



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 05:59 AM GMT

PDB ID : 2VJP  
Title : FORMYL-COA TRANSFERASE MUTANT VARIANT W48F  
Authors : Toyota, C.G.; Berthold, C.L.; Gruez, A.; Jonsson, S.; Lindqvist, Y.; Cambillau, C.; Richards, N.G.J.  
Deposited on : 2007-12-11  
Resolution : 1.95 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  
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

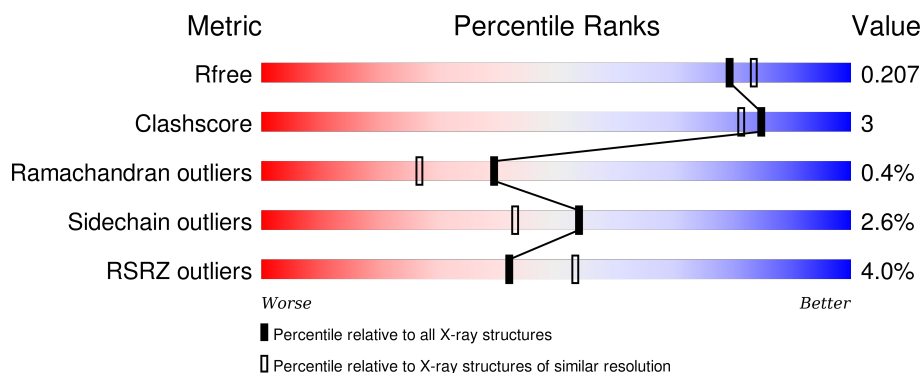
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

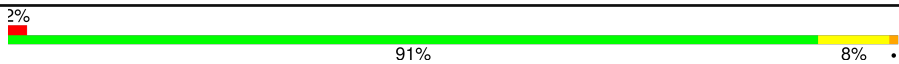
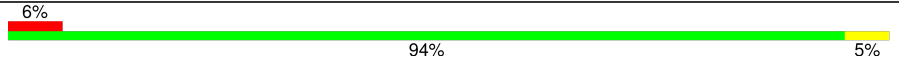
The reported resolution of this entry is 1.95 Å.

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	1833 (1.96-1.96)
Clashscore	102246	1953 (1.96-1.96)
Ramachandran outliers	100387	1936 (1.96-1.96)
Sidechain outliers	100360	1936 (1.96-1.96)
RSRZ outliers	91569	1835 (1.96-1.96)

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	A	428	
1	B	428	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7617 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 FORMYL-COENZYME A TRANSFERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	427	Total	C	N	O	S	0	3	0
			3324	2103	567	631	23			
1	B	427	Total	C	N	O	S	0	5	0
			3335	2110	570	632	23			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	48	PHE	TRP	ENGINEERED MUTATION	UNP O06644
B	48	PHE	TRP	ENGINEERED MUTATION	UNP O06644
A	186	ILE	MET	SEE REMARK 999	UNP O06644
B	186	ILE	MET	SEE REMARK 999	UNP O06644

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Na	0	0
			1	1		

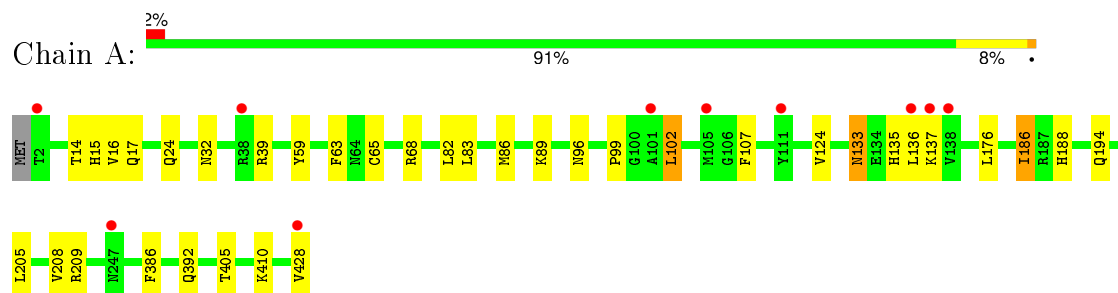
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	514	Total	O	0	0
			514	514		
3	B	443	Total	O	0	0
			443	443		

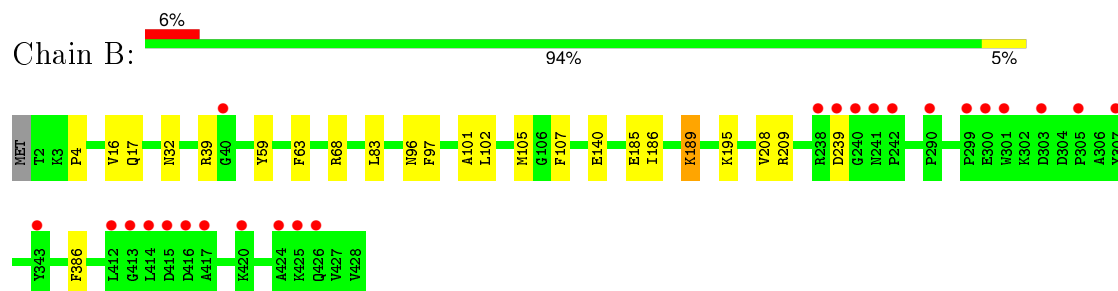
### 3 Residue-property plots

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: FORMYL-COENZYME A TRANSFERASE



#### • Molecule 1: FORMYL-COENZYME A TRANSFERASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 4	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	152.72Å 152.72Å 99.45Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 1.95 49.72 – 1.95	Depositor EDS
% Data completeness (in resolution range)	97.3 (30.00-1.95) 97.3 (49.72-1.95)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.36 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.166 , 0.201 0.172 , 0.207	Depositor DCC
$R_{free}$ test set	4030 reflections (5.25%)	DCC
Wilson B-factor (Å <sup>2</sup> )	17.4	Xtriage
Anisotropy	0.194	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 52.9	EDS
Estimated twinning fraction	0.025 for -k,-h,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 80836 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7617	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.05% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NA

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	0/3406	0.59	0/4604
1	B	0.43	0/3424	0.56	0/4629
All	All	0.45	0/6830	0.57	0/9233

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

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	3324	0	3266	28	0
1	B	3335	0	3277	16	0
2	A	1	0	0	0	0
3	A	514	0	0	2	0
3	B	443	0	0	2	0
All	All	7617	0	6543	38	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 (38) 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:186:ILE:HG23	1:B:4:PRO:HA	1.25	1.14
1:A:186:ILE:HD11	1:A:194:GLN:NE2	1.75	1.02
1:A:186:ILE:HD11	1:A:194:GLN:HE21	1.27	0.97
1:A:405:THR:HG21	1:A:428:VAL:HG13	1.60	0.81
1:A:186:ILE:CG2	1:B:4:PRO:HA	2.11	0.77
1:A:186:ILE:CD1	1:A:194:GLN:NE2	2.55	0.69
1:A:15:HIS:H	1:A:96:ASN:HD21	1.39	0.68
1:A:133:ASN:HD21	1:A:136:LEU:HD11	1.62	0.63
1:B:83:LEU:HD23	1:B:107:PHE:CE1	2.38	0.58
1:A:83:LEU:HA	1:A:86:MET:HE3	1.89	0.55
1:A:99:PRO:HG2	1:A:135:HIS:HA	1.89	0.54
1:B:208:VAL:O	1:B:208:VAL:HG12	2.08	0.53
1:A:405:THR:HG21	1:A:428:VAL:CG1	2.35	0.52
1:B:32:ASN:HD21	1:B:68:ARG:HH21	1.55	0.52
1:A:24:GLN:HE21	1:A:65:CYS:H	1.58	0.51
1:A:17:GLN:HG3	1:A:63:PHE:CE2	2.46	0.51
1:A:136:LEU:HD22	3:B:2167:HOH:O	2.12	0.49
1:A:102:LEU:HD22	1:A:107:PHE:HB2	1.95	0.48
1:A:136:LEU:HD23	3:A:2062:HOH:O	2.14	0.47
1:A:63:PHE:CZ	1:B:209:ARG:HD2	2.51	0.46
1:A:208:VAL:HG12	1:A:208:VAL:O	2.15	0.45
1:A:83:LEU:HD23	1:A:86:MET:CE	2.45	0.45
1:A:124:VAL:CG2	1:A:176:LEU:HD22	2.45	0.45
1:A:32:ASN:ND2	3:A:2031:HOH:O	2.49	0.45
1:B:96:ASN:O	1:B:96:ASN:ND2	2.49	0.45
1:A:209:ARG:HD2	1:B:63:PHE:CZ	2.53	0.44
1:A:32:ASN:HD21	1:A:68:ARG:HH21	1.65	0.43
1:A:133:ASN:ND2	1:A:136:LEU:HD11	2.31	0.43
1:B:102:LEU:HA	1:B:105:MET:HG2	1.99	0.42
1:B:97:PHE:HB3	1:B:101:ALA:HB3	2.02	0.42
1:A:124:VAL:HG22	1:A:176:LEU:HD22	2.01	0.42
1:A:209:ARG:HD2	1:B:63:PHE:CE2	2.55	0.42
1:A:14:THR:HB	1:A:96:ASN:ND2	2.35	0.42
1:B:186:ILE:O	1:B:189:LYS:HG3	2.21	0.40
1:B:32:ASN:ND2	3:B:2034:HOH:O	2.53	0.40
1:A:188:HIS:HE1	1:B:185:GLU:OE1	2.05	0.40
1:B:208:VAL:O	1:B:208:VAL:CG1	2.69	0.40
1:B:17:GLN:HG3	1:B:63:PHE:CE2	2.56	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	A	428/428 (100%)	417 (97%)	10 (2%)	1 (0%)	52	43
1	B	430/428 (100%)	422 (98%)	6 (1%)	2 (0%)	34	21
All	All	858/856 (100%)	839 (98%)	16 (2%)	3 (0%)	39	35

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	140	GLU
1	A	16	VAL
1	B	16	VAL

### 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	353/351 (101%)	341 (97%)	12 (3%)	44	30
1	B	355/351 (101%)	348 (98%)	7 (2%)	63	55
All	All	708/702 (101%)	689 (97%)	19 (3%)	54	41

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	39	ARG
1	A	59	TYR
1	A	82	LEU

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	89	LYS
1	A	102	LEU
1	A	133	ASN
1	A	137	LYS
1	A	186	ILE
1	A	205	LEU
1	A	386	PHE
1	A	392	GLN
1	A	410	LYS
1	B	39	ARG
1	B	59	TYR
1	B	189	LYS
1	B	195	LYS
1	B	239[A]	ASP
1	B	239[B]	ASP
1	B	386	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	9	ASN
1	A	24	GLN
1	A	32	ASN
1	A	54	ASN
1	A	96	ASN
1	A	113	GLN
1	A	133	ASN
1	A	188	HIS
1	A	206	ASN
1	A	392	GLN
1	B	32	ASN

### 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 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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 ⓘ

### 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	427/428 (99%)	-0.13	10 (2%) 64 73	7, 14, 35, 45	0
1	B	427/428 (99%)	0.13	24 (5%) 28 39	7, 17, 41, 48	0
All	All	854/856 (99%)	-0.00	34 (3%) 42 53	7, 15, 38, 48	0

All (34) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	415	ASP	5.9
1	B	417	ALA	5.9
1	B	424	ALA	4.4
1	B	412	LEU	4.4
1	B	301	TRP	3.9
1	B	414	LEU	3.9
1	A	101	ALA	3.7
1	B	241	ASN	3.5
1	A	136	LEU	3.4
1	B	413	GLY	3.2
1	A	138	VAL	3.1
1	B	303	ASP	3.1
1	B	307	TYR	3.0
1	B	343	TYR	3.0
1	B	420	LYS	2.9
1	B	242	PRO	2.8
1	B	299	PRO	2.5
1	A	38	ARG	2.5
1	B	240	GLY	2.3
1	B	238	ARG	2.3
1	A	428	VAL	2.3
1	B	239[A]	ASP	2.3
1	B	40	GLY	2.3
1	A	111	TYR	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	105	MET	2.2
1	B	425	LYS	2.2
1	A	137	LYS	2.1
1	B	300	GLU	2.1
1	B	290	PRO	2.1
1	B	305	PRO	2.1
1	A	2	THR	2.1
1	B	426	GLN	2.0
1	A	247	ASN	2.0
1	B	416	ASP	2.0

## 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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	NA	A	1429	1/1	0.99	0.08	-0.83	23,23,23,23	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.