



# wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 02:08 PM GMT

PDB ID : 1O04  
Title : Cys302Ser mutant of human mitochondrial aldehyde dehydrogenase complexed with NAD<sup>+</sup> and Mg<sup>2+</sup>  
Authors : Perez-Miller, S.J.; Hurley, T.D.  
Deposited on : 2003-02-20  
Resolution : 1.42 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

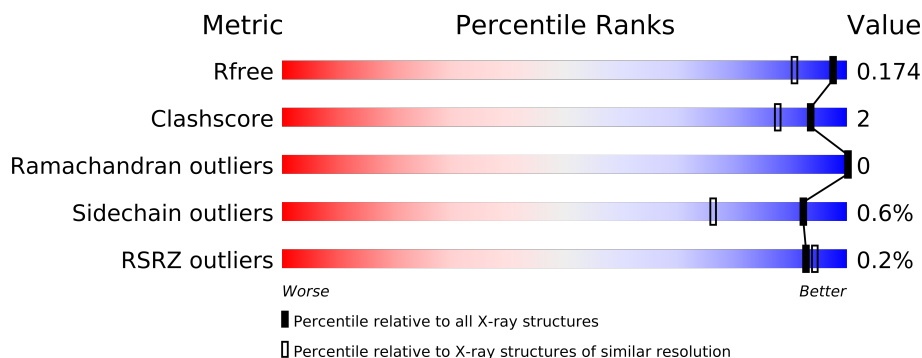
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 1.42 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))
$R_{free}$	66092	1110 (1.44-1.40)
Clashscore	79885	1263 (1.44-1.40)
Ramachandran outliers	78287	1226 (1.44-1.40)
Sidechain outliers	78261	1225 (1.44-1.40)
RSRZ outliers	66119	1110 (1.44-1.40)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	500	
1	B	500	
1	C	500	
1	D	500	
1	E	500	
1	F	500	
1	G	500	
1	H	500	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	NA	E	6605	-	X

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Geometry	Electron density
3	MG	A	6701	-	X
3	MG	B	6702	-	X
3	MG	C	6703	-	X
3	MG	D	6704	-	X
3	MG	E	6705	-	X
3	MG	F	6706	-	X
3	MG	G	6707	-	X
3	MG	H	6708	-	X
4	NAD	A	6501	-	X
4	NAD	B	6502	-	X
4	NAD	C	6503	-	X
4	NAD	D	6504	-	X
4	NAD	E	6505	-	X
4	NAD	F	6506	-	X
4	NAD	G	6507	-	X
4	NAD	H	6508	-	X
5	GAI	A	6801	-	X
5	GAI	A	6811	-	X
5	GAI	B	6802	-	X
5	GAI	B	6812	-	X
5	GAI	C	6813	-	X
5	GAI	C	6823	-	X
5	GAI	D	6804	-	X
5	GAI	D	6814	-	X
5	GAI	E	6805	-	X
5	GAI	E	6815	-	X
5	GAI	F	6806	-	X
5	GAI	F	6816	-	X
5	GAI	F	6826	-	X
5	GAI	G	6817	-	X
5	GAI	H	6808	-	X
5	GAI	H	6818	-	X
6	EDO	A	6921	-	X
6	EDO	A	6941	-	X
6	EDO	B	6922	-	X
6	EDO	B	6942	-	X
6	EDO	C	6923	-	X
6	EDO	C	6943	-	X
6	EDO	D	6914	-	X
6	EDO	D	6944	-	X
6	EDO	E	6925	-	X
6	EDO	F	6946	-	X

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Geometry	Electron density
6	EDO	F	6956	-	X
6	EDO	G	6917	-	X
6	EDO	H	6918	-	X
6	EDO	H	6928	-	X
6	EDO	H	6948	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 36161 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Aldehyde dehydrogenase, mitochondrial precursor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	494	Total	C	N	O	S	0	3	0
			3821	2428	652	722	19			
1	B	494	Total	C	N	O	S	0	3	0
			3821	2428	652	722	19			
1	C	494	Total	C	N	O	S	0	3	0
			3821	2428	652	722	19			
1	D	494	Total	C	N	O	S	0	3	0
			3821	2428	652	722	19			
1	E	494	Total	C	N	O	S	0	3	0
			3821	2428	652	722	19			
1	F	494	Total	C	N	O	S	0	3	0
			3821	2428	652	722	19			
1	G	494	Total	C	N	O	S	0	3	0
			3821	2428	652	722	19			
1	H	494	Total	C	N	O	S	0	3	0
			3821	2428	652	722	19			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	302	SER	CYS	ENGINEERED	UNP P05091
B	302	SER	CYS	ENGINEERED	UNP P05091
C	302	SER	CYS	ENGINEERED	UNP P05091
D	302	SER	CYS	ENGINEERED	UNP P05091
E	302	SER	CYS	ENGINEERED	UNP P05091
F	302	SER	CYS	ENGINEERED	UNP P05091
G	302	SER	CYS	ENGINEERED	UNP P05091
H	302	SER	CYS	ENGINEERED	UNP P05091

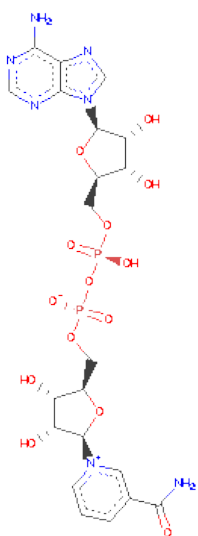
- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	G	1	Total Na 1 1	0	0
2	D	1	Total Na 1 1	0	0
2	E	1	Total Na 1 1	0	0
2	H	1	Total Na 1 1	0	0
2	B	1	Total Na 1 1	0	0
2	C	1	Total Na 1 1	0	0
2	A	1	Total Na 1 1	0	0
2	F	1	Total Na 1 1	0	0

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

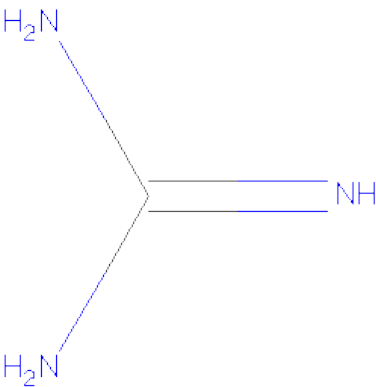
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	G	1	Total Mg 1 1	0	0
3	D	1	Total Mg 1 1	0	0
3	E	1	Total Mg 1 1	0	0
3	H	1	Total Mg 1 1	0	0
3	B	1	Total Mg 1 1	0	0
3	C	1	Total Mg 1 1	0	0
3	A	1	Total Mg 1 1	0	0
3	F	1	Total Mg 1 1	0	0

- Molecule 4 is NICOTINAMIDE-ADENINE-DINUCLEOTIDE (three-letter code: NAD) (formula: C<sub>21</sub>H<sub>27</sub>N<sub>7</sub>O<sub>14</sub>P<sub>2</sub>).



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

- Molecule 5 is GUANIDINE (three-letter code: GAI) (formula:  $\text{CH}_5\text{N}_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	N	0	0
			4	1	3		
5	A	1	Total	C	N	0	0
			4	1	3		
5	B	1	Total	C	N	0	0
			4	1	3		
5	B	1	Total	C	N	0	0
			4	1	3		
5	C	1	Total	C	N	0	0
			4	1	3		
5	C	1	Total	C	N	0	0
			4	1	3		
5	C	1	Total	C	N	0	0
			4	1	3		
5	D	1	Total	C	N	0	0
			4	1	3		
5	D	1	Total	C	N	0	0
			4	1	3		
5	E	1	Total	C	N	0	0
			4	1	3		
5	E	1	Total	C	N	0	0
			4	1	3		
5	F	1	Total	C	N	0	0
			4	1	3		
5	F	1	Total	C	N	0	0
			4	1	3		
5	F	1	Total	C	N	0	0
			4	1	3		

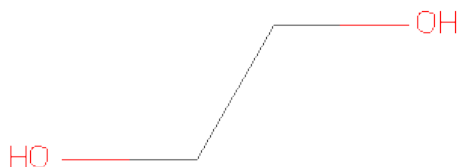
Continued on next page...



*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	G	1	Total	C	N	0	0
			4	1	3		
5	G	1	Total	C	N	0	0
			4	1	3		
5	H	1	Total	C	N	0	0
			4	1	3		
5	H	1	Total	C	N	0	0
			4	1	3		

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



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

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total 4	C 2	O 2	0	0
6	C	1	Total 4	C 2	O 2	0	0
6	C	1	Total 4	C 2	O 2	0	0
6	C	1	Total 4	C 2	O 2	0	0
6	C	1	Total 4	C 2	O 2	0	0
6	C	1	Total 4	C 2	O 2	0	0
6	D	1	Total 4	C 2	O 2	0	0
6	D	1	Total 4	C 2	O 2	0	0
6	D	1	Total 4	C 2	O 2	0	0
6	E	1	Total 4	C 2	O 2	0	0
6	E	1	Total 4	C 2	O 2	0	0
6	E	1	Total 4	C 2	O 2	0	0
6	F	1	Total 4	C 2	O 2	0	0
6	F	1	Total 4	C 2	O 2	0	0
6	F	1	Total 4	C 2	O 2	0	0
6	F	1	Total 4	C 2	O 2	0	0
6	F	1	Total 4	C 2	O 2	0	0
6	F	1	Total 4	C 2	O 2	0	0
6	G	1	Total 4	C 2	O 2	0	0
6	G	1	Total 4	C 2	O 2	0	0
6	G	1	Total 4	C 2	O 2	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	H	1	Total	C	O	0	0
			4	2	2		
6	H	1	Total	C	O	0	0
			4	2	2		
6	H	1	Total	C	O	0	0
			4	2	2		
6	H	1	Total	C	O	0	0
			4	2	2		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	602	Total	O	0	0
			602	602		
7	B	645	Total	O	0	0
			645	645		
7	C	627	Total	O	0	0
			627	627		
7	D	598	Total	O	0	0
			598	598		
7	E	665	Total	O	0	0
			665	665		
7	F	632	Total	O	0	0
			632	632		
7	G	628	Total	O	0	0
			628	628		
7	H	628	Total	O	0	0
			628	628		

### 3 Residue-property plots

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.

- Molecule 1: Aldehyde dehydrogenase, mitochondrial precursor

Chain A: 



- Molecule 1: Aldehyde dehydrogenase, mitochondrial precursor

Chain B: 



- Molecule 1: Aldehyde dehydrogenase, mitochondrial precursor

Chain C: 



- Molecule 1: Aldehyde dehydrogenase, mitochondrial precursor

Chain D: 



- Molecule 1: Aldehyde dehydrogenase, mitochondrial precursor

Chain E: 

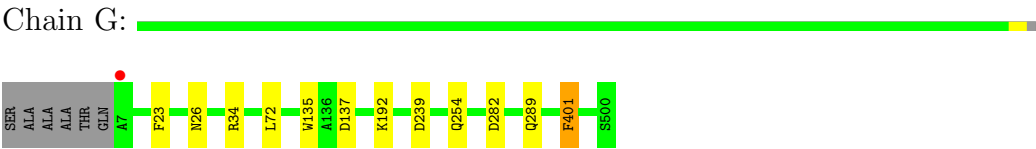


- Molecule 1: Aldehyde dehydrogenase, mitochondrial precursor

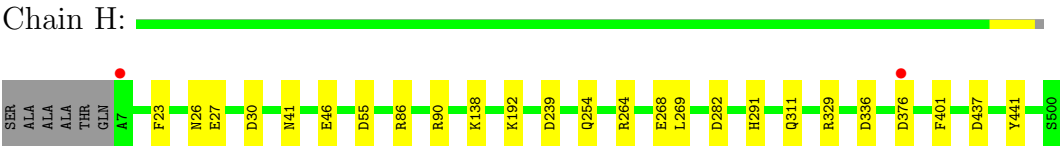
Chain F: 



- Molecule 1: Aldehyde dehydrogenase, mitochondrial precursor



- Molecule 1: Aldehyde dehydrogenase, mitochondrial precursor



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	141.21Å 152.49Å 177.20Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 1.42 45.49 – 1.42	Depositor EDS
% Data completeness (in resolution range)	95.9 (50.00-1.42) 95.9 (45.49-1.42)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.89 (at 1.42Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, $R_{free}$	0.145 , 0.171 0.148 , 0.174	Depositor DCC
$R_{free}$ test set	34234 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	10.1	Xtriage
Anisotropy	0.280	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 37.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 682783 reflections	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	36161	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	12.0	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 38.51 % 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 3.6455e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NA, MG, EDO, NAD, GAI

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	0.64	0/3905	0.80	7/5297 (0.1%)
1	B	0.67	0/3905	0.80	5/5297 (0.1%)
1	C	0.67	0/3905	0.80	7/5297 (0.1%)
1	D	0.67	0/3905	0.79	6/5297 (0.1%)
1	E	0.68	1/3905 (0.0%)	0.82	6/5297 (0.1%)
1	F	0.67	0/3905	0.81	5/5297 (0.1%)
1	G	0.65	0/3905	0.79	3/5297 (0.1%)
1	H	0.67	0/3905	0.82	9/5297 (0.2%)
All	All	0.67	1/31240 (0.0%)	0.80	48/42376 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	363	GLU	CD-OE1	5.24	1.31	1.25

The worst 5 of 48 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	H	336	ASP	CB-CG-OD2	9.55	126.90	118.30
1	C	282	ASP	CB-CG-OD2	9.49	126.84	118.30
1	A	282	ASP	CB-CG-OD2	9.26	126.63	118.30
1	B	336	ASP	CB-CG-OD2	9.17	126.55	118.30
1	E	282	ASP	CB-CG-OD2	9.10	126.49	118.30

There are no chirality outliers.

There are no planarity outliers.

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3821	0	3764	13	0
1	B	3821	0	3764	15	0
1	C	3821	0	3764	12	0
1	D	3821	0	3764	10	0
1	E	3821	0	3764	7	0
1	F	3821	0	3764	16	0
1	G	3821	0	3764	8	0
1	H	3821	0	3764	19	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
3	G	1	0	0	0	0
3	H	1	0	0	0	0
4	A	44	0	26	2	0
4	B	44	0	26	2	0
4	C	44	0	26	1	0
4	D	44	0	26	0	0
4	E	44	0	26	1	0
4	F	44	0	26	4	0
4	G	44	0	26	1	0
4	H	44	0	26	2	0
5	A	8	0	8	0	0
5	B	8	0	8	0	0
5	C	12	0	12	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	D	8	0	8	0	0
5	E	8	0	8	0	0
5	F	12	0	12	0	0
5	G	8	0	8	0	0
5	H	8	0	8	0	0
6	A	16	0	24	5	0
6	B	16	0	24	4	0
6	C	20	0	30	5	0
6	D	12	0	18	2	0
6	E	12	0	18	0	0
6	F	24	0	36	3	0
6	G	12	0	18	0	0
6	H	16	0	24	11	0
7	A	602	0	0	4	0
7	B	645	0	0	3	0
7	C	627	0	0	2	0
7	D	598	0	0	2	0
7	E	665	0	0	2	0
7	F	632	0	0	5	0
7	G	628	0	0	3	0
7	H	628	0	0	8	0
All	All	36161	0	30584	100	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 2.

The worst 5 of 100 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
7:G:4087:HOH:O	1:H:254:GLN:HG3	1.38	1.21
7:E:3898:HOH:O	1:F:254:GLN:HG3	1.36	1.21
1:E:254:GLN:HG3	7:F:3897:HOH:O	1.38	1.20
1:G:254:GLN:HG3	7:H:4086:HOH:O	1.45	1.16
1:H:46:GLU:HB2	6:H:6918:EDO:H21	1.26	1.13

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	495/500 (99%)	483 (98%)	12 (2%)	0	100	100
1	B	495/500 (99%)	483 (98%)	12 (2%)	0	100	100
1	C	495/500 (99%)	482 (97%)	13 (3%)	0	100	100
1	D	495/500 (99%)	482 (97%)	13 (3%)	0	100	100
1	E	495/500 (99%)	483 (98%)	12 (2%)	0	100	100
1	F	495/500 (99%)	482 (97%)	13 (3%)	0	100	100
1	G	495/500 (99%)	482 (97%)	13 (3%)	0	100	100
1	H	495/500 (99%)	483 (98%)	12 (2%)	0	100	100
All	All	3960/4000 (99%)	3860 (98%)	100 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	402/402 (100%)	399 (99%)	3 (1%)	91	73
1	B	402/402 (100%)	397 (99%)	5 (1%)	82	56
1	C	402/402 (100%)	400 (100%)	2 (0%)	94	80
1	D	402/402 (100%)	400 (100%)	2 (0%)	94	80
1	E	402/402 (100%)	400 (100%)	2 (0%)	94	80
1	F	402/402 (100%)	400 (100%)	2 (0%)	94	80
1	G	402/402 (100%)	400 (100%)	2 (0%)	94	80
1	H	402/402 (100%)	400 (100%)	2 (0%)	94	80

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	3216/3216 (100%)	3196 (99%)	20 (1%)	92	76

5 of 20 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	401	PHE
1	D	192	LYS
1	G	192	LYS
1	B	422	ASN
1	C	192	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	291	HIS
1	C	358	ASN
1	G	254	GLN
1	C	29	HIS
1	D	444	GLN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

Of 74 ligands modelled in this entry, 16 are monoatomic - leaving 58 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
4	NAD	A	6501	3	48,48,48	1.33	4 (8%)	73,73,73	1.53	7 (9%)
5	GAI	A	6801	-	3,3,3	0.90	0	3,3,3	1.18	0
5	GAI	A	6811	-	3,3,3	1.76	1 (33%)	3,3,3	0.94	0
6	EDO	A	6901	-	3,3,3	0.65	0	2,2,2	0.22	0
6	EDO	A	6911	-	3,3,3	0.48	0	2,2,2	0.26	0
6	EDO	A	6921	-	3,3,3	0.43	0	2,2,2	0.77	0
6	EDO	A	6941	-	3,3,3	0.51	0	2,2,2	0.50	0
4	NAD	B	6502	3	48,48,48	1.26	2 (4%)	73,73,73	1.51	8 (10%)
5	GAI	B	6802	-	3,3,3	1.06	0	3,3,3	0.42	0
5	GAI	B	6812	-	3,3,3	1.94	2 (66%)	3,3,3	1.15	0
6	EDO	B	6902	-	3,3,3	0.51	0	2,2,2	0.10	0
6	EDO	B	6912	-	3,3,3	0.52	0	2,2,2	0.23	0
6	EDO	B	6922	-	3,3,3	0.47	0	2,2,2	0.21	0
6	EDO	B	6942	-	3,3,3	0.84	0	2,2,2	0.41	0
4	NAD	C	6503	3	48,48,48	1.26	3 (6%)	73,73,73	1.57	9 (12%)
5	GAI	C	6803	-	3,3,3	1.33	0	3,3,3	0.61	0
5	GAI	C	6813	-	3,3,3	1.64	1 (33%)	3,3,3	1.43	1 (33%)
5	GAI	C	6823	-	3,3,3	1.50	0	3,3,3	1.21	0
6	EDO	C	6903	-	3,3,3	0.51	0	2,2,2	0.64	0
6	EDO	C	6913	-	3,3,3	0.49	0	2,2,2	0.34	0
6	EDO	C	6923	-	3,3,3	0.37	0	2,2,2	0.63	0
6	EDO	C	6943	-	3,3,3	0.45	0	2,2,2	0.13	0
6	EDO	C	6963	-	3,3,3	0.42	0	2,2,2	0.72	0
4	NAD	D	6504	3	48,48,48	1.33	3 (6%)	73,73,73	1.68	8 (10%)
5	GAI	D	6804	-	3,3,3	0.93	0	3,3,3	1.55	1 (33%)
5	GAI	D	6814	-	3,3,3	1.98	2 (66%)	3,3,3	0.94	0
6	EDO	D	6904	-	3,3,3	0.55	0	2,2,2	0.32	0
6	EDO	D	6914	-	3,3,3	0.67	0	2,2,2	0.53	0
6	EDO	D	6944	-	3,3,3	0.40	0	2,2,2	1.01	0
4	NAD	E	6505	3	48,48,48	1.26	3 (6%)	73,73,73	1.49	9 (12%)
5	GAI	E	6805	-	3,3,3	0.76	0	3,3,3	1.18	0
5	GAI	E	6815	-	3,3,3	2.06	2 (66%)	3,3,3	1.02	0
6	EDO	E	6905	-	3,3,3	0.59	0	2,2,2	0.30	0
6	EDO	E	6915	-	3,3,3	0.50	0	2,2,2	0.54	0
6	EDO	E	6925	-	3,3,3	0.29	0	2,2,2	1.10	0
4	NAD	F	6506	3	48,48,48	1.37	4 (8%)	73,73,73	1.49	8 (10%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	GAI	F	6806	-	3,3,3	1.16	0	3,3,3	0.79	0
5	GAI	F	6816	-	3,3,3	1.78	1 (33%)	3,3,3	0.86	0
5	GAI	F	6826	-	3,3,3	1.36	0	3,3,3	1.11	0
6	EDO	F	6906	-	3,3,3	0.63	0	2,2,2	0.36	0
6	EDO	F	6916	-	3,3,3	0.52	0	2,2,2	0.37	0
6	EDO	F	6926	-	3,3,3	0.43	0	2,2,2	1.05	0
6	EDO	F	6946	-	3,3,3	0.54	0	2,2,2	0.43	0
6	EDO	F	6956	-	3,3,3	0.57	0	2,2,2	0.42	0
6	EDO	F	6966	-	3,3,3	0.51	0	2,2,2	0.38	0
4	NAD	G	6507	3	48,48,48	1.28	3 (6%)	73,73,73	1.58	9 (12%)
5	GAI	G	6807	-	3,3,3	0.96	0	3,3,3	0.33	0
5	GAI	G	6817	-	3,3,3	1.55	0	3,3,3	1.53	1 (33%)
6	EDO	G	6907	-	3,3,3	0.52	0	2,2,2	0.34	0
6	EDO	G	6917	-	3,3,3	0.47	0	2,2,2	0.26	0
6	EDO	G	6927	-	3,3,3	0.44	0	2,2,2	0.64	0
4	NAD	H	6508	3	48,48,48	1.27	4 (8%)	73,73,73	1.71	8 (10%)
5	GAI	H	6808	-	3,3,3	0.57	0	3,3,3	1.62	1 (33%)
5	GAI	H	6818	-	3,3,3	1.61	1 (33%)	3,3,3	0.86	0
6	EDO	H	6908	-	3,3,3	0.66	0	2,2,2	0.26	0
6	EDO	H	6918	-	3,3,3	0.42	0	2,2,2	0.45	0
6	EDO	H	6928	-	3,3,3	0.39	0	2,2,2	0.48	0
6	EDO	H	6948	-	3,3,3	0.51	0	2,2,2	0.54	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
4	NAD	A	6501	3	-	0/30/62/62	0/3/5/5
5	GAI	A	6801	-	-	0/0/0/0	0/0/0/0
5	GAI	A	6811	-	-	0/0/0/0	0/0/0/0
6	EDO	A	6901	-	-	0/1/1/1	0/0/0/0
6	EDO	A	6911	-	-	0/1/1/1	0/0/0/0
6	EDO	A	6921	-	-	0/1/1/1	0/0/0/0
6	EDO	A	6941	-	-	0/1/1/1	0/0/0/0
4	NAD	B	6502	3	-	0/30/62/62	0/3/5/5
5	GAI	B	6802	-	-	0/0/0/0	0/0/0/0
5	GAI	B	6812	-	-	0/0/0/0	0/0/0/0
6	EDO	B	6902	-	-	0/1/1/1	0/0/0/0
6	EDO	B	6912	-	-	0/1/1/1	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	EDO	B	6922	-	-	0/1/1/1	0/0/0/0
6	EDO	B	6942	-	-	0/1/1/1	0/0/0/0
4	NAD	C	6503	3	-	0/30/62/62	0/3/5/5
5	GAI	C	6803	-	-	0/0/0/0	0/0/0/0
5	GAI	C	6813	-	-	0/0/0/0	0/0/0/0
5	GAI	C	6823	-	-	0/0/0/0	0/0/0/0
6	EDO	C	6903	-	-	0/1/1/1	0/0/0/0
6	EDO	C	6913	-	-	0/1/1/1	0/0/0/0
6	EDO	C	6923	-	-	0/1/1/1	0/0/0/0
6	EDO	C	6943	-	-	0/1/1/1	0/0/0/0
6	EDO	C	6963	-	-	0/1/1/1	0/0/0/0
4	NAD	D	6504	3	-	0/30/62/62	0/3/5/5
5	GAI	D	6804	-	-	0/0/0/0	0/0/0/0
5	GAI	D	6814	-	-	0/0/0/0	0/0/0/0
6	EDO	D	6904	-	-	0/1/1/1	0/0/0/0
6	EDO	D	6914	-	-	0/1/1/1	0/0/0/0
6	EDO	D	6944	-	-	0/1/1/1	0/0/0/0
4	NAD	E	6505	3	-	0/30/62/62	0/3/5/5
5	GAI	E	6805	-	-	0/0/0/0	0/0/0/0
5	GAI	E	6815	-	-	0/0/0/0	0/0/0/0
6	EDO	E	6905	-	-	0/1/1/1	0/0/0/0
6	EDO	E	6915	-	-	0/1/1/1	0/0/0/0
6	EDO	E	6925	-	-	0/1/1/1	0/0/0/0
4	NAD	F	6506	3	-	0/30/62/62	0/3/5/5
5	GAI	F	6806	-	-	0/0/0/0	0/0/0/0
5	GAI	F	6816	-	-	0/0/0/0	0/0/0/0
5	GAI	F	6826	-	-	0/0/0/0	0/0/0/0
6	EDO	F	6906	-	-	0/1/1/1	0/0/0/0
6	EDO	F	6916	-	-	0/1/1/1	0/0/0/0
6	EDO	F	6926	-	-	0/1/1/1	0/0/0/0
6	EDO	F	6946	-	-	0/1/1/1	0/0/0/0
6	EDO	F	6956	-	-	0/1/1/1	0/0/0/0
6	EDO	F	6966	-	-	0/1/1/1	0/0/0/0
4	NAD	G	6507	3	-	0/30/62/62	0/3/5/5
5	GAI	G	6807	-	-	0/0/0/0	0/0/0/0
5	GAI	G	6817	-	-	0/0/0/0	0/0/0/0
6	EDO	G	6907	-	-	0/1/1/1	0/0/0/0
6	EDO	G	6917	-	-	0/1/1/1	0/0/0/0
6	EDO	G	6927	-	-	0/1/1/1	0/0/0/0
4	NAD	H	6508	3	-	0/30/62/62	0/3/5/5
5	GAI	H	6808	-	-	0/0/0/0	0/0/0/0
5	GAI	H	6818	-	-	0/0/0/0	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	EDO	H	6908	-	-	0/1/1/1	0/0/0/0
6	EDO	H	6918	-	-	0/1/1/1	0/0/0/0
6	EDO	H	6928	-	-	0/1/1/1	0/0/0/0
6	EDO	H	6948	-	-	0/1/1/1	0/0/0/0

The worst 5 of 36 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	6504	NAD	O7N-C7N	7.01	1.40	1.24
4	F	6506	NAD	O7N-C7N	6.91	1.40	1.24
4	A	6501	NAD	O7N-C7N	6.80	1.40	1.24
4	B	6502	NAD	O7N-C7N	6.56	1.39	1.24
4	G	6507	NAD	O7N-C7N	6.49	1.39	1.24

The worst 5 of 70 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	6504	NAD	N3A-C2A-N1A	-10.34	120.06	128.71
4	H	6508	NAD	N3A-C2A-N1A	-10.06	120.30	128.71
4	G	6507	NAD	N3A-C2A-N1A	-8.54	121.57	128.71
4	C	6503	NAD	N3A-C2A-N1A	-8.54	121.57	128.71
4	A	6501	NAD	N3A-C2A-N1A	-7.94	122.07	128.71

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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<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	494/500 (98%)	-0.59	1 (0%) 93 94	6, 11, 21, 35	0
1	B	494/500 (98%)	-0.63	1 (0%) 93 94	6, 10, 17, 32	0
1	C	494/500 (98%)	-0.65	1 (0%) 93 94	6, 9, 17, 30	0
1	D	494/500 (98%)	-0.63	1 (0%) 93 94	6, 10, 19, 34	0
1	E	494/500 (98%)	-0.69	1 (0%) 93 94	6, 9, 16, 38	0
1	F	494/500 (98%)	-0.69	1 (0%) 93 94	5, 9, 16, 26	0
1	G	494/500 (98%)	-0.57	1 (0%) 93 94	5, 10, 20, 40	0
1	H	494/500 (98%)	-0.61	2 (0%) 90 92	5, 10, 19, 37	0
All	All	3952/4000 (98%)	-0.63	9 (0%) 93 94	5, 10, 18, 40	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	7	ALA	6.7
1	H	7	ALA	4.7
1	A	7	ALA	4.2
1	B	7	ALA	3.8
1	D	7	ALA	3.5

### 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 carbohydrates in this entry.



## 6.4 Ligands

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
3	MG	F	6706	1/1	0.23	41.00	18,18,18,18	0
3	MG	C	6703	1/1	0.15	28.33	18,18,18,18	0
5	GAI	C	6813	4/4	0.17	23.51	20,21,22,23	0
3	MG	A	6701	1/1	0.24	20.06	18,18,18,18	0
5	GAI	B	6812	4/4	0.15	16.68	20,21,21,23	0
6	EDO	B	6922	4/4	0.17	16.68	15,27,29,30	0
6	EDO	C	6943	4/4	0.24	14.85	12,24,26,36	0
5	GAI	G	6817	4/4	0.13	13.25	23,23,25,25	0
5	GAI	D	6804	4/4	0.10	10.86	13,15,16,19	0
3	MG	B	6702	1/1	0.19	10.25	16,16,16,16	0
6	EDO	G	6917	4/4	0.15	9.08	17,22,23,26	0
3	MG	H	6708	1/1	0.19	8.62	17,17,17,17	0
3	MG	G	6707	1/1	0.17	8.52	15,15,15,15	0
6	EDO	H	6948	4/4	0.17	8.44	10,23,23,31	0
6	EDO	B	6942	4/4	0.17	8.20	11,27,27,31	0
5	GAI	E	6815	4/4	0.09	8.11	18,20,20,22	0
5	GAI	D	6814	4/4	0.13	7.26	21,22,23,25	0
5	GAI	B	6802	4/4	0.10	7.05	13,14,14,15	0
6	EDO	A	6921	4/4	0.13	7.03	26,28,31,34	0
3	MG	D	6704	1/1	0.18	6.87	15,15,15,15	0
3	MG	E	6705	1/1	0.19	6.76	15,15,15,15	0
4	NAD	A	6501	44/44	0.13	6.48	12,16,19,21	0
6	EDO	H	6918	4/4	0.19	6.28	13,14,30,31	0
5	GAI	E	6805	4/4	0.12	6.20	12,17,19,19	0
6	EDO	A	6941	4/4	0.10	6.07	11,20,24,26	0
6	EDO	E	6925	4/4	0.15	5.25	22,25,28,32	0
5	GAI	F	6826	4/4	0.17	5.14	27,27,30,31	0
4	NAD	E	6505	44/44	0.12	5.07	10,13,16,17	0
4	NAD	B	6502	44/44	0.12	4.84	11,15,18,19	0
6	EDO	F	6946	4/4	0.14	4.77	11,24,25,29	0
6	EDO	H	6928	4/4	0.09	4.75	20,24,25,26	0
4	NAD	G	6507	44/44	0.11	4.52	10,14,16,18	0
5	GAI	H	6818	4/4	0.10	4.50	21,22,22,23	0
4	NAD	D	6504	44/44	0.13	4.47	12,15,19,20	0

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NAD	H	6508	44/44	0.13	4.30	10,15,17,18	0
4	NAD	C	6503	44/44	0.11	4.26	11,14,17,19	0
5	GAI	A	6811	4/4	0.11	3.98	22,22,22,24	0
6	EDO	C	6923	4/4	0.09	3.96	16,18,21,26	0
4	NAD	F	6506	44/44	0.10	3.86	10,13,17,19	0
5	GAI	F	6816	4/4	0.10	3.85	19,19,23,24	0
5	GAI	H	6808	4/4	0.10	3.82	12,15,15,17	0
6	EDO	F	6956	4/4	0.09	3.53	13,14,18,19	0
5	GAI	A	6801	4/4	0.09	3.01	12,14,14,15	0
5	GAI	F	6806	4/4	0.09	2.97	11,12,12,13	0
5	GAI	C	6823	4/4	0.13	2.46	22,24,27,28	0
6	EDO	D	6944	4/4	0.10	2.42	24,24,26,28	0
2	NA	E	6605	1/1	0.06	2.26	10,10,10,10	0
6	EDO	D	6914	4/4	0.10	2.23	17,17,20,21	0
5	GAI	G	6807	4/4	0.09	1.98	11,14,14,16	0
6	EDO	F	6906	4/4	0.07	1.14	12,13,13,18	0
6	EDO	F	6966	4/4	0.09	0.96	20,22,23,24	0
6	EDO	B	6902	4/4	0.07	0.85	13,14,17,17	0
6	EDO	H	6908	4/4	0.06	0.82	15,15,16,20	0
6	EDO	F	6926	4/4	0.07	0.80	12,15,17,22	0
6	EDO	G	6927	4/4	0.15	0.73	26,27,28,31	0
6	EDO	E	6915	4/4	0.10	0.60	17,20,23,29	0
6	EDO	C	6903	4/4	0.06	0.47	13,17,17,20	0
6	EDO	C	6913	4/4	0.07	0.43	14,18,20,24	0
6	EDO	C	6963	4/4	0.08	0.37	20,23,27,29	0
5	GAI	C	6803	4/4	0.07	0.34	11,12,13,14	0
6	EDO	A	6901	4/4	0.06	0.03	15,15,16,19	0
6	EDO	B	6912	4/4	0.07	-0.09	16,21,23,24	0
6	EDO	D	6904	4/4	0.05	-0.31	15,17,18,19	0
6	EDO	G	6907	4/4	0.05	-0.64	13,15,16,17	0
6	EDO	F	6916	4/4	0.05	-0.66	14,17,19,22	0
6	EDO	A	6911	4/4	0.07	-0.68	23,24,24,30	0
2	NA	H	6608	1/1	0.05	-0.75	11,11,11,11	0
6	EDO	E	6905	4/4	0.05	-0.77	13,18,19,19	0
2	NA	D	6604	1/1	0.04	-1.02	11,11,11,11	0
2	NA	A	6601	1/1	0.05	-1.33	14,14,14,14	0
2	NA	B	6602	1/1	0.03	-1.88	10,10,10,10	0
2	NA	G	6607	1/1	0.04	-2.51	12,12,12,12	0
2	NA	C	6603	1/1	0.02	-2.84	10,10,10,10	0
2	NA	F	6606	1/1	0.02	-4.01	11,11,11,11	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.