doi:10.1016/j.jmb.2007.04.023 J. Mol. Biol. (2007) 369, 1307–1317

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

The Role of High Affinity Non-specific DNA Binding by Lrp in Transcriptional Regulation and DNA Organization

Stacey N. Peterson[1], Frederick W. Dahlquist[1,2] and Norbert O. Reich[1,2]*

[1]*Program in Biomolecular Science and Engineering University of California Santa Barbara, CA 93106, USA*

[2]*Department of Chemistry and Biochemistry, University of California, Santa Barbara CA 93106, USA*

**Corresponding author*

Transcriptional regulatory proteins typically bind specific DNA sequences with ~10^3–10^7-fold higher affinity than non-specific DNA and this discrimination is essential for their *in vivo* function. Here we show that the bacterial leucine-responsive regulatory protein (Lrp) does not follow this trend and has a ~20–400-fold binding discrimination between specific and non-specific DNA sequences. We suggest that the dual function of Lrp to regulate genes and to organize DNA utilizes this unique property. A ~20-fold decrease in binding affinity from specific DNA is dependent upon cryptic binding sites, including the sequence $GN_{2\text{-}3}TTT$ and A-tracts. Removal of these sites still results in high binding affinity, only ~70-fold weaker than that of specific sites. Similar to Lrp's binding of specific sites in the *pap* and *ilvIH* promoters, Lrp binds cooperatively to non-specific DNA; thus, protein/protein interactions are important for both specific and non-specific DNA binding. When considering this cooperativity of Lrp binding, the binding selectivity to specific sites may increase to a maximum of ~400-fold. Neither leucine nor the *pap*-specific local regulator PapI alter Lrp's non-specific binding affinity or cooperative binding of non-specific DNA. We hypothesize that Lrp combines low sequence discrimination and relatively high intracellular protein concentrations to ensure its ability to regulate the transcription of specific genes while also functioning as a nucleoid-associated protein. Modeling of Lrp binding data and comparison to other proteins with regulatory and nucleoid-associated properties suggests similar mechanisms.

Published by Elsevier Ltd.

Keywords: Lrp; non-specific binding; transcriptional regulation; nucleoid-associated proteins; cooperative

Introduction

Escherichia coli leucine-responsive regulatory protein (Lrp) is a positive and negative regulator of several *E. coli* operons, including those involved in pili formation, amino acid biosynthesis and catabolism, and nutrient transport.[1,2] Leucine influences the regulatory effect of Lrp in some operons, including *ilvIH*, *sdaC*, and *fan*, whereas no such regulation occurs in others like *pap* and *fae*.[1,2] The systematic evolution of ligands by exponential enrichment (SELEX)-selected degenerate Lrp consensus recognition sequence, (C/T)AG(A/C/T)A(A/T)ATT(A/T)T(A/G/T)CT(A/G), has been useful for identifying binding sites in bacterial promoters,[3,4] but is not always predictive of Lrp binding sites. A portion of this consensus sequence, $GN_{2\text{-}3}TTT$, exists in the six *pap* Lrp binding sites where N is any nucleotide.[5] The Lrp homolog LrpC prefers binding to bent regions containing A-tracts rather than sequences similar to the SELEX-selected consensus sequence.[6]

Lrp is an 18.8 kDa protein that forms higher order structures (octamers and hexadecamers) at micromolar concentrations and in the presence of leucine tends to favor the octamer configuration.[7,8] Leucine increases Lrp's affinity for certain DNA sequences *in vivo*.[9] Lrp binds its consensus sequence and site 2 in the *ilvIH* promoter as a dimer at nanomolar concentrations.[10] Due to its structural organization

Abbreviations used: Lrp, leucine-responsive regulatory protein; H-NS, histone-like nucleoid structuring protein; NS, non-specific; Dam, DNA adenine methyltransferase; SELEX, systematic evolution of ligands by exponential enrichment.

E-mail address of the corresponding author: reich@chem.ucsb.edu

0022-2836/$ - see front matter. Published by Elsevier Ltd.

(higher-order oligomer formation),[8] large and variable DNA recognition sequence,[3] and high cellular concentrations,[11,12] Lrp also functions as a nucleoid-associated protein, similar to the histone-like nucleoid structuring protein (H-NS) and the factor for inversion stimulation (Fis) protein.[2,13] Similar to H-NS, Lrp causes significant DNA conformational changes, observed in the *ilvIH* promoter.[14] Structural models of Lrp/DNA interactions based on crystal structures of the Lrp-like proteins AsnC and LrpC, suggest that the Lrp octamer binds a ~100 bp DNA segment by looping the DNA around the protein core.[15]

In an effort to understand how Lrp and the DNA adenine methyltransferase (Dam) regulate transcription of the *pap* operon, our preliminary data suggested that Lrp bound non-specific and specific DNA with comparable affinities (unpublished data). This relatively tight binding to non-specific DNA in conjunction with Lrp's role in DNA organization led us to examine the non-specific binding affinity of Lrp more quantitatively. Here we show that Lrp binds cooperatively to non-specific DNA with only ~20–400-fold less selectivity than specific binding, supporting its role as a nucleoid-associated protein. The non-specific binding affinity does not change in the presence of leucine or the *pap*-specific regulator PapI. Using models developed for classic gene regulatory proteins,[16] we show how a protein with less than a 400-fold preference for its specific site can still regulate transcription within specific promoters.

Results

Here we describe Lrp's interaction with non-specific DNA and its discrimination between specific and non-specific DNA. We determined the non-specific binding properties of Lrp using three different DNA substrates from two plasmid sources. Each substrate had a different length and a different percentage of A/T and G/C base-pairs (see Materials and Methods). Before choosing these sequences, we verified that the SELEX-selected consensus sequence or any similar sequence was not present.

Verification of the absence of an Lrp consensus sequence through BLAST search and alignments

Two non-specific DNA substrates were initially selected for use in the binding analyses: a 246 base-pair sequence from pTZ19U (246mer NS) and a 326 base-pair sequence from pBR322 (326mer NS). We verified that the SELEX-selected 15 base-pair Lrp consensus sequence[3] was not present in these non-specific substrates by performing BLAST searches, which resulted in no significant hits. Identification of potential specific Lrp binding sites has been determined previously by calculating total binding energies of sequences that are similar to the 15 base-pair SELEX-selected consensus sequence.[3,4] To confirm that no potential specific sites were present in our DNA substrates, the 15 base-pair regions of each substrate that best aligned to the consensus sequence were selected and their binding energies were calculated as previously.[3] The aligned regions that showed the highest binding energies (positive values) are shown in Table 1. The maximum binding energy calculated was 0.44 arbitrary units, much less than the three to four arbitrary units predicted for a specific site.[3]

Lrp binds non-specific DNA cooperatively with apparent binding affinities only ~20–70 times weaker than specific DNA

To better understand Lrp's role as a transcriptional regulator, particularly in the context of *pap* regulation by Dam,[17,18] and to explain the presence of several supershifted bands in Lrp/*pap* DNA gel mobility shift assays at high Lrp concentrations (data not shown) we determined the Lrp concentration dependence of its interaction with non-specific DNA. Gel mobility shifts using two non-specific DNA substrates (246mer NS and 326mer NS) revealed two distinct Lrp/DNA complexes, one that appeared at lower Lrp concentrations and a supershifted complex seen at higher Lrp concentrations (Figure 1(a) and (c)). The binding of Lrp to form these complexes was analyzed using a phenomenological Hill analysis where the binding of Lrp to DNA to form the lower band was treated as:

$$\mathrm{DNA} + n\,\mathrm{Lrp} \rightarrow \mathrm{DNA(Lrp)_n}$$
$$\text{where } (K_{D1})^n = [\mathrm{DNA}][\mathrm{Lrp}]^n/\mathrm{DNA(Lrp)_n} \quad (1)$$

and the Hill coefficient n corresponds to the lower Lrp/DNA band. The supershifted band was treated with a similar phenomenological approach:

$$\mathrm{DNA} + m\,\mathrm{Lrp} \rightarrow \mathrm{DNA(Lrp)_m}$$
$$\text{where } (K_{D2})^m = [\mathrm{DNA}][\mathrm{Lrp}]^m/\mathrm{DNA(Lrp)_{nm}} \quad (2)$$

and m corresponds to the Hill coefficient for forming the supershifted band. The overall binding was subject to the material balance condition:

$$\mathrm{DNA_{total}} = \mathrm{DNA_{free}} + \mathrm{DNA(Lrp)_n} + \mathrm{DNA(Lrp)_m} \quad (3)$$

Since the DNA concentration was always at least tenfold lower than the Lrp concentration, the free

Table 1. 15 bp sequences from non-specific DNA substrates that best aligned to Lrp consensus sequence

		Identical bases[a]	Binding energy[b] (arb. units)
Consensus	YAGHAWATTWTDCTR	15	5.36–6.12
246mer NS	TTACACTTTATGCTT	10	0.04
326mer NS	TAGCAATTTAACTGT	9	0.44

Y = C or T, H = not G, W = A or T, D = not C, and R = A or G.
[a] Number of common bases to consensus.
[b] Determined by addition of individual estimated binding energies from Cui *et al.*;[3] values >3 are considered specific binding sites.[3]