



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 11:45 AM GMT

PDB ID : 3PUO
Title : Crystal structure of dihydrodipicolinate synthase from *Pseudomonas aeruginosa* (PsDHDPS) complexed with L-lysine at 2.65 Å resolution
Authors : Kaur, N.; Kumar, M.; Kumar, S.; Gautam, A.; Sinha, M.; Kaur, P.; Sharma, S.; Sharma, R.; Tewari, R.; Singh, T.P.
Deposited on : 2010-12-06
Resolution : 2.65 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
<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

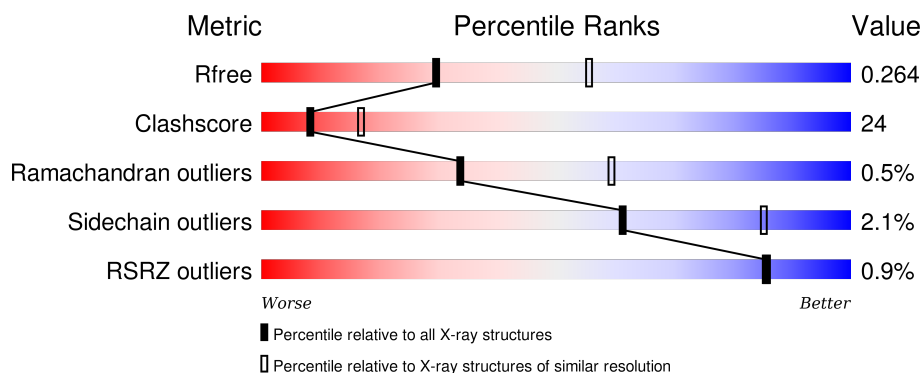
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.65 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}	91344	3152 (2.70-2.62)
Clashscore	102246	3524 (2.70-2.62)
Ramachandran outliers	100387	3469 (2.70-2.62)
Sidechain outliers	100360	3469 (2.70-2.62)
RSRZ outliers	91569	3161 (2.70-2.62)

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. 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	292	 61% 38% •
1	B	292	 59% 39% •

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	LYS	A	295	-	-	-	X

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4762 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.

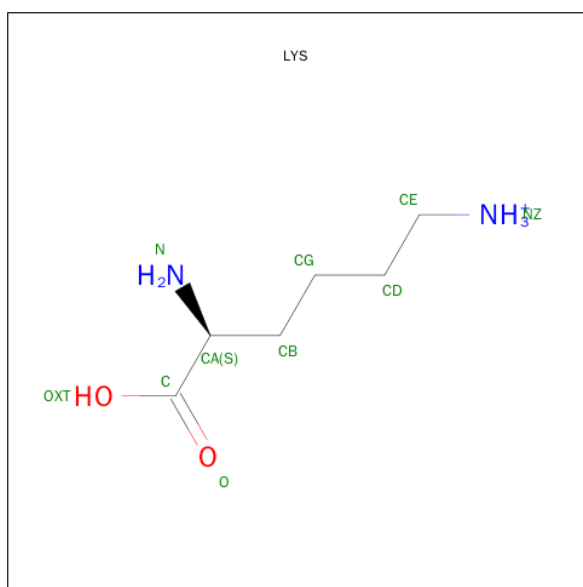
- Molecule 1 is a protein called Dihydrodipicolinate synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	292	Total	C	N	O	S	0	0	0
			2200	1382	392	412	14			
1	B	292	Total	C	N	O	S	0	0	0
			2200	1382	392	412	14			

There are 8 discrepancies between the modelled and reference sequences:

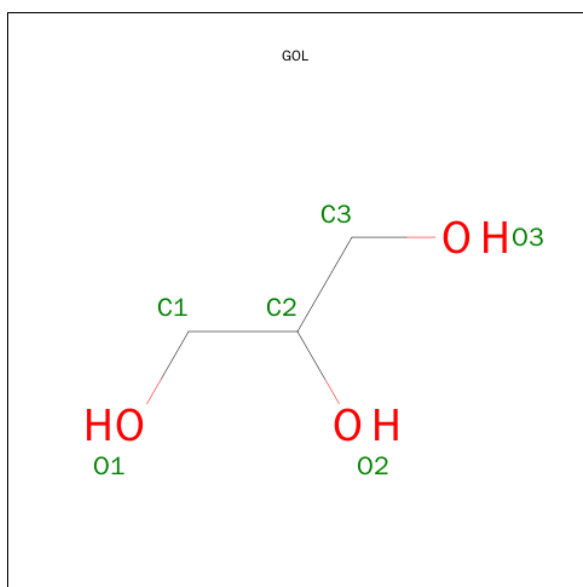
Chain	Residue	Modelled	Actual	Comment	Reference
A	34	ASP	GLU	ENGINEERED MUTATION	UNP D1MH64
A	60	VAL	ILE	ENGINEERED MUTATION	UNP D1MH64
A	234	ASP	GLU	ENGINEERED MUTATION	UNP D1MH64
A	279	ASP	GLU	ENGINEERED MUTATION	UNP D1MH64
B	34	ASP	GLU	ENGINEERED MUTATION	UNP D1MH64
B	60	VAL	ILE	ENGINEERED MUTATION	UNP D1MH64
B	234	ASP	GLU	ENGINEERED MUTATION	UNP D1MH64
B	279	ASP	GLU	ENGINEERED MUTATION	UNP D1MH64

- Molecule 2 is LYSINE (three-letter code: LYS) (formula: C₆H₁₅N₂O₂).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			10	6	2	2		
2	A	1	Total	C	N	O	0	0
			10	6	2	2		
2	B	1	Total	C	N	O	0	0
			10	6	2	2		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	153	Total 153	O 153	0	0
4	B	173	Total 173	O 173	0	0

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

Chain A:

61% 38%

Item	Category
G199	Green
M202	Green
I203	Green
S204	Green
A207	Green
R212	Green
L217	Green
I232	Green
M237	Green
P238	Green
L239	Green
H240	Green
K241	Green
F244	Green
I245	Green
E246	Green
S247	Green
M248	Green
P249	Green
I250	Green
P251	Green
W254	Green
H257	Green
P263	Green
E264	Green
G265	Green
I266	Green
L267	Green
L268	Green
P269	Green
L270	Green
T271	Green
W272	Green
L273	Green
H276	Green
C277	Green
H278	Green
D279	Green
P280	Green
L281	Green
A284	Green
T288	Green
G289	Green
V290	Green
L291	Green
A292	Green
C100	Yellow
L101	Yellow
V103	Yellow
T104	Yellow
F105	Yellow
F106	Yellow
M108	Yellow
T111	Yellow
G114	Yellow
R120	Yellow
A123	Yellow
E124	Yellow
A125	Yellow
V126	Yellow
Q130	Yellow
F131	Yellow
L132	Yellow
V133	Yellow
M134	Yellow
V135	Yellow
R138	Yellow
C141	Yellow
D142	Yellow
M143	Yellow
E146	Yellow
L151	Yellow
S152	Yellow
V154	Yellow
I160	Yellow
K161	Yellow
Q168	Yellow
V173	Yellow
I174	Yellow
E175	Yellow
R176	Yellow
V177	Yellow
D180	Yellow
F181	Yellow
L182	Yellow
D188	Yellow
L194	Yellow
M1	Yellow
I2	Yellow
S5	Yellow
M6	Yellow
V7	Yellow
D14	Yellow
R18	Yellow
L19	Yellow
D20	Yellow
K21	Yellow
V28	Yellow
H31	Yellow
L32	Yellow
Q33	Yellow
D34	Yellow
N37	Yellow
A38	Yellow
V42	Yellow
T45	Yellow
L51	Yellow
D52	Yellow
E55	Yellow
H56	Yellow
I57	Yellow
R61	Yellow
R62	Yellow
R70	Yellow
I71	Yellow
P72	Yellow
V73	Yellow
I74	Yellow
A79	Yellow
R83	Yellow
E84	Yellow
L88	Yellow
T89	Yellow
E90	Yellow
A91	Yellow
A92	Yellow
A93	Yellow
S94	Yellow
G95	Yellow
D98	Yellow
V99	Yellow

Chain B:

Category	Value
M1	1
I2	1
S5	1
M6	1
V7	1
A8	1
L9	1
V10	1
L19	1
D22	1
S23	1
L24	1
V28	1
H31	1
L32	1
Q33	1
D34	1
N37	1
A38	1
L39	1
V40	1
A41	1
V42	1
T45	1
G46	1
E47	1
L51	1
D52	1
V53	1
E54	1
E55	1
H56	1
I57	1
Q58	1
V59	1
V60	1
V67	1
R70	1
I71	1
P72	1
V73	1
I74	1
R83	1
A91	1
A92	1
K93	1
S94	1
O95	1
D98	1
V103	1
T104	1
F105	1
Y106	1
Y107	1
P110	1
T111	1
G114	1
R120	1
H121	1
I122	1
A123	1
I128	1
P129	1
Q130	1
I131	1
L132	1
Y133	1
I134	1
V135	1
C141	1
L144	1
V154	1
P155	1
N156	1
T157	1
I158	1
G159	1
I160	1
K161	1
E162	1
A163	1
L167	1
Q168	1
R169	1
A170	1
K171	1
E172	1
V173	1
I174	1
E175	1
R176	1
D180	1
F181	1
L182	1
E193	1
L194	1
G199	1
K200	1
I203	1
S204	1
V205	1
T206	1
A207	1
N208	1
R212	1
D216	1
L217	1
R235	1
L236	1
M237	1
P238	1
L239	1
L243	1
P244	1
I245	1
E246	1
S247	1
N248	1
P249	1
L250	1
P251	1
V252	1
K253	1
W254	1
A255	1
L256	1
H257	1
E258	1
P263	1
E264	1
G265	1
L266	1
R267	1
L268	1
P269	1
L270	1
L271	1
H272	1
L273	1
H276	1
C277	1
H278	1
D279	1
P280	1
L281	1
M285	1
L286	1
L291	1
A292	1

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	81.34Å 81.84Å 81.78Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.27 – 2.65 33.27 – 2.65	Depositor EDS
% Data completeness (in resolution range)	99.1 (33.27-2.65) 99.2 (33.27-2.65)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.12	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.54 (at 2.65Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.216 , 0.255 0.227 , 0.264	Depositor DCC
R_{free} test set	830 reflections (5.37%)	DCC
Wilson B-factor (Å ²)	25.2	Xtriage
Anisotropy	0.348	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 59.1	EDS
Estimated twinning fraction	0.000 for -h,l,k 0.000 for -k,-h,l 0.000 for l,-k,h 0.000 for l,h,k 0.000 for k,l,h	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 16314 reflections (0.006%)	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	4762	wwPDB-VP
Average B, all atoms (Å ²)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 50.81 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 6.1392e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹ Intensities estimated from amplitudes.

² 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: GOL

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	0.68	5/2239 (0.2%)	0.77	4/3043 (0.1%)
1	B	0.71	5/2239 (0.2%)	0.77	2/3043 (0.1%)
All	All	0.70	10/4478 (0.2%)	0.77	6/6086 (0.1%)

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	106	TYR	CD2-CE2	-8.37	1.26	1.39
1	B	106	TYR	CD1-CE1	-6.84	1.29	1.39
1	A	106	TYR	CD1-CE1	-6.79	1.29	1.39
1	B	106	TYR	CD2-CE2	-6.39	1.29	1.39
1	B	107	TYR	CD2-CE2	-6.35	1.29	1.39

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	279	ASP	CB-CG-OD1	6.85	124.47	118.30
1	A	106	TYR	N-CA-C	6.62	128.88	111.00
1	A	138	ARG	O-C-N	6.52	133.14	122.70
1	A	138	ARG	C-N-CA	-6.03	106.62	121.70
1	A	138	ARG	CA-C-N	-5.23	105.69	117.20

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	2200	0	2235	111	0
1	B	2200	0	2235	109	0
2	A	20	0	24	2	0
2	B	10	0	12	0	0
3	A	6	0	8	1	0
4	A	153	0	0	8	0
4	B	173	0	0	7	0
All	All	4762	0	4514	217	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 24.

The worst 5 of 217 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:5:SER:N	1:B:37:ASN:HD22	1.44	1.14
1:B:5:SER:H	1:B:37:ASN:ND2	1.47	1.12
1:B:111:THR:HG23	1:B:114:GLY:H	1.24	1.03
1:A:237:MET:HG3	1:A:241:LYS:HE3	1.48	0.95
1:A:103:VAL:HA	1:A:133:TYR:HB3	1.49	0.94

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	290/292 (99%)	273 (94%)	16 (6%)	1 (0%)	46 72
1	B	290/292 (99%)	278 (96%)	10 (3%)	2 (1%)	26 51
All	All	580/584 (99%)	551 (95%)	26 (4%)	3 (0%)	34 59

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	95	GLY
1	B	42	VAL
1	B	95	GLY

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	232/232 (100%)	227 (98%)	5 (2%)	60	84
1	B	232/232 (100%)	227 (98%)	5 (2%)	60	84
All	All	464/464 (100%)	454 (98%)	10 (2%)	61	84

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	279	ASP
1	B	141	CYS
1	B	180	ASP
1	A	182	LEU
1	B	162	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	278	HIS
1	B	31	HIS
1	B	276	HIS
1	A	276	HIS
1	B	130	GLN

5.3.3 RNA ⓘ

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](#)

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	$\# Z > 2$
2	LYS	A	293	-	6,9,9	0.18	0	4,10,10	0.47	0
3	GOL	A	294	-	5,5,5	0.26	0	5,5,5	0.34	0
2	LYS	A	295	-	6,9,9	0.80	0	4,10,10	0.41	0
2	LYS	B	293	-	6,9,9	0.29	0	4,10,10	0.26	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
2	LYS	A	293	-	-	0/5/9/9	0/0/0/0
3	GOL	A	294	-	-	0/4/4/4	0/0/0/0
2	LYS	A	295	-	-	0/5/9/9	0/0/0/0
2	LYS	B	293	-	-	0/5/9/9	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 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	293	LYS	1	0
3	A	294	GOL	1	0
2	A	295	LYS	1	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, 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	292/292 (100%)	-0.09	4 (1%) 78 76	5, 20, 36, 48	0
1	B	292/292 (100%)	-0.11	1 (0%) 94 95	6, 19, 33, 43	0
All	All	584/584 (100%)	-0.10	5 (0%) 85 86	5, 20, 34, 48	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	276	HIS	2.6
1	A	168	GLN	2.4
1	B	292	ALA	2.3
1	A	180	ASP	2.3
1	A	153	LYS	2.1

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, 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	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	LYS	A	295	10/10	0.66	0.36	2.93	15,24,29,31	0
2	LYS	A	293	10/10	0.79	0.25	1.87	33,33,34,35	0
2	LYS	B	293	10/10	0.77	0.25	1.86	30,33,34,35	0
3	GOL	A	294	6/6	0.79	0.34	-	28,28,29,30	0

6.5 Other polymers [i](#)

There are no such residues in this entry.