



Full wwPDB X-ray Structure Validation Report ⓘ

Aug 9, 2020 – 10:53 AM BST

PDB ID : 6WRK
Title : Crystal structure of 3rd-generation Mj 3-nitro-tyrosine tRNA synthetase ("A7") bound to 3-nitro-tyrosine
Authors : Beyer, J.N.; Hosseinzadeh, P.; Karplus, P.A.; Mehl, R.A.; Cooley, R.B.
Deposited on : 2020-04-29
Resolution : 1.95 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.13.1
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) : 2.13.1

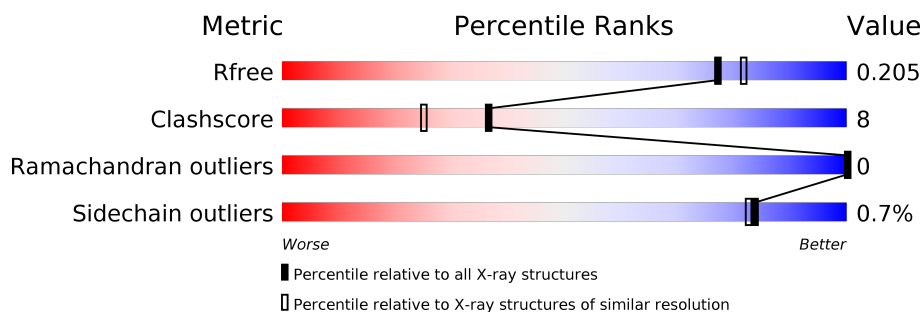
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.95 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	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)

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 on 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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	314	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	GOL	A	401	-	-	X	-
2	GOL	A	402	-	X	-	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 5565 atoms, of which 2731 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 Tyrosine–tRNA ligase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	311	5320	1676	2692	452	486	14	0	26	0

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	32	HIS	TYR	engineered mutation	UNP Q57834
A	70	THR	HIS	engineered mutation	UNP Q57834
A	158	HIS	ASP	engineered mutation	UNP Q57834
A	159	ALA	ILE	engineered mutation	UNP Q57834
A	162	ARG	LEU	engineered mutation	UNP Q57834
A	307	LEU	-	expression tag	UNP Q57834
A	308	GLU	-	expression tag	UNP Q57834
A	309	HIS	-	expression tag	UNP Q57834
A	310	HIS	-	expression tag	UNP Q57834
A	311	HIS	-	expression tag	UNP Q57834
A	312	HIS	-	expression tag	UNP Q57834
A	313	HIS	-	expression tag	UNP Q57834
A	314	HIS	-	expression tag	UNP Q57834

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).

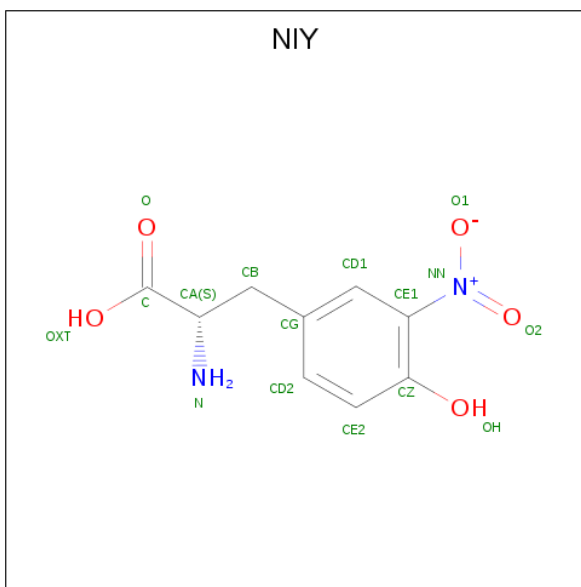


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	H	O	0	0
			14	3	8	3		
2	A	1	Total	C	H	O	0	0
			12	3	6	3		
2	A	1	Total	C	H	O	0	0
			14	3	8	3		

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

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

- Molecule 4 is META-NITRO-TYROSINE (three-letter code: NIY) (formula: C₉H₁₀N₂O₅) (labeled as "Ligand of Interest" by author).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	H	N	O	0	0
			24	9	8	2	5		
4	A	1	Total	C	H	N	O	0	0
			25	9	9	2	5		

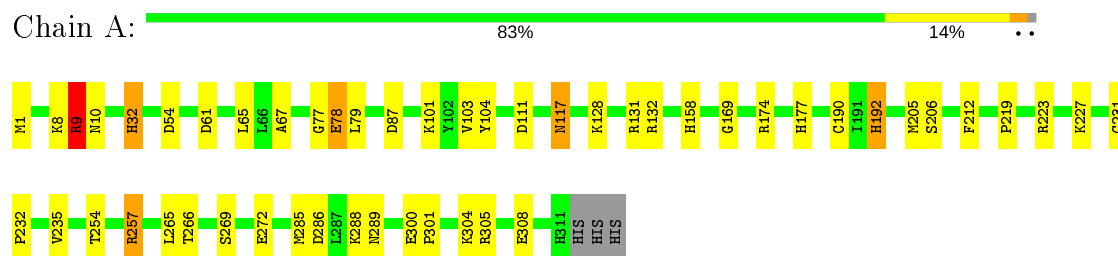
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	155	Total	O	0	0
			155	155		

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. 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. 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: Tyrosine-tRNA ligase



4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	101.43 Å 101.43 Å 71.50 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.36 – 1.95 45.36 – 1.34	Depositor EDS
% Data completeness (in resolution range)	99.9 (45.36-1.95) 80.1 (45.36-1.34)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.93 (at 1.34 Å)	Xtriage
Refinement program	PHENIX 1.17.1-3660	Depositor
R, R_{free}	0.181 , 0.205 0.181 , 0.205	Depositor DCC
R_{free} test set	2000 reflections (2.96%)	wwPDB-VP
Wilson B-factor (Å ²)	23.3	Xtriage
Anisotropy	0.121	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.41 , 56.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	5565	wwPDB-VP
Average B, all atoms (Å ²)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.14% 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: GOL, NIY, NA

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	0.86	13/2766 (0.5%)	0.83	5/3713 (0.1%)

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	3

All (13) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	9[A]	ARG	CZ-NH1	6.82	1.42	1.33
1	A	9[B]	ARG	CZ-NH1	6.82	1.42	1.33
1	A	9[A]	ARG	CZ-NH2	6.60	1.41	1.33
1	A	9[B]	ARG	CZ-NH2	6.60	1.41	1.33
1	A	257	ARG	CZ-NH1	6.19	1.41	1.33
1	A	78[A]	GLU	N-CA	6.06	1.58	1.46
1	A	78[B]	GLU	N-CA	6.06	1.58	1.46
1	A	257	ARG	CZ-NH2	6.04	1.41	1.33
1	A	9[A]	ARG	CA-C	5.50	1.67	1.52
1	A	9[B]	ARG	CA-C	5.50	1.67	1.52
1	A	192	HIS	C-N	-5.35	1.21	1.34
1	A	78[A]	GLU	CA-C	5.00	1.66	1.52
1	A	78[B]	GLU	CA-C	5.00	1.66	1.52

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	286	ASP	CB-CG-OD2	6.62	124.26	118.30
1	A	78[A]	GLU	CA-C-O	6.05	132.80	120.10
1	A	78[B]	GLU	CA-C-O	6.05	132.80	120.10
1	A	77	GLY	C-N-CA	-5.72	107.40	121.70
1	A	111	ASP	CB-CG-OD2	5.04	122.83	118.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	117[B]	ASN	Mainchain
1	A	9[A]	ARG	Mainchain
1	A	9[B]	ARG	Mainchain

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	2628	2692	2606	42	2
2	A	18	22	24	6	0
3	A	1	0	0	0	0
4	A	32	17	18	3	0
5	A	155	0	0	4	1
All	All	2834	2731	2648	43	3

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 8.

All (43) 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:285:MET:HE2	1:A:289[A]:ASN:OD1	1.75	0.86
1:A:87[B]:ASP:OD1	1:A:104:TYR:OH	1.94	0.84
1:A:285:MET:HE1	2:A:401:GOL:O1	1.82	0.80
1:A:304:LYS:O	1:A:308:GLU:HG3	1.85	0.77
1:A:78[A]:GLU:HG3	5:A:570:HOH:O	1.85	0.74

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:78[A]:GLU:CG	5:A:570:HOH:O	2.35	0.73
1:A:205:MET:HE3	1:A:212:PHE:HB2	1.72	0.71
1:A:9[B]:ARG:NH2	1:A:54:ASP:OD2	2.24	0.70
1:A:285:MET:CE	1:A:289[A]:ASN:OD1	2.40	0.69
1:A:285:MET:CE	2:A:401:GOL:O1	2.41	0.69
1:A:78[A]:GLU:CG	1:A:79:LEU:H	2.10	0.64
1:A:257:ARG:NH1	1:A:265:LEU:HD13	2.14	0.63
1:A:65[A]:LEU:HD12	1:A:103:VAL:HG13	1.82	0.62
1:A:78[A]:GLU:HG3	1:A:79:LEU:N	2.13	0.62
1:A:131:ARG:NH2	5:A:502:HOH:O	2.32	0.62
1:A:78[A]:GLU:HG3	1:A:79:LEU:H	1.65	0.61
1:A:206:SER:OG	4:A:405:NIY:OH	2.14	0.60
1:A:223:ARG:HB3	2:A:401:GOL:H2	1.87	0.57
1:A:289[B]:ASN:HA	2:A:401:GOL:H31	1.86	0.57
1:A:254:THR:OG1	1:A:266:THR:HG22	2.06	0.54
1:A:78[A]:GLU:CG	1:A:79:LEU:N	2.69	0.54
1:A:289[A]:ASN:HA	2:A:401:GOL:H31	1.90	0.54
1:A:300:GLU:HB3	1:A:301:PRO:HD3	1.92	0.52
1:A:65[B]:LEU:HD22	1:A:67:ALA:HA	1.93	0.51
1:A:78[A]:GLU:OE2	5:A:501:HOH:O	2.19	0.51
1:A:301:PRO:O	1:A:305:ARG:HG3	2.11	0.50
1:A:206:SER:HG	4:A:405:NIY:HH	1.53	0.49
1:A:61:ASP:OD2	1:A:101:LYS:HE3	2.14	0.47
1:A:177:HIS:HB3	1:A:190:CYS:SG	2.54	0.47
1:A:227:LYS:HB2	2:A:401:GOL:H11	1.96	0.47
4:A:405:NIY:CD2	4:A:405:NIY:C	2.89	0.47
1:A:128:LYS:O	1:A:132:ARG:HG3	2.16	0.46
1:A:232:PRO:HG2	1:A:235:VAL:HB	1.98	0.45
1:A:169:GLY:O	1:A:192:HIS:HA	2.16	0.45
1:A:269:SER:OG	1:A:272:GLU:HG3	2.18	0.44
1:A:65[B]:LEU:HD12	1:A:158:HIS:CE1	2.52	0.43
1:A:78[A]:GLU:CD	1:A:79:LEU:H	2.21	0.43
1:A:219:PRO:O	1:A:223:ARG:HG3	2.20	0.42
1:A:9[B]:ARG:HH22	1:A:54:ASP:CG	2.22	0.41
1:A:32:HIS:CD2	1:A:32:HIS:C	2.93	0.41
1:A:231:CYS:HB3	1:A:288:LYS:HE3	2.03	0.40

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1:MET:O	1:A:174[B]:ARG:HH12[4_565]	1.37	0.23
5:A:622:HOH:O	5:A:623:HOH:O[8_665]	2.11	0.09
1:A:1:MET:O	1:A:174[B]:ARG:NH1[4_565]	2.14	0.06

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	335/314 (107%)	328 (98%)	7 (2%)	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	297/275 (108%)	294 (99%)	3 (1%)	76	74

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	32	HIS
1	A	117[A]	ASN
1	A	117[B]	ASN

Some 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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 6 ligands modelled in this entry, 1 is monoatomic - leaving 5 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	GOL	A	402	-	5,5,5	1.03	1 (20%)	5,5,5	1.52	1 (20%)
4	NIY	A	404	-	12,16,16	1.18	2 (16%)	13,22,22	0.77	0
4	NIY	A	405	-	12,16,16	1.13	2 (16%)	13,22,22	1.07	1 (7%)
2	GOL	A	406	-	5,5,5	0.40	0	5,5,5	0.58	0
2	GOL	A	401	-	5,5,5	1.06	0	5,5,5	1.25	1 (20%)

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	GOL	A	402	-	-	4/4/4/4	-
4	NIY	A	404	-	-	2/6/12/12	0/1/1/1
4	NIY	A	405	-	-	2/6/12/12	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GOL	A	406	-	-	2/4/4/4	-
2	GOL	A	401	-	-	2/4/4/4	-

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	404	NIY	CE1-NN	-3.23	1.39	1.45
4	A	405	NIY	CE1-NN	-2.70	1.40	1.45
4	A	405	NIY	O2-NN	-2.15	1.19	1.22
2	A	402	GOL	O3-C3	2.05	1.51	1.42
4	A	404	NIY	CA-N	2.03	1.51	1.47

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	405	NIY	O2-NN-CE1	3.01	124.19	119.03
2	A	402	GOL	C3-C2-C1	-2.54	101.82	111.70
2	A	401	GOL	C3-C2-C1	-2.50	101.99	111.70

There are no chirality outliers.

All (12) torsion outliers are listed below:

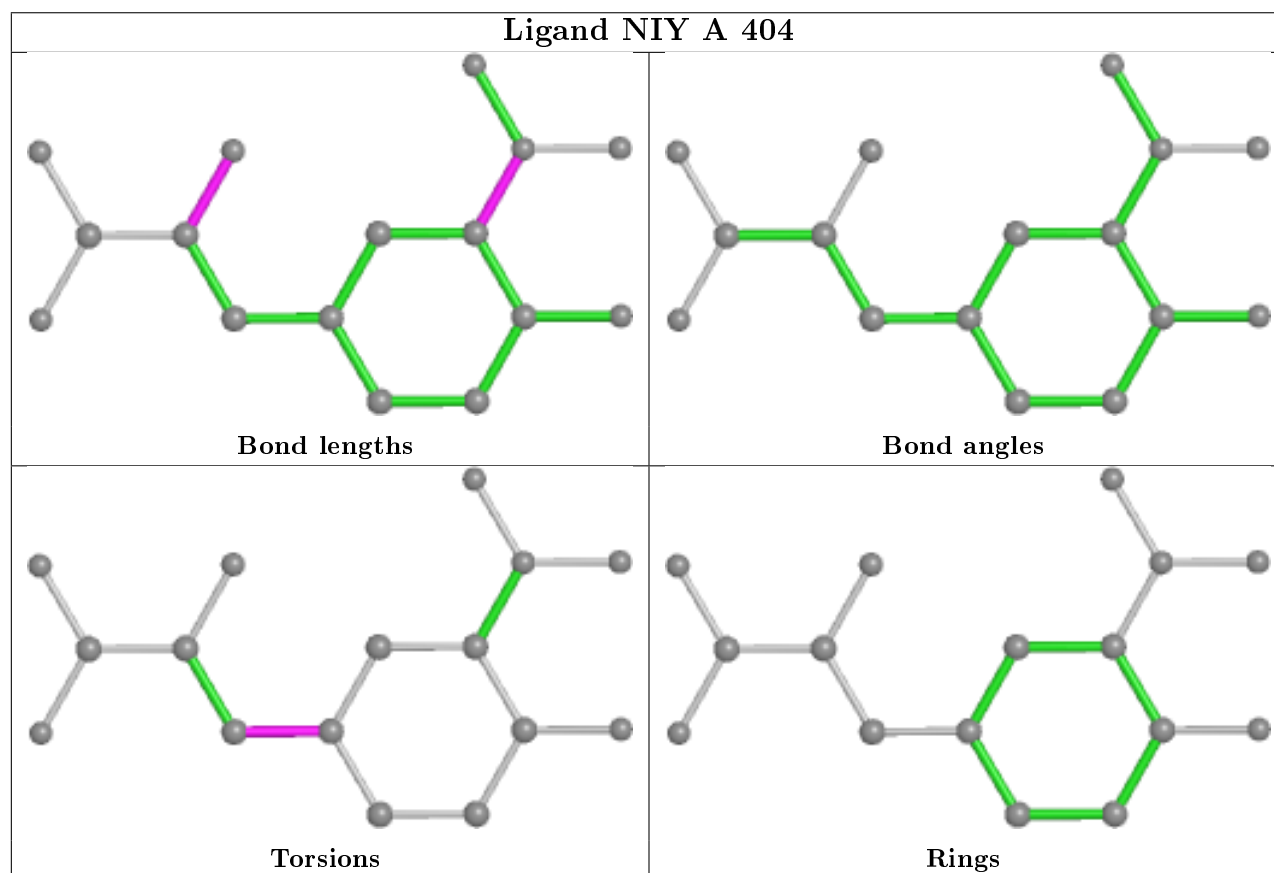
Mol	Chain	Res	Type	Atoms
2	A	402	GOL	O1-C1-C2-C3
2	A	402	GOL	C1-C2-C3-O3
2	A	401	GOL	C1-C2-C3-O3
2	A	402	GOL	O1-C1-C2-O2
2	A	406	GOL	C1-C2-C3-O3
2	A	402	GOL	O2-C2-C3-O3
2	A	401	GOL	O2-C2-C3-O3
4	A	405	NIY	CA-CB-CG-CD1
4	A	405	NIY	CA-CB-CG-CD2
2	A	406	GOL	O2-C2-C3-O3
4	A	404	NIY	CA-CB-CG-CD2
4	A	404	NIY	CA-CB-CG-CD1

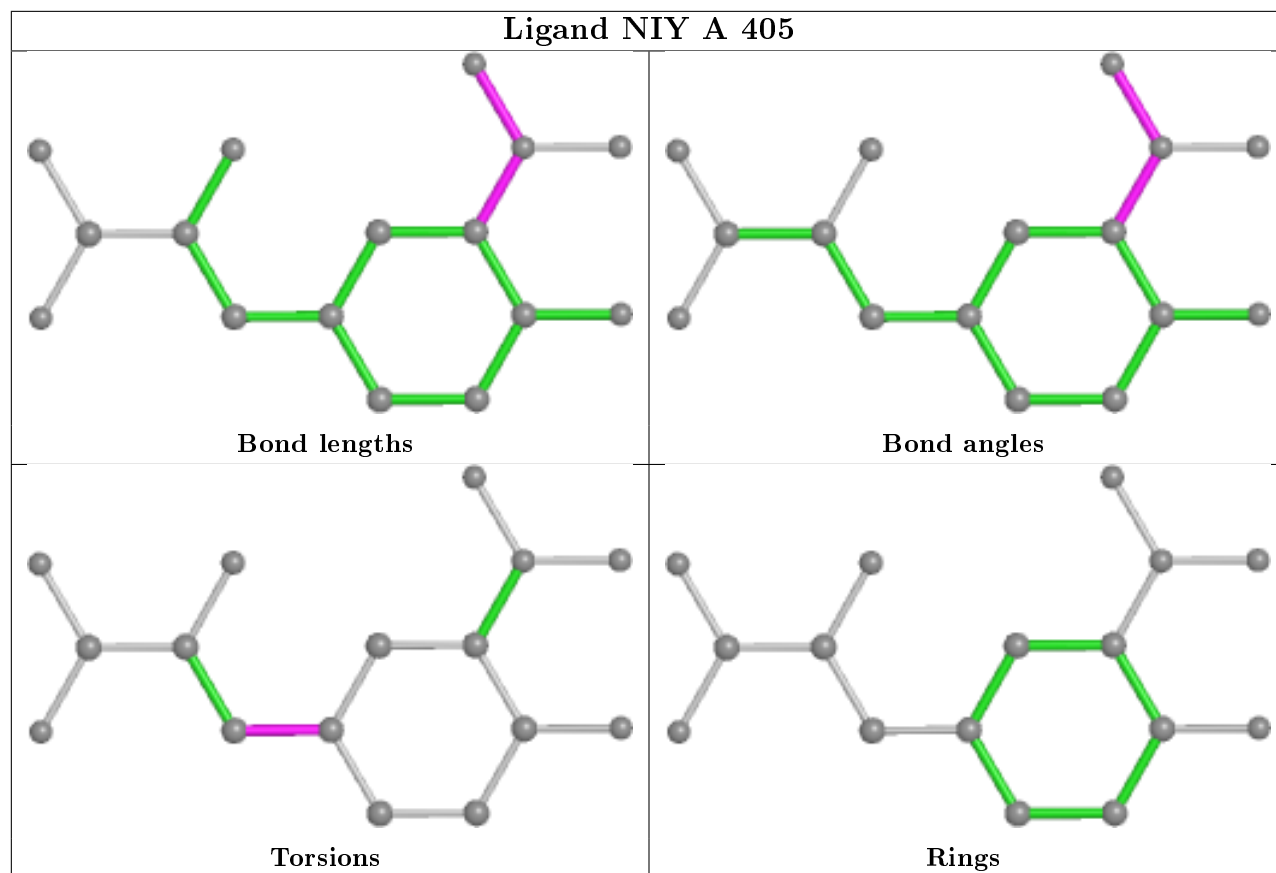
There are no ring outliers.

2 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	405	NIY	3	0
2	A	401	GOL	6	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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

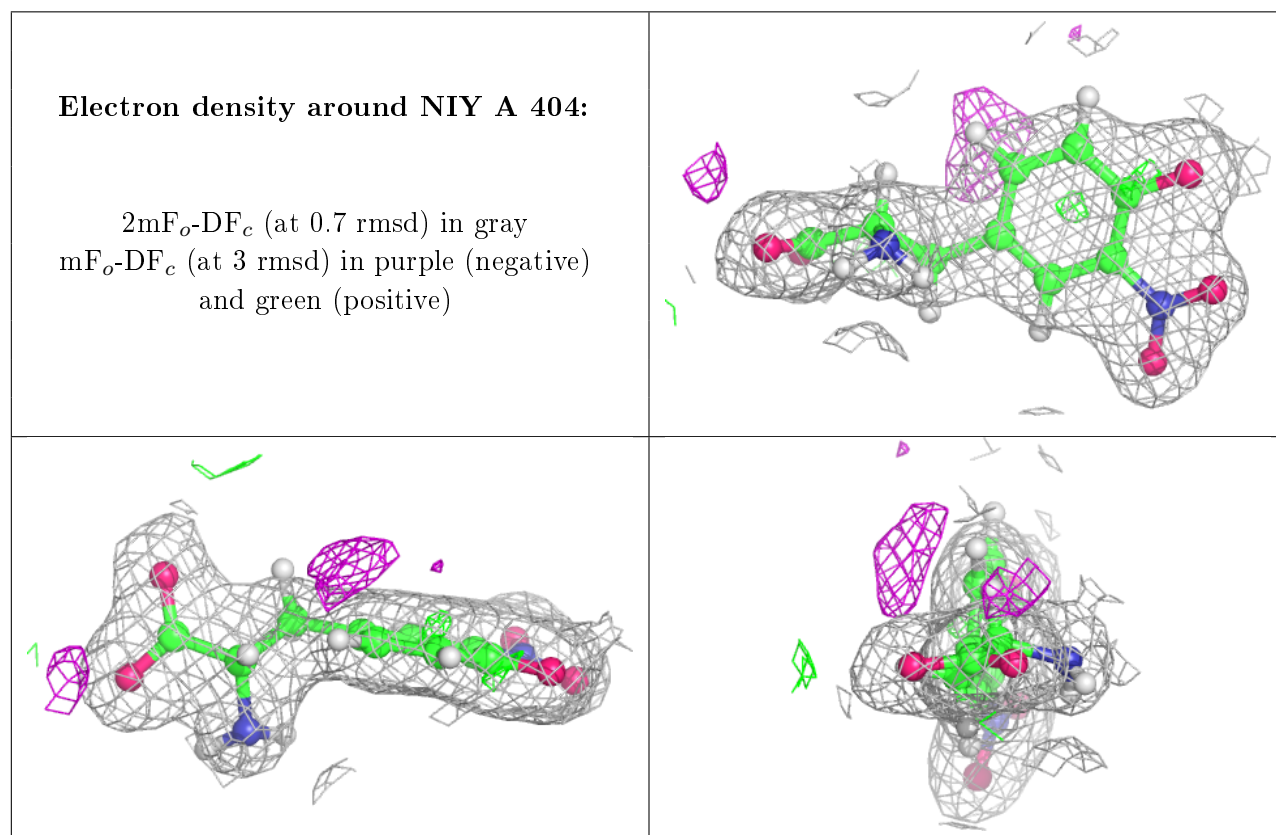
6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

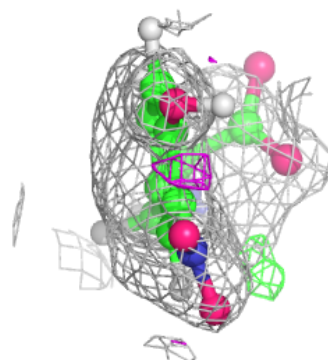
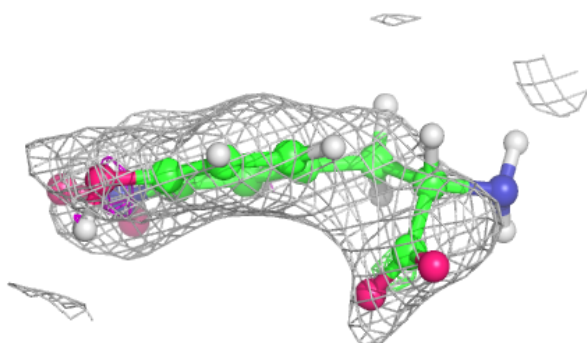
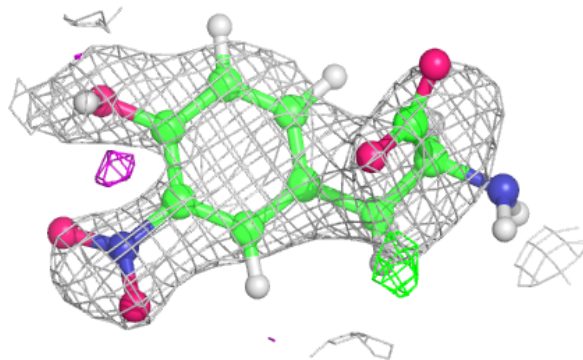
Unable to reproduce the depositors R factor - this section is therefore empty.

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 NIY A 405:

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



6.5 Other polymers [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.