



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 10:01 PM UTC

PDB ID : 8WY1 / pdb_00008wy1
Title : The structure of cyclization domain in cyclic beta-1,2-glucan synthase from *Thermoanaerobacter italicus*
Authors : Tanaka, N.; Saito, R.; Kobayashi, K.; Nakai, H.; Kamo, S.; Kuramochi, K.; Taguchi, H.; Nakajima, M.; Masaike, T.
Deposited on : 2023-10-30
Resolution : 3.90 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 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	:	4-5-2 with Phenix2.0
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

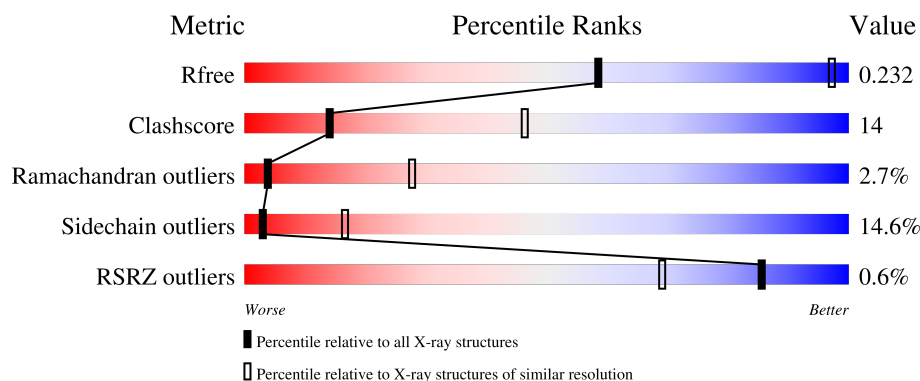
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.90 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	1270 (4.10-3.70)
Clashscore	190562	1034 (4.08-3.72)
Ramachandran outliers	187476	1251 (4.10-3.70)
Sidechain outliers	187428	1243 (4.10-3.70)
RSRZ outliers	180081	1269 (4.10-3.70)

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 $\leq 5\%$. 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	596	58% 33% 5% .
1	B	596	57% 34% 5% .
1	C	596	57% 31% 7% .
1	D	596	% 54% 36% 6% .

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 18842 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 Glycosyltransferase 36.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	574	Total	C	N	O	S	0	0	0
			4715	3047	772	881	15			
1	B	573	Total	C	N	O	S	0	0	0
			4706	3042	771	878	15			
1	C	573	Total	C	N	O	S	0	0	0
			4706	3042	771	878	15			
1	D	574	Total	C	N	O	S	0	0	0
			4715	3047	772	881	15			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1004	MET	-	initiating methionine	UNP D3T4C1
A	1592	LEU	-	expression tag	UNP D3T4C1
A	1593	GLU	-	expression tag	UNP D3T4C1
A	1594	HIS	-	expression tag	UNP D3T4C1
A	1595	HIS	-	expression tag	UNP D3T4C1
A	1596	HIS	-	expression tag	UNP D3T4C1
A	1597	HIS	-	expression tag	UNP D3T4C1
A	1598	HIS	-	expression tag	UNP D3T4C1
A	1599	HIS	-	expression tag	UNP D3T4C1
B	1004	MET	-	initiating methionine	UNP D3T4C1
B	1592	LEU	-	expression tag	UNP D3T4C1
B	1593	GLU	-	expression tag	UNP D3T4C1
B	1594	HIS	-	expression tag	UNP D3T4C1
B	1595	HIS	-	expression tag	UNP D3T4C1
B	1596	HIS	-	expression tag	UNP D3T4C1
B	1597	HIS	-	expression tag	UNP D3T4C1
B	1598	HIS	-	expression tag	UNP D3T4C1
B	1599	HIS	-	expression tag	UNP D3T4C1
C	1004	MET	-	initiating methionine	UNP D3T4C1
C	1592	LEU	-	expression tag	UNP D3T4C1
C	1593	GLU	-	expression tag	UNP D3T4C1

Continued on next page...

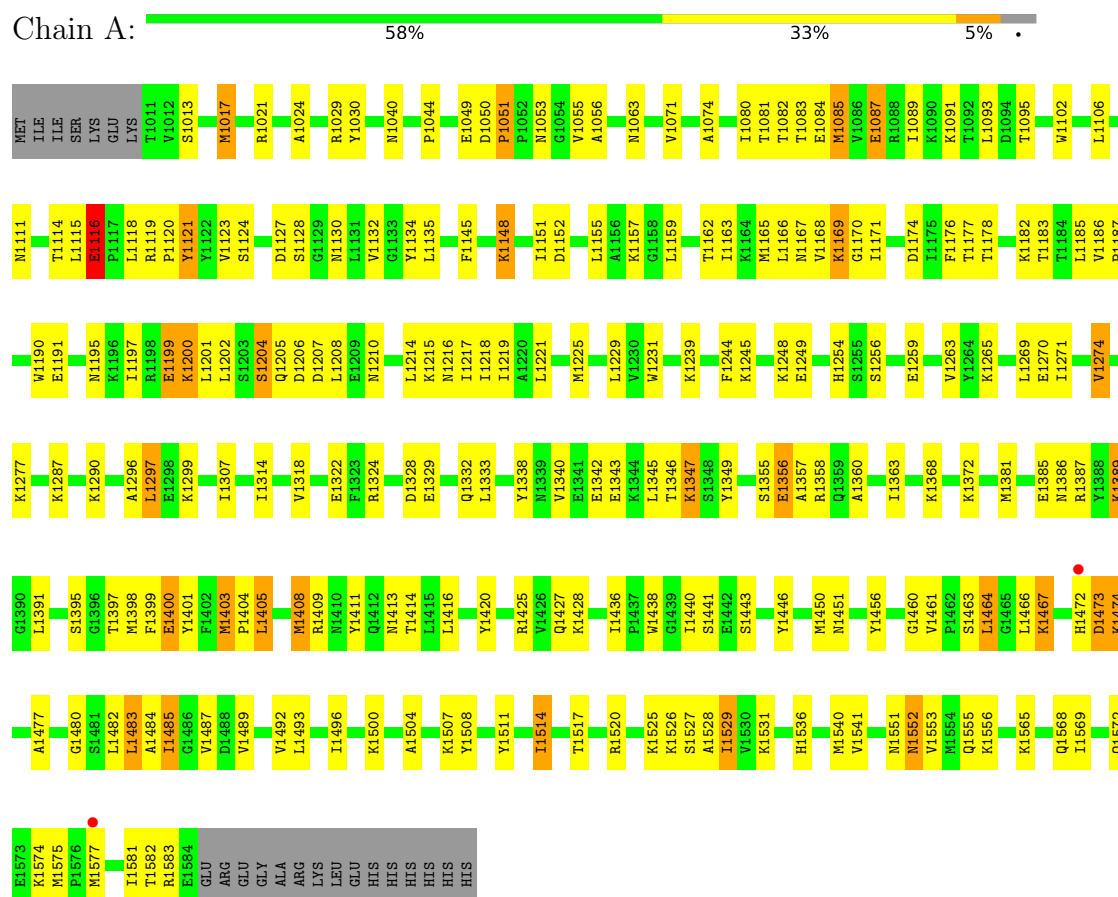
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	1594	HIS	-	expression tag	UNP D3T4C1
C	1595	HIS	-	expression tag	UNP D3T4C1
C	1596	HIS	-	expression tag	UNP D3T4C1
C	1597	HIS	-	expression tag	UNP D3T4C1
C	1598	HIS	-	expression tag	UNP D3T4C1
C	1599	HIS	-	expression tag	UNP D3T4C1
D	1004	MET	-	initiating methionine	UNP D3T4C1
D	1592	LEU	-	expression tag	UNP D3T4C1
D	1593	GLU	-	expression tag	UNP D3T4C1
D	1594	HIS	-	expression tag	UNP D3T4C1
D	1595	HIS	-	expression tag	UNP D3T4C1
D	1596	HIS	-	expression tag	UNP D3T4C1
D	1597	HIS	-	expression tag	UNP D3T4C1
D	1598	HIS	-	expression tag	UNP D3T4C1
D	1599	HIS	-	expression tag	UNP D3T4C1

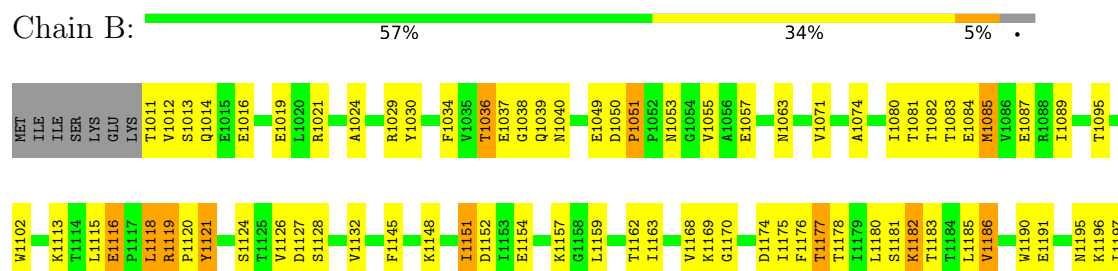
3 Residue-property plots

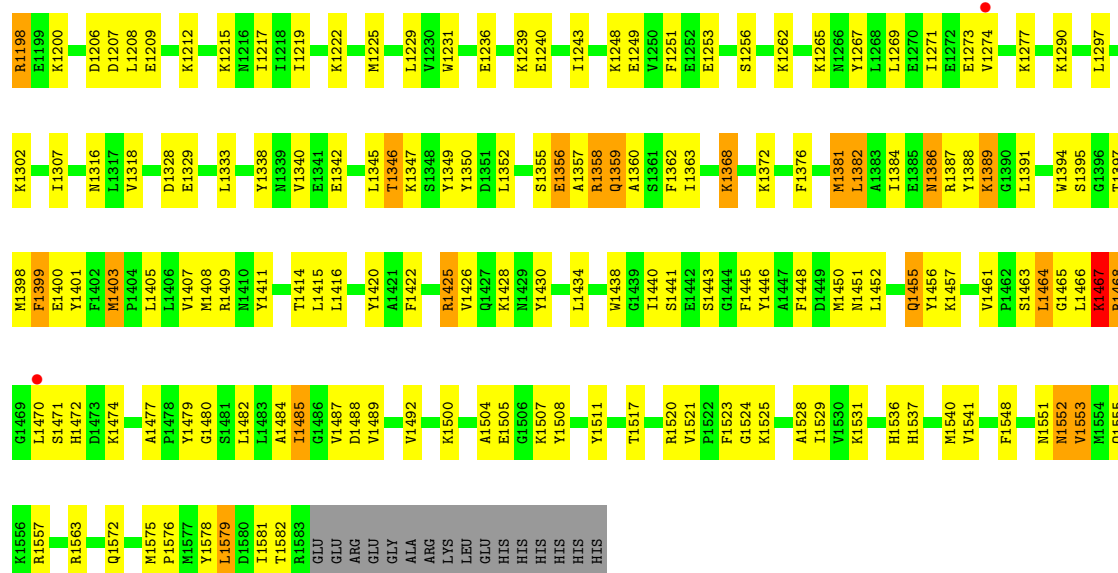
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Glycosyltransferase 36

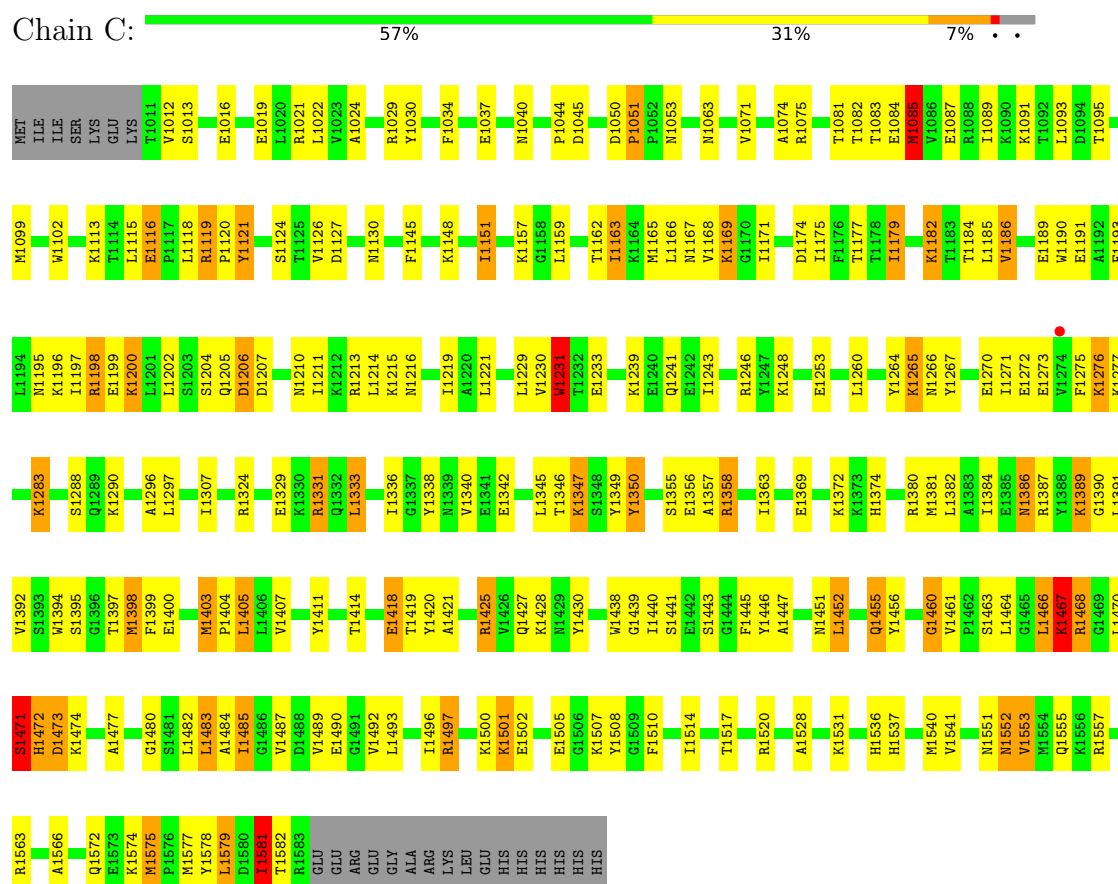


• Molecule 1: Glycosyltransferase 36





• Molecule 1: Glycosyltransferase 36



• Molecule 1: Glycosyltransferase 36



R1520	V1521	K1525	A1528	I1529	V1530	K1531	H1536	M1540	V1541	N1551	M1552	V1553	M1554	Q1555	A1556	R1557	F1558	H1559	K1565	Q1568	Q1572	M1577	Y1578	L1579	T1582	R1583	E1584	GLU	ARG	GLU	GLY	ALA	ARG	LYS	LEU	GLU	HIS	HIS	HIS	HIS											
I1436	I1440	S1441	S1442	S1443	G1444	F1445	Y1446	A1447	F1448	D1449	M1450	N1451	L1452	Q1455	Y1456	K1457	V1461	P1462	S1463	L1464	G1465	L1466	K1467	R1468	G1469	L1470	S1471	K1474	A1477	G1480	S1481	L1482	L1483	A1484	I1485	G1486	V1487	D1488	V1489	V1492	L1493	K1500	A1504	E1505	G1506	K1507	Y1508	T1517			
R1358	I1363	A1364	I1365	A1366	K1367	K1368	E1369	V1370	D1371	K1372	K1373	M1381	L1382	A1383	I1384	E1385	M1386	R1387	Y1388	K1389	G1390	L1391	W1394	S1395	M1398	F1399	M1403	P1404	L1405	L1406	V1407	M1408	R1409	N1410	Y1411	T1414	E1418	T1419	Y1420	A1421	F1422	V1426	Q1427	K1428	N1429	Y1430	A1431	K1432	E1433	L1434	G1435
L1269	E1270	I1271	E1272	E1273	V1274	K1277	E1280	E1281	L1285	L1286	K1287	S1288	Q1289	K1290	D1291	Q1295	A1296	L1297	E1298	I1307	S1312	T1321	R1324	D1328	E1329	K1330	R1331	Q1332	L1333	Y1338	N1339	V1340	E1341	E1342	E1343	K1344	L1345	T1346	K1347	S1348	Y1349	L1352	L1353	A1354	S1355	E1356	A1357				
I1179	L1180	S1181	T1184	L1185	V1186	W1190	F1193	L1194	N1195	K1196	I1197	R1198	L1201	D1206	D1207	I1211	K1212	R1213	L1214	K1215	N1216	I1217	I1218	I1219	A1220	L1221	K1222	M1225	L1229	V1230	W1231	K1157	G1158	L1159	K1160	D1161	T1162	I1163	K1164	M1165	L1166	K1169	T1172	E1173	D1174	I1175	F1176	T1177	T1178		
MET	ILE	ILE	SER	LYS	GLU	LYS	T1011	V1012	S1013	Q1014	E1015	E1016	E1019	L1020	R1021	L1022	R1119	V1023	A1024	R1029	Y1030	E1037	N1040	P1044	D1045	E1049	D1050	P1051	N1052	G1054	V1055	A1056	E1057	N1063	Y1067	V1071	I1072	G1073	A1074	R1075	T1080	T1081	T1082	T1083	E1084	M1085	V1086	E1087	R1088		

4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α , β , γ	172.72Å 172.72Å 395.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	89.60 – 3.90 89.60 – 3.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (89.60-3.90) 100.0 (89.60-3.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.32 (at 3.89Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.198 , 0.232 0.204 , 0.232	Depositor DCC
R_{free} test set	2738 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å ²)	104.0	Xtriage
Anisotropy	0.133	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 107.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	18842	wwPDB-VP
Average B, all atoms (Å ²)	120.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.09% 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](#)

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	1.01	0/4814	1.58	9/6485 (0.1%)
1	B	1.01	0/4805	1.61	15/6473 (0.2%)
1	C	1.02	0/4805	1.59	12/6473 (0.2%)
1	D	1.03	0/4814	1.59	7/6485 (0.1%)
All	All	1.02	0/19238	1.59	43/25916 (0.2%)

There are no bond length outliers.

The worst 5 of 43 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	1280	GLU	CA-C-N	7.97	130.96	120.28
1	D	1280	GLU	C-N-CA	7.97	130.96	120.28
1	A	1152	ASP	CA-CB-CG	7.31	119.91	112.60
1	D	1050	ASP	N-CA-CB	7.11	117.14	110.39
1	B	1051	PRO	N-CA-C	-6.64	102.60	110.70

There are no chirality outliers.

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	4715	0	4756	113	0
1	B	4706	0	4750	152	0
1	C	4706	0	4750	170	0
1	D	4715	0	4756	130	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	18842	0	19012	542	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.

The worst 5 of 542 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:1036:THR:HG23	1:B:1039:GLN:HG2	1.40	1.00
1:C:1050:ASP:HB3	1:C:1051:PRO:HD3	1.51	0.93
1:B:1387:ARG:HB3	1:C:1387:ARG:HB3	1.52	0.89
1:A:1050:ASP:HB3	1:A:1051:PRO:HD3	1.59	0.83
1:B:1395:SER:OG	1:B:1400:GLU:OE2	1.99	0.78

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	572/596 (96%)	478 (84%)	79 (14%)	15 (3%)	4	28
1	B	571/596 (96%)	483 (85%)	74 (13%)	14 (2%)	4	29
1	C	571/596 (96%)	468 (82%)	88 (15%)	15 (3%)	4	28
1	D	572/596 (96%)	463 (81%)	91 (16%)	18 (3%)	3	25
All	All	2286/2384 (96%)	1892 (83%)	332 (14%)	62 (3%)	4	28

5 of 62 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1356	GLU
1	A	1473	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	1356	GLU
1	B	1471	SER
1	C	1356	GLU

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	512/532 (96%)	436 (85%)	76 (15%)	3	16
1	B	511/532 (96%)	445 (87%)	66 (13%)	4	20
1	C	511/532 (96%)	436 (85%)	75 (15%)	3	16
1	D	512/532 (96%)	431 (84%)	81 (16%)	2	15
All	All	2046/2128 (96%)	1748 (85%)	298 (15%)	3	16

5 of 298 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	1128	SER
1	D	1466	LEU
1	D	1175	ILE
1	D	1341	GLU
1	B	1183	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	1559	HIS
1	D	1451	ASN
1	B	1538	GLN
1	C	1039	GLN
1	C	1046	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](#)

There are no ligands in this entry.

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	574/596 (96%)	-0.33	2 (0%) 90 78	60, 104, 157, 203	0
1	B	573/596 (96%)	-0.31	2 (0%) 90 78	65, 105, 187, 230	0
1	C	573/596 (96%)	-0.17	1 (0%) 91 81	60, 113, 212, 269	0
1	D	574/596 (96%)	-0.06	8 (1%) 73 52	68, 119, 215, 279	0
All	All	2294/2384 (96%)	-0.22	13 (0%) 85 69	60, 109, 199, 279	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	1267	TYR	3.7
1	D	1274	VAL	3.1
1	D	1470	LEU	2.8
1	D	1269	LEU	2.7
1	B	1274	VAL	2.6

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.