

Table 8: Summary of Integrative Structure Determination of RNAPII Rpb1-Rpb2 dimer

<b>1) Gathering information</b>	
<i>Prior models</i>	rpb1: comparative model, template 6GMH:A rpb2: comparative model, template 4AYB:B None: sequence
<i>Physical principles and statistical preferences</i>	Excluded volume Sequence connectivity
<i>Experimental data</i>	123 pE-MAP derived distance restraints
<b>2) Representing the system</b>	
<i>Composition (number of copies)</i>	rpb1: 1 rpb2: 1
<i>Atomic (structured) components</i>	rpb1: 13-58, 82-105, 120-143, 173-184, 200-551, 565-589, 606-1077, 1098-1105, 1119-1171, 1191-1240, 1256-1286, 1304-1404 rpb2: 24-68, 88-139, 162-301, 314-565, 579-639, 655-661, 681-709, 740-871, 887-917, 934-1099
<i>Unstructured components</i>	rpb1: 1-12, 59-81, 106-119, 144-172, 185-199, 552-564, 590-605, 1078-1097, 1106-1112, 1113-1118, 1172-1190, 1241-1255, 1287-1303 rpb2: 1-23, 69-87, 140-161, 302-313, 566-578, 640-654, 662-680, 710-739, 872-886, 918-933
<i>Resolution of structured components</i>	1 [R1] residue per bead
<i>Resolution of unstructured components</i>	10 [R10] residues per bead
<i>Structural coverage</i>	85.07 %
<i>Rigid body (RB) definitions</i>	RB1: rpb1 <sub>1-1105</sub> RB2: rpb1 <sub>1113-1404</sub> RB3: rpb2 <sub>1-1099</sub>
<i>Spatial restraints encoded into scoring function</i>	Excluded volume; applied to the R1 representation Sequence connectivity; applied to the R1 representation pE-MAP MIC pair-restraints; applied to the R1 representation
<b>3) Structural Sampling</b>	
<i>Sampling method</i>	Replica Exchange Gibbs sampling, based on Metropolis Monte Carlo
<i>Replica exchange temperature range</i>	1.0 - 2.5
<i>Number of replicas</i>	8
<i>Number of runs</i>	60
<i>Number of structures generated</i>	3600000
<i>Movers for flexible string of bead</i>	Random translation up to 4.0 Å
<i>CPU time</i>	36 hours on 20 processors
<b>4) Validating the model</b>	
<b>Models selected for validation</b>	
<i>Number of models after equilibration</i>	3600000
<i>Number of models that satisfy the input information</i>	425380
<i>Number of structures in samples A/B</i>	210367/215013
<i>p-value of non-parametric Kolmogorov-Smirnov two-sample test</i>	0.012 (threshold p-value > 0.05)
<i>Kolmogorov-Smirnov two-sample test statistic, D</i>	0.23
<b>Thoroughness of the structural sampling</b>	
<i>Sampling precision</i>	14.84 Å
<i>Homogeneity of proportions <math>\chi^2</math> test (p-value)/Cramers V value</i>	0.000/0.099 (thresholds: p-value>0.05 OR Cramer's V<0.1)
<i>Number of clusters</i>	2
<i>Cluster populations</i>	cluster 1 : 81.7 % cluster 2 : 14.8 %
<i>Cluster precisions</i>	cluster 1 : 9.78 Å cluster 2 : 9.92 Å
<i>Average cross-correlation between localization probability densities of samples A and B</i>	cluster 1: 0.76 cluster 2: 0.84
<b>Validation by information used for modeling</b>	
<i>Percent of sequence connectivity restraints satisfied per structure</i>	99 %
<i>Percent pE-MAP restraints satisfied per structure</i>	95 %
<i>Percent of excluded volume restraints satisfied per structure</i>	99 %

<b>5) Benchmark</b>	
<i>Structural accuracy (95 % CI)</i> <i>PDB used for benchmark</i>	16.8 (12.5-23.2) Å 1I3Q
<b>6) Software and data availability</b>	
<i>Modeling programs</i>	IMP PMI module, version develop-39c22a602 Integrative Modeling Platform (IMP), version develop-39c22a602 MODELLER, version 9.21 <a href="https://integrativemodeling.org/systems/pemap">https://integrativemodeling.org/systems/pemap</a>
<i>Modeling scripts and data</i>	HHPred, version 2.0.16
<i>Homology detection and structure prediction</i>	UCSF Chimera, version 1.10
<i>Visualization and plotting</i>	Matplotlib, version 3.0.3