2.3
1	A	866	ALA	2.3
1	A	725	THR	2.2
1	A	1004	GLU	2.2
1	A	862	LEU	2.2
1	A	865	ALA	2.0
1	A	871	ALA	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](#)

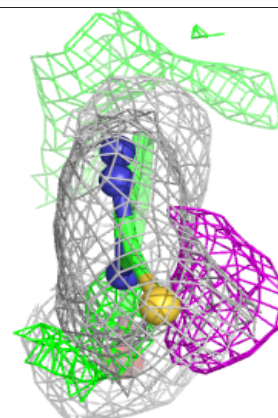
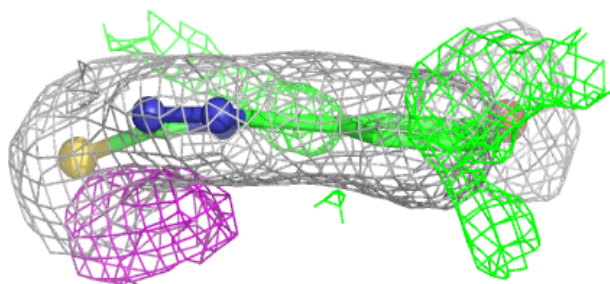
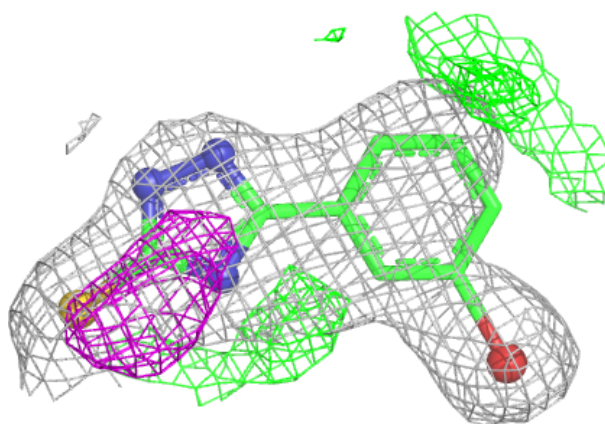
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists 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
3	N6N	A	1102	13/13	0.84	0.16	65,79,108,109	1
2	PE5	A	1101	26/27	0.87	0.18	49,66,139,147	26
4	NA	A	1103	1/1	0.96	0.24	77,77,77,77	0

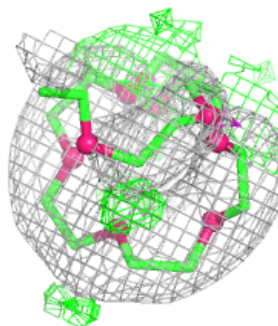
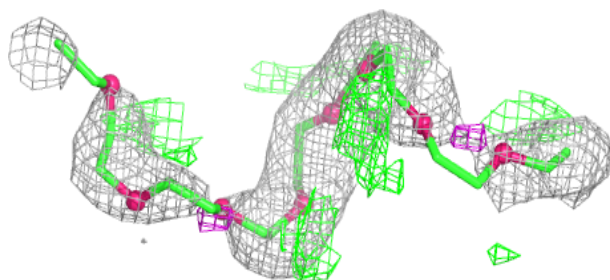
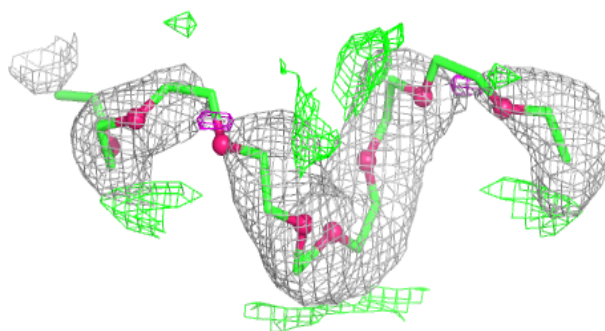
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around N6N A 1102:

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

**Electron density around PE5 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)



6.5 Other polymers [i](#)

There are no such residues in this entry.