



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 06:02 am BST

PDB ID : 4YFX  
Title : Escherichia coli RNA polymerase in complex with Myxopyronin B  
Authors : Molodtsov, V.; Fleming, P.R.; Eyermann, C.J.; Ferguson, A.D.; Foulk, M.A.;  
McKinney, D.C.; Masse, C.E.; Buurman, E.T.; Murakami, K.S.  
Deposited on : 2015-02-25  
Resolution : 3.84 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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
buster-report : 1.1.7 (2018)  
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

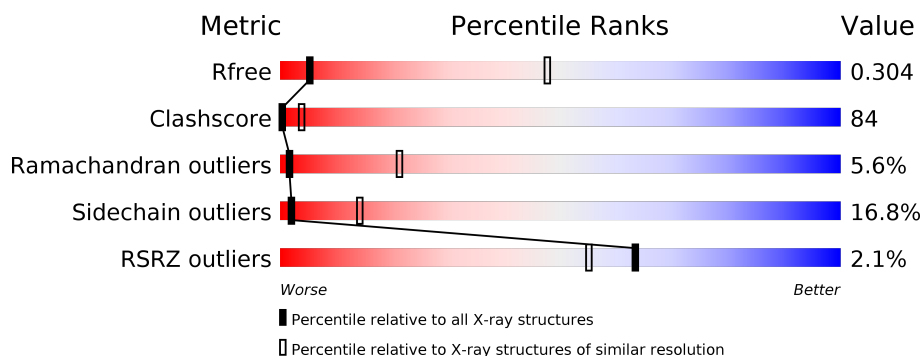
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.84 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))
$R_{free}$	130704	1242 (4.08-3.60)
Clashscore	141614	1004 (4.04-3.64)
Ramachandran outliers	138981	1003 (4.06-3.62)
Sidechain outliers	138945	1266 (4.08-3.60)
RSRZ outliers	127900	1149 (4.08-3.60)

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 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	329	<div> <div>13%</div> <div>44%</div> <div>12%</div> <div>•</div> <div>30%</div> </div>
1	B	329	<div> <div>2%</div> <div>8%</div> <div>48%</div> <div>28%</div> <div>•</div> <div>12%</div> </div>
1	G	329	<div> <div>•</div> <div>14%</div> <div>47%</div> <div>8%</div> <div>31%</div> </div>
1	H	329	<div> <div>7%</div> <div>13%</div> <div>42%</div> <div>11%</div> <div>34%</div> </div>
2	C	1342	<div> <div>17%</div> <div>66%</div> <div>15%</div> <div>•</div> </div>
2	I	1342	<div> <div>2%</div> <div>20%</div> <div>65%</div> <div>14%</div> <div>•</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	1407	<div><div></div><div></div><div></div><div></div><div></div></div> <div>12%52%17%•17%</div>
3	J	1407	<div>%<div></div><div></div><div></div><div></div><div></div></div> <div>13%53%15%•18%</div>
4	E	91	<div><div></div><div></div><div></div><div></div><div></div></div> <div>22%67%9%•</div>
4	K	91	<div><div></div><div></div><div></div><div></div><div></div></div> <div>8%19%60%8%13%</div>
5	F	613	<div><div></div><div></div><div></div><div></div><div></div></div> <div>4%16%46%12%•25%</div>
5	L	613	<div><div></div><div></div><div></div><div></div><div></div></div> <div>3%17%48%9%25%</div>

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 55388 atoms, of which 33 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 subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	229	Total	C	N	O	S	0	0	0
			1779	1108	316	349	6			
1	B	289	Total	C	N	O	S	0	0	0
			2239	1403	393	435	8			
1	G	227	Total	C	N	O	S	0	0	0
			1755	1093	311	345	6			
1	H	216	Total	C	N	O	S	0	0	0
			1662	1038	292	326	6			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1340	Total	C	N	O	S	0	0	0
			10564	6628	1838	2055	43			
2	I	1340	Total	C	N	O	S	0	0	0
			10552	6621	1835	2053	43			

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1166	Total	C	N	O	S	0	0	0
			9062	5701	1622	1693	46			
3	J	1155	Total	C	N	O	S	0	0	0
			9021	5675	1617	1683	46			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	89	Total	C	N	O	S	0	0	0
			691	421	129	140	1			
4	K	79	Total	C	N	O	S	0	0	0
			627	382	118	126	1			

- Molecule 5 is a protein called RNA polymerase sigma factor RpoD.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	458	Total	C	N	O	S	0	0	0
			3726	2332	668	703	23			
5	L	458	Total	C	N	O	S	0	0	0
			3640	2282	647	690	21			

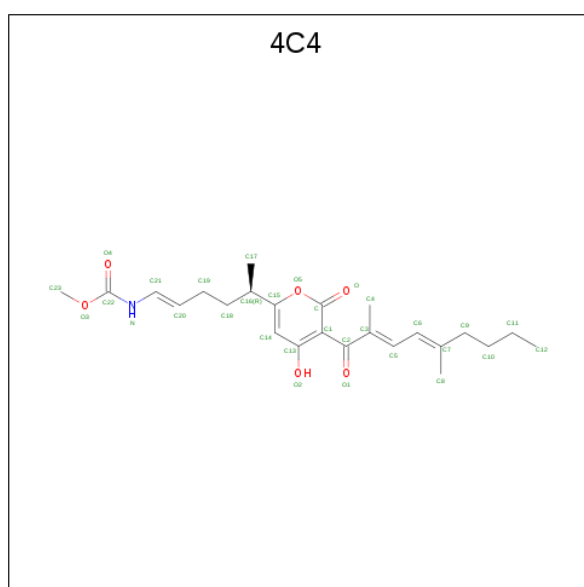
- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	J	1	Total	Mg	0	0
			1	1		
6	D	1	Total	Mg	0	0
			1	1		

- Molecule 7 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	J	2	Total	Zn	0	0
			2	2		
7	D	2	Total	Zn	0	0
			2	2		

- Molecule 8 is Myxopyronin B (three-letter code: 4C4) (formula: C<sub>24</sub>H<sub>33</sub>NO<sub>6</sub>).

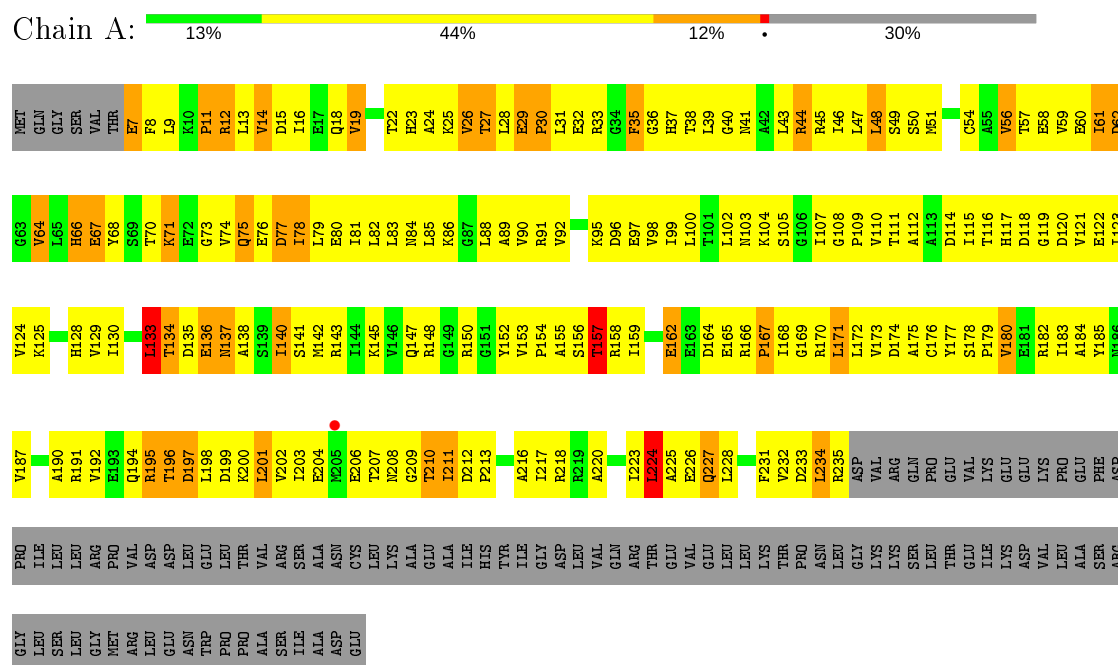


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
8	D	1	Total	C	H	N	O	0	0
			64	24	33	1	6		

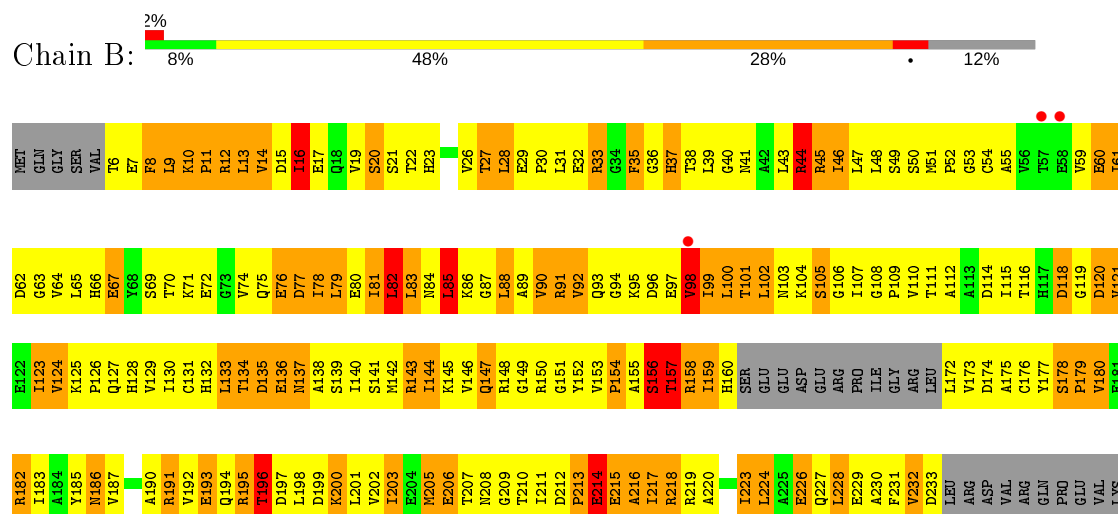
### 3 Residue-property plots

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: DNA-directed RNA polymerase subunit alpha



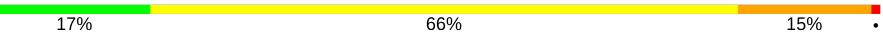
#### • Molecule 1: DNA-directed RNA polymerase subunit alpha



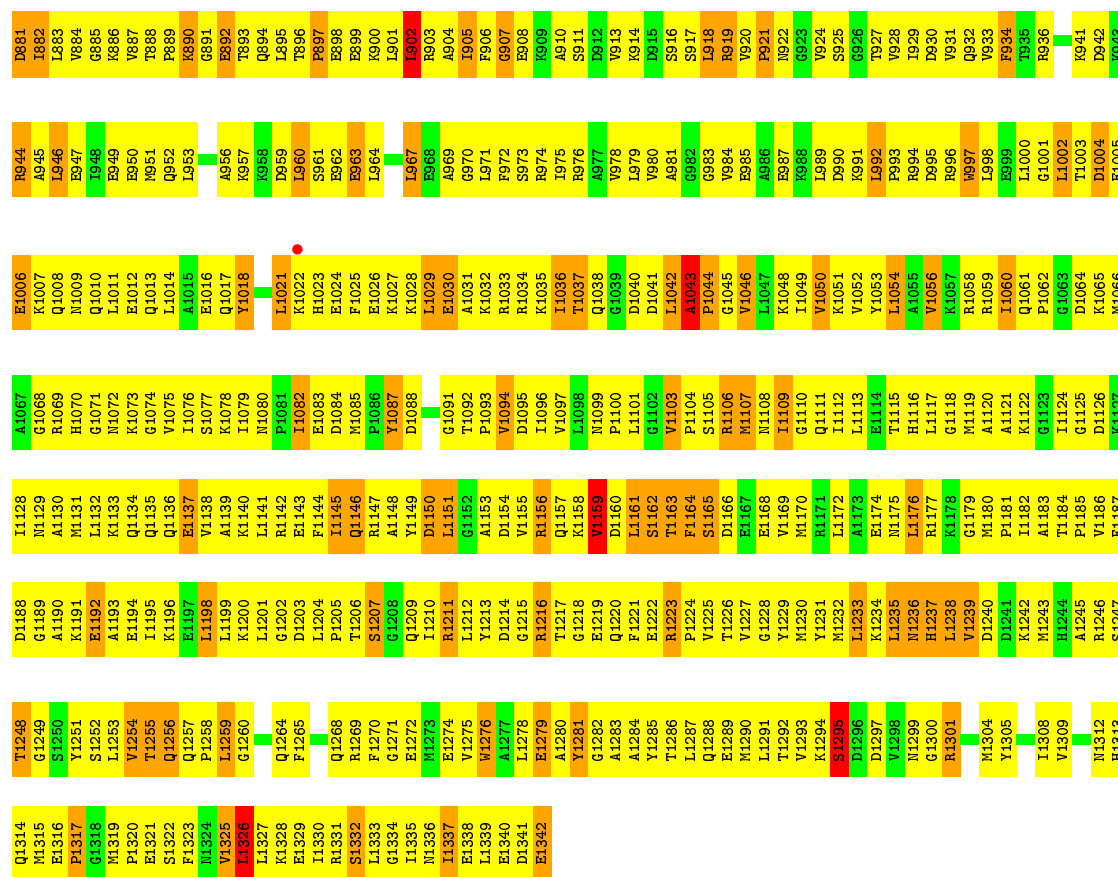




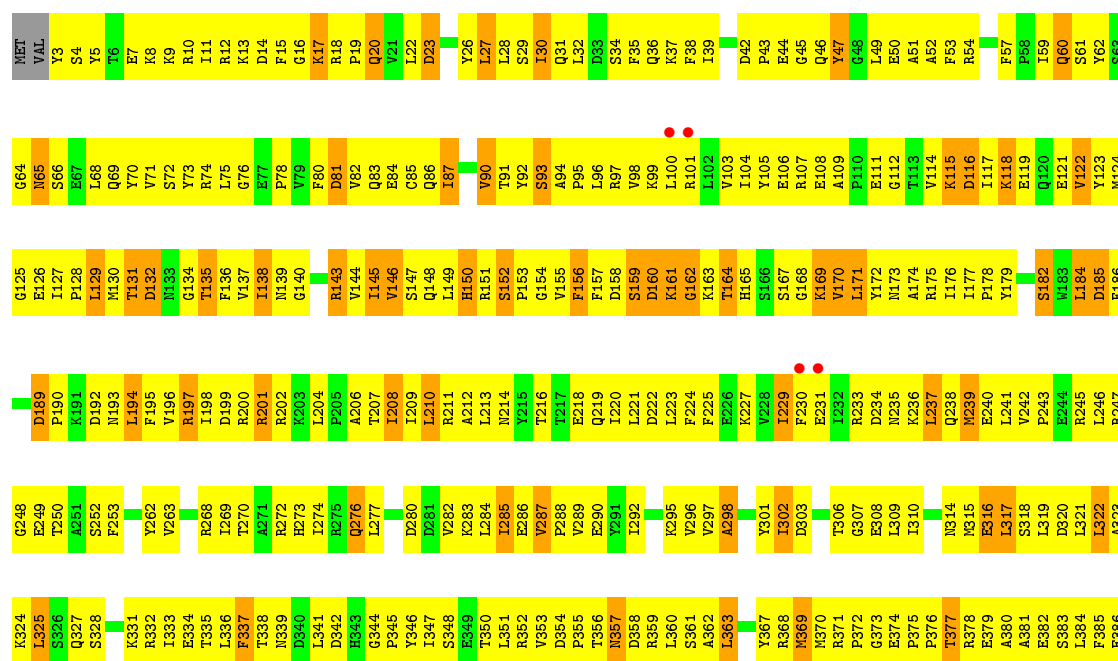
● Molecule 2: DNA-directed RNA polymerase subunit beta

Chain C:  17% 66% 15%

MET	VAL	S4	Y5	T6	E7	K8	K9	R10	F11	I12	R13	D14	F15	G16	K17	R18	P19	Q20	V21	D22	D23	Y26	L27	L28	S29	T30	Q31	L32	F35	Q36	K37	F38	T39	E40	Q41	D42	P43	E44	G45	Q46	Y47	G48	L49	E50	A51	F52	R54	V56	Q55	E57	P58	L59	Q60	G61	E62						
S63	G64	N65	S66	E67	L68	Q69	Y70	V71	S72	F73	R74	L75	F80	D81	V82	Q83	E84	G85	Q86	L87	R88	G89	Y90	T91	Y92	S93	A94	P95	L96	R97	N98	K99	L100	L102	V103	I104	Y105	E106	R107	G108	A109	G112	T113	L114	E115	K116	T117	R118	E119	S120	Q121	S122	V123	M124	G125	S126					
I127	P128	M129	M130	T131	D132	N133	G134	T135	V136	V137	I138	N139	G140	T141	R142	V143	A144	I145	V146	Q147	Q148	L149	R150	S152	V155	F156	F157	D158	S159	D160	K161	G162	L163	T164	H165	K166	S167	G168	K169	V170	L171	A174	R175	D176	N177	M178	P179	V180	E181	S182	Q183	W184	D185	F186	E187	F188					
D189	P190	K191	D192	N193	L194	F195	V196	R197	I198	D199	R200	R201	R202	K203	L204	P205	A206	T207	L208	I209	L210	R211	A212	L213	N214	Y215	T216	T217	E218	Q219	I220	D221	D222	L223	F224	F225	E226	K227	V228	I229	G230	D231	N232	R233	Q234	N235	L236	Q237	Q238	N239	E240	L241	V242	E243	E244	R245	L246	R247	G248	E249	T250
A251	S252	D253	D254	E255	E256	K260	V261	V262	V263	G266	R267	R268	L269	T270	T271	R272	H273	L274	Q275	Q276	L277	E278	K279	D280	D281	V282	L285	E286	V287	P288	V289	E290	Y291	L292	V296	V297	A298	K299	D300	Y301	I302	D303	E304	S305	T306	G307	E308	L309	I310	N311	A312	A313	N314	M315	E316						
L317	S318	L319	D320	L321	L322	A323	K324	L325	S326	Q327	S328	G329	H330	K331	R332	L333	E334	T335	L336	F337	T338	N339	D340	L341	D342	R343	G344	L345	Y346	L347	S348	E349	T350	L351	V352	V353	D354	F355	T356	N357	D358	R359	L360	S361	A362	L363	V364	E365	L366	Y367	R368	M369	N370	R371	P372	T377	R378	E379			
A380	A381	E382	S383	L384	F385	E386	R387	L388	F389	F390	D393	R394	Y395	D396	L397	E398	A399	V400	G401	R402	L403	L404	V405	N406	R407	S408	L409	L410	R411	E412	L413	S414	L415	L416	L417	S418	L419	L420	S421	K422	D423	D424	D427	V428	M429	K430	K431	L432	L433	D434	A500	A501	L502	K503	P504	E505	F506				
D443	G507	S508	S509	Q510	L511	S512	Q513	F514	M515	D516	S517	V518	E519	F520	L521	A522	E523	D524	T525	H526	K527	R528	P529	L530	S531	E532	L533	G534	P535	G536	G537	L538	T539	R540	E541	L542	M543	D544	F545	E546	V547	R548	D549	V550	H551	P552	T553	H554	A555	K556	R557	V558	C559	P560	L561	T562	K563	B564	E565	G566	
P567	R568	L569	G570	L571	L572	N573	S574	L575	S576	V577	T581	N582	E583	L584	F585	G586	L587	E588	T589	P590	Y591	R592	K593	V594	T595	D596	G597	V598	V599	T600	P601	E602	L603	H604	V605	L606	S607	A608	B610	B611	Y614	V615	L616	A617	D618	L619	N620	S621	N622	L623	D624	L625	B626	T627	H628	F629					
V630	E631	D632	L633	V634	T635	C636	R637	S638	V639	G640	E641	S642	S643	L644	F645	S646	D647	L648	Q649	V650	D651	V652	V653	D654	V655	S656	T657	Q658	Q659	V660	V661	S662	A665	S666	L667	L668	F669	F670	L671	E672	H673	D674	D675	L676	A677	N677	R678	A679	L680	H681	G682	A683	H684	V685	Q686	R687	Q688	A689	V690		
P691	T692	L693	R694	L695	D696	L697	L698	L699	G700	V701	V702	G703	H704	R705	T706	A707	V708	D711	S712	G713	V714	V715	A718	R719	R720	V721	G722	G723	V724	Q725	V726	T727	D728	A729	S730	R731	L732	K735	V736	V737	E738	D739	E740	N741	V742	P743	E744	E745	A746	G747	N748	V749	L750	V751	N752	L753					
V754	K755	V756	T757	R758	S759	N760	Q761	N762	T763	G764	V765	N766	Q767	V768	L769	P770	V771	S772	L773	G774	E775	V776	V777	E778	R779	D780	R781	V782	L783	P784	V785	S788	T789	D790	L791	G792	E793	L794	A795	L796	V797	Q798	N799	N800	R801	V802	A803	R804	N805	P806	G809	L810	V811	H812	N813	E814	S815	G816	L817		
L817	V818	S819	E820	R821	V822	V823	Q824	R827	F828	T829	V830	L831	H832	L833	Q834	E835	L836	L837	C838	V839	S840	R841	D842	T843	R844	L845	V846	R847	V848	E849	L850	T851	A852	D853	L854	P855	N856	V857	G858	E859	A860	L861	L862	S863	R864	L865	D866	E867	S868	G869	L870	V871	H872	L873	T878	G879	G880				



• Molecule 2: DNA-directed RNA polymerase subunit beta







● Molecule 3: DNA-directed RNA polymerase subunit beta'

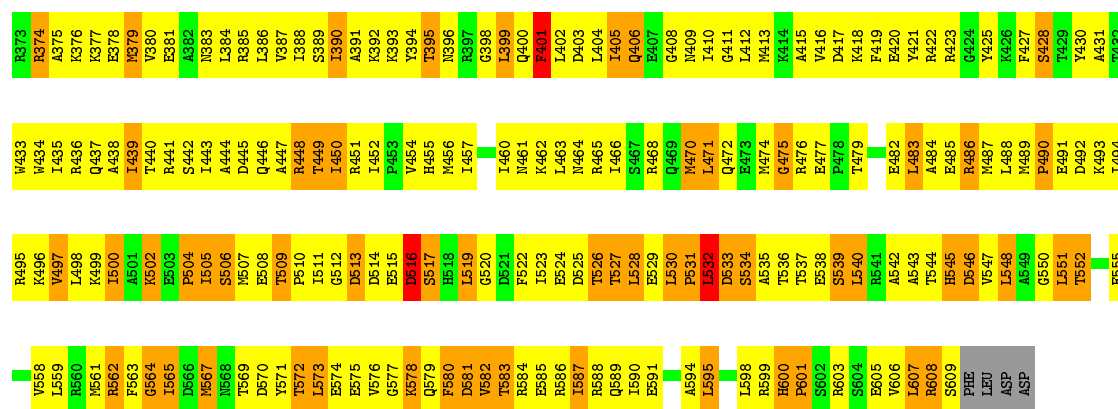
Chain D: 12% 52% 17% 17%

MET	LYS	ASP	LEU	LEU	LYS	PHE	L8	K9	A10	C70	Q11	Q11	T12	K12	T14	T14	E15	E15	F17	R77	D18	A19	I20	K21	I22	A23	A25	A25	S26	P27	D28	M29	V90	I30	R31	S32	W33	S34	F35	G36	E37	V38	R39	K40	P41	E42	T43	L44	L45	Y46	R47	T48	F49	K50	P51	E52	R53	D54	L55	G56	F57	C58	A59	R60																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		











## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	188.52Å 205.18Å 310.14Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.95 – 3.84 44.64 – 3.84	Depositor EDS
% Data completeness (in resolution range)	92.4 (29.95-3.84) 79.3 (44.64-3.84)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.08 (at 3.88Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.236 , 0.300 0.241 , 0.304	Depositor DCC
$R_{free}$ test set	1859 reflections (1.75%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	136.1	Xtriage
Anisotropy	0.197	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 117.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	55388	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	170.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.12% of the height of the origin peak. No significant pseudotranslation is detected.*

<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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, MG, 4C4

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	0.75	0/1801	1.11	5/2440 (0.2%)
1	B	0.52	0/2265	0.90	3/3066 (0.1%)
1	G	0.49	0/1777	0.88	0/2408
1	H	0.49	0/1681	0.93	4/2278 (0.2%)
2	C	0.79	4/10733 (0.0%)	1.13	39/14482 (0.3%)
2	I	0.63	0/10721	0.98	23/14468 (0.2%)
3	D	0.77	3/9202 (0.0%)	1.15	45/12424 (0.4%)
3	J	0.62	1/9161 (0.0%)	1.02	17/12366 (0.1%)
4	E	0.67	0/693	1.01	0/935
4	K	0.36	0/629	0.72	0/847
5	F	0.57	0/3777	0.93	6/5076 (0.1%)
5	L	0.48	0/3689	0.83	4/4969 (0.1%)
All	All	0.67	8/56129 (0.0%)	1.03	146/75759 (0.2%)

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
1	B	0	1
2	C	0	3
2	I	0	3
3	D	0	5
3	J	0	2
All	All	0	14

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	377	PHE	CE1-CZ	6.64	1.50	1.37

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1276	TRP	CB-CG	-6.57	1.38	1.50
2	C	1270	PHE	CE1-CZ	6.17	1.49	1.37
3	D	1319	PHE	CE2-CZ	6.01	1.48	1.37
3	J	686	TRP	CB-CG	-5.50	1.40	1.50

The worst 5 of 146 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	L	540	LEU	CA-CB-CG	-9.96	92.38	115.30
2	C	32	LEU	CA-CB-CG	-9.67	93.07	115.30
2	I	575	LEU	CA-CB-CG	-9.22	94.10	115.30
2	C	1161	LEU	CA-CB-CG	-9.16	94.23	115.30
3	D	788	LEU	CA-CB-CG	-9.10	94.38	115.30

There are no chirality outliers.

5 of 14 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	196	THR	Peptide
2	C	236	LYS	Peptide
2	C	600	THR	Peptide
2	C	658	GLN	Peptide
3	D	14	THR	Peptide

## 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	1779	0	1806	347	0
1	B	2239	0	2300	482	0
1	G	1755	0	1773	325	0
1	H	1662	0	1687	285	0
2	C	10564	0	10571	1842	1
2	I	10552	0	10548	1775	0
3	D	9062	0	9227	1784	1
3	J	9021	0	9213	1774	1
4	E	691	0	695	88	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	K	627	0	634	104	0
5	F	3726	0	3798	623	1
5	L	3640	0	3650	576	0
6	D	1	0	0	0	0
6	J	1	0	0	0	0
7	D	2	0	0	0	0
7	J	2	0	0	0	0
8	D	31	33	32	17	0
All	All	55355	33	55934	9319	2

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 84.

The worst 5 of 9319 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:B:88:LEU:HD12	1:B:89:ALA:H	1.07	1.18
2:I:280:ASP:HB3	2:I:282:VAL:HG23	1.24	1.18
3:D:850:LYS:HG2	3:D:857:LEU:HD11	1.24	1.18
2:C:798:GLN:HB2	2:C:828:PHE:HE1	1.04	1.17
3:D:1372:ARG:HH21	3:J:854:ALA:HB3	1.07	1.15

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:170:GLU:OE1	3:J:165:TYR:OH[3_444]	2.04	0.16
2:C:44:GLU:OE1	5:F:599:ARG:NE[4_445]	2.09	0.11

## 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	227/329 (69%)	179 (79%)	35 (15%)	13 (6%)	1	21
1	B	283/329 (86%)	167 (59%)	60 (21%)	56 (20%)	0	2
1	G	225/329 (68%)	177 (79%)	40 (18%)	8 (4%)	3	29
1	H	212/329 (64%)	170 (80%)	35 (16%)	7 (3%)	4	31
2	C	1338/1342 (100%)	1019 (76%)	256 (19%)	63 (5%)	2	24
2	I	1338/1342 (100%)	1031 (77%)	261 (20%)	46 (3%)	3	30
3	D	1162/1407 (83%)	850 (73%)	234 (20%)	78 (7%)	1	18
3	J	1151/1407 (82%)	851 (74%)	232 (20%)	68 (6%)	1	20
4	E	87/91 (96%)	69 (79%)	14 (16%)	4 (5%)	2	24
4	K	77/91 (85%)	61 (79%)	13 (17%)	3 (4%)	3	27
5	F	454/613 (74%)	342 (75%)	89 (20%)	23 (5%)	2	23
5	L	454/613 (74%)	336 (74%)	94 (21%)	24 (5%)	2	22
All	All	7008/8222 (85%)	5252 (75%)	1363 (19%)	393 (6%)	2	21

5 of 393 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	136	GLU
1	A	162	GLU
1	B	11	PRO
1	B	12	ARG
1	B	44	ARG

### 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	197/286 (69%)	164 (83%)	33 (17%)	2	14
1	B	250/286 (87%)	169 (68%)	81 (32%)	0	2
1	G	193/286 (68%)	171 (89%)	22 (11%)	5	27
1	H	183/286 (64%)	149 (81%)	34 (19%)	1	11

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	C	1154/1157 (100%)	982 (85%)	172 (15%)	3	18
2	I	1151/1157 (100%)	987 (86%)	164 (14%)	3	20
3	D	963/1168 (82%)	768 (80%)	195 (20%)	1	9
3	J	965/1168 (83%)	780 (81%)	185 (19%)	1	10
4	E	72/75 (96%)	65 (90%)	7 (10%)	8	33
4	K	67/75 (89%)	60 (90%)	7 (10%)	7	30
5	F	406/540 (75%)	346 (85%)	60 (15%)	3	18
5	L	385/540 (71%)	339 (88%)	46 (12%)	5	25
All	All	5986/7024 (85%)	4980 (83%)	1006 (17%)	2	14

5 of 1006 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	1194	ARG
1	G	157	THR
3	J	1242	ARG
3	D	1278	GLU
5	F	448	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 114 such sidechains are listed below:

Mol	Chain	Res	Type
5	F	317	ASN
2	I	357	ASN
3	J	1366	HIS
5	F	345	GLN
1	G	84	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 7 ligands modelled in this entry, 6 are monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
8	4C4	D	2004	-	27,31,31	1.35	2 (7%)	30,40,40	1.18	4 (13%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	4C4	D	2004	-	-	10/29/31/31	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	D	2004	4C4	C1-C2	-5.63	1.39	1.50
8	D	2004	4C4	O5-C15	2.12	1.38	1.35

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	D	2004	4C4	O5-C15-C14	2.67	122.29	119.70
8	D	2004	4C4	C13-C14-C15	-2.59	118.94	122.17
8	D	2004	4C4	O1-C2-C1	2.32	125.33	119.91
8	D	2004	4C4	C4-C3-C2	2.01	119.22	115.53

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

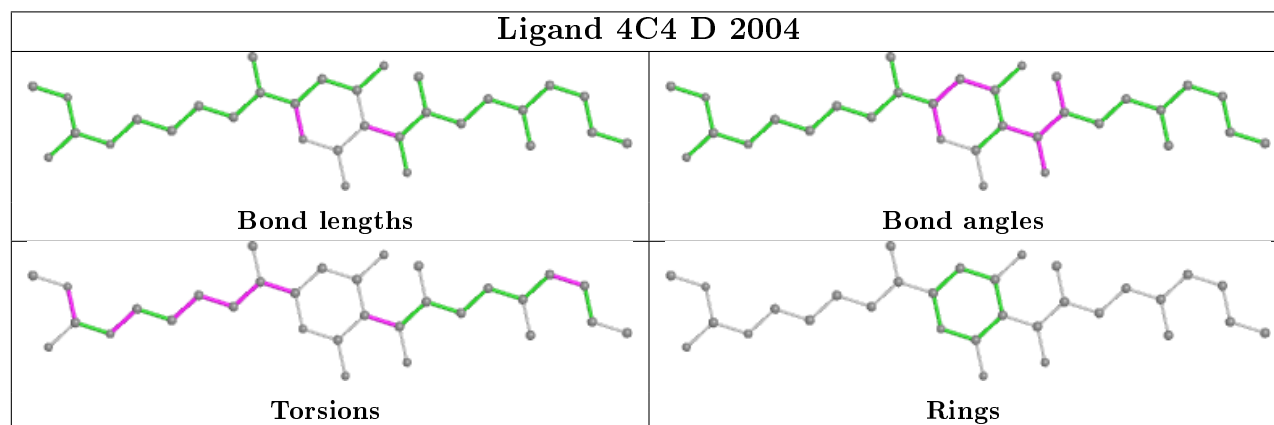
Mol	Chain	Res	Type	Atoms
8	D	2004	4C4	C14-C15-C16-C17
8	D	2004	4C4	C17-C16-C18-C19
8	D	2004	4C4	O4-C22-O3-C23
8	D	2004	4C4	N-C22-O3-C23
8	D	2004	4C4	C15-C16-C18-C19

There are no ring outliers.

1 monomer is involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	D	2004	4C4	17	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



## 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	229/329 (69%)	-0.33	1 (0%) 92 88	85, 130, 177, 246	0
1	B	289/329 (87%)	-0.20	8 (2%) 53 43	100, 194, 249, 268	0
1	G	227/329 (68%)	-0.19	4 (1%) 68 61	184, 217, 237, 248	0
1	H	216/329 (65%)	0.43	22 (10%) 6 6	179, 244, 274, 286	0
2	C	1340/1342 (99%)	-0.31	4 (0%) 94 91	75, 125, 228, 303	0
2	I	1340/1342 (99%)	-0.20	30 (2%) 62 53	124, 163, 252, 431	0
3	D	1166/1407 (82%)	-0.30	7 (0%) 89 85	84, 126, 211, 261	0
3	J	1155/1407 (82%)	-0.22	20 (1%) 70 62	122, 169, 233, 316	0
4	E	89/91 (97%)	-0.28	0 100 100	118, 140, 168, 177	0
4	K	79/91 (86%)	0.51	7 (8%) 9 7	223, 285, 342, 352	0
5	F	458/613 (74%)	-0.09	24 (5%) 27 23	115, 176, 321, 349	0
5	L	458/613 (74%)	-0.03	21 (4%) 32 27	153, 197, 350, 376	0
All	All	7046/8222 (85%)	-0.20	148 (2%) 63 55	75, 159, 278, 431	0

The worst 5 of 148 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	L	319	ALA	7.3
5	L	315	TRP	6.8
5	F	305	LEU	6.3
1	H	24	ALA	6.2
5	F	304	THR	6.1

### 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

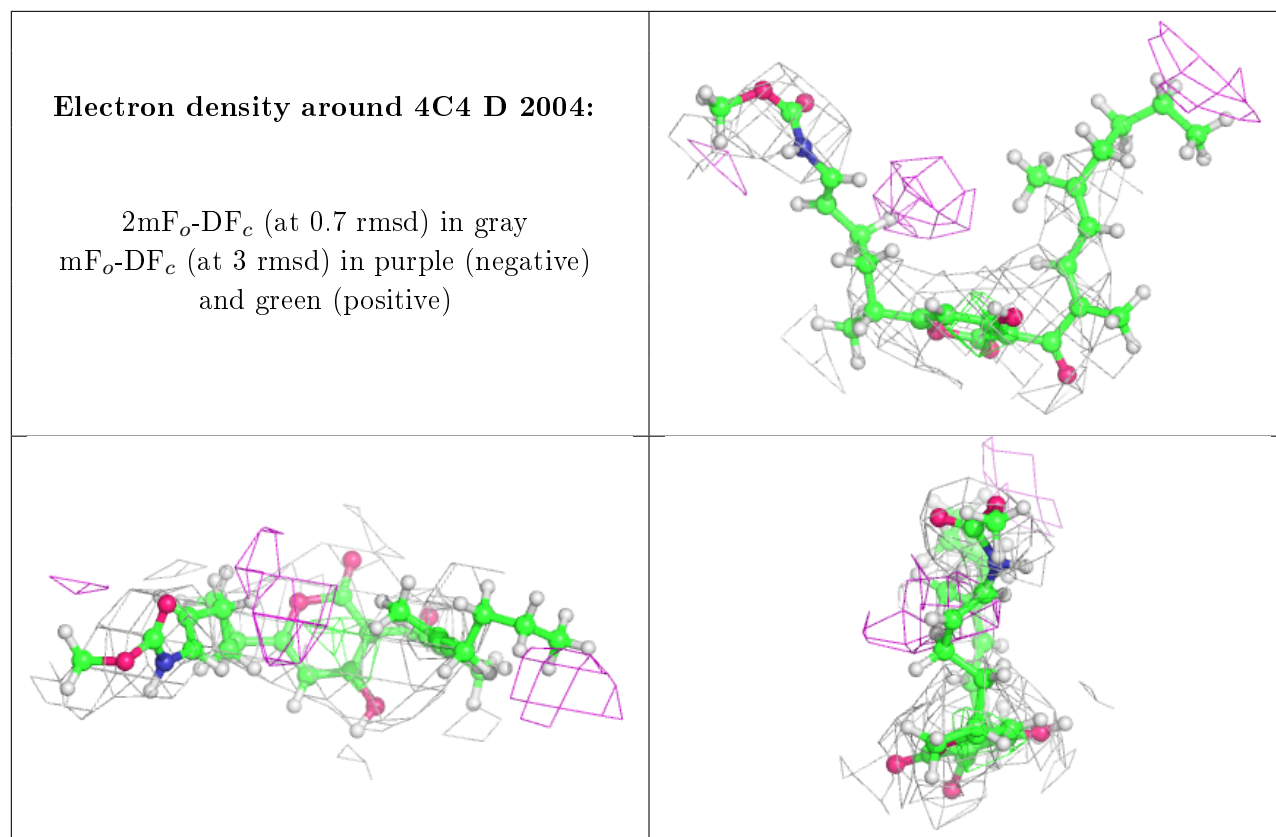
There are no carbohydrates in this entry.

### 6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
8	4C4	D	2004	31/31	0.89	0.43	128,154,155,155	0
7	ZN	J	2003	1/1	0.92	0.14	130,130,130,130	0
6	MG	J	2001	1/1	0.95	0.26	127,127,127,127	0
6	MG	D	2001	1/1	0.98	0.42	126,126,126,126	0
7	ZN	D	2003	1/1	0.99	0.23	125,125,125,125	0
7	ZN	D	2002	1/1	0.99	0.08	126,126,126,126	0
7	ZN	J	2002	1/1	0.99	0.04	143,143,143,143	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.



## 6.5 Other polymers [i](#)

There are no such residues in this entry.