



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 4, 2017 – 03:55 PM EDT

PDB ID : 2A92
Title : Crystal structure of lactate dehydrogenase from Plasmodium vivax: complex with NADH
Authors : Chaikuad, A.; Fairweather, V.; Conners, R.; Joseph-Horne, T.; Turgut-Balik, D.; Brady, R.L.
Deposited on : unknown
Resolution : 2.04 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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20030345
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
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) : rb-20030345

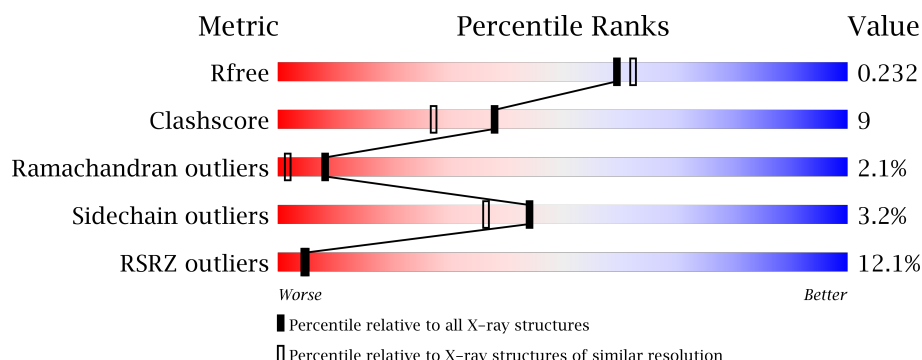
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.04 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}	100719	1316 (2.04-2.04)
Clashscore	112137	1394 (2.04-2.04)
Ramachandran outliers	110173	1383 (2.04-2.04)
Sidechain outliers	110143	1383 (2.04-2.04)
RSRZ outliers	101464	1319 (2.04-2.04)

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	321	<div> <div>12%</div> <div>87%</div> <div>8%</div> <div>...</div> </div>
1	B	321	<div> <div>5%</div> <div>85%</div> <div>10%</div> <div>..</div> </div>
1	C	321	<div> <div>12%</div> <div>85%</div> <div>12%</div> <div>..</div> </div>
1	D	321	<div> <div>19%</div> <div>83%</div> <div>12%</div> <div>..</div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 10257 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 L-lactate dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	316	Total	C	N	O	S	0	4	0
			2404	1535	404	451	14			
1	B	315	Total	C	N	O	S	0	3	0
			2389	1526	400	449	14			
1	C	318	Total	C	N	O	S	0	4	0
			2419	1544	409	452	14			
1	D	315	Total	C	N	O	S	0	3	0
			2389	1526	400	449	14			

There are 24 discrepancies between the modelled and reference sequences:

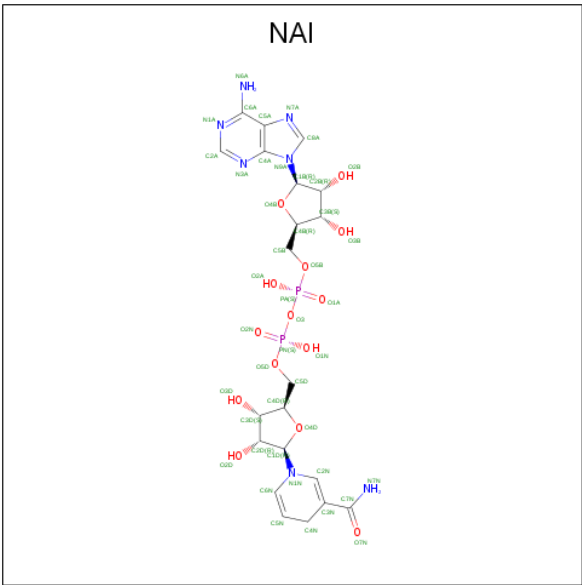
Chain	Residue	Modelled	Actual	Comment	Reference
A	330	HIS	-	EXPRESSION TAG	GB 66967948
A	331	HIS	-	EXPRESSION TAG	GB 66967948
A	332	HIS	-	EXPRESSION TAG	GB 66967948
A	333	HIS	-	EXPRESSION TAG	GB 66967948
A	334	HIS	-	EXPRESSION TAG	GB 66967948
A	335	HIS	-	EXPRESSION TAG	GB 66967948
B	330	HIS	-	EXPRESSION TAG	GB 66967948
B	331	HIS	-	EXPRESSION TAG	GB 66967948
B	332	HIS	-	EXPRESSION TAG	GB 66967948
B	333	HIS	-	EXPRESSION TAG	GB 66967948
B	334	HIS	-	EXPRESSION TAG	GB 66967948
B	335	HIS	-	EXPRESSION TAG	GB 66967948
C	330	HIS	-	EXPRESSION TAG	GB 66967948
C	331	HIS	-	EXPRESSION TAG	GB 66967948
C	332	HIS	-	EXPRESSION TAG	GB 66967948
C	333	HIS	-	EXPRESSION TAG	GB 66967948
C	334	HIS	-	EXPRESSION TAG	GB 66967948
C	335	HIS	-	EXPRESSION TAG	GB 66967948
D	330	HIS	-	EXPRESSION TAG	GB 66967948
D	331	HIS	-	EXPRESSION TAG	GB 66967948
D	332	HIS	-	EXPRESSION TAG	GB 66967948

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	333	HIS	-	EXPRESSION TAG	GB 66967948
D	334	HIS	-	EXPRESSION TAG	GB 66967948
D	335	HIS	-	EXPRESSION TAG	GB 66967948

- Molecule 2 is 1,4-DIHYDRONICOTINAMIDE ADENINE DINUCLEOTIDE (three-letter code: NAI) (formula: C₂₁H₂₉N₇O₁₄P₂).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	B	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	C	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	D	1	Total	C	N	O	P	0	0
			44	21	7	14	2		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	148	Total	O	0	0
			148	148		
3	B	130	Total	O	0	0
			130	130		
3	C	112	Total	O	0	0
			112	112		

Continued on next page...

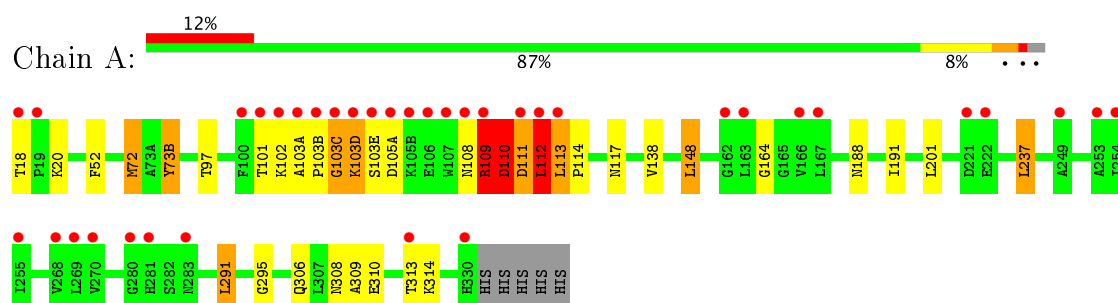
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	D	90	Total	O	0	0
			90	90		

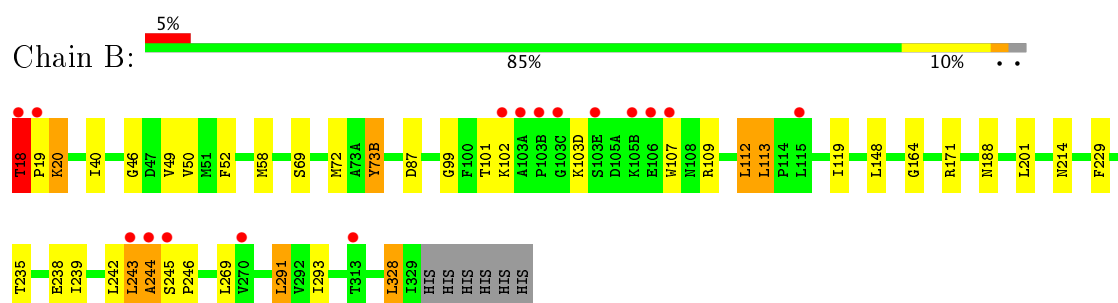
3 Residue-property plots [i](#)

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 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.

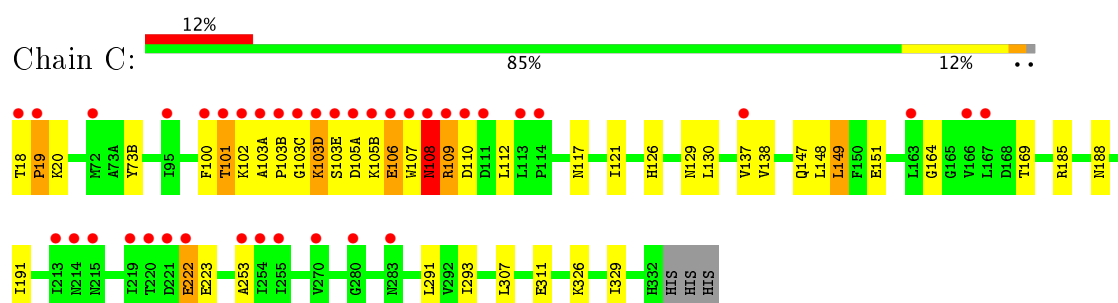
• Molecule 1: L-lactate dehydrogenase



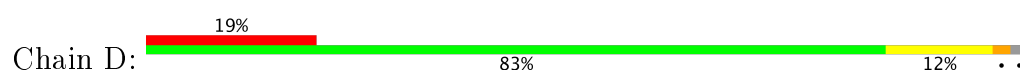
• Molecule 1: L-lactate dehydrogenase

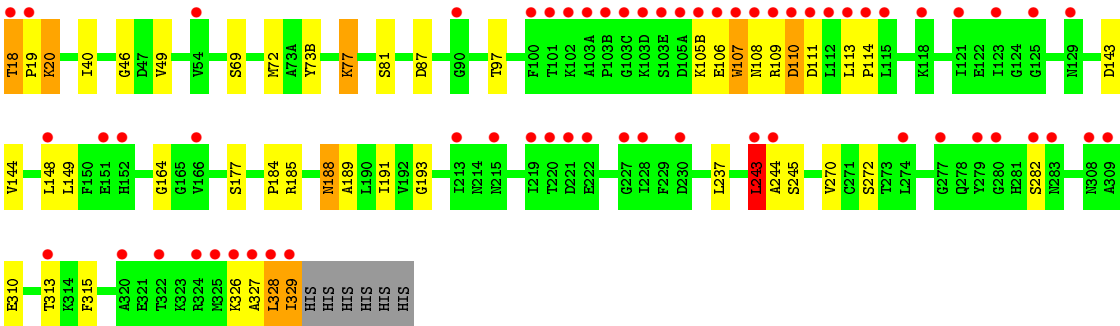


• Molecule 1: L-lactate dehydrogenase



• Molecule 1: L-lactate dehydrogenase





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	81.59 Å 128.64 Å 130.58 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	91.29 – 2.04 32.97 – 2.04	Depositor EDS
% Data completeness (in resolution range)	97.8 (91.29-2.04) 97.9 (32.97-2.04)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.07 (at 2.05 Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.194 , 0.236 0.192 , 0.232	Depositor DCC
R_{free} test set	4299 reflections (5.27%)	DCC
Wilson B-factor (Å ²)	27.1	Xtriage
Anisotropy	0.346	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 42.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.013 for -h,l,k	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	10257	wwPDB-VP
Average B, all atoms (Å ²)	9.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.12% 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: NAI

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	0.73	0/2443	0.74	2/3308 (0.1%)
1	B	0.67	0/2424	0.69	2/3282 (0.1%)
1	C	0.67	0/2457	0.63	0/3327
1	D	0.59	0/2424	0.63	0/3282
All	All	0.67	0/9748	0.67	4/13199 (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	3
1	B	0	2
1	D	0	1
All	All	0	6

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	291	LEU	CA-CB-CG	9.46	137.06	115.30
1	B	291	LEU	CA-CB-CG	7.43	132.40	115.30
1	A	113	LEU	N-CA-C	-6.92	92.33	111.00
1	B	112	LEU	CA-CB-CG	6.01	129.13	115.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	109	ARG	Peptide
1	A	111	ASP	Peptide
1	A	112	LEU	Peptide
1	B	18	THR	Peptide
1	B	244	ALA	Peptide

5.2 Too-close contacts ⓘ

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	2404	0	2496	43	0
1	B	2389	0	2483	37	0
1	C	2419	0	2502	68	0
1	D	2389	0	2483	33	0
2	A	44	0	27	3	0
2	B	44	0	27	1	0
2	C	44	0	27	1	0
2	D	44	0	27	1	0
3	A	148	0	0	2	0
3	B	130	0	0	2	0
3	C	112	0	0	1	0
3	D	90	0	0	1	0
All	All	10257	0	10072	173	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 9.

The worst 5 of 173 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:100:PHE:CG	1:C:101:THR:HB	1.42	1.55
1:C:103(C):GLY:HA3	1:C:103(D):LYS:CB	1.26	1.49
1:C:103(C):GLY:CA	1:C:103(D):LYS:HB2	1.27	1.39
1:C:100:PHE:HA	1:C:101:THR:CB	1.60	1.25
1:C:100:PHE:HA	1:C:101:THR:OG1	1.35	1.22

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	315/321 (98%)	297 (94%)	12 (4%)	6 (2%)	9	2
1	B	313/321 (98%)	304 (97%)	6 (2%)	3 (1%)	18	8
1	C	316/321 (98%)	294 (93%)	13 (4%)	9 (3%)	6	1
1	D	313/321 (98%)	295 (94%)	10 (3%)	8 (3%)	6	1
All	All	1257/1284 (98%)	1190 (95%)	41 (3%)	26 (2%)	8	2

5 of 26 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	103(C)	GLY
1	A	103(D)	LYS
1	A	110	ASP
1	A	112	LEU
1	B	243	LEU

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	269/273 (98%)	260 (97%)	9 (3%)	43	36
1	B	267/273 (98%)	258 (97%)	9 (3%)	42	34
1	C	270/273 (99%)	263 (97%)	7 (3%)	51	44
1	D	267/273 (98%)	258 (97%)	9 (3%)	42	34
All	All	1073/1092 (98%)	1039 (97%)	34 (3%)	44	36

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	188	ASN
1	C	102	LYS
1	D	243	LEU
1	B	328	LEU
1	A	188	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	188	ASN
1	B	210(B)	GLN
1	C	234	ASN
1	A	306	GLN
1	B	178	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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

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	NAI	A	601	-	40,48,48	1.79	5 (12%)	41,73,73	2.00	5 (12%)
2	NAI	B	701	-	40,48,48	1.65	5 (12%)	41,73,73	1.82	4 (9%)
2	NAI	C	501	-	40,48,48	1.70	5 (12%)	41,73,73	2.09	5 (12%)
2	NAI	D	801	-	40,48,48	1.67	5 (12%)	41,73,73	1.85	3 (7%)

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	NAI	A	601	-	-	0/25/72/72	0/5/5/5
2	NAI	B	701	-	-	0/25/72/72	0/5/5/5
2	NAI	C	501	-	-	0/25/72/72	0/5/5/5
2	NAI	D	801	-	-	0/25/72/72	0/5/5/5

The worst 5 of 20 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	601	NAI	C4N-C5N	-4.16	1.40	1.49
2	D	801	NAI	C4N-C5N	-4.08	1.40	1.49
2	B	701	NAI	C4N-C5N	-3.99	1.40	1.49
2	C	501	NAI	C4N-C5N	-3.63	1.41	1.49
2	C	501	NAI	C2A-N1A	2.57	1.38	1.33

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	501	NAI	N3A-C2A-N1A	-11.06	119.22	128.86
2	A	601	NAI	N3A-C2A-N1A	-9.66	120.44	128.86
2	D	801	NAI	N3A-C2A-N1A	-9.62	120.48	128.86
2	B	701	NAI	N3A-C2A-N1A	-8.87	121.13	128.86
2	B	701	NAI	C1D-N1N-C2N	-3.82	114.61	121.09

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	601	NAI	3	0
2	B	701	NAI	1	0
2	C	501	NAI	1	0
2	D	801	NAI	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	316/321 (98%)	0.93	37 (11%) 5 5	4, 9, 14, 20	31 (9%)
1	B	315/321 (98%)	0.42	16 (5%) 29 31	3, 9, 15, 22	34 (10%)
1	C	318/321 (99%)	1.14	39 (12%) 5 4	4, 9, 15, 21	52 (16%)
1	D	315/321 (98%)	1.62	61 (19%) 1 1	2, 10, 16, 22	55 (17%)
All	All	1264/1284 (98%)	1.03	153 (12%) 5 4	2, 9, 15, 22	172 (13%)

The worst 5 of 153 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	107	TRP	29.4
1	D	100	PHE	22.6
1	A	103(B)	PRO	22.6
1	A	103(A)	ALA	21.9
1	D	108	ASN	20.1

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	NAI	B	701	44/44	0.90	0.16	0.37	8,13,22,25	0
2	NAI	D	801	44/44	0.90	0.15	-0.28	12,17,22,25	0
2	NAI	A	601	44/44	0.93	0.12	-0.34	3,7,12,17	0
2	NAI	C	501	44/44	0.94	0.11	-0.47	2,6,9,11	0

6.5 Other polymers [i](#)

There are no such residues in this entry.