



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 11:32 pm GMT

PDB ID : 4CUK  
Title : Structure of Salmonella D-Lactate Dehydrogenase in complex with NADH  
Authors : Attarataya, J.; Zaccai, N.R.; Brady, R.L.  
Deposited on : 2014-03-19  
Resolution : 2.18 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](mailto: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 : trunk28620  
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) : recalc28949

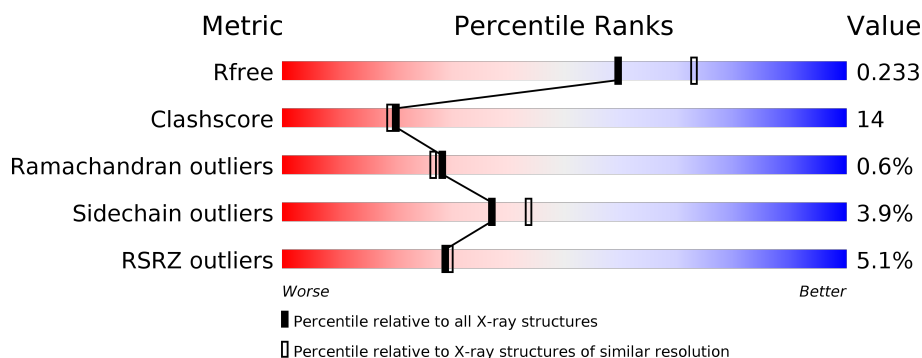
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.18 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	5526 (2.20-2.16)
Clashscore	112137	6386 (2.20-2.16)
Ramachandran outliers	110173	6282 (2.20-2.16)
Sidechain outliers	110143	6282 (2.20-2.16)
RSRZ outliers	101464	5562 (2.20-2.16)

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  $\geq 3$ , 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	338	<div> <div>9%</div> <div> <div></div> <div>72%</div> <div>21%</div> <div>5%</div> </div> </div>
1	B	338	<div> <div>4%</div> <div> <div></div> <div>80%</div> <div>17%</div> </div> </div>
1	C	338	<div> <div>4%</div> <div> <div></div> <div>84%</div> <div>13%</div> </div> </div>
1	D	338	<div> <div>4%</div> <div> <div></div> <div>80%</div> <div>15%</div> </div> </div>

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	NAI	A	1331	-	-	-	X
2	NAI	D	1331	-	-	-	X

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 10931 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 D-LACTATE DEHYDROGENASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	330	Total	C	N	O	S	0	0	0
			2565	1629	438	483	15			
1	B	335	Total	C	N	O	S	0	0	0
			2615	1659	453	488	15			
1	C	335	Total	C	N	O	S	0	0	0
			2615	1659	453	488	15			
1	D	330	Total	C	N	O	S	0	0	0
			2565	1629	438	483	15			

There are 40 discrepancies between the modelled and reference sequences:

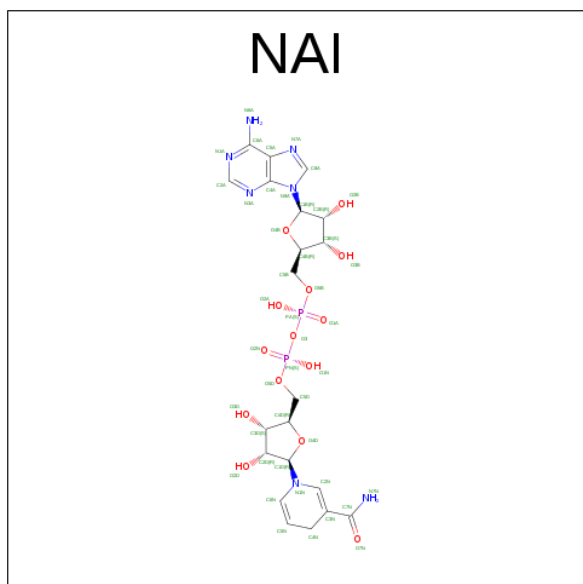
Chain	Residue	Modelled	Actual	Comment	Reference
A	330	LYS	-	EXPRESSION TAG	UNP Q8Z780
A	331	HIS	-	EXPRESSION TAG	UNP Q8Z780
A	332	HIS	-	EXPRESSION TAG	UNP Q8Z780
A	333	HIS	-	EXPRESSION TAG	UNP Q8Z780
A	334	HIS	-	EXPRESSION TAG	UNP Q8Z780
A	335	HIS	-	EXPRESSION TAG	UNP Q8Z780
A	336	HIS	-	EXPRESSION TAG	UNP Q8Z780
A	337	HIS	-	EXPRESSION TAG	UNP Q8Z780
A	338	HIS	-	EXPRESSION TAG	UNP Q8Z780
A	273	VAL	ASN	CONFLICT	UNP Q8Z780
B	330	LYS	-	EXPRESSION TAG	UNP Q8Z780
B	331	HIS	-	EXPRESSION TAG	UNP Q8Z780
B	332	HIS	-	EXPRESSION TAG	UNP Q8Z780
B	333	HIS	-	EXPRESSION TAG	UNP Q8Z780
B	334	HIS	-	EXPRESSION TAG	UNP Q8Z780
B	335	HIS	-	EXPRESSION TAG	UNP Q8Z780
B	336	HIS	-	EXPRESSION TAG	UNP Q8Z780
B	337	HIS	-	EXPRESSION TAG	UNP Q8Z780
B	338	HIS	-	EXPRESSION TAG	UNP Q8Z780
B	273	VAL	ASN	CONFLICT	UNP Q8Z780
C	330	LYS	-	EXPRESSION TAG	UNP Q8Z780

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	331	HIS	-	EXPRESSION TAG	UNP Q8Z780
C	332	HIS	-	EXPRESSION TAG	UNP Q8Z780
C	333	HIS	-	EXPRESSION TAG	UNP Q8Z780
C	334	HIS	-	EXPRESSION TAG	UNP Q8Z780
C	335	HIS	-	EXPRESSION TAG	UNP Q8Z780
C	336	HIS	-	EXPRESSION TAG	UNP Q8Z780
C	337	HIS	-	EXPRESSION TAG	UNP Q8Z780
C	338	HIS	-	EXPRESSION TAG	UNP Q8Z780
C	273	VAL	ASN	CONFLICT	UNP Q8Z780
D	330	LYS	-	EXPRESSION TAG	UNP Q8Z780
D	331	HIS	-	EXPRESSION TAG	UNP Q8Z780
D	332	HIS	-	EXPRESSION TAG	UNP Q8Z780
D	333	HIS	-	EXPRESSION TAG	UNP Q8Z780
D	334	HIS	-	EXPRESSION TAG	UNP Q8Z780
D	335	HIS	-	EXPRESSION TAG	UNP Q8Z780
D	336	HIS	-	EXPRESSION TAG	UNP Q8Z780
D	337	HIS	-	EXPRESSION TAG	UNP Q8Z780
D	338	HIS	-	EXPRESSION TAG	UNP Q8Z780
D	273	VAL	ASN	CONFLICT	UNP Q8Z780

- Molecule 2 is 1,4-DIHYDRONICOTINAMIDE ADENINE DINUCLEOTIDE (three-letter code: NAI) (formula: C<sub>21</sub>H<sub>29</sub>N<sub>7</sub>O<sub>14</sub>P<sub>2</sub>).



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

*Continued on next page...*

*Continued from previous page...*

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

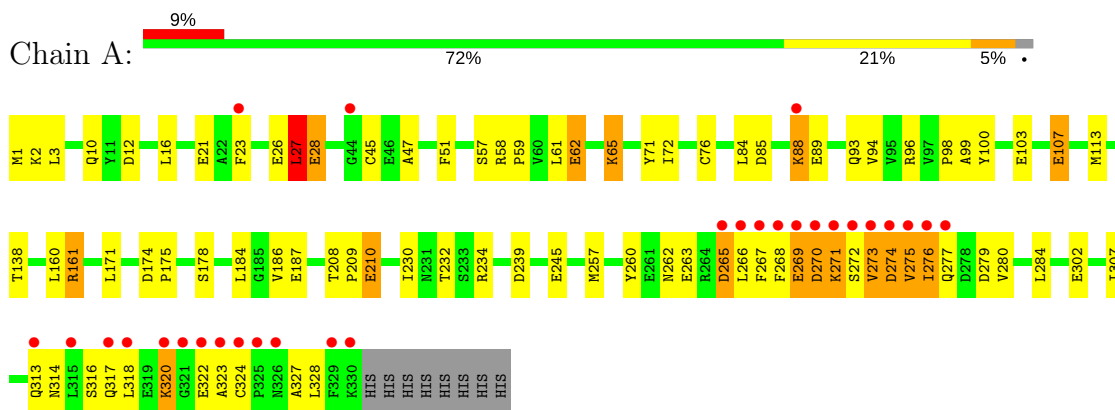
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	104	Total	O	0	0
			104	104		
3	B	125	Total	O	0	0
			125	125		
3	C	148	Total	O	0	0
			148	148		
3	D	106	Total	O	0	0
			106	106		

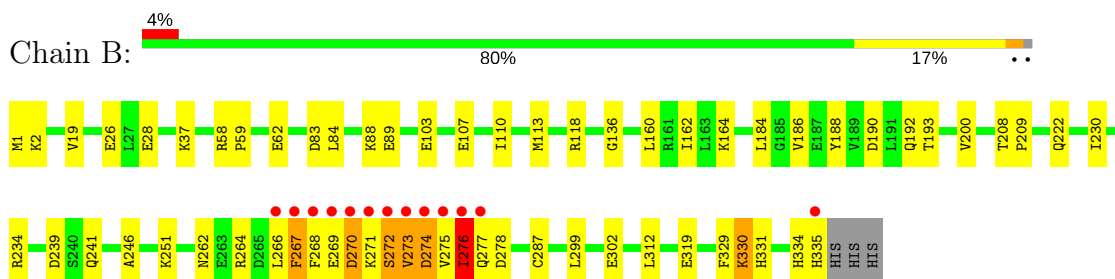
### 3 Residue-property plots [i](#)

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.

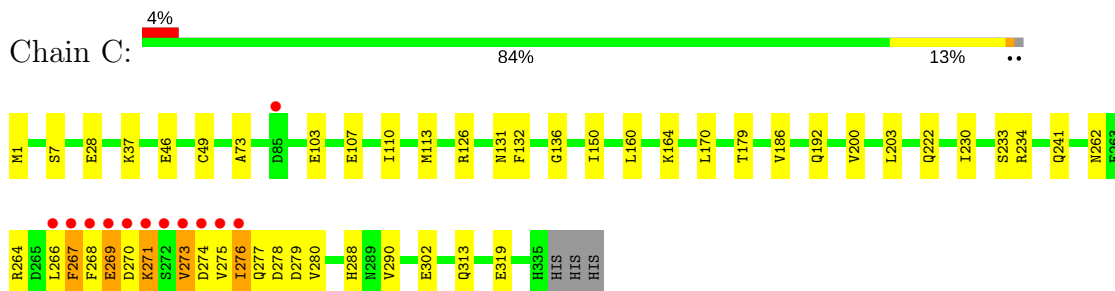
#### • Molecule 1: D-LACTATE DEHYDROGENASE



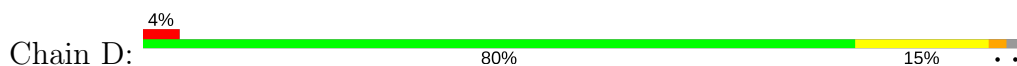
#### • Molecule 1: D-LACTATE DEHYDROGENASE

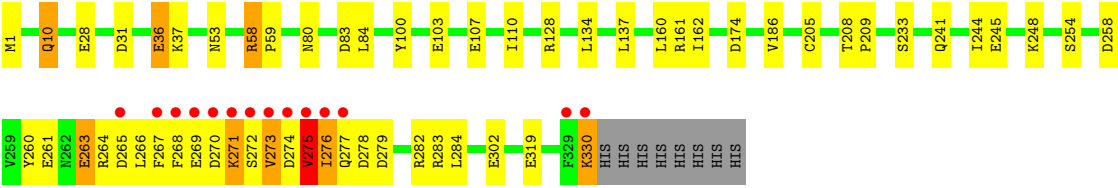


#### • Molecule 1: D-LACTATE DEHYDROGENASE



#### • Molecule 1: D-LACTATE DEHYDROGENASE







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	110.35Å 60.70Å 130.39Å 90.00° 106.92° 90.00°	Depositor
Resolution (Å)	105.57 – 2.18 51.22 – 2.18	Depositor EDS
% Data completeness (in resolution range)	95.8 (105.57-2.18) 95.8 (51.22-2.18)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.01 (at 2.18Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.177 , 0.224 0.197 , 0.233	Depositor DCC
$R_{free}$ test set	4173 reflections (5.29%)	DCC
Wilson B-factor (Å <sup>2</sup> )	22.0	Xtriage
Anisotropy	0.041	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 46.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	10931	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 13.25% 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: 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	1.34	2/2609 (0.1%)	0.89	4/3520 (0.1%)
1	B	1.41	6/2664 (0.2%)	0.87	3/3595 (0.1%)
1	C	1.43	7/2664 (0.3%)	0.87	0/3595
1	D	1.45	9/2609 (0.3%)	0.91	6/3520 (0.2%)
All	All	1.41	24/10546 (0.2%)	0.89	13/14230 (0.1%)

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	1	1
1	D	0	1
All	All	1	2

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	302	GLU	CD-OE1	-7.81	1.17	1.25
1	B	302	GLU	CD-OE2	-7.49	1.17	1.25
1	B	302	GLU	CD-OE1	-7.45	1.17	1.25
1	D	302	GLU	CD-OE2	-7.08	1.17	1.25
1	C	302	GLU	CD-OE1	-6.81	1.18	1.25

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	28	GLU	N-CA-C	7.66	131.68	111.00
1	D	58	ARG	NE-CZ-NH1	6.69	123.65	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	31	ASP	CB-CG-OD1	6.44	124.09	118.30
1	A	27	LEU	C-N-CA	6.42	137.75	121.70
1	D	83	ASP	CB-CG-OD1	6.29	123.96	118.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	28	GLU	CA

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	27	LEU	Peptide
1	D	275	VAL	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	2565	0	2553	108	0
1	B	2615	0	2592	62	0
1	C	2615	0	2591	62	0
1	D	2565	0	2557	69	3
2	A	44	0	27	3	0
2	D	44	0	27	6	0
3	A	104	0	0	2	0
3	B	125	0	0	0	0
3	C	148	0	0	3	0
3	D	106	0	0	1	0
All	All	10931	0	10347	294	3

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 14.

The worst 5 of 294 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:317:GLN:NE2	1:A:324:CYS:CA	1.68	1.29
1:D:265:ASP:O	1:D:266:LEU:HD12	1.29	1.25
1:D:269:GLU:OE1	1:D:283:ARG:NH2	1.78	1.15
1:B:274:ASP:N	1:B:275:VAL:HA	1.55	1.12
1:A:269:GLU:HG2	1:A:280:VAL:HG21	1.35	1.08

All (3) 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:D:1:MET:N	1:D:58:ARG:NH1[2_858]	0.74	1.46
1:D:1:MET:N	1:D:58:ARG:CZ[2_858]	2.07	0.13
1:D:1:MET:CA	1:D:58:ARG:NH1[2_858]	2.11	0.09

## 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	328/338 (97%)	321 (98%)	6 (2%)	1 (0%)	44	47
1	B	333/338 (98%)	324 (97%)	6 (2%)	3 (1%)	20	17
1	C	333/338 (98%)	328 (98%)	3 (1%)	2 (1%)	28	27
1	D	328/338 (97%)	322 (98%)	4 (1%)	2 (1%)	28	27
All	All	1322/1352 (98%)	1295 (98%)	19 (1%)	8 (1%)	28	27

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	274	ASP
1	C	276	ILE
1	D	275	VAL
1	B	267	PHE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	276	ILE

### 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	271/281 (96%)	252 (93%)	19 (7%)	18	17
1	B	276/281 (98%)	267 (97%)	9 (3%)	43	51
1	C	276/281 (98%)	270 (98%)	6 (2%)	57	68
1	D	271/281 (96%)	262 (97%)	9 (3%)	43	51
All	All	1094/1124 (97%)	1051 (96%)	43 (4%)	37	43

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	37	LYS
1	B	274	ASP
1	D	271	LYS
1	B	84	LEU
1	B	264	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	277	GLN
1	D	53	ASN
1	B	334	HIS
1	B	241	GLN
1	B	331	HIS

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

2 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	1331	-	40,48,48	1.20	4 (10%)	41,73,73	2.16	11 (26%)
2	NAI	D	1331	-	40,48,48	1.40	5 (12%)	41,73,73	2.01	10 (24%)

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	1331	-	-	0/25/72/72	0/5/5/5
2	NAI	D	1331	-	-	0/25/72/72	0/5/5/5

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	1331	NAI	C2B-C1B	2.22	1.57	1.53
2	A	1331	NAI	O4B-C1B	2.32	1.44	1.41
2	D	1331	NAI	C2N-C3N	2.41	1.41	1.34
2	A	1331	NAI	C6N-C5N	2.46	1.37	1.33
2	A	1331	NAI	C2N-C3N	2.49	1.42	1.34

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1331	NAI	N3A-C2A-N1A	-7.05	122.72	128.86
2	D	1331	NAI	N3A-C2A-N1A	-6.82	122.91	128.86
2	A	1331	NAI	C1B-N9A-C4A	-4.77	118.40	126.64
2	A	1331	NAI	C4B-O4B-C1B	-4.58	104.89	109.77
2	D	1331	NAI	C4B-O4B-C1B	-4.17	105.33	109.77

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1331	NAI	3	0
2	D	1331	NAI	6	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, 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	330/338 (97%)	0.44	29 (8%)	11 11	14, 29, 75, 137	1 (0%)
1	B	335/338 (99%)	0.03	13 (3%)	40 40	13, 24, 56, 142	1 (0%)
1	C	335/338 (99%)	-0.07	12 (3%)	43 44	9, 21, 57, 151	1 (0%)
1	D	330/338 (97%)	-0.04	14 (4%)	37 37	11, 21, 54, 134	1 (0%)
All	All	1330/1352 (98%)	0.09	68 (5%)	29 30	9, 23, 66, 151	4 (0%)

The worst 5 of 68 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	275	VAL	14.8
1	C	275	VAL	14.5
1	B	274	ASP	13.3
1	B	275	VAL	12.5
1	C	273	VAL	12.4

### 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, 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	NAI	D	1331	44/44	0.79	0.34	9.07	34,49,85,89	0
2	NAI	A	1331	44/44	0.90	0.19	2.47	30,40,63,64	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.