



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 12, 2017 – 09:42 pm GMT

PDB ID : 1R4I
Title : Crystal Structure of Androgen Receptor DNA-Binding Domain Bound to a Direct Repeat Response Element
Authors : Shaffer, P.L.; Jivan, A.; Dollins, D.E.; Claessens, F.; Gewirth, D.T.
Deposited on : 2003-10-06
Resolution : 3.10 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	:	trunk28620
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)	:	recalc28949

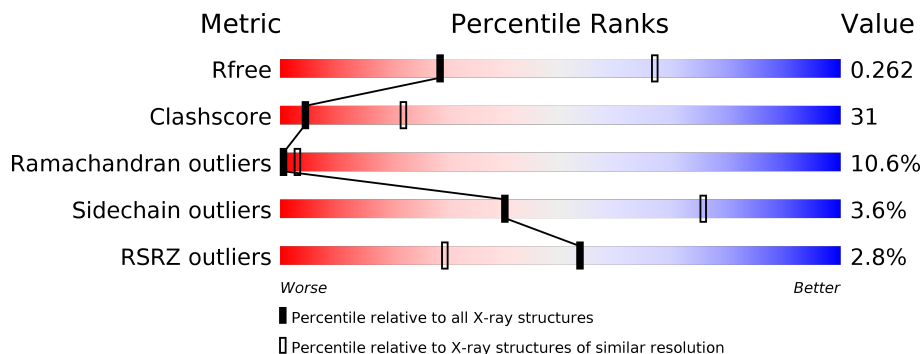
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.10 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	1001 (3.12-3.08)
Clashscore	112137	1099 (3.12-3.08)
Ramachandran outliers	110173	1057 (3.12-3.08)
Sidechain outliers	110143	1057 (3.12-3.08)
RSRZ outliers	101464	1006 (3.12-3.08)

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	C	18	<div> <div>17%</div> <div>83%</div> </div>
2	D	18	<div> <div>22%</div> <div>78%</div> </div>
3	A	105	<div> <div>%</div> <div>30%</div> <div>31%</div> <div>10%</div> <div>30%</div> </div>
3	B	105	<div> <div>4%</div> <div>32%</div> <div>31%</div> <div>•</div> <div>32%</div> </div>

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 1813 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 5'-D(*CP*CP*AP*GP*AP*AP*CP*AP*TP*CP*AP*AP*GP*AP*AP*CP*AP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	18	Total	C	N	O	P	0	0	0
			367	175	77	98	17			

- Molecule 2 is a DNA chain called 5'-D(*CP*TP*GP*TP*TP*CP*TP*TP*GP*AP*TP*GP*TP*TP*CP*TP*GP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	18	Total	C	N	O	P	0	0	0
			365	177	57	114	17			

- Molecule 3 is a protein called Androgen receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	74	Total	C	N	O	S	0	0	0
			553	338	104	101	10			
3	B	71	Total	C	N	O	S	0	0	0
			524	319	98	97	10			

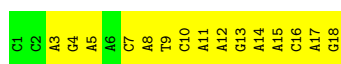
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	552	ALA	CYS	ENGINEERED	UNP P15207
B	552	ALA	CYS	ENGINEERED	UNP P15207

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	2	Total	Zn	0	0
			2	2		
4	A	2	Total	Zn	0	0
			2	2		

● Molecule 1: 5'-D(*CP*CP*AP*GP*AP*AP*CP*AP*TP*CP*AP*AP*GP*AP*AP*CP*AP*G)-3'



C19	T20	G21	T22	T23	C24	T25	T26	G27	A28	T29	G30	T31	T32	C33	T34	G35	G36
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

ASP	TTR	TTR	PHE	PRO	PS38	PS39	CS42	CS43	CS44	CS45	CS46	DS47	HS53	HS54	GS55	AS56	LS57	TS58	CS59	GS60	SS61	CS62	KS63	VS64	FS65	FS66	KS67	RS68	AS69	AS70	ES71	QS74	AS75	YS76	LS77	SS80	RS81	IS86	DS87	KS88	RS91	GS92	NS93	CS94	PS95	RS98	LS99	RS00	KS01	CS02	YS03
MS07	TS08	LS09	GS10	AS11	ARG	LYS	LEU	LVS	LVS	GLN	GLU	GLY	GLY	GLU	ASN	SER	SER	ALA	GLY	SER	PRO	THR	GLU	ASP																											

Protein	Residue	Score
ASP	Tyr	1543
	Tyr	1544
	Phe	C545
	Pro	C546
	Pro	D547
	Pro	E548
	Gln	H553
	Gln	Y554
	Glu	G555
	Glu	A556
ASP	Glu	L557
	Asn	T558
	Asn	C559
	Ser	S560
	Ser	S561
	Ala	V564
	Gly	F565
	Pro	F566
	Thr	K567
	Glu	R568
ASP	Asp	E571
	Thr	D572
	Glu	K573
	Glu	Q574
	Glu	K575
	Asp	Y576
	Ala	R581
	Ser	N582
	Ser	D583
	Ser	C584
ASP	Thr	T585
	Ser	L586
	Asp	D587
	Ser	R590
	Ser	R591
	Ser	K592
	Ser	N593
	Asp	C597
	Ser	R598
	Ser	L599
ASP	Glu	R600
	Glu	K601
	Asp	C602
	Asp	Y603
	Asp	C604
	Asp	C605
	Asp	C606
	Asp	C607
	Asp	C608
	Asp	C609

4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	137.89Å 137.89Å 85.71Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 3.10 40.34 – 3.10	Depositor EDS
% Data completeness (in resolution range)	85.4 (50.00-3.10) 85.5 (40.34-3.10)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.16 (at 3.12Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.246 , 0.264 0.243 , 0.262	Depositor DCC
R_{free} test set	725 reflections (4.89%)	DCC
Wilson B-factor (Å ²)	113.9	Xtriage
Anisotropy	0.217	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 54.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.019 for -h,-k,l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	1813	wwPDB-VP
Average B, all atoms (Å ²)	117.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.59% 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: ZN

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	0.40	0/414	0.75	0/636
2	D	0.46	0/406	0.77	0/626
3	A	0.50	0/561	0.68	0/748
3	B	0.44	0/531	0.67	0/710
All	All	0.46	0/1912	0.72	0/2720

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	C	367	0	201	19	0
2	D	365	0	209	15	0
3	A	553	0	534	38	0
3	B	524	0	491	34	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0
All	All	1813	0	1435	102	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 31.

The worst 5 of 102 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:D:28:DA:H1'	2:D:29:DT:H5''	1.51	0.91
1:C:10:DC:H1'	1:C:11:DA:H5'	1.58	0.85
3:A:542:CYS:HB3	3:A:546:GLY:H	1.42	0.84
1:C:9:DT:H2''	1:C:10:DC:H5'	1.67	0.76
3:B:586:ILE:HD13	3:B:598:ARG:HA	1.72	0.69

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	72/105 (69%)	47 (65%)	17 (24%)	8 (11%)	0	2
3	B	69/105 (66%)	48 (70%)	14 (20%)	7 (10%)	1	4
All	All	141/210 (67%)	95 (67%)	31 (22%)	15 (11%)	0	3

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	554	TYR
3	B	547	ASP
3	B	556	ALA
3	A	566	PHE
3	A	588	LYS

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	57/86 (66%)	54 (95%)	3 (5%)	26	63
3	B	53/86 (62%)	52 (98%)	1 (2%)	62	87
All	All	110/172 (64%)	106 (96%)	4 (4%)	40	75

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	539	GLN
3	A	577	LEU
3	A	586	ILE
3	B	581	ARG

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

Mol	Chain	Res	Type
3	A	574	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](#)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

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

There are no such residues in this entry.

5.8 Polymer linkage issues

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	18/18 (100%)	-0.93	0	100 100	98, 116, 144, 146	0
2	D	18/18 (100%)	-0.94	0	100 100	99, 125, 139, 142	0
3	A	74/105 (70%)	0.03	1 (1%)	75 57	86, 106, 136, 147	0
3	B	71/105 (67%)	0.07	4 (5%)	25 11	99, 126, 143, 151	0
All	All	181/246 (73%)	-0.15	5 (2%)	53 29	86, 116, 142, 151	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	B	571	GLU	3.0
3	B	566	PHE	2.7
3	B	610	GLY	2.5
3	B	609	LEU	2.4
3	A	538	PRO	2.1

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains.

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	LLDF	B-factors(Å ²)	Q<0.9
4	ZN	A	651	1/1	1.00	0.15	-0.15	89,89,89,89	0
4	ZN	B	652	1/1	1.00	0.13	-0.65	95,95,95,95	0
4	ZN	A	650	1/1	0.99	0.14	-0.87	86,86,86,86	0
4	ZN	B	653	1/1	0.99	0.10	-0.87	88,88,88,88	0

6.5 Other polymers [i](#)

There are no such residues in this entry.