



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2018 – 10:12 pm GMT

PDB ID : 2O7P  
Title : The crystal structure of RibD from Escherichia coli in complex with the oxidised NADP<sup>+</sup> cofactor in the active site of the reductase domain  
Authors : Moche, M.; Stenmark, P.; Gurmu, D.; Nordlund, P.; Structural Proteomics in Europe (SPINE)  
Deposited on : 2006-12-11  
Resolution : 3.00 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

<https://www.wwpdb.org/validation/2017/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 : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk30967  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

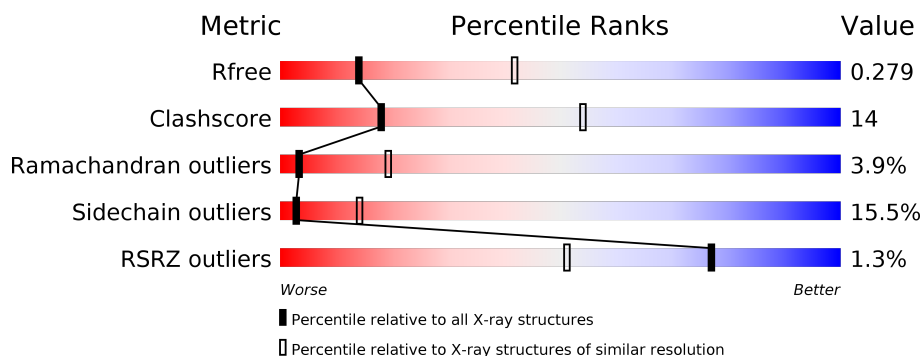
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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))
$R_{free}$	111664	1851 (3.00-3.00)
Clashscore	122126	2167 (3.00-3.00)
Ramachandran outliers	120053	2101 (3.00-3.00)
Sidechain outliers	120020	2104 (3.00-3.00)
RSRZ outliers	108989	1751 (3.00-3.00)

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	380	<div> <div> <div></div> <div>57%</div> <div>31%</div> <div>6%</div> <div>6%</div> </div> </div>
1	B	380	<div> <div> <div>2%</div> <div>61%</div> <div>26%</div> <div>6%</div> <div>7%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5523 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 Riboflavin biosynthesis protein ribD.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	358	Total	C	N	O	S	Se	0	0	0
			2733	1720	495	503	6	9			
1	B	353	Total	C	N	O	S	Se	0	0	0
			2694	1698	489	496	3	8			

There are 28 discrepancies between the modelled and reference sequences:

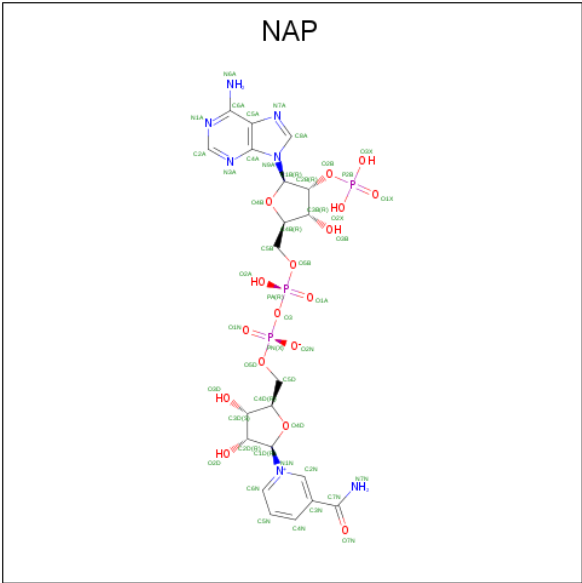
Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	THR	-	CLONING ARTIFACT	UNP P25539
A	-3	LEU	-	CLONING ARTIFACT	UNP P25539
A	-2	TYR	-	CLONING ARTIFACT	UNP P25539
A	-1	ILE	-	CLONING ARTIFACT	UNP P25539
A	0	GLN	-	CLONING ARTIFACT	UNP P25539
A	1	GLY	-	CLONING ARTIFACT	UNP P25539
A	368	SER	-	CLONING ARTIFACT	UNP P25539
A	369	THR	-	CLONING ARTIFACT	UNP P25539
A	370	HIS	-	EXPRESSION TAG	UNP P25539
A	371	HIS	-	EXPRESSION TAG	UNP P25539
A	372	HIS	-	EXPRESSION TAG	UNP P25539
A	373	HIS	-	EXPRESSION TAG	UNP P25539
A	374	HIS	-	EXPRESSION TAG	UNP P25539
A	375	HIS	-	EXPRESSION TAG	UNP P25539
B	-4	THR	-	CLONING ARTIFACT	UNP P25539
B	-3	LEU	-	CLONING ARTIFACT	UNP P25539
B	-2	TYR	-	CLONING ARTIFACT	UNP P25539
B	-1	ILE	-	CLONING ARTIFACT	UNP P25539
B	0	GLN	-	CLONING ARTIFACT	UNP P25539
B	1	GLY	-	CLONING ARTIFACT	UNP P25539
B	368	SER	-	CLONING ARTIFACT	UNP P25539
B	369	THR	-	CLONING ARTIFACT	UNP P25539
B	370	HIS	-	EXPRESSION TAG	UNP P25539
B	371	HIS	-	EXPRESSION TAG	UNP P25539
B	372	HIS	-	EXPRESSION TAG	UNP P25539

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	373	HIS	-	EXPRESSION TAG	UNP P25539
B	374	HIS	-	EXPRESSION TAG	UNP P25539
B	375	HIS	-	EXPRESSION TAG	UNP P25539

- Molecule 2 is NADP NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NAP) (formula: C<sub>21</sub>H<sub>28</sub>N<sub>7</sub>O<sub>17</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
2	B	1	Total	C	N	O	P	0	0
			48	21	7	17	3		



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	173.00Å 173.00Å 77.80Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	28.93 – 3.00 28.92 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.9 (28.93-3.00) 99.9 (28.92-3.00)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.28 (at 3.00Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.212 , 0.267 0.234 , 0.279	Depositor DCC
$R_{free}$ test set	1359 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	93.2	Xtriage
Anisotropy	0.014	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 41.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.047 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	5523	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	85.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.75% 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

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

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	0.82	4/2776 (0.1%)	0.87	3/3753 (0.1%)
1	B	0.82	7/2736 (0.3%)	0.83	4/3703 (0.1%)
All	All	0.82	11/5512 (0.2%)	0.85	7/7456 (0.1%)

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	2
1	B	0	3
All	All	0	5

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	116	GLN	CD-OE1	8.29	1.42	1.24
1	B	116	GLN	CD-NE2	8.20	1.53	1.32
1	B	62	LYS	CD-CE	7.45	1.69	1.51
1	A	297	TRP	CD2-CE3	6.82	1.50	1.40
1	A	297	TRP	CD1-NE1	6.46	1.49	1.38

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	116	GLN	CG-CD-OE1	-7.85	105.90	121.60
1	A	67	THR	N-CA-C	7.03	129.99	111.00
1	B	116	GLN	OE1-CD-NE2	5.72	135.05	121.90
1	A	143	ARG	NE-CZ-NH1	5.53	123.07	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	153	LEU	CA-CB-CG	5.19	127.23	115.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	145	GLY	Peptide
1	A	217	ALA	Peptide
1	B	116	GLN	Sidechain
1	B	217	ALA	Peptide
1	B	366	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	2733	0	2727	83	0
1	B	2694	0	2686	78	0
2	A	48	0	25	1	0
2	B	48	0	25	1	0
All	All	5523	0	5463	158	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 14.

The worst 5 of 158 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:212:ASP:OD1	1:A:214:GLN:HG2	1.54	1.07
1:B:213:GLU:OE1	1:B:213:GLU:HA	1.59	1.03
1:A:194:SER:HB3	1:A:233:ASP:HB3	1.46	0.97
1:B:206:VAL:H	1:B:226:GLN:NE2	1.63	0.96
1:A:170:TRP:HH2	1:A:200:ASP:OD1	1.47	0.96

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	350/380 (92%)	309 (88%)	29 (8%)	12 (3%)	4	22
1	B	345/380 (91%)	295 (86%)	35 (10%)	15 (4%)	3	17
All	All	695/760 (91%)	604 (87%)	64 (9%)	27 (4%)	3	19

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	67	THR
1	A	95	ARG
1	A	217	ALA
1	A	328	SER
1	B	95	ARG

### 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	284/302 (94%)	238 (84%)	46 (16%)	2	13
1	B	277/302 (92%)	236 (85%)	41 (15%)	3	16
All	All	561/604 (93%)	474 (84%)	87 (16%)	3	14

5 of 87 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	278	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	20	THR
1	B	260	SER
1	A	291	GLN
1	A	349	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	294	ASN
1	B	104	ASN
1	B	226	GLN
1	A	226	GLN
1	A	247	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](#)

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	NAP	A	402	-	44,52,52	1.68	3 (6%)	53,80,80	1.60	3 (5%)
2	NAP	B	1402	-	44,52,52	1.67	3 (6%)	53,80,80	1.62	3 (5%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAP	A	402	-	-	0/27/67/67	0/5/5/5
2	NAP	B	1402	-	-	0/27/67/67	0/5/5/5

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	402	NAP	C2A-N1A	2.75	1.39	1.33
2	B	1402	NAP	C2A-N1A	2.79	1.39	1.33
2	B	1402	NAP	C2A-N3A	3.43	1.37	1.32
2	A	402	NAP	C2A-N3A	4.08	1.38	1.32
2	A	402	NAP	O7N-C7N	8.59	1.41	1.24

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1402	NAP	N3A-C2A-N1A	-9.39	120.83	128.86
2	A	402	NAP	N3A-C2A-N1A	-8.75	121.37	128.86
2	B	1402	NAP	C4D-O4D-C1D	-3.90	105.76	109.83
2	A	402	NAP	C4D-O4D-C1D	-2.12	107.62	109.83
2	B	1402	NAP	C2N-C3N-C4N	2.16	120.75	118.26

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	402	NAP	1	0
2	B	1402	NAP	1	0

## 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	349/380 (91%)	-0.16	3 (0%)	84 62	68, 82, 108, 122	0
1	B	345/380 (90%)	-0.14	6 (1%)	70 41	63, 81, 109, 118	0
All	All	694/760 (91%)	-0.15	9 (1%)	77 51	63, 81, 109, 122	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	-4	THR	3.9
1	A	75	CYS	3.0
1	B	217	ALA	2.6
1	B	118	ALA	2.6
1	B	123	SER	2.4

### 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. 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	B-factors( $\text{\AA}^2$ )	Q<0.9
2	NAP	B	1402	48/48	0.94	0.14	95,107,141,143	0
2	NAP	A	402	48/48	0.94	0.16	71,81,89,93	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.