

## SUPPLEMENTARY MATERIAL

### Estimating the difference between the observed and the apparent DNA affinity and DNA cleavage equilibrium

As described in the Results, the data were analyzed in terms of a simplified reaction framework (Scheme 2). This reaction scheme is derived from a more complex scheme that contains all steps required by the data (Scheme 1), but whose individual equilibrium constants cannot be disentangled with the available data. Reduction of Scheme 1 to Scheme 2 entails grouping together the non-covalent enzyme-DNA species ( $[E_{SS}]$  and  $[E'_{SS}]$ ; Scheme 1) and the covalent species ( $[E'_{SP}]$  and  $[E'_{PP}]$ ), giving rise to the definition of an apparent affinity and DNA cleavage equilibrium (Eqs. 4 & 5; Scheme 2).

$$K_{d,app} = [E] * [SS] / ([E_{SS}] + [E'_{SS}]) \quad (4)$$

$$K_{clvg,app} = ([E'_{SP}] + [E'_{PP}]) / ([E_{SS}] + [E'_{SS}]) \quad (5)$$

Two equilibrium constants, one affinity term and one cleavage equilibrium term, are available from the cleavage data presented. However, because single strand breaks introduced by the enzyme (species  $E'_{SP}$ ) produce only half of the signal of double strand breaks (species  $E'_{PP}$ ) in the DNA cleavage assay (see Materials and Methods), only approximations of values for  $K_{d,app}$  and  $K_{clvg,app}$  (Scheme 2) can be obtained from the DNA cleavage assay. We termed these approximations  $K_{d,obs}$  and  $K_{clvg,obs}$  in the Results. Below, we first derive equations that define how good these approximations are. Using additional thermodynamic data, we then estimate that the observed (i.e. measured) equilibria underestimate the apparent equilibria (Scheme 2) by at most two-fold.

It is possible that additional hitherto unidentified reaction intermediates exist, requiring an expansion of the minimal reaction scheme (Scheme 1). The apparent association and cleavage equilibria can be analogously defined by pooling all non-covalent and all covalent species in such an expanded scheme together, reducing it to Scheme 2. As before, the fact that single cuts in the DNA produce only half the signal of double cuts prevents us from directly measuring these apparent equilibria. The deviation of the observed from the apparent equilibria for such an expanded scheme, however, would be the same as for the minimal reaction depicted in Scheme 1 (derivation not shown). Conclusions in this manuscript are therefore unaffected by an expansion of the reaction scheme.

#### Estimation of the deviation of $K_{clvg,obs}$ from $K_{clvg,app}$

We first derive an expression for the *apparent* cleavage equilibrium ( $K_{clvg,app}$ ) and the experimentally *observed* cleavage equilibrium ( $K_{clvg,obs}$ ) based on the minimal reaction shown in Scheme 1. By relating the expressions for the apparent and the observed cleavage equilibria to one another, we could calculate their deviation, and we could estimate the magnitude of this deviation for the duplexes we used. We then follow the same approach for the apparent and the observed affinities in the subsequent section.

According to Scheme 1, the observed fraction of cleavage ( $fr$ ) is given by

$$fr = ([E_{SP}] + 2 [E_{PP}]) / 2 [DNA]_{tot} \quad (6)$$

For saturating enzyme concentrations  $fr$  approaches  $maxcleavage$  (Eq. 7; Scheme 1).

$$maxcleavage = 0.5 K_{iso} K_{clvg,1} (1 + 2 K_{clvg,2}) / (1 + K_{iso} (1 + K_{clvg,1} + K_{clvg,1} K_{clvg,2})) \quad (7)$$

The observed cleavage equilibrium can be calculated from  $maxcleavage$  (Eq. 2; see Results), yielding

$$K_{clvg,obs} = (K_{iso} K_{clvg,1} + 2 K_{iso} K_{clvg,1} K_{clvg,2}) / (2 + 2 K_{iso} + K_{iso} K_{clvg,1}) \quad (8)$$

We rewrite Eq. 5 to obtain an expression for  $K_{clvg,app}$  only dependent on the equilibrium constants in Scheme 1 (Eq. 9).

$$K_{clvg,app} = (K_{iso} K_{clvg,1} + K_{iso} K_{clvg,1} K_{clvg,2}) / (1 + K_{iso}) \quad (9)$$

The ratio of  $K_{clvg,obs}$  and  $K_{clvg,app}$  describes the deviation of these constants from one another ( $ErrK_{clvg}$ ; Eq. 10).

$$\begin{aligned} ErrK_{clvg} &= K_{clvg,app} / K_{clvg,obs} \\ &= (1 + K_{iso} + 0.5 K_{iso} K_{clvg,1}) (1 + K_{clvg,2}) / [(1 + K_{iso})(0.5 + K_{clvg,2})] \end{aligned} \quad (10)$$

Because equilibrium constants can by definition never take on negative values,  $ErrK_{clvg}$  can according to Eq. 10 never assume values below one. Thus,  $K_{clvg,obs}$  is always an underestimate of  $K_{clvg,app}$ .

With the help of additional thermodynamic information, we can estimate the magnitude of  $ErrK_{clvg}$ . We obtained quantitative information about the equilibria in Scheme 1 for one particular DNA duplex, Dup1<sub>34 bp</sub>; G<sub>2T</sub> (manuscript in preparation). For Dup1<sub>34 bp</sub>; G<sub>2T</sub>, we directly measured  $K_{clvg,2}$  and an observed equilibrium ( $K_{obs}$ ) describing the isomerization step ( $K_{iso}$ ; Scheme 1) together with the first strand cleavage event ( $K_{clvg,1}$ ; Eq. 11).

$$K_{obs} = [E'_{SP}] / ([E_{ss}] + [E'_{ss}]) = K_{clvg,1} / (1 / K_{iso} + 1) \quad (11)$$

Single strand breaks relative to double strand breaks did not strongly accumulate. Thus,  $K_{clvg,2} \gg K_{obs}$ . With this additional thermodynamic information,  $ErrK_{clvg}$  can be approximated from Eqs. 10 and 11 according to Eq. 12.

$$ErrK_{clvg} \sim (2 + K_{obs}) / 2 \quad (12)$$

For Dup1<sub>34 bp; G2T</sub>,  $K_{obs} = 0.26$  (manuscript in preparation) and, according to Eq. 12,  $ErrK_{clvg} \sim 1.13$ . In other words, the observed cleavage equilibrium underestimates the apparent equilibrium by  $\sim 13\%$  for this duplex.

To calculate  $ErrK_{clvg}$  also for the other DNA duplexes, values for their corresponding  $K_{obs}$  must be estimated. Dup1<sub>34 bp; G2T</sub> belongs to one of the best cleavage substrates we have identified (Table 1A). Therefore,  $K_{obs}$  for the other DNA duplexes would be expected to assume values that are smaller than 0.26. Assuming a conservative upper limit of 2 for  $K_{obs}$ , the observed cleavage equilibria would underestimate the apparent equilibria by at most two-fold (Eq. 12).

#### Estimation of the deviation of $K_{d,obs}$ from $K_{d,app}$

We now derive an expression describing the deviation of the apparent DNA affinity ( $K_{d,app}$ ) from the experimentally observed affinity ( $K_{d,obs}$ ) and estimate the magnitude of this deviation for the duplexes we used. In analogy to the derivation above, we first derive expressions for  $K_{d,app}$  and  $K_{d,obs}$  based on the mechanism in Scheme 1. We then relate these two expressions to one another to calculate their deviation. Through substitutions, we can rewrite Eq. 4 and obtain

$$K_{d,app} = (K_{1/2} + K_{1/2} K_{iso} + K_{1/2} K_{iso} K_{clvg,1} + K_{1/2} K_{iso} K_{clvg,1} K_{clvg,2}) / (1 + K_{iso}) \quad (13)$$

with  $K_{1/2}$  representing the enzyme concentration at which DNA cleavage is half maximal at equilibrium (see Fig. 2). With Eq. 3 (see Results) and Eq. 8 we obtain an expression for  $K_{d,obs}$  (Eq. 14).

$$K_{d,obs} = K_{1/2} [1 + (K_{iso} K_{clvg,1} + 2 K_{iso} K_{clvg,1} K_{clvg,2}) / (2 + 2 K_{iso} + K_{iso} K_{clvg,1})] \quad (14)$$

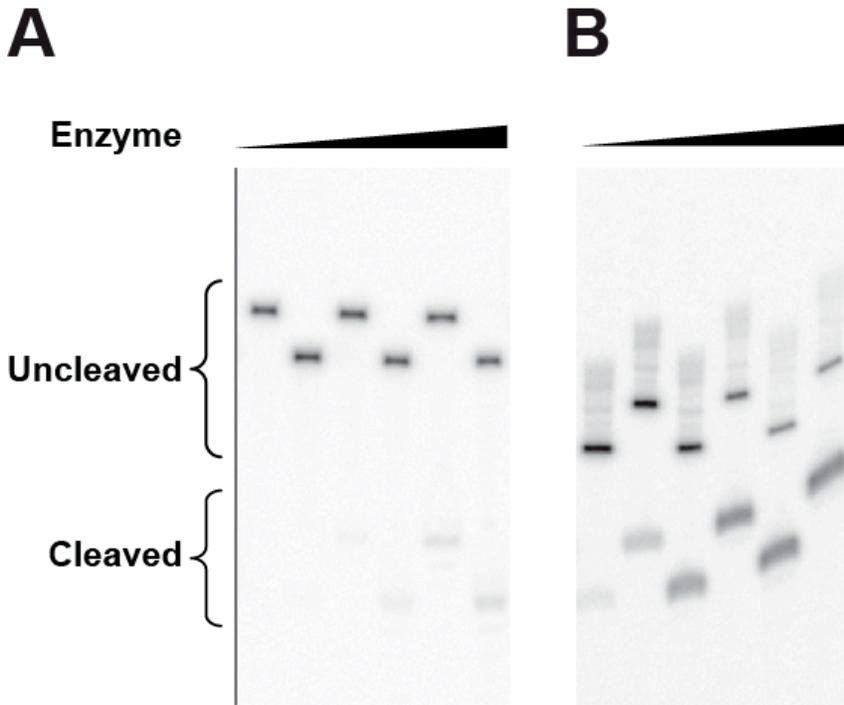
We define the deviation of the observed from the apparent affinity analogously to the deviation of the observed from the apparent cleavage equilibrium (see above):

$$ErrK_d = K_{d,app} / K_{d,obs} = (2 + 2 K_{iso} + K_{iso} K_{clvg,1}) / (2 + 2 K_{iso}) \quad (15)$$

Inspection of Eq. 15 shows that, like  $ErrK_{clvg}$ ,  $ErrK_d$  can never assume values below one. Using Eq. 11, Eq. 15 can be rewritten (Eq. 16).

$$ErrK_d = (2 + K_{obs}) / 2 \quad (16)$$

Comparing Eq. 16 with Eq. 12, we note that  $ErrK_d$  equals the approximation for  $ErrK_{clvg}$ . For the same reasons described in the previous section, we suggest that  $ErrK_d$  never exceeds two for the duplexes used. The observed DNA affinities ( $K_{d,obs}$ ) therefore underestimate the apparent affinities ( $K_{d,app}$ ) by at most two-fold.



**Fig. S1:** DNA cleavage at equilibrium with varying enzyme concentrations for a good (Dup<sub>134 bp</sub>; A) and a poor DNA cleavage substrate (Dup<sub>134 bp</sub>; G<sub>2T</sub>, C<sub>6T</sub>, C<sub>9T</sub>; B). DNA cleavage was measured as in Fig. 2A. Every second lane was loaded on the gel after a time delay to increase the space between the bands for optimal quantification. Quantification of the data is shown in Fig. 2B. Smearing of the band belonging to the uncleaved DNA oligonucleotide indicates residual secondary structure, a result of the self-complementarity of the oligonucleotide. This residual structure can be broken using strongly denaturing conditions (40% formamide, 7 M urea; data not shown).