



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 26, 2023 – 08:50 PM EDT

PDB ID : 3QJO  
Title : Refined Structure of the functional unit (KLH1-H) of keyhole limpet hemo-  
cyanin  
Authors : Jaenicke, E.; Buchler, K.; Decker, H.; Markl, J.; Schroder, G.F.  
Deposited on : 2011-01-30  
Resolution : 4.00 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  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

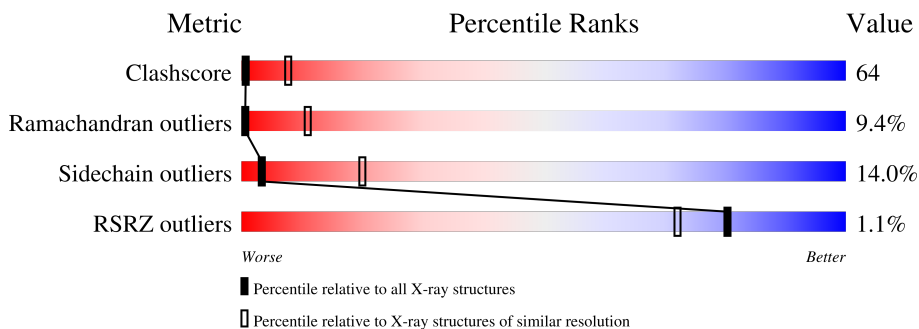
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.00 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))
Clashscore	141614	1148 (4.30-3.70)
Ramachandran outliers	138981	1108 (4.30-3.70)
Sidechain outliers	138945	1099 (4.30-3.70)
RSRZ outliers	127900	1028 (4.34-3.66)

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	491	 26% 58% 15%
1	B	491	 25% 58% 15%

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 7994 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 Hemocyanin 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	491	3995	2558	683	736	18	0	0	0
1	B	491	3995	2558	683	736	18	0	0	0

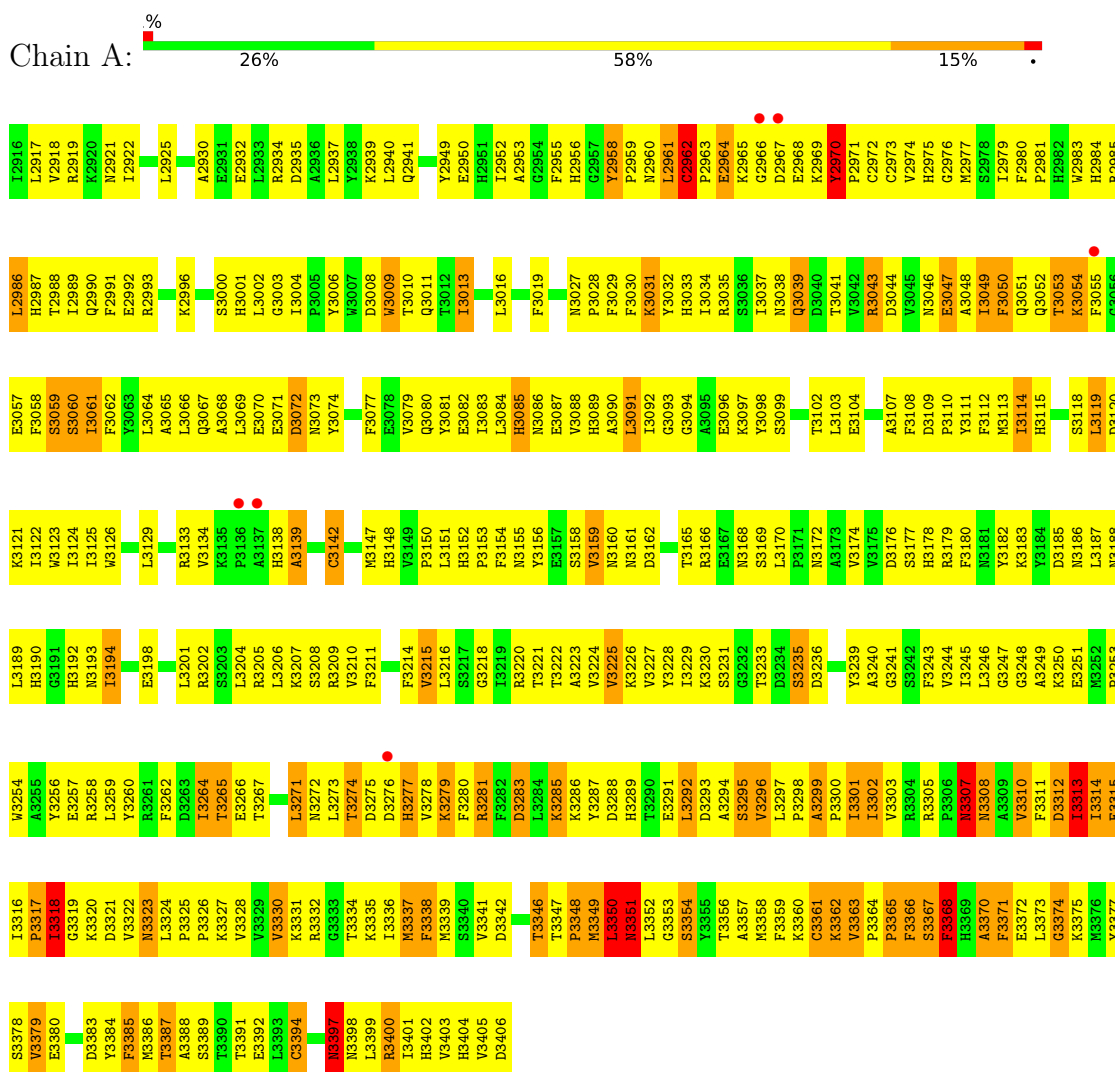
- Molecule 2 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Cu	0	0
			2	2		
2	B	2	Total	Cu	0	0
			2	2		

### 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: Hemocyanin 1



- Molecule 1: Hemocyanin 1



V3379	E3380	S3381	G3382	D3383	Y3384	F3385	K3386	T3387	A3388	S3389	T3391	E3392	C3393	C3394	<b>K3397</b>	K3398	L3399	R3400	I3401	H3402	V3403	H3404	V3405	D3406																																				
P3317	<b>I3316</b>	G3319	K3320	K3321	D3322	V3323	L3324	P3325	P3326	K3327	V3328	V3329	V3330	K3331	R3332	G3333	T3334	K3335	I3336	M3337	F3338	M3339	S3340	V3341	D3342	T3346	T3347	K3348	M3349	L3350	K3351	L3352	G3353	S3354	Y3355	T3356	A3357	M3358	F3359	K3360	C3361	K3362	V3363	P3364	P3365	F3366	S3367	<b>F3368</b>	H3369	A3370	F3371	E3372	L3373	G3374	K3375	<b>K3376</b>	M3377	Y3378	S3379	
H3254	E3257	R3258	L3259	V3260	<b>K3261</b>	F3262	D3263	L3264	T3265	E3266	T3267	L3271	N3272	L3273	T3274	D3275	D3276	H3277	V3278	K3279	F3280	K3281	<b>F3282</b>	D3283	L3284	K3285	K3286	Y3287	D3288	H3289	T3290	E3291	L3292	D3293	A3294	S3295	V3296	L3297	P3298	A3299	P3300	L3301	V3302	V3303	<b>K3304</b>	K3305	<b>P3306</b>	N3307	N3308	A3309	V3310	V3311	G3312	F3313	V3314	<b>K3315</b>	L3316	E3317	L3318	S3319
G3191	H3192	N3193	I3194	E3195	E3196	<b>L3197</b>	E3198	L3201	R3202	S3203	R3204	R3205	L3206	K3207	S3208	R3209	V3210	F3211	F3214	P3150	L3151	H3152	P3153	F3154	N3155	Y3156	E3157	S3158	V3159	H3160	K3161	D3162	T3165	R3166	S3169	L3170	<b>F3171</b>	N3172	<b>K3173</b>	V3174	<b>V3175</b>	D3176	S3177	H3178	R3179	Y3182	K3183	<b>V3184</b>	D3185	N3186	L3187	N3188	M3252	H3190						
H2984	R2985	L2986	<b>H2987</b>	S2988	T2989	L2990	F2991	E2992	K2996	S3000	H3001	L3002	G3003	I3004	D2935	F3005	Y3006	<b>H3007</b>	D3008	W3009	L2940	Q3011	<b>T3012</b>	L3013	L3016	F3019	N3027	P3028	F3029	F3030	K3031	Y3032	H3033	I3034	R3035	<b>S3036</b>	K3037	I3037	N3038	<b>Q3039</b>	D3040	V3041	<b>V3042</b>	R3043	D3044	<b>D3045</b>	N3046	E3047	A3048	G2976	I3049	<b>F3050</b>	M3113	Q3051	Q3052	T3053	<b>K3054</b>	<b>F3055</b>		
G3056	E3057	F3058	S3059	S3060	I3061	<b>F3062</b>	<b>T3063</b>	L3064	A3065	L3066	<b>Q3067</b>	<b>A3068</b>	L3069	E3070	E3071	R2934	<b>D3072</b>	N3073	Y3074	<b>F3077</b>	<b>E3078</b>	V3079	Q3080	Y3081	E3082	I3083	L3084	H3085	N3086	E3087	V3088	H3089	A3090	L3092	G3093	G3094	<b>A3095</b>	E3096	K3097	Y3098	S3099	T3102	L3103	E3104	A3107	F3108	D3109	P3110	Y3111	F3112	M3113	<b>T3114</b>	H3115	S3118	<b>L3119</b>					
L2916	L2917	V2918	R2919	<b>K2920</b>	L2921	L2922	<b>L2925</b>	E2929	A2930	<b>E2931</b>	E2932	L2933	D2934	D2935	L2937	L2938	K2939	L2940	Q2941	N2942	Y2949	E2950	<b>H2951</b>	L2952	A2953	<b>G2954</b>	F2955	H2956	<b>G2957</b>	Y2958	P2959	N2960	L2961	<b>C2962</b>	P2963	E2964	K2965	G2966	<b>D2967</b>	E2968	K2969	<b>V2970</b>	P2971	C2972	C2973	V2974	H2975	G2976	M2977	C2978	C2979	Q3051	F2980	P2981	<b>K3054</b>	<b>W2983</b>				

## 4 Data and refinement statistics i

Property	Value	Source
Space group	I 21 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	251.02Å 251.02Å 251.02Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 4.00 29.58 – 4.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) (30.00-4.00) 100.0 (29.58-4.00)	Depositor EDS
$R_{merge}$	0.18	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.73 (at 3.98Å)	Xtrriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.271 , 0.293 0.269 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	122.7	Xtrriage
Anisotropy	0.000	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.24 , 149.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	0.055 for -l,-k,-h	Xtrriage
$F_o, F_c$ correlation	0.86	EDS
Total number of atoms	7994	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	142.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.70% 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](#)

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

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	0.26	0/4110	0.46	0/5575
1	B	0.26	0/4110	0.46	0/5575
All	All	0.26	0/8220	0.46	0/11150

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 [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	3995	0	3831	510	0
1	B	3995	0	3831	502	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
All	All	7994	0	7662	1000	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 64.

The worst 5 of 1000 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:3318:ILE:HA	1:A:3339:MET:HB2	1.35	1.08
1:B:3318:ILE:HA	1:B:3339:MET:HB2	1.36	1.06
1:A:3061:ILE:HG23	1:A:3062:PHE:H	1.24	1.03
1:B:3061:ILE:HG23	1:B:3062:PHE:H	1.23	1.00
1:B:3326:PRO:HB3	1:B:3402:HIS:HB3	1.44	0.96

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	489/491 (100%)	341 (70%)	102 (21%)	46 (9%)	0	11
1	B	489/491 (100%)	342 (70%)	101 (21%)	46 (9%)	0	11
All	All	978/982 (100%)	683 (70%)	203 (21%)	92 (9%)	0	11

5 of 92 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2962	CYS
1	A	2970	TYR
1	A	3061	ILE
1	A	3139	ALA
1	A	3158	SER

### 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	437/437 (100%)	376 (86%)	61 (14%)	3	20
1	B	437/437 (100%)	376 (86%)	61 (14%)	3	20
All	All	874/874 (100%)	752 (86%)	122 (14%)	3	20

5 of 122 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	3387	THR
1	B	3338	PHE
1	B	3019	PHE
1	B	3332	ARG
1	B	3387	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	3161	ASN
1	B	3351	ASN
1	B	3168	ASN
1	B	3272	ASN
1	B	3397	ASN

### 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](#)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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 [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	491/491 (100%)	-0.30	6 (1%) 79 70	57, 135, 209, 279	0
1	B	491/491 (100%)	-0.32	5 (1%) 82 74	58, 134, 208, 279	0
All	All	982/982 (100%)	-0.31	11 (1%) 80 72	57, 135, 209, 279	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	3055	PHE	3.8
1	B	3054	LYS	3.6
1	B	3055	PHE	3.2
1	B	3052	GLN	2.9
1	A	2967	ASP	2.6

### 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( $\text{\AA}^2$ )	Q<0.9
2	CU	A	3408	1/1	0.95	0.13	86,86,86,86	0
2	CU	A	3407	1/1	0.96	0.13	79,79,79,79	0
2	CU	B	3407	1/1	0.98	0.15	79,79,79,79	0
2	CU	B	3408	1/1	0.98	0.17	74,74,74,74	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.