DNA charge transport: correlation with pathogenic mutations

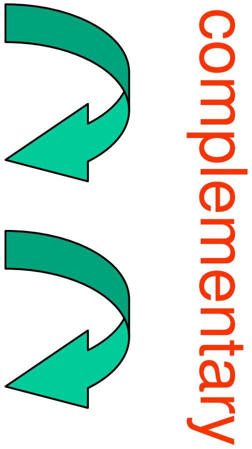
Stephen A. Wells
(s.a.wells@warwick.ac.uk)
Rudolph A. Roemer, Chi-tin Shih,
Stephan Roche

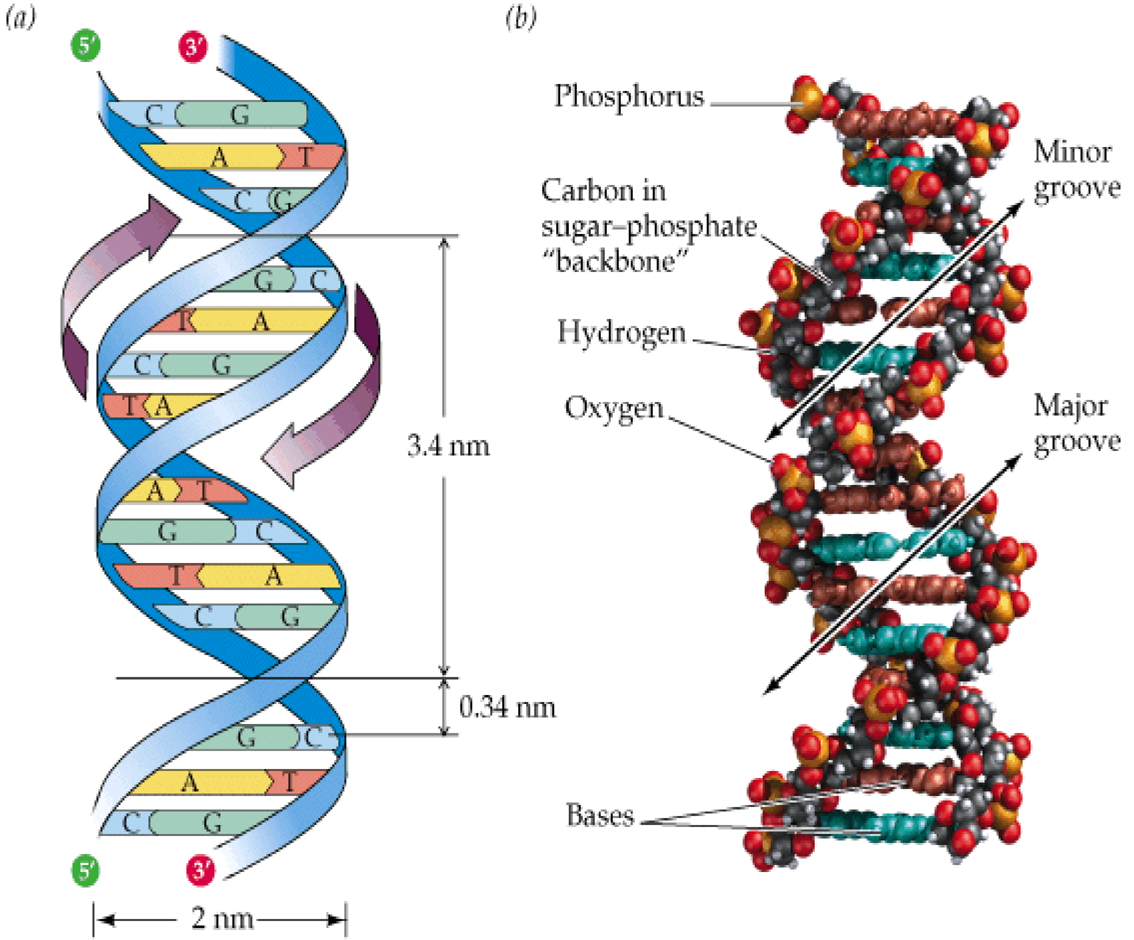
Centre for Scientific Computing
THE UNIVERSITY OF WARWICK
Centre for Scientific Computing

DNA (Deoxyribonucleicacid)

Linear bio-polymer, backbone of repeated sugar-phosphate units, with paired “bases”

- G uanine
- C ytosine
- A denine
- T hymine


complementary



DNA & electronic transport

Conductor:

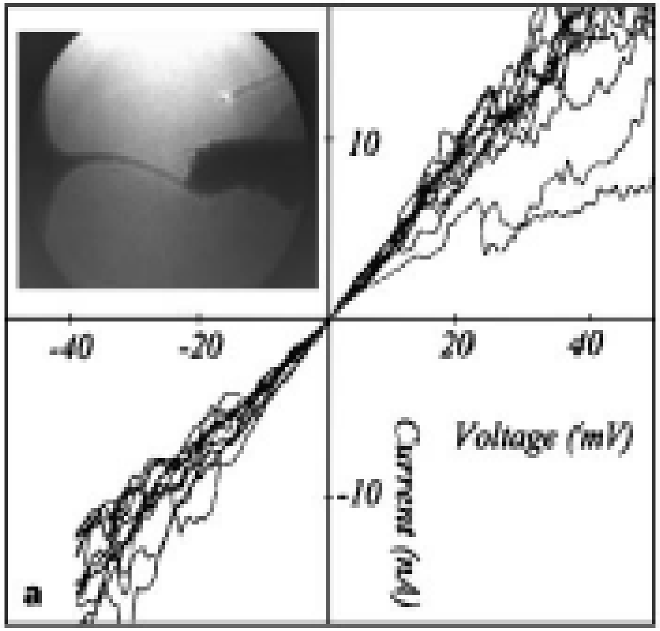
[Fink/Schoenenberger, *Nature* **398**, 407 (1999)]

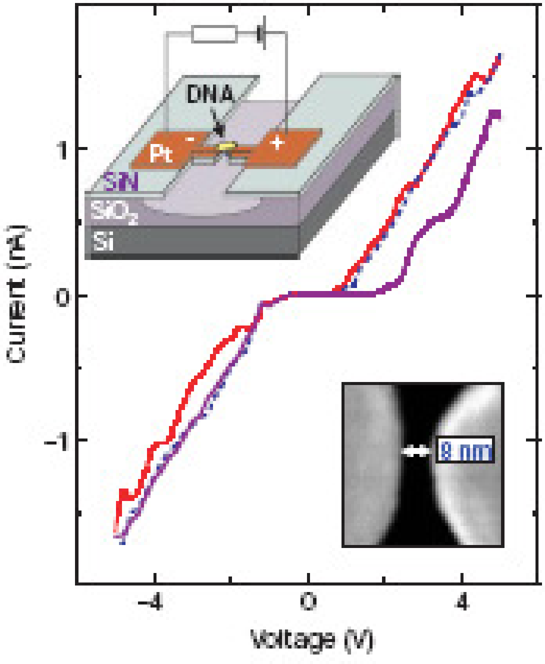
Semiconductor:

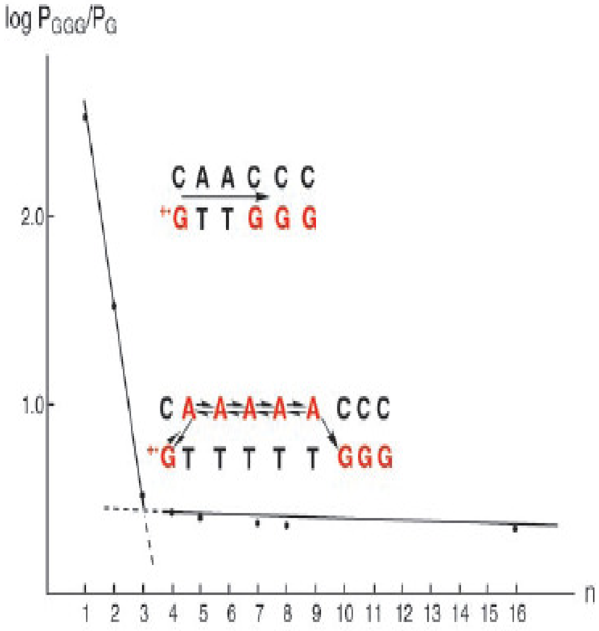
[Porath et al., *Nature* **403**, 635 - 638 (10 Feb 2000)]

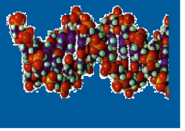
Insulator:

B. Giese, Annu. Rev. Biochem. **71**, 51 (2002)



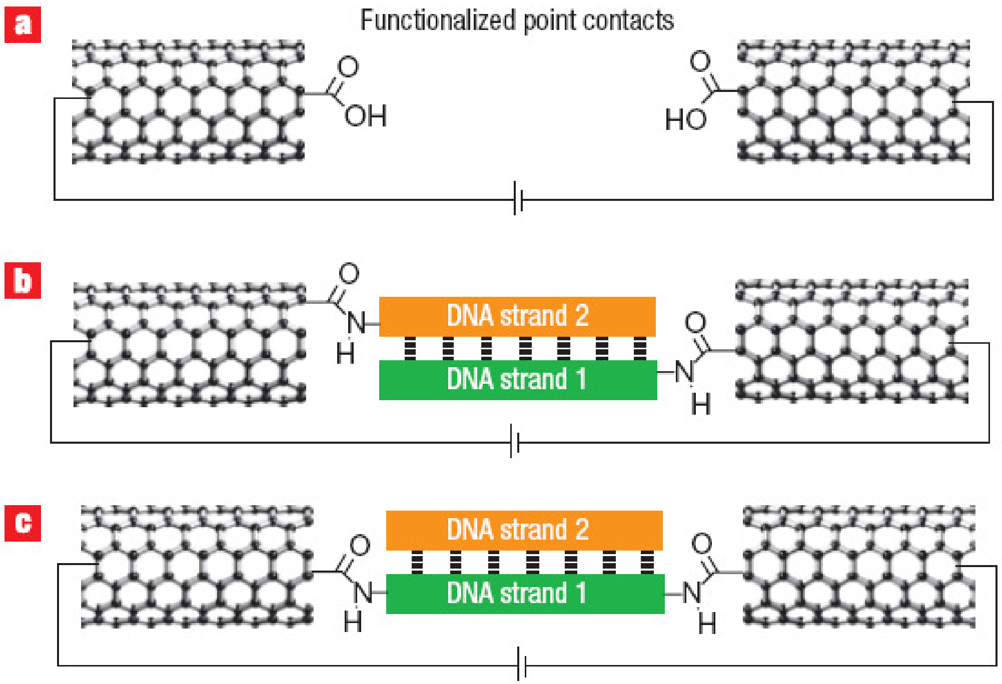


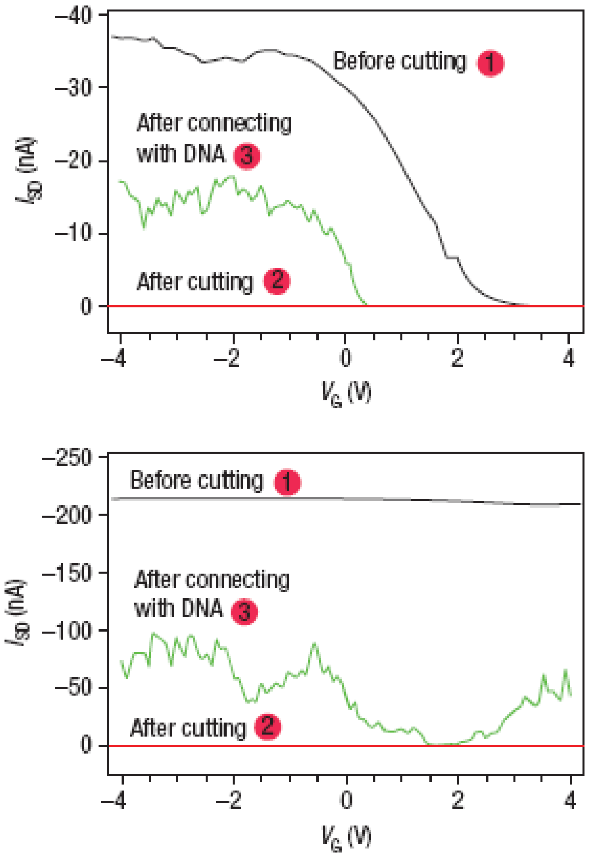


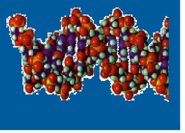


Combining DNA & electronics

Xuefeng Guo et al., *Nature Nanotech.* 2008



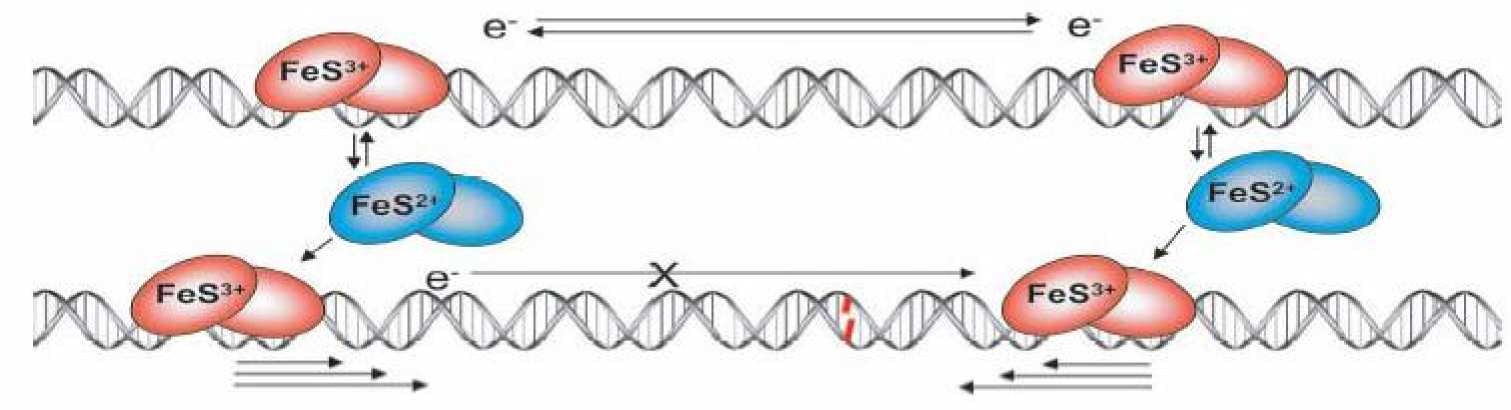




Charge transport and DNA repair

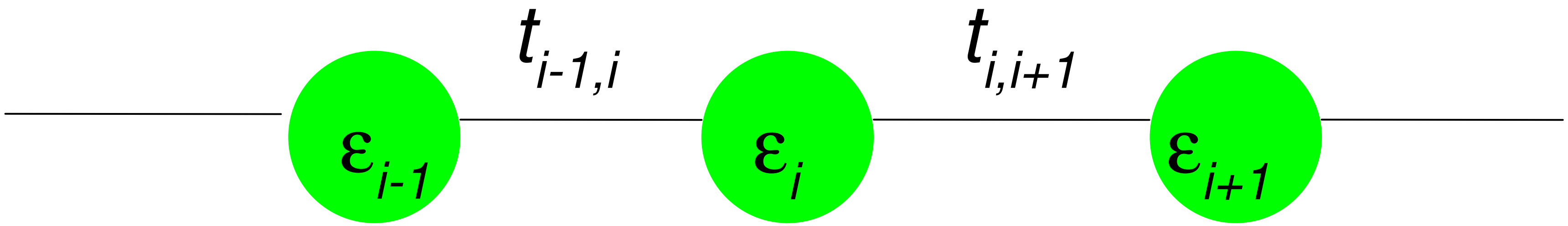
BER (base excision repair) enzyme with $[Fe_4S_4]^{2+}$ cluster – robust to oxidation in the absence of DNA

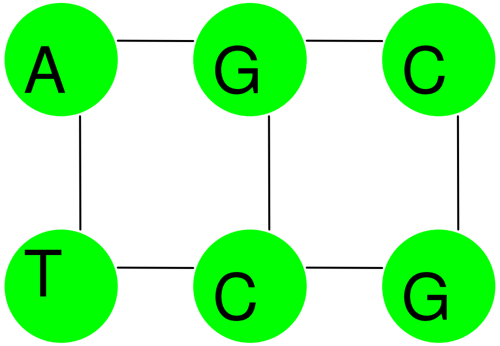
BER binding to DNA – oxidation activated ($[Fe_4S_4]^{2+} \rightarrow [Fe_4S_4]^{3+}$)

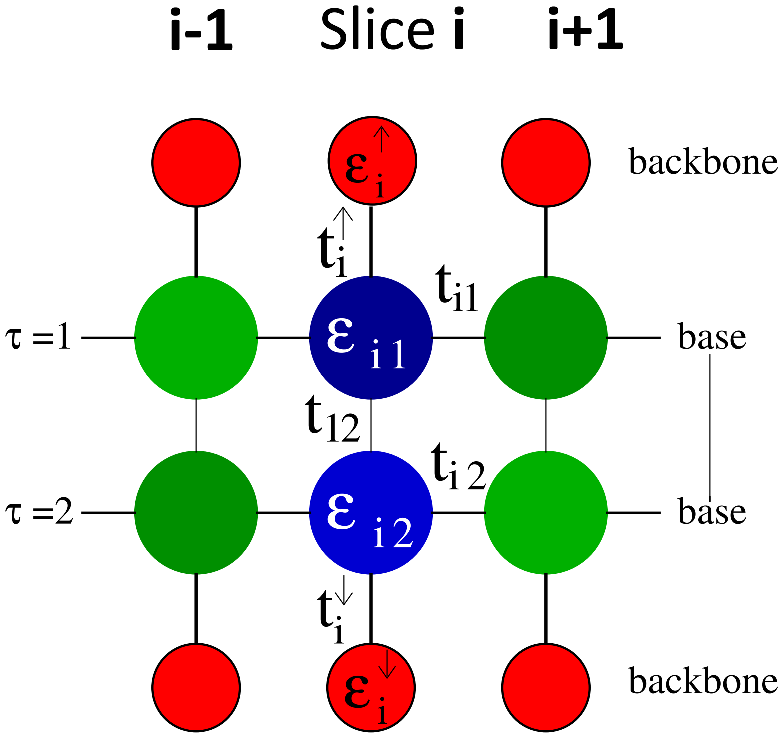


E. Yavin et al. (JK Barton group), PNAS **103**, 3610 (2006).

Tight-binding models







Tight-binding model parameters

Use ionization potentials for onsite energies:

$$\varepsilon_G = 7.75\,e\mathrm{V}$$

$$\varepsilon_C = 8.87\,e\mathrm{V}$$

$$\varepsilon_A = 8.24\,e\mathrm{V}$$

$$\varepsilon_T = 9.14\,e\mathrm{V}$$

Phenomenology/guessing for transfer parameters:

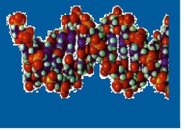
Example for 1-D ladder: hopping=0.4 eV

For ladder model: “like” hopping 0.35 eV,

“unlike” hopping 0.17 eV

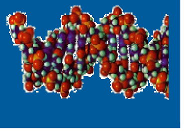
interchain hopping 0.1 eV

Transfer matrix method to extract localisation lengths, Lyapunov exponents, charge transmission.

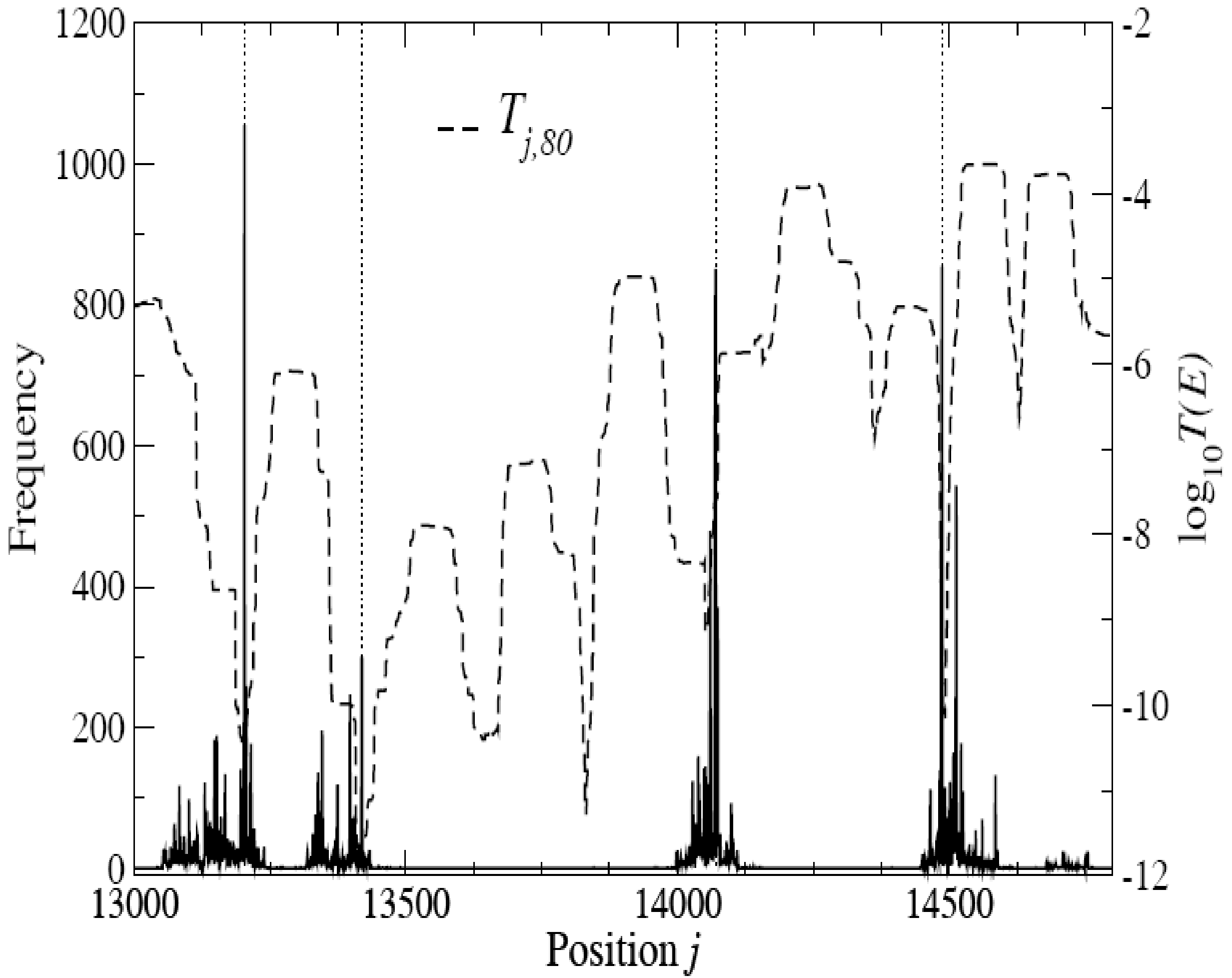


Biology

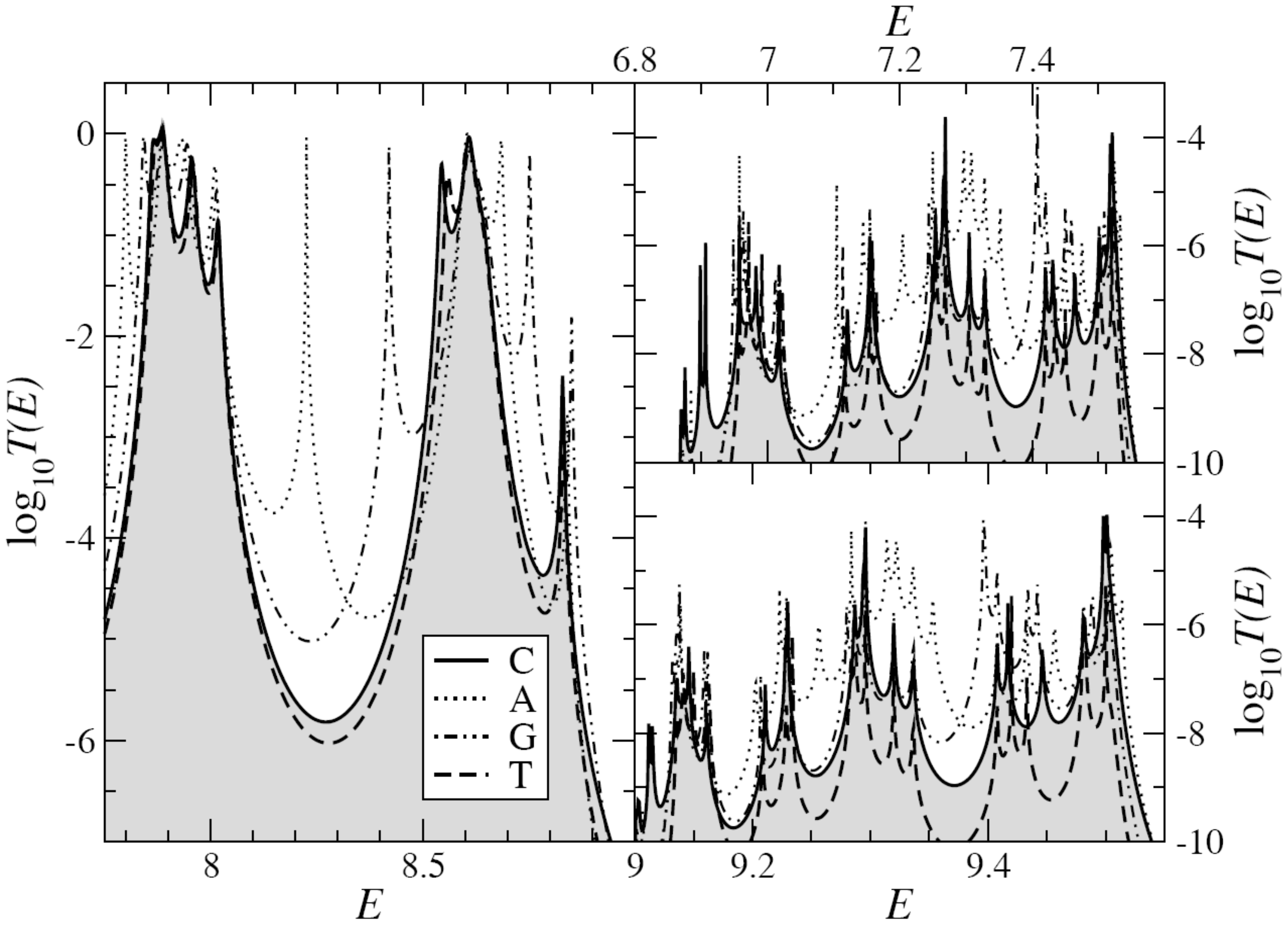
- Mutations in an important oncogene, p53
- “Guardian of the genome”
- “Point Mutations Effects on Charge Transport Properties of the Tumor-Suppressor Gene p53”; Chi-tin Shih *et al*., PRL 2008.

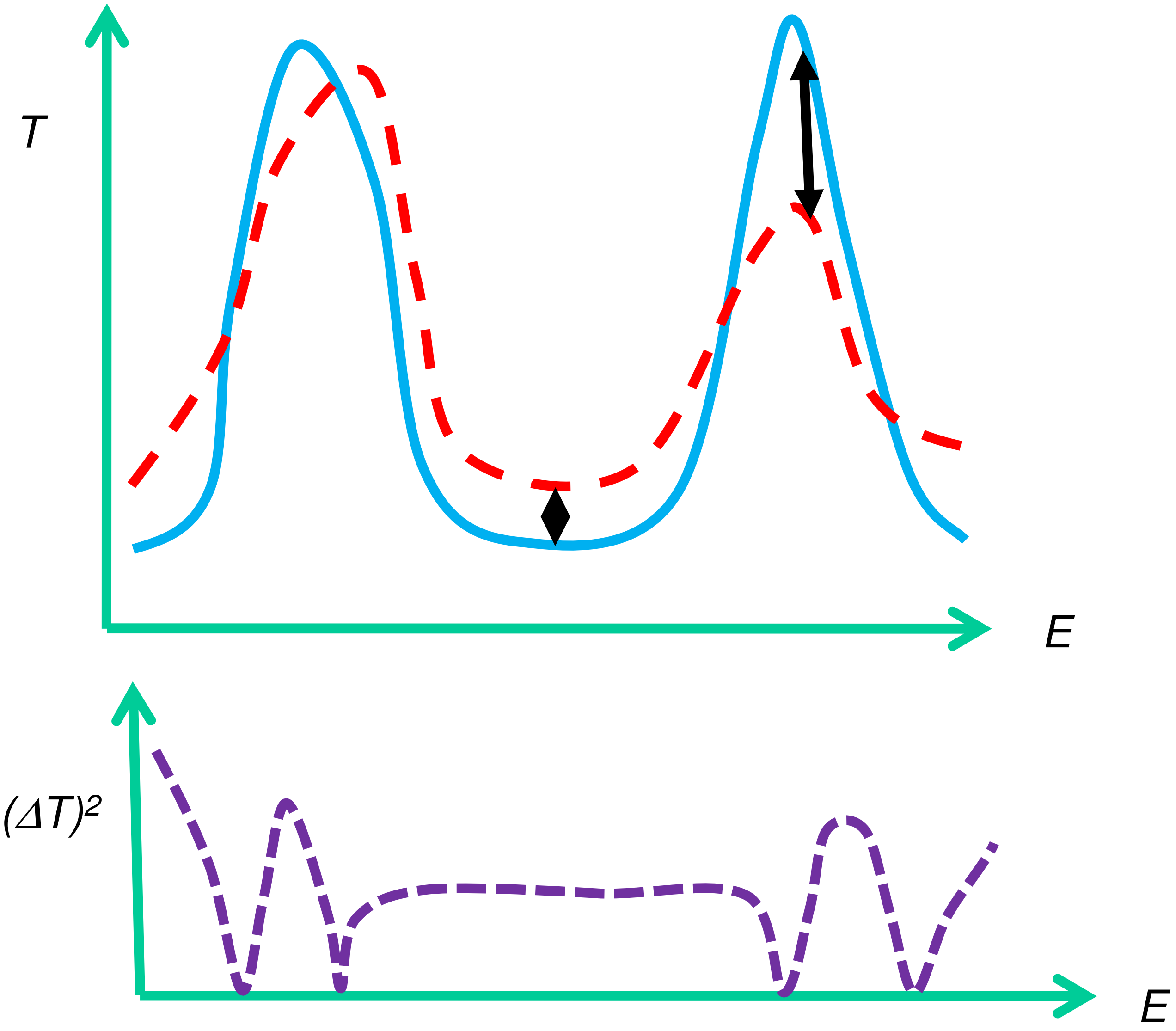


Mutation "hotspots" correlate with low charge transmission.

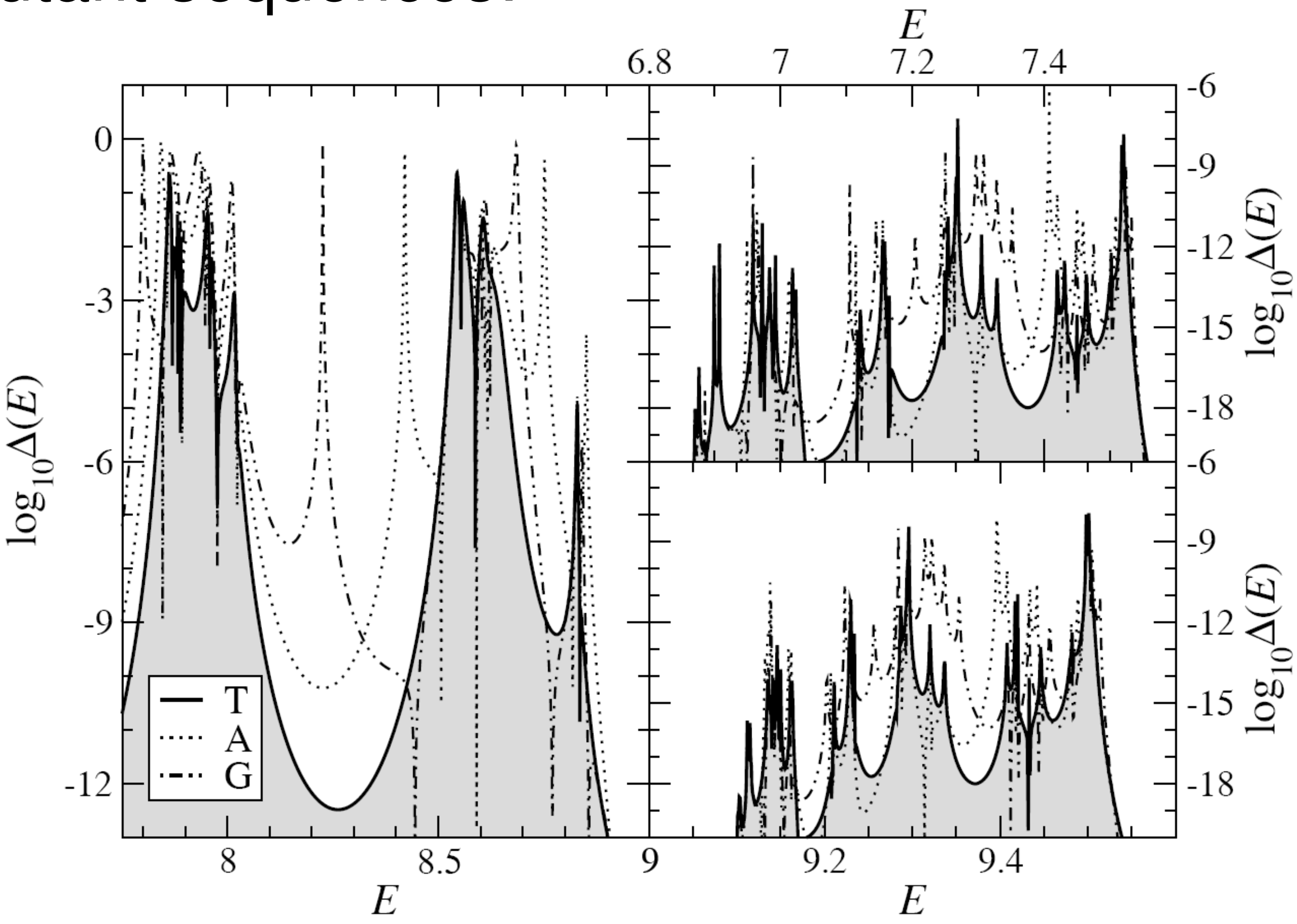


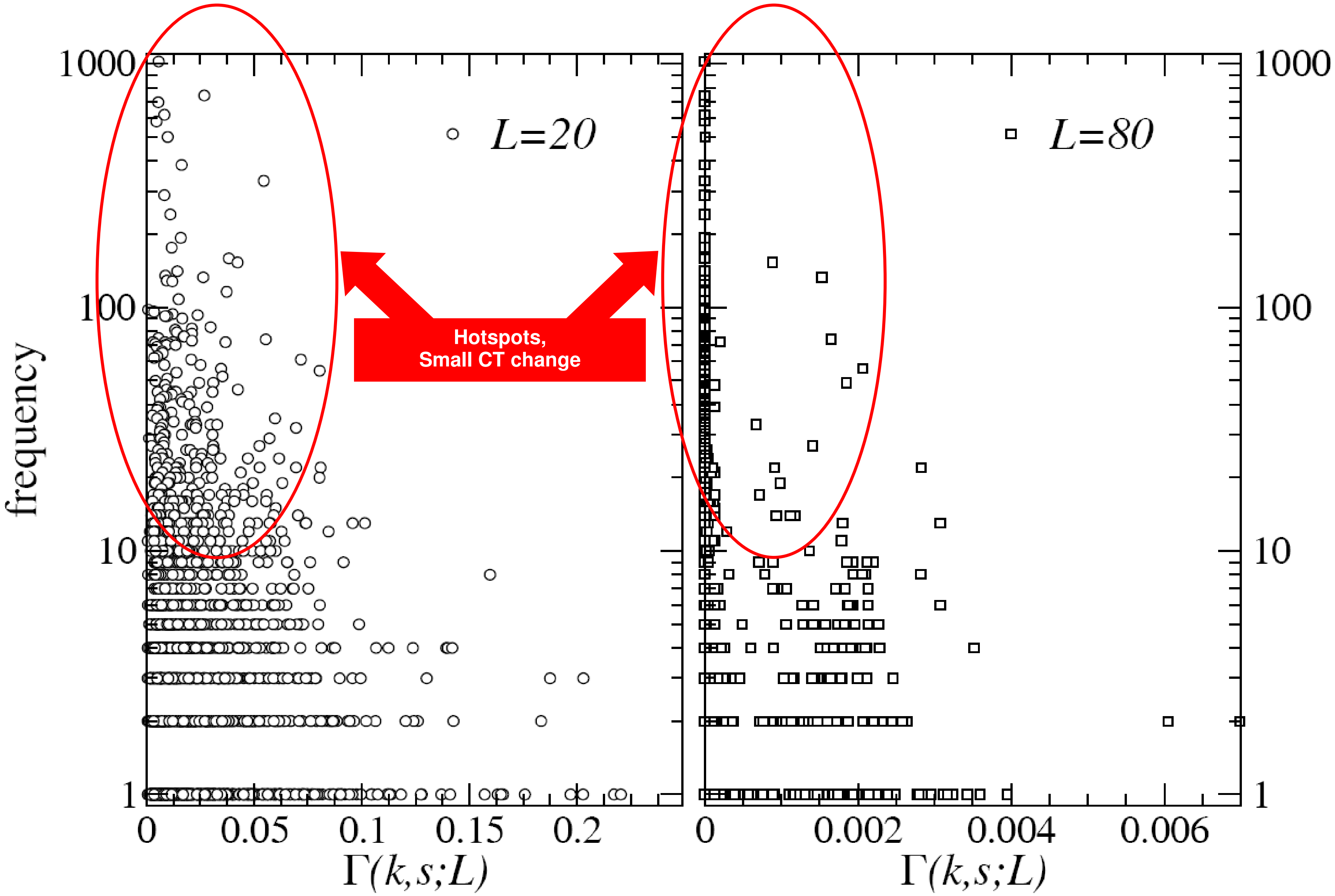
T(E) in a section of p53 for native and mutant sequences:




T
E
(ΔT)²
E

Squared difference in charge transmission for mutant sequences:





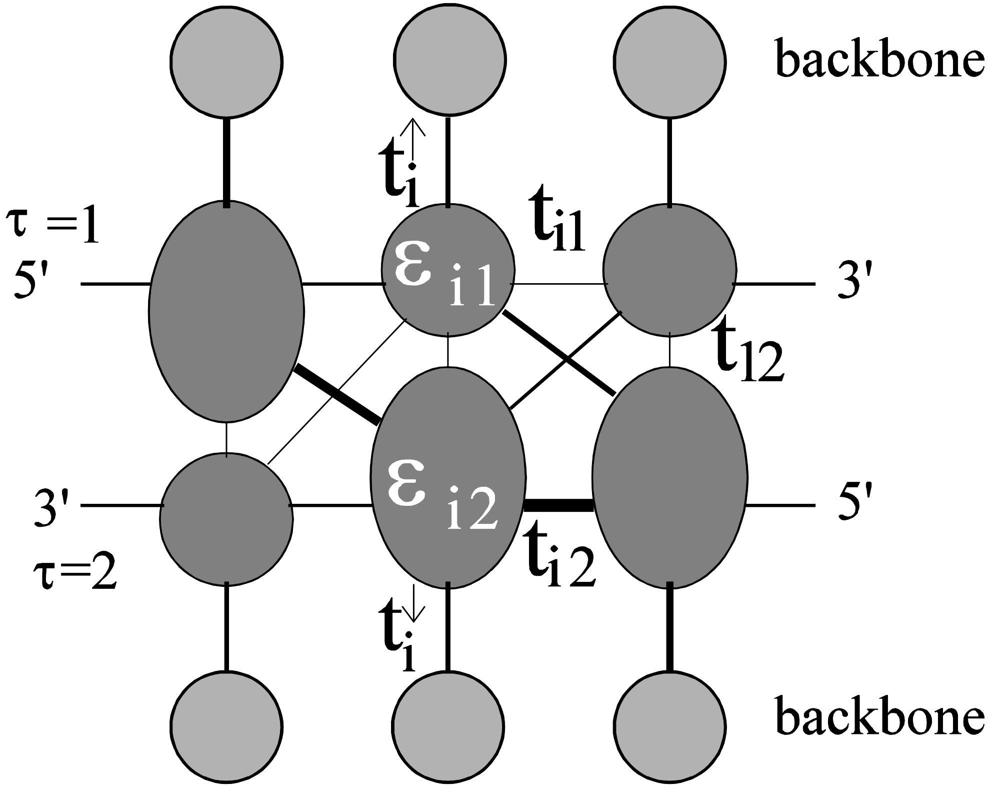
Hotspots,
Small CT change

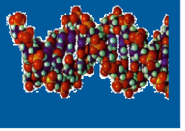
Diagonal model

Ladder model without diagonal hopping includes an unphysically large term for hopping across the hydrogen bond.

Better to explicitly include diagonal terms.

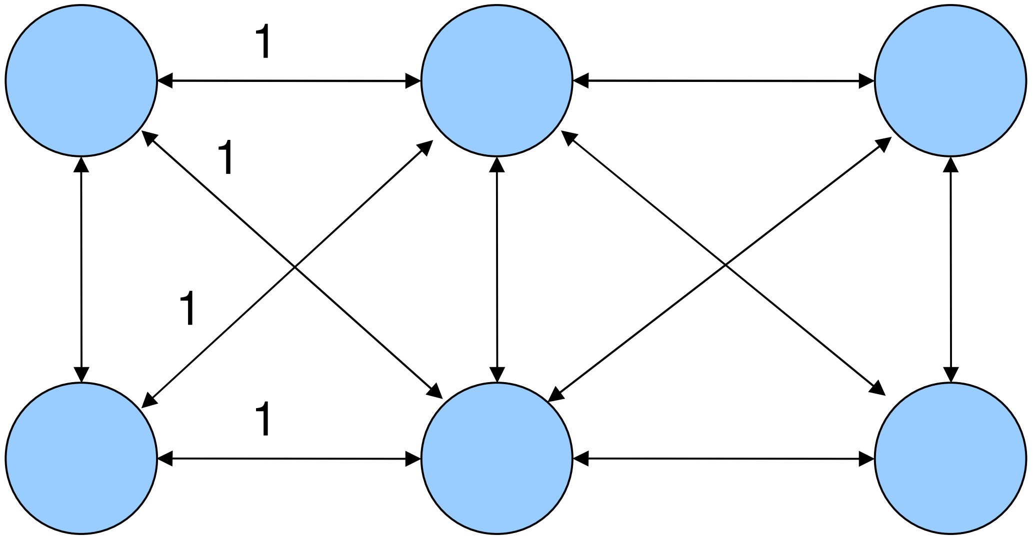
We use: 0.1 eV purine-purine
0.01 eV purine-pyrimidine
0.001 eV pyr-pyr





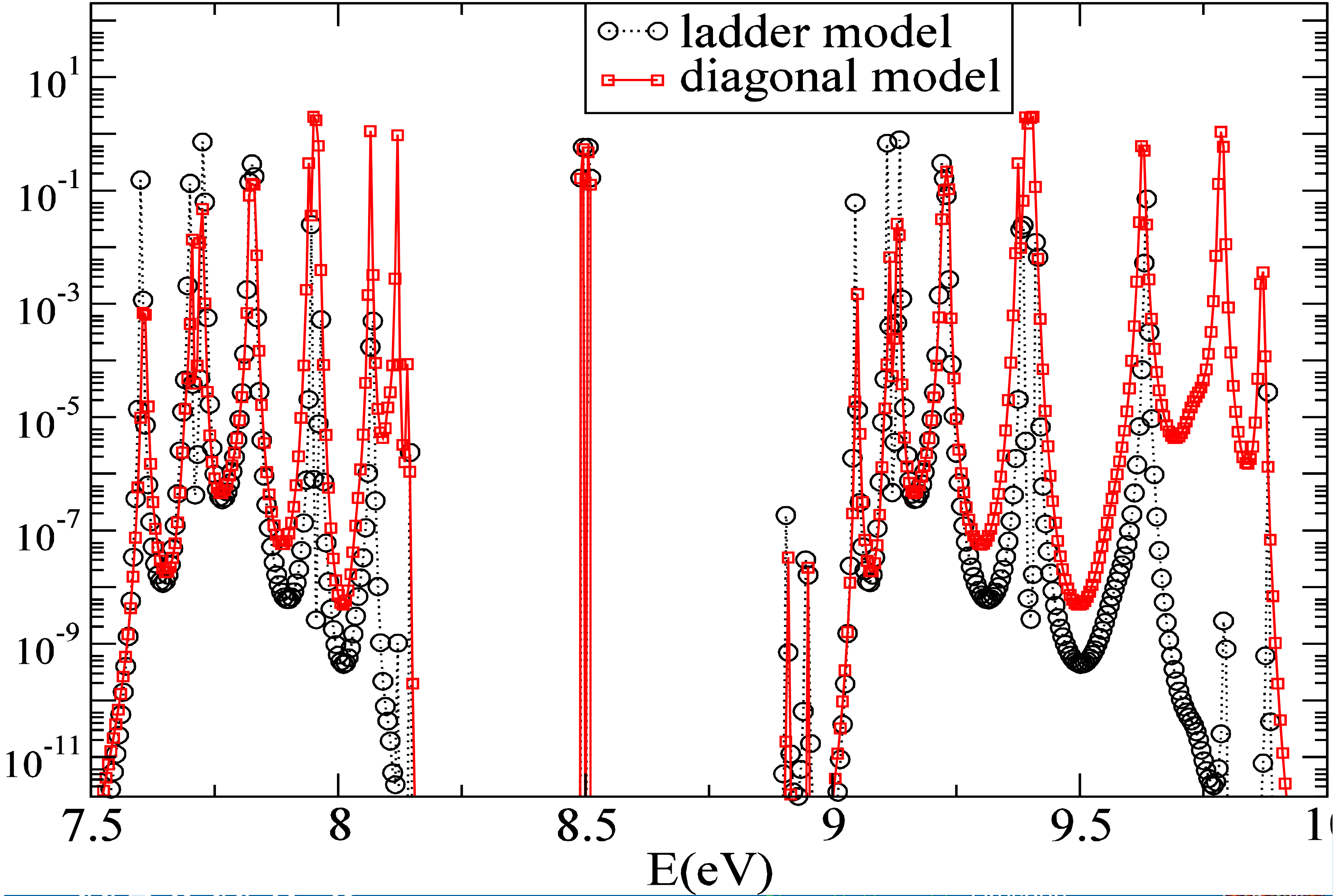
An interesting way to break the transfer-matrix method

Set along-strand and diagonal elements all equal:

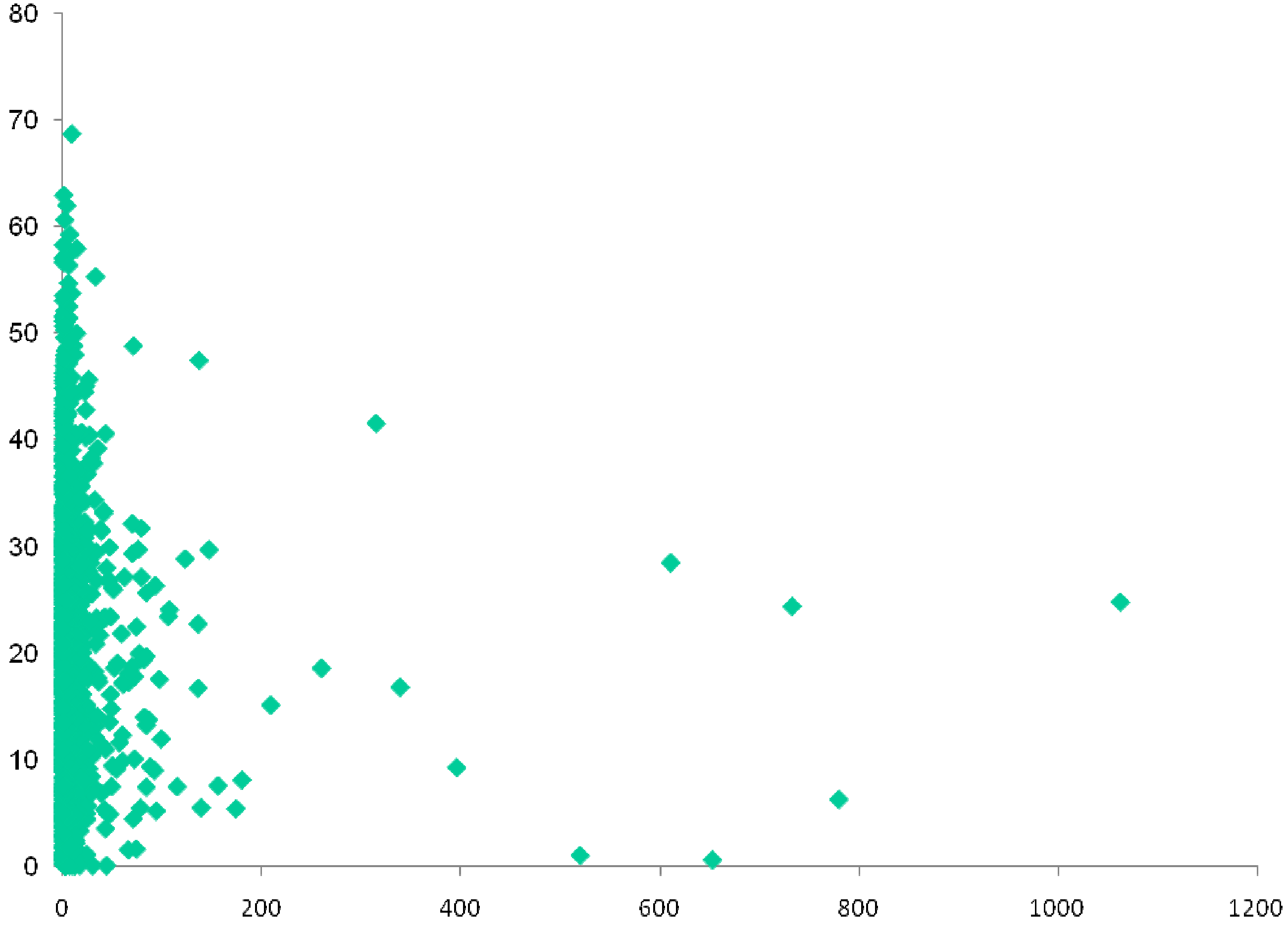


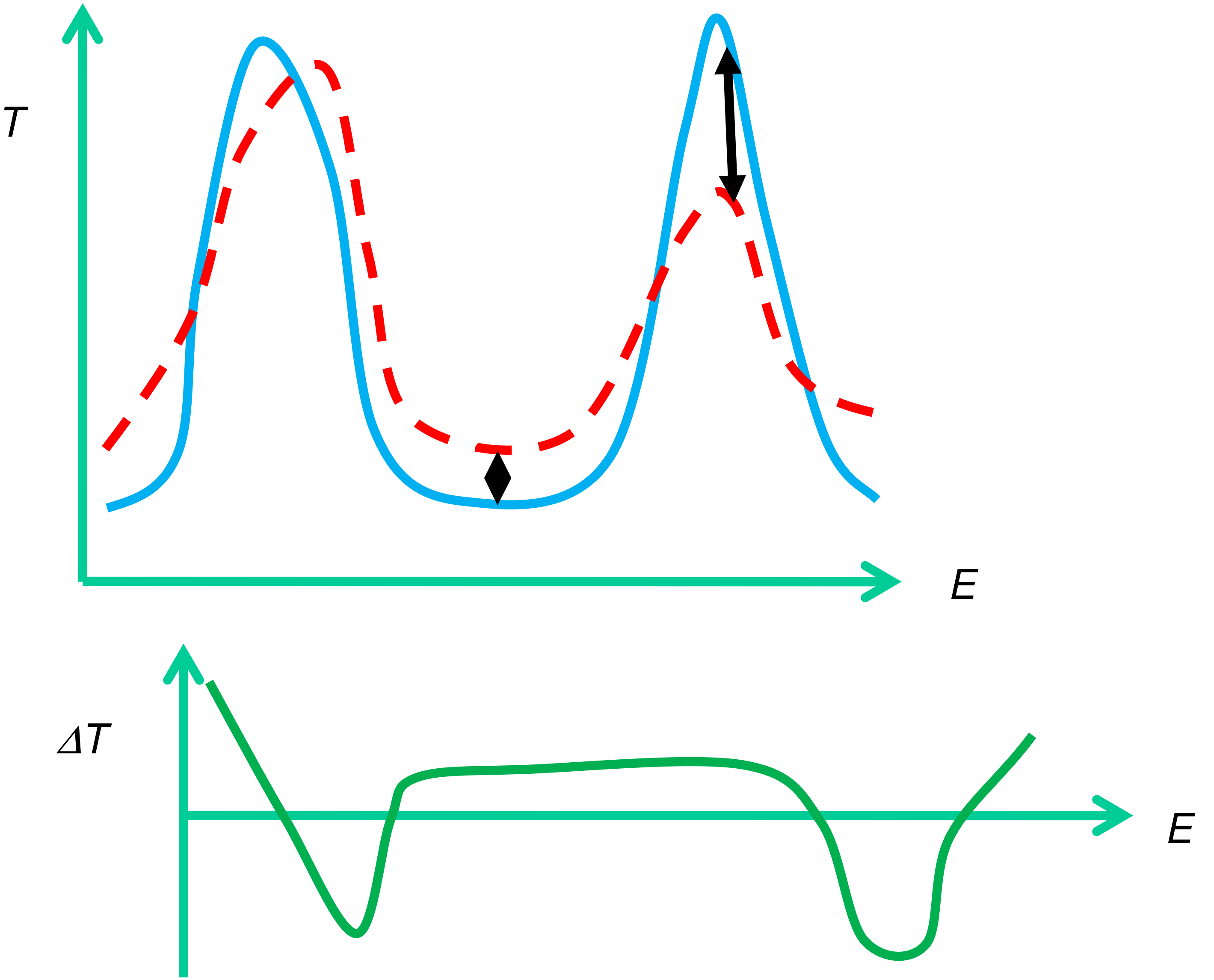
Transfer matrix involves inverse of a singular matrix:

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

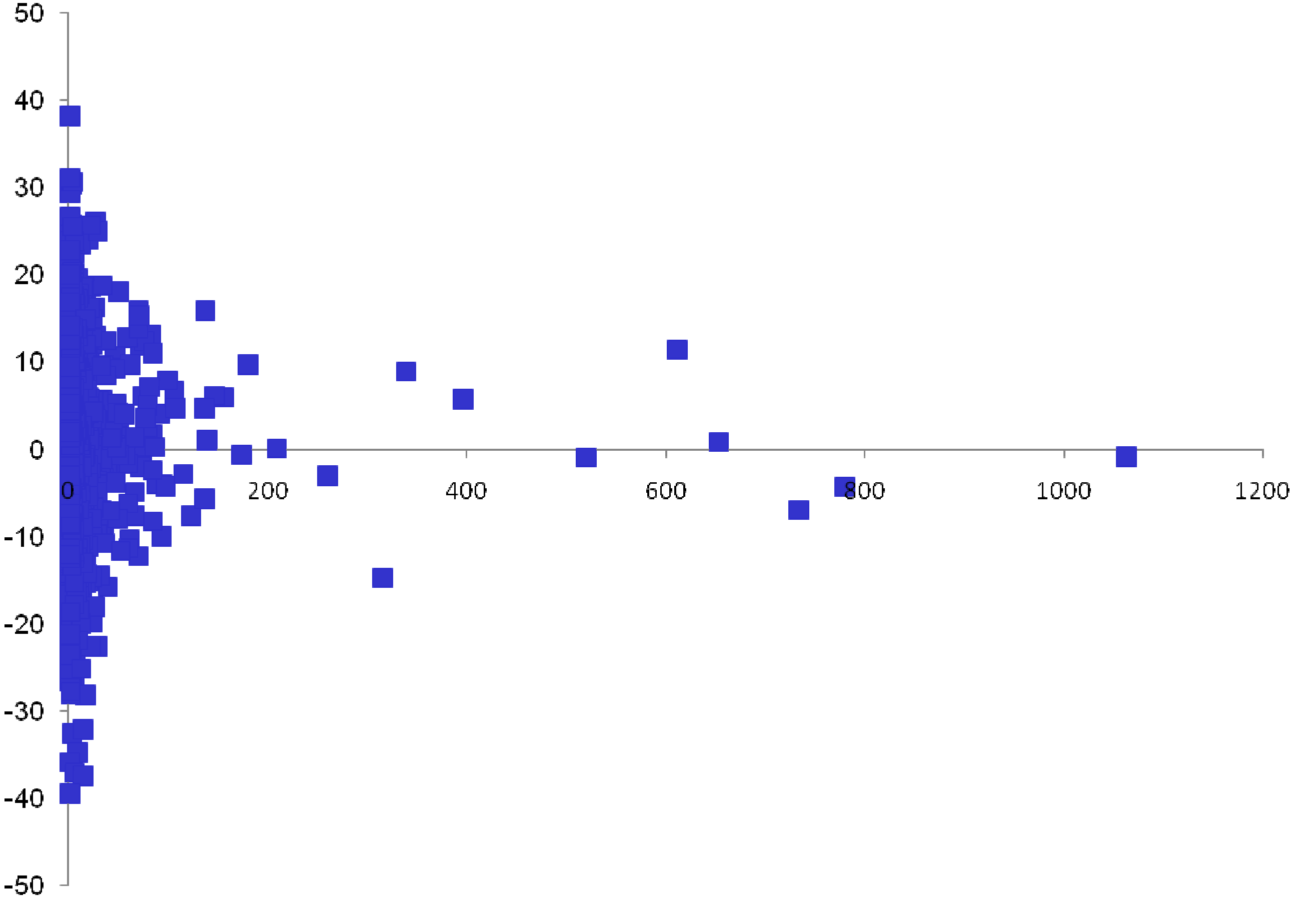

ladder model
diagonal model
10^{1}
10^{-1}
10^{-3}
10^{-5}
10^{-7}
10^{-9}
10^{-11}
7.5
8
8.5
9
9.5
1
E(eV)

Mutation frequency versus change-squared measure:



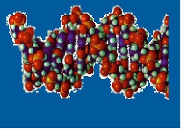

T
E
ΔT
E

Mutation frequency versus linear change measure:



Really?

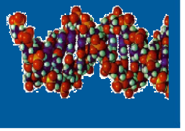
- The eye is notoriously good at spotting non-existent patterns.
- Check: variance and kurtosis of distribution with and without weighting by frequency.
- Unweighted : var 85.9027, kurt -1.92281
- Weighted : var 66.686, kurt 4.07172



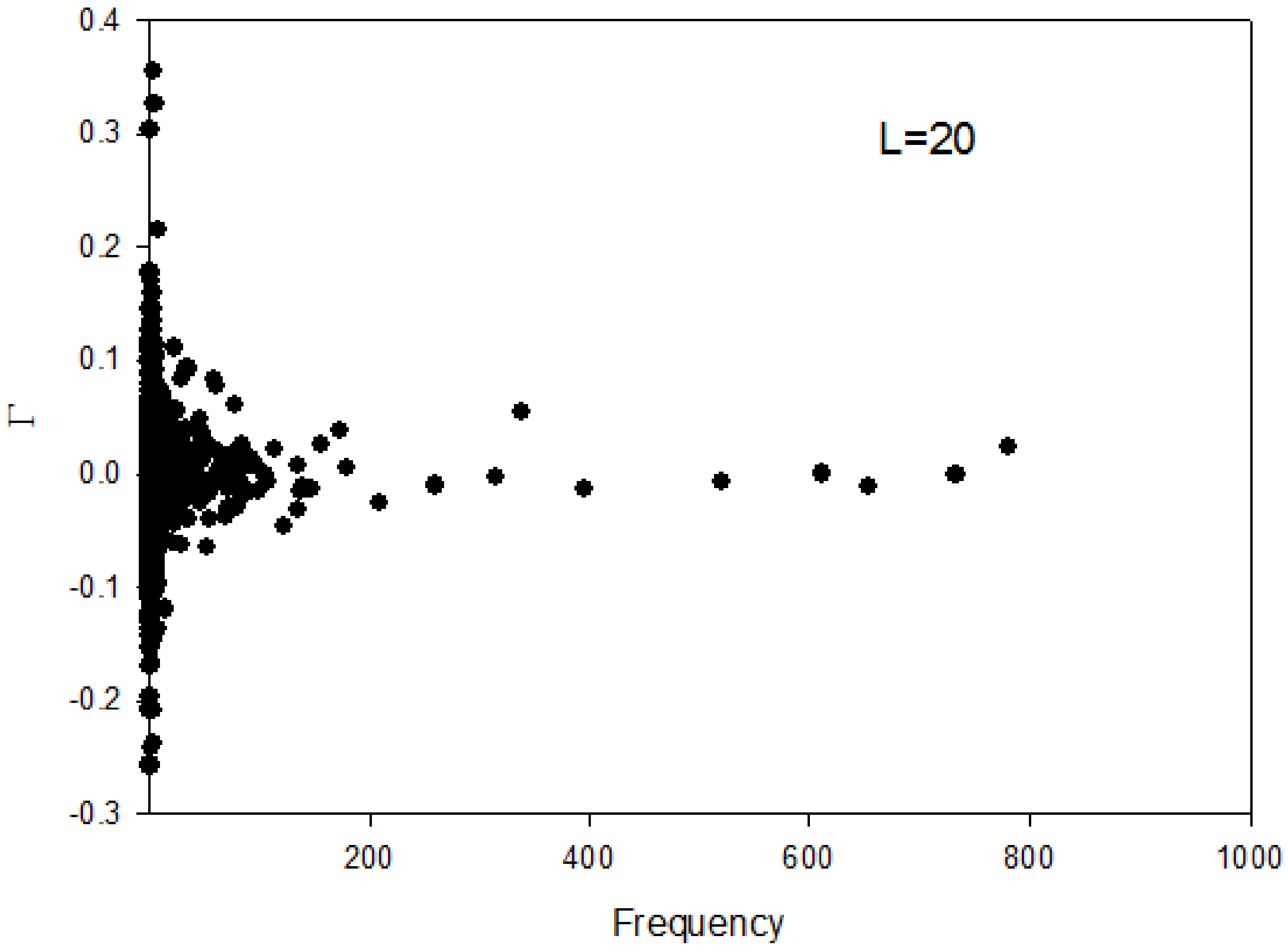
- Array size: 2002
- First pass: unweighted mean 0.530162 from 2002 entries;
- First pass: weighted mean 0.758528 from 21361 total weight.

- Estimated standard deviation for higher moments:
- For unweighted data: stdev of skew estimate root 15/N = 0.0865593
- For unweighted data: stdev of kurt estimate root 96/N = 0.21898
- For weighted data: stdev of skew estimate root 15/N = 0.0264993
- For weighted data: stdev of kurt estimate root 96/N = 0.0670386

- Moments: weighting: mean: var: stdev: skew: kurt
- Moments: unweighted : 0.530162 85.9027 9.26837 -0.108417 -1.92281
- Moments: weighted : 0.758528 66.686 8.16615 -0.0386236 4.07172

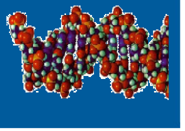


Mutation frequency versus linear change measure in 1-D model:



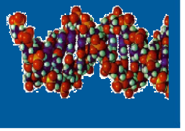
Intepretation 1:

- Carcinogenesis is intimately related to DNA charge transport and we are modelling it.
- Problems:
  - Unrealistic model parameters
  - Model of unreal situation!



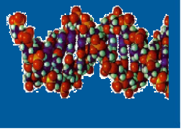
Intepretation 2:

- Whether a mutation appears in a genetic disease depends on: the likelihood of DNA ***damage***, the likelihood of damage ***detection and repair***, and the effect of the mutation: coding, ***regulatory***.
- These depend on ***sequence*** and our models are probing the properties of the sequence.



Outlook

- Search for more genes with good statistics on mutation frequency and disease
- Improve model parameters: parameters for bioinformatics may differ from those for physical CT
- Medically useful predictions: identify high-risk mutations for screening.
- Understanding: what is CT model telling us about sequences?

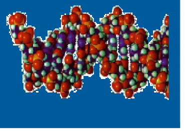


References

C.T. Shih, S. Roche, R. Roemer, **Phys. Rev. Lett.** 100, 018105 (2008)

S. Roche, D. Bicout, E. Macia, E. Kats, **Phys. Rev. Lett.** 92, 109901 (2004)
S. Roche and E. Macia, **Modern Physics Letters B** 18, 847 (2004)

+ see Stephan's poster for more details!



Acknowledgements

- Leverhulme trust for funding

- DNATEC09 for invitation and travel

- Your attention.