



# Full wwPDB X-ray Structure Validation Report

Mar 9, 2018 – 05:31 pm GMT

PDB ID : 5GTO  
Title : Human PPARgamma ligand binding dmain complexed with S35  
Authors : Jang, J.Y.; Suh, S.W.  
Deposited on : 2016-08-22  
Resolution : 2.10 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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)  
Xtrriage (Phenix) : 1.13  
EDS : trunk30967  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
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

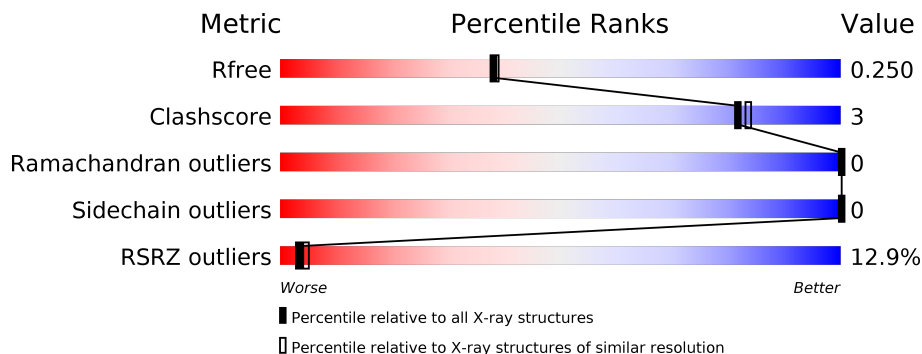
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 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))
$R_{free}$	111664	4608 (2.10-2.10)
Clashscore	122126	5109 (2.10-2.10)
Ramachandran outliers	120053	5059 (2.10-2.10)
Sidechain outliers	120020	5060 (2.10-2.10)
RSRZ outliers	108989	4497 (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 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.

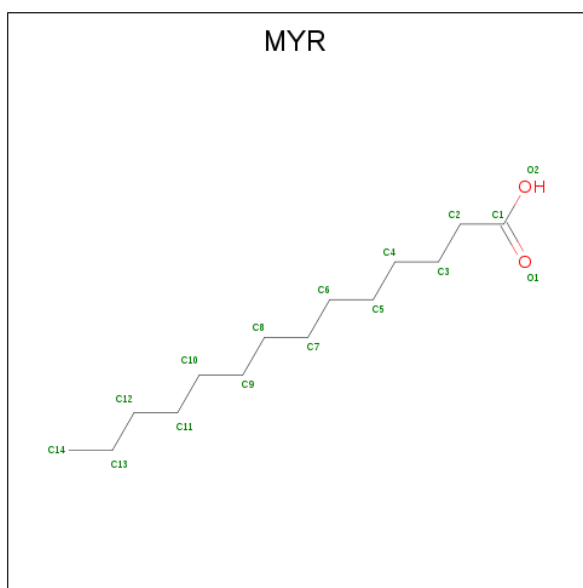
Mol	Chain	Length	Quality of chain
1	A	283	
2	B	16	

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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	MYR	A	502	-	-	-	X

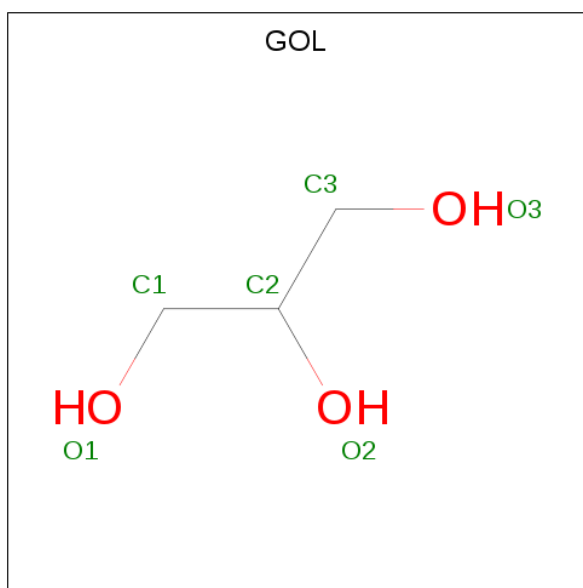


- Molecule 4 is MYRISTIC ACID (three-letter code: MYR) (formula:  $C_{14}H_{28}O_2$ ).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	
4	A	1	Total	C	O	0	0
			16	14	2		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	
5	A	1	Total	C	O	0	0
			6	3	3		

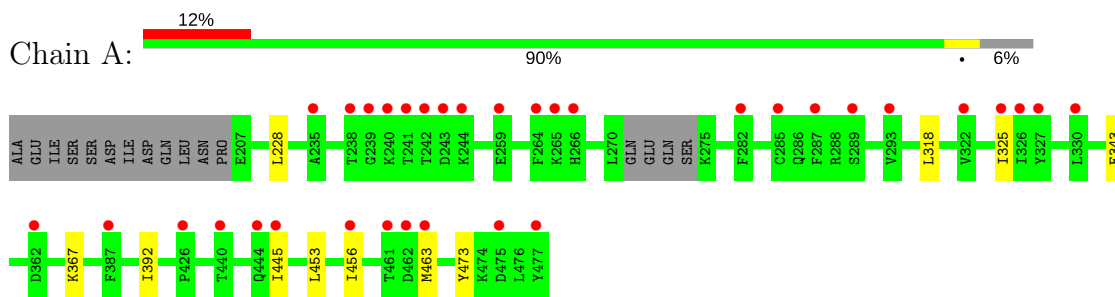
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	106	Total 106	O 106	0	0
6	B	2	Total 2	O 2	0	0

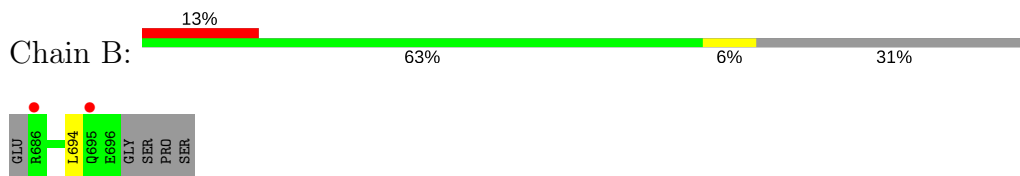
### 3 Residue-property plots [i](#)

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.

- Molecule 1: Peroxisome proliferator-activated receptor gamma



- Molecule 2: Nuclear receptor coactivator 1



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	131.82Å 53.29Å 53.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.10 28.67 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.5 (30.00-2.10) 99.5 (28.67-2.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.44 (at 2.10Å)	Xtrriage
Refinement program	REFMAC 5.8.0103	Depositor
R, $R_{free}$	0.211 , 0.242 0.218 , 0.250	Depositor DCC
$R_{free}$ test set	1119 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	41.4	Xtrriage
Anisotropy	0.213	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 60.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.029 for -h,l,k	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	2428	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	55.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.86% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: T35, GOL, MYR

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	0.50	0/2192	0.67	0/2952
2	B	0.47	0/102	0.77	0/134
All	All	0.50	0/2294	0.68	0/3086

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

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	2152	0	2226	11	0
2	B	101	0	110	1	0
3	A	45	0	0	1	0
4	A	16	0	27	2	0
5	A	6	0	8	0	0
6	A	106	0	0	0	0
6	B	2	0	0	0	0
All	All	2428	0	2371	13	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 3.

All (13) 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)
1:A:453:LEU:HD23	1:A:456:ILE:HD11	1.88	0.54
1:A:228:LEU:O	4:A:502:MYR:H21	2.09	0.52
1:A:453:LEU:CD1	1:A:473:TYR:CZ	2.95	0.50
1:A:456:ILE:HD13	1:A:463:MET:CE	2.43	0.49
1:A:325:ILE:HD11	1:A:392:ILE:HG13	1.97	0.46
1:A:367:LYS:CD	1:A:445:ILE:HD11	2.47	0.45
1:A:228:LEU:HD22	1:A:343:GLU:HA	1.99	0.45
1:A:456:ILE:HD13	1:A:463:MET:HE3	2.01	0.43
3:A:501:T35:OAH	3:A:501:T35:NAZ	2.52	0.42
1:A:325:ILE:HD11	1:A:392:ILE:CG1	2.50	0.42
1:A:318:LEU:HD11	2:B:694:LEU:HD21	2.03	0.41
1:A:367:LYS:HD2	1:A:445:ILE:HD11	2.03	0.41
4:A:502:MYR:H143	4:A:502:MYR:H112	1.90	0.40

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	264/283 (93%)	259 (98%)	5 (2%)	0	100	100
2	B	9/16 (56%)	9 (100%)	0	0	100	100
All	All	273/299 (91%)	268 (98%)	5 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	241/255 (94%)	241 (100%)	0	100	100
2	B	11/15 (73%)	11 (100%)	0	100	100
All	All	252/270 (93%)	252 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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](#)

3 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	T35	A	501	-	41,48,48	1.89	7 (17%)	49,67,67	1.71	7 (14%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	MYR	A	502	-	12,15,15	0.35	0	11,15,15	1.12	2 (18%)
5	GOL	A	503	-	5,5,5	0.36	0	5,5,5	0.31	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
3	T35	A	501	-	-	0/35/45/45	0/4/4/4
4	MYR	A	502	-	-	0/11/13/13	0/0/0/0
5	GOL	A	503	-	-	0/4/4/4	0/0/0/0

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	501	T35	CBN-CBO	-8.67	1.35	1.48
3	A	501	T35	CAX-CBI	-4.25	1.43	1.51
3	A	501	T35	CAV-CBO	-2.75	1.35	1.40
3	A	501	T35	CAV-CBP	-2.62	1.34	1.39
3	A	501	T35	CBJ-NAZ	-2.26	1.37	1.41
3	A	501	T35	OAH-CBH	2.08	1.24	1.21
3	A	501	T35	OBC-CBH	2.83	1.40	1.34

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	502	MYR	C7-C6-C5	-2.64	100.06	114.41
3	A	501	T35	CAS-CBJ-CAT	-2.20	116.47	119.56
3	A	501	T35	OBC-CBH-OAH	-2.11	120.45	124.86
4	A	502	MYR	C5-C4-C3	-2.01	103.52	114.41
3	A	501	T35	CAT-CBJ-NAZ	2.28	127.58	120.16
3	A	501	T35	OBC-CBQ-CBP	3.40	113.31	108.88
3	A	501	T35	CBS-OBQ-CBK	4.07	128.63	121.10
3	A	501	T35	NAZ-CBG-NBR	5.75	120.71	115.27
3	A	501	T35	CBQ-OBC-CBH	5.77	125.82	116.04

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	501	T35	1	0
4	A	502	MYR	2	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

### 6.1 Protein, DNA and RNA chains

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	267/283 (94%)	0.75	34 (12%) <b>3</b> <b>5</b>	29, 48, 92, 106	0
2	B	11/16 (68%)	0.95	2 (18%) <b>1</b> <b>1</b>	45, 58, 93, 101	0
All	All	278/299 (92%)	0.76	36 (12%) <b>3</b> <b>4</b>	29, 49, 92, 106	0

All (36) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	242	THR	6.9
1	A	239	GLY	6.3
2	B	686	ARG	5.5
1	A	477	TYR	5.2
1	A	240	LYS	4.5
1	A	461	THR	3.9
1	A	243	ASP	3.9
1	A	235	ALA	3.8
1	A	327	TYR	3.6
1	A	462	ASP	3.5
1	A	426	PRO	3.3
1	A	285	CYS	3.3
1	A	266	HIS	3.3
1	A	293	VAL	3.2
1	A	326	ILE	3.2
1	A	325	ILE	3.1
1	A	244	LYS	3.0
1	A	241	THR	3.0
1	A	322	VAL	2.9
1	A	330	LEU	2.8
1	A	475	ASP	2.8
1	A	362	ASP	2.7
1	A	264	PHE	2.7
1	A	259	GLU	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	265	LYS	2.2
1	A	287	PHE	2.2
1	A	387	PHE	2.2
2	B	695	GLN	2.2
1	A	463	MET	2.1
1	A	445	ILE	2.1
1	A	456	ILE	2.1
1	A	289	SER	2.1
1	A	440	THR	2.1
1	A	238	THR	2.1
1	A	444	GLN	2.1
1	A	282	PHE	2.0

## 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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
4	MYR	A	502	16/16	0.59	0.63	69,73,84,88	0
3	T35	A	501	45/45	0.80	0.22	60,72,93,100	0
5	GOL	A	503	6/6	0.91	0.51	67,73,75,76	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.