Full wwPDB X-ray Structure Validation Report ⓘ

Aug 15, 2023 – 04:35 PM EDT

PDB ID	:	1XKY
Title	:	Crystal Structure of Dihydrodipicolinate Synthase DapA-2 (BA3935) from Bacillus Anthracis at 1.94A Resolution.
Authors	:	Levdikov, V.; Blagova, E.; Fogg, M.J.; Brannigan, J.A.; Milioti, N.; Wilkinson, A.J.; Wilson, K.S.
Deposited on	:	2004-09-30
Resolution	:	1.94 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*
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:

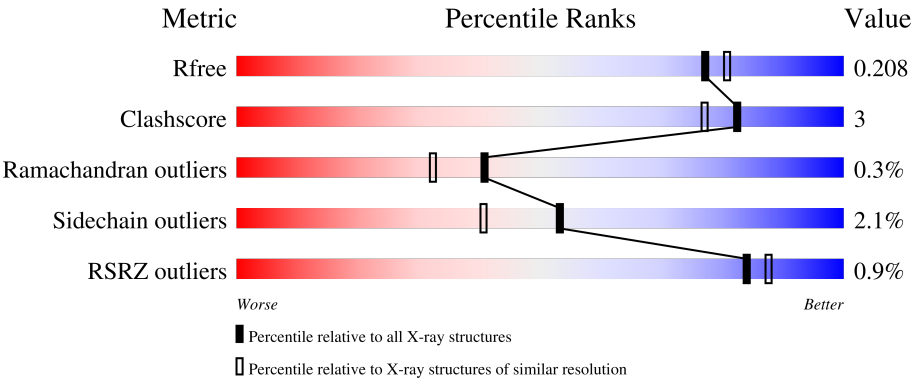
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.94 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	4310 (1.96-1.92)
Clashscore	141614	1023 (1.94-1.94)
Ramachandran outliers	138981	1007 (1.94-1.94)
Sidechain outliers	138945	1007 (1.94-1.94)
RSRZ outliers	127900	4250 (1.96-1.92)

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 >=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 <=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	301	91% 5% • •
1	B	301	% 91% 5% • •
1	C	301	% 90% 6% • •
1	D	301	% 90% 5% • •

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 10516 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
			2188	1388	361	427	12			
1	B	292	Total	C	N	O	S	0	0	0
			2188	1388	361	427	12			
1	C	292	Total	C	N	O	S	0	1	0
			2194	1392	361	429	12			
1	D	292	Total	C	N	O	S	0	0	0
			2188	1388	361	427	12			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-8	GLY	-	cloning artifact	UNP Q81WN7
A	-7	SER	-	cloning artifact	UNP Q81WN7
A	-6	SER	-	cloning artifact	UNP Q81WN7
A	-5	HIS	-	expression tag	UNP Q81WN7
A	-4	HIS	-	expression tag	UNP Q81WN7
A	-3	HIS	-	expression tag	UNP Q81WN7
A	-2	HIS	-	expression tag	UNP Q81WN7
A	-1	HIS	-	expression tag	UNP Q81WN7
A	0	HIS	-	expression tag	UNP Q81WN7
B	-8	GLY	-	cloning artifact	UNP Q81WN7
B	-7	SER	-	cloning artifact	UNP Q81WN7
B	-6	SER	-	cloning artifact	UNP Q81WN7
B	-5	HIS	-	expression tag	UNP Q81WN7
B	-4	HIS	-	expression tag	UNP Q81WN7
B	-3	HIS	-	expression tag	UNP Q81WN7
B	-2	HIS	-	expression tag	UNP Q81WN7
B	-1	HIS	-	expression tag	UNP Q81WN7
B	0	HIS	-	expression tag	UNP Q81WN7
C	-8	GLY	-	cloning artifact	UNP Q81WN7
C	-7	SER	-	cloning artifact	UNP Q81WN7
C	-6	SER	-	cloning artifact	UNP Q81WN7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-5	HIS	-	expression tag	UNP Q81WN7
C	-4	HIS	-	expression tag	UNP Q81WN7
C	-3	HIS	-	expression tag	UNP Q81WN7
C	-2	HIS	-	expression tag	UNP Q81WN7
C	-1	HIS	-	expression tag	UNP Q81WN7
C	0	HIS	-	expression tag	UNP Q81WN7
D	-8	GLY	-	cloning artifact	UNP Q81WN7
D	-7	SER	-	cloning artifact	UNP Q81WN7
D	-6	SER	-	cloning artifact	UNP Q81WN7
D	-5	HIS	-	expression tag	UNP Q81WN7
D	-4	HIS	-	expression tag	UNP Q81WN7
D	-3	HIS	-	expression tag	UNP Q81WN7
D	-2	HIS	-	expression tag	UNP Q81WN7
D	-1	HIS	-	expression tag	UNP Q81WN7
D	0	HIS	-	expression tag	UNP Q81WN7

- Molecule 2 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	2	Total K 2 2	0	0
2	B	3	Total K 3 3	0	0
2	C	3	Total K 3 3	0	0
2	D	2	Total K 2 2	0	0

- Molecule 3 is water.

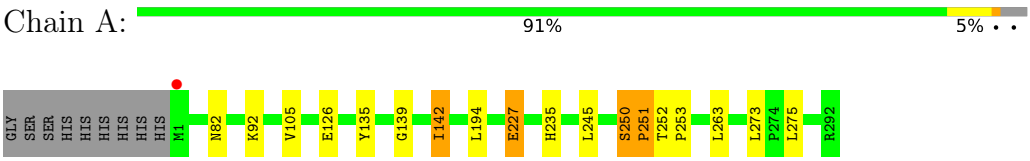
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	429	Total O 429 429	0	0
3	B	408	Total O 408 408	0	0
3	C	463	Total O 463 463	0	0
3	D	448	Total O 448 448	0	0

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.

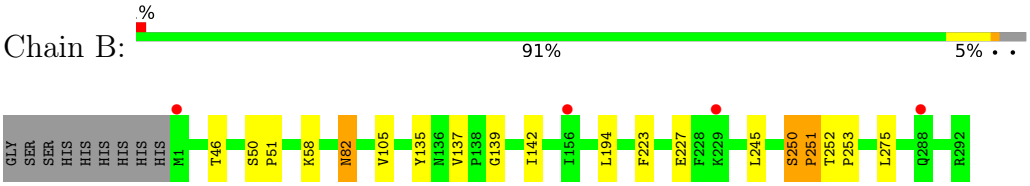
• Molecule 1: dihydrodipicolinate synthase

Chain A:



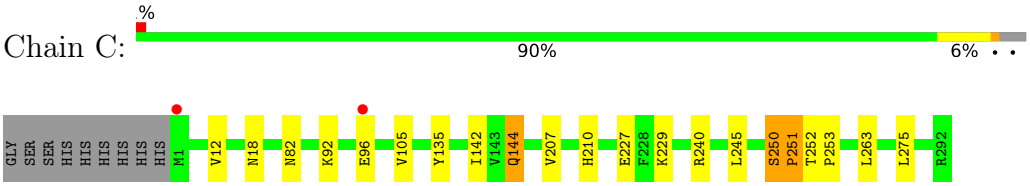
• Molecule 1: dihydrodipicolinate synthase

Chain B:



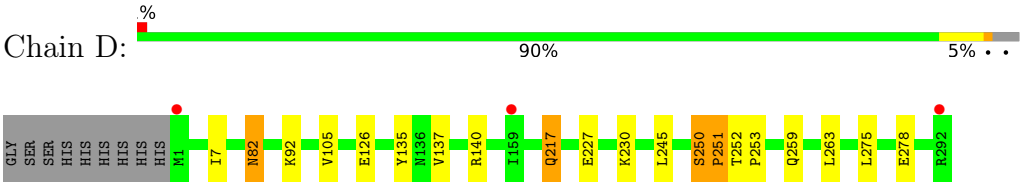
• Molecule 1: dihydrodipicolinate synthase

Chain C:



• Molecule 1: dihydrodipicolinate synthase

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	55.20Å 126.88Å 173.32Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.96 – 1.94 19.96 – 1.94	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.96-1.94) 95.7 (19.96-1.94)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.58 (at 1.94Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.153 , 0.197 0.167 , 0.208	Depositor DCC
R_{free} test set	4360 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	19.5	Xtriage
Anisotropy	0.029	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 56.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	10516	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.60% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: K

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	1/2222 (0.0%)	0.54	0/3025
1	B	0.45	1/2222 (0.0%)	0.55	0/3025
1	C	0.44	1/2231 (0.0%)	0.56	0/3037
1	D	0.44	1/2222 (0.0%)	0.57	0/3025
All	All	0.44	4/8897 (0.0%)	0.55	0/12112

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	1
1	B	0	1
1	C	0	1
1	D	0	1
All	All	0	4

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	227	GLU	CD-OE2	7.56	1.33	1.25
1	C	227	GLU	CD-OE2	7.55	1.33	1.25
1	B	227	GLU	CD-OE2	6.87	1.33	1.25
1	A	227	GLU	CD-OE2	6.36	1.32	1.25

There are no bond angle outliers.

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	250	SER	Peptide
1	B	250	SER	Peptide
1	C	250	SER	Peptide
1	D	250	SER	Peptide

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	2188	0	2245	11	0
1	B	2188	0	2245	13	0
1	C	2194	0	2251	12	0
1	D	2188	0	2245	12	0
2	A	2	0	0	0	0
2	B	3	0	0	0	0
2	C	3	0	0	0	0
2	D	2	0	0	0	0
3	A	429	0	0	2	0
3	B	408	0	0	1	0
3	C	463	0	0	3	0
3	D	448	0	0	2	0
All	All	10516	0	8986	45	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 3.

All (45) 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:C:144:GLN:HG2	3:C:2220:HOH:O	1.65	0.97
1:B:252:THR:HG22	1:B:275:LEU:HD13	1.72	0.71
1:B:252:THR:HB	1:B:253:PRO:HD3	1.77	0.66
1:D:250:SER:OG	1:D:251:PRO:HA	1.96	0.65
1:C:252:THR:HG22	1:C:275:LEU:HD13	1.78	0.65
1:A:252:THR:HG22	1:A:275:LEU:HD13	1.79	0.64
1:C:250:SER:OG	1:C:251:PRO:HA	1.98	0.62
1:A:250:SER:OG	1:A:251:PRO:HA	1.98	0.62

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:250:SER:OG	1:B:251:PRO:HA	2.00	0.61
1:C:252:THR:HB	1:C:253:PRO:HD3	1.83	0.60
1:D:217:GLN:HE21	1:D:217:GLN:HA	1.68	0.59
1:D:252:THR:HG22	1:D:275:LEU:HD13	1.84	0.58
1:D:252:THR:HB	1:D:253:PRO:HD3	1.84	0.58
1:A:252:THR:HB	1:A:253:PRO:HD3	1.87	0.56
1:D:126:GLU:HG3	3:D:2042:HOH:O	2.09	0.52
1:A:252:THR:N	1:A:253:PRO:CD	2.73	0.52
1:A:194:LEU:HD22	1:A:235:HIS:HB2	1.94	0.50
1:D:252:THR:N	1:D:253:PRO:CD	2.75	0.50
1:C:252:THR:N	1:C:253:PRO:CD	2.74	0.49
1:C:92:LYS:HG3	3:C:1987:HOH:O	2.13	0.48
1:A:105:VAL:HA	1:A:135:TYR:HB3	1.96	0.47
1:B:194:LEU:HD11	1:B:223:PHE:CD1	2.50	0.47
1:C:142:ILE:HD13	1:D:140:ARG:HA	1.96	0.47
1:B:252:THR:N	1:B:253:PRO:CD	2.79	0.46
1:A:142:ILE:HD12	1:B:139:GLY:O	2.16	0.46
1:A:139:GLY:O	1:B:142:ILE:HD13	2.16	0.44
1:D:92:LYS:HE2	3:D:1980:HOH:O	2.17	0.44
1:B:58:LYS:NZ	3:B:1857:HOH:O	2.44	0.43
1:B:105:VAL:HA	1:B:135:TYR:HB3	2.00	0.43
1:C:207:VAL:O	1:C:210:HIS:HD2	2.00	0.43
1:D:105:VAL:HA	1:D:135:TYR:HB3	2.01	0.43
1:D:7:ILE:HD12	1:D:217:GLN:HG2	2.00	0.43
1:A:126:GLU:HG3	3:A:2108:HOH:O	2.18	0.42
1:A:252:THR:HG21	1:A:273:LEU:O	2.20	0.42
1:B:252:THR:HG22	1:B:275:LEU:CD1	2.44	0.42
1:B:137:VAL:HG13	1:B:137:VAL:O	2.19	0.42
1:C:18:ASN:ND2	3:C:2085:HOH:O	2.46	0.42
1:A:92:LYS:HG2	3:A:2029:HOH:O	2.19	0.42
1:D:82:ASN:N	1:D:82:ASN:HD22	2.17	0.42
1:B:82:ASN:N	1:B:82:ASN:HD22	2.18	0.41
1:C:92:LYS:O	1:C:96:GLU:HG3	2.20	0.41
1:C:12:VAL:H	1:C:210:HIS:CE1	2.38	0.41
1:C:105:VAL:HA	1:C:135:TYR:HB3	2.02	0.41
1:D:137:VAL:O	1:D:137:VAL:HG13	2.20	0.41
1:B:50:SER:N	1:B:51:PRO:CD	2.84	0.41

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	290/301 (96%)	280 (97%)	9 (3%)	1 (0%)	41	32
1	B	290/301 (96%)	284 (98%)	5 (2%)	1 (0%)	41	32
1	C	291/301 (97%)	283 (97%)	7 (2%)	1 (0%)	41	32
1	D	290/301 (96%)	283 (98%)	6 (2%)	1 (0%)	41	32
All	All	1161/1204 (96%)	1130 (97%)	27 (2%)	4 (0%)	41	32

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	251	PRO
1	B	251	PRO
1	D	251	PRO
1	A	251	PRO

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	245/253 (97%)	240 (98%)	5 (2%)	55	42
1	B	245/253 (97%)	242 (99%)	3 (1%)	71	64
1	C	246/253 (97%)	240 (98%)	6 (2%)	49	36
1	D	245/253 (97%)	238 (97%)	7 (3%)	42	28
All	All	981/1012 (97%)	960 (98%)	21 (2%)	53	41

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	82	ASN
1	A	142	ILE
1	A	227	GLU
1	A	245	LEU
1	A	263	LEU
1	B	46	THR
1	B	82	ASN
1	B	245	LEU
1	C	82	ASN
1	C	144	GLN
1	C	229	LYS
1	C	240	ARG
1	C	245	LEU
1	C	263	LEU
1	D	82	ASN
1	D	217	GLN
1	D	230	LYS
1	D	245	LEU
1	D	259	GLN
1	D	263	LEU
1	D	278	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (22) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	82	ASN
1	A	119	GLN
1	A	210	HIS
1	A	236	GLN
1	A	288	GLN
1	B	20	ASN
1	B	82	ASN
1	B	119	GLN
1	B	210	HIS
1	B	217	GLN
1	B	236	GLN
1	C	18	ASN
1	C	82	ASN
1	C	119	GLN
1	C	210	HIS
1	C	288	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	20	ASN
1	D	82	ASN
1	D	119	GLN
1	D	210	HIS
1	D	217	GLN
1	D	288	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

Of 10 ligands modelled in this entry, 10 are monoatomic - leaving 0 for Mogul analysis.

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.

No monomer is involved in short contacts.

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, 95^{th} 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	292/301 (97%)	-0.14	1 (0%)	94	96	13, 18, 25, 34	2 (0%)
1	B	292/301 (97%)	-0.11	4 (1%)	75	80	14, 18, 30, 34	3 (1%)
1	C	292/301 (97%)	-0.17	2 (0%)	87	91	12, 16, 25, 36	3 (1%)
1	D	292/301 (97%)	-0.10	3 (1%)	82	86	12, 17, 25, 41	3 (1%)
All	All	1168/1204 (97%)	-0.13	10 (0%)	84	87	12, 17, 27, 41	11 (0%)

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	292	ARG	5.1
1	A	1	MET	5.1
1	C	1	MET	4.7
1	B	1	MET	4.2
1	D	1	MET	3.4
1	B	156	ILE	3.1
1	B	229	LYS	2.9
1	D	159	ILE	2.5
1	B	288	GLN	2.3
1	C	96	GLU	2.2

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

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^{th} 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^2)	Q<0.9
2	K	C	1810	1/1	0.96	0.06	27,27,27,27	0
2	K	B	1809	1/1	0.98	0.09	28,28,28,28	0
2	K	B	1802	1/1	0.99	0.05	19,19,19,19	0
2	K	A	1806	1/1	0.99	0.04	21,21,21,21	0
2	K	C	1801	1/1	0.99	0.03	17,17,17,17	0
2	K	C	1805	1/1	0.99	0.04	19,19,19,19	0
2	K	A	1807	1/1	0.99	0.04	17,17,17,17	0
2	K	D	1804	1/1	0.99	0.03	19,19,19,19	0
2	K	D	1808	1/1	0.99	0.04	19,19,19,19	0
2	K	B	1803	1/1	1.00	0.03	19,19,19,19	0

6.5 Other polymers (i)

There are no such residues in this entry.