Localization of DNA damage by current exchanging repair enzymes:
effects of cooperativity on detection time

Kasper Astrup Eriksen

Department of Theoretical Physics, Lund University,
Sövegatan 14A, SE-223 62 Lund, Sweden.
(Dated: September 23, 2013)

How DNA repair enzymes find the relatively rare sites of damage is not known in great detail. Recent experiments and molecular data suggest that the individual repair enzymes do not work independently of each other, but rather interact with each other through currents exchanged along DNA. A damaged site in DNA hinders this exchange and this makes it possible to quickly free up resources from error free stretches of DNA. Here the size of the speedup gained from this current exchange mechanism is calculated and the characteristic length and time scales are identified. In particular for Escherichia coli we estimate the speedup to be $50000/N$, where $N$ is the number of repair enzymes participating in the current exchange mechanism. Even though $N$ is not exactly known a speedup of order $10$ is not entirely unreasonable. Furthermore upon over expression of repair enzymes the detection time only varies as $N^{-1/2}$ and not as $1/N$. This behavior is of interest in assessing the impact of stressful and radioactive environments on individual cell mutation rates.

PACS numbers:

I. MODEL

The model is pictorially presented in Figure 1. See also [2, 3]. The repair enzyme MutY contains an evolutionary well-conserved [4Fe4S] cluster that is suspected to change its charge configuration from +2 to +3 upon binding to DNA [3]. Upon binding to DNA an electron is thus emitted into DNA, while upon receipt of an electron from DNA the MutY-DNA binding is destabilized. As only error free stretches of DNA are able to transport the electron from one MutY enzyme to a neighboring MutY enzyme [4] this charge exchange enables MutY to quickly free up scanning resources from error free stretches of DNA [5] — see Figure 1. In the original proposal only MutY enzymes participate in the current exchange and to illustrate the line of thought we first consider this scenario. Afterward, we then extend the consideration to the case where there are many different kinds of repair enzymes each specialized to fix a specific kind of damage. Furthermore, we also consider the effect of a finite scan length before MutY spontaneously drops...
The rate with which a MutY enzyme randomly docks onto DNA within a distance of the order \( k v t \) onto a specific base pair and start scanning is denoted by \( t_{\text{detect}} \). After a base pair get damaged, how long time does it take before a MutY enzyme has located the error? In the regime where current scanning yields a considerable speedup of the order \( vT \), the average time it take for any repair enzyme to find a damaged point in the traditional scenario without any cooperation between MutY enzymes. In terms of \( T \) and \( \tau \) the detection time is

\[
t_{\text{detect}} \approx \sqrt{T \tau} = T \sqrt{\frac{\tau}{T}}.
\]

A more detailed model calculation in the supplementary material yields the same result apart from a factor 1.3. From standard Poisson statistics \( T \) is also the detection time of a damaged point in the traditional scenario without any cooperation between MutY enzymes. The cooperation thus give a speed up of order \( \sqrt{T / \tau} \). Before I get down to consider numbers and orders of magnitude I want to make the model a tad more realistic.

**B. Full model**

The functionally central [4Fe4S] cluster is also present in other repair enzymes e.g. endonuclease III. It is thus very likely that other repair enzymes also are able to inject currents into DNA and participate in the electrical scanning of DNA. Furthermore all of these repair enzymes get ’attracted’ to the damaged DNA pair in exactly the same way as MutY. In the above model and calculation we thus have to replace ’MutY’ by ’any repair enzyme participating in the DNA mediated charge transport’ — for short just a repair enzyme below. Likewise the calculated detection time \( t_{\text{detect}} \) becomes the time before the first repair enzyme finds the damaged site and \( T \) the average time it take for any repair enzyme to find the site without using currents. I have here implicitly assumed that both the scan velocity \( v \) and the single repair enzyme attempt frequency \( 1 / \tau \) are of the same order of magnitude for all repair enzymes i.e. MutY is a typical repair enzyme. Biologically the time \( t_{\text{detect}} \) is not the most relevant one as the first repair enzyme that arrives at the damaged base pair very likely is unable to fix the damage. For instance is MutY’s primary target oxidation damage resulting in G:A and 8-oxo-G:A mismatched pairs. In other words the calculation in the supplementary material yields the same result apart from a factor 1.3. From standard Poisson statistics \( T \) is also the detection time of a damaged point in the traditional scenario without any cooperation between MutY enzymes. The cooperation thus give a speed up of order \( \sqrt{T / \tau} \). Before I get down to consider numbers and orders of magnitude I want to make the model a tad more realistic.

**A. Only MutY and it scans forever**

After a a base pair get damaged, how long time does it take before a MutY enzyme has located the error? In the regime where current scanning yields a considerable speedup in the location of the error, the damaged base pair typically is located by a MutY enzyme that binds to DNA next to it and then scans all the way up to and including the faulty base pair (and not by a MutY enzyme that just happened to be nearby at the time of damage). Let \( t_{\text{detect}} \) denote the typical detection time and \( v \) the scan velocity of MutY. The MutY enzyme then typically docks onto DNA within a distance of order \( vt_{\text{detect}} \). The rate with which a MutY enzyme docks onto a specific base pair and start scanning is denoted by \( k \). The probability for a MutY enzyme to land within a distance of \( vt_{\text{detect}} \) of the error in the time interval \( t_{\text{detect}} \) can be estimated as \( k v t_{\text{detect}} \). As at time \( t_{\text{detect}} \) a MutY enzyme typically arrives at the faulty base pair this probability is of order 1 and

\[
t_{\text{detect}} \approx \frac{1}{\sqrt{k v}}.
\]

where \( \tau \) is the time between two successive binding events for a single MutY enzyme. \( N(N(\text{MutY})) \) is the total number of MutY enzymes and \( L \) is the total number of base pairs in DNA. \( T = L / v / N(N(\text{MutY})) \) is the time it takes for all the MutY enzymes to scan all the bases of DNA once. The probability for a MutY enzyme to land within a distance of order \( k v t \) onto a specific base pair and start scanning is denoted by \( t_{\text{detect}} \). As at time \( t_{\text{detect}} \) a MutY enzyme typically arrives at the faulty base pair this probability is of order 1 and

\[
t_{\text{detect}} \approx \frac{1}{\sqrt{k v}}.
\]

A more detailed model calculation in the supplementary material yields the same result apart from a factor 1.3. From standard Poisson statistics \( T \) is also the detection time of a damaged point in the traditional scenario without any cooperation between MutY enzymes. In terms of \( T \) and \( \tau \) the detection time is

\[
t_{\text{detect}} \approx \sqrt{T \tau} = T \sqrt{\frac{\tau}{T}}.
\]

The cooperation thus give a speed up of order \( \sqrt{T / \tau} \). Before I get down to consider numbers and orders of magnitude I want to make the model a tad more realistic.
as $T/T(\text{MutY})$, where $T(\text{MutY})$ is the time it takes for a MutY enzyme to find the site by scanning alone. The MutY detection time is thus

$$t_{\text{detect}}(\text{MutY}) \approx T(\text{MutY})\sqrt{\frac{\tau}{T}}, \quad (4)$$

again corresponding to a speedup of size $\sqrt{\tau/T}$, but this time $T$ is the time it take for any repair enzyme to locate the damage.

I have not yet considered that MutY is known spontaneously to drop off DNA after having scanned of the order 100 base pairs (bp) [1]. In order to estimate the effect of this I am going to derive the above result Eq. (4) in a slightly different manner. The MutY enzyme that eventually localizes the damage typically docks onto DNA within a distance $\Delta$ from the faulty base pair, where $\Delta$ both has to be small enough to allow scanning all the way up to the error i.e. $< 100$bp and also so small that it is unlikely another repair enzyme dock in front of the MutY enzyme while it scans. The last length scale can be estimated to be of the order $\Delta = \sqrt{v/k} = v\sqrt{T/\tau}$. The MutY detection time $t_{\text{detect}}(\text{MutY})$ is then determined as above by setting the probability for a MutY enzyme to dock within a distance $\Delta$ in the time interval $t_{\text{detect}}(\text{MutY})$ equal to 1 i.e.

$$t_{\text{detect}}(\text{MutY}) = T(\text{MutY})\frac{v\tau}{\Delta}, \quad (5)$$

where $\Delta = \min(100 \text{ bp}, v\sqrt{T/\tau})$. Notice that the choice $\Delta = v\sqrt{T/\tau}$ leads to Eq. (4).

C. Estimating order of magnitude

I have not been able to find any direct experimental measurements of $\tau$ and $T$, so I am instead going to use a somewhat more uncertain way to estimate the size of the reduction $\Delta_{\text{scanning}}$. The distance $v\tau$ is the average scan length of a repair enzyme (MutY) in the presence of the current interactions. The numerator $\Delta$ is the smallest of the maximal scan length 100bp and the docking distance $\sqrt{v\tau T}$. In fact, I believe these two distances are of the same order of magnitude as anything else seems inefficient, so I assume $\sqrt{v\tau T} \leq 100$bp, with equality as the most likely option. The distance $v\tau = L/N$ is the average distance between repair enzymes. From this I can estimate $v\tau$. The reduction is thus $\frac{\Delta_{\text{scanning}}}{\tau} \geq v\tau/100 \text{ bp} = 5 \cdot 10^4/N$, where $N$ is the total number of repair enzymes with a current exchange mechanism similar to MutY and I have used that the length of E.Coli’s DNA is $5 \cdot 10^6$ base pairs. Unfortunately $N$ is unknown. In [2] the number of the two $[4\text{Fe}4\text{S}]^{2+}$ containing repair enzymes MutY and endonuclease III is estimated to be 30 and 400 respectively. In the same paper the number of formamidopyrimidine glycosylase (FAPy or MutM) repair enzymes is estimated to 400. FAPy does not contain the $[4\text{Fe}4\text{S}]^{2+}$ cluster. The target of FAPy is 8oxoG which is estimated to constitute 5% of all adducts due to oxidative damage in E. coli. All in all it seems reasonable that the total number of repair enzymes participating in the current exchange mechanism is significantly smaller than 50000, and that a speed up of order 10 is realistic. Notice this would correspond to a typical scan length 10 times smaller than the maximal one (100 bp) due to the current exchange mechanism ($v\tau$ is of the order 10 bp).

II. CONCLUSION

In this paper we have considered some of the implications of a very interesting proposal for cooperation between repair enzymes in the localization of defects in single base pairs. First we have pointed out that the mechanism is likely to speed up the localization by a factor of order 10 compared with independent scanning by the repair enzymes. If an error is detected by the 30 MutY enzymes in say 20 minutes (it has to be considerably shorter than the replication time [7]) this corresponds to a reduction in scanning speed from 125 bp/sec to 13 bp/sec. For comparison, the scan velocity for RNA polymerase is 50 bp/sec, while for DNA polymerase it is 1000 bp/sec. Another in principle testable prediction is that upon over expression of all the repair enzymes with say a factor 4 the detection time is only decreased by a factor 2 (both $T(\text{MutY})$ and $T$ are 4 times smaller), because information is not carried as efficiently along the DNA any more. Physiologically oxidative and radiative environments may result in an increased expression of repair enzymes [3], so the square root behavior and the coupling of the effectiveness of different kinds of repair enzymes is potentially of huge importance for the mutation rates in these kinds of stress full environments. Summing up, only further experimentation can finally confirm this charge transport mechanism, while we here have demonstrated that the mechanism indeed offers a great benefit for the cell. In addition, the model is a nice toy model for protein cooperativity and one might wonder if the underlying principles behind could be of practical use in apparently unrelated engineering problems.

Acknowledgments

Kasper Astrup Eriksen acknowledges support from both the Danish Natural Science Research Council grant number 21-03-0284 and the Bio+IT programme under the Øresund Science Region and Øforsk.
[1] A.W. Francis and S.S. David Escherichia coli MutY and Fpg Utilize a Processive Mechanism for Target Location. Biochemistry 2003 42: 801-810

[2] A. Ananthaswamy, Enzymes scan DNA using electric pulse. New Scientist 2003 180 no. 2417: 10-10

[3] E.M. Boon, A.L. Livingston, N.H. Chmiel, S.S. David, and J.K. Barton DNA-mediated charge transport for DNA repair PNAS 2003 100: 12543-12547

[4] S. Delaney and J.K. Barton Long-Range DNA Charge Transport J. Org. Chem. 2003 68: 6475-6483

[5] B. Demple, and L. Harrison Repair of Oxidative Damage to DNA: Enzymology and Biology Annu. Rev. Biochem. 1994 63: 915-948

[6] K.B. Beckman and B.N. Ames Oxidative Decay of DNA J. Biol. Chem. 1997 272: 19633-19636

[7] Y. Nghiem, M. Cabrera, C.G. Cupples, and J.H. Miller The mutY gene: A mutator locus in Escherichia coli that generates G·C→T·A transversions PNAS 1988 85: 2709-2713

[8] H.S. Kim, Y.W. Park, H. Kasai, S. Nishimura, C.W. Park, K.H. Choi, and M.H. Chung Induction of E. coli oh^Gua endonuclease by oxidative stress: its significance in aerobic life Mutat. Res. 1996 363: 115-123