Principles of dimer-specific gene regulation revealed by a comprehensive characterization of NF-κB family DNA binding

Trevor Siggers[1,7], Abraham B Chang[2,7], Ana Teixeira[3], Daniel Wong[3], Kevin J Williams[2], Bilal Ahmed[1,4], Jiannis Ragoussis[3], Irina A Udalova[5], Stephen T Smale[2] & Martha L Bulyk[1,4,6]

The unique DNA-binding properties of distinct NF-κB dimers influence the selective regulation of NF-κB target genes. To more thoroughly investigate these dimer-specific differences, we combined protein-binding microarrays and surface plasmon resonance to evaluate DNA sites recognized by eight different NF-κB dimers. We observed three distinct binding-specificity classes and clarified mechanisms by which dimers might regulate distinct sets of genes. We identified many new nontraditional NF-κB binding site (κB site) sequences and highlight the plasticity of NF-κB dimers in recognizing κB sites with a single consensus half-site. This study provides a database that can be used in efforts to identify NF-κB target sites and uncover gene regulatory circuitry.

The transcription factor NF-κB regulates a broad range of genes central to the body's immune and inflammatory responses[1–4]. NF-κB represents homo- and heterodimers of five different family members: c-Rel (REL), RelA/p65 (RELA), RelB (RELB), p50/p105 (NFKB1) and p52/p100 (NFKB2)[5–7]. Studies of knockout mice have shown that each NF-κB family member carries out unique biological functions[8–11]. At a molecular level, DNA-binding differences of individual NF-κB dimers have been linked to dimer-specific roles in gene regulation[6,7]; however, much remains unclear regarding the full scope of these differences and how they affect dimer-specific functions *in vivo*.

Protein-DNA crystal structures[6,12] and DNA-binding studies[12–14] have led to a basic partition of NF-κB family members: p50 and p52 recognize a 5-bp 5′-GGGRN-3′ half-site, whereas c-Rel, RelA and RelB recognize a 4-bp 5′-GGRR-3′ half-site (where R is A or G, and N is A, C, G or T). These half-sites, separated by a 1-bp spacer, have led to the consensus κB site 5′-GGGRN(Y)YYCC-3′ (ref. 15). Despite the appeal of this paradigm, reports of additional dimer-specific DNA-binding preferences[16] and noncanonical κB site sequences[5,17] suggest complications to this picture.

Dimer-specific DNA recognition provides a mechanism for disentangling the *in vivo* functions of NF-κB heterodimers and closely related homodimers. One such example is specific recognition of the mouse B lymphocyte chemoattractant (BLC)-κB site reported for the RelB-p52 heterodimer, the primary dimer mediating the alternative NF-κB signaling pathway[16,18–20]. However, contradictory results for RelB-p52-specific binding have been reported[21]. Binding of other NF-κB dimers to noncanonical binding site sequences has also been reported and is suggestive of plasticity in DNA binding. Examples include the c-Rel target site in the *Il12b* gene promoter (5′-GGGGAATTTT-3′)[17] and the CD28 response element (CD28RE) from the *Il2* and *Csf2* gene promoters (5′-GGAATTTCT-3′)[5]. Both sites deviate from the consensus sequence and score poorly according to the standard position weight matrices (PWMs) derived from binding site selections[13]. Structural analyses of NF-κB dimers in complex with different κB site sequences have also demonstrated plasticity in amino acid–base interactions[12]. Together, these observations indicate that the consensus sequence and PWM descriptions of NF-κB DNA binding may be too limited.

To address these issues, we used protein-binding microarrays (PBMs)[22–24] and surface plasmon resonance (SPR) to carry out an unbiased characterization of potential κB binding-site sequences using multiple NF-κB dimers. Earlier large-scale analyses of NF-κB DNA binding have been biased either to certain κB site sequences[14,25] or to only the few highest-affinity sites[13]. We observed three distinct NF-κB binding-specificity classes, identified many new, nontraditional κB site sequences and highlight the plasticity of NF-κB dimer binding for shorter κB sites with one consensus half-site. We provide a new data set that could be useful in genomic analysis of NF-κB regulatory elements and the interpretation of *in vivo* binding experiments. The data set and online tools with DNA sequence search capabilities are online (http://thebrain.bwh.harvard.edu/nfkb/).

[1]Division of Genetics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA. [2]Molecular Biology Institute and Department of Microbiology, Immunology, and Molecular Genetics, University of California Los Angeles, Los Angeles, California, USA. [3]Wellcome Trust Centre for Human Genetics, Oxford University, Oxford, UK. [4]Harvard-MIT Division of Health Sciences and Technology, Harvard Medical School, Boston, Massachusetts, USA. [5]Kennedy Institute of Rheumatology, Imperial College, London, UK. [6]Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA. [7]These authors contributed equally to this work. Correspondence should be addressed to M.L.B. (mlbulyk@receptor.med.harvard.edu).
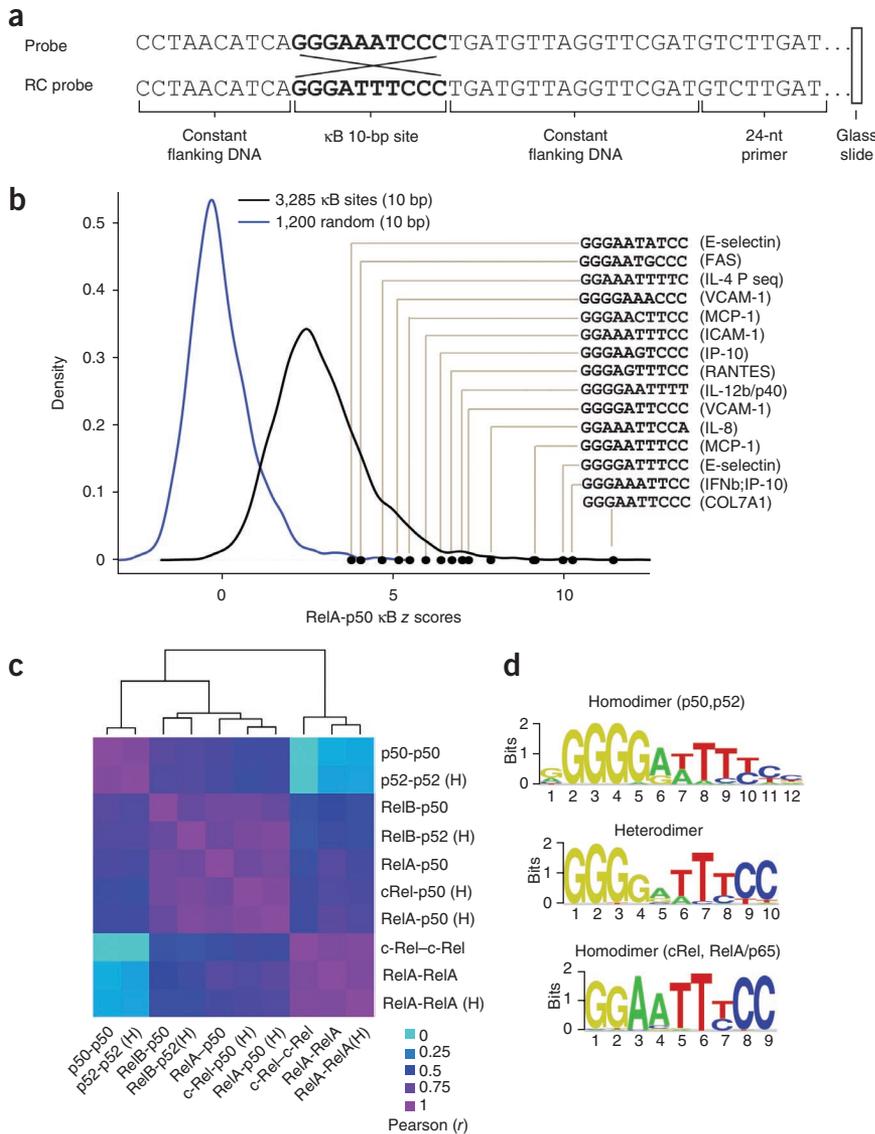
Figure 1 Examining NF-κB dimer binding by custom NF-κB PBMs. (**a**) Design of 60-bp DNA sequence probes on custom NF-κB PBM. κB sites (10 bp) are at a fixed position along the probe (relative to the glass slide surface) within constant flanking sequence. Each 10-bp κB site is present at four replicate spots in both the forward (Probe) and reverse complement (RC probe) orientation (eight spots in total). (**b**) Distributions of PBM-derived binding site z scores for mouse RelA-p50 binding to 3,285 κB sites and to a background set of 1,200 random 10-bp sequences. z scores for 15 κB sites described in literature. (**c**) Pairwise comparison of κB site binding for ten NF-κB dimers. Pairwise binding similarity was assessed by Pearson correlation of κB site z scores, and hierarchical clustering was carried out on the comparison matrix (Online Methods). Representative DNA-binding site motifs were determined for each dimer class using the top 25 highest-scoring κB sites bound by each group member (Online Methods; see **Supplementary Fig. 2** for individual motifs). Data are representative of single experiments (**b**,**c**; median of eight replicates) or derived from pooled data of two (**d**, top logo), five (**d**, middle logo), or three (**d**, bottom logo) experiments.

RESULTS

Designing an NF-κB-specific protein-binding microarray

To examine the DNA-binding specificities of NF-κB dimers in a systematic and unbiased manner, we used PBM technology. PBMs are double-stranded DNA microarrays that allow the *in vitro* characterization of protein binding to tens of thousands of unique DNA sequences in a single experiment[24,26,27]. The universal PBM (uPBM) developed earlier[23] allows a comprehensive, unbiased assessment of protein-DNA binding to all ungapped and gapped 8-bp sequences. We carried out uPBM experiments with six human and mouse NF-κB dimers (c-Rel–c-Rel, RelA-RelA, p52-p52, p50-p50, c-Rel–p50, RelB-p52) for an initial, comprehensive survey of potential κB site sequences. DNA-binding site motifs derived from the uPBM experiments were in agreement with published SELEX data on cRel-cRel, RelA-RelA and p50-p50 homodimers[13] (**Supplementary Fig. 1**), demonstrating highly specific binding in our assay.

The uPBM platform assesses binding to 8-bp sequences; however, the canonical κB site is 10 bp long[6,7]. Therefore, we created a custom NF-κB PBM containing 10-bp sequences prioritized according to the uPBM 8-bp data (**Supplementary Methods**). We compiled the 1,000 top-scoring 10-bp sequences determined for the six NF-κB dimers

into a list of 3,285 nonredundant sequences that represent the top-scoring set of potential κB site sequences. We incorporated these 10-bp κB sites into a custom NF-κB PBM with each site situated within constant flanking sequence (**Fig. 1a**, Online Methods).

We initially examined the binding of RelA-p50 to our custom NF-κB PBM. To assess significance of the results, we transformed the natural log of the median PBM signal intensity of each 10-bp site into a z score using the scores from 1,200 randomly chosen 10-bp sites as a background distribution (**Fig. 1b**; Online Methods). Many potential κB sites, including a set of validated κB sites, scored significantly higher than the background distribution (z score > 4), indicating that the custom PBMs show the specific binding sites on DNA for NF-κB dimers. Thus, our custom NF-κB PBM provides a platform for assessing the DNA-binding specificities of different NF-κB dimers for a large set of potential κB site sequences.

Three distinct DNA-binding classes

To examine the DNA-binding preferences of different NF-κB dimers, we carried out custom NF-κB PBM experiments for ten dimers from mouse or human. We compared the DNA-binding specificities of different dimers by correlating their κB site z scores. Hierarchical clustering showed that the NF-κB dimers separated into three distinct classes: p50 or p52 homodimers; heterodimers; and c-Rel or RelA homodimers (**Fig. 1c**). This subdivision is similar to the basic division of the NF-κB family members into two subclasses on the basis of protein sequence of the Rel-homology domains: p50 and p52 (subclass 1) and c-Rel, RelB and RelA (subclass 2). The common binding specificity we observed for the heterodimers is suggestive of a canonical DNA-binding contribution from members of each NF-κB subclass.
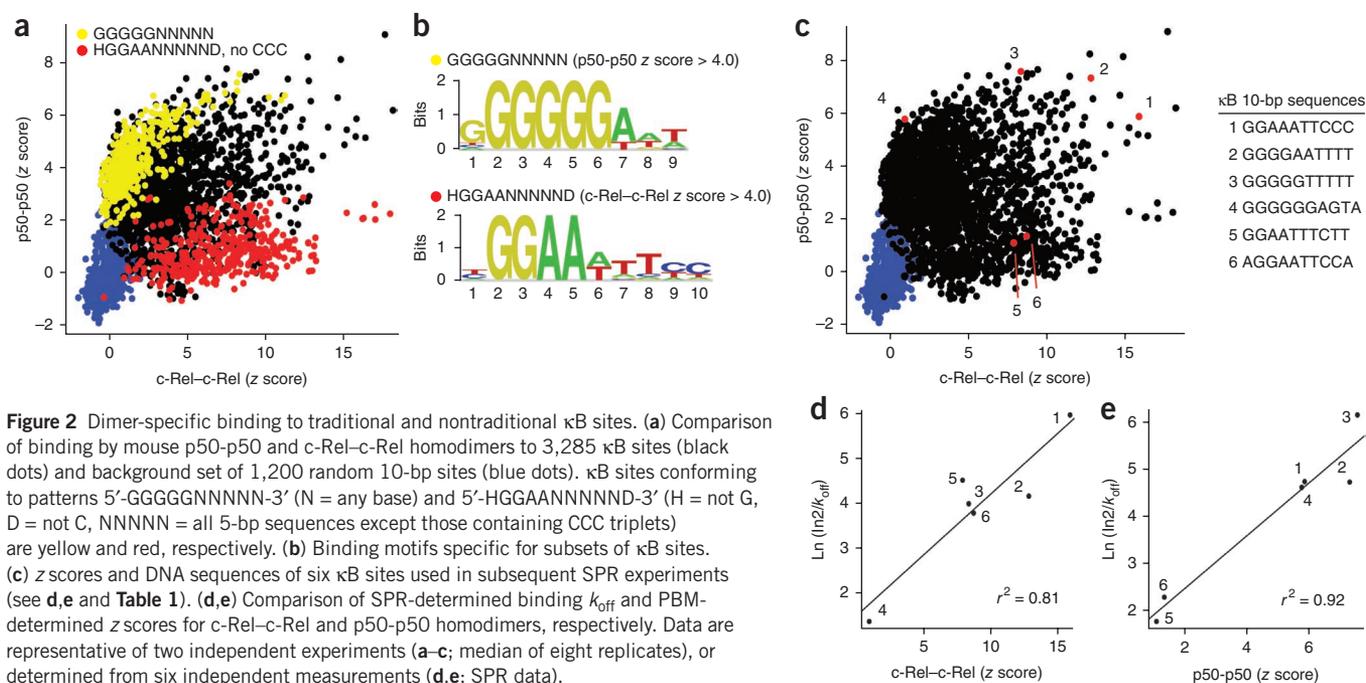
Figure 2 Dimer-specific binding to traditional and nontraditional κB sites. (**a**) Comparison of binding by mouse p50-p50 and c-Rel–c-Rel homodimers to 3,285 κB sites (black dots) and background set of 1,200 random 10-bp sites (blue dots). κB sites conforming to patterns 5′-GGGGGNNNNN-3′ (N = any base) and 5′-HGGAANNNNND-3′ (H = not G, D = not C, NNNNN = all 5-bp sequences except those containing CCC triplets) are yellow and red, respectively. (**b**) Binding motifs specific for subsets of κB sites. (**c**) z scores and DNA sequences of six κB sites used in subsequent SPR experiments (see **d,e** and **Table 1**). (**d,e**) Comparison of SPR-determined binding k_{off} and PBM-determined z scores for c-Rel–c-Rel and p50-p50 homodimers, respectively. Data are representative of two independent experiments (**a–c**; median of eight replicates), or determined from six independent measurements (**d,e**; SPR data).

To highlight the differences between these three NF-κB classes, we constructed a representative DNA binding-site motif for each class (**Fig. 1d**; **Supplementary Fig. 2** for all individual motifs). We identified different DNA binding-site motif lengths for each class: 9 bp for c-Rel and RelA homodimers, 10 bp for heterodimers and 11 or 12 bp for p50 and p52 homodimers. Although our custom NF-κB PBM was designed to assay binding to a large collection of 10-bp sequences, our *de novo* motif–finding approach identified a longer motif for p50 and p52 homodimers; we examine length preferences more directly below. The 10-bp motif for the heterodimer class is in agreement with the known NF-κB consensus sequence 5′-GGGRNWYYCC-3′, demonstrating that we correctly identified the known high-affinity binding sites. The variant 9-bp motif used for homodimer classes c-Rel and RelA and the 11-bp motif used for homodimer classes p50 and p52 also agree with the reported DNA-binding preferences of these different homodimers[6]. We summarized the binding landscape for all dimers to the 10-bp κB sites (**Supplementary Fig. 3a**), along with example genomic regions in which our PBM data are used to annotate dimer preferences for putative κB sites (**Supplementary Fig. 3b,c**).

In our pairwise comparisons, all heterodimers had a common DNA-binding specificity. Notably, RelB-containing heterodimers had DNA-binding specificity similar to that of c-Rel- and RelA-containing heterodimers. We discuss the biological relevance of this finding below (Discussion and **Supplementary Discussion**).
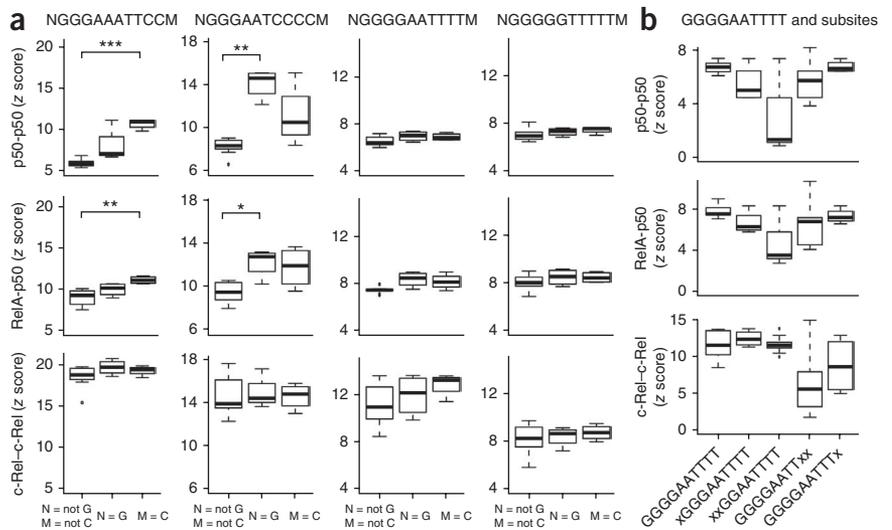
Dimer preferences for traditional and nontraditional κB sites
Dimer-specific DNA-binding preferences provide a mechanism for NF-κB dimers to target distinct binding sites and thus to regulate distinct target genes. We observed the most distinct DNA-binding preferences (the lowest z-score correlation) between members of the two homodimer classes (**Fig. 1c**). To investigate these differences further, we compared the binding specificities of the most dissimilar dimers, p50-p50 and c-Rel–c-Rel (z-score correlation r= −0.13; **Fig. 2a**). We observed many off-diagonal features that correspond to sites bound preferentially by one of the dimers ('dimer-preferred' κB sites).

To identify sequence features that could explain the relative dimer preferences, we examined the c-Rel–c-Rel-preferred κB sites (κB sites with p50-p50 z score < 2 and c-Rel–c-Rel z score > 4). We found that a majority had a 5′-HGGAA-3′ half-site (**Fig. 2a**, red dots). Many of these sequences conform to the canonical c-Rel–RelA-preferred 9-bp binding site with two 5′-HGGAA-3′ half-sites separated by a 1-bp spacer[12,28]. However, we also found many nontraditional κB site sequences with only one 5′-HGGAA-3′ half-site. We use the term "nontraditional" in lieu of "noncanonical" to avoid potential confusion with variant κB sites (referred to as noncanonical) reported to be downstream of the noncanonical NF-κB signaling pathway[16]. Nontraditional sites are sites that score poorly using the widely used NF-κB PWMs[13,29] (**Supplementary Methods**); examples include CD28RE from the *Il2* and *Csf2* gene promoters (5′-GGAATTTCT-3′, c-Rel–c-Rel z score = 8.5) and the κB site from the mouse *Plau* gene promoter (5′-GGAAAGTAC-3′, c-Rel–c-Rel z score = 12.9)[5]. We also found that a motif constructed from the c-Rel–c-Rel-preferred κB sequences had a degenerate half-site (**Fig. 2b**). Therefore, although the highest scoring c-Rel–c-Rel-preferred sites are pseudo-symmetric (**Fig. 1d**), many nontraditional, c-Rel–c-Rel-preferred sites (and RelA-RelA-preferred sites, **Supplementary Fig. 4**) scored significantly above background yet have only a single canonical 5′-HGGAA-3′ half-site.

We examined the p50-p50-preferred κB sites (p50-p50 z score > 4, c-Rel–c-Rel z score < 2) and found several κB sites with a G-rich 5′ half-site. Highlighting the κB sites that conform to the pattern 5′-GGGGGNNNNN-3′ (**Fig. 2a**, yellow dots), we observed a strong p50-p50 preference, although a subset of the sites were also bound well by c-Rel–c-Rel (discussed further below). A motif constructed from the G-rich sites bound well by p50-p50 (z score > 4) showed a 5′-GGGGG-3′ half-site and a degenerate 3′ half-site (**Fig. 2b**), although we observed a moderate preference for adenine and thymine bases 3′ to the guanine run. Therefore, as in the c-Rel–c-Rel-preferred sites, we observed statistically significant binding to a large group of κB sites defined by a single half-site sequence (z score > 4; $P < 10^{-4}$).

Figure 3 Preferences for flanking DNA bases and κB site length. (**a**) z score distributions for 10-bp κB sites with different flanking bases (for example, identity of N and M in NGGGAATCCCCM). In each panel, left bar, scores for κB sites with no 5′ guanine (forward orientation, N = not G; reverse complement orientation, M = not C); middle bar, scores for κB sites with 5′ guanine (N = G); right bar, scores for κB sites with 5′ guanine in reverse complement orientation (M = C). κB sites for which a 5′ guanine flanking base (column 2 or 3) led to significantly higher z scores (P value < 0.01, one-tailed Student's t-test) are indicated (***$P < 10^{-4}$, **$P < 10^{-3}$, *$P < 10^{-2}$). Data are representative of PBM experiments carried out for p50-p50, RelA-p50 and c-Rel–c-Rel. (**b**) z score distributions for nontraditional 10-bp κB site 5′-GGGGAATTTT-3′ and shorter variant sites. Score distribution for 10-bp sites are as in **a**. Score distributions for shorter sites were determined by examining scores from all κB sites in our data set that contained the subsite sequence. For example, column 2, labeled xGGGAATTTT, has scores from the four κB sites where x = A, C, G or T. Data are representative of single experiments (**a**,**b**; median of eight replicates) for each unique sequence. Mean and s.d. determined for four (**a**, middle and right bars; **b**, columns 2,5), nine (**a**, left bar), or sixteen (**b**, columns 3,4) sequence variants, or eight sequence replicates (**b**, column 1).

To ensure that the observed dimer-specific binding to nontraditional κB sites was not an artifact of our PBM approach, we examined binding to a set of traditional and nontraditional κB sites using SPR (**Fig. 2b**–**d** and **Supplementary Fig. 5**). Owing to the very fast on-rates (k_{on}) of some dimers, we did not obtain reliable k_{on} measurements. However, we obtained reliable off-rate (k_{off}) values and found agreement between the SPR-determined k_{off} values and our PBM-determined z scores (**Fig. 2d**,**e**). These results are consistent with earlier reports[25] showing differential off-rates as the major contributor to differences in binding affinity between κB sites. Our data demonstrate that our PBM-determined z scores reflect equilibrium binding measurements and lend further support to the potential regulatory importance of the many nontraditional κB sites in our data set.

Dimer preferences for κB sites of different lengths
DNA-binding studies[13] and X-ray crystal structures[6,12] have shown different κB site lengths for c-Rel and RelA homodimers (9 bp), heterodimers (10 bp), and p50 and p52 homodimers (11 bp). These length preferences are consistent with the DNA binding-site motifs we determined for each dimer class (**Fig. 1a**). However, because of the number of nontraditional binding sites in our data set, we sought to determine whether binding site length preferences depend on the binding site sequence itself.

We examined how the DNA bases flanking 10-bp κB sites affect binding to different dimers. Because p50-p50 binds an 11-bp site, we expected to observe a strong effect owing to flanking base identity. We measured binding by PBM to all 16 κB site variants in which the bases immediately 5′ and 3′ of the 10-bp site were exhaustively sampled (**Fig. 3a**). We examined binding of p50-p50, RelA-p50 and c-Rel–c-Rel to traditional κB site sequences, and we observed greater binding by p50-p50 and RelA-p50 with the addition of 5′ guanine to one strand (**Fig. 3a**, columns 1 and 2). We interpret these differences in terms of known half-site preferences, with high-affinity binding occurring on 11-bp sites with symmetrically opposed, optimal 5-bp half-sites (5′-GGGGA(A)TCCCC-3′ and 5′-GGGAA(A)TTCCC-3′). However, for p50-p50, the highest-affinity binding occurred with 5′ guanines flanking both half-sites,

indicating a preference beyond the 5-bp half-site and showing that a 12-bp site can be differentiated from an 11-bp site.

In contrast, we observed that binding of all three dimers was unaffected by the identity of the bases flanking nontraditional κB sites (**Fig. 3a**, columns 3 and 4). This indicated that there may be a different mode of protein-DNA interaction for nontraditional κB sites. To determine their lengths, we used our PBM data set to examine binding to shorter κB sites. For example, to interrogate binding to a 9-bp subsequence of the 5′-GGGGAATTTT-3′ site, we examined binding to the four κB sites in our data set of the form 5′-NGGGAATTTT-3′. Binding of p50-p50 and RelA-p50 was insensitive to base identity at position 5 (positions numbered as 5′-$G_{-5}G_{-4}G_{-3}A_{-2}A_{-1}T_{+1}T_{+2}T_{+3}T_{+4}T_{+5}$-3′) and only moderately sensitive to the base identity at position –5 (**Fig. 3b**). Therefore, the length of this nontraditional site is 9 bp (5′-GGGGAATTT-3′), which differs from the 9- to 11-bp traditional κB site sequences (**Fig. 3a**). In contrast, c-Rel–c-Rel binding is insensitive to positions –5 and –4 but is sensitive to base identity at position 5, demonstrating a similarly short 8-bp site length but with binding to a different subsequence (5′-GGAATTTT-3′). We observed the same preferences for the 5′-GGGGGTTTTT-3′ site (**Supplementary Fig. 6**). Our data indicate that fundamentally different modes of binding may mediate the recognition of traditional versus nontraditional sites and that this difference may translate into κB sites of different lengths. Furthermore, we observed that the length of the binding site is dimer-specific.

We analyzed the role of flanking bases to 17 additional κB sites and similarly found that 5′ guanine bases were the most predictive of greater binding affinity for all NF-κB dimers, and that the effect of a 5′ guanine depended on the G content in adjacent bases. We generated a linear model to predict PBM scores for all 12-bp κB sites on the basis of our set of 10-bp κB sites (**Supplementary Methods**). We demonstrate the efficacy of this extended data set below in our comparison of PBM data with chromatin immunoprecipitation (ChIP) data, and we make the full data set available for use in genomic analysis (see **Supplementary Spreadsheet 1**).

Affinity versus specificity of c-Rel and RelA homodimers
The Rel homology regions (RHRs) of c-Rel and RelA are more similar to each other than are the RHRs of any other pair of NF-κB family members[30].

Table 1 SPR-determined dissociation half-life values ($t_{1/2}$) for different NF-κB dimers and κB sites

	Sequence	Probe ID	p50-p50 (s)	c-Rel–c-Rel (s)	RelA-RelA (s)	RelA/N3,4–RelA/N3,4 (s)
1	GGAAATTCCC	p65-7	120 (27)	367 (54)	45 (1)	115 (17)
2	GGGGGAATTTT	p40 original	113 (12)	64 (2)	7 (2)	25 (1)
3	GGGGGTTTTT	PBM2	437 (77)	58 (11)	7 (1)	28 (6)
4	GGGGGGAGTA	PBM3	91 (27)	4 (1)	4 (1)	8 (5)
5	GGAATTTCTT	CD28RE	6 (0.2)	92 (10)	3 (1)	39 (2)
6	AGGAATTCCA	PBM1	9 (2)	46 (8)	1 (0.3)	22 (3)

Half-life values, directly proportional to dissociation off-rates ($t_{1/2} = \ln(2)/k_{off}$, s.d. in parentheses) for six different 10-bp κB site sequences (**Fig. 2**), and for four different mouse NF-κB dimers (columns 3–6). The variant RelA/N3,4–RelA/N3,4 is a homodimer of the described RelA mutant[17].

In vitro DNA-binding studies have demonstrated highly similar binding specificities, although c-Rel homodimers seem to bind a broader range of κB site sequences than RelA[13]. We observed highly correlated binding of c-Rel and RelA homodimers over our large set of ~3,300 κB sites (**Figs. 1** and **3** and **Supplementary Fig. 2**). Despite these similarities, c-Rel and RelA can elicit distinct biological functions *in vivo*[6,17]. c-Rel homodimers can preferentially activate the mouse *Il12b* gene by binding nontraditional NF-κB sites with much higher affinity than RelA homodimers[17]. Within the c-Rel RHR, 46 residues were found responsible for enhanced binding affinity, and a chimeric RelA protein containing these residues rescued *Il12b* expression in *Rel*[-/-] macrophages.

To determine the relationship between our PBM profiles and binding affinity, we carried out SPR with six different DNA sequences. For all sequences tested, except for one that bound poorly to both dimers, we observed much slower k_{off} values for c-Rel homodimers than for RelA homodimers (**Table 1**). Notably, swapping these 46 residues of the c-Rel RHR domain into RelA (protein RelA/N3,4) led to substantially slower dissociation rates (**Table 1**). The binding specificity of RelA/N3,4 homodimers also was more highly correlated with that of c-Rel homodimers (Pearson $r = 0.87$) than with that of RelA homodimers (Pearson $r = 0.82$; **Fig. 4**). However, these specificity differences are subtle in comparison to the global difference in binding affinity distinguishing c-Rel from RelA homodimers (median fold difference of 8.7 for c-Rel versus RelA k_{off} values). Thus, although the DNA-binding specificities of RelA and c-Rel homodimers are highly correlated, c-Rel homodimers have much slower off-rates than RelA homodimers, leading to a higher overall affinity and contributing to the selective regulation of c-Rel-dependent genes.

These results raised the question whether binding affinities could be used to discriminate other NF-κB dimers with correlated binding specificities (**Table 2**). We carried out SPR experiments with the six different DNA sequences using mouse p50-p50 homodimers and c-Rel–p50, RelA-p50, RelB-p50 and RelB-p52 heterodimers (**Supplementary Table 1**). We did not observe differences of the same magnitude as those found for c-Rel and RelA homodimers. The most notable difference was that c-Rel–p50 heterodimers showed slower off-rates with some DNA sequences than the other heterodimers. Although the magnitudes of these differences were smaller than those observed with c-Rel and RelA homodimers (median k_{off} fold difference of 8.7 for the c-Rel versus RelA homodimer, compared with pairwise heterodimer differences of 1.1–5.0), the results raise the possibility that enhanced binding affinity allows c-Rel–p50 heterodimers to selectively regulate some genes.

Comparison with *in vivo* binding data

We examined the relationship between our PBM-derived binding data and available genome-scale ChIP data sets with respect to *in vivo* occupancy of RelA and p50 (refs. 31–33). We found highly significant enrichment of high-scoring PBM-determined κB site sequences within ChIP-enriched (that is, dimer-bound) regions (**Supplementary Fig. 7**).

Moreover, we found significant enrichment when traditional κB sites were masked from the genomic sequence. These results demonstrate that both traditional and nontraditional κB sites in our PBM data set represent binding sequences used *in vivo*.

To determine whether PBM-determined dimer-specific differences are correlated with dimer-specific binding differences *in vivo*, we examined an NF-κB ChIP data set[33] in which array-based chromatin immunoprecipitation (ChIP-chip) was done on lipopolysaccharide (LPS)-stimulated human macrophages for all five NF-κB proteins. Focusing on p50, which had the largest number of bound regions, we separated regions into those bound by p50 only (regions bound by p50-p50 homodimers, **Fig. 5a**) and those also bound by at least one of RelA, c-Rel or RelB (regions bound by p50 heterodimers or multiple dimers, **Fig. 5a**). We used this analysis to examine whether particular κB sequences distinguish regions bound only by p50-p50 homodimers and whether our PBM data for different dimers captured these sequence differences.

The ~8,000 human promoter regions in the ChIP data set[33] were scanned with our PBM-determined 12-bp κB site sequences and were assigned the z score of the top-scoring κB site (**Supplementary Methods**). We used receiver-operating characteristic (ROC) curve analyses to quantify whether the p50-bound regions (true positives) scored higher than the unbound regions (true negatives). We observed large area under the ROC curve (AUC) enrichment scores for all three dimers tested (**Fig. 5b–d**). However, we observed that p50-p50 PBM data yielded significantly higher enrichment scores for the regions bound only by p50 than for the regions bound by additional NF-κB members (AUC = 0.84 versus 0.67, **Fig. 5b**).
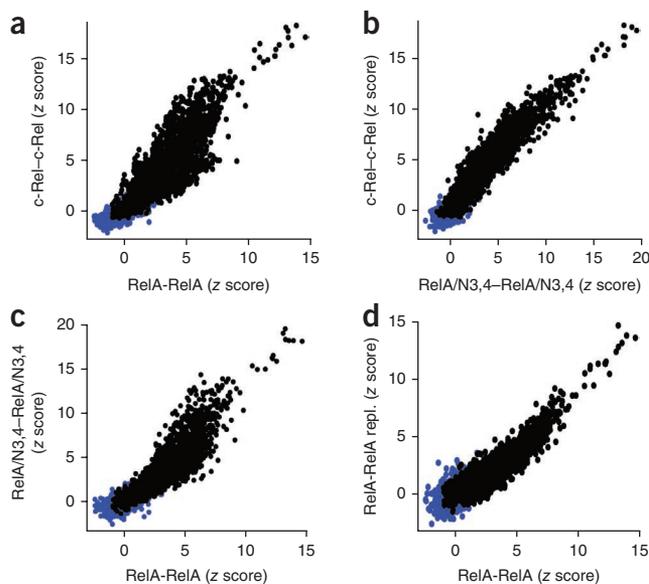

Figure 4 Comparison of c-Rel, RelA/N3,4 and RelA homodimer DNA-binding specificity. (**a**) Comparison of binding by mouse c-Rel–c-Rel and RelA-RelA homodimers to 3,285 κB sites (black dots) and background set of 1,200 random 10-bp sites (blue dots). (**b**) Comparison for c-Rel–c-Rel and RelA/N3,4–RelA/N3,4. (**c**) Comparison for RelA/N3,4–RelA/N3,4 and RelA-RelA. (**d**) Comparison for RelA-RelA (replicate experiment) and RelA-RelA. Data are representative of two independent experiments (**a–d**; median of eight replicates).
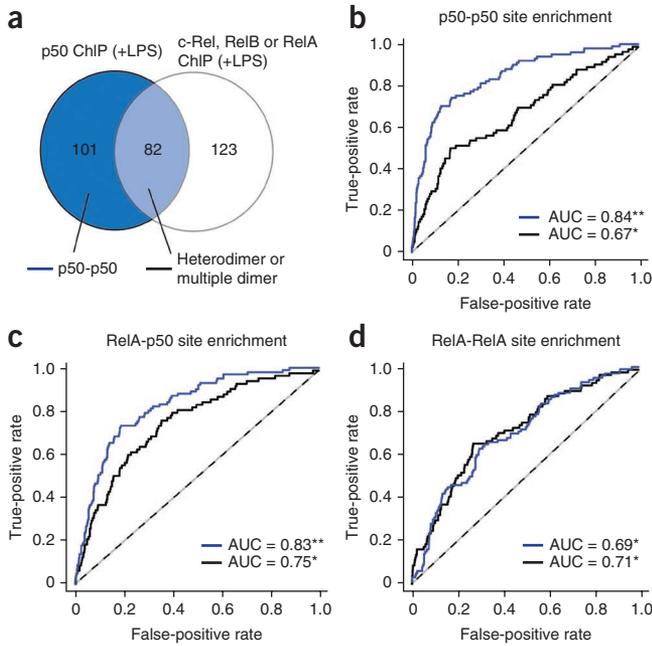
Figure 5 Enrichment of PBM-determined κB sites in published data set of p50-bound genomic regions from LPS-stimulated human macrophages[33]. (**a**) Venn diagram of overlap of 183 p50-bound regions with 205 regions bound by c-Rel, RelB or RelA. Bound regions are ChIP-enriched regions ($P < 0.002$) reported in **Figure 1** of ref. 33. (**b–d**) ROC curve analyses quantifying enrichment within p50-bound regions of PBM-determined κB sites for p50 (**b**), RelA-p50 (**c**) and RelA (**d**). ROC curves describe enrichment within p50-specifically bound regions (blue line) and within regions bound by p50 and at least one of c-Rel, RelB or RelA (black line). AUC values quantify enrichment (Wilcoxon-Mann-Whitney U test; $**P < 10^{-29}$, $*P < 10^{-7}$). Data are representative of analyses performed using measurements from single experiments (**b–d**; median of eight replicates).

In contrast, the RelA PBM data showed no discrimination between the two types of regions (**Fig. 5d**). This indicates that κB sequence features can discriminate the regions that specifically bound p50 and that these features are most highly correlated with p50-p50 homodimer PBM data. We carried out the same analysis with PBM-determined 10-bp or 11-bp κB sites but did not observe the same discriminatory capacity for p50-p50 (data not shown), indicating that the p50-bound sites are discriminated primarily by p50-p50 preferences for 12-bp κB sites. These results demonstrate that the PBM-derived, dimer-specific binding differences (summarized in **Table 2**) relate directly to dimer-specific binding differences *in vivo*.

DISCUSSION

A complete understanding of NF-κB dimer DNA-binding specificities and affinities would provide insight into mechanisms available for dimer-specific function in the cell. In this study, we examined the DNA-binding preferences of ten NF-κB dimers from mouse and human for a wide-ranging set of 3,285 potential κB site sequences. We anticipate that this large and detailed data set of κB sites will prove useful for analyses of NF-κB regulatory elements at a genome scale.

Our results have immediate biological and mechanistic implications for each of the three dimer classes. We found that c-Rel homodimers bound with substantially higher affinity than RelA homodimers to all κB sites, despite highly correlated binding profiles. This is suggestive of an affinity-dependent mechanism for discriminating these homodimers

in which c-Rel homodimers outcompete RelA homodimers for κB sites *in vivo* on the basis of DNA-binding affinity. Under these conditions, RelA homodimers do not preferentially bind to any κB sites. Therefore, to selectively regulate genes in cells that also express c-Rel homodimers (primarily hematopoietic cells), RelA homodimers need to rely on mechanisms other than selective DNA-binding, such as RelA-dependent coactivator interactions[34,35].

For the heterodimer class, we found that the selective functions of each heterodimer may not be achieved via dimer-specific recognition of κB motifs in target genes. The PBM data for all heterodimers was highly correlated, indicating that they recognize the same sequences. Notably, binding data for RelB-p52 was highly correlated with those of the other heterodimers. RelB-p52, but not RelB-p50 and RelA-p50, has been suggested to bind well to the nontraditional mouse BLC-κB site[16]. RelB-p52 is less discriminatory than RelA-p50 and can bind to a broader set of κB site sequences[20]. Binding sites unique to RelB-p52, the primary dimer activated in response to the alternative NF-κB pathway[18,19], would provide a mechanism for cells to differentiate target genes of the alternative NF-κB pathway from those of the classical pathway activating RelA-p50 (ref. 6).

However, additional studies report that RelB-p52 and RelA-p50 share highly similar binding specificities, with no clear preference shown by RelB-p52 (ref. 21). We found that RelB-p52 and RelB-p50 do not differ significantly in their DNA-binding preferences. Furthermore, we found that all NF-κB heterodimers have common binding preferences to the ~3,300 κB site sequences examined in this study. Modest binding differences reported for heterodimers[20] may be below the resolution of our approach, may prove functionally important *in vivo* and should be examined in greater depth in the future. However, this and other work[21] indicate that the regulation of distinct sets of genes by different heterodimers is likely to be achieved primarily through alternative mechanisms, such as dimer-specific interactions with co-regulatory proteins[34], dimer-specific synergy with other transcription factors or dimer-specific conformational differences[20,36,37].

We defined a subset of κB site sequences bound preferentially by the p50 and p52 homodimers. Notably, these sequences include a recently described G-rich p50 homodimer recognition motif found upstream of the interferon-inducible *Gbp1* gene (5′-GGGGGAAAAA-3′, p50-p50 z score = 6.2; c-Rel–c-Rel z score = 4.5) that mediates

Table 2 Principles of regulatory specificity for NF-κB dimer classes

Dimer class	Specificity principles
c-Rel–c-Rel, RelA-RelA homodimers	• Highly correlated binding profiles • Selective activation by c-Rel–c-Rel achieved via enhanced binding affinity[17] • Selective activation by RelA-RelA may require interactions with co-regulatory proteins[34]
Heterodimers	• Highly correlated binding profiles • Selective activation by each heterodimer may require interactions with co-regulatory proteins
p50-p50, p52-p52 homodimers	• Highly correlated binding profiles • Nontraditional, G-rich sites support preferential binding by these dimers and confer dimer-specific regulatory functions[38]
All dimers	• Binding to varied DNA sites (for example, sites of different lengths or one degenerate half-site) that is correlated with structural differences of DNA-bound complexes may facilitate allosteric regulatory mechanisms[36,37]

p50 homodimer-dependent repression[38]. This indicates that many other nontraditional, G-rich, p50-p50-preferred κB sites in our data set may similarly function as p50-p50-specific target sites *in vivo*.

In addition to the broad principles summarized above, our results highlight complexity in NF-κB–DNA interactions. First, we observed that DNA binding-site motifs for statistically significantly bound sites (z score > 4; $P < 10^{-4}$) showed one strong half-site but a degenerate preference for the opposing half-site. This contrasts with more symmetric motifs derived from the highest-affinity κB sites. Second, we observed that nontraditional κB sites seem shorter (8–9 bp long) than traditional κB sites (9–11 bp long). These findings suggest a more modest requirement for a κB site: one traditional half-site sequence is recognized via a stereotyped pattern of amino acid–base contacts, whereas a second half-site can show considerable plasticity. These results are consistent with structural analyses showing considerable plasticity in both the global conformation of the protein–DNA complex and the amino acid–base contacts mediated by dimer subunits[12,28,39]. Analyses of c-Rel and RelA homodimers bound to different κB sequences showed stereotyped amino acid–base interactions with the consensus 5′-GGAA-3′ half-site common to each structure, but highly variable contacts with the half-site sequences that differed between the structures. We propose that structural plasticity afforded by the ability of NF-κB dimers to bind to many κB sites with only a single strong half-site provides a mechanism to partially disentangle DNA binding from structural conformation. This, in turn, may allow greater structural diversity and the potential for allosteric mechanisms in transcriptional control, as have been reported[36,37].

As well as highlighting the challenge of understanding how DNA sequence may influence NF-κB conformation and the functional consequences of NF-κB binding, our results emphasize the importance of the relationship among binding specificity, affinity and function. Our data set shows NF-κB binding to a notably diverse range of sequences, and suggests that many functionally important sequences (for example, the *Il2* CD28RE) may diverge considerably from the optimal κB site. Furthermore, we observed highly overlapping binding specificity of NF-κB dimers and considerable potential for competitive binding. In reporter assays, high-affinity binding sites for NF-κB and other factors lead to stronger transcription than low-affinity sites[40]. However, in a physiological setting in the context of native chromatin, it is unknown whether an affinity threshold must be achieved for function, whether a simple relationship exists between affinity and transcriptional output, or what may be the role of co-regulatory proteins and other DNA-bound transcription factors. Our results take a step toward addressing these fundamental questions. Unlike basic consensus sequences and PWMs, which show preferences at each position of a recognition motif, data provided by PBMs and other high-throughput methods[41,42] show preferences throughout the continuum of possible binding sequences. These data sets will be useful for detailed analyses of the DNA sequence dependence of transcriptional regulatory control.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/natureimmunology/.

Note: Supplementary information is available on the Nature Immunology website.

AUTHOR CONTRIBUTIONS
T.S. designed and carried out PBM experiments and carried out ChIP data analysis. T.S. and B.A. carried out PBM data analyses. A.B.C. and K.J.W. made mouse protein samples. A.B.C. carried out SPR experiments. A.T., D.W., J.R. and I.A.U. provided human protein samples. The manuscript was written by T.S., A.B.C., S.T.S. and M.L.B.

1. Baldwin, A.S. Jr. Series introduction: the transcription factor NF-κB and human disease. *J. Clin. Invest.* **107**, 3–6 (2001).
2. Tak, P.P. & Firestein, G.S. NF-κB: a key role in inflammatory diseases. *J. Clin. Invest.* **107**, 7–11 (2001).
3. Zhang, G. & Ghosh, S. Toll-like receptor-mediated NF-κB activation: a phylogenetically conserved paradigm in innate immunity. *J. Clin. Invest.* **107**, 13–19 (2001).
4. Hiscott, J., Kwon, H. & Genin, P. Hostile takeovers: viral appropriation of the NF-κB pathway. *J. Clin. Invest.* **107**, 143–151 (2001).
5. Natoli, G., Saccani, S., Bosisio, D. & Marazzi, I. Interactions of NF-κB with chromatin: the art of being at the right place at the right time. *Nat. Immunol.* **6**, 439–445 (2005).
6. Hoffmann, A., Natoli, G. & Ghosh, G. Transcriptional regulation via the NF-κB signaling module. *Oncogene* **25**, 6706–6716 (2006).
7. Natoli, G. Tuning up inflammation: how DNA sequence and chromatin organization control the induction of inflammatory genes by NF-κB. *FEBS Lett.* **580**, 2843–2849 (2006).
8. Bonizzi, G. & Karin, M. The two NF-κB activation pathways and their role in innate and adaptive immunity. *Trends Immunol.* **25**, 280–288 (2004).
9. Hayden, M.S. & Ghosh, S. Signaling to NF-κB. *Genes Dev.* **18**, 2195–2224 (2004).
10. Gerondakis, S. *et al.* Unravelling the complexities of the NF-κB signalling pathway using mouse knockout and transgenic models. *Oncogene* **25**, 6781–6799 (2006).
11. Hoffmann, A., Leung, T.H. & Baltimore, D. Genetic analysis of NF-κB/Rel transcription factors defines functional specificities. *EMBO J.* **22**, 5530–5539 (2003).
12. Chen, F.E. & Ghosh, G. Regulation of DNA binding by Rel/NF-κB transcription factors: structural views. *Oncogene* **18**, 6845–6852 (1999).
13. Kunsch, C., Ruben, S.M. & Rosen, C.A. Selection of optimal κB/Rel DNA-binding motifs: interaction of both subunits of NF-κB with DNA is required for transcriptional activation. *Mol. Cell. Biol.* **12**, 4412–4421 (1992).
14. Udalova, I.A., Mott, R., Field, D. & Kwiatkowski, D. Quantitative prediction of NF-κB DNA-protein interactions. *Proc. Natl. Acad. Sci. USA* **99**, 8167–8172 (2002).
15. Hoffmann, A. & Baltimore, D. Circuitry of nuclear factor κB signaling. *Immunol. Rev.* **210**, 171–186 (2006).
16. Bonizzi, G. *et al.* Activation of IKKα target genes depends on recognition of specific κB binding sites by RelB-p52 dimers. *EMBO J.* **23**, 4202–4210 (2004).
17. Sanjabi, S. *et al.* A c-Rel subdomain responsible for enhanced DNA-binding affinity and selective gene activation. *Genes Dev.* **19**, 2138–2151 (2005).
18. Senftleben, U., Li, Z.W., Baud, V. & Karin, M. IKKβ is essential for protecting T cells from TNFα-induced apoptosis. *Immunity* **14**, 217–230 (2001).
19. Xiao, G., Harhaj, E.W. & Sun, S.C. NF-κB-inducing kinase regulates the processing of NF-κB2 p100. *Mol. Cell* **7**, 401–409 (2001).
20. Fusco, A.J. *et al.* NF-κB p52:RelB heterodimer recognizes two classes of κB sites with two distinct modes. *EMBO Rep.* **10**, 152–159 (2009).
21. Britanova, L.V., Makeev, V.J. & Kuprash, D.V. *In vitro* selection of optimal RelB/p52 DNA-binding motifs. *Biochem. Biophys. Res. Commun.* **365**, 583–588 (2008).
22. Berger, M.F. & Bulyk, M.L. Universal protein-binding microarrays for the comprehensive characterization of the DNA-binding specificities of transcription factors. *Nat. Protoc.* **4**, 393–411 (2009).
23. Berger, M.F. *et al.* Compact, universal DNA microarrays to comprehensively determine transcription-factor binding site specificities. *Nat. Biotechnol.* **24**, 1429–1435 (2006).
24. Mukherjee, S. *et al.* Rapid analysis of the DNA-binding specificities of transcription factors with DNA microarrays. *Nat. Genet.* **36**, 1331–1339 (2004).
25. Linnell, J. *et al.* Quantitative high-throughput analysis of transcription factor binding specificities. *Nucleic Acids Res.* **32**, e44 (2004).
26. Berger, M.F. & Bulyk, M.L. Protein binding microarrays (PBMs) for rapid, high-throughput characterization of the sequence specificities of DNA binding proteins. *Methods Mol. Biol.* **338**, 245–260 (2006).
27. Bulyk, M.L., Huang, X., Choo, Y. & Church, G.M. Exploring the DNA-binding specificities of zinc fingers with DNA microarrays. *Proc. Natl. Acad. Sci. USA* **98**, 7158–7163 (2001).

28. Chen, Y.Q., Sengchanthalangsy, L.L., Hackett, A. & Ghosh, G. NF-κB p65 (RelA) homodimer uses distinct mechanisms to recognize DNA targets. *Structure* **8**, 419–428 (2000).

29. Grilli, M., Chiu, J.J. & Lenardo, M.J. NF-κB and Rel: participants in a multiform transcriptional regulatory system. *Int. Rev. Cytol.* **143**, 1–62 (1993).

30. Li, Q. & Verma, I.M. NF-κB regulation in the immune system. *Nat. Rev. Immunol.* **2**, 725–734 (2002).

31. Lim, C.A. *et al.* Genome-wide mapping of RELA(p65) binding identifies E2F1 as a transcriptional activator recruited by NF-κB upon TLR4 activation. *Mol. Cell* **27**, 622–635 (2007).

32. Kasowski, M. *et al.* Variation in transcription factor binding among humans. *Science* **328**, 232–235 (2010).

33. Schreiber, J. *et al.* Coordinated binding of NF-κB family members in the response of human cells to lipopolysaccharide. *Proc. Natl. Acad. Sci. USA* **103**, 5899–5904 (2006).

34. Wang, J. *et al.* Distinct roles of different NF-κB subunits in regulating inflammatory and T cell stimulatory gene expression in dendritic cells. *J. Immunol.* **178**, 6777–6788 (2007).

35. Merika, M., Williams, A.J., Chen, G., Collins, T. & Thanos, D. Recruitment of CBP/p300 by the IFNβ enhanceosome is required for synergistic activation of transcription. *Mol. Cell* **1**, 277–287 (1998).

36. Fujita, T., Nolan, G.P., Ghosh, S. & Baltimore, D. Independent modes of transcriptional activation by the p50 and p65 subunits of NF-κB. *Genes Dev.* **6**, 775–787 (1992).

37. Leung, T.H., Hoffmann, A. & Baltimore, D. One nucleotide in a κB site can determine cofactor specificity for NF-κB dimers. *Cell* **118**, 453–464 (2004).

38. Cheng, C.S. *et al.* The specificity of innate immune responses is enforced by repression of interferon response elements by NF-κB p50. *Sci. Signal.* **4**, ra11 (2011).

39. Chen, Y.Q., Ghosh, S. & Ghosh, G. A novel DNA recognition mode by the NF-κB p65 homodimer. *Nat. Struct. Biol.* **5**, 67–73 (1998).

40. Mauxion, F., Jamieson, C., Yoshida, M., Arai, K. & Sen, R. Comparison of constitutive and inducible transcriptional enhancement mediated by κB-related sequences: modulation of activity in B cells by human T-cell leukemia virus type I tax gene. *Proc. Natl. Acad. Sci. USA* **88**, 2141–2145 (1991).

41. Wong, D. *et al.* Extensive characterization of NF-κB binding uncovers noncanonical motifs and advances the interpretation of genetic functional traits. *Genome Biol.* **12**, R70 (2011).

42. Fordyce, P.M. *et al.* *De novo* identification and biophysical characterization of transcription-factor binding sites with microfluidic affinity analysis. *Nat. Biotechnol.* **28**, 970–975 (2010).

ONLINE METHODS

Preparation of protein samples. Mouse sequences for RelA, c-Rel, p50, p52 and RelB were cloned into a modified pET11a expression vector for purification. Constructs contained the RHR of each subunit: RelA(1–314), c-Rel(1–282), p50(1–429) and RelB(1–400). The p50 subunit had a C-terminal FLAG tag. Proteins were expressed in BL21 (DE3) *Escherichia coli* cells (0.1 mM IPTG induction) for 16 h at 25 °C. Heterodimer subunits were coexpressed using a bicistronic expression plasmid[43]. Protein purification was done on a Q-Sepharose High Performance anion-exchange column (GE Healthcare) and a SP Sepharose High Performance cation-exchange column (GE Healthcare) and was analyzed by SDS-PAGE and electrophoretic mobility shift assay. The final purified protein samples were then frozen in aliquots in a storage buffer of 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5 mM DTT and 10% (vol/vol) glycerol.

Expression constructs for the human NF-κB dimers were created as described[44]. Briefly, His-tagged recombinant proteins were produced using pET vectors in BL21 (DE3) *E. coli* (Merck). Constructs contained the RHR of each subunit: RelA(1–307), c-Rel(1–285), p50(7–356), p52(4–332) and RelB(120–401). Proteins expressed were induced with 0.2 mM IPTG at 30 °C for 5 h. Cell pellets were harvested in Ni-NTA binding buffer with added EDTA-free Protease Inhibitor (Roche); they were pulse-sonicated for 2 min and debris was removed via centrifugation at 16,000*g*. We used a two-step purification procedure, first with the Ni-NTA His-Bind Resin system (Merck 70666) and then with DNA-affinity isolation of functional, DNA-binding protein. Ni-NTA purification was carried according to manufacturer's guidelines. For DNA-affinity isolation, a sample derived from 250 ml bacterial culture was processed with 0.128 µM oligonucleotides TNF promoter (biotinylated) and TNF-promoter complementary. Oligonucleotides were annealed via incubation in NEB Buffer 3 (New England BioLabs) at 94 °C for 1 min, followed by 69 cycles of 1 min with stepwise decrease of 1 °C. Preannealed oligonucleotide mixture (712.5 µl) was conjugated with streptavidin-agarose (Sigma).

Protein-binding microarray experiments and analysis. PBM experiments were carried out using custom-designed oligonucleotide arrays (Agilent Technologies). Two different PBM designs were used: all 10-bp site universal PBM (Agilent Technologies, AMADID no. 015681, 4 × 44K array format) as described[26] and a custom NF-κB PBM developed as part of this study (AMADID no. 025227, Agilent Technologies). DNA probe sequences synthesized on the custom-designed arrays are in **Supplementary Spreadsheet 1**.

Custom-designed oligonucleotide arrays (Agilent Technologies) were converted to double-stranded DNA arrays by primer extension and used in PBM experiments as described[22,26]. Protein samples were incubated on microarrays (concentrations in **Supplementary Tables 2** and **3**) for 1 h in binding buffer (10 mM Tris-HCl, pH 7.4; 0.2 µg/µl BSA, New England Biolabs B9001S; 0.3 ng/µl salmon testes DNA, Sigma, D7656; 2% nonfat dry milk, Stop & Shop brand; 0.02% (vol/vol) Triton X-100; 3 mM DTT; NaCl or KCl, salt concentrations in **Supplementary Tables 2** and **3**). Protein-bound arrays were then washed and incubated with primary antibody (see **Supplementary Tables 2** and **3**, column 4) for 20 min. For PBM experiments in which a secondary antibody was used (see **Supplementary Tables 2** and **3**, column 4) we deviated from the published protocol[22] and applied an additional wash step (0.05% (vol/vol) Tween-20/PBS for 3 min; 0.01% (vol/vol) Triton X-100 for 2 min) before 20 min secondary antibody incubation (**Supplementary Tables 2** and **3**, column 5).

Microarray scanning, quantification and data normalization were done using GenePix Pro ver. 6 (Axon) and masliner (MicroArray LINEar Regression) software as described[22,26]. For the custom NF-κB PBM, median fluorescence intensities for each 10-bp κB site were determined from the eight corresponding probes (forward and reverse complement orientations, four replicates each, **Fig. 1a**). For each PBM experiment, the median fluorescence intensity (MI) for each of the 3,285 10-bp κB sites was transformed into a *z* score using the mean (μ) and s.d. derived from the median intensity values of a background set of 1,200 randomly selected 10-bp sequences also present on the PBM; that is, $z \text{ score} = (MI - \mu)/\text{s.d.}$ Scoring of 10-bp sequences from the uPBM data is described in detail in **Supplementary Methods** and schematized in **Supplementary Figures 8** and **9**. Calculation of 12-bp sequence scores from custom PBM 10-bp sequence scores is described in **Supplementary Methods** and **Supplementary Figures 10** and **11**.

Surface plasmon resonance experiments. Sensorgrams were recorded on a Biacore T100 (GE Healthcare) using streptavidin chips (Sensor Chip SA). Biotinylated oligonucleotide probes were immobilized on the surface of the streptavidin sensor chip in running buffer (10 mM HEPES, pH 7.5, 150 mM NaCl, 3 mM EDTA, 0.005% (vol/vol) Tween 20). Protein samples were applied to the sensor chip at 50 µl/min at 10 °C and were referenced to an unmodified surface. Binding data were collected in the running buffer described above; sensor chip surface was regenerated with a 90-s pulse of 2 M NaCl followed by a 180-s pulse of the running buffer. Dissociation rates were obtained by global fitting of the real-time kinetic data using the Scrubber2 software (BioLogic Software) and a simple 1:1 binding model. Six concentrations of each NF-κB protein were used, ranging from 1 nM to 1 µM.

Comparison and clustering of NF-κB protein-binding microarray data. NF-κB dimer binding specificities were compared using Pearson correlation coefficient of κB site *z* scores. Only κB sites with a *z* score > 1 in at least one experiment were included in these calculations. Possibly redundant κB sites were ignored in the calculation if they could be explained by a higher-scoring κB site (that is, if a higher-scoring κB 10-bp site matched its probe sequence); ~500 κB sites met this criterion. Calculations were made using the R statistical software package. Hierarchical clustering and visualization of the comparison matrix (**Fig. 1c**) were done using the heatmap function in R, with a 'euclidean' distance function and a 'complete' clustering function.

DNA binding-site motif analysis. Binding motifs for universal PBM experiments were derived using the Seed-and-Wobble algorithm[22,26]. DNA binding-site motifs from top-scoring κB sites identified by custom NF-κB PBM experiments were determined by running the PRIORITY 2.1.0 motif finding algorithm[45] on the 10-bp sequences. Graphical sequence logos were generated using enoLOGOS[46].

43. Rucker, P., Torti, F.M. & Torti, S.V. Recombinant ferritin: modulation of subunit stoichiometry in bacterial expression systems. *Protein Eng.* **10**, 967–973 (1997).
44. Field, S., Udalova, I. & Ragoussis, J. Accuracy and reproducibility of protein-DNA microarray technology. *Adv. Biochem. Eng. Biotechnol.* **104**, 87–110 (2007).
45. Gordân, R., Narlikar, L. & Hartemink, A.J. Finding regulatory DNA motifs using alignment-free evolutionary conservation information. *Nucleic Acids Res.* **38**, e90 (2010).
46. Workman, C.T. *et al.* enoLOGOS: a versatile web tool for energy normalized sequence logos. *Nucleic Acids Res.* **33**, W389–392 (2005).