



Full wwPDB X-ray Structure Validation Report ⓘ

Aug 17, 2026 – 05:26 PM EDT

PDB ID : 7IMB / pdb_00007imb
Title : Group deposition for crystallographic fragment screening of SARS-CoV-2 nucleocapsid protein (CTD) – Crystal Structure of SARS-CoV-2 nucleocapsid protein (CTD) in complex with Z235341991 (Nprot-x0372)
Authors : Aschenbrenner, J.C.; Luptak, J.; Balcomb, B.H.; Marples, P.G.; Bellini, D.; Yu, C.W.; Douangamath, A.; Dias, A.; Powell, A.; Fearon, D.; James, L.; von Delft, F.
Deposited on : 2025-08-11
Resolution : 1.66 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	:	9.0.015 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)

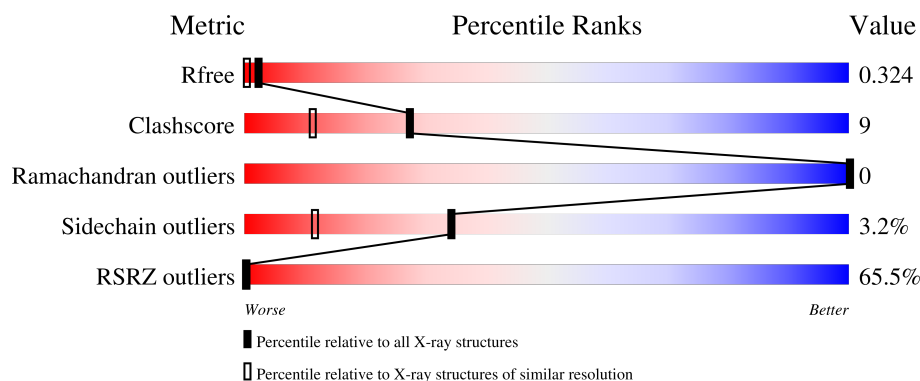
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.66 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	2563 (1.66-1.66)
Clashscore	190562	2662 (1.66-1.66)
Ramachandran outliers	187476	2621 (1.66-1.66)
Sidechain outliers	187428	2621 (1.66-1.66)
RSRZ outliers	180081	2564 (1.66-1.66)

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	117	65% 79% 19% ..

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:

Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.50

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	NY7	A	401	-	-	-	X

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 1068 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.

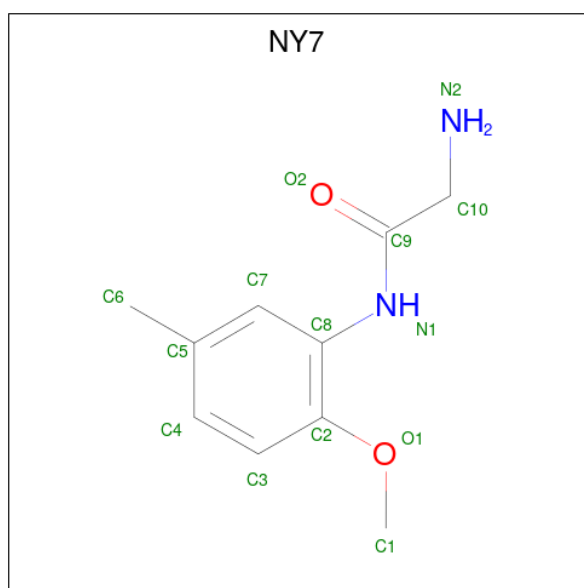
- Molecule 1 is a protein called Nucleoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	116	Total	C	N	O	S	0	1	0
			927	586	165	173	3			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	248	SER	-	expression tag	UNP P0DTC9
A	249	MET	-	expression tag	UNP P0DTC9

- Molecule 2 is N-(2-methoxy-5-methylphenyl)glycinamide (CCD ID: NY7) (formula: $C_{10}H_{14}N_2O_2$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	10	2	2		

- Molecule 3 is DI(HYDROXYETHYL)ETHER (CCD ID: PEG) (formula: $C_4H_{10}O_3$) (labeled

as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			7	4	3		

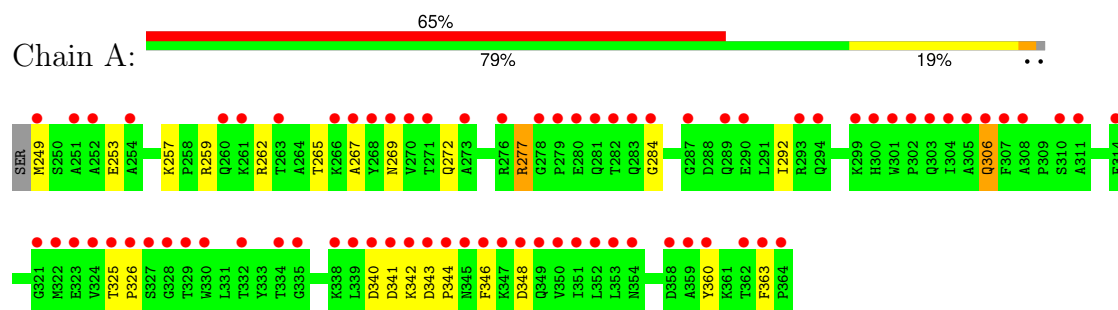
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	120	Total	O	0	0
			120	120		

3 Residue-property plots [i](#)

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: Nucleoprotein



4 Data and refinement statistics

Property	Value	Source
Space group	I 41	Depositor
Cell constants a, b, c, α , β , γ	88.31Å 88.31Å 40.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	37.12 – 1.66 37.12 – 1.66	Depositor EDS
% Data completeness (in resolution range)	49.0 (37.12-1.66) 49.2 (37.12-1.66)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.19 (at 1.66Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.203 , 0.284 0.270 , 0.324	Depositor DCC
R_{free} test set	1014 reflections (5.37%)	wwPDB-VP
Wilson B-factor (Å ²)	27.0	Xtriage
Anisotropy	0.410	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 53.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.036 for -k,-h,-l	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	1068	wwPDB-VP
Average B, all atoms (Å ²)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.12% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NY7, PEG

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	1.05	0/950	1.34	0/1282

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

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	927	0	906	17	0
2	A	14	0	0	1	0
3	A	7	0	10	0	0
4	A	120	0	0	2	0
All	All	1068	0	916	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 9.

All (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:341:ASP:HA	1:A:346:PHE:CG	2.36	0.60
1:A:340:ASP:OD1	1:A:342:LYS:HB3	2.02	0.60
1:A:341:ASP:HA	1:A:346:PHE:CD1	2.41	0.56
1:A:272:GLN:HG3	4:A:536:HOH:O	2.07	0.54
1:A:249:MET:HE1	1:A:257:LYS:NZ	2.23	0.53
1:A:249:MET:HE1	1:A:257:LYS:HZ2	1.73	0.52
1:A:360:TYR:HA	1:A:363:PHE:CE2	2.45	0.52
1:A:306:GLN:HE22	2:A:401:NY7:C10	2.23	0.51
1:A:360:TYR:HA	1:A:363:PHE:CD2	2.50	0.46
1:A:277:ARG:HA	1:A:284:GLY:O	2.15	0.46
1:A:269:ASN:HB2	1:A:292:ILE:O	2.16	0.46
1:A:249:MET:HE2	1:A:253:GLU:C	2.42	0.45
1:A:267:ALA:N	4:A:505:HOH:O	2.36	0.44
1:A:259:ARG:HA	1:A:262:ARG:HD2	2.00	0.42
1:A:265:THR:O	1:A:269:ASN:HA	2.20	0.41
1:A:325:THR:HB	1:A:326:PRO:HD2	2.03	0.41
1:A:343:ASP:OD1	1:A:344:PRO:HD2	2.21	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	115/117 (98%)	113 (98%)	2 (2%)	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	96/96 (100%)	93 (97%)	3 (3%)	35	12

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

Mol	Chain	Res	Type
1	A	277	ARG
1	A	306	GLN
1	A	348	ASP

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

Mol	Chain	Res	Type
1	A	306	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	NY7	A	401	-	13,14,14	0.44	0	16,18,18	0.40	0
3	PEG	A	402	-	6,6,6	0.11	0	5,5,5	0.08	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	NY7	A	401	-	-	2/8/8/8	0/1/1/1
3	PEG	A	402	-	-	0/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

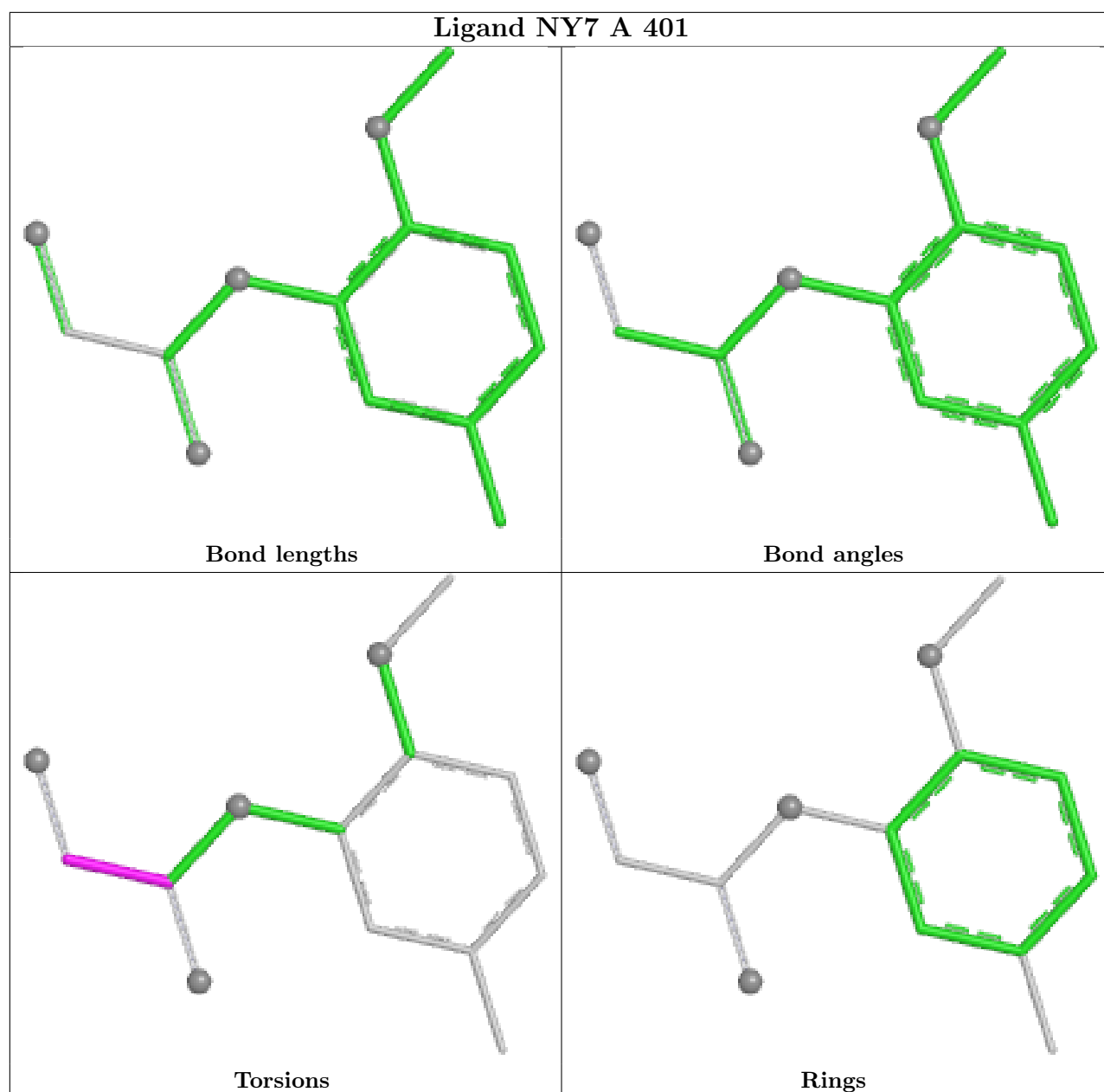
Mol	Chain	Res	Type	Atoms
2	A	401	NY7	N2-C10-C9-N1
2	A	401	NY7	N2-C10-C9-O2

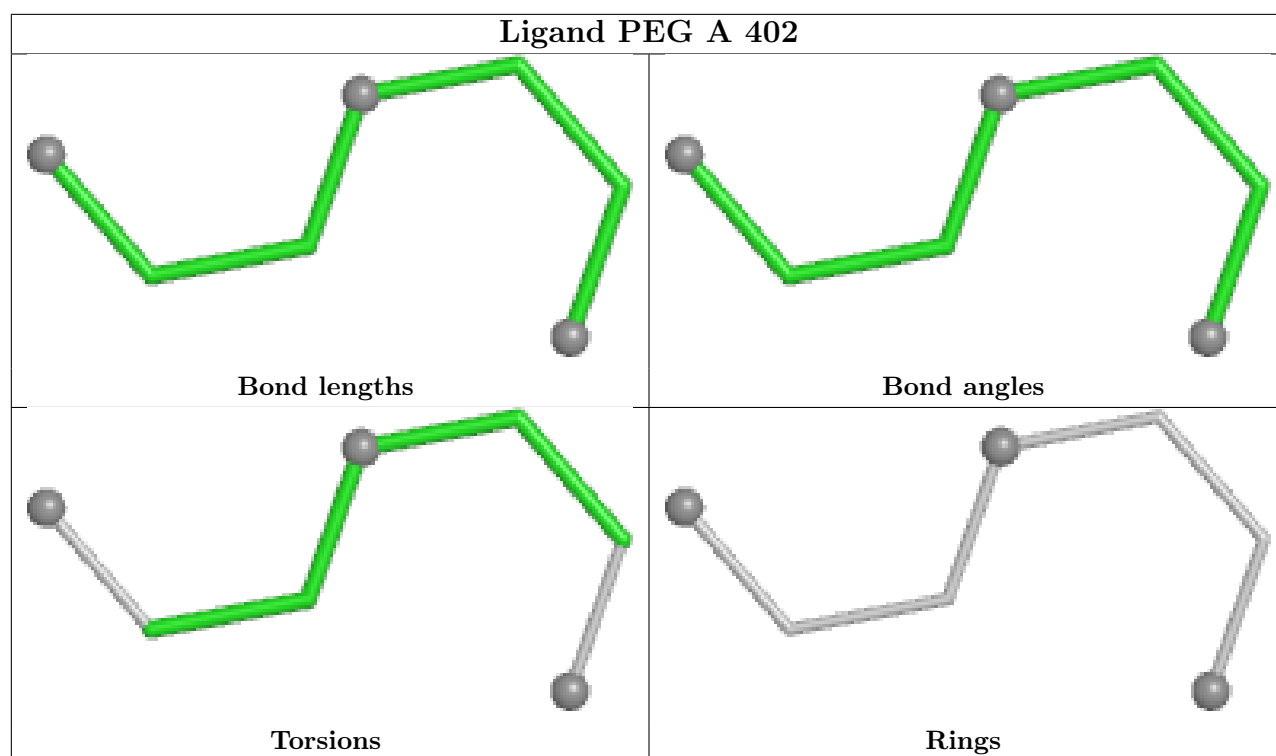
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	401	NY7	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	116/117 (99%)	3.27	76 (65%) 0 0	11, 24, 46, 51	55 (47%)

All (76) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	346	PHE	11.3
1	A	348	ASP	11.2
1	A	345	ASN	10.3
1	A	324	VAL	9.1
1	A	344	PRO	8.1
1	A	281[A]	GLN	7.9
1	A	339	LEU	7.7
1	A	353	LEU	6.9
1	A	343	ASP	6.9
1	A	330	TRP	6.8
1	A	351	ILE	6.8
1	A	350	VAL	6.6
1	A	360	TYR	6.5
1	A	303	GLN	6.3
1	A	347	LYS	6.2
1	A	326	PRO	6.0
1	A	266	LYS	5.9
1	A	341	ASP	5.8
1	A	364	PRO	5.8
1	A	342	LYS	5.6
1	A	325	THR	5.5
1	A	363	PHE	5.3
1	A	352	LEU	5.2
1	A	349	GLN	4.9
1	A	340	ASP	4.9
1	A	334	THR	4.9
1	A	322	MET	4.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	332	THR	4.6
1	A	289	GLN	4.6
1	A	282	THR	4.6
1	A	323	GLU	4.5
1	A	252	ALA	4.5
1	A	301	TRP	4.3
1	A	276	ARG	4.2
1	A	278	GLY	4.1
1	A	279	PRO	4.1
1	A	306	GLN	4.0
1	A	305	ALA	4.0
1	A	304	ILE	4.0
1	A	338	LYS	4.0
1	A	354	ASN	3.9
1	A	307	PHE	3.8
1	A	302	PRO	3.8
1	A	327	SER	3.7
1	A	358	ASP	3.7
1	A	280	GLU	3.6
1	A	268	TYR	3.2
1	A	273	ALA	3.2
1	A	249	MET	3.2
1	A	359	ALA	3.1
1	A	283	GLN	3.1
1	A	328	GLY	3.1
1	A	263	THR	3.1
1	A	329	THR	3.0
1	A	311	ALA	3.0
1	A	299	LYS	3.0
1	A	335	GLY	2.9
1	A	267	ALA	2.8
1	A	314	PHE	2.8
1	A	269	ASN	2.8
1	A	294	GLN	2.7
1	A	254	ALA	2.7
1	A	271	THR	2.6
1	A	290	GLU	2.5
1	A	260	GLN	2.5
1	A	284	GLY	2.5
1	A	287	GLY	2.5
1	A	261	LYS	2.4
1	A	321	GLY	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	308	ALA	2.3
1	A	251	ALA	2.3
1	A	362	THR	2.2
1	A	310	SER	2.1
1	A	293	ARG	2.1
1	A	300	HIS	2.1
1	A	270	VAL	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](#)

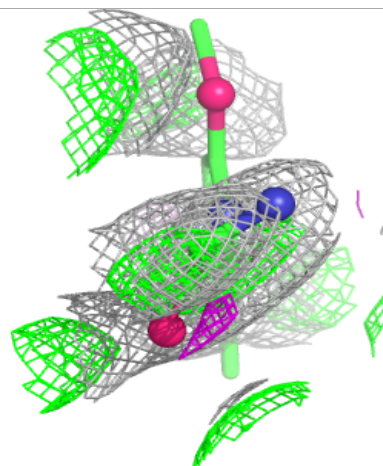
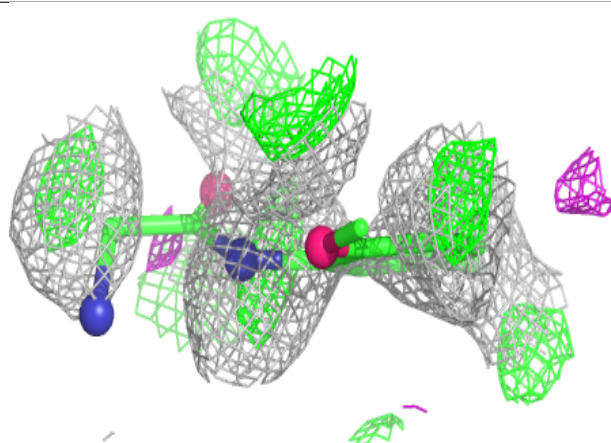
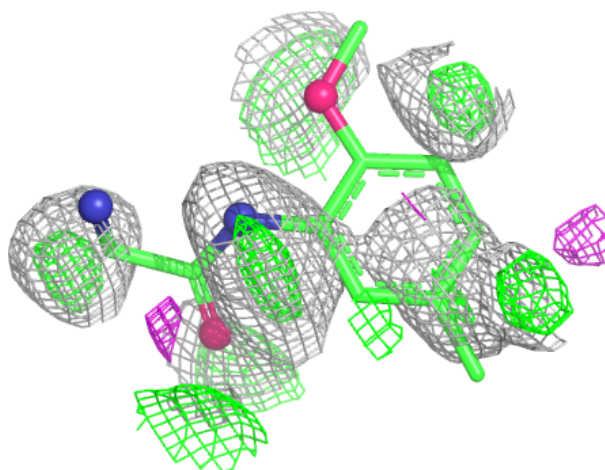
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists 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	NY7	A	401	14/14	0.56	0.50	74,88,91,93	14
3	PEG	A	402	7/7	0.66	0.23	63,65,69,71	7

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.

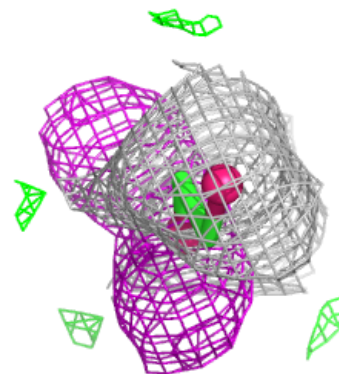
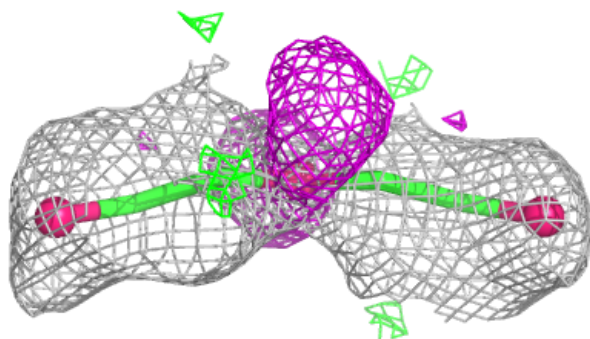
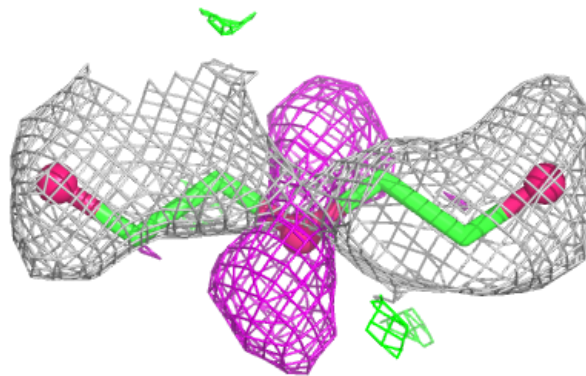
Electron density around NY7 A 401:

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



Electron density around PEG A 402:

$2mF_o-DF_c$ (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.