



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:30 AM GMT

PDB ID : 2QZO
Title : Crystal Structure of the Estrogen Receptor Alpha Ligand Binding Domain
Complexed with WAY-169916
Authors : Bruning, J.B.; Gil, G.; Nowak, J.; Katzenellenbogen, J.; Nettles, K.W.
Deposited on : 2007-08-16
Resolution : 1.72 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
<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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : trunk26865

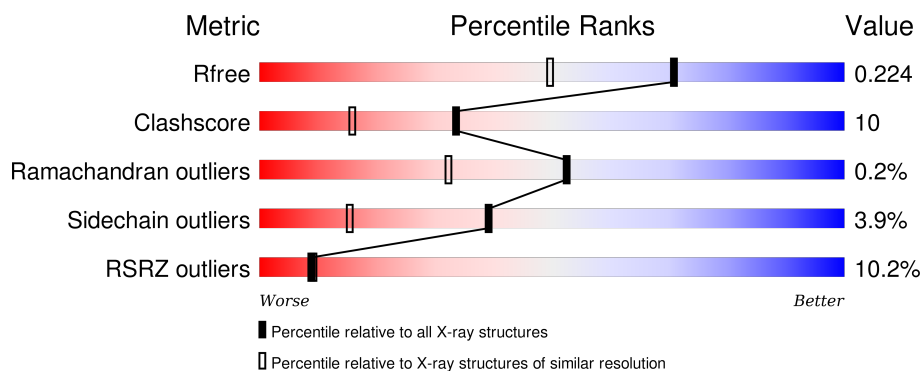
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.72 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}	91344	3998 (1.74-1.70)
Clashscore	102246	4425 (1.74-1.70)
Ramachandran outliers	100387	4360 (1.74-1.70)
Sidechain outliers	100360	4360 (1.74-1.70)
RSRZ outliers	91569	4010 (1.74-1.70)

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 ≥ 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	258	<div> <div>12%</div> <div> <div></div> <div>72%</div> <div>13%</div> <div>•</div> <div>12%</div> </div> </div>
2	B	258	<div> <div>5%</div> <div> <div></div> <div>80%</div> <div>10%</div> <div>•</div> <div>8%</div> </div> </div>
3	C	13	<div> <div>23%</div> <div> <div></div> <div>54%</div> <div>8%</div> <div>8%</div> <div>31%</div> </div> </div>
3	D	13	<div> <div></div> <div> <div></div> <div>46%</div> <div>23%</div> <div>31%</div> </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
2	CME	B	381	-	-	X	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 4364 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	228	Total	C	N	O	S	0	7	0
			1823	1168	313	323	19			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	297	SER	-	EXPRESSION TAG	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	237	Total	C	N	O	S	0	10	0
			1920	1237	319	341	23			

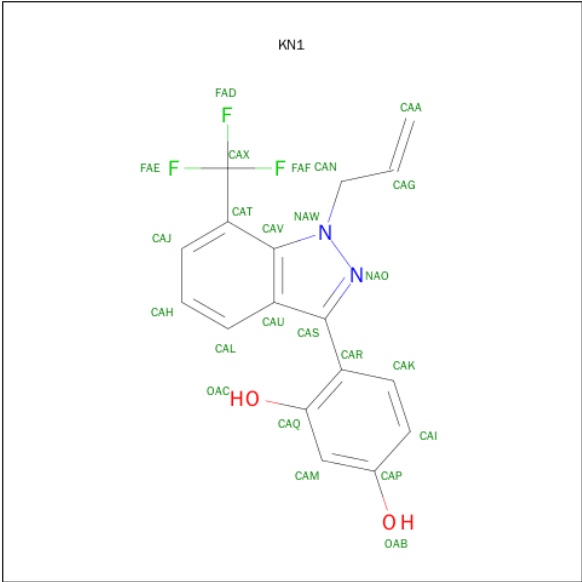
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	297	SER	-	EXPRESSION TAG	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	9	Total	C	N	O	0	0	0
			72	47	15	10			
3	D	9	Total	C	N	O	0	0	0
			79	51	16	12			

- Molecule 4 is 4-[1-ALLYL-7-(TRIFLUOROMETHYL)-1H-INDAZOL-3-YL]BENZENE-1,3-DIOL (three-letter code: KN1) (formula: C₁₇H₁₃F₃N₂O₂).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	F	N	O	0	0
			24	17	3	2	2		
4	B	1	Total	C	F	N	O	0	0
			24	17	3	2	2		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	178	Total	O	0	0
			178	178		
5	B	236	Total	O	0	0
			236	236		
5	C	1	Total	O	0	0
			1	1		
5	D	7	Total	O	0	0
			7	7		

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	55.81Å 82.03Å 58.52Å 90.00° 109.04° 90.00°	Depositor
Resolution (Å)	9.91 – 1.72 9.84 – 1.72	Depositor EDS
% Data completeness (in resolution range)	91.6 (9.91-1.72) 91.6 (9.84-1.72)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	0.05	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.85 (at 1.72Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.178 , 0.221 0.182 , 0.224	Depositor DCC
R_{free} test set	2446 reflections (5.35%)	DCC
Wilson B-factor (Å ²)	28.7	Xtriage
Anisotropy	0.062	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.43 , 70.2	EDS
Estimated twinning fraction	0.021 for l,-k,h	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	1 of 48123 reflections (0.002%)	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	4364	wwPDB-VP
Average B, all atoms (Å ²)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.26% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CME, KN1

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	0.57	0/1866	0.70	0/2521
2	B	0.66	0/1963	0.70	1/2652 (0.0%)
3	C	0.44	0/72	0.64	0/96
3	D	0.55	0/79	0.72	0/104
All	All	0.61	0/3980	0.70	1/5373 (0.0%)

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	A	0	6

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
2	B	436	ARG	NE-CZ-NH2	-5.06	117.77	120.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	332	ASP	Peptide
1	A	333	PRO	Peptide
1	A	334	THR	Peptide
1	A	339	GLU	Peptide
1	A	456	SER	Peptide

5.2 Too-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 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	1823	0	1858	51	0
2	B	1920	0	1981	33	0
3	C	72	0	75	3	0
3	D	79	0	88	2	0
4	A	24	0	11	5	0
4	B	24	0	12	0	0
5	A	178	0	0	10	0
5	B	236	0	0	6	0
5	C	1	0	0	0	0
5	D	7	0	0	0	0
All	All	4364	0	4025	81	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 10.

The worst 5 of 81 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:381:CME:HB2	2:B:381:CME:OH	1.26	1.29
1:A:513[B]:HIS:NE2	2:B:459:TYR:CE2	2.14	1.15
1:A:334:THR:HG23	1:A:335:ARG:CA	1.78	1.13
2:B:381:CME:CB	2:B:381:CME:OH	1.99	1.10
1:A:513[B]:HIS:NE2	2:B:459:TYR:HE2	1.50	1.09

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	
1	A	228/258 (88%)	224 (98%)	3 (1%)	1 (0%)	39	20
2	B	241/258 (93%)	238 (99%)	3 (1%)	0	100	100
3	C	7/13 (54%)	7 (100%)	0	0	100	100
3	D	7/13 (54%)	7 (100%)	0	0	100	100
All	All	483/542 (89%)	476 (99%)	6 (1%)	1 (0%)	52	32

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	333	PRO

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	201/232 (87%)	193 (96%)	8 (4%)	38	15
2	B	215/231 (93%)	208 (97%)	7 (3%)	45	21
3	C	7/13 (54%)	6 (86%)	1 (14%)	4	0
3	D	9/13 (69%)	8 (89%)	1 (11%)	8	1
All	All	432/489 (88%)	415 (96%)	17 (4%)	39	15

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	541	LEU
2	B	416	LYS
2	B	534[B]	VAL
1	A	532	ASN
2	B	541	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	373	HIS
1	A	474	HIS
2	B	398	HIS
2	B	513	HIS
3	C	691	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

3 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	381	1	8,9,10	0.82	0	6,9,11	2.16	3 (50%)
2	CME	B	381	2	8,9,10	0.87	0	6,9,11	2.08	1 (16%)
2	CME	B	530	2	8,9,10	0.79	0	6,9,11	1.34	0

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

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	381	CME	CZ-CE-SD	-3.11	105.55	113.16
1	A	381	CME	O-C-CA	-2.25	119.63	125.49
1	A	381	CME	CE-SD-SG	3.59	122.15	103.56
2	B	381	CME	CB-SG-SD	4.20	112.13	103.95

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	381	CME	1	0
2	B	381	CME	6	0

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
4	KN1	A	1	-	23,26,26	2.48	4 (17%)	28,39,39	4.01	11 (39%)
4	KN1	B	1	-	23,26,26	2.57	4 (17%)	28,39,39	2.16	9 (32%)

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	KN1	A	1	-	-	0/9/13/13	0/3/3/3
4	KN1	B	1	-	-	0/9/13/13	0/3/3/3

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1	KN1	CAR-CAS	-9.92	1.39	1.49
4	B	1	KN1	CAR-CAS	-9.64	1.39	1.49
4	B	1	KN1	CAN-CAG	-4.60	1.40	1.48
4	A	1	KN1	CAN-CAG	-4.44	1.40	1.48
4	A	1	KN1	CAS-NAO	-2.38	1.33	1.35

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1	KN1	FAD-CAX-CAT	-8.18	104.43	112.39
4	B	1	KN1	FAE-CAX-CAT	-5.90	106.64	112.39
4	A	1	KN1	FAE-CAX-CAT	-4.99	107.53	112.39
4	A	1	KN1	CAJ-CAT-CAX	-4.41	114.90	119.80
4	A	1	KN1	CAQ-CAR-CAS	-4.18	116.67	121.48

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1	KN1	5	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, 95th 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 ²)	Q<0.9
1	A	227/258 (87%)	0.49	32 (14%) 4 4	12, 22, 41, 47	4 (1%)
2	B	235/258 (91%)	-0.07	14 (5%) 25 27	12, 18, 36, 51	3 (1%)
3	C	9/13 (69%)	1.49	3 (33%) 0 0	18, 22, 25, 25	0
3	D	9/13 (69%)	0.30	0 100 100	22, 23, 29, 30	0
All	All	480/542 (88%)	0.23	49 (10%) 9 9	12, 19, 40, 51	7 (1%)

The worst 5 of 49 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	334	THR	11.9
1	A	549	LEU	10.8
1	A	469	LEU	8.3
2	B	306	LEU	6.3
1	A	458	VAL	5.6

6.2 Non-standard residues in protein, DNA, RNA chains [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, 95th 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(Å ²)	Q<0.9
2	CME	B	381	10/11	0.83	0.15	-	12,15,30,37	3
1	CME	A	381	10/11	0.89	0.14	-	17,18,38,42	0
2	CME	B	530	10/11	0.90	0.12	-	12,15,36,40	0

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, 95th 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(\AA^2)	Q<0.9
4	KN1	A	1	24/24	0.94	0.07	-0.67	30,39,46,49	0
4	KN1	B	1	24/24	0.97	0.05	-1.54	20,25,31,32	0

6.5 Other polymers [i](#)

There are no such residues in this entry.