



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 15, 2020 – 06:54 am BST

PDB ID : 4R88  
Title : Crystal structure of 5-methylcytosine deaminase from *Klebsiella pneumoniae*  
liganded with 5-fluorocytosine  
Authors : Fedorov, A.A.; Fedorov, E.V.; Hitchcock, D.S.; Raushel, F.M.; Almo, S.C.  
Deposited on : 2014-08-29  
Resolution : 1.95 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 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	:	2.11
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)	:	2.11

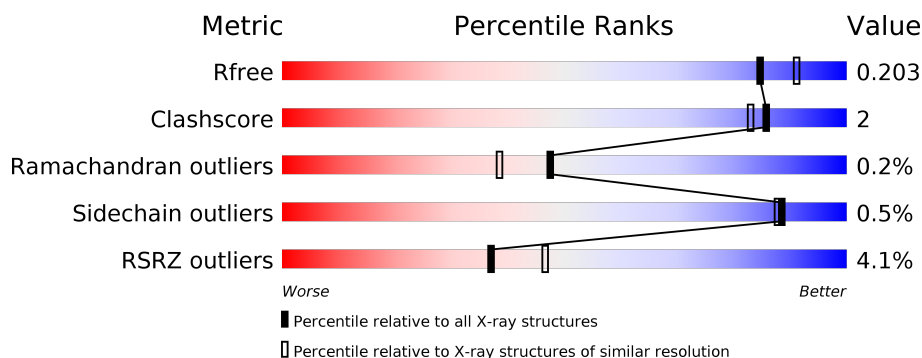
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.95 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	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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	431	<div> <div>%</div> <div> <div></div> <div>90%</div> <div>5%</div> <div>5%</div> </div> </div>
1	B	431	<div> <div>3%</div> <div> <div></div> <div>90%</div> <div>6%</div> <div>•</div> </div> </div>
1	C	431	<div> <div>6%</div> <div> <div></div> <div>91%</div> <div>•</div> <div>5%</div> </div> </div>
1	D	431	<div> <div>3%</div> <div> <div></div> <div>92%</div> <div>•</div> <div>•</div> </div> </div>
1	E	431	<div> <div>%</div> <div> <div></div> <div>89%</div> <div>7%</div> <div>5%</div> </div> </div>
1	F	431	<div> <div>10%</div> <div> <div></div> <div>91%</div> <div>•</div> <div>5%</div> </div> </div>

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 20876 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 Cytosine deaminase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	411	Total	C	N	O	S	0	4	0
			3254	2029	601	606	18			
1	B	413	Total	C	N	O	S	0	3	0
			3259	2029	603	609	18			
1	C	411	Total	C	N	O	S	0	0	0
			3217	2006	592	601	18			
1	D	412	Total	C	N	O	S	0	1	0
			3233	2015	597	603	18			
1	E	411	Total	C	N	O	S	0	0	0
			3217	2006	592	601	18			
1	F	411	Total	C	N	O	S	0	2	0
			3236	2016	597	605	18			

There are 114 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	413	ASN	-	EXPRESSION TAG	UNP W8V4R8
A	414	SER	-	EXPRESSION TAG	UNP W8V4R8
A	415	SER	-	EXPRESSION TAG	UNP W8V4R8
A	416	SER	-	EXPRESSION TAG	UNP W8V4R8
A	417	VAL	-	EXPRESSION TAG	UNP W8V4R8
A	418	ASP	-	EXPRESSION TAG	UNP W8V4R8
A	419	LYS	-	EXPRESSION TAG	UNP W8V4R8
A	420	LEU	-	EXPRESSION TAG	UNP W8V4R8
A	421	ALA	-	EXPRESSION TAG	UNP W8V4R8
A	422	ALA	-	EXPRESSION TAG	UNP W8V4R8
A	423	ALA	-	EXPRESSION TAG	UNP W8V4R8
A	424	LEU	-	EXPRESSION TAG	UNP W8V4R8
A	425	GLU	-	EXPRESSION TAG	UNP W8V4R8
A	426	HIS	-	EXPRESSION TAG	UNP W8V4R8
A	427	HIS	-	EXPRESSION TAG	UNP W8V4R8
A	428	HIS	-	EXPRESSION TAG	UNP W8V4R8
A	429	HIS	-	EXPRESSION TAG	UNP W8V4R8

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	430	HIS	-	EXPRESSION TAG	UNP W8V4R8
A	431	HIS	-	EXPRESSION TAG	UNP W8V4R8
B	413	ASN	-	EXPRESSION TAG	UNP W8V4R8
B	414	SER	-	EXPRESSION TAG	UNP W8V4R8
B	415	SER	-	EXPRESSION TAG	UNP W8V4R8
B	416	SER	-	EXPRESSION TAG	UNP W8V4R8
B	417	VAL	-	EXPRESSION TAG	UNP W8V4R8
B	418	ASP	-	EXPRESSION TAG	UNP W8V4R8
B	419	LYS	-	EXPRESSION TAG	UNP W8V4R8
B	420	LEU	-	EXPRESSION TAG	UNP W8V4R8
B	421	ALA	-	EXPRESSION TAG	UNP W8V4R8
B	422	ALA	-	EXPRESSION TAG	UNP W8V4R8
B	423	ALA	-	EXPRESSION TAG	UNP W8V4R8
B	424	LEU	-	EXPRESSION TAG	UNP W8V4R8
B	425	GLU	-	EXPRESSION TAG	UNP W8V4R8
B	426	HIS	-	EXPRESSION TAG	UNP W8V4R8
B	427	HIS	-	EXPRESSION TAG	UNP W8V4R8
B	428	HIS	-	EXPRESSION TAG	UNP W8V4R8
B	429	HIS	-	EXPRESSION TAG	UNP W8V4R8
B	430	HIS	-	EXPRESSION TAG	UNP W8V4R8
B	431	HIS	-	EXPRESSION TAG	UNP W8V4R8
C	413	ASN	-	EXPRESSION TAG	UNP W8V4R8
C	414	SER	-	EXPRESSION TAG	UNP W8V4R8
C	415	SER	-	EXPRESSION TAG	UNP W8V4R8
C	416	SER	-	EXPRESSION TAG	UNP W8V4R8
C	417	VAL	-	EXPRESSION TAG	UNP W8V4R8
C	418	ASP	-	EXPRESSION TAG	UNP W8V4R8
C	419	LYS	-	EXPRESSION TAG	UNP W8V4R8
C	420	LEU	-	EXPRESSION TAG	UNP W8V4R8
C	421	ALA	-	EXPRESSION TAG	UNP W8V4R8
C	422	ALA	-	EXPRESSION TAG	UNP W8V4R8
C	423	ALA	-	EXPRESSION TAG	UNP W8V4R8
C	424	LEU	-	EXPRESSION TAG	UNP W8V4R8
C	425	GLU	-	EXPRESSION TAG	UNP W8V4R8
C	426	HIS	-	EXPRESSION TAG	UNP W8V4R8
C	427	HIS	-	EXPRESSION TAG	UNP W8V4R8
C	428	HIS	-	EXPRESSION TAG	UNP W8V4R8
C	429	HIS	-	EXPRESSION TAG	UNP W8V4R8
C	430	HIS	-	EXPRESSION TAG	UNP W8V4R8
C	431	HIS	-	EXPRESSION TAG	UNP W8V4R8
D	413	ASN	-	EXPRESSION TAG	UNP W8V4R8
D	414	SER	-	EXPRESSION TAG	UNP W8V4R8

*Continued on next page...*

*Continued from previous page...*

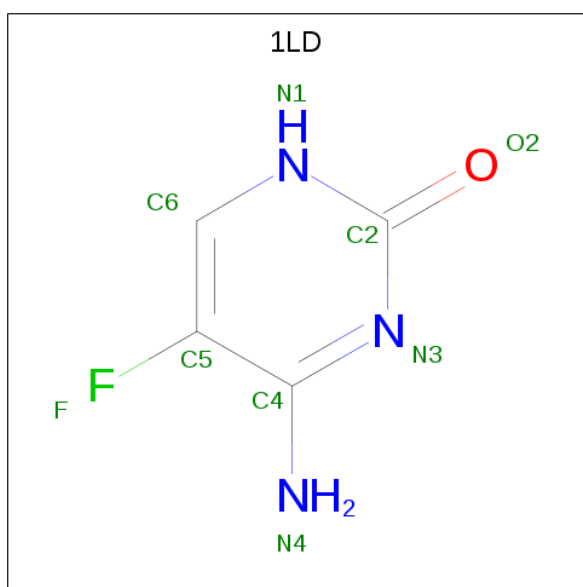
Chain	Residue	Modelled	Actual	Comment	Reference
D	415	SER	-	EXPRESSION TAG	UNP W8V4R8
D	416	SER	-	EXPRESSION TAG	UNP W8V4R8
D	417	VAL	-	EXPRESSION TAG	UNP W8V4R8
D	418	ASP	-	EXPRESSION TAG	UNP W8V4R8
D	419	LYS	-	EXPRESSION TAG	UNP W8V4R8
D	420	LEU	-	EXPRESSION TAG	UNP W8V4R8
D	421	ALA	-	EXPRESSION TAG	UNP W8V4R8
D	422	ALA	-	EXPRESSION TAG	UNP W8V4R8
D	423	ALA	-	EXPRESSION TAG	UNP W8V4R8
D	424	LEU	-	EXPRESSION TAG	UNP W8V4R8
D	425	GLU	-	EXPRESSION TAG	UNP W8V4R8
D	426	HIS	-	EXPRESSION TAG	UNP W8V4R8
D	427	HIS	-	EXPRESSION TAG	UNP W8V4R8
D	428	HIS	-	EXPRESSION TAG	UNP W8V4R8
D	429	HIS	-	EXPRESSION TAG	UNP W8V4R8
D	430	HIS	-	EXPRESSION TAG	UNP W8V4R8
D	431	HIS	-	EXPRESSION TAG	UNP W8V4R8
E	413	ASN	-	EXPRESSION TAG	UNP W8V4R8
E	414	SER	-	EXPRESSION TAG	UNP W8V4R8
E	415	SER	-	EXPRESSION TAG	UNP W8V4R8
E	416	SER	-	EXPRESSION TAG	UNP W8V4R8
E	417	VAL	-	EXPRESSION TAG	UNP W8V4R8
E	418	ASP	-	EXPRESSION TAG	UNP W8V4R8
E	419	LYS	-	EXPRESSION TAG	UNP W8V4R8
E	420	LEU	-	EXPRESSION TAG	UNP W8V4R8
E	421	ALA	-	EXPRESSION TAG	UNP W8V4R8
E	422	ALA	-	EXPRESSION TAG	UNP W8V4R8
E	423	ALA	-	EXPRESSION TAG	UNP W8V4R8
E	424	LEU	-	EXPRESSION TAG	UNP W8V4R8
E	425	GLU	-	EXPRESSION TAG	UNP W8V4R8
E	426	HIS	-	EXPRESSION TAG	UNP W8V4R8
E	427	HIS	-	EXPRESSION TAG	UNP W8V4R8
E	428	HIS	-	EXPRESSION TAG	UNP W8V4R8
E	429	HIS	-	EXPRESSION TAG	UNP W8V4R8
E	430	HIS	-	EXPRESSION TAG	UNP W8V4R8
E	431	HIS	-	EXPRESSION TAG	UNP W8V4R8
F	413	ASN	-	EXPRESSION TAG	UNP W8V4R8
F	414	SER	-	EXPRESSION TAG	UNP W8V4R8
F	415	SER	-	EXPRESSION TAG	UNP W8V4R8
F	416	SER	-	EXPRESSION TAG	UNP W8V4R8
F	417	VAL	-	EXPRESSION TAG	UNP W8V4R8
F	418	ASP	-	EXPRESSION TAG	UNP W8V4R8

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	419	LYS	-	EXPRESSION TAG	UNP W8V4R8
F	420	LEU	-	EXPRESSION TAG	UNP W8V4R8
F	421	ALA	-	EXPRESSION TAG	UNP W8V4R8
F	422	ALA	-	EXPRESSION TAG	UNP W8V4R8
F	423	ALA	-	EXPRESSION TAG	UNP W8V4R8
F	424	LEU	-	EXPRESSION TAG	UNP W8V4R8
F	425	GLU	-	EXPRESSION TAG	UNP W8V4R8
F	426	HIS	-	EXPRESSION TAG	UNP W8V4R8
F	427	HIS	-	EXPRESSION TAG	UNP W8V4R8
F	428	HIS	-	EXPRESSION TAG	UNP W8V4R8
F	429	HIS	-	EXPRESSION TAG	UNP W8V4R8
F	430	HIS	-	EXPRESSION TAG	UNP W8V4R8
F	431	HIS	-	EXPRESSION TAG	UNP W8V4R8

- Molecule 2 is 5-fluorocytosine (three-letter code: 1LD) (formula:  $C_4H_4FN_3O$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	F	N	O	0	0
			9	4	1	3	1		
2	B	1	Total	C	F	N	O	0	0
			9	4	1	3	1		
2	C	1	Total	C	F	N	O	0	0
			9	4	1	3	1		
2	D	1	Total	C	F	N	O	0	0
			9	4	1	3	1		
2	E	1	Total	C	F	N	O	0	0
			9	4	1	3	1		

Continued on next page...

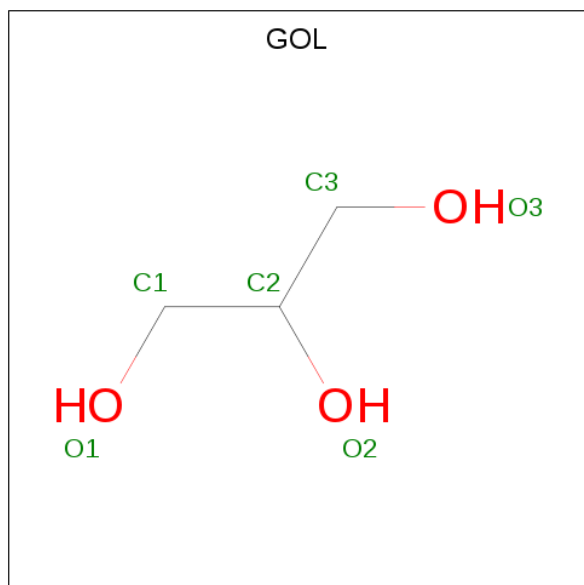
*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	F	1	Total	C	F	N	O	0	0
			9	4	1	3	1		

- Molecule 3 is FE (II) ION (three-letter code: FE2) (formula: Fe).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	D	1	Total	Fe	0	0
			1	1		
3	E	1	Total	Fe	0	0
			1	1		
3	B	1	Total	Fe	0	0
			1	1		
3	C	1	Total	Fe	0	0
			1	1		
3	A	1	Total	Fe	0	0
			1	1		
3	F	1	Total	Fe	0	0
			1	1		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



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

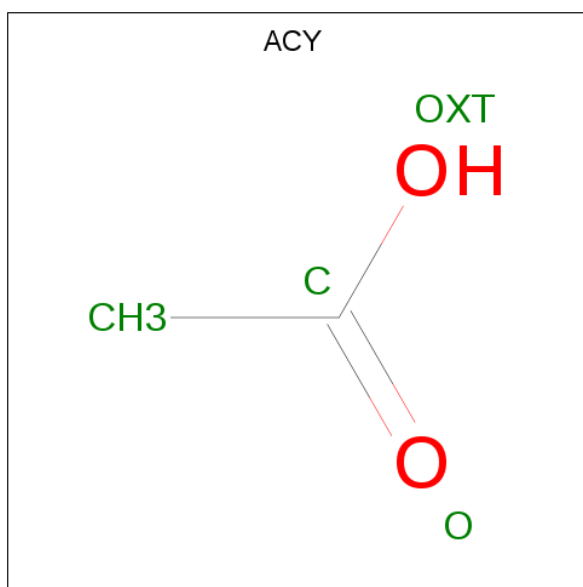
*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		
4	C	1	Total	C	O	0	0
			6	3	3		
4	C	1	Total	C	O	0	0
			6	3	3		
4	C	1	Total	C	O	0	0
			6	3	3		
4	D	1	Total	C	O	0	0
			6	3	3		
4	D	1	Total	C	O	0	0
			6	3	3		
4	D	1	Total	C	O	0	0
			6	3	3		
4	E	1	Total	C	O	0	0
			6	3	3		
4	E	1	Total	C	O	0	1
			12	6	6		
4	E	1	Total	C	O	0	0
			6	3	3		
4	E	1	Total	C	O	0	0
			6	3	3		
4	F	1	Total	C	O	0	0
			6	3	3		
4	F	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is ACETIC ACID (three-letter code: ACY) (formula: C<sub>2</sub>H<sub>4</sub>O<sub>2</sub>).





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

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



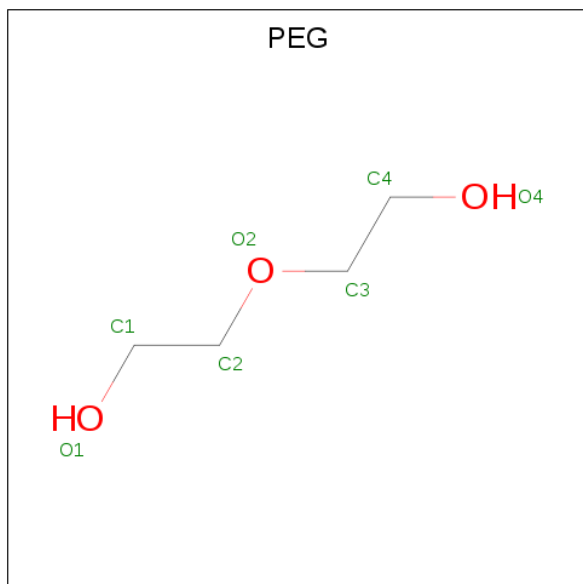
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			4	2	2		
6	A	1	Total	C	O	0	0
			4	2	2		
6	A	1	Total	C	O	0	0
			4	2	2		
6	A	1	Total	C	O	0	0
			4	2	2		
6	B	1	Total	C	O	0	0
			4	2	2		
6	B	1	Total	C	O	0	0
			4	2	2		
6	B	1	Total	C	O	0	0
			4	2	2		
6	B	1	Total	C	O	0	0
			4	2	2		
6	C	1	Total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

*Continued on next page...*

*Continued from previous page...*

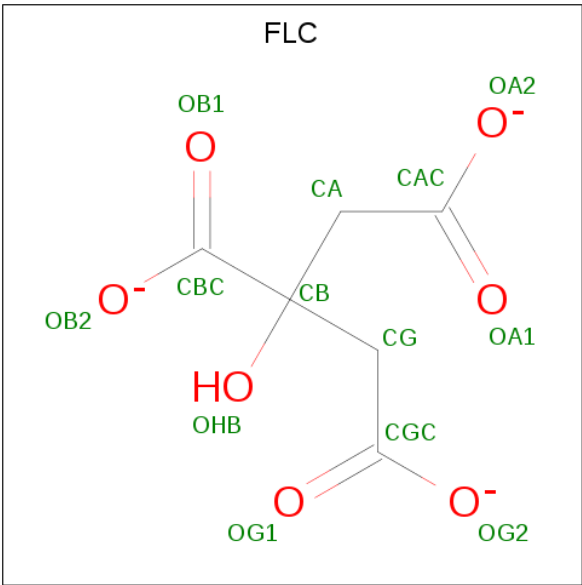
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	E	1	Total	C	O	0	0
			4	2	2		

- Molecule 7 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	B	1	Total	C	O	0	0
			7	4	3		
7	F	1	Total	C	O	0	0
			7	4	3		

- Molecule 8 is CITRATE ANION (three-letter code: FLC) (formula:  $C_6H_5O_7$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	E	1	Total	C	O	0	0
			13	6	7		

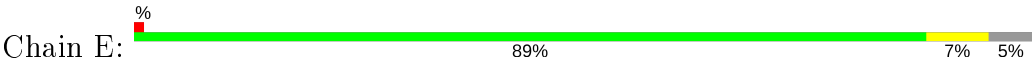
- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	209	Total	O	0	2
			211	211		
9	B	197	Total	O	0	0
			197	197		
9	C	182	Total	O	0	2
			184	184		
9	D	191	Total	O	0	2
			193	193		
9	E	204	Total	O	0	4
			208	208		
9	F	165	Total	O	0	1
			166	166		



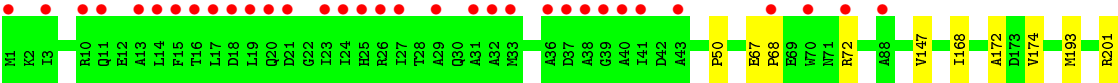
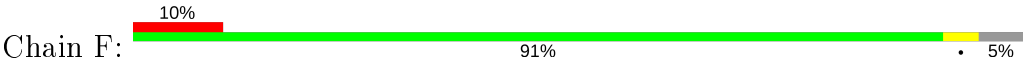
LEU  
ALA  
ALA  
ALA  
LEU  
GLU  
HIS  
HIS  
HIS  
HIS  
HIS

● Molecule 1: Cytosine deaminase



ALA  
ASN  
SER  
SER  
VAL  
ASP  
LYS  
LEU  
ALA  
ALA  
ALA  
LEU  
GLU  
HIS  
HIS  
HIS  
HIS  
HIS

● Molecule 1: Cytosine deaminase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	101.95Å 162.42Å 184.05Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.66 – 1.95 46.66 – 1.95	Depositor EDS
% Data completeness (in resolution range)	95.4 (46.66-1.95) 95.2 (46.66-1.95)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.73 (at 1.95Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8_1069)	Depositor
R, $R_{free}$	0.165 , 0.201 0.166 , 0.203	Depositor DCC
$R_{free}$ test set	6377 reflections (3.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.4	Xtriage
Anisotropy	0.441	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 55.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	20876	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.58% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, 1LD, EDO, FE2, ACY, PEG, FLC

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	0.38	0/3312	0.54	0/4484
1	B	0.37	0/3317	0.53	0/4492
1	C	0.35	0/3275	0.51	0/4436
1	D	0.36	0/3291	0.53	0/4457
1	E	0.38	0/3275	0.52	0/4436
1	F	0.35	0/3294	0.51	0/4461
All	All	0.36	0/19764	0.52	0/26766

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	3254	0	3234	15	0
1	B	3259	0	3230	18	0
1	C	3217	0	3193	11	0
1	D	3233	0	3210	14	0
1	E	3217	0	3193	17	0
1	F	3236	0	3208	13	0
2	A	9	0	4	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	9	0	4	0	0
2	C	9	0	4	0	0
2	D	9	0	4	0	0
2	E	9	0	4	0	0
2	F	9	0	4	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
4	A	24	0	32	5	0
4	B	12	0	16	3	0
4	C	18	0	24	1	0
4	D	18	0	24	0	0
4	E	30	0	40	4	0
4	F	12	0	16	0	0
5	A	8	0	6	0	0
5	B	4	0	3	0	0
5	C	4	0	3	1	0
5	D	4	0	3	0	0
5	E	12	0	9	1	0
5	F	8	0	6	1	0
6	A	16	0	24	0	0
6	B	20	0	30	0	0
6	C	8	0	12	2	0
6	D	12	0	18	1	0
6	E	4	0	6	0	0
7	B	7	0	10	0	0
7	F	7	0	10	1	0
8	E	13	0	5	0	0
9	A	211	0	0	0	0
9	B	197	0	0	2	0
9	C	184	0	0	3	0
9	D	193	0	0	2	0
9	E	208	0	0	1	0
9	F	166	0	0	2	0
All	All	20876	0	19589	87	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.

The worst 5 of 87 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:B:407:HIS:HB3	1:B:409:ARG:HH12	1.56	0.70
1:A:376:ASP:O	1:A:386:ARG:NH1	2.25	0.69
1:A:409:ARG:HH22	1:B:404:GLU:HG3	1.58	0.67
1:C:201:ARG:NH2	9:C:685:HOH:O	2.29	0.66
1:D:1:MET:HE3	1:D:19:LEU:HD12	1.79	0.65

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	413/431 (96%)	400 (97%)	12 (3%)	1 (0%)	47	38
1	B	414/431 (96%)	402 (97%)	11 (3%)	1 (0%)	47	38
1	C	409/431 (95%)	397 (97%)	11 (3%)	1 (0%)	47	38
1	D	411/431 (95%)	396 (96%)	14 (3%)	1 (0%)	47	38
1	E	409/431 (95%)	396 (97%)	12 (3%)	1 (0%)	47	38
1	F	411/431 (95%)	399 (97%)	12 (3%)	0	100	100
All	All	2467/2586 (95%)	2390 (97%)	72 (3%)	5 (0%)	47	38

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	241	HIS
1	C	241	HIS
1	B	241	HIS
1	D	241	HIS
1	E	241	HIS

### 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	343/355 (97%)	342 (100%)	1 (0%)	92	92
1	B	343/355 (97%)	341 (99%)	2 (1%)	86	85
1	C	339/355 (96%)	337 (99%)	2 (1%)	86	85
1	D	340/355 (96%)	339 (100%)	1 (0%)	92	92
1	E	339/355 (96%)	335 (99%)	4 (1%)	71	68
1	F	341/355 (96%)	340 (100%)	1 (0%)	92	92
All	All	2045/2130 (96%)	2034 (100%)	11 (0%)	88	88

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	343	ARG
1	D	94	ASP
1	E	343	ARG
1	C	94	ASP
1	E	94	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 59 ligands modelled in this entry, 6 are monoatomic - leaving 53 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
4	GOL	C	503	-	5,5,5	0.45	0	5,5,5	0.58	0
4	GOL	A	503	-	5,5,5	0.46	0	5,5,5	0.69	0
2	ILD	A	501	-	8,9,9	1.73	1 (12%)	9,12,12	3.34	4 (44%)
4	GOL	D	504	-	5,5,5	0.34	0	5,5,5	0.35	0
4	GOL	F	504	-	5,5,5	0.38	0	5,5,5	0.55	0
4	GOL	E	506	-	5,5,5	0.36	0	5,5,5	0.44	0
5	ACY	E	507	-	1,3,3	1.86	0	0,3,3	0.00	-
6	EDO	D	509	-	3,3,3	0.44	0	2,2,2	0.34	0
4	GOL	B	506	-	5,5,5	0.39	0	5,5,5	0.38	0
4	GOL	A	504	-	5,5,5	0.31	0	5,5,5	0.56	0
4	GOL	C	504	-	5,5,5	0.38	0	5,5,5	0.23	0
6	EDO	B	511	-	3,3,3	0.40	0	2,2,2	0.53	0
4	GOL	B	503	-	5,5,5	0.40	0	5,5,5	0.52	0
4	GOL	E	503	-	5,5,5	0.35	0	5,5,5	0.49	0
5	ACY	A	505	-	1,3,3	1.66	0	0,3,3	0.00	-
4	GOL	D	503	-	5,5,5	0.36	0	5,5,5	0.33	0
4	GOL	E	504[B]	-	5,5,5	0.34	0	5,5,5	0.47	0
6	EDO	D	508	-	3,3,3	0.50	0	2,2,2	0.20	0
6	EDO	E	511	-	3,3,3	0.42	0	2,2,2	0.46	0
4	GOL	E	504[A]	-	5,5,5	0.37	0	5,5,5	0.30	0
7	PEG	B	505	-	6,6,6	0.50	0	5,5,5	0.26	0
4	GOL	E	505	-	5,5,5	0.35	0	5,5,5	0.45	0
6	EDO	D	507	-	3,3,3	0.48	0	2,2,2	0.36	0
4	GOL	D	505	-	5,5,5	0.33	0	5,5,5	0.37	0
2	ILD	B	501	-	8,9,9	1.72	1 (12%)	9,12,12	3.26	4 (44%)
5	ACY	E	509	-	1,3,3	1.51	0	0,3,3	0.00	-
5	ACY	F	506	-	1,3,3	1.36	0	0,3,3	0.00	-
2	ILD	D	501	-	8,9,9	1.79	2 (25%)	9,12,12	3.31	5 (55%)
4	GOL	A	507	-	5,5,5	0.35	0	5,5,5	0.39	0
2	ILD	E	501	-	8,9,9	1.79	1 (12%)	9,12,12	3.23	5 (55%)
2	ILD	C	501	-	8,9,9	1.84	1 (12%)	9,12,12	3.37	5 (55%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	ACY	B	504	-	1,3,3	1.73	0	0,3,3	0.00	-
6	EDO	B	508	-	3,3,3	0.51	0	2,2,2	0.17	0
6	EDO	A	510	-	3,3,3	0.45	0	2,2,2	0.29	0
6	EDO	B	509	-	3,3,3	0.49	0	2,2,2	0.30	0
4	GOL	A	508	-	5,5,5	0.33	0	5,5,5	0.27	0
5	ACY	E	508	-	1,3,3	1.57	0	0,3,3	0.00	-
8	FLC	E	510	-	3,12,12	3.41	3 (100%)	3,17,17	2.51	1 (33%)
6	EDO	A	509	-	3,3,3	0.49	0	2,2,2	0.23	0
5	ACY	F	505	-	1,3,3	1.82	0	0,3,3	0.00	-
7	PEG	F	507	-	6,6,6	0.45	0	5,5,5	0.89	0
5	ACY	C	506	-	1,3,3	1.49	0	0,3,3	0.00	-
6	EDO	C	507	-	3,3,3	0.47	0	2,2,2	0.36	0
5	ACY	A	506	-	1,3,3	1.28	0	0,3,3	0.00	-
6	EDO	B	507	-	3,3,3	0.48	0	2,2,2	0.49	0
4	GOL	C	505	-	5,5,5	0.38	0	5,5,5	0.51	0
4	GOL	F	501	-	5,5,5	0.34	0	5,5,5	0.21	0
6	EDO	C	508	-	3,3,3	0.45	0	2,2,2	0.31	0
2	1LD	F	502	-	8,9,9	1.85	2 (25%)	9,12,12	3.41	4 (44%)
6	EDO	A	512	-	3,3,3	0.40	0	2,2,2	0.58	0
5	ACY	D	506	-	1,3,3	2.10	1 (100%)	0,3,3	0.00	-
6	EDO	A	511	-	3,3,3	0.42	0	2,2,2	0.48	0
6	EDO	B	510	-	3,3,3	0.43	0	2,2,2	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
4	GOL	C	503	-	-	2/4/4/4	-
4	GOL	A	503	-	-	2/4/4/4	-
4	GOL	B	503	-	-	0/4/4/4	-
2	1LD	A	501	-	-	-	0/1/1/1
4	GOL	D	504	-	-	2/4/4/4	-
4	GOL	F	504	-	-	2/4/4/4	-
4	GOL	E	506	-	-	4/4/4/4	-
6	EDO	D	509	-	-	0/1/1/1	-
4	GOL	B	506	-	-	1/4/4/4	-
4	GOL	A	504	-	-	0/4/4/4	-
6	EDO	B	511	-	-	1/1/1/1	-
4	GOL	A	508	-	-	1/4/4/4	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	E	503	-	-	0/4/4/4	-
4	GOL	D	503	-	-	1/4/4/4	-
4	GOL	E	504[B]	-	-	2/4/4/4	-
6	EDO	D	508	-	-	0/1/1/1	-
6	EDO	E	511	-	-	0/1/1/1	-
4	GOL	E	504[A]	-	-	0/4/4/4	-
7	PEG	B	505	-	-	2/4/4/4	-
4	GOL	E	505	-	-	0/4/4/4	-
6	EDO	D	507	-	-	0/1/1/1	-
4	GOL	D	505	-	-	0/4/4/4	-
2	ILD	B	501	-	-	-	0/1/1/1
6	EDO	B	507	-	-	0/1/1/1	-
2	ILD	D	501	-	-	-	0/1/1/1
4	GOL	A	507	-	-	0/4/4/4	-
2	ILD	E	501	-	-	-	0/1/1/1
2	ILD	C	501	-	-	-	0/1/1/1
6	EDO	B	508	-	-	0/1/1/1	-
6	EDO	A	510	-	-	0/1/1/1	-
6	EDO	B	509	-	-	0/1/1/1	-
6	EDO	A	509	-	-	0/1/1/1	-
8	FLC	E	510	-	-	3/6/16/16	-
7	PEG	F	507	-	-	2/4/4/4	-
4	GOL	C	504	-	-	0/4/4/4	-
6	EDO	C	507	-	-	0/1/1/1	-
4	GOL	C	505	-	-	4/4/4/4	-
4	GOL	F	501	-	-	2/4/4/4	-
6	EDO	C	508	-	-	1/1/1/1	-
2	ILD	F	502	-	-	-	0/1/1/1
6	EDO	A	512	-	-	0/1/1/1	-
6	EDO	A	511	-	-	0/1/1/1	-
6	EDO	B	510	-	-	0/1/1/1	-

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	501	ILD	C4-N4	4.06	1.44	1.34
8	E	510	FLC	CA-CB	-3.99	1.49	1.54
2	E	501	ILD	C4-N4	3.94	1.44	1.34
2	F	502	ILD	C4-N4	3.84	1.43	1.34
2	B	501	ILD	C4-N4	3.80	1.43	1.34

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	501	1LD	N1-C2-N3	-7.93	122.12	128.43
2	C	501	1LD	N1-C2-N3	-7.91	122.14	128.43
2	F	502	1LD	N1-C2-N3	-7.86	122.19	128.43
2	D	501	1LD	N1-C2-N3	-7.76	122.26	128.43
2	B	501	1LD	N1-C2-N3	-7.66	122.34	128.43

There are no chirality outliers.

5 of 32 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	503	GOL	O1-C1-C2-C3
4	E	506	GOL	O1-C1-C2-O2
4	E	506	GOL	O1-C1-C2-C3
4	E	506	GOL	C1-C2-C3-O3
8	E	510	FLC	CA-CB-CG-CGC

There are no ring outliers.

13 monomers are involved in 20 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	D	509	EDO	1	0
4	B	506	GOL	3	0
4	A	504	GOL	3	0
4	C	504	GOL	1	0
4	E	504[B]	GOL	3	0
4	E	504[A]	GOL	1	0
5	F	506	ACY	1	0
4	A	507	GOL	1	0
4	A	508	GOL	1	0
5	E	508	ACY	1	0
7	F	507	PEG	1	0
5	C	506	ACY	1	0
6	C	508	EDO	2	0

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.



## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	411/431 (95%)	-0.12	6 (1%) 73 81	14, 20, 40, 72	0
1	B	413/431 (95%)	-0.07	14 (3%) 45 55	13, 22, 50, 77	0
1	C	411/431 (95%)	0.02	25 (6%) 21 29	15, 24, 49, 79	0
1	D	412/431 (95%)	-0.09	12 (2%) 51 60	14, 23, 42, 75	0
1	E	411/431 (95%)	-0.13	3 (0%) 87 92	15, 22, 39, 61	0
1	F	411/431 (95%)	0.29	42 (10%) 6 11	16, 26, 57, 89	0
All	All	2469/2586 (95%)	-0.02	102 (4%) 37 46	13, 23, 46, 89	0

The worst 5 of 102 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	38	ALA	7.3
1	B	38	ALA	6.4
1	F	38	ALA	5.6
1	C	39	GLY	5.2
1	F	26	ARG	5.1

### 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 carbohydrates 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
5	ACY	E	509	4/4	0.76	0.14	53,56,58,60	0
8	FLC	E	510	13/13	0.78	0.17	41,58,67,67	0
4	GOL	E	504[A]	6/6	0.80	0.29	27,32,36,38	6
4	GOL	E	504[B]	6/6	0.80	0.29	29,34,37,38	6
6	EDO	B	507	4/4	0.80	0.18	49,52,52,56	0
4	GOL	C	504	6/6	0.81	0.22	35,46,49,50	0
7	PEG	F	507	7/7	0.82	0.29	44,45,46,48	0
5	ACY	F	506	4/4	0.83	0.21	43,53,56,57	0
6	EDO	B	509	4/4	0.84	0.15	47,47,48,48	0
5	ACY	B	504	4/4	0.84	0.14	33,40,42,43	0
6	EDO	C	508	4/4	0.84	0.18	49,55,57,60	0
6	EDO	D	509	4/4	0.85	0.18	51,55,56,57	0
4	GOL	D	505	6/6	0.85	0.21	39,44,47,49	0
6	EDO	A	510	4/4	0.85	0.14	47,47,48,54	0
4	GOL	E	506	6/6	0.85	0.15	42,50,53,54	0
7	PEG	B	505	7/7	0.85	0.23	36,53,59,61	0
5	ACY	E	507	4/4	0.85	0.17	23,31,37,40	0
5	ACY	E	508	4/4	0.86	0.20	54,55,59,64	0
6	EDO	D	507	4/4	0.86	0.10	53,53,58,58	0
4	GOL	B	506	6/6	0.86	0.27	52,56,57,65	0
6	EDO	B	510	4/4	0.86	0.21	57,58,59,60	0
4	GOL	A	504	6/6	0.87	0.24	29,40,46,49	0
5	ACY	A	505	4/4	0.87	0.13	50,53,56,57	0
6	EDO	D	508	4/4	0.88	0.13	44,46,49,50	0
4	GOL	C	505	6/6	0.88	0.16	41,51,53,54	0
3	FE2	B	502	1/1	0.88	0.16	34,34,34,34	1
5	ACY	C	506	4/4	0.88	0.13	28,34,35,42	0
6	EDO	B	508	4/4	0.89	0.12	44,46,49,53	0
4	GOL	D	503	6/6	0.89	0.14	49,52,57,60	0
5	ACY	F	505	4/4	0.90	0.12	29,35,38,44	0
5	ACY	D	506	4/4	0.90	0.12	26,32,33,41	0
6	EDO	A	509	4/4	0.90	0.12	40,44,45,45	0
3	FE2	A	502	1/1	0.91	0.16	27,27,27,27	1
4	GOL	A	507	6/6	0.91	0.15	28,43,44,50	0
6	EDO	A	511	4/4	0.91	0.11	53,55,55,59	0
4	GOL	F	501	6/6	0.92	0.11	36,47,48,51	0
5	ACY	A	506	4/4	0.92	0.10	46,47,51,56	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	ILD	F	502	9/9	0.94	0.11	22,28,33,35	0
6	EDO	A	512	4/4	0.94	0.12	48,49,51,51	0
3	FE2	F	503	1/1	0.94	0.12	31,31,31,31	1
6	EDO	C	507	4/4	0.94	0.12	47,48,50,51	0
2	ILD	C	501	9/9	0.95	0.08	16,23,28,29	0
6	EDO	B	511	4/4	0.95	0.11	49,51,55,56	0
4	GOL	E	505	6/6	0.95	0.27	33,37,40,40	0
6	EDO	E	511	4/4	0.95	0.13	32,38,42,49	0
3	FE2	E	502	1/1	0.95	0.11	27,27,27,27	1
3	FE2	C	502	1/1	0.95	0.09	31,31,31,31	1
2	ILD	E	501	9/9	0.96	0.07	20,22,31,31	0
2	ILD	B	501	9/9	0.96	0.09	16,23,29,30	0
4	GOL	A	508	6/6	0.96	0.13	26,43,45,48	0
4	GOL	A	503	6/6	0.96	0.12	25,25,29,29	0
4	GOL	C	503	6/6	0.96	0.12	22,27,29,32	0
4	GOL	E	503	6/6	0.97	0.07	20,24,28,30	0
2	ILD	A	501	9/9	0.97	0.09	19,22,26,30	0
3	FE2	D	502	1/1	0.97	0.10	35,35,35,35	1
2	ILD	D	501	9/9	0.97	0.07	18,20,28,29	0
4	GOL	F	504	6/6	0.97	0.09	22,25,26,26	0
4	GOL	B	503	6/6	0.98	0.06	19,22,27,29	0
4	GOL	D	504	6/6	0.98	0.10	20,23,26,28	0

## 6.5 Other polymers

There are no such residues in this entry.