



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 12:37 PM GMT

PDB ID : 2CE9  
Title : A WRPW PEPTIDE BOUND TO THE GROUCHO-TLE WD40 DOMAIN.  
Authors : Pickles, L.M.; Roe, S.M.; Pearl, L.H.  
Deposited on : 2006-02-03  
Resolution : 2.12 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

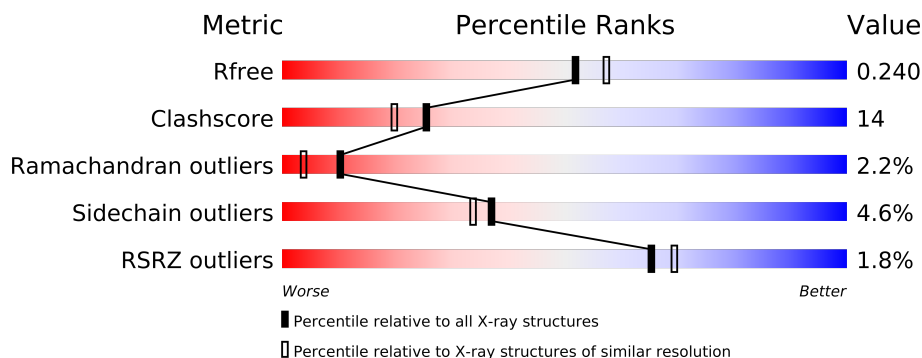
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.12 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))
$R_{free}$	66092	3409 (2.14-2.10)
Clashscore	79885	4090 (2.14-2.10)
Ramachandran outliers	78287	4048 (2.14-2.10)
Sidechain outliers	78261	4049 (2.14-2.10)
RSRZ outliers	66119	3410 (2.14-2.10)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	337	
1	B	337	
1	C	337	
1	D	337	
2	X	5	
2	Y	5	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 11285 atoms, of which 0 are hydrogen and 0 are deuterium.

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 TRANSDUCIN-LIKE ENHANCER PROTEIN 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	337	Total	C	N	O	S	0	0	0
			2587	1628	445	497	17			
1	B	337	Total	C	N	O	S	0	0	0
			2587	1628	445	497	17			
1	C	337	Total	C	N	O	S	0	0	0
			2587	1628	445	497	17			
1	D	332	Total	C	N	O	S	0	0	0
			2554	1610	439	488	17			

- Molecule 2 is a protein called WRPW PEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	X	5	Total	C	N	O	S	0	0	0
			55	38	10	6	1			
2	Y	5	Total	C	N	O	S	0	0	0
			54	38	10	5	1			

- Molecule 3 is water.

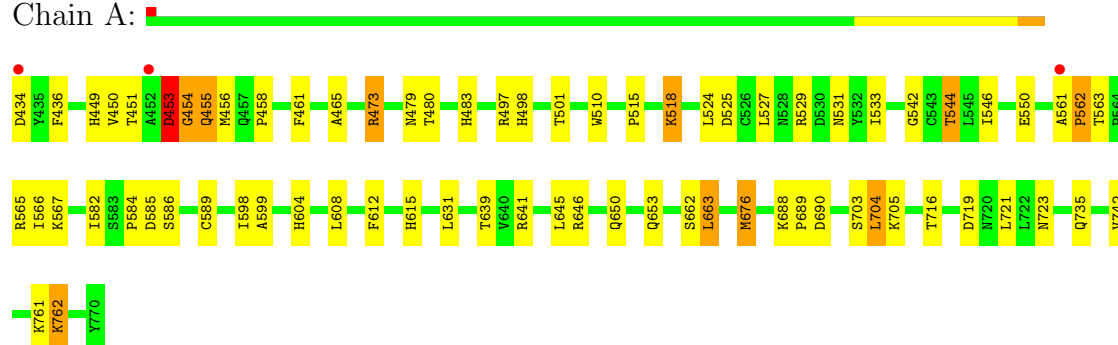
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	246	Total	O	0	0
			246	246		
3	B	228	Total	O	0	0
			228	228		
3	C	206	Total	O	0	0
			206	206		
3	D	176	Total	O	0	0
			176	176		
3	X	3	Total	O	0	0
			3	3		
3	Y	2	Total	O	0	0
			2	2		

### 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 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 ( $RSRZ > 2$ ). 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.

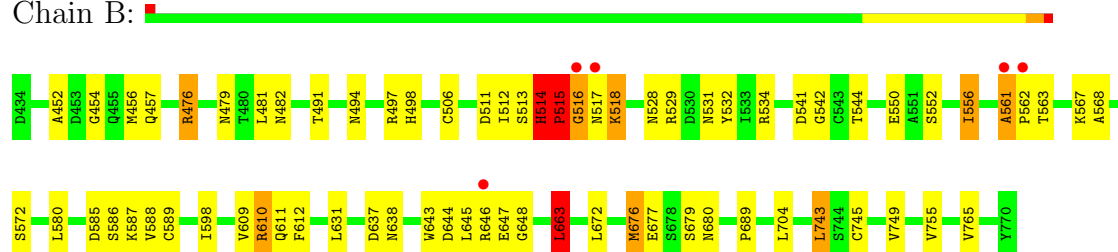
#### • Molecule 1: TRANSDUCIN-LIKE ENHANCER PROTEIN 1

Chain A:



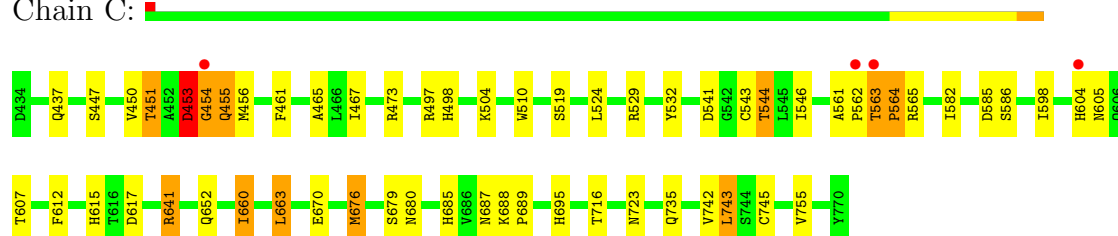
#### • Molecule 1: TRANSDUCIN-LIKE ENHANCER PROTEIN 1

Chain B:



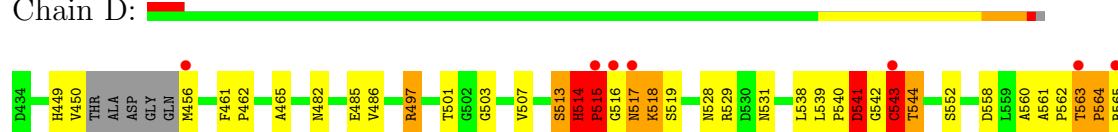
#### • Molecule 1: TRANSDUCIN-LIKE ENHANCER PROTEIN 1

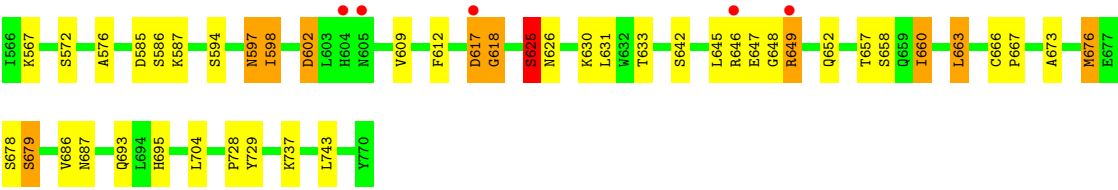
Chain C:



#### • Molecule 1: TRANSDUCIN-LIKE ENHANCER PROTEIN 1

Chain D:





• Molecule 2: WRPW PEPTIDE

Chain X:



• Molecule 2: WRPW PEPTIDE

Chain Y:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	107.81Å 56.48Å 126.64Å 90.00° 112.68° 90.00°	Depositor
Resolution (Å)	117.04 – 2.12 42.48 – 2.12	Depositor EDS
% Data completeness (in resolution range)	97.7 (117.04-2.12) 97.7 (42.48-2.12)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.39 (at 2.12Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.178 , 0.244 0.178 , 0.240	Depositor DCC
$R_{free}$ test set	3924 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	25.8	Xtriage
Anisotropy	0.168	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 36.7	EDS
Estimated twinning fraction	0.000 for h,-k,-h-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	3 of 77993 reflections (0.004%)	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	11285	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 83.94 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 1.9016e-07. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

## 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	0.98	3/2653 (0.1%)	0.98	10/3612 (0.3%)
1	B	0.94	1/2653 (0.0%)	1.00	8/3612 (0.2%)
1	C	0.90	0/2653	0.89	6/3612 (0.2%)
1	D	0.87	1/2619 (0.0%)	0.95	11/3564 (0.3%)
2	X	0.87	0/59	0.68	0/79
2	Y	0.82	0/58	0.83	0/79
All	All	0.92	5/10695 (0.0%)	0.95	35/14558 (0.2%)

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	2
1	B	0	2
1	C	0	1
1	D	0	4
All	All	0	9

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	506	CYS	CB-SG	-6.45	1.71	1.82
1	A	762	LYS	CD-CE	6.00	1.66	1.51
1	A	762	LYS	CE-NZ	5.74	1.63	1.49
1	A	589	CYS	CB-SG	-5.20	1.73	1.81
1	D	485	GLU	CG-CD	5.03	1.59	1.51

The worst 5 of 35 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	514	HIS	C-N-CD	-18.97	78.87	120.60

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	514	HIS	C-N-CA	13.80	179.96	122.00
1	A	704	LEU	CA-CB-CG	11.45	141.63	115.30
1	D	514	HIS	C-N-CD	-11.14	96.08	120.60
1	A	473	ARG	NE-CZ-NH2	-9.94	115.33	120.30

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	453	ASP	Peptide
1	A	454	GLY	Peptide
1	B	514	HIS	Peptide
1	B	561	ALA	Peptide
1	C	453	ASP	Peptide

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2587	0	2503	62	0
1	B	2587	0	2503	72	0
1	C	2587	0	2503	52	0
1	D	2554	0	2475	92	0
2	X	55	0	48	4	0
2	Y	54	0	48	1	0
3	A	246	0	0	29	0
3	B	228	0	0	16	0
3	C	206	0	0	12	0
3	D	176	0	0	13	0
3	X	3	0	0	0	0
3	Y	2	0	0	0	0
All	All	11285	0	10080	279	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 14.



The worst 5 of 279 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:517:ASN:HA	1:D:518:LYS:CB	1.46	1.38
1:D:517:ASN:CA	1:D:518:LYS:HB2	1.63	1.27
1:C:563:THR:HB	1:C:564:PRO:CD	1.64	1.24
1:C:454:GLY:HA2	1:C:455:GLN:CB	1.66	1.23
1:A:762:LYS:HD2	3:A:2219:HOH:O	1.36	1.22

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	335/337 (99%)	325 (97%)	8 (2%)	2 (1%)	33	27
1	B	335/337 (99%)	313 (93%)	15 (4%)	7 (2%)	11	4
1	C	335/337 (99%)	319 (95%)	10 (3%)	6 (2%)	13	5
1	D	328/337 (97%)	302 (92%)	11 (3%)	15 (5%)	4	1
2	X	3/5 (60%)	3 (100%)	0	0	100	100
2	Y	3/5 (60%)	3 (100%)	0	0	100	100
All	All	1339/1358 (99%)	1265 (94%)	44 (3%)	30 (2%)	10	4

5 of 30 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	514	HIS
1	B	515	PRO
1	B	562	PRO
1	C	563	THR
1	D	514	HIS

### 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	288/288 (100%)	278 (96%)	10 (4%)	48	48
1	B	288/288 (100%)	274 (95%)	14 (5%)	35	31
1	C	288/288 (100%)	276 (96%)	12 (4%)	40	38
1	D	285/288 (99%)	269 (94%)	16 (6%)	30	25
2	X	5/5 (100%)	4 (80%)	1 (20%)	2	1
2	Y	5/5 (100%)	5 (100%)	0	100	100
All	All	1159/1162 (100%)	1106 (95%)	53 (5%)	37	34

5 of 53 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	743	LEU
1	C	524	LEU
1	D	657	THR
1	C	447	SER
1	C	453	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	680	ASN
1	C	613	GLN
1	D	687	ASN
1	B	695	HIS
1	C	531	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

### 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	337/337 (100%)	-0.27	3 (0%) 81 85	13, 20, 40, 57	0
1	B	337/337 (100%)	-0.11	5 (1%) 70 74	14, 23, 44, 57	0
1	C	337/337 (100%)	-0.23	4 (1%) 75 79	15, 23, 44, 57	0
1	D	332/337 (98%)	0.01	12 (3%) 41 44	15, 25, 45, 59	0
2	X	5/5 (100%)	0.02	0 100 100	28, 29, 34, 37	0
2	Y	5/5 (100%)	0.00	0 100 100	31, 31, 37, 40	0
All	All	1353/1358 (99%)	-0.15	24 (1%) 65 69	13, 23, 44, 59	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	516	GLY	5.8
1	B	517	ASN	4.7
1	D	517	ASN	4.5
1	C	562	PRO	2.9
1	B	561	ALA	2.8

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

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

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

There are no ligands in this entry.

## 6.5 Other polymers ⓘ

There are no such residues in this entry.