



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 02:04 am BST

PDB ID : 2O2K  
Title : Crystal Structure of the Activation Domain of Human Methionine Synthase Isoform/Mutant D963E/K1071N  
Authors : Wolthers, K.R.; Toogood, H.S.; Jowitt, T.A.; Marshall, K.R.; Leys, D.; Scrutton, N.S.  
Deposited on : 2006-11-30  
Resolution : 1.60 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
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

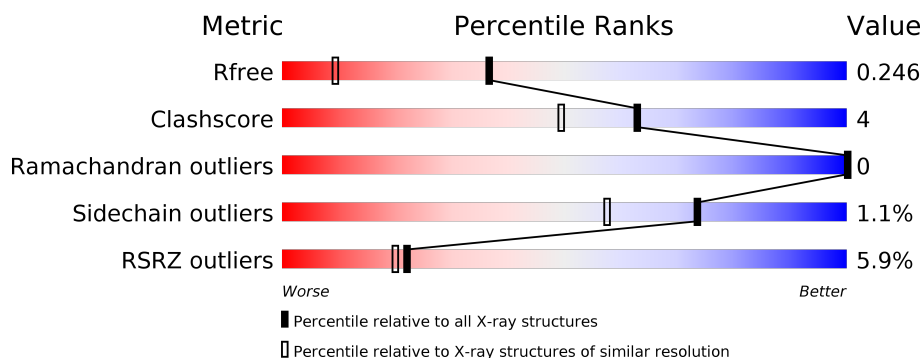
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.60 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))
$R_{free}$	130704	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)
RSRZ outliers	127900	3321 (1.60-1.60)

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 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	355	<div> <div>6%</div> <div> <div></div> <div>86%</div> <div>8%</div> <div>6%</div> </div> </div>
1	B	355	<div> <div>5%</div> <div> <div></div> <div>85%</div> <div>9%</div> <div>6%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6655 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 Methionine synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	332	Total	C	N	O	S	0	16	0
			2705	1740	450	506	9			
1	B	333	Total	C	N	O	S	0	26	0
			2755	1764	465	517	9			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	924	MET	-	INITIATING METHIONINE	UNP Q99707
A	963	GLU	ASP	SEE REMARK 999	UNP Q99707
A	1071	ASN	LYS	SEE REMARK 999	UNP Q99707
A	1266	LYS	-	EXPRESSION TAG	UNP Q99707
A	1267	LEU	-	EXPRESSION TAG	UNP Q99707
A	1268	ALA	-	EXPRESSION TAG	UNP Q99707
A	1269	ALA	-	EXPRESSION TAG	UNP Q99707
A	1270	ALA	-	EXPRESSION TAG	UNP Q99707
A	1271	LEU	-	EXPRESSION TAG	UNP Q99707
A	1272	GLU	-	EXPRESSION TAG	UNP Q99707
A	1273	HIS	-	EXPRESSION TAG	UNP Q99707
A	1274	HIS	-	EXPRESSION TAG	UNP Q99707
A	1275	HIS	-	EXPRESSION TAG	UNP Q99707
A	1276	HIS	-	EXPRESSION TAG	UNP Q99707
A	1277	HIS	-	EXPRESSION TAG	UNP Q99707
A	1278	HIS	-	EXPRESSION TAG	UNP Q99707
B	924	MET	-	INITIATING METHIONINE	UNP Q99707
B	963	GLU	ASP	SEE REMARK 999	UNP Q99707
B	1071	ASN	LYS	SEE REMARK 999	UNP Q99707
B	1266	LYS	-	EXPRESSION TAG	UNP Q99707
B	1267	LEU	-	EXPRESSION TAG	UNP Q99707
B	1268	ALA	-	EXPRESSION TAG	UNP Q99707
B	1269	ALA	-	EXPRESSION TAG	UNP Q99707
B	1270	ALA	-	EXPRESSION TAG	UNP Q99707
B	1271	LEU	-	EXPRESSION TAG	UNP Q99707

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	1272	GLU	-	EXPRESSION TAG	UNP Q99707
B	1273	HIS	-	EXPRESSION TAG	UNP Q99707
B	1274	HIS	-	EXPRESSION TAG	UNP Q99707
B	1275	HIS	-	EXPRESSION TAG	UNP Q99707
B	1276	HIS	-	EXPRESSION TAG	UNP Q99707
B	1277	HIS	-	EXPRESSION TAG	UNP Q99707
B	1278	HIS	-	EXPRESSION TAG	UNP Q99707

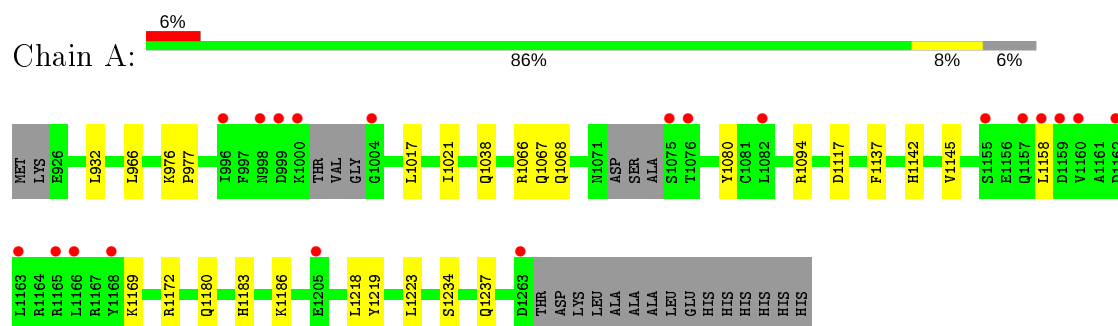
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	638	Total O 638 638	0	0
2	B	557	Total O 557 557	0	0

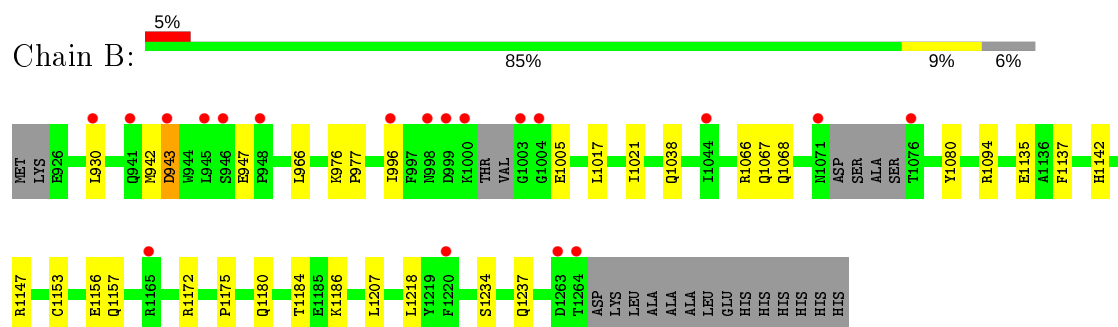
### 3 Residue-property plots [i](#)

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 ( $\text{RSRZ} > 2$ ). 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: Methionine synthase



#### • Molecule 1: Methionine synthase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	77.85Å 90.05Å 123.01Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.83 – 1.60 29.16 – 1.60	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.83-1.60) 95.9 (29.16-1.60)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.66 (at 1.60Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.209 , 0.248 0.210 , 0.246	Depositor DCC
$R_{free}$ test set	5481 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	21.7	Xtriage
Anisotropy	0.208	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 41.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	6655	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	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 42.41 % 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 2.0475e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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](#)

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	0.35	0/2829	0.52	0/3837
1	B	0.36	0/2920	0.53	0/3962
All	All	0.36	0/5749	0.53	0/7799

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	2705	0	2554	23	0
1	B	2755	0	2589	23	0
2	A	638	0	0	2	0
2	B	557	0	0	2	0
All	All	6655	0	5143	46	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 4.

The worst 5 of 46 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:1066:ARG:H	1:A:1237:GLN:HE22	1.08	0.95

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:1066:ARG:H	1:B:1237:GLN:HE22	1.10	0.92
1:A:1017:LEU:O	1:A:1021[A]:ILE:HG12	1.81	0.81
1:A:1142:HIS:HD2	1:A:1172:ARG:HH11	1.37	0.72
1:B:1038:GLN:HE22	1:B:1094:ARG:HA	1.56	0.69

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	342/355 (96%)	332 (97%)	10 (3%)	0	100	100
1	B	353/355 (99%)	342 (97%)	11 (3%)	0	100	100
All	All	695/710 (98%)	674 (97%)	21 (3%)	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	280/298 (94%)	276 (99%)	4 (1%)	67	47
1	B	289/298 (97%)	286 (99%)	3 (1%)	76	61
All	All	569/596 (96%)	562 (99%)	7 (1%)	73	54



5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1180	GLN
1	B	1180	GLN
1	B	943	ASP
1	A	1117[B]	ASP
1	B	1137	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1183	HIS
1	B	941	GLN
1	B	1068	GLN
1	A	1142	HIS
1	B	1142	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 ⓘ

There are no ligands in this entry.

## 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 [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, 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	332/355 (93%)	0.14	20 (6%) 21 19	14, 21, 34, 43	3 (0%)
1	B	333/355 (93%)	0.19	19 (5%) 23 21	14, 22, 34, 42	0
All	All	665/710 (93%)	0.17	39 (5%) 22 20	14, 21, 34, 43	3 (0%)

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1075	SER	7.9
1	B	1264	THR	6.0
1	B	1071	ASN	4.8
1	B	945	LEU	4.7
1	B	1076	THR	4.2

### 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.