



Full wwPDB X-ray Structure Validation Report ⓘ

Jun 14, 2020 – 10:40 am BST

PDB ID : 2D3H
Title : Crystal structures of collagen model peptides (Pro-Pro-Gly)4-Hyp-Hyp-Gly-(Pro-Pro-Gly)4
Authors : Wu, G.; Noguchi, K.; Okuyama, K.; Mizuno, K.; Bachinger, H.P.
Deposited on : 2005-09-28
Resolution : 1.22 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
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

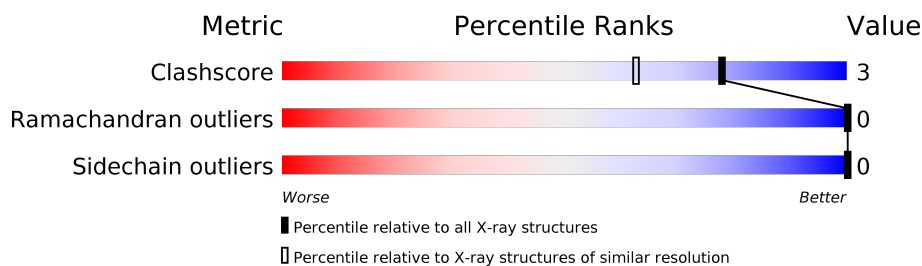
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.22 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	1294 (1.24-1.20)
Ramachandran outliers	138981	1251 (1.24-1.20)
Sidechain outliers	138945	1250 (1.24-1.20)

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 $\leq 5\%$.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	27	 85% • 11%
1	B	27	 89% 11%
1	C	27	 89% • 7%
1	D	27	 85% 7% 7%
1	E	27	 89% • 7%
1	F	27	 85% • 11%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 1158 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 COLLAGEN MODEL PEPTIDES (PRO-PRO-GLY)4-HYP-HYP-GLY-(PRO-PRO-GLY)4.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	24	Total	C	N	O	0	0	0
			146	96	24	26			
1	B	24	Total	C	N	O	0	0	0
			146	96	24	26			
1	C	25	Total	C	N	O	0	0	0
			153	101	25	27			
1	D	25	Total	C	N	O	0	0	0
			150	98	25	27			
1	E	25	Total	C	N	O	0	0	0
			153	101	25	27			
1	F	24	Total	C	N	O	0	0	0
			147	96	24	27			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	50	Total	O	0	0
			50	50		
2	B	45	Total	O	0	0
			45	45		
2	C	34	Total	O	0	0
			34	34		
2	D	38	Total	O	0	0
			38	38		
2	E	48	Total	O	0	0
			48	48		
2	F	48	Total	O	0	0
			48	48		

i

Note EDS was not executed.

-
- A diagram showing a protein segment. A green bar labeled 'P1' is connected to a yellow bar labeled 'G24'. To the right of 'G24' is a grey bar containing the text 'PRO', 'PRO', and 'GLY' stacked vertically.

- | | | | | |
|----|-----|-----|-----|-----|
| P1 | G24 | PRO | PRO | GLY |
|----|-----|-----|-----|-----|


-
- P1
P25
PRO
GLY

-
- PRO
PRO
G3
P22
G27

- 



- Molecule 1: COLLAGEN MODEL PEPTIDES (PRO-PRO-GLY)₄-HYP-HYP-GLY-(PRO-PRO-GLY)₄

Chain F:  85% • 11%



4 Data and refinement statistics

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	25.99Å 26.67Å 79.84Å 90.00° 90.03° 90.00°	Depositor
Resolution (Å)	8.00 – 1.22	Depositor
% Data completeness (in resolution range)	(Not available) (8.00-1.22)	Depositor
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	SHELXL-97	Depositor
R, R_{free}	0.193 , 0.255	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1158	wwPDB-VP
Average B, all atoms (Å ²)	18.0	wwPDB-VP

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: HYP

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	0.53	0/142	0.86	0/201
1	B	0.55	0/142	0.90	0/201
1	C	0.56	0/150	0.89	0/213
1	D	0.61	0/146	0.76	0/207
1	E	0.50	0/150	0.80	0/213
1	F	0.52	0/143	0.93	0/201
All	All	0.55	0/873	0.86	0/1236

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 [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	146	0	138	1	0
1	B	146	0	138	0	0
1	C	153	0	145	1	0
1	D	150	0	138	2	0
1	E	153	0	143	2	0
1	F	147	0	136	2	0
2	A	50	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	45	0	0	0	0
2	C	34	0	0	0	0
2	D	38	0	0	0	0
2	E	48	0	0	0	0
2	F	48	0	0	2	0
All	All	1158	0	838	5	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 3.

All (5) 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:E:2:PRO:HA	1:F:4:PRO:O	1.87	0.74
1:D:27:GLY:HA3	2:F:1223:HOH:O	2.01	0.61
1:D:22:PRO:HG3	2:F:1226:HOH:O	2.09	0.53
1:E:2:PRO:N	1:F:4:PRO:HD2	2.34	0.42
1:A:24:GLY:HA3	1:C:25:PRO:O	2.19	0.42

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	20/27 (74%)	20 (100%)	0	0	100	100
1	B	20/27 (74%)	20 (100%)	0	0	100	100
1	C	21/27 (78%)	20 (95%)	1 (5%)	0	100	100
1	D	21/27 (78%)	21 (100%)	0	0	100	100
1	E	21/27 (78%)	21 (100%)	0	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	20/27 (74%)	20 (100%)	0	0	100	100
All	All	123/162 (76%)	122 (99%)	1 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	14/16 (88%)	14 (100%)	0	100	100
1	B	14/16 (88%)	14 (100%)	0	100	100
1	C	15/16 (94%)	15 (100%)	0	100	100
1	D	14/16 (88%)	14 (100%)	0	100	100
1	E	15/16 (94%)	15 (100%)	0	100	100
1	F	14/16 (88%)	14 (100%)	0	100	100
All	All	86/96 (90%)	86 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

12 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The

Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
1	HYP	A	13	1	6,8,9	1.61	1 (16%)	5,10,12	2.20	2 (40%)
1	HYP	C	13	1	6,8,9	0.79	0	5,10,12	2.33	2 (40%)
1	HYP	B	14	1	6,8,9	1.00	0	5,10,12	1.34	0
1	HYP	C	14	1	6,8,9	1.11	0	5,10,12	1.49	1 (20%)
1	HYP	A	14	1	6,8,9	0.79	0	5,10,12	1.25	0
1	HYP	F	14	1	6,8,9	0.96	0	5,10,12	3.05	1 (20%)
1	HYP	D	14	1	6,8,9	0.80	0	5,10,12	1.45	1 (20%)
1	HYP	E	14	1	6,8,9	0.88	0	5,10,12	0.78	0
1	HYP	D	13	1	6,8,9	1.26	1 (16%)	5,10,12	2.01	2 (40%)
1	HYP	F	13	1	6,8,9	1.14	1 (16%)	5,10,12	1.20	0
1	HYP	B	13	1	6,8,9	1.10	0	5,10,12	2.35	2 (40%)
1	HYP	E	13	1	6,8,9	1.00	0	5,10,12	1.50	1 (20%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	HYP	A	13	1	-	0/0/11/13	0/1/1/1
1	HYP	C	13	1	-	0/0/11/13	0/1/1/1
1	HYP	B	14	1	-	0/0/11/13	0/1/1/1
1	HYP	C	14	1	-	0/0/11/13	0/1/1/1
1	HYP	A	14	1	-	0/0/11/13	0/1/1/1
1	HYP	F	14	1	-	0/0/11/13	0/1/1/1
1	HYP	D	14	1	-	0/0/11/13	0/1/1/1
1	HYP	E	14	1	-	0/0/11/13	0/1/1/1
1	HYP	D	13	1	-	0/0/11/13	0/1/1/1
1	HYP	F	13	1	-	0/0/11/13	0/1/1/1
1	HYP	B	13	1	-	0/0/11/13	0/1/1/1
1	HYP	E	13	1	-	0/0/11/13	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	13	HYP	CB-CG	-3.17	1.46	1.52

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	13	HYP	CB-CG	-2.72	1.47	1.52
1	F	13	HYP	CB-CG	-2.33	1.48	1.52

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	14	HYP	CB-CG-CD	6.19	110.86	103.27
1	B	13	HYP	CB-CG-CD	3.25	107.25	103.27
1	C	13	HYP	OD1-CG-CD	-3.19	103.38	110.35
1	A	13	HYP	CB-CG-CD	3.13	107.10	103.27
1	D	13	HYP	CG-CB-CA	3.04	107.81	103.96
1	B	13	HYP	OD1-CG-CB	-2.97	102.69	110.03
1	A	13	HYP	OD1-CG-CD	2.80	116.47	110.35
1	C	14	HYP	CB-CG-CD	2.72	106.59	103.27
1	C	13	HYP	O-C-CA	-2.69	117.73	124.78
1	E	13	HYP	CG-CB-CA	2.63	107.28	103.96
1	D	13	HYP	CB-CG-CD	2.24	106.01	103.27
1	D	14	HYP	CB-CG-CD	-2.11	100.68	103.27

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no carbohydrates 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 ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.