JOURNAL OF PROTEOMICS XX (2013) XXX–XXX

Available online at www.sciencedirect.com

SciVerse ScienceDirect

www.elsevier.com/locate/jprot

Proteomic survey of the *Streptomyces coelicolor* nucleoid

*Elizabeth Bradshaw**, *Gerhard Saalbach*, *Michael McArthur*

John Innes Centre, Norwich, UK

ARTICLE INFO

Article history:
Received 9 November 2012
Accepted 28 February 2013

Keywords:
DNA-binding
LC–MS/MS
Nucleoid
H-NS
IHF
Streptomyces

ABSTRACT

Nucleoid-associated proteins (NAPs) are small, highly abundant transcriptional regulators with low sequence specificity which are involved in multiple DNA-related processes including gene expression, DNA protection, recombination/repair and nucleoid structuring. Through these functions they are able to regulate important phenotypic properties including virulence, secondary metabolism and stress resistance. However the set of NAPs known within the Actinobacteria is small and incomplete. The missing proteins are likely to be key regulators of virulence in pathogens such as *Mycobacterium tuberculosis* and also of development and secondary metabolism in industrially-important species such as *Streptomyces*. Here, we use label-free LC–MS/MS to systematically search for novel NAPs in isolated nucleoids of the model actinomycete *Streptomyces coelicolor*. Based on the criteria of high abundance (emPAI score) and predicted DNA-binding ability (DNAbinder score) we identified a set of 24 proteins with a high predicted likelihood of being NAPs. The approach was deemed successful as the set included known major NAPs HupA, HupS, sIHF and Lsr2 as well as the global transcriptional regulators BldD and CRP and the pleiotropic response regulator AfsQ1. It also included a number of proteins whose functions are not yet known from recognisable classes of transcription factor (SCO2140, SCO4493, SCO1839, SCO1210, SCO5405, SCO4229, SCO3198) or from uncharacterised protein families (SCO5783, SCO5592, SCO3793, SCO6482) which comprise a valuable set of candidates for further study.

Biological significance

In this paper we establish a robust protocol for preparing *S. coelicolor* nucleoids for mass spectrometric analysis and develop a workflow for identifying novel nucleoid-associated proteins (NAPs) by combining LC–MS/MS with a bioinformatical analysis. The nucleoid-associated proteins of many species are known to be key regulators of virulence, stress tolerance and global patterns of gene expression. Identifying the "missing" nucleoid proteins of *S. coelicolor* is likely to have important implications for manipulating the production of secondary metabolites such as antibiotics. Candidate NAPs were identified. Several of these are highly conserved in clinically important species such as *Mycobacterium* and in many commercially important species such as *Salinispora* and *Micromonospora* which represent a vital source of novel drugs such as antibiotics, antifungals and anticancer agents.

Abbreviations: GR, Global regulator; NAP, Nucleoid-associated proteins.

* *Corresponding author at:* 219 Chatt building, John Innes Centre, Norwich Research Park, Colney Lane, Norwich, NR4 7UH, United Kingdom. Tel.: +44 1603 452764; fax: +44 1603 450018.
E-mail address: Beth.bradshaw@jic.ac.uk (E. Bradshaw).

1874-3919/$ – see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.jprot.2013.02.033

1. Introduction

The bacterial nucleoid is a non-membrane-bound region in the middle of the cell which contains the chromosome and its associated protein and RNA species. Nucleoid-associated proteins (NAPs) are generally defined as small, highly abundant DNA-binding proteins with low sequence specificity which have widespread effects on both gene expression and chromatin structure [1]. They will typically be present in tens of thousands of monomers per cell and affect the expression of 5–10% of the genome [2–4]. Proteins such as cAMP receptor protein (CRP) may be described as global regulators (GRs) since they have very large regulons [5] but have not been shown to substantially contribute towards nucleoid structure. NAPs are a highly heterogeneous group, representing a variety of protein folds and mechanisms of action. Some are involved in DNA repair (e.g. HU, [6]) or post-transcriptional processes (e.g. StpA, [7]) in addition to roles in nucleoid structure and the global transcriptional programme. Other NAPs such as Dps may have minimal roles in regulating gene expression but are involved in physical and chemical protection of the chromosome, with some homologs (e.g. from *Agrobacterium tumifaciens*) having lost the ability to bind DNA over the course of evolution [8]. In addition to a "core set" of around 10–20 NAPs and GRs the nucleoid also contains many dozens of low-abundance local transcription factors, as well as proteins which interact indirectly with the nucleoid by protein–protein interactions and large molecular machines or enzymes which act on DNA (polymerases, helicases, DNA methyltransferases).

NAPs are commonly involved in regulation of clusters of genes acquired by horizontal gene transfer, such as pathogenicity islands [9]. We hypothesize that NAPs will similarly be important in Actinomycetes for regulation of genes in secondary metabolic clusters, whose expression levels are also specific to certain growth conditions or developmental stages and may be found in genomic islands [10]. The aim of this study was to identify novel NAPs/GRs in *Streptomyces coelicolor* which could subsequently be tested for pleiotropic effects on secondary metabolism, in particular on expression from cryptic clusters for which no product has been observed.

The majority of research on NAPs has been carried out in members of the Enterobacteriaceae but these proteins are often absent, altered or duplicated in other lineages. A number have been identified in the *S. coelicolor* genome by sequence similarity with known NAPs from other species but more remain to be found. The number remaining to be found in *Streptomyces* is not known, but may be higher than in other species as it has a large genome with a particularly high proportion of sigma factors and transcriptional regulators [11]. Two copies of HU are present, the vegetative version of which resembles *Escherichia coli* HU and the spore-associated version of which uses a novel lysine-rich tail to compact and protect DNA specifically in aerial mycelium and spores [12,13]. Two copies of the H-NS-like protein Lsr2 and one copy of sIHF [14], both originally identified in *Mycobacterium*, are present but not yet well characterised. Multiple genes encoding transcription factors annotated as belonging to the Lrp/AsnC family can be found but it is not known whether any of these are truly global regulators while equivalents of Fis and Hfq have never been identified.

Discovery of novel NAPs is not trivial as conventional genetic screens are not possible due to the absence of a readily-screenable phenotype. Many of the canonical NAPs of *E. coli* were originally discovered during studies on bacteriophages, for example IHF (integration host factor), Fis (factor for inversion stimulation) and Hfq (host factor for phage Qβ). Therefore these constitute an important set of proteins which are difficult to identify on the basis of shared structure of common function. Instead, a direct biochemical approach must be taken. We reasoned that the most direct approach would be to survey the protein content of intact nucleoids which are obtained by gentle lysis of cells into a buffer containing spermidine and a physiological salt concentration, then recovered by sucrose gradient centrifugation. Nucleoids isolated in this manner from other species have previously been shown to contain large quantities of NAPs such as HU/H-NS/Fis and also co-sediment with post-transcriptional regulators of gene expression such as the global RNA chaperone Hfq and the translation apparatus [15,16]. A protocol for isolating intact nucleoids from *Streptomyces hygroscopicus* had previously been described by Sarfert et al. [17] but they were unable to identify the proteins present.

In this paper we use label-free LC–MS/MS to identify the abundant proteins found in the nucleoid of the model Actinomycete *S. coelicolor* and a bioinformatic analysis was performed to determine which of these were most likely to be NAPs, generating a list of candidates for further study. It is not possible to separate NAPs from global regulators from these data as effects on chromatin structure are not measured, therefore they will be considered together here.

2. Materials and methods

2.1. Experimental design

In order to study a given subcellular fraction, a comparison is commonly made between its protein composition and that of a whole-cell lysate or dissimilar fraction. However this approach is not necessarily appropriate when studying NAPs because of their great abundance in such a large, loosely-defined cellular structure and because the amounts of unbound NAPs normally present in the cytoplasmic "pool" are unknown. In a similar study by Ohniwa et al. [15], suspected contaminants were subtracted by comparing the composition of the nucleoids against cell envelope and "top matter" fractions (material remaining at the top of a preparative sucrose gradient following ultra-centrifugation), then classifying proteins which were at least three times more abundant in the nucleoid fractions than the others as "csNAPs" (contaminant-subtracted NAPs). While this removed a proportion of the probable contaminants, such as porins and metabolic enzymes, this method excluded the major NAPs H-NS and Hfq. In this study we aimed to identify promising candidate NAPs so we chose not to compare the nucleoid samples with other fractions (a "ruling out" approach) but rather to prioritise the most promising candidates from the lists of detected proteins as they stand (a "ruling in" approach).

Please cite this article as: Bradshaw E, et al, Proteomic survey of the *Streptomyces coelicolor* nucleoid, J Prot (2013), http://dx.doi.org/10.1016/j.jprot.2013.02.033