Full wwPDB X-ray Structure Validation Report (i)

Mar 4, 2024 – 06:42 AM EST

PDB ID	:	3P6Z
Title	:	Structural basis of thrombin mediated factor V activation: essential role of the hirudin-like sequence Glu666-Glu672 for processing at the heavy chain-B domain junction
Authors	:	Corral-Rodriguez, M.A.; Bock, P.E.; Hernandez-Carvajal, E.; Gutierrez-Gallego, R.; Fuentes-Prior, P.
Deposited on	:	2010-10-11
Resolution	:	1.70 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:

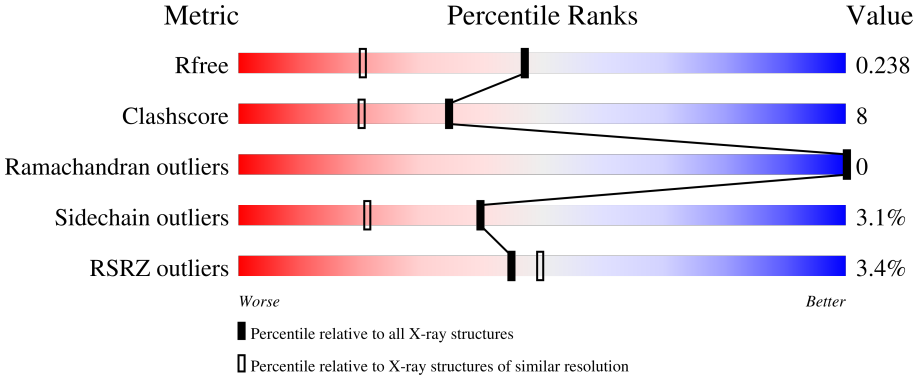
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 1.70 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	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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	36	3%, 69%, 8%, 6%, 17%
1	G	36	3%, 72%, 8%, 19%
2	B	259	3%, 85%, 12%, • •
2	H	259	3%, 85%, 12%, • •
3	C	71	6%, • •, 90%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	I	71	3% 7% 89%

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
5	0G6	B	707	X	-	-	-
5	0G6	H	707	X	-	-	-

2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 5445 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 Thrombin light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	30	254	156	44	52	2	0	5	0
1	G	29	246	151	43	51	1	0	5	0

- Molecule 2 is a protein called Thrombin heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	255	2091	1331	372	372	16	0	10	0
2	H	255	2100	1332	375	378	15	0	12	0

- Molecule 3 is a protein called Coagulation factor V.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	7	60	40	7	13	0	0	0
3	I	8	66	43	8	15	0	0	0

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	639	ALA	-	expression tag	UNP P12259
C	640	HIS	-	expression tag	UNP P12259
C	641	HIS	-	expression tag	UNP P12259
C	642	HIS	-	expression tag	UNP P12259
C	643	HIS	-	expression tag	UNP P12259
C	644	HIS	-	expression tag	UNP P12259
C	645	HIS	-	expression tag	UNP P12259
C	646	VAL	-	expression tag	UNP P12259

Continued on next page...

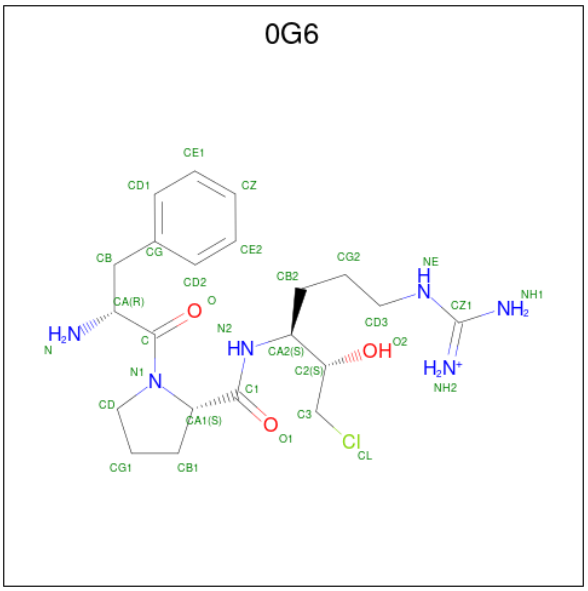
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	647	GLY	-	expression tag	UNP P12259
C	648	THR	-	expression tag	UNP P12259
C	649	TRP	-	expression tag	UNP P12259
C	650	GLU	-	expression tag	UNP P12259
C	651	ASN	-	expression tag	UNP P12259
C	652	LEU	-	expression tag	UNP P12259
C	653	TYR	-	expression tag	UNP P12259
C	654	PHE	-	expression tag	UNP P12259
C	655	GLN	-	expression tag	UNP P12259
C	656	SER	-	expression tag	UNP P12259
I	639	ALA	-	expression tag	UNP P12259
I	640	HIS	-	expression tag	UNP P12259
I	641	HIS	-	expression tag	UNP P12259
I	642	HIS	-	expression tag	UNP P12259
I	643	HIS	-	expression tag	UNP P12259
I	644	HIS	-	expression tag	UNP P12259
I	645	HIS	-	expression tag	UNP P12259
I	646	VAL	-	expression tag	UNP P12259
I	647	GLY	-	expression tag	UNP P12259
I	648	THR	-	expression tag	UNP P12259
I	649	TRP	-	expression tag	UNP P12259
I	650	GLU	-	expression tag	UNP P12259
I	651	ASN	-	expression tag	UNP P12259
I	652	LEU	-	expression tag	UNP P12259
I	653	TYR	-	expression tag	UNP P12259
I	654	PHE	-	expression tag	UNP P12259
I	655	GLN	-	expression tag	UNP P12259
I	656	SER	-	expression tag	UNP P12259

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

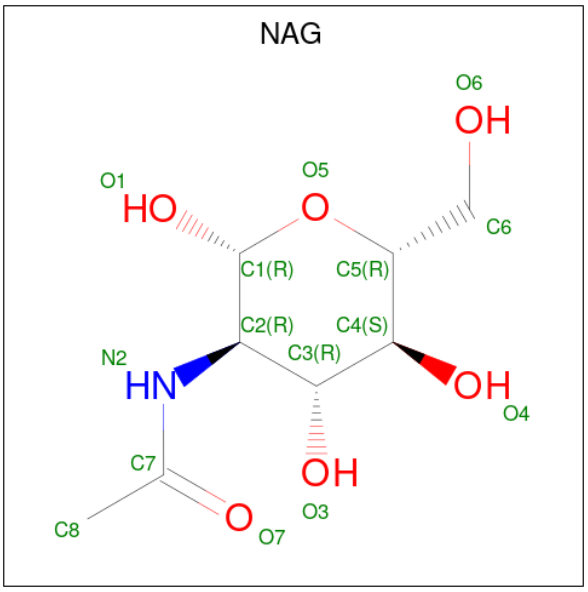
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Na		
4	A	1	1	1	0	0
			Total	Na		
4	B	2	2	2	0	0
			Total	Na		
4	G	1	1	1	0	0
			Total	Na		
4	H	5	5	5	0	0

- Molecule 5 is D-phenylalanyl-N-[(2S,3S)-6-{[amino(iminio)methyl]amino}-1-chloro-2-hydroxyhexan-3-yl]-L-prolinamide (three-letter code: 0G6) (formula: $C_{21}H_{34}ClN_6O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
5	B	1	30	21	6	3	0	0
			Total	C	N	O		
5	H	1	30	21	6	3	0	0

- Molecule 6 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



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

Continued on next page...

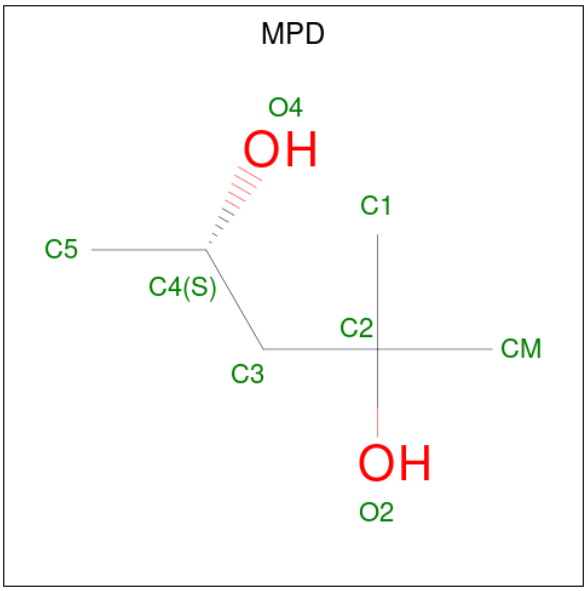
Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
6	H	1	14	8	1	5	0	0

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cl		
7	B	3	3	3	0	0

- Molecule 8 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: $C_6H_{14}O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
8	B	1	8	6	2	0	0

- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
9	A	31	33	33	0	2
9	B	222	232	232	0	10
9	C	1	1	1	0	0

Continued on next page...

Continued from previous page...

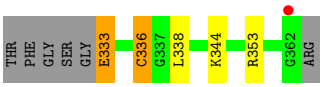
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	G	36	Total O 38 38	0	4
9	H	213	Total O 214 214	0	1
9	I	2	Total O 2 2	0	0

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.

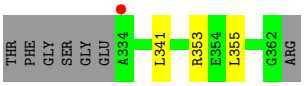
• Molecule 1: Thrombin light chain

Chain A:



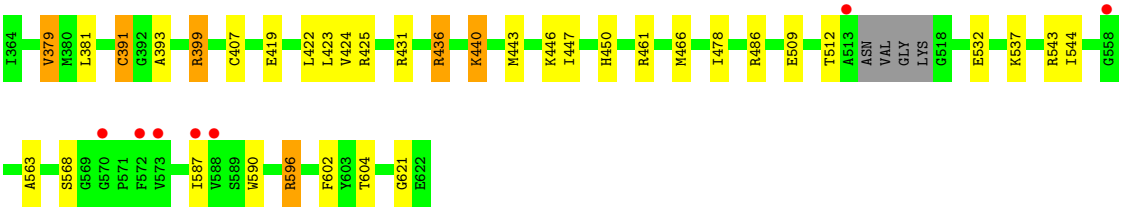
• Molecule 1: Thrombin light chain

Chain G:



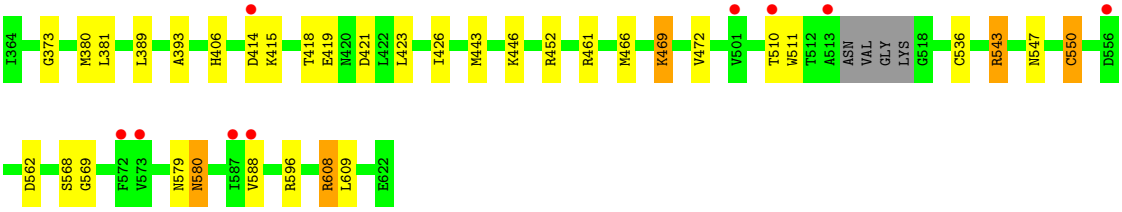
• Molecule 2: Thrombin heavy chain

Chain B:



• Molecule 2: Thrombin heavy chain

Chain H:



• Molecule 3: Coagulation factor V

Chain C: 6% 90%

ALA HIS HIS HIS HIS HIS HIS VAL GLY THR TRP GLU ASN LEU TYR PHE GLN SER ILE PRO ASP ASP ASP GLU ASP SER TYR E666 P671 E672 SER THR VAL MET ALA THR ARG LYS MET HIS ASP ARG LEU GLU PRO GLU ASP GLU GLU SER ASP ALA ASP TYR ASP TYR GLN ASN ARG

LEU ALA ALA ALA LEU GLY ILE ARG

• Molecule 3: Coagulation factor V

Chain I: 3% 7% 89%

ALA HIS HIS HIS HIS HIS HIS VAL GLY THR TRP GLU ASN LEU TYR PHE GLN SER ILE PRO ASP ASP ASP GLU ASP SER TYR E666 E672 S673 THR VAL MET ALA THR ARG LYS MET HIS ASP ARG LEU GLU PRO GLU ASP GLU GLU SER ASP ALA ASP TYR ASP TYR GLN ASN ARG LEU

ALA ALA ALA LEU GLY ILE ARG

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	52.21Å 62.26Å 67.51Å 99.38° 110.46° 92.26°	Depositor
Resolution (Å)	31.92 – 1.70 31.92 – 1.70	Depositor EDS
% Data completeness (in resolution range)	94.5 (31.92-1.70) 94.5 (31.92-1.70)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.17 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.5.0072	Depositor
R, R_{free}	0.200 , 0.239 0.200 , 0.238	Depositor DCC
R_{free} test set	2047 reflections (2.50%)	wwPDB-VP
Wilson B-factor (Å^2)	22.3	Xtriage
Anisotropy	0.143	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 50.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.118 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	5445	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.28% 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: 0G6, MPD, NAG, CL, NA

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	1.29	2/288 (0.7%)	0.99	0/381
1	G	1.14	0/282	1.00	1/373 (0.3%)
2	B	1.08	3/2202 (0.1%)	1.00	6/2972 (0.2%)
2	H	1.12	3/2221 (0.1%)	1.03	3/2998 (0.1%)
3	C	0.90	0/62	0.99	0/84
3	I	0.79	0/68	0.68	0/92
All	All	1.11	8/5123 (0.2%)	1.01	10/6900 (0.1%)

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
3	C	0	1

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	336[A]	CYS	CB-SG	8.27	1.96	1.82
1	A	336[B]	CYS	CB-SG	8.27	1.96	1.82
2	H	588	VAL	CB-CG2	6.11	1.65	1.52
2	B	563	ALA	CA-CB	5.88	1.64	1.52
2	H	536	CYS	CB-SG	5.48	1.91	1.82
2	B	590	TRP	CB-CG	5.45	1.60	1.50
2	B	379	VAL	CB-CG1	5.26	1.64	1.52
2	H	550	CYS	CB-SG	5.20	1.91	1.82

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	425	ARG	NE-CZ-NH1	-9.30	115.65	120.30
2	B	391[A]	CYS	CA-CB-SG	9.28	130.70	114.00
2	B	391[B]	CYS	CA-CB-SG	9.28	130.70	114.00
2	B	425	ARG	NE-CZ-NH2	7.07	123.83	120.30
2	H	543	ARG	NE-CZ-NH2	-6.77	116.92	120.30
2	H	562	ASP	CB-CG-OD1	6.74	124.36	118.30
1	G	355	LEU	CB-CG-CD1	-5.90	100.96	111.00
2	B	596	ARG	NE-CZ-NH2	-5.61	117.50	120.30
2	B	596	ARG	NE-CZ-NH1	5.45	123.02	120.30
2	H	596	ARG	NE-CZ-NH2	-5.28	117.66	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	671	PRO	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	254	0	242	3	0
1	G	246	0	229	3	0
2	B	2091	0	2060	43	0
2	H	2100	0	2043	31	0
3	C	60	0	51	1	0
3	I	66	0	56	0	0
4	A	1	0	0	0	0
4	B	2	0	0	0	0
4	G	1	0	0	0	0
4	H	5	0	0	0	0
5	B	30	0	30	6	0
5	H	30	0	30	10	0
6	B	14	0	13	0	0
6	H	14	0	13	0	0
7	B	3	0	0	0	0
8	B	8	0	14	1	0
9	A	33	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	B	232	0	0	2	0
9	C	1	0	0	0	0
9	G	38	0	0	2	0
9	H	214	0	0	2	0
9	I	2	0	0	0	0
All	All	5445	0	4781	78	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 8.

All (78) 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:B:436:ARG:HG3	2:B:436:ARG:HH11	1.29	0.94
2:B:399[B]:ARG:NH2	9:B:1214:HOH:O	1.99	0.94
2:B:381[B]:LEU:HD12	2:B:424:VAL:HG12	1.52	0.90
2:B:423:LEU:CD2	2:B:443[A]:MET:SD	2.62	0.88
2:B:436:ARG:HH11	2:B:436:ARG:CG	1.87	0.86
2:B:381[B]:LEU:HD23	2:B:391[B]:CYS:SG	2.17	0.84
2:H:568:SER:CB	5:H:707:0G6:C2	2.58	0.81
2:H:547:ASN:OD1	2:H:608[A]:ARG:CD	2.31	0.78
2:H:406:HIS:CD2	5:H:707:0G6:C3	2.66	0.78
2:H:568:SER:OG	5:H:707:0G6:CA2	2.32	0.77
2:B:423:LEU:HD21	2:B:443[A]:MET:SD	2.27	0.74
2:H:547:ASN:OD1	2:H:608[A]:ARG:HD2	1.89	0.73
2:B:381[B]:LEU:HD11	2:B:466:MET:CE	2.19	0.72
2:B:509:GLU:O	2:B:512:THR:HB	1.92	0.70
2:B:621:GLY:HA3	2:H:452:ARG:HH22	1.57	0.69
2:B:568:SER:OG	5:B:707:0G6:CA2	2.42	0.68
2:B:568:SER:CB	5:B:707:0G6:C2	2.73	0.67
2:B:381[B]:LEU:CD1	2:B:424:VAL:HG12	2.22	0.67
2:B:568:SER:OG	5:B:707:0G6:O2	2.14	0.64
2:B:381[B]:LEU:HD11	2:B:466:MET:HE1	1.82	0.61
1:A:333:GLU:O	1:A:336[B]:CYS:SG	2.54	0.61
2:H:579:ASN:O	2:H:580[B]:ASN:HB2	2.01	0.60
2:H:426:ILE:HD12	2:H:472[B]:VAL:HG11	1.85	0.59
2:H:547:ASN:OD1	2:H:608[A]:ARG:HD3	2.01	0.58
2:H:418[B]:THR:HG23	2:H:421:ASP:OD2	2.05	0.57
2:H:568:SER:CB	5:H:707:0G6:O2	2.52	0.57
2:B:419:GLU:HG2	2:B:446:LYS:HA	1.88	0.56
2:B:440:LYS:HD3	2:B:478:ILE:CD1	2.35	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:568:SER:OG	5:H:707:0G6:CB2	2.55	0.54
1:A:338:LEU:O	1:A:344[A]:LYS:NZ	2.40	0.54
2:B:568:SER:OG	5:B:707:0G6:C3	2.56	0.53
2:H:415[A]:LYS:NZ	9:H:983:HOH:O	2.41	0.52
2:H:579:ASN:O	2:H:580[B]:ASN:CB	2.55	0.52
2:B:381[B]:LEU:HD11	2:B:466:MET:HE3	1.89	0.51
2:H:568:SER:HB2	5:H:707:0G6:O2	2.09	0.51
2:B:587:ILE:HB	2:B:604:THR:HB	1.93	0.50
1:G:353[A]:ARG:HD2	9:G:1284:HOH:O	2.11	0.50
2:B:391[B]:CYS:SG	2:B:407:CYS:HB3	2.52	0.50
1:G:341:LEU:HD12	2:H:373:GLY:HA3	1.93	0.50
2:B:568:SER:CB	5:B:707:0G6:O2	2.60	0.50
2:B:436:ARG:HG3	2:B:436:ARG:NH1	2.06	0.50
2:H:568:SER:OG	5:H:707:0G6:O2	2.22	0.50
2:B:423:LEU:HD22	2:B:443[A]:MET:SD	2.49	0.48
2:B:440:LYS:HD3	2:B:478:ILE:HD11	1.95	0.48
2:B:436:ARG:CG	2:B:436:ARG:NH1	2.58	0.48
1:G:353[B]:ARG:HD2	9:G:1284:HOH:O	2.13	0.48
2:B:440:LYS:HE2	2:B:478:ILE:HD12	1.96	0.47
2:B:422:LEU:HD12	2:B:447:ILE:HD11	1.96	0.47
2:B:537:LYS:HA	2:B:544:ILE:HD12	1.96	0.47
2:H:568:SER:HB2	5:H:707:0G6:C2	2.42	0.46
2:H:423:LEU:CD2	2:H:443[B]:MET:SD	3.04	0.46
2:B:486:ARG:HG2	9:B:1403:HOH:O	2.15	0.46
2:B:544:ILE:HD11	2:B:602:PHE:CZ	2.50	0.46
2:B:391[B]:CYS:SG	2:B:407:CYS:SG	3.14	0.45
2:B:568:SER:OG	5:B:707:0G6:CB2	2.65	0.45
2:H:461:ARG:HG2	2:H:609:LEU:HD21	1.98	0.45
2:H:469:LYS:HE3	2:H:469:LYS:HB2	1.80	0.45
2:B:431:ARG:HH22	3:C:666:GLU:HG3	1.81	0.45
1:A:353[B]:ARG:HD3	9:A:1366:HOH:O	2.17	0.44
2:B:440:LYS:CE	2:B:478:ILE:HD12	2.48	0.44
2:B:450:HIS:CE1	2:B:461:ARG:HD3	2.52	0.44
2:B:621:GLY:HA3	2:H:452:ARG:NH2	2.30	0.43
2:B:544:ILE:CD1	2:B:602:PHE:CE2	3.02	0.43
2:H:510:THR:O	2:H:511:TRP:HB2	2.18	0.43
2:B:381[B]:LEU:HD23	2:B:391[B]:CYS:HG	1.84	0.42
2:H:543:ARG:NH2	9:H:1061:HOH:O	2.53	0.42
2:B:379:VAL:HB	2:B:393:ALA:HB3	2.01	0.42
2:H:406:HIS:CE1	5:H:707:0G6:C3	2.81	0.41
2:B:509:GLU:OE1	2:B:596:ARG:HD2	2.20	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:419:GLU:OE2	2:H:446:LYS:HD2	2.20	0.41
2:B:532:GLU:H	2:B:532:GLU:CD	2.24	0.41
2:B:544:ILE:HD13	2:B:602:PHE:HE2	1.84	0.41
2:H:568:SER:CB	5:H:707:0G6:C3	2.98	0.41
8:B:901:MPD:H4	8:B:901:MPD:H12	1.69	0.41
2:H:393:ALA:HA	2:H:569:GLY:O	2.21	0.41
2:H:426:ILE:HD12	2:H:472[A]:VAL:HG21	2.03	0.40
2:H:380:MET:HG3	2:H:389:LEU:HG	2.04	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	33/36 (92%)	32 (97%)	1 (3%)	0	100	100
1	G	32/36 (89%)	31 (97%)	1 (3%)	0	100	100
2	B	261/259 (101%)	252 (97%)	9 (3%)	0	100	100
2	H	263/259 (102%)	252 (96%)	11 (4%)	0	100	100
3	C	5/71 (7%)	5 (100%)	0	0	100	100
3	I	6/71 (8%)	6 (100%)	0	0	100	100
All	All	600/732 (82%)	578 (96%)	22 (4%)	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	32/31 (103%)	31 (97%)	1 (3%)	40	21
1	G	31/31 (100%)	31 (100%)	0	100	100
2	B	232/225 (103%)	227 (98%)	5 (2%)	52	34
2	H	234/225 (104%)	225 (96%)	9 (4%)	33	14
3	C	7/63 (11%)	5 (71%)	2 (29%)	0	0
3	I	8/63 (13%)	5 (62%)	3 (38%)	0	0
All	All	544/638 (85%)	524 (96%)	20 (4%)	40	15

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	333	GLU
2	B	399[A]	ARG
2	B	399[B]	ARG
2	B	436	ARG
2	B	440	LYS
2	B	543	ARG
3	C	666	GLU
3	C	672	GLU
2	H	381	LEU
2	H	414[A]	ASP
2	H	414[B]	ASP
2	H	469	LYS
2	H	550	CYS
2	H	580[A]	ASN
2	H	580[B]	ASN
2	H	608[A]	ARG
2	H	608[B]	ARG
3	I	666	GLU
3	I	672	GLU
3	I	673	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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)

Of 17 ligands modelled in this entry, 12 are monoatomic - leaving 5 for Mogul analysis.

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	$\#\|Z\| > 2$
6	NAG	B	1416	2	14,14,15	0.65	0	17,19,21	1.90	5 (29%)
8	MPD	B	901	-	7,7,7	0.44	0	9,10,10	0.97	0
5	0G6	B	707	2	30,31,32	2.15	5 (16%)	37,41,42	1.68	5 (13%)
6	NAG	H	1416	2	14,14,15	0.61	0	17,19,21	3.39	8 (47%)
5	0G6	H	707	2	30,31,32	1.69	3 (10%)	37,41,42	1.35	6 (16%)

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
6	NAG	B	1416	2	-	3/6/23/26	0/1/1/1
8	MPD	B	901	-	-	0/5/5/5	-
5	0G6	B	707	2	1/1/8/10	2/31/41/43	0/2/2/2
6	NAG	H	1416	2	-	5/6/23/26	0/1/1/1
5	0G6	H	707	2	1/1/8/10	2/31/41/43	0/2/2/2

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	707	0G6	C1-N2	7.44	1.50	1.34
5	B	707	0G6	O2-C2	-7.17	1.23	1.43
5	H	707	0G6	O2-C2	-6.72	1.24	1.43
5	H	707	0G6	C1-N2	2.97	1.40	1.34
5	B	707	0G6	CZ-CE1	2.27	1.44	1.38
5	B	707	0G6	CE1-CD1	2.27	1.43	1.38
5	H	707	0G6	CD3-NE	2.12	1.51	1.46
5	B	707	0G6	O-C	2.06	1.26	1.22

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	H	1416	NAG	C2-N2-C7	10.20	137.43	122.90
5	B	707	0G6	CB2-CA2-N2	-6.04	102.39	110.33
5	B	707	0G6	CB2-CA2-C2	4.76	120.78	112.51
6	H	1416	NAG	C8-C7-N2	4.63	123.94	116.10
6	H	1416	NAG	C1-O5-C5	4.12	117.78	112.19
6	B	1416	NAG	C4-C3-C2	-3.91	105.29	111.02
6	H	1416	NAG	C3-C4-C5	3.75	116.93	110.24
5	B	707	0G6	CA2-N2-C1	-3.65	116.61	123.07
5	H	707	0G6	CG2-CB2-CA2	-3.63	106.53	113.93
5	H	707	0G6	CB2-CA2-C2	3.60	118.77	112.51
6	B	1416	NAG	C1-O5-C5	3.40	116.80	112.19
6	B	1416	NAG	C2-N2-C7	3.34	127.66	122.90
6	H	1416	NAG	O5-C5-C4	3.23	118.69	110.83
6	H	1416	NAG	C1-C2-N2	3.10	115.78	110.49
5	H	707	0G6	O2-C2-C3	2.97	118.55	109.74
5	H	707	0G6	NH1-CZ1-NE	2.90	125.90	119.19
6	H	1416	NAG	O7-C7-C8	-2.61	117.21	122.06
5	B	707	0G6	O2-C2-C3	2.56	117.34	109.74
6	B	1416	NAG	C1-C2-N2	2.56	114.86	110.49
6	H	1416	NAG	C6-C5-C4	-2.35	107.49	113.00
5	H	707	0G6	NE-CZ1-NH2	-2.31	116.63	120.70
6	B	1416	NAG	O5-C1-C2	-2.15	107.89	111.29
5	H	707	0G6	CB-CA-C	2.02	114.12	109.27
5	B	707	0G6	NH1-CZ1-NE	2.02	123.85	119.19

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	B	707	0G6	C2
5	H	707	0G6	C2

All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	707	0G6	C3-C2-CA2-N2
5	H	707	0G6	C3-C2-CA2-N2
6	B	1416	NAG	O5-C5-C6-O6
6	H	1416	NAG	C8-C7-N2-C2
6	H	1416	NAG	O7-C7-N2-C2
6	H	1416	NAG	C4-C5-C6-O6
5	B	707	0G6	C3-C2-CA2-CB2
6	B	1416	NAG	C4-C5-C6-O6
5	H	707	0G6	C3-C2-CA2-CB2
6	H	1416	NAG	C3-C2-N2-C7
6	B	1416	NAG	C3-C2-N2-C7
6	H	1416	NAG	O5-C5-C6-O6

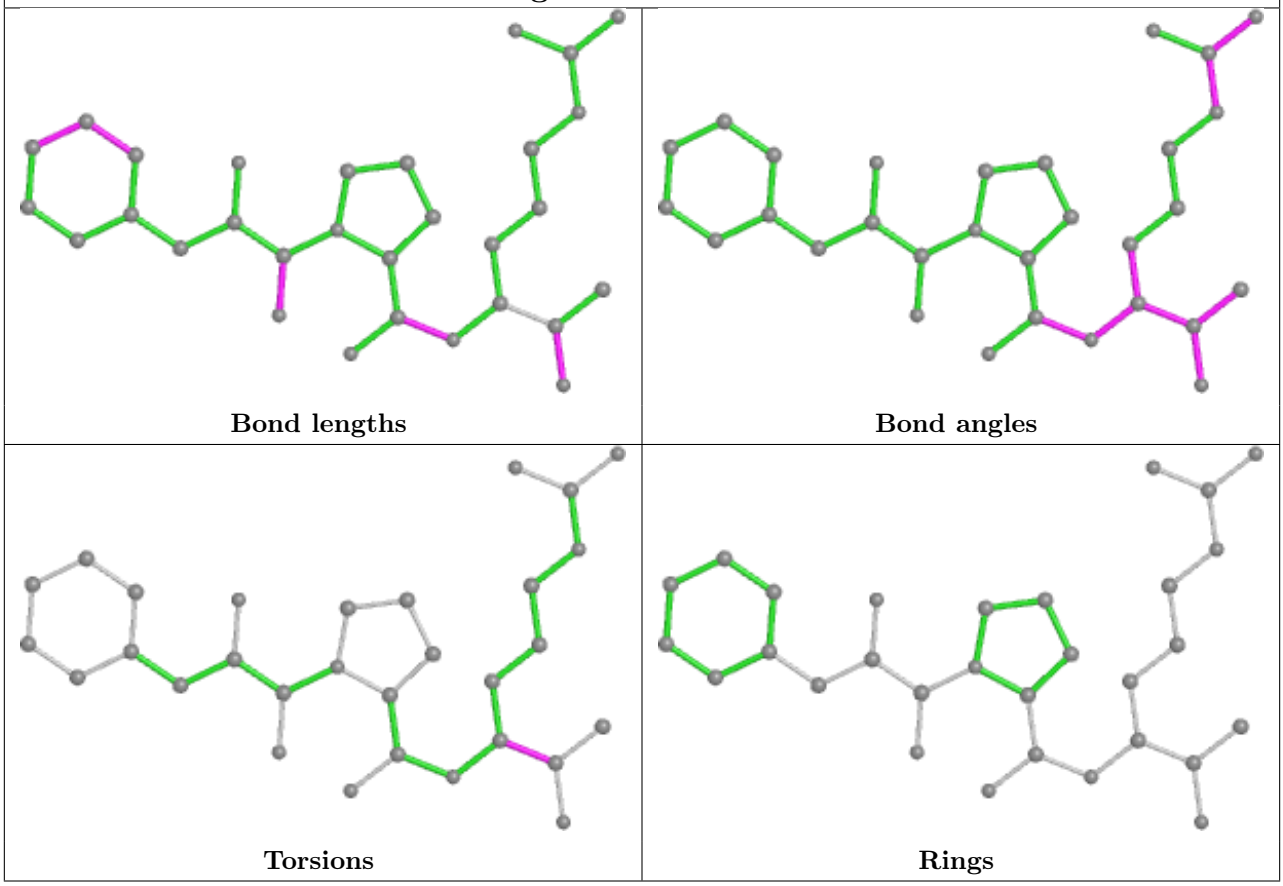
There are no ring outliers.

3 monomers are involved in 17 short contacts:

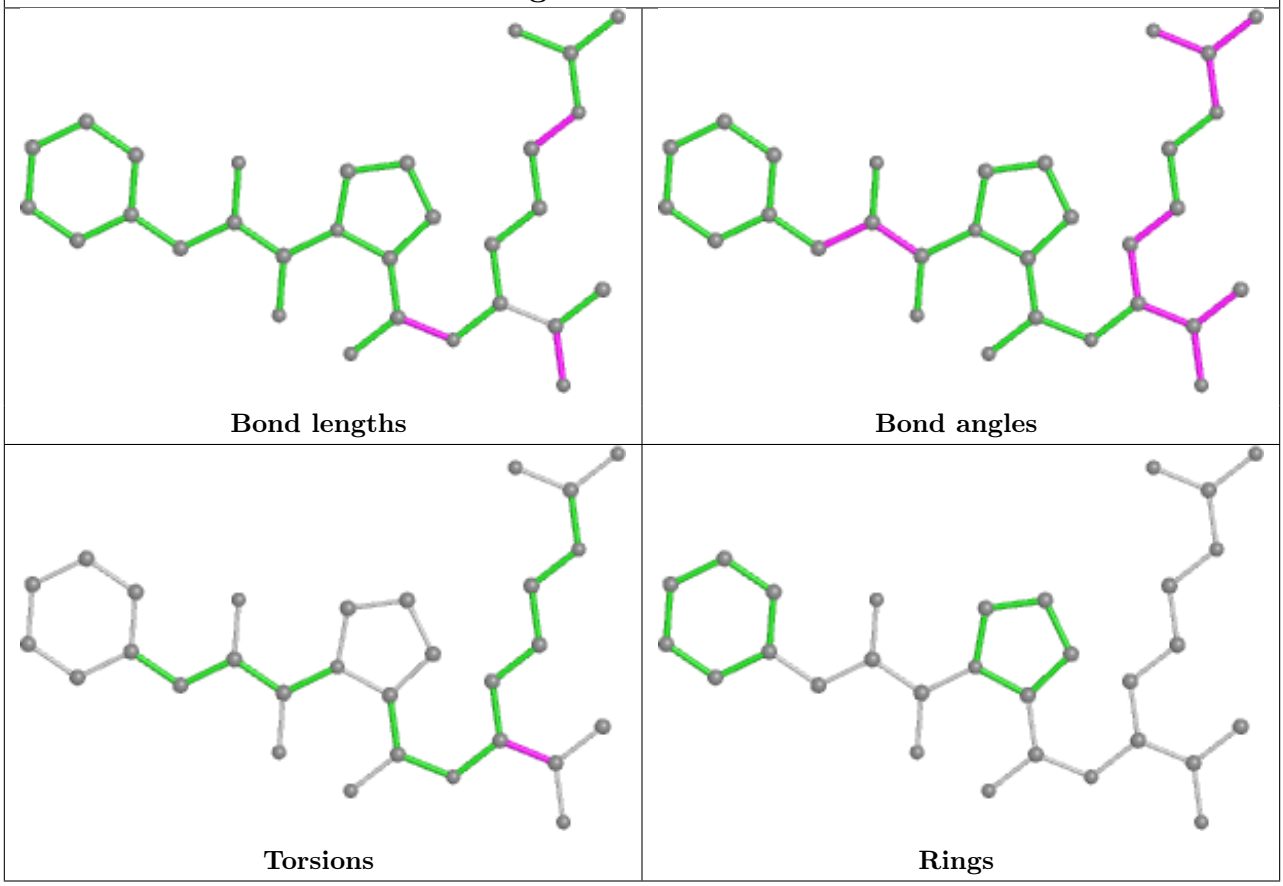
Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	B	901	MPD	1	0
5	B	707	0G6	6	0
5	H	707	0G6	10	0

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 0G6 B 707



Ligand 0G6 H 707



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	30/36 (83%)	-0.14	1 (3%)	46	51	19, 24, 43, 54	0
1	G	29/36 (80%)	-0.06	1 (3%)	45	50	17, 24, 37, 42	0
2	B	255/259 (98%)	-0.01	7 (2%)	54	58	16, 24, 40, 57	0
2	H	255/259 (98%)	0.05	9 (3%)	44	49	14, 24, 38, 51	0
3	C	7/71 (9%)	0.80	0	100	100	30, 34, 53, 56	0
3	I	8/71 (11%)	0.82	2 (25%)	0	0	32, 42, 55, 55	0
All	All	584/732 (79%)	0.03	20 (3%)	45	50	14, 25, 41, 57	0

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	513	ALA	4.8
1	G	334	ALA	4.1
2	B	513	ALA	3.9
2	H	556[A]	ASP	3.0
3	I	666	GLU	3.0
2	H	573	VAL	2.8
2	H	510	THR	2.8
3	I	673	SER	2.7
2	H	588	VAL	2.7
2	B	570	GLY	2.6
2	H	572	PHE	2.5
2	B	558	GLY	2.5
2	B	573	VAL	2.3
2	B	587	ILE	2.2
2	B	588	VAL	2.2
1	A	362	GLY	2.2
2	B	572	PHE	2.1
2	H	501	VAL	2.1
2	H	414[A]	ASP	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	H	587	ILE	2.1

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)

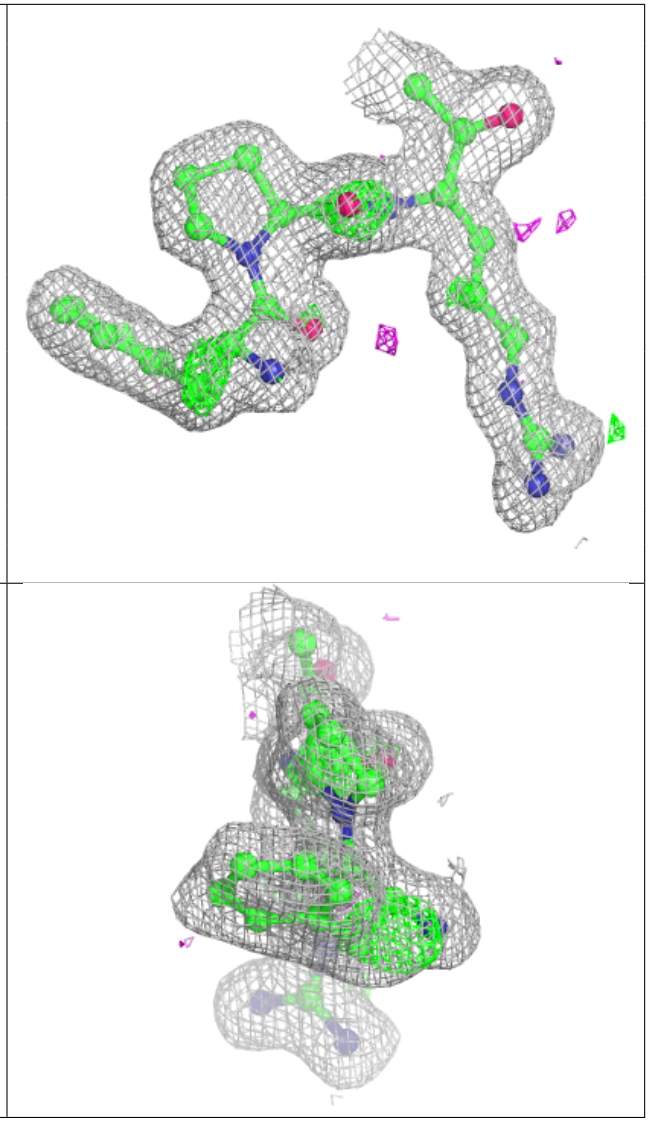
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	NA	H	808	1/1	0.65	0.18	60,60,60,60	0
6	NAG	B	1416	14/15	0.70	0.26	59,65,69,72	0
6	NAG	H	1416	14/15	0.80	0.28	54,58,61,62	0
7	CL	B	852	1/1	0.87	0.13	45,45,45,45	0
7	CL	B	851	1/1	0.90	0.06	63,63,63,63	0
4	NA	H	806	1/1	0.92	0.23	50,50,50,50	0
4	NA	A	809	1/1	0.92	0.12	57,57,57,57	0
4	NA	B	805	1/1	0.93	0.11	52,52,52,52	0
5	0G6	B	707	30/31	0.93	0.16	21,31,38,40	0
8	MPD	B	901	8/8	0.93	0.12	32,41,45,46	0
5	0G6	H	707	30/31	0.95	0.10	13,21,28,29	0
4	NA	H	803	1/1	0.96	0.12	42,42,42,42	0
4	NA	H	807	1/1	0.96	0.28	53,53,53,53	0
4	NA	B	802	1/1	0.97	0.05	24,24,24,24	0
7	CL	B	850	1/1	0.98	0.03	32,32,32,32	0
4	NA	H	801	1/1	0.98	0.06	22,22,22,22	0
4	NA	G	804	1/1	0.99	0.11	31,31,31,31	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.

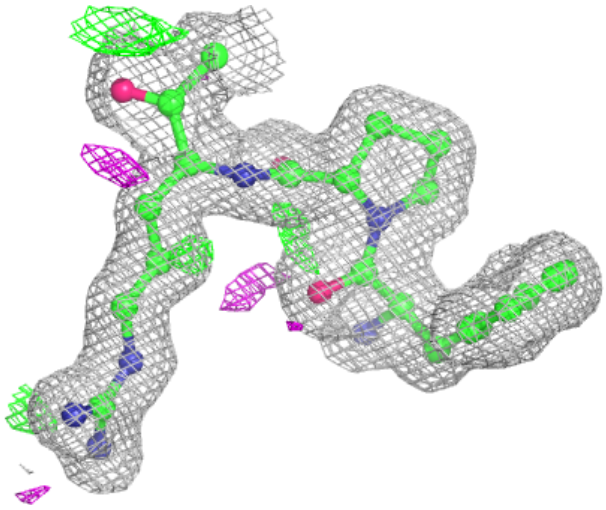
Electron density around 0G6 B 707:

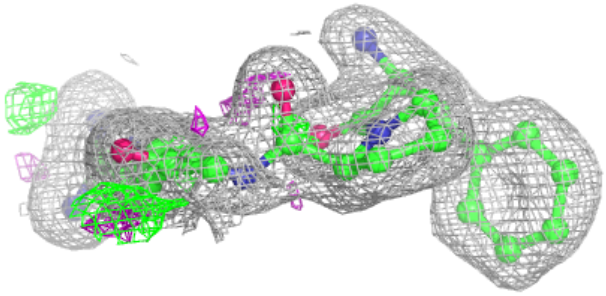
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

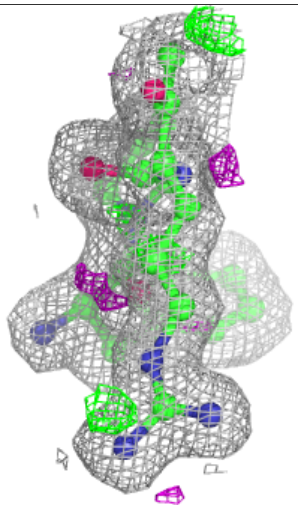


Electron density around 0G6 H 707:

$2mF_o$-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.