



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 05:55 am BST

PDB ID : 6D3I  
Title : ftv7 dioxygenase with 2,4-D bound  
Authors : Rydel, T.J.; Halls, C.E.  
Deposited on : 2018-04-16  
Resolution : 3.20 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

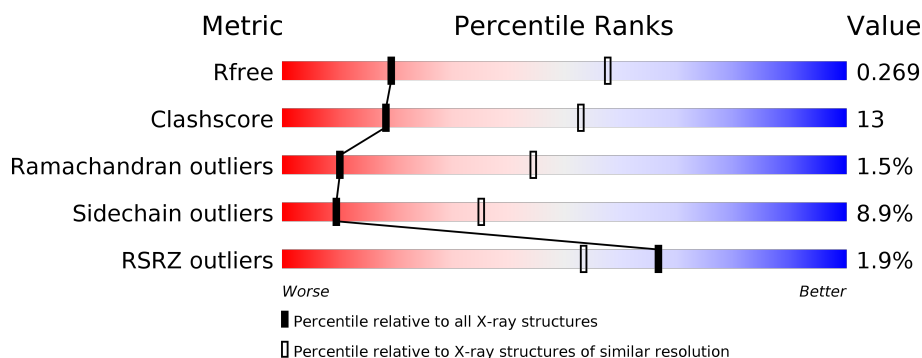
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.20 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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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	295	<div> <div>4%</div> <div> <div></div> <div>55%</div> <div>34%</div> <div>7%</div> <div></div> </div> </div>
1	D	295	<div> <div>2%</div> <div> <div></div> <div>71%</div> <div>23%</div> <div></div> </div> </div>
1	G	295	<div> <div></div> <div> <div>64%</div> <div>29%</div> <div></div> </div> </div>
1	J	295	<div> <div>%</div> <div> <div>70%</div> <div>22%</div> <div></div> </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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	CFA	G	303	-	-	X	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 9247 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 ftv7 dioxygenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	286	Total	C	N	O	S	0	1	0
			2291	1455	394	431	11			
1	G	285	Total	C	N	O	S	0	0	0
			2280	1449	392	428	11			
1	D	285	Total	C	N	O	S	0	0	0
			2280	1449	392	428	11			
1	J	285	Total	C	N	O	S	0	0	0
			2280	1449	392	428	11			

- Molecule 2 is COBALT (II) ION (three-letter code: CO) (formula: Co).

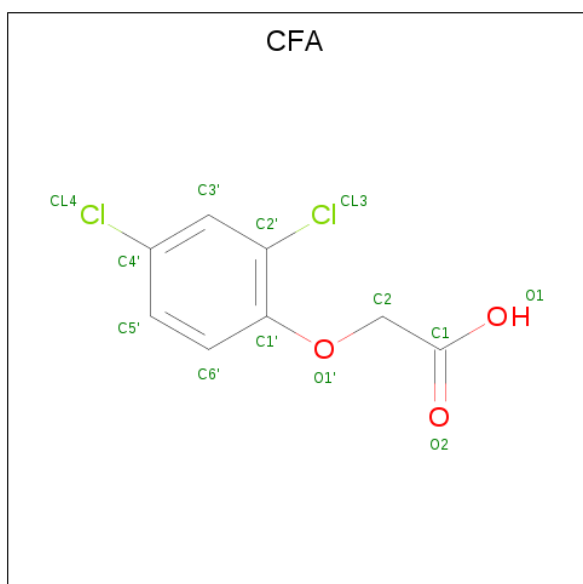
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	1	Total	Co	0	0
			1	1		
2	J	1	Total	Co	0	0
			1	1		
2	A	1	Total	Co	0	0
			1	1		
2	D	1	Total	Co	0	0
			1	1		

- Molecule 3 is 2-OXOGLUTARIC ACID (three-letter code: AKG) (formula: C<sub>5</sub>H<sub>6</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			10	5	5		
3	G	1	Total	C	O	0	0
			10	5	5		
3	D	1	Total	C	O	0	0
			10	5	5		
3	J	1	Total	C	O	0	0
			10	5	5		

- Molecule 4 is (2,4-DICHLOROPHENOXY)ACETIC ACID (three-letter code: CFA) (formula:  $C_8H_6Cl_2O_3$ ).



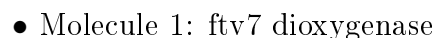
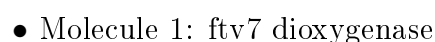
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	G	1	Total 13	C 8	Cl 2	O 3	0	0
4	D	1	Total 13	C 8	Cl 2	O 3	0	0

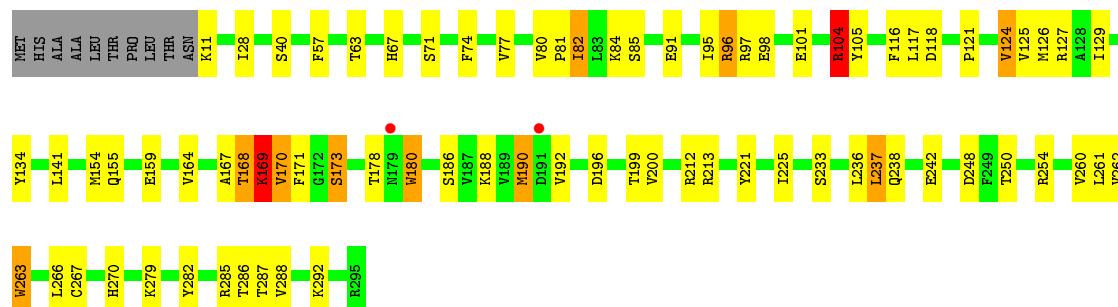
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	15	Total 15	O 15	0	0
5	G	10	Total 10	O 10	0	0
5	D	9	Total 9	O 9	0	0
5	J	12	Total 12	O 12	0	0



- Molecule 1: ftv7 dioxygenase







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	94.93Å 109.98Å 116.14Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.70 – 3.20 49.70 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.2 (49.70-3.20) 94.2 (49.70-3.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.11	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.21 (at 3.19Å)	Xtriage
Refinement program	PHENIX (1.11.1 _2575: ???)	Depositor
R, $R_{free}$	0.176 , 0.273 0.182 , 0.269	Depositor DCC
$R_{free}$ test set	1985 reflections (9.65%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	53.2	Xtriage
Anisotropy	0.512	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 34.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	9247	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 24.38 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 3.8760e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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

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

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CO, AKG, CFA

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	0.68	2/2354 (0.1%)	0.85	3/3206 (0.1%)
1	D	0.67	0/2340	0.84	2/3187 (0.1%)
1	G	0.75	2/2340 (0.1%)	0.91	5/3187 (0.2%)
1	J	0.75	1/2340 (0.0%)	0.85	0/3187
All	All	0.71	5/9374 (0.1%)	0.86	10/12767 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	G	0	1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	J	267	CYS	CB-SG	-7.65	1.69	1.82
1	G	222	CYS	CB-SG	-7.24	1.70	1.82
1	A	222	CYS	CB-SG	-6.79	1.70	1.82
1	G	110	TRP	CB-CG	-6.04	1.39	1.50
1	A	42	TRP	CB-CG	6.00	1.61	1.50

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	181	ARG	NE-CZ-NH1	11.59	126.10	120.30
1	A	20	LEU	CB-CG-CD2	7.08	123.03	111.00
1	G	181	ARG	NE-CZ-NH2	-6.35	117.13	120.30
1	D	212	ARG	NE-CZ-NH1	6.05	123.33	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	124	VAL	CG1-CB-CG2	6.04	120.56	110.90

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	G	178	THR	Peptide

## 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	2291	0	2217	79	0
1	D	2280	0	2206	50	0
1	G	2280	0	2206	67	0
1	J	2280	0	2206	47	0
2	A	1	0	0	0	0
2	D	1	0	0	0	0
2	G	1	0	0	0	0
2	J	1	0	0	0	0
3	A	10	0	4	2	0
3	D	10	0	4	2	0
3	G	10	0	4	3	0
3	J	10	0	4	0	0
4	D	13	0	5	1	0
4	G	13	0	5	6	0
5	A	15	0	0	5	0
5	D	9	0	0	0	0
5	G	10	0	0	0	0
5	J	12	0	0	0	0
All	All	9247	0	8861	240	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 13.

The worst 5 of 240 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:30:GLY:O	5:A:401:HOH:O	1.82	0.95
1:A:286:THR:O	1:A:287:THR:OG1	1.89	0.91
1:A:148:GLU:OE2	1:A:212:ARG:NH1	2.09	0.85
1:A:118:ASP:OD1	1:A:197:ARG:NH2	2.12	0.83
1:D:147:TRP:CZ3	1:D:204:VAL:HG21	2.21	0.75

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	285/295 (97%)	250 (88%)	28 (10%)	7 (2%)	5	32
1	D	283/295 (96%)	260 (92%)	18 (6%)	5 (2%)	8	41
1	G	283/295 (96%)	265 (94%)	16 (6%)	2 (1%)	22	61
1	J	283/295 (96%)	263 (93%)	17 (6%)	3 (1%)	14	51
All	All	1134/1180 (96%)	1038 (92%)	79 (7%)	17 (2%)	10	44

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	287	THR
1	G	82	ILE
1	D	180	TRP
1	J	82	ILE
1	A	11	LYS

### 5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	249/255 (98%)	219 (88%)	30 (12%)	5	22
1	D	247/255 (97%)	232 (94%)	15 (6%)	18	54
1	G	247/255 (97%)	229 (93%)	18 (7%)	14	46
1	J	247/255 (97%)	222 (90%)	25 (10%)	7	29
All	All	990/1020 (97%)	902 (91%)	88 (9%)	9	35

5 of 88 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	180	TRP
1	D	18	GLN
1	J	212	ARG
1	G	181	ARG
1	G	190	MET

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

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 10 ligands modelled in this entry, 4 are monoatomic - leaving 6 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	CFA	D	303	-	10,13,13	1.64	3 (30%)	14,17,17	2.52	4 (28%)
3	AKG	G	302	2	3,9,9	1.64	1 (33%)	4,11,11	3.21	1 (25%)
3	AKG	A	302	2	3,9,9	1.64	2 (66%)	4,11,11	1.76	1 (25%)
4	CFA	G	303	-	10,13,13	1.50	3 (30%)	14,17,17	2.19	5 (35%)
3	AKG	J	302	2	3,9,9	2.02	1 (33%)	4,11,11	1.40	1 (25%)
3	AKG	D	302	2	3,9,9	1.65	1 (33%)	4,11,11	1.20	1 (25%)

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	CFA	D	303	-	-	1/3/5/5	0/1/1/1
3	AKG	G	302	2	-	3/3/9/9	-
3	AKG	A	302	2	-	3/3/9/9	-
4	CFA	G	303	-	-	2/3/5/5	0/1/1/1
3	AKG	J	302	2	-	1/3/9/9	-
3	AKG	D	302	2	-	1/3/9/9	-

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	303	CFA	C2'-CL3	3.17	1.81	1.73
3	J	302	AKG	O5-C2	-2.88	1.17	1.22
4	G	303	CFA	C4'-CL4	2.72	1.80	1.74
3	G	302	AKG	O5-C2	-2.60	1.18	1.22
4	G	303	CFA	C2'-CL3	2.35	1.79	1.73

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	G	302	AKG	C4-C3-C2	-6.05	100.08	113.14
4	D	303	CFA	C3'-C4'-CL4	5.35	125.84	119.15
4	D	303	CFA	C3'-C2'-CL3	4.71	126.10	118.49

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	303	CFA	C1'-C2'-CL3	-4.23	114.46	119.43
4	D	303	CFA	C1'-C2'-CL3	-4.13	114.58	119.43

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	G	302	AKG	C2-C3-C4-C5
3	A	302	AKG	C2-C3-C4-C5
3	J	302	AKG	C2-C3-C4-C5
4	G	303	CFA	C2'-C1'-O1'-C2
3	G	302	AKG	O5-C2-C3-C4

There are no ring outliers.

5 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	303	CFA	1	0
3	G	302	AKG	3	0
3	A	302	AKG	2	0
4	G	303	CFA	6	0
3	D	302	AKG	2	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 [i](#)

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

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	286/295 (96%)	0.13	13 (4%) 33 21	37, 65, 100, 130	0
1	D	285/295 (96%)	-0.32	7 (2%) 57 43	30, 46, 88, 110	0
1	G	285/295 (96%)	-0.53	0 100 100	27, 44, 67, 99	0
1	J	285/295 (96%)	-0.41	2 (0%) 87 81	24, 39, 80, 120	0
All	All	1141/1180 (96%)	-0.28	22 (1%) 66 53	24, 48, 91, 130	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	174	LEU	5.4
1	A	191	ASP	3.4
1	D	183	SER	2.8
1	A	86	ILE	2.8
1	D	185	THR	2.7

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

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

### 6.3 Carbohydrates [i](#)

There are no 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. 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( $\text{\AA}^2$ )	Q<0.9
4	CFA	D	303	13/13	0.86	0.30	64,71,89,92	0
4	CFA	G	303	13/13	0.91	0.28	41,57,64,98	0
3	AKG	A	302	10/10	0.94	0.20	32,55,72,73	0
2	CO	G	301	1/1	0.94	0.07	52,52,52,52	0
2	CO	D	301	1/1	0.96	0.09	40,40,40,40	0
3	AKG	J	302	10/10	0.97	0.20	31,37,46,47	0
3	AKG	D	302	10/10	0.97	0.15	38,48,56,58	0
2	CO	J	301	1/1	0.98	0.05	40,40,40,40	0
3	AKG	G	302	10/10	0.98	0.14	15,33,43,43	0
2	CO	A	301	1/1	1.00	0.05	45,45,45,45	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.