wwPDB X-ray Structure Validation Summary Report (i)

Sep 5, 2023 – 01:22 PM EDT

PDB ID	:	3U4J
Title	:	Crystal structure of NAD-dependent aldehyde dehydrogenase from Sinorhizobium meliloti
Authors	:	Eswaramoorthy, S.; Chamala, S.; Evans, B.; Foti, R.; Gizzi, A.; Hillerich, B.; Kar, A.; LaFleur, J.; Seidel, R.; Villigas, G.; Zencheck, W.; Almo, S.C.; Swaminathan, S.; New York Structural Genomics Research Consortium (NYSGRC)
Deposited on	:	2011-10-08
Resolution	:	2.00 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

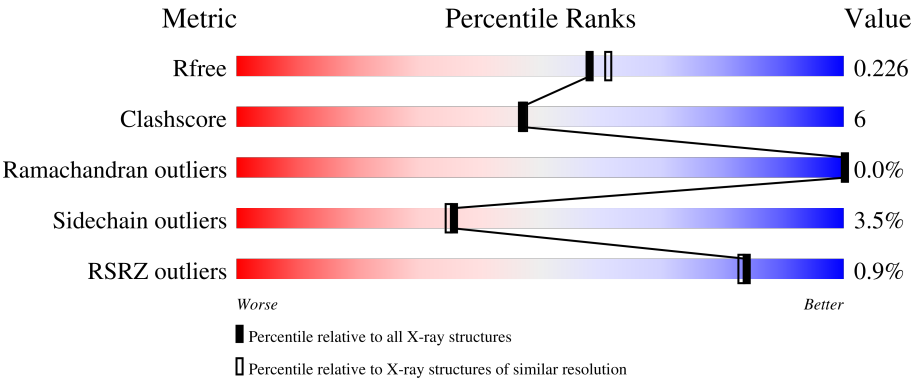
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)

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 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	528	81% 13%
1	B	528	2% 80% 13% 5%
1	C	528	% 84% 11%
1	D	528	82% 13% 5%

Validation Pipeline (wwPDB-VP) : 2.35

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 15736 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 NAD-dependent aldehyde dehydrogenase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	505	Total	C	N	O	S	Se	0	0	0
			3785	2388	667	717	5	8			
1	B	502	Total	C	N	O	S	Se	0	0	0
			3765	2375	664	714	5	7			
1	C	505	Total	C	N	O	S	Se	0	0	0
			3791	2391	670	717	5	8			
1	D	503	Total	C	N	O	S	Se	0	0	0
			3776	2381	668	715	5	7			

There are 96 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MSE	-	expression tag	UNP Q930S8
A	0	VAL	-	expression tag	UNP Q930S8
A	505	ALA	-	expression tag	UNP Q930S8
A	506	GLU	-	expression tag	UNP Q930S8
A	507	ASN	-	expression tag	UNP Q930S8
A	508	LEU	-	expression tag	UNP Q930S8
A	509	TYR	-	expression tag	UNP Q930S8
A	510	PHE	-	expression tag	UNP Q930S8
A	511	GLN	-	expression tag	UNP Q930S8
A	512	SER	-	expression tag	UNP Q930S8
A	513	HIS	-	expression tag	UNP Q930S8
A	514	HIS	-	expression tag	UNP Q930S8
A	515	HIS	-	expression tag	UNP Q930S8
A	516	HIS	-	expression tag	UNP Q930S8
A	517	HIS	-	expression tag	UNP Q930S8
A	518	HIS	-	expression tag	UNP Q930S8
A	519	TRP	-	expression tag	UNP Q930S8
A	520	SER	-	expression tag	UNP Q930S8
A	521	HIS	-	expression tag	UNP Q930S8
A	522	PRO	-	expression tag	UNP Q930S8
A	523	GLN	-	expression tag	UNP Q930S8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	524	PHE	-	expression tag	UNP Q930S8
A	525	GLU	-	expression tag	UNP Q930S8
A	526	LYS	-	expression tag	UNP Q930S8
B	-1	MSE	-	expression tag	UNP Q930S8
B	0	VAL	-	expression tag	UNP Q930S8
B	505	ALA	-	expression tag	UNP Q930S8
B	506	GLU	-	expression tag	UNP Q930S8
B	507	ASN	-	expression tag	UNP Q930S8
B	508	LEU	-	expression tag	UNP Q930S8
B	509	TYR	-	expression tag	UNP Q930S8
B	510	PHE	-	expression tag	UNP Q930S8
B	511	GLN	-	expression tag	UNP Q930S8
B	512	SER	-	expression tag	UNP Q930S8
B	513	HIS	-	expression tag	UNP Q930S8
B	514	HIS	-	expression tag	UNP Q930S8
B	515	HIS	-	expression tag	UNP Q930S8
B	516	HIS	-	expression tag	UNP Q930S8
B	517	HIS	-	expression tag	UNP Q930S8
B	518	HIS	-	expression tag	UNP Q930S8
B	519	TRP	-	expression tag	UNP Q930S8
B	520	SER	-	expression tag	UNP Q930S8
B	521	HIS	-	expression tag	UNP Q930S8
B	522	PRO	-	expression tag	UNP Q930S8
B	523	GLN	-	expression tag	UNP Q930S8
B	524	PHE	-	expression tag	UNP Q930S8
B	525	GLU	-	expression tag	UNP Q930S8
B	526	LYS	-	expression tag	UNP Q930S8
C	-1	MSE	-	expression tag	UNP Q930S8
C	0	VAL	-	expression tag	UNP Q930S8
C	505	ALA	-	expression tag	UNP Q930S8
C	506	GLU	-	expression tag	UNP Q930S8
C	507	ASN	-	expression tag	UNP Q930S8
C	508	LEU	-	expression tag	UNP Q930S8
C	509	TYR	-	expression tag	UNP Q930S8
C	510	PHE	-	expression tag	UNP Q930S8
C	511	GLN	-	expression tag	UNP Q930S8
C	512	SER	-	expression tag	UNP Q930S8
C	513	HIS	-	expression tag	UNP Q930S8
C	514	HIS	-	expression tag	UNP Q930S8
C	515	HIS	-	expression tag	UNP Q930S8
C	516	HIS	-	expression tag	UNP Q930S8
C	517	HIS	-	expression tag	UNP Q930S8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	518	HIS	-	expression tag	UNP Q930S8
C	519	TRP	-	expression tag	UNP Q930S8
C	520	SER	-	expression tag	UNP Q930S8
C	521	HIS	-	expression tag	UNP Q930S8
C	522	PRO	-	expression tag	UNP Q930S8
C	523	GLN	-	expression tag	UNP Q930S8
C	524	PHE	-	expression tag	UNP Q930S8
C	525	GLU	-	expression tag	UNP Q930S8
C	526	LYS	-	expression tag	UNP Q930S8
D	-1	MSE	-	expression tag	UNP Q930S8
D	0	VAL	-	expression tag	UNP Q930S8
D	505	ALA	-	expression tag	UNP Q930S8
D	506	GLU	-	expression tag	UNP Q930S8
D	507	ASN	-	expression tag	UNP Q930S8
D	508	LEU	-	expression tag	UNP Q930S8
D	509	TYR	-	expression tag	UNP Q930S8
D	510	PHE	-	expression tag	UNP Q930S8
D	511	GLN	-	expression tag	UNP Q930S8
D	512	SER	-	expression tag	UNP Q930S8
D	513	HIS	-	expression tag	UNP Q930S8
D	514	HIS	-	expression tag	UNP Q930S8
D	515	HIS	-	expression tag	UNP Q930S8
D	516	HIS	-	expression tag	UNP Q930S8
D	517	HIS	-	expression tag	UNP Q930S8
D	518	HIS	-	expression tag	UNP Q930S8
D	519	TRP	-	expression tag	UNP Q930S8
D	520	SER	-	expression tag	UNP Q930S8
D	521	HIS	-	expression tag	UNP Q930S8
D	522	PRO	-	expression tag	UNP Q930S8
D	523	GLN	-	expression tag	UNP Q930S8
D	524	PHE	-	expression tag	UNP Q930S8
D	525	GLU	-	expression tag	UNP Q930S8
D	526	LYS	-	expression tag	UNP Q930S8

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
2	A	1	1	1	0	0

- Molecule 3 is water.

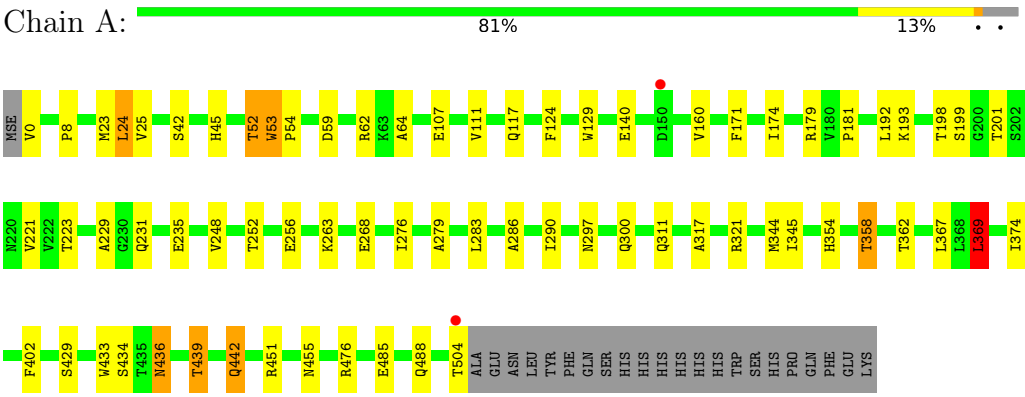
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	193	Total O 193 193	0	0
3	B	108	Total O 108 108	0	0
3	C	176	Total O 176 176	0	0
3	D	141	Total O 141 141	0	0

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.

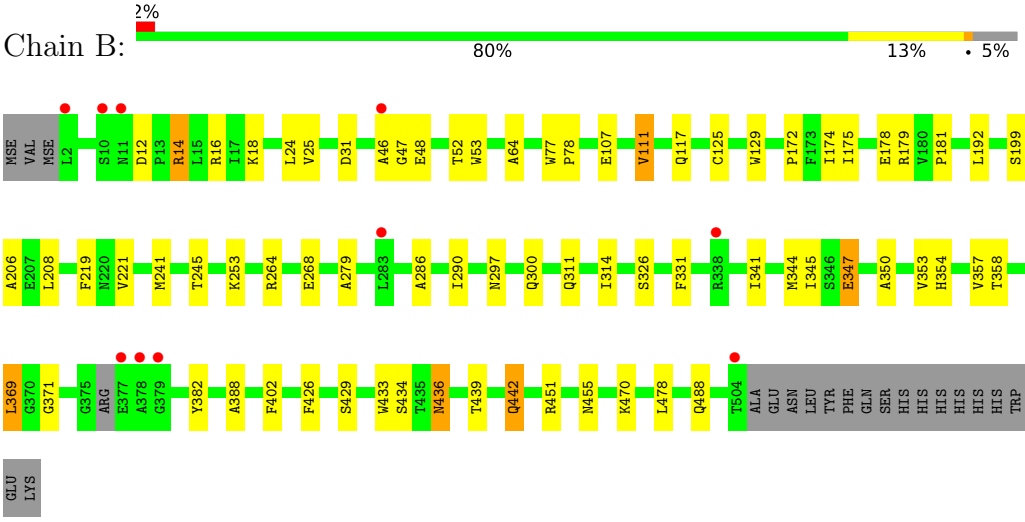
• Molecule 1: NAD-dependent aldehyde dehydrogenase

Chain A:



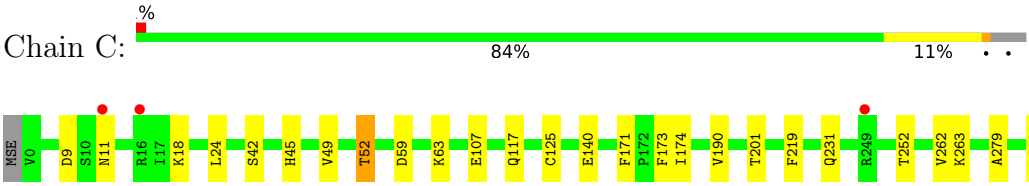
• Molecule 1: NAD-dependent aldehyde dehydrogenase

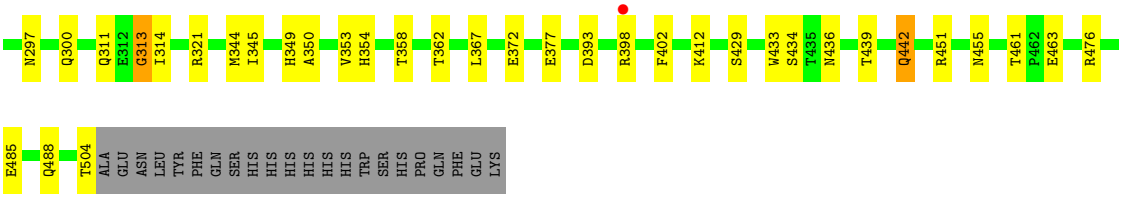
Chain B:



• Molecule 1: NAD-dependent aldehyde dehydrogenase

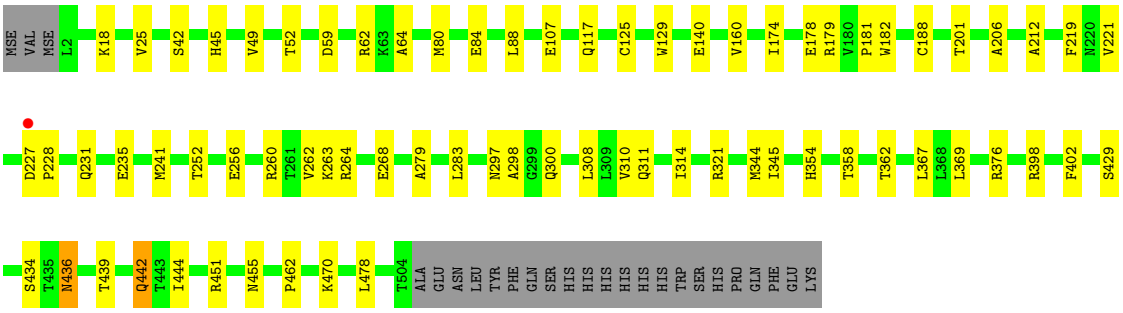
Chain C:





• Molecule 1: NAD-dependent aldehyde dehydrogenase

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	79.11Å 176.66Å 83.66Å 90.00° 98.49° 90.00°	Depositor
Resolution (Å)	47.98 – 2.00 47.98 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.9 (47.98-2.00) 99.9 (47.98-2.00)	Depositor EDS
R_{merge}	0.92	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.71 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.5.0110	Depositor
R, R_{free}	0.187 , 0.224 0.190 , 0.226	Depositor DCC
R_{free} test set	7629 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	17.4	Xtriage
Anisotropy	0.043	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 42.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	15736	wwPDB-VP
Average B, all atoms (Å^2)	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.69% 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: CA

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	0/3851	0.69	2/5213 (0.0%)
1	B	0.62	0/3831	0.67	0/5186
1	C	0.66	0/3857	0.68	0/5220
1	D	0.64	0/3843	0.68	1/5203 (0.0%)
All	All	0.65	0/15382	0.68	3/20822 (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	C	0	1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	369	LEU	CA-CB-CG	5.42	127.76	115.30
1	D	212	ALA	N-CA-C	5.33	125.38	111.00
1	A	24	LEU	CA-CB-CG	5.05	126.93	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	313	GLY	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	3785	0	3777	53	0
1	B	3765	0	3753	51	0
1	C	3791	0	3788	40	0
1	D	3776	0	3767	45	0
2	A	1	0	0	0	0
3	A	193	0	0	6	0
3	B	108	0	0	1	0
3	C	176	0	0	4	0
3	D	141	0	0	4	0
All	All	15736	0	15085	187	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 6.

The worst 5 of 187 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:344:MSE:HE1	1:B:353:VAL:HG21	1.34	1.05
1:C:344:MSE:HE1	1:C:353:VAL:HG21	1.45	0.98
1:D:80:MSE:HE3	1:D:84:GLU:HB3	1.46	0.96
1:C:279:ALA:H	1:C:311:GLN:HE21	1.12	0.94
1:A:279:ALA:H	1:A:311:GLN:HE21	1.11	0.90

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	503/528 (95%)	490 (97%)	13 (3%)	0	100	100
1	B	498/528 (94%)	482 (97%)	15 (3%)	1 (0%)	47	44
1	C	503/528 (95%)	490 (97%)	13 (3%)	0	100	100
1	D	501/528 (95%)	486 (97%)	15 (3%)	0	100	100
All	All	2005/2112 (95%)	1948 (97%)	56 (3%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	47	GLY

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	387/401 (96%)	372 (96%)	15 (4%)	32	30
1	B	385/401 (96%)	368 (96%)	17 (4%)	28	25
1	C	388/401 (97%)	375 (97%)	13 (3%)	37	36
1	D	386/401 (96%)	377 (98%)	9 (2%)	50	53
All	All	1546/1604 (96%)	1492 (96%)	54 (4%)	36	35

5 of 54 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	402	PHE
1	C	219	PHE
1	D	369	LEU
1	B	436	ASN
1	C	24	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	297	ASN
1	D	300	GLN
1	D	442	GLN
1	B	436	ASN
1	B	311	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 1 ligands modelled in this entry, 1 is 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 (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

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	497/528 (94%)	-0.53	2 (0%)	92	92	7, 13, 24, 41	0
1	B	495/528 (93%)	-0.14	10 (2%)	65	63	7, 19, 41, 57	0
1	C	497/528 (94%)	-0.49	4 (0%)	86	85	6, 14, 27, 39	0
1	D	496/528 (93%)	-0.37	1 (0%)	95	94	9, 17, 29, 43	0
All	All	1985/2112 (93%)	-0.38	17 (0%)	84	83	6, 15, 33, 57	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	46	ALA	5.9
1	D	227	ASP	4.1
1	B	283	LEU	2.8
1	B	378	ALA	2.7
1	A	504	THR	2.6

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

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	CA	A	601	1/1	0.99	0.06	15,15,15,15	0

6.5 Other polymers (i)

There are no such residues in this entry.