



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 12:52 AM GMT

PDB ID : 2C0M  
Title : APO FORM OF THE TPR DOMAIN OF THE PEX5P RECEPTOR  
Authors : Stanley, W.A.; Kursula, P.; Wilmanns, M.  
Deposited on : 2005-09-05  
Resolution : 2.50 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  
<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



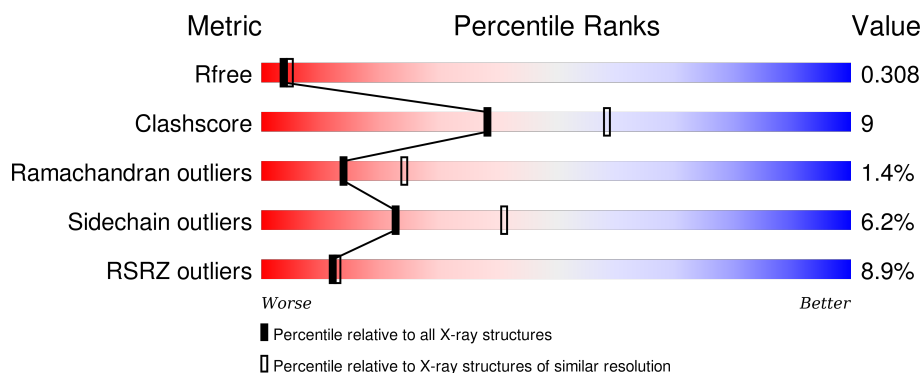
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (2.50-2.50)

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	319	<div> <div> <div>0%</div> <div>65%</div> <div>25%</div> <div>7%</div> </div> </div>
1	B	319	<div> <div>2%</div> <div>71%</div> <div>20%</div> <div>7%</div> </div>
1	C	319	<div> <div>16%</div> <div>76%</div> <div>16%</div> <div>5%</div> </div>
1	F	319	<div> <div>15%</div> <div>77%</div> <div>15%</div> <div>5%</div> </div>



## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 9629 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 PEROXISOMAL TARGETING SIGNAL 1 RECEPTOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	297	Total	C	N	O	S	0	3	0
			2364	1482	418	453	11			
1	B	297	Total	C	N	O	S	0	4	0
			2374	1489	419	455	11			
1	C	302	Total	C	N	O	S	0	1	0
			2377	1493	417	456	11			
1	F	302	Total	C	N	O	S	0	0	0
			2368	1488	416	453	11			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	388	ILE	THR	CONFLICT	UNP P50542
B	388	ILE	THR	CONFLICT	UNP P50542
C	388	ILE	THR	CONFLICT	UNP P50542
F	388	ILE	THR	CONFLICT	UNP P50542

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	54	Total	O	0	0
			54	54		
2	B	57	Total	O	0	0
			57	57		
2	C	14	Total	O	0	0
			14	14		
2	F	21	Total	O	0	0
			21	21		



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

65% 25% 7%

Residue	State	Percentage
D284	Green	65%
L285	Green	65%
I287	Green	65%
Y287	Green	65%
E291	Green	65%
E292	Green	65%
H298	Green	65%
P299	Green	65%
Q300	Green	65%
F301	Green	65%
E303	Green	65%
E304	Green	65%
R307	Green	65%
R308	Green	65%
L309	Green	65%
Q310	Green	65%
E311	Green	65%
G312	Green	65%
N313	Green	65%
L314	Green	65%
P315	Green	65%
N316	Green	65%
A317	Green	65%
F321	Green	65%
Q336	Green	65%
E344	Green	65%
R356	Green	65%
P363	Green	65%
D364	Green	65%
N365	Green	65%
T377	Green	65%
Q382	Green	65%
L393	Green	65%
R394	Green	65%
Y395	Green	65%
T396	Green	65%
A398	Green	65%
Y399	Green	65%
A400	Green	65%
V403	Green	65%
R404	Green	65%
P405	Green	65%
A410	Green	65%
G411	Green	65%
G412	Green	65%
G413	Green	65%
G414	Green	65%
G415	Green	65%
G416	Green	65%
G417	Green	65%
G418	Green	65%
G419	Green	65%
G420	Green	65%
G421	Green	65%
G422	Green	65%
G423	Green	65%
G424	Green	65%
G425	Green	65%
G426	Green	65%
G427	Green	65%
G428	Green	65%
G429	Green	65%
G430	Green	65%
G431	Green	65%
G432	Green	65%
G433	Green	65%
G434	Green	65%
G435	Green	65%
G436	Green	65%
G437	Green	65%
G438	Green	65%
G439	Green	65%
G440	Green	65%
D445	Green	65%
S448	Green	65%
V459	Green	65%
L460	Green	65%
L463	Green	65%
C473	Green	65%
L486	Green	65%
L487	Green	65%
G492	Green	65%
E502	Green	65%
V505	Green	65%
Y508	Green	65%
R509	Green	65%
Q515	Green	65%
I519	Green	65%
R520	Green	65%
S521	Green	65%
R522	Green	65%
Y523	Green	65%
N524	Green	65%
L525	Green	65%
G526	Green	65%
I527	Green	65%
S528	Green	65%
C529	Green	65%
H535	Green	65%
R536	Green	65%
E537	Green	65%
A538	Green	65%
V539	Green	65%
F542	Green	65%
L543	Green	65%
E544	Green	65%
N547	Green	65%
Y546	Green	65%
Q549	Green	65%
R550	Green	65%
K551	Green	65%
S552	Green	65%
R553	Green	65%
G554	Green	65%
P555	Green	65%
R556	Green	65%
G557	Green	65%
E558	Green	65%
G559	Green	65%
W567	Green	65%
S568	Green	65%
R571	Green	65%
L572	Green	65%
M576	Green	65%
L577	Green	

Chain B:

2% 71% 20% 7%

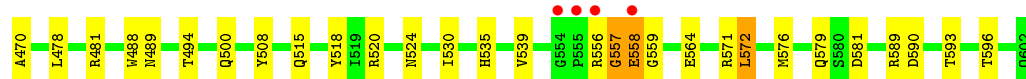
D284 D285 Q296 E291 E292 H298 F299 Q300 F302 E303 E304 G305 L306 R307 R308 A317 L320 F321 Q336 L359 R365 Q366 F367 A368 L369 F377 Q382 R383 E387 R390 V403 T404 P405 A415 GLU GLY ALA GLY GLY ALA GLY LEU GLY PRO SER TVS

**Chain C:**

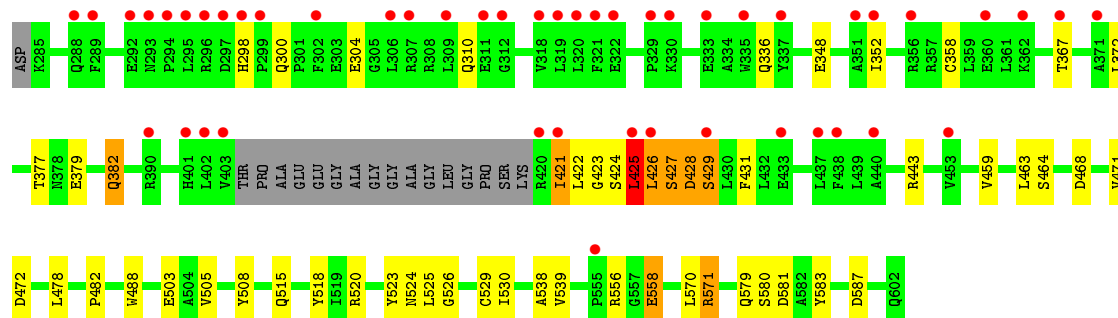
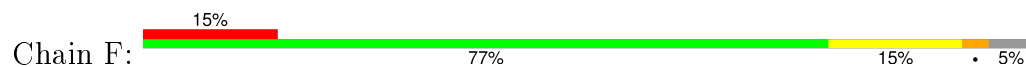
ASP K285 Q286 Y287 Q288 E291 E292 N293 D294 L295 R296 D297 H298 P299 Q300 P301 F302 E303 E304 G305 L306 R307 R308 L309 L314 P315 V318 L319 L320 Q326 Q327 R330 W335 Q336 Y337 E348 I352 S353 R356 R357 C358 L359 D364 T367 L372 F376 T377

16% 76% 16% 5%





● Molecule 1: PEROXISOMAL TARGETING SIGNAL 1 RECEPTOR





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	53.47Å 85.55Å 88.89Å 71.17° 89.99° 73.43°	Depositor
Resolution (Å)	19.54 – 2.50 19.53 – 2.50	Depositor EDS
% Data completeness (in resolution range)	96.5 (19.54-2.50) 79.1 (19.53-2.50)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.31 (at 2.50Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.263 , 0.309 0.296 , 0.308	Depositor DCC
$R_{free}$ test set	2363 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	27.9	Xtriage
Anisotropy	0.782	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , -4.0	EDS
Estimated twinning fraction	0.229 for h,h-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtriage
Outliers	0 of 47256 reflections	Xtriage
$F_o, F_c$ correlation	0.84	EDS
Total number of atoms	9629	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.75% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.



## 5 Model quality

### 5.1 Standard geometry

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	0.86	3/2409 (0.1%)	0.85	6/3265 (0.2%)
1	B	0.69	1/2419 (0.0%)	0.73	1/3278 (0.0%)
1	C	0.50	0/2421	0.62	0/3280
1	F	0.51	1/2412 (0.0%)	0.66	1/3268 (0.0%)
All	All	0.66	5/9661 (0.1%)	0.72	8/13091 (0.1%)

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	B	0	1
1	C	0	1
All	All	0	2

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	473	CYS	CB-SG	-6.50	1.71	1.82
1	A	544	GLU	CG-CD	6.45	1.61	1.51
1	A	529	CYS	CB-SG	-6.00	1.72	1.82
1	A	473	CYS	CB-SG	-5.80	1.72	1.81
1	F	529	CYS	CB-SG	-5.15	1.73	1.81

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	308	ARG	NE-CZ-NH1	9.72	125.16	120.30
1	A	308	ARG	NE-CZ-NH2	-6.19	117.21	120.30
1	A	486	LEU	CA-CB-CG	5.67	128.34	115.30
1	F	429	SER	N-CA-CB	-5.60	102.10	110.50
1	A	308	ARG	CG-CD-NE	5.38	123.10	111.80



There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	557	GLY	Peptide
1	C	425	LEU	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	2364	0	2312	64	0
1	B	2374	0	2321	47	0
1	C	2377	0	2327	29	0
1	F	2368	0	2322	33	0
2	A	54	0	0	4	0
2	B	57	0	0	0	0
2	C	14	0	0	2	0
2	F	21	0	0	1	0
All	All	9629	0	9282	171	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 9.

The worst 5 of 171 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:F:426:LEU:CD1	1:F:427:SER:H	1.59	1.13
1:F:426:LEU:HD13	1:F:427:SER:H	1.35	0.91
1:F:426:LEU:HD12	1:F:427:SER:H	1.37	0.90
1:C:572:LEU:HD22	1:C:576:MET:HE3	1.54	0.90
1:F:426:LEU:CD1	1:F:427:SER:N	2.41	0.82

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	296/319 (93%)	263 (89%)	29 (10%)	4 (1%)	14	24
1	B	297/319 (93%)	276 (93%)	21 (7%)	0	100	100
1	C	299/319 (94%)	275 (92%)	18 (6%)	6 (2%)	9	15
1	F	298/319 (93%)	269 (90%)	23 (8%)	6 (2%)	9	15
All	All	1190/1276 (93%)	1083 (91%)	91 (8%)	16 (1%)	14	26

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	287	TYR
1	C	557	GLY
1	C	558	GLU
1	F	425	LEU
1	F	427	SER

### 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	248/258 (96%)	232 (94%)	16 (6%)	21	39
1	B	249/258 (96%)	232 (93%)	17 (7%)	20	36
1	C	249/258 (96%)	237 (95%)	12 (5%)	31	55
1	F	248/258 (96%)	232 (94%)	16 (6%)	21	39
All	All	994/1032 (96%)	933 (94%)	61 (6%)	23	42



5 of 61 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	515	GLN
1	C	292	GLU
1	F	530	ILE
1	B	520	ARG
1	B	579	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	331	HIS
1	C	298	HIS
1	F	336	GLN
1	B	326	GLN
1	F	489	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

There are no ligands in this entry.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.



## 5.8 Polymer linkage issues ⓘ

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	297/319 (93%)	0.39	3 (1%) 84 86	13, 22, 26, 29	0
1	B	297/319 (93%)	0.42	6 (2%) 68 72	17, 23, 27, 31	0
1	C	302/319 (94%)	0.92	50 (16%) 2 2	20, 23, 25, 28	1 (0%)
1	F	302/319 (94%)	0.85	48 (15%) 3 2	20, 23, 25, 28	0
All	All	1198/1276 (93%)	0.65	107 (8%) 12 13	13, 23, 26, 31	1 (0%)

The worst 5 of 107 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	320	LEU	6.2
1	C	403	VAL	5.8
1	C	337	TYR	5.6
1	F	402	LEU	5.4
1	C	335	TRP	4.9

### 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.