Biochimica et Biophysica Acta xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagen

Generating novel recombinant prokaryotic lectins with altered carbohydrate binding properties through mutagenesis of the PA-IL protein from *Pseudomonas aeruginosa*

Damien Keogh [a,1], Roisin Thompson [a,b,1], Ruth Larragy [a,b], Kenneth McMahon [a], Michael O'Connell [a], Brendan O'Connor [a,b], Paul Clarke [a,b,*]

[a] School of Biotechnology, Dublin City University, Glasnevin, Dublin 9, Ireland
[b] Irish Separation Science Cluster (ISSC), National Centre for Sensor Research (NCSR), Dublin City University, Glasnevin, Dublin 9, Ireland

ARTICLE INFO

Article history:
Received 29 July 2013
Received in revised form 17 December 2013
Accepted 13 January 2014
Available online xxxx

Keywords:
LecA
N-acetyl-lactosamine
Galectins
Lectinology
Glycoprotein analysis
Glycoprotein purification

ABSTRACT

Background: Prokaryotic lectins offer significant advantages over eukaryotic lectins for the development of enhanced glycoselective tools. Amenability to recombinant expression in *Escherichia coli* simplifies their production and presents opportunities for further genetic manipulation to create novel recombinant prokaryotic lectins (RPLs) with altered or enhanced carbohydrate binding properties. This study explored the potential of the α-galactophilic PA-IL lectin from *Pseudomonas aeruginosa* for use as a scaffold structure for the generation of novel RPLs.
Method: Specific amino acid residues in the carbohydrate binding site of a recombinant PA-IL protein were randomly substituted by site-directed mutagenesis. The resulting expression clones were then functionally screened to identify clones expressing rPA-IL proteins with altered carbohydrate binding properties.
Results: This study generated RPLs exhibiting diverse carbohydrate binding activities including specificity and high affinity for β-linked galactose and N-acetyl-lactosamine (LacNAc) displayed by N-linked glycans on glycoprotein targets. Key amino acid substitutions were identified and linked with specific carbohydrate binding activities. Ultimately, the utility of these novel RPLs for glycoprotein analysis and for selective fractionation and isolation of glycoproteins and their glycoforms was demonstrated.
Conclusions: The carbohydrate binding properties of the PA-IL protein can be significantly altered using site-directed mutagenesis strategies to generate novel RPLs with diverse carbohydrate binding properties.
General significance: The novel RPLs reported would find a broad range of applications in glycobiology, diagnostics and in the analysis of biotherapeutics. The ability to readily produce these RPLs in gram quantities could enable them to find larger scale applications for glycoprotein or biotherapeutic purification.

1. Introduction

Lectins are proteins that are capable of recognizing and binding reversibly to specific oligosaccharide (glycan) structures displayed on cell surfaces and glycoproteins [1–3]. Since their initial discovery, lectins have found diverse applications [4], but with the growing interest in the field of glycobiology lectins have come to the forefront as important analytical and diagnostic tools helping to elucidate the basis of complex biological processes and disease states [5,6]. Lectins display exquisite specificity for cognate ligands, and their ability to bind to glycans *in situ* facilitates interrogation of the glycosylation status of a sample without the prior need for glycan release and derivatization [1]. Labelled lectins are widely used in glycoanalytical procedures such as lectin blotting and ELLAs that are analogous to the standard ELISA [7–10]. Lectin arrays, which represent powerful tools enabling high throughput glycoanalysis, are increasingly being reported in the literature for a diverse range of applications including cancer diagnostics [6,11–16]. Through immobilization onto solid support media, the ability of lectins to discriminate between different glycan structures can be harnessed to effect the separation and selective purification of glycoproteins and their glycoforms. Lectin affinity chromatography (LAC) is now commonly used in laboratories for glycoprotein purification and is often used as an initial step to pre-concentrate glycopeptides, or separate glycoforms, prior to MS-based glycoanalysis [17–19].

The most commonly used lectins are plant derived, but these have a number of significant limitations. These eukaryotic lectins can be structurally complex and often require post-translational modifications, including glycosylation, or disulfide bond formation [1,2]. As a result of

Abbreviations: PA-IL, *Pseudomonas aeruginosa* lectin 1; PA-ILU, untagged native PA-IL; rPA-IL, recombinant PA-IL; 6HIS, affinity purification tag comprised of 6 tandem histidine residues; ELLA, enzyme linked lectin assay; iGb3, isoglobotriaosylceramide (Galα1–3Galβ1–4Glc); PBS, Phosphate Buffered Saline; TBS, Tris Buffered Saline; TBST, Tris Buffered Saline with Tween 20; IPTG, Isopropyl-β-D-thiogalactopyranoside

* Corresponding author at: School of Biotechnology, Dublin City University (DCU), Glasnevin, Dublin 9, Ireland. Tel.: +353 1 700 5961; fax: +353 1 7005412.
E-mail address: Paul.Clarke3@gmail.com (P. Clarke).
[1] Authors made equal contribution to the work reported.

0304-4165/$ – see front matter © 2014 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.bbagen.2014.01.020

this, eukaryotic lectins are often not amenable to recombinant production, particularly in *Escherichia coli* based systems which are the systems of choice for such purposes [20]. While many attempts have been made to produce plant lectins recombinantly, these frequently result in the production of insoluble proteins and very low yields [20,21]. Consequently, plant lectins are usually purified from source materials by conventional liquid chromatography techniques, and it is from here that many of their limitations stem. The quality, specificity and activity of these lectins can vary from one supplier to another and from batch to batch depending on the quality of the starting material and the purification methods used in their preparation [6,20]. These inconsistencies in performance complicate applications for glycoprotein analysis, particularly in relation to their use for diagnostic purposes. In addition, the need to purify from source material means that only limited quantities of these lectins can be obtained, and the final lectin preparations can be expensive [20]. These issues have restricted the use of lectins to analytical scale applications where large quantities are not generally required.

While lectins are ubiquitous in nature, prokaryotic lectins present unique opportunities for the development of new glycoselective bioaffinity tools. Most of these advantages stem from their greater amenability to production in *E. coli* as recombinant prokaryotic lectins (RPLs). In addition to facilitating simplified production, recombinant approaches enable the modification and optimization of lectin binding specificities and affinities [22–24] as well as opportunities to facilitate orientation specific immobilization to generate enhanced glycoanalytical platforms [25]. Despite their many advantages, prokaryotic lectins have to date remained largely underexploited, but there is now growing interest in the exploitation of these new lectin sources as they have the potential to redress the many shortcomings of the traditional plant lectins and lead to more sensitive and robust glycoanalytical and diagnostic tools [20].

The PA-IL protein is one of two soluble lectins produced by the opportunistic pathogen *Pseudomonas aeruginosa* (Fig. 1) [26,27]. The binding specificity of this lectin has been examined extensively using a broad spectrum of methodologies, and it has been shown to bind preferentially to glycans with terminal α-linked galactose and not to bind significantly to glycans with terminal β-linked galactose [28]. The PA-IL protein is known to exhibit relative resistance to heating, extreme pH and proteolysis [27] and therefore embodies ideal physiochemical properties as a template scaffold protein structure for the generation of novel robust glycoanalytical tools. This study explored the potential of the PA-IL protein for the generation of novel RPLs with altered or enhanced carbohydrate binding properties. It demonstrated that the carbohydrate binding specificity of the PA-IL protein could be significantly altered by introducing random substitutions at specific amino acid residues in the protein's carbohydrate binding site. Novel RPLs displaying diverse carbohydrate binding properties, including specificity and high affinity for glycans with terminal β-linked galactose and N-acetyl-lactosamine (LacNAc) displayed by N-linked glycans on glycoprotein targets, were successfully generated. Some of these RPLs displayed affinities for these glycan epitopes significantly greater than that of the equivalent commercially available plant lectin ECL (*Erythrina cristagalli* Lectin) [21,29]. The functional characterization of a collection of these RPLs enabled the identification of specific amino acid substitutions linked with observed carbohydrate binding properties. Ultimately, the utility of these novel RPLs for simple and highly sensitive glycoprotein analysis, without the need for prior labelling, was demonstrated. Through immobilization onto sepharose resins and magnetic particles, we demonstrated the utility of these RPLs for efficient glycoselective fractionation and isolation of glycoproteins and their glycoforms. The ability to produce these RPLs in large quantities and to very high levels of purity via cost-effective and scalable protocols could enable them to overcome the limitations of plant and other eukaryotic lectins and to potentially find large-scale applications for the selective purification of glycoproteins and glycosylated biotherapeutics.

2. Materials and methods

2.1. Plasmid construction—pQE30PA-IL & pQE60PA-IL

All strains and plasmids used or constructed as part of this study are listed and described in Table 1. The *lecA* gene encoding the PA-IL protein was amplified from *P. aeruginosa* PAO1 genomic DNA by PCR to facilitate cloning into the pQE series of *E. coli* expression vectors from Qiagen. PCR reactions were carried out using high fidelity Phusion Taq and PCR conditions recommended by the manufacturer (New England BioLabs). The *lecA* gene was amplified using primers PA-IL-F1 [5′- AAAAGGATCCATGG CTTGGAAAGGTGAGG-3′] and PA-IL-R1 [5′-AAAAAAGCTTTCACGACT GATCCTTTCCAATATT-3′] which generated a product that could be cloned as a BamHI-HindIII fragment into the pQE30 expression vector. The resulting plasmid, pQE30PA-IL (Supplementary Fig. 1A), expressed a recombinant PA-IL protein (rPA-IL) with an amino terminal 6HIS tag (rPA-ILN). The *lecA* gene was also amplified using primers PA-IL-F2

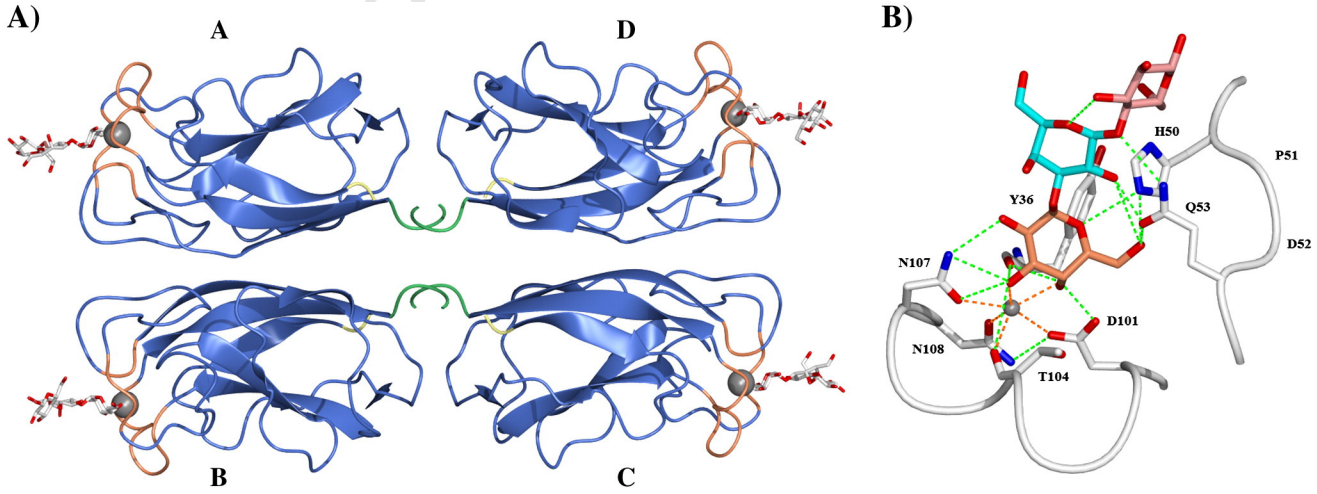

Fig. 1. Structure of the PA-IL protein and its carbohydrate binding site. (A) Tetrameric PA-IL protein with bound iGb3 trisaccharide Galα1–3Galβ1–3Glc (PDB code 2VXJ) [28]. Each monomer subunit contains a single carbohydrate binding site. A single calcium ion is coordinated within each binding site (grey sphere) and is essential for sugar binding. Residues involved in calcium coordination and carbohydrate binding are highlighted (Orange). The C-terminus and N-terminus of each monomer are highlighted (Green and Yellow respectively). (B) The PA-IL binding site showing coordination of the calcium ion and binding of iGb3. Amino acid residues involved in calcium coordination (Red lines) and contributing to the formation of hydrogen bonds (Green lines) with bound iGb3 are indicated. Images were generated using Deep View (Swiss Model) [50] and rendered using CCP4MG software [51].

Please cite this article as: D. Keogh, et al., Generating novel recombinant prokaryotic lectins with altered carbohydrate binding properties through mutagenesis of the PA-IL pro..., Biochim. Biophys. Acta (2014), http://dx.doi.org/10.1016/j.bbagen.2014.01.020