wwPDB X-ray Structure Validation Summary Report (i)

Mar 17, 2026 – 06:21 PM UTC

PDB ID	:	4B6F / pdb_00004b6f
Title	:	Discovery of an allosteric mechanism for the regulation of HCV NS3 protein function
Authors	:	Saalau-Bethell, S.M.; Woodhead, A.J.; Chessari, G.; Carr, M.G.; Coyle, J.; Graham, B.; Hiscock, S.D.; Murray, C.W.; Pathuri, P.; Rich, S.J.; Richardson, C.J.; Williams, P.A.; Jhoti, H.
Deposited on	:	2012-08-09
Resolution	:	2.89 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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:

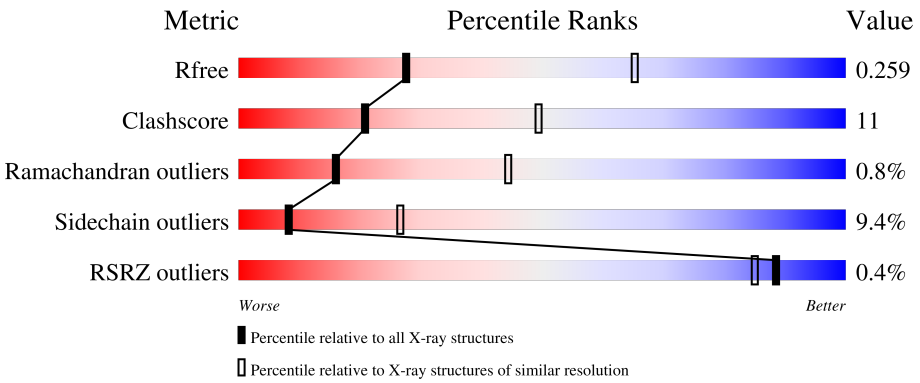
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.89 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	2481 (2.90-2.90)
Clashscore	190562	2690 (2.90-2.90)
Ramachandran outliers	187476	2623 (2.90-2.90)
Sidechain outliers	187428	2625 (2.90-2.90)
RSRZ outliers	180081	2481 (2.90-2.90)

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	683	69% 23% • 5%
1	B	683	% 68% 22% • 6%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 9828 atoms, of which 28 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 NON-STRUCTURAL PROTEIN 4A, SERINE PROTEASE NS3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	646	4808	3026	835	917	30	0	0	1
1	B	643	4794	3019	832	913	30	0	0	1

There are 58 discrepancies between the modelled and reference sequences:

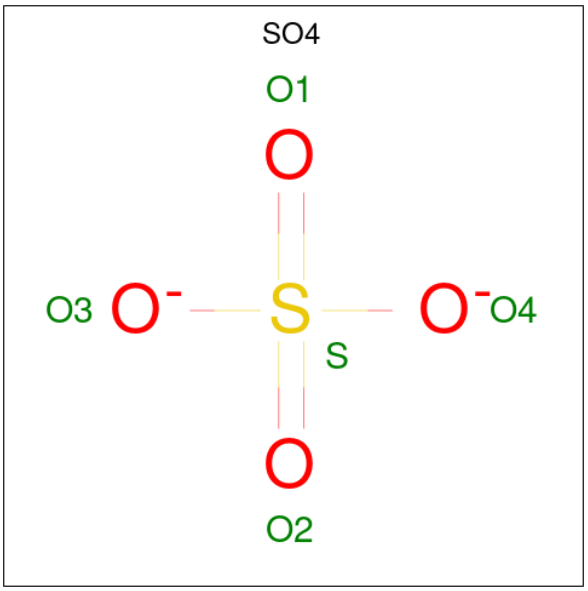
Chain	Residue	Modelled	Actual	Comment	Reference
A	-34	MET	-	expression tag	UNP P26663
A	-33	GLY	-	expression tag	UNP P26663
A	-32	SER	-	expression tag	UNP P26663
A	-31	SER	-	expression tag	UNP P26663
A	-30	HIS	-	expression tag	UNP P26663
A	-29	HIS	-	expression tag	UNP P26663
A	-28	HIS	-	expression tag	UNP P26663
A	-27	HIS	-	expression tag	UNP P26663
A	-26	HIS	-	expression tag	UNP P26663
A	-25	HIS	-	expression tag	UNP P26663
A	-24	SER	-	expression tag	UNP P26663
A	-23	SER	-	expression tag	UNP P26663
A	-22	GLY	-	expression tag	UNP P26663
A	-21	LEU	-	expression tag	UNP P26663
A	-20	VAL	-	expression tag	UNP P26663
A	-19	PRO	-	expression tag	UNP P26663
A	-18	ARG	-	expression tag	UNP P26663
A	-17	GLY	-	expression tag	UNP P26663
A	-16	SER	-	expression tag	UNP P26663
A	-15	HIS	-	expression tag	UNP P26663
A	-14	MET	-	expression tag	UNP P26663
A	-1	GLY	-	linker	UNP P26663
A	0	SER	-	linker	UNP P26663
A	1	GLY	-	linker	UNP P26663

Continued on next page...

Continued from previous page...

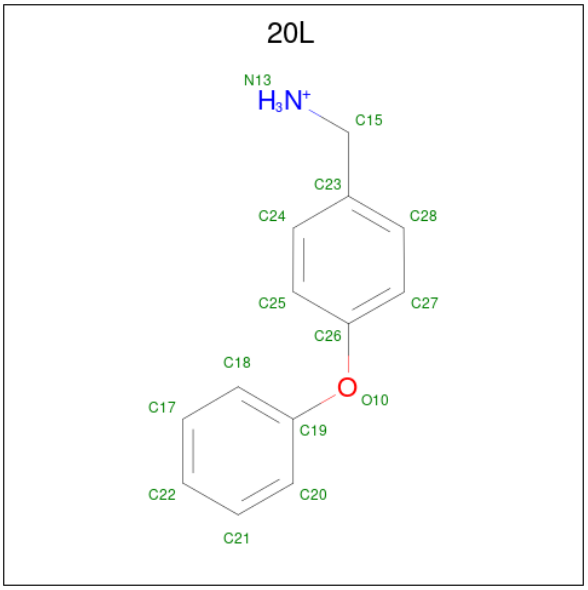
Chain	Residue	Modelled	Actual	Comment	Reference
A	2	SER	-	linker	UNP P26663
A	66	GLY	ALA	conflict	UNP P26663
A	86	GLN	PRO	conflict	UNP P26663
A	87	ALA	LYS	conflict	UNP P26663
A	147	SER	PHE	conflict	UNP P26663
B	-34	MET	-	expression tag	UNP P26663
B	-33	GLY	-	expression tag	UNP P26663
B	-32	SER	-	expression tag	UNP P26663
B	-31	SER	-	expression tag	UNP P26663
B	-30	HIS	-	expression tag	UNP P26663
B	-29	HIS	-	expression tag	UNP P26663
B	-28	HIS	-	expression tag	UNP P26663
B	-27	HIS	-	expression tag	UNP P26663
B	-26	HIS	-	expression tag	UNP P26663
B	-25	HIS	-	expression tag	UNP P26663
B	-24	SER	-	expression tag	UNP P26663
B	-23	SER	-	expression tag	UNP P26663
B	-22	GLY	-	expression tag	UNP P26663
B	-21	LEU	-	expression tag	UNP P26663
B	-20	VAL	-	expression tag	UNP P26663
B	-19	PRO	-	expression tag	UNP P26663
B	-18	ARG	-	expression tag	UNP P26663
B	-17	GLY	-	expression tag	UNP P26663
B	-16	SER	-	expression tag	UNP P26663
B	-15	HIS	-	expression tag	UNP P26663
B	-14	MET	-	expression tag	UNP P26663
B	-1	GLY	-	linker	UNP P26663
B	0	SER	-	linker	UNP P26663
B	1	GLY	-	linker	UNP P26663
B	2	SER	-	linker	UNP P26663
B	66	GLY	ALA	conflict	UNP P26663
B	86	GLN	PRO	conflict	UNP P26663
B	87	ALA	LYS	conflict	UNP P26663
B	147	SER	PHE	conflict	UNP P26663

- Molecule 2 is SULFATE ION (CCD ID: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
2	A	1	5	4	1	0	0
			Total	O	S		
2	B	1	5	4	1	0	0

- Molecule 3 is (4-phenoxyphenyl)methylazanium (CCD ID: 20L) (formula: $C_{13}H_{14}NO$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	N	O		
3	A	1	29	13	14	1	1	0	0
			Total	C	H	N	O		
3	B	1	29	13	14	1	1	0	0

- Molecule 4 is water.

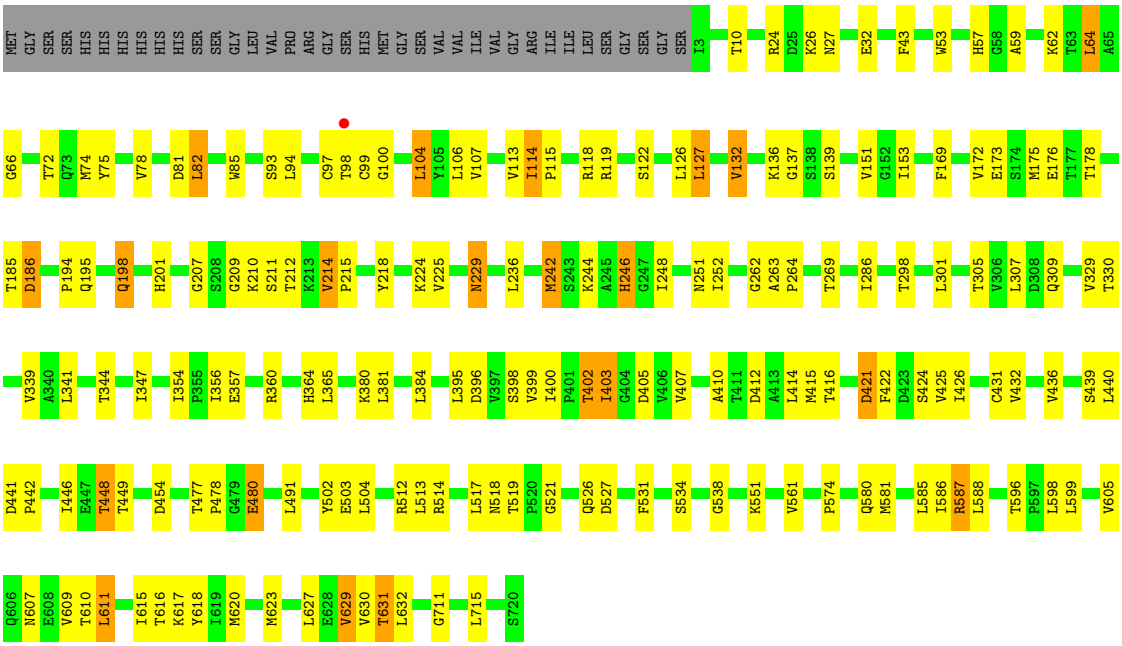
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	82	82	82	0	0
			Total	O		
4	B	76	76	76	0	0

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.

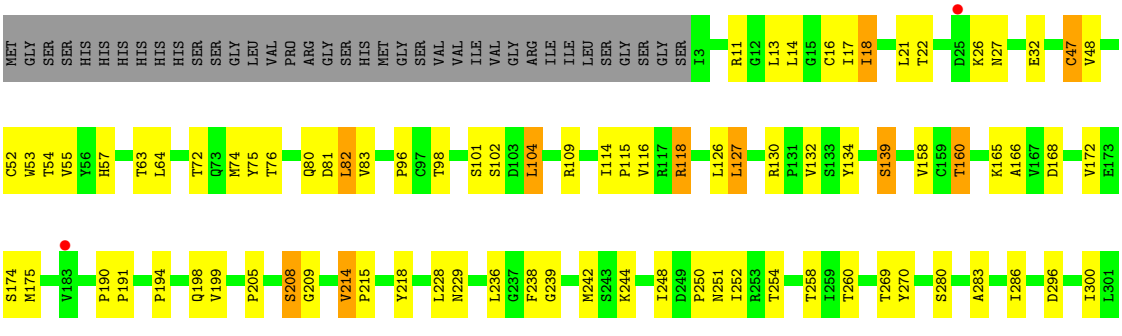
• Molecule 1: NON-STRUCTURAL PROTEIN 4A, SERINE PROTEASE NS3

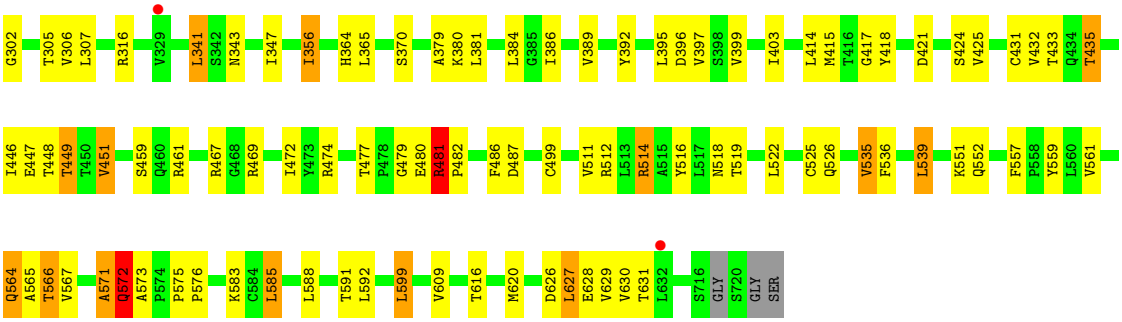
Chain A: 69% 23% • 5%



• Molecule 1: NON-STRUCTURAL PROTEIN 4A, SERINE PROTEASE NS3

Chain B: % 68% 22% • 6%




G302
T305
V306
L307
R316
V329
L341
S342
N343
I347
I356
H364
L365
S370
A379
K380
L381
L384
G385
I386
V389
Y392
L395
D396
V397
S398
V399
I403
L414
M415
T416
G417
Y418
D421
S424
V425
C431
V432
T433
Q434
T435
I446
E447
T448
T449
T450
V451
S459
Q460
R461
R467
G468
R469
I472
Y473
R474
T477
P478
G479
E480
R481
P482
F486
D487
C499
V511
R512
L513
R514
A515
Y516
L517
N518
T519
L522
C525
Q526
V535
F536
L539
K551
Q552
F557
P558
Y559
L560
V561
Q564
A565
T566
V567
A571
Q572
A573
P574
P575
P576
K583
C584
L585
L588
T591
L592
L599
V609
T616
M620
D626
L627
E628
V629
V630
T631
L632
S716
GLY
S720
GLY
SER

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	91.31Å 111.29Å 139.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	87.03 – 2.89 87.03 – 2.89	Depositor EDS
% Data completeness (in resolution range)	95.0 (87.03-2.89) 95.0 (87.03-2.89)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.36 (at 2.91Å)	Xtriage
Refinement program	REFMAC 5.2.0019I	Depositor
R, R_{free}	0.170 , 0.259 0.179 , 0.259	Depositor DCC
R_{free} test set	1655 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å^2)	45.2	Xtriage
Anisotropy	0.080	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 53.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	9828	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.17% 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: 20L, SO4

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	0.76	2/4916 (0.0%)	0.93	6/6713 (0.1%)
1	B	0.77	2/4901 (0.0%)	0.90	3/6692 (0.0%)
All	All	0.77	4/9817 (0.0%)	0.92	9/13405 (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
1	A	1	0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	631	THR	C-N	-6.83	1.23	1.33
1	B	631	THR	C-N	-6.32	1.24	1.33
1	B	76	THR	CA-CB	6.05	1.60	1.52
1	A	114	ILE	CA-CB	5.23	1.58	1.54

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	561	VAL	N-CA-C	-6.75	104.18	110.53
1	A	66	GLY	CA-C-N	6.33	126.53	119.32
1	A	66	GLY	C-N-CA	6.33	126.53	119.32
1	B	479	GLY	N-CA-C	6.15	118.56	111.36
1	A	402	THR	N-CA-C	6.07	120.85	112.90

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	403	ILE	CA

There are no planarity outliers.

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	4808	0	4783	103	0
1	B	4794	0	4771	108	0
2	A	5	0	0	0	0
2	B	5	0	0	0	0
3	A	15	14	14	2	0
3	B	15	14	14	3	0
4	A	82	0	0	4	0
4	B	76	0	0	2	0
All	All	9800	28	9582	213	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 11.

The worst 5 of 213 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:139:SER:OG	1:A:632:LEU:N	1.94	1.00
1:A:616:THR:HG22	1:A:620:MET:HE2	1.53	0.90
1:B:74:MET:HE3	1:B:75:TYR:CE2	2.13	0.83
1:A:356:ILE:HD11	1:A:384:LEU:HD23	1.60	0.83
1:A:210:LYS:HA	1:A:214:VAL:HG13	1.62	0.81

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	642/683 (94%)	592 (92%)	44 (7%)	6 (1%)	14 41
1	B	638/683 (93%)	596 (93%)	38 (6%)	4 (1%)	21 51
All	All	1280/1366 (94%)	1188 (93%)	82 (6%)	10 (1%)	16 44

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	246	HIS
1	A	403	ILE
1	B	572	GLN
1	A	100	GLY
1	A	244	LYS

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	526/557 (94%)	477 (91%)	49 (9%)	8 27
1	B	525/557 (94%)	475 (90%)	50 (10%)	8 26
All	All	1051/1114 (94%)	952 (91%)	99 (9%)	8 26

5 of 99 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	98	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	356	ILE
1	B	109	ARG
1	B	172	VAL
1	B	403	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	549	GLN
1	B	606	GLN
1	B	9	GLN
1	B	518	ASN
1	A	607	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

4 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	SO4	B	1721	-	4,4,4	0.28	0	6,6,6	0.20	0
3	20L	B	1722	-	16,16,16	0.34	0	20,20,20	0.58	1 (5%)
2	SO4	A	1721	-	4,4,4	0.23	0	6,6,6	0.21	0
3	20L	A	1722	-	16,16,16	0.35	0	20,20,20	0.51	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
3	20L	B	1722	-	-	0/6/6/6	0/2/2/2
3	20L	A	1722	-	-	0/6/6/6	0/2/2/2

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	1722	20L	C26-O10-C19	2.06	123.48	118.78

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1722	20L	3	0
3	A	1722	20L	2	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 (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	646/683 (94%)	-0.23	1 (0%)	91	89	10, 36, 66, 80	0
1	B	643/683 (94%)	-0.21	4 (0%)	85	81	11, 38, 67, 95	0
All	All	1289/1366 (94%)	-0.22	5 (0%)	88	85	10, 38, 67, 95	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	183	VAL	4.2
1	B	632	LEU	3.0
1	A	98	THR	2.5
1	B	25	ASP	2.2
1	B	329	VAL	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 oligosaccharides in this entry.

6.4 Ligands (i)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	SO4	B	1721	5/5	0.65	0.19	119,120,121,121	0
3	20L	B	1722	15/15	0.93	0.12	73,74,80,80	0
3	20L	A	1722	15/15	0.95	0.10	62,64,66,66	0
2	SO4	A	1721	5/5	0.97	0.14	54,55,57,57	0

6.5 Other polymers (i)

There are no such residues in this entry.