



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 13, 2017 – 04:29 pm GMT

PDB ID : 4DM8  
Title : Crystal structure of RARb LBD in complex with 9cis retinoic acid  
Authors : Osz, J.; Br livet, Y.; Peluso-Iltis, C.; Cura, V.; Eiler, S.; Ruff, M.; Bourguet, W.; Rochel, N.; Moras, D.  
Deposited on : 2012-02-07  
Resolution : 2.30 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

<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  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
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

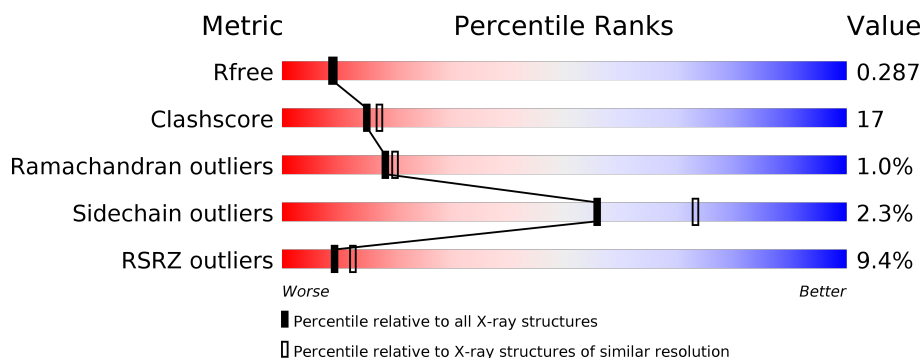
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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	4130 (2.30-2.30)
Clashscore	112137	4751 (2.30-2.30)
Ramachandran outliers	110173	4705 (2.30-2.30)
Sidechain outliers	110143	4704 (2.30-2.30)
RSRZ outliers	101464	4156 (2.30-2.30)

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	267	<div> <div>11%</div> <div> <div></div> <div>62%</div> <div>26%</div> <div>•</div> <div>11%</div> </div> </div>
1	B	267	<div> <div>6%</div> <div> <div></div> <div>64%</div> <div>24%</div> <div>•</div> <div>10%</div> </div> </div>
2	C	25	<div> <div>12%</div> <div>12%</div> <div>8%</div> <div>68%</div> </div>
2	D	25	<div> <div>16%</div> <div>16%</div> <div>68%</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	REA	A	501	-	-	-	X
3	REA	B	1501	-	-	-	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4083 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 Retinoic acid receptor beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	238	Total	C	N	O	S	36	0	0
			1879	1196	315	353	15			
1	B	239	Total	C	N	O	S	2	0	0
			1890	1204	317	354	15			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	157	MET	-	INITIATING METHIONINE	UNP P10826
A	158	GLY	-	EXPRESSION TAG	UNP P10826
A	159	SER	-	EXPRESSION TAG	UNP P10826
A	160	SER	-	EXPRESSION TAG	UNP P10826
A	161	HIS	-	EXPRESSION TAG	UNP P10826
A	162	HIS	-	EXPRESSION TAG	UNP P10826
A	163	HIS	-	EXPRESSION TAG	UNP P10826
A	164	HIS	-	EXPRESSION TAG	UNP P10826
A	165	HIS	-	EXPRESSION TAG	UNP P10826
A	166	HIS	-	EXPRESSION TAG	UNP P10826
A	167	SER	-	EXPRESSION TAG	UNP P10826
A	168	SER	-	EXPRESSION TAG	UNP P10826
A	169	GLY	-	EXPRESSION TAG	UNP P10826
A	170	LEU	-	EXPRESSION TAG	UNP P10826
A	171	VAL	-	EXPRESSION TAG	UNP P10826
A	172	PRO	-	EXPRESSION TAG	UNP P10826
A	173	ARG	-	EXPRESSION TAG	UNP P10826
A	174	GLY	-	EXPRESSION TAG	UNP P10826
A	175	SER	-	EXPRESSION TAG	UNP P10826
A	176	HIS	-	EXPRESSION TAG	UNP P10826
A	177	MET	-	EXPRESSION TAG	UNP P10826
B	1157	MET	-	INITIATING METHIONINE	UNP P10826
B	1158	GLY	-	EXPRESSION TAG	UNP P10826
B	1159	SER	-	EXPRESSION TAG	UNP P10826
B	1160	SER	-	EXPRESSION TAG	UNP P10826

*Continued on next page...*

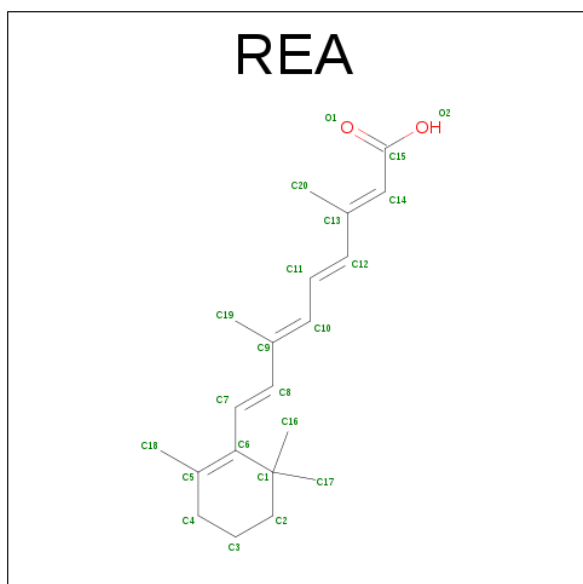
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	1161	HIS	-	EXPRESSION TAG	UNP P10826
B	1162	HIS	-	EXPRESSION TAG	UNP P10826
B	1163	HIS	-	EXPRESSION TAG	UNP P10826
B	1164	HIS	-	EXPRESSION TAG	UNP P10826
B	1165	HIS	-	EXPRESSION TAG	UNP P10826
B	1166	HIS	-	EXPRESSION TAG	UNP P10826
B	1167	SER	-	EXPRESSION TAG	UNP P10826
B	1168	SER	-	EXPRESSION TAG	UNP P10826
B	1169	GLY	-	EXPRESSION TAG	UNP P10826
B	1170	LEU	-	EXPRESSION TAG	UNP P10826
B	1171	VAL	-	EXPRESSION TAG	UNP P10826
B	1172	PRO	-	EXPRESSION TAG	UNP P10826
B	1173	ARG	-	EXPRESSION TAG	UNP P10826
B	1174	GLY	-	EXPRESSION TAG	UNP P10826
B	1175	SER	-	EXPRESSION TAG	UNP P10826
B	1176	HIS	-	EXPRESSION TAG	UNP P10826
B	1177	MET	-	EXPRESSION TAG	UNP P10826

- Molecule 2 is a protein called Nuclear receptor coactivator 1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	8	Total	C	N	O	0	0	0
			71	47	15	9			
2	D	8	Total	C	N	O	4	0	0
			71	47	15	9			

- Molecule 3 is RETINOIC ACID (three-letter code: REA) (formula: C<sub>20</sub>H<sub>28</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			22	20	2		
3	B	1	Total	C	O	0	0
			22	20	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	31	Total	O	0	0
			31	31		
4	B	91	Total	O	0	0
			91	91		
4	C	2	Total	O	0	0
			2	2		
4	D	4	Total	O	0	0
			4	4		



CYS	PRO	SER	SER	HIS	SER	SER	LEU	THR	GLU	ARG	HIS	K1631	L1632	H1633	H1634	R1635	L1636	L1637	Q1638	GLU	GLY	SER	PRO	SER
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------	-------	-------	-------	-------	-------	-------	-------	-----	-----	-----	-----	-----



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	58.48Å 83.62Å 108.78Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.30 29.24 – 2.20	Depositor EDS
% Data completeness (in resolution range)	(Not available) (20.00-2.30) 94.5 (29.24-2.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.22 (at 2.20Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.224 , 0.277 0.231 , 0.287	Depositor DCC
$R_{free}$ test set	1102 reflections (4.76%)	DCC
Wilson B-factor (Å <sup>2</sup> )	31.2	Xtriage
Anisotropy	0.361	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 52.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	4083	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.19% 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 [i](#)

### 5.1 Standard geometry [i](#)

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

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	0.31	0/1908	0.56	1/2576 (0.0%)
1	B	0.35	0/1921	0.57	0/2595
2	C	0.24	0/71	0.48	0/93
2	D	0.28	0/71	0.63	0/93
All	All	0.33	0/3971	0.56	1/5357 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	373	PRO	N-CA-CB	5.65	110.08	103.30

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	1879	0	1937	57	0
1	B	1890	0	1956	76	0
2	C	71	0	84	4	0
2	D	71	0	84	6	0
3	A	22	0	27	1	0
3	B	22	0	27	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	31	0	0	3	0
4	B	91	0	0	7	0
4	C	2	0	0	0	0
4	D	4	0	0	0	0
All	All	4083	0	4115	135	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 17.

The worst 5 of 135 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)
4:B:1601:HOH:O	2:D:1631:LYS:HB3	1.69	0.91
1:A:191:GLU:HG3	1:A:195:LYS:HE3	1.60	0.83
1:A:341:ARG:HH22	1:A:387:ARG:HH22	1.25	0.81
1:A:375:MET:O	1:A:379:ILE:HG12	1.81	0.80
1:B:1376:PHE:HB3	1:B:1377:PRO:HD3	1.65	0.78

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	234/267 (88%)	212 (91%)	18 (8%)	4 (2%)	11	9
1	B	237/267 (89%)	229 (97%)	7 (3%)	1 (0%)	38	47
2	C	6/25 (24%)	6 (100%)	0	0	100	100
2	D	6/25 (24%)	6 (100%)	0	0	100	100
All	All	483/584 (83%)	453 (94%)	25 (5%)	5 (1%)	18	20

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	213	ASN
1	A	215	SER
1	A	405	PRO
1	A	214	SER
1	B	1215	SER

### 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	211/238 (89%)	208 (99%)	3 (1%)	71	85
1	B	213/238 (90%)	209 (98%)	4 (2%)	62	78
2	C	8/24 (33%)	5 (62%)	3 (38%)	0	0
2	D	8/24 (33%)	8 (100%)	0	100	100
All	All	440/524 (84%)	430 (98%)	10 (2%)	56	73

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1318	LEU
1	B	1369	ARG
2	C	633	LEU
1	B	1186	LEU
1	B	1395	GLU

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

Mol	Chain	Res	Type
1	B	1198	GLN
1	B	1206	GLN
1	B	1316	ASN
1	A	413	GLN
1	B	1298	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	$\# Z  > 2$
3	REA	A	501	-	19,22,22	2.23	2 (10%)	26,30,30	1.39	4 (15%)
3	REA	B	1501	-	19,22,22	2.23	2 (10%)	26,30,30	1.59	5 (19%)

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
3	REA	A	501	-	-	0/13/32/32	0/1/1/1
3	REA	B	1501	-	-	0/13/32/32	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1501	REA	C5-C6	4.17	1.41	1.34

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	501	REA	C5-C6	4.18	1.41	1.34
3	A	501	REA	C1-C6	7.12	1.63	1.53
3	B	1501	REA	C1-C6	7.20	1.63	1.53

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	1501	REA	C10-C11-C12	-3.13	113.63	123.23
3	B	1501	REA	C19-C9-C10	-3.06	118.64	122.92
3	A	501	REA	C10-C11-C12	-2.77	114.74	123.23
3	B	1501	REA	C18-C5-C4	-2.55	108.62	113.45
3	A	501	REA	C18-C5-C4	-2.34	109.00	113.45

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
3	A	501	REA	1	0
3	B	1501	REA	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	235/267 (88%)	0.80	29 (12%) 5 7	24, 43, 70, 83	5 (2%)
1	B	239/267 (89%)	0.33	17 (7%) 17 22	15, 33, 56, 80	1 (0%)
2	C	8/25 (32%)	0.34	0 100 100	37, 42, 48, 57	0
2	D	8/25 (32%)	-0.04	0 100 100	27, 32, 40, 42	1 (12%)
All	All	490/584 (83%)	0.55	46 (9%) 9 12	15, 38, 66, 83	7 (1%)

The worst 5 of 46 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	205	CYS	7.5
1	B	1370	PRO	6.5
1	B	1205	CYS	5.9
1	A	178	GLU	5.5
1	B	1206	GLN	4.8

### 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, 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	REA	B	1501	22/22	0.88	0.31	3.19	25,33,41,47	0
3	REA	A	501	22/22	0.88	0.33	2.93	33,37,40,40	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.