

wwPDB X-ray Structure Validation Summary Report (i)

Mar 4, 2024 – 11:04 PM EST

PDB ID	:	1Y1V
Title	:	Refined RNA Polymerase II-TFIIS complex
Authors	:	Kettenberger, H.; Armache, KJ.; Cramer, P.
Deposited on	:	2004-11-19
Resolution	:	3.80 Å(reported)

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

We welcome your comments at validation@mail.wwpdb.org A user guide is available at https://www.wwpdb.org/validation/2017/XrayValidationReportHelp with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
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)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: $X\text{-}RAY\;DIFFRACTION$

The reported resolution of this entry is 3.80 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Motria	Whole archive	Similar resolution
wietric	$(\# {\rm Entries})$	$(\# { m Entries}, { m resolution} { m range}({ m \AA}))$
R _{free}	130704	1212 (4.00-3.60)
Clashscore	141614	1288 (4.00-3.60)
Ramachandran outliers	138981	1243 (4.00-3.60)
Sidechain outliers	138945	1237 (4.00-3.60)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Mol	Chain	Length			Quality of chair	1	
1	А	1733	21%		47%	13%	• 18%
2	В	1224	24%		52%	1	.3% • 9%
3	С	318	22%		48%	12%	• 16%
4	D	221	33%		40%	6%	20%
5	Е	215	34%		55%		11%
6	F	155	15%	30%	8%	46%	
7	G	171	33%		59%	0	8%



Mol	Chain	Length		Quality of chain	
8	Н	146	30%	53%	8% 9%
9	Ι	122	31%	48%	13% 5% ·
10	J	70	23%	51%	17% • 7%
11	K	120	37%	52%	5% • 5%
12	L	70	6% 41	% 19%	34%
13	S	179		62%	28% •••



2 Entry composition (i)

There are 15 unique types of molecules in this entry. The entry contains 31803 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 DNA-directed RNA polymerase II largest subunit.

Mol	Chain	Residues		Α	toms		ZeroOcc	AltConf	Trace	
1	А	1426	Total 11214	C 7069	N 1959	O 2124	S 62	0	0	0

• Molecule 2 is a protein called DNA-directed RNA polymerase II 140 kDa polypeptide.

Mol	Chain	Residues		A	toms		ZeroOcc	AltConf	Trace	
2	В	1112	Total 8837	C 5594	N 1548	O 1640	${ m S}{55}$	58	0	0

• Molecule 3 is a protein called DNA-directed RNA polymerase II 45 kDa polypeptide.

Mol	Chain	Residues		At	oms		ZeroOcc	AltConf	Trace	
3	С	266	Total 2095	C 1317	N 348	0 417	S 13	0	0	0

• Molecule 4 is a protein called DNA-directed RNA polymerase II 32 kDa polypeptide.

Mol	Chain	Residues		At	oms		ZeroOcc	AltConf	Trace	
4	D	177	Total 1356	C 840	N 241	0 273	${ m S} { m 2}$	0	0	0

• Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III 27 kDa polypep-tide.

Mol	Chain	Residues		At	oms			ZeroOcc	AltConf	Trace
5	Е	214	Total 1752	C 1111	N 309	0 321	S 11	0	0	0

• Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide.



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	84	Total 679	C 434	N 115	O 127	${ m S} { m 3}$	0	0	0

• Molecule 7 is a protein called DNA-directed RNA polymerase II 19 kDa polypeptide.

Mol	Chain	Residues		At	oms		ZeroOcc	AltConf	Trace	
7	G	171	Total 1340	C 861	N 222	O 249	S 8	0	0	0

• Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide.

Mol	Chain	Residues		At	oms			ZeroOcc	AltConf	Trace
8	Н	133	Total 1068	C 673	N 180	0 211	${f S}{4}$	0	0	0

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

Mol	Chain	Residues		\mathbf{A}^{\dagger}	toms			ZeroOcc	AltConf	Trace
9	Ι	119	Total	C	N 170	0	S 10	0	0	0
			971	596	179	180	10			

• Molecule 10 is a protein called DNA-directed RNA polymerases I/II/III subunit 10.

Mol	Chain	Residues		Ato	\mathbf{ms}			ZeroOcc	AltConf	Trace
10	J	65	Total 532	C 339	N 93	0 94	S 6	0	0	0

• Molecule 11 is a protein called DNA-directed RNA polymerase II 13.6 kDa polypeptide.

Mol	Chain	Residues		At	oms			ZeroOcc	AltConf	Trace
11	K	114	Total 919	$\begin{array}{c} \mathrm{C} \\ 590 \end{array}$	N 156	0 171	${ m S} { m 2}$	0	0	0

• Molecule 12 is a protein called DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide.

Mol	Chain	Residues		Ato	\mathbf{ms}			ZeroOcc	AltConf	Trace
12	L	46	Total 364	C 224	N 72	O 64	${S \atop 4}$	0	0	0

• Molecule 13 is a protein called Transcription elongation factor S-II.



Mol	Chain	Residues		Ate	\mathbf{oms}			ZeroOcc	AltConf	Trace
13	S	174	Total 666	C 454	N 99	0 108	${ m S}{ m 5}$	0	0	104

• Molecule 14 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
14	А	2	Total Zn 2 2	0	0
14	В	1	Total Zn 1 1	0	0
14	С	1	Total Zn 1 1	0	0
14	Ι	2	Total Zn 2 2	0	0
14	J	1	Total Zn 1 1	0	0
14	L	1	Total Zn 1 1	0	0
14	S	1	Total Zn 1 1	0	0

• Molecule 15 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
15	S	1	Total Mg 1 1	0	0



3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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 largest subunit



V743	K744 N745	M746	V747	M748	A/49 G750	S751	K752	G754 S754	F755	1756	N757	I758	A/ 59 0760	M761	S762	W7.65	G766	Q767	Q7 68	S769	E771		R774	1775 1776	ET77	G778	F779 V780	D781	R782	T783	L/ 04 P7 85	H786	K789	D7 90	D791	P794	E795	S796	K797 C708	F7 99	V800	E801 NR02	5803 5803	Y804	
L808	T809 D810	0811 0811	E812	F813	F815	H816	A817	01818 0810	G820	R821	E822	G823	L825 T825	D826	T827	A828 V820	K830	T831	A832	E833 T024	1004	1837	<mark>0838</mark>	R839 D040	L841	V842	K843	L845	E846	D847	1040 M849	V850	H851 Y852	D853	N854	1000 T856	R857	N858	5859 1860	G861	N862	V863 1864	1865 0865	F866	1867 1868
G 869	E870 D871	G872	M873	D874	A876 A876	H877	1878 7070	E879 K880	0881	S882	L883		1887 1887	G888	<mark>S889</mark>	D890	A 892	F893	E894	K895 P806	7897	R898	V899	D900	L902	N9 03	T904	906H	T907	L908	L913	E914	29162	1919	L920	D922	L923	K924	L925 D926	V927	L928	L929	E931	E932	Y933 K934
0 935	L936 V037	K938	D939	R940	F942	L943	R944	E945 V946	F947	V948	D949		1.956	P957	V958	N959 T960	R961	R962	I 963	1964 0065	N966	A967	Q968	0969 1070	1970 F971		5979 D080	L981	T982	1983 V004	D985	1986	V 987 L 988	G 989	V990 V001	D992	L993	Q 994	E995 Noor	L997	L998	B1001	G1002	K1003	N1004 E1005
I1006	11007	N1009	A1010	Q1011	V1015	T1016	L1017	F1018 C1019	C1020	L1021	L1022	R1023	81025 R1025	L1026	A1027	T1028 B1029	R1030	V1031	L1032	Q1033	E1035 Y1035	R1036	L1037	T1038 V1030	01040	A1041		V1045	L1046	S1047 N1048	N1040	E1050	A1051 01052	F1053	L1054	81056	V1057	V1058	H1059 P1060	G1061	E1062	M1063 V1064	G1065	V1066	L1067 A1068
A1069	Q1070	11072	G1073	E1074	A1076	T1077	Q1078	M1079 T1080	L1081	N1082	T1083	F1084	F1086	A1087	G1088	V1089	S1091	K1092	K1093	V1094 T1005	S1096	G1097	V1098	P1099	L1101	K1102	E1103	L1105	N1106	V1107	N1110	M1111	K1112 T1113	P1114	S1115 1116	T1117	V1118	Y1119	L1120 F1121	P1122	G1123	H1124 A1125	A1126	D1127	u1128 E1129
Q1130	A1131 K1132	L1133	I1134	R1135	A1137	11138	E1139 11130	H1140 T1141	T1142	L1143	K1144	S1145	V1140 T1147	11148	A1149	S1150 F1161	11152 11152	Y1153	Y1154	D1155	D1157	P1158	R1159	S1160 T1161	V1162	I1163	P1164 E1165	D1166	E1167	E1168 T1160	11109 11170	Q1171	L1172 H1173	F1174	S1175 1176	TEU	ASP	GLU	GLU	GLU	GLN	SER DHF	ASP	Q1187	u1188 S1189
P1190	W1191 11192	L1193	R1194		D1198 R1199		M1202	N1203	K1205	D1206	L1207	T1208	G1210	Q1211	V1212	G1213 E1214	R1215	11216		T1219 E1220	K1221	N1222	D1223	L1224 E1225	V1226	11227	W1228 81228	E1230	D1231	N1232	L1236	11237	11238 R1239	C1240	R1241 V1243	V1243	ARG	PRO	LYS SFR	TEU	ASP	ALA	THR	GLU	A1254 E1255
E1256	T 1 DEO	K1261	K1262	11263	E1 264 N1 265	T1266	M1267	T1971	T1272	L1273	R1274	G1275	V1276 F1277	N1278	11279	E1280 B1281	V1282	V1283	M1284	M1285	R1289	K1290	V1291	P1292	51233 P1294	T1295	G1296 E1207	Y1298	V1299	K1300 E1301	F1301		V1305 L1306	E1307	T1308	G1310	V1311	N1312	L1313 C1 21 4	E1315	V1316	M1317 T1318	V1319		11322 D1323
P1324	T1325 B1376	11327	Y1328	T1329	N1330 S1331	F1332	I 1333	D1334 T1335	M1336	E1337	V1338	L1339	61340 11341	E1342	A1343	G1344 B1345	A1346	A1347	L1348	Y1349	E1351	V1352	Y1353	N1354		<mark>S1358</mark>	D1359 C1360	S1361	Y1362	V1363 N1364	Y1365	R1366	A1369	L1370	L1371 V1372	D1373	V1374	M1375	T1376 T1377	Q1378	G1379	G1380 11381		V1384	T1385 R1386
H1387	G1388	R1391	S1392	N1393	11394 G1395	A1396	L1397	M1398 R1399	C1400	S1401	F1402	E1403	E1404 T1405	V1406		L1409 F1410	E1411	A1412	G1413	0 7 7 7	E1417	L1418	D1419	D1420	C1421 R1422	G1423	V1424	E1426	N1427	V1428	L1429 L1430	G1431	41432 M1433	A1434	P1435	G1437	T1438	G1439	A1440 F1221	D1442	V1443	M1444	D1446	E1447	E1448 S1449
L1450	V1451 K1452	Y1453	M1454	P1455	GL.N	LYS	ILE	THK GI II	ILE	GLU	ASP	GLY	ASP	GLY	GLY	VAL THP	PRO	TYR	SER	ASN	SER	GLY	LEU	VAL	ALA	ASP	LEU	VAL	LYS	ASP	LEU	MET	PHE SER	PRO	LEU	ASP	SER	GLY	SER	ASP	ALA	MET at a	GLY	GLY	THR
ALA	TYR	GLY	ALA	ASP	1 T.K. GT.Y	GLU	ALA	THK	PRO	PHE	GLY	ALA	1 T.K. GT.Y	GLU	ALA	PRO THB	SER	PRO	GLY	PHE	VAL.	SER	SER	PRO	PHE	SER	PRO THD	SER	PRO	THR	SER	PRO	SER	PRO	ALA TVD	SER	PRO	THR	SER	SER	TYR	SER	THR	SER	PRU
TYR	SER	THR	SER	PRO	TYR	SER	PRO	THK	PRO	SER	TYR	SER	THR	SER	PRO	SER	SER	PRO	THR	SER	SER	TYR	SER	PRO	SER	PRO	SER	SER	PRO	THR	PRO	SER	TYR	PRO	THR	PRO	SER	TYR	SER	THR	SER	PRO	TYR	SER	THR
SER	PRO	TYR	SER	PRO	SER	PRO	SER	TYR	PRO	THR	SER	PRO	TYR	SER	PRO	THR	PRO	SER	TYR	SER	THR	SER	PRO	ALA	SER	PRO	THR	PRO	SER	TYR	PRO	THR	PRO	SER	TYR	PRO	THR	SER	PRO	TYR	SER	PRO THR	SER	PRO	TYR
SER	PR.0 THR	SER	PRO	ASN	SER	PRO	THR	DRU	SER	TYR	SER	PRO	SER	PRO	GLY	TYR	PRO	GLY	SER	PRO AT A	TYR	SER	PRO	LYS	ASP	GLU	GLN	SIH	ASN	GLU	GLU	ASN	ARG												

• Molecule 2: DNA-directed RNA polymerase II 140 kDa polypeptide



1	Y	1	ľ	Τ
-	-	_		

С	ha	in	В	: •				24	%				•									52	%										13	%		•	9	%	-				
MET	SER ASP	LEU	ALA ASN	SER	LYS	TYR	T YR ASP	GLU	ASP	PRO TVD	GLY	PHE	GLU	E21	S22	A23	r 24 125	T26	A27	D29	S30	W31	134	535 S35	A36	F37 F38	R39	E40 V/1	G42	L43	V 44 S 45	Q46	u447 L48	D49	850 F51	N52	053 	F54 VFF	D56	Y57 #50	1.50 L59	<mark>д60</mark> Л61	TOT
162	I63 C64	E65	S67	T68	170 I70	LEU	GLU	TEU	ALA	GLN	THR	THR	GLU	ASP	ASN	ILE	ARG	LYS	TYR	פרט 190	S91		K94 T95	Y96	797 797	198 Kqq	P100	M101 W103		V108	60 L.I.	L112	P113	Q115	E116 A117	R118	L119	R120 N121	L122	T123	1124 S125	S126 C127	1715
L128	F129 V130	D131	V132 K133	K134 ABC	THR	TYR	GLU AT.A	ILE	ASP	VAL	GLY	ARG	GLU	LYS	TYR	GLU	ILE	ALA	GLU	SER	GLU	ASP	ASP SFR	GLU	SER	GLY K164	V165	F166 T167	G168	R169	P171	1172	M1/3 L174	R175	S176 K177	N178	C179	Y180 1.181	S182	0 074	E160 S187	D188 1.189	PT02
	L192 K193	E194	P196	F197	G200	G201	Y202 F203	1204	I 205	N206	S208	E209	K210 V211	1212 L212	1213	A214	4213 E216	R217	S218 A710	G220	N221	I 222	V223	V225	F226	K227 K738	A229	A230 D231	r 231 S232	P233	1234 S235	H236	V 23/ A 238	E239	I240 R241	S242		G247 S248	R249	C LO D	5202 T253	L254 0255	007h
V256	K257 L258	Y259	4260 R261	1900 000	2203 A266	R267	T268 T269	K270	A271	T272	P274	Y275	1276 V277	0278	D279	1280 D284	1282 1282	V283	1284 1705	1205 F286	R287	A288	1001	1292	P293	D294 C295	E296	1297 1209	E299	H300	Y303		0309 0309	M310	M313	L314	K315	P316 C317		F322	V323 I324	Q325 D326	0707
R327	E328 T329	A330	L331 D332	F333 T334	1335 G335	ARG	ARG GLY	THR	ALA	CI V	ILE	LYS	K345 F346	K347	R348	I349	¥351		1355 1356	F-200	E359	F360	L361 D360	H363	1364	1365 1366	L367	E368 C360	6309 F370	E371	8372 R373	K374	A3/5 F376	F377	L378 G379	Y380	M381	I 382	L385	L386	L30/ C388	A389 1.390	F020
D391	R392 K393	D394	4395 D396	107 107 107	G402	K403	K404 R405	L406	D407	L408	6410 6410	P411	L412 1413	A414		F417 17440	T419 T419	L420	F421	L424		D427	I428 F429	R430	Y431	M432	E437	GLU AT A	ALA	ASP	ASN	MET	LYS L446	A447	A450		1453	T454	L457	K458	1453 A460	L461 4462	707V
T463	G464 N465	1/466 CAET	G407 E468	Q469	S474	S475	R476	S480	Q481	V482 1 /83	N484	R485	Y486 TAB7	1488 Y488	S489	5490 7404	1491 L492	S493	H494	L490 R496	R497	T498	N499	I 502	G503	R504	K507	L508 A 509	K510	P511	K512 0513	L514	GICH N516	T517	H518 W519	G520	L521	V522 C523	P524	A525 TECE	E520 T527	P528	E072
G530	0531 A532	C533	4034 L535	V536 V536	N537 N538	L539	S540 1.541	M542	S543	C544 TEAE	S546	V547	G548 TEAO	D550	P551	M552	1554	1555	T556 EEE7	r 55/ L558	S559	E560	W561 CERO	M563	E564	P565 L566		Y569 VE70	P571	H572	Ub/3 S574	P575		T578	R579 V580	F581	V582	N583 G584	V585	W586	G588	V589 H590	Decu
R591	N592 P593	A594	M597	E598	1009	R601	T602 L603	R604	R605	K606	D608	1609	N610 D611	E612	V613	S614 M61E	1616 1616	R617	D618	LOI9 R620	E621	K622	E623 1624	K625	1626 1626	F627 T628	D629	A630 C631	R632	V633	Y634 R635	P636	L03/ F638	1639	V640 F641	D642	D643	E644 S645	L646	G647	но4 о К649	VESS	COOA
R654	H657	I 658	K660 K660	L661 MEED	A663	T664	E665	D668	ILE	GLU	GLY	PHE	GLU	VAL	GLU	E678 ve70	16/9 T680	W681	S682	2003 L684	L685	N686	E687 C688	L689	V690	Е691 V692	I 693	D694	CEON	E698	E699 S700	1701	L/02 1703	A704	M705	E708	D709	L710 F711	P712	u F	ASN	GLU	ALU
ASN	ASP LEU	D722 V723	V/ 23 D7 24	P725	A/ 20 K7 27		R730 V731		T737	F738 T730	H740	C741	E742	H7 44	P745	S746 M747	I748	L749	G750 V7551	A752	A753	S754	1755 1756	P757	F758	P759	H761	N762 0763	47 03 S764	P765	R/66 N767	T768	1/69/1 0770	S771	A772 M773	G774	K775	0776 ≜777	M778	G779	V 60 F781	L782 T783	CO 1 T
N784	Y785 N786	V787 5766	R/80 M789	COTM	A793	N794	1795 1.796	797	Y798		K801	P802	TODE	1806 T806	R807	A808 W800	E810	Y811	L812 V012	F814	R815	E816	L817 D818	A819	G820	0821 N822	A823	I824 Vere	4 0 2 3 A 8 2 6	1827	A828 C829	Y830	6832 G832	Y833	N834 D835	E836	D837	5838 M839	1840	M841	N 04 2 0 843	5844 5845	0400
1846 	D847 R848	G849 1 0 5 0	F851	R852	2003 L854	F855	F856 R857	2858	Y859	M860 D861	000 0862	E863	K864 V065	Y866	<mark>G867</mark>	M868	1870	T871	E872 T073	10/3 F874		P877	0878 8879	T880	N881	1882 1883	R884		TEOT	D895	T898	1899	P901	G902	V903 R904	V905	2906S	G907 F908	D909	V910 7011	1911 1912	G913 K914	4T CU
	P917 1918	S919 Dag	ASP	GLU	GLU	LEU	GLY	ARG	THR	ALA TVD	HIS	S933	K934 D026	D936	A937	S938	1939 P940	L941	R942	7944	E945	N946	G947 T948	V949	D950	(1951 V952	L953	V954 TOEE	1900 T956	N957	0959 D959	0960 1001	L961 K962	F963	V964 K965	V966	-	T970 T971	K972	1973	r9/4 0975	1976 1977	1 100







 \bullet Molecule 5: DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide







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





Y282 1/286 1/286 7/286 7/286 7/286 7/286 7/286 7/286 7/286 7/286 7/296 7/200 7/206 7/206 7/206 7/206 7/206 7/206 7/206 7/206 7/206 7/206 7/206 7/206 7/206 7/206 7/2000 7/2000 7/2000 7/2000 7/2000 7/2000 7/2000 7/2000 7/2000 7/20



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants	218.90Å 395.30Å 281.00Å	Deneriten
a, b, c, α , β , γ	90.00° 90.00° 90.00°	Depositor
Bosolution(A)	50.00 - 3.80	Depositor
Resolution (A)	49.41 - 3.74	EDS
% Data completeness	(Not available) $(50.00-3.80)$	Depositor
(in resolution range)	86.6(49.41-3.74)	EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	4.52 (at 3.77 Å)	Xtriage
Refinement program	CNS	Depositor
D D	0.282 , 0.294	Depositor
Λ, Λ_{free}	0.257 , 0.268	DCC
R_{free} test set	2439 reflections $(1.99%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	68.0	Xtriage
Anisotropy	0.336	Xtriage
Bulk solvent $k_{sol}(e/A^3)$, $B_{sol}(A^2)$	0.28 , 42.2	EDS
L-test for twinning ²	$< L > = 0.30, < L^2 > = 0.13$	Xtriage
Estimated twinning fraction	0.199 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l	Vtriago
Estimated twinning fraction	0.206 for 1/2 *h + 1/2 *k, 3/2 *h - 1/2 *k, -1	Atnage
F_o, F_c correlation	0.81	EDS
Total number of atoms	31803	wwPDB-VP
Average B, all atoms $(Å^2)$	72.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 1.74% of the height of the origin peak. No significant pseudotranslation is detected.

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



¹Intensities estimated from amplitudes.

5 Model quality (i)

5.1 Standard geometry (i)

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).

Mal	Chain	Bo	ond lengths	B	ond angles
MIOI	Unam	RMSZ	# Z > 5	RMSZ	# Z > 5
1	А	0.50	2/11417~(0.0%)	0.78	4/15442~(0.0%)
2	В	0.51	4/9009~(0.0%)	0.76	8/12146~(0.1%)
3	С	0.48	0/2133	0.77	1/2891~(0.0%)
4	D	0.41	0/1365	0.64	0/1837
5	Е	0.43	0/1788	0.66	0/2406
6	F	0.52	0/691	0.77	0/933
7	G	0.49	0/1368	0.72	0/1844
8	Н	0.38	0/1086	0.65	1/1470~(0.1%)
9	Ι	0.46	0/989	0.77	1/1331~(0.1%)
10	J	0.48	0/541	0.75	0/727
11	Κ	0.45	0/937	0.67	0/1265
12	L	0.54	0/366	0.79	0/485
13	S	1.31	4/571~(0.7%)	1.64	7/765~(0.9%)
All	All	0.51	10/32261~(0.0%)	0.77	22/43542~(0.1%)

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
2	В	0	3
13	S	0	2
All	All	0	5

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
13	S	269	PHE	C-N	-16.91	0.95	1.34
2	В	467	GLY	C-O	-11.91	1.04	1.23
13	S	260	THR	CA-CB	10.48	1.80	1.53



Continued from previous page...

Mol	Chain	Res	Type	Atoms	Ζ	Observed(Å)	Ideal(Å)
2	В	468	GLU	CB-CG	8.39	1.68	1.52
13	S	268	ARG	CG-CD	6.05	1.67	1.51

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
13	S	269	PHE	O-C-N	-19.02	92.27	122.70
13	S	269	PHE	C-N-CA	16.73	163.51	121.70
13	S	269	PHE	CA-C-N	16.08	152.57	117.20
1	А	195	ASP	N-CA-C	9.35	136.25	111.00
2	В	510	LYS	CB-CA-C	-7.63	95.14	110.40

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	В	510	LYS	Mainchain
2	В	785	TYR	Sidechain
2	В	833	TYR	Sidechain
13	S	269	PHE	Sidechain,Peptide

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	А	11214	0	11281	1514	0
2	В	8837	0	8871	1206	0
3	С	2095	0	2052	260	0
4	D	1356	0	1319	101	0
5	Е	1752	0	1776	200	0
6	F	679	0	701	82	0
7	G	1340	0	1357	159	0
8	Н	1068	0	1040	115	0
9	Ι	971	0	929	110	0
10	J	532	0	542	103	0
11	K	919	0	929	96	0
12	L	364	0	387	68	0



Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
13	S	666	0	553	105	0
14	А	2	0	0	0	0
14	В	1	0	0	0	0
14	С	1	0	0	0	0
14	Ι	2	0	0	0	0
14	J	1	0	0	0	0
14	L	1	0	0	0	0
14	S	1	0	0	0	0
15	S	1	0	0	0	0
All	All	31803	0	31737	3774	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 60.

The worst 5 of 3774 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)
13:S:260:THR:CA	13:S:260:THR:CB	1.80	1.53
13:S:269:PHE:CZ	13:S:297:CYS:SG	2.04	1.50
13:S:269:PHE:CE2	13:S:297:CYS:SG	2.14	1.39
1:A:1230:GLU:OE2	13:S:201:ILE:CA	1.75	1.32
1:A:1283:VAL:CG1	13:S:256:ALA:O	1.78	1.31

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	Perce	ntiles
1	А	1418/1733 (82%)	914 (64%)	316 (22%)	188 (13%)	0	4
2	В	1096/1224~(90%)	726 (66%)	223 (20%)	147 (13%)	0	4



Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
3	С	264/318~(83%)	169 (64%)	62 (24%)	33 (12%)	0 6
4	D	173/221~(78%)	129 (75%)	27 (16%)	17 (10%)	0 10
5	Е	212/215~(99%)	141 (66%)	50 (24%)	21 (10%)	0 10
6	F	82/155~(53%)	60 (73%)	15 (18%)	7 (8%)	1 12
7	G	169/171~(99%)	123 (73%)	34 (20%)	12 (7%)	1 17
8	Н	129/146~(88%)	93 (72%)	26 (20%)	10 (8%)	1 15
9	Ι	117/122~(96%)	80 (68%)	22 (19%)	15 (13%)	0 5
10	J	63/70~(90%)	36 (57%)	14 (22%)	13 (21%)	0 2
11	К	112/120~(93%)	82 (73%)	25 (22%)	5 (4%)	2 25
12	L	44/70~(63%)	18 (41%)	14 (32%)	12 (27%)	0 0
13	S	68/179~(38%)	51 (75%)	10 (15%)	7 (10%)	0 9
All	All	3947/4744~(83%)	2622 (66%)	838 (21%)	487 (12%)	0 6

5 of 487 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	А	4	GLN
1	А	48	ALA
1	А	55	ASP
1	А	58	LEU
1	А	62	ASP

Protein sidechains (i) 5.3.2

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	А	1246/1520~(82%)	1133 (91%)	113 (9%)	9	36
2	В	964/1061~(91%)	880 (91%)	84 (9%)	10	38
3	С	234/274~(85%)	205~(88%)	29 (12%)	4	24
4	D	140/200~(70%)	126 (90%)	14 (10%)	7	32
5	Ε	196/197~(100%)	184 (94%)	12 (6%)	18	50





Mol	Chain	Analysed	Rotameric	Outliers	Per	Percentiles	
6	F	74/137~(54%)	63~(85%)	11 (15%)	3	18	
7	G	152/152~(100%)	143 (94%)	9 (6%)	19) 51	
8	Н	117/128~(91%)	110 (94%)	7~(6%)	19) 50	
9	Ι	113/116~(97%)	97~(86%)	16 (14%)	3	21	
10	J	60/65~(92%)	55~(92%)	5 (8%)	1	40	
11	Κ	99/102~(97%)	91~(92%)	8 (8%)	1	41	
12	L	40/57~(70%)	33~(82%)	7 (18%)	2	13	
13	S	62/156~(40%)	55(89%)	7 (11%)	6	28	
All	All	3497/4165~(84%)	3175 (91%)	322 (9%)	9	35	

5 of 322 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	D	148	LEU
9	Ι	46	HIS
4	D	197	SER
6	F	143	PHE
11	Κ	10	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 91 such sidechains are listed below:

Mol	Chain	\mathbf{Res}	Type
2	В	1179	GLN
4	D	137	ASN
3	С	65	HIS
3	С	167	HIS
5	Е	143	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

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



5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 10 ligands modelled in this entry, 10 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 (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
13	S	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	S	269:PHE	С	270:THR	Ν	0.95



6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

