WORLDWIDE PDB
PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Dec 13, 2023 – 06:09 am GMT

PDB ID	:	2WZQ
Title	:	Insertion Mutant E173GP174 of the NS3 protease-helicase from dengue virus
Authors	:	Luo, D.; Wei, N.; Doan, D.; Paradkar, P.; Chong, Y.; Davidson, A.; Kotaka, M.; Lescar, J.; Vasudevan, S.
Deposited on	:	2009-12-02
Resolution	:	2.80 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
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 ⓘ

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

The reported resolution of this entry is 2.80 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	Percentile Ranks	Value
Rfree		0.301
Clashscore		9
Ramachandran outliers		1.4%
Sidechain outliers		11.7%

Worse   *Better*

▮ Percentile relative to all X-ray structures
⫾ Percentile relative to X-ray structures of similar resolution
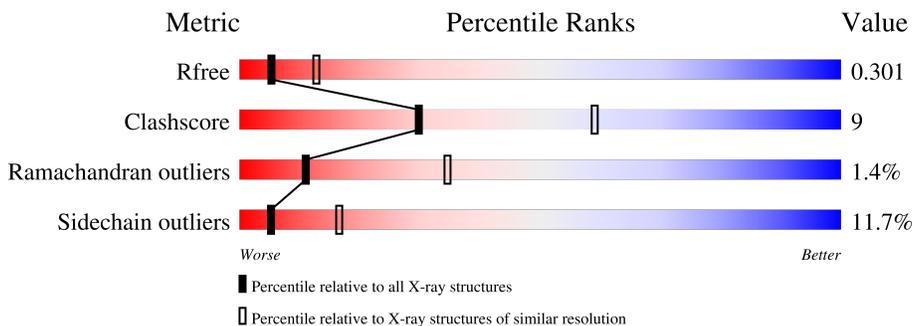
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)

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%

Mol	Chain	Length	Quality of chain
1	A	619	67% 22% • 6%
2	C	31	39% • 58%

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 4661 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 NS3 PROTEASE-HELICASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	581	Total 4537	C 2861	N 809	O 846	S 21	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	251	ASP	GLU	conflict	UNP Q2TN89

- Molecule 2 is a protein called NS3 PROTEASE-HELICASE.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	13	Total 92	C 56	N 16	O 20	0	0	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	58	ASN	SER	conflict	UNP Q2TN89
C	45	GLY	-	expression tag	UNP Q2TN89
C	46	SER	-	expression tag	UNP Q2TN89
C	47	ALA	-	expression tag	UNP Q2TN89
C	48	MET	-	expression tag	UNP Q2TN89
C	67	GLY	-	expression tag	UNP Q2TN89
C	68	GLY	-	expression tag	UNP Q2TN89
C	69	GLY	-	expression tag	UNP Q2TN89
C	70	GLY	-	expression tag	UNP Q2TN89
C	71	SER	-	expression tag	UNP Q2TN89
C	72	GLY	-	expression tag	UNP Q2TN89
C	73	GLY	-	expression tag	UNP Q2TN89
C	74	GLY	-	expression tag	UNP Q2TN89
C	75	GLY	-	expression tag	UNP Q2TN89

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total Cl 1 1		0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).
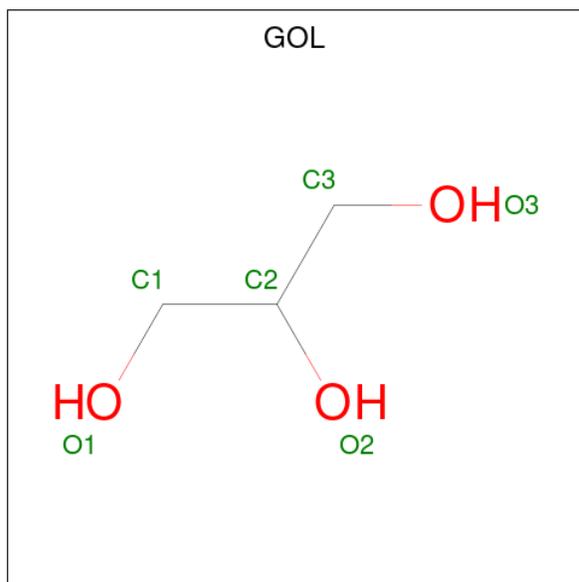
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total C O 6 3 3			0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	25	Total O 25 25		0	0

3 Residue-property plots ⓘ

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. 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. 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: NS3 PROTEASE-HELICASE

Chain A:
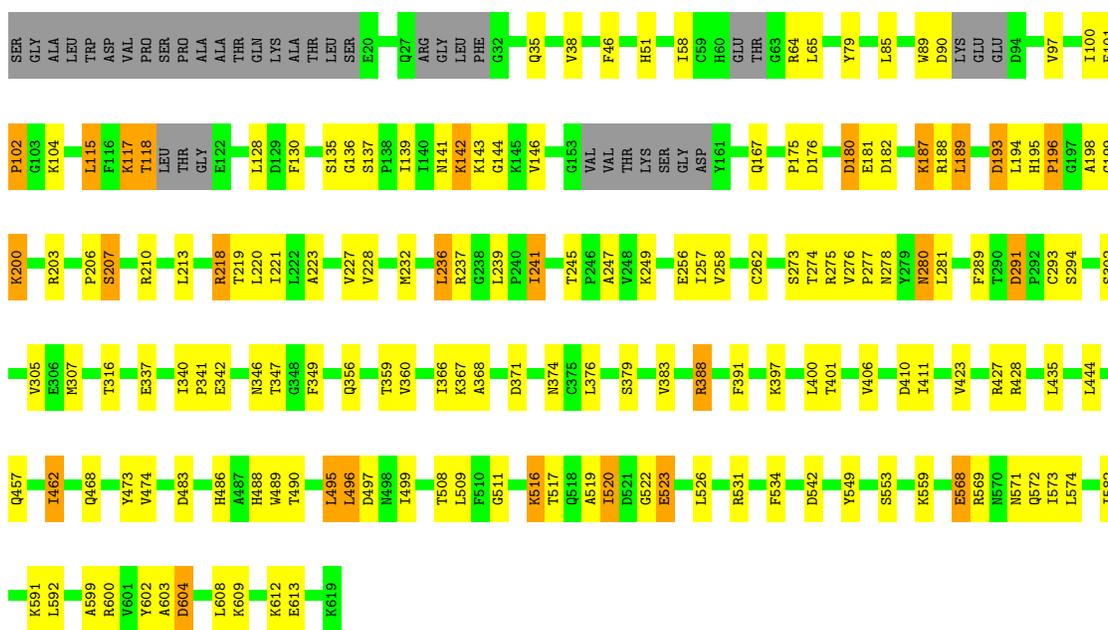
- Molecule 2: NS3 PROTEASE-HELICASE

Chain C:
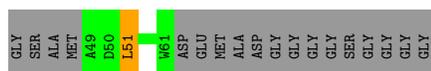
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	52.42Å 87.72Å 75.78Å 90.00° 92.90° 90.00°	Depositor
Resolution (Å)	30.00 – 2.80 29.59 – 2.76	Depositor EDS
% Data completeness (in resolution range)	100.0 (30.00-2.80) 98.6 (29.59-2.76)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.21 (at 2.76Å)	Xtriage
Refinement program	REFMAC 5.0	Depositor
R, R_{free}	0.220 , 0.288 0.268 , 0.301	Depositor DCC
R_{free} test set	881 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	60.0	Xtriage
Anisotropy	0.070	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 63.9	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.041 for h,-k,-l	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	4661	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.70% 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: GOL, 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	A	0.47	0/4634	0.65	1/6272 (0.0%)
2	C	0.46	0/91	0.60	0/122
All	All	0.47	0/4725	0.65	1/6394 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	175	PRO	N-CA-CB	5.62	110.04	103.30

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	4537	0	4468	86	2
2	C	92	0	89	1	0
3	A	1	0	0	0	0
4	A	6	0	8	1	0
5	A	25	0	0	2	0
All	All	4661	0	4565	86	2

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 9.

All (86) 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:A:181:GLU:HB3	1:A:182:ASP:CB	1.59	1.30
1:A:181:GLU:CB	1:A:182:ASP:HB3	1.69	1.23
1:A:423:VAL:HG23	1:A:462:ILE:HD13	1.54	0.90
1:A:520:ILE:O	1:A:523:GLU:HG2	1.84	0.78
1:A:486:HIS:HD2	1:A:488:HIS:H	1.35	0.74
1:A:599:ALA:HA	1:A:602:TYR:CE1	2.21	0.74
1:A:46:PHE:HB3	1:A:79:TYR:HB2	1.70	0.72
1:A:427:ARG:HD2	1:A:473:TYR:OH	1.94	0.67
1:A:573:ILE:HG13	1:A:592:LEU:HG	1.78	0.65
1:A:280:ASN:ND2	1:A:280:ASN:H	1.95	0.64
1:A:181:GLU:HB3	1:A:182:ASP:HB3	0.73	0.63
1:A:180:ASP:OD2	1:A:180:ASP:N	2.29	0.62
1:A:388:ARG:HG3	1:A:411:ILE:HD13	1.82	0.62
1:A:65:LEU:HD22	1:A:79:TYR:HB3	1.82	0.62
1:A:486:HIS:CD2	1:A:488:HIS:H	2.18	0.62
1:A:366:ILE:H	1:A:366:ILE:HD12	1.64	0.61
1:A:368:ALA:O	1:A:371:ASP:HB2	2.01	0.61
1:A:196:PRO:HA	1:A:200:LYS:HD3	1.82	0.60
1:A:280:ASN:H	1:A:280:ASN:HD22	1.48	0.59
1:A:51:HIS:HE1	1:A:135:SER:HB3	1.69	0.58
1:A:496:LEU:HA	1:A:499:ILE:HD12	1.85	0.57
1:A:38:VAL:O	1:A:136:GLY:HA3	2.03	0.57
1:A:526:LEU:O	1:A:531:ARG:HB2	2.04	0.57
1:A:245:THR:HG22	1:A:247:ALA:H	1.71	0.55
1:A:203:ARG:O	1:A:207:SER:HB2	2.06	0.55
1:A:360:VAL:HA	1:A:406:VAL:HG13	1.88	0.55
1:A:97:VAL:HG22	1:A:139:ILE:HG12	1.89	0.54
1:A:276:VAL:HG22	1:A:278:ASN:ND2	2.23	0.54
1:A:291:ASP:O	1:A:294:SER:HB2	2.09	0.53
1:A:218:ARG:NH1	1:A:280:ASN:OD1	2.42	0.53
1:A:206:PRO:HB2	1:A:210:ARG:HH12	1.74	0.53
1:A:388:ARG:NH2	1:A:410:ASP:OD1	2.42	0.52
1:A:236:LEU:HB3	1:A:241:ILE:HD13	1.91	0.52
1:A:293:CYS:HB2	5:A:2008:HOH:O	2.10	0.51
1:A:508:THR:HG22	1:A:509:LEU:N	2.26	0.51
1:A:89:TRP:CG	1:A:90:ASP:N	2.79	0.50
1:A:341:PRO:HD3	1:A:349:PHE:CE1	2.47	0.49
1:A:51:HIS:CE1	1:A:135:SER:HB3	2.47	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:218:ARG:N	1:A:280:ASN:HD21	2.09	0.49
1:A:346:ASN:O	1:A:347:THR:HG23	2.13	0.47
1:A:227:VAL:HG21	1:A:391:PHE:CE2	2.49	0.47
1:A:35:GLN:HB3	1:A:102:PRO:HB3	1.96	0.47
1:A:199:GLY:HA3	1:A:203:ARG:H	1.80	0.47
1:A:280:ASN:HD22	1:A:280:ASN:N	2.10	0.47
1:A:115:LEU:H	1:A:115:LEU:HD13	1.80	0.46
1:A:193:ASP:HA	1:A:316:THR:O	2.14	0.46
1:A:181:GLU:HB3	1:A:182:ASP:CA	2.36	0.46
1:A:141:ASN:O	1:A:143:LYS:N	2.47	0.46
1:A:397:LYS:O	1:A:401:THR:HG22	2.16	0.46
1:A:280:ASN:ND2	1:A:280:ASN:N	2.60	0.46
1:A:600:ARG:HD2	5:A:2022:HOH:O	2.16	0.46
1:A:228:VAL:HG12	1:A:232:MET:CE	2.47	0.45
1:A:187:LYS:HA	1:A:305:VAL:O	2.16	0.45
1:A:603:ALA:O	1:A:604:ASP:HB3	2.17	0.45
1:A:236:LEU:O	1:A:237:ARG:C	2.55	0.45
1:A:495:LEU:O	1:A:499:ILE:HD12	2.16	0.45
1:A:568:GLU:O	1:A:571:ASN:HB2	2.17	0.45
1:A:206:PRO:HB2	1:A:210:ARG:NH1	2.30	0.45
1:A:468:GLN:H	1:A:468:GLN:CD	2.21	0.45
1:A:142:LYS:HD2	1:A:142:LYS:HA	1.86	0.44
1:A:223:ALA:O	1:A:262:CYS:HA	2.17	0.44
1:A:572:GLN:OE1	1:A:591:LYS:HB3	2.17	0.44
1:A:51:HIS:CE1	1:A:135:SER:CB	3.01	0.44
1:A:117:LYS:O	1:A:118:THR:HG22	2.18	0.44
1:A:483:ASP:OD1	4:A:1621:GOL:O1	2.35	0.44
1:A:489:TRP:CE2	1:A:516:LYS:HG2	2.53	0.44
1:A:534:PHE:CD1	1:A:549:TYR:HA	2.53	0.44
1:A:198:ALA:HA	1:A:199:GLY:HA2	1.67	0.43
1:A:213:LEU:HD21	1:A:258:VAL:HG23	1.99	0.43
1:A:181:GLU:CB	1:A:182:ASP:CB	2.55	0.43
1:A:275:ARG:O	1:A:277:PRO:HD3	2.19	0.43
1:A:609:LYS:O	1:A:613:GLU:HG2	2.18	0.43
1:A:101:GLU:HB2	1:A:104:LYS:HB2	2.00	0.43
1:A:526:LEU:HD11	1:A:553:SER:HB3	2.01	0.43
1:A:128:LEU:HB3	1:A:130:PHE:CE1	2.54	0.42
1:A:340:ILE:HD11	1:A:474:VAL:HG12	2.00	0.42
1:A:519:ALA:HB1	1:A:523:GLU:OE1	2.19	0.42
1:A:219:THR:HG23	1:A:281:LEU:HD23	2.02	0.42
1:A:189:LEU:HD12	1:A:305:VAL:HG21	2.02	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:497:ASP:HB2	1:A:522:GLY:HA2	2.02	0.41
1:A:457:GLN:NE2	1:A:457:GLN:HA	2.36	0.41
1:A:38:VAL:HG21	1:A:100:ILE:HB	2.02	0.41
1:A:582:ILE:HD13	1:A:612:LYS:HB3	2.01	0.41
1:A:534:PHE:HD1	1:A:549:TYR:HA	1.86	0.41
1:A:58:ILE:HG23	2:C:51:LEU:HD12	2.02	0.41
1:A:359:THR:HG22	1:A:360:VAL:N	2.36	0.41

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:273:SER:OG	1:A:337:GLU:OE1[1_655]	2.18	0.02
1:A:249:LYS:N	1:A:337:GLU:O[1_655]	2.19	0.01

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	569/619 (92%)	509 (90%)	52 (9%)	8 (1%)	11	34
2	C	11/31 (36%)	10 (91%)	1 (9%)	0	100	100
All	All	580/650 (89%)	519 (90%)	53 (9%)	8 (1%)	11	34

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	196	PRO
1	A	200	LYS
1	A	142	LYS
1	A	176	ASP
1	A	144	GLY

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	194	LEU
1	A	218	ARG
1	A	511	GLY

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	478/524 (91%)	422 (88%)	56 (12%)	5	16
2	C	9/17 (53%)	8 (89%)	1 (11%)	6	19
All	All	487/541 (90%)	430 (88%)	57 (12%)	5	16

All (57) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	64	ARG
1	A	85	LEU
1	A	102	PRO
1	A	115	LEU
1	A	117	LYS
1	A	118	THR
1	A	137	SER
1	A	146	VAL
1	A	167	GLN
1	A	180	ASP
1	A	187	LYS
1	A	188	ARG
1	A	189	LEU
1	A	193	ASP
1	A	195	HIS
1	A	207	SER
1	A	220	LEU
1	A	221	ILE
1	A	236	LEU
1	A	239	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	241	ILE
1	A	256	GLU
1	A	257	ILE
1	A	274	THR
1	A	280	ASN
1	A	289	PHE
1	A	291	ASP
1	A	302	SER
1	A	307	MET
1	A	342	GLU
1	A	356	GLN
1	A	367	LYS
1	A	374	ASN
1	A	376	LEU
1	A	379	SER
1	A	383	VAL
1	A	388	ARG
1	A	400	LEU
1	A	428	ARG
1	A	435	LEU
1	A	444	LEU
1	A	462	ILE
1	A	490	THR
1	A	495	LEU
1	A	496	LEU
1	A	516	LYS
1	A	517	THR
1	A	520	ILE
1	A	523	GLU
1	A	542	ASP
1	A	559	LYS
1	A	568	GLU
1	A	569	ARG
1	A	574	LEU
1	A	604	ASP
1	A	608	LEU
2	C	51	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	51	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	167	GLN
1	A	278	ASN
1	A	280	ASN
1	A	288	HIS
1	A	330	ASN
1	A	356	GLN
1	A	385	GLN
1	A	457	GLN
1	A	486	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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	GOL	A	1621	-	5,5,5	0.36	0	5,5,5	0.36	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
4	GOL	A	1621	-	-	4/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1621	GOL	O1-C1-C2-O2
4	A	1621	GOL	O1-C1-C2-C3
4	A	1621	GOL	C1-C2-C3-O3
4	A	1621	GOL	O2-C2-C3-O3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1621	GOL	1	0

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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.