



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 14, 2017 – 04:50 PM EDT

PDB ID : 2NS8
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 : 2.55 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

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

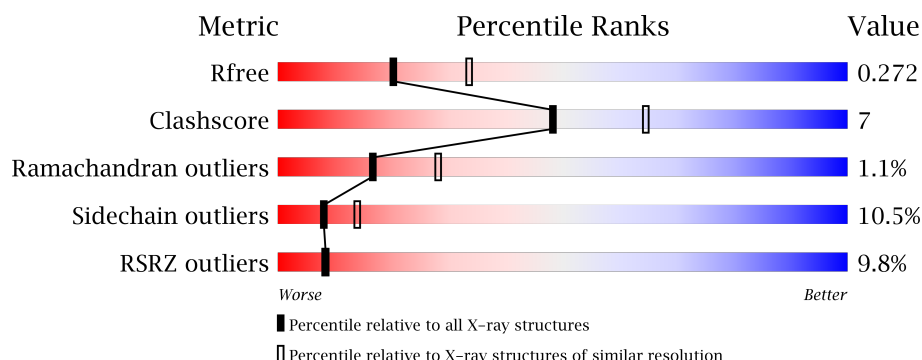
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.55 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	4993 (2.58-2.50)
Clashscore	112137	5755 (2.58-2.50)
Ramachandran outliers	110173	5652 (2.58-2.50)
Sidechain outliers	110143	5654 (2.58-2.50)
RSRZ outliers	101464	5026 (2.58-2.50)

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 ≥ 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	<div> <div>8%</div> <div>76%</div> <div>17%</div> <div>• •</div> </div>
1	B	208	<div> <div>6%</div> <div>72%</div> <div>24%</div> <div>• •</div> </div>
1	C	208	<div> <div>11%</div> <div>76%</div> <div>19%</div> <div>• •</div> </div>
1	D	208	<div> <div>9%</div> <div>75%</div> <div>18%</div> <div>• 5%</div> </div>
2	E	17	<div> <div>12%</div> <div>71%</div> <div>29%</div> </div>

Continued on next page...

Mol	Chain	Length	Quality of chain
2	F	17	<p>24% 59% 18% 24%</p>
2	G	17	<p>6% 53% 18% 29%</p>
2	H	17	<p>24% 65% 12% 24%</p>
2	Z	17	<p>18% 18% 6% 76%</p>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6884 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	200	Total	C	N	O	S	0	0	0
			1599	1009	285	303	2			
1	B	204	Total	C	N	O	S	0	0	0
			1626	1029	287	308	2			
1	C	206	Total	C	N	O	S	0	0	0
			1643	1038	292	311	2			
1	D	198	Total	C	N	O	S	0	0	0
			1577	996	280	299	2			

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

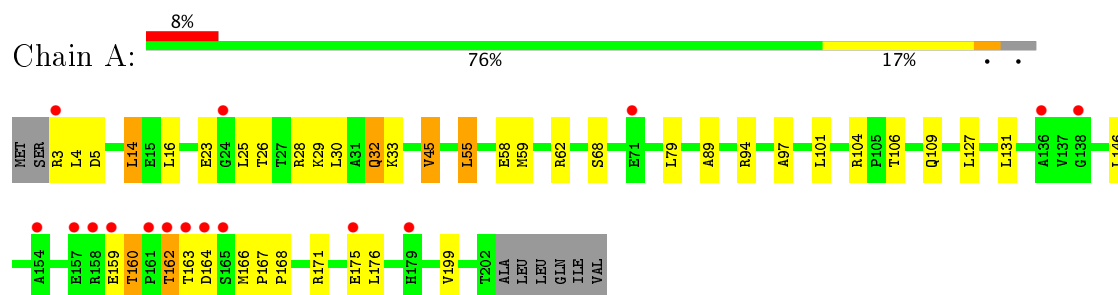
- Molecule 2 is a protein called 16 residue peptide Tip (Transcription inducing peptide).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	H	13	Total	C	N	O	0	0	0
			102	70	15	17			
2	E	12	Total	C	N	O	0	0	0
			96	67	14	15			
2	F	13	Total	C	N	O	0	0	0
			102	70	15	17			
2	G	12	Total	C	N	O	0	0	0
			96	67	14	15			
2	Z	4	Total	C	N	O	0	0	0
			43	30	7	6			

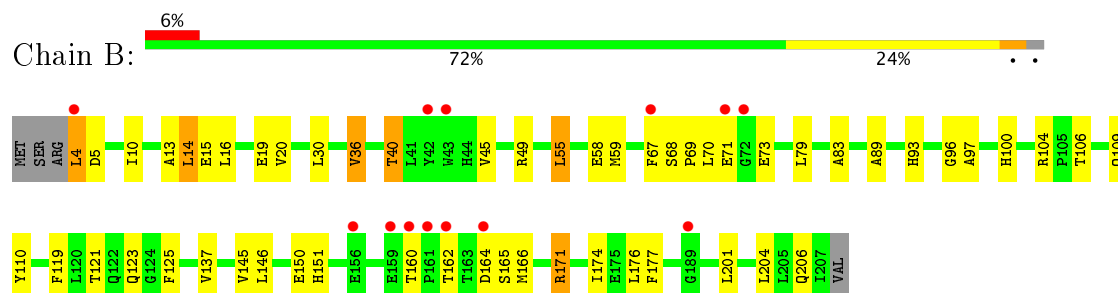
3 Residue-property plots [i](#)

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.

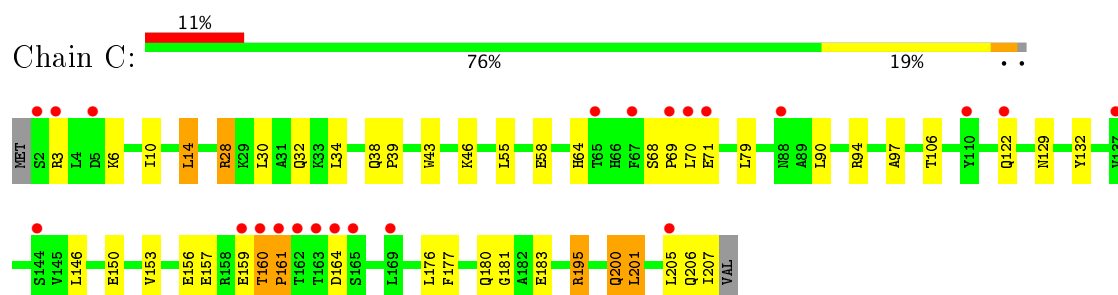
- Molecule 1: Tetracycline repressor protein



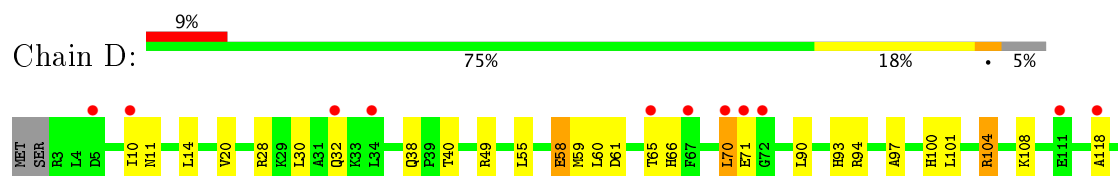
- Molecule 1: Tetracycline repressor protein

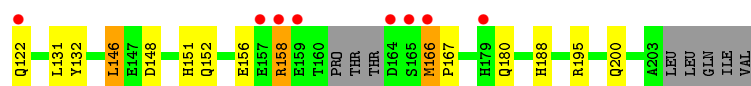


- Molecule 1: Tetracycline repressor protein



- Molecule 1: Tetracycline repressor protein

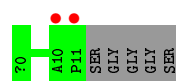




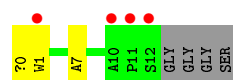
- Molecule 2: 16 residue peptide Tip (Transcription inducing peptide)



- Molecule 2: 16 residue peptide Tip (Transcription inducing peptide)



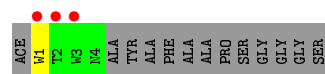
- Molecule 2: 16 residue peptide Tip (Transcription inducing peptide)



- Molecule 2: 16 residue peptide Tip (Transcription inducing peptide)



- Molecule 2: 16 residue peptide Tip (Transcription inducing peptide)



4 Data and refinement statistics

Property	Value	Source
Space group	P 2 ₁ 2 ₁ 2	Depositor
Cell constants a, b, c, α , β , γ	77.89 Å 200.59 Å 65.28 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.55 38.23 – 2.55	Depositor EDS
% Data completeness (in resolution range)	100.0 (30.00-2.55) 99.5 (38.23-2.55)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.57 (at 2.54 Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.206 , 0.270 0.214 , 0.272	Depositor DCC
R_{free} test set	2713 reflections (8.69%)	DCC
Wilson B-factor (Å ²)	39.1	Xtriage
Anisotropy	0.439	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 46.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	6884	wwPDB-VP
Average B, all atoms (Å ²)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.38% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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

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

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	0.66	0/1631	0.70	0/2206
1	B	0.70	0/1658	0.79	2/2244 (0.1%)
1	C	0.61	0/1675	0.71	1/2266 (0.0%)
1	D	0.60	0/1607	0.69	0/2171
2	E	0.60	0/100	0.66	0/140
2	F	0.58	0/106	0.53	0/148
2	G	0.53	0/100	0.63	0/140
2	H	0.70	0/106	0.64	0/148
2	Z	0.89	0/46	0.59	0/64
All	All	0.64	0/7029	0.72	3/9527 (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	B	0	1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	171	ARG	NE-CZ-NH2	-6.69	116.95	120.30
1	C	195	ARG	NE-CZ-NH1	5.82	123.21	120.30
1	B	171	ARG	NE-CZ-NH1	5.46	123.03	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	164	ASP	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	1599	0	1583	19	0
1	B	1626	0	1616	30	0
1	C	1643	0	1634	26	0
1	D	1577	0	1555	26	0
2	E	96	0	81	0	0
2	F	102	0	86	4	0
2	G	96	0	81	6	0
2	H	102	0	86	2	0
2	Z	43	0	35	0	0
All	All	6884	0	6757	94	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.

The worst 5 of 94 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)
2:F:0:ACE:H1	2:F:1:TRP:HB2	1.53	0.91
1:C:10:ILE:HG21	1:C:58:GLU:HG3	1.61	0.83
1:C:206:GLN:O	1:C:207:ILE:HG23	1.88	0.74
1:A:26:THR:OG1	1:A:29:LYS:HD2	1.90	0.71
1:C:200:GLN:HE21	1:D:188:HIS:HE1	1.39	0.70

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	198/208 (95%)	183 (92%)	13 (7%)	2 (1%)	18	31
1	B	202/208 (97%)	197 (98%)	3 (2%)	2 (1%)	18	31
1	C	204/208 (98%)	197 (97%)	5 (2%)	2 (1%)	18	31
1	D	194/208 (93%)	181 (93%)	10 (5%)	3 (2%)	12	20
2	E	10/17 (59%)	9 (90%)	1 (10%)	0	100	100
2	F	11/17 (65%)	9 (82%)	2 (18%)	0	100	100
2	G	10/17 (59%)	9 (90%)	1 (10%)	0	100	100
2	H	11/17 (65%)	11 (100%)	0	0	100	100
2	Z	2/17 (12%)	2 (100%)	0	0	100	100
All	All	842/917 (92%)	798 (95%)	35 (4%)	9 (1%)	17	28

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	160	THR
1	B	165	SER
1	A	162	THR
1	C	64	HIS
1	C	161	PRO

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	172/179 (96%)	155 (90%)	17 (10%)	9	16
1	B	175/179 (98%)	156 (89%)	19 (11%)	7	13
1	C	177/179 (99%)	157 (89%)	20 (11%)	7	11
1	D	168/179 (94%)	149 (89%)	19 (11%)	7	11

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	E	7/9 (78%)	7 (100%)	0	100	100
2	F	8/9 (89%)	8 (100%)	0	100	100
2	G	7/9 (78%)	7 (100%)	0	100	100
2	H	8/9 (89%)	8 (100%)	0	100	100
2	Z	4/9 (44%)	3 (75%)	1 (25%)	1	1
All	All	726/761 (95%)	650 (90%)	76 (10%)	8	14

5 of 76 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	201	LEU
1	C	94	ARG
1	D	146	LEU
1	B	204	LEU
1	C	30	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	109	GLN
1	B	206	GLN
1	D	180	GLN
1	B	172	GLN
1	C	100	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 [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, 95th 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 ²)	Q<0.9
1	A	200/208 (96%)	0.56	16 (8%) 13 13	38, 48, 63, 73	0
1	B	204/208 (98%)	0.58	13 (6%) 20 21	42, 47, 56, 60	0
1	C	206/208 (99%)	0.74	22 (10%) 7 6	36, 46, 58, 70	0
1	D	198/208 (95%)	0.80	19 (9%) 9 9	38, 47, 59, 71	0
2	E	11/17 (64%)	1.27	2 (18%) 1 1	48, 50, 57, 59	0
2	F	12/17 (70%)	1.49	4 (33%) 0 0	49, 53, 60, 61	0
2	G	11/17 (64%)	0.95	1 (9%) 10 10	48, 50, 54, 58	0
2	H	12/17 (70%)	1.63	4 (33%) 0 0	45, 47, 54, 55	0
2	Z	4/17 (23%)	2.94	3 (75%) 0 0	53, 54, 55, 56	0
All	All	858/917 (93%)	0.72	84 (9%) 8 8	36, 47, 59, 73	0

The worst 5 of 84 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	163	THR	6.3
1	A	161	PRO	6.1
2	E	10	ALA	6.0
1	C	2	SER	5.6
2	Z	2	THR	5.6

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.