



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 24, 2024 – 09:13 AM EDT

PDB ID : 3EPM  
Title : Crystal structure of Caulobacter crescentus ThiC  
Authors : Li, S.; Chatterjee, A.; Zhang, Y.; Grove, T.L.; Lee, M.; Krebs, C.; Booker, S.J.; Begley, T.P.; Ealick, S.E.  
Deposited on : 2008-09-29  
Resolution : 2.79 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 : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.20.1  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.003 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

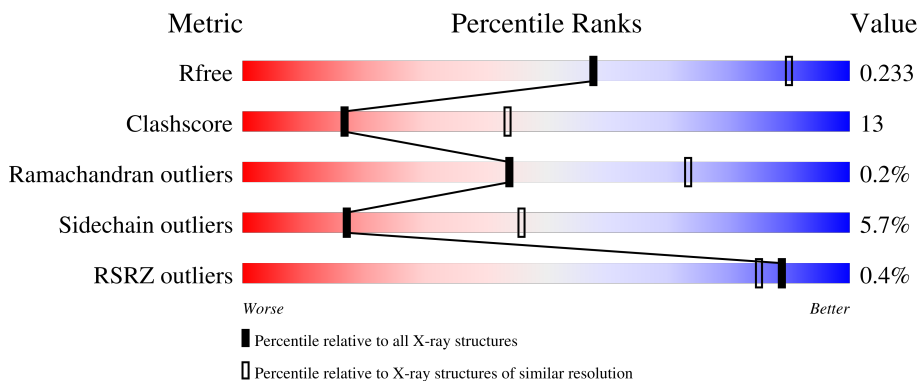
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.79 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))
$R_{free}$	164625	3657 (2.80-2.80)
Clashscore	180529	4123 (2.80-2.80)
Ramachandran outliers	177936	4071 (2.80-2.80)
Sidechain outliers	177891	4073 (2.80-2.80)
RSRZ outliers	164620	3659 (2.80-2.80)

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  $\leq 5\%$ . 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	612	 64% 18% • 16%
1	B	612	 58% 24% • 16%

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 8210 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.

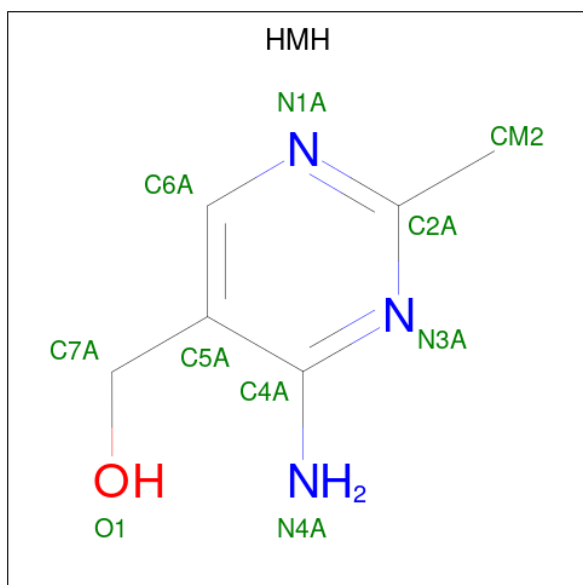
- Molecule 1 is a protein called Thiamine biosynthesis protein thiC.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	515	Total 4042	C 2558	N 715	O 751	S 6	Se 12	0	0	0
1	B	512	Total 4018	C 2543	N 711	O 746	S 6	Se 12	0	0	0

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

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

- Molecule 3 is 4-AMINO-5-HYDROXYMETHYL-2-METHYLPYRIMIDINE (three-letter code: HMH) (formula: C<sub>6</sub>H<sub>9</sub>N<sub>3</sub>O).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	10	6	3	1	0	0

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



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

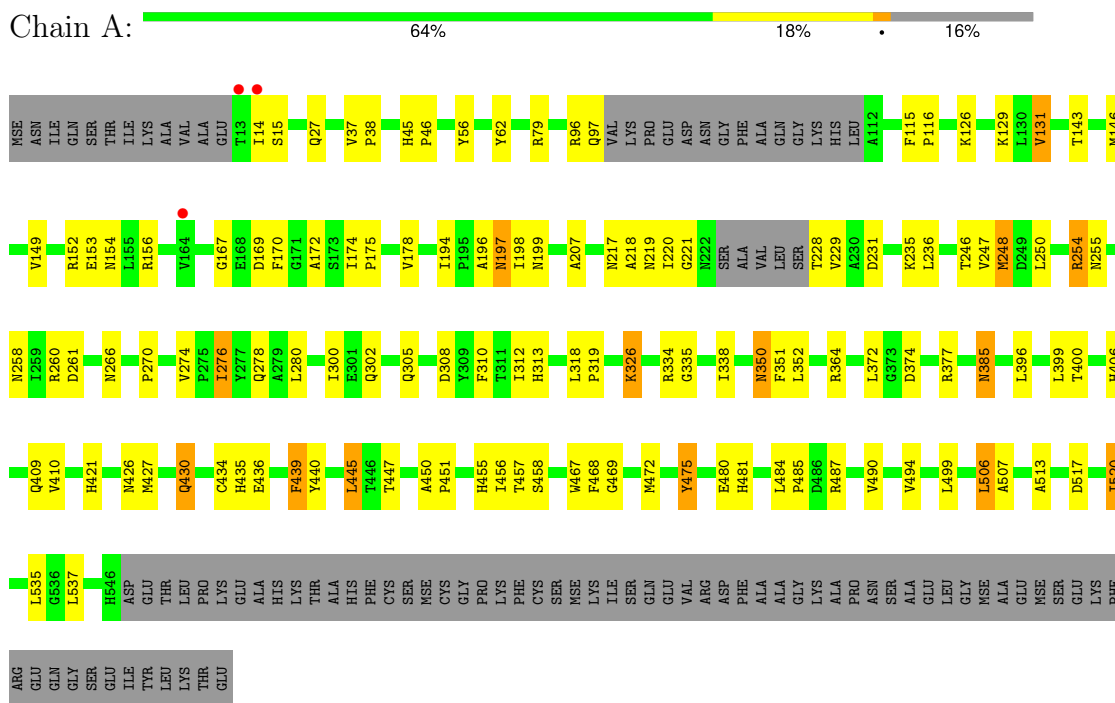
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	91	91	91	0	0
5	B	42	42	42	0	0

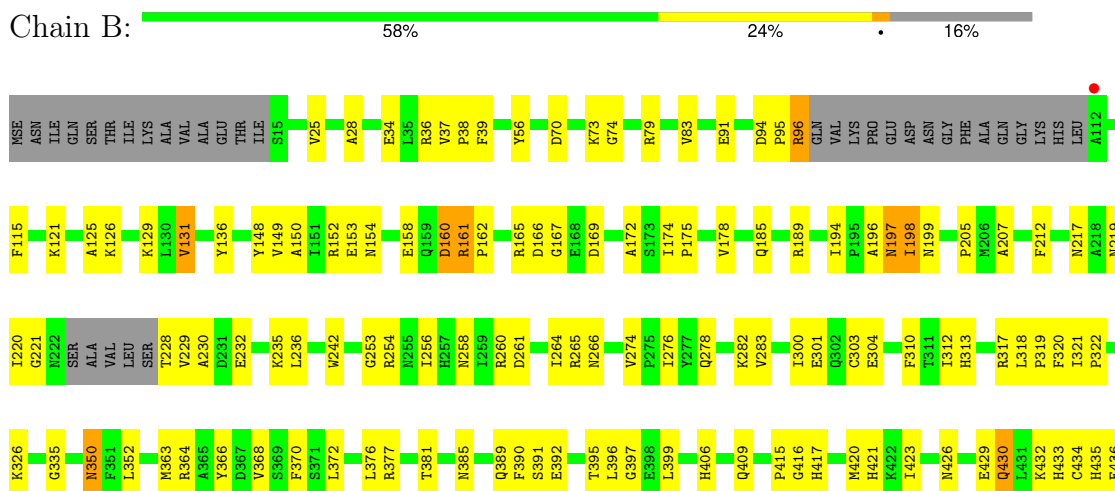
### 3 Residue-property plots

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 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: Thiamine biosynthesis protein thiC



- Molecule 1: Thiamine biosynthesis protein thiC



PHE	F439	ASP	H546
CYS	Y440	GLU	GLU
SER	T441	THR	THR
NSE	CYS	LEU	LEU
CYS	P444	PRO	PRO
PRO	L446	LYS	LYS
LYS	D448	GLU	GLU
PHE	CYS	ALA	ALA
CYS	T457	HIS	HIS
SER	S458	LYS	LYS
NSE	M459	THR	THR
LYS	I460	ALA	ALA
ILE	G461	LYS	LYS
SER	GLN	THR	THR
GLN	M464	ALA	ALA
GLU	I465	ALA	ALA
VAL	ARG	ALA	ALA
ARG	T470	GLU	GLU
ASP	CYS	LEU	LEU
ASP	PHE	GLY	GLY
PHE	ALA	NSE	NSE
ALA	Y475	ALA	ALA
ALA	V476	GLU	GLU
GLY	T477	GLU	GLU
LYS	P478	NSE	NSE
ALA	K479	SER	SER
PRO	E480	GLU	GLU
ASN	H481	LYS	LYS
SER	ALA	ARG	ARG
ALA	L484	GLU	GLU
GLU	GLU	GLN	GLN
LEU	R487	GLY	GLY
GLY	W490	SER	SER
NSE	K491	GLU	GLU
ALA	GLU	LYS	LYS
GLU	V494	PHE	PHE
NSE	L499	ARG	ARG
SER	L506	GLU	GLU
GLU	D517	GLN	GLN
LYS	I520	GLY	GLY
PHE	THR	SER	SER
ARG	THR	GLU	GLU
GLU	THR	ILE	ILE
GLU	THR	TYR	TYR
GLU	THR	LEU	LEU
GLU	THR	LYS	LYS
GLU	THR	THR	THR
GLU	THR	ALA	ALA
GLU	THR	ALA	ALA
GLU	THR	HIS	HIS

## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	63.25Å 103.44Å 95.37Å 90.00° 91.62° 90.00°	Depositor
Resolution (Å)	43.29 – 2.79 43.29 – 2.79	Depositor EDS
% Data completeness (in resolution range)	99.5 (43.29-2.79) 99.5 (43.29-2.79)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	0.14	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.57 (at 2.81Å)	Xtrriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.181 , 0.244 0.173 , 0.233	Depositor DCC
$R_{free}$ test set	1538 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	44.4	Xtrriage
Anisotropy	0.568	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 36.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.032 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	8210	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.67% 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: SO4, HMH, 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	0.35	0/4128	0.53	0/5587
1	B	0.34	0/4104	0.52	0/5554
All	All	0.34	0/8232	0.52	0/11141

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	4042	0	3962	94	0
1	B	4018	0	3936	120	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	10	0	9	1	0
4	A	5	0	0	0	0
5	A	91	0	0	2	0
5	B	42	0	0	2	0
All	All	8210	0	7907	202	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 13.



All (202) 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:372:LEU:HB3	1:A:396:LEU:HD22	1.52	0.91
1:A:197:ASN:HD22	1:A:199:ASN:H	1.16	0.91
1:B:197:ASN:ND2	1:B:199:ASN:H	1.70	0.89
1:A:318:LEU:HB3	1:A:319:PRO:HD3	1.54	0.88
1:A:457:THR:HG22	1:B:499:LEU:HD13	1.58	0.84
1:B:197:ASN:HD22	1:B:199:ASN:H	1.22	0.84
1:A:197:ASN:ND2	1:A:199:ASN:H	1.76	0.83
1:A:154:ASN:HD21	1:A:178:VAL:H	1.25	0.82
1:B:350:ASN:HD21	1:B:352:LEU:HB2	1.43	0.81
1:B:313:HIS:CE1	1:B:335:GLY:HA3	2.17	0.80
1:A:194:ILE:H	1:A:409:GLN:HE22	1.29	0.78
1:B:154:ASN:HD21	1:B:178:VAL:H	1.32	0.78
1:A:350:ASN:ND2	1:A:352:LEU:H	1.83	0.77
1:A:350:ASN:C	1:A:350:ASN:HD22	1.89	0.75
1:A:499:LEU:HD13	1:B:457:THR:HG22	1.71	0.72
1:B:219:ASN:O	1:B:220:ILE:HD13	1.89	0.71
1:B:465:ILE:HG12	1:B:470:THR:OG1	1.92	0.69
1:B:350:ASN:ND2	1:B:352:LEU:H	1.92	0.68
1:A:149:VAL:HG11	1:A:207:ALA:HB2	1.76	0.67
1:B:480:GLU:O	1:B:481:HIS:HB2	1.94	0.67
1:B:199:ASN:ND2	1:B:304:GLU:HA	2.10	0.66
1:B:158:GLU:HG3	1:B:160:ASP:HB2	1.78	0.66
1:B:197:ASN:HD22	1:B:199:ASN:N	1.92	0.66
1:A:480:GLU:O	1:A:481:HIS:HB2	1.97	0.65
1:A:430:GLN:HG3	1:A:439:PHE:CE1	2.31	0.65
1:B:153:GLU:HB3	5:B:631:HOH:O	1.96	0.64
1:B:199:ASN:HD22	1:B:304:GLU:HA	1.63	0.64
1:B:25:VAL:HG23	1:B:39:PHE:HB2	1.80	0.63
1:A:313:HIS:CE1	1:A:335:GLY:HA3	2.34	0.63
1:B:242:TRP:CZ2	1:B:491:LYS:HE3	2.34	0.62
1:B:350:ASN:ND2	1:B:352:LEU:HB2	2.15	0.61
1:A:219:ASN:O	1:A:220:ILE:HD13	2.02	0.60
1:A:350:ASN:HD22	1:A:351:PHE:N	1.98	0.60
1:A:457:THR:HG22	1:B:499:LEU:CD1	2.31	0.60
1:B:318:LEU:HB3	1:B:319:PRO:HD3	1.85	0.59
1:B:131:VAL:O	1:B:136:TYR:HE2	1.86	0.59
1:B:96:ARG:HH21	1:B:300:ILE:HG21	1.68	0.58
1:B:430:GLN:O	1:B:434:CYS:HB2	2.02	0.58
1:B:445:LEU:HD22	1:B:458:SER:HB3	1.85	0.58
1:A:236:LEU:HD12	1:A:247:VAL:HG11	1.85	0.58

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:260:ARG:HD2	1:A:305:GLN:OE1	2.04	0.58
1:B:465:ILE:HG12	1:B:470:THR:HG1	1.68	0.57
1:B:283:VAL:O	1:B:283:VAL:HG23	2.02	0.57
1:A:280:LEU:HD22	1:A:338:ILE:HD13	1.86	0.57
1:B:149:VAL:HG11	1:B:207:ALA:HB2	1.86	0.57
1:A:197:ASN:HD22	1:A:199:ASN:N	1.97	0.57
1:B:160:ASP:HB3	1:B:161:ARG:HD2	1.87	0.57
1:A:350:ASN:HD22	1:A:352:LEU:H	1.49	0.57
1:B:160:ASP:O	1:B:162:PRO:HD3	2.05	0.57
1:A:27:GLN:OE1	1:A:156:ARG:HG2	2.05	0.56
1:A:302:GLN:HE21	1:A:310:PHE:HE1	1.54	0.56
1:A:400:THR:HG23	1:A:410:VAL:HG11	1.87	0.56
1:B:312:ILE:HG23	1:B:352:LEU:HD13	1.88	0.56
1:A:334:ARG:O	1:A:338:ILE:HG13	2.05	0.56
1:B:490:VAL:O	1:B:494:VAL:HG23	2.06	0.55
1:B:372:LEU:HB3	1:B:396:LEU:HD22	1.89	0.55
1:B:28:ALA:HB1	1:B:36:ARG:HH21	1.71	0.55
1:B:392:GLU:O	1:B:396:LEU:HG	2.07	0.55
1:B:161:ARG:HD2	1:B:161:ARG:N	2.21	0.55
1:A:246:THR:HG22	1:A:270:PRO:HD2	1.88	0.55
1:B:70:ASP:OD2	1:B:73:LYS:HG2	2.06	0.55
1:A:350:ASN:ND2	1:A:350:ASN:C	2.60	0.55
1:B:432:LYS:HD3	1:B:433:HIS:CE1	2.43	0.54
1:B:310:PHE:CD2	1:B:368:VAL:HG11	2.42	0.54
1:B:499:LEU:HD12	1:B:499:LEU:C	2.28	0.54
1:B:153:GLU:OE2	1:B:212:PHE:HB3	2.08	0.53
1:B:220:ILE:HG23	1:B:235:LYS:NZ	2.23	0.53
1:B:350:ASN:HD22	1:B:352:LEU:H	1.55	0.53
1:A:430:GLN:HG3	1:A:439:PHE:CD1	2.43	0.53
1:A:79:ARG:NH1	1:A:196:ALA:HB1	2.22	0.53
1:A:318:LEU:HB3	1:A:319:PRO:CD	2.33	0.53
1:B:350:ASN:HD22	1:B:352:LEU:N	2.07	0.53
1:A:455:HIS:CE1	1:A:456:ILE:HG13	2.44	0.53
1:B:385:ASN:H	1:B:389:GLN:HE22	1.55	0.53
1:B:460:ILE:O	1:B:464:MSE:HG3	2.08	0.53
1:A:364:ARG:HG3	5:A:2065:HOH:O	2.09	0.52
1:A:153:GLU:HB3	5:A:2060:HOH:O	2.09	0.52
1:A:167:GLY:HA2	1:B:421:HIS:CE1	2.44	0.52
1:A:467:TRP:HB2	1:B:420:MSE:HE1	1.92	0.52
1:B:444:PRO:HD2	1:B:461:GLY:C	2.29	0.52
1:B:242:TRP:CE2	1:B:491:LYS:HE3	2.44	0.51

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:421:HIS:CE1	1:B:167:GLY:HA2	2.46	0.51
1:A:302:GLN:NE2	1:A:310:PHE:CE1	2.77	0.51
1:B:115:PHE:CE1	1:B:265:ARG:HG3	2.45	0.51
1:A:254:ARG:O	1:A:255:ASN:HB2	2.10	0.51
1:A:334:ARG:H	3:A:1001:HMH:C6A	2.24	0.50
1:B:321:ILE:HB	1:B:322:PRO:HD3	1.94	0.50
1:A:467:TRP:HB2	1:B:420:MSE:CE	2.41	0.50
1:B:194:ILE:H	1:B:409:GLN:NE2	2.10	0.50
1:B:148:TYR:CE2	1:B:205:PRO:HD2	2.47	0.49
1:A:440:TYR:CD1	1:A:472:MSE:HG2	2.47	0.49
1:A:197:ASN:HD22	1:A:197:ASN:C	2.16	0.49
1:B:420:MSE:HA	1:B:423:ILE:HD12	1.94	0.49
1:A:56:TYR:CE1	1:A:152:ARG:HG2	2.47	0.49
1:B:197:ASN:ND2	1:B:199:ASN:N	2.51	0.49
1:A:96:ARG:HE	1:A:300:ILE:HG21	1.78	0.49
1:A:507:ALA:HA	1:B:421:HIS:CD2	2.48	0.49
1:B:350:ASN:HD21	1:B:352:LEU:CB	2.21	0.49
1:A:364:ARG:HD2	1:A:406:HIS:O	2.13	0.49
1:B:91:GLU:OE2	1:B:121:LYS:HD3	2.13	0.49
1:B:350:ASN:ND2	1:B:352:LEU:N	2.59	0.49
1:B:420:MSE:HG3	5:B:617:HOH:O	2.12	0.48
1:A:520:ILE:HG12	1:A:535:LEU:CB	2.43	0.48
1:A:435:HIS:O	1:A:436:GLU:HB2	2.13	0.48
1:B:310:PHE:CE2	1:B:368:VAL:HG11	2.49	0.48
1:B:282:LYS:NZ	1:B:301:GLU:OE1	2.45	0.48
1:B:260:ARG:O	1:B:264:ILE:HG13	2.14	0.47
1:A:520:ILE:HG12	1:A:535:LEU:HB3	1.96	0.47
1:B:198:ILE:HG23	1:B:366:TYR:O	2.13	0.47
1:A:45:HIS:CD2	1:A:46:PRO:HD2	2.50	0.47
1:A:220:ILE:HG23	1:A:235:LYS:NZ	2.30	0.47
1:A:385:ASN:HD22	1:A:385:ASN:N	2.14	0.46
1:B:395:THR:O	1:B:399:LEU:HG	2.15	0.46
1:A:219:ASN:HB3	1:A:250:LEU:HD12	1.97	0.46
1:A:260:ARG:HH22	1:A:278:GLN:NE2	2.12	0.46
1:A:274:VAL:HG12	1:A:276:ILE:HG22	1.97	0.46
1:A:475:TYR:CD2	1:A:475:TYR:C	2.89	0.46
1:B:435:HIS:O	1:B:436:GLU:HB2	2.14	0.46
1:B:194:ILE:H	1:B:409:GLN:HE22	1.62	0.46
1:B:441:THR:HG21	1:B:465:ILE:HD13	1.97	0.46
1:B:479:LYS:HD2	1:B:484:LEU:O	2.16	0.46
1:A:126:LYS:HB2	1:A:129:LYS:HG3	1.96	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:248:MSE:HE2	1:A:248:MSE:HB3	1.87	0.45
1:A:231:ASP:O	1:A:235:LYS:HG3	2.16	0.45
1:B:197:ASN:HD22	1:B:197:ASN:C	2.20	0.45
1:A:326:LYS:HD2	1:A:326:LYS:H	1.82	0.45
1:A:445:LEU:HD22	1:A:458:SER:HB3	1.98	0.45
1:B:320:PHE:CE1	1:B:391:SER:HB3	2.52	0.45
1:A:350:ASN:HD21	1:A:352:LEU:HB2	1.82	0.45
1:B:149:VAL:CG1	1:B:207:ALA:HB2	2.46	0.45
1:B:417:HIS:ND1	1:B:448:ASP:OD2	2.39	0.45
1:A:197:ASN:HD22	1:A:198:ILE:N	2.16	0.44
1:A:484:LEU:HA	1:A:485:PRO:HD3	1.76	0.44
1:B:363:MSE:HE1	1:B:370:PHE:CD1	2.52	0.44
1:B:364:ARG:HD2	1:B:406:HIS:O	2.17	0.44
1:A:62:TYR:OH	1:A:266:ASN:HB3	2.17	0.44
1:B:169:ASP:HB2	1:B:172:ALA:HB3	1.99	0.44
1:B:174:ILE:HA	1:B:175:PRO:HD3	1.84	0.44
1:B:256:ILE:HD11	1:B:278:GLN:HA	1.98	0.44
1:A:276:ILE:HD12	1:A:276:ILE:HA	1.76	0.44
1:A:490:VAL:O	1:A:494:VAL:HG23	2.18	0.44
1:B:165:ARG:NH2	1:B:174:ILE:O	2.50	0.44
1:B:350:ASN:HD22	1:B:350:ASN:C	2.21	0.44
1:A:535:LEU:HD23	1:A:535:LEU:HA	1.82	0.44
1:B:303:CYS:SG	1:B:368:VAL:HG21	2.58	0.44
1:B:232:GLU:OE2	1:B:235:LYS:HE3	2.18	0.43
1:B:258:ASN:O	1:B:261:ASP:HB2	2.18	0.43
1:A:427:MSE:HG2	1:A:468:PHE:HB3	2.00	0.43
1:B:220:ILE:HG23	1:B:235:LYS:HZ1	1.82	0.43
1:B:376:LEU:HD12	1:B:392:GLU:HB2	1.99	0.43
1:A:506:LEU:HD23	1:A:506:LEU:HA	1.75	0.43
1:B:34:GLU:CD	1:B:34:GLU:H	2.21	0.43
1:A:37:VAL:HA	1:A:38:PRO:HD3	1.75	0.43
1:A:450:ALA:N	1:A:451:PRO:HD3	2.34	0.43
1:A:507:ALA:HA	1:B:421:HIS:HD2	1.84	0.43
1:B:242:TRP:CD1	1:B:491:LYS:HG3	2.54	0.43
1:B:390:PHE:CZ	1:B:429:GLU:HG3	2.53	0.43
1:A:169:ASP:HB2	1:A:172:ALA:HB3	2.00	0.42
1:A:174:ILE:HA	1:A:175:PRO:HD3	1.77	0.42
1:A:312:ILE:HG23	1:A:352:LEU:HD13	2.01	0.42
1:B:185:GLN:O	1:B:189:ARG:HG2	2.19	0.42
1:B:475:TYR:CD2	1:B:475:TYR:C	2.92	0.42
1:B:74:GLY:HA2	1:B:266:ASN:OD1	2.20	0.42

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:276:ILE:HD12	1:B:276:ILE:HA	1.93	0.42
1:A:506:LEU:HD23	1:A:513:ALA:CB	2.50	0.42
1:A:427:MSE:HE2	1:A:469:GLY:HA3	2.01	0.42
1:B:480:GLU:O	1:B:481:HIS:CB	2.66	0.42
1:A:220:ILE:HG22	1:A:221:GLY:N	2.34	0.42
1:A:421:HIS:ND1	1:B:167:GLY:HA2	2.35	0.42
1:A:440:TYR:CE1	1:A:472:MSE:HG2	2.55	0.42
1:B:79:ARG:NH1	1:B:196:ALA:HB1	2.35	0.42
1:B:477:THR:HB	1:B:478:PRO:HD2	2.02	0.42
1:A:167:GLY:HA2	1:B:421:HIS:ND1	2.34	0.41
1:A:421:HIS:HE1	1:B:166:ASP:OD1	2.03	0.41
1:B:125:ALA:HB2	1:B:131:VAL:HG21	2.02	0.41
1:B:229:VAL:HG23	1:B:230:ALA:H	1.86	0.41
1:B:260:ARG:HH12	1:B:278:GLN:HE22	1.68	0.41
1:B:317:ARG:HB3	1:B:319:PRO:HD2	2.02	0.41
1:A:258:ASN:O	1:A:261:ASP:HB2	2.21	0.41
1:A:218:ALA:O	1:A:247:VAL:HA	2.20	0.41
1:A:115:PHE:HA	1:A:116:PRO:HD3	1.79	0.41
1:A:270:PRO:HA	1:A:308:ASP:OD2	2.20	0.41
1:B:37:VAL:HA	1:B:38:PRO:HD3	1.80	0.41
1:B:149:VAL:HG12	1:B:150:ALA:N	2.36	0.41
1:A:374:ASP:OD1	1:A:374:ASP:N	2.53	0.41
1:B:126:LYS:HB2	1:B:129:LYS:HG3	2.01	0.41
1:A:399:LEU:HD23	1:A:399:LEU:HA	1.91	0.41
1:B:415:PRO:HA	1:B:416:GLY:HA3	1.91	0.41
1:A:170:PHE:HE2	1:A:467:TRP:CE2	2.39	0.41
1:B:79:ARG:O	1:B:83:VAL:HG23	2.21	0.41
1:B:131:VAL:O	1:B:136:TYR:CE2	2.69	0.41
1:B:397:GLY:HA3	1:B:433:HIS:O	2.21	0.41
1:B:303:CYS:HB3	1:B:366:TYR:CD2	2.56	0.41
1:B:56:TYR:CE1	1:B:152:ARG:HG2	2.57	0.40
1:B:94:ASP:HA	1:B:95:PRO:HD3	1.79	0.40
1:B:253:GLY:O	1:B:256:ILE:HG23	2.21	0.40
1:B:320:PHE:HB3	1:B:376:LEU:HD11	2.03	0.40
1:A:430:GLN:O	1:A:434:CYS:HB2	2.21	0.40
1:B:220:ILE:HG22	1:B:221:GLY:N	2.36	0.40
1:A:143:THR:OG1	1:A:146:MSE:HE3	2.21	0.40
1:A:537:LEU:HA	1:A:537:LEU:HD23	1.82	0.40
1:B:274:VAL:HG12	1:B:276:ILE:HG22	2.04	0.40

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	509/612 (83%)	483 (95%)	25 (5%)	1 (0%)	44	73
1	B	506/612 (83%)	479 (95%)	26 (5%)	1 (0%)	44	73
All	All	1015/1224 (83%)	962 (95%)	51 (5%)	2 (0%)	44	73

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	160	ASP
1	A	131	VAL

### 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	423/486 (87%)	398 (94%)	25 (6%)	16	44
1	B	420/486 (86%)	397 (94%)	23 (6%)	18	47
All	All	843/972 (87%)	795 (94%)	48 (6%)	17	46

All (48) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	14	ILE
1	A	15	SER
1	A	97	GLN
1	A	131	VAL

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
1	A	197	ASN
1	A	217	ASN
1	A	228	THR
1	A	229	VAL
1	A	248	MSE
1	A	254	ARG
1	A	276	ILE
1	A	326	LYS
1	A	350	ASN
1	A	377	ARG
1	A	385	ASN
1	A	426	ASN
1	A	430	GLN
1	A	439	PHE
1	A	445	LEU
1	A	447	THR
1	A	475	TYR
1	A	487	ARG
1	A	506	LEU
1	A	517	ASP
1	A	520	ILE
1	B	96	ARG
1	B	131	VAL
1	B	161	ARG
1	B	197	ASN
1	B	198	ILE
1	B	217	ASN
1	B	228	THR
1	B	236	LEU
1	B	254	ARG
1	B	326	LYS
1	B	350	ASN
1	B	377	ARG
1	B	381	THR
1	B	426	ASN
1	B	430	GLN
1	B	439	PHE
1	B	445	LEU
1	B	475	TYR
1	B	487	ARG
1	B	499	LEU
1	B	506	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	517	ASP
1	B	520	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (34) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	45	HIS
1	A	97	GLN
1	A	154	ASN
1	A	159	GLN
1	A	197	ASN
1	A	217	ASN
1	A	255	ASN
1	A	257	HIS
1	A	278	GLN
1	A	302	GLN
1	A	350	ASN
1	A	385	ASN
1	A	389	GLN
1	A	409	GLN
1	A	421	HIS
1	A	430	GLN
1	A	532	GLN
1	A	546	HIS
1	B	45	HIS
1	B	154	ASN
1	B	159	GLN
1	B	197	ASN
1	B	255	ASN
1	B	278	GLN
1	B	302	GLN
1	B	313	HIS
1	B	350	ASN
1	B	385	ASN
1	B	389	GLN
1	B	409	GLN
1	B	421	HIS
1	B	430	GLN
1	B	433	HIS
1	B	546	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 oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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	2001	-	4,4,4	0.26	0	6,6,6	0.09	0
3	HMH	A	1001	-	10,10,10	1.11	1 (10%)	12,13,13	2.21	7 (58%)

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	HMH	A	1001	-	-	1/2/2/2	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1001	HMH	C5A-C4A	-2.26	1.39	1.42

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1001	HMH	N1A-C2A-N3A	-3.54	119.63	125.53
3	A	1001	HMH	C6A-N1A-C2A	3.43	121.70	116.07
3	A	1001	HMH	N4A-C4A-N3A	2.78	120.77	117.03
3	A	1001	HMH	CM2-C2A-N3A	2.40	120.72	117.13
3	A	1001	HMH	C5A-C4A-N4A	-2.39	118.92	122.14
3	A	1001	HMH	CM2-C2A-N1A	2.30	119.65	117.20
3	A	1001	HMH	C5A-C6A-N1A	-2.08	120.45	123.83

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1001	HMH	C6A-C5A-C7A-O1

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1001	HMH	1	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	503/612 (82%)	-0.33	3 (0%) 85 81	21, 39, 67, 121	0
1	B	500/612 (81%)	-0.37	1 (0%) 92 89	22, 41, 67, 119	0
All	All	1003/1224 (81%)	-0.35	4 (0%) 89 85	21, 40, 67, 121	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	164	VAL	2.4
1	A	13	THR	2.3
1	B	112	ALA	2.1
1	A	14	ILE	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 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<sup>th</sup> 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(Å <sup>2</sup> )	Q<0.9
3	HMH	A	1001	10/10	0.64	0.21	55,80,88,94	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	SO4	A	2001	5/5	0.85	0.08	59,68,96,111	0
2	ZN	A	613	1/1	0.96	0.04	56,56,56,56	0
2	ZN	B	613	1/1	0.96	0.05	69,69,69,69	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.