Supplementary Table S1

MFPT From δ to α

	Wild	No Gamma	No Beta	No Gamma, Beta
Dense Inference	1.077405	0.96467	1.74165	2.537465
Dense Naive	1.112255	1.077785	1.345095	1.905235
Sparse Inference	1.135825	NA	NA	NA

MFPT From α to δ

	Wild	No Gamma	No Beta	No Gamma, Beta
Dense Inference	0.553195	1.037265	0.854175	1.465245
Dense Naive	0.59218	0.855405	0.767095	1.22421
Sparse Inference	0.823185	NA	NA	NA

Supplemental Table 2. Fission yeast strains used in this study

Strain no.	***Strain genotype***	***Source***	***Related to***
KR342	*h^+ leu1-32 otr1R (SphI)::ura4+ ura4-D18 ade6-M210*	*Moazed lab*	*Figure 1B*
KR407	*h^+ leu1-32 otr1R (SphI)::ura4+ ura4-D18 ade6-M210 clr4Δ::kanMX6*	*Moazed lab*	*Figure 1B*
KR823	*h^+ leu1-32 otr1R (SphI)::ura4+ ura4-D18 ade6-M210 PAmCherry-swi6*	*This study*	*Figure 1B,*
KR778	*h^{90} leu1-32 ade6-M216 ura4-D18 swi6Δ: PAmCherry-swi6*	*This study*	*Figure 1E-H, 5B,D, S1A*
KR836	*h^{90} leu1-32 ade6-M216 ura4-D18 swi6Δ: PAmCherry-swi6 clr4Δ::kanMX6*	*This study*	*Figure 2A-B and S2B*
KR1425	*h^{90} leu1-32 ade6-M216 ura4-D18 swi6Δ: PAmCherry-swi6^{hinge}-hphMX6*	*This study*	*Figure 2C-D*
KR1665	*h^{90} leu1-32 ade6-M216 ura4-D18 swi6Δ: PAmCherry-swi6 clr4Δ:13-myc-clr4 F449Y-natMX6*	*This study*	*Figure 2E-F*
KR1428	*h^{90} leu1-32 ade6-M216 ura4-D18 swi6Δ: PAmCherry-swi6 mst2Δ::ura4+ epe1Δ::kanMX6 cut11-mCitrine-natMX6*	*This study*	*Figure 2G-H*
KR1096	*h^{90} leu1$^+$:mNeonGreen-swi6-kanMX6 ade6-M216 ura4-D18 swi6Δ:PAmCherry-swi6 CD^{mut}-hphMX6*	*This study*	*Figure 4B*
KR1097	*h^{90} leu1$^+$:mNeonGreen-swi6-kanMX6 ade6-M216 ura4-D18 swi6Δ: PAmCherry-swi6 CSD^{mut}-hphMX6*	*This study*	*Figure 4C*
KR1884	*h^{90} leu1$^+$:mNeonGreen-swi6-kanMX6 ade6-M216 ura4-D18 swi6Δ:PAmCherry-swi6^{hinge} CD^{mut}-hphMX6*	*This study*	*Figure 4D*
KR1711	*h^{90} leu1+:PAmCherry-2xCD GST-kanMX6 ade6-M216 ura4-D18 swi6Δ::natMX6*	*This study*	*Figure 5C-F*
KR1092	*h^{90} leu1-32 ade6-M216 ura4-D18 swi6Δ: PAmCherry-swi6 CD^{mut}-hphMX6*	*This study*	*Figure S2A*

KR858	*h^{90} leu1-32 ade6-M216 ura4D-18 swi6Δ: PAmCherry-swi6 mst2Δ::kanMX6 cut11-mCitrine-natMX6*	*This study*	*Figure S2C*
KR856	*h^{90} leu1-32 ade6-M216 ura4D-18 swi6Δ: PAmCherry-swi6 epe1Δ::kanMX6 cut11-mCitrine-natMX6*	*This study*	*Figure S2D*
KR1432	*h^{90} leu1-32 ade6-M216 ura4-D18 swi6Δ: PAmCherry-swi6 clr4Δ::hphMX6 mst2Δ::ura4+ epe1Δ::kanMX6 cut11-mCitrine-natMX6*	*This study*	*Figure S2E*
KR1882	*h^{90} leu1-32 ade6-M216 ura4-D18 swi6Δ: PAmCherry-1xCD GST-natMX6*	*This study*	*Figure S5A-C*
KR1878	*h^{90} leu1^{+}:PAmCherry-2xCD GST-kanMX6 ade6-M216 ura4-D18 swi6Δ::natMX6 clr4Δ::hphMX6*	*This study*	*Figure S5D*
KR1880	*h^{90} leu1^{+}:PAmCherry-2xCD^{ARK} GST-kanMX6 ade6-M216 ura4-D18 swi6Δ::natMX6*	*This study*	*Figure S5E*

Supplementary text- Fine-grained chemical rate constant inference

Our Monte Carlo chain used five move proposal types to sample the range of possible kinetic parameters:

(**i**) Single move: change a single rate constant by multiplying it by $e^{T*\sigma 1}$, where T is a t-distributed variable with 10 degrees of freedom and σ_1 is an arbitrary scaling factor (currently equal to 0.02) which we tune for optimal convergence.

(**ii**) Paired move: change a pair of rate constants which encode opposite reactions (e.g., the rate constant of a Swi6 molecule moving from the α state to the β state and the rate constant of a Swi6 molecule moving from the β state to the α state) by multiplying them both by $e^{T*\sigma 2}$ (σ_2 was optimized to 0.02 for the calculations given here to yield ideal transition rates; note that the value of σ_2 and similar parameters on move sizes affect only sampling efficiency, and not the equilibrium distribution of parameters). T is a t-distributed variable with 10 degrees of freedom.

(**iii**) Global rescale: multiply every rate constant by a factor $e^{T*\sigma 3}$. T is a t-distributed variable with 10 degrees of freedom and σ_3 was optimized to be 0.00125.

(**iv**) Local randomize: pick one rate constant and replace it from a draw from a Gamma distribution. This distribution was picked based on the experimental data. For a rate constant denoting transition from state X to state Y, where M is the experimental probability that a Swi6 in state X would stay in state X, the shape parameter for this distribution was $(-\ln(M)/0.04)^{1/2}$ and the scale parameter was $(-\ln(M)/0.04)^{3/2}$.

(**v**) Global randomize: replace all rate constants with random draws from Gamma distributions identical to the ones in move **iv**.

For the *clr4*Δ model, we discovered that the reaction rate constants corresponding to opposite reactions can be highly correlated: in particular, the rates to and from the δ state and the β state were correlated to the point of the inference being unable to resolve after a disproportionate amount of computation. We therefore decided to transform the rates using the following bijection: instead of conceiving of each pair of reactions as having independent rate constants, we consider the reactions as having a magnitude and an equilibrium constant. We considered the forward direction of reaction to be the one which increases the level of binding (e.g., from δ to α or γ to β) The magnitude parameter was set as the natural logarithm of the rate of the reverse reaction, and the equilibrium was set as the natural logarithm of the K_{eq}. We used the following Monte Carlo moves for this parameterization:

(**i**) Move one magnitude: change a single magnitude by adding T*σ_1, where T is a t-distributed variable with 10 degrees of freedom and σ_1 is an arbitrary scaling factor (currently equal to 0.2) which we tune for optimal convergence.

(**ii**) Move all magnitudes: add T*σ_3 to all magnitudes. T is once more a t-distributed variable with 10 degrees of freedom, and σ_3 was optimized to be 0.0222.

(**iii**) Move one equilibrium constant: change a single equilibrium by adding T*σ_1, where T is a t-distributed variable with 10 degrees of freedom and σ_1 is an arbitrary scaling factor (currently equal to 0.2) which we tune for optimal convergence.

(**iv**) Move all equilibrium constants: add T*σ_3 to all equilibria. T is a t-distributed variable with 10 degrees of freedom, and σ_3 was optimized to be 0.0222.

(**v**) Randomize one magnitude: pick one magnitude and replace it from a draw from a t distribution with σ = 1 and 10 degrees of freedom.

(**vi**) Randomize all magnitudes: replace all magnitudes from a draw from a t distribution with σ = 1 and 10 degrees of freedom.

(**vii**) Randomize one equilibrium: pick one equilibrium constant and replace it from a draw from a t distribution with σ = 1 and 10 degrees of freedom.

(**viii**) Randomize all magnitudes: replace all equilibrium constants from a draw from a t distribution with σ = 1 and 10 degrees of freedom.

To evaluate the prior likelihoods, we transformed the state of our model back to the rate constant formulation described previously and evaluated using the same priors as for the WT model. Thus, the equilibrium distribution of parameters will be unaffected by the reparameterization; only the convergence properties of the Monte Carlo sampling are changed.

For any of the model parameterizations described above, at each Monte Carlo move, we accept or reject the new set of constants based on the Metropolis-Hastings criterion: we divide the probability of the new rates co-occurring with the data by the probability of the old rates co-occurring with the data. To implement the Metropolis-Hastings criterion for symmetrical moves, we set this equal to the odds ratio. For asymmetrical moves, we must also multiply the probability we would have proposed the old rates and divide by the probability that we propose the new rates to derive our odds ratio. If our odds ratio is greater than 1, we accept the new rates. If our odds ratio is less than 1, we accept the new rates still with a probability equal to the

odds ratio. The sampling process was repeated with four independent until convergence of the posterior distribution, when the Ř metric for the parameters is less than 1.2 (Gelman and Rubin, 1992). We discarded the first part of the inference as a burn-in period, where the parameter set was influenced by initial conditions more than the posterior. We determined this burn-in based on plotting the likelihood evaluations: the likelihoods approached a final value roughly asymptotically, and so we removed the sharply increasing likelihoods as burn-in.

For the five-state model, we discovered that the remaining rate constants could not be easily directly deduced by Monte Carlo inference: they were very highly correlated. When we analyzed the transition rates mathematically, the issue was found to be that there were only four parameters which could vary independently, given the information we had: *p,* the equilibrium proportion of β-Swi6 in wild type which was bound to nonmethylated histone (a state we will call β_u-Swi6); q, the proportion of the rate of α-Swi6 transitioning to β-Swi6 which was the result of transitions to β_u-Swi6; *R*, the rate constant in units 1/sec, at which β_u-Swi6 transitions to β-Swi6 bound to methylated histone (which we will call β_s-Swi6); and *S*, the rate constant in units 1/sec, at which β_u-Swi6 transitions to α-Swi6. These four parameters plus the parameters given by both the wild type simulation and the *clr4Δ* simulation recapitulate all 20 of the rate constants in the five-state model by the following calculation, which preserves all of our observed transitions and equilibria from the four-state wild type inference:

1) The rates of transition involving the α, γ, and δ states exclusively are taken directly from the wild type simulation

2) The rates of transition between δ and β_u and γ and β_u were taken directly from the *clr4Δ* simulation

3) The rate of transition from β_u to α was set as *S*, and the rate of transition from β_u to β_s was set as *R.*

4) The rate of transition from α to β_u was set as *q**the rate of transition from α to β from the wild type inference, and the rate of transition from α to β_s was set as (1-*q*)*the rate of transition from α to β from the wild type inference

5) The rate of transition from β_s to α (γ, δ) was set as 1/(1-*p*) times difference of the observed rate of transition from β to α (γ, δ) and *p* times the rate of transition from β_u to α (γ, δ)

6) The rate of transition from β_s to β_u is set as the unique number which preserves the equilibrium proportions of the α, γ, and δ states from the wild type transitions, and which sets the equilibrium proportions of the β_u and β_s populations at *p* * the equilibrium proportion of β from the wild type transitions and (1-*p*)* the equilibrium proportion of β, respectively.

So long as any parameter set *p*, *q*, *S*, and *R* induces a set of rate constants which are all nonnegative, they are considered a valid parameter set. This constraint forces this set into a space which is bounded above and below for all four parameters. With the new information enforced by this constraint, we have no special *a priori* reason to prefer any set of four parameters over any others, so we choose a new joint prior which has uniform finite positive density *K* for all valid parameter sets, and zero density for invalid sets. We do not need to explicitly calculate the prior density *K*: it will be cancelled out for any comparison between any two valid parameter sets, and the program will automatically reject any parameter set forbidden by the prior.

Because of the smaller dimensionality of the model, instead of exploring the posterior with MCMC, we instead calculated the likelihood function explicitly across the grid of valid parameters. The parameter *p* was calculated from 0.50 to 0.555 in increments of 0.005, *q* from 0.35 to 0.79 in increments of 0.01, *S* from 0.10/sec to 3.05/sec in increments of 0.01/sec, and *R* from 0.15/sec to 0.45/sec in increments of 0.05/sec and from 1/sec to 12/sec in increments of 0.5/sec. These grid points mostly avoided parameters which resulted in invalid rates or whose likelihood was sufficiently small that we could safely ignore that part of the landscape. Some invalid parameter sets did occur near the boundaries of this grid, but were omitted. Because of the approximate nature of our likelihood computation, we used the R package “polspline” to smooth the likelihood landscape.