



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 05:11 AM GMT

PDB ID : 3DPN  
Title : Crystal Structure of cpaf s499a mutant  
Authors : Chai, J.; Huang, Z.  
Deposited on : 2008-07-09  
Resolution : 3.30 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>

---

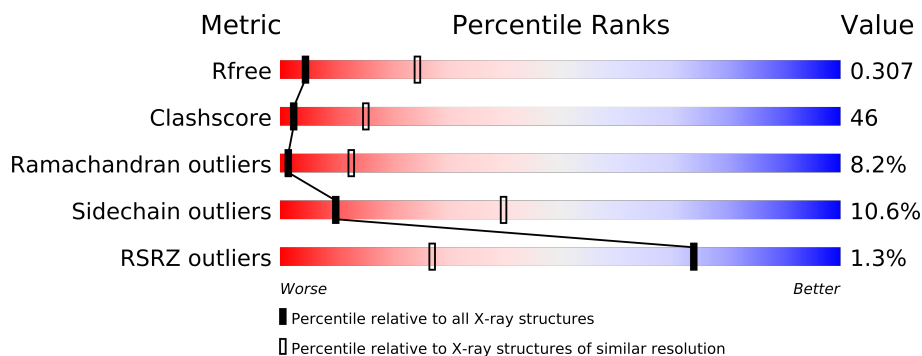
The following versions of software and data (see [references](#)) were used in the production of this report:

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 3.30 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	1341 (3.40-3.20)
Clashscore	79885	1696 (3.40-3.20)
Ramachandran outliers	78287	1664 (3.40-3.20)
Sidechain outliers	78261	1662 (3.40-3.20)
RSRZ outliers	66119	1342 (3.40-3.20)

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	583	
1	B	583	

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 8476 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 Protein CT\_858.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	537	Total	C	N	O	S	0	0	0
			4238	2719	709	796	14			
1	B	537	Total	C	N	O	S	0	0	0
			4238	2719	709	796	14			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	499	ALA	SER	ENGINEERED	UNP O84866
A	610	HIS	-	EXPRESSION TAG	UNP O84866
A	611	HIS	-	EXPRESSION TAG	UNP O84866
A	612	HIS	-	EXPRESSION TAG	UNP O84866
A	613	HIS	-	EXPRESSION TAG	UNP O84866
A	614	HIS	-	EXPRESSION TAG	UNP O84866
A	615	HIS	-	EXPRESSION TAG	UNP O84866
B	499	ALA	SER	ENGINEERED	UNP O84866
B	610	HIS	-	EXPRESSION TAG	UNP O84866
B	611	HIS	-	EXPRESSION TAG	UNP O84866
B	612	HIS	-	EXPRESSION TAG	UNP O84866
B	613	HIS	-	EXPRESSION TAG	UNP O84866
B	614	HIS	-	EXPRESSION TAG	UNP O84866
B	615	HIS	-	EXPRESSION TAG	UNP O84866



T600	V631	T467	Q404	V325	SER	E174	G107
L601	Q532	P468	D405	G326	PHE	E175	V108
L602	P533	T469	E406	F327	PHE	S176	T109
L603	P534	P470	V407	L328	PRO	A177	F110
A604	N635	L471	D408	R329	LYS	A178	F111
R636	R636	F472	V409	T332	LYS	L179	A112
T637	T637	G473	A410	Y333	ASP	R180	I113
G638	G638	F474	L411	Y334	ASP	T181	E114
I639	I639	E475	L414	S334	ALA	L182	S115
K640	K640	R476	T415	W335	PHE	F183	A116
T641	T641	I477	L416	Q336	HIS	S184	Y117
C642	C642	H480	L417	F341	ARG	R185	L118
S643	S643	F481	V420	P348	SER	M186	P119
L644	L644	R482	D421	W349	SER	A187	Y120
T645	T645	V483	T422	F260	S258	S188	T121
G646	G646	Y484	N423	E350	L259	L189	I122
S647	S647	Y485	E424	E351	F260	G190	Q123
L648	L648	S486	V424	F352	W264	H191	K124
A649	A649	K487	E425	A353	V265	K192	S125
V650	V650	P488	S426	I356	P266	V193	S126
R651	R651	T489	R427	F359	H267	P194	G127
E652	E652	C490	L428	T363	F268	L200	G128
H653	H653	I493	A429	L366	W269	K201	R129
G654	G654	M494	L430	I367	A270	I202	F130
A655	A655	E495	G431	L368	E271	R203	D134
F656	F656	Q496	D432	I369	L272	R204	I135
I657	I657	D497	M434	I368	Y275	G207	M136
V662	V662	F498	E435	I369	H276	T208	T137
E663	E663	A499	G436	I369	Y276	F138	F138
P664	P664	C500	Y437	Q370	S279	T269	S139
H665	H665	A501	T438	N373	G280	R210	S140
L668	L668	D502	V439	P374	L281	W216	E141
P669	P669	F503	D440	G375	K282	R217	I142
F670	F670	F504	L441	G376	S283	R218	R143
T671	T671	P505	Q442	G376	G284	Y218	V144
A672	A672	V506	V443	S377	Y285	V219	G145
N673	N673	V507	A444	V378	N286	G222	L148
D674	D674	L508	E445	L379	T290	VAL	L148
I675	I675	K509	Y446	Y380	D291	GLY	V151
K578	K578	D510	L447	L381	T291	ASP	D152
G579	G579	N511	K448	Y382	D292	LEU	G153
Y580	Y580	D512	S449	A383	F293	ALA	G154
S681	S681	R513	F450	L384	F296	THR	P155
E682	E682	A514	G451	L385	I297	ILE	V156
Y683	Y683	L515	R452	T389	Q298	ALA	Q157
K686	K686	I516	Q453	P392	D299	PRO	D158
V691	V691	V517	V454	L393	P299	SER	V159
C692	C692	G518	L455	L394	V300	ILE	L160
Q693	Q693	A522	M456	E394	I301	ARG	A161
L694	L694	G523	C457	L395	W302	ALA	T162
I695	I695	A524	W458	H398	E303	PRO	L163
G526	G526	G525	S459	R399	S304	GLN	Y164
N696	N696	G527	K460	M400	A310	LEU	H168
N697	N697	F527	T463	I401	Y311	LYS	K169
D698	D698	V528	E464	L402	I312	SER	G170
G599	G599	N530	L465	T403	S313	MET	T171

## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	192.67Å 192.67Å 338.87Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 3.30 49.90 – 3.22	Depositor EDS
% Data completeness (in resolution range)	99.6 (20.00-3.30) 98.7 (49.90-3.22)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.12 (at 3.19Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.257 , 0.305 0.262 , 0.307	Depositor DCC
$R_{free}$ test set	1800 reflections (4.96%)	DCC
Wilson B-factor (Å <sup>2</sup> )	100.6	Xtriage
Anisotropy	0.607	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 107.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	3 of 39050 reflections (0.008%)	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	8476	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	106.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.00% of the height of the origin peak. No significant pseudotranslation is detected.*

<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	0.47	0/4344	0.76	4/5905 (0.1%)
1	B	0.48	0/4344	0.79	6/5905 (0.1%)
All	All	0.47	0/8688	0.78	10/11810 (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	A	0	1

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	138	PHE	N-CA-C	11.85	143.00	111.00
1	B	556	PHE	N-CA-C	7.79	132.03	111.00
1	A	554	GLY	N-CA-C	6.73	129.92	113.10
1	B	281	LEU	N-CA-C	-6.38	93.78	111.00
1	B	137	THR	C-N-CA	-5.65	107.58	121.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	138	PHE	Sidechain

## 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	4238	0	4157	386	0
1	B	4238	0	4157	384	0
All	All	8476	0	8314	768	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 46.

The worst 5 of 768 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:550:VAL:HG12	1:A:554:GLY:HA2	1.29	1.12
1:B:54:TYR:CZ	1:B:56:PRO:HG2	1.90	1.05
1:A:557:ILE:HG22	1:A:558:GLU:N	1.71	1.04
1:B:47:GLU:HG3	1:B:68:LEU:HD11	1.39	1.04
1:A:398:HIS:HD2	1:A:471:LEU:HD21	1.22	1.01

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	533/583 (91%)	397 (74%)	90 (17%)	46 (9%)	1	11
1	B	533/583 (91%)	404 (76%)	88 (16%)	41 (8%)	1	14
All	All	1066/1166 (91%)	801 (75%)	178 (17%)	87 (8%)	1	13

5 of 87 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	60	LYS
1	A	136	MET
1	A	139	SER
1	A	261	TYR
1	A	268	PHE

### 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	463/505 (92%)	418 (90%)	45 (10%)	12	45
1	B	463/505 (92%)	410 (89%)	53 (11%)	8	36
All	All	926/1010 (92%)	828 (89%)	98 (11%)	10	40

5 of 98 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	556	PHE
1	B	113	ILE
1	B	550	VAL
1	B	41	GLN
1	B	70	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	565	HIS
1	B	102	ASN
1	B	530	ASN
1	A	596	ASN
1	B	84	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	537/583 (92%)	0.16	9 (1%)	67 21	54, 107, 187, 198	0
1	B	537/583 (92%)	0.07	5 (0%)	81 37	43, 90, 169, 198	0
All	All	1074/1166 (92%)	0.12	14 (1%)	74 27	43, 98, 181, 198	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	460	LYS	3.8
1	A	604	ALA	3.3
1	B	554	GLY	2.8
1	B	536	ARG	2.6
1	A	600	THR	2.5

### 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.