



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2018 – 09:31 pm GMT

PDB ID : 3RYZ
Title : Fluoroalkyl and Alkyl Chains Have Similar Hydrophobicities in Binding to the Hydrophobic Wall of Carbonic Anhydrase
Authors : Snyder, P.W.; Bai, S.; Heroux, A.; Whitesides, G.W.
Deposited on : 2011-05-11
Resolution : 1.37 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

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
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

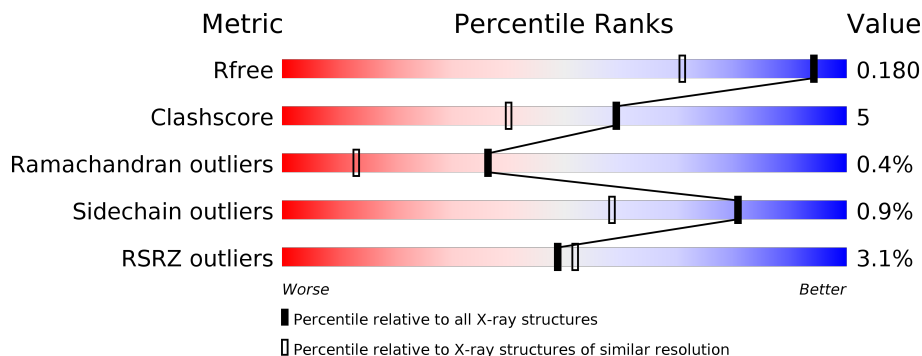
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.37 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	2404 (1.40-1.36)
Clashscore	122126	2520 (1.40-1.36)
Ramachandran outliers	120053	2464 (1.40-1.36)
Sidechain outliers	120020	2463 (1.40-1.36)
RSRZ outliers	108989	2346 (1.40-1.36)

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 ≥ 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	259	

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 2474 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.

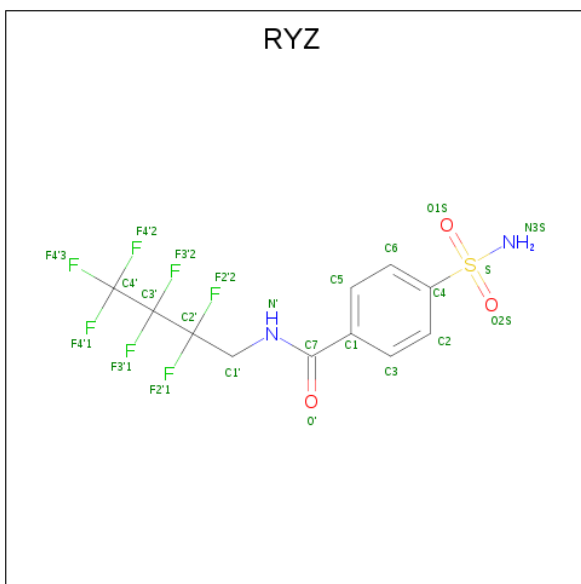
- Molecule 1 is a protein called Carbonic anhydrase 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	258	2065	1328	352	383	2	0	4	0

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
2	A	1	1	1	0	0

- Molecule 3 is N-(2,2,3,3,4,4,4-heptafluorobutyl)-4-sulfamoylbenzamide (three-letter code: RYZ) (formula: C₁₁H₉F₇N₂O₃S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	F	N	O			S
3	A	1	24	11	7	2	3	1	0	0

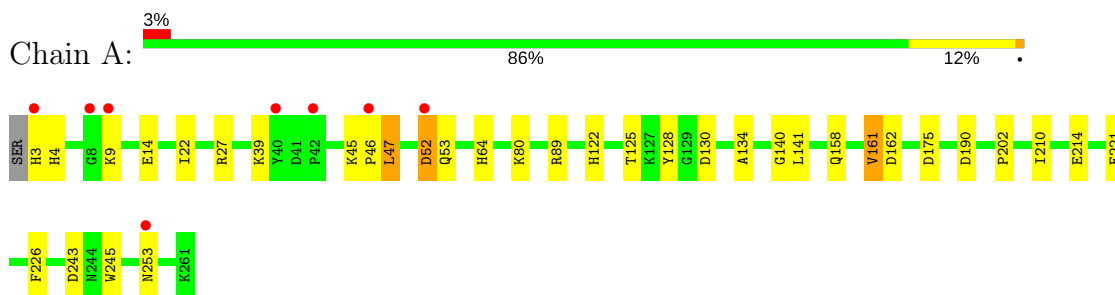
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	384	Total 384	O 384	0	0

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: Carbonic anhydrase 2



4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	42.41Å 41.70Å 72.36Å 90.00° 104.81° 90.00°	Depositor
Resolution (Å)	30.78 – 1.37 30.78 – 1.37	Depositor EDS
% Data completeness (in resolution range)	92.3 (30.78-1.37) 92.3 (30.78-1.37)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.18 (at 1.37Å)	Xtrriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.138 , 0.183 0.135 , 0.180	Depositor DCC
R_{free} test set	2413 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	10.6	Xtrriage
Anisotropy	0.035	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 39.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.026 for h,-k,-h-l	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	2474	wwPDB-VP
Average B, all atoms (Å ²)	14.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.56% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: ZN, RYZ

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	1.19	5/2138 (0.2%)	1.11	10/2903 (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	1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	161	VAL	CB-CG2	-8.37	1.35	1.52
1	A	214	GLU	CD-OE1	-7.86	1.17	1.25
1	A	253	ASN	C-O	6.36	1.35	1.23
1	A	14	GLU	CG-CD	6.07	1.61	1.51
1	A	128	TYR	CG-CD1	5.58	1.46	1.39

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	141	LEU	CB-CG-CD2	10.80	129.37	111.00
1	A	162	ASP	CB-CG-OD1	7.01	124.61	118.30
1	A	47	LEU	CA-CB-CG	6.81	130.97	115.30
1	A	190	ASP	CB-CG-OD2	-6.50	112.45	118.30
1	A	226	PHE	CB-CG-CD2	-6.43	116.30	120.80
1	A	130	ASP	CB-CG-OD1	6.08	123.78	118.30
1	A	162	ASP	CB-CG-OD2	-5.83	113.05	118.30
1	A	226	PHE	CB-CG-CD1	5.69	124.78	120.80

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	27	ARG	NE-CZ-NH1	-5.09	117.76	120.30
1	A	221	GLU	OE1-CD-OE2	-5.04	117.26	123.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	3	HIS	Peptide

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	2065	0	2017	18	0
2	A	1	0	0	0	0
3	A	24	0	9	4	0
4	A	384	0	0	5	1
All	All	2474	0	2026	21	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 5.

All (21) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (\AA)	Clash overlap (\AA)
3:A:262:RYZ:F2'1	3:A:262:RYZ:F4'2	2.57	0.80
1:A:64[B]:HIS:HE1	4:A:295:HOH:O	1.67	0.76
3:A:262:RYZ:F2'2	3:A:262:RYZ:F4'3	2.63	0.74
1:A:80:LYS:HE2	4:A:305:HOH:O	1.93	0.68
1:A:158:GLN:HA	1:A:161:VAL:CG2	2.33	0.59
1:A:158:GLN:HA	1:A:161:VAL:HG22	1.85	0.58
1:A:45:LYS:HB3	1:A:46:PRO:HD2	1.89	0.55
1:A:64[B]:HIS:CE1	4:A:295:HOH:O	2.50	0.54
1:A:202:PRO:HD2	3:A:262:RYZ:F2'1	2.38	0.53
1:A:89:ARG:HG3	1:A:125:THR:CG2	2.40	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:47:LEU:HD21	1:A:210:ILE:HD13	1.93	0.51
1:A:134:ALA:O	1:A:140:GLY:HA3	2.11	0.49
1:A:158:GLN:O	1:A:161:VAL:CG2	2.66	0.44
1:A:89:ARG:O	1:A:122:HIS:HA	2.19	0.42
1:A:9:LYS:C	4:A:291:HOH:O	2.58	0.42
3:A:262:RYZ:F2'2	3:A:262:RYZ:F4'2	2.89	0.41
1:A:158:GLN:O	1:A:161:VAL:HG23	2.21	0.41
1:A:243:ASP:HA	1:A:245:TRP:CD1	2.56	0.41
1:A:22[A]:ILE:HA	1:A:22[A]:ILE:HD12	1.98	0.40
1:A:39:LYS:HE2	4:A:419:HOH:O	2.22	0.40
1:A:52:ASP:OD2	1:A:53:GLN:HG2	2.22	0.40

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)
4:A:381:HOH:O	4:A:643:HOH:O[1_565]	1.99	0.21

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	260/259 (100%)	248 (95%)	11 (4%)	1 (0%)	36 13

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	4	HIS

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	224/224 (100%)	222 (99%)	2 (1%)	81 58

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	52	ASP
1	A	175	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	253	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 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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
3	RYZ	A	262	2	24,24,24	2.01	6 (25%)	39,39,39	1.56	6 (15%)

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
3	RYZ	A	262	2	-	0/34/34/34	0/1/1/1

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	262	RYZ	C4-S	-5.99	1.67	1.77
3	A	262	RYZ	F2'2-C2'	-2.05	1.30	1.36
3	A	262	RYZ	O1S-S	2.09	1.47	1.43
3	A	262	RYZ	C6-C5	2.50	1.43	1.38
3	A	262	RYZ	C2-C3	3.36	1.44	1.38
3	A	262	RYZ	C6-C4	4.39	1.45	1.38

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	262	RYZ	F2'2-C2'-F2'1	-4.02	96.02	105.98
3	A	262	RYZ	C2'-C1'-N'	-2.03	109.59	112.09
3	A	262	RYZ	F3'2-C3'-C4'	2.11	115.75	106.65
3	A	262	RYZ	F4'2-C4'-F4'1	2.41	115.01	107.54
3	A	262	RYZ	C6-C4-S	2.88	124.13	119.75
3	A	262	RYZ	F2'1-C2'-C3'	3.07	116.99	107.19

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	262	RYZ	4	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, 95th 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 ²)	Q<0.9
1	A	258/259 (99%)	-0.01	8 (3%) 49 52	5, 11, 22, 39	0

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	3	HIS	10.3
1	A	9	LYS	3.8
1	A	253	ASN	3.6
1	A	42	PRO	3.3
1	A	40	TYR	3.2
1	A	46	PRO	3.0
1	A	52	ASP	2.1
1	A	8	GLY	2.0

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, 95th 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(\AA^2)	Q<0.9
3	RYZ	A	262	24/24	0.97	0.08	5,10,22,26	0
2	ZN	A	1	1/1	1.00	0.05	5,5,5,5	0

6.5 Other polymers [i](#)

There are no such residues in this entry.