



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 26, 2024 – 02:03 PM EST

PDB ID : 9DK1  
Title : Lexapeptide dehydratase complex LxmKY apo  
Authors : Randall, G.T.; Bashiri, G.  
Deposited on : 2024-09-07  
Resolution : 2.40 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 : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.21  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.004 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.40

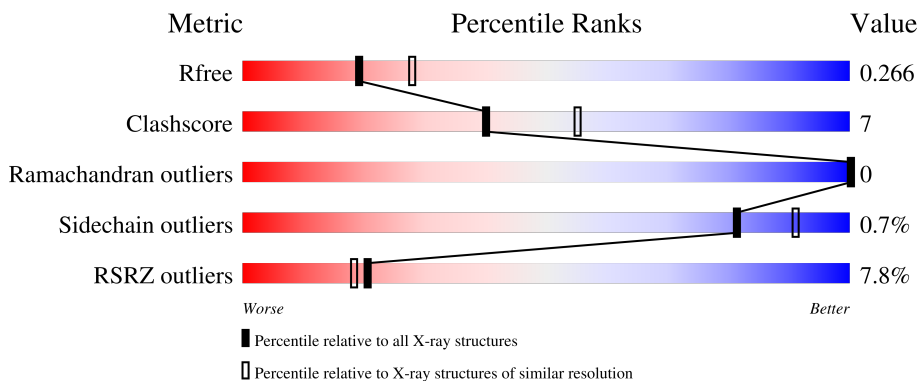
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	4642 (2.40-2.40)
Clashscore	180529	5218 (2.40-2.40)
Ramachandran outliers	177936	5158 (2.40-2.40)
Sidechain outliers	177891	5159 (2.40-2.40)
RSRZ outliers	164620	4642 (2.40-2.40)

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	403	 6% 67% 15% 17%
2	B	378	 7% 72% 11% 16%

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 4899 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 Kinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	335	2500	1579	452	462	7	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP A0A0K1TP15
A	0	ALA	-	expression tag	UNP A0A0K1TP15

- Molecule 2 is a protein called Lyase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	316	2304	1442	439	420	3	0	1	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-1	GLY	-	expression tag	UNP A0A0K1TP21
B	0	ALA	-	expression tag	UNP A0A0K1TP21

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

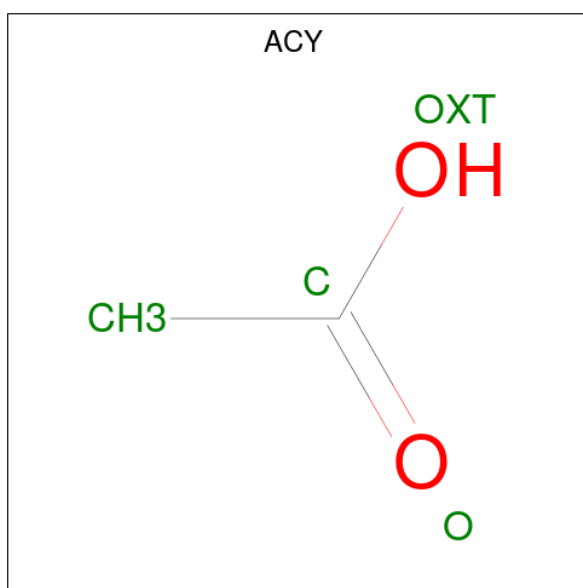
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Na	0	0
			1	1		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



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

- Molecule 5 is ACETIC ACID (three-letter code: ACY) (formula:  $C_2H_4O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			4	2	2		

- Molecule 6 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
6	A	45	Total 45	O 45	0	0
6	B	33	Total 33	O 33	0	0



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	57.70Å 90.28Å 148.95Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.62 – 2.40 48.62 – 2.40	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.62-2.40) 100.0 (48.62-2.40)	Depositor EDS
$R_{merge}$	0.43	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.41 (at 2.39Å)	Xtrriage
Refinement program	REFMAC 5.8.0415	Depositor
R, $R_{free}$	0.211 , 0.265 0.215 , 0.266	Depositor DCC
$R_{free}$ test set	1613 reflections (5.17%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	47.7	Xtrriage
Anisotropy	0.214	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 37.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	4899	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	53.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.55% 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: GOL, NA, ACY

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	0.52	0/2557	0.76	0/3489
2	B	0.50	0/2361	0.81	1/3230 (0.0%)
All	All	0.51	0/4918	0.78	1/6719 (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	6
2	B	0	4
All	All	0	10

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	B	181	PRO	N-CA-C	-5.74	97.18	112.10

There are no chirality outliers.

All (10) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	182	ARG	Sidechain
1	A	207	ARG	Sidechain
1	A	211	ARG	Sidechain
1	A	229	ARG	Sidechain
1	A	308	ARG	Sidechain

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	361	ARG	Sidechain
2	B	105	ARG	Sidechain
2	B	162	ARG	Sidechain
2	B	55	ARG	Sidechain
2	B	90	ARG	Sidechain

## 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	2500	0	2432	40	1
2	B	2304	0	2251	22	1
3	A	1	0	0	0	0
4	A	6	0	8	0	0
4	B	6	0	8	0	0
5	A	4	0	3	0	0
6	A	45	0	0	0	0
6	B	33	0	0	0	0
All	All	4899	0	4702	62	1

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.

All (62) 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:290:THR:HG22	1:A:292:GLU:H	1.57	0.69
2:B:180:ARG:CB	2:B:181:PRO:HD3	2.26	0.66
2:B:221:LEU:HD11	2:B:229:ARG:HD3	1.77	0.65
1:A:37:ARG:HG2	1:A:58:THR:CG2	2.27	0.64
1:A:290:THR:HB	1:A:293:GLU:HG3	1.82	0.62
1:A:79:GLU:HB3	1:A:243:GLU:HB3	1.84	0.59
1:A:149:LEU:HD12	1:A:152:LEU:HD21	1.85	0.57
1:A:149:LEU:O	1:A:152:LEU:HG	2.05	0.57
2:B:190:LEU:HB2	2:B:233:VAL:HB	1.87	0.56
2:B:302:ALA:O	2:B:306:ILE:HG12	2.07	0.55

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:37:ARG:HG2	1:A:58:THR:HG21	1.88	0.54
1:A:53:TRP:HZ3	1:A:67:THR:CG2	2.20	0.54
2:B:355:ARG:HG3	2:B:368:LEU:HD11	1.89	0.54
1:A:80:LEU:HD22	1:A:112:GLY:HA2	1.90	0.54
1:A:178:ALA:O	1:A:182:ARG:HG2	2.07	0.53
1:A:37:ARG:HG2	1:A:58:THR:HG23	1.90	0.52
2:B:351:VAL:CG1	2:B:355:ARG:NH1	2.73	0.52
2:B:30:LEU:HD21	2:B:301:LEU:HD11	1.92	0.52
1:A:80:LEU:O	1:A:84:LEU:HG	2.10	0.51
1:A:182:ARG:HG2	1:A:182:ARG:HH11	1.76	0.51
1:A:35:LEU:O	1:A:57:THR:HB	2.11	0.51
1:A:176:TRP:O	1:A:180:GLN:HG3	2.11	0.50
1:A:53:TRP:HZ3	1:A:67:THR:HG21	1.76	0.50
1:A:327:ASP:OD2	1:A:333:ARG:NH2	2.43	0.50
1:A:51:ASP:O	1:A:52:ASN:ND2	2.45	0.49
1:A:90:ALA:O	1:A:93:LEU:HB2	2.13	0.49
1:A:152:LEU:C	1:A:152:LEU:HD12	2.33	0.49
1:A:209:LEU:O	1:A:213:VAL:HG13	2.12	0.49
2:B:22:LEU:HD22	2:B:26:LEU:HD23	1.94	0.48
2:B:176:ARG:CZ	2:B:213:ALA:HA	2.43	0.48
2:B:193:ARG:HD2	2:B:258:THR:O	2.13	0.48
1:A:331:PRO:HB3	1:A:383:LEU:O	2.13	0.48
1:A:28:ALA:O	1:A:32:ARG:HG3	2.14	0.48
2:B:249:THR:O	2:B:253:SER:HB2	2.14	0.47
2:B:22:LEU:HD11	2:B:350:VAL:CG1	2.44	0.47
2:B:252:VAL:O	2:B:255:TYR:N	2.38	0.47
1:A:149:LEU:HA	1:A:152:LEU:HD23	1.97	0.46
1:A:53:TRP:HB2	1:A:65:VAL:HB	1.97	0.46
1:A:226:GLN:O	1:A:239:LEU:HD12	2.15	0.46
2:B:170:PHE:O	2:B:219:LYS:HA	2.17	0.44
2:B:351:VAL:CG1	2:B:355:ARG:HH12	2.31	0.43
1:A:226:GLN:HG3	1:A:241:ASP:HB2	2.00	0.43
1:A:139:CYS:O	1:A:312:PHE:HA	2.18	0.43
2:B:176:ARG:HA	2:B:177:PRO:HD2	1.89	0.43
1:A:264:HIS:HA	1:A:268:SER:HB2	2.01	0.42
1:A:29:LEU:O	1:A:33:LEU:HG	2.19	0.42
1:A:143:GLY:O	1:A:316:TYR:HA	2.20	0.42
1:A:348:ALA:O	1:A:351:GLU:HB2	2.18	0.42
2:B:22:LEU:HG	2:B:351:VAL:CG2	2.50	0.42
1:A:373:GLY:N	1:A:374:PRO:HD3	2.35	0.42
2:B:297:ARG:CZ	2:B:359:ILE:HG12	2.49	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:210:GLU:O	1:A:215:LEU:HD21	2.20	0.42
1:A:114:MET:HB2	1:A:114:MET:HE2	1.95	0.41
2:B:187:ARG:NE	2:B:280:ASP:OD1	2.51	0.41
1:A:227:PHE:CD1	1:A:239:LEU:HD13	2.56	0.41
1:A:37:ARG:H	1:A:58:THR:HG1	1.68	0.41
1:A:300:ALA:O	1:A:304:ARG:HG3	2.21	0.41
2:B:244:GLY:O	2:B:248:VAL:HG23	2.22	0.40
1:A:29:LEU:HD23	1:A:115:VAL:HG23	2.04	0.40
1:A:262:LEU:HD13	1:A:342:MET:HE2	2.03	0.40
2:B:14:SER:HB3	2:B:17:ALA:HB2	2.04	0.40
2:B:191:HIS:O	2:B:260:THR:HA	2.21	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:293:GLU:OE2	2:B:145:SER:OG[1_455]	1.61	0.59

## 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	325/403 (81%)	315 (97%)	10 (3%)	0	100	100
2	B	311/378 (82%)	295 (95%)	16 (5%)	0	100	100
All	All	636/781 (81%)	610 (96%)	26 (4%)	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	243/302 (80%)	241 (99%)	2 (1%)	79	90
2	B	219/277 (79%)	218 (100%)	1 (0%)	86	94
All	All	462/579 (80%)	459 (99%)	3 (1%)	81	92

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	80	LEU
1	A	134	PHE
2	B	263	ASP

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

Mol	Chain	Res	Type
1	A	52	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](#)

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

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
5	ACY	A	503	-	3,3,3	1.15	0	3,3,3	1.08	0
4	GOL	B	401	-	5,5,5	0.20	0	5,5,5	0.56	0
4	GOL	A	502	-	5,5,5	0.10	0	5,5,5	0.31	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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	B	401	-	-	2/4/4/4	-
4	GOL	A	502	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	502	GOL	C1-C2-C3-O3
4	B	401	GOL	O1-C1-C2-C3
4	A	502	GOL	O2-C2-C3-O3
4	B	401	GOL	O1-C1-C2-O2

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

### 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	335/403 (83%)	0.37	25 (7%) 22 20	30, 48, 82, 115	0
2	B	316/378 (83%)	0.48	26 (8%) 19 17	26, 53, 91, 118	1 (0%)
All	All	651/781 (83%)	0.42	51 (7%) 20 18	26, 51, 86, 118	1 (0%)

All (51) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	42	ALA	6.0
1	A	94	THR	5.8
1	A	384	SER	5.7
1	A	50	ASN	5.4
2	B	13	ALA	5.3
1	A	51	ASP	5.3
2	B	344	PRO	5.1
1	A	39	ASP	4.8
2	B	79	VAL	4.5
2	B	88	THR	4.5
1	A	21	ASP	4.4
1	A	154	PRO	4.4
1	A	374	PRO	4.3
1	A	152	LEU	4.2
2	B	143	SER	4.1
1	A	95	PRO	4.0
2	B	288	GLN	3.8
1	A	76	GLY	3.7
1	A	288	GLY	3.5
1	A	235	GLY	3.2
2	B	182	ALA	3.2
1	A	38	LEU	3.2
1	A	289	LEU	3.1
1	A	326	LEU	3.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	63	ASN	2.9
2	B	120	PRO	2.7
2	B	286	PRO	2.7
2	B	119	ARG	2.6
1	A	40	ALA	2.6
2	B	287	GLY	2.6
1	A	92	ARG	2.6
2	B	16	VAL	2.6
1	A	70	PRO	2.6
2	B	145	SER	2.5
2	B	180	ARG	2.5
2	B	311	ALA	2.5
2	B	146	VAL	2.5
2	B	121	ASP	2.5
2	B	144	PRO	2.3
1	A	96	ALA	2.2
2	B	75	ARG	2.2
2	B	181	PRO	2.2
1	A	234	ALA	2.2
1	A	291	HIS	2.2
2	B	253	SER	2.1
2	B	15	PRO	2.1
1	A	271	ALA	2.1
2	B	301	LEU	2.1
2	B	73	ARG	2.1
2	B	289	SER	2.0
1	A	41	ALA	2.0

## 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 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(Å <sup>2</sup> )	Q<0.9
5	ACY	A	503	4/4	0.84	0.17	62,63,72,77	0
4	GOL	A	502	6/6	0.85	0.22	47,59,62,76	0
4	GOL	B	401	6/6	0.90	0.18	68,71,81,84	0
3	NA	A	501	1/1	0.95	0.12	39,39,39,39	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.