

# Full wwPDB X-ray Structure Validation Report (i)

Jan 5, 2024 - 01:18 am GMT

PDB ID : 4V2K

Title : Crystal structure of the thiosulfate dehydrogenase TsdA in complex with thio-

sulfate

Authors: Grabarczyk, D.B.; Chappell, P.E.; Eisel, B.; Johnson, S.; Lea, S.M.; Berks,

B.C.

Deposited on : 2014-10-10

Resolution : 1.29 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 (i) 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 (i)) were used in the production of this report:

MolProbity: 4.02b-467

Mogul : 1.8.4, CSD as 541be (2020)

Xtriage (Phenix) : 1.13

EDS : 2.36

buster-report : 1.1.7 (2018)

Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)

 $Refmac \quad : \quad 5.8.0158$ 

CCP4 : 7.0.044 (Gargrove)

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

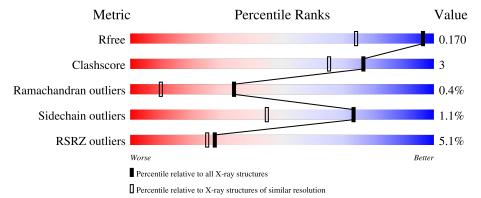
Validation Pipeline (wwPDB-VP) : 2.36

## 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: X- $RAY\ DIFFRACTION$ 

The reported resolution of this entry is 1.29 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	Similar resolution
Metric	$(\#  ext{Entries})$	$(\#  ext{Entries},  ext{ resolution range}( ext{Å}))$
$R_{free}$	130704	1058 (1.30-1.30)
Clashscore	141614	1101 (1.30-1.30)
Ramachandran outliers	138981	1058 (1.30-1.30)
Sidechain outliers	138945	1058 (1.30-1.30)
RSRZ outliers	127900	1029 (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 >=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% 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	
			5%	
1	A	253	83%	9% • 7%



# 2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2080 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 THIOSULFATE DEHYDROGENASE.

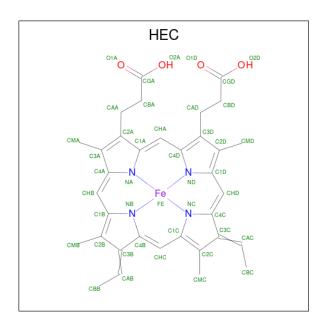
Mo	l Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	A	235	Total 1784	C 1122	N 325	O 326	S 11	0	5	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	271	ALA	-	expression tag	UNP D3RVD4
A	272	SER	-	expression tag	UNP D3RVD4
A	273	TRP	-	expression tag	UNP D3RVD4
A	274	SER	-	expression tag	UNP D3RVD4
A	275	HIS	-	expression tag	UNP D3RVD4
A	276	PRO	_	expression tag	UNP D3RVD4
A	277	GLN	-	expression tag	UNP D3RVD4
A	278	PHE	-	expression tag	UNP D3RVD4
A	279	GLU	-	expression tag	UNP D3RVD4
A	280	LYS	-	expression tag	UNP D3RVD4

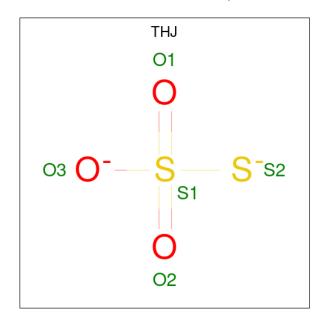
• Molecule 2 is HEME C (three-letter code: HEC) (formula: C<sub>34</sub>H<sub>34</sub>FeN<sub>4</sub>O<sub>4</sub>).





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf		
2	Λ	1	Total	С	Fe	N	О	0	0	
	Z A	1	43	34	1	4	4	0		
2	Λ	1	Total	С	Fe	N	О	0	1	
	A	1	47	36	1	4	6	0		

 $\bullet$  Molecule 3 is THIOSULFATE (three-letter code: THJ) (formula:  $\mathrm{O_{3}S_{2}}).$ 



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total 6	O S 3 2	0	0

 $\bullet$  Molecule 4 is water.



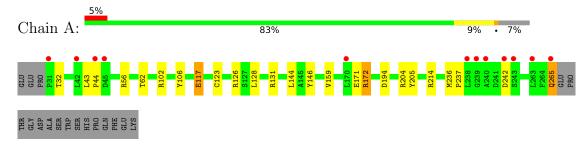
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	201	Total O 201 201	0	0



# 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 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: THIOSULFATE DEHYDROGENASE





# 4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants	79.28Å 70.42Å 57.91Å	Donositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $129.34^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	44.52 - 1.29	Depositor
Resolution (A)	44.52 - 1.29	EDS
% Data completeness	96.3 (44.52-1.29)	Depositor
(in resolution range)	96.3 (44.52-1.29)	EDS
$R_{merge}$	0.03	Depositor
$R_{sym}$	(Not available)	Depositor
$< I/\sigma(I) > 1$	3.05 (at 1.29Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
D D.	0.143 , 0.155	Depositor
$R, R_{free}$	0.162 , $0.170$	DCC
$R_{free}$ test set	3003  reflections  (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	14.2	Xtriage
Anisotropy	0.996	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.36, 44.7	EDS
L-test for twinning <sup>2</sup>	$ < L > = 0.48, < L^2> = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	2080	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 7.78% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;sup>2</sup>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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 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: CSS, HEC, THJ

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		Boı	nd lengths	Bond angles		
IVIOI	Chain	RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	A	1.25	2/1834 (0.1%)	1.26	14/2493 (0.6%)	

#### All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(A)	$Ideal(\AA)$
1	A	205	TYR	CE2-CZ	-6.38	1.30	1.38
1	A	146	TYR	CE2-CZ	-6.24	1.30	1.38

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
1	A	56	ARG	NE-CZ-NH2	-10.12	115.24	120.30
1	A	102	ARG	NE-CZ-NH2	-8.32	116.14	120.30
1	A	102	ARG	NE-CZ-NH1	8.04	124.32	120.30
1	A	128	LEU	CB-CG-CD1	7.76	124.19	111.00
1	A	242	ASP	CB-CG-OD1	-7.19	111.83	118.30
1	A	117	GLU	OE1-CD-OE2	-6.79	115.15	123.30
1	A	214	ARG	NE-CZ-NH1	6.69	123.64	120.30
1	A	204	ARG	NE-CZ-NH1	-6.49	117.05	120.30
1	A	126	ARG	NE-CZ-NH2	-6.27	117.17	120.30
1	A	194	ASP	CB-CG-OD1	6.15	123.84	118.30
1	A	126	ARG	NE-CZ-NH1	5.47	123.04	120.30
1	A	106	TYR	CB-CG-CD2	-5.41	117.75	121.00
1	A	171	GLU	N-CA-C	-5.29	96.71	111.00
1	A	131	ARG	NE-CZ-NH1	5.08	122.84	120.30

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	1784	0	1737	9	0
2	A	90	0	38	3	0
3	A	5	0	0	1	0
4	A	201	0	0	4	3
All	All	2080	0	1775	11	3

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	$\begin{array}{c} {\rm Interatomic} \\ {\rm distance} \ ({\rm \AA}) \end{array}$	Clash overlap (Å)
2:A:601[B]:HEC:O1D	4:A:2110:HOH:O	1.92	0.85
1:A:172:ARG:HG3	4:A:2147:HOH:O	1.92	0.69
1:A:32:THR:HA	1:A:159:VAL:O	1.98	0.63
2:A:601[B]:HEC:O1D	3:A:700:THJ:S2	2.60	0.59
1:A:117:GLU:HG3	1:A:144:LEU:HD11	1.84	0.58
1:A:265:GLN:NE2	4:A:2201:HOH:O	2.44	0.49
1:A:237:PRO:HB3	2:A:600:HEC:HMC2	2.01	0.43
1:A:236:MET:HA	1:A:237:PRO:C	2.40	0.42
1:A:265:GLN:HE21	1:A:265:GLN:HB3	1.61	0.42
1:A:172:ARG:NH1	4:A:2147:HOH:O	2.29	0.42
1:A:43:LEU:HD23	1:A:44:PRO:HD2	2.03	0.40

All (3) 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	$\begin{array}{c} {\rm Interatomic} \\ {\rm distance} \ ({\rm \AA}) \end{array}$	Clash overlap (Å)
4:A:2160:HOH:O	4:A:2160:HOH:O[2_455]	1.59	0.61
4:A:2106:HOH:O	4:A:2106:HOH:O[2_555]	1.67	0.53
4:A:2105:HOH:O	4:A:2106:HOH:O[2_555]	2.09	0.11



#### 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.

Mo	l Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	237/253 (94%)	231 (98%)	5 (2%)	1 (0%)	34 10	

#### All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	172	ARG

#### 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	181/192 (94%)	179 (99%)	2 (1%)	73 45	

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

Mol	Chain	Res	Type
1	A	62	THR
1	A	265	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	265	GLN



#### 5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	oe Chain	Res	Link	Bond lengths			Bond angles		
MIOI	Туре		nes	LIIIK	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
1	CSS	A	123	2,1	4,6,7	1.32	1 (25%)	1,6,8	0.06	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
1	CSS	A	123	2,1	-	0/1/5/7	-

#### All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\operatorname{Observed}(\text{\AA})$	$\operatorname{Ideal}( ext{\AA})$
1	A	123	CSS	O-C	2.08	1.28	1.19

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates (i)

There are no monosaccharides in this entry.



## 5.6 Ligand geometry (i)

4 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	Во	ond leng	ths	Bond angles		
IVIOI	Type	Chain	nes	LIIIK	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	HEC	A	601[B]	-	32,50,50	1.27	5 (15%)	24,82,82	2.05	10 (41%)
2	HEC	A	601[A]	-	32,50,50	1.25	4 (12%)	24,82,82	2.14	11 (45%)
3	THJ	A	700	-	2,4,4	1.86	1 (50%)	2,6,6	0.52	0
2	HEC	A	600	1	32,50,50	1.52	6 (18%)	24,82,82	2.59	11 (45%)

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	HEC	A	601[B]	-	-	2/10/54/54	-
2	HEC	A	601[A]	-	-	2/10/54/54	-
2	HEC	A	600	1	-	2/10/54/54	-

All (16) bond length outliers are listed below:

Mol	Chain	Res	Type	pe Atoms		Observed(A)	Ideal(A)
2	A	600	HEC	C3C-C2C	-3.90	1.36	1.40
2	A	601[A]	HEC	C2B-C3B	-3.37	1.37	1.40
2	A	601[B]	HEC	C2B-C3B	-3.37	1.37	1.40
2	A	600	HEC	CBC-CAC	-3.20	1.37	1.49
2	A	601[A]	HEC	C3C-C2C	-2.54	1.38	1.40
2	A	601[B]	HEC	C3C-C2C	-2.54	1.38	1.40
3	A	700	THJ	O2-S1	2.50	1.53	1.45
2	A	600	HEC	O1A-CGA	2.49	1.30	1.22
2	A	600	HEC	O2A-CGA	-2.46	1.22	1.30
2	A	600	HEC	CBD-CGD	2.44	1.56	1.50
2	A	601[A]	HEC	O1A-CGA	2.30	1.29	1.22
2	A	601[B]	HEC	O1A-CGA	2.30	1.29	1.22

Continued on next page...



Continued from previous page...

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	Observed(A)	$oxed{Ideal(A)}$
2	A	601[A]	HEC	O2A-CGA	-2.23	1.23	1.30
2	A	601[B]	HEC	O2A-CGA	-2.23	1.23	1.30
2	A	601[B]	HEC	O2D-CGD	-2.10	1.23	1.30
2	A	600	HEC	C4B-C3B	2.09	1.46	1.43

All (32) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$Observed(^o)$	$\operatorname{Ideal}(^{o})$
2	A	600	HEC	O2D-CGD-O1D	7.19	141.21	123.30
2	A	600	HEC	O1D-CGD-CBD	-4.21	109.55	123.08
2	A	601[A]	HEC	CMB-C2B-C3B	3.69	130.16	125.82
2	A	601[B]	HEC	CMB-C2B-C3B	3.69	130.16	125.82
2	A	600	HEC	C1D-C2D-C3D	-3.64	104.46	107.00
2	A	600	HEC	CMA-C3A-C2A	3.48	131.51	124.94
2	A	600	HEC	CMD-C2D-C3D	3.47	131.49	124.94
2	A	601[A]	HEC	CMC-C2C-C3C	3.41	129.83	125.82
2	A	601[B]	HEC	CMC-C2C-C3C	3.41	129.83	125.82
2	A	601[A]	HEC	CMB-C2B-C1B	-3.41	123.22	128.46
2	A	601[B]	HEC	CMB-C2B-C1B	-3.41	123.22	128.46
2	A	601[A]	HEC	O2A-CGA-O1A	-3.16	115.43	123.30
2	A	601[B]	HEC	O2A-CGA-O1A	-3.16	115.43	123.30
2	A	601[A]	HEC	O2A-CGA-CBA	3.08	123.92	114.03
2	A	601[B]	HEC	O2A-CGA-CBA	3.08	123.92	114.03
2	A	600	HEC	CMD-C2D-C1D	-2.89	124.03	128.46
2	A	601[A]	HEC	CMC-C2C-C1C	-2.82	124.13	128.46
2	A	601[B]	HEC	CMC-C2C-C1C	-2.82	124.13	128.46
2	A	601[A]	HEC	CBD-CAD-C3D	2.81	117.41	112.62
2	A	600	HEC	CMC-C2C-C3C	2.79	129.10	125.82
2	A	600	HEC	O2A-CGA-CBA	2.52	122.11	114.03
2	A	601[A]	HEC	CMD-C2D-C1D	-2.49	124.64	128.46
2	A	601[B]	HEC	CMD-C2D-C1D	-2.49	124.64	128.46
2	A	601[A]	HEC	CMA-C3A-C2A	2.35	129.37	124.94
2	A	601[B]	HEC	CMA-C3A-C2A	2.35	129.37	124.94
2	A	600	HEC	O1A-CGA-CBA	-2.31	115.65	123.08
2	A	600	HEC	O2D-CGD-CBD	-2.23	106.86	114.03
2	A	601[B]	HEC	CBD-CAD-C3D	-2.22	108.83	112.62
2	A	601[A]	HEC	CBA-CAA-C2A	-2.10	109.07	112.60
2	A	601[B]	HEC	CBA-CAA-C2A	-2.10	109.07	112.60
2	A	600	HEC	CMC-C2C-C1C	-2.09	125.26	128.46
2	A	601[A]	HEC	O2D-CGD-O1D	-2.08	118.11	123.30

There are no chirality outliers.



All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	601[A]	HEC	CAD-CBD-CGD-O2D
2	A	601[B]	HEC	CAD-CBD-CGD-O1D
2	A	601[B]	HEC	CAD-CBD-CGD-O2D
2	A	601[A]	HEC	CAD-CBD-CGD-O1D
2	A	600	HEC	CAD-CBD-CGD-O1D
2	A	600	HEC	CAD-CBD-CGD-O2D

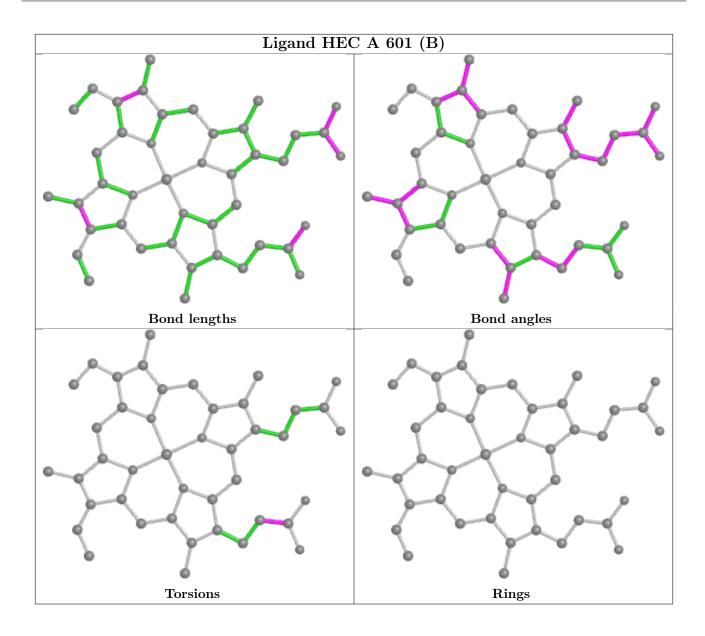
There are no ring outliers.

3 monomers are involved in 3 short contacts:

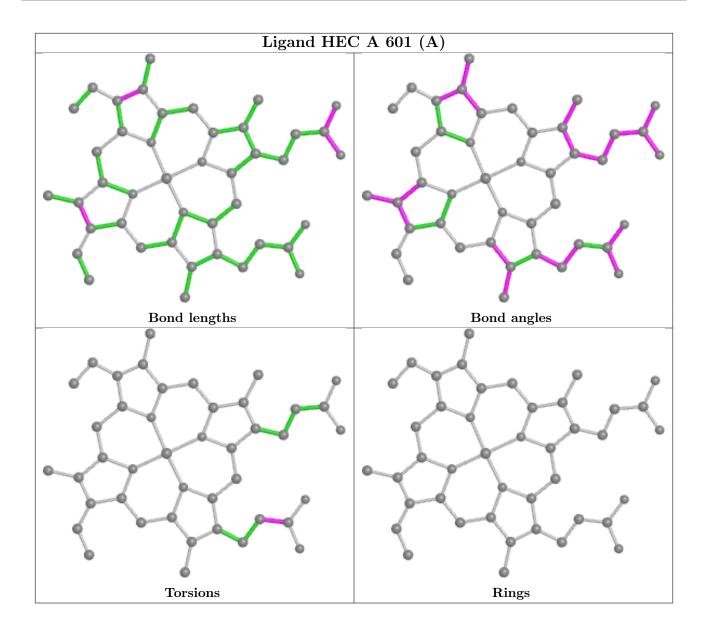
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	601[B]	HEC	2	0
3	A	700	THJ	1	0
2	A	600	HEC	1	0

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 then 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.

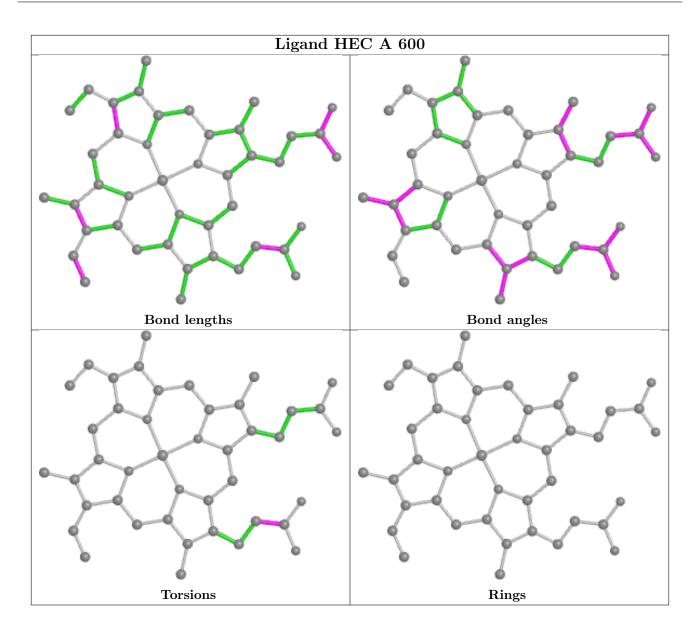












## 5.7 Other polymers (i)

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 (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,  $95^{th}$  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 $>$	$\#\mathrm{RSRZ}{>}2$		$OWAB(A^2)$	Q < 0.9	
1	A	234/253 (92%)	0.24	12 (5%)	28	25	13, 20, 43, 80	0

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	265	GLN	5.1
1	A	42	LEU	4.8
1	A	242	ASP	3.3
1	A	238	LEU	3.1
1	A	31	PRO	3.0
1	A	263	LEU	2.9
1	A	45	ASP	2.8
1	A	170	LEU	2.7
1	A	239	GLY	2.6
1	A	240	ALA	2.6
1	A	44	PRO	2.4
1	A	243	SER	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains (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,  $95^{th}$  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	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q<0.9
1	CSS	A	123	7/8	0.97	0.07	15,15,17,22	0



## 6.3 Carbohydrates (i)

There are no monosaccharides 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,  $95^{th}$  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.

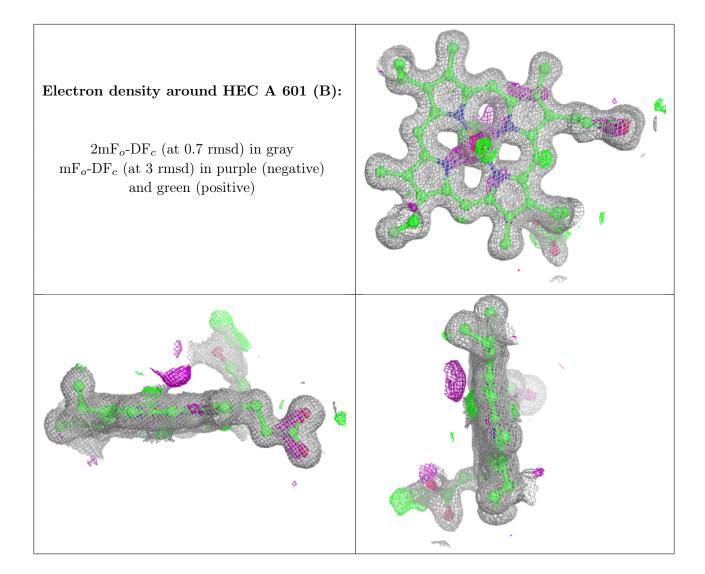
Mol	Type	Chain	Res	Atoms	RSCC	RSR	$oxed{f B-factors({ m \AA}^2)}$	Q<0.9
3	THJ	A	700	5/5	0.93	0.11	11,13,15,17	5
2	HEC	A	600	43/43	0.98	0.07	8,9,12,21	0
2	HEC	A	601[B]	43/43	0.99	0.06	9,10,15,24	4
2	HEC	A	601[A]	43/43	0.99	0.06	9,10,13,22	4

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

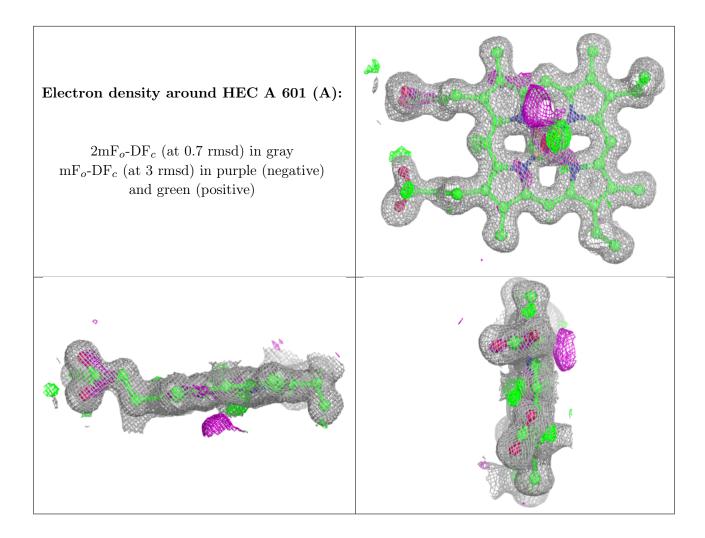


# Electron density around HEC A 600: $2 {\rm mF}_o\text{-}{\rm DF}_c$ (at 0.7 rmsd) in gray ${\rm mF}_o\text{-}{\rm DF}_c$ (at 3 rmsd) in purple (negative) and green (positive)









# 6.5 Other polymers (i)

There are no such residues in this entry.

