SUPPLEMENTARY INFORMATION

The kinetic landscape of an RNA binding protein in cells

Deepak Sharma, Leah L. Zagore, Matthew M. Brister, Xuan Ye, Carlos E. Crespo-Hernández, Donny D. Licatalosi & Eckhard Jankowsky

Supplementary Methods

Calculation of kinetic parameters
   Numerical Approach
   Analytical Approach
   Multiple Linear Regression Analysis

Supplementary Tables S1 – S7
Supplementary Figures S1 – S7
Supplementary Schemes S1, S2
SUPPLEMENTARY METHODS

Calculation of kinetic parameters

Numerical approach

The numerical approach to calculate kinetic parameters is based on numerically fitting crosslinking timecourses to the differential equations describing the Dazl-RNA binding and crosslinking process (Fig. 1a), according to:

\[
\frac{d(DR)}{dt} = k_{on}(D)(R) - k_{diss.}(DR) - k_{XL}(DR) \quad \text{(Eq. 17)}
\]

\[
\frac{d(DR^*)}{dt} = k_{XL}(DR) \quad \text{(Eq. 18)}
\]

(DR: concentration of non-crosslinked Dazl-RNA complex (for each binding site), DR*: concentration of crosslinked Dazl-RNA complex (for each binding site), D: Dazl concentration, R: RNA concentration (binding site), \(k_{on}\): association rate constant, \(k_{diss.}\): dissociation rate constant, \(k_{XL}\): crosslinking rate constant).

Because concentrations of free Dazl and RNA in the cell are experimentally inaccessible, the second order association process \((k_{on})\) was treated as pseudo-first order reaction at each of the two Dazl concentrations. Accordingly, we calculated a pseudo first order rate constant for each Dazl concentration \((k_{on}(1xDazl), k_{on}(4.2xDazl))\), and \(k_{diss.}, k_{XL}(1mW)\) and \(k_{XL}(2.6mW)\) for each binding site. Numerical fitting of timecourses of normalized read coverage for each binding site (Fig. 2c) was performed in R with packages deSolve (with ODE function) \(^65\), ggplot2 \(^66\), reshape2 \(^67\) and rmarkdown \(^68\).

The fitting strategy encompassed two steps: (i) estimation of parameter ranges following a sequential parameter estimation procedure \(^69\) and (ii) fitting the timecourses using estimated parameter ranges as input (Supplementary Material, Scheme 1). Estimation of parameter ranges was also performed in two steps, (i,a) initial parameter range estimation for \(k_{on}(1xDazl), k_{on}(4.2xDazl), k_{diss.}, k_{XL}(1mW)\) and \(k_{XL}(2.6mW)\), and (i,b) refinement of initial parameter range estimates to obtain final parameter range estimates (Supplementary Material, Scheme 1). To estimate
*initial* parameter ranges, timecourses from reactions with 4.2xDazl at high laser power (2.6 mW) and low laser power (1mW) were fit separately. Starting values were based on the kinetic parameters measured *in vitro* (Fig. 1; $k_{on}(1\times\text{Dazl}) = 0.0001 \text{ s}^{-1}$, $k_{on}(4.2\times\text{Dazl}) = 0.0001 \text{ s}^{-1}$, $k_{diss} = 1 \text{ s}^{-1}$, $k_{XL}(1\text{mW}) = 1 \text{ s}^{-1}$ and $k_{XL}(2.6\text{mW}) = 10 \text{ s}^{-1}$. Use of significantly different starting values did not yield acceptable fits for the majority of binding sites). This step provided average initial values for $k_{on}(4.2\times\text{Dazl})$ and $k_{diss}$ as well as initial values for $k_{XL}(1\text{mW})$ and $k_{XL}(2.6\text{mW})$. Next, timecourses at 1xDazl at high laser power (2.6mW) and low laser power (1mW) were fit separately, yielding average initial values for $k_{on}(1\times\text{Dazl})$ and $k_{diss}$ and initial values for $k_{XL}(1\text{mW})$ and $k_{XL}(2.6\text{mW})$. This process was performed for each binding site until the $X^2$ was minimized (no change in $X^2$ for 4 consecutive cycles) or 1,000 fitting cycles were completed. The process provided 10,341 x 5 parameter values, which were plotted as distribution (10,341 values for each parameter). The *initial* parameter range estimate represents the 95% confidence interval from the mean of the distribution for $k_{on}(1\times\text{Dazl})$, $k_{on}(4.2\times\text{Dazl})$, $k_{diss}$, $k_{XL}(1\text{mW})$ and $k_{XL}(2.6\text{mW})$.

To obtain *final* parameter range estimates, the initial parameter range estimates were used as input to fit multiple, random subsets of 2,000 randomly selected binding sites. 10,000 iterations, each with a unique random subset of 2,000 binding were performed. Each iteration yielded a distribution. All 10,000 distributions were superimposed, and the median apex of all distributions was identified. The *final* parameter range estimates represent the 95% confidence interval from the median apex of the averaged distributions. The *final* parameter range estimate was about 35% smaller than the *initial* parameter range estimate.

The estimated parameter ranges were used as input for fitting of the timecourses ([Supplementary Material, Scheme 1](#)). We fitted timecourses for reactions at 4.2xDazl at the different laser powers (1 mW, 2.6 mW), varying linked $k_{on}(4.2\times\text{Dazl})$ and $k_{diss}$ (which do not scale with laser power), and differing $k_{XL}(1\text{mW})$ and $k_{XL}(2.6\text{mW})$. We then fit timecourses at 1xDazl at both laser powers, varying linked $k_{on}(1\times\text{Dazl})$ and $k_{diss}$, and differing $k_{XL}(1\text{mW})$ and $k_{XL}(2.6\text{mW})$. Utilizing parameters obtained from these two steps, we fit all 4 timecourses linking $k_{on}(4.2\times\text{Dazl})$ and $k_{on}(1\times\text{Dazl})$ for differing laser powers, linking $k_{XL}(2.6\text{mW})$, $k_{XL}(1\text{mW})$ for differing Dazl concentrations and linking $k_{diss}$ for all conditions. The process of fitting all 4 timecourses for each binding site was repeated 642 times, after which $\chi^2$ did not show significant fluctuation (< 5% for 4 consecutive cycles). Obtained rate constants were used as final kinetic parameters for the numerical approach ([Extended Data Fig.3b-d](#)).
Fitting quality was assessed by calculating chi-squared ($\chi^2$) for each binding site, the overall cumulative reduced chi-squared ($\chi^2_v$) and the coefficient of determination/R² (COD) according to:

$$\chi^2 = \sum_i \frac{(O_i - C_i)^2}{\sigma_i^2} \quad \text{(Eq.19)}$$

($O$: observed value, $C$: calculated value for each binding site (i). $\sigma_i^2$ is the squared variance between data points $O$, $C$);

$$\chi^2_v = \frac{\chi^2_v}{v} \quad \text{(Eq.20)}$$

($v$: degree of free; equals $(n - m)$, with $n$: number of observations ($n = 16$), $m$: number of fitted parameters ($m = 5$)].

The coefficient of determination/R² (COD) was calculated using the standard method as described 70. The COD describes correlation between calculated and observed timecourses. For the last fitting cycle, COD = 0.92, $\chi^2_v = 0.043$ (Extended Data Fig.3c).

**Analytical approach**

The analytical approach to calculate kinetic parameters is based on fitting of crosslinking timecourses to explicit solutions of the system of differential equations (Eqs.17,18) for the kinetic scheme (Fig.1a). To solve the system of differential equations, we considered that at any given time ($t$) during crosslinking, the accessible fraction of a given Dazl binding site is either free ($R$), occupied ($DR$) or crosslinked ($DR^*$):

$$(R)_t + (DR)_t + (DR^*)_t = 1 \quad \text{(Eq.21)}$$

In addition, at $t \to \infty$, 100% of the accessible fraction of a given Dazl binding site is crosslinked. As described for the numerical approach, the second order association process ($k_{on}$) was treated as pseudo-first order process at each Dazl concentration.
Treating second order association process \( (k_{on}) \) as pseudo-first order process, considering Eq.21 and rearranging Eq.17 yields:

\[
\frac{d(DR)}{dt} = k_{on}[1 - [DR]_t - DR^*_t] - k_{diss}[DR]_t - k_{XL}[DR]_t
\]  
(Eq.22)

Before crosslinking (t = 0), at steady-state of the binding reaction,

\[
\frac{d(DR^*)}{dt} = 0
\]  
(Eq.23)

Because

\[
k_{XL} = 0
\]  
(Eq.24)

From Eq.17, we thus obtain:

\[
0 = k_{on}[R] - k_{diss}[DR]
\]  
(Eq.25)

which yields, after rearranging,

\[
[DR]_t = \frac{k_{on}}{k_{on}+k_{diss}}
\]  
(Eq.26)

At t = \( \infty \), crosslinking is complete, and thus

\[
\frac{d(DR)}{dt} = 0
\]  
(Eq.27)

\[
\frac{d(DR^*)}{dt} = 0
\]  
(Eq.28)
The boundary limits are:

\[ \lim_{t \to 0} DR^* \sim 0 \quad (\text{Eq.29}) \]

\[ \lim_{t \to \infty} DR^* \sim 0 \quad (\text{Eq.30}) \]

Equations 23-30 define the boundary conditions.

Crosslinking timecourses represent amount of crosslinked material at a given time \((t)\), expressed as normalized coverage value for each binding site \([DR^*](t)\). \([DR^*](t)\) depends on amount of Dazl-RNA complex \([DR]\) at the time \((t)\) (Eq.18) and thus on \(k_{on}, k_{diss.}\) and \(k_{XL}\). Absolute concentrations of \([D],[R]\) and \([DR]\) are not known in our system. To extract \(k_{on}, k_{diss.}\) and \(k_{XL}\) for each binding site from the crosslinking timecourses, we integrate Eq.18 after appropriate substitution of \([DR]\). To accomplish this, we take a second differential of Eq.22, considering the boundary conditions (Eq.23-30). We obtain the general solution of the second order differential equation:

\[ \frac{d^2(DR)}{dt^2} = k_{on} \frac{d(DR)}{dt} - k_{on} \frac{d(DR^*)}{dt} - k_{diss.} \frac{d(DR)}{dt} - k_{XL} \frac{d(DR)}{dt} \quad (\text{Eq.31}) \]

\[ \frac{d^2(DR)}{dt^2} = -(k_{on} + k_{diss.} + k_{XL}) \frac{d(DR)}{dt} + (k_{XL}k_{on})[DR](t) \quad (\text{Eq.32}) \]

Equation 32 is a constant coefficient, homogenous, linear, second order differential equation with two independent solutions \((y_1, y_2)\):

\[ y(t) = c_1 y_1(t) + c_2 y_2(t) \quad (\text{Eq.33}) \]

The coefficients \(c_1\) and \(c_2\) (by the principle of superposition) \(^{72}\) are obtained after providing the boundary conditions from equations 23-30. We identify a function \(y\) where a constant multiplied by its second derivative \(y''\) plus another constant times \(y'\) plus a third constant multiplied by \(y\) equals zero \(^{72}\).
The exponential function

\[ y = e^{rx} \text{ (} r \text{ : constant).} \quad \text{(Eq.34)} \]

has the property that its derivative is a constant multiple of itself:

\[ y' = re^{rx} \quad \text{(Eq.35)} \]

Furthermore,

\[ y'' = r^2e^{rx} \quad \text{(Eq.36)} \]

Substituting these expressions into (Eq.32), we obtain:

\[ ar^2 + br + c = 0 \quad \text{(Eq.37)} \]

Equation 37 is the auxiliary (characteristic) equation of the differential equation 32 (ref. 73). The equation is transformed into an algebraic equation by replacing

\[ \frac{d^2(\text{DR})}{dt^2} = r^2, \quad \text{(Eq.38)} \]

\[ \frac{d(\text{DR})}{dt} = r \quad \text{(Eq.39)} \]

and [DR] by 1.

The roots of Eq.37 are found by factoring 73:

\[ r_1 = \frac{(k_{on} + k_{diss} + k_{XL}) + \sqrt{(k_{on} + k_{diss} + k_{XL})^2 - 4(k_{XL}k_{on}[P])}}{2} \quad \text{(Eq.40)} \]
\[ r_2 = \frac{(k_{on} + k_{diss} + k_{XL}) - \sqrt{(k_{on} + k_{diss} + k_{XL})^2 - 4(k_{XL}k_{on}[P])}}{2} \quad \text{(Eq.41)} \]

With Eq.33-41, the general solution of Eq.32 is 74:

\[ [DR]_t = c_1 e^{r_1 t} + c_2 e^{r_2 t} \quad \text{(Eq.42)} \]

To obtain our observable \([DR^*]_0\), we integrate Eq.18 under consideration of the boundary conditions (Eqs.23-30):

\[ [DR^*]_t - [DR^*]_0 = k_{XL} \int_0^t [DR]_t \, dt \quad \text{(Eq.43)} \]

Substituting \([DR]_t\) from Eq.42 yields

\[ [DR^*]_t - [DR^*]_0 = k_{XL} [r_1 c_1 (1 - e^{r_1 t}) + r_2 c_2 (1 - e^{r_2 t})] \quad \text{(Eq.44)} \]

Substituting \(c_1\) and \(c_2\) by providing the boundary conditions (Eqs.23-30) and considering (Eqs.33-41), we obtain:

\[ [DR^*]_t = k_{XL} \left[ \frac{1}{k_{XL}} - r_1 (1 - \frac{k_{on}}{k_{on} + k_{diss}})(1 - e^{r_1 t}) + r_2 (1 - \frac{k_{on}}{k_{on} + k_{diss}})(1 - e^{r_2 t}) \right] \quad \text{(Eq.45)} \]

Equation 45 is an explicit nonlinear equation of the form:

\[ Y = f(t, \beta) + \epsilon \quad \text{(Eq.46)} \]

\(t = (t_1, t_2, \ldots \ldots \ldots \ldots t_n)\) are the independent variables (the normalized read coverage values at different timepoints), \(\beta = (\beta_1, \beta_2, \ldots \ldots \ldots \beta_n)^t\) are the parameters (\(k_{on} = k_{on}^{4.2x \#1}, k_{on}^{1x \#1}, k_{XL} = \ldots \ldots \ldots \))
$k_{XL}^{2.6mW \#i}, k_{XL}^{1mW \#i}$ and $k_{diss.} = k_{diss. \#i}$, where \#i represents the crosslinking conditions. $\epsilon$ is the fitting error between observed and expected timecourses. $f(t, \beta)$ represents the functional relationship between t, $\beta$ and Y.

Equation 45, adapted to the different Dazl concentrations and different laser powers was used to fit the crosslinking timecourses for each binding site. The resulting equations represent the non-linear model:

For 4.2xDazl, 2.6 mW laser:

$$\left[D_{XL}^* \right]'(t) = k_{XL}^{2.6mW \#1} \left( \frac{1}{k_{XL}^{2.6mW \#1}} - r_1 \left(1 - \frac{k_{XL}^{2.6mW \#1}}{k_{XL}^{2.6mW \#1} + k_{diss. \#1}} \right)(1 - e^{r_1 t})\right) + r_2 \left(1 - \frac{k_{XL}^{4.2x \#1}}{k_{XL}^{4.2x \#1} + k_{diss. \#1}} \right)(1 - e^{r_2 t})$$

(Eq.47)

For 4.2xDazl, 1 mW laser:

$$\left[D_{XL}^* \right]'(t) = k_{XL}^{1mW \#2} \left( \frac{1}{k_{XL}^{1mW \#2}} - r_1 \left(1 - \frac{k_{XL}^{1mW \#2}}{k_{XL}^{1mW \#2} + k_{diss. \#2}} \right)(1 - e^{r_1 t})\right) + r_2 \left(1 - \frac{k_{XL}^{4.2x \#2}}{k_{XL}^{4.2x \#2} + k_{diss. \#2}} \right)(1 - e^{r_2 t})$$

(Eq.48)

For 1xDazl, 2.6 mW laser:

$$\left[D_{XL}^* \right]'(t) = k_{XL}^{2.6mW \#3} \left( \frac{1}{k_{XL}^{2.6mW \#3}} - r_1 \left(1 - \frac{k_{XL}^{2.6mW \#3}}{k_{XL}^{2.6mW \#3} + k_{diss. \#3}} \right)(1 - e^{r_1 t})\right) + r_2 \left(1 - \frac{k_{XL}^{1x \#3}}{k_{XL}^{1x \#3} + k_{diss. \#3}} \right)(1 - e^{r_2 t})$$

(Eq.49)

For 1xDazl, 1 mW laser:

$$\left[D_{XL}^* \right]'(t) = k_{XL}^{1mW \#4} \left( \frac{1}{k_{XL}^{1mW \#4}} - r_1 \left(1 - \frac{k_{XL}^{1mW \#4}}{k_{XL}^{1mW \#4} + k_{diss. \#4}} \right)(1 - e^{r_1 t})\right) + r_2 \left(1 - \frac{k_{XL}^{1x \#4}}{k_{XL}^{1x \#4} + k_{diss. \#4}} \right)(1 - e^{r_2 t})$$

(Eq.50)
r1 and r2 are:

\[
\begin{align*}
    r_1 &= \left( k_{on}^h + k_{diss}^L \right) + \frac{\left( k_{on}^h + k_{diss}^L \right)^2 - 4 \left( k_{on}^h k_{diss}^L \right)}{2} \\
    r_2 &= \left( k_{on}^h + k_{diss}^L \right) - \frac{\left( k_{on}^h + k_{diss}^L \right)^2 - 4 \left( k_{on}^h k_{diss}^L \right)}{2}
\end{align*}
\]

(Eq.51) (Eq.52)

h represents 4.2xDazl #1 (Eq.47), 4.2xDazl #2 (Eq.48), 1xDazl #3 (Eq.49) and 1xDazl #4 (Eq.50). i represents #1 (Eq.47), #2 (Eq.48), #3 (Eq.49) and #4 (Eq.50). j represents 2.6 mW #1 (Eq.47), 1 mW #2 (Eq.48), 2.6 mW #3 (Eq.49) and 1 mW #4 (Eq.50).

Timecourses for 4.2xDazl at high laser (2.6 mW), 4.2xDazl at low laser (1mW), 1xDazl at high laser power (2.6 mW) and 1xDazl at low laser power (1mW) were separately fit to the non-linear model (Supplementary Material Scheme 2). A matrix of initial parameters was obtained,

\[
\begin{align*}
    4.2x Dazl: 2.6 \text{ mW laser} & \quad k_{on}^{(4.2x Dazl)} #1 \\
    4.2x Dazl: 1 \text{ mW laser} & \quad k_{on}^{(4.2x Dazl)} #2 \\
    1x Dazl: 2.6 \text{ mW laser} & \quad k_{on}^{(1x Dazl)} #3 \\
    1x Dazl: 1 \text{ mW laser} & \quad k_{on}^{(1x Dazl)} #4 \\
\end{align*}
\]

Next, a global datafit for all four timecourses (#1-4) for an individual binding site was performed. Initial parameters were iteratively adjusted, considering the following criteria:

\[
\begin{align*}
    k_{on}^{(4.2x Dazl)} #1 & \cong k_{on}^{(4.2x Dazl)} #2 (\text{at different laser powers}) \\
    k_{on}^{(1x Dazl)} #3 & \cong k_{on}^{(1x Dazl)} #4 (\text{at different laser powers}) \\
    k_{diss} #1 & \cong k_{diss} #2 \cong k_{diss} #3 \cong k_{diss} #4 \\
    k_{XL}^{(2.6mW)} #1 & \cong k_{XL}^{(2.6mW)} #3 (\text{at 2.6 mW laser power}) \\
    k_{XL}^{(1mW)} #2 & \cong k_{XL}^{(1mW)} #4 (\text{at 2.6 mW laser power})
\end{align*}
\]
Fits were repeated until the best fit was reached (no change in $\chi^2$ for 4 successive fittings), as measured by Chi-squared $\chi^2$ minimization, according to:

$$
\chi^2 = \sum_{i=1}^{n} \left[ \frac{Y_i - f(x'_i, \beta)}{\sigma_i} \right]^2
$$

(Eq.53)

$x'_i$ is the row vector for the $ith$ ($i = 1, 2, ..., n; n = 10,341$) observation. $\beta$ is the parameter under consideration. $Y_i$ is the estimated parameter value for the $ith$ ($i = 1, 2, ..., n; n = 10,341$) observation. $\sigma_i$ is the variance between observed and estimated parameter values. $f(x'_i, \beta)$ represents the function for which $x'_i$ and $\beta$ are measured.

Obtained parameters were further refined by additional rounds of fitting using the analytical, Levenberg-Marquardt (L-M) least squares algorithm, which combines the Gauss-Newton and the steepest descent method 75. Utilizing the values obtained above, parameters for timecourses at 4.2xDazl at high (2.6 mW) and low laser power (1mW) were adjusted together. $k_{on}^{(4.2xDazl)} #2$ was increased or decreased (depending on initial values for a given binding site) in small increments ($\partial b$) in order to move $k_{on}^{(4.2xDazl)} #2$ closer to $k_{on}^{(4.2xDazl)} #1$. $\partial b$ was set as 5% of $k_{on}^{(4.2xDazl)} #2$ for a given binding site. Following each increment, the timecourse was fitted to the non-linear model and $\chi^2$ calculated. $k_{diss.} #2$ and $k_{XL} (1mW) #2$ were floated during the fitting. If $\chi^2 (b + \partial b) \geq \chi^2 (b)$ for >3 consecutive fitting cycles, $k_{on}^{(4.2xDazl)} #1$ was increased or decreased (depending on initial values) in small increments to improve fitting. This fitting procedure was repeated for $N = 642$ cycles.

Next, the parameters for timecourses at 1xDazl at high (2.6 mW) and low laser power (1mW) were adjusted, providing $k_{on}^{(4.2xDazl)} #1$, $k_{on}^{(4.2xDazl)} #2$, $k_{on}^{(1xDazl)} #3$ and $k_{on}^{(1xDazl)} #4$. Keeping the adjusted $k_{on}$ constant (floating $k_{XL}$), $k_{diss.} #1$, $k_{diss.} #2$, $k_{diss.} #3$ and $k_{diss.} #4$ were subsequently adjusted (within 25% range of each other). Finally, $k_{XL} (2.6mW) #1$ and $k_{XL} (2.6mW) #3$ were adjusted by increasing or decreasing $k_{on}^{(4.2xDazl)} #1$ and $k_{on}^{(4.2xDazl)} #3$ in small increments ($\partial b \leq 5\%$ of parameter values) while maintaining $k_{on}^{(4.2xDazl)} #1 > k_{on}^{(4.2xDazl)} #3$. Additionally, $k_{diss.} #1$ and $k_{diss.} #3$ were increased or decreased in increments of $\partial b \leq 1\%$. The same process was performed for adjusting $k_{on}^{(4.2xDazl)} #2$ and $k_{on}^{(4.2xDazl)} #4$. Every parameter adjustment cycle was repeated 642 times after which $\chi^2$ values computed in 4 successive iterations showed fluctuations of less than 5% for > 95% of binding sites.
Multiple Linear Regression Analysis

Multiple linear regression (MLR) analysis was performed with “dummy coding”, e.g. transformation of categorical independent variables into dichotomous variables 76. The dependent variables, ΔRPF and ΔRNA, were used as continuous data, either separately or merged (Extended Data Fig.10). 45 models were formulated describing Dazl binding and corresponding mRNA characteristics for various combinations of “dummy coded” independent variables, “continuous” independent variables, “continuous” dependent variables (separate ΔRPF and ΔRNA) and “merged” dependent variables (Extended Data Fig.10). Models were progressively shortlisted and the best performing model (M1) was selected after 4 steps.

Step 1

We utilized the best subsets regression procedure to identify all possible model permutations of parameters (N = 45) that satisfied the following criteria:

1. Models contain n ≥ 3 independent variables
2. Models account for Dazl kinetics and binding pattern along with RNA features.
3. Selected independent variables do not show multi collinearity (assessed by pairwise correlation).

The data was randomly divided into training (70%, N = 699) and test set (30%, N = 492). The training set was utilized to evaluate, estimate and identify the optimal models and cross-validation was performed using the test set. Each model was regressed on associated independent variables and adjusted $R^2$ and root mean standard errors (RMSE) were calculated according to:

$$Adjusted \ R^2 = 1 - \left( \frac{n-1}{n-(k+1)} \right)(1 - R^2) \quad (Eq.54)$$

(n = 699, number of observations; k=7: number of independent variable terms). The root mean standard error (i.e. estimated standard deviation; $\sigma^2$ of the error term u) was obtained as:
\[ RMSE = \sqrt{\frac{SSE}{n-(k+1)}} \]  
(Eq.55)

(n = 699, number of observations; k = 7: number of independent variable terms; SSE: sum of squares error, difference between observed and predicted value). As expected, the adjusted \( R^2 \) showed inverse correlation with RMSE.

We selected the models with the highest adjusted \( R^2 \) (≥ 0.5) and lowest root mean standard errors (RMSE; top 50%). We also examined models with \( R^2 \) ≥ 0.5 despite low adjusted \( R^2 \), high RMSE according to:

\[ R^2 = \frac{SSR}{SSTO} = 1 - \frac{SSE}{SSTO} \]  
(Eq.56)

SSR (sum of squares due to regression; the sum of the differences between the predicted value and the mean of the dependent variable, measures unexplained variance) is equivalent to the distance from each point to the regression line. SSR was calculated according to:

\[ SSR = \sum_i (y_i - y')^2 \]  
(Eq.57)

(\( y_i \) = predicted value; \( y' \) = mean)

SSTO (sample variance) was calculated according to:

\[ SSTO = \sum_i (x_i - y')^2 \]  
(Eq.58)

(\( y_i \) = observed value; \( y' \) = mean)
With this approach, we shortlisted 24 models according to adjusted $R^2$, RMSE and $R^2$ values (Supplementary Material Fig.S4). We next determined information criterion statistics (ICS) for these models. ICS combines the SSE, number of parameters in the model, and sample size. We utilized three established information criterion parameters: Akaike’s Information Criterion (AIC), the Bayesian Information Criterion (BIC) and Amemiya’s Prediction Criterion (APC), which were calculated according to:

$$AIC_k = n \ln(SSE) - n \ln(n) + 2(k + 1) \quad (Eq.59)$$

$$BIC_k = n \ln(SSE) - n \ln(n) + (k + 1) \ln(n) \quad (Eq.60)$$

$$APC_k = \frac{(n+k+1)}{n(n-k-1)} \frac{SSE}{n(n-k-1)} \quad (Eq.61)$$

($n$: sample size, $k$: number of predictor terms, e. g. $k+1$ = number of regression parameters in the model, including the intercept). We compared all 24 models and ranked the models according to values for AIC, BIC and APC (lowest value – highest rank). At this stage, no model was removed.

**Step 2**

Further shortlisting was performed by comparing information criteria with model fitness parameters. To determine the fitness of the shortlisted models, two different hypothesis tests for slopes were conducted. We first tested the hypothesis that at least one slope parameter is 0:

$$H_0: \beta_1 = \beta_2 = \beta_{(n-)} = 0$$

$$H_\alpha: At \ least \ one \ \beta_i \neq 0 \ (for \ i = 1, 2, n, ...) \ where \ \alpha = 0.05 \quad (Eq.62)$$
using the general linear F test (ANOVA F statistic) by obtaining error sum of squares (the squared distances between the observed and predicted responses) for full (with all independent variables) and reduced models (with intercept only). p values were computed.

We next tested the hypothesis that only one of the slope parameters is 0:

\[ H_0: \beta_1 = 0 \]
\[ H_\alpha: \beta_1 \neq 0 \text{ where } \alpha = 0.05 \] (Eq.63)

using t-test statistics for each independent variable in the model. p values were computed.

Next, we compared information criterion parameters (AIC, BIC and APC), general linear F statistic and t-test statistic values for all 24 models. We shortlisted the models with the lowest AIC, BIC and APC values, most significant general linear F statistic and significant t-test statistic for all associated independent variables (Supplementary Material Fig.S5). All models satisfied the general linear F statistic condition, indicating that addition of selected independent variables (i.e. features) increased the explanatory power of the models. 13 out of 24 models had significantly lower information criterion parameters (Supplementary Material Fig.S5). We further assessed these 13 models according to obtained coefficients, standard errors, t-statistic, p-value and confidence intervals for all the independent variables. 6 out of 13 models showed significant t-statistics (p-values) for all coefficient terms and the smallest confidence interval ranges (Supplementary Material Fig.S5).

**Step 3**

To estimate the quality of the remaining 6 models, we tested 4 multiple linear regression conditions (LINE conditions):

1. The mean of the response, \( E(Y_i) \), at each set of values of predictors, \((x_{i1}, x_{i2}, x_{in})\) is a Linear function of the predictors.
2. The errors, \( \varepsilon_i \), are Independent.
3. The errors, \( \varepsilon_i \), at each set of values of the predictors are Normally distributed.
4. The errors, \( \varepsilon_i \), at each set of values of predictors have Equal variance (\( \sigma^2 \)).
To visually validate the LINE conditions (assessment of the distribution of errors), we recorded residuals vs. predicted values, and plotted a histogram of residuals for each model (Supplementary Material Fig.S6). We also performed the Kolmogorov-Smirnov Test (K-S test) for all 6 models. Three models, M1, M19 and M24 showed normal distribution of error residuals, absence of outliers and equal variance and hence were selected for cross-validation (Supplementary Material Fig.S6).

Step 4

These three models were validated using the test dataset (N = 492) and model M1 was identified as the optimal model on the basis of smallest Mean Squared Prediction Error value (MSPE) (Extended Data Figure 10e, Supplementary Material Fig.S7). This model (M1) consisted of seven independent variables: number of clusters in 3'UTR, ΣB, ΔΣB, number of binding sites in a cluster, UTR length, proximity to PAS, transcript level all expressed as dummy coded variables in terciles of their respective distributions. Multiple regression on a training data set of N = 699 was performed according to:

\[ Y'_i = b_0 + b_1X_{1i} + b_2X_{2i} + b_3X_{3i} + b_4X_{4i} + b_5X_{5i} + b_6X_{6i} + b_7X_{7i} + u \] (Eq.64)

(Y': predicted dependent, continuous variable (ΔRPF and ΔRNA) or predicted dependent, merged continuous variable, \( b_{(i=0...7)} \): differential intercept linear coefficients, \( X_{(n)} \): independent variables, \( u \): error term). The differential intercept linear coefficients (DILC) associated with each dummy coded/continuous independent variable terms are the expected difference in the mean of the outcome for that variable, compared to the reference group (TMRM class), with all other predictors constant. The "\( b_n \)" values represent regression weights that were computed by minimization of the sum of squared deviations:

\[ \sum_{i=1}^{n}(Y_i - Y'_i)^2 \] (Eq.65)
(n = 699, sample size of training data set, Y_i: observed value for the dependent variable ΔRPF and ΔRNA). The optimal regression model was:

\[ \Delta RPF = 1.01 + (cluster)^{+0.02H_i -0.03L_o} + (\text{bind. prob.})^{+0.05H_i -0.02L_o} + (\Delta \text{bind. prob.})^{+0.11H_i +0.05L_o} + \\
(\# \text{bind. sites})^{+0.03H_i -0.15L_o} + (\text{dist. PAS})^{-0.005H_i +0.01L_o} + (UTR \text{ len})^{+0.06} + (\text{RPKM})(-0.00004) + 0.07 \\
\]

\[ \Delta RNA = 1.01 + (cluster)^{+0.03H_i -0.003L_o} + (\text{bind. prob.})^{+0.05H_i -0.02L_o} + (\Delta \text{bind. prob.})^{+0.03H_i +0.01L_o} + \\
(\# \text{bind. sites})^{+0.04H_i -0.11L_o} + (\text{dist. PAS})^{+0.04H_i +0.007L_o} + (UTR \text{ len})^{+0.07} + (\text{RPKM})(-0.000007) + 0.06 \\
\]

(Eq.66)

\[
(\Delta RPF = \frac{\text{RPF at high Dazl}}{\text{RPF at low Dazl}}; \Delta RNA = \frac{\text{RNA at high Dazl}}{\text{RNA at low Dazl}}).
\]

The model was evaluated on a test data set (N = 492, 30% of the data; Fig.4). Regression analysis was performed using Scikit-learn and Statsmodels modules in Python 3.6.5.

References, Supplementary Methods

65 Soetaert, K. deSolve: Solvers for Initial Value Problems of Differential Equations <https://cran.r-project.org/web/packages/deSolve/index.html>

66 <https://www.r-graph-gallery.com/79-levelplot-with-ggplot2.html>

67 <https://www.rdocumentation.org/packages/reshape2/versions/1.4.3>

68 <https://rmarkdown.rstudio.com/>

69 Schön, T. B., Wills, A. & Ninness, B. System identification of nonlinear state-space models. Automatica 47, 39-49 (2011).

70 Ross, S. M. Introductory Statistics. (Elsevier Inc., 2010).

71 Boyce, W. E. & DiPrima, R. C. Elementary differential equations and boundary value problems. (Wiley, 2012).

72 Polyanin, A. D. & Zaitsev, V. F. Handbook of Ordinary Differential Equations: Exact Solutions, Methods, and Problems. (CRC Press, 2017).

73 Edwards, C. H. & Penney, D. E. Differential Equations: Computing and Modeling., 156–170 (Pearson Education).
Kreyszig, E. *Advanced Engineering Mathematics* (Wiley, 1972).

Gill, P. E. & Murray, W. Algorithms for the solution of the nonlinear least-squares problem. *SIAM Journal on Numerical Analysis* **15**, 977–979 (1978).

Draper, N. R. & Smith, H. *Applied Regression Analysis*. (Wiley & Sons, Inc., 1998).

Zhang Z. (2016). Variable selection with stepwise and best subset approaches. *Ann Transl Med* **4**(7), 136

Mishra P, Pandey CM, Singh U, Gupta A, Sahu C, Keshri A. Descriptive statistics and normality tests for statistical data. *Ann Card Anaesth* **22**, 67-72 (2019)

<https://scikit-learn.org/stable/>

<https://www.statsmodels.org/stable/index.html>
Supplementary Table S1 | Codon optimized *Mus musculus* Dazl (RRM) DNA construct (amino acids 32-117) and primers for cloning.

---

**Dazl (RRM) DNA construct**

SacI and XhoI restriction sites are underlined. Complete DNA construct was purchased from Genscript.

```
GGAAATATAGAGCTCTTGCCGGAAGGCAAGATCATGCATGATTCTGTAGGAGGAATAG
ACGTACGCATGGACGAAACCAGAAATCCGCTCTTTTTTCGACGCTACGCTCTCTGTAAAGGAGGT
TTAAAAATACCCGAGACAGAAGGGGTTTTCGAAAGGCTACGGATTCGTCTCTTTCTACACGAT
GTTGACGTTCAAGAAAAATAGTAGAGTCCTCAGATAAACTTTTCATGGGAAAGAAAACTGAAGCTGGCC
CGGCTATCCGCAAACAAATAATGACCTCGAGGGCTGCAA
```

**Primers for cloning**

SacI and XhoI restriction sites are underlined.

**Dazl Forward**

5′-GGAAATATAGAGCTCTTGCCGGAAGGCAAGATCATGC

**Dazl Reverse**

5′-TTGCAGCCCTCGAGGTCTATTTATTTGGGCGATA
**Supplementary Table S2 | Sequencing adapters and primers.**

| RNA linkers (Dharmacon) |
|-------------------------|
| RL5: 5’-OH AGG GAG GAC GAU GCG G 3’-OH |
| RL5D: 5’-OH AGG GAG GAC GAU GCG G(r(N)r(N) r(N)r(N)G 3’-OH |
| RL3: 5’-P GUG UCA GUC ACU UCC AGC GG 3’-puromycin |

| DNA primers (Operon) |
|---------------------|
| DP5: 5’-AGG GAG GAC GAT GCG G-3’ |
| DP3: 5’-CCG CTG GAA GTG ACT GAC AC-3’ |

| Solexa Fusion Primers (Operon) |
|--------------------------------|
| SSP1: 5’-CTA TGG ATA CTT AGT CAG GGA GGA CGA TGC GG-3’ |

| Circularization RT primer (Dharmacon) |
|--------------------------------------|
| 5’Phos/(GGTTA)(CCGCTAGTCCTCCCT)(CCCTATAGTGAGTCGTATTA)/iSp18/CACTCA/iSp18/(CCGCTGAA GTGACTGACAC)/3’ |

| Antisense DP5 Antisense T7 Promoter DP3 |
|----------------------------------------|
| 1) 5’Phos-GNNNN CGT GAT CCGCATCGTCCTCCCTC CCTATAGTGAGTCGTATTA - iSp18 - CACTCA -iSp18 – CCGCTGGAAGTGACTGACAC |
| 2) 5’Phos-GNNNN ACATCG CCGCATCGTCCTCCCTC CCTATAGTGAGTCGTATTA - iSp18 - CACTCA -iSp18 – CCGCTGGAAGTGACTGACAC |
| 3) 5’Phos-GNNNN GCCCTA CCGCATCGTCCTCCCTC CCTATAGTGAGTCGTATTA - iSp18 - CACTCA -iSp18 – CCGCTGGAAGTGACTGACAC |
| 4) 5’Phos-GNNNN TGCTCA CCGCATCGTCCTCCCTC CCTATAGTGAGTCGTATTA - iSp18 - CACTCA -iSp18 – CCGCTGGAAGTGACTGACAC |
| 5) 5’Phos-GNNNN CACAGT CCGCATCGTCCTCCCTC CCTATAGTGAGTCGTATTA - iSp18 - CACTCA -iSp18 – CCGCTGGAAGTGACTGACAC |
| 6) 5’Phos-GNNNN ATTGGC CCGCATCGTCCTCCCTC CCTATAGTGAGTCGTATTA - iSp18 - CACTCA -iSp18 – CCGCTGGAAGTGACTGACAC |

| Complementary barcode sequence |
|-------------------------------|
| 1) ATCACGNNNNNG…………… |
| 2) CGATGTNNNNNG…………… |
| 3) TAGGGCNNNNG…………….. |
| 4) TGACCANNNNG…………….. |
| 5) ACTGTGNNNNG…………….. |
| 6) GCCAATNNNNG…………….. |
| Time (s) | Dazl: 4.2x L: 2.6 mW | Dazl: 4.2x L: 1 mW | Dazl: 1x L: 2.6 mW | Dazl: 1x L: 1 mW | Stratalinker |
|---------|---------------------|---------------------|---------------------|---------------------|--------------|
| 0       | 5·10^6              | 6·10^6              | 4·10^6              | 3·10^5              | 5·10^6       |
| 30      | 3·10^6              | 3.6·10^6            | 4·10^6              | 8·10^6              | 5·10^6       |
| 180     | 1.9·10^6            | 2.4·10^6            | 4·10^6              | 5·10^6              | 5·10^6       |
| 680     | 0.6·10^6            | 1.2·10^6            | 2·10^6              | 3·10^6              | 5·10^6       |

**Supplementary Table S3 | Number of cells used in each crosslinking experiment**
(L: laser power)
| Time (s) | Dazl: 4.2x L: 2.6 mW | Dazl: 4.2x L: 1 mW | Dazl: 1x L: 2.6 mW | Dazl: 1x L: 1 mW | Stratalinker |
|---------|---------------------|---------------------|---------------------|---------------------|-------------|
| 30      | 88%                 | 98%                 | 80%                 | 91%                 | 91%         |
| 180     | 79%                 | 92%                 | 82%                 | 87%                 | 84%         |
| 680     | 87%                 | 81%                 | 93%                 | 91%                 | 83%         |

**Supplementary Table S4 | Cell Viability after each crosslinking experiment**

(L: laser power). Cell viability was measured by Trypan-blue staining and cell counting in a hemocytometer (Materials and Methods).
## Supplementary Table S5 | Sequencing and read processing statistics.

(a) Post processed reads: Reads remaining after de-multiplexing, adapter removal and PCR duplicate collapsing.

(b) Mapped reads: Reads mapped to mouse genome (mm10).

(c) Correction factor: Intensity per read obtained by normalizing number of reads per condition with total crosslinked RNA.

(d) Reads-Peak intersection: Number of reads corresponding to Dazl binding site peaks common to all KIN-CLIP conditions.
| Conditions | 680 s | 180 s | 30 s | 0 | 680 s | 180 s | 30 s | 0 | 680 s | 180 s | 30 s | 0 | 680 s | 180 s | 30 s | 0 |
|------------|-------|-------|-----|---|-------|-------|-----|---|-------|-------|-----|---|-------|-------|-----|---|
| Dazl: 4.2x |       |       |     |   | Dazl: 1x |       |       |   | Dazl: 4.2x |       |       |   | Dazl: 1x |       |       |   |
| Laser: 2.6 mW |       |       |     |   | Laser: 2.6 mW |       |       |   | Laser: 1 mW |       |       |   | Laser: 1 mW |       |       |   |
| Bulk Crosslinking Intensity (10<sup>5</sup>) | 1.012 ± 0.25 | 0.775 ± 0.22 | 0.537 ± 0.07 | 10<sup>-5</sup> | 0.722 ± 0.19 | 0.384 ± 0.11 | 0.346 ± 0.07 | 10<sup>-5</sup> | 0.343 ± 0.10 | 0.403 ± 0.07 | 0.199 ± 0.07 | 10<sup>-5</sup> | 0.311 ± 0.11 | 0.392 ± 0.07 | 0.336 ± 0.06 | 10<sup>-5</sup> |

**Supplementary Table S6 | Bulk crosslinking intensity for each crosslinking condition.**

Bulk crosslinking (AU; pixel density as described in Image J) was measured as described in Materials and Methods. The errors associated with intensity represent deviation in bulk cross linking as obtained by measuring bulk cross linking for at least three replicates for each time point.
|                          | GC1 Replicate 1 | GC1 Replicate 2 | GC1 Replicate 3 |
|--------------------------|-----------------|-----------------|-----------------|
| Post processed Reads\(^{(a)}\) | 1,351,295       | 910,651         | 996,650         |
| Mapped Reads \(^{(b)}\)     | 123,851         | 59,674          | 71,288          |

**Supplementary Table S7 | Sequencing and read processing statistics for iCLIP experiments.**

\(^{(a)}\) Reads remaining after adapter removal and PCR duplicate collapsing.

\(^{(b)}\) Reads mapped to mouse genome (mm9).
Supplementary Material Figure S1 | Gel Source data for images shown in Extended Data Figures 1d,e and Extended Data Fig.2a,c.
Boxes: gel regions shown in the Extended Data Figures.
**a**

RbFox(RRM)

Relative $\chi^2$ min $(\chi^2/\chi^2_{\text{min}})$

- $k_\text{on}$ ($10^7 \text{M}^{-1}\text{s}^{-1}$)
- $k_d^{(2.6\text{mW})}$ ($\text{s}^{-1}$)
- $k_d^{(1\text{mW})}$ ($\text{s}^{-1}$)

**b**

RbFox$^{\text{mut}}$(RRM)

Relative $\chi^2$ min $(\chi^2/\chi^2_{\text{min}})$

- $k_\text{on}$ ($10^7 \text{M}^{-1}\text{s}^{-1}$)
- $k_d^{(2.6\text{mW})}$ ($\text{s}^{-1}$)
- $k_d^{(1\text{mW})}$ ($\text{s}^{-1}$)

**c**

Dazl(RRM)

Relative $\chi^2$ min $(\chi^2/\chi^2_{\text{min}})$

- $k_\text{on}$ ($10^7 \text{M}^{-1}\text{s}^{-1}$)
- $k_d^{(2.6\text{mW})}$ ($\text{s}^{-1}$)
- $k_d^{(1\text{mW})}$ ($\text{s}^{-1}$)

**d**

- RbFox$^{\text{WT}}$(RRM)
- RbFox$^{\text{mut}}$(RRM)
- Dazl(RRM)

$k_d$ ($\text{mW}$)
Supplementary Figure S2 | fs laser crosslinking fit space parameters.

1D Fit space analysis (KINTEK) for obtained kinetic parameters ($k_{on}$, $k_{on}^{2.6mW}$, $k_{off}$ and $k_{off}^{1mW}$) for (a). RbFox(RRM), (b) RbFox$^{Mut}$(RRM) and (c) Dazl(RRM). (Fig.1e). The relative $X^2$ represents the smallest (optimal) $X^2$ divided by the $X^2$ obtained for the entire thermodynamic model. For the optimal parameter value, the relative $X^2 = 1$. Horizontal lines mark the 95% confidence interval.

d. 2D Fit space analysis of the relative $X^2$ of co-varying $k_{on}$ and $k_{off}$. Both rate constants are constrained for all 3 proteins with a well-defined local minimum (red).
Supplementary Figure S3 | Determination of fractional occupancy ($\Phi_{\text{max}}$)

Maximal amplitude ($\alpha_{\text{max}}$, probability of Dazl bound to the fraction of a given binding site that is accessible during the course of the experiment, extrapolated to saturating concentrations of Dazl) plotted vs. level of the corresponding transcript (RPKM). Eq.44 (Materials and Methods) is used to calculate the maximal fractional occupancy ($\Phi_{\text{max}}$).
Supplementary Figure S4 | Impact of rate constant variation on crosslinking time courses.

(a) Time courses for Dazl binding sites with differing $k_{off}$ values (highlighted; high, medium and low range of the distribution of $k_{off}$ values) and similar values for other rate constants. (b) Time courses for Dazl binding sites with differing values for $k_{on}^{(1xDazl)}$ (s$^{-1}$) (left) and $k_{on}^{(4.2xDazl)}$ (s$^{-1}$) (right), and similar values for other rate constants. (c) Time courses for Dazl binding sites with differing values for $k_{xl}^{(2.6mW)}$ (s$^{-1}$) (left) and $k_{xl}^{(1.0mW)}$ (s$^{-1}$) (right) and similar values for other rate constants. Points mark the experimental normalized peak coverage value (error bars: 95% confidence interval for normalized peak coverage value, determined by minimizing $X^2$), lines show the curves with calculated rate constants. For panels (a-c), points mark the experimental normalized peak coverage value, error bars mark 95% confidence interval for normalized peak coverage value, determined by minimizing $X^2$), and lines show the curves with calculated rate constants.
**Supplementary Figure S5 | Generation of the Multiple Linear Regression Models.**

(a) Flowchart for the development of the multiple linear regression (MLR) models. (b) Adjusted $R^2$ values for all selected candidate models (N = 45 models). (c) Root Mean Squared Error (RMSE) values for all selected candidate models (N = 45 models). (d) $R^2$ values for all selected candidate models (N = 45 models). Models with adjusted $R^2 \geq 0.5$, lowest 50% RMSE and/or $R^2 > 0.5$ were shortlisted (N = 24 models, grey area). Red dots: ΔRNA; Black dots: ΔRPF. (e-g) Information criterion statistics (ICS) for models with separate ΔRNA and ΔRPF terms (N = 15 models; Extended Data Figure 10b-e). ICS for models with merged ΔRNA and ΔRPF (N = 8 models; not shown) was carried out in the similar manner. (e) Models with lowest Akaike’s Information Criterion (AIC) and (f) Bayesian Information Criterion (BIC) are marked (N = 9 models; arrows).
Models with lowest Amemiya’s Prediction Criterion (APC) are selected (lowest 30%). 13 models remain after ICS criterion (9 models with separate ΔRNA and ΔRPF and 4 models with merged ΔRNA and ΔRPF, not shown). (h) F-statistic (one-sided) for models with separate ΔRNA and ΔRPF. All models satisfied general linear F-statistic condition (F-statistic > 15). Heatmap on the right show shortlisted models (N = 6 models) with significant t-tests for majority of independent variable terms (at least 60%, p < 0.05 to 0.005, black, p > 0.05 to 0.5, white), lowest ICS and significant F-test statistic (N = 6 models).
**Supplementary Figure S6 | Shortlisting of Multiple Linear Regression Models.** (a-f) Upper panels: Standardized residuals versus average predicted ΔRNA and ΔRPF values for models remaining after significance testing (N = 6 models, Supplementary Materials Fig.S4). p-value (one-sided): Kolmogorov-Smirnov Test (K-S test) for error normality. p < 0.05 indicates normal distribution of error residuals (Models M1, M19, M24). Lower panels: Histogram of error residuals for models remaining after significance testing (N = 6 models, Supplementary Materials Fig.S4). (g) Correlation between experimental values for ΔRPF (top panel) and ΔRNA (bottom panel) (training data set, N = 699 transcripts; 60%) and values calculated with the linear regression model ($R^2$: adjusted linear correlation coefficient) for models shortlisted in panels a-f.
Supplementary Figure S7 | MLR models for merged ΔRPF and ΔRNA terms

(a) Linear Regression models tested (M28 – M45). (Yellow: dummy coding, using terciles of the variables, Extended Data Fig.8. Red: no dummy coding; use of continuous data. Grey: variable was omitted. ΔRPF and ΔRNA were merged by normalizing both, ΔRPF and ΔRNA to a scale of 0-1 and then multiplying [ΔRPF X (ΔRNA – 0.01]. The merged terms are distinct from the translation efficiency (ΔTE). (b) Adjusted R2 for each model. (c) Differential Intercept Linear Coefficients (DILC) for each model. (d) P-values (one-sided) of t-test for each independent variable (N = 7 features) for all models. p < 0.05 to 0.005, black, p > 0.05 to 0.5, white).
Supplementary Scheme S1 | Numerical data fitting process

Steps for the numerical fitting of crosslinking timecourses to calculate kinetic parameters. Square boxes represent KIN-CLIP conditions (red).
Supplementary Scheme S2 | Analytical data fitting process

Steps for the numerical fitting of crosslinking timecourses to calculate kinetic parameters. Square boxes represent KIN-CLIP conditions (red).