5. Discussion

The global transcriptional regulator and partitioning protein, KorB, is known to bind DNA specifically, to function as a repressor and to interact with IncC proteins. I have determined the crystal structure of the DNA-binding domain of KorB, KorB-O, in complex with its operator, O_B. The structure reveals how two monomers bind the operator and how each monomer docks into the major groove. It provides a plausible model for other ParB homologs that bind DNA.

5.1. Thr211 and Arg240 are essential for specific KorB-O_B complex formation

To gain further insight into the determinants of KorB-DNA recognition, a number of site-specific mutations were introduced into the full-length protein in which residues from the O_B-binding site were exchanged against alanine as suggested by the crystal structure. The mutated proteins were expressed in *E. coli* and their ability to bind to the consensus target site was tested by electrophoretic mobility shift assays (Fig. 27).

Both wildtype KorB and KorB-O, containing residues 117-294, were able to shift the restriction fragment containing the consensus O_B site (Fig. 27A, lanes b-d). At the protein concentration causing specific binding to the O_B-containing fragment, prominent non-specific binding to either smaller or larger DNA fragments was not observed. Mutation of the two residues that form direct hydrogen bonds to bases in the operator site, Thr211 and Arg240, completely abolished binding to the operator DNA. In contrast, KorB proteins with single alanines substituted for residues that form non-specific contacts to the DNA backbone, either direct or water-mediated, retain the ability to shift the operator fragment (Table 6). At higher protein concentrations (15 pmol rather than 6 pmol in the reaction volume), the KorB mutant Thr211Ala partly regains its operator-binding capacity, while the Arg240Ala variant remains deficient in specific DNA binding (not shown). A Lys171Ala/Arg208Ala KorB double mutant has lost its ability for operator binding which was retained in the two protein variants carrying the single mutations.

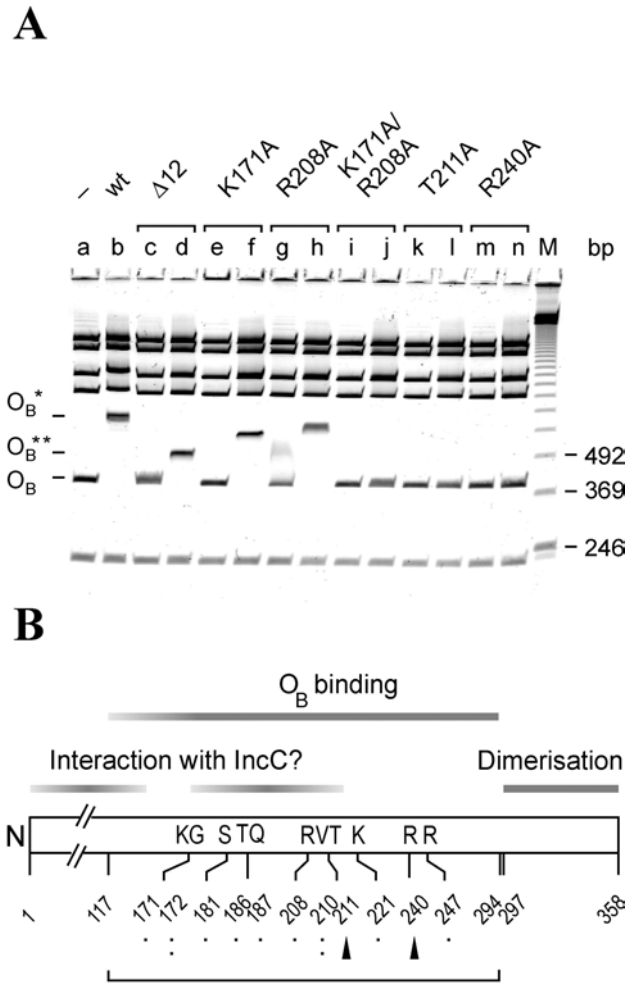

Figure 27. Mutational study of KorB-O_B binding. (**A**) Complex formation of KorB proteins with O_B-DNA was done essentially as described (Balzer *et al*., 1992). For all KorB variants, extracts of induced KorB-overproducing cells were used. Lanes are: a, without protein; b, 6 pmol KorB wt; c and d, KorBΔ12 (aa 101-294), e-n, KorB containing amino-acid substitutions as indicated. For each mutant protein, 1 pmol and 6 pmol KorB were applied. O_B, DNA fragment containing O_B; O_B*, complexes of O_B with wild-type KorB or mutant KorB containing amino acid substitutions; O_B**, complexes of O_B with KorBΔ12. M, 123-bp ladder as DNA fragment size marker. (**B**) Functional domains of KorB and schematic presentation of mutations. The bar represents KorB consisting of 358 amino acids. The bracket denotes the KorB portion used for crystallization. The positions of mutated residues are given. Dots mark amino acids contacting the DNA backbone *via* side chains. Single colons highlight residues that form hydrogen bridges between the peptide backbone and the sugar phosphate backbone, whereas wedges mark residues making side chain-base contacts. The location of known and proposed domains are shown above the bar. The scheme is not drawn to scale.

Table 6. Specific complex formation of KorB proteins with O_B DNA

KorB variant	Specific O_B binding	Location of mutation
wild type	+	
Lys171Ala	+	HTH-α3
Gly172Ala	+	HTH-α3
Ser181Ala	+	HTH-turn
Thr186Ala	+	HTH-α4
Gln187Ala	+	HTH-α4
Arg208Ala	+	turn α5-α6
Lys171Ala/Arg208Ala	–	turn α5-α6/ HTH-α3
Val210Ala	+	α6
Thr211Ala	(–) [1]	α6
Lys221Ala	+	α6
Arg240Ala	–	α8
Arg247Ala	+	α8

[1]Increasing the amount of KorB T211Ala in the reaction mixtures threefold results in a detectable and stable complex with O_B DNA.

The specific DNA binding of these mutant proteins as revealed by the gel retardation assay is completely in line with the KorB-O—DNA crystal structure if one assumes that all observed protein-DNA contacts contribute to operator binding and that specificity-determining direct protein-base contacts are of primary importance.

5.2. KorB-O—bound operator DNA adopts standard B-conformation

Recently we characterized the mode of KorB-O binding to O_B by means of Raman spectroscopy demonstrating that in solution 17-bp O_B DNA shows B-DNA character and that the secondary structure of the KorB-O fragment is mainly α-helical (Dostál *et al*., 2003). As seen in Fig. 28 the Raman difference spectrum of the complex reveals significant spectral changes, suggesting alterations in the secondary structure of KorB upon interaction with DNA and showing that the protein binds into the major groove. These data are in agreement with the crystal structure, both regarding the operator DNA structure and the general mode of KorB-O binding.

The central base pair of the operator can be either CG or GC, since it is not involved in any specific interaction. The two flanking base pairs, GC and CG on each half-site are important specificity elements in the complex. Within a standard B-DNA environment, the outer base pairs (G7-C11) show a unique geometry with near-zero propeller twist which may favour the hydrogen bonding between the guanine base and the Thr211 side chain. Whether this conformation is pre-formed in the DNA or adopted upon KorB-O binding remains unclear in the absence of a structure of the free operator fragment.

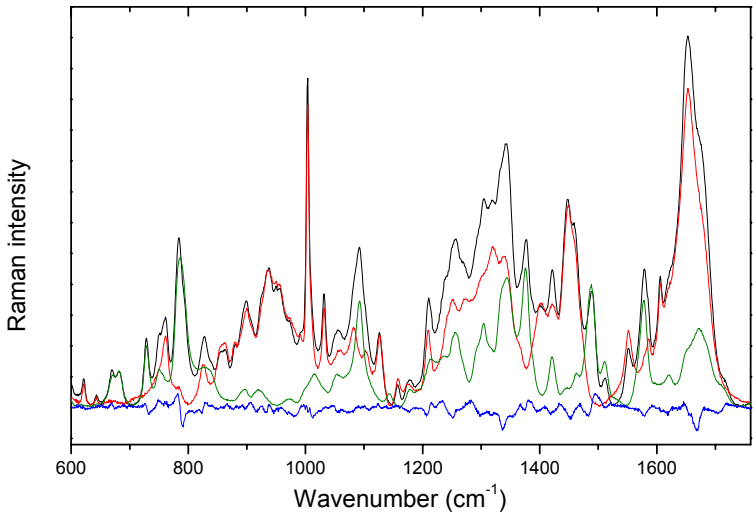

Figure 28: Raman spectrum of KorB-O in complex with 17-bp O_B (black trace). The red trace is from KorB-O, the green trace belongs to O_B. The blue trace is the difference spectrum obtained by subtraction of the individual isolated component spectra from the spectra of the complex. The spectra are normalized to represent the same amounts of protein and DNA in the complex and in free form. The features of the difference spectrum provide information about conformational changes, structural rearrangements and interactions between DNA and protein. (Dostál *et al*., 2003).

5.3. Specific DNA binding occurs outside the standard HTH motif

According to the DALI server (Holm and Sander, 1993), the KorB-O structure presents a new combination of structural domains with previously observed folds. No homologs were found for the intact monomer; however, fold similarities were observed for the N-terminal (HTH) subdomain which closely resembles members of the LuxR-FixJ family of bacterial transcription factors, such as the *Bacillus* transcription factor GerE (1FSE [Ducros *et al*., 2001]) with a Z score of 4.5 and an rms deviation of 2.0 Å, and the DNA binding domain of NarL (1A04 [Baikalov *et al*., 1996]; DALI Z-score of 4.2 and rms deviation of 2.1 Å). The N-terminal subdomain of *Drosophila* Pax6 (residues 12-58), a transcription factor resembling KorB residues 147-193 that encompass the HTH motif was amongst the results (6PAX [Xu *et al*., 1999]; DALI Z-score of 4.3 and rms deviation of 3.2 Å) (Fig. 29).

The standard DNA-binding HTH motif is highly conserved in its three-dimensional structure, with the "recognition helix" binding along the major groove of the DNA (Pabo and Sauer, 1984; Brennan and Matthews, 1989; Harrison and Aggarwal, 1990). The recognition helix may approach the major groove in various geometries, but hydrogen bonding between its side chains and DNA bases is in nearly all cases primarily important for sequence-specific protein binding. A search for structurally KorB-O related HTH motifs involved in DNA binding at http://www.ebi.ac.uk/thornton-srv/databases/DNA-motifs/ yielded a number of close matches including the Hin recombinase (1HCR), Tet repressor (2TCT) and Tc3 transposase (1TC3), all with rmsd's of about 1 Å (Fig. 29). The α3-turn-α4 segment of KorB has a sequence signature and conformation typical for an HTH motif generally involved in specific DNA-major groove interaction. However, the KorB HTH motif is involved only in unspecific DNA binding through numerous non-specific contacts to the DNA backbone (Fig. 25 and 26). Strikingly, the residue Gln187 which is conserved at the central position in the "recognition helix" in many bacterial repressors, is dispensable for DNA recognition by KorB as seen in the structure of the KorB-O—operator complex and corroborated by mutagenesis. The same applies to residues Lys171, Gly172, Lys180, Ser181, and Thr186 (Table 6). Since five of these six residues are conserved in the four KorB homologs (Fig. 33) the HTH motif is responsible for the non-sequence specific DNA binding. The structure reveals that mainly two residues also conserved in KorB homologs, Thr211 and Arg240 (Fig. 33), recognize the O_B sequence through direct hydrogen bonding. In agreement with these data, mutant KorB

proteins Thr211Ala and Arg240Ala do not recognize O_B specifically, strongly suggesting that the mode of operator binding by the KorB-O fragment observed in the crystal reflects the specific DNA binding of the wildtype protein in solution.

In Fig. 29 a structural overlay KorB-O with HTH motifs from some transcription factors described above is shown.

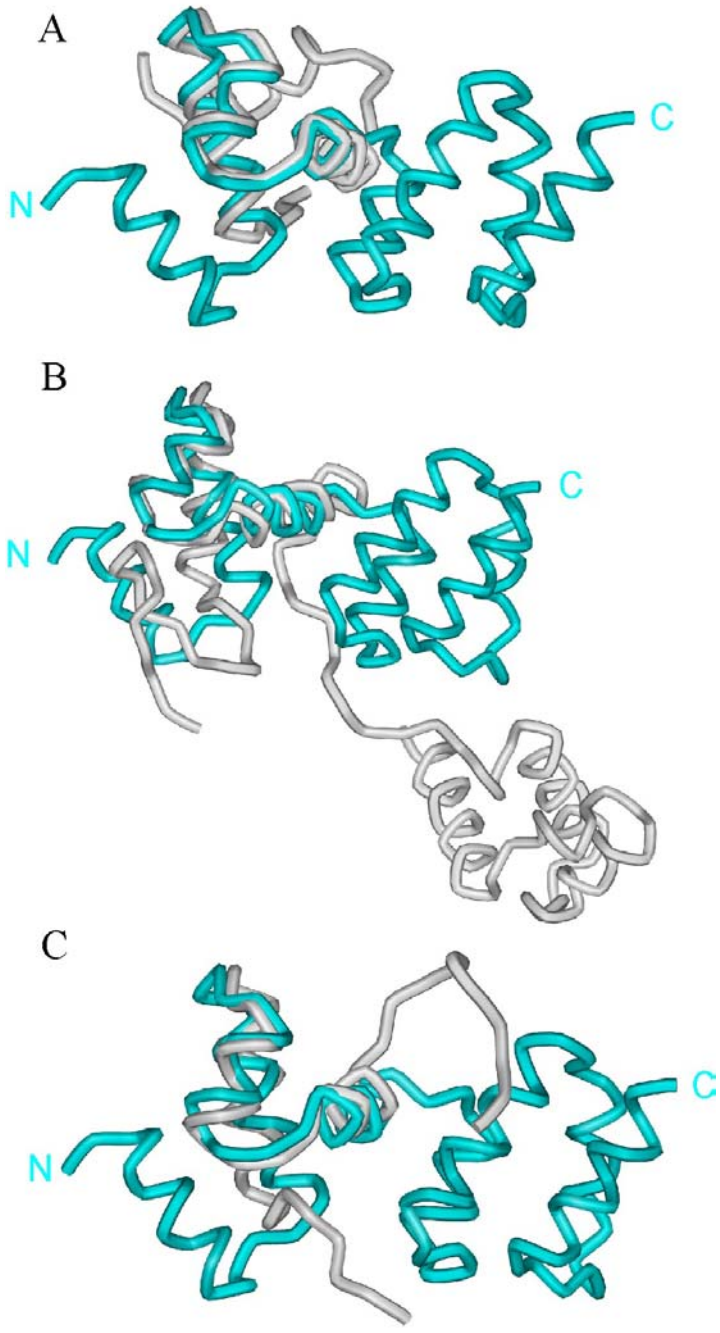

Figure 29. Structural similarity of the N-terminal HTH motif of KorB-O and three other HTH motifs of DNA-binding proteins. The KorB-O monomer (blue) is aligned with *Bacillus* transcription factor GerE (**A**), *Drosophila* Pax6 (**B**) and Hin recombinase (**C**) in grey.

5.4. KorB traps the operator sequence

The oligonucleotide containing the palindromic O_B site is bound symmetrically by two KorB-O subunits that contact the major groove and sugar-phosphate backbone of the DNA. This general structural organization has been observed in many bacterial repressor-operator complexes (Pabo and Sauer, 1984; Brennan and Matthews, 1989; Harrison and Aggarwal, 1990). Since KorB-O represents only a fragment of KorB, it is worth considering the possible DNA-binding mode of the full-size protein (Fig. 30). Each of the two KorB-O molecules in the complex binds to the upper and lower half-site of O_B, respectively. In this arrangement, the DNA is trapped like in a clamshell, and each KorB-O covers the DNA helix of each half-site from the major-groove side only. Previous foot-printing data using full-length KorB demonstrated protection of the entire O_B sequence (Balzer *et al*., 1992; Williams *et al*., 1993). The KorB-N domain for which no structural information is available is likely to be connected to helix 1 of KorB-O *via* a flexible linker as indicated by residues 117-137 which are disordered in the crystal structure of the KorB-O—DNA complex. Therefore KorB-N could be arranged in a way that allows it to cover the accessible portion of O_B, thus causing full protection of the DNA.

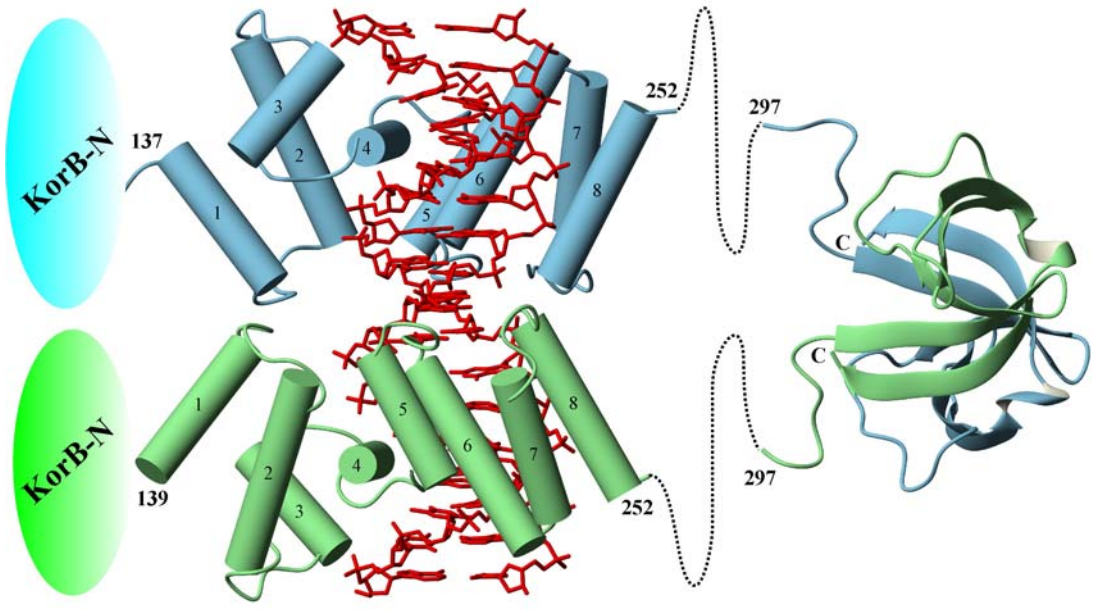

Figure 30. Schematic model for operator binding by intact KorB. The crystal structure of the dimerization domain KorB-C is known (Delbrück *et al*., 2002; 1IGQ). The KorB-O structure is connected by dotted lines to the KorB-C dimer. No structural model exists for the N-terminal domains, KorB-N, shown as ellipsoids.

In the KorB-O—operator complex, the protein subunits weakly interact through contacts centered around Gln148 located at the C-terminus of helix α1 and within the loop connecting α1 and α2, but do not interact C-terminally (Fig. 24). As expected from the crystal structure, KorB-O runs as a monomer over a gel filtration column and *in vitro* chemical cross-linking with glutaraldehyde fails to provide evidence for multimeric protein forms. Thus, the weak contacts between operator-bound KorB-O subunits apparently do not favour dimerization in the absence of DNA.

Intact KorB dimerizes *via* KorB-C (Balzer *et al*., 1992; Delbrück *et al*., 2002). KorB-O and KorB-C are connected *via* a linker which is probably flexible, since residues 253-296 are disordered in the KorB-O—DNA crystal. As a consequence of the dimeric KorB-C domain and the interactions between DNA-bound KorB-O subunits described above, the operator DNA is completely enclosed within the dimeric KorB. In addition, the KorB-N domain may engage in additional DNA contacts as discussed above, thus further tightening the grip of KorB on the bound DNA. For this complex to form on a circular plasmid, flexibility of the KorB-O—KorB-C linker is clearly required, since the protein has to undergo a conformational change upon binding that results in complete engulfment of the DNA helix (Fig. 30). As we have no information regarding the position and orientation of KorB-C relative to KorB-O and the operator DNA, we cannot exclude the possibility that the C-terminal domain may also be involved in DNA protection, either in combination with KorB-O or with KorB-O and KorB-N.

5.5. Model for IncP plasmid segregation

The current model for active partition of bacterial plasmids assumes that a nucleoprotein complex at the centromer site is formed containing the site-specific DNA-binding protein (ParB/KorB) and the ATP-hydrolyzing protein (ParA/IncC) (Austin 1988, Møller-Jensen *et al*., 2000) to facilitate the formation of plasmid pairs or the positioning of the plasmids in the cell. How does the IncP partitioning system with ParA and ParB homologs fit with these observations? In two separate studies the IncP ParA homolog, the IncC protein with putative NTPase activity, was shown to interact with KorB both by yeast two-hybrid analysis and by *in vitro* studies with partially purified proteins (Lukaszewicz *et al*., 2002; Rosche *et al*., 2000). There does not seem to be a general rule for prediction of the location of the domain of ParB which interacts with ParA, because in different systems it has been

located at the N-terminal, central or C-terminal parts of various ParB proteins (Surtees and Funnell, 1999; Kim and Shim, 1999). Lukaszewicz *et al*. (2002) localized the 'IncC-binding domain' on KorB to the 45 amino-acid patch from Ile174 to Thr218 that belongs to the DNA-binding region. The surface of KorB-O that is not involved in DNA interaction is acidic (Fig. 23 & 24). Since the IncC polypeptides are highly basic with calculated pIs higher than 10, the basic side chains are likely to form stable ion pairs (or salt bridges) to the predominantly acidic KorB-O domain (Fig. 31).

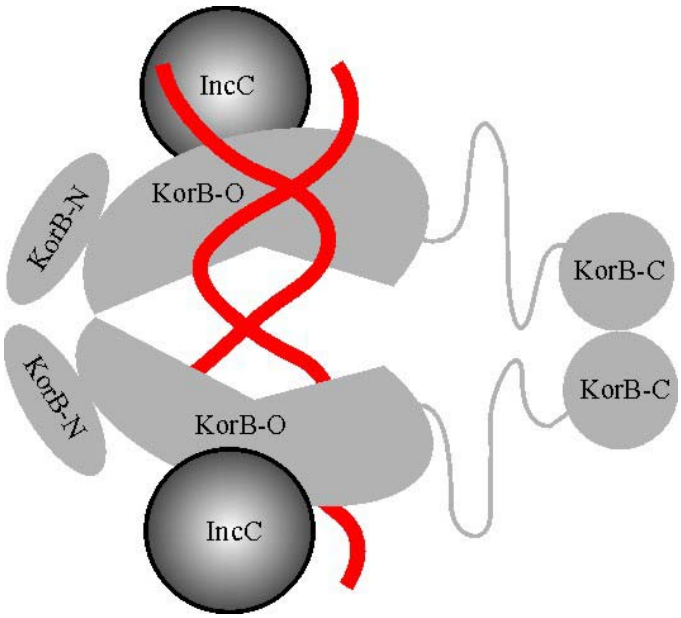

Figure 31. Interaction of the ParA homolog, IncC, of RP4 with O_B bound KorB. The IncC binding surfaces of DNA bound KorB point in opposite directions allowing two IncC molecules to bind one KorB-O--O_B complex.

A working model for the plasmid partitioning cycle was suggested by Bignell & Thomas (2001) and is described in Fig. 32. However, whether the nucleoprotein complex forms at the membrane or in the cytoplasm and which role NTP binding/hydrolysis plays in the partition process remain open questions for the IncP system.

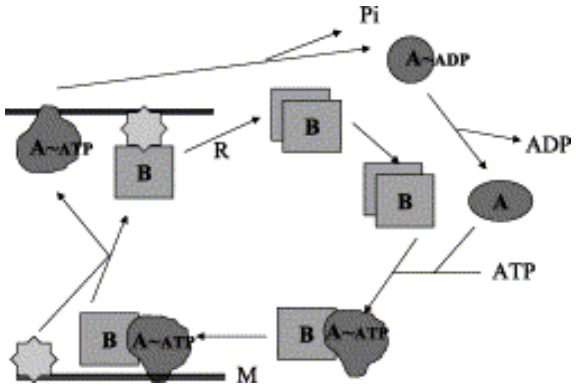

Figure 32. Working model of the partition cycle involving ParA (A) and ParB (B). Plasmid DNA attached to B is not shown. Replication at the membrane (R) generates duplicated and possibly paired B partition complexes. A binds to ATP and then binds to B, adopting a conformation which favours binding to the membrane (M). Facilitated diffusion takes place until the complex encounters the unknown protein (star), which defines the new replication zone. B becomes anchored to this protein and A~ATP is displaced. A~ATP hydrolyses to A~ADP and dissociates from the membrane. ADP dissociates and A is ready to re-enter the cycle. (Adopted from Bignell & Thomas, 2001)

5.6. KorB-O structure a model for other ParB homologues

Comparison of KorB proteins from IncPα (RP4) and IncPβ reveals that the proteins are very similar (Fig. 33). The two residues determining specificity Thr211 and Arg240 are conserved in the IncPα and IncPβ members. Based on such high degree of similarity, it appears reasonable to assume that these KorB homologues might have similar structures and similar mode of binding. KorB when compared with other ParB members (chromosomal homologues) shows similarity but not to such high extent, and only Arg240 seems to be conserved (Fig. 19). Whether these ParB members have similar folds like KorB and interact with DNA in the same way remains to be determined.

Figure 33: Sequence comparison of KorB proteins and secondary structure of RP4 KorB-O. The primary structures of homologs from four IncP plasmids are shown. Residues identical to those of RP4 KorB are represented by dots. Amino acids with side chains of identical chemical character conserved in all four proteins are highlighted in light blue, whereas residues identical in three KorB proteins are marked yellow. The segment 137-252 observed in the RP4 KorB-O crystal contains helices α1 to α8 schematically given above the RP4 KorB amino acid sequence. α3 and α4 (blue) comprise the helix-turn-helix motif, the residues Thr211 and Arg240 (marked red) are essential for specific operator binding. Accession numbers are: RP4 (IncPα), L27758; R751, U67194; pADP-1, U66917; pB4, AJ431260. The latter three plasmids belong to the IncPβ group.