



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:02 PM GMT

PDB ID : 4GUO  
Title : structure of p73 DNA binding domain complex with 12 bp DNA  
Authors : Ethayathulla, A.S; Viadiu, H.  
Deposited on : 2012-08-29  
Resolution : 3.19 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

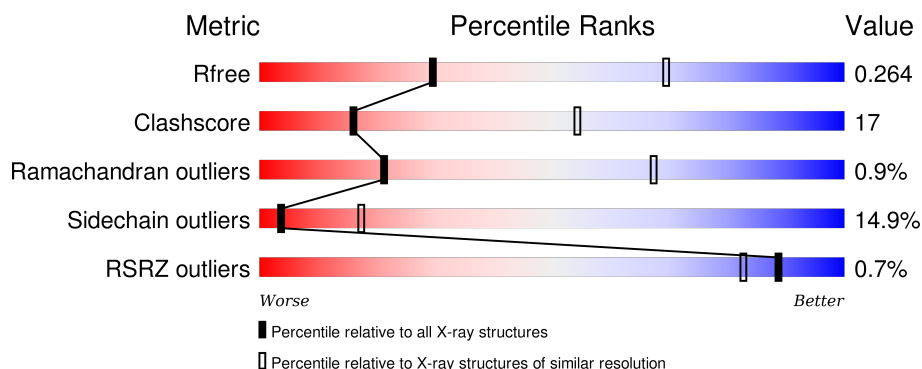
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.19 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	1124 (3.24-3.16)
Clashscore	102246	1024 (3.22-3.18)
Ramachandran outliers	100387	1004 (3.22-3.18)
Sidechain outliers	100360	1003 (3.22-3.18)
RSRZ outliers	91569	1129 (3.24-3.16)

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	210	 60% 31% . . .
1	B	210	 55% 36% . 6%
1	C	210	 63% 28% 5% .
1	D	210	 60% 29% 5% 6%
1	I	210	 59% 30% 5% 6%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	J	210	
1	K	210	
1	L	210	
2	E	12	
2	G	12	
2	M	12	
2	O	12	
3	F	12	
3	H	12	
3	N	12	
3	P	12	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	ZN	C	401	-	-	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 14607 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 Tumor protein p73.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	201	Total	C	N	O	S	0	0	0
			1571	982	283	295	11			
1	B	198	Total	C	N	O	S	0	0	0
			1553	975	279	288	11			
1	C	201	Total	C	N	O	S	0	0	0
			1586	993	285	297	11			
1	D	198	Total	C	N	O	S	0	0	0
			1551	971	279	290	11			
1	I	198	Total	C	N	O	S	0	0	0
			1556	973	280	292	11			
1	J	199	Total	C	N	O	S	0	0	0
			1561	976	281	293	11			
1	K	201	Total	C	N	O	S	0	0	0
			1576	988	282	295	11			
1	L	201	Total	C	N	O	S	0	0	0
			1581	991	284	295	11			

There are 96 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	103	MET	-	INITIATING METHIONINE	UNP O15350
A	104	GLY	-	EXPRESSION TAG	UNP O15350
A	105	HIS	-	EXPRESSION TAG	UNP O15350
A	106	HIS	-	EXPRESSION TAG	UNP O15350
A	107	HIS	-	EXPRESSION TAG	UNP O15350
A	108	HIS	-	EXPRESSION TAG	UNP O15350
A	109	HIS	-	EXPRESSION TAG	UNP O15350
A	110	HIS	-	EXPRESSION TAG	UNP O15350
A	111	HIS	-	EXPRESSION TAG	UNP O15350
A	112	HIS	-	EXPRESSION TAG	UNP O15350
A	113	GLU	-	EXPRESSION TAG	UNP O15350
A	114	PHE	-	EXPRESSION TAG	UNP O15350
B	103	MET	-	INITIATING METHIONINE	UNP O15350

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	104	GLY	-	EXPRESSION TAG	UNP O15350
B	105	HIS	-	EXPRESSION TAG	UNP O15350
B	106	HIS	-	EXPRESSION TAG	UNP O15350
B	107	HIS	-	EXPRESSION TAG	UNP O15350
B	108	HIS	-	EXPRESSION TAG	UNP O15350
B	109	HIS	-	EXPRESSION TAG	UNP O15350
B	110	HIS	-	EXPRESSION TAG	UNP O15350
B	111	HIS	-	EXPRESSION TAG	UNP O15350
B	112	HIS	-	EXPRESSION TAG	UNP O15350
B	113	GLU	-	EXPRESSION TAG	UNP O15350
B	114	PHE	-	EXPRESSION TAG	UNP O15350
C	103	MET	-	INITIATING METHIONINE	UNP O15350
C	104	GLY	-	EXPRESSION TAG	UNP O15350
C	105	HIS	-	EXPRESSION TAG	UNP O15350
C	106	HIS	-	EXPRESSION TAG	UNP O15350
C	107	HIS	-	EXPRESSION TAG	UNP O15350
C	108	HIS	-	EXPRESSION TAG	UNP O15350
C	109	HIS	-	EXPRESSION TAG	UNP O15350
C	110	HIS	-	EXPRESSION TAG	UNP O15350
C	111	HIS	-	EXPRESSION TAG	UNP O15350
C	112	HIS	-	EXPRESSION TAG	UNP O15350
C	113	GLU	-	EXPRESSION TAG	UNP O15350
C	114	PHE	-	EXPRESSION TAG	UNP O15350
D	103	MET	-	INITIATING METHIONINE	UNP O15350
D	104	GLY	-	EXPRESSION TAG	UNP O15350
D	105	HIS	-	EXPRESSION TAG	UNP O15350
D	106	HIS	-	EXPRESSION TAG	UNP O15350
D	107	HIS	-	EXPRESSION TAG	UNP O15350
D	108	HIS	-	EXPRESSION TAG	UNP O15350
D	109	HIS	-	EXPRESSION TAG	UNP O15350
D	110	HIS	-	EXPRESSION TAG	UNP O15350
D	111	HIS	-	EXPRESSION TAG	UNP O15350
D	112	HIS	-	EXPRESSION TAG	UNP O15350
D	113	GLU	-	EXPRESSION TAG	UNP O15350
D	114	PHE	-	EXPRESSION TAG	UNP O15350
I	103	MET	-	INITIATING METHIONINE	UNP O15350
I	104	GLY	-	EXPRESSION TAG	UNP O15350
I	105	HIS	-	EXPRESSION TAG	UNP O15350
I	106	HIS	-	EXPRESSION TAG	UNP O15350
I	107	HIS	-	EXPRESSION TAG	UNP O15350
I	108	HIS	-	EXPRESSION TAG	UNP O15350
I	109	HIS	-	EXPRESSION TAG	UNP O15350

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
I	110	HIS	-	EXPRESSION TAG	UNP 015350
I	111	HIS	-	EXPRESSION TAG	UNP 015350
I	112	HIS	-	EXPRESSION TAG	UNP 015350
I	113	GLU	-	EXPRESSION TAG	UNP 015350
I	114	PHE	-	EXPRESSION TAG	UNP 015350
J	103	MET	-	INITIATING METHIONINE	UNP 015350
J	104	GLY	-	EXPRESSION TAG	UNP 015350
J	105	HIS	-	EXPRESSION TAG	UNP 015350
J	106	HIS	-	EXPRESSION TAG	UNP 015350
J	107	HIS	-	EXPRESSION TAG	UNP 015350
J	108	HIS	-	EXPRESSION TAG	UNP 015350
J	109	HIS	-	EXPRESSION TAG	UNP 015350
J	110	HIS	-	EXPRESSION TAG	UNP 015350
J	111	HIS	-	EXPRESSION TAG	UNP 015350
J	112	HIS	-	EXPRESSION TAG	UNP 015350
J	113	GLU	-	EXPRESSION TAG	UNP 015350
J	114	PHE	-	EXPRESSION TAG	UNP 015350
K	103	MET	-	INITIATING METHIONINE	UNP 015350
K	104	GLY	-	EXPRESSION TAG	UNP 015350
K	105	HIS	-	EXPRESSION TAG	UNP 015350
K	106	HIS	-	EXPRESSION TAG	UNP 015350
K	107	HIS	-	EXPRESSION TAG	UNP 015350
K	108	HIS	-	EXPRESSION TAG	UNP 015350
K	109	HIS	-	EXPRESSION TAG	UNP 015350
K	110	HIS	-	EXPRESSION TAG	UNP 015350
K	111	HIS	-	EXPRESSION TAG	UNP 015350
K	112	HIS	-	EXPRESSION TAG	UNP 015350
K	113	GLU	-	EXPRESSION TAG	UNP 015350
K	114	PHE	-	EXPRESSION TAG	UNP 015350
L	103	MET	-	INITIATING METHIONINE	UNP 015350
L	104	GLY	-	EXPRESSION TAG	UNP 015350
L	105	HIS	-	EXPRESSION TAG	UNP 015350
L	106	HIS	-	EXPRESSION TAG	UNP 015350
L	107	HIS	-	EXPRESSION TAG	UNP 015350
L	108	HIS	-	EXPRESSION TAG	UNP 015350
L	109	HIS	-	EXPRESSION TAG	UNP 015350
L	110	HIS	-	EXPRESSION TAG	UNP 015350
L	111	HIS	-	EXPRESSION TAG	UNP 015350
L	112	HIS	-	EXPRESSION TAG	UNP 015350
L	113	GLU	-	EXPRESSION TAG	UNP 015350
L	114	PHE	-	EXPRESSION TAG	UNP 015350

- Molecule 2 is a DNA chain called DNA (5'-D(\*CP\*GP\*GP\*GP\*CP\*AP\*AP\*GP\*CP\*CP\*

CP\*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	12	Total	C	N	O	P	0	0	0
			244	115	50	68	11			
2	G	11	Total	C	N	O	P	0	0	0
			228	106	47	64	11			
2	M	12	Total	C	N	O	P	0	0	0
			244	115	50	68	11			
2	O	12	Total	C	N	O	P	0	0	0
			244	115	50	68	11			

- Molecule 3 is a DNA chain called DNA (5'-D(\*CP\*GP\*GP\*GP\*CP\*TP\*TP\*GP\*CP\*CP\*CP\*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	F	12	Total	C	N	O	P	0	0	0
			242	115	44	72	11			
3	H	12	Total	C	N	O	P	0	0	0
			242	115	44	72	11			
3	N	12	Total	C	N	O	P	0	0	0
			242	115	44	72	11			
3	P	12	Total	C	N	O	P	0	0	0
			242	115	44	72	11			

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	J	1	Total	Zn	0	0
			1	1		
4	D	1	Total	Zn	0	0
			1	1		
4	K	1	Total	Zn	0	0
			1	1		
4	B	1	Total	Zn	0	0
			1	1		
4	I	1	Total	Zn	0	0
			1	1		
4	C	1	Total	Zn	0	0
			1	1		
4	A	1	Total	Zn	0	0
			1	1		
4	L	1	Total	Zn	0	0
			1	1		

- Molecule 5 is water.

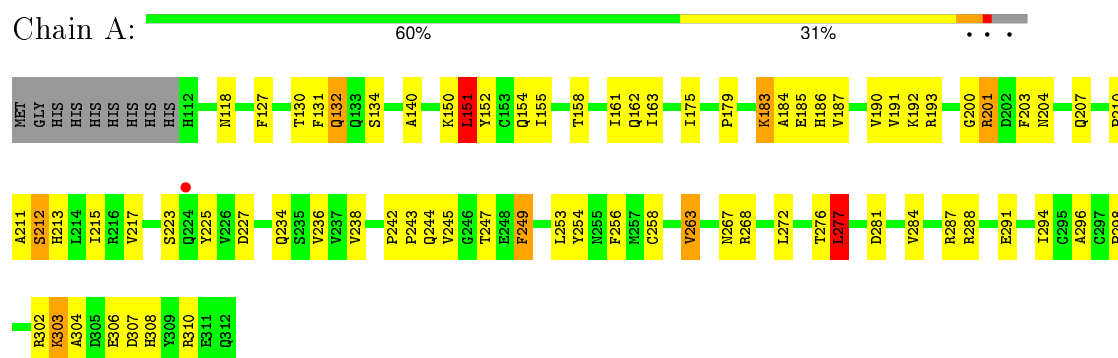
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	16	Total 16	O 16	0	0
5	B	15	Total 15	O 15	0	0
5	C	21	Total 21	O 21	0	0
5	D	18	Total 18	O 18	0	0
5	I	14	Total 14	O 14	0	0
5	J	10	Total 10	O 10	0	0
5	K	18	Total 18	O 18	0	0
5	L	20	Total 20	O 20	0	0
5	E	1	Total 1	O 1	0	0
5	H	1	Total 1	O 1	0	0
5	N	1	Total 1	O 1	0	0
5	P	1	Total 1	O 1	0	0



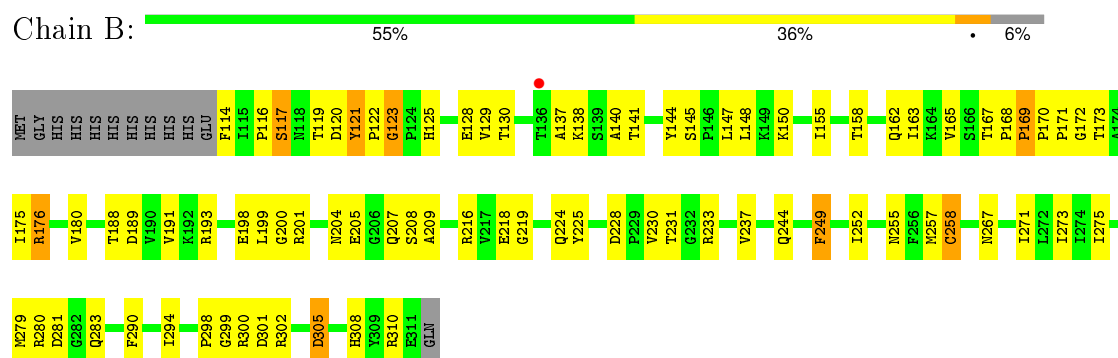
### 3 Residue-property plots [i](#)

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

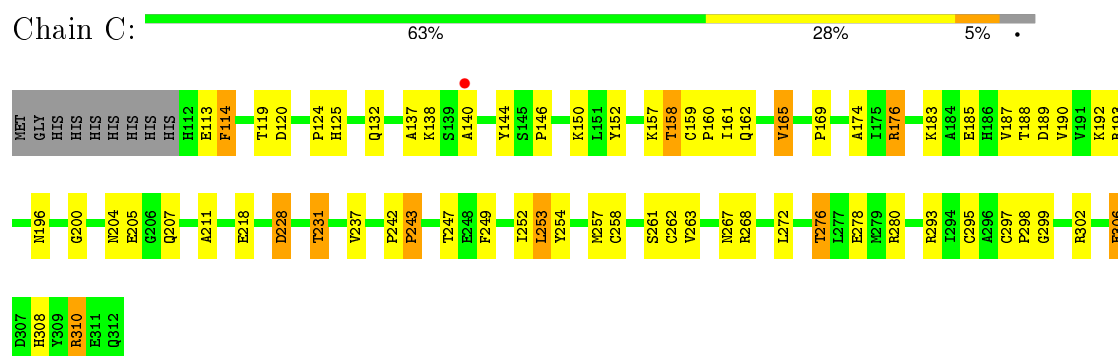
#### • Molecule 1: Tumor protein p73



#### • Molecule 1: Tumor protein p73



#### • Molecule 1: Tumor protein p73

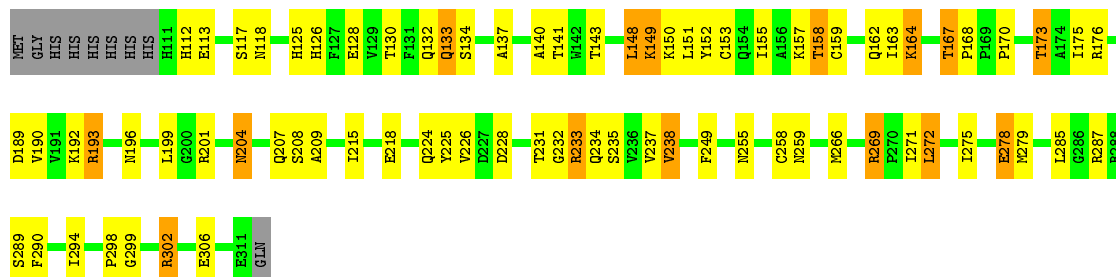


Y181	E185	E186	V190	V191	K192	K193	E198	F203	N204	E205	G206	G207	S208	A209	P210	A211	L214	L215	R216	V217	E218	S223	V226	T231	G232	R233	G234	S235	V236	V237	Y240	E241	F249	T250	L253	F256	I259	S260	V263	L267	R268	P269	L272	L273	L274	L275	L276	L277	L278	L279	L280	L281	L282	L283	L284	L285	L286	L287	L288	L289	L290	L291	L292	L293	L294	L295	L296	L297	L298	L299	L300	L301	L302	L303	L304	L305	L306	L307	L308	L309	L310	L311	L312	L313	L314	L315	L316	L317	L318	L319	L320	L321	L322	L323	L324	L325	L326	L327	L328	L329	L330	L331	L332	L333	L334	L335	L336	L337	L338	L339	L340	L341	L342	L343	L344	L345	L346	L347	L348	L349	L350	L351	L352	L353	L354	L355	L356	L357	L358	L359	L360	L361	L362	L363	L364	L365	L366	L367	L368	L369	L370	L371	L372	L373	L374	L375	L376	L377	L378	L379	L380	L381	L382	L383	L384	L385	L386	L387	L388	L389	L390	L391	L392	L393	L394	L395	L396	L397	L398	L399	L400	L401	L402	L403	L404	L405	L406	L407	L408	L409	L410	L411	L412	L413	L414	L415	L416	L417	L418	L419	L420	L421	L422	L423	L424	L425	L426	L427	L428	L429	L430	L431	L432	L433	L434	L435	L436	L437	L438	L439	L440	L441	L442	L443	L444	L445	L446	L447	L448	L449	L450	L451	L452	L453	L454	L455	L456	L457	L458	L459	L460	L461	L462	L463	L464	L465	L466	L467	L468	L469	L470	L471	L472	L473	L474	L475	L476	L477	L478	L479	L480	L481	L482	L483	L484	L485	L486	L487	L488	L489	L490	L491	L492	L493	L494	L495	L496	L497	L498	L499	L500	L501	L502	L503	L504	L505	L506	L507	L508	L509	L510	L511	L512	L513	L514	L515	L516	L517	L518	L519	L520	L521	L522	L523	L524	L525	L526	L527	L528	L529	L530	L531	L532	L533	L534	L535	L536	L537	L538	L539	L540	L541	L542	L543	L544	L545	L546	L547	L548	L549	L550	L551	L552	L553	L554	L555	L556	L557	L558	L559	L560	L561	L562	L563	L564	L565	L566	L567	L568	L569	L570	L571	L572	L573	L574	L575	L576	L577	L578	L579	L580	L581	L582	L583	L584	L585	L586	L587	L588	L589	L590	L591	L592	L593	L594	L595	L596	L597	L598	L599	L600	L601	L602	L603	L604	L605	L606	L607	L608	L609	L610	L611	L612	L613	L614	L615	L616	L617	L618	L619	L620	L621	L622	L623	L624	L625	L626	L627	L628	L629	L630	L631	L632	L633	L634	L635	L636	L637	L638	L639	L640	L641	L642	L643	L644	L645	L646	L647	L648	L649	L650	L651	L652	L653	L654	L655	L656	L657	L658	L659	L660	L661	L662	L663	L664	L665	L666	L667	L668	L669	L670	L671	L672	L673	L674	L675	L676	L677	L678	L679	L680	L681	L682	L683	L684	L685	L686	L687	L688	L689	L690	L691	L692	L693	L694	L695	L696	L697	L698	L699	L700	L701	L702	L703	L704	L705	L706	L707	L708	L709	L710	L711	L712	L713	L714	L715	L716	L717	L718	L719	L720	L721	L722	L723	L724	L725	L726	L727	L728	L729	L730	L731	L732	L733	L734	L735	L736	L737	L738	L739
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------



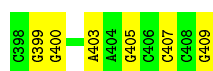
- Molecule 1: Tumor protein p73

Chain L: 59% 30% 7%



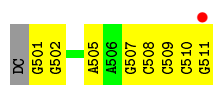
- Molecule 2: DNA (5'-D(\*CP\*GP\*GP\*GP\*CP\*AP\*AP\*GP\*CP\*CP\*CP\*G)-3')

Chain E: 50% 50%



- Molecule 2: DNA (5'-D(\*CP\*GP\*GP\*GP\*CP\*AP\*AP\*GP\*CP\*CP\*CP\*G)-3')

Chain G: 8% 25% 67% 8%



- Molecule 2: DNA (5'-D(\*CP\*GP\*GP\*GP\*CP\*AP\*AP\*GP\*CP\*CP\*CP\*G)-3')

Chain M: 83% 8% 8%



- Molecule 2: DNA (5'-D(\*CP\*GP\*GP\*GP\*CP\*AP\*AP\*GP\*CP\*CP\*CP\*G)-3')

Chain O: 75% 25%



- Molecule 3: DNA (5'-D(\*CP\*GP\*GP\*GP\*CP\*TP\*TP\*GP\*CP\*CP\*CP\*G)-3')

Chain F: 42% 50% 8%



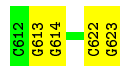
- Molecule 3: DNA (5'-D(\*CP\*GP\*GP\*GP\*CP\*TP\*TP\*GP\*CP\*CP\*CP\*G)-3')

Chain H:  8% 58% 42%



- Molecule 3: DNA (5'-D(\*CP\*GP\*GP\*GP\*CP\*TP\*TP\*GP\*CP\*CP\*CP\*G)-3')

Chain N:  67% 33%



- Molecule 3: DNA (5'-D(\*CP\*GP\*GP\*GP\*CP\*TP\*TP\*GP\*CP\*CP\*CP\*G)-3')

Chain P:  58% 33% 8%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	82.25Å 104.33Å 123.32Å 90.00° 96.38° 90.00°	Depositor
Resolution (Å)	44.01 – 3.19 43.97 – 3.19	Depositor EDS
% Data completeness (in resolution range)	99.1 (44.01-3.19) 99.1 (43.97-3.19)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.64 (at 3.19Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.208 , 0.270 0.202 , 0.264	Depositor DCC
$R_{free}$ test set	1738 reflections (5.32%)	DCC
Wilson B-factor (Å <sup>2</sup> )	80.8	Xtriage
Anisotropy	0.441	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.25 , 42.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 34474 reflections (0.003%)	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	14607	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	95.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.16% 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 ⓘ

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

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	0.55	0/1610	0.85	2/2188 (0.1%)
1	B	0.57	0/1593	0.93	4/2166 (0.2%)
1	C	0.59	0/1627	0.86	1/2210 (0.0%)
1	D	0.57	0/1590	0.91	5/2162 (0.2%)
1	I	0.57	0/1595	0.90	2/2167 (0.1%)
1	J	0.57	0/1600	0.89	2/2174 (0.1%)
1	K	0.59	0/1616	0.95	5/2197 (0.2%)
1	L	0.59	0/1622	0.89	0/2205
2	E	0.67	1/274 (0.4%)	0.77	0/421
2	G	0.33	0/256	0.66	0/393
2	M	0.76	1/274 (0.4%)	0.87	1/421 (0.2%)
2	O	0.51	0/274	0.82	1/421 (0.2%)
3	F	0.58	0/270	0.88	1/415 (0.2%)
3	H	0.47	0/270	1.00	1/415 (0.2%)
3	N	0.58	0/270	0.92	0/415
3	P	0.61	0/270	1.01	1/415 (0.2%)
All	All	0.57	2/15011 (0.0%)	0.89	26/20785 (0.1%)

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	K	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	M	600	DC	O3'-P	-6.87	1.52	1.61
2	E	407	DC	O3'-P	-5.18	1.54	1.61

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	K	216	ARG	NE-CZ-NH2	7.74	124.17	120.30
1	K	288	ARG	NE-CZ-NH2	7.01	123.81	120.30
1	A	277	LEU	CA-CB-CG	6.93	131.25	115.30
1	D	228	ASP	CB-CG-OD2	6.81	124.43	118.30
3	P	716	DC	O5'-P-OP2	-6.29	100.04	105.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	K	283	GLN	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	1571	0	1535	66	0
1	B	1553	0	1526	54	0
1	C	1586	0	1551	45	0
1	D	1551	0	1523	54	0
1	I	1556	0	1529	51	0
1	J	1561	0	1531	65	0
1	K	1576	0	1540	68	0
1	L	1581	0	1545	58	0
2	E	244	0	134	6	0
2	G	228	0	122	7	0
2	M	244	0	134	1	0
2	O	244	0	134	2	0
3	F	242	0	136	16	0
3	H	242	0	136	4	0
3	N	242	0	136	3	0
3	P	242	0	136	4	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	I	1	0	0	0	0
4	J	1	0	0	0	0
4	K	1	0	0	0	0
4	L	1	0	0	0	0
5	A	16	0	0	2	0
5	B	15	0	0	1	0
5	C	21	0	0	4	0
5	D	18	0	0	0	0
5	E	1	0	0	1	0
5	H	1	0	0	0	0
5	I	14	0	0	5	0
5	J	10	0	0	2	0
5	K	18	0	0	0	0
5	L	20	0	0	2	0
5	N	1	0	0	0	0
5	P	1	0	0	0	0
All	All	14607	0	13348	471	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 17.

The worst 5 of 471 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:L:233:ARG:O	1:L:233:ARG:HD3	1.47	1.14
1:J:255:ASN:HB3	1:J:257:MET:HE1	1.30	1.11
1:L:279:MET:HG3	1:L:285:LEU:HD11	1.32	1.09
1:J:255:ASN:HB3	1:J:257:MET:CE	1.84	1.08
1:K:169:PRO:HB2	1:K:170:PRO:HD2	1.37	1.07

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	199/210 (95%)	189 (95%)	10 (5%)	0	100	100
1	B	196/210 (93%)	179 (91%)	15 (8%)	2 (1%)	19	65
1	C	199/210 (95%)	188 (94%)	9 (4%)	2 (1%)	19	65
1	D	196/210 (93%)	187 (95%)	6 (3%)	3 (2%)	13	55
1	I	196/210 (93%)	188 (96%)	7 (4%)	1 (0%)	34	78
1	J	197/210 (94%)	188 (95%)	8 (4%)	1 (0%)	34	78
1	K	199/210 (95%)	184 (92%)	10 (5%)	5 (2%)	7	41
1	L	199/210 (95%)	187 (94%)	11 (6%)	1 (0%)	34	78
All	All	1581/1680 (94%)	1490 (94%)	76 (5%)	15 (1%)	21	67

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	123	GLY
1	B	169	PRO
1	C	243	PRO
1	D	116	PRO
1	J	122	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	175/186 (94%)	157 (90%)	18 (10%)	9	36
1	B	174/186 (94%)	150 (86%)	24 (14%)	4	21
1	C	178/186 (96%)	155 (87%)	23 (13%)	5	24
1	D	174/186 (94%)	148 (85%)	26 (15%)	4	17
1	I	175/186 (94%)	151 (86%)	24 (14%)	4	21
1	J	175/186 (94%)	144 (82%)	31 (18%)	2	11

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	K	176/186 (95%)	140 (80%)	36 (20%)	1	7
1	L	177/186 (95%)	150 (85%)	27 (15%)	3	17
All	All	1404/1488 (94%)	1195 (85%)	209 (15%)	4	17

5 of 209 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	I	147	LEU
1	J	117	SER
1	L	167	THR
1	I	183	LYS
1	I	260	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	255	ASN
1	I	133	GLN
1	L	213	HIS
1	D	267	ASN
1	I	154	GLN

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 8 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 [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	201/210 (95%)	-0.21	1 (0%) 91 87	59, 98, 147, 172	1 (0%)
1	B	198/210 (94%)	-0.36	1 (0%) 91 87	60, 92, 126, 145	2 (1%)
1	C	201/210 (95%)	-0.40	1 (0%) 91 87	55, 84, 126, 141	0
1	D	198/210 (94%)	-0.37	0 100 100	63, 89, 132, 160	1 (0%)
1	I	198/210 (94%)	-0.31	1 (0%) 91 87	64, 87, 129, 156	0
1	J	199/210 (94%)	-0.00	3 (1%) 76 63	75, 111, 158, 195	0
1	K	201/210 (95%)	-0.24	3 (1%) 76 63	63, 87, 144, 164	0
1	L	201/210 (95%)	-0.48	0 100 100	48, 73, 108, 138	0
2	E	12/12 (100%)	-0.15	0 100 100	73, 100, 152, 178	0
2	G	11/12 (91%)	-0.01	1 (9%) 11 6	85, 106, 135, 142	0
2	M	12/12 (100%)	-0.57	0 100 100	72, 90, 103, 106	0
2	O	12/12 (100%)	-0.60	0 100 100	92, 106, 140, 146	0
3	F	12/12 (100%)	-0.28	0 100 100	69, 102, 138, 158	0
3	H	12/12 (100%)	0.29	1 (8%) 14 7	79, 115, 157, 170	0
3	N	12/12 (100%)	-0.65	0 100 100	72, 84, 105, 117	0
3	P	12/12 (100%)	-0.33	0 100 100	80, 99, 150, 160	0
All	All	1692/1776 (95%)	-0.30	12 (0%) 89 83	48, 90, 141, 195	4 (0%)

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	J	286	GLY	5.4
3	H	523	DG	4.6
1	K	137	ALA	4.6
1	J	137	ALA	4.4
1	I	139	SER	3.9

## 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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	ZN	C	401	1/1	0.99	0.18	3.70	59,59,59,59	0
4	ZN	B	401	1/1	0.99	0.16	0.18	63,63,63,63	0
4	ZN	L	401	1/1	0.99	0.11	-1.31	76,76,76,76	0
4	ZN	J	401	1/1	0.99	0.12	-1.63	84,84,84,84	0
4	ZN	A	401	1/1	0.98	0.11	-2.36	87,87,87,87	0
4	ZN	K	401	1/1	0.99	0.07	-2.65	89,89,89,89	0
4	ZN	D	401	1/1	0.99	0.14	-	71,71,71,71	0
4	ZN	I	401	1/1	0.96	0.13	-	98,98,98,98	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.