



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 12:23 PM GMT

PDB ID : 3R64
Title : Crystal structure of a NAD-dependent benzaldehyde dehydrogenase from *Corynebacterium glutamicum*
Authors : Agarwal, R.; Almo, S.C.; Swaminathan, S.; New York Structural Genomics Research Consortium (NYSGRC)
Deposited on : 2011-03-21
Resolution : 2.57 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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : trunk26865

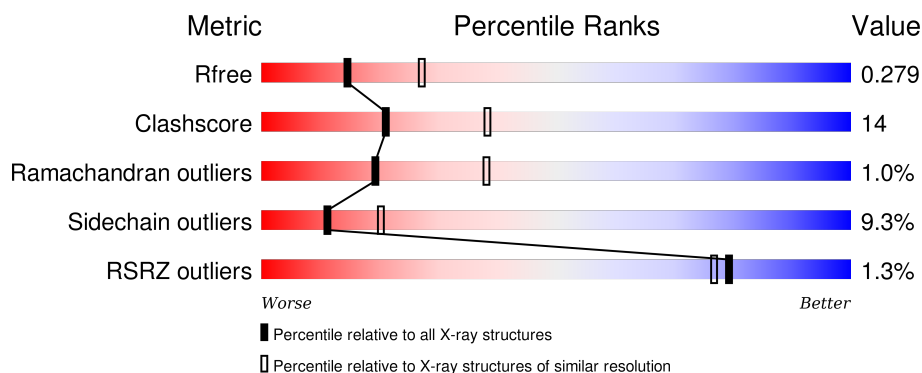
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.57 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}	91344	2636 (2.60-2.56)
Clashscore	102246	3003 (2.60-2.56)
Ramachandran outliers	100387	2956 (2.60-2.56)
Sidechain outliers	100360	2956 (2.60-2.56)
RSRZ outliers	91569	2642 (2.60-2.56)

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	508	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, green 61%, yellow 27%, orange 8%, grey 4%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> 61% 27% • 8% </div> </div>
1	B	508	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, green 64%, yellow 25%, orange 10%, grey 2%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> 64% 25% • 10% </div> </div>
1	C	508	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 3%, green 61%, yellow 25%, orange 11%, grey 2%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> 61% 25% • 11% </div> </div>
1	D	508	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green 60%, yellow 26%, orange 10%, grey 4%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> 60% 26% • 10% </div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 13674 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 benzaldehyde dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	468	Total	C	N	O	S	0	0	0
			3411	2137	596	674	4			
1	B	459	Total	C	N	O	S	0	0	0
			3349	2100	580	664	5			
1	C	453	Total	C	N	O	S	0	0	0
			3243	2030	571	638	4			
1	D	459	Total	C	N	O	S	0	0	0
			3384	2126	591	662	5			

There are 96 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	EXPRESSION TAG	UNP Q6M2H6
A	2	VAL	-	EXPRESSION TAG	UNP Q6M2H6
A	487	ALA	-	EXPRESSION TAG	UNP Q6M2H6
A	488	GLU	-	EXPRESSION TAG	UNP Q6M2H6
A	489	ASN	-	EXPRESSION TAG	UNP Q6M2H6
A	490	LEU	-	EXPRESSION TAG	UNP Q6M2H6
A	491	TYR	-	EXPRESSION TAG	UNP Q6M2H6
A	492	PHE	-	EXPRESSION TAG	UNP Q6M2H6
A	493	GLN	-	EXPRESSION TAG	UNP Q6M2H6
A	494	SER	-	EXPRESSION TAG	UNP Q6M2H6
A	495	HIS	-	EXPRESSION TAG	UNP Q6M2H6
A	496	HIS	-	EXPRESSION TAG	UNP Q6M2H6
A	497	HIS	-	EXPRESSION TAG	UNP Q6M2H6
A	498	HIS	-	EXPRESSION TAG	UNP Q6M2H6
A	499	HIS	-	EXPRESSION TAG	UNP Q6M2H6
A	500	HIS	-	EXPRESSION TAG	UNP Q6M2H6
A	501	TRP	-	EXPRESSION TAG	UNP Q6M2H6
A	502	SER	-	EXPRESSION TAG	UNP Q6M2H6
A	503	HIS	-	EXPRESSION TAG	UNP Q6M2H6
A	504	PRO	-	EXPRESSION TAG	UNP Q6M2H6
A	505	GLN	-	EXPRESSION TAG	UNP Q6M2H6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	506	PHE	-	EXPRESSION TAG	UNP Q6M2H6
A	507	GLU	-	EXPRESSION TAG	UNP Q6M2H6
A	508	LYS	-	EXPRESSION TAG	UNP Q6M2H6
B	1	MET	-	EXPRESSION TAG	UNP Q6M2H6
B	2	VAL	-	EXPRESSION TAG	UNP Q6M2H6
B	487	ALA	-	EXPRESSION TAG	UNP Q6M2H6
B	488	GLU	-	EXPRESSION TAG	UNP Q6M2H6
B	489	ASN	-	EXPRESSION TAG	UNP Q6M2H6
B	490	LEU	-	EXPRESSION TAG	UNP Q6M2H6
B	491	TYR	-	EXPRESSION TAG	UNP Q6M2H6
B	492	PHE	-	EXPRESSION TAG	UNP Q6M2H6
B	493	GLN	-	EXPRESSION TAG	UNP Q6M2H6
B	494	SER	-	EXPRESSION TAG	UNP Q6M2H6
B	495	HIS	-	EXPRESSION TAG	UNP Q6M2H6
B	496	HIS	-	EXPRESSION TAG	UNP Q6M2H6
B	497	HIS	-	EXPRESSION TAG	UNP Q6M2H6
B	498	HIS	-	EXPRESSION TAG	UNP Q6M2H6
B	499	HIS	-	EXPRESSION TAG	UNP Q6M2H6
B	500	HIS	-	EXPRESSION TAG	UNP Q6M2H6
B	501	TRP	-	EXPRESSION TAG	UNP Q6M2H6
B	502	SER	-	EXPRESSION TAG	UNP Q6M2H6
B	503	HIS	-	EXPRESSION TAG	UNP Q6M2H6
B	504	PRO	-	EXPRESSION TAG	UNP Q6M2H6
B	505	GLN	-	EXPRESSION TAG	UNP Q6M2H6
B	506	PHE	-	EXPRESSION TAG	UNP Q6M2H6
B	507	GLU	-	EXPRESSION TAG	UNP Q6M2H6
B	508	LYS	-	EXPRESSION TAG	UNP Q6M2H6
C	1	MET	-	EXPRESSION TAG	UNP Q6M2H6
C	2	VAL	-	EXPRESSION TAG	UNP Q6M2H6
C	487	ALA	-	EXPRESSION TAG	UNP Q6M2H6
C	488	GLU	-	EXPRESSION TAG	UNP Q6M2H6
C	489	ASN	-	EXPRESSION TAG	UNP Q6M2H6
C	490	LEU	-	EXPRESSION TAG	UNP Q6M2H6
C	491	TYR	-	EXPRESSION TAG	UNP Q6M2H6
C	492	PHE	-	EXPRESSION TAG	UNP Q6M2H6
C	493	GLN	-	EXPRESSION TAG	UNP Q6M2H6
C	494	SER	-	EXPRESSION TAG	UNP Q6M2H6
C	495	HIS	-	EXPRESSION TAG	UNP Q6M2H6
C	496	HIS	-	EXPRESSION TAG	UNP Q6M2H6
C	497	HIS	-	EXPRESSION TAG	UNP Q6M2H6
C	498	HIS	-	EXPRESSION TAG	UNP Q6M2H6
C	499	HIS	-	EXPRESSION TAG	UNP Q6M2H6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	500	HIS	-	EXPRESSION TAG	UNP Q6M2H6
C	501	TRP	-	EXPRESSION TAG	UNP Q6M2H6
C	502	SER	-	EXPRESSION TAG	UNP Q6M2H6
C	503	HIS	-	EXPRESSION TAG	UNP Q6M2H6
C	504	PRO	-	EXPRESSION TAG	UNP Q6M2H6
C	505	GLN	-	EXPRESSION TAG	UNP Q6M2H6
C	506	PHE	-	EXPRESSION TAG	UNP Q6M2H6
C	507	GLU	-	EXPRESSION TAG	UNP Q6M2H6
C	508	LYS	-	EXPRESSION TAG	UNP Q6M2H6
D	1	MET	-	EXPRESSION TAG	UNP Q6M2H6
D	2	VAL	-	EXPRESSION TAG	UNP Q6M2H6
D	487	ALA	-	EXPRESSION TAG	UNP Q6M2H6
D	488	GLU	-	EXPRESSION TAG	UNP Q6M2H6
D	489	ASN	-	EXPRESSION TAG	UNP Q6M2H6
D	490	LEU	-	EXPRESSION TAG	UNP Q6M2H6
D	491	TYR	-	EXPRESSION TAG	UNP Q6M2H6
D	492	PHE	-	EXPRESSION TAG	UNP Q6M2H6
D	493	GLN	-	EXPRESSION TAG	UNP Q6M2H6
D	494	SER	-	EXPRESSION TAG	UNP Q6M2H6
D	495	HIS	-	EXPRESSION TAG	UNP Q6M2H6
D	496	HIS	-	EXPRESSION TAG	UNP Q6M2H6
D	497	HIS	-	EXPRESSION TAG	UNP Q6M2H6
D	498	HIS	-	EXPRESSION TAG	UNP Q6M2H6
D	499	HIS	-	EXPRESSION TAG	UNP Q6M2H6
D	500	HIS	-	EXPRESSION TAG	UNP Q6M2H6
D	501	TRP	-	EXPRESSION TAG	UNP Q6M2H6
D	502	SER	-	EXPRESSION TAG	UNP Q6M2H6
D	503	HIS	-	EXPRESSION TAG	UNP Q6M2H6
D	504	PRO	-	EXPRESSION TAG	UNP Q6M2H6
D	505	GLN	-	EXPRESSION TAG	UNP Q6M2H6
D	506	PHE	-	EXPRESSION TAG	UNP Q6M2H6
D	507	GLU	-	EXPRESSION TAG	UNP Q6M2H6
D	508	LYS	-	EXPRESSION TAG	UNP Q6M2H6

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	69	Total O 69 69	0	0
2	B	86	Total O 86 86	0	0
2	C	51	Total O 51 51	0	0

Continued on next page...

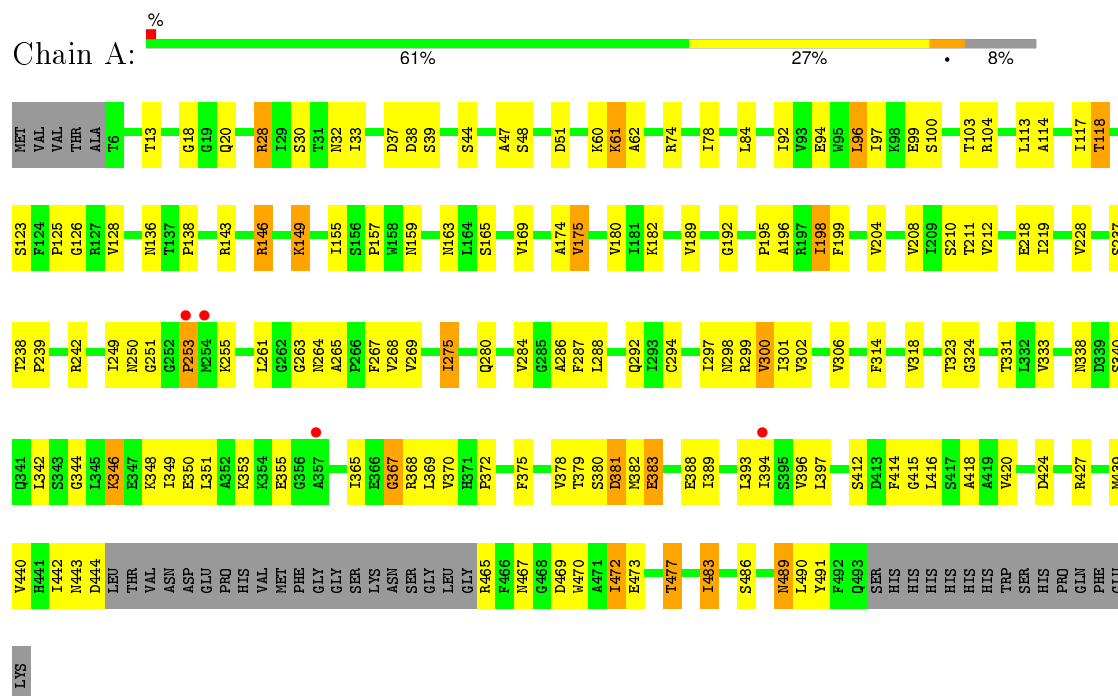
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	D	81	Total	O	0	0
			81	81		

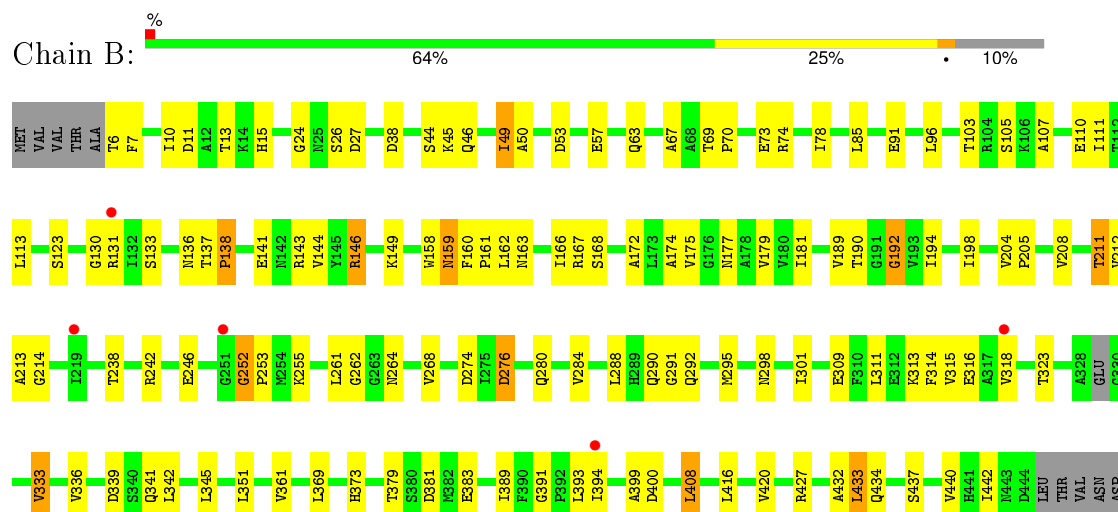
3 Residue-property plots

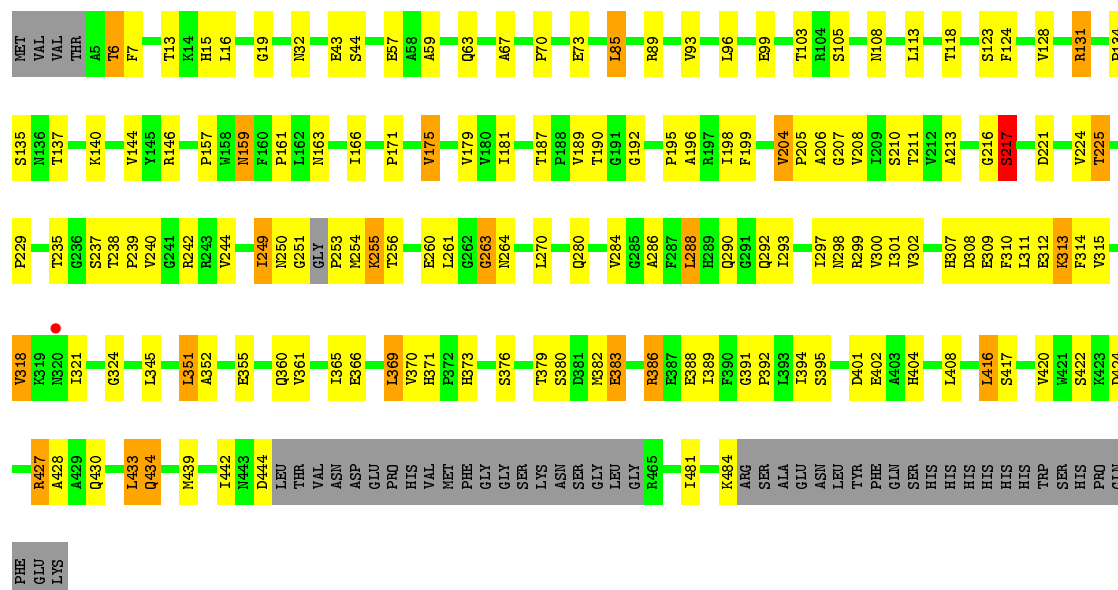
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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: NAD dependent benzaldehyde dehydrogenase



- Molecule 1: NAD dependent benzaldehyde dehydrogenase





4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	140.92Å 99.83Å 160.97Å 90.00° 113.94° 90.00°	Depositor
Resolution (Å)	50.00 – 2.57 50.00 – 2.57	Depositor EDS
% Data completeness (in resolution range)	97.2 (50.00-2.57) 97.3 (50.00-2.57)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.62 (at 2.58Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.235 , 0.282 0.235 , 0.279	Depositor DCC
R_{free} test set	1929 reflections (3.04%)	DCC
Wilson B-factor (Å ²)	42.2	Xtriage
Anisotropy	0.262	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 48.2	EDS
Estimated twinning fraction	0.015 for h,-k,-h-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 65060 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	13674	wwPDB-VP
Average B, all atoms (Å ²)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.91% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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	0.46	0/3474	0.69	1/4734 (0.0%)
1	B	0.46	0/3410	0.70	2/4645 (0.0%)
1	C	0.43	0/3302	0.64	2/4505 (0.0%)
1	D	0.48	0/3446	0.73	3/4691 (0.1%)
All	All	0.46	0/13632	0.69	8/18575 (0.0%)

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	217	SER	N-CA-C	5.70	126.38	111.00
1	D	416	LEU	CA-CB-CG	5.34	127.58	115.30
1	B	138	PRO	N-CA-CB	5.25	109.60	103.30
1	C	273	ALA	N-CA-C	5.20	125.05	111.00
1	A	138	PRO	N-CA-CB	5.17	109.50	103.30

There are no chirality outliers.

There are no planarity outliers.

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	3411	0	3268	109	0
1	B	3349	0	3220	84	0
1	C	3243	0	3039	97	0
1	D	3384	0	3305	98	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	69	0	0	2	0
2	B	86	0	0	4	0
2	C	51	0	0	1	0
2	D	81	0	0	1	0
All	All	13674	0	12832	375	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 14.

The worst 5 of 375 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:286:ALA:CB	1:A:394:ILE:HD11	1.74	1.17
1:A:286:ALA:HB1	1:A:394:ILE:CD1	1.76	1.14
1:B:393:LEU:O	1:B:394:ILE:HD13	1.68	0.92
1:C:46:GLN:HG3	1:C:214:GLY:HA2	1.56	0.87
1:B:13:THR:HB	1:B:44:SER:HB2	1.58	0.86

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	464/508 (91%)	429 (92%)	28 (6%)	7 (2%)	13	25
1	B	453/508 (89%)	422 (93%)	28 (6%)	3 (1%)	26	49
1	C	445/508 (88%)	411 (92%)	30 (7%)	4 (1%)	21	41
1	D	453/508 (89%)	428 (94%)	21 (5%)	4 (1%)	21	41
All	All	1815/2032 (89%)	1690 (93%)	107 (6%)	18 (1%)	19	38

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	380	SER
1	A	415	GLY
1	B	416	LEU
1	C	10	ILE
1	D	216	GLY

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	338/398 (85%)	308 (91%)	30 (9%)	12	22
1	B	335/398 (84%)	303 (90%)	32 (10%)	10	19
1	C	308/398 (77%)	286 (93%)	22 (7%)	18	35
1	D	343/398 (86%)	304 (89%)	39 (11%)	7	12
All	All	1324/1592 (83%)	1201 (91%)	123 (9%)	11	20

5 of 123 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	383	GLU
1	C	179	VAL
1	D	366	GLU
1	B	408	LEU
1	C	46	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 48 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	250	ASN
1	B	360	GLN
1	D	264	ASN
1	B	264	ASN
1	B	338	ASN

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](#)

There are no ligands in this entry.

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	468/508 (92%)	0.24	4 (0%) 85 84	21, 48, 63, 71	0
1	B	459/508 (90%)	0.06	5 (1%) 82 80	18, 42, 57, 67	0
1	C	453/508 (89%)	0.34	14 (3%) 52 48	26, 51, 67, 76	0
1	D	459/508 (90%)	0.06	1 (0%) 95 95	16, 39, 56, 66	0
All	All	1839/2032 (90%)	0.18	24 (1%) 79 77	16, 45, 62, 76	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	331	THR	3.8
1	B	251	GLY	3.5
1	C	470	TRP	3.5
1	A	394	ILE	3.5
1	C	394	ILE	3.4

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.