



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 21, 2020 – 03:02 pm BST

PDB ID : 3CTB  
Title : Tethered PXR-LBD/SRC-1p apoprotein  
Authors : Lesburg, C.A.  
Deposited on : 2008-04-11  
Resolution : 2.00 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)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

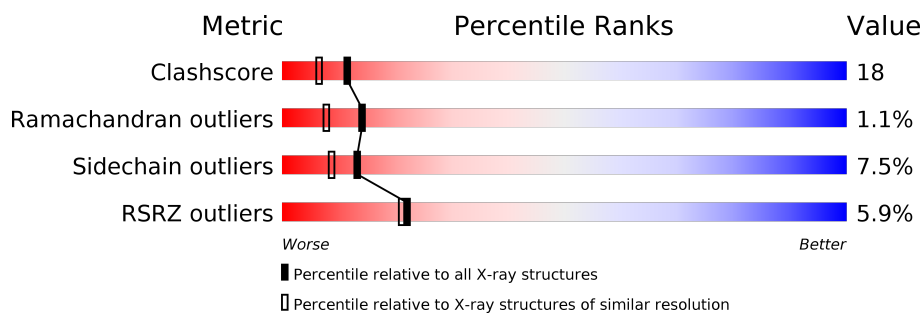
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 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  $\leq 5\%$ . 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	344	
1	B	344	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5078 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 Pregnane X receptor, Linker, Steroid receptor coactivator 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	287	Total	C	N	O	S	0	2	0
			2369	1520	410	420	19			
1	B	287	Total	C	N	O	S	0	1	0
			2353	1510	406	418	19			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	119	MET	-	EXPRESSION TAG	UNP O75469
A	120	LYS	-	EXPRESSION TAG	UNP O75469
A	121	LYS	-	EXPRESSION TAG	UNP O75469
A	122	GLY	-	EXPRESSION TAG	UNP O75469
A	123	HIS	-	EXPRESSION TAG	UNP O75469
A	124	HIS	-	EXPRESSION TAG	UNP O75469
A	125	HIS	-	EXPRESSION TAG	UNP O75469
A	126	HIS	-	EXPRESSION TAG	UNP O75469
A	127	HIS	-	EXPRESSION TAG	UNP O75469
A	128	HIS	-	EXPRESSION TAG	UNP O75469
A	129	GLY	-	EXPRESSION TAG	UNP O75469
A	435	GLY	-	LINKER	UNP O75469
A	436	GLY	-	LINKER	UNP O75469
A	437	SER	-	LINKER	UNP O75469
A	438	GLY	-	LINKER	UNP O75469
A	439	GLY	-	LINKER	UNP O75469
B	119	MET	-	EXPRESSION TAG	UNP O75469
B	120	LYS	-	EXPRESSION TAG	UNP O75469
B	121	LYS	-	EXPRESSION TAG	UNP O75469
B	122	GLY	-	EXPRESSION TAG	UNP O75469
B	123	HIS	-	EXPRESSION TAG	UNP O75469
B	124	HIS	-	EXPRESSION TAG	UNP O75469
B	125	HIS	-	EXPRESSION TAG	UNP O75469
B	126	HIS	-	EXPRESSION TAG	UNP O75469
B	127	HIS	-	EXPRESSION TAG	UNP O75469

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	128	HIS	-	EXPRESSION TAG	UNP O75469
B	129	GLY	-	EXPRESSION TAG	UNP O75469
B	435	GLY	-	LINKER	UNP O75469
B	436	GLY	-	LINKER	UNP O75469
B	437	SER	-	LINKER	UNP O75469
B	438	GLY	-	LINKER	UNP O75469
B	439	GLY	-	LINKER	UNP O75469

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	184	Total 184	O 184	0	0
2	B	172	Total 172	O 172	0	0

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 ( $\text{RSRZ} > 2$ ). 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.

- Chain A:
- 
- | Amino Acid | Percentage |
|------------|------------|
| GLN        | 54%        |
| GLY        | 25%        |
| ALA        | 4%         |
| PRO        | 17%        |

- Chain B:
- 
- 6% 60% 19% 5% 17%
- Q201 R204 D205 L206 C207 S208 L209 L213 Q214 L215 R216 V222 W223 N224 P228 G232 G233 R234 E235 I236 F237 L239 M246 T254 R265 E270 A279 L283 R287 T296 G297 T298 R299 E300 G301 R303 R304 S305 Y306 C307 L308 E309 ASP THR A
- GLY GLY PHE GLN Q317 L318 L319 L320 E321 P322 M323 L324 K325 K332 F349 Q358 H359 R360 V361 Q364 L365 Q366 R381 P382 Q383 P384 A385 H386 R387 F388 L389 R410 L411 L417 F420 E427 G430 I431 THR GLY SER GLY GLY SER GLY GLY SER HIS SER
- S444 L445 T446 E447 R448 H449 K450 I451 L452 H453 S460 PRO SER

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	85.24Å 89.24Å 105.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.31 – 2.00 45.32 – 2.00	Depositor EDS
% Data completeness (in resolution range)	98.5 (45.31-2.00) 96.4 (45.32-2.00)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	0.04	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.82 (at 2.00Å)	Xtriage
Refinement program	BUSTER-TNT 1.3.1	Depositor
R, $R_{free}$	0.204 , 0.244 0.204 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	39.3	Xtriage
Anisotropy	0.095	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 56.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	5078	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.74% 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](#)

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	0.66	0/2419	0.70	5/3253 (0.2%)
1	B	0.64	1/2402 (0.0%)	0.70	0/3229
All	All	0.65	1/4821 (0.0%)	0.70	5/6482 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	159	MET	SD-CE	-7.58	1.35	1.77

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	159	MET	CG-SD-CE	-8.91	85.94	100.20
1	A	383	GLN	C-N-CD	-6.23	106.89	120.60
1	A	396	MET	CG-SD-CE	-5.48	91.42	100.20
1	A	308	LEU	CB-CG-CD1	5.07	119.61	111.00
1	A	239	LEU	CB-CG-CD1	5.01	119.52	111.00

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	2369	0	2387	101	0
1	B	2353	0	2374	72	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	184	0	0	11	0
2	B	172	0	0	11	0
All	All	5078	0	4761	167	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 18.

The worst 5 of 167 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:B:170:LYS:HD3	1:B:303:ARG:HH22	1.12	1.09
1:B:317:GLN:HG2	1:B:318:LEU:H	0.93	1.04
1:A:383:GLN:HB2	1:A:384:PRO:HD3	1.38	1.02
1:B:317:GLN:HG2	1:B:318:LEU:N	1.72	0.98
1:A:383:GLN:CB	1:A:384:PRO:HD3	2.01	0.89

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	281/344 (82%)	272 (97%)	6 (2%)	3 (1%)	14	8
1	B	280/344 (81%)	267 (95%)	10 (4%)	3 (1%)	14	8
All	All	561/688 (82%)	539 (96%)	16 (3%)	6 (1%)	14	8

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	383	GLN
1	A	384	PRO

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	B	193	ARG
1	A	234	LYS
1	B	302	GLY

### 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	262/302 (87%)	245 (94%)	17 (6%)	17	12
1	B	260/302 (86%)	238 (92%)	22 (8%)	10	6
All	All	522/604 (86%)	483 (92%)	39 (8%)	13	9

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	159	MET
1	B	200	SER
1	B	410	ARG
1	B	178	LEU
1	B	192	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	449	HIS
1	A	453	HIS
1	B	383	GLN
1	A	418	HIS
1	B	386	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](#)

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](#)

There are no ligands in this entry.

## 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, 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	287/344 (83%)	0.21	13 (4%) 33 32	29, 46, 81, 123	0
1	B	287/344 (83%)	0.34	21 (7%) 15 14	30, 47, 87, 143	0
All	All	574/688 (83%)	0.27	34 (5%) 22 21	29, 47, 85, 143	0

The worst 5 of 34 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	209	LEU	7.1
1	B	193	ARG	5.5
1	B	192	SER	5.1
1	B	194	GLU	5.0
1	A	177	VAL	4.3

### 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](#)

There are no ligands in this entry.

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

There are no such residues in this entry.