

# The Role of Electrostatics in Discrimination of Adenine and Guanine by Proteins

Dakshanamurthy Sivanesan<sup>1</sup>

siva@is.aist-nara.ac.jp

Gautam Basu<sup>1</sup>

gautam@is.aist-nara.ac.jp

Nobuhiro Go<sup>1,2</sup>

ngo@is.aist-nara.ac.jp

<sup>1</sup> Graduate School of Information Sciences, Nara Institute of Science & Technology, Nara 630-01, Japan

<sup>2</sup> CCSE, Japan Atomic Energy Research Institute, 8-1 Umemidai, Kizu-cho, Sourakugun, Kyoto 619-0215, Japan

**Keywords:** adenine, guanine, non-cognate binding, binding energy, electrostatic energy

## 1 Introduction

Discrimination of Adenine (A) and Guanine (G) (see Fig. 1) binding to proteins is crucial for proper functioning of biological systems. Therefore, it is important to clarify the recognition mechanisms that lead to their binding or discrimination. Electrostatic interaction is known to be an important contributor to the protein-ligand interactions [2]. Although previous studies have focussed on local protein-ligand atomistic interactions [1, 3, 4] (H-bonding or spatial arrangement of residues) in A and G binding proteins, the role of electrostatics have not been looked at carefully. In this paper we present our work on electrostatic contribution to A and G binding to proteins. The focus of our analysis is two-fold: 1) assessing the relative contribution of electrostatics in the native G or A binding, and, 2) the role of electrostatics in discriminating between A and G by proteins.

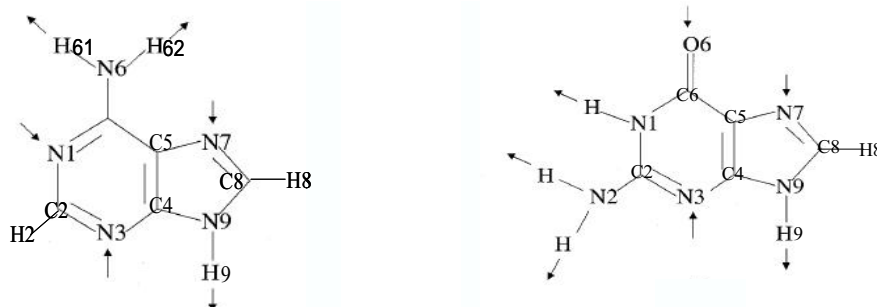


Figure 1: a: The adenine moiety.

b: The guanine moiety.

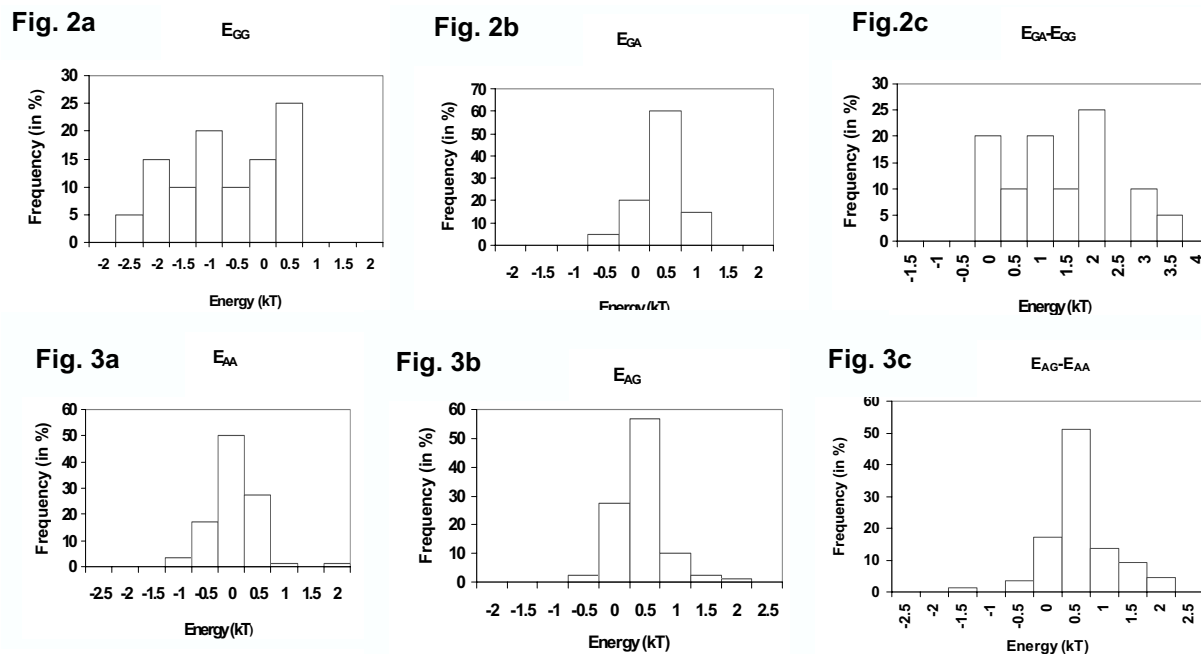
## 2 Method and Results

The non-redundant data set used in this work consists of 88 A-binding proteins and 20 G-binding proteins, identical to that used by Nobeli *et al.* [4]. Electrostatic potentials, associated with individual structures, in absence of any bound ligand (and water), were calculated numerically by solving the linearized Poisson-Boltzmann equation (DelPhi from Biosym Inc.). The following parameters were used:  $\epsilon_{in}$ : 2,  $\epsilon_{out}$ : 80, solvent probe radius: 1.4Å grid resolution: 0.6Å solute extent: 70.0. Partial charges were taken from the AMBER parameter set. The electrostatic energies of binding were calculated using

$$E_i = \sum_{j=1}^N \phi_{ij} q_j \quad (1)$$

where  $\phi_{ij}$  represents electrostatic potential arising from the *i*-th protein structure at the *j*-th A (or G) site and  $q_j$  represents the partial charge at the *j*-th site on A (G).

The central result is summarized in Figs. 2 & 3 as histograms of energies defined in Eq. (1). Both cognate ( $E_{AA}$  and  $E_{GG}$ ) and non-cognate ( $E_{GA}$  and  $E_{AG}$ ) energies were calculated, where the subscripts AA, GG, GA and AG represent the energy of A in an A-binding site, the energy of G in a G-binding site, the energy of A in a G-binding site and the energy of G in an A-binding site, respectively.



### 3 Discussion

The distribution of cognate electrostatic binding energies for both G ( $E_{GG}$ , Fig. 2a) and A ( $E_{AA}$ , Fig. 3a), show a clear preference for negative values, indicating that for the majority of the cases ( $\approx 75\%$ ), electrostatics play a favorable role in the binding energetics. However, the mean value of the negative trend is only nominal ( $\approx -0.5 k_B T$  for A and  $\approx -1.0 k_B T$  for G) with a small number of cases where the energy is slightly positive or near zero. Therefore along with a favorable electrostatic picture for binding, the role of other non-bonded interactions cannot be ruled out. The distribution of non-cognate electrostatic binding energies,  $E_{GA}$  (Fig. 2b) and  $E_{AG}$  (Fig. 3b), on the other hand, show a clear preference for positive values, indicating that for the majority of the cases, electrostatics disfavors non-cognate binding. This positive trend is further augmented when the  $E_{GA} - E_{GG}$  (Fig. 2c) and  $E_{AG} - E_{AA}$  (Fig. 3c), representing the electrostatic cost of replacing G by A in a G-binding protein and the electrostatic cost of replacing A by G in an A-binding protein respectively, distributions are considered. Our results clearly demonstrate that electrostatics play an important role not only in cognate binding of A and G but also in discriminating between cognate and non-cognate binding of A and G. A previous study [4] on the same set of proteins identified a number of “fuzzy” structural features that were identified as potential contributors to the differential binding. In contrast, our work, which is based on energetics, rather than structure, indicates an important role for electrostatics.

### References

- [1] Cappello, V., Tramontano, A., and Koch, U., Classification of proteins based on the properties of the ligand-binding site: The case of adenine-binding proteins, *Proteins*, 47:106–115, 2002.
- [2] Honig, B. and Nicholls, A., Classical electrostatics in biology and chemistry, *Science*, 8:1144–1149, 1995.
- [3] Kobayashi, N. and Go, N., ATP binding proteins with different folds share a common ATP-binding structural motif, *Nature Str. Biol.*, 4:6–7, 1997.
- [4] Nobeli, I., Laskowski, R.A., Valdar, W.S.J., and Thornton J.M., On the molecular discrimination between adenine and guanine by proteins, *Nucleic Acids Res.*, 29(21):4294–4309, 2001.