The "complex" RNA post-transcriptional element of a "simple" retrovirus

Katarzyna J Purzycka[1], Guy R Pilkington[2], Barbara K Felber[2,*], and Stuart F J Le Grice[3,*]

[1]Institute of Bioorganic Chemistry; Polish Academy of Sciences; Poznan, Poland; [2]Human Retrovirus; Pathogenesis Section; Vaccine Branch; Frederick, MD USA; [3]RT Biochemistry Section; Basic Research Laboratory; National Cancer Institute Frederick; Frederick, MD USA

Replication of retroviruses and transposition of endogenous retroelements exploits a unique mechanism of post-transcriptional regulation as a means of exporting their incompletely-spliced mRNAs (which serve as both the genomic RNA and the template for protein synthesis). Following discovery of the Rev response element (RRE) that mediates nucleocytoplasmic export of the full-length and singly-spliced human immunodeficiency virus type 1 (HIV-1) genome, equivalent *cis*-acting regulatory elements have been characterized for both complex and simple retroviruses and retroelements, together with the obligate viral and host proteins with which they interact. The exception to this is the gammaretrovirus family of simple retroviruses, exemplified by reticuloendotheliosis virus (REV), murine leukemia virus (MLV) and xenotropic MLV-related retrovirus (XMRV). In this commentary, we discuss our recent data that reported structural and functional data on the MLV/XMRV post-transcriptional regulatory element (designated the PTE). The PTE was characterized by a highly-structured region of multiple stem-loops (SL1 – SL7) overlapping the *pro* and 5′ portion of the *pol* open reading frames, comprising a bipartite export signal whose structures are separated by ∼1400 nt. In addition, structural probing suggested that SL3 nucleotides were involved in pseudoknot formation. These data, when compared with RNA transport elements of complex retroviruses (HIV) and simple murine retrotransposons (*musD*), collectively present an emerging picture that long-range tertiary interactions are critical mediators of their biological function.

For retroviruses and retroelements, both intra- and intermolecular long-range tertiary interactions, mediated by their genomic RNAs, are crucial to interactions with regulatory proteins or other nucleic acids, and exploited to control critical events such as transcription, genome dimerization and packaging, ribosomal frameshifting and nuclear export of unspliced or partially-spliced transcripts. With respect to nuclear export, detailed structural information for the cis-acting elements of complex retroviruses (Human Immunodeficiency Virus (HIV) and Human T-cell Leukemia Virus (HTLV)), human retrotransposons (long interspersed nuclear element-1 or LINE-1), murine retroelements (intracisternal A particles and the MusD retroelement) and simple retroviruses (Rous sarcoma virus (RSV), Mason-Pfizer monkey virus (MPMV)/simian retrovirus (SRV) type 1 and 2) has been contrasted by the lack of information on the counterparts from gammaretroviruses exemplified by reticuloendotheliosis virus (REV) murine leukemia virus (MLV) and xenotropic MLV-related retrovirus (XMRV). The recent study of Pilkington et al. [1] has now filled this gap, reporting the location, biological function, secondary structure and potential tertiary interactions mediated by the MLV/XMRV post-transcriptional element, or PTE. Preliminary analysis involving deletion mutagenesis of both the MLV and XMRV genomes led to the unexpected observation that, in contrast to other retroviruses and retroelements, the PTE was not located in the untranslated region of the genome between the envelope (*env*) open reading frame and the 3′ long terminal repeat (LTR), but

rather overlaps sequences encoding its protease (PR) and a portion of the reverse transcriptase (RT) coding region (**Fig. 1**). Equally surprising was the length of the PTE (~1400 nt) which is significantly larger in comparison to RNA transport elements of complex and simple retroviruses, most of which are ~500 nt. Despite these unexpected findings, support for the biological role of the PTE was provided by the ability of this 1400 nt RNA to confer functionality to heterologous and poorly expressed genes that are dependent on post-transcriptional regulation.

Chemical probing of the XMRV PTE by selective 2′ hydroxyl acylation monitored by primer extension (SHAPE) indicated a complex collection of stem-loops (designated SL-1 SL7) and a potential pseudoknot specified by SL3 nucleotides, which collectively suggested a picture of a significantly more complex RNA export element than might have been envisaged for a simple retrovirus. Since the PTE-specifying regions of MLV and XMRV differed in only 32 of 1500 nucleotides (with many in single-stranded regions), this structure is clearly shared by both viruses. By performing a series of internal deletions that systematically removed structural elements, we concluded that PTE function was dependent upon the presence of SL1 and SL7. Separation of the minimal PTE components by ~1100 nucleotides predicted that removing the intervening sequences would be unlikely to compromise RNA transport function. This was clearly not the case, evidenced by the residual activity displayed by a mutant that lacked sequences spanning SL2 – SL6. Since chemical probing (a) found no evidence for an intramolecular SL1/SL7 interaction and (b) with the exception of the SL3 pseudoknot, did not predict additional, extensive long-range interactions, altering the SL1-SL7 spatial separation in a manner that was incompatible with sequestering obligate host factors would appear to be the most likely explanation for loss of activity. Stated differently, the overall 3-dimensional fold of the PTE may be a critical requirement, possibly to coordinate stepwise assembly of the export complex. In this respect, recent data for the HIV-1 Rev/RRE interaction, which is clearly the most extensively-studied RNA export system, has provided evidence for a hitherto unidentified Rev binding site in Stem 1 that may play a critical role in modulating Rev multimerization and RRE conformation[2]. An important extension of our analysis of the PTE would be to determine whether SL1 and/or SL7 constitute recognition sites for obligate cellular factors of the NXF1 RNA export machinery. Although the use of isolated components might provide a working model, *in vivo* analysis of the PTE export complex, using structural approaches such as SHAPE-seq [3] in combination with photocrosslinking strategies e.g PAR-CLIP [4], would be more instructive. Clearly the mechanistic basis for XMRV/MLV PTE function will require additional experimentation. However, our work [1] has provided additional evidence that long-range tertiary interactions are a common feature of RNA export elements. Additional examples of this are discussed in the following sections.

Tertiary interactions mediate activity of the murine *musD* transport element (MTE). The MTE, an ~400 nt RNA located adjacent to the 3′ long-terminal repeat region, mediates post-transcriptional control of genome export in the type D murine retrotransposon *musD*, a distant relative of retroviruses.
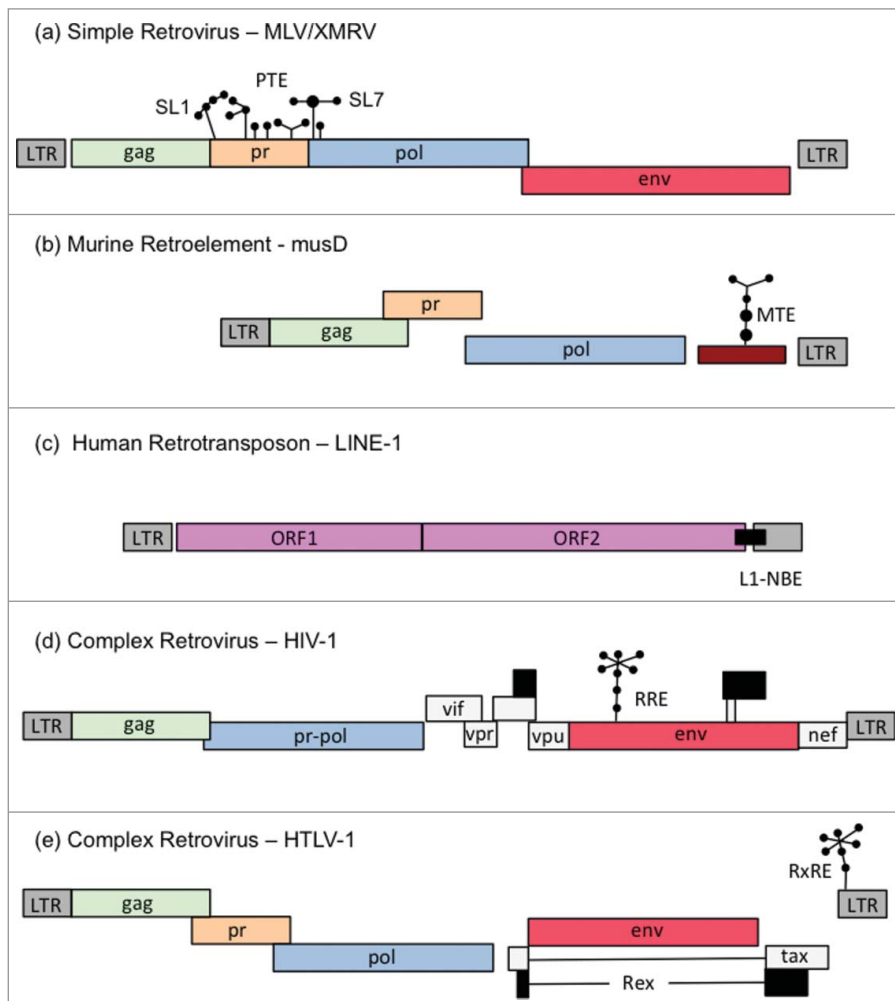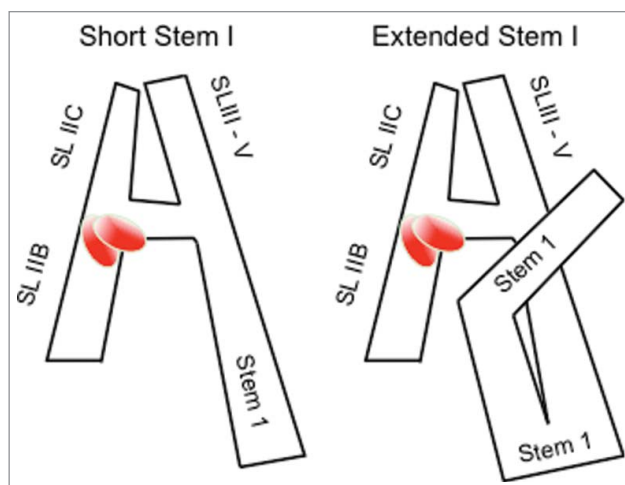

Figure 1. Localization of RNA export elements in retroviruses and retroelements. Murine retroelements, e.g. musD, have the MTE located adjacent to the 3′ LTR (**B**), while the counterpart from LINE-1 is located between ORF2 and the 3′ LTR (**C**). Complex retroviruses, such as HIV-1 have this element overlapping the env open reading frame (**D**), while the HTLV-1 RxRE (**E**) is located in the 3′ LTR. In contrast, the MLV PTE (**A**) is not only significantly larger (~1500 nt), but is located further upstream, overlapping the *Pr* and a portion of the *Pol* open reading frames.
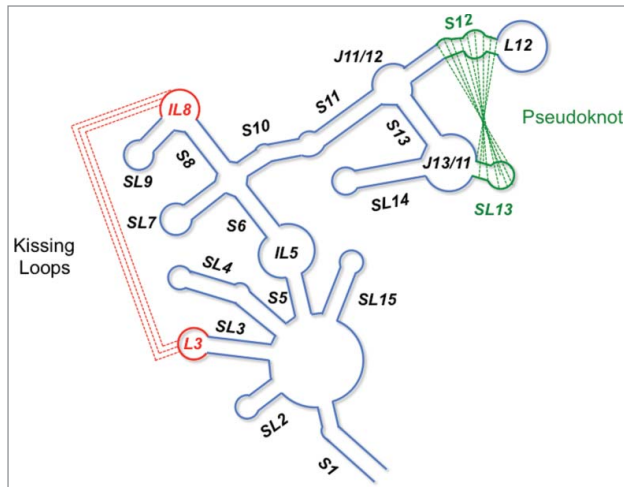
Figure 2. SHAPE-predicted secondary structure model of the musD RNA export element MTE, indicating long-range, tertiary interactions mediated by either a kissing complex (L3/IL8) or a pseudoknot (SL12/SL13). Adapted from [5]. *Insert,* ai-SHAPE analysis of the L3/IL8 interaction, depicting chemical reactivity of L3 nucleotides in the absence (blue trace) and presence (gold trace) of an oligobucleotide hybridized to IL8. Structural elements are defined by: S, stem; SL, stem-loop; IL, internal loop; L, loop; J, junction.

Chemoenzymatic probing of the MTE [5] revealed a highly organized structure comprising multiple stem-loops. Two potential tertiary interactions were also predicted in the form of (a) a "kissing" interaction involving nucleotides of Loop 3 (L3) and Internal Loop 8 (IL8), and (b) a complex pseudoknot structure between nucleotides G217 and C312 (**Fig. 2**). Biochemically, the L3/IL8 interaction could be verified using an "antisense(ai)-interfered" SHAPE approach. This involved hybridizing an oligonucleotide to one sequence of the complex and determining whether this induces chemical reactivity of the displaced partner (**Fig. 2**, insert). More importantly, using a *gag* reporter assay, *in vivo* experiments elegantly demonstrated that mutations disrupting the kissing complex impaired RNA transport,

while this could be restored by compensatory mutations. With respect to the more complex G217-C312 pseudoknot, and in support of ai-SHAPE, MTE mutations that retained base pairing as a simple secondary structure (i.e. removed the potential for pseudoknot formation) disrupted biological function, while replacing these with unrelated sequences that were predicted to support pseudoknot formation restored export activity to ∼30% of wild type. Although the dependence on these tertiary interactions for MTE function could be clearly demonstrated, the mechanistic basis for this was less evident, and a role for assuming the appropriate structure for recognition by obligate host protein(s) could once more be invoked. In combination with recent structural data for the HIV-1 Rev response element RRE discussed in the next section [2], such long-range tertiary interactions may represent a "decision-making" mechanism that controls the cytoplasmic distribution of retroelement RNA with respect to its requirement as the template for translation or as genetic material to be packaged into virus-like particles or virions.

A new "twist" to the HIV Rev responsive element (RRE). The notion that additional intramolecular interactions in RNA transport elements play a pivotal role their nucleocytoplasmic transport has now been extended to the HIV-1 RRE, whose "A-like" topology, determined by small angle X-ray scattering (SAXS), was reported in 2013 by Fang et al. [6]. The RRE construct of a later SAXS study conducted by Bai et al. [2], differed in size through an extension of Stem I but, somewhat surprisingly, concurred with respect to topology and particle dimension. Through an elegant series of experiments combining ai-SHAPE [7] and SAXS, Bai et al. could clearly demonstrated that long-range interactions were responsible for folding and re-alignment of the extended Stem I onto the multi-stem junction of the RRE, the compact nature of which accounted for the similarity in overall dimensions obtained by the 2 SAXS studies. At the same time, studies of Bal et al. [2], have also provided an explanation for observations that truncating Stem I reduced HIV-1 RNA nuclear transport, the structural basis of which was

Figure 3. Alternative configurations of the HIV-1 RRE are dependent on the length of Stem 1. Although the A-like topology is preserved between both structures [2, 6], recent SAXS of Bai et al. [2] has demonstrated that an extended Stem 1 is capable of forming tertiary interactions with upstream nucleotides that implicate a cryptic Rev binding site. The primary Rev binding site on Stem 2 is indicated in red. SL, stem-loop.

previously unclear. Nucleotides within the extended Stem I have now been proposed to interact with junction nucleotides immediately preceding Stem II, which houses the primary Rev binding site (**Fig. 3**). Lack of complementarity between these 2 regions suggested participation of long-range tertiary interactions. The extended Stem I was also shown to contain a cryptic Rev binding site that contributes to efficient tertiary folding of the RRE, thereby coordinating accurate assembly of the ribonucleoprotein complex. Although it remains to be determined whether equivalently-compacted RRE structures must be adopted as a general prerequisite to accelerated RNP formation in related retroviruses such as simian immunodeficiency virus or equine infectious anemia virus, the examples of long-range interactions briefly reviewed here clearly challenge our previous definitions of "simple" and "complex" retroviruses and retroelements.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Pilkington GR, Purzycka KJ, Bear J, Le Grice SF, Felber BK. Gammaretrovirus mRNA expression is mediated by a novel, bipartite post-transcriptional regulatory element. Nucleic Acids Res 2014; 42:11092-106; PMID:25190459; http://dx.doi.org/10.1093/nar/gku798
2. Bai Y, Tambe A, Zhou K, Doudna JA. RNA-guided assembly of Rev-RRE nuclear export complexes. eLife 2014; 3:e03656; PMID:25163983; http://dx.doi.org/10.7554/eLife.03656
3. Loughrey D, Watters KE, Settle AH, Lucks JB. SHAPE-Seq 2.0: systematic optimization and extension of high-throughput chemical probing of RNA secondary structure with next generation sequencing. Nucleic Acids Res 2014; PMID:25303992
4. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M, Jungkamp AC, Munschauer M, et al. PAR-CliP–a method to identify transcriptome-wide the binding sites of RNA binding proteins. J Vis Exp 2010; PMID:20644507
5. Legiewicz M, Zolotukhin AS, Pilkington GR, Purzycka KJ, Mitchell M, Uranishi H, Bear J, Pavlakis GN, Le Grice SF, Felber BK. The RNA transport element of the murine musD retrotransposon requires long-range intramolecular interactions for function. J Biol Chem 2010; 285:42097-104; PMID:20978285; http://dx.doi.org/10.1074/jbc.M110.182840
6. Fang X, Wang J, O'Carroll IP, Mitchell M, Zuo X, Wang Y, Yu P, Liu Y, Rausch JW, Dyba MA, et al. An unusual topological structure of the HIV-1 Rev response element. Cell 2013; 155:594-605; PMID:24243017; http://dx.doi.org/10.1016/j.cell.2013.10.008
7. Legiewicz M, Badorrek CS, Turner KB, Fabris D, Hamm TE, Rekosh D, Hammarskjold ML, Le Grice SF. Resistance to RevM10 inhibition reflects a conformational switch in the HIV-1 Rev response element. Proc Natl Acad Sci U S A 2008; 105:14365-70; PMID:18776047