W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

May 11, 2016 – 05:24 PM EDT

PDB ID : 5HCE
Title : Ternary complex of human Complement C5 with Ornithodoros moubata OmCI and Rhipicephalus appendiculatus RaCI1
Authors : Jore, M.M.; Johnson, S.; Lea, S.M.
Deposited on : 2016-01-04
Resolution : 3.12 Å(reported)

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

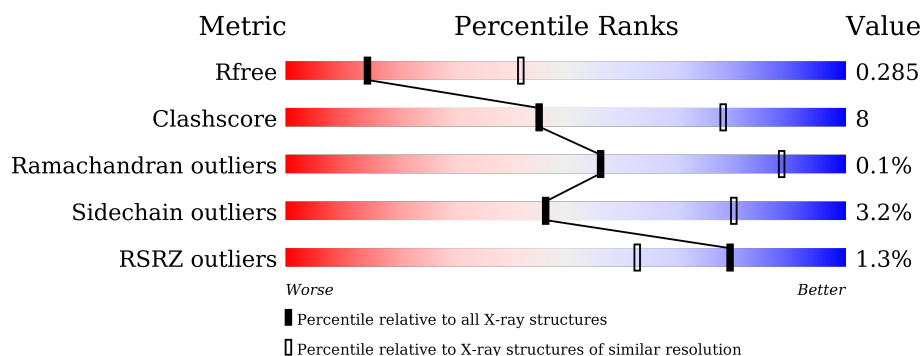|		
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027457
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027457

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 3.12 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		0.285
Clashscore		8
Ramachandran outliers		0.1%
Sidechain outliers		3.2%
RSRZ outliers		1.3%

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}	91344	1112 (3.16-3.08)
Clashscore	102246	1218 (3.16-3.08)
Ramachandran outliers	100387	1175 (3.16-3.08)
Sidechain outliers	100360	1175 (3.16-3.08)
RSRZ outliers	91569	1114 (3.16-3.08)

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 on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. 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	B	656	.% 77% 21% ••
2	A	998	2% 79% 18% ••
3	C	165	.% 78% 12% 10%
4	D	81	33% 20% 5% 42%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 14418 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 Complement C5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	647	Total	C	N	O	S	0	0	0
			5116	3281	819	1003	13			

- Molecule 2 is a protein called Complement C5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	982	Total	C	N	O	S	0	0	0
			7775	4973	1298	1463	41			

- Molecule 3 is a protein called Complement inhibitor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	148	Total	C	N	O	S	0	0	0
			1160	715	195	239	11			

There are 17 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	4	MET	-	initiating methionine	UNP Q5YD59
C	5	ALA	-	expression tag	UNP Q5YD59
C	6	SER	-	expression tag	UNP Q5YD59
C	7	HIS	-	expression tag	UNP Q5YD59
C	8	HIS	-	expression tag	UNP Q5YD59
C	9	HIS	-	expression tag	UNP Q5YD59
C	10	HIS	-	expression tag	UNP Q5YD59
C	11	HIS	-	expression tag	UNP Q5YD59
C	12	HIS	-	expression tag	UNP Q5YD59
C	13	HIS	-	expression tag	UNP Q5YD59
C	14	HIS	-	expression tag	UNP Q5YD59
C	15	HIS	-	expression tag	UNP Q5YD59
C	16	HIS	-	expression tag	UNP Q5YD59
C	17	SER	-	expression tag	UNP Q5YD59

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	18	GLY	-	expression tag	UNP Q5YD59
C	78	GLN	ASN	engineered mutation	UNP Q5YD59
C	102	GLN	ASN	engineered mutation	UNP Q5YD59

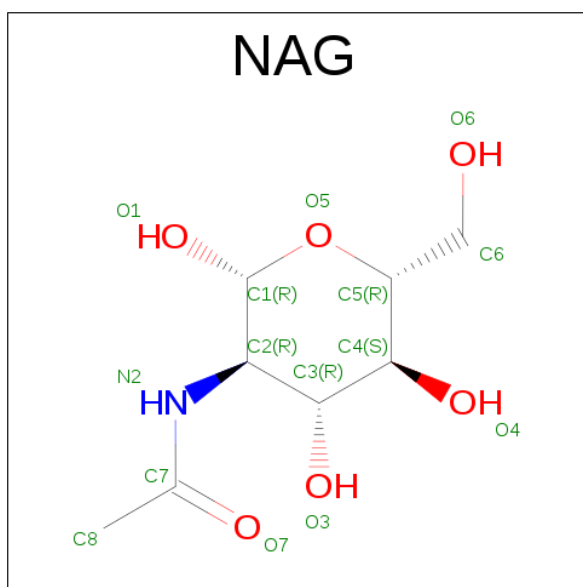• Molecule 4 is a protein called Rhipicephalus appendiculatus RaCI1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	47	333	203	62	62	6	0	0	0

• Molecule 5 is CYSTEINE (three-letter code: CYS) (formula: $C_3H_7NO_2S$).
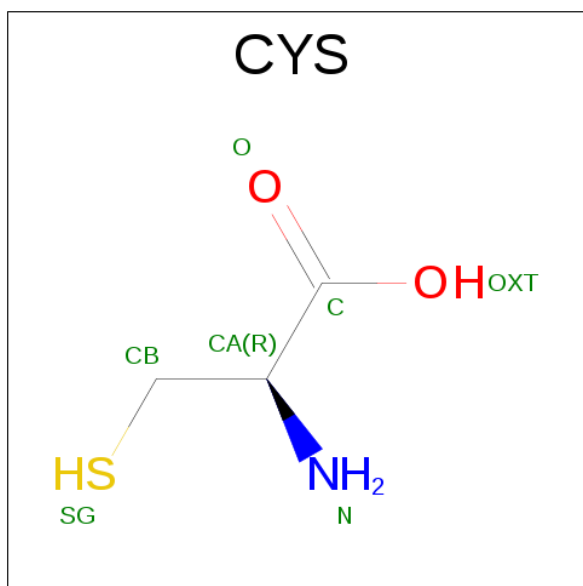
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
5	A	1	6	3	1	1	1	0	0

• Molecule 6 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

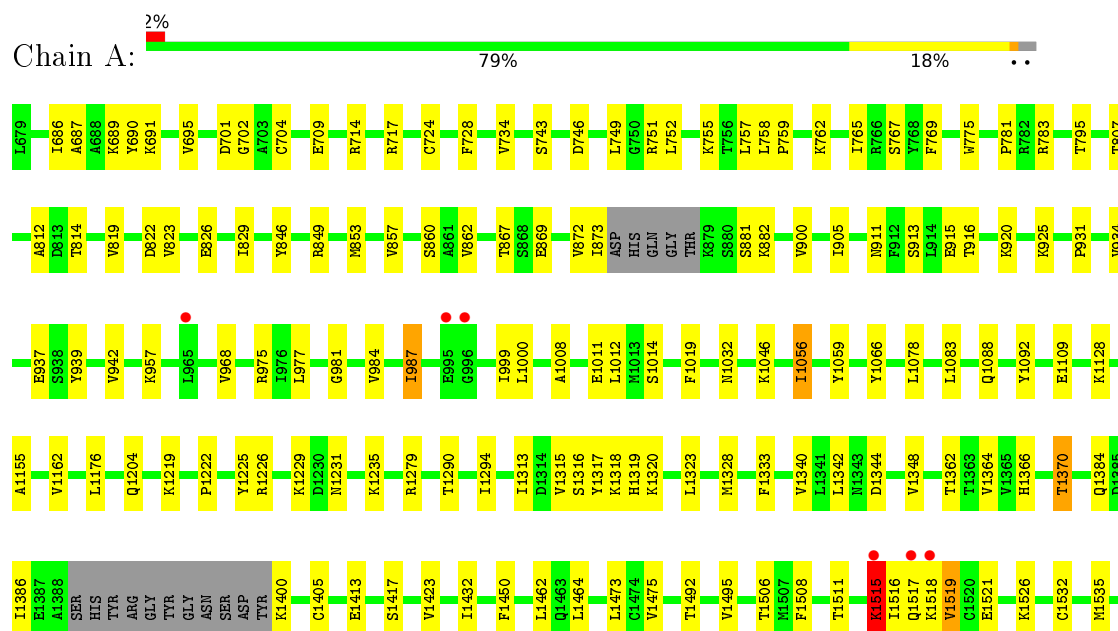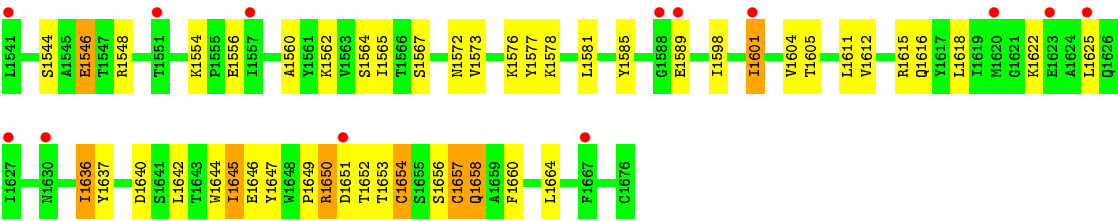| Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	A	1	Total	C	N	O	0	0
			14	8	1	5		

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.
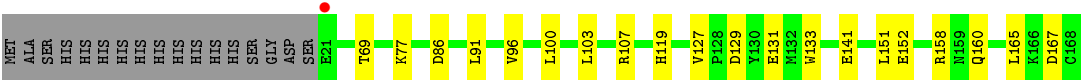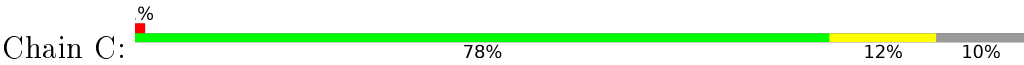
- Molecule 1: Complement C5
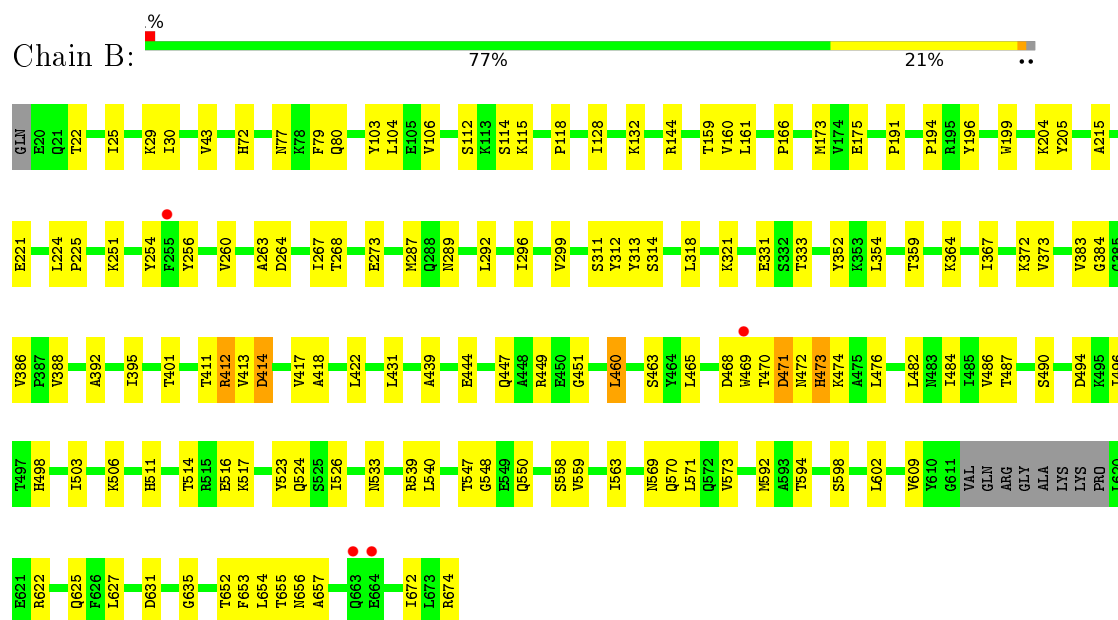
- Molecule 2: Complement C5

L1541 S1544 A1545 E1546 T1547 R1548 T1551 K1554 P1555 E1556 I1557 A1560 Y1561 K1562 V1563 S1564 I1565 T1566 S1567 N1572 V1573 K1576 Y1577 K1578 L1581 Y1585 G1588 E1589 I1598 I1601 V1604 T1605 L1611 V1612 R1615 Q1616 Y1617 L1618 I1619 M1620 G1621 K1622 E1623 A1624 L1625 Q1626

I1627 N1630 I1636 Y1637 D1640 S1641 L1642 T1643 W1644 I1645 E1646 Y1647 W1648 P1649 R1650 D1651 T1652 T1653 C1654 S1655 S1656 C1657 Q1658 A1659 F1660 L1664 F1667 C1676

• Molecule 3: Complement inhibitor

Chain C:

%

78% 12% 10%

MET ALA SER HIS HIS HIS HIS HIS HIS HIS HIS HIS SER GLY ASP SER E21 T69 K77 D86 L91 V96 L100 L103 R107 H119 V127 P128 D129 Y130 E131 M132 W133 E141 L151 E152 R158 N159 Q160 L165 K166 D167 C168

• Molecule 4: Rhipicephalus appendiculatus RaCI1

Chain D:

33% 20% 5% 42%

GLY PRO MET GLU GLU VAL LYS THR THR PRO ILE PRO ASN HIS Q13 C14 C19 E20 R21 K22 L23 D24 A25 I26 G27 N28 T32 Q36 L39 R43 G44 A45 I48 A51 Q57 L58 A59 THR THR LYS PRO PRO MET ALA PRO GLY ASP ASN LYS ASP ASN LYS

GLU GLU GLU SER ASN
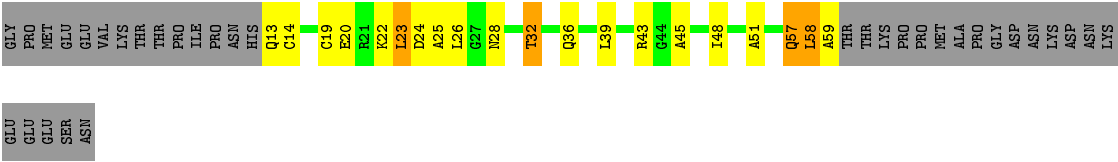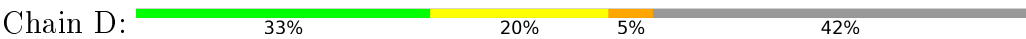
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	105.26Å 140.72Å 210.37Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	74.40 – 3.12 94.13 – 3.12	Depositor EDS
% Data completeness (in resolution range)	98.9 (74.40-3.12) 99.1 (94.13-3.12)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	0.15	Depositor
$< I/\sigma(I) >$ [1]	1.63 (at 3.13Å)	Xtriage
Refinement program	PHENIX (1.10_2155: ???)	Depositor
R, R_{free}	0.267 , 0.281 0.274 , 0.285	Depositor DCC
R_{free} test set	5264 reflections (4.95%)	DCC
Wilson B-factor (Å^2)	70.2	Xtriage
Anisotropy	0.329	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 23.9	EDS
L-test for twinning[2]	$< |L| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	14418	wwPDB-VP
Average B, all atoms (Å^2)	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.53% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: NAG

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	B	0.34	0/5234	0.51	0/7119
2	A	0.36	0/7932	0.50	1/10740 (0.0%)
3	C	0.30	0/1183	0.47	0/1599
4	D	0.39	0/335	0.55	0/454
All	All	0.35	0/14684	0.51	1/19912 (0.0%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1515	LYS	N-CA-C	-5.83	95.27	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	468	ASP	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	B	5116	0	5022	87	0
2	A	7775	0	7822	124	0
3	C	1160	0	1085	10	0
4	D	333	0	335	14	0
5	A	6	0	3	0	0
6	A	28	0	25	0	0
All	All	14418	0	14292	225	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.

The worst 5 of 225 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:B:469:TRP:HA	1:B:484:ILE:HA	1.60	0.84
3:C:77:LYS:NZ	3:C:167:ASP:O	2.17	0.77
1:B:359:THR:HG21	1:B:372:LYS:H	1.49	0.77
2:A:1625:LEU:HB2	2:A:1636:ILE:HG23	1.64	0.77
1:B:384:GLY:HA3	1:B:413:VAL:HG23	1.68	0.76

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	B	643/656 (98%)	616 (96%)	27 (4%)	0	100 100

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	A	976/998 (98%)	924 (95%)	52 (5%)	0	100	100
3	C	146/165 (88%)	138 (94%)	8 (6%)	0	100	100
4	D	45/81 (56%)	39 (87%)	5 (11%)	1 (2%)	8	37
All	All	1810/1900 (95%)	1717 (95%)	92 (5%)	1 (0%)	56	88

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	14	CYS

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	B	574/581 (99%)	559 (97%)	15 (3%)	54	83
2	A	872/885 (98%)	845 (97%)	27 (3%)	47	80
3	C	128/143 (90%)	125 (98%)	3 (2%)	58	85
4	D	37/68 (54%)	31 (84%)	6 (16%)	3	12
All	All	1611/1677 (96%)	1560 (97%)	51 (3%)	46	80

5 of 51 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	1515	LYS
2	A	1562	LYS
4	D	32	THR
2	A	1519	VAL
2	A	1581	LEU

Some 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 carbohydrates in this entry.

5.6   Ligand geometry (i)

3 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	$\#|Z| > 2$
5	CYS	A	2101	2	3,5,6	0.72	0	3,5,7	1.31	1 (33%)
6	NAG	A	2102	2,6	14,14,15	0.47	0	15,19,21	0.50	0
6	NAG	A	2103	6	14,14,15	0.26	0	15,19,21	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
5	CYS	A	2101	2	-	0/1/4/6	0/0/0/0
6	NAG	A	2102	2,6	-	0/6/23/26	0/1/1/1
6	NAG	A	2103	6	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	2101	CYS	O-C-CA	-2.04	120.26	125.72

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	B	647/656 (98%)	0.18	4 (0%)	90	81	49, 76, 118, 157	0
2	A	982/998 (98%)	0.30	19 (1%)	70	48	44, 66, 160, 215	0
3	C	148/165 (89%)	0.14	1 (0%)	89	79	61, 85, 123, 138	0
4	D	47/81 (58%)	0.01	0	100	100	63, 78, 98, 121	0
All	All	1824/1900 (96%)	0.24	24 (1%)	79	63	44, 72, 144, 215	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	1518	LYS	9.1
2	A	1517	GLN	7.5
2	A	996	GLY	4.6
3	C	21	GLU	3.9
2	A	995	GLU	3.5

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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
5	CYS	A	2101	6/7	0.75	0.25	0.25	106,111,116,122	0
6	NAG	A	2102	14/15	0.91	0.18	-0.98	60,67,87,98	0
6	NAG	A	2103	14/15	0.86	0.17	-	89,98,108,110	0

6.5   Other polymers (i)

There are no such residues in this entry.