Journal of Cell Signaling
ISSN: 2576-1471
Ryan et al., J Cell Signal 2017, 2:2
Commentary
Open Access

The Potential Consequences for Cell Signaling by a Class of NOD-Like Receptor Proteins (NLRs) Bearing an N-terminal Signal Sequence

Martin D Ryan[1*], Claire Roulston[1], Pablo de Felipe[2], Valerie Odon[3], Jens Tilsner[4] and Garry A Luke[1]

[1]*Biomedical Sciences Research Complex, School of Biology, University of St Andrews, North Haugh, St Andrews KY16 9ST, UK*

[2]*Spanish Medicines Agency (AEMPS), Parque Empresarial "Las Mercedes", Campezo 1 –Edificio 8, 28022 Madrid, Spain*

[3]*Department of Microbial Sciences, School of Biosciences and Medicine, University of Surrey, Guildford GU2 7XH, UK*

[4]*James Hutton Institute, Invergowrie, Dundee, DD2 5DA, UK*

[*]**Corresponding author:** Martin D Ryan, Biomedical Sciences Research Complex, School of Biology, University of St Andrews, North Haugh, St Andrews KY16 9ST, UK, Tel: 01334 463403; E-mail: martin.ryan@st-andrews.ac.uk

Received date: April 17, 2017; **Accepted date:** May 10, 2017; **Published date:** May 19, 2017

Commentary

Toll-like and NOD-like receptors (TLRs/NLRs) recognise pathogen- or damage-associated molecular patterns (PAMPs/DAMPs) and are key initiators of cell signalling pathways by which the cell mounts an appropriate defence response. Whilst TLRs are type I transmembrane proteins, NLRs are cytosolic proteins. Determination of the genome sequence of the purple sea urchin (*Strongylocentrotus purpuratus*) revealed this echinoderm had a greatly expanded repertoire of TLR/NLR innate immunity genes in comparison with vertebrates possessing an acquired immune system [1]. The open reading frame (ORF) of a number of NLR genes in *S. purpuratus* (and organisms from other phyla) commence with a leader sequence with sequence similarity to '2A/2A-like' oligopeptide sequences found in many virus genomes, which mediate an unusual translational 'recoding' event in which translation arrests at the C-terminus of 2A (mid-ORF; no stop codon present), but then may recommence the synthesis of the downstream protein as a separate, entirely discrete, translation product [2-5]. Bioinformatic analyses also suggested, however, that these *S. purpuratus* NLR 2A-like sequences might also function as co-translational signal sequences. We showed this type of 2A-like leader sequence possessed both activities [6], and in this commentary we speculate upon the possibility that such a single gene could produce both cytoplasmic and membrane-bound/secreted forms of such NLRs.

'2A' and '2A-like' oligopeptide sequences were first characterised from a group of single-stranded positive (mRNA) sense RNA viruses called picornaviruses. These short (~20aa) sequences were shown to mediate a co-translational 'cleavage' of a polyprotein at their own C-termini. This apparent 'cleavage' was, however, not a product of proteolysis, but a newly-characterised form of translational 'recoding' in which the synthesis of a specific peptide bond at the C-terminus of 2A was 'skipped'. Briefly, the model of this mechanism states that when the elongating ribosome translates the 2A sequence, the nascent polypeptide sequence within the exit tunnel of the ribosome (2A) interacts with the tunnel to pause elongation and inhibit peptide bond formation. This is proposed to bring about the release of the nascent peptide synthesised to that point, but then translation of the sequences downstream of 2A may recommence. In this manner the translation products are actually synthesised as discrete products rather than as a single product which is then cleaved apart [2-5].

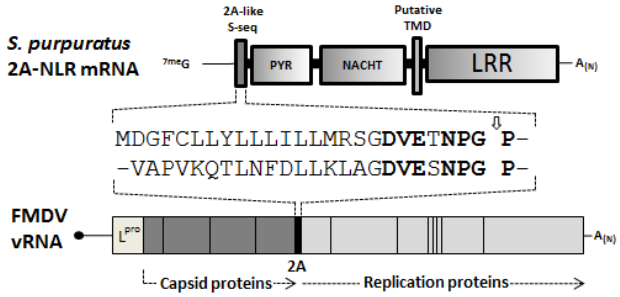

Figure 1: NLR and Virus 2A sequences. The structure of the mRNA encoding the *S. purpuratus* NLR is shown encoding a single ORF comprising (i) a 2A-like sequence at the N-terminus, then (ii) pyrin (PYR), (iii) nucleotide-binding (NACHT), (iv) putative transmembrane (TMD) and (v) leucine-rich repeat (LRR; PAMP/DAMP recognition) domains (boxed areas). The virus RNA (vRNA) genome of the positive-stranded RNA Foot-and-Mouth Disease Virus (FMDV) encodes a single ORF (a ~2,300aa 'polyprotein'; boxed areas) comprising a virus-encoded proteinase at the N-terminus (Lpro), the capsid proteins domain, 2A, then the domains comprising the various genome replication proteins. The genome bears an oligopeptide cap structure (unlike cellular mRNAs) and is polyadenylated. The 2A-like sequence of the *S. purpuratus* NLR is shown together with FMDV 2A. The C-terminal motif conserved amongst all 2As characterised to date is shown in bold, and the site of translational recoding is shown by the vertical arrow.

Many 2As possess only partial recoding activity: if, for example, a particular 2A sequence recodes translation only 50% of the time, then in 50% of the translation products the peptide bond is formed and the resulting translation profile is a mixture of the recoded ('cleaved') products (50%) plus the fusion protein (50%).

Cellular 2A-like Sequences

Database probing with a motif characteristic of 2A ([-D(V/I)ExNPGP-]) revealed the presence of '2A-like' sequences in the genomes of many other RNA viruses but, surprisingly, also multiple occurrences within the genome of *S. purpuratus*. Here, 2A-like

sequences were observed within two main types of cellular sequence: non-LTR retrotransposons [7,8] and certain NLRs [6]. In the latter case our surprise was compounded in that these particular 2A-like sequences formed the N-termini of certain NLR proteins: until that point all of the (virus) 2A/2A-like sequences we had studied were internal features used for the co-expression of protein domains up- and down-stream of 2A (Figure 1). These cellular NLR-2As were of approximately the same length as virus 2As (~25-30aa), but comprised a higher proportion of hydrophobic residues and were predominantly flanked by charged residues – suggestive of a signal sequence and, indeed, bioinformatic analyses confirmed these NLR-2As could be signal sequences. But how to show these NLR-2As could (i) mediate translational recoding and (ii) could function as N-terminal signal sequences? Given the lack of immortalised *S. purpuratus* cell-lines, the limited availability of primary cell-lines, the lack of anti-NLR antibody probes etc., we chose to recapitulate the system using fluorescent reporter proteins (to study the effects upon protein localisation) and, since the function of N-terminal sequences is conserved across kingdoms, we used both immortalised, readily transfectable, mammalian (HeLa) cells and plant (*Nicotiana benthamiana*) leaf epidermal cells. Our data showed that these NLR-2As could both mediate (partial) translational recoding and act as N-terminal signal sequences [6]. When the NLR-2A recoded itself from the fluorescent reporter protein (mCherryFP), this protein localised to the cytoplasm. In the proportion of translation products where the NLR-2A did not recode – remained as an N-terminal feature of mCherryFP - then mCherryFP was secreted from the cell. Our data showed that in mammalian cells the NLR-2A was removed from the secreted mCherryFP, presumably by signal peptide peptidase.

Innate Immune Signalling in *S. purpuratus*

The systems outlined above allowed us to demonstrate that these NLR-2As were a new type of bi-functional translational recoding/signal sequence – but what are the implications for cell signalling in the innate immune system of *S. purpuratus*? TLRs are type I transmembrane proteins. Upon activation by PAMPs/DAMPs binding to the leucine-rich repeat domain (LRR), TLRs recruit adapter proteins within the cytosol of the immune cell in order to initiate the propagation of signal transduction pathways (Figure 2, Panel A). NLRs are, however, entirely cytosolic proteins detecting and signalling the presence of PAMPs/DAMPs present within the cytoplasm (Figure 2, Panel B). Recognition of the ligand by the LRR leads to the NOD/NBD – nucleotide-binding (NACHT) domain mediating ATP-dependent self-oligomerisation and initiates cell signalling by the N-terminal effector binding region comprising a protein-protein interaction domain such as a caspase recruitment domain (CARD), pyrin (PYD), baculovirus inhibitor repeat (BIR) or transactivator domain (AD). The end result of NLR signalling is the cell producing an inflammatory response, undergoing autophagy or cell death.

Database annotation of the *S. purpuratus* genome is based upon detection of similarities with other, previously characterised, protein domains: for example, NCBI entry XP011670951.1 (predicted NACHT, LRR and PYD domains-containing protein 3-like [S. purpuratus]) indicates a predicted Death Domain (aa's 38-177), a P-loop NTPase (NACHT) domain (aa's 222-414) prior to the C-terminal LRR domain binding PAMPs/DAMPs. Interestingly, the transmembrane protein structure prediction algorithm MemBrain [9] predicts three main transmembrane domains (TMDs) for this protein; the first centring on residues 9/10 (propensity=0.97) – corresponding to the N-terminal 2A-like signal sequence, the second centring on residues 441/442 (propensity=0.92) – immediately after the NACHT domain, and the third on residues 474/475 (propensity=0.78) – potentially prior to the LRR domain: could either/both of these putative TMDs between the NACHT and LRR domains act as internal signal-anchor sequences?

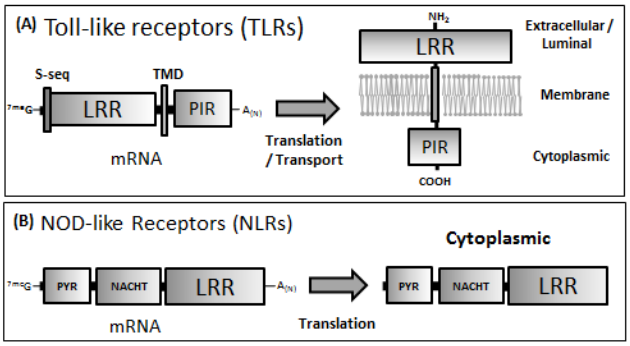

Figure 2: TLR and NLR Innate immunity receptors. Toll-like receptors bear an N-terminal signal sequence targeting the leucine-rich repeat domain (LRR; PAMP/DAMP recognition) of the protein to either the extracellular space or the lumen of endosomes. The internal signal-sequence/stop-transfer domain arrests protein translocation with the signalling (PIR) domain of the protein being cytoplasmic – the orientation of a type I transmembrane protein (Panel A). NOD-like receptors do not possess any sub-cellular signalling sequences: the arrangement of protein domains shows the opposite polarity to NLRs, and comprises an additional central NOD/NBD – nucleotide-binding (NACHT) domain (Panel B).

The presence of a functional N-terminal signal sequence (functional at least in mammalian and plant cells) together with putative transmembrane/internal signal-anchor/stop-transfer domains between the NACHT and LLR domains raises a number of questions with regards the sub-cellular localisation, potential membrane association (and, if this is the case, the protein orientation in the membrane) of these *S. purpuratus* NLR proteins. As shown in (Figure 3) (left panels), if the 2A-like signal sequence does not mediate a translational recoding event (remaining fused to the NLR), then the protein could either be secreted from the cell (no effect of putative TMD/internal signal-anchor/stop-transfer sequences), or, become a type 1 transmembrane protein (putative TMD/internal signal-anchor/stop-transfer sequence functional), although it is difficult to envisage a signalling function of such a molecule either being completely secreted from the cell or with this particular orientation within the membrane. As shown in (Figure 3) (right panels) shows the possibilities if the 2A-like signal sequence does mediate a translational recoding event (2A-like signal sequence not present at the N-terminus of the NLR): the NLR could become cytoplasmic (no effect of putative TMD/internal signal-anchor/stop-transfer sequences), or, perhaps, become either a type I or a type II transmembrane protein (internal signal-anchor/stop-transfer sequence functional). The orientation of the protein in the membrane is determined by the difference in the charges of the 15 residues flanking the first internal signal-anchor/stop-transfer sequence, with the more positive portion facing the cytosol [10].

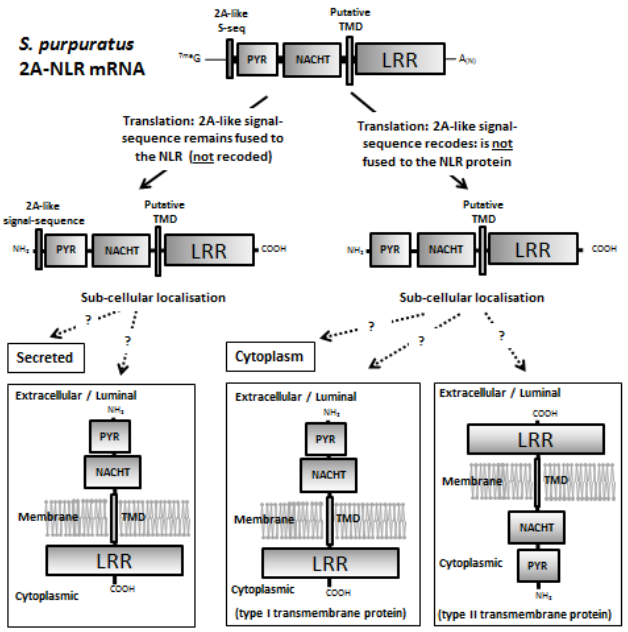

Figure 3: *S. Purpuratus* NLR Innate immunity receptors. The organisation of protein domains encoded by the mRNA is shown at the top of the figure. Translation of the mRNA, however, can produce 2, distinct, translation products. In the first case, the NLR 2A-like signal sequence does not mediate translational recoding (remains fused to the NLR protein), the NLR could either be secreted from the cell, or, if the putative transmembrane domain (TMD) is recognised by the exocytic pathway, adopt a type I transmembrane orientation with the LRR cytoplasmic and the NACHT oligomerisation and pyrin (PYR) domains oriented to the exterior of the cell (left panels). In this illustration we have assumed that the 2A-like signal sequence is removed by signal peptide peptidase (as is the case in mammalian cells). In the second case, where the 2A-like S-seq does mediate translational recoding, the NLR could be localised to the cytoplasm (putative TMD not recognised by the exocytic pathway), or, if the putative TMD is recognised by the exocytic pathway then adopts the orientation of either a type I or type II transmembrane protein – depending upon the charge distribution surrounding the putative TMD (right panels).

Again, it is difficult to envisage the function of such an NLR oriented as a type I transmembrane protein, but if it became a type II trans membrane protein, then could it fulfil the same type of function as a TLR? Interestingly, the 2A-like signal sequence could be acquired by these *S. purpuratus* NLR genes by a process of 'exon-shuffling' since it is encoded by its own mini-exon preceding the bulk of the NLR gene: in the case of the NLR described above (XP011670951.1), encoding the entire signal plus six amino acids downstream MDGFCLLYLLLILLMRSGDVETNPG↓PNTAVS; 2A-like signal sequence underlined, site of translational recoding indicated by vertical arrow).

Our paper presented analyses of the function of these N-terminal 2A-like signal sequences of NLR genes in *S. purpuratus* [6]. It should be noted, however, that we have detected a similar 2A-NLR gene organisation in other phyla; the cephalochordate *Branchiostoma floridae*, the sponge *Amphimedon queenslandica*, the mollusc *Lottia gigantea* and the arthropod *Ixodes scapularis* [11].

These data, plus subsequent bioinformatic analyses presented here, raise a number of important questions as to the function of these 'NLR-like' genes in these organisms. It may prove to be the case that they may have not only expanded the number of innate immunity genes, but also expanded the functionality of some of these NLR proteins – potentially involving novel signalling pathways.

Acknowledgement

The authors gratefully acknowledge the support of the UK Biotechnology and Biological Sciences Research Council (BBSRC) for the work on cellular 2A-like sequences.

References

1. Hibino T, Loza-Coll M, Messier C, Majeske AJ, Cohen AH, et al. (2006) The immune gene repertoire encoded in the purple sea urchin genome. Dev Biol 300: 349-365.
2. Ryan MD, Drew J (1994) Foot-and-mouth disease virus 2A oligopeptide mediated cleavage of an artificial polyprotein. EMBO J 134: 928-933.
3. Donnelly MLL, Luke G, Mehrotra A, Li X, Hughes LE, et al. (2001) Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. J Gen Virol 82: 1013-1025.
4. de Felipe P, Hughes LE, Ryan MD, Brown JD (2003) Co-translational, intraribosomal cleavage of polypeptides by the foot-and-mouth disease virus 2A peptide. J Biol Chem 278: 11441-11448.
5. Sharma P, Yan F, Doronina V, Escuin-Ordinas H, Ryan MD, et al. (2012) 2A peptides provide distinct solutions to driving stop-carry on translational recoding. Nuc Acids Res 40: 1-9.
6. Roulston C, Luke GA, de Felipe P, Ruan L, Cope J, et al., (2016) '2A-Like' signal sequences mediating translational recoding: A novel form of dual protein targeting. Traffic 17: 923-939.
7. Odon V, Luke GA, de Felipe P, Ruan L, Roulston C, et al. (2013) APE-type non-LTR retrotransposons of multicellular organisms encode virus-like 2A oligopeptide sequences which mediate translational recoding during protein synthesis. Mol Biol Evol 30: 1955-1965.
8. Luke GA, Roulston C, Odon V, de Felipe P, Sukhodub A, et al. (2013) Lost in translation: The biogenesis of non-LTR retrotransposon proteins. Mob Genet Elements 3: e27525.
9. http://www.csbio.sjtu.edu.cn/cgi-bin/MEMBRAIN.cgi
10. Hartmann E, Rapoport TA, Lodish HF (1989) Predicting the orientation of eukaryotic membrane-spanning proteins. Proc Natl Acad Sci USA 86: 5786-5790.
11. Roulston C (2015) Occurrence and function of cellular 2A sequences. PhD Thesis University of St Andrews.