



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 23, 2020 – 01:35 pm BST

PDB ID : 6ABH  
Title : Structure of a natural red emitting luciferase from Phrixothrix hirtus (P1 crystal form)  
Authors : Carrasco-Lopez, C.; Panjekar, S.; Naumov, P.; Rabeh, W.  
Deposited on : 2018-07-21  
Resolution : 3.05 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
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

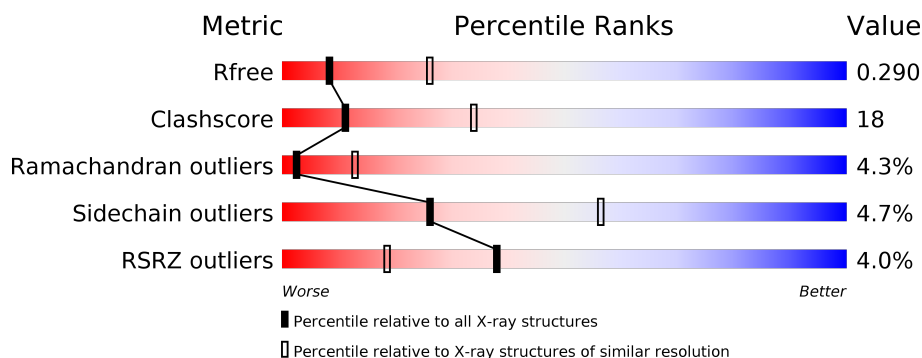
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.05 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	1754 (3.10-3.02)
Clashscore	141614	1864 (3.10-3.02)
Ramachandran outliers	138981	1794 (3.10-3.02)
Sidechain outliers	138945	1793 (3.10-3.02)
RSRZ outliers	127900	1713 (3.10-3.02)

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  $\leq 5\%$ . 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	546	<div> <div>4%</div> <div> <div></div> <div>44%</div> <div>23%</div> <div>•</div> <div>31%</div> </div> </div>
1	B	546	<div> <div>4%</div> <div> <div></div> <div>48%</div> <div>23%</div> <div>•</div> <div>26%</div> </div> </div>
1	C	546	<div> <div>2%</div> <div> <div></div> <div>47%</div> <div>25%</div> <div>•</div> <div>25%</div> </div> </div>
1	D	546	<div> <div>3%</div> <div> <div></div> <div>50%</div> <div>27%</div> <div>• •</div> <div>19%</div> </div> </div>
1	E	546	<div> <div>4%</div> <div> <div></div> <div>43%</div> <div>27%</div> <div>•</div> <div>27%</div> </div> </div>
1	F	546	<div> <div>3%</div> <div> <div></div> <div>51%</div> <div>23%</div> <div>•</div> <div>24%</div> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	546	<div><div><div>3%</div><div>47%</div><div>24%</div><div>•</div><div>27%</div></div></div>
1	H	546	<div><div><div>3%</div><div>43%</div><div>27%</div><div>•</div><div>27%</div></div></div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 25451 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.

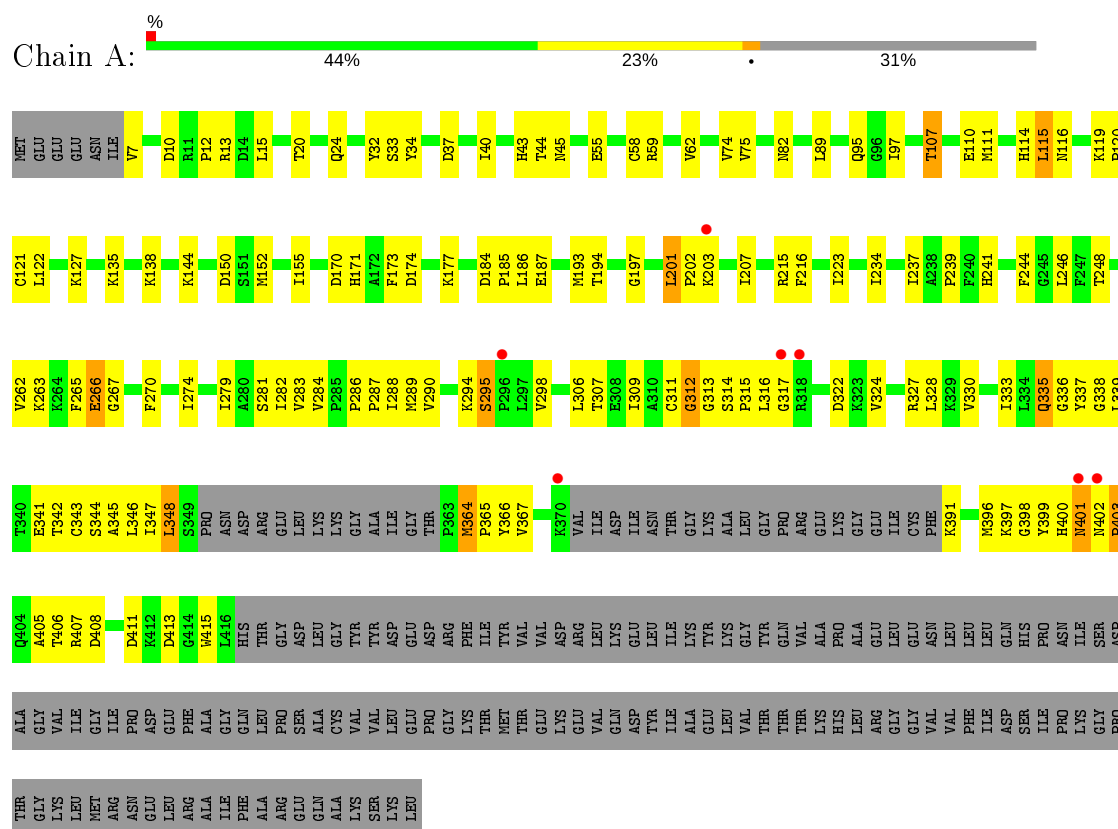
- Molecule 1 is a protein called Red-bioluminescence eliciting luciferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	377	Total	C	N	O	S	0	0	0
			2955	1915	487	535	18			
1	B	405	Total	C	N	O	S	0	0	0
			3170	2047	527	577	19			
1	C	409	Total	C	N	O	S	0	0	0
			3199	2069	531	580	19			
1	D	442	Total	C	N	O	S	0	0	0
			3480	2251	576	634	19			
1	E	398	Total	C	N	O	S	0	0	0
			3110	2011	515	565	19			
1	F	417	Total	C	N	O	S	0	0	0
			3262	2107	541	595	19			
1	G	401	Total	C	N	O	S	0	0	0
			3137	2030	519	570	18			
1	H	400	Total	C	N	O	S	0	0	0
			3138	2028	517	575	18			

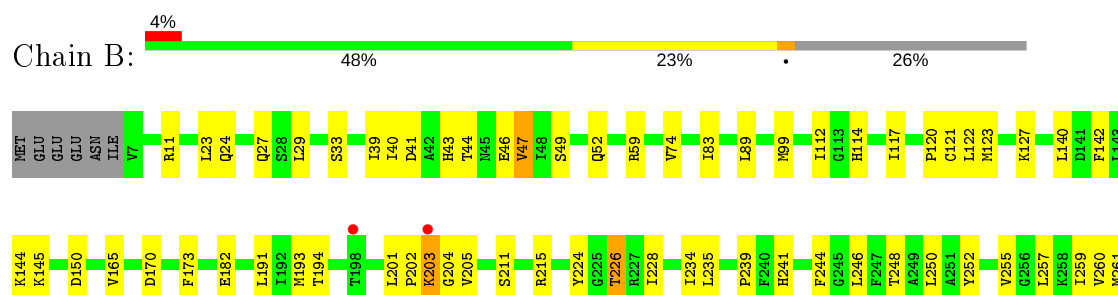
### 3 Residue-property plots [i](#)

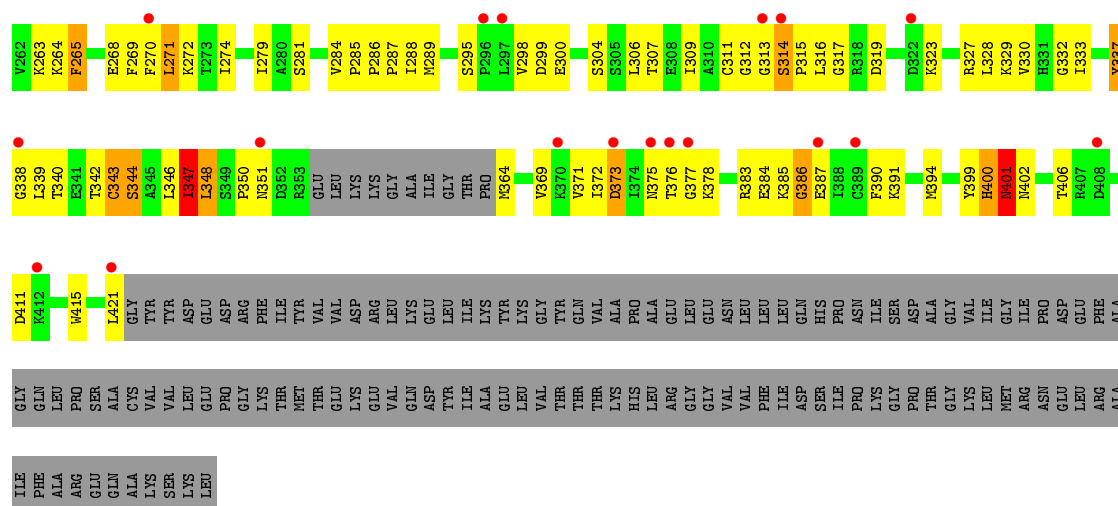
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Red-bioluminescence eliciting luciferase

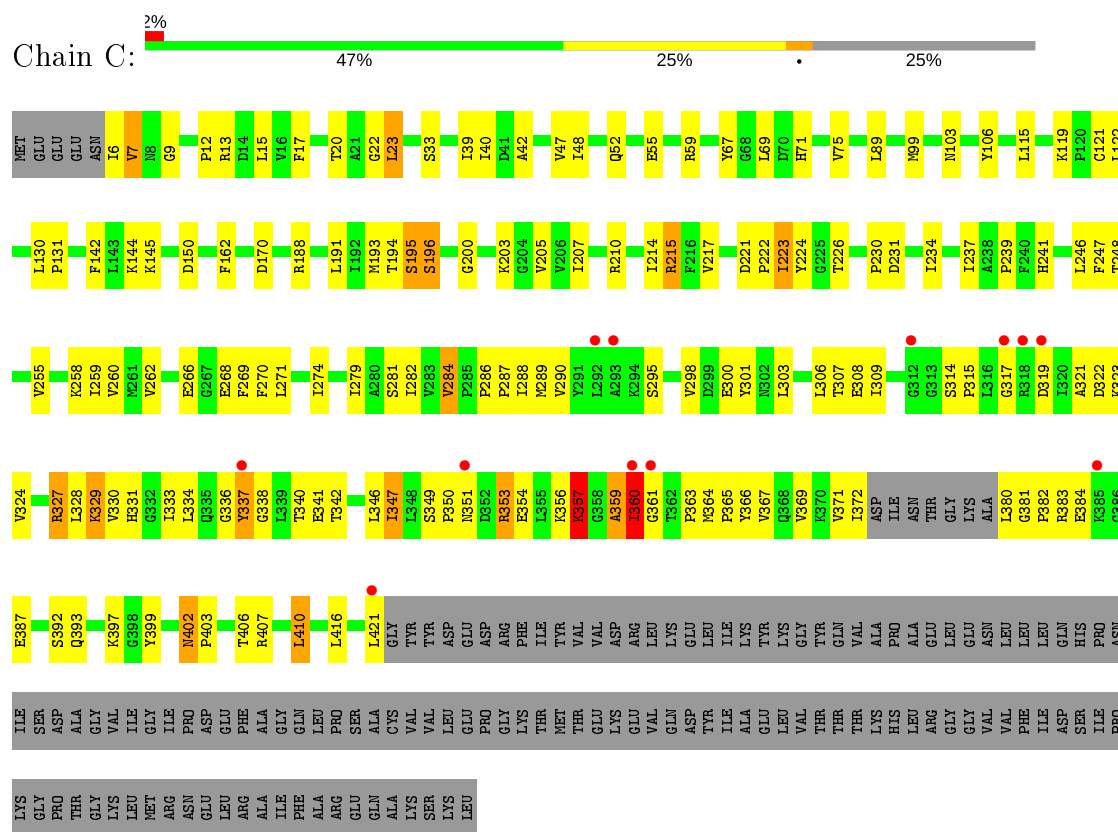


- Molecule 1: Red-bioluminescence eliciting luciferase

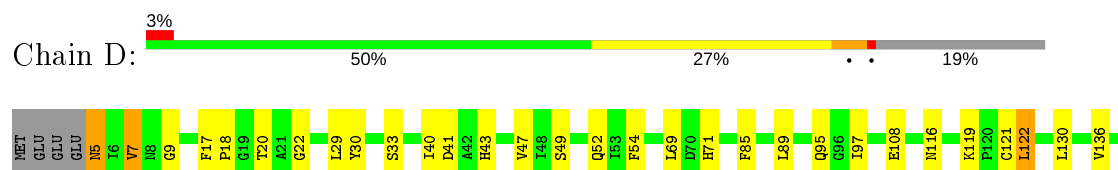


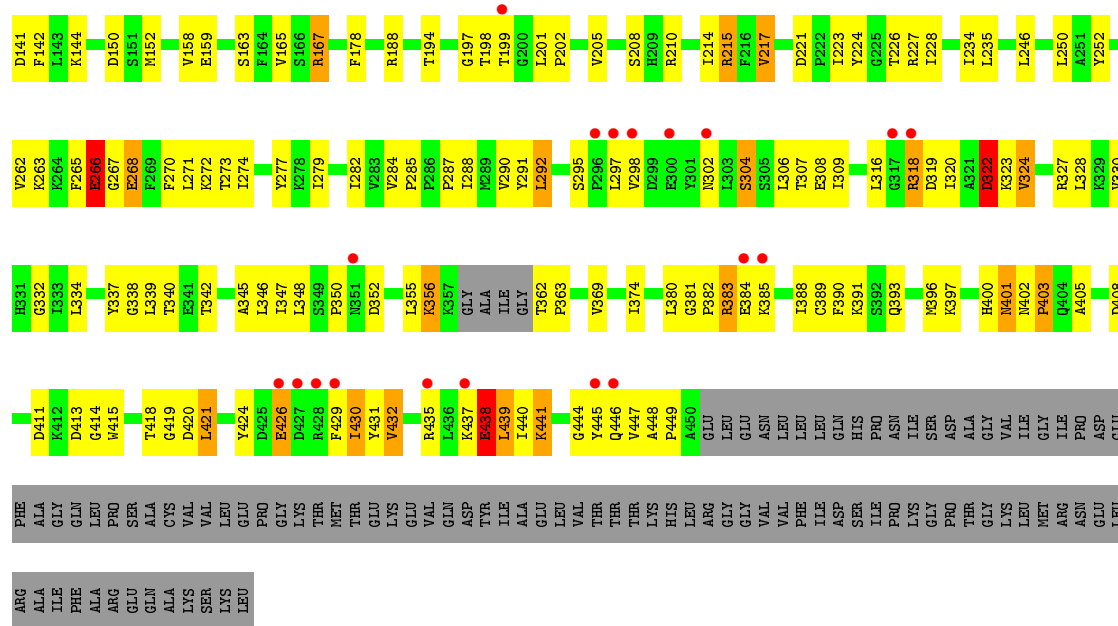


• Molecule 1: Red-bioluminescence eliciting luciferase

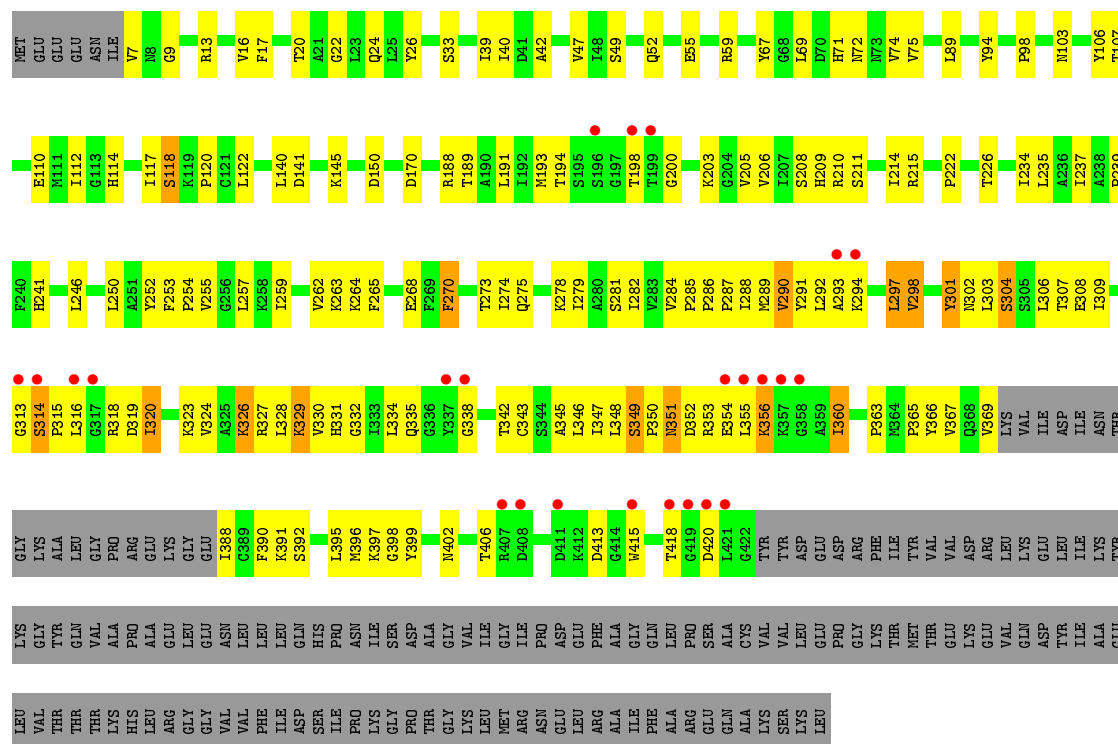
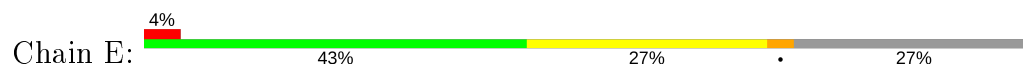


• Molecule 1: Red-bioluminescence eliciting luciferase





• Molecule 1: Red-bioluminescence eliciting luciferase



• Molecule 1: Red-bioluminescence eliciting luciferase







THR	THR	THR	THR	THR	THR
ILE	GLY	S314	I228	I112	MET
ALA	LYS	P315	T232	L115	GLU
GLU	ALA	L316	S233	K119	GLU
LEU	LEU	G317	I234	P120	ASN
VAL	GLY	R318	L235	L140	ILE
THR	PRO	D319	A236	D141	V7
THR	GLN	I320		L140	I8
THR	VAL			D141	G9
LYS	ALA	K323	L246	F142	
HIS	GLY	V324	F247	L143	P12
LEU	GLU	A325	T248	K144	R13
ARG	ILE	K326	K249	K145	D14
GLY	CYS	R327	L250	V146	L15
GLY	F390	L328	A251	I147	V16
VAL	ASN	K329	Y252	V148	F17
VAL	LEU	V330		I149	
PHE	LEU	H331	K258	D150	T20
ILE	LEU	G332	I259		A21
ASP	GLN	I333	V260	I155	G22
SER	HIS	L334	F261		L25
ILE	PRO	Q335	V262	V161	
PRO	ASN	G336		F162	
LYS	ILE	Y337	F265	V165	S33
GLY	SER	M401	E266		
PRO	ASP	M402	G267		D87
THR	ALA	P403	E268	D170	I40
GLY	GLY	Q404	F269	H171	D41
LYS	VAL	A405	F270	A172	A42
LEU	ILE	T406	C343	F173	
MET	GLY	R407	S344		
ARG	ILE	D408		D184	V47
ASN	PRO	A409	I274	F185	I48
GLU	ASP	L410	I279	L186	
LEU	GLU	L348	A280	E187	Q52
ARG	PHE	D413	S281		
ALA	ALA	G414	I282	L191	E55
ILE	GLY	M415	V283	I192	
PHE	GLN	L416	V284	M193	
ALA	LEU	H417	E354		R59
ARG	PRO		L355	S196	V74
GLU	SER	D420	K356	G197	
ALA	ALA	L421	I357	T198	F85
LYS	CYS	G422	G358	T199	F86
SER	VAL	Y423	A359	G200	
LYS	VAL	Y424	I360	L201	I89
SER	LEU	D425	G361	P202	I90
LEU	GLU	E426	T362	K203	
	PRO	D427	P363	G204	Q95
GLY	GLY	ARG	K364		G96
LYS	LYS	PHE		R210	I97
THR	ILE	I367	V298	S211	
MET	THR	Q368	L303	T212	N103
GLU	THR	VAL	S304	T213	
LYS	GLU	LYS	S305	I214	Y106
ASP	LYS	VAL	L306	R215	T107
ARG	ARG	ILE		F216	E108
LEU	LEU	ASP	I309	V217	R109
LYS	GLN	ILE	A310		E110
ASP	ASP	ASN	C311	P222	M111

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	105.70Å 121.17Å 129.44Å 61.86° 68.35° 74.17°	Depositor
Resolution (Å)	48.76 – 3.05 48.76 – 3.05	Depositor EDS
% Data completeness (in resolution range)	94.1 (48.76-3.05) 94.2 (48.76-3.05)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	0.11	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.69 (at 3.07Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.240 , 0.285 0.258 , 0.290	Depositor DCC
$R_{free}$ test set	955 reflections (1.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	73.0	Xtriage
Anisotropy	0.397	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 66.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	25451	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	90.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.17% 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

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	0.59	0/3026	0.77	0/4100
1	B	0.54	0/3245	0.72	0/4398
1	C	0.57	0/3275	0.75	0/4438
1	D	0.62	0/3563	0.78	0/4829
1	E	0.53	0/3185	0.74	0/4318
1	F	0.58	0/3340	0.81	2/4528 (0.0%)
1	G	0.55	0/3213	0.75	0/4356
1	H	0.60	0/3215	0.79	1/4359 (0.0%)
All	All	0.57	0/26062	0.76	3/35326 (0.0%)

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	B	0	1
1	D	0	3
All	All	0	4

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	341	GLU	N-CA-C	-12.20	78.05	111.00
1	H	398	GLY	N-CA-C	-5.69	98.88	113.10
1	F	341	GLU	N-CA-CB	5.18	119.92	110.60

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	11	ARG	Sidechain
1	D	167	ARG	Sidechain
1	D	215	ARG	Sidechain
1	D	318	ARG	Sidechain

## 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	2955	0	2990	107	0
1	B	3170	0	3205	96	0
1	C	3199	0	3246	116	0
1	D	3480	0	3513	141	0
1	E	3110	0	3143	126	0
1	F	3262	0	3293	102	0
1	G	3137	0	3167	103	0
1	H	3138	0	3150	137	0
All	All	25451	0	25707	915	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 18.

The worst 5 of 915 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:201:LEU:HD22	1:A:202:PRO:CD	1.45	1.47
1:A:201:LEU:CD2	1:A:202:PRO:HD3	1.59	1.30
1:F:340:THR:OG1	1:F:344:SER:HA	1.29	1.29
1:G:370:LYS:HD2	1:G:415:TRP:CE3	1.72	1.24
1:D:319:ASP:O	1:D:323:LYS:HG3	1.43	1.17

There are no symmetry-related clashes.

## 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	371/546 (68%)	329 (89%)	29 (8%)	13 (4%)	3	17
1	B	401/546 (73%)	352 (88%)	28 (7%)	21 (5%)	2	10
1	C	405/546 (74%)	359 (89%)	27 (7%)	19 (5%)	2	12
1	D	438/546 (80%)	374 (85%)	44 (10%)	20 (5%)	2	12
1	E	394/546 (72%)	349 (89%)	31 (8%)	14 (4%)	3	16
1	F	413/546 (76%)	362 (88%)	36 (9%)	15 (4%)	3	16
1	G	395/546 (72%)	339 (86%)	35 (9%)	21 (5%)	2	10
1	H	396/546 (72%)	345 (87%)	36 (9%)	15 (4%)	3	15
All	All	3213/4368 (74%)	2809 (87%)	266 (8%)	138 (4%)	2	13

5 of 138 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	266	GLU
1	A	335	GLN
1	A	401	ASN
1	A	403	PRO
1	B	316	LEU

### 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	328/471 (70%)	313 (95%)	15 (5%)	27	57

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	352/471 (75%)	338 (96%)	14 (4%)	31	62
1	C	355/471 (75%)	339 (96%)	16 (4%)	27	58
1	D	384/471 (82%)	365 (95%)	19 (5%)	25	55
1	E	345/471 (73%)	330 (96%)	15 (4%)	29	59
1	F	361/471 (77%)	346 (96%)	15 (4%)	30	60
1	G	348/471 (74%)	328 (94%)	20 (6%)	20	49
1	H	347/471 (74%)	328 (94%)	19 (6%)	21	50
All	All	2820/3768 (75%)	2687 (95%)	133 (5%)	26	56

5 of 133 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	421	LEU
1	E	326	LYS
1	H	270	PHE
1	D	438	GLU
1	E	226	THR

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

Mol	Chain	Res	Type
1	C	331	HIS
1	D	5	ASN
1	F	8	ASN
1	B	95	GLN
1	E	351	ASN

### 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 [i](#)

There are no carbohydrates 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 [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	377/546 (69%)	-0.20	7 (1%) 66 43	49, 71, 150, 244	0
1	B	405/546 (74%)	0.01	20 (4%) 29 13	43, 86, 182, 250	0
1	C	409/546 (74%)	-0.17	12 (2%) 51 26	50, 76, 169, 239	0
1	D	442/546 (80%)	-0.07	19 (4%) 35 16	41, 68, 155, 256	0
1	E	398/546 (72%)	0.01	24 (6%) 21 9	58, 84, 179, 239	0
1	F	417/546 (76%)	-0.14	14 (3%) 45 22	45, 72, 160, 221	0
1	G	401/546 (73%)	-0.13	16 (3%) 38 18	30, 74, 163, 223	0
1	H	400/546 (73%)	-0.07	17 (4%) 35 16	50, 74, 170, 260	0
All	All	3249/4368 (74%)	-0.09	129 (3%) 38 18	30, 77, 170, 260	0

The worst 5 of 129 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	402	ASN	13.1
1	D	429	PHE	8.2
1	H	317	GLY	6.9
1	D	427	ASP	6.9
1	D	317	GLY	6.5

### 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 carbohydrates in this entry.



## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.