

Reply to a review of

Modelling of DNA Mismatch Repair with reversible process calculus

by Stefan Kuhn and Irek Ulidowski submitted for publication in the Journal of Theoretical Computer Science, 2021

The authors would like to thank the reviewer for their comments and suggestions. We have taken them all into account while improving the paper.

The text of the review is included for completeness and is displayed using normal font. Our responses and comments are displayed here in italic font.

In this paper the authors show that an abstraction of an important biological mechanism, i.e., the DNA Mismatch Repair, may be modelled in a reversible process calculus. The abstraction here ignores the complicated geometrical structure of the DNA and also the mechanisms required for making local part of interest of the DNA available the reactions (e.g., gene expressions).

Nevertheless, it is an good step forward in linking process calculi to biological processes. To this aim the authors reshape a former process calculus CCB (the Calculus of Covalent Bonding), previously introduced and use by the authors for a related correction mechanism - Base Excision Repair. The extension requires allowing prefixing with collections of bonding sits and a more careful operational semantics.

The paper is interesting, correct, and of interest for this journal. I is also well written - an extra value of the paper is the good survey of the process algebra based approach to model biological processes. I strongly recommend its publication in this journal. A few more comments and suggestions for improving the presentation are inserted below.

General comments:

1. The process terms describing the systems are long and somehow difficult to follows. For instance, the given example, involving only 6 pairs of bases, results in a half a page formula (page 22). While there is some effort to simplify the term, the notation is still long and much of the text in pages 23-29 contains redundant information. A suggestion here is the use something similar to the global-local handling notation used in the K-framework (see, e.g., <https://kframework.org/>).

We have studied kframework and thought how to use it to improve display of our reactions. However, we have decided that it would be too disruptive and it would not shorten the paper considerably as some new background material would need to be added to explain kframework and its usage. We only list those processes on pages 24 to 31 that are at some point involved in the reactions. So, having their syntax repeated several times to some extent repetitive but the benefit is that it shows the full context of the reactions.

Instead of using kframework, we have used different colours and fonts, both in the syntax of processes (that represent molecules) in our transitions and in the supporting figures that show how the bonds change over the lifetime of the reactions. See a fuller description of the use of colours, fonts and line styles below.

2. Quite often, graphs are useful in grasping intuition on process terms. Maybe developing a graphical representation of the CCB processes may simplify understanding and pave a way towards including geometrical information on DNA shape.

We have improved graphical representation of the main transitions and diagrammatic representation of transitions and the resulting structure of the DNA stand on pages 24-31. We have a new notation for the bonds. In our updated figures we have three types of lines:

- a) Normal black lines (unchanged bonds)*
- b) Blue lines (recently created bonds)*
- c) Red dotted lines (recently broken bonds)*

To help readability further we include keys of the blue and red bonds in Figures 10 to 14.

We have also used colours in the syntax of processes on pages 24 to 29. We use red to indicate bonds that will break next, namely the pairs of synchronised actions and their keys. Also, we highlight in blue the pairs of actions that will synchronise next thus producing new bonds. Finally, after each transition (reaction) we use bold font to indicate the actions that took part in bond creation or bond breaking.

Detailed comments (page, global line numbers):

p2,25: DNA errors often occur when DNA is used for reading the information about the proteins to be synthesized in the cells.

There are indeed many different sources of DNA damage, including damage during transcription, i. e. production of mRNA from DNA. We have decided here to focus on DNA replication errors. There are certainly more potential issues to discuss, including DNA repair as part of transcription (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1470299/>). We believe that for this paper focusing on one case is justified.

A clarification has been added on page 3.

p5,110: bee

Corrected.

p6,130: Maybe a short clarification of the distinction strong-weak actions is useful at this point.

A clarification has been added which, together with a slight adjustment of text on pages 7-8, makes matters clearer.

p6,140: : "use α , s and s, s' to denote a concatenation" - the notation is unclear; include some parentheses (), or " " to separate the terms from the text

We have use quotation marks on line 147 on page 7.

P7, 150: "A collection μ is either σ or $\sigma \mid \sigma$ " - if the is a BN-like syntax, should be $\sigma \mid \mu$

This has now been changed.

p7,155: The syntax includes also recursive definition, not included in that one-line description

No, it is not meant that way. We have changed the explanation.

p8,185: "(s; b).C | P | C | C | P | C \ L" - confusion of two roles of "|"

The display of this BNF has been improved.

p10, Fig3: b does not appear in this rule

b is in the s rule.

p11, Fig4: similar

We have added a clarification.

p12, Fig5: similar somehow, it is unclear why do you use condition $*$ and $**$; why not simply include those particular cases explicitly?

We have simplified presentation of the conditions in Figure 5.

p12, Ex3: it seems to be a confusion here between b, d, f - check this example again

Corrected. We have uses b and d (not b and b) since b is weak and d is not.

p13, Fig6: what is "t" in concert2 act?

It has been changed to eta.

p16,275: Is this strong-weak action explanation useful at p6,130?

We have added an explanation on p 6 and pointed to page 16 for fuller explanation, which can only be made once all SOS and promotion rules have been given and explained.

p22, top: "The MMR system in full detail..." Actually, this is the result after all reactions in the previously described MMR system are used. It shows an order of firing transitions in the original system to arrive at this form, directly linked to the picture in Fig.9

We believe there are three descriptions the reviewer is considering here, two of them as processes, one graphical. Firstly, there is the process following the sentence “The system shown in Figure 9 is modelled in CCB as follows:”, secondly the process following the sentence “The MMR system in full detail (but without restriction) is given below”, and thirdly Figure 9. All of those actually describe the same system. The first process description just has the process names from the previous pages written, whereas the second process description has the full definitions, including bonds, which are in the figures as well. It does not imply there is a development from the first to the second process. We do not consider how we arrived at this state, and give no order of transitions. The first process description is intended to give an overview, the second the details, which are shown in the Figure as well.

An appropriate explanation has been added on pages 22-23.

p30, Fig14: "the offending base C_3" - is G_3

Corrected.