



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 25, 2020 – 04:03 PM BST

PDB ID : 6J9U  
Title : Complex structure of Lactobacillus casei lactate dehydrogenase penta mutant with pyruvate  
Authors : Arai, K.; Miyanaga, A.; Uchikoba, H.; Fushinobu, S.; Taguchi, H.  
Deposited on : 2019-01-24  
Resolution : 2.79 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 : 2.13  
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) : 2.13

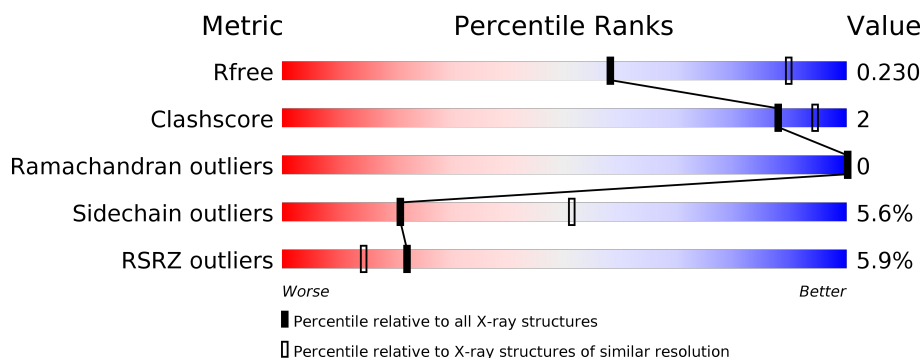
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.79 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 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	326	<div> <div>10%</div> <div> <div></div> <div>84%</div> <div>10%</div> <div>6%</div> </div> </div>
1	B	326	<div> <div>4%</div> <div> <div></div> <div>89%</div> <div>8%</div> <div>.</div> </div> </div>
1	C	326	<div> <div>5%</div> <div> <div></div> <div>85%</div> <div>11%</div> <div>.</div> </div> </div>
1	D	326	<div> <div>6%</div> <div> <div></div> <div>86%</div> <div>10%</div> <div>..</div> </div> </div>
1	E	326	<div> <div>5%</div> <div> <div></div> <div>88%</div> <div>8%</div> <div>..</div> </div> </div>
1	F	326	<div> <div>4%</div> <div> <div></div> <div>87%</div> <div>10%</div> <div>..</div> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 14596 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 L-lactate dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	306	Total	C	N	O	S	0	0	0
			2354	1505	392	451	6			
1	B	315	Total	C	N	O	S	0	0	0
			2422	1547	403	466	6			
1	C	315	Total	C	N	O	S	0	0	0
			2422	1547	403	466	6			
1	D	315	Total	C	N	O	S	0	0	0
			2422	1547	403	466	6			
1	E	315	Total	C	N	O	S	0	0	0
			2421	1547	403	465	6			
1	F	316	Total	C	N	O	S	0	0	0
			2427	1550	404	467	6			

There are 30 discrepancies between the modelled and reference sequences:

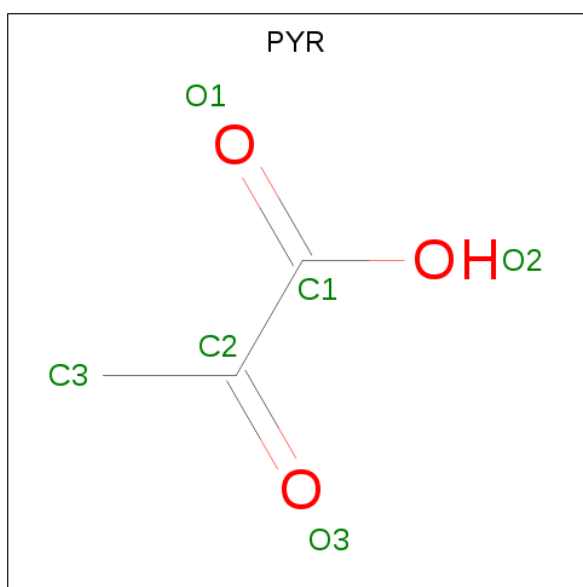
Chain	Residue	Modelled	Actual	Comment	Reference
A	55	GLU	SER	engineered mutation	UNP S6C1Z0
A	56	ASP	ASN	engineered mutation	UNP S6C1Z0
A	164	LYS	GLU	engineered mutation	UNP S6C1Z0
A	223	ASN	ASP	engineered mutation	UNP S6C1Z0
A	224	LYS	ALA	engineered mutation	UNP S6C1Z0
B	55	GLU	SER	engineered mutation	UNP S6C1Z0
B	56	ASP	ASN	engineered mutation	UNP S6C1Z0
B	164	LYS	GLU	engineered mutation	UNP S6C1Z0
B	223	ASN	ASP	engineered mutation	UNP S6C1Z0
B	224	LYS	ALA	engineered mutation	UNP S6C1Z0
C	55	GLU	SER	engineered mutation	UNP S6C1Z0
C	56	ASP	ASN	engineered mutation	UNP S6C1Z0
C	164	LYS	GLU	engineered mutation	UNP S6C1Z0
C	223	ASN	ASP	engineered mutation	UNP S6C1Z0
C	224	LYS	ALA	engineered mutation	UNP S6C1Z0
D	55	GLU	SER	engineered mutation	UNP S6C1Z0
D	56	ASP	ASN	engineered mutation	UNP S6C1Z0

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	164	LYS	GLU	engineered mutation	UNP S6C1Z0
D	223	ASN	ASP	engineered mutation	UNP S6C1Z0
D	224	LYS	ALA	engineered mutation	UNP S6C1Z0
E	55	GLU	SER	engineered mutation	UNP S6C1Z0
E	56	ASP	ASN	engineered mutation	UNP S6C1Z0
E	164	LYS	GLU	engineered mutation	UNP S6C1Z0
E	223	ASN	ASP	engineered mutation	UNP S6C1Z0
E	224	LYS	ALA	engineered mutation	UNP S6C1Z0
F	55	GLU	SER	engineered mutation	UNP S6C1Z0
F	56	ASP	ASN	engineered mutation	UNP S6C1Z0
F	164	LYS	GLU	engineered mutation	UNP S6C1Z0
F	223	ASN	ASP	engineered mutation	UNP S6C1Z0
F	224	LYS	ALA	engineered mutation	UNP S6C1Z0

- Molecule 2 is PYRUVIC ACID (three-letter code: PYR) (formula:  $C_3H_4O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			6	3	3		
2	C	1	Total	C	O	0	0
			6	3	3		
2	D	1	Total	C	O	0	0
			6	3	3		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula:  $O_4S$ ).



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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	3	Total	O	0	0
			3	3		
4	B	4	Total	O	0	0
			4	4		
4	C	11	Total	O	0	0
			11	11		

*Continued on next page...*

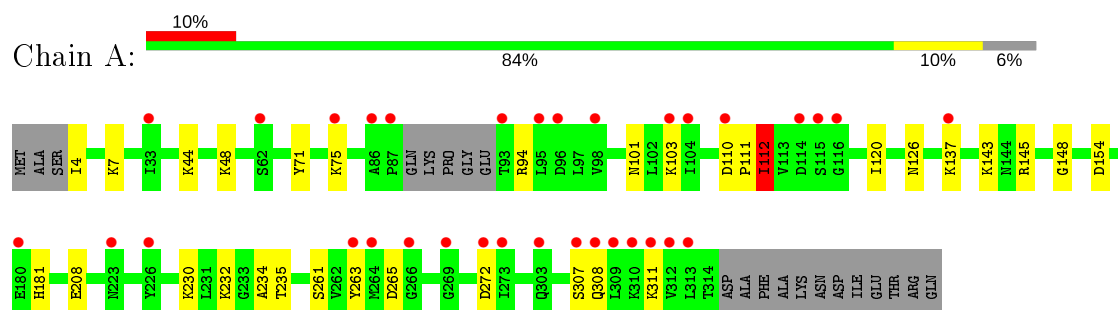
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	20	Total 20	O 20	0	0
4	E	13	Total 13	O 13	0	0
4	F	14	Total 14	O 14	0	0

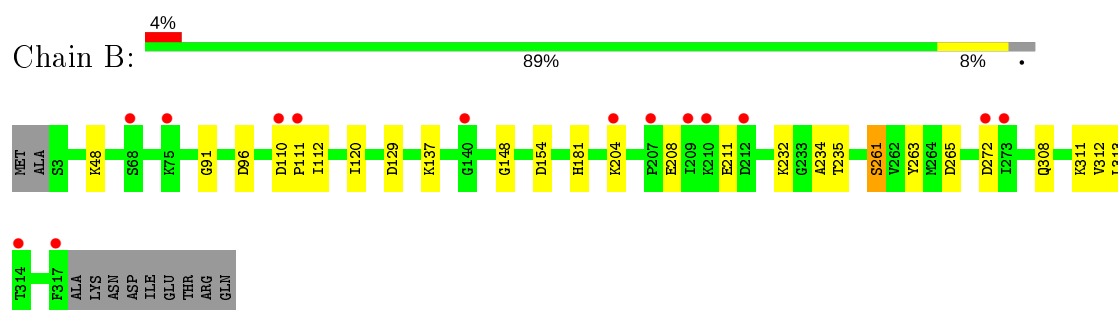
### 3 Residue-property plots [i](#)

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.

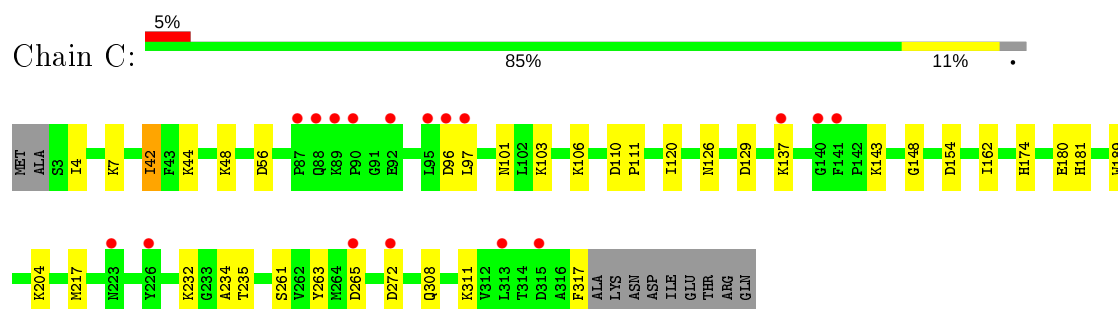
- Molecule 1: L-lactate dehydrogenase



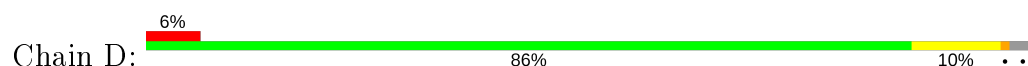
- Molecule 1: L-lactate dehydrogenase

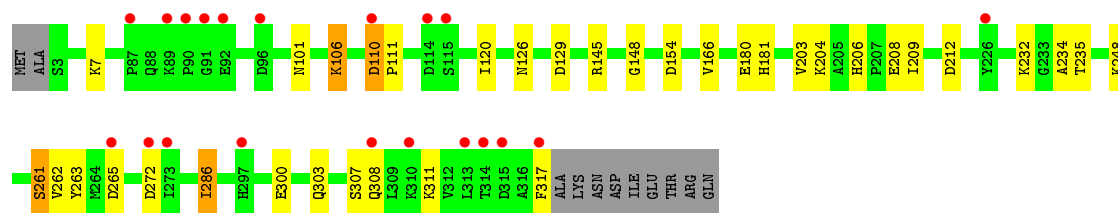


- Molecule 1: L-lactate dehydrogenase

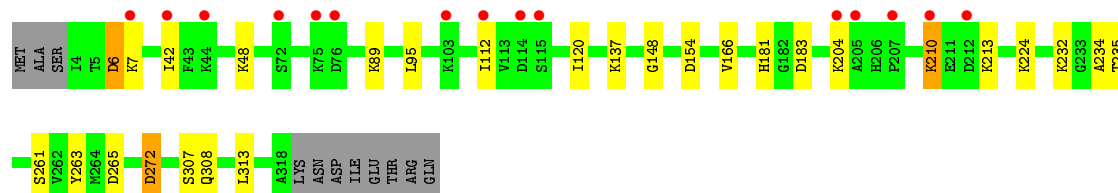
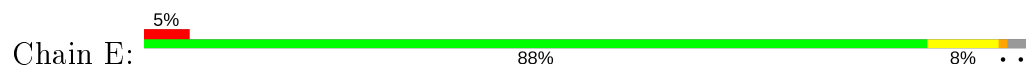


- Molecule 1: L-lactate dehydrogenase

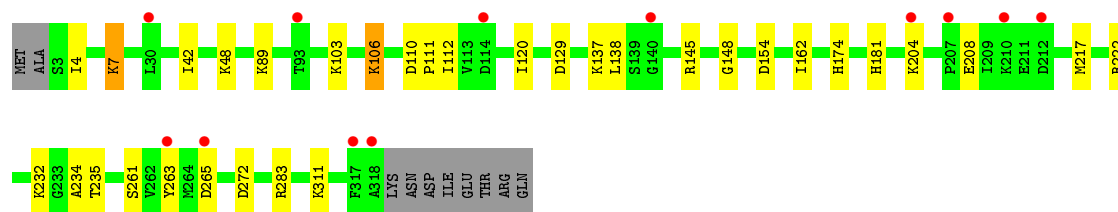
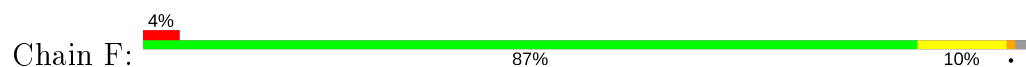




• Molecule 1: L-lactate dehydrogenase



• Molecule 1: L-lactate dehydrogenase





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	165.32Å 82.80Å 180.15Å 90.00° 91.52° 90.00°	Depositor
Resolution (Å)	46.90 – 2.79 46.90 – 2.79	Depositor EDS
% Data completeness (in resolution range)	97.0 (46.90-2.79) 97.1 (46.90-2.79)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.37 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.191 , 0.227 0.197 , 0.230	Depositor DCC
$R_{free}$ test set	2998 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	50.0	Xtriage
Anisotropy	0.133	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 49.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.015 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	14596	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 3.19% 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 ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: PYR, 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	0.60	0/2395	0.77	3/3246 (0.1%)
1	B	0.61	0/2466	0.79	2/3343 (0.1%)
1	C	0.64	0/2466	0.78	2/3343 (0.1%)
1	D	0.69	0/2466	0.84	6/3343 (0.2%)
1	E	0.68	0/2465	0.81	8/3342 (0.2%)
1	F	0.57	0/2471	0.79	5/3350 (0.1%)
All	All	0.63	0/14729	0.80	26/19967 (0.1%)

There are no bond length outliers.

All (26) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	112	ILE	CG1-CB-CG2	-9.36	90.81	111.40
1	B	96	ASP	CB-CG-OD2	-8.53	110.62	118.30
1	B	96	ASP	CB-CG-OD1	8.16	125.65	118.30
1	F	222	ARG	NE-CZ-NH1	-7.66	116.47	120.30
1	F	222	ARG	CG-CD-NE	6.67	125.81	111.80
1	F	145	ARG	NE-CZ-NH2	6.66	123.63	120.30
1	E	7	LYS	CB-CG-CD	6.37	128.16	111.60
1	C	96	ASP	CB-CG-OD2	-6.23	112.70	118.30
1	F	222	ARG	NE-CZ-NH2	5.91	123.26	120.30
1	D	212	ASP	CB-CG-OD1	-5.89	113.00	118.30
1	E	183	ASP	CB-CG-OD2	5.84	123.56	118.30
1	D	248	LYS	CD-CE-NZ	5.83	125.10	111.70
1	E	313	LEU	CB-CG-CD1	-5.78	101.18	111.00
1	E	183	ASP	CB-CG-OD1	-5.66	113.21	118.30
1	E	6	ASP	CB-CG-OD1	-5.63	113.23	118.30
1	E	272	ASP	CB-CG-OD1	-5.57	113.29	118.30
1	C	42	ILE	CB-CA-C	-5.55	100.50	111.60
1	D	286	ILE	CG1-CB-CG2	-5.50	99.31	111.40

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	112	ILE	CA-CB-CG1	5.48	121.41	111.00
1	D	145	ARG	NE-CZ-NH1	-5.38	117.61	120.30
1	E	7	LYS	CA-CB-CG	-5.32	101.69	113.40
1	A	145	ARG	NE-CZ-NH2	5.30	122.95	120.30
1	D	106	LYS	CA-CB-CG	5.25	124.94	113.40
1	F	89	LYS	CA-CB-CG	5.20	124.83	113.40
1	E	6	ASP	CB-CG-OD2	5.11	122.90	118.30
1	D	145	ARG	NE-CZ-NH2	5.10	122.85	120.30

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2354	0	2390	9	0
1	B	2422	0	2451	8	0
1	C	2422	0	2451	14	0
1	D	2422	0	2451	10	0
1	E	2421	0	2451	8	0
1	F	2427	0	2456	13	0
2	A	6	0	3	0	0
2	C	6	0	3	0	0
2	D	6	0	3	0	0
3	A	5	0	0	0	0
3	B	10	0	0	0	0
3	C	5	0	0	1	0
3	D	5	0	0	0	0
3	E	10	0	0	1	0
3	F	10	0	0	0	0
4	A	3	0	0	0	0
4	B	4	0	0	0	0
4	C	11	0	0	0	0
4	D	20	0	0	0	0
4	E	13	0	0	1	0
4	F	14	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	14596	0	14659	58	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 (58) 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:A:4:ILE:O	1:A:7:LYS:HG3	1.98	0.63
1:C:4:ILE:O	1:C:7:LYS:HG3	2.00	0.62
1:A:154:ASP:OD2	1:A:181:HIS:ND1	2.38	0.55
1:F:4:ILE:O	1:F:7:LYS:HG3	2.06	0.55
1:F:162:ILE:HG12	1:F:217:MET:HE1	1.90	0.53
1:C:162:ILE:HG12	1:C:217:MET:HE1	1.90	0.53
1:E:154:ASP:OD2	1:E:181:HIS:ND1	2.40	0.53
1:C:162:ILE:HG12	1:C:217:MET:CE	2.40	0.52
1:B:154:ASP:OD2	1:B:181:HIS:ND1	2.42	0.51
1:D:154:ASP:OD2	1:D:181:HIS:ND1	2.38	0.51
1:F:162:ILE:HG12	1:F:217:MET:CE	2.41	0.51
1:E:6:ASP:OD2	1:F:283:ARG:NH1	2.45	0.50
1:C:154:ASP:OD2	1:C:181:HIS:ND1	2.38	0.50
1:F:154:ASP:OD2	1:F:181:HIS:ND1	2.39	0.50
1:F:232:LYS:HE3	1:F:234:ALA:O	2.13	0.49
1:F:106:LYS:HG2	1:F:138:LEU:HD22	1.95	0.49
1:C:7:LYS:HD3	1:D:7:LYS:HB3	1.94	0.48
1:B:91:GLY:HA3	1:C:189:TRP:CD1	2.48	0.48
1:C:232:LYS:HE3	1:C:234:ALA:O	2.13	0.48
1:E:232:LYS:HE3	1:E:234:ALA:O	2.13	0.48
1:D:232:LYS:HE3	1:D:234:ALA:O	2.14	0.48
1:B:232:LYS:HE3	1:B:234:ALA:O	2.13	0.48
1:E:263:TYR:CZ	1:E:272:ASP:HA	2.49	0.47
1:A:232:LYS:HE3	1:A:234:ALA:O	2.14	0.47
1:D:101:ASN:ND2	1:D:126:ASN:O	2.47	0.47
3:E:401:SO4:O4	4:E:501:HOH:O	2.15	0.47
1:A:263:TYR:CZ	1:A:272:ASP:HA	2.50	0.46
1:A:4:ILE:O	1:A:7:LYS:CG	2.63	0.46
1:B:263:TYR:CZ	1:B:272:ASP:HA	2.49	0.46
1:D:262:VAL:HG12	1:D:286:ILE:CD1	2.47	0.45
1:C:4:ILE:O	1:C:7:LYS:CG	2.65	0.45
1:F:311:LYS:HB2	1:F:311:LYS:HE3	1.60	0.45
1:F:110:ASP:HB2	1:F:111:PRO:HD3	1.99	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:110:ASP:HB2	1:C:111:PRO:HD3	2.00	0.44
1:E:148:GLY:HA3	1:E:261:SER:HB2	2.00	0.44
1:B:148:GLY:HA3	1:B:261:SER:HB2	2.00	0.43
1:F:148:GLY:HA3	1:F:261:SER:HB2	2.00	0.43
1:A:101:ASN:ND2	1:A:126:ASN:O	2.51	0.43
1:E:210:LYS:HE2	1:E:213:LYS:HD3	2.00	0.43
1:B:110:ASP:HB2	1:B:111:PRO:HD3	2.01	0.43
1:C:148:GLY:HA3	1:C:261:SER:HB2	2.00	0.43
1:E:210:LYS:HE2	1:E:213:LYS:CD	2.49	0.43
1:F:263:TYR:CZ	1:F:272:ASP:HA	2.53	0.43
1:C:101:ASN:ND2	1:C:126:ASN:O	2.51	0.42
1:A:71:TYR:CD1	1:A:112:ILE:CD1	3.02	0.42
1:D:148:GLY:HA3	1:D:261:SER:HB2	2.00	0.42
1:A:148:GLY:HA3	1:A:261:SER:HB2	2.00	0.42
1:A:110:ASP:HB2	1:A:111:PRO:HD3	2.01	0.42
1:B:211:GLU:O	1:B:211:GLU:HG3	2.18	0.42
1:C:56:ASP:OD1	1:E:224:LYS:NZ	2.53	0.42
1:D:263:TYR:CZ	1:D:272:ASP:HA	2.55	0.42
1:C:174:HIS:CE1	1:F:174:HIS:CE1	3.08	0.41
3:C:402:SO4:O4	1:F:174:HIS:NE2	2.40	0.41
1:D:206:HIS:ND1	1:D:208:GLU:OE2	2.54	0.41
1:C:263:TYR:CZ	1:C:272:ASP:HA	2.56	0.41
1:D:203:VAL:HG22	1:D:209:ILE:HG22	2.03	0.41
1:D:110:ASP:HB2	1:D:111:PRO:HD3	2.03	0.41
1:B:308:GLN:O	1:B:312:VAL:HG23	2.21	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	302/326 (93%)	292 (97%)	10 (3%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	313/326 (96%)	306 (98%)	7 (2%)	0	100	100
1	C	313/326 (96%)	305 (97%)	8 (3%)	0	100	100
1	D	313/326 (96%)	304 (97%)	9 (3%)	0	100	100
1	E	313/326 (96%)	304 (97%)	9 (3%)	0	100	100
1	F	314/326 (96%)	306 (98%)	8 (2%)	0	100	100
All	All	1868/1956 (96%)	1817 (97%)	51 (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	254/270 (94%)	238 (94%)	16 (6%)	18	46
1	B	261/270 (97%)	249 (95%)	12 (5%)	27	60
1	C	261/270 (97%)	244 (94%)	17 (6%)	17	44
1	D	261/270 (97%)	245 (94%)	16 (6%)	18	48
1	E	260/270 (96%)	246 (95%)	14 (5%)	22	53
1	F	261/270 (97%)	248 (95%)	13 (5%)	24	56
All	All	1558/1620 (96%)	1470 (94%)	88 (6%)	21	51

All (88) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	44	LYS
1	A	48	LYS
1	A	75	LYS
1	A	94	ARG
1	A	103	LYS
1	A	112	ILE
1	A	120	ILE
1	A	137	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	143	LYS
1	A	208	GLU
1	A	230	LYS
1	A	235	THR
1	A	265	ASP
1	A	307	SER
1	A	308	GLN
1	A	311	LYS
1	B	48	LYS
1	B	112	ILE
1	B	120	ILE
1	B	129	ASP
1	B	137	LYS
1	B	204	LYS
1	B	208	GLU
1	B	235	THR
1	B	261	SER
1	B	265	ASP
1	B	311	LYS
1	B	313	LEU
1	C	42	ILE
1	C	44	LYS
1	C	48	LYS
1	C	97	LEU
1	C	103	LYS
1	C	106	LYS
1	C	120	ILE
1	C	129	ASP
1	C	137	LYS
1	C	143	LYS
1	C	180	GLU
1	C	204	LYS
1	C	235	THR
1	C	265	ASP
1	C	308	GLN
1	C	311	LYS
1	C	317	PHE
1	D	106	LYS
1	D	110	ASP
1	D	120	ILE
1	D	129	ASP
1	D	166	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	180	GLU
1	D	204	LYS
1	D	235	THR
1	D	261	SER
1	D	265	ASP
1	D	300	GLU
1	D	303	GLN
1	D	307	SER
1	D	308	GLN
1	D	311	LYS
1	D	317	PHE
1	E	42	ILE
1	E	48	LYS
1	E	89	LYS
1	E	95	LEU
1	E	112	ILE
1	E	120	ILE
1	E	137	LYS
1	E	166	VAL
1	E	204	LYS
1	E	210	LYS
1	E	235	THR
1	E	265	ASP
1	E	307	SER
1	E	308	GLN
1	F	7	LYS
1	F	42	ILE
1	F	48	LYS
1	F	103	LYS
1	F	106	LYS
1	F	112	ILE
1	F	120	ILE
1	F	129	ASP
1	F	137	LYS
1	F	204	LYS
1	F	208	GLU
1	F	235	THR
1	F	265	ASP

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



Mol	Chain	Res	Type
1	B	191	HIS
1	B	288	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

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	PYR	C	401	-	2,5,5	0.80	0	2,6,6	3.35	1 (50%)
3	SO4	B	402	-	4,4,4	0.40	0	6,6,6	0.36	0
3	SO4	F	402	-	4,4,4	0.44	0	6,6,6	0.36	0
3	SO4	B	401	-	4,4,4	0.50	0	6,6,6	0.64	0
2	PYR	D	401	-	2,5,5	1.25	0	2,6,6	3.69	1 (50%)
2	PYR	A	401	-	2,5,5	1.08	0	2,6,6	2.28	1 (50%)
3	SO4	F	401	-	4,4,4	0.43	0	6,6,6	0.65	0
3	SO4	E	401	-	4,4,4	0.54	0	6,6,6	0.53	0
3	SO4	C	402	-	4,4,4	0.48	0	6,6,6	0.53	0
3	SO4	A	402	-	4,4,4	0.32	0	6,6,6	0.69	0
3	SO4	E	402	-	4,4,4	0.50	0	6,6,6	0.50	0
3	SO4	D	402	-	4,4,4	0.58	0	6,6,6	0.52	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	PYR	C	401	-	-	0/0/4/4	-
2	PYR	D	401	-	-	0/0/4/4	-
2	PYR	A	401	-	-	0/0/4/4	-

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	401	PYR	O3-C2-C3	5.00	131.43	120.17
2	C	401	PYR	O3-C2-C3	4.72	130.78	120.17
2	A	401	PYR	O3-C2-C3	2.83	126.55	120.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	401	SO4	1	0
3	C	402	SO4	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<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	306/326 (93%)	0.70	33 (10%) 5 3	29, 53, 81, 101	0
1	B	315/326 (96%)	0.42	14 (4%) 34 24	28, 44, 65, 83	0
1	C	315/326 (96%)	0.48	17 (5%) 25 17	29, 48, 78, 93	0
1	D	315/326 (96%)	0.47	20 (6%) 20 12	24, 41, 72, 98	0
1	E	315/326 (96%)	0.31	15 (4%) 30 21	26, 39, 59, 83	0
1	F	316/326 (96%)	0.39	12 (3%) 40 30	31, 48, 68, 79	0
All	All	1882/1956 (96%)	0.46	111 (5%) 22 14	24, 45, 73, 101	0

All (111) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	96	ASP	5.2
1	F	212	ASP	4.4
1	A	95	LEU	4.4
1	A	263	TYR	4.1
1	A	96	ASP	4.1
1	A	272	ASP	4.1
1	C	97	LEU	3.9
1	A	93	THR	3.9
1	D	90	PRO	3.8
1	A	310	LYS	3.8
1	E	205	ALA	3.7
1	A	110	ASP	3.7
1	A	103	LYS	3.7
1	E	114	ASP	3.6
1	A	115	SER	3.5
1	D	89	LYS	3.5
1	F	204	LYS	3.5
1	A	180	GLU	3.5
1	A	114	ASP	3.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	313	LEU	3.4
1	B	212	ASP	3.3
1	D	272	ASP	3.3
1	E	212	ASP	3.3
1	A	269	GLY	3.3
1	C	88	GLN	3.2
1	A	311	LYS	3.2
1	A	226	TYR	3.1
1	C	89	LYS	3.1
1	D	92	GLU	3.1
1	C	315	ASP	3.1
1	B	209	ILE	3.0
1	D	91	GLY	3.0
1	A	303	GLN	3.0
1	C	92	GLU	3.0
1	B	210	LYS	3.0
1	E	210	LYS	2.9
1	A	309	LEU	2.9
1	D	313	LEU	2.9
1	C	90	PRO	2.9
1	A	307	SER	2.9
1	A	137	LYS	2.9
1	E	115	SER	2.8
1	E	72	SER	2.8
1	D	114	ASP	2.8
1	D	314	THR	2.8
1	A	308	GLN	2.7
1	C	313	LEU	2.7
1	D	317	PHE	2.7
1	A	266	GLY	2.7
1	A	75	LYS	2.7
1	A	273	ILE	2.7
1	A	86	ALA	2.7
1	D	297	HIS	2.7
1	D	265	ASP	2.7
1	E	207	PRO	2.6
1	B	75	LYS	2.6
1	C	265	ASP	2.6
1	D	310	LYS	2.6
1	E	7	LYS	2.5
1	F	207	PRO	2.5
1	F	210	LYS	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	F	265	ASP	2.5
1	E	204	LYS	2.5
1	D	115	SER	2.5
1	C	140	GLY	2.4
1	C	137	LYS	2.4
1	A	116	GLY	2.4
1	E	75	LYS	2.4
1	D	87	PRO	2.4
1	F	140	GLY	2.4
1	E	76	ASP	2.4
1	F	93	THR	2.4
1	D	273	ILE	2.4
1	A	104	ILE	2.3
1	B	273	ILE	2.3
1	A	312	VAL	2.3
1	B	317	PHE	2.3
1	C	223	ASN	2.3
1	B	314	THR	2.3
1	E	42	ILE	2.3
1	B	204	LYS	2.2
1	B	272	ASP	2.2
1	A	87	PRO	2.2
1	C	272	ASP	2.2
1	D	315	ASP	2.2
1	B	140	GLY	2.2
1	E	112	ILE	2.1
1	A	62	SER	2.1
1	A	33	ILE	2.1
1	F	263	TYR	2.1
1	D	110	ASP	2.1
1	F	317	PHE	2.1
1	B	111	PRO	2.1
1	E	103	LYS	2.1
1	D	96	ASP	2.1
1	B	207	PRO	2.1
1	F	318	ALA	2.1
1	F	30	LEU	2.1
1	A	264	MET	2.1
1	D	308	GLN	2.0
1	A	223	ASN	2.0
1	F	114	ASP	2.0
1	C	141	PHE	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	95	LEU	2.0
1	D	226	TYR	2.0
1	B	68	SER	2.0
1	A	98	VAL	2.0
1	C	87	PRO	2.0
1	C	226	TYR	2.0
1	B	110	ASP	2.0
1	E	44	LYS	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. 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
2	PYR	C	401	6/6	0.88	0.22	49,50,52,53	0
3	SO4	C	402	5/5	0.91	0.19	67,72,77,80	0
2	PYR	A	401	6/6	0.92	0.16	46,49,54,54	0
3	SO4	F	402	5/5	0.92	0.21	74,76,80,83	0
3	SO4	E	402	5/5	0.92	0.24	66,68,76,85	0
3	SO4	E	401	5/5	0.93	0.18	63,66,74,78	0
3	SO4	D	402	5/5	0.93	0.18	58,59,69,69	0
3	SO4	B	401	5/5	0.94	0.14	62,70,74,77	0
2	PYR	D	401	6/6	0.95	0.17	29,36,39,41	0
3	SO4	F	401	5/5	0.95	0.17	63,65,76,79	0
3	SO4	B	402	5/5	0.96	0.18	86,86,87,90	0
3	SO4	A	402	5/5	0.96	0.17	67,69,72,79	0

## 6.5 Other polymers

There are no such residues in this entry.