



wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2020 – 05:50 pm BST

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 X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at 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

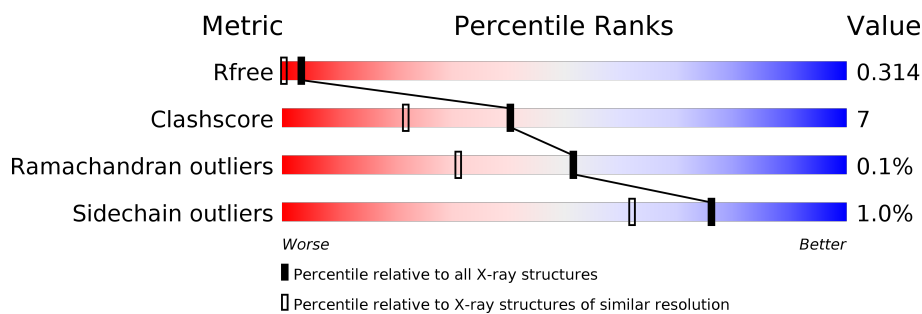
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.65 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}	130704	1827 (1.66-1.66)
Clashscore	141614	1931 (1.66-1.66)
Ramachandran outliers	138981	1891 (1.66-1.66)
Sidechain outliers	138945	1891 (1.66-1.66)

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 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 $\leq 5\%$.

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

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 9955 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 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...

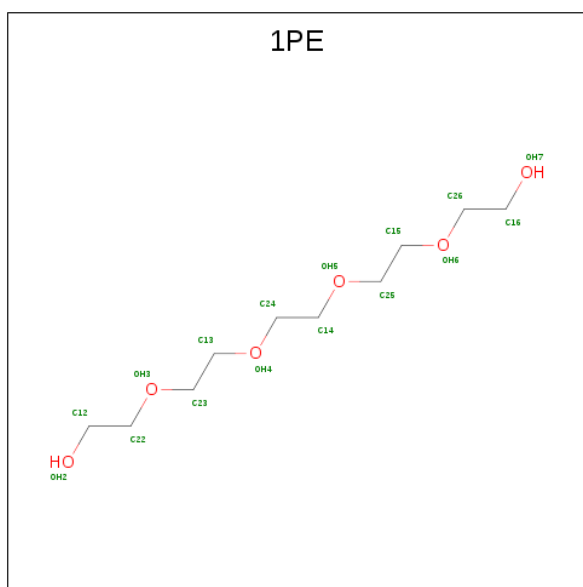
Continued from previous page...

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		

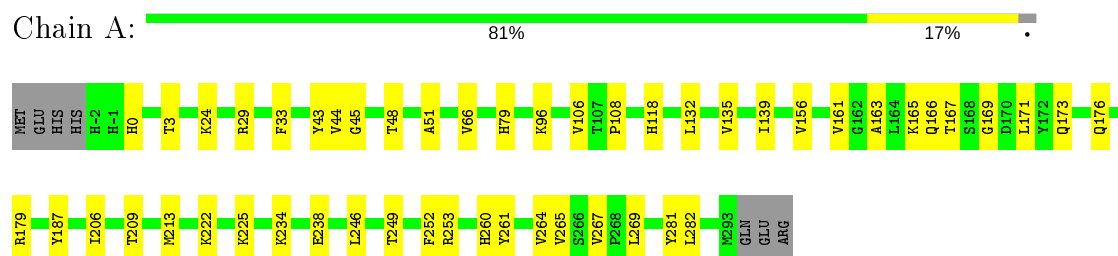
- Molecule 4 is water.

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		

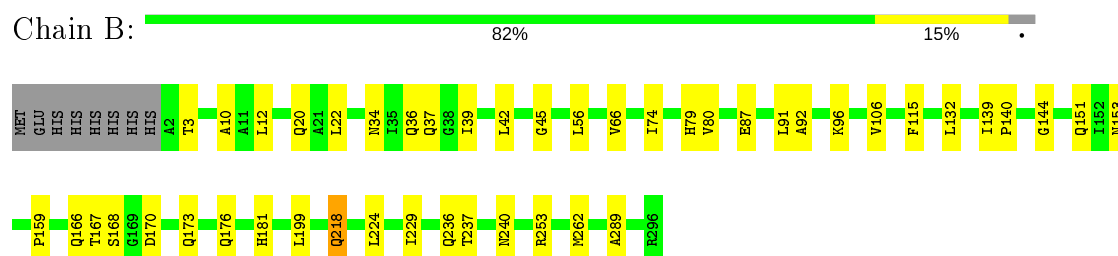
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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. 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. 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.

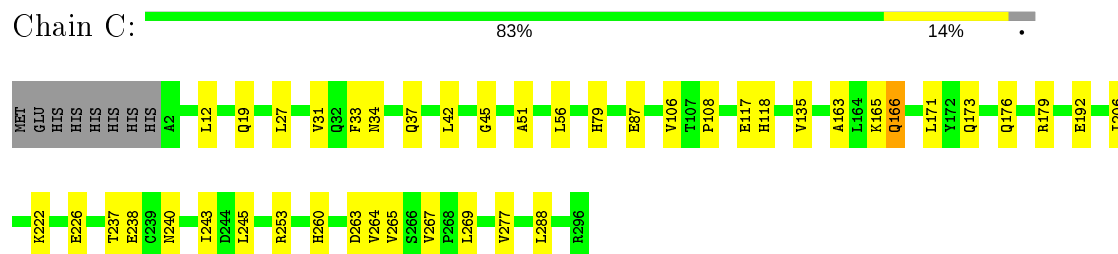
• Molecule 1: N-ACETYLNEURAMINATE LYASE



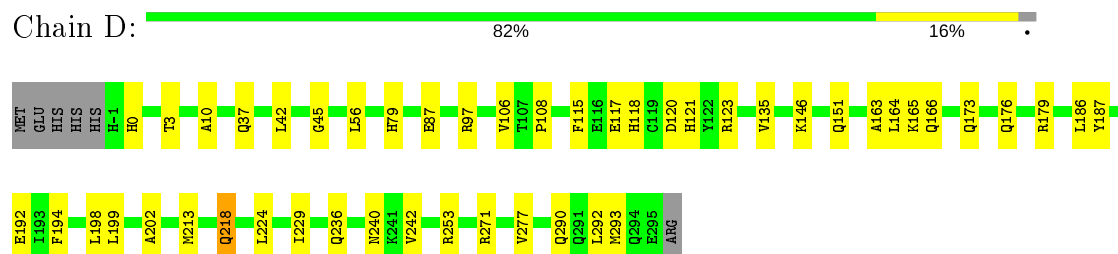
• Molecule 1: N-ACETYLNEURAMINATE LYASE



• Molecule 1: N-ACETYLNEURAMINATE LYASE



• Molecule 1: N-ACETYLNEURAMINATE LYASE



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	6825 reflections (5.00%)	wwPDB-VP
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 , 44.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.53$, $\langle L^2 \rangle = 0.37$	Xtriage
Estimated twinning fraction	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
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.

²Theoretical 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: 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 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	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
2	B	1	0	0	0	0
3	B	23	0	28	7	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
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

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 7.

The worst 5 of 133 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: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 [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	295/303 (97%)	290 (98%)	5 (2%)	0	100	100
1	B	297/303 (98%)	287 (97%)	9 (3%)	1 (0%)	41	22
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%)	51	31

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%)	71	53
1	B	245/250 (98%)	241 (98%)	4 (2%)	62	41
1	C	239/250 (96%)	237 (99%)	2 (1%)	81	70
1	D	243/250 (97%)	241 (99%)	2 (1%)	81	70
All	All	970/1000 (97%)	959 (99%)	11 (1%)	76	57

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 molecules 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	B	165	1	10,13,14	0.93	1 (10%)	6,15,17	1.67	1 (16%)
1	KPI	A	165	1	10,13,14	0.76	0	6,15,17	1.78	1 (16%)
1	KPI	D	165	1	10,13,14	0.77	0	6,15,17	1.46	1 (16%)
1	KPI	C	165	1	10,13,14	1.21	1 (10%)	6,15,17	1.62	1 (16%)

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	B	165	1	-	2/9/14/16	-
1	KPI	A	165	1	-	2/9/14/16	-
1	KPI	D	165	1	-	0/9/14/16	-
1	KPI	C	165	1	-	1/9/14/16	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	165	KPI	CX1-NZ	2.29	1.36	1.29
1	B	165	KPI	C1-CX1	2.02	1.54	1.50

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	165	KPI	C1-CX1-CX2	3.75	122.08	117.92
1	B	165	KPI	C1-CX1-CX2	3.70	122.02	117.92
1	D	165	KPI	C1-CX1-CX2	3.25	121.52	117.92
1	C	165	KPI	C1-CX1-CX2	3.23	121.50	117.92

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	B	165	KPI	C1-CX1-NZ-CE
1	B	165	KPI	CX2-CX1-NZ-CE
1	A	165	KPI	C1-CX1-NZ-CE
1	A	165	KPI	CX2-CX1-NZ-CE
1	C	165	KPI	C1-CX1-NZ-CE

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	165	KPI	1	0
1	D	165	KPI	1	0
1	C	165	KPI	1	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

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	-	6,6,15	0.58	0	5,5,14	0.38	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	1PE	B	1299	-	15,15,15	0.85	0	14,14,14	0.94	1 (7%)
3	1PE	C	1297	-	6,6,15	0.67	0	5,5,14	0.63	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	-	-	2/4/4/13	-
3	1PE	B	1299	-	-	8/13/13/13	-
3	1PE	C	1297	-	-	3/4/4/13	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	1299	1PE	OH6-C15-C25	2.00	119.42	110.39

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	1297	1PE	OH4-C13-C23-OH3
3	B	1299	1PE	OH5-C14-C24-OH4
3	B	1298	1PE	OH4-C13-C23-OH3
3	B	1299	1PE	OH2-C12-C22-OH3
3	C	1297	1PE	C12-C22-OH3-C23

There are no ring outliers.

3 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1298	1PE	2	0
3	B	1299	1PE	5	0
3	C	1297	1PE	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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.