SUPPORTING INFORMATION

Pharmacologic efficacy of PU.1 inhibition by heterocyclic dications: a mechanistic analysis

March 10, 2016

Dominique C. Stephens,[1] Hye Mi Kim,[1] Arvind Kumar,[1] Abdelbasset A. Farahat,[1,a] David W. Boykin,[1] and Gregory M. K. Poon[1,2,*]

[1] Department of Chemistry, Georgia State University, Atlanta, GA 30303

[2] Center for Diagnostics and Therapeutics, Georgia State University, Atlanta, GA 30303

[a] Present address: Department of Pharmaceutical Organic Chemistry, Faculty of Pharmacy, Mansoura University, Mansoura 35516, Egypt

* To whom correspondence should be addressed at: P.O. Box 3965, Atlanta, GA 30302-3965. Email: gpoon@gsu.edu. Tel. (404) 413-5491.

SUPPLEMENT INDEX **PAGE**

SUPPLEMENTAL METHODS

Non-denaturing polyacrylamide electrophoresis. Mixtures containing 1 μg of PU.1 and various amounts of DB270 in a total volume of 12 μL (containing 10% glycerol) were resolved on a 6% polyacrylamide gel running in standard Tris-glycine (25:250 mM) buffer at ambient temperature. No denaturant was present in the samples, gel, or running buffer. The electric field polarity was reversed (anode in the lower buffer) relative to the conventional configuration to drive cationic species (such as PU.1, pI 10.6) into the gel matrix. After electrophoresis, the gel was stained with Coomassie Blue and digitized with a Typhoon 9400 imager (GE) at 50 μm resolution.

Data analysis — mechanistic titration models. Models for describing titration data at thermodynamic equilibrium were formulated generally as prescribed by Wells (1) with additional details as described (2). In all cases, the independent variable was the *total* (not free or unbound) titrant concentration as taken, to account for depleting effects of the titrate(s) which were present at concentrations close to or above the expected dissociation constants (10^{-9} to 10^{-8} M). The general strategy was to formulate the model as a binding polynomial in the concentration of a specific bound or unbound species of interest, then solve for the polynomial's roots numerically. The concentration of other species can then be computed directly from related equilibrium expressions and/or equations of state (i.e., mass balance). Mapping of concentrations to anisotropies was according to Eq. (1) in the main text.

Notation. For ease of expression, we adopt a three-digit numeric code *xyz* for referring to the various species of interest, where *x*, *y*, and *z* refer to stoichiometric equivalents of DB270, DNA, and PU.1, receptively. Thus for single species: DB270 ≡ 100; DNA ≡ 010; PU.1 ≡ 001;

for complexes: DB270:DNA ≡ 110, DNA:PU.1 ≡ 011, DB270/PU.1 ≡ 101, etc. (In the main text, we occasionally depart from this convention e.g., the Schemes, when the meaning is clear.)

1:1 direct binding. This model describes the binding of DB270 (100) to oligomeric DNA duplexes (010). The stoichiometry has been previously determined by biosensor-plasmon surface resonance (3). The solution (1), which is quadratic in [110], is:

$$\begin{aligned} 0 &= \varphi_0 + \varphi_1[110] + \varphi_2[110]^2 \\ &\begin{cases} \varphi_0 = [100]_t[010]_t \\ \varphi_1 = -K_{110} - [100]_t - [010]_t \\ \varphi_2 = 1 \end{cases}, \end{aligned} \tag{S1}$$

where $K_{100} = \dfrac{[100][010]}{[110]}$. While Eq. (S1) may be solved analytically via the quadratic formula, it is more advantageous to compute the roots numerically (4).

2:1 direct binding (Scheme 1). We previously reported the formation of a 2:1 complex by the ETS domain of PU.1 to a single DNA binding site in an isothermal titration calorimetric (ITC) study of PU.1/DNA binding (5) in which, among other observations, an oligomeric DNA site functionally behaved as a bivalent receptor that bound two equivalents of PU.1 with strong negative cooperativity. The concentration requirements of ITC measurements (>10^{-5} M titrate) prohibited direct determination of the strong binding affinities. As embodied by **Scheme 1** in the main text, the two stepwise dissociation constants describing this interaction are:

$$\begin{aligned} K_{011} &= \frac{[010][001]}{[011]} \\ K_{012} &= \frac{[011][001]}{[012]} = \omega K_{011} \end{aligned}, \tag{S2}$$

where K_{012} may be alternatively expressed by the cooperativity parameter ω. In direct titrations of Cy3-[5']AGC by PU.1, since the observed anisotropy change represented the summed contributions of the two complexes, the most efficient approach is to build the binding polynomial in terms of the unbound titrant, 001. The solution, which is cubic in [001], is:

$$0 = \varphi_0 + \varphi_1[001] + \varphi_2[001]^2 + \varphi_3[001]^3$$
$$\begin{cases} \varphi_0 = K_{011}K_{012}[001]_t \\ \varphi_1 = K_{011}K_{012} - K_{012}[001]_t + K_{012}[010]_t \\ \varphi_2 = K_{012} - [001]_t + 2[010]_t \\ \varphi_3 = 1 \end{cases} . \quad \text{(S3)}$$

As in the case of 1:1 binding, [001] was solved numerically from Eq. (S3), rather than analytically via the cubic formula, to avoid failure due to loss of significance. With [001] in hand, [010], [011], and [012] were computed from Eq. (S2) and the corresponding equations of state.

In the limit of no binding of the second equivalent of PU.1 (i.e., $K_{012} \to \infty$), such as observed with Cy5-[5']AGC as probe, Eq. (S3) simplifies to a quadratic:

$$0 = K_{012}(\varphi'_0 + \varphi'_1[001] + \varphi'_2[001]^2) + [001]^3 \equiv P_{2:1}$$
$$\lim_{K_{012}\to\infty} P_{2:1} = \varphi'_0 + \varphi'_1[001] + \varphi'_2[001]^2 \quad \text{(S4)}$$
$$\begin{cases} \varphi'_0 = K_{011}[001]_t \\ \varphi'_1 = K_{011} - [001]_t + [010]_t \\ \varphi'_2 = 1 \end{cases}$$

Eq. (S4) is the analog of the 1:1 model i.e., Eq. (S1) in terms of [001]. For data fitting, this limiting situation is handled numerically by setting log K_{012} = -3.

Single-site competition. Titration of (Cy3-)labeled DNA and PU.1 with unlabeled DNA is described by a competitive model PU.1 (001) which binds either the labeled probe (01*0; the asterisk denotes probe) or unlabeled competitor (010), but not both. We previously applied this model to describe the competition of PU.1 and other ETS proteins between a polymeric DNA fragment and oligomeric DNA harboring various sequences (6,7). The binding polynomial (1), which is cubic in [011] is,

$$0 = \varphi_0 + \varphi_1[011] + \varphi_2[011]^2 + \varphi_3[011]^3$$

$$\begin{cases} \varphi_0 = -K_{01*1}[010]_t^{\,2}[001]_t \\ \varphi_1 = K_{011}K_{01*1}[010]_t + K_{011}[01*0]_t[010]_t + K_{01*1}[010]_t^{\,2} + 2K_{01*1}[010]_t[001]_t - K_{011}[010]_t[001]_t \\ \varphi_2 = -K_{011}K_{01*1} + K_{011}^{\,2} - K_{011}[01*0]_t - 2K_{01*1}[010]_t + K_{011}[010]_t - K_{01*1}[001]_t + K_{011}[001]_t \\ \varphi_3 = K_{01*1} - K_{011} \end{cases}$$

(S5)

The two dissociation constants tend to be highly correlated if both are allowed to float in the fit. We therefore *always* fix K_{01*0} with the value independently determined from direct titration of the probes alone.

Equal-*k* *n*:1 binding (Scheme 2). Given the depleting effects of 10 nM of DB270 on the titrant, we interpreted the "anomalous" binding of DB270 to PU.1 as the independent binding of *n* stoichiometric equivalents of DB270 per PU.1 with a microscopic dissociation constant *k*. In this formulation, the macroscopic dissociation constant is $K_{101} = k_{101}/n$. Since PU.1 (001) represents the titrant as well as the multivalent "receptor," the required model differs from a simple adaptation

of the usual treatment involving a single class of site. Specifically, the relevant equations of state are:

$$\begin{aligned}[100]_t &= [100] + n[n01] \\ [001]_t &= [001] + [n01]\end{aligned} \quad . \tag{S6}$$

The solution for [n01] in terms of [001] is therefore

$$[n01] = \frac{[100]_t[001]}{k_{101} + n[001]} \quad . \tag{S7}$$

Substitution of Eq. (S6) for [001], followed by rearrangement, gives the following quadratic, c.f. Eq. (S1):

$$\begin{aligned}0 &= \varphi_0 + \varphi_1[n01] + \varphi_2[n01]^2 \\ &\begin{cases}\varphi_0 = [100]_t[001]_t \\ \varphi_1 = -k_{101} - [100]_t - n[001]_t \\ \varphi_2 = n\end{cases} \quad .\end{aligned} \tag{S8}$$

DB270/DNA/PU.1 competition (Scheme 3). The combined competitive interactions of pairwise equilibria among drug, DNA, and protein are represented by **Scheme 3** in the main text. As with PU.1/DNA binding, the most efficient strategy is to solve for the unbound concentrations of the three species, from which all of the bound species can be subsequently computed. The relevant equations of state are:

$$
\begin{aligned}
[100]_t &= [100] + [110] + n[n01] \\
[010]_t &= [010] + [110] + [011] + [012] \\
[001]_t &= [001] + [n01] + [011] + 2[012]
\end{aligned}
\tag{S9}
$$

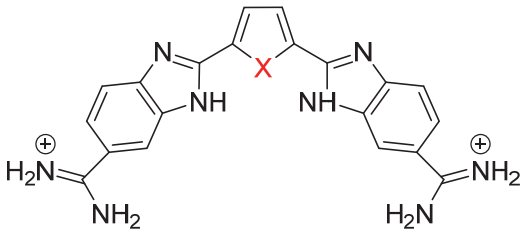
Substitution of the bound species with their equilibrium expressions yield the following system of three non-linear equations:

$$
\begin{cases}
0 = [100]_t - [100]\left(1 + \dfrac{[010]}{K_{110}} + \dfrac{n^2[001]}{k_{101}}\right) \\
0 = [010]_t - [010]\left(1 + \dfrac{[001]}{K_{011}} + \dfrac{[001]^2}{K_{011}K_{012}} + \dfrac{[100]}{K_{110}}\right) \\
0 = [001]_t - \dfrac{2[001]^2[010]}{K_{011}K_{012}} - [001]\left(1 + \dfrac{n[100]}{k_{101}} + \dfrac{[010]}{K_{011}}\right)
\end{cases}
\tag{S10}
$$

Eqs. (S10) are solved by numerical root-finding starting with iterative trials of initial estimates. The set of convergent solutions is then used to compute concentrations of the bound species using values of the total concentrations $[xyz]_t$ and dissociation constants K_{xyz} used in solving the system.

Table S1

Calculated physicochemical properties of the isosteric analogues DB270 and DB1976



X = O (DB270)
X = Se (DB1976)

Physicochemical properties were calculated as follows: *n*-octanol/water partition coefficient (Log P_{ow}), solubility (Log *S*, *S* in molar), polar surface area (PSA), molecular volume (V), van der Waals surface area (SA), polar surface area (PSA).

	DB270 (furan)		**DB1976** (selenophene)		Source
	Dication	Uncharged	Dication	Uncharged	
log P_{ow}	-0.68	1.14	-0.28	1.54	1
log *S*		-2.45		-2.44	2
MR, cm^3 mol^{-1}	110	108	119	117	3
V, A^3	331	325	345	340	1
SA, A^2	382	378	390	387	4
PSA, A^2	174	170	161	157	1

1. Molinspiration Cheminformatics (http://www.molinspiration.com/cgi-bin/properties)
2. pkCSM (http://bleoberis.bioc.cam.ac.uk/pkcsm/)
3. ChemDraw (PerkinElmer Informatics)
4. Volume Assessor (http://3vee.molmovdb.org/volumeCalc.php)

INDEX OF SUPPLEMENTAL FIGURES

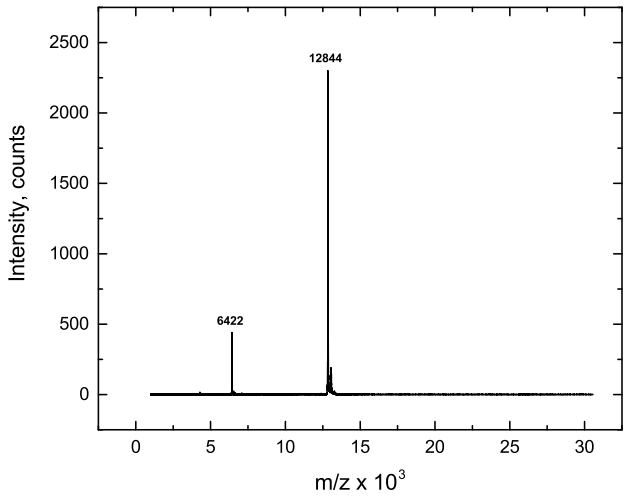


FIGURE S1. **Mass spectrometric analysis of purified recombinant PU.1 ETS domain (PU.1ΔN167).** Eight pmol of protein was mixed with matrix at a 1:1 v/v ratio and analyzed by MALDI-ToF mass spectrometry in linear positive-ion mode. Peaks represent the +1 and +2 ions. The expected MW (12,978) exceeds the observed value by 134, which is consistent with post-translational removal of the N-terminal methionine (-131) *in vivo*. Hydrolytic cleavage by the *E. coli* methionine aminopeptidase is common with Gly in the P1' position (8), which is the case for PU.1ΔN167.

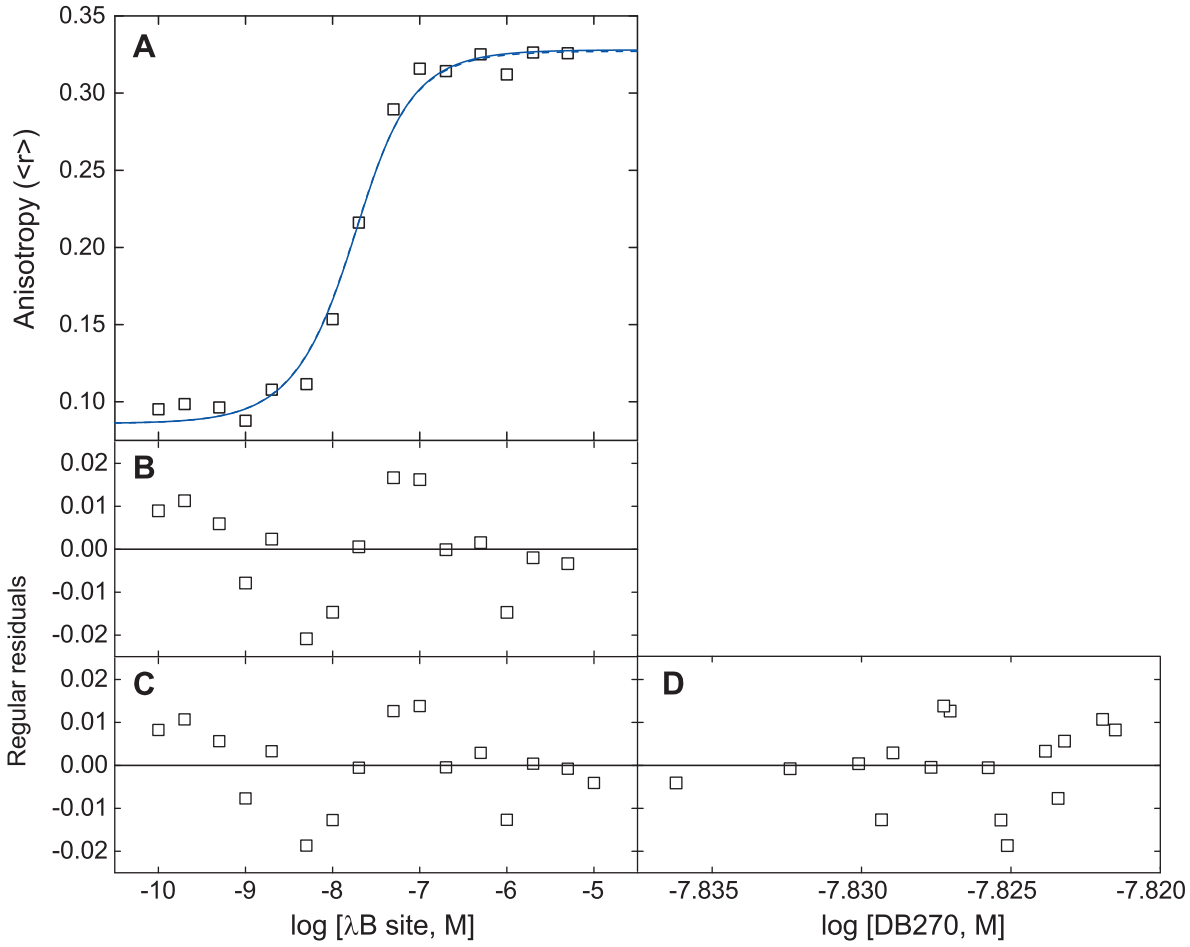


FIGURE S2. **Direct titration of DB270 with the λB motif in the presence of phosphate as counter-ion.** *A*, 15 nM of DB270 was titrated with a duplex DNA oligo harboring the λB motif as specified in **Table 1** of the main text. The λB motif contains a 7-bp A-track that is strongly targeted by AT-selective heterocyclic dications such as DB270. The buffer condition was 25 mM NaH_2PO_4/Na_2HPO_4 (pH 7.4) and sufficient NaCl to give a Na^+ concentration of 150 mM. These conditions differed from those reported in the main text in the substitution of $TrisH^+$/Tris as the buffering species; the Na^+ concentration was identical. *Curves* represent a fit of the data to a 1:1 binding model treated as a univariate function in the titrant concentration (*i.e.,* $[010]_t$; *solid*) or a bivariate function in the titrant as well as the titrate concentrations (*i.e.,* $[100]_t$, *dashed*), the latter being slightly diluted (<10%) during the titration. *B*, Regular residuals of the univariate fit. *C and D*, Residuals of the bivariate fit as a function of the two independent variables. The statistics and estimated parameters of the two fits are:

	Residual sum of squares	**Parameters**		
		$\log K_{110}$	$\langle r_{100} \rangle$	$\langle r_{110} \rangle$
Univariate	0.00131	-7.99 (-8.31, -7.72)	0.086 (0.068, 0.10)	0.33 (0.31, 0.35)
Bivariate	0.00133	-8.00 (-8.29, -7.74)	0.086 (0.069, 0.10)	0.33 (0.31, 0.34)

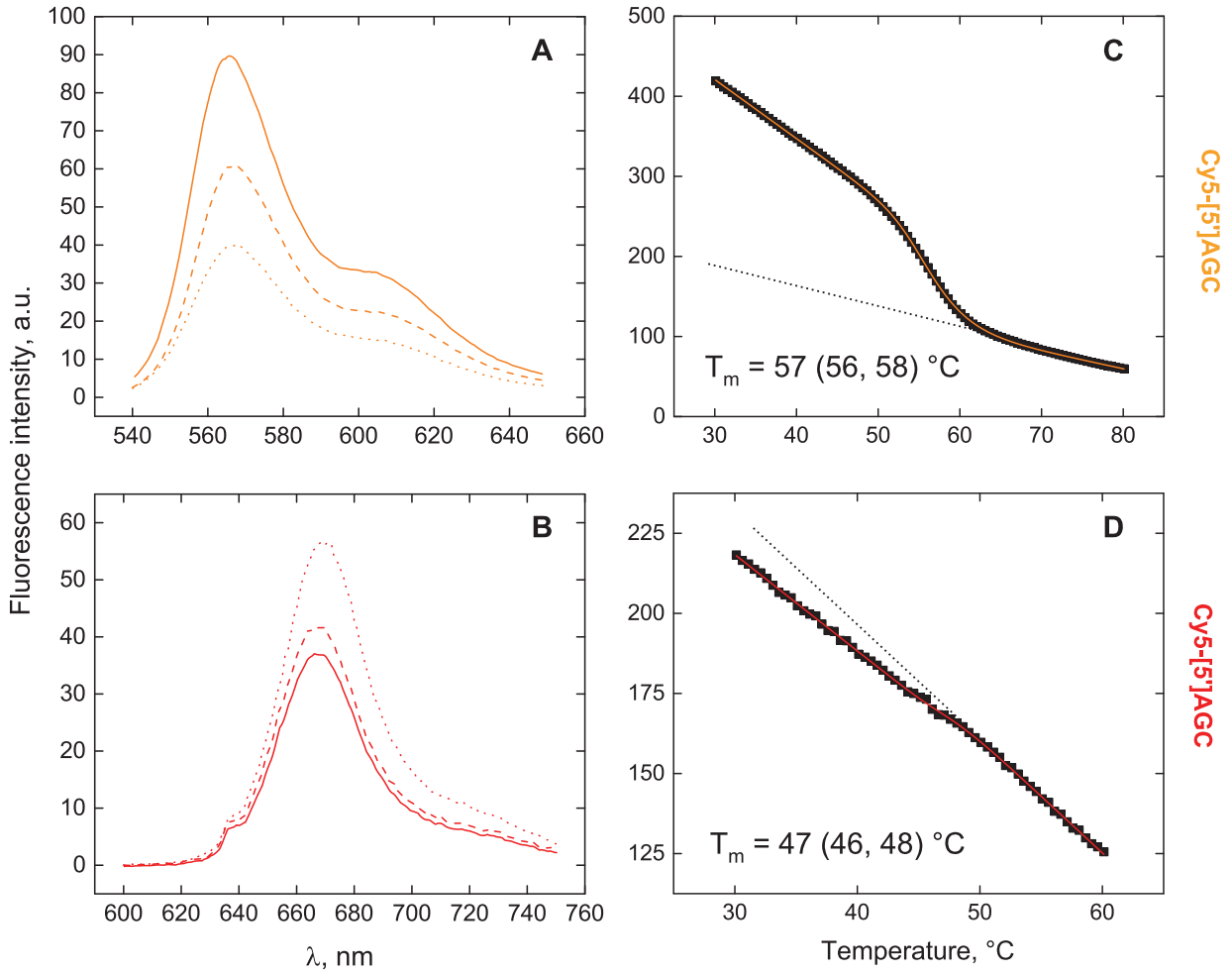


FIGURE S3. **Spectroscopic and thermal characterization of internally Cy3- and Cy5-labeled DNA probe.** *A and B*, Fluorescence emission spectra of Cy3- and Cy5- probes bearing the PU.1 binding sequence [5']AGC (*dotted*), after annealing with a complementary strand (*dashed*), and after saturating the duplexes with 1 μM of PU.1 ETS domain (*solid*). Samples were excited at 522 and 637 nm for Cy3 and Cy5, respectively, at ambient temperature in binding buffer with excitation and emission slit widths of 2.5 and 1.5 nm, respectively. *C and D*, Thermal melting of 3 nM of duplex Cy3- and Cy5-[5']AGC in 10 mM NaH_2PO_4/Na_2HPO_4 (pH 7.4) buffer containing 150 mM Na^+. Fluorescence was measured at 563 and 672 nm for Cy3 and Cy5, respectively, with emission slit width now set at 20 nm. Samples were heated at 0.5°C intervals, with a 60 s holding period at each temperature, before total fluorescence intensity (F) was measured. *Curves* represent fits of the melting data to a two-state helix-coil transition model with linear pre- and post-tranistion baselines F_{helix} and F_{coil}:

$$F(T) = \alpha(F_{coil} - F_{helix}) + F_{helix} . \quad \text{(S11)}$$

α is the fraction melted for non-self-complementary strands as described by Marky and Breslauer (9). Extrapolation of the post-transition baseline (*dotted lines*) back to ambient temperature confirms the opposite trends in fluorescence intensity between the two probes in their single-stranded and duplex states as indicated by their emission spectra in *Panels A and B*.

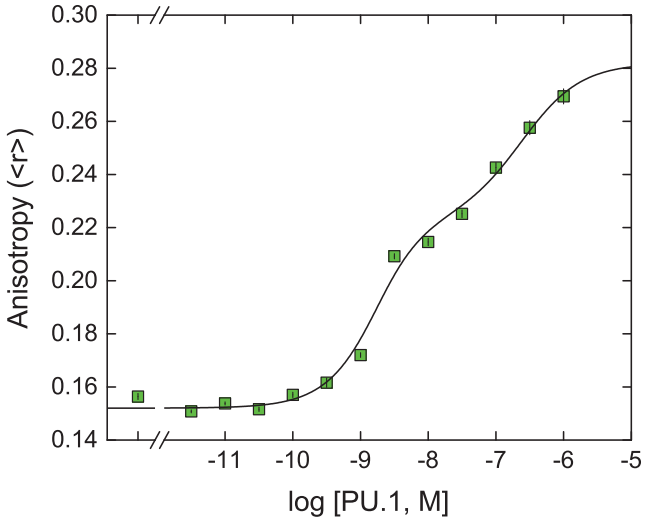


FIGURE S4. **Direct titration of an end-labeled duplex DNA probe by PU.1.** A synthetic oligo 5'-CAAGC**GGAA**GTGAG-3' was labeled at the 5' end with Alexa Fluor 488. After annealing with a fully-matched complementary strand, 1 nM of duplex probe was titrated with PU.1 ETS domain. The conditions of the experiment were identical to those described in the main text using internally Cy3- and Cy5-labeled probes bearing the same PU.1 binding site. *Curve* represents a fit of the data to a 2:1 binding model, as was performed with the Cy data in **Figure 3B** in the main text. The fitted dissociation constants (c.f. **Table 2**) are log K_{011} = -8.94 (-9.14, -8.74) and K_{012} = -6.61 (-6.75, -6.47).

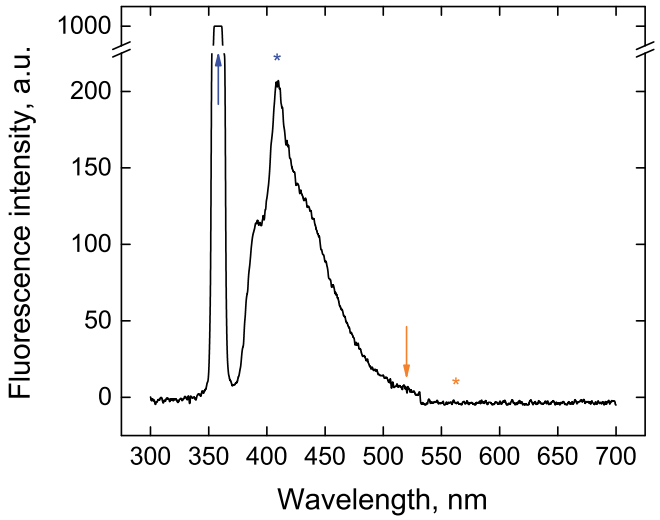


FIGURE S5. **DB270 and internally Cy3-labeled duplex DNA are spectrally isolated.** Emission spectrum of a solution containing 10 nM of DB270 and 1 nM of Cy3-[5']AGC. The sample was irradiated at 358 nm (*blue arrow*), which excites DB270 with an emission maximum at 420 nm (*blue asterisk*), but did not detectably transfer to the Cy3 label (at 522 nm; *orange arrow*), which emitted at 563 nm (*orange asterisk*).

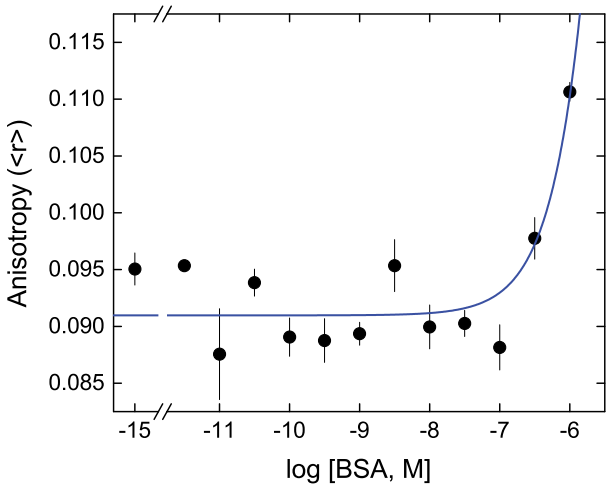


FIGURE S6. **DB270 binds bovine serum albumin with very low affinity.** DB270 (10 nM) was titrated with bovine serum albumin (66 kDa) across the same concentration range and under the same solution conditions as with PU.1 (13 kDa) in **Figure 5A** in the main text (150 mM Na^+, without DNA). The *Curve* represents an emprical fit of the data to the Hill equation only to guide the eye.

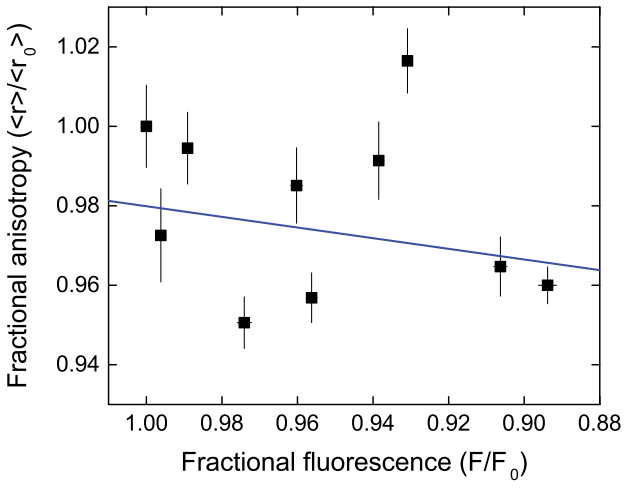


FIGURE S7. **Photobleaching of the DB270/PU.1 complex reveals no homotropic FRET among PU.1-bound DB270.** The anticipated high spatial density of fluorescent DB270 in the DB270/PU.1 complex may be subject to homotropic resonance fluorescence transfer. To detect possible homotropic FRET, the DB270/PU.1 complex was photobleached, by irradiation at the excitation maximum of DB270 (358 nm, slit width 1 nm) to reduce the density of fluorescing DB270. If FRET is operative, the observed anisotropy is expected to rise significantly ($\langle r \rangle / \langle r_0 \rangle > 1$) as total DB270 fluorescence becomes quenched ($F/F_0 < 1$) (10). Starting with fully-bound DB270 (10 nM DB270, 1 μM PU.1; c.f. **Figure 4A**), as total fluorescence intensity was suppressed, no increase in DB270 anisotropy was observed. The *Line* represents a least square fit to the data, but the slope was not statistically different from zero ($p = 0.05$).

SUPPLEMENTAL REFERENCES

1. Wells, J.W. (1992) In Hulme, E. C. (ed.), *Receptor-Ligand Interactions: a Practical Approach*. IRL Press at Oxford University Press, Oxford [England]; New York, pp. 289-395.

2. Poon, G.M. (2013) Quantitative analysis of affinity enhancement by noncovalently oligomeric ligands. *Anal Biochem*, **433**, 19-27.

3. Wang, L., Bailly, C., Kumar, A., Ding, D., Bajic, M., Boykin, D.W. and Wilson, W.D. (2000) Specific molecular recognition of mixed nucleic acid sequences: An aromatic dication that binds in the DNA minor groove as a dimer. *Proc Natl Acad Sci U S A*, **97**, 12-16.

4. Press, W.H. (2007) *Numerical recipes : the art of scientific computing*. 3rd ed. Cambridge University Press, Cambridge, UK ; New York.

5. Poon, G.M.K. (2012) DNA Binding Regulates the Self-Association of the ETS Domain of PU.1 in a Sequence-Dependent Manner. *Biochemistry*, **51**, 4096-4107.

6. Poon, G.M.K. (2012) Sequence Discrimination by DNA-binding Domain of ETS Family Transcription Factor PU.1 Is Linked to Specific Hydration of Protein-DNA Interface. *J Biol Chem*, **287**, 18297-18307.

7. Wang, S., Linde, M.H., Munde, M., Carvalho, V.D., Wilson, W.D. and Poon, G.M. (2014) Mechanistic heterogeneity in site recognition by the structurally homologous DNA-binding domains of the ETS family transcription factors Ets-1 and PU.1. *J Biol Chem*, **289**, 21605-21616.

8. Xiao, Q., Zhang, F., Nacev, B.A., Liu, J.O. and Pei, D. (2010) Protein N-terminal processing: substrate specificity of Escherichia coli and human methionine aminopeptidases. *Biochemistry*, **49**, 5588-5599.

9. Marky, L.A. and Breslauer, K.J. (1987) Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers*, **26**, 1601-1620.

10. Chan, F.T.S., Kaminski, C.F. and Kaminski Schierle, G.S. (2011) HomoFRET Fluorescence Anisotropy Imaging as a Tool to Study Molecular Self-Assembly in Live Cells. *ChemPhysChem*, **12**, 500-509.