



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 7, 2018 – 08:05 am GMT

PDB ID : 2QA8  
Title : Crystal Structure of the Estrogen Receptor Alpha Ligand Binding Domain Mutant 537S Complexed with Genistein  
Authors : Nettles, K.W.; Bruning, J.B.; Nowak, J.; Sharma, S.K.; Hahm, J.B.; Shi, Y.; Kulp, K.; Hochberg, R.B.; Zhou, H.; Katzenellenbogen, J.A.; Katzenellenbogen, B.S.; Kim, Y.; Joachmiak, A.; Greene, G.L.  
Deposited on : 2007-06-14  
Resolution : 1.85 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  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk30967  
Percentile statistics : (not set)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

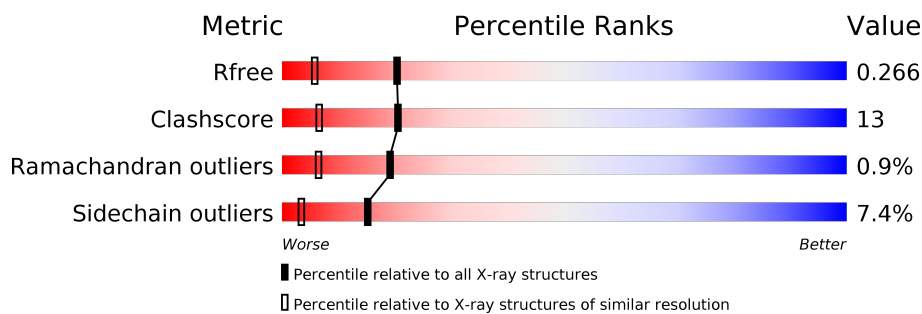
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.85 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))
$R_{free}$	111664	2111 (1.86-1.86)
Clashscore	122126	2258 (1.86-1.86)
Ramachandran outliers	120053	2234 (1.86-1.86)
Sidechain outliers	120020	2234 (1.86-1.86)

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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	258	
2	B	258	
3	C	13	
3	D	13	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 4015 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 Estrogen receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	230	Total	C	N	O	S	0	2	0
			1843	1178	314	330	21			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	297	SER	-	CLONING ARTIFACT	UNP P03372
A	530	CME	CYS	MODIFIED RESIDUE	UNP P03372
A	537	SER	TYR	ENGINEERED	UNP P03372

- Molecule 2 is a protein called Estrogen receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	235	Total	C	N	O	S	0	4	0
			1894	1215	322	336	21			

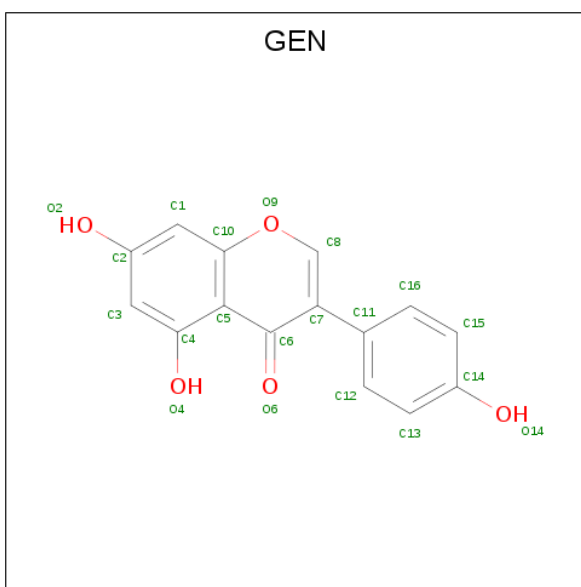
There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	297	SER	-	CLONING ARTIFACT	UNP P03372
B	381	CME	CYS	MODIFIED RESIDUE	UNP P03372
B	537	SER	TYR	ENGINEERED	UNP P03372

- Molecule 3 is a protein called nuclear receptor coactivator 2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	10	Total	C	N	O	0	0	0
			89	57	19	13			
3	D	9	Total	C	N	O	0	0	0
			79	51	16	12			

- Molecule 4 is GENISTEIN (three-letter code: GEN) (formula: C<sub>15</sub>H<sub>10</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			20	15	5		
4	B	1	Total	C	O	0	0
			20	15	5		

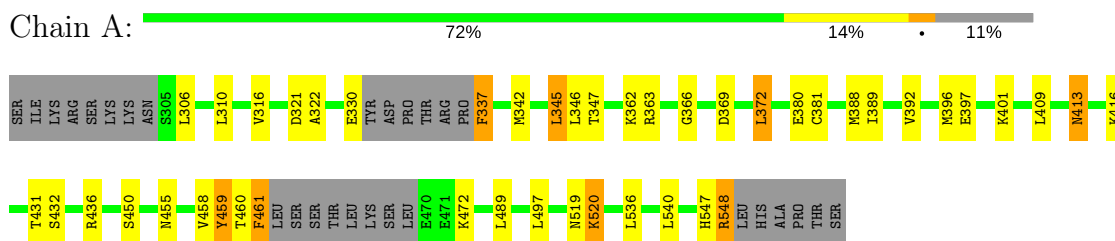
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	38	Total	O	0	0
			38	38		
5	B	32	Total	O	0	0
			32	32		

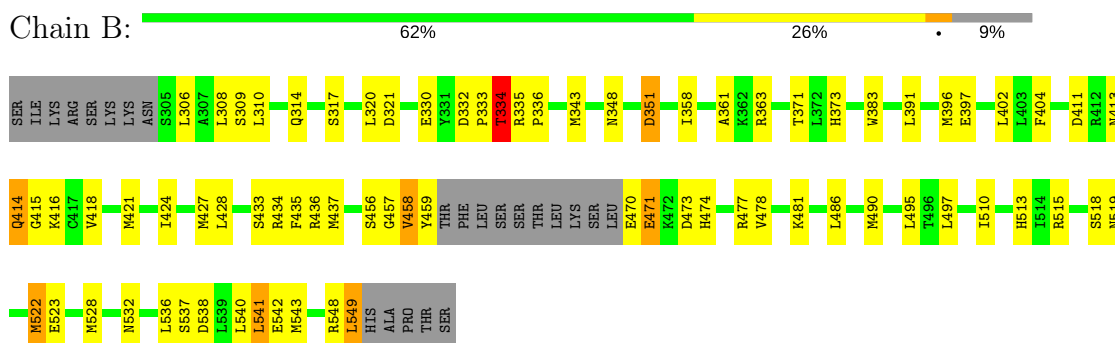
### 3 Residue-property plots

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

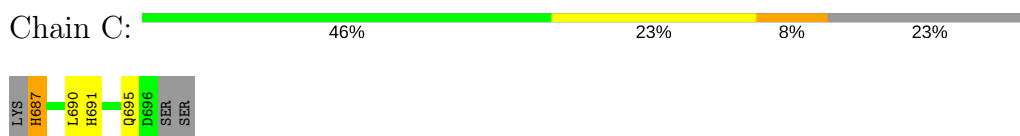
#### • Molecule 1: Estrogen receptor



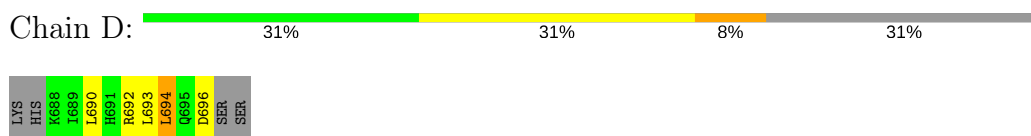
#### • Molecule 2: Estrogen receptor



#### • Molecule 3: nuclear receptor coactivator 2



#### • Molecule 3: nuclear receptor coactivator 2



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	55.67Å 77.96Å 58.00Å 90.00° 109.05° 90.00°	Depositor
Resolution (Å)	19.89 – 1.85 19.89 – 1.85	Depositor EDS
% Data completeness (in resolution range)	96.8 (19.89-1.85) 96.8 (19.89-1.85)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.18 (at 1.85Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.211 , 0.264 0.211 , 0.266	Depositor DCC
$R_{free}$ test set	1935 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	26.5	Xtriage
Anisotropy	0.237	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 61.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.035 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	4015	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.97% 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: CME, GEN

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	0.48	0/1870	0.66	0/2518
2	B	0.43	0/1933	0.60	0/2609
3	C	0.48	0/90	0.61	0/119
3	D	0.36	0/79	0.63	0/104
All	All	0.46	0/3972	0.63	0/5350

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	A	1843	0	1887	37	0
2	B	1894	0	1953	74	0
3	C	89	0	95	4	0
3	D	79	0	88	4	0
4	A	20	0	8	0	0
4	B	20	0	9	0	0
5	A	38	0	0	1	0
5	B	32	0	0	4	0
All	All	4015	0	4040	106	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 13.

The worst 5 of 106 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:B:391[B]:LEU:CD1	2:B:402[B]:LEU:HD22	1.57	1.35
1:A:547:HIS:O	1:A:548:ARG:HG3	1.60	1.00
1:A:519:ASN:ND2	2:B:519:ASN:HD22	1.60	1.00
1:A:519:ASN:HD22	2:B:519:ASN:ND2	1.61	0.99
2:B:391[B]:LEU:HD11	2:B:402[B]:LEU:CD2	1.94	0.97

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	224/258 (87%)	222 (99%)	1 (0%)	1 (0%)	36	21
2	B	234/258 (91%)	223 (95%)	8 (3%)	3 (1%)	13	3
3	C	8/13 (62%)	8 (100%)	0	0	100	100
3	D	7/13 (54%)	7 (100%)	0	0	100	100
All	All	473/542 (87%)	460 (97%)	9 (2%)	4 (1%)	19	8

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	334	THR
2	B	458	VAL
1	A	460	THR
2	B	336	PRO



### 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	
1	A	207/232 (89%)	190 (92%)	17 (8%)	12	2
2	B	214/232 (92%)	203 (95%)	11 (5%)	26	10
3	C	10/13 (77%)	8 (80%)	2 (20%)	1	0
3	D	9/13 (69%)	5 (56%)	4 (44%)	0	0
All	All	440/490 (90%)	406 (92%)	34 (8%)	15	3

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	520	LYS
2	B	334	THR
3	D	693	LEU
2	B	317	SER
1	A	345	LEU

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	A	519	ASN
2	B	398	HIS
2	B	532	ASN
1	A	498	GLN
2	B	513	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues 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$
1	CME	A	530	1	9,9,10	0.94	1 (11%)	6,9,11	1.40	1 (16%)
2	CME	B	381	2	9,9,10	0.77	0	6,9,11	1.57	2 (33%)

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
1	CME	A	530	1	-	0/5/8/10	0/0/0/0
2	CME	B	381	2	-	0/5/8/10	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	530	CME	CA-C	2.20	1.53	1.50

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	381	CME	CZ-CE-SD	-2.06	106.22	113.37
2	B	381	CME	CE-SD-SG	2.03	113.24	103.47
1	A	530	CME	CB-SG-SD	2.15	108.03	103.85

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

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
4	GEN	A	600	-	19,22,22	1.76	1 (5%)	27,32,32	1.11	2 (7%)
4	GEN	B	600	-	19,22,22	1.61	1 (5%)	27,32,32	1.35	5 (18%)

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
4	GEN	A	600	-	-	0/4/4/4	0/3/3/3
4	GEN	B	600	-	-	0/4/4/4	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	600	GEN	C6-C5	6.10	1.50	1.41
4	A	600	GEN	C6-C5	6.90	1.51	1.41

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	600	GEN	C1-C10-C5	-2.59	120.19	123.05
4	A	600	GEN	C1-C10-C5	-2.56	120.21	123.05
4	B	600	GEN	C16-C11-C7	-2.43	117.06	120.91
4	B	600	GEN	C7-C6-C5	-2.10	116.98	120.45
4	B	600	GEN	C4-C5-C10	2.04	120.30	117.90

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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.