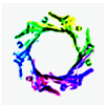
IJCB
International Journal of Computational Biology
ISSN: 2229–6700, E-ISSN: 2229–6719, Vol. 2, Issue 1, 2011, pp-14-23
Available online at http://www.bioinfo.in/contents.php?id=96



AN IN SILICO ANALYSIS OF TELOMERASE AND TELOMERE BINDING PROTEINS

SAYANI MITRA, SAYAK GANGULI* AND ABHIJIT DATTA
DBT-Centre for Bioinformatics, Presidency University, Kolkata
*Corresponding author. E-mail: sayakbif@yahoo.com

Received: December 13, 2010; Accepted: December 28, 2010

Abstract- Telomerase, an enzyme which plays a key role in the maintenance of the ends of chromosome in eukaryotes have been found activated in many cancer cells. Telomerase is an enzyme that adds specific DNA sequence repeats *i.e.,* "TTAGGG" in all vertebrates to the telomere region and thereby maintains the telomere length along with other telomere binding proteins. A group of 6 proteins make up the Shelterin complex which is responsible for telomere capping. The objective of this work is to detect the homologues with the query sequences, identify evolutionarily conserved regions in the query, find patterns of concerted evolution by detecting conserved domain architectures in the query and subject proteins, phylogenetic relationships amongst the query and subject proteins were analyzed to match them with the concerted evolution pattern and Telomerase inhibitors were screened to detect the best fit molecule which could possibly act to prevent carcinogenesis.
Keywords- Telomerase, Shelterin Complex, The Lagging Strand Mechanism, Curcumin

Introduction

The ends of linear chromosomes i.e the telomeric region in the eukaryotic cells undergoes major problems during DNA replication. The lagging strand mechanism of DNA replication cannot copy all the way to the end of the linear molecule resulting in short telomeric regions. The shortening of telomeric region and gradually its disintegration accounts for causing replicative aging and limits the proliferative capacity of normal human cells.

Telomeric sequences have found to be conserved among eukaryotic vertebrates throughout evolution. However the length of telomeres differs between species. Humans telomeric region is restricted up to 20 kb in length while Mus spretus shares nearly same size of the telomeric region of that of human by having a telomere upto 30 kb in size [1]. Telomeres of rodents on the other hand have been reported to be heterogeneous in length with rat telomere lengths ranging from 20 to 100 kb[2]. Mus musculus has been reported to have telomeres up to 150 kb in size [3].

Telomerase is an enzyme that helps to maintain the telomere length by adding specific DNA sequence repeats to the 3' end of the DNA strand in the telomeric region. Being a Reverse transcriptase, it carries its own RNA molecules which functions as a template during the elongation of chromosome in the replication mechanism. Telomerase was discovered by Carol.W.Greider and Elizabeth Blackburn in 1984 in the ciliate Tetrahymena [4] and the protein composition was identified by Scott Cohen and his team in 2007.The protein consists of two molecules, each of telomerase reverse transcriptase (TERT), telomerase RNA (TR or TERC), and dyskerin (DKC1) [5]. The TERC(Telomere RNA component) provides an AAUCCC template to guide the insertion of TTAGGG [6]. The TERT(Telomere Reverse Transcriptase) provides with the catalytic action.

Other than the telomerase other telomere binding protein also provide essential functions in the chromosome maintenance and is thereby suited for analysis in the context of evolution. The first telomere binding protein was identified in the ciliate Oxytricha nova. Telomere binding protein can either be double as well as single stranded ,both of which has been conserved throughout the evolution. Identified in budding and fission yeast and even in humans, double stranded telomere binding proteins are negative regulators of telomerase .They function by counting the number of existing telomere repeats on each chromosome end [7]. While double stranded protein functions in only negative regulation, single stranded telomere binding proteins performs multiple function, both negatively and positively. They influence the extension of telomere telomerase as well as provide a direct link between leading and lagging strand DNA in the replication machinery . The proteins utilizes the oligosaccharide/oligonucleotide fold(OB fold) for efficient binding to telomeric DNA [8]. Some telomere binding proteins are – TRF1(Telomere

Repeat Binding Factor1), TRF2(Telomere Repeat Binding Factor 2), hRAP1, TIN2(TRF interacting nuclear factor2), TANK1(Tankyrase,TRF1 interacting ankyrin related polymerase), TANK2 (Tankyrase2).

Other telomere binding proteins like Pot1 and Pot2 are also involved in regulations of the telomeric length, might possibly function by regulating telomerase access to telomeres [9]. Pot1 protein also functions by protecting telomeres from nuclease attack [10]. Pot2 protein which is another component of the protective cap, restricts the access of the non- homologous end joining(NHEJ) machinery to the chromosomes ends [11]. Depending on the functions of these telomere binding proteins, a complex comprising of 6 proteins- TRF1, TRF2, POT1, RAP1, TIN2 and TPP1 have been formed, which is known as the Shelterin Complex or the Telosome [12].

The presence of telomerase in various human cancers and its absence in many normal cells has raised a possibility of the enzyme serving as a good target for anticancer drugs. Agents which can restrict the activity of telomerase might kill tumor cells(by allowing telomeres to shrink and gradually disintegrate) without disrupting the functioning of the normal cells. This is in contrast to most anticancer therapies which are not only toxic but at the same time disturbs normal cells as well as malignant ones .Since telomerase is essential for the immortality of different cancer types , it is thereby thought to be a potential drug target. Moreover, Blackburn has also shown that most of the harmful cancer-related effects of telomerase are dependent on an intact RNA template .This provides a worthy target for drug development [13].

Curcumin, among the various anti cancer agents , exhibit therapeutic potentials against variety of different cancers including leukemia and lymphoma; gastrointestinal cancers, genitourinary cancers, breast cancer, ovarian cancer ,head and neck squamous cell carcinoma, lung cancer, melanoma, neurological cancers and sarcoma [14].

It is also reported to be an effective inhibitor of angiogenesis (a fundamental step in the transition of tumors from a dormant state to a malignant one) both in vitro and *in vivo* [15]. In addition to anti cancer treatment, curcumin has also exhibited functions in boosting the immune system by increasing the production of proteins that cause immune cells to proliferate and reduces the production of proteins that destroy immune cells [16]. Curcumin is one of the major components of the Indian spice turmeric which is a member of the ginger family (Zingiberaceae).Cui et al(2006) has reported that Curcumin inhibits telomerase activity in cancer cell lines and thus induces apoptosis [17]. Thus Curcumin has a high prospect of serving as an anticancer drug.

Materials and Method

Since the work was performed absolutely in silico hence the materials were database sequences which were downloaded from the public database such as SWISS-PROT and GenBank of NCBI. Nucleotide and protein sequences of all Telomere Binding Proteins were taken from GenBank of NCBI by performing BLASTn and BLASTp respectively. Conserved Domains were searched using the Conserved Domain search (CD search) tool to identify evolutionarily conserved regions in the query. Conserved domain architecture retrieval tool (CDART) was used to find patterns of concerted evolution by detecting conserved domain architectures in the query and subject proteins. Clustal W was used to analyze the phylogenetic relationships amongst the query and subject proteins. Finally, telomerase inhibitors were screened to detect the best fit molecule which could possibly act to prevent carcinogenesis. Docking was the performed using AUTODOCK and were visualized using RASMOL .Following this QSITEFINDER was used to generate the residues of the binding sites.

Results and Discussion

BLAST analysis revealed that the telomere binding proteins had paralogues in Protozoa, Yeasts, Algae, and Vertebrates. Results are obtained by taking the lowest expected values or with e values near to 0.0. Conserved domains were observed in which RNA binding domains were predominant. This proves the essentiality of being a part of the end protection maintenance-replication machinery is to be able to bind RNA. 21 common domains were found in different query sequences taken respectively(Fig 1) These domain architectures were investigated and the most predominant domains were found to be those associated with viruses and DNA replication machinery. The presence of many domains of unknown functions was also detected in the results indicative of new domain acquisition in the proteins an evidence of continuous evolution. Multiple catalytic motifs detected are indicative of the fact the molecules are probably RNA world remnants and in pre-biotic conditions these catalytic RNA motifs would probably have performed the functions which coded proteins are performing today. Phylogenetic analysis showed that mammalian sequences are clustered in a single node and the plant sequences served as outgroups to the other sequences. (Fig 3)UPGMA trees generated with the Shelterin complex proteins as query, produced significant alignment among all the homologues. This leads us to conclude that the proteins follow the pattern of speciation as well. Cladograms show that sister groups belong to the unique class of species such as primates, protists etc.(Fig 2) Curcumin which is a pleiotropic molecule inhibits telomerase activity in cancer cell lines and thus

induces apoptosis. Through our analysis we it was founded that curcumin binds specifically to the RNA binding domain of telomerase and thus inhibits its activity. The E-Total observed during docking was -435.75,which is low as compared to other existing telomerase inhibitors in the DrugBank. This proves that Curcumin has a potential to serve as a possible telomerase inhibitor.(Fig 4,5,6)

References

[1] Meyne J., Ratliff R.L. & Moysis R.K. (1989) *Proc. Natl Acad. Sci. USA*, 86, 7049–7053.

[2] Prowse K.R. & Greider C.W. (1995) *Proc. Natl Acad. Sci. USA*, 92, 4818–4822.

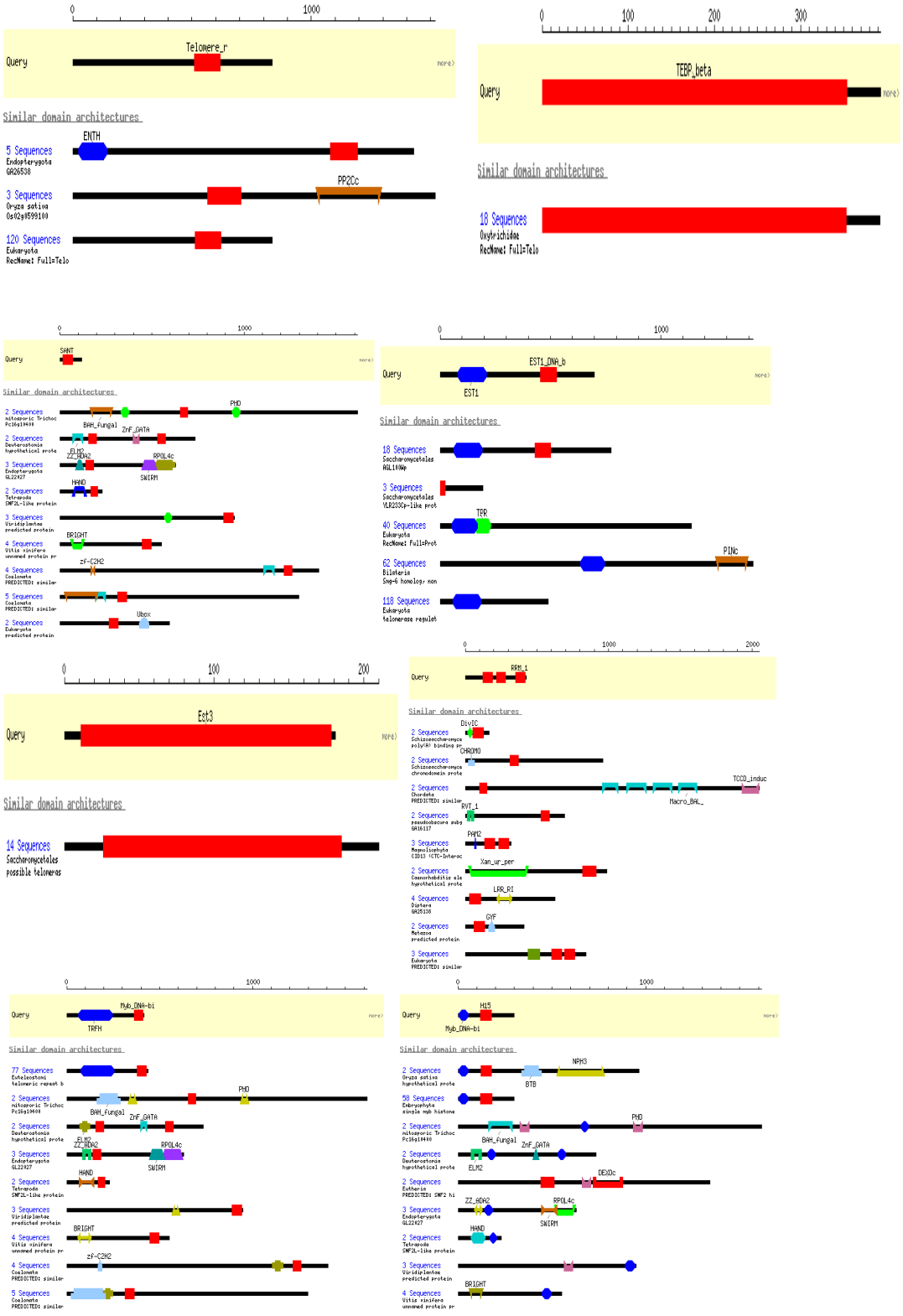
[3] Zijlmans J.M., Martens U.M., Poon S.S., Raap A.K., Tanke H.J., Ward R.K. and Lansdorp P.M. (1997) *Proc. Natl Acad. Sci. USA*, 94, 7423–7428.

[4] Greider C.W and Blackburn E.H. (1985) *Cell* 43,405-413

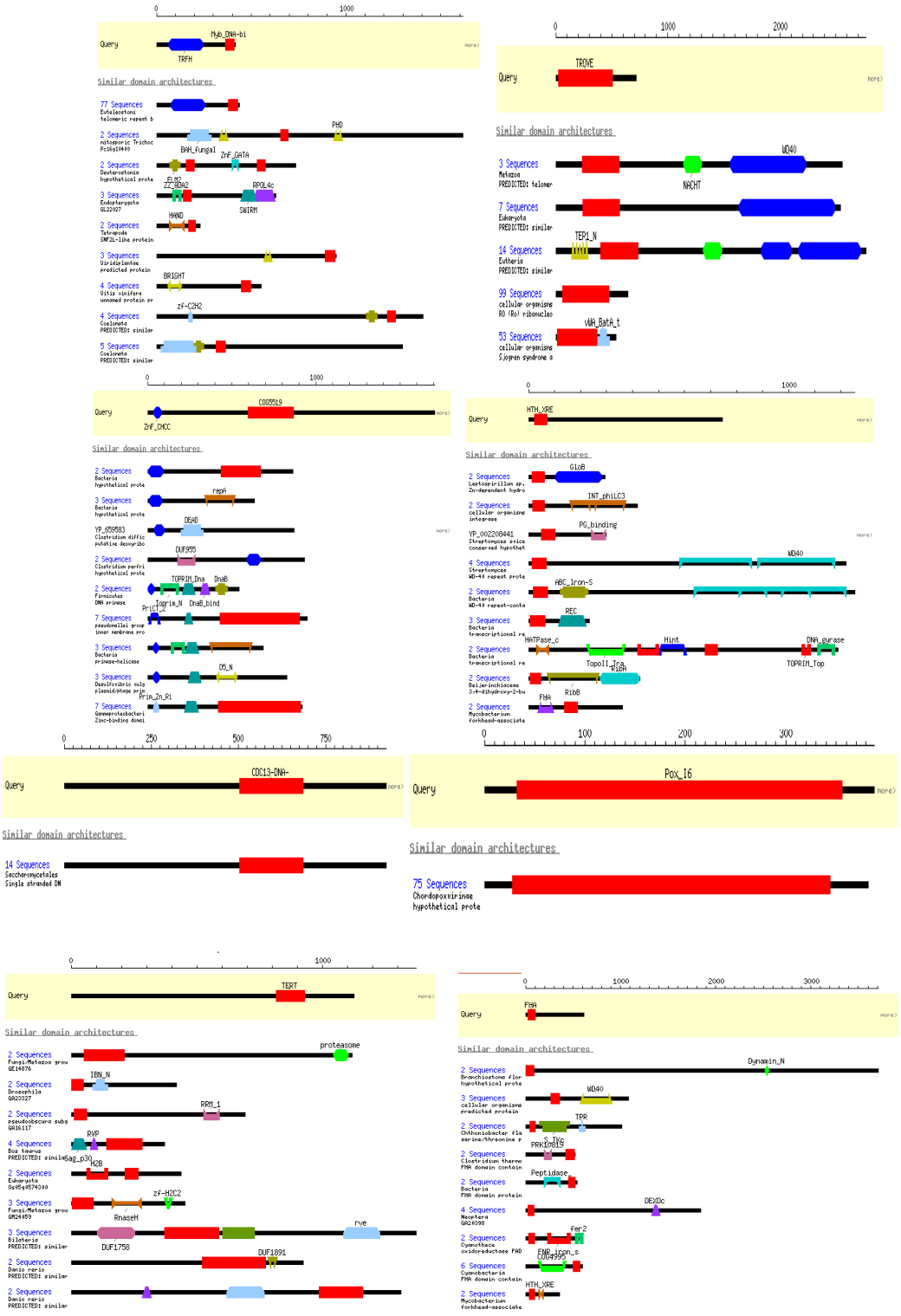
[5] Cohen S., Graham M., Lovrecz G., Bache N., Robinson P., Reddel R. (2007) *Science* 315, 5820.

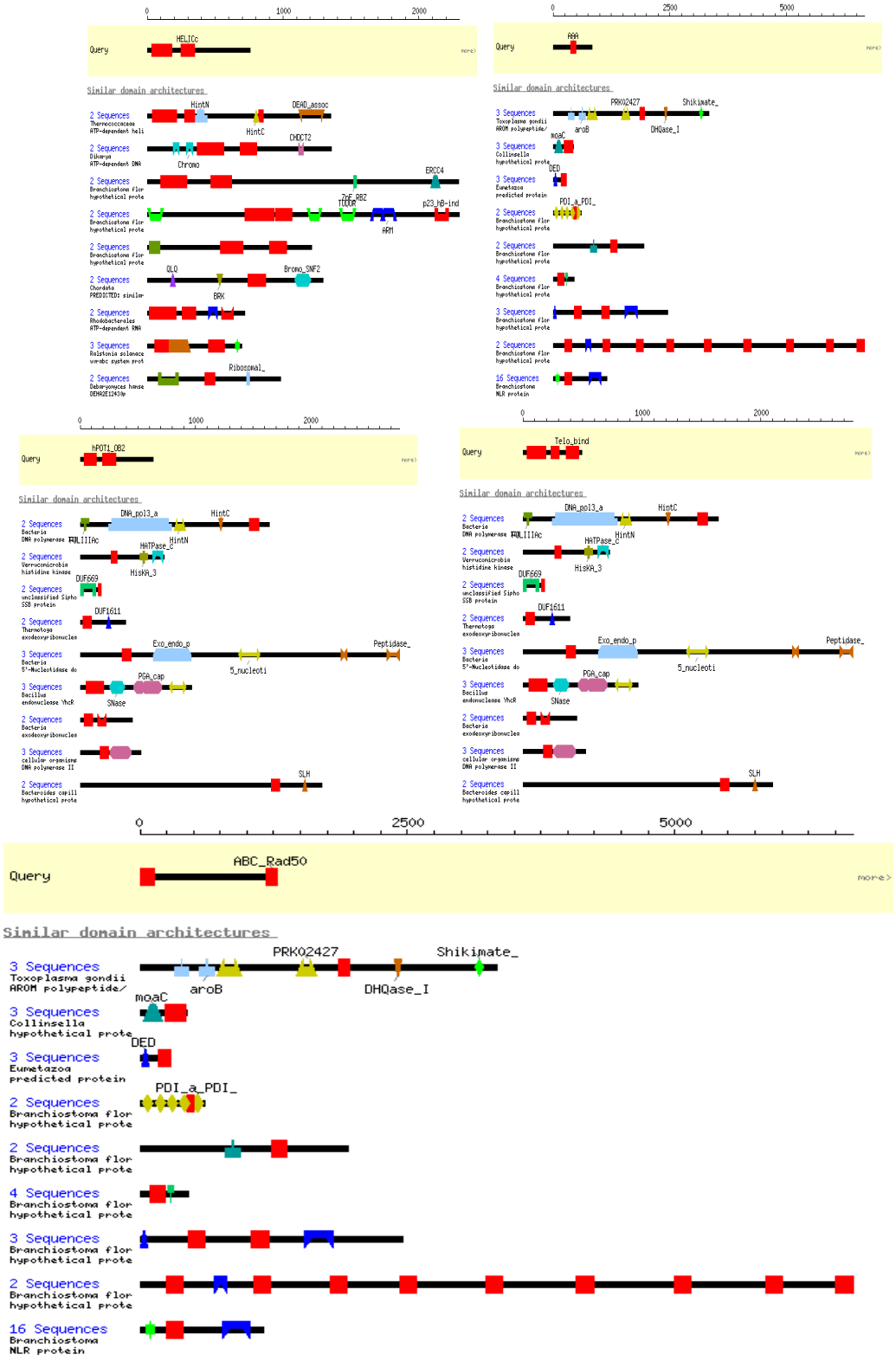
[6] Leticia R. Vega, Maria K. Mateyak & Virginia A. Zakian (2003) *Nature Reviews Molecular Cell Biology* 4, 948-959

[7] Reuter S., Eifes S., Dicato M., Aggarwal B.B, Diederich M. (2008) *Biochem Pharmacol*., 76, 1340–1351.



International Journal of Computational Biology
ISSN: 2229–6700 , E-ISSN: 2229–6719, Vol. 2, Issue 1, 2011





Similar domain architectures

3 Sequences
Toxoplasma gondii
AROM polypeptide/
PRK02427
Shikimate_
aroB
DHQase_I

3 Sequences
Collinsella
hypothetical prote
moaC

3 Sequences
Eumetazoa
predicted protein
DED

2 Sequences
Branchiostoma flor
hypothetical prote
PDI_a_PDI_

2 Sequences
Branchiostoma flor
hypothetical prote

4 Sequences
Branchiostoma flor
hypothetical prote

3 Sequences
Branchiostoma flor
hypothetical prote

2 Sequences
Branchiostoma flor
hypothetical prote

16 Sequences
Branchiostoma
NLR protein

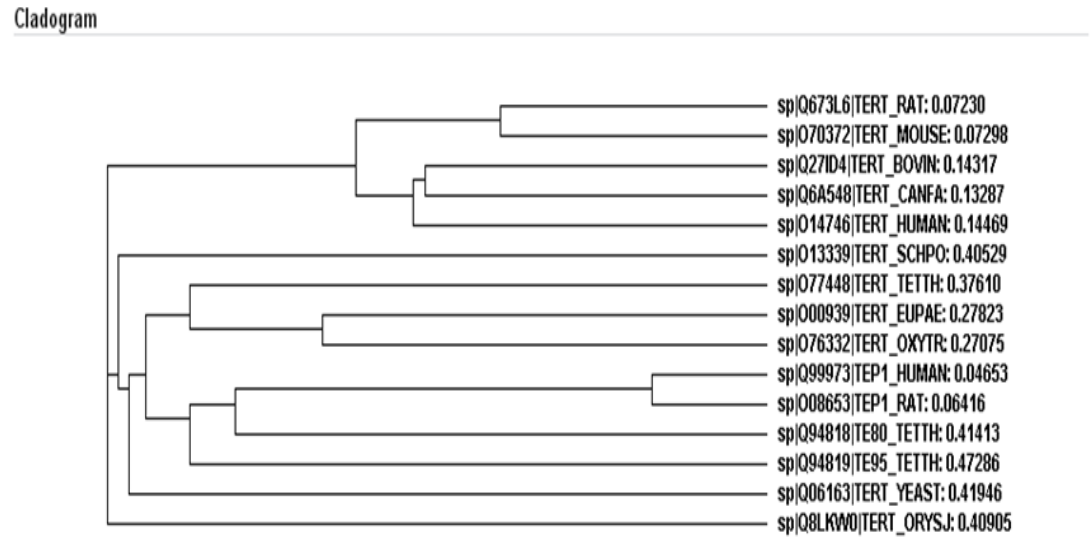
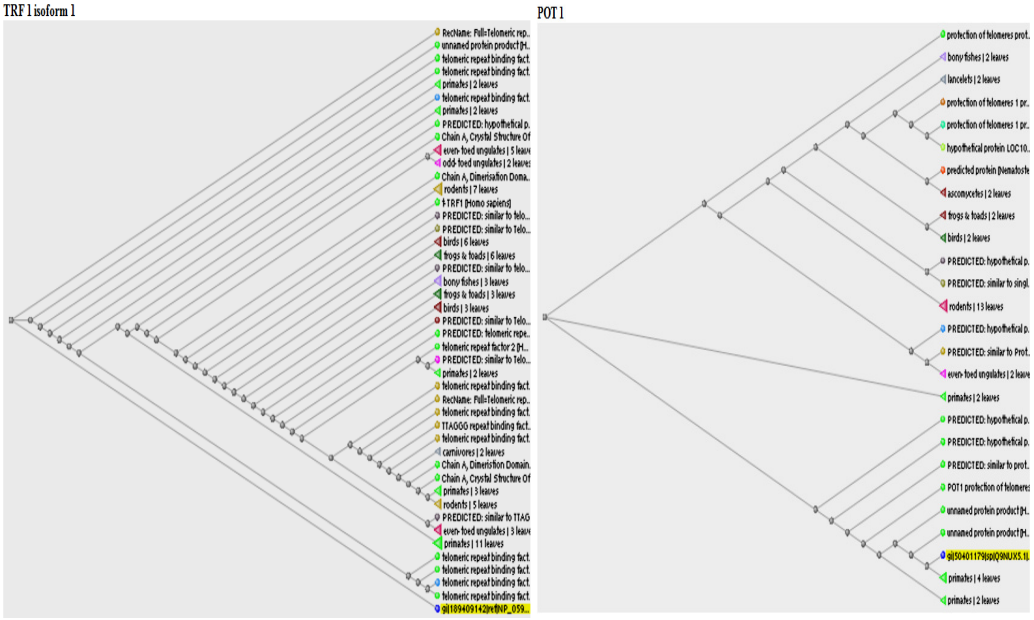
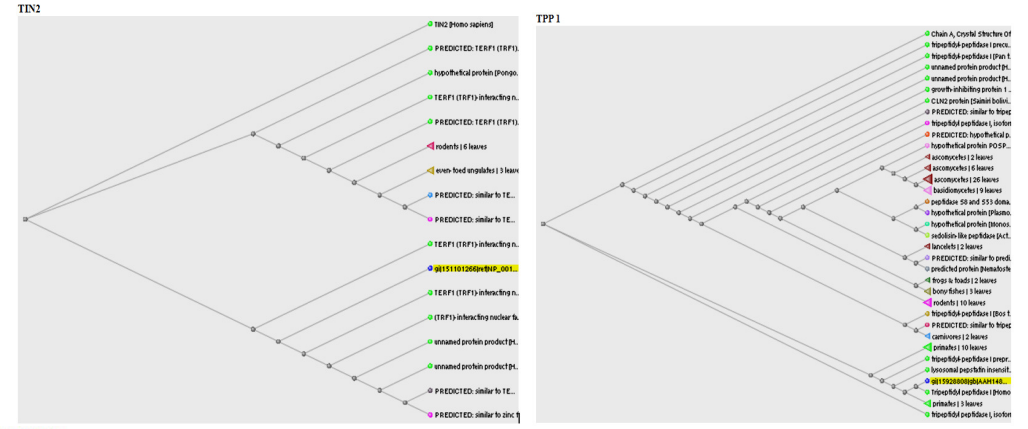


Fig. 2- Cladogram depicting evolutionary relationship among Telomerase enzyme





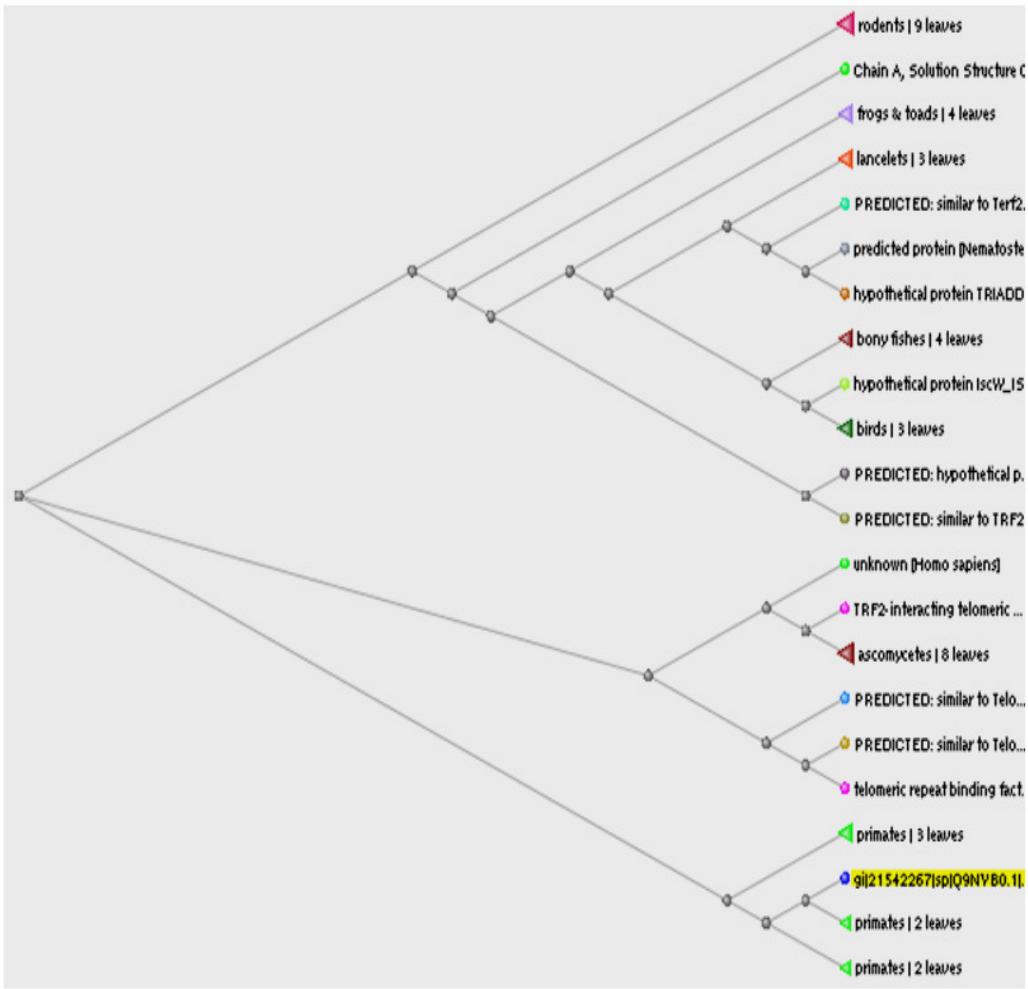


Fig 3: Phylogenetic profiles of the Shelterin Complex

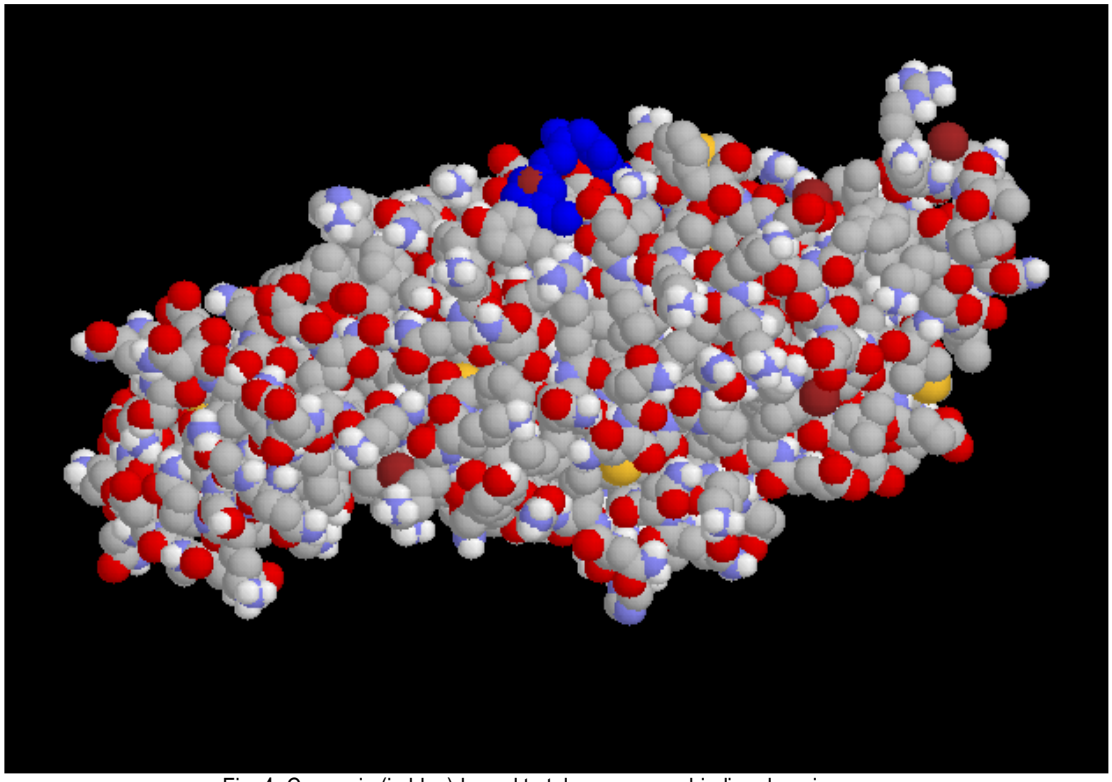


Fig. 4- Curcumin (in blue) bound to telomerase rna binding domain

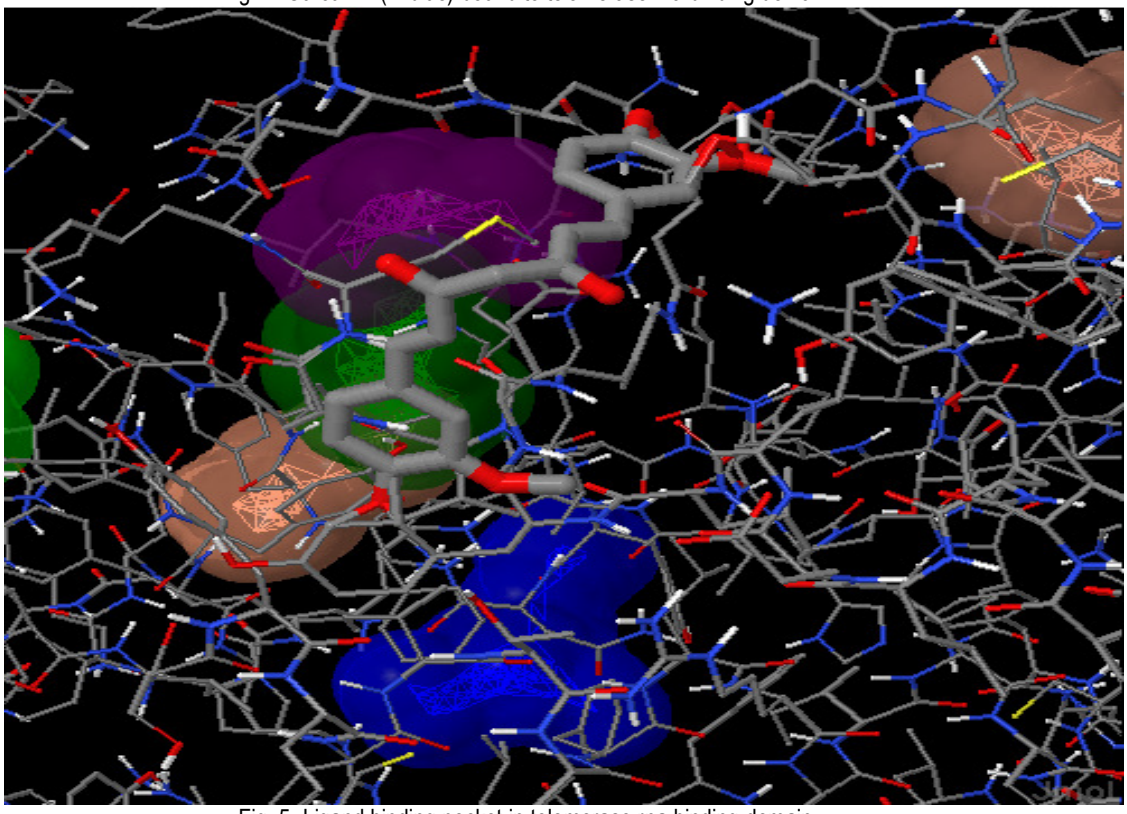


Fig. 5- Ligand binding pocket in telomerase rna binding domain

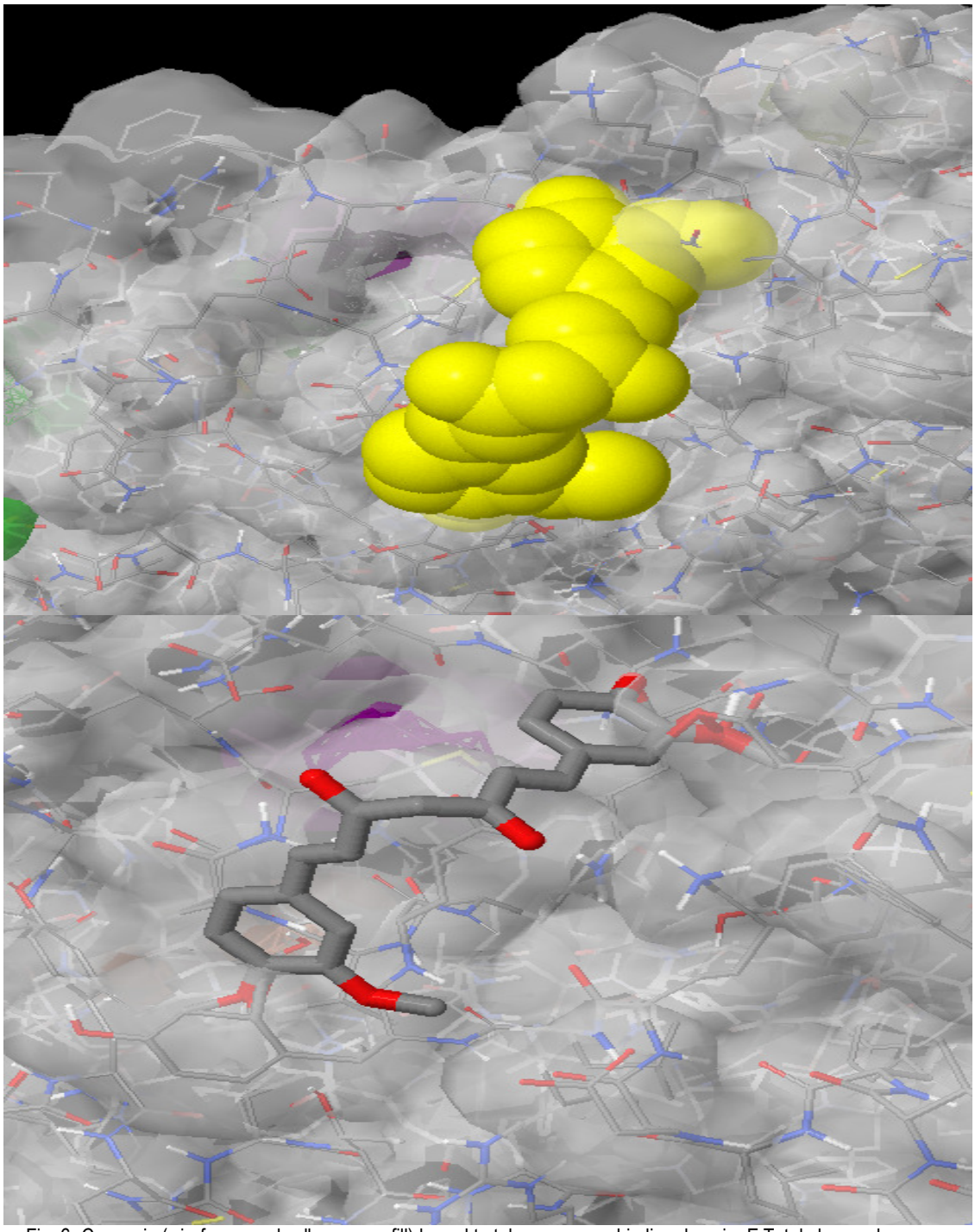


Fig. 6- Curcumin (wireframe and yellow space fill) bound to telomerase rna binding domain. E Total observed during docking was -435.73