wwPDB X-ray Structure Validation Summary Report (i)

Oct 9, 2023 – 04:56 PM EDT

PDB ID	:	6WVV
Title	:	Plasmodium vivax M17 leucyl aminopeptidase
Authors	:	Malcolm, T.R.; Drinkwater, N.; McGowan, S.
Deposited on	:	2020-05-07
Resolution	:	2.33 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

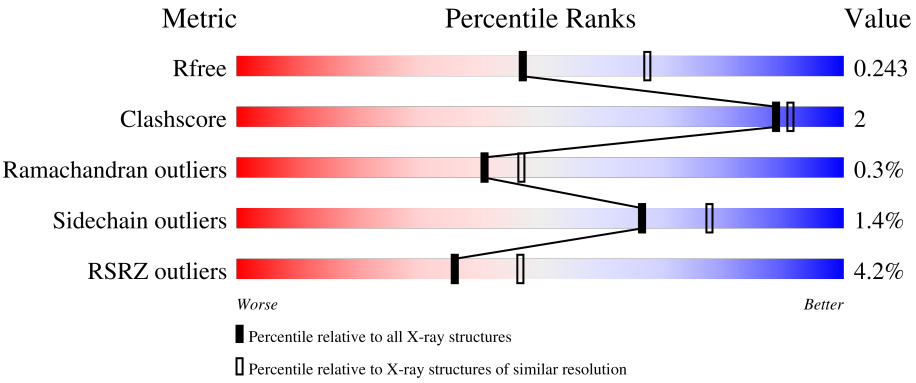
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.33 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}	130704	2096 (2.36-2.32)
Clashscore	141614	2193 (2.36-2.32)
Ramachandran outliers	138981	2159 (2.36-2.32)
Sidechain outliers	138945	2160 (2.36-2.32)
RSRZ outliers	127900	2067 (2.36-2.32)

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% 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	528	92%
1	B	528	93%
1	C	528	92%
1	D	528	92%
1	E	528	4% 87% 8% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	528	4% 89% 5% 6%
1	G	528	7% 85% 12%
1	H	528	7% 86% 9%
1	I	528	6% 85% 12%
1	J	528	4% 79% 7% 13%
1	K	528	6% 78% 5% 17%
1	L	528	6% 78% 7% 14%

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
3	PEG	F	704	-	-	-	X

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 86464 atoms, of which 41597 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 M17 leucyl aminopeptidase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	509	Total	C	H	N	O	S	0	0	0
			7847	2523	3920	637	751	16			
1	B	508	Total	C	H	N	O	S	0	0	0
			7839	2519	3919	634	751	16			
1	C	507	Total	C	H	N	O	S	0	0	0
			7837	2518	3918	633	752	16			
1	D	507	Total	C	H	N	O	S	0	1	0
			7806	2512	3896	632	750	16			
1	E	503	Total	C	H	N	O	S	0	0	0
			7240	2377	3523	611	715	14			
1	F	497	Total	C	H	N	O	S	0	0	0
			7259	2366	3565	608	706	14			
1	G	466	Total	C	H	N	O	S	0	0	0
			6543	2159	3161	561	648	14			
1	H	478	Total	C	H	N	O	S	0	0	0
			6772	2228	3280	579	671	14			
1	I	467	Total	C	H	N	O	S	0	0	0
			6564	2164	3163	568	656	13			
1	J	458	Total	C	H	N	O	S	0	0	0
			6328	2095	3027	549	643	14			
1	K	440	Total	C	H	N	O	S	0	0	0
			6076	2003	2920	531	609	13			
1	L	453	Total	C	H	N	O	S	0	0	0
			6175	2053	2946	543	619	14			

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	622	HIS	-	expression tag	UNP A0A1G4HHP8
A	623	HIS	-	expression tag	UNP A0A1G4HHP8
A	624	HIS	-	expression tag	UNP A0A1G4HHP8
A	625	HIS	-	expression tag	UNP A0A1G4HHP8
A	626	HIS	-	expression tag	UNP A0A1G4HHP8

Continued on next page...

Continued from previous page...

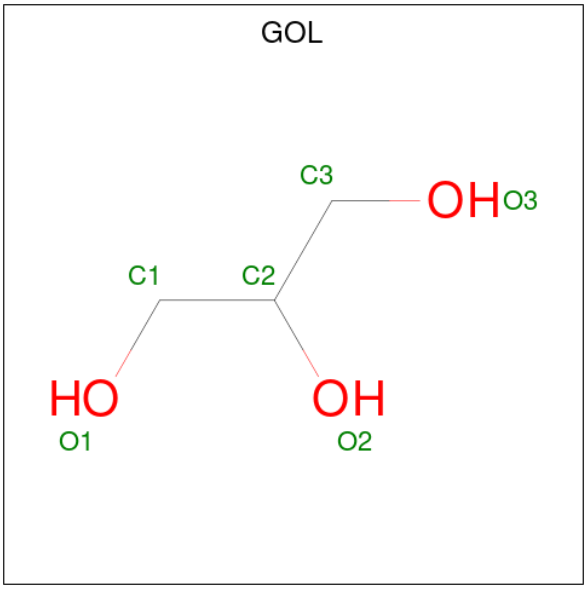
Chain	Residue	Modelled	Actual	Comment	Reference
A	627	HIS	-	expression tag	UNP A0A1G4HHP8
B	622	HIS	-	expression tag	UNP A0A1G4HHP8
B	623	HIS	-	expression tag	UNP A0A1G4HHP8
B	624	HIS	-	expression tag	UNP A0A1G4HHP8
B	625	HIS	-	expression tag	UNP A0A1G4HHP8
B	626	HIS	-	expression tag	UNP A0A1G4HHP8
B	627	HIS	-	expression tag	UNP A0A1G4HHP8
C	622	HIS	-	expression tag	UNP A0A1G4HHP8
C	623	HIS	-	expression tag	UNP A0A1G4HHP8
C	624	HIS	-	expression tag	UNP A0A1G4HHP8
C	625	HIS	-	expression tag	UNP A0A1G4HHP8
C	626	HIS	-	expression tag	UNP A0A1G4HHP8
C	627	HIS	-	expression tag	UNP A0A1G4HHP8
D	622	HIS	-	expression tag	UNP A0A1G4HHP8
D	623	HIS	-	expression tag	UNP A0A1G4HHP8
D	624	HIS	-	expression tag	UNP A0A1G4HHP8
D	625	HIS	-	expression tag	UNP A0A1G4HHP8
D	626	HIS	-	expression tag	UNP A0A1G4HHP8
D	627	HIS	-	expression tag	UNP A0A1G4HHP8
E	622	HIS	-	expression tag	UNP A0A1G4HHP8
E	623	HIS	-	expression tag	UNP A0A1G4HHP8
E	624	HIS	-	expression tag	UNP A0A1G4HHP8
E	625	HIS	-	expression tag	UNP A0A1G4HHP8
E	626	HIS	-	expression tag	UNP A0A1G4HHP8
E	627	HIS	-	expression tag	UNP A0A1G4HHP8
F	622	HIS	-	expression tag	UNP A0A1G4HHP8
F	623	HIS	-	expression tag	UNP A0A1G4HHP8
F	624	HIS	-	expression tag	UNP A0A1G4HHP8
F	625	HIS	-	expression tag	UNP A0A1G4HHP8
F	626	HIS	-	expression tag	UNP A0A1G4HHP8
F	627	HIS	-	expression tag	UNP A0A1G4HHP8
G	622	HIS	-	expression tag	UNP A0A1G4HHP8
G	623	HIS	-	expression tag	UNP A0A1G4HHP8
G	624	HIS	-	expression tag	UNP A0A1G4HHP8
G	625	HIS	-	expression tag	UNP A0A1G4HHP8
G	626	HIS	-	expression tag	UNP A0A1G4HHP8
G	627	HIS	-	expression tag	UNP A0A1G4HHP8
H	622	HIS	-	expression tag	UNP A0A1G4HHP8
H	623	HIS	-	expression tag	UNP A0A1G4HHP8
H	624	HIS	-	expression tag	UNP A0A1G4HHP8
H	625	HIS	-	expression tag	UNP A0A1G4HHP8
H	626	HIS	-	expression tag	UNP A0A1G4HHP8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	627	HIS	-	expression tag	UNP A0A1G4HHP8
I	622	HIS	-	expression tag	UNP A0A1G4HHP8
I	623	HIS	-	expression tag	UNP A0A1G4HHP8
I	624	HIS	-	expression tag	UNP A0A1G4HHP8
I	625	HIS	-	expression tag	UNP A0A1G4HHP8
I	626	HIS	-	expression tag	UNP A0A1G4HHP8
I	627	HIS	-	expression tag	UNP A0A1G4HHP8
J	622	HIS	-	expression tag	UNP A0A1G4HHP8
J	623	HIS	-	expression tag	UNP A0A1G4HHP8
J	624	HIS	-	expression tag	UNP A0A1G4HHP8
J	625	HIS	-	expression tag	UNP A0A1G4HHP8
J	626	HIS	-	expression tag	UNP A0A1G4HHP8
J	627	HIS	-	expression tag	UNP A0A1G4HHP8
K	622	HIS	-	expression tag	UNP A0A1G4HHP8
K	623	HIS	-	expression tag	UNP A0A1G4HHP8
K	624	HIS	-	expression tag	UNP A0A1G4HHP8
K	625	HIS	-	expression tag	UNP A0A1G4HHP8
K	626	HIS	-	expression tag	UNP A0A1G4HHP8
K	627	HIS	-	expression tag	UNP A0A1G4HHP8
L	622	HIS	-	expression tag	UNP A0A1G4HHP8
L	623	HIS	-	expression tag	UNP A0A1G4HHP8
L	624	HIS	-	expression tag	UNP A0A1G4HHP8
L	625	HIS	-	expression tag	UNP A0A1G4HHP8
L	626	HIS	-	expression tag	UNP A0A1G4HHP8
L	627	HIS	-	expression tag	UNP A0A1G4HHP8

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



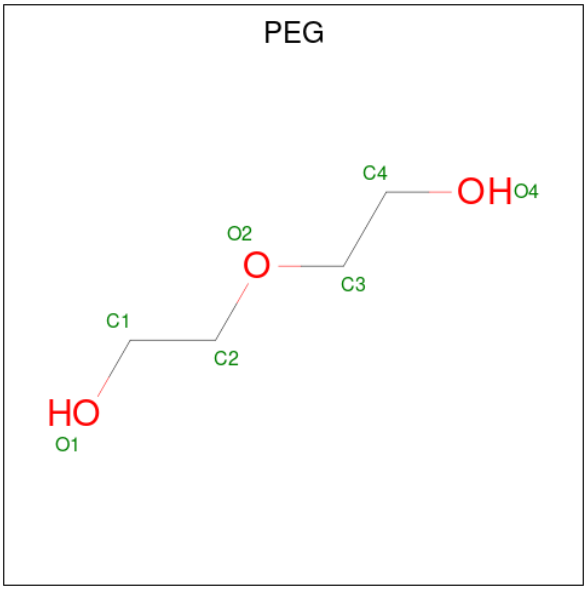
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	H	O	0	0
			14	3	8	3		
2	A	1	Total	C	H	O	0	0
			14	3	8	3		
2	A	1	Total	C	H	O	0	0
			14	3	8	3		
2	A	1	Total	C	H	O	0	0
			13	3	7	3		
2	A	1	Total	C	H	O	0	0
			14	3	8	3		
2	A	1	Total	C	H	O	0	0
			14	3	8	3		
2	B	1	Total	C	H	O	0	0
			14	3	8	3		
2	B	1	Total	C	H	O	0	0
			14	3	8	3		
2	B	1	Total	C	H	O	0	0
			14	3	8	3		
2	C	1	Total	C	H	O	0	0
			14	3	8	3		
2	C	1	Total	C	H	O	0	0
			14	3	8	3		
2	D	1	Total	C	H	O	0	0
			14	3	8	3		
2	D	1	Total	C	H	O	0	0
			14	3	8	3		
2	D	1	Total	C	H	O	0	0
			14	3	8	3		
2	E	1	Total	C	H	O	0	0
			14	3	8	3		
2	E	1	Total	C	H	O	0	0
			14	3	8	3		
2	E	1	Total	C	H	O	0	0
			14	3	8	3		
2	E	1	Total	C	H	O	0	0
			14	3	8	3		
2	E	1	Total	C	H	O	0	0
			14	3	8	3		
2	F	1	Total	C	H	O	0	0
			14	3	8	3		
2	F	1	Total	C	H	O	0	0
			14	3	8	3		
2	F	1	Total	C	H	O	0	0
			14	3	8	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	G	1	Total C H O 14 3 8 3	0	0
2	G	1	Total C H O 14 3 8 3	0	0
2	H	1	Total C H O 14 3 8 3	0	0
2	I	1	Total C H O 14 3 8 3	0	0
2	I	1	Total C H O 14 3 8 3	0	0
2	J	1	Total C H O 14 3 8 3	0	0
2	J	1	Total C H O 14 3 8 3	0	0
2	J	1	Total C H O 14 3 8 3	0	0

- Molecule 3 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C H O 17 4 10 3	0	0
3	B	1	Total C H O 17 4 10 3	0	0
3	B	1	Total C H O 17 4 10 3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	C	1	Total C H O 17 4 10 3	0	0
3	D	1	Total C H O 17 4 10 3	0	0
3	E	1	Total C H O 17 4 10 3	0	0
3	E	1	Total C H O 17 4 10 3	0	0
3	E	1	Total C H O 17 4 10 3	0	0
3	F	1	Total C H O 17 4 10 3	0	0
3	F	1	Total C H O 17 4 10 3	0	0
3	G	1	Total C H O 17 4 10 3	0	0
3	H	1	Total C H O 17 4 10 3	0	0

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

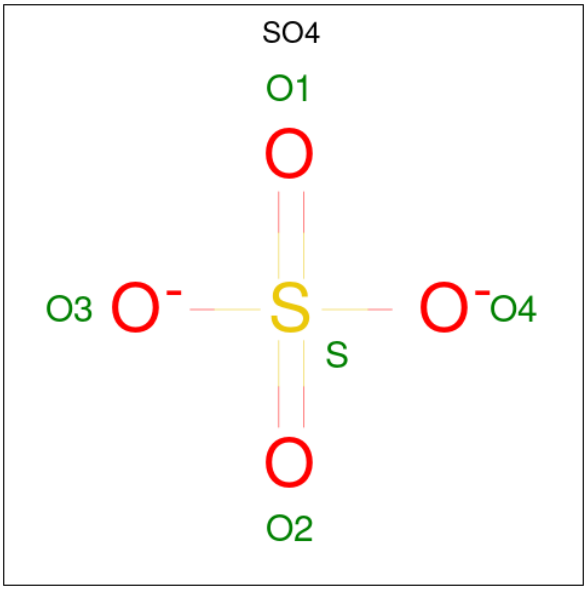
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	2	Total Zn 2 2	0	0
4	B	2	Total Zn 2 2	0	0
4	C	2	Total Zn 2 2	0	0
4	D	2	Total Zn 2 2	0	0
4	E	2	Total Zn 2 2	0	0
4	F	2	Total Zn 2 2	0	0
4	G	2	Total Zn 2 2	0	0
4	H	2	Total Zn 2 2	0	0
4	I	2	Total Zn 2 2	0	0
4	J	2	Total Zn 2 2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	K	2	Total Zn 2 2	0	0
4	L	2	Total Zn 2 2	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total O S 5 4 1	0	0
5	A	1	Total O S 5 4 1	0	0
5	A	1	Total O S 5 4 1	0	0
5	B	1	Total O S 5 4 1	0	0
5	B	1	Total O S 5 4 1	0	0
5	B	1	Total O S 5 4 1	0	0
5	C	1	Total O S 5 4 1	0	0
5	D	1	Total O S 5 4 1	0	0
5	E	1	Total O S 5 4 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	F	1	Total O S 5 4 1	0	0
5	G	1	Total O S 5 4 1	0	0
5	G	1	Total O S 5 4 1	0	0
5	H	1	Total O S 5 4 1	0	0
5	I	1	Total O S 5 4 1	0	0
5	J	1	Total O S 5 4 1	0	0
5	K	1	Total O S 5 4 1	0	0
5	L	1	Total O S 5 4 1	0	0
5	L	1	Total O S 5 4 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	206	Total O 206 206	0	0
6	B	182	Total O 182 182	0	0
6	C	201	Total O 201 201	0	0
6	D	179	Total O 179 179	0	0
6	E	154	Total O 154 154	0	0
6	F	128	Total O 128 128	0	0
6	G	108	Total O 108 108	0	0
6	H	83	Total O 83 83	0	0
6	I	56	Total O 56 56	0	0
6	J	47	Total O 47 47	0	0

Continued on next page...

Continued from previous page...

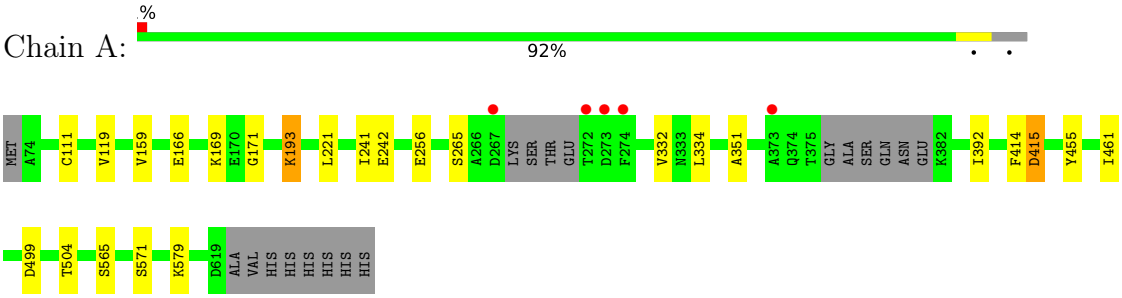
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	K	53	Total O 53 53	0	0
6	L	44	Total O 44 44	0	0

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 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.

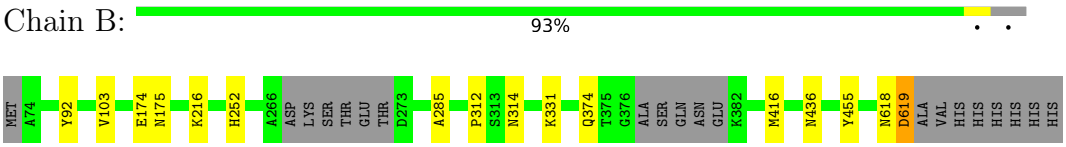
• Molecule 1: M17 leucyl aminopeptidase

Chain A:



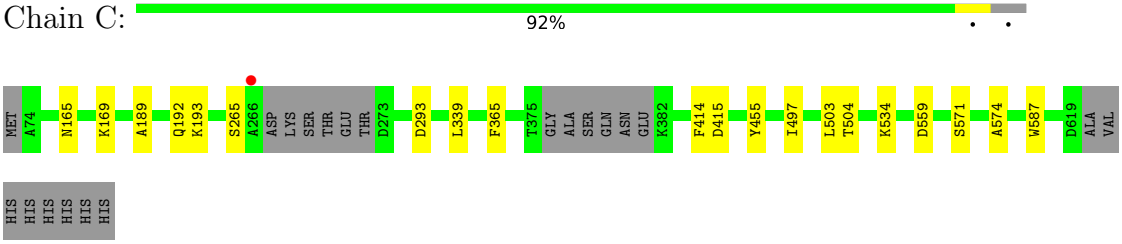
• Molecule 1: M17 leucyl aminopeptidase

Chain B:



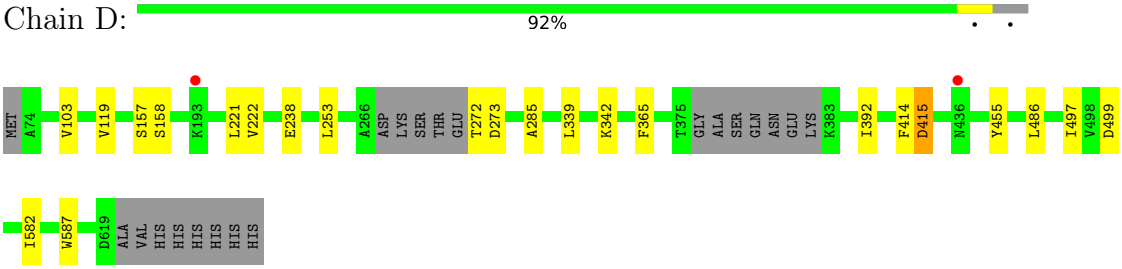
• Molecule 1: M17 leucyl aminopeptidase

Chain C:



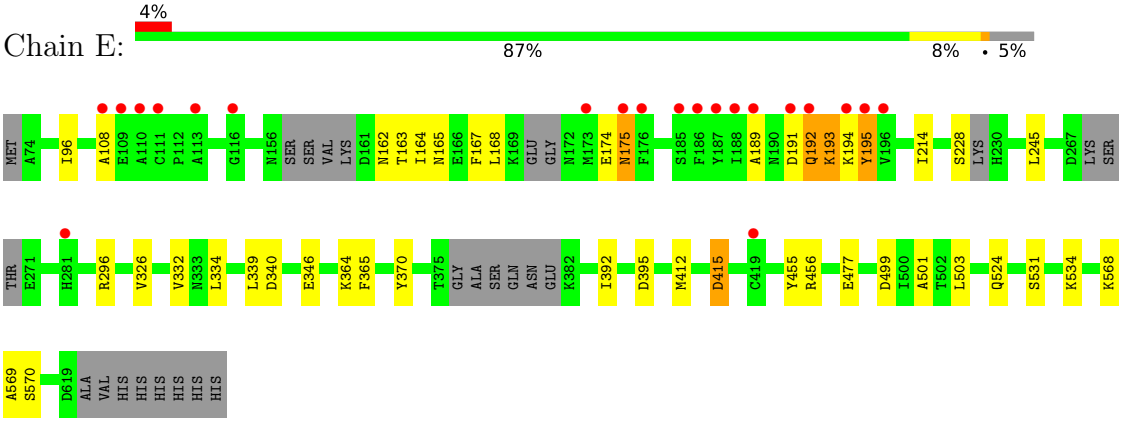
• Molecule 1: M17 leucyl aminopeptidase

Chain D:



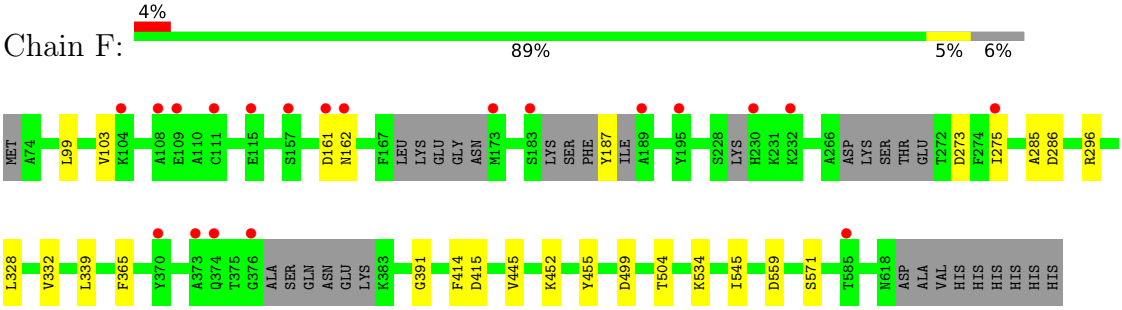
• Molecule 1: M17 leucyl aminopeptidase

Chain E:



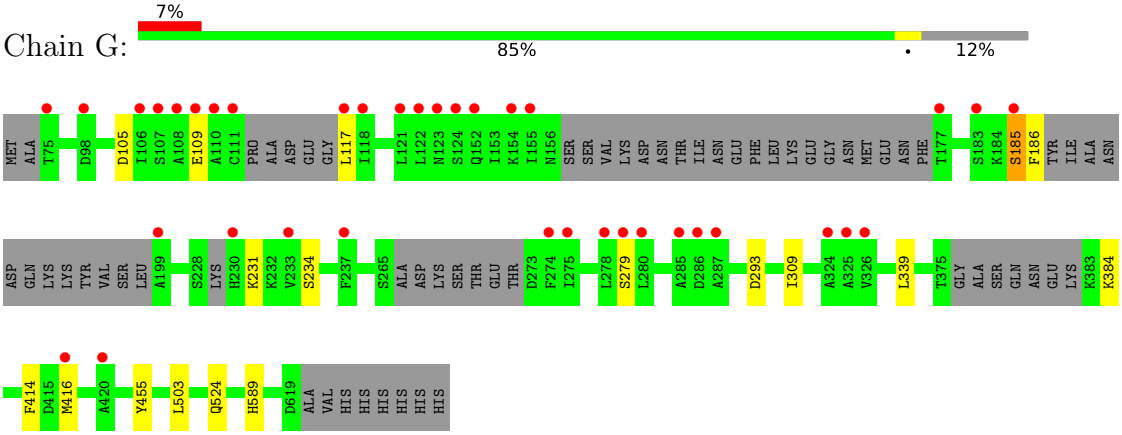
• Molecule 1: M17 leucyl aminopeptidase

Chain F:



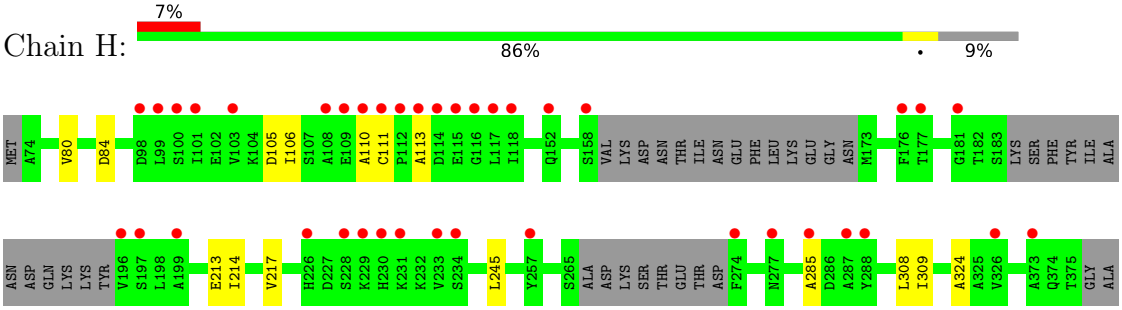
• Molecule 1: M17 leucyl aminopeptidase

Chain G:



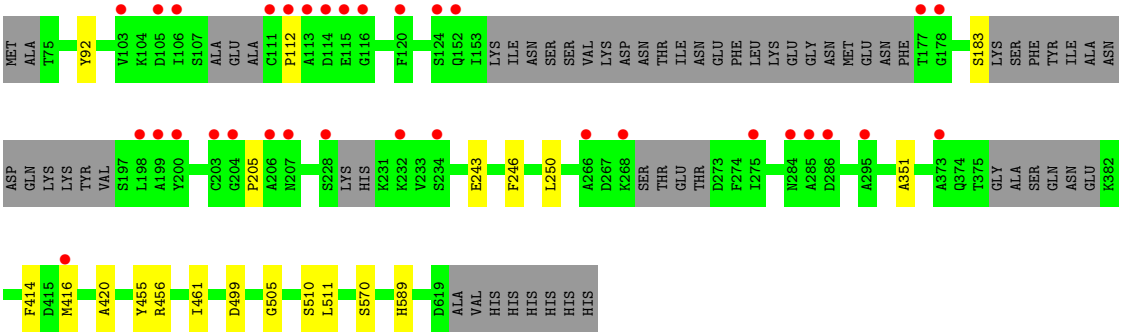
• Molecule 1: M17 leucyl aminopeptidase

Chain H:



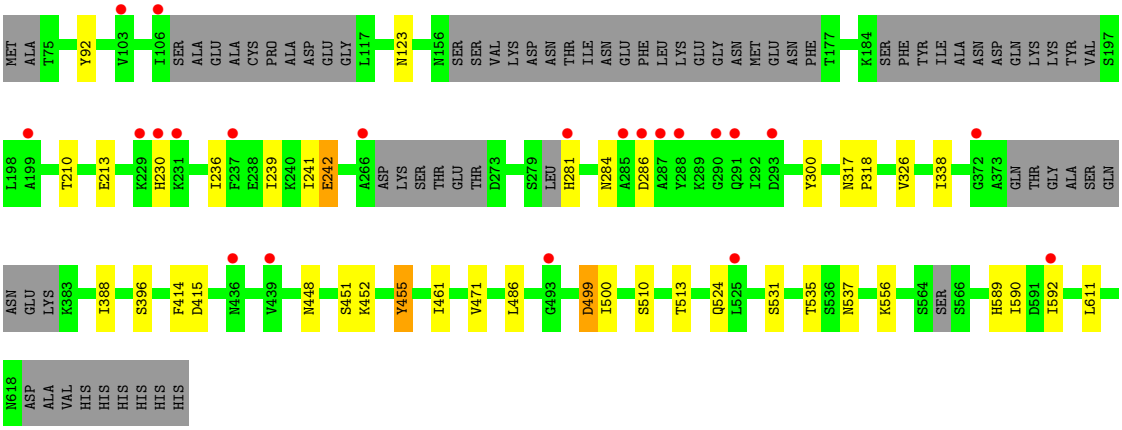
• Molecule 1: M17 leucyl aminopeptidase

Chain I:



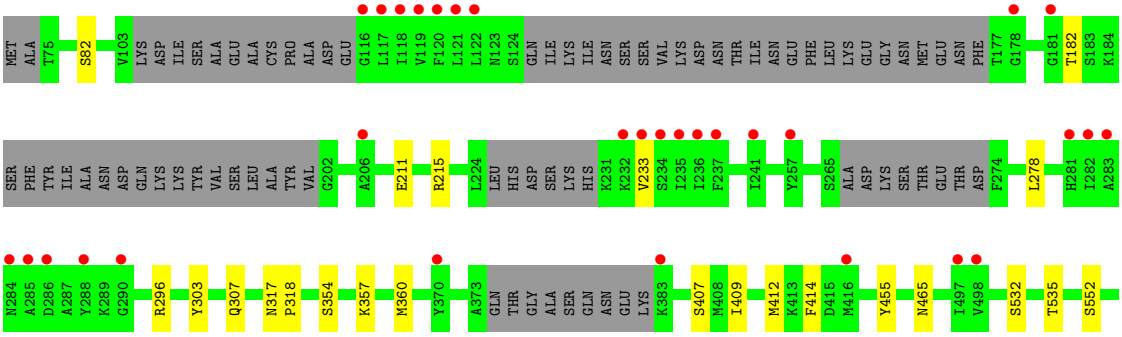
• Molecule 1: M17 leucyl aminopeptidase

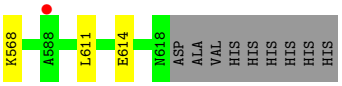
Chain J:



• Molecule 1: M17 leucyl aminopeptidase

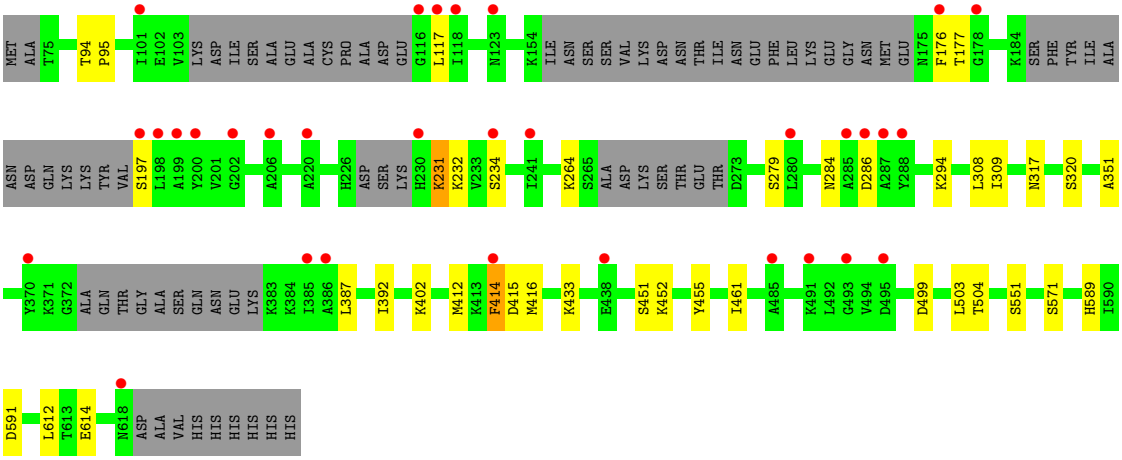
Chain K:





• Molecule 1: M17 leucyl aminopeptidase

Chain L:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	117.03Å 201.55Å 166.22Å 90.00° 106.01° 90.00°	Depositor
Resolution (Å)	44.32 – 2.33 44.32 – 2.33	Depositor EDS
% Data completeness (in resolution range)	99.9 (44.32-2.33) 99.9 (44.32-2.33)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.61 (at 2.34Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.205 , 0.243 0.205 , 0.243	Depositor DCC
R_{free} test set	15702 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	34.1	Xtriage
Anisotropy	0.323	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 52.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	86464	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.90% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: GOL, ZN, PEG, SO4

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	0.26	0/4003	0.43	0/5421
1	B	0.26	0/3996	0.43	0/5412
1	C	0.26	0/3995	0.43	0/5411
1	D	0.26	0/3989	0.43	0/5406
1	E	0.26	0/3783	0.46	0/5137
1	F	0.26	0/3759	0.43	0/5101
1	G	0.26	0/3440	0.43	0/4673
1	H	0.26	0/3554	0.43	0/4830
1	I	0.25	0/3460	0.43	0/4697
1	J	0.25	0/3355	0.42	0/4560
1	K	0.26	0/3207	0.44	0/4356
1	L	0.25	0/3284	0.42	0/4458
All	All	0.26	0/43825	0.43	0/59462

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	E	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	195	TYR	Peptide

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	3927	3920	3924	13	0
1	B	3920	3919	3919	8	0
1	C	3919	3918	3918	9	0
1	D	3910	3896	3896	11	0
1	E	3717	3523	3540	27	0
1	F	3694	3565	3561	11	0
1	G	3382	3161	3162	9	0
1	H	3492	3280	3281	12	0
1	I	3401	3163	3160	10	0
1	J	3301	3027	3039	20	0
1	K	3156	2920	2918	13	0
1	L	3229	2946	2948	23	0
2	A	36	47	48	0	0
2	B	18	24	24	1	0
2	C	12	16	16	0	0
2	D	18	24	24	0	0
2	E	30	40	40	1	0
2	F	18	24	24	1	0
2	G	12	16	16	0	0
2	H	6	8	8	0	0
2	I	12	16	16	2	0
2	J	18	24	24	0	0
3	A	7	10	10	0	0
3	B	14	20	20	1	0
3	C	7	10	10	0	0
3	D	7	10	10	0	0
3	E	21	30	30	1	0
3	F	14	20	20	0	0
3	G	7	10	10	1	0
3	H	7	10	10	0	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0
4	C	2	0	0	0	0
4	D	2	0	0	0	0
4	E	2	0	0	0	0
4	F	2	0	0	0	0
4	G	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	H	2	0	0	0	0
4	I	2	0	0	0	0
4	J	2	0	0	0	0
4	K	2	0	0	0	0
4	L	2	0	0	0	0
5	A	15	0	0	0	0
5	B	15	0	0	0	0
5	C	5	0	0	0	0
5	D	5	0	0	0	0
5	E	5	0	0	0	0
5	F	5	0	0	0	0
5	G	10	0	0	0	0
5	H	5	0	0	0	0
5	I	5	0	0	0	0
5	J	5	0	0	0	0
5	K	5	0	0	0	0
5	L	10	0	0	0	0
6	A	206	0	0	2	0
6	B	182	0	0	3	0
6	C	201	0	0	1	0
6	D	179	0	0	1	0
6	E	154	0	0	3	0
6	F	128	0	0	1	0
6	G	108	0	0	3	0
6	H	83	0	0	1	0
6	I	56	0	0	2	0
6	J	47	0	0	2	0
6	K	53	0	0	1	0
6	L	44	0	0	0	0
All	All	44867	41597	41626	164	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 2.

The worst 5 of 164 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)
3:G:702:PEG:O4	6:G:801:HOH:O	1.98	0.82
1:L:433:LYS:HA	1:L:433:LYS:HE2	1.62	0.79
1:E:534:LYS:NZ	6:E:802:HOH:O	2.17	0.77
1:F:187:TYR:OH	6:F:801:HOH:O	2.02	0.76

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:165:ASN:O	1:E:168:LEU:CB	2.36	0.73

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	503/528 (95%)	490 (97%)	12 (2%)	1 (0%)	47 55
1	B	502/528 (95%)	490 (98%)	12 (2%)	0	100 100
1	C	501/528 (95%)	488 (97%)	13 (3%)	0	100 100
1	D	502/528 (95%)	490 (98%)	12 (2%)	0	100 100
1	E	491/528 (93%)	464 (94%)	19 (4%)	8 (2%)	9 7
1	F	484/528 (92%)	466 (96%)	17 (4%)	1 (0%)	47 55
1	G	452/528 (86%)	435 (96%)	15 (3%)	2 (0%)	34 38
1	H	468/528 (89%)	443 (95%)	23 (5%)	2 (0%)	34 38
1	I	453/528 (86%)	433 (96%)	18 (4%)	2 (0%)	34 38
1	J	442/528 (84%)	420 (95%)	20 (4%)	2 (0%)	29 31
1	K	426/528 (81%)	410 (96%)	15 (4%)	1 (0%)	47 55
1	L	439/528 (83%)	415 (94%)	24 (6%)	0	100 100
All	All	5663/6336 (89%)	5444 (96%)	200 (4%)	19 (0%)	41 47

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	163	THR
1	E	192	GLN
1	F	273	ASP
1	I	112	PRO

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	I	205	PRO

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	425/448 (95%)	419 (99%)	6 (1%)	67	78
1	B	425/448 (95%)	421 (99%)	4 (1%)	78	87
1	C	426/448 (95%)	421 (99%)	5 (1%)	71	82
1	D	423/448 (94%)	418 (99%)	5 (1%)	71	82
1	E	372/448 (83%)	366 (98%)	6 (2%)	62	74
1	F	373/448 (83%)	366 (98%)	7 (2%)	57	68
1	G	325/448 (72%)	322 (99%)	3 (1%)	78	87
1	H	337/448 (75%)	334 (99%)	3 (1%)	78	87
1	I	325/448 (72%)	321 (99%)	4 (1%)	71	82
1	J	315/448 (70%)	308 (98%)	7 (2%)	52	63
1	K	298/448 (66%)	291 (98%)	7 (2%)	50	61
1	L	296/448 (66%)	291 (98%)	5 (2%)	60	72
All	All	4340/5376 (81%)	4278 (99%)	62 (1%)	67	78

5 of 62 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	415	ASP
1	K	465	ASN
1	G	455	TYR
1	K	455	TYR
1	L	414	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	G	207	ASN
1	L	537	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 84 ligands modelled in this entry, 24 are monoatomic - leaving 60 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	SO4	B	710	-	4,4,4	0.14	0	6,6,6	0.10	0
2	GOL	J	702	-	5,5,5	0.82	0	5,5,5	1.07	0
2	GOL	F	702	-	5,5,5	1.01	0	5,5,5	1.05	0
3	PEG	F	703	4	6,6,6	0.51	0	5,5,5	0.41	0
2	GOL	A	704	-	5,5,5	0.83	0	5,5,5	1.09	0
2	GOL	E	704	-	5,5,5	0.85	0	5,5,5	1.06	0
5	SO4	K	703	-	4,4,4	0.13	0	6,6,6	0.15	0
2	GOL	B	705	-	5,5,5	0.87	0	5,5,5	1.05	0
2	GOL	I	701	-	5,5,5	0.91	0	5,5,5	1.07	0
5	SO4	A	710	-	4,4,4	0.16	0	6,6,6	0.07	0
2	GOL	E	705	-	5,5,5	0.80	0	5,5,5	0.97	0
2	GOL	E	703	-	5,5,5	0.73	0	5,5,5	0.95	0
3	PEG	E	706	-	6,6,6	0.48	0	5,5,5	0.38	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
2	GOL	G	701	-	5,5,5	0.81	0	5,5,5	0.95	0
2	GOL	A	706	-	5,5,5	0.83	0	5,5,5	0.96	0
5	SO4	G	707	-	4,4,4	0.13	0	6,6,6	0.12	0
3	PEG	C	702	-	6,6,6	0.48	0	5,5,5	0.37	0
5	SO4	A	711	-	4,4,4	0.13	0	6,6,6	0.17	0
5	SO4	I	705	-	4,4,4	0.17	0	6,6,6	0.18	0
2	GOL	E	702	-	5,5,5	0.84	0	5,5,5	1.02	0
3	PEG	H	702	-	6,6,6	0.49	0	5,5,5	0.30	0
3	PEG	A	703	-	6,6,6	0.48	0	5,5,5	0.42	0
2	GOL	J	701	-	5,5,5	0.88	0	5,5,5	0.98	0
2	GOL	J	703	-	5,5,5	0.90	0	5,5,5	1.02	0
3	PEG	G	702	-	6,6,6	0.49	0	5,5,5	0.27	0
5	SO4	B	709	-	4,4,4	0.17	0	6,6,6	0.14	0
2	GOL	C	703	-	5,5,5	0.85	0	5,5,5	0.98	0
5	SO4	A	712	-	4,4,4	0.13	0	6,6,6	0.07	0
5	SO4	H	705	-	4,4,4	0.13	0	6,6,6	0.08	0
2	GOL	H	701	-	5,5,5	0.85	0	5,5,5	1.05	0
2	GOL	B	702	-	5,5,5	0.87	0	5,5,5	1.06	0
5	SO4	G	706	-	4,4,4	0.17	0	6,6,6	0.19	0
5	SO4	E	711	-	4,4,4	0.15	0	6,6,6	0.13	0
5	SO4	F	708	-	4,4,4	0.15	0	6,6,6	0.19	0
2	GOL	E	701	-	5,5,5	0.84	0	5,5,5	0.97	0
5	SO4	D	707	-	4,4,4	0.16	0	6,6,6	0.10	0
2	GOL	A	707	-	5,5,5	0.85	0	5,5,5	1.03	0
3	PEG	E	707	-	6,6,6	0.48	0	5,5,5	0.26	0
2	GOL	A	702	-	5,5,5	0.88	0	5,5,5	1.07	0
2	GOL	D	704	-	5,5,5	0.89	0	5,5,5	1.00	0
2	GOL	F	701	-	5,5,5	0.88	0	5,5,5	0.99	0
2	GOL	I	702	-	5,5,5	0.86	0	5,5,5	0.99	0
2	GOL	C	701	-	5,5,5	0.82	0	5,5,5	1.02	0
2	GOL	A	701	-	5,5,5	0.87	0	5,5,5	0.98	0
3	PEG	B	704	-	6,6,6	0.49	0	5,5,5	0.44	0
2	GOL	D	703	-	5,5,5	0.81	0	5,5,5	0.93	0
5	SO4	L	704	-	4,4,4	0.14	0	6,6,6	0.13	0
5	SO4	B	708	-	4,4,4	0.15	0	6,6,6	0.32	0
2	GOL	D	701	-	5,5,5	0.87	0	5,5,5	1.00	0
2	GOL	B	701	-	5,5,5	0.97	0	5,5,5	1.12	1 (20%)
2	GOL	F	705	-	5,5,5	0.92	0	5,5,5	0.87	0
3	PEG	F	704	-	6,6,6	0.48	0	5,5,5	0.43	0
3	PEG	D	702	-	6,6,6	0.49	0	5,5,5	0.26	0
2	GOL	G	703	-	5,5,5	0.79	0	5,5,5	1.01	0
3	PEG	E	708	-	6,6,6	0.48	0	5,5,5	0.25	0
5	SO4	J	706	-	4,4,4	0.13	0	6,6,6	0.10	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
3	PEG	B	703	-	6,6,6	0.49	0	5,5,5	0.22	0
2	GOL	A	705	-	5,5,5	0.83	0	5,5,5	0.99	0
5	SO4	L	703	-	4,4,4	0.15	0	6,6,6	0.15	0
5	SO4	C	706	-	4,4,4	0.12	0	6,6,6	0.11	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
2	GOL	J	702	-	-	2/4/4/4	-
2	GOL	F	702	-	-	0/4/4/4	-
3	PEG	F	703	4	-	3/4/4/4	-
2	GOL	A	704	-	-	0/4/4/4	-
2	GOL	E	704	-	-	0/4/4/4	-
2	GOL	B	705	-	-	2/4/4/4	-
2	GOL	I	701	-	-	0/4/4/4	-
2	GOL	E	705	-	-	2/4/4/4	-
2	GOL	E	703	-	-	2/4/4/4	-
3	PEG	E	706	-	-	1/4/4/4	-
2	GOL	G	701	-	-	2/4/4/4	-
2	GOL	A	706	-	-	0/4/4/4	-
3	PEG	C	702	-	-	3/4/4/4	-
2	GOL	E	702	-	-	0/4/4/4	-
3	PEG	H	702	-	-	1/4/4/4	-
3	PEG	A	703	-	-	2/4/4/4	-
2	GOL	J	701	-	-	2/4/4/4	-
2	GOL	J	703	-	-	0/4/4/4	-
3	PEG	G	702	-	-	2/4/4/4	-
2	GOL	C	703	-	-	0/4/4/4	-
2	GOL	H	701	-	-	0/4/4/4	-
2	GOL	B	702	-	-	0/4/4/4	-
2	GOL	E	701	-	-	0/4/4/4	-
2	GOL	A	707	-	-	2/4/4/4	-
3	PEG	E	707	-	-	2/4/4/4	-
2	GOL	A	702	-	-	2/4/4/4	-
2	GOL	D	704	-	-	1/4/4/4	-
2	GOL	F	701	-	-	0/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GOL	I	702	-	-	2/4/4/4	-
2	GOL	C	701	-	-	2/4/4/4	-
2	GOL	A	701	-	-	0/4/4/4	-
3	PEG	B	704	-	-	3/4/4/4	-
2	GOL	D	703	-	-	3/4/4/4	-
2	GOL	D	701	-	-	2/4/4/4	-
2	GOL	B	701	-	-	2/4/4/4	-
2	GOL	F	705	-	-	2/4/4/4	-
3	PEG	F	704	-	-	3/4/4/4	-
3	PEG	D	702	-	-	2/4/4/4	-
2	GOL	G	703	-	-	0/4/4/4	-
3	PEG	E	708	-	-	0/4/4/4	-
3	PEG	B	703	-	-	1/4/4/4	-
2	GOL	A	705	-	-	2/4/4/4	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	701	GOL	C3-C2-C1	-2.04	103.78	111.70

There are no chirality outliers.

5 of 55 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	702	GOL	C1-C2-C3-O3
2	B	701	GOL	C1-C2-C3-O3
2	B	705	GOL	C1-C2-C3-O3
2	E	703	GOL	O1-C1-C2-C3
2	F	705	GOL	C1-C2-C3-O3

There are no ring outliers.

8 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	I	701	GOL	1	0
2	E	705	GOL	1	0
3	G	702	PEG	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	707	PEG	1	0
2	I	702	GOL	1	0
2	B	701	GOL	1	0
2	F	705	GOL	1	0
3	B	703	PEG	1	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^{th} 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^2)	Q<0.9
1	A	509/528 (96%)	0.01	5 (0%) 82 88	22, 30, 52, 64	0
1	B	508/528 (96%)	-0.02	0 100 100	22, 31, 50, 65	0
1	C	507/528 (96%)	-0.08	1 (0%) 95 97	21, 30, 46, 64	0
1	D	507/528 (96%)	-0.01	2 (0%) 92 96	22, 33, 49, 69	0
1	E	503/528 (95%)	0.32	21 (4%) 36 47	23, 34, 72, 85	0
1	F	497/528 (94%)	0.38	20 (4%) 38 49	23, 36, 70, 82	0
1	G	466/528 (88%)	0.53	37 (7%) 12 19	28, 39, 73, 88	0
1	H	478/528 (90%)	0.55	39 (8%) 11 17	31, 42, 73, 89	0
1	I	467/528 (88%)	0.51	33 (7%) 16 23	34, 49, 75, 93	0
1	J	458/528 (86%)	0.46	22 (4%) 30 41	36, 53, 77, 87	0
1	K	440/528 (83%)	0.60	32 (7%) 15 22	35, 52, 74, 83	0
1	L	453/528 (85%)	0.64	32 (7%) 16 23	38, 54, 79, 89	0
All	All	5793/6336 (91%)	0.31	244 (4%) 36 47	21, 40, 72, 93	0

The worst 5 of 244 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	107	SER	6.6
1	K	178	GLY	6.4
1	G	110	ALA	5.7
1	H	110	ALA	5.7
1	E	195	TYR	5.6

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 (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

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. The B-factors column lists the minimum, median, 95^{th} 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	B-factors(Å^2)	Q<0.9
3	PEG	F	704	7/7	0.62	0.43	46,63,76,81	0
2	GOL	J	702	6/6	0.65	0.26	65,83,91,102	0
2	GOL	I	701	6/6	0.68	0.40	62,78,94,94	0
2	GOL	A	701	6/6	0.72	0.35	51,62,71,71	0
3	PEG	E	708	7/7	0.73	0.32	52,63,76,76	0
3	PEG	B	704	7/7	0.74	0.27	46,62,74,77	0
3	PEG	H	702	7/7	0.75	0.19	53,64,74,79	0
3	PEG	B	703	7/7	0.76	0.24	47,57,62,62	0
2	GOL	A	707	6/6	0.77	0.26	58,69,74,74	0
2	GOL	G	701	6/6	0.77	0.32	46,55,68,82	0
3	PEG	E	707	7/7	0.77	0.36	37,47,63,63	0
3	PEG	G	702	7/7	0.78	0.34	56,67,76,76	0
3	PEG	A	703	7/7	0.80	0.22	50,64,76,82	0
2	GOL	E	703	6/6	0.81	0.32	44,55,63,66	0
3	PEG	C	702	7/7	0.82	0.18	40,54,66,66	0
2	GOL	A	705	6/6	0.82	0.28	44,55,67,71	0
2	GOL	F	705	6/6	0.83	0.30	40,52,73,73	0
2	GOL	B	701	6/6	0.83	0.20	37,44,53,57	0
3	PEG	E	706	7/7	0.83	0.19	52,62,72,72	0
2	GOL	A	706	6/6	0.83	0.24	46,63,72,76	0
2	GOL	D	703	6/6	0.84	0.31	59,75,79,93	0
2	GOL	J	703	6/6	0.84	0.20	58,69,80,81	0
4	ZN	L	701	1/1	0.84	0.16	53,53,53,53	1
2	GOL	E	704	6/6	0.85	0.46	44,53,60,62	0
3	PEG	F	703	7/7	0.85	0.20	44,53,61,74	0
2	GOL	I	702	6/6	0.86	0.33	44,57,64,69	0
3	PEG	D	702	7/7	0.88	0.25	47,57,68,68	0
2	GOL	B	702	6/6	0.88	0.21	44,53,64,64	0
2	GOL	F	702	6/6	0.88	0.30	39,53,64,64	0
2	GOL	E	705	6/6	0.89	0.28	30,48,59,67	0
2	GOL	H	701	6/6	0.90	0.23	47,56,64,67	0
2	GOL	C	701	6/6	0.90	0.17	42,51,61,71	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	GOL	G	703	6/6	0.90	0.16	37,53,57,65	0
4	ZN	I	703	1/1	0.91	0.12	46,46,46,46	1
2	GOL	E	702	6/6	0.91	0.30	39,53,63,63	0
2	GOL	D	701	6/6	0.92	0.36	44,57,64,68	0
2	GOL	J	701	6/6	0.92	0.25	43,52,59,60	0
2	GOL	F	701	6/6	0.92	0.17	33,45,51,55	0
2	GOL	A	702	6/6	0.94	0.16	40,51,57,61	0
2	GOL	B	705	6/6	0.94	0.19	34,41,48,56	0
4	ZN	A	708	1/1	0.94	0.12	34,34,34,34	1
4	ZN	D	706	1/1	0.94	0.12	30,30,30,30	1
4	ZN	E	709	1/1	0.94	0.13	32,32,32,32	1
4	ZN	H	704	1/1	0.94	0.19	44,44,44,44	1
2	GOL	E	701	6/6	0.94	0.26	35,44,51,54	0
4	ZN	K	702	1/1	0.94	0.10	52,52,52,52	1
2	GOL	A	704	6/6	0.94	0.16	34,42,50,60	0
5	SO4	L	703	5/5	0.94	0.16	45,46,49,50	5
4	ZN	F	707	1/1	0.95	0.10	31,31,31,31	1
2	GOL	D	704	6/6	0.95	0.16	35,42,47,47	0
2	GOL	C	703	6/6	0.95	0.33	38,46,50,57	0
4	ZN	K	701	1/1	0.96	0.09	48,48,48,48	1
4	ZN	G	704	1/1	0.96	0.19	38,38,38,38	1
4	ZN	D	705	1/1	0.97	0.12	44,44,44,44	1
4	ZN	I	704	1/1	0.97	0.10	46,46,46,46	1
4	ZN	J	705	1/1	0.97	0.10	42,42,42,42	1
4	ZN	A	709	1/1	0.97	0.12	37,37,37,37	1
4	ZN	G	705	1/1	0.97	0.15	52,52,52,52	1
4	ZN	H	703	1/1	0.97	0.12	41,41,41,41	1
5	SO4	A	712	5/5	0.97	0.37	57,62,72,75	0
5	SO4	B	710	5/5	0.97	0.21	56,58,65,67	0
4	ZN	B	707	1/1	0.97	0.07	33,33,33,33	1
4	ZN	J	704	1/1	0.98	0.13	46,46,46,46	1
4	ZN	L	702	1/1	0.98	0.13	56,56,56,56	1
4	ZN	E	710	1/1	0.98	0.09	43,43,43,43	1
5	SO4	B	708	5/5	0.98	0.15	28,31,32,34	0
4	ZN	F	706	1/1	0.98	0.13	36,36,36,36	1
5	SO4	F	708	5/5	0.98	0.16	35,35,36,38	0
5	SO4	G	706	5/5	0.98	0.22	30,33,36,42	5
5	SO4	K	703	5/5	0.98	0.12	48,50,52,56	0
4	ZN	B	706	1/1	0.98	0.10	38,38,38,38	1
5	SO4	L	704	5/5	0.98	0.15	37,41,42,46	0
5	SO4	C	706	5/5	0.99	0.08	28,30,38,38	0
5	SO4	D	707	5/5	0.99	0.12	31,34,37,37	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	SO4	E	711	5/5	0.99	0.16	32,33,38,39	0
5	SO4	A	711	5/5	0.99	0.13	23,27,29,31	0
4	ZN	C	704	1/1	0.99	0.09	28,28,28,28	1
5	SO4	G	707	5/5	0.99	0.12	31,32,35,35	0
5	SO4	H	705	5/5	0.99	0.20	36,37,39,43	5
5	SO4	I	705	5/5	0.99	0.16	34,39,42,49	5
5	SO4	J	706	5/5	0.99	0.17	39,45,46,49	5
4	ZN	C	705	1/1	0.99	0.08	47,47,47,47	1
5	SO4	B	709	5/5	0.99	0.15	25,28,28,29	0
5	SO4	A	710	5/5	0.99	0.14	32,32,33,35	0

6.5 Other polymers (i)

There are no such residues in this entry.