



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 11, 2022 – 06:04 PM EDT

PDB ID : 8D57  
Title : Crystal Structure of dihydrodipicolinate reductase from *Acinetobacter baumannii*  
Authors : Seattle Structural Genomics Center for Infectious Disease; Seattle Structural Genomics Center for Infectious Disease (SSGCID)  
Deposited on : 2022-06-04  
Resolution : 2.65 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

<https://www.wwpdb.org/validation/2017/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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.28.1  
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)  
Validation Pipeline (wwPDB-VP) : 2.28.1

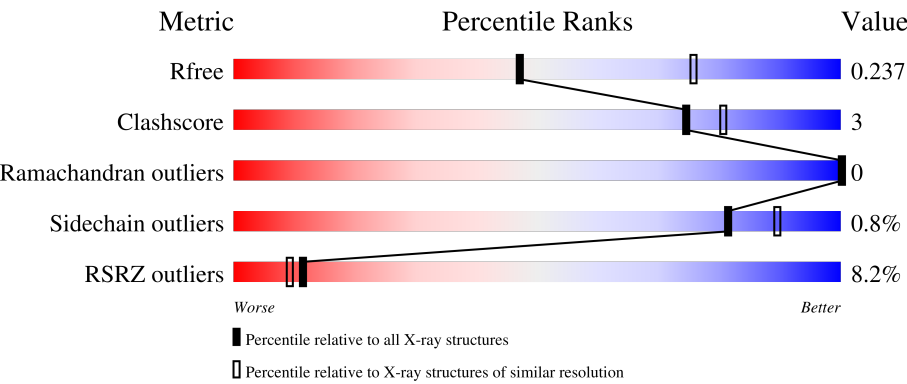
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.65 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	1332 (2.68-2.64)
Clashscore	141614	1374 (2.68-2.64)
Ramachandran outliers	138981	1349 (2.68-2.64)
Sidechain outliers	138945	1349 (2.68-2.64)
RSRZ outliers	127900	1318 (2.68-2.64)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	281	<div><div>2%</div><div></div><div>85%</div><div>10%</div><div>5%</div></div>
1	B	281	<div><div>12%</div><div></div><div>89%</div><div>6%</div><div>5%</div></div>
1	C	281	<div><div>2%</div><div></div><div>88%</div><div>7%</div><div>5%</div></div>
1	D	281	<div><div>14%</div><div></div><div>86%</div><div>9%</div><div>5%</div></div>
1	E	281	<div><div>15%</div><div></div><div>89%</div><div>6%</div><div>5%</div></div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	281	<div> <div>%</div> <div>  <div>89%6%5%</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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	CIT	B	301	-	-	-	X

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 11302 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 4-hydroxy-tetrahydrodipicolinate reductase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	267	Total	C	N	O	S	0	0	0
			1907	1190	337	371	9			
1	B	268	Total	C	N	O	S	0	0	0
			1826	1138	325	355	8			
1	C	267	Total	C	N	O	S	0	1	0
			1903	1190	337	367	9			
1	D	266	Total	C	N	O	S	0	1	0
			1830	1142	326	353	9			
1	E	267	Total	C	N	O	S	0	0	0
			1770	1098	316	348	8			
1	F	267	Total	C	N	O	S	0	0	0
			1901	1187	334	371	9			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	MET	-	initiating methionine	UNP A0A0D5YCZ0
A	-6	ALA	-	expression tag	UNP A0A0D5YCZ0
A	-5	HIS	-	expression tag	UNP A0A0D5YCZ0
A	-4	HIS	-	expression tag	UNP A0A0D5YCZ0
A	-3	HIS	-	expression tag	UNP A0A0D5YCZ0
A	-2	HIS	-	expression tag	UNP A0A0D5YCZ0
A	-1	HIS	-	expression tag	UNP A0A0D5YCZ0
A	0	HIS	-	expression tag	UNP A0A0D5YCZ0
B	-7	MET	-	initiating methionine	UNP A0A0D5YCZ0
B	-6	ALA	-	expression tag	UNP A0A0D5YCZ0
B	-5	HIS	-	expression tag	UNP A0A0D5YCZ0
B	-4	HIS	-	expression tag	UNP A0A0D5YCZ0
B	-3	HIS	-	expression tag	UNP A0A0D5YCZ0
B	-2	HIS	-	expression tag	UNP A0A0D5YCZ0
B	-1	HIS	-	expression tag	UNP A0A0D5YCZ0
B	0	HIS	-	expression tag	UNP A0A0D5YCZ0
C	-7	MET	-	initiating methionine	UNP A0A0D5YCZ0

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	-6	ALA	-	expression tag	UNP A0A0D5YCZ0
C	-5	HIS	-	expression tag	UNP A0A0D5YCZ0
C	-4	HIS	-	expression tag	UNP A0A0D5YCZ0
C	-3	HIS	-	expression tag	UNP A0A0D5YCZ0
C	-2	HIS	-	expression tag	UNP A0A0D5YCZ0
C	-1	HIS	-	expression tag	UNP A0A0D5YCZ0
C	0	HIS	-	expression tag	UNP A0A0D5YCZ0
D	-7	MET	-	initiating methionine	UNP A0A0D5YCZ0
D	-6	ALA	-	expression tag	UNP A0A0D5YCZ0
D	-5	HIS	-	expression tag	UNP A0A0D5YCZ0
D	-4	HIS	-	expression tag	UNP A0A0D5YCZ0
D	-3	HIS	-	expression tag	UNP A0A0D5YCZ0
D	-2	HIS	-	expression tag	UNP A0A0D5YCZ0
D	-1	HIS	-	expression tag	UNP A0A0D5YCZ0
D	0	HIS	-	expression tag	UNP A0A0D5YCZ0
E	-7	MET	-	initiating methionine	UNP A0A0D5YCZ0
E	-6	ALA	-	expression tag	UNP A0A0D5YCZ0
E	-5	HIS	-	expression tag	UNP A0A0D5YCZ0
E	-4	HIS	-	expression tag	UNP A0A0D5YCZ0
E	-3	HIS	-	expression tag	UNP A0A0D5YCZ0
E	-2	HIS	-	expression tag	UNP A0A0D5YCZ0
E	-1	HIS	-	expression tag	UNP A0A0D5YCZ0
E	0	HIS	-	expression tag	UNP A0A0D5YCZ0
F	-7	MET	-	initiating methionine	UNP A0A0D5YCZ0
F	-6	ALA	-	expression tag	UNP A0A0D5YCZ0
F	-5	HIS	-	expression tag	UNP A0A0D5YCZ0
F	-4	HIS	-	expression tag	UNP A0A0D5YCZ0
F	-3	HIS	-	expression tag	UNP A0A0D5YCZ0
F	-2	HIS	-	expression tag	UNP A0A0D5YCZ0
F	-1	HIS	-	expression tag	UNP A0A0D5YCZ0
F	0	HIS	-	expression tag	UNP A0A0D5YCZ0

- Molecule 2 is CITRIC ACID (three-letter code: CIT) (formula: C<sub>6</sub>H<sub>8</sub>O<sub>7</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			13	6	7		
2	B	1	Total	C	O	0	0
			13	6	7		
2	C	1	Total	C	O	0	0
			13	6	7		
2	D	1	Total	C	O	0	0
			13	6	7		
2	E	1	Total	C	O	0	0
			13	6	7		
2	F	1	Total	C	O	0	0
			13	6	7		

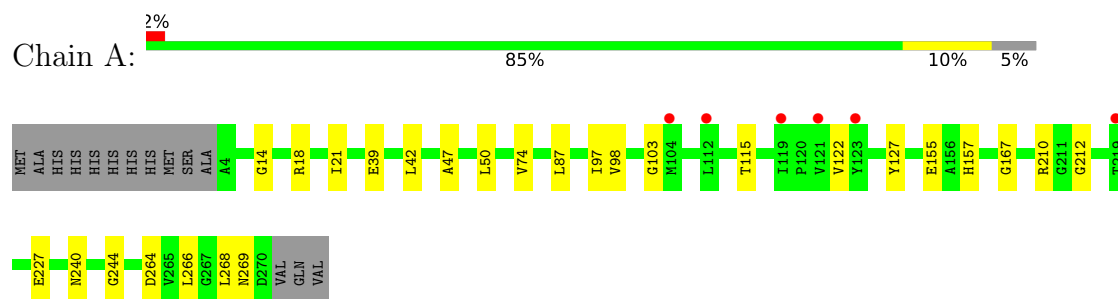
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	32	Total	O	0	0
			32	32		
3	B	8	Total	O	0	0
			8	8		
3	C	19	Total	O	0	0
			19	19		
3	D	18	Total	O	0	0
			18	18		
3	E	1	Total	O	0	0
			1	1		
3	F	9	Total	O	0	0
			9	9		

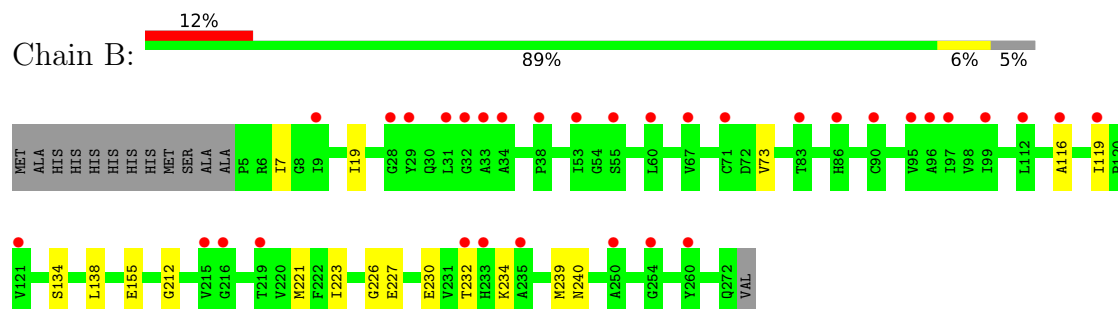
### 3 Residue-property plots [i](#)

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.

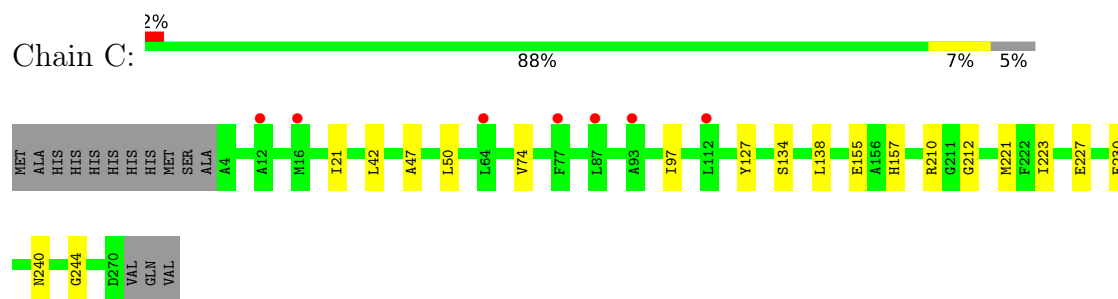
- Molecule 1: 4-hydroxy-tetrahydronicotinate reductase



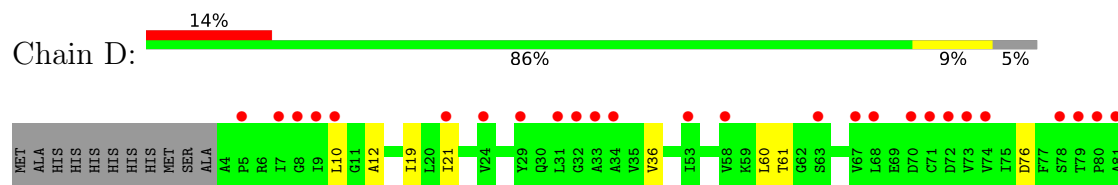
- Molecule 1: 4-hydroxy-tetrahydronicotinate reductase



- Molecule 1: 4-hydroxy-tetrahydronicotinate reductase

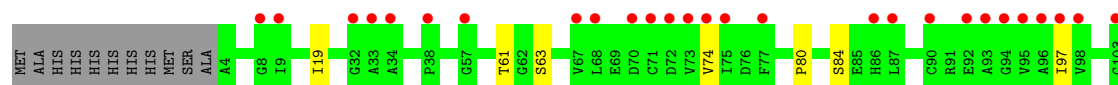
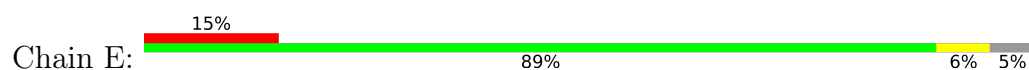


- Molecule 1: 4-hydroxy-tetrahydronicotinate reductase

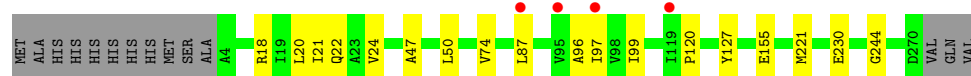
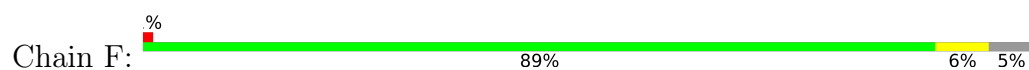




- Molecule 1: 4-hydroxy-tetrahydronicotinate reductase



- Molecule 1: 4-hydroxy-tetrahydronicotinate reductase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	79.80Å 79.80Å 519.13Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	41.51 – 2.65 47.31 – 2.65	Depositor EDS
% Data completeness (in resolution range)	99.2 (41.51-2.65) 99.3 (47.31-2.65)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.75 (at 2.65Å)	Xtriage
Refinement program	PHENIX 1.20.1	Depositor
R, $R_{free}$	0.208 , 0.235 0.210 , 0.237	Depositor DCC
$R_{free}$ test set	1977 reflections (3.45%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	68.0	Xtriage
Anisotropy	0.438	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 57.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.024 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	11302	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	86.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.59% 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: CIT

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	0.31	0/1932	0.55	0/2626
1	B	0.30	0/1849	0.54	0/2519
1	C	0.31	0/1931	0.54	0/2626
1	D	0.29	0/1857	0.52	0/2530
1	E	0.27	0/1793	0.51	0/2451
1	F	0.31	0/1926	0.54	0/2619
All	All	0.30	0/11288	0.53	0/15371

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	1907	0	1861	20	0
1	B	1826	0	1710	13	0
1	C	1903	0	1862	14	0
1	D	1830	0	1736	15	0
1	E	1770	0	1598	9	0
1	F	1901	0	1850	9	0
2	A	13	0	5	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	13	0	5	0	0
2	C	13	0	5	0	0
2	D	13	0	5	0	0
2	E	13	0	5	0	0
2	F	13	0	5	1	0
3	A	32	0	0	2	0
3	B	8	0	0	0	0
3	C	19	0	0	0	0
3	D	18	0	0	0	0
3	E	1	0	0	0	0
3	F	9	0	0	0	0
All	All	11302	0	10647	70	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 3.

All (70) 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:240:ASN:HD22	1:D:227:GLU:HB3	1.41	0.85
1:A:240:ASN:HD22	1:C:227:GLU:HB3	1.44	0.83
1:A:227:GLU:HB3	1:C:240:ASN:HD22	1.54	0.72
1:A:42:LEU:HD13	1:A:50:LEU:HD21	1.78	0.66
1:A:21:ILE:HD13	1:A:47:ALA:HB1	1.79	0.65
1:E:61:THR:HG22	1:E:63:SER:H	1.62	0.63
1:A:39:GLU:OE1	1:A:39:GLU:N	2.33	0.61
1:C:42:LEU:HD13	1:C:50:LEU:HD21	1.82	0.60
1:F:21:ILE:HD13	1:F:47:ALA:HB1	1.84	0.59
1:F:221:MET:HG2	1:F:230:GLU:HG2	1.86	0.58
1:C:21:ILE:HD13	1:C:47:ALA:HB1	1.85	0.57
2:F:301:CIT:O3	2:F:301:CIT:O7	2.20	0.55
1:E:127:TYR:O	1:E:244:GLY:HA3	2.08	0.54
1:A:167:GLY:H	2:A:301:CIT:C5	2.21	0.53
1:F:96:ALA:HB2	1:F:120:PRO:HG2	1.90	0.53
1:B:227:GLU:HB3	1:D:240:ASN:HD22	1.74	0.53
1:D:10:LEU:HD12	1:D:76:ASP:HB2	1.91	0.53
1:E:221:MET:HG2	1:E:230:GLU:HG2	1.91	0.52
1:E:74:VAL:HG23	1:E:97:ILE:HG13	1.92	0.51
1:F:18:ARG:O	1:F:22:GLN:HG3	2.10	0.51
1:E:134:SER:O	1:E:138:LEU:HG	2.10	0.51
1:C:127:TYR:O	1:C:244:GLY:HA3	2.11	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:87:LEU:HD21	1:F:99:ILE:HD11	1.93	0.50
1:A:14:GLY:O	1:A:18:ARG:HG3	2.11	0.50
1:B:19:ILE:HD13	1:B:239:MET:HG2	1.93	0.50
1:C:221:MET:HG2	1:C:230:GLU:HG2	1.92	0.50
1:E:84:SER:OG	1:E:108:GLN:NE2	2.44	0.49
1:D:127:TYR:O	1:D:244:GLY:HA3	2.13	0.49
1:E:19:ILE:HD13	1:E:239:MET:HG2	1.94	0.49
1:D:134:SER:O	1:D:138:LEU:HG	2.13	0.48
1:B:7:ILE:HG23	1:B:73:VAL:HG13	1.95	0.48
1:B:221:MET:HG2	1:B:230:GLU:HG2	1.94	0.48
1:A:240:ASN:ND2	1:C:227:GLU:HB3	2.20	0.47
1:A:103:GLY:N	3:A:402:HOH:O	2.47	0.47
1:A:227:GLU:HB3	1:C:240:ASN:ND2	2.26	0.47
1:A:240:ASN:OD1	1:A:240:ASN:N	2.49	0.46
1:C:157:HIS:HB2	1:C:210:ARG:NH1	2.30	0.46
1:D:168:THR:HB	1:D:218:HIS:NE2	2.31	0.46
1:C:74:VAL:HG23	1:C:97:ILE:HG13	1.97	0.46
1:D:36:VAL:HG11	1:D:60:LEU:HD23	1.99	0.45
1:D:12:ALA:HB1	1:D:21:ILE:HD11	1.97	0.45
1:D:221:MET:HG2	1:D:230:GLU:HG2	1.99	0.45
1:A:98:VAL:HG22	1:A:122:VAL:HB	1.99	0.45
1:B:240:ASN:ND2	1:D:227:GLU:HB3	2.21	0.45
1:A:212:GLY:HA3	1:D:223:ILE:HG22	1.97	0.45
1:A:74:VAL:HG23	1:A:97:ILE:HG13	1.98	0.44
1:A:266:LEU:HB2	1:A:268:LEU:HG	1.99	0.44
1:A:127:TYR:O	1:A:244:GLY:HA3	2.18	0.44
1:F:127:TYR:O	1:F:244:GLY:HA3	2.18	0.44
1:B:223:ILE:HG22	1:C:212:GLY:HA3	2.00	0.43
1:D:143:LYS:HA	1:D:143:LYS:HD3	1.90	0.43
1:F:74:VAL:HG23	1:F:97:ILE:HG13	2.01	0.43
1:A:157:HIS:HB2	1:A:210:ARG:NH1	2.33	0.43
1:E:80:PRO:O	1:E:108:GLN:NE2	2.52	0.43
1:C:134:SER:O	1:C:138:LEU:HG	2.19	0.42
1:A:87:LEU:HD22	1:A:115:THR:HG21	2.00	0.42
1:B:116:ALA:HA	1:B:119:ILE:O	2.19	0.42
1:B:240:ASN:OD1	1:B:240:ASN:N	2.53	0.42
1:F:18:ARG:NH1	1:F:50:LEU:HG	2.35	0.42
1:B:212:GLY:HA3	1:C:223:ILE:HG22	2.03	0.41
1:C:240:ASN:N	1:C:240:ASN:OD1	2.53	0.41
1:A:240:ASN:HA	3:A:412:HOH:O	2.20	0.41
1:B:226:GLY:HA2	1:D:236:THR:HB	2.02	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:134:SER:O	1:B:138:LEU:HG	2.22	0.40
1:F:20:LEU:O	1:F:24:VAL:HG23	2.20	0.40
1:A:264:ASP:OD1	1:A:269:ASN:ND2	2.54	0.40
1:D:87:LEU:HD22	1:D:115:THR:HG21	2.04	0.40
1:B:232:THR:HG21	1:B:234:LYS:HE3	2.03	0.40
1:D:19:ILE:HD13	1:D:239:MET:HG2	2.04	0.40
1:E:240:ASN:OD1	1:E:240:ASN:N	2.54	0.40

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	265/281 (94%)	261 (98%)	4 (2%)	0	100	100
1	B	266/281 (95%)	262 (98%)	4 (2%)	0	100	100
1	C	266/281 (95%)	261 (98%)	5 (2%)	0	100	100
1	D	265/281 (94%)	261 (98%)	4 (2%)	0	100	100
1	E	265/281 (94%)	260 (98%)	5 (2%)	0	100	100
1	F	265/281 (94%)	260 (98%)	5 (2%)	0	100	100
All	All	1592/1686 (94%)	1565 (98%)	27 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	185/212 (87%)	184 (100%)	1 (0%)	88	94
1	B	162/212 (76%)	161 (99%)	1 (1%)	86	92
1	C	184/212 (87%)	183 (100%)	1 (0%)	88	94
1	D	166/212 (78%)	163 (98%)	3 (2%)	59	75
1	E	149/212 (70%)	148 (99%)	1 (1%)	84	91
1	F	184/212 (87%)	183 (100%)	1 (0%)	88	94
All	All	1030/1272 (81%)	1022 (99%)	8 (1%)	81	89

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	155	GLU
1	B	155	GLU
1	C	155	GLU
1	D	61	THR
1	D	143	LYS
1	D	155	GLU
1	E	155	GLU
1	F	155	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	183	ASN
1	B	183	ASN
1	C	195	HIS

### 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 [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

6 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	CIT	E	301	-	3,12,12	1.40	0	3,17,17	2.76	1 (33%)
2	CIT	F	301	-	3,12,12	1.23	0	3,17,17	1.81	1 (33%)
2	CIT	C	301	-	3,12,12	1.00	0	3,17,17	1.95	1 (33%)
2	CIT	D	301	-	3,12,12	1.30	0	3,17,17	3.22	2 (66%)
2	CIT	B	301	-	3,12,12	1.17	0	3,17,17	2.30	2 (66%)
2	CIT	A	301	-	3,12,12	1.26	0	3,17,17	2.53	2 (66%)

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	CIT	E	301	-	-	0/6/16/16	-
2	CIT	F	301	-	-	3/6/16/16	-
2	CIT	C	301	-	-	0/6/16/16	-
2	CIT	D	301	-	-	6/6/16/16	-
2	CIT	B	301	-	-	3/6/16/16	-
2	CIT	A	301	-	-	6/6/16/16	-

There are no bond length outliers.

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	301	CIT	C3-C2-C1	-4.53	107.72	114.98
2	E	301	CIT	C3-C4-C5	-4.44	107.87	114.98
2	D	301	CIT	C3-C4-C5	-3.23	109.82	114.98
2	C	301	CIT	C3-C4-C5	-3.10	110.02	114.98
2	A	301	CIT	C3-C2-C1	-3.06	110.08	114.98
2	A	301	CIT	C3-C4-C5	-3.01	110.16	114.98
2	F	301	CIT	C3-C4-C5	-2.96	110.24	114.98
2	B	301	CIT	C3-C4-C5	-2.90	110.34	114.98
2	B	301	CIT	C3-C2-C1	-2.09	111.63	114.98

There are no chirality outliers.

All (18) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	301	CIT	C1-C2-C3-O7
2	A	301	CIT	C1-C2-C3-C6
2	A	301	CIT	C2-C3-C4-C5
2	A	301	CIT	O7-C3-C4-C5
2	A	301	CIT	C6-C3-C4-C5
2	B	301	CIT	C6-C3-C4-C5
2	D	301	CIT	C1-C2-C3-O7
2	D	301	CIT	C1-C2-C3-C4
2	D	301	CIT	C1-C2-C3-C6
2	D	301	CIT	C6-C3-C4-C5
2	F	301	CIT	C2-C3-C4-C5
2	F	301	CIT	C6-C3-C4-C5
2	A	301	CIT	C1-C2-C3-C4
2	D	301	CIT	O7-C3-C4-C5
2	F	301	CIT	O7-C3-C4-C5
2	B	301	CIT	C2-C3-C4-C5
2	B	301	CIT	O7-C3-C4-C5
2	D	301	CIT	C2-C3-C4-C5

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	F	301	CIT	1	0
2	A	301	CIT	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

### 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	267/281 (95%)	0.17	6 (2%) 62 57	47, 64, 106, 146	0
1	B	268/281 (95%)	0.56	33 (12%) 4 2	52, 97, 142, 168	0
1	C	267/281 (95%)	0.13	7 (2%) 56 52	49, 74, 109, 131	0
1	D	266/281 (94%)	0.63	39 (14%) 2 1	48, 91, 140, 178	0
1	E	267/281 (95%)	0.83	43 (16%) 1 1	73, 115, 151, 189	0
1	F	267/281 (95%)	0.10	4 (1%) 73 71	53, 76, 111, 139	0
All	All	1602/1686 (95%)	0.40	132 (8%) 11 9	47, 82, 138, 189	0

All (132) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	121	VAL	10.9
1	E	93	ALA	7.8
1	E	94	GLY	7.8
1	E	119	ILE	7.7
1	E	86	HIS	6.7
1	D	67	VAL	6.6
1	D	71	CYS	5.7
1	B	90	CYS	5.4
1	E	95	VAL	5.4
1	E	73	VAL	5.3
1	E	96	ALA	5.3
1	E	120	PRO	5.1
1	E	249	ALA	4.9
1	D	7	ILE	4.8
1	E	260	TYR	4.6
1	D	68	LEU	4.5
1	B	95	VAL	4.5
1	E	253	VAL	4.5
1	E	122	VAL	4.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	33	ALA	4.4
1	E	8	GLY	4.3
1	B	119	ILE	4.0
1	E	74	VAL	4.0
1	D	87	LEU	4.0
1	B	67	VAL	4.0
1	E	67	VAL	3.9
1	D	29	TYR	3.8
1	B	28	GLY	3.8
1	E	33	ALA	3.8
1	E	87	LEU	3.7
1	D	84	SER	3.6
1	E	71	CYS	3.5
1	B	97	ILE	3.5
1	E	254	GLY	3.5
1	E	9	ILE	3.5
1	E	97	ILE	3.5
1	E	75	ILE	3.5
1	A	104	MET	3.5
1	E	92	GLU	3.4
1	E	34	ALA	3.4
1	A	112	LEU	3.4
1	D	81	ALA	3.4
1	E	72	ASP	3.4
1	D	95	VAL	3.3
1	B	86	HIS	3.2
1	E	118	HIS	3.2
1	E	98	VAL	3.2
1	D	33	ALA	3.2
1	E	258	ARG	3.2
1	F	87	LEU	3.2
1	B	60	LEU	3.1
1	B	9	ILE	3.1
1	D	21	ILE	3.1
1	D	96	ALA	3.1
1	B	121	VAL	3.1
1	D	253	VAL	3.1
1	B	250	ALA	3.1
1	D	58	VAL	3.0
1	D	9	ILE	3.0
1	E	90	CYS	3.0
1	A	119	ILE	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	74	VAL	3.0
1	D	24	VAL	2.9
1	D	86	HIS	2.9
1	A	121	VAL	2.9
1	E	124	ALA	2.8
1	B	38	PRO	2.8
1	D	31	LEU	2.7
1	E	32	GLY	2.7
1	D	82	ALA	2.6
1	B	55	SER	2.6
1	B	254	GLY	2.6
1	B	116	ALA	2.6
1	E	103	GLY	2.6
1	D	79	THR	2.6
1	B	31	LEU	2.5
1	B	34	ALA	2.5
1	B	83	THR	2.5
1	B	235	ALA	2.5
1	E	266	LEU	2.5
1	C	16	MET	2.5
1	B	232	THR	2.5
1	D	92	GLU	2.4
1	D	32	GLY	2.4
1	C	12	ALA	2.4
1	C	77	PHE	2.4
1	C	87	LEU	2.4
1	D	110	ALA	2.4
1	D	72	ASP	2.4
1	B	260	TYR	2.4
1	C	93	ALA	2.3
1	B	29	TYR	2.3
1	B	216	GLY	2.3
1	D	219	THR	2.3
1	B	53	ILE	2.3
1	B	99	ILE	2.3
1	E	68	LEU	2.3
1	B	219	THR	2.3
1	B	71	CYS	2.3
1	D	5	PRO	2.2
1	E	70	ASP	2.2
1	E	77	PHE	2.2
1	D	63	SER	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	10	LEU	2.2
1	F	119	ILE	2.2
1	C	64	LEU	2.2
1	E	57	GLY	2.2
1	D	53	ILE	2.2
1	D	78	SER	2.2
1	F	95	VAL	2.2
1	F	97	ILE	2.2
1	B	112	LEU	2.2
1	D	70	ASP	2.1
1	E	250	ALA	2.1
1	E	38	PRO	2.1
1	A	219	THR	2.1
1	D	34	ALA	2.1
1	B	233	HIS	2.1
1	C	112	LEU	2.1
1	D	80	PRO	2.1
1	D	90	CYS	2.1
1	E	104	MET	2.1
1	D	73	VAL	2.1
1	B	32	GLY	2.1
1	D	99	ILE	2.0
1	A	123	TYR	2.0
1	D	251	TRP	2.0
1	E	265	VAL	2.0
1	D	8	GLY	2.0
1	E	251	TRP	2.0
1	B	96	ALA	2.0
1	B	215	VAL	2.0

## 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 [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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
2	CIT	B	301	13/13	0.76	0.41	68,72,76,78	13
2	CIT	C	301	13/13	0.80	0.20	71,72,78,81	13
2	CIT	A	301	13/13	0.81	0.19	65,70,72,79	13
2	CIT	E	301	13/13	0.81	0.20	96,99,102,107	13
2	CIT	F	301	13/13	0.84	0.15	84,88,91,91	13
2	CIT	D	301	13/13	0.86	0.16	60,67,71,72	13

## 6.5 Other polymers [i](#)

There are no such residues in this entry.