Molecular & Biochemical Parasitology 175 (2011) 169–180

Contents lists available at ScienceDirect

Molecular & Biochemical Parasitology

An atypical proprotein convertase in *Giardia lamblia* differentiation

B.J. Davids[a,*], M.A. Gilbert[a,1], Q. Liu[a], D.S. Reiner[a,2], A.J. Smith[a], T. Lauwaet[a], C. Lee[a], A.G. McArthur[b], F.D. Gillin[a]

[a] *Department of Pathology, University of California, San Diego, CA 92103-8416, USA*
[b] *Andrew McArthur Consulting, Hamilton, Ontario, Canada L8S 3P6*

ARTICLE INFO

Article history:
Received 25 November 2009
Received in revised form 29 October 2010
Accepted 5 November 2010
Available online 12 November 2010

Keywords:
Giardia lamblia
Protease
Subtilisin-like proprotein convertase
Encystation
Excystation
Serial analysis of gene expression (SAGE)

ABSTRACT

Proteolytic activity is important in the lifecycles of parasites and their interactions with hosts. Cysteine proteases have been best studied in *Giardia*, but other protease classes have been implicated in growth and/or differentiation. In this study, we employed bioinformatics to reveal the complete set of putative proteases in the *Giardia* genome. We identified 73 peptidase homologs distributed over 5 catalytic classes in the genome. Serial analysis of gene expression of the *G. lamblia* lifecycle found thirteen protease genes with significant transcriptional variation over the lifecycle, with only one serine protease transcript upregulated late in encystation. The translated gene sequence of this encystation-specific transcript was most similar to eukaryotic subtilisin-like proprotein convertases (SPC), although the typical catalytic triad was not identified. Epitope-tagged gSPC protein expressed in *Giardia* under its own promoter was upregulated during encystation with highest expression in cysts and it localized to encystation-specific secretory vesicles (ESV). Total gSPC from encysting cells produced proteolysis in gelatin gels that co-migrated with the epitope-tagged protease in immunoblots. Immuno-purified gSPC also had gelatinase activity. To test whether endogenous gSPC activity is involved in differentiation, trophozoites and cysts were exposed to the specific serine proteinase inhibitor 4-(2-aminoethyl)-benzenesulfonyl fluoride hydrochloride (AEBSF). After 21 h encystation, a significant decrease in ESV was observed with 1 mM AEBSF and by 42 h the number of cysts was significantly reduced, but trophozoite growth was not inhibited. Concurrently, levels of cyst wall proteins 1 and 2, and AU1-tagged gSPC protein itself were decreased. Excystation of *G. muris* cysts was also significantly reduced in the presence of AEBSF. These results support the idea that serine protease activity is essential for *Giardia* encystation and excystation.

1. Introduction

Giardia lamblia is a major cause of waterborne diarrheal disease worldwide [1,2]. *Giardia* cycles between two developmental and morphological forms, the trophozoite and the cyst. Trophozoites colonize the upper small intestine by attaching to epithelial cells where they can cause disease. Infection with *Giardia* does not typically cause inflammation and parasites do not invade the host mucosa or damage intestinal tissues. When trophozoites descend through the small intestine, they generate a rigid extracellular matrix, or cyst wall (CW), that protects them from the external environment (reviewed in [3]). The CW also protects the parasites from gastric acid as they pass through the stomach of a new host before reaching the small intestine where they excyst. Since cysts are the infective stage for mammalian hosts, these differentiations are key to *Giardia*'s survival and its success as a pathogen.

Formation of the resistant extracellular CW begins with intestinal signals that induce the transcriptional upregulation and expression of cyst wall proteins (CWP) [4–7] and the enzymes needed for their processing or post-transcriptional modifications [8–12]. Four structural CWPs have been identified, CWP1, 2, and 3, and high cysteine non-variant cyst protein (HCNCp). All CWPs are transported via encystation-specific secretory vesicles (ESV) and are released at the site of CW assembly. Each of these proteins contains cysteine residues that form extensive intermolecular disulfide bonds during maturation and transport. Additionally, CWP2 and HCNCp in the mature CW are proteolytically processed. It is not known which peptidase(s) cleave HCNCp, but cysteine proteinase

Abbreviations: AEBSF, 4-(2-aminoethyl)-benzenesulfonyl fluoride hydrochloride; CP2, cysteine protease 2; CW, cyst wall; CWG, cyst wall glycopolymer; CWPs, cyst wall proteins; ESVs, encystation-specific secretory vesicles; gSPC, giardial subtilisin-like proprotein convertase; HCNCp, high cysteine non-variant cyst protein; SPC, subtilisin-like proprotein convertase; SAGE, serial analysis of gene expression; WGA, wheat germ agglutinin.

* Corresponding author at: Department of Pathology, University of California, UCSD Medical Center, CTF-C 403, 214 Dickinson Street, San Diego, CA 92103-8416, USA. Tel.: +1 619 543 7729; fax: +1 619 543 6614.
E-mail address: bdavids@ucsd.edu (B.J. Davids).
[1] Current address: Department of Biology, The University of Montana Western, Dillon, MT 59725, USA.
[2] Current address: Signal Transduction Group, Sanford-Burnham Medical Research Institute, La Jolla, CA 92037, USA.

0166-6851/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.molbiopara.2010.11.008

2 (CP2) [9] within the ESV and perhaps lysosomal encystation-specific cysteine protease (ESCP) [8] can cleave CWP-2 into its mature form. The physiological signal(s) that leads to autocatalytic activation of CP2 is not known. Export of mature CWP may be calcium dependent based on evidence provided by Touz and colleagues [13] who were able to show that knockdown of granule-specific calcium-binding protein (gGSP) prevents ESV from releasing CWP. Like CWP2 and HCNCp, gGSP is proteolytically processed during encystation; however, the pathways responsible have not been identified. The mature CW also contains fibrils of poly N-acetylgalactosamine or CW glycopolymer (CWG) [10,14,15] that contribute to the strength of the cyst wall, although its architecture is not well-understood.

Encystation must be tightly regulated to ensure the synthesis, trafficking, and processing of encystation-specific gene products, resulting in their deposition to the nascent CW matrix at the proper time and place [3]. Proteases play an essential role in these processes and may be drug targets to interfere with transmission or replication of parasites [16]. To date, CP2 [9], encystation-specific cysteine protease (ESCP) [8], and the bestatin-binding membrane-associated dipeptidyl [17] peptidases are the best-characterized proteases in *Giardia*. Unlike CP2 and ESCP, inhibition with the protease inhibitor bestatin causes a complete block of transcriptional activation of CWP1 and 2 during encystation. Activity of other protease types has been detected in the life cycle of *Giardia* (e.g. [18–23]), but the enzymes have yet to be identified. Notably, cysteine protease activity is also reported to be critical during excystation of *Giardia* [24].

Cysteine protease activity is the most abundant proteolytic activity detected in *Giardia* as in many other protozoan parasites [25–28]. Through transcriptome analyses of the *Giardia* life cycle [29] we identified a putative prohormone peptidase, gene ID GL50803_2897, with increased expression late in encystation. This protein was recently identified as one of the most abundant wheat germ agglutinin (WGA)-binding glycoproteins in encysting *Giardia* [30]. Sequence comparisons revealed that it is most similar to subtilisin-like proprotein convertases (SPC) of the serine protease family. SPCs are calcium dependent serine endopeptidases that cleave diverse pro-peptides into molecules that are frequently biologically active [31–33]. Cleavage of substrates typically occurs after a pair or series of basic residues. SPCs have conserved structural and functional regions, and the catalytic domain has positionally conserved amino acid residues. These regions are [31,34,35]: the N-terminal signal peptide for transport through the secretory system, the partially conserved pro-domain that assists in intramolecular chaperone folding within the ER and is usually autocatalytically cleaved [36,37], the conserved subtilisin-like catalytic segment with three positional active site residues, and the highly conserved P-domain that stabilizes the catalytic domain and may regulate pH and calcium dependence [34,38]. A variable carboxyl-terminal extension follows the P-domain of SPCs. Several subtilisin type serine proteases have been detected and characterized in protozoan parasites (e.g. [39–44]), however none of those can be classified as a SPC.

SPCs are critical in activating precursor proteins into biologically active forms through regulated proteolysis and are considered novel targets for drug design [32,34]. We found a single SPC-like gene in the *Giardia* degradome (GL50803_2897; gSPC) and we evaluated its role in the life cycle. We demonstrate that AU1-tagged gSPC protein is upregulated in encystation and follows traffic of CWPs to the ESV, but not to the CW. Exposure to the serine protease inhibitor AEBSF resulted in a dramatic block of encystation and excystation, although growth was not affected. The natural substrates for gSPC have not yet been defined. gSPC is catalytically active even though the classic catalytic triad is not readily apparent in this enzyme. Although gSPC was classified as a non-peptidase homolog in MEROPS we detected proteolytic activity. Because of the extensive divergence of many giardial sequences [45], we included other protease homologs considered as "non-peptidase" in our examination of the *Giardia* degradome.

2. Materials and methods

2.1. Giardial cultivation and encystation

Giardia lamblia isolate WB clone C6 (ATCC 50803; [46]) was grown in modified TYI-S33 medium and encysted as described [47] with the modification that bovine bile (Ox gall, unfractionated; Sigma B3883) was substituted for porcine bile at a final concentration of 1.3 mg/ml.

2.2. Detection of G. lamblia peptidase homologs

To achieve maximum sensitivity for the detection of *Giardia* peptidases, the non-redundant library of peptidase units and inhibitor unit proteins (pepunit.lib; 120,595 annotated sequences) from the MEROPS Release 8.4 (April 3, 2009) [48] was used to query the GiardiaDB (Release 1.1, May 12, 2008) database (http://www.giardiadb.org/giardiadb/). The searches were restricted to only the peptidase and inhibitor units library to decrease the possibility of matches to false positives or non-peptidase domains. To assign current GiardiaDB sequence identifier to the existing MEROPS identifiers and to discover any new peptidases, BLASTP (cutoff *E*-value $\leq e^{-6}$) was used on locally installed WU-Blast2 [49] software. To confirm the presence of relevant peptidase functional domains, any peptidase homolog detected using BLASTP was tested locally using the hmmscan algorithm in HMMER3 [50] against the Pfam 23.0 (July 2008, 10,340 families) profile HMM database [51]. Any putative *Giardia* peptidase not found in the current MEROPS release was assigned the Clan, Family or MEROPS identifiers using the highest hit from the MEROPS BLAST search engine. All results were cataloged in an Excel spreadsheet. MEROPS annotation is as defined previously [52,53]. For example, a non-peptidase homolog is defined as similar in tertiary structure to a peptidase family gene product, but may lack one or more critical catalytic residues thought to be essential for peptide hydrolysis. A protein that is labeled as "unassigned peptidases" is structurally catalytic, but its species classification within a specific peptidase family could not be determined at the time of our analysis.

To obtain the nucleotide sequence of the GS gSPC homolog we first used BLASTn to search the Assemblage B GS genome, using the WB gSPC sequence as the query. The BLASTn search was done through the GiardiaDB website (http://giardiadb.org/giardiadb). The top hit was from contig ACGJ01002589, which contained the gSPC sequence from the start codon to position 2217 of the nucleotide sequence (positions 1–2217 of the contig) on the reverse strand. The second hit from contig ACGJ01002292 contained position 1812 to the stop codon of the GS gSPC sequence (positions 35,222–35,733 of the contig). The overlapping sequence matched 100%. The resulting full length GS gSPC gene was 2322 nucleotides, encoding a 773 amino acid protein. BLASTp was used to obtain the Assemblage E P15 gSPC homolog from the GiardiaDB website. The protein sequences of the gSPC homologs from WB, GS and P15 were aligned using CLUSTALW and the BLOSUM weight matrix.

2.3. Gene identification: mRNA expression patterns

SAGE libraries were constructed as described previously [54]. The unique tag sequences were mapped to all available *Giardia* genome assembly contigs (http://www.giardiadb.org/giardiadb/) to determine the identity of expressed genes. As mRNA transcripts