



HARVARD MEDICAL SCHOOL

Alan Engelman, Ph.D.

Professor of Medicine

Fang-Fang Ji

Science Editor, Editorial Office

Baishideng Publishing Group Inc.

12/23/2016

Dear Dr. Ji,

Re: Manuscript No. 31153

Please find attached the revised version of our manuscript “Retroviral Integrase Protein and Intasome Nucleoprotein Complex Structures”, which we resubmit for consideration for publication as a Review in World Journal of Biological Chemistry.

We thank you and the reviewers for your comments, which enabled us to improve the quality of our manuscript. Follows is a point-by-point discussion of the changes that were made in response to the reviewers’ comments. Please note that changes are marked in red font in the text of 31153-Revised manuscript.docx.

Reviewer #1

1. very good review, thanks for putting this structural analysis in a nice way so anyone can understand.

Response: We thank the reviewer for his/her positive feedback.

Reviewer #2

1. This review on retroviral intasome structure by Grawenhoff and Engelman is in continuation of excellent reviews articles from Engelman’s group in recent years and it particularly provides discussion recently resolved structures of MMTV SSC and RSV STC structures. The citation of several original research articles (not just the recent ones and only the review articles) in this review article is highly appreciated. I have included few suggestions to improve the technical language and couple missing relevant information. Even though several articles (and this one too) have used the term : integration is “catalyzed” by integrase (in Abstract, first paragraph of introduction, title of second heading on page 4, Fig 1 legend and other places too). The correct terminology is integration of viral DNA in host DNA is “mediated” by integrase (also including the 3-processing step). Technically, catalysis means



DANA-FARBER CANCER INSTITUTE

Department of Cancer Immunology & Virology

450 Brookline Avenue, CLS 1010 Boston MA 02215

Tel: 617 632 4361 Fax: 617 632 4338

alan_engelman@dfci.harvard.edu

enhancing the reaction rate. However, in this case integrase is essential for above two reactions. Similar language change is warranted for reverse-transcriptase (RT). Reverse transcription is mediated (not catalyzed) by RT (first paragraph in Introduction).

Response: We thank reviewer #2 for his/her positive feedback, and appreciate the insightful comments.

According to our understanding, reverse transcriptase (RT) and integrase (IN) are enzymes, and hence by the very nature of the definition, are catalysts. The reactions that they mediate would not take place, or at best would occur at infinitesimally slow rates, in their absence. Formally, the catalytic turnover per IN dimer is estimated to be 0.52/min (ref. 50 herein). Thus, IN is not consumed during 3'-processing or strand transfer, which confirms it is a catalyst. We understand that technically neither RT nor IN "catalyze" reverse transcription or integration, respectively, as these are processes, and not reactions. We accordingly amended the manuscript in the following places, as recommended by the reviewer: abstract line 2, "catalyzed" was replaced by "mediated"; Introduction line 7, "RT catalyzes" was rephrased "RT mediates"; Introduction line 10, "Is catalyzed by IN" was rewritten "is promoted by IN". However, for the reasons discussed above, we have kept the page 5 section heading "Reactions Catalyzed by IN". For clarity, the Figure 1 legend heading "Reactions catalyzed by IN: 3' processing and strand transfer" was reworded: "IN catalytic functions and intasome complexes."

2. Page-12, couple sentences on the top of the page. Authors mentioned that HIV-1 IN require vDNA of several hundred bp to perform concerted integration while the PFV is able to carry out concerted integration with short DNAs (15-30 bp). This statement would have been correct couple of years back; however recent publications have reported efficient concerted integration by HIV-1 IN using similar sized short DNAs (18-40 bp). It was first reported by Pandey/Grandgenett group et al. in Biochemistry, 2011 (PMID 21992419) with wild type full-length HIV IN and later by Li/Craigie group et al., PLoSOne 2014 (PMID 25119883) using a modified HIV-1 IN (Sso7d fusion).

Response: We very much appreciate the Reviewer's comment. Accordingly, this part of the paper was reworked. Specifically, "Although the insertion of only one vDNA end, termed half-site integration, is commonly observed in in vitro reactions with HIV-1 IN^[77], PFV IN efficiently catalyzed the concerted integration of two vDNA ends into tDNA^[120]. Whereas HIV-1 IN required vDNA substrates of several hundred bp to perform concerted integration^[122-127], PFV IN catalyzed efficient concerted integration of vDNA ends as short as 15-30 bp^[120]." was expanded (page 12 line 23 to page 13 line 11): "Though early work had revealed that relatively short oligonucleotide substrates, which modeled the vDNA ends, supported IN 3' processing and strand transfer activities^[72-77], not all enzymes behaved similarly. Most critical for intasome structural biology was the ability for the IN multimer to coordinate the binding of two vDNA ends, and insert these in concerted fashion into opposing strands of tDNA. Critically, PFV IN was discovered to promote efficient concerted integration of oligonucleotide vDNA ends^[121]. By contrast, HIV-1 IN had revealed the tendency to insert just one vDNA end at a time^[77]. Subsequent modifications of HIV-1 IN expression systems, including protein purification under

relatively dilute conditions to prevent IN aggregation^[123], or by fusing the small Sso7d DNA binding domain from *Sulfolobus solfataricus* to the IN N-terminus to mimic the NED that naturally exists in PFV IN^[124], yielded proteins that supported efficient concerted integration activity. Such modifications might eventually prove useful to characterize HIV-1 intasomes structurally^[123, 124].” As suggested by the reviewer, please note the inclusion of new references 123 and 124 on pages 28, 29. The remainder of the original page 12 paragraph is now a subsequent paragraph on page 13 (lines 12-22).

3. The small section on IN-LEDGF co-crystal structure (page 10) seem to be out of place in this review which mainly focuses on Intasome structures (IN-DNA nucleoprotein complexes). It might be advisable to take this section out to make it coherent reading.

Response: We respectfully disagree with the reviewer’s comment, for the following reasons. First, in addition to discussing intasome structures, we provide a historical overview of preceding IN domain/ two-domain fragment crystallography, which necessitated the discussion of among other changes the critical solubility enhancing F185K mutation in HIV-1 IN (page 9, lines 3-22). As a solubility-enhancing cofactor, LEDGF/p75 enabled the structures of wild-type NTD-CCD 2-domain HIV-2 and MVV IN structures to be solved in the absence of solubility enhancing mutations (page 11, lines 16-19): “Though HIV-2 and MVV INs harbor hydrophobic residues at the positions analogous to Phe185 in HIV-1 IN (Phe and Ile, respectively), the favorable solubility properties of lentiviral IN-LEDGF/p75 complexes likely dispelled the need for solubility-enhancing mutations for the crystallization of these constructs.” Secondly, because LEDGF/p75-IN binding underlies the propensity for lentiviruses to integrate into active genes (page 11, lines 6 and 7), it remains a critical IN-binding structural biology substrate. As lentiviral intasome structures are solved moving forward, we can accordingly predict that LEDGF/p75 will be a critical component. Omitting the discussion of prior IN-LEDGF structures we feel would seriously disservice the review, as it would a priori omit the citation of the paper from these predictable breakthrough structural biology papers. Third, and most importantly, the LEDGF-MVV IN NTD-CCD structure provided the initial glimpses of the NTD-CCD linker region, which in the context of the intasome either extends to situate the two domains from separate functional molecules in *trans*, or contracts in the remaining IN molecules that do not contribute catalytic active sites (page 15, lines 20-24): “The NTD-CCD linker that is extended in IN1 and IN3 to contact the opposing CCD *in trans* is contracted in the other IN monomers^[26]. This observation highlights the necessity for NTD-CCD linker flexibility: though principally contracted, it must also possess the ability to extend when situated at the IN1 and IN3 positions to support IN catalytic function.” To clarify the critical contribution of this LEDGF-IN structure to our understanding of the functional intasome architecture, page 10 lines 17-24 [“The resulting X-ray structure revealed interactions between two of the NTDs with two CCDs of opposing NTD-CCD molecules, which seemed at the time to potentially represent a physiologically relevant interaction due to the fact that it was known from prior biochemical studies that the NTD functioned *in trans* with the CCD^[51-53]. Indeed, the domain sharing arrangement revealed in this structure was later confirmed through mutagenesis^[104] and the elucidation of intasome structures (see below)] was reworded: “The resulting X-ray structure revealed possible interactions between two of the NTDs with two CCDs of opposing NTD-CCD molecules, which was of potential physiological relevance due to the fact that it was known from prior biochemical studies that the NTD functioned *in trans* with the CCD^[51-53]. However, the inability to trace the

NTD-CCD linker regions limited the confidence of this interpretation. Importantly, the domain sharing arrangement suggested by this structure was later confirmed by additional NTD-CCD structures and mutagenesis^[104], and ultimately through the elucidation of intasome structures (see below).” We moreover added: “In hindsight, it is not surprising that the linker regions in the original HIV-1 IN NTD-CCD structure, which lacked LEDGF/p75 or DNA binding partners, were untraceable^[103].” to the bottom of page 15.

4. Page 11-first paragraph: following sentence could be modified as- “dead HIV-1 viruses.....” dead HIV-1 particles carrying NTD reverse charge substitutions in IN regained....., (since HIV-1 itself implies virus)

Response: as suggested by the reviewer, “such that otherwise dead HIV-1 viruses carrying NTD reverse charge substitutions regained partial activity only in the presence of the complementary reverse charge LEDGF/p75 partner protein” was rewritten (page 12 lines 6-8): “such that HIV-1 particles carrying NTD reverse charge substitutions that were otherwise dead regained partial activity in the presence of the complementary reverse charge LEDGF/p75 partner protein.”

5. Fig 1 legend- SSC- should be stable synaptic complex (not contact) Page 7- second sentence “SCC”- possibly it should be SSC.

Response: we thank the reviewer for catching these typographical errors. We changed “contact” to “complex” in the figure 1 and in the main text (page 7 line 24), and corrected “SCC” to “SSC” (bottom page 7).

6. Fig 3b- it might be helpful to point out in figure legend how the multiple intasome alignment was performed (software and other technical details) to get the flexibility values of 30-40 angstrom.

Response: this part of the legend was expanded to include (page 32, lines 18-20): “was performed using Chimera. For the MMTV intasome, flanking dimers were unambiguously positioned into the intasome core of the cryo-EM map via rigid-body docking.”

Reviewer #3

1. The review for retroviral integration with structural aspect by Grawenhoff & Engelman is well-organized and written, therefore, is highly recommended to publication.

For the readers of the journal, the reviewer recommend to edit the introduction. The reviewer well-understand that the authors describe RT step in the introduction. For researchers who do not focus on retroviral integration(including HIV), it is too simple for RT description (P4 line 5). On the other hand, in this review article, description for RT step is not required.

Response: the reviewer suggests to expand the line 5 description of reverse transcription (now on page 5). Please see page 5 lines 9-11: “In the cytoplasm, RT mediates the synthesis of a linear viral DNA (vDNA) molecule that harbors a copy of the viral long-terminal repeat (LTR) at each end^[8-10]”.

2. Moreover, it is difficult to distinguish RT and preintegration complexes. Therefore, the reviewer would like to rewrite these phrases.

Response: The phrases “reverse transcription complex” and “preintegration complex” were coined in 1999 and 1987 to describe the retroviral replication intermediates, with over 5,000 and 234 Pubmed articles since retrieved using these search terms, respectively. We therefore feel it would be highly inappropriate for us to redefine such long-standing, heavily entrenched terminology in this review article.

We trust we have satisfactorily addressed the reviewers’ comments, and that the revised manuscript may now be considered acceptable for publication in the World Journal of Biological Chemistry. We look forward to hearing your decision.

Additionally, as I am a native English speaker, I have not provided a No. 8 document (Language Certificate) as part of the resubmission package.

Sincerely,



Alan Engelman, Ph.D.
Professor of Medicine
Harvard Medical School