



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 05:07 PM GMT

PDB ID : 4H8L  
Title : Crystal structure of a parallel 6-helix coiled coil CC-Hex-D24-A5/7C  
Authors : Chi, B.; Zaccai, N.R.; Brady, R.L.; Woolfson, D.N.  
Deposited on : 2012-09-23  
Resolution : 1.75 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  
<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

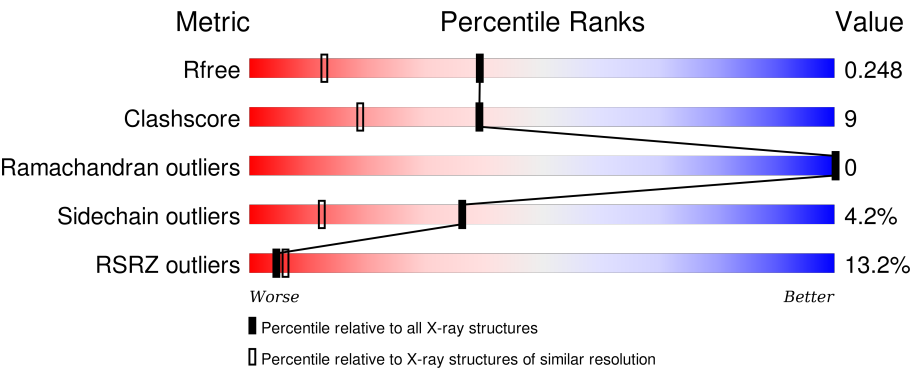


# 1 Overall quality at a glance i

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

The reported resolution of this entry is 1.75 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 <sub>free</sub>	91344	1609 (1.76-1.76)
Clashscore	102246	1730 (1.76-1.76)
Ramachandran outliers	100387	1711 (1.76-1.76)
Sidechain outliers	100360	1711 (1.76-1.76)
RSRZ outliers	91569	1610 (1.76-1.76)

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	34	<div><div>9%</div><div><div></div><div>71%</div><div>9%</div><div>21%</div></div></div>
1	B	34	<div><div>12%</div><div><div></div><div>65%</div><div>9%</div><div>6%</div><div>21%</div></div></div>
1	C	34	<div><div>12%</div><div><div></div><div>65%</div><div>15%</div><div>21%</div></div></div>
1	D	34	<div><div>15%</div><div><div></div><div>62%</div><div>18%</div><div>21%</div></div></div>
1	E	34	<div><div>6%</div><div><div></div><div>53%</div><div>21%</div><div>26%</div></div></div>

Continued on next page...



Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	34	<p>6% 62% 12% 26%</p>



## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 1318 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 CC-Hex-D24-A5/7C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	27	Total	C	N	O	S	0	0	0
			204	131	33	38	2			
1	B	27	Total	C	N	O	S	0	1	0
			210	135	33	40	2			
1	C	27	Total	C	N	O	S	0	0	0
			204	131	33	38	2			
1	D	27	Total	C	N	O	S	0	0	0
			204	131	33	38	2			
1	E	25	Total	C	N	O	S	0	1	0
			196	126	30	38	2			
1	F	25	Total	C	N	O	S	0	0	0
			190	122	30	36	2			

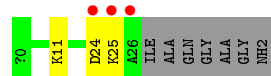
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	18	Total	O	0	0
			18	18		
2	B	19	Total	O	0	0
			19	19		
2	C	22	Total	O	0	0
			22	22		
2	D	20	Total	O	0	0
			20	20		
2	E	17	Total	O	0	0
			17	17		
2	F	14	Total	O	0	0
			14	14		





- Molecule 1: CC-Hex-D24-A5/7C



- 

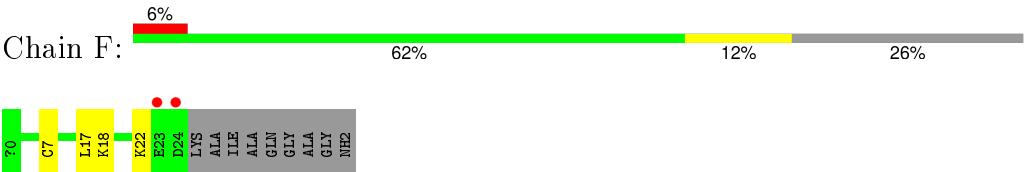
- 
- Diagram illustrating the schematic representation of the protein structure of the 100 kDa subunit of the 20S proteasome. The structure is shown as a vertical bar with various regions labeled. The regions are: ?0 (yellow), K4 (yellow), C5 (yellow), K11 (yellow), A12 (green), I13 (yellow), E23 (green), D24 (green), K25 (green), A26 (green), ILE (grey), ALA (grey), GLN (grey), GLY (grey), ALA (grey), GLY (grey), and NH2 (grey). Four red dots are positioned above the E23, D24, K25, and A26 regions.

- 
- ?O  
 G1  
 C7  
 Q8  
 E9  
 L10  
 K11  
 W15  
 K18  
 K22  
 E23  
 D24  
 K25  
 A26  
 ILE  
 ALA  
 GLN  
 GLY  
 ALA  
 GLY  
 NH2

- 
- Sequence logo for the 10th position. The y-axis represents information content in bits, ranging from 0 to 1. The x-axis shows amino acids: ?0, C7, Q8, E9, L10, K11, K18, A19, K22, E23, D24, LYS, ALA, ILE, ALA, GLN, GLY, ALA, GLY, NH2. The bar for D24 is the tallest, reaching approximately 0.8 bits. Red dots are placed above the D24 bar and the LYS bar.

- 
- WORLD WIDE  
PDB  
PROTEIN DATA BANK







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	31.54Å 54.33Å 98.28Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.03 – 1.75 30.03 – 1.75	Depositor EDS
% Data completeness (in resolution range)	92.0 (30.03-1.75) 91.2 (30.03-1.75)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.23 (at 1.75Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7_650)	Depositor
R, $R_{free}$	0.200 , 0.254 0.196 , 0.248	Depositor DCC
$R_{free}$ test set	816 reflections (5.01%)	DCC
Wilson B-factor (Å <sup>2</sup> )	21.4	Xtriage
Anisotropy	0.398	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 73.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	1 of 16772 reflections (0.006%)	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	1318	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 25.29 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 3.2596e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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

Bond lengths and bond angles in the following residue types are not validated in this section: ACE

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	0.70	0/203	0.61	0/270
1	B	0.78	0/212	0.60	0/282
1	C	0.67	0/203	0.67	0/270
1	D	0.72	0/203	0.68	0/270
1	E	0.59	0/198	0.69	0/264
1	F	0.62	0/189	0.63	0/252
All	All	0.68	0/1208	0.65	0/1608

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	204	0	216	2	0
1	B	210	0	222	6	0
1	C	204	0	216	4	0
1	D	204	0	216	7	0
1	E	196	0	204	7	0
1	F	190	0	198	2	0
2	A	18	0	0	1	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	19	0	0	0	0
2	C	22	0	0	1	0
2	D	20	0	0	0	1
2	E	17	0	0	0	0
2	F	14	0	0	0	1
All	All	1318	0	1272	22	1

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 9.

All (22) 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:18:LYS:O	1:F:22:LYS:HG2	1.71	0.87
1:E:18:LYS:O	1:E:22:LYS:HD3	1.76	0.85
1:D:18:LYS:O	1:D:22:LYS:HG2	1.82	0.80
1:E:7:CYS:O	1:E:11:LYS:HG3	2.00	0.61
1:D:18:LYS:HG3	1:D:22:LYS:HE3	1.82	0.61
1:E:18:LYS:HG3	1:E:22:LYS:HE2	1.82	0.60
1:D:18:LYS:CG	1:D:22:LYS:HE3	2.34	0.58
1:A:24:ASP:HB3	1:B:23:GLU:CD	2.23	0.58
1:C:0:ACE:H1	1:C:4:LYS:H	1.71	0.56
1:B:20:ILE:HA	1:B:23:GLU:OE1	2.09	0.51
1:A:11:LYS:HG2	1:B:9:GLU:CD	2.33	0.50
1:D:18:LYS:O	1:D:22:LYS:HE2	2.14	0.47
1:B:24:ASP:HB3	1:E:23:GLU:OE2	2.14	0.47
1:B:11:LYS:HG2	1:E:9:GLU:CD	2.36	0.46
1:D:7:CYS:O	1:D:11:LYS:HG3	2.16	0.45
1:B:23:GLU:O	1:B:23:GLU:HG2	2.15	0.45
1:E:18:LYS:HG3	1:E:22:LYS:CE	2.46	0.45
2:A:110:HOH:O	1:D:1:GLY:N	2.46	0.43
1:C:11:LYS:HG2	1:D:9:GLU:CD	2.40	0.42
1:E:19:ALA:O	1:E:23:GLU:HB2	2.20	0.41
1:C:13:ILE:HG23	1:F:17:LEU:CD1	2.51	0.41
1:C:4:LYS:HE3	2:C:116:HOH:O	2.20	0.40

All (1) 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	Interatomic distance (Å)	Clash overlap (Å)
2:D:116:HOH:O	2:F:112:HOH:O[2_555]	2.18	0.02

### 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	25/34 (74%)	25 (100%)	0	0	100	100
1	B	26/34 (76%)	26 (100%)	0	0	100	100
1	C	25/34 (74%)	24 (96%)	1 (4%)	0	100	100
1	D	25/34 (74%)	25 (100%)	0	0	100	100
1	E	24/34 (71%)	23 (96%)	1 (4%)	0	100	100
1	F	23/34 (68%)	23 (100%)	0	0	100	100
All	All	148/204 (72%)	146 (99%)	2 (1%)	0	100	100

There are no Ramachandran outliers to report.

#### 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	20/22 (91%)	19 (95%)	1 (5%)	30	8
1	B	21/22 (96%)	19 (90%)	2 (10%)	11	1
1	C	20/22 (91%)	19 (95%)	1 (5%)	30	8
1	D	20/22 (91%)	20 (100%)	0	100	100

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	E	20/22 (91%)	20 (100%)	0	100	100
1	F	19/22 (86%)	18 (95%)	1 (5%)	28	7
All	All	120/132 (91%)	115 (96%)	5 (4%)	36	12

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	25	LYS
1	B	23	GLU
1	B	24	ASP
1	C	5	CYS
1	F	7	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

There are no ligands in this entry.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.



## 5.8 Polymer linkage issues ⓘ

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	26/34 (76%)	0.64	3 (11%) 6 8	11, 21, 61, 75	0
1	B	26/34 (76%)	0.92	4 (15%) 3 4	10, 21, 73, 86	0
1	C	26/34 (76%)	0.81	4 (15%) 3 4	14, 25, 84, 89	0
1	D	26/34 (76%)	0.67	5 (19%) 2 2	13, 26, 67, 73	0
1	E	24/34 (70%)	0.48	2 (8%) 14 17	13, 28, 55, 82	0
1	F	24/34 (70%)	0.62	2 (8%) 14 17	13, 25, 46, 91	0
All	All	152/204 (74%)	0.69	20 (13%) 4 6	10, 25, 77, 91	0

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	24	ASP	10.5
1	B	24	ASP	8.4
1	B	26	ALA	8.1
1	A	24	ASP	7.6
1	C	26	ALA	5.5
1	C	24	ASP	5.2
1	C	25	LYS	5.2
1	B	23	GLU	4.6
1	A	26	ALA	4.6
1	D	24	ASP	4.1
1	C	23	GLU	4.0
1	F	23	GLU	3.8
1	A	25	LYS	3.8
1	D	25	LYS	3.6
1	B	25	LYS	3.5
1	E	24	ASP	3.4
1	D	26	ALA	2.8
1	D	23	GLU	2.5
1	E	23	GLU	2.5

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	15	TRP	2.1

## 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.