



# wwPDB X-ray Structure Validation Summary Report (i)

Feb 28, 2014 – 12:32 AM GMT

PDB ID : 3PYV  
Title : Crystal structure of a complex containing domain 3 of CrPV IGR IRES RNA bound to the 70S ribosome. This file contains the 50S subunit of the second 70S ribosome.  
Authors : Zhu, J.; Korostelev, A.; Costantino, D.; Noller, H.F.; Kieft, J.S.  
Deposited on : 2010-12-13  
Resolution : 3.40 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

The following versions of software and data (see [references](#)) were used in the production of this report:

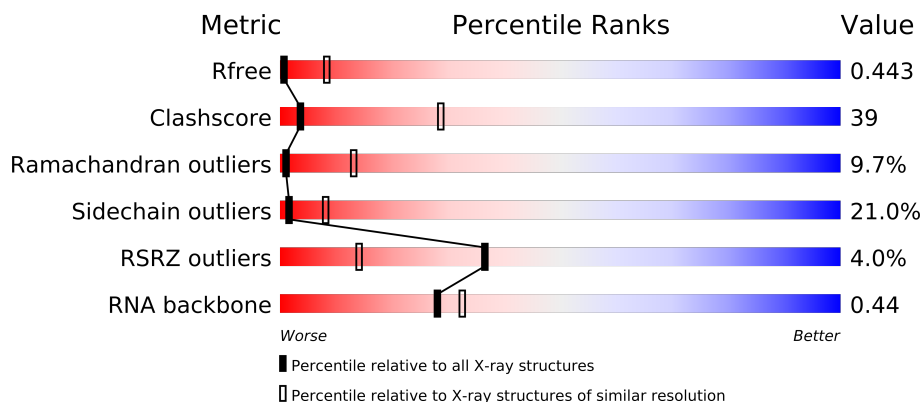
MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683



# 1 Overall quality at a glance

The reported resolution of this entry is 3.40 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))
$R_{free}$	66092	1017 (3.52-3.28)
Clashscore	79885	1214 (3.50-3.30)
Ramachandran outliers	78287	1177 (3.50-3.30)
Sidechain outliers	78261	1177 (3.50-3.30)
RSRZ outliers	66119	1017 (3.52-3.28)
RNA backbone	1838	1002 (4.02-2.76)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	2879	
2	B	119	
3	C	271	
4	D	204	
5	E	202	
6	F	181	
7	G	159	
8	H	145	
9	I	65	
10	J	137	
11	K	122	
12	L	146	

Continued on next page...



*Continued from previous page...*

Mol	Chain	Length	Quality of chain
13	M	136	
14	N	117	
15	O	98	
16	P	137	
17	Q	116	
18	R	101	
19	S	112	
20	T	92	
21	U	100	
22	V	188	
23	W	76	
24	X	88	
25	Y	62	
26	Z	59	
27	1	30	
28	2	52	
29	3	44	
30	4	48	
31	5	63	



## 2 Entry composition

There are 31 unique types of molecules in this entry. The entry contains 88276 atoms, of which 0 are hydrogen and 0 are deuterium.

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 RNA chain called 23S ribosomal RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	2760	Total	C	N	O	P	0	0	0
			59442	26456	11114	19113	2759			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1142	U	C	SEE REMARK 999	GB AE017221.1
A	2825	U	G	SEE REMARK 999	GB AE017221.1

- Molecule 2 is a RNA chain called 5S ribosomal RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	119	Total	C	N	O	P	0	0	0
			2551	1136	471	826	118			

- Molecule 3 is a protein called 50S ribosomal protein L2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	271	Total	C	N	O	S	0	0	0
			2105	1329	416	357	3			

- Molecule 4 is a protein called 50S ribosomal protein L3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	204	Total	C	N	O	S	0	0	0
			1564	988	299	271	6			

- Molecule 5 is a protein called 50S ribosomal protein L4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	202	Total	C	N	O	S	0	0	0
			1587	1011	297	276	3			



- Molecule 6 is a protein called 50S ribosomal protein L5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	181	Total	C	N	O	S	0	0	0
			1475	943	268	260	4			

- Molecule 7 is a protein called 50S ribosomal protein L6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	159	Total	C	N	O	S	0	0	0
			1223	773	228	221	1			

- Molecule 8 is a protein called 50S ribosomal protein L9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	145	Total	C	N	O	S	0	0	0
			1133	724	200	208	1			

- Molecule 9 is a protein called 50S ribosomal protein L10.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
9	I	32	Total	C	N	O	0	0	0
			254	157	49	48			

- Molecule 10 is a protein called 50S ribosomal protein L13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	137	Total	C	N	O	S	0	0	0
			1097	707	205	182	3			

- Molecule 11 is a protein called 50S ribosomal protein L14.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	122	Total	C	N	O	S	0	0	0
			932	587	171	170	4			

- Molecule 12 is a protein called 50S ribosomal protein L15.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	146	Total	C	N	O	S	0	0	0
			1114	692	227	193	2			

- Molecule 13 is a protein called 50S ribosomal protein L16.



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	M	136	Total	C	N	O	S	0	0	0
			1079	688	204	182	5			

- Molecule 14 is a protein called 50S ribosomal protein L17.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	N	117	Total	C	N	O		0	0	0
			960	599	202	159				

- Molecule 15 is a protein called 50S ribosomal protein L18.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	O	98	Total	C	N	O		0	0	0
			771	486	154	131				

- Molecule 16 is a protein called 50S ribosomal protein L19.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
16	P	137	Total	C	N	O	S	0	0	0
			1144	713	234	196	1			

- Molecule 17 is a protein called 50S ribosomal protein L20.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
17	Q	116	Total	C	N	O	S	0	0	0
			953	601	201	150	1			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
Q	?	-	PHE	DELETION	UNP Q72L76

- Molecule 18 is a protein called 50S ribosomal protein L21.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
18	R	101	Total	C	N	O	S	0	0	0
			779	501	142	135	1			

- Molecule 19 is a protein called 50S ribosomal protein L22.



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
19	S	112	Total	C	N	O	S	0	0	0
			891	560	175	154	2			

- Molecule 20 is a protein called 50S ribosomal protein L23.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
20	T	92	Total	C	N	O		0	0	0
			726	471	131	124				

- Molecule 21 is a protein called 50S ribosomal protein L24.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
21	U	100	Total	C	N	O	S	0	0	0
			776	500	148	124	4			

- Molecule 22 is a protein called 50S ribosomal protein L25.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
22	V	188	Total	C	N	O	S	0	0	0
			1492	950	265	275	2			

- Molecule 23 is a protein called 50S ribosomal protein L27.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
23	W	76	Total	C	N	O	S	0	0	0
			605	376	126	102	1			

- Molecule 24 is a protein called 50S ribosomal protein L28.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
24	X	88	Total	C	N	O		0	0	0
			695	435	141	119				

- Molecule 25 is a protein called 50S ribosomal protein L29.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
25	Y	62	Total	C	N	O	S	0	0	0
			521	325	102	92	2			

- Molecule 26 is a protein called 50S ribosomal protein L30.



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
26	Z	59	Total	C	N	O	S	0	0	0
			468	298	90	79	1			

- Molecule 27 is a protein called 50S ribosomal protein L31.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
27	1	30	Total	C	N	O	S	0	0	0
			226	142	36	44	4			

- Molecule 28 is a protein called 50S ribosomal protein L32.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
28	2	52	Total	C	N	O	S	0	0	0
			405	255	79	66	5			

- Molecule 29 is a protein called 50S ribosomal protein L33.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
29	3	44	Total	C	N	O	S	0	0	0
			381	235	77	65	4			

- Molecule 30 is a protein called 50S ribosomal protein L34.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
30	4	48	Total	C	N	O	S	0	0	0
			419	257	104	56	2			

- Molecule 31 is a protein called 50S ribosomal protein L35.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
31	5	63	Total	C	N	O	S	0	0	0
			508	326	101	79	2			

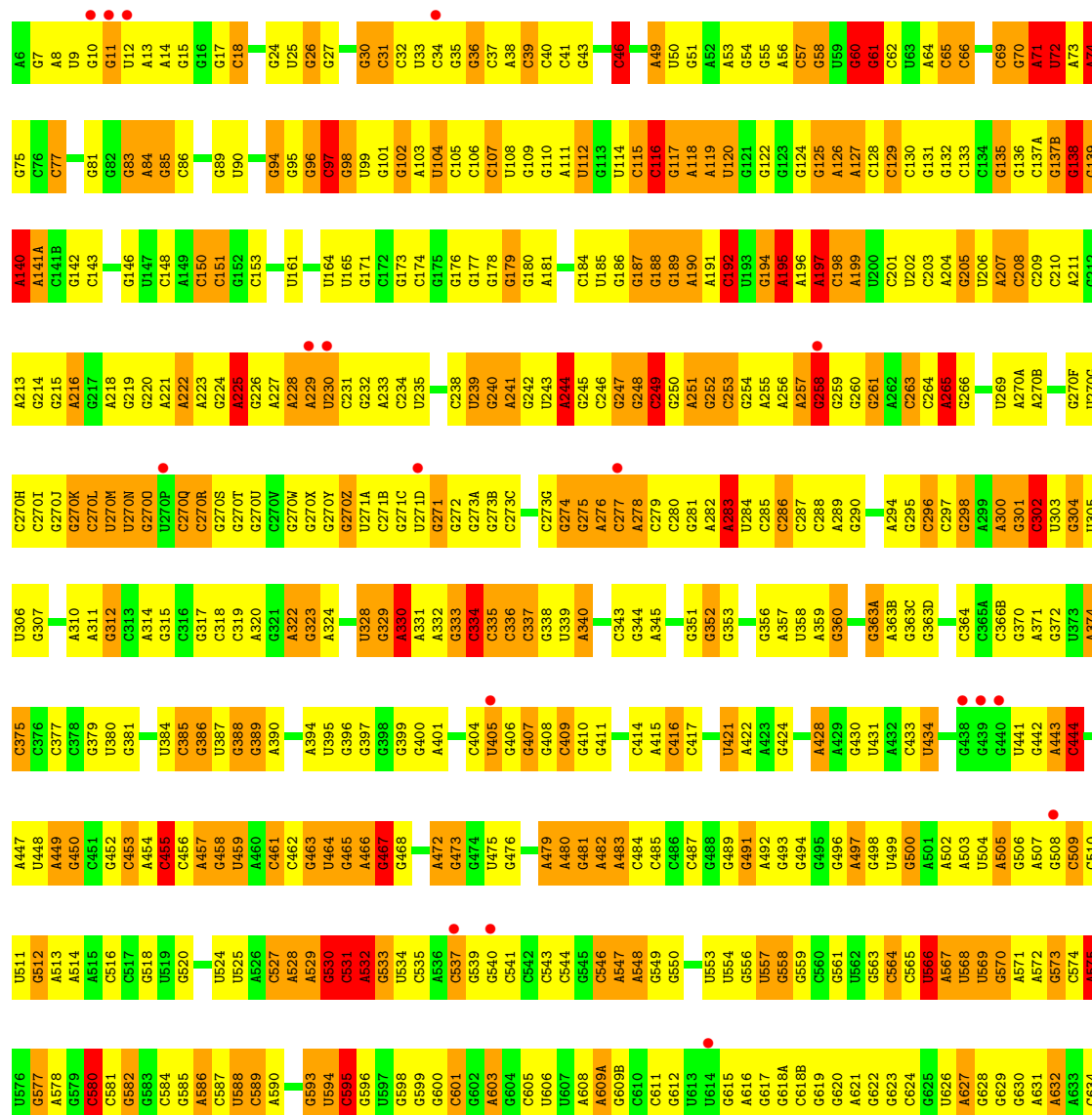


### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

#### • Molecule 1: 23S ribosomal RNA

Chain A: 



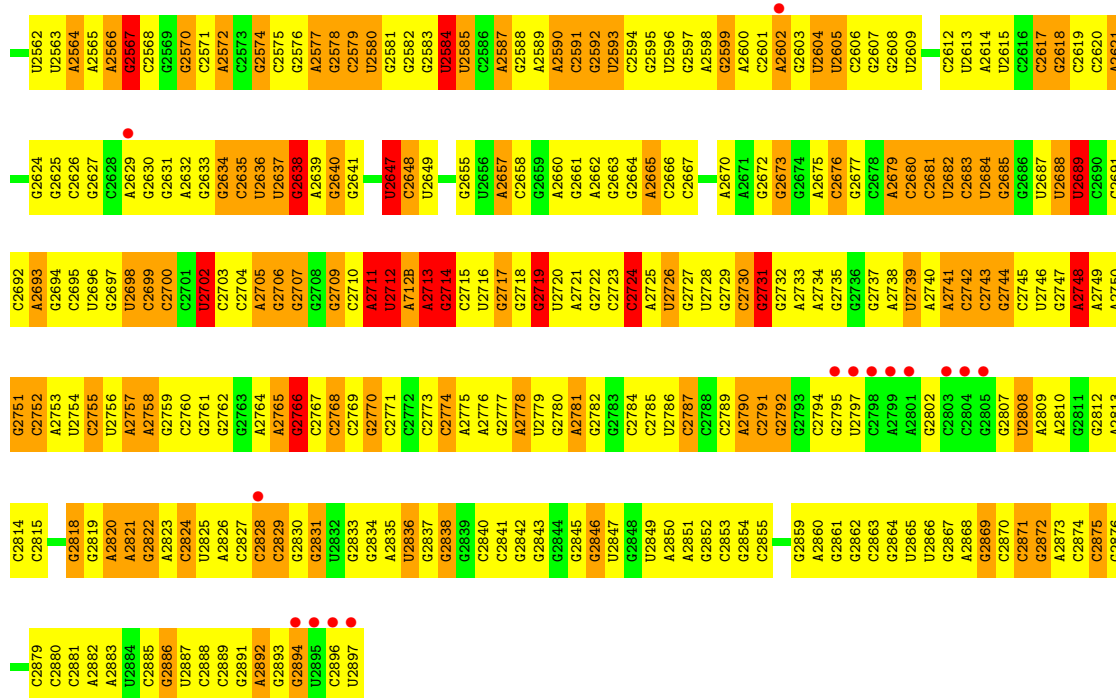






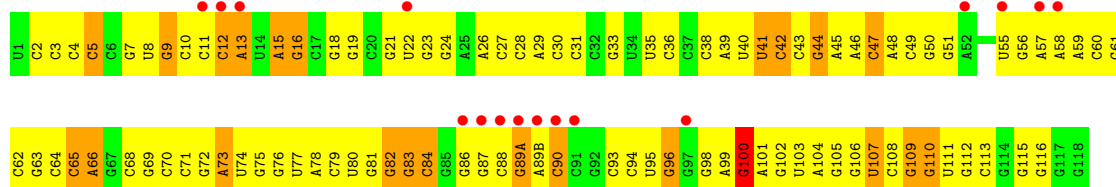
A2497	C2486	G2437	U2312	G2375	U2313	G2251	C	U	U1995	U1931	G1857	U1794	G1725	A1852	G1593
C2498	U2437	A2376	C2313	A2377	C2314	G2252	C	A	C1996	A1932	G1858	C1795	G1726	G1653	G1594
C2499	A2438	A2378	G2315	A2379	G2316	G2253	U	G	G1997	G1933	A1859	U1796	U1727	A1854	G1595
U2500	A2439	A2379	G2316	A2380	C2317	G2254	G	G	G1998	C1934	G1860	C1797	G1728	A1855	A1596
G2501	C2440	G2379	C2317	G2380	G2318	G2255	U	U	G1999	G1935	G1861	U1798	A1729	C1656	A1597
G2502	C2441	C2380	G2318	U2357	G2319	G2256	G	G	G2000	A1936	G1862	G1799	U1730	C1657	C1598
A2503	C2442	G2383	A2320	U2358	G2320	U2359	G	A	A2001	A1937	G1863	C1800	G1731	C1658	C1599
U2504	G2443	G2384	G2321	G2260	G2321	G2261	A	A	G2002	U1938	C1870	G1801	A1732	U1659	G1600
G2505	C2444	C2385	A2322	C2260	G2322	C2261	G	G	G2003	A1802	A1871	A1802	G1733	C1660	G1601
U2506	G2445	C2386	G2323	C2262	G2323	C2263	C	C	G2004	A1803	A1872	C1804	U1734	G1661	A1602
C2507	G2446	C2387	C2324	C2263	G2324	C2264	U	U	A2005	U1873	G1878	U1805	U1735	G1662	A1603
G2510	A2448	U2387	C2324	C2263	G2324	C2264	C	C	C2006	U1946	C1879	C1806	G1742	A1663	C1604
U2511	U2449	G2389	G2325	C2264	G2325	C2265	G	G	C2007	C1947	G1880		G1743	A1664	C1605
G2512	A2450	U2390	C2326		G2326	C2266	U	U	G2008	G1946	C1881		G1744	A1665	G1606
C2513	A2451	G2391	A2327	A2267	G2327	A2268	U	U	G2009	G1949	C1882	G1811	G1746	A1666	G1607
U2514	C2452	A2392	A2328	A2268	G2328	A2269	G	G	G2010	G1950	C1883	A1812	U1747	G1667	A1608
G2515	A2453	A2393	G2329	A2269	G2329	A2270	A	A	U2011	U1951	G1884	U1813	U1748	A1668	A1609
C2516	C2454	C2394	G2330	G2270	G2330	G2271	C	C	G2012	A1952	G1885	G1814	A1749	A1669	A1610
G2517	G2455	G2395	G2331	G2271	G2331	G2272	C	C	A2013	A1953	C1886	G1815	G1750	C1670	G1611
U2518	C2456	G2396	G2332	U2272	G2332	U2273	C	C	A2014	G1954	C1887	G1816		U1671	C1612
C2519	A2459	G2397	A2333	A2273	G2333	A2274	C	C	A2015	U1955	C1888	G1817	C1754	G1672	G1613
U2520	C2460	U2398	G2334	A2274	G2334	A2275	C	C	U2016	U1956	G1889	U1818	A1755	U1673	A1614
G2521	U2461	G2400	A2335	G2275	G2335	G2276	G	G	U2017	C1957	A1889	A1819	G1756	A1674	C1615
U2522	C2462	G2401	A2336	G2276	G2336	G2277	C	C	G1958	C1958	A1890	U1820	U1757	A1675	A1616
G2523	U2463	C2402	G2337	G2277	G2337	G2278	U	U	A2018	A1959	C1891	A1821	G1758	A1677	C1617
C2524	C2464	C2403	G2338	A2278	G2338	A2279	C	C	A2019	A1960	G1892	G1822	A1759	A1678	A1618
G2525	U2465	G2404	G2339	A2279	G2339	A2280	U	U	U2020	C1961	C1893	G1823	A1760	G1679	G1619
C2526	G2466	C2405	G2340	G2281	G2340	G2282	C	C	U2021	C1962	C1894	G1824	A1761	U1680	G1620
G2527	U2467	U2406	C2342	G2282	G2342	G2283	C	C	G2022	U1963	C1895	A1825	U1681	U1681	U1621
U2528	C2467	G2407	G2343	G2283	G2343	G2284	G	G	G2023	G1964	C1896	G1826	G1763	G1682	G1622
G2529	U2468	U2408	U2344	G2284	G2344	G2285	U	U	G2024	C1965	C1897	C1827	G1764	G1683	G1623
A2530	A2469	G2409	G2345	C2284	G2345	C2285	C	C	C2025	A1966	C1898	G1828	G1765	C1684	G1624
U2531	C2470	U2410	A2346	C2285	G2346	C2286	G	G	G2026	U1967	A1900	A1829	U1766	C1685	C1625
G2532	G2472	A2411	G2347	A2287	G2347	A2288	U	U	U2027	C1968	A1901	C1830	U1767	C1686	G1626
U2535	U2473	G2412	C2350	A2288	G2350	A2289	G	G	G2028	A1969	C1902	G1831	U1768	C1687	G1627
G2536	C2474	G2413	G2351	G2289	G2351	G2290	C	C	A2030	G1970	G1903	C1832	G1769	U1688	G1628
U2537	A2475	G2414	A2352	C2290	G2352	U2291	U	U	G2031	A1971	G1904	U1833	G1770	A1689	U1629
C2538	C2476	G2415	G2353	U2291	G2353	U2292	G	G	G2032	A1972	C1905	U1834	C1771	A1690	G1630
G2539	U2477	C2416	G2354	C2292	G2354	C2293	C	C	U2033	G1973	G1906	G1835	G1772	G1691	C1631
C2540	A2478	G2417	U2356	C2293	G2356	C2294	G	G	G2034	A1974		C1836	U1773	U1692	G1632
A2541	U2479	U2418	G2357	C2294	U2357	C2295	C	C	G2035	G1975	C1909	C1837	C1774	G1696	A1633
G2543	C2480	C2420	C2358	C2295	G2358	C2296	U	U	C2036	C1976		C1838	U1775	G1697	G1634
C2544	G2482	G2421	A2360	C2296	G2359	U2297	C	C	G2037	A1977	A1913	G1839	U1776	G1698	A1635
U2547	C2483	A2422	A2361	C2297	G2360	C2298	G	G	U2038	C1978	C1914	U1840	U1777	A1699	C1636
G2548	G2484	U2423	G2362	C2298	A2362	G2299	C	C	C2039	G1979	U1915	G1841	U1778	G1699	C1637
U2549	C2485	C2424	C2363	G2300	G2363	G2301	G	G	U2040	A1981	A1916	G1842	U1779	A1700	C1638
G2550	U2486	A2425	A2364	C2301	G2364	C2302	U	U	A2041	C1982	A1918	G1844	A1780	G1703	U1639
C2551	G2487	G2426	G2365	G2302	C2365	G2303	G	G	C2042	C1983	A1919	G1845	C1782	G1704	C1640
U2552	A2488	C2427	A2366	C2303	G2366	G2304	A	A	C2043	A1984	C1920	G1846	A1783	G1705	A1641
G2553	G2489	G2428	G2367	G2304	A2367	G2305	C	C	G2044	G1985	G1922	A1847	A1784	U1706	G1642
U2554	U2490	C2429	C2368	C2305	G2368	G2306	A	A	U2045	C1986	U1923	A1848	A1785	G1707	G1643
C2555	C2491	G2430	A2369	G2306	G2369	G2307	U	U	G2046	G1987	G1924	G1849	A1786	C1708	G1644
U2556	U2492	U2431	G2370	C2307	G2370	G2308	C	C	G2047	C1988	C1925	U1851	A1787	U1709	G1645
U2493	U2493	A2432	G2371	G2308	G2371	G2309	A	A	G2048	G1989	U1926	G1852	C1788	C1710	G1646
G2494	G2494	A2433	G2372	A2309	G2372	A2310	C	C	A2049	C1990	A1927	A1853	C1789	C1711	G1647
C2495	C2495	A2434	G2373	G2310	G2373	A2311	C	C	A2051	U1991	A1928	A1854	C1790	C1712	G1648
G2496	C2496	A2435	G2374	G2311	G2374	A2312	A	A	G2052	G1992	A1929	G1855	U1716	U1716	G1649
									G2053	U1993	G1929	G1856	G1792	G1717	G1650
									A2054	C1994	G1930	G1856	C1793	G1718	G1651





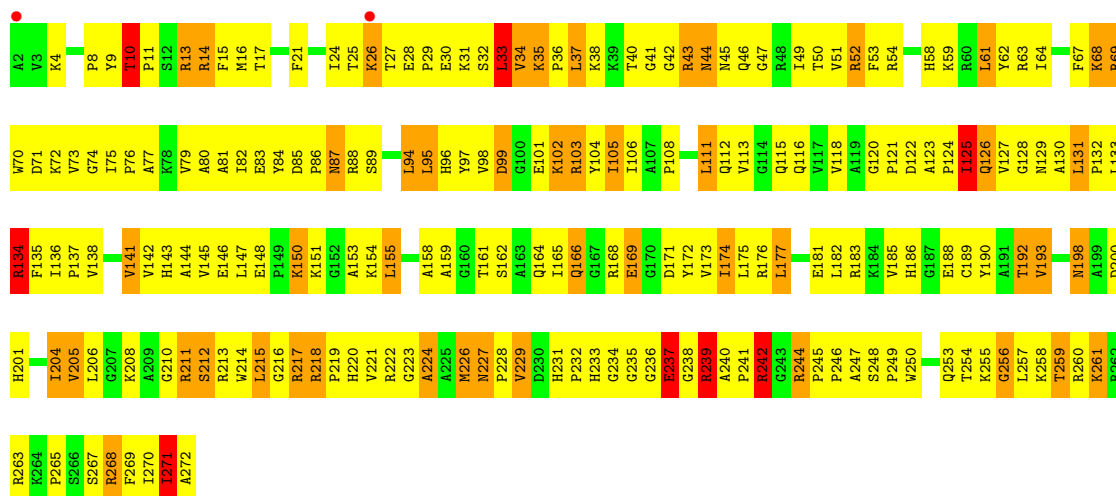
• Molecule 2: 5S ribosomal RNA

Chain B:



• Molecule 3: 50S ribosomal protein L2

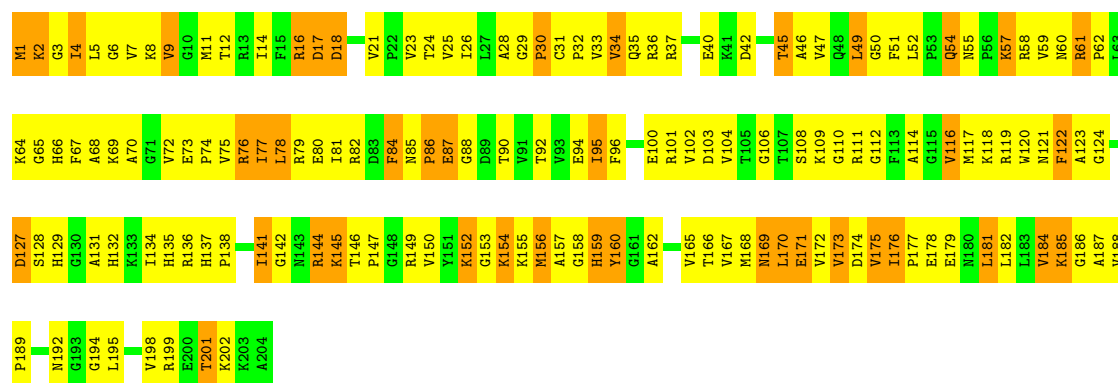
Chain C:



• Molecule 4: 50S ribosomal protein L3

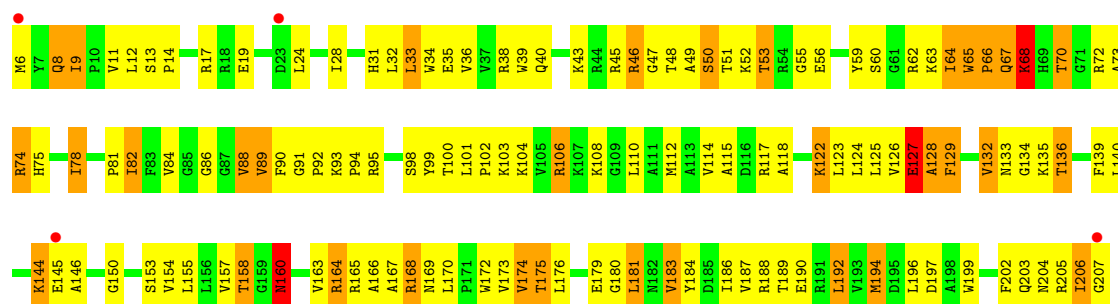


## Chain D:



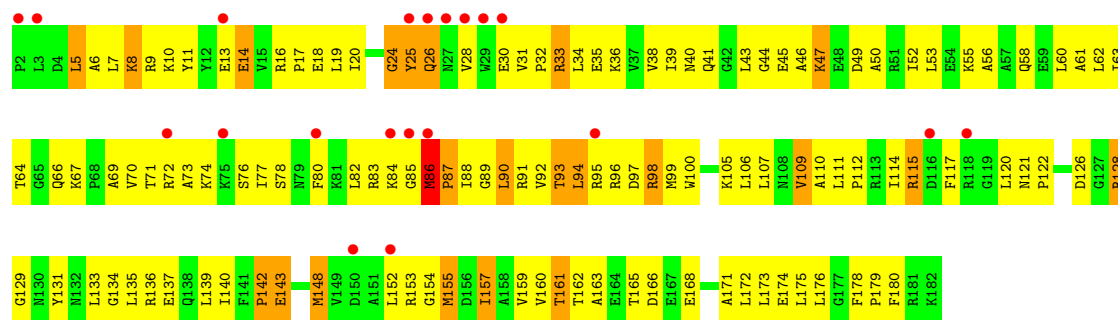
## • Molecule 5: 50S ribosomal protein L4

## Chain E:



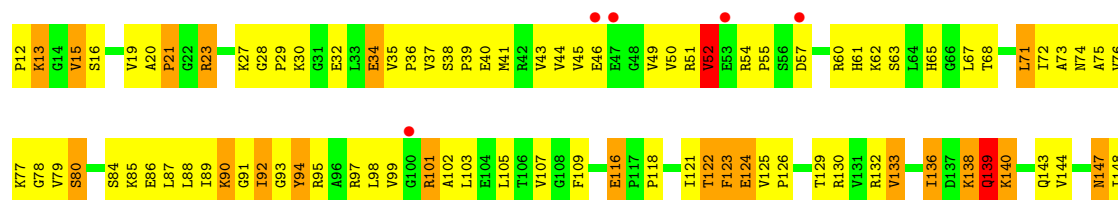
## • Molecule 6: 50S ribosomal protein L5

## Chain F:



## • Molecule 7: 50S ribosomal protein L6

## Chain G:

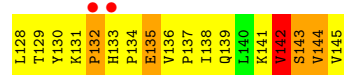






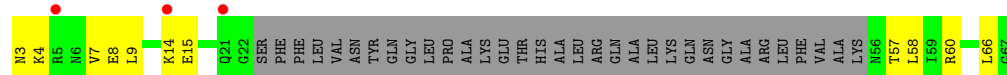
• Molecule 8: 50S ribosomal protein L9

Chain H:



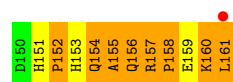
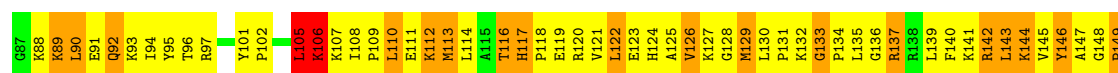
• Molecule 9: 50S ribosomal protein L10

Chain I:



• Molecule 10: 50S ribosomal protein L13

Chain J:



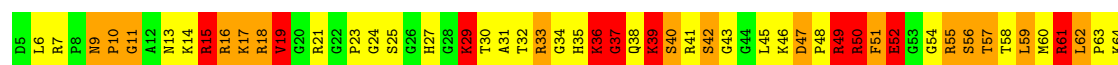
• Molecule 11: 50S ribosomal protein L14

Chain K:

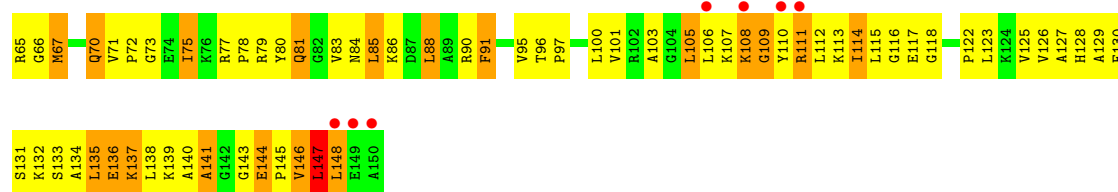


• Molecule 12: 50S ribosomal protein L15

Chain L:

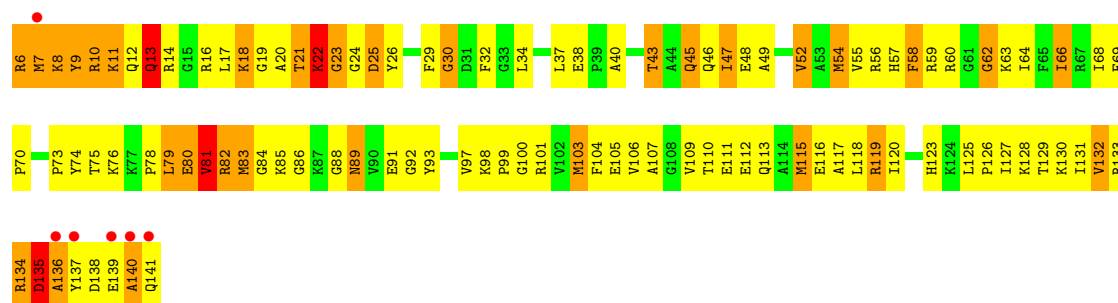






• Molecule 13: 50S ribosomal protein L16

Chain M:



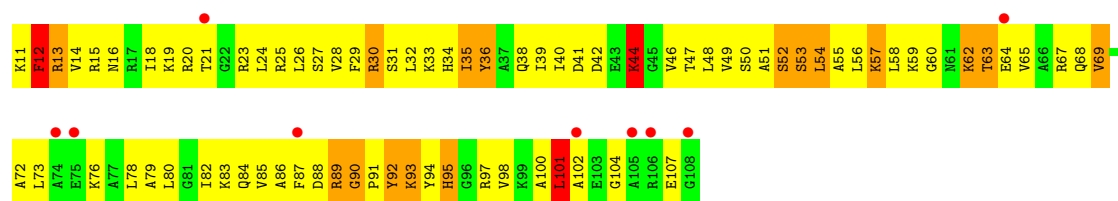
• Molecule 14: 50S ribosomal protein L17

Chain N:



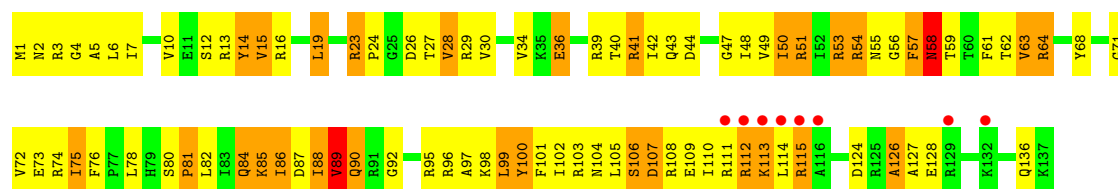
• Molecule 15: 50S ribosomal protein L18

Chain O:



• Molecule 16: 50S ribosomal protein L19

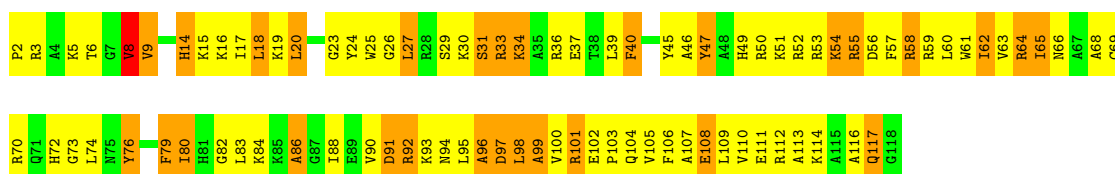
Chain P:



• Molecule 17: 50S ribosomal protein L20

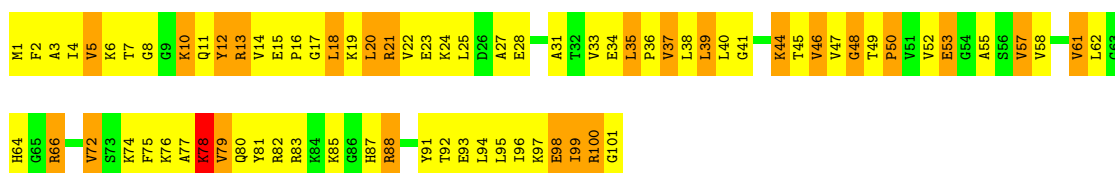
Chain Q:





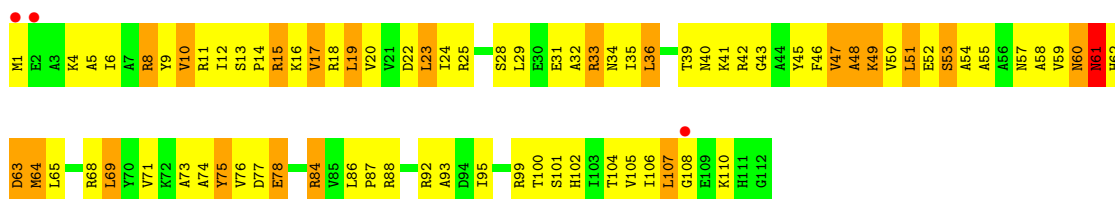
• Molecule 18: 50S ribosomal protein L21

Chain R:



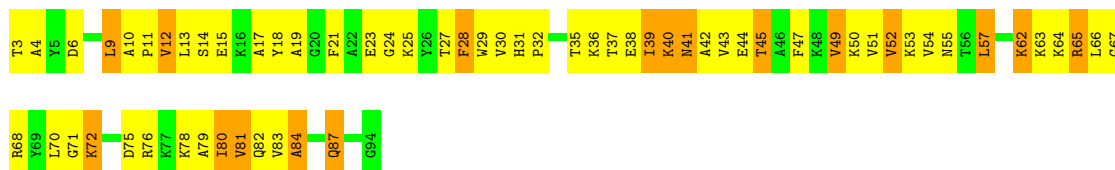
• Molecule 19: 50S ribosomal protein L22

Chain S:



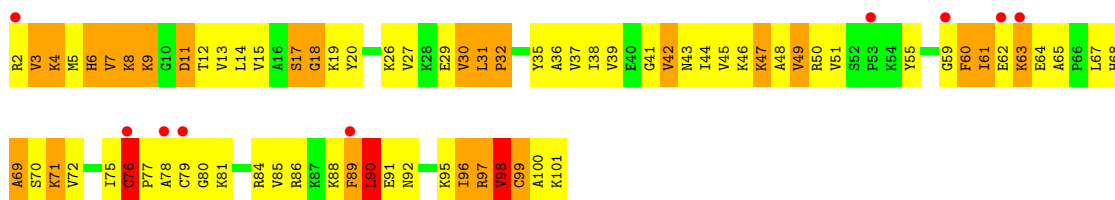
• Molecule 20: 50S ribosomal protein L23

Chain T:



• Molecule 21: 50S ribosomal protein L24

Chain U:

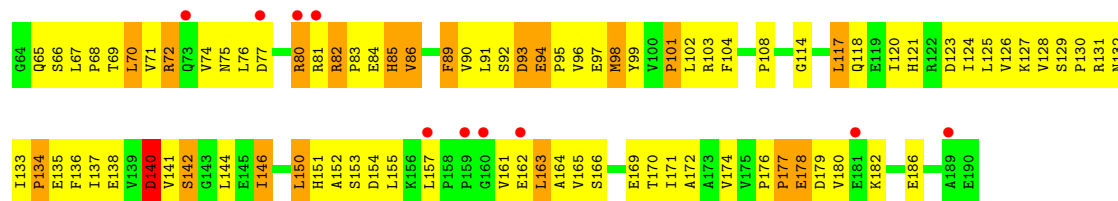


• Molecule 22: 50S ribosomal protein L25

Chain V:

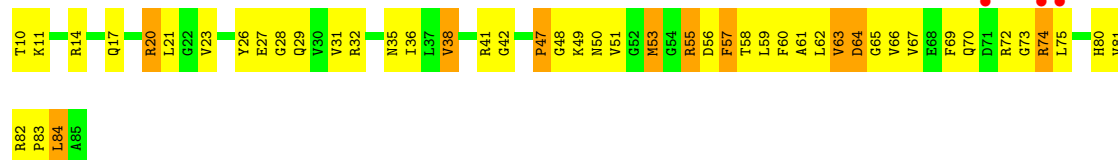






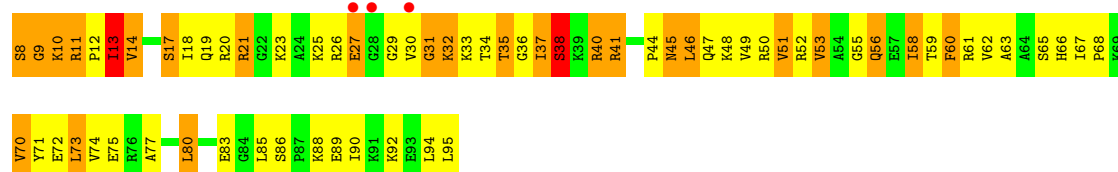
• Molecule 23: 50S ribosomal protein L27

Chain W:



• Molecule 24: 50S ribosomal protein L28

Chain X:



• Molecule 25: 50S ribosomal protein L29

Chain Y:



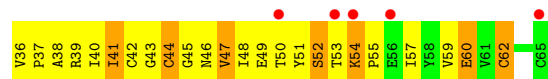
• Molecule 26: 50S ribosomal protein L30

Chain Z:



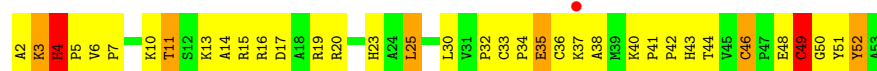
• Molecule 27: 50S ribosomal protein L31

Chain 1:



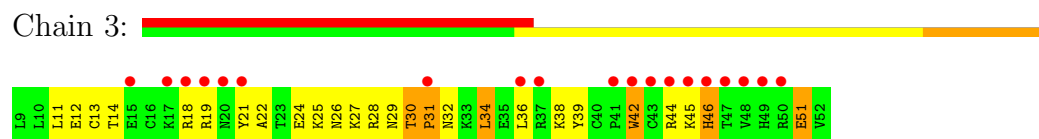
• Molecule 28: 50S ribosomal protein L32

Chain 2:





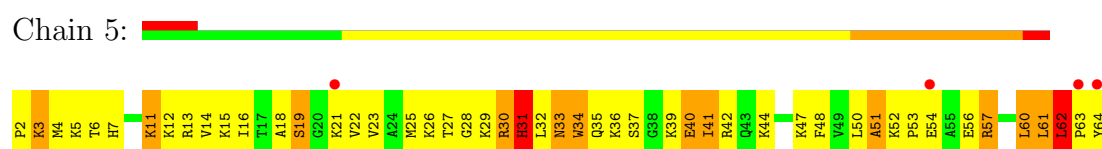
- Molecule 29: 50S ribosomal protein L33



- Molecule 30: 50S ribosomal protein L34



- Molecule 31: 50S ribosomal protein L35





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	210.69Å 451.66Å 614.25Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.52 – 3.40 49.52 – 3.40	Depositor EDS
% Data completeness (in resolution range)	97.5 (49.52-3.40) 97.6 (49.52-3.40)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.28 (at 3.40Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.1_357)	Depositor
R, $R_{free}$	0.228 , 0.266 0.440 , 0.443	Depositor DCC
$R_{free}$ test set	7680 reflections (0.99%)	DCC
Wilson B-factor (Å <sup>2</sup> )	86.0	Xtriage
Anisotropy	0.391	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 54.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtriage
Outliers	0 of 775950 reflections	Xtriage
$F_o, F_c$ correlation	0.61	EDS
Total number of atoms	88276	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	88.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.68% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.



## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

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	1.19	253/66575 (0.4%)	1.59	1756/103930 (1.7%)
2	B	0.59	0/2853	1.04	3/4451 (0.1%)
3	C	0.74	1/2155 (0.0%)	0.91	5/2905 (0.2%)
4	D	0.62	1/1597 (0.1%)	0.81	0/2153
5	E	0.67	0/1622	0.78	0/2194
6	F	0.28	0/1500	0.49	0/2017
7	G	0.44	0/1246	0.64	0/1682
8	H	0.38	0/1148	0.56	0/1552
9	I	0.27	0/252	0.46	0/333
10	J	0.59	0/1124	0.76	0/1515
11	K	0.61	0/942	0.77	0/1268
12	L	0.75	2/1131 (0.2%)	1.03	5/1504 (0.3%)
13	M	0.60	0/1099	0.83	1/1468 (0.1%)
14	N	0.59	0/974	0.83	1/1302 (0.1%)
15	O	0.39	0/779	0.61	0/1036
16	P	0.51	0/1158	0.69	0/1544
17	Q	0.67	0/970	0.81	0/1290
18	R	0.61	0/790	0.74	1/1057 (0.1%)
19	S	0.66	0/902	0.76	0/1209
20	T	0.74	0/740	0.84	0/993
21	U	0.56	0/789	0.76	0/1051
22	V	0.38	0/1524	0.57	0/2068
23	W	0.52	0/613	0.72	0/816
24	X	0.82	0/702	1.04	2/932 (0.2%)
25	Y	0.72	0/523	0.98	3/690 (0.4%)
26	Z	0.50	0/473	0.65	0/634
27	1	0.22	0/229	0.41	0/309
28	2	0.58	0/419	0.79	0/567
29	3	0.27	0/388	0.46	0/518
30	4	0.84	0/427	1.05	1/561 (0.2%)
31	5	0.69	0/516	0.88	1/679 (0.1%)
All	All	1.05	257/96160 (0.3%)	1.42	1779/144228 (1.2%)

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
3	C	0	1
5	E	0	1
12	L	0	5
13	M	0	1
14	N	0	1
17	Q	0	2
All	All	0	11

The worst 5 of 257 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	774	A	N9-C4	-13.87	1.29	1.37
1	A	1332	G	N9-C4	-11.99	1.28	1.38
1	A	1602	U	C4-O4	11.31	1.32	1.23
1	A	2249	U	C4-O4	10.67	1.32	1.23
1	A	1614	A	N9-C4	-10.65	1.31	1.37

The worst 5 of 1779 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	761	A	N1-C6-N6	30.82	137.09	118.60
1	A	1332	G	N3-C4-N9	-24.42	111.35	126.00
1	A	1332	G	N3-C4-C5	23.85	140.52	128.60
1	A	761	A	C6-C5-N7	-21.62	117.17	132.30
1	A	1602	U	N3-C4-C5	-19.94	102.64	114.60

There are no chirality outliers.

5 of 11 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	237	GLU	Peptide
5	E	47	GLY	Peptide
12	L	29	LYS	Peptide
12	L	37	GLY	Peptide
12	L	9	ASN	Peptide



## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	59442	0	29966	2583	0
2	B	2551	0	1295	148	0
3	C	2105	0	2182	345	0
4	D	1564	0	1629	224	0
5	E	1587	0	1632	155	0
6	F	1475	0	1537	150	0
7	G	1223	0	1282	121	0
8	H	1133	0	1220	133	0
9	I	254	0	275	8	0
10	J	1097	0	1168	158	0
11	K	932	0	994	95	0
12	L	1114	0	1187	279	0
13	M	1079	0	1127	172	0
14	N	960	0	1021	142	0
15	O	771	0	832	100	0
16	P	1144	0	1211	122	0
17	Q	953	0	1013	155	0
18	R	779	0	852	128	0
19	S	891	0	951	110	0
20	T	726	0	778	92	0
21	U	776	0	870	139	0
22	V	1492	0	1513	171	0
23	W	605	0	628	63	0
24	X	695	0	764	106	0
25	Y	521	0	575	81	0
26	Z	468	0	523	46	0
27	1	226	0	225	24	0
28	2	405	0	420	64	0
29	3	381	0	391	26	0
30	4	419	0	467	48	0
31	5	508	0	576	110	0
All	All	88276	0	59104	5676	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 39.



The worst 5 of 5676 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
12:L:59:LEU:HA	12:L:61:ARG:NE	1.55	1.20
12:L:57:THR:HG23	12:L:59:LEU:HD22	1.22	1.20
30:4:8:ASN:HD22	30:4:8:ASN:C	1.42	1.18
25:Y:2:LYS:HE2	25:Y:2:LYS:H	1.02	1.15
12:L:114:ILE:H	12:L:114:ILE:HD12	1.11	1.15

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

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	C	269/271 (99%)	210 (78%)	39 (14%)	20 (7%)	2	22
4	D	202/204 (99%)	155 (77%)	32 (16%)	15 (7%)	2	22
5	E	200/202 (99%)	155 (78%)	30 (15%)	15 (8%)	2	22
6	F	179/181 (99%)	136 (76%)	31 (17%)	12 (7%)	2	25
7	G	157/159 (99%)	111 (71%)	36 (23%)	10 (6%)	2	26
8	H	143/145 (99%)	91 (64%)	31 (22%)	21 (15%)	0	5
9	I	28/65 (43%)	25 (89%)	3 (11%)	0	100	100
10	J	135/137 (98%)	97 (72%)	24 (18%)	14 (10%)	1	11
11	K	120/122 (98%)	98 (82%)	14 (12%)	8 (7%)	2	25
12	L	144/146 (99%)	86 (60%)	35 (24%)	23 (16%)	0	4
13	M	134/136 (98%)	86 (64%)	30 (22%)	18 (13%)	0	6
14	N	115/117 (98%)	90 (78%)	15 (13%)	10 (9%)	1	17
15	O	96/98 (98%)	54 (56%)	25 (26%)	17 (18%)	0	3
16	P	135/137 (98%)	100 (74%)	19 (14%)	16 (12%)	1	9
17	Q	114/116 (98%)	82 (72%)	20 (18%)	12 (10%)	1	11
18	R	99/101 (98%)	70 (71%)	20 (20%)	9 (9%)	1	15
19	S	110/112 (98%)	87 (79%)	17 (16%)	6 (6%)	3	31

Continued on next page...



Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
20	T	90/92 (98%)	67 (74%)	18 (20%)	5 (6%)	3	30
21	U	98/100 (98%)	58 (59%)	21 (21%)	19 (19%)	0	2
22	V	186/188 (99%)	135 (73%)	34 (18%)	17 (9%)	1	15
23	W	74/76 (97%)	60 (81%)	10 (14%)	4 (5%)	3	31
24	X	86/88 (98%)	54 (63%)	19 (22%)	13 (15%)	0	4
25	Y	60/62 (97%)	41 (68%)	12 (20%)	7 (12%)	1	9
26	Z	57/59 (97%)	50 (88%)	6 (10%)	1 (2%)	13	65
27	1	28/30 (93%)	15 (54%)	7 (25%)	6 (21%)	0	2
28	2	50/52 (96%)	39 (78%)	7 (14%)	4 (8%)	1	19
29	3	42/44 (96%)	26 (62%)	11 (26%)	5 (12%)	1	9
30	4	46/48 (96%)	42 (91%)	3 (6%)	1 (2%)	10	60
31	5	61/63 (97%)	44 (72%)	10 (16%)	7 (12%)	1	9
All	All	3258/3351 (97%)	2364 (73%)	579 (18%)	315 (10%)	1	13

5 of 315 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	26	LYS
3	C	33	LEU
3	C	34	VAL
3	C	271	ILE
4	D	2	LYS

### 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	
3	C	213/213 (100%)	162 (76%)	51 (24%)	1	5
4	D	165/165 (100%)	129 (78%)	36 (22%)	1	7
5	E	161/161 (100%)	124 (77%)	37 (23%)	1	6
6	F	155/155 (100%)	134 (86%)	21 (14%)	6	29
7	G	132/132 (100%)	107 (81%)	25 (19%)	2	12

Continued on next page...



Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
8	H	122/122 (100%)	103 (84%)	19 (16%)	4	23
9	I	27/53 (51%)	25 (93%)	2 (7%)	20	65
10	J	116/116 (100%)	85 (73%)	31 (27%)	1	4
11	K	100/100 (100%)	78 (78%)	22 (22%)	1	7
12	L	112/112 (100%)	76 (68%)	36 (32%)	0	2
13	M	106/106 (100%)	81 (76%)	25 (24%)	1	6
14	N	100/100 (100%)	76 (76%)	24 (24%)	1	5
15	O	77/77 (100%)	63 (82%)	14 (18%)	2	13
16	P	121/121 (100%)	94 (78%)	27 (22%)	1	7
17	Q	92/92 (100%)	71 (77%)	21 (23%)	1	6
18	R	82/82 (100%)	61 (74%)	21 (26%)	1	4
19	S	91/91 (100%)	65 (71%)	26 (29%)	0	3
20	T	74/74 (100%)	60 (81%)	14 (19%)	2	12
21	U	84/84 (100%)	67 (80%)	17 (20%)	2	9
22	V	163/163 (100%)	141 (86%)	22 (14%)	6	29
23	W	61/61 (100%)	53 (87%)	8 (13%)	6	31
24	X	73/73 (100%)	50 (68%)	23 (32%)	0	3
25	Y	58/58 (100%)	46 (79%)	12 (21%)	2	8
26	Z	51/51 (100%)	43 (84%)	8 (16%)	4	23
27	1	27/27 (100%)	26 (96%)	1 (4%)	45	85
28	2	45/45 (100%)	39 (87%)	6 (13%)	6	30
29	3	43/43 (100%)	38 (88%)	5 (12%)	8	37
30	4	41/41 (100%)	28 (68%)	13 (32%)	0	3
31	5	53/53 (100%)	43 (81%)	10 (19%)	2	12
All	All	2745/2771 (99%)	2168 (79%)	577 (21%)	1	8

5 of 577 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
12	L	57	THR
14	N	79	LEU
25	Y	62	THR
12	L	85	LEU
13	M	79	LEU



Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 84 such sidechains are listed below:

Mol	Chain	Res	Type
10	J	154	GLN
16	P	90	GLN
28	2	23	HIS
12	L	70	GLN
14	N	61	HIS

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	2757/2879 (95%)	589 (21%)	29 (1%)
2	B	118/119 (99%)	27 (22%)	0
All	All	2875/2998 (95%)	616 (21%)	29 (1%)

5 of 616 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	10	G
1	A	11	G
1	A	34	C
1	A	35	G
1	A	46	C

5 of 29 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	1379	A
1	A	1609	A
1	A	2435	A
1	A	1558	A
1	A	1617	C

## 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 ⓘ

There are no ligands in this entry.

## 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.



## 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, 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	2760/2879 (95%)	0.06	77 (2%) 50 21	25, 63, 178, 410	0
2	B	119/119 (100%)	0.77	16 (13%) 4 2	78, 129, 184, 236	0
3	C	271/271 (100%)	-0.05	2 (0%) 84 52	18, 57, 109, 177	0
4	D	204/204 (100%)	-0.01	0 100 100	33, 71, 145, 347	0
5	E	202/202 (100%)	0.06	4 (1%) 62 28	25, 73, 155, 192	0
6	F	181/181 (100%)	0.60	20 (11%) 6 4	104, 185, 268, 331	0
7	G	159/159 (100%)	0.21	6 (3%) 38 15	79, 136, 186, 235	0
8	H	145/145 (100%)	0.96	20 (13%) 4 2	64, 236, 379, 480	0
9	I	32/65 (49%)	0.70	3 (9%) 9 4	168, 253, 310, 334	0
10	J	137/137 (100%)	-0.04	3 (2%) 59 25	52, 81, 146, 194	0
11	K	122/122 (100%)	-0.02	1 (0%) 83 49	41, 69, 111, 162	0
12	L	146/146 (100%)	0.23	7 (4%) 29 12	32, 97, 163, 293	0
13	M	136/136 (100%)	0.36	6 (4%) 33 13	48, 88, 205, 406	0
14	N	117/117 (100%)	0.20	3 (2%) 53 21	43, 73, 134, 235	0
15	O	98/98 (100%)	0.51	9 (9%) 9 5	80, 136, 190, 215	0
16	P	137/137 (100%)	0.30	8 (5%) 22 8	55, 92, 190, 273	0
17	Q	116/116 (100%)	-0.03	0 100 100	26, 74, 126, 248	0
18	R	101/101 (100%)	0.16	0 100 100	41, 110, 156, 259	0
19	S	112/112 (100%)	-0.11	3 (2%) 52 21	43, 59, 134, 255	0
20	T	92/92 (100%)	0.04	0 100 100	36, 73, 127, 169	0
21	U	100/100 (100%)	0.44	9 (9%) 10 5	61, 102, 251, 408	0
22	V	188/188 (100%)	0.31	10 (5%) 25 9	83, 139, 194, 230	0
23	W	76/76 (100%)	0.19	3 (3%) 37 14	59, 84, 135, 256	0
24	X	88/88 (100%)	0.33	3 (3%) 43 17	39, 70, 153, 326	0

*Continued on next page...*



Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
25	Y	62/62 (100%)	-0.01	2 (3%) 45 18	51, 96, 212, 328	0
26	Z	59/59 (100%)	-0.02	1 (1%) 67 31	45, 85, 157, 305	0
27	1	30/30 (100%)	0.96	5 (16%) 2 2	183, 261, 306, 358	0
28	2	52/52 (100%)	-0.03	1 (1%) 64 29	21, 72, 197, 229	0
29	3	44/44 (100%)	2.06	19 (43%) 1 1	141, 245, 312, 333	0
30	4	48/48 (100%)	0.10	1 (2%) 60 27	21, 41, 91, 200	0
31	5	63/63 (100%)	0.60	4 (6%) 19 8	45, 70, 132, 216	0
All	All	6197/6349 (97%)	0.17	246 (3%) 36 14	18, 77, 220, 480	0

The worst 5 of 246 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
29	3	47	THR	12.1
1	A	2801	A	11.2
12	L	150	ALA	10.4
1	A	10	G	9.8
1	A	2798	C	9.8

## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

There are no ligands in this entry.

## 6.5 Other polymers ⓘ

There are no such residues in this entry.