



wwPDB EM Map/Model Validation Report ⓘ

Sep 27, 2016 – 08:06 PM EDT

PDB ID : 5IY6
EMDB ID: : EMD-3307
Title : Human holo-PIC in the closed state
Authors : He, Y.; Yan, C.; Fang, J.; Inouye, C.; Tjian, R.; Ivanov, I.; Nogales, E.
Deposited on : 2016-03-24
Resolution : 7.20 Å(reported)
Based on PDB ID : ?

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
<http://wwpdb.org/validation/2016/EMValidationReportHelp>

MolProbity : 4.02b-467
Mogul : unknown
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20027939

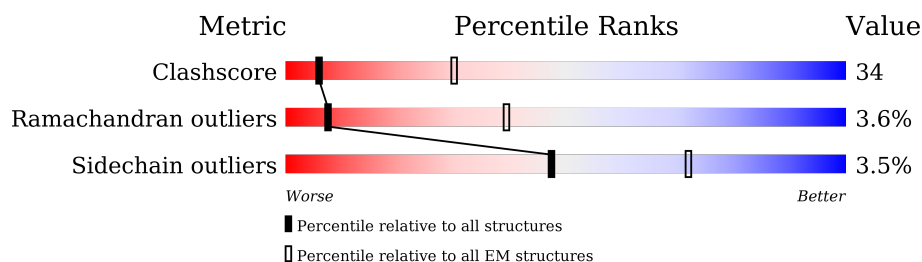
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 7.20 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.





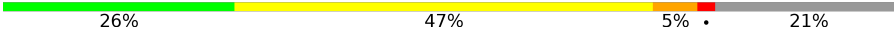








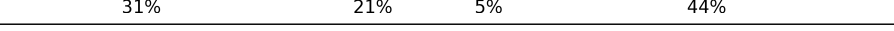
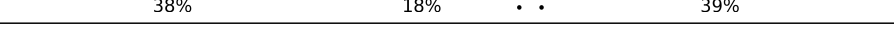
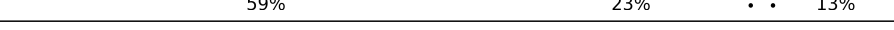

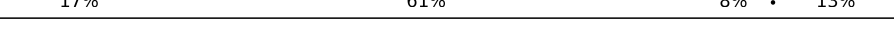




Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for ≥ 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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	1970	
2	B	1174	
3	C	275	
4	D	142	
5	E	210	
6	F	127	
7	G	172	
8	H	150	
9	I	125	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
10	J	67	
11	K	117	
12	L	58	
13	M	316	
14	N	376	
15	O	109	
16	P	339	
17	Q	439	
18	R	291	
19	S	517	
20	T	249	
21	U	301	
22	V	782	
23	W	760	
24	0	395	
25	1	71	
26	2	462	
27	3	308	
28	X	65	
29	Y	65	

2 Entry composition

There are 31 unique types of molecules in this entry. The entry contains 61839 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 DNA-directed RNA polymerase II subunit RPB1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1454	Total	C	N	O	S	0	0
			11515	7234	2058	2150	73		

- Molecule 2 is a protein called DNA-directed RNA polymerase II subunit RPB2.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	1165	Total	C	N	O	S	0	0
			9317	5878	1637	1738	64		

- Molecule 3 is a protein called DNA-directed RNA polymerase II subunit RPB3.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	275	Total	C	N	O	S	0	0
			2213	1386	380	440	7		

- Molecule 4 is a protein called DNA-directed RNA polymerase II subunit RPB4.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	129	Total	C	N	O	S	0	0
			1062	665	179	214	4		

- Molecule 5 is a protein called DNA-directed RNA polymerase II subunit RPB5.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	210	Total	C	N	O	S	0	0
			1723	1088	301	325	9		

- Molecule 6 is a protein called DNA-directed RNA polymerase II subunit RPB6.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	86	Total	C	N	O	S	0	0
			689	437	120	127	5		

- Molecule 7 is a protein called DNA-directed RNA polymerase II subunit RPB7.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	G	171	Total	C	N	O	S	0	0
			1351	875	219	249	8		

- Molecule 8 is a protein called DNA-directed RNA polymerase II subunit RPB8.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	H	150	Total	C	N	O	S	0	0
			1205	764	196	239	6		

- Molecule 9 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	I	125	Total	C	N	O	S	0	0
			1013	626	177	198	12		

- Molecule 10 is a protein called DNA-directed RNA polymerase II subunit RPB10.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	J	67	Total	C	N	O	S	0	0
			533	345	90	92	6		

- Molecule 11 is a protein called DNA-directed RNA polymerase II subunit RPB11-a.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	K	117	Total	C	N	O	S	0	0
			937	604	154	177	2		

- Molecule 12 is a protein called DNA-directed RNA polymerase II subunit RPB12.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	46	Total	C	N	O	S	0	0
			388	241	75	66	6		

- Molecule 13 is a protein called Transcription initiation factor IIB.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	M	260	Total	C	N	O	S	0	0
			2018	1265	360	376	17		

- Molecule 14 is a protein called Transcription initiation factor IIA subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	N	113	Total	C	N	O	S	0	0
			930	585	152	189	4		

- Molecule 15 is a protein called Transcription initiation factor IIA subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	O	99	Total	C	N	O	S	0	0
			806	510	142	151	3		

- Molecule 16 is a protein called TATA-box-binding protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	P	185	Total	C	N	O	S	0	0
			1462	946	257	252	7		

- Molecule 17 is a protein called General transcription factor IIE subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	Q	180	Total	C	N	O	S	0	0
			1484	938	262	273	11		

- Molecule 18 is a protein called Transcription initiation factor IIE subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	R	165	Total	C	N	O	S	0	0
			1357	865	235	253	4		

- Molecule 19 is a protein called General transcription factor IIF subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	S	138	Total	C	N	O	S	0	0
			1138	719	208	208	3		

- Molecule 20 is a protein called General transcription factor IIF subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	T	222	Total	C	N	O	S	0	0
			1788	1127	320	338	3		

- Molecule 21 is a protein called Transcription elongation factor A protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	U	170	Total	C	N	O	S	0	0
			1343	818	247	263	15		

- Molecule 22 is a protein called TFIIF basal transcription factor complex helicase XPB subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	V	475	Total	C	N	O	S	0	0
			3855	2454	663	712	26		

- Molecule 23 is a protein called TFIIF basal transcription factor complex helicase XPD subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	W	665	Total	C	N	O	S	0	0
			5348	3415	932	975	26		

- Molecule 24 is a protein called General transcription factor IIF subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	0	188	Total	C	N	O	S	0	0
			1479	935	258	276	10		

- Molecule 25 is a protein called General transcription factor IIF subunit 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	1	62	Total	C	N	O	S	0	0
			491	317	77	93	4		

- Molecule 26 is a protein called General transcription factor IIF subunit 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	2	274	Total	C	N	O	S	0	0
			2196	1417	377	392	10		

- Molecule 27 is a protein called General transcription factor IIF subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	3	193	Total	C	N	O	S	0	0
			1526	978	252	284	12		

- Molecule 28 is a DNA chain called SCP-X.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	X	65	Total	C	N	O	P	0	0
			1343	633	261	385	64		

- Molecule 29 is a DNA chain called SCP-Y.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	Y	65	Total	C	N	O	P	0	0
			1316	625	236	391	64		

- Molecule 30 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
30	B	1	Total	Mg	0
			1	1	
30	A	1	Total	Mg	0
			1	1	

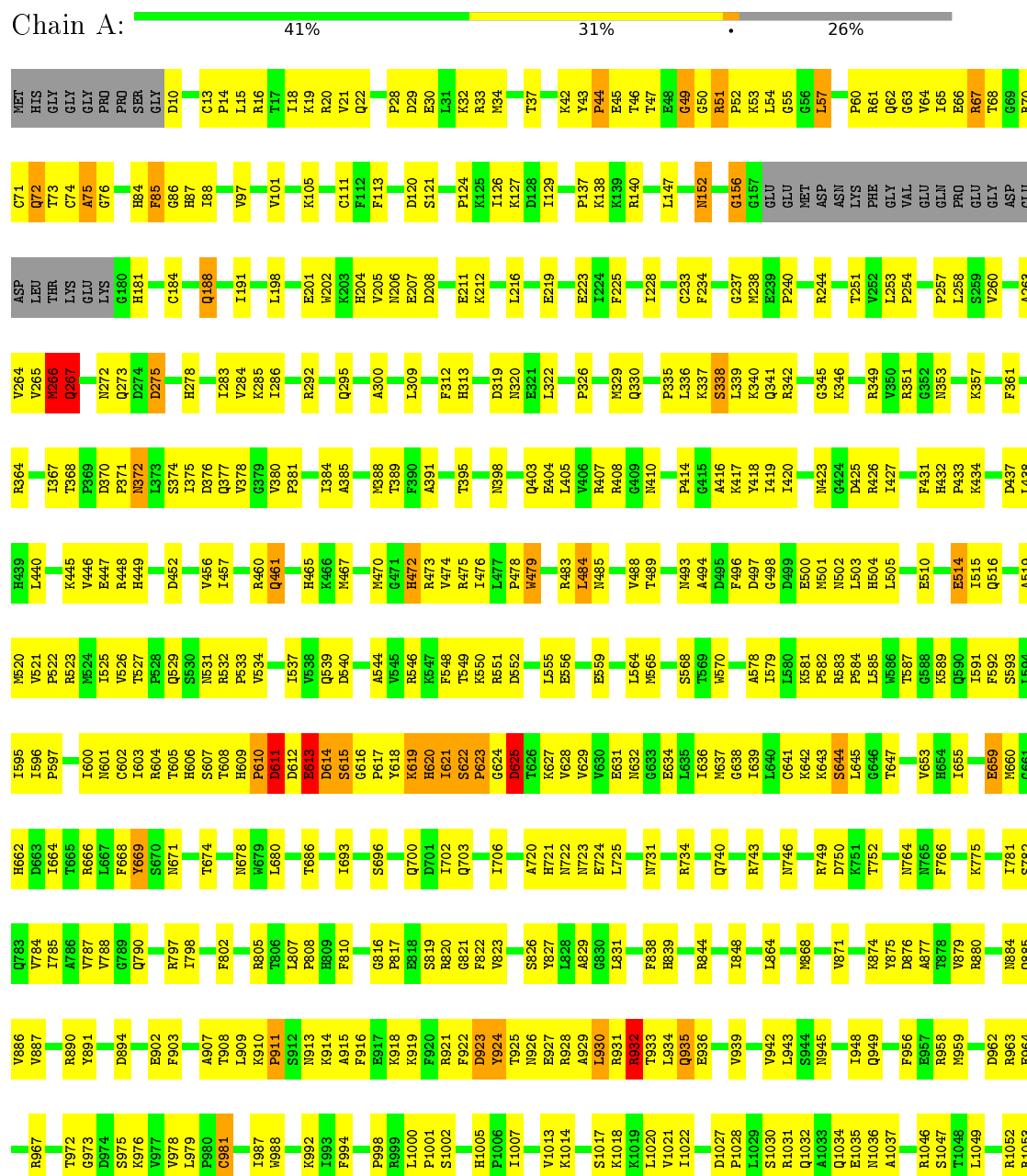
- Molecule 31 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
31	J	1	Total	Zn	0
			1	1	
31	Q	1	Total	Zn	0
			1	1	
31	B	1	Total	Zn	0
			1	1	
31	I	2	Total	Zn	0
			2	2	
31	C	1	Total	Zn	0
			1	1	
31	A	2	Total	Zn	0
			2	2	
31	U	1	Total	Zn	0
			1	1	
31	L	1	Total	Zn	0
			1	1	
31	M	1	Total	Zn	0
			1	1	

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. 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. 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: DNA-directed RNA polymerase II subunit RPB1

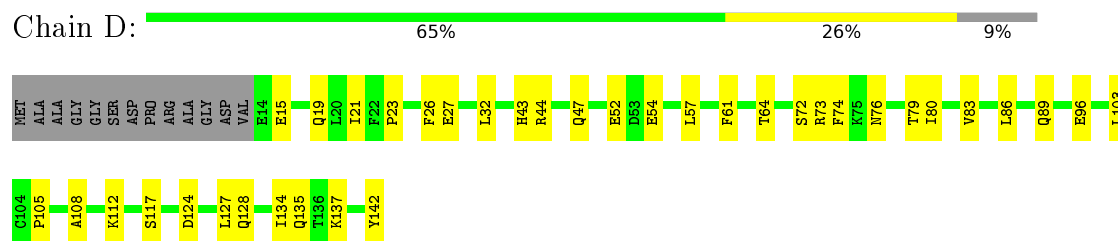


- Molecule 2: DNA-directed RNA polymerase II subunit RPB2

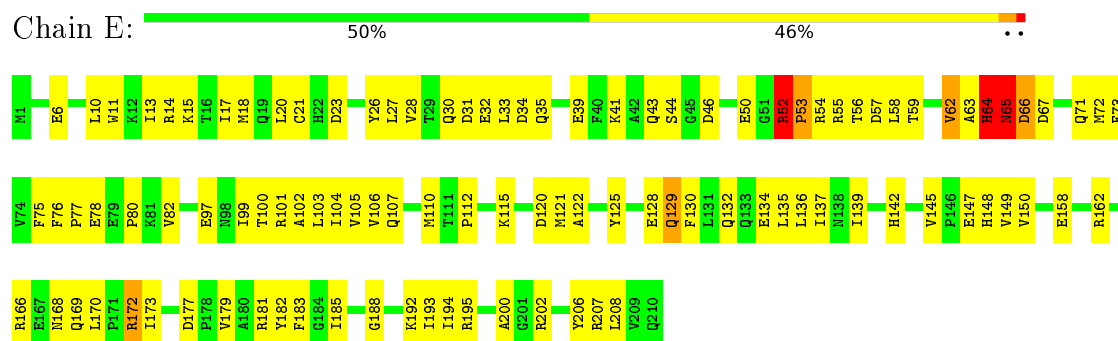
I152	P153	I154	M155	L156	L157	Y160	C161	M164	G165	L166	R169	L170	L171	C172	E173	L174	M175	E176	P181	G182	G183	Y184	G189	K192	V193	L194	E198	A201	V205	A209	K210	K211	A216	R222	S223	C224	L225	E226	N227	S228	R230	T232	K236	P237										
V78	E79	P81	P82	R83	Y84	L85	K86	F88	E89	Q90	I91	Y92	L93	S94	T97	H98	W99	E100	R101	D102	P107	M108	M109	P110	N111	E112	R116	N117	Y120	S121	A122	P123	L124	Y125	V126	D127	L128	T129	K130	T131	V132	K134	E135	G136	E137	E138	Q139	Q145	F148					
MET	TVR	ASP	ALA	ASP	GLU	ASP	ASP	MET	MET	GLN	Y10	D11	E12	L21	W22	Q23	E24	A25	C26	W27	L28	V29	I30	E36	K37	V40	R41	Q42	S46	F50	I51	Q52	M53	S54	V55	Q56	R57	E60	D61	A62	P63	P64	I65	D66	K67	Q68	A69	E70	A71	H72	A74	S75	G76	E77

S238	L318	D428	Q500	1596	1665	V759	R834	T901	Q988	Q1068	N1142	L149	A218
R239	G322	F429	L501	1597	D666	T760	E835	G902	R988	R1072	Q1142	I150	N223
L240	A241	I430	H502	V598	E669	T761	R836	D905	Y989	Q1073	Q1145	Y151	
G243	A326	E432	L505	E670	E670	S763	C838	T910	A991	P1074	I1146	R228	
G244	T331	L433	V509	E600	E671	M764	G839	T910	A992	M1075	S1147	R229	
Q245	K334	L434	A512	E602	T672	T765	R840	G915	Y993	S1079	L1148	Y230	
K248	K336	L435	E513	S602	V673	T766	R841	G916	G994	R1080	P1152	T231	
S250	V342	L439	E519	1604	A674	L767	H842	Y916	A999	G1080	Y1153	N232	
A251	K345	L440	A519	1607	Q683	P773	Y845	1921	T1000	G1081	L1158	R234	
L252	E346	S441	E521	E608	E684	T777	L848	R922	P1001	G1082	L1163	S235	
G253	E346	D442	L522	E609	K695	T778	D849	R924	F1083	L1084	M1167	S238	
Q254	P349	K445	V523	H610	K696	T779	D850	R927	F1086	L1085	R1170	L239	
R255	T359	Y446	K524	1612	A688	T780	D851	P929	G1087	E1090	M1171	R240	
T256	K360	G451	N525	1614	I698	V780	G852	P932	V1008	A1098	M1172	P241	
V257	Y363	K457	L526	1615	T692	Y788	P856	G932	L1015	A1099	S1173	I244	
A258	Y363	K458	A527	1620	H695	Y789	V860	G933	H1021	G1093	M1174	G168	
L259	K368	K459	E527	1621	C696	T790	S861	D933	L1022	Q1094	Y1174	F169	
K264	V369	D460	Y531	1622	E697	T791	G862	K934	R1023	H1097	K1177	K171	
P268	M368	K467	S539	1623	I698	T792	D863	A936	G1024	G1098	R1178	K172	
L270	E370	K468	L542	1626	I711	T795	V865	H939	E1026	A1100	R1179	H173	
L271	R371	H460	E543	1627	P712	T796	L866	K942	Y1027	Q1101	S1179	W176	
V272	L372	H461	F544	1633	P714	T797	G868	K943	L1028	F1102	Y1180	N177	
F273	L373	R463	L545	1634	P715	T798	G869	T944	Y1029	L1103	Y1181	T179	
R274	L374	R463	E546	1635	P716	T799	K869	C945	M1030	L1106	Y1182	A183	
L275	G378	K467	E547	1636	P717	T800	T870	G946	G1031	L1107	Y1183	F184	
L276	R379	K467	E548	1637	P718	T801	L873	1947	T1033	A1110	Y1184	R185	
G277	D384	K472	S549	1640	P719	T802	P874	P955	G1034	S1111	Y1185	R186	
V279	R385	L473	E551	1643	P720	T803	E875	P956	K1035	D1112	Y1186	Y196	
S280	D386	T474	E552	1646	P721	T804	N876	G960	K1036	P1113	Y1187	P198	
D281	H387	S477	E553	1649	P722	T805	E877	G965	Q1040	Y1114	Y1188	K199	
L284	Y388	H481	E554	1651	P723	T806	E878	1965	I1041	G1123	Y1189	P200	
L289	R392	L482	E555	1652	P724	T807	L880	1968	Y1045	I1124	Y1190	E201	
Y290	L395	R483	E556	1653	P725	T808	S881	1969	T1046	M1125	Y1191	E202	
D291	R405	L485	E557	1654	P726	T809	T883	1970	Y1047	A1126	Y1192	E203	
P295	M407	S497	F568	1655	P727	T810	E884	1971	Y1048	M1129	Y1193	E204	
E299	F408	P498	W573	1656	P728	T811	R885	1972	Y1049	T1130	Y1194	E205	
K300	M410	L489	W574	1657	P729	T812	R886	1973	Y1050	R1131	Y1195	S209	
V301	L412	G490	G575	1658	P730	T813	E887	1974	L1051	T1132	Y1196	E210	
P303	L413	D492	I576	1659	P731	T814	T888	1975	K1052	H1133	Y1197	L211	
E307	Y418	G493	L563	1660	P732	T815	K889	1976	H1053	H1134	Y1198	E212	
A308	A419	L494	E568	1661	P733	T816	R890	1977	S1061	Y1135	Y1199	E213	
F309	R425	L495	R497	1662	P734	T817	D891	1978	A1063	Y1136	Y1200	D214	
Q312	K427	R498	R499	1663	P735	T818	C892	1979	E1066	E1137	Y1201	E215	
				1664	P736	T819	S893	1980	P1066	R1138	Y1202	S216	
				1665	P737	T820	S894	1981	I1067	G1139	Y1203	Q217	
				1666	P738	T821	R895	1982			Y1204		
				1667	P739	T822	F896	1983			Y1205		
				1668	P740	T823	R897	1984			Y1206		
				1669	P741	T824	A898	1985			Y1207		
				1670	P742	T825	R899	1986			Y1208		
				1671	P743	T826	T890	1987			Y1209		
				1672	P744	T827	D891	1988			Y1210		
				1673	P745	T828	C892	1989			Y1211		
				1674	P746	T829	S893	1990			Y1212		
				1675	P747	T830	R894	1991			Y1213		
				1676	P748	T831	F895	1992			Y1214		
				1677	P749	T832	R896	1993			Y1215		
				1678	P750	T833	R897	1994			Y1216		
				1679	P751	T834	A898	1995			Y1217		
				1680	P752	T835	R899	1996			Y1218		
				1681	P753	T836	T890	1997			Y1219		
				1682	P754	T837	D891	1998			Y1220		
				1683	P755	T838	C892	1999			Y1221		
				1684	P756	T839	S893	2000			Y1222		
				1685	P757	T840	R894	2001			Y1223		
				1686	P758	T841	F895	2002			Y1224		
				1687	P759	T842	R896	2003			Y1225		
				1688	P760	T843	R897	2004			Y1226		
				1689	P761	T844	A898	2005			Y1227		
				1690	P762	T845	R899	2006			Y1228		
				1691	P763	T846	T890	2007			Y1229		
				1692	P764	T847	D891	2008			Y1230		
				1693	P765	T848	C892	2009			Y1231		
				1694	P766	T849	S893	2010			Y1232		
				1695	P767	T850	R894	2011			Y1233		
				1696	P768	T851	F895	2012			Y1234		
				1697	P769	T852	R896	2013			Y1235		
				1698	P770	T853	R897	2014			Y1236		
				1699	P771	T854	A898	2015			Y1237		
				1700	P772	T855	R899	2016			Y1238		
				1701	P773	T856	T890	2017			Y1239		
				1702	P774	T857	D891	2018			Y1240		
				1703	P775	T858	C892	2019			Y1241		
				1704	P776	T859	S893	2020			Y1242		
				1705	P777	T860	R894	2021			Y1243		
				1706	P778	T861	F895	2022			Y1244		
				1707	P779	T862	R896	2023			Y1245		
				1708	P780	T863	R897	2024			Y1246		
				1709	P781	T864	A898	2025			Y1247		
				1710	P782	T865	R899	2026			Y1248		
				1711	P783	T866	T890	2027			Y1249		
				1712	P784	T867	D891	2028			Y1250		
				1713	P785	T868	C892	2029			Y1251		
				1714	P786	T869	S893	2030			Y1252		
				1715	P787	T870	R894	2031			Y1253		
				1716	P788	T871	F895	2032			Y1254		
				1717	P789	T872	R896	2033			Y1255		
				1718	P790	T873	R897	2034			Y1256		
				1719	P791	T874	A898	2035			Y1257		
				1720	P792	T875	R899	2036			Y1258		
				1721	P793	T876	T890	2037			Y1259		
				1722	P794	T877	D891	2038			Y1260		
				1723	P795	T878	C892	2039			Y1261		
				1724	P796	T879	S893	2040			Y1262		
				1725	P797	T880	R894	2041			Y1263		
				1726	P798	T881	F895	2042			Y1264		
				1727	P799	T882	R896	2043			Y1265		
				1728	P800	T883	R897	2044			Y1266		
				1729	P801	T884	A898	2045			Y1267		
				1730	P802	T885	R899	2046			Y1268		
				1731	P803	T886	T890	2047			Y1269		
				1732	P804	T887	D891	2048			Y1270		
				1733	P805	T888	C892	2049			Y1271		
				1734	P806	T889	S893	2050			Y1272		
				1735	P807	T890	R894	2051			Y1273		
				1736	P808	T891	F895	2052			Y1274		
				1737	P809	T892	R896	2053			Y1275		
				1738	P810	T893	R897	2054			Y1276		
				1739	P811	T894	A898	2055			Y1277		
				1740	P812	T895	R899	2056			Y1278		
				1741	P813	T896	T890	2057			Y1279		
				1742	P814	T897	D891	2058			Y1280		
				1743	P815	T898	C892	2059			Y1281		
				1744	P816	T899	S893	2060			Y1282		
				1745	P817	T900	R894	2061			Y1283		
				1746	P818	T901	F895	2062			Y1284		
				1747	P819	T902	R896	2063			Y1285		
				1748	P820	T903	R897	2064			Y1286		
				1749	P821	T904	A898	20					

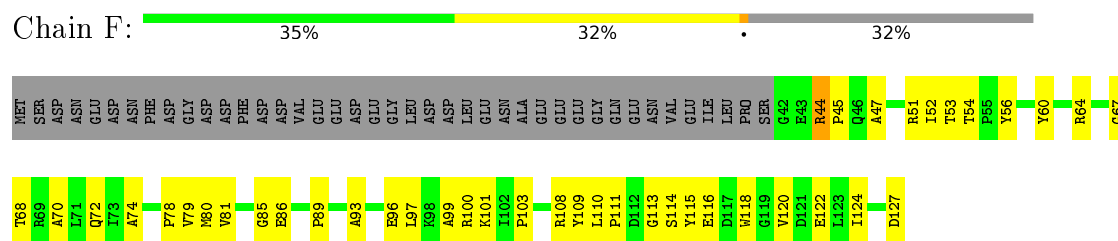
- Molecule 4: DNA-directed RNA polymerase II subunit RPB4



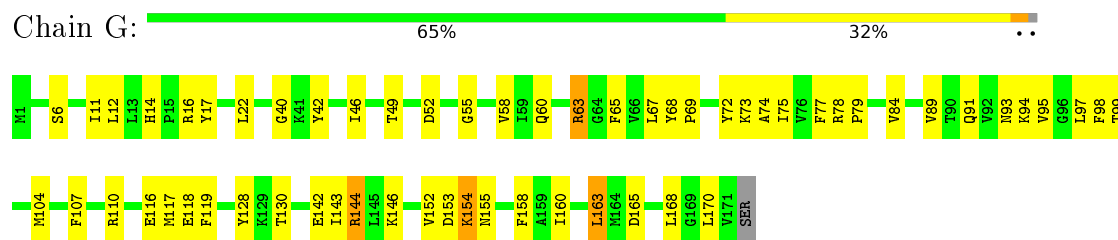
- Molecule 5: DNA-directed RNA polymerase II subunit RPB5



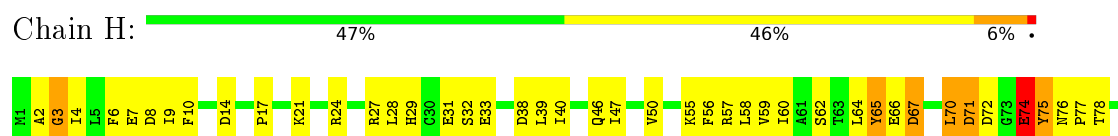
- Molecule 6: DNA-directed RNA polymerase II subunit RPB6



- Molecule 7: DNA-directed RNA polymerase II subunit RPB7



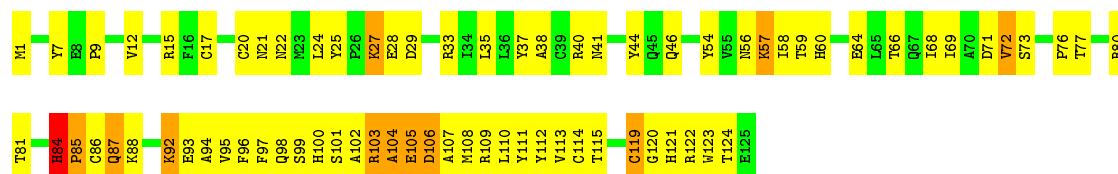
- Molecule 8: DNA-directed RNA polymerase II subunit RPB8





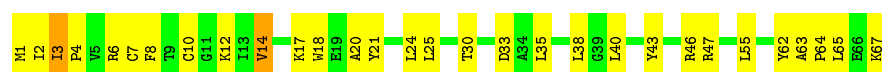
• Molecule 9: DNA-directed RNA polymerase II subunit RPB9

Chain I: 41% 50% 9% .



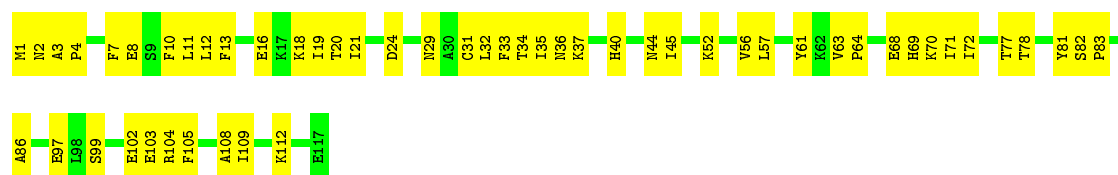
• Molecule 10: DNA-directed RNA polymerase II subunit RPB10

Chain J: 55% 42% .



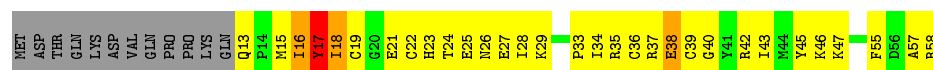
• Molecule 11: DNA-directed RNA polymerase II subunit RPB11-a

Chain K: 55% 45%



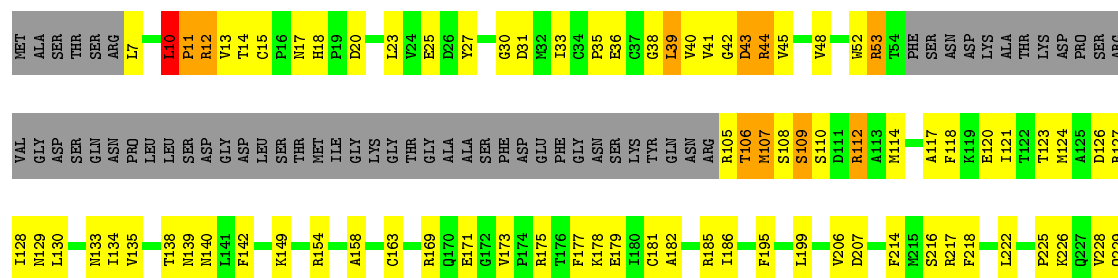
• Molecule 12: DNA-directed RNA polymerase II subunit RPB12

Chain L: 26% 47% 5% . 21%

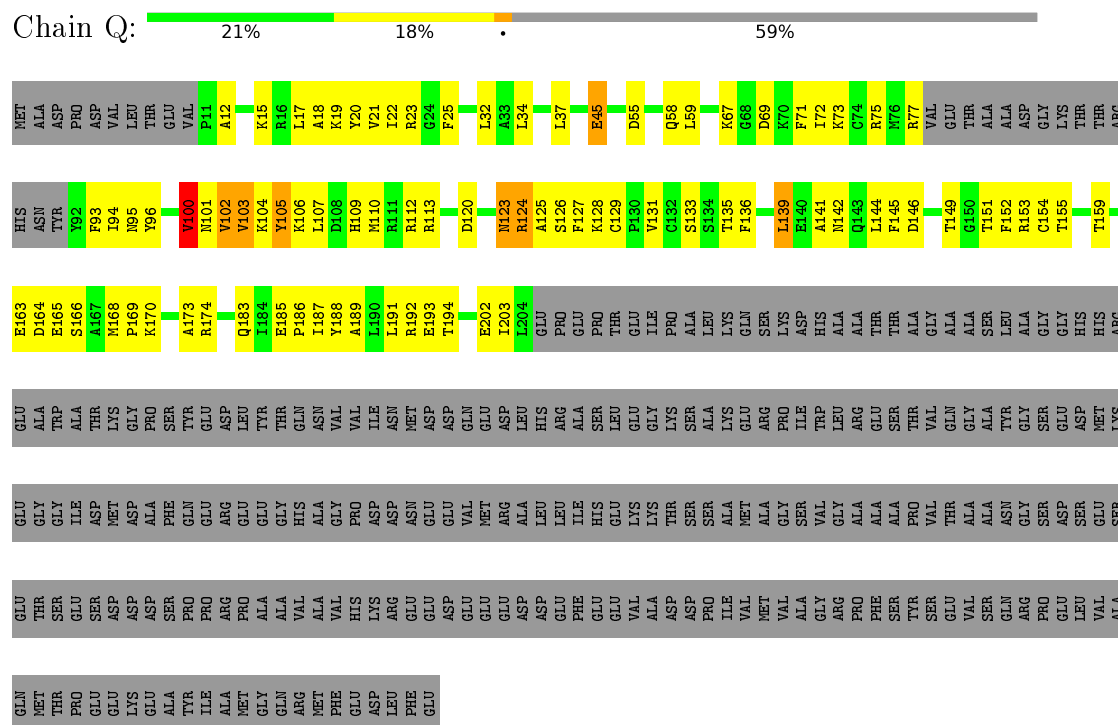


• Molecule 13: Transcription initiation factor IIB

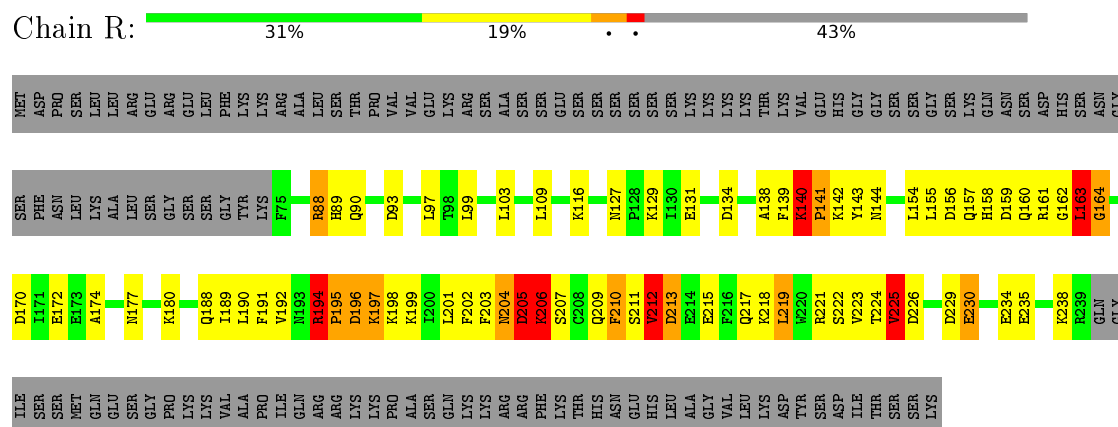
Chain M: 48% 30% . 18%



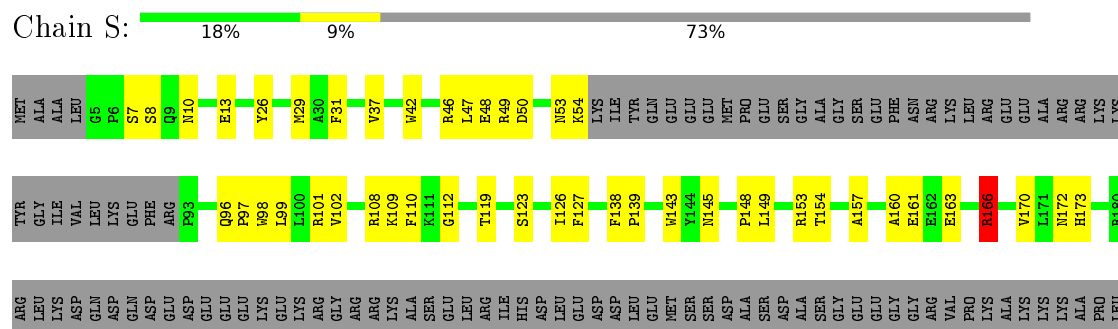
- Molecule 17: General transcription factor IIE subunit 1



- Molecule 18: Transcription initiation factor IIE subunit beta



- Molecule 19: General transcription factor IIF subunit 1



[illegible]

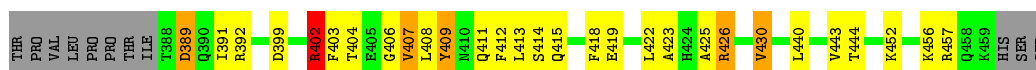
Chain 1:

Amino Acid	Percentage
M1	17%
V8	61%
L9	61%
I10	61%
E11	61%
C12	61%
D13	61%
P14	61%
A15	61%
M16	61%
K17	61%
Q18	61%
F19	61%
L20	61%
L21	61%
Y22	61%
L23	61%
D24	61%
E25	61%
A28	61%
L29	61%
G30	61%
K31	61%
K32	61%
F33	61%
I34	61%
I35	61%
Q36	61%
D37	61%
I38	61%
D39	61%
D40	61%
V43	61%
F44	61%
V45	61%
I46	61%
A47	61%
E48	61%
L49	61%
V50	61%
N51	61%
V52	61%
L53	61%
Q54	61%
E55	61%
R56	61%
V57	61%
G58	61%
E59	61%
D62	61%
G63	61%
A64	61%

Chain 2:

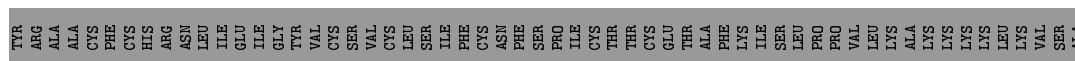
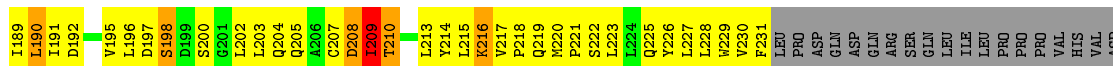
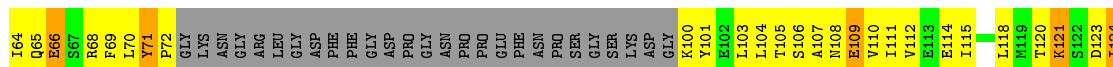
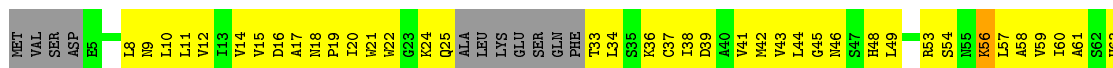
21% 35% 41%

Amino Acid	Percentage
MET	1%
GLU	1%
SER	1%
THR	1%
PRO	1%
SER	1%
ARG	1%
GLY	1%
LEU	1%
ASN	1%
ARG	1%
VAL	1%
HIS	1%
LEU	1%
GLN	1%
CYS	1%
ARG	1%
ASN	1%
P28	1%
G29	1%
V30	1%
L31	1%
L34	1%
Y35	1%
G36	1%
H37	1%
P38	1%
C41	1%
L42	1%
A43	1%
V44	1%
F45	1%
R46	1%
E47	1%
L48	1%
P49	1%
S50	1%
L51	1%
A52	1%
K53	1%
N54	1%
V55	1%
V56	1%
M57	1%
R58	1%
M59	1%
L60	1%
R61	1%
L62	1%



• Molecule 27: General transcription factor IIH subunit 3

Chain 3: 15% 42% 5% 37%



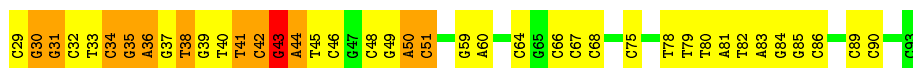
• Molecule 28: SCP-X

Chain X: 45% 34% 18%



• Molecule 29: SCP-Y

Chain Y: 38% 43% 17%



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	34728	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	Not provided	Depositor
Minimum defocus (nm)	2000	Depositor
Maximum defocus (nm)	4000	Depositor
Magnification	27500	Depositor
Image detector	Not provided	Depositor

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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

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	$\# Z > 2$	RMSZ	$\# Z > 2$
1	A	0.33	5/11727 (0.0%)	0.66	23/15833 (0.1%)
10	J	0.27	0/542	0.56	0/730
11	K	0.26	0/956	0.52	0/1294
12	L	0.28	0/394	0.65	1/524 (0.2%)
13	M	0.26	0/2049	0.69	2/2769 (0.1%)
14	N	0.31	0/945	0.58	1/1274 (0.1%)
15	O	0.26	0/816	0.54	0/1105
16	P	0.26	0/1489	0.60	2/2005 (0.1%)
17	Q	0.28	0/1507	0.59	2/2023 (0.1%)
18	R	0.44	0/1380	0.87	2/1854 (0.1%)
19	S	0.29	0/1167	0.53	1/1576 (0.1%)
2	B	0.29	1/9503 (0.0%)	0.63	4/12831 (0.0%)
20	T	0.27	0/1817	0.57	0/2445
21	U	0.33	0/1358	0.66	2/1820 (0.1%)
22	V	1.40	13/3931 (0.3%)	1.87	96/5298 (1.8%)
23	W	1.50	22/5460 (0.4%)	2.00	154/7390 (2.1%)
24	0	1.49	5/1506 (0.3%)	1.95	43/2038 (2.1%)
25	1	0.83	0/496	1.15	1/669 (0.1%)
26	2	0.88	0/2243	1.18	8/3024 (0.3%)
27	3	0.85	0/1548	1.22	6/2090 (0.3%)
28	X	1.32	18/1510 (1.2%)	1.75	58/2332 (2.5%)
29	Y	1.26	15/1472 (1.0%)	1.71	52/2267 (2.3%)
3	C	0.27	0/2259	0.67	2/3073 (0.1%)
4	D	0.28	0/1077	0.51	0/1446
5	E	0.27	0/1753	0.66	2/2368 (0.1%)
6	F	0.25	0/700	0.51	0/946
7	G	0.27	0/1382	0.55	0/1874
8	H	0.26	0/1227	0.64	1/1654 (0.1%)
9	I	0.25	0/1038	0.90	1/1407 (0.1%)
All	All	0.75	79/63252 (0.1%)	1.09	464/85959 (0.5%)

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	A	0	1
17	Q	0	1
18	R	0	8
20	T	0	1
22	V	0	8
23	W	0	11
24	0	0	1
25	1	0	1
26	2	0	8
28	X	0	8
29	Y	0	6
All	All	0	54

The worst 5 of 79 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
29	Y	51	DC	O3'-P	-14.92	1.43	1.61
28	X	53	DA	P-O5'	-9.87	1.49	1.59
28	X	61	DA	C5'-C4'	8.65	1.60	1.51
29	Y	38	DT	P-O5'	8.07	1.67	1.59
23	W	158	TYR	CE1-CZ	8.04	1.49	1.38

The worst 5 of 464 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	I	84	HIS	C-N-CD	-24.67	66.32	120.60
24	0	77	LYS	C-N-CD	-21.82	72.60	120.60
3	C	6	GLN	C-N-CD	-21.15	74.06	120.60
13	M	10	LEU	C-N-CD	-20.90	74.63	120.60
27	3	71	TYR	C-N-CD	-20.67	75.13	120.60

There are no chirality outliers.

5 of 54 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	85	PHE	Peptide
17	Q	100	VAL	Mainchain
18	R	204	ASN	Mainchain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
18	R	205	ASP	Mainchain,Peptide
18	R	206	LYS	Mainchain

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	11515	0	11609	697	0
2	B	9317	0	9305	579	0
3	C	2213	0	2153	156	0
4	D	1062	0	1042	24	0
5	E	1723	0	1744	117	0
6	F	689	0	715	46	0
7	G	1351	0	1358	53	0
8	H	1205	0	1167	91	0
9	I	1013	0	932	93	0
10	J	533	0	553	51	0
11	K	937	0	959	47	0
12	L	388	0	393	70	0
13	M	2018	0	2059	132	0
14	N	930	0	888	68	0
15	O	806	0	818	50	0
16	P	1462	0	1549	112	0
17	Q	1484	0	1496	230	0
18	R	1357	0	1377	299	0
19	S	1138	0	1103	39	0
20	T	1788	0	1819	171	0
21	U	1343	0	1338	100	0
22	V	3855	0	3872	221	0
23	W	5348	0	5372	179	0
24	0	1479	0	1524	39	0
25	1	491	0	507	239	0
26	2	2196	0	2206	595	0
27	3	1526	0	1561	471	0
28	X	1343	0	725	35	0
29	Y	1316	0	730	36	0
30	A	1	0	0	0	0
30	B	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
31	A	2	0	0	0	0
31	B	1	0	0	0	0
31	C	1	0	0	0	0
31	I	2	0	0	0	0
31	J	1	0	0	0	0
31	L	1	0	0	0	0
31	M	1	0	0	0	0
31	Q	1	0	0	0	0
31	U	1	0	0	0	0
All	All	61839	0	60874	4132	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 34.

The worst 5 of 4132 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)
23:W:421:PHE:CD1	23:W:431:PRO:HG3	1.15	1.65
5:E:27:LEU:HD12	5:E:64:HIS:CD2	1.29	1.64
27:3:59:VAL:HG12	27:3:71:TYR:CD1	1.24	1.64
5:E:27:LEU:HB2	5:E:64:HIS:CD2	1.33	1.63
22:V:315:VAL:HG13	23:W:500:ASP:CB	1.21	1.63

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 PDB entries followed by that with respect to all EM entries.

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	1450/1970 (74%)	1305 (90%)	97 (7%)	48 (3%)	5	40
2	B	1163/1174 (99%)	1049 (90%)	76 (6%)	38 (3%)	5	40

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	C	273/275 (99%)	241 (88%)	18 (7%)	14 (5%)	2	30
4	D	127/142 (89%)	118 (93%)	8 (6%)	1 (1%)	24	69
5	E	208/210 (99%)	195 (94%)	7 (3%)	6 (3%)	6	43
6	F	84/127 (66%)	78 (93%)	4 (5%)	2 (2%)	7	47
7	G	169/172 (98%)	158 (94%)	10 (6%)	1 (1%)	30	74
8	H	148/150 (99%)	123 (83%)	13 (9%)	12 (8%)	1	19
9	I	123/125 (98%)	100 (81%)	14 (11%)	9 (7%)	1	21
10	J	65/67 (97%)	53 (82%)	9 (14%)	3 (5%)	3	32
11	K	115/117 (98%)	112 (97%)	3 (3%)	0	100	100
12	L	44/58 (76%)	37 (84%)	3 (7%)	4 (9%)	1	17
13	M	256/316 (81%)	236 (92%)	12 (5%)	8 (3%)	5	42
14	N	109/376 (29%)	100 (92%)	5 (5%)	4 (4%)	4	37
15	O	97/109 (89%)	90 (93%)	7 (7%)	0	100	100
16	P	183/339 (54%)	170 (93%)	8 (4%)	5 (3%)	6	45
17	Q	176/439 (40%)	159 (90%)	11 (6%)	6 (3%)	5	40
18	R	163/291 (56%)	128 (78%)	22 (14%)	13 (8%)	1	19
19	S	134/517 (26%)	123 (92%)	7 (5%)	4 (3%)	5	42
20	T	218/249 (88%)	191 (88%)	17 (8%)	10 (5%)	3	32
21	U	168/301 (56%)	136 (81%)	21 (12%)	11 (6%)	1	25
22	V	473/782 (60%)	400 (85%)	46 (10%)	27 (6%)	2	27
23	W	661/760 (87%)	567 (86%)	69 (10%)	25 (4%)	4	37
24	0	186/395 (47%)	168 (90%)	13 (7%)	5 (3%)	6	45
25	1	60/71 (84%)	53 (88%)	5 (8%)	2 (3%)	5	40
26	2	264/462 (57%)	246 (93%)	14 (5%)	4 (2%)	13	57
27	3	187/308 (61%)	175 (94%)	9 (5%)	3 (2%)	12	56
All	All	7304/10302 (71%)	6511 (89%)	528 (7%)	265 (4%)	7	38

5 of 265 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	67	ARG
1	A	205	VAL
1	A	266	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	267	GLN
1	A	531	ASN

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 PDB entries followed by that with respect to all EM entries.

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	1279/1748 (73%)	1247 (98%)	32 (2%)	55	81
2	B	1020/1028 (99%)	997 (98%)	23 (2%)	58	82
3	C	252/252 (100%)	246 (98%)	6 (2%)	57	82
4	D	119/126 (94%)	118 (99%)	1 (1%)	86	94
5	E	192/192 (100%)	186 (97%)	6 (3%)	47	77
6	F	74/111 (67%)	74 (100%)	0	100	100
7	G	152/153 (99%)	149 (98%)	3 (2%)	63	85
8	H	131/131 (100%)	127 (97%)	4 (3%)	47	77
9	I	112/112 (100%)	106 (95%)	6 (5%)	27	64
10	J	56/56 (100%)	56 (100%)	0	100	100
11	K	106/106 (100%)	105 (99%)	1 (1%)	84	93
12	L	43/55 (78%)	43 (100%)	0	100	100
13	M	222/268 (83%)	212 (96%)	10 (4%)	34	69
14	N	105/324 (32%)	104 (99%)	1 (1%)	82	92
15	O	90/98 (92%)	89 (99%)	1 (1%)	80	91
16	P	159/293 (54%)	154 (97%)	5 (3%)	47	77
17	Q	164/373 (44%)	157 (96%)	7 (4%)	35	70
18	R	150/261 (58%)	138 (92%)	12 (8%)	15	50
19	S	121/448 (27%)	118 (98%)	3 (2%)	55	81
20	T	196/218 (90%)	187 (95%)	9 (5%)	33	68
21	U	148/266 (56%)	139 (94%)	9 (6%)	23	60
22	V	422/688 (61%)	403 (96%)	19 (4%)	34	69

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
23	W	577/664 (87%)	541 (94%)	36 (6%)	23	60
24	0	171/352 (49%)	163 (95%)	8 (5%)	32	68
25	1	56/64 (88%)	52 (93%)	4 (7%)	18	55
26	2	238/399 (60%)	229 (96%)	9 (4%)	40	73
27	3	171/272 (63%)	159 (93%)	12 (7%)	19	56
All	All	6526/9058 (72%)	6299 (96%)	227 (4%)	47	74

5 of 227 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
17	Q	123	ASN
20	T	162	ASN
26	2	202	GLN
17	Q	193	GLU
18	R	212	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 61 such sidechains are listed below:

Mol	Chain	Res	Type
9	I	98	GLN
22	V	286	HIS
27	3	148	ASN
9	I	121	HIS
18	R	177	ASN

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 carbohydrates in this entry.

5.6 Ligand geometry

Of 13 ligands modelled in this entry, 13 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.