



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 15, 2017 – 04:29 am GMT

PDB ID : 4E1O  
Title : Human histidine decarboxylase complex with Histidine methyl ester (HME)  
Authors : Komori, H.; Nitta, Y.; Ueno, H.; Higuchi, Y.  
Deposited on : 2012-03-06  
Resolution : 1.80 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](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

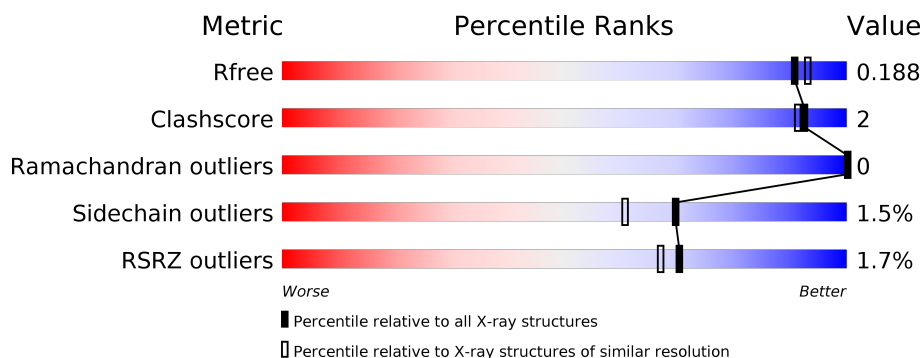
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	100719	4827 (1.80-1.80)
Clashscore	112137	5742 (1.80-1.80)
Ramachandran outliers	110173	5676 (1.80-1.80)
Sidechain outliers	110143	5675 (1.80-1.80)
RSRZ outliers	101464	4906 (1.80-1.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments 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	481	<div> <div>2%</div> <div> <div></div> <div>94%</div> <div>5%</div> </div> </div>
1	B	481	<div> <div></div> <div>95%</div> <div></div> </div>
1	C	481	<div> <div>2%</div> <div> <div></div> <div>94%</div> <div>6%</div> </div> </div>
1	D	481	<div> <div>2%</div> <div> <div></div> <div>95%</div> <div></div> </div> </div>
1	E	481	<div> <div>2%</div> <div> <div></div> <div>93%</div> <div>7%</div> </div> </div>
1	F	481	<div> <div>%</div> <div> <div></div> <div>94%</div> <div>5%</div> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 25270 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 Histidine decarboxylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	477	Total	C	N	O	S	0	5	0
			3826	2448	653	702	23			
1	B	477	Total	C	N	O	S	0	5	0
			3828	2449	657	699	23			
1	C	481	Total	C	N	O	S	0	11	0
			3904	2503	672	706	23			
1	D	478	Total	C	N	O	S	0	8	0
			3858	2470	669	696	23			
1	E	478	Total	C	N	O	S	0	3	0
			3821	2446	654	698	23			
1	F	478	Total	C	N	O	S	0	4	0
			3832	2453	658	698	23			

There are 42 discrepancies between the modelled and reference sequences:

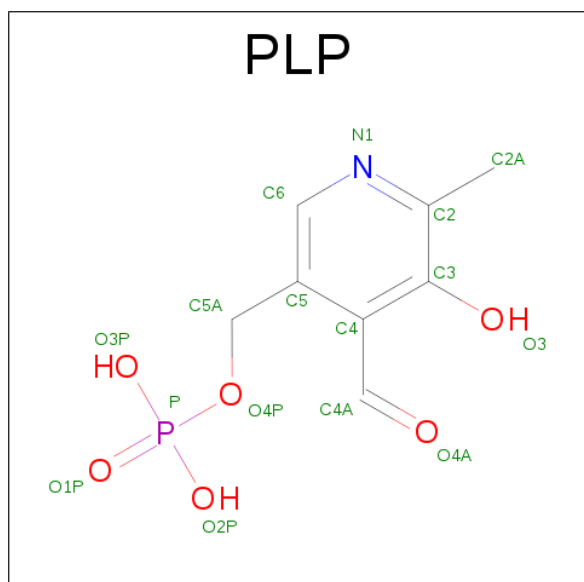
Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	EXPRESSION TAG	UNP P19113
A	-2	PRO	-	EXPRESSION TAG	UNP P19113
A	-1	LEU	-	EXPRESSION TAG	UNP P19113
A	0	GLY	-	EXPRESSION TAG	UNP P19113
A	1	SER	-	EXPRESSION TAG	UNP P19113
A	180	SER	CYS	ENGINEERED MUTATION	UNP P19113
A	418	SER	CYS	ENGINEERED MUTATION	UNP P19113
B	-3	GLY	-	EXPRESSION TAG	UNP P19113
B	-2	PRO	-	EXPRESSION TAG	UNP P19113
B	-1	LEU	-	EXPRESSION TAG	UNP P19113
B	0	GLY	-	EXPRESSION TAG	UNP P19113
B	1	SER	-	EXPRESSION TAG	UNP P19113
B	180	SER	CYS	ENGINEERED MUTATION	UNP P19113
B	418	SER	CYS	ENGINEERED MUTATION	UNP P19113
C	-3	GLY	-	EXPRESSION TAG	UNP P19113
C	-2	PRO	-	EXPRESSION TAG	UNP P19113
C	-1	LEU	-	EXPRESSION TAG	UNP P19113

*Continued on next page...*

Continued from previous page...

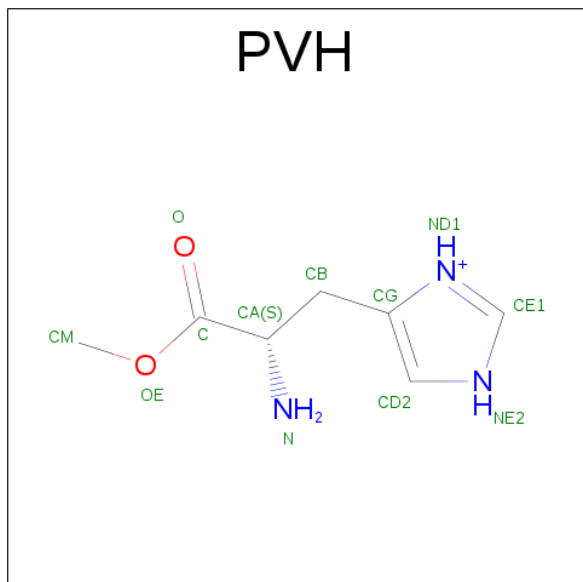
Chain	Residue	Modelled	Actual	Comment	Reference
C	0	GLY	-	EXPRESSION TAG	UNP P19113
C	1	SER	-	EXPRESSION TAG	UNP P19113
C	180	SER	CYS	ENGINEERED MUTATION	UNP P19113
C	418	SER	CYS	ENGINEERED MUTATION	UNP P19113
D	-3	GLY	-	EXPRESSION TAG	UNP P19113
D	-2	PRO	-	EXPRESSION TAG	UNP P19113
D	-1	LEU	-	EXPRESSION TAG	UNP P19113
D	0	GLY	-	EXPRESSION TAG	UNP P19113
D	1	SER	-	EXPRESSION TAG	UNP P19113
D	180	SER	CYS	ENGINEERED MUTATION	UNP P19113
D	418	SER	CYS	ENGINEERED MUTATION	UNP P19113
E	-3	GLY	-	EXPRESSION TAG	UNP P19113
E	-2	PRO	-	EXPRESSION TAG	UNP P19113
E	-1	LEU	-	EXPRESSION TAG	UNP P19113
E	0	GLY	-	EXPRESSION TAG	UNP P19113
E	1	SER	-	EXPRESSION TAG	UNP P19113
E	180	SER	CYS	ENGINEERED MUTATION	UNP P19113
E	418	SER	CYS	ENGINEERED MUTATION	UNP P19113
F	-3	GLY	-	EXPRESSION TAG	UNP P19113
F	-2	PRO	-	EXPRESSION TAG	UNP P19113
F	-1	LEU	-	EXPRESSION TAG	UNP P19113
F	0	GLY	-	EXPRESSION TAG	UNP P19113
F	1	SER	-	EXPRESSION TAG	UNP P19113
F	180	SER	CYS	ENGINEERED MUTATION	UNP P19113
F	418	SER	CYS	ENGINEERED MUTATION	UNP P19113

- Molecule 2 is PYRIDOXAL-5'-PHOSPHATE (three-letter code: PLP) (formula:  $C_8H_{10}NO_6P$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	B	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	C	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	D	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	E	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	F	1	Total	C	N	O	P	0	0
			15	8	1	5	1		

- Molecule 3 is HISTIDINE-METHYL-ESTER (three-letter code: PVH) (formula:  $C_7H_{12}N_3O_2$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			12	7	3	2		
3	B	1	Total	C	N	O	0	0
			12	7	3	2		
3	C	1	Total	C	N	O	0	0
			12	7	3	2		
3	D	1	Total	C	N	O	0	0
			12	7	3	2		
3	E	1	Total	C	N	O	0	0
			12	7	3	2		
3	F	1	Total	C	N	O	0	0
			12	7	3	2		

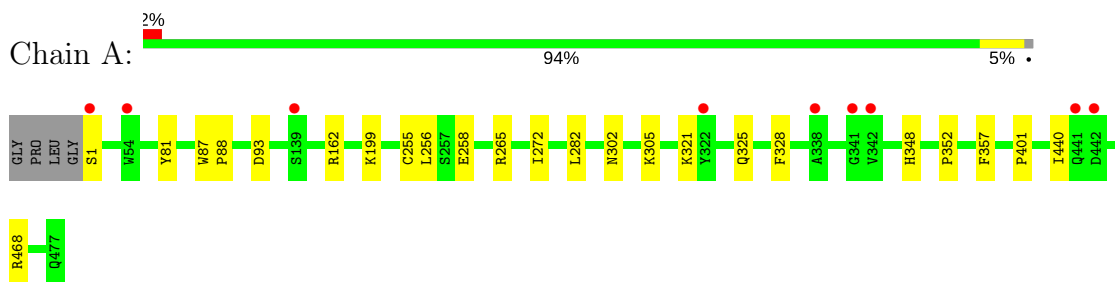
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	357	Total 357	O 357	0	0
4	B	355	Total 355	O 355	0	0
4	C	352	Total 352	O 352	0	0
4	D	353	Total 353	O 353	0	0
4	E	292	Total 292	O 292	0	0
4	F	330	Total 330	O 330	0	0

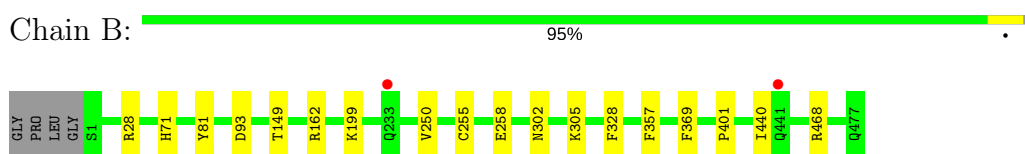
### 3 Residue-property plots [i](#)

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.

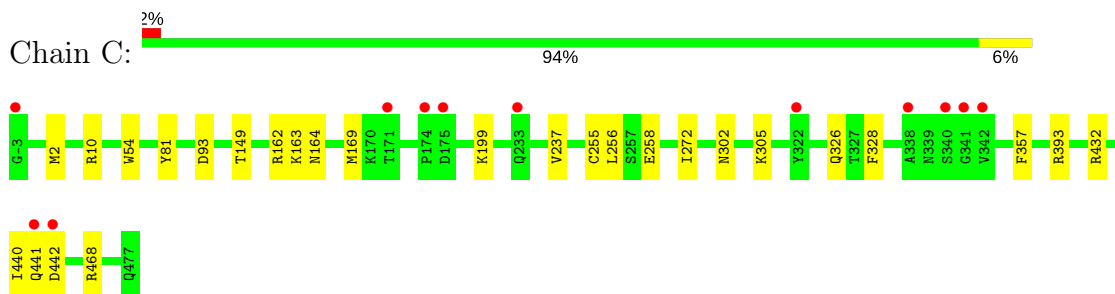
- Molecule 1: Histidine decarboxylase



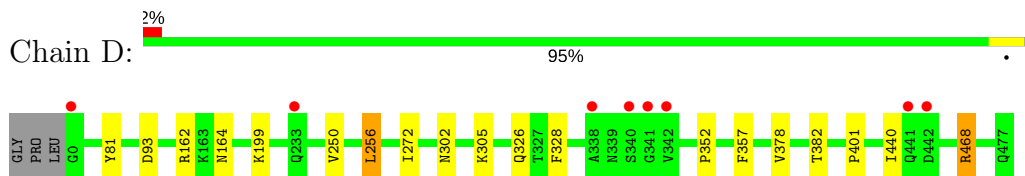
- Molecule 1: Histidine decarboxylase



- Molecule 1: Histidine decarboxylase

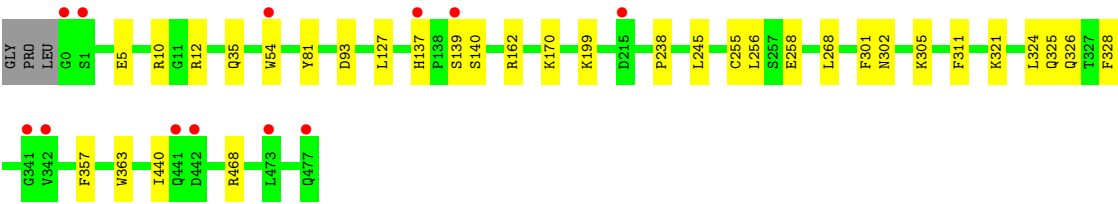


- Molecule 1: Histidine decarboxylase

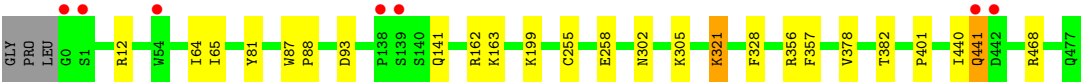
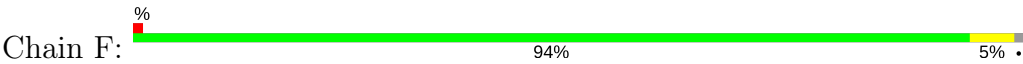


- Molecule 1: Histidine decarboxylase





● Molecule 1: Histidine decarboxylase





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	215.16Å 112.72Å 171.39Å 90.00° 110.30° 90.00°	Depositor
Resolution (Å)	37.59 – 1.80 37.59 – 1.80	Depositor EDS
% Data completeness (in resolution range)	100.0 (37.59-1.80) 99.8 (37.59-1.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.27 (at 1.79Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.162 , 0.181 0.168 , 0.188	Depositor DCC
$R_{free}$ test set	17637 reflections (5.24%)	DCC
Wilson B-factor (Å <sup>2</sup> )	17.2	Xtriage
Anisotropy	0.440	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 43.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	25270	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.42% 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: PVH, CSX, PLP

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	0.37	0/3931	0.57	0/5330
1	B	0.36	0/3933	0.54	0/5332
1	C	0.36	0/4028	0.55	0/5459
1	D	0.37	0/3973	0.55	0/5382
1	E	0.36	0/3921	0.54	0/5317
1	F	0.36	0/3932	0.55	0/5330
All	All	0.36	0/23718	0.55	0/32150

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	3826	0	3787	13	0
1	B	3828	0	3796	11	0
1	C	3904	0	3885	25	0
1	D	3858	0	3847	12	0
1	E	3821	0	3780	18	0
1	F	3832	0	3795	18	0
2	A	15	0	6	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	15	0	7	0	0
2	C	15	0	6	0	0
2	D	15	0	7	0	0
2	E	15	0	7	0	0
2	F	15	0	7	0	0
3	A	12	0	10	1	0
3	B	12	0	10	1	0
3	C	12	0	10	1	0
3	D	12	0	10	1	0
3	E	12	0	10	1	0
3	F	12	0	10	1	0
4	A	357	0	0	2	0
4	B	355	0	0	0	0
4	C	352	0	0	1	0
4	D	353	0	0	0	0
4	E	292	0	0	0	0
4	F	330	0	0	0	0
All	All	25270	0	22990	84	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 (84) 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:C:2:MET:HE2	1:C:10[B]:ARG:HD2	1.63	0.80
1:C:81:TYR:H	3:C:1001:PVH:HD1	1.38	0.72
1:C:2:MET:HE2	1:C:10[B]:ARG:CD	2.21	0.71
1:C:432[B]:ARG:NH2	4:C:1417:HOH:O	2.24	0.71
1:B:81:TYR:H	3:B:1001:PVH:HD1	1.38	0.71
1:D:81:TYR:H	3:D:1001:PVH:HD1	1.40	0.70
1:E:363:TRP:CH2	1:F:65:ILE:HD11	2.29	0.67
1:C:199:LYS:HE3	1:D:328:PHE:HB3	1.80	0.63
1:C:164:ASN:HD22	1:C:326:GLN:HE22	1.45	0.63
1:A:81:TYR:H	3:A:1001:PVH:HD1	1.47	0.63
1:E:363:TRP:CZ3	1:F:65:ILE:HD11	2.34	0.62
1:F:81:TYR:H	3:F:1001:PVH:HD1	1.46	0.62
1:D:164:ASN:HD22	1:D:326:GLN:HE22	1.46	0.61
1:E:81:TYR:H	3:E:1001:PVH:HD1	1.48	0.61
1:F:64:ILE:HG13	1:F:65:ILE:HD12	1.84	0.59
1:C:302:ASN:HD22	1:C:305:LYS:NZ	1.99	0.59

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:401:PRO:HG2	1:D:440:ILE:HD11	1.85	0.58
1:D:302:ASN:HD22	1:D:305:LYS:NZ	2.02	0.58
1:B:255:CSX:HG	1:B:258:GLU:HG3	1.68	0.58
1:E:245:LEU:HD21	1:E:256:LEU:HD13	1.86	0.57
1:C:255:CSX:HG	1:C:258:GLU:HG3	1.70	0.56
1:D:378:VAL:O	1:D:382:THR:HG23	2.05	0.56
1:A:328:PHE:HB3	1:B:199:LYS:HE3	1.87	0.55
1:D:250:VAL:HG11	1:D:440:ILE:HG21	1.88	0.55
1:F:302:ASN:HD22	1:F:305:LYS:NZ	2.05	0.55
1:A:401:PRO:HG2	1:A:440:ILE:HD11	1.89	0.54
1:B:302:ASN:HD22	1:B:305:LYS:NZ	2.06	0.54
1:C:169:MET:HE3	1:C:237:VAL:HG11	1.89	0.53
1:C:10[B]:ARG:HG2	1:C:54[B]:TRP:CD1	2.45	0.52
1:E:302:ASN:HD22	1:E:305:LYS:NZ	2.08	0.52
1:A:256:LEU:HD21	1:A:272:ILE:HD11	1.91	0.51
1:C:2:MET:HE3	1:C:54[A]:TRP:CE2	2.45	0.51
1:C:169:MET:CE	1:C:237:VAL:HG11	2.39	0.51
1:C:2:MET:CE	1:C:54[A]:TRP:CE2	2.94	0.50
1:E:328:PHE:HB3	1:F:199:LYS:HE3	1.92	0.50
1:D:256:LEU:HD11	1:D:272:ILE:HD12	1.92	0.50
1:F:378:VAL:O	1:F:382:THR:HG23	2.12	0.50
1:A:325:GLN:NE2	4:A:1268:HOH:O	2.34	0.50
1:A:255:CSX:HG	1:A:258:GLU:HG3	1.76	0.49
1:C:440:ILE:HG22	1:C:441:GLN:OE1	2.14	0.48
1:A:302:ASN:HD22	1:A:305:LYS:NZ	2.11	0.48
1:F:65:ILE:HD12	1:F:65:ILE:N	2.29	0.48
1:A:265:ARG:NH2	4:A:1248:HOH:O	2.47	0.47
1:C:328:PHE:HB3	1:D:199:LYS:HE3	1.96	0.47
1:F:401:PRO:HG2	1:F:440:ILE:HD11	1.97	0.46
1:E:255:CSX:HG	1:E:258:GLU:HG3	1.80	0.46
1:E:12:ARG:HG2	1:F:12:ARG:HG2	1.97	0.45
1:C:302:ASN:HD22	1:C:305:LYS:HZ3	1.65	0.45
1:C:2:MET:HE3	1:C:54[A]:TRP:NE1	2.32	0.45
1:A:1:SER:HB2	1:B:369:PHE:CD2	2.52	0.45
1:A:87:TRP:HB2	1:A:88:PRO:HD3	1.99	0.45
1:F:440:ILE:O	1:F:441:GLN:HG2	2.17	0.45
1:F:255:CSX:HG	1:F:258:GLU:HG3	1.83	0.44
1:C:393[B]:ARG:HG3	1:C:393[B]:ARG:HH11	1.83	0.44
1:C:10[B]:ARG:HG2	1:C:54[B]:TRP:NE1	2.32	0.44
1:E:238:PRO:HB2	1:E:268:LEU:HD21	2.00	0.44
1:B:302:ASN:HD22	1:B:305:LYS:HZ2	1.65	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:87:TRP:HB2	1:F:88:PRO:HD3	2.00	0.43
1:F:141:GLN:O	1:F:321:LYS:HD2	2.18	0.43
1:C:302:ASN:HD22	1:C:305:LYS:HZ2	1.63	0.43
1:E:137:HIS:HB3	1:E:140:SER:HB2	2.01	0.43
1:E:199:LYS:HE3	1:F:328:PHE:HB3	1.99	0.43
1:D:468[A]:ARG:HA	1:D:468[A]:ARG:HD3	1.70	0.43
1:C:256:LEU:HD21	1:C:272:ILE:HD11	2.02	0.42
1:E:311:PHE:HD2	1:F:356:ARG:HB3	1.84	0.42
1:E:324:LEU:HA	1:E:324:LEU:HD12	1.90	0.42
1:C:2:MET:HE2	1:C:10[B]:ARG:HD3	1.99	0.42
1:B:28:ARG:HD2	1:B:71:HIS:CE1	2.55	0.42
1:C:164:ASN:HD22	1:C:326:GLN:NE2	2.15	0.42
1:E:321:LYS:O	1:E:325:GLN:HG3	2.20	0.42
1:A:325:GLN:HE21	1:A:348:HIS:HB2	1.85	0.42
1:C:149:THR:HG21	1:D:352:PRO:HB2	2.02	0.42
1:E:5:GLU:H	1:E:5:GLU:CD	2.22	0.42
1:C:10[B]:ARG:HG2	1:C:54[B]:TRP:CE2	2.55	0.41
1:D:302:ASN:HD22	1:D:305:LYS:HZ3	1.66	0.41
1:E:127:LEU:HD11	1:E:301:PHE:CE1	2.55	0.41
1:B:250:VAL:HG11	1:B:440:ILE:HG21	2.03	0.41
1:B:401:PRO:HG2	1:B:440:ILE:HD11	2.02	0.41
1:F:302:ASN:HD22	1:F:305:LYS:HZ2	1.69	0.41
1:E:440:ILE:O	1:E:440:ILE:HG23	2.21	0.41
1:A:352:PRO:HB2	1:B:149:THR:HG21	2.03	0.41
1:E:10:ARG:HD2	1:E:54:TRP:CD1	2.56	0.40
1:A:199:LYS:HE3	1:B:328:PHE:HB3	2.02	0.40
1:F:401:PRO:CG	1:F:440:ILE:HD11	2.51	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	479/481 (100%)	465 (97%)	14 (3%)	0	100	100
1	B	479/481 (100%)	467 (98%)	12 (2%)	0	100	100
1	C	489/481 (102%)	475 (97%)	14 (3%)	0	100	100
1	D	483/481 (100%)	470 (97%)	13 (3%)	0	100	100
1	E	478/481 (99%)	466 (98%)	12 (2%)	0	100	100
1	F	479/481 (100%)	465 (97%)	14 (3%)	0	100	100
All	All	2887/2886 (100%)	2808 (97%)	79 (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	416/413 (101%)	410 (99%)	6 (1%)	71	64
1	B	416/413 (101%)	412 (99%)	4 (1%)	80	75
1	C	424/413 (103%)	418 (99%)	6 (1%)	71	64
1	D	419/413 (102%)	413 (99%)	6 (1%)	71	64
1	E	414/413 (100%)	406 (98%)	8 (2%)	62	50
1	F	415/413 (100%)	408 (98%)	7 (2%)	66	55
All	All	2504/2478 (101%)	2467 (98%)	37 (2%)	70	61

All (37) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	93	ASP
1	A	162	ARG
1	A	282	LEU
1	A	321	LYS
1	A	357	PHE
1	A	468	ARG
1	B	93	ASP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	162	ARG
1	B	357	PHE
1	B	468	ARG
1	C	93	ASP
1	C	162	ARG
1	C	163	LYS
1	C	357	PHE
1	C	442	ASP
1	C	468	ARG
1	D	93	ASP
1	D	162	ARG
1	D	256	LEU
1	D	357	PHE
1	D	468[A]	ARG
1	D	468[B]	ARG
1	E	35	GLN
1	E	93	ASP
1	E	139	SER
1	E	162	ARG
1	E	170	LYS
1	E	326	GLN
1	E	357	PHE
1	E	468	ARG
1	F	93	ASP
1	F	162	ARG
1	F	163	LYS
1	F	321	LYS
1	F	357	PHE
1	F	441	GLN
1	F	468	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	164	ASN
1	C	302	ASN
1	D	302	ASN
1	D	326	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

6 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	CSX	A	255	1	4,6,7	0.64	0	2,6,8	0.98	0
1	CSX	B	255	1	4,6,7	0.55	0	2,6,8	1.62	0
1	CSX	C	255	1	4,6,7	0.56	0	2,6,8	1.14	0
1	CSX	D	255	1	4,6,7	0.62	0	2,6,8	1.28	0
1	CSX	E	255	1	4,6,7	0.56	0	2,6,8	1.23	0
1	CSX	F	255	1	4,6,7	0.43	0	2,6,8	1.38	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	CSX	A	255	1	-	0/1/5/7	0/0/0/0
1	CSX	B	255	1	-	0/1/5/7	0/0/0/0
1	CSX	C	255	1	-	0/1/5/7	0/0/0/0
1	CSX	D	255	1	-	0/1/5/7	0/0/0/0
1	CSX	E	255	1	-	0/1/5/7	0/0/0/0
1	CSX	F	255	1	-	0/1/5/7	0/0/0/0

There are no bond length outliers.

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 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	255	CSX	1	0
1	B	255	CSX	1	0
1	C	255	CSX	1	0
1	E	255	CSX	1	0
1	F	255	CSX	1	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

12 ligands 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$
2	PLP	A	1000	3	15,15,16	0.96	1 (6%)	20,22,23	1.11	2 (10%)
3	PVH	A	1001	2	8,12,12	1.51	1 (12%)	8,15,15	1.42	2 (25%)
2	PLP	B	1000	3	15,15,16	1.02	1 (6%)	20,22,23	1.06	1 (5%)
3	PVH	B	1001	2	8,12,12	1.63	1 (12%)	8,15,15	1.41	2 (25%)
2	PLP	C	1000	3	15,15,16	1.09	2 (13%)	20,22,23	1.31	2 (10%)
3	PVH	C	1001	2	8,12,12	1.61	1 (12%)	8,15,15	1.22	0
2	PLP	D	1000	3	15,15,16	1.09	2 (13%)	20,22,23	1.14	2 (10%)
3	PVH	D	1001	2	8,12,12	1.59	1 (12%)	8,15,15	1.55	2 (25%)
2	PLP	E	1000	3	15,15,16	0.98	1 (6%)	20,22,23	1.06	1 (5%)
3	PVH	E	1001	2	8,12,12	1.54	1 (12%)	8,15,15	1.52	3 (37%)
2	PLP	F	1000	3	15,15,16	1.08	2 (13%)	20,22,23	1.28	2 (10%)
3	PVH	F	1001	2	8,12,12	1.62	1 (12%)	8,15,15	1.45	1 (12%)

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	PLP	A	1000	3	-	0/6/6/8	0/1/1/1
3	PVH	A	1001	2	-	0/10/10/10	0/1/1/1
2	PLP	B	1000	3	-	0/6/6/8	0/1/1/1
3	PVH	B	1001	2	-	0/10/10/10	0/1/1/1
2	PLP	C	1000	3	-	0/6/6/8	0/1/1/1
3	PVH	C	1001	2	-	0/10/10/10	0/1/1/1
2	PLP	D	1000	3	-	0/6/6/8	0/1/1/1
3	PVH	D	1001	2	-	0/10/10/10	0/1/1/1
2	PLP	E	1000	3	-	0/6/6/8	0/1/1/1
3	PVH	E	1001	2	-	0/10/10/10	0/1/1/1
2	PLP	F	1000	3	-	0/6/6/8	0/1/1/1
3	PVH	F	1001	2	-	0/10/10/10	0/1/1/1

All (15) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	1000	PLP	C3-C2	-2.31	1.39	1.40
2	C	1000	PLP	C3-C2	-2.20	1.39	1.40
2	B	1000	PLP	C3-C2	-2.02	1.39	1.40
2	F	1000	PLP	C6-N1	2.03	1.38	1.34
2	A	1000	PLP	C2-N1	2.05	1.38	1.33
2	D	1000	PLP	C2-N1	2.08	1.38	1.33
2	E	1000	PLP	C2-N1	2.17	1.38	1.33
2	C	1000	PLP	C2-N1	2.22	1.38	1.33
2	F	1000	PLP	C2-N1	2.40	1.38	1.33
3	A	1001	PVH	OE-C	4.00	1.43	1.33
3	E	1001	PVH	OE-C	4.08	1.43	1.33
3	D	1001	PVH	OE-C	4.19	1.43	1.33
3	C	1001	PVH	OE-C	4.23	1.43	1.33
3	F	1001	PVH	OE-C	4.28	1.44	1.33
3	B	1001	PVH	OE-C	4.28	1.44	1.33

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	1000	PLP	C4A-C4-C3	-2.93	115.49	120.54
2	C	1000	PLP	C4A-C4-C3	-2.52	116.18	120.54
2	D	1000	PLP	C4A-C4-C3	-2.50	116.22	120.54
2	E	1000	PLP	C4A-C4-C3	-2.34	116.50	120.54
3	B	1001	PVH	OE-C-O	-2.28	119.23	123.82

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1001	PVH	OE-C-O	-2.28	119.24	123.82
3	E	1001	PVH	OE-C-O	-2.22	119.34	123.82
2	A	1000	PLP	C4A-C4-C3	-2.16	116.81	120.54
3	D	1001	PVH	OE-C-O	-2.04	119.72	123.82
2	B	1000	PLP	C5-C6-N1	-2.02	120.46	123.87
2	D	1000	PLP	C4A-C4-C5	2.04	122.92	120.86
3	E	1001	PVH	CM-OE-C	2.07	120.82	115.97
3	B	1001	PVH	OE-C-CA	2.28	117.57	111.55
3	F	1001	PVH	OE-C-CA	2.31	117.66	111.55
3	E	1001	PVH	OE-C-CA	2.33	117.70	111.55
3	A	1001	PVH	OE-C-CA	2.38	117.84	111.55
2	A	1000	PLP	C4A-C4-C5	2.39	123.28	120.86
3	D	1001	PVH	OE-C-CA	2.58	118.36	111.55
2	F	1000	PLP	C4A-C4-C5	2.89	123.78	120.86
2	C	1000	PLP	C4A-C4-C5	3.35	124.24	120.86

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1001	PVH	1	0
3	B	1001	PVH	1	0
3	C	1001	PVH	1	0
3	D	1001	PVH	1	0
3	E	1001	PVH	1	0
3	F	1001	PVH	1	0

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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	476/481 (98%)	-0.48	9 (1%) 67 63	10, 16, 29, 41	0
1	B	476/481 (98%)	-0.37	2 (0%) 92 90	12, 18, 29, 36	0
1	C	480/481 (99%)	-0.31	12 (2%) 58 53	10, 17, 33, 42	0
1	D	477/481 (99%)	-0.38	8 (1%) 70 67	10, 17, 30, 40	0
1	E	477/481 (99%)	-0.24	12 (2%) 58 53	13, 21, 34, 46	0
1	F	477/481 (99%)	-0.38	7 (1%) 74 70	11, 18, 31, 46	0
All	All	2863/2886 (99%)	-0.36	50 (1%) 70 67	10, 18, 31, 46	0

All (50) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	1	SER	6.1
1	F	0	GLY	6.1
1	A	342	VAL	5.8
1	E	139	SER	5.6
1	C	442	ASP	5.3
1	C	-3	GLY	4.9
1	C	441	GLN	4.7
1	D	442	ASP	4.7
1	A	1	SER	4.6
1	F	1	SER	4.2
1	D	0	GLY	4.0
1	C	233	GLN	3.9
1	C	171	THR	3.8
1	F	442	ASP	3.8
1	C	342	VAL	3.8
1	E	0	GLY	3.6
1	E	442	ASP	3.6
1	F	54	TRP	3.6
1	D	233	GLN	3.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	342	VAL	3.5
1	C	175	ASP	3.5
1	D	441	GLN	3.4
1	F	139	SER	3.2
1	E	342	VAL	3.1
1	D	340	SER	3.0
1	D	338	ALA	2.8
1	B	233	GLN	2.8
1	E	341	GLY	2.8
1	E	441	GLN	2.7
1	C	341	GLY	2.7
1	E	473	LEU	2.6
1	D	341	GLY	2.6
1	A	441	GLN	2.6
1	C	340	SER	2.5
1	A	338	ALA	2.5
1	F	138	PRO	2.5
1	A	139	SER	2.5
1	E	54	TRP	2.5
1	F	441	GLN	2.5
1	E	137	HIS	2.4
1	B	441	GLN	2.4
1	A	322	TYR	2.3
1	E	215	ASP	2.3
1	E	477	GLN	2.3
1	C	338	ALA	2.3
1	C	322	TYR	2.2
1	C	174	PRO	2.2
1	A	442	ASP	2.0
1	A	341	GLY	2.0
1	A	54	TRP	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [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. 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( $\text{\AA}^2$ )	Q<0.9
1	CSX	A	255	7/8	0.93	0.07	-	16,17,22,26	0
1	CSX	B	255	7/8	0.90	0.09	-	21,23,27,29	0
1	CSX	E	255	7/8	0.92	0.09	-	24,25,28,29	0
1	CSX	C	255	7/8	0.94	0.08	-	20,22,26,28	0
1	CSX	F	255	7/8	0.92	0.09	-	19,20,24,27	0
1	CSX	D	255	7/8	0.94	0.08	-	21,23,26,29	0

### 6.3 Carbohydrates [i](#)

There are no carbohydrates 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. 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( $\text{\AA}^2$ )	Q<0.9
3	PVH	D	1001	12/12	0.95	0.13	0.84	13,17,19,19	0
3	PVH	A	1001	12/12	0.96	0.10	0.07	11,14,15,15	0
2	PLP	D	1000	15/16	0.98	0.12	0.00	13,15,16,17	0
2	PLP	B	1000	15/16	0.98	0.12	-0.05	13,15,18,18	0
2	PLP	E	1000	15/16	0.99	0.10	-0.08	14,17,19,20	0
3	PVH	B	1001	12/12	0.95	0.10	-0.18	14,17,19,19	0
2	PLP	F	1000	15/16	0.98	0.09	-0.34	12,13,15,16	0
3	PVH	C	1001	12/12	0.96	0.08	-0.37	13,15,17,18	0
3	PVH	E	1001	12/12	0.96	0.10	-0.37	13,17,18,19	0
2	PLP	C	1000	15/16	0.98	0.09	-0.58	12,14,15,17	0
2	PLP	A	1000	15/16	0.99	0.07	-0.70	10,12,13,14	0
3	PVH	F	1001	12/12	0.95	0.08	-0.71	13,15,17,17	0

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

There are no such residues in this entry.