



# Full wwPDB X-ray Structure Validation Report ⓘ

May 27, 2024 – 07:46 PM EDT

PDB ID : 6GCY  
Title : Joint neutron and x-ray crystal structure of human carbonic anhydrase IX mimic (saccharin-sugar conjugate complex)  
Authors : Fisher, S.Z.; Koruza, K.  
Deposited on : 2018-04-20  
Resolution : 1.30 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	<b>FAILED</b>
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

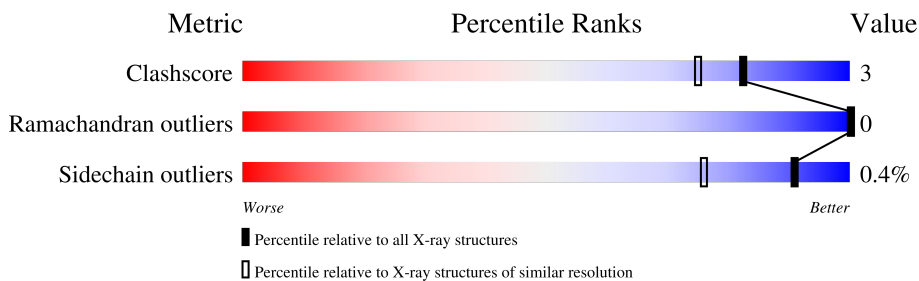
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION, NEUTRON DIFFRACTION*

The reported resolution of this entry is 1.30 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))
Clashscore	141614	1101 (1.30-1.30)
Ramachandran outliers	138981	1058 (1.30-1.30)
Sidechain outliers	138945	1058 (1.30-1.30)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. 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  $\leq 5\%$ .

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	260	97%

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
3	3QR	A	302	X	-	-	-

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5307 atoms, of which 2067 are hydrogens and 859 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
1	A	258	Total	C	D	H	N	O	S	77	196	0
			4635	1384	439	2054	359	397	2			

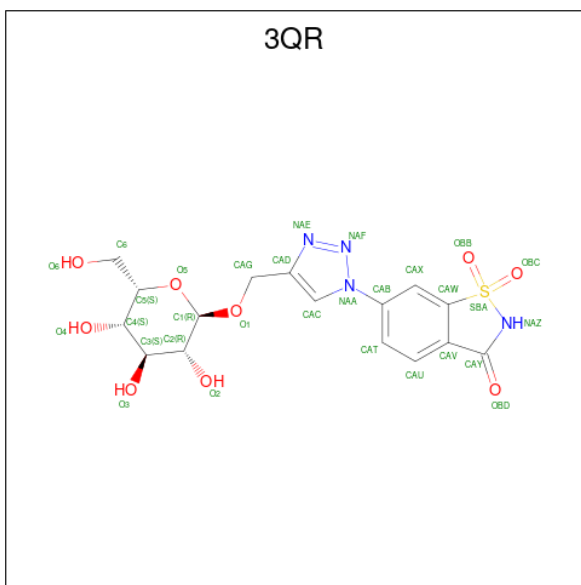
There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	65	SER	ALA	engineered mutation	UNP P00918
A	67	GLN	ASN	engineered mutation	UNP P00918
A	69	THR	GLU	engineered mutation	UNP P00918
A	91	LEU	ILE	engineered mutation	UNP P00918
A	130	VAL	PHE	engineered mutation	UNP P00918
A	169	GLU	LYS	engineered mutation	UNP P00918
A	203	ALA	LEU	engineered mutation	UNP P00918

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

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

- Molecule 3 is [1-(1,1-dioxido-3-oxo-2,3-dihydro-1,2-benzothiazol-6-yl)-1H-1,2,3-triazol-4-yl] methyl alpha-L-idopyranoside (three-letter code: 3QR) (formula: C<sub>16</sub>H<sub>18</sub>N<sub>4</sub>O<sub>9</sub>S).



Mol	Chain	Residues	Atoms							ZeroOcc	AltConf
3	A	1	Total	C	D	H	N	O	S	0	0
			47	16	4	13	4	9	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	208	Total	D	O	0	0
			624	416	208		

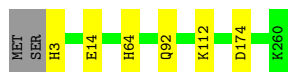
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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. 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. 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 failed to run properly.

- Molecule 1: Carbonic anhydrase 2

Chain A:  97% ..



## 4 Data and refinement statistics

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	42.61Å 41.94Å 72.77Å 90.00° 104.01° 90.00°	Depositor
Resolution (Å)	20.97 – 1.30	Depositor
% Data completeness (in resolution range)	97.6 (20.97-1.30)	Depositor
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.21 (at 1.30Å)	Xtriage
Refinement program	PHENIX (1.12_2829: ???)	Depositor
R, $R_{free}$	0.148 , 0.155	Depositor
Wilson B-factor (Å <sup>2</sup> )	11.3	Xtriage
Anisotropy	0.137	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.025 for h,-k,-h-l	Xtriage
Total number of atoms	5307	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.97% 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: DOD, 3QR, ZN

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	0.52	0/3840	0.70	0/5202

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	2581	2054	768	8	0
2	A	1	0	0	0	0
3	A	34	13	17	2	0
4	A	624	0	0	7	1
All	All	3240	2067	785	11	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 3.

All (11) 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:64[B]:HIS:CE1	4:A:408:DOD:O	2.32	0.83
1:A:112[B]:LYS:NZ	4:A:401:DOD:O	1.98	0.79
1:A:64[A]:HIS:HD2	4:A:563:DOD:O	1.95	0.45
1:A:3:HIS:CE1	4:A:448:DOD:O	2.71	0.43
1:A:64[B]:HIS:ND1	4:A:408:DOD:O	2.41	0.42
1:A:64[A]:HIS:CD2	4:A:563:DOD:O	2.72	0.42

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:415:DOD:O	4:A:435:DOD:O[1_455]	2.10	0.10

## 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	466/260 (179%)	451 (97%)	15 (3%)	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	403/225 (179%)	401 (100%)	2 (0%)	88	69



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

Mol	Chain	Res	Type
1	A	92[A]	GLN
1	A	92[B]	GLN

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

Mol	Chain	Res	Type
1	A	3	HIS

### 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 monosaccharides 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	3QR	A	302	2	32,33,33	3.93	14 (43%)	43,50,50	4.64	28 (65%)

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	3QR	A	302	2	1/1/7/8	8/9/46/46	0/4/4/4

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	302	3QR	OBB-SBA	11.45	1.56	1.43
3	A	302	3QR	OBC-SBA	9.25	1.54	1.43
3	A	302	3QR	CAV-CAY	7.84	1.60	1.48
3	A	302	3QR	CAW-SBA	5.98	1.82	1.75
3	A	302	3QR	CAU-CAV	-5.32	1.31	1.39
3	A	302	3QR	OBD-CAY	-5.17	1.12	1.23
3	A	302	3QR	CAD-NAE	-4.98	1.27	1.34
3	A	302	3QR	CAV-CAW	4.98	1.44	1.38
3	A	302	3QR	CAY-NAZ	4.02	1.44	1.37
3	A	302	3QR	NAF-NAA	-3.72	1.31	1.37
3	A	302	3QR	CAT-CAU	-3.10	1.33	1.38
3	A	302	3QR	O3-C3	2.34	1.48	1.43
3	A	302	3QR	CAC-CAD	2.08	1.39	1.36
3	A	302	3QR	CAX-CAB	-2.03	1.36	1.38

All (28) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	302	3QR	CAX-CAW-CAV	-17.73	115.72	123.79
3	A	302	3QR	OBD-CAY-NAZ	-10.44	118.89	124.37
3	A	302	3QR	O1-C1-C2	10.19	124.22	108.30
3	A	302	3QR	CAU-CAV-CAW	7.58	127.78	119.91
3	A	302	3QR	C1-O5-C5	5.74	124.95	113.69
3	A	302	3QR	O4-C4-C3	5.29	122.58	110.35
3	A	302	3QR	C3-C4-C5	-4.87	101.56	110.24
3	A	302	3QR	O2-C2-C3	-4.55	99.82	110.35
3	A	302	3QR	OBC-SBA-NAZ	4.50	113.67	110.39
3	A	302	3QR	CAG-O1-C1	4.25	120.85	113.31
3	A	302	3QR	CAX-CAW-SBA	4.13	133.20	127.09
3	A	302	3QR	O6-C6-C5	-4.03	97.46	111.29
3	A	302	3QR	O2-C2-C1	3.80	119.27	110.05
3	A	302	3QR	CAT-CAU-CAV	-3.68	114.97	120.94
3	A	302	3QR	CAB-CAX-CAW	-3.66	116.26	120.88
3	A	302	3QR	CAW-CAV-CAY	-3.63	108.62	112.08
3	A	302	3QR	CAB-NAA-NAF	3.44	124.17	117.19
3	A	302	3QR	OBC-SBA-CAW	-3.21	106.58	111.59

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	302	3QR	O5-C5-C6	-3.17	98.55	106.44
3	A	302	3QR	O5-C5-C4	3.17	115.44	109.69
3	A	302	3QR	O5-C1-O1	-3.16	102.50	109.97
3	A	302	3QR	CAX-CAB-NAA	-2.99	115.25	119.07
3	A	302	3QR	CAU-CAV-CAY	-2.88	123.60	129.37
3	A	302	3QR	O3-C3-C2	-2.77	103.95	110.35
3	A	302	3QR	O3-C3-C4	2.56	116.27	110.35
3	A	302	3QR	OBC-SBA-OB	2.48	119.50	116.27
3	A	302	3QR	CAT-CAB-CAX	2.23	124.38	121.74
3	A	302	3QR	OB-SBA-NAZ	-2.16	108.81	110.39

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	A	302	3QR	C5

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	302	3QR	O5-C5-C6-O6
3	A	302	3QR	O5-C1-O1-CAG
3	A	302	3QR	C2-C1-O1-CAG
3	A	302	3QR	C4-C5-C6-O6
3	A	302	3QR	CAX-CAB-NAA-NAF
3	A	302	3QR	CAT-CAB-NAA-NAF
3	A	302	3QR	CAT-CAB-NAA-CAC
3	A	302	3QR	CAX-CAB-NAA-CAC

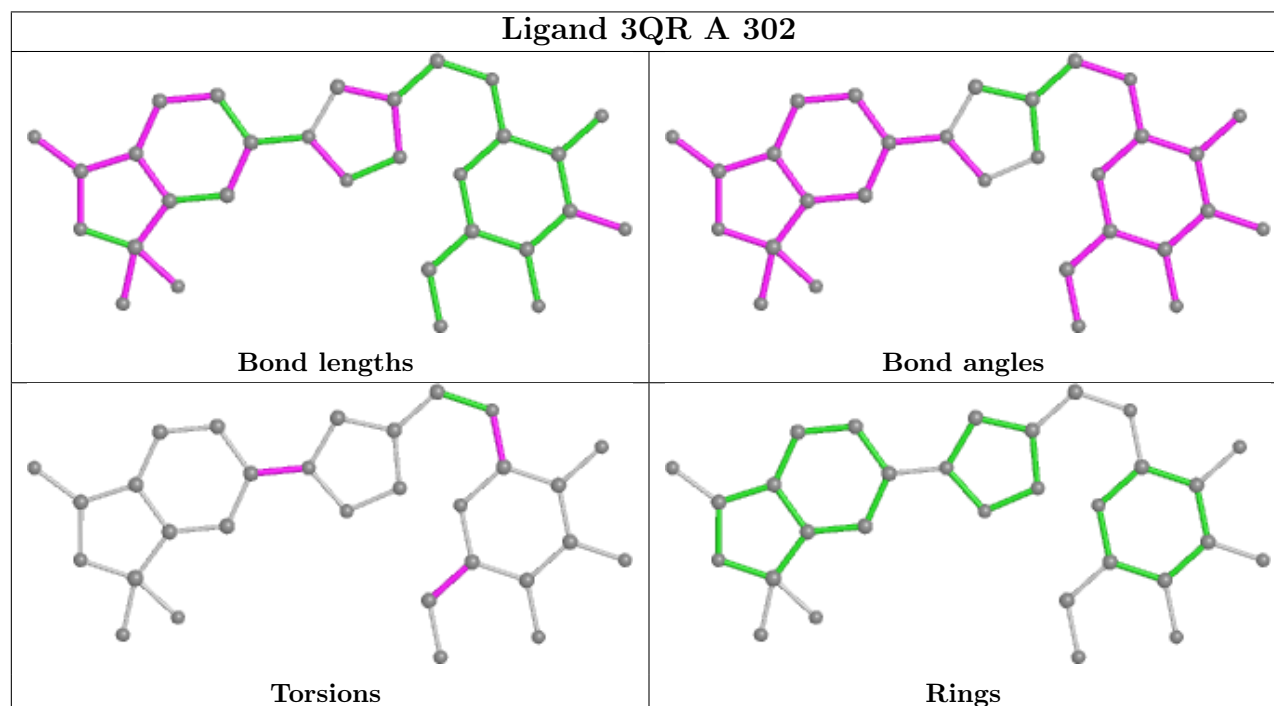
There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	302	3QR	2	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the

average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



## 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 failed to run properly - this section is therefore empty.

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

EDS failed to run properly - this section is therefore empty.

### 6.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

### 6.4 Ligands

EDS failed to run properly - this section is therefore empty.

### 6.5 Other polymers

EDS failed to run properly - this section is therefore empty.