J. Biosci., Vol. 21, Number 5, September 1996, pp 613–629. © Printed in India.

Alignment and phylogenetic analysis of type II DNA topoisomerases

K MADHUSUDAN and V NAGARAJA*
Centre for Genetic Engineering, Indian Institute of Science, Bangalore 560 012, India

MS received 6 February 1996; revised 14 August 1996

Abstract. **DNA topoisomerases have been evolved to solve the topological problems of DNA during replication, transcription, recombination and segregation. Discovery of several new enzymes and their characterization has necessitated this compilation. This analysis shows the distinct evolutionary relatedness of type II DNA topoisomerases. A striking feature is the absence of a contiguous stretch of about 160 amino acids in one of the subunits of prokaryotic type II enzymes, which might have important implications to their structure and function.**

Keywords. **Type II DNA topoisomerases; DNA gyrase; multiple alignment; phylogenetic trees.**

1. Introduction

DNA topoisomerases catalyse topological interconversions: supercoiling-relaxation, catenation-decatenation and knotting-unknotting of DNA. These topological events occur during important cellular processes such as replication, transcription, recombination and chromosome segregation. Thus, the enzymes are essential for the cell survival, and hence are ubiquitous. The topoisomerases are classified into two distinct subclasses based on the mechanistics of the reaction (Wang 1985; Maxwell and Gellert 1986). The type I topoisomerases break one strand of DNA and pass the other stand through the nick created and change the linking number in steps of one. On the other hand, type II enzymes cleave both stands of DNA and pass the duplex through the 'DNA gate' resulting in the change of linking number in steps of two. All known topoisomerases form a transient covalent intermediate with DNA through a phosphotyrosine linkage and reseal after strand passage. Both prokaryotes and eukaryotes have been shown to possess multiple topoisomerases, possibly evolved to provide division of labour and in certain cases as backup strategies to take care of important cellular functions. The bacterium, *Escherichia coli* contains two type II topoisomerases bes ides two type I enzymes; the yeast, *Saccharomyces cerevisiae* has two type I activities (Wallis *et al* 1989), and in humans, two isozymes of topoisomerase II have been reported (Jenkins *et al* 1992). Amongst all type II topoisomerases, only DNA gyrase has the ability to introduce negative supercoils into DNA in an ATP driven reaction (Gellert *et al* 1976). The heteromeric enzyme has been the subject of extensive study (Reece and Maxwell 1991). The second bacterial type II enzyme, topoisomerase IV, has strong decatenation and weak relaxation activities. Both the bacterial type II topoisomerases have a similar architecture and also, share considerable sequence similarity.

*Corresponding author (Fax, 91-80-3341683; E-mail, vraj@cge.iisc.ernet.in).

In this compilation, we have focused our attention on type II topoisomerases. This is due to the wealth of information available on these enzymes, their indispensability and the degree of conservation amongst the genes from variety of organisms. On the other hand, type I topoisomerases are less conserved and only few genes are characterized. Since an elaborate discussion of sequence conservation of topoisomerases has earlier been presented (Caron and Wang 1993), we have emphasized on the evolutionary relationship.

2. Sequence alignment and generation of evolutionary trees

The topoisomerase II genes have been characterized from several bacteria, yeast, protozoans and higher animals. Table 1 summarizes the source and the length of the derived polypeptides. The polypeptide sequences were aligned by MACAW ver 2.0.3 using BLOSSUM62 (Schuler *et al* 1991; Lawrence *et al* 1993). This software allows manual editing of the alignment. Also, the GyrB and GyrA polypeptide sequences were fused manually and aligned with eukaryotic type II topoisomerase using Multalin (Corpet 1988). The multiply aligned sequences were subjected to PHYLIP analysis (Felsenstein 1989). The distance matrix was generated by PROTDIST of PHYLIP employing Kimura-2 parameter. The output was then analysed by NEIGHBOR applying Neighbor-joining method. The UPGMA method was used for NEIGHBOR analysis of gyrase and eukaryotic topoisomerase II alignment. The unrooted trees were generated using DRAWGRAM and DRAWTREE.

3. Results and discussion

This compilation and alignment of type II topoisomerases is an attempt to compile complete sequences, identify subclasses and determine the extent of phylogenetic relationships. Sequence information on DNA gyrase and eukaryotic type II topoisomerase genes has been accumulating in the databank. These reports show conservation of amino acid sequence in gyrase and also its partial homology with eukaryotic type II topoisomerases. Hence, we have presented the alignment of all deduced polypeptide sequences of type II topoisomerases in figures 1 and 2. In order to avoid errors in alignment and phylogeny analyses, we have omitted partial sequences. The alignment of A subunits of gyrase and topoisomerase IV, given in figure 1A, shows the high sequence conservation predominantly in the amino terminal region. The DNA breakage-reunion site of subunit A has the sequence AAMRYTE common to all the members. The residue Tyr-122 of *E. coli* GyrA, present in this sequence gets covalently attached to DNA through phospodiester bond. On the other hand, the C-terminal region does not show such extensive conservation. The dot matrix analysis, however, showed repeated sequences within this region in all GyrA sequences (Madhusudan and Nagaraja 1995). The C-terminal 33 kDa domain of *E. coli* GyrA has been shown to bind DNA (Reece and Maxwell 1991).

The subunit B of bacterial type II topoisomerases shows identical patches of amino acids scattered through out the sequence. The N-terminal 43 kDa fragment of *E. coli* GyrB is known to retain ATPase activity, a characteristic of all type II topoisomerases. The crystal structure of this domain complexed with ADPNP has revealed the direct interaction between the protein and the cofactor (Wigely *et al* 1991). These contact

Table 1. Polypeptide sequences of type II DNA topoisomerases.

Protein	Acronym	length	SwissProt/ *EmBL Acc. no.	Reference
GyrA				
Escherichia coli	EcoA	875	P09097	Swanberg and Wang 1987
Klebsiella pneumoniae	KpnA	876	P14829	Dimri and Das 1990
Campylobacter jejuni	CajA	863	Q03470	Wang *et al* 1993
Rickettsia prowazekii	RprA	905	P41080	Wood and Waite 1994
Bacillus subtilis	BsuA	821	P05653	Moriya *et al* 1985
Staphylococcus aureus	SauA	889	P20831	Margerrison *et al* 1992
Streptomyces coelicolor	ScoA	864	P35885	Calcutt 1994
Haloferax Aa2.2	HalA	858	*X60178	Holmes and Dyall-Smith 1994
Mycobacterium tuberculosis	MtbA	838	*L27512	Takiff *et al* 1994
Mycobacterium smegmatis	MsmA	854	*X84077	Madhusudan and Nagaraja 1995
GyrB				
Escherichia coli	EcoB	803	P06982	Yamagishi *et al* 1986; Adachi *et al* 1987
Neisseria gonorrhoeae	NgoB	781	P22118	Stein *et al* 1991
Pseudomonas putida	PpuB	806	P13364	Parales and Harwood 1990
Bacillus subtilis	BsuB	638	P05652	Moriya *et al* 1985
Staphylococcus aureus	SauB	640	P20832	Margerrison *et al* 1992
Haloferax sps.	HalB	639	P21558	Holmes and Dyall-Smith 1991
Mycoplasma pneumoniae	MpnB	650	P22447	Colman *et al* 1990
Spiroplasma citri	SciB	640	P34031	—
Streptomyces sphaeroides (novobiocin resistant)	SspBr	677	*Z17304	Thiara and Cundliffe 1993
Streptomyces sphaeroides (novobiocin sensitive)	SspBs	684	*Z17305	Thiara and Cundliffe 1993
Streptomyces coelicolor	ScoB	676	P35886	Calcutt 1994
Mycobacterium tuberculosis	MtbB	675	*X78888	Madhusudan *et al* 1994
Mycobacterium smegmatis	MsmB	675	*X84077	Madhusudan and Nagaraja 1995
ParC				
Escherichia coli	EcoC	752	P20082	Kato *et al* 1990
Salmonella typhimurium	StyC	752	P26973	Luttinger *et al* 1991
Staphycoccus aureus	SauE	800	*L25288	Ferrero *et al* 1994
ParE				
Escherichia coli	EcoE	630	P20083	Kato *et al* 1990
Salmonella typhimurium	StyE	630	P31598	Springer and Schmid 1991
Staphycoccus aureus	SauE	663	*L25288	Ferrero *et al* 1994
Eukaryotic topoisomerase II				
ASF Virus (BA71V)	Top2_AS	1192	Q00942	Garcia-Beato *et al* 1992
ASF Virus (Malawi)	Top2_AF	1191	P34203	Baylis *et al* 1992
Crithidia fasciculata	Top2_CR	1239	P27570	Pasion *et al* 1992
Trypanosoma brucei	Top2_TR	1221	P12531	Strauss and Wang 1990
Trypanosoma cruzi	Top2_TY	1232	P30190	Fragoso and Goldberg 1992
Caenorhabditis elegans	Top2_CA	1198	P34534	Wilson *et al* 1994
Homo sapiens (A)	TopA_HU	1530	P11388	Tsai-Pflugfelder *et al* 1988
Homo sapiens (B)	TopB_HU	1626	QO2880	Jenkins *et al* 1992
Mus musculus (Mouse)	Top2_MO	1528	Q01320	Adachi *et al* 1992
Drosophila melanogaster	Top2_DR	1447	P15348	Wykoff *et al* 1989
Plasmodium falciparum	Top2_PL	1398	P41001	Ridley and Kilbey 1994
Saccharomyces cerevisiae	Top2_YE	1429	P06786	Giaever *et al* 1986
Schizosaccharomyces pombe	Top2_SC	1431	P08906	Uemura *et al* 1986

```
                                                                                                                                   *
MsmA   mtdttlppege-----AHDRIEPVDIQQEMQRSYIDYAMSVIVGRALPEVRDGLKPVHRRVLYAMYDSASSDRSHAKSARSVAETMGNYHPHGDASIYDTLVRMAQPWSLRYPLVDGQGNFGSPGn-DPP--AAMRYTEARLTPLAMEML   142
MtbA   mtdttlppdds-----LDRIEPVDIEQEMQRSYIDYAMSVIVGRALPEVRDGLKPVHRRVLYAMFDSGFRPDRSHAKSARSVAETMGNYHPHGDASIYDSLVRMAQPWSLRYPLVDGQGNFGSPDn-DPP--AAMRYTEARLTPLAMEML   142
ScoA   mtdentpvttpegdalAMRVEPVGLETEMQRSYLDYAMSVIVSRALPVVRDGLKPVHRRVLYAMYDGGYRPERGFYKCARVVGDVMGNYHPHGDSSIYDALVRLAQPWSMRMPLVDSNGNFGSPGn-DPA--AAMRYTECKMAPLSMEMV   147
BsuA   mseqntp---------QVREINISQEMRTSFLDYAMSVIVSRALPSDVRDGLKPVHRRILYAMNDLGMTSDKPYKKSARIVGSEVIGKYHPHGDSAVYESMVRMAQDFNYRYMLVDGHGSNFGSVdgDSA--AAMRYTEARMSKISMEIL   139
HalA   mssdapdsfepgagi-AAEVKNARIEDEMEQSYIDYAMSVIAGRALPDVRDGLKPVHRRILYAMHQAGVTSNSSHRKSSSIVGETMGDYHPHGDSAIYDTLARMAEDFSMRYPLVDGQGNFGSVDg-DPP--AAMRYTEARMSPIAEELL   146
SauA   maelp-----------QSRINERNITSEMRESFLDYAMSVIVARALPDVRDGLKPVHRRILYGLNEQGMTPDKSYKKSARIVGDVMGKYHPHGDSSIYEAMVRMAQDFSYRYPLVDGQGNFGSMDg-DGA--AAMRYTEARMTKITLELL   136
RprA   midky-----------SSNLVPVNIEDEMKVSYLDYAMSVIVSRAIPDVRDGLKPVHRRIIYSMYEAGNHASKPYRKSARIVGDVMGKYHPHGDSAIYDSLVRMAQDFSLRLPLVDGQGNFGSMDg-DAA--AAMRYTESRMAKVAHKLV   136
EcoA   msdl------------AREITPVNIEEELKSSYLDYAMSVIVGRALPDVRDGLKPVHRRVLYAMNVLGNDWNKAYKKSARVVGDVIGKYHPHGDSAVYDTIVRMAQPFSLRYMLVDGQGNFGSIDg-DSA--AAMRYTEIRLAKIAHELM   135
KpnA   msdl------------AREITPVNIEEELKSSYLDYAMSVIVGRALPDVRDGLKPVHRRVLYAMNVLGNDWNKAYKKSARVVGDVIGKYHPHGDTAVYDTIVRMAQPFSLRYMLVDGQGNFGSVDg-DSA--AAMRYTEIRMSKIAHELM   135
CajA   menifsk---------DSDIELVDIENSIKSSYLDYSMSVIIGRALPDARDGLKPVHRRILYAMQNDEAKSRTDFVKSARIVGAVIGRYHPHGDTAVYDALVRMAQDFSMRYPSITGQGNFGSIDg-DSA--AAMRYTEAKMSKLSHELL   138
SauC   m---------------SEIIQDLSLEDVLGDRFGRYSKYIIQERALPDVRDGLKPVQRRILYAMYSSGNYHDKNFRKSAKTVGDVIGQYHPHGDSSVYEAMVRLSQDWKLRHVLIEMHGNNGSIDn-DPP--AAMRYTEAKLSLLAEELL   132
StyC   m---------------SDMAERLALHEFTENAYLNYSMYVIMDRALPFIGDGLKPVQRRIVYAMSELGLNATAKFKKSARTVGDVLGKYHPHGDSACYEAMVLMAQPFSYRYPLVDGQGNWGAPD--DPKsfAAMRYTESRLSKYAELLL   133
EcoC   m---------------SDMAERLALHEFTENAYLNYSMYVIMDRALPFIGDGLKPVQRRIVYAMSELGLNASAKFKKSARTVGDVLGKYHPHGDSACYEAMVLMAQPFSYRYPLVDGQGNWGAPD--DPKsfAAMRYTESRLSKYSELLL   133

MsmA   REIDEETV-DFIPNYDGRVQEPTVLPSRFPNLLANGSRGIAVGMATNNPPHNLGELAEAVYWCLENYEADEEATCEAVMERVKGPDFPTSGLIVGTQGIEDTYKTGRGSIKMRGVVEIEEDSRGRTSIVITELPYQVNHDNFITSIAEQV   291
MtbA   REIDEETV-DFIPNYDGRVQEPTVLPSRFPNLLANGSGGIAVGMATNIPPHNLRELADAVFWALENHDADEEETLAAVMGRVKGPDFPTAGLIVGSQGTADAYKTGRGSIRMRGVVEVEEDSRGRTSLVITELPYQVNHDNFITSIAEQV   291
ScoA   RDIDEETV-DFTDNYDGRSQEPTVLPARFPNLLINGSAGIAVGMATNIPPHNLREVAAGAQWYLENYEASHEELLDALIERIKGPDFPTGALVVGRKGIEEAYRTGRGSITMRAV-VEVEEIQNRQCLVVTELPYQTNPDNLAQKIADLV   295
BsuA   RDITKDTIsDYQDNYDGSEREPVVMPSRFPNLLVNGAAGIAVGMASTNIPPHQLGEIIDGVLAVSENPDITIPELMEVIPGPSDFPTAGQILGRSGIRKAYESGRGSITIRAKAE-IEQTSSSGKERIIVTELPYQVNKAKLIEKIADLV   288
HalA   DDIDKDTV-DFQSNYDDRKQEPTVLPSSFPNLLVNGSSGIAVGMSTNIPPHNLGEVVDAAVELIENPDATVADLMEHIKGPDFPTGANIVARNAVHKAYKTGRGRVRVR-------ADYDVFEEEGRIVINELPYQENKARLIERIADDV   288
SauA   RDINKDTI-DFIDNYDGNEREPSVLPARFPNLLANGASGIAVGMATNIPPHNLTELINGVLSLSKNPDISIAELMEDIEGPDFPTAGLILGKSGIRRAYETGRGSIQMRSRA----VIEERGGGRQRIVVTEIPFQVNKARMIEKIAELV   281
RprA   EDIDKGTV-SFNINYDGSEEEPSVLPAMFPNLLVNGSGGIAVGMATNIPPHNLGEIIDACCLYIDNHDIEILDLLEVVKGPDFPTGSMILGISGIRSAYLTGRGSIIMRGKA----EIENVGNSRQAIIITEIPYMVNKARLVEKIAEMV   281
EcoA   ADLEKETV-DFVDNYDGTEKIPDVMPTKIPNLLVNGSSGIAVGMATNIPPHNLTEVINGCLAYIDDEDISIEGLMEHIPEPDFPTAAIINGRRGIEEAYRTGRGKVYIRARAE---VEVDAKTGRETIIVHEIPYQVNKARLIEKIAELV   281
KpnA   ADLEKETV-DFVDNYDGTEKIPDVMPTKIPNLLVNGSFGIAVGMATNIPPHNLTEVINGRLAYVEDEEISIEGLMEHIPGPDFPTAAIINGRRGIEEAYRTGRGKVYICARAE---VEADAKTGRETIIVHEIPYQVNKARLIEKIAELV   281
CajA   KDIDKDTV-DFVPNYDGSESEPDVLPSRVPNLLLNGSSGIAVGMATNIPPHSLNELIDGLLYLLDNKDASLEEIMQFIKGPDFPTGGIIYGKKGIIEAYRTGRGRVKVRAK-----THIEKKTNKDVIVIDELPYQTNKARLIEQIAELV   282
SauC   RDINKETV-SFIPNYDDTTLEPMVLPSRFPNLLVNGSTGISAGYATDIPPHNLAEVIQATLKYIDNPDITVNQLMKYIKGPDFPTGGIIQGIDGIKKAYESGKGRIIVRSKV----EEETLRNGRKQLIITEIPYEVNKGSLVKRIDELR   277
StyC   SELGQGT--ADWVPNFDGTMQEPKMLPARLPNILLNGTTGIAVGMATDIPPHNLREVAKAAITLIEQPKTTLDQLLDIVQGPDYPTEAEIITPRAEIRKIYENGRGSVR-------IGAVWTKEDGAVVISALPHQVSGAKVLEQIAAQM   274
EcoC   SELGQGT--ADWVPNFDGTLQEPKMLPARLPNILLNGTTGIAVGMATDIPPHNLREVAQAAIALIDQPKTTLDQLLDIVQGPDYPTEAEIITSRAEIRKIYENGRGSVR-------MRAVWKKEDGAVVISALPHQVSGARVLEQIAAQM   274

MsmA   RDGKLAGISNIEDQSSDRVGLRIVVELKRDAVAKVVLNNLYKH--TQLQTSFGANMLSIVDGVPRTLRLDQLIRLYVDHQLDVIVRRTR-YRLRKANERAHILRGLVKALDALDEVIALIRASQTVDIAR--------------------   418
MtbA   RDGKLAGISNIEDQSSDRVGLRIVIEIKRDAVAKVVINNLYKH--TQLQTSFGANMLAIVDGVPRTLRLDQLIRYYVDHQLDVIVRRTT-YRLRKANERAHILRGLVKALDALDEVIALIRASETVDIAR--------------------   418
ScoA   KDGKVGGIADVRDETSSRTGQRLVIVLKRDAVAKVVLNNLYKH--TDLQSNFGANMLALVDGVPRTLSLDAFIRHWVNHQIEVIVRRTR-FRLRKAEERAHILRGLLKALDAIDEVIALIRRSDTVEIAR--------------------   422
BsuA   RDKKIEGSITDLRDESDRTGMRIVIEIRRDANANVILNNLYKQTsALQTSFGI-NLLALVDGQPKVLTLKQCLEHYLDHQKVSVIRRRTaYELRKAEARAHILEGLRVALDHLDAVISLSIRNSQTAEIART------------------   419
HalA   NEGKIEGIRDIRDESERDGIRVVIELKRGAMAEVVKNQLLD----NHLESTFGVINLALVDGQPQVLTLKETLEHYLDHRRDVVRRRSE-YELAEAEDRAHILDGRLKALDNIDDVVETIRNSESRDDAKAALRGEvevevd--------   425
SauA   RDKKIDGITDLRDETSLRTGVRVVIDVRKDANASVILNNLYKQ--TPLQTSFGVNMIALVNGRPKLINLKEALVHYLEHQKTVVRRRTQ-YNLRKAKDRAHILEGLRIALDHIDEIISTIRESDTDKVAM--------------------   408
RprA   KEKRIEGISDLRDESNKNGIRIFIELKKDVVAEVVLNQIYAC---TQLQTNFGVIMLALKDGLPKVMNLKEVIAAFVSFREVVITNRTI-YLLNKARDRAHILLGLTIAISNIDEIIYIIKASNDTNLAKQELMARqwevldilpliklv   427
EcoA   KEKRVEGISALRDESDKDGMRIVIEVKRDAVGEVVLNNLYSQ---TQLQVSFGINMVALHHGQPKIMNLKDIIAAFVRHRREVVTRRTI-FELRKARDRAHILEALAVALANIDPIIELIRHAPTPAEAKTALVANpwqlgnvaamlera   427
KpnA   KEKRVEGISALRDESDKDGMRIVIEVKRDAVGRVVLNNLYSQ---TQLQVSFGINMVALHHGQPKIMNLKEIIAAFVRHRREVVTRRTI-LALRKARDRADILEALSIALANIDPIIELIRRAPTPAEAKAGLIARswdlgnvsamleag   427
CajA   KERQIEGISEVRDESNKEGIRVVIELKREAMSEIVLNNLFKS---TTMESTFGVIMLAIHNKEPKIFSLLELLNLFLTHRKTVIIRRTI-FELQKARARAHILEGLKIALDNIDEVIALIKNSSDNNTAR--------------------   408
SauC   ADKKVDGIVEVRDETDRTGLRIAIELKKDVNSESIKNYLYKN---SDLQISYNFNMVAISDGRPKLMGIRQIIDSYLNHQIEVVANRTK-FELDNAEKRMHIVEGLIKALSILDKVIELIRSSKNKRDAK--------------------   403
StyC   RNKKLPMVDDLRDESDHENPTRLVIVPRSNRVDMEQVMNHLFATtDLEKSYRINLNMIGLDGRPAVKNLLEILTEWLAFRRDTVRRRLN-YRLEKVLKRLHILEGLLVAFLNIDEVIEIIRSEDEPK---------------------   400
EcoC   RNKKLPMVDDLRDESDHENPTRLVIVPRSNRVDMDQVMNHLFATtDLEKSYRINLNMIGLDGRPAVKNLLEILSEWLVFRRDTVRRRLN-YRLEKVLKRLHILEGLLVAFLNIDEVIEIIRNEDEPK---------------------   400

MsmA   --------------AGLIELLDIDDIQAQAILDMQLRRLAALERQKIVDDLAKIEAEIADLEDILAKPERQRGIVRDELKEIVDKHGDARRTRIVPADGQVSDEDLIAREDVVVTITETGYAKRTKTDLYHSQKRGGKGVQGAGLKQDDM   554
MtbA   --------------AGLIELLDIDEIQAQAILDMQLRRLAALERQRIIDDLAKIEAEIADLEDILAKPERQRGIVRDELAEIVDRHGDDRRTRIIAADGDVSDEDLIAREDVVVTITETGYAKRTKTDLYRSQKRGGKGVQGAGLKQDDI   554
ScoA   --------------GGLMDLLEIDEIQANAILEMQLRRLAALERQKIVREHDELQAKITEYNEILASPVRQRGIVSEELTALVEKYGDDRKTKLIPYEGDMSIEDLIAEEDIVVTVTRGGYIKRTKTDDYRAQKRGGKGVRGTKLKEDDI   558
BsuA   --------------GLIEQFSLTEKQAQAILDMRLQRLSTGLEREKIEEEYQSLVKLIAELKDILANEYKVLEIISREELTEIKERFNDERRTEIVTSGLETIEDEDLIERESNIVVTLTHNGYVKRLPASTYRSQKRGGKGVQGMGTNS   555
HalA   --------------GEPLPTFDFSEEQANHIVSMQLGSLTSMEAAEIEAEYEDVQATIERLETILGDQSELDAVIESELLDIKDEYADDRRTSFVANTGEVTRADLIPEEDVVVVVSEDDYIKRMPVSRFRAQHRGGKGIIGTDLKEGDN   561
SauA   --------------ESLQQRFKLSEKQAQAILDMRLRRLTGLERDKIEAEYNELLNYISELEAILADEEVLLQLVRDELTEIRDRFGDDRRTEIQLGGFEDLEDEDLIPEEQIVITLSHNNYIKRLPVSTYRAQNRGGRGVQGMNTLEED   544
RprA   ddkvi---------LNERGTLSFTEVQAKAILEMKLQRLTAMEKEKLEQDLKHLATDIAEYLNILASRTRLLEILKEELIKVKEEFASPRLTSIEFGEFDQDIEDLIQREEMVVTVTLGGYIKRVPLSSYRSQKRGGKGRSGLSMRDEDI   568
EcoA   gddaarpewlepefGVRDGLYYLTEQQAQAILDLRLQKLTGLEHEKLLDEYKELLDQIAELLRILGSADRLMEVIREELELVREQFGDKRRTEITANSADINLEDLITQEDVVVTLSHQGYVKYQPLSEYEAQRRGGKGKSAARIKEEDF   577
KpnA   ddaarpewlepef-GVRDGQYYLTEQQAQAILDLRLQKLTGLEHEKLLDEYKELLEQIAELLHILGSADRLMEVIREELELVREQFGDARRTDITANSVDINIEDLITQEDVVVTLSHEGYVKYQPVNDYEAQRRGGKGKSAPRIKEEDF   576
CajA   --------------DSLVAKFGLSELQANAILDMKLGRLTGLEREKIENELAELMKEIARLEEILKSETLLENLIRDELKEIRSKFDVPRITQIEDDYDDIDIEDLIPNENMVVTITHRGYIKRVPSKQYEKQKRGGKGKLAVTTYDDDF   544
SauC   --------------ENLIEVYEFTEEQAEAIVMLQLYRLTNTDIVALEGEHKELEALIKQLRHILDNHDALLNVIKEELNEIKKKFKSERLSLIEAEIEEIKIDKEVMVPSEEVILSMTRHGYIKRTSIRSFNASGVEDIGLKDGDSLLK   539
StyC   --------------PALMSRFGISETQAEAILELKLRHLAKLEEMKIRGEQDELEKERDQLQGILASERKMNTLLKKELQADADAYGDDRRSPLREREEAKAMSEHDMLPSEPVTIVLSQMGWVRSAKGHDIDAPGLNYKAGDSFKAAVK   536
EcoC   --------------PALMSRFGLTETQAEAILELKLRHLAKLEEMKIRGEQSELEKERDQLQGILASERKMNNLLKKELQADAQAYGDDRRSPLQEREEAKAMSEHDMLPSEPVTIVLSQMGWVRSAKGHDIDAPGLNYKAGDSFKAAVK   536
```

Figure 1A.

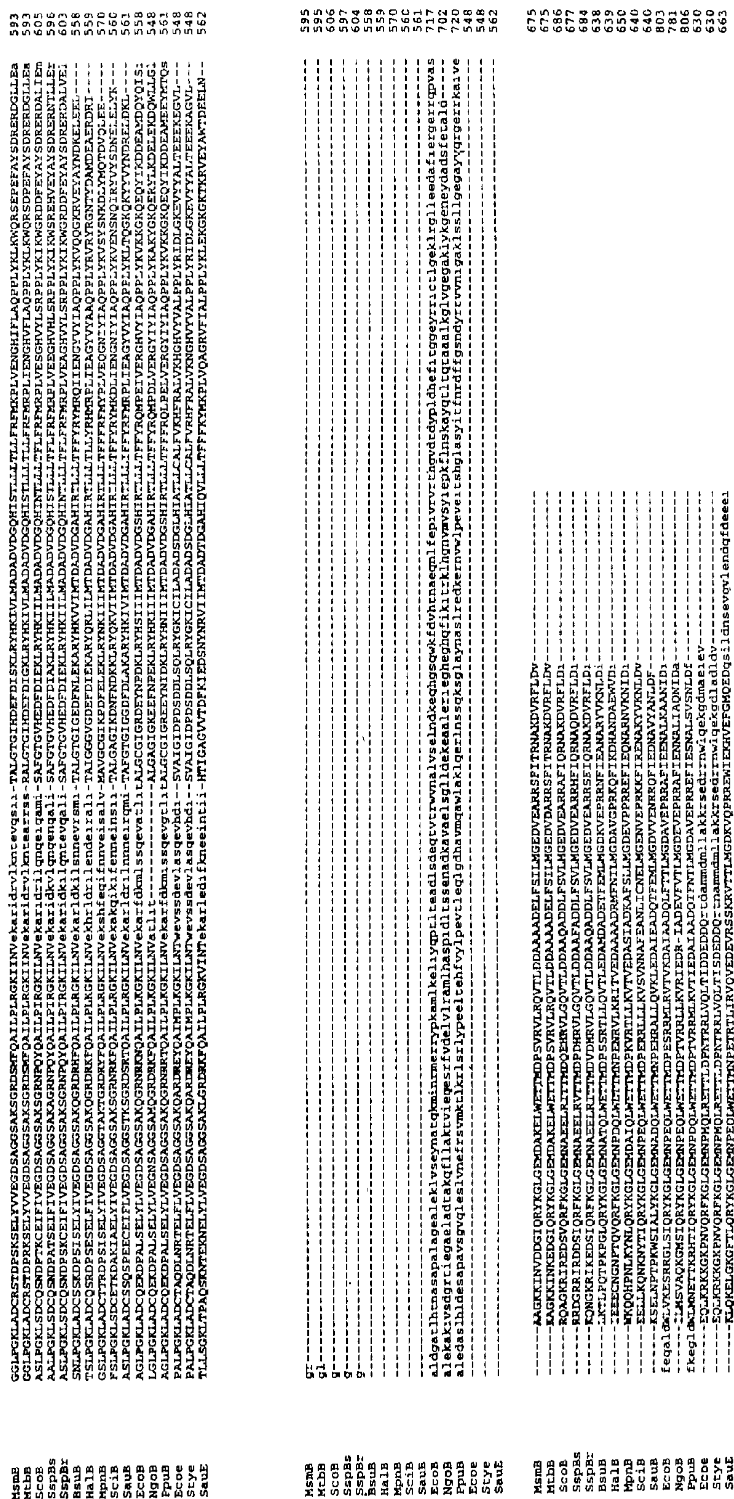


Figure 1B.

```
TOP2_AS   meafeisdfkehakkksmwagalnkvtisglmgvftededlmalp-----------------------------------------------------IHRDHCPALLKIFDELIVNATDHERACHSKTKKVTYIKISFDKGVFS   92
TOP2_AF   meafeisdfkehakkksmwagalnkvtisglmgvftededlmalp-----------------------------------------------------IHRDHCPALLKIFDELIVNATDHERACHSKTKKVTYIKISFDKGVFA   92
TOP2_CA   mydthdlrlpwvafftrkyvkagdeltwdyqytqdqtattqltchcgaenctgrllKKAAAKYEKKSPTEHVLLRPDTYIGGVAMREDQIIWLRDSENrkmiaKEVTYPPGLLKIFDEILVN-------AADNKARDSSMNRLEVWLDR   142
TOP2_CR   -----------------------------------------------------MTDASKYQKLTPIDHVLLRPEMYVGSIETQSIPMFVFDPAKGkmvw-ESMQVNQGLLKIVDEILLNAA------DNINNSVRGARMTYISIKIS   87
TOP2_DR   mengnka----------------------------------------------LSIEQMYQKKSQLEHILLRPDSYIGSVEFTKELMWVYDNSQNrmvq-KEISFVPGLYKIFDEILVN-------AADNKQRDKSMNTIKIDIDP   92
TOP2_MO   melsplqpvnenmlmnkkknedgkkr---------------------------LSIERIYQKKTQLEHILLRPDTYIGSVELVTQQMWVYDEDVGiny--REVTFVPGLYKIFDEILVN-------AADNKQRDPKMSCIRVTIDP   110
TOP2_PL   makn-------------------------------------------------KTIEERYQKKSQIEHILLRPDTYIGSVEMHTQLLWVWNKEKNrmvq-KNITYVPGLYKIFDEIIVNAADVK---AREKEKSENPMTCIKIEINK   94
TOP2_SC   mtaseqiplvtnngngn------------------------------------SNVSTQYQRLTPREHVLRRPDTYIGSIEPTTSEMWVFDSEKNkldy-KAVTYVPGLYKIFDEIIVN-------AADNKVRDPNMNTLKVTLDP   102
TOP2_TR   -----------------------------------------------------MAEAHKYKKLTPIEHVLTRPEMYIGSLDTTATPMFIYDEQKGhmvw-ETVKLNHGLLKIVDEILLN-------ASDNISNRSARMTYIRVTIT   85
TOP2_TY   -----------------------------------------------------MAEASKYKKLTPIDHVLIRPEMYVGSVDTSSSSMFVFDHEKGrmvw-ESLKVNHGLLKIVDEILLN-------ASDNIANKGGRMTYIRVHIT   85
TOP2_YE   mstep------------------------------------------------VSASDKYQKISQLEHILKRPDTYIGSVETQEQLQWIYDEETDcmie-KNVTIVPGLFKIFDEILVNA------ADNNKVRDPSMKRIDVNIHA   91
TOPA_HU   mevsplqpvnenmqvnkikknedakkr--------------------------LSVERIYQKKTQLEHILLRPDTYIGSVELVTQQMWVYDEDVGiny--REVTFVPGLYKIFDEILV--------NAADNKQRDPKMSCIRVTMI   110
TOPB_HU   maksggcgagagvgggngaltwvtlfdqnnaakkeesetankndsskk-------LSVERVYQKKTQLEHILLRPDTYIGSVEPLTQFMWVYDEDVGmnc--REVTFVPGLYKIFDEILVN-------AADNKQRDKNMTCIKVSIDP   132

TOP2_AS   CENDGPGIPIAKHEQASLIAKRDVYVPEVASCFFLAGTNINKAKDCIKGGTNGVGLKLAMVHSQWAILTTADGAQKYVVQQINQRLDIIEPPTITPSREMFTRIELMPVYQELGYAEPLSETEQA--------DLSAWIYLRACQCAAYVG   234
TOP2_AF   CENDGPGIPIAKHEQASLIAKRDVYVPEVASCHFLAGTNINKAKDCIKGGTNGVGLKLAMVHSQWAILTTADGAQKYVVQHINQRLDIIEPPTITPSREMFTRIELMPVYQELGYAQPLSETEQA--------DLSAWIYLRACQCAAYVG   234
TOP2_CA   ETARISVWNNGSGLPVEIHPTEGIYVPTLVFGNLFTSSNYDDSEIKTVGGRNGYGAKLCNIFSKEFIVETVDTRIKRRFRQKWYDNMKKCNEAEVveildetVKDYTKVEFVPDLERFQIDKLSDDVIDLIGRRVFEVAATLPRDVDVYL   292
TOP2_CR   DSGEIMVENDGAGLPIVKSKEHKMYIPEMVFGHLLTSSNYNNDASSTTAGRHGYGAKLTNILSTKFSVVCRTAGREFHMSWTDHMRMATTPRVSNvdpke--KNVTRVTFMPDYAHFGFPTAAISLDMKRVLHKRIMDLAAMFSKIEVRL   235
TOP2_DR   ERNMVSVWNNGQGIPVTMHKEQKMYVPTMIFGHLLTSSNYNDDEKKVTGGRNGYGAKLCNIFSTSFTVETATREYKKSFKQTWGNNMGKASDVQIkdf----NGTDYTRITFSPDLAKFKMDRLDEDIVALMSRRAYDVAASSKGVSVFL   238
TOP2_MO   ENNVISIWNNGKGIPVVEHKVEKIYVPALIFGQLLTSSNYDDDEKKVTGGRNGYGAKLCNIFSTKFTVETASRAYKKMFKQTWMDNMGRAGDMELkpf----SGEDYTCITFQPDLSKFKMQSLDKDIVALMVRRAYDIAGSTKDVKVFL   256
TOP2_PL   ENKRISVYNDGEGIPVDIHKEMNIYVPHMIFGELLTSDNYDDAEDRITGGRNGFGAKLTNIFSKEFIVQCGDSSRKKEFKMTWSDNMSKFSEPHIkn-----YNGKDYVKVTFKPDLNKFGMTEMDDDIESLLFKRVYDLAGTCSVRVYL   239
TOP2_SC   EANVISIYNNGKGIPIEIHDKEKIYIPELIFGNLLTSSNYDDNQKKVTGGRNGYGAKLCNIFSTEFVVETADKERMKKYKQTWYDNMSRKSEPVItslk---KPDEYTKITFKPDLAKFGMDKIDDDMVSIIKRRIYDMAGTVRETKVYL   249
TOP2_TR   DTGEITIENDGAGIPIVRSREHKLYIPEMVFGHLLTSSNYDDDNQNAVAGRHGYGAKLTNILSLSFSVCCRTNGREFHMSWQDHMRKATAPRVSNvgtke--KNVTRVKFLPDYERFGMKEKKISNDMKRVLYKRIMDLSAMFPNIQITL   233
TOP2_TY   EAGEITIENDGAGIPIVRSKEHKLYIPEMVFGHLLTSSNYDDTSQNAVAGRHGYGAKLTNILSHRFSVCCRTKGKEFHMSWHDHMRRATAPRVSNvdpke--KNLTRVKFLPDYERFGLDANKISHDMKRVLHKRIMDLAAMFPSIEISL   233
TOP2_YE   EEHTIEVKNDGKGIPIEIHNKENIYIPEMIFGHLLTSSNYDDDEKKVTGGRNGYGAKLCNIFSTEFILETADLNVGQKYVQKWENNMSICHPPKItsyk---KGPSYTKVTFKPDLTRFGMKELDNDILGVMRRRVYDINGSVRDINVYL   238
TOPA_HU   RKQLISIWNNGKGIPVVEHKVEKMYVPALIFGQLLTSSNYDDDEKKVTGGRNGYGAKLCNIFSTKFTVETASREYKKMFKQTWMDNMGRAGEMELkpf----NGEDYTCITFQPDLSKFKMQSLDKDIVALMVRRAYDIAGSTKDVKVFL   256
TOPB_HU   ESNIISIWNNGKGIPVVEHKVEKVYVPALIFGQLLTSSNYDDDEKKVTGGRNGYGAKLCNIFSTKFTVETACKEYKHSFKQTWMNNMMKTSEAKIkhf----DGEDYTCITFQPDLSKFKMEKLDKDIVALMTRRAYDLAGSCRGVKVMF   278

TOP2_AS   KGTTIYYNDKPCRTGSVMALAKMYTLLSAPNSTIHTATIKADAKPYSLHPLQVAAVVSPKFKKFEHVSIINGVNCVKGEHVTFLKKTINEMVIKKFQQTIKDKNRKTTLRDSCSNIFVVIVGSIPGIEWTGQRKDELSIAENVFKTHYSI   384
TOP2_AF   KGTTIYYNDKPCNTGSVMALAKMYTLLSAPNSTIHTTAIKADAKPYSLHPLQVAAVVSPKFKKFEHVSIINGVNCVKGEHVTFLKKTINEMVVKKFQQTIKDKNRKTTLRDSCSNIFVVIVGSIPGIEWTGQRKDELSIAENVFKTHYSI   384
TOP2_CA   NGQKCDVDGFEDYVKMFN-----------------------------------------------DSSSLLFLHPTPRWHVGVAKRNNFFGESHVVLPKIVSFVNNINTEKGGSHVDYVMDKIVNIIKPIVDSKlgdptksvkp   389
TOP2_CR   NNVPFGFQTFNDYARLYS-----------------------------------------------LPGADGAMPPEPFVHTGPNGSIAFVPQLTQSPKRIVGVVNGVVTYNGGTHCTSAMEILETGLDSLSRSLkkdgkvidtn   332
TOP2_DR   NGNKLGVRNFKDYIDLHI-----------------------------------------------KNTDDDSGPPIKIVHEVANERWEVACCPSDRGFQQVSFVNSIATYKGGRHVDHVVDNLIKQLLEVLKKKnkgginikpf   335
TOP2_MO   NGNMLPVKGFRSYVDLYL-----------------------------------------------KDKVDETGNSLKVIHEQVNPRWEVCLTMSERGFQQISFVNSIATSKGGRHVDYVADQIVSKLVDVVKKKnkggvavkah   353
TOP2_PL   NGQRLAVKDFKSYVDLYLKDNSNDNKNNkggndnnnnnnnnndenanqnndnldvslsnepadgtptknnnNNNNNNDEDEIVKIHEKQHRWEIVVSKSDGSQFQQVSFVNSICTTKGGSHVNYIVEQLLSSLSKKANAKnkggmeiksg   389
TOP2_SC   NNERISISGFKKYVEMYL-----------------------------------------------ASDTKPDEEPPRVIYEHVNDRWDVAFAVSDGQFKQVSFVNNISTIRGGTHVNYVANKIVDAIDEVVKKEnkkapvkafq   346
TOP2_TR   NGSSFGFKSFKDYATLYS-----------------------------------------------AMTPKGEKPPPPYVYESKSGCVAFIPSVVPGVRRMFGVVNGVVTYNGGTHCNAAQDILTGCLDGVERELkkenkvmdtn   330
TOP2_TY   NGVPFAFKSFADYAMLYS-----------------------------------------------SPSSSGEMPPAPFVYESRNGAIAFIPSLTAGTRRIFGVVNGVVTHNGGTHCNAAQEVLQSSLESVEKALkkdnkvidtn   330
TOP2_YE   NGKSLKIRNFKNYVELYLKSLEKKRQLDngedgaaksdiptilyerinnrwevafavsdisfqqisfvnsiattmggthvnyitdqivkkiseilkkkkkksvksfqiknnmfifincliienpaftsqtkeqlttrvkdfgsrceipley   388
TOPA_HU   NGNKLPVKGFRSYVDMYL-----------------------------------------------KDKLDETGNSLKVIHEQVNHRWEVCLTMSEKGFQQISFVNSIATSKGGRHVDYVADQIVTKLVDVVKKKnkggvavkah   353
TOPB_HU   NGKKLPVNGFRSYVDLYV-----------------------------------------------KDKLDETGVALKVIHELANERWDVCLTLSEKGFQQISFVNSIATTKGGRHVDYVVDQVVGKLIEVVKKKnkagvsvkpf   375

TOP2_AS   PSSFLTSMTRSIVDILLQSISKKDNHKQVDVDKYTRArNAGGKRAQDCMLLAAEGDSALSLLRTGLTLGKSNPSGPSFDFCGMISLGGVIMNACKKVTNITTDSGETimvr-----------------------------------NEQL   499
TOP2_AF   PSSFLTNMTRSIVDILLQSISKKDNHKQVDVDKYTRA-NAGGKKAQDCMLLAAEGDSALSLVRAGLTLGKSNPSGPSFDFCGMISLGGVIMNACKKVTNITTDSGETimvr-----------------------------------NEQL   498
TOP2_CA   avikknnlsifincliienpsfesqtketlttkaknfgsifecdakktaewaeqsgliedivevlnmkkkkl--PGKRVSVSSVRDIVKLEDAEWAGITGTAEKCTLILTEGDSAKALALAGLEVLGRETYGVFPLKGKLLNVSNLDDARA   537
TOP2_CR   rvarhftvlvfliqsqpkfdsqskarlvstvtmprvprtaldqylaampfleahmnsmddqlaaelnkeig--TGKRLSSRSLISSITKLVDATSSRSDGKNIRTLIVTEGDSAKALALNSLSSEQKKFCGVFPLRGKLLNVR---NKNL   477
TOP2_DR   qvrnhlwvfvncllienptfdsqtkenmtlqqkgfgskctlsekfinnmsksgivesvlawakfkaqndi----AKTCGRKSSKIKGIPKLEDANEAGGKNSIKCTLILTEGDSAKSLAVSGLGVIGRDLYGVFPLRGKLLNVREANFKQL   481
TOP2_MO   qvknhmwifvnalienptfdsqtkenmtlqaksfgstcqlsekfikaaigcgivesilnwvkfkaqiqln---KKCSAVKHTKIKGIPKLDDANDAGSRNSTECTLILTEGDSAKTLAVSGLGVVGRDKYGVFPLRGKILNVREASHKQI   500
TOP2_PL   hirnhlwvfvnclivnptfdsqtketlttkpvkfgskcilsdktinnvlkspilsnillwaqakaqvelkkkKAGSSKARERIIGIPKLEDANDAGSKYSQECTLILTEGDSAKTSCLAGLSIVGRDKYGVFPLKGKLLNVR---DASF   536
TOP2_SC   iknyvqvfvncqienpsfdsqtketlttkvsafgsqctlsdkflkaikkssvveevlkfatakadqql----SKGDGGLRSRITGLTKLEDANKAGTKESHKCVLILTEGDSAKSLAVSGLSVVGRDYYGVFPLRGKLLNVR---EASH   488
TOP2_TR   rvlrhftilvflvqvqpkfdsqnkarlvstptmprvprqdvmkyllrmpfleahvstitgqlaqelnkeig--TGRRMSSKTLLTSITKLVDATSTRRDPKHTRTLIVTEGDSAKALAQNSLSSDQKRYTGVFPLRGKLLNVRNKNLKRL   478
TOP2_TY   rvlrhfmilvflvqvqpkfdsqnkarlvsvptmprvprqelmdfllrmpfleahvntvtgqladelnkemg--AGRRMSSKSLISSITKLVDATTTRRDPRFVRTLIVTEGDSAKALAQNSLSSDQKRYTGVFPLRGKLLNVR---NKNL   475
TOP2_YE   inkimktdlatrmfeiadaneenal-----------------------------------------------KKSDGTRKSRITNYPKLEDANKAGTKEGYKCTLVLTEGDSALSLAVAGLAVVGRDYYGCYPLRGKMLNVR---EASA   487
TOPA_HU   qvknhmwifvnalienptfdsqtkenmtlqpksfgstcqlsekfikaaigcgivesilnwvkfkaqvqln---KKCSAVKHNRIKGIPKLDDANDAGGRNSTECTLILTEGDSAKTLAVSGLGVVGRDKYGVFPLRGKILNVREASHKQI   500
TOPB_HU   qvknhiwvfincllienptfdsqtkenmtlqpksfgskcqlsekffkaasncgivesilnwvkfkaqtqln---KKCSSVKYSKIKGIPKLDDANDAGGKHSLECTLILTEGDSAKSLAVSGLGVIGRDRYGVFPLRGKILNVREASHKQI   522
```

```
TOP2_AS   TNNKVLQGIVQVLGLDFNCHYKTQEERAKLRYGCIVACVDQDLDGCGKILGLLLAYFHLFWPQLIIh-GFVKRLLTPLIRVYEKGKTMPVEFYYEQEFDAWAKKQTSLVNHTVKYYKGLAAHDTHEVKSMFKHFDNMVYTFTLDDSAKEl   648
TOP2_AF   TNNKVLQGIVQVLGLDFNCHYKTQEERAKLRYGCIVACVDQDLDGCGKILGLLLAYFHLFWPQLIIh-GFVKRLLTPLIRVYEKGKTVPVEFYYEQEFDAWAKKQTSLANHTVKYYKGLAAHDTHEVKSMFKHFDNMVYTFTLDDSAKEl   647
TOP2_CA   SKNEEISNLLRILGLKFEDSNSITRESLRYGrllliladqdedgshikglivnfihkfwpslv------HTDGFIQSFRTPLLKAKKGDKVRSFFSMNEYRKWADVEEGGKWKIKYYKGLGTSTSNEAREYFSDLDHHTVNFKYTGTTDDd   681
TOP2_CR   KRLKTCKELQDLFLALGLELGKTYKSPAELRYQRLLVMTDQDADGSHIKGLVINAFESLWPSLLQHnpGYISLFSTPIVKIKVNGKAKEVVAFHSFRDFHRWQRANPNARYSAKYYKGLGTSTTAEGKEYFADMERNVMRLVVEPKDHRl   627
TOP2_DR   SENAEINNLCKIIGLQYKKKYLTEDDLKTLRygkvmimtdqdqdgshikgllinfihtnwpell----RLPFLEEFITPIVKATKKNEELSFYSLPEFEEWKNDTANHHTYNIKYYKGLGTSTSKEAKEYFQDMDRHRILFKYDGSVDDe   627
TOP2_MO   MENAEINNIIKIVGLQYKKNYEDEDSLKTLRYGKIMIMTDQDQDGSHIKGLLINFIHHNWPSLL----RHRFLEEFITPIVKVSKNKQEIAFYSLPEFEEWKSSTPNHKKWKVKYYKGLGTSTSKEAKEYFADMKRHRIQFKYPGPEDDa   646
TOP2_PL   KQLMDNKEIQNIFRIMGLDITDKNKDDIKGLRYGSLMIMTDQDYDGSHIKGLLINMIHKFWPSLL---KHKGFLSEFVTPIVKVQKGSQEYSFFTIAEYEQWKENTNLLGWKIKYYKGLGTSTDREFKQYFSDIKNHKIMFLWTGDRDGd   683
TOP2_SC   SQILNNKEIQAIKKIMGFTHKKTYTDVKGLRYGHLMIMTDQDHDGSHIKGLIINYLESSYPSLLQ---IPGFLIQFITPIIKCTRGNQVQAFYTLPEYEYWKEANNNGRGWKIKYYKGLGTSDHDDMKSYFSDLDRHMKYFHAMQEKDAe   635
TOP2_TR   RNCKELQELFCALGLELDKDYTDADELRYQRilimtdqdadgshikglvinafeslwpsllvr-----NPGFISIFSTPIVKARLRDKSVVSFFSMKEFHKWQRSNANTPYTCKYYKGLGTSTTAEGKEYFKDMEKHTMRLLVDRSDHKl   623
TOP2_TY   KRLKNCKELQELFCALGLELGKIYKDAEELRYQRLLVMTDQDADGSHIKGLVINAFEALWPSLLNR--NPGFISIFSTPIVKVRLRDKSTHSFFSLKEFHKWQKTHGNVSYTAKYYKALGTSTTAEGKEYFKDMDKHTMRLVVERNDHKl   623
TOP2_YE   DQILKNAEIQAIKKIMGLQHRKKYEDTKSLRYGHLMIMTDQDHDGSHIKGLIINFLESSFLGLLDIqgFLLEFITPIIKVSITKPTKNTIAFYNMPDYEKWREEESHKFTWKQKYYKGLGTSLAQEVREYFSNLDRHLKIFHSLQGNDKd   637
TOPA_HU   MENAEINNIIKIVGLQYKKNYEDEDSLKTLRYGKIMIMTDQDQDGSHIKGLLINFIHHNWPSLL----RHRFLEEFITPIVKVSKNKQEMAFYSLPEFEEWKSSTPNHKKWKVKYYKGLGTSTSKEAKEYFADMKRHRIQFKYSGPEDDa   646
TOPB_HU   MENAEINNIIKIVGLQYKKSYDDAESLKTLRYGKIMIMTDQDQDGSHIKGLLINFIHHNWPSLL----KHGFLEEFITPIVKASKNKQELSFYSIPEFDEWKKHIENQKAWKIKYYKGLGTSTAKEAKEYFADMERHRILFRYAGPEDDa   668

TOP2_AS   fhiyfggeselrkrelctgvvpltetqtqsihs-------VRRIPCSLHLQVDTKAYKLDAIERQIPNFLDGMTRARRKILAGGVKCFASNNRERKVFQFGGYVADHMFYHHGDMSLNTSIIKAAQYYPGSSHLYPVFIGIGSFGSRHLG   791
TOP2_AF   fhiyfggeselrkrelctgvvpltetqtqsihs-------DRQIPCSLHLQVDTKAYKLDAIERQIPNFLDGMTRARRKILAGGLKCFASNNRERKVFQFGGYVADHMFYHHGDMSLNTSIIKAAQYYPGSSHLYPVFIGIGSFGSRHLG   790
TOP2_CA   airmafdrdksderkewirrseneitn-------------EDDEQKKARMSDGVEDLQRFESQRIQSIPSLIDGLKPSQRKILWTLLNNMDESTEIKVSQLAGAVAHRQSYHHGEESLVRTIIRMGQTFCGSSNLPLLQPIGQFGTRHEG   818
TOP2_CR   ldsvfdsaevewrkewmskanafqgevdid----------RSKKLLTIGDFVHKEMVHFALVGNARAIPHCVDGLKPSQRKILWAMLKRHSSEAAKVAQLSGYISEVSSFHHGEASLQETIVKMAQNFTGGNNINLLVPEGQFG-SRQQL   766
TOP2_DR   sivmafskkhiesrkvwltnhmdevkrrkelglperylytKGTKSITYADFINLELVLFSNADNERSIPSLVDGLKPGQRKVMFTCFKRNDKREVKVAQLSGSVAEMSAYHHGEVSLQMTIVNLAQNFVGANNINLLEPRGQFG-TRLSG   776
TOP2_MO   aislafskkqvddrkewltnfmedrrqrkllglpedylygQSTSYLTYNDFINKELILFSNSDNERSIPSMVDGLKPGQRKVLFTCFKRNDKREVKVAQLAGSVAEMSSYHHGEMSLMMTIINLAQNFVGSNNLNLLQPIGQF-GTRLHG   795
TOP2_PL   sidmafskkriedrklwlqnfilgsyvd------------HKEKDLSYYDFVNKELIYYSRYDTERSIPNIMDGWKPGQRKVLYGCFKRNLRNECKVAQLVGYIAEHSAYHHHGESSLQQTIINMAQTFVGSNNINFLEPCGQFGSRKEG   821
TOP2_SC   liemafakkkadvrkewlrtyrpgiymd------------YTQPQIPIDDFINRELIQFSMADNIRSIPSVVDGLKPGQRKVVYYCFKRNLVHETKVSRLAGYVASETAYHHGEVSMEQTIVNLAQNFVGSNNINLLMPNGQFG-TRSEG   772
TOP2_TR   ldnvfdsqevewrkdwmtkanaftgevdid----------RSKKMLTVTDFVHKEMVHFALVGNARALAHSVDGLKPSQRKIIWALMRRSGNEAAKVAQLSGYISEASAFHHGETSLQETMIKMAQSFTGGNNVNLLVPEGQFG-SRQQL   762
TOP2_TY   ldsvfdsqevewrkdwrtkanaytgevdid----------RSKKTLTVPDFVHKEMVHFALAGNARALAHAVDGLKPSQRKILWAIMRRSGNESAKVAQLSGYISEVSAFHHGEMSLQETIIKMAQNFTGGNNINLLIPEGQFG-SRQQL   762
TOP2_YE   yidlafskkkaddrkewlrqyepgtvld------------PTLKEIPISDFINKELILFSLADNIRSIPNVLDGFKPGQRKVLYGCFKKNLKSELKVAQLAPYVSECTAYHHGEQSLAQTIIGLAQNFVGSNNIYLLLPNGAF-GTRATG   774
TOPA_HU   aislafskkqiddrkewltnfmedrrqrkllglpedylygQTTTYLTYNDFINKELILFSNSDNERSIPSMVDGLKPGQRKVLFTCFKRNDKREVKVAQLAGSVAEMSSYHHGEMSLMMTIINLAQNFVGSNNLNLLQPIGQF-GTRLHG   795
TOPB_HU   aitlafskkkiddrkewltnfmedrrqrrlhglpeqflygTATKHLTYNDFINKELILFSNSDNERSIPSLVDGFKPGQRKVLFTCFKRNDKREVKVAQLAGSVAEMSAYHHGEQALMMTIVNLAQNFVGSNNINLLQPIGQF-GTRLHG   817

TOP2_AS   GKDAGSPRYI-SVQLASEFIKTMFPAEDSWLLPYVFEDGQRAEPEYYVPVLPLAIMEYGANPSEGWKYTTWARQLEDILALVRAYVDKDNPKHEllhyaikhkitILPLRPSNYNFKGHLKRFGQYYYSYGTYDISEQRNIITITELPLR   940
TOP2_AF   GKDAGSPRYI-SVQLASEFIKTMFPAEDSWLLPYVFEDGQRAEPEYYVPVLPLAIMEYGANPSEGWKYTTWARQLEDILALVRAYVDKNNPKHEllhyaierkitVLPLRPSNYNFKGHLKRFGQYYYSYGTYVVSEQRNIITITELPLR   939
TOP2_CA   GNDAASARYI-F--TALAPTTRLLFPQADDDLLQKNVEEGMVVEPTLCPIVPLILINGTEGIGTGWSTKIANRNPIDIIDMIRRKIDSISTEYEippfyeefrgklevvtptkfissgkiqlirperknastfsieivelpigiwtskyk   965
TOP2_CR   GNDHAAPRYIFT--KLSRFARLLFPEDDDPLLDYIDEEGTMVEPNHYVPILPLLLCNGAVGIGFGFATNIPSFHPLDVSAAVRAMINGESAKQVvRNLVPWAVGFQGTVRRGPEKEYIAVGKYTAHRNGRLHVSELPWMTSIEAFRSHIS   914
TOP2_DR   GKDCASARYIFT--IMSPLTRLIYHPLDDPLLDYQVDDGQKIEPLWYLPIIPMVLVNGAEGIGTGWSTKISNYNPREIMKNLRKMINGQEP----SVMHPWYKNFLGRMEYVSDGRYIQTGNIQILSGNRLEISELPVGVWTQNYKENVL   920
TOP2_MO   GKDSASPRYIFT--MLSPLARLLFPPKDDHTLRFLYDDNQRVEPEWYNPINTMVLINGAEGIGTGWSCKIPNFDVREVVNNIRRLLDGEEP----LPMLPSYKNFKGTIEELASNQYVINGEVAILDSTTIEISELPIRTWTQTYKEQVL   939
TOP2_PL   GKDASAARYI-FT-KLASSTRSIFNEYDDPILKYLNEEGQKIEPQYYIPVIPTILVNGCEGIGTGYSSFIPNYNYKDIIDNIKRYINKEPL----IPMVPWYKDFKGRIESNGKTGYETIGIINKIDNDTLEITELPIKKWTQDYKEFLE   965
TOP2_SC   GKNASASRYLNT--ALSPLARVLFNSNDDQLLNYQNDEGQWIEPEYYVPILPMVLVNGAEGIGTGWSTFIPNYNPKDITANLRHMLNGEPL----EIMTPWYRGFRGSITKVAPDRYKISGIINQIGENKVEITELPIRFWTQDMKEYLE   916
TOP2_TR   GNDHAAPRYIFT--KLSKVARLLFPSEDDPLLDYIVEEGQQVEPNHYVPILPLLLCNGSVGIGFGFSSNIPPFHRLDVSAAVRAMISGERAKSVvrrlvpwavgfqgeirrgpegefiavgtytyckggrvhvtelpwtcsveafrehis   910
TOP2_TY   GNDHAAARYIFT--KLSSLARILFPSEDEPLLDYVTEEGQQVEPNHYVPILPLLLCNGSVGIGFGFASNIPPFHPLDVSAAVRSMINGEAAKVVvrrlvpwavgyqgevrrgpegefiaagsyqyyvdgrvhvtelpwtlsieafrdhis   910
TOP2_YE   GKDAAAARYIYT--ELNKLTRKIFHPADDPLYKYIQEDEKTVEPEWYLPILPMILVNGAEGIGTGRSTYIPPFNPLEIIKNIRHLMNDEEL----EQMHPWFRGWTGTIEEIEPLRYRMYGRIEQIGDNVLEITELPARTWTSTIKEYLL   918
TOPA_HU   GKDSASPRYIFT--MLSSLARLLFPPKDDHTLKFLYDDNQRVEPEWYIPIIPMVLINGAEGIGTGWSCKIPNFDVREIVNNIRRLMDGEEP----LPMLPSYKNFKGTIEELAPNQYVISGEVAILNSTTIEISELPVRTWTQTYKEQVL   939
TOPB_HC   GKDAASPRYIFT--MLSTLARLLFPAVDDNLLKFLYDDNQRVEPEWYIPIIPMVLINGAEGIGTGWACKLPNYDAREIVNNVRRMLDGLDP----HPMLPNYKNFKGTIQELGQNQYAVSGEIFVVDRNTVEITELPVRTWTQVYKEQVL   961

TOP2_AS   VPTVAYIESIKKSSNRMTFIEEIIDYSSSETIEILVKLKPNSLNRIVEEFKETEEQDSIENFLRLRNCLHSHLNFVKPKGGIIEFNSYYEILYAWLPYRRELYQKRLMREHAVLKLRIIMETAIVRYINESAELNLSHYEDEKEASRILS   1090
TOP2_AF   VPTVAYIESIKKSSNRMAFIEEIIDYSSSETIEILVKLKPNSLNRIVEEFKETEEQDSIENFLRLRNCLHSHLNFVKPKGGIIEFNSYYEILYAWLPYRRDVYQKRLMRERAVLKLRIIMETAIVRYINESADLNLSHYEDEKEASRILS   1089
TOP2_CA   eklskivetlpvlefser--HTEKRVHFRITLDRKKSSRFLQKSNSDLLNYFKLRTSLTENRVLFDRNGELKEFGNISEIAAEFFEVRRDLYEKRLKIQKEECEAKLIYVENQLNFIEMVTNGTIEIRSMGRNQLEEKLQEMGFRVDPMA   1113
TOP2_CR   SLASsdvvqriadysganhididlivreqsmttwaecetdlalsqriyingtvfspdgtlspida----------DLSPVLQWHYDRRLDLYKRRRTRQIGLLEMDLARLQSTRKFVEHFRQGHIDFLA------------------   1033
TOP2_DR   EPLSngtekvkgiiseyreyHTDTTVRFVISFAPGEFERIHAEEGGFYRVFKLTTTLSTNQMHAFDQNNCLRRFPTAIDILKEYYKLRREYYARRRDFLVGQLTAQADRLSDQARFILEKCEKKLVVENKQRKAMCDELLKRGYRPDPVK   1070
TOP2_MO   EPMLngtektpslitdyr--EYHTDTTVKFVIKMTEEKLAEAERVGLHKVFKLQSSLTCNSMVLFDHVGCLKKYDTVLDILRDFFELRLKYYGLRKEWLLGMLGAESSKLNNQARFILEKIDGKIVIENKPKKELIKVLIQRGYDSDPVK   1087
TOP2_PL   ELLTdekhqlildyi-----DNSSHEDICFTIKMDPAKLQKAEEEGLEKVFKLKSTLTTTNMTLFDPNLKLQRYSTELDILKEFCYQRLKAYENRKSYLISKLEKEKRIISNKTKFILAIVNNELIVNKKKKKVLVEELYRKGYDPYKDI   1110
TOB2_SC   AGLVgtekirkfivdye---SHHGEGIVHFNVTLTEAGMKEALNESLEVKFKLSRTQATSNMIAFDASGRIKKYDSVEDILTEFYEVRLRTYQRRKEHMVNELEKRFDRFSNQARFIHMIIEGELVVSKKKKKDLIVELKEKKFQPISKP   1063
TOP2_TR   ylatkdivnriadysganhvdidvevaqgavntyaeceselgltqrihingtvfspngtlsplesdltpvlqwhydrrldlykkrrqrnltlleqelarekstlkfvqhfgaghidf----------------------------------   1027
TOP2_TY   vlaskdvvqriadysganhvdidleltngamttyaeceselsltqriyingtvfsptgvltpleg----------DLAPVLQWHYDRRLDLYKKRRQRNLGLLEAELAREKSTLKFVTHFREGKIDIVN-------------------   1029
TOP2_YE   LGLSgndkikpwikdm----EEQHDDNIKFIITLSPEEMAKTRKIGFYERFKLISPISLMNMVAFDPHGKIKKYNSVNEILSEFYYVRLEYYQKRKDHMSERLQWEVEKYSFQVKFIKMIIEKELTVTNKPRNAIIQELENLGFPRFNKE   1064
TOPA_HU   EPMLngtektpplitdyr--EYHTDTTVKFVVKMTEEKLAEAERVGLHKVFKLQTSLTCNSMVLFDHVGCLKKYDTVLDILRDLFELRLKYYGLRKEWLLGMLGAESAKLNNQARFILEKIDGKIIIENKPKKELIKVLIQRGYDSDPVK   1087
TOPB_HU   EPMLngtdktpalisdyk--EYHTDTTVKFVVKMTEEKLAQAEAAGLHKVFKLQTTLTCNSMVLFDHMGCLKKYETVQDILKEFFDLRLSYYGLRKEWLVGMLGAESTKLNNQARFILEKIQGKITIENRSKKDLIQMLVQRGYESDPVK   1109
```

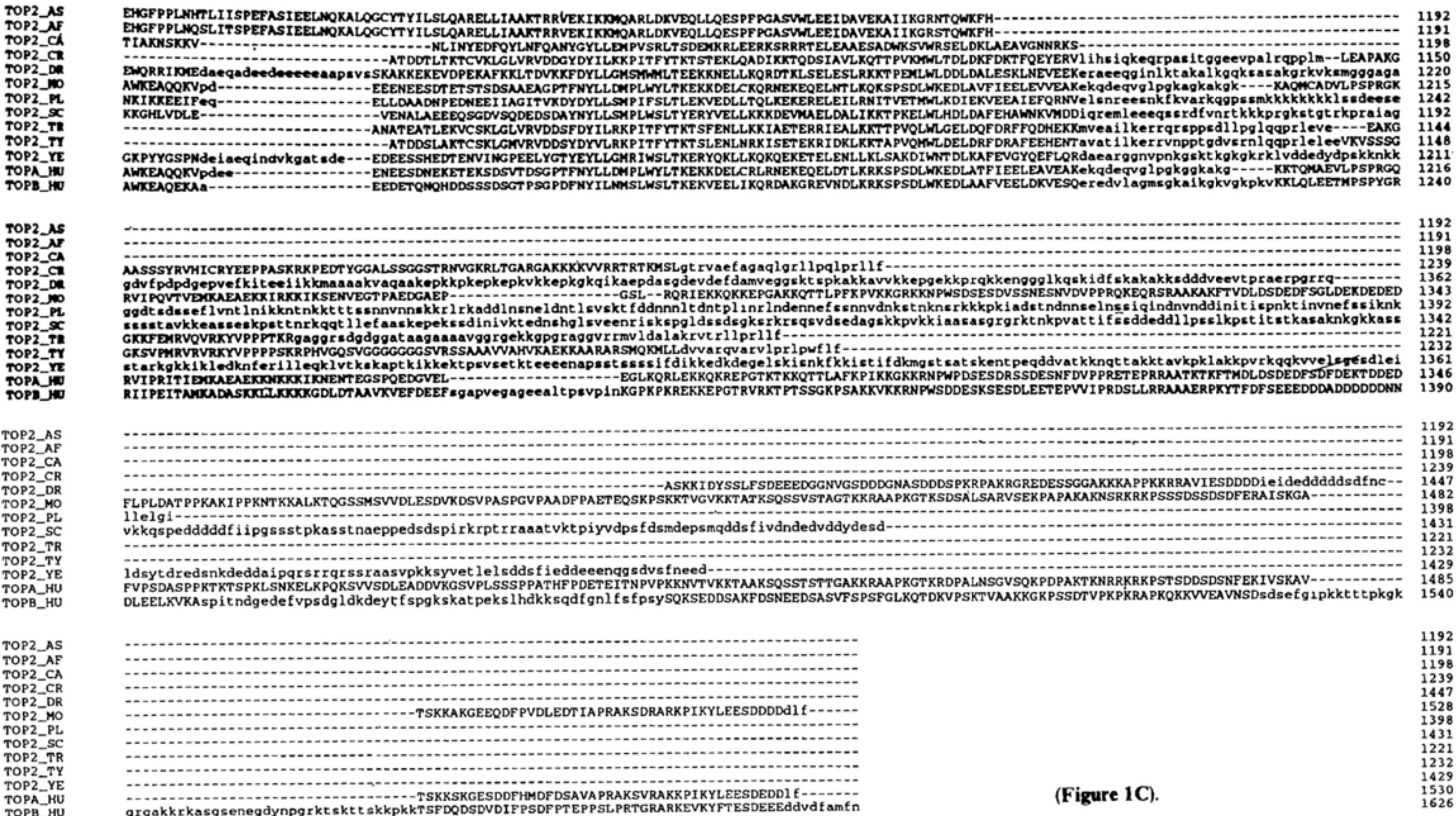


Figure 1. Alignment of type II topoisomerases. **(A)** Subunit A of bacterial type II topoisomerases. The conserved tyrosine involved in DNA breakage-reunion reaction is marked by an asterisk (corresponding to *E. coli* sequence). **(B)** Subunit B of bacterial type II topoisomerases. The asterisks mark the residues involved in contacting the nucleotide cofactor. All the positions correspond to *E. coli* sequence. **(C)** Eukaryotic type II topoisomerases. The sequence in the upper case represent the region with high conservation. (Figure 1C).

points, Tyr-5, Asn-46, Asp-73, Lys-103, Tyr-109, Gln-335 and Lys-337 are positionally conserved in all the members, except *Spiroplasma citri* wherein Lys-337 is replaced by Asn. The two important residues, Glu-42 and His-38 (Jackson and Maxwell 1994) implicated in ATPase activity of the subunit B of the *E. coli* enzyme are also present in all the other B sequences examined. A very significant difference is the absence of a long stretch of amino acids (158–163 amino acids) in GyrB proteins of Gram positive bacteria and *Mycoplasma* (figure 2B). The same stretch of amino acids is also absent in ParE sequences (B subunits of topoisomerase IV) of both Gram negative bacteria and *S. aureus*. It should also be noted here that gyrase and topoisomerase IV differ not only in their supercoiling ability but also in their potency of decatenation.

The subunits of bacterial type II topoisomerases were further analysed to understand the evolutionary relatedness. The unrooted trees are shown in figures 3 and 4. The ParC proteins share a branch with GyrA of higher bacteria (figure 3). The ParE proteins are located closer to GyrB polypeptides of Gram negative bacteria (figure 4) in spite of sharing a common character (the absence of a long stretch of amino acids) with Gram positive bacteria and *Mycoplasma* (figure 2B). The B subunit of bacterial type II topoisomerases shows a distinct feature. All the GyrB proteins of *Mycoplasma* and Gram positive bacteria form a monophyletic group while other GyrB polypeptides and ParE sequences diverge into another group (figure 3). In case of A subunits, no such clear separation could be observed (figure 3).

The eukaryotic type II topoisomerases also show primary conservation to some extent. Although the sequences have diverged, functionally significant domains (like

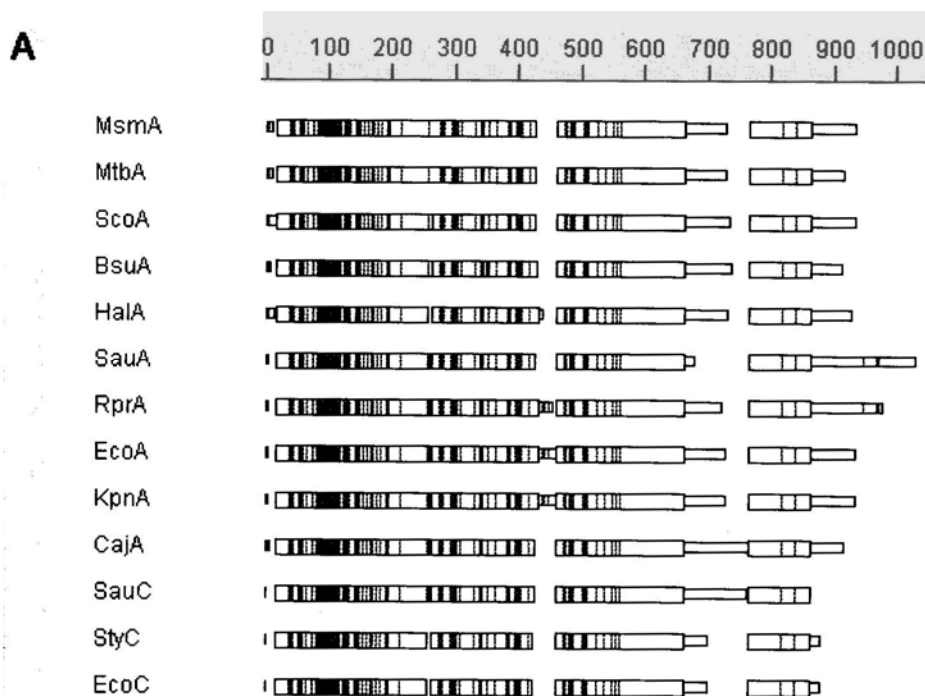


Figure 2A.

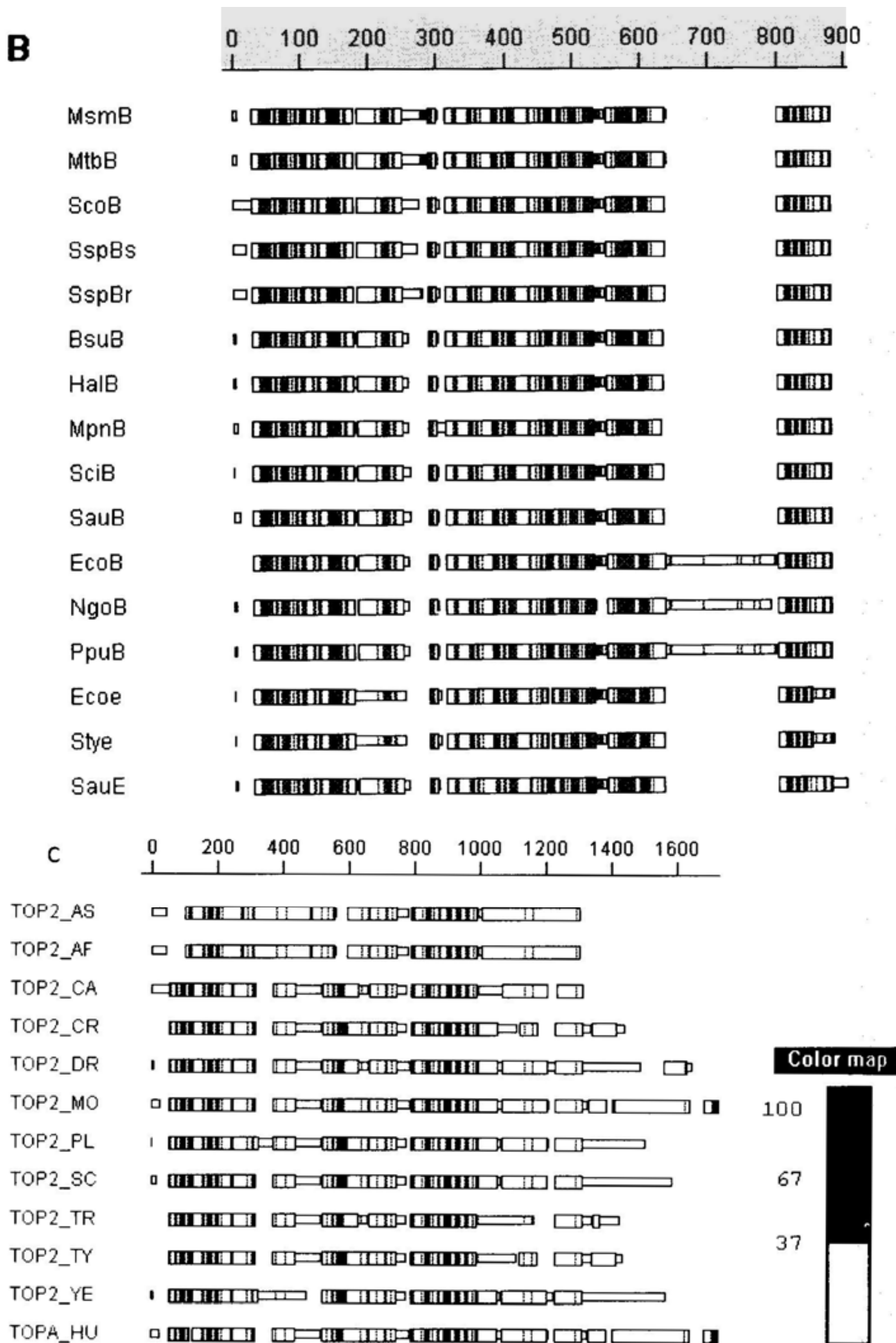
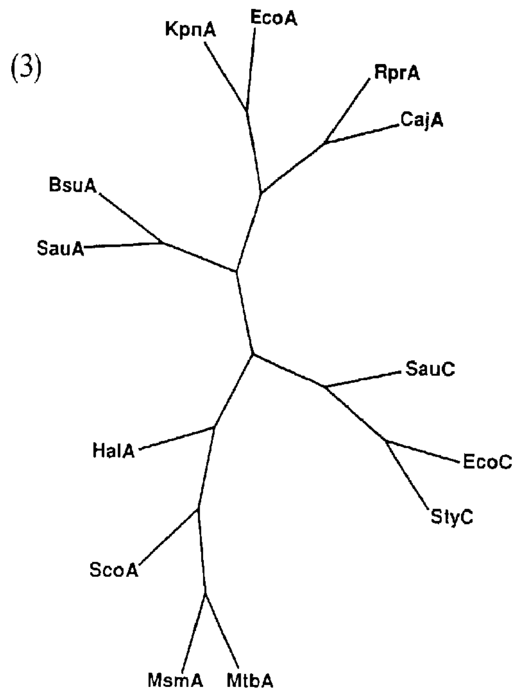
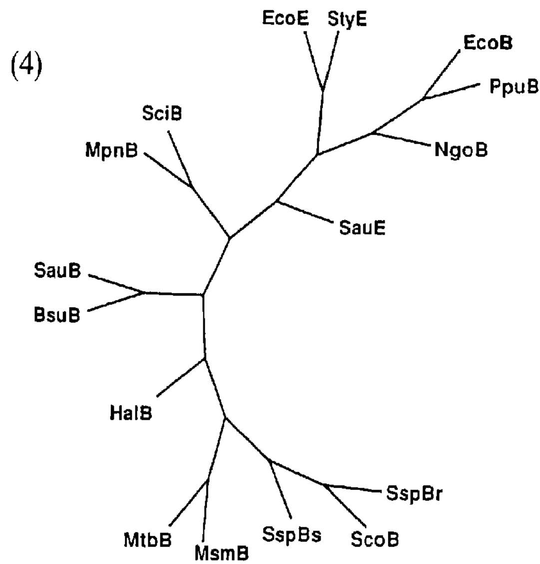


Figure 2. Schematic representation of multiple alignment derived from figure 1. The panels (**A**) (**B**) and (**C**) correspond to that of figure 1. The thick sequences correspond to blocks sharing significant similarity with the shaded regions representing the conserved sequence. The thin blocks represent the regions of the sequence not sharing statistically significant homology in the multiple alignment.





Figures 3 and 4. Evolutionary relationship among bacterial topoisomerase II subunits. Unrooted phylogenetic trees produced from the alignment of A subunits (**3**) in figure 1A and B subunits (**4**) in figure 1B, using DRAWTREE.

ATPase and DNA breakage-reunion regions) have retained the residues important for the activities. The crystal structure of 92 kDa domain of yeast DNA topoisomerase II at 2·7 Å has been reported recently (Berger *et al* 1996). Whereas the N-terminal 409 amino acids constitute ATPase domain that shares homology with bacterial DNA gyrase subunit B ATPase region, this domain has been implicated in DNA cleavage and strand passage reactions required for the topological interconversion (Berger and Wang 1996). This

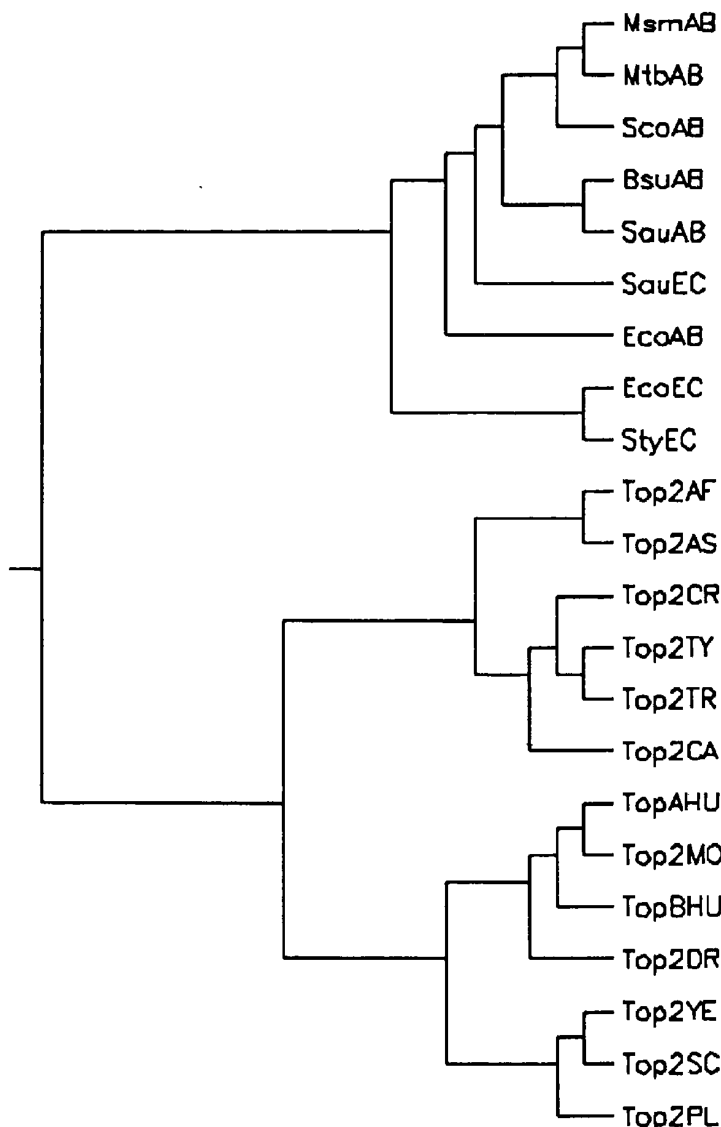


Figure 5. Phylogenetic tree of type II topoisomerases. The bacterial type II topoisomerase subunits were fused prior to the alignment. The abbreviations correspond to those given in the table 1, with AB representing fused sequences.

domain contains GyrB and GyrA like sequences (Caron and Wang 1993). Also, this region shows high conservation among eukaryotic type II DNA topoisomerases (figures 1C and 2C). Beyond this region, the amino acid sequences of eukaryotic type II DNA topoisomerases display less conservation. The cluster analysis of all type II topoisomerases, presented in figure 5, shows the monophyletic separation of bacterial sequences from the distinct diphyletic groups of eukaryotic enzymes.

Acknowledgements

We thank Shamala Prasad and Bindu D Paul for the assistance and C D Nager (Freidrich Miescher Institute, Basel) for introducing to MACAW. The sequence analysis was carried out at Bioinformatics Centre, Indian Institute of Science. The infrastructural facility was provided by Department of Biotechnology, New Delhi. KM is a recipient of Council of Scientifics and Industrial Research, New Delhi Senior Research Fellowship.

References

Adachi N, Miyaike M, Ikeda H and Kikuchi A 1992 Characterization of cDNA encoding the mouse DNA topoisomerase II that can complement the budding yeast top2 mutation; *Nucleic Acids Res.* **20** 5297–5303

Adachi T, Mizuuchi M, Robinson E A, Apella E, O'Dea M H, Gellert M and Mizuuchi K 1987 DNA sequence of the *E. coli gyrB* gene: application of a new sequencing strategy; *Nucleic Acids Res.* **15** 771–784

Baylis S A, Dixon L K, Vydelingum S and Smith G L 1992 African Swine Fever virus encodes a gene with extensive homology to type II topoisomerase; *J. Mol. Biol.* **228** 1003–1010

Berger J M, Gamblin S J, Harrison S C and Wang J C 1996 Structure and mechanism of DNA topoisomerase II; *Nature* (*London*) **379** 225–232

Berger J M and Wang J C 1996 Recent developments in DNA topoisomerase 11 structure and mechanism; *Curr. Op. Structural Biol.* **6** 84–90

Calcutt M J 1994 Improved shuttle vector for *Haloferax volcanii* including a dual-resistance plasmid; *Gene* **146** 117–121

Caron P R and Wang J C 1993 DNA topoisomerases as target of therapeutics: A structural overview; in *Molecular biology of DNA topoisomerases* (Proceedings of International Symposium on DNA Topoisomerases in Chemotherapy, Nagoya, Japan) (Nov. 18–20, 1991) (New York: CRC press) pp 1–21

Colman S D, Hu P C and Bott K F 1990 *Mycoplasma pneumoniae* DNA gyrase genes; *Mol. Microbiol.* **4** 1129–1134

Corpet J 1988 Multiple sequence alignment with heirarchical clustering; *Nucleic Acids Res.* **16** 10881–10890

Dimri G P and Das H 1990 Cloning and sequence analysis of gyrA gene of *Kiebsiella pneumoniae*; *Nucleic Acids Res.* **18** 151–156

Felsenstein J 1989 PHYLIP-phylogeny inference package (version 3.2); *Cladistics* **5** 164–166

Ferrero L, Cameron B, Manse B, Lagneaux D, Crouzet Famechon A and Blanche F 1994 Cloning and primary structure analysis of *Staphylococcus aureus* DNA topoisomerase IV: a primary target of fluoroquinolones; *Mol. Microbiol.* **13** 641–653

Fragoso S P and Goldenberg S 1992 Cloning and characterization of the gene encoding *Trypanosoma cruzi* DNA topoisomerase II; *Mol. Biochem. Parasitol.* **55** 127–134

Garcia-Beato R, Freije J M, Lopez-Otin C, Blasco R, Vinuela E and Salas M L 1992 A gene homologous to topoisomerase 11 in African swine fever virus; *Virology* **188** 938–947

Gellert M, Mizuuchi K, O'Dea M H and Nash H 1976 DNA gyrase: an enzyme that introduces superhelical turns into DNA; *Proc. Natl. Acad. Sci. USA* **73** 3872–3875

Giaever F, Lynn R, Goto T and Wang J C 1986 The complete nucleotide sequence of the structural gene TOP2 of yeast DNA topoisomerase II; *J. Biol. Chem.* **261** 12448–12454

Holmes M L and Dyall-Smith M L 1991 Mutations in DNA gyrase result in novobiocin resistance in halophilic archaebacteria; *J. Bacteriol.* **173** 642–648

Holmes M L and Dyall-Smith M L 1994 Gene organization in the *duaA-gyrA* region of the *Streptomyces coelicolor* chromosome; *Gene* **146** 117–121

Jackson A P and Maxwell A 1994 Identifying the catalytic residue of the ATPase reaction of DNA gyrase; *Proc. Natl. Acad. Sci. USA* **90** 11232–11236

Jenkins J R, Ayton P, Jones T, Davies S L, Simmons D L, Harris A L, Sheer D and Hickson I D 1992 Isolation of eDNA clones encoding the β isozyme of human DNA topoisomerase II and localization of the gene to chromosome; *Nucleic Acids Res.* **20** 5587–5592

Kato J-I, Nishimura Y, Immure R, Niki H Iraga S and Suzuki H 1990 New topoisomerase essential for chromosome segregation in *E. coli*; *Cell* **63** 393–404

Lawrence C E, Altschul S F, Boguski M S, Liu J S, Neuwald A F and Wootton J C 1993 Detecting subtle sequence signals: a Gibbs sampling strategy for multiple alignment; *Science* **62** 208–214

Luttinger A L, Springer A L and Schmid M B 1991 A cluster of genes that affects nucleoid segregation in *Salmonella typhimurium*; *New Biol.* **3** 687–697

Madhusudan K and Nagaraja V 1995 *Mycobacterium smegmatis* DNA gyrase: cloning and overexpression in *Escherichia coli*; *Microbiology* **140** 3029–3037

Madhusudan K, Ramesh V and Nagaraja V 1994 Molecular cloning of *gyrA and gyrB* genes of *Mycobacterium tuberculosis*: analysis of nucleotide sequence; *Biochem. Mol. Biol. Int.* **33** 651–660

Margerrison E E C, Hopewell R and Fisher L M 1992 Nucleotide sequence of the *Staphylococcus aureus* gyrB-gyrA Locus encoding the DNA gyrase A and B proteins; *J. Bacteriol.* **174** 1596–1603

Maxwell A and Gellert M 1986 Mechanistic aspects of DNA topoisomerses; *Adv. Prot. Chem.* **38** 69–107

Moriya S, Ogasawara N and Yoshikawa H 1985 Structure and function of the region of the replication origin of the *Bacillus subtilis* chromosome. III: nucleotide sequence of some 10,000 base pairs in the origin region; *Nucleic Acids Res.* **13** 2251–2265

Parales R E and Harwood C S 1990 Nucleotide sequence of the *gyrB* gene of *Pseudomonas putida; Nucleic Acids Res.* **18** 5880

Pasion S G, Hines J C, Aebersold R and Ray D S 1992 Molecular cloning arid expression of the genes encoding kinetoplast-associated type II topoisomerase of *Crithidia fasciculata*; *Mol. Biochem. Parasitol.* **50** 57–68

Reece R J and Maxwell A 1991 DNA gyrase: structure and function; *CRC Rev. Biochem. Mol. Biol.* **26** 335–375

Ridley R G and Kilbey B J 1994 The gene encoding topoisomerase II from *Plasmodium falciparum*; *Nucleic Acids Res.* **22** 2547–2551

Schuler G D, Altshul S F and Lipman D J 1991 A workbench for multiple alignment construction and analysis; *Proteins* **9** 180–190

Springer A L and Schmid M B 1991 Molecular characterization of *Salmonella typhimurium parE* gene; *Nucleic Acids Res.* **21** 1805–1809

Stein D C, Danaher R J and Cook T M 1991 Characterization of *gyrB* mutation responsible for low-level Nalidixic acid resistance in *Neisseria gonorrhoeae*; *Antimicrob Agents Chemother.* **35** 622–626

Strauss P R and Wang J C 1990 The TOP2 gene of *Trypanosoma brucei*: a single copy gene that shares extensive homology with other TOP2 genes encodng eukaryotic DNA topoisomerase II; *Mol. Biochem. Parasitol.* **38** 141–150

Swanberg S L and Wang J C 1987 Cloning and sequencing of the *Escherichia coli gyrA* gene coding for the A subunit of DNA gyrase; *J. Mol. Biol.* **197** 729–736

Takiff H E, Salazar L, Guerrero C, Philipp W, Huang W M, Kreisworth B, Cole S T, Jacobs W R Jr and Telenti A 1994 Cloning and nucleotide sequence of *Mycobacterium tuberculosis gyrA and gyrB* genes and detection of quinolone resistance mutations; *Antimicrob Agents Chemother.* **38** 773–780

Thiara A S and Cundliffe E 1993 Expression and analysis of two *gyrB* genes from the novobiocin producer, *Streptomyces sphaeroides*; *Mol. Microbiol.* **8** 495–506

Tsai-Pflugfelder M, Liu L F, Liu A A, Tewey K M, Whang-Peng J, Knutsen T, Huebner K, Croce C M and Wang J C 1988 Cloning and sequencing of cDNA encoding human topoisomerase II and localisation of the gene to chromosome; *Proc. Natl. Acad. Sci. USA* **85** 7177–7181

Uemura T, Morikawa K and Yanagida M 1986 The nucleotide sequence of the fission yeast DNA topoisomerase II gene: structural and functional relationships to other topoisomerases; *EMBO J.* **5** 2355–2361

Wallis J W, Chrebet G, Brodsky G, Rolfe M and Rothstein R 1989 A hyper-recombination mutation is *S. cerevisiae* identifies a novel eukaryotic topoisomerase; *Cell* **58** 409–419

Wang J C 1985 DNA topoisomerases; *Annu. Rev. Biochem.* **54** 665–697

Wang Y, Huang W M and Taylor D E 1993 Cloning and nucleotide sequence of the *Campylobacter jejuni gyrA* gene and characterization of quinolone resistance mutations; *Antimicrob Agents Chemother.* **37** 457–463

Wigley D B, Davies G J, Dodson E J, Maxwell A and Dodson G 1991 Crystal structure of the N-terminal domain of the DNA gyrase B protein; *Nature* (*London*) **351** 624–629

Wilson R, Ainscough R, Anderson K *et al* 1994 2.2 Mb of contiguous nucleotide from chromosome III of *C. elegans*; *Nature* (*London*) **368** 32–38

Wood D O and Waite R T 1994 Sequence analysis of the *Ricketsia prowazekii gyrA* gene; *Gene* **151** 191–196

Wykoff E, Natalie D, Nolan J M, Lee M and Hsieh T-S 1989 Structure of the *Drosophila* DNA topoisomerase II gene: nucleotide sequence and homology among topoisomerases II; *J. Mol. Biol.* **205** 1–13

Yamagishi J, Yoshida H, Yamayoshi M and Nakamura S 1986 Nalidixic acid-resistant mutations of the *gyrB* gene of *Escherichia coli*; *Mol. Gen. Genet.* **204** 367–373

Corresponding editor: VIDYANAND NANJUNDIAH