



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 24, 2020 – 05:09 PM BST

PDB ID : 6Y4Z
Title : The crystal structure of human MACROD2 in space group P43212
Authors : Wazir, S.; Maksimainen, M.M.; Lehtio, L.
Deposited on : 2020-02-24
Resolution : 1.90 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.14.6
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) : 2.14.6

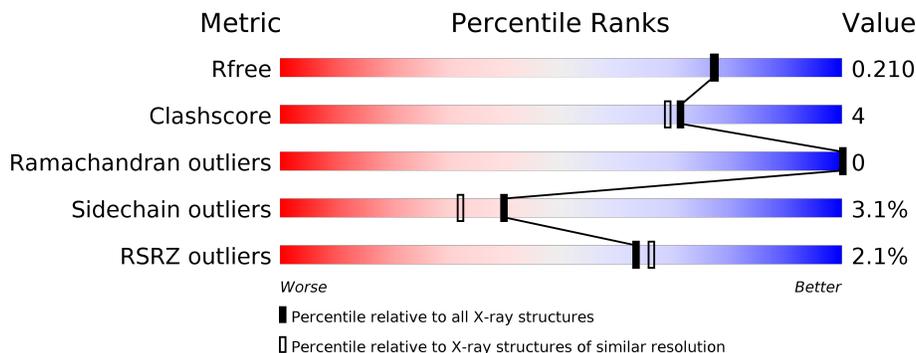
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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 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	A	366	
1	B	366	
1	C	366	
1	D	366	

2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 7372 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 Thioredoxin 1,ADP-ribose glycohydrolase MACROD2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	221	Total 1778	C 1132	N 314	O 319	S 13	0	4	0
1	B	216	Total 1713	C 1092	N 299	O 310	S 12	0	1	0
1	C	214	Total 1702	C 1086	N 298	O 307	S 11	0	1	0
1	D	215	Total 1729	C 1102	N 305	O 310	S 12	0	4	0

There are 80 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-122	MET	-	initiating methionine	UNP P0AA25
A	-121	HIS	-	expression tag	UNP P0AA25
A	-120	HIS	-	expression tag	UNP P0AA25
A	-119	HIS	-	expression tag	UNP P0AA25
A	-118	HIS	-	expression tag	UNP P0AA25
A	-117	HIS	-	expression tag	UNP P0AA25
A	-116	HIS	-	expression tag	UNP P0AA25
A	-115	SER	-	expression tag	UNP P0AA25
A	-114	SER	-	expression tag	UNP P0AA25
A	-113	GLY	-	expression tag	UNP P0AA25
A	-3	GLY	-	linker	UNP P0AA25
A	-2	THR	-	linker	UNP P0AA25
A	-1	GLU	-	linker	UNP P0AA25
A	0	ASN	-	linker	UNP P0AA25
A	1	LEU	-	linker	UNP P0AA25
A	2	TYR	-	linker	UNP P0AA25
A	3	PHE	-	linker	UNP P0AA25
A	4	GLN	-	linker	UNP P0AA25
A	5	SER	-	linker	UNP P0AA25
A	6	MET	-	linker	UNP P0AA25
B	-122	MET	-	initiating methionine	UNP P0AA25

Continued on next page...

Continued from previous page...

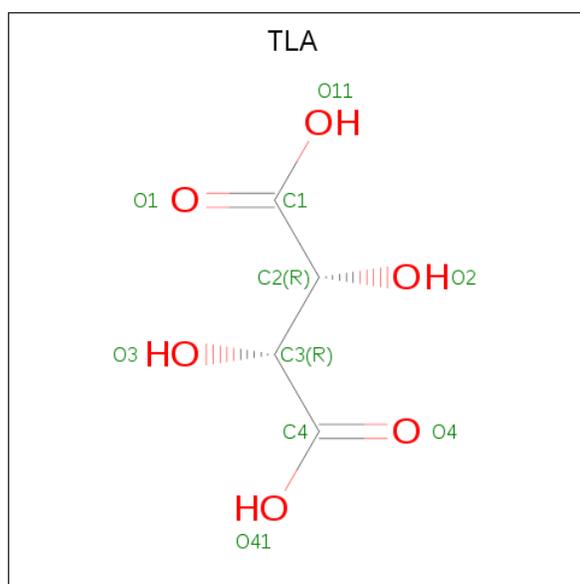
Chain	Residue	Modelled	Actual	Comment	Reference
B	-121	HIS	-	expression tag	UNP P0AA25
B	-120	HIS	-	expression tag	UNP P0AA25
B	-119	HIS	-	expression tag	UNP P0AA25
B	-118	HIS	-	expression tag	UNP P0AA25
B	-117	HIS	-	expression tag	UNP P0AA25
B	-116	HIS	-	expression tag	UNP P0AA25
B	-115	SER	-	expression tag	UNP P0AA25
B	-114	SER	-	expression tag	UNP P0AA25
B	-113	GLY	-	expression tag	UNP P0AA25
B	-3	GLY	-	linker	UNP P0AA25
B	-2	THR	-	linker	UNP P0AA25
B	-1	GLU	-	linker	UNP P0AA25
B	0	ASN	-	linker	UNP P0AA25
B	1	LEU	-	linker	UNP P0AA25
B	2	TYR	-	linker	UNP P0AA25
B	3	PHE	-	linker	UNP P0AA25
B	4	GLN	-	linker	UNP P0AA25
B	5	SER	-	linker	UNP P0AA25
B	6	MET	-	linker	UNP P0AA25
C	-122	MET	-	initiating methionine	UNP P0AA25
C	-121	HIS	-	expression tag	UNP P0AA25
C	-120	HIS	-	expression tag	UNP P0AA25
C	-119	HIS	-	expression tag	UNP P0AA25
C	-118	HIS	-	expression tag	UNP P0AA25
C	-117	HIS	-	expression tag	UNP P0AA25
C	-116	HIS	-	expression tag	UNP P0AA25
C	-115	SER	-	expression tag	UNP P0AA25
C	-114	SER	-	expression tag	UNP P0AA25
C	-113	GLY	-	expression tag	UNP P0AA25
C	-3	GLY	-	linker	UNP P0AA25
C	-2	THR	-	linker	UNP P0AA25
C	-1	GLU	-	linker	UNP P0AA25
C	0	ASN	-	linker	UNP P0AA25
C	1	LEU	-	linker	UNP P0AA25
C	2	TYR	-	linker	UNP P0AA25
C	3	PHE	-	linker	UNP P0AA25
C	4	GLN	-	linker	UNP P0AA25
C	5	SER	-	linker	UNP P0AA25
C	6	MET	-	linker	UNP P0AA25
D	-122	MET	-	initiating methionine	UNP P0AA25
D	-121	HIS	-	expression tag	UNP P0AA25
D	-120	HIS	-	expression tag	UNP P0AA25

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-119	HIS	-	expression tag	UNP P0AA25
D	-118	HIS	-	expression tag	UNP P0AA25
D	-117	HIS	-	expression tag	UNP P0AA25
D	-116	HIS	-	expression tag	UNP P0AA25
D	-115	SER	-	expression tag	UNP P0AA25
D	-114	SER	-	expression tag	UNP P0AA25
D	-113	GLY	-	expression tag	UNP P0AA25
D	-3	GLY	-	linker	UNP P0AA25
D	-2	THR	-	linker	UNP P0AA25
D	-1	GLU	-	linker	UNP P0AA25
D	0	ASN	-	linker	UNP P0AA25
D	1	LEU	-	linker	UNP P0AA25
D	2	TYR	-	linker	UNP P0AA25
D	3	PHE	-	linker	UNP P0AA25
D	4	GLN	-	linker	UNP P0AA25
D	5	SER	-	linker	UNP P0AA25
D	6	MET	-	linker	UNP P0AA25

- Molecule 2 is L(+)-TARTARIC ACID (three-letter code: TLA) (formula: C₄H₆O₆).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	D	1	Total	C	O	0	0
			10	4	6		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	127	Total 127	O 127	0	0
3	B	104	Total 105	O 105	0	1
3	C	109	Total 109	O 109	0	0
3	D	99	Total 99	O 99	0	0

GLN
GLY
LYS
LEU
THR
VAL
ALA
LYS
LEU
ASN
ASP
GLN
ASN
PRO
GLY
THR
ALA
PRO
LYS
TYR
ILE
ILE
ARG
GLY
GLY
ILE
PRO
THR
LEU
LEU
LEU
LEU
PHE
LYS
ASN
GLN
GLY
GLU
VAL
VAL
ALA
ALA
THR
LYS
LYS
VAL
GLY
GLY
ALA
LEU
LYS
SER

THR
GLU
ASN
LEU
TYR
PHE
GLN
SER
MET
LYS
K9
E24
K28
R32
D33
E46
M47
K48
GLY
LYS
GLY
GLN
ASN
LEU
GLU
GLU
ASN
THR
GLN
VAL
VAL
LYS
S67
V72
R76
A108
C132
G133
Y134
E195
P196
E215
I219
GLY

F224
F229
K230
I231
Y232
K235
M236
F240
SER
VAL
ASP

- Molecule 1: Thioredoxin 1,ADP-ribose glycohydrolase MACROD2



MET
HIS
HIS
HIS
HIS
HIS
HIS
SER
SER
MET
SER
ASP
GLN
LYS
ILE
ILE
HIS
HIS
LEU
THR
ASP
SER
PHE
ASP
THR
VAL
LEU
LYS
ALA
GLY
ALA
ILE
ILE
VAL
ASP
PHE
TRP
ALA
GLU
TRP
CYS
GLY
PRO
LYS
MET
ILE
ALA
PRO
ILE
LEU
ASP
GLU
ILE
ALA
ASP
ALA
TYR

GLN
GLY
LYS
LEU
THR
VAL
ALA
LYS
LEU
ASN
ASP
GLN
ASN
PRO
GLY
THR
ALA
PRO
LYS
TYR
GLY
PHE
ILE
ARG
GLY
ILE
PRO
VAL
LEU
LEU
LEU
LEU
PHE
LYS
ASN
GLY
GLU
VAL
ALA
THR
LYS
VAL
GLY
ALA
LEU
SER
LYS
GLN
GLY
LEU
LYS
PHE
LEU
ASP
ALA
ASN
ALA
GLY

THR
GLU
ASN
LEU
TYR
PHE
GLN
SER
MET
LYS
K8
E14
R17
L23
R26
R32
M47
K48
GLY
LYS
GLY
GLN
THR
ASN
GLU
THR
SER
GLN
VAL
LYS
S67
I68
I69
I79
V64
D85
A86
C132
G133
Y134
A138
G146

C164
K170
E174
S179
V180
A181
E195
P196
I205
R218
C222
I231
Y232
K233
K234
K235
M236
N237
F240
SER
VAL
ASP

4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	95.36Å 95.36Å 261.92Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.73 – 1.90 47.68 – 1.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (47.73-1.90) 100.0 (47.68-1.90)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.36 (at 1.90Å)	Xtrriage
Refinement program	REFMAC 5.8.0257	Depositor
R, R_{free}	0.201 , 0.228 0.210 , 0.210	Depositor DCC
R_{free} test set	4799 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	36.0	Xtrriage
Anisotropy	0.499	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 42.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	7372	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 4.79% 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](#)

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

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	0.72	0/1817	0.86	0/2443
1	B	0.71	0/1749	0.83	0/2356
1	C	0.69	0/1739	0.84	0/2344
1	D	0.71	0/1771	0.84	0/2384
All	All	0.71	0/7076	0.84	0/9527

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 [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	1778	0	1817	12	0
1	B	1713	0	1740	10	0
1	C	1702	0	1724	8	0
1	D	1729	0	1765	21	0
2	D	10	0	4	0	0
3	A	127	0	0	0	0
3	B	105	0	0	0	0
3	C	109	0	0	0	0
3	D	99	0	0	0	0
All	All	7372	0	7050	49	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 4.

The worst 5 of 49 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:206:LYS:HE2	1:A:239:PHE:O	1.65	0.97
1:D:47:MET:CE	1:D:237:ASN:HD21	1.81	0.93
1:D:26[A]:ARG:HH11	1:D:26[A]:ARG:HG3	1.57	0.68
1:D:47:MET:HE3	1:D:237:ASN:HD21	1.62	0.64
1:C:232:TYR:O	1:C:236:MET:HG2	2.02	0.59

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	221/366 (60%)	218 (99%)	3 (1%)	0	100	100
1	B	213/366 (58%)	207 (97%)	6 (3%)	0	100	100
1	C	211/366 (58%)	208 (99%)	3 (1%)	0	100	100
1	D	215/366 (59%)	210 (98%)	5 (2%)	0	100	100
All	All	860/1464 (59%)	843 (98%)	17 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	
1	A	193/310 (62%)	187 (97%)	6 (3%)	40	32
1	B	186/310 (60%)	179 (96%)	7 (4%)	33	24
1	C	184/310 (59%)	180 (98%)	4 (2%)	52	47
1	D	188/310 (61%)	180 (96%)	8 (4%)	29	19
All	All	751/1240 (61%)	726 (97%)	25 (3%)	40	29

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	233	LYS
1	C	24	GLU
1	D	205	ILE
1	B	241	SER
1	C	33	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	237	ASN

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](#)

1 ligand is 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	TLA	D	301	-	3,9,9	1.61	1 (33%)	6,12,12	2.11	2 (33%)

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	TLA	D	301	-	-	0/4/12/12	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	301	TLA	O2-C2	2.52	1.47	1.42

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	301	TLA	C1-C2-C3	-3.66	105.23	113.11
2	D	301	TLA	C4-C3-C2	-3.06	106.51	113.11

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 [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	221/366 (60%)	-0.15	2 (0%) 84 85	29, 40, 67, 108	0
1	B	216/366 (59%)	0.12	7 (3%) 47 50	27, 40, 75, 96	0
1	C	214/366 (58%)	0.12	4 (1%) 66 69	31, 41, 74, 91	0
1	D	215/366 (58%)	0.24	5 (2%) 60 63	28, 41, 74, 106	0
All	All	866/1464 (59%)	0.08	18 (2%) 63 66	27, 40, 74, 108	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	242	VAL	5.2
1	C	224	PHE	3.9
1	D	240	PHE	3.6
1	B	8	LYS	3.4
1	D	68	LEU	3.0

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](#)

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, 95th 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(\AA^2)	Q<0.9
2	TLA	D	301	10/10	0.95	0.08	36,41,46,46	0

6.5 Other polymers [i](#)

There are no such residues in this entry.