



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 10:27 AM GMT

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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
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) : trunk26865

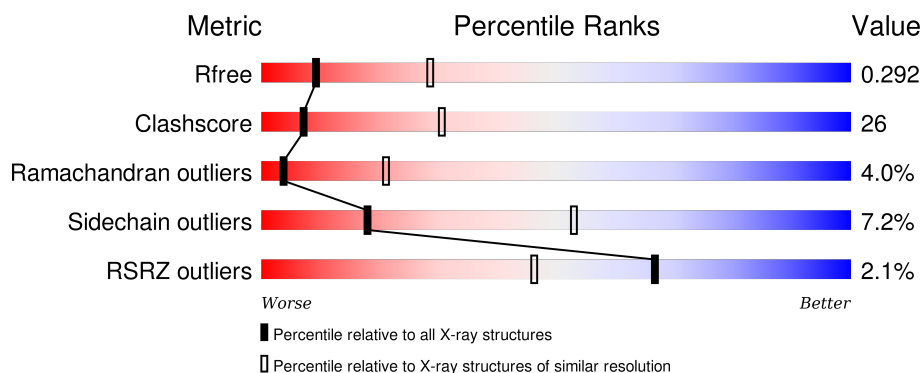
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.14 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))
$R_{free}$	91344	1095 (3.18-3.10)
Clashscore	102246	1202 (3.18-3.10)
Ramachandran outliers	100387	1162 (3.18-3.10)
Sidechain outliers	100360	1162 (3.18-3.10)
RSRZ outliers	91569	1097 (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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	X	1354	

## 2 Entry composition

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.

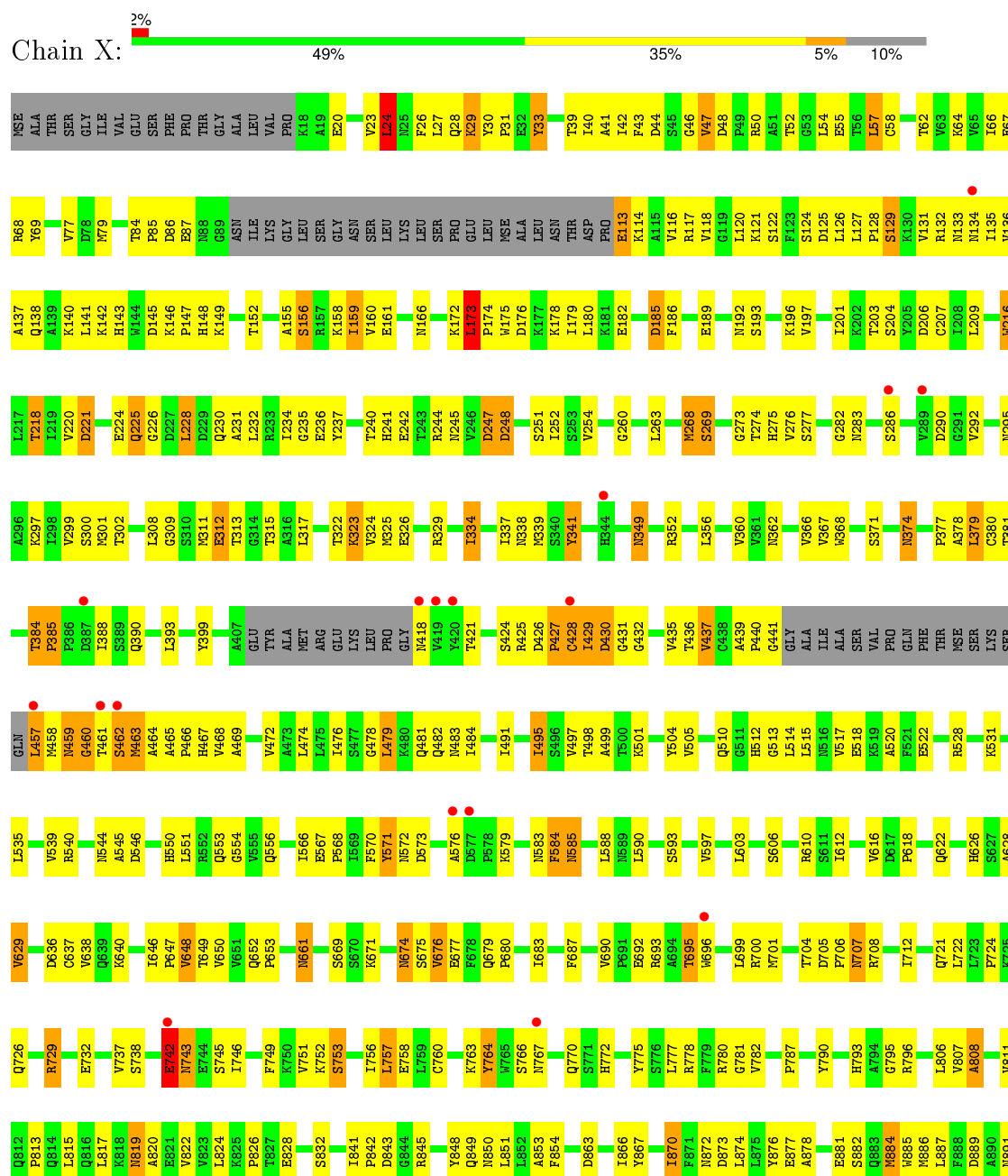
- Molecule 1 is a protein called Tripeptidyl-peptidase 2.

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

### 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: Tripeptidyl-peptidase 2



I1311	T1239	L1160	ALA	G1010	K892
H1319	L1240	P1161	ALA	R1011	A893
F1320	I1241	L1162	GLY	R1012	
E1321	E1242	T1163	ASP	V1013	T897
V1322	A1243	T1163	GLY	A1014	G898
L1323	L1244	T1170	ILE	F1018	K908
	S1245	S1171	SER	T1019	
	K1246	P1172	VAL	Y1020	Y914
	K1247	P1173	GLN	I1021	T915
	G1248	E1174	ASN	L1022	I916
	I1249	A1175	ASP	E1026	R917
	A1250	S1178	PRO	LYS	L918
	V1251	Q1186	VAL	LYS	R921
	L1254	V1189	ASP	LYS	H922
	L1257	R1190	SER	HIS	R925
	L1260	S1191	GLY	THR	
	I1261	A1192	SER	ASN	
	K1262	I1196	PRO	GLY	L928
	D1263	V1197	ALA	SER	E929
	S1264	K1198	SER	K930	I931
	L1265	L1199	PRO	ASN	
	A1266	A1200	LYS	GLY	A834
	E1267	D1201	SER	SER	
	I1268	K1202	ALA	ALA	K941
	N1269	V1203	A1101	L942	
	E1270	I1204	GLY	THR	L948
	L1271	E1206	SER	ALA	Q956
	Y1272	D1208	D1105	THR	
	T1273	L1212	R1111	ALA	G961
	E1274	L1213	D1112	ALA	R962
	K1277	Y1216	F1113	ALA	K963
	A1281	G1217	Q1114	VAL	
	N1282	L1218	S1116	THR	L969
	K1285	G1219	Q1117	THR	R970
	A1286	K1220	I1118	ALA	L971
	I1287	N1221	V1119	ASN	
	Q1288	T1222	K1120	GLY	I978
	F1289		C1121	ALA	
	W1292		E1122	LYS	Q983
	H1297		L1123	PRO	
	G1298		E1124	LYS	T987
	H1299		M1125	ALA	
	Y1300		I1129	ALA	A993
	M1303		A1227	PRO	Q994
	Y1304		K1228	THR	C995
	K1305		I1229	PRO	
	Y1306		L1145	GLN	L998
	V1307		L1146	ALA	
	V1308		L1147	ALA	V1003
	K1309		I1151	THR	F1004
	L1310		E1152	SER	P1005
			S1153	VAL	Q1006
			N1154	THR	D1007
			Q1155	ASN	E1008
			L1156	PRO	V1009

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	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
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.77 (at 3.12Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.241 , 0.293 0.246 , 0.292	Depositor DCC
$R_{free}$ test set	1548 reflections (5.31%)	DCC
Wilson B-factor (Å <sup>2</sup> )	88.1	Xtriage
Anisotropy	0.303	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 44.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 30700 reflections	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	9464	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	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.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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:HG2	1:X:729:ARG:HH11	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%)	4	21

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%)	18 53

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

Mol	Chain	Res	Type
1	X	535	LEU
1	X	649	THR
1	X	1204	ILE
1	X	544	ASN
1	X	588	LEU

Some 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	679	GLN
1	X	788	ASN
1	X	1288	GLN
1	X	743	ASN
1	X	793	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	X	1192/1354 (88%)	0.15	25 (2%) 67 46	53, 86, 137, 179	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	X	462	SER	5.3
1	X	286	SER	4.2
1	X	577	ASP	4.1
1	X	576	ALA	4.1
1	X	1227	ALA	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](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.