



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2026 – 09:46 AM UTC

PDB ID : 6VMZ / pdb_00006vmz
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 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	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

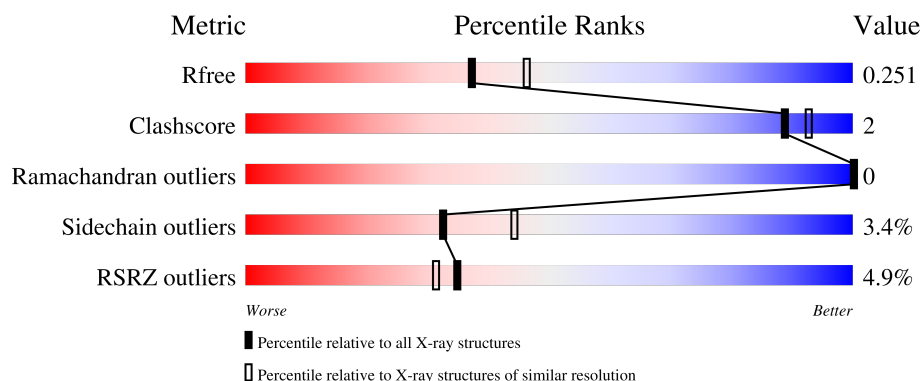
1 Overall quality at a glance

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}	180053	6164 (2.20-2.20)
Clashscore	190562	6851 (2.20-2.20)
Ramachandran outliers	187476	6768 (2.20-2.20)
Sidechain outliers	187428	6769 (2.20-2.20)
RSRZ outliers	180081	6166 (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 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	334	87% 8% ••
1	C	334	91% 5% •
1	E	334	89% 7% •
2	B	181	6% 86% 9% ••
2	D	181	13% 86% 9% ••

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	181	 A horizontal bar chart showing the quality of chain F. The bar is divided into four segments: a red segment at the beginning labeled '14%', followed by a green segment labeled '70%', then a yellow segment labeled '9%', and finally a grey segment at the end labeled '22%'.

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:

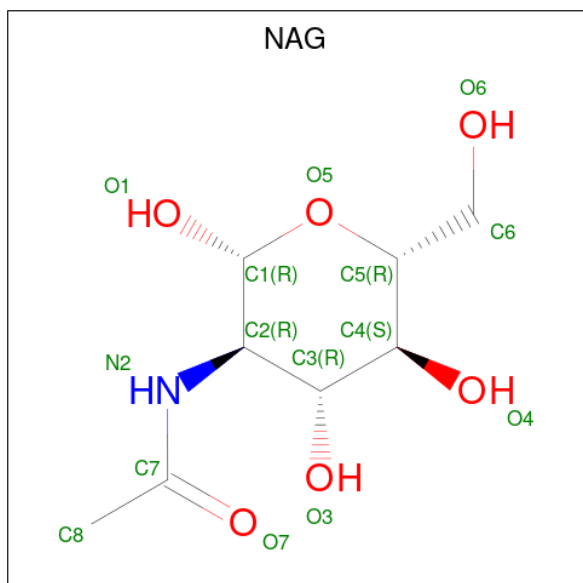
Chain	Residue	Modelled	Actual	Comment	Reference
B	176	SER	-	expression tag	UNP Q1KHK7
B	177	ARG	-	expression tag	UNP Q1KHK7
B	178	LEU	-	expression tag	UNP Q1KHK7
B	179	VAL	-	expression tag	UNP Q1KHK7
B	180	PRO	-	expression tag	UNP Q1KHK7
B	181	ARG	-	expression tag	UNP Q1KHK7
D	176	SER	-	expression tag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

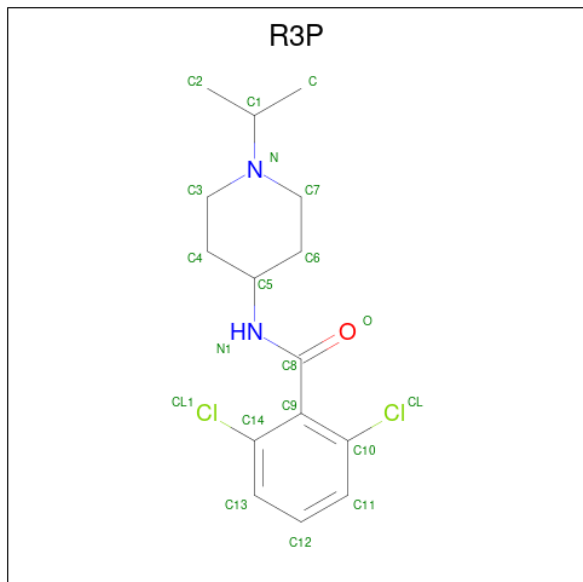
Chain	Residue	Modelled	Actual	Comment	Reference
F	176	SER	-	expression tag	UNP Q1KHK7
F	177	ARG	-	expression tag	UNP Q1KHK7
F	178	LEU	-	expression tag	UNP Q1KHK7
F	179	VAL	-	expression tag	UNP Q1KHK7
F	180	PRO	-	expression tag	UNP Q1KHK7
F	181	ARG	-	expression tag	UNP Q1KHK7

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (CCD ID: 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 (CCD ID: R3P) (formula: $C_{15}H_{20}Cl_2N_2O$) (labeled as "Ligand of Interest" by depositor).



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		

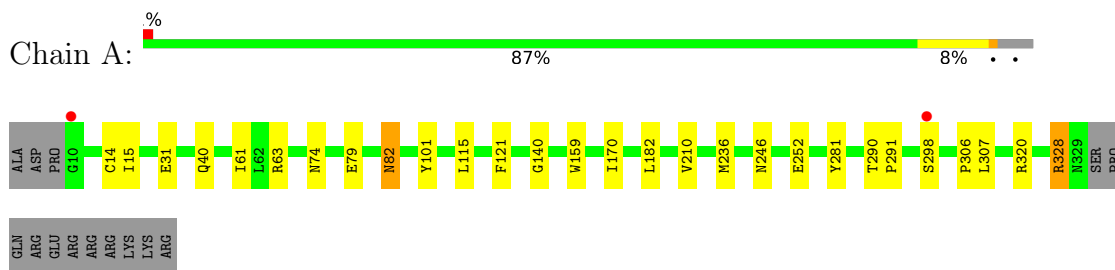
- Molecule 5 is water.

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		

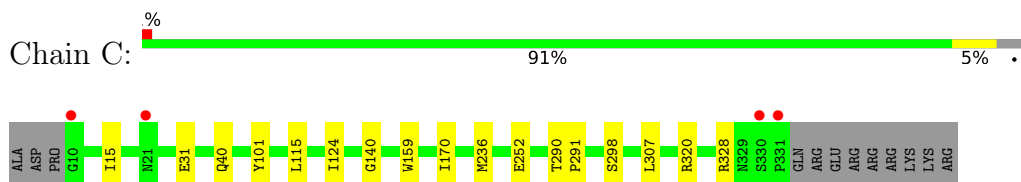
3 Residue-property plots [i](#)

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.

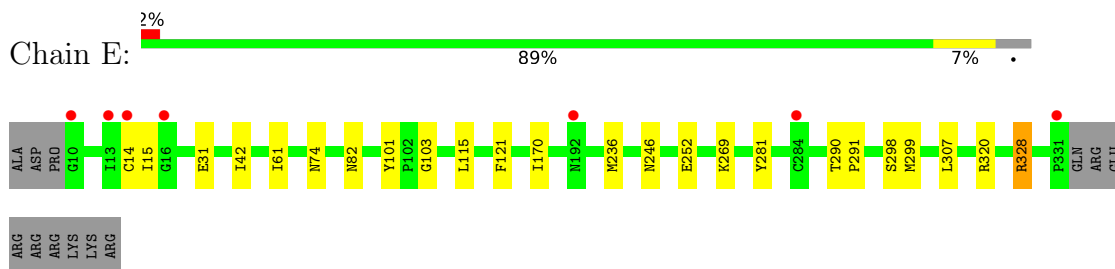
- Molecule 1: Hemagglutinin



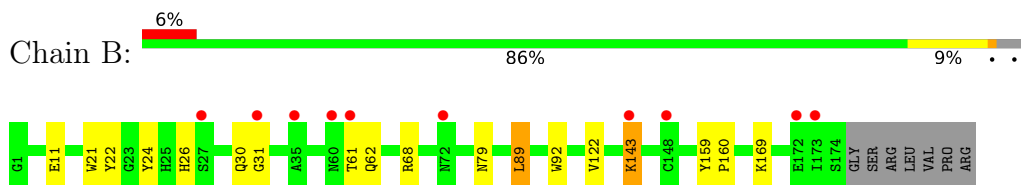
- Molecule 1: Hemagglutinin



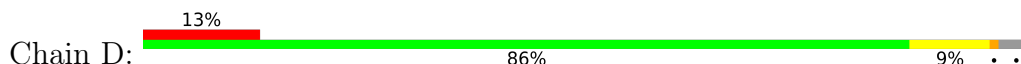
- Molecule 1: Hemagglutinin

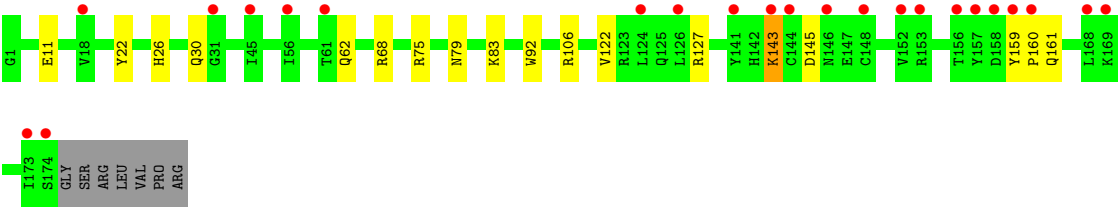


- Molecule 2: Hemagglutinin

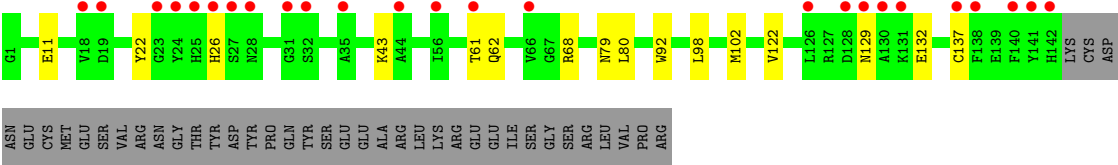


- Molecule 2: Hemagglutinin





● 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 29.47 – 2.20	Depositor EDS
% Data completeness (in resolution range)	93.5 (29.47-2.20) 93.4 (29.47-2.20)	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 (4.63%)	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)	0.34 , 38.0	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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, R3P

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	1.08	0/2612	1.25	4/3545 (0.1%)
1	C	1.06	0/2634	1.25	0/3575
1	E	1.05	0/2621	1.25	2/3559 (0.1%)
2	B	1.06	0/1439	1.51	1/1934 (0.1%)
2	D	1.05	0/1439	1.52	0/1934
2	F	1.03	0/1169	1.47	1/1572 (0.1%)
All	All	1.06	0/11914	1.34	8/16119 (0.0%)

There are no bond length outliers.

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	31	GLY	CA-C-O	-5.63	118.26	122.37
1	A	210	VAL	CA-C-N	5.56	125.43	121.65
1	A	210	VAL	C-N-CA	5.56	125.43	121.65
1	E	74	ASN	CB-CA-C	5.41	117.77	109.50
2	F	129	ASN	CA-CB-CG	5.40	118.00	112.60
1	A	121	PHE	CA-CB-CG	5.16	118.96	113.80
1	E	121	PHE	CA-CB-CG	5.11	118.91	113.80
1	A	74	ASN	CB-CA-C	5.09	117.28	109.50

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	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	8	0
3	A	42	0	39	0	0
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	44	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 (44) 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
1:A:82:ASN:HD22	1:A:82:ASN:C	2.05	0.62
2:B:68:ARG:HE	2:D:79:ASN:HD22	1.49	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

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
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:82:ASN:C	1:A:82:ASN:ND2	2.71	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:42:ILE:C	1:E:299:MET:HG3	2.38	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
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:HE3	2:D:143:LYS:HA	1.99	0.44
2:B:143:LYS:HE3	2:B:143:LYS:HA	2.00	0.44
2:F:98:LEU:HG	2:F:102:MET:HE2	2.00	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: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
2:F:80:LEU:C	2:F:80:LEU:HD23	2.46	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

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	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
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 ⓘ

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%)	279 (97%)	9 (3%)	35	48
1	C	291/300 (97%)	285 (98%)	6 (2%)	47	63
1	E	290/300 (97%)	282 (97%)	8 (3%)	38	52
2	B	149/155 (96%)	140 (94%)	9 (6%)	17	21
2	D	149/155 (96%)	143 (96%)	6 (4%)	28	38
2	F	119/155 (77%)	113 (95%)	6 (5%)	22	28
All	All	1286/1365 (94%)	1242 (97%)	44 (3%)	32	44

All (44) 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	182	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	246	ASN
1	A	290	THR
1	A	298	SER
1	A	320	ARG
1	A	328	ARG
1	C	40	GLN
1	C	115	LEU
1	C	124	ILE
1	C	290	THR
1	C	298	SER
1	C	320	ARG
1	E	14	CYS
1	E	82	ASN
1	E	115	LEU
1	E	246	ASN
1	E	290	THR
1	E	298	SER
1	E	320	ARG
1	E	328	ARG
2	B	11	GLU
2	B	21	TRP
2	B	22	TYR
2	B	26	HIS
2	B	30	GLN
2	B	61	THR
2	B	89	LEU
2	B	143	LYS
2	B	169	LYS
2	D	11	GLU
2	D	22	TYR
2	D	26	HIS
2	D	83	LYS
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

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	82	ASN
1	A	178	ASN
1	C	25	GLN
1	C	40	GLN
1	C	120	HIS
1	C	246	ASN
1	E	25	GLN
1	E	178	ASN
2	B	50	ASN
2	B	79	ASN
2	D	28	ASN
2	D	30	GLN
2	D	79	ASN
2	F	79	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 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	E	402	1	14,14,15	0.62	0	17,19,21	1.62	4 (23%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	R3P	C	404	-	21,21,21	0.88	1 (4%)	29,29,29	1.44	3 (10%)
3	NAG	A	403	1	14,14,15	0.83	0	17,19,21	2.11	4 (23%)
3	NAG	C	403	1	14,14,15	0.48	0	17,19,21	1.38	3 (17%)
3	NAG	C	402	1	14,14,15	0.60	0	17,19,21	1.68	4 (23%)
3	NAG	A	401	1	14,14,15	0.69	0	17,19,21	1.95	1 (5%)
4	R3P	E	403	-	21,21,21	0.95	2 (9%)	29,29,29	1.68	8 (27%)
4	R3P	A	404	-	21,21,21	0.98	2 (9%)	29,29,29	1.55	6 (20%)
3	NAG	C	401	1	14,14,15	0.45	0	17,19,21	0.99	1 (5%)
3	NAG	E	401	1	14,14,15	0.69	0	17,19,21	2.08	5 (29%)
3	NAG	A	402	1	14,14,15	1.10	1 (7%)	17,19,21	2.97	8 (47%)

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	E	402	1	-	2/6/23/26	0/1/1/1
4	R3P	C	404	-	-	11/12/22/22	0/2/2/2
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
3	NAG	C	402	1	-	1/6/23/26	0/1/1/1
3	NAG	A	401	1	-	0/6/23/26	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	C	401	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	402	1	-	4/6/23/26	0/1/1/1

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	402	NAG	C1-C2	2.89	1.56	1.52
4	A	404	R3P	C10-CL	2.74	1.80	1.73
4	A	404	R3P	C14-CL1	2.71	1.80	1.73
4	E	403	R3P	C10-CL	2.63	1.79	1.73
4	C	404	R3P	C14-CL1	2.58	1.79	1.73
4	E	403	R3P	C14-CL1	2.41	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	7.33	122.02	112.19
3	A	402	NAG	C1-O5-C5	6.55	120.97	112.19
3	A	402	NAG	C2-N2-C7	5.71	130.55	122.90
3	A	403	NAG	C1-O5-C5	5.67	119.78	112.19
3	A	402	NAG	C4-C3-C2	4.60	117.76	111.02
3	A	403	NAG	C2-N2-C7	4.30	128.66	122.90
4	C	404	R3P	C6-C5-N1	4.04	118.64	110.57
3	E	401	NAG	O5-C1-C2	-4.03	105.05	111.29
3	C	402	NAG	C1-O5-C5	3.84	117.33	112.19
3	E	401	NAG	C1-O5-C5	3.81	117.29	112.19
3	A	402	NAG	C3-C4-C5	3.60	116.76	110.23
4	E	403	R3P	C4-C3-N	3.59	116.95	111.42
3	A	402	NAG	C8-C7-N2	3.51	121.94	116.12
3	E	401	NAG	C3-C4-C5	-3.51	103.87	110.23
4	A	404	R3P	C6-C5-N1	3.51	117.58	110.57
3	A	402	NAG	O5-C1-C2	3.31	116.42	111.29
3	E	402	NAG	C1-C2-N2	3.30	115.63	110.43
4	A	404	R3P	C3-N-C1	3.22	118.15	113.01
3	E	402	NAG	C3-C4-C5	3.13	115.90	110.23
3	E	401	NAG	C4-C3-C2	3.06	115.50	111.02
3	C	403	NAG	O5-C5-C6	3.01	113.52	107.66
3	C	403	NAG	C1-C2-N2	2.99	115.15	110.43
3	C	402	NAG	C1-C2-N2	2.96	115.10	110.43
4	A	404	R3P	C5-N1-C8	2.93	128.09	122.56
4	E	403	R3P	C5-N1-C8	2.78	127.82	122.56
4	C	404	R3P	C7-C6-C5	-2.76	106.31	110.67
4	E	403	R3P	C10-C9-C8	-2.76	117.96	121.29
3	C	402	NAG	O5-C1-C2	-2.69	107.12	111.29
4	E	403	R3P	C3-N-C1	2.68	117.28	113.01
4	E	403	R3P	C7-N-C3	2.65	114.18	109.13
3	E	402	NAG	C4-C3-C2	2.57	114.79	111.02
3	A	403	NAG	C1-C2-N2	2.57	114.48	110.43
3	A	403	NAG	C3-C4-C5	2.54	114.84	110.23
3	C	402	NAG	C3-C4-C5	2.53	114.83	110.23
4	A	404	R3P	C6-C7-N	2.49	115.25	111.42
4	E	403	R3P	C7-C6-C5	-2.48	106.75	110.67
4	A	404	R3P	C-C1-N	-2.41	106.80	112.26
3	A	402	NAG	C1-C2-N2	2.37	114.17	110.43
3	C	403	NAG	C1-O5-C5	2.36	115.34	112.19
4	E	403	R3P	C-C1-N	-2.23	107.20	112.26
4	C	404	R3P	C3-C4-C5	2.22	114.18	110.67
3	A	402	NAG	O7-C7-C8	-2.21	118.11	122.05

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	404	R3P	C7-N-C1	-2.20	109.49	113.01
3	E	402	NAG	C2-N2-C7	2.12	125.74	122.90
4	E	403	R3P	C6-C5-N1	2.09	114.75	110.57
3	E	401	NAG	O7-C7-N2	2.07	125.64	121.98
3	C	401	NAG	C1-O5-C5	2.07	114.96	112.19

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	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	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
4	E	403	R3P	O-C8-C9-C14
4	E	403	R3P	C6-C5-N1-C8
4	E	403	R3P	C2-C1-N-C3
4	E	403	R3P	C-C1-N-C3
4	E	403	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
3	C	402	NAG	O5-C5-C6-O6
4	A	404	R3P	C6-C5-N1-C8
4	E	403	R3P	N1-C8-C9-C14
3	E	401	NAG	C4-C5-C6-O6
3	E	402	NAG	C4-C5-C6-O6
3	E	401	NAG	O5-C5-C6-O6
3	C	403	NAG	O5-C5-C6-O6
4	C	404	R3P	O-C8-C9-C14
4	E	403	R3P	C2-C1-N-C7
3	C	401	NAG	C4-C5-C6-O6

Continued on next page...

Continued from previous page...

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

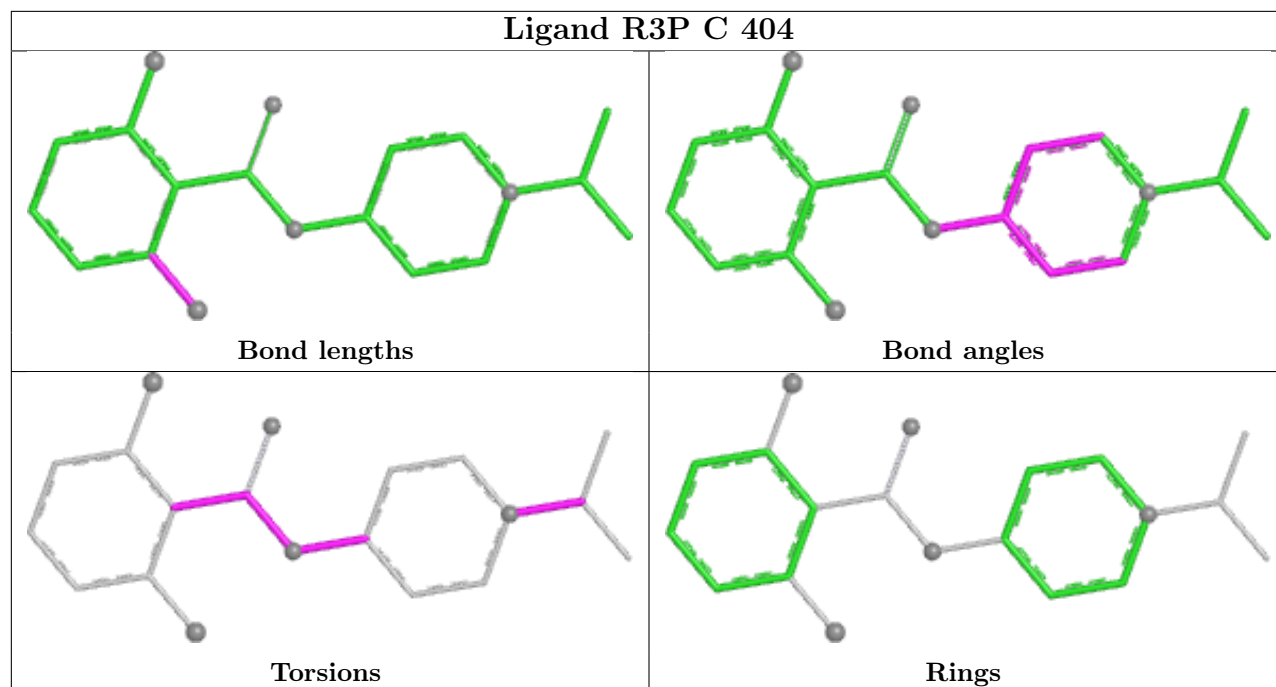
All (1) ring outliers are listed below:

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

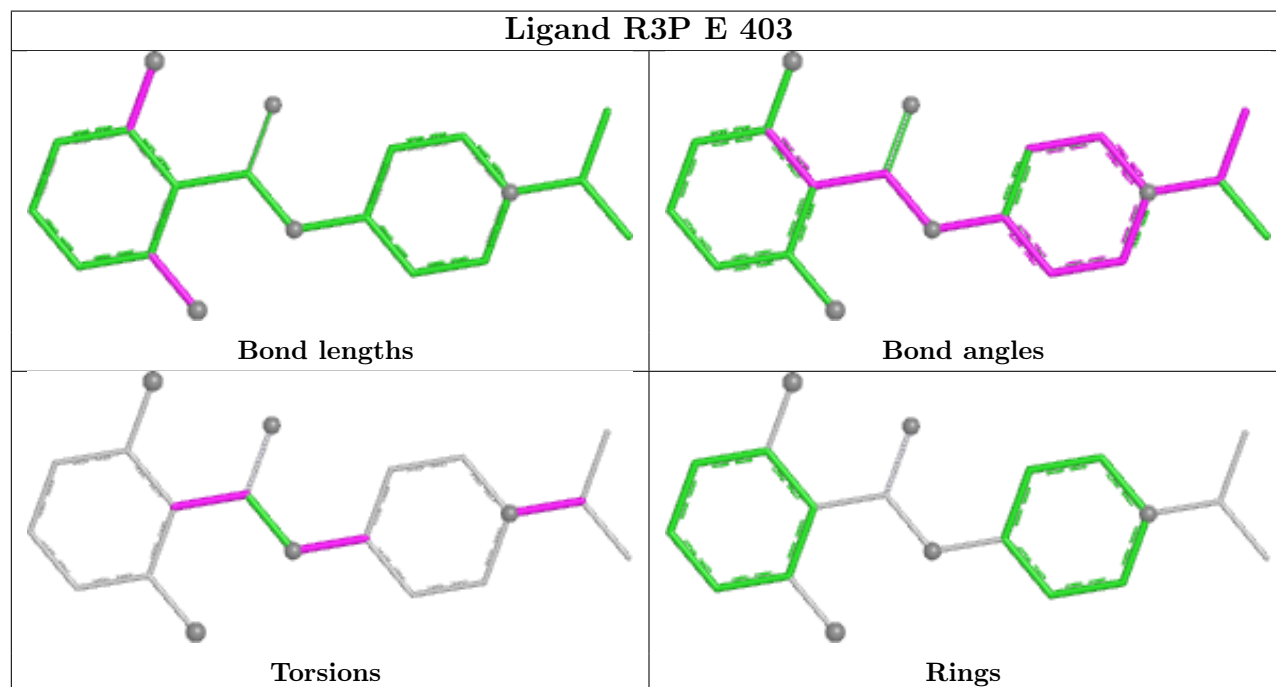
No monomer is involved in short contacts.

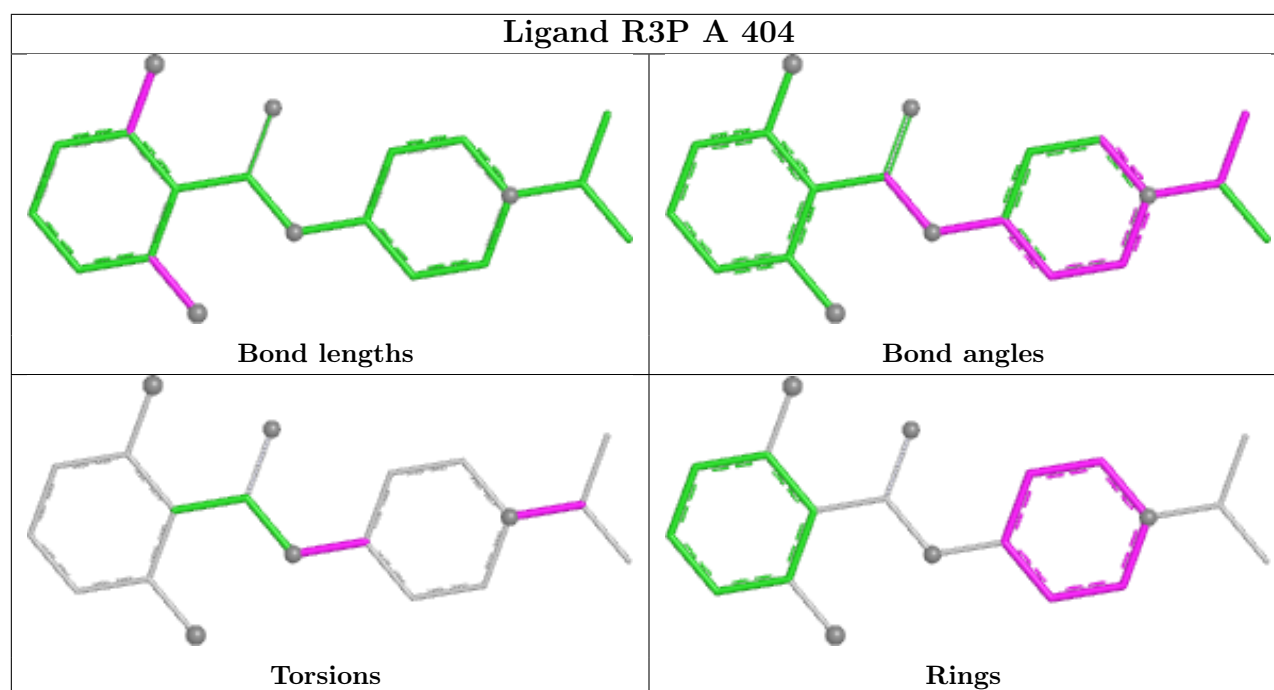
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.

Ligand R3P C 404



Ligand R3P E 403





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%)	0.03	2 (0%) 85 83	20, 44, 69, 90	1 (0%)
1	C	322/334 (96%)	-0.09	4 (1%) 76 74	19, 40, 67, 139	2 (0%)
1	E	322/334 (96%)	0.14	7 (2%) 62 59	23, 45, 78, 126	1 (0%)
2	B	174/181 (96%)	0.61	10 (5%) 29 26	26, 60, 97, 110	0
2	D	174/181 (96%)	0.83	23 (13%) 7 5	26, 66, 102, 120	0
2	F	142/181 (78%)	0.99	25 (17%) 4 3	24, 71, 115, 146	0
All	All	1454/1545 (94%)	0.29	71 (4%) 35 31	19, 49, 96, 146	4 (0%)

All (71) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	10	GLY	4.5
1	C	331	PRO	4.4
2	D	173	ILE	4.1
1	C	10	GLY	4.0
2	F	140	PHE	3.8
1	A	10	GLY	3.8
2	F	35	ALA	3.8
2	D	31	GLY	3.3
2	D	160	PRO	3.3
2	D	174	SER	3.3
2	B	173	ILE	3.2
2	F	44	ALA	3.2
1	E	14	CYS	3.1
2	F	27	SER	3.1
2	F	126	LEU	3.1
2	B	27	SER	3.0
2	D	18	VAL	3.0
2	D	158	ASP	3.0
2	D	157	TYR	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	153	ARG	2.9
2	B	72	ASN	2.9
1	E	13	ILE	2.9
2	B	31	GLY	2.7
2	D	143	LYS	2.7
2	F	130	ALA	2.7
2	F	18	VAL	2.7
2	B	61	THR	2.7
2	F	142	HIS	2.6
2	F	137	CYS	2.6
2	F	128	ASP	2.6
2	D	61	THR	2.6
2	F	138	PHE	2.6
2	D	141	TYR	2.5
2	F	129	ASN	2.5
2	D	126	LEU	2.5
2	F	26	HIS	2.5
2	F	19	ASP	2.4
2	D	169	LYS	2.4
1	E	284	CYS	2.4
2	F	141	TYR	2.4
2	D	144	CYS	2.4
2	D	168	LEU	2.4
2	F	61	THR	2.3
2	D	148	CYS	2.3
1	E	331	PRO	2.3
2	D	146	ASN	2.3
2	F	25	HIS	2.3
2	F	28	ASN	2.3
2	B	172	GLU	2.3
2	F	56	ILE	2.3
2	F	23	GLY	2.2
1	C	330	SER	2.2
1	E	192	ASN	2.2
1	A	298	SER	2.2
2	D	152	VAL	2.2
2	B	148	CYS	2.2
2	F	66	VAL	2.1
2	D	45	ILE	2.1
2	F	24	TYR	2.1
2	D	156	THR	2.1
2	F	32	SER	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	60	ASN	2.1
2	D	56	ILE	2.1
2	D	159	TYR	2.1
1	E	16	GLY	2.1
2	F	31	GLY	2.0
2	B	35	ALA	2.0
1	C	21	ASN	2.0
2	D	124	LEU	2.0
2	B	143	LYS	2.0
2	F	131	LYS	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.

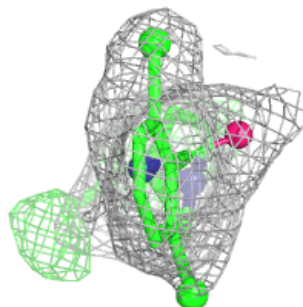
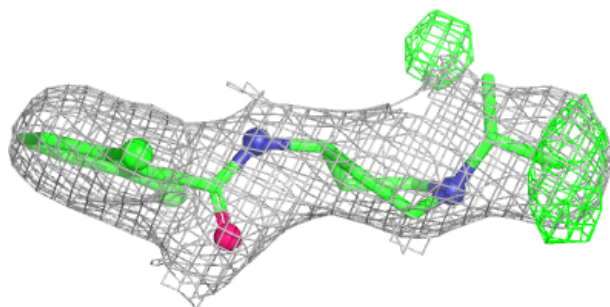
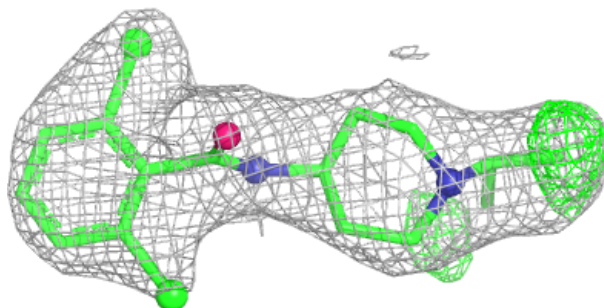
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
3	NAG	A	403	14/15	0.51	0.18	98,116,126,127	0
3	NAG	A	402	14/15	0.67	0.16	112,121,128,128	0
3	NAG	E	402	14/15	0.67	0.16	100,116,123,124	0
3	NAG	C	403	14/15	0.77	0.15	99,115,118,120	0
3	NAG	C	402	14/15	0.79	0.15	87,104,117,118	0
4	R3P	A	404	20/20	0.81	0.18	67,85,94,114	0
4	R3P	C	404	20/20	0.83	0.20	81,89,97,107	0
4	R3P	E	403	20/20	0.85	0.16	86,103,108,131	0
3	NAG	C	401	14/15	0.88	0.11	63,70,81,82	0
3	NAG	A	401	14/15	0.93	0.08	44,49,52,53	0
3	NAG	E	401	14/15	0.94	0.09	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.

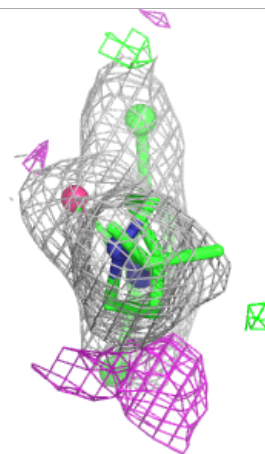
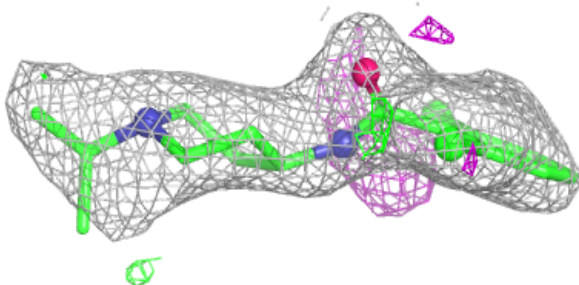
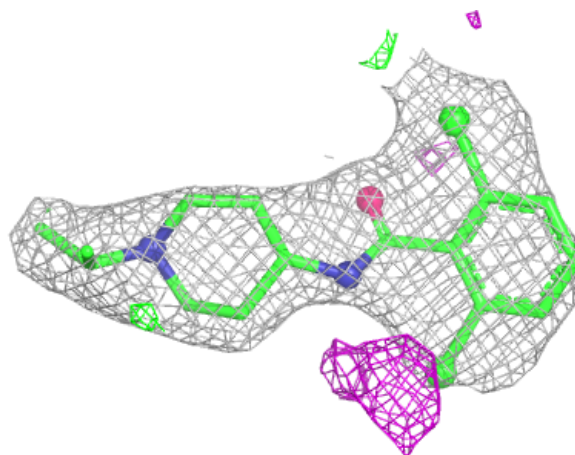
Electron density around R3P A 404:

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



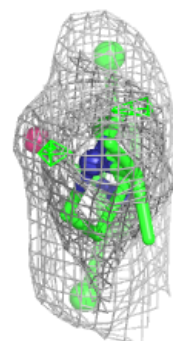
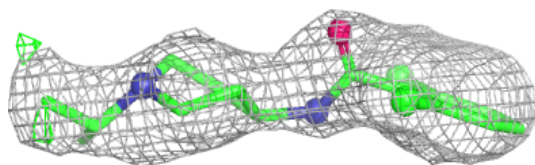
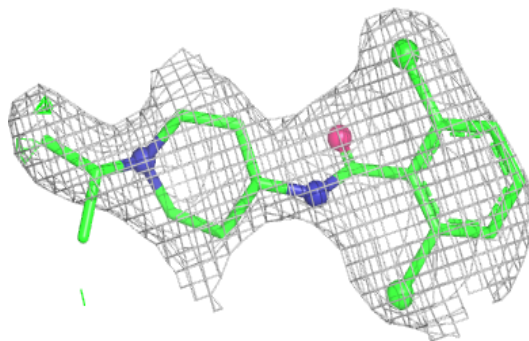
Electron density around R3P C 404:

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



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)



6.5 Other polymers [i](#)

There are no such residues in this entry.