



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 9, 2017 – 02:10 PM EDT

PDB ID : 2NS7  
Title : How an in vitro selected peptide mimics the antibiotic tetracycline to induce TET repressor  
Authors : Luckner, S.R.; Klotzsche, M.; Berens, C.; Hillen, W.; Muller, Y.A.  
Deposited on : unknown  
Resolution : 2.40 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

<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-20030345  
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-20030345

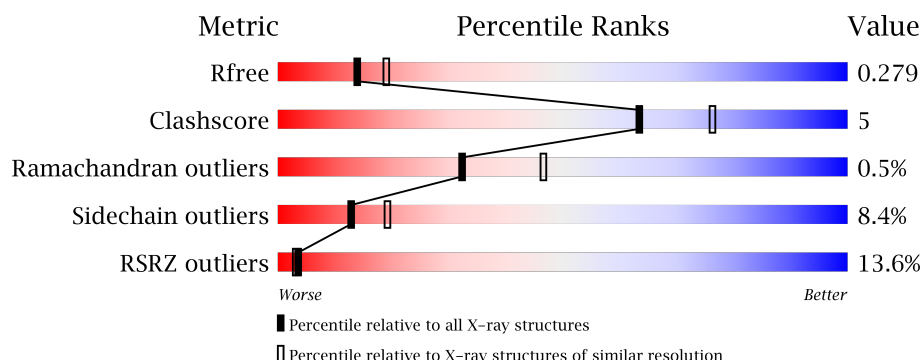
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3166 (2.40-2.40)
Clashscore	112137	3674 (2.40-2.40)
Ramachandran outliers	110173	3616 (2.40-2.40)
Sidechain outliers	110143	3617 (2.40-2.40)
RSRZ outliers	101464	3195 (2.40-2.40)

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	208	
1	B	208	
1	C	208	
1	D	208	

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 6027 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 Tetracycline repressor protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	196	Total	C	N	O	S	0	0	0
			1558	987	276	293	2			
1	B	194	Total	C	N	O	S	0	0	0
			1546	980	276	288	2			
1	C	190	Total	C	N	O	S	0	0	0
			1511	957	268	284	2			
1	D	178	Total	C	N	O	S	0	0	0
			1412	896	250	265	1			

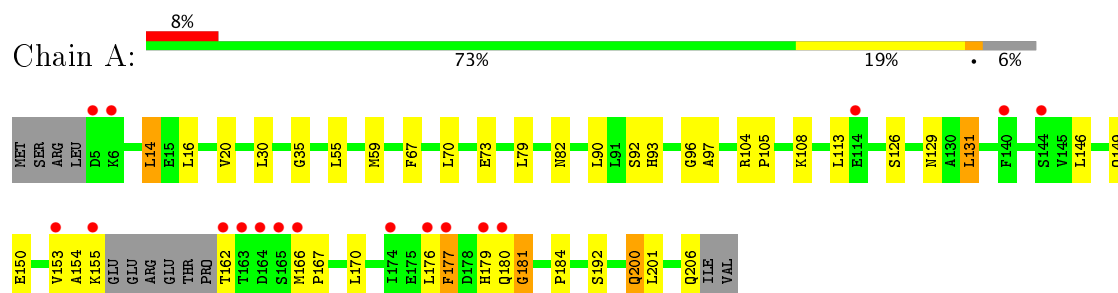
There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	68	SER	CYS	ENGINEERED	UNP P04483
A	88	ASN	CYS	ENGINEERED	UNP P04483
A	121	THR	CYS	ENGINEERED	UNP P04483
A	144	SER	CYS	ENGINEERED	UNP P04483
B	68	SER	CYS	ENGINEERED	UNP P04483
B	88	ASN	CYS	ENGINEERED	UNP P04483
B	121	THR	CYS	ENGINEERED	UNP P04483
B	144	SER	CYS	ENGINEERED	UNP P04483
C	68	SER	CYS	ENGINEERED	UNP P04483
C	88	ASN	CYS	ENGINEERED	UNP P04483
C	121	THR	CYS	ENGINEERED	UNP P04483
C	144	SER	CYS	ENGINEERED	UNP P04483
D	68	SER	CYS	ENGINEERED	UNP P04483
D	88	ASN	CYS	ENGINEERED	UNP P04483
D	121	THR	CYS	ENGINEERED	UNP P04483
D	144	SER	CYS	ENGINEERED	UNP P04483

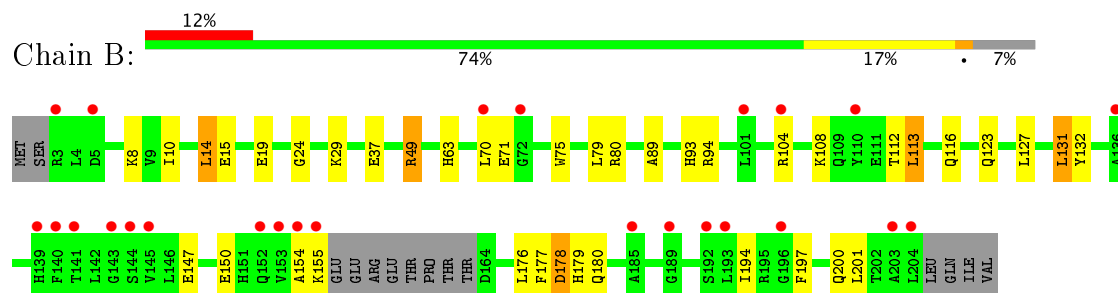
### 3 Residue-property plots [i](#)

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.

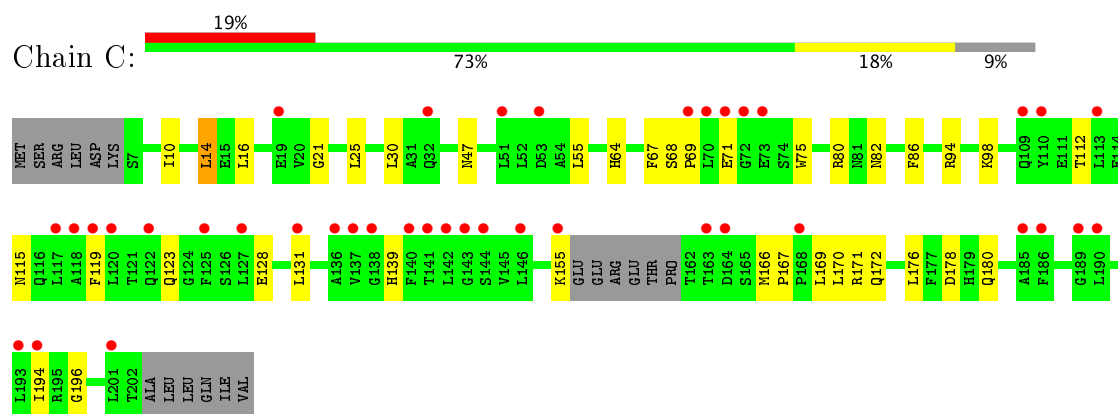
#### • Molecule 1: Tetracycline repressor protein



#### • Molecule 1: Tetracycline repressor protein

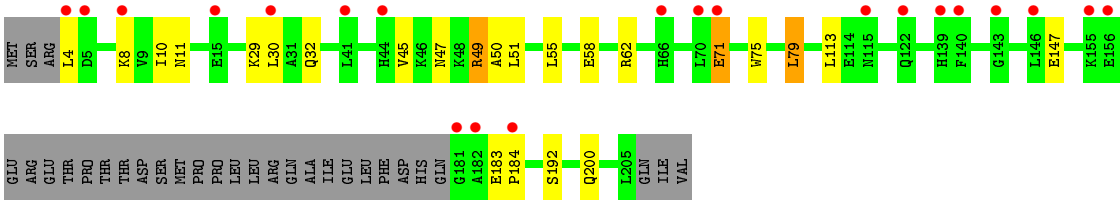


#### • Molecule 1: Tetracycline repressor protein



#### • Molecule 1: Tetracycline repressor protein





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	107.91Å 107.91Å 303.94Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	37.04 – 2.40 37.05 – 2.40	Depositor EDS
% Data completeness (in resolution range)	100.0 (37.04-2.40) 98.5 (37.05-2.40)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.90 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.230 , 0.277 0.237 , 0.279	Depositor DCC
$R_{free}$ test set	3302 reflections (8.70%)	DCC
Wilson B-factor (Å <sup>2</sup> )	52.6	Xtriage
Anisotropy	0.540	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 82.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	6027	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	64.0	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 30.60 % 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 1.2720e-03. 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

### 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	A	0.52	0/1588	0.63	0/2147
1	B	0.65	1/1576 (0.1%)	0.68	0/2129
1	C	0.61	2/1541 (0.1%)	0.59	1/2084 (0.0%)
1	D	0.46	0/1438	0.57	0/1942
All	All	0.57	3/6143 (0.0%)	0.62	1/8302 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	128	GLU	CD-OE1	14.10	1.41	1.25
1	B	37	GLU	CD-OE2	13.57	1.40	1.25
1	C	128	GLU	CD-OE2	8.06	1.34	1.25

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	128	GLU	OE1-CD-OE2	5.48	129.88	123.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	35	GLY	Peptide

## 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	1558	0	1548	21	0
1	B	1546	0	1539	17	0
1	C	1511	0	1496	14	0
1	D	1412	0	1407	15	0
All	All	6027	0	5990	59	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 5.

The worst 5 of 59 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:D:29:LYS:HA	1:D:32:GLN:HE21	1.52	0.75
1:A:16:LEU:O	1:A:20:VAL:HG22	1.95	0.67
1:A:180:GLN:OE1	1:B:132:TYR:OH	2.14	0.66
1:A:97:ALA:H	1:A:149:GLN:HE22	1.47	0.63
1:A:177:PHE:CD2	1:B:131:LEU:HD11	2.35	0.62

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	192/208 (92%)	185 (96%)	4 (2%)	3 (2%)	11	15
1	B	190/208 (91%)	187 (98%)	3 (2%)	0	100	100

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	186/208 (89%)	184 (99%)	2 (1%)	0	100	100
1	D	174/208 (84%)	169 (97%)	4 (2%)	1 (1%)	28	41
All	All	742/832 (89%)	725 (98%)	13 (2%)	4 (0%)	32	46

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	154	ALA
1	A	181	GLY
1	D	71	GLU
1	A	179	HIS

### 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	167/179 (93%)	154 (92%)	13 (8%)	15	23
1	B	165/179 (92%)	146 (88%)	19 (12%)	6	9
1	C	162/179 (90%)	149 (92%)	13 (8%)	14	21
1	D	150/179 (84%)	141 (94%)	9 (6%)	22	35
All	All	644/716 (90%)	590 (92%)	54 (8%)	13	19

5 of 54 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	131	LEU
1	B	180	GLN
1	D	55	LEU
1	B	147	GLU
1	B	155	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	109	GLN
1	B	122	GLN
1	D	123	GLN
1	B	115	ASN
1	B	152	GLN

### 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	A	196/208 (94%)	0.69	17 (8%) 11 10	58, 63, 77, 88	0
1	B	194/208 (93%)	0.90	25 (12%) 4 3	54, 62, 76, 88	0
1	C	190/208 (91%)	1.22	40 (21%) 1 1	52, 62, 74, 80	0
1	D	178/208 (85%)	0.85	21 (11%) 5 4	56, 64, 73, 86	0
All	All	758/832 (91%)	0.92	103 (13%) 3 3	52, 63, 76, 88	0

The worst 5 of 103 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	72	GLY	9.1
1	A	164	ASP	7.2
1	B	72	GLY	7.2
1	C	117	LEU	6.9
1	B	110	TYR	6.6

### 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.