

Table 3: Summary of the NPC spoke orphan structural elements integrative threading and the Pom152 ring integrative modeling

1) Gathering information	
<i>Prior models</i>	Structure of the isolated and <i>in situ</i> NPCs
<i>Physical principles and statistical preferences</i>	Excluded volume Sequence connectivity Predicted secondary structure (RaptorXProperties) Predicted transmembrane domains (http://yeastgenome.org and HeliQuest)
<i>Experimental data</i>	1425 DSS and EDC chemical crosslinks Cryo-EM map; EMDB TBD FIB-milled CET map; EMDB TBD <i>in vitro</i> protein-protein interactions
2) Representing the system	
<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>	71 %
<i>Number of structural elements (SEs)</i>	30
<i>Composition (number of copies of Pom152)</i>	8
<i>Atomic (structured) components (Pom152)</i>	260-362, 379-472, 520-611, 616-714, 722-818, 824-918, 931-1026, 1036-1141, 1150-1229, 1244-1337
<i>Unstructured components (Pom152)</i>	1-259, 361-378, 471-519, 610-615, 713-721, 819-823, 919-930, 1027-1035, 1142-1149, 1230-1243
<i>Spatial restraints encoded into scoring function (Pom152)</i>	Excluded volume; applied to the R1 representation Sequence connectivity; applied to the R1 representation Cross-link restraints; applied to the R1 representation EM density restraint using Gaussian Mixture Model (GMMs) representations Transmembrane domain restraint (Pom152 (111-200)) Peri-nuclear restraint (Pom152 (201-1337)) Sequence connectivity and excluded volume
<i>Spatial restraints encoded into threading scoring function</i>	Secondary structure restraint; applied to the SEs Loop end-to-end distance restraint; applied to the SEs Cross-link restraints; applied to the SEs and R1 representation
3.1) Enumeration of threading of degrees of freedom	
<i>Sequences used for threading</i>	Nic96 (residues 1-205), Nup53 (residues 1-247), Nup59 (residues 1-265), Nup100 (residues 551-815), Nup116 (residues 751-965), and Nup145N (residues 201-458)
3.2) Structural Sampling	
<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>	50
<i>Number of structures generated</i>	300000
<i>Movers for flexible string of bead</i>	Random translation up to 4.0 Å
<i>CPU time</i>	14 hours on 42 processors
4.1) Validating the threading models	
<i>Protein, sequence range, and start residue standard deviation</i>	SE1: Nic96, 15-30 (1) SE2: Nic96, 38-43 (2) SE3: Nic96, 46-61 (0.3) SE4: Nic96, 66-78 (2) SE5: Nic96, 84-99 (1) SE6: Nic96, 118-141 and 147-166 (1.4) SE7: Nic96, 118-153 and 155-165 (0) SE8: undefined
4.2) Validating the Pom152 ring models	
Models selected for validation	
<i>Number of models after equilibration (isolated/in situ)</i>	300000/300000
<i>Number of models that satisfy the input information (isolated/in situ)</i>	92281/51183
<i>Number of structures in samples A,B (isolated/in situ)</i>	52831,39450/18722,32461

<i>p-value of non-parametric Kolmogorov-Smirnov two-sample test (isolated/in situ)</i>	0.446/0.34
<i>Kolmogorov-Smirnov two-sample test statistic (D) (isolated/in situ)</i>	0.0,1E-16
Thoroughness of the structural sampling	
<i>Sampling precision (isolated/in situ)</i>	14.66/12.6 Å
<i>Homogeneity of proportions χ^2 test p-value (Cramers V value) (isolated/in situ)</i>	1.000 (0.000)/ 1.000 (0.000) (thresholds: p-value>0.05 OR Cramer's V<0.1)
<i>Number of clusters (isolated/in situ)</i>	1/1
<i>Cluster populations (isolated/in situ)</i>	Cluster 1: 99/98 %
<i>Cluster precisions (isolated/in situ)</i>	Cluster 1: 15.42/10.1 Å
<i>Average cross-correlation between localization probability densities of samples A and B (isolated/in situ)</i>	Cluster 1: 0.84/0.92
Validation by information used for modeling	
<i>Percent of sequence connectivity restraints satisfied per structure (isolated/in situ)</i>	99/99 %
<i>Percent cross-link restraints satisfied by ensemble (isolated/in situ)</i>	89/89 %
<i>Percent of transmembrane domain restraints satisfied by ensemble (isolated/in situ)</i>	98/99 %
<i>Percent of excluded volume restraints satisfied per structure (isolated/in situ)</i>	99/99 %
6) Software and data availability	
Software	
<i>Modeling programs</i>	IMP PMI module, version develop-548de65454 Integrative Modeling Platform (IMP), version develop-548de65454
<i>Modeling scripts</i>	https://github.com/integrativemodeling/NPC_v3.0
<i>Homology detection and structure prediction</i>	HHPred, version 2.0.16
<i>Visualization and plotting</i>	UCSF Chimera Matplotlib, version 3.0.3
Data	
<i>PDB-dev accession code</i>	TBD