wwPDB X-ray Structure Validation Summary Report (i)

Jun 17, 2024 – 01:31 PM EDT

PDB ID	:	3LXU
Title	:	Crystal Structure of Tripeptidyl Peptidase 2 (TPP II)
Authors	:	Chuang, C.K.
Deposited on	:	2010-02-25
Resolution	:	3.14 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:

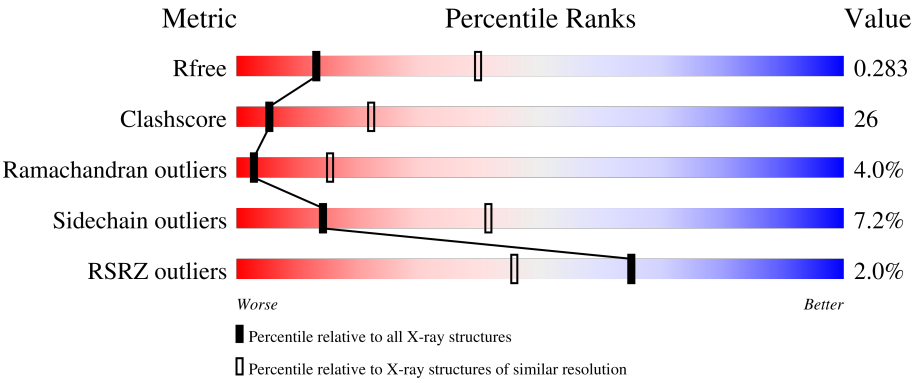
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.14 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	1626 (3.18-3.10)
Clashscore	141614	1735 (3.18-3.10)
Ramachandran outliers	138981	1677 (3.18-3.10)
Sidechain outliers	138945	1677 (3.18-3.10)
RSRZ outliers	127900	1588 (3.18-3.10)

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	X	1354	2% 49% 35% 5% 10%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 9464 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 Tripeptidyl-peptidase 2.

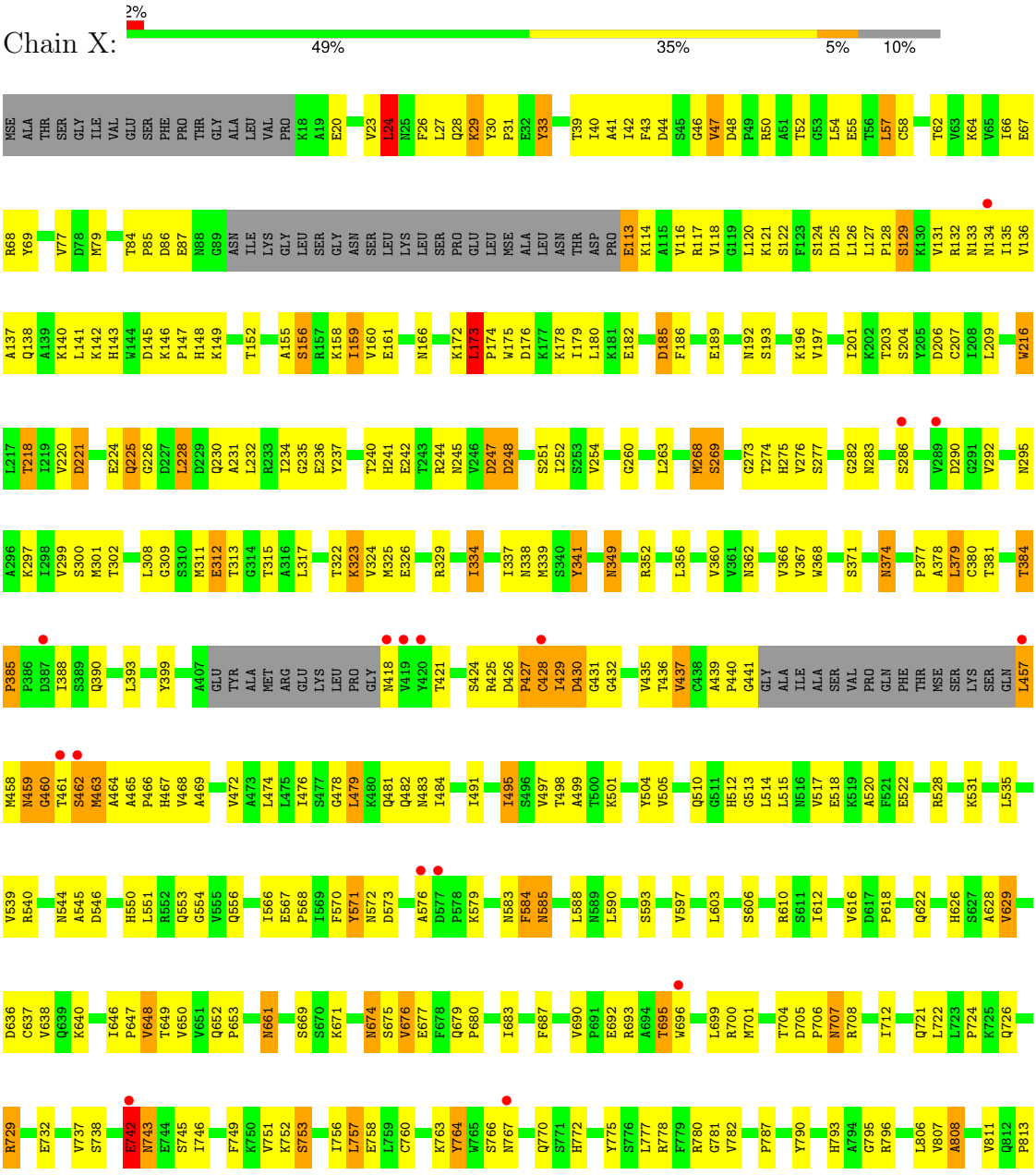
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	X	1217	9464	5958	1648	1816	17	25	0	0	0

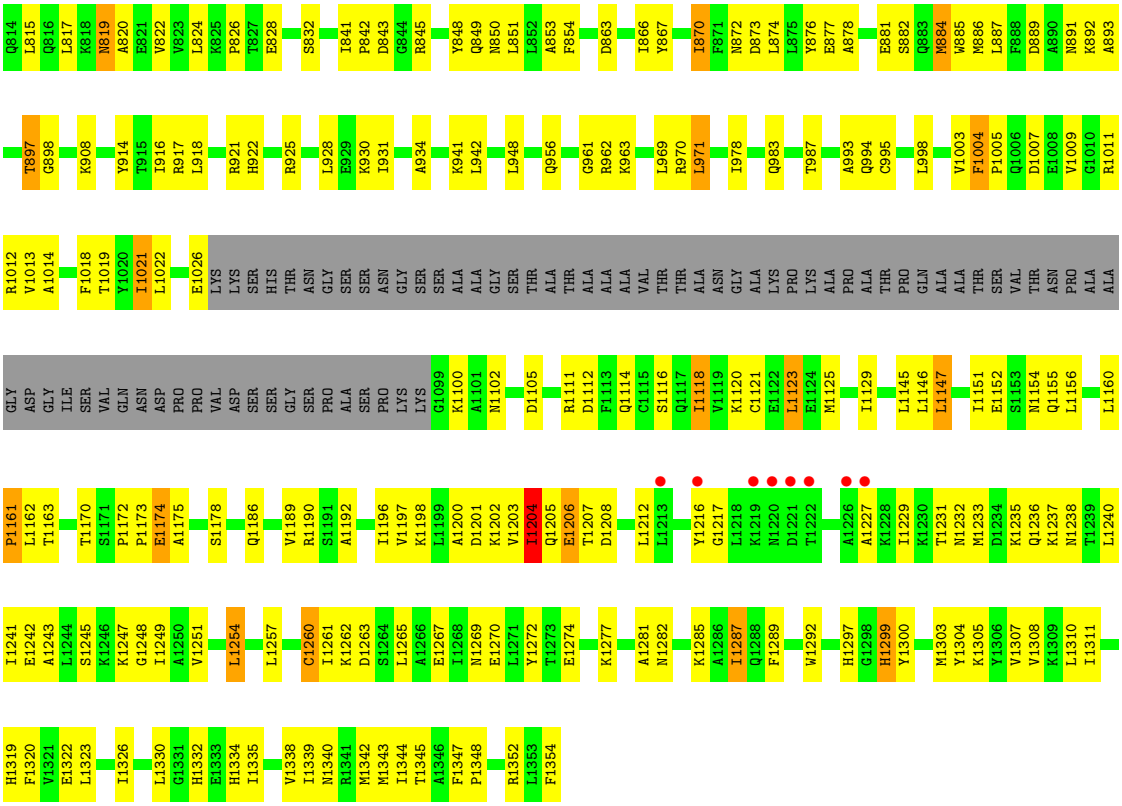
3 Residue-property plots (i)

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: Tripeptidyl-peptidase 2

Chain X:




Q814 L815 Q816 L817 K818 N819 A820 E821 V822 V823 L824 K825 P826 T827 E828 S832 I841 P842 D843 G844 R845 Y848 Q849 N850 L851 L852 A853 F854 D863 I866 Y867 I870 F871 N872 D873 L874 L875 Y876 E877 A878 E881 S882 Q883 M884 W885 M886 L887 F888 D889 A890 N891 K892 A893
T897 G898 K908 Y914 T915 I916 R917 L918 R921 H922 R925 L928 E929 K930 I931 A934 K941 L942 L948 Q956 G961 R962 K963 L969 R970 L971 I978 Q983 T987 A993 Q994 C995 L998 V1003 F1004 P1005 Q1006 D1007 E1008 V1009 G1010 R1011
R1012 V1013 A1014 F1018 T1019 Y1020 I1021 L1022 E1026 LYS LYS SER HIS THR ASN GLY SER SER ASN GLY SER SER ALA ALA GLY SER THR ALA THR ALA ALA ALA VAL THR THR ALA ASN GLY ALA LYS PRO LYS ALA PRO ALA THR PRO GLN ALA ALA THR SER VAL THR ASN PRO ALA ALA
GLY ASP GLY ILE SER VAL GLN ASN ASP PRO PRO VAL ASP SER SER GLY SER PRO ALA SER PRO LYS LYS G1099 K1100 A1101 N1102 D1105 R1111 D1112 F1113 Q1114 C1115 S1116 Q1117 I1118 V1119 K1120 C1121 E1122 L1123 E1124 M1125 I1129 L1145 L1146 L1147 I1151 E1152 S1153 N1154 Q1155 L1156 L1160
P1161 L1162 T1163 T1170 S1171 P1172 P1173 E1174 A1175 S1178 Q1186 V1189 R1190 S1191 A1192 I1196 V1197 K1198 L1199 A1200 D1201 K1202 V1203 I1204 Q1205 E1206 T1207 D1208 L1212 L1213 Y1216 G1217 L1218 K1219 N1220 D1221 T1222 A1226 A1227 K1228 I1229 K1230 T1231 N1232 M1233 D1234 K1235 Q1236 K1237 N1238 T1239 L1240
I1241 E1242 A1243 L1244 S1245 K1246 K1247 G1248 I1249 A1250 V1251 L1254 L1257 C1260 I1261 K1262 D1263 S1264 L1265 A1266 E1267 I1268 N1269 E1270 L1271 Y1272 T1273 E1274 K1277 A1281 N1282 K1285 A1286 I1287 Q1288 F1289 W1292 H1297 G1298 H1299 Y1300 M1303 Y1304 K1305 Y1306 V1307 V1308 K1309 L1310 I1311
H1319 F1320 V1321 E1322 L1323 I1326 L1330 G1331 H1332 E1333 H1334 I1335 V1338 I1339 N1340 R1341 M1342 M1343 I1344 T1345 A1346 F1347 P1348 R1352 L1353 F1354

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	126.37Å 126.37Å 213.98Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.14 49.97 – 3.14	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-3.14) 99.2 (49.97-3.14)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.77 (at 3.12Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.241 , 0.293 0.240 , 0.283	Depositor DCC
R_{free} test set	1548 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	88.1	Xtriage
Anisotropy	0.303	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 44.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	9464	wwPDB-VP
Average B, all atoms (Å^2)	92.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.85% 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	X	0.22	0/9621	0.41	0/12999

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	X	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	X	384	THR	Peptide
1	X	462	SER	Peptide

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	X	9464	0	9342	492	0
All	All	9464	0	9342	492	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 26.

The worst 5 of 492 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:X:429:ILE:CD1	1:X:722:LEU:HD13	1.87	1.03
1:X:457:LEU:HD13	1:X:458:MSE:H	1.30	0.97
1:X:729:ARG:HH11	1:X:729:ARG:HG2	1.31	0.94
1:X:793:HIS:HD2	1:X:795:GLY:H	1.10	0.90
1:X:116:VAL:HG22	1:X:117:ARG:H	1.35	0.88

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	X	1207/1354 (89%)	1001 (83%)	158 (13%)	48 (4%)	3 15

5 of 48 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	X	129	SER
1	X	225	GLN
1	X	385	PRO
1	X	429	ILE
1	X	571	TYR

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	X	1027/1116 (92%)	953 (93%)	74 (7%)	14 41

5 of 74 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	X	851	LEU
1	X	1265	LEU
1	X	897	THR
1	X	1118	ILE
1	X	349	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 47 such sidechains are listed below:

Mol	Chain	Res	Type
1	X	812	GLN
1	X	1117	GLN
1	X	849	GLN
1	X	956	GLN
1	X	1142	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	X	1192/1354 (88%)	0.12	24 (2%) 65 46	53, 86, 137, 179	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	X	462	SER	5.2
1	X	286	SER	4.1
1	X	576	ALA	4.0
1	X	577	ASP	3.9
1	X	1227	ALA	3.4

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.