



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 16, 2019 – 03:04 AM EDT

PDB ID : 1E3C  
Title : Crystal structure of an Arylsulfatase A mutant C69S soaked in synthetic substrate  
Authors : von Buelow, R.; Schmidt, B.; Dierks, T.; von Figura, K.; Uson, I.  
Deposited on : 2000-06-13  
Resolution : 2.65 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  
Mogul : 1.8.0 (224370), CSD as540be (2019)  
Xtriage (Phenix) : 1.13  
EDS : 2.3.2  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.3.2

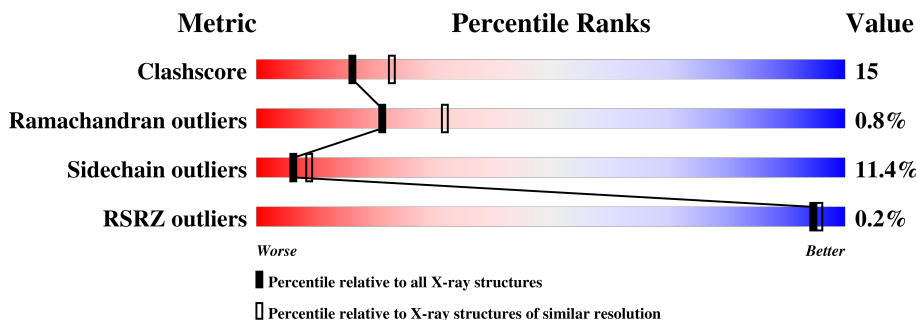
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*


The reported resolution of this entry is 2.65 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	122126	1151 (2.68-2.64)
Ramachandran outliers	120053	1133 (2.68-2.64)
Sidechain outliers	120020	1133 (2.68-2.64)
RSRZ outliers	108989	1098 (2.68-2.64)

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	P	489	

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	NAG	P	601	X	-	-	-
2	NAG	P	602	X	-	-	-

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 3784 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.

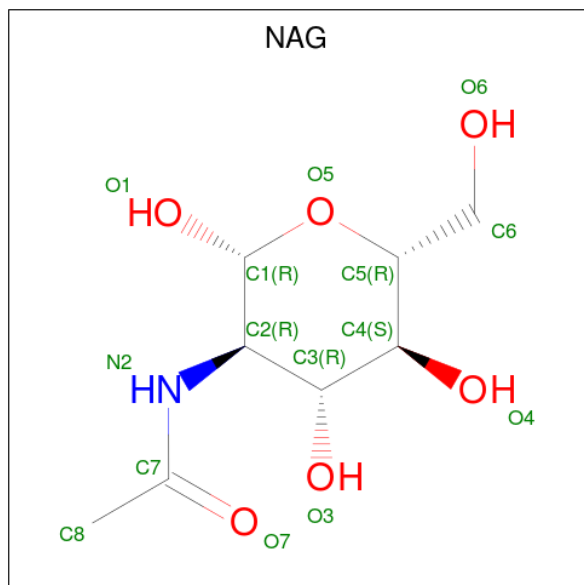
- Molecule 1 is a protein called Arylsulfatase A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	P	481	Total	C	N	O	S	0	0	0
			3582	2282	610	667	23			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
P	69	SER	CYS	engineered mutation	UNP P15289
P	215	GLU	GLN	conflict	UNP P15289

- Molecule 2 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	P	1	Total	C	N	O	0	0
			14	8	1	5		
2	P	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	P	1	Total 1	Mg 1	0	0

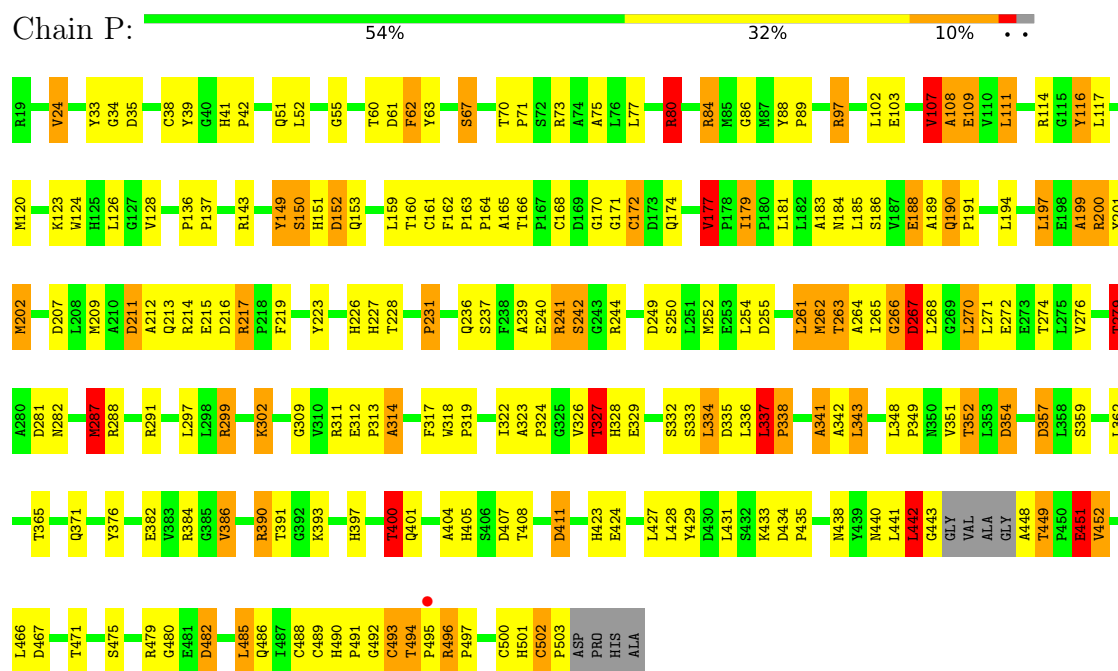
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	P	173	Total 173	O 173	0	0

### 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 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 ( $RSRZ > 2$ ). 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: Arylsulfatase A



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 4 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	131.80Å 131.80Å 192.10Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 2.65 29.47 – 2.65	Depositor EDS
% Data completeness (in resolution range)	99.8 (40.00-2.65) 99.4 (29.47-2.65)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.15 (at 2.64Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.173 , 0.243 0.150 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	41.1	Xtriage
Anisotropy	0.597	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 68.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.015 for -1/2*h-1/2*k-1/2*l,-1/2*h-1/2*k+1/2*l,-h+k 0.005 for -1/2*h+1/2*k-1/2*l,1/2*h-1/2*k-1/2*l,-h-k	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	3784	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.40% 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

### 5.1 Standard geometry

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

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	P	0.97	1/3689 (0.0%)	2.23	165/5042 (3.3%)

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	P	0	6

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	P	503	PRO	N-CD	7.20	1.57	1.47

The worst 5 of 165 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	P	291	ARG	NE-CZ-NH2	-23.56	108.52	120.30
1	P	354	ASP	CB-CG-OD1	17.63	134.17	118.30
1	P	281	ASP	CB-CG-OD1	14.53	131.38	118.30
1	P	217	ARG	NE-CZ-NH2	13.96	127.28	120.30
1	P	384	ARG	NE-CZ-NH2	-13.93	113.33	120.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	P	181	LEU	Mainchain

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
1	P	194	LEU	Mainchain
1	P	241	ARG	Mainchain
1	P	287	MET	Mainchain
1	P	42	PRO	Mainchain

## 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	P	3582	0	3453	109	1
2	P	28	0	25	0	0
3	P	1	0	0	0	0
4	P	173	0	0	18	1
All	All	3784	0	3478	109	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 15.

The worst 5 of 109 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:P:102:LEU:O	4:P:2030:HOH:O	1.65	1.12
1:P:109:GLU:OE1	4:P:2034:HOH:O	1.90	0.88
1:P:423:HIS:HD2	4:P:2150:HOH:O	1.61	0.84
1:P:449:THR:HG22	1:P:451:GLU:H	1.53	0.74
1:P:24:VAL:HG13	1:P:276:VAL:HG22	1.71	0.73

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)
4:P:2148:HOH:O	4:P:2148:HOH:O[5_555]	1.84	0.36
1:P:490:HIS:NE2	1:P:490:HIS:NE2[16_554]	2.17	0.03



## 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	P	477/489 (98%)	438 (92%)	35 (7%)	4 (1%)	21	32

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	P	493	CYS
1	P	267	ASP
1	P	172	CYS
1	P	266	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	P	377/385 (98%)	334 (89%)	43 (11%)	6	9

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	P	262	MET
1	P	324	PRO
1	P	466	LEU
1	P	272	GLU
1	P	279	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	P	328	HIS
1	P	465	GLN
1	P	423	HIS
1	P	226	HIS
1	P	440	ASN

### 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 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	NAG	P	601	1,2	14,14,15	1.97	4 (28%)	17,19,21	3.51	11 (64%)
2	NAG	P	602	2	14,14,15	1.28	2 (14%)	17,19,21	2.49	6 (35%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	P	601	1,2	1/1/5/7	2/6/23/26	0/1/1/1
2	NAG	P	602	2	1/1/5/7	2/6/23/26	0/1/1/1

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	P	601	NAG	O7-C7	-4.16	1.13	1.23
2	P	601	NAG	C2-N2	4.03	1.53	1.46
2	P	602	NAG	O7-C7	-2.91	1.16	1.23
2	P	601	NAG	O4-C4	-2.82	1.36	1.43
2	P	601	NAG	O5-C5	-2.58	1.38	1.43

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	P	601	NAG	C8-C7-N2	6.85	127.92	116.10
2	P	601	NAG	O7-C7-N2	-6.38	110.06	121.95
2	P	601	NAG	O4-C4-C3	5.34	122.75	110.34
2	P	602	NAG	C1-O5-C5	5.14	119.19	112.20
2	P	602	NAG	O3-C3-C2	-5.00	98.84	109.38

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	P	601	NAG	C1
2	P	602	NAG	C1

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	P	601	NAG	C8-C7-N2-C2
2	P	601	NAG	O7-C7-N2-C2
2	P	602	NAG	O5-C5-C6-O6
2	P	602	NAG	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

## 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	P	481/489 (98%)	-0.79	1 (0%) 94 95	15, 34, 61, 95	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	P	495	PRO	2.5

### 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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
2	NAG	P	602	14/15	0.90	0.25	70,84,97,98	0
2	NAG	P	601	14/15	0.96	0.17	43,54,65,70	0
3	MG	P	603	1/1	0.98	0.27	48,48,48,48	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.