



# wwPDB X-ray Structure Validation Summary Report

Oct 5, 2023 – 04:33 AM EDT

PDB ID : 6VMZ  
Title : Crystal Structure of a H5N1 influenza virus hemagglutinin with CBS1117  
Authors : Antanasijevic, A.; Durst, M.A.; Lavie, A.; Caffrey, M.  
Deposited on : 2020-01-28  
Resolution : 2.20 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](mailto: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

<https://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 : **FAILED**  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : **FAILED**  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.35.1

## 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 2.20 Å.

There are no overall percentile quality scores available for this entry.

MolProbity and EDS failed to run properly - the sequence quality summary graphics cannot be shown.

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 12247 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 Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	320	Total 2548	C 1610	N 442	O 481	S 15	0	1	0
1	C	322	Total 2566	C 1622	N 444	O 484	S 16	0	2	0
1	E	322	Total 2556	C 1615	N 441	O 485	S 15	0	1	0

There are 39 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	ALA	-	expression tag	UNP Q1KHJ8
A	8	ASP	-	expression tag	UNP Q1KHJ8
A	9	PRO	-	expression tag	UNP Q1KHJ8
A	10	GLY	-	expression tag	UNP Q1KHJ8
A	332	GLN	-	expression tag	UNP Q1KHJ8
A	333	ARG	-	expression tag	UNP Q1KHJ8
A	334	GLU	-	expression tag	UNP Q1KHJ8
A	335	ARG	-	expression tag	UNP Q1KHJ8
A	336	ARG	-	expression tag	UNP Q1KHJ8
A	337	ARG	-	expression tag	UNP Q1KHJ8
A	338	LYS	-	expression tag	UNP Q1KHJ8
A	339	LYS	-	expression tag	UNP Q1KHJ8
A	340	ARG	-	expression tag	UNP Q1KHJ8
C	7	ALA	-	expression tag	UNP Q1KHJ8
C	8	ASP	-	expression tag	UNP Q1KHJ8
C	9	PRO	-	expression tag	UNP Q1KHJ8
C	10	GLY	-	expression tag	UNP Q1KHJ8
C	332	GLN	-	expression tag	UNP Q1KHJ8
C	333	ARG	-	expression tag	UNP Q1KHJ8
C	334	GLU	-	expression tag	UNP Q1KHJ8
C	335	ARG	-	expression tag	UNP Q1KHJ8
C	336	ARG	-	expression tag	UNP Q1KHJ8
C	337	ARG	-	expression tag	UNP Q1KHJ8

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	338	LYS	-	expression tag	UNP Q1KHJ8
C	339	LYS	-	expression tag	UNP Q1KHJ8
C	340	ARG	-	expression tag	UNP Q1KHJ8
E	7	ALA	-	expression tag	UNP Q1KHJ8
E	8	ASP	-	expression tag	UNP Q1KHJ8
E	9	PRO	-	expression tag	UNP Q1KHJ8
E	10	GLY	-	expression tag	UNP Q1KHJ8
E	332	GLN	-	expression tag	UNP Q1KHJ8
E	333	ARG	-	expression tag	UNP Q1KHJ8
E	334	GLU	-	expression tag	UNP Q1KHJ8
E	335	ARG	-	expression tag	UNP Q1KHJ8
E	336	ARG	-	expression tag	UNP Q1KHJ8
E	337	ARG	-	expression tag	UNP Q1KHJ8
E	338	LYS	-	expression tag	UNP Q1KHJ8
E	339	LYS	-	expression tag	UNP Q1KHJ8
E	340	ARG	-	expression tag	UNP Q1KHJ8

- Molecule 2 is a protein called Hemagglutinin.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
			Total	C	N	O				S
2	B	174	1412	878	245	281	8	0	0	0
2	D	174	1412	878	245	281	8	0	0	0
2	F	142	1146	719	199	223	5	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

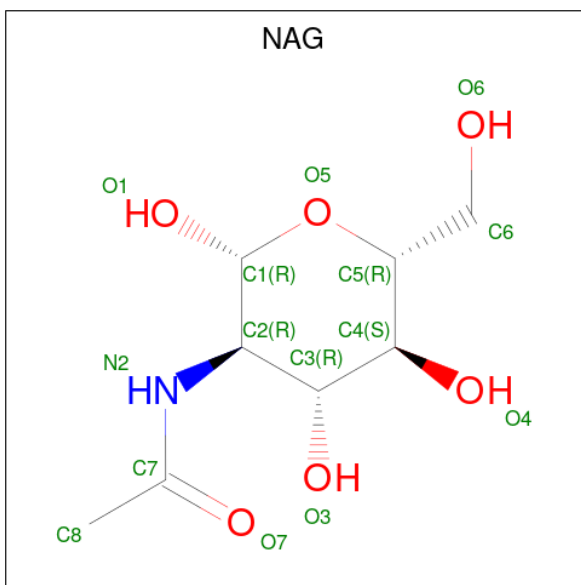
Chain	Residue	Modelled	Actual	Comment	Reference
B	176	SER	-	expression tag	UNP Q1KHK7
B	177	ARG	-	expression tag	UNP Q1KHK7
B	178	LEU	-	expression tag	UNP Q1KHK7
B	179	VAL	-	expression tag	UNP Q1KHK7
B	180	PRO	-	expression tag	UNP Q1KHK7
B	181	ARG	-	expression tag	UNP Q1KHK7
D	176	SER	-	expression tag	UNP Q1KHK7
D	177	ARG	-	expression tag	UNP Q1KHK7
D	178	LEU	-	expression tag	UNP Q1KHK7
D	179	VAL	-	expression tag	UNP Q1KHK7
D	180	PRO	-	expression tag	UNP Q1KHK7
D	181	ARG	-	expression tag	UNP Q1KHK7

*Continued on next page...*

Continued from previous page...

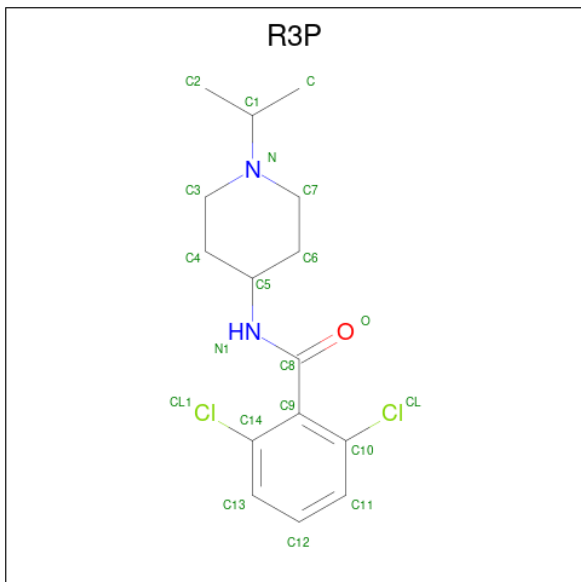
Chain	Residue	Modelled	Actual	Comment	Reference
F	176	SER	-	expression tag	UNP Q1KHK7
F	177	ARG	-	expression tag	UNP Q1KHK7
F	178	LEU	-	expression tag	UNP Q1KHK7
F	179	VAL	-	expression tag	UNP Q1KHK7
F	180	PRO	-	expression tag	UNP Q1KHK7
F	181	ARG	-	expression tag	UNP Q1KHK7

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	E	1	Total	C	N	O	0	0
			14	8	1	5		
3	E	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is 2,6-dichloro-N-[1-(propan-2-yl)piperidin-4-yl]benzamide (three-letter code: R3P) (formula:  $C_{15}H_{20}Cl_2N_2O$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	Cl	N	O		
4	A	1	Total	C	Cl	N	O	0	0
			20	15	2	2	1		
4	C	1	Total	C	Cl	N	O	0	0
			20	15	2	2	1		
4	E	1	Total	C	Cl	N	O	0	0
			20	15	2	2	1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	114	Total	O	0	0
			114	114		
5	C	132	Total	O	0	0
			132	132		
5	E	109	Total	O	0	0
			109	109		
5	B	29	Total	O	0	0
			29	29		
5	D	27	Total	O	0	0
			27	27		
5	F	24	Total	O	0	0
			24	24		

MolProbity and EDS failed to run properly - this section is therefore empty.

### 3 Data and refinement statistics

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	72.71Å 126.08Å 249.64Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.47 – 2.20	Depositor
% Data completeness (in resolution range)	93.5 (29.47-2.20)	Depositor
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.20 (at 2.20Å)	Xtrriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.208 , 0.252	Depositor
Wilson B-factor (Å <sup>2</sup> )	43.3	Xtrriage
Anisotropy	0.071	Xtrriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	12247	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	53.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.96% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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.

## 4 Model quality [i](#)

### 4.1 Standard geometry [i](#)

MolProbity failed to run properly - this section is therefore empty.

### 4.2 Too-close contacts [i](#)

MolProbity failed to run properly - this section is therefore empty.

### 4.3 Torsion angles [i](#)

#### 4.3.1 Protein backbone [i](#)

MolProbity failed to run properly - this section is therefore empty.

#### 4.3.2 Protein sidechains [i](#)

MolProbity failed to run properly - this section is therefore empty.

#### 4.3.3 RNA [i](#)

MolProbity failed to run properly - this section is therefore empty.

### 4.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 4.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 4.6 Ligand geometry [i](#)

11 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
3	NAG	C	401	1	14,14,15	0.45	0	17,19,21	1.02	1 (5%)
3	NAG	E	402	1	14,14,15	0.60	0	17,19,21	1.59	3 (17%)
3	NAG	C	402	1	14,14,15	0.59	0	17,19,21	1.64	4 (23%)
3	NAG	C	403	1	14,14,15	0.48	0	17,19,21	1.49	4 (23%)
4	R3P	E	403	-	21,21,21	0.94	2 (9%)	29,29,29	1.67	7 (24%)
3	NAG	A	402	1	14,14,15	1.05	1 (7%)	17,19,21	2.90	8 (47%)
4	R3P	A	404	-	21,21,21	0.97	2 (9%)	29,29,29	1.50	4 (13%)
3	NAG	E	401	1	14,14,15	0.69	0	17,19,21	2.11	6 (35%)
3	NAG	A	403	1	14,14,15	0.80	0	17,19,21	2.08	4 (23%)
3	NAG	A	401	1	14,14,15	0.68	0	17,19,21	1.97	2 (11%)
4	R3P	C	404	-	21,21,21	0.88	1 (4%)	29,29,29	1.39	4 (13%)

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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	C	401	1	-	1/6/23/26	0/1/1/1
3	NAG	E	402	1	-	2/6/23/26	0/1/1/1
3	NAG	C	402	1	-	1/6/23/26	0/1/1/1
3	NAG	C	403	1	-	1/6/23/26	0/1/1/1
4	R3P	E	403	-	-	9/12/22/22	0/2/2/2
3	NAG	A	402	1	-	4/6/23/26	0/1/1/1
4	R3P	A	404	-	-	5/12/22/22	1/2/2/2
3	NAG	E	401	1	-	2/6/23/26	0/1/1/1
3	NAG	A	403	1	-	3/6/23/26	0/1/1/1
3	NAG	A	401	1	-	0/6/23/26	0/1/1/1
4	R3P	C	404	-	-	11/12/22/22	0/2/2/2

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	404	R3P	C10-CL	2.73	1.80	1.73

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	404	R3P	C14-CL1	2.69	1.80	1.73
3	A	402	NAG	C1-C2	2.64	1.56	1.52
4	E	403	R3P	C10-CL	2.61	1.79	1.73
4	C	404	R3P	C14-CL1	2.56	1.79	1.73

The worst 5 of 47 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	401	NAG	C1-O5-C5	7.25	122.02	112.19
3	A	402	NAG	C1-O5-C5	6.48	120.97	112.19
3	A	403	NAG	C1-O5-C5	5.60	119.78	112.19
3	A	402	NAG	C2-N2-C7	5.37	130.55	122.90
3	A	402	NAG	C4-C3-C2	4.60	117.76	111.02

There are no chirality outliers.

5 of 39 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	403	NAG	C3-C2-N2-C7
4	A	404	R3P	C2-C1-N-C3
4	A	404	R3P	C-C1-N-C3
4	A	404	R3P	C2-C1-N-C7
4	A	404	R3P	C-C1-N-C7

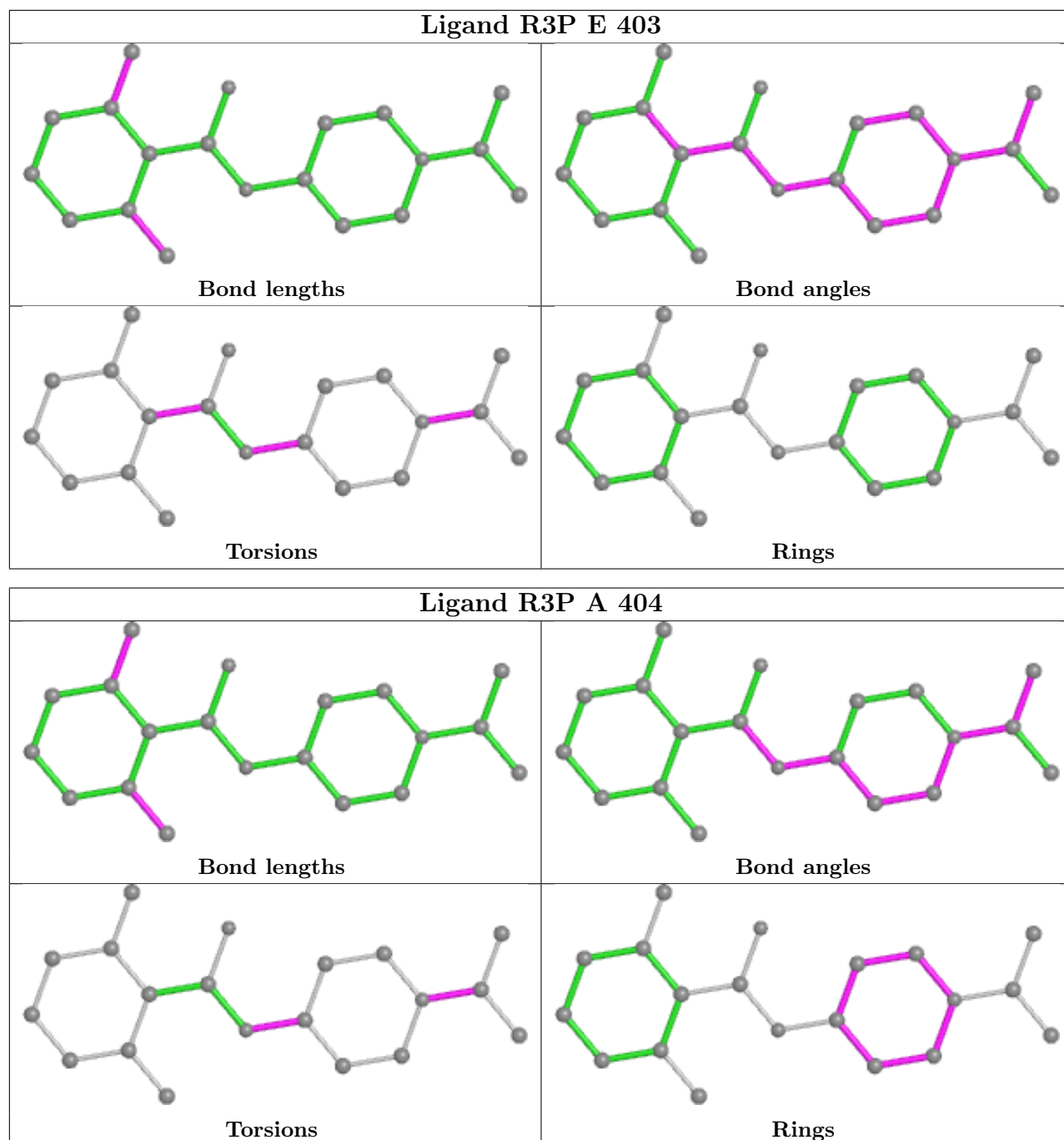
All (1) ring outliers are listed below:

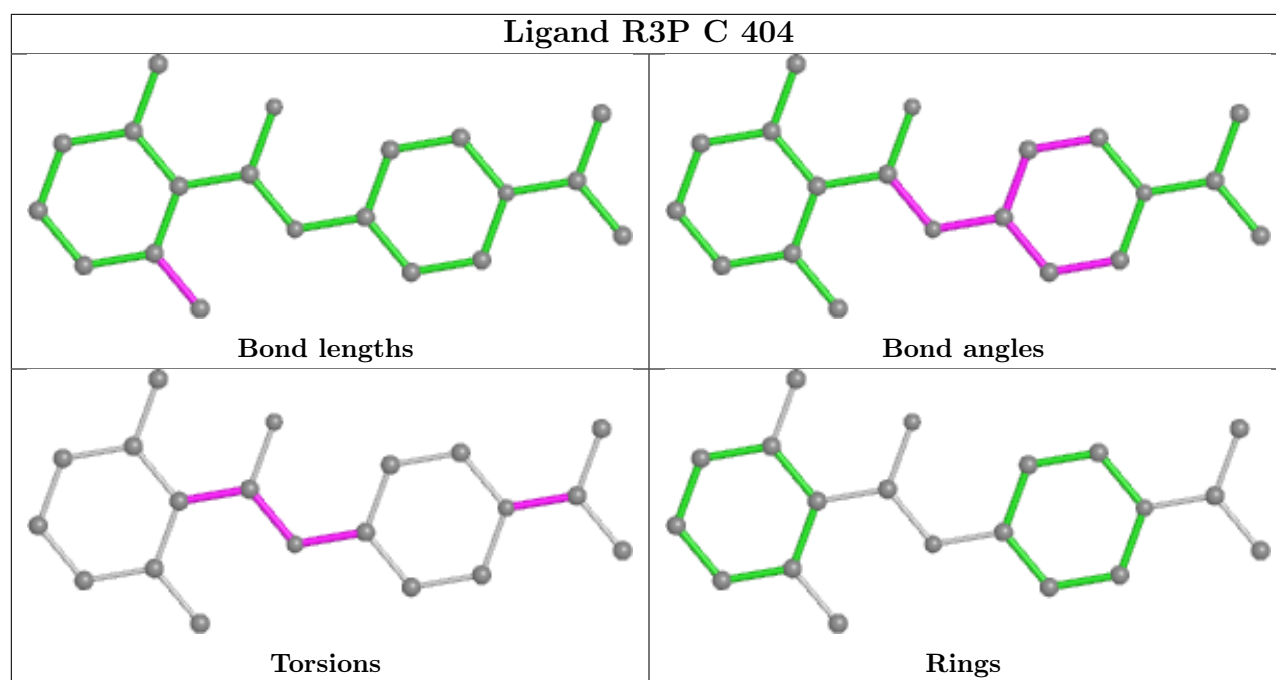
Mol	Chain	Res	Type	Atoms
4	A	404	R3P	C3-C4-C5-C6-C7-N

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 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.





#### 4.7 Other polymers [i](#)

There are no such residues in this entry.

#### 4.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 5 Fit of model and data

### 5.1 Protein, DNA and RNA chains

EDS failed to run properly - this section is therefore empty.

### 5.2 Non-standard residues in protein, DNA, RNA chains

EDS failed to run properly - this section is therefore empty.

### 5.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

### 5.4 Ligands

EDS failed to run properly - this section is therefore empty.

### 5.5 Other polymers

EDS failed to run properly - this section is therefore empty.