



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 17, 2020 – 09:54 am BST

PDB ID : 1DCH  
Title : CRYSTAL STRUCTURE OF DCOH, A BIFUNCTIONAL, PROTEIN-BINDING TRANSCRIPTION COACTIVATOR  
Authors : Endrizzi, J.A.; Cronk, J.D.; Wang, W.; Crabtree, G.R.; Alber, T.  
Deposited on : 1995-01-24  
Resolution : 3.00 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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
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) : 2.11

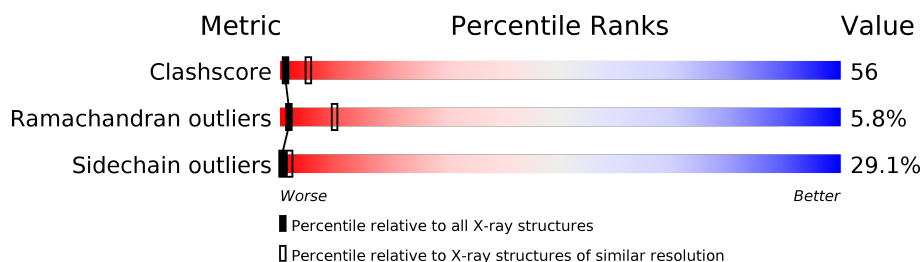
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)

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 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 was not executed.

Mol	Chain	Length	Quality of chain
1	A	104	
1	B	104	
1	C	104	
1	D	104	
1	E	104	
1	F	104	
1	G	104	
1	H	104	

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
2	SO4	A	200	-	-	X	-
2	SO4	F	200	-	-	X	-
2	SO4	G	200	-	-	X	-

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6494 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 DCOH (DIMERIZATION COFACTOR OF HNF-1).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	99	Total	C	N	O	S	0	0	0
			806	512	145	146	3			
1	B	99	Total	C	N	O	S	0	0	0
			806	512	145	146	3			
1	C	99	Total	C	N	O	S	0	0	0
			810	514	146	147	3			
1	D	99	Total	C	N	O	S	0	0	0
			803	511	146	143	3			
1	E	99	Total	C	N	O	S	0	0	0
			802	510	146	143	3			
1	F	99	Total	C	N	O	S	0	0	0
			807	513	146	145	3			
1	G	99	Total	C	N	O	S	0	0	0
			810	514	146	147	3			
1	H	99	Total	C	N	O	S	0	0	0
			810	514	146	147	3			

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



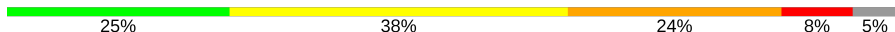
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		
2	E	1	Total	O	S	0	0
			5	4	1		
2	F	1	Total	O	S	0	0
			5	4	1		
2	G	1	Total	O	S	0	0
			5	4	1		
2	H	1	Total	O	S	0	0
			5	4	1		

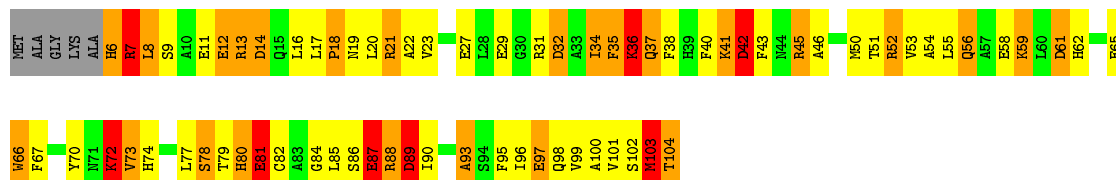
### 3 Residue-property plots

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. 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 was not executed.

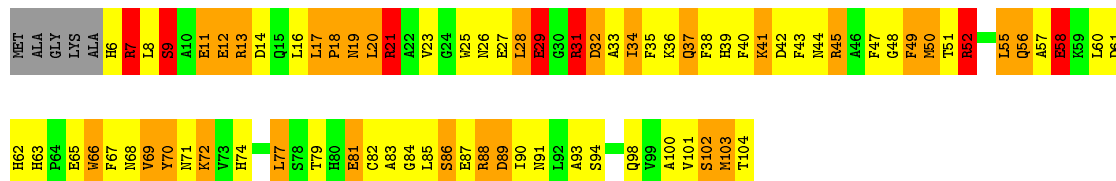
#### • Molecule 1: DCOH (DIMERIZATION COFACTOR OF HNF-1)

Chain A:  25% 38% 24% 8% 5%




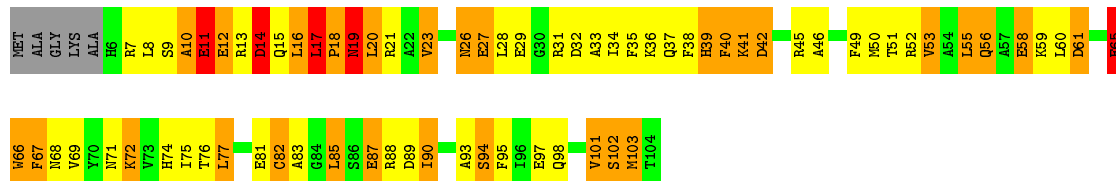
#### • Molecule 1: DCOH (DIMERIZATION COFACTOR OF HNF-1)

Chain B:  19% 42% 27% 7% 5%

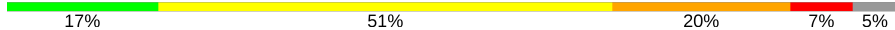


#### • Molecule 1: DCOH (DIMERIZATION COFACTOR OF HNF-1)

Chain C:  26% 37% 28% 5% 5%



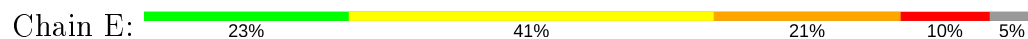
#### • Molecule 1: DCOH (DIMERIZATION COFACTOR OF HNF-1)

Chain D:  17% 51% 20% 7% 5%

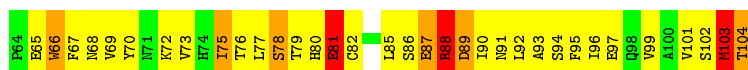
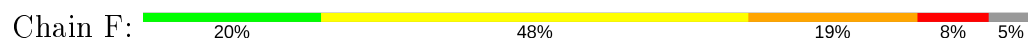




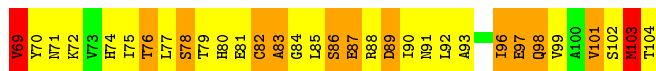
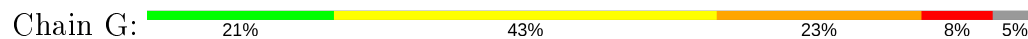
- Molecule 1: DCOH (DIMERIZATION COFACTOR OF HNF-1)



- Molecule 1: DCOH (DIMERIZATION COFACTOR OF HNF-1)



- Molecule 1: DCOH (DIMERIZATION COFACTOR OF HNF-1)



- Molecule 1: DCOH (DIMERIZATION COFACTOR OF HNF-1)



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	105.65Å 105.65Å 196.23Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	6.00 – 3.00	Depositor
% Data completeness (in resolution range)	(Not available) (6.00-3.00)	Depositor
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	TNT	Depositor
R, $R_{free}$	0.184 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	6494	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	12.0	wwPDB-VP



## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 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	1.33	11/825 (1.3%)	1.71	22/1114 (2.0%)
1	B	1.32	7/825 (0.8%)	1.73	22/1114 (2.0%)
1	C	1.33	10/829 (1.2%)	1.72	16/1119 (1.4%)
1	D	1.26	7/822 (0.9%)	1.65	12/1110 (1.1%)
1	E	1.24	6/821 (0.7%)	1.66	18/1109 (1.6%)
1	F	1.26	7/826 (0.8%)	1.73	23/1115 (2.1%)
1	G	1.25	7/829 (0.8%)	1.72	21/1119 (1.9%)
1	H	1.36	7/829 (0.8%)	1.64	17/1119 (1.5%)
All	All	1.29	62/6606 (0.9%)	1.70	151/8919 (1.7%)

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
1	C	1	1
1	D	0	1
1	F	0	1
1	G	0	3
1	H	0	3
All	All	1	10

The worst 5 of 62 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	11	GLU	CD-OE1	11.23	1.38	1.25
1	E	81	GLU	CD-OE1	11.13	1.37	1.25
1	H	27	GLU	CD-OE2	11.11	1.37	1.25
1	G	81	GLU	CD-OE1	10.92	1.37	1.25

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	87	GLU	CD-OE1	10.22	1.36	1.25

The worst 5 of 151 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	52	ARG	NE-CZ-NH1	10.98	125.79	120.30
1	E	42	ASP	CB-CG-OD2	10.33	127.60	118.30
1	G	89	ASP	CB-CG-OD1	-10.28	109.05	118.30
1	D	42	ASP	CB-CG-OD1	-10.19	109.13	118.30
1	F	31	ARG	NE-CZ-NH1	10.09	125.35	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	C	17	LEU	CA

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	103	MET	Peptide
1	C	11	GLU	Mainchain
1	D	81	GLU	Mainchain
1	F	34	ILE	Mainchain
1	G	26	ASN	Mainchain

## 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	806	0	771	113	0
1	B	806	0	771	99	0
1	C	810	0	777	87	0
1	D	803	0	771	83	0
1	E	802	0	769	105	0
1	F	807	0	775	93	0
1	G	810	0	777	84	0
1	H	810	0	777	89	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	5	0	0	2	0
2	B	5	0	0	1	0
2	C	5	0	0	0	0
2	D	5	0	0	0	0
2	E	5	0	0	1	0
2	F	5	0	0	3	0
2	G	5	0	0	2	0
2	H	5	0	0	1	0
All	All	6494	0	6188	711	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 56.

The worst 5 of 711 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:E:8:LEU:HD21	1:E:90:ILE:HD11	1.20	1.15
1:D:34:ILE:HG23	1:D:85:LEU:HD11	1.23	1.14
1:A:8:LEU:HB3	1:A:13:ARG:HB3	1.26	1.14
1:E:8:LEU:HB3	1:E:13:ARG:HD2	1.27	1.09
1:E:19:ASN:H	1:E:19:ASN:ND2	1.53	1.05

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	97/104 (93%)	76 (78%)	16 (16%)	5 (5%)	2	12
1	B	97/104 (93%)	77 (79%)	15 (16%)	5 (5%)	2	12
1	C	97/104 (93%)	71 (73%)	22 (23%)	4 (4%)	3	16

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	97/104 (93%)	78 (80%)	13 (13%)	6 (6%)	1	8
1	E	97/104 (93%)	75 (77%)	13 (13%)	9 (9%)	0	3
1	F	97/104 (93%)	78 (80%)	11 (11%)	8 (8%)	1	4
1	G	97/104 (93%)	82 (84%)	11 (11%)	4 (4%)	3	16
1	H	97/104 (93%)	82 (84%)	11 (11%)	4 (4%)	3	16
All	All	776/832 (93%)	619 (80%)	112 (14%)	45 (6%)	1	10

5 of 45 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	7	ARG
1	A	78	SER
1	B	9	SER
1	C	17	LEU
1	D	10	ALA

### 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	84/88 (96%)	62 (74%)	22 (26%)	0	2
1	B	84/88 (96%)	58 (69%)	26 (31%)	0	1
1	C	85/88 (97%)	57 (67%)	28 (33%)	0	1
1	D	83/88 (94%)	59 (71%)	24 (29%)	0	2
1	E	83/88 (94%)	55 (66%)	28 (34%)	0	1
1	F	84/88 (96%)	64 (76%)	20 (24%)	0	3
1	G	85/88 (97%)	62 (73%)	23 (27%)	0	2
1	H	85/88 (97%)	60 (71%)	25 (29%)	0	1
All	All	673/704 (96%)	477 (71%)	196 (29%)	0	2

5 of 196 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	42	ASP
1	E	31	ARG
1	H	41	LYS
1	D	60	LEU
1	D	104	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 39 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	80	HIS
1	E	39	HIS
1	H	26	ASN
1	D	91	ASN
1	E	19	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](#)

8 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	SO4	C	200	-	4,4,4	0.82	0	6,6,6	0.37	0
2	SO4	A	200	-	4,4,4	0.61	0	6,6,6	0.39	0
2	SO4	G	200	-	4,4,4	0.57	0	6,6,6	0.37	0
2	SO4	E	200	-	4,4,4	0.95	0	6,6,6	0.24	0
2	SO4	D	200	-	4,4,4	0.78	0	6,6,6	0.32	0
2	SO4	B	200	-	4,4,4	0.91	0	6,6,6	0.45	0
2	SO4	H	200	-	4,4,4	0.89	0	6,6,6	0.54	0
2	SO4	F	200	-	4,4,4	0.63	0	6,6,6	0.21	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	200	SO4	2	0
2	G	200	SO4	2	0
2	E	200	SO4	1	0
2	B	200	SO4	1	0
2	H	200	SO4	1	0
2	F	200	SO4	3	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 ⓘ

### 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.