Protein Expression and Purification 63 (2009) 147–157

Contents lists available at ScienceDirect

Protein Expression and Purification

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

Amalgam, an axon guidance *Drosophila* adhesion protein belonging to the immunoglobulin superfamily: Over-expression, purification and biophysical characterization

Tzviya Zeev-Ben-Mordehai [a,b], Aviv Paz [a,b], Yoav Peleg [a,c], Lilly Toker [b], Sharon G. Wolf [d], Edwin H. Rydberg [e], Joel L. Sussman [a,c], Israel Silman [b,c,*]

[a] *Department of Structural Biology, Weizmann Institute of Science, Rehovoth 76100, Israel*
[b] *Department of Neurobiology, Weizmann Institute of Science, Rehovoth 76100, Israel*
[c] *Israel Structural Proteomics Center, Weizmann Institute of Science, Rehovoth 76100, Israel*
[d] *Chemical Research Services, Weizmann Institute of Science, Rehovoth 76100, Israel*
[e] *Instituto di Ricerche di Biologia Molecolare P. Angeletti, Pomezia, Italy*

ARTICLE INFO

Article history:
Received 4 August 2008
and in revised form 17 September 2008
Available online 8 October 2008

Keywords:
Immunoglobulin
Expression
Secretion
Pichia
Detergent
Multimers

ABSTRACT

Amalgam, a multi-domain member of the immunoglobulin superfamily, possesses homophilic and heterophilic cell adhesion properties. It is required for axon guidance during *Drosophila* development in which it interacts with the extracellular domain of the transmembrane protein, neurotactin, to promote adhesion. Amalgam was heterologously expressed in *Pichia pastoris*, and the secreted protein product, bearing an NH_2-terminal His_6Tag, was purified from the growth medium by metal affinity chromatography. Size exclusion chromatography separated the purified protein into two fractions: a major, multimeric fraction and a minor, dimeric one. Two protocols to reduce the percentage of multimers were tested. In one, protein induction was performed in the presence of the zwitterionic detergent CHAPS, yielding primarily the dimeric form of amalgam. In a second protocol, agitation was gradually reduced during the course of the induction and antifoam was added daily to reduce the air/liquid interfacial foam area. This latter protocol lowered the percentage of multimer 2-fold, compared to constant agitation. Circular dichroism measurements showed that the dimeric fraction had a high β-sheet content, as expected for a protein with an immunoglobulin fold. Dynamic light scattering and sedimentation velocity measurements showed that the multimeric fraction displays a monodisperse distribution, with R_H = 16 nm. When co-expressed together with amalgam the ectodomain of neurotactin copurified with it. Furthermore, both purified fractions of amalgam were shown to interact with *Torpedo californica* acetylcholinesterase, a structural homolog of neurotactin.

Introduction

Multi-domain proteins, which account for over 70% of all eukaryotic proteins, are involved in a wide range of biological processes [1]. While the multiple domains can either be tandem repeats of the same fold or originate from different fold families, bioinformatics analysis suggests that most adjacent domains in multi-domain proteins belong to the latter category [2]. When tandem repeats of the same fold do occur, the two adjacent domains generally have a sequence identity of less than 40%. Interestingly, greater identity has been found to promote co-aggregation [2]. Often, structural characterization of full-length multi-domain proteins is a prerequisite for understanding their function(s); however, such studies require substantial amounts of pure protein. In most cases, this can only be obtained by over-expression in a heterologous system, which can be challenging for a tandem repeat multi-domain protein, since the effective local protein concentration in the vicinity of each domain is high, making co-aggregation an especial risk [3].

Amalgam (Ama)[1] is a secreted cell adhesion protein in the Antennapedia complex of *Drosophila melanogaster* [4], which is

* Corresponding author. Address: Department of Neurobiology, Weizmann Institute of Science, Rehovoth 76100, Israel. Fax: +972 8 934 6017.
E-mail address: israel.silman@weizmann.ac.il (I. Silman).

[1] *Abbreviations used:* Ama, Amalgam; IgSF, immunoglobulin superfamily; S2, Schneider 2; Nrt, neurotactin; *Dm*AChE, *Drosophila melanogaster* acetylcholinesterase; *Tc*AChE, *Torpedo californica* acetylcholinesterase; TEV, tobacco etch virus; MD, minimal dextrose; BMMY, buffered methanol complex medium; BMM, buffered minimal methanol; SEC, size-exclusion chromatography; MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight; Endo F1, endoglycosidase F1; GST, glutathione S-transferase; FPLC, fast performance liquid chromatography; MW, molecular weight; CD, Circular dichroism; DLS, Dynamic light scattering; EM, Electron microscopy.

1046-5928/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2008.09.019

responsible for specifying correct embryonic head and thoracic segmental identity. Immunostaining studies localized it to the extracellular surface of various mesodermal and neuronal cells during embryogenesis [4]. Ama is a member of the immunoglobulin superfamily (IgSF), a large group of proteins involved in many biological processes, including the immune response and various aspects of cell surface recognition [5]. As implied by the name, all proteins in this superfamily contain a conserved Ig domain, which contains ~100 amino acids, and forms a sandwich of two β-sheets that are typically stabilized by a single conserved intrachain disulfide [6]. IgSF members often contain more than one domain; indeed, Ama consists of three Ig domains. These domains are 18–32% identical to each other in pairwise comparisons, and the sequences displaying the greatest identity are centered around the pairs of conserved cysteines that form the intrachain disulfide bonds. Sequence analysis predicts three N-linked glycosylation sites, two in the first and one in the third domain.

The *ama* gene encodes a 333-amino-acid polypeptide, which is predicted to have an NH_2-terminal signal sequence, three Ig domains, and a short COOH-terminal segment. Schneider 2 (S2) cells transfected with the *ama* gene secrete Ama into the medium [7]. Use of this medium in a cell aggregation assay showed that Ama possesses heterophilic adhesion properties, and serves as a ligand for the neuronal adhesion protein, neurotactin (Nrt) [7], a type-II transmembrane glycoprotein which functions as a heterophilic cell adhesion molecule in axon pathfinding and fasciculation [8,9]. Nrt has a 500-amino-acid extracellular domain with 30% and 31% identity, respectively, to *Drosophila melanogaster* acetylcholinesterase (*Dm*AChE) and *Torpedo californica* acetylcholinesterase (*Tc*AChE) [8], and a 324-amino-acid cytoplasmic domain which has been assigned to the class of intrinsically disordered proteins [10]. Ama is required both for primary cultures of embryonic cells to associate and for Nrt-expressing S2 cells to aggregate [7]. Further aggregation assays performed with Ama engineered to be expressed on the plasma membrane of S2 cells showed that Ama also possesses homophilic adhesion properties [11]. Using an Ama-deficient stock it was shown that Ama function is not essential for viability, but that pupae deficient in Ama exhibit defasciculation defects, i.e. defects in the dissociation of axon bundles of the ocellar nerves similar to those found in Nrt mutants [11].

Here we report the over-expression of Ama, a multi-domain protein with three immunoglobulin (Ig) domain repeats, in the methylotrophic yeast, *Pichia pastoris*. Its purification, and biochemical and biophysical characterization are also presented.

Materials and methods

Materials

The *Pichia* expression vector and strain were from Invitrogen (Carlsbad, CA). Yeast extract and peptone were from Pronadisa Conda Laboratories (Madrid, Spain). Yeast nitrogen base was from Difco (Lawrence, KS). Anti-His-tag antibodies, antifoam C, biotin, BSA, sorbitol, KH_2PO_4, K_2HPO_4, NaH_2PO_4, Na_2HPO_4, imidazole and decamethonium bromide were from Sigma (St. Louis, MO). Gentamicin (G418) was from Gibco (Grand Island, NY). Zwittergent 3–12 and CHAPS (3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate) were from Calbiochem (Darmstadt, Germany). The analytical Superdex200 HR 10/30 column, HiLoad 26/60, Superdex200pg, CNBr-activated Sepharose 4B and the Ni Sepharose 6 Fast Flow column were all from General Electric Healthcare (Uppsala, Sweden). GelCode Blue Stain and GelCode glycoprotein staining kit were from Pierce (Rockford, IL). *Tc*AChE was the dimeric (G_2) glycosylphosphatidylinositol-anchored form purified from electric organ tissue of *Torpedo californica* by affinity chromatography on a *m*PTA-Sepharose column subsequent to solubilization with phosphatidylinositol-specific phospholipase C [12]. Anti-Ama antibodies were raised by the Weizmann Institute Biological Services Antibody Unit.

Cloning

The gene fragment encoding full-length *D. melanogaster* Ama (GenBank accession no. CG2198, residues 24–333, without the putative signal peptide) and the gene fragment encoding the ectodomain of *D. melanogaster* Nrt (Nrt-ext; GenBank accession no. CG9704, residues 347–846) were amplified using standard PCR methods from *ama* cDNA and *nrt* cDNA, respectively, both of which were a gift from Dr. Michel Piovant (Université de la Méditerranée, Marseille, France). The primers used for amplification introduced EcoRI and NotI restriction sites at the 5′ and 3′ ends, respectively, and, in the case of Ama, also introduced a 5′-His_6-tag coding sequence, followed by a recognition site for cleavage by the tobacco etch virus (TEV) protease [13] (Table 1). The PCR amplification products of *ama* and *nrt* were cloned in-frame into the corresponding EcoRI and NotI sites on the *Pichia* expression vectors pPIC9K and pPICZα, respectively, containing the α factor secretion signal [14]. The resulting expression vectors were designated pPIC9K/*ama* and *pPICZα/nrt-ext*. Integrity of the coding sequences was confirmed by the Weizmann Institute Biological Services DNA Sequencing Facility.

Establishment of stable Pichia recombinant clones for the expression of Ama

pPIC9K/*ama* was linearized with SacI, and transformed by electroporation according to the Invitrogen protocol. Briefly, electrocompetent *P. pastoris* GS115 cells were freshly prepared before transformation by consecutive washing with ice-cold sterile water and 1 M sorbitol. For transformation, 120 μL of electro-competent GS115 cells were incubated for 5 min with 1 μg of the linearized vector in a 0.2 cm electroporation cuvette. A single 1.5 kV pulse was employed, with a capacitance of 25 μF, a resistance of 200 Ω, and a time constant of 4.5 ms. Following the pulse, 1 mL of 1 M ice-cold sorbitol was added, and the cells were plated on minimal dextrose (MD) plates. Single colonies isolated from the MD plates were directly screened for the *ama* gene by colony PCR, using 5′ and 3′ AOX1 primers (Invitrogen, Carlsbad, CA) (Haaning 1997).

Screening for multiple integration events and for the highest expression clone

Multiple plasmid integration events occur spontaneously in *Pichia* at a low but detectable frequency. The *in vivo* method of screening for such events utilizes hyper-resistance to the antibiotic G418 [14]. The vector used, pPIC9K, contains the bacterial kanamycin gene that confers resistance to G418 in *Pichia*. The level of G418 resistance roughly correlates with the number of kanamycin cassettes integrated. GS115/pPIC9K/*ama* clones were thus collected and re-plated on YPD plates containing increasing concentrations of G418, namely 0.25, 0.5, 1.0, 2.0, 3.0 and 4.0 mg/mL. All clones

Table 1
List of primers.

Ama	Upper	5′-CGTAGAATTC*CATCATCATCATCATCAT*AGCAGCGAAAACCTG TACTTCCAGGGT GCCCCAGTGATCAGCCAGATC
	Lower	5′-CTAATTCGCGGCCGCTTACGACAAGGAGGGCACTGGGAT
Nrt-ext	Upper	5′-CGTAGAATTCCACGAGACTTTGACCTCGCCG
	Lower	5′-CTAATTCGCGGCCGCTTAATCGACGCGCGCATACCGCG