Full wwPDB X-ray Structure Validation Report ⓘ

Sep 4, 2023 – 09:38 PM EDT

PDB ID	:	3TH3
Title	:	Mg2+ Is Required for Optimal Folding of the Gamma-Carboxyglutamic Acid (Gla) Domains of Vitamin K-Dependent Clotting Factors At Physiological Ca2+
Authors	:	Vadivel, K.; Agah, S.; Cascio, D.; Padmanabhan, K.; Bajaj, S.P.
Deposited on	:	2011-08-18
Resolution	:	2.70 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 ⓘ 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:

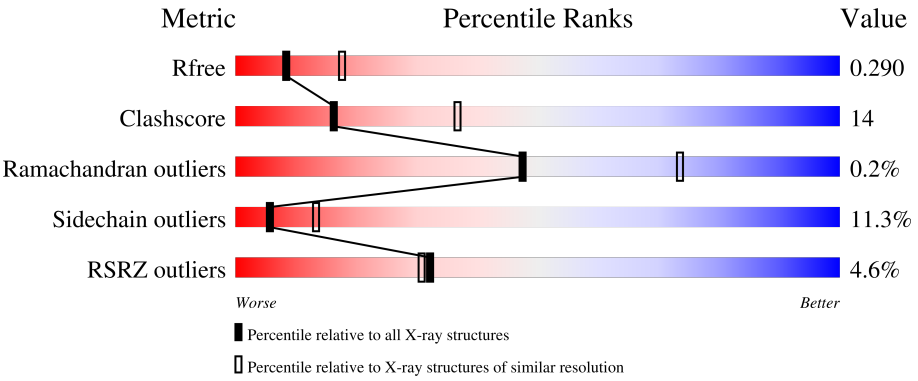
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	L	142	2% 51% 12% · 32%
2	H	254	% 68% 29% ·
3	T	205	9% 46% 24% · 25%

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 4082 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 Coagulation factor VII light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	L	96	718	431	125	149	13	0	0	0

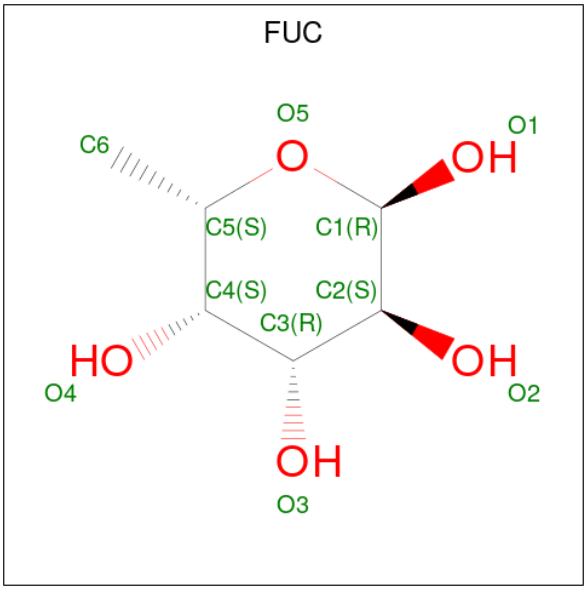
- Molecule 2 is a protein called Coagulation factor VII heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	H	254	1974	1253	351	357	13	0	0	0

- Molecule 3 is a protein called Tissue factor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	T	154	1242	787	205	246	4	0	0	0

- Molecule 4 is alpha-L-fucopyranose (three-letter code: FUC) (formula: $C_6H_{12}O_5$).

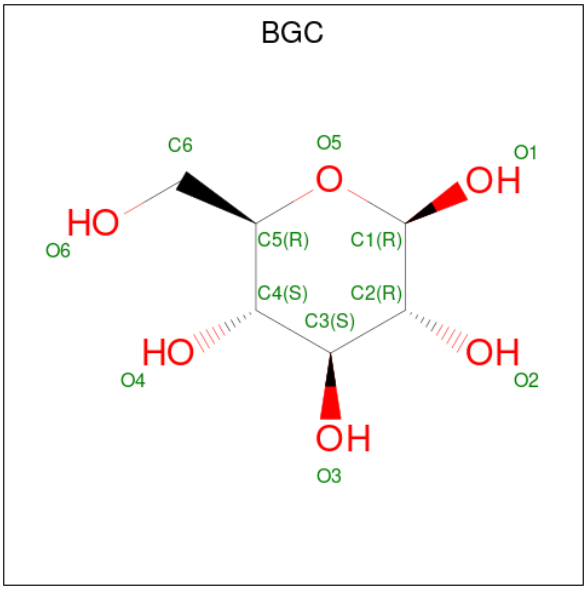


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	L	1	Total C O 10 6 4	0	0

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

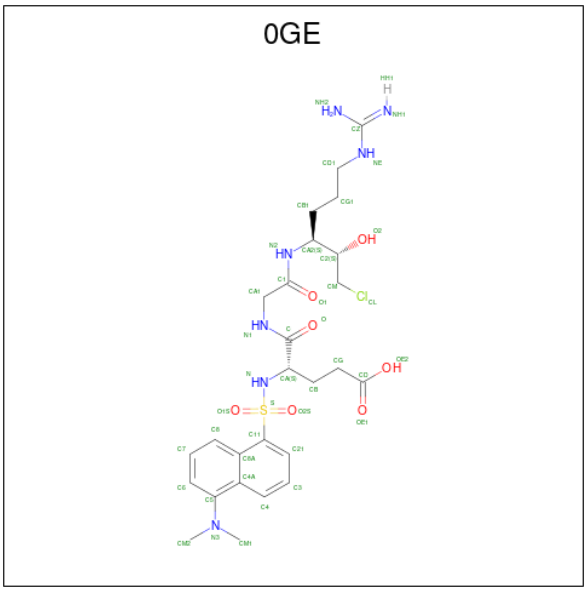
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	L	1	Total Ca 1 1	0	0
5	H	1	Total Ca 1 1	0	0

- Molecule 6 is beta-D-glucopyranose (three-letter code: BGC) (formula: $C_6H_{12}O_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	L	1	Total C O 11 6 5	0	0

- Molecule 7 is N-{[5-(dimethylamino)naphthalen-1-yl]sulfonyl}-L-alpha-glutamyl-N-[(2S,3S)-6-carbamimidamido-1-chloro-2-hydroxyhexan-3-yl]glycinamide (three-letter code: 0GE) (formula: $C_{26}H_{38}ClN_7O_7S$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	H	1	Total C N O S 41 26 7 7 1	0	0

- Molecule 8 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	H	1	Total Cl 1 1	0	0

- Molecule 9 is water.

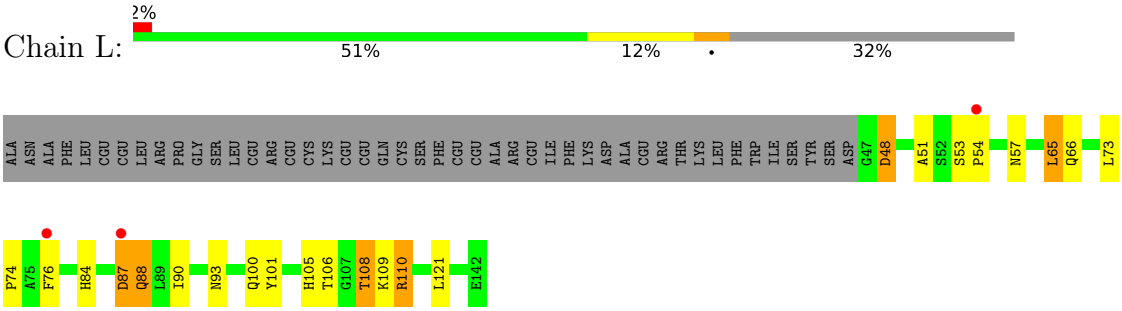
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	L	19	Total O 19 19	0	0
9	H	54	Total O 54 54	0	0
9	T	10	Total O 10 10	0	0

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.

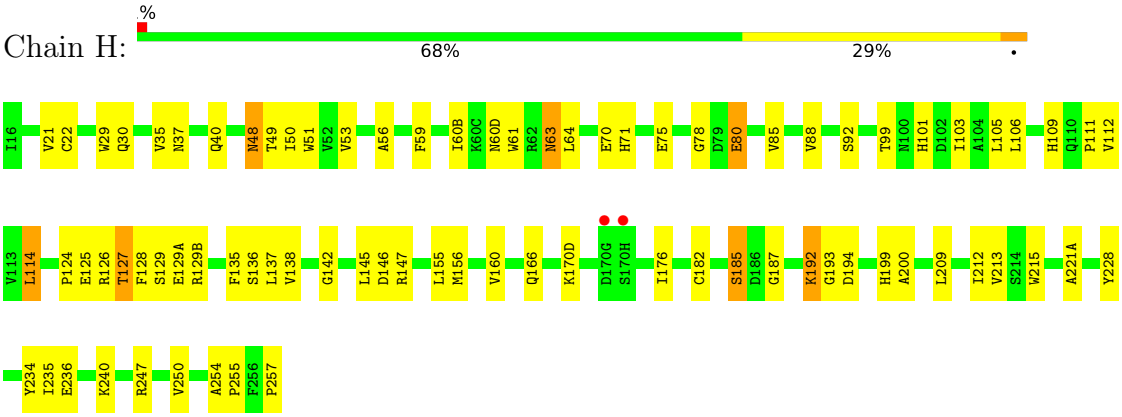
• Molecule 1: Coagulation factor VII light chain

Chain L:



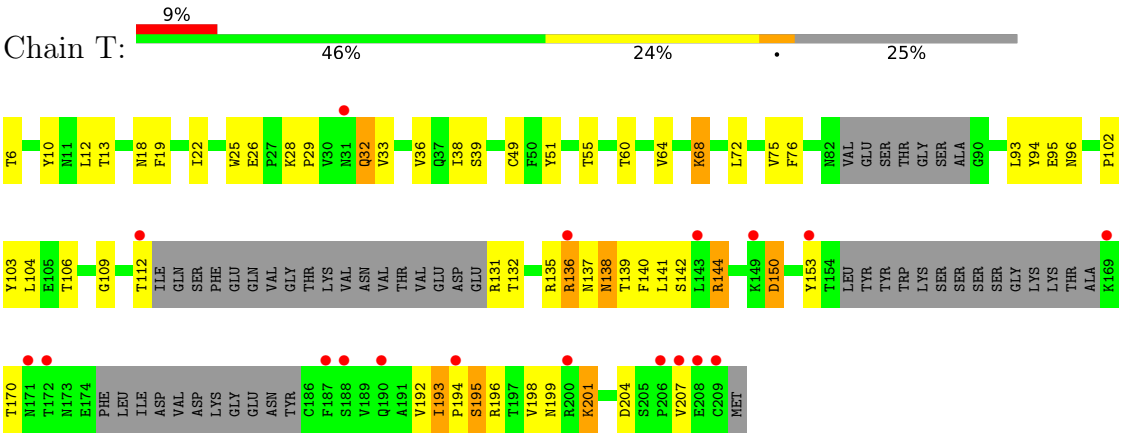
• Molecule 2: Coagulation factor VII heavy chain

Chain H:



• Molecule 3: Tissue factor

Chain T:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	78.08Å 68.89Å 79.15Å 90.00° 91.19° 90.00°	Depositor
Resolution (Å)	55.01 – 2.70 55.01 – 2.70	Depositor EDS
% Data completeness (in resolution range)	98.4 (55.01-2.70) 98.4 (55.01-2.70)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	0.10	Depositor
$< I/\sigma(I) >$ [1]	2.57 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.228 , 0.298 0.225 , 0.290	Depositor DCC
R_{free} test set	1184 reflections (5.17%)	wwPDB-VP
Wilson B-factor (Å^2)	44.3	Xtriage
Anisotropy	0.280	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 51.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.000 for l,k,-h 0.033 for h,-k,-l 0.016 for l,-k,h	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	4082	wwPDB-VP
Average B, all atoms (Å^2)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.30% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: FUC, BGC, CA, 0GE, CL

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	L	0.71	0/731	0.80	0/987
2	H	0.88	1/2024 (0.0%)	0.90	1/2755 (0.0%)
3	T	0.69	1/1267 (0.1%)	0.73	1/1722 (0.1%)
All	All	0.79	2/4022 (0.0%)	0.83	2/5464 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	236	GLU	CG-CD	5.78	1.60	1.51
3	T	49	CYS	CB-SG	5.39	1.91	1.82

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	T	93	LEU	CA-CB-CG	5.75	128.53	115.30
2	H	137	LEU	CB-CG-CD2	5.18	119.80	111.00

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	L	718	0	645	18	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	H	1974	0	1949	59	0
3	T	1242	0	1206	42	0
4	L	10	0	10	1	0
5	H	1	0	0	0	0
5	L	1	0	0	0	0
6	L	11	0	10	1	0
7	H	41	0	33	1	0
8	H	1	0	0	1	0
9	H	54	0	0	5	0
9	L	19	0	0	0	0
9	T	10	0	0	0	0
All	All	4082	0	3853	113	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 14.

All (113) 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:H:142:GLY:CA	2:H:193:GLY:HA3	1.83	1.08
2:H:142:GLY:HA2	2:H:193:GLY:HA3	1.26	1.07
1:L:73:LEU:HB3	1:L:74:PRO:HD2	1.54	0.87
3:T:12:LEU:H	3:T:12:LEU:HD12	1.52	0.74
2:H:142:GLY:HA3	2:H:193:GLY:HA3	1.69	0.74
2:H:30:GLN:HG2	2:H:155:LEU:HD13	1.71	0.72
1:L:101:TYR:HE1	2:H:128:PHE:CD2	2.08	0.71
2:H:60(D):ASN:HB3	2:H:63:ASN:OD1	1.90	0.71
2:H:127:THR:HG23	9:H:446:HOH:O	1.91	0.71
2:H:136:SER:HB2	2:H:199:HIS:CE1	2.27	0.69
2:H:30:GLN:HG2	2:H:155:LEU:CD1	2.22	0.69
3:T:144:ARG:HH21	3:T:153:TYR:HE2	1.38	0.69
3:T:192:VAL:HG22	3:T:201:LYS:HB3	1.77	0.67
2:H:50:ILE:HD11	2:H:111:PRO:HD3	1.76	0.66
3:T:12:LEU:HD12	3:T:12:LEU:N	2.11	0.65
3:T:138:ASN:O	3:T:138:ASN:ND2	2.30	0.65
1:L:87:ASP:N	1:L:87:ASP:OD1	2.30	0.64
3:T:18:ASN:OD1	3:T:109:GLY:N	2.31	0.64
3:T:193:ILE:HB	3:T:196:ARG:HD2	1.81	0.63
2:H:59:PHE:HA	2:H:60(B):ILE:HG12	1.81	0.63
1:L:57:ASN:OD1	1:L:76:PHE:CE1	2.52	0.62
2:H:92:SER:HB3	2:H:255:PRO:HA	1.81	0.62

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:T:18:ASN:O	3:T:19:PHE:HB2	2.01	0.61
1:L:73:LEU:HB3	1:L:74:PRO:CD	2.31	0.60
2:H:176:ILE:HD13	2:H:182:CYS:SG	2.41	0.60
2:H:101:HIS:CD2	2:H:234:TYR:OH	2.55	0.60
2:H:136:SER:HB3	2:H:200:ALA:O	2.02	0.59
2:H:78:GLY:HA2	9:H:405:HOH:O	2.03	0.59
2:H:124:PRO:O	2:H:235:ILE:HD12	2.04	0.58
3:T:12:LEU:H	3:T:12:LEU:CD1	2.18	0.57
3:T:25:TRP:O	3:T:55:THR:HB	2.05	0.57
3:T:60:THR:O	3:T:64:VAL:HB	2.04	0.57
3:T:132:THR:O	3:T:140:PHE:HB3	2.04	0.57
1:L:101:TYR:CE2	2:H:125:GLU:HG3	2.40	0.57
3:T:36:VAL:HG22	3:T:75:VAL:HG22	1.87	0.56
2:H:48:ASN:HB3	2:H:51:TRP:HB2	1.87	0.56
1:L:88:GLN:HB2	1:L:93:ASN:HD22	1.71	0.55
3:T:33:VAL:HG21	3:T:51:TYR:HB3	1.88	0.55
3:T:137:ASN:O	3:T:138:ASN:HB3	2.04	0.55
2:H:142:GLY:CA	2:H:193:GLY:CA	2.74	0.55
3:T:29:PRO:HB2	3:T:32:GLN:O	2.06	0.54
2:H:50:ILE:HD11	2:H:111:PRO:CD	2.37	0.54
2:H:99:THR:HB	2:H:215:TRP:CD1	2.43	0.54
3:T:131:ARG:HA	3:T:142:SER:HA	1.90	0.53
1:L:54:PRO:HA	6:L:203:BGC:O6	2.10	0.52
2:H:129(B):ARG:NH2	9:H:406:HOH:O	2.43	0.52
1:L:48:ASP:OD2	1:L:51:ALA:HB2	2.09	0.52
2:H:142:GLY:HA3	2:H:193:GLY:CA	2.37	0.52
3:T:196:ARG:NH1	3:T:199:ASN:O	2.43	0.52
1:L:101:TYR:CZ	2:H:125:GLU:HG3	2.45	0.51
3:T:76:PHE:CD1	3:T:94:TYR:HB3	2.45	0.51
2:H:49:THR:HG22	2:H:114:LEU:HD13	1.93	0.51
3:T:10:TYR:CZ	3:T:26:GLU:CB	2.94	0.51
3:T:95:GLU:OE2	3:T:96:ASN:N	2.30	0.51
3:T:10:TYR:CZ	3:T:26:GLU:HB2	2.46	0.50
3:T:137:ASN:O	3:T:138:ASN:CB	2.59	0.50
1:L:101:TYR:CE1	2:H:128:PHE:CD2	2.94	0.50
3:T:135:ARG:HG3	3:T:135:ARG:O	2.12	0.50
2:H:48:ASN:OD1	2:H:51:TRP:HD1	1.95	0.49
3:T:75:VAL:O	3:T:94:TYR:HA	2.11	0.49
3:T:193:ILE:HG22	3:T:194:PRO:HD2	1.94	0.49
3:T:28:LYS:HD2	3:T:55:THR:HG21	1.95	0.49
3:T:103:TYR:HD1	3:T:104:LEU:HD23	1.77	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:L:84:HIS:HB2	1:L:87:ASP:OD1	2.13	0.48
2:H:49:THR:HG22	2:H:114:LEU:CD1	2.43	0.48
1:L:105:HIS:HB2	1:L:108:THR:HG22	1.95	0.48
4:L:201:FUC:H61	3:T:140:PHE:HE2	1.77	0.48
3:T:193:ILE:CB	3:T:196:ARG:HD2	2.44	0.47
1:L:100:GLN:O	2:H:128:PHE:CE1	2.68	0.47
2:H:135:PHE:HA	2:H:160:VAL:O	2.15	0.47
3:T:136:ARG:HD3	3:T:141:LEU:HD21	1.97	0.46
2:H:80:GLU:H	2:H:80:GLU:HG2	1.45	0.46
2:H:146:ASP:O	2:H:147:ARG:HB2	2.15	0.46
2:H:21:VAL:HG22	2:H:156:MET:CE	2.46	0.46
2:H:88:VAL:HG22	2:H:106:LEU:HD22	1.97	0.46
2:H:138:VAL:HG11	2:H:228:TYR:CE2	2.51	0.46
2:H:257:PRO:HB3	9:H:441:HOH:O	2.15	0.45
1:L:53:SER:N	1:L:54:PRO:CD	2.80	0.45
3:T:10:TYR:CE1	3:T:26:GLU:HB2	2.51	0.45
2:H:53:VAL:HG11	2:H:212:ILE:HD11	1.98	0.45
2:H:92:SER:N	2:H:254:ALA:O	2.50	0.45
2:H:194:ASP:HB2	2:H:213:VAL:HG11	1.99	0.44
3:T:150:ASP:O	3:T:194:PRO:HD3	2.17	0.44
1:L:88:GLN:HB2	1:L:93:ASN:ND2	2.33	0.44
2:H:56:ALA:HB2	2:H:103:ILE:O	2.18	0.44
3:T:194:PRO:CD	3:T:195:SER:H	2.30	0.44
2:H:61:TRP:HB3	2:H:250:VAL:HG11	1.99	0.44
3:T:193:ILE:HG13	3:T:196:ARG:HD2	1.98	0.44
1:L:88:GLN:H	1:L:88:GLN:HG2	1.35	0.44
3:T:29:PRO:CB	3:T:32:GLN:O	2.66	0.44
2:H:64:LEU:HB3	2:H:85:VAL:HB	1.99	0.43
2:H:48:ASN:HD22	2:H:48:ASN:HA	1.64	0.43
2:H:136:SER:CB	2:H:199:HIS:CE1	2.99	0.43
2:H:78:GLY:C	9:H:405:HOH:O	2.57	0.43
2:H:92:SER:HB3	2:H:254:ALA:O	2.18	0.43
2:H:30:GLN:HG2	2:H:155:LEU:HD11	1.98	0.43
2:H:105:LEU:C	2:H:106:LEU:HD23	2.39	0.42
2:H:215:TRP:HA	7:H:301:0GE:HG21	2.00	0.42
2:H:192:LYS:HE2	2:H:192:LYS:HB3	1.72	0.42
2:H:70:GLU:HG3	2:H:71:HIS:N	2.35	0.42
2:H:126:ARG:O	2:H:129(A):GLU:HG3	2.20	0.42
2:H:166:GLN:HB2	3:T:94:TYR:OH	2.20	0.42
2:H:53:VAL:CG1	2:H:212:ILE:HD11	2.50	0.41
1:L:65:LEU:HA	1:L:65:LEU:HD23	1.46	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:187:GLY:HA2	2:H:221(A):ALA:O	2.20	0.41
3:T:102:PRO:O	3:T:106:THR:HG22	2.20	0.41
2:H:50:ILE:HD13	2:H:50:ILE:HA	1.98	0.41
3:T:10:TYR:CZ	3:T:26:GLU:HB3	2.55	0.41
2:H:185:SER:C	2:H:187:GLY:H	2.25	0.41
2:H:109:HIS:HB3	8:H:303:CL:CL	2.57	0.41
3:T:38:ILE:HG12	3:T:39:SER:N	2.36	0.40
3:T:68:LYS:H	3:T:68:LYS:HG2	1.69	0.40
3:T:194:PRO:CG	3:T:195:SER:N	2.85	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	L	94/142 (66%)	86 (92%)	7 (7%)	1 (1%)	14 34
2	H	252/254 (99%)	234 (93%)	18 (7%)	0	100 100
3	T	144/205 (70%)	134 (93%)	10 (7%)	0	100 100
All	All	490/601 (82%)	454 (93%)	35 (7%)	1 (0%)	47 73

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	L	110	ARG

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	L	83/114 (73%)	72 (87%)	11 (13%)	4 9
2	H	216/216 (100%)	196 (91%)	20 (9%)	9 21
3	T	144/189 (76%)	125 (87%)	19 (13%)	4 9
All	All	443/519 (85%)	393 (89%)	50 (11%)	6 13

All (50) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	L	48	ASP
1	L	65	LEU
1	L	66	GLN
1	L	87	ASP
1	L	88	GLN
1	L	90	ILE
1	L	106	THR
1	L	108	THR
1	L	109	LYS
1	L	110	ARG
1	L	121	LEU
2	H	22	CYS
2	H	29	TRP
2	H	35	VAL
2	H	37	ASN
2	H	40	GLN
2	H	48	ASN
2	H	63	ASN
2	H	75	GLU
2	H	80	GLU
2	H	112	VAL
2	H	114	LEU
2	H	127	THR
2	H	129	SER
2	H	145	LEU
2	H	170(D)	LYS
2	H	185	SER
2	H	192	LYS
2	H	209	LEU
2	H	240	LYS
2	H	247	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	T	6	THR
3	T	13	THR
3	T	22	ILE
3	T	32	GLN
3	T	68	LYS
3	T	72	LEU
3	T	112	THR
3	T	136	ARG
3	T	138	ASN
3	T	139	THR
3	T	144	ARG
3	T	150	ASP
3	T	170	THR
3	T	193	ILE
3	T	195	SER
3	T	198	VAL
3	T	201	LYS
3	T	204	ASP
3	T	207	VAL

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

Mol	Chain	Res	Type
1	L	57	ASN
1	L	84	HIS
2	H	48	ASN
2	H	101	HIS
2	H	175	ASN
2	H	224	HIS
3	T	138	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)

Of 6 ligands modelled in this entry, 3 are monoatomic - leaving 3 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	$\#\|Z\| > 2$
4	FUC	L	201	1	10,10,11	0.69	0	14,14,16	1.71	3 (21%)
7	0GE	H	301	-	41,42,43	2.87	3 (7%)	53,58,59	1.67	6 (11%)
6	BGC	L	203	1	11,11,12	0.62	0	15,15,17	1.08	1 (6%)

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
4	FUC	L	201	1	-	-	0/1/1/1
7	0GE	H	301	-	-	6/44/44/46	0/2/2/2
6	BGC	L	203	1	-	2/2/19/22	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	H	301	0GE	C11-S	-17.17	1.60	1.77
7	H	301	0GE	C11-C8A	3.43	1.48	1.43
7	H	301	0GE	C8A-C4A	2.77	1.48	1.43

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	H	301	0GE	O2S-S-O1S	-8.18	109.49	119.55

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	L	201	FUC	O5-C1-C2	-4.62	103.64	110.77
7	H	301	0GE	CB1-CA2-N2	-4.05	105.00	110.33
7	H	301	0GE	C21-C11-S	3.12	120.68	116.98
7	H	301	0GE	CG1-CD1-NE	-2.53	104.99	112.21
4	L	201	FUC	C3-C4-C5	2.42	113.55	109.77
6	L	203	BGC	O5-C5-C6	2.39	110.94	107.20
7	H	301	0GE	C21-C11-C8A	-2.38	118.61	121.00
4	L	201	FUC	C6-C5-C4	-2.27	108.88	113.07
7	H	301	0GE	CA1-N1-C	2.26	126.96	121.37

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	H	301	0GE	N-CA-CB-CG
7	H	301	0GE	O2-C2-CA2-N2
7	H	301	0GE	CM-C2-CA2-N2
7	H	301	0GE	O2-C2-CA2-CB1
7	H	301	0GE	CM-C2-CA2-CB1
6	L	203	BGC	O5-C5-C6-O6
6	L	203	BGC	C4-C5-C6-O6
7	H	301	0GE	CA-CB-CG-CD

There are no ring outliers.

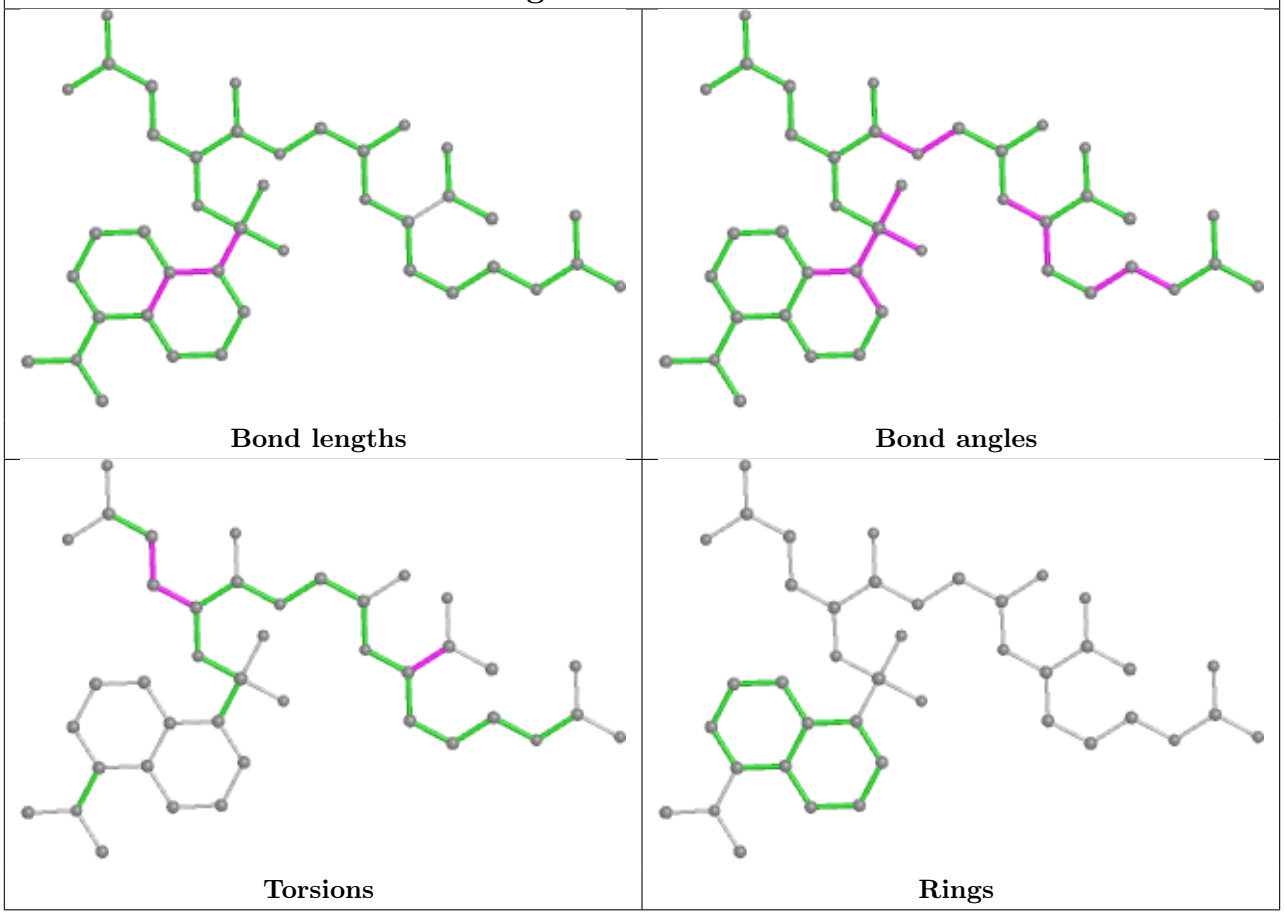
3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	L	201	FUC	1	0
7	H	301	0GE	1	0
6	L	203	BGC	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 0GE H 301



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, 95^{th} 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	L	96/142 (67%)	0.08	3 (3%) 49 49	27, 43, 88, 93	0
2	H	254/254 (100%)	-0.12	2 (0%) 86 87	15, 31, 65, 73	0
3	T	154/205 (75%)	0.50	18 (11%) 4 3	26, 53, 103, 109	0
All	All	504/601 (83%)	0.11	23 (4%) 32 31	15, 40, 96, 109	0

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	T	187	PHE	5.0
3	T	136	ARG	4.4
2	H	170(G)	ASP	4.3
3	T	208	GLU	4.0
3	T	112	THR	3.7
3	T	169	LYS	3.5
1	L	54	PRO	2.9
2	H	170(H)	SER	2.7
3	T	153	TYR	2.7
3	T	31	ASN	2.7
1	L	76	PHE	2.7
3	T	149	LYS	2.7
3	T	209	CYS	2.5
3	T	207	VAL	2.5
3	T	194	PRO	2.4
3	T	172	THR	2.4
3	T	171	ASN	2.4
3	T	200	ARG	2.3
3	T	143	LEU	2.2
1	L	87	ASP	2.2
3	T	190	GLN	2.1
3	T	206	PRO	2.1
3	T	188	SER	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)

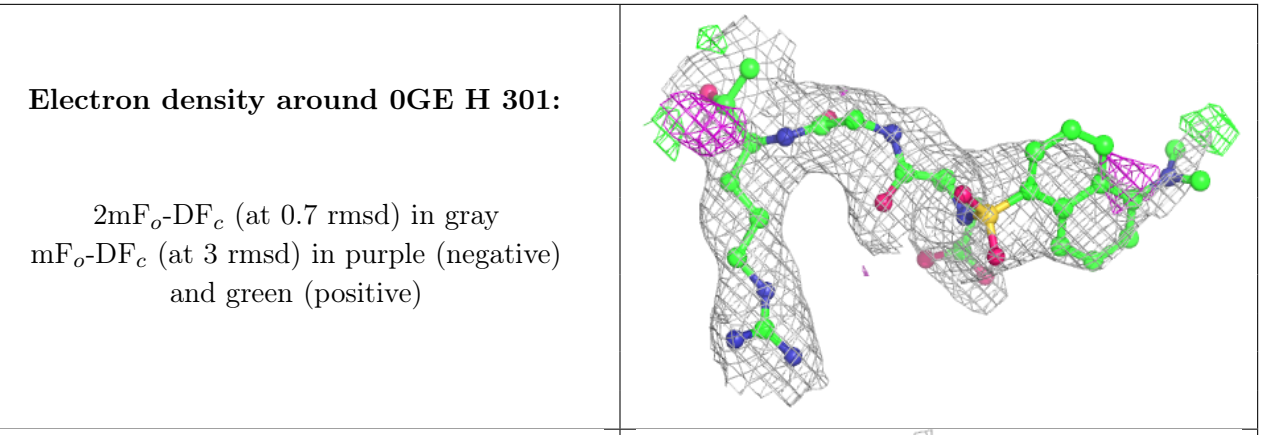
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

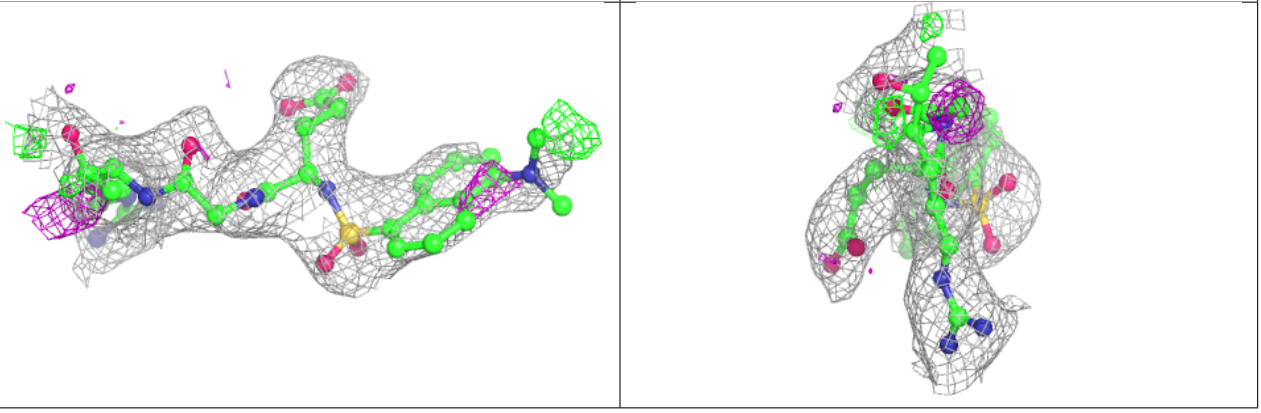
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	CA	L	202	1/1	0.77	0.26	98,98,98,98	0
7	0GE	H	301	41/42	0.91	0.20	24,52,72,74	0
6	BGC	L	203	11/12	0.92	0.15	76,78,80,81	0
5	CA	H	302	1/1	0.93	0.08	63,63,63,63	0
4	FUC	L	201	10/11	0.95	0.19	85,86,87,88	0
8	CL	H	303	1/1	0.97	0.12	39,39,39,39	0

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 0GE H 301:

$2mF_o$-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.