



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:01 AM GMT

PDB ID : 2F7K  
Title : Crystal Structure of Human Pyridoxal Kinase  
Authors : Jiang, T.; Cao, P.; Gong, Y.; Tang, L.  
Deposited on : 2005-12-01  
Resolution : 2.80 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

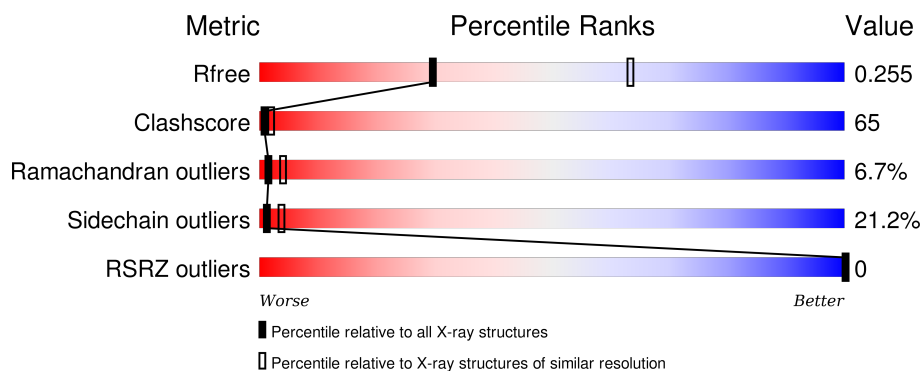
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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	327	 21% 59% 16% ..
1	B	327	 26% 57% 15% ..

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5327 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 Pyridoxal kinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	323	Total	C	N	O	S	0	0	0
			2560	1608	458	478	16			
1	B	323	Total	C	N	O	S	0	0	0
			2560	1608	458	478	16			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-14	SER	-	CLONING ARTIFACT	UNP O00764
A	-13	TYR	-	CLONING ARTIFACT	UNP O00764
A	-12	TYR	-	CLONING ARTIFACT	UNP O00764
A	-11	HIS	-	EXPRESSION TAG	UNP O00764
A	-10	HIS	-	EXPRESSION TAG	UNP O00764
A	-9	HIS	-	EXPRESSION TAG	UNP O00764
A	-8	HIS	-	EXPRESSION TAG	UNP O00764
A	-7	HIS	-	EXPRESSION TAG	UNP O00764
A	-6	HIS	-	EXPRESSION TAG	UNP O00764
A	-5	HIS	-	EXPRESSION TAG	UNP O00764
A	-4	GLU	-	CLONING ARTIFACT	UNP O00764
A	-3	GLY	-	CLONING ARTIFACT	UNP O00764
A	-2	VAL	-	CLONING ARTIFACT	UNP O00764
A	-1	ARG	-	CLONING ARTIFACT	UNP O00764
A	0	THR	-	CLONING ARTIFACT	UNP O00764
B	-14	SER	-	CLONING ARTIFACT	UNP O00764
B	-13	TYR	-	CLONING ARTIFACT	UNP O00764
B	-12	TYR	-	CLONING ARTIFACT	UNP O00764
B	-11	HIS	-	EXPRESSION TAG	UNP O00764
B	-10	HIS	-	EXPRESSION TAG	UNP O00764
B	-9	HIS	-	EXPRESSION TAG	UNP O00764
B	-8	HIS	-	EXPRESSION TAG	UNP O00764
B	-7	HIS	-	EXPRESSION TAG	UNP O00764
B	-6	HIS	-	EXPRESSION TAG	UNP O00764
B	-5	HIS	-	EXPRESSION TAG	UNP O00764

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-4	GLU	-	CLONING ARTIFACT	UNP O00764
B	-3	GLY	-	CLONING ARTIFACT	UNP O00764
B	-2	VAL	-	CLONING ARTIFACT	UNP O00764
B	-1	ARG	-	CLONING ARTIFACT	UNP O00764
B	0	THR	-	CLONING ARTIFACT	UNP O00764

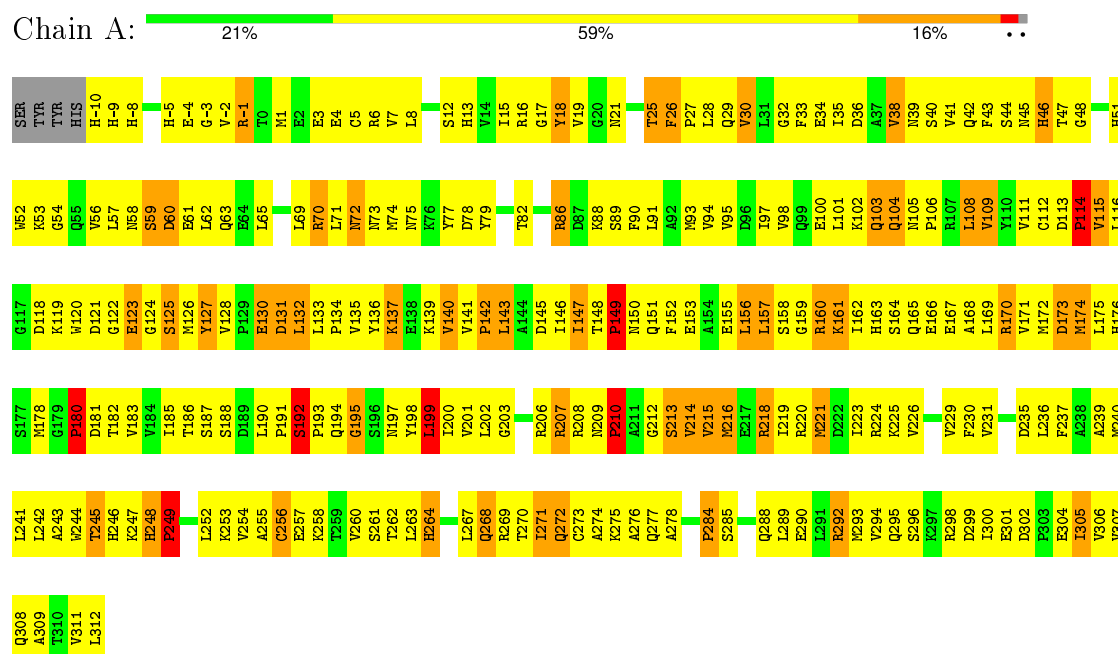
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	96	Total O 96 96	0	0
2	B	111	Total O 111 111	0	0

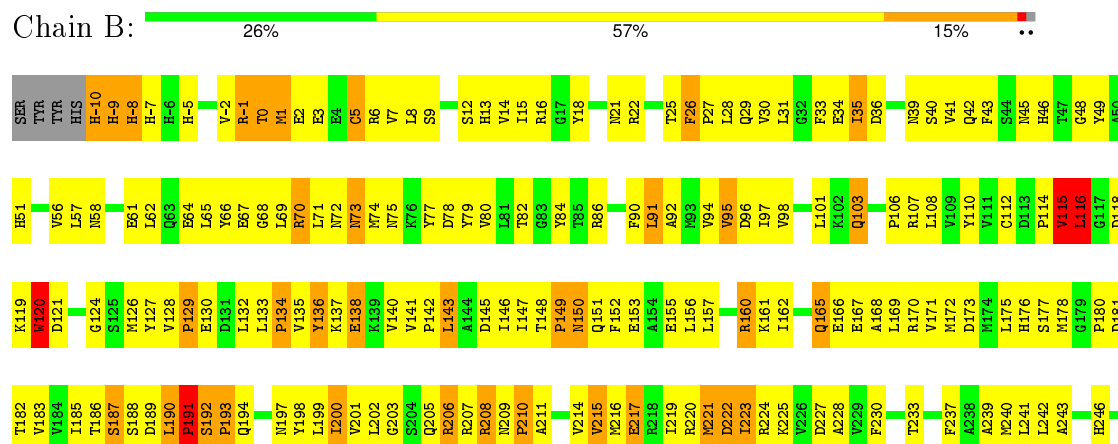
### 3 Residue-property plots

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: Pyridoxal kinase



#### • Molecule 1: Pyridoxal kinase



P249	P250	P251	P252	P253	P254	P255	P256	P257	P258	P259	P260	P261	P262	P263	P264	P265	P266	P267	P268	P269	P270	P271	P272	P273	P274	P275	P276	P277	P278	P279	P280	P281	P282	P283	P284	P285	P286	P287	P288	P289	P290	P291	P292	P293	P294	P295	P296	P297	P298	P299	P300	P301	P302	P303	P304	P305	P306	P307	P308	P309	P310	P311	P312	P313	P314	P315	P316	P317	P318	P319	P320	P321	P322	P323	P324	P325	P326	P327	P328	P329	P330	P331	P332	P333	P334	P335	P336	P337	P338	P339	P340	P341	P342	P343	P344	P345	P346	P347	P348	P349	P350	P351	P352	P353	P354	P355	P356	P357	P358	P359	P360	P361	P362	P363	P364	P365	P366	P367	P368	P369	P370	P371	P372	P373	P374	P375	P376	P377	P378	P379	P380	P381	P382	P383	P384	P385	P386	P387	P388	P389	P390	P391	P392	P393	P394	P395	P396	P397	P398	P399	P400	P401	P402	P403	P404	P405	P406	P407	P408	P409	P410	P411	P412	P413	P414	P415	P416	P417	P418	P419	P420	P421	P422	P423	P424	P425	P426	P427	P428	P429	P430	P431	P432	P433	P434	P435	P436	P437	P438	P439	P440	P441	P442	P443	P444	P445	P446	P447	P448	P449	P450	P451	P452	P453	P454	P455	P456	P457	P458	P459	P460	P461	P462	P463	P464	P465	P466	P467	P468	P469	P470	P471	P472	P473	P474	P475	P476	P477	P478	P479	P480	P481	P482	P483	P484	P485	P486	P487	P488	P489	P490	P491	P492	P493	P494	P495	P496	P497	P498	P499	P500	P501	P502	P503	P504	P505	P506	P507	P508	P509	P510	P511	P512	P513	P514	P515	P516	P517	P518	P519	P520	P521	P522	P523	P524	P525	P526	P527	P528	P529	P530	P531	P532	P533	P534	P535	P536	P537	P538	P539	P540	P541	P542	P543	P544	P545	P546	P547	P548	P549	P550	P551	P552	P553	P554	P555	P556	P557	P558	P559	P560	P561	P562	P563	P564	P565	P566	P567	P568	P569	P570	P571	P572	P573	P574	P575	P576	P577	P578	P579	P580	P581	P582	P583	P584	P585	P586	P587	P588	P589	P590	P591	P592	P593	P594	P595	P596	P597	P598	P599	P600	P601	P602	P603	P604	P605	P606	P607	P608	P609	P610	P611	P612	P613	P614	P615	P616	P617	P618	P619	P620	P621	P622	P623	P624	P625	P626	P627	P628	P629	P630	P631	P632	P633	P634	P635	P636	P637	P638	P639	P640	P641	P642	P643	P644	P645	P646	P647	P648	P649	P650	P651	P652	P653	P654	P655	P656	P657	P658	P659	P660	P661	P662	P663	P664	P665	P666	P667	P668	P669	P670	P671	P672	P673	P674	P675	P676	P677	P678	P679	P680	P681	P682	P683	P684	P685	P686	P687	P688	P689	P690	P691	P692	P693	P694	P695	P696	P697	P698	P699	P700	P701	P702	P703	P704	P705	P706	P707	P708	P709	P710	P711	P712	P713	P714	P715	P716	P717	P718	P719	P720	P721	P722	P723	P724	P725	P726	P727	P728	P729	P730	P731	P732	P733	P734	P735	P736	P737	P738	P739	P740	P741	P742	P743	P744	P745	P746	P747	P748	P749	P750	P751	P752	P753	P754	P755	P756	P757	P758	P759	P760	P761	P762	P763	P764	P765	P766	P767	P768	P769	P770	P771	P772	P773	P774	P775	P776	P777	P778	P779	P780	P781	P782	P783	P784	P785	P786	P787	P788	P789	P790	P791	P792	P793	P794	P795	P796	P797	P798	P799	P800	P801	P802	P803	P804	P805	P806	P807	P808	P809	P810	P811	P812	P813	P814	P815	P816	P817	P818	P819	P820	P821	P822	P823	P824	P825	P826	P827	P828	P829	P830	P831	P832	P833	P834	P835	P836	P837	P838	P839	P840	P841	P842	P843	P844	P845	P846	P847	P848	P849	P850	P851	P852	P853	P854	P855	P856	P857	P858	P859	P860	P861	P862	P863	P864	P865	P866	P867	P868	P869	P870	P871	P872	P873	P874	P875	P876	P877	P878	P879	P880	P881	P882	P883	P884	P885	P886	P887	P888	P889	P890	P891	P892	P893	P894	P895	P896	P897	P898	P899	P900	P901	P902	P903	P904	P905	P906	P907	P908	P909	P910	P911	P912	P913	P914	P915	P916	P917	P918	P919	P920	P921	P922	P923	P924	P925	P926	P927	P928	P929	P930	P931	P932	P933	P934	P935	P936	P937	P938	P939	P940	P941	P942	P943	P944	P945	P946	P947	P948	P949	P950	P951	P952	P953	P954	P955	P956	P957	P958	P959	P960	P961	P962	P963	P964	P965	P966	P967	P968	P969	P970	P971	P972	P973	P974	P975	P976	P977	P978	P979	P980	P981	P982	P983	P984	P985	P986	P987	P988	P989	P990	P991	P992	P993	P994	P995	P996	P997	P998	P999
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	52.50 Å 52.50 Å 301.00 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 2.80 7.99 – 2.80	Depositor EDS
% Data completeness (in resolution range)	81.0 (8.00-2.80) 81.0 (7.99-2.80)	Depositor EDS
$R_{merge}$	0.03	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.94 (at 2.78 Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.228 , 0.271 0.213 , 0.255	Depositor DCC
$R_{free}$ test set	707 reflections (4.58%)	DCC
Wilson B-factor (Å <sup>2</sup> )	68.3	Xtriage
Anisotropy	0.089	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 71.6	EDS
Estimated twinning fraction	0.190 for h,-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.52$ , $\langle L^2 \rangle = 0.36$	Xtriage
Outliers	0 of 15454 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	5327	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.72% 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 [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	0.46	0/2612	0.87	10/3540 (0.3%)
1	B	0.45	0/2612	0.83	3/3540 (0.1%)
All	All	0.45	0/5224	0.85	13/7080 (0.2%)

There are no bond length outliers.

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	142	PRO	CA-N-CD	-7.94	100.39	111.50
1	A	149	PRO	CA-N-CD	-7.75	100.66	111.50
1	B	120	TRP	CA-CB-CG	6.54	126.12	113.70
1	A	-3	GLY	N-CA-C	-6.51	96.83	113.10
1	A	192	SER	N-CA-C	-5.85	95.22	111.00

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	2560	0	2551	354	0
1	B	2560	0	2551	334	0
2	A	96	0	0	0	0
2	B	111	0	0	3	0
All	All	5327	0	5102	660	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 65.

The worst 5 of 660 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:199:LEU:HB2	1:A:223:ILE:HB	1.38	1.05
1:A:72:ASN:HD21	1:A:74:MET:HG3	1.22	1.04
1:A:221:MET:HG2	1:A:309:ALA:HA	1.39	1.02
1:A:147:ILE:H	1:A:147:ILE:CD1	1.74	1.00
1:A:147:ILE:H	1:A:147:ILE:HD13	1.24	1.00

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	321/327 (98%)	229 (71%)	65 (20%)	27 (8%)	1	2
1	B	321/327 (98%)	240 (75%)	65 (20%)	16 (5%)	3	8
All	All	642/654 (98%)	469 (73%)	130 (20%)	43 (7%)	1	4

5 of 43 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	-9	HIS
1	A	59	SER
1	A	109	VAL
1	A	195	GLY
1	A	210	PRO

### 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	285/289 (99%)	226 (79%)	59 (21%)	1	4
1	B	285/289 (99%)	223 (78%)	62 (22%)	1	3
All	All	570/578 (99%)	449 (79%)	121 (21%)	1	4

5 of 121 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	292	ARG
1	B	5	CYS
1	B	273	CYS
1	A	305	ILE
1	B	-8	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	268	GLN
1	B	-10	HIS
1	B	209	ASN
1	A	288	GLN
1	A	308	GLN

### 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	323/327 (98%)	-0.27	0 100 100	36, 44, 50, 56	0
1	B	323/327 (98%)	-0.27	0 100 100	35, 44, 50, 55	0
All	All	646/654 (98%)	-0.27	0 100 100	35, 44, 50, 56	0

There are no RSRZ outliers to report.

### 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.