



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 07:12 AM GMT

PDB ID : 2ZWO  
Title : Crystal structure of Ca2 site mutant of Pro-S324A  
Authors : Takeuchi, Y.; Tanaka, S.; Matsumura, H.; Koga, Y.; Takano, K.; Kanaya, S.  
Deposited on : 2008-12-17  
Resolution : 2.07 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

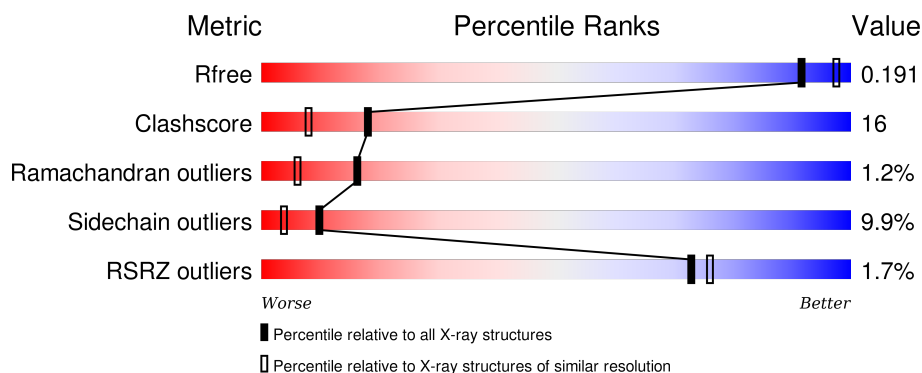
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.07 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	1799 (2.08-2.04)
Clashscore	102246	1910 (2.08-2.04)
Ramachandran outliers	100387	1893 (2.08-2.04)
Sidechain outliers	100360	1893 (2.08-2.04)
RSRZ outliers	91569	1802 (2.08-2.04)

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	398	<div> <div>2%</div> <div>73%</div> <div>22%</div> <div>• •</div> </div>
1	B	398	<div> <div>3%</div> <div>66%</div> <div>26%</div> <div>6%</div> <div>• •</div> </div>
1	C	398	<div> <div>%</div> <div>62%</div> <div>29%</div> <div>6%</div> <div>• •</div> </div>

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
2	CA	B	403	-	-	-	X

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 9389 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 Tk-subtilisin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	395	Total	C	N	O	S	0	0	0
			2890	1830	484	569	7			
1	B	393	Total	C	N	O	S	0	0	0
			2888	1828	483	570	7			
1	C	389	Total	C	N	O	S	0	0	0
			2857	1811	478	561	7			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	226	ALA	ASP	ENGINEERED	UNP P58502
A	324	ALA	SER	ENGINEERED	UNP P58502
B	226	ALA	ASP	ENGINEERED	UNP P58502
B	324	ALA	SER	ENGINEERED	UNP P58502
C	226	ALA	ASP	ENGINEERED	UNP P58502
C	324	ALA	SER	ENGINEERED	UNP P58502

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	5	Total	Ca	0	0
			5	5		
2	A	6	Total	Ca	0	0
			6	6		
2	C	5	Total	Ca	0	0
			5	5		

- Molecule 3 is water.

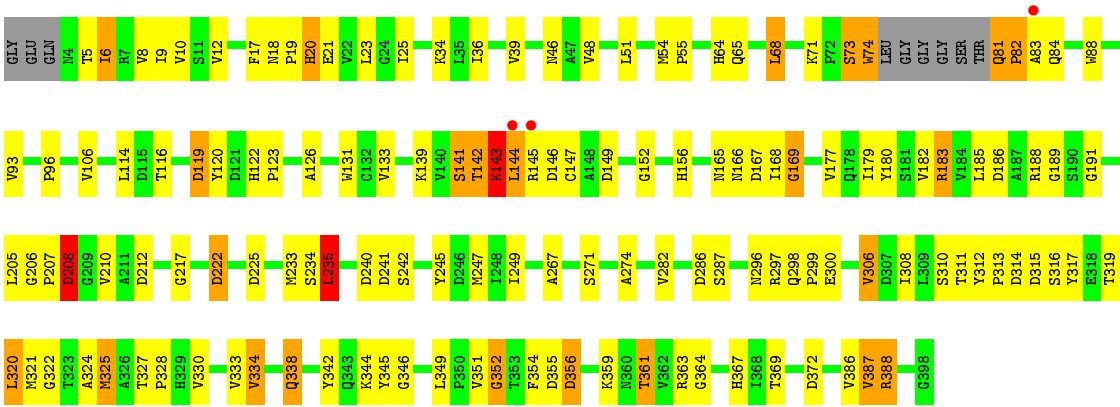
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	267	Total	O	0	0
			267	267		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	251	Total	O	0	0
			251	251		
3	C	220	Total	O	0	0
			220	220		





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	93.73Å 118.31Å 120.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.63 – 2.07 46.86 – 2.07	Depositor EDS
% Data completeness (in resolution range)	100.0 (46.63-2.07) 96.3 (46.86-2.07)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.08 (at 2.07Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.184 , 0.227 0.185 , 0.191	Depositor DCC
$R_{free}$ test set	3973 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	34.5	Xtriage
Anisotropy	0.084	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 57.6	EDS
Estimated twinning fraction	0.015 for -h,l,k	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	2 of 79345 reflections (0.003%)	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	9389	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.11% 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

Bond lengths and bond angles in the following residue types are not validated in this section: CA

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	1.02	1/2951 (0.0%)	1.01	8/4038 (0.2%)
1	B	0.95	1/2948 (0.0%)	1.01	16/4033 (0.4%)
1	C	1.00	1/2917 (0.0%)	1.05	13/3991 (0.3%)
All	All	0.99	3/8816 (0.0%)	1.03	37/12062 (0.3%)

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	14
1	B	0	7
1	C	1	10
All	All	1	31

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	282	VAL	CB-CG2	6.92	1.67	1.52
1	A	230	VAL	CB-CG2	5.47	1.64	1.52
1	B	210	VAL	C-O	5.35	1.33	1.23

The worst 5 of 37 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	183	ARG	NE-CZ-NH1	10.44	125.52	120.30
1	A	183	ARG	NE-CZ-NH2	-10.38	115.11	120.30
1	C	183	ARG	NE-CZ-NH1	9.26	124.93	120.30
1	C	351	VAL	C-N-CA	-8.06	105.38	122.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	183	ARG	NE-CZ-NH2	-8.03	116.29	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	C	6	ILE	CB

5 of 31 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	12	VAL	Peptide
1	A	53	LYS	Peptide
1	A	55	PRO	Peptide
1	A	65	GLN	Mainchain
1	A	75	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	2890	0	2851	67	0
1	B	2888	0	2844	106	0
1	C	2857	0	2818	103	0
2	A	6	0	0	0	0
2	B	5	0	0	0	0
2	C	5	0	0	0	0
3	A	267	0	0	7	1
3	B	251	0	0	12	1
3	C	220	0	0	3	2
All	All	9389	0	8513	269	2

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 16.

The worst 5 of 269 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:338:GLN:HA	1:B:338:GLN:HE21	1.05	1.10
1:C:361:THR:HG22	1:C:364:GLY:H	1.21	1.03
1:B:352:GLY:HA3	1:B:363:ARG:HB2	1.41	0.98
1:C:321:MET:C	1:C:325:MET:CE	2.33	0.96
1:B:321:MET:C	1:B:325:MET:HE2	1.89	0.93

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:518:HOH:O	3:C:649:HOH:O[2_565]	1.53	0.67
3:B:634:HOH:O	3:C:555:HOH:O[1_556]	1.84	0.36

## 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	393/398 (99%)	364 (93%)	26 (7%)	3 (1%)	24	12
1	B	389/398 (98%)	356 (92%)	27 (7%)	6 (2%)	13	4
1	C	385/398 (97%)	354 (92%)	26 (7%)	5 (1%)	15	5
All	All	1167/1194 (98%)	1074 (92%)	79 (7%)	14 (1%)	16	6

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	76	GLY
1	B	51	LEU
1	B	206	GLY
1	C	119	ASP
1	C	169	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	304/306 (99%)	278 (91%)	26 (9%)	13	6
1	B	305/306 (100%)	276 (90%)	29 (10%)	11	4
1	C	301/306 (98%)	266 (88%)	35 (12%)	7	2
All	All	910/918 (99%)	820 (90%)	90 (10%)	10	4

5 of 90 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	133	VAL
1	B	334	VAL
1	C	334	VAL
1	B	147	CYS
1	B	235	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	65	GLN
1	B	296	ASN
1	C	338	GLN
1	B	122	HIS
1	B	150	GLN

### 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](#)

Of 16 ligands modelled in this entry, 16 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	395/398 (99%)	-0.31	7 (1%) 71 74	18, 32, 53, 67	0
1	B	393/398 (98%)	-0.17	10 (2%) 61 65	24, 35, 52, 81	0
1	C	389/398 (97%)	-0.47	3 (0%) 87 89	19, 35, 57, 80	0
All	All	1177/1194 (98%)	-0.32	20 (1%) 73 76	18, 34, 54, 81	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	80	THR	4.3
1	B	82	PRO	3.5
1	A	345	TYR	3.4
1	B	81	GLN	3.4
1	B	83	ALA	3.2

### 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	CA	B	403	1/1	0.98	0.10	3.02	35,35,35,35	0
2	CA	C	401	1/1	0.99	0.11	1.50	26,26,26,26	0
2	CA	B	400	1/1	0.98	0.10	0.50	34,34,34,34	0
2	CA	A	399	1/1	0.99	0.12	0.40	24,24,24,24	0
2	CA	A	404	1/1	0.99	0.07	-0.61	35,35,35,35	0
2	CA	C	400	1/1	0.99	0.07	-1.18	39,39,39,39	0
2	CA	C	403	1/1	0.97	0.07	-1.24	36,36,36,36	0
2	CA	C	399	1/1	0.98	0.05	-1.48	44,44,44,44	0
2	CA	B	401	1/1	0.98	0.07	-1.55	28,28,28,28	0
2	CA	A	402	1/1	0.99	0.04	-1.98	36,36,36,36	0
2	CA	A	400	1/1	0.99	0.06	-2.58	52,52,52,52	0
2	CA	B	402	1/1	0.99	0.06	-2.78	47,47,47,47	0
2	CA	C	402	1/1	0.99	0.05	-2.79	39,39,39,39	0
2	CA	B	399	1/1	0.97	0.03	-3.60	44,44,44,44	0
2	CA	A	401	1/1	0.98	0.05	-	60,60,60,60	0
2	CA	A	403	1/1	0.99	0.07	-	37,37,37,37	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.