



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 21, 2023 – 12:29 pm GMT

PDB ID : 2WNZ  
Title : Structure of the E192N mutant of E. coli N-acetylneuraminic acid lyase in complex with pyruvate in space group P21 crystal form I  
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 : 2009-07-21  
Resolution : 1.85 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 : 2.36  
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) : 2.36

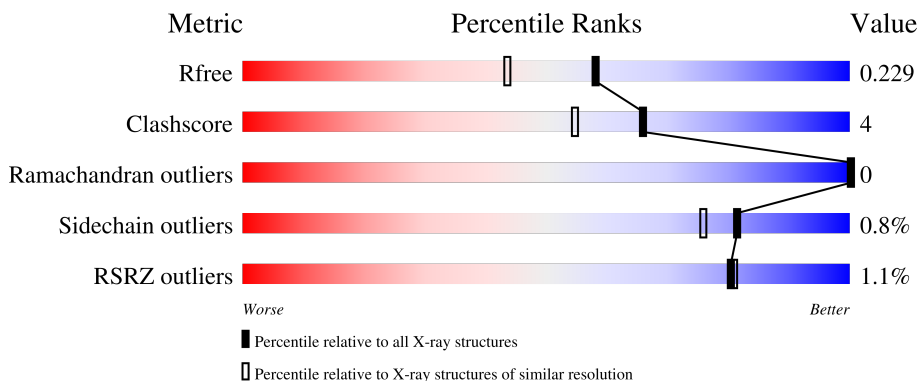
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*



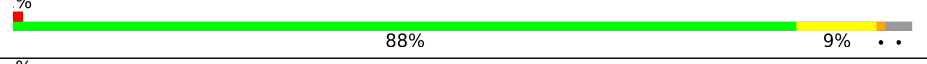

The reported resolution of this entry is 1.85 Å.

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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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	304	
1	B	304	
1	C	304	
1	D	304	

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:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	LAC	C	1297	X	X	-	-
4	2OP	D	1296	X	X	-	-

## 2 Entry composition i

There are 5 unique types of molecules in this entry. The entry contains 9970 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
			Total	C	N	O	S			
1	A	302	2368	1507	407	442	12	0	4	0
1	B	295	2293	1461	387	434	11	0	2	0
1	C	295	2304	1467	393	434	10	0	3	0
1	D	297	2327	1483	396	437	11	0	4	0

There are 36 discrepancies between the modelled and reference sequences:

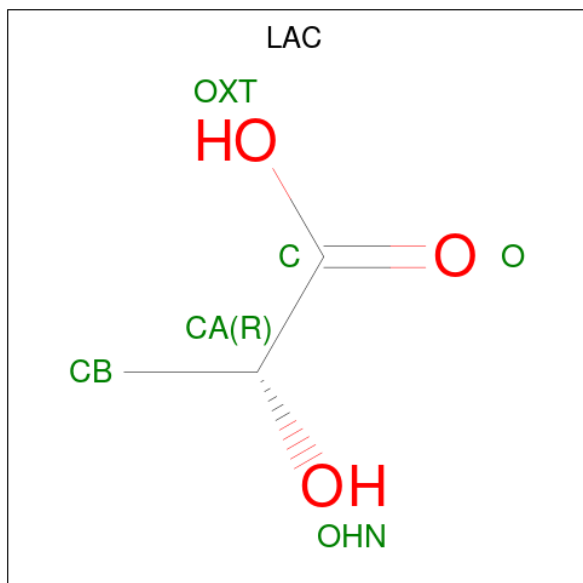
Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	MET	-	expression tag	UNP P0A6L4
A	-5	GLU	-	expression tag	UNP P0A6L4
A	-4	HIS	-	expression tag	UNP P0A6L4
A	-3	HIS	-	expression tag	UNP P0A6L4
A	-2	HIS	-	expression tag	UNP P0A6L4
A	-1	HIS	-	expression tag	UNP P0A6L4
A	0	HIS	-	expression tag	UNP P0A6L4
A	1	HIS	-	expression tag	UNP P0A6L4
A	192	ASN	GLU	conflict	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
B	192	ASN	GLU	conflict	UNP P0A6L4
C	-6	MET	-	expression tag	UNP P0A6L4
C	-5	GLU	-	expression tag	UNP P0A6L4
C	-4	HIS	-	expression tag	UNP P0A6L4

*Continued on next page...*

Continued from previous page...

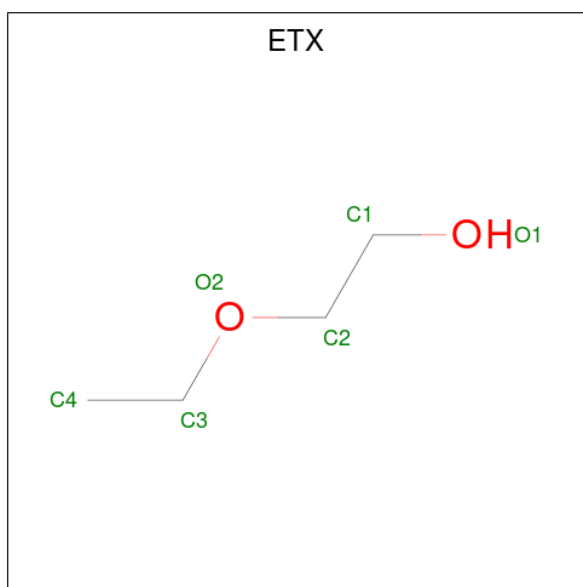
Chain	Residue	Modelled	Actual	Comment	Reference
C	-3	HIS	-	expression tag	UNP P0A6L4
C	-2	HIS	-	expression tag	UNP P0A6L4
C	-1	HIS	-	expression tag	UNP P0A6L4
C	0	HIS	-	expression tag	UNP P0A6L4
C	1	HIS	-	expression tag	UNP P0A6L4
C	192	ASN	GLU	conflict	UNP P0A6L4
D	-6	MET	-	expression tag	UNP P0A6L4
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
D	192	ASN	GLU	conflict	UNP P0A6L4

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



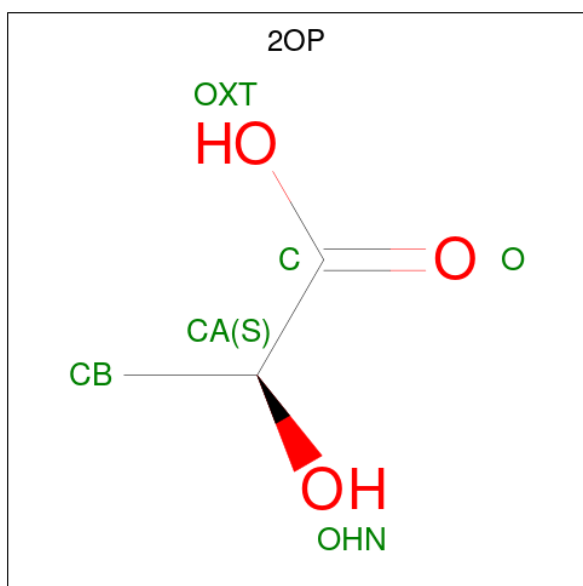
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	C	1	Total	C	O	0	0
			6	3	3		

- Molecule 3 is 2-ETHOXYETHANOL (three-letter code: ETX) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	C	1	Total	C	O	0	0
			6	4	2		
3	D	1	Total	C	O	0	0
			6	4	2		

- Molecule 4 is (2S)-2-HYDROXYPROPANOIC ACID (three-letter code: 2OP) (formula:  $C_3H_6O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	D	1	Total	C	O	0	0
			6	3	3		

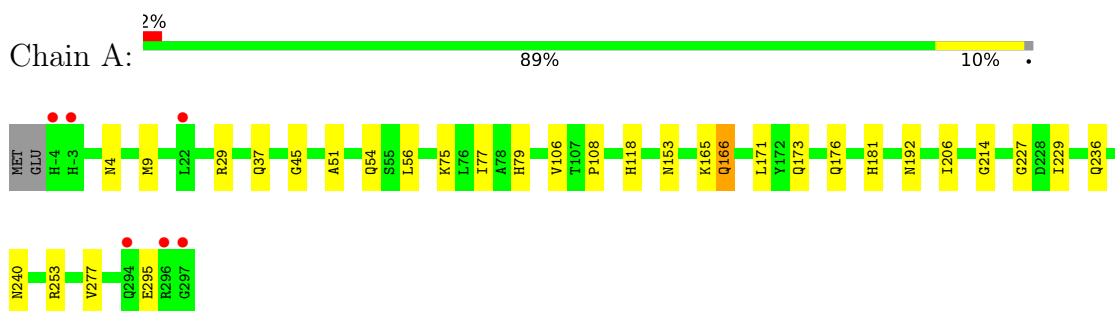
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	150	Total 150	O 150	0	0
5	B	169	Total 169	O 169	0	0
5	C	163	Total 163	O 163	0	0
5	D	172	Total 172	O 172	0	0

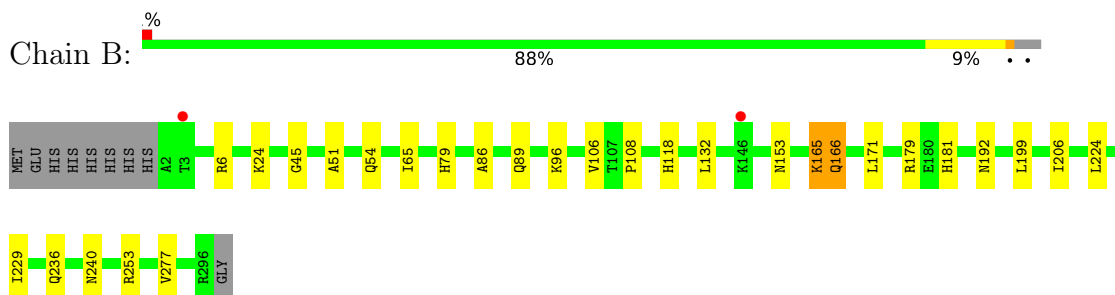
### 3 Residue-property plots [i](#)

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.

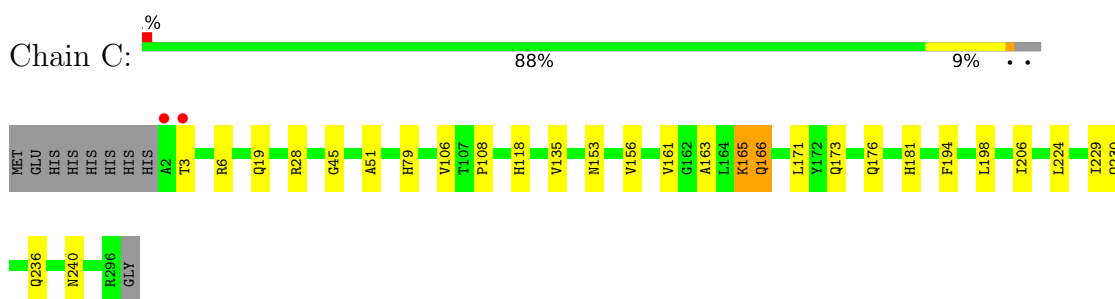
- Molecule 1: N-ACETYLNEURAMINATE LYASE



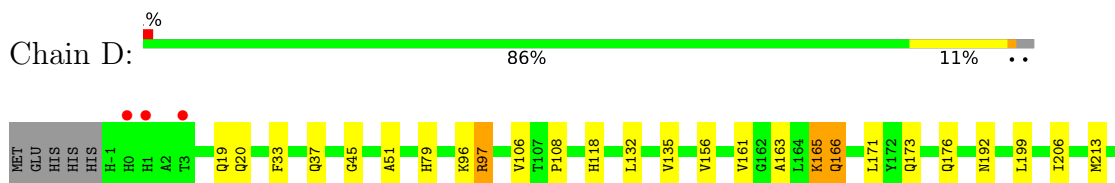
- Molecule 1: N-ACETYLNEURAMINATE LYASE



- 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, $\alpha$ , $\beta$ , $\gamma$	56.92Å 143.04Å 83.92Å 90.00° 109.79° 90.00°	Depositor
Resolution (Å)	78.96 – 1.85 71.52 – 1.85	Depositor EDS
% Data completeness (in resolution range)	100.0 (78.96-1.85) 99.9 (71.52-1.85)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.19 (at 1.86Å)	Xtrriage
Refinement program	REFMAC 5.5.0097	Depositor
R, $R_{free}$	0.188 , 0.225 0.194 , 0.229	Depositor DCC
$R_{free}$ test set	5386 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	22.6	Xtrriage
Anisotropy	0.052	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 25.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.187 for h,-k,-h-l	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	9970	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.54% 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, 0.333 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: LAC, 2OP, KPI, ETX

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	0.42	0/2411	0.57	0/3262
1	B	0.44	0/2324	0.59	0/3145
1	C	0.45	0/2338	0.60	0/3163
1	D	0.45	0/2366	0.62	2/3200 (0.1%)
All	All	0.44	0/9439	0.59	2/12770 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	97	ARG	NE-CZ-NH2	-8.08	116.26	120.30
1	D	97	ARG	NE-CZ-NH1	7.23	123.92	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	2368	0	2377	21	0
1	B	2293	0	2312	20	0
1	C	2304	0	2329	17	0
1	D	2327	0	2346	26	0
2	C	6	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	6	0	10	0	0
3	D	6	0	10	0	0
4	D	6	0	3	0	0
5	A	150	0	0	2	0
5	B	169	0	0	1	0
5	C	163	0	0	1	0
5	D	172	0	0	1	0
All	All	9970	0	9387	75	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 4.

All (75) 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:D:20[B]:GLN:HE22	1:D:272:LYS:H	1.13	0.95
1:D:97:ARG:NH2	5:D:2074:HOH:O	2.09	0.85
1:D:213[B]:MET:HE3	1:D:242:VAL:HG11	1.68	0.75
1:B:229:ILE:HD13	1:D:227:GLY:HA2	1.68	0.74
1:D:20[B]:GLN:NE2	1:D:272:LYS:H	1.85	0.73
1:A:79:HIS:HE1	1:A:106:VAL:H	1.36	0.72
1:C:28[A]:ARG:NH2	5:C:2016:HOH:O	2.09	0.72
1:A:173:GLN:HA	1:A:176:GLN:HE21	1.55	0.70
1:C:173:GLN:HA	1:C:176:GLN:HE21	1.58	0.69
1:D:173:GLN:HA	1:D:176:GLN:HE21	1.58	0.68
1:D:79:HIS:HE1	1:D:106:VAL:H	1.43	0.66
1:B:224:LEU:O	1:D:229:ILE:HD12	1.97	0.64
1:A:229:ILE:HD12	1:C:224:LEU:O	1.97	0.63
1:B:79:HIS:HE1	1:B:106:VAL:H	1.46	0.63
1:C:166:GLN:HE21	1:C:166:GLN:HA	1.63	0.63
1:C:79:HIS:HE1	1:C:106:VAL:H	1.48	0.61
1:A:166:GLN:HE21	1:A:166:GLN:HA	1.65	0.61
1:A:227:GLY:HA2	1:C:229:ILE:HD12	1.83	0.60
1:B:24:LYS:HD3	1:B:65:ILE:HD11	1.86	0.57
1:A:171:LEU:HD12	1:C:171:LEU:HD12	1.85	0.57
1:B:45:GLY:O	1:B:79:HIS:HD2	1.88	0.57
1:D:108:PRO:HD2	1:D:118:HIS:HD2	1.70	0.56
1:B:96:LYS:HD3	1:B:132:LEU:HD13	1.88	0.55
1:C:194:PHE:CE2	1:C:198:LEU:HD11	2.43	0.54
1:B:166:GLN:HA	1:B:166:GLN:HE21	1.73	0.54
1:A:51:ALA:HA	1:A:54:GLN:HE21	1.73	0.54

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:171:LEU:HD12	1:D:171:LEU:HD12	1.90	0.53
1:A:253:ARG:HG3	1:A:277:VAL:HG22	1.90	0.53
1:D:135:VAL:HG22	1:D:163:ALA:HB3	1.93	0.51
1:B:6:ARG:NH2	5:B:2004:HOH:O	2.44	0.50
1:B:86:ALA:HA	1:B:89:GLN:HE21	1.76	0.50
1:A:9[C]:MET:HE2	1:A:77:ILE:HD12	1.93	0.50
1:A:108:PRO:HD2	1:A:118:HIS:HD2	1.77	0.49
1:B:192:ASN:HA	1:B:240:ASN:HD21	1.77	0.49
1:C:236:GLN:HE21	1:C:240:ASN:ND2	2.11	0.48
1:A:192:ASN:HA	1:A:240:ASN:HD21	1.79	0.48
1:D:236:GLN:HE21	1:D:240:ASN:ND2	2.11	0.48
1:A:236:GLN:HE21	1:A:240:ASN:ND2	2.12	0.47
1:C:153:ASN:OD1	1:C:181:HIS:HE1	1.98	0.47
1:D:253:ARG:HG3	1:D:277:VAL:HG22	1.96	0.47
1:A:45:GLY:O	1:A:79:HIS:HD2	1.97	0.47
1:C:108:PRO:HD2	1:C:118:HIS:HD2	1.80	0.47
1:D:213[B]:MET:CE	1:D:292:LEU:HD11	2.45	0.47
1:A:236:GLN:HE21	1:A:240:ASN:HD21	1.63	0.47
1:D:96:LYS:HD3	1:D:132:LEU:HD13	1.97	0.46
1:C:165:KPI:H1A	1:C:206:ILE:HG22	1.97	0.46
1:D:165:KPI:H1A	1:D:206:ILE:HG22	1.97	0.46
1:B:51:ALA:HA	1:B:54:GLN:HE21	1.80	0.45
1:C:236:GLN:HE21	1:C:240:ASN:HD21	1.64	0.45
1:A:181:HIS:HD2	5:A:2003:HOH:O	2.00	0.45
1:B:108:PRO:HD2	1:B:118:HIS:HD2	1.82	0.44
1:D:192:ASN:HA	1:D:240:ASN:HD21	1.83	0.44
1:D:236:GLN:HE21	1:D:240:ASN:HD21	1.65	0.44
1:A:153:ASN:OD1	1:A:181:HIS:HE1	2.01	0.43
1:D:166:GLN:HE21	1:D:166:GLN:HA	1.83	0.43
1:B:199:LEU:HD22	1:D:199:LEU:HD22	2.01	0.43
1:A:37:GLN:HE22	1:A:214:GLY:H	1.67	0.42
1:B:179:ARG:HD3	1:D:237:THR:HG23	2.00	0.42
1:B:224:LEU:O	1:D:229:ILE:CD1	2.66	0.42
1:B:165:KPI:H1A	1:B:206:ILE:HG22	2.02	0.42
1:C:156:VAL:HA	1:C:161:VAL:HG11	2.02	0.42
1:C:3:THR:HG23	1:C:6[A]:ARG:HE	1.85	0.42
1:D:45:GLY:O	1:D:51:ALA:HB2	2.20	0.41
1:C:135:VAL:HG22	1:C:163:ALA:HB3	2.02	0.41
1:C:45:GLY:O	1:C:51:ALA:HB2	2.20	0.41
1:A:29:ARG:NH1	5:A:2021:HOH:O	2.52	0.41
1:B:153:ASN:OD1	1:B:181:HIS:HE1	2.04	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:236:GLN:HE21	1:B:240:ASN:ND2	2.18	0.41
1:B:253:ARG:HG3	1:B:277:VAL:HG22	2.02	0.41
1:A:37:GLN:NE2	1:A:214:GLY:H	2.19	0.41
1:D:19:GLN:O	1:D:20[B]:GLN:HB2	2.21	0.41
1:D:156:VAL:HA	1:D:161:VAL:HG11	2.03	0.41
1:D:33:PHE:CE2	1:D:37:GLN:OE1	2.74	0.41
1:A:4:ASN:HD22	1:A:75:LYS:NZ	2.19	0.40

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	303/304 (100%)	297 (98%)	6 (2%)	0	100	100
1	B	294/304 (97%)	288 (98%)	6 (2%)	0	100	100
1	C	295/304 (97%)	289 (98%)	6 (2%)	0	100	100
1	D	297/304 (98%)	289 (97%)	8 (3%)	0	100	100
All	All	1189/1216 (98%)	1163 (98%)	26 (2%)	0	100	100

There are no Ramachandran outliers to report.

#### 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	252/250 (101%)	249 (99%)	3 (1%)	71	62
1	B	244/250 (98%)	243 (100%)	1 (0%)	91	89
1	C	245/250 (98%)	242 (99%)	3 (1%)	71	62
1	D	248/250 (99%)	247 (100%)	1 (0%)	91	89
All	All	989/1000 (99%)	981 (99%)	8 (1%)	81	76

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	56	LEU
1	A	166	GLN
1	A	295	GLU
1	B	166	GLN
1	C	19	GLN
1	C	166	GLN
1	C	230	GLN
1	D	166	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (43) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	0	HIS
1	A	4	ASN
1	A	37	GLN
1	A	54	GLN
1	A	79	HIS
1	A	118	HIS
1	A	138	ASN
1	A	166	GLN
1	A	176	GLN
1	A	181	HIS
1	A	240	ASN
1	B	32	GLN
1	B	37	GLN
1	B	54	GLN
1	B	79	HIS
1	B	89	GLN
1	B	118	HIS
1	B	138	ASN
1	B	166	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	181	HIS
1	B	240	ASN
1	B	260	HIS
1	C	79	HIS
1	C	89	GLN
1	C	118	HIS
1	C	138	ASN
1	C	166	GLN
1	C	176	GLN
1	C	181	HIS
1	C	230	GLN
1	C	240	ASN
1	C	260	HIS
1	C	291	GLN
1	C	294	GLN
1	D	37	GLN
1	D	79	HIS
1	D	118	HIS
1	D	138	ASN
1	D	166	GLN
1	D	176	GLN
1	D	181	HIS
1	D	230	GLN
1	D	240	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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	11,13,14	1.03	0	10,15,17	1.01	1 (10%)
1	KPI	B	165	1	11,13,14	1.05	1 (9%)	10,15,17	1.67	2 (20%)
1	KPI	C	165	1	11,13,14	1.23	1 (9%)	10,15,17	1.09	1 (10%)
1	KPI	D	165	1	11,13,14	1.11	1 (9%)	10,15,17	0.73	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
1	KPI	A	165	1	-	0/13/14/16	-
1	KPI	B	165	1	-	0/13/14/16	-
1	KPI	C	165	1	-	0/13/14/16	-
1	KPI	D	165	1	-	0/13/14/16	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	165	KPI	CX2-CX1	2.30	1.52	1.49
1	D	165	KPI	O1-CX2	-2.21	1.24	1.30
1	B	165	KPI	CX2-CX1	2.19	1.52	1.49

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	165	KPI	O2-CX2-CX1	-3.91	116.38	121.38
1	C	165	KPI	O2-CX2-CX1	-2.75	117.87	121.38
1	B	165	KPI	C1-CX1-CX2	2.41	120.50	118.17
1	A	165	KPI	O2-CX2-CX1	-2.16	118.62	121.38

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 3 short contacts:

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

4 ligands 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
4	2OP	D	1296	-	5,5,5	3.08	3 (60%)	4,6,6	2.77	3 (75%)
2	LAC	C	1297	-	5,5,5	3.05	3 (60%)	4,6,6	2.64	3 (75%)
3	ETX	D	1297	-	5,5,5	0.59	0	4,4,4	0.35	0
3	ETX	C	1298	-	5,5,5	0.71	0	4,4,4	0.21	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
4	2OP	D	1296	-	1/1/2/2	2/4/4/4	-
2	LAC	C	1297	-	1/1/2/2	2/4/4/4	-
3	ETX	D	1297	-	-	2/3/3/3	-
3	ETX	C	1298	-	-	2/3/3/3	-

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	1296	2OP	CA-C	-5.20	1.38	1.51
2	C	1297	LAC	CA-C	-5.15	1.38	1.51
4	D	1296	2OP	OHN-CA	-3.35	1.31	1.42
2	C	1297	LAC	OHN-CA	-3.18	1.31	1.42
2	C	1297	LAC	O-C	3.12	1.31	1.22
4	D	1296	2OP	O-C	3.02	1.31	1.22

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	1296	2OP	CB-CA-C	4.12	119.96	110.63
2	C	1297	LAC	CB-CA-C	3.60	118.79	110.63
2	C	1297	LAC	OHN-CA-CB	2.85	122.64	110.54
4	D	1296	2OP	OHN-CA-CB	2.74	122.16	110.54
2	C	1297	LAC	OHN-CA-C	2.40	118.56	109.75
4	D	1296	2OP	OHN-CA-C	2.22	117.89	109.75

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	C	1297	LAC	CA
4	D	1296	2OP	CA

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	1297	LAC	O-C-CA-CB
4	D	1296	2OP	OXT-C-CA-OHN
3	D	1297	ETX	O1-C1-C2-O2
2	C	1297	LAC	OXT-C-CA-CB
4	D	1296	2OP	O-C-CA-OHN
3	C	1298	ETX	C1-C2-O2-C3
3	D	1297	ETX	C4-C3-O2-C2
3	C	1298	ETX	C4-C3-O2-C2

There are no ring outliers.

No monomer is involved in short contacts.

## 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	301/304 (99%)	-0.37	6 (1%) 65 64	15, 23, 38, 62	0
1	B	294/304 (96%)	-0.47	2 (0%) 87 88	15, 22, 34, 48	0
1	C	294/304 (96%)	-0.55	2 (0%) 87 88	13, 19, 33, 50	0
1	D	296/304 (97%)	-0.55	3 (1%) 82 82	13, 19, 32, 52	0
All	All	1185/1216 (97%)	-0.49	13 (1%) 80 81	13, 21, 35, 62	0

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	-4	HIS	4.4
1	D	1	HIS	4.3
1	D	0	HIS	4.1
1	A	297	GLY	3.7
1	C	3	THR	3.1
1	C	2	ALA	3.0
1	A	296	ARG	3.0
1	B	3	THR	2.9
1	A	-3	HIS	2.8
1	B	146	LYS	2.6
1	A	22	LEU	2.5
1	D	3	THR	2.4
1	A	294	GLN	2.2

### 6.2 Non-standard residues in protein, DNA, RNA chains [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
1	KPI	B	165	14/15	0.96	0.09	14,16,17,17	0
1	KPI	D	165	14/15	0.96	0.08	14,15,17,18	0
1	KPI	C	165	14/15	0.97	0.07	15,16,17,18	0
1	KPI	A	165	14/15	0.97	0.07	15,17,18,19	0

### 6.3 Carbohydrates [i](#)

There are no monosaccharides 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
3	ETX	D	1297	6/6	0.70	0.16	49,50,51,51	0
3	ETX	C	1298	6/6	0.72	0.16	49,49,50,50	0
2	LAC	C	1297	6/6	0.72	0.20	32,35,36,37	0
4	2OP	D	1296	6/6	0.79	0.24	45,46,47,49	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.