Minireviews:

Trypanosoma cruzi surface mucins with exposed variant epitopes

Guido D. Pollevick, Javier M. Di Noia, Maria L. Salto, Carlos Lima, M. Susana Leguizamón, Rosa M. de Lederkremer and Alberto C.C. Frasch

J. Biol. Chem. published online June 7, 2000

JBC Papers in Press. Published on June 7, 2000 as Manuscript M000253200

Trypanosoma cruzi Surface Mucins With Exposed Variant Epitopes.

Guido D. Pollevick*^, Javier M. Di Noia*^, Maria L. Salto+, Carlos Lima+, M. Susana Leguizamón#, Rosa M. de Lederkremer+ and Alberto C. C. Frasch*[1].

*Instituto de Investigaciones Biotecnológicas - Instituto Tecnológico de Chascomús, (CONICET) - Universidad Nacional de General San Martín, Av. Gral. Paz s/n, INTI, Edificio 24, 1650 – San Martín, Pcia. de Buenos Aires, Argentina.
+CIHIDECAR (CONICET) Departamento de Química Orgánica, Facultad de Ciencias Exactas y Naturales; #Departamento de Microbiología, Facultad de Medicina, Universidad de Buenos Aires.

^ These two authors contributed equally to this work

RUNNING TITLE: gene identification and structure of *T. cruzi* mucins

[1] To whom correspondence should be addressed: Instituto de Investigaciones Biotecnológicas –Av. Gral. Paz s/n, INTI, Edificio 24, 1650 – San Martín, Pcia. de Buenos Aires, Argentina.TE 54-11-4752-9639 Fax: 54-11-4752-0021 E-mail: cfrasch@iib.unsam.edu.ar

SUMMARY

The protozoan parasite *Trypanosoma cruzi*, the agent of Chagas' disease, has a large number of mucin molecules on its surface, whose expression is regulated during the life cycle. These mucins are the main acceptors of sialic acid, a monosaccharide that is required by the parasite to infect and survive in the mammalian host. A large mucin-like gene family named TcMUC of about 500 members has been previously identified in *T. cruzi*. TcMUC can be divided into two subfamilies according to the presence or absence of tandem repeats in the central region of the genes. In this work, *T. cruzi* parasites were transfected with one tagged member of each subfamily. Only the product from the gene with repeats was highly *O*-glycosylated *in vivo*. The *O*-linked oligosaccharides consisted mainly of ß-D-Gal*p*(1→4)GlcNAc and ß-D-Gal*p*(1→4)[ß-D-Gal*p*(1→6)]–D-GlcNAc. The same glycosyl moieties were found in endogenous mucins. The mature product was anchored by glycosylphosphatidylinositol to the plasma membrane and exposed to the medium. Sera from infected mice recognized the recombinant product of one repeats-containing gene thus showing that they are expressed during the infection. TcMUC genes encode a hypervariable (HV) region at the N-terminus. We now show that the HV region is indeed present in the exposed mature N-termini of the mucins because sera from infected hosts recognized peptides having sequences from this region. The results are discussed in comparison with the mucins from the insect stages of the parasite (Di Noia et al. (2000) J. Biol. Chem. 275, 10218-27) that do not have variable regions.

INTRODUCTION

Mucins in vertebrate cells are highly *O*-glycosylated proteins having relevant roles in protection and in cell-cell interactions. The former function is mainly accomplished by high molecular weight epithelial mucins and the latter by the much smaller endothelial and leukocyte mucins, all of them encoded by a few dozens of genes (1,2). In the lower eukaryote *Trypanosoma cruzi*, the protozoan agent of Chagas disease, mucins seem to have essential functions and unusual characteristics. The parasite mucins, located on the surface membrane, are the main acceptors of sialic acid transferred from host's glycoconjugates by the Trypanosomatid-specific enzyme trans-sialidase (3). Surface mucins seem to be involved, after acquiring sialic acid, in the invasion of host cells by the parasite and in the protection against the alternative complement pathway (4-7). Both steps are essential for parasite replication and survival in mammals.

Mucins from *T. cruzi* were first isolated with aqueous-phenol and called bands A, B and C (8,9). These glycoproteins were strongly labeled after intact parasites were treated with galactose oxidase/NaB^3H_4, thus proving their surface location (10). More recently, the structure of the *O*-linked sugars (11-13) was described. The oligosaccharides are linked through GlcNAc, rather than the GalNAc commonly found in vertebrate mucins. Most of the *O*-linked GlcNAc residues are substituted with one to five galactosyl units. The structures of the *O*-linked oligosaccharides are conserved between epimastigotes (the form of the parasite present in the insect vector midgut) and metacyclic trypomastigotes (the infective form in the feces of the insect). However, there are some polymorphism among the strains being the most important one the presence of galactofuranose in the G strain (12,13), whereas in the Y strain the *O*-linked oligosaccharides contain only galactopyranose (11). The mucins present in these forms of the parasite are glycoproteins of about 35-50 kDa whereas those in cell culture derived trypomastigotes are considerably larger (70-200 kDa). The latter molecules have *O*-linked oligosaccharides terminating with α-Gal*p* residues, epitopes that elicit a lytic antibody response in the mammal (14).

The largest gene family encoding mucin-like genes described so far is present in *T. cruzi*, showing diversity among and within strains of the parasite (15,16). Five hundred mucin-like genes per

haploid genome have been estimated to be present in this parasite (16), all members of a family named TcMUC because their overall structure resembled that of mucin genes in higher eukaryotic cells (17). They all have conserved 5' and 3' regions encoding a predicted signal peptide at the N-terminus, and a C-terminal region that includes a putative sequence for a GPI-anchor. Between the conserved ends, there is a variable central domain that characterizes two subfamilies within the mucin-like gene family. In one of these subfamilies, the central region encodes repetitive motifs with the consensus sequence T_8KP_2. In the second subfamily, the central domain codes for a region lacking amino acid repeats, that is rich in codons for Thr, Ser and Pro residues, but that is variant in sequence among the members (16). Due to its organization and amino acid composition, these central regions were predicted to contain the target sites for *O*-glycosylation (17). This assumption is supported by the facts that a repeats-containing member became highly glycosylated when transfected in Vero cells (15) and that the synthetic peptide $KP_2T_8KP_2$ is a good substrate for the enzyme UDP-N-acetylglucosamine:Peptide N-acetylglucosaminyl transferase, that starts *O*-glycosylation in *T. cruzi* (18). However, a direct demonstration that these genes indeed code for the core protein of parasite mucins is still lacking. The small sequence (8-16 amino acids) between the signal peptide and the repeats, likely to code for the mature N-terminus of mucins, is highly variable among members of the family (16). Comparison of this hypervariable (HV) region from 32 cDNA clones showed 22 different variants (16).

A second gene family encoding mucin-type polypeptides (TcSMUG) was recently reported in *T. cruzi*. It is less complex than TcMUC and does not present any variable region. Several evidences suggest that TcSMUG probably encodes mucins from the insect stages of the parasite (19).

Three main questions about the TcMUC family are addressed in this work. First, do any of the mucin-like gene subfamilies code for parasite mucins and what is the structure of these glycoproteins? Second, at which stages are the products expressed? Third, are the HV regions present in the mature products and do they interact with the host immune system? Here we show that a gene representative of those encoding the repetitive T_8KP_2 central domain resulted in a highly *O*-glycosylated product, linked to the surface membrane through a GPI-anchor. Sera reactivity with recombinant proteins indicates that these kinds of products are expressed during the vertebrate infection and that antibodies from natural and experimental infections can recognize the HV regions. Taken together, these results

suggest a structural model for TcMUC-encoded mucins on the surface of the parasite and that the hypervariability in their exposed N-terminus is driven by natural selection.

MATERIALS AND METHODS

Parasites. Epimastigotes of the *T. cruzi* RA strain (20) and the CL-Brener cloned stock (21) were axenically grown with shaking at 28°C in BHT medium (22) supplemented with 10% fetal calf serum (Life Technologies, Gaithersburg, MD.). Cell-derived trypomastigotes, equivalent to the bloodstream forms of the parasite were obtained by centrifugation from the supernatant of infected Vero cells at day 6-7 post infection (19).

Construction of epitope tagged mucin genes and T. cruzi transformation. Since sugars might hinder the access of antibodies to protein epitopes of molecules expected to be heavily *O*-glycosylated like TcMUC products in *T. cruzi*; a tag made up of the amino acid residues 21 to 44 of the HCV core protein (23), which lacks *N*- and *O*-glycosylation sites, was inserted within the sequence predicted to encode the mature *N*-terminus region of *MUC-CA3* and *MUC-RA2* (17). Gene fusions were generated by PCR following standard protocols (24) and using the oligonucleotides indicated in Table 1. The region spanning the repeats and C-terminus domains of *MUC-CA3* was amplified using primers MREcoRI and P2 and cloned into pBlueScript KS II+ (Stratagene, San Diego, CA). Primers MNREcoRI and P2 were used to do alike with the central and C-terminus domains of *MUC-RA2.* Oligonucleotides C21-44A and C21-44B were annealed to create a double-strand insert encoding the HCV core epitope that was cloned in frame with both fragments mentioned above. Finally the corresponding N-termini regions of *MUC-CA3*, amplified with primers MRatg and MRBglII, and *MUC-RA2*, amplified using primers MNRatg and MNRBglII, were added to the corresponding construct in frame with the tag. These tagged versions of *MUC-CA3* and *MUC-RA2* were named *MUC-R* and *MUC-NR*, respectively and subcloned into both the *T. cruzi* expression vector pRIBOTEX (25), kindly provided by Dr. R. Hernández (Universidad Nacional Autónoma de México, México), and the pET-25b(+) (Novagen, Madison, WI) *E. coli* expression vector.

Epimastigotes of *T. cruzi* were transformed by electroporation with the Qiagen-purified (QIAGEN Inc, Chatworth, CA) recombinant pRIBOTEX plasmids, using a slight modification of described procedures (26). The transfected parasites were selected by the pRIBOTEX-encoded G418 resistance using 500 μg/ml Genetecin (Sigma) and used as populations after 40 days of selection. MUC-R and MUC-NR were transfected into epimastigotes of both RA strain and of CL-Brener clone of *T. cruzi***.** A similar result was obtained when each gene was expressed in both strains of the parasite, although the

levels of product expression were very different (results not shown). Transcription of the recombinant genes was analyzed by Northern blot washed at 65 °C in 0.1X SSC, 0.1% SDS as described (19).

Antibodies and Serum Samples. Anti-tag is a polyclonal serum raised in rabbit against the purified recombinant 21 to 44 amino acid residues of HCV core protein (23) expressed in pGEX-1 (Pharmacia Biotech, Uppsala, Sweden) as a fusion with the *S. japonicum* glutathione-S-transferase (GST). This serum was preadsorbed with GST before use to prevent background signals from *T. cruzi* homologous enzymes. Anti-Cross Reacting Determinant (CRD) was obtained from Oxford GlycoSystem, Rosedale, NY. Anti-GST is a polyclonal serum raised in rabbit against purified recombinant GST expressed from the pGEX-1 plasmid vector. Monoclonal antibody 3F5 was a kindly gift from Dr R. Mortara (Escola Paulista de Medicina, Sao Paulo, Brazil).

Sera from 10 rabbits infected with different *T. cruzi* strains and bled at 15, 30 and 60 days post infection (dpi) were kindly given by Dr. D. Sánchez (Instituto de Investigaciones Biotecnológicas, San Martín, Argentina). Sera from 12 infected humans in the chronic stage of the disease, as assessed by their differential reactivity against cruzipain and SAPA *T. cruzi* antigens (27), were used. Mice sera used in Table 2 were from 42 individuals each one belonging to any of 4 different mice strains, infected with one of 7 different *T. cruzi* strains and bled at post-infection times ranging from 8 to 210 days and so being from the chronic or acute stages of the infection (see details in table 2 legend). Sera used in Table 3 were obtained from 13 one-month-old Rockland mice intraperitoneally infected with 10^5 blood trypomastigotes of the Acosta population of *T. cruzi* and bled at three different times post-infection. Sera from non-infected individuals were used as controls (4 from rabbits, 8 from humans and 4 from mice).

Immunoprecipitation and Endoglycosidase H treatment of recombinant mucin. Conditioned culture medium from transfected *T. cruzi* epimastigotes was clarified for 10 min at 10,000 g, and concentrated 8 times by polyethyleneglycol 8000 dialysis. Aliquots were immunoprecipitated using a 1/500 dilution of either anti-tag or anti-GST sera and protein-A-agarose beads (Life Technologies, Gaithersburg, MD), following standard protocols (28). Immunoconjugates were analyzed by Western blot with anti-tag or anti-CRD sera. In another experiment, equally anti-tag immunoprecipitated proteins were eluted from protein A-Sepharose beads in 0.1M sodium acetate pH5.5, 0.1% SDS, by boiling 5 min. The supernatant was treated with 5 mU Endoglycosidase H (New England Biolabs, Beverly, MA) for

16 h at 37ºC or mock treated as control. Reactions were analyzed with anti-tag by Western blot.

Expression and purification of recombinant proteins. a) Protein M76: The TcMUC repeats containing gene *MUC-M76* (GenBank accession number L20809) was cloned in pGEX-1λT vector (Pharmacia Biotech, Uppsala, Sweden). The resulting recombinant protein spanned from the HV region to the conserved C-terminus fused with GST. b) Proteins GST-E13, T15, T18 and NCA: Two complementary oligonucleotides having the sequence encoding the HV region from clones EMUCe-13, EMUCt-15, EMUCt-18 (16) and MUC-CA2 (17) were synthesized. Sense (S) and antisense (A) oligonucleotides (Table 1) were annealed and cloned in pGEX-2T. Their deduced amino acid sequences are: AAEGGGQKQENT, SEEGKQET, TASGQKAEQDT and AESVSQNN, respectively. All the GST fusion proteins were purified identically by affinity chromatography on glutathione-sepharose beads (Sigma) following manufacturers instructions. Their purity was assessed by SDS-PAGE and they were quantified using Bradford reagent (Bio-Rad Laboratories, Richmond, CA).

Gel electrophoresis and Western blots. Gel electrophoresis was performed in 7.5 or 12.5 % polyacrylamide in the presence of 0.1 % SDS (SDS-PAGE). When needed, gels were prepared for fluorography (29), dried and exposed to Kodak X-Omat AR-5 films at -70 °C. For Western blot, parasites were resuspended in lysis buffer (150 mM NaCl, 1% Nonidet P-40, 50 mM Tris-HCl, pH 7.5, in the presence of 50 μM E64 and 1 mM phenylmethylsulfonyl fluoride), at 10^6 parasites/μl and 5 x 10^7 epimastigotes were loaded in each lane. After SDS-PAGE, lysates were transferred to nitrocellulose filters (Life Technologies, Gaithersburg, MD), reacted with the appropriate sera and developed with ^{125}I-Protein A (DuPont NEN, Boston, MA).

Dot-spot immunoassays. One microliter containing 200 ng of each indicated recombinant protein in TBS (50 mM Tris-HCl pH7.6, 150 mM NaCl) was spotted on nitrocellulose filters and air-dried. Filters were processed as in Western blots using sera from different infected animals diluted 1/100 and ^{125}I-Protein A. Recombinant GST-SAPA (30) and epimastigote purified cruzipain (31) antigens were always included as controls of positive infection. Sera from non-infected individuals were included in each assay as background controls. Internal negative controls consisting of GST derived from pGEX-1, 2T and 1λT were also included. Only those signals clearly above those obtained with the non-infected sera and GST were recorded as positive.

Oligosaccharide standards. β-D-Gal*p*(1→4)GlcNAc and β-D-Gal*p*(1→3)GlcNAc were from Sigma. The disaccharides β-D-Gal*f*(1→4)GlcNAc (32), β-D-Gal*f*(1→6)GlcNAc and β-D-

Gal*f*(1→3)GlcNAc (33), β-D-Gal*p*(1→6)GlcNAc (34) and the trisaccharides β-D-Gal*f*(1→4)[β-D-Gal*p*(1→6)]-D-GlcNAc (34) and β-D-Gal*p*(1→3)[β-D-Gal*p*(1→6)]-D-GlcNAc were kindly given by C. Gallo-Rodriguez. Labeled β-D-Gal*p*(1→4)[β-D-Gal*p*(1→6)]-D-GlcNAcol was kindly provided by Drs J. O. Previato and Lucia Mendoça Previato (Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil).

Preparation of unlabeled and labeled alditols. The corresponding alditols were prepared by reduction of 0.2 mmol of disaccharide or trisaccharide in 9:1 methanol:water (10 ml) with $NaBH_4$ (2 mM). The mixture was left overnight at room temperature and the solution was decationized by elution through a column of Bio Rad AG 50 W-X12 (H^+ form) resin. The solvent was evaporated and the boric acid was eliminated by five successive coevaporations with methanol. The purity of the alditol was checked by TLC. For the preparation of the labeled alditols, 100 μg of the sugars were reduced with 500 μCi of NaB^3H_4 (DuPont NEN) in water for 1 h at room temperature, followed by the addition of unlabeled $NaBH_4$ and processing as above.

Thin layer chromatography and high-pH anion exchange chromatography with pulse amperometric detection (HPAEC-PAD). Analysis by TLC and HPAEC-PAD were performed as described (35).

Galactose oxidase/NaB^3H_4. a) *In vivo* labeling: For the galactose oxidase-labeling experiments 10^8 *T. cruzi* cells were resuspended in 0.5 ml PBS containing 22 U galactose oxidase (Sigma) and incubated for 90 min at 30°C with gentle agitation. The cells were washed three times with PBS, suspended in 0.5 ml PBS and 1mCi NaB^3H_4 was added and incubation proceeded for 30 min at room temperature. The reaction was stopped with 0.5 ml cold PBS and the cells were washed with a solution of 1% $NaBH_4$ in NaCl 7% and then twice with NaCl 7%. The final pellet was directly lysed with the electrophoresis sample buffer under reducing conditions and applied on to polyacrylamide gels. b) Labeling of isolated glycoconjugates: A sample of glycoconjugate (5 μg) was suspended in 0.5 ml PBS containing 20 U galactose oxidase (Sigma) and incubated overnight at 37°C. The solution was adjusted to pH 8 with 2.5 M NH_4OH, 1mCi NaB^3H_4 was added and incubation proceeded for 3 h at room temperature. Reduction was completed with $NaBH_4$ for 4 h and dialyzed against water with

gentle agitation. The sample was lyophilized and applied to a column of Octyl-Sepharose CL 4B with 50 μg of non-labeled material.

Purification of mucins from Trypanosoma cruzi a) Mucins from wild type epimastigotes: Mucins were extracted from lyophilized cells (~10^{11}) and purified as described (35). b) Mucins from transfected epimastigotes: The pellet obtained after extraction with water/butanol as in a) was further extracted with 44% aqueous phenol as previously described (8). The aqueous extracts in each case were lyophilized and purified by affinity chromatography, using rabbit IgG anti-tag bound to Sepharose 4B.

PI-PLC treatment: The pellet obtained after extraction with water/butanol was digested with 1U PI-PLC from Bacillus thuringiensis (Oxford GlycoSciences, Inc) in 50 μl Tris-HCl, pH 7.2 containing 0.1% deoxycholate at 37ºC. After centrifugation the supernatant was analyzed by Western blot.

Reductive β-elimination of O-linked oligosaccharides. The glycoproteins were treated with a solution (0.2 ml) of 0.05 M NaOH containing 500 μCi NaB^3H_4 at 37°C. After 1 h, a solution of 0.6 M $NaBH_4$ (0.2 ml) was added and the reaction mixture was kept for 24 h at 37°C, acidified with 1 M acetic acid, *N*-acetylated (36) decationized by passage through AG 50-X_{12} (H^+) (Bio-Rad Laboratories, Richmond, CA) and dried *in vacuo* at room temperature. Boric acid was removed by repeated coevaporations with methanol. The labeled sugar alditols were purified by Biogel P-2 chromatography.

Acetolysis. A sample was acetylated with acetic anhydride:pyridine (1:1) for 30 min at 100°C and dried under a stream of N_2. A mixture of acetic anhydride:acetic acid:concentrated H_2SO_4 (10:10:1) was added and incubation proceeded for 8 h at 37°C. After the addition of 40 μl of pyridine and evaporation (three times with additions of toluene), partition was performed with 1 ml of water and chloroform. The organic phase was evaporated and desacetylated with sodium methoxide at room temperature for 1 h. The mixture was decationized by passage through AG 50-X_{12} (H^+) (Bio-Rad Laboratories, Richmond, CA).

Metabolic labeling of T. cruzi. CL-Brener epimastigotes in the logarithmic growth phase, and cell-derived trypomastigotes from one-week-infected Vero cells monolayers were harvested, washed

twice with PBS and resuspended at 10^8/ml in MEM-Select-amine (Life Technologies, Gaithersburg, MD) Thr and FCS free, supplemented with 1.5 mg/ml glucose. After 30 min incubation at 37ºC, a pulse of (^{14}C) Thr (DuPont NEN, Boston, MA) was added and further incubated for 10 min, after which a 2 h chase was done by adding Thr to 0.5 mM and 2% FCS. Parasites were harvested, washed twice in PBS, resuspended in lysis buffer (50mM Tris-HCl pH7.6, 150 mM NaCl, 1% Triton X-100, 1% Sodium deoxycholate with 1mM phenylmethylsulfonyl fluoride and 0.5 mM Nα-p-tosyl-L-lysine chloromethyl ketone) and submitted to aqueous phenol extraction. The aqueous layer was dialyzed, lyophilized, resuspended in water and analyzed by SDS-PAGE stained by PAS (36) after which fluorography was performed using EN^{3}HANCE (DuPont NEN, Boston, MA).

RESULTS

Expression of two tagged TcMUC-genes in T. cruzi. - Two representative members of the TcMUC gene family, one belonging to subfamily encoding T_8KP_2 repeats (MUC-R) and a second one coding for a non-repetitive central domain but still rich in Ser, Thr, and Pro (MUC-NR) (Fig. 1A), were tagged and cloned in the vector pRIBOTEX (25) and transfected into *T. cruzi* epimastigotes. After selection with G418, the tag probe detected bands in Northern blots of the transfected parasite population (Fig. 1B). The repetitive TcMUC genes like MUC-R show their highest level of mRNA at the trypomastigote stage in wild type parasites (17,37). The non-repetitive genes should be considered individually as they widely differ in sequence and their mRNA level can be different for each member at a given stage (16). *MUC-RA2*, the gene used for MUC-NR construction, mRNA level is similar, although very low, in both epimastigotes and trypomastigotes (unpublished results). Notwithstanding this, the coding regions cloned into pRIBOTEX, devoid of any extragenic region are expressed irrespectively of the expression of their endogenous homologous.

Antibodies directed to the tag detected two broad bands of 85 and 90 kDa for the pRMUC-R repetitive gene expressed in *T. cruzi* (Fig. 2 A lane 2). The product from this same gene expressed in *E. coli* migrated as a band of 35 kDa (Fig. 2 A lane 1), thus suggesting that post-translational modifications were occurring when expressed in *T. cruzi*. Moreover, MUC-R products were probed with sera raised against the untagged recombinant protein fused to GST and expressed in *E. coli*. None was detected after expression in *T. cruzi* whereas the *E. coli* expressed protein was readily detected (not shown), further supporting the presence of extensive post-translational modifications. The product of pRMUC-NR (non-repetitive gene) transfected in *T. cruzi* showed 3 to 4 bands from 35 to 55 kDa after reaction with the anti-tag serum (Fig. 2 A lane 7). The lower molecular weight band run in the same position as the product of this construct expressed in *E. coli*, but the presence of bands migrating at higher positions suggested that this product has also undergone some post-translational modifications when expressed in *T. cruzi*. Wild type epimastigotes used as controls did not produce any signal when probed with the anti-tag serum (Fig. 2 A lane 4).

Only the product from the repetitive gene MUC-R was detectable in the culture medium from transfected epimastigotes as a 90 kDa band (Fig. 2 A lane 3). Furthermore, whereas two products were detected in the parasite lysate, only the upper band was observed in conditioned medium from

transfected parasites. This behavior is analogous to that reported for soluble variable surface glycoprotein, which lacking the lipid anchor, showed a decreased electrophoretic migration (38).

The repetitive member of the TcMUC family is expressed in transfected T. cruzi epimastigotes as a surface mucin - To analyze if the recombinant proteins behaved as mucins, they were isolated from delipidated cells by extraction with water saturated with 1-butanol as previously reported (35). The extract was analyzed by Western blotting, using the anti-tag serum. The product of the repetitive gene (MUC-R), but not the one from the non-repetitive gene (MUC-NR), was detected in the extract, as expected for a highly *O*-glycosylated molecule. Nonetheless, the majority remained in the pellet (data not shown). Therefore, a second extraction with 44% phenol/water was conducted. Only the MUC-R product, but not MUC-NR, was detected in the aqueous phase and after this second step only traces remained in the phenol phase (Fig. 2 B and C). From these experiments, we concluded that only the product of the MUC-gene having the repetitive central domain behaved as a typical mucin and so, this one was further purified and characterized. MUC-NR could still have a mucin domain in the Thr rich degenerated repeats (Fig 1 A) and/or N glycosylated as it has two *NXT* consensus sequences, probably causing the observed shift when expressed in *T. cruzi*. However as our goal was to identify mucin genes, MUC-NR product was not further analyzed in this work.

The second purification step for MUC-R was an affinity column chromatography on immobilized antibodies directed against the HCV tag. The aqueous phase of both, water/butanol and phenol extraction were loaded. As shown in figure 2 D, the antibodies directed to the molecular tag detected the recombinant mucin only in the samples retained in the column. Possible contamination of the samples with epimastigote endogenous mucins was checked using the monoclonal antibody 3F5, that specifically detects the epimastigote endogenous 35/50 kDa mucins (39). This antibody did not detect the 35-50 kDa mucins in the sample retained by the anti-tag antibody column, indicating that the recombinant mucin was not contaminated with endogenous mucins (data not shown).

Parasite surface mucins were labeled *in vivo* using galactose oxidase and NaB^3H_4, in the wild type and pRMUC-R transfected cells, as previously reported (10). Besides the endogenous mucins of 35/50 kDa, a parasite surface glycoprotein of about 85 kDa was slightly and differentially labeled in the transfected parasites (Fig. 3 A). Labeling was not improved by previous treatment with neuraminidase.

The apparent molecular mass observed corresponded with that of the lower band detected by anti-tag in extracts from pMUC-R epimastigotes in Figure 2A. The reason for the different behavior in SDS-PAGE migration of MUC-R and endogenous epimastigote mucins will be discussed later.

Comparison of the carbohydrate structure between endogenous and recombinant mucins of T. cruzi - Analysis of the sugars obtained from endogenous and recombinant MUC-R mucins by reductive β-elimination was performed by HPAEC and TLC. The mucins, obtained from wild type cells labeled by the galactose oxidase / NaB^3H_4 method were purified on octyl-Sepharose as previously reported (35). The recombinant MUC-R mucin was purified using the affinity column. In this case a low amount of radioactivity was recovered; thus, reductive β-elimination was performed in the presence of NaB^3H_4. The products obtained were separated from the radioactive contaminants on a BioGel P-2 column (not shown). From the recombinant mucin an included peak was obtained, eluting between maltotriose and maltose (Fig. 4, lane 1). An excluded peak was also eluted. Analysis by paper electrophoresis of the void radioactivity showed that no negatively charged molecules were present, excluding the presence of sialic acid. This fraction was not further analyzed.

On the other hand, it was interesting to investigate the presence of galactofuranose containing oligosaccharides, as this sugar is important in antibody recognition (40). Galactofuranose was detected in the mucins of one parasite strain (12). Analysis by HPAEC (Dionex) allows the easy identification of galactofuranose containing disaccharitols and trisaccharitols since they elute later than the galactopyranose containing isomeric oligosaccharitols (35). No galactofuranose containing disaccharide or trisaccharide was detected (Fig. 5). The presence of ß-D-Gal*p*(1→4)GlcNAcol was shown by TLC (lanes 1, 2; Fig. 4) and HPAEC (Fig. 5), by comparison with an authentic sample of the disaccharitol (Fig. 4 lane 3 and Fig. 5 standard 3). The sugar alditols obtained from the recombinant mucin were separately eluted from the plate (major compounds in Fig. 4 lane 1) and treated with ß-galactosidase. Both oligosaccharitols were hydrolysed by the enzyme (not shown). The fast compound (Fig. 4 lane 1) was identified as glucitol by HPAEC, utilising the column MA-1, which differentiates alditols (Dionex, Application Note 117). Glucose, originating labelled glucitol, is a common contaminant in the analysis of glycoconjugates. When analysing the oligosaccharitols of the recombinant mucin by HPAEC in the PA-10 column (Fig. 5), besides the ß-D-Gal*p*(1→4)-D-

GlcNAcol, a peak with the elution position of ß-D-Gal*p*(1→4)[ß-D-Gal*p*(1→6)]-D-GlcNAcol was detected. This trisaccharitol was previously obtained from the mucins of the Y strain (11). The higher oligosaccharitols obtained from the wild type mucins (Fig. 4 lane 2) were extracted from the TLC (except the origin) and subjected to partial acetolysis which selectively hydrolysed 1→6 linkages. A main product comigrating with the disaccharitol ß-D-Gal*p*(1→4)-D-GlcNAcol was obtained (not shown). These data suggested that they all have the same disaccharide core.

The MUC-R encoded mucin is anchored by GPI and N-glycosylated. - To find out if the transfected repetitive mucin was GPI-anchored, as predicted from primary sequence, conditioned culture media from pRMUC-R transfected epimastigotes was immunoprecipitated with anti-tag serum and probed with an anti-CRD polyclonal antibody. This antibody reacts with a carbohydrate epitope whose key structural feature is an inositol 1,2-cyclic phosphate moiety which is uniquely found in PI-PLC cleaved GPI glycoproteins (41). As shown in figure 6 A, antibodies directed to the tag and CRD moiety detected the same protein. Furthermore, MUC-R was significantly enriched in supernatants of transfected parasite lysates treated with PI-PLC in comparison with mock treated lysates (Fig 6B). Thus, the recombinant mucin MUC-R is anchored by GPI to the membrane and shed into the medium by the action of some endogenous phospholipase of the parasite.

As three consensus *N*-glycosylation sites were present in the C-terminus of the MUC-R product (Fig. 1), we analyzed if they had an attached oligosaccharide. MUC-R immunoprecipitated with anti-tag from conditioned culture medium was treated with Endoglycosidase H and then analyzed by Western blot with anti-tag serum. This treatment increased the electrophoretic mobility of the tag-containing band by 10 kDa (Fig. 6 C), suggesting that all the *N*-glycosylation sites have attached oligosaccharides.

TcMUC genes encode mucins present in the mammalian stage of the parasite - From the experiments described above, it can be concluded that the product of the MUC-R tagged gene encodes the protein core of a mucin glycoprotein in *T. cruzi*. However, the mobility of the tagged product (about 90kDa) is different from that of the endogenous mucins observed in the epimastigote stage (35/50 kDa). When mucins extracted by aqueous-phenol from (^{14}C)Thr metabolically labeled *T. cruzi* CL-Brener were compared, the radioactive bands present in epimastigotes showed the same pattern than those labeled by galactose oxidase on parasites (see Fig. 3 A and B). On the other hand, the Thr-rich mucins from cell derived trypomastigotes presented a different pattern composed of bands of slower mobility, encompassing the size of the MUC-R product (Fig. 3 B). This, along with the level of the mRNA of TcMUC encoding T_8KP_2 repeats, that is much higher in cell-derived trypomastigotes (37), suggested that MUC-R encoded mucins present in this stage.

To further analyze if the natural products of TcMUC repetitive members, like MUC-R, were present in the stages associated to the vertebrate infection we used a second strategy. A recombinant TcMUC product (M76) containing four T_8KP_2 repeats was expressed in *E. coli* fused to GST, affinity purified and probed in dot spots with sera from human and experimental (rabbits and mice) infections. We employed a panel of sera from animals infected with different parasite

strains and from a wide range of times post infection. The results indicated a different response to M76 depending on the species analyzed. It was recognized by almost all sera from infected mice tested (43/45), independently of the parasite strain or the genetic background of mice used. Sera obtained as early as 25 days post-infection (dpi) to up to 6 months post-infection recognized M76. Only 3 out of 10 rabbit's sera detected M76 and none of the 12 human sera used here did, despite the fact that M76 was isolated by immunological screening of an expression library using human infection sera (42). These results showed that TcMUC repetitive products are present in the stages of the parasite related to the infection (trypomastigote and/or amastigote) as they can elicit an antibody response in mice although their antigenic properties can vary with the host.

Hypervariable regions present in mature mucins are antigenic during the infection - The above mentioned results showed that membrane located mucins, encoded by TcMUC genes, might be present in the trypomastigote stage, that is exposed to the vertebrate's immune system. The repetitive TcMUC genes encoding these molecules are all highly similar, except for the region predicted to be non-glycosylated and at the mature N-terminus, this is the HV region (16). Two mice infected with parasites expressing MUC-R produced antibodies against the HCV tag as assessed by dot spot (not shown), suggesting that the HV region was present in the mature protein and could be immunogenic during the infection. To determine if the natural HV region was indeed expressed and remained in the mature product during natural and experimental infections and give rise to an antibody response, GST fusion proteins having a single randomly chosen HV region (proteins E13, T15, T18 and NCA2, see Materials and Methods) were generated and assayed by dot spot against sera obtained from different infected hosts. Because of the hypervariability there was the possibility that some strain-specific expression or temporal regulation could influence the response to different HV regions, as well as host species and/or haplotipe dependence that could not be known *a priori*. So, in a first set of assays, a heterogeneous sera collection from human and experimental infections (mice and rabbits) covering many of these possibilities were employed and the results are summarized in Table 2. The four HV regions chosen were detected by more than one serum. Human sera were all obained from patients during the chronic stage of the disease and both rabbit positive sera were collected from 60 dpi. This tendency of anti-HV regions antibodies to appear late during the infection is clearly seen in mice infection, with all the positive sera observed at times over 60 dpi. Some sera that detected more than one HV region were all obtained from chronic infections. None of the 8 human, 4 rabbit and 4 mice non-infected sera tested detected any of the HV regions. Furthermore, three GST, from plasmids pGEX-1, pGEX-2T and pGEX-1λT, with different C-termini having similar size than, but unrelated sequences to, HV regions were used as specificity controls and none of them was detected by any sera.

To further study the humoral response against these antigens, sera from a follow up in infected mice were analyzed. Infected mice were bled at 15, 52 and 150 dpi and sera analyzed by dot spot. No reactivity was detected early during the infection, but antibodies against HV regions were detected at 52 dpi and remained at 150 dpi (Table 3). Most of the sera (9/11) detected more than one fusion protein

by day 150. Signals against HV regions were positive but weaker than those obtained against M76 protein. Taken together, these results confirmed the presence of the TcMUC repetitive products during the infection and showed that the HV region was retained in the mature product *in vivo*. They also showed that more than one HV region could be expressed in a parasite population during the infection and that they could be targets of specific antibodies.

DISCUSSION

There are a few examples in protozoan parasites of large and variable gene families coding for proteins with similar function, as is the case in African trypanosomes and *Plasmodium falciparum* (43,44). In both, the large number of genes is directly related to a requirement for parasite survival in the host and sequence variability is the naturally selected character. The N-terminal region of these proteins, the one exposed to the host antibodies, largely differs in sequence among members of these families. The finding that *T. cruzi* contains a large family of mucin-type genes with their putative mature N-terminus having HV regions (16), raised the question of the reason for this diversity. Tempting possibilities could be antigenic variation or some related immunoevasive role, and ligand diversity generation. However, several necessary prerequisites must be demonstrated to sustain any of these hypotheses, including the mucin nature of TcMUC gene products, the persistence of the HV region at the protein level and their exposure to the host antibody response. In this paper, structural and serology data were obtained answering these questions.

Two different representative TcMUC products (MUC-R and MUC-NR) were traced *in vivo* by the use of *T. cruzi* transfection with tagged genes. The MUC-R gene product behaved as a mucin in phenol/water extractions and its *O*-glycosylation was studied, consisting mainly of ß-D-Gal*p*(1→4)GlcNAc and ß-D-Gal*p*(1→4)[ß-D-Gal*p*(1→6)]–D-GlcNAc. This post-translational modification, along with the presence of N-glycosylation and GPI anchor and the finding of a secreted form, are indications that the transfected MUC-R yielded a mucin able to traverse the secretory pathway when expressed under these conditions. The MUC-R product is likely to be highly O-glycosylated in the repeats region since the reactivity with a serum recognizing the recombinant protein expressed in *E. coli* is lost when expressed in *T. cruzi*. All the products from TcMUC repetitive genes, like MUC-R, are highly conserved, differing only in the number of repeats and the sequence of the HV region (16). So they would be expected to be equally processed *in vivo*. Differences in the repeat number may contribute to the heterogeneity in apparent molecular mass of *T. cruzi* mucins.

A second transfected gene which lacks repeats (MUC-NR) is translated into a product that, albeit not behaving as a typical mucin in the phenol/water extraction, is likely to have undergone some post-translational modifications *in vivo* (Fig. 2). The high identity in the leader sequence and GPI-anchor signal of MUC-NR with those of MUC-R suggests that they should also be functional. MUC-NR may still have a mucin domain since its mature C-terminal region is very similar to the one in MUC-R product (see Fig. 1). It has been suggested that the non-repetitive TcMUC genes would not be translated (37). Our results are not conclusive to this respect and the nature of non-repetitive TcMUC products is still undetermined. It is difficult to imagine that *T. cruzi* would have conserved a subfamily

composed of over 200 genes, most of them found as mature mRNAs with an intact open reading frame (16), without any function. In relation to the different behavior between MUC-R and MUC-NR, it is worth noting that the ratio of Thr/Ser codons in the central domain is around 2 in deduced products from non-repeated genes, far from the value around 7 in repeats containing genes. Thr seems to be the main target for *O*-glycosylation in *T. cruzi*, but Ser was also found to be thus modified (3,14). The T_8KP_2 motif is an excellent substrate for an UDP-N-acetylglucosamine: peptide: N-acetylglucosaminyl transferase, the only so far described O-glycosylation initiating enzyme of *T. cruzi* present in both, epimastigotes and cell-derived trypomastigotes (18). MUC-NR, that is rich in Ser (Fig. 1) could be O-glycosylated by a different, still non-characterized, transferase expressed in other developmental stage but not in epimastigotes. In this work, the same di- and trisaccharide structures were obtained from the endogenous mucins present in the epimastigote stage and the MUC-R product, even though evidences are provided that MUC-R encode mucins from the trypomastigote stage (see below) whose O-linked oligosaccharides have a more complex structure (14). This is in agreement with state of the art knowledge about *O*-glycosylation in vertebrate cells, where *O*-linked structures are in general tissue or cell specific depending on the expressed transferases (45) even when some sequence influence could exist due to secondary structures (46).

The evidences indicating that MUC-R like products are expressed in the stages of the vertebrate infection, most probably in the trypomastigote are: 1) The apparent molecular mass of MUC-R expressed in epimastigotes, that is different from endogenous mucins expressed in this stage (35-50 kDa) but within the size-range of mucins expressed in cell-derived trypomastigotes (60-200 kDa); 2) The molecular mass of the mature polypeptide deduced from TcMUC repetitive genes is around 13–18 kDa, in agreement with that calculated for the trypomastigotes apo-mucins (14) and not with the 5-7 kDa determined for epimastigote mucins (3,15); 3) T_8KP_2-containing TcMUC transcripts have their highest level in the trypomastigote stage (17,37); 4) The amino acid sequence of the predicted GPI anchor point of TcMUC genes is consistent with the mature C-termini of trypomastigote mucins, that was recently determined to be *Ile-Asp* with the GPI attached to the latter amino acid (47) 5) A repeats-containing recombinant protein encoded by one of these genes (M76) was recognized by 43 out of 45 sera from infected mice; 6) HV regions present at the mature N-terminus of TcMUC repetitive products were recognized by sera from different infected hosts.

This last mentioned evidence showed not only that the HV region is present in the mature mucin but that it is able to elicit a humoral response. Differences were found in reactivity between the complete protein and HV regions when probed with mice sera. Antibodies directed to M76 showed an early and strong reactivity and remain positive through the chronic phase of the infection. On the other hand, antibodies recognizing HV regions showed weaker signals and were displayed later during the infection, something not unexpected given the difference in size and abundance of each region. While repeats are present in several copies in each protein and, along with the mature C-termini, are conserved in all members of the family, HV regions are small epitopes restricted to a subset of TcMUC proteins. The four HV regions tested, chosen at random out of the many described (16), were detected by more than one sera. As the sera used were from different animals infected with different strains of the parasite, it is possible that the same HV regions could be expressed by different populations or, more likely, that closely related HV regions elicit cross-reacting antibodies. At least one identical HV region exists in the Y (37) and CL-Brener (16) populations of the parasite. Also, several variant but clearly related HV regions were described within one parasite clone (16). Many sera detected more than one HV region, indicating that multiple variants are expressed in the same parasite population during the infection. However, if they appear sequentially or simultaneously or if the distribution is homogeneous or heterogeneous among individual parasites can not be said by now.

O-glycosylation has predictable consequences on the glycoprotein structure. Sugar proximity to the protein backbone and hydrophilic interactions between sugars, cause structural restrictions resulting in protein backbone stiffness (45,48). The rod-like structure of densely *O*-glycosylated domains is due to the first attached sugar and independent of the composition of the oligosaccharide, and its length has been measured for many mucins (49-51). An average length of 0.25 nm per amino acid residue was established (45,48). It is reasonable to assume that O-GlcNAc in *T. cruzi* would have a similar structural effect as O-GalNAc in vertebrate mucins. Therefore, we can estimate the length of a TcMUC repetitive product with the average number of three T_8KP_2 repeats, plus the 37 amino acids that follows up to the likely GPI-addition site and a 10 residues HV region. No secondary structure is predicted for these regions using the available algorithms, supporting their extended or random coil nature. The result would be a rod with a length of about 20 nm, highly glycosylated, GPI-anchored to the membrane and exposing to the medium the hypervariable N-terminus. In agreement with this

prediction, *T. cruzi* trypomastigotes are covered by a dense 20 nm thick coat of mucins (52).

When a protein from a pathogen has a variant region exposed to the medium that is associated to the presence of specific antibodies or lymphocytes, this is usually taken as evidence that the immune system pressure is causing the appearance of variants as a way of evading the immune response. This was described in *Plasmodium spp* (53,54), *Streptococcus* (55), the Hepatitis C Virus (56) and the Simian Immunodeficiency Virus (57) among others. The presence of antibodies against the exposed N-terminus HV region of TcMUC encoded mucins in sera from infected individual suggests that variation is being selected by the immune system of the vertebrate.

We have recently identified a novel mucin-type gene family in *T. cruzi*, named TcSMUG, whose deduced products have all the characteristics of epimastigote apo-mucins (19). Beside the presence of two groups of genes, all members within each group showed no sequence variability. The sequence of one of the groups of TcSMUG deduced proteins coincided with peptide sequences obtained from the 35-50 kDa mucins purified from epimastigotes (Dr I. Almeida, Universidade de Sao Paulo, Brazil, personal communication), suggesting that this new family encodes the mucins from the insect stages of the parasite. Mucins are also the main surface glycoprotein in epimastigotes (52) but variant sequences would not represent an adaptive advantage in an insect host, which has a non-specific immune response. On the other hand, a large number of identical *N*-termini in molecules covering the parasite surface would be dangerous for the trypomastigote under a high affinity specific immune response, and having to survive in the blood long enough to invade different tissues and for being available to the insect vector. We presented herein evidences that the parasite expresses different mucins in this stage, encoded by TcMUC repeats-containing genes, with variant exposed HV regions probably selected to delay the maturation of the immune response. In relation with this, trypomastigote specific surface glycoproteins belonging to the trans-sialidase superfamily (58), displayed variant but related epitopes at defined regions of different family members (59). These variants are expressed simultaneously in a single parasite causing that a potentially protective CD4+ response becomes anergic (60). So, the presence of variant but related epitopes in the surface of the trypomastigote might be a general strategy of *T. cruzi* to evade an early immune response and allow the infection to be established.

REFERENCES

1. Van Klinken, B. J. W., Dekker, J., Buller, H. A., and Einerhand, A. W. C. (1995) *Am J Physiol* **269,** g613-27
2. Van Klinken, B. J. W., Einerhand, A. W. C., Buller, H. A., and Dekker, J. (1998) *Anal Biochem* **265,** 103-16
3. Schenkman, S., Ferguson, M. A., Heise, N., de Almeida, M. L., Mortara, R. A., and Yoshida, N. (1993) *Mol Biochem Parasitol* **59,** 293-303
4. Tomlinson, S., and Raper, J. (1998) *Parasitol Today* **14,** 354-9
5. Schenkman, S., Jiang, M. S., Hart, G. W., and Nussenzweig, V. (1991) *Cell* **65,** 1117-25
6. Schenkman, S., Kurosaki, T., Ravetch, J. V., and Nussenzweig, V. (1992) *J Exp Med* **175,** 1635-41
7. Ruiz, R. d. C., Rigoni, V. L., Gonzalez, J., and Yoshida, N. (1993) *Parasite Immunol* **15,** 121-5
8. Alves, M. J. M., and Colli, W. (1975) *Febs Lett* **52,** 188-90
9. De Lederkremer, R. M., Alves, M. J., Fonseca, G. C., and Colli, W. (1976) *Biochim Biophys Acta* **444,** 85-96
10. Zingales, B., Martin, N. F., de Lederkremer, R. M., and Colli, W. (1982) *FEBS Lett* **142,** 238-42
11. Previato, J. O., Jones, C., Xavier, M. T., Wait, R., Travassos, L. R., Parodi, A. J., and Mendonca-Previato, L. (1995) *J Biol Chem* **270,** 7241-50
12. Previato, J. O., Jones, C., Goncalves, L. P., Wait, R., Travassos, L. R., and Mendonca-Previato, L. (1994) *Biochem J* **301,** 151-9
13. Serrano, A. A., Schenkman, S., Yoshida, N., Mehlert, A., Richardson, J. M., and Ferguson, M. A. (1995) *J Biol Chem* **270,** 27244-53
14. Almeida, I. C., Ferguson, M. A., Schenkman, S., and Travassos, L. R. (1994) *Biochem J* **304,** 793-802
15. Di Noia, J. M., Pollevick, G. D., Xavier, M. T., Previato, J. O., Mendoca-Previato, L., Sanchez, D. O., and Frasch, A. C. (1996) *J Biol Chem* **271,** 32078-83
16. Di Noia, J. M., D'Orso, I., Aslund, L., Sanchez, D. O., and Frasch, A. C. (1998) *J Biol Chem* **273,** 10843-50
17. Di Noia, J. M., Sanchez, D. O., and Frasch, A. C. (1995) *J Biol Chem* **270,** 24146-9
18. Previato, J. O., Sola-Penna, M., Agrellos, O. A., Jones, C., Oeltmann, T., Travassos, L. R., and Mendonca-Previato, L. (1998) *J Biol Chem* **273,** 14982-8
19. Di Noia, J. M., D'Orso, I., Sánchez, D. O., and Frasch, A. C. C. (2000) *J Biol Chem* **275,** 10218-27
20. Gonzalez Cappa, S. M., Bijovsky, A. T., Freilij, H., Muller, L., and Katzin, A. M. (1981) *Medicina* **41,** 119-20
21. Zingales, B., Pereira, M. E., Oliveira, R. P., Almeida, K. A., Umezawa, E. S., Souto, R. P., Vargas, N., Cano, M. I., da Silveira, J. F., Nehme, N. S., Morel, C. M., Brener, Z., and Macedo, A. (1997) *Acta Trop* **68,** 159-73
22. Cazzulo, J. J., Franke de Cazzulo, B. M., Engel, J. C., and Cannata, J. J. (1985) *Mol Biochem Parasitol* **16,** 329-43
23. Ishida, C., Matsumoto, K., Fukada, K., Matsushita, K., Shiraki, H., and Maeda, Y. (1993) *J Clin Microbiol* **31,** 936-40
24. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular cloning: A laboratory manual*,

Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
25. Martinez-Calvillo, S., Lopez, I., and Hernandez, R. (1997) *Gene* **199,** 71-6
26. Hariharan, S., Ajioka, J., and Swindle, J. (1993) *Mol Biochem Parasitol* **57,** 15-30
27. Cazzulo, J. J., and Frasch, A. C. (1992) *Faseb J* **6,** 3259-64
28. Harlow, E., and Lane, D. (1988) *Antibodies. A laboratory manual*, Cold Spring Harbour Laboratory
29. Chamberlain, J. P. (1979) *Anal Biochem* **98,** 132-5
30. Pastini, A. C., Iglesias, S. R., Carricarte, V. C., Guerin, M. E., Sánchez, D. O., and Frasch, A. C. (1994) *Clin Chem* **40,** 1893-7
31. Labriola, C., Sousa, M., and Cazzulo, J. J. (1993) *Biological Res* **26,** 101-7
32. Gallo-Rodriguez, C., Varela, O., and de Lederkremer, R. M. (1996) *J Org Chem* **61,** 1886-9
33. Gallo-Rodriguez, C., Gandolfi, L., and de Lederkremer , R. M. (1999) *Org Lett* **1,** 245-7
34. Gallo-Rodriguez, C., Varela, O., and de Lederkremer, R. M. (1998) *Carbohydr Res* **305,** 163-70
35. Salto, M. L., Gallo-Rodriguez, C., Lima, C., and de Lederkremer, R. M. (2000) *Anal Biochem* **279,** 79-84
36. Fukuda, M., and Kobata, A. (eds) (1993) *Glycobiology. A Practical Approach*, Oxford University Press
37. Freitas-Junior, L. H., Briones, M. R., and Schenkman, S. (1998) *Mol Biochem Parasitol* **93,** 101-14
38. Cardoso de Almeida, M. L., and Turner, M. J. (1983) *Nature* **302,** 349-52
39. Mortara, R. A., da Silva, S., Araguth, M. F., Blanco, S. A., and Yoshida, N. (1992) *Infect Immun* **60,** 4673-8
40. De Arruda, M. V., Colli, W., and Zingales, B. (1989) *Eur J Biochem* **182,** 413-21
41. Mensa-Wilmot, K., Morris, J. C., Al-Qahtani, A., and Englund, P. T. (1995) in *Meth Enzymol* Vol. 250, pp. 641-55
42. Reyes, M. B., Pollevick, G. D., and Frasch, A. C. (1994) *Gene* **140,** 139-40
43. Borst, P., and Rudenko, G. (1994) *Science* **264,** 1872-3
44. Wahlgren, M., Fernandez, V., Chen, Q., Svard, S., and Hagblom, P. (1999) *Cell* **96,** 603-6
45. Van den Steen, P., Rudd, P. M., Dwek, R. A., and Opdenakker, G. (1998) *Cril Rev Biochem Mol Biol* **33,** 151-208
46. Varki, A. (1998) *Trends Cell Biol* **8,** 34-9
47. Almeida, I. C., Camargo, M. M., Procópio, D. O., Silva, L. S., Mehlert, A., Travassos, L. R., Gazzinelli, R. T., and Ferguson, M. A. J. (2000) *EMBO J* **19,** 1476-85
48. Jentoft, N. (1990) *Trends Biochem Sci* **15,** 291-4
49. Shogren, R., Gerken, T. A., and Jentoft, N. (1989) *Biochemistry* **28,** 5525-36
50. Gerken, T. A., Butenhof, K. J., and Shogren, R. (1989) *Biochemistry* **28,** 5536-43
51. Cyster, J. G., Shotton, D. M., and Williams, A. F. (1991) *EMBO J* **10,** 893-902
52. Pereira-Chioccola, V. L., Acosta-Serrano, A., Correia de Almeida, I., Ferguson, M. A., Souto-Padron, T., Rodriguez, M. M., Travassos, L. R., and Schenkman, S. (2000) *J Cell Sci* **113,** 1299-1307
53. Fernandez, V., Hommel, M., Chen, Q., Hagblom, P., and Wahlgren, M. (1999) *J Exp Med* **190,** 1393-1403
54. Plebanski, M., Flanagan, K. L., Lee, E. A. M., Reece, W. H. H., Hart, K., Gelder, C., Gillespie, G., Pinder, M., and Hill, A. V. S. (1999) *Immunity* **10,** 651-60
55. Stockbauer, K. E., Grigsby, D., Pan, X., Fu, Y., Perea Mejia, L. M., Cravioto, A., and Musser, J.

M. (1998) *Proc Natl Acad Sci USA* **95,** 3128-33
56. Weiner, A. J., Geysen, H. M., Christopherson, C., Hall, J. E., Mason, T. J., Saracco, G., Bonino, H., Crawford, K., Marion, C. D., Crawford, K. A., Brunetto, M., Barr, P. J., Miyamura, T., McHutchinson, J., and Houghton, M. (1992) *Proc Natl Acad Sci USA* **89**(3468-3472)
57. Evans, D. T., O´Connor, D. H., Jing, P., Dzuris, J. L., Sidney, J., da Silva, J., Allen, T. A., Horton, H., Venham, J. E., Rudersdorf, R. A., Vogel, T., Pauza, C. D., Bontrop, R. E., DeMars, R., Sette, A., Hughes, A. L., and Watkins, D. I. (1999) *Nat Med* **5,** 1270-6
58. Campetella, O., Sánchez, D. O., Cazzulo, J. J., and Frasch, A. C. C. (1992) *Parasitology Today* **8,** 378-81
59. Kahn, S. J., and Wleklinski, M. (1997) *J Immunol* **159,** 4444-51
60. Millar, A. E., Wleklinski-Lee, M., and Kahn, S. J. (1999) *J immunol* **162,** 6092-9

Footnotes

Abbreviations used are CRD, cross-reacting determinant, dpi, days post infection, DTT, dithiothreitol, GPI, glycophosphatidylinositol, GST, glutathione-S-transferase, HV, hypervariable, LPPG, lipopeptidophosphoglycan.

Acknowledgments

Drs J.J. Cazzulo, A. Parodi and O. Campetella are acknowledged for critically reading the manuscript. We thank Dr D. Sánchez for helpful discussions and suggestions and for sera from infected rabbits, Dr R. Mortara for providing us the 3F5 antibody, Dr. R. Hernández for pRIBOTEX plasmid, and Dr I.C. Almeida for letting us include unpublished results. We thank L. Sferco and B.Franke de Cazzulo for technical assistance and I. D´Orso and C. Ascencio for their help in some electrophoresis assays.

This work was supported by grants from the Agencia Nacional de Promoción Científica y Tecnológica, Argentina, the Swedish Agency for Research Cooperation with Developing Countries (SAREC-SIDA), the International Atomic Energy Agency, the Consejo Nacional de Investigaciones Científicas y Técnicas and the University of Buenos Aires. M.L.S. is a fellow from the University of Buenos Aires. J.D.N. is a fellow, C.L. is a principal professional and R.M. de L., G.D.P, M.S.L and A.C.C.F. are researchers from the CONICET, Argentina. The research of A.C.C.F. was supported in part by an International Research Scholar Grant from the Howard Hughes Medical Institute.

FIGURE LEGENDS

Figure 1 - A, **schematic representation of proteins derived from the transfected repetitive (MUC-R) and non-repeated (MUC-NR) MUC genes of *T. cruzi*.** The structure of the two deduced products is represented with their different regions indicated. The dashed line indicates a gap for best alignment. Degenerate repeats are indicated as black boxes. The deduced sequence of the central and Thr-rich regions is indicated above or below the corresponding representative box. The triangles indicate *N*-glycosylation consensus sequences. HV, hypervariable region. The percentage of similarity is indicated for each region below the scheme. Calculations were done using the program DNASTAR (DNASTAR Inc, Madison, WI). The drawing is not in scale. B, **Northern blot analysis of transfected populations**. Total RNA from transfected (MUC-R and MUC-NR transfected with pMUC-R and pMUC-NR respectively) and wild type (wt) *T. cruzi* populations were hybridized with an antisense oligonucleotide of the tag.

Figure. 2 - **Expression and purification of tagged products from transfected *T. cruzi* epimastigotes.** *A*, Western blot analysis of: lane 1, total lysate of *E. coli* expressing MUC-R; lane 2, total lysate of $5x10^7$ pRMUC-R transfected parasites; lane 3, medium from pRMUC-R transfected parasites; lane 4, total lysate of $5x10^7$ wild type (wt) parasites; lane 5, medium from wild type (wt) parasites; lane 6, total lysate of *E. coli* expressing MUC-NR; lane 7, total lysate of $5x10^7$ pRMUC-NR transfected parasites; lane 8, medium from pRMUC-NR transfected parasites. Molecular mass markers are indicated at the left. *B*, phenol/water extraction of parasites expressing MUC-R. Phenolic (Ph) and aqueous (Aq) phases were loaded on SDS-PAGE, transferred to nylon membrane and, after probing with the anti-tag serum, revealed with ^{125}I-Protein A. *C*, identical to B for MUC-NR expressing parasites. *D*, affinity purification of MUC-R product. The aqueous phase of phenolic extraction was applied to an anti-tag affinity column. The eluted (lane 1) and percolated material (lane 2) were analyzed by Western blot with the anti-tag serum. The calculated molecular mass of the observed bands is indicated in panels *B, C* and *D*.

Figure 3 - **Labeling of pRMUC-R transfected and wild type *T. cruzi* epimastigotes**. A, intact transfected (MUC-R) or wild type (wt) epimastigotes were surface labeled by galactose oxidase/NaB^{3}H$_4$. Total lysates were separated by 10 % SDS/PAGE and revealed by fluorography. NA, *Clostridium*

perfringens neuraminidase. B, *T. cruzi* epimastigotes (E) and trypomastigotes (T) were metabolically labeled with (^{14}C) Thr and submitted to phenol /water extraction. Extracted material was separated by 10 % SDS/PAGE and revealed by fluorography. The molecular mass of standards is indicated in the right.

Figure 4 - **Thin layer chromatography of *O*-linked sugar alditols obtained from the recombinant and wild type mucins of *T. cruzi* epimastigotes.** Lane 1, *O*-linked sugars released from the recombinant mucins by β-elimination in the presence of NaB^3H_4; lane 2, *O*-linked sugars released from galactose oxidase/ NaB^3H_4 labeled mucins; lane 3, authentic sample of [1-^{3}H]-β-D-Gal*p*(1→4)GlcNAcol; lane 4, [1-^{3}H]- GlcNAcol. The products were analysed by TLC with propanol:NH_3:water (7:1:2), double development and fluorography. The migration of nonradioactive standards are shown in the right. **I**, β-D-Gal*f*(1→3)GlcNAcol; **II**, glucitol; **III**, β-D-Gal*p*(1→3)GlcNAcol; **IV**, β-D-Gal*p*(1→4)GlcNAcol; **V**, β-D-Gal*p*(1→6)GlcNAcol. Origin is indicated by an arrow.

Figure 5 - **HPAEC-PAD analysis of the radioactive *O*-linked sugars obtained by NaB^3H_4 reductive β-elimination of the mucins from MUC-R transfected epimastigotes.** The sample was desalted on BioGel P-2 and chromatographed on a CarboPac PA-10 column under the conditions indicated under Materials and Methods. The numbers correspond to the sugar alditol standards. **1**, Glucitol; **2**, β-D-Gal*p*(1→3)GlcNAcol; **3**, β-D-Gal*p*(1→4)GlcNAcol; **4**, β-D-Gal*p*(1→3)[β-D-Gal*p*(1→6)]GlcNAcol; **5**, β-D-Gal*p*(1→4)[β-D-Gal*p*(1→6)]GlcNAcol; **6**, β-D-Gal*f*(1→3)GlcNAcol; **7**, β-D-Gal*f*(1→4)GlcNAcol; **8**, β-D-Gal*p*(1→4)[β-D-Gal*f*(1→6)]GlcNAcol.

Figure 6 - **The MUC-R mucin is anchored by GPI and N-glycosylated.** Immunoprecipitation of culture medium from MUC-R transfected epimastigotes of *T. cruzi*. A, the immunoprecipitated material was revealed with ^{125}I-Protein A, after SDS-PAGE separation and Western blotting. Left panel, probed with anti-tag serum; right panel, probed with anti-CRD serum. Lanes 1, immunoprecipitated with anti-tag serum; lanes 2, immunoprecipitated with anti-GST control serum. B, MUC-R transfected parasite extracts were incubated in the absence (lane 1) of presence (lane 2) of *B. thuringiensis* PI-PLC and the supernatants probed with anti-tag after Western blot. Lysate from non-transfected parasites treated with PI-PLC was included as a further control (lane 3). C, anti-tag

immunoprecipitated proteins were eluted from protein A-Sepharose beads, treated with Endoglycosidase H and probed with anti-tag serum after Western blot. Apparent molecular weight calculated from the markers migration is indicated by arrows.

Table 1 – **Oligonucleotides used**. Sequences from the template are in uppercase; added sequences are in lowercase with restriction enzyme sites underlined. The gene or clone used as template is indicated.

Table 2. **Detection of *TcMUC* encoded hypervariable regions by *T. cruzi* infection sera**. A panel of sera from humans, mice and rabbits infected with different *T. cruzi* strains and bled at different times post-infection were used. Each positive represents one serum from one distinct infected individual. Mice positive sera were divided in two groups: <60 dpi represents 6 sera from 8 dpi, 3 from 30 dpi, 4 from 60 dpi; >60 dpi represents 4 sera from 90 dpi, 3 from 120 dpi, 12 from 150 dpi and 7 from 210 dpi. Sera were reacted with dot spots of recombinant proteins having HV regions from different TcMUC genes (E13, T15, T18 and NCA2 have a single HV region fused to GST). The ratio of positive sera to total assayed sera is indicated for each antigen. nd, non determined.

Table 3. **Follow up of antibodies against hypervariable regions of TcMUC proteins in infected mice**. Mice were infected with *T. cruzi* and bled at the indicated days post infection (dpi). Sera were used to probe dot spots of recombinant proteins as in Table 2. Positive signals for each antigen is indicated. n indicates the number of sera assayed at each time post infection. nd, non determined, dpi, days post-infection. The sum of the positives is greater than the total sera tested because most of the sera reacted with more than one antigen by day 150.

Table 1

name	sequence	template
MRatg	tgcgcggccgggatccATGAATACACTCACGATGATG	*MUC-CA3*
MRBglII	gaagatctACACTTTCTGCCACGCACACGGA	*MUC-CA3*
MREcoRI	ccggaattcAGTCAGAACAATACCACTACG	*MUC-CA3*
P2	cccaagctttctagaACATCGGACCACGGTAGAAGTA	*MUC-CA3* *MUC-RA2*
MNRatg	tgcgcggccgggatccATGACGACGTGCCGTCTGCTG	*MUC-RA2*
MNRBglII	gaagatctACGGTCGCCATCACGCACAC	*MUC-RA2*
MNREcoRI	ccggaattcGTCGCGGAAGAAGATGGCCCC	*MUC-RA2*
C21-44A	gatcccGATGTGAAATTTCCGGGCGGCCAGATTGTG GGCGGCGTGTATCTGCTGCCGCGTCTGGGCCCGC GTCTGGGCg	HCV Core
C21-44B	aattcGCCCAGACGCGGGCCCAGACGCGGCAGCAGA TACACGCCGCCCACAATCTGGCCGCCCGGAAATTT CACATCgg	HCV Core
E13S	gatccCCCTGCGGCAGAAGGTGGTGGTCAAAAGCAAGAAAATACAg	EMUCe-13
E13A	aattcTGTATTTTCTTGCTTTTGACCACCACCTTCTGCCGCAGGg	EMUCe-13
T15S	gatccCCTAGCGAAGAGGGTAAGCAGGAAACAg	EMUCt-15
T15A	aattcTGTTTCCTGCTTACCCTCTTCGCTAGGg	EMUCt-15
T18S	gatccCCTACAGCAAGTGGTCAGAAGGCTGAGCAGGACACAg	EMUCt-18
T18A	aattcTGGTGTCCTGCTCAGCCTTCTGACCACTTGCTGTAGGg	EMUCt-18
NCA2S	gatccCCGGCAGAAAGTGTTAGTCAGAACAATACCTGAg	*MUC-CA2*
NCA2A	aattcTCAGGTATTGTTCTGACTAACACTTTCTGCCGGg.	*MUC-CA2*

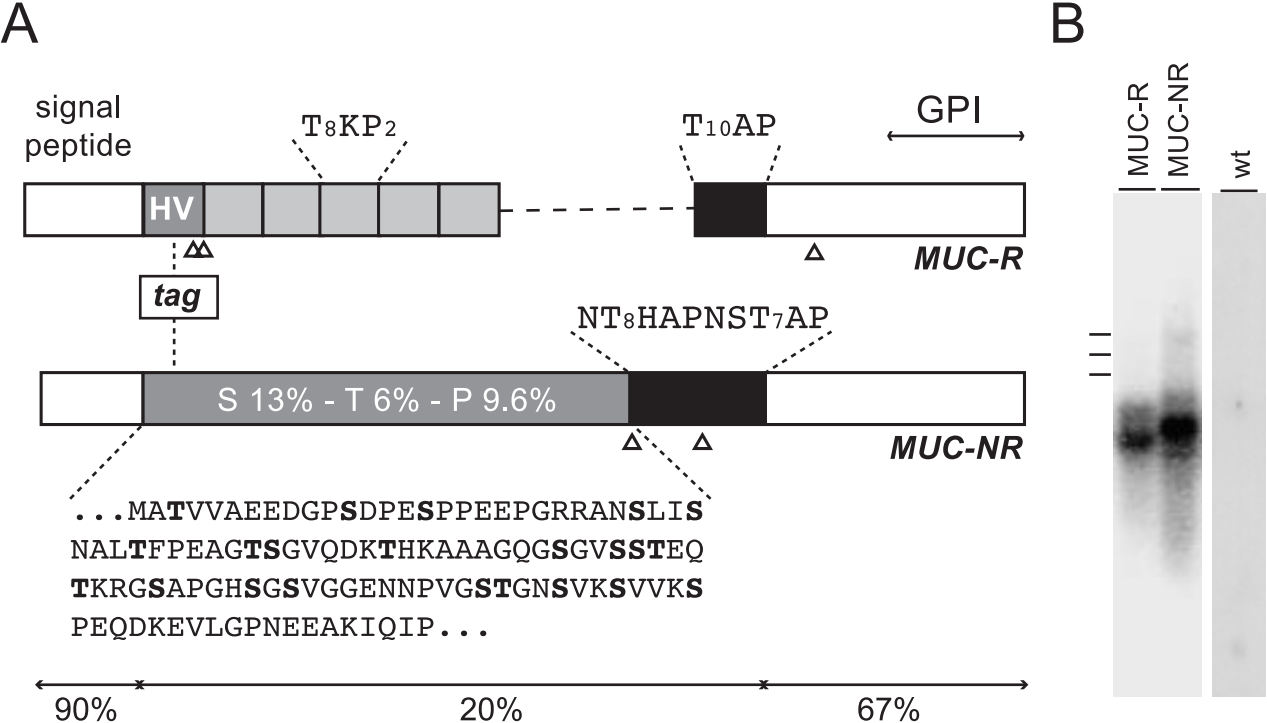
Table 2

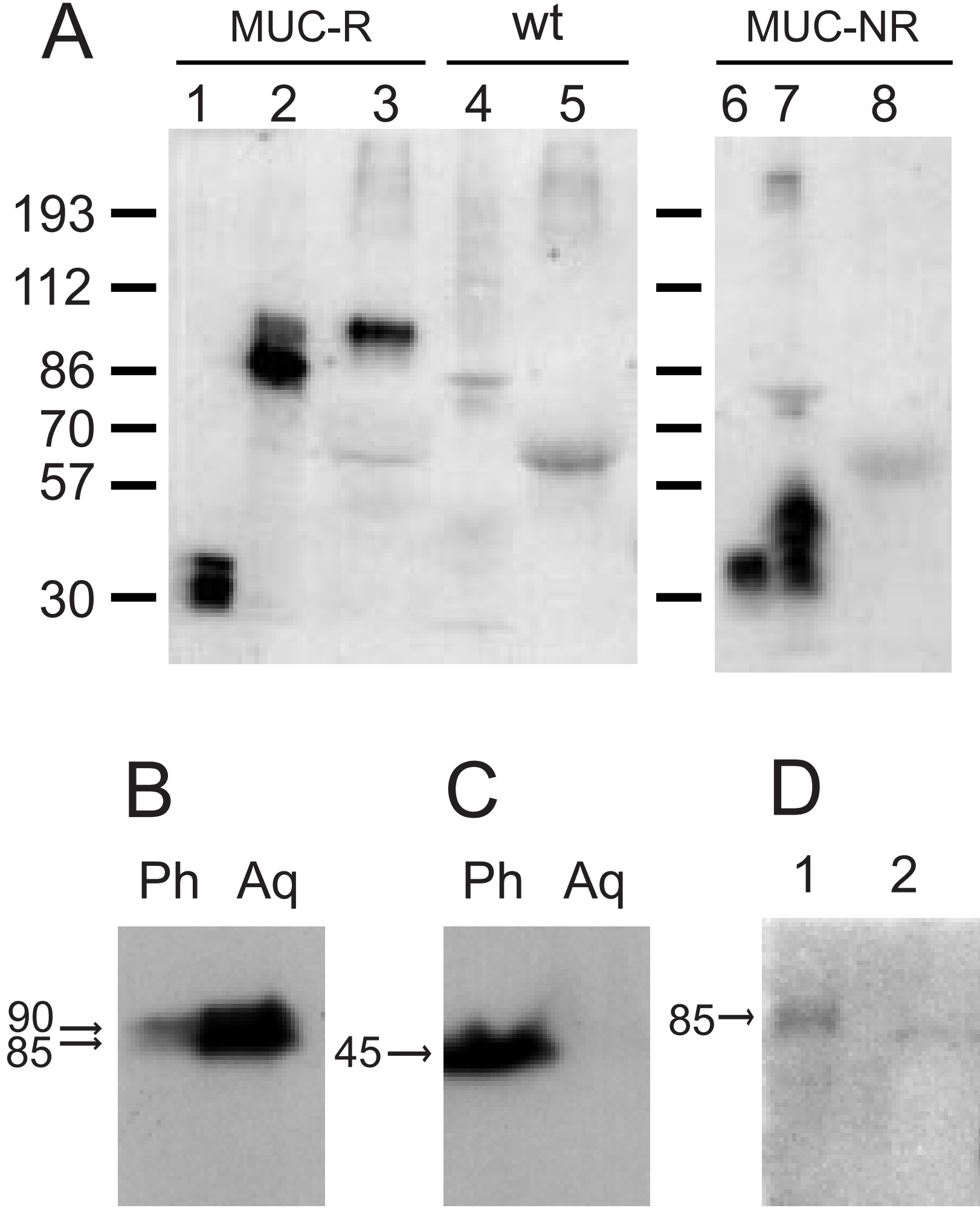
	positive sera/assayed sera			
	human	mouse		rabbit
		< 60 dpi	> 60 dpi	
E13	0/12	0/16	7/21	0/10
T15	7/12	0/16	3/21	1/10
T18	0/12	0/16	5/21	1/10
NCA2	2/12	nd	0/3	nd

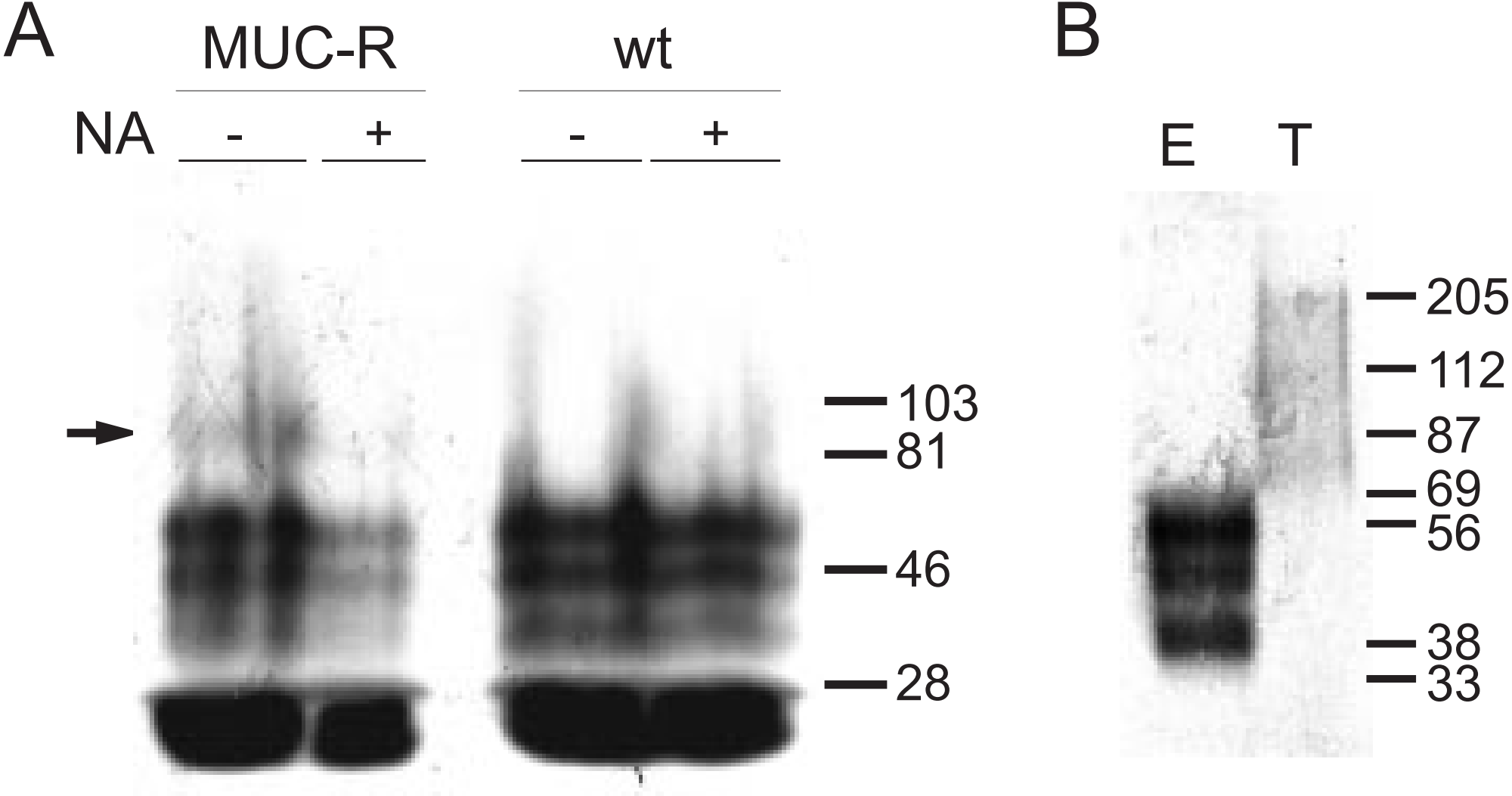
Table 3

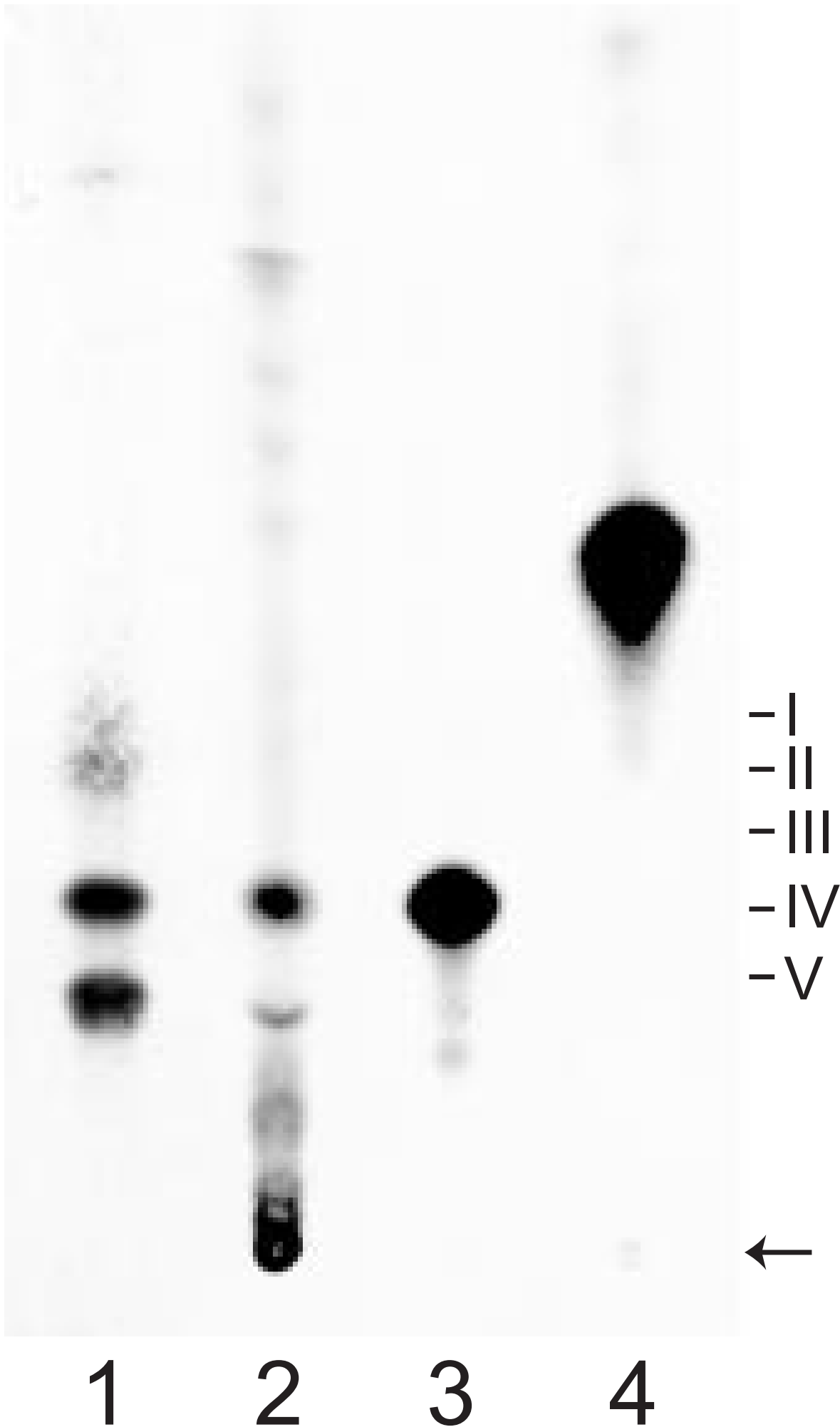
positive sera

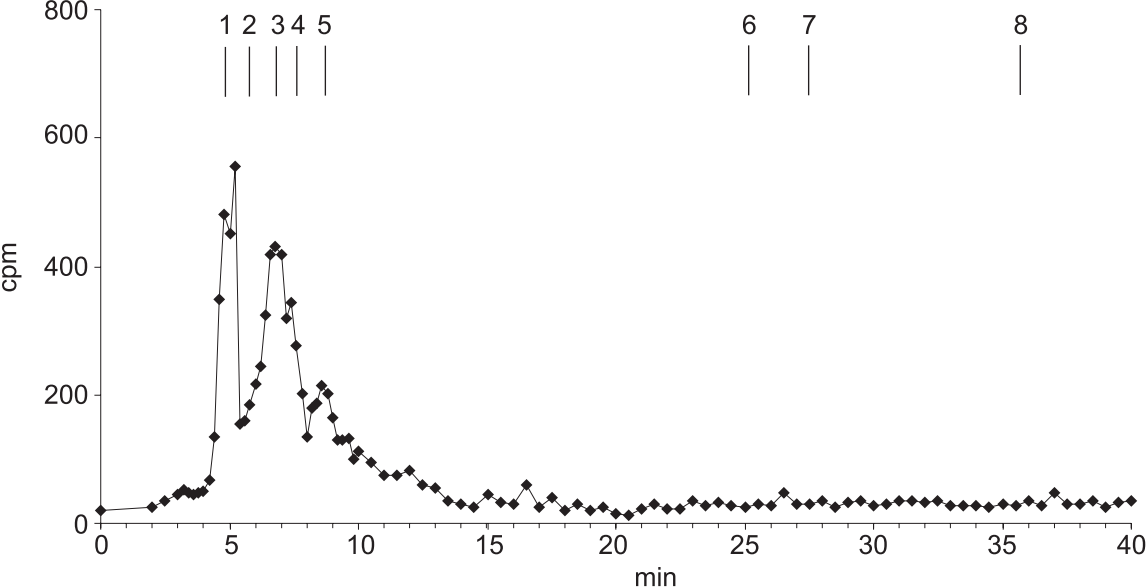
	dpi	15 (n=13)	52 (n=13)	150 (n=11)
antigens	E13	0	4	9
	T15	0	2	11
	T18	0	3	nd
	NCA	nd	nd	2

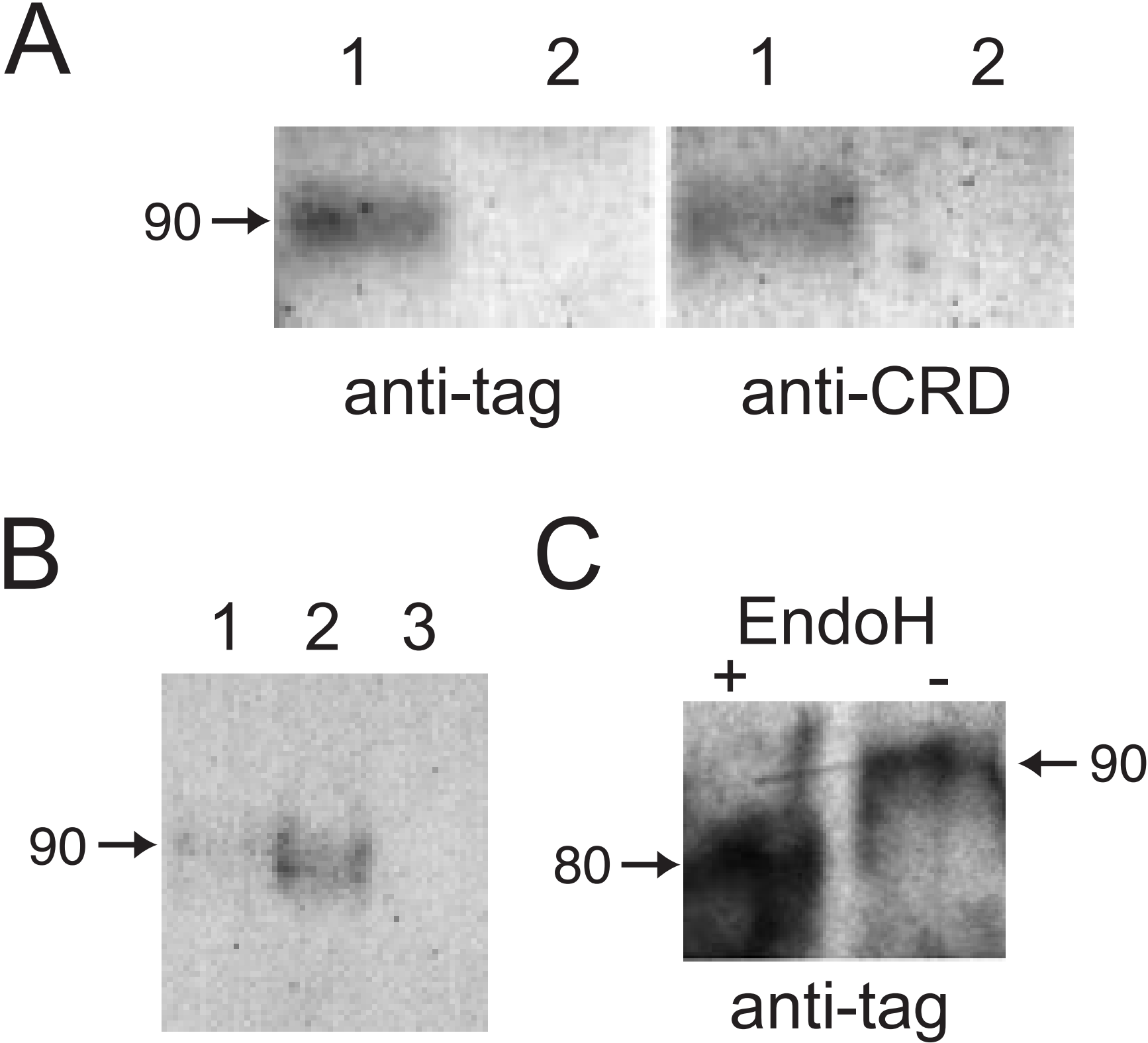

A
signal peptide
T8KP2
T10AP
GPI
HV
MUC-R
tag
NT8HAPNST7AP
S 13% - T 6% - P 9.6%
MUC-NR
...MATVVAEEDGPSDPESPPEEPGRRANSLIS
NALTFPEAGTSGVQDKTHKAAAGQGSGVSSTEQ
TKRGSAPGHSGSVGGENNPVGSTGNSVKSVVKS
PEQDKEVLGPNEEAKIQIP...
90%
20%
67%
B
MUC-R
MUC-NR
wt


A
MUC-R
wt
MUC-NR
1 2 3 4 5 6 7 8
193
112
86
70
57
30
B
Ph Aq
90
85
C
Ph Aq
45
D
1 2
85


A
MUC-R
wt
NA
-
+
-
+
103
81
46
28
B
E
T
205
112
87
69
56
38
33


I
II
III
IV
V
1
2
3
4


800
600
400
200
0
cpm
1 2 3 4 5
6
7
8
0
5
10
15
20
25
30
35
40
min


A
1
2
1
2
90
anti-tag
anti-CRD
B
1
2
3
90
C
EndoH
+
-
90
80
anti-tag