Full wwPDB X-ray Structure Validation Report (i)

Nov 13, 2023 – 04:25 PM JST

PDB ID	:	5XL9
Title	:	The structure of hemagglutinin G228S mutant from an avian-origin H4N6 influenza virus in complex with avian receptor analog LSTa
Authors	:	Song, H.; Qi, J.; Gao, F.G.
Deposited on	:	2017-05-10
Resolution	:	2.39 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 (i) 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 (i)) were used in the production of this report:

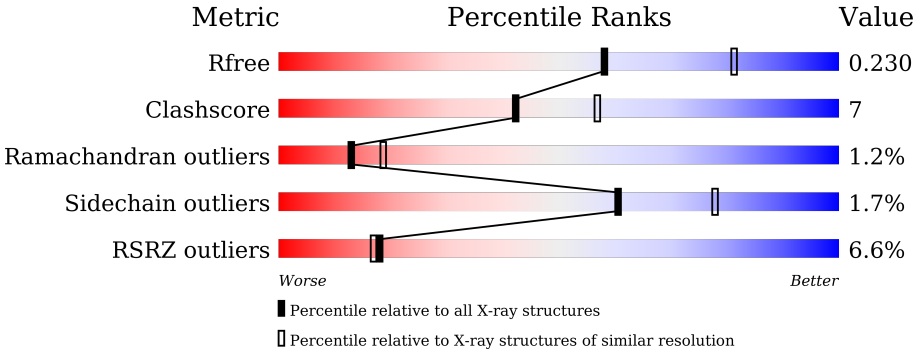
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.39 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 <=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	503	6% 80% 14% • • •
1	B	503	7% 82% 14% • •
1	C	503	6% 83% 12% • •

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 12019 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.

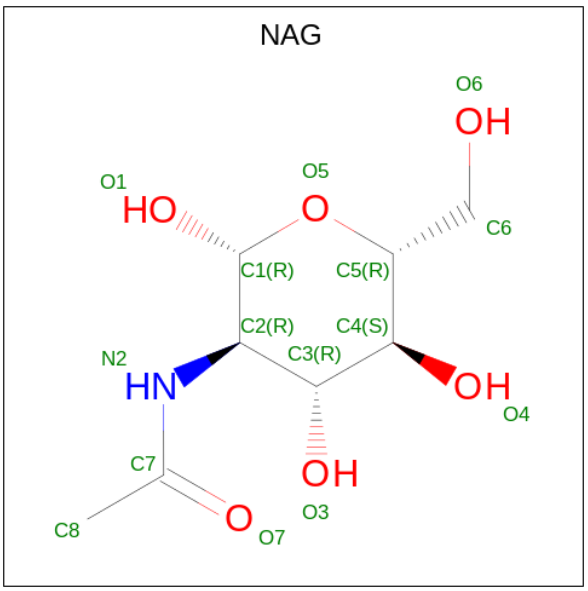
- Molecule 1 is a protein called Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	486	3833	2390	682	745	16	0	0	0
1	B	486	3833	2390	682	745	16	0	0	0
1	C	486	3833	2390	682	745	16	0	0	0

There are 3 discrepancies between the modelled and reference sequences:

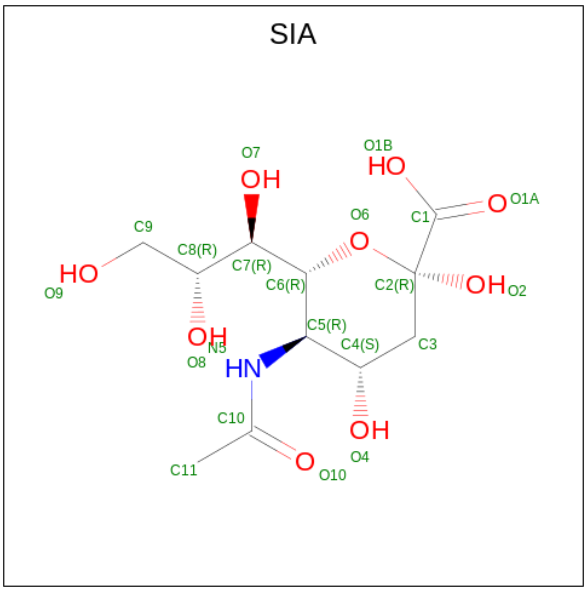
Chain	Residue	Modelled	Actual	Comment	Reference
A	225	SER	GLY	engineered mutation	UNP A3KF09
B	225	SER	GLY	engineered mutation	UNP A3KF09
C	225	SER	GLY	engineered mutation	UNP A3KF09

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



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

- Molecule 3 is N-acetyl-alpha-neuraminic acid (three-letter code: SIA) (formula: $C_{11}H_{19}NO_9$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			21	11	1	9		

- Molecule 4 is water.

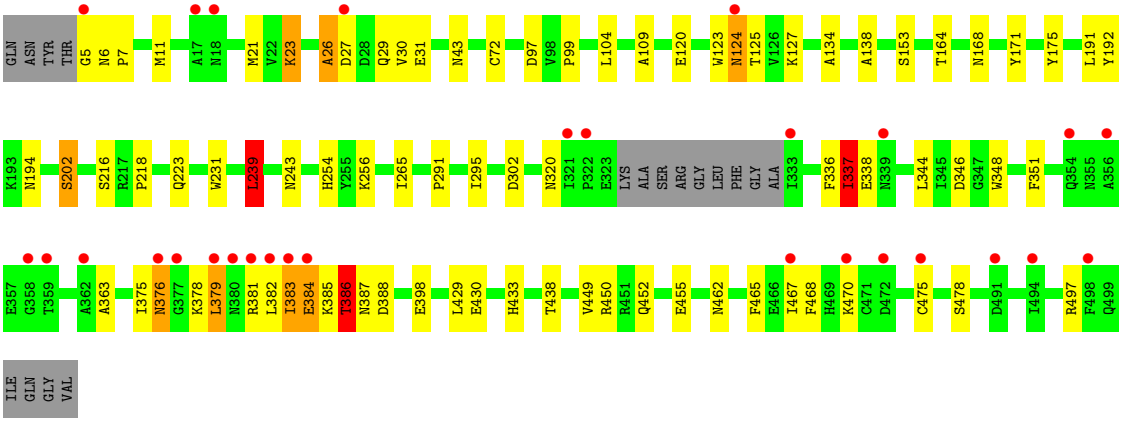
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	131	Total	O	0	0
			131	131		
4	B	144	Total	O	0	0
			144	144		
4	C	140	Total	O	0	0
			140	140		

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.

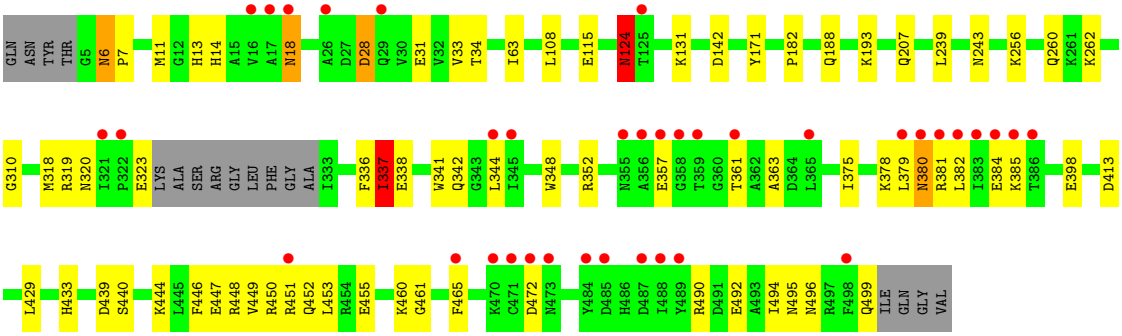
• Molecule 1: Hemagglutinin

Chain A:



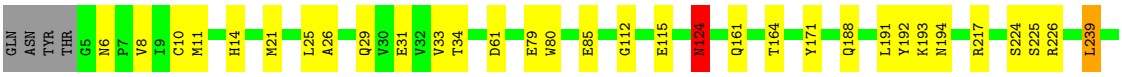
• Molecule 1: Hemagglutinin

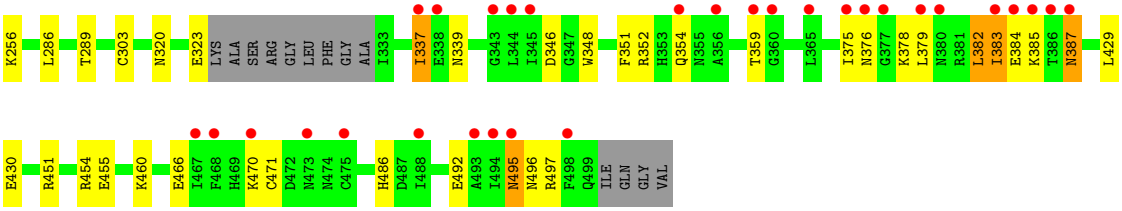
Chain B:



• Molecule 1: Hemagglutinin

Chain C:




K256
L286
T289
C303
N320
E323
LYS
ALA
SER
ARG
GLY
LEU
PHE
GLY
ALA
I333
I337
E338
N339
G343
L344
I345
D346
G347
W348
F351
R352
H353
Q354
N355
A356
T359
G360
L365
I375
N376
G377
K378
L379
N380
R381
L382
I383
E384
K385
T386
N387
L429
E430
R451
R454
E455
K460
E466
I467
F468
H469
K470
C471
D472
N473
N474
C475
H486
D487
I488
E492
A493
I494
N495
N496
R497
F498
Q499
ILE
GLN
GLY
VAL

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	65.12Å 115.16Å 132.03Å 90.00° 102.23° 90.00°	Depositor
Resolution (Å)	46.66 – 2.39 46.66 – 2.39	Depositor EDS
% Data completeness (in resolution range)	99.5 (46.66-2.39) 99.5 (46.66-2.39)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.11 (at 2.39Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.183 , 0.229 0.186 , 0.230	Depositor DCC
R_{free} test set	3794 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	39.5	Xtriage
Anisotropy	0.111	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 45.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.021 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	12019	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.88% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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: NAG, SIA

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	0.33	0/3910	0.54	1/5302 (0.0%)
1	B	0.34	0/3910	0.55	1/5302 (0.0%)
1	C	0.34	0/3910	0.51	0/5302
All	All	0.34	0/11730	0.53	2/15906 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	1
All	All	0	2

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	239	LEU	CA-CB-CG	5.33	127.57	115.30
1	B	453	LEU	CB-CA-C	5.19	120.05	110.20

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	383	ILE	Peptide
1	B	380	ASN	Peptide

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	3833	0	3700	60	0
1	B	3833	0	3700	55	0
1	C	3833	0	3700	50	0
2	A	28	0	26	1	0
2	B	28	0	26	2	0
2	C	28	0	26	0	0
3	A	21	0	18	1	0
4	A	131	0	0	9	0
4	B	144	0	0	6	0
4	C	140	0	0	3	0
All	All	12019	0	11196	158	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 7.

All (158) 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:C:21:MET:HB3	1:C:29:GLN:HG3	1.47	0.96
1:A:123:TRP:O	4:A:702:HOH:O	1.95	0.84
1:C:14:HIS:HD2	1:C:348:TRP:HA	1.45	0.81
1:C:14:HIS:CD2	1:C:348:TRP:HA	2.17	0.78
1:B:413:ASP:OD2	4:B:702:HOH:O	2.01	0.78
2:A:602:NAG:O4	4:A:701:HOH:O	1.93	0.77
1:C:323:GLU:HG2	1:C:339:ASN:HD21	1.50	0.77
1:A:302:ASP:OD1	4:A:703:HOH:O	2.04	0.74
1:B:492:GLU:O	1:B:496:ASN:ND2	2.21	0.73
1:B:495:ASN:ND2	1:B:499:GLN:OE1	2.22	0.72
1:C:79:GLU:OE1	4:C:702:HOH:O	2.09	0.70
1:C:188:GLN:O	4:C:701:HOH:O	2.09	0.68
1:B:379:LEU:HG	1:B:379:LEU:O	1.91	0.68
1:A:376:ASN:HB2	1:A:379:LEU:HD13	1.75	0.67
1:C:470:LYS:HD2	1:C:471:CYS:N	2.10	0.67
1:B:461:GLY:HA2	1:C:451:ARG:HD2	1.77	0.67
1:A:43:ASN:HB3	1:A:295:ILE:HD13	1.78	0.66

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:33:VAL:HG12	1:C:34:THR:HG23	1.78	0.65
1:C:191:LEU:N	4:C:701:HOH:O	2.22	0.65
1:C:171:TYR:CZ	1:C:256:LYS:HD2	2.31	0.65
1:C:495:ASN:CG	1:C:496:ASN:H	2.00	0.64
1:A:216:SER:O	1:B:243:ASN:ND2	2.31	0.63
1:A:171:TYR:CZ	1:A:256:LYS:HD2	2.34	0.63
1:A:497:ARG:NH2	1:B:455:GLU:OE2	2.32	0.62
1:B:352:ARG:HD2	1:B:361:THR:OG1	2.00	0.62
1:A:72:CYS:O	4:A:704:HOH:O	2.16	0.61
1:B:31:GLU:HG2	1:B:320:ASN:HB3	1.84	0.60
1:A:99:PRO:O	4:A:706:HOH:O	2.17	0.60
1:A:375:ILE:O	1:A:378:LYS:HG2	2.03	0.58
1:A:386:THR:HG22	1:A:388:ASP:HB2	1.85	0.58
1:A:376:ASN:O	1:A:379:LEU:HB2	2.04	0.58
2:B:602:NAG:O4	4:B:701:HOH:O	1.97	0.58
1:A:455:GLU:HG3	1:A:497:ARG:NH1	2.19	0.57
1:B:447:GLU:O	1:B:451:ARG:HG3	2.05	0.56
1:A:202:SER:OG	4:A:705:HOH:O	2.17	0.56
1:A:337:ILE:HG13	1:A:462:ASN:HA	1.87	0.55
1:B:429:LEU:HD21	1:C:429:LEU:HD13	1.88	0.55
1:C:31:GLU:HG2	1:C:320:ASN:HB3	1.89	0.55
1:C:384:GLU:HG2	1:C:385:LYS:HG3	1.88	0.54
1:A:398:GLU:HG3	4:A:762:HOH:O	2.07	0.54
1:C:354:GLN:HG2	1:C:359:THR:HG22	1.88	0.54
1:C:384:GLU:HG2	1:C:385:LYS:N	2.23	0.53
1:B:375:ILE:O	1:B:378:LYS:HG2	2.08	0.53
1:C:164:THR:HG23	1:C:239:LEU:HD23	1.90	0.53
1:B:336:PHE:O	1:B:337:ILE:HG12	2.09	0.53
1:B:375:ILE:HA	1:B:378:LYS:HE3	1.90	0.53
1:C:31:GLU:OE2	1:C:320:ASN:ND2	2.42	0.53
1:A:429:LEU:HD13	1:C:429:LEU:HD21	1.90	0.53
1:B:131:LYS:HG2	1:B:142:ASP:HA	1.91	0.52
1:A:138:ALA:O	4:A:707:HOH:O	2.19	0.52
1:A:31:GLU:HG2	1:A:320:ASN:HB3	1.92	0.52
1:B:193:LYS:NZ	4:B:706:HOH:O	2.29	0.52
1:C:303:CYS:O	1:C:387:ASN:ND2	2.38	0.52
1:A:455:GLU:HG3	1:A:497:ARG:HH12	1.75	0.52
1:C:11:MET:HE1	1:C:351:PHE:CE2	2.44	0.52
1:C:192:TYR:O	1:C:194:ASN:N	2.40	0.51
1:A:433:HIS:HE1	1:C:26:ALA:HB2	1.76	0.51
1:C:454:ARG:HD3	1:C:486:HIS:CD2	2.46	0.51

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:11:MET:HE1	1:B:449:VAL:HG21	1.91	0.51
1:C:495:ASN:OD1	1:C:496:ASN:N	2.36	0.50
1:C:492:GLU:O	1:C:492:GLU:HG2	2.11	0.50
1:B:319:ARG:NH1	1:B:439:ASP:OD2	2.45	0.50
1:A:192:TYR:O	1:A:194:ASN:N	2.45	0.49
1:A:134:ALA:HB2	1:A:223:GLN:HG2	1.94	0.49
1:C:455:GLU:O	1:C:497:ARG:NH1	2.43	0.49
1:B:341:TRP:CH2	1:B:352:ARG:HB2	2.48	0.49
1:A:337:ILE:O	1:A:338:GLU:HG3	2.13	0.49
1:B:11:MET:HG3	1:B:446:PHE:HD1	1.78	0.49
1:C:323:GLU:HG2	1:C:339:ASN:ND2	2.22	0.49
1:C:217:ARG:HB2	1:C:224:SER:O	2.13	0.49
1:A:97:ASP:OD2	1:B:207:GLN:NE2	2.44	0.49
1:C:378:LYS:O	1:C:382:LEU:HB2	2.13	0.49
1:B:323:GLU:H	1:B:342:GLN:HE21	1.61	0.48
1:C:375:ILE:O	1:C:378:LYS:HG2	2.14	0.48
1:C:495:ASN:CG	1:C:496:ASN:N	2.66	0.48
1:A:104:LEU:HB2	1:A:231:TRP:CZ3	2.50	0.47
1:A:348:TRP:HZ3	1:A:375:ILE:HD12	1.78	0.47
1:B:450:ARG:HB2	1:B:465:PHE:CZ	2.48	0.47
1:A:109:ALA:HB1	1:A:265:ILE:HB	1.96	0.47
1:A:127:LYS:HG3	1:A:153:SER:HA	1.95	0.47
1:B:344:LEU:HD11	1:B:363:ALA:HB2	1.95	0.47
1:A:191:LEU:HD11	3:A:603:SIA:H91	1.96	0.47
1:A:375:ILE:HD11	1:A:438:THR:HG23	1.95	0.47
1:B:260:GLN:HB3	1:B:262:LYS:HE2	1.97	0.47
1:B:379:LEU:HA	1:B:382:LEU:HB2	1.96	0.47
1:B:455:GLU:HG3	1:B:490:ARG:HH21	1.80	0.47
1:B:460:LYS:HE2	1:B:460:LYS:HB3	1.68	0.47
1:A:344:LEU:HD11	1:A:363:ALA:HB2	1.96	0.47
1:C:375:ILE:HA	1:C:378:LYS:HE3	1.96	0.46
1:B:450:ARG:HB2	1:B:465:PHE:HZ	1.80	0.46
1:B:14:HIS:ND1	1:B:348:TRP:HA	2.30	0.46
1:A:375:ILE:HA	1:A:378:LYS:HE3	1.96	0.46
1:B:18:ASN:OD1	1:B:18:ASN:C	2.53	0.46
1:C:10:CYS:N	1:C:352:ARG:O	2.43	0.46
1:A:337:ILE:HD13	1:A:337:ILE:HA	1.82	0.46
1:A:348:TRP:CZ3	1:A:375:ILE:HD12	2.51	0.46
1:C:376:ASN:HA	1:C:379:LEU:HG	1.98	0.46
1:A:336:PHE:O	1:A:337:ILE:HG12	2.16	0.46
1:B:381:ARG:HA	1:B:384:GLU:HB3	1.97	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:475:CYS:O	1:A:478:SER:OG	2.27	0.45
1:B:337:ILE:O	1:B:338:GLU:HG3	2.17	0.45
1:B:357:GLU:OE2	1:B:472:ASP:HB2	2.16	0.45
1:C:378:LYS:HD3	1:C:430:GLU:OE1	2.16	0.45
1:C:460:LYS:HD2	1:C:466:GLU:HB2	1.99	0.45
1:B:243:ASN:ND2	4:B:712:HOH:O	2.38	0.45
1:B:413:ASP:OD1	4:B:703:HOH:O	2.21	0.45
1:A:26:ALA:HB2	1:B:433:HIS:HE1	1.83	0.44
1:A:378:LYS:HB3	1:A:378:LYS:HE2	1.85	0.44
1:A:164:THR:HG23	1:A:239:LEU:HD23	1.98	0.44
1:A:124:ASN:HB3	1:A:125:THR:H	1.45	0.44
1:A:218:PRO:HB3	1:B:239:LEU:HD11	2.00	0.44
1:B:440:SER:O	1:B:444:LYS:HG3	2.18	0.43
1:C:124:ASN:HB3	1:C:161:GLN:HE22	1.83	0.43
1:A:21:MET:HB3	1:A:29:GLN:HG3	2.00	0.43
1:C:225:SER:O	1:C:226:ARG:NH1	2.41	0.43
1:A:27:ASP:HB2	1:A:30:VAL:HG23	2.01	0.43
1:A:378:LYS:HB2	1:A:382:LEU:HD23	2.00	0.43
1:C:460:LYS:HE2	1:C:460:LYS:HB3	1.88	0.43
1:A:7:PRO:HD2	1:A:467:ILE:O	2.19	0.43
1:A:375:ILE:HD11	1:A:438:THR:CG2	2.48	0.43
1:B:171:TYR:CZ	1:B:256:LYS:HD2	2.53	0.43
1:A:291:PRO:HB3	1:A:383:ILE:HB	2.01	0.43
1:B:33:VAL:HG12	1:B:34:THR:HG23	2.00	0.42
1:A:23:LYS:HE2	1:A:23:LYS:HB2	1.82	0.42
1:B:13:HIS:HB2	1:B:318:MET:SD	2.59	0.42
1:B:378:LYS:HE2	1:B:378:LYS:HB3	1.80	0.42
1:A:382:LEU:HD21	1:A:430:GLU:OE2	2.19	0.42
1:B:28:ASP:OD1	1:B:28:ASP:N	2.51	0.42
1:A:450:ARG:HB2	1:A:465:PHE:HZ	1.85	0.42
1:A:11:MET:HE1	1:A:351:PHE:CE2	2.53	0.42
1:A:175:TYR:CE2	1:A:254:HIS:HB3	2.55	0.42
1:C:124:ASN:CB	1:C:161:GLN:HE22	2.32	0.42
1:B:310:GLY:O	2:B:602:NAG:O6	2.37	0.42
1:B:448:ARG:O	1:B:452:GLN:HG3	2.20	0.42
1:B:182:PRO:HG2	1:B:188:GLN:OE1	2.19	0.41
1:A:5:GLY:N	1:A:468:PHE:HA	2.35	0.41
1:C:346:ASP:N	1:C:346:ASP:OD1	2.51	0.41
1:B:6:ASN:HB3	1:B:7:PRO:HD3	2.01	0.41
1:C:80:TRP:CE2	1:C:112:GLY:HA2	2.56	0.41
1:B:115:GLU:CD	1:B:256:LYS:HE2	2.41	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:25:LEU:HD11	1:C:429:LEU:HD23	2.00	0.41
1:B:14:HIS:CD2	1:B:33:VAL:HG21	2.55	0.41
1:C:61:ASP:HB2	1:C:85:GLU:OE2	2.21	0.41
1:A:385:LYS:HD2	1:A:387:ASN:H	1.86	0.41
1:A:125:THR:HG22	1:A:125:THR:O	2.21	0.41
1:A:168:ASN:ND2	4:A:724:HOH:O	2.53	0.41
1:A:381:ARG:HA	1:A:384:GLU:HG3	2.02	0.41
1:B:384:GLU:HG3	1:B:385:LYS:N	2.35	0.41
1:B:490:ARG:O	1:B:494:ILE:HG12	2.20	0.41
1:C:8:VAL:HG22	1:C:466:GLU:HA	2.02	0.41
1:C:115:GLU:CD	1:C:256:LYS:HE2	2.42	0.41
1:A:449:VAL:HA	1:A:452:GLN:HG3	2.03	0.40
1:C:289:THR:HB	1:C:383:ILE:HD11	2.03	0.40
1:B:63:ILE:HD13	1:B:108:LEU:HD12	2.03	0.40
1:C:470:LYS:HD2	1:C:470:LYS:C	2.41	0.40
1:A:376:ASN:OD1	1:A:376:ASN:C	2.60	0.40
1:B:6:ASN:ND2	4:B:726:HOH:O	2.53	0.40
1:B:124:ASN:HD22	1:B:124:ASN:HA	1.70	0.40

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	482/503 (96%)	455 (94%)	20 (4%)	7 (2%)	10 14
1	B	482/503 (96%)	454 (94%)	25 (5%)	3 (1%)	25 36
1	C	482/503 (96%)	457 (95%)	18 (4%)	7 (2%)	10 14
All	All	1446/1509 (96%)	1366 (94%)	63 (4%)	17 (1%)	13 19

All (17) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	124	ASN
1	A	337	ILE
1	A	379	LEU
1	A	386	THR
1	B	124	ASN
1	C	124	ASN
1	C	337	ILE
1	B	337	ILE
1	C	387	ASN
1	A	6	ASN
1	A	384	GLU
1	B	6	ASN
1	C	193	LYS
1	C	495	ASN
1	A	26	ALA
1	C	6	ASN
1	C	383	ILE

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	421/433 (97%)	411 (98%)	10 (2%)	49	68
1	B	421/433 (97%)	415 (99%)	6 (1%)	67	82
1	C	421/433 (97%)	416 (99%)	5 (1%)	71	85
All	All	1263/1299 (97%)	1242 (98%)	21 (2%)	60	78

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	23	LYS
1	A	120	GLU
1	A	202	SER
1	A	239	LEU
1	A	243	ASN
1	A	337	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	346	ASP
1	A	376	ASN
1	A	386	THR
1	A	470	LYS
1	B	18	ASN
1	B	28	ASP
1	B	124	ASN
1	B	337	ILE
1	B	380	ASN
1	B	398	GLU
1	C	124	ASN
1	C	239	LEU
1	C	286	LEU
1	C	337	ILE
1	C	382	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	14	HIS
1	A	243	ASN
1	A	452	GLN
1	B	124	ASN
1	B	495	ASN
1	C	14	HIS
1	C	339	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

7 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	NAG	A	602	1	14,14,15	0.22	0	17,19,21	0.76	1 (5%)
2	NAG	C	602	1	14,14,15	0.22	0	17,19,21	0.78	1 (5%)
2	NAG	B	602	1	14,14,15	0.43	0	17,19,21	0.55	0
2	NAG	C	601	1	14,14,15	0.35	0	17,19,21	0.37	0
2	NAG	A	601	1	14,14,15	0.30	0	17,19,21	0.34	0
3	SIA	A	603	-	21,21,21	0.95	1 (4%)	25,31,31	1.08	1 (4%)
2	NAG	B	601	1	14,14,15	0.72	0	17,19,21	0.49	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	NAG	A	602	1	-	2/6/23/26	0/1/1/1
2	NAG	C	602	1	-	2/6/23/26	0/1/1/1
2	NAG	B	602	1	-	4/6/23/26	0/1/1/1
2	NAG	C	601	1	-	2/6/23/26	0/1/1/1
2	NAG	A	601	1	-	0/6/23/26	0/1/1/1
3	SIA	A	603	-	-	7/20/38/38	0/1/1/1
2	NAG	B	601	1	-	1/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	603	SIA	O2-C2	3.53	1.44	1.39

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	603	SIA	O1A-C1-C2	-3.87	117.73	123.59
2	C	602	NAG	C1-O5-C5	2.77	115.94	112.19
2	A	602	NAG	C1-O5-C5	2.52	115.60	112.19

There are no chirality outliers.

All (18) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	603	SIA	C11-C10-N5-C5
3	A	603	SIA	O10-C10-N5-C5
3	A	603	SIA	O8-C8-C9-O9
2	B	602	NAG	O5-C5-C6-O6
2	B	602	NAG	C4-C5-C6-O6
2	C	601	NAG	O5-C5-C6-O6
2	A	602	NAG	C8-C7-N2-C2
2	A	602	NAG	O7-C7-N2-C2
2	B	602	NAG	C8-C7-N2-C2
2	B	602	NAG	O7-C7-N2-C2
2	C	602	NAG	C8-C7-N2-C2
2	C	602	NAG	O7-C7-N2-C2
2	C	601	NAG	C4-C5-C6-O6
3	A	603	SIA	C7-C8-C9-O9
3	A	603	SIA	O1A-C1-C2-O6
2	B	601	NAG	O5-C5-C6-O6
3	A	603	SIA	O1A-C1-C2-O2
3	A	603	SIA	O1B-C1-C2-O6

There are no ring outliers.

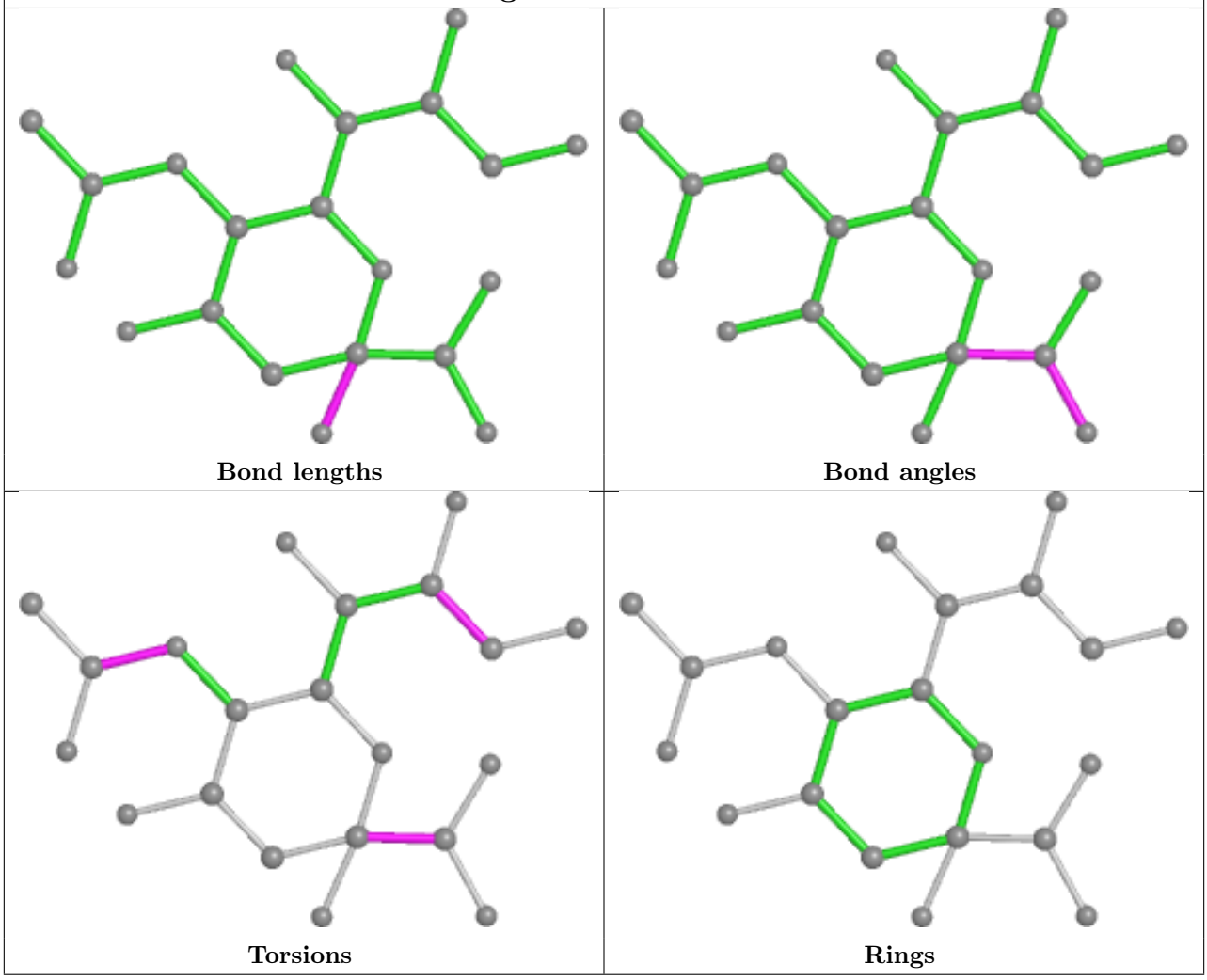
3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	602	NAG	1	0
2	B	602	NAG	2	0
3	A	603	SIA	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 then 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 SIA A 603



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 (i)

6.1 Protein, DNA and RNA chains (i)

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	486/503 (96%)	0.13	29 (5%)	21	20	23, 42, 94, 158	0
1	B	486/503 (96%)	0.22	37 (7%)	13	12	21, 41, 107, 156	0
1	C	486/503 (96%)	0.16	30 (6%)	20	19	21, 44, 97, 160	0
All	All	1458/1509 (96%)	0.17	96 (6%)	18	17	21, 43, 99, 160	0

All (96) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	383	ILE	11.0
1	B	498	PHE	6.7
1	A	377	GLY	6.2
1	B	356	ALA	5.8
1	B	470	LYS	5.6
1	C	383	ILE	4.9
1	A	383	ILE	4.8
1	C	386	THR	4.7
1	A	18	ASN	4.5
1	B	380	ASN	4.4
1	B	386	THR	4.1
1	A	17	ALA	4.0
1	C	380	ASN	3.9
1	C	377	GLY	3.9
1	A	379	LEU	3.9
1	B	488	ILE	3.8
1	B	358	GLY	3.8
1	B	17	ALA	3.8
1	B	384	GLU	3.8
1	A	380	ASN	3.7
1	B	322	PRO	3.7
1	B	361	THR	3.7
1	B	487	ASP	3.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	356	ALA	3.5
1	C	337	ILE	3.4
1	A	354	GLN	3.4
1	B	382	LEU	3.4
1	A	498	PHE	3.3
1	C	359	THR	3.3
1	B	471	CYS	3.2
1	A	467	ILE	3.2
1	C	379	LEU	3.1
1	B	16	VAL	3.1
1	A	356	ALA	3.1
1	A	5	GLY	3.1
1	B	379	LEU	3.1
1	A	27	ASP	3.0
1	B	484	TYR	3.0
1	B	385	LYS	3.0
1	B	359	THR	3.0
1	B	485	ASP	3.0
1	A	382	LEU	2.9
1	A	376	ASN	2.9
1	C	385	LYS	2.8
1	C	343	GLY	2.8
1	B	451	ARG	2.8
1	B	473	ASN	2.8
1	C	493	ALA	2.8
1	B	321	ILE	2.8
1	A	322	PRO	2.7
1	C	470	LYS	2.7
1	B	344	LEU	2.7
1	B	381	ARG	2.6
1	B	489	TYR	2.6
1	C	387	ASN	2.6
1	C	494	ILE	2.6
1	C	498	PHE	2.5
1	A	491	ASP	2.5
1	A	494	ILE	2.5
1	C	360	GLY	2.5
1	B	357	GLU	2.5
1	A	359	THR	2.5
1	C	384	GLU	2.4
1	A	333	ILE	2.4
1	C	354	GLN	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	345	ILE	2.4
1	C	495	ASN	2.4
1	C	468	PHE	2.4
1	B	355	ASN	2.4
1	C	473	ASN	2.4
1	A	384	GLU	2.4
1	C	338	GLU	2.4
1	A	362	ALA	2.3
1	A	321	ILE	2.3
1	A	358	GLY	2.3
1	C	345	ILE	2.3
1	C	375	ILE	2.2
1	A	470	LYS	2.2
1	C	488	ILE	2.2
1	B	26	ALA	2.2
1	A	472	ASP	2.2
1	B	29	GLN	2.2
1	A	475	CYS	2.1
1	B	472	ASP	2.1
1	C	467	ILE	2.1
1	C	344	LEU	2.1
1	C	475	CYS	2.1
1	B	465	PHE	2.1
1	A	124	ASN	2.1
1	B	365	LEU	2.1
1	C	365	LEU	2.1
1	B	18	ASN	2.1
1	B	125	THR	2.1
1	A	381	ARG	2.0
1	A	339	ASN	2.0
1	C	376	ASN	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 monosaccharides in this entry.

6.4 Ligands (i)

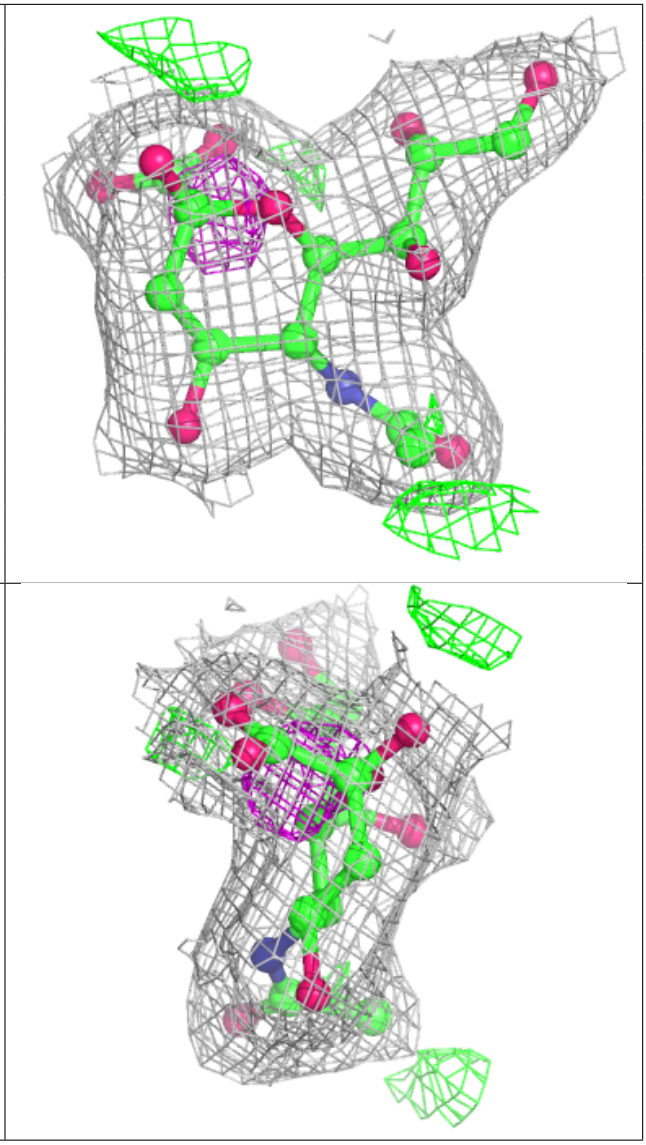
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	NAG	C	601	14/15	0.73	0.24	73,86,93,96	0
2	NAG	A	601	14/15	0.74	0.26	85,98,104,110	0
2	NAG	B	601	14/15	0.83	0.27	84,92,100,100	0
3	SIA	A	603	21/21	0.88	0.12	46,52,59,67	0
2	NAG	C	602	14/15	0.91	0.12	38,45,51,54	0
2	NAG	A	602	14/15	0.94	0.11	37,45,51,56	0
2	NAG	B	602	14/15	0.96	0.13	40,46,51,52	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 SIA A 603:

$2mF_o$-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.