



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 15, 2024 – 10:55 AM EST

PDB ID : 6CHT
Title : HNF4alpha in complex with the corepressor EBP1 fragment
Authors : Chi, Y.I.; Singh, P.; Lee, I.K.
Deposited on : 2018-02-22
Resolution : 3.17 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	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

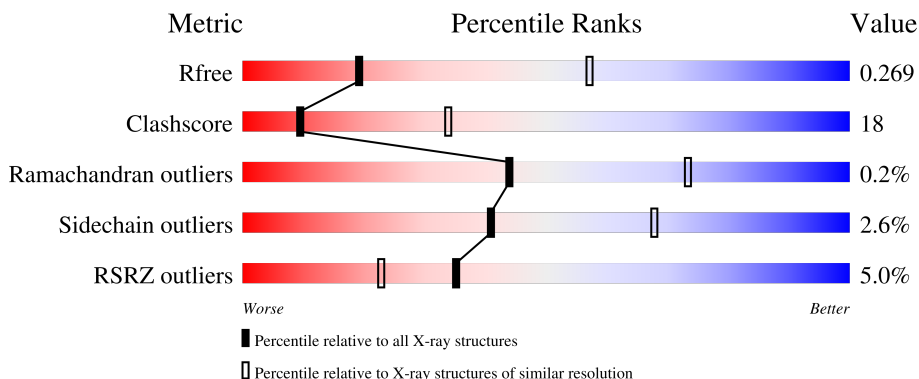
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.17 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}	164625	1851 (3.20-3.16)
Clashscore	180529	1999 (3.20-3.16)
Ramachandran outliers	177936	1961 (3.20-3.16)
Sidechain outliers	177891	1960 (3.20-3.16)
RSRZ outliers	164620	1852 (3.20-3.16)

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 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	A	245	<div> <div>53%</div> <div>31%</div> <div>•</div> <div>14%</div> </div>
1	B	245	<div> <div>54%</div> <div>29%</div> <div></div> <div>16%</div> </div>
1	D	245	<div> <div>11%</div> <div>55%</div> <div>30%</div> <div>•</div> <div>13%</div> </div>
1	E	245	<div> <div>12%</div> <div>53%</div> <div>30%</div> <div></div> <div>17%</div> </div>
1	G	245	<div> <div>2%</div> <div>52%</div> <div>33%</div> <div>•</div> <div>15%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	H	245	
1	J	245	
1	K	245	
1	M	245	
1	N	245	
1	P	245	
1	Q	245	
1	S	245	
1	T	245	
1	V	245	
1	W	245	
2	C	20	
2	F	20	
2	I	20	
2	L	20	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 26381 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 Hepatocyte nuclear factor 4-alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	210	Total	C	N	O	S	0	0	0
			1681	1086	278	308	9			
1	B	205	Total	C	N	O	S	0	0	0
			1647	1062	275	301	9			
1	D	214	Total	C	N	O	S	0	0	0
			1712	1104	285	314	9			
1	E	204	Total	C	N	O	S	0	0	0
			1639	1056	274	300	9			
1	G	209	Total	C	N	O	S	0	0	0
			1678	1085	279	305	9			
1	H	202	Total	C	N	O	S	0	0	0
			1627	1050	271	297	9			
1	J	209	Total	C	N	O	S	0	0	0
			1678	1085	279	305	9			
1	K	202	Total	C	N	O	S	0	0	0
			1627	1050	271	297	9			
1	M	201	Total	C	N	O	S	0	0	0
			1619	1046	270	294	9			
1	N	193	Total	C	N	O	S	0	0	0
			1557	1005	257	286	9			
1	P	200	Total	C	N	O	S	0	0	0
			1611	1040	269	293	9			
1	Q	196	Total	C	N	O	S	0	0	0
			1582	1023	261	289	9			
1	S	201	Total	C	N	O	S	0	0	0
			1615	1042	270	294	9			
1	T	192	Total	C	N	O	S	0	0	0
			1554	1005	256	284	9			
1	V	202	Total	C	N	O	S	0	0	0
			1627	1050	271	297	9			
1	W	192	Total	C	N	O	S	0	0	0
			1554	1005	256	284	9			

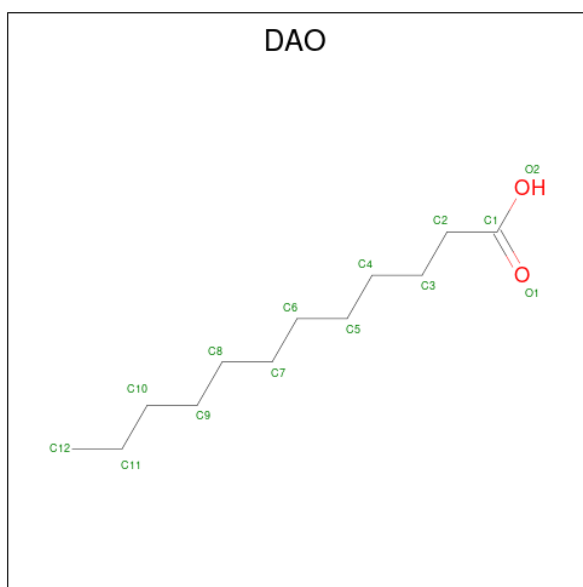
There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	138	GLY	-	expression tag	UNP P41235
B	138	GLY	-	expression tag	UNP P41235
D	138	GLY	-	expression tag	UNP P41235
E	138	GLY	-	expression tag	UNP P41235
G	138	GLY	-	expression tag	UNP P41235
H	138	GLY	-	expression tag	UNP P41235
J	138	GLY	-	expression tag	UNP P41235
K	138	GLY	-	expression tag	UNP P41235
M	138	GLY	-	expression tag	UNP P41235
N	138	GLY	-	expression tag	UNP P41235
P	138	GLY	-	expression tag	UNP P41235
Q	138	GLY	-	expression tag	UNP P41235
S	138	GLY	-	expression tag	UNP P41235
T	138	GLY	-	expression tag	UNP P41235
V	138	GLY	-	expression tag	UNP P41235
W	138	GLY	-	expression tag	UNP P41235

- Molecule 2 is a protein called Proliferation-associated protein 2G4.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	7	Total	C	N	O	0	0	0
			52	35	8	9			
2	F	7	Total	C	N	O	0	0	0
			52	35	8	9			
2	I	7	Total	C	N	O	0	0	0
			52	35	8	9			
2	L	8	Total	C	N	O	0	0	0
			61	40	10	11			

- Molecule 3 is LAURIC ACID (three-letter code: DAO) (formula: C₁₂H₂₄O₂).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			14	12	2		
3	B	1	Total	C	O	0	0
			14	12	2		
3	D	1	Total	C	O	0	0
			14	12	2		
3	E	1	Total	C	O	0	0
			14	12	2		
3	G	1	Total	C	O	0	0
			14	12	2		
3	H	1	Total	C	O	0	0
			14	12	2		
3	J	1	Total	C	O	0	0
			14	12	2		
3	K	1	Total	C	O	0	0
			14	12	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	6	Total	O	0	0
			6	6		
4	B	6	Total	O	0	0
			6	6		
4	D	1	Total	O	0	0
			1	1		
4	E	1	Total	O	0	0
			1	1		

Continued on next page...

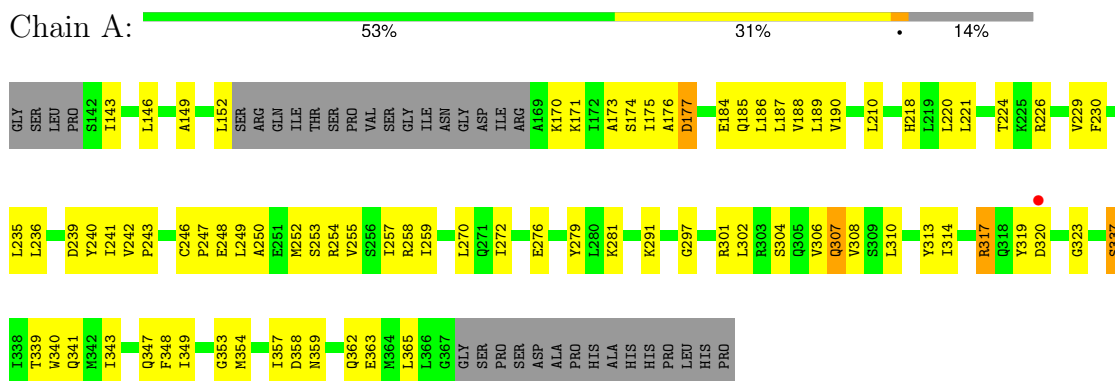
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	F	1	Total 1	O 1	0	0
4	G	5	Total 5	O 5	0	0
4	H	4	Total 4	O 4	0	0
4	J	2	Total 2	O 2	0	0
4	K	7	Total 7	O 7	0	0
4	L	1	Total 1	O 1	0	0
4	M	1	Total 1	O 1	0	0
4	N	2	Total 2	O 2	0	0
4	Q	1	Total 1	O 1	0	0
4	S	2	Total 2	O 2	0	0
4	V	2	Total 2	O 2	0	0
4	W	2	Total 2	O 2	0	0

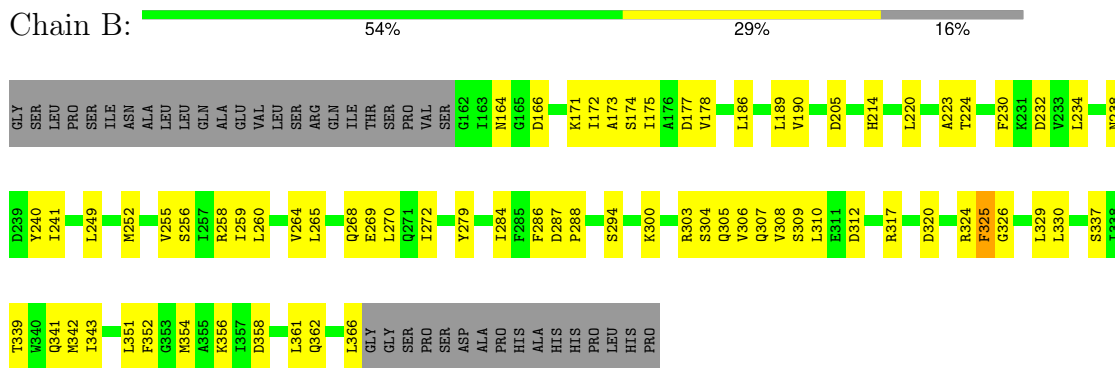
3 Residue-property plots

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.

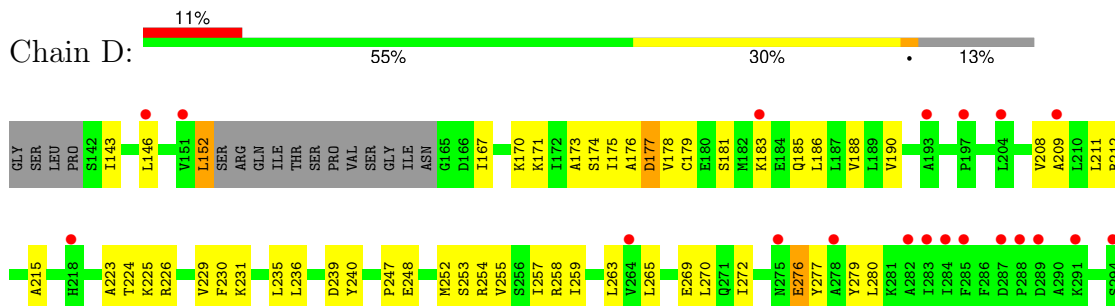
• Molecule 1: Hepatocyte nuclear factor 4-alpha

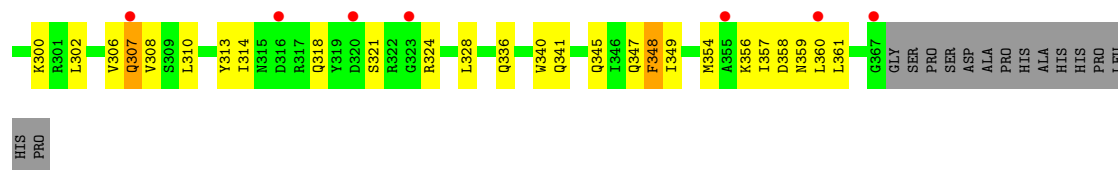


• Molecule 1: Hepatocyte nuclear factor 4-alpha

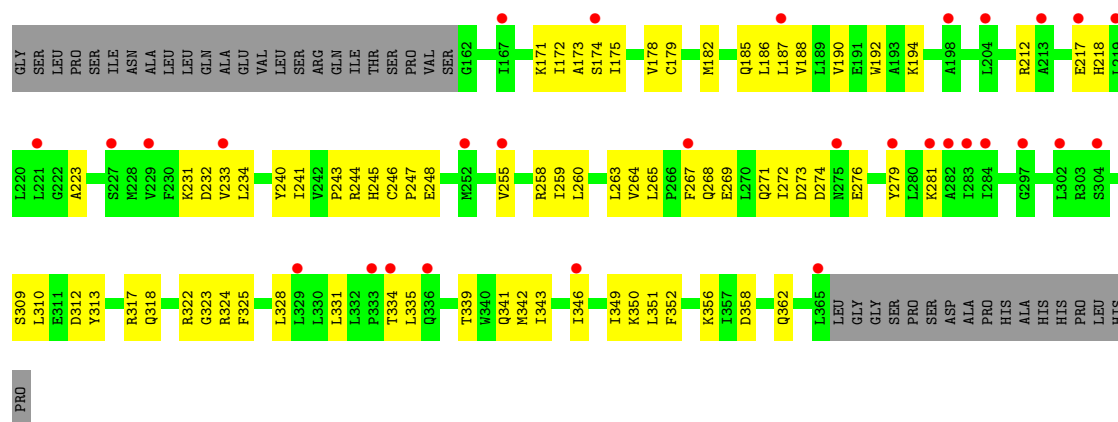


• Molecule 1: Hepatocyte nuclear factor 4-alpha

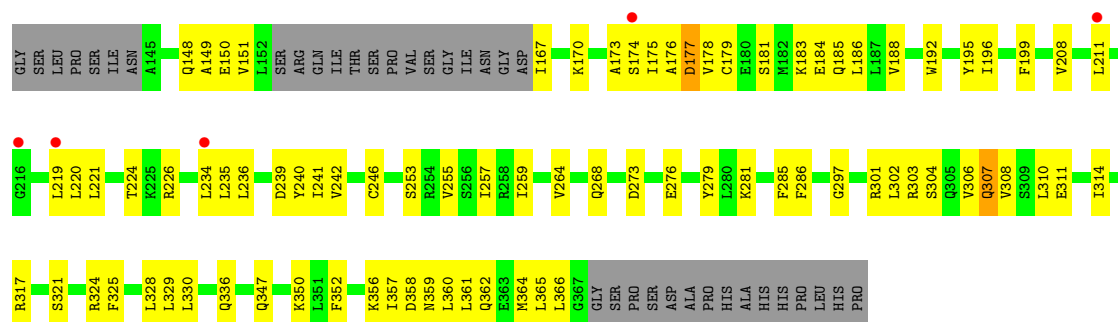




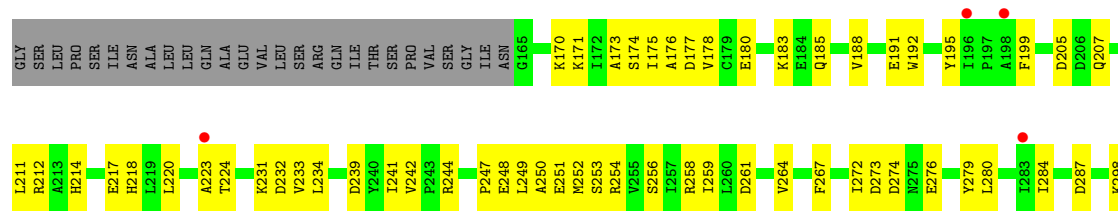
• Molecule 1: Hepatocyte nuclear factor 4-alpha



• Molecule 1: Hepatocyte nuclear factor 4-alpha



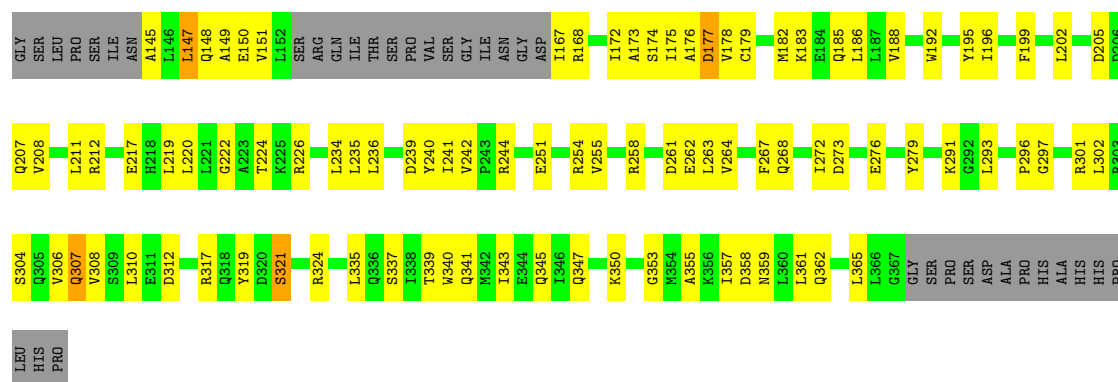
• Molecule 1: Hepatocyte nuclear factor 4-alpha





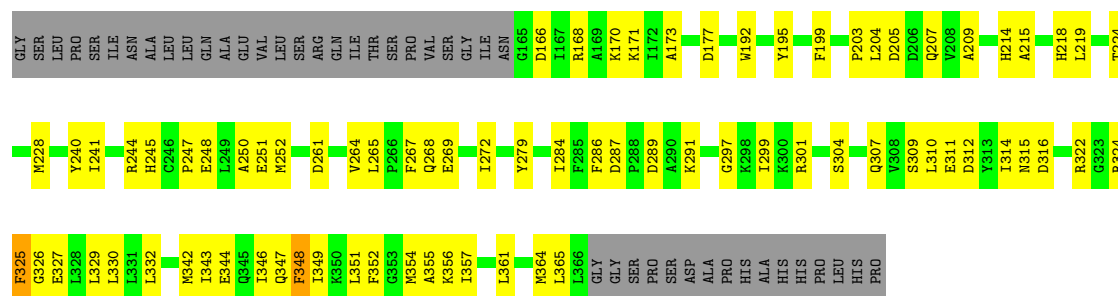
• Molecule 1: Hepatocyte nuclear factor 4-alpha

Chain J: 48% 36% 15%



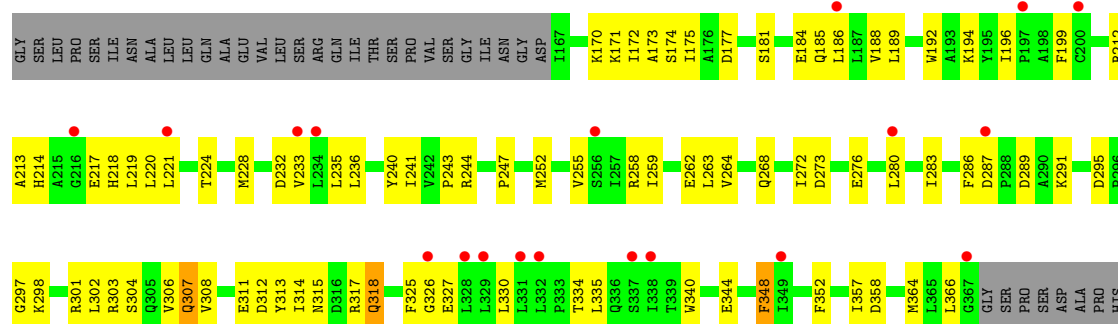
• Molecule 1: Hepatocyte nuclear factor 4-alpha

Chain K: 51% 31% 18%



• Molecule 1: Hepatocyte nuclear factor 4-alpha

Chain M: 8% 48% 33% 18%

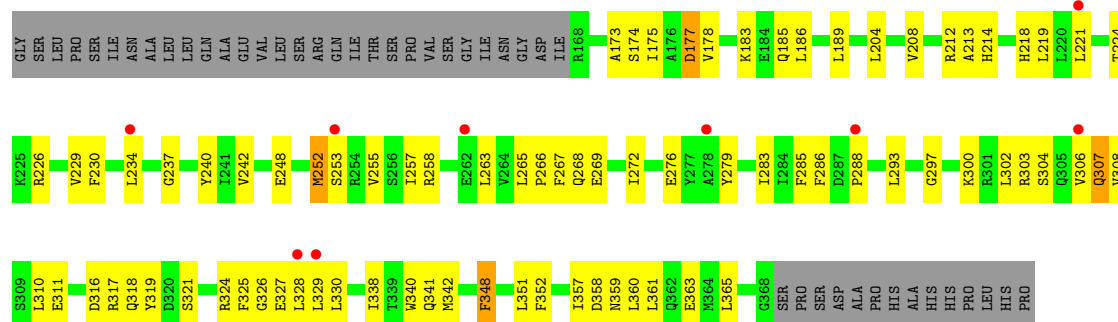


GLY	H214	K291	LEU	A215	PRO	L220	SER	L221	THR	V242	ASP	L251	ALA	L182	LYS	K171	GLY	S181	LEU	L226	GLN	L235	LEU	L236	GLU	F230	VAL	P243	SER	H245	ASN	E248	GLY	L249	ASP	E250	LEU	L252	ARG	A253	ALA	L183	LEU	L257	LEU	L258	LEU	L259	LEU	L260	LEU	L261	LEU	L262	LEU	L263	LEU	L264	LEU	L265	LEU	L266	LEU	L267	LEU	L268	LEU	L269	LEU	L270	LEU	L271	LEU	L272	LEU	L273	LEU	L274	LEU	L275	LEU	L276	LEU	L277	LEU	L278	LEU	L279	LEU	L280	LEU	L281	LEU	L282	LEU	L283	LEU	L284	LEU	L285	LEU	L286	LEU	L287	LEU	L288	LEU	L289	LEU	L290	LEU	L291	LEU	L292	LEU	L293	LEU	L294	LEU	L295	LEU	L296	LEU	L297	LEU	L298	LEU	L299	LEU	L300	LEU	L301	LEU	L302	LEU	L303	LEU	L304	LEU	L305	LEU	L306	LEU	L307	LEU	L308	LEU	L309	LEU	L310	LEU	L311	LEU	L312	LEU	L313	LEU	L314	LEU	L315	LEU	L316	LEU	L317	LEU	L318	LEU	L319	LEU	L320	LEU	L321	LEU	L322	LEU	L323	LEU	L324	LEU	L325	LEU	L326	LEU	L327	LEU	L328	LEU	L329	LEU	L330	LEU	L331	LEU	L332	LEU	L333	LEU	L334	LEU	L335	LEU	L336	LEU	L337	LEU	L338	LEU	L339	LEU	L340	LEU	L341	LEU	L342	LEU	L343	LEU	L344	LEU	L345	LEU	L346	LEU	L347	LEU	L348	LEU	L349	LEU	L350	LEU	L351	LEU	L352	LEU	L353	LEU	L354	LEU	L355	LEU	L356	LEU	L357	LEU	L358	LEU	L359	LEU	L360	LEU	L361	LEU	L362	LEU	L363	LEU	L364	LEU	L365	LEU	L366	LEU	L367	LEU	L368	LEU	L369	LEU	L370	LEU	L371	LEU	L372	LEU	L373	LEU	L374	LEU	L375	LEU	L376	LEU	L377	LEU	L378	LEU	L379	LEU	L380	LEU	L381	LEU	L382	LEU	L383	LEU	L384	LEU	L385	LEU	L386	LEU	L387	LEU	L388	LEU	L389	LEU	L390	LEU	L391	LEU	L392	LEU	L393	LEU	L394	LEU	L395	LEU	L396	LEU	L397	LEU	L398	LEU	L399	LEU	L400	LEU	L401	LEU	L402	LEU	L403	LEU	L404	LEU	L405	LEU	L406	LEU	L407	LEU	L408	LEU	L409	LEU	L410	LEU	L411	LEU	L412	LEU	L413	LEU	L414	LEU	L415	LEU	L416	LEU	L417	LEU	L418	LEU	L419	LEU	L420	LEU	L421	LEU	L422	LEU	L423	LEU	L424	LEU	L425	LEU	L426	LEU	L427	LEU	L428	LEU	L429	LEU	L430	LEU	L431	LEU	L432	LEU	L433	LEU	L434	LEU	L435	LEU	L436	LEU	L437	LEU	L438	LEU	L439	LEU	L440	LEU	L441	LEU	L442	LEU	L443	LEU	L444	LEU	L445	LEU	L446	LEU	L447	LEU	L448	LEU	L449	LEU	L450	LEU	L451	LEU	L452	LEU	L453	LEU	L454	LEU	L455	LEU	L456	LEU	L457	LEU	L458	LEU	L459	LEU	L460	LEU	L461	LEU	L462	LEU	L463	LEU	L464	LEU	L465	LEU	L466	LEU	L467	LEU	L468	LEU	L469	LEU	L470	LEU	L471	LEU	L472	LEU	L473	LEU	L474	LEU	L475	LEU	L476	LEU	L477	LEU	L478	LEU	L479	LEU	L480	LEU	L481	LEU	L482	LEU	L483	LEU	L484	LEU	L485	LEU	L486	LEU	L487	LEU	L488	LEU	L489	LEU	L490	LEU	L491	LEU	L492	LEU	L493	LEU	L494	LEU	L495	LEU	L496	LEU	L497	LEU	L498	LEU	L499	LEU	L500	LEU	L501	LEU	L502	LEU	L503	LEU	L504	LEU	L505	LEU	L506	LEU	L507	LEU	L508
-----	------	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------	-----	------

GLY
GLY
SER
LEU
PRO
PRO
SER
SER
ASP
ALA
PRO
HIS
HIS
HIS
HIS
LEU
HIS
PRO

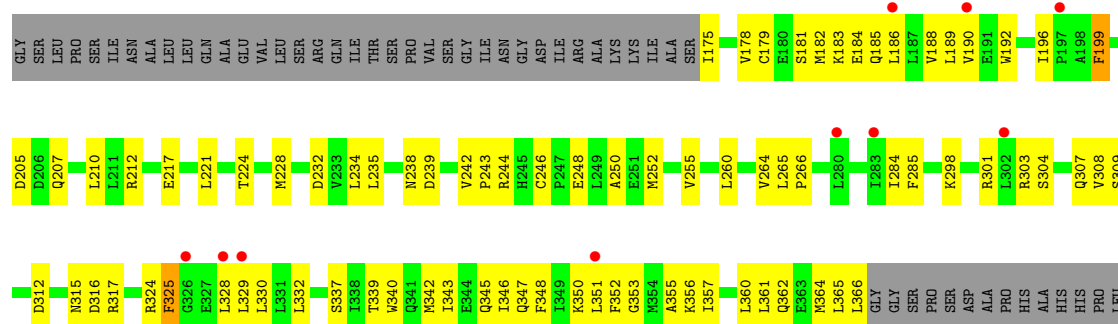
• Molecule 1: Hepatocyte nuclear factor 4-alpha

Chain S: 4% 49% 31% 18%



• Molecule 1: Hepatocyte nuclear factor 4-alpha

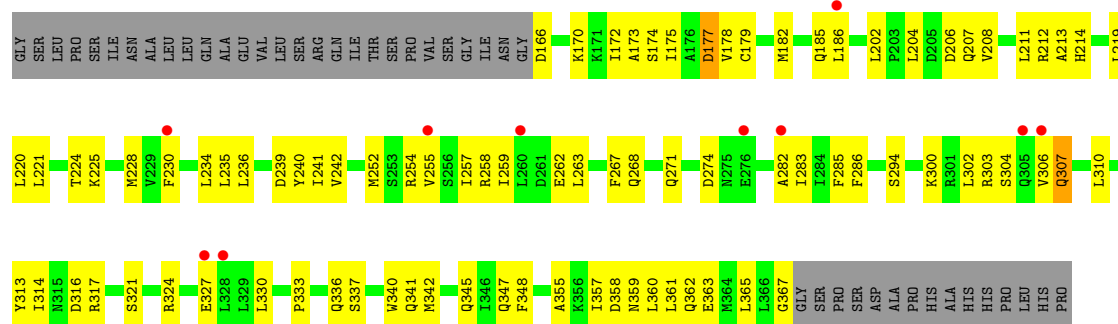
Chain T: 4% 45% 32% 22%



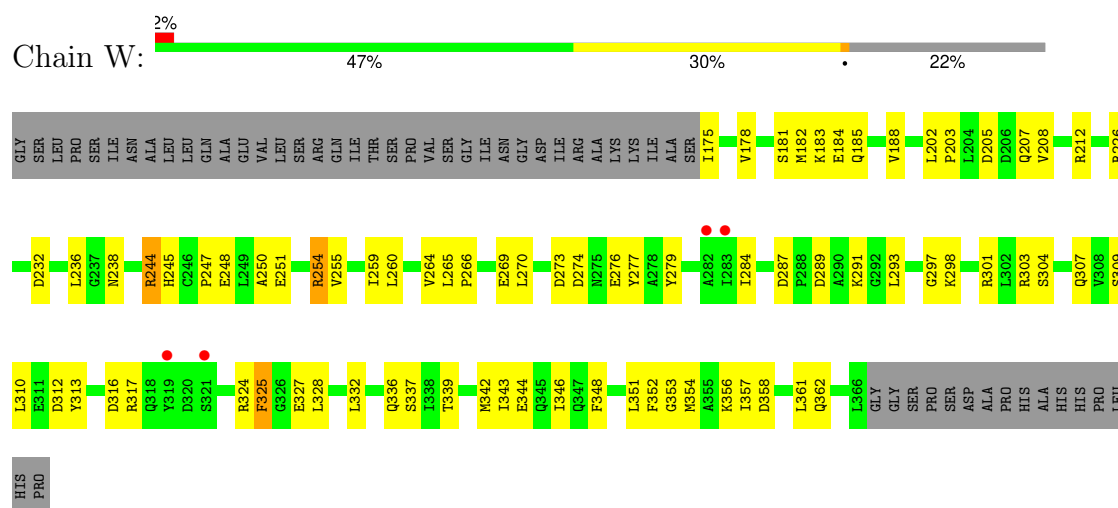
HIS
PRO

• Molecule 1: Hepatocyte nuclear factor 4-alpha

Chain V: 4% 47% 34% 18%



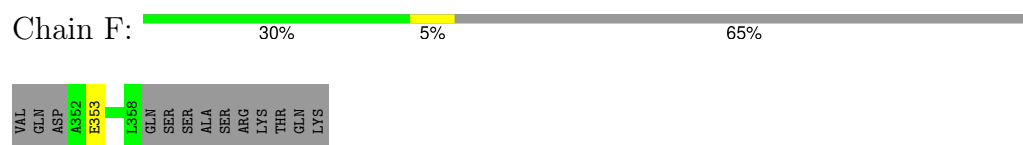
• Molecule 1: Hepatocyte nuclear factor 4-alpha



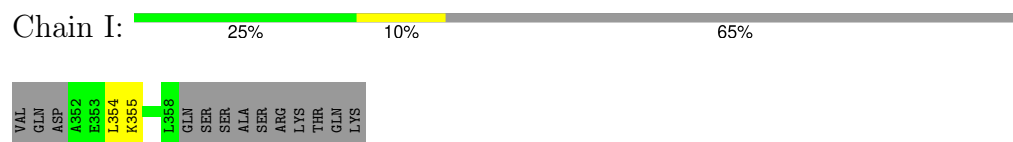
- Molecule 2: Proliferation-associated protein 2G4



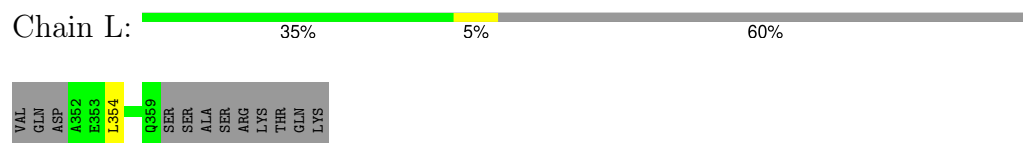
- Molecule 2: Proliferation-associated protein 2G4



- Molecule 2: Proliferation-associated protein 2G4



- Molecule 2: Proliferation-associated protein 2G4



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	139.72Å 104.95Å 139.56Å 90.00° 90.61° 90.00°	Depositor
Resolution (Å)	49.63 – 3.17 49.63 – 3.17	Depositor EDS
% Data completeness (in resolution range)	71.6 (49.63-3.17) 71.6 (49.63-3.17)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	644.16 (at 3.19Å)	Xtriage
Refinement program	PHENIX (1.12_2829: ???)	Depositor
R, R_{free}	0.262 , 0.274 0.259 , 0.269	Depositor DCC
R_{free} test set	2505 reflections (5.24%)	wwPDB-VP
Wilson B-factor (Å ²)	66.1	Xtriage
Anisotropy	0.153	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 689.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.35$, $\langle L^2 \rangle = 0.18$	Xtriage
Estimated twinning fraction	0.046 for -l,k,h 0.056 for -h,-k,l 0.390 for l,-k,h	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	26381	wwPDB-VP
Average B, all atoms (Å ²)	124.0	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 77.42 % 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 8.5463e-07. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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: DAO

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	0.30	0/1708	0.52	0/2307
1	B	0.31	0/1675	0.54	0/2262
1	D	0.28	0/1739	0.50	0/2348
1	E	0.29	0/1667	0.52	0/2251
1	G	0.28	0/1705	0.52	0/2302
1	H	0.30	0/1655	0.53	0/2235
1	J	0.29	0/1705	0.53	0/2302
1	K	0.31	0/1655	0.55	0/2235
1	M	0.27	0/1647	0.51	0/2224
1	N	0.26	0/1585	0.47	0/2143
1	P	0.27	0/1639	0.50	0/2213
1	Q	0.27	0/1610	0.47	0/2176
1	S	0.28	0/1643	0.52	0/2218
1	T	0.28	0/1582	0.47	0/2139
1	V	0.29	0/1655	0.53	0/2235
1	W	0.28	0/1582	0.47	0/2139
2	C	0.26	0/51	0.45	0/67
2	F	0.24	0/51	0.48	0/67
2	I	0.24	0/51	0.67	0/67
2	L	0.24	0/60	0.65	0/79
All	All	0.28	0/26665	0.51	0/36009

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	1681	0	1733	69	1
1	B	1647	0	1694	60	0
1	D	1712	0	1765	59	1
1	E	1639	0	1683	58	0
1	G	1678	0	1735	67	0
1	H	1627	0	1673	66	0
1	J	1678	0	1736	63	0
1	K	1627	0	1673	63	0
1	M	1619	0	1671	64	0
1	N	1557	0	1591	69	0
1	P	1611	0	1660	73	0
1	Q	1582	0	1626	70	0
1	S	1615	0	1663	61	0
1	T	1554	0	1592	66	1
1	V	1627	0	1675	69	0
1	W	1554	0	1592	68	1
2	C	52	0	61	3	0
2	F	52	0	61	1	0
2	I	52	0	61	4	0
2	L	61	0	69	2	0
3	A	14	0	23	1	0
3	B	14	0	23	2	0
3	D	14	0	23	3	0
3	E	14	0	23	2	0
3	G	14	0	23	3	0
3	H	14	0	23	2	0
3	J	14	0	23	2	0
3	K	14	0	23	1	0
4	A	6	0	0	0	0
4	B	6	0	0	0	0
4	D	1	0	0	0	0
4	E	1	0	0	0	0
4	F	1	0	0	0	0
4	G	5	0	0	0	0
4	H	4	0	0	0	0
4	J	2	0	0	0	0
4	K	7	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	L	1	0	0	0	0
4	M	1	0	0	0	0
4	N	2	0	0	0	0
4	Q	1	0	0	1	0
4	S	2	0	0	0	0
4	V	2	0	0	1	0
4	W	2	0	0	0	0
All	All	26381	0	27198	941	2

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 18.

The worst 5 of 941 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:M:348:PHE:HE1	1:M:352:PHE:CE2	1.32	1.46
1:M:348:PHE:CE1	1:M:352:PHE:CE2	2.19	1.31
1:A:317:ARG:NE	1:A:319:TYR:CD2	2.25	1.04
1:A:317:ARG:CZ	1:A:319:TYR:CE2	2.51	0.93
1:A:317:ARG:CZ	1:A:319:TYR:CD2	2.53	0.91

All (2) 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)
1:D:240:TYR:OH	1:T:316:ASP:O[2_544]	2.15	0.05
1:A:171:LYS:NZ	1:W:316:ASP:O[2_445]	2.16	0.04

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	206/245 (84%)	189 (92%)	17 (8%)	0	100	100
1	B	203/245 (83%)	189 (93%)	13 (6%)	1 (0%)	25	58
1	D	210/245 (86%)	190 (90%)	18 (9%)	2 (1%)	13	45
1	E	202/245 (82%)	190 (94%)	12 (6%)	0	100	100
1	G	205/245 (84%)	182 (89%)	23 (11%)	0	100	100
1	H	200/245 (82%)	192 (96%)	7 (4%)	1 (0%)	25	58
1	J	205/245 (84%)	187 (91%)	17 (8%)	1 (0%)	25	58
1	K	200/245 (82%)	185 (92%)	15 (8%)	0	100	100
1	M	199/245 (81%)	178 (89%)	20 (10%)	1 (0%)	25	58
1	N	191/245 (78%)	186 (97%)	5 (3%)	0	100	100
1	P	198/245 (81%)	178 (90%)	20 (10%)	0	100	100
1	Q	194/245 (79%)	183 (94%)	11 (6%)	0	100	100
1	S	199/245 (81%)	173 (87%)	24 (12%)	2 (1%)	13	45
1	T	190/245 (78%)	181 (95%)	9 (5%)	0	100	100
1	V	200/245 (82%)	181 (90%)	19 (10%)	0	100	100
1	W	190/245 (78%)	179 (94%)	11 (6%)	0	100	100
2	C	5/20 (25%)	5 (100%)	0	0	100	100
2	F	5/20 (25%)	5 (100%)	0	0	100	100
2	I	5/20 (25%)	5 (100%)	0	0	100	100
2	L	6/20 (30%)	6 (100%)	0	0	100	100
All	All	3213/4000 (80%)	2964 (92%)	241 (8%)	8 (0%)	44	73

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	J	321	SER
1	M	318	GLN
1	S	248	GLU
1	H	232	ASP
1	B	232	ASP

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	185/214 (86%)	178 (96%)	7 (4%)	28	58
1	B	181/214 (85%)	177 (98%)	4 (2%)	47	71
1	D	188/214 (88%)	182 (97%)	6 (3%)	34	63
1	E	180/214 (84%)	178 (99%)	2 (1%)	70	84
1	G	184/214 (86%)	180 (98%)	4 (2%)	47	71
1	H	179/214 (84%)	176 (98%)	3 (2%)	56	77
1	J	184/214 (86%)	178 (97%)	6 (3%)	33	62
1	K	179/214 (84%)	175 (98%)	4 (2%)	47	71
1	M	178/214 (83%)	171 (96%)	7 (4%)	27	57
1	N	172/214 (80%)	169 (98%)	3 (2%)	56	77
1	P	177/214 (83%)	169 (96%)	8 (4%)	23	53
1	Q	175/214 (82%)	173 (99%)	2 (1%)	70	84
1	S	177/214 (83%)	170 (96%)	7 (4%)	27	56
1	T	172/214 (80%)	169 (98%)	3 (2%)	56	77
1	V	179/214 (84%)	173 (97%)	6 (3%)	32	61
1	W	172/214 (80%)	169 (98%)	3 (2%)	56	77
2	C	5/17 (29%)	4 (80%)	1 (20%)	1	5
2	F	5/17 (29%)	5 (100%)	0	100	100
2	I	5/17 (29%)	5 (100%)	0	100	100
2	L	6/17 (35%)	6 (100%)	0	100	100
All	All	2883/3492 (83%)	2807 (97%)	76 (3%)	41	68

5 of 76 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	S	177	ASP
1	V	313	TYR
1	S	252	MET
1	T	315	ASN
1	W	325	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	N	307	GLN
1	P	307	GLN
1	V	207	GLN
1	Q	185	GLN
1	K	347	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 ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

8 ligands are 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	$\# Z > 2$
3	DAO	K	500	1	13,13,13	0.64	0	13,13,13	1.09	0
3	DAO	J	500	-	13,13,13	0.68	0	13,13,13	1.20	1 (7%)
3	DAO	H	500	1	13,13,13	0.70	0	13,13,13	1.22	1 (7%)
3	DAO	D	500	-	13,13,13	0.66	0	13,13,13	1.13	0
3	DAO	B	500	1	13,13,13	0.70	0	13,13,13	1.28	1 (7%)
3	DAO	A	500	1	13,13,13	0.67	0	13,13,13	1.13	0
3	DAO	G	500	1	13,13,13	0.68	0	13,13,13	1.18	1 (7%)
3	DAO	E	500	1	13,13,13	0.68	0	13,13,13	1.26	1 (7%)

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
3	DAO	K	500	1	-	6/11/11/11	-
3	DAO	J	500	-	-	7/11/11/11	-
3	DAO	H	500	1	-	4/11/11/11	-
3	DAO	D	500	-	-	6/11/11/11	-
3	DAO	B	500	1	-	3/11/11/11	-
3	DAO	A	500	1	-	3/11/11/11	-
3	DAO	G	500	1	-	2/11/11/11	-
3	DAO	E	500	1	-	3/11/11/11	-

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	500	DAO	C3-C2-C1	-2.68	107.51	114.51
3	E	500	DAO	C3-C2-C1	-2.66	107.56	114.51
3	J	500	DAO	C3-C2-C1	-2.40	108.24	114.51
3	H	500	DAO	C3-C2-C1	-2.28	108.57	114.51
3	G	500	DAO	O2-C1-C2	2.20	120.94	114.00

There are no chirality outliers.

5 of 34 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	500	DAO	C1-C2-C3-C4
3	J	500	DAO	C1-C2-C3-C4
3	J	500	DAO	C3-C4-C5-C6
3	D	500	DAO	C6-C7-C8-C9
3	B	500	DAO	C5-C6-C7-C8

There are no ring outliers.

8 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	K	500	DAO	1	0
3	J	500	DAO	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	H	500	DAO	2	0
3	D	500	DAO	3	0
3	B	500	DAO	2	0
3	A	500	DAO	1	0
3	G	500	DAO	3	0
3	E	500	DAO	2	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 ⓘ

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, 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	A	210/245 (85%)	-0.20	1 (0%) 87 78	18, 46, 87, 107	0
1	B	205/245 (83%)	-0.50	0 100 100	12, 28, 86, 118	0
1	D	214/245 (87%)	1.11	27 (12%) 9 6	73, 127, 199, 208	0
1	E	204/245 (83%)	1.17	30 (14%) 7 4	85, 120, 145, 157	0
1	G	209/245 (85%)	0.59	5 (2%) 59 43	63, 95, 117, 128	0
1	H	202/245 (82%)	0.28	6 (2%) 52 36	25, 62, 101, 124	0
1	J	209/245 (85%)	0.06	0 100 100	41, 70, 103, 141	0
1	K	202/245 (82%)	-0.23	0 100 100	18, 40, 81, 99	0
1	M	201/245 (82%)	0.83	19 (9%) 15 10	84, 193, 319, 333	0
1	N	193/245 (78%)	0.78	13 (6%) 25 17	58, 207, 454, 485	0
1	P	200/245 (81%)	0.73	11 (5%) 32 20	118, 152, 235, 277	0
1	Q	196/245 (80%)	0.78	17 (8%) 17 12	69, 209, 377, 395	0
1	S	201/245 (82%)	0.65	9 (4%) 39 25	69, 139, 198, 244	0
1	T	192/245 (78%)	0.64	10 (5%) 34 22	101, 153, 174, 180	0
1	V	202/245 (82%)	0.62	10 (4%) 35 23	89, 129, 164, 192	0
1	W	192/245 (78%)	0.47	4 (2%) 63 47	54, 142, 162, 167	0
2	C	7/20 (35%)	-0.24	0 100 100	76, 79, 90, 95	0
2	F	7/20 (35%)	0.09	0 100 100	127, 137, 150, 152	0
2	I	7/20 (35%)	0.61	0 100 100	188, 189, 191, 191	0
2	L	8/20 (40%)	-0.33	0 100 100	84, 87, 89, 92	0
All	All	3261/4000 (81%)	0.48	162 (4%) 35 23	12, 120, 260, 485	0

The worst 5 of 162 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	N	253	SER	5.7
1	M	329	LEU	5.4
1	Q	253	SER	4.8
1	P	328	LEU	4.6
1	W	282	ALA	4.4

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, 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(\AA^2)	Q<0.9
3	DAO	D	500	14/14	0.95	0.17	78,95,114,119	0
3	DAO	G	500	14/14	0.96	0.13	75,96,103,105	0
3	DAO	J	500	14/14	0.96	0.16	62,80,95,96	0
3	DAO	H	500	14/14	0.97	0.14	41,45,54,56	0
3	DAO	A	500	14/14	0.97	0.12	42,52,70,74	0
3	DAO	E	500	14/14	0.98	0.19	94,108,118,119	0
3	DAO	B	500	14/14	0.98	0.10	21,26,46,47	0
3	DAO	K	500	14/14	0.98	0.10	15,34,44,48	0

6.5 Other polymers [i](#)

There are no such residues in this entry.