



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 16, 2020 – 12:22 am BST

PDB ID : 6ECF  
Title : Vlm2 thioesterase domain with genetically encoded 2,3-diaminopropionic acid bound with a dodecadepsipeptide, space group P1  
Authors : Alonzo, D.A.; Huguenin-Dezot, N.; Heberlig, G.W.; Mahesh, M.; Nguyen, D.P.; Dornan, M.H.; Boddy, C.N.; Chin, J.W.; Schmeing, T.M.  
Deposited on : 2018-08-07  
Resolution : 2.50 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](mailto: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) : 1.13  
EDS : 2.11  
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) : 2.11

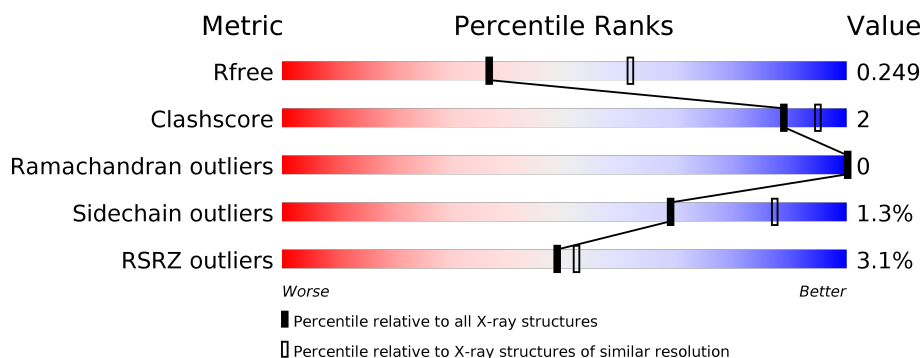
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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))
$R_{free}$	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 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  $\leq 5\%$ . 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	303	<div> <div>0%</div> <div> <div></div> <div>81%</div> <div>17%</div> </div> </div>
1	B	303	<div> <div>3%</div> <div> <div></div> <div>77%</div> <div>6%</div> <div>17%</div> </div> </div>
1	C	303	<div> <div>2%</div> <div> <div></div> <div>79%</div> <div>5%</div> <div>17%</div> </div> </div>
1	D	303	<div> <div>3%</div> <div> <div></div> <div>84%</div> <div>13%</div> </div> </div>
1	E	303	<div> <div>2%</div> <div> <div></div> <div>80%</div> <div>6%</div> <div>14%</div> </div> </div>
1	F	303	<div> <div>4%</div> <div> <div></div> <div>82%</div> <div>15%</div> </div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	G	12	 17% 17% 67%
2	H	12	 17% 83%
2	I	12	 25% 75%
2	J	12	 8% 8% 83%
2	K	12	 17% 83%
2	L	12	 8% 8% 83%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 23003 atoms, of which 11242 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 Vlm2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	251	Total	C	H	N	O	S	0	0	0
			3675	1201	1805	324	337	8			
1	B	252	Total	C	H	N	O	S	0	0	0
			3738	1212	1847	330	341	8			
1	C	252	Total	C	H	N	O	S	0	0	0
			3745	1215	1854	332	336	8			
1	D	265	Total	C	H	N	O	S	0	0	0
			3863	1257	1905	340	352	9			
1	E	262	Total	C	H	N	O	S	0	0	0
			3758	1235	1839	326	349	9			
1	F	258	Total	C	H	N	O	S	0	2	0
			3821	1246	1888	333	345	9			

There are 90 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2353	MET	-	expression tag	UNP Q1PSF3
A	2354	HIS	-	expression tag	UNP Q1PSF3
A	2355	HIS	-	expression tag	UNP Q1PSF3
A	2356	HIS	-	expression tag	UNP Q1PSF3
A	2357	HIS	-	expression tag	UNP Q1PSF3
A	2358	HIS	-	expression tag	UNP Q1PSF3
A	2359	HIS	-	expression tag	UNP Q1PSF3
A	2360	HIS	-	expression tag	UNP Q1PSF3
A	2361	HIS	-	expression tag	UNP Q1PSF3
A	2362	GLU	-	expression tag	UNP Q1PSF3
A	2363	ASN	-	expression tag	UNP Q1PSF3
A	2364	LEU	-	expression tag	UNP Q1PSF3
A	2365	TYR	-	expression tag	UNP Q1PSF3
A	2366	PHE	-	expression tag	UNP Q1PSF3
A	2367	GLN	-	expression tag	UNP Q1PSF3
B	2353	MET	-	expression tag	UNP Q1PSF3
B	2354	HIS	-	expression tag	UNP Q1PSF3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	2355	HIS	-	expression tag	UNP Q1PSF3
B	2356	HIS	-	expression tag	UNP Q1PSF3
B	2357	HIS	-	expression tag	UNP Q1PSF3
B	2358	HIS	-	expression tag	UNP Q1PSF3
B	2359	HIS	-	expression tag	UNP Q1PSF3
B	2360	HIS	-	expression tag	UNP Q1PSF3
B	2361	HIS	-	expression tag	UNP Q1PSF3
B	2362	GLU	-	expression tag	UNP Q1PSF3
B	2363	ASN	-	expression tag	UNP Q1PSF3
B	2364	LEU	-	expression tag	UNP Q1PSF3
B	2365	TYR	-	expression tag	UNP Q1PSF3
B	2366	PHE	-	expression tag	UNP Q1PSF3
B	2367	GLN	-	expression tag	UNP Q1PSF3
C	2353	MET	-	expression tag	UNP Q1PSF3
C	2354	HIS	-	expression tag	UNP Q1PSF3
C	2355	HIS	-	expression tag	UNP Q1PSF3
C	2356	HIS	-	expression tag	UNP Q1PSF3
C	2357	HIS	-	expression tag	UNP Q1PSF3
C	2358	HIS	-	expression tag	UNP Q1PSF3
C	2359	HIS	-	expression tag	UNP Q1PSF3
C	2360	HIS	-	expression tag	UNP Q1PSF3
C	2361	HIS	-	expression tag	UNP Q1PSF3
C	2362	GLU	-	expression tag	UNP Q1PSF3
C	2363	ASN	-	expression tag	UNP Q1PSF3
C	2364	LEU	-	expression tag	UNP Q1PSF3
C	2365	TYR	-	expression tag	UNP Q1PSF3
C	2366	PHE	-	expression tag	UNP Q1PSF3
C	2367	GLN	-	expression tag	UNP Q1PSF3
D	2353	MET	-	expression tag	UNP Q1PSF3
D	2354	HIS	-	expression tag	UNP Q1PSF3
D	2355	HIS	-	expression tag	UNP Q1PSF3
D	2356	HIS	-	expression tag	UNP Q1PSF3
D	2357	HIS	-	expression tag	UNP Q1PSF3
D	2358	HIS	-	expression tag	UNP Q1PSF3
D	2359	HIS	-	expression tag	UNP Q1PSF3
D	2360	HIS	-	expression tag	UNP Q1PSF3
D	2361	HIS	-	expression tag	UNP Q1PSF3
D	2362	GLU	-	expression tag	UNP Q1PSF3
D	2363	ASN	-	expression tag	UNP Q1PSF3
D	2364	LEU	-	expression tag	UNP Q1PSF3
D	2365	TYR	-	expression tag	UNP Q1PSF3
D	2366	PHE	-	expression tag	UNP Q1PSF3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	2367	GLN	-	expression tag	UNP Q1PSF3
E	2353	MET	-	expression tag	UNP Q1PSF3
E	2354	HIS	-	expression tag	UNP Q1PSF3
E	2355	HIS	-	expression tag	UNP Q1PSF3
E	2356	HIS	-	expression tag	UNP Q1PSF3
E	2357	HIS	-	expression tag	UNP Q1PSF3
E	2358	HIS	-	expression tag	UNP Q1PSF3
E	2359	HIS	-	expression tag	UNP Q1PSF3
E	2360	HIS	-	expression tag	UNP Q1PSF3
E	2361	HIS	-	expression tag	UNP Q1PSF3
E	2362	GLU	-	expression tag	UNP Q1PSF3
E	2363	ASN	-	expression tag	UNP Q1PSF3
E	2364	LEU	-	expression tag	UNP Q1PSF3
E	2365	TYR	-	expression tag	UNP Q1PSF3
E	2366	PHE	-	expression tag	UNP Q1PSF3
E	2367	GLN	-	expression tag	UNP Q1PSF3
F	2353	MET	-	expression tag	UNP Q1PSF3
F	2354	HIS	-	expression tag	UNP Q1PSF3
F	2355	HIS	-	expression tag	UNP Q1PSF3
F	2356	HIS	-	expression tag	UNP Q1PSF3
F	2357	HIS	-	expression tag	UNP Q1PSF3
F	2358	HIS	-	expression tag	UNP Q1PSF3
F	2359	HIS	-	expression tag	UNP Q1PSF3
F	2360	HIS	-	expression tag	UNP Q1PSF3
F	2361	HIS	-	expression tag	UNP Q1PSF3
F	2362	GLU	-	expression tag	UNP Q1PSF3
F	2363	ASN	-	expression tag	UNP Q1PSF3
F	2364	LEU	-	expression tag	UNP Q1PSF3
F	2365	TYR	-	expression tag	UNP Q1PSF3
F	2366	PHE	-	expression tag	UNP Q1PSF3
F	2367	GLN	-	expression tag	UNP Q1PSF3

- Molecule 2 is a protein called dodecadepsipeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	G	4	Total 56	C 18	H 30	N 2	O 6	0	0	0
2	I	3	Total 41	C 13	H 22	N 2	O 4	0	0	0
2	H	2	Total 25	C 8	H 13	N 1	O 3	0	0	0
2	K	2	Total 25	C 8	H 13	N 1	O 3	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	J	2	Total	C	H	N	O	0	0	0
			25	8	13	1	3			
2	L	2	Total	C	H	N	O	0	0	0
			25	8	13	1	3			

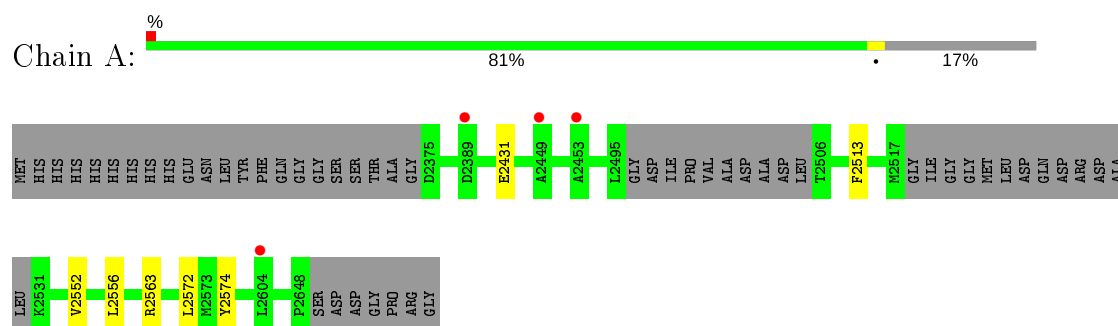
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	50	Total	O	0	0
			50	50		
3	B	23	Total	O	0	0
			23	23		
3	C	35	Total	O	0	0
			35	35		
3	D	30	Total	O	0	0
			30	30		
3	E	34	Total	O	0	0
			34	34		
3	F	34	Total	O	0	0
			34	34		

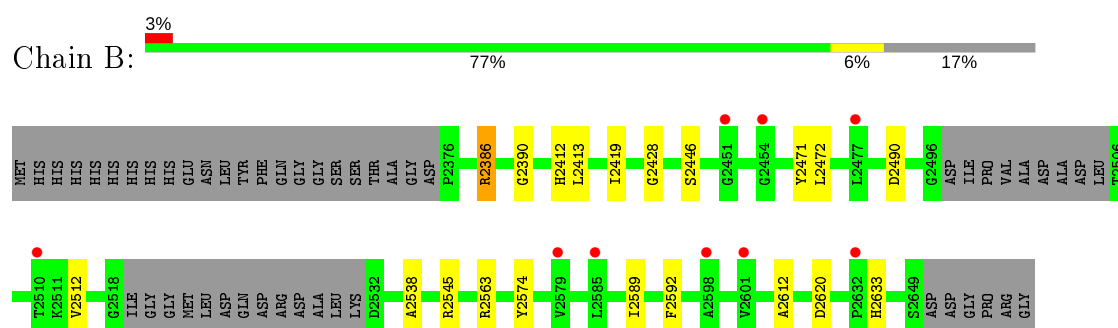
### 3 Residue-property plots [i](#)

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.

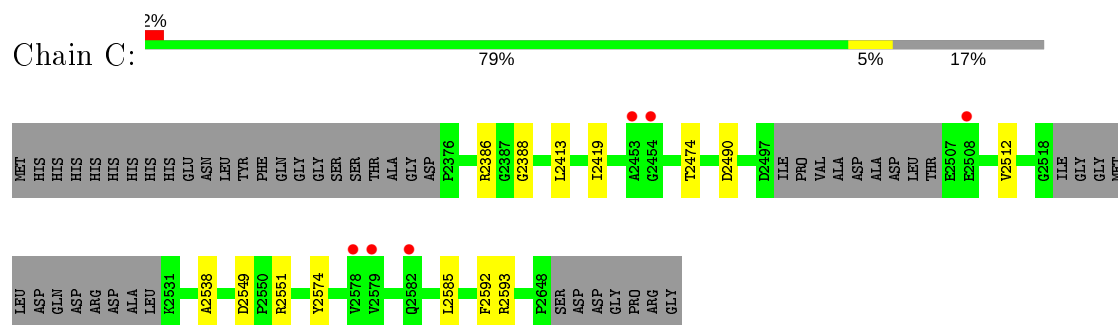
#### • Molecule 1: Vlm2



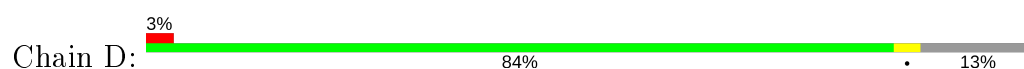
#### • Molecule 1: Vlm2



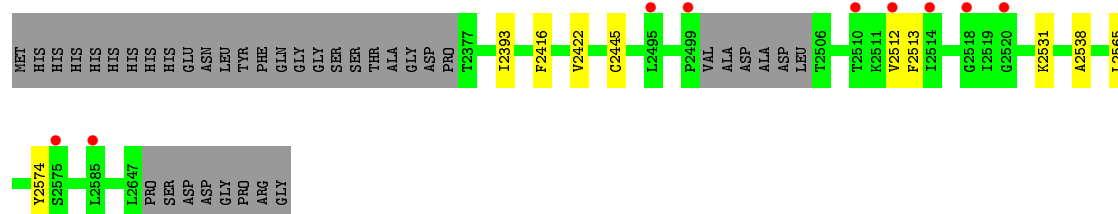
#### • Molecule 1: Vlm2



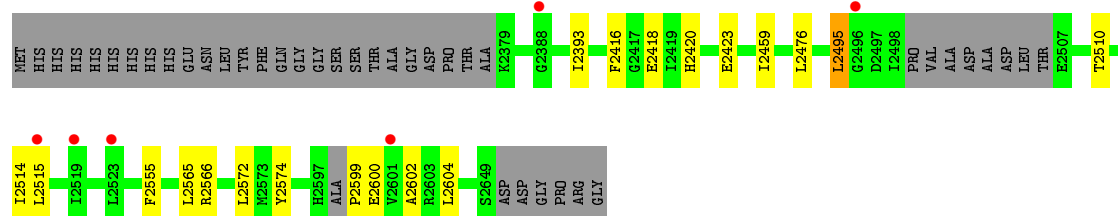
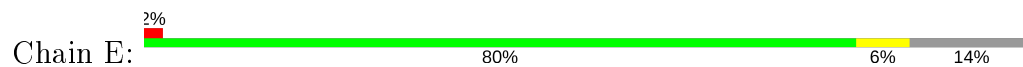
#### • Molecule 1: Vlm2



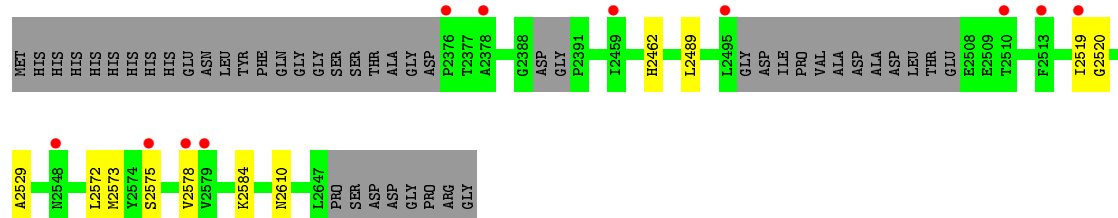
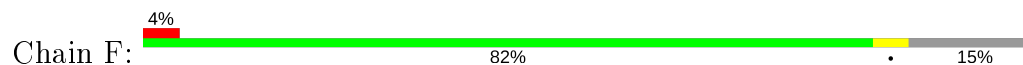




- Molecule 1: Vlm2



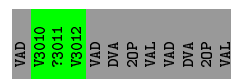
- Molecule 1: Vlm2



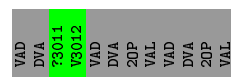
- Molecule 2: dodecadepsipeptide



- Molecule 2: dodecadepsipeptide



- Molecule 2: dodecadepsipeptide



- Molecule 2: dodecadepsiptide

Chain K:  17% 83%

VAD	DVA	V3011	VAD	DVA	20P	VAL	VAD	DVA	20P	VAL

- Molecule 2: dodecadepsiptide

Chain J:  8% 8% 83%

VAD	DVA	V3011	VAD	DVA	20P	VAL	VAD	DVA	20P	VAL

- Molecule 2: dodecadepsiptide

Chain L:  8% 8% 83%

VAD	DVA	V3011	VAD	DVA	20P	VAL	VAD	DVA	20P	VAL

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	77.00Å 77.13Å 90.33Å 91.77° 114.89° 117.94°	Depositor
Resolution (Å)	78.55 – 2.50 78.56 – 2.50	Depositor EDS
% Data completeness (in resolution range)	97.6 (78.55-2.50) 92.5 (78.56-2.50)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.77 (at 2.51Å)	Xtriage
Refinement program	PHENIX 1.13_2998	Depositor
R, $R_{free}$	0.197 , 0.249 0.197 , 0.249	Depositor DCC
$R_{free}$ test set	2683 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	37.6	Xtriage
Anisotropy	0.354	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 40.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.076 for k,-h-k,h+l 0.076 for -h-k,h,h+k+l 0.031 for h,-h-k,-h-l 0.023 for -h-k,k,-l 0.035 for k,h,-h-k-l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	23003	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.79% of the height of the origin peak. No significant pseudotranslation is detected.*

<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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: 2OP, DVA, VAD, DPP

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	0.25	0/1910	0.41	0/2600
1	B	0.25	0/1931	0.41	0/2623
1	C	0.25	0/1931	0.42	0/2622
1	D	0.26	0/1998	0.42	0/2717
1	E	0.25	0/1958	0.41	0/2664
1	F	0.25	0/1979	0.41	0/2687
2	G	0.26	0/6	0.77	0/7
2	H	0.34	0/6	0.33	0/7
2	I	0.23	0/6	0.33	0/7
2	J	0.25	0/6	0.42	0/7
2	K	0.20	0/6	0.50	0/7
2	L	0.24	0/6	0.40	0/7
All	All	0.25	0/11743	0.41	0/15955

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	1870	1805	1803	5	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1891	1847	1846	10	0
1	C	1891	1854	1852	6	0
1	D	1958	1905	1902	5	0
1	E	1919	1839	1836	11	0
1	F	1933	1888	1881	7	0
2	G	26	30	30	3	0
2	H	12	13	14	0	0
2	I	19	22	23	0	0
2	J	12	13	14	2	0
2	K	12	13	14	0	0
2	L	12	13	14	2	0
3	A	50	0	0	0	0
3	B	23	0	0	3	0
3	C	35	0	0	0	0
3	D	30	0	0	0	0
3	E	34	0	0	1	0
3	F	34	0	0	0	0
All	All	11761	11242	11229	44	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.

All (44) 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:2572:LEU:HD21	2:J:3012:VAL:HG21	1.58	0.83
1:F:2572:LEU:HD21	2:L:3012:VAL:HG21	1.65	0.77
1:C:2474:THR:HG22	1:C:2585:LEU:HD21	1.69	0.73
1:F:2584:LYS:NZ	1:F:2610:ASN:O	2.24	0.70
1:E:2572:LEU:CD2	2:J:3012:VAL:HG21	2.22	0.69
1:A:2572:LEU:HD21	2:G:3012:VAL:HG21	1.79	0.64
1:A:2431:GLU:O	1:A:2563:ARG:NH2	2.34	0.61
1:D:2393:ILE:HD11	1:D:2416:PHE:CD2	2.40	0.57
1:E:2599:PRO:HA	1:E:2602:ALA:HB3	1.89	0.55
1:D:2512:VAL:HG23	1:D:2538:ALA:HB1	1.90	0.54
1:F:2572:LEU:CD2	2:L:3012:VAL:HG21	2.35	0.54
1:F:2575:SER:O	1:F:2578:VAL:HG12	2.09	0.53
1:B:2633:HIS:ND1	3:B:2703:HOH:O	2.34	0.52
1:B:2413:LEU:HD12	1:B:2419:ILE:HD13	1.92	0.51
1:B:2512:VAL:HG13	1:B:2538:ALA:HB1	1.93	0.51
1:E:2600:GLU:O	1:E:2604:LEU:HD13	2.11	0.51

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:2519:ILE:HG23	1:F:2529:ALA:HB3	1.93	0.50
1:C:2386:ARG:NH2	1:C:2388:GLY:O	2.46	0.49
1:B:2412:HIS:O	3:B:2701:HOH:O	2.20	0.48
1:D:2393:ILE:HD11	1:D:2416:PHE:CE2	2.49	0.48
1:A:2572:LEU:CD2	2:G:3012:VAL:HG21	2.43	0.47
1:E:2514:ILE:HG23	1:E:2515:LEU:HD12	1.97	0.47
1:E:2510:THR:O	1:E:2514:ILE:HG22	2.14	0.47
1:A:2513:PHE:CZ	2:G:3010:DVA:N	2.83	0.47
1:C:2549:ASP:OD2	1:C:2551:ARG:NH1	2.48	0.46
1:C:2413:LEU:HD12	1:C:2419:ILE:HD13	1.98	0.46
1:E:2495:LEU:H	1:E:2495:LEU:HD12	1.81	0.45
1:D:2512:VAL:HG13	1:D:2513:PHE:H	1.81	0.45
1:E:2459:ILE:HD11	1:E:2476:LEU:HD12	1.98	0.45
1:E:2393:ILE:HD11	1:E:2416:PHE:CE2	2.52	0.45
1:E:2555:PHE:O	3:E:2701:HOH:O	2.21	0.44
1:B:2446:SER:HA	1:B:2472:LEU:HD11	1.99	0.44
1:A:2552:VAL:HG13	1:A:2556:LEU:HD23	1.99	0.44
1:B:2386:ARG:NH1	1:B:2390:GLY:O	2.51	0.44
1:B:2620:ASP:OD2	3:B:2702:HOH:O	2.21	0.44
1:B:2428:GLY:O	1:B:2563:ARG:NH1	2.51	0.44
1:F:2520:GLY:N	1:F:2573:MET:O	2.49	0.44
1:C:2512:VAL:HG13	1:C:2538:ALA:HB1	2.00	0.43
1:B:2490:ASP:N	1:B:2592:PHE:O	2.50	0.43
1:F:2462:HIS:HA	1:F:2489:LEU:HB2	1.99	0.43
1:D:2422:VAL:HG21	1:D:2445:CYS:SG	2.59	0.43
1:E:2418:GLU:HB3	1:E:2420:HIS:CE1	2.55	0.41
1:B:2589:ILE:HD12	1:B:2612:ALA:HB2	2.03	0.41
1:C:2490:ASP:N	1:C:2592:PHE:O	2.52	0.41

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	244/303 (80%)	237 (97%)	7 (3%)	0	100	100
1	B	245/303 (81%)	239 (98%)	6 (2%)	0	100	100
1	C	245/303 (81%)	238 (97%)	7 (3%)	0	100	100
1	D	260/303 (86%)	250 (96%)	10 (4%)	0	100	100
1	E	255/303 (84%)	248 (97%)	7 (3%)	0	100	100
1	F	253/303 (84%)	241 (95%)	12 (5%)	0	100	100
All	All	1502/1818 (83%)	1453 (97%)	49 (3%)	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	185/241 (77%)	184 (100%)	1 (0%)	88	96
1	B	190/241 (79%)	186 (98%)	4 (2%)	53	78
1	C	188/241 (78%)	186 (99%)	2 (1%)	73	89
1	D	192/241 (80%)	189 (98%)	3 (2%)	62	84
1	E	187/241 (78%)	182 (97%)	5 (3%)	44	71
1	F	192/241 (80%)	192 (100%)	0	100	100
2	G	1/3 (33%)	1 (100%)	0	100	100
2	H	1/3 (33%)	1 (100%)	0	100	100
2	I	1/3 (33%)	1 (100%)	0	100	100
2	J	1/3 (33%)	1 (100%)	0	100	100
2	K	1/3 (33%)	1 (100%)	0	100	100
2	L	1/3 (33%)	1 (100%)	0	100	100
All	All	1140/1464 (78%)	1125 (99%)	15 (1%)	69	87

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2574	TYR
1	B	2386	ARG
1	B	2471	TYR
1	B	2545	ARG
1	B	2574	TYR
1	C	2574	TYR
1	C	2593	ARG
1	D	2531	LYS
1	D	2565	LEU
1	D	2574	TYR
1	E	2423	GLU
1	E	2495	LEU
1	E	2565	LEU
1	E	2566	ARG
1	E	2574	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	E	2420	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

9 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	DPP	B	2463	1,2	3,5,6	0.98	0	1,5,7	0.39	0
1	DPP	D	2463	1,2	3,5,6	1.02	0	1,5,7	0.21	0
1	DPP	F	2463	1,2	3,5,6	0.97	0	1,5,7	0.06	0



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
1	DPP	C	2463	1,2	3,5,6	0.91	0	1,5,7	0.39	0
1	DPP	A	2463	1,2	3,5,6	0.92	0	1,5,7	0.32	0
2	VAD	G	3009	2	5,6,7	0.48	0	6,7,9	1.51	1 (16%)
1	DPP	E	2463	1,2	3,5,6	1.05	0	1,5,7	0.57	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
1	DPP	B	2463	1,2	-	2/2/4/6	-
1	DPP	D	2463	1,2	-	1/2/4/6	-
1	DPP	F	2463	1,2	-	2/2/4/6	-
1	DPP	C	2463	1,2	-	0/2/4/6	-
1	DPP	A	2463	1,2	-	1/2/4/6	-
2	VAD	G	3009	2	-	5/5/6/8	-
1	DPP	E	2463	1,2	-	2/2/4/6	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	G	3009	VAD	CG1-CB-CA	2.03	114.54	111.33

There are no chirality outliers.

All (13) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	B	2463	DPP	N-CA-CB-NG
1	F	2463	DPP	N-CA-CB-NG
1	F	2463	DPP	C-CA-CB-NG
1	A	2463	DPP	N-CA-CB-NG
2	G	3009	VAD	C-CA-CB-CG1
2	G	3009	VAD	C-CA-CB-CG2
2	G	3009	VAD	OE-CA-CB-CG1
2	G	3009	VAD	OE-CA-CB-CG2
1	E	2463	DPP	N-CA-CB-NG
1	B	2463	DPP	C-CA-CB-NG
1	E	2463	DPP	C-CA-CB-NG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
2	G	3009	VAD	O-C-CA-CB
1	D	2463	DPP	N-CA-CB-NG

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 [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	250/303 (82%)	0.08	4 (1%) 72 74	25, 39, 58, 68	0
1	B	251/303 (82%)	0.14	9 (3%) 42 46	30, 43, 64, 80	0
1	C	251/303 (82%)	0.15	6 (2%) 59 62	28, 41, 63, 79	0
1	D	264/303 (87%)	0.16	9 (3%) 45 48	26, 41, 66, 92	0
1	E	261/303 (86%)	0.09	6 (2%) 60 63	30, 44, 74, 83	0
1	F	257/303 (84%)	0.18	11 (4%) 35 38	28, 42, 70, 86	0
2	G	1/12 (8%)	1.06	0 100 100	45, 45, 45, 45	0
2	H	1/12 (8%)	0.97	0 100 100	51, 51, 51, 51	0
2	I	1/12 (8%)	-0.42	0 100 100	46, 46, 46, 46	0
2	J	1/12 (8%)	3.50	1 (100%) 0 0	60, 60, 60, 60	0
2	K	1/12 (8%)	0.54	0 100 100	51, 51, 51, 51	0
2	L	1/12 (8%)	2.33	1 (100%) 0 0	62, 62, 62, 62	0
All	All	1540/1890 (81%)	0.14	47 (3%) 49 52	25, 41, 67, 92	0

All (47) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	2579	VAL	5.6
1	C	2454	GLY	4.4
1	F	2578	VAL	4.4
1	E	2496	GLY	4.2
1	B	2579	VAL	4.1
1	C	2453	ALA	4.0
1	C	2508	GLU	3.6
2	J	3012	VAL	3.5
1	F	2510	THR	3.5
1	D	2499	PRO	3.3
1	B	2451	GLY	3.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	2510	THR	3.1
1	C	2578	VAL	3.1
1	A	2453	ALA	2.9
1	B	2477	LEU	2.9
1	E	2523	LEU	2.9
1	E	2388	GLY	2.9
1	F	2519	ILE	2.8
1	A	2389	ASP	2.8
1	A	2604	LEU	2.8
1	B	2632	PRO	2.6
1	D	2512	VAL	2.5
1	F	2513	PHE	2.5
1	D	2575	SER	2.5
1	C	2579	VAL	2.5
1	E	2515	LEU	2.4
1	B	2598	ALA	2.4
1	F	2495	LEU	2.4
1	C	2582	GLN	2.3
1	B	2510	THR	2.3
2	L	3012	VAL	2.3
1	F	2376	PRO	2.3
1	D	2518	GLY	2.2
1	F	2378	ALA	2.2
1	D	2514	ILE	2.2
1	F	2459	ILE	2.2
1	D	2495	LEU	2.2
1	A	2449	ALA	2.1
1	D	2585	LEU	2.1
1	B	2454	GLY	2.1
1	B	2601	VAL	2.1
1	E	2601	VAL	2.1
1	F	2548	ASN	2.0
1	D	2520	GLY	2.0
1	F	2575	SER	2.0
1	B	2585	LEU	2.0
1	E	2519	ILE	2.0

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

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<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	B-factors( $\text{\AA}^2$ )	Q<0.9
2	DVA	I	3010	7/8	0.61	0.26	68,84,91,97	0
2	VAD	G	3009	7/8	0.62	0.24	69,83,93,94	0
2	DVA	G	3010	7/8	0.89	0.26	63,69,83,83	0
1	DPP	F	2463	6/7	0.92	0.16	32,39,48,54	0
1	DPP	E	2463	6/7	0.94	0.13	32,38,48,51	0
1	DPP	A	2463	6/7	0.94	0.15	29,33,39,45	0
1	DPP	D	2463	6/7	0.96	0.18	29,35,42,48	0
1	DPP	C	2463	6/7	0.96	0.12	32,36,45,54	0
1	DPP	B	2463	6/7	0.97	0.10	33,36,43,45	0

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.