INSIGHTS INTO DISEASES OF HUMAN TELOMERASE FROM DYNAMICAL MODELING

SAMUEL COULBOURN FLORES
Cell and Molecular Biology Department, Uppsala University, Biomedical Center, Box 596, 75124 Uppsala, Sweden
Email: samuel.flores@icm.uu.se

GEORGETA ZEMORA
Max F. Perutz Laboratories, University of Vienna, Dr. Bohrgasse 9/5, 1030 Vienna, Austria

CHRISTINA WALDSICH
Max F. Perutz Laboratories, University of Vienna, Dr. Bohrgasse 9/5, 1030 Vienna, Austria

Mutations in the telomerase complex disrupt either nucleic acid binding or catalysis, and are the cause of numerous human diseases. Despite its importance, the structure of the human telomerase complex has not been observed crystallographically, nor are its dynamics understood in detail. Fragments of this complex from Tetrahymena thermophila and Tribolium castaneum have been crystallized. Biochemical probes provide important insight into dynamics. In this work we summarize evidence that the T. castaneum structure is Telomerase Reverse Transcriptase. We use this structure to build a partial model of the human Telomerase complex. The model suggests an explanation for the structural role of several disease-associated mutations. We then generate a 3D kinematic trajectory of telomere elongation to illustrate a "typewriter" mechanism: the RNA template moves to keep the end of the growing telomeric primer in the active site, disengaging after every 6-residue extension to execute a "carriage return" and go back to its starting position. A hairpin can easily form in the primer, from DNA residues leaving the primer-template duplex. The trajectory is consistent with available experimental evidence. The methodology is extensible to many problems in structural biology in general and personalized medicine in particular.

1. Introduction

Telomere maintenance has broad implications for human health and longevity. Positive lifestyle changes including exercise and smoking cessation result in increased telomerase activity in the immune system.(1) Cancer cells also have vigorous telomerase activity (1) as required for immortality. (2) Despite its importance, the human Telomerase complex has not been solved crystallographically. However various components of the complex from other organisms are available. We use these to predict the structure of key components of human Telomerase by a novel homology modeling approach and illustrate its functional mechanism. We also give a possible explanation for the role of several disease associated mutations.

The primary enzymatic task of Telomerase is the extension of the leading telomeric DNA strand following the sequence of an RNA template. This is done one DNA residue at a time at the active site of the Reverse Transcriptase (RT). The RT domain shares motifs with RNA and DNA polymerases, suggesting similar function and mechanism.(3) The RNA Binding Domain (RBD) appears to help position the primer-template duplex for extension.(3) The carboxy-terminal extension (CTE) is implicated in DNA binding.(3)

In this work we focus on the RT, RBD, and the primer-template duplex. Key regions of RBD and most of the RT are conserved across organisms. In particular a published multiple sequence alignment (MSE)(4) shows considerable sequence identity between human and *T.castaneum* ("beetle") in a substantial portion of TERT. A crystallographic structure of beetle TERT with a putative primer-template duplex now exists, but considerable debate exists as to whether it is truly telomerase;(5) in this work we present the evidence that it is. We then use the beetle structure as the template for a partial homology model of human Telomerase. The model explains the probable role of disease associated residues and is consistent with biochemical probes of structure and dynamics.

Considerable insight into the dynamics of primer elongation exists from biochemical experiments. The primer-template duplex is always about 7 base pairs (bp) long as determined by dimethyl sulfate footprinting assays in yeast,(6) because bp's are denatured at the distal end as they form at the proximal end. In human, the template region is 6 bases long while an alignment region on the 3' side adds 5 bases for a total of 11; however it is believed that 11 base pairs would never form simultaneously during elongation as subsequently denaturing multiple base pairs would require too much energy. (7) Dissociation would also be difficult for an excessively long duplex. (7)

When the extending primer reaches the 5' end of the template, the template disengages and reattaches to the primer having shifted by six residues, ready for another six-residue extension to be be added (6). Meanwhile as DNA residues exit the primer-template duplex, they queue to join a hairpin or quadruplex which may help drive processivity.(8)

The N-Terminal domain (TEN) is a low-conservation region of TERT (3) known to be important for primer positioning and elongation. (9) *T.Castaneum* telomerase has no TEN domain, a point we will return to.
Additional insight comes from prior structural modeling. The TEN, RBD, RT and primer-template duplex domains were ambitiously predicted by homology modeling and docking by Steczkiewicz and collaborators. (10) However the primer-template duplex in that model is about 15 (10) rather than 7 bp long as it is in yeast (6) and is even longer than the 11 bp discussed above. Also, TEN residues 170-175 which have been experimentally implicated in the active site (9) are about 9 RNA (not DNA) nt away from the active site in the model. (10) That model also includes the CTE, which we did not model due to the low sequence identity between beetle and human in that domain. Lastly, the mechanism of processivity proposed in that work is based on a low-order normal mode expansion. The RBD and RT domains of that model, on the other hand, agree with those presented in this work.

In this work we thus present a knowledge-based structural and dynamical model of human telomerase, including much of TERT, the template, and a telomeric extension. We generate the structural model of TERT by homology modeling. The dynamical model incorporates additional biochemical information. We address the debate on the function of the beetle structure. (11) The results provide an explanation for the role of various disease associated mutations. Our model also supports a role in processivity for the primer extension hairpin. We show that MMB (formerly RNABuilder) (12) is a structural and dynamical modeling code with many potential applications in molecular biology. Its economy, versatility, and ease of use make it a good tool to use for examining the effect of individual genetic variation on disease phenotype.

2. Methods

2.1 *Validating human-beetle sequence alignment using secondary structure*

Generating the structure of the human TERT by threading to the existing beetle TERT structure (13) requires a multiple sequence alignment, which is available from the telomerase database.(4). Note that not all domains have sufficient sequence identity for alignment. As a first validation of the alignment, we predicted the secondary structure of the human TERT using the Jpred server, which is not biased by the use of the known *T.thermophila* or beetle TERT structures.(14)

2.2 *Aligning T. thermophila RNA Binding Domain to beetle TERT*

For further validation of the MSE and to support telomerase function of the beetle structure, we used the former as the basis for a rigid-body structural alignment (15,16) of the crystallographically observed *T.thermophila* RNA Binding Domain (RBD) (13) onto the beetle TERT. (11)

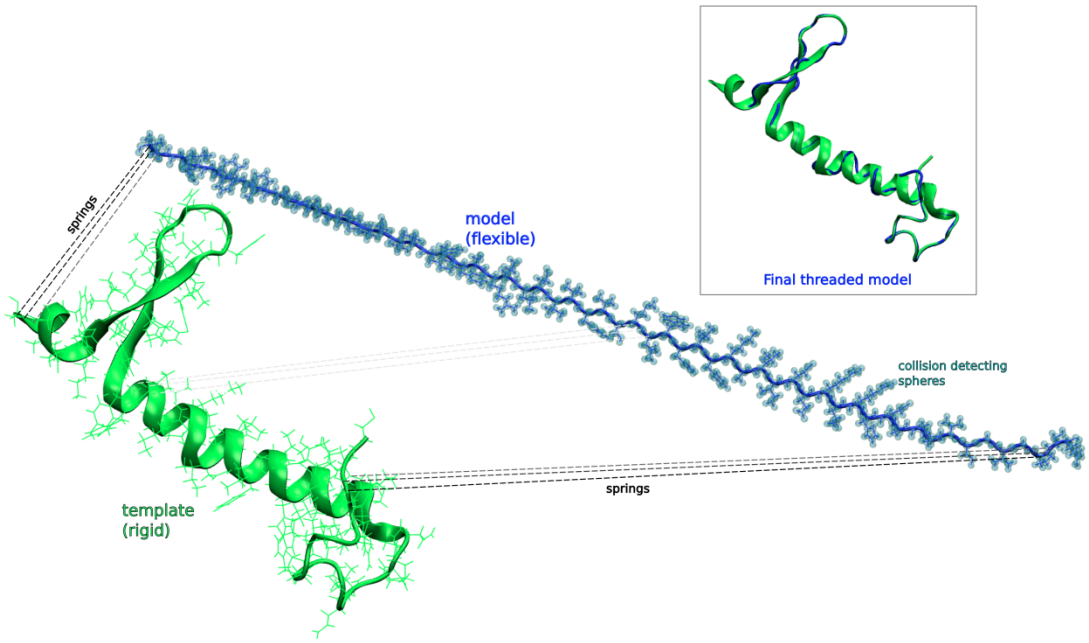


Figure 1. Illustration of internal coordinate threading procedure.
Threading is done as follows. The template (here a fragment of the beetle RBD is used, in green) is made rigid. In the model (here the corresponding human peptide fragment, in blue), bond lengths and angles are fixed, but torsion angles are free to vary (except for proline which has additional freedom for ring closure). Springs connect backbone atoms in the template with backbone atoms in the model which correspond according to sequence alignment (a representative set are shown as black dashed lines). Collision detecting spheres (transparent cyan) prevent steric clashes within the model. The system dynamics are allowed to proceed until the backbones are aligned. The final threaded model (inset) has a backbone RMSD of 1.002Å with respect to the template in this example.

2.3 *Threading human to beetle TERT*

In prior work we showed how MMB/RNABuilder can be used for RNA threading. (16) In this work we show that the package can also do protein threading.(17) The threading forces can be combined with other forces, constraints, and coordinate matching features available in this full-featured modeling package. We aligned a flexible human TERT protein backbone to the rigid beetle TERT(11) template by connecting corresponding backbone atoms with springs, while steric clashes were economically prevented by means of collision detecting spheres.(15) The approach of using internal coordinate dynamics(18) to align the backbone in this way is unique to this work;(17) most other protein threading algorithms work by rigid fragment assembly, segment matching, spatial restraint, and artificial evolution.(19) The approach is preferred for this work because it is economical, gives us full control over the alignment, conserves chemistry and sterics, allows protein, DNA, and RNA(16) to be threaded simultaneously, and permits dynamical (16) rather than only static modeling of the mechanism of primer extension as we will show. An illustrative example of this process is shown in Figure 1 above.

The basis for the correspondence for all threaded TERT fragments was the Telomerase Database MSE.(4) Note that much of TERT is highly diverged and therefore not all residues are aligned (Figure 2). (4).

```
525                  545 (RBD)          563    570   576
GVGCVPAAEHRLREEILAKF LHWLMSVYVVELLRSFFYVTETTFQKNRLFFYRKSVWSKLQSIGIRQHLKRVQLREL
------HHHHHHHHHHHHH HHHHH--HHHHHH---EEEEEE----EEEEEEEHHHHHHHHHHHHHHHHH-------
YDAIPWLQNVEPNLRPKLLL HNLFLLDNIVKPIIAFYYKPIKTLNGHEIKFIRKEEYISFESKVFHKLKKMKYLVEV
---HHHHH-----HHHHHHH HHHHHHHHHHHHHHHHHEEEEE------EEEEEHHHHHHHHHHHHHHHHH--EEE-

602   617                    648                    681                704
SEAEV LTSRLRFIPKPDGLRPIVNMDY EKRAERLTSRVKALFSVLNYERA LGLDDIHRAWRTFVLRVR PELYFVKVD
-HHHH ---EEEEEE----EEEEEE--- --HHHHHHHHHHHHHHHHHHHH ---HHHHHHHHHHHHHH --EEEEEE-
QDEVK PRGVLNIIPKQDNFRAIVSIFP DSARKPFFKLLTSKIYKVLEEKY KTSGSLYTCWSEFTQKTQ GQIYGIKVD
----- --EEEEEE----EEEEEE-- ---HHHHHHHHHHHHHHHHHH ------HHHHHHHHHH- ---EEEEEE

713
VTGAYDTIPQDRLTEVIASIIKP
---EE----HHHHHHHHHH----
IRDAYGNVKIPVLCKLIQSIPTH
E-------HHHHHHHHHH-----

803     811                                                865 869
SGLFDVFLKFMCHHAVRIRGKSYVQCQGIPQGSILSTLLCSLCYGDMENKLFAGIRRDGLLLRLVDDFLLVTPHLTH
HHHHHHHHHHH---EEE---EEEEE--------HHHHHHHHHHHHHHHHH-------EEEEEE--EEEEE--HHH
SEKKNFIVDHISNQFVAFRRKIYKWNHGLLQGDPLSGCLCELYMAFMDRLYFSNLDKDAFIHRTVDDYFFCSPHPHK
HHHHHHHHHHHHHEEEE----EEEE---------HHHHHHHHHHHHHHHHH-------EEEE---EEEEEE-HHHH

880      889        898 902
AKTFLRTL RGV        VNLRKTVVNFPVEDEAL
HHHHHHHH HHH        E---EEEEE--------
VYDFELLI KGV        VNPTKTRTNLPTHRHPQ
HHHHHHHH HHH        EEEEEEEE---------
```

Key:

```
Black numbers: human residue number range of aligned fragment
Dark green:  human sequence
Light green: predicted human secondary structure (Jpred)
Red: beetle sequence
Orange: observed beetle secondary structure
```

Figure 2. Alignment of human to beetle sequence and comparison of predicted human to observed beetle secondary structure. Note wide agreement between the two secondary structures, in particular in the RBD. Human residues of special interest include the active site (712,868,869, cyan highlight – note conservation), positively charged disease associated residues (570, 811, 865, 901, 902, green highlight – note only 865 and 902 are charged in beetle), and highest OPRA propensity from beetle (563, 576, yellow highlight).

As mentioned the beetle telomerase has no TEN. We found it difficult to dock the *Tetrahymena* TEN to the beetle TERT. The beetle TERT has a very crowded RT active site, with little room for an additional protein to be involved. We found no clear shape complementarity with the *T.Tetrahymena* TEN; significant structural rearrangement would be needed just to get the crucial TEN residues 170-175 near the 3' end of the primer. Our model for this reason does not include TEN.

2.4 *Predicting RNA binding interfaces on beetle TERT*

The correct position of the RNA component of Telomerase is unknown except for the location of the primer-template duplex (11). We therefore used OPRA (Optimal Protein RNA Area)(20) to predict points of high RNA-binding propensity on beetle TERT.

2.5 *Generating the primer elongation trajectory*

We used the available biochemical knowledge to generate a dynamical trajectory of elongation as follows.

In a preparatory stage, we threaded (16) the human Telomerase RNA (TR) template portion spanning residues 53 to 59 (18) onto the putative RNA component of the beetle primer-template duplex. (11) We attached a spring to pull the 5' end of the modeled template flanking region (residue 38) so as to pass near the predicted RNA binding hotspot in the T-domain. This was motivated by biochemical evidence suggesting the double-helical template boundary element (which includes residues 32-37) binds to the RBD. The T-domain is therefore one candidate for the location of the boundary element. We threaded seven residues of the modeled primer onto the DNA portion of the beetle duplex.

To model the addition of a single residue, we then shifted the primer by one residue in the 5' direction and attached a new residue at the 3' end. We shifted the template by one residue position in its 3' direction. We released the most distal base pair of the duplex. We repeated this process for a total of seven residues. Once sufficient primer residues exited the duplex, we enforced a stem-loop.

As a final step, following the addition of a nucleotide conjugate to the 5'-most template RNA residue, we released the template and reattached it 6 residue positions in the 5' direction, ready for another round of extension.

2.5 *Functional assay*

The plasmids pcDNA6hTERT and pBSU1hTR were used for transient transfections of HEK cells and were a kind gift from J. Lingner. The hTERT mutations K570N, R865C and K902N were generated in the pcDNA6 vector as well. Cell extracts and direct telomerase assays were done as described in (21) and (22), respectively.

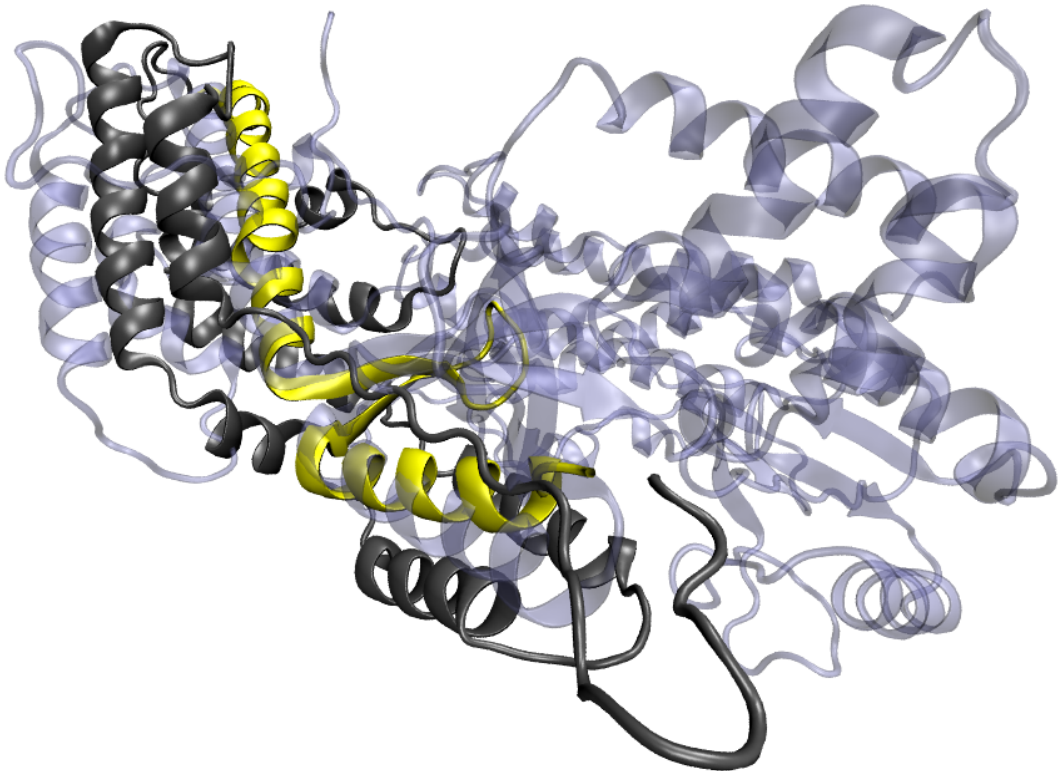
3. Results

3.1. *Validating human-beetle sequence alignment*

We compared to the predicted human secondary structure(14) to the observed beetle secondary structure(11) and found that they agreed in 232 of 283 aligned residues (Figure 2).

3.2. *Aligning Tetrahymena RBD to beetle*

We found that the *T.thermophila* RBD T-domain was aligned incorrectly in the telomerase database(4) as evidenced by a lack of secondary structure correspondence and the failure of the sequence alignment to guide correct 3D alignment between beetle and *tetrahymena* TERT. Once we corrected this the secondary structure aligned properly and we were able to align structurally (Figure 3)



```
CVPAAEHRLREEILAKF-LHWLMSVYVVELLRSFFYVTETTFQKNRLFFYRKSVWSKLQSIGIRQLKRVQLREL
---HHHHHHHHHHHHH HHHHH--HHHHHH---EEEEEE----EEEEEEEHHHHHHHHHHHHHHHHH------
IPWLQNVEPNLRPKLLLKHNLFLLDNIVKPIIAFYYKPIKTLNGHEIKFIRKEEYISFESKVFHKKKMKYLVEVQDEVK
HHHHH-----HHHHHHHHHHHHHHHHHHHHHHHHEEEEE------EEEEEHHHHHHHHHHHHHHHH--EEE------
TQKRKYYISDKRKILGDLIVFIINKIVIPVLRYNFYITEKHKEGSQIFYYRKPIWKLVSKLTIVKLEE
-HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHEEEEE-------EEEEEHHHHHHHHHHHHHHHH
```

Dark green: human sequence
Light green: human predicted secondary structure (Jpred)
Red: Beetle sequence
Orange: Beetle observed secondary structure
Yellow highlight: Tetrahymena sequence
Gray: Tetrahymena observed secondary structure

Figure 3. Alignments of the T-motif

Upper panel: Structural alignment of Tetrahymena T-motif (yellow) and remainder of Tetrahymena TRBD (gray). Beetle TERT is shown in transparent gunmetal. Rigid structural alignment is based on the known beetle-Tetrahymena sequence alignment. The T-motif is a unique fold, and the close structural alignment is evidence that the beetle structure is TERT. Therefore, we should be able to predict human structure based on alignment to beetle, in regions of high sequence identity.

Lower panel: Sequence alignment of the RNA Binding Domain T-motif. Note that previously published alignments are incorrect for Tetrahymena in this region.

3.3. *RNA binding interfaces on beetle TERT*

OPRA predicted a strong binding propensity in the RBD, in particular at the base of the β-hairpin. It also predicted a second binding region in the C-terminal domain (Figure 4) recapitulating its known DNA binding function.(3)

3.4. *Basic disease associated residues near the primer-template duplex*

A novel feature of MMB called "Physics where you want it" allows flexibilizing a selection of residues, leaving the remainder of the system fixed and rigid. An MD (in this case Amber99) force field is turned on for another (typically larger and enclosing) selection of residues. We flexibilized residues 570, 902, and 865 and turned on physics for those residues and for the primer-template duplex. We found that residues 570 and 902 rapidly gravitated towards the duplex, settling within 2Å of the DNA strand at the closest point. 901 was in a similar position to 902. Residue 865 came within 6Å and 8Å, respectively, of the RNA and DNA strands (Figure 4).

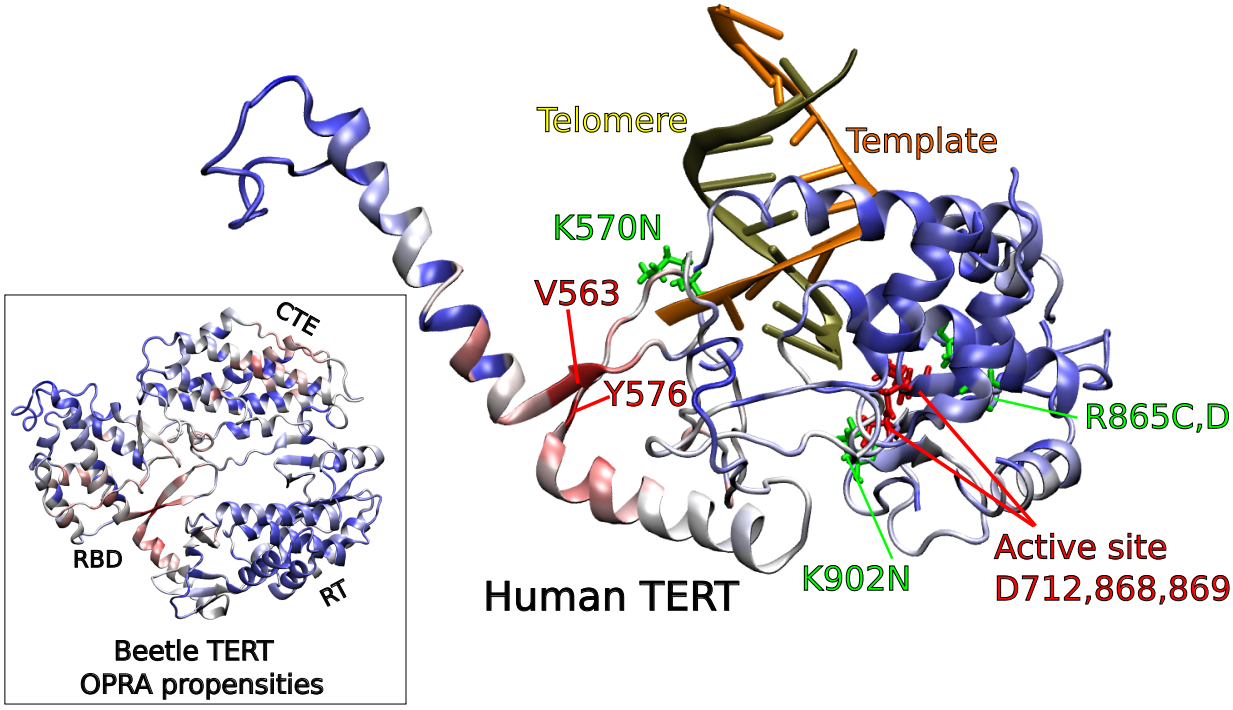


Figure 4. Human TERT charged disease-associated residues and OPRA RNA binding propensities in the vicinity of the duplex.

Labeled green residues are disease associated. Residue 570 lies very close to the primer backbone and is very likely to coordinate it. Residues 865 and 902 can come within a few Ångströms of the primer strand, and are also close to the active site residues (conserved ASP 712,868, and 869).

OPRA propensities were calculated in beetle TERT (inset) and then transferred to human. Highest propensities (red shading) were in the C-terminal extension (or Thumb) and RBD (see human residues 563 and 576). Lowest propensities are in blue shading.

3.5. *Basic beetle residues near the primer-template duplex*

For the beetle structure (PDB accession: 3KYL) we found that of 20 protein residues within 5Å of the 7 DNA residues in the 7-bp duplex, seven (144,194,406,416,418,437,477) were basic. Further, they were mostly in range to

make contact. On the other hand, we found that 30 protein residues were within 5Å of the 7 RNA residues in the primer-template duplex. Of the 30, only two (206,210) were within the 5Å range, but at its very edge and apparently interacting with a portion of the DNA strand outside of the 7-bp duplex (Figure 5).

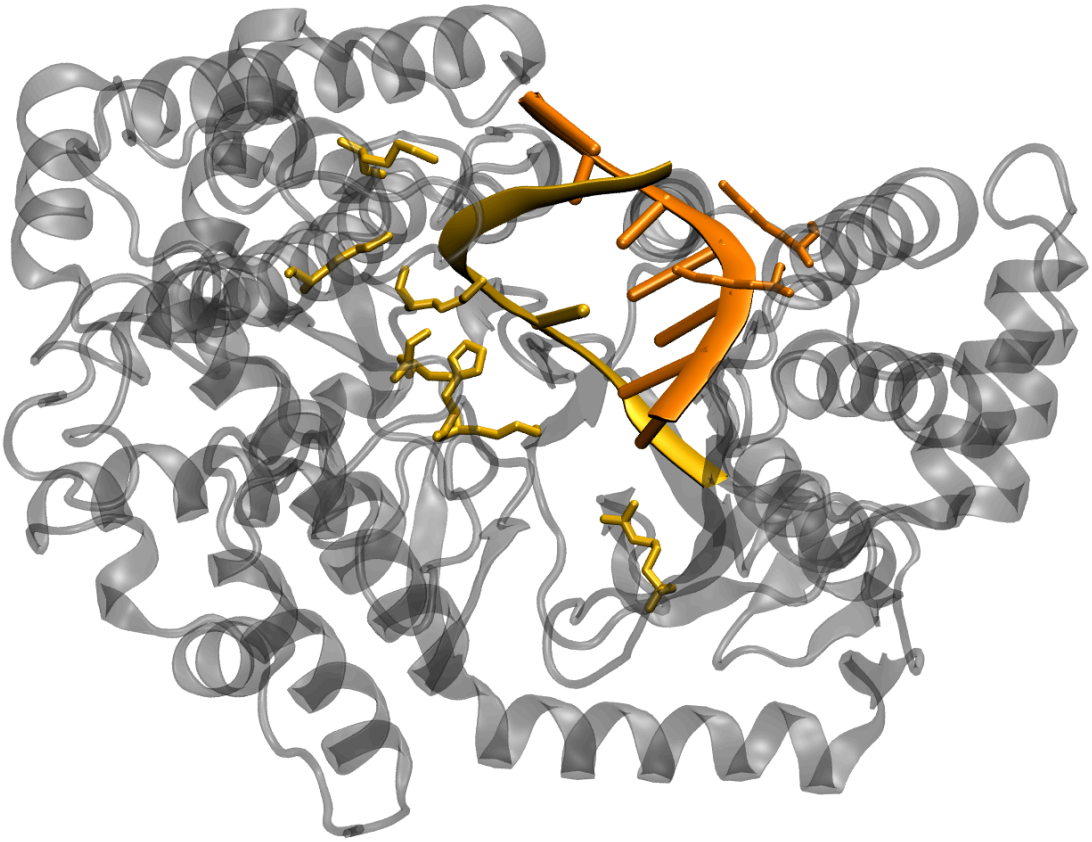


Figure 5 Basic residues within 5Å of the template-primer duplex in T. Castaneum.

Seven basic amino acid residues (gold) coordinate the seven primer residues (also gold) of the duplex. Two additional basic residues (orange) are within 5Å of the template portion of the duplex (also orange), but appear to be coordinating a portion of the primer outside the duplex. Thus all residues within 5Å of the duplex In beetle are binding the primer, not the template.

3.5. *Functional assay*

The hTERT variants carrying mutations K570N, R865C, K902N had been identified in patients with DC, AA or IPF (23-25). Measuring the activity of these mutant telomerase complexes revealed that the K570N and K902N variants showed no telomerase activity, while R865C showed a 20% telomerase activity (25). We now attempted to rescue this strong phenotype by increasing the telomeric primer concentration 10-fold in the extension assay; however this did not significantly improve the activity of mutant telomerases (data not shown).

3.6. *The dynamical trajectory of primer elongation*

We noted several possible routes for the 5' template flanking region to exit the primer-template duplex. The route used in our trajectory follows the β-hairpin, contacting the RNA binding hotspot at its base, and takes a direction tangent to its duplex portion (supplementary materials).

There is sufficient space to form a harpin in bursts as sufficient residues emerge from the primer-template duplex. We note that one or more hairpins may form, each with 8 base pairs in the stem. Alternatively, a single long hairpin may form, with its end loop migrating to 3' in six-residue increments as the primer elongates, or some combination of the two may occur.

4. Discussion

4.1. *Validating human-beetle sequence alignment*

Sequence identity alone does not necessarily imply structural homology; structural divergence is possible. In particular controversy surrounds the beetle structure used, with some authors pointing out that it may not be TERT(5). We present evidence that the beetle-human sequence alignment is correct, that the beetle structure is in fact TERT, and that there is a basis for homology modeling.

First, the beetle – *T.thermophila* sequence alignment was used as the basis for aligning the less controversial *T.thermophila* RBD to its putative beetle counterpart. The key conserved T-motif, a β-hairpin with flanking α-helices aligns well (Figure 4). Since this motif has not been observed outside of TERT to our knowledge, this observation indicates that the beetle structure is likely to be TERT, and that the sequence alignment is correct in this region.

Second, we ran the Jpred secondary structure prediction algorithm on the human sequence and found that the predicted human secondary structure matches the observed beetle secondary structure for 232 of 283 residues (Figure 2). Jpred(14) is not biased by the use of the beetle or *T.thermophila* TERT structures (C. Cole, personal communication). Therefore there is probably considerable structural homology in the aligned regions. The human sequence and predicted secondary structure also aligns to that of *T.thermophila* in the RBD (Figure 3), further bolstering the structural and functional correspondence (non-telomerase reverse transcriptases do not have this RBD).

As mentioned the TEN domain has been directly implicated in the active site (9) and yet is absent in beetle. We speculate that the role of the human TEN residues involved in elongation may be played in beetle by residues in the RT domain, leading to a crowded active site with little room for an additional protein subunit. This may explain why it is difficult to dock the *Tetrahymena* TEN to the beetle TERT.

4.2. *RNA binding interfaces on beetle TERT*

The results of the OPRA RNA binding interface prediction strengthen the case that the beetle structure is TERT. The binding hotspot in the CTE is consistent with its role in DNA binding.(3) The strong signal at the base of the β-hairpin confirms that this is the RNA Binding Domain.

4.3. *Basic beetle residues near the primer-template duplex*

The evidence of DNA binding being more important than RNA binding near the duplex prompted us to examine the beetle structure (3KIY) more closely. The finding that 7 of the 20 residues near the DNA part of the 7-bp duplex were basic, while only 2 of 30 the residues near the RNA part were basic, is strong indication that DNA binding is more crucial in this region (Figure 4). Experimentalists may find it profitable to mutate beetle TERT residues 144,194,406,416,418,437, and/or 477 to try to produce reduced DNA binding in vitro or, more interestingly, a beetle with a telomerase deficiency disease.

4.4. *Charged disease associated residues near the primer-template duplex*

Disease associated residue 570 is very likely in contact with the primer strand, a few residues from the terminus. Its charge, location in a highly conserved RNA binding region, and clear proximity to the DNA backbone make this prediction strong.

Residue 902 (and by extension 901) is also positioned to contact DNA. We verified this with a simple equilibration using MMB.

Residue 865 is spatially near the primer strand. It is also close in sequence to the active site residues, further suggesting a DNA binding role. However an equilibration using MMB, shows that 865 could contact either DNA or RNA without significant backbone motion.

4.5. *Insight into disease from structure*

Our rather conservative model leaves out diverged regions of TERT such as the CTE, and for that reason contains only five basic disease associated residues: 570, 811, 865, 901, and 902. Of these, all except 811 are within 9Å of DNA. A very limited "Physics where you want it" simulation showed that 865 is positioned such that it could also be contacting RNA, but that 570, 901, and 902 are very likely binding DNA only. 811 is about 19Å from the duplex. Our model therefore predicts a likely DNA-binding role for disease-associated residues 570, 901, 902, and perhaps 865. All basic residues within 9Å of the DNA in our model (with the exception of some at the periphery of that cutoff) are documented disease-associated residues, again suggesting that basic residues near the duplex have an essential DNA-binding function. If this model is correct, TERT with these mutations should have reduced affinity for DNA. A functional assay should show reduced activity, which should be recovered with a saturating concentration of DNA.

Inspection of the Steczkiewicz model also leads to interesting findings. Due to differences in the modeling approach the findings are overlapping but not identical. In particular, since their model included highly diverged regions, it includes many residues which ours did not. First, of the seven basic residues known to be disease-associated (486, 570, 811, 865, 901, 902, and 979), three (570, 865, and 979) are within 9Å of the DNA strand. Two of the latter (570 and 979) were also within 9Å of the RNA strand. Within 9Å of the DNA, there were 22 positively charged residues in total. In contrast, there were only 13 within the same radius of the RNA, again suggesting the greater importance of binding the stretch of DNA.

One interpretation of our computational model is that basic residues which coordinate the DNA strand are more likely to be essential, while those coordinating the RNA may be dispensable. This could be explained by the fact that while TERT has many points of contact with RNA, it has very few with DNA. In accordance with this idea, other basic residues within about 9Å of the DNA strand, when mutated, could be as-yet undiscovered causes of telomerase diseases. However in that case saturating concentration of DNA would be expected to recover function in mutants K507N, R965C and K902N, and it did not. This may mean that the residues contribute to the stability of TERT, and that a loss of charge leads to misfolding. Alternatively, it may be that the contribution of the residues to DNA binding is very strong, and we did not reach saturating conditions of DNA.

The Steczkiewicz model by our interpretation predicts that residues 499, 500, 570, 626, 631, 643, 646, 647, 649, 650, 710, 865, 955, 962, 968, 971, 972, 973, 979, 981, 983, and 1011 are within 9Å of the DNA part of the 7 bp duplex; accordingly clinicians should be on the lookout for mutations in these residues. Note that these include known disease-associated residues 570, 865, and 979. For costlier experimental assays, we suggest starting with residues within 5Å: 643, 649, 962, 972, 973, 970, 979, and 981). Note that the latter list includes known disease-associated residue 979. However we remind experimentalists that the Steczkiewicz model aligns highly diverged domains.

4.6. *Primer elongation movie*

We created a movie and figure (supplementary material) which shows a single cycle of primer elongation. One residue is added at a time, with the rest of the primer-template duplex shifting one residue position towards the distal end with each addition. Simultaneously, the distal base pair denatures.(6) Once sufficient residues emerge from the duplex, a hairpin forms.(8) In our model there is room for a series of such hairpins, or for a single long hairpin with and end loop that migrates to 3' in six-residue bursts, or for some combination of the two. We look forward to future workers elucidating how this is denatured to extend the lagging strand of the primer. At the end of one cycle of elongation, the template shifts six residues in the proximal direction, like an old-style typewriter carriage preparing to write another line.(6,26) The trajectory shows that this mechanism is sterically and geometrically

feasible and consistent with existing structural and biochemical data. It further provides a structural basis for designing focused experiments to test specific steps of this process. We encourage other workers to modify the MMB/RNABuilder command file we provide to extend or modify our simulation or add more Telomerase subunits.

5. Conclusion

Considerable progress has been made on Telomerase structure and function, but this had not been turned into a 3D dynamical model. In this work we first presented the evidence that a recent *T.Castaneum* structure (3) is in fact TERT, addressing a topic of current debate. (11) We then built a threaded model of part of the RT and RBD domains and the primer-template duplex of human telomerase. We find that all four basic residues within 9Å of the primer-template duplex are disease associated, and further that at least three of them appear to be DNA binding. Similarly, we find that in the published beetle structure, basic residues cluster near the DNA and not the RNA strand of the duplex. This may suggest that several positively-charged, disease-associated residues are involved in coordinating DNA in human telomerase. We propose that such DNA-binding residues are more likely to be essential, whereas RNA-binding residues may be dispensable, since TERT has fewer points of contact with DNA than RNA. Our functional assay does not rule out the possibility that these residues also contribute to the stability of TERT. We used biochemical and biophysical results as constraints to generate a 3D kinematic model of primer extension as an illustration of this important process.

6. Distribution and supplementary materials

The telomere extension movie is distributed as a .mpg and a figure file at https://simtk.org/home/telomerase . The entire trajectory in .pdb format, the command file (in RNABuilder 2.2 syntax), and initial structure file are also available on request. RNABuilder 2.2 and more recent MMB distributions for Windows, OSX, and Linux are available for download from https://simtk.org/home/rnatoolbox .

7. Acknowledgements

We thank Juan Fernández-Recio and Laura Perez-Cano for running OPRA on the beetle TERT structure for us. S. Flores acknowledges funding from eSSENCE (essenceofscience.se). Joachim Lingner is acknowledged for his generous gift of plasmids. This work was supported by the Austrian Science Foundation (FWF) grant Y401 to C. Waldsich.

8. References

1. Sekaran, V.G., Soares, J. and Jarstfer, M.B. (2010) Structures of telomerase subunits provide functional insights. *Biochim Biophys Acta*, 1804, 1190-1201.
2. Kim, N.W., Piatyszek, M.A., Prowse, K.R., Harley, C.B., West, M.D., Ho, P.L., Coviello, G.M., Wright, W.E., Weinrich, S.L. and Shay, J.W. (1994) Specific association of human telomerase activity with immortal cells and cancer. *Science (New York, N.Y*, 266, 2011-2015.
3. Gillis, A.J., Schuller, A.P. and Skordalakes, E. (2008) Structure of the Tribolium castaneum telomerase catalytic subunit TERT. *Nature*, 455, 633-637.
4. Podlevsky, J.D., Bley, C.J., Omana, R.V., Qi, X. and Chen, J.J. (2008) The telomerase database. *Nucleic acids research*, 36, D339-343.
5. Wyatt, H.D., West, S.C. and Beattie, T.L. (2010) InTERTpreting telomerase structure and function. *Nucleic acids research*, 38, 5609-5622.
6. Forstemann, K. and Lingner, J. (2005) Telomerase limits the extent of base pairing between template RNA and telomeric DNA. *EMBO Rep*, 6, 361-366.
7. Hammond, P.W. and Cech, T.R. (1998) Euplotes telomerase: evidence for limited base-pairing during primer elongation and dGTP as an effector of translocation. *Biochemistry*, 37, 5162-5172.
8. Lue, N.F. (2004) Adding to the ends: what makes telomerase processive and how important is it? *Bioessays*, 26, 955-962.

9. Jurczyluk, J., Nouwens, A.S., Holien, J.K., Adams, T.E., Lovrecz, G.O., Parker, M.W., Cohen, S.B. and Bryan, T.M. (2011) Direct involvement of the TEN domain at the active site of human telomerase. *Nucleic acids research*, 39, 1774-1788.
10. Steczkiewicz, K., Zimmermann, M.T., Kurcinski, M., Lewis, B.A., Dobbs, D., Kloczkowski, A., Jernigan, R.L., Kolinski, A. and Ginalski, K. (2011) Human telomerase model shows the role of the TEN domain in advancing the double helix for the next polymerization step. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 9443-9448.
11. Mitchell, M., Gillis, A., Futahashi, M., Fujiwara, H. and Skordalakes, E. (2010) Structural basis for telomerase catalytic subunit TERT binding to RNA template and telomeric DNA. *Nat Struct Mol Biol*, 17, 513-518.
12. Flores, S. and Altman, R. (2010) Turning limited experimental information into 3D models of RNA *RNA (New York, N.Y*, 16, 1769-1778.
13. Rouda, S. and Skordalakes, E. (2007) Structure of the RNA-binding domain of telomerase: implications for RNA recognition and binding. *Structure*, 15, 1403-1412.
14. Cole, C., Barber, J.D. and Barton, G.J. (2008) The Jpred 3 secondary structure prediction server. *Nucleic acids research*, 36, W197-201.
15. Flores, S.C. (2011) RNABuilder 2.2 Tutorial.
16. Flores, S., Wan, Y., Russell, R. and Altman, R. (2010) Predicting RNA structure by multiple template homology modeling. *Proceedings of the Pacific Symposium on Biocomputing*, 216-227.
17. Wallner, B. and Elofsson, A. (2005) All are not equal: a benchmark of different homology modeling programs. *Protein Sci*, 14, 1315-1327.
18. Flores, S., Sherman, M., Bruns, C., Eastman, P. and Altman, R. (2010) Fast flexible modeling of macromolecular structure using internal coordinates. *IEEE Transactions in Computational Biology and Bioinformatics*, 8, 1247-1257.
19. Xiang, Z. (2006) Advances in homology protein structure modeling. *Curr Protein Pept Sci*, 7, 217-227.
20. Perez-Cano, L. and Fernandez-Recio, J. (2010) Optimal protein-RNA area, OPRA: a propensity-based method to identify RNA-binding sites on proteins. *Proteins*, 78, 25-35.
21. Cristofari, G., Adolf, E., Reichenbach, P., Sikora, K., Terns, R.M., Terns, M.P. and Lingner, J. (2007) Human telomerase RNA accumulation in Cajal bodies facilitates telomerase recruitment to telomeres and telomere elongation. *Molecular cell*, 27, 882-889.
22. Cristofari, G. and Lingner, J. (2006) Telomere length homeostasis requires that telomerase levels are limiting. *The EMBO journal*, 25, 565-574.
23. Armanios, M., Chen, J.L., Chang, Y.P., Brodsky, R.A., Hawkins, A., Griffin, C.A., Eshleman, J.R., Cohen, A.R., Chakravarti, A., Hamosh, A. *et al.* (2005) Haploinsufficiency of telomerase reverse transcriptase leads to anticipation in autosomal dominant dyskeratosis congenita. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 15960-15964.
24. Xin, Z.T., Beauchamp, A.D., Calado, R.T., Bradford, J.W., Regal, J.A., Shenoy, A., Liang, Y., Lansdorp, P.M., Young, N.S. and Ly, H. (2007) Functional characterization of natural telomerase mutations found in patients with hematologic disorders. *Blood*, 109, 524-532.
25. Tsakiri, K.D., Cronkhite, J.T., Kuan, P.J., Xing, C., Raghu, G., Weissler, J.C., Rosenblatt, R.L., Shay, J.W. and Garcia, C.K. (2007) Adult-onset pulmonary fibrosis caused by mutations in telomerase. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 7552-7557.
26. Zaug, A.J., Podell, E.R. and Cech, T.R. (2008) Mutation in TERT separates processivity from anchor-site function. *Nat Struct Mol Biol*, 15, 870-872.