



Full wwPDB X-ray Structure Validation Report ⓘ

Aug 7, 2020 – 04:11 AM BST

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 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

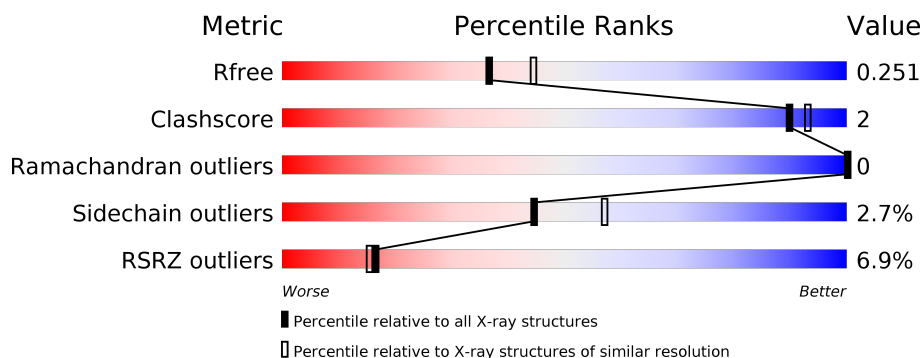
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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 $\leq 5\%$. 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	334	<div> <div>3%</div> <div>89%</div> <div>6%</div> <div>.</div> </div>
1	C	334	<div> <div>3%</div> <div>92%</div> <div>5%</div> <div>.</div> </div>
1	E	334	<div> <div>7%</div> <div>90%</div> <div>6%</div> <div>.</div> </div>
2	B	181	<div> <div>6%</div> <div>88%</div> <div>7%</div> <div>.</div> </div>
2	D	181	<div> <div>10%</div> <div>87%</div> <div>8%</div> <div>.</div> </div>
2	F	181	<div> <div>14%</div> <div>72%</div> <div>7%</div> <div>22%</div> </div>

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
3	NAG	A	402	-	-	-	X
3	NAG	A	403	-	-	-	X
3	NAG	C	403	-	-	-	X

2 Entry composition

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
1	A	320	Total	C	N	O	S	0	1	0
			2548	1610	442	481	15			
1	C	322	Total	C	N	O	S	0	2	0
			2566	1622	444	484	16			
1	E	322	Total	C	N	O	S	0	1	0
			2556	1615	441	485	15			

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
2	B	174	Total	C	N	O	S	0	0	0
			1412	878	245	281	8			
2	D	174	Total	C	N	O	S	0	0	0
			1412	878	245	281	8			
2	F	142	Total	C	N	O	S	0	0	0
			1146	719	199	223	5			

There are 18 discrepancies between the modelled and reference sequences:

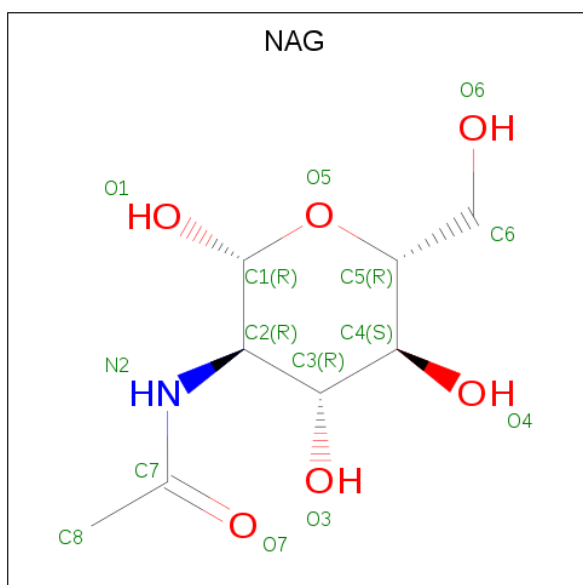
Chain	Residue	Modelled	Actual	Comment	Reference
B	176	SER	-	expression tag	UNP Q1KHK7
B	177	ARG	-	expression tag	UNP Q1KHK7
B	178	LEU	-	expression tag	UNP Q1KHK7
B	179	VAL	-	expression tag	UNP Q1KHK7
B	180	PRO	-	expression tag	UNP Q1KHK7
B	181	ARG	-	expression tag	UNP Q1KHK7
D	176	SER	-	expression tag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

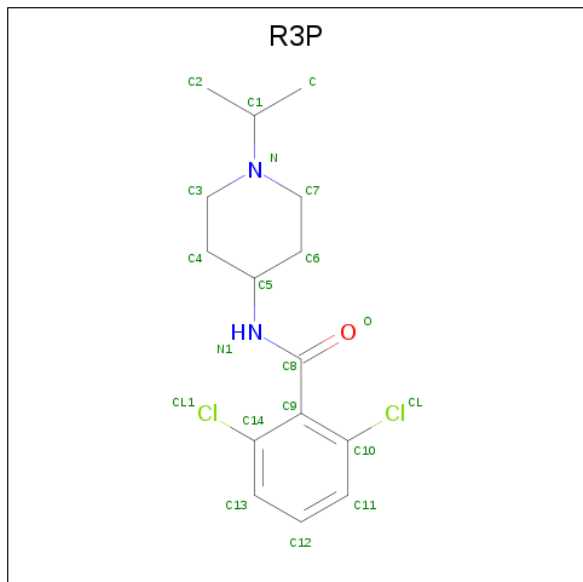
Chain	Residue	Modelled	Actual	Comment	Reference
F	176	SER	-	expression tag	UNP Q1KHK7
F	177	ARG	-	expression tag	UNP Q1KHK7
F	178	LEU	-	expression tag	UNP Q1KHK7
F	179	VAL	-	expression tag	UNP Q1KHK7
F	180	PRO	-	expression tag	UNP Q1KHK7
F	181	ARG	-	expression tag	UNP Q1KHK7

- 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
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 author).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
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		

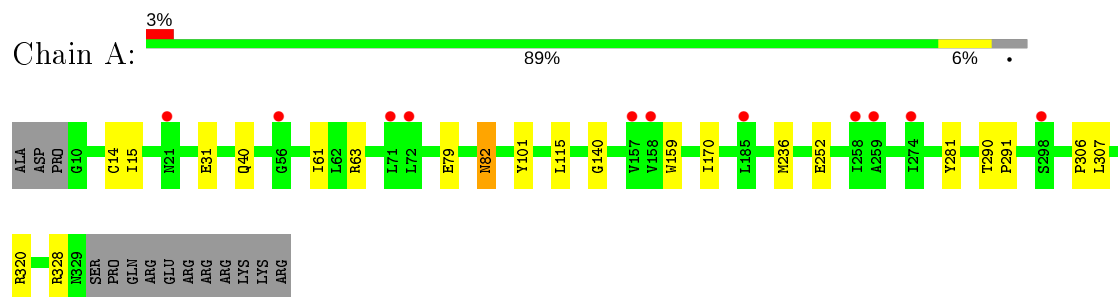
- Molecule 5 is water.

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		

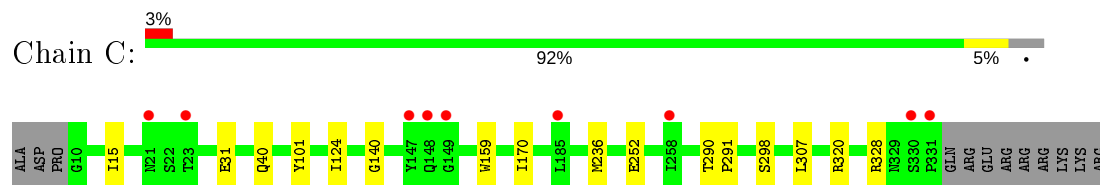
3 Residue-property plots [i](#)

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.

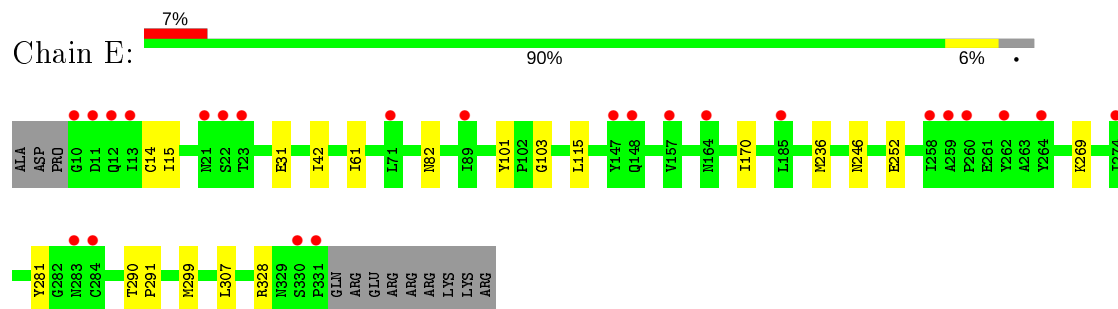
• Molecule 1: Hemagglutinin



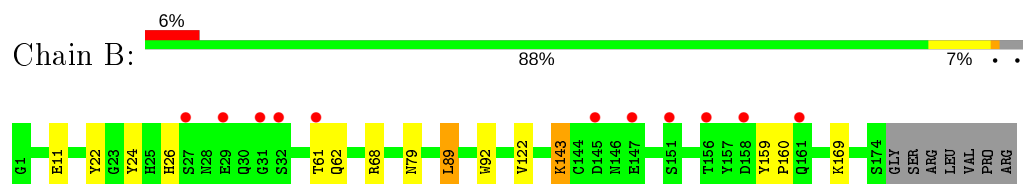
• Molecule 1: Hemagglutinin



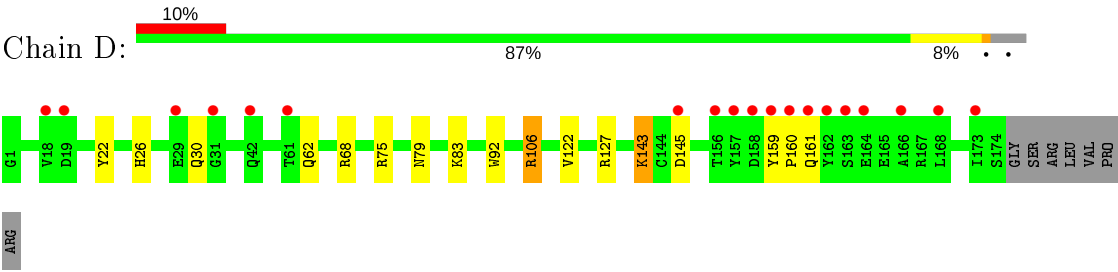
• Molecule 1: Hemagglutinin



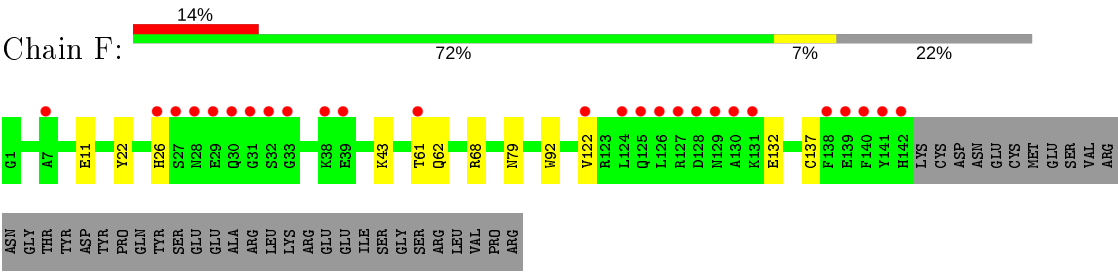
• Molecule 2: Hemagglutinin



• Molecule 2: Hemagglutinin



• Molecule 2: Hemagglutinin



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	72.71Å 126.08Å 249.64Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.47 – 2.20 29.41 – 2.20	Depositor EDS
% Data completeness (in resolution range)	93.5 (29.47-2.20) 93.5 (29.41-2.20)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.20 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.208 , 0.252 0.211 , 0.251	Depositor DCC
R_{free} test set	5426 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å ²)	43.3	Xtriage
Anisotropy	0.071	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 38.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	12247	wwPDB-VP
Average B, all atoms (Å ²)	53.0	wwPDB-VP

Xtriage'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.*

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: R3P, NAG

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	0.72	0/2612	0.80	0/3545
1	C	0.71	0/2634	0.82	0/3575
1	E	0.69	0/2621	0.79	0/3559
2	B	0.72	0/1439	0.80	0/1934
2	D	0.71	0/1439	0.81	0/1934
2	F	0.69	0/1169	0.78	0/1572
All	All	0.71	0/11914	0.80	0/16119

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 [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	2548	0	2496	12	0
1	C	2566	0	2517	6	0
1	E	2556	0	2501	9	0
2	B	1412	0	1319	8	0
2	D	1412	0	1319	11	0
2	F	1146	0	1078	6	0
3	A	42	0	39	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	42	0	39	0	0
3	E	28	0	26	0	0
4	A	20	0	0	0	0
4	C	20	0	0	0	0
4	E	20	0	0	0	0
5	A	114	0	0	0	0
5	B	29	0	0	0	0
5	C	132	0	0	0	0
5	D	27	0	0	0	0
5	E	109	0	0	1	0
5	F	24	0	0	1	0
All	All	12247	0	11334	42	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 2.

All (42) 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:D:30:GLN:HE22	2:D:145:ASP:HB2	1.46	0.79
2:D:68:ARG:HE	2:F:79:ASN:HD22	1.41	0.68
2:D:127:ARG:NH1	2:F:132:GLU:O	2.27	0.67
2:B:68:ARG:HE	2:D:79:ASN:HD22	1.49	0.58
1:A:82:ASN:C	1:A:82:ASN:HD22	2.05	0.58
1:A:15:ILE:HD11	2:B:122:VAL:HG21	1.87	0.56
1:E:269:LYS:NZ	5:E:501:HOH:O	2.42	0.53
1:C:15:ILE:HD11	2:D:122:VAL:HG21	1.90	0.52
1:E:101:TYR:CD1	1:E:236:MET:HG2	2.45	0.52
1:A:101:TYR:CD1	1:A:236:MET:HG2	2.45	0.52
2:B:79:ASN:HD22	2:F:68:ARG:HE	1.58	0.52
1:A:63:ARG:HD2	1:A:79:GLU:OE2	2.12	0.49
2:D:68:ARG:HH21	2:F:79:ASN:ND2	2.10	0.49
1:E:31:GLU:OE2	1:E:328:ARG:NH2	2.46	0.49
1:C:31:GLU:OE2	1:C:328:ARG:NH2	2.45	0.49
1:A:31:GLU:OE2	1:A:328:ARG:NH2	2.46	0.48
1:C:101:TYR:CD1	1:C:236:MET:HG2	2.48	0.48
1:E:15:ILE:HD11	2:F:122:VAL:HG21	1.96	0.47
2:F:62:GLN:HG3	2:F:92:TRP:CG	2.50	0.47
1:C:170:ILE:O	1:C:252:GLU:HA	2.15	0.46
1:A:170:ILE:O	1:A:252:GLU:HA	2.15	0.46
1:E:170:ILE:O	1:E:252:GLU:HA	2.15	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:61:ILE:HD12	1:E:281:TYR:HB2	1.99	0.45
2:B:62:GLN:HG3	2:B:92:TRP:CG	2.51	0.45
1:A:306:PRO:HB3	2:B:89:LEU:HD21	1.99	0.45
2:D:106:ARG:HD3	5:F:301:HOH:O	2.16	0.44
2:D:143:LYS:HA	2:D:143:LYS:HE3	1.99	0.44
2:B:143:LYS:HA	2:B:143:LYS:HE3	2.00	0.44
1:E:42:ILE:C	1:E:299:MET:HG3	2.38	0.43
2:D:62:GLN:HG3	2:D:92:TRP:CG	2.53	0.43
2:B:159:TYR:HB3	2:B:160:PRO:HD3	2.00	0.43
1:A:82:ASN:ND2	1:A:82:ASN:C	2.71	0.43
1:C:291:PRO:HD3	1:C:307:LEU:O	2.18	0.43
1:E:291:PRO:HD3	1:E:307:LEU:O	2.18	0.42
2:D:159:TYR:HB3	2:D:160:PRO:HD3	1.99	0.42
1:A:61:ILE:HD12	1:A:281:TYR:HB2	2.02	0.42
1:A:140:GLY:HA3	1:A:159:TRP:HB3	2.02	0.41
1:A:291:PRO:HD3	1:A:307:LEU:O	2.20	0.41
1:E:103:GLY:HA3	1:E:236:MET:O	2.21	0.41
2:D:75:ARG:HA	2:D:75:ARG:HD3	1.92	0.41
1:A:14:CYS:O	2:B:24:TYR:HA	2.21	0.41
1:C:140:GLY:HA3	1:C:159:TRP:HB3	2.03	0.41

There are no symmetry-related clashes.

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	319/334 (96%)	313 (98%)	6 (2%)	0	100	100
1	C	322/334 (96%)	316 (98%)	6 (2%)	0	100	100
1	E	321/334 (96%)	315 (98%)	6 (2%)	0	100	100
2	B	172/181 (95%)	164 (95%)	8 (5%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	D	172/181 (95%)	166 (96%)	6 (4%)	0	100	100
2	F	140/181 (77%)	135 (96%)	5 (4%)	0	100	100
All	All	1446/1545 (94%)	1409 (97%)	37 (3%)	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	288/300 (96%)	283 (98%)	5 (2%)	60	74
1	C	291/300 (97%)	286 (98%)	5 (2%)	60	74
1	E	290/300 (97%)	285 (98%)	5 (2%)	60	74
2	B	149/155 (96%)	142 (95%)	7 (5%)	26	33
2	D	149/155 (96%)	143 (96%)	6 (4%)	31	40
2	F	119/155 (77%)	113 (95%)	6 (5%)	24	30
All	All	1286/1365 (94%)	1252 (97%)	34 (3%)	44	58

All (34) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	40	GLN
1	A	82	ASN
1	A	115	LEU
1	A	290	THR
1	A	320	ARG
1	C	40	GLN
1	C	124	ILE
1	C	290	THR
1	C	298	SER
1	C	320	ARG
1	E	14	CYS
1	E	82	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	115	LEU
1	E	246	ASN
1	E	290	THR
2	B	11	GLU
2	B	22	TYR
2	B	26	HIS
2	B	61	THR
2	B	89	LEU
2	B	143	LYS
2	B	169	LYS
2	D	22	TYR
2	D	26	HIS
2	D	83	LYS
2	D	106	ARG
2	D	143	LYS
2	D	161	GLN
2	F	11	GLU
2	F	22	TYR
2	F	26	HIS
2	F	43	LYS
2	F	61	THR
2	F	137	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	82	ASN
1	A	178	ASN
1	C	40	GLN
1	C	246	ASN
1	E	82	ASN
2	B	50	ASN
2	B	79	ASN
2	D	30	GLN
2	D	79	ASN
2	F	79	ASN
2	F	129	ASN

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 ⓘ

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	A	401	1	14,14,15	0.68	0	17,19,21	1.97	2 (11%)
3	NAG	C	402	1	14,14,15	0.59	0	17,19,21	1.64	4 (23%)
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	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%)
4	R3P	A	404	-	21,21,21	0.97	2 (9%)	29,29,29	1.50	4 (13%)
3	NAG	A	402	1	14,14,15	1.05	1 (7%)	17,19,21	2.90	8 (47%)
4	R3P	C	404	-	21,21,21	0.88	1 (4%)	29,29,29	1.39	4 (13%)
3	NAG	E	402	1	14,14,15	0.60	0	17,19,21	1.59	3 (17%)
3	NAG	C	401	1	14,14,15	0.45	0	17,19,21	1.02	1 (5%)

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
3	NAG	A	401	1	-	0/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

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

All (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
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
4	E	403	R3P	C14-CL1	2.39	1.79	1.73

All (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
3	A	403	NAG	C2-N2-C7	4.05	128.66	122.90
3	C	403	NAG	O5-C5-C6	4.03	113.52	107.20
4	A	404	R3P	C5-N1-C8	4.00	128.09	122.55
3	E	401	NAG	O5-C1-C2	-3.95	105.05	111.29
4	C	404	R3P	C6-C5-N1	3.90	118.64	110.56
4	E	403	R3P	C5-N1-C8	3.81	127.82	122.55
3	C	402	NAG	C1-O5-C5	3.79	117.33	112.19
3	E	401	NAG	C1-O5-C5	3.76	117.29	112.19
3	A	402	NAG	C3-C4-C5	3.66	116.76	110.24
4	E	403	R3P	C4-C3-N	3.60	116.95	111.43
3	E	401	NAG	C3-C4-C5	-3.57	103.87	110.24

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	402	NAG	C8-C7-N2	3.45	121.94	116.10
4	A	404	R3P	C6-C5-N1	3.39	117.58	110.56
3	A	402	NAG	O5-C1-C2	3.25	116.42	111.29
3	E	402	NAG	C3-C4-C5	3.17	115.90	110.24
3	E	401	NAG	C4-C3-C2	3.06	115.50	111.02
3	E	402	NAG	C1-C2-N2	3.01	115.63	110.49
4	E	403	R3P	C7-N-C3	2.79	114.18	109.08
3	C	403	NAG	C1-C2-N2	2.73	115.15	110.49
3	C	402	NAG	C1-C2-N2	2.70	115.10	110.49
4	E	403	R3P	C10-C9-C8	-2.65	117.96	121.24
4	C	404	R3P	C5-N1-C8	2.65	126.22	122.55
3	C	402	NAG	O5-C1-C2	-2.64	107.12	111.29
3	E	401	NAG	O5-C5-C6	2.58	111.25	107.20
3	A	403	NAG	C3-C4-C5	2.58	114.84	110.24
3	E	402	NAG	C4-C3-C2	2.57	114.79	111.02
3	C	402	NAG	C3-C4-C5	2.57	114.83	110.24
4	A	404	R3P	C6-C7-N	2.49	115.25	111.43
4	A	404	R3P	C-C1-N	-2.48	106.80	112.20
4	C	404	R3P	C7-C6-C5	-2.39	106.31	110.50
3	A	403	NAG	C1-C2-N2	2.34	114.48	110.49
3	C	403	NAG	C1-O5-C5	2.33	115.34	112.19
4	E	403	R3P	C-C1-N	-2.30	107.20	112.20
3	A	401	NAG	O5-C5-C6	2.28	110.78	107.20
3	A	402	NAG	C1-C2-N2	2.16	114.17	110.49
4	E	403	R3P	C7-C6-C5	-2.14	106.75	110.50
3	A	402	NAG	O7-C7-C8	-2.12	118.11	122.06
4	C	404	R3P	C3-C4-C5	2.10	114.18	110.50
3	C	403	NAG	C6-C5-C4	-2.07	108.15	113.00
3	C	401	NAG	C1-O5-C5	2.04	114.96	112.19
4	E	403	R3P	C6-C5-N1	2.02	114.75	110.56
3	E	401	NAG	O7-C7-N2	2.01	125.64	121.95

There are no chirality outliers.

All (39) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	403	NAG	C3-C2-N2-C7
4	E	403	R3P	O-C8-C9-C14
4	E	403	R3P	O-C8-C9-C10
4	E	403	R3P	C6-C5-N1-C8
4	E	403	R3P	C2-C1-N-C3
4	E	403	R3P	C-C1-N-C3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
4	E	403	R3P	C2-C1-N-C7
4	E	403	R3P	C-C1-N-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
4	C	404	R3P	O-C8-C9-C14
4	C	404	R3P	N1-C8-C9-C14
4	C	404	R3P	O-C8-C9-C10
4	C	404	R3P	N1-C8-C9-C10
4	C	404	R3P	C6-C5-N1-C8
4	C	404	R3P	C2-C1-N-C3
4	C	404	R3P	C-C1-N-C3
4	C	404	R3P	C-C1-N-C7
3	A	403	NAG	O5-C5-C6-O6
4	C	404	R3P	C9-C8-N1-C5
3	A	403	NAG	C4-C5-C6-O6
3	A	402	NAG	C8-C7-N2-C2
3	A	402	NAG	O7-C7-N2-C2
4	A	404	R3P	C6-C5-N1-C8
3	C	402	NAG	O5-C5-C6-O6
3	E	402	NAG	C4-C5-C6-O6
4	E	403	R3P	N1-C8-C9-C14
3	E	401	NAG	C4-C5-C6-O6
3	E	401	NAG	O5-C5-C6-O6
3	C	403	NAG	O5-C5-C6-O6
3	C	401	NAG	C4-C5-C6-O6
4	C	404	R3P	O-C8-N1-C5
4	C	404	R3P	C2-C1-N-C7
4	E	403	R3P	N1-C8-C9-C10
3	A	402	NAG	C3-C2-N2-C7
3	E	402	NAG	O5-C5-C6-O6
3	A	402	NAG	O5-C5-C6-O6

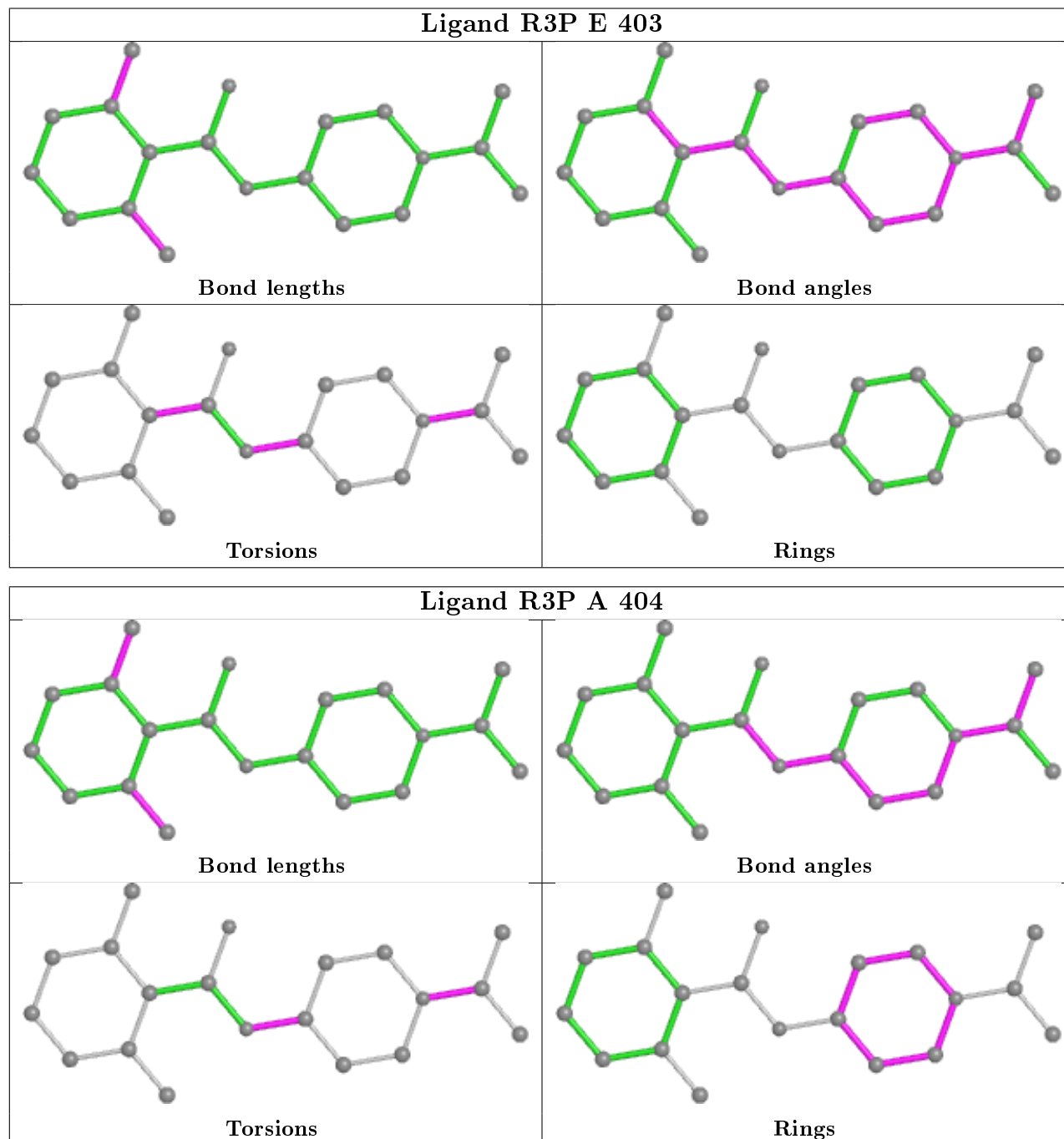
All (1) ring outliers are listed below:

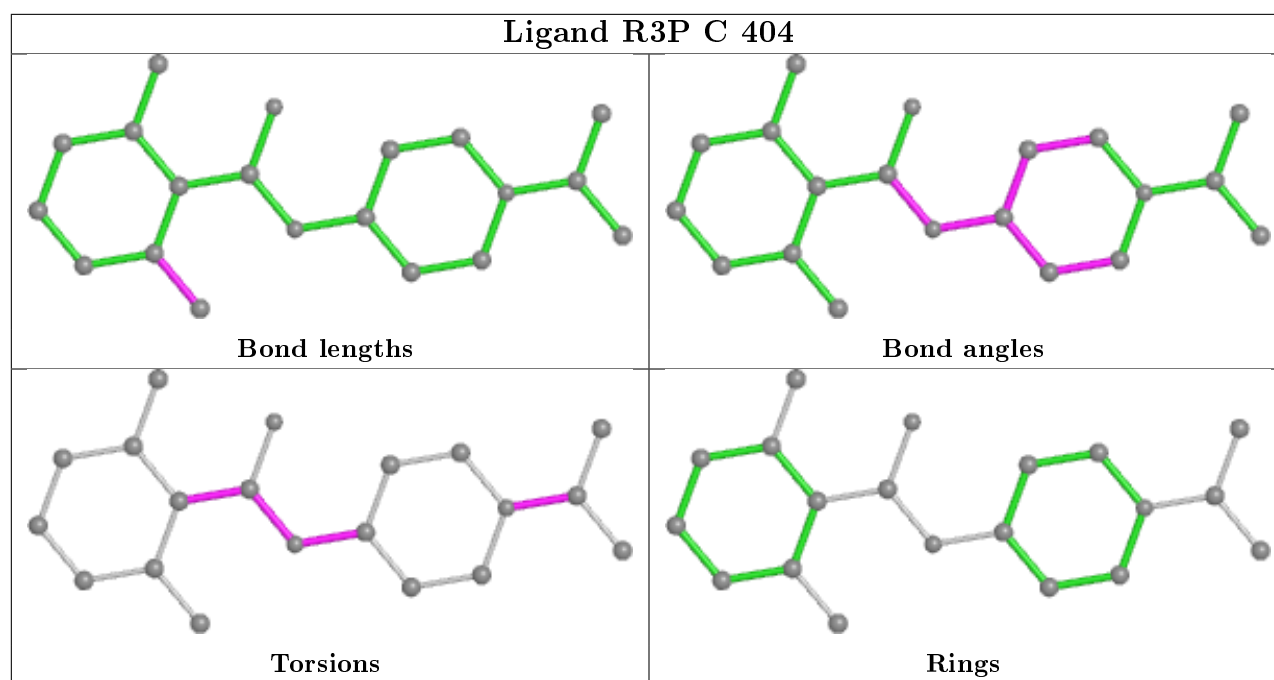
Mol	Chain	Res	Type	Atoms
4	A	404	R3P	C3-C4-C5-C6-C7-N

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.





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	320/334 (95%)	-0.04	11 (3%) 45 43	25, 44, 69, 90	0
1	C	322/334 (96%)	-0.08	9 (2%) 53 51	25, 40, 67, 139	0
1	E	322/334 (96%)	0.24	24 (7%) 14 13	23, 45, 78, 126	0
2	B	174/181 (96%)	0.21	11 (6%) 20 19	26, 60, 97, 110	0
2	D	174/181 (96%)	0.43	19 (10%) 5 5	26, 66, 102, 120	0
2	F	142/181 (78%)	0.80	26 (18%) 1 1	24, 71, 115, 146	0
All	All	1454/1545 (94%)	0.18	100 (6%) 16 15	23, 49, 96, 146	0

All (100) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	140	PHE	7.9
1	C	331	PRO	6.5
2	D	158	ASP	6.0
2	F	128	ASP	5.6
2	D	173	ILE	5.6
2	F	141	TYR	5.3
2	F	129	ASN	5.3
1	E	331	PRO	5.1
2	F	130	ALA	5.0
2	F	27	SER	4.8
2	B	145	ASP	4.8
2	F	31	GLY	4.7
2	F	32	SER	4.3
1	E	10	GLY	4.3
2	D	161	GLN	4.2
2	F	127	ARG	4.2
2	F	142	HIS	4.0
2	F	29	GLU	4.0
1	E	71	LEU	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	158	ASP	3.9
2	B	29	GLU	3.9
1	A	21	ASN	3.9
2	F	139	GLU	3.8
2	F	125	GLN	3.7
1	C	258	ILE	3.6
2	F	38	LYS	3.6
1	E	11	ASP	3.6
1	E	330	SER	3.6
2	D	29	GLU	3.5
2	D	164	GLU	3.5
1	E	147	TYR	3.5
2	F	126	LEU	3.4
2	D	159	TYR	3.4
2	F	124	LEU	3.4
2	B	156	THR	3.3
1	E	258	ILE	3.3
1	E	12	GLN	3.2
2	D	156	THR	3.2
1	E	164	ASN	3.1
2	B	31	GLY	3.1
2	D	157	TYR	3.0
2	F	28	ASN	3.0
2	D	31	GLY	3.0
2	D	160	PRO	3.0
2	F	39	GLU	3.0
2	F	33	GLY	3.0
2	B	27	SER	3.0
1	A	157	VAL	2.9
2	D	162	TYR	2.9
2	F	30	GLN	2.8
1	E	157	VAL	2.8
1	E	260	PRO	2.8
2	D	42	GLN	2.8
2	F	7	ALA	2.7
1	E	148	GLN	2.7
2	D	166	ALA	2.7
2	F	122	VAL	2.7
2	F	61	THR	2.6
1	C	149	GLY	2.6
1	E	284	CYS	2.6
1	E	21	ASN	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	19	ASP	2.6
2	D	168	LEU	2.6
1	E	23	THR	2.6
1	A	298	SER	2.6
1	E	283	ASN	2.6
2	F	138	PHE	2.6
1	E	22	SER	2.5
1	E	89	ILE	2.5
1	A	259	ALA	2.5
1	C	185	LEU	2.5
1	E	259	ALA	2.5
1	E	264	TYR	2.5
1	E	185	LEU	2.4
2	B	32	SER	2.4
1	A	258	ILE	2.4
1	C	147	TYR	2.4
1	A	72	LEU	2.3
1	A	56	GLY	2.3
1	A	158	VAL	2.3
1	A	185	LEU	2.3
1	E	13	ILE	2.3
2	F	26	HIS	2.2
1	C	21	ASN	2.2
1	E	262	TYR	2.2
1	C	330	SER	2.2
2	F	131	LYS	2.2
2	D	18	VAL	2.2
2	D	145	ASP	2.2
2	B	147	GLU	2.1
1	C	23	THR	2.1
2	D	61	THR	2.1
2	B	61	THR	2.1
1	A	274	ILE	2.1
1	E	274	ILE	2.1
2	B	151	SER	2.1
1	C	148	GLN	2.0
2	B	161	GLN	2.0
1	A	71	LEU	2.0
2	D	163	SER	2.0

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 ⓘ

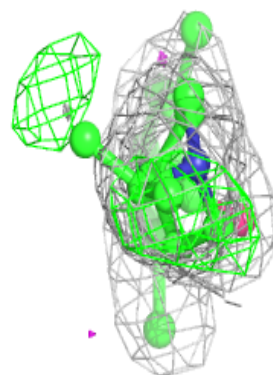
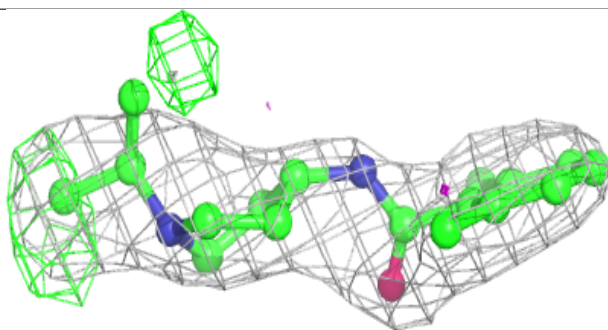
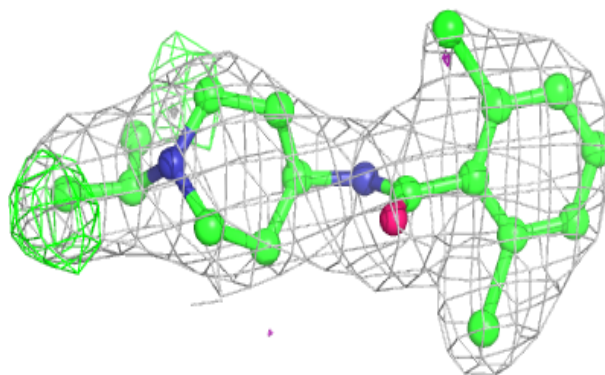
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
3	NAG	A	403	14/15	0.72	0.51	98,116,126,127	0
3	NAG	A	402	14/15	0.75	0.40	112,121,128,128	0
3	NAG	E	402	14/15	0.76	0.38	100,116,123,124	0
4	R3P	A	404	20/20	0.77	0.20	67,85,94,114	0
3	NAG	C	402	14/15	0.78	0.27	87,104,117,118	0
4	R3P	E	403	20/20	0.80	0.20	86,103,108,131	0
3	NAG	C	403	14/15	0.80	0.43	99,115,118,120	0
4	R3P	C	404	20/20	0.83	0.20	81,89,97,107	0
3	NAG	C	401	14/15	0.91	0.21	63,70,81,82	0
3	NAG	A	401	14/15	0.93	0.14	44,49,52,53	0
3	NAG	E	401	14/15	0.94	0.18	42,48,57,59	0

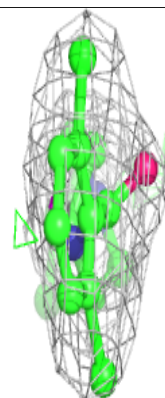
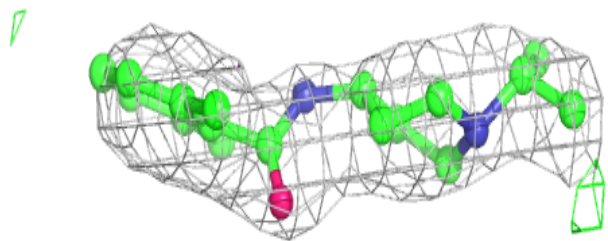
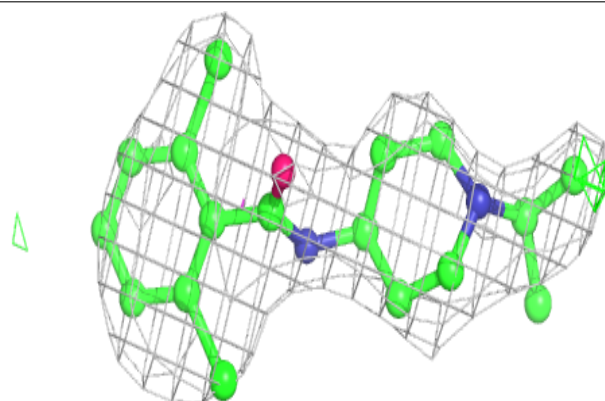
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around R3P A 404:

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

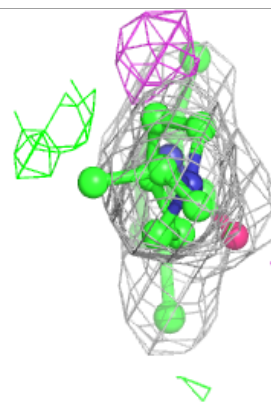
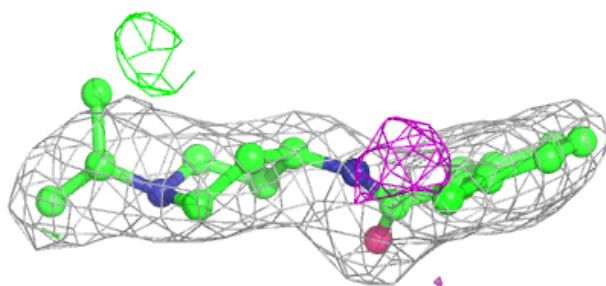
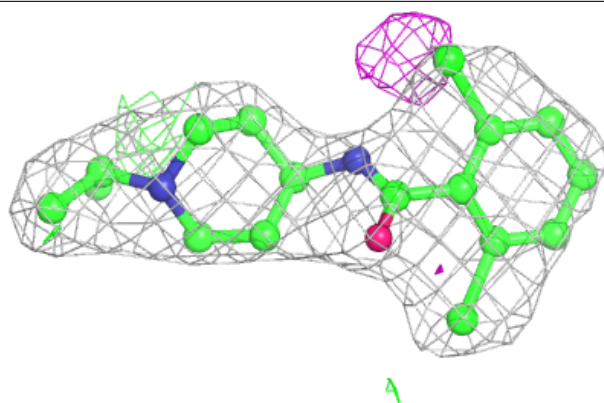
**Electron density around R3P E 403:**

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



Electron density around R3P C 404:

$2mF_o-DF_c$ (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.