



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:50 PM GMT

PDB ID : 4ANI  
Title : Structural basis for the intermolecular communication between DnaK and GrpE in the DnaK chaperone system from *Geobacillus kaustophilus* HTA426  
Authors : Wu, C.-C.; Naveen, V.; Chien, C.-H.; Chang, Y.-W.; Hsiao, C.-D.  
Deposited on : 2012-03-19  
Resolution : 4.09 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  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

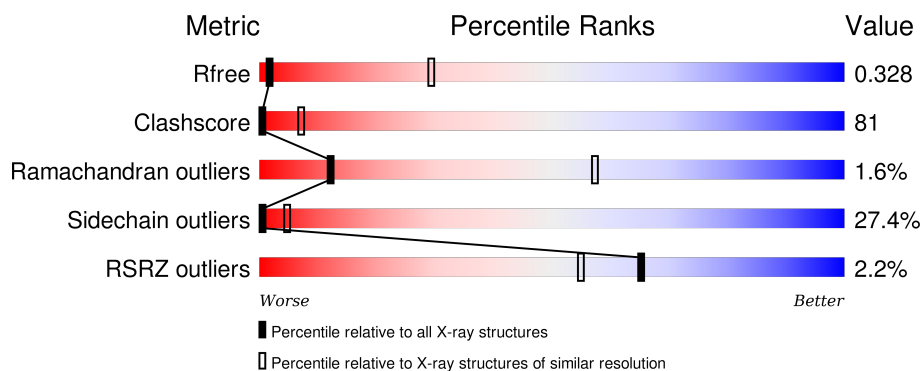
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 4.09 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))
$R_{free}$	91344	1018 (4.60-3.60)
Clashscore	102246	1117 (4.60-3.60)
Ramachandran outliers	100387	1063 (4.60-3.60)
Sidechain outliers	100360	1049 (4.60-3.60)
RSRZ outliers	91569	1022 (4.60-3.60)

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. 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	213	
1	B	213	
1	E	213	
1	F	213	
2	C	509	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	509	<div><div></div><div>19%60%21%</div></div>
2	G	509	<div><div></div><div>7%34%54%11%•</div></div>
2	H	509	<div><div></div><div>3%34%47%17%••</div></div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 18889 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 PROTEIN GRPE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	156	Total	C	N	O	S	0	0	0
			1242	778	215	243	6			
1	B	154	Total	C	N	O	S	0	0	0
			1240	779	216	239	6			
1	E	135	Total	C	N	O	S	0	0	0
			820	498	160	160	2			
1	F	130	Total	C	N	O	S	0	0	0
			815	496	159	158	2			

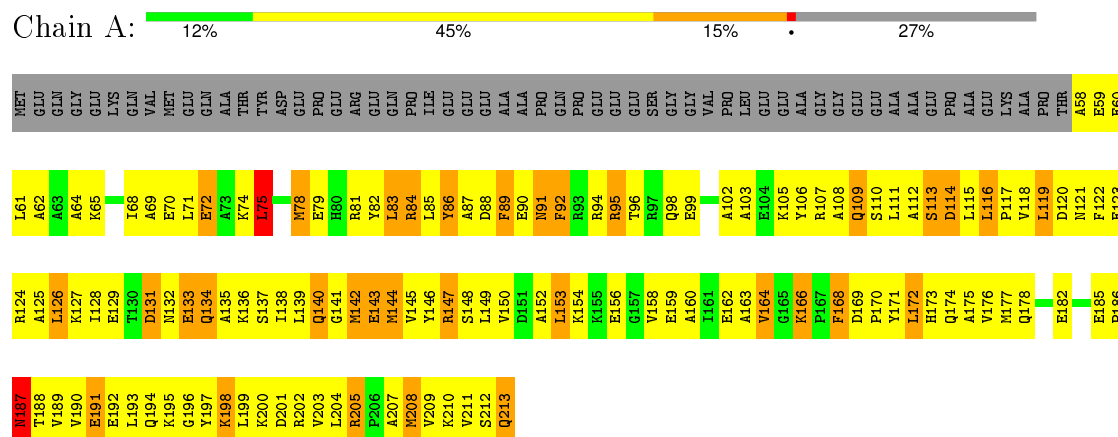
- Molecule 2 is a protein called CHAPERONE PROTEIN DNAK.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	509	Total	C	N	O	S	0	0	0
			3862	2410	670	773	9			
2	D	509	Total	C	N	O	S	0	0	0
			3864	2413	671	771	9			
2	G	502	Total	C	N	O	S	0	0	0
			3527	2178	627	716	6			
2	H	501	Total	C	N	O	S	0	0	0
			3519	2176	624	712	7			

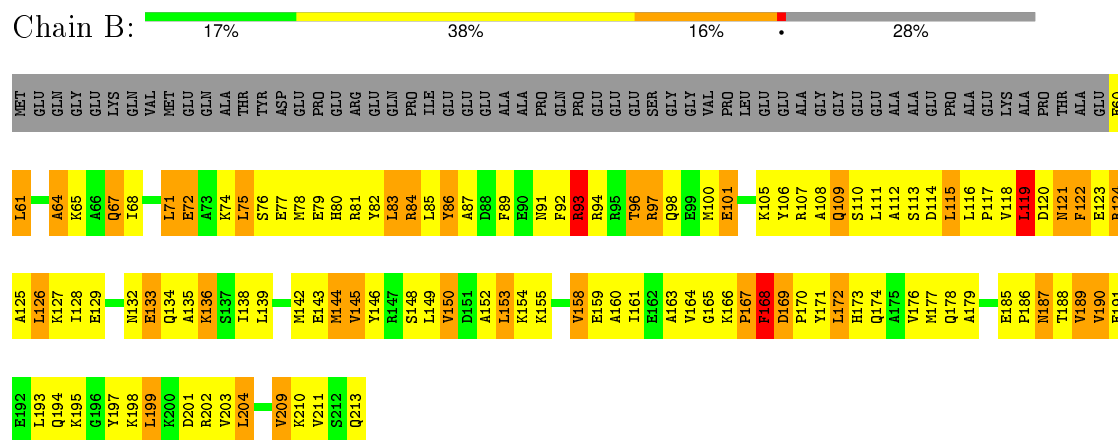
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

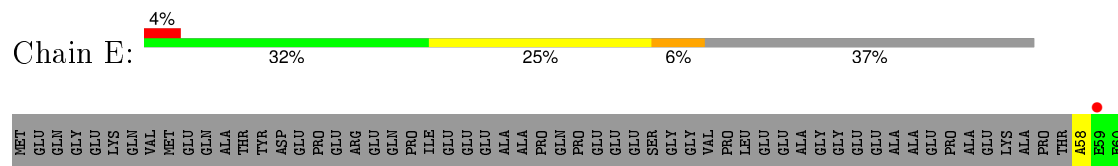
#### • Molecule 1: PROTEIN GRPE



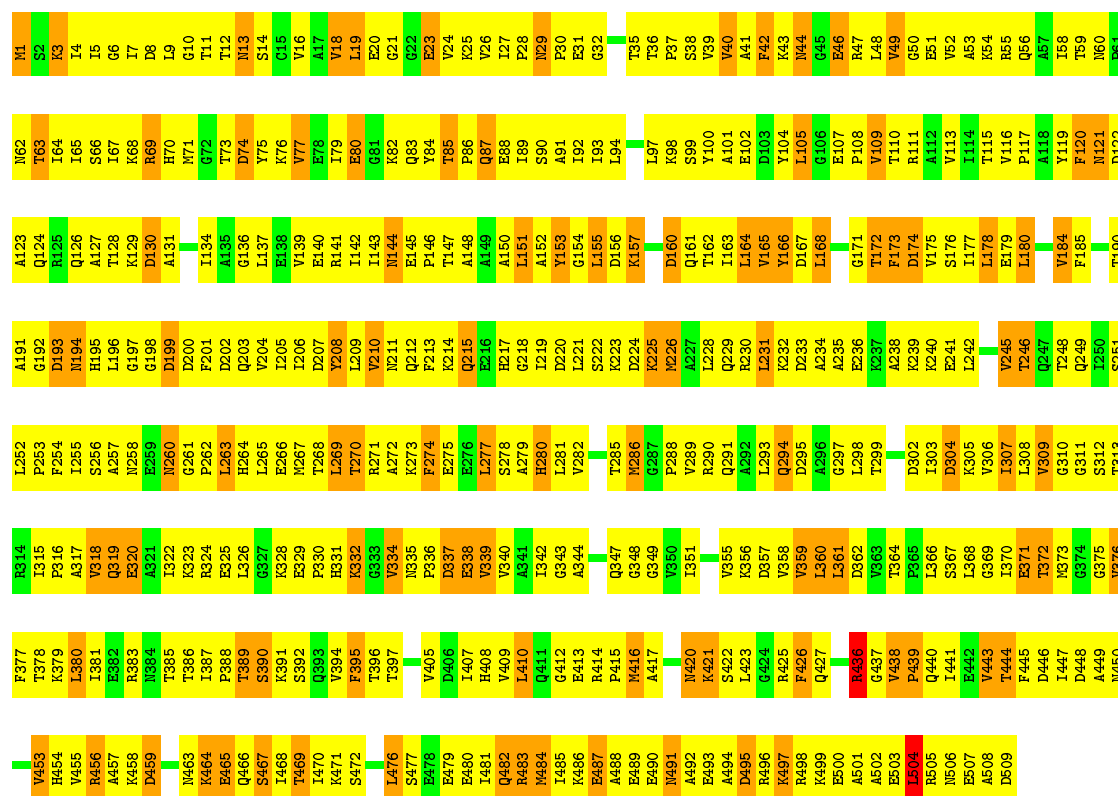
#### • Molecule 1: PROTEIN GRPE



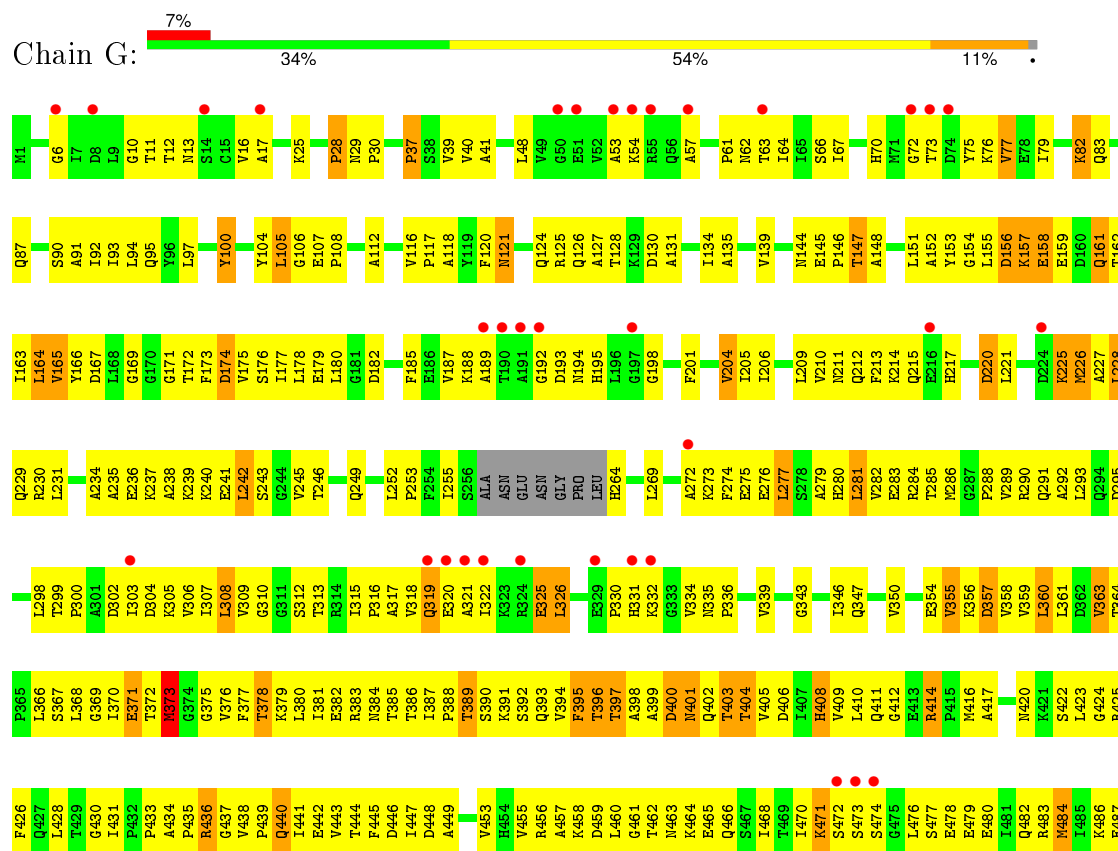
#### • Molecule 1: PROTEIN GRPE





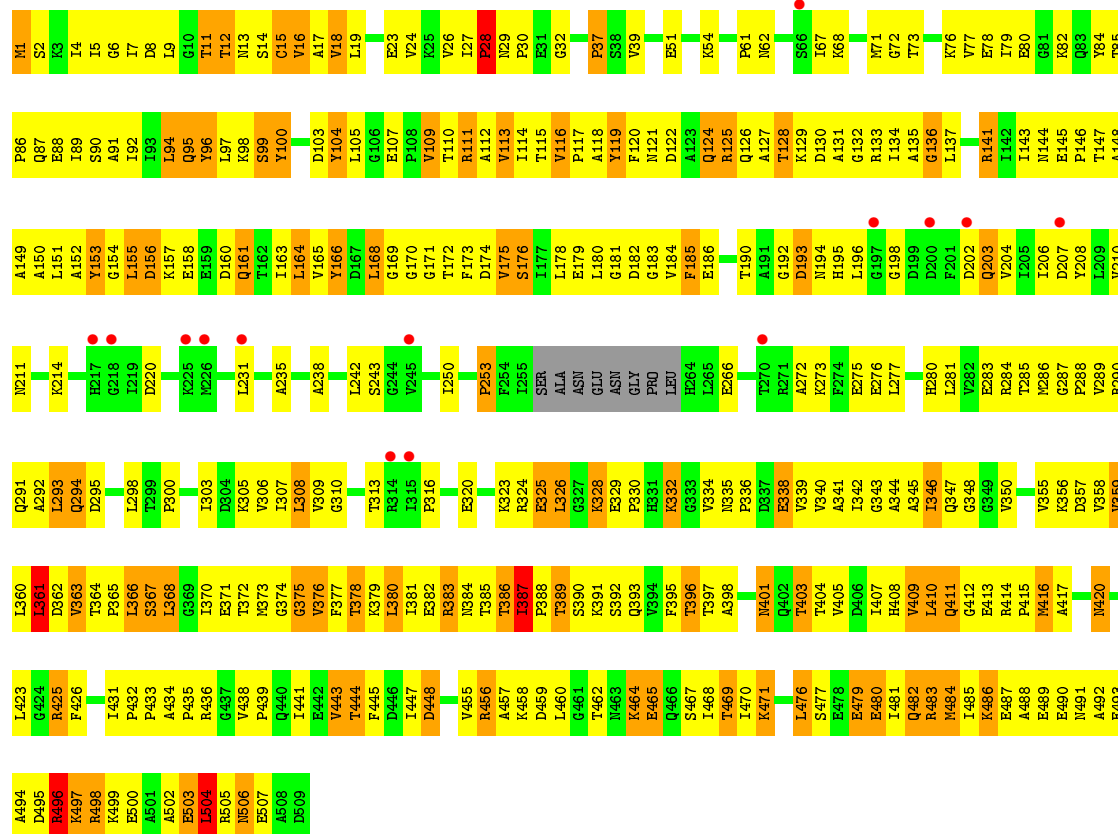


• Molecule 2: CHAPERONE PROTEIN DNAK





## • Molecule 2: CHAPERONE PROTEIN DNAK





## 4 Data and refinement statistics

Property	Value	Source
Space group	I 41 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	279.98Å 279.98Å 278.82Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	26.22 – 4.09 26.22 – 4.09	Depositor EDS
% Data completeness (in resolution range)	94.8 (26.22-4.09) 94.8 (26.22-4.09)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.25 (at 4.10Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE: 1.7_650)	Depositor
R, $R_{free}$	0.274 , 0.347 0.251 , 0.328	Depositor DCC
$R_{free}$ test set	2083 reflections (5.05%)	DCC
Wilson B-factor (Å <sup>2</sup> )	143.4	Xtriage
Anisotropy	0.406	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , 133.3	EDS
Estimated twinning fraction	0.027 for l,-k,h 0.016 for -h,-l,-k	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	0 of 41252 reflections	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	18889	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	207.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.31% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.75	0/1258	0.90	1/1691 (0.1%)
1	B	0.74	0/1257	0.93	2/1688 (0.1%)
1	E	0.64	0/821	0.84	5/1117 (0.4%)
1	F	0.64	0/818	0.84	5/1110 (0.5%)
2	C	0.70	1/3912 (0.0%)	0.86	3/5296 (0.1%)
2	D	0.72	0/3914	0.87	0/5299
2	G	0.67	0/3564	0.85	8/4834 (0.2%)
2	H	0.70	1/3556 (0.0%)	0.92	15/4832 (0.3%)
All	All	0.70	2/19100 (0.0%)	0.88	39/25867 (0.2%)

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
2	D	0	2
2	H	0	2
All	All	0	5

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	141	ARG	CZ-NH1	-6.64	1.24	1.33
2	C	64	ILE	CB-CG2	-5.03	1.37	1.52

The worst 5 of 39 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	141	ARG	NE-CZ-NH2	10.76	125.68	120.30
2	H	361	LEU	CB-CG-CD1	-7.35	98.50	111.00
2	H	387	ILE	CG1-CB-CG2	-7.13	95.71	111.40

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	G	117	PRO	N-CA-CB	7.10	111.82	103.30
2	G	108	PRO	N-CA-CB	6.86	111.53	103.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	167	PRO	Peptide
2	D	394	VAL	Peptide
2	D	436	ARG	Peptide
2	H	477	SER	Peptide
2	H	496	ARG	Peptide

## 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	1242	0	1239	264	0
1	B	1240	0	1246	243	0
1	E	820	0	567	85	0
1	F	815	0	588	79	0
2	C	3862	0	3901	638	0
2	D	3864	0	3909	766	0
2	G	3527	0	3217	518	0
2	H	3519	0	3231	544	0
All	All	18889	0	17898	2971	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 81.

The worst 5 of 2971 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)
2:C:67:ILE:CD1	2:C:89:ILE:HG21	1.51	1.40
2:G:152:ALA:HA	2:G:346:ILE:CD1	1.51	1.39

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:141:ARG:NH1	2:H:143:ILE:HD11	1.40	1.36
2:H:67:ILE:CD1	2:H:89:ILE:HG21	1.57	1.31
2:D:426:PHE:HB3	2:D:468:ILE:CD1	1.61	1.30

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	154/213 (72%)	125 (81%)	26 (17%)	3 (2%)	10	54
1	B	152/213 (71%)	119 (78%)	27 (18%)	6 (4%)	4	38
1	E	127/213 (60%)	103 (81%)	22 (17%)	2 (2%)	12	57
1	F	124/213 (58%)	100 (81%)	22 (18%)	2 (2%)	12	57
2	C	507/509 (100%)	458 (90%)	47 (9%)	2 (0%)	39	79
2	D	507/509 (100%)	426 (84%)	72 (14%)	9 (2%)	11	55
2	G	498/509 (98%)	430 (86%)	62 (12%)	6 (1%)	16	62
2	H	497/509 (98%)	423 (85%)	63 (13%)	11 (2%)	8	51
All	All	2566/2888 (89%)	2184 (85%)	341 (13%)	41 (2%)	12	57

5 of 41 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	359	VAL
1	E	170	PRO
2	H	253	PRO
1	B	168	PHE
2	D	361	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	128/174 (74%)	89 (70%)	39 (30%)	0	3
1	B	129/174 (74%)	88 (68%)	41 (32%)	0	3
1	E	35/174 (20%)	24 (69%)	11 (31%)	0	3
1	F	39/174 (22%)	26 (67%)	13 (33%)	0	2
2	C	416/418 (100%)	302 (73%)	114 (27%)	0	5
2	D	416/418 (100%)	305 (73%)	111 (27%)	0	5
2	G	316/418 (76%)	251 (79%)	65 (21%)	1	12
2	H	319/418 (76%)	221 (69%)	98 (31%)	0	3
All	All	1798/2368 (76%)	1306 (73%)	492 (27%)	0	5

5 of 492 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	144	ASN
2	D	410	LEU
2	H	382	GLU
2	D	165	VAL
2	D	260	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 55 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	161	GLN
2	D	466	GLN
2	H	212	GLN
2	D	260	ASN
2	D	319	GLN

### 5.3.3 RNA ⓘ

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](#)

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 ⓘ

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	156/213 (73%)	-0.67	0 100 100	108, 165, 240, 281	0
1	B	154/213 (72%)	-0.62	0 100 100	108, 151, 213, 293	0
1	E	135/213 (63%)	0.27	9 (6%) 21 14	193, 261, 363, 476	0
1	F	130/213 (61%)	-0.01	0 100 100	185, 244, 324, 370	0
2	C	509/509 (100%)	-0.60	0 100 100	109, 152, 216, 283	0
2	D	509/509 (100%)	-0.61	0 100 100	109, 163, 217, 333	0
2	G	502/509 (98%)	0.01	34 (6%) 20 14	141, 283, 373, 481	0
2	H	501/509 (98%)	-0.21	14 (2%) 56 45	115, 237, 370, 516	0
All	All	2596/2888 (89%)	-0.34	57 (2%) 65 55	108, 193, 341, 516	0

The worst 5 of 57 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	205	ARG	10.6
1	E	206	PRO	10.0
2	H	225	LYS	6.1
2	G	320	GLU	6.1
2	H	245	VAL	5.4

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.