Journal of Proteomics 152 (2017) 131–137

Contents lists available at ScienceDirect

Journal of Proteomics

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

The *Aotus nancymaae* erythrocyte proteome and its importance for biomedical research

D.A. Moreno-Pérez [a,b], R. García-Valiente [c], N. Ibarrola [c], A. Muro [d], M.A. Patarroyo [a,e,*]

[a] Molecular Biology and Immunology Department, Fundación Instituto de Inmunología de Colombia (FIDIC), Carrera 50 No. 26–20, Bogotá, Colombia
[b] PhD Programme in Biomedical and Biological Sciences, Universidad del Rosario, Carrera 24 No. 63C-69, Bogotá, Colombia
[c] Unidad de Proteómica, Instituto de Biología Molecular y Celular del Cáncer -USAL-CSIC, ProteoRed ISCIII, Campus Unamuno, E-37007 Salamanca, Spain
[d] Unidad de Investigación Enfermedades Infecciosas y Tropicales (e-INTRO), Instituto de Investigación Biomédica de Salamanca-Centro de Investigación de Enfermedades Tropicales de la Universidad de Salamanca (IBSAL-CIETUS), Facultad de Farmacia, Universidad de Salamanca, Campus Universitario Miguel de Unamuno s/n, E-37007 Salamanca, Spain
[e] Basic Sciences Department, Universidad del Rosario, Carrera 24 No. 63C-69, Bogotá, Colombia

ARTICLE INFO

Article history:
Received 9 August 2016
Received in revised form 20 October 2016
Accepted 25 October 2016
Available online 29 October 2016

Keywords:
Non-human primate
Experimental model
Aotus
Erythrocyte
Proteome

ABSTRACT

The *Aotus nancymaae* species has been of great importance in researching the biology and pathogenesis of malaria, particularly for studying *Plasmodium* molecules for including them in effective vaccines against such microorganism. In spite of the forgoing, there has been no report to date describing the biology of parasite target cells in primates or their biomedical importance. This study was thus designed to analyse *A. nancymaae* erythrocyte protein composition using MS data collected during a previous study aimed at characterising the *Plasmodium vivax* proteome and published in the pertinent literature. Most peptides identified were similar to those belonging to 1189 *Homo sapiens* molecules; >95% of them had orthologues in New World primates. GO terms revealed a correlation between categories having the greatest amount of proteins and vital cell function. Integral membrane molecules were also identified which could be possible receptors facilitating interaction with *Plasmodium* species. The *A. nancymaae* erythrocyte proteome is described here for the first time, as a starting point for more in-depth/extensive studies. The data reported represents a source of invaluable information for laboratories interested in carrying out basic and applied biomedical investigation studies which involve using this primate.
Significance: An understanding of the proteomics characteristics of *A. nancymaae* erythrocytes represents a fascinating area for research regarding the study of the pathogenesis of malaria since these are the main target for *Plasmodium* invasion. However, and even though *Aotus* is one of the non-human primate models considered most appropriate for biomedical research, knowledge of its proteome, particularly its erythrocytes, remains unknown. According to the above and bearing in mind the lack of information about the *A. nancymaae* species genome and transcriptome, this study involved a search for primate proteins for comparing their MS/MS spectra with the available information for *Homo sapiens*. The great similarity found between the primate's molecules and those for humans supported the use of the monkeys or their cells for continuing assays involved in studying malaria. Integral membrane receptors used by *Plasmodium* for invading cells were also found; this required timely characterisation for evaluating their therapeutic role. The list of erythrocyte protein composition reported here represents a useful source of basic knowledge for advancing biomedical investigation in this field.

1. Introduction

Animal research has been essential for understanding and studying some human diseases, particularly those having the greatest impact around the world, such as malaria. For example, using rodents (BALB/c, C57BL/6, NOD/SCID or humanised strains) has led to obtaining valuable information about this parasite pathogenesis [1,2]. Rodent parasite species (*Plasmodium chabaudi, Plasmodium vinckei, Plasmodium berghei and Plasmodium yoelii*) are different to those infecting humans (*Plasmodium falciparum, Plasmodium vivax, Plasmodium malariae, Plasmodium ovale* and *Plasmodium knowlesi*), therefore having differences regarding their biology and immune response [3]; this means that extrapolating such studies in humans is not always reliable.

Non-human primates represent another model; they have been shown to be the most suitable for studying pathogenesis, immunology and anti-malarial vaccine development, given that they are genetically and immunologically more similar to humans [1]. It is worth noting that some of these primates (mainly *Saimiri* sp. [4,5] and *Aotus* sp. [6,7]) have been widely used in basic and applied biomedical research.

Aotus spp. has been used since its susceptibility to experimental infection by parasites from the genus *Plasmodium* was shown in the

* Corresponding author at: Molecular Biology and Immunology Department, Fundación Instituto de Inmunología de Colombia (FIDIC), Carrera 50 No. 26–20, Bogotá, Colombia.
E-mail addresses: darandmorper@gmail.com (D.A. Moreno-Pérez), rodrigo.garcia.valiente@gmail.com (R. García-Valiente), nibarrola@usal.es (N. Ibarrola), ama@usal.es (A. Muro), mapatarr.fidic@gmail.com (M.A. Patarroyo).

http://dx.doi.org/10.1016/j.jprot.2016.10.018
1874-3919/

1960s [8]; several parasite strains have been adapted since then in this model for studying malaria and developing possible pharmacological treatments or vaccines [9]. Within the genus, the *Aotus nancymaae* species has been infected with different *Plasmodium* strains (*P. falciparum*: Santa Lucia, Indochina I/CDC and Uganda Palo Alto strains; *P. vivax*: Chesson, ONG, Vietnam Palo Alto, Salvador I and Honduran I/CDC; *P. malariae*: Uganda I/CDC), as reported by Collins and his group several years ago [10].

The *Aotus* species has led to an enormous advance regarding preclinical studies highlighting the immunological and protective role of various molecules or parts of them from the *P. falciparum* FVO strain; taking into account that the complex machinery involved in erythrocyte invasion used by this parasite for infecting cells is partly known today, the *Aotus* model has been essential for describing the fundamental basis when identifying vaccine components against this parasite species [11,12]. On the other hand, these primates develop very reproducible infection following experimental infection with the *P. vivax* VCG-1 (Vivax Colombia Guaviare-1) strain, having high levels of parasitaemia (>5%) after 22 passages [13]; this has been of great importance for advancing molecular (MSP-7, Pv38, RAP-1 and RBP-1 between others) [14–17] and immunological (MSP-1_{20} and MSP-1_{14} from MSP-1_{33} fragment) [18,19] characterisation studies of some molecules from the *P. vivax* species and evaluating their usefulness in developing an effective vaccine. These findings highlight the fact that using *A. nancymaae* in combination with the *P. falciparum* FVO or *P. vivax* VCG-1 strains is valuable for screening suitable vaccine candidates for later testing in humans.

In spite of *A. nancymaae* species having led to a promising advance in developing an anti-malarial vaccine, the biology of its erythrocytes still remains unknown (these being vital hosts for *Plasmodium*). Most studies have focused on establishing the similarity between primate and human genes encoding proteins related to the immune response [20–22]. The revolution in omic sciences represented by Baylor College of Medicine's Human Genome Sequencing Centre (BCM-HGSC) has led to the genome and transcriptome of species being studied through the Owl Monkey Genome Project. However, no study describing primate protein composition has been carried out to date. Taking the importance of studying *A. nancymaae* erythrocytes into account, our group was thus interested in obtaining the greatest amount of information possible about the proteome of these cells using data obtained from a previous study by our research group [23] and evaluating it in terms of protein composition and function.

2. Material and methods

2.1. Reanalysing proteome data

Tandem mass spectrometry (MS/MS) data came from a sample consisting of a mixture of mature erythrocytes and reticulocytes (*P. vivax* infected and non-infected ones); two samples had a 50:1 ratio (mature erythrocytes:reticulocytes) whilst the other had a 1.11:1 ratio as it had previously been subjected to a Percoll gradient to enrich infected reticulocytes (preferential invasion target for *P. vivax*) [23]. Data were used for searching for similar peptides, using the human proteome reported in the UniProt database [24]. Mascot [25] and SEQUEST algorithms [26] and Thermo Scientific Proteome Discoverer software were used with astringent search parameters. In brief, the most recent UniProt *H. sapiens* (AUP000005640), *P. vivax* (AUP000008333) and *P. falciparum* (AUP000001450) proteomes were used for compiling a FASTA file containing common non-human contaminants (trypsin, Lys-C and BSA). Thermo's Proteome Discoverer (version 1.4.0.288) was then used for analysing each file in batch and in MudPit [27] for replicates from the same sample; the latter led to identifying low quantity proteins. Results having a <0.01 (high confidence) q-value were filtered using a Mascot Score threshold above 20 and 1.5, 2.0, 2.25, 2.5, 2.75, 3, 3.2, 3.4 for SEQUEST HT (XCorr) for charge states from 1 to 7 and from 3.4 for values >7. An Excel file was generated for each filter showing protein identification details (accession code, description and coverage), including all scores and identified peptides. Redundant UniProt access codes were manually eliminated so that the total list of molecules identified here could be reported.

2.2. Searching orthologous genes in New World monkeys

The search strategy for *H. sapiens* orthologous molecules with New World primates involved using the biological DataBase network (bioDBnet) [28], an online web resource enabling the search for orthologue identifiers in different species. UniProt accession codes from *A. nancymaae-H. sapiens* analysis were used for searching for orthologous molecules in *Callithrix jacchus*, the only species from the primate family phylogenetically related to *Aotus* for which proteome data is available to date. Molecules identified as non-orthologous were analysed again using the OrthoDB database [29].

2.3. Identifying erythrocyte proteins

The proteins identified here were compared to the most extensive profiling of human erythrocyte RNAs published to date [30]. UniProt accession codes were converted into Ensembl gene ID codes with bioDBnet [28] and then compared to 8092 genes expressed as a >0.5 threshold according to an erythrocyte transcriptome study [30].

2.4. Protein annotation according to gene ontology terms

Gene ontology (GO) annotations available in the UniProt database were analysed using the Software Tool for Rapid Annotation of Proteins (STRAP, version 1.5) [31], developed by Boston University School of Medicine's Cardiovascular Proteomics Centre (Boston, MA). The National Institute of Allergy and Infectious Diseases (NIAID) Database for Annotation, Visualization, and Integrated Discovery (DAVID) [32,33] was also used for categorising molecules according to GO terms; stringent parameters were used to ensure statistical significance (thresholds: EASE value = 0.001 and Count = 2).

2.5. Predicting cell membrane molecules

The Red Blood Cell Collection (RBCC) database was used for predicting Surface molecules; RBCC integrates the proteome of human RBC proteins identified to date [34]. The search parameters involved being a highly confident match in both hRBCD and BSc_CH, a blood group or CD marker, experimentally tested in the Sarkadi-lab. The UniProt accession codes for each protein so identified were then manually downloaded for compiling the supplementary table.

3. Results

3.1. Aotus protein prediction

The flow chart in Fig. 1 shows how *A. nancymaae* proteins were identified and analysed. *A. nancymaae* mass spectra were initially imported in Proteome Discoverer software and compared to the information available for the *H. sapiens* proteome by data Mascot and SEQUEST search algorithms; 1084 fully-tryptic and 1052 semi-tryptic molecules were identified using digestion search parameters; 901 were recognised by both parameters, whilst 183 fully-tryptic and 151 semi-tryptic ones were only recognised by one parameter (Supplementary data 1). Primate peptides had great similarity with those from humans, representing 1189 molecules (SD 2). The bioDBnet and OrthoDB tools were used for confirming *A. nancymaae* proteins by comparing the orthology of molecules found in this study with the available information regarding New World monkey proteomes in the UniProt database (Fig. 1). It was found that 95.7% (1138 proteins) of the proteins