



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2018 – 11:42 pm GMT

PDB ID : 5N2S  
Title : Crystal structure of stabilized A1 receptor in complex with PSB36 at 3.3Å resolution  
Authors : Cheng, R.K.Y.; Segala, E.; Robertson, N.; Deflorian, F.; Dore, A.S.; Errey, J.C.; Fiez-Vandal, C.; Marshall, F.H.; Cooke, R.M.  
Deposited on : 2017-02-08  
Resolution : 3.30 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 : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk30967  
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) : trunk30967

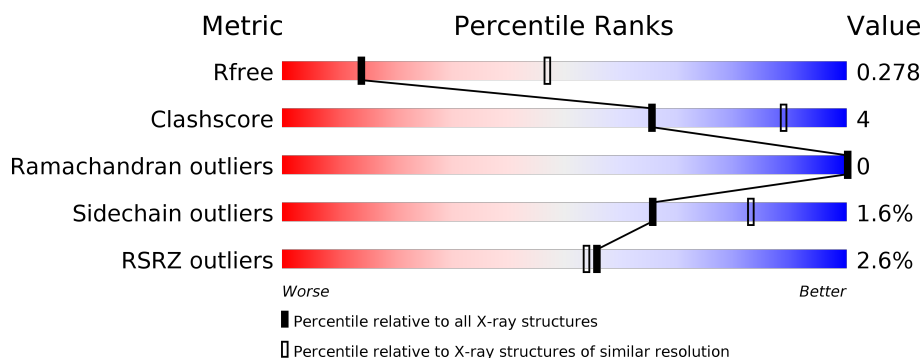
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 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))
$R_{free}$	111664	1168 (3.36-3.24)
Clashscore	122126	1022 (3.34-3.26)
Ramachandran outliers	120053	1004 (3.34-3.26)
Sidechain outliers	120020	1003 (3.34-3.26)
RSRZ outliers	108989	1133 (3.36-3.24)

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	436	<div> <div>2%</div> <div>76%</div> <div>11%</div> <div>12%</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3029 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 Soluble cytochrome b562,Adenosine receptor A1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	383	Total	C	N	O	S	0	0	0
			2989	1969	489	513	18			

There are 28 discrepancies between the modelled and reference sequences:

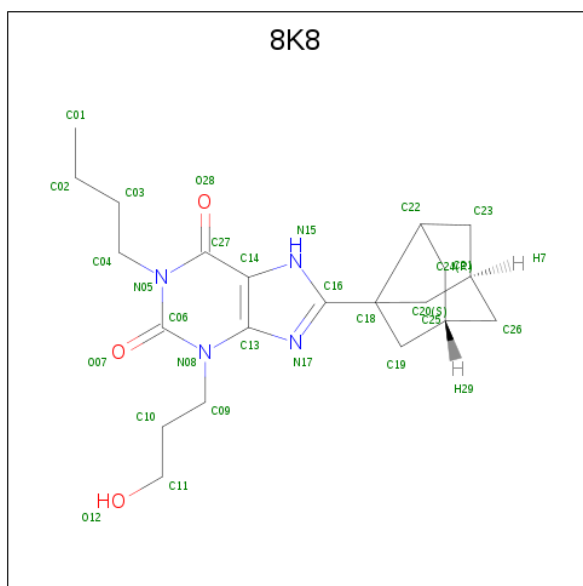
Chain	Residue	Modelled	Actual	Comment	Reference
A	999	ALA	-	expression tag	UNP P0ABE7
A	1000	SER	ALA	engineered mutation	UNP P0ABE7
A	1007	TRP	MET	engineered mutation	UNP P0ABE7
A	1102	ILE	HIS	engineered mutation	UNP P0ABE7
A	?	LEU	-	linker	UNP P0ABE7
A	?	ALA	-	linker	UNP P0ABE7
A	?	SER	-	linker	UNP P0ABE7
A	1162	LEU	ALA	engineered mutation	UNP P30542
A	1196	ALA	THR	engineered mutation	UNP P30542
A	1264	ALA	ASN	engineered mutation	UNP P30542
A	1310	ALA	TYR	engineered mutation	UNP P30542
A	1341	ALA	LEU	engineered mutation	UNP P30542
A	1345	ALA	LEU	engineered mutation	UNP P30542
A	1382	ALA	THR	engineered mutation	UNP P30542
A	1414	SER	CYS	engineered mutation	UNP P30542
A	1422	ALA	-	expression tag	UNP P30542
A	1423	ALA	-	expression tag	UNP P30542
A	1424	ALA	-	expression tag	UNP P30542
A	1425	HIS	-	expression tag	UNP P30542
A	1426	HIS	-	expression tag	UNP P30542
A	1427	HIS	-	expression tag	UNP P30542
A	1428	HIS	-	expression tag	UNP P30542
A	1429	HIS	-	expression tag	UNP P30542
A	1430	HIS	-	expression tag	UNP P30542
A	1431	HIS	-	expression tag	UNP P30542
A	1432	HIS	-	expression tag	UNP P30542
A	1433	HIS	-	expression tag	UNP P30542

*Continued on next page...*

Continued from previous page...

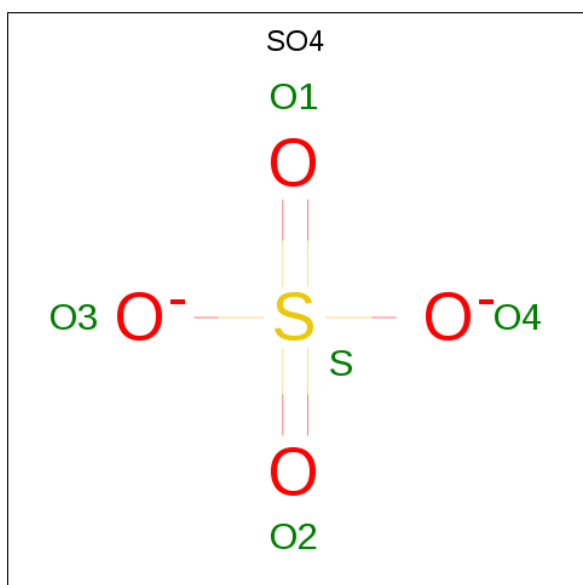
Chain	Residue	Modelled	Actual	Comment	Reference
A	1434	HIS	-	expression tag	UNP P30542

- Molecule 2 is 1-butyl-3-(3-oxidanylpropyl)-8-[(1 {R},5 {S})-3-tricyclo[3.3.1.0<sup>^</sup>{3,7}]nonanyl]-7 {H}-purine-2,6-dione (three-letter code: 8K8) (formula: C<sub>21</sub>H<sub>30</sub>N<sub>4</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			28	21	4	3		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		

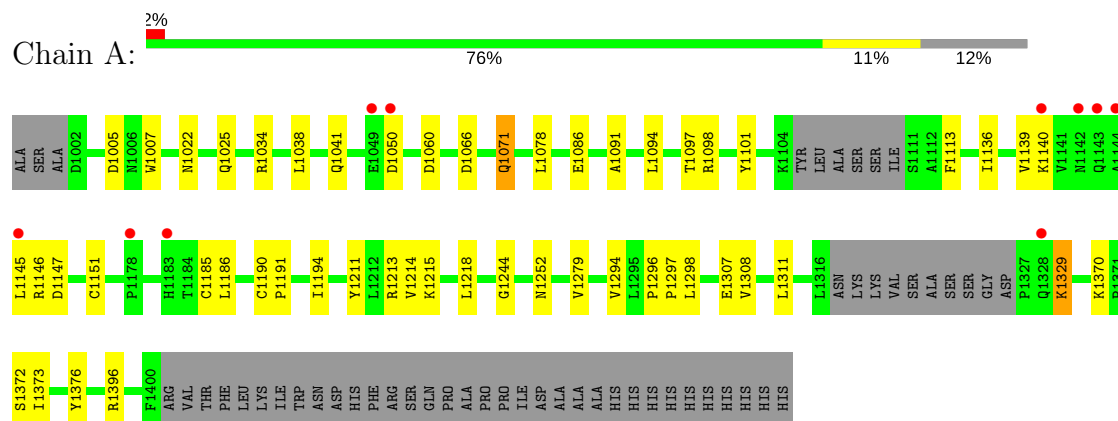
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total	O	0	0
			2	2		

### 3 Residue-property plots [i](#)

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: Soluble cytochrome b562, Adenosine receptor A1



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	70.27Å 111.66Å 160.54Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	32.89 – 3.30 33.27 – 3.30	Depositor EDS
% Data completeness (in resolution range)	97.2 (32.89-3.30) 97.5 (33.27-3.30)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.58 (at 3.32Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, $R_{free}$	0.256 , 0.278 0.256 , 0.278	Depositor DCC
$R_{free}$ test set	460 reflections (4.81%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	94.9	Xtriage
Anisotropy	0.771	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 65.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	3029	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	107.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.90% 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 [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: 8K8, SO4

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	0.25	0/3055	0.38	1/4154 (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	1329	LYS	CD-CE-NZ	-5.02	100.16	111.70

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	2989	0	3119	27	1
2	A	28	0	0	0	0
3	A	10	0	0	1	0
4	A	2	0	0	0	0
All	All	3029	0	3119	27	1

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 4.

The worst 5 of 27 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:1136:ILE:O	1:A:1140:LYS:HB2	1.94	0.68
1:A:1186:LEU:HD11	1:A:1252:ASN:HD22	1.64	0.61
1:A:1139:VAL:HA	1:A:1145:LEU:HD12	1.83	0.61
1:A:1194:ILE:HG12	1:A:1244:GLY:HA3	1.85	0.58
1:A:1140:LYS:O	1:A:1146:ARG:NH2	2.39	0.54

All (1) 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)
1:A:1329:LYS:CE	1:A:1329:LYS:NZ[3_855]	1.50	0.70

## 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	377/436 (86%)	359 (95%)	18 (5%)	0	100	100

There are no Ramachandran outliers to report.

### 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	318/362 (88%)	313 (98%)	5 (2%)	65	83

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1005	ASP
1	A	1060	ASP
1	A	1071	GLN
1	A	1151	CYS
1	A	1218	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1103	GLN
1	A	1252	ASN
1	A	1315	GLN
1	A	1385	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](#)

3 ligands are 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	8K8	A	2001	-	25,32,32	3.77	11 (44%)	21,49,49	2.82	8 (38%)
3	SO4	A	2002	-	4,4,4	0.17	0	6,6,6	0.07	0
3	SO4	A	2003	-	4,4,4	0.16	0	6,6,6	0.07	0

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	8K8	A	2001	-	-	0/8/42/42	0/6/5/5
3	SO4	A	2002	-	-	0/0/0/0	0/0/0/0
3	SO4	A	2003	-	-	0/0/0/0	0/0/0/0

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	2001	8K8	O28-C27	-2.75	1.17	1.24
2	A	2001	8K8	C09-N08	2.03	1.55	1.47
2	A	2001	8K8	C23-C24	2.22	1.62	1.52
2	A	2001	8K8	C18-C22	2.37	1.59	1.55
2	A	2001	8K8	C19-C18	2.77	1.58	1.54

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	2001	8K8	C14-C27-N05	-6.45	114.40	120.40
2	A	2001	8K8	C26-C20-C19	-5.71	100.64	110.32
2	A	2001	8K8	C23-C24-C25	-3.87	93.59	101.61
2	A	2001	8K8	C13-C14-N15	-3.39	106.29	109.47
2	A	2001	8K8	C21-C20-C19	2.40	106.57	101.61

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	2003	SO4	1	0

## 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	383/436 (87%)	0.02	10 (2%) 56 54	53, 101, 160, 206	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1144	ALA	7.6
1	A	1143	GLN	6.1
1	A	1142	ASN	2.9
1	A	1049	GLU	2.5
1	A	1328	GLN	2.3

### 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
3	SO4	A	2002	5/5	0.85	0.17	113,115,124,136	0
3	SO4	A	2003	5/5	0.89	0.16	140,144,148,153	0
2	8K8	A	2001	28/28	0.92	0.27	66,81,95,97	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.