Available online at www.sciencedirect.com
ScienceDirect
Biochimie 89 (2007) 456–467
BIOCHIMIE
www.elsevier.com/locate/biochi

Type IA topoisomerases: A simple puzzle?

Thierry Viard [a,*], Claire Bouthier de la Tour [b]

[a] Nicholas Cozzarelli Laboratory, Molecular and Cell Biology Department, 16 Barker Hall, University of California, Berkeley, CA 94720-3204, USA
[b] Laboratoire d'Enzymologie des Acides Nucléiques, Institut de Génétique et Microbiologie, Université Paris-Sud, Unité Mixte de Recherche 8621, Centre National de la Recherche Scientifique, 91405 Orsay, France

Received 5 August 2006; accepted 20 October 2006
Available online 15 November 2006

Abstract

Type IA topoisomerases are enzymes that can modify DNA topology. They form a distinct family of proteins present in all domains of life, from bacteria to archaea and higher eukaryotes. They are composed of two domains: a core domain containing all the conserved motifs involved in the trans-esterification reactions, and a carboxyl-terminal domain that is highly variable in size and sequence. The latter appears to interact with other proteins, defining the physiological use of the topoisomerase activity. The evolutionary relevance of this topoisomerase-cofactor complex, also known as the ''toposome'', as well as its enzymatic consequences are discussed in this review.

Keywords: Type IA topoisomerase; Topoisomerase I; Topoisomerase III; Three-dimensional structure; Protein folding; Evolution; DNA topology

''As enzymes, the DNA topoisomerases are magicians among magicians; they open and close gates in DNA without leaving a trace, and they enable two DNA strands or duplexes to pass each other as if the physical laws of spatial exclusion do not exist'' (James C. Wang, 1982). Written in the entry door of the Enzymologie des Acides Nucléiques laboratory.

1. Introduction

The year 1953 can be regarded as year zero of molecular biology. Before this period, it was known that chromosomes consist of DNA and that this molecule is the genetic data carrier. But nothing was known about the molecular structure of DNA. The model proposed by the Watson–Crick duo (personified since by many students as one and only one person) put forth the following theory: DNA is composed of two anti-parallel strands interwoven to form a ladder in the shape of a right-handed helix [1,2]. This model remarkably explained the parity of bases A-T and G-C in the chemical composition of the DNA that is found in all living organisms.

But in 1953, many disagreed with this rigid and too well adjusted structure. At that time, the principal objection related to the orientation of the bases towards the interior of the double helix and the inaccessibility of the bases from the outside of the helix. Such an organization implied necessarily the opening of the double helix to reach the nucleotide sequence. However, because of their interlaced organization, the two strands were impossible to separate and any attempt to do so would break the molecule, limiting the physiological relevance of the model. For that reason, 1971 should probably also be considered, even among molecular biologists, as a very important year in DNA structure history. The discovery of the ω protein [3] in 1971 showed that evolution had indeed provided tools for manipulating the topology of DNA, the DNA topoisomerases. This discovery placed this universal family of proteins at the forefront of molecular biology, bringing a decisive justification to the double helix structure.

Topoisomerases solve all the topological problems that are related to the physical structure of the double helix of the

* Corresponding author. Tel.: +1 510 642 7108.
E-mail address: viard@berkeley.edu (T. Viard).

0300-9084/$ - see front matter © 2006 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.biochi.2006.10.013

DNA by coordinately cleaving, manipulating, and religating DNA strands together in the same catalytic event. They can modify DNA superhelicity to allow specific binding of proteins or to allow local and temporary unwinding of the two DNA strands as routinely required during replication or transcription [4]. They can also decatenate circular DNAs or disentangle large DNA segments, allowing plasmids or chromosomes to be separated without any breakage or loss of genetic material [5], as needed for correct transmission of genetic material into daughter cells during cell division [6,7]. As expected, topoisomerase mutant or deficient cells indeed exhibit defects in genome replication [6,7] and chromosome partitioning [8,9], slow growth and low survival rates [10], leading to serious pathologies in humans [11,12].

The double helix model really found its best ally in topoisomerases. The existence of topoisomerases was indeed, almost by itself, sufficient to validate the double helix structure, as DNA unfolding and disentanglement during fundamental metabolic processes of the DNA were no longer an obstacle. In addition, this argument also could be used reciprocally: ''If the double helix structure is wrong, why do enzymes able to solve problems which would result from it exist?'' (Michel Duguet, Informative Macromolecules Technologies course, September 1999).

Topoisomerases are classified according to the number of strands they cleave: the enzymes cleaving only one strand of DNA are called type I topoisomerases, while type II topoisomerases cleave both strands of the DNA. Each type can be subdivided into two groups based on their structural homologies, and for type I topoisomerases only, on the way they covalently bind to their substrate.

In this review, we will focus on type IA topoisomerases, mainly on their structures and their potential in vivo roles. Type II and IB (sometimes also referred as type I-3') enzymes will not be described. We encourage the reader to refer to comprehensive reviews for more complete coverage of these two other very important and fascinating topoisomerase families [13–16].

The best example of type IA topoisomerases is the well studied and previously mentioned ω protein [3], now known as *Escherichia coli* topoisomerase I (*Ec*TopoI). The enzymatic activities of these enzymes are based on a succession of events that modify the topological state of the substrate by changing the linking number (the number of times the two strands of a duplex are wrapped around each other [17]) in discrete steps of one [18,19]. First, type IA topoisomerases (with the noticeable exception of the Reverse Gyrase) require an exposed single-stranded region within the substrate DNA [20]. Once bound, they cleave one of the two strands of DNA, defined as the G-segment (for Gate). During the cleavage process, both DNA extremities are bound to the protein: the 5′ end is covalently attached to the catalytic tyrosine while the other is non-covalently but tightly bound to the protein [21,22]. Then, using the torsional stress stored in the DNA molecule as energy, type IA topoisomerases pass the other strand (T-segment for Transported) through the previously described gap. Finally, the breach is religated and the enzyme resets. After resetting, the enzyme may initiate another round of catalysis or, because type IA topoisomerases preferentially bind single-stranded DNA, dissociates as the affinity has been lowered by the last reaction cycle. This latter attribute explains why type IA topoisomerases cannot complete a full relaxation: before reaching its completely relaxed state, the supercoiling is reduced below the level that allows single-stranded DNA regions to exist. The enzyme is then ejected [23].

Because type IA topoisomerases can conduct such amazingly complex DNA manipulations, they are considered as wonderful biomolecular machines. How they are able to coordinately manage such a complex series of steps cleaving, manipulating, and religating DNA strands together at such high speeds and with remarkable fidelity, just by using the energy of DNA supercoiling is really puzzling, but it seems that the answer might just be found by looking at how the parts of this machine are put together.

2. All the same

Type IA topoisomerases form a vast family of conserved enzymes present in all the domains of life, from bacteria to archaea, and higher eukaryotes. Sequence analysis indicates that all type IA topoisomerases are generally composed of two parts as follows: a core, containing all the conserved motifs, particularly those forming the active site of the protein [13,14,24,25], and a carboxyl-terminal end, highly variable in size and sequence (Fig. 1a). Structural data confirmed that all type IA topoisomerase cores share the same common toroidal architecture [26–29] formed by several conserved protein domains (Fig. 1b). In *E. coli* topoisomerase I (*Ec*TopoI), domain I is composed of the first 158 amino acids (129 for *Thermotoga maritima* topoisomerase I, *Tm*TopoI). Domain II consists of a discontinuous amino acid sequence from residues 213 to 279 and 407 to 477 (186–239 and 376–443, in *Tm*TopoI). Domain III corresponds to the amino acids 280–406 (240–375, in *Tm*TopoI). Domain IV links domain I to domain II with the amino acids 159–213 (130–185, in *Tm*TopoI) and ends the toroidal structure from amino acid 477 to 596 (443–542, in *Tm*TopoI). Recently, a three-dimensional structure determination of the full-length topoisomerase I from *T. maritima* has revealed a new domain, domain V, comprised of the last amino acids of the protein (from 543 to 633) including the unique zinc motif of the protein [30]. This domain is tightly associated with domain IV and is composed of a four-stranded β sheet resembling a zinc ribbon fold found in transcription factors [31]. Domain V is absent in some topoisomerases. Domains IV and V present important variations between species and will be extensively discussed later.

All these five domains exhibit very specific and well-defined interactions with each other: domains III, II and IV form a claw whose central width measures approximately 27.5 Å (25 Å in *Tm*TopoI). The surface of the cavity is peppered with lysines and arginines, giving an overall positive charge potentially able to interact with DNA [26]. The size of this cavity, importantly, is sufficient to hold either a single-stranded or a double-stranded molecule of DNA.