wwPDB X-ray Structure Validation Summary Report (i)

Aug 2, 2023 – 04:34 AM EDT

PDB ID	:	1BY4
Title	:	STRUCTURE AND MECHANISM OF THE HOMODIMERIC ASSEMBLY OF THE RXR ON DNA
Authors	:	Zhao, Q.; Chasse, S.A.; Devarakonda, S.; Sierk, M.L.; Ahvazi, B.; Sigler, P.B.; Rastinejad, F.
Deposited on	:	1998-10-22
Resolution	:	2.10 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

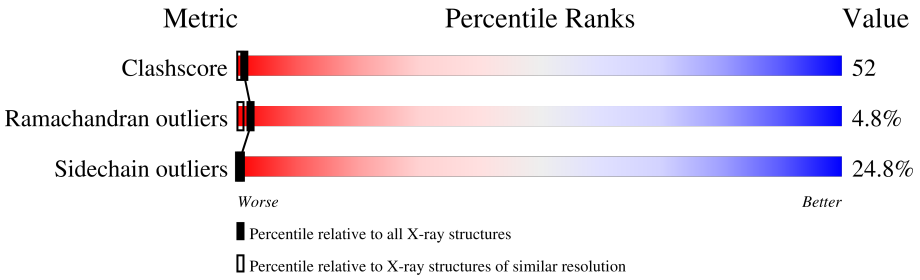
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.34

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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))
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)

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 of 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 <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	E	16	12% 62% 19% 6%
1	G	16	12% 50% 31% 6%
2	F	15	7% 67% 27%
2	H	15	93% 7%
3	A	82	27% 49% 17% • •
3	B	82	41% 50% 9%
3	C	82	17% 61% 15% • 5%
3	D	82	40% 43% 12% • •

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 4062 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 DNA (5'-D(*C*TP*AP*GP*GP*TP*CP*AP*AP*AP*GP*GP*TP*CP*AP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	15	Total	C	N	O	P	0	0	0
			310	148	62	86	14			
1	G	16	Total	C	N	O	P	0	0	0
			329	157	65	92	15			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	15	Total	C	N	O	P	0	0	0
			299	145	50	90	14			
2	H	15	Total	C	N	O	P	0	0	0
			299	145	50	90	14			

- Molecule 3 is a protein called PROTEIN (RETINOIC ACID RECEPTOR RXR-ALPHA).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	79	Total	C	N	O	S	0	0	0
			640	387	127	115	11			
3	B	82	Total	C	N	O	S	0	0	0
			667	405	131	120	11			
3	C	78	Total	C	N	O	S	0	0	0
			629	381	123	114	11			
3	D	80	Total	C	N	O	S	0	0	0
			651	397	126	117	11			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1128	GLY	ALA	conflict	UNP P19793
B	1228	GLY	ALA	conflict	UNP P19793

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	2128	GLY	ALA	conflict	UNP P19793
D	2228	GLY	ALA	conflict	UNP P19793

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

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

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	E	34	Total O 34 34	0	0
5	F	43	Total O 43 43	0	0
5	G	27	Total O 27 27	0	0
5	H	39	Total O 39 39	0	0
5	A	15	Total O 15 15	0	0
5	B	32	Total O 32 32	0	0
5	C	16	Total O 16 16	0	0
5	D	24	Total O 24 24	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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.

Note EDS was not executed.

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

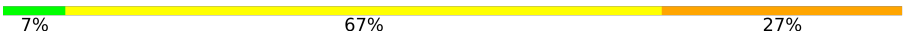
Chain E:

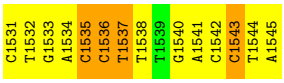
• Molecule 1: DNA (5'-D(*C*TP*AP*GP*GP*TP*CP*AP*AP*AP*GP*GP*TP*CP*AP*G)-3')

Chain G:

• Molecule 2: DNA (5'-D(*CP*TP*GP*AP*CP*CP*TP*TP*TP*GP*AP*CP*CP*TP*A)-3')

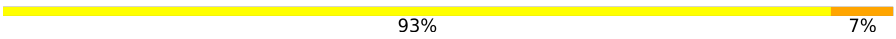
Chain F:

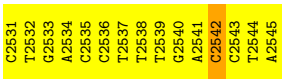




• Molecule 2: DNA (5'-D(*CP*TP*GP*AP*CP*CP*TP*TP*TP*GP*AP*CP*CP*TP*A)-3')

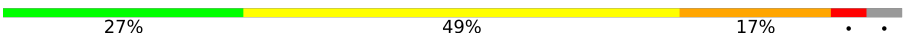
Chain H:

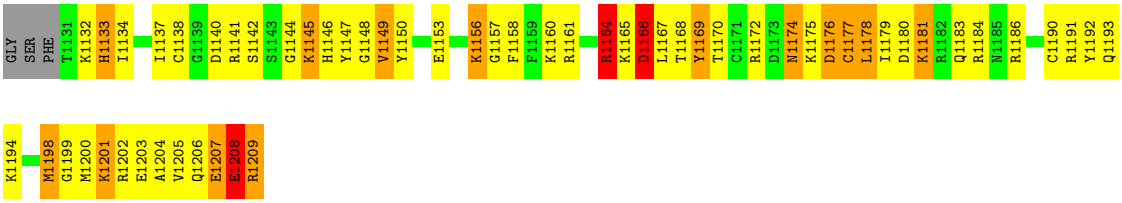




• Molecule 3: PROTEIN (RETINOIC ACID RECEPTOR RXR-ALPHA)

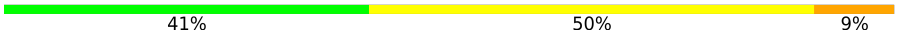
Chain A:





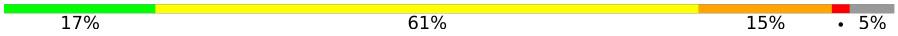
• Molecule 3: PROTEIN (RETINOIC ACID RECEPTOR RXR-ALPHA)

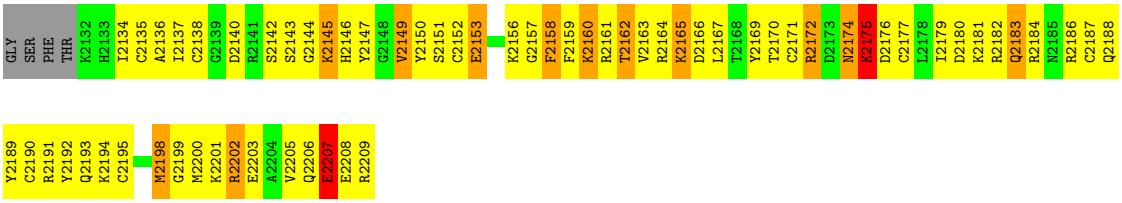
Chain B:



• Molecule 3: PROTEIN (RETINOIC ACID RECEPTOR RXR-ALPHA)

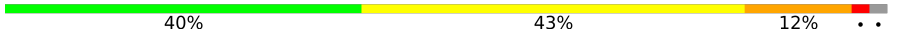
Chain C:

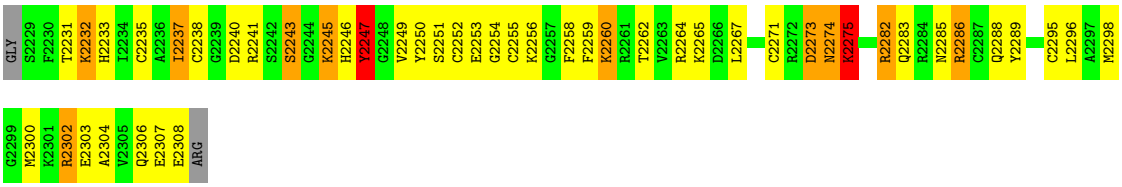




• Molecule 3: PROTEIN (RETINOIC ACID RECEPTOR RXR-ALPHA)

Chain D:





4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	70.93Å 33.65Å 101.40Å 90.00° 89.95° 90.00°	Depositor
Resolution (Å)	10.00 – 2.10	Depositor
% Data completeness (in resolution range)	72.3 (10.00-2.10)	Depositor
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR 3.851	Depositor
R, R_{free}	0.234 , 0.284	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4062	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

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	E	1.10	0/349	1.07	1/538 (0.2%)
1	G	1.27	0/370	1.20	4/570 (0.7%)
2	F	1.09	0/333	1.05	0/511
2	H	1.11	0/333	1.00	1/511 (0.2%)
3	A	0.85	0/648	0.98	0/857
3	B	0.92	0/676	0.99	0/893
3	C	0.85	0/637	1.08	2/843 (0.2%)
3	D	0.85	0/660	1.02	2/874 (0.2%)
All	All	0.98	0/4006	1.05	10/5597 (0.2%)

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	E	0	2
1	G	0	4
2	F	0	4
3	C	0	1
3	D	0	1
All	All	0	12

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	2144	GLY	N-CA-C	-6.38	97.14	113.10
2	H	2542	DC	C1'-O4'-C4'	-6.26	103.84	110.10
1	G	2494	DC	O4'-C1'-C2'	-6.18	100.95	105.90

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	G	2505	DG	O5'-P-OP1	-5.91	100.38	105.70
3	D	2247	TYR	N-CA-C	-5.73	95.53	111.00

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	1503	DA	Sidechain
1	E	1508	DA	Sidechain
2	F	1535	DC	Sidechain
2	F	1536	DC	Sidechain
2	F	1537	DT	Sidechain

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	E	310	0	170	38	0
1	G	329	0	181	33	0
2	F	299	0	172	29	0
2	H	299	0	172	32	0
3	A	640	0	619	69	0
3	B	667	0	652	60	0
3	C	629	0	606	105	0
3	D	651	0	636	50	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0
4	C	2	0	0	0	0
4	D	2	0	0	0	0
5	A	15	0	0	2	0
5	B	32	0	0	4	0
5	C	16	0	0	2	0
5	D	24	0	0	6	0
5	E	34	0	0	4	0
5	F	43	0	0	3	0
5	G	27	0	0	1	0
5	H	39	0	0	5	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	4062	0	3208	366	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 52.

The worst 5 of 366 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:H:2535:DC:H2"	2:H:2536:DC:H5"	1.13	1.09
3:C:2201:LYS:HA	3:C:2202:ARG:HH21	1.02	1.08
3:C:2202:ARG:NE	3:C:2202:ARG:H	1.54	1.04
3:C:2160:LYS:O	3:C:2164:ARG:HG2	1.57	1.04
2:F:1531:DC:H2"	2:F:1532:DT:H5'	1.38	1.01

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	77/82 (94%)	57 (74%)	13 (17%)	7 (9%)	1 0
3	B	80/82 (98%)	67 (84%)	9 (11%)	4 (5%)	2 0
3	C	76/82 (93%)	61 (80%)	12 (16%)	3 (4%)	3 1
3	D	78/82 (95%)	66 (85%)	11 (14%)	1 (1%)	12 7
All	All	311/328 (95%)	251 (81%)	45 (14%)	15 (5%)	2 0

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	1132	LYS
3	A	1166	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	B	1308	GLU
3	C	2166	ASP
3	D	2275	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	68/72 (94%)	45 (66%)	23 (34%)	0	0
3	B	72/72 (100%)	59 (82%)	13 (18%)	1	1
3	C	67/72 (93%)	50 (75%)	17 (25%)	0	0
3	D	71/72 (99%)	55 (78%)	16 (22%)	1	0
All	All	278/288 (96%)	209 (75%)	69 (25%)	0	0

5 of 69 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	2245	LYS
3	D	2253	GLU
3	D	2286	ARG
3	B	1241	ARG
3	B	1232	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	2283	GLN
3	D	2288	GLN
3	C	2183	GLN
3	C	2185	ASN
3	C	2188	GLN

5.3.3 RNA

There are no RNA molecules 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 monosaccharides in this entry.

5.6 Ligand geometry

Of 8 ligands modelled in this entry, 8 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

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.