doi:10.1016/j.jmb.2006.01.064 J. Mol. Biol. (2006) 357, 1511–1521

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Solution Structure and Backbone Dynamics of the *Trypanosoma cruzi* Cysteine Protease Inhibitor Chagasin

Didier Salmon[1], Rodolpho do Aido-Machado[1], Anne Diehl[2] Martina Leidert[2], Oliver Schmetzer[3], Ana P. C. de A. Lima[4] Julio Scharfstein[4], Hartmut Oschkinat[2] and José R. Pires[1]*

[1]*Instituto de Bioquímica Médica CCS, Universidade Federal do Rio de Janeiro, Av. Brigadeiro Trompowiski s/n, Rio de Janeiro RJ 21941-590, Brazil*

[2]*Forschungsinstitut für molekulare Pharmakologie Robert-Rössle-Str. 10, 13125 Berlin, Germany*

[3]*Max Delbrück Center for Molecular Medicine, Robert-Rössle-Str. 10, 13125 Berlin, Germany*

[4]*Instituto de Biofísica Carlos Chagas Filho, CCS, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ 21949-900, Brazil*

**Corresponding author*

A *Trypanosoma cruzi* cysteine protease inhibitor, termed chagasin, is the first characterized member of a new family of tight-binding cysteine protease inhibitors identified in several lower eukaryotes and prokaryotes but not present in mammals. In the protozoan parasite *T. cruzi*, chagasin plays a role in parasite differentiation and in mammalian host cell invasion, due to its ability to modulate the endogenous activity of cruzipain, a lysosomal-like cysteine protease. In the present work, we determined the solution structure of chagasin and studied its backbone dynamics by NMR techniques. Structured as a single immunoglobulin-like domain in solution, chagasin exerts its inhibitory activity on cruzipain through conserved residues placed in three loops in the same side of the structure. One of these three loops, L4, predicted to be of variable length among chagasin homologues, is flexible in solution as determined by measurements of ^{15}N relaxation. The biological implications of structural homology between chagasin and other members of the immunoglobulin super-family are discussed.

Keywords: chagasin; *T. cruzi*; cysteine protease inhibitors; NMR structure; dynamics

Introduction

The unicellular protozoan pathogen *Trypanosoma cruzi* is the causative agent of Chagas disease, a human parasitic illness endemic in Central and South America that affects 16–18 million people while leaving over 100 million at risk†. To date, no vaccines are available and drugs for the treatment are inadequate. New perspectives for drug treatment of Chagas disease came from analysis of the structure and function of the major lysosomal *T. cruzi* cysteine protease, cruzipain.[1–4] The recent sequencing of the *T. cruzi* genome revealed that Clan CA cysteine proteases (CP) belonging to the C1 papain-like family are well represented in the parasite genome.[5] Accordingly, members of the C1 CP family in *T. cruzi* include a single copy gene encoding a cathepsin B-like protease and at least 11 polymorphic genes encoding the cathepsin L-like cruzipain. Synthesized as zymogens, pro-cruzipain is converted into active enzyme following proteolytic cleavage of the N-terminal pro-peptide domain. The maturation process, initiated during trafficking through the Golgi[6] is completed by delivery of cruzipain into the lysosomal compartment.[7] Cruzipain is expressed in all parasite life stages and responsible for the major proteolytic activity in the insect-stage of the parasite.[8,9] Multiple lines of evidence indicated that cruzipain function is crucial for parasite infectivity and survival in mammalian host cells.[1,8,10,11]

Abbreviations used: CP, cysteine protease; ICP, inhibitor of cysteine peptidase; Ig, immunoglobulin; HSQC, heteronuclear single quantum coherence; NOE, nuclear Overhauser enhancement; NOESY, NOE spectroscopy; CSP, chemical shift perturbation; TOCSY, total correlation spectroscopy.

E-mail address of the corresponding author: jrmpires@gmx.net

† http://www.who.int/ctd/chagas

The determination of the X-ray structure of the catalytic domain of cruzipain[2] bound to an irreversible inhibitor paved the way for the development of non-toxic drug analogues, some of which proved capable of protecting mice from lethal *T. cruzi* infections.[3]

To date, 48 families of protease inhibitors are described, containing not less than 200 members in mammalian cells (e.g. serpins and cystatins).[12] In trypanosomatids, genes coding for CP inhibitors homologous to the mammalian cystatins are absent.[13] Instead, trypanosomatids, amoeba[14] and some prokaryotes[15,16] express CP inhibitor proteins that share between 20% and 40% sequence identity with chagasin,[15] a single-chained 110 amino acid residue protein originally identified in *T. cruzi*. This family is termed as ICP (inhibitors of cysteine peptidases).[17] Studies of the functional role of chagasin demonstrated that it modulates the endogenous activity of cruzipain, thus indirectly interfering with *T. cruzi* ability to differentiate and/or to invade mammalian host cells.[18] Along similar lines, recent data suggest that chagasin homologues in *Leishmania mexicana* modulate the outcome of host–parasite interactions.[19]

Threading and comparative modelling provided predictions of the structure of chagasin-like proteins.[20] Based on these theoretical studies, it was hypothesized that chagasin-like ICPs adopt an immunoglobulin-like (Ig-like) fold. It was further proposed that some conserved residues in loop regions (L2, L4 and L6; Figure 1) of chagasin could be implicated in the inhibitory interaction with CPs.[15] Nevertheless, due to the low level of sequence identity of chagasin with the templates used in the abovementioned modelling studies (ca 12–17%), it was important to validate these predictions by direct structural data. Here, we determined the solution structure of *T. cruzi* chagasin, studied its backbone dynamics and mapped its interaction with cruzipain by ^{15}N-heteronuclear single quantum coherence (HSQC) based experiments.

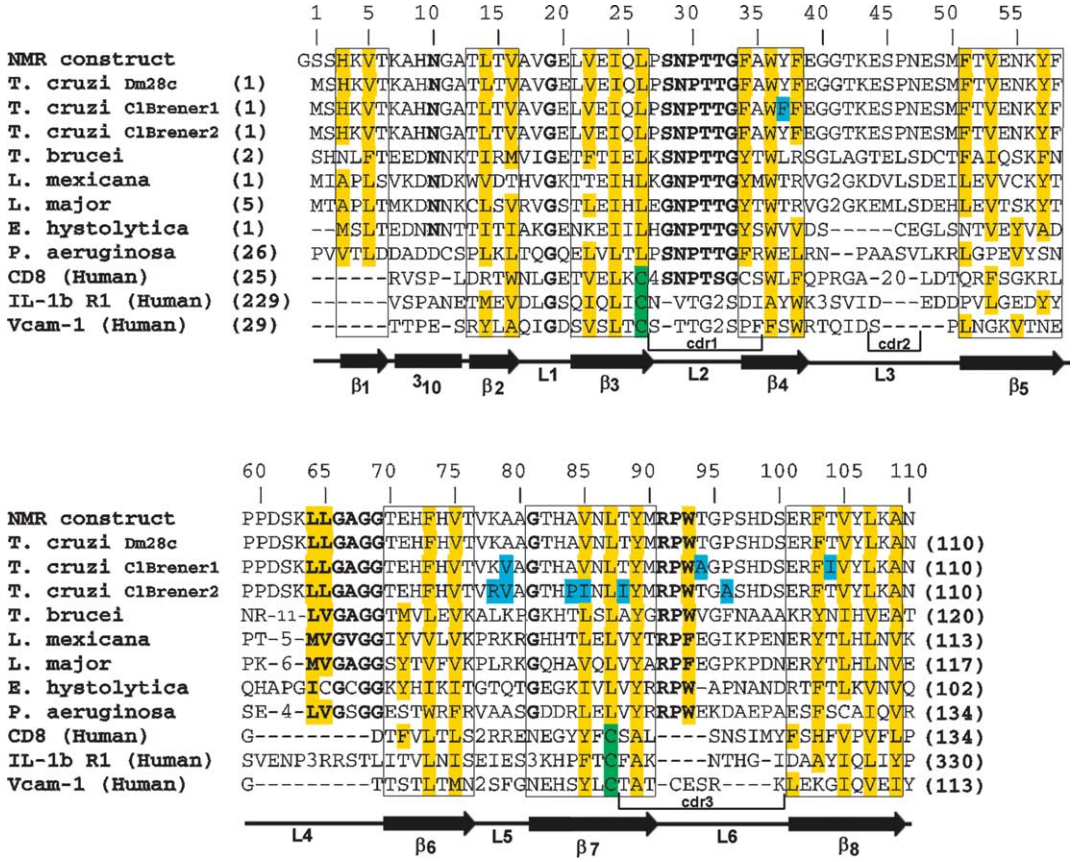

Figure 1. Sequence alignment of chagasin-like cysteine protease inhibitors from different pathogen bacteria and protozoa and of structurally homologous human proteins. Residue numbers refer to *T. cruzi* chagasin. β-Strands and 3_{10}-helix are displayed by arrows and cylinders, respectively, as found in the *T. cruzi* chagasin NMR structure, and the sequences relative to the β-strands are boxed. The positions of the most conserved residues in loops are highlighted in bold and conserved hydrophobic (aromatic and aliphatic) residues are coloured yellow. *T. cruzi* non-conserved residues are highlighted in blue. Complementarity-determining regions (CDR 1–3) in CD8 are indicated. Swiss-Prot accession codes are as follows: chagasin-like ICPs from *T. cruzi* (Dm28c clone, Q966X9; CL Brenner strain, isoform 1, Q4DH32; CL Brenner strain, isoform 2, Q4DY71); *T. brucei* (Q868H0), *L. mexicana* (Q868H1), *L. major* (Q868G9), *E. hystolytica* (Q6KCA4), *P. aeruginosa* (Q9I5G0) and human structural homologue proteins CD8 T-cell surface glycoprotein (P01732), IL-1 R1, interleukine-1 type1 receptor (P14778); Vcam-1, vascular cell adhesion molecule-1 (P19320).