



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 10:31 AM GMT

PDB ID : 3M82  
Title : Crystal structure of Acetyl xylan esterase (TM0077) from THERMOTOGA MARITIMA at 2.40 Å resolution (PMSF inhibitor complex structure)  
Authors : Joint Center for Structural Genomics (JCSG)  
Deposited on : 2010-03-17  
Resolution : 2.40 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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

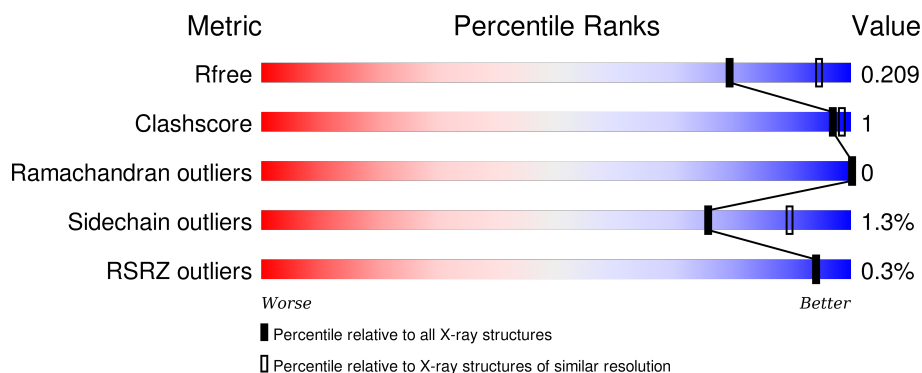
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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))
$R_{free}$	91344	2919 (2.40-2.40)
Clashscore	102246	3407 (2.40-2.40)
Ramachandran outliers	100387	3351 (2.40-2.40)
Sidechain outliers	100360	3352 (2.40-2.40)
RSRZ outliers	91569	2928 (2.40-2.40)

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	337	 88% 7% 5%
1	B	337	 91% • 5%
1	C	337	 90% 6% 5%
1	D	337	 91% • 5%
1	E	337	 89% 6% 5%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	337	 92% 5%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	ACT	E	411	-	-	-	X
4	EDO	B	414	-	-	-	X
4	EDO	C	415	-	-	-	X
4	EDO	E	416	-	-	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 16721 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 Acetyl xylan esterase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	321	Total	C	N	O	S	0	2	0
			2620	1703	432	474	11			
1	B	321	Total	C	N	O	S	0	3	0
			2631	1709	437	474	11			
1	C	321	Total	C	N	O	S	0	1	0
			2611	1698	431	471	11			
1	D	321	Total	C	N	O	S	0	1	0
			2611	1698	431	471	11			
1	E	321	Total	C	N	O	S	0	4	0
			2635	1712	436	476	11			
1	F	321	Total	C	N	O	S	0	2	0
			2621	1704	432	474	11			

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	MET	-	leader sequence	UNP Q9WXT2
A	-10	GLY	-	leader sequence	UNP Q9WXT2
A	-9	SER	-	leader sequence	UNP Q9WXT2
A	-8	ASP	-	leader sequence	UNP Q9WXT2
A	-7	LYS	-	leader sequence	UNP Q9WXT2
A	-6	ILE	-	leader sequence	UNP Q9WXT2
A	-5	HIS	-	leader sequence	UNP Q9WXT2
A	-4	HIS	-	leader sequence	UNP Q9WXT2
A	-3	HIS	-	leader sequence	UNP Q9WXT2
A	-2	HIS	-	leader sequence	UNP Q9WXT2
A	-1	HIS	-	leader sequence	UNP Q9WXT2
A	0	HIS	-	leader sequence	UNP Q9WXT2
B	-11	MET	-	leader sequence	UNP Q9WXT2
B	-10	GLY	-	leader sequence	UNP Q9WXT2
B	-9	SER	-	leader sequence	UNP Q9WXT2
B	-8	ASP	-	leader sequence	UNP Q9WXT2
B	-7	LYS	-	leader sequence	UNP Q9WXT2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-6	ILE	-	leader sequence	UNP Q9WXT2
B	-5	HIS	-	leader sequence	UNP Q9WXT2
B	-4	HIS	-	leader sequence	UNP Q9WXT2
B	-3	HIS	-	leader sequence	UNP Q9WXT2
B	-2	HIS	-	leader sequence	UNP Q9WXT2
B	-1	HIS	-	leader sequence	UNP Q9WXT2
B	0	HIS	-	leader sequence	UNP Q9WXT2
C	-11	MET	-	leader sequence	UNP Q9WXT2
C	-10	GLY	-	leader sequence	UNP Q9WXT2
C	-9	SER	-	leader sequence	UNP Q9WXT2
C	-8	ASP	-	leader sequence	UNP Q9WXT2
C	-7	LYS	-	leader sequence	UNP Q9WXT2
C	-6	ILE	-	leader sequence	UNP Q9WXT2
C	-5	HIS	-	leader sequence	UNP Q9WXT2
C	-4	HIS	-	leader sequence	UNP Q9WXT2
C	-3	HIS	-	leader sequence	UNP Q9WXT2
C	-2	HIS	-	leader sequence	UNP Q9WXT2
C	-1	HIS	-	leader sequence	UNP Q9WXT2
C	0	HIS	-	leader sequence	UNP Q9WXT2
D	-11	MET	-	leader sequence	UNP Q9WXT2
D	-10	GLY	-	leader sequence	UNP Q9WXT2
D	-9	SER	-	leader sequence	UNP Q9WXT2
D	-8	ASP	-	leader sequence	UNP Q9WXT2
D	-7	LYS	-	leader sequence	UNP Q9WXT2
D	-6	ILE	-	leader sequence	UNP Q9WXT2
D	-5	HIS	-	leader sequence	UNP Q9WXT2
D	-4	HIS	-	leader sequence	UNP Q9WXT2
D	-3	HIS	-	leader sequence	UNP Q9WXT2
D	-2	HIS	-	leader sequence	UNP Q9WXT2
D	-1	HIS	-	leader sequence	UNP Q9WXT2
D	0	HIS	-	leader sequence	UNP Q9WXT2
E	-11	MET	-	leader sequence	UNP Q9WXT2
E	-10	GLY	-	leader sequence	UNP Q9WXT2
E	-9	SER	-	leader sequence	UNP Q9WXT2
E	-8	ASP	-	leader sequence	UNP Q9WXT2
E	-7	LYS	-	leader sequence	UNP Q9WXT2
E	-6	ILE	-	leader sequence	UNP Q9WXT2
E	-5	HIS	-	leader sequence	UNP Q9WXT2
E	-4	HIS	-	leader sequence	UNP Q9WXT2
E	-3	HIS	-	leader sequence	UNP Q9WXT2
E	-2	HIS	-	leader sequence	UNP Q9WXT2
E	-1	HIS	-	leader sequence	UNP Q9WXT2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	0	HIS	-	leader sequence	UNP Q9WXT2
F	-11	MET	-	leader sequence	UNP Q9WXT2
F	-10	GLY	-	leader sequence	UNP Q9WXT2
F	-9	SER	-	leader sequence	UNP Q9WXT2
F	-8	ASP	-	leader sequence	UNP Q9WXT2
F	-7	LYS	-	leader sequence	UNP Q9WXT2
F	-6	ILE	-	leader sequence	UNP Q9WXT2
F	-5	HIS	-	leader sequence	UNP Q9WXT2
F	-4	HIS	-	leader sequence	UNP Q9WXT2
F	-3	HIS	-	leader sequence	UNP Q9WXT2
F	-2	HIS	-	leader sequence	UNP Q9WXT2
F	-1	HIS	-	leader sequence	UNP Q9WXT2
F	0	HIS	-	leader sequence	UNP Q9WXT2

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

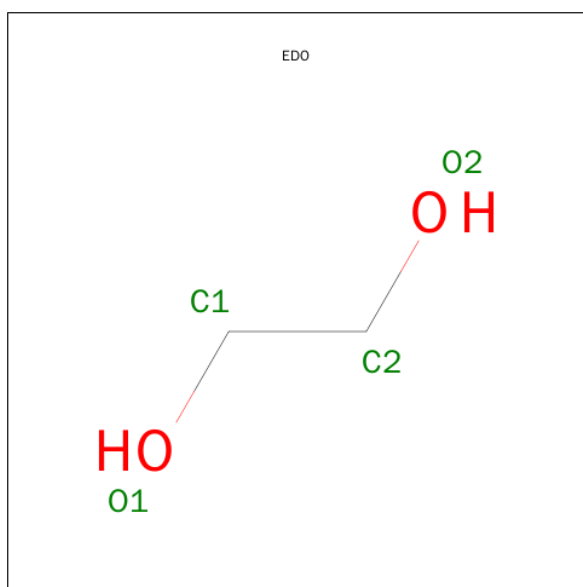
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	D	2	Total 2 Ca 2	0	0
2	E	1	Total 1 Ca 1	0	0
2	B	1	Total 1 Ca 1	0	0
2	C	1	Total 1 Ca 1	0	0
2	A	2	Total 2 Ca 2	0	0
2	F	1	Total 1 Ca 1	0	0

- Molecule 3 is ACETATE ION (three-letter code: ACT) (formula: C<sub>2</sub>H<sub>3</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			4	2	2		
3	B	1	Total	C	O	0	0
			4	2	2		
3	E	1	Total	C	O	0	0
			4	2	2		

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			4	2	2		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 4	C 2	O 2	0	0
4	B	1	Total 4	C 2	O 2	0	0
4	C	1	Total 4	C 2	O 2	0	0
4	E	1	Total 4	C 2	O 2	0	0
4	F	1	Total 4	C 2	O 2	0	0

- Molecule 5 is water.

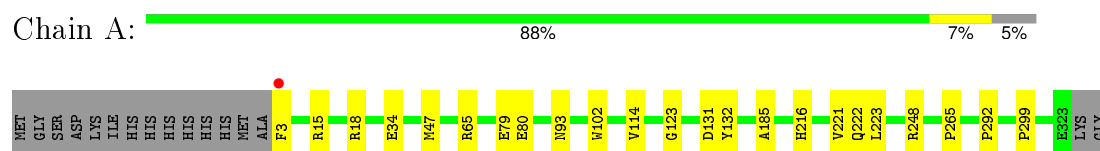
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	174	Total 174	O 174	0	0
5	B	163	Total 164	O 164	0	1
5	C	164	Total 165	O 165	0	1
5	D	161	Total 161	O 161	0	0
5	E	141	Total 141	O 141	0	0
5	F	143	Total 143	O 143	0	0



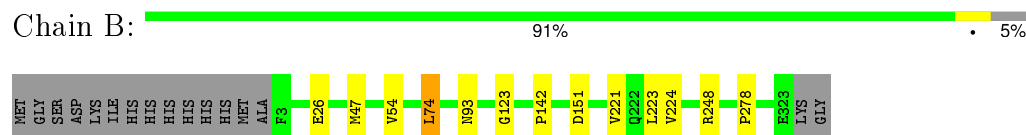
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

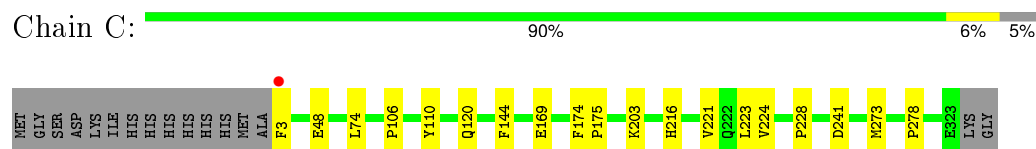
- Molecule 1: Acetyl xylan esterase



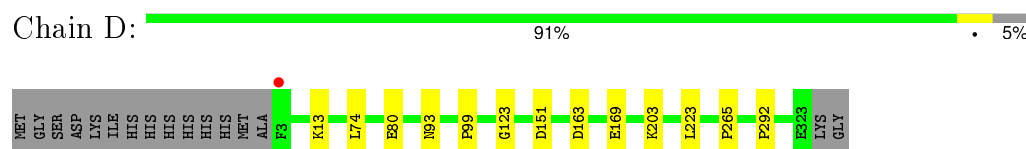
- Molecule 1: Acetyl xylan esterase



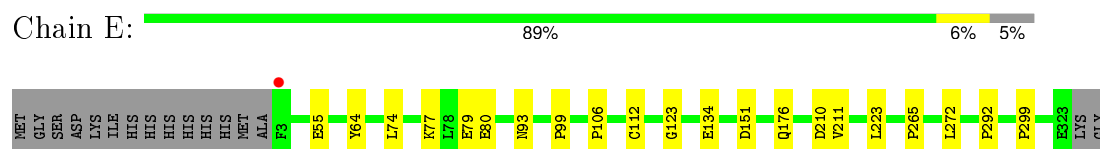
- Molecule 1: Acetyl xylan esterase



- Molecule 1: Acetyl xylan esterase



- Molecule 1: Acetyl xylan esterase



- Molecule 1: Acetyl xylan esterase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	103.57Å 104.50Å 221.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.97 – 2.40 48.95 – 2.40	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.97-2.40) 100.0 (48.95-2.40)	Depositor EDS
$R_{merge}$	0.18	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.21 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.160 , 0.208 0.164 , 0.209	Depositor DCC
$R_{free}$ test set	4742 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	25.4	Xtriage
Anisotropy	0.436	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 40.8	EDS
Estimated twinning fraction	0.025 for k,h,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 94681 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	16721	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.35% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: CA, ACT, SEB, EDO

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	0.78	0/2683	0.77	4/3637 (0.1%)
1	B	0.83	0/2694	0.78	2/3651 (0.1%)
1	C	0.82	0/2674	0.80	1/3625 (0.0%)
1	D	0.84	1/2674 (0.0%)	0.79	3/3625 (0.1%)
1	E	0.86	0/2698	0.75	1/3657 (0.0%)
1	F	0.86	1/2684 (0.0%)	0.77	1/3638 (0.0%)
All	All	0.83	2/16107 (0.0%)	0.78	12/21833 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	80	GLU	CG-CD	5.46	1.60	1.51
1	F	80	GLU	CG-CD	5.43	1.60	1.51

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	131	ASP	CB-CG-OD1	6.06	123.75	118.30
1	D	163	ASP	CB-CG-OD1	5.95	123.66	118.30
1	D	151	ASP	CB-CG-OD1	5.62	123.36	118.30
1	A	248	ARG	NE-CZ-NH1	5.41	123.00	120.30
1	D	163	ASP	CB-CG-OD2	-5.32	113.52	118.30
1	A	65	ARG	NE-CZ-NH1	5.27	122.94	120.30
1	F	248	ARG	NE-CZ-NH1	5.24	122.92	120.30
1	B	151	ASP	CB-CG-OD1	5.22	123.00	118.30
1	E	151	ASP	CB-CG-OD1	5.16	122.94	118.30
1	C	241	ASP	CB-CG-OD1	5.16	122.94	118.30
1	B	248	ARG	NE-CZ-NH2	-5.15	117.72	120.30
1	A	18	ARG	NE-CZ-NH1	5.01	122.81	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	2620	0	2517	8	0
1	B	2631	0	2531	7	0
1	C	2611	0	2512	7	0
1	D	2611	0	2512	4	0
1	E	2635	0	2529	10	0
1	F	2621	0	2519	5	0
2	A	2	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	2	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
3	A	4	0	3	0	0
3	B	4	0	3	1	0
3	E	4	0	3	1	0
4	A	8	0	12	2	0
4	B	4	0	6	0	0
4	C	4	0	6	0	0
4	E	4	0	6	0	0
4	F	4	0	6	1	0
5	A	174	0	0	1	0
5	B	164	0	0	1	0
5	C	165	0	0	0	0
5	D	161	0	0	0	0
5	E	141	0	0	0	0
5	F	143	0	0	1	0
All	All	16721	0	15165	41	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 1.

All (41) 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:F:185:ALA:HB1	4:F:417:EDO:H12	1.71	0.71
1:A:34:GLU:OE1	4:A:412:EDO:H12	1.99	0.61
1:B:142:PRO:HA	3:B:409:ACT:H1	1.84	0.58
1:E:93:ASN:HB3	1:E:123:GLY:HA3	1.85	0.57
1:E:64:TYR:OH	1:E:134:GLU:OE1	2.26	0.52
1:C:3:PHE:CD1	1:C:273:MET:HE2	2.45	0.51
1:F:93:ASN:HB3	1:F:123:GLY:HA3	1.94	0.50
1:E:210:ASP:O	1:E:211:VAL:C	2.51	0.49
1:B:26:GLU:OE1	1:B:26:GLU:N	2.45	0.49
3:E:411:ACT:H1	1:F:142:PRO:HA	1.95	0.48
1:C:169:GLU:OE1	1:C:203:LYS:NZ	2.47	0.46
1:D:93:ASN:HB3	1:D:123:GLY:HA3	1.97	0.46
1:E:272:LEU:HD12	1:E:299:PRO:HA	1.98	0.45
1:E:74:LEU:HD21	1:E:99:PRO:HB3	1.97	0.45
1:E:79:GLU:O	1:E:80:GLU:C	2.55	0.45
1:B:47:MET:HE1	5:B:1131:HOH:O	2.17	0.45
1:A:79[A]:GLU:O	1:A:80:GLU:C	2.55	0.44
1:C:224:VAL:HG11	1:C:278:PRO:HG3	1.99	0.44
1:B:54:VAL:HG11	1:B:74:LEU:HD12	1.99	0.44
1:E:55:GLU:HG3	1:E:77:LYS:HE2	1.99	0.43
1:B:224:VAL:CG1	1:B:278:PRO:HG3	2.48	0.43
1:A:93:ASN:HB3	1:A:123:GLY:HA3	2.01	0.43
1:A:132:TYR:HE1	5:A:1621:HOH:O	2.02	0.43
1:D:265:PRO:HA	1:D:292:PRO:O	2.19	0.43
1:B:224:VAL:HG11	1:B:278:PRO:HG3	2.00	0.42
1:B:93:ASN:HB3	1:B:123:GLY:HA3	2.01	0.42
1:A:3:PHE:HE1	1:A:299:PRO:O	2.02	0.42
1:C:120:GLN:HG2	1:C:144:PHE:CZ	2.54	0.42
1:D:169:GLU:OE2	1:D:203:LYS:HG3	2.19	0.42
1:F:265:PRO:HA	1:F:292:PRO:O	2.20	0.42
1:F:176:GLN:NE2	5:F:1653:HOH:O	2.51	0.42
1:D:74:LEU:HD11	1:D:99:PRO:O	2.19	0.42
1:A:185:ALA:HB1	4:A:413:EDO:H22	2.02	0.41
1:C:106:PRO:HA	1:C:110:TYR:O	2.20	0.41
1:A:265:PRO:HA	1:A:292:PRO:O	2.21	0.41
1:A:102:TRP:CG	1:A:114:VAL:HG21	2.56	0.41
1:E:265:PRO:HA	1:E:292:PRO:O	2.20	0.41
1:C:174:PHE:HA	1:C:175:PRO:HD3	1.99	0.40
1:E:106:PRO:HD3	1:E:112:CYS:HB2	2.04	0.40
1:C:3:PHE:CE1	1:C:273:MET:HE3	2.56	0.40
1:E:80:GLU:O	1:E:176[B]:GLN:NE2	2.55	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	320/337 (95%)	309 (97%)	11 (3%)	0	100	100
1	B	321/337 (95%)	311 (97%)	10 (3%)	0	100	100
1	C	319/337 (95%)	309 (97%)	10 (3%)	0	100	100
1	D	319/337 (95%)	310 (97%)	9 (3%)	0	100	100
1	E	322/337 (96%)	311 (97%)	11 (3%)	0	100	100
1	F	320/337 (95%)	312 (98%)	8 (2%)	0	100	100
All	All	1921/2022 (95%)	1862 (97%)	59 (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	269/284 (95%)	263 (98%)	6 (2%)	60	79
1	B	270/284 (95%)	267 (99%)	3 (1%)	80	92
1	C	268/284 (94%)	262 (98%)	6 (2%)	60	79
1	D	268/284 (94%)	266 (99%)	2 (1%)	88	95
1	E	270/284 (95%)	269 (100%)	1 (0%)	93	98
1	F	269/284 (95%)	266 (99%)	3 (1%)	80	92
All	All	1614/1704 (95%)	1593 (99%)	21 (1%)	76	89

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	15	ARG
1	A	47	MET
1	A	216	HIS
1	A	221	VAL
1	A	222	GLN
1	A	223	LEU
1	B	74	LEU
1	B	221	VAL
1	B	223	LEU
1	C	48	GLU
1	C	74	LEU
1	C	216	HIS
1	C	221	VAL
1	C	223	LEU
1	C	228	PRO
1	D	13	LYS
1	D	223	LEU
1	E	223	LEU
1	F	15	ARG
1	F	216	HIS
1	F	223	LEU

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

Mol	Chain	Res	Type
1	D	50	HIS
1	F	176	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	SEB	A	188	1	14,16,17	1.32	1 (7%)	15,21,23	1.50	3 (20%)
1	SEB	B	188	1	14,16,17	1.45	1 (7%)	15,21,23	1.79	4 (26%)
1	SEB	C	188	1	14,16,17	1.30	1 (7%)	15,21,23	1.06	3 (20%)
1	SEB	D	188	1	14,16,17	1.40	1 (7%)	15,21,23	1.66	4 (26%)
1	SEB	E	188	1	14,16,17	1.50	1 (7%)	15,21,23	1.72	3 (20%)
1	SEB	F	188	1	14,16,17	1.34	1 (7%)	15,21,23	2.01	3 (20%)

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	SEB	A	188	1	-	0/10/13/15	0/1/1/1
1	SEB	B	188	1	-	0/10/13/15	0/1/1/1
1	SEB	C	188	1	-	0/10/13/15	0/1/1/1
1	SEB	D	188	1	-	0/10/13/15	0/1/1/1
1	SEB	E	188	1	-	0/10/13/15	0/1/1/1
1	SEB	F	188	1	-	0/10/13/15	0/1/1/1

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	188	SEB	OG-SD	3.52	1.67	1.56
1	C	188	SEB	OG-SD	3.63	1.67	1.56
1	F	188	SEB	OG-SD	3.70	1.68	1.56
1	B	188	SEB	OG-SD	4.04	1.69	1.56
1	E	188	SEB	OG-SD	4.10	1.69	1.56
1	D	188	SEB	OG-SD	4.22	1.69	1.56

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	188	SEB	O-C-CA	-2.88	117.98	125.49
1	C	188	SEB	O-C-CA	-2.58	118.77	125.49
1	D	188	SEB	O-C-CA	-2.48	119.02	125.49
1	A	188	SEB	O-C-CA	-2.37	119.32	125.49

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	188	SEB	O-C-CA	-2.22	119.72	125.49
1	B	188	SEB	O-C-CA	-2.04	120.17	125.49
1	C	188	SEB	CZ-CE-SD	2.05	114.91	112.58
1	C	188	SEB	CB-OG-SD	2.15	124.00	119.10
1	D	188	SEB	OG-CB-CA	2.57	114.16	108.03
1	A	188	SEB	CB-OG-SD	2.71	125.29	119.10
1	E	188	SEB	CB-OG-SD	2.72	125.31	119.10
1	B	188	SEB	OG-CB-CA	2.74	114.56	108.03
1	D	188	SEB	CB-OG-SD	3.00	125.95	119.10
1	B	188	SEB	CZ-CE-SD	3.86	116.98	112.58
1	B	188	SEB	CB-OG-SD	3.89	127.99	119.10
1	D	188	SEB	CZ-CE-SD	3.93	117.05	112.58
1	F	188	SEB	CB-OG-SD	4.14	128.54	119.10
1	A	188	SEB	CZ-CE-SD	4.37	117.56	112.58
1	E	188	SEB	CZ-CE-SD	4.83	118.08	112.58
1	F	188	SEB	CZ-CE-SD	5.87	119.26	112.58

There are no chirality outliers.

There are no torsion outliers.

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](#)

Of 17 ligands modelled in this entry, 8 are monoatomic - leaving 9 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	$\# Z  > 2$
3	ACT	A	410	-	1,3,3	3.87	1 (100%)	0,3,3	0.00	-

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	EDO	A	412	-	3,3,3	0.70	0	2,2,2	0.17	0
4	EDO	A	413	-	3,3,3	0.49	0	2,2,2	0.27	0
3	ACT	B	409	-	1,3,3	2.27	1 (100%)	0,3,3	0.00	-
4	EDO	B	414	-	3,3,3	0.54	0	2,2,2	0.44	0
4	EDO	C	415	-	3,3,3	0.50	0	2,2,2	0.37	0
3	ACT	E	411	-	1,3,3	2.77	1 (100%)	0,3,3	0.00	-
4	EDO	E	416	-	3,3,3	0.65	0	2,2,2	0.21	0
4	EDO	F	417	-	3,3,3	0.55	0	2,2,2	0.18	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
3	ACT	A	410	-	-	0/0/0/0	0/0/0/0
4	EDO	A	412	-	-	0/1/1/1	0/0/0/0
4	EDO	A	413	-	-	0/1/1/1	0/0/0/0
3	ACT	B	409	-	-	0/0/0/0	0/0/0/0
4	EDO	B	414	-	-	0/1/1/1	0/0/0/0
4	EDO	C	415	-	-	0/1/1/1	0/0/0/0
3	ACT	E	411	-	-	0/0/0/0	0/0/0/0
4	EDO	E	416	-	-	0/1/1/1	0/0/0/0
4	EDO	F	417	-	-	0/1/1/1	0/0/0/0

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	409	ACT	CH3-C	2.27	1.52	1.48
3	E	411	ACT	CH3-C	2.77	1.52	1.48
3	A	410	ACT	CH3-C	3.87	1.54	1.48

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
4	A	412	EDO	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	413	EDO	1	0
3	B	409	ACT	1	0
3	E	411	ACT	1	0
4	F	417	EDO	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	320/337 (94%)	-0.78	1 (0%) 94 94	15, 20, 37, 54	0
1	B	320/337 (94%)	-0.71	0 100 100	14, 20, 38, 55	0
1	C	320/337 (94%)	-0.76	1 (0%) 94 94	14, 20, 38, 54	0
1	D	320/337 (94%)	-0.81	1 (0%) 94 94	14, 20, 38, 55	0
1	E	320/337 (94%)	-0.84	1 (0%) 94 94	14, 21, 38, 54	0
1	F	320/337 (94%)	-0.78	1 (0%) 94 94	15, 21, 38, 54	0
All	All	1920/2022 (94%)	-0.78	5 (0%) 94 94	14, 20, 38, 55	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	3	PHE	2.9
1	D	3	PHE	2.7
1	F	3	PHE	2.7
1	E	3	PHE	2.5
1	C	3	PHE	2.2

### 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. 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(Å <sup>2</sup> )	Q<0.9
1	SEB	D	188	16/17	0.90	0.16	-	19,42,53,54	0
1	SEB	F	188	16/17	0.93	0.16	-	19,41,54,54	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
1	SEB	B	188	16/17	0.94	0.17	-	19,40,53,53	0
1	SEB	C	188	16/17	0.95	0.13	-	19,39,54,54	0
1	SEB	E	188	16/17	0.92	0.19	-	20,41,53,54	0
1	SEB	A	188	16/17	0.93	0.15	-	19,39,53,54	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
4	EDO	B	414	4/4	0.97	0.14	3.27	20,25,28,32	0
4	EDO	E	416	4/4	0.96	0.12	3.27	25,25,32,34	0
4	EDO	C	415	4/4	0.97	0.13	2.32	23,24,28,30	0
3	ACT	E	411	4/4	0.95	0.13	2.29	34,34,37,37	0
4	EDO	F	417	4/4	0.97	0.11	1.66	24,32,36,38	0
3	ACT	A	410	4/4	0.94	0.14	1.59	23,25,28,28	0
4	EDO	A	413	4/4	0.98	0.09	-0.92	21,24,24,31	0
3	ACT	B	409	4/4	0.98	0.09	-1.32	18,19,19,20	0
2	CA	F	406	1/1	0.67	0.09	-1.72	65,65,65,65	0
2	CA	E	407	1/1	0.85	0.09	-2.04	65,65,65,65	0
2	CA	B	408	1/1	0.80	0.08	-	64,64,64,64	0
2	CA	A	404	1/1	0.83	0.09	-	64,64,64,64	0
2	CA	A	401	1/1	0.99	0.12	-	21,21,21,21	0
2	CA	D	405	1/1	0.84	0.13	-	65,65,65,65	0
2	CA	D	402	1/1	0.99	0.08	-	22,22,22,22	0
2	CA	C	403	1/1	0.91	0.12	-	60,60,60,60	0
4	EDO	A	412	4/4	0.79	0.36	-	45,53,55,56	0

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

There are no such residues in this entry.