wwPDB X-ray Structure Validation Summary Report (i)

Jun 16, 2021 – 04:05 PM EDT

PDB ID	:	7JSR
Title	:	Crystal structure of the large glutamate dehydrogenase composed of 180 kDa subunits from Mycobacterium smegmatis
Authors	:	Lazaro, M.; Melero, R.; Huet, C.; Lopez-Alonso, J.P.; Delgado, S.; Dodu, A.; Bruch, E.M.; Abriata, L.A.; Alzari, P.M.; Valle, M.; Lisa, M.N.
Deposited on	:	2020-08-15
Resolution	:	6.27 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 following versions of software and data (see references (1)) were used in the production of this report:

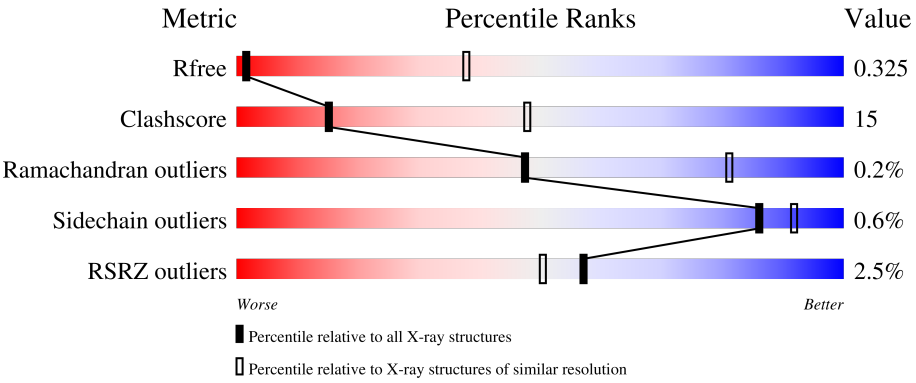
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.20
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.20

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 6.27 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}	130704	1007 (8.50-3.88)
Clashscore	141614	1056 (8.50-3.90)
Ramachandran outliers	138981	1004 (8.50-3.88)
Sidechain outliers	138945	1005 (8.70-3.84)
RSRZ outliers	127900	1018 (8.70-3.78)

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	1611	2% 63% 30% 7%
1	B	1611	3% 61% 30% 8%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 22997 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 NAD-specific glutamate dehydrogenase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	1496	Total	C	N	O	S	Se	0	0	0
			11563	7280	2058	2193	7	25			
1	B	1477	Total	C	N	O	S	Se	0	0	0
			11434	7200	2037	2165	7	25			

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-16	MSE	-	initiating methionine	UNP A0R1C2
A	-15	HIS	-	expression tag	UNP A0R1C2
A	-14	HIS	-	expression tag	UNP A0R1C2
A	-13	HIS	-	expression tag	UNP A0R1C2
A	-12	HIS	-	expression tag	UNP A0R1C2
A	-11	HIS	-	expression tag	UNP A0R1C2
A	-10	HIS	-	expression tag	UNP A0R1C2
A	-9	GLU	-	expression tag	UNP A0R1C2
A	-8	ASN	-	expression tag	UNP A0R1C2
A	-7	LEU	-	expression tag	UNP A0R1C2
A	-6	TYR	-	expression tag	UNP A0R1C2
A	-5	PHE	-	expression tag	UNP A0R1C2
A	-4	GLN	-	expression tag	UNP A0R1C2
A	-3	GLY	-	expression tag	UNP A0R1C2
A	-2	ALA	-	expression tag	UNP A0R1C2
A	-1	ALA	-	expression tag	UNP A0R1C2
A	0	SER	-	expression tag	UNP A0R1C2
B	-16	MSE	-	initiating methionine	UNP A0R1C2
B	-15	HIS	-	expression tag	UNP A0R1C2
B	-14	HIS	-	expression tag	UNP A0R1C2
B	-13	HIS	-	expression tag	UNP A0R1C2
B	-12	HIS	-	expression tag	UNP A0R1C2
B	-11	HIS	-	expression tag	UNP A0R1C2
B	-10	HIS	-	expression tag	UNP A0R1C2
B	-9	GLU	-	expression tag	UNP A0R1C2

Continued on next page...

Continued from previous page...

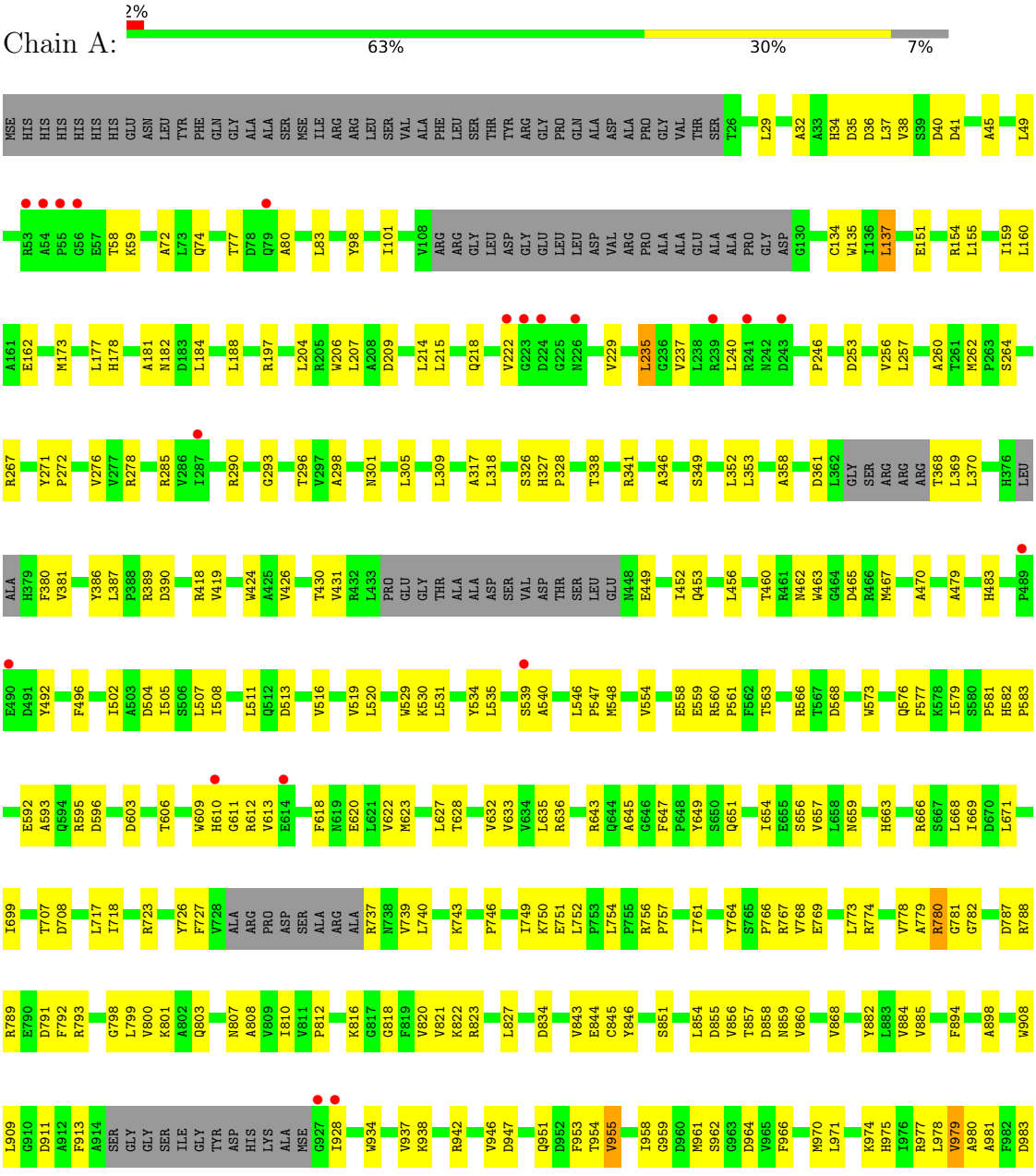
Chain	Residue	Modelled	Actual	Comment	Reference
B	-8	ASN	-	expression tag	UNP A0R1C2
B	-7	LEU	-	expression tag	UNP A0R1C2
B	-6	TYR	-	expression tag	UNP A0R1C2
B	-5	PHE	-	expression tag	UNP A0R1C2
B	-4	GLN	-	expression tag	UNP A0R1C2
B	-3	GLY	-	expression tag	UNP A0R1C2
B	-2	ALA	-	expression tag	UNP A0R1C2
B	-1	ALA	-	expression tag	UNP A0R1C2
B	0	SER	-	expression tag	UNP A0R1C2

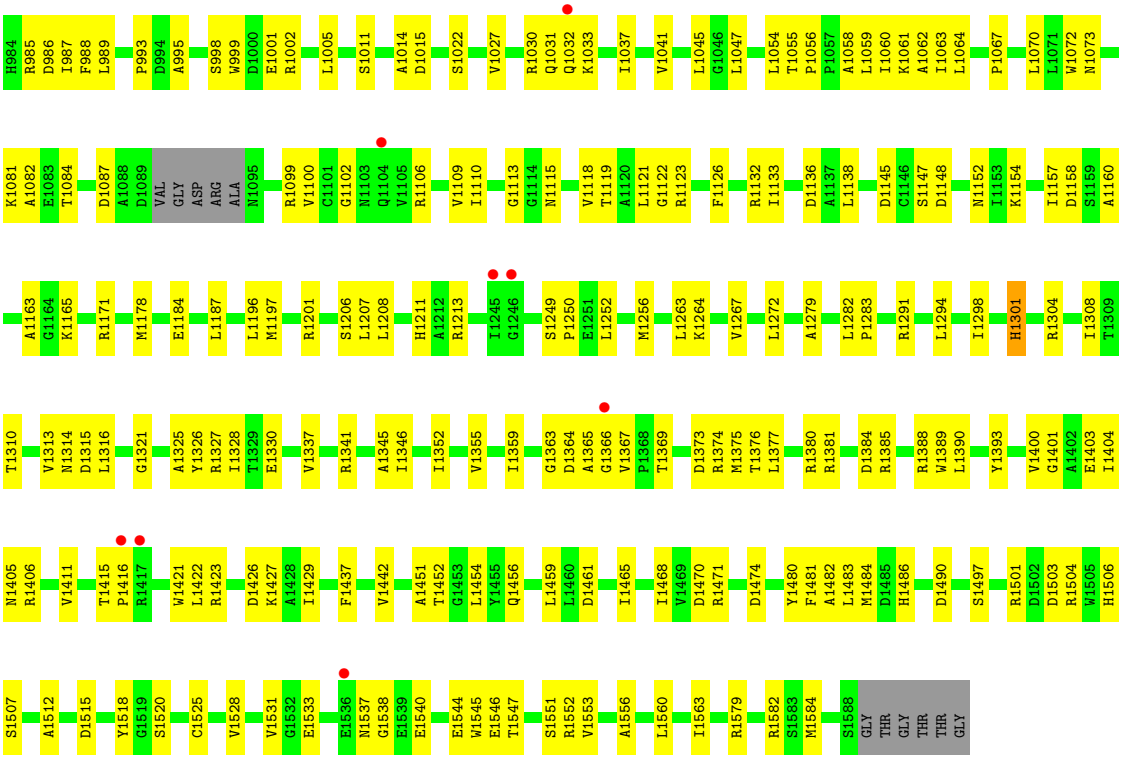
3 Residue-property plots (i)

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: NAD-specific glutamate dehydrogenase

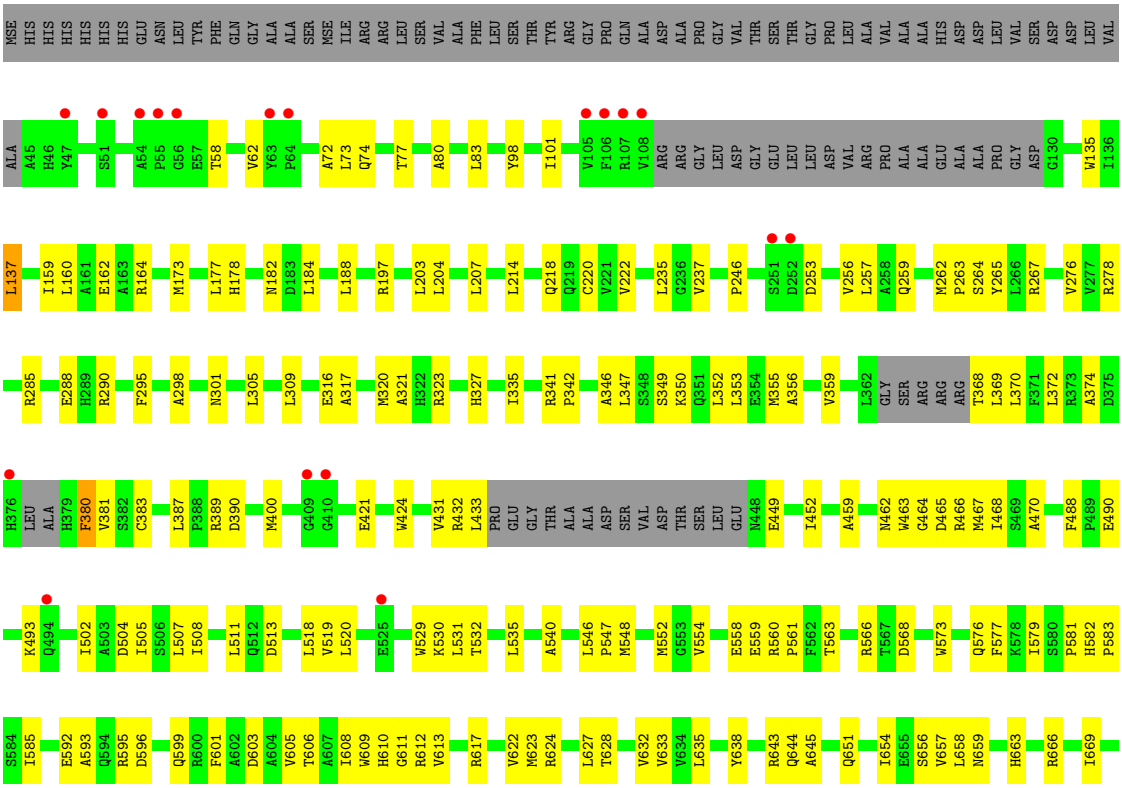
Chain A:

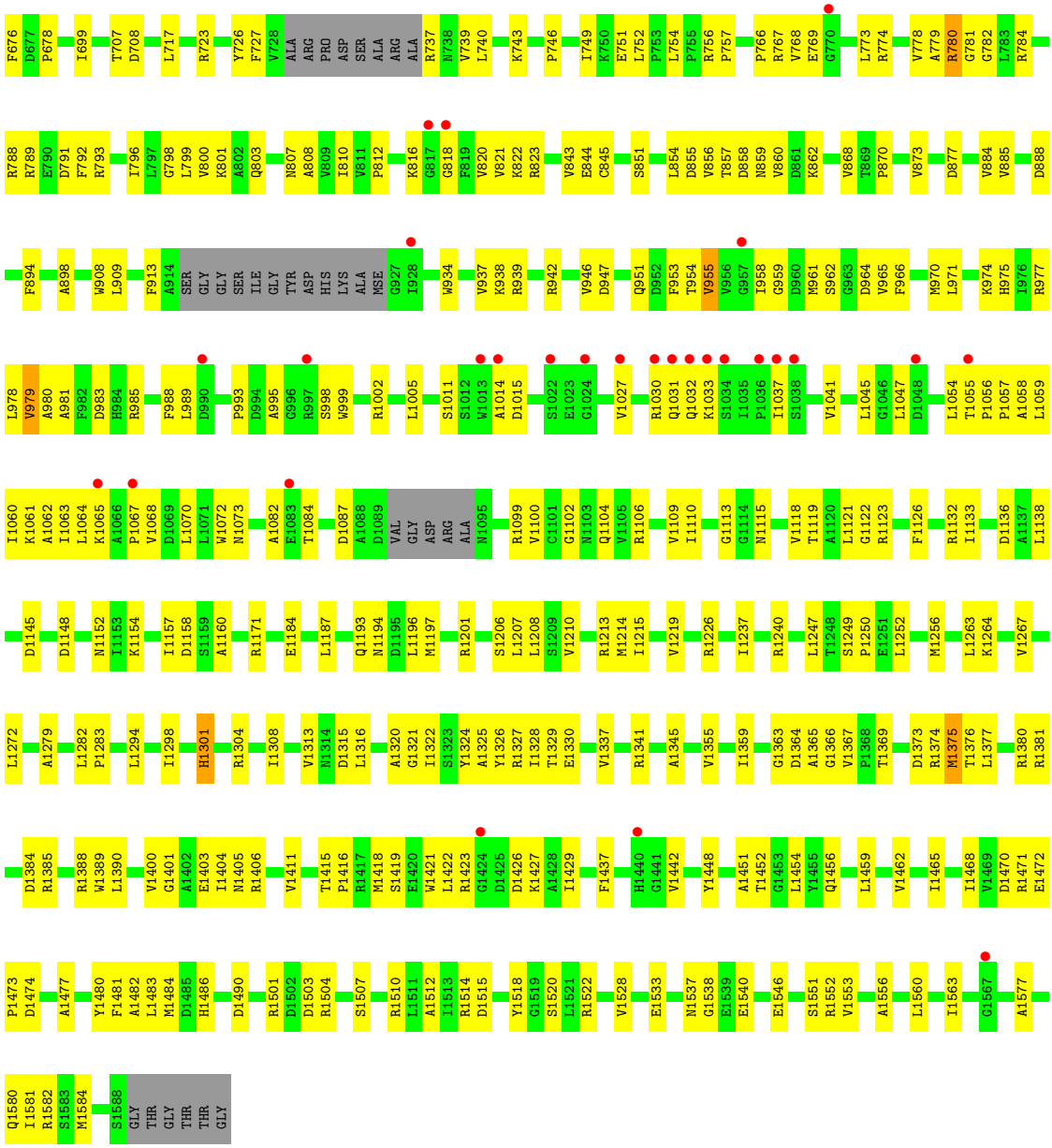




• Molecule 1: NAD-specific glutamate dehydrogenase

Chain B:




F676 D677 P678 I699 T707 D708 L717 R723 Y726 F727 V728 ALA ARG PRO ASP SER ALA ARG ALA R737 N738 V739 L740 K743 P746 I749 K750 E751 L752 P753 L754 P755 R756 P757 P766 R767 V768 E769 G770 L773 R774 V778 A779 R780 G781 G782 L783 R784
R788 R789 E790 D791 F792 R793 I796 L797 G798 L799 V800 K801 A802 Q803 N807 A808 V809 I810 V811 P812 K816 G817 G818 F819 V820 V821 K822 R823 V843 E844 C845 S851 L854 D855 V856 T857 D858 N859 V860 D861 K862 V868 T869 P870 V873 D877 V884 V885 D888
F894 A898 W908 L909 F913 A914 SER GLY GLY SER ILE GLY TYR ASP HIS LYS ALA MSE G927 I928 W934 V937 K938 R939 R942 V946 D947 Q951 D952 F953 T954 V955 V956 G957 I958 G959 D960 M961 S962 G963 D964 V965 F966 M970 L971 K974 H975 I976 R977
L978 V979 A980 A981 F982 D983 H984 R985 F988 L989 D990 P993 D994 A995 G996 R997 S998 W999 R1002 L1005 S1011 S1012 W1013 A1014 D1015 S1022 E1023 G1024 V1027 R1030 Q1031 Q1032 K1033 S1034 I1035 P1036 I1037 S1038 V1041 L1045 G1046 L1047 D1048 L1054 T1055 P1056 P1057 A1058 L1059
I1060 K1061 A1062 I1063 L1064 K1065 A1066 P1067 V1068 D1069 L1070 L1071 W1072 N1073 A1082 E1083 T1084 D1087 A1088 D1089 VAL GLY ASP ARG ALA N1095 R1099 V1100 C1101 G1102 N1103 Q1104 V1105 R1106 V1109 I1110 G1113 G1114 N1115 V1118 T1119 A1120 L1121 G1122 R1123 F1126 R1132 I1133 D1136 A1137 L1138
D1145 D1148 N1152 I1153 K1154 I1157 D1158 S1159 A1160 R1171 E1184 L1187 Q1193 N1194 D1196 L1196 M1197 R1201 S1206 L1207 L1208 S1209 V1210 R1213 M1214 I1215 V1219 R1226 I1237 R1240 L1247 T1248 S1249 P1250 E1251 L1252 M1256 L1263 K1264 V1267
L1272 A1279 L1282 P1283 L1294 I1298 H1301 R1304 I1308 V1313 N1314 D1315 L1316 A1320 G1321 I1322 S1323 Y1324 A1325 Y1326 R1327 I1328 T1329 E1330 V1337 R1341 A1345 V1355 I1359 G1363 D1364 A1365 G1366 V1367 P1368 T1369 D1373 R1374 M1375 T1376 L1377 R1380 R1381
D1384 R1385 R1388 W1389 L1390 V1400 G1401 A1402 E1403 I1404 N1405 R1406 V1411 T1415 P1416 R1417 M1418 S1419 E1420 W1421 L1422 R1423 G1424 D1425 D1426 K1427 A1428 I1429 F1437 H1440 G1441 V1442 Y1448 A1451 T1452 G1453 L1454 Y1455 Q1456 L1459 V1462 I1465 I1468 V1469 D1470 R1471 E1472
P1473 D1474 A1477 Y1480 F1481 A1482 L1483 M1484 D1485 H1486 D1490 R1501 D1502 D1503 R1504 S1507 R1510 L1511 A1512 I1513 R1514 D1515 Y1518 G1519 S1520 L1521 R1522 V1528 E1533 N1537 G1538 E1539 E1540 E1546 S1551 R1552 V1553 A1556 L1560 I1563 G1567 A1577
Q1580 I1581 R1582 S1583 M1584 S1588 GLY THR GLY THR THR GLY

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	151.56Å 253.54Å 399.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	24.98 – 6.27 24.98 – 6.27	Depositor EDS
% Data completeness (in resolution range)	98.4 (24.98-6.27) 86.4 (24.98-6.27)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.67 (at 6.51Å)	Xtriage
Refinement program	PHENIX 1.14rc3_3199	Depositor
R, R_{free}	0.277 , 0.325 0.278 , 0.325	Depositor DCC
R_{free} test set	860 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å^2)	329.5	Xtriage
Anisotropy	0.347	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.20 , 299.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.44$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	0.046 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.058 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o,F_c correlation	0.86	EDS
Total number of atoms	22997	wwPDB-VP
Average B, all atoms (Å^2)	472.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.07% 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

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	0.24	0/11753	0.45	0/15937
1	B	0.24	0/11622	0.45	0/15755
All	All	0.24	0/23375	0.45	0/31692

There are no bond length outliers.

There are no bond angle outliers.

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	11563	0	11503	351	0
1	B	11434	0	11383	349	0
All	All	22997	0	22886	688	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 15.

The worst 5 of 688 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:162:GLU:HA	1:A:237:VAL:HG21	1.56	0.85
1:A:1546:GLU:HG3	1:A:1553:VAL:HG11	1.60	0.84

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:774:ARG:HD2	1:A:778:VAL:HB	1.60	0.83
1:B:162:GLU:HA	1:B:237:VAL:HG21	1.63	0.80
1:A:370:LEU:HD22	1:A:452:ILE:HD13	1.65	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	1480/1611 (92%)	1377 (93%)	100 (7%)	3 (0%)	47	81
1	B	1461/1611 (91%)	1359 (93%)	99 (7%)	3 (0%)	47	81
All	All	2941/3222 (91%)	2736 (93%)	199 (7%)	6 (0%)	47	81

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1366	GLY
1	B	1366	GLY
1	A	780	ARG
1	A	1470	ASP
1	B	780	ARG

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	1210/1266 (96%)	1203 (99%)	7 (1%)	86 92
1	B	1197/1266 (94%)	1189 (99%)	8 (1%)	84 90
All	All	2407/2532 (95%)	2392 (99%)	15 (1%)	86 92

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	137	LEU
1	B	1301	HIS
1	B	235	LEU
1	B	1375	MSE
1	B	979	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	1193	GLN
1	B	1456	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 monosaccharides 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

6.1 Protein, DNA and RNA chains

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, 95^{th} 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	1471/1611 (91%)	-0.15	28 (1%)	66	59	296, 461, 569, 656	0
1	B	1452/1611 (90%)	-0.06	46 (3%)	47	41	299, 472, 603, 685	0
All	All	2923/3222 (90%)	-0.11	74 (2%)	57	50	296, 466, 588, 685	0

The worst 5 of 74 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	55	PRO	7.8
1	B	54	ALA	7.7
1	B	107	ARG	7.7
1	B	106	PHE	5.9
1	B	108	VAL	5.5

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

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.