WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Dec 11, 2024 – 06:02 PM EST

PDB ID	:	8UGW
Title	:	Computational design of highly signaling active membrane receptors through de novo solvent-mediated allosteric networks
Authors	:	Wang, J.; Chen, K.Y.; Lai, J.K.; Russell, A.M.; Conners, K.; Rutter, M.E.; Condon, B.; Tung, F.; Kodandapani, L.; Chau, B.; Zhao, X.; Benach, J.; Baker, K.; Hembre, E.J.; Barth, P.
Deposited on	:	2023-10-06
Resolution	:	3.90 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 (i) 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 ①) were used in the production of this report:

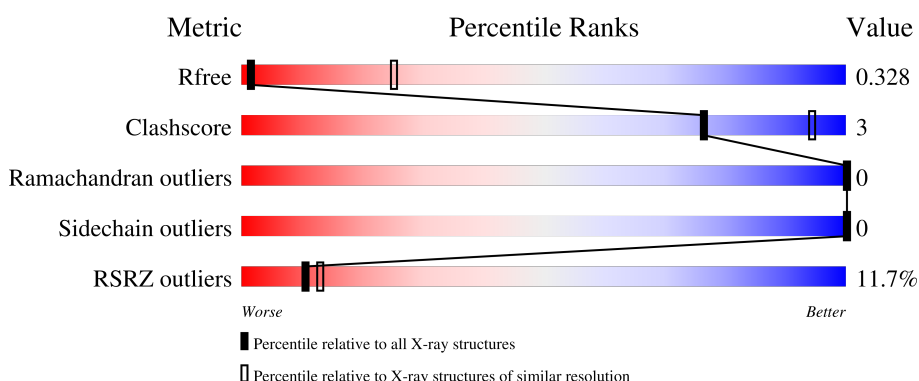|		
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.21
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

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 3.90 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	Percentile Ranks	Value
Rfree		0.328
Clashscore		3
Ramachandran outliers		0
Sidechain outliers		0
RSRZ outliers		11.7%

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	1157 (4.10-3.70)
Clashscore	180529	1219 (4.10-3.70)
Ramachandran outliers	177936	1177 (4.10-3.70)
Sidechain outliers	177891	1169 (4.10-3.70)
RSRZ outliers	164620	1157 (4.10-3.70)

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 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 <=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	472	10% 77% • 18%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 3056 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 Endolysin,Adenosine receptor A2a.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	385	Total 3020	C 1973	N 511	O 510	S 26	0	0	0

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1001	MET	-	initiating methionine	UNP P00720
A	1002	LYS	-	expression tag	UNP P00720
A	1003	THR	-	expression tag	UNP P00720
A	1004	ILE	-	expression tag	UNP P00720
A	1005	ILE	-	expression tag	UNP P00720
A	1006	ALA	-	expression tag	UNP P00720
A	1007	LEU	-	expression tag	UNP P00720
A	1008	SER	-	expression tag	UNP P00720
A	1009	TYR	-	expression tag	UNP P00720
A	1010	ILE	-	expression tag	UNP P00720
A	1011	PHE	-	expression tag	UNP P00720
A	1012	CYS	-	expression tag	UNP P00720
A	1013	LEU	-	expression tag	UNP P00720
A	1014	VAL	-	expression tag	UNP P00720
A	1015	PHE	-	expression tag	UNP P00720
A	1016	ALA	-	expression tag	UNP P00720
A	1017	ASP	-	expression tag	UNP P00720
A	1018	TYR	-	expression tag	UNP P00720
A	1019	LYS	-	expression tag	UNP P00720
A	1020	ASP	-	expression tag	UNP P00720
A	1021	ASP	-	expression tag	UNP P00720
A	1022	ASP	-	expression tag	UNP P00720
A	1023	ASP	-	expression tag	UNP P00720
A	1024	GLY	-	expression tag	UNP P00720
A	1025	ALA	-	expression tag	UNP P00720
A	1026	PRO	-	expression tag	UNP P00720
A	1064	ALA	CYS	conflict	UNP P00720

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	1104	ARG	ILE	conflict	UNP P00720
A	1132	SER	-	linker	UNP P00720
A	1133	GLY	-	linker	UNP P00720
A	1134	GLY	-	linker	UNP P00720
A	1135	GLY	-	linker	UNP P00720
A	1136	GLY	-	linker	UNP P00720
A	1137	ALA	-	linker	UNP P00720
A	1138	MET	-	linker	UNP P00720
A	1139	ASP	-	linker	UNP P00720
A	1140	ILE	-	linker	UNP P00720
A	1141	PHE	-	linker	UNP P00720
A	1142	GLU	-	linker	UNP P00720
A	1143	MET	-	linker	UNP P00720
A	1144	LEU	-	linker	UNP P00720
A	1145	ARG	-	linker	UNP P00720
A	1146	ILE	-	linker	UNP P00720
A	1147	ASP	-	linker	UNP P00720
A	1148	GLU	-	linker	UNP P00720
A	1149	GLY	-	linker	UNP P00720
A	?	ALA	-	linker	UNP P00720
A	?	PRO	-	linker	UNP P00720
A	48	ALA	LEU	conflict	UNP P29274
A	91	THR	SER	conflict	UNP P29274
A	95	MET	LEU	conflict	UNP P29274
A	194	MET	LEU	conflict	UNP P29274
A	238	TYR	ILE	conflict	UNP P29274
A	239	LEU	VAL	conflict	UNP P29274
A	243	LEU	ALA	conflict	UNP P29274
A	317	GLY	ALA	conflict	UNP P29274
A	318	SER	GLY	conflict	UNP P29274
A	319	LEU	THR	conflict	UNP P29274
A	320	VAL	SER	conflict	UNP P29274
A	321	PRO	ALA	conflict	UNP P29274

- Molecule 2 is 2-[P-(2-CARBOXYETHYL)PHENYLETHYL-AMINO]-5'-N-ETHYLCA
  RBOXAMIDO ADENOSINE (three-letter code: NGI) (formula: $C_{23}H_{29}N_7O_6$) (labeled as
  "Ligand of Interest" by depositor).

NGI
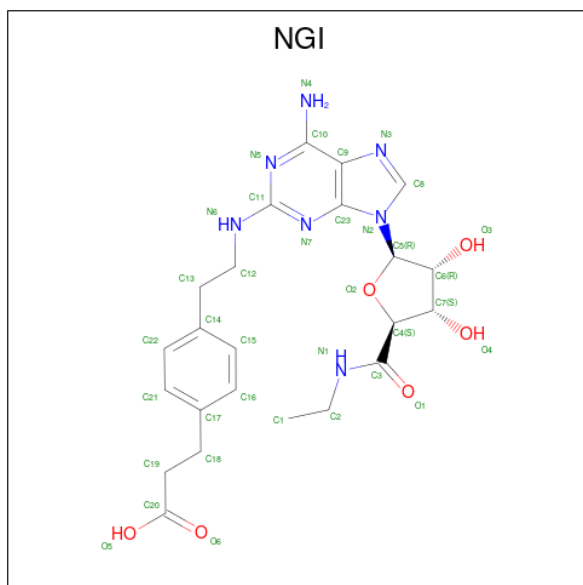
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	36	23	7	6	0	0

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: Endolysin,Adenosine receptor A2a

4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	60.18Å 73.67Å 149.85Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	66.00 – 3.90 66.00 – 3.90	Depositor EDS
% Data completeness (in resolution range)	99.3 (66.00-3.90) 99.0 (66.00-3.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.13	Depositor
$< I/\sigma(I) >$ [1]	2.10 (at 3.89Å)	Xtriage
Refinement program	REFMAC 5.8	Depositor
R, R_{free}	0.300 , 0.330 0.307 , 0.328	Depositor DCC
R_{free} test set	308 reflections (4.80%)	wwPDB-VP
Wilson B-factor (Å^2)	101.0	Xtriage
Anisotropy	0.756	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 62.4	EDS
L-test for twinning[2]	$< |L| > = 0.45, < L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.80	EDS
Total number of atoms	3056	wwPDB-VP
Average B, all atoms (Å^2)	117.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.68% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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: NGI

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	0.53	0/3086	0.64	1/4194 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	197	TYR	CB-CG-CD1	-5.12	117.93	121.00

There are no chirality outliers.

There are no planarity outliers.

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	3020	0	3075	17	0
2	A	36	0	28	0	0
All	All	3056	0	3103	17	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.

The worst 5 of 17 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:1131:LEU:O	1:A:1131:LEU:HD23	1.79	0.83
1:A:1131:LEU:HD23	1:A:1131:LEU:C	2.17	0.64
1:A:283:VAL:O	1:A:283:VAL:HG12	2.02	0.59
1:A:284:ASN:HB2	1:A:285:PRO:HD3	1.87	0.55
1:A:108:ILE:O	1:A:108:ILE:HG22	2.06	0.55

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	375/472 (79%)	363 (97%)	12 (3%)	0	100 100

There are no Ramachandran outliers to report.

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	319/392 (81%)	319 (100%)	0	100 100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 ⓘ

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	$\#|Z| > 2$
2	NGI	A	1601	-	35,39,39	1.81	9 (25%)	42,55,55	1.43	6 (14%)

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	NGI	A	1601	-	-	4/18/38/38	0/4/4/4

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1601	NGI	O6-C20	4.53	1.36	1.22
2	A	1601	NGI	C11-N6	3.76	1.40	1.34
2	A	1601	NGI	C10-N5	3.15	1.39	1.33

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1601	NGI	C23-N7	2.97	1.40	1.35
2	A	1601	NGI	C11-N7	2.65	1.42	1.34

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1601	NGI	O5-C20-C19	4.14	127.07	114.00
2	A	1601	NGI	O5-C20-O6	-4.05	112.93	123.33
2	A	1601	NGI	C4-C3-N1	3.58	119.91	115.47
2	A	1601	NGI	C5-O2-C4	2.58	113.59	109.11
2	A	1601	NGI	C23-C9-N3	2.24	111.70	109.34

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1601	NGI	N1-C3-C4-O2
2	A	1601	NGI	O1-C3-C4-O2
2	A	1601	NGI	N6-C12-C13-C14
2	A	1601	NGI	C18-C19-C20-O5

There are no ring outliers.

No monomer is involved in short contacts.

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 then 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.

Ligand NGI A 1601

Bond lengths

Bond angles

Torsions

Rings
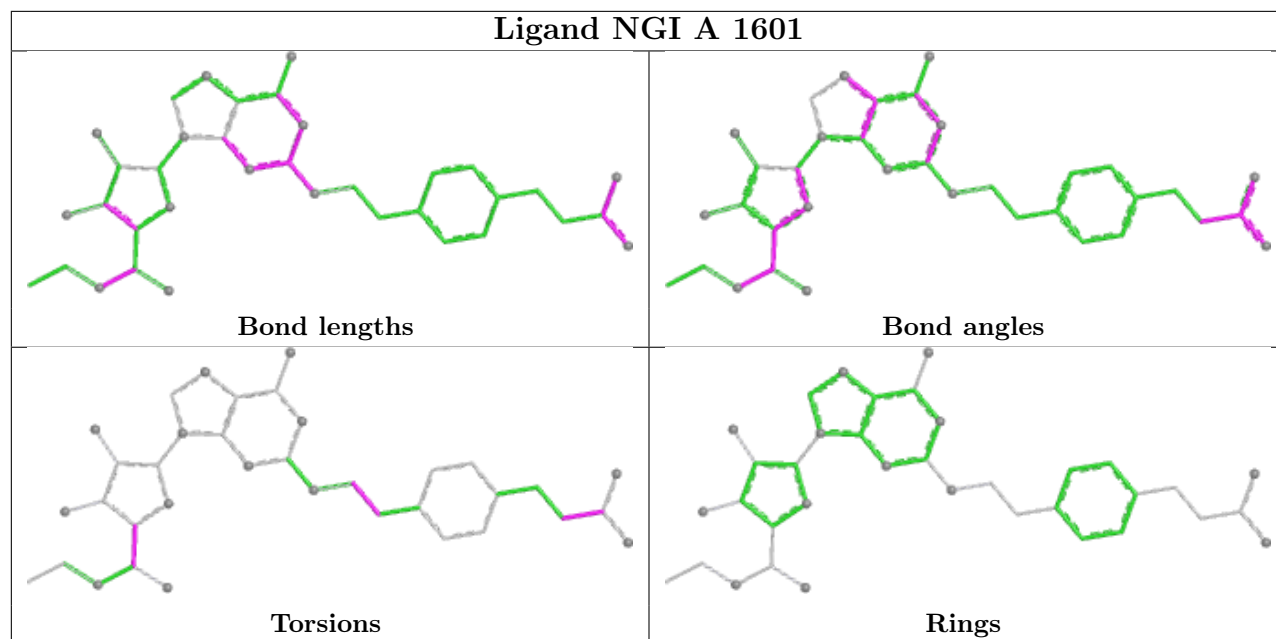
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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	385/472 (81%)	0.71	45 (11%)	10	13	68, 107, 182, 282	0

The worst 5 of 45 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	222	ARG	5.0
1	A	1041	ALA	4.7
1	A	1045	ILE	4.5
1	A	1080	GLY	4.2
1	A	221	ALA	4.1

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

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

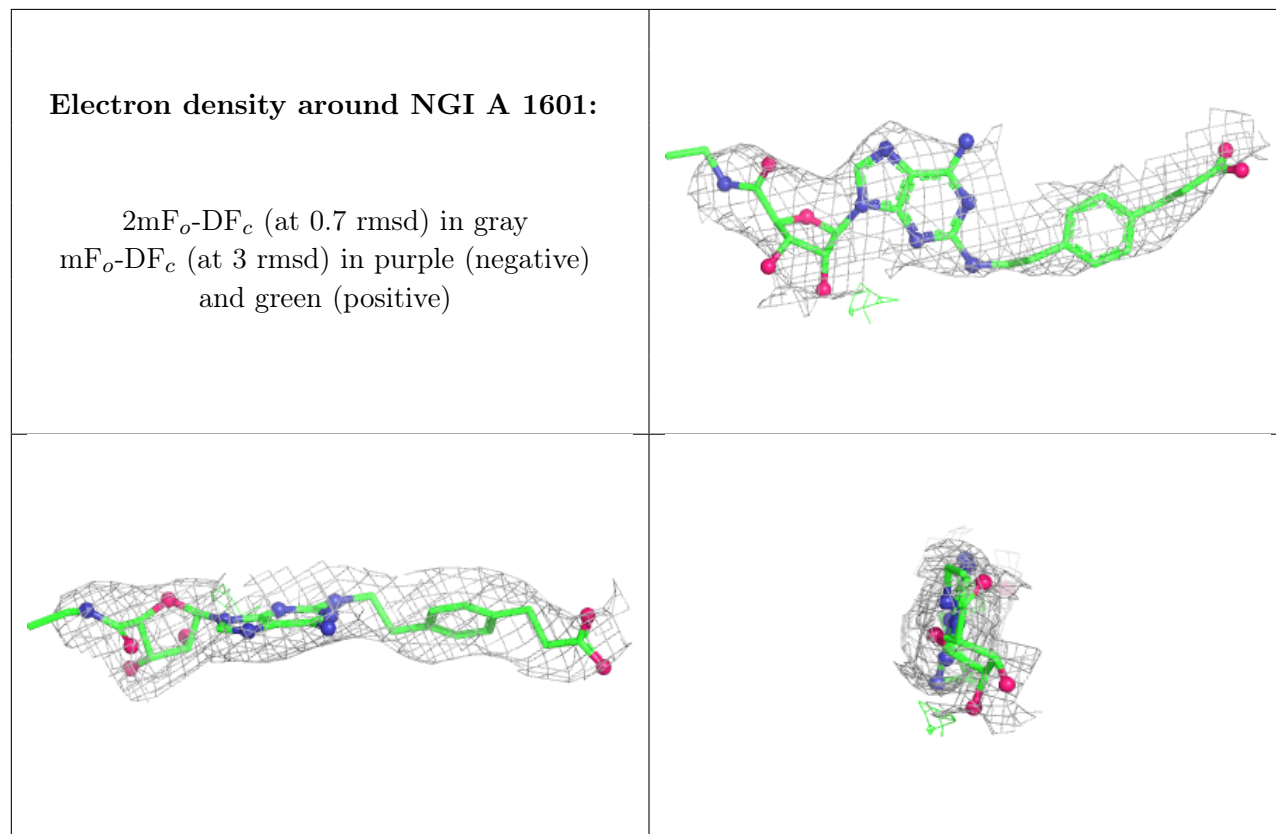## 6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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
2	NGI	A	1601	36/36	0.79	0.17	80,113,130,143	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 NGI A 1601:

$2mF_o$-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.