W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 3, 2020 – 02:30 pm BST

PDB ID	:	1DPP
Title	:	DIPEPTIDE BINDING PROTEIN COMPLEX WITH GLYCYL-L-LEUCINE
Authors	:	Dunten, P.; Mowbray, S.L.
Deposited on	:	1995-08-11
Resolution	:	3.20 Å(reported)

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

|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 3.20 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.
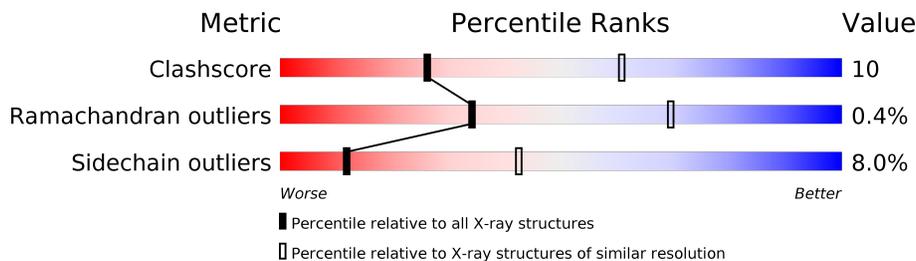
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)

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 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%

Note EDS was not executed.

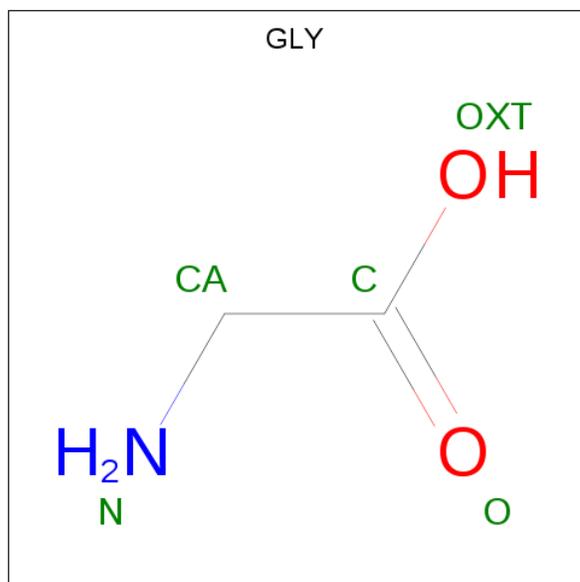| Mol	Chain	Length	Quality of chain
1	A	507	77% 19% •
1	C	507	77% 19% •
1	E	507	76% 20% •
1	G	507	77% 19% •

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 16244 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 DIPEPTIDE BINDING PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	507	Total	C	N	O	S	0	0	0
			4048	2584	676	769	19			
1	C	507	Total	C	N	O	S	0	0	0
			4048	2584	676	769	19			
1	E	507	Total	C	N	O	S	0	0	0
			4048	2584	676	769	19			
1	G	507	Total	C	N	O	S	0	0	0
			4048	2584	676	769	19			

- Molecule 2 is GLYCINE (three-letter code: GLY) (formula: $C_2H_5NO_2$).

GLY

OXT

OH

CA   C

H_2N

N   O

O

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			4	2	1	1		
2	C	1	Total	C	N	O	0	0
			4	2	1	1		

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	E	1	Total	C	N	O	0	0
			4	2	1	1		
2	G	1	Total	C	N	O	0	0
			4	2	1	1		

• Molecule 3 is LEUCINE (three-letter code: LEU) (formula: $C_6H_{13}NO_2$).
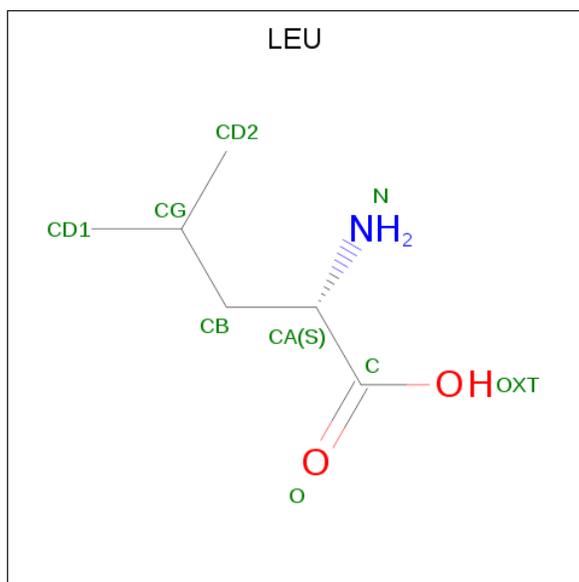
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			9	6	1	2		
3	C	1	Total	C	N	O	0	0
			9	6	1	2		
3	E	1	Total	C	N	O	0	0
			9	6	1	2		
3	G	1	Total	C	N	O	0	0
			9	6	1	2		

# 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. 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. 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.

Note EDS was not executed.
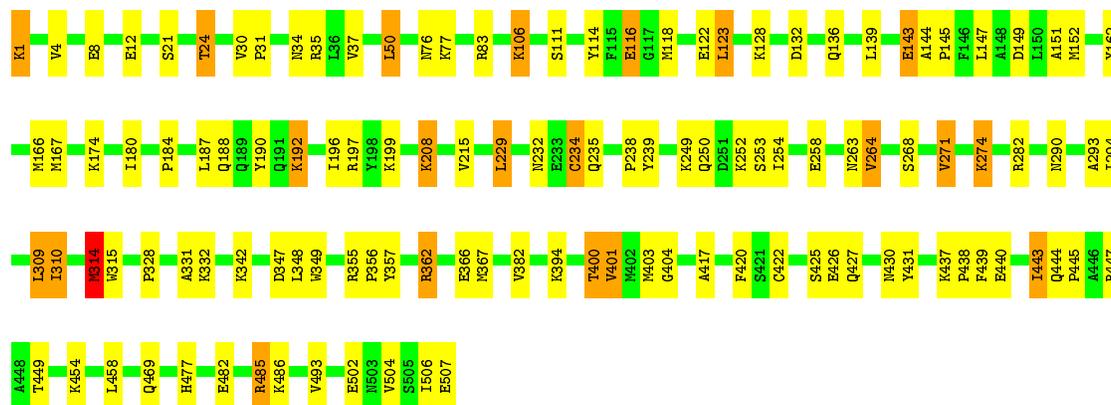
• Molecule 1: DIPEPTIDE BINDING PROTEIN
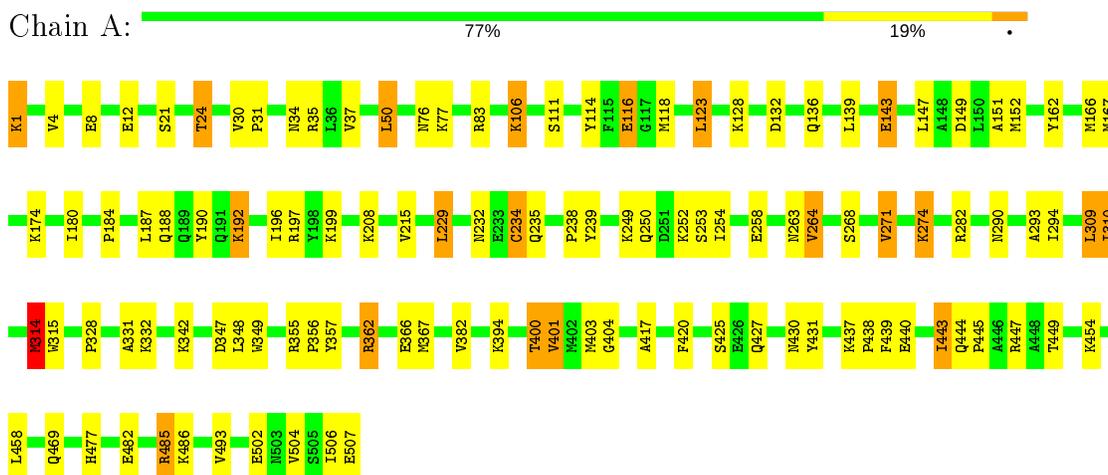
Chain A:
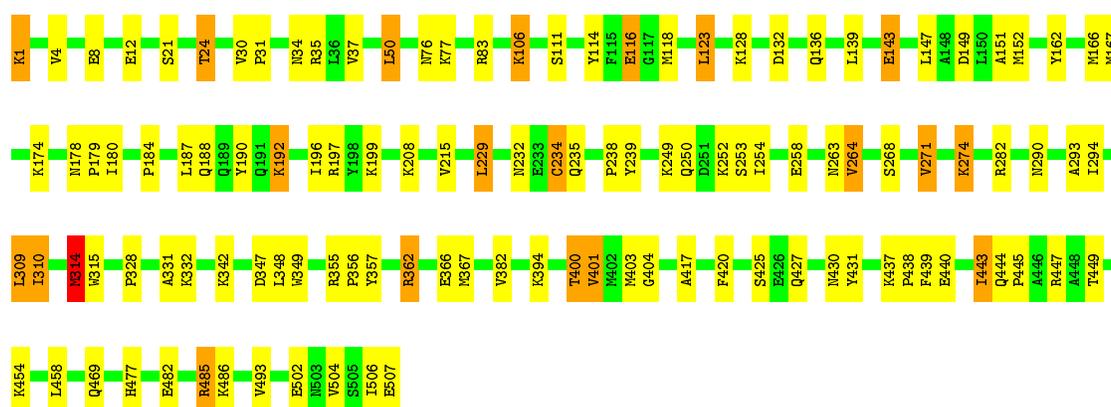

• Molecule 1: DIPEPTIDE BINDING PROTEIN
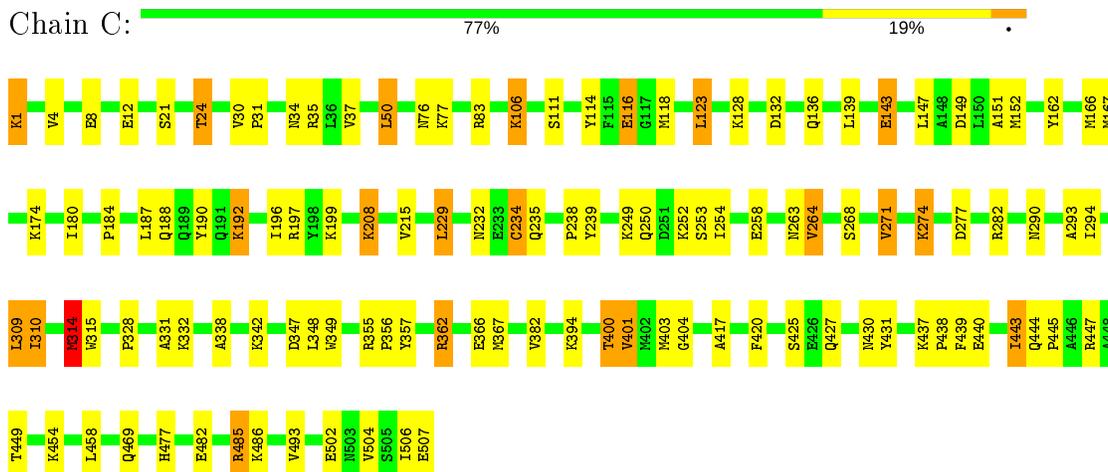
Chain C:

• Molecule 1: DIPEPTIDE BINDING PROTEIN

Chain E:

• Molecule 1: DIPEPTIDE BINDING PROTEIN

Chain G:

4   Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α, β, γ	182.57Å 182.57Å 211.88Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	25.00 – 3.20	Depositor
% Data completeness (in resolution range)	87.0 (25.00-3.20)	Depositor
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR 3.0	Depositor
R, R_{free}	0.223 , 0.240	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	16244	wwPDB-VP
Average B, all atoms (Å^2)	21.0	wwPDB-VP

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	0.42	0/4153	0.66	1/5630 (0.0%)
1	C	0.42	0/4153	0.66	1/5630 (0.0%)
1	E	0.42	0/4153	0.66	1/5630 (0.0%)
1	G	0.42	0/4153	0.66	1/5630 (0.0%)
All	All	0.42	0/16612	0.66	4/22520 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	314	MET	CG-SD-CE	6.10	109.96	100.20
1	G	314	MET	CG-SD-CE	6.10	109.96	100.20
1	A	314	MET	CG-SD-CE	6.07	109.92	100.20
1	E	314	MET	CG-SD-CE	6.07	109.91	100.20

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	4048	0	3946	81	0
1	C	4048	0	3946	94	0
1	E	4048	0	3946	94	0
1	G	4048	0	3946	81	0
2	A	4	0	2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	4	0	2	0	0
2	E	4	0	2	0	0
2	G	4	0	2	0	0
3	A	9	0	11	0	0
3	C	9	0	11	0	0
3	E	9	0	11	0	0
3	G	9	0	11	0	0
All	All	16244	0	15836	333	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 10.

The worst 5 of 333 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:C:274:LYS:HD2	1:E:426:GLU:CB	1.28	1.59
1:C:274:LYS:CD	1:E:426:GLU:HB2	1.63	1.28
1:C:274:LYS:CD	1:E:426:GLU:CB	2.23	1.14
1:C:274:LYS:HD2	1:E:426:GLU:CG	1.82	1.08
1:A:1:LYS:HE2	1:A:507:GLU:HA	1.48	0.95

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	505/507 (100%)	476 (94%)	27 (5%)	2 (0%)	34	69
1	C	505/507 (100%)	476 (94%)	27 (5%)	2 (0%)	34	69
1	E	505/507 (100%)	476 (94%)	27 (5%)	2 (0%)	34	69
1	G	505/507 (100%)	476 (94%)	27 (5%)	2 (0%)	34	69

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	2020/2028 (100%)	1904 (94%)	108 (5%)	8 (0%)	34 69

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	132	ASP
1	C	132	ASP
1	E	132	ASP
1	G	132	ASP
1	A	271	VAL

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	434/434 (100%)	399 (92%)	35 (8%)	11 42
1	C	434/434 (100%)	399 (92%)	35 (8%)	11 42
1	E	434/434 (100%)	399 (92%)	35 (8%)	11 42
1	G	434/434 (100%)	399 (92%)	35 (8%)	11 42
All	All	1736/1736 (100%)	1596 (92%)	140 (8%)	12 42

5 of 140 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	400	THR
1	E	143	GLU
1	G	314	MET
1	C	443	ILE
1	E	24	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	477	HIS
1	E	136	GLN
1	G	178	ASN
1	C	232	ASN
1	C	467	HIS

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)

8 ligands are modelled in this entry.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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)

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains (i)

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.