Protein Expression and Purification 64 (2009) 39–46

Contents lists available at ScienceDirect

Protein Expression and Purification

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

Expression and purification of human full-length N Oct-3, a transcription factor involved in melanoma growth

Béatrice Cabos-Siguier [a,1], Anne-Lise Steunou [a,1], Gérard Joseph [a], Robert Alazard [a], Manuelle Ducoux-Petit [a], Laurence Nieto [a], Bernard Monsarrat [b], Monique Erard [a], Eric Clottes [a,*]

[a] Group "Interactions acides nucléiques/protéines comme cibles pharmacologiques", Université de Toulouse UPS/CNRS-IPBS (Institut de Pharmacologie et Biologie Structurale), 205 route de Narbonne, F-31077 Toulouse, France
[b] Group "Protéomique et spectrométrie de masse des biomolécules", Université de Toulouse UPS/CNRS-IPBS (Institut de Pharmacologie et Biologie Structurale), 205 route de Narbonne, F-31077 Toulouse, France

ARTICLE INFO

Article history:
Received 1 August 2008
and in revised form 9 October 2008
Available online 25 October 2008

Keywords:
POU proteins
Circular dichroïsm
Fluorescence
Footprinting
EMSA

ABSTRACT

This report describes the first purification procedure of the human full-length N Oct-3 protein in amounts suitable for structural studies and proteomic investigations. N Oct-3 is a transcription factor member of the POU protein family. It possesses a large N-terminal transactivation domain and a DNA-binding domain (DBD) which is composed of two subdomains, POUs and POUh, which are joined by a linker peptide. N Oct-3 is a master gene for central nervous system development but also for melanoma progression. Previous structural studies have all been performed using N Oct-3 DBD only. In this study, the full-length N Oct-3 protein was bacterially expressed and purified to homogeneity. The purified protein gave a single band at approximately 53 kDa on SDS–PAGE, while cDNA sequence analysis revealed a calculated molecular mass of 47 kDa confirmed by mass spectroscopy. Size-exclusion chromatography experiments indicated that in solution, full-length N Oct-3 was a monomer. Circular dichroïsm and intrinsic tryptophan fluorescence showed that full-length N Oct-3 was folded, with a significant α-helix content probably located in its DBD. Comparison with the purified N Oct-3 DBD demonstrated that, at least *in vitro*, the affinity of the protein for its DNA targets was similar. This suggests that the transactivation domain of N Oct-3 was not involved in N Oct-3 DNA interaction.

Transcriptional regulators are clustered into large families of proteins defined by highly homologous DNA-binding domains (DBD)[2] that have the ability to bind the same or analogous DNA sequences. POU transcription factors present the particularity of being able to bind, at least *in vitro*, very different sets of DNA target sequences [1]. POU is the acronym for Pit-1, Oct-1, Unc-86, the first three proteins of this structural family.

The ability of the POU proteins to recognise different DNA sequences is due to the presence of two DNA-binding units (POUs for POU specific and POUh for POU homeodomain) connected by a flexible linker [2]. Until now, three distinct types of N Oct-3 DNA targets have been described: "PORE" for Palindromic Oct Recognition Element, "MORE" for More palindromic Oct Recognition Element and "NORE" for N Oct-3 Recognition Element [3]. The molecular structure of their DBD allows POU proteins not only to distinguish a large set of DNA targets, but also to bind different transactivator proteins, depending on the spacing and the positioning adopted by the two subdomains of the POU DBD [3].

N Oct-3 (termed Brn-2 in mice), is a POU protein of about 47 kDa made up of an N-terminus transactivation domain (TAD) and a C-terminus DNA-binding domain (DBD). The protein is expressed during central nervous system development but also in adult brain [4–6]. Although N Oct-3 is not expressed in melanocytes, it has been demonstrated to be involved in melanoma growth [7]. N Oct-3 is directly or indirectly implicated in molecular events leading to a malignant transformation of melanocytes to melanoma, such as β-catenin overexpression [8,9], BRAF activation [10] and microphthalmia-associated transcription factor dysregulation [10].

Most, if not all, of the *in vitro* N Oct-3 DNA interaction studies have been performed so far with the DNA-binding domain of the protein [3,11–13]. N Oct-3 DBD is an 18 kDa domain encountered *in vitro* as a monomer [14]. However, N Oct-3 DBD can be recovered as a dimer interacting with various DNA sequences [13]. The transactivation domain exhibits structural features such as extended stretches of glycine and glutamine residues and clusters of histidine residues (Fig. 1A). With the exception of a fish isoform of Pit-1 [15,16], few attempts have been made to produce and purify functional full-length POU family transcription factors. This paper describes the production of the human

* Corresponding author. Fax: +33 561 175 994.
E-mail address: eric.clottes@ipbs.fr (E. Clottes).
[1] These authors contributed equally to this work.
[2] *Abbreviations used:* DBD, DNA-binding domains; TAD, transactivation domain; GST, glutathione-*S*-transferase; TeV, Tobacco Etch Virus; LB, Luria–Broth; TCEP, tricarboxyethylphosphine.

1046-5928/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2008.10.009

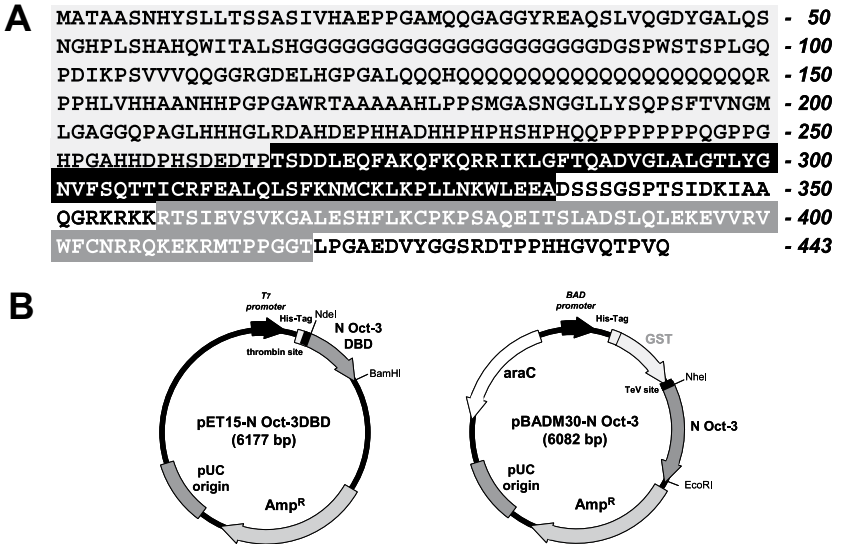

Fig. 1. Protein sequence and plasmid constructions. (A) The human full-length N Oct-3 amino acid sequence is presented. The N-terminal transactivation domain is highlighted in grey. The N Oct-3 DBD is a bipartite domain composed of a POUs (highlighted in black) and a POUh (highlighted in dark grey) subdomain located in the C-terminal part of the protein. The two POU subdomains are connected by a flexible linker. Due to the removing of the His-GST tag by the TeV protease, the N-terminal sequence of the full-length protein was GASMATAA… After thrombin-removing of the histidine tag, the N-terminal end of N Oct-3 DBD was GSHMTPTSDD... (B) Schematic representation of the plasmid constructions used for bacterial expression of His-N Oct-3 DBD and His-GST-N Oct-3 proteins. In both vectors, a sequence coding for a protease cleavage site was present between the sequences encoding the tag and the protein of interest.

full-length N Oct-3 protein in *Escherichia coli,* its purification to homogeneity and the comparison of its DNA recognition properties with those of its DBD alone. The full-length N Oct-3 protein exhibited an aptitude to recognise and bind to DNA targets comparable to the N Oct-3 DBD alone suggesting that, at least *in vitro*, the transactivation domain of the protein was not involved in N Oct-3 DNA-binding properties.

Materials and methods

Expression plasmids

A 468-bp PCR fragment corresponding to the N Oct-3 DBD (amino acids 264–420; GenBank Accession No. P20265) was subcloned into a pET15b plasmid allowing the production of a histidine-tagged protein. The DNA fragment was inserted downstream and in-frame with a hexahistidine tag followed by a sequence encoding a thrombin cleavage site [11,13]. A fragment of 1360-bp encoding the human full-length N Oct-3 protein was amplified by PCR using human cDNA as a template. The amplified cDNA was first inserted into a pET15b vector and did not allow any bacterial expression of the protein. The cDNA was therefore inserted into a pGex plasmid at *Bam*HI and *Xba*I sites. However, only very low protein expression was observed in the *E. coli* strain BL21(DE3)pLysS. Therefore, the 1360-bp fragment was amplified using the following primers: F/5′-cgcgtggagctagcatggcgaccgcagc-3′ (*Nhe*I); R/5′-agct catggagtcgaattcactggacgggc-3′ (*Eco*RI), containing the indicated restriction sites. The amplified cDNA was digested for 1 h at 37 °C by *Nhe*I and *Eco*RI and ligated into a pBADM-30 plasmid in which the original *Nco*I site had been replaced by an *Nhe*I one ("Quick change mutagenesis kit", Stratagene). The fragment was inserted downstream and in-frame with a hexahistidine tag followed by a glutathione-*S*-transferase (GST) coding sequence and a Tobacco Etch Virus (TeV) protease cleavage site. Fidelity of PCR and identity of the constructs were confirmed by sequencing the inserts. *E. coli* strain BL21(DE3)pLysS was transformed with the obtained plasmids named pET15-N Oct-3 DBD and pBADM30-N Oct-3, respectively (Fig. 1B).

Expression and purification of N Oct-3 DBD and full-length N Oct-3 protein

pET15-N Oct-3 DBD-transformed BL21(DE3)pLysS bacteria cells were grown overnight in Luria–Broth (LB) medium supplemented with ampicillin at 100 μg/ml. The cell suspension was then diluted 10-fold in fresh LB and incubated for 15 min at 37 °C. Then, five volumes of M9 medium (16 mM Na_2HPO_4, 9 mM KH_2PO_4, 9 mM NaCl, 18 mM NH_4Cl, 2 mM $MgSO_4$, 1 mM $CaCl_2$, 1% glucose containing ampicillin at 100 μg/ml) were added and bacteria were grown to an $OD_{600\,nm}$ of 0.6. The cell suspension was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 4 h at 30 °C. One hour after IPTG induction, 100 μg/ml rifampicin was added to the cell suspension. Cells were harvested by centrifugation at 8000 rpm (JA10 rotor) for 15 min. The cell pellet from 2 l of culture was suspended in 40 ml of sonication buffer (50 mM Tris–HCl, pH 8.0, 500 mM NaCl, 10 mM β-mercaptoethanol, 1% betaine, 10% glycerol, 0.4 mg/ml of lysozyme). The cells were disrupted by sonication (6 cycles, 60 s each) on ice. Tween-20 was added at a final concentration of 0.2% and the suspension was incubated whilst shaking at 4 °C for 30 min. After centrifugation (1 h at 4 °C and 48,000*g*), the supernatant was passed through a HisTrap 1 ml HP column (GE Healthcare, France) at a 1 ml/min flow rate. The column was washed with 20 volumes of buffer A (50 mM Tris–HCl, pH 8.0, 500 mM NaCl, 2 mM β-mercaptoethanol) followed by two steps of 20 volumes of 10% and 20% of buffer B (buffer A containing 250 mM imidazole). The fusion protein was eluted with 20 volumes of buffer B and concentrated to a final volume of 2 ml using 20 ml Vivaspin concentrators with a cut-off value of 3000 Da (Vivascience, France). The eluted protein fractions were pooled and treated for 2 h at room temperature with 8 U of thrombin (Novagen, France) in order to remove the hexahistidine tag. After incubation, thrombin was inactivated by adding 4 mM EDTA and 2 mM of a freshly prepared phenylmethylsulfonyl fluoride solution. Finally, the protein solution was loaded onto a HiLoad 16/60 Superdex 75 column (GE Healthcare, France) eluted with 50 mM Tris–HCl, pH 8.0, 250 mM NaCl, 5% glycerol, 0.5 mM tricarboxyethylphosphine (TCEP) at a flow rate of 0.5 ml/min.