



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 13, 2024 – 07:12 pm BST

PDB ID : 5JP3  
Title : Structure of Xanthomonas campestris effector protein XopD bound to ubiquitin  
Authors : Pruneda, J.N.; Komander, D.  
Deposited on : 2016-05-03  
Resolution : 2.90 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

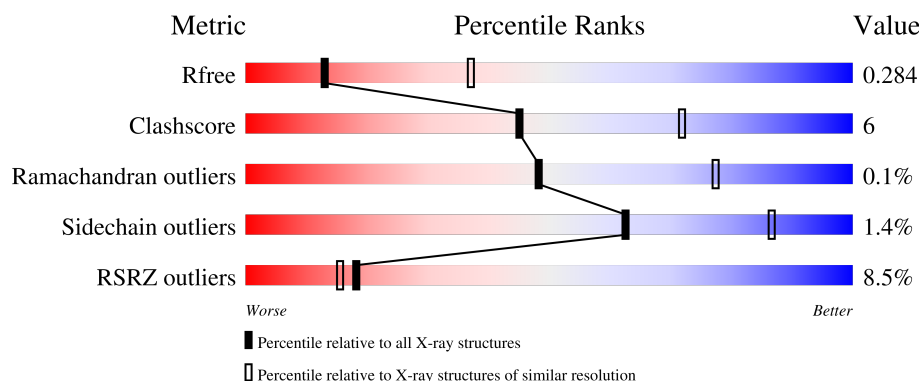
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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))
$R_{free}$	164625	2335 (2.90-2.90)
Clashscore	180529	2564 (2.90-2.90)
Ramachandran outliers	177936	2514 (2.90-2.90)
Sidechain outliers	177891	2516 (2.90-2.90)
RSRZ outliers	164620	2337 (2.90-2.90)

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 of 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	220	<div> <div>3%</div> <div>82%</div> <div>10%</div> <div>7%</div> </div>
1	C	220	<div> <div>2%</div> <div>80%</div> <div>9%</div> <div>11%</div> </div>
1	E	220	<div> <div>13%</div> <div>76%</div> <div>10%</div> <div>12%</div> </div>
1	G	220	<div> <div>5%</div> <div>70%</div> <div>18%</div> <div>11%</div> </div>
2	B	76	<div> <div>97%</div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	D	76	<div><div>%</div><div><div></div><div>83%</div><div>17%</div></div></div>
2	F	76	<div><div>21%</div><div><div></div><div>80%</div><div>20%</div></div></div>
2	H	76	<div><div>33%</div><div><div></div><div>76%</div><div>24%</div></div></div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 8236 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 Xanthomonas outer protein D.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	205	Total	C	N	O	S	0	0	0
			1560	977	280	298	5			
1	C	196	Total	C	N	O	S	0	0	0
			1478	928	268	277	5			
1	E	193	Total	C	N	O	S	0	0	0
			1426	894	253	274	5			
1	G	196	Total	C	N	O	S	0	0	0
			1510	945	274	286	5			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	296	GLY	-	expression tag	UNP Q3BYJ5
A	297	PRO	-	expression tag	UNP Q3BYJ5
C	296	GLY	-	expression tag	UNP Q3BYJ5
C	297	PRO	-	expression tag	UNP Q3BYJ5
E	296	GLY	-	expression tag	UNP Q3BYJ5
E	297	PRO	-	expression tag	UNP Q3BYJ5
G	296	GLY	-	expression tag	UNP Q3BYJ5
G	297	PRO	-	expression tag	UNP Q3BYJ5

- Molecule 2 is a protein called Polyubiquitin-B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	76	Total	C	N	O	S	0	0	0
			586	371	104	110	1			
2	D	76	Total	C	N	O	S	0	0	0
			570	361	96	112	1			
2	H	76	Total	C	N	O	S	0	0	0
			514	325	91	97	1			
2	F	76	Total	C	N	O	S	0	0	0
			548	349	93	105	1			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	76	AYE	-	expression tag	UNP P0CG47
D	76	AYE	-	expression tag	UNP P0CG47
H	76	AYE	-	expression tag	UNP P0CG47
F	76	AYE	-	expression tag	UNP P0CG47

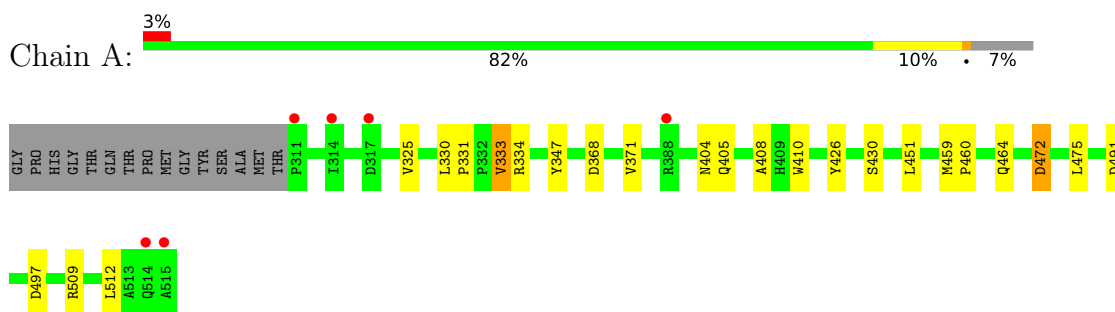
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	25	Total O 25 25	0	0
3	C	6	Total O 6 6	0	0
3	E	3	Total O 3 3	0	0
3	B	4	Total O 4 4	0	0
3	D	1	Total O 1 1	0	0
3	H	1	Total O 1 1	0	0
3	G	4	Total O 4 4	0	0

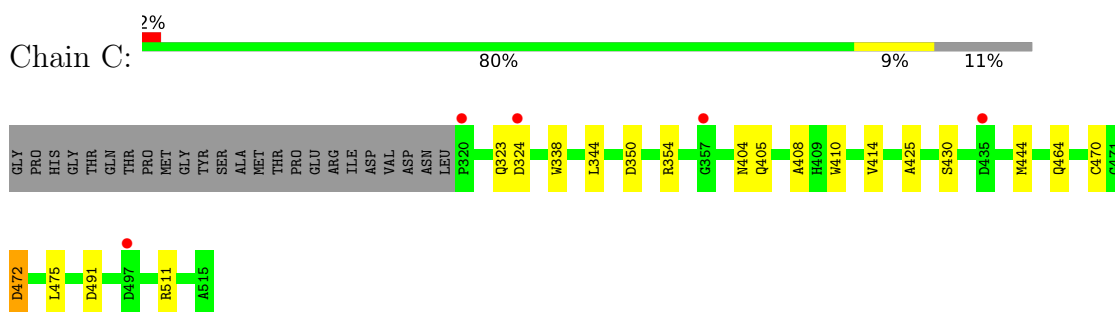
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

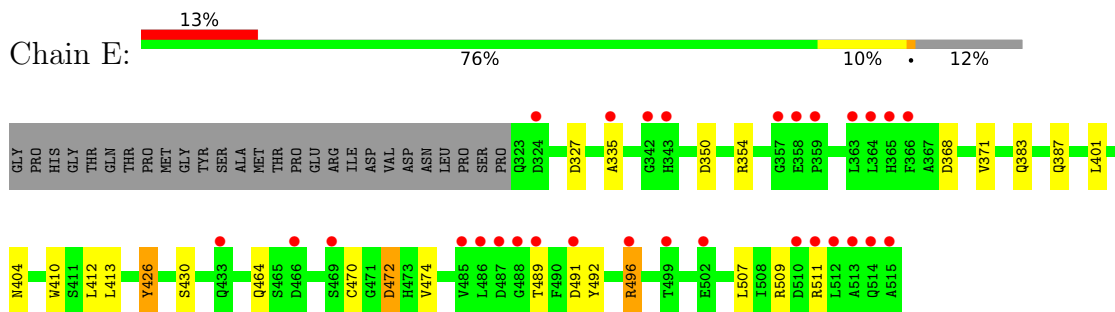
- Molecule 1: Xanthomonas outer protein D



- Molecule 1: Xanthomonas outer protein D

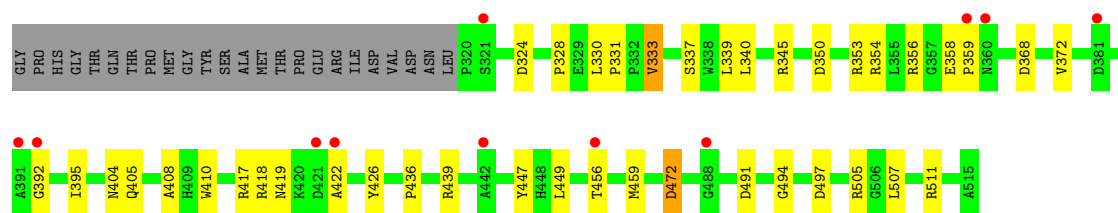


- Molecule 1: Xanthomonas outer protein D

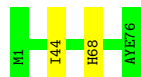


- Molecule 1: Xanthomonas outer protein D

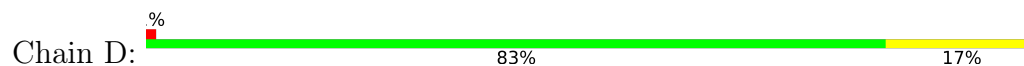




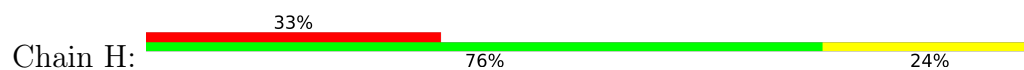
● Molecule 2: Polyubiquitin-B



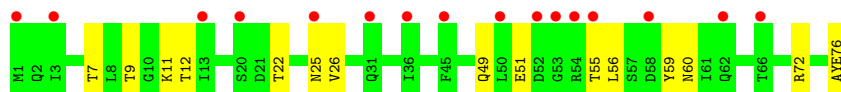
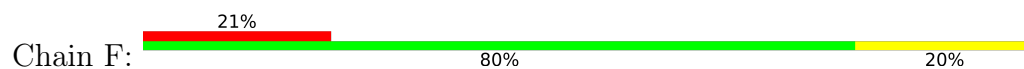
● Molecule 2: Polyubiquitin-B



● Molecule 2: Polyubiquitin-B



● Molecule 2: Polyubiquitin-B



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	117.75Å 132.28Å 117.31Å 90.00° 105.84° 90.00°	Depositor
Resolution (Å)	43.30 – 2.90 43.30 – 2.90	Depositor EDS
% Data completeness (in resolution range)	98.8 (43.30-2.90) 98.8 (43.30-2.90)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.40 (at 2.90Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, $R_{free}$	0.241 , 0.285 0.240 , 0.284	Depositor DCC
$R_{free}$ test set	1836 reflections (4.84%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	40.8	Xtriage
Anisotropy	0.502	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 65.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.87	EDS
Total number of atoms	8236	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	46.0	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 30.76 % 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 1.2397e-03.*

<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: AYE

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	0.32	0/1598	0.56	0/2186
1	C	0.31	0/1515	0.51	0/2074
1	E	0.41	1/1460 (0.1%)	0.67	2/2001 (0.1%)
1	G	0.30	0/1547	0.51	0/2113
2	B	0.30	0/588	0.60	0/793
2	D	0.28	0/572	0.61	1/777 (0.1%)
2	F	0.27	0/550	0.57	1/748 (0.1%)
2	H	0.29	0/515	0.64	0/702
All	All	0.32	1/8345 (0.0%)	0.57	4/11394 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	426	TYR	CG-CD1	7.93	1.49	1.39

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	496	ARG	NE-CZ-NH1	-16.16	112.22	120.30
1	E	496	ARG	NH1-CZ-NH2	6.28	126.31	119.40
2	F	72	ARG	NE-CZ-NH2	-5.34	117.63	120.30
2	D	72	ARG	NE-CZ-NH2	-5.25	117.68	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	1560	0	1459	13	0
1	C	1478	0	1385	14	0
1	E	1426	0	1295	18	0
1	G	1510	0	1435	27	0
2	B	586	0	609	2	0
2	D	570	0	559	6	0
2	F	548	0	530	9	0
2	H	514	0	464	14	0
3	A	25	0	0	0	0
3	B	4	0	0	0	0
3	C	6	0	0	0	0
3	D	1	0	0	0	0
3	E	3	0	0	0	0
3	G	4	0	0	0	0
3	H	1	0	0	0	0
All	All	8236	0	7736	92	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 6.

The worst 5 of 92 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:G:392:GLY:H	1:G:395:ILE:HD12	1.48	0.76
2:H:27:LYS:NZ	2:H:52:ASP:OD2	2.18	0.75
1:C:344:LEU:HG	1:C:475:LEU:HD11	1.70	0.72
1:G:337:SER:O	1:G:505:ARG:NH2	2.23	0.68
1:G:472:ASP:OD1	1:G:505:ARG:HG2	1.93	0.67

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	203/220 (92%)	195 (96%)	8 (4%)	0	100	100
1	C	194/220 (88%)	190 (98%)	4 (2%)	0	100	100
1	E	191/220 (87%)	186 (97%)	5 (3%)	0	100	100
1	G	194/220 (88%)	188 (97%)	6 (3%)	0	100	100
2	B	73/76 (96%)	73 (100%)	0	0	100	100
2	D	73/76 (96%)	73 (100%)	0	0	100	100
2	F	73/76 (96%)	72 (99%)	1 (1%)	0	100	100
2	H	73/76 (96%)	70 (96%)	2 (3%)	1 (1%)	9	31
All	All	1074/1184 (91%)	1047 (98%)	26 (2%)	1 (0%)	48	77

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	H	17	VAL

### 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	156/182 (86%)	152 (97%)	4 (3%)	41	74
1	C	146/182 (80%)	145 (99%)	1 (1%)	81	94
1	E	135/182 (74%)	134 (99%)	1 (1%)	81	94
1	G	154/182 (85%)	151 (98%)	3 (2%)	52	81

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	64/68 (94%)	64 (100%)	0	100	100
2	D	59/68 (87%)	57 (97%)	2 (3%)	32	67
2	F	54/68 (79%)	54 (100%)	0	100	100
2	H	43/68 (63%)	43 (100%)	0	100	100
All	All	811/1000 (81%)	800 (99%)	11 (1%)	62	86

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

Mol	Chain	Res	Type
2	D	52	ASP
1	G	333	VAL
1	G	472	ASP
1	G	447	TYR
1	C	472	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	H	49	GLN

### 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 oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 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	205/220 (93%)	-0.11	6 (2%) 54 48	14, 29, 54, 91	0
1	C	196/220 (89%)	0.16	5 (2%) 57 51	19, 34, 56, 95	0
1	E	193/220 (87%)	0.91	29 (15%) 6 6	22, 56, 90, 117	0
1	G	196/220 (89%)	0.29	11 (5%) 31 27	21, 39, 70, 104	0
2	B	75/76 (98%)	-0.21	0 100 100	18, 31, 50, 59	0
2	D	75/76 (98%)	0.47	1 (1%) 74 69	31, 56, 76, 86	0
2	F	75/76 (98%)	1.36	16 (21%) 3 3	46, 85, 124, 139	0
2	H	75/76 (98%)	1.58	25 (33%) 1 1	27, 74, 103, 115	0
All	All	1090/1184 (92%)	0.44	93 (8%) 18 15	14, 41, 91, 139	0

The worst 5 of 93 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	13	ILE	6.0
2	H	45	PHE	5.4
2	H	64	GLU	5.3
1	E	487	ASP	5.1
1	E	496	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 monosaccharides in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.