Supporting Information

Protein-Dendron Conjugates for DNA Binding: Understanding the Effect of the Protein Core on Multivalency

Giovanni M. Pavan,[a] Mauri A. Kostiainen,[b] Andrea Danani*[a]

a: University for Applied Science of Southern Switzerland (SUPSI) – Laboratory of Applied Mathematics and Physics (LamFI), Centro Galleria 2, Manno, CH-6928 (Switzerland)

giovanni.pavan@supsi.ch

b: Aalto University School of Science and Technology – Department of Applied Physics,. FI-00076 AALTO, Finland

TABLE OF CONTENTS

Construction of the Dendron-Protein Conjugate Models

The construction of HFBI-dendron conjugates were described in our previous work,[S1] here the key steps are reported for comparison with the creation of the BSA-dendron conjugates, which are created *de novo* in this study. The models for bovine serum albumin (BSA) and a class I hydrophobin protein (HFBI) were taken from the SWISS-MODEL repository (code: P02769) and from the Protein Data Bank (file: 2FZ6.pdb, chain A only) respectively. We created the molecular models for HFBI- and BSA- dendron conjugates by grafting the first and second generation UV-degradable dendrons (**pllG1** and **pllG2**) to HFBI and BSA in the same way as presented in our previous experimental work.[S2,S3] In the BSA-based conjugates the dendrons were attached to the free Cysteine-34 present in the BSA amino acid chain. On the other hand, the N terminus of HFBI was modified by site-directed mutagenesis to add a short peptide with a free Cysteine residue necessary to graft **pllG1** and **pllG2** to the protein. The resulting modified HFBI had the following sequence:

SCPATTTGSS PGPSNGNGNV CPPGLFSNPQ CCATQVLGLI GLDCKVPSQN VYDGTDFRNV CAKTGAQPLC CVAPVAGQAL LCQTAVGA

where the added N-terminus is underlined.

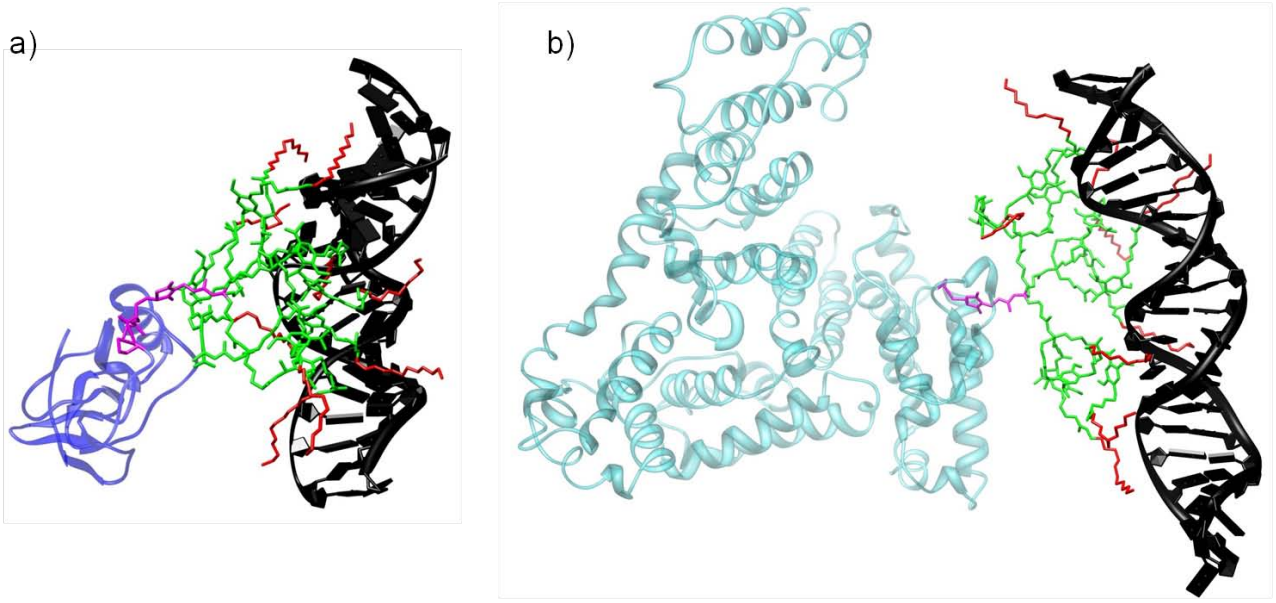

Figure S1 Molecular models of HFBI-**pllG2** (a) and BSA-**pllG2** (b) in complex with DNA used for simulations. The HFBI and BSA proteins are represented as purple and cyan ribbons respectively. Within the Dendron, the spermine (SPM) residues are colored in red, the repetitive (REP) in green and the central (CEN) unit that bridges dendron and protein together in purple.

Simulation Procedure

All calculations were conducted using the AMBER 10 suite of programs.[S4] For the DNA model we used a 21 base-pair double-stranded *β*-DNA,[S5,S1] which was generated with the *nucgen* module of AMBER 10. The force field parameters for the residue types of the dendrons were obtained by *ab initio* techniques[S6] with the *antechamber* module of AMBER 10.

The HFBI-dendrons were created and studied in our previous work.[S1] The BSA-dendrons conjugates were solvated in a TIP3P[S7] water box extending 12 Å from the solute and a suitable number of counterions were added using the *leap* module of AMBER 10 in order to guarantee the system neutrality. Particularly, the salt ions were added in the periodic systems with the standard *addIons* utility of *leap* – Na^+ and Cl^- ions were placed onto a shell around the solute using a Coulombic potential on a grid – water molecule was replaced with the ion if eventual superposition occurred. Systems were minimized and equilibrated for other 10 ns in NPT conditions to obtain a reliable configuration for the BSA–dendron conjugates in solution. From the corresponding equilibrated systems, water molecules and counterions were removed, and the DNA model was placed with its major groove in the proximity of the BSA-**pllG1** and BSA-**pllG2** conjugates following a well validated procedure.[S8,S5,S6] Thus, two final molecular complexes were obtained. The resulting structures were again solvated in a water box extending 12 Å from the solute and the proper amount of Na^+ and Cl^- ions was added in two steps – firstly to guarantee the system neutrality and then to reproduce the relevant salt concentrations of 9.4 and 150 mM NaCl in the

solution (the number of ions was calculated with respect to the volumes of the water boxes). Table S1 summarizes the main characteristics of the simulated systems.

Table S1 Details of molecular systems simulated in this work.

Complex	[NaCl][a] (mM)	Conjugate charge[b]	DNA charge[c]	Water box volume (Å^3)	Number of Na^+ and Cl^- atoms[d] in the system	Number of water molecules in the system	Total number of atom in the system
BSA-**pllG1**	9.4	+0	-40	1100470	46	33261	110317
BSA -**pllG2**	9.4	+18	-40	1571377	48	46743	151357
BSA -**pllG1**	150	+0	-40	1098526	196	33107	110005
BSA -**pllG2**	150	+18	-40	1566651	262	46509	150886
HFBI-**pllG**[e]	9.4	+9	-40	530306	35	16668	52849
HFBI-**pllG2**[e]	9.4	+27	-40	634944	19	19945	63271
HFBI-**pllG1**[e]	150	+9	-40	529499	95	16610	52733
HFBI-**pllG2**[e]	150	+27	-40	633612	115	19848	63079

[a] Experimental ionic concentration in solution. [b] The protein charge for BSA and HFBI was assumed to be -9 and 0 respectively, thus the resulting conjugate charge is the sum of the one of the protein and the one of dendrons (**pllG1**: +9 and **pllG2**: +27). [c] The 21 base-pair DNA has an overall charge of -40 because the terminal nucleotides do not carry a charge in the model. [d] The total amount of counterions is the sum of the Na^+ and Cl^- atoms required for system neutralization and to reproduce the experimental ionic concentration reported in the 2nd column. [e] Data for HFBI-**pllG1** and HFBI-**pllG2** conjugates are taken from our previous work.[S1]

The systems were initially minimized and then equilibrated at 300 K by 50 ps MD simulation in NVT conditions. A density equilibration MD run (other 50 ps) followed in NPT conditions. The production MD lasted for 10 ns in NPT periodic boundary condition at 300 K and 1 atm, using a time step of 2 femtoseconds, the Langevin thermostat and a cutoff of 10 Å. The particle mesh Ewald[S9] (PME) approach was used for long-range electrostatic effects, and all bonds involving hydrogen atoms were constrained using the SHAKE algorithm.[S10]

All of the molecular dynamics simulations were conducted using the *sander.MPI* and *pmemd* module of AMBER 10 and the *parm99* all-atom force field by Cornell et al.[S11] working in parallel on 128 processors of the Cray XT5 calculation cluster of the CSCS Swiss National Supercomputer Centre of Manno (Switzerland).

Free energy of binding

The energetic analysis for each molecular system was performed for 200 unbound protein-dendron conjugates and DNA snapshots taken from the equilibrated phase of a single 10 ns MD trajectory. The Root Mean Square Deviation (RMSD) data were obtained from the molecular dynamics trajectories in order to verify that all of the systems converged to the equilibrium with good stability. The binding free energy for each ligand/receptor systems, ΔG_{bind}, was calculated with the Molecular Mechanics/Poisson-Boltzmann Surface Area method (MM-PBSA)[S12] as:

$$\Delta G_{bind} = \Delta H_{bind} - T\Delta S_{bind} \quad (1)$$

$$\Delta H_{bind} = \Delta E_{gas} + \Delta G_{sol} \quad (2)$$

The average enthalpic contribution (ΔH_{bind}) were calculated as the sum of the gas-phase *in vacuo* non-bond energies ($\Delta E_{gas} = \Delta E_{ele} + \Delta E_{vdw}$) and the solvation free energies ($\Delta G_{solv} = \Delta G_{PB} + \Delta G_{NP}$).[S13]

The polar term of ΔG_{solv} (ΔG_{PB}) was calculated with the Poisson-Boltzmann (PB) approach,[S14] and the non-polar contribution to the solvation energy was calculated as $\Delta G_{NP} = \gamma$ (SASA) + β, where $\gamma = 0.00542$ kcal/Å^2, $\beta = 0.92$ kcal/mol, and SASA is the solvent-accessible surface calculated with the MSMS program.[S15] Finally, the normal-mode analysis approach was used to estimate the entropic contributions ($-T\Delta S$).[S16]

All of the enthalpic, entropic and total free energies of binding were then normalized per-charge (ΔH_{bind}, $-T\Delta S_{bind}$ and ΔG_{bind}, Table 1 in the paper) in order to allow a direct comparison between first and second generation protein-dendron conjugates. Similarly, to explore the role played by the protein core in the multivalent recognition between the conjugates and the DNA we calculated the averaged interaction energy (ΔE_{intra}) between each SPM ligand of the dendron and the HFBI and BSA protein cores (Table 2 in the paper). In fact, the charged residues of the protein can subtract few spermines from the bind with DNA, similarly to what is done by the ions in solution at 150 mM. The average SPM-protein intramolecular interaction energies (ΔE_{intra}) were calculated as the

sum of the gas-phase *in vacuo* non-bond energies ($\Delta E_{gas} = \Delta E_{ele} + \Delta E_{vdw}$) and corrected for solvation with the MM-PBSA approach.[S11]

References

S1 M. A. Kostiainen, J. Kotimaa, M.-L. Laukkanen, G. M. Pavan, *Chem. Eur. J.* **2010**, *16*, 6912-6918.

S2 M. A. Kostiainen, G. R. Szilvay, J. Lehtinen, D. K. Smith, M. B. Linder, A. Urtti, O. Ikkala, *ACS Nano* **2007**, *1*, 103-113.

S3 M. A. Kostiainen, G. R. Szilvay, D. K. Smith, M. B. Linder, O. Ikkala, *Angew. Chem. Int. Ed.* **2006**, *45*, 3538-3542.

S4 Case, D.A.; Darden, T.A.; Cheatham, T.E. *et al.* AMBER 10, University of California, San Francisco, CA USA, **2008**.

S5 (a) G. M. Pavan, A. Danani, S. Pricl, D. K. Smith, *J. Am. Chem. Soc.*, **2009**, *131*, 9686-9694. (b) S. P. Jones, G. M. Pavan, A. Danani, S. Pricl, D. K. Smith, *Chem. Eur. J.*, **2010**, *16*, 4519-4532.

S6 (a) G. M. Pavan, M. A. Kostiainen, A. Danani, *J. Phys. Chem. B.*, **2010**, *114*, 5686–5693. (b) G. Doni, M.A. Kostiainen, A. Danani, G.M. Pavan, *Nano Lett.*, **2011**, *11*, 723-728.

S7 Jorgensen W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–35.

S8 (a) G. M. Pavan, L. Albertazzi, A. Danani, *J Phys Chem B*, 2010, **114**, 2667-2675. (b) G. M. Pavan, M. A. Mintzer, E. E. Simanek, O. M. Merkel, T. Kissel, A. Danani, *Biomacromolecules*, 2010, **11**, 721–730. (c) L.B. Jensen, K. Mortensen, G.M. Pavan, M.R. Kasimova, D.K. Jensen, V. Gadzhyeva, H.M. Nielsen, C. Foged, *Biomacromolecules*, **2010**, *11*, 3571-3577.

S9 T. Darden, D. York, L. Pedersen, *J. Chem. Phys.*, **1998**, *98*, 10089-10092.

S10 a) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys*. **1977**, *23*, 327. b) Krautler, V.; van Gunsteren, W. F.; Hanenberger, P. H. *J. Comput. Chem*. **2001**, *5*, 501.

S11 Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.

S12 Srinivasan, J.; Cheatham, T. E.Cieplak, P.; Kollman, P. A.; Case, D. A. *J. Am. Chem. Soc.* **1998**, *120*, 9401-09.

S13 Andricioaei, I.; Karplus, M. J. *Chem. Phys.* **2001**, *115*, 6289-92.

S14 Jayaram, B.; Sprous, D.; Beveridge, D. L.; *J. Phys. Chem.* **1998**, *102*, 9571-9576.

S15 Sitkoff, D.; Sharp, K. A.; Honig, B. *J. Phys. Chem.* **1994**, *98*, 1978-1988.

S16 Sanner, M. F.; Olson, A. J.; Spehner, J. C. *Biopolymers*. **1996**, *38*, 305-20.