Isolation, characterization and expression of a *GRA2* homologue from *Neospora caninum*

J. T. ELLIS[1]*, C. RYCE[1], R. ATKINSON[1], S. BALU[1], P. JONES[2] *and* P. A. W. HARPER[3]

[1] *Molecular Parasitology Unit, Department of Cell and Molecular Biology, University of Technology, Sydney, Westbourne Street, Gore Hill, NSW 2065, Australia*
[2] *Department of Cell and Molecular Biology, University of Technology, Sydney, Westbourne Street, Gore Hill, NSW 2065, Australia*
[3] *Grafton Agricultural Research and Advisory Station (NSW Agriculture), Grafton, NSW 2460, Australia*

(*Received 22 July 1999; revised 11 November 1999; accepted 11 November 1999*)

SUMMARY

A cDNA library derived from mRNA of tachyzoites of *Neospora caninum* (NC-Liverpool strain) was screened with antisera from a cow naturally infected with *N. caninum*. The DNA sequence of 1 recombinant isolated predicted a significant protein sequence homology of the gene product to the 28 kDa (*GRA2*) antigen of *Toxoplasma gondii*. Studies on the *N. caninum* gene coding for this antigen demonstrated the presence of a single intron flanked by 2 exons; the gene was also highly expressed in culture-derived tachyzoites. The antigen was expressed in *Escherichia coli*; when injected into mice it stimulated the production of antibodies which detected a 29 kDa antigen of *N. caninum*. Secondary structure predictions made for the *N. caninum* protein showed support for several amphipathic helices separated by loops and turns. The available evidence indicates maintenance of protein secondary structure, and not DNA or amino acid sequence, has occurred during the evolution of *GRA2* proteins in *N. caninum* and *T. gondii*.

Key words: *Neospora caninum*, *Toxoplasma gondii*, *GRA2*, dense granule, antigen.

INTRODUCTION

The genus *Neospora* was established in the family Sarcocystidae because of the close similarity in morphology between *N. caninum* and other cyst-forming coccidia such as *Toxoplasma gondii*. The complete life-cycle of *N. caninum* is not known but congenital transmission appears common in naturally infected cows and a number of experimentally infected animals (Dubey & Lindsay, 1996). Horizontal transmission has not been recorded; however, dogs excrete oocysts after experimental infection (McAllister *et al.* 1998) suggesting transmission from dog to cow and other intermediate hosts may be possible. *N. caninum* is recognized as one of the main causes of abortion in livestock around the world including, for example, the United States of America (Anderson *et al.* 1995), Europe (Wouda *et al.* 1997; Gottstein *et al.* 1998), New Zealand (Reichel, 1998) and Australia (Boulton *et al.* 1995). Vaccines for the control of *N. caninum* are not available, although infections in dogs (if caught early enough) may be treated with clindamycin (Knowler & Skerritt, 1994). Therapy is not considered practical for cattle herds, and a vaccine is believed to represent one potential form of control.

In this study we describe the isolation, by immunoscreening a cDNA expression library, of a gene sequence coding for an antigen of the tachyzoite life-cycle stage of *N. caninum*†. The gene sequence isolated predicted a significant level of protein sequence homology of the gene product to the *GRA2* antigen of *T. gondii*.

MATERIALS AND METHODS

Parasite culture

N. caninum isolates NC-Liverpool (Barber *et al.* 1995) and NC-SweB1 (Stenlund *et al.* 1997) were propagated *in vitro* in Vero host cells according to established procedures (Barber *et al.* 1995).

Immunoscreening of expression libraries

The approach used involved screening of a recombinant cDNA expression library with sera obtained from a naturally-infected cow. Sera were obtained from cows in a herd where *Neospora*-associated abortion was common (Atkinson *et al.* 1999*a*). Serum

* Corresponding author: Department of Cell and Molecular Biology, University of Technology, Sydney, Westbourne Street, Gore Hill, NSW 2065, Australia. Tel: +61 2 9514 4161. Fax: +61 2 9514 4003. E-mail: j.ellis@uts.edu.au

† The nucleotide sequence data reported in this paper are available from the GenBank database under the accession number AF196293.

 © 2000 Cambridge University Press
https://doi.org/10.1017/S0031182099005673 Published online by Cambridge University Press

from one cow (cow X), reactive to *N. caninum* tachyzoite antigen in immunoblots, was prepared for immunoscreening by pseudoscreening in order to remove non-specific cross-reactive antibodies from it prior to use (Sambrook *et al*. 1989).

A cDNA library, prepared from mRNA of *in vitro*-derived tachyzoites of NC-Liverpool, was screened with prepared antisera from cow X using standard procedures. Briefly, filters containing plaques were incubated in Tris-buffered saline supplemented with 5% skim milk powder either overnight at 4 °C, or for 1 h at room temperature (RT) in order to prevent non-specific antibody binding. Filters were then incubated for 45 min at RT with either a negative bovine control serum or from cow X, diluted 1/50 or 1/100. Filters were then washed in Tris-buffered saline–Tween and further incubated for 45 min at RT in a 1/500 dilution of anti-bovine IgG conjugated to alkaline phosphatase (Sigma). Washing was repeated and membranes were placed in the developing solution of nitroblue tetrozoleum and 5-bromo-4-chloro-3-indolyl-phosphate (Sigma) for 20 min at RT. Recombinant clones expressing *N. caninum* antigens were picked and rescreened until a pure population of phages was produced.

Characterization of cloned sequences

The cloned DNAs coding for *N. caninum*-specific antigens were characterized as follows. A recombinant phage plaque was picked into double distilled water and subject to PCR amplification using primers FpB (5′-GTAAAACGACGGCCAGT-3′) and RpB2 (5′-GCCGCTCTAGAACTA-3′). A 50 μl PCR reaction was used with 2·5 mM $MgCl_2$, 200 μM dNTP, 25 pmol primer with cycling conditions, 1 cycle, 95 °C, 3 min; 25 cycles, 95 °C, 1 min, 52 °C, 1 min, 72 °C, 2·5 min and 1 cycle, 72 °C, 5 min. Five μl of the PCR product was run on a 1% agarose gel to estimate size and amount of product obtained. The PCR product was then purified using a Qiagen column and sequenced by cycle sequencing and the aid of an ABI automated sequencer. The non-redundant nucleotide sequence database maintained by the National Centre for Bioinformatics (NCBI) and the Apicomplexa nucleotide sequence database at the Parasite Genome Blast Server (PGBS; http://www.ebi.ac.uk/parasite_blast_server.html) were searched with the sequences obtained using the program BlastN and TBlastX in order to detect homologies with nucleotide sequences currently in the nucleotide sequence databases. The recombinants were then grouped according to their database matches. Further searches were also made of the *Toxoplasma* Database of Clustered ESTS (ToxoDB; http://www.cibil.upenn.edu/agi-bin/ParaDBs/Toxoplasma/index.html).

In order to obtain the complete sequence of the *N. caninum* cDNAs isolated, a PCR product derived for each cloned insert was cloned into the plasmid vector pGEM-T and the inserts were sequenced by cycle sequencing and a LiCOR sequencer. The sequences obtained were compiled using AssemblyAlign.

Northern blotting

Total RNA was extracted from *N. caninum* tachyzoites using a Qiagen RNeasy Mini kit following the manufacturer's instructions. The quality of the RNA was checked by agarose gel electrophoresis. For Northern blotting, 5 μg of total RNA was mixed with formaldehyde, formamide, 10 × MOPS buffer and DEPC-treated water. This mixture was heated to 65 °C for 10 min and gel loading buffer added (50% glycerol, 1 mM EDTA, 0·25% bromophenol blue, 0·25% xylene cyanol). Samples were then loaded onto a 1% agarose gel containing 5% formaldehyde and 1 × MOPS buffer. RNA markers (0·28–6·58 kb range) from Promega were used. The gel was run overnight at 30 V with buffer recirculating. After electrophoresis, RNA markers were cut off, stained with ethidium bromide and photographed. The remaining gel was Northern blotted as detailed by Sambrook *et al*. (1989). Membranes carrying RNA were pre-hybridized for 1 h at 65 °C in hybridization solution (6 × SSC, 5 × Denhardts, 0·5% SDS, 20 μg/ml salmon sperm DNA). Finally, 150 ng of DNA (for probe) was labelled using the Amersham Megaprime kit and added to the membrane which was hybridized overnight at 65 °C. The membrane was then washed 3 times at room temperature (2 × SSC) for 10 min each. Two further washes were done for 30 min in 0·1 × SSC, 0·1% SDS. Membranes were then rinsed in 0·2 × SSC, wrapped in Gladwrap and exposed to Fuji film for the required time.

Expressed sequence tag analysis

Individual, random, recombinant phage plaques were picked from the cDNA library, placed in 100 μl of sterile water and boiled for 3 min before being put on ice. Then 5 μl of this was used as a template for a PCR reaction using primers FpB and RpB2 as described above. PCR products were then purified using the Qiaquick (Qiagen) PCR purification kit and cycle sequenced with the RpB2 primer and the aid of an ABI automated sequencer.

All DNA sequences were manually inspected and edited to remove vector sequences and sequences of poor quality normally close to the primer binding sites. The poly A tail, if present, was also removed and the *N. caninum* data set was then compiled using CreateDB into a local database (MyDB:Ncaninum) on the Australian Genome Information Service (ANGIS). BlastN was used to search MyDB:Ncaninum for sequences homologous to those under

https://doi.org/10.1017/S0031182099005673 Published online by Cambridge University Press

study. Matches were considered significant if scores were returned with a probability of greater than or equal to 10^6.

Isolation of GRA2*-like sequences from genomic DNA of* N. caninum

PCR was performed using total cellular DNA from NC-Liverpool and NC-SweB1 using primers 12F2 (5′-CGAGCACCCACAAGTAA-3′) and 12R2 (5′-GACCATAACGGATGCAAC-3′). PCR and DNA sequencing was also performed with primers P28F (5′-CAGCGGTTATTCCGGATA-3′) and P28R (5′-GCCTCAAGAATTTCCTCAGC-3′). PCR products were then purified using the Qiaquick (Qiagen) PCR purification kit and sequenced by cycle sequencing and the aid of an ABI automated sequencer.

*GRA*2 intron sequences were amplified by PCR using primers CRIF (5′-GGTAGGTTACCACA-ACTTGC-3′) and CRIR (5′-GCAATTGCATT-GAGCATC-3′) that were designed from within the intron sequence of *GRA*2. The PCR cycling conditions used were: 95 °C, 3 min, 1 cycle; 95 °C, 45 sec; 50 °C, 45 sec; 72 °C, 1 min, 28 cycles; 72 °C, 5 min, 1 cycle. Five μl of PCR product was run on a 1 % agarose gel to check for amplification and size.

Expression of NcGra2 in E. coli

The open reading frame (ORF) of *GRA*2 was PCR amplified from clone 12 with primers pTrcHis-IDA12F2 (5′-ACGGATGGATCCGTTCACGG-GGAAACGTTGG-3′) and pTrcHisIDA12R2 (5′-ACGTCAGAATTCTAACGCCATACACACC-GT-3′). These primers place unique *Bam*HI and *Eco*RI restriction sites on the 5 and 3 prime sides of the *GRA*2 ORF respectively. PCR product was checked on a 1 % agarose gel for size and purified using a Qiaquick PCR purification kit. DNA from the purified PCR product and pTrcHisB vector (Invitrogen) were then digested with both *Bam*HI and *Eco*RI restriction enzymes for 3 h at 37 °C. The digested DNA were purified using a Qiaquick column and checked on a 1 % agarose gel. The ORF of *GRA*2 was then ligated into the pTrcHisB vector and transformed into *E. coli* DH5α. Individual recombinants were screened for inserts by PCR using primers pTrcHisFwd (5′-GAGGTATATA-TTAATGTATCG-3′) and pTrcHisIDA12R2. The sequence of the constructs made were confirmed by cycle sequencing. This strategy ensures the initiation codon of *GRA*2 is cloned in-frame into the pTrcHisB vector, which following transcription and translation should produce a polypeptide of 26 kDa. Subsequently, *E. coli* containing recombinant DNA were grown in LB medium containing ampicillin and at mid-log phase were induced with 1-mM IPTG. After several hours, the bacteria were collected by centrifugation and solubilized in guanidinium lysis buffer. His-tagged protein was purified using Ni-NTA (Qiagen) resin following the manufacturer's instructions for preparation of denatured *E. coli* cell lysate. Proteins were analysed on 14 % SDS–PAGE gels by either staining with Coomassie blue or by Western blotting after transfer to PVDF membrane (Atkinson *et al*. 1999*a*). Antigen expression was detected using pooled, mouse serum from animals made resistant to a lethal challenge of NC-Liverpool. This serum was produced in female in-bred Balb/C mice using 2 infections of NC-SweB1 tachyzoites as described by Atkinson *et al*. (1999*b*).

Native tachyzoite antigen corresponding to NcGra2

Purified recombinant NcGra2 was injected into 5 QS mice with complete and incomplete Freund's adjuvant 4 weeks apart. Serum was collected from the mice 3 weeks following the booster inoculation and used in immunoblot analysis of antigen extracts from NC-Liverpool and NC-SweB1.

Secondary structure predictions for Gra2

Signal peptides were predicted using the SIG-CLEAVE program of von Heijne (1986). The protein sequence of NcGra2 was also submitted to the PSA server and a secondary structure prediction made using a Type-1 analysis and the DSM model of Stultz, White & Smith (1993) which presumes that the protein is a monomeric, single-domain, globular, water-soluble protein. The following algorithms were subsequently used to investigate the location of potential helical structures in NcGra2: SSPRED (Mehta, Heringa & Argos, 1995), nnSSP (Salamov & Solovyev, 1995), PHDsec (Rost & Sander, 1993; Rost, 1996), GOR 1 (Garnier, Osguthorpe & Robson, 1978), 2 (Gibrat, Garnier & Robson, 1987) and 4 (Garnier, Gibrat & Robson, 1996), SIMPA96 (Levin, 1997), LEV (Levin, Robson & Garnier, 1986), DPM (Deleage & Roux, 1987), predator (Frishman & Argos, 1996), SOPM (Geourjon & Deleage, 1994), SOPMA (Geourjon & Deleage, 1995) and HNN (Guermeur, 1997). Solvent accessibility was performed using PHDacc (Rost & Sander, 1994).

RESULTS

*Isolation of NC*GRA2

Twenty-five independent bacteriophage clones were isolated that express antigen which was recognized by antibody from a cow that was naturally infected with *Neospora*. The sequence of one recombinant (clone 12) was found to predict significant protein sequence homology of the gene product to the amino

https://doi.org/10.1017/S0031182099005673 Published online by Cambridge University Press

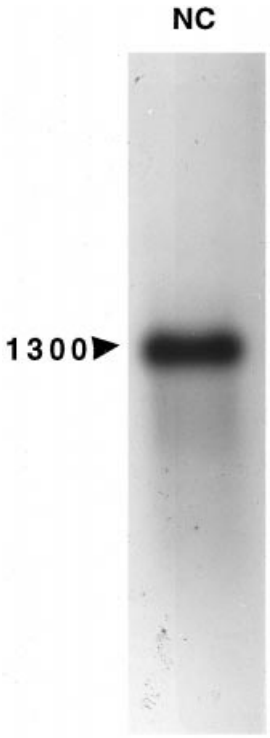


Fig. 1. Northern blot of total cellular RNA probed with *GRA*2.

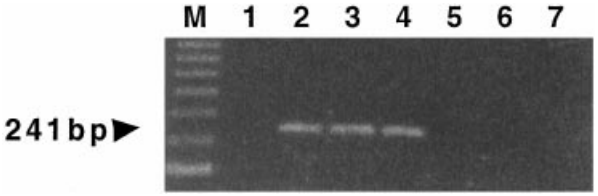


Fig. 2. PCR amplification of the *GRA*2 intron using primers CRIF and CRIR. Lane 1, 100 bp DNA ladder; Lane 2, no DNA control; Lane 3, NC-Liverpool genomic DNA; Lane 4, NC-SweB1 DNA; Lane 5, NC-1 DNA; Lane 6, *Toxoplasma gondii* (RH) DNA; Lane 7, Vero DNA.

acid sequence of the 28 kDa antigen (Gra2) of *T. gondii* (Prince *et al*. 1989; hereafter called TgGra2) and so was studied further. The sequence of clone 12 (hereafter called NC*GRA*2) clustered, using the Tblast X algorithm, in the ToxoDB database with the cluster Ctoxqual2_1721 and Ctoxqual2_289 which contains sequences coding for Gra2. Thus we conclude this clone represents an *N. caninum* homologue of TG*GRA*2, which has not been described previously.

*Expression and gene organization of NC*GRA*2 in tachyzoites*

RNA was extracted from tachyzoites of NC-Liverpool and subjected to Northern blotting using clone 12 as probe. A single transcript of approximately 1300 bases was detected (Fig. 1). Twelve (2·3 %) ESTs were homologous to NC*GRA*2. This represented the most abundant transcript detected in the data set. The EST sequences and the sequence of clone 12 were compiled to yield a consensus sequence for the mRNA of NC*GRA*2.

PCR amplification of total cellular DNA using primers 12F2 and 12R2 from both NC-Liverpool and NC-SweB1 yielded 2 PCR products (approximately 800 and 1200 bps). These were both sequenced and subsequent BlastN searches revealed the 1200 bp fragment contained the desired *GRA*2-like sequences. The 800 bp fragment was found to be homologous to cytochrome B of *T. gondii* (GenBank accession number AF023246). The 1200 bp product from NC-SweBI and NC-Liverpool was almost identical in sequence (98 %).

Genomic and cDNA sequences for NC*GRA*2 were compared. The gene encoding NcGra2 was found to possess 2 exons separated by an intron of 242 bp. The intron showed no sequence similarity to any sequence in GenBank including the intron of the *T. gondii* gene. In order to confirm this observation PCR was performed with primers CRIF and CRIR using both *N. caninum* and *T. gondii* genomic DNA as template. A PCR product of 228 bp was produced only from *N. caninum* DNA (NC-Liverpool, NC-SweBI and NC1 strains) but not from DNA of Vero or *T. gondii* (RH strain) (Fig. 2). DNA sequencing confirmed the PCR product was derived from the *N. caninum* intron.

A comparison of the *N. caninum* and *T. gondii* (M993921) coding sequences revealed (excluding the 3 prime end) a 56 % sequence similarity between them. The nucleotide differences between the two sequences were manifest as a range of indels and nucleotide substitutions. In addition, the 3 prime end of NC*GRA*2 encoded 19 additional amino acids not present in TgGra2.

*Expression of NC*GRA2 *in* E. coli *and identification of native tachyzoite antigen*

Western blotting detected a 45 kDa antigen in both soluble and insoluble denatured, reduced extracts of *E. coli*. Consequently, bacteria expressing NC*GRA*2 were collected by centrifugation and solubilized in guanidinium lysis buffer and His-tagged protein purified using Ni-NTA resin following the manufacturer's instructions. After purification, Coomassie blue staining of an SDS–PAGE gel revealed the presence of a 35 kDa protein that was also detected specifically by mouse anti-*N. caninum* antiserum (data not shown). Injection of this protein into mice, resulted in the production of IgG antibodies to an *N. caninum* tachyzoite antigen of 29 kDa in NC-Liverpool (Fig. 3) which was also detected in NC-SweBI. Reduction of tachyzoite proteins did not alter the size of the antigen detected by Western blotting (data not shown).

https://doi.org/10.1017/S0031182099005673 Published online by Cambridge University Press

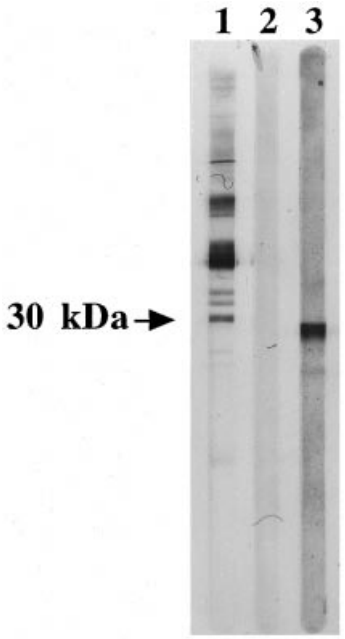


Fig. 3. Western blot analysis of non-reduced tachyzoite proteins of NC-Liverpool probed with pooled sera from mice injected with NcGra2 produced in *E. coli*. Lane 1, pooled sera from mice inoculated with NC-Liverpool 15–21 days following infection; Lane 2, negative control; Lane 3, pooled sera from mice injected with recombinant NcGra2. The immunodominant 30 kDa antigen of *Neospora caninum* is arrowed.

Secondary structure predictions

Analysis of the predicted protein sequence encoded by NC*GRA*2 (excluding the unique C-terminal domain of NcGra2) revealed the amino acid sequence was 52% similar to TgGra2. The amino terminus was particularly conserved. The SIGCLEAVE program predicted a signal peptide (WILVVAVGALVGA) in this region that was almost identitical to that present in *T. gondii*. Mercier *et al*. (1993) predicted that the secondary structure of Gra2 in *T. gondii* was a globular protein with 2 amphipathic helices separated by a 10 amino acid linker. Consequently, the protein sequence of NcGra2 was submitted to the PSA server and a secondary structure prediction made using a Type-1 analysis and the DSM model of Stultz *et al*. (1993). The analyses showed that NcGra2 most probably belongs to the protein superclass which predominantly contains alpha helices (probability 0·95574). The algorithm also predicted the presence of 2 major and 1 minor helical regions in NcGra2 and that the most plausible explanation for the structure of the remaining residues of NcGra2 was in the form of loops or turns. In NcGra2 the helical regions spanned residues 70–110, 110–150 and 170–190. This secondary structure prediction was investigated further using 14 additional algorithms, the results from which are presented in Fig. 4. A consensus opinion from this alignment is considerable support for 4, and not 3, helices (H1-4) spanning residues P69-V92, E95-D116, K120-G132 and N176-G193. The amphipathic nature of H1, H2 and H4 is clearly evident by the distribution of hydrophobic and hydrophilic residues on alternative sides of a helical wheel derived from these amino acid sequences (not shown).

DISCUSSION

A gene from *N. caninum*, homologous to the *GRA*2 gene of *T. gondii*, has been cloned and sequenced. The *N. caninum* gene, like that of *T. gondii*, is composed of 2 exons and a single intron and is highly expressed in tachzyoites (as detected by Northern blotting and EST sequencing) as a very abundant messenger RNA.

Early research on the *T. gondii* antigen indicated a submembraneous location in the tachyzoite (Sibley & Sharma, 1987) although more recent work has now shown that TgGra2 is located in the dense granules of the tachyzoite. Upon infection TgGra2 is secreted into the parasite-containing vacuole where it is rapidly and specifically targeted to a network of membranous tubules which connect with the vacuolar membrane (Mercier *et al*. 1998). The subcellular location and function of NcGra2 is currently not known; however, it is likely to fullfill a similar role to TgGra2.

The protein sequence of NcGra2 is approximately 50% similar to that of TgGra2. Although our study shows that the protein sequence of Gra2 is not highly conserved between *N. caninum* and *T. gondii*, it would appear that maintenance of secondary structure has occurred during the evolution of these molecules. Secondary structure predictions made, using a wide variety of algorithms, indicate a high degree of support for NcGra2 containing 3 or 4 amphipathic helices separated by loops and turns. The algorithms used do not agree on the precise location of the helices, yet they indicate a similar location for the helices within both NcGra2 and TgGra2. In passing we report the structure of TgGra2 was also re-evaluated using these computer algorithms. H1 in TgGra2 was assigned to residues P70-V94, H2 to S97-K115 plus an additional helix (H3) was predicted at R121-G133. These results differ from those of Mercier *et al*. (1993) by the location of the start of H1 at P plus the existence of H3 not previously identified nor suggested. Proline has long been known to be a 'helix breaker' (Chou & Fasman, 1973) and so is generally restricted to the first turn of the N-terminal helix. Its location as the first residue in an amphipathic helix, or in a turn leading into the helix, is suggested by the fact that proline has no free NH group and therefore cannot form the conventional intra-helical NH...O=C hydrogen bond (Chakrabarti & Chakrabarti, 1998). The subtle differences proposed here in the helical

https://doi.org/10.1017/S0031182099005673 Published online by Cambridge University Press

```
          ....,....1....,....2....,....3....,....4....,....5....,....6
AA        MFTGKRWILVVAVGALVGASVKAADFSGRGTVNGQPVGSGYSGYPRGDDVRESMAAPEDL
PHD sec       EEEEEEEEE   E EEEEEEEEE
Nnpred    ------EEEEEHH-HHHHHHH-HH-----------------------------------
PredicSS        EEEEEEE                                    HHHHHHHHH
GOR1      HHHHHHHEEEEEEEEHHHHHHHHHHHEEETTTCCEEEEEEEEETEEETEEEEHHHHHHHHHC
GOR4      CCCCCEEEEEEEEEHHHHCHHHHHHCCCCCCEECCCCCCCCCCCCCCCCCCHHHHHCCCCCC
GOR2      HHHHCHHHEEEHHHHHHHHHHHEEHEEECCEEEECEEECECCCCCCCCCCCEHHHHCCCCCC
SIMPA96   CCCCCEEEEEEEHHHHHHHHHHHHHCCCCCCCCCCCCCCCCCCCCCCCCCHHHHHCCCCCCC
LEV       HHCCCHHEEEEEHHHHHCHCCCHHCCCTCCEECTCCECTCCESCCTTHHHHHHHHHCCSCC
DPM       CCHCHHHEEEEHEHHHEHHHHHHHHHTTCCCCCCTCCTTTCTCCCCTCCHHHHHHHHCHHH
PREDATOR  CCCCCCEEEEEECCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCHHHHHCCCCCC
SOPM      TCCCCCEEEEEEHHHHHHHHHHHHTCCCCCEEEETCCCCCCCCCCCCCCCHHHHHHCCCCCC
SOPMA     TCCCCCEEEEEEHHHHHHHHHHHHTCCCCCEEEETCCCCCCCCCCCCCCCHHHHHHCCCCCC
HNN       CCCCHHHHHHHHHHHHHCCCEEHCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
P3        bbeb bbbbbbbbbbbebbbebbebbbebeeeeeebebbeeeee beebeeeeeeeeee

          ....,....7....,....8....,....9....,....10...,....11...,....12
AA        PGERQPETPTAEAVKQAAAKAYRLLKQFTAKVGQETENAYYHVKKATMKGFDVAKDQSYK
PHD sec           HHHHHHHHHHHHHHHHHHHHHH    HHHHHHHHHHHHHHHHE     HH
Nnpred    ----------HHHHHHHHHHHHHHHHHHHH----HHHHHHHHHH--HHHH--------H
Predic          HHHHHHHHHHHHHHH HHHHHHHHHH                          HH
GOR1      HTCCCCCCHHHHHHHHHHHHHHHHHHHHHHHHCCHHHHHHHHHHHHHHHHHHHHHHTTE
GOR4      CCCCCCCCCHHHHHHHHHHHHHHHHHHHHHHHCCCHHHHHHHHHHHHHCCCEEECHHHHH
GOR2      CCCCCCCCCHHHHHHHHHHHHHHHHHHHHHHHHCHHHHHHHHHHHHHHHHHHHHHHHHH
SIMPA96   CCCCCCCCHHHHHHHHHHHHHHHHHHHHHHHCCCCCHHHHHHHHHHHCCCCCCCCCCCC
LEV       CTCCCTCCCHHHHHHHHHHHHHHHHHHHHHHHCCHHHHHHHHHHHHHHTCCCCCCTCCH
DPM       CCHHCCHCCHHHHHHHHHHHHHHHHHHHHHHHCHHHHHHHHHHHHHHHHHHHHHHHCCC
PREDATOR  CCCCCCCCCCCCHHHHHHHHHHHHHHHHHHHCCCCHHHHHHHHHHHCCCCCCCCCCCCH
SOPM      CCCCCCCCCHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHTTEEECCCCCHH
SOPMA     CCCCCCCCCHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHTTEEECCCCCHH
HNN       CCCCCCCCCCHHHHHHHHHHHHHHHHHHHCCCCCHHHHHHHHHHHHCCCHHCHCHHHC
P3        eeeeeeeebeeeebeebbeebb ebeebbeebbeebeebbeebeebbeebbebbeeeb e

          ....,....13...,....14...,....15...,....16...,....17...,....18
AA        GYLAVRKATAKGLQSAGKSLELKESAPTGTTTAAPTEKVPPSGPSGEVQRTRKEQNDVQQ
PHD sec   HHHHHHHHHHHHHHH   EEEE                                HHHHHH    H
Nnpred    -HEEHHHHHHHHHHH---HHH-----------------------------------HHH
Predic    HHHHHHHHHHHHHH                                    HHHHHHHHH  HH
GOR1      EEEEHHHHHHHHHHHHHHHHHHHHHCCTCEEEEEEEEEETCCCCCEEEEHHTTTCHHHH
GOR4      CHHHHHHHHHHHHHHHHHHHHHHCCCCCCCCEECCCCCCCCCCCCCCCCHHHHHHHCHHHH
GOR2      HHHHHHHHHHHHHHHHHCHHHHEEECCCCCCCEECCCCCCECCCCCCCCCEEEEHHHCCCHHH
SIMPA96   HHHHHHHHHHHCHHHHHCCEEEECCCCCCCCCCCCCCCCCCCCCCCCCHHHHHHHHHHHHHH
LEV       HHHEHHHHHHHCCCHTTCHHHECCCCCCCCCCCCCCCCCCCCCCTSCEECCCHCHHHHHH
DPM       CHHHHHHHHHHCHHTHTHHHHHHHHHCCCCCCHHCCHHCCCCTTCCCHHHHHHHHHHHHH
PREDATOR  HHHHHHHHHHHHHCCCCCHHHHHCCCCCCCCCCCCCEEECCCCCCCCCCCCCCCCCHHHH
SOPM      HHEEHHHHHHHHHHHTTHHHHHHCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCHHHH
SOPMA     HHEEHHHHHHHHHHHTTHHHHHHCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCHHHH
HNN       HHHHHHHHHHHCCCCCCCEEEECCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCHH
P3        bbebbeebbbebbeebbeebeeeebeeeebbbeb eeeeeeeeeeeeeeee eeeeebe

          ....,....19...,....20...,....21
AA        TAEMLAEEILEAGLKKDDGEGRGTPEAEVN
PHD sec   HHHHHHHHHHHHHH
Nnpred    HHHHHHHHHHHHH-----------------
Pred SS   HHHHHHHHHHHH
GOR1      HHHHHHHHHHHHHHHHHTTTTCCCCCHHHHH
GOR4      HHHHHHHHHHHHHHCCCCCCCCCCCCEEEEC
GOR2      HHHHHHHHHHHHHHHHCCCCCCCCCCCCHEEH
SIMPA96   HHHHHHHHHHHHHCCCCCCCCCCCCCCCEEC
LEV       HHHHHHHHHHHHHCCCTSSTCCCCCHHHHH
DPM       HHHHHHHHHHHHHHHHTTTTCTCCCCHHHCC
PREDATOR  HHHHHHHHHHHHCCCCCCCCCCCCCCCCCC
SOPM      HHHHHHHHHHHHHHCTTCCCCCCCCHHHEE
SOPMA     HHHHHHHHHHHHHHCTTCCCCCCCCHHHEE
HNN       HHHHHHHHHHHHHCCCCCCCCCCCCCCCCC
P3        ebbebbbeebbeebbeeeeeeeee eebeb
```

Fig. 4. Secondary structure predictions for NcGra2. The algorithm predictions are shown aligned for each residue. The amino acid sequence of NcGra2 is shown in line AA; the numbers above this line represent residue position (in 10's); the following annotation is used: α-helix; E, β-strand; C,T or S, loop or turn; b, buried; e, exposed.

https://doi.org/10.1017/S0031182099005673 Published online by Cambridge University Press

structures from those reported by Mercier *et al*. (1993) result from the extensive number of refined algorithms which exist today.

Sufficient dissimilarity exists between the *T. gondii* and *N. caninum* proteins for us to hypothesize that they are antigenically distinct. For example, we note the carboxy termini differs between NcGra2 and TgGra2. This region contains, in *T. gondii*, an epitope recognized by antibodies from naturally infected humans (Cesbron-Delauw *et al*. 1992).

Expression of the entire *GRA*2 ORF in a prokaryotic expression vector (the plasmid pTrcHisB) was achieved. A simple purification procedure was used to isolate the recombinant protein, of apparent molecular weight 35 kDa, from *E. coli*. The molecular weights reported here are somewhat anomalous, because predictions of the protein size from the open reading frame for the bacterially expressed protein (including vector-encoded amino acids) suggest a size of 26, and not 35 kDa. We do not have a rational explanation for the observed mobility of the protein produced in *E. coli*. The anomalous mobility may simply be the influence of the protein's shape or the protein may be modified in some way. Although post-translational modifications, such as glycosylation, have been shown to occur in *T. gondii* (Odenthal-Schnittler *et al*. 1993; Schwarz & Tomavo, 1993; Dieckmann-Schuppert, Bause & Schwarz, 1994) such modifications were discounted because they do not occur in *E. coli*. Despite the anomalous mobility, the purified recombinant protein maintained its antigenicity as determined by western blotting with sera from mice resistant to an acute *N. caninum* infection. Sera from mice injected with recombinant NcGra2 produced in *E. coli* recognized a 29 kDa antigen present in NC-Liverpool and NC-SweBI tachyzoite extracts. Thus we conclude the native antigen of NcGra2 is 29 kDa.

Research suggests that TgGra2 may be worthy of including in a vaccine against toxoplasmosis. Using the techniques of passive transfer of antibody and immunization of affinity purified antigen into mice, strong evidence for the induction of protective immunity was demonstrated for this antigen (Sharma, Araujo & Remington, 1984; Sibley & Sharma, 1987). More recently, a *GRA*2 knockout strain of *T. gondii* was shown to have reduced virulence compared to a wildtype organism (Mercier *et al*. 1998). We suggest NcGra2 may therefore warrant testing as a potential vaccine against neosporosis and *N. caninum*-associated abortion.

This work was supported by grants from the Australian Research Council; Cattle Compensation Fund (NSW Agriculture) and the University of Technology, Sydney.

REFERENCES

ANDERSON, M. L., PALMER, C. W., THURMOND, M. C., PICANSO, J. P., BLANCHARD, P. C., BREITMEYER, R. E., LAYTON, A. W., McALLISTER, M., DAFT, B., KINDE, H., READ, D. H., DUBEY, J. P., CONRAD, P. A. & BARR, B. C. (1995). Evaluation of abortions in cattle attributable to neosporosis in selected dairy herds in California. *Journal of the American Veterinary Medicine Association* **207**, 1206–1210.

ATKINSON, R. A., COOK, R., REDDACLIFF, L., ROTHWELL, J., BROADY, K. W., HARPER, P. A. W. & ELLIS, J. T. (1999*a*). Seroprevalence of *Neospora caninum* infection following an abortion outbreak in a New South Wales dairy cattle herd. *Australian Veterinary Journal* (in the Press).

ATKINSON, R. A., HARPER, P. A. W., RYCE, C., MORRISON, D. A. & ELLIS, J. T. (1999*b*). Comparison of the biological characteristics of two isolates of *Neospora caninum*. *Parasitology* **118**, 363–370.

BARBER, J. S., HOLMDAHL, O. J. M., OWEN, M. R., GUY, F., UGGLA, A. & TREES, A. J. (1995). Characterization of the first European isolate of *Neospora caninum* (Dubey, Carpenter, Speer, Topper and Uggla). *Parasitology* **111**, 563–568.

BOULTON, J. G., GILL, P. A., COOK, R. W., FRASER, G. C., HARPER, P. A. W. & DUBEY, J. P. (1995). Bovine *Neospora* abortion in north-eastern New South Wales. *Australian Veterinary Journal* **72**, 119–120.

CESBRON-DELAUW, M. F., BOUTILLON, C., MERCIER, C., FOURMAUX, M. P., MURRAY, A., MIQUEY, F., TARTAR, A. & CAPRON, A. (1992). Amino acid sequence requirements for the epitope recognised by a monoclonal antibody reacting with the secreted antigen GP28.5 of *Toxoplasma gondii*. *Molecular Immunology* **29**, 1375–1382.

CHAKRABARTI, P. & CHAKRABARTI, S. (1998). C-H center dot center dot center dot O hydrogen bond involving proline residues in alpha-helices. *Journal of Molecular Biology* **284**, 867–873.

CHOU, P. T. & FASMAN, G. D. (1973). Structural and functional role of leucine residues in proteins. *Journal of Molecular Biology* **74**, 263–281.

DELEAGE, G. & ROUX, B. (1987). An algorithm for protein secondary structure prediction based on class prediction. *Protein Engineering* **1**, 289–294.

DIECKMANN-SCHUPPERT, A., BAUSE, E. & SCHWARZ, R. T. (1994). Glycosylation reactions in *Plasmodium falciparum*, *Toxoplasma gondii*, and *Trypanosoma brucei brucei* probed by the use of synthetic peptides. *Biochimica et Biophysica Acta* **1199**, 37–44.

DUBEY, J. P. & LINDSAY, D. S. (1996). A review of *Neospora caninum* and neosporosis. *Veterinary Parasitology* **67**, 1–52.

FRISHMAN, D. & ARGOS, P. (1996). Incorporation of non-local interactions in protein secondary structure prediction from the amino acid sequence. *Protein Engineering* **9**, 133–142.

GARNIER, J., GIBRAT, J-F. & ROBSON, B. (1996). GOR secondary structure prediction method version IV. *Methods in Enzymology* **266**, 540–553.

GARNIER, J, OSGUTHORPE, D. J. & ROBSON, B. (1978). Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *Journal of Molecular Biology* **120**, 97–120.

GEOURJON, C. & DELEAGE, G. (1994). SOPM: a self-

https://doi.org/10.1017/S0031182099005673 Published online by Cambridge University Press

optimized method for protein secondary structure prediction. *Protein Engineering* **7**, 157–164.

GEOURJON, C. & DELEAGE, G. (1995). SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Computer Applications in the Biosciences* **11**, 681–684.

GIBRAT, J. F., GARNIER, J. & ROBSON, B. (1987). Further developments of protein secondary structure prediction using information theory. New parameters and consideration of residue pairs. *Journal of Molecular Biology* **198**, 425–443.

GOTTSTEIN, B., HENTRICH, B., WYSS, R., THUR, B., BUSATO, A., STARK, K. D. C. & MULLER, N. (1998). Molecular and immunodiagnostic investigations on bovine neosporosis in Switzerland. *International Journal for Parasitology* **28**, 679–691.

GUERMEUR, Y. (1997). Combinaison de classifieurs statistiques, Application à la prediction de structure secondaire des proteines. Ph.D. thesis, University of Lyon.

KNOWLER, C. & SKERRITT, G. (1994). How do I treat? Canine neosporosis and toxoplasmosis. *Progress in Veterinary Neurology* **5**, 167–169.

LEVIN, J. (1997). Exploring the limits of nearest neighbour secondary structure prediction. *Protein Engineering* **7**, 771–776.

LEVIN, J., ROBSON, B. & GARNIER, J. (1986). An algorithm for secondary structure determination in proteins based on sequence similarity. *FEBS* **205**, 303–308.

McALLISTER, M. M., DUBEY, J. P., LINDSAY, D. S., JOLLEY, W. R., WILLS, R. A. & McGUIRE, A. M. (1998). Dogs are definitive hosts of *Neospora caninum*. *International Journal for Parasitology* **28**, 1473–1478.

MEHTA, P. K., HERINGA, J. & ARGOS, P. (1995). A simple and fast approach to prediction of protein secondary structure from multiply aligned sequences with accuracy above 70%. *Protein Science* **4**, 2517–2525.

MERCIER, C., HOWE, D. K., MORDUE, D., LINGNAU, M. & SIBLEY, L. D. (1998). Targeted disruption of the GRA2 locus in *Toxoplasma gondii* decreases acute virulence in mice. *Infection and Immunity* **66**, 4176–4182.

MERCIER, C., LECORDIER, L., DARCY, D., DESLEE, A., MURRAY, B., TOURVIELLE, P., MAES, A., CAPRON, A. & CESBRON-DELAUW, M. F. (1993). *Toxoplasma gondii*: molecular characterisation of a dense granule antigen (GRA2) associated with the network of the parasitophorous vacuole. *Molecular and Biochemical Parasitology* **58**, 71–82.

ODENTHAL-SCHNITTLER, M., TOMAVO, S., BECKER, D., DUBREMETZ, J. F. & SCHWARZ, R. T. (1993). Evidence for N-linked glycosylation in *Toxoplasma gondii*. *The Biochemical Journal* **291**, 713–721.

PRINCE, J. B., ARAUJO, F. G., REMINGTON, J. S., BURG, J. L., BOOTHROYD, J. C. & SHARMA, S. D. (1989). Cloning of cDNAs encoding a 28 kilodalton antigen of *Toxoplasma gondii*. *Molecular and Biochemical Parasitology* **34**, 3–14.

REICHEL, M. P. (1998). Prevalence of *Neospora* antibodies in New Zealand dairy cattle and dogs. *New Zealand Veterinary Journal* **46**, 38.

ROST, B. & SANDER, C. (1993). Prediction of protein structure at better than 70% accuracy. *Journal of Molecular Biology* **232**, 584–599.

ROST, B. & SANDER, C. (1994). Conservation and prediction of solvent accessibility in protein families. *Proteins* **20**, 216–226.

ROST, B. (1996). PHD: predicting one-dimensional protein structure by profile based neural networks. *Methods in Enzymology* **266**, 525–539.

SALAMOV, A. A. & SOLOVYEV, V. V. (1995). Prediction of protein secondary structure by combining nearest-neighbor algorithms and multiple sequence alignments. *Journal of Molecular Biology* **247**, 11–15.

SAMBROOK, J., FRITSCH, E. F. & MANIATIS, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor Press, New York.

SCHWARZ, R. T. & TOMAVO, S. (1993). The current status of the glycobiology of *Toxoplasma gondii*: glycosylphosphatidylinositols, N- and O-linked glycans. *Research in Immunology* **144**, 24–31.

SHARMA, S. D., ARAUJO, F. G. & REMINGTON, J. S. (1984). *Toxoplasma* antigen isolated by affinity chromatography with monoclonal antibody protects mice against lethal infection with *Toxoplasma gondii*. *Journal of Immunology* **133**, 2818–2830.

SIBLEY, L. D. & SHARMA, S. D. (1987). Ultrastructural localisation of an intracellular *Toxoplasma* protein that induces protection in mice. *Infection and Immunity* **55**, 2137–2141.

STENLUND, S., BJÖRKMAN, C., HOLMDAHL, O. J. M., KINDAHL, H. & UGGLA, A. (1997). Characterisation of a Swedish bovine isolate of *Neospora caninum*. *Parasitology Research* **83**, 214–219.

STULTZ, C. M., WHITE, J. V. & SMITH, T. F. (1993). Structural analysis based on state-space modelling. *Protein Science* **2**, 305–314.

VON HEIJNE, G. (1986). A new method for predicting signal sequences cleavage sites. *Nucleic Acids Research* **14**, 4683–4690.

WOUDA, W., MOEN, A. R., VISSER, I. J. R. & VAN KNAPEN, F. (1997). Bovine fetal neosporosis: a comparison of epizootic and sporadic abortion cases and different age classes with regard to lesion severity and immunohistochemical identification of organism in brain, heart and liver. *Journal of Veterinary Diagnostic Investigation* **9**, 180–185.

https://doi.org/10.1017/S0031182099005673 Published online by Cambridge University Press