



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 15, 2017 – 10:28 AM EDT

PDB ID : 2QP3  
Title : Identification and Characterization of Two Amino Acids Critical for the Substrate Inhibition of SULT2A1  
Authors : Hsieh, Y.C.  
Deposited on : unknown  
Resolution : 2.60 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030345

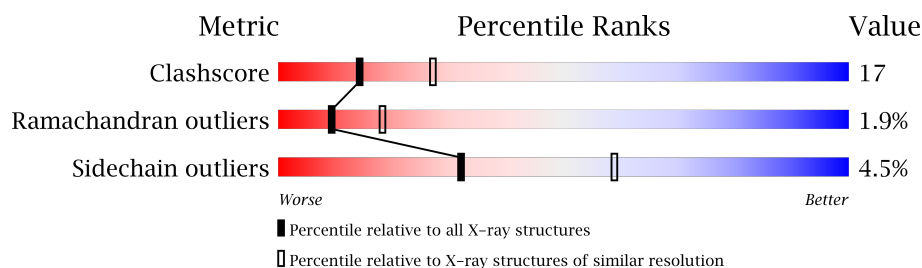
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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	112137	2895 (2.60-2.60)
Ramachandran outliers	110173	2848 (2.60-2.60)
Sidechain outliers	110143	2848 (2.60-2.60)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	284	

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 2269 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.

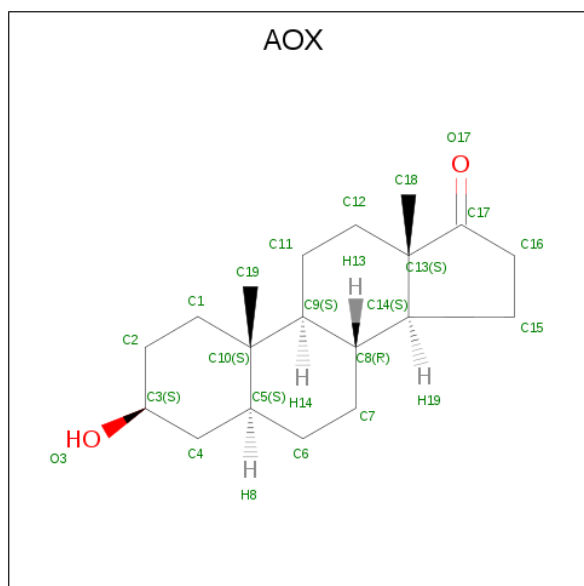
- Molecule 1 is a protein called Bile salt sulfotransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	267	2248	1472	360	406	10	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	137	ILE	MET	ENGINEERED	UNP Q06520
A	243	THR	ALA	CONFLICT	UNP Q06520

- Molecule 2 is (3Beta,5alpha)-3-Hydroxyandrostane-17-one (three-letter code: AOX) (formula:  $C_{19}H_{30}O_2$ ).



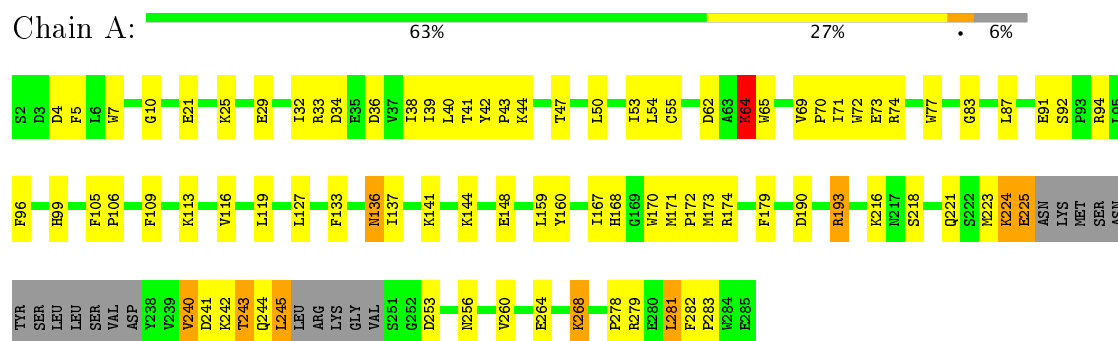
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			21	19	2		

### 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 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.

Note EDS was not executed.

- Molecule 1: Bile salt sulfotransferase



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	76.28 Å   131.72 Å   44.55 Å 90.00°   90.00°   90.00°	Depositor
Resolution (Å)	20.00 – 2.60	Depositor
% Data completeness (in resolution range)	98.3 (20.00-2.60)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS	Depositor
R, $R_{free}$	0.226 , 0.264	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2269	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	46.0	wwPDB-VP

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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

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	0.39	0/2317	0.57	1/3129 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	42	TYR	N-CA-C	-5.07	97.31	111.00

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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	2248	0	2186	78	0
2	A	21	0	30	0	0
All	All	2269	0	2216	78	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 17.

All (78) 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:A:70:PRO:HG2	1:A:73:GLU:HG3	1.46	0.95
1:A:242:LYS:HD2	1:A:242:LYS:H	1.44	0.82
1:A:53:ILE:HD13	1:A:218:SER:HB2	1.60	0.81
1:A:136:ASN:HD22	1:A:240:VAL:HB	1.50	0.76
1:A:65:TRP:HB3	1:A:74:ARG:NH2	2.02	0.73
1:A:39:ILE:HD11	1:A:54:LEU:CD1	2.20	0.72
1:A:71:ILE:HG13	1:A:72:TRP:H	1.56	0.70
1:A:221:GLN:HA	1:A:224:LYS:NZ	2.08	0.68
1:A:190:ASP:OD2	1:A:193:ARG:HB2	1.93	0.67
1:A:242:LYS:CD	1:A:242:LYS:H	2.07	0.67
1:A:136:ASN:ND2	1:A:240:VAL:HB	2.11	0.66
1:A:242:LYS:N	1:A:242:LYS:HD2	2.11	0.65
1:A:221:GLN:HA	1:A:224:LYS:HZ2	1.61	0.65
1:A:39:ILE:HD11	1:A:54:LEU:HD11	1.79	0.64
1:A:21:GLU:O	1:A:25:LYS:HG2	1.98	0.63
1:A:33:ARG:HH11	1:A:92:SER:HB2	1.62	0.62
1:A:70:PRO:CG	1:A:73:GLU:HG3	2.25	0.62
1:A:33:ARG:NH1	1:A:92:SER:H	1.97	0.62
1:A:216:LYS:NZ	1:A:216:LYS:HB2	2.18	0.59
1:A:268:LYS:HB2	1:A:268:LYS:NZ	2.17	0.59
1:A:221:GLN:CD	1:A:221:GLN:H	2.09	0.56
1:A:39:ILE:HD11	1:A:54:LEU:HD12	1.86	0.56
1:A:33:ARG:NH1	1:A:92:SER:N	2.53	0.56
1:A:260:VAL:O	1:A:264:GLU:HG3	2.06	0.55
1:A:71:ILE:HG13	1:A:72:TRP:N	2.21	0.55
1:A:40:LEU:HD12	1:A:116:VAL:CG1	2.37	0.54
1:A:136:ASN:HD21	1:A:241:ASP:H	1.53	0.54
1:A:136:ASN:HD21	1:A:241:ASP:N	2.05	0.53
1:A:41:THR:HG22	1:A:119:LEU:HB3	1.90	0.53
1:A:62:ASP:OD2	1:A:64:LYS:HB2	2.09	0.53
1:A:221:GLN:NE2	1:A:221:GLN:H	2.06	0.53
1:A:167:ILE:O	1:A:171:MET:HG2	2.08	0.53
1:A:144:LYS:HG2	1:A:148:GLU:OE1	2.10	0.52
1:A:223:MET:O	1:A:225:GLU:N	2.43	0.51
1:A:44:LYS:HE3	1:A:133:PHE:CG	2.45	0.51
1:A:32:ILE:HD12	1:A:109:PHE:HB2	1.93	0.51
1:A:7:TRP:CZ2	1:A:10:GLY:HA2	2.46	0.51
1:A:99:HIS:HB2	1:A:160:TYR:CE2	2.46	0.51
1:A:278:PRO:O	1:A:281:LEU:HB2	2.12	0.50
1:A:40:LEU:HD12	1:A:116:VAL:HG11	1.93	0.50
1:A:40:LEU:CD1	1:A:116:VAL:HG11	2.42	0.50
1:A:53:ILE:CD1	1:A:218:SER:HB2	2.37	0.49

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:53:ILE:HD13	1:A:218:SER:CB	2.37	0.49
1:A:240:VAL:HG12	1:A:241:ASP:N	2.28	0.49
1:A:4:ASP:HB2	1:A:141:LYS:HB2	1.95	0.49
1:A:69:VAL:HG21	1:A:74:ARG:HH21	1.78	0.49
1:A:43:PRO:HD3	1:A:99:HIS:HB3	1.94	0.49
1:A:77:TRP:O	1:A:83:GLY:HA3	2.13	0.49
1:A:170:TRP:HZ3	1:A:173:MET:HE1	1.77	0.48
1:A:25:LYS:HE2	1:A:25:LYS:N	2.29	0.48
1:A:105:PHE:CG	1:A:106:PRO:HD2	2.48	0.48
1:A:41:THR:O	1:A:99:HIS:HA	2.13	0.48
1:A:133:PHE:CE2	1:A:137:ILE:HD11	2.49	0.48
1:A:25:LYS:HE2	1:A:25:LYS:CA	2.43	0.48
1:A:38:ILE:HG12	1:A:96:PHE:CZ	2.48	0.47
1:A:268:LYS:HZ3	1:A:268:LYS:HB2	1.79	0.47
1:A:25:LYS:HA	1:A:25:LYS:HE2	1.97	0.47
1:A:216:LYS:HZ2	1:A:216:LYS:HB2	1.80	0.46
1:A:243:THR:O	1:A:245:LEU:N	2.47	0.46
1:A:159:LEU:O	1:A:160:TYR:HB2	2.14	0.46
1:A:253:ASP:H	1:A:256:ASN:HD22	1.64	0.46
1:A:168:HIS:O	1:A:172:PRO:HD3	2.16	0.46
1:A:36:ASP:OD1	1:A:94:ARG:HB2	2.16	0.45
1:A:39:ILE:HG23	1:A:50:LEU:HD23	1.99	0.44
1:A:242:LYS:CD	1:A:242:LYS:N	2.75	0.44
1:A:136:ASN:ND2	1:A:241:ASP:H	2.13	0.44
1:A:40:LEU:CD1	1:A:116:VAL:CG1	2.96	0.43
1:A:25:LYS:HD3	1:A:29:GLU:OE2	2.19	0.43
1:A:4:ASP:O	1:A:5:PHE:HB3	2.18	0.43
1:A:282:PHE:HA	1:A:283:PRO:HD2	1.85	0.43
1:A:25:LYS:HD3	1:A:29:GLU:CD	2.40	0.42
1:A:33:ARG:HH12	1:A:92:SER:N	2.17	0.42
1:A:34:ASP:OD2	1:A:113:LYS:HB2	2.20	0.41
1:A:190:ASP:OD2	1:A:193:ARG:CB	2.64	0.41
1:A:174:ARG:HA	1:A:179:PHE:CD1	2.56	0.41
1:A:127:LEU:HD23	1:A:127:LEU:C	2.41	0.41
1:A:43:PRO:O	1:A:44:LYS:HB2	2.21	0.41
1:A:55:CYS:SG	1:A:74:ARG:HD2	2.62	0.40

There are no symmetry-related clashes.



## 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	A	261/284 (92%)	242 (93%)	14 (5%)	5 (2%)	9	18

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	64	LYS
1	A	240	VAL
1	A	243	THR
1	A	244	GLN
1	A	224	LYS

### 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	247/263 (94%)	236 (96%)	11 (4%)	32	59

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	47	THR
1	A	64	LYS
1	A	87	LEU
1	A	91	GLU
1	A	136	ASN
1	A	193	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	225	GLU
1	A	245	LEU
1	A	268	LYS
1	A	279	ARG
1	A	281	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	48	ASN
1	A	136	ASN
1	A	189	GLN
1	A	212	ASN
1	A	221	GLN
1	A	256	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](#)

1 ligand is modelled in this entry.

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	AOX	A	286	-	24,24,24	4.37	17 (70%)	39,39,39	3.71	20 (51%)

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	AOX	A	286	-	-	0/0/58/58	0/4/4/4

All (17) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	286	AOX	C8-C14	2.07	1.57	1.53
2	A	286	AOX	C16-C17	2.36	1.55	1.51
2	A	286	AOX	C12-C13	2.51	1.59	1.54
2	A	286	AOX	C10-C9	2.81	1.61	1.56
2	A	286	AOX	C6-C7	3.02	1.60	1.52
2	A	286	AOX	C12-C11	3.93	1.61	1.53
2	A	286	AOX	C2-C3	4.16	1.62	1.51
2	A	286	AOX	C1-C2	4.21	1.62	1.53
2	A	286	AOX	C8-C9	4.29	1.62	1.53
2	A	286	AOX	C7-C8	4.66	1.62	1.53
2	A	286	AOX	C11-C9	4.86	1.62	1.53
2	A	286	AOX	C6-C5	5.22	1.66	1.53
2	A	286	AOX	C18-C13	5.38	1.64	1.54
2	A	286	AOX	C19-C10	6.55	1.66	1.54
2	A	286	AOX	C4-C5	6.58	1.64	1.53
2	A	286	AOX	C4-C3	6.99	1.64	1.51
2	A	286	AOX	C10-C5	10.64	1.73	1.55

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	286	AOX	C5-C4-C3	-12.30	94.82	112.87
2	A	286	AOX	C1-C2-C3	-11.64	95.57	110.42
2	A	286	AOX	C18-C13-C12	-6.07	103.64	111.12
2	A	286	AOX	C4-C3-C2	-4.67	104.75	110.55
2	A	286	AOX	C19-C10-C5	-4.32	102.86	110.30
2	A	286	AOX	C16-C15-C14	-3.98	96.60	103.05
2	A	286	AOX	C11-C9-C10	-3.74	109.36	113.91
2	A	286	AOX	C4-C5-C10	-3.11	109.26	112.66

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	286	AOX	C14-C13-C17	-2.94	97.07	100.56
2	A	286	AOX	C6-C5-C10	-2.91	106.73	112.32
2	A	286	AOX	C2-C1-C10	-2.41	108.61	112.80
2	A	286	AOX	C15-C14-C8	-2.04	115.82	119.07
2	A	286	AOX	O17-C17-C16	-2.03	123.27	125.76
2	A	286	AOX	C1-C10-C5	2.02	110.93	107.79
2	A	286	AOX	O17-C17-C13	2.15	128.92	125.96
2	A	286	AOX	C4-C5-C6	2.34	116.24	111.73
2	A	286	AOX	C12-C13-C17	2.62	120.69	116.61
2	A	286	AOX	C13-C14-C8	2.95	116.31	113.15
2	A	286	AOX	C18-C13-C14	3.26	118.01	112.94
2	A	286	AOX	O3-C3-C2	6.58	126.05	110.10

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

### 6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.