Regions of the Cry1Ac Toxin Predicted to be under Positive Selection are shown to be the Carbohydrate Binding Sites and can be Altered in their Glycoprotein Target Specificity

Norah Cassidy[1, 2], Roisin Thompson [2] ,Paul Clarke [2], Damien Keogh [2], Brendan O'Connor[1, 2] , Michael O'Connell[1, 2]

1. *Irish Separation Science Cluster, Dublin City University, Dublin, Ireland*
2. *School of Biotechnology, Dublin City University, Dublin, Ireland*

Introduction

The *cry* gene family, is a large family of homologous genes from *Bacillus thuringiensis*. Studies have examined the structural and functional relationships of the Cry proteins. They have revealed several residues in domains II and III that are important for target recognition and receptor attachment. In 2007 Wu, Jin-Yu et al employed a maximum likelihood method to detect evidence of adaptive evolution in Cry proteins. They identified positively selected residues, which are all located in Domain II or III. Figure 1 shows a protein sequence alignment between domain II and III of Cry1Ac and Cry1Aa. This highlights the areas which are thought to be under positive selection. Cry1Ac and Cry1Aa are structurally very similar and they both bind to a variety of N-aminopeptidases (APN's) in different insect species. However Cry1Aa has a higher specificity for the cadherin like receptor HevCalP and Cry1Ac binds to N-acetylgalactosamine (GalNAc) on the surface of APN's. Differences in the binding of the two toxins has been shown in an in-direct toxin-binding assay where GalNAc completely abolished toxin binding of Cry1Ac but had no effect on the binding of Cry1Aa. The binding site has been shown to be located in the third domain of Cry1Ac. Some of these sites correlate with the positively selected residues found by Wu et al 2007 in Cry1Aa. Our aim was to use the comparison of the toxins to analyse the potential to alter the binding specificity of Cry1Ac and its domains. In this work we identified critical amino acid residues for this objective.

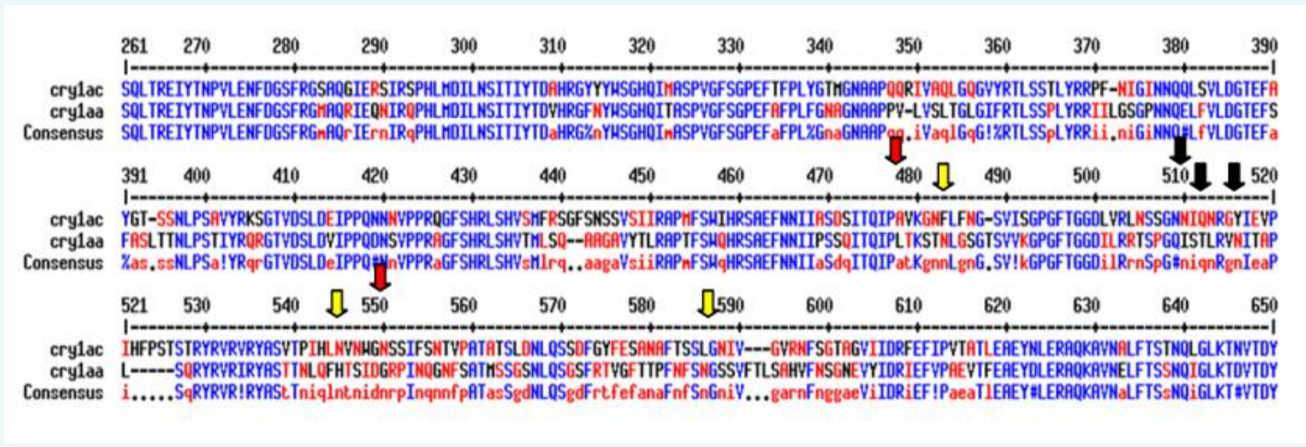

Figure 1: ***Protein sequence alignment of Cry1Ac and Cry1Aa***
This sequence alignment spans the second and third domains of both proteins. It shows the amino acids which are positively selected and undergoing adaptive evolution represented by yellow arrows. The black arrows indicate the amino acids involved in GalNAc binding in Cry1Ac. While the red arrows show the amino acids which are positively selected and are predicted to be involved in GalNac and Galactose binding according to the protein structure models (Figure 2).

Results

Two truncated forms of Cry1Ac (tCry1Ac and CryD3) have been cloned, expressed in *Escherichia coli* and purified by Immobilised Metal Affinity Chromatography (IMAC) (Figure. 3). The binding abilities of both proteins have been analysed in an Enzyme Linked Lectin Assay (ELLA) (Figure. 4).

1. Predicted protein structure for the recombinant tCry1Ac and CryD3

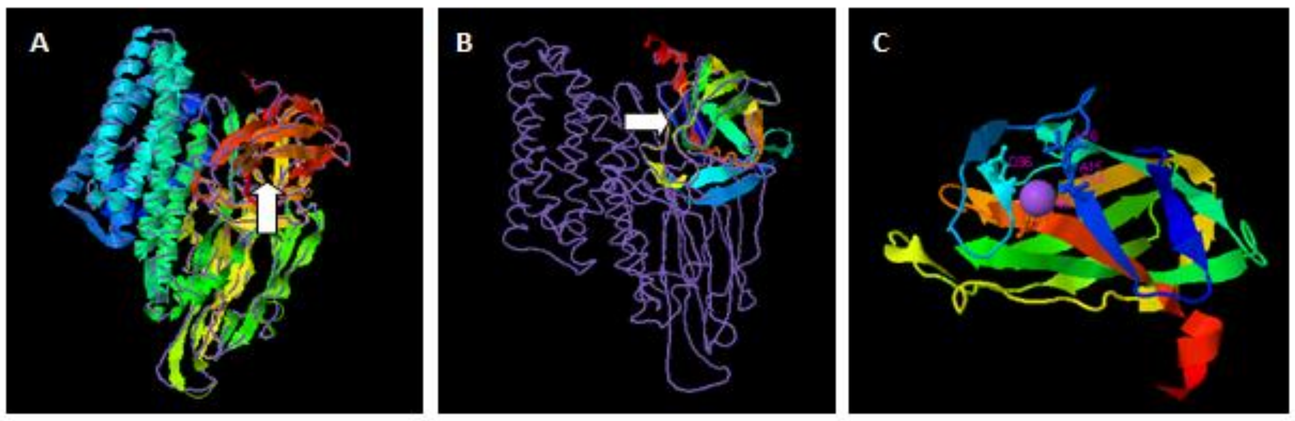

Figure 2: ***Expected protein structure models from I-TASSER (protein structure and function predictions) of cloned tCry1Ac and CryD3****.*

A. *The expected structure of tCry1Ac; the white arrow points to the known binding site of GalNAc on the third domain (red).* ***B*** *shows the expected structure of the cloned CryD3 fitting into the Cry1Ac model, the white arrow points to the expected binding site of Galactose.* ***C*** *shows a more detailed view of the predicted binding site of CryD3 with a salt ion in purple. The 6 X histidine tag is represented in red in all pictures.*

2. Purification of tCry1Ac and CryD3 by IMAC

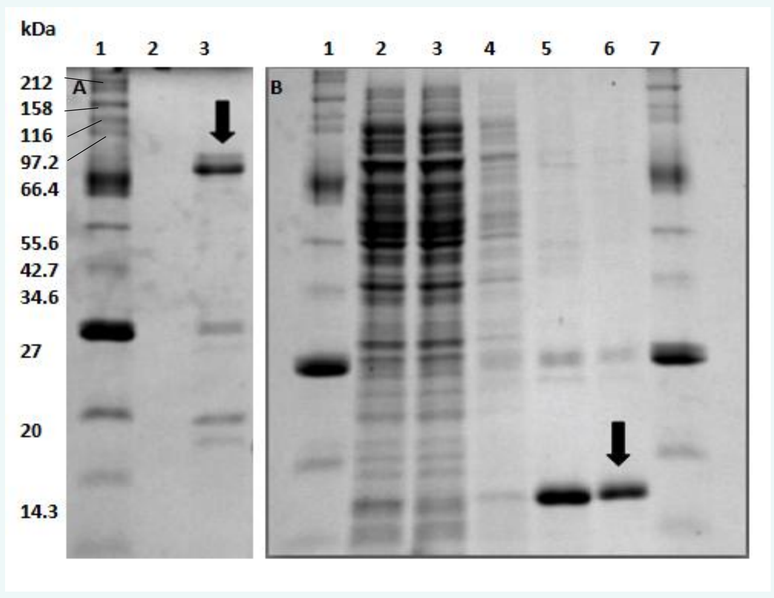

Figure 3: ***Purification of tCry1Ac (A) and CryD3 (B)***

A: 1*; Broad range protein ladder,* ***3****; Elution of tCry1Ac at 69kDa (lower bands are cleaved protein).* ***B: 1 and 7;*** *Broad range protein ladder,* ***2****; Cleared lysate,* ***3****; Flow through of contaminant proteins,* ***4****; Wash at 100Mm imidazole,* ***6 and 7****; Elution of CryD3 at 15kDa (CryD3 dimer at 30kDa)*

3. ELLA of tCry1Ac, CryD3 and ECL

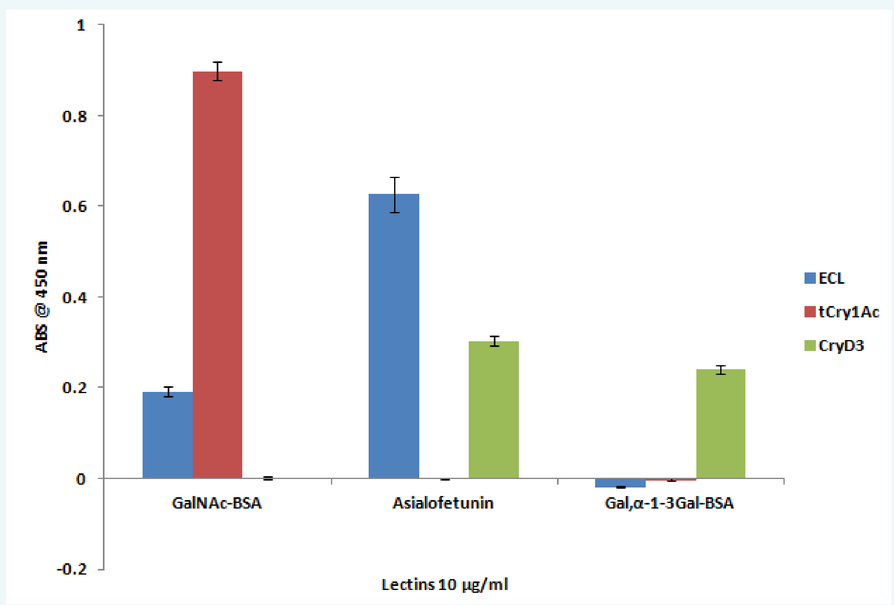

Figure 4: ***ELLA showing binding of tCry1ac, CryD3 and the commercial lectin ECL to neo-glycoproteins; GalNAc-BSA and Gal,α1-3gal-BSA and the glycoprotein Asialofetuin***

Significantly tCry1Ac shows binding to GalNac, while CryD3 binds to Asialofetunin and α1-3 galactose. This result correlates with the predicted binding models from I-TASSER.

Conclusions and future work

We have cloned, expressed, and purified truncated forms of Cry1Ac, which we have shown to be active, binding to GalNac and Galactose in ELLA's. Significantly, the recombinant CryD3 has a specificity for Galactose rather than GalNac. This shows that small alterations of Cry1Ac have effects on the toxin specificity. Further ELLA's will be carried out on the recombinant tCry1Ac and CryD3 to examine the stability of the proteins and the specificity if any to the T (Gal-beta1,3-GalNAc-alpha1,O-Ser/Thr) and T_n (GalNAc-alpha1,O-Ser/Thr) antigen and other glycoproteins. Site directed mutagenesis is being carried out on the predicted binding sites of both tCry1Ac and CryD3 to investigate the structural basis for binding.

These results show that novel lectins with altered binding specificities, that may have bio-analytical applications, can be formed by the truncation of these toxins.

This research has been funded by Science Foundation Ireland under grant number 08/SRC/B1412