



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 04:42 AM GMT

PDB ID : 2NWX  
Title : Crystal structure of GltPh in complex with L-aspartate and sodium ions  
Authors : Gouaux, E.; Boudker, O.; Ryan, R.; Yernool, D.; Shimamoto, K.  
Deposited on : 2006-11-16  
Resolution : 3.29 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  
<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

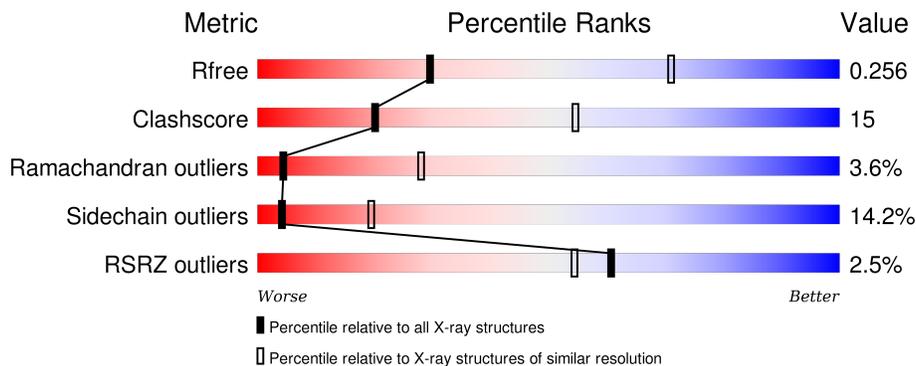
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.29 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	2060 (3.40-3.20)
Clashscore	102246	1058 (3.38-3.22)
Ramachandran outliers	100387	1038 (3.38-3.22)
Sidechain outliers	100360	1037 (3.38-3.22)
RSRZ outliers	91569	2070 (3.40-3.20)

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. 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	422	 58% 31% 6% 5%
1	B	422	 5% 59% 30% 5% 6%
1	C	422	 60% 29% 6% 5%

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

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Chirality</b>	<b>Geometry</b>	<b>Clashes</b>	<b>Electron density</b>
2	NA	A	424	-	-	-	X
4	PLM	A	428	-	-	-	X
4	PLM	B	425	-	-	-	X
4	PLM	C	425	-	-	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 8596 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 425aa long hypothetical proton glutamate symport protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	400	2846	1871	456	502	17	0	0	0
1	B	396	2816	1853	450	496	17	0	0	0
1	C	401	2847	1872	456	502	17	0	0	0

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	37	HIS	ASP	ENGINEERED	UNP O59010
A	40	HIS	LYS	ENGINEERED	UNP O59010
A	125	HIS	LYS	ENGINEERED	UNP O59010
A	132	HIS	LYS	ENGINEERED	UNP O59010
A	223	HIS	LYS	ENGINEERED	UNP O59010
A	264	HIS	LYS	ENGINEERED	UNP O59010
A	368	HIS	GLU	ENGINEERED	UNP O59010
A	418	THR	-	CLONING ARTIFACT	UNP O59010
A	419	LEU	-	CLONING ARTIFACT	UNP O59010
A	420	VAL	-	CLONING ARTIFACT	UNP O59010
A	421	PRO	-	CLONING ARTIFACT	UNP O59010
A	422	ARG	-	CLONING ARTIFACT	UNP O59010
B	37	HIS	ASP	ENGINEERED	UNP O59010
B	40	HIS	LYS	ENGINEERED	UNP O59010
B	125	HIS	LYS	ENGINEERED	UNP O59010
B	132	HIS	LYS	ENGINEERED	UNP O59010
B	223	HIS	LYS	ENGINEERED	UNP O59010
B	264	HIS	LYS	ENGINEERED	UNP O59010
B	368	HIS	GLU	ENGINEERED	UNP O59010
B	418	THR	-	CLONING ARTIFACT	UNP O59010
B	419	LEU	-	CLONING ARTIFACT	UNP O59010
B	420	VAL	-	CLONING ARTIFACT	UNP O59010
B	421	PRO	-	CLONING ARTIFACT	UNP O59010

*Continued on next page...*

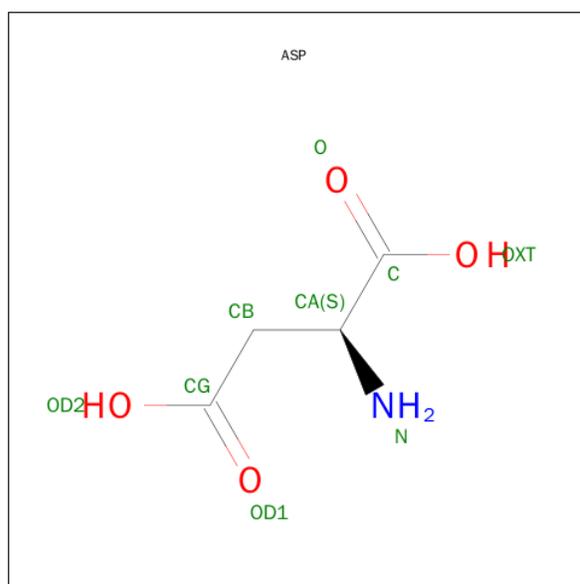
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	422	ARG	-	CLONING ARTIFACT	UNP O59010
C	37	HIS	ASP	ENGINEERED	UNP O59010
C	40	HIS	LYS	ENGINEERED	UNP O59010
C	125	HIS	LYS	ENGINEERED	UNP O59010
C	132	HIS	LYS	ENGINEERED	UNP O59010
C	223	HIS	LYS	ENGINEERED	UNP O59010
C	264	HIS	LYS	ENGINEERED	UNP O59010
C	368	HIS	GLU	ENGINEERED	UNP O59010
C	418	THR	-	CLONING ARTIFACT	UNP O59010
C	419	LEU	-	CLONING ARTIFACT	UNP O59010
C	420	VAL	-	CLONING ARTIFACT	UNP O59010
C	421	PRO	-	CLONING ARTIFACT	UNP O59010
C	422	ARG	-	CLONING ARTIFACT	UNP O59010

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

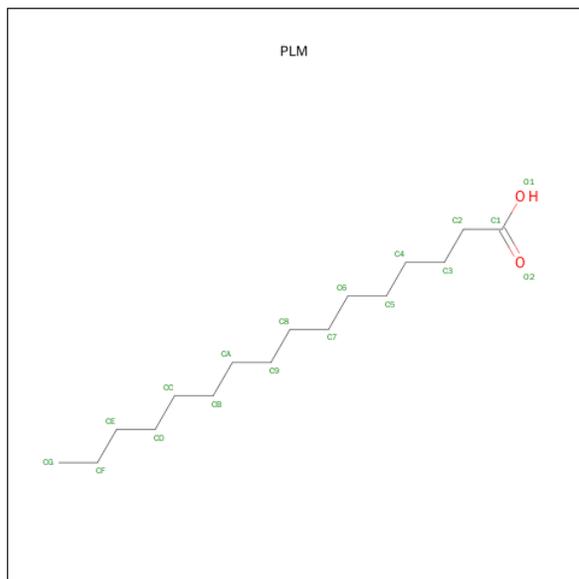
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	2	Total Na 2 2	0	0
2	A	2	Total Na 2 2	0	0
2	C	2	Total Na 2 2	0	0

- Molecule 3 is ASPARTIC ACID (three-letter code: ASP) (formula: C<sub>4</sub>H<sub>7</sub>NO<sub>4</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	3	Total	C	N	O	0	0
			27	12	3	12		

- Molecule 4 is PALMITIC ACID (three-letter code: PLM) (formula: C<sub>16</sub>H<sub>32</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			17	16	1		
4	B	1	Total	C	O	0	0
			17	16	1		
4	C	1	Total	C	O	0	0
			17	16	1		

- Molecule 5 is water.

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





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	115.47Å 115.47Å 324.23Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	100.00 – 3.29 39.60 – 3.29	Depositor EDS
% Data completeness (in resolution range)	69.2 (100.00-3.29) 69.3 (39.60-3.29)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	9.40	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.31 (at 3.32Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.263 , 0.286 0.250 , 0.256	Depositor DCC
$R_{free}$ test set	1155 reflections (5.13%)	DCC
Wilson B-factor (Å <sup>2</sup> )	123.4	Xtriage
Anisotropy	0.416	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.21 , 73.5	EDS
Estimated twinning fraction	0.087 for h,-h-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	0 of 25514 reflections	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	8596	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	134.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.23% 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: NA, PLM

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	0.75	0/2896	0.84	1/3960 (0.0%)
1	B	0.66	0/2866	0.77	0/3920
1	C	0.66	0/2897	0.76	0/3962
All	All	0.69	0/8659	0.79	1/11842 (0.0%)

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	A	0	2
1	B	0	1
1	C	0	2
All	All	0	5

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	363	LEU	CA-CB-CG	5.03	126.88	115.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	119	GLY	Peptide
1	A	70	ALA	Peptide
1	B	70	ALA	Peptide

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Group
1	C	376	ASP	Peptide
1	C	70	ALA	Peptide

## 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	2846	0	2902	93	0
1	B	2816	0	2872	84	0
1	C	2847	0	2898	93	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
3	A	27	0	9	8	0
4	A	17	0	31	0	0
4	B	17	0	31	0	0
4	C	17	0	31	0	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
5	C	1	0	0	0	0
All	All	8596	0	8774	255	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 15.

The worst 5 of 255 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:155:ILE:HD11	1:B:304:PRO:HB2	1.44	0.98
1:A:155:ILE:HD11	1:A:304:PRO:HB2	1.47	0.95
1:C:155:ILE:HD11	1:C:304:PRO:HB2	1.51	0.91
1:A:136:ASP:HA	1:C:52:ARG:NH2	1.91	0.85
1:A:81:VAL:HG21	1:A:298:ILE:HD12	1.60	0.83

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	396/422 (94%)	309 (78%)	71 (18%)	16 (4%)	4	24
1	B	392/422 (93%)	319 (81%)	60 (15%)	13 (3%)	5	30
1	C	397/422 (94%)	317 (80%)	66 (17%)	14 (4%)	4	29
All	All	1185/1266 (94%)	945 (80%)	197 (17%)	43 (4%)	4	28

5 of 43 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	72	SER
1	B	72	SER
1	B	115	LEU
1	C	72	SER
1	C	114	HIS

### 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	277/330 (84%)	238 (86%)	39 (14%)	4	20
1	B	274/330 (83%)	234 (85%)	40 (15%)	4	18
1	C	276/330 (84%)	238 (86%)	38 (14%)	4	20
All	All	827/990 (84%)	710 (86%)	117 (14%)	4	20

5 of 117 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	192	GLU
1	B	285	THR
1	C	338	GLN
1	B	198	VAL
1	B	245	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	40	HIS
1	C	327	ASN
1	B	327	ASN
1	A	327	ASN
1	C	40	HIS

### 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 12 ligands modelled in this entry, 6 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
3	ASP	A	425	-	2,8,8	0.30	0	0,10,10	0.00	-
3	ASP	A	426	-	2,8,8	0.26	0	0,10,10	0.00	-
3	ASP	A	427	-	2,8,8	0.29	0	0,10,10	0.00	-
4	PLM	A	428	-	16,16,17	0.60	0	14,15,17	0.54	0
4	PLM	B	425	-	16,16,17	0.56	0	14,15,17	0.56	0
4	PLM	C	425	-	16,16,17	0.69	0	14,15,17	0.46	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	ASP	A	425	-	-	0/2/8/8	0/0/0/0
3	ASP	A	426	-	-	0/2/8/8	0/0/0/0
3	ASP	A	427	-	-	0/2/8/8	0/0/0/0
4	PLM	A	428	-	-	0/13/14/15	0/0/0/0
4	PLM	B	425	-	-	0/13/14/15	0/0/0/0
4	PLM	C	425	-	-	0/13/14/15	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	425	ASP	3	0
3	A	426	ASP	2	0
3	A	427	ASP	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 [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	400/422 (94%)	-0.37	4 (1%) 84 80	98, 126, 170, 185	0
1	B	396/422 (93%)	-0.26	20 (5%) 32 25	100, 133, 176, 189	0
1	C	401/422 (95%)	-0.44	6 (1%) 76 71	106, 134, 178, 189	0
All	All	1197/1266 (94%)	-0.36	30 (2%) 61 54	98, 132, 176, 189	0

The worst 5 of 30 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	33	TYR	4.9
1	B	267	ASP	4.9
1	C	252	LYS	4.4
1	B	336	GLY	3.7
1	C	370	VAL	3.3

### 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. 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
4	PLM	A	428	17/18	0.76	0.59	14.88	151,154,160,160	0
4	PLM	C	425	17/18	0.49	0.65	7.47	149,152,157,157	0
2	NA	A	424	1/1	0.99	0.29	3.62	84,84,84,84	0
4	PLM	B	425	17/18	0.69	0.45	3.12	150,152,157,157	0
2	NA	B	424	1/1	0.98	0.21	0.76	118,118,118,118	0
3	ASP	A	426	9/9	0.89	0.27	0.69	124,125,126,126	0
2	NA	C	424	1/1	0.97	0.22	0.55	90,90,90,90	0
3	ASP	A	425	9/9	0.94	0.18	0.00	111,113,114,114	0
2	NA	C	423	1/1	0.99	0.14	-0.85	78,78,78,78	0
3	ASP	A	427	9/9	0.94	0.13	-0.91	121,123,123,123	0
2	NA	B	423	1/1	0.94	0.10	-1.52	65,65,65,65	0
2	NA	A	423	1/1	0.97	0.08	-1.77	57,57,57,57	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.