Full wwPDB X-ray Structure Validation Report ⓘ

May 29, 2024 – 05:52 PM EDT

PDB ID	:	1HYL
Title	:	THE 1.8 A STRUCTURE OF COLLAGENASE FROM HYPODERMA LINEATUM
Authors	:	Broutin, I.; Arnoux, B.; Riche, C.; Lecroisey, A.; Keil, B.; Pascard, C.; Ducruix, A.
Deposited on	:	1995-05-02
Resolution	:	1.80 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

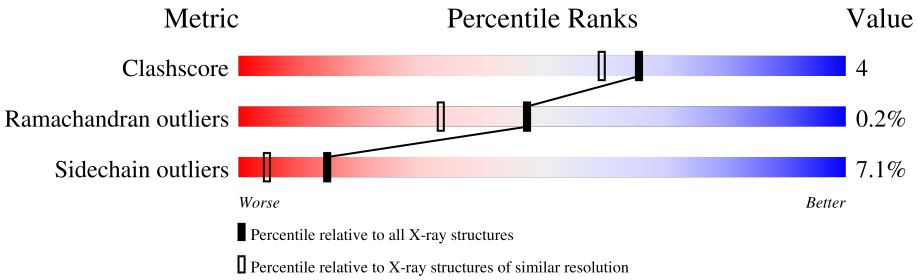
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 1.80 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	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	230	82% 14%
1	B	230	80% 16%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3851 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 HYPODERMA LINEATUM COLLAGENASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	230	Total	C	N	O	S	0	0	0
			1778	1128	292	350	8			
1	B	230	Total	C	N	O	S	0	0	0
			1778	1128	292	350	8			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	186	SER	CYS	conflict	UNP P08897
A	245	LYS	ILE	conflict	UNP P08897
B	186	SER	CYS	conflict	UNP P08897
B	245	LYS	ILE	conflict	UNP P08897

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	147	Total	O	0	0
			147	147		
2	B	148	Total	O	0	0
			148	148		

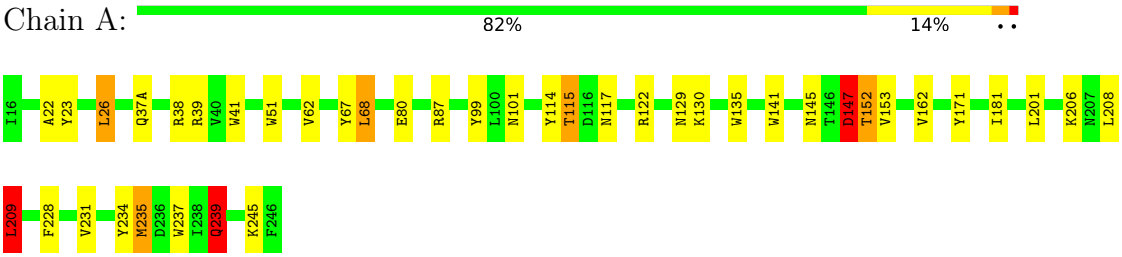
3 Residue-property plots (i)

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. 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.

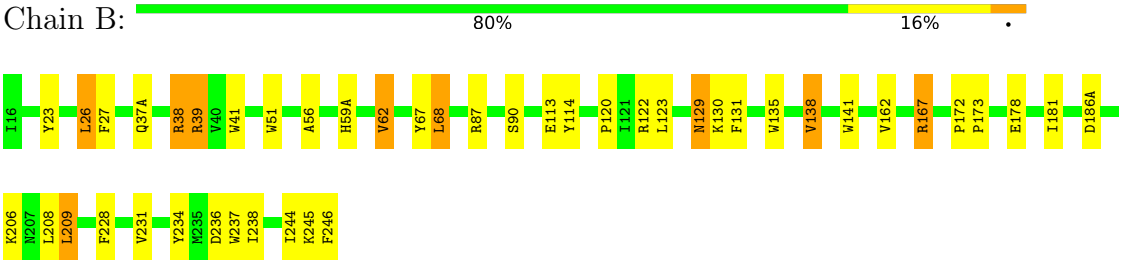
• Molecule 1: HYPODERMA LINEATUM COLLAGENASE

Chain A:



• Molecule 1: HYPODERMA LINEATUM COLLAGENASE

Chain B:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	I 4 2 2	Depositor
Cell constants a, b, c, α, β, γ	111.70Å 111.70Å 165.80Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 1.80	Depositor
% Data completeness (in resolution range)	80.7 (8.00-1.80)	Depositor
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR 2.1	Depositor
R, R_{free}	0.183 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3851	wwPDB-VP
Average B, all atoms (Å^2)	31.0	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	0.83	0/1820	1.56	32/2482 (1.3%)
1	B	0.83	0/1820	1.57	33/2482 (1.3%)
All	All	0.83	0/3640	1.56	65/4964 (1.3%)

There are no bond length outliers.

All (65) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	38	ARG	NE-CZ-NH1	12.83	126.72	120.30
1	A	38	ARG	NE-CZ-NH2	-12.73	113.94	120.30
1	A	38	ARG	NE-CZ-NH1	12.19	126.39	120.30
1	B	38	ARG	NE-CZ-NH2	-10.82	114.89	120.30
1	A	171	TYR	CB-CG-CD2	-9.68	115.19	121.00
1	B	39	ARG	NE-CZ-NH1	9.52	125.06	120.30
1	A	122	ARG	NE-CZ-NH1	9.31	124.96	120.30
1	B	51	TRP	CD1-CG-CD2	9.14	113.61	106.30
1	B	39	ARG	NE-CZ-NH2	-9.01	115.80	120.30
1	A	122	ARG	NE-CZ-NH2	-8.78	115.91	120.30
1	B	41	TRP	CD1-CG-CD2	8.18	112.84	106.30
1	B	135	TRP	CD1-CG-CD2	8.15	112.82	106.30
1	B	237	TRP	CD1-CG-CD2	8.03	112.72	106.30
1	B	141	TRP	CD1-CG-CD2	8.01	112.71	106.30
1	B	167	ARG	NE-CZ-NH1	7.92	124.26	120.30
1	B	41	TRP	CE2-CD2-CG	-7.86	101.01	107.30
1	B	87	ARG	NE-CZ-NH2	-7.84	116.38	120.30
1	A	114	TYR	CB-CG-CD2	-7.67	116.40	121.00
1	A	141	TRP	CD1-CG-CD2	7.60	112.38	106.30
1	A	87	ARG	NE-CZ-NH2	-7.59	116.51	120.30
1	A	41	TRP	CD1-CG-CD2	7.55	112.34	106.30
1	B	51	TRP	CE2-CD2-CG	-7.48	101.32	107.30
1	B	141	TRP	CE2-CD2-CG	-7.46	101.33	107.30
1	A	237	TRP	CD1-CG-CD2	7.46	112.27	106.30
1	B	122	ARG	NE-CZ-NH2	7.28	123.94	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	38	ARG	CB-CG-CD	-7.24	92.78	111.60
1	A	237	TRP	CE2-CD2-CG	-7.17	101.56	107.30
1	A	51	TRP	CD1-CG-CD2	7.15	112.02	106.30
1	B	237	TRP	CE2-CD2-CG	-7.15	101.58	107.30
1	A	99	TYR	CB-CG-CD1	-7.14	116.72	121.00
1	B	135	TRP	CE2-CD2-CG	-6.93	101.75	107.30
1	B	167	ARG	NE-CZ-NH2	-6.92	116.84	120.30
1	A	135	TRP	CD1-CG-CD2	6.73	111.69	106.30
1	B	41	TRP	CG-CD2-CE3	6.69	139.92	133.90
1	A	135	TRP	CE2-CD2-CG	-6.67	101.97	107.30
1	A	141	TRP	CE2-CD2-CG	-6.64	101.99	107.30
1	A	51	TRP	CE2-CD2-CG	-6.58	102.04	107.30
1	A	239	GLN	CA-CB-CG	6.42	127.52	113.40
1	A	41	TRP	CE2-CD2-CG	-6.30	102.26	107.30
1	A	209	LEU	CA-CB-CG	6.26	129.70	115.30
1	B	51	TRP	CG-CD1-NE1	-6.14	103.96	110.10
1	A	171	TYR	CB-CG-CD1	5.98	124.58	121.00
1	A	68	LEU	CA-CB-CG	5.88	128.81	115.30
1	B	138	VAL	CB-CA-C	-5.83	100.32	111.40
1	A	237	TRP	CG-CD2-CE3	5.83	139.15	133.90
1	A	237	TRP	CB-CG-CD1	-5.79	119.47	127.00
1	B	206	LYS	CB-CG-CD	-5.54	97.21	111.60
1	B	122	ARG	NH1-CZ-NH2	-5.53	113.32	119.40
1	A	62	VAL	CA-CB-CG2	-5.50	102.65	110.90
1	A	235	MET	CG-SD-CE	5.47	108.96	100.20
1	B	135	TRP	CG-CD1-NE1	-5.41	104.69	110.10
1	B	23	TYR	CB-CG-CD2	-5.35	117.79	121.00
1	B	67	TYR	CB-CG-CD1	-5.35	117.79	121.00
1	B	237	TRP	CB-CG-CD1	-5.26	120.16	127.00
1	B	68	LEU	CA-CB-CG	5.20	127.26	115.30
1	B	237	TRP	CG-CD1-NE1	-5.20	104.90	110.10
1	B	62	VAL	CA-CB-CG2	-5.16	103.16	110.90
1	A	141	TRP	CG-CD1-NE1	-5.14	104.96	110.10
1	B	41	TRP	CB-CG-CD1	-5.12	120.35	127.00
1	A	67	TYR	CB-CG-CD1	-5.11	117.94	121.00
1	A	147	ASP	O-C-N	5.10	130.86	122.70
1	A	23	TYR	CB-CG-CD1	-5.07	117.96	121.00
1	B	138	VAL	CG1-CB-CG2	5.06	119.00	110.90
1	A	141	TRP	CB-CG-CD1	-5.03	120.46	127.00
1	A	237	TRP	CG-CD1-NE1	-5.02	105.08	110.10

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	1778	0	1700	12	0
1	B	1778	0	1700	16	0
2	A	147	0	0	0	0
2	B	148	0	0	1	0
All	All	3851	0	3400	28	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 4.

All (28) 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:B:114:TYR:CZ	1:B:120:PRO:HG3	1.96	1.01
1:A:80:GLU:OE2	1:A:115:THR:HG21	1.87	0.75
1:A:115:THR:HG22	1:A:117:ASN:H	1.60	0.67
1:A:209:LEU:HD13	1:A:231:VAL:HG21	1.86	0.58
1:B:234:TYR:O	1:B:238:ILE:HG13	2.05	0.56
1:B:114:TYR:OH	1:B:120:PRO:HG3	2.05	0.55
1:B:129:ASN:HD22	1:B:131:PHE:H	1.53	0.54
1:A:22:ALA:HA	1:A:26:LEU:HD12	1.89	0.53
1:B:181:ILE:CG2	1:B:228:PHE:HB2	2.42	0.50
1:A:201:LEU:HD13	1:A:206:LYS:HD2	1.92	0.49
1:B:56:ALA:O	1:B:59(A):HIS:HB3	2.13	0.49
1:B:244:ILE:HG13	1:B:246:PHE:CE2	2.49	0.48
1:B:129:ASN:ND2	1:B:131:PHE:H	2.12	0.47
1:A:115:THR:HG22	1:A:117:ASN:N	2.27	0.47
1:A:147:ASP:CG	1:A:152:THR:H	2.19	0.46
1:A:239:GLN:HE21	1:A:245:LYS:NZ	2.13	0.46
1:B:26:LEU:HD13	1:B:27:PHE:CE2	2.50	0.45
1:A:235:MET:O	1:A:239:GLN:HB2	2.16	0.45
1:B:172:PRO:HA	1:B:173:PRO:HD2	1.65	0.44
1:B:181:ILE:HG23	1:B:228:PHE:HB2	1.99	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:245:LYS:N	1:B:245:LYS:HD2	2.32	0.43
1:A:181:ILE:HG23	1:A:228:PHE:HB2	2.01	0.43
1:B:37(A):GLN:HG3	1:B:62:VAL:HG23	1.99	0.43
1:B:236:ASP:HB2	2:B:320:HOH:O	2.19	0.42
1:A:130:LYS:HA	1:A:130:LYS:HD3	1.74	0.41
1:B:130:LYS:HA	1:B:130:LYS:HD3	1.92	0.41
1:A:101:ASN:HA	1:A:234:TYR:OH	2.21	0.41
1:B:209:LEU:HD13	1:B:231:VAL:HG21	2.03	0.40

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	228/230 (99%)	225 (99%)	2 (1%)	1 (0%)	34 21
1	B	228/230 (99%)	224 (98%)	4 (2%)	0	100 100
All	All	456/460 (99%)	449 (98%)	6 (1%)	1 (0%)	47 33

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	152	THR

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	197/197 (100%)	184 (93%)	13 (7%)	16 5
1	B	197/197 (100%)	182 (92%)	15 (8%)	13 4
All	All	394/394 (100%)	366 (93%)	28 (7%)	14 5

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	26	LEU
1	A	37(A)	GLN
1	A	39	ARG
1	A	68	LEU
1	A	115	THR
1	A	129	ASN
1	A	145	ASN
1	A	147	ASP
1	A	153	VAL
1	A	162	VAL
1	A	208	LEU
1	A	209	LEU
1	A	239	GLN
1	B	26	LEU
1	B	38	ARG
1	B	39	ARG
1	B	68	LEU
1	B	90	SER
1	B	113	GLU
1	B	123	LEU
1	B	129	ASN
1	B	138	VAL
1	B	162	VAL
1	B	167	ARG
1	B	178	GLU
1	B	186(A)	ASP
1	B	208	LEU
1	B	209	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	49	ASN
1	A	129	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	156	GLN
1	A	207	ASN
1	A	239	GLN
1	B	49	ASN
1	B	128	ASN
1	B	129	ASN

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 monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

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.