Structural insights into NF-κB/IκB signaling

Review Article

Gourisankar Ghosh, De-Bin Huang, and Tom Huxford
Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0359 USA.

Correspondence: Gourisankar Ghosh, Ph.D., Department of Chemistry and Biochemistry, University of California, San Diego 9500 Gilman Drive, Urey Hall, La Jolla, CA. 92093, USA. Tel: (619)-822-0469; Fax: (619)-534-7042; E-mail: gghosh@ucsd.edu

Abbreviations: RHR, rel homology region; TAD, transcriptional activation domain; IKK, IκB kinase complex; ARD, ankyrin repeat-containing domain; PEST sequence, amino acids sequence rich in proline, glutamic acid, serine and threonine; SRD, amino-terminal signal receiving domains; NLS, nuclear localization signals.

Key Words: NFκB/IκB signaling, apoptosis, mitogen-induced signaling, transcription factors,

Received: 7 August 1999; accepted: 15 September 1999

Summary

The vital cellular activities of growth, differentiation, reaction to stimuli, and apoptosis are controlled by the coordinated expression of a vast number of genes. Regulation of gene expression occurs primarily at the level of transcription. The process begins as one of a multitude of mitogen-induced signaling events triggers the intricate and exquisitely regulated signal transduction pathways. These lead ultimately to the recruitment of transcription factors on specific promoter/enhancer sites. One example of how complex cell signaling can lead to the temporal activation of transcription in a specific set of genes is illustrated by the transcription factor NF-κB. Since its discovery almost a decade and a half ago, NF-κB has fascinated researchers because of the complexity of the NF-κB signaling pathway. Several recently determined crystal structures of a number of NF-κB complexes have given a new dimension of understanding of the biochemistry behind NF-κB. We will review these structures in light of their functions.

I. Introduction

Members of the Rel/NF-κB transcription factor family are involved in diverse biological activities ranging from the regulation of inflammatory and immune responses to development and apoptosis (Ghosh et al, 1998; Baldwin, 1996; Baeuerle and Henkel, 1994; Baeuerle and Baltimore, 1996). Rel/NF-κB proteins modulate transcription by binding distinct DNA target sites that are collectively known as κB DNA sequences (Baeuerle and Henkel, 1994). In mammals, five homologous polypeptides, p50, p52, p65 (or RelA), c-Rel, and RelB constitute the Rel/NF-κB family (Figure 1a) (Ghosh et al, 1998). Common to each of the Rel/NF-κB polypeptides is an amino-terminal region of approximately 300 amino acids with high sequence homology known as the rel homology region (RHR). This segment is responsible for nuclear translocation and κB DNA binding. Similar to most other site-specific DNA-binding transcription factors, the Rel/NF-κB proteins function as dimers. The Rel/NF-κB dimerization domain is also contained within the RHR. Interestingly, the existence of many, but not all, of the possible homo- and heterodimer combinations has been observed in cells. Polypeptides of the Rel/NF-κB family can be divided into two subgroups based on the presence or absence of a transcriptional activation domain (TAD). Two family members, p50 and p52, do not contain distinct TADs and are therefore categorized as belonging to class I. Evidence exists suggesting that the homodimers of p50 and p52 and the p50/p52 heterodimer can function as repressors of gene transcription (Franzoso et al, 1992). The other three Rel/NF-κB family members, p65, c-Rel, and RelB, constitute the class II subgroup. NF-κB dimers containing one or two of these polypeptides act as activators by virtue of the presence of at least one transactivation domain. The two most abundant and biologically well characterized of the Rel/NF-κB dimers are the p50 homodimer and p50/p65 heterodimer.

In most resting cells, Rel/NF-κB dimers with transactivation potential are localized in the cytoplasm in complex with inhibitor proteins IκBα, IκBβ, and IκBε (Whiteside et al, 1997; Li and Nabel, 1997). These macromolecules belong to the larger IκB family of transcription factor inhibitor proteins which also includes
IkBγ, Bcl-3, p105, and p100 (Figure 1b) (Verma et al, 1995). The p105 and p100 proteins contain the RHR of p50 and p52, respectively, in their amino-termini and inhibit NF-κB proteins by a different mechanism than do IkBα, -β, and -ε. The inhibitory functions of IkBγ and Bcl-3 are at present not entirely clear.

Various physiological and environmental signals promote nuclear translocation of NF-κB proteins by removing inhibitory IkBα, IkBβ, and IkBε from NF-κB/IkB complexes (Figure 2a). Each of these three IkB proteins contains two conserved serines within a homologous segment in their amino-terminal signal receiving domains (SRD). These serines are phosphorylated by the multisubunit IkB kinase complex (IKK) in a signal dependent manner (DiDonato et al, 1997; Lee et al, 1998). Specific lysine residues of IkB located near the phosphorylated serines are then polyubiquinated, marking IkB for rapid degradation by the proteosome (Baeuerle and Henkel, 1994).

Figure 1. Domain organization of the NF-κB and IkB protein families.

a) The Rel/NF-κB transcription factor family polypeptides are organized according to class I and class II depending on the presence or absence of transactivation domains. The Rel homology region is indicated with the amino-terminal domain in red and dimerization domain in green. Other structural elements of interest are labeled. b) The IkB family of transcription factor inhibitor proteins are aligned according to ankyrin repeat-containing domains. The amino-terminal signal receiving domain and carboxy-terminal PEST sequence are indicated. The p105 and p100 polypeptides are special cases which contain IkBγ-like carboxy-terminal domains as well as their own amino-terminal rel homology regions.
**Figure 2.** IκBα regulates NF-κB transcriptional activity by dual mechanisms. **a)** A host of cellular stimuli activate signal transduction cascades which lead to the phosphorylation and subsequent proteolysis of NF-κB-associated IκBα. Free NFκB then enters the nucleus where it binds gene enhancers and activates transcription. **b)** Shortly after NF-κB induction (1), newly synthesized IκBα (2) accumulates in the cell cytoplasm (3). This free IκBα then enters the nucleus (4) where it removes NF-κB from target genes and shuttles it back to the cytosol (5).
Removal of IκB proteins from the NF-κB/IκB complex activates nuclear localization signals (NLS) on the NF-κB subunits. This allows for rapid nuclear translocation of the active NF-κB dimer. Interestingly, among the host of genes regulated by NF-κB activation is the gene encoding IκBα. As a result of this arrangement, a pool of newly synthesized IκBα begins to accumulate in the cell cytosol shortly after NF-κB induction (Figure 2b) (Baldwin, 1996). In the absence of NF-κB binding partner, this newly synthesized IκBα translocates to the nucleus where it binds to and dissociates pre-formed NF-κB/DNA complexes (Ghosh et al, 1998). Nuclear NF-κB/IκB complexes are then shuttled back to the cytoplasm, restoring the pre-induction state (Arenzana-Seisdedos et al, 1997). Therefore, IκBα plays an important role in the regulation of NF-κB activity by regulating NF-κB both in the cytoplasm and in the nucleus.

High resolution X-ray crystal structures containing the RHR of various Rel/NF-κB dimers in complex with DNA and the structure of the NF-κB p50/p65 heterodimer/IκBα complex have significantly extended our knowledge of the chemistry behind the NF-κB and IκB proteins.

II. Structure of the RHR

High resolution X-ray crystal structures have been determined containing the RHR of four of the five NF-κB family members namely, p50 (Ghosh et al, 1995; Müller et al, 1995; Cramer et al, 1997), p52 (Chen et al, 1998), p65 (Cramer et al, 1997; Chen Y-Q et al, 1998), and c-Rel (Chen Y-Q, unpublished data). As expected from their high degree of sequence similarity, these structures are also highly homologous. The RHR is comprised of two independent structural modules or domains and a flexible 25 amino acid segment at the carboxy-terminal (Figure 3a). This carboxy-terminal flexible segment contains a cluster of basic amino acids responsible for nuclear localization of NF-κB proteins. We refer this segment as the NLS polypeptide. Ordered polypeptide structure is not observed for this NLS polypeptide in the absence of its IκBα protein binding partner (Ghosh et al, 1995; Cramer et al, 1997; Huang et al, 1997). The two structured domains are folded into immunoglobulin-like folds. Like other immunoglobulin domains, the roughly 100 amino acids long carboxy-terminal domain exhibits only beta structure whereby two beta sheets form a globular beta-sandwich. This immunoglobulin-like domain is solely responsible for subunit association and so is also commonly referred to as the dimerization domain. Residues from the carboxy-terminal dimerization domain also participate in non-specific binding to the sugar-phosphate backbone of target DNA.

The larger, approximately 180 amino acids long amino-terminal domain determines the sequence specificity of DNA binding by mediating base-specific DNA contacts. Although this domain also bears an apparent immunoglobulin-like fold, finer analysis of the domain structure shows deviation from the canonical immunoglobulin domain. This domain also represents the region of highest sequence variation among different Rel/NF-κB family members. The differences stem primarily from a non-homologous insertion of varying lengths within the RHR amino-terminal domain. The insert is largest in the p50 and p52 subunits where it forms a two α-helix bundle connected by a large loop. In p65 and c-Rel, on the other hand, the insert is smaller forming only a single small alpha-helix. This helical insert may play a role in transcriptional activation by mediating interactions with other transcription factors or coactivators. The overall structure of the core of the RHR amino-terminal domain shows distant similarity to the DNA binding domains of the tumor suppressor p53 and the STAT family transcription factors.

The RHR amino- and carboxy-terminal domains are joined by a 10 amino acid linker. In the DNA bound conformation, these two domains contact each other. However, interdomain interactions are not conserved between RHR polypeptides and depend highly on the target DNA sequences. In the absence of any structural information for free RHR (in the absence of DNA) one can only speculate if these two domains interact. High proteolytic sensitivity of the linker and the variations of interactions between two domains in the DNA bound forms suggest that in the absence of DNA the two domains are flexible about the linker polypeptide.

III. Dimerization of NF-κB proteins

Rel/NF-κB dimers are formed by the stacking of two symmetrical beta sheets, one from each immunoglobulin-like carboxy-terminal dimerization domain, onto one another (Huang et al, 1997). The residues that participate in intersubunit contacts are highly homologous among the family members. In fact, seven of twelve dimer forming residues are identical while the other five represent conservative substitutions. Immunochemical and other in vivo and in vitro experiments suggest that Rel/NF-κB proteins form dimers in a combinatorial fashion. Interestingly, the Rel/NF-κB polypeptides show a drastic specificity for homo- and heterodimer formation. The most dramatic example is the inability of the RelB polypeptide to self associate or to form heterodimers with c-Rel and p65 subunits (Ryseck et al, 1995). In another case of NF-κB dimer selectivity, the p50/p65 heterodimer forms with greater stability than do either homodimers of p50 or p65. Against the backdrop of such a global similarity the slight energetic differences at the dimer interfaces invoke interesting regulatory interactions that probably dictate differential dimerization affinity. Alanine scanning mutagenesis of p50 dimer interface residues shows that only four residues, Tyr267, Leu269, Asp302, and Val310, contribute
significantly to the binding energy of p50 dimerization (Sengchanthalangsy LL, unpublished data). All four of these residues are identical in RelB. Results from the mutagenesis experiments, therefore, suggest that identical residues play different roles in the context of different partners.

**IV. NF-κB DNA binding**

There are over 100 DNA sequences that NF-κB dimers have been known to recognize (Baldwin, 1996). The DNA targets are collectively known as κB DNA which display strong sequence similarity. Most of the sequences are pseudosymmetric. The identifiable feature of κB DNA is the strong conservation of flanking G:C base pairs. The G:C base pair at the 5' end is even more conserved. Crystal structures of NF-κB/DNA complexes have revealed the detailed mechanism of DNA complex formation.

---

**Figure 3.** Ribbon diagrams of NF-κB complex x-ray crystal structures.

**a)** The Rel homology regions of the NF-κB p50/p65 heterodimer bound to the Ig κB DNA target site. The p50 subunit is represented in green, the p65 subunit is shown in red, the DNA double helix is depicted in cyan and magenta. Each subunit displays the characteristic Rel/NF-κB folds with two immunoglobulin-like domains connected by a short, flexible linker.

**b)** The x-ray crystal structure of the IκBα/NF-κB p50/p65 heterodimer complex. The color scheme for NF-κB is consistent with part a, IκBα is represented in purple. Note the large change in conformation made by the amino-terminal domain of p65 in the presence of the IκBα inhibitor. The p50 amino-terminal domain is modeled into this figure in its DNA-bound conformation and is not present in the x-ray crystal structure.
All NF-κB dimers recognize κB DNA targets in an overall similar mode. The most visible feature of these complexes are: (i) dimers interact with DNA targets through loops connecting secondary structures (beta strands). Five such loops are contributed by each subunit. (ii) Both domains and the flexible linker polypeptide interact with DNA targets. (iii) One turn of the DNA major groove is contacted intimately. (iv) Symmetrical homodimers recognize both symmetric and asymmetric DNA targets asymmetrically. The recognition loop, which arises from the amino-terminal domain, provides three to four residues for sequence specific DNA recognition. A cluster of basic residues from a second loop of the amino-terminal domain binds the sugar phosphate backbone from the minor groove side of the DNA. The interdomain connector loop contributes one basic residue for sequence specific DNA binding and other residues for backbone interactions. Two loops from the dimerization domain primarily recognize the sugar phosphate backbone of the DNA targets. These structures also show that each monomer recognizes a half site of DNA with sequence specificity. p50 subunit optimally binds to a five base pair half site of sequence 5'-GGGAA-3', whereas p65 prefers a four base pair 5'-GGAA-3' half site. The two half sites are separated by a non-contacted central base pair. This central base pair is often A:T. The ideal target sites of the p50 homodimer and the p65 homodimer are the 11-mer nucleotide sequence (5'-GGGGAATTTCC-3') and 9 base pair sequence (5'-GGATTTCC-3'), respectively. The p50/p65 heterodimer preferably interacts with 10 base pair targets.

In p50 the first three G:C base pairs are recognized by four conserved amino acids, Arg54, Arg56, Glu60, and His64, of which the first base pair is contacted by His64. Whereas Arg54, Arg56, and Glu60 are identical throughout the Rel/NF-κB polyepitides, His64 is replaced by an Alanine in p65. This helps to explain why p65 recognizes and binds to a shorter DNA target. The pseudodyad axis of DNA passing through the central base pair aligns with the dyad axis of the dimerization domain.

Not all the physiological DNA targets, however, follow these general rules of κB DNA sequence consensus. Some of the known physiological targets, such as the P-element of the IL-4 promoter, has only partial sequence homology to that of the consensus κB DNA. NF-κB dimers are also capable of binding to these targets with affinities comparable to "ideal" κB DNA (Rao et al, 1997). The crystal structure of p65 homodimer bound to the P-element-like target (Chen et al, 1998) shows that one subunit recognizes the non-consensus half site with an altered conformation with no sequence specific interactions with DNA. NF-κB p65 homodimer thus can bind to DNA targets with as few as only four base pairs. The modular architecture of the DNA binding domains of Rel/NF-κB factors enables them to bind DNA with such diversity. The amino-terminal DNA specificity region of a monomer, in the absence of proper nucleotide sequence in the DNA half site, may move about to find an alternate mode of DNA interaction without sacrificing too much binding energy. Consequently, Rel/NF-κB dimers are capable of binding DNA using multiple conformations. This unusual mechanism of DNA target selection allows the relatively few Rel/NF-κB family dimers to regulate a large number of κB DNA enhancers.

V. Structure of IκB proteins

The IκB family proteins contain either six or seven ankyrin repeats in a centrally located ankyrin repeat-containing domain (ARD) (Verma et al, 1995). Ankyrin repeats were first identified as a repeated 33 amino acid element in the erythrocyte membrane protein ankyrin and have since been identified in many proteins of diverse biological functions (Bork et al, 1986). The ankyrin domain in IκB is flanked by two segments. In three of the IκB family proteins, IκBα, IκBβ and IκBε, these flanking segments are homologous and likely to perform similar functions. The amino-terminal segments in IκBα, -β and -ε contain two serines that are phosphorylated by the IκB kinase. The carboxy-terminal segments are rich in the amino acids proline, glutamic acid, serine and threonine (PEST). The acidic PEST sequence is common to many proteins which display rapid turnover in the cell (Rogers et al, 1986). Serines and threonines of the IκB PEST sequences are constitutively phosphorylated by casein kinase II (McElhinny, 1996). Whereas, the role of the ankyrin domain and the carboxy-terminal PEST sequence have been shown to be essential for interactions with NF-κB dimers (Ernst et al, 1995; Malek et al, 1998), it is not yet clear what role, if any, the amino-terminal signal response domain plays in NF-κB recognition. The other IκB family members, specifically IκBγ, p105, p100, and Bcl-3, exhibit significant differences from this domain arrangement.

The three dimensional structures of the ankyrin repeat domains from several different proteins have been solved (Gorina and Pavletich, 1996; Luh et al, 1997; Venkataramani et al, 1998; Batchelor et al, 1998). Their overall structures are similar in that each 33 amino acid sequence forms the repeating ankyrin structural unit. This structural motif consists of two alpha helices, followed by a loop of variable length, and a short beta turn. Each repeat forms a layer in the stacked, approximately cylindrical domain. The beta turn in each repeat is projected in an orientation roughly perpendicular from the helices and extends like a finger. Ankyrin domains are curved displaying two distinct surfaces. The helical parts lie on the concave face and the stacked fingers form the convex surface.

The recent solution by two laboratories of the IκBα crystal structure in complex with the NF-κB p50/p65
heterodimer (Jacobs and Harrison, 1998; Huxford et al., 1998) reveals its structural similarity to other ankyrin repeat-containing structures (Figure 3b). The six ankyrin repeats of IkBa are stacked with approximately 10 Å spacing between them with a slight superhelical twist. The helical bundle in each layer interacts with the helices in the layer above and below. A conserved set of residues present in all the inner helices are critical for maintaining the structure of ankyrin repeat domains. Residues in the finger regions vary among IkB and other ankyrin proteins. Therefore, these segments are likely to be responsible for the discriminatory interactions with NF-κB dimers. In complex with NF-κB, the highly charged PEST sequence of IkBa assumes an extended conformation devoid of any distinct secondary structure.

VI. Interactions between IkBa and NF-κB p50/p65 heterodimer

Results from biochemical experiments suggest that the discriminatory and affinity-determining interactions between IkBa and NF-κB p50/p65 heterodimer are highly complex (Ernst et al., 1995; Malek et al., 1998; Latimer et al., 1998). The complicated nature of these interactions can be explained by the fact that each of the independent structural elements of both IkBa and NF-κB participate in the extensive protein-protein interface. Two interacting segments of IkBa; the ankyrin domain and the PEST sequence, do not form a single structural unit but are rather flexible with respect to each other. Similarly, five flexibly linked structural units of the heterodimer, the dimerization domains and the carboxy-terminal NLS polypeptides of p50 and p65 and the amino-terminal domain of p65, are engaged in IkBa recognition. Results from the biochemical experiments strongly imply that the amino-terminal domain of p50 is not essential for IkBa recognition.

The x-ray crystal structures of the IkBa/NF-κB p50/p65 complex determined independently in two laboratories support the model generated from biochemical experiments. The combined results of these crystallographic analyses of the complex also help to explain genetic and biochemical experiments. The most important insights that these structures provide involve the mechanisms of NF-κB cytoplasmic retention and its dissociation from κB DNA in the nucleus.

Upon binding to IkBa, the NLS polypeptide of p65 forms two successive alpha helices which contact the first two ankyrin repeats of IkBa (Jacobs and Harrison, 1998). Three of the central four basic residues comprising the NLS of p65 mediate direct salt bridges with acidic residues of IkBa. A similar type I basic NLS sequence from SV40 has been shown to interact with nuclear transport protein α-karyopherin in an extended conformation (Conti et al., 1998). The free NLS sequence can thus adopt different conformations in the context of differing protein binding partners. The NLS polypeptide of p50 also lies in close proximity to IkBa without making any direct interactions. It is important to note that Latimer et al. have shown that the SRD of IkBa is essential for masking the p50 NLS (Latimer et al., 1998). The SRD is absent in both structures which may explain the lack of direct interactions between the p50 NLS polypeptide and IkBa. Whereas, IkBa ankyrin repeats one and two are engaged in mediating interactions with the NLS polypeptides of NF-κB, the sixth repeat and the acidic carboxy-terminal PEST sequence of IkBa contact the p65 amino-terminal domain primarily through electrostatic interactions (Huxford et al., 1998). A field of negatively charged residues from these segments of IkBa interacts with the positively charged DNA binding face of the p65 amino-terminal domain. This interaction is augmented by interdomain contacts between negatively charged residues of p65 dimerization domain with its amino-terminal domain. Therefore, both intra- and intermolecular interactions involving the amino-terminal domain of p65 serve to force the NF-κB p65 subunit into a closed conformation. Adoption of this conformation renders NF-κB incapable of binding to DNA as key charged specificity and affinity-determining amino acid side chains are buried in an electrostatic sea.

VII. Conclusion and future direction

While tremendous progress has been made over the last few years in our understanding of the NF-κB signaling pathway and transcriptional regulation, a number of questions remain to be answered. Chief among these are: How does IkBa discriminate for p65 and c-Rel-containing NF-κB dimers? What is the role of the p50 amino-terminal domain in the NF-κB/IkBa complex? Does the IkBa signal response domain contribute to NF-κB binding? What is the nature of the IkBa kinase/IkBa interaction? What controls IkBa proteolytic processing? What is the exact mode for nuclear trafficking of NF-κB and IkBa? How does NF-κB interact with other transcriptional coactivators and ultimately influence basal transcription machinery? And, what roles do the remaining IkB family members play? Further biochemical and biophysical testing of the elements involved coupled with careful structural analyses will be required before questions like these can be properly addressed.

Acknowledgements

Work performed in this laboratory is supported by grants from the National Institutes of Health/NCI, the University of California AIDS Research Program (UARP), and California Career Research Program. G.G. is a Alfred P. Sloan fellow. T.H. is supported by a dissertation award from UARP. The authors wish to thank Devin Drew and Chris Phelps for help preparing the manuscript.