wwPDB X-ray Structure Validation Summary Report (i)

Dec 7, 2023 – 12:31 pm GMT

PDB ID	:	2CGN
Title	:	FACTOR INHIBITING HIF-1 ALPHA with succinate
Authors	:	McDonough, M.A.; Clifton, I.J.; Schofield, C.J.
Deposited on	:	2006-03-09
Resolution	:	2.40 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.40 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.
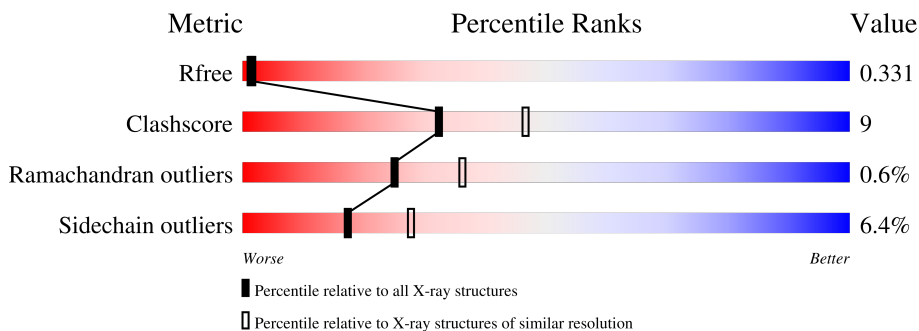
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)

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 >=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 <=5%

Mol	Chain	Length	Quality of chain
1	A	349	73%　　　　21%　　·　·

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 2858 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 HYPOXIA-INDUCIBLE FACTOR 1 ALPHA INHIBITOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	338	Total 2768	C 1772	N 469	O 516	S 11	0	1	0

- Molecule 2 is FE (III) ION (three-letter code: FE) (formula: Fe).

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

- Molecule 3 is SUCCINIC ACID (three-letter code: SIN) (formula: $C_4H_6O_4$).
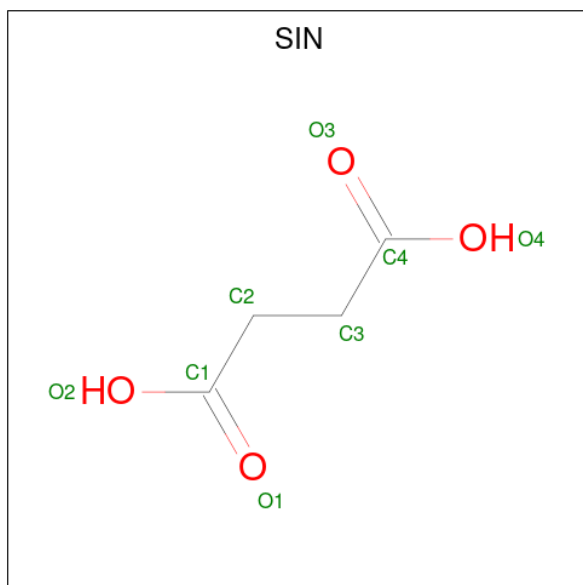
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 8	C 4	O 4	0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

SO4

O1

O3 O⁻ — S — O⁻ O4

S

O2

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

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	61	Total 61	O 61	0	0

3 Residue-property plots ⓘ

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.

• Molecule 1: HYPOXIA-INDUCIBLE FACTOR 1 ALPHA INHIBITOR

Chain A:
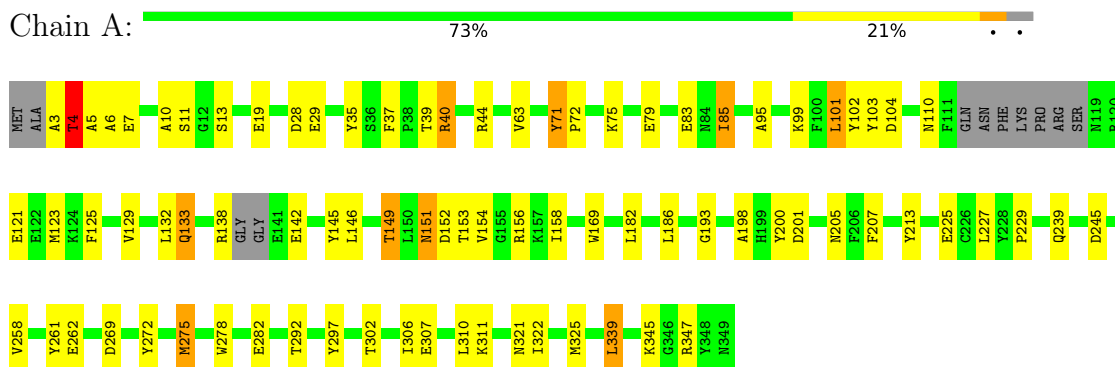
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	86.61Å 86.61Å 145.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	74.54 – 2.40	Depositor
	34.21 – 2.40	EDS
% Data completeness	98.3 (74.54-2.40)	Depositor
(in resolution range)	98.4 (34.21-2.40)	EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.79 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.205 , 0.263	Depositor
	0.286 , 0.331	DCC
R_{free} test set	2191 reflections (9.94%)	wwPDB-VP
Wilson B-factor (Å^2)	59.4	Xtriage
Anisotropy	0.364	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 62.9	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	2858	wwPDB-VP
Average B, all atoms (Å^2)	69.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.57% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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: SIN, FE, 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	1.36	11/2850 (0.4%)	1.11	13/3869 (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

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	71	TYR	CE1-CZ	-7.65	1.28	1.38
1	A	282	GLU	CG-CD	6.87	1.62	1.51
1	A	200	TYR	CE1-CZ	6.81	1.47	1.38
1	A	35	TYR	CD1-CE1	-6.20	1.30	1.39
1	A	198	ALA	CA-CB	5.92	1.64	1.52

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	347	ARG	NE-CZ-NH2	-12.77	113.92	120.30
1	A	347	ARG	NE-CZ-NH1	10.66	125.63	120.30
1	A	201	ASP	CB-CG-OD2	9.85	127.16	118.30
1	A	201	ASP	CB-CG-OD1	-9.80	109.48	118.30
1	A	44	ARG	NE-CZ-NH2	6.90	123.75	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Chain	Res	Type	Group
1	A	4	THR	Peptide

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	2768	0	2647	48	1
2	A	1	0	0	0	0
3	A	8	0	4	0	0
4	A	20	0	0	2	0
5	A	61	0	0	0	0
All	All	2858	0	2651	48	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 9.

The worst 5 of 48 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:4:THR:HG23	1:A:5:ALA:H	1.46	0.80
1:A:4:THR:CG2	1:A:5:ALA:H	2.01	0.73
1:A:101:LEU:HD12	1:A:101:LEU:O	1.95	0.67
1:A:149:THR:N	1:A:186:LEU:HD12	2.12	0.65
1:A:71:TYR:HB3	1:A:72:PRO:HD3	1.79	0.65

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:71:TYR:OH	1:A:225[A]:GLU:OE1[5_545]	2.17	0.03

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	333/349 (95%)	313 (94%)	18 (5%)	2 (1%)	25	36

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	4	THR
1	A	110	ASN

5.3.2 Protein sidechains ⓘ

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	296/304 (97%)	277 (94%)	19 (6%)	17	28

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	205	ASN
1	A	302	THR
1	A	310	LEU
1	A	275	MET
1	A	138	ARG

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

Mol	Chain	Res	Type
1	A	151	ASN
1	A	205	ASN
1	A	233	HIS
1	A	313	HIS

5.3.3　RNA ⓘ

There are no RNA molecules in this entry.

5.4　Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5　Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6　Ligand geometry ⓘ

Of 6 ligands modelled in this entry, 1 is monoatomic - leaving 5 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	$\#|Z| > 2$
4	SO4	A	1354	2	4,4,4	0.48	0	6,6,6	0.65	0
4	SO4	A	1355	-	4,4,4	0.29	0	6,6,6	0.38	0
4	SO4	A	1352	-	4,4,4	0.13	0	6,6,6	0.57	0
3	SIN	A	1351	2	7,7,7	1.22	1 (14%)	8,8,8	1.49	2 (25%)
4	SO4	A	1353	-	4,4,4	0.15	0	6,6,6	0.55	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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	SIN	A	1351	2	-	2/5/5/5	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1351	SIN	O3-C4	2.38	1.30	1.22

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1351	SIN	O4-C4-C3	2.56	122.25	114.03
3	A	1351	SIN	O3-C4-C3	-2.54	114.93	123.08

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1351	SIN	C2-C3-C4-O4
3	A	1351	SIN	C2-C3-C4-O3

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1355	SO4	1	0
4	A	1353	SO4	1	0

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 ⓘ

6.1   Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3   Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4   Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5   Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.