ACS OMEGA

Cite This: ACS Omega 2019, 4, 9904−9910

http://pubs.acs.org/journal/acsodf

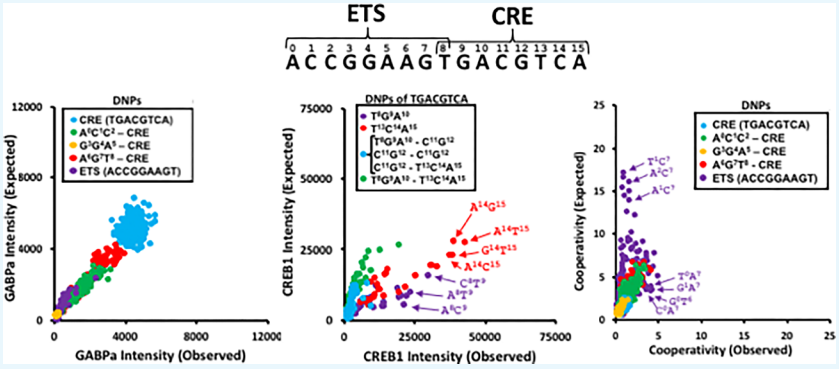
GABPα and CREB1 Binding to Double Nucleotide Polymorphisms of Their Consensus Motifs and Cooperative Binding to the Composite ETS ⇔ CRE Motif (ACCGGAAGTGACGTCA)

Nima Assad,[†] Desiree Tillo,[†] Sreejana Ray,[†] Alexa Dzienny,[†] Peter C. FitzGerald,[‡] and Charles Vinson*[,†]

[†]Laboratory of Metabolism, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, United States

[‡]Genome Analysis Unit, Genetics Branch, National Cancer Institute, National Institutes of Health, Building 37, Bethesda, Maryland 20892, United States

Supporting Information



ABSTRACT: Previously, cooperative binding of the bZIP domain of CREB1 and the ETS domain of GABPα was observed for the composite DNA ETS ⇔ CRE motif ($\mathbf{A^0C^1C^2G^3G^4A^5A^6G^7T^8G^9A^{10}C^{11}G^{12}T^{13}C^{14}A^{15}}$). Single nucleotide polymorphisms (SNPs) at the beginning and end of the ETS motif (**ACCGGAAGT**) increased cooperative binding. Here, we use an Agilent microarray of 60-mers containing all double nucleotide polymorphisms (DNPs) of the ETS ⇔ CRE motif to explore GABPα and CREB1 binding to their individual motifs and their cooperative binding. For GABPα, all DNPs were bound as if each SNP acted independently. In contrast, CREB1 binding to some DNPs was stronger or weaker than expected, depending on the locations of each SNP. CREB1 binding to DNPs where both SNPs were in the same half site, $\mathbf{T^8G^9A^{10}}$ or $\mathbf{T^{13}C^{14}A^{15}}$, was greater than expected, indicating that an additional SNP cannot destroy binding as much as expected, suggesting that an individual SNP is enough to abolish sequence-specific DNA binding of a single bZIP monomer. If a DNP contains SNPs in each half site, binding is weaker than expected. Similar results were observed for additional ETS and bZIP family members. Cooperative binding between GABPα and CREB1 to the ETS ⇔ CRE motif was weaker than expected except for DNPs containing $\mathbf{A^7}$ and SNPs at the beginning of the ETS motif.

INTRODUCTION

In eukaryotic genomes, sequence-specific DNA binding proteins often cooperate to bind composite DNA motifs.[1−11] An example is the ETS ⇔ CRE motif (**ACCGGAAGTGACGTCA**), which localizes to proximal promoters in mammals[12,13] and contains the ETS motif (**ACCGGAAGT**) and the overlapping CRE motif (**GTGACGTCA**) with the **GT** dinucleotide occurring in each motif. The dimeric bZIP domain of CREB1[14] strengthens binding of the monomeric ETS domain of GABPα[15−17] to the ETS ⇔ CRE motif only when the two motifs are spaced in the configuration shown above, as described in ref 12

Previously, we investigated the sequence-specific cooperative binding of GABPα and CREB1 using a custom protein binding microarray (PBM) platform containing 177 440 DNA features consisting of the ETS ⇔ CRE motif and variants.[18] The single nucleotide polymorphisms (SNPs) at the beginning and end of the ETS motif (**ACCGGAAGT**) were more cooperatively bound by CREB1 and GABPα−glutathione S-transferase (GST) than the composite motif. Here, we evaluate GABPα and CREB1 binding to double nucleotide polymorphisms (DNPs) of the composite ETS ⇔ CRE motif (**ACCGGAAGTGACGTCA**) to further explore the nature of GABPα and CREB1 binding to their consensus motifs, as well as to examine their effect on cooperative binding.

Received: February 26, 2019
Accepted: May 24, 2019
Published: June 6, 2019

DOI: 10.1021/acsomega.9b00540
ACS Omega 2019, 4, 9904−9910

Table 1. Design of the 165 384 Feature Custom Agilent ETS ⇔ CRE DNP Microarray[a]

category (SNPs and DNPs)		solvent—variable 36-mer|constant 24-mer—glass	GABPα intensity	CREB1 intensity	cooperativity
solvent	ETS ⇔ CRE	**ACCGGAAGTGACGTCA**GTCCTCAAGAGACTCAGGTG| GGACACACTTTAACACATGGAGAG	930	49 000	3.8
	CRE ⇔ ETS	**TGACGTCACTTCCGGT**GTCCTCAAGAGACTCAGGTG| GGACACACTTTAACACATGGAGAG	8300	18 000	1.2
central	ETS ⇔ CRE	GTCCTCAAGA**ACCGGAAGTGACGTCA**GACTCAGGTG| GGACACACTTTAACACATGGAGAG	4200	50 000	1.9
	CRE ⇔ ETS	GTCCTCAAGA**TGACGTCACTTCCGGT**GACTCAGGTG| GGACACACTTTAACACATGGAGAG	4300	61 000	1.7
glass	ETS ⇔ CRE	GTCCTCAAGAGACTCAGGTG**ACCGGAAGTGACGTCA**| GGACACACTTTAACACATGGAGAG	3300	36 000	1.5
	CRE ⇔ ETS	GTCCTCAAGAGACTCAGGTG**TGACGTCACTTCCGGT**| GGACACACTTTAACACATGGAGAG	3000	57 000	1.3

[a]The experimental microarray DNA probes for every SNP and DNP for the 16-mer, CG dinucleotide containing composite ETS ⇔ CRE motif, **ACCGGAAGTGACGTCA**, with the motif placed either in the center, near the solvent, or near the glass surface of the slide. The ETS ⇔ CRE motif is represented in both orientations on the microarray.

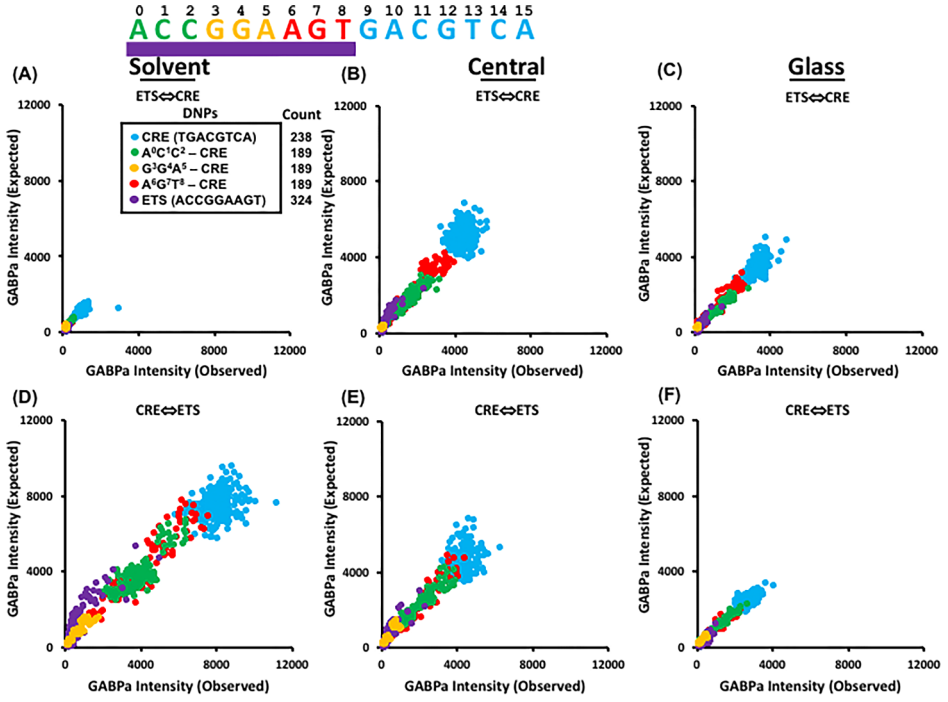


Figure 1. Observed vs Expected GABPα–GST binding to ETS ⇔ CRE DNPs. Scatter plot comparison of observed vs expected GABPα–GST binding intensity to ETS ⇔ CRE motif DNPs for (A) solvent, (B) central, and (C) glass positions on the ETS ⇔ CRE DNP array. Expected DNP binding intensities are calculated as the product of the fold-change in binding intensity observed for each SNP relative to consensus. (D–F) Same as (A–C) for the CRE ⇔ ETS orientation of the motif.

MATERIAL AND METHODS

Cloning and Expression of Mouse bZIP Proteins. We obtained a GABPα–GST plasmid from the Tim Hughes lab, in which the DNA binding domain of GABPα is fused to GST at the C-terminal end to produce the chimeric protein GABPα–GST.[19] The CREB1 bZIP domain without GST was expressed from a pT5 plasmid.[20] The proteins were expressed in in vitro translation (IVT) system reactions using PURExpress an in vitro Protein Synthesis Kit (NEB) as described in ref 19. For the GABPα–GST and CREB1–GST IVT reactions, 570 ng of plasmid was added to 250 μL of IVT solution. For analysis of cooperativity between GABPα–GST and CREB1, 570 ng of GABPα–GST plasmid and 66 ng of CREB1 plasmid (determined by serial dilution for highest cooperativity values; Figure S1) were used in IVT reactions in a final volume of 250 μL. IVT reactions were carried out at 37 °C for 2 h, and then 230 μL of the IVT solution was added to the arrays.

PBM Experiments. The single-stranded DNA 60-mer ETS ⇔ CRE DNP microarrays were double-stranded by primer extension and protein binding reactions were performed as previously described (ref 18) All proteins in this study were assayed twice (Figures S2–S4″), with high agreement between replicates (R = 0.97–0.98) and little to no saturation of spots on the arrays. Arrays with the least number of saturated spots were used for further analysis. Data (raw probe intensities) are available at the NCBI GEO database under accession GSE125613.

Analysis of ETS and bZIP Family DNPs. We obtained PBM Z-score data for all ETS and bZIP family transcription factors (TFs) from v1.02 of the Catalogue of Inferred Sequence Binding Preferences (CISBP[21]) and from other

DOI: 10.1021/acsomega.9b00540
ACS Omega 2019, 4, 9904–9910

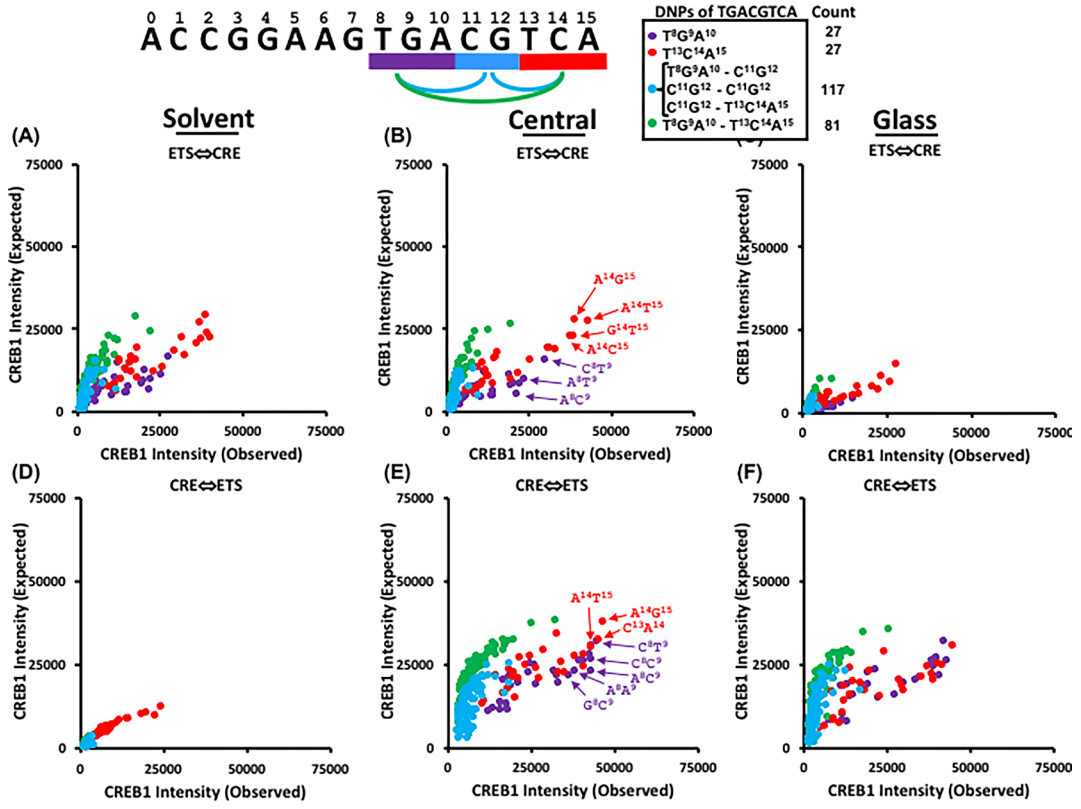


Figure 2. Observed vs Expected CREB1−GST binding to ETS ⇔ CRE DNPs. Scatter plot comparison of observed vs expected CREB1−GST binding intensity to ETS ⇔ CRE motif DNPs for (A) solvent, (B) central, and (C) glass positions on the ETS ⇔ CRE DNP array. Expected DNP binding intensities are calculated as the product of the fold-change in binding intensity observed for each SNP relative to consensus. (D−F) Same as (A−C) for the CRE ⇔ ETS orientation of the motif.

PBM experiments.[22,23] For the 8-mer with the highest *Z*-score, the expected *Z*-score of a DNP was computed from the Z-scores for each individual SNP using the following formula: expected Z-score = [(Z-score SNP1)/(*Z*-score Top8mer)] × [(Z-score SNP2)/(*Z*-score Top8mer)] × *Z*-score Top8mer.

Examination of Cooperative Binding of GABPα and CREB1 to DNPs in Vivo. We used ENCODE ChIP-seq data available for both GABPα and CREB1 for the A549 cell line.[24] We divided the GABPα ChIP-seq peaks into two groups based on the presence of CREB1 binding: "GABPα + CREB1" (regions bound by both GABPα and CREB1) and "GABPα − CREB1" (GABPα peaks that do not overlap CREB1 peaks). These peaks were further subdivided into promoter and nonpromoter sets based on their overlap with a set of promoters (−1000 to +500 bp from the transcription start site) using the refSeq gene annotations for UCSC genome build hg19. For each set of peaks (all, promoter, and nonpromoter GABPα ± CREB1), we computed an "enrichment score", $E = OCC_{obs}/OCC_{exp}$, for the ETS variant $\mathbf{A^0C^1C^2G^3G^4A^5A^6\underline{A}^7}$, an ETS motif containing a SNP at $\underline{\mathbf{A}}^7$, and all of its 1-bp variants. OCC_{obs} is the number of observed occurrences in each set of peaks, and OCC_{exp} is the number of expected occurrences of the motif. OCC_{exp} was calculated as: $OCC_{exp} = N \times L_r/L_g$, where N is the total number of motifs in the whole genome, L_r is the total length (in base pairs) in the set of peaks, and L_g is the total length (in base pairs) of the human genome.

RESULTS

Design of ETS ⇔ CRE DNP Microarray. We designed an array, the ETS ⇔ CRE DNP microarray, which contains all DNPs of the ETS ⇔ CRE (**ACCGGAAGTGACGTCA**) composite DNA motif (Table 1A). The microarray contains 6891 DNA sequences, each occurring 24 times for a total of 165 384 features. The sequences include the composite motif, the 48 SNPs, and the 1080 DNPs for each of the three positions of the ETS ⇔ CRE motif and the three positions of CRE ⇔ ETS, and the reverse orientation of the motif in the 60-mer DNA on the microarray. 117 control probes are included (Table S1). A 24-bp sequence (**GGACACACTTTAACACATGGAGAG**) is nearest the glass and in all features and is complementary to the DNA primer used to make double-stranded DNA (dsDNA) before the binding experiment (see Methods). This microarray was used: (1) to examine the dsDNA binding specificity of GABPα−GST and CREB1−GST, each a chimeric protein containing the GST domain at the C-terminal, and (2) to measure binding of GABPα−GST in the presence of the bZIP domain of CREB1 (cooperative binding). The binding of a fluorescent antibody to the GST domain was used as a measure of the strength of binding to DNA.[19]

GABPα−GST Binding to SNPs and DNPs. We first examined the effect of DNPs of the ETS motif on GABPα−GST binding. Figure 1A−F shows six comparisons of observed versus expected GABPα−GST binding to DNPs of the ETS motif **ACCGGAAGT**. Expected DNP binding intensities are calculated as the product of the fold-change in binding intensity observed for each SNP relative to the consensus

DOI: 10.1021/acsomega.9b00540
ACS Omega 2019, 4, 9904−9910

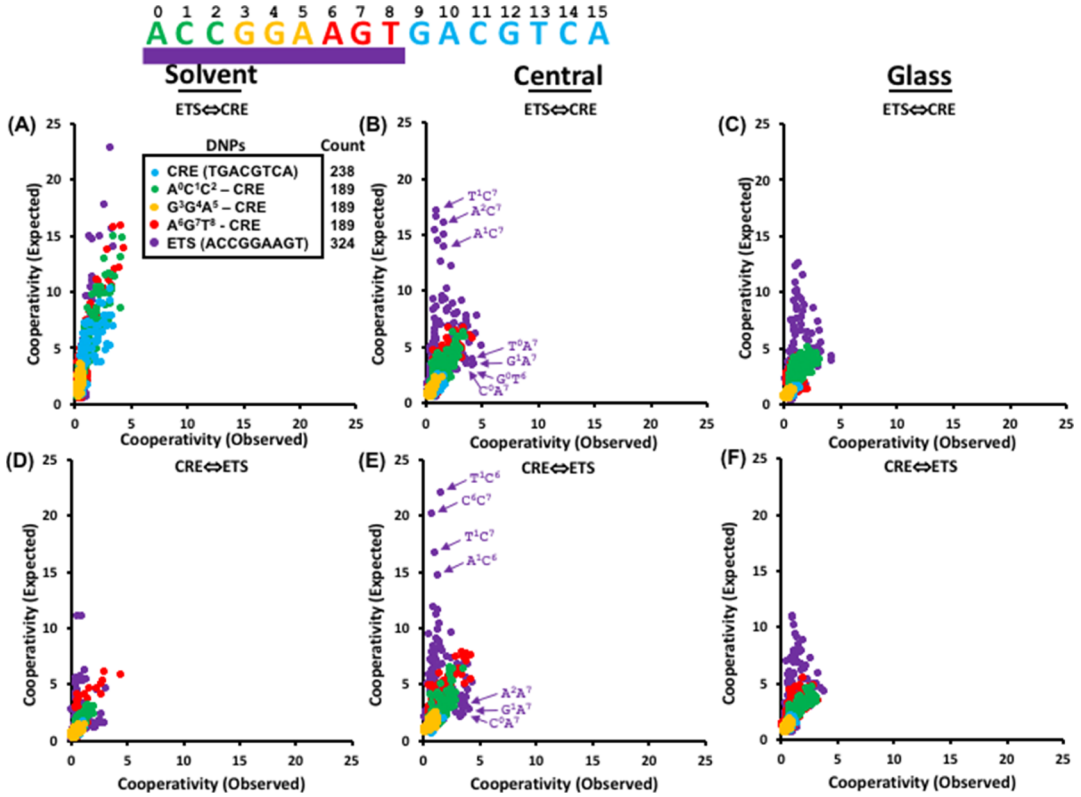


Figure 3. Observed vs Expected GABPα cooperativity with CREB1. Scatter plot comparison of observed vs expected GABPα−GST and CREB1 cooperative binding to the ETS ⇔ CRE motif DNPs at the (A) solvent, (B) central, and (C) glass positions on the ETS ⇔ CRE DNP array. Cooperativity is defined as the ratio of GABPα−GST binding intensity in the presence of CREB1 to GABPα−GST binding intensity in the absence of CREB1.[18] Expected DNP binding intensities are calculated as the product of the fold-change in binding intensity observed for each SNP relative to consensus. (D−F) Same as (A−C) for the CRE ⇔ ETS orientation of the motif.

ACCGGAAGT. For instance, the SNP $\mathbf{C^0}$ reduces GABPα−GST binding two-fold relative to the consensus. The SNP $\mathbf{A^1}$ reduces binding four-fold. Assuming the SNPs contribute independently to binding, the DNP $\mathbf{C^0A^1}$ is expected to reduce binding eight-fold relative to the consensus, which is what is observed at all six positions. Therefore, SNPs of the ETS ⇔ CRE motif contribute independently to GABPα binding. The experimental and observed binding intensities are similar, indicating no cooperative or anticooperative interactions between nucleotides in the ETS motif. Observed GABPα−GST binding intensities are consistent with degeneracy of the GABPα binding site at the flanks of the motif ($\mathbf{A^0C^1C^2}$, green spots; $\mathbf{A^6G^7T^8}$, red spots).[19,25] DNPs with at least one SNP in the central $\mathbf{G^3G^4A^5}$ of the ETS motif (gold and purple spots) reduce GABPα−GST binding several-fold, as observed for GABPα−GST binding to SNPs of the ETS motif.[18] GABPα−GST binding intensity is reduced when the ETS motif is placed at the solvent or when it is far buried in the 60-mer DNA probe (Figure S5).

CREB1−GST Binding to SNPs and DNPs. CREB1 binds to the CRE as a dimer, with each monomer binding different overlapping 5-mers ($\mathbf{T^8G^9A^{10}C^{11}G^{12}}$ or $\mathbf{C^{11}G^{12}T^{13}C^{14}A^{15}}$) with the central CG dinucleotide being bound by each monomer. SNPs at positions $\mathbf{T^8}$, $\mathbf{G^9}$, $\mathbf{C^{14}}$, and $\mathbf{A^{15}}$ at the flanks of the palindromic motif reduce CREB1 binding 2-fold relative to consensus, whereas SNPs at the central $\mathbf{A^{10}C^{11}G^{12}T^{13}}$ of the CRE motif reduce binding 8-fold (Figure S6). Unlike GABPα, DNPs are either stronger or weaker bound than expected for all six positions and orientations of the motif on the microarray (Figure 2, off-diagonal points). In other words, CREB1−GST does not always bind to DNPs of the CRE motif as the independent product of its binding to SNPs of the motif. We have color-coded four groups of DNPs according to the location of each SNP within the CRE and observe opposite effects. DNPs in either half site ($\mathbf{T^8G^9A^{10}}$ or $\mathbf{T^{13}C^{14}A^{15}}$) which exclude the central CG dinucleotide are better bound than expected, as if one SNP abolishes binding and a second SNP in the same half site cannot weaken binding even more. These DNPs tend to be in $\mathbf{T^8G^9}$ and $\mathbf{C^{14}A^{15}}$, which are equivalent positions in the palindromic motif. DNPs with one SNP in each half site ($\mathbf{T^8G^9A^{10}}$−$\mathbf{T^{13}C^{14}A^{15}}$) and DNPs with a SNP in the central CG dinucleotide ($\mathbf{T^8G^9A^{10}}$−$\mathbf{C^{11}G^{12}}$ and $\mathbf{C^{11}G^{12}}$−$\mathbf{T^{13}C^{14}A^{15}}$) are worse bound than expected. In other words, a SNP in each half site compromises both binding half sites and are weaker bound than expected.

Binding of ETS and bZIP Family Members to DNPs. To examine whether the observed binding activities of GABPα and CREB1 to DNPs is a general property of ETS and bZIP families, we analyzed all publicly available mouse bZIP (64) and ETS (23) PBM data sets[21−23] (Figures S7,S8). For this analysis, we examined Z-scores,[26] a measure of relative binding which has a near linear relationship with fluorescent intensity of our custom PBMs.[18] For ETS family members, each SNP contributes independently to binding of DNPs (Figure S7), as observed for GABPα-GST. For most bZIP family members, several DNPs in one-half site are better bound than expected as observed for CREB1 (Figure S8). The DNPs that are better bound than expected are both in the same half site as observed

DOI: 10.1021/acsomega.9b00540
ACS Omega 2019, 4, 9904−9910

for CREB1−GST. This is particularly true for bZIPs binding the CRE (**TGACGTCA**) as opposed to those binding the PAR motif (**TTACGTAA**).

Cooperative Binding of GABPα−GST and CREB1 to DNPs. Figure 3 compares the observed and expected cooperativity of GABPα−GST binding in the presence of CREB1 to DNPs of the ETS ⇔ CRE and the CRE ⇔ ETS motif at all six locations. Cooperativity is defined as the ratio of GABPα−GST binding in the presence of CREB1 to GABPα−GST binding.[18] The expected cooperativity of an ETS ⇔ CRE DNP is defined as the product of cooperativity observed for each SNP. Different patterns of observed versus expected cooperativity are obtained depending on the orientation and location of the ETS ⇔ CRE motif.

Focusing on the ETS ⇔ CRE at the central position (Figure 3B), the DNPs in the ETS motif (purple) are better or worse bound than expected compared to DNPs with only one SNP in the ETS motif (green, yellow, and red). DNPs within the ETS motif $\mathbf{A^0C^1C^2G^3G^4A^5A^6\underline{G}^7\underline{T}^8}$ (purple) may be divided into two classes: those which are less cooperatively bound than expected (points on the upper left of the plot) and those which are more cooperatively bound than expected (points on the lower right of the plot). The ETS DNPs that are less cooperatively bound than expected have SNPs with high cooperativity (e.g., $\mathbf{T^1C^7}$, $\mathbf{C^6C^7}$, $\mathbf{A^2C^7}$, $\mathbf{T^1C^6}$, $\mathbf{T^1A^2}$, $\mathbf{A^2C^6}$, $\mathbf{A^1C^7}$, $\mathbf{A^1C^6}$, and $\mathbf{A^1A^2}$) (Figure S9). The second class of DNPs within the ETS motif that are more cooperatively bound than expected often involve the SNP $\mathbf{A^7}$ (e.g., $\mathbf{G^1A^7}$ and $\mathbf{G^6A^7}$). A histogram for DNPs of the SNP $\mathbf{A^7}$ (Figure S9B) shows that the SNP $\mathbf{A^7}$ has a cooperativity value of only 2, whereas DNPs at positions 0, 1, 2, and 6 have cooperativity values between 3 and 5. This indicates that the SNP $\mathbf{A^7}$ enhances cooperativity of SNPs at the beginning of the ETS motif.

GABPα and CREB1 Cooperative Binding to ETS ⇔ CRE DNPs in Vivo. We examined publicly available ChIP-seq data in A549 cells to determine if GABPα and CREB1 preferentially colocalized to genomic regions containing DNPs of the ETS motif (an in-depth microarray and ChIP-seq comparison of GABPα and CREB1 cooperative binding to SNPs of the composite ETS ⇔ CRE motif may be found in ref 18). For this analysis, we chose the sequence $\mathbf{A^0C^1C^2G^3G^4A^5A^6\underline{A}^7}$, an ETS motif containing a SNP at $\mathbf{\underline{A}^7}$, which produces the highest observed cooperativity with additional SNPs (Figure 3B,E). The ETS ⇔ CRE motif SNP $\mathbf{\underline{A}^7}$ is 2-fold more enriched in genomic regions containing overlapping GABPα and CREB1 ChIP-seq peaks (GABPα + CREB1) versus genomic regions in which GABPα is bound alone (GABPα − CREB1) (Figure S10A). Examination of specific DNPs highlight $\mathbf{G^1A^7}$ (2.6), $\mathbf{C^5A^7}$ (2.2), and $\mathbf{G^6A^7}$ (2.4) DNPs with greater enrichment in cobound GABPα and CREB1 ChIP-seq peaks than those containing a single $\mathbf{\underline{A}^7}$ SNP. In particular, $\mathbf{G^1A^7}$ and $\mathbf{G^6A^7}$ DNPs show the most cooperativity, providing evidence that preferential binding of these DNPs occurs only when CREB1 is colocalized in vivo (Figure 3). The remaining DNP ($\mathbf{C^5A^7}$), however, shows little cooperativity in our in vitro experiments. This suggests that cooperative binding with other family members or mechanisms other than intrinsic transcription factor-DNA binding affinity (e.g., chromatin posttranslational modifications, recruitment of cofactors and other protein complexes) can drive cooperativity in vivo. Examination of enrichment scores of GABPα and CREB1 ChIP-seq peaks separated by nonpromoter and promoter status (Figure S10B,C) indicates that the cooperativity of binding the $\mathbf{G^1A^7}$ and $\mathbf{G^6A^7}$ DNPs is strongest for GABPα + CREB1 peaks in nonpromoter regions, whereas cooperativity of binding the $\mathbf{C^5A^7}$ DNP is strongest in promoter-associated GABPα + CREB1 peaks, suggesting that genomic or regulatory (e.g., promoter or enhancer) context may also affect cooperativity in vivo.

DISCUSSION

GABPα and CREB1 cooperatively bind to the composite ETS ⇔ CRE motif **ACCGGAA<u>GT</u>GACGTCA**.[12] Cooperativity is enhanced for several SNPs at the beginning and end of the ETS motif (**<u>ACC</u>GGA<u>AG</u>**),[18] suggesting an intricate allostery.[27] To explore this cooperativity in more detail, we designed the ETS ⇔ CRE DNP microarray that contains all DNPs of the ETS ⇔ CRE motif. For GABPα−GST, SNPs contributed independently to binding DNPs of the canonical ETS motif. For CREB1, DNPs with both SNPs in the same half site are better bound than expected and DNPs with a SNP in each half site are worse bound than expected. This suggests that the CRE motif can sustain a SNP and still maintain a functional binding site for CREB1. In other words, both half sites of the CRE motif must be compromised to abolish sequence-specific DNA binding to the motif. Examination of publicly available PBM data indicates that these differences in binding DNPs appear to be a general property of ETS and bZIP family members.

The preferential binding of bZIPs to DNPs occurring in the same half site of the palindromic motif may be explained by the second SNP failing to compromise binding because sequence-specific DNA binding was destroyed by the first SNP. This is in stark contrast to DNPs in which each SNP occurs in different half sites, which would destroy optimal binding of each bZIP monomer. This result highlights the cooperative binding of the two monomers in the bZIP dimer. One SNP is sufficient to break the specificity of a CREB1 monomer to its DNA binding half-site. In contrast, ETS proteins bind DNA as monomers and are unable to bind DNPs better than expected. This highlights how a single SNP does not destroy sequence specific DNA binding. For ETS proteins, multiple changes to the DNA are necessary to break the specificity of the protein to its DNA binding site. These properties are general for these two families of sequence-specific DNA binding proteins.

GABPα and CREB1 cooperatively bind some DNPs of the ETS ⇔ CRE motif. Many DNPs containing the SNP $\mathbf{A^7}$ are stronger bound than expected. In contrast, DNPs containing SNPs at opposite ends of the ETS motif ($\mathbf{\underline{A}^0\underline{C}^1\underline{C}^2G^3G^4A^5\underline{A}^6\underline{G}^7}$) are worse bound than expected. Examination of ChIP-seq data in A549 cells shows GABPα and CREB1 preferentially colocalized to genomic regions containing DNPs of the ETS motif $\mathbf{A^0C^1C^2G^3G^4A^5A^6\underline{A}^7}$. Several of the most cooperatively bound DNPs in our in vitro experiment are also preferentially bound by GABPα only when CREB1 is colocalized in vivo. A few DNPs which are not cooperatively bound in our in vitro experiment are well bound in vivo, particularly at promoters, suggesting the role of genomic contexts, other family members, or mechanisms other than intrinsic transcription factor-DNA binding affinity driving cooperativity in vivo.

Cooperative binding of ETS and bZIP family members to DNA has previously been described. The bZIP heterodimer AP-1 has been shown to cooperatively bind DNA in the presence of NFAT.[28] Cooperative DNA binding has also been shown between members of the ETS family, such as C/EBP,

DOI: 10.1021/acsomega.9b00540
ACS Omega 2019, 4, 9904−9910

and NF-kB,[29] PAX,[16,30,31] and bZIP family members.[32] Cooperative TF binding is thought to be a critical mechanism for fine-tuning genetic regulation.[1,8,16,33,34] For GABPα and CREB1 colocalization in vivo, accumulation to some genomic positions can be driven by intrinsic DNA binding specificity, as we have identified in vitro.

Although the understanding of protein−protein interactions has matured,[35,36] interactions driving protein−DNA complex formations remain less explored. These intricate data sets provide insight into the interconnected protein−DNA interactions employed by genetic regulatory processes. It would be interesting to study the cooperative interactions of other transcription factors known to bind the ETS ⇔ CRE and ETS ⇔ AP1 motifs using the custom DNA microarray platform.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsomega.9b00540.

ETS ⇔ CRE DNP microarray control probe descriptions, intensities of GABPα-GST binding to the ETS ⇔ CRE motif and SNPs with varying amounts of CREB1, duplicate GABPα−GST, GABPα−GST + CREB1−PT5, and CREB1−GST PBM comparisons, control probe intensity histograms, CREB1−GST binding intensities to ETS ⇔ CRE SNPs at 6 positions, observed versus expected mouse ETS and bZIP binding intensities to DNPs of consensus motifs, cooperative binding of GABPα and CREB1 to central ETS ⇔ CRE, SNPs, and position 7 DNPs, and ratio of enrichment scores of GABPα in ENCODE ChIP data of A549 cells of DNPs at $\mathbf{A}^7$ (PDF)

Accession Codes

Data (raw probe intensities) are available at the NCBI GEO database under accession GSE125613.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: vinsonc@mail.nih.gov. Phone: 1-301-496-8783. Fax: 1-301-496-8419.

ORCID

Nima Assad: 0000-0002-2751-2327

Desiree Tillo: 0000-0003-3568-6148

Funding

This work is supported by the intramural research project of the National Cancer Institute, National Institutes of Health (NIH), Bethesda, USA.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) McGhee, J. D.; von Hippel, P. H. Theoretical aspects of DNA-protein interactions: co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice. *J. Mol. Biol.* **1974**, *86*, 469−489.

(2) Arnosti, D. N.; Kulkarni, M. M. Transcriptional enhancers: Intelligent enhanceosomes or flexible billboards? *J. Cell. Biochem.* **2005**, *94*, 890−898.

(3) Panne, D. The enhanceosome. *Curr. Opin. Struct. Biol.* **2008**, *18*, 236−242.

(4) Wunderlich, Z.; Mirny, L. A. Different gene regulation strategies revealed by analysis of binding motifs. *Trends Genet.* **2009**, *25*, 434−440.

(5) Martinez, G. J.; Rao, A. Cooperative transcription factor complexes in control. *Science* **2012**, *338*, 891−892.

(6) Ptashne, M. Epigenetics: core misconcept. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 7101−7103.

(7) Ptashne, M. The chemistry of regulation of genes and other things. *J. Biol. Chem.* **2014**, *289*, 5417−5435.

(8) Jolma, A.; Yin, Y.; Nitta, K. R.; Dave, K.; Popov, A.; Taipale, M.; Enge, M.; Kivioja, T.; Morgunova, E.; Taipale, J. DNA-dependent formation of transcription factor pairs alters their binding specificity. *Nature* **2015**, *527*, 384−388.

(9) Kasahara, K.; Shiina, M.; Fukuda, I.; Ogata, K.; Nakamura, H. Molecular mechanisms of cooperative binding of transcription factors Runx1-CBFbeta-Ets1 on the TCRalpha gene enhancer. *PLoS One* **2017**, *12*, No. e0172654.

(10) Morgunova, E.; Taipale, J. Structural perspective of cooperative transcription factor binding. *Curr. Opin. Struct. Biol.* **2017**, *47*, 1−8.

(11) Lambert, S. A.; Jolma, A.; Campitelli, L. F.; Das, P. K.; Yin, Y.; Albu, M.; Chen, X.; Taipale, J.; Hughes, T. R.; Weirauch, M. T. The Human Transcription Factors. *Cell* **2018**, *172*, 650−665.

(12) Chatterjee, R.; Zhao, J.; He, X.; Shlyakhtenko, A.; Mann, I.; Waterfall, J. J.; Meltzer, P.; Sathyanarayana, B. K.; FitzGerald, P. C.; Vinson, C. Overlapping ETS and CRE Motifs ((G/C)CGGAAGTGACGTCA) preferentially bound by GABPα and CREB proteins. *G3: Genes, Genomes, Genet.* **2011**, *2*, 1243−1256.

(13) Rozenberg, J. M.; Bhattacharya, P.; Chatterjee, R.; Glass, K.; Vinson, C. Combinatorial recruitment of CREB, C/EBPβ and c-Jun determines activation of promoters upon keratinocyte differentiation. *PLoS One* **2013**, *8*, No. e78179.

(14) Vinson, C.; Myakishev, M.; Acharya, A.; Mir, A. A.; Moll, J. R.; Bonovich, M. Classification of human B-ZIP proteins based on dimerization properties. *Mol. Cell. Biol.* **2002**, *22*, 6321−6335.

(15) Batchelor, A. H.; Piper, D. E.; de la Brousse, F. C.; McKnight, S. L.; Wolberger, C. The structure of GABPα/β: an ETS domain-ankyrin repeat heterodimer bound to DNA. *Science* **1998**, *279*, 1037−1041.

(16) Garvie, C. W.; Hagman, J.; Wolberger, C. Structural studies of Ets-1/Pax5 complex formation on DNA. *Mol. Cell* **2001**, *8*, 1267−1276.

(17) Hollenhorst, P. C.; Ferris, M. W.; Hull, M. A.; Chae, H.; Kim, S.; Graves, B. J. Oncogenic ETS proteins mimic activated RAS/MAPK signaling in prostate cells. *Genes Dev.* **2011**, *25*, 2147−2157.

(18) He, X.; Syed, K. S.; Tillo, D.; Mann, I.; Weirauch, M. T.; Vinson, C. GABPα Binding to Overlapping ETS and CRE DNA Motifs Is Enhanced by CREB1: Custom DNA Microarrays. *G3: Genes, Genomes, Genet.* **2015**, *5*, 1909−1918.

(19) Badis, G.; Berger, M. F.; Philippakis, A. A.; Talukder, S.; Gehrke, A. R.; Jaeger, S. A.; Chan, E. T.; Metzler, G.; Vedenko, A.; Chen, X.; Kuznetsov, H.; Wang, C.-F.; Coburn, D.; Newburger, D. E.; Morris, Q.; Hughes, T. R.; Bulyk, M. L. Diversity and complexity in DNA recognition by transcription factors. *Science* **2009**, *324*, 1720−1723.

(20) Ahn, S.; Olive, M.; Aggarwal, S.; Krylov, D.; Ginty, D. D.; Vinson, C. A dominant-negative inhibitor of CREB reveals that it is a general mediator of stimulus-dependent transcription of c-fos. *Mol. Cell. Biol.* **1998**, *18*, 967−977.

(21) Weirauch, M. T.; Yang, A.; Albu, M.; Cote, A. G.; Montenegro-Montero, A.; Drewe, P.; Najafabadi, H. S.; Lambert, S. A.; Mann, I.; Cook, K.; Zheng, H.; Goity, A.; van Bakel, H.; Lozano, J.-C.; Galli, M.; Lewsey, M. G.; Huang, E.; Mukherjee, T.; Chen, X.; Reece-Hoyes, J. S.; Govindarajan, S.; Shaulsky, G.; Walhout, A. J. M.; Bouget, F.-Y.; Ratsch, G.; Larrondo, L. F.; Ecker, J. R.; Hughes, T. R. Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* **2014**, *158*, 1431−1443.

(22) Syed, K. S.; He, X.; Tillo, D.; Wang, J.; Durell, S. R.; Vinson, C. 5-Methylcytosine (5mC) and 5-Hydroxymethylcytosine (5hmC)

DOI: 10.1021/acsomega.9b00540
ACS Omega 2019, 4, 9904−9910

Enhance the DNA Binding of CREB1 to the C/EBP Half-Site Tetranucleotide GCAA. *Biochemistry* **2016**, *55*, 6940−6948.
(23) Tillo, D.; Ray, S.; Syed, K.-S.; Gaylor, M. R.; He, X.; Wang, J.; Assad, N.; Durell, S. R.; Porollo, A.; Weirauch, M. T.; Vinson, C. The Epstein-Barr Virus B-ZIP Protein Zta Recognizes Specific DNA Sequences Containing 5-Methylcytosine and 5-Hydroxymethylcytosine. *Biochemistry* **2017**, *56*, 6200−6210.
(24) The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **2012**, *489*, 57−74. DOI: 10.1038/nature11247
(25) Valouev, A.; Johnson, D. S.; Sundquist, A.; Medina, C.; Anton, E.; Batzoglou, S.; Myers, R. M.; Sidow, A. Genome-wide analysis of transcription factor binding sites based on ChIP-Seq data. *Nat. Methods* **2008**, *5*, 829−834.
(26) Berger, M. F.; Bulyk, M. L. Universal protein-binding microarrays for the comprehensive characterization of the DNA-binding specificities of transcription factors. *Nat. Protoc.* **2009**, *4*, 393−411.
(27) Lefstin, J. A.; Yamamoto, K. R. Allosteric effects of DNA on transcriptional regulators. *Nature* **1998**, *392*, 885−888.
(28) Erlanson, D. A.; Chytil, M.; Verdine, G. L. The leucine zipper domain controls the orientation of AP-1 in the NFAT·AP-1·DNA complex. *Chem. Biol.* **1996**, *3*, 981−991.
(29) Stein, B.; Baldwin, A. S., Jr. Distinct mechanisms for regulation of the interleukin-8 gene involve synergism and cooperativity between C/EBP and NF-kappa B. *Mol. Cell. Biol.* **1993**, *13*, 7191−7198.
(30) Fitzsimmons, D.; Hodsdon, W.; Wheat, W.; Maira, S. M.; Wasylyk, B.; Hagman, J. Pax-5 (BSAP) recruits Ets proto-oncogene family proteins to form functional ternary complexes on a B-cell-specific promoter. *Genes Dev.* **1996**, *10*, 2198−2211.
(31) Maier, H.; Ostraat, R.; Parenti, S.; Fitzsimmons, D.; Abraham, L. J.; Garvie, C. W.; Hagman, J. Requirements for selective recruitment of Ets proteins and activation of mb-1/Ig-α gene transcription by Pax-5 (BSAP). *Nucleic Acids Res.* **2003**, *31*, 5483−5489.
(32) Kim, S.; Denny, C. T.; Wisdom, R. Cooperative DNA binding with AP-1 proteins is required for transformation by EWS-Ets fusion proteins. *Mol. Cell. Biol.* **2006**, *26*, 2467−2478.
(33) Kazemian, M.; Pham, H.; Wolfe, S. A.; Brodsky, M. H.; Sinha, S. Widespread evidence of cooperative DNA binding by transcription factors in Drosophila development. *Nucleic Acids Res.* **2013**, *41*, 8237−8252.
(34) Deplancke, B.; Alpern, D.; Gardeux, V. The Genetics of Transcription Factor DNA Binding Variation. *Cell* **2016**, *166*, 538−554.
(35) Rual, J.-F.; Venkatesan, K.; Hao, T.; Hirozane-Kishikawa, T.; Dricot, A.; Li, N.; Berriz, G. F.; Gibbons, F. D.; Dreze, M.; Ayivi-Guedehoussou, N.; Klitgord, N.; Simon, C.; Boxem, M.; Milstein, S.; Rosenberg, J.; Goldberg, D. S.; Zhang, L. V.; Wong, S. L.; Franklin, G.; Li, S.; Albala, J. S.; Lim, J.; Fraughton, C.; Llamosas, E.; Cevik, S.; Bex, C.; Lamesch, P.; Sikorski, R. S.; Vandenhaute, J.; Zoghbi, H. Y.; Smolyar, A.; Bosak, S.; Sequerra, R.; Doucette-Stamm, L.; Cusick, M. E.; Hill, D. E.; Roth, F. P.; Vidal, M. Towards a proteome-scale map of the human protein-protein interaction network. *Nature* **2005**, *437*, 1173−1178.
(36) Szklarczyk, D.; Franceschini, A.; Wyder, S.; Forslund, K.; Heller, D.; Huerta-Cepas, J.; Simonovic, M.; Roth, A.; Santos, A.; Tsafou, K. P.; Kuhn, M.; Bork, P.; Jensen, L. J.; von Mering, C. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **2015**, *43*, D447−D452.

DOI: 10.1021/acsomega.9b00540
ACS Omega 2019, 4, 9904−9910