



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2018 – 12:44 am GMT

PDB ID : 2B39
Title : Structure of mammalian C3 with an intact thioester at 3Å resolution
Authors : Fredslund, F.; Jenner, L.; Husted, L.B.; Nyborg, J.; Andersen, G.R.; Sottrup-Jensen, L.
Deposited on : 2005-09-20
Resolution : 3.00 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : trunk30967
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30967

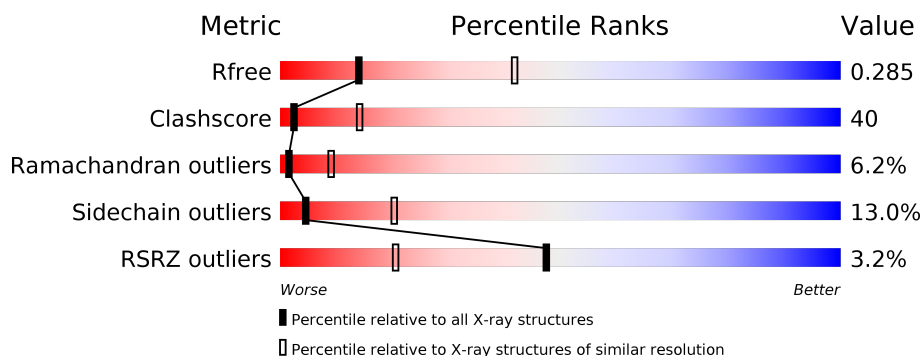
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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}	111664	1851 (3.00-3.00)
Clashscore	122126	2167 (3.00-3.00)
Ramachandran outliers	120053	2101 (3.00-3.00)
Sidechain outliers	120020	2104 (3.00-3.00)
RSRZ outliers	108989	1751 (3.00-3.00)

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 $\leq 5\%$. 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	1661	<div> <div>3%</div> <div>39%</div> <div>46%</div> <div>11%</div> <div>••</div> </div>
1	B	1661	<div> <div>3%</div> <div>38%</div> <div>48%</div> <div>10%</div> <div>••</div> </div>

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
2	NAG	A	2001	X	-	-	-
2	NAG	B	2001	X	-	-	-
3	BMA	A	2003	-	-	-	X

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 25624 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 C3.

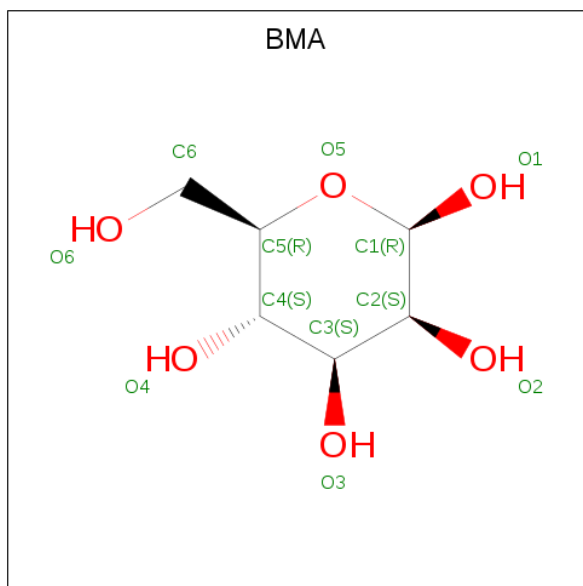
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1610	Total	C	N	O	S	0	0	0
			12773	8093	2187	2439	54			
1	B	1610	Total	C	N	O	S	0	0	0
			12773	8093	2187	2439	54			

- Molecule 2 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



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

- Molecule 3 is BETA-D-MANNOSE (three-letter code: BMA) (formula: $C_6H_{12}O_6$).

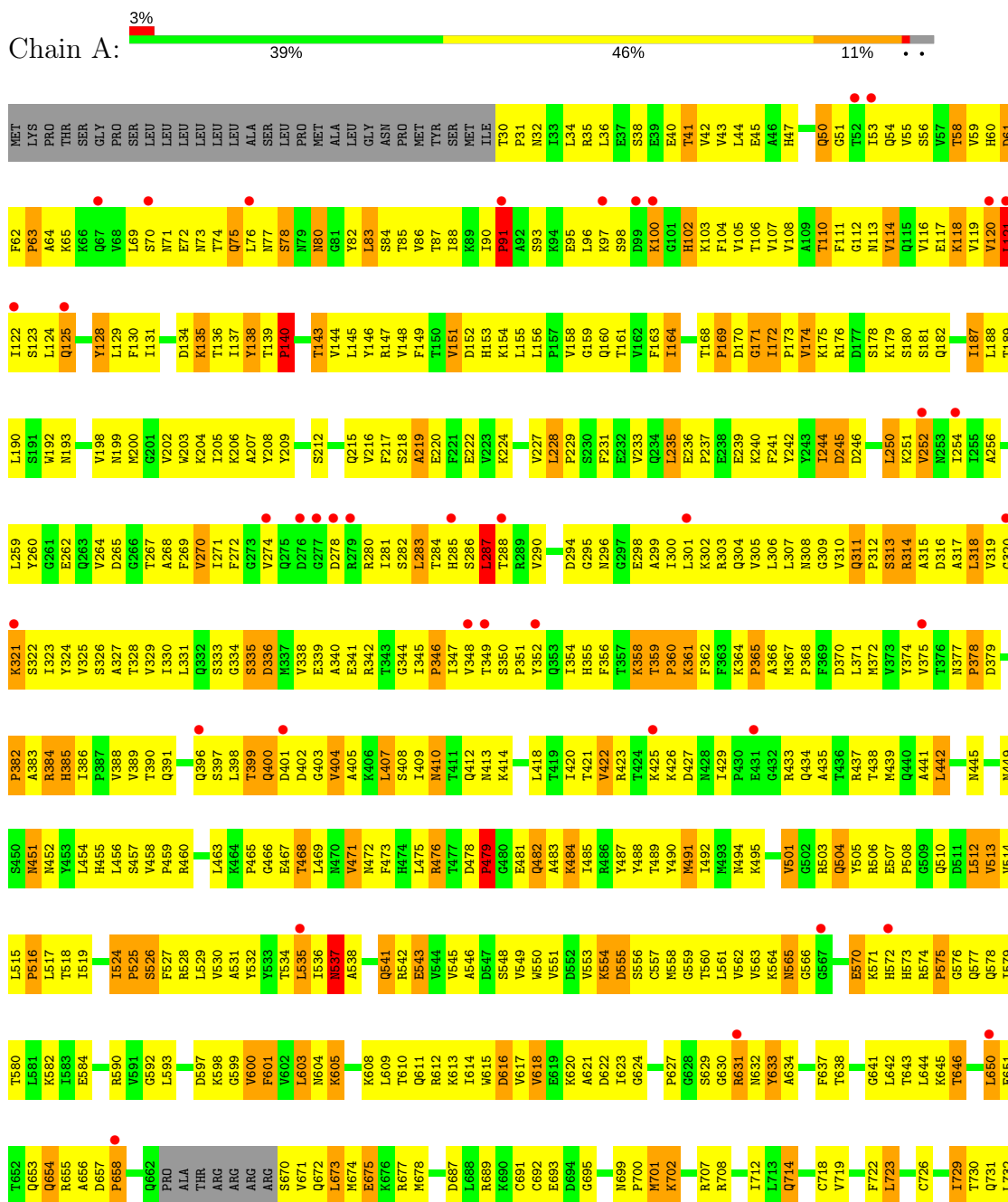


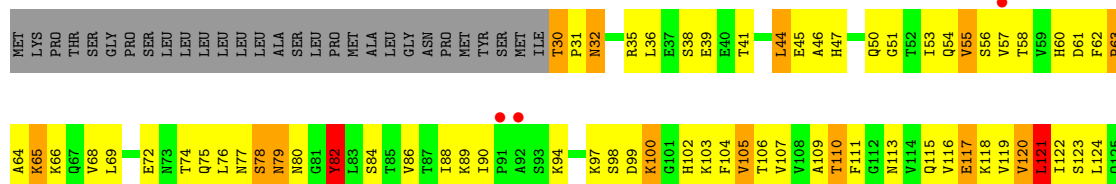
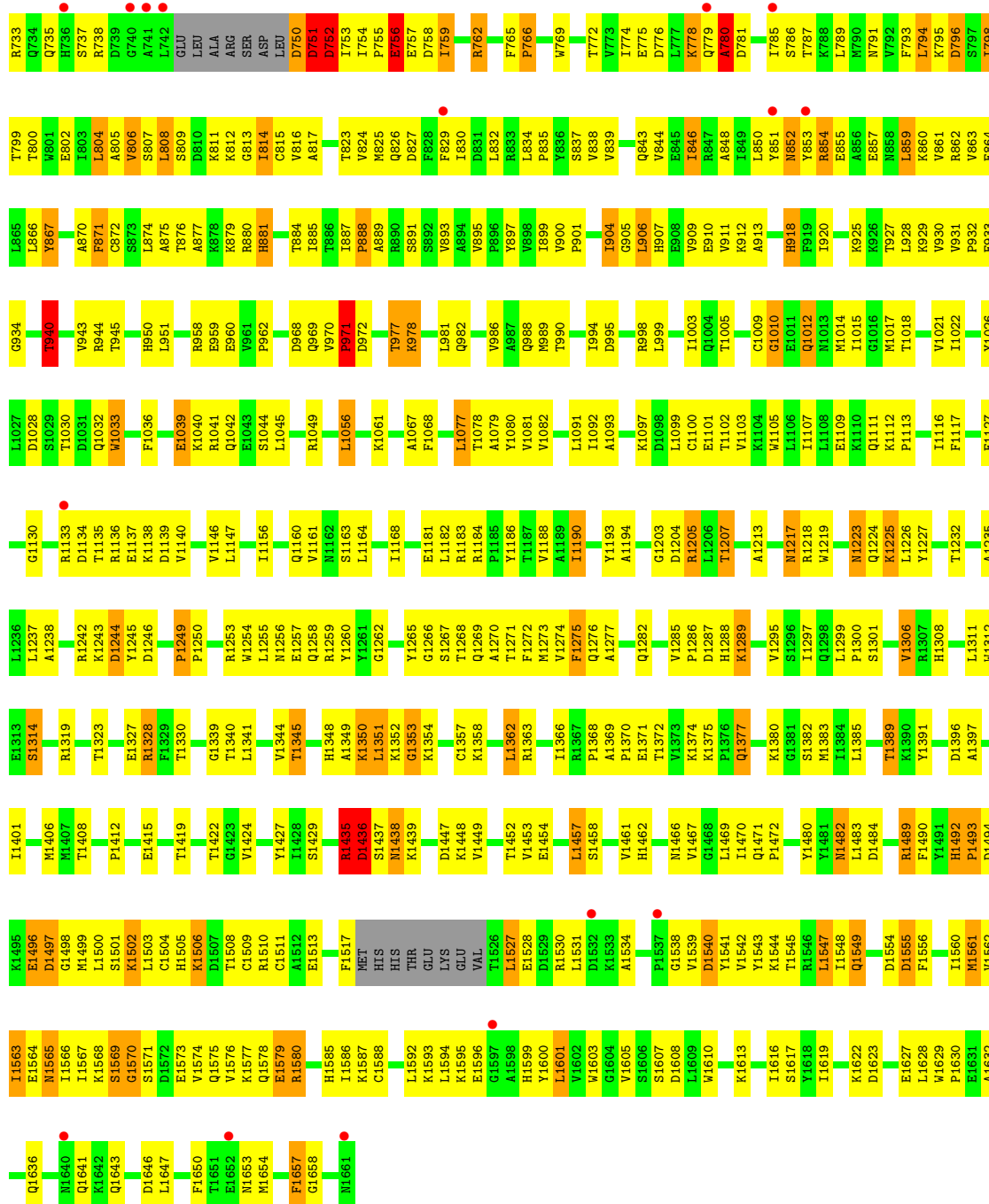
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			11	6	5		
3	B	1	Total	C	O	0	0
			11	6	5		

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 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: C3







E1612	K1613	P1614	K1615	I1616	S1617	I1620	G1621	K1622	D1623	E1627	L1628	W1629	P1630	E1631	A1632	E1633	E1634	C1635	Q1636	D1637	E1638	E1639	N1640	Q1641	K1642	Q1643	D1646	L1647	A1648	N1649	F1650	T1651	E1652	N1653	M1654	V1655	V1656	F1657	G1658	C1659	P1660	N1661															
Q1549	K1550	K1551	L1552	E1553	D1554	I1558	Y1559	I1560	M1561	V1562	I1563	E1564	N1565	I1566	K1567	K1568	S1569	G1570	S1571	D1572	E1573	V1574	Q1575	V1576	K1577	Q1578	E1579	R1580	K1581	F1582	I1583	S1584	H1585	I1586	K1587	C1588	R1589	E1590	A1591	L1592	K1593	E1596	G1597	A1598	H1599	Y1600	L1601	V1602	W1603	G1604	V1605	S1606	S1607	D1608	L1609	W1610	G1611
Y1480	Y1481	M1482	L1483	D1484	E1485	I1488	R1489	F1490	P1493	D1494	K1495	E1496	D1497	G1498	M1499	L1500	S1501	K1502	L1503	C1504	H1505	K1506	D1507	T1508	C1509	R1510	E1513	F1517	MET	HIS	THR	GLU	LYS	GLU	VAL	T1526	L1527	R1530	L1531	D1532	K1533	A1534	C1535	E1536	Y1541	V1542	Y1543	K1544	T1545	R1546	L1547	T1548					
D1396	A1397	I1401	I1404	S1405	M1406	M1407	T1408	G1409	F1410	S1411	P1412	E1415	K1418	T1419	L1420	S1421	T1422	G1423	V1424	N1434	R1435	D1436	S1437	M1438	K1439	H1440	T1441	Y1445	L1446	D1447	K1448	V1449	S1450	H1451	T1452	V1453	E1454	D1455	C1456	L1457	S1458	V1461	H1462	Q1463	V1467	G1468	L1469	A1474	V1475								
W1312	E1313	S1314	L1318	R1319	E1322	T1323	R1328	K1332	Q1338	G1339	T1340	L1341	S1342	V1343	V1344	T1345	V1346	Y1347	H1348	A1349	A1350	L1351	K1352	G1353	C1357	K1358	K1359	L1362	R1363	I1366	R1367	P1368	A1369	P1370	E1371	T1372	K1375	P1376	Q1377	S1382	M1383	L1384	L1385	Y1391	L1392	G1393	D1394	Q1395									

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	254.25Å 246.86Å 113.35Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 3.00 37.85 – 3.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) (40.00-3.00) 97.2 (37.85-3.00)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.73 (at 3.01Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.278 , 0.286 0.273 , 0.285	Depositor DCC
R_{free} test set	1392 reflections (1.00%)	wwPDB-VP
Wilson B-factor (Å ²)	69.3	Xtriage
Anisotropy	0.479	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 75.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.008 for k,h,-l	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	25624	wwPDB-VP
Average B, all atoms (Å ²)	97.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 18.36% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: BMA, 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	A	0.58	1/13020 (0.0%)	0.82	15/17632 (0.1%)
1	B	0.57	0/13020	0.81	11/17632 (0.1%)
All	All	0.57	1/26040 (0.0%)	0.82	26/35264 (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	0	1
1	B	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	102	HIS	CB-CG	6.64	1.61	1.50

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	121	LEU	CA-CB-CG	-9.75	92.87	115.30
1	A	91	PRO	N-CA-C	7.91	132.66	112.10
1	B	1362	LEU	CA-CB-CG	6.57	130.40	115.30
1	A	1362	LEU	CA-CB-CG	6.54	130.34	115.30
1	A	1130	GLY	N-CA-C	-6.49	96.87	113.10

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	138	TYR	Sidechain
1	B	820	TYR	Sidechain

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	12773	0	12782	1027	0
1	B	12773	0	12782	1002	0
2	A	28	0	24	3	0
2	B	28	0	24	2	0
3	A	11	0	10	0	0
3	B	11	0	10	1	0
All	All	25624	0	25632	2029	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 40.

The worst 5 of 2029 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:117:GLU:O	1:B:118:LYS:HG2	1.47	1.14
1:B:369:PHE:HB3	1:B:409:ILE:HD12	1.32	1.10
1:A:44:LEU:HD13	1:A:55:VAL:HG11	1.34	1.09
1:B:244:ILE:HD11	1:B:319:VAL:CG2	1.84	1.08
1:B:116:VAL:HG13	1:B:645:LYS:HG2	1.28	1.06

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	1602/1661 (96%)	1285 (80%)	213 (13%)	104 (6%)	1	7
1	B	1602/1661 (96%)	1268 (79%)	239 (15%)	95 (6%)	2	10
All	All	3204/3322 (96%)	2553 (80%)	452 (14%)	199 (6%)	1	9

5 of 199 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	63	PRO
1	A	91	PRO
1	A	187	ILE
1	A	212	SER
1	A	244	ILE

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	1420/1465 (97%)	1238 (87%)	182 (13%)	5	21
1	B	1420/1465 (97%)	1234 (87%)	186 (13%)	4	20
All	All	2840/2930 (97%)	2472 (87%)	368 (13%)	4	20

5 of 368 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1447	ASP
1	B	151	VAL
1	B	1351	LEU
1	A	1492	HIS
1	B	54	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 85 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	71	ASN
1	B	296	ASN
1	B	1298	GLN
1	B	79	ASN
1	B	184	GLN

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](#)

6 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	NAG	A	2001	1,2	14,14,15	0.56	0	17,19,21	1.08	1 (5%)
2	NAG	A	2002	3,2	14,14,15	0.61	0	17,19,21	0.94	1 (5%)
3	BMA	A	2003	2	11,11,12	0.46	0	15,15,17	0.55	0
2	NAG	B	2001	1,2	14,14,15	0.56	0	17,19,21	1.00	0
2	NAG	B	2002	3,2	14,14,15	0.60	0	17,19,21	1.09	2 (11%)
3	BMA	B	2003	2	11,11,12	0.47	0	15,15,17	0.69	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
2	NAG	A	2001	1,2	1/1/5/7	0/6/23/26	0/1/1/1
2	NAG	A	2002	3,2	-	0/6/23/26	0/1/1/1
3	BMA	A	2003	2	-	0/2/19/22	0/1/1/1
2	NAG	B	2001	1,2	1/1/5/7	0/6/23/26	0/1/1/1
2	NAG	B	2002	3,2	-	0/6/23/26	0/1/1/1
3	BMA	B	2003	2	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	2001	NAG	O5-C1-C2	-2.89	107.54	111.52
2	B	2002	NAG	O5-C1-C2	-2.52	108.04	111.52
2	A	2002	NAG	C3-C4-C5	-2.34	106.06	110.24
2	B	2002	NAG	O5-C5-C6	2.07	110.42	107.15

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	2001	NAG	C1
2	B	2001	NAG	C1

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	2001	NAG	3	0
2	A	2002	NAG	2	0
2	B	2001	NAG	1	0
2	B	2002	NAG	2	0
3	B	2003	BMA	1	0

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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	1610/1661 (96%)	-0.07	55 (3%)	45	19	15, 101, 169, 200	0
1	B	1610/1661 (96%)	-0.12	47 (2%)	51	23	17, 96, 169, 200	0
All	All	3220/3322 (96%)	-0.09	102 (3%)	47	20	15, 99, 169, 200	0

The worst 5 of 102 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	671	VAL	6.0
1	A	742	LEU	5.6
1	A	741	ALA	5.2
1	A	276	ASP	5.0
1	A	348	VAL	4.9

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. 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(\AA^2)	Q<0.9
3	BMA	A	2003	11/12	0.72	0.41	81,81,82,82	0
3	BMA	B	2003	11/12	0.79	0.23	80,81,81,82	0
2	NAG	A	2002	14/15	0.80	0.32	80,81,82,82	0
2	NAG	A	2001	14/15	0.83	0.24	74,80,81,82	0
2	NAG	B	2002	14/15	0.88	0.26	80,80,81,82	0
2	NAG	B	2001	14/15	0.93	0.15	73,80,81,82	0

6.5 Other polymers [i](#)

There are no such residues in this entry.