



wwPDB X-ray Structure Validation Summary Report i

Feb 27, 2014 – 12:14 AM GMT

PDB ID : 2WNN
Title : STRUCTURE OF WILD TYPE E. COLI N-ACETYLNEURAMINIC ACID
LYASE IN COMPLEX WITH PYRUVATE IN SPACE GROUP P21
Authors : Campeotto, I.; Bolt, A.H.; Harman, T.A.; Trinh, C.H.; Dennis, C.A.; Phillips,
S.E.V.; Pearson, A.R.; Nelson, A.; Berry, A.
Deposited on : 2009-07-13
Resolution : 1.65 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

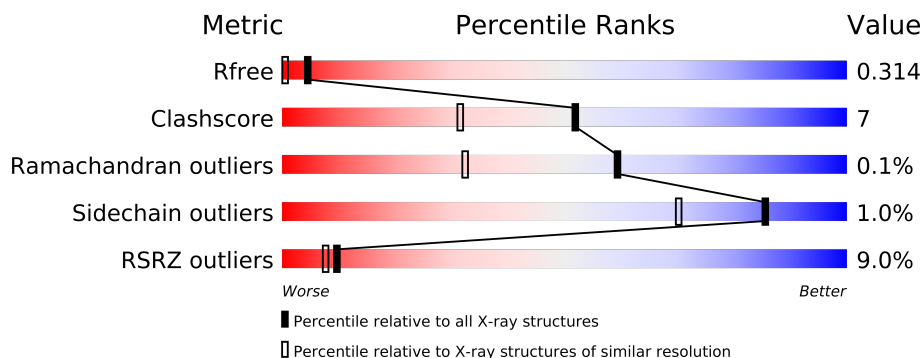
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 1.65 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}	66092	1404 (1.68-1.64)
Clashscore	79885	1001 (1.66-1.66)
Ramachandran outliers	78287	1581 (1.68-1.64)
Sidechain outliers	78261	1580 (1.68-1.64)
RSRZ outliers	66119	1404 (1.68-1.64)

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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	303	
1	B	303	
1	C	303	
1	D	303	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
3	1PE	B	1299	-	X
3	1PE	C	1297	-	X

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 9955 atoms, of which 0 are hydrogen and 0 are deuterium.

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 N-ACETYLNEURAMINATE LYASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	296	Total	C	N	O	S	0	2	0
			2298	1467	389	431	11			
1	B	295	Total	C	N	O	S	0	5	0
			2300	1467	384	438	11			
1	C	295	Total	C	N	O	S	0	0	0
			2271	1448	383	430	10			
1	D	297	Total	C	N	O	S	0	2	0
			2297	1464	387	435	11			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	MET	-	EXPRESSION TAG	UNP P0A6L4
A	-5	GLU	-	EXPRESSION TAG	UNP P0A6L4
A	-4	HIS	-	EXPRESSION TAG	UNP P0A6L4
A	-3	HIS	-	EXPRESSION TAG	UNP P0A6L4
A	-2	HIS	-	EXPRESSION TAG	UNP P0A6L4
A	-1	HIS	-	EXPRESSION TAG	UNP P0A6L4
A	0	HIS	-	EXPRESSION TAG	UNP P0A6L4
A	1	HIS	-	EXPRESSION TAG	UNP P0A6L4
B	-6	MET	-	EXPRESSION TAG	UNP P0A6L4
B	-5	GLU	-	EXPRESSION TAG	UNP P0A6L4
B	-4	HIS	-	EXPRESSION TAG	UNP P0A6L4
B	-3	HIS	-	EXPRESSION TAG	UNP P0A6L4
B	-2	HIS	-	EXPRESSION TAG	UNP P0A6L4
B	-1	HIS	-	EXPRESSION TAG	UNP P0A6L4
B	0	HIS	-	EXPRESSION TAG	UNP P0A6L4
B	1	HIS	-	EXPRESSION TAG	UNP P0A6L4
C	-6	MET	-	EXPRESSION TAG	UNP P0A6L4
C	-5	GLU	-	EXPRESSION TAG	UNP P0A6L4
C	-4	HIS	-	EXPRESSION TAG	UNP P0A6L4
C	-3	HIS	-	EXPRESSION TAG	UNP P0A6L4
C	-2	HIS	-	EXPRESSION TAG	UNP P0A6L4

Continued on next page...

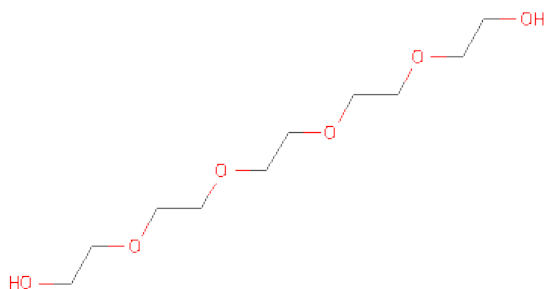
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-1	HIS	-	EXPRESSION TAG	UNP P0A6L4
C	0	HIS	-	EXPRESSION TAG	UNP P0A6L4
C	1	HIS	-	EXPRESSION TAG	UNP P0A6L4
D	-6	MET	-	EXPRESSION TAG	UNP P0A6L4
D	-5	GLU	-	EXPRESSION TAG	UNP P0A6L4
D	-4	HIS	-	EXPRESSION TAG	UNP P0A6L4
D	-3	HIS	-	EXPRESSION TAG	UNP P0A6L4
D	-2	HIS	-	EXPRESSION TAG	UNP P0A6L4
D	-1	HIS	-	EXPRESSION TAG	UNP P0A6L4
D	0	HIS	-	EXPRESSION TAG	UNP P0A6L4
D	1	HIS	-	EXPRESSION TAG	UNP P0A6L4

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total Na 1 1	0	0

- Molecule 3 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: C₁₀H₂₂O₆).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total C O 7 5 2	0	0
3	B	1	Total C O 16 10 6	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	C	1	Total	C	O	0	0
			7	5	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	172	Total	O	0	0
			172	172		
4	B	195	Total	O	0	0
			195	195		
4	C	188	Total	O	0	0
			188	188		
4	D	203	Total	O	0	0
			203	203		

i

- Molecule 1: N-ACETYLNEURAMINATE LYASE

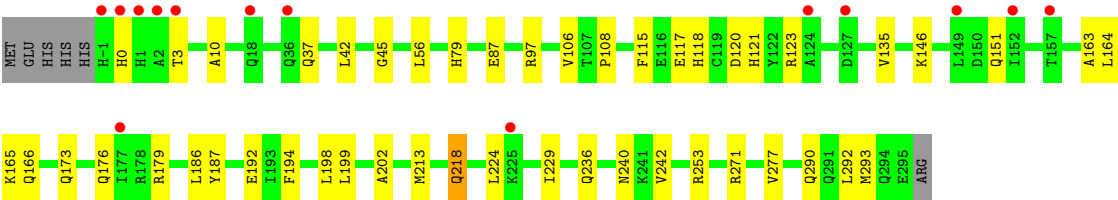
Age Group	Percentage
18-24	10%
25-34	85%
35-44	3%
45-54	2%
55-64	1%
65-74	1%
75-84	1%
85+	1%



Age Group	Percentage
18-29	80%
30-49	70%
50-64	60%
65-74	50%
75 and older	40%



Government	Percentage
Current government	95%
Previous government	5%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	54.69Å 142.46Å 83.63Å 90.00° 109.16° 90.00°	Depositor
Resolution (Å)	79.07 – 1.65 78.99 – 1.65	Depositor EDS
% Data completeness (in resolution range)	90.6 (79.07-1.65) 93.9 (78.99-1.65)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.44 (at 1.64Å)	Xtriage
Refinement program	REFMAC 5.5.0097	Depositor
R, R_{free}	0.210 , 0.249 0.266 , 0.314	Depositor DCC
R_{free} test set	6816 reflections (5.00%)	DCC
Wilson B-factor (Å ²)	20.9	Xtriage
Anisotropy	0.096	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 32.8	EDS
Estimated twinning fraction	0.666 for H,K,L 0.334 for -H,-K,H+L 0.086 for h,-k,-h-l	Xtriage
Reported twinning fraction	0.666 for H,K,L 0.334 for -H,-K,H+L	Depositor
L-test for twinning	$\langle L \rangle = 0.53$, $\langle L^2 \rangle = 0.37$	Xtriage
Outliers	0 of 136539 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	9955	wwPDB-VP
Average B, all atoms (Å ²)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.62% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: KPI, 1PE, NA

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	0.56	0/2332	0.68	0/3157
1	B	0.59	0/2340	0.72	0/3166
1	C	0.64	0/2296	0.73	0/3108
1	D	0.61	0/2329	0.71	1/3151 (0.0%)
All	All	0.60	0/9297	0.71	1/12582 (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	D	271	ARG	NE-CZ-NH2	-5.17	117.71	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2298	0	2311	37	0
1	B	2300	0	2315	38	0
1	C	2271	0	2278	31	0
1	D	2297	0	2309	38	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	1	0	0	0	0
3	B	23	0	28	7	0
3	C	7	0	6	2	0
4	A	172	0	0	3	0
4	B	195	0	0	3	0
4	C	188	0	0	3	0
4	D	203	0	0	3	0
All	All	9955	0	9247	133	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 7.

The worst 5 of 133 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:33:PHE:CD2	1:A:264:VAL:HG13	2.05	0.91
1:D:97:ARG:NH2	4:D:2094:HOH:O	2.12	0.82
1:C:79:HIS:HE1	1:C:106:VAL:H	1.27	0.80
1:A:179:ARG:HD3	1:C:237:THR:HG23	1.67	0.77
1:D:0:HIS:O	1:D:3:THR:HG22	1.84	0.76

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	295/303 (97%)	290 (98%)	5 (2%)	0	100	100
1	B	297/303 (98%)	287 (97%)	9 (3%)	1 (0%)	50	24
1	C	292/303 (96%)	286 (98%)	6 (2%)	0	100	100
1	D	296/303 (98%)	290 (98%)	6 (2%)	0	100	100
All	All	1180/1212 (97%)	1153 (98%)	26 (2%)	1 (0%)	59	32

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	3	THR

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	243/250 (97%)	240 (99%)	3 (1%)	82	63
1	B	245/250 (98%)	241 (98%)	4 (2%)	75	51
1	C	239/250 (96%)	237 (99%)	2 (1%)	89	78
1	D	243/250 (97%)	241 (99%)	2 (1%)	89	78
All	All	970/1000 (97%)	959 (99%)	11 (1%)	85	66

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

Mol	Chain	Res	Type
1	B	22	LEU
1	B	166	GLN
1	C	263	ASP
1	B	20	GLN
1	C	166	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	240	ASN
1	C	90	GLN
1	D	230	GLN
1	C	20	GLN
1	C	118	HIS

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

4 non-standard protein/DNA/RNA residues 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$
1	KPI	A	165	1	13,13,14	5.47	3 (23%)	13,15,17	2.86	4 (30%)
1	KPI	B	165	1	13,13,14	4.94	4 (30%)	13,15,17	2.76	3 (23%)
1	KPI	C	165	1	13,13,14	5.70	4 (30%)	13,15,17	2.82	4 (30%)
1	KPI	D	165	1	13,13,14	5.04	2 (15%)	13,15,17	2.06	4 (30%)

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
1	KPI	A	165	1	-	0/12/14/16	0/0/0/0
1	KPI	B	165	1	-	0/12/14/16	0/0/0/0
1	KPI	C	165	1	-	0/12/14/16	0/0/0/0
1	KPI	D	165	1	-	0/12/14/16	0/0/0/0

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	165	KPI	O-C	19.88	1.25	1.11
1	A	165	KPI	O-C	19.25	1.24	1.11
1	D	165	KPI	O-C	17.70	1.23	1.11
1	B	165	KPI	O-C	17.13	1.23	1.11
1	C	165	KPI	CX1-NZ	3.26	1.36	1.28

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	165	KPI	C-CA-N	-7.95	105.89	113.83
1	C	165	KPI	C-CA-N	-7.82	106.02	113.83

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	165	KPI	C-CA-N	-7.63	106.21	113.83
1	A	165	KPI	C1-CX1-CX2	4.47	122.08	118.39
1	B	165	KPI	C1-CX1-CX2	4.40	122.02	118.39

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 1 is monoatomic - leaving 3 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	1PE	B	1298	-	5,6,15	2.06	1 (20%)	3,5,14	0.34	0
3	1PE	B	1299	-	15,15,15	0.85	0	14,14,14	0.92	0
3	1PE	C	1297	-	5,6,15	2.33	1 (20%)	3,5,14	0.38	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	1PE	B	1298	-	-	0/4/4/13	0/0/0/0
3	1PE	B	1299	-	-	0/13/13/13	0/0/0/0
3	1PE	C	1297	-	-	0/4/4/13	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	1297	1PE	C12-C22	-4.94	1.53	1.55
3	B	1298	1PE	C12-C22	-4.45	1.53	1.55

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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, 95th 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 ²)	Q<0.9
1	A	296/303 (97%)	0.97	39 (13%) 4 3	11, 25, 37, 44	0
1	B	295/303 (97%)	0.75	30 (10%) 7 5	9, 22, 33, 40	0
1	C	295/303 (97%)	0.62	24 (8%) 12 10	9, 21, 33, 47	0
1	D	297/303 (98%)	0.56	14 (4%) 30 28	11, 21, 32, 47	0
All	All	1183/1212 (97%)	0.72	107 (9%) 10 8	9, 22, 36, 47	0

The worst 5 of 107 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	33	PHE	8.5
1	A	268	PRO	5.7
1	B	2	ALA	5.5
1	A	285	LEU	5.1
1	D	1	HIS	5.1

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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, 95th 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
1	KPI	A	165	14/15	0.11	-0.02	16,18,20,20	0
1	KPI	D	165	14/15	0.10	-0.17	11,16,17,17	0
1	KPI	B	165	14/15	0.10	-0.21	13,15,17,17	0
1	KPI	C	165	14/15	0.07	-1.76	11,13,15,15	0

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. 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, 95th 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
3	1PE	B	1299	16/16	0.26	11.04	22,34,39,40	0
3	1PE	C	1297	7/16	0.16	2.84	11,25,28,29	0
3	1PE	B	1298	7/16	0.13	0.14	34,35,37,37	0
2	NA	B	1297	1/1	0.15	-0.63	43,43,43,43	0

6.5 Other polymers ⓘ

There are no such residues in this entry.