



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 19, 2018 – 10:46 am GMT

PDB ID : 1HCQ  
Title : THE CRYSTAL STRUCTURE OF THE ESTROGEN RECEPTOR DNA-BINDING DOMAIN BOUND TO DNA: HOW RECEPTORS DISCRIMINATE BETWEEN THEIR RESPONSE ELEMENTS  
Authors : Schwabe, J.W.R.; Chapman, L.; Finch, J.T.; Rhodes, D.  
Deposited on : 1995-01-04  
Resolution : 2.40 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

<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  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30686

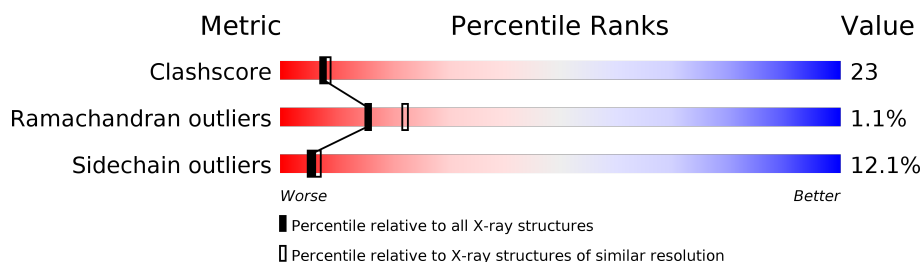
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	122078	3953 (2.40-2.40)
Ramachandran outliers	120005	3894 (2.40-2.40)
Sidechain outliers	119972	3895 (2.40-2.40)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	C	18	
1	G	18	
2	D	18	
2	H	18	
3	A	84	
3	B	84	
3	E	84	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	84	 <div> <div>26%</div> <div>45%</div> <div>13%</div> <div>15%</div> </div>

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 3864 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(\*CP\*CP\*AP\*GP\*GP\*TP\*CP\*AP\*CP\*AP\*GP\*TP\*GP\*AP\*CP\*CP\*T P\*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	18	Total	C	N	O	P	0	0	0
			365	174	69	105	17			
1	G	18	Total	C	N	O	P	0	0	0
			365	174	69	105	17			

- Molecule 2 is a DNA chain called DNA (5'-D(\*CP\*CP\*AP\*GP\*GP\*TP\*CP\*AP\*CP\*TP\*GP\*TP\*GP\*AP\*CP\*CP\*T P\*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	18	Total	C	N	O	P	0	0	0
			364	174	66	107	17			
2	H	18	Total	C	N	O	P	0	0	0
			364	174	66	107	17			

- Molecule 3 is a protein called PROTEIN (ESTROGEN RECEPTOR).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	74	Total	C	N	O	S	0	0	0
			572	350	107	103	12			
3	B	71	Total	C	N	O	S	0	0	0
			562	344	106	101	11			
3	E	74	Total	C	N	O	S	0	0	1
			565	346	106	101	12			
3	F	71	Total	C	N	O	S	0	0	0
			541	332	98	99	12			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	INITIAL METHIONINE	UNP P03372
B	1	MET	-	INITIAL METHIONINE	UNP P03372

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	1	MET	-	INITIAL METHIONINE	UNP P03372
F	1	MET	-	INITIAL METHIONINE	UNP P03372

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

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

- Molecule 5 is water.

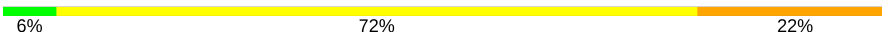
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	21	Total O 21 21	0	0
5	D	14	Total O 14 14	0	0
5	G	13	Total O 13 13	0	0
5	H	16	Total O 16 16	0	0
5	A	33	Total O 33 33	0	0
5	B	29	Total O 29 29	0	0
5	E	18	Total O 18 18	0	0
5	F	14	Total O 14 14	0	0

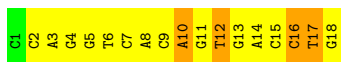
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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 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.

Note EDS was not executed.

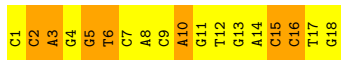
- Molecule 1: DNA (5'-D(\*CP\*CP\*AP\*GP\*GP\*TP\*CP\*AP\*CP\*AP\*GP\*TP\*GP\*AP\*CP\*CP\*T P\*G)-3')

Chain C: 



- Molecule 1: DNA (5'-D(\*CP\*CP\*AP\*GP\*GP\*TP\*CP\*AP\*CP\*AP\*GP\*TP\*GP\*AP\*CP\*CP\*T P\*G)-3')

Chain G: 



- Molecule 2: DNA (5'-D(\*CP\*CP\*AP\*GP\*GP\*TP\*CP\*AP\*CP\*TP\*GP\*TP\*GP\*AP\*CP\*CP\*T P\*G)-3')

Chain D: 



- Molecule 2: DNA (5'-D(\*CP\*CP\*AP\*GP\*GP\*TP\*CP\*AP\*CP\*TP\*GP\*TP\*GP\*AP\*CP\*CP\*T P\*G)-3')

Chain H: 

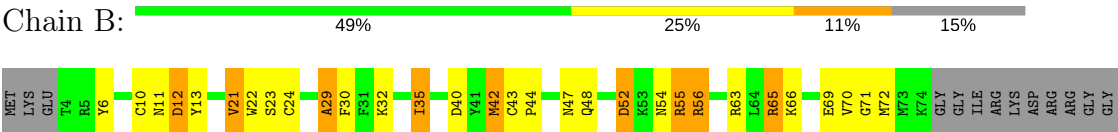


- Molecule 3: PROTEIN (ESTROGEN RECEPTOR)

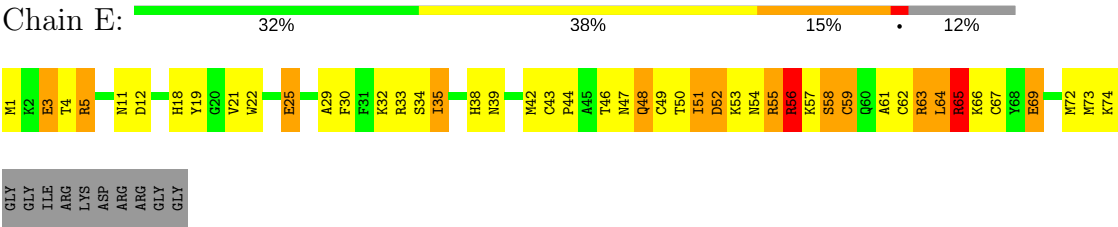
Chain A: 



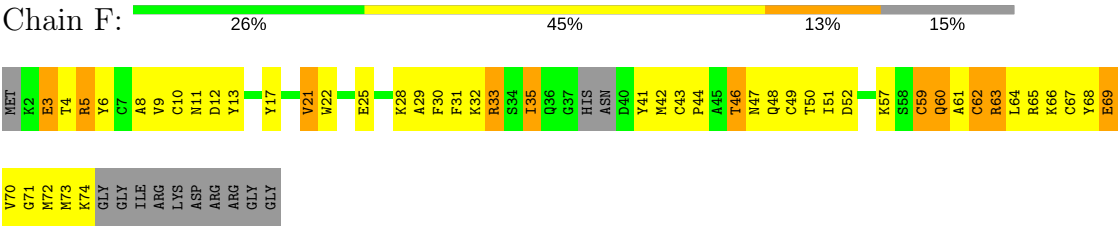
● Molecule 3: PROTEIN (ESTROGEN RECEPTOR)



● Molecule 3: PROTEIN (ESTROGEN RECEPTOR)



● Molecule 3: PROTEIN (ESTROGEN RECEPTOR)



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	51.60Å 90.80Å 114.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	(Not available) – 2.40	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-2.40)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	X-PLOR, TNT	Depositor
R, $R_{free}$	0.204 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3864	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP



## 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	2.30	18/409 (4.4%)	2.91	45/629 (7.2%)
1	G	2.10	14/409 (3.4%)	3.03	42/629 (6.7%)
2	D	2.04	10/407 (2.5%)	2.77	41/626 (6.5%)
2	H	1.94	8/407 (2.0%)	2.85	37/626 (5.9%)
3	A	1.16	3/583 (0.5%)	1.35	5/777 (0.6%)
3	B	1.10	1/574 (0.2%)	1.51	10/769 (1.3%)
3	E	1.16	3/575 (0.5%)	1.51	7/764 (0.9%)
3	F	1.04	3/551 (0.5%)	1.44	5/735 (0.7%)
All	All	1.60	60/3915 (1.5%)	2.22	192/5555 (3.5%)

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	1

The worst 5 of 60 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	10	DA	N9-C4	-9.64	1.32	1.37
1	C	10	DA	N3-C4	-8.73	1.29	1.34
1	C	6	DT	C5-C6	-8.50	1.28	1.34
1	C	13	DG	C5-C4	-8.15	1.32	1.38
3	E	69	GLU	CD-OE2	8.06	1.34	1.25

The worst 5 of 192 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	9	DC	C2-N1-C1'	-22.48	94.07	118.80
2	H	20	DC	C2-N1-C1'	21.59	142.55	118.80
1	C	16	DC	C2-N1-C1'	20.30	141.13	118.80
1	C	16	DC	C6-N1-C1'	-20.09	96.69	120.80
2	H	27	DC	C2-N1-C1'	-19.70	97.12	118.80

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	12	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	C	365	0	203	5	0
1	G	365	0	203	7	0
2	D	364	0	204	14	0
2	H	364	0	204	7	0
3	A	572	0	531	11	0
3	B	562	0	517	25	0
3	E	565	0	524	41	0
3	F	541	0	486	39	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0
4	E	2	0	0	0	0
4	F	2	0	0	0	0
5	A	33	0	0	1	0
5	B	29	0	0	2	0
5	C	21	0	0	3	0
5	D	14	0	0	2	0
5	E	18	0	0	1	0
5	F	14	0	0	0	0
5	G	13	0	0	1	0
5	H	16	0	0	1	0
All	All	3864	0	2872	138	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 23.

The worst 5 of 138 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)
3:E:38:HIS:CA	3:E:39:ASN:N	2.27	0.97
3:A:55:ARG:HD2	3:B:44:PRO:HA	1.50	0.92
3:A:65:ARG:HH11	3:A:65:ARG:HG2	1.38	0.88
2:D:19:DC:H6	2:D:19:DC:HO5'	1.21	0.88
2:D:33:DC:H2''	2:D:34:DC:C5'	2.11	0.81

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	72/84 (86%)	67 (93%)	5 (7%)	0	100	100
3	B	69/84 (82%)	61 (88%)	8 (12%)	0	100	100
3	E	70/84 (83%)	55 (79%)	13 (19%)	2 (3%)	5	4
3	F	67/84 (80%)	59 (88%)	7 (10%)	1 (2%)	11	15
All	All	278/336 (83%)	242 (87%)	33 (12%)	3 (1%)	16	22

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	E	56	ARG
3	E	58	SER
3	F	60	GLN

### 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	59/70 (84%)	54 (92%)	5 (8%)	12	18
3	B	59/70 (84%)	55 (93%)	4 (7%)	17	27
3	E	58/70 (83%)	48 (83%)	10 (17%)	2	2
3	F	55/70 (79%)	46 (84%)	9 (16%)	2	3
All	All	231/280 (82%)	203 (88%)	28 (12%)	5	7

5 of 28 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	E	48	GLN
3	E	56	ARG
3	F	46	THR
3	E	51	ILE
3	E	53	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
3	B	38	HIS
3	B	48	GLN
3	F	47	ASN
3	F	48	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 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 [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](#)

EDS was not executed - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section is therefore empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section is therefore empty.