



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 27, 2020 – 02:19 am BST

PDB ID : 3E1K  
Title : Crystal structure of Kluyveromyces lactis Gal80p in complex with the acidic activation domain of Gal4p  
Authors : Thoden, J.B.; Holden, H.M.  
Deposited on : 2008-08-04  
Resolution : 3.00 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

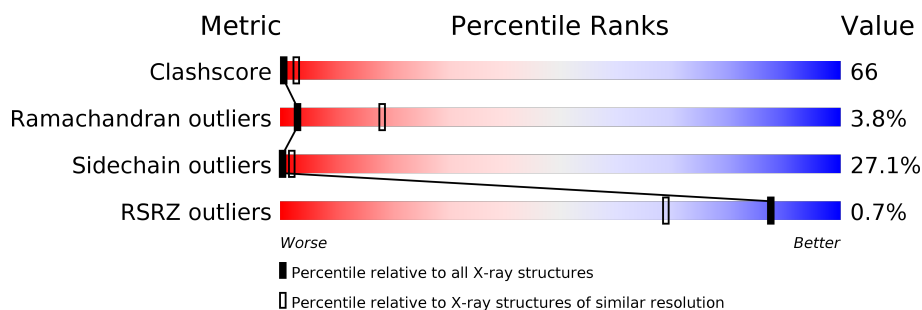
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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))
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 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	465	
1	C	465	
1	E	465	
1	G	465	
1	I	465	
1	K	465	
1	M	465	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	O	465	<div><div><div></div><div></div><div></div><div></div></div><div>% 15%47%19%15%</div></div>
2	B	22	<div><div><div></div><div></div><div></div><div></div></div><div>5% 18%36%9%36%</div></div>
2	D	22	<div><div><div></div><div></div><div></div><div></div></div><div>14%45%5%36%</div></div>
2	F	22	<div><div><div></div><div></div><div></div><div></div></div><div>5% 32%27%5%36%</div></div>
2	H	22	<div><div><div></div><div></div><div></div><div></div></div><div>27%23%14%36%</div></div>
2	J	22	<div><div><div></div><div></div><div></div><div></div></div><div>5% 18%27%18%36%</div></div>
2	L	22	<div><div><div></div><div></div><div></div><div></div></div><div>18%41%5%36%</div></div>
2	N	22	<div><div><div></div><div></div><div></div><div></div></div><div>5% 14%41%9%36%</div></div>
2	P	22	<div><div><div></div><div></div><div></div><div></div></div><div>18%36%9%36%</div></div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 26057 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 Galactose/lactose metabolism regulatory protein GAL80.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	395	Total	C	N	O	S	0	0	0
			3153	2024	527	593	9			
1	C	393	Total	C	N	O	S	0	0	0
			3136	2013	525	589	9			
1	E	393	Total	C	N	O	S	0	0	0
			3136	2013	525	589	9			
1	G	393	Total	C	N	O	S	0	0	0
			3136	2013	525	589	9			
1	I	393	Total	C	N	O	S	0	0	0
			3136	2013	525	589	9			
1	K	393	Total	C	N	O	S	0	0	0
			3136	2013	525	589	9			
1	M	393	Total	C	N	O	S	0	0	0
			3136	2013	525	589	9			
1	O	393	Total	C	N	O	S	0	0	0
			3136	2013	525	589	9			

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	458	LEU	-	EXPRESSION TAG	UNP Q06433
A	459	GLU	-	EXPRESSION TAG	UNP Q06433
A	460	HIS	-	EXPRESSION TAG	UNP Q06433
A	461	HIS	-	EXPRESSION TAG	UNP Q06433
A	462	HIS	-	EXPRESSION TAG	UNP Q06433
A	463	HIS	-	EXPRESSION TAG	UNP Q06433
A	464	HIS	-	EXPRESSION TAG	UNP Q06433
A	465	HIS	-	EXPRESSION TAG	UNP Q06433
C	458	LEU	-	EXPRESSION TAG	UNP Q06433
C	459	GLU	-	EXPRESSION TAG	UNP Q06433
C	460	HIS	-	EXPRESSION TAG	UNP Q06433
C	461	HIS	-	EXPRESSION TAG	UNP Q06433
C	462	HIS	-	EXPRESSION TAG	UNP Q06433

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	463	HIS	-	EXPRESSION TAG	UNP Q06433
C	464	HIS	-	EXPRESSION TAG	UNP Q06433
C	465	HIS	-	EXPRESSION TAG	UNP Q06433
E	458	LEU	-	EXPRESSION TAG	UNP Q06433
E	459	GLU	-	EXPRESSION TAG	UNP Q06433
E	460	HIS	-	EXPRESSION TAG	UNP Q06433
E	461	HIS	-	EXPRESSION TAG	UNP Q06433
E	462	HIS	-	EXPRESSION TAG	UNP Q06433
E	463	HIS	-	EXPRESSION TAG	UNP Q06433
E	464	HIS	-	EXPRESSION TAG	UNP Q06433
E	465	HIS	-	EXPRESSION TAG	UNP Q06433
G	458	LEU	-	EXPRESSION TAG	UNP Q06433
G	459	GLU	-	EXPRESSION TAG	UNP Q06433
G	460	HIS	-	EXPRESSION TAG	UNP Q06433
G	461	HIS	-	EXPRESSION TAG	UNP Q06433
G	462	HIS	-	EXPRESSION TAG	UNP Q06433
G	463	HIS	-	EXPRESSION TAG	UNP Q06433
G	464	HIS	-	EXPRESSION TAG	UNP Q06433
G	465	HIS	-	EXPRESSION TAG	UNP Q06433
I	458	LEU	-	EXPRESSION TAG	UNP Q06433
I	459	GLU	-	EXPRESSION TAG	UNP Q06433
I	460	HIS	-	EXPRESSION TAG	UNP Q06433
I	461	HIS	-	EXPRESSION TAG	UNP Q06433
I	462	HIS	-	EXPRESSION TAG	UNP Q06433
I	463	HIS	-	EXPRESSION TAG	UNP Q06433
I	464	HIS	-	EXPRESSION TAG	UNP Q06433
I	465	HIS	-	EXPRESSION TAG	UNP Q06433
K	458	LEU	-	EXPRESSION TAG	UNP Q06433
K	459	GLU	-	EXPRESSION TAG	UNP Q06433
K	460	HIS	-	EXPRESSION TAG	UNP Q06433
K	461	HIS	-	EXPRESSION TAG	UNP Q06433
K	462	HIS	-	EXPRESSION TAG	UNP Q06433
K	463	HIS	-	EXPRESSION TAG	UNP Q06433
K	464	HIS	-	EXPRESSION TAG	UNP Q06433
K	465	HIS	-	EXPRESSION TAG	UNP Q06433
M	458	LEU	-	EXPRESSION TAG	UNP Q06433
M	459	GLU	-	EXPRESSION TAG	UNP Q06433
M	460	HIS	-	EXPRESSION TAG	UNP Q06433
M	461	HIS	-	EXPRESSION TAG	UNP Q06433
M	462	HIS	-	EXPRESSION TAG	UNP Q06433
M	463	HIS	-	EXPRESSION TAG	UNP Q06433
M	464	HIS	-	EXPRESSION TAG	UNP Q06433

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
M	465	HIS	-	EXPRESSION TAG	UNP Q06433
O	458	LEU	-	EXPRESSION TAG	UNP Q06433
O	459	GLU	-	EXPRESSION TAG	UNP Q06433
O	460	HIS	-	EXPRESSION TAG	UNP Q06433
O	461	HIS	-	EXPRESSION TAG	UNP Q06433
O	462	HIS	-	EXPRESSION TAG	UNP Q06433
O	463	HIS	-	EXPRESSION TAG	UNP Q06433
O	464	HIS	-	EXPRESSION TAG	UNP Q06433
O	465	HIS	-	EXPRESSION TAG	UNP Q06433

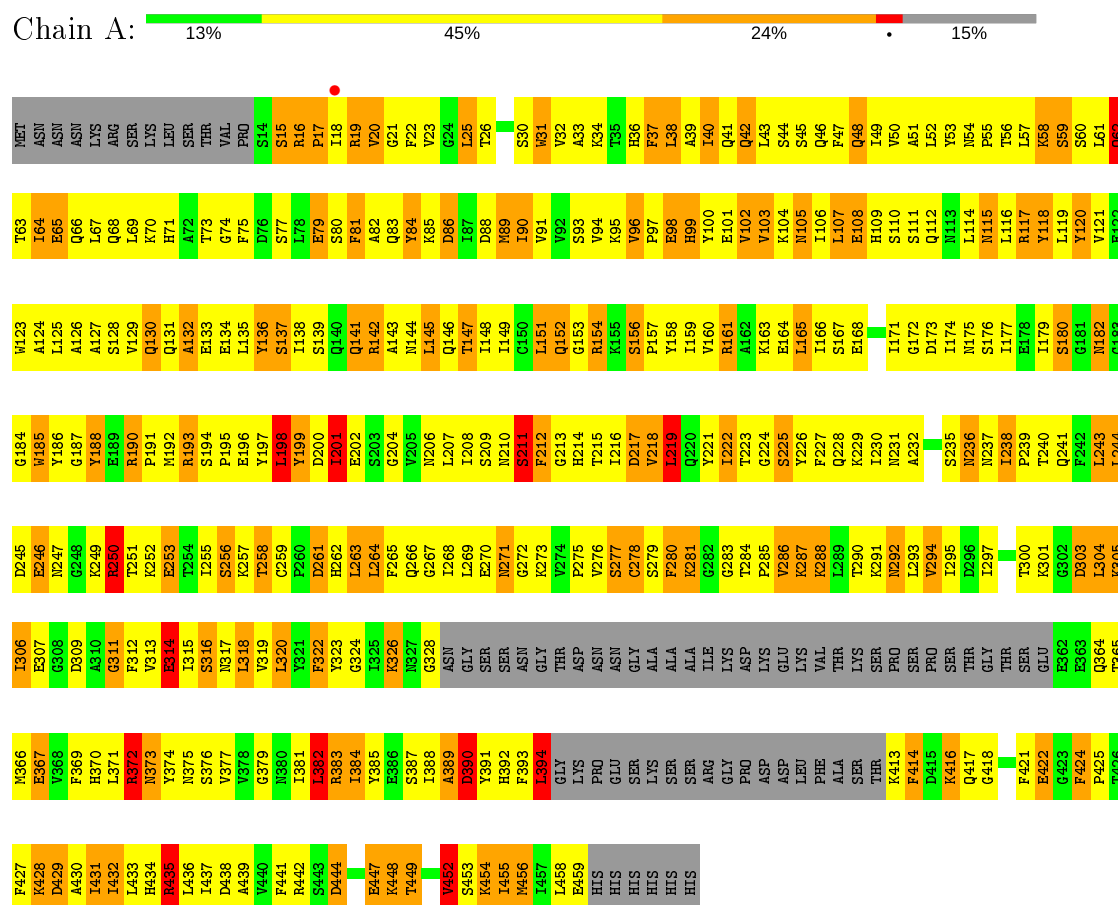
- Molecule 2 is a protein called Lactose regulatory protein LAC9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	14	Total	C	N	O	S	0	0	0
			119	77	16	25	1			
2	D	14	Total	C	N	O	S	0	0	0
			119	77	16	25	1			
2	F	14	Total	C	N	O	S	0	0	0
			119	77	16	25	1			
2	H	14	Total	C	N	O	S	0	0	0
			119	77	16	25	1			
2	J	14	Total	C	N	O	S	0	0	0
			119	77	16	25	1			
2	L	14	Total	C	N	O	S	0	0	0
			119	77	16	25	1			
2	N	14	Total	C	N	O	S	0	0	0
			119	77	16	25	1			
2	P	14	Total	C	N	O	S	0	0	0
			119	77	16	25	1			

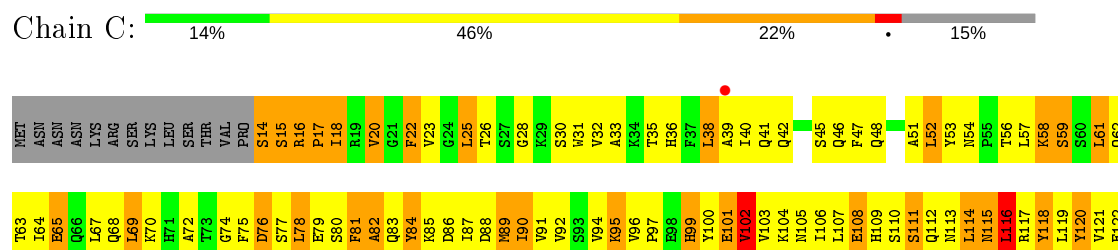
### 3 Residue-property plots

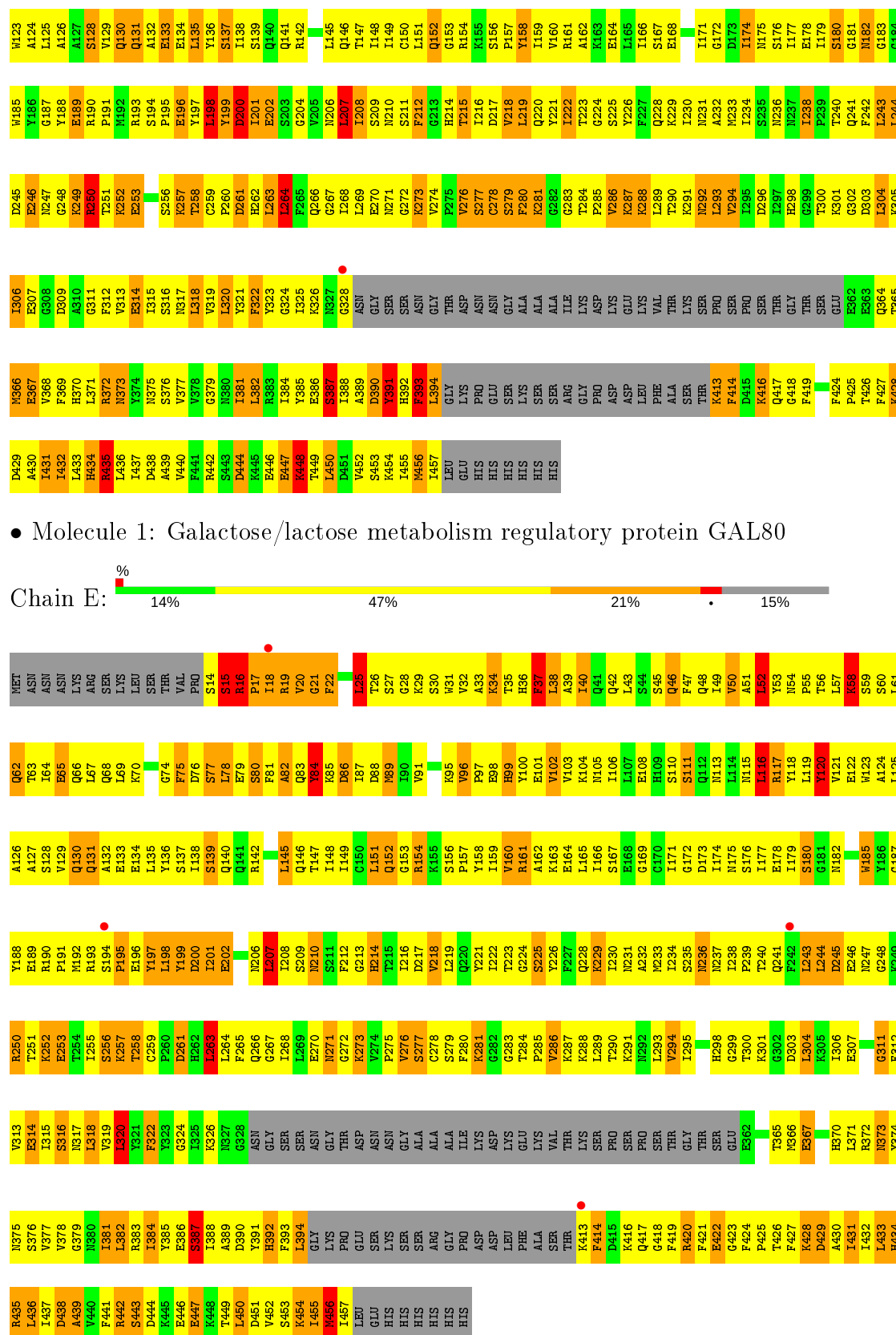
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Galactose/lactose metabolism regulatory protein GAL80



- Molecule 1: Galactose/lactose metabolism regulatory protein GAL80

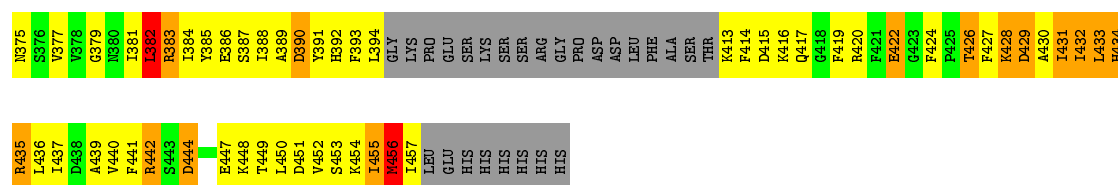






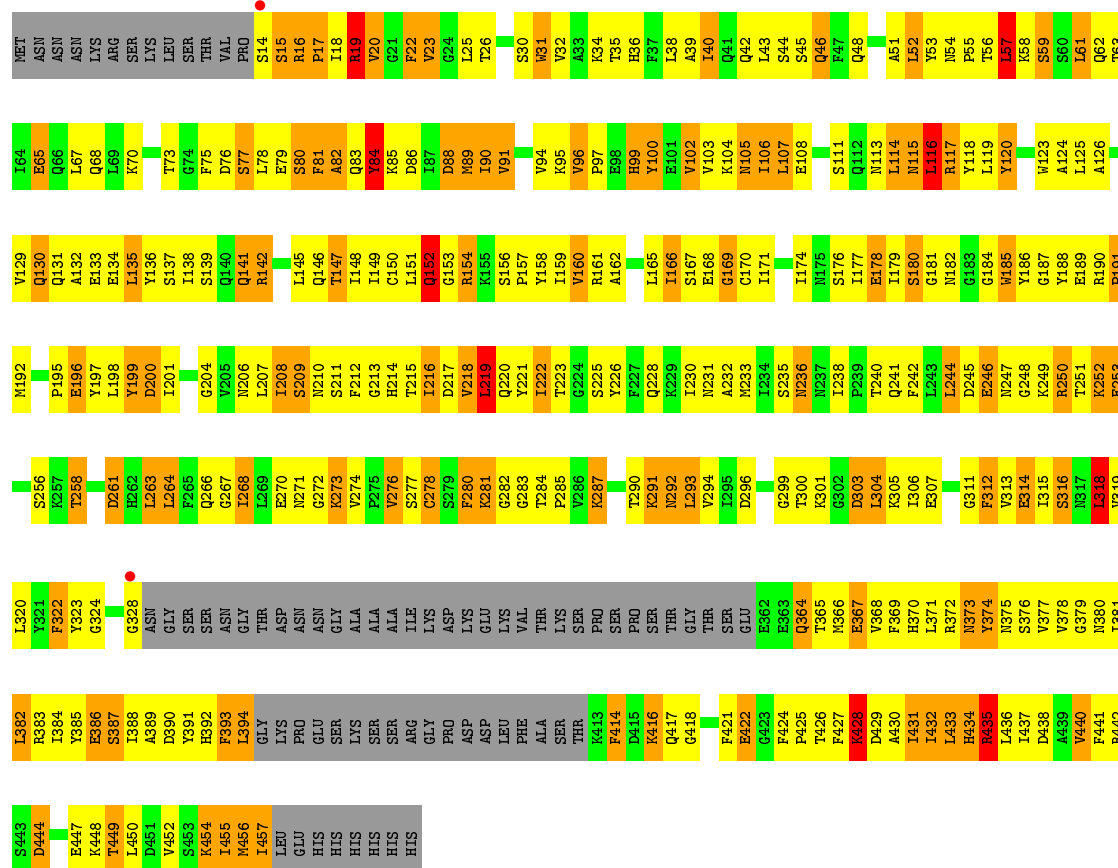
Chain G: 

MET	ASN	ASN	ASN	ASN	LYS	ARG	SER	LYS	LEU	SER	THR	VAL	PRO	S14	S15	R16	P17	I18	R19	V20	G21	F22	V23	G24	L25	T26	S27	G28	K29	S30	N31	V32	A33	T35	R36	F37	L38	A39	I40	Q41	Q42	Q43	L44	S45	Q46	F47	Q48	I49	V50	A51	L52	Y53	N54	P55	T56	L57	K58	S59	S60																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
L61	Q62	T63	I64	E65	Q66	L67	Q68	L69	K70	H71	A72	F75	D76	S77	Q78	K85	Y84	D86	L87	D88	S89	V91	V92	S93	V94	K95	V96	P97	E98	H99	Y100	E101	V102	V103	K104	N105	I106	E108	H109	S110	S111	Q112	N113	N115	L116	R117	Y119	S180	G181	Y120	Y121	G183																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
E122	W123	A124	L125	A126	A127	S128	V129	Q130	Q131	A132	E133	E134	L135	Y136	S137	I138	L139	S140	S141	S142	S143	S144	S145	S146	S147	S148	S149	S150	S151	S152	S153	S154	S155	S156	S157	S158	S159	S160	S161	S162	S163	S164	S165	S166	S167	S168	S169	S170	S171	S172	S173	S174	S175	S176	S177	S178	S179	S180	S181	S182	S183	S184	S185	S186	S187	S188	S189	S190	S191	S192	S193	S194	S195	S196	S197	S198	S199	S200	S201	S202	S203	S204	S205	S206	S207	S208	S209	S210	S211	S212	S213	S214	S215	S216	S217	S218	S219	S220	S221	S222	S223	S224	S225	S226	S227	S228	S229	S230	S231	S232	S233	S234	S235	S236	S237	S238	S239	S240	S241	S242	S243	S244	S245	S246	S247	S248	S249	S250	S251	S252	S253	S254	S255	S256	S257	S258	S259	S260	S261	S262	S263	S264	S265	S266	S267	S268	S269	S270	S271	S272	S273	S274	S275	S276	S277	S278	S279	S280	S281	S282	S283	S284	S285	S286	S287	S288	S289	S290	S291	S292	S293	S294	S295	S296	S297	S298	S299	S300	S301	S302	S303	S304	S305	S306	S307	S308	S309	S310	S311	S312	S313	S314	S315	S316	S317	S318	S319	S320	S321	S322	S323	S324	S325	S326	S327	S328	S329	S330	S331	S332	S333	S334	S335	S336	S337	S338	S339	S340	S341	S342	S343	S344	S345	S346	S347	S348	S349	S350	S351	S352	S353	S354	S355	S356	S357	S358	S359	S360	S361	S362	S363	S364	S365	S366	S367	S368	S369	S370	S371	S372	S373	S374	S375	S376	S377	S378	S379	S380	S381	S382	S383	S384	S385	S386	S387	S388	S389	S390	S391	S392	S393	S394	S395	S396	S397	S398	S399	S400	S401	S402	S403	S404	S405	S406	S407	S408	S409	S410	S411	S412	S413	S414	S415	S416	S417	S418	S419	S420	S421	S422	S423	S424	S425	S426	S427	S428	S429	S430	S431	S432	S433	S434	S435	S436	S437	S438	S439	S440	S441	S442	S443	S444	S445	S446	S447	S448	S449	S450	S451	S452	S453	S454	S455	S456	S457	S458	S459	S460	S461	S462	S463	S464	S465	S466	S467	S468	S469	S470	S471	S472	S473	S474	S475	S476	S477	S478	S479	S480	S481	S482	S483	S484	S485	S486	S487	S488	S489	S490	S491	S492	S493	S494	S495	S496	S497	S498	S499	S500	S501	S502	S503	S504	S505	S506	S507	S508	S509	S510	S511	S512	S513	S514	S515	S516	S517	S518	S519	S520	S521	S522	S523	S524	S525	S526	S527	S528	S529	S530	S531	S532	S533	S534	S535	S536	S537	S538	S539	S540	S541	S542	S543	S544	S545	S546	S547	S548	S549	S550	S551	S552	S553	S554	S555	S556	S557	S558	S559	S560	S561	S562	S563	S564	S565	S566	S567	S568	S569	S570	S571	S572	S573	S574	S575	S576	S577	S578	S579	S580	S581	S582	S583	S584	S585	S586	S587	S588	S589	S590	S591	S592	S593	S594	S595	S596	S597	S598	S599	S600	S601	S602	S603	S604	S605	S606	S607	S608	S609	S610	S611	S612	S613	S614	S615	S616	S617	S618	S619	S620	S621	S622	S623	S624	S625	S626	S627	S628	S629	S630	S631	S632	S633	S634	S635	S636	S637	S638	S639	S640	S641	S642	S643	S644	S645	S646	S647	S648	S649	S650	S651	S652	S653	S654	S655	S656	S657	S658	S659	S660	S661	S662	S663	S664	S665	S666	S667	S668	S669	S670	S671	S672	S673	S674	S675	S676	S677	S678	S679	S680	S681	S682	S683	S684	S685	S686	S687	S688	S689	S690	S691	S692	S693	S694	S695	S696	S697	S698	S699	S700	S701	S702	S703	S704	S705	S706	S707	S708	S709	S710	S711	S712	S713	S714	S715	S716	S717	S718	S719	S720	S721	S722	S723	S724	S725	S726	S727	S728	S729	S730	S731	S732	S733	S734	S735	S736	S737	S738	S739	S740	S741	S742	S743	S744	S745	S746	S747	S748	S749	S750	S751	S752	S753	S754	S755	S756	S757	S758	S759	S760	S761	S762	S763	S764	S765	S766	S767	S768	S769	S770	S771	S772	S773	S774	S775	S776	S777	S778	S779	S780	S781	S782	S783	S784	S785	S786	S787	S788	S789	S790	S791	S792	S793	S794	S795	S796	S797	S798	S799	S800	S801	S802	S803	S804	S805	S806	S807	S808	S809	S810	S811	S812	S813	S814	S815	S816	S817	S818	S819	S820	S821	S822	S823	S824	S825	S826	S827	S828	S829	S830	S831	S832	S833	S834	S835	S836	S837	S838	S839	S840	S841	S842	S843	S844	S845	S846	S847	S848	S849	S850	S851	S852	S853	S854	S855	S856	S857	S858	S859	S860	S861	S862	S863	S864	S865	S866	S867	S868	S869	S870	S871	S872	S873	S874	S875	S876	S877	S878	S879	S880	S881	S882	S883	S884	S885	S886	S887	S888	S889	S890	S891	S892	S893	S894	S895	S896	S897	S898	S899	S900	S901	S902	S903	S904	S905	S906	S907	S908	S909	S910	S911	S912	S913	S914	S915	S916	S917	S918	S919	S920	S921	S922	S923	S924	S925	S926	S927	S928	S929	S930	S931	S932	S933	S934	S935	S936	S937	S938	S939	S940	S941	S942	S943	S944	S945	S946	S947	S948	S949	S950	S951	S952	S953	S954	S955	S956	S957	S958	S959	S960	S961	S962	S963	S964	S965	S966	S967	S968	S969	S970	S971	S972	S973	S974	S975	S976	S977	S978	S979	S980	S981	S982	S983	S984	S985	S986	S987	S988	S989	S990	S991	S992	S993	S994	S995	S996	S997	S998	S999	S1000	S1001	S1002	S1003	S1004	S1005	S1006	S1007	S1008	S1009	S1010	S1011	S1012	S1013	S1014	S1015	S1016	S1017	S1018	S1019	S1020	S1021	S1022	S1023	S1024	S1025	S1026	S1027	S1028	S1029	S1030	S1031	S1032	S1033	S1034	S1035	S1036	S1037	S1038	S1039	S1040	S1041	S1042	S1043	S1044	S1045	S1046	S1047	S1048	S1049	S1050	S1051	S1052	S1053	S1054	S1055	S1056	S1057	S1058	S1059	S1060	S1061	S1062	S1063	S1064	S1065	S1066	S1067	S1068	S1069	S1070	S1071	S1072	S1073	S1074	S1075	S1076	S1077	S1078	S1079	S1080	S1081	S1082	S1083	S1084	S1085	S1086	S1087	S1088	S1089	S1090	S1091	S1092	S1093	S1094	S1095	S1096	S1097	S1098	S1099	S1100	S1101	S1102	S1103	S1104	S1105	S1106	S1107	S1108	S1109	S1110	S1111	S1112	S1113	S1114	S1115	S1116	S1117	S1118	S1119	S1120	S1121	S1122	S1123	S1124	S1125	S1126	S1127	S1128	S1129	S1130	S1131	S1132	S1133	S1134	S1135	S1136	S1137	S1138	S1139	S1140	S1141	S1142	S1143	S1144	S1145	S1146	S1147	S1148	S1149	S1150	S1151	S1152	S1153	S1154	S1155	S1156	S1157	S1158	S1159	S1160	S1161	S1162	S1163	S1164	S1165	S1166	S1167	S1168	S1169	S1170	S1171	S1172	S1173	S1174	S1175	S1176	S1177	S1178	S1179	S1180	S1181	S1182	S1183	S1184	S1185	S1186	S1187	S1188	S1189	S1190	S1191	S1192	S1193	S1194	S1195	S1196	S1197	S1198	S1199	S1200	S1201	S1202	S1203	S1204	S1205	S1206	S1207	S1208	S1209	S1210	S1211	S1212	S1213	S1214	S1215	S1216	S1217	S1218	S1219	S1220	S1221	S1222	S1223	S1224	S1225	S1226	S1227	S1228	S1229	S1230	S1231	S1232	S1233	S1234	S1235	S1236	S1237	S1238	S1239	S1240	S1241	S1242	S1243	S1244	S1245	S1246	S1247	S1248	S1249	S1250	S1251	S1252	S1253	S1254	S1255	S1256	S1257	S1258	S1259	S1260	S1261	S1262	S1263	S1264	S1265	S1266	S1267	S1268	S1269	S1270	S1271	S1272	S1273	S1274	S1275	S1276	S1277	S1278	S1279	S1280	S1281	S1282	S1283	S1284	S1285	S1286	S1287	S1288	S1289	S1290	S1291	S1292	S1293	S1294	S1295	S1296	S1297	S1298	S1299	S1300	S1301	S1302	S1303	S1304	S1305	S1306	S1307	S1308	S1309	S1310	S1311	S1312	S1313	S1314	S1315	S1316	S1317	S1318	S1319	S1320	S1321	S1322	S1323	S1324	S1325	S1326	S1327	S1328	S1329	S1330	S1331	S1332	S1333	S1334	S1335	S1336	S1337	S1338	S1339	S1340	S1341	S1342	S1343	S1344	S1345	S1346	S1347	S1348	S1349	S1350	S1351	S1352	S1353	S1354	S1355	S1356	S1357	S1358	S1359	S1360	S1361	S1362	S1363	S1364	S1365	S1366	S1367	S1368	S1369	S1370	S1371	S1372	S1373	S1374	S1375	S1376	S1377	S1378	S1379	S1380	S1381	S1382	S1383	S1384	S1385	S1386	S1387	S1388	S1389	S1390	S1391	S1392	S1393	S1394	S1395	S1396	S1397	S1398	S1399	S1400	S1401	S1402	S1403	S1404	S1405	S1406	S1407	S1408	S1409	S1410	S1411	S1412	S1413	S1414	S1415	S1416	S1417	S1418	S1419	S1420	S1421	S1422	S1423	S1424	S1425	S1426	S1427	S1428	S1429	S1430	S1431	S1432	S1433	S1434	S1435	S1436	S1437	S1438	S1439	S1440	S1441	S1442	S1443	S1444	S1445	S1446	S1447	S1448	S1449	S1450	S1451	S1452	S1453	S1454	S1455	S1456	S1457	S1458	S1459	S1460	S1461	S1462	S1463	S1464	S1465	S1466	S1467	S1468	S1469	S1470	S1471	S1472	S1473	S1474	S1475	S1476	S147



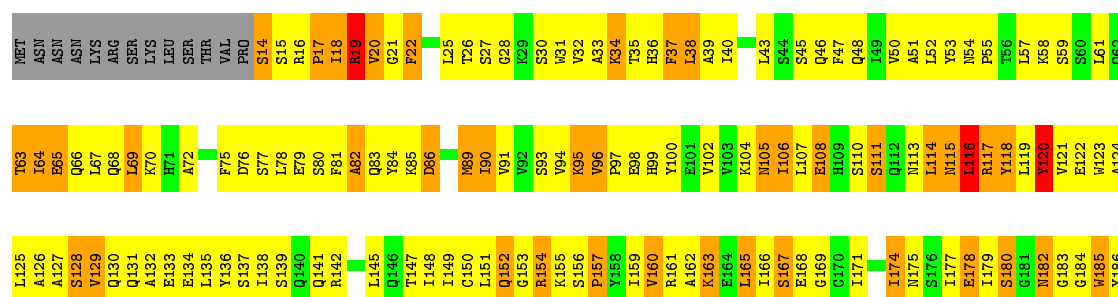
• Molecule 1: Galactose/lactose metabolism regulatory protein GAL80

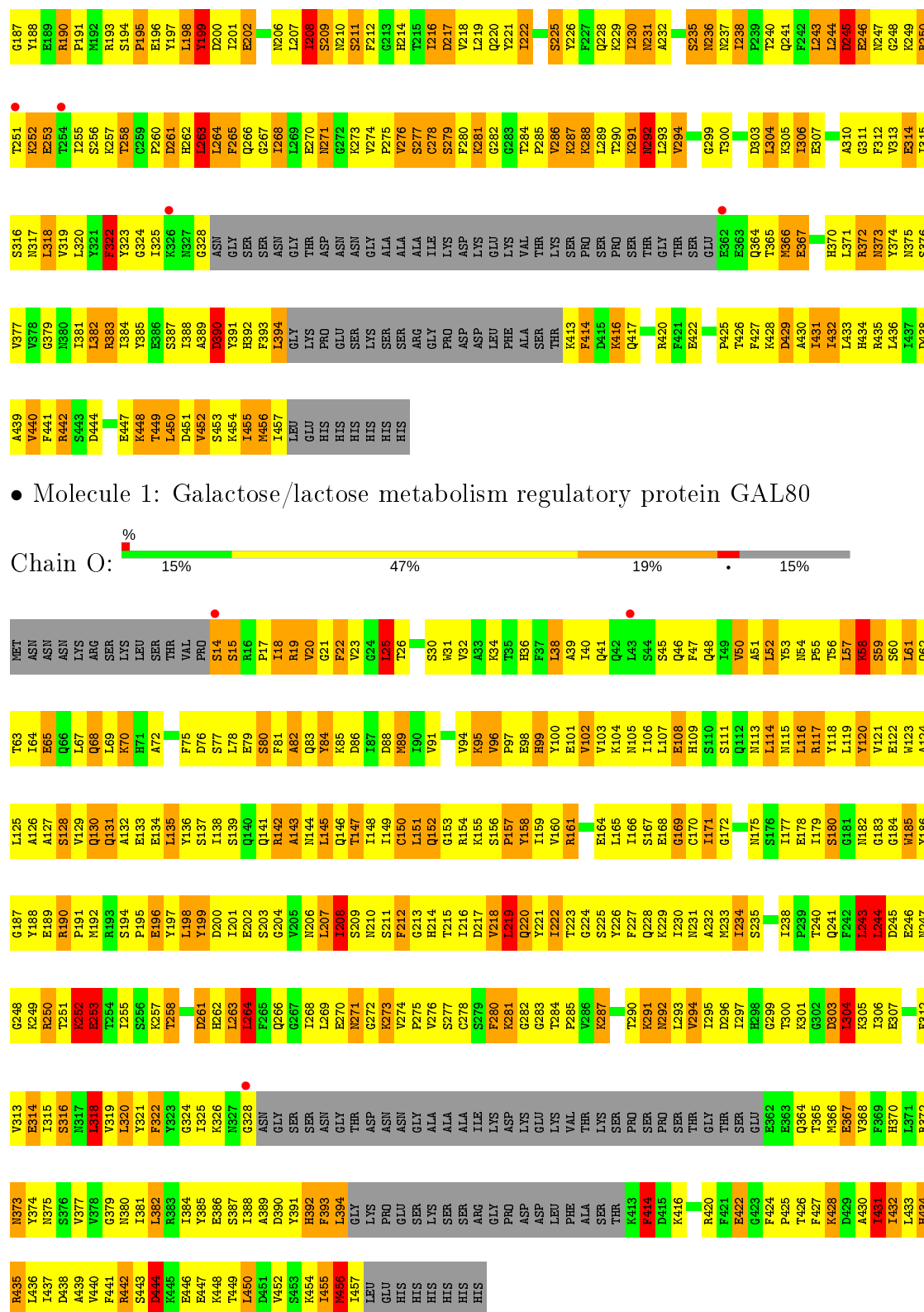
Chain K: 19% 41% 22% 15%



• Molecule 1: Galactose/lactose metabolism regulatory protein GAL80

Chain M: 17% 44% 22% 15%



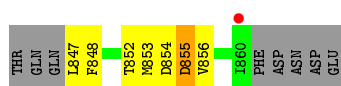




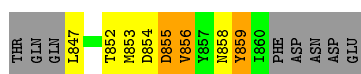
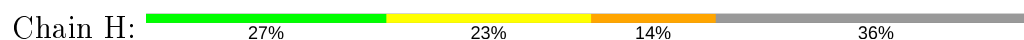
- Molecule 2: Lactose regulatory protein LAC9



- Molecule 2: Lactose regulatory protein LAC9



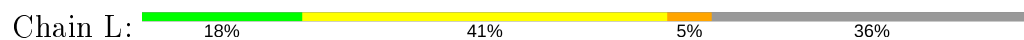
- Molecule 2: Lactose regulatory protein LAC9



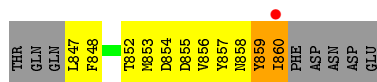
- Molecule 2: Lactose regulatory protein LAC9



- Molecule 2: Lactose regulatory protein LAC9



- Molecule 2: Lactose regulatory protein LAC9



- Molecule 2: Lactose regulatory protein LAC9



THR	GLN	GLN	L847	T852	M853	D854	D855	V856	Y857	N858	Y859	I860	PHE	ASP	ASN	ASP	GLU
-----	-----	-----	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	101.10Å 160.50Å 132.60Å 90.00° 94.70° 90.00°	Depositor
Resolution (Å)	30.00 – 3.00 48.41 – 2.99	Depositor EDS
% Data completeness (in resolution range)	88.8 (30.00-3.00) 88.3 (48.41-2.99)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.68 (at 3.01Å)	Xtriage
Refinement program	TNT	Depositor
R, $R_{free}$	0.228 , 0.289 0.221 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	36.5	Xtriage
Anisotropy	0.317	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 123.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.87	EDS
Total number of atoms	26057	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	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 19.86 % 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 9.8688e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 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	1.11	7/3215 (0.2%)	1.66	65/4340 (1.5%)
1	C	0.98	7/3198 (0.2%)	1.59	46/4317 (1.1%)
1	E	1.12	5/3198 (0.2%)	1.62	45/4317 (1.0%)
1	G	1.09	11/3198 (0.3%)	1.56	37/4317 (0.9%)
1	I	1.14	8/3198 (0.3%)	1.61	57/4317 (1.3%)
1	K	1.06	4/3198 (0.1%)	1.58	50/4317 (1.2%)
1	M	1.05	4/3198 (0.1%)	1.62	53/4317 (1.2%)
1	O	0.96	1/3198 (0.0%)	1.60	48/4317 (1.1%)
2	B	0.95	0/121	1.63	0/165
2	D	1.07	0/121	1.23	0/165
2	F	0.97	0/121	1.30	1/165 (0.6%)
2	H	0.88	0/121	1.41	0/165
2	J	1.09	0/121	1.32	0/165
2	L	1.01	0/121	1.29	0/165
2	N	0.79	0/121	1.33	0/165
2	P	0.72	0/121	1.39	0/165
All	All	1.06	47/26569 (0.2%)	1.60	402/35879 (1.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	O	0	1

The worst 5 of 47 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	158	TYR	CD2-CE2	-6.72	1.29	1.39
1	G	158	TYR	CD2-CE2	-6.69	1.29	1.39
1	C	158	TYR	CE2-CZ	-6.55	1.30	1.38
1	A	326	LYS	CE-NZ	6.50	1.65	1.49

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	84	TYR	CD2-CE2	-6.30	1.29	1.39

The worst 5 of 402 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	119	LEU	CB-CG-CD2	-14.83	85.79	111.00
1	A	165	LEU	CB-CG-CD2	-12.70	89.41	111.00
1	G	38	LEU	CA-CB-CG	-12.12	87.41	115.30
1	O	38	LEU	CA-CB-CG	-10.97	90.07	115.30
1	M	165	LEU	CB-CG-CD2	-10.97	92.36	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	O	321	TYR	Sidechain

## 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	3153	0	3158	482	0
1	C	3136	0	3141	458	0
1	E	3136	0	3139	465	0
1	G	3136	0	3141	463	0
1	I	3136	0	3141	420	0
1	K	3136	0	3141	391	0
1	M	3136	0	3141	373	0
1	O	3136	0	3141	435	0
2	B	119	0	107	22	0
2	D	119	0	107	9	0
2	F	119	0	107	6	0
2	H	119	0	107	9	0
2	J	119	0	107	12	0
2	L	119	0	107	12	0
2	N	119	0	107	11	0
2	P	119	0	107	10	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	26057	0	25999	3442	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 66.

The worst 5 of 3442 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:C:452:VAL:HG13	1:C:455:ILE:CD1	1.46	1.43
1:C:452:VAL:CG1	1:C:455:ILE:HD12	1.47	1.43
1:E:198:LEU:HD13	1:E:199:TYR:CE1	1.57	1.38
1:E:194:SER:HB3	1:E:199:TYR:OH	1.20	1.27
1:I:281:LYS:HD2	1:I:282:GLY:N	1.51	1.23

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	389/465 (84%)	301 (77%)	73 (19%)	15 (4%)	3	17
1	C	387/465 (83%)	299 (77%)	71 (18%)	17 (4%)	2	15
1	E	387/465 (83%)	300 (78%)	67 (17%)	20 (5%)	2	12
1	G	387/465 (83%)	301 (78%)	75 (19%)	11 (3%)	5	25
1	I	387/465 (83%)	303 (78%)	70 (18%)	14 (4%)	3	19
1	K	387/465 (83%)	305 (79%)	66 (17%)	16 (4%)	3	16
1	M	387/465 (83%)	308 (80%)	68 (18%)	11 (3%)	5	25
1	O	387/465 (83%)	309 (80%)	67 (17%)	11 (3%)	5	25
2	B	12/22 (54%)	9 (75%)	2 (17%)	1 (8%)	1	4

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	D	12/22 (54%)	10 (83%)	1 (8%)	1 (8%)	1	4
2	F	12/22 (54%)	10 (83%)	2 (17%)	0	100	100
2	H	12/22 (54%)	10 (83%)	1 (8%)	1 (8%)	1	4
2	J	12/22 (54%)	11 (92%)	1 (8%)	0	100	100
2	L	12/22 (54%)	10 (83%)	1 (8%)	1 (8%)	1	4
2	N	12/22 (54%)	11 (92%)	1 (8%)	0	100	100
2	P	12/22 (54%)	9 (75%)	2 (17%)	1 (8%)	1	4
All	All	3194/3896 (82%)	2506 (78%)	568 (18%)	120 (4%)	3	18

5 of 120 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	17	PRO
1	A	37	PHE
1	A	42	GLN
1	A	82	ALA
1	A	201	ILE

### 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	350/411 (85%)	251 (72%)	99 (28%)	0	2
1	C	348/411 (85%)	259 (74%)	89 (26%)	0	3
1	E	348/411 (85%)	251 (72%)	97 (28%)	0	2
1	G	348/411 (85%)	247 (71%)	101 (29%)	0	2
1	I	348/411 (85%)	254 (73%)	94 (27%)	0	2
1	K	348/411 (85%)	258 (74%)	90 (26%)	0	2
1	M	348/411 (85%)	248 (71%)	100 (29%)	0	2
1	O	348/411 (85%)	254 (73%)	94 (27%)	0	2
2	B	14/22 (64%)	13 (93%)	1 (7%)	14	46

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	D	14/22 (64%)	11 (79%)	3 (21%)	1	5
2	F	14/22 (64%)	13 (93%)	1 (7%)	14	46
2	H	14/22 (64%)	11 (79%)	3 (21%)	1	5
2	J	14/22 (64%)	9 (64%)	5 (36%)	0	1
2	L	14/22 (64%)	12 (86%)	2 (14%)	3	15
2	N	14/22 (64%)	11 (79%)	3 (21%)	1	5
2	P	14/22 (64%)	12 (86%)	2 (14%)	3	15
All	All	2898/3464 (84%)	2114 (73%)	784 (27%)	0	2

5 of 784 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	276	VAL
1	I	214	HIS
1	O	185	TRP
1	G	314	GLU
1	I	22	PHE

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

Mol	Chain	Res	Type
1	G	41	GLN
1	I	105	ASN
1	O	152	GLN
1	G	105	ASN
1	G	220	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 [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 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, 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	395/465 (84%)	-0.19	1 (0%) 94 84	5, 40, 76, 100	0
1	C	393/465 (84%)	-0.38	2 (0%) 91 75	2, 34, 76, 100	0
1	E	393/465 (84%)	-0.12	4 (1%) 82 59	2, 42, 78, 100	0
1	G	393/465 (84%)	-0.43	1 (0%) 94 84	4, 35, 72, 100	0
1	I	393/465 (84%)	-0.17	2 (0%) 91 75	4, 41, 76, 100	0
1	K	393/465 (84%)	-0.31	2 (0%) 91 75	1, 39, 74, 100	0
1	M	393/465 (84%)	-0.24	4 (1%) 82 59	5, 41, 76, 100	0
1	O	393/465 (84%)	-0.29	3 (0%) 86 65	4, 39, 78, 93	0
2	B	14/22 (63%)	0.12	1 (7%) 16 5	42, 59, 84, 85	0
2	D	14/22 (63%)	-0.14	0 100 100	21, 55, 77, 80	0
2	F	14/22 (63%)	0.08	1 (7%) 16 5	29, 59, 77, 91	0
2	H	14/22 (63%)	-0.46	0 100 100	25, 42, 75, 78	0
2	J	14/22 (63%)	0.09	1 (7%) 16 5	13, 54, 74, 90	0
2	L	14/22 (63%)	-0.12	0 100 100	23, 59, 88, 100	0
2	N	14/22 (63%)	0.09	1 (7%) 16 5	15, 53, 71, 81	0
2	P	14/22 (63%)	-0.22	0 100 100	19, 50, 86, 98	0
All	All	3258/3896 (83%)	-0.26	23 (0%) 87 69	1, 39, 77, 100	0

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	328	GLY	5.2
1	C	328	GLY	4.9
1	K	328	GLY	4.3
1	O	328	GLY	3.3
2	N	860	ILE	3.2

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.