



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 27, 2026 – 05:19 PM EDT

PDB ID : 8PIA / pdb_00008pia
Title : DNA binding domain of HNF-1A bound to P2-HNF4A promoter DNA variant (P2 -181G>T)
Authors : Kind, L.; Myllykoski, M.; Raasakka, A.; Kursula, P.
Deposited on : 2023-06-21
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

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 : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

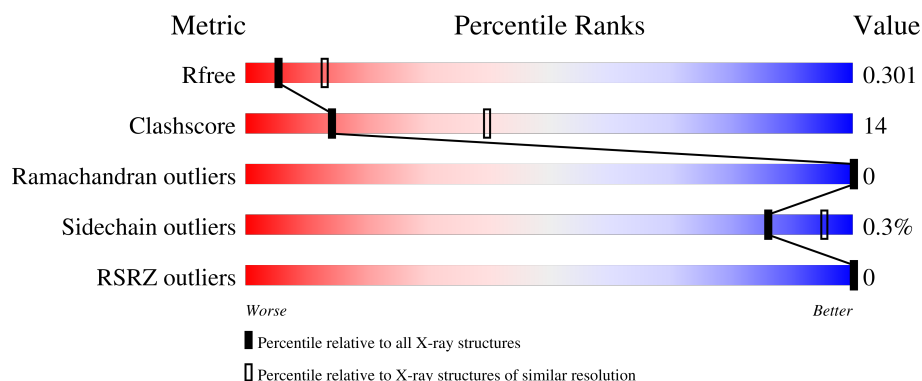
1 Overall quality at a glance

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}	180053	3866 (2.80-2.80)
Clashscore	190562	4276 (2.80-2.80)
Ramachandran outliers	187476	4196 (2.80-2.80)
Sidechain outliers	187428	4198 (2.80-2.80)
RSRZ outliers	180081	3869 (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 of the lower bar indicate the fraction of residues that contain outliers for ≥ 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	E	21	52% 48%
2	F	21	48% 52%
3	A	198	60% 22% 18%
3	B	198	53% 29% • 18%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3545 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 DNA chain called Chains: E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	21	Total	C	N	O	P	0	0	0
			422	206	67	129	20			

- Molecule 2 is a DNA chain called Chains: F.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	21	Total	C	N	O	P	0	0	0
			433	208	86	119	20			

- Molecule 3 is a protein called Hepatocyte nuclear factor 1-alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	163	Total	C	N	O	S	0	0	0
			1342	833	257	248	4			
3	B	163	Total	C	N	O	S	0	0	0
			1342	833	257	248	4			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	82	GLY	-	expression tag	UNP P20823
B	82	GLY	-	expression tag	UNP P20823

- Molecule 4 is GLYCEROL (CCD ID: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 1: Chains: E

A301	C302		T307	A308		T314	A315	A316	C317	G318	T319	A320	T321
------	------	--	------	------	--	------	------	------	------	------	------	------	------

A401	T402	A403	T406	T407	A408	A412	G413	T414	A415	A416	A417	G420	T421
------	------	------	------	------	------	------	------	------	------	------	------	------	------

GLY	PRO	PRO	PRO	LEU	LEU	LYS	GLU	LEU	GLU	ASN	SER	P94	E96	A97	A98	H99	I100	K101	A102	V103	E110	D111	P112	V119	L123	P129	E132	T137	G138	L139	M140	Q141	S142	H143	M154	K155	T156	R159	A160	A161	L162	T163	T164	V165	V166	F177	A180	C181
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

GLN	GLY	GLY	LEU	ILE	GLU	GLU	PRO	THR	GLY	ASP	GLU	LEU	PRO	THR	LYS	LYS	GLY	ARG	R201	W206	L214	Y218	Q221	K222	N223	K226	E228	R229	E230	V233	E240	R244	Q250	V259	R263	V264	V265	A276	PHE	ARG	THR
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----

GLY	PRO	PRO	PRO	ILE	LEU	LYS	GLU	LEU	GLU	ASN	LEU	SER	P94	E95	E96	A97	A98	H99	Q100	K101	A102	V103	P112	V115	M118	V119	L123	Q124	Q125	H126	M127	L128	P129	E132	V133	V134	L139	M140	Q141	S142	H143	L144	S145	Q146	H147	L148	T152	P153	M154	K155	T156	M157	V158
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

R159	R160	A161	L162	Y163	T164	W165	R168	K169	V173	A174	Q175	F177	T178	H179	A180	G180	G181	G182	G183	G184	L185	L186	L187	L188	L189	L190	L191	L192	L193	L194	L195	L196	L197	L198	L199	L200	L201	L202	L203	L204	L205	L206	L207	L208	L209	L210	L211	L212	L213	L214	L215	L216	L217	L218	L219	L220	L221	L222	L223	L224	L225	L226	L227	L228	L229	L230	L231	L232	L233	L234	L235	L236	L237	L238	L239	L240	L241	L242	L243	L244	L245	L246	L247	L248	L249	L250	L251	L252	L253	L254	L255	L256	L257	L258	L259	L260	L261	L262	L263	L264	L265	L266	L267	L268	L269	L270	L271	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	L67
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----

R271	+	A276	PHE
R272		ARG	
			HIS

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	48.70Å 48.75Å 215.81Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.55 – 2.80 47.55 – 2.80	Depositor EDS
% Data completeness (in resolution range)	91.3 (47.55-2.80) 76.9 (47.55-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.29 (at 2.81Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.245 , 0.300 0.245 , 0.301	Depositor DCC
R_{free} test set	613 reflections (2.61%)	wwPDB-VP
Wilson B-factor (Å ²)	64.8	Xtriage
Anisotropy	0.456	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.24 , 96.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.428 for k,h,-l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	3545	wwPDB-VP
Average B, all atoms (Å ²)	99.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.33% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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

5.1 Standard geometry

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

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	E	0.21	0/470	0.45	0/723
2	F	0.18	0/488	0.35	0/752
3	A	0.15	0/1369	0.34	0/1846
3	B	0.15	0/1369	0.37	0/1846
All	All	0.17	0/3696	0.37	0/5167

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	E	422	0	243	7	0
2	F	433	0	238	10	0
3	A	1342	0	1321	31	0
3	B	1342	0	1321	53	0
4	B	6	0	8	1	0
All	All	3545	0	3131	95	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 14.

The worst 5 of 95 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)
3:B:229:ARG:HH21	3:B:265:TYR:HB2	1.38	0.87
3:A:229:ARG:HH21	3:A:265:TYR:HB2	1.40	0.84
3:A:156:THR:HA	3:A:159:ARG:HD3	1.59	0.84
3:B:156:THR:HA	3:B:159:ARG:HD3	1.63	0.81
3:A:226:LYS:HA	3:A:229:ARG:HD3	1.70	0.74

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	
3	A	159/198 (80%)	157 (99%)	2 (1%)	0	100	100
3	B	159/198 (80%)	155 (98%)	4 (2%)	0	100	100
All	All	318/396 (80%)	312 (98%)	6 (2%)	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	
3	A	143/172 (83%)	143 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	B	143/172 (83%)	142 (99%)	1 (1%)	76	91
All	All	286/344 (83%)	285 (100%)	1 (0%)	86	95

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	B	250	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	100	GLN
3	B	127	ASN
3	B	141	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	GOL	B	301	-	5,5,5	0.92	0	5,5,5	1.10	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	B	301	-	-	0/4/4/4	-

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	301	GOL	1	0

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, 95th 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 ²)	Q<0.9
1	E	21/21 (100%)	-0.60	0 100 100	83, 100, 127, 139	0
2	F	21/21 (100%)	-0.47	0 100 100	82, 105, 142, 158	0
3	A	163/198 (82%)	-0.46	0 100 100	68, 93, 121, 139	0
3	B	163/198 (82%)	-0.61	0 100 100	64, 93, 129, 141	0
All	All	368/438 (84%)	-0.53	0 100 100	64, 95, 129, 158	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 oligosaccharides 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, 95th 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(Å ²)	Q<0.9
4	GOL	B	301	6/6	0.96	0.07	67,85,90,90	0

6.5 Other polymers [i](#)

There are no such residues in this entry.