



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 18, 2024 – 12:31 AM EST

PDB ID : 4DDF  
Title : Computationally Designed Self-assembling Octahedral Cage protein, O333,  
Crystallized in space group P4  
Authors : Sawaya, M.R.; King, N.P.; Sheffler, W.; Baker, D.; Yeates, T.O.  
Deposited on : 2012-01-18  
Resolution : 3.15 Å (reported)

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

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

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

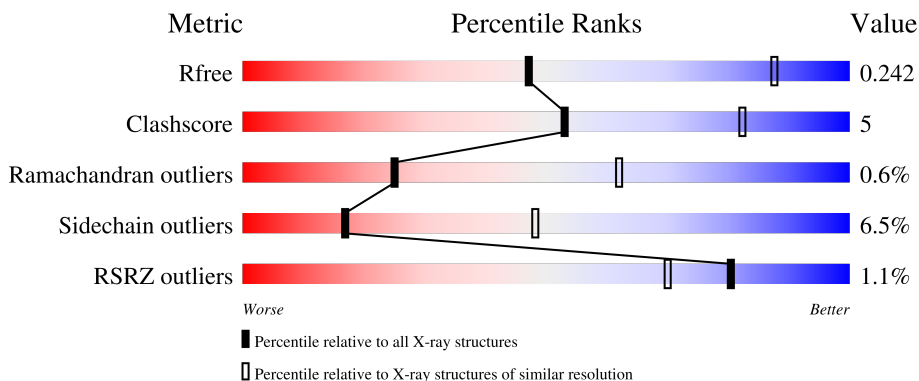
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.15 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	1665 (3.20-3.12)
Clashscore	141614	1804 (3.20-3.12)
Ramachandran outliers	138981	1770 (3.20-3.12)
Sidechain outliers	138945	1769 (3.20-3.12)
RSRZ outliers	127900	1616 (3.20-3.12)

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  $\geq 3$ , 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	192	77% 17% • 5%
1	B	192	80% 15% • 5%
1	C	192	76% 19% • 5%
1	D	192	76% 19% • 5%
1	E	192	78% 17% • 5%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	192	<p>2% 83% 11% • 5%</p>
1	G	192	<p>79% 15% • 5%</p>
1	H	192	<p>3% 82% 13% • 5%</p>
1	I	192	<p>3% 78% 16% • 5%</p>
1	J	192	<p>76% 18% • 5%</p>
1	K	192	<p>76% 16% • 5%</p>
1	L	192	<p>0% 82% 12% • 5%</p>

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 15798 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 Propanediol utilization polyhedral body protein PduT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	183	1310	828	222	251	9	0	0	0
1	B	183	1310	828	222	251	9	0	0	0
1	C	183	1310	828	222	251	9	0	0	0
1	D	183	1310	828	222	251	9	0	0	0
1	E	183	1310	828	222	251	9	0	0	0
1	F	183	1310	828	222	251	9	0	0	0
1	G	183	1310	828	222	251	9	0	0	0
1	H	183	1310	828	222	251	9	0	0	0
1	I	183	1310	828	222	251	9	0	0	0
1	J	183	1310	828	222	251	9	0	0	0
1	K	183	1310	828	222	251	9	0	0	0
1	L	183	1310	828	222	251	9	0	0	0

There are 216 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	15	ALA	LYS	engineered mutation	UNP E7V033
A	38	SER	CYS	engineered mutation	UNP E7V033
A	67	LEU	MET	engineered mutation	UNP E7V033
A	148	ALA	ASN	engineered mutation	UNP E7V033
A	149	LEU	ASN	engineered mutation	UNP E7V033

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	156	SER	GLU	engineered mutation	UNP E7V033
A	160	ALA	GLU	engineered mutation	UNP E7V033
A	161	TYR	LYS	engineered mutation	UNP E7V033
A	167	ALA	ARG	engineered mutation	UNP E7V033
A	169	LEU	VAL	engineered mutation	UNP E7V033
A	185	LEU	-	expression tag	UNP E7V033
A	186	GLU	-	expression tag	UNP E7V033
A	187	HIS	-	expression tag	UNP E7V033
A	188	HIS	-	expression tag	UNP E7V033
A	189	HIS	-	expression tag	UNP E7V033
A	190	HIS	-	expression tag	UNP E7V033
A	191	HIS	-	expression tag	UNP E7V033
A	192	HIS	-	expression tag	UNP E7V033
B	15	ALA	LYS	engineered mutation	UNP E7V033
B	38	SER	CYS	engineered mutation	UNP E7V033
B	67	LEU	MET	engineered mutation	UNP E7V033
B	148	ALA	ASN	engineered mutation	UNP E7V033
B	149	LEU	ASN	engineered mutation	UNP E7V033
B	156	SER	GLU	engineered mutation	UNP E7V033
B	160	ALA	GLU	engineered mutation	UNP E7V033
B	161	TYR	LYS	engineered mutation	UNP E7V033
B	167	ALA	ARG	engineered mutation	UNP E7V033
B	169	LEU	VAL	engineered mutation	UNP E7V033
B	185	LEU	-	expression tag	UNP E7V033
B	186	GLU	-	expression tag	UNP E7V033
B	187	HIS	-	expression tag	UNP E7V033
B	188	HIS	-	expression tag	UNP E7V033
B	189	HIS	-	expression tag	UNP E7V033
B	190	HIS	-	expression tag	UNP E7V033
B	191	HIS	-	expression tag	UNP E7V033
B	192	HIS	-	expression tag	UNP E7V033
C	15	ALA	LYS	engineered mutation	UNP E7V033
C	38	SER	CYS	engineered mutation	UNP E7V033
C	67	LEU	MET	engineered mutation	UNP E7V033
C	148	ALA	ASN	engineered mutation	UNP E7V033
C	149	LEU	ASN	engineered mutation	UNP E7V033
C	156	SER	GLU	engineered mutation	UNP E7V033
C	160	ALA	GLU	engineered mutation	UNP E7V033
C	161	TYR	LYS	engineered mutation	UNP E7V033
C	167	ALA	ARG	engineered mutation	UNP E7V033
C	169	LEU	VAL	engineered mutation	UNP E7V033
C	185	LEU	-	expression tag	UNP E7V033

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	186	GLU	-	expression tag	UNP E7V033
C	187	HIS	-	expression tag	UNP E7V033
C	188	HIS	-	expression tag	UNP E7V033
C	189	HIS	-	expression tag	UNP E7V033
C	190	HIS	-	expression tag	UNP E7V033
C	191	HIS	-	expression tag	UNP E7V033
C	192	HIS	-	expression tag	UNP E7V033
D	15	ALA	LYS	engineered mutation	UNP E7V033
D	38	SER	CYS	engineered mutation	UNP E7V033
D	67	LEU	MET	engineered mutation	UNP E7V033
D	148	ALA	ASN	engineered mutation	UNP E7V033
D	149	LEU	ASN	engineered mutation	UNP E7V033
D	156	SER	GLU	engineered mutation	UNP E7V033
D	160	ALA	GLU	engineered mutation	UNP E7V033
D	161	TYR	LYS	engineered mutation	UNP E7V033
D	167	ALA	ARG	engineered mutation	UNP E7V033
D	169	LEU	VAL	engineered mutation	UNP E7V033
D	185	LEU	-	expression tag	UNP E7V033
D	186	GLU	-	expression tag	UNP E7V033
D	187	HIS	-	expression tag	UNP E7V033
D	188	HIS	-	expression tag	UNP E7V033
D	189	HIS	-	expression tag	UNP E7V033
D	190	HIS	-	expression tag	UNP E7V033
D	191	HIS	-	expression tag	UNP E7V033
D	192	HIS	-	expression tag	UNP E7V033
E	15	ALA	LYS	engineered mutation	UNP E7V033
E	38	SER	CYS	engineered mutation	UNP E7V033
E	67	LEU	MET	engineered mutation	UNP E7V033
E	148	ALA	ASN	engineered mutation	UNP E7V033
E	149	LEU	ASN	engineered mutation	UNP E7V033
E	156	SER	GLU	engineered mutation	UNP E7V033
E	160	ALA	GLU	engineered mutation	UNP E7V033
E	161	TYR	LYS	engineered mutation	UNP E7V033
E	167	ALA	ARG	engineered mutation	UNP E7V033
E	169	LEU	VAL	engineered mutation	UNP E7V033
E	185	LEU	-	expression tag	UNP E7V033
E	186	GLU	-	expression tag	UNP E7V033
E	187	HIS	-	expression tag	UNP E7V033
E	188	HIS	-	expression tag	UNP E7V033
E	189	HIS	-	expression tag	UNP E7V033
E	190	HIS	-	expression tag	UNP E7V033
E	191	HIS	-	expression tag	UNP E7V033

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	192	HIS	-	expression tag	UNP E7V033
F	15	ALA	LYS	engineered mutation	UNP E7V033
F	38	SER	CYS	engineered mutation	UNP E7V033
F	67	LEU	MET	engineered mutation	UNP E7V033
F	148	ALA	ASN	engineered mutation	UNP E7V033
F	149	LEU	ASN	engineered mutation	UNP E7V033
F	156	SER	GLU	engineered mutation	UNP E7V033
F	160	ALA	GLU	engineered mutation	UNP E7V033
F	161	TYR	LYS	engineered mutation	UNP E7V033
F	167	ALA	ARG	engineered mutation	UNP E7V033
F	169	LEU	VAL	engineered mutation	UNP E7V033
F	185	LEU	-	expression tag	UNP E7V033
F	186	GLU	-	expression tag	UNP E7V033
F	187	HIS	-	expression tag	UNP E7V033
F	188	HIS	-	expression tag	UNP E7V033
F	189	HIS	-	expression tag	UNP E7V033
F	190	HIS	-	expression tag	UNP E7V033
F	191	HIS	-	expression tag	UNP E7V033
F	192	HIS	-	expression tag	UNP E7V033
G	15	ALA	LYS	engineered mutation	UNP E7V033
G	38	SER	CYS	engineered mutation	UNP E7V033
G	67	LEU	MET	engineered mutation	UNP E7V033
G	148	ALA	ASN	engineered mutation	UNP E7V033
G	149	LEU	ASN	engineered mutation	UNP E7V033
G	156	SER	GLU	engineered mutation	UNP E7V033
G	160	ALA	GLU	engineered mutation	UNP E7V033
G	161	TYR	LYS	engineered mutation	UNP E7V033
G	167	ALA	ARG	engineered mutation	UNP E7V033
G	169	LEU	VAL	engineered mutation	UNP E7V033
G	185	LEU	-	expression tag	UNP E7V033
G	186	GLU	-	expression tag	UNP E7V033
G	187	HIS	-	expression tag	UNP E7V033
G	188	HIS	-	expression tag	UNP E7V033
G	189	HIS	-	expression tag	UNP E7V033
G	190	HIS	-	expression tag	UNP E7V033
G	191	HIS	-	expression tag	UNP E7V033
G	192	HIS	-	expression tag	UNP E7V033
H	15	ALA	LYS	engineered mutation	UNP E7V033
H	38	SER	CYS	engineered mutation	UNP E7V033
H	67	LEU	MET	engineered mutation	UNP E7V033
H	148	ALA	ASN	engineered mutation	UNP E7V033
H	149	LEU	ASN	engineered mutation	UNP E7V033

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
H	156	SER	GLU	engineered mutation	UNP E7V033
H	160	ALA	GLU	engineered mutation	UNP E7V033
H	161	TYR	LYS	engineered mutation	UNP E7V033
H	167	ALA	ARG	engineered mutation	UNP E7V033
H	169	LEU	VAL	engineered mutation	UNP E7V033
H	185	LEU	-	expression tag	UNP E7V033
H	186	GLU	-	expression tag	UNP E7V033
H	187	HIS	-	expression tag	UNP E7V033
H	188	HIS	-	expression tag	UNP E7V033
H	189	HIS	-	expression tag	UNP E7V033
H	190	HIS	-	expression tag	UNP E7V033
H	191	HIS	-	expression tag	UNP E7V033
H	192	HIS	-	expression tag	UNP E7V033
I	15	ALA	LYS	engineered mutation	UNP E7V033
I	38	SER	CYS	engineered mutation	UNP E7V033
I	67	LEU	MET	engineered mutation	UNP E7V033
I	148	ALA	ASN	engineered mutation	UNP E7V033
I	149	LEU	ASN	engineered mutation	UNP E7V033
I	156	SER	GLU	engineered mutation	UNP E7V033
I	160	ALA	GLU	engineered mutation	UNP E7V033
I	161	TYR	LYS	engineered mutation	UNP E7V033
I	167	ALA	ARG	engineered mutation	UNP E7V033
I	169	LEU	VAL	engineered mutation	UNP E7V033
I	185	LEU	-	expression tag	UNP E7V033
I	186	GLU	-	expression tag	UNP E7V033
I	187	HIS	-	expression tag	UNP E7V033
I	188	HIS	-	expression tag	UNP E7V033
I	189	HIS	-	expression tag	UNP E7V033
I	190	HIS	-	expression tag	UNP E7V033
I	191	HIS	-	expression tag	UNP E7V033
I	192	HIS	-	expression tag	UNP E7V033
J	15	ALA	LYS	engineered mutation	UNP E7V033
J	38	SER	CYS	engineered mutation	UNP E7V033
J	67	LEU	MET	engineered mutation	UNP E7V033
J	148	ALA	ASN	engineered mutation	UNP E7V033
J	149	LEU	ASN	engineered mutation	UNP E7V033
J	156	SER	GLU	engineered mutation	UNP E7V033
J	160	ALA	GLU	engineered mutation	UNP E7V033
J	161	TYR	LYS	engineered mutation	UNP E7V033
J	167	ALA	ARG	engineered mutation	UNP E7V033
J	169	LEU	VAL	engineered mutation	UNP E7V033
J	185	LEU	-	expression tag	UNP E7V033

*Continued on next page...*



*Continued from previous page...*

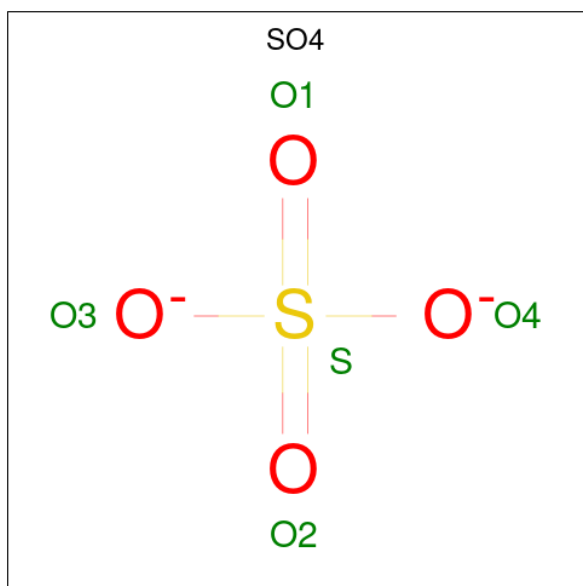
Chain	Residue	Modelled	Actual	Comment	Reference
J	186	GLU	-	expression tag	UNP E7V033
J	187	HIS	-	expression tag	UNP E7V033
J	188	HIS	-	expression tag	UNP E7V033
J	189	HIS	-	expression tag	UNP E7V033
J	190	HIS	-	expression tag	UNP E7V033
J	191	HIS	-	expression tag	UNP E7V033
J	192	HIS	-	expression tag	UNP E7V033
K	15	ALA	LYS	engineered mutation	UNP E7V033
K	38	SER	CYS	engineered mutation	UNP E7V033
K	67	LEU	MET	engineered mutation	UNP E7V033
K	148	ALA	ASN	engineered mutation	UNP E7V033
K	149	LEU	ASN	engineered mutation	UNP E7V033
K	156	SER	GLU	engineered mutation	UNP E7V033
K	160	ALA	GLU	engineered mutation	UNP E7V033
K	161	TYR	LYS	engineered mutation	UNP E7V033
K	167	ALA	ARG	engineered mutation	UNP E7V033
K	169	LEU	VAL	engineered mutation	UNP E7V033
K	185	LEU	-	expression tag	UNP E7V033
K	186	GLU	-	expression tag	UNP E7V033
K	187	HIS	-	expression tag	UNP E7V033
K	188	HIS	-	expression tag	UNP E7V033
K	189	HIS	-	expression tag	UNP E7V033
K	190	HIS	-	expression tag	UNP E7V033
K	191	HIS	-	expression tag	UNP E7V033
K	192	HIS	-	expression tag	UNP E7V033
L	15	ALA	LYS	engineered mutation	UNP E7V033
L	38	SER	CYS	engineered mutation	UNP E7V033
L	67	LEU	MET	engineered mutation	UNP E7V033
L	148	ALA	ASN	engineered mutation	UNP E7V033
L	149	LEU	ASN	engineered mutation	UNP E7V033
L	156	SER	GLU	engineered mutation	UNP E7V033
L	160	ALA	GLU	engineered mutation	UNP E7V033
L	161	TYR	LYS	engineered mutation	UNP E7V033
L	167	ALA	ARG	engineered mutation	UNP E7V033
L	169	LEU	VAL	engineered mutation	UNP E7V033
L	185	LEU	-	expression tag	UNP E7V033
L	186	GLU	-	expression tag	UNP E7V033
L	187	HIS	-	expression tag	UNP E7V033
L	188	HIS	-	expression tag	UNP E7V033
L	189	HIS	-	expression tag	UNP E7V033
L	190	HIS	-	expression tag	UNP E7V033
L	191	HIS	-	expression tag	UNP E7V033

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
L	192	HIS	-	expression tag	UNP E7V033

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
2	A	1	5	4	1	0	0
2	B	1	5	4	1	0	0
2	C	1	5	4	1	0	0
2	C	1	5	4	1	0	0
2	D	1	5	4	1	0	0
2	E	1	5	4	1	0	0
2	G	1	5	4	1	0	0
2	H	1	5	4	1	0	0
2	I	1	5	4	1	0	0
2	J	1	5	4	1	0	0
2	K	1	5	4	1	0	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	L	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Cl	0	0
			1	1		
3	B	2	Total	Cl	0	0
			2	2		
3	C	1	Total	Cl	0	0
			1	1		
3	E	1	Total	Cl	0	0
			1	1		
3	F	1	Total	Cl	0	0
			1	1		
3	G	1	Total	Cl	0	0
			1	1		
3	H	1	Total	Cl	0	0
			1	1		
3	I	1	Total	Cl	0	0
			1	1		
3	J	1	Total	Cl	0	0
			1	1		
3	K	1	Total	Cl	0	0
			1	1		
3	L	1	Total	Cl	0	0
			1	1		

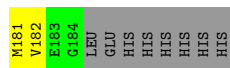
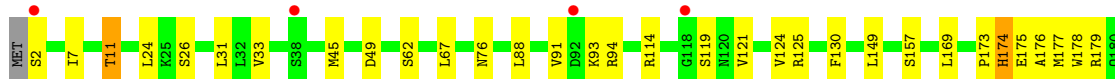
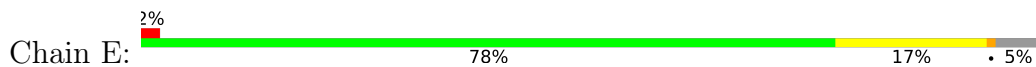
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total	O	0	0
			2	2		
4	D	1	Total	O	0	0
			1	1		
4	E	1	Total	O	0	0
			1	1		
4	K	1	Total	O	0	0
			1	1		
4	L	1	Total	O	0	0
			1	1		

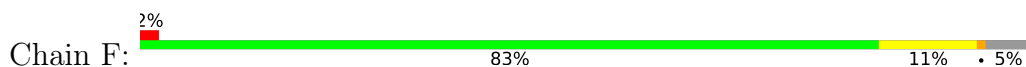




- Molecule 1: Propanediol utilization polyhedral body protein PduT

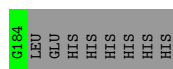
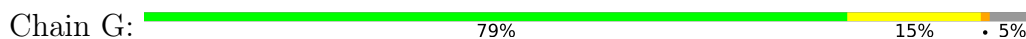


- Molecule 1: Propanediol utilization polyhedral body protein PduT

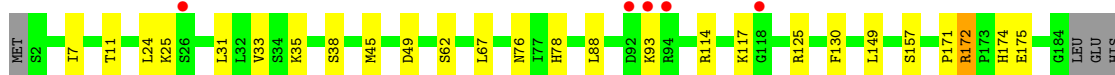
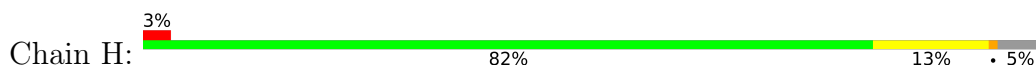


HIS

- Molecule 1: Propanediol utilization polyhedral body protein PduT

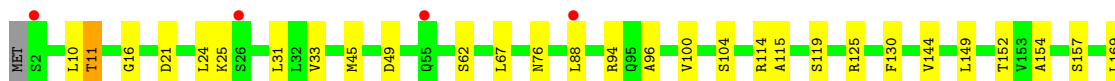
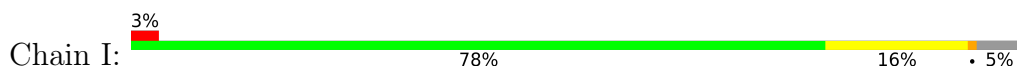


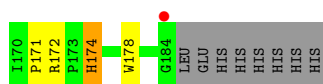
- Molecule 1: Propanediol utilization polyhedral body protein PduT



HIS  
HIS  
HIS  
HIS  
HIS

- Molecule 1: Propanediol utilization polyhedral body protein PduT





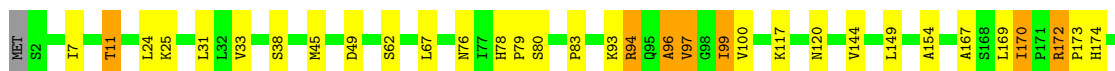
- Molecule 1: Propanediol utilization polyhedral body protein PduT

Chain J: 76% 18% • 5%



- Molecule 1: Propanediol utilization polyhedral body protein PduT

Chain K: 76% 16% • 5%



- Molecule 1: Propanediol utilization polyhedral body protein PduT

Chain L: 82% 12% • 5%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 4	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	169.46Å 169.46Å 119.93Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.84 – 3.15 19.84 – 3.15	Depositor EDS
% Data completeness (in resolution range)	90.3 (19.84-3.15) 90.3 (19.84-3.15)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.93 (at 3.15Å)	Xtrriage
Refinement program	BUSTER-TNT, BUSTER 2.10.0	Depositor
R, $R_{free}$	0.183 , 0.221 0.202 , 0.242	Depositor DCC
$R_{free}$ test set	2704 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	63.7	Xtrriage
Anisotropy	0.026	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 74.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	0.043 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	15798	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	75.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 45.02 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.4041e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: CL, 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	0.59	0/1326	0.77	0/1801
1	B	0.60	0/1326	0.75	0/1801
1	C	0.53	0/1326	0.74	0/1801
1	D	0.48	0/1326	0.71	0/1801
1	E	0.49	0/1326	0.72	0/1801
1	F	0.45	0/1326	0.70	0/1801
1	G	0.44	0/1326	0.70	0/1801
1	H	0.43	0/1326	0.69	0/1801
1	I	0.45	0/1326	0.70	0/1801
1	J	0.61	0/1326	0.81	0/1801
1	K	0.55	0/1326	0.74	0/1801
1	L	0.49	0/1326	0.70	0/1801
All	All	0.51	0/15912	0.73	0/21612

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	1310	0	1355	18	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1310	0	1355	14	0
1	C	1310	0	1355	21	0
1	D	1310	0	1355	24	0
1	E	1310	0	1355	14	0
1	F	1310	0	1355	12	0
1	G	1310	0	1355	16	0
1	H	1310	0	1355	13	0
1	I	1310	0	1355	16	0
1	J	1310	0	1355	20	0
1	K	1310	0	1355	25	0
1	L	1310	0	1355	12	0
2	A	5	0	0	0	0
2	B	5	0	0	0	0
2	C	10	0	0	0	0
2	D	5	0	0	0	0
2	E	5	0	0	0	0
2	G	5	0	0	0	0
2	H	5	0	0	0	0
2	I	5	0	0	0	0
2	J	5	0	0	0	0
2	K	5	0	0	0	0
2	L	5	0	0	0	0
3	A	1	0	0	1	0
3	B	2	0	0	0	0
3	C	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
3	G	1	0	0	0	0
3	H	1	0	0	1	0
3	I	1	0	0	0	0
3	J	1	0	0	0	0
3	K	1	0	0	0	0
3	L	1	0	0	0	0
4	A	2	0	0	0	0
4	D	1	0	0	0	0
4	E	1	0	0	0	0
4	K	1	0	0	0	0
4	L	1	0	0	0	0
All	All	15798	0	16260	174	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 5.

The worst 5 of 174 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:99:ILE:HD11	1:C:181:MET:SD	2.09	0.93
1:C:152:THR:HG23	1:D:152:THR:HG23	1.52	0.91
1:B:24:LEU:HD21	1:B:31:LEU:HB2	1.57	0.87
1:I:152:THR:HG23	1:J:152:THR:HG23	1.62	0.82
1:A:148:ALA:O	1:A:152:THR:HG23	1.88	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	181/192 (94%)	167 (92%)	13 (7%)	1 (1%)	25	62
1	B	181/192 (94%)	174 (96%)	6 (3%)	1 (1%)	25	62
1	C	181/192 (94%)	172 (95%)	8 (4%)	1 (1%)	25	62
1	D	181/192 (94%)	172 (95%)	8 (4%)	1 (1%)	25	62
1	E	181/192 (94%)	167 (92%)	12 (7%)	2 (1%)	14	48
1	F	181/192 (94%)	171 (94%)	10 (6%)	0	100	100
1	G	181/192 (94%)	169 (93%)	12 (7%)	0	100	100
1	H	181/192 (94%)	169 (93%)	10 (6%)	2 (1%)	14	48
1	I	181/192 (94%)	171 (94%)	9 (5%)	1 (1%)	25	62
1	J	181/192 (94%)	172 (95%)	7 (4%)	2 (1%)	14	48
1	K	181/192 (94%)	170 (94%)	10 (6%)	1 (1%)	25	62
1	L	181/192 (94%)	172 (95%)	8 (4%)	1 (1%)	25	62
All	All	2172/2304 (94%)	2046 (94%)	113 (5%)	13 (1%)	25	62

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	175	GLU
1	K	96	ALA
1	E	94	ARG
1	L	175	GLU
1	J	24	LEU

### 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	139/148 (94%)	129 (93%)	10 (7%)	14	43
1	B	139/148 (94%)	129 (93%)	10 (7%)	14	43
1	C	139/148 (94%)	131 (94%)	8 (6%)	20	52
1	D	139/148 (94%)	132 (95%)	7 (5%)	24	57
1	E	139/148 (94%)	125 (90%)	14 (10%)	7	27
1	F	139/148 (94%)	131 (94%)	8 (6%)	20	52
1	G	139/148 (94%)	131 (94%)	8 (6%)	20	52
1	H	139/148 (94%)	130 (94%)	9 (6%)	17	48
1	I	139/148 (94%)	132 (95%)	7 (5%)	24	57
1	J	139/148 (94%)	131 (94%)	8 (6%)	20	52
1	K	139/148 (94%)	126 (91%)	13 (9%)	8	30
1	L	139/148 (94%)	132 (95%)	7 (5%)	24	57
All	All	1668/1776 (94%)	1559 (94%)	109 (6%)	17	48

5 of 109 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	11	THR
1	H	149	LEU
1	K	170	ILE
1	G	62	SER
1	H	11	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	J	120	ASN
1	L	120	ASN
1	E	180	GLN
1	F	89	ASN
1	F	120	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 24 ligands modelled in this entry, 12 are monoatomic - leaving 12 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	$\# Z  > 2$
2	SO4	J	201	-	4,4,4	0.36	0	6,6,6	0.33	0
2	SO4	C	202	-	4,4,4	0.33	0	6,6,6	0.25	0
2	SO4	C	201	-	4,4,4	0.11	0	6,6,6	0.20	0
2	SO4	H	201	-	4,4,4	0.27	0	6,6,6	0.22	0
2	SO4	G	201	-	4,4,4	0.16	0	6,6,6	0.19	0
2	SO4	B	201	-	4,4,4	0.15	0	6,6,6	0.42	0
2	SO4	K	201	-	4,4,4	0.20	0	6,6,6	0.15	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	SO4	D	201	-	4,4,4	0.10	0	6,6,6	0.28	0
2	SO4	A	201	-	4,4,4	0.24	0	6,6,6	0.25	0
2	SO4	L	201	-	4,4,4	0.16	0	6,6,6	0.18	0
2	SO4	I	201	-	4,4,4	0.17	0	6,6,6	0.15	0
2	SO4	E	201	-	4,4,4	0.15	0	6,6,6	0.21	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.

No monomer is involved in short contacts.

## 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	183/192 (95%)	-0.58	1 (0%) 91 86	24, 50, 93, 116	0
1	B	183/192 (95%)	-0.65	0 100 100	21, 45, 85, 103	0
1	C	183/192 (95%)	-0.46	2 (1%) 80 70	32, 58, 95, 117	0
1	D	183/192 (95%)	-0.26	2 (1%) 80 70	45, 81, 122, 131	0
1	E	183/192 (95%)	-0.29	4 (2%) 62 47	51, 77, 110, 122	0
1	F	183/192 (95%)	-0.22	3 (1%) 72 59	55, 83, 123, 141	0
1	G	183/192 (95%)	-0.16	0 100 100	65, 89, 124, 143	0
1	H	183/192 (95%)	-0.06	5 (2%) 54 38	63, 94, 128, 149	0
1	I	183/192 (95%)	-0.10	5 (2%) 54 38	53, 89, 123, 136	0
1	J	183/192 (95%)	-0.54	0 100 100	28, 58, 95, 109	0
1	K	183/192 (95%)	-0.50	0 100 100	35, 58, 91, 112	0
1	L	183/192 (95%)	-0.27	2 (1%) 80 70	38, 74, 107, 121	0
All	All	2196/2304 (95%)	-0.34	24 (1%) 80 70	21, 74, 116, 149	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	L	2	SER	3.9
1	L	26	SER	3.4
1	H	26	SER	3.0
1	C	118	GLY	2.7
1	F	26	SER	2.7

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

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<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(Å <sup>2</sup> )	Q<0.9
3	CL	C	203	1/1	0.67	0.16	79,79,79,79	0
3	CL	I	202	1/1	0.81	0.10	80,80,80,80	0
2	SO4	C	202	5/5	0.82	0.33	144,144,145,145	0
3	CL	E	202	1/1	0.84	0.11	65,65,65,65	0
3	CL	H	202	1/1	0.85	0.10	114,114,114,114	0
3	CL	G	202	1/1	0.86	0.11	90,90,90,90	0
3	CL	J	202	1/1	0.87	0.28	48,48,48,48	0
2	SO4	D	201	5/5	0.88	0.25	124,124,125,125	0
3	CL	K	202	1/1	0.89	0.06	58,58,58,58	0
3	CL	F	201	1/1	0.90	0.11	79,79,79,79	0
3	CL	L	202	1/1	0.90	0.13	79,79,79,79	0
2	SO4	B	201	5/5	0.91	0.19	112,112,113,113	0
2	SO4	I	201	5/5	0.91	0.15	147,147,148,148	0
2	SO4	J	201	5/5	0.91	0.15	127,127,127,128	0
2	SO4	H	201	5/5	0.92	0.21	116,116,116,116	0
2	SO4	G	201	5/5	0.94	0.16	134,134,134,135	0
2	SO4	K	201	5/5	0.95	0.16	113,113,114,115	0
3	CL	B	202	1/1	0.96	0.15	78,78,78,78	0
2	SO4	E	201	5/5	0.96	0.10	116,116,117,118	0
2	SO4	A	201	5/5	0.96	0.15	109,109,109,110	0
2	SO4	C	201	5/5	0.96	0.10	120,122,123,123	0
3	CL	A	202	1/1	0.96	0.06	53,53,53,53	0
2	SO4	L	201	5/5	0.97	0.19	96,97,98,98	0
3	CL	B	203	1/1	0.98	0.22	35,35,35,35	0

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

There are no such residues in this entry.