



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 5, 2017 – 05:21 AM EDT

PDB ID : 5CKY  
Title : Crystal Structure of the MTERF1 R162A substitution bound to the termination sequence.  
Authors : Byrnes, J.; Hauser, K.; Norona, L.; Mejia, E.; Simmerling, C.; Garcia-Diaz, M.  
Deposited on : unknown  
Resolution : 2.62 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

<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
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20029824
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	rb-20029824

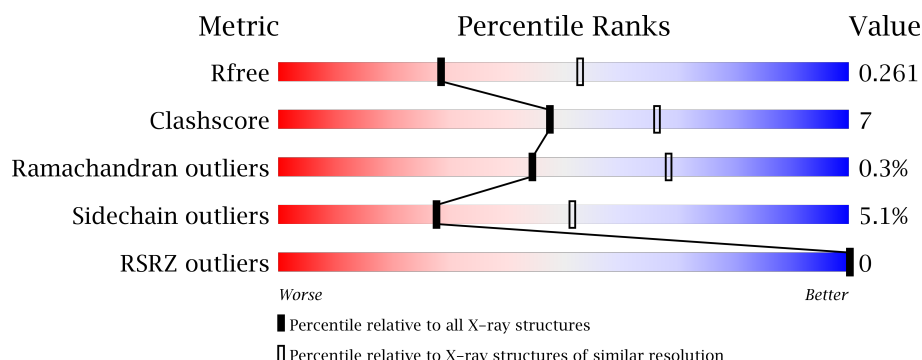
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.62 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))
$R_{free}$	100719	2983 (2.64-2.60)
Clashscore	112137	3351 (2.64-2.60)
Ramachandran outliers	110173	3298 (2.64-2.60)
Sidechain outliers	110143	3298 (2.64-2.60)
RSRZ outliers	101464	2992 (2.64-2.60)

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.

Mol	Chain	Length	Quality of chain
1	O	324	<div> <div>85%</div> <div>12%</div> <div>.</div> </div>
2	D	22	<div> <div>55%</div> <div>27%</div> <div>18%</div> </div>
3	E	22	<div> <div>64%</div> <div>36%</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3501 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 Transcription termination factor 1, mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	O	324	Total	C	N	O	S	0	0	0
			2597	1655	451	479	12			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
O	162	ALA	ARG	engineered mutation	UNP B4DPR9

- Molecule 2 is a DNA chain called 5'-D(\*AP\*TP\*TP\*AP\*CP\*CP\*GP\*GP\*GP\*CP\*TP\*C  
P\*TP\*GP\*CP\*CP\*AP\*TP\*CP\*TP\*TP\*A)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	22	Total	C	N	O	P	0	0	0
			442	213	75	133	21			

- Molecule 3 is a DNA chain called 5' -D (\*TP\*AP\*AP\*GP\*AP\*TP\*GP\*GP\*CP\*AP\*GP\*  
AP\*GP\*CP\*CP\*CP\*GP\*GP\*TP\*AP\*AP\*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	22	Total	C	N	O	P	0	0	0
			454	216	90	127	21			

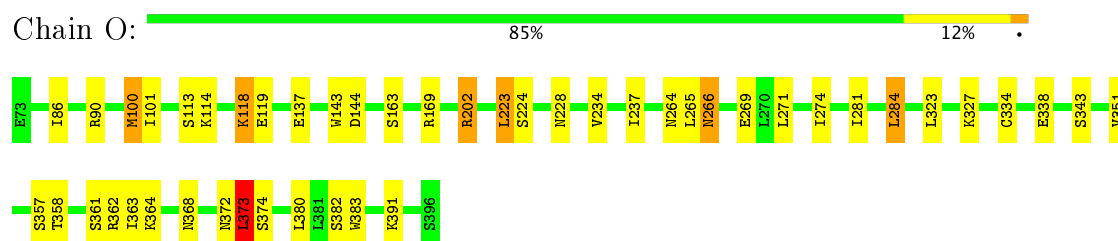
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	O	7	Total	O	0	0
			7	7		
4	D	1	Total	O	0	0
			1	1		

### 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 ( $RSRZ > 2$ ). 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: Transcription termination factor 1, mitochondrial



- Molecule 2: 5'-D(\*AP\*TP\*TP\*AP\*CP\*CP\*GP\*GP\*GP\*CP\*TP\*CP\*TP\*GP\*CP\*CP\*AP\*T  
P\*CP\*TP\*TP\*A)-3'



- Molecule 3: 5' -D (\*TP\*AP\*AP\*GP\*AP\*TP\*GP\*GP\*CP\*AP\*GP\*AP\*GP\*CP\*CP\*CP\*GP\*  
GP\*TP\*AP\*AP\*T)-3'



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	88.22Å 91.16Å 159.10Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	79.55 – 2.62 79.55 – 2.62	Depositor EDS
% Data completeness (in resolution range)	99.9 (79.55-2.62) 99.9 (79.55-2.62)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.24 (at 2.62Å)	Xtriage
Refinement program	REFMAC 5.8.0071	Depositor
R, $R_{free}$	0.214 , 0.259 0.219 , 0.261	Depositor DCC
$R_{free}$ test set	1008 reflections (5.41%)	DCC
Wilson B-factor (Å <sup>2</sup> )	54.1	Xtriage
Anisotropy	0.038	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 30.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.033 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	3501	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.75% 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

### 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	O	0.68	0/2636	0.87	6/3551 (0.2%)
2	D	0.84	3/493 (0.6%)	1.08	3/758 (0.4%)
3	E	0.62	0/511	1.06	5/788 (0.6%)
All	All	0.70	3/3640 (0.1%)	0.93	14/5097 (0.3%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	12	DC	O3'-P	-9.19	1.50	1.61
2	D	13	DT	O3'-P	-6.54	1.53	1.61
2	D	20	DT	O3'-P	-5.09	1.55	1.61

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	O	202	ARG	CG-CD-NE	-9.78	91.27	111.80
1	O	169	ARG	NE-CZ-NH2	-8.11	116.25	120.30
2	D	15	DC	O5'-P-OP1	7.60	119.82	110.70
3	E	6	DT	O5'-P-OP1	-7.28	99.15	105.70
3	E	21	DA	O5'-P-OP1	-6.88	99.51	105.70
1	O	169	ARG	NE-CZ-NH1	6.64	123.62	120.30
2	D	11	DT	C1'-O4'-C4'	-6.41	103.69	110.10
3	E	9	DC	O5'-P-OP1	6.31	118.27	110.70
3	E	14	DC	O5'-P-OP2	-6.26	100.06	105.70
1	O	144	ASP	CB-CG-OD1	6.12	123.81	118.30
2	D	11	DT	O5'-P-OP2	-5.81	100.47	105.70
3	E	9	DC	O5'-P-OP2	-5.52	100.73	105.70
1	O	100	MET	CB-CG-SD	5.48	128.85	112.40
1	O	202	ARG	NE-CZ-NH2	-5.27	117.66	120.30

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	O	2597	0	2695	35	0
2	D	442	0	251	10	0
3	E	454	0	248	7	0
4	D	1	0	0	1	0
4	O	7	0	0	2	0
All	All	3501	0	3194	47	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 7.

All (47) 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)
3:E:3:DA:H2''	3:E:4:DG:O5'	1.65	0.96
1:O:266:ASN:H	1:O:266:ASN:HD22	1.24	0.85
1:O:391:LYS:NZ	2:D:3:DT:OP1	2.10	0.85
1:O:264:ASN:CB	1:O:265:LEU:HA	2.11	0.80
1:O:266:ASN:ND2	1:O:269:GLU:OE1	2.14	0.79
1:O:372:ASN:OD1	1:O:374:SER:HB3	1.84	0.76
1:O:266:ASN:HD21	1:O:269:GLU:CD	1.89	0.75
3:E:3:DA:H2'	3:E:4:DG:C8	2.23	0.74
1:O:266:ASN:HD22	1:O:266:ASN:N	1.86	0.72
1:O:372:ASN:O	1:O:373:LEU:HB2	1.91	0.71
2:D:12:DC:H2''	2:D:13:DT:OP1	1.92	0.69
1:O:264:ASN:HB2	1:O:265:LEU:HD12	1.74	0.68
1:O:338:GLU:HG3	1:O:363:ILE:HG21	1.76	0.67
1:O:264:ASN:HB2	1:O:265:LEU:HA	1.78	0.64
1:O:86:ILE:O	1:O:90:ARG:HG3	1.99	0.63
1:O:364:LYS:O	1:O:368:ASN:HB2	1.98	0.62
2:D:19:DC:H4'	2:D:20:DT:OP1	2.01	0.60
3:E:3:DA:H2'	3:E:4:DG:H8	1.66	0.60
2:D:19:DC:H2'	2:D:20:DT:C6	2.37	0.60
1:O:343:SER:HB2	4:O:404:HOH:O	2.03	0.59
1:O:234:VAL:O	1:O:237:ILE:HG22	2.03	0.58
1:O:264:ASN:HB3	1:O:265:LEU:HA	1.84	0.57

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:18:DT:H2''	2:D:19:DC:OP2	2.06	0.56
1:O:163:SER:HB3	2:D:13:DT:OP1	2.05	0.56
1:O:266:ASN:N	1:O:266:ASN:ND2	2.54	0.56
1:O:351:VAL:HA	1:O:382:SER:HB3	1.87	0.55
3:E:3:DA:H2''	3:E:4:DG:C5'	2.35	0.55
1:O:114:LYS:HD3	1:O:143:TRP:CE2	2.43	0.54
2:D:19:DC:H2''	2:D:20:DT:O5'	2.07	0.54
1:O:264:ASN:HB2	1:O:265:LEU:CD1	2.38	0.53
1:O:264:ASN:CB	1:O:265:LEU:CA	2.87	0.52
2:D:15:DC:H2'	2:D:16:DC:C6	2.45	0.51
1:O:163:SER:HB3	2:D:13:DT:H5''	1.92	0.51
1:O:266:ASN:H	1:O:266:ASN:ND2	2.01	0.50
1:O:264:ASN:H	1:O:265:LEU:HB2	1.77	0.48
1:O:202:ARG:HB2	4:D:101:HOH:O	2.12	0.48
1:O:223:LEU:HD12	1:O:224:SER:N	2.30	0.47
3:E:3:DA:C2'	3:E:4:DG:O5'	2.52	0.47
1:O:380:LEU:HA	1:O:383:TRP:CD2	2.50	0.46
1:O:327:LYS:NZ	3:E:13:DG:OP1	2.49	0.44
2:D:4:DA:C2	3:E:20:DA:C2	3.07	0.43
1:O:343:SER:CB	4:O:404:HOH:O	2.61	0.43
1:O:281:ILE:HA	1:O:284:LEU:HD22	2.02	0.41
1:O:358:THR:O	1:O:362:ARG:HG3	2.20	0.41
1:O:372:ASN:O	1:O:373:LEU:CB	2.64	0.41
1:O:118:LYS:HE3	1:O:119:GLU:OE2	2.20	0.40
1:O:228:ASN:OD1	1:O:228:ASN:N	2.55	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	O	322/324 (99%)	311 (97%)	10 (3%)	1 (0%)	44 68

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	O	373	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	O	296/298 (99%)	281 (95%)	15 (5%)	28	51

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	O	100	MET
1	O	101	ILE
1	O	113	SER
1	O	118	LYS
1	O	137	GLU
1	O	223	LEU
1	O	266	ASN
1	O	271	LEU
1	O	274	ILE
1	O	284	LEU
1	O	323	LEU
1	O	334	CYS
1	O	357	SER
1	O	361	SER
1	O	373	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	O	266	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, 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	O	324/324 (100%)	-0.11	0 <a href="#">100</a> <a href="#">100</a>	31, 50, 81, 99	1 (0%)
2	D	22/22 (100%)	-0.24	0 <a href="#">100</a> <a href="#">100</a>	39, 50, 67, 84	0
3	E	22/22 (100%)	-0.32	0 <a href="#">100</a> <a href="#">100</a>	39, 48, 65, 91	0
All	All	368/368 (100%)	-0.13	0 <a href="#">100</a> <a href="#">100</a>	31, 50, 81, 99	1 (0%)

There are no RSRZ outliers to report.

### 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.