



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 17, 2018 – 09:09 am GMT

PDB ID : 1EFK  
Title : STRUCTURE OF HUMAN MALIC ENZYME IN COMPLEX WITH KETOMALONATE  
Authors : Yang, Z.; Floyd, D.L.; Loeber, G.; Tong, L.  
Deposited on : 2000-02-09  
Resolution : 2.60 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<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 : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30686

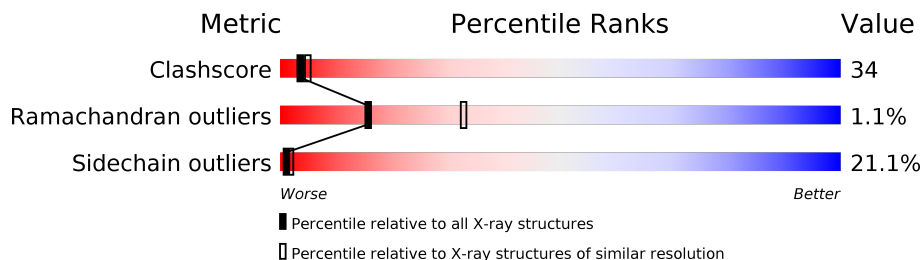
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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))
Clashscore	122078	3109 (2.60-2.60)
Ramachandran outliers	120005	3061 (2.60-2.60)
Sidechain outliers	119972	3061 (2.60-2.60)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	584	
1	B	584	
1	C	584	
1	D	584	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 17931 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 MALIC ENZYME.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	553	Total	C	N	O	S	Se	0	0	0
			4367	2796	744	804	9	14			
1	B	553	Total	C	N	O	S	Se	0	0	0
			4367	2796	744	804	9	14			
1	C	553	Total	C	N	O	S	Se	0	0	0
			4367	2796	744	804	9	14			
1	D	553	Total	C	N	O	S	Se	0	0	0
			4367	2796	744	804	9	14			

There are 56 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	29	MSE	MET	MODIFIED RESIDUE	UNP P23368
A	38	MSE	MET	MODIFIED RESIDUE	UNP P23368
A	47	MSE	MET	MODIFIED RESIDUE	UNP P23368
A	75	MSE	MET	MODIFIED RESIDUE	UNP P23368
A	86	MSE	MET	MODIFIED RESIDUE	UNP P23368
A	108	MSE	MET	MODIFIED RESIDUE	UNP P23368
A	177	MSE	MET	MODIFIED RESIDUE	UNP P23368
A	219	MSE	MET	MODIFIED RESIDUE	UNP P23368
A	239	MSE	MET	MODIFIED RESIDUE	UNP P23368
A	325	MSE	MET	MODIFIED RESIDUE	UNP P23368
A	327	MSE	MET	MODIFIED RESIDUE	UNP P23368
A	343	MSE	MET	MODIFIED RESIDUE	UNP P23368
A	407	MSE	MET	MODIFIED RESIDUE	UNP P23368
A	539	MSE	MET	MODIFIED RESIDUE	UNP P23368
B	29	MSE	MET	MODIFIED RESIDUE	UNP P23368
B	38	MSE	MET	MODIFIED RESIDUE	UNP P23368
B	47	MSE	MET	MODIFIED RESIDUE	UNP P23368
B	75	MSE	MET	MODIFIED RESIDUE	UNP P23368
B	86	MSE	MET	MODIFIED RESIDUE	UNP P23368
B	108	MSE	MET	MODIFIED RESIDUE	UNP P23368
B	177	MSE	MET	MODIFIED RESIDUE	UNP P23368

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	219	MSE	MET	MODIFIED RESIDUE	UNP P23368
B	239	MSE	MET	MODIFIED RESIDUE	UNP P23368
B	325	MSE	MET	MODIFIED RESIDUE	UNP P23368
B	327	MSE	MET	MODIFIED RESIDUE	UNP P23368
B	343	MSE	MET	MODIFIED RESIDUE	UNP P23368
B	407	MSE	MET	MODIFIED RESIDUE	UNP P23368
B	539	MSE	MET	MODIFIED RESIDUE	UNP P23368
C	29	MSE	MET	MODIFIED RESIDUE	UNP P23368
C	38	MSE	MET	MODIFIED RESIDUE	UNP P23368
C	47	MSE	MET	MODIFIED RESIDUE	UNP P23368
C	75	MSE	MET	MODIFIED RESIDUE	UNP P23368
C	86	MSE	MET	MODIFIED RESIDUE	UNP P23368
C	108	MSE	MET	MODIFIED RESIDUE	UNP P23368
C	177	MSE	MET	MODIFIED RESIDUE	UNP P23368
C	219	MSE	MET	MODIFIED RESIDUE	UNP P23368
C	239	MSE	MET	MODIFIED RESIDUE	UNP P23368
C	325	MSE	MET	MODIFIED RESIDUE	UNP P23368
C	327	MSE	MET	MODIFIED RESIDUE	UNP P23368
C	343	MSE	MET	MODIFIED RESIDUE	UNP P23368
C	407	MSE	MET	MODIFIED RESIDUE	UNP P23368
C	539	MSE	MET	MODIFIED RESIDUE	UNP P23368
D	29	MSE	MET	MODIFIED RESIDUE	UNP P23368
D	38	MSE	MET	MODIFIED RESIDUE	UNP P23368
D	47	MSE	MET	MODIFIED RESIDUE	UNP P23368
D	75	MSE	MET	MODIFIED RESIDUE	UNP P23368
D	86	MSE	MET	MODIFIED RESIDUE	UNP P23368
D	108	MSE	MET	MODIFIED RESIDUE	UNP P23368
D	177	MSE	MET	MODIFIED RESIDUE	UNP P23368
D	219	MSE	MET	MODIFIED RESIDUE	UNP P23368
D	239	MSE	MET	MODIFIED RESIDUE	UNP P23368
D	325	MSE	MET	MODIFIED RESIDUE	UNP P23368
D	327	MSE	MET	MODIFIED RESIDUE	UNP P23368
D	343	MSE	MET	MODIFIED RESIDUE	UNP P23368
D	407	MSE	MET	MODIFIED RESIDUE	UNP P23368
D	539	MSE	MET	MODIFIED RESIDUE	UNP P23368

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

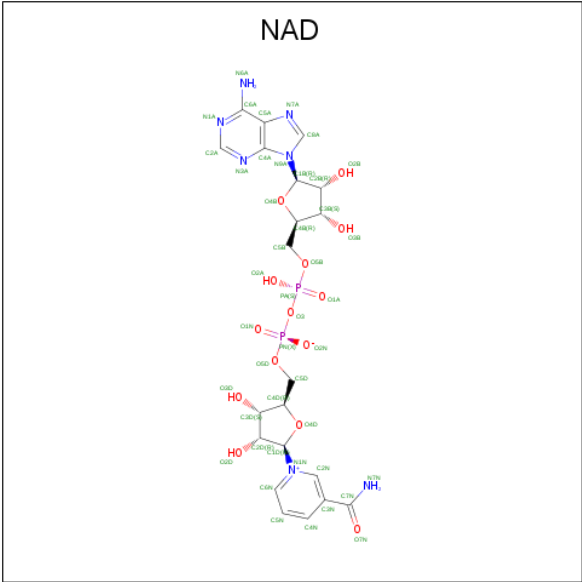
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total Mg 1 1	0	0
2	A	1	Total Mg 1 1	0	0

*Continued on next page...*

Continued from previous page...

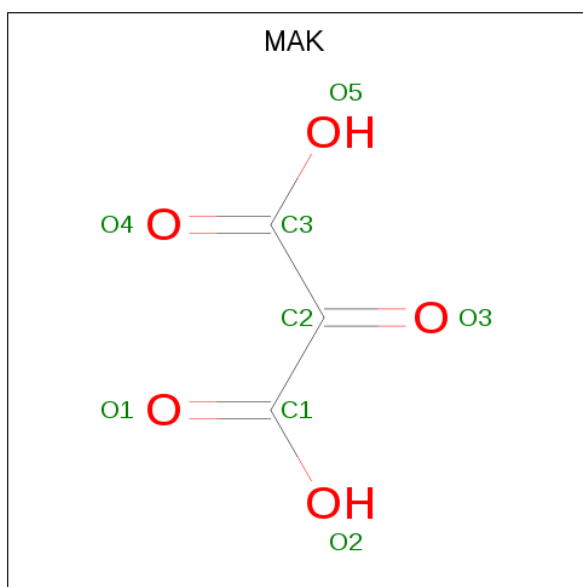
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	D	1	Total	Mg	0	0
			1	1		
2	C	1	Total	Mg	0	0
			1	1		

- Molecule 3 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
3	A	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
3	A	1	Total	C	N	O	P	9	0
			44	21	7	14	2		
3	B	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
3	B	1	Total	C	N	O	P	9	0
			44	21	7	14	2		
3	C	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
3	C	1	Total	C	N	O	P	9	0
			44	21	7	14	2		
3	D	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
3	D	1	Total	C	N	O	P	9	0
			44	21	7	14	2		

- Molecule 4 is ALPHA-KETOMALONIC ACID (three-letter code: MAK) (formula: C<sub>3</sub>H<sub>2</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			8	3	5		
4	B	1	Total	C	O	0	0
			8	3	5		
4	C	1	Total	C	O	0	0
			8	3	5		
4	D	1	Total	C	O	0	0
			8	3	5		

- Molecule 5 is water.

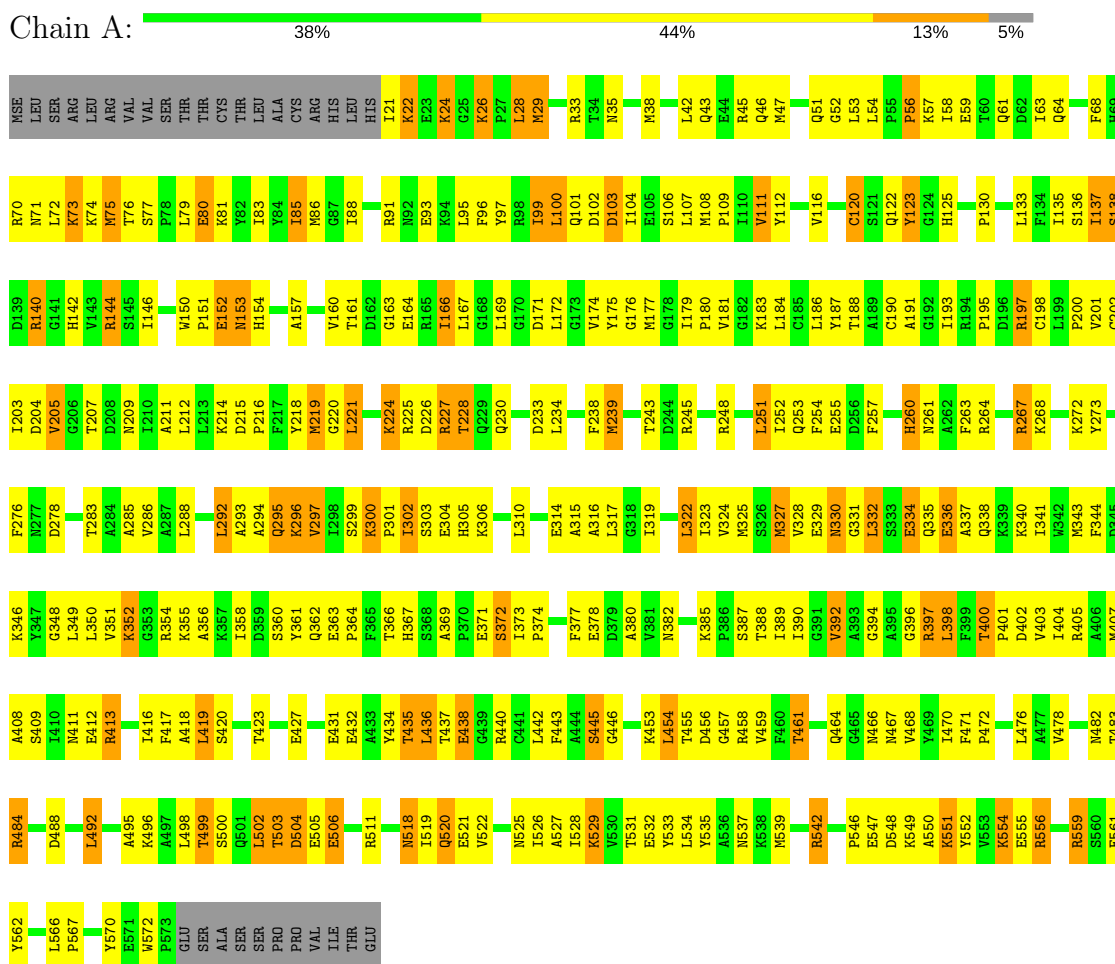
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	25	Total	O	0	0
			25	25		
5	B	18	Total	O	0	0
			18	18		
5	C	16	Total	O	0	0
			16	16		
5	D	16	Total	O	0	0
			16	16		

### 3 Residue-property plots

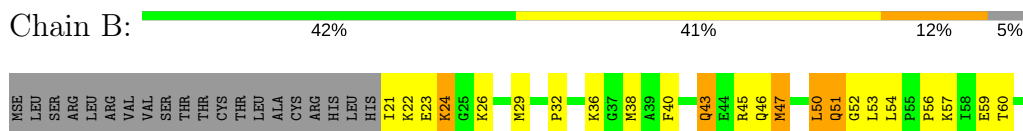
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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.

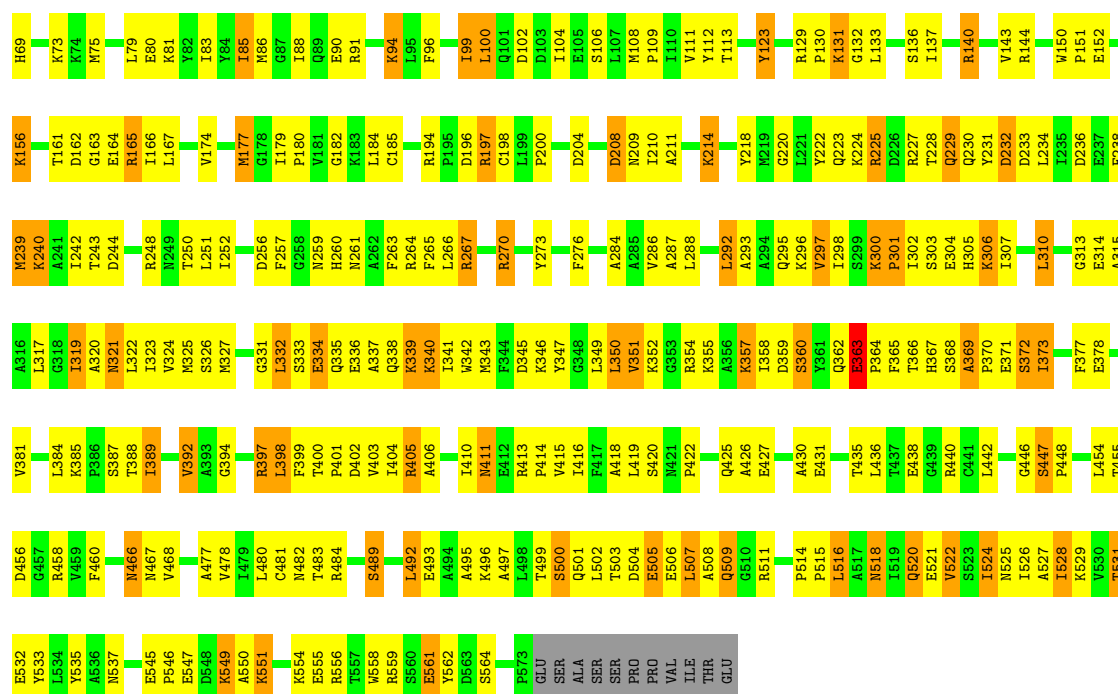
Note EDS was not executed.

#### • Molecule 1: MALIC ENZYME



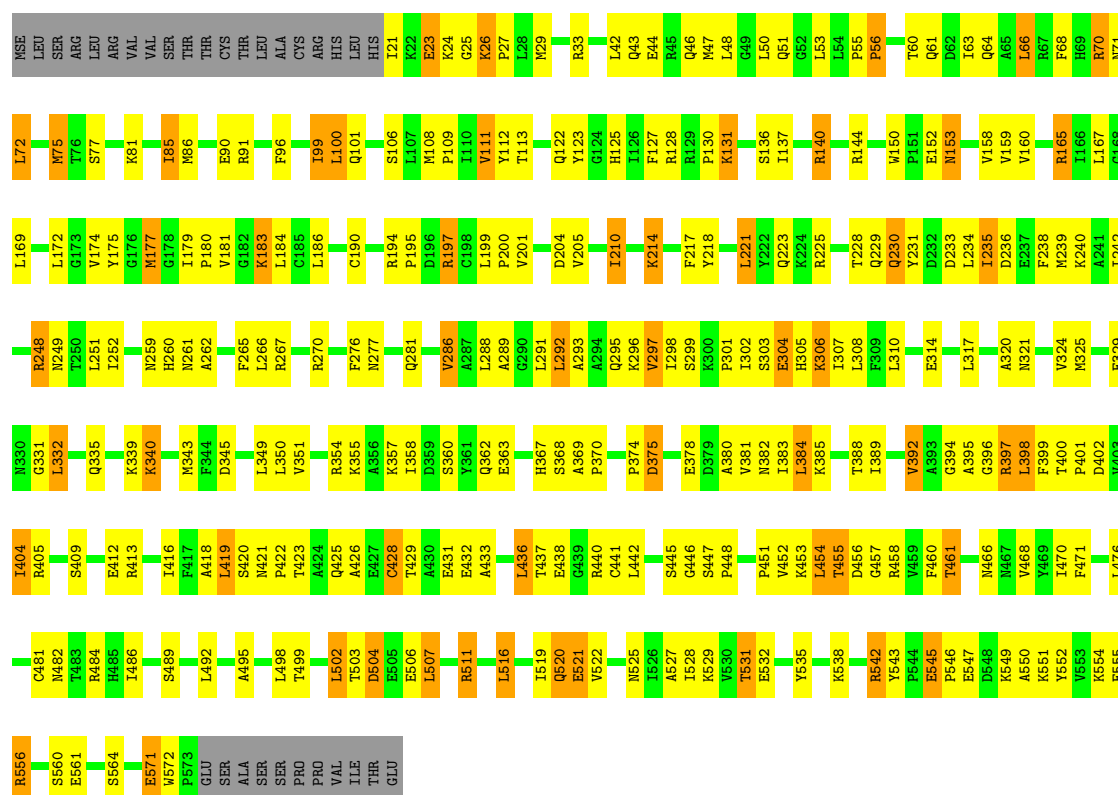
#### • Molecule 1: MALIC ENZYME





### • Molecule 1: MALIC ENZYME

Chain C: 47% 38% 9% 5%



### • Molecule 1: MALIC ENZYME

Chain D: 41% 39% 14% 5%



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	229.60Å 118.60Å 113.10Å 90.00° 109.60° 90.00°	Depositor
Resolution (Å)	20.00 – 2.60	Depositor
% Data completeness (in resolution range)	94.0 (20.00-2.60)	Depositor
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	X-PLOR 3.851	Depositor
R, $R_{free}$	0.218 , 0.301	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	17931	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

## 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: MG, MAK, NAD

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	0.44	0/4447	0.64	0/5998
1	B	0.45	0/4447	0.64	0/5998
1	C	0.46	0/4447	0.65	0/5998
1	D	0.45	1/4447 (0.0%)	0.64	0/5998
All	All	0.45	1/17788 (0.0%)	0.64	0/23992

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	185	CYS	CB-SG	-5.57	1.72	1.81

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	4367	0	4407	359	0
1	B	4367	0	4407	313	0
1	C	4367	0	4407	217	0
1	D	4367	0	4407	334	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	88	0	52	4	0
3	B	88	0	52	5	0
3	C	88	0	52	3	0
3	D	88	0	52	3	0
4	A	8	0	0	1	0
4	B	8	0	0	1	0
4	C	8	0	0	1	0
4	D	8	0	0	2	0
5	A	25	0	0	4	0
5	B	18	0	0	4	0
5	C	16	0	0	3	0
5	D	16	0	0	1	0
All	All	17931	0	17836	1198	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 34.

The worst 5 of 1198 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:C:210:ILE:H	1:C:210:ILE:HD13	1.00	1.10
1:D:140:ARG:HH22	1:D:233:ASP:HB3	1.15	1.09
1:D:177:MSE:HE1	1:D:181:VAL:HG23	1.34	1.06
1:A:327:MSE:HE3	1:A:337:ALA:HB1	1.38	1.06
1:B:354:ARG:HE	1:B:358:ILE:HD11	1.21	1.06

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	551/584 (94%)	512 (93%)	34 (6%)	5 (1%)	19	38
1	B	551/584 (94%)	512 (93%)	32 (6%)	7 (1%)	13	27
1	C	551/584 (94%)	513 (93%)	34 (6%)	4 (1%)	24	46
1	D	551/584 (94%)	499 (91%)	43 (8%)	9 (2%)	11	21
All	All	2204/2336 (94%)	2036 (92%)	143 (6%)	25 (1%)	16	33

5 of 25 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	332	LEU
1	B	332	LEU
1	C	332	LEU
1	D	270	ARG
1	D	332	LEU

### 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	469/483 (97%)	371 (79%)	98 (21%)	1	2
1	B	469/483 (97%)	374 (80%)	95 (20%)	1	2
1	C	469/483 (97%)	380 (81%)	89 (19%)	1	2
1	D	469/483 (97%)	356 (76%)	113 (24%)	1	1
All	All	1876/1932 (97%)	1481 (79%)	395 (21%)	1	2

5 of 395 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	511	ARG
1	C	223	GLN
1	D	445	SER
1	B	522	VAL
1	C	75	MSE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 52 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	518	ASN
1	C	229	GLN
1	D	425	GLN
1	B	520	GLN
1	C	69	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 ⓘ

There are no carbohydrates in this entry.

### 5.6 Ligand geometry ⓘ

Of 16 ligands modelled in this entry, 4 are monoatomic - leaving 12 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
3	NAD	A	601	-	40,48,48	1.75	9 (22%)	44,73,73	1.75	3 (6%)
3	NAD	A	602	-	40,48,48	1.79	8 (20%)	44,73,73	1.77	3 (6%)
4	MAK	A	603	2	1,7,7	3.27	1 (100%)	0,9,9	0.00	-
3	NAD	B	1601	-	40,48,48	1.64	7 (17%)	44,73,73	1.91	4 (9%)
3	NAD	B	1602	-	40,48,48	1.86	10 (25%)	44,73,73	1.74	3 (6%)
4	MAK	B	1603	2	1,7,7	3.13	1 (100%)	0,9,9	0.00	-

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	NAD	C	2601	-	40,48,48	1.65	8 (20%)	44,73,73	1.82	6 (13%)
3	NAD	C	2602	-	40,48,48	1.94	10 (25%)	44,73,73	1.74	3 (6%)
4	MAK	C	2603	2	1,7,7	3.15	1 (100%)	0,9,9	0.00	-
3	NAD	D	3601	-	40,48,48	1.86	8 (20%)	44,73,73	1.82	3 (6%)
3	NAD	D	3602	-	40,48,48	1.73	8 (20%)	44,73,73	1.81	4 (9%)
4	MAK	D	3603	2	1,7,7	3.09	1 (100%)	0,9,9	0.00	-

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
3	NAD	A	601	-	-	0/22/62/62	0/5/5/5
3	NAD	A	602	-	-	0/22/62/62	0/5/5/5
4	MAK	A	603	2	-	0/0/8/8	0/0/0/0
3	NAD	B	1601	-	-	0/22/62/62	0/5/5/5
3	NAD	B	1602	-	-	0/22/62/62	0/5/5/5
4	MAK	B	1603	2	-	0/0/8/8	0/0/0/0
3	NAD	C	2601	-	-	0/22/62/62	0/5/5/5
3	NAD	C	2602	-	-	0/22/62/62	0/5/5/5
4	MAK	C	2603	2	-	0/0/8/8	0/0/0/0
3	NAD	D	3601	-	-	0/22/62/62	0/5/5/5
3	NAD	D	3602	-	-	0/22/62/62	0/5/5/5
4	MAK	D	3603	2	-	0/0/8/8	0/0/0/0

The worst 5 of 72 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	3601	NAD	C5A-C4A	-3.33	1.33	1.40
3	B	1601	NAD	C5A-C4A	-3.18	1.33	1.40
3	C	2601	NAD	C5A-C4A	-3.18	1.33	1.40
3	A	601	NAD	C5A-C4A	-2.98	1.33	1.40
3	D	3602	NAD	C5A-C4A	-2.88	1.34	1.40

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	1601	NAD	N3A-C2A-N1A	-9.96	120.34	128.86
3	D	3602	NAD	N3A-C2A-N1A	-9.53	120.71	128.86
3	D	3601	NAD	N3A-C2A-N1A	-9.47	120.76	128.86

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	2601	NAD	N3A-C2A-N1A	-9.26	120.94	128.86
3	A	602	NAD	N3A-C2A-N1A	-9.23	120.96	128.86

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

10 monomers are involved in 20 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	601	NAD	3	0
3	A	602	NAD	1	0
4	A	603	MAK	1	0
3	B	1601	NAD	4	0
3	B	1602	NAD	1	0
4	B	1603	MAK	1	0
3	C	2601	NAD	3	0
4	C	2603	MAK	1	0
3	D	3601	NAD	3	0
4	D	3603	MAK	2	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

EDS was not executed - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section is therefore empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section is therefore empty.