



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 03:37 PM GMT

PDB ID : 4CU4  
Title : FhuA from E. coli in complex with the lasso peptide microcin J25 (MccJ25)  
Authors : Mathavan, I.; Rebuffat, S.; Beis, K.  
Deposited on : 2014-03-17  
Resolution : 2.30 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

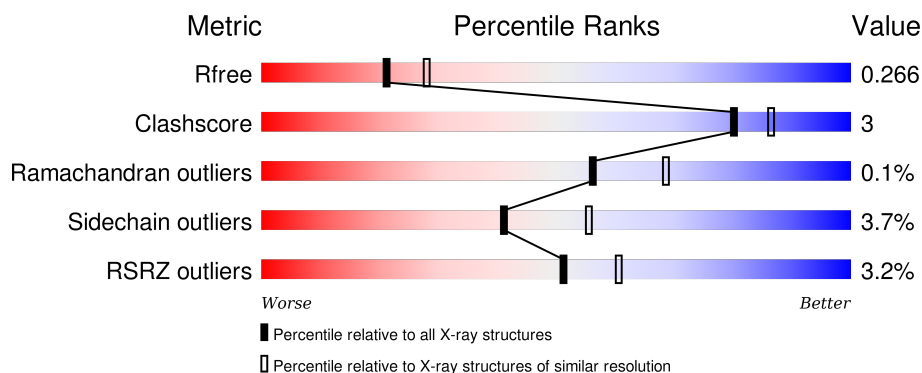
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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	706	<div> <div>3%</div> <div>87%</div> <div>10%</div> <div>••</div> </div>
2	B	21	<div> <div>14%</div> <div>86%</div> <div>14%</div> </div>

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
4	FTT	A	1009	-	-	-	X
6	MYR	A	1014	-	-	-	X
7	LDA	A	1726	-	-	-	X
7	LDA	A	1728	-	-	-	X
7	LDA	A	1729	-	-	-	X
7	LDA	A	1730	-	-	-	X
7	LDA	A	1731	-	-	-	X
7	LDA	A	1732	-	-	-	X
7	LDA	A	1735	-	-	-	X
7	LDA	A	1737	-	-	-	X
7	LDA	A	1739	-	-	-	X
8	3PH	A	1744	-	-	-	X

## 2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 6260 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 FERRICHRONE-IRON RECEPTOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	693	Total	C	N	O	S	0	3	0
			5430	3423	920	1072	15			

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	406	SER	-	EXPRESSION TAG	UNP P06971
A	407	SER	-	EXPRESSION TAG	UNP P06971
A	408	HIS	-	EXPRESSION TAG	UNP P06971
A	409	HIS	-	EXPRESSION TAG	UNP P06971
A	410	HIS	-	EXPRESSION TAG	UNP P06971
A	411	HIS	-	EXPRESSION TAG	UNP P06971
A	412	HIS	-	EXPRESSION TAG	UNP P06971
A	413	HIS	-	EXPRESSION TAG	UNP P06971
A	414	GLY	-	EXPRESSION TAG	UNP P06971
A	415	SER	-	EXPRESSION TAG	UNP P06971
A	416	SER	-	EXPRESSION TAG	UNP P06971

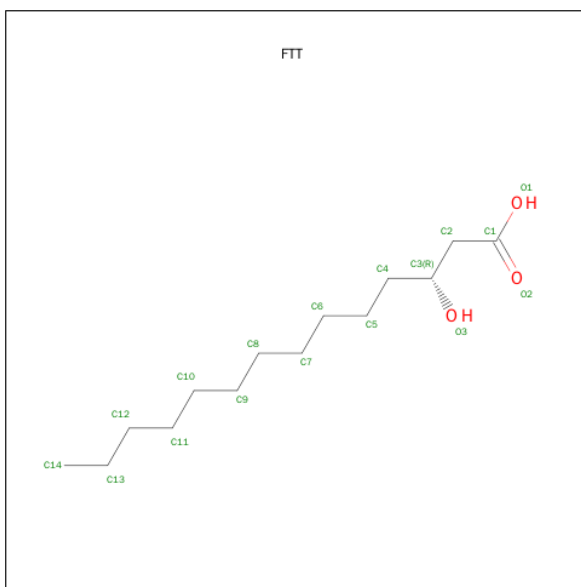
- Molecule 2 is a protein called MICROCIN J25.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	21	Total	C	N	O	0	0	0
			150	101	23	26			

- Molecule 3 is a polymer of unknown type called SUGAR (6-MER).

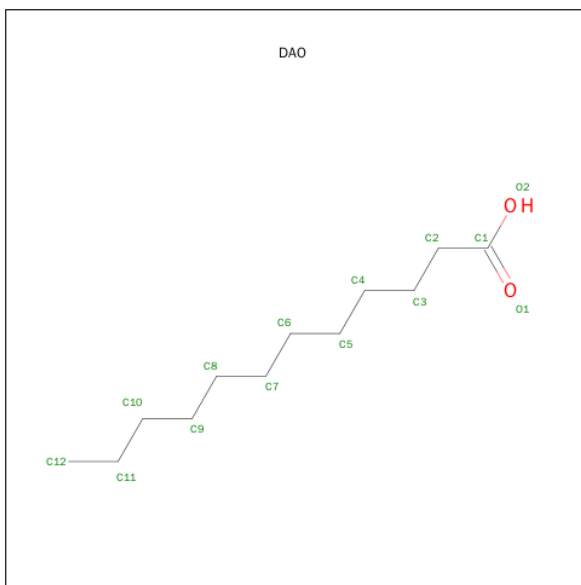
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	6	Total	C	N	O	0	0
			78	42	2	34		

- Molecule 4 is 3-HYDROXY-TETRADECANOIC ACID (three-letter code: FTT) (formula: C<sub>14</sub>H<sub>28</sub>O<sub>3</sub>).



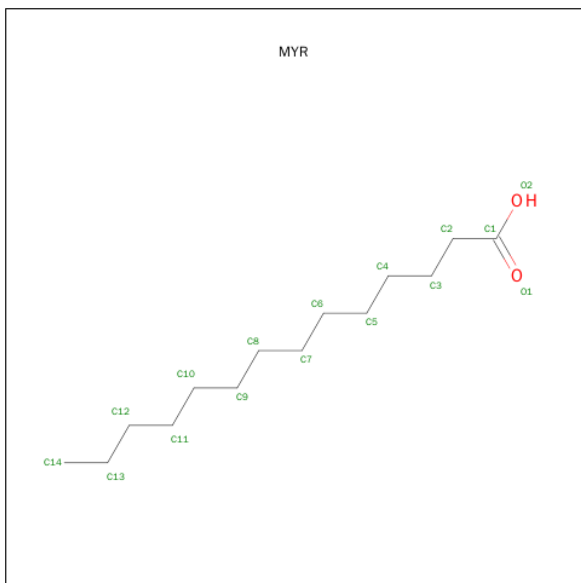
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			16	14	2		
4	A	1	Total	C	O	0	0
			16	14	2		
4	A	1	Total	C	O	0	0
			16	14	2		
4	A	1	Total	C	O	0	0
			16	14	2		

- Molecule 5 is LAURIC ACID (three-letter code: DAO) (formula:  $C_{12}H_{24}O_2$ ).



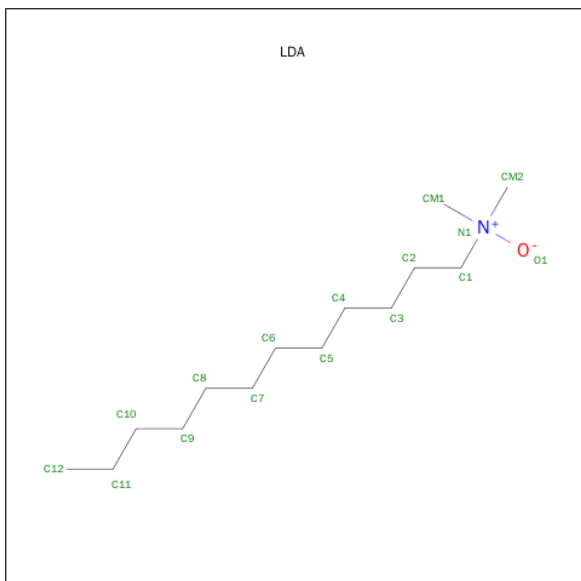
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			13	12	1		

- Molecule 6 is MYRISTIC ACID (three-letter code: MYR) (formula:  $C_{14}H_{28}O_2$ ).



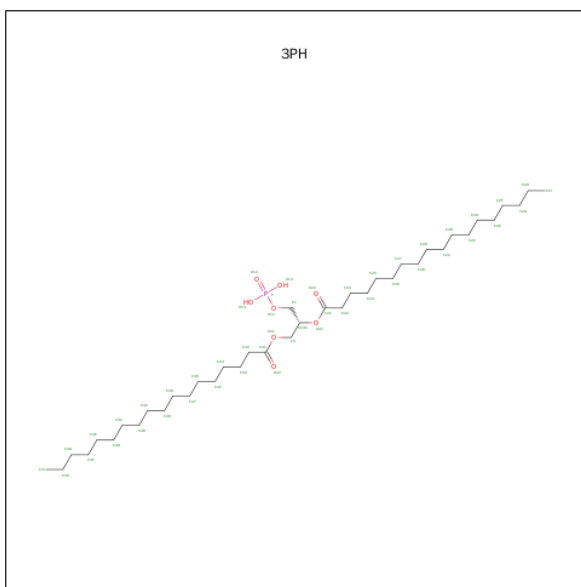
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			15	14	1		

- Molecule 7 is LAURYL DIMETHYLAMINE-N-OXIDE (three-letter code: LDA) (formula:  $C_{14}H_{31}NO$ ).



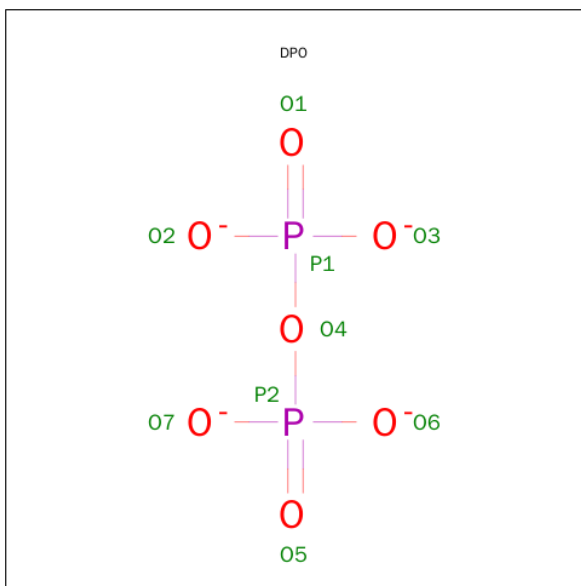
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	A	1	Total	C	N	O	0	0
			16	14	1	1		
7	A	1	Total	C	N	O	0	0
			16	14	1	1		
7	A	1	Total	C	N	O	0	0
			16	14	1	1		
7	A	1	Total	C	N	O	0	0
			16	14	1	1		
7	A	1	Total	C	N	O	0	0
			16	14	1	1		
7	A	1	Total	C	N	O	0	0
			16	14	1	1		
7	A	1	Total	C	N	O	0	0
			16	14	1	1		
7	A	1	Total	C	N	O	0	0
			12	10	1	1		
7	A	1	Total	C	N	O	0	0
			16	14	1	1		
7	A	1	Total	C	N	O	0	0
			13	11	1	1		
7	A	1	Total	C	N	O	0	0
			14	12	1	1		
7	A	1	Total	C			0	0
			9	9				
7	A	1	Total	C			0	0
			10	10				
7	A	1	Total	C	N	O	0	0
			11	9	1	1		
7	A	1	Total	C	N	O	0	0
			11	9	1	1		
7	A	1	Total	C	N	O	0	0
			16	14	1	1		
7	A	1	Total	C	N	O	0	0
			14	12	1	1		
7	A	1	Total	C	N	O	0	0
			12	10	1	1		
7	A	1	Total	C	N	O	0	0
			16	14	1	1		
7	A	1	Total	C	N	O	0	0
			16	14	1	1		

- Molecule 8 is 1,2-DIACYL-GLYCEROL-3-SN-PHOSPHATE (three-letter code: 3PH) (formula: C<sub>39</sub>H<sub>77</sub>O<sub>8</sub>P).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	A	1	Total	C	O	P	0	0
			35	26	8	1		

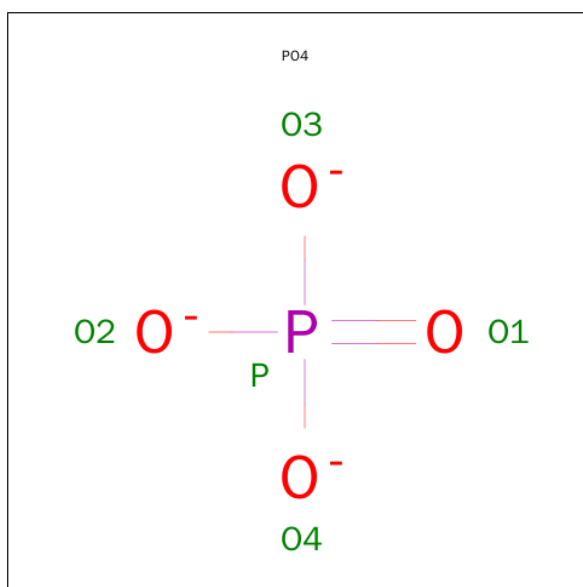
- Molecule 9 is DIPHOSPHATE (three-letter code: DPO) (formula:  $O_7P_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	A	1	Total	O	P	0	0
			9	7	2		
9	A	1	Total	O	P	0	0
			8	6	2		

- Molecule 10 is PHOSPHATE ION (three-letter code: PO4) (formula:  $O_4P$ ).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	A	1	Total	O	P	0	0
			4	3	1		
10	A	1	Total	O	P	0	0
			4	3	1		

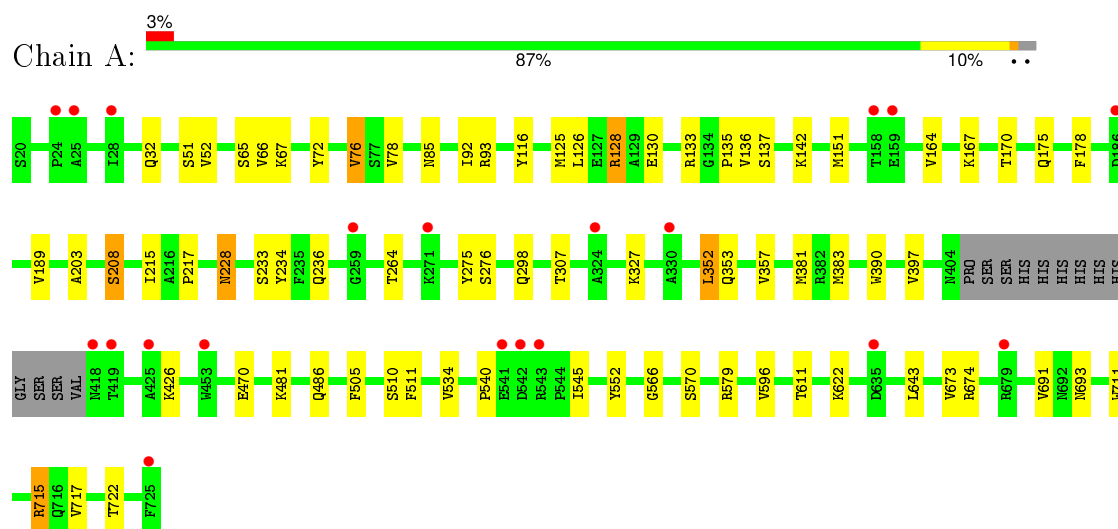
- Molecule 11 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	167	Total	O	0	0
			167	167		
11	B	1	Total	O	0	0
			1	1		

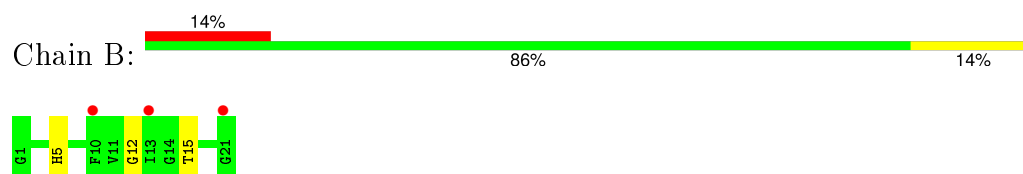
### 3 Residue-property plots [i](#)

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: FERRICHRONE-IRON RECEPTOR



#### • Molecule 2: MICROCIN J25



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	95.98Å 277.05Å 106.76Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.46 – 2.30 30.46 – 2.30	Depositor EDS
% Data completeness (in resolution range)	98.5 (30.46-2.30) 98.5 (30.46-2.30)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.95 (at 2.31Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, $R_{free}$	0.212 , 0.255 0.219 , 0.266	Depositor DCC
$R_{free}$ test set	3172 reflections (5.34%)	DCC
Wilson B-factor (Å <sup>2</sup> )	45.6	Xtriage
Anisotropy	0.853	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 43.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 62614 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	6260	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.34% 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: PA1, LDA, GMH, FTT, PO4, MYR, DAO, 3PH, GCS, KDO, DPO

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	0.53	0/5572	0.71	1/7569 (0.0%)
2	B	0.62	0/156	0.62	0/211
All	All	0.53	0/5728	0.71	1/7780 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	352	LEU	CA-CB-CG	6.30	129.78	115.30

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	5430	0	5157	36	0
2	B	150	0	137	2	0
3	A	78	0	56	0	0
4	A	64	0	106	2	0
5	A	13	0	23	1	0
6	A	15	0	27	0	0
7	A	282	0	517	4	0
8	A	35	0	43	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	A	17	0	0	0	0
10	A	8	0	0	0	0
11	A	167	0	0	2	0
11	B	1	0	0	0	0
All	All	6260	0	6066	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 3.

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:A:228:ASN:C	1:A:228:ASN:HD22	1.90	0.74
1:A:203:ALA:HB3	1:A:208:SER:HB3	1.76	0.68
1:A:66:VAL:HG23	11:A:3015:HOH:O	1.97	0.64
7:A:1732:LDA:H52	8:A:1744:3PH:H222	1.82	0.62
1:A:51:SER:OG	1:A:133:ARG:NH2	2.32	0.62
1:A:505:PHE:HE2	7:A:1741:LDA:H72	1.65	0.60
1:A:643:LEU:HD22	1:A:673:VAL:HG12	1.84	0.60
1:A:52:VAL:HG22	1:A:130:GLU:HG2	1.86	0.57
1:A:470:GLU:HG3	1:A:481:LYS:HG2	1.89	0.55
2:B:12:GLY:H	2:B:15:THR:HB	1.70	0.55
1:A:135:PRO:HB3	1:A:510:SER:HB3	1.89	0.55
8:A:1744:3PH:H342	8:A:1744:3PH:H252	1.90	0.54
1:A:298:GLN:HE22	4:A:1009:FTT:H143	1.72	0.54
1:A:137:SER:HB2	1:A:510:SER:HA	1.90	0.54
1:A:234:TYR:OH	1:A:236:GLN:NE2	2.42	0.53
1:A:505:PHE:HD1	1:A:534:VAL:HG22	1.74	0.53
1:A:164:VAL:HG13	7:A:1726:LDA:H121	1.91	0.52
1:A:228:ASN:ND2	1:A:228:ASN:C	2.59	0.52
8:A:1744:3PH:H322	8:A:1744:3PH:H231	1.92	0.51
1:A:65:SER:OG	1:A:67:LYS:HG2	2.11	0.51
1:A:32:GLN:O	1:A:128:ARG:NH2	2.44	0.50
1:A:217:PRO:HD2	1:A:233:SER:OG	2.11	0.50
1:A:167:LYS:HB2	1:A:175:GLN:HB3	1.93	0.49
4:A:1009:FTT:H101	5:A:1012:DAO:H61	1.94	0.49
1:A:381[B]:MET:HE1	1:A:383:MET:HB2	1.96	0.48
1:A:486:GLN:HG2	1:A:511:PHE:CG	2.49	0.48
1:A:72:TYR:HB3	1:A:596:VAL:HG12	1.95	0.47
1:A:276:SER:O	1:A:307:THR:HG23	2.16	0.46
1:A:126:LEU:HD22	1:A:151:MET:HB3	1.98	0.44

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:691:VAL:HG13	1:A:717:VAL:HG22	1.99	0.44
1:A:65:SER:HB3	1:A:85:ASN:HD21	1.83	0.43
1:A:76:VAL:HG13	1:A:92:ILE:HG12	1.99	0.43
1:A:116:TYR:CZ	2:B:5:HIS:HB3	2.54	0.43
1:A:540:PRO:HD2	1:A:545:ILE:O	2.19	0.42
1:A:693:ASN:O	1:A:715:ARG:HB2	2.19	0.42
1:A:67:LYS:HD3	1:A:78:VAL:O	2.19	0.42
1:A:264:THR:HA	1:A:711:TRP:CD1	2.54	0.42
1:A:142:LYS:NZ	11:A:3066:HOH:O	2.48	0.42
1:A:93:ARG:HG2	1:A:552:TYR:CZ	2.55	0.42
1:A:390:TRP:CE2	1:A:426:LYS:HB2	2.55	0.41
1:A:178:PHE:HD2	7:A:1726:LDA:H92	1.86	0.41

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	692/706 (98%)	667 (96%)	24 (4%)	1 (0%)	56	68
2	B	19/21 (90%)	18 (95%)	1 (5%)	0	100	100
All	All	711/727 (98%)	685 (96%)	25 (4%)	1 (0%)	56	68

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	566	GLY

### 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	577/586 (98%)	555 (96%)	22 (4%)	40	54
2	B	14/14 (100%)	14 (100%)	0	100	100
All	All	591/600 (98%)	569 (96%)	22 (4%)	41	55

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	76	VAL
1	A	125	MET
1	A	128	ARG
1	A	136	VAL
1	A	170	THR
1	A	189	VAL
1	A	208	SER
1	A	215	ILE
1	A	228	ASN
1	A	275	TYR
1	A	327	LYS
1	A	352	LEU
1	A	353	GLN
1	A	357	VAL
1	A	397	VAL
1	A	570	SER
1	A	579	ARG
1	A	611	THR
1	A	622	LYS
1	A	674	ARG
1	A	715	ARG
1	A	722	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	85	ASN
1	A	100	GLN
1	A	228	ASN
1	A	236	GLN
1	A	298	GLN
1	A	299	ASN
1	A	309	GLN
1	A	692	ASN
1	A	716	GLN

### 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 ⓘ

6 carbohydrates 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	$\# Z  > 2$
3	PA1	A	1000	9,3,4	11,11,12	0.72	1 (9%)	13,15,17	1.54	3 (23%)
3	GCS	A	1001	10,3,4	11,11,12	1.04	1 (9%)	13,15,17	1.32	2 (15%)
3	KDO	A	1002	3	12,15,16	0.39	0	12,21,24	0.99	1 (8%)
3	KDO	A	1003	3	12,15,16	0.34	0	12,21,24	0.69	0
3	GMH	A	1004	9,3	13,13,14	0.37	0	17,18,20	1.11	1 (5%)
3	GMH	A	1005	10,3	13,13,14	0.41	0	17,18,20	1.13	1 (5%)

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	PA1	A	1000	9,3,4	-	0/2/19/22	0/1/1/1
3	GCS	A	1001	10,3,4	-	0/2/19/22	0/1/1/1
3	KDO	A	1002	3	-	0/6/26/30	0/1/1/1
3	KDO	A	1003	3	-	0/6/26/30	0/1/1/1
3	GMH	A	1004	9,3	-	0/6/23/26	0/1/1/1
3	GMH	A	1005	10,3	-	0/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1000	PA1	C1-C2	2.32	1.55	1.52
3	A	1001	GCS	C1-C2	3.10	1.56	1.52

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1000	PA1	O3-C3-C2	-3.32	103.70	109.87
3	A	1001	GCS	O3-C3-C4	-2.76	104.13	110.34
3	A	1000	PA1	O3-C3-C4	-2.41	104.92	110.34
3	A	1000	PA1	C3-C2-N2	2.00	114.34	110.78
3	A	1002	KDO	C3-C4-C5	2.07	113.38	110.56
3	A	1001	GCS	O3-C3-C2	3.28	115.96	109.87
3	A	1004	GMH	C1-O5-C5	3.89	117.84	111.52
3	A	1005	GMH	C1-O5-C5	4.13	118.23	111.52

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.6 Ligand geometry [i](#)

31 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	$\# Z  > 2$
4	FTT	A	1009	3	14,15,16	0.26	0	15,15,17	0.40	0
4	FTT	A	1010	3	14,15,16	0.21	0	15,15,17	0.68	1 (6%)
4	FTT	A	1011	3,5	14,15,16	0.27	0	15,15,17	0.47	0
5	DAO	A	1012	4	12,12,13	0.56	0	10,11,13	0.18	0
4	FTT	A	1013	3,6	14,15,16	0.18	0	15,15,17	0.28	0
6	MYR	A	1014	4	14,14,15	0.54	0	12,13,15	0.24	0
7	LDA	A	1726	-	15,15,15	0.27	0	16,17,17	0.43	0
7	LDA	A	1727	-	15,15,15	0.24	0	16,17,17	0.40	0
7	LDA	A	1728	-	15,15,15	0.30	0	16,17,17	0.21	0
7	LDA	A	1729	-	15,15,15	0.35	0	16,17,17	0.30	0
7	LDA	A	1730	-	15,15,15	0.33	0	16,17,17	0.24	0
7	LDA	A	1731	-	15,15,15	0.31	0	16,17,17	0.37	0
7	LDA	A	1732	-	15,15,15	0.26	0	16,17,17	1.08	2 (12%)
7	LDA	A	1733	-	11,11,15	0.34	0	12,13,17	0.42	0
7	LDA	A	1734	-	15,15,15	0.39	0	16,17,17	0.24	0
7	LDA	A	1735	-	12,12,15	0.34	0	13,14,17	0.39	0
7	LDA	A	1736	-	13,13,15	0.42	0	14,15,17	0.23	0
7	LDA	A	1737	-	8,8,15	0.21	0	7,7,17	0.09	0
7	LDA	A	1738	-	9,9,15	0.20	0	8,8,17	0.11	0
7	LDA	A	1739	-	10,10,15	0.29	0	11,12,17	0.33	0
7	LDA	A	1740	-	10,10,15	0.37	0	11,12,17	0.54	0
7	LDA	A	1741	-	15,15,15	0.21	0	16,17,17	0.41	0
7	LDA	A	1742	-	13,13,15	0.35	0	14,15,17	0.39	0
7	LDA	A	1743	-	11,11,15	0.38	0	12,13,17	0.42	0
8	3PH	A	1744	-	34,34,47	0.46	0	37,39,52	0.52	0
7	LDA	A	1745	-	15,15,15	0.23	0	16,17,17	0.30	0
7	LDA	A	1746	-	15,15,15	0.19	0	16,17,17	0.47	0
9	DPO	A	2000	3	8,8,8	1.66	3 (37%)	12,13,13	3.51	3 (25%)
10	PO4	A	2001	3	0,3,4	0.00	-	0,3,6	0.00	-
9	DPO	A	2004	3	4,7,8	1.47	0	6,10,13	0.96	0
10	PO4	A	2005	3	0,3,4	0.00	-	0,3,6	0.00	-

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
4	FTT	A	1009	3	-	0/14/14/15	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	FTT	A	1010	3	-	0/14/14/15	0/0/0/0
4	FTT	A	1011	3,5	-	0/14/14/15	0/0/0/0
5	DAO	A	1012	4	-	0/9/10/11	0/0/0/0
4	FTT	A	1013	3,6	-	0/14/14/15	0/0/0/0
6	MYR	A	1014	4	-	0/11/12/13	0/0/0/0
7	LDA	A	1726	-	-	0/13/13/13	0/0/0/0
7	LDA	A	1727	-	-	0/13/13/13	0/0/0/0
7	LDA	A	1728	-	-	0/13/13/13	0/0/0/0
7	LDA	A	1729	-	-	0/13/13/13	0/0/0/0
7	LDA	A	1730	-	-	0/13/13/13	0/0/0/0
7	LDA	A	1731	-	-	0/13/13/13	0/0/0/0
7	LDA	A	1732	-	-	0/13/13/13	0/0/0/0
7	LDA	A	1733	-	-	0/9/9/13	0/0/0/0
7	LDA	A	1734	-	-	0/13/13/13	0/0/0/0
7	LDA	A	1735	-	-	0/10/10/13	0/0/0/0
7	LDA	A	1736	-	-	0/11/11/13	0/0/0/0
7	LDA	A	1737	-	-	0/6/6/13	0/0/0/0
7	LDA	A	1738	-	-	0/7/7/13	0/0/0/0
7	LDA	A	1739	-	-	0/8/8/13	0/0/0/0
7	LDA	A	1740	-	-	0/8/8/13	0/0/0/0
7	LDA	A	1741	-	-	0/13/13/13	0/0/0/0
7	LDA	A	1742	-	-	0/11/11/13	0/0/0/0
7	LDA	A	1743	-	-	0/9/9/13	0/0/0/0
8	3PH	A	1744	-	-	0/36/36/49	0/0/0/0
7	LDA	A	1745	-	-	0/13/13/13	0/0/0/0
7	LDA	A	1746	-	-	0/13/13/13	0/0/0/0
9	DPO	A	2000	3	-	0/6/6/6	0/0/0/0
10	PO4	A	2001	3	-	0/0/0/0	0/0/0/0
9	DPO	A	2004	3	-	0/3/5/6	0/0/0/0
10	PO4	A	2005	3	-	0/0/0/0	0/0/0/0

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	A	2000	DPO	P2-O6	2.18	1.55	1.51
9	A	2000	DPO	P2-O5	2.21	1.56	1.51
9	A	2000	DPO	P1-O1	2.49	1.56	1.51

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	A	2000	DPO	O2-P1-O1	-10.47	88.43	112.76

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	A	1732	LDA	CM1-N1-C1	-2.53	101.63	109.77
4	A	1010	FTT	C3-C2-C1	-2.32	106.86	111.90
9	A	2000	DPO	O3-P1-O4	2.50	112.97	105.19
7	A	1732	LDA	CM2-N1-C1	2.95	119.27	109.77
9	A	2000	DPO	O3-P1-O1	4.85	124.02	112.76

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1009	FTT	2	0
5	A	1012	DAO	1	0
7	A	1726	LDA	2	0
7	A	1732	LDA	1	0
7	A	1741	LDA	1	0
8	A	1744	3PH	3	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	693/706 (98%)	-0.11	20 (2%) 55 64	36, 53, 75, 104	0
2	B	21/21 (100%)	0.53	3 (14%) 4 6	56, 70, 80, 95	0
All	All	714/727 (98%)	-0.09	23 (3%) 51 60	36, 53, 76, 104	0

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	725	PHE	3.9
1	A	158	THR	3.9
1	A	28	ILE	3.6
1	A	418	ASN	3.1
1	A	635	ASP	3.0
2	B	10	PHE	3.0
1	A	542	ASP	2.7
1	A	419	THR	2.6
1	A	453	TRP	2.6
1	A	25	ALA	2.6
1	A	159	GLU	2.6
1	A	543	ARG	2.4
1	A	425	ALA	2.4
1	A	24	PRO	2.4
1	A	186	ASP	2.3
1	A	541	GLU	2.3
1	A	259	GLY	2.3
2	B	21	GLY	2.2
1	A	330	ALA	2.2
1	A	679	ARG	2.2
1	A	324	ALA	2.2
2	B	13	ILE	2.1
1	A	271	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 ⓘ

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
3	GMH	A	1004	13/14	0.94	0.15	1.76	74,81,92,100	0
3	KDO	A	1002	15/16	0.95	0.12	0.08	66,73,77,79	0
3	GCS	A	1001	11/12	0.95	0.08	-	50,54,67,72	0
3	PA1	A	1000	11/12	0.97	0.07	-	47,47,51,51	0
3	GMH	A	1005	13/14	0.85	0.28	-	109,117,120,127	0
3	KDO	A	1003	15/16	0.90	0.21	-	69,81,89,90	0

## 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. 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
6	MYR	A	1014	15/16	0.76	0.40	18.43	81,86,92,92	0
7	LDA	A	1739	11/16	0.72	0.32	8.85	55,67,85,86	0
7	LDA	A	1735	13/16	0.73	0.29	8.67	53,65,91,92	0
7	LDA	A	1732	16/16	0.74	0.27	8.33	64,77,91,91	0
7	LDA	A	1726	16/16	0.86	0.25	7.09	67,72,76,78	0
8	3PH	A	1744	35/48	0.71	0.25	4.90	55,85,114,114	0
7	LDA	A	1728	16/16	0.73	0.21	4.19	46,60,89,89	0
7	LDA	A	1729	16/16	0.55	0.28	3.95	70,74,97,99	0
7	LDA	A	1730	16/16	0.82	0.25	3.71	71,75,89,90	0
7	LDA	A	1737	9/16	0.81	0.32	3.48	79,81,83,84	0
4	FTT	A	1009	16/17	0.94	0.15	2.08	46,63,73,73	0
7	LDA	A	1731	16/16	0.82	0.32	2.05	81,83,91,92	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
5	DAO	A	1012	13/14	0.95	0.14	1.55	45,61,69,69	0
7	LDA	A	1727	16/16	0.93	0.17	1.32	51,58,69,71	0
7	LDA	A	1734	16/16	0.80	0.23	0.94	64,77,96,96	0
4	FTT	A	1011	16/17	0.96	0.11	0.06	50,53,57,58	0
7	LDA	A	1741	16/16	0.90	0.14	-0.25	64,68,79,82	0
7	LDA	A	1738	10/16	0.78	0.40	-	87,92,94,94	0
9	DPO	A	2004	8/9	0.90	0.14	-	88,98,105,107	0
10	PO4	A	2001	4/5	0.98	0.05	-	60,66,66,71	0
7	LDA	A	1745	16/16	0.69	0.27	-	83,85,104,106	0
7	LDA	A	1742	14/16	0.62	0.31	-	78,88,103,104	0
4	FTT	A	1010	16/17	0.94	0.13	-	52,60,74,74	0
10	PO4	A	2005	4/5	0.84	0.29	-	133,134,134,134	0
7	LDA	A	1746	16/16	0.60	0.34	-	76,79,99,101	0
7	LDA	A	1743	12/16	0.78	0.22	-	68,74,90,91	0
7	LDA	A	1736	14/16	0.61	0.30	-	63,77,91,92	0
4	FTT	A	1013	16/17	0.91	0.15	-	59,68,73,79	0
7	LDA	A	1733	12/16	0.65	0.27	-	91,99,106,106	0
9	DPO	A	2000	9/9	0.89	0.19	-	47,56,71,74	0
7	LDA	A	1740	11/16	0.72	0.34	-	80,84,95,95	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.