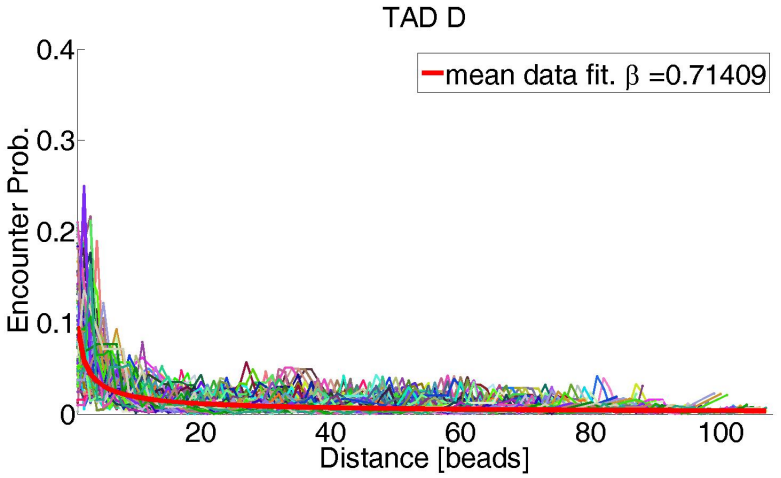
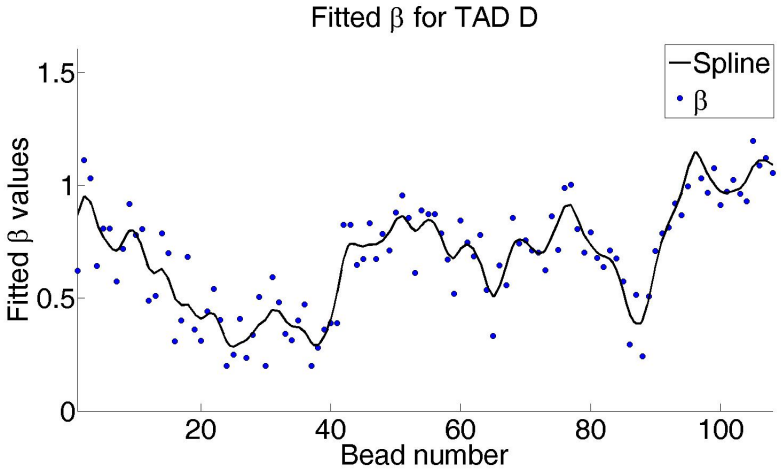
0.1 Peak Calling

To accurately identify pair of beads that interact more frequently than expected, we ran a peak calling procedure on the 5C probability data of each TAD separately. For each bead $i = 1..N$ of the chain, we fit an encounter probability curve of the form $\alpha_i d^{-\beta_i}$. We then use the values of the fitted curves at each distance $d = 1..(N-1)$ to estimate the expected encounter probability, $\mu(d) = E[\alpha_i d^{-\beta_i}]$, and the expected standard-deviation, $\sigma(d) = \sqrt{E[(\mu(d) - \alpha_i d^{-\beta_i})^2]}$, for each d, and $i = 1...N$.

For each observation i at distance d we calculate a z-score by $z_i(d) = \frac{|\alpha_i d^{-\beta_i} - \mu(d)|}{\sigma(d)}$ and fit a Weibull distribution to it. For each z-score a p-value is calculated based on the fitted Weibull CDF and then transformed into a q-value to obtain false discovery rate. We then threshold the data using false discovery rate of 0.01.

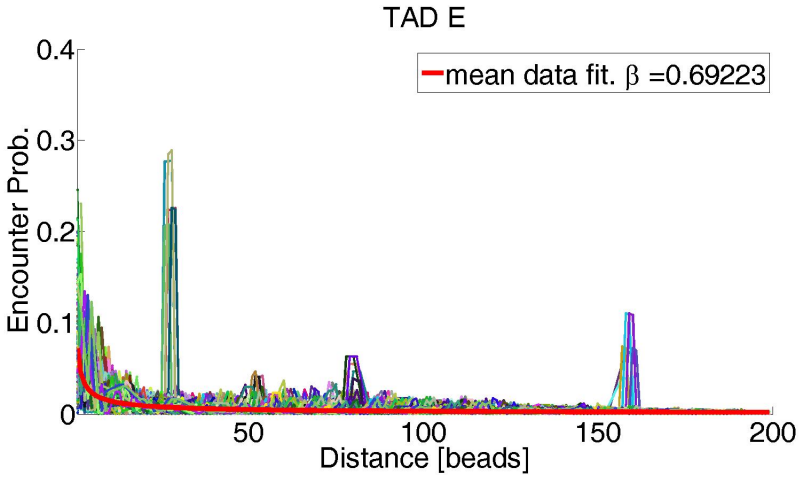


(a)

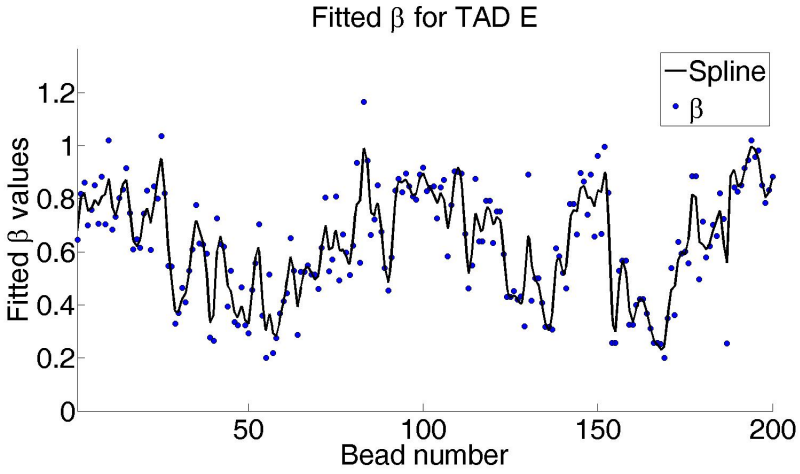


(b)

Figure 1: The encounter probability and the fitted β values for TAD D.



(a)



(b)

Figure 2: The encounter probability and the fitted β values for TAD E.