



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 16, 2024 – 01:59 AM EDT

PDB ID : 1TT7  
Title : Crystal structure of Bacillus subtilis protein yhfP  
Authors : Min, T.; Gorman, J.; Shapiro, L.; Burley, S.K.; New York SGX Research Center for Structural Genomics (NYSGXRC)  
Deposited on : 2004-06-22  
Resolution : 2.70 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 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	:	2.37.1
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)	:	2.37.1

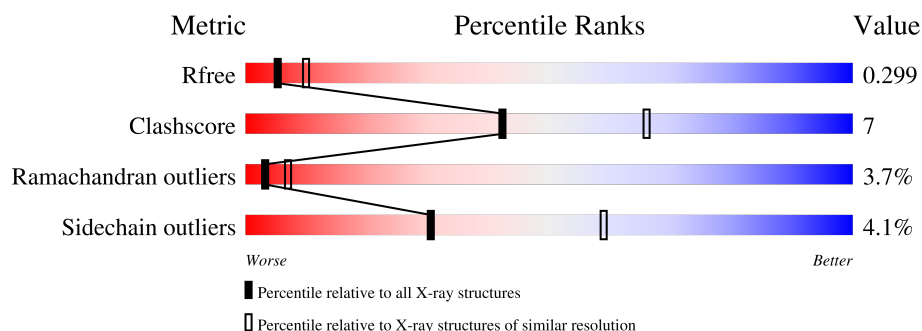
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)

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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	330	79% 18% .
1	B	330	77% 21% .
1	C	330	75% 22% .
1	D	330	78% 18% .
1	E	330	78% 19% .
1	F	330	76% 20% . .

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 14777 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 YHFP.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	329	Total	C	N	O	S	Se	0	0	0
			2436	1533	414	484	1	4			
1	B	329	Total	C	N	O	S	Se	0	0	0
			2436	1533	414	484	1	4			
1	C	329	Total	C	N	O	S	Se	0	0	0
			2436	1533	414	484	1	4			
1	D	329	Total	C	N	O	S	Se	0	0	0
			2436	1533	414	484	1	4			
1	E	329	Total	C	N	O	S	Se	0	0	0
			2436	1533	414	484	1	4			
1	F	329	Total	C	N	O	S	Se	0	0	0
			2437	1533	414	485	1	4			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	126	MSE	MET	modified residue	UNP O07615
A	170	MSE	MET	modified residue	UNP O07615
A	278	MSE	MET	modified residue	UNP O07615
A	288	MSE	MET	modified residue	UNP O07615
B	126	MSE	MET	modified residue	UNP O07615
B	170	MSE	MET	modified residue	UNP O07615
B	278	MSE	MET	modified residue	UNP O07615
B	288	MSE	MET	modified residue	UNP O07615
C	126	MSE	MET	modified residue	UNP O07615
C	170	MSE	MET	modified residue	UNP O07615
C	278	MSE	MET	modified residue	UNP O07615
C	288	MSE	MET	modified residue	UNP O07615
D	126	MSE	MET	modified residue	UNP O07615
D	170	MSE	MET	modified residue	UNP O07615
D	278	MSE	MET	modified residue	UNP O07615
D	288	MSE	MET	modified residue	UNP O07615
E	126	MSE	MET	modified residue	UNP O07615

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	170	MSE	MET	modified residue	UNP O07615
E	278	MSE	MET	modified residue	UNP O07615
E	288	MSE	MET	modified residue	UNP O07615
F	126	MSE	MET	modified residue	UNP O07615
F	170	MSE	MET	modified residue	UNP O07615
F	278	MSE	MET	modified residue	UNP O07615
F	288	MSE	MET	modified residue	UNP O07615

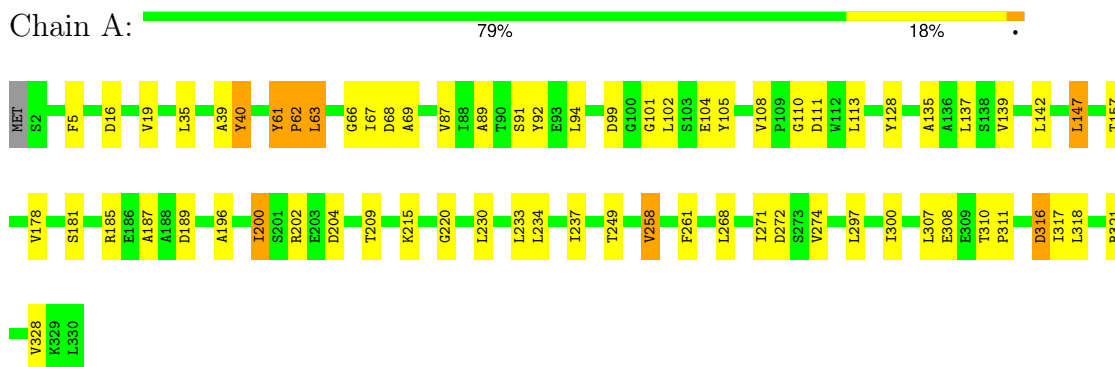
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	25	Total O 25 25	0	0
2	B	26	Total O 26 26	0	0
2	C	29	Total O 29 29	0	0
2	D	19	Total O 19 19	0	0
2	E	30	Total O 30 30	0	0
2	F	31	Total O 31 31	0	0

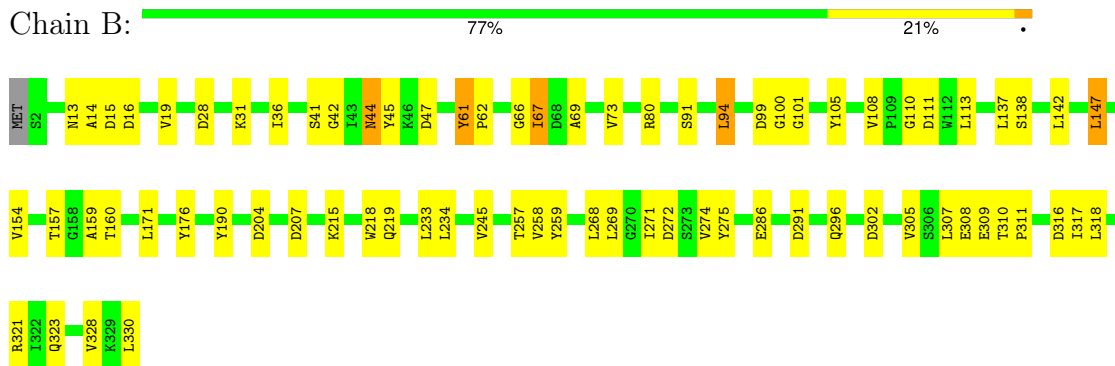
### 3 Residue-property plots

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.

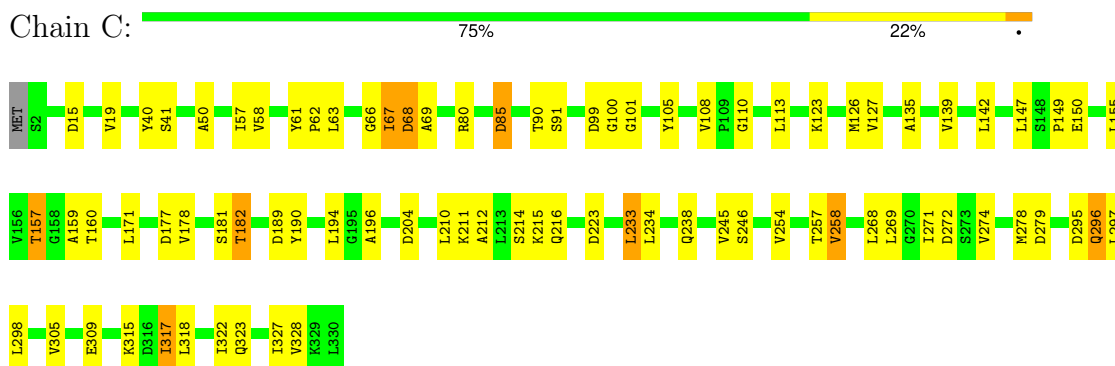
- Molecule 1: YHFP




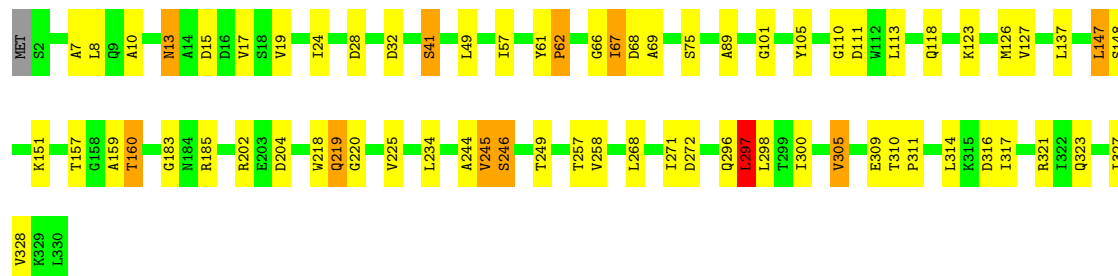
- Molecule 1: YHFP




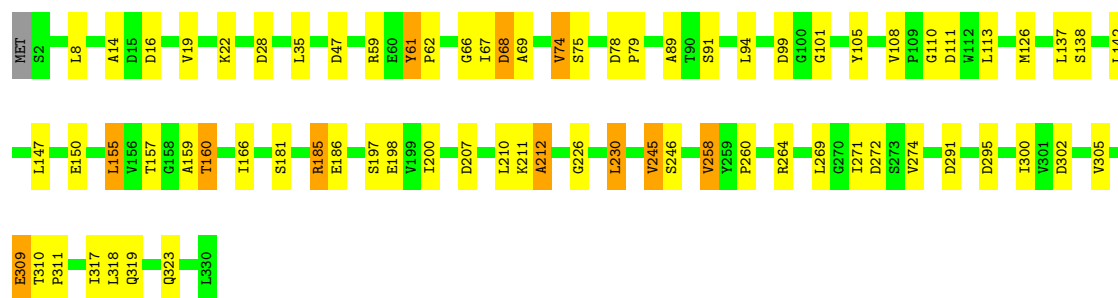
- Molecule 1: YHFP



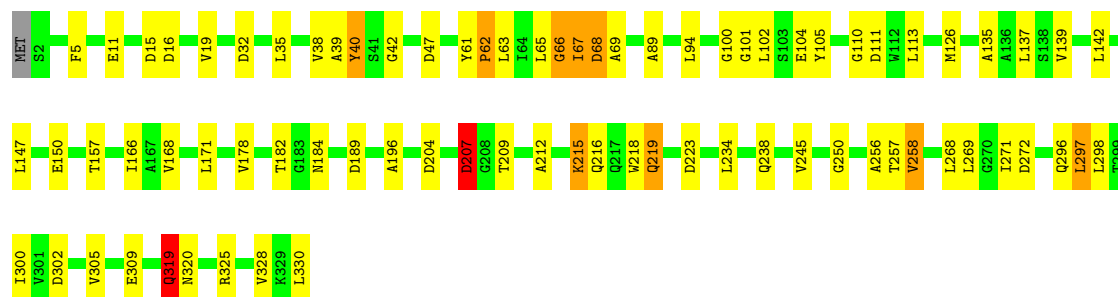
## ● Molecule 1: YHFP

Chain D:  78% 18%

## ● Molecule 1: YHFP

Chain E:  78% 19%

## ● Molecule 1: YHFP

Chain F:  76% 20%

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	102.22Å 133.55Å 166.34Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.70 19.85 – 2.70	Depositor EDS
% Data completeness (in resolution range)	98.7 (20.00-2.70) 98.7 (19.85-2.70)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.07 (at 2.71Å)	Xtriage
Refinement program	REFMAC 5.2.0003	Depositor
R, $R_{free}$	0.187 , 0.262 0.282 , 0.299	Depositor DCC
$R_{free}$ test set	3148 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.5	Xtriage
Anisotropy	0.828	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 28.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.85	EDS
Total number of atoms	14777	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.98% 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 [i](#)

### 5.1 Standard geometry [i](#)

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	0.29	0/2468	0.69	6/3343 (0.2%)
1	B	0.29	0/2468	0.71	11/3343 (0.3%)
1	C	0.29	0/2468	0.70	9/3343 (0.3%)
1	D	0.29	0/2468	0.69	7/3343 (0.2%)
1	E	0.29	0/2468	0.69	10/3343 (0.3%)
1	F	0.29	0/2469	0.69	10/3343 (0.3%)
All	All	0.29	0/14809	0.69	53/20058 (0.3%)

There are no bond length outliers.

The worst 5 of 53 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	16	ASP	CB-CG-OD2	6.28	123.95	118.30
1	B	291	ASP	CB-CG-OD2	6.16	123.84	118.30
1	C	15	ASP	CB-CG-OD2	6.10	123.79	118.30
1	C	177	ASP	CB-CG-OD2	6.01	123.71	118.30
1	A	316	ASP	CB-CG-OD2	5.99	123.69	118.30

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	2436	0	2462	32	0
1	B	2436	0	2462	35	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2436	0	2462	37	0
1	D	2436	0	2462	38	0
1	E	2436	0	2462	32	0
1	F	2437	0	2462	33	0
2	A	25	0	0	1	0
2	B	26	0	0	1	0
2	C	29	0	0	0	0
2	D	19	0	0	1	0
2	E	30	0	0	0	0
2	F	31	0	0	0	0
All	All	14777	0	14772	203	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 7.

The worst 5 of 203 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:F:68:ASP:HB3	1:F:126:MSE:HE1	1.57	0.86
1:D:66:GLY:HA3	1:D:101:GLY:H	1.43	0.83
1:E:68:ASP:HB3	1:E:126:MSE:HE1	1.63	0.81
1:D:245:VAL:O	1:D:246:SER:HB3	1.83	0.79
1:C:257:THR:O	1:C:258:VAL:HB	1.85	0.75

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	327/330 (99%)	295 (90%)	22 (7%)	10 (3%)	<b>4</b> <b>9</b>
1	B	327/330 (99%)	282 (86%)	36 (11%)	9 (3%)	<b>5</b> <b>11</b>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	327/330 (99%)	280 (86%)	36 (11%)	11 (3%)	3	8
1	D	327/330 (99%)	287 (88%)	26 (8%)	14 (4%)	2	5
1	E	327/330 (99%)	285 (87%)	31 (10%)	11 (3%)	3	8
1	F	327/330 (99%)	285 (87%)	25 (8%)	17 (5%)	2	3
All	All	1962/1980 (99%)	1714 (87%)	176 (9%)	72 (4%)	3	7

5 of 72 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	61	TYR
1	A	62	PRO
1	A	67	ILE
1	A	258	VAL
1	B	258	VAL

### 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	261/258 (101%)	253 (97%)	8 (3%)	40	69
1	B	261/258 (101%)	252 (97%)	9 (3%)	37	66
1	C	261/258 (101%)	249 (95%)	12 (5%)	27	54
1	D	261/258 (101%)	252 (97%)	9 (3%)	37	66
1	E	261/258 (101%)	248 (95%)	13 (5%)	24	51
1	F	261/258 (101%)	248 (95%)	13 (5%)	24	51
All	All	1566/1548 (101%)	1502 (96%)	64 (4%)	30	59

5 of 64 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	207	ASP
1	F	219	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	233	LEU
1	C	182	THR
1	F	257	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	219	GLN
1	D	238	GLN
1	F	219	GLN
1	E	238	GLN
1	F	217	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](#)

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](#)

There are no ligands in this entry.

## 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 ⓘ

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.