



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 25, 2023 – 03:40 AM EDT

PDB ID : 2Z9O
Title : Crystal structure of the dimeric form of RepE in complex with the repE operator DNA
Authors : Nakamura, A.; Wada, C.; Miki, K.
Deposited on : 2007-09-21
Resolution : 3.14 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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 : 2.36
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) : 2.36

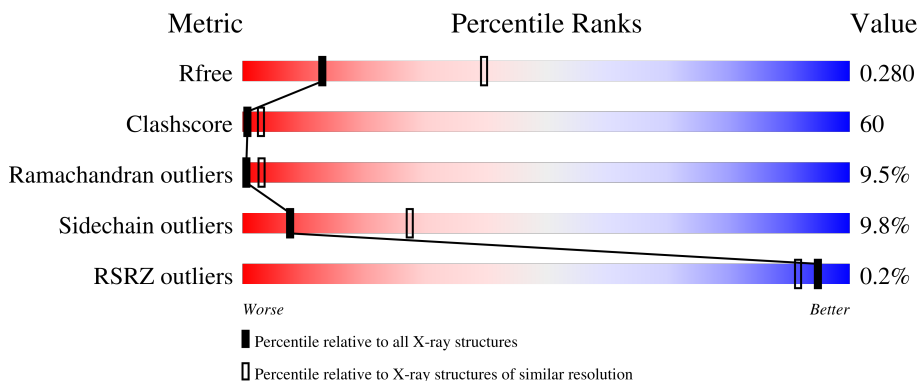
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

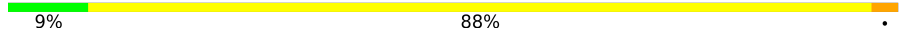
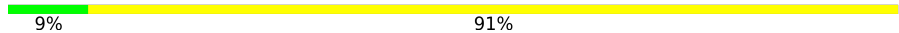
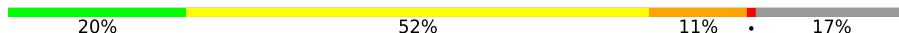
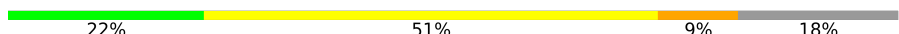
The reported resolution of this entry is 3.14 Å.

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	1626 (3.18-3.10)
Clashscore	141614	1735 (3.18-3.10)
Ramachandran outliers	138981	1677 (3.18-3.10)
Sidechain outliers	138945	1677 (3.18-3.10)
RSRZ outliers	127900	1588 (3.18-3.10)

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 of the lower bar indicate the fraction of residues that contain outliers for ≥ 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	C	33	 9% 88%
2	D	33	 9% 91%
3	A	266	 20% 52% 11% 17%
3	B	266	 22% 51% 9% 18%

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 4926 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 DNA chain called DNA (33-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	C	33	669	323	118	196	32	0	0	0

- Molecule 2 is a DNA chain called DNA (33-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	D	33	678	326	124	196	32	0	0	0

- Molecule 3 is a protein called Replication initiation protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	222	1796	1151	312	325	8	0	0	0
3	B	218	1783	1142	310	324	7	0	0	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-14	MET	-	expression tag	UNP P03856
A	-13	ARG	-	expression tag	UNP P03856
A	-12	GLY	-	expression tag	UNP P03856
A	-11	SER	-	expression tag	UNP P03856
A	-10	HIS	-	expression tag	UNP P03856
A	-9	HIS	-	expression tag	UNP P03856
A	-8	HIS	-	expression tag	UNP P03856
A	-7	HIS	-	expression tag	UNP P03856
A	-6	HIS	-	expression tag	UNP P03856
A	-5	HIS	-	expression tag	UNP P03856
A	-4	GLY	-	expression tag	UNP P03856
A	-3	SER	-	expression tag	UNP P03856

Continued on next page...

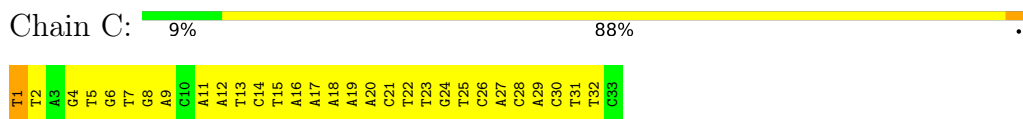
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	ILE	-	expression tag	UNP P03856
A	-1	GLU	-	expression tag	UNP P03856
A	0	GLY	-	expression tag	UNP P03856
A	1	ARG	-	expression tag	UNP P03856
B	-14	MET	-	expression tag	UNP P03856
B	-13	ARG	-	expression tag	UNP P03856
B	-12	GLY	-	expression tag	UNP P03856
B	-11	SER	-	expression tag	UNP P03856
B	-10	HIS	-	expression tag	UNP P03856
B	-9	HIS	-	expression tag	UNP P03856
B	-8	HIS	-	expression tag	UNP P03856
B	-7	HIS	-	expression tag	UNP P03856
B	-6	HIS	-	expression tag	UNP P03856
B	-5	HIS	-	expression tag	UNP P03856
B	-4	GLY	-	expression tag	UNP P03856
B	-3	SER	-	expression tag	UNP P03856
B	-2	ILE	-	expression tag	UNP P03856
B	-1	GLU	-	expression tag	UNP P03856
B	0	GLY	-	expression tag	UNP P03856
B	1	ARG	-	expression tag	UNP P03856

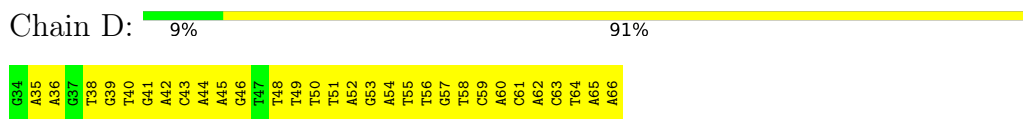
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

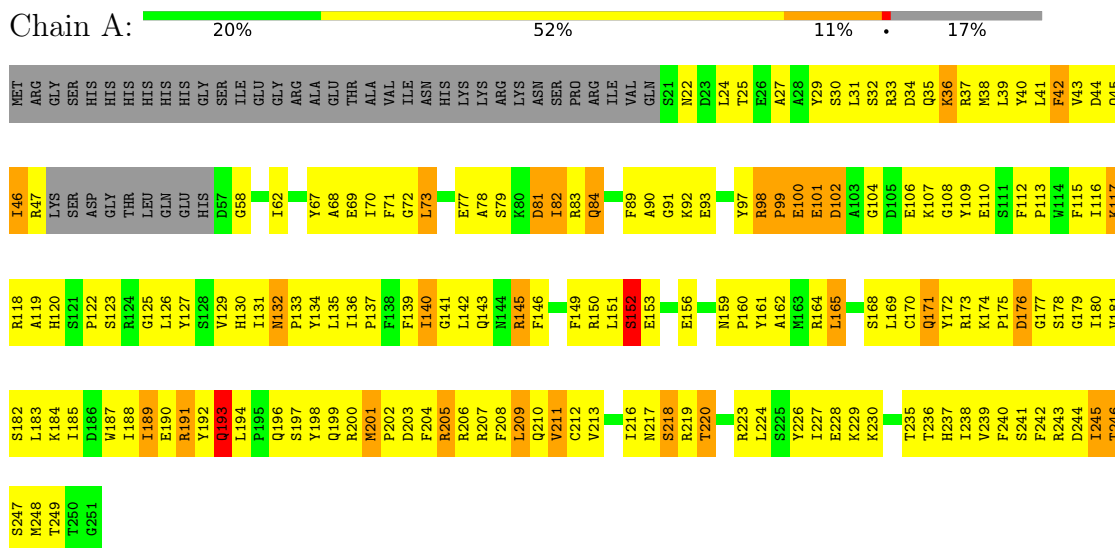
- Molecule 1: DNA (33-MER)



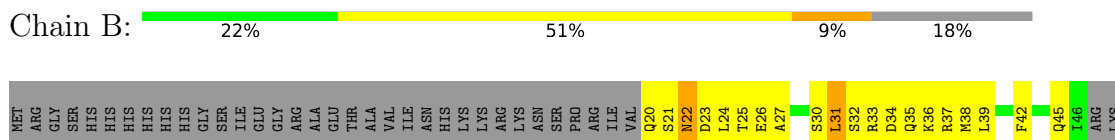
- Molecule 2: DNA (33-MER)



- Molecule 3: Replication initiation protein



- Molecule 3: Replication initiation protein



SER	R118	L183	D244
ASP	A119	K184	I245
GLY	H120	I185	T246
THR	S121	D186	S247
LEU	P122	W187	MET
GLN	S123	I188	THR
GLU	R124	I189	THR
HIS	G125	E190	GLY
D57	L126	R191	
G58	Y127	Y192	
I59	S128	Q193	
C60	V129	L194	
E61	H130	P195	
I62	I131	Q196	
H63	M132	S197	
V64	P133	Y198	
A65	Y134	Q199	
K66	L135	R200	
Y67	I136	K201	
A68	P137	F202	
E69	F138	D203	
I70	F139	F204	
F71		R205	
	L142	R206	
		R207	
T74		F208	
	R145	L209	
E77	F146	Q210	
A78	T147	V211	
S79	Q148	C212	
K80	F149	V213	
D81	R150	N214	
I82	L151	E215	
R83	S152	I216	
Q84	E153		
A85	T154	R219	
	K155	T220	
F89	E156	P221	
A90		F222	
	M159	R223	
E93	P160	L224	
	Y161	S225	
R98	A162	Y226	
P99	M163	L227	
E100	R164	E228	
E101	L165	K229	
D102	Y166	K230	
A103	E167	K231	
	S168	G232	
E106	L169	R233	
K107	C170	Q234	
G108	Q171	T235	
Y109	Y172	T236	
E110	R173	H237	
S111	K174	I238	
F112	P175	V239	
P113		F240	
W114	S178	S241	
F115		F242	
I116	V181	R243	
I117	S182		

4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	60.73Å 99.32Å 95.00Å 90.00° 108.55° 90.00°	Depositor
Resolution (Å)	45.03 – 3.14 45.03 – 3.14	Depositor EDS
% Data completeness (in resolution range)	87.8 (45.03-3.14) 87.8 (45.03-3.14)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.01 (at 3.12Å)	Xtrriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.266 , 0.313 0.247 , 0.280	Depositor DCC
R_{free} test set	850 reflections (5.15%)	wwPDB-VP
Wilson B-factor (Å ²)	51.2	Xtrriage
Anisotropy	1.080	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , -6.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.42$, $\langle L^2 \rangle = 0.25$	Xtrriage
Estimated twinning fraction	0.088 for h,-k,-h-l	Xtrriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	4926	wwPDB-VP
Average B, all atoms (Å ²)	45.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.43% 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 [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	C	0.46	0/749	0.72	0/1153
2	D	0.43	0/761	0.69	0/1174
3	A	0.47	0/1839	0.71	0/2480
3	B	0.47	0/1826	0.68	0/2462
All	All	0.46	0/5175	0.70	0/7269

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	C	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	1	DT	Sidechain

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	C	669	0	376	70	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	678	0	376	67	0
3	A	1796	0	1742	226	0
3	B	1783	0	1735	230	0
All	All	4926	0	4229	552	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 60.

The worst 5 of 552 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:B:200:ARG:HG2	3:B:202:PRO:HD2	1.15	1.10
2:D:39:DG:H2''	2:D:40:DT:H5''	1.25	1.10
2:D:41:DG:H2''	2:D:42:DA:H5'	1.38	1.04
1:C:6:DG:C2'	1:C:7:DT:H5''	1.86	1.04
2:D:60:DA:H2''	2:D:61:DC:H5'	1.40	1.04

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	
3	A	218/266 (82%)	158 (72%)	40 (18%)	20 (9%)	1	3
3	B	214/266 (80%)	158 (74%)	35 (16%)	21 (10%)	0	3
All	All	432/532 (81%)	316 (73%)	75 (17%)	41 (10%)	0	3

5 of 41 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	91	GLY

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	A	99	PRO
3	B	68	ALA
3	B	147	THR
3	B	209	LEU

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	
3	A	188/238 (79%)	162 (86%)	26 (14%)	3	15
3	B	189/238 (79%)	178 (94%)	11 (6%)	20	49
All	All	377/476 (79%)	340 (90%)	37 (10%)	8	27

5 of 37 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	B	89	PHE
3	B	223	ARG
3	B	98	ARG
3	B	154	THR
3	A	132	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
3	B	84	GLN
3	B	130	HIS
3	B	237	HIS
3	B	148	GLN
3	A	148	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 monosaccharides 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	C	33/33 (100%)	-0.88	0	100 100	22, 65, 95, 99	0
2	D	33/33 (100%)	-0.85	0	100 100	39, 59, 103, 109	0
3	A	222/266 (83%)	-0.31	0	100 100	3, 33, 66, 79	0
3	B	218/266 (81%)	-0.24	1 (0%)	91 83	5, 41, 69, 80	0
All	All	506/598 (84%)	-0.35	1 (0%)	95 91	3, 40, 75, 109	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	B	246	THR	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 monosaccharides 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.