



# Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 11:49 AM GMT

PDB ID : 1KB2  
Title : Crystal Structure of VDR DNA-binding Domain Bound to Mouse Osteopontin  
(SPP) Response Element  
Authors : Shaffer, P.L.; Gewirth, D.T.  
Deposited on : 2001-11-05  
Resolution : 2.70 Å(reported)

This is a full wwPDB validation 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

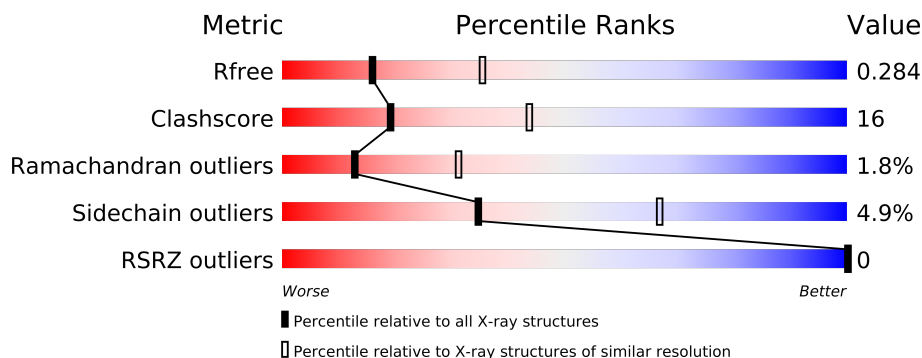
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.70 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))
$R_{free}$	66092	1557 (2.70-2.70)
Clashscore	79885	1939 (2.70-2.70)
Ramachandran outliers	78287	1905 (2.70-2.70)
Sidechain outliers	78261	1905 (2.70-2.70)
RSRZ outliers	66119	1559 (2.70-2.70)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	C	18	<div><div></div><div><div></div><div></div><div></div></div></div>
2	D	18	<div><div></div><div><div></div><div></div><div></div></div></div>
3	A	110	<div><div></div><div><div></div><div></div><div></div></div></div>
3	B	110	<div><div></div><div><div></div><div></div><div></div></div></div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 2135 atoms, of which 0 are hydrogen and 0 are deuterium.

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\*AP\*CP\*GP\*GP\*TP\*TP\*CP\*AP\*CP\*GP\*A P\*GP\*GP\*TP\*TP\*CP\*A)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	18	Total	C	N	O	P	0	0	0
			366	175	68	106	17			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	18	Total	C	N	O	P	0	0	0
			366	175	68	106	17			

- Molecule 3 is a protein called Vitamin D3 Receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	89	Total	C	N	O	S	0	0	0
			704	431	139	120	14			
3	B	85	Total	C	N	O	S	0	0	0
			660	406	128	112	14			

- 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 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	17	Total	O	0	0
			17	17		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	8	Total	O	0	0
			8	8		
5	C	5	Total	O	0	0
			5	5		
5	D	5	Total	O	0	0
			5	5		

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

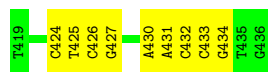
- Molecule 1: 5'-D(\*CP\*AP\*CP\*GP\*GP\*TP\*TP\*CP\*AP\*CP\*GP\*AP\*GP\*GP\*TP\*TP\*CP\*A)-3'

Chain C: 



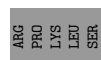
- Molecule 2: 5'-D(\*TP\*GP\*AP\*AP\*CP\*CP\*TP\*CP\*GP\*TP\*GP\*AP\*AP\*CP\*CP\*GP\*TP\*G)-3'

Chain D: 



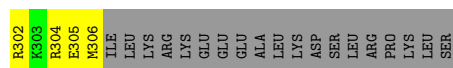
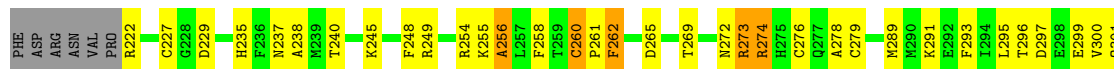
- Molecule 3: Vitamin D3 Receptor

Chain A: 



- Molecule 3: Vitamin D3 Receptor

Chain B: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	62.14Å 62.14Å 241.75Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.70 32.52 – 2.70	Depositor EDS
% Data completeness (in resolution range)	90.2 (50.00-2.70) 90.1 (32.52-2.70)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.83 (at 2.68Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.226 , 0.289 0.228 , 0.284	Depositor DCC
$R_{free}$ test set	1277 reflections (11.39%)	DCC
Wilson B-factor (Å <sup>2</sup> )	73.5	Xtriage
Anisotropy	0.025	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 46.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 12538 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	2135	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	75.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 17.74% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

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	0.83	0/410	0.91	0/631
2	D	0.86	0/410	0.97	0/631
3	A	0.82	1/713 (0.1%)	0.87	0/943
3	B	0.76	1/669 (0.1%)	0.83	0/887
All	All	0.81	2/2202 (0.1%)	0.89	0/3092

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	4
2	D	0	1
All	All	0	5

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	260	CYS	CB-SG	-6.22	1.71	1.82
3	A	27	CYS	CB-SG	-5.41	1.73	1.81

There are no bond angle outliers.

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	404	DG	Sidechain
1	C	405	DG	Sidechain

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
1	C	413	DG	Sidechain
1	C	414	DG	Sidechain
2	D	432	DC	Sidechain

## 5.2 Close contacts ⓘ

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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	366	0	204	10	0
2	D	366	0	204	6	0
3	A	704	0	689	17	0
3	B	660	0	640	32	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0
5	A	17	0	0	2	0
5	B	8	0	0	1	0
5	C	5	0	0	1	0
5	D	5	0	0	0	0
All	All	2135	0	1737	63	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 16.

All (63) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:B:272:ASN:O	3:B:274:ARG:N	2.07	0.87
3:A:96:THR:O	3:A:100:VAL:HG23	1.80	0.81
3:B:237:ASN:HB2	3:B:293:PHE:HE1	1.46	0.80
3:B:237:ASN:HB2	3:B:293:PHE:CE1	2.19	0.75
3:A:42:GLU:HG2	5:A:630:HOH:O	1.91	0.71
3:A:109:LYS:O	3:A:109:LYS:HE2	1.91	0.69
1:C:404:DG:H2''	1:C:405:DG:H5'	1.75	0.69
3:B:222:ARG:HH11	3:B:222:ARG:HG2	1.59	0.67
3:B:227:CYS:SG	3:B:229:ASP:HB2	2.35	0.65
3:B:235:HIS:CE1	3:B:245:LYS:HD2	2.35	0.61

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:B:254:ARG:C	3:B:256:ALA:H	2.01	0.61
2:D:426:DC:H1'	2:D:427:DG:C8	2.36	0.60
1:C:403:DC:H1'	1:C:404:DG:C8	2.37	0.58
3:B:272:ASN:OD1	3:B:273:ARG:N	2.38	0.57
2:D:424:DC:H1'	2:D:425:DT:H5''	1.87	0.56
3:B:272:ASN:CG	3:B:273:ARG:N	2.59	0.56
3:B:301:GLN:O	3:B:302:ARG:C	2.44	0.55
3:B:296:THR:OG1	3:B:299:GLU:HG3	2.07	0.54
3:B:235:HIS:ND1	3:B:245:LYS:HD2	2.24	0.53
1:C:415:DT:H2''	1:C:416:DT:H5'	1.92	0.52
3:A:109:LYS:O	3:A:109:LYS:CE	2.58	0.52
3:B:301:GLN:O	3:B:304:ARG:N	2.43	0.51
3:B:222:ARG:NH1	3:B:222:ARG:HG2	2.23	0.51
3:A:105:GLU:O	3:A:109:LYS:HB2	2.11	0.50
3:B:254:ARG:HG2	3:B:254:ARG:O	2.12	0.50
3:A:38:ALA:O	3:A:40:THR:HG23	2.12	0.50
3:B:248:PHE:CG	3:B:289:MET:HE2	2.47	0.49
3:B:269:THR:O	3:B:273:ARG:HB2	2.13	0.49
3:B:255:LYS:O	3:B:256:ALA:C	2.52	0.48
3:B:258:PHE:HB2	3:B:278:ALA:HA	1.96	0.48
3:A:37:ASN:HB2	3:A:93:PHE:CE1	2.49	0.48
3:B:254:ARG:C	3:B:256:ALA:N	2.68	0.48
3:A:50:ARG:NE	5:A:633:HOH:O	2.07	0.48
3:A:35:HIS:ND1	3:A:45:LYS:HD2	2.29	0.47
2:D:424:DC:H2''	2:D:425:DT:H5'	1.97	0.47
3:B:265:ASP:O	3:B:265:ASP:CG	2.52	0.47
2:D:426:DC:H4'	2:D:427:DG:OP1	2.14	0.47
3:B:262:PHE:CD1	3:B:262:PHE:N	2.83	0.47
3:A:91:LYS:HG3	3:A:92:GLU:OE2	2.17	0.45
3:B:296:THR:O	3:B:297:ASP:C	2.55	0.44
3:A:42:GLU:OE2	3:A:45:LYS:HD3	2.17	0.44
1:C:412:DA:H1'	1:C:413:DG:C8	2.53	0.44
2:D:430:DA:H2''	2:D:431:DA:H5'	2.00	0.44
3:A:58:PHE:HB2	3:A:78:ALA:HA	2.00	0.44
1:C:405:DG:H5''	3:A:95:LEU:HD12	2.01	0.43
3:B:300:VAL:HG12	3:B:300:VAL:O	2.19	0.43
3:B:238:ALA:O	3:B:240:THR:HG23	2.19	0.43
3:A:29:ASP:OD2	3:A:30:ARG:N	2.49	0.42
1:C:417:DC:H5''	5:C:611:HOH:O	2.19	0.42
3:A:105:GLU:HB3	3:A:106:MET:CE	2.49	0.42
3:B:272:ASN:CG	3:B:273:ARG:H	2.22	0.42
2:D:433:DC:H2''	2:D:434:DG:H5'	2.01	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:B:260:CYS:HA	3:B:261:PRO:HD3	1.93	0.42
1:C:414:DG:OP2	3:B:249:ARG:HD3	2.20	0.42
1:C:416:DT:H5'	1:C:416:DT:C6	2.55	0.42
1:C:404:DG:C2'	1:C:405:DG:H5'	2.48	0.42
3:B:222:ARG:CG	3:B:222:ARG:NH1	2.82	0.42
3:B:274:ARG:O	3:B:274:ARG:CG	2.68	0.41
3:B:254:ARG:O	3:B:256:ALA:N	2.54	0.41
3:A:65:ASP:OD1	3:A:65:ASP:N	2.49	0.41
1:C:413:DG:H5''	5:B:620:HOH:O	2.20	0.41
3:B:276:CYS:SG	3:B:279:CYS:HB2	2.61	0.41
3:A:36:PHE:O	3:A:37:ASN:CB	2.69	0.40

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	
3	A	87/110 (79%)	80 (92%)	7 (8%)	0	100	100
3	B	83/110 (76%)	67 (81%)	13 (16%)	3 (4%)	5	11
All	All	170/220 (77%)	147 (86%)	20 (12%)	3 (2%)	13	31

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	B	273	ARG
3	B	256	ALA
3	B	305	GLU

### 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	
3	A	74/98 (76%)	72 (97%)	2 (3%)	57	87
3	B	69/98 (70%)	64 (93%)	5 (7%)	21	45
All	All	143/196 (73%)	136 (95%)	7 (5%)	35	67

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	54	ARG
3	A	70	LYS
3	B	262	PHE
3	B	274	ARG
3	B	291	LYS
3	B	295	LEU
3	B	306	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
3	B	237	ASN
3	B	272	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

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.

## 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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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	C	18/18 (100%)	-0.55	0 100 100	63, 79, 92, 97	0
2	D	18/18 (100%)	-0.64	0 100 100	59, 79, 86, 87	0
3	A	89/110 (80%)	-0.19	0 100 100	50, 63, 97, 100	0
3	B	85/110 (77%)	0.00	0 100 100	61, 80, 105, 110	0
All	All	210/256 (82%)	-0.18	0 100 100	50, 74, 99, 110	0

There are no RSRZ outliers to report.

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

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, 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	ZN	A	150	1/1	0.16	0.14	58,58,58,58	0
4	ZN	A	151	1/1	0.11	-0.39	58,58,58,58	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	ZN	B	351	1/1	0.11	-0.40	81,81,81,81	0
4	ZN	B	350	1/1	0.12	-0.52	71,71,71,71	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.