



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 07:00 am GMT

PDB ID : 1T60  
Title : Crystal structure of Type IV collagen NC1 domain from bovine lens capsule  
Authors : Vanacore, R.M.; Shanmugasundararaj, S.; Friedman, D.B.; Bondar, O.; Hudson, B.G.; Sundaramoorthy, M.  
Deposited on : 2004-05-05  
Resolution : 1.50 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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : recalc28949

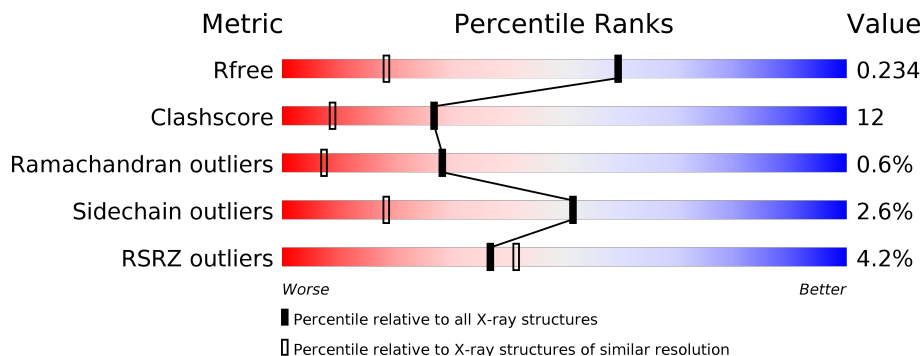
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.50 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))
$R_{free}$	100719	2279 (1.50-1.50)
Clashscore	112137	2503 (1.50-1.50)
Ramachandran outliers	110173	2445 (1.50-1.50)
Sidechain outliers	110143	2443 (1.50-1.50)
RSRZ outliers	101464	2305 (1.50-1.50)

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	A	229	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 1%, green 1%, green 78%, yellow 78%, yellow 97%, grey 97%);"></div> <div style="display: flex; justify-content: space-between; width: 90%; margin: 0 auto;"> <span>78%</span> <span>19%</span> </div> </div>
1	B	229	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 7%, green 7%, green 75%, yellow 75%, yellow 97%, grey 97%);"></div> <div style="display: flex; justify-content: space-between; width: 90%; margin: 0 auto;"> <span>75%</span> <span>22%</span> </div> </div>
1	D	229	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 2%, green 2%, green 76%, yellow 76%, yellow 96%, grey 96%);"></div> <div style="display: flex; justify-content: space-between; width: 90%; margin: 0 auto;"> <span>76%</span> <span>20%</span> </div> </div>
1	E	229	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 3%, green 3%, green 79%, yellow 79%, yellow 97%, grey 97%);"></div> <div style="display: flex; justify-content: space-between; width: 90%; margin: 0 auto;"> <span>79%</span> <span>18%</span> </div> </div>
1	G	229	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 3%, green 3%, green 81%, yellow 81%, yellow 98%, grey 98%);"></div> <div style="display: flex; justify-content: space-between; width: 90%; margin: 0 auto;"> <span>81%</span> <span>17%</span> </div> </div>
1	H	229	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 5%, green 5%, green 76%, yellow 76%, yellow 98%, grey 98%);"></div> <div style="display: flex; justify-content: space-between; width: 90%; margin: 0 auto;"> <span>76%</span> <span>22%</span> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	J	229	
1	K	229	
1	M	229	
1	N	229	
1	P	229	
1	Q	229	
1	S	229	
1	T	229	
1	V	229	
1	W	229	
2	C	227	
2	F	227	
2	I	227	
2	L	227	
2	O	227	
2	R	227	
2	U	227	
2	X	227	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	MPD	D	5007	-	-	-	X
5	MPD	J	5009	-	-	-	X
5	MPD	M	5008	-	-	-	X
5	MPD	P	5006	-	-	X	X
5	MPD	S	5004	-	-	-	X
5	MPD	S	5005	-	-	-	X
5	MPD	T	5002	-	-	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 44599 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 type iv collagen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	223	Total	C	N	O	S	0	0	0
			1724	1085	300	318	21			
1	B	225	Total	C	N	O	S	0	0	0
			1745	1097	307	320	21			
1	D	224	Total	C	N	O	S	0	0	0
			1734	1091	303	319	21			
1	E	224	Total	C	N	O	S	0	0	0
			1735	1091	304	319	21			
1	G	228	Total	C	N	O	S	0	0	0
			1767	1110	310	326	21			
1	H	227	Total	C	N	O	S	0	0	0
			1760	1105	309	325	21			
1	J	224	Total	C	N	O	S	0	0	0
			1734	1091	303	319	21			
1	K	224	Total	C	N	O	S	0	0	0
			1734	1091	303	319	21			
1	M	223	Total	C	N	O	S	0	0	0
			1724	1085	300	318	21			
1	N	224	Total	C	N	O	S	0	0	0
			1738	1093	304	320	21			
1	P	225	Total	C	N	O	S	0	0	0
			1745	1097	307	320	21			
1	Q	224	Total	C	N	O	S	0	0	0
			1734	1091	303	319	21			
1	S	225	Total	C	N	O	S	0	0	0
			1745	1097	307	320	21			
1	T	227	Total	C	N	O	S	0	0	0
			1760	1105	309	325	21			
1	V	224	Total	C	N	O	S	0	0	0
			1734	1091	303	319	21			
1	W	224	Total	C	N	O	S	0	0	0
			1734	1091	303	319	21			

- Molecule 2 is a protein called type iv collagen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	223	Total	C	N	O	S	0	0	0
			1724	1093	291	321	19			
2	F	222	Total	C	N	O	S	0	0	0
			1720	1089	291	321	19			
2	I	225	Total	C	N	O	S	0	0	0
			1738	1100	294	325	19			
2	L	221	Total	C	N	O	S	0	0	0
			1708	1080	290	319	19			
2	O	224	Total	C	N	O	S	0	0	0
			1732	1097	293	323	19			
2	R	224	Total	C	N	O	S	0	0	0
			1732	1097	293	323	19			
2	U	224	Total	C	N	O	S	0	0	0
			1732	1097	293	323	19			
2	X	223	Total	C	N	O	S	0	0	0
			1724	1091	292	322	19			

- Molecule 3 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	P	1	Total	K	0	0
			1	1		
3	G	1	Total	K	0	0
			1	1		
3	D	1	Total	K	0	0
			1	1		
3	H	1	Total	K	0	0
			1	1		
3	B	1	Total	K	0	0
			1	1		
3	I	1	Total	K	0	0
			1	1		
3	C	2	Total	K	0	0
			2	2		
3	V	1	Total	K	0	0
			1	1		
3	T	1	Total	K	0	0
			1	1		
3	N	1	Total	K	0	0
			1	1		
3	U	2	Total	K	0	0
			2	2		

*Continued on next page...*

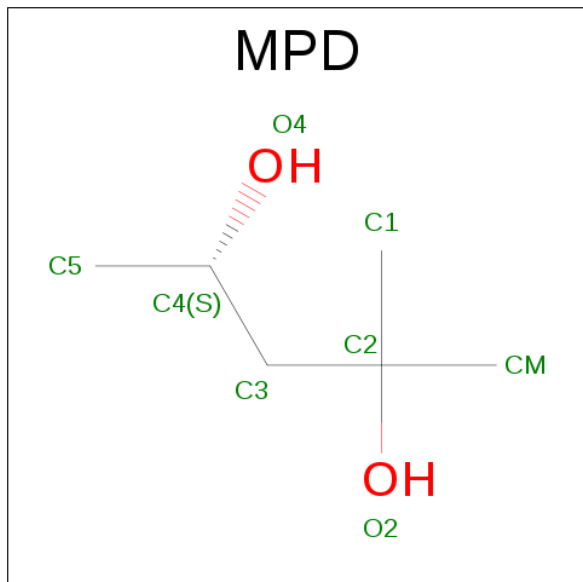
*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	O	1	Total K 1 1	0	0
3	L	1	Total K 1 1	0	0
3	M	1	Total K 1 1	0	0

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	P	1	Total Cl 1 1	0	0
4	G	1	Total Cl 1 1	0	0
4	J	1	Total Cl 1 1	0	0
4	D	1	Total Cl 1 1	0	0
4	I	1	Total Cl 1 1	0	0
4	C	1	Total Cl 1 1	0	0
4	V	1	Total Cl 1 1	0	0
4	A	1	Total Cl 1 1	0	0
4	U	1	Total Cl 1 1	0	0
4	X	1	Total Cl 1 1	0	0
4	O	1	Total Cl 1 1	0	0
4	R	1	Total Cl 1 1	0	0
4	L	1	Total Cl 1 1	0	0
4	S	1	Total Cl 1 1	0	0
4	F	1	Total Cl 1 1	0	0
4	M	1	Total Cl 1 1	0	0

- Molecule 5 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			8	6	2		
5	T	1	Total	C	O	0	0
			8	6	2		
5	M	1	Total	C	O	0	0
			8	6	2		
5	S	1	Total	C	O	0	0
			8	6	2		
5	S	1	Total	C	O	0	0
			8	6	2		
5	P	1	Total	C	O	0	0
			8	6	2		
5	D	1	Total	C	O	0	0
			8	6	2		
5	M	1	Total	C	O	0	0
			8	6	2		
5	J	1	Total	C	O	0	0
			8	6	2		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	112	Total	O	0	0
			112	112		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	101	Total 101	O 101	0	0
6	C	116	Total 116	O 116	0	0
6	D	134	Total 134	O 134	0	0
6	E	127	Total 127	O 127	0	0
6	F	117	Total 117	O 117	0	0
6	G	144	Total 144	O 144	0	0
6	H	104	Total 104	O 104	0	0
6	I	121	Total 121	O 121	0	0
6	J	97	Total 97	O 97	0	0
6	K	102	Total 102	O 102	0	0
6	L	111	Total 111	O 111	0	0
6	M	118	Total 118	O 118	0	0
6	N	113	Total 113	O 113	0	0
6	O	125	Total 125	O 125	0	0
6	P	133	Total 133	O 133	0	0
6	Q	108	Total 108	O 108	0	0
6	R	129	Total 129	O 129	0	0
6	S	137	Total 137	O 137	0	0
6	T	127	Total 127	O 127	0	0
6	U	139	Total 139	O 139	0	0
6	V	110	Total 110	O 110	0	0

*Continued on next page...*



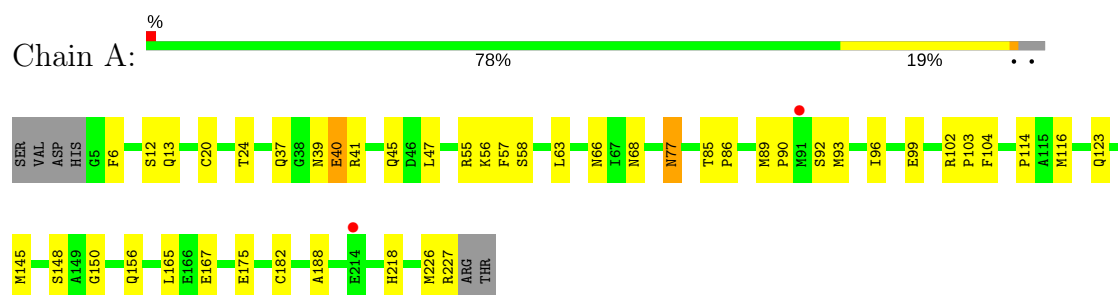
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	W	103	Total 103	O 103	0	0
6	X	110	Total 110	O 110	0	0

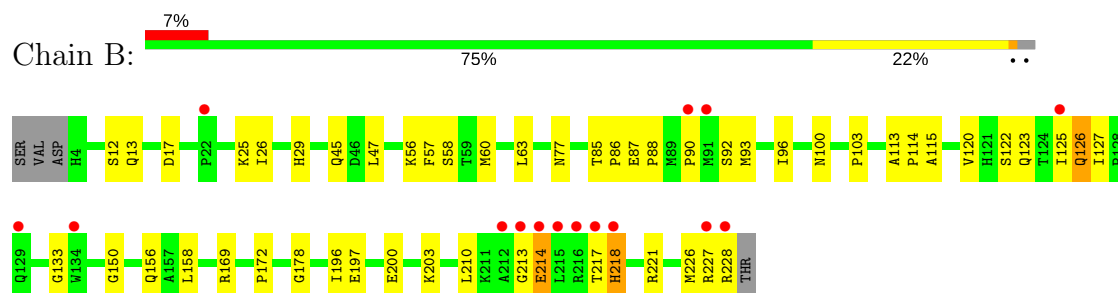
### 3 Residue-property plots [i](#)

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 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.

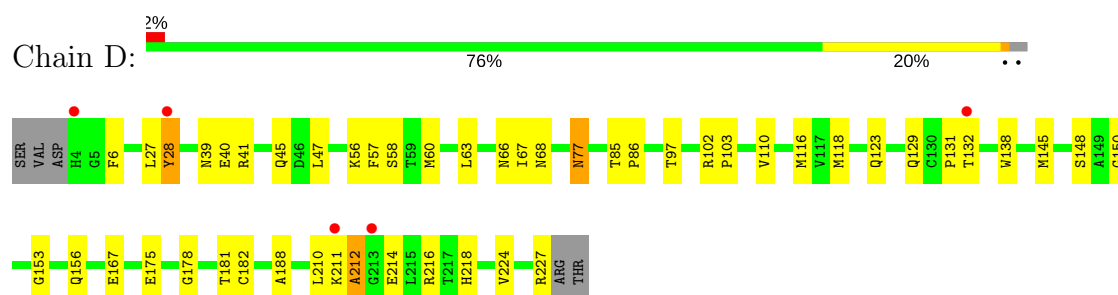
- Molecule 1: type iv collagen



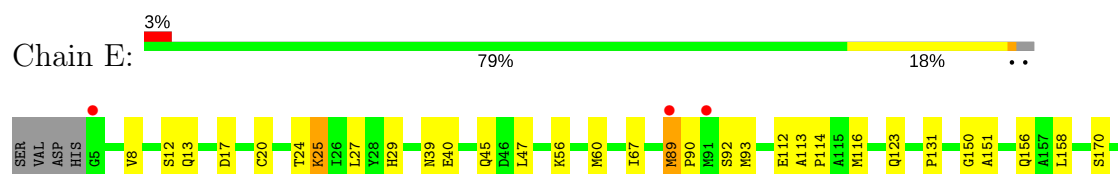
- Molecule 1: type iv collagen

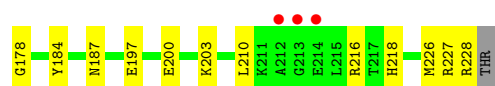


- Molecule 1: type iv collagen

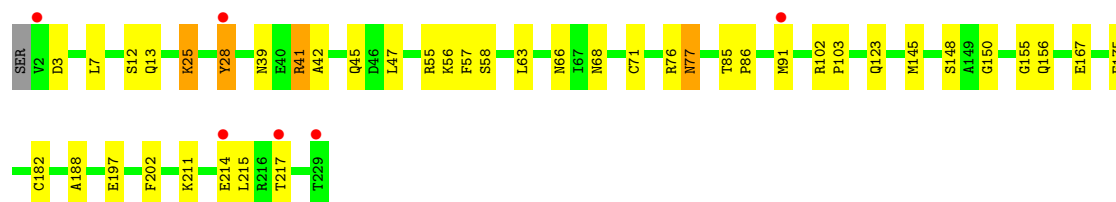
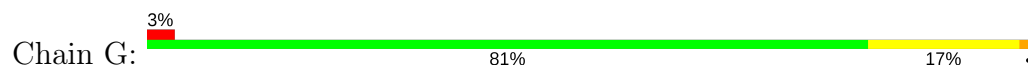


- Molecule 1: type iv collagen

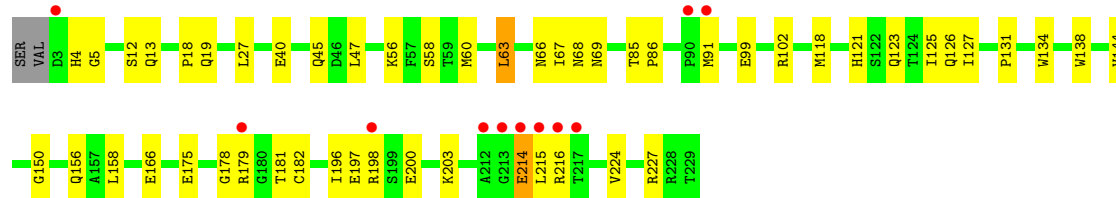
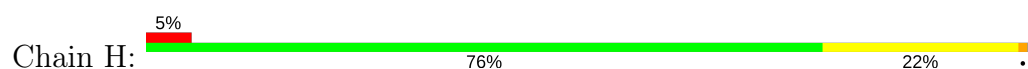




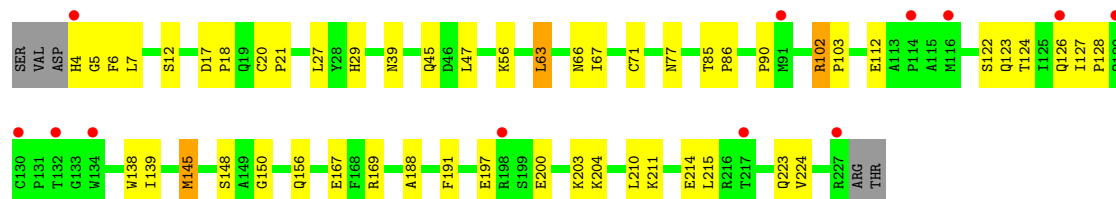
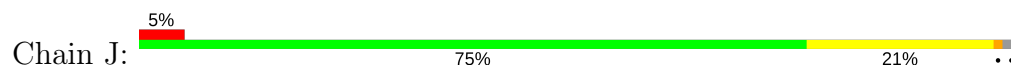
- Molecule 1: type iv collagen



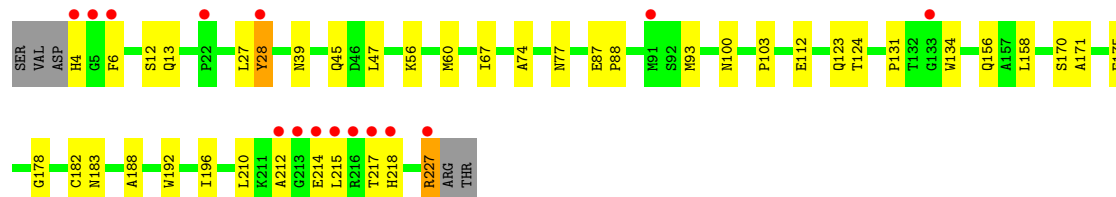
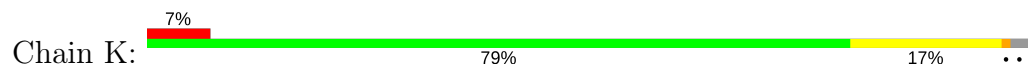
- Molecule 1: type iv collagen



- Molecule 1: type iv collagen

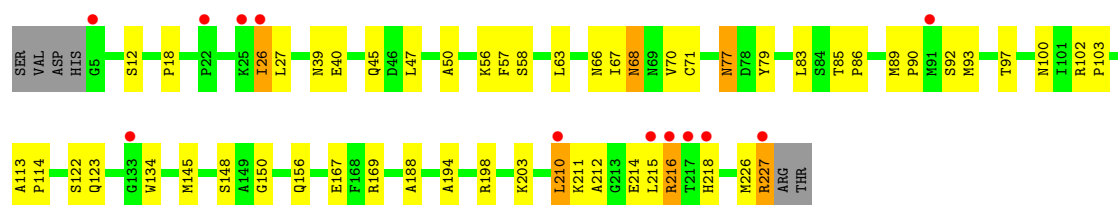


- Molecule 1: type iv collagen

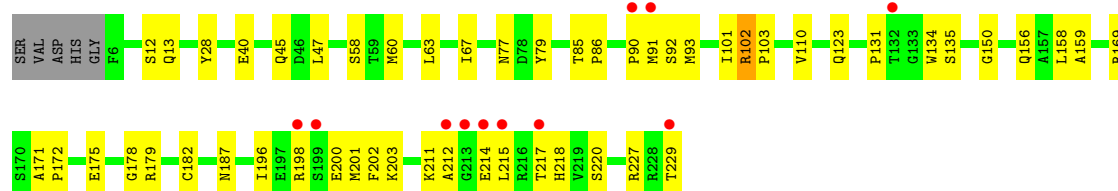
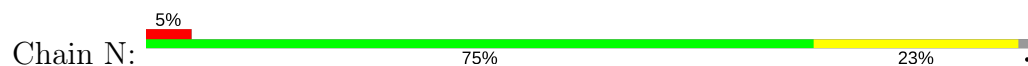


- Molecule 1: type iv collagen

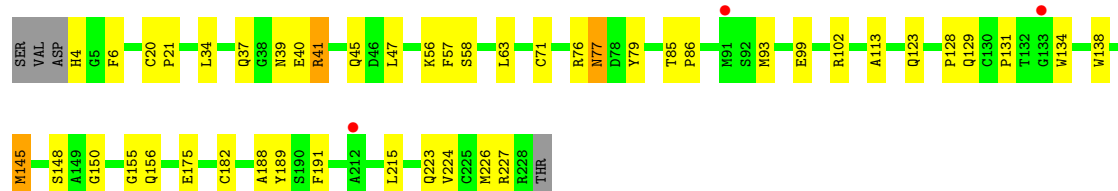
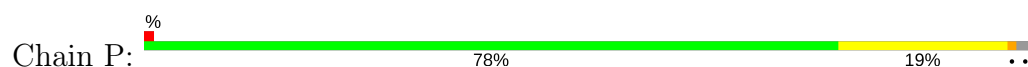




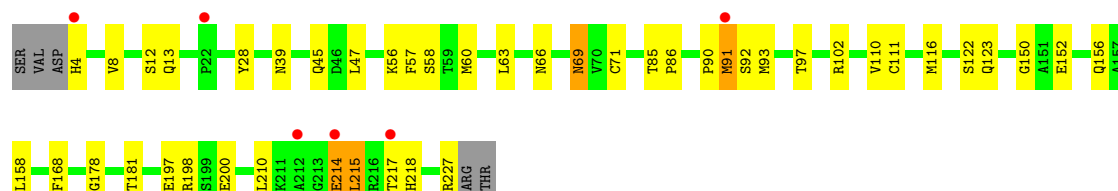
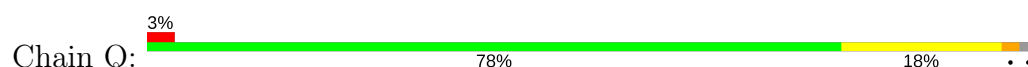
- Molecule 1: type iv collagen



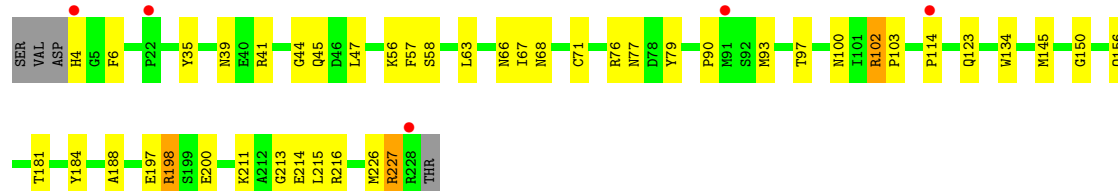
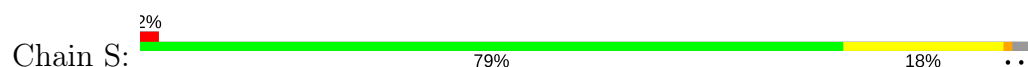
- Molecule 1: type iv collagen



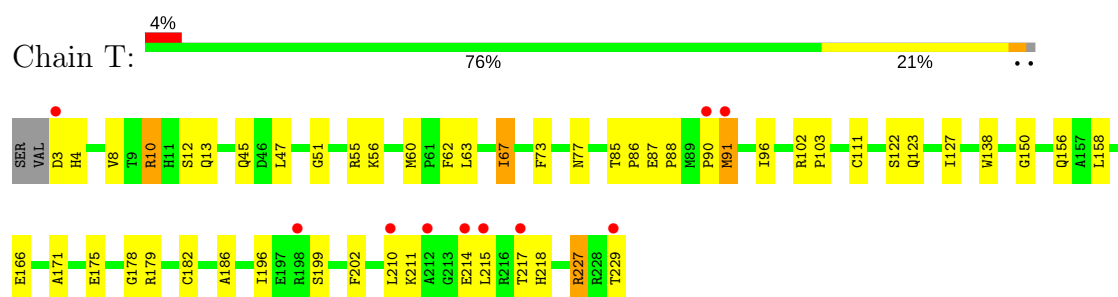
- Molecule 1: type iv collagen



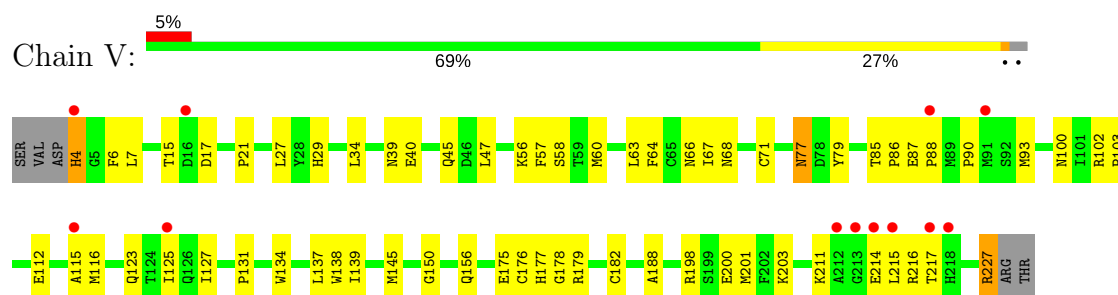
- Molecule 1: type iv collagen



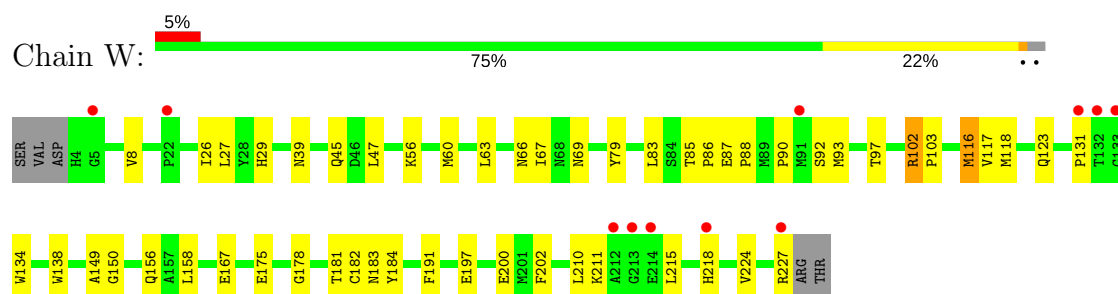
- Molecule 1: type iv collagen



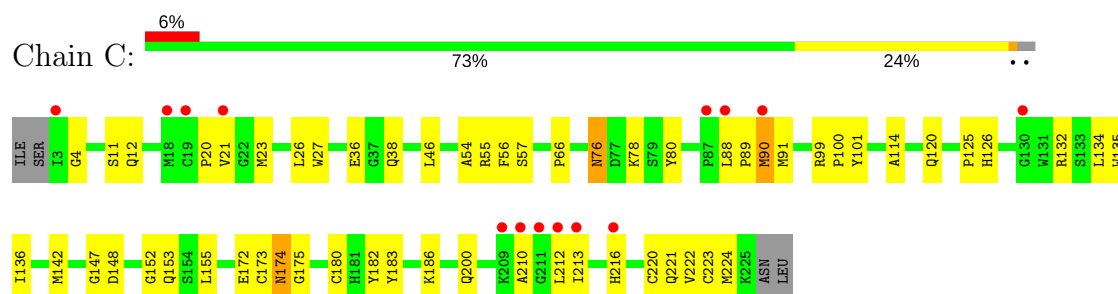
- Molecule 1: type iv collagen



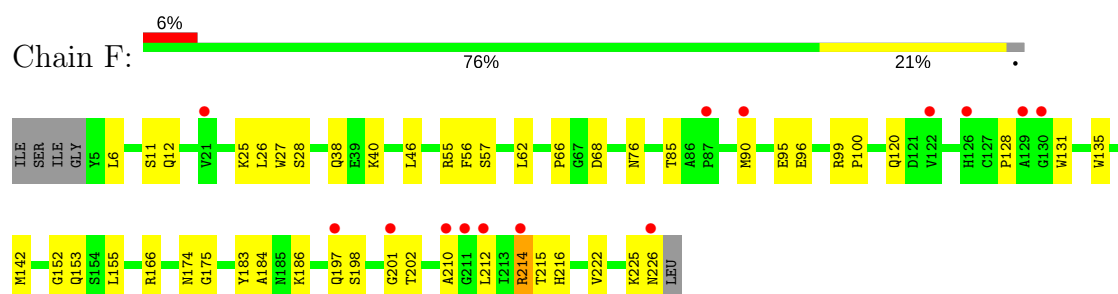
- Molecule 1: type iv collagen



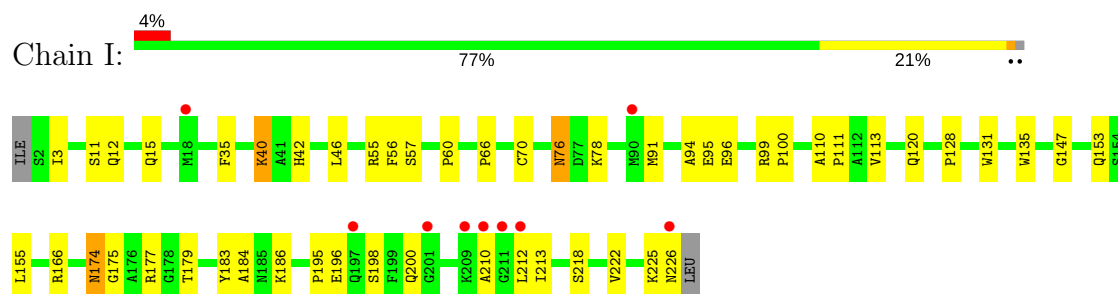
- Molecule 2: type iv collagen



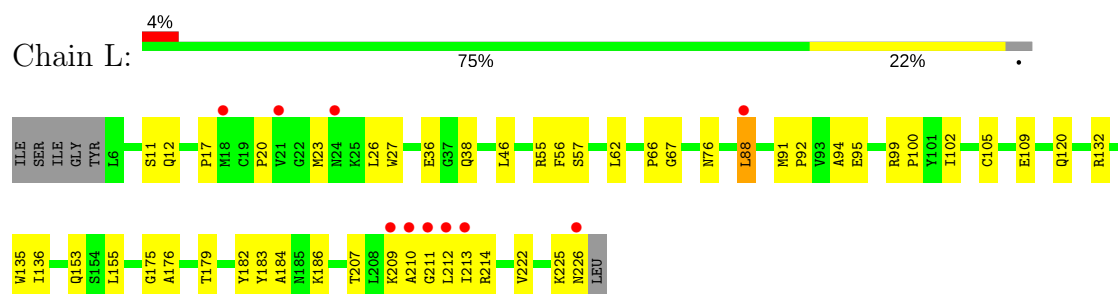
- Molecule 2: type iv collagen



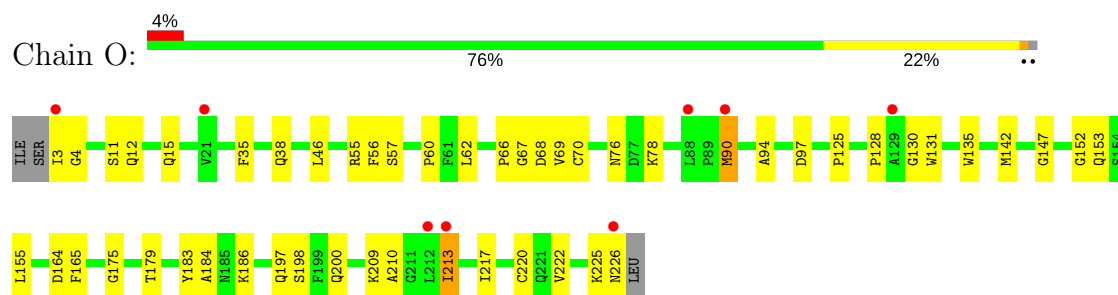
- Molecule 2: type iv collagen



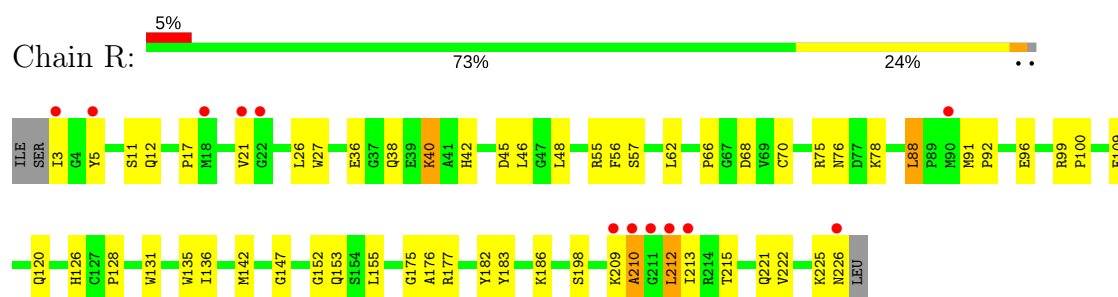
- Molecule 2: type iv collagen



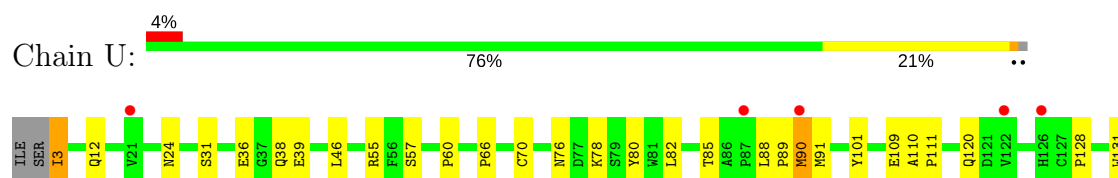
- Molecule 2: type iv collagen

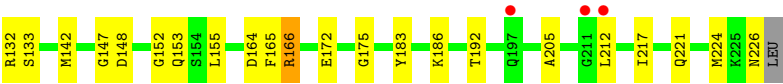


- Molecule 2: type iv collagen

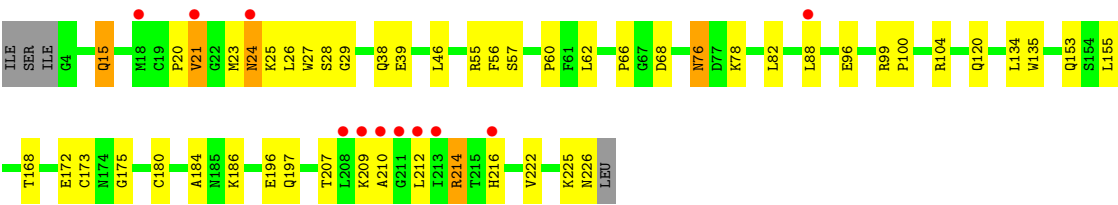
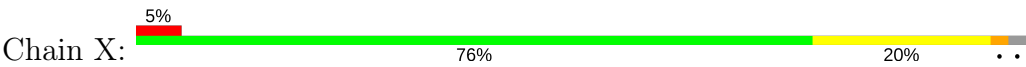


- Molecule 2: type iv collagen





• Molecule 2: type iv collagen



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	127.88Å 140.13Å 160.69Å 90.00° 91.26° 90.00°	Depositor
Resolution (Å)	8.00 – 1.50 50.02 – 1.50	Depositor EDS
% Data completeness (in resolution range)	(Not available) (8.00-1.50) 89.5 (50.02-1.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.09 (at 1.50Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.189 , 0.216 0.208 , 0.234	Depositor DCC
$R_{free}$ test set	39655 reflections (4.97%)	DCC
Wilson B-factor (Å <sup>2</sup> )	13.9	Xtriage
Anisotropy	0.228	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 62.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.018 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	44599	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 32.08 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 9.9581e-04.*

<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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: K, MPD, CL

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	0.35	0/1774	0.62	0/2410
1	B	0.34	0/1796	0.61	0/2439
1	D	0.36	0/1785	0.63	0/2425
1	E	0.35	0/1785	0.61	0/2424
1	G	0.36	0/1818	0.62	0/2470
1	H	0.34	0/1811	0.60	0/2460
1	J	0.34	0/1785	0.62	0/2425
1	K	0.34	0/1785	0.61	0/2425
1	M	0.35	0/1774	0.61	0/2410
1	N	0.33	0/1788	0.61	0/2429
1	P	0.35	0/1796	0.61	0/2439
1	Q	0.35	0/1785	0.61	0/2425
1	S	0.34	0/1796	0.60	0/2439
1	T	0.34	0/1811	0.61	0/2460
1	V	0.33	0/1785	0.59	0/2425
1	W	0.33	0/1785	0.59	0/2425
2	C	0.36	0/1775	0.65	0/2415
2	F	0.36	0/1771	0.63	0/2410
2	I	0.34	0/1789	0.63	0/2434
2	L	0.35	0/1758	0.64	0/2392
2	O	0.35	0/1783	0.63	0/2426
2	R	0.35	0/1783	0.63	0/2426
2	U	0.35	0/1783	0.64	0/2426
2	X	0.34	0/1775	0.63	0/2415
All	All	0.35	0/42876	0.62	0/58274

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

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	1724	0	1640	45	0
1	B	1745	0	1661	56	0
1	D	1734	0	1647	56	0
1	E	1735	0	1654	39	0
1	G	1767	0	1680	41	0
1	H	1760	0	1672	48	0
1	J	1734	0	1647	55	0
1	K	1734	0	1648	49	0
1	M	1724	0	1640	65	0
1	N	1738	0	1658	49	0
1	P	1745	0	1660	44	0
1	Q	1734	0	1648	41	0
1	S	1745	0	1661	48	0
1	T	1760	0	1672	55	0
1	V	1734	0	1647	57	0
1	W	1734	0	1648	49	0
2	C	1724	0	1637	54	0
2	F	1720	0	1629	39	0
2	I	1738	0	1648	54	0
2	L	1708	0	1619	40	0
2	O	1732	0	1643	43	0
2	R	1732	0	1643	56	0
2	U	1732	0	1643	48	0
2	X	1724	0	1631	51	0
3	B	1	0	0	0	0
3	C	2	0	0	0	0
3	D	1	0	0	0	0
3	G	1	0	0	0	0
3	H	1	0	0	0	0
3	I	1	0	0	0	0
3	L	1	0	0	0	0
3	M	1	0	0	0	0
3	N	1	0	0	0	0
3	O	1	0	0	0	0
3	P	1	0	0	0	0
3	T	1	0	0	0	0
3	U	2	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	V	1	0	0	0	0
4	A	1	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
4	F	1	0	0	0	0
4	G	1	0	0	0	0
4	I	1	0	0	0	0
4	J	1	0	0	0	0
4	L	1	0	0	0	0
4	M	1	0	0	0	0
4	O	1	0	0	0	0
4	P	1	0	0	0	0
4	R	1	0	0	0	0
4	S	1	0	0	0	0
4	U	1	0	0	0	0
4	V	1	0	0	0	0
4	X	1	0	0	0	0
5	A	8	0	14	0	0
5	D	8	0	14	1	0
5	J	8	0	14	3	0
5	M	16	0	28	1	0
5	P	8	0	14	6	0
5	S	16	0	28	5	0
5	T	8	0	14	0	0
6	A	112	0	0	4	0
6	B	101	0	0	1	0
6	C	116	0	0	0	0
6	D	134	0	0	3	0
6	E	127	0	0	1	0
6	F	117	0	0	0	0
6	G	144	0	0	1	0
6	H	104	0	0	2	0
6	I	121	0	0	2	0
6	J	97	0	0	1	0
6	K	102	0	0	1	0
6	L	111	0	0	0	0
6	M	118	0	0	2	0
6	N	113	0	0	5	0
6	O	125	0	0	1	0
6	P	133	0	0	3	0
6	Q	108	0	0	1	0
6	R	129	0	0	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	S	137	0	0	4	0
6	T	127	0	0	5	0
6	U	139	0	0	1	0
6	V	110	0	0	3	0
6	W	103	0	0	2	0
6	X	110	0	0	5	0
All	All	44599	0	39702	939	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 12.

The worst 5 of 939 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:203:LYS:HE2	1:B:203:LYS:HA	1.40	1.03
1:E:27:LEU:HD21	1:E:112:GLU:HG3	1.40	1.00
1:V:123:GLN:HE22	2:X:55:ARG:H	1.20	0.89
1:N:200:GLU:HA	1:N:203:LYS:HE2	1.56	0.87
1:J:45:GLN:HE22	1:K:158:LEU:H	1.22	0.85

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	221/229 (96%)	214 (97%)	6 (3%)	1 (0%)	32 10
1	B	223/229 (97%)	213 (96%)	9 (4%)	1 (0%)	38 14
1	D	222/229 (97%)	211 (95%)	9 (4%)	2 (1%)	20 4
1	E	222/229 (97%)	213 (96%)	9 (4%)	0	100 100
1	G	226/229 (99%)	218 (96%)	6 (3%)	2 (1%)	20 4

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	H	225/229 (98%)	215 (96%)	9 (4%)	1 (0%)	38	14
1	J	222/229 (97%)	213 (96%)	8 (4%)	1 (0%)	32	10
1	K	222/229 (97%)	211 (95%)	10 (4%)	1 (0%)	32	10
1	M	221/229 (96%)	210 (95%)	10 (4%)	1 (0%)	32	10
1	N	222/229 (97%)	213 (96%)	7 (3%)	2 (1%)	20	4
1	P	223/229 (97%)	217 (97%)	4 (2%)	2 (1%)	20	4
1	Q	222/229 (97%)	212 (96%)	10 (4%)	0	100	100
1	S	223/229 (97%)	215 (96%)	7 (3%)	1 (0%)	38	14
1	T	225/229 (98%)	212 (94%)	13 (6%)	0	100	100
1	V	222/229 (97%)	209 (94%)	12 (5%)	1 (0%)	32	10
1	W	222/229 (97%)	210 (95%)	12 (5%)	0	100	100
2	C	221/227 (97%)	203 (92%)	17 (8%)	1 (0%)	32	10
2	F	220/227 (97%)	205 (93%)	13 (6%)	2 (1%)	20	4
2	I	223/227 (98%)	211 (95%)	10 (4%)	2 (1%)	20	4
2	L	219/227 (96%)	203 (93%)	15 (7%)	1 (0%)	32	10
2	O	222/227 (98%)	208 (94%)	11 (5%)	3 (1%)	13	2
2	R	222/227 (98%)	208 (94%)	12 (5%)	2 (1%)	20	4
2	U	222/227 (98%)	212 (96%)	9 (4%)	1 (0%)	32	10
2	X	221/227 (97%)	200 (90%)	19 (9%)	2 (1%)	20	4
All	All	5333/5480 (97%)	5056 (95%)	247 (5%)	30 (1%)	28	7

5 of 30 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	214	GLU
1	D	212	ALA
2	F	175	GLY
1	H	214	GLU
2	L	175	GLY

### 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	189/195 (97%)	185 (98%)	4 (2%)	59	26
1	B	191/195 (98%)	189 (99%)	2 (1%)	80	60
1	D	190/195 (97%)	186 (98%)	4 (2%)	59	26
1	E	190/195 (97%)	185 (97%)	5 (3%)	51	18
1	G	194/195 (100%)	188 (97%)	6 (3%)	45	13
1	H	193/195 (99%)	188 (97%)	5 (3%)	51	18
1	J	190/195 (97%)	186 (98%)	4 (2%)	59	26
1	K	190/195 (97%)	188 (99%)	2 (1%)	78	56
1	M	189/195 (97%)	181 (96%)	8 (4%)	34	6
1	N	191/195 (98%)	187 (98%)	4 (2%)	59	26
1	P	191/195 (98%)	187 (98%)	4 (2%)	59	26
1	Q	190/195 (97%)	184 (97%)	6 (3%)	44	12
1	S	191/195 (98%)	184 (96%)	7 (4%)	39	9
1	T	193/195 (99%)	188 (97%)	5 (3%)	51	18
1	V	190/195 (97%)	184 (97%)	6 (3%)	44	12
1	W	190/195 (97%)	186 (98%)	4 (2%)	59	26
2	C	187/191 (98%)	182 (97%)	5 (3%)	50	17
2	F	187/191 (98%)	182 (97%)	5 (3%)	50	17
2	I	189/191 (99%)	185 (98%)	4 (2%)	59	26
2	L	186/191 (97%)	181 (97%)	5 (3%)	50	17
2	O	188/191 (98%)	184 (98%)	4 (2%)	59	26
2	R	188/191 (98%)	182 (97%)	6 (3%)	44	12
2	U	188/191 (98%)	183 (97%)	5 (3%)	50	17
2	X	187/191 (98%)	180 (96%)	7 (4%)	39	9
All	All	4552/4648 (98%)	4435 (97%)	117 (3%)	51	18

5 of 117 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	M	68	ASN
2	O	164	ASP
1	W	102	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	M	79	TYR
1	N	79	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 190 such sidechains are listed below:

Mol	Chain	Res	Type
2	L	15	GLN
1	N	129	GLN
1	W	37	GLN
2	L	120	GLN
1	M	66	ASN

### 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](#)

Of 41 ligands modelled in this entry, 32 are monoatomic - leaving 9 for Mogul analysis.

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	$\# Z  > 2$
5	MPD	A	5001	-	7,7,7	0.50	0	9,10,10	0.50	0
5	MPD	D	5007	-	7,7,7	0.90	0	9,10,10	0.48	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	MPD	J	5009	-	7,7,7	0.54	0	9,10,10	0.55	0
5	MPD	M	5003	-	7,7,7	0.58	0	9,10,10	0.53	0
5	MPD	M	5008	-	7,7,7	0.59	0	9,10,10	0.46	0
5	MPD	P	5006	-	7,7,7	0.58	0	9,10,10	0.52	0
5	MPD	S	5004	-	7,7,7	0.50	0	9,10,10	0.47	0
5	MPD	S	5005	-	7,7,7	0.76	0	9,10,10	0.65	0
5	MPD	T	5002	-	7,7,7	1.53	2 (28%)	9,10,10	1.04	0

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
5	MPD	A	5001	-	-	0/5/5/5	0/0/0/0
5	MPD	D	5007	-	-	0/5/5/5	0/0/0/0
5	MPD	J	5009	-	-	0/5/5/5	0/0/0/0
5	MPD	M	5003	-	-	0/5/5/5	0/0/0/0
5	MPD	M	5008	-	-	0/5/5/5	0/0/0/0
5	MPD	P	5006	-	-	0/5/5/5	0/0/0/0
5	MPD	S	5004	-	-	0/5/5/5	0/0/0/0
5	MPD	S	5005	-	-	0/5/5/5	0/0/0/0
5	MPD	T	5002	-	-	0/5/5/5	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	T	5002	MPD	CM-C2	-2.36	1.44	1.52
5	T	5002	MPD	C3-C2	-2.04	1.48	1.53

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	D	5007	MPD	1	0
5	J	5009	MPD	3	0
5	M	5008	MPD	1	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	P	5006	MPD	6	0
5	S	5005	MPD	5	0

## 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<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	A	223/229 (97%)	-0.01	2 (0%) 84 87	10, 18, 33, 42	0
1	B	225/229 (98%)	0.20	15 (6%) 19 21	11, 19, 41, 49	0
1	D	224/229 (97%)	-0.17	5 (2%) 62 68	8, 14, 33, 49	0
1	E	224/229 (97%)	-0.08	6 (2%) 55 61	9, 16, 35, 45	0
1	G	228/229 (99%)	-0.11	6 (2%) 56 62	9, 15, 39, 47	0
1	H	227/229 (99%)	0.09	11 (4%) 31 35	10, 18, 39, 50	0
1	J	224/229 (97%)	0.14	12 (5%) 26 30	10, 20, 41, 49	0
1	K	224/229 (97%)	0.13	15 (6%) 19 21	10, 18, 41, 50	0
1	M	223/229 (97%)	0.17	12 (5%) 26 30	10, 20, 38, 50	0
1	N	224/229 (97%)	0.24	11 (4%) 30 34	10, 19, 41, 49	0
1	P	225/229 (98%)	-0.15	3 (1%) 77 81	10, 18, 32, 45	0
1	Q	224/229 (97%)	-0.04	6 (2%) 55 61	10, 18, 39, 47	0
1	S	225/229 (98%)	-0.05	5 (2%) 62 68	10, 17, 32, 42	0
1	T	227/229 (99%)	0.13	10 (4%) 35 40	10, 18, 42, 49	0
1	V	224/229 (97%)	0.30	12 (5%) 26 30	12, 22, 42, 50	0
1	W	224/229 (97%)	0.17	11 (4%) 30 34	10, 21, 43, 49	0
2	C	223/227 (98%)	0.27	14 (6%) 21 24	10, 20, 40, 50	0
2	F	222/227 (97%)	0.25	14 (6%) 21 24	9, 19, 38, 50	0
2	I	225/227 (99%)	0.19	9 (4%) 39 44	9, 20, 36, 48	0
2	L	221/227 (97%)	0.21	10 (4%) 34 38	11, 20, 41, 50	0
2	O	224/227 (98%)	0.05	8 (3%) 43 49	9, 17, 34, 45	0
2	R	224/227 (98%)	0.04	12 (5%) 26 30	9, 17, 37, 50	0
2	U	224/227 (98%)	0.09	8 (3%) 43 49	9, 16, 37, 47	0
2	X	223/227 (98%)	0.20	11 (4%) 30 34	10, 20, 41, 50	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
All	All	5381/5480 (98%)	0.09	228 (4%) 37 41	8, 18, 39, 50	0

The worst 5 of 228 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	212	LEU	12.5
2	X	21	VAL	11.6
2	L	212	LEU	10.7
2	X	212	LEU	9.9
2	R	211	GLY	8.6

## 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 carbohydrates in this entry.

## 6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
5	MPD	P	5006	8/8	0.59	0.34	12.19	42,45,49,50	0
5	MPD	J	5009	8/8	0.64	0.47	10.24	46,48,50,50	0
5	MPD	M	5008	8/8	0.70	0.22	7.27	38,40,44,46	0
5	MPD	S	5004	8/8	0.75	0.18	7.14	45,45,46,47	0
5	MPD	D	5007	8/8	0.71	0.30	7.13	42,44,44,46	0
5	MPD	S	5005	8/8	0.62	0.33	4.37	47,49,49,50	0
5	MPD	T	5002	8/8	0.84	0.33	3.59	45,46,47,47	0
5	MPD	A	5001	8/8	0.82	0.14	1.09	37,40,41,44	0
5	MPD	M	5003	8/8	0.92	0.08	0.19	24,29,31,33	0
3	K	M	5105	1/1	0.90	0.08	0.03	42,42,42,42	0
3	K	D	5102	1/1	0.95	0.06	-0.60	35,35,35,35	0

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	K	P	5106	1/1	0.98	0.07	-0.74	33,33,33,33	0
3	K	G	5103	1/1	0.99	0.06	-0.97	34,34,34,34	0
4	CL	L	5209	1/1	0.99	0.03	-0.98	15,15,15,15	0
4	CL	G	5205	1/1	0.99	0.05	-1.28	14,14,14,14	0
3	K	V	5108	1/1	0.96	0.06	-1.37	36,36,36,36	0
4	CL	O	5210	1/1	0.99	0.05	-1.45	13,13,13,13	0
4	CL	U	5214	1/1	1.00	0.04	-1.46	12,12,12,12	0
4	CL	X	5216	1/1	0.99	0.05	-1.47	18,18,18,18	0
4	CL	F	5204	1/1	0.99	0.05	-1.63	14,14,14,14	0
4	CL	R	5212	1/1	0.99	0.04	-1.73	14,14,14,14	0
4	CL	D	5203	1/1	1.00	0.03	-1.91	13,13,13,13	0
3	K	C	5101	1/1	0.98	0.04	-2.18	26,26,26,26	0
3	K	U	5107	1/1	0.99	0.05	-2.25	30,30,30,30	0
3	K	C	5110	1/1	0.99	0.03	-2.29	19,19,19,19	0
4	CL	S	5213	1/1	0.99	0.03	-2.34	16,16,16,16	0
4	CL	P	5211	1/1	0.99	0.04	-2.45	13,13,13,13	0
4	CL	I	5206	1/1	1.00	0.03	-2.48	11,11,11,11	0
3	K	B	5109	1/1	0.98	0.05	-2.50	20,20,20,20	0
4	CL	C	5202	1/1	0.99	0.05	-2.55	15,15,15,15	0
3	K	L	5104	1/1	0.99	0.04	-2.72	27,27,27,27	0
3	K	T	5115	1/1	0.99	0.04	-2.99	17,17,17,17	0
4	CL	V	5215	1/1	0.99	0.04	-3.01	17,17,17,17	0
4	CL	J	5207	1/1	1.00	0.03	-3.60	13,13,13,13	0
4	CL	A	5201	1/1	0.99	0.03	-3.68	13,13,13,13	0
3	K	H	5111	1/1	0.98	0.04	-3.99	21,21,21,21	0
3	K	O	5114	1/1	1.00	0.03	-4.19	16,16,16,16	0
3	K	N	5113	1/1	1.00	0.03	-4.22	15,15,15,15	0
4	CL	M	5208	1/1	0.99	0.03	-4.59	15,15,15,15	0
3	K	U	5116	1/1	0.99	0.04	-5.25	18,18,18,18	0
3	K	I	5112	1/1	0.99	0.04	-5.79	21,21,21,21	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.