Synthetic peptides

in the diagnosis of human echinococcosis

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Claudia List

aus

Arvigo (GR)

Basel, 2010

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Prof. Dr. Marcel Tanner
Dr. Ingrid Felger
Prof. Dr. Alexander Mathis

Basel, den 19.Oktober 2010

Prof. Dr. Martin Spiess
Dekan

Acknowledgements

This PhD thesis was the biggest challenge in my career so far and I am deeply grateful to those people who helped me to accomplish it with their constant support in so many different ways.

- Ingrid Felger who gave me the chance to carry out this PhD project under her supervision and who kept on motivating me with her ideas and enthusiasm.
- Norbert Müller and Bruno Gottstein from the Institute of Parasitology in Bern for the provision of precious patients' sera together with invaluable information about basic and advanced *Echinococcus*-knowledge. Thank you for explaining me facts I simply could not have grasped from literature.
- Hanspeter Marti, Stefanie Kramme and Eva Maag from the Diagnostics Services unit of the Swiss TPH for providing valuable sera and also theoretical and practical guidance in the lab.
- Marcel Tanner and Alexander Mathis for joining my PhD committee.
- Hans-Peter Beck who always had a good advice, especially in lab-related emergencies.
- Valentin Pflüger and Mabritec who solved my biggest mysteries about inconsistent ELISA results with an emergency run on their MALDI-TOF.
- Christoph Schmid and Pascal Mäser for their patience in explaining bioinformatics.
- Yvette Endriss who always helped with words and deeds.
- Dania Müller who many times saved my experiments by obtaining missing consumables within no time at all and who always had a comforting word when desperately needed.
- Sebastian Rusch whose real name is actually Sebastian Maniatis
- All the MolPar and GR lab members who created a pleasant working atmosphere.
- The library team who managed to find all the odd articles I ordered.

- Caroline Kulangara, Anna Perchuc, Annette Gaida, Sophie Oehring, Sonja Schöpflin, Serej Ley, Eva Maria Hodel and all the other numerous members of the “Damenkränzchen”: your friendship and support was simply invaluable. Thank you very much for all the good discussions, the tips and tricks and all your sympathy!
- My family and friends who never ceased to believe in me.

Summary

Echinococcus granulosus and *Echinococcus multilocularis* are the two tapeworm species responsible for the majority of human echinococcosis cases. Infection of this zoonosis can be acquired all over the world by accidental ingestion of eggs. The prevalence of echinococcosis is highest in areas with extensive animal husbandry.

Most diagnostic tests for the detection of antibodies against echinococcosis in humans are based on crude or partly purified native antigen extracts. The production of standardized diagnostic-grade native antigen is not possible. Native antigen suffers from batch-to-batch variation depending on purity, origin and developmental stage of the parasite material. In addition, native antigen is limited in availability.

Substantial progress in the development of standardized reagents has been achieved by the production of recombinant antigens. Consequent further development of antigens might lead to the production of synthetic antigen.

This PhD thesis investigated the applicability of synthetic peptides ranging from 24 to 47 amino acids in length for the use as synthetic antigens in serological diagnosis of cystic and alveolar echinococcosis. A bioinformatic selection procedure was established for identification of potentially antigenic protein sequences. From each of these parent proteins one or several peptides were selected for chemical synthesis. Peptides were designed from predicted domains of alpha-helical coiled-coils and intrinsically unstructured regions. These two motifs are capable of adopting their natural conformation even if synthesized as isolated fragments. The diagnostic performance of a synthetic antigen depends on its ability to fold into its native structure and thus be recognized by antibodies that had been naturally induced in the course of an infection.

Peptide candidates were designed from different sets of parent proteins which had been selected by three different approaches: i) a genomics approach, ii) a proteomics approach, and iii) a transcriptomics approach.

These three approaches identified 6 promising peptide candidates. The synthetic peptide performing best obtained a sensitivity of 74% for the detection of CE infection and 43% for the detection of AE infection in adult patients. Specificity was 94%. All identified peptides, apart from one candidate, were not able to discriminate between CE and AE infection.

The major drawback of using peptides for serodiagnosis was their reduced sensitivity compared to EgHF, the native antigen used in routine diagnostics. Our results indicated that test sensitivity could be increased by combination of several peptides into a pool of synthetic antigen. This strategy can solve the problem of decreased sensitivity in future assay development.

The most promising candidate, peptide longD8-9, was investigated as potential marker for serological follow-up of treatment success in young patients. This represents a novelty, as synthetic peptides have not yet been tested as marker for serological follow-up of CE patients.

The performance of longD8-9-ELISA compared well to that of EgHF-ELISA, but it was superior in that within three years after treatment, more cured CE patients reached seronegativity in peptide ELISA compared to EgHF ELISA. Non-cured CE patients produced highly positive test results until the end of the follow-up period.

This thesis provides proof-of-principle for the discovery of diagnostically relevant peptides by bioinformatic selection complemented with screening on a high-throughput microarray platform. It also showed the value of synthetic peptides as potential markers not only in primary diagnosis, but also in monitoring of treatment outcome in follow-up patients.

Abbreviations

ABZ	Albendazole
AE	Alveolar echinococcosis caused by *E. multilocularis*
AHX	Aminohexanoic acid
CC	Alpha-helical coiled-coil
CCE	Cured cystic echinococcosis
CE	Cystic echinococcosis caused by *E. granulosus*
cMDH	Cytosolic malate dehydrogenase
CT	Computed tomography
EgHF	*E. granulosus* hydatid cyst fluid
ELISA	enzyme-linked-immunosorbent assay
EmVF	*E. multilocularis* vesicle fluid
EST	Expressed sequence tag
FI	Fluorescence intensity
GPI	Glycosylphosphatidylinositol
IB	Immunoblotting
IFAT	Indirect fluoresence antibody test
IHA	Indirect hem agglutination
IUR	Intrinsically unstructured region
LSP	Long synthetic peptide
MRI	Magnetic resonance imaging
MS	Mass spectrometry
NCCE	Non-cured cystic echinococcosis
PAIR	Puncture, aspiration, injection, re-aspiration
PEG	Polyethylene glycol
TM	Transmembrane domain
US	Ultrasound imaging

General introduction

Echinococcosis in humans

Human echinococcosis is a zoonosis caused by tapeworms of the genus *Echinococcus*. There are several species infecting man, *E. granulosus* and *E. multilocularis* being the two most important. The disease in humans is caused by the growth of the metacestode larva in the liver, lungs or any other organ. Infection with *E. granulosus* causes cystic echinococcosis (CE) which is characterized by unilocular cysts, while the infection with *E. multilocularis* causing alveolar echinococcosis (AE) is characterized by the growth of infiltrating, tumor-like metacestode tissue scattered with multiple vesicles. Figure 1 depicts the life cycle of both tapeworm species and illustrates the appearance of the metacestode larvae (from Wen et al. [1]).

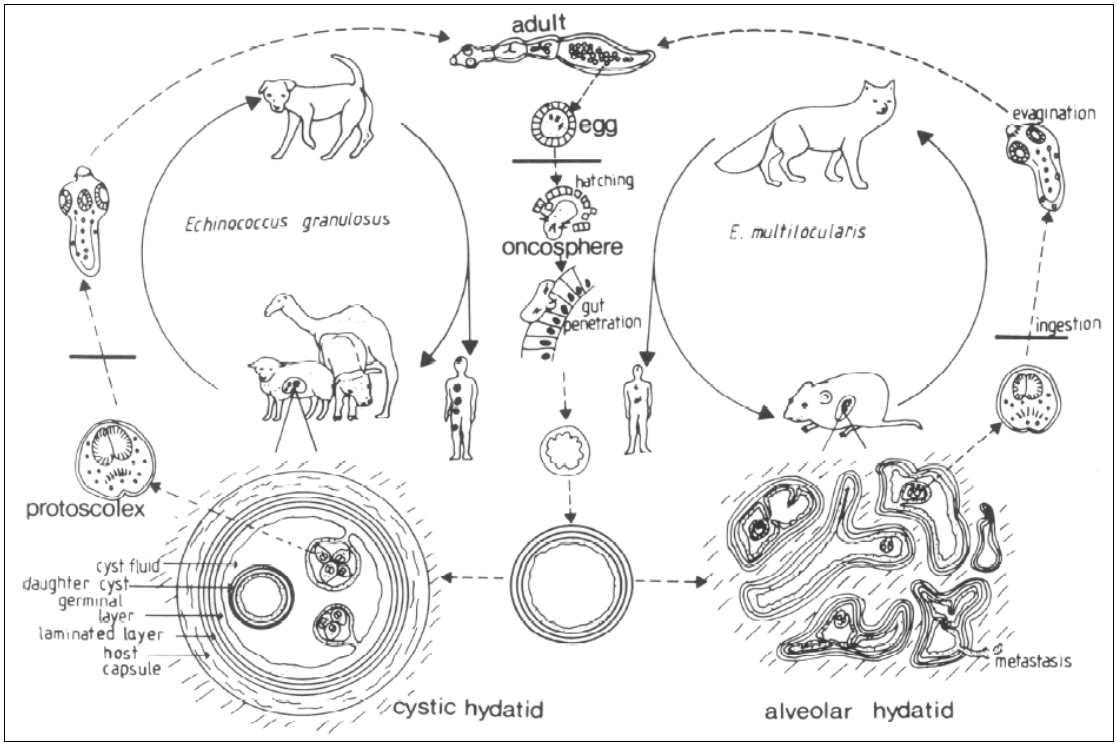

Figure 1: Life cycle of *E. granulosus* (left-hand side) and *E. multilocularis* (right-hand side), picture from Wen et al. [1].

E. granulosus has a worldwide distribution. CE is transmitted in domestic life cycles involving dog and livestock and constitutes an emerging public health problem especially in regions with extensive livestock husbandry and non-supervised slaughter [2]. AE is endemic in the northern hemisphere and is

transmitted in sylvatic life cycles involving foxes and small rodents. AE was considered to be rather rare because of limited contact between humans and wildlife. However, human AE can be extremely frequent particularly in foci where domestic dogs became involved in the life cycle as definitive hosts. Such reports exist from Alaska [3], People's Republic of China [4] and Europe [5]. Recent reports demonstrate the spread of *E. multilocularis* in Europe after distinct increase in fox populations [5-7].

Humans acquire infection by accidental ingestion of eggs from contaminated environment. The disease develops slowly. The latency period of AE is approximately 10-15 years [8]. Symptoms mainly occur when active cysts or metacestode tissue expand in size resulting in dysfunction of adjacent tissues or organs [9].

The diagnosis of echinococcosis is based on the detection of cyst structures by imaging procedures, mainly ultrasound and computed tomography, and confirmation by immunodiagnostic tests demonstrating *Echinococcus*-specific antibodies. Hydatid cyst fluid is the antigenic source of reference for immunodiagnosis of human hydatidosis [10].

ELISA using crude hydatid cyst fluid (EgHF) has a high sensitivity (75 - 95%), but its specificity is often unsatisfactory [10]. Cross-reactivity with sera from patients infected with other helminthic diseases is commonly observed (89% with other cestode infections, 30% with trematode infections and 39% with nematode infections) [11]. Specificity can be improved by using purified native antigen and/or alternative diagnostic techniques, such as immunoblotting.

EgHF is a complex mixture of glyco- and lipoproteins, carbohydrates and salts [10]. Most of these molecules are of parasite origin, but also host components were found, e.g. host albumin and immunoglobulins. Antigen B and antigen 5 are the diagnostically most important components of EgHF.

Antigen 5 is synthesized as single polypeptide chain and is processed into single disulphide-bridged 22 and 38kDa subunits [12]. Applied in

immunoelectrophoresis, it produces the diagnostically relevant precipitation line termed “arc-5” [13]. A study comparing recombinant, native, and deglycosylated native antigen 5 showed the importance of post-translational glycosylations for the antigenicity of this protein [14]. Although the diagnostic value of antigen 5 has been thoroughly evaluated, there is only little knowledge in the biological function of this protein. Lorenzo and co-workers showed high similarity of antigen 5 with serine proteases of the trypsin family in terms of sequence conservation and secondary structure [12].

Antigen B is a polymeric lipoprotein of 160kDa that aggregates from 8kDa subunits which are encoded by a multigene family. A recent study showed that this multigene family comprises at least five subfamilies and identified 10 unique family members which were differentially expressed throughout the life-cycle [15]. This study also showed that antigen B genes were highly conserved even in isolates from distant geographical locations or from different hosts [15]. These findings contradicted previous reports claiming significant variation in antigen B family members isolated from different hosts [16-18]. The implication of this recently detected sequence conservation is that antigen B subunits can be considered valuable antigens for immunodiagnosis.

The biological function of antigen B is no yet fully understood. A possible role of antigen B could be in the evasion of host immune response. Antigen B was shown to inhibit neutrophil chemotaxis [19] and to elicit a non-protective Th2 immune response [20]. Applied in SDS-PAGE under reducing and non-reducing conditions, antigen B dissociates into 8-12, 16 and 20-24kDa subunits [10]. Its smallest 8kDa subunit is highly antigenic and recognized in immunoblot by about 90% of patients with CE infection and 40% of patients with AE infections [21].

The major issue of using native parasite material as main source of antigen is the lack of standardization due to the varying content of its components. It was shown that carbohydrate and lipid content were significantly different

between fertile and non-fertile cysts [22]. Thus source and fertility of cysts seem to be critical for test outcome.

To overcome problems associated with native antigen, several recombinant antigens from both, *E. granulosus* and *E. multilocularis* have been produced and tested, for example recombinant *E. multilocularis* antigen II/3-10, which has a long history of successful application in diagnosis of human AE [23-25]. Recombinant *E. granulosus* B2t showed most promising results in the diagnosis of primary CE, but also in follow-up monitoring of treated CE patients in ELISA [26]. Recombinant *E. granulosus* elongation factor 1 beta/delta was applied in IgE immunoblot and was found positive in 33% of surgically confirmed CE cases, which previously had been tested negative in routine diagnostic assays [27].

The next step of further antigen development might consist in the chemical production of synthetic antigen. Through constant improvements of synthesis protocols, reagents and instruments, it has become possible to synthesize peptides of 100 - 150 amino acids in length on a routine basis [28]. Already much shorter peptides of about 20-25 amino acids in length were proven suitable antigens for the diagnosis of various infectious diseases, mainly of viral origin [29-31], but also of parasitic diseases [32-34].

Chemical synthesis of antigens may result in complete independence of native parasite material and biological expression systems and permits the production of GMP-grade test reagents. Benefits of synthetic peptides used as antigens are unlimited availability, stability and reproducibility.

About 90% of antibodies induced by natural infection are directed against conformational epitopes, i.e. epitopes made up of amino acids that are not necessarily contiguous in the sequence, but are brought into spatial proximity by the folding of the polypeptide chain [35]. An estimated 10% of antibodies are directed against linear epitopes made of continuous amino acids. Linear epitopes usually have flexible unfolded conformations and are located in large unstructured loops or terminal protein regions [28].

Prerequisite for antibody recognition by synthetic antigens is their ability to adopt native antigen conformation. Among the short and structurally stable protein domains that are able to fold into their native structure even as isolated peptides are the alpha-helical coiled-coils (CC) and the intrinsically unstructured regions (IUR). IURs do not exhibit any particular folding, while CC have been shown *in vitro* to readily assemble into two- or three-stranded structures [36,37].

Both, alpha-helical coiled-coils and intrinsically unstructured domains are frequent motifs. Possibly up to 30% of eukaryotic proteins are either completely or partially disordered [38]. Current estimates indicate that approximately 5-10% of sequences from the various genome projects encode coiled-coil regions [39]. A survey of GenBank entries revealed that in pathogens and parasites many extended coiled-coil domains (>75 amino acids in length) were external and binding to host proteins [40]. Detailed immunological studies showed alpha-helical coiled-coil proteins to possess antigenic characteristics [28,41,42]. Furthermore, the locations of alpha-helical coiled-coils and intrinsically unstructured regions can both readily be predicted by bioinformatic algorithms from the primary amino acid sequence of proteins.

Taken together, CC and IUR domains predicted from *Echinococcus granulosus* and *E. multilocularis* proteins appeared as ideal candidate motifs to be tested as synthetic peptide antigens during this PhD thesis.

General aim of this thesis

The principal aim of this study was the establishment of a new discovery pathway for diagnostic peptides. The peptides were selected *in silico* by bioinformatics means complemented with immunological analysis on peptide microarray, a high-throughput screening method. Using a microarray format offers the advantages of low reagent consumption and rapid multiplexed analysis. In particular for the development and validation of new diagnostic tests, serum-saving methods are essential, because well characterized sera are rare and thus highly precious.

The project made use of publicly available sequences from *Echinococcus granulosus* and *E. multilocularis*.

Specific objectives

i. Development of bioinformatic procedures to assess and select peptides from a genome-wide, a proteome-wide and a transcriptome-wide analysis.
ii. Establishment of a peptide microarray platform for the screening of diagnostic targets.
iii. Determination of optimal length of synthetic peptide antigens.
iv. Providing proof-of-principle for the suitability of the peptides selected for serological diagnosis of echinococcosis in ELISA format.

References

1. Wen H, New RRC, Craig PS: **Diagnosis and Treatment of Human Hydatidosis.** *British Journal of Clinical Pharmacology* 1993, **35:**565-574.

2. Romig T: **Epidemiology of echinococcosis.** *Langenbecks Archives of Surgery* 2003, **388:**209-217.

3. Rausch RL, Wilson JF, Schantz PM: **A Program to reduce the risk of infection by *Echinococcus multilocularis* - the use of praziquantel to control the cestode in a village in the hyperendemic region of alaska.** *Annals of Tropical Medicine and Parasitology* 1990, **84:**239-250.

4. Craig PS, Echinococcosis Working Group in China: **Epidemiology of human alveolar echinococcosis in China.** *Parasitology International* 2006, **55:**S221-S225.

5. Deplazes P, Hegglin D, Gloor S, Romig T: **Wilderness in the city: the urbanization of *Echinococcus multilocularis*.** *Trends in Parasitology* 2004, **20:**77-84.

6. Schweiger A, Ammann RW, Candinas D, Clavien PA, Eckert J, Gottstein B, Halkic N, Muellhaupt B, Prinz BM, Reichen J et al.: **Human alveolar echinococcosis after fox population increase, Switzerland.** *Emerging Infectious Diseases* 2007, **13:**878-882.

7. Berke O, Romig T, von Keyserlingk M: **Emergence of *Echinococcus multilocularis* among red foxes in northern Germany, 1991-2005.** *Veterinary Parasitology* 2008, **155:**319-322.

8. Torgerson PR, Schweiger A, Deplazes P, Pohar M, Reichen J, Ammann RW, Tarr PE, Halkik N, Müllhaupt B: **Alveolar echinococcosis: From a deadly disease to a well-controlled infection. Relative survival and economic analysis in Switzerland over the last 35 years.** *Journal of Hepatology* 2008, **49:**72-77.

9. Siles-Lucas M, Gottstein B: **Review: Molecular tools for the diagnosis of cystic and alveolar echinococcosis.** *Tropical Medicine & International Health* 2001, **6:**463-475.

10. Carmena D, Benito A, Eraso E: **Antigens for the immunodiagnosis of *Echinococcus granulosus* infection: An update.** *Acta Tropica* 2006, **98:**74-86.

11. Eckert J, Deplazes P: **Biological, epidemiological and clinical aspects of echinococcosis, a zoonosis of increasing concern.** *Clin Microbiol Rev* 2004, **17:**107-135.

12. Lorenzo C, Salinas G, Brugnini A, Wernstedt C, Hellman U, Gonzalez-Sapienza G: ***Echinococcus granulosus* antigen 5 is closely related to proteases of the trypsin family.** *Biochem J* 2003, **369:**191-198.

13. Dottorini S, Tassi C: ***Echinococcus granulosus*: Characterization of the main antigenic component (arc 5) of hydatid fluid.** *Experimental Parasitology* 1977, **43:**307-314.

14. Lorenzo C, Last JA, Gonzalez-Sapienza G: **The immunogenicity of *Echinococcus granulosus* antigen 5 is determined by its post-translational modifications.** *Parasitology* 2005, **131:**669-677.

15. Zhang WB, Li J, Jones MK, Zhang ZZ, Zhao L, Blair D, McManus DP: **The *Echinococcus granulosus* antigen B gene family comprises at least 10 unique genes in five subclasses which are differentially expressed.** *PLoS Neglected Tropical Diseases* 2010, **4**.

16. Frosch P, Hartmann M, Mühlschlegel F, Frosch M: **Sequence heterogeneity of the echinococcal antigen B.** *Molecular and Biochemical Parasitology* 1994, **64:**171-175.

17. Chemale G, Haag KL, Ferreira HB, Zaha A: ***Echinococcus granulosus* antigen B is encoded by a gene family.** *Molecular and Biochemical Parasitology* 2001, **116:**233-237.

18. Kamenetzky L, Muzulin PM, Gutierrez AM, Angel SO, Zaha A, Guarnera EA, Rosenzvit MC: **High polymorphism in genes encoding antigen B from human infecting strains of *Echinococcus granulosus*.** *Parasitology* 2005, **131:**805-815.

19. Shepherd JC, Aitken A, McManus DP: **A protein secreted in vivo by *Echinococcus granulosus* inhibits elastase activity and neutrophil chemotaxis.** *Molecular and Biochemical Parasitology* 1991, **44:**81-90.

20. Rigano R, Profumo E, Bruschi F, Carulli G, Azzara A, Ioppolo S, Buttari B, Ortona E, Margutti P, Teggi A et al.: **Modulation of human immune response by *Echinococcus granulosus* antigen B and its possible role in evading host defenses.** *Infection and Immunity* 2001, **69:**288-296.

21. Leggatt GR, Yang W, McManus DP: **Serological evaluation of the 12 kDa subunit of antigen B in *Echinococcus granulosus* cyst fluid by immunoblot analysis.** *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2003, **86:**189-192.

22. Irabuena O, Nieto A, Ferreira AM, Battistoni J, Ferragut G: **Characterization and optimization of bovine *Echinococcus granulosus* cyst fluid to be used in immunodiagnosis of hydatid disease by ELISA.** *Revista do Instituto de Medicina Tropical de São Paulo* 2000, **42:**255-262.

23. Vogel M, Gottstein B, Müller N, Seebeck T: **Production of a recombinant antigen of *Echinococcus multilocularis* with high immunodiagnostic sensitivity and specificity.** *Molecular and Biochemical Parasitology* 1988, **31:**117-125.

24. Müller N, Gottstein B, Vogel M, Flury K, Seebeck T: **Application of a recombinant *Echinococcus multilocularis* antigen in an enzyme-linked immunosorbent assay for immunodiagnosis of human alveolar echinococcosis.** *Molecular and Biochemical Parasitology* 1989, **36:**151-160.

25. Gottstein B, Jacquier P, Bresson-Hadni S, Eckert J: **Improved primary immunodiagnosis of alveolar echinococcosis in humans by an enzyme-linked immunosorbent assay using the Em2^{plus} antigen.** *J Clin Microbiol* 1993, **31:**373-376.

26. Hernandez-Gonzalez A, Muro A, Barrera I, Ramos G, Orduna A, Siles-Lucas M: **Usefulness of four different *Echinococcus granulosus* recombinant antigens for serodiagnosis of unilocular hydatid disease (UHD) and postsurgical follow-up of patients treated for UHD.** *Clin Diagn Lab Immunol* 2008, **15:**147-153.

27. Ortona E, Rigano R, Margutti P, Notargiacomo S, Ioppolo S, Vaccari S, Barca S, Buttari B, Profumo E, Teggi A et al.: **Native and recombinant antigens in the immunodiagnosis of human cystic echinococcosis.** *Parasite Immunology* 2000, **22:**553-559.

28. Corradin G, Villard V, Kajava AV: **Protein structure based strategies for antigen discovery and vaccine development against malaria and other pathogens.** *Endocrine Metabolic & Immune Disorders-Drug Targets* 2007, **7:**259-265.

29. Chan PKS, To WK, Liu EYM, Ng TK, Tam JS, Sung JJY, Lacroix JM, Houde M: **Evaluation of a peptide-based enzyme immunoassay for anti-SARS coronavirus IgG antibody.** *Journal of Medical Virology* 2004, **74:**517-520.

30. Zrein M, Joncas JH, Pedneault L, Robillard L, Dwyer RJ, Lacroix M: **Comparison of A Whole-Virus Enzyme-Immunoassay (Eia) with A Peptide-Based Eia for Detecting Rubella-Virus Immunoglobulin-G Antibodies Following Rubella Vaccination.** *Journal of Clinical Microbiology* 1993, **31:**1521-1524.

31. Alcaro MC, Peroni E, Rovero P, Papini AM: **Synthetic peptides in the diagnosis of HIV infection.** *Current Protein & Peptide Science* 2003, **4:**285-290.

32. Kong JT, Grigg ME, Uyetake L, Parmley S, Boothroyd JC: **Serotyping of *Toxoplasma gondii* infections in humans using synthetic peptides.** *Journal of Infectious Diseases* 2003, **187:**1484-1495.

33. Madhumathi J, Pradiba D, Prince P, Jeyaprita P, Rao D, Kaliraj P: **Crucial epitopes of *Wuchereria bancrofti* abundant larval transcript recognized in natural infection.** *European Journal of Clinical Microbiology & Infectious Diseases* 2010,1-6.

34. Intapan PM, Tantrawatpan C, Maleewong W, Wongkham S, Wongkham C, Nakashima K: **Potent epitopes derived from *Fasciola gigantica* cathepsin L1 in peptide-based immunoassay for the serodiagnosis of human fascioliasis.** *Diagnostic Microbiology and Infectious Disease* 2005, **53:**125-129.

35. Van Regenmortel MHV: **Antigenicity and immunogenicity of synthetic peptides.** *Biologicals* 2001, **29:**209-213.

36. Betz S, Fairman R, Oneil K, Lear J, Degrado W: **Design of two-stranded and three-stranded coiled-coil peptides.** *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 1995, **348:**81-88.

37. Su JY, Hodges RS, Kay CM: **Effect of chain-length on the formation and stability of synthetic alpha-helical coiled coils.** *Biochemistry* 1994, **33:**15501-15510.

38. Fink AL: **Natively unfolded proteins.** *Current Opinion in Structural Biology* 2005, **15:**35-41.

39. Walshaw J, Woolfson DN: **SOCKET: A program for identifying and analysing coiled-coil motifs within protein structures.** *Journal of Molecular Biology* 2001, **307:**1427-1450.

40. Odgren PR, Harvie LW, Fey EG: **Phylogenetic occurrence of coiled coil proteins: implications for tissue structure in metazoa via a coiled coil tissue matrix.** *Proteins-Structure Function and Genetics* 1996, **24:**467-484.

41. Jiz M, Wu HW, Meng R, Pond-Tor S, Reynolds M, Friedman JF, Olveda R, Acosta L, Kurtis JD: **Pilot-scale production and characterization of paramyosin, a vaccine candidate for schistosomiasis japonica.** *Infection and Immunity* 2008, **76:**3164-3169.

42. Tripet B, Kao DJ, Jeffers SA, Holmes KV, Hodges RS: **Template-based coiled-coil antigens elicit neutralizing antibodies to the SARS-coronavirus.** *Journal of Structural Biology* 2006, **155:**176-194.

PLoS Negl Trop Dis. 2010 August; 4(8): e771.
Published online 2010 August 3. doi: 10.1371/journal.pntd.0000771

Chapter 1:

Serodiagnosis of *Echinococcus* spp. infection: explorative selection of diagnostic antigen by peptide microarray

Claudia List[1,2], Weihong Qi[3], Eva Maag[4,2], Bruno Gottstein[5], Norbert Müller[5], Ingrid Felger[1,2]

1 Department of Medical Parasitology and Infection Biology, Swiss Tropical and Public Health Institute, Basel, Switzerland
2 University of Basel, Basel, Switzerland
3 Functional Genomics Center Zurich, ETH Zurich, Zurich, Switzerland
4 Department of Medical and Diagnostic Services, Swiss Tropical and Public Health Institute, Basel, Switzerland
5 Institute of Parasitology, University of Bern, Bern, Switzerland

OPEN ACCESS Freely available online

PLoS NEGLECTED TROPICAL DISEASES

Serodiagnosis of *Echinococcus* spp. Infection: Explorative Selection of Diagnostic Antigens by Peptide Microarray

Claudia List[1,2], Weihong Qi[3], Eva Maag[2,4], Bruno Gottstein[5], Norbert Müller[5], Ingrid Felger[1,2]*

1 Department of Medical Parasitology and Infection Biology, Swiss Tropical and Public Health Institute, Basel, Switzerland, **2** University of Basel, Basel, Switzerland, **3** Functional Genomics Center Zurich, ETH Zurich, Zurich, Switzerland, **4** Department of Medical and Diagnostic Services, Swiss Tropical and Public Health Institute, Basel, Switzerland, **5** Institute of Parasitology, University of Bern, Bern, Switzerland

Abstract

Background: Production of native antigens for serodiagnosis of helminthic infections is laborious and hampered by batch-to-batch variation. For serodiagnosis of echinococcosis, especially cystic disease, most screening tests rely on crude or purified *Echinococcus granulosus* hydatid cyst fluid. To resolve limitations associated with native antigens in serological tests, the use of standardized and highly pure antigens produced by chemical synthesis offers considerable advantages, provided appropriate diagnostic sensitivity and specificity is achieved.

Methodology/Principal Findings: Making use of the growing collection of genomic and proteomic data, we applied a set of bioinformatic selection criteria to a collection of protein sequences including conceptually translated nucleotide sequence data of two related tapeworms, *Echinococcus multilocularis* and *Echinococcus granulosus*. Our approach targeted alpha-helical coiled-coils and intrinsically unstructured regions of parasite proteins potentially exposed to the host immune system. From 6 proteins of *E. multilocularis* and 5 proteins of *E. granulosus*, 45 peptides between 24 and 30 amino acids in length were designed. These peptides were chemically synthesized, spotted on microarrays and screened for reactivity with sera from infected humans. Peptides reacting above the cut-off were validated in enzyme-linked immunosorbent assays (ELISA). Peptides identified failed to differentiate between *E. multilocularis* and *E. granulosus* infection. The peptide performing best reached 57% sensitivity and 94% specificity. This candidate derived from *Echinococcus multilocularis* antigen B8/1 and showed strong reactivity to sera from patients infected either with *E. multilocularis* or *E. granulosus*.

Conclusions/Significance: This study provides proof of principle for the discovery of diagnostically relevant peptides by bioinformatic selection complemented with screening on a high-throughput microarray platform. Our data showed that a single peptide cannot provide sufficient diagnostic sensitivity whereas pooling several peptide antigens improved sensitivity; thus combinations of several peptides may lead the way to new diagnostic tests that replace, or at least complement conventional immunodiagnosis of echinococcosis. Our strategy could prove useful for diagnostic developments in other pathogens.

Citation: List C, Qi W, Maag E, Gottstein B, Müller N, et al. (2010) Serodiagnosis of *Echinococcus* spp. Infection: Explorative Selection of Diagnostic Antigens by Peptide Microarray. PLoS Negl Trop Dis 4(8): e771. doi:10.1371/journal.pntd.0000771

Editor: Ana Flisser, Universidad Nacional Autónoma de México, Mexico

Received August 28, 2009; **Accepted** June 22, 2010; **Published** August 3, 2010

Funding: This study was supported by Velux Foundation (http://www.veluxstiftung.ch/), Roche Research Foundation (http://www.research-foundation.org/rrf/), Rudolf Geigy Foundation (http://www.sti.ch/en/about-us/foundations.html) and Fonds zur Förderung von Lehre und Forschung (http://www.fag-basel.ch/frameset.html). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: ingrid.felger@unibas.ch

Introduction

For serodiagnosis of human helminthic infections, many currently used tests rely on native antigens, either extracted from whole worms (somatic antigens) maintained in laboratory animals, or cultivated *in vitro* to obtain excretory/secretory products (metabolic antigens). These natural antigens are limited in availability and suffer from batch-to-batch variation. A number of *Echinococcus* proteins have been recombinantly produced and tested for use in serodiagnosis [1–6] but to our knowledge, only recombinant EmII/3-10 [7,8] and its related sequence Em18 [9] are successfully applied in commercial test kits. Recombinantly expressed antigens used in diagnostic tests require a high degree of purification to avoid cross-reactivity due to contaminants from the expression system. Unspecific binding and cross-reactivity are major problems with both, extracts of whole worms [10,11] and recombinant proteins [12]. For improving diagnostic test performance, it is desirable to identify highly specific and highly reactive epitopes from the proteome of the pathogen in question and synthetically produce the corresponding peptide antigens. Synthetic peptides are advantageous for diagnostic applications since they are well defined, easily produced in large amounts, highly pure and often cost-saving if compared to the production of natural antigen in animal models or *in vitro* culture. Applications of peptides in immunodiagnosis of different parasitic diseases were given by Noya et al. [13].

The availability of an increasing number of pathogen genomes is boosting basic research as well as applied science. In the field of parasitological diagnostics, sequencing of parasites genomes also creates new opportunities. Annotated genomes and proteomes are available for some of the medically important protozoan pathogens, such as *Plasmodium* species and some Kinetoplastida

Author Summary

Crude or purified, somatic or metabolic extracts of native antigens are routinely used for the serodiagnosis of human helminthic infections. These antigens are often cross-reactive, i.e., recognized by sera from patients infected with heterologous helminth species. To overcome limitations in antigen production, test sensitivity and specificity, chemically synthesized peptides offer a pure and standardized alternative, provided they yield acceptable operative characteristics. Ongoing genome and proteome work create new resources for the identification of antigens. Making use of the growing amount of genomic and proteomic data available in public databases, we tested a bioinformatic procedure for the selection of potentially antigenic peptides from a collection of protein sequences including conceptually translated nucleotide sequence data of *Echinococcus multilocularis and E. granulosus* (Plathyhelminthes, Cestoda). The *in silico* selection was combined with high-throughput screening of peptides on microarray and systematic validation of reactive candidates in enzyme-linked immunosorbent assay. Our study proved the applicability of this approach for selection of peptide antigens with good diagnostic characteristics. Our results suggested the pooling of several peptides to reach a high level of sensitivity required for reliable immunodiagnosis.

species. Following the genome of the parasitic nematode *Brugia malayi* [14], the genome of *Schistosoma mansoni*, a blood vessel dwelling trematode, has recently been released [15] as well as a draft genomic sequence for *Schistosoma japonicum* [16] and more genome data from various helminth species are to be expected in the near future (see for example http://www.sanger.ac.uk/Projects/Helminths/ or http://www.nematode.net/). Currently, proteins predicted from ESTs, whole genomes or contigs are available for data mining already from a considerable number of helminth species. Using the sequence data available in the public domain, we have designed and tested a pathway for identifying novel antigens for serodiagnostic test development.

Our selection procedure for peptide antigens relied on the application of bioinformatic filters to collections of protein sequences, including conceptually translated nucleotide sequence data. The selection criteria aimed at proteins located to the host-parasite interface. Such proteins are potentially seen by the host immune system and may elicit an immune response and thus represent good candidates of diagnostic antigens. Our *in silico* analysis therefore prioritized proteins containing a sorting signal directing the protein to the extracellular space (PSORT II [17]), transmembrane domains (TMHMM II [18]), or a C-terminal signature sequence for addition of glycosyl phosphatidylinositol (GPI) anchor (GPI-som [19]). These prediction algorithms require full length amino acid sequences with complete N- and C-termini, thus excluding the use of most EST libraries. The majority of proteins generally display well defined three dimensional structures, which is mostly globular. Our selection procedure preferred sequences in which the chemically synthesized peptide also adopts a natural conformation *in vitro*. There are two structural motifs meeting this claim. Firstly, the intrinsically unstructured regions (IUR) that lack a well defined three dimensional structure displaying an extended conformation that can be identified for example by IUPred [20]. Secondly the alpha-helical coiled-coil motif predicted for example by Paircoil2 [21]. Alpha-helical coiled-coils are believed to readily fold into a stable structure in aqueous solution [22]. A microarray platform served for testing the diagnostic potential of the peptides selected *in silico*. Using a high-throughput microarray format brings together the advantages of low reagent consumption and rapid multiplexed analysis. Particularly in diagnostics, the possibility of testing the reactivity of a given serum sample with multiple antigens simultaneously harbors great benefits. Promising candidates identified on microarray were further explored and validated for use in ELISA.

We have carried out the proof-of-principle for bioinformatic selection of suitable long synthetic peptides (LSPs) for the tapeworms *Echinococcus multilocularis* and *Echinococcus granulosus*, both of major medical importance causing alveolar (AE) and cystic echinococcosis (CE), respectively. CE and particularly AE, due to its infiltrative and/or space-occupying growth, are severe diseases with high fatality rate and poor prognosis if managed incorrectly [23]. The importance of alveolar echinococcosis is not represented by the number of reported cases but rather by the severity of the disease in the individual patient [24]. Standard diagnostic tests for human echinococcosis, AE as well as CE, imply imaging techniques such as ultrasonography, x-ray and computer tomography, as well as serological tests based on enzyme-linked immunosorbent assay (ELISA) and immunoblots [25]. Antigens used in these serodiagnostic tests were developed from crude extracts to purified fractions to recombinant antigens and for *E. multilocularis* to vesicular fluid originating from *in vitro* cultivated metacestodes [26]. Despite substantial improvements have been achieved, the main screening test still relies on the availability and quality of native antigens, i.e. hydatid fluid of *Echinococcus granulosus* cysts collected from naturally infected intermediate hosts at the slaughterhouse. One study reported that source and quality or fertility of cysts are critical for test outcome. This called for standardization of antigens and test methods [10]. To produce a robust and reproducible test for the routine diagnostic laboratory, we have evaluated the performance of chemically synthesized peptides in comparison to natural (EgHF, EM2) [26,27] and recombinant antigen (EmII/3-10) [7].

Materials and Methods

Ethics statement

Ethical clearance for retrospective use of anonymized patient sera for test development and quality control was obtained from the ethical committee (Ethikkommission beider Basel).

Human serum samples

Sera of healthy blood donors living in Switzerland were used to define a cut-off for distinguishing between positive and negative test results. In ELISA, 50 blood donor sera were used. In microarray, a single serum and 2 pools made of 5 sera each were used. For testing diagnostic peptide reactivity in ELISA, 44 sera from *E. multilocularis*- and 35 sera from *E. granulosus*-infected patients from Central Europe were used. All echinococcosis patients had active hepatic lesions of either CE1 or CE2 type (WHO-IWGE standardized classification) and all sera were sampled prior to any therapeutic intervention, i.e. before surgery and/or chemotherapy. Diagnoses were confirmed serologically as described by Müller et al. [26], in complementation to the clinical diagnosis based on imaging procedures and, if available, retrospective histopathological investigations. For testing cross-reactivity of the peptides, sera from patients with following infections were used (concomitant echinococcosis was ruled out by clinical and serological criteria): 2 trichinellosis, 2 trichuriasis, 10 toxocariasis, 8 ascariasis, 1 anisakiasis, 2 hookworm infection, 8 strongyloidiasis, 10 filariasis (*Loa loa*, *Mansonella perstans*, *Onchocerca volvulus*), 10 fascioliasis, 1 paragonimiasis, 11 schistosomiasis (*Schistosoma mansoni*, S. *haematobium*, S. *mekongi*), 12 neurocysticercosis

(*Taenia solium*), 1 taeniasis (*Taenia saginata*), 1 diphyllobothriasis and 26 amoebic liver abscess (*Entamoeba histolytica*).

Sequence selection and peptide design

Protein sequences of *Echinococcus multilocularis* and *E. granulosus* were retrieved from NCBI Entrez Protein Database on October 16, 2007. NCBI Entrez Protein Database are compiled from a variety of sources with daily updates, including SwissProt, PIR, PRF, PDB, and translations from annotated coding regions in GenBank and RefSeq. To identify putative membrane or extracellular proteins, the protein sequences were analyzed using TMHMM II [18] for prediction of transmembrane domains, GPI-SOM [19] for prediction of glycosylphosphatidylinisotol (GPI) anchor signals, and PSORT II [17] for prediction of protein subcellular localization. All protein sequences were further screened for stretches of alpha-helical coiled-coils using Paircoil2 [21] and intrinsically unstructured regions using IUPred [20]. Both of these structures are believed to adopt native conformation in aqueous solution and therefore constitute major selection criteria. Due to the limited number of *Echinococcus* sequences in NCBI Entrez Protein Database and few predicted surface proteins containing stretches of alpha-helical CC and/or IURs, the analysis was extended to all proteins with alpha-helical coiled-coil and IUR predictions, irrespective of their predicted locations. This included known proteins previously tested in immunodiagnosis. To narrow down protein regions for selection of 30mer peptides, protein sequences were subjected to prediction of coiled-coil stability by STABLECOIL (http://biomol.uchsc.edu/cores/biophysics/stablecoil). In case of IURs antigenicity predictions were performed using BepiPred [28]. Thus detected regions of high predicted CC stability or antigenicity were favored. To increase solubility, peptides were selected to start and end with a hydrophilic or neutral amino acid and to harbour on average one charged amino acid, either positive or negative, per five residues. In addition, the guidelines suggested by the manufacturer were followed (http://www.altabioscience.bham.ac.uk/pdfs/Intro_to_series_SYNTHETIC_PEPTIDES.pdf). From 11 sequences listed in table S1, 45 peptides were selected according to the criteria mentioned above.

Peptides were produced by Fmoc solid phase synthesis (Alta Bioscience, University of Birmingham, UK). The length of the peptides was limited by the EpiScan synthesis procedure (Alta Bioscience, University of Birmingham, UK) to a maximum of 30 amino acids. Thus the length of the peptides used for spotting onto microarrays ranged between 24 and 30 amino acids with an additional aminohexanoid acid (AHX) spacer and a biotin at the N-terminus. Biotin was required to bind peptides to a streptavidin-coated solid phase, i.e. microscope glass slide for microarrays (Alta Bioscience, University of Birmingham, UK) or to 96 well plates for ELISA. In order to remove electric charges from the free C-terminus, the synthetic peptides were modified by carboxy-terminal amidation. Thus, both N-terminal and C-terminal peptide ends were uncharged, mimicking natural segments of internal protein sequences.

For defining optimal peptide length, we designed extended variants of the 8 most reactive candidates from the microarray. The length of extended peptides ranged between 40 and 47 residues (table 1). Modifications at the N- and C-terminus were identical to those of the microarray peptides. The extension of peptides was chosen to increase alpha-helical coiled-coil stability or to combine epitopes from two single peptides. To target these improved peptides to *E. multilocularis* diagnosis, we chose the *E. multilocularis* sequences for designing longer peptides. Sequences derived from antigen B8/1 (accession number BAC77657) and antigen B8/2 (accession number BAD89809). To our knowledge, these were the first synthetic peptides from any *E. multilocularis* antigen B sequence that were serologically evaluated. Alignments of homologous antigen B8/1- and antigen B8/2-sequences of the two *Echinococcus* species are shown in figures S1 and S2, respectively. Antigen B8/1-sequences of *E. granulosus* and *E. multilocularis* showed 86% identity and antigen B8/2-sequences 93% identity. Peptide longD8-9 coincided with the region of a previously published *E. granulosus* epitope EVKYFFER [29], but differed from another published synthetic peptide p176 [30]: p176 spans the N-terminal region of antigen B8/1 (amino acids 17–54), while longD8-9 covered the central part of the protein (amino acids 33–74). Within the overlap of 26 amino acids, three amino acids differed.

Microarray

A microarray platform was used for screening *in silico* selected peptides for reactivity with patients' sera. The low-density peptide microarray (Alta Bioscience, University of Birmingham, UK) featured two blocks, each containing 45 *Echinococcus* peptide spots, a spotting control allowing positioning of grid for analysis (unrelated, TAMRA-labeled peptide), a control for serological detection (human IgG) and a blank (spotting solution). All microarray work was done at room temperature. Microarray slides were blocked for one hour in assay buffer (1×PBS pH 7.4, 0.05% Tween20, 3% milk powder). Individual human sera or serum pools were diluted 1:50 in assay buffer and incubated for two hours in a moist chamber. Slides were washed for 3×5 minutes in assay buffer and incubated for one hour with a 1:100 dilution of Cy5-labeled goat anti-human IgG (H+L) secondary antibody (Jackson ImmunoResearch Laboratories product number 109-175-088). Slides were washed for 3×5 min-

Table 1. Original and extended peptides of the 8 most reactive candidates from microarray.

Protein	Accession number	Microarray peptide	Length	Accession number	Extended peptide	Length	Peptide sequence
EmII/3	AAA50580	D1	29	AAA50580	longD1	43	EQKLRELRAQMVEKESDLADMKNKASAYESKIAELEMLLQQER
EmII/3	AAA50580	D12	25	AAA50580	longD12	40	DEVQREVEAQKVAMAKKEAEKAQAEAELRRMREKHDAKHK
AgB8/1	AAD38373	D8, D9	29	BAC77657	longD8-9	42	KMLGEMKYFFERDPLGQKLVDLLKELEEVFQMLRKKLRTALK
AgB8/2	AAC47169	D11	27	BAD89809	longD11	45	DPLGQRLVALGNDLTAICQKLQLKIREVLKKYVKNLVEEKDDDSK
EM13	Q07840	A9	29	Q07840	longA9	45	QVQNAKNEPFGTPEQLRKIEDKLRKGIMEEEKTRKAYEEALSSLS
PSCCP*	CAD44854	B6	25	CAD44854	longB6	47	ETIQSLCEHNAALQQKLDEANQSVTEVSVQMKVMQQHLHTARVAIQS
EG19	ABI24154	D3	27	ABI24154	longD3	42	EAEAKCLRRPHQRVVKEGEVSKGDEVDGEDRDDECVGGDEGR

*PSCCP: protoscolex-specific coiled-coil protein.
doi:10.1371/journal.pntd.0000771.t001

utes in assay buffer and quickly rinsed with deionized water before drying with compressed air. Slides were scanned at 532 and 635 nm using a GenePix 4100A microarray scanner (Bucher Biotec AG, Basel, Switzerland). Cy5 and TAMRA images and local background corrected fluorescence intensities (FI) were acquired using the Axon GenePix Pro 6.0 software. Further analysis was done in Microsoft Excel. Duplicate FI-values were averaged. Peptides were determined reactive if FI>1009. This cut-off was determined by testing single or pooled blood donor samples (pools made of 5 sera each), calculating average FI intensities plus 4 times the standard deviation. We chose 4 standard deviations rather than 3 because the majority of duplicate values differed in more than 20% from each other. This variation can be explained by unequal spot morphology. Peptides from reactive spots were forwarded to evaluation of test sensitivity and specificity performed by ELISA, the current standard format of routine serodiagnostics.

ELISA

96-well plates (NUNC Immobilizer Streptavidin) were pre-washed three times using an ELISA plate washer (deionized H_2O, 0.05% Tween20) and coated overnight at 4°C with 100 µl/well of synthetic peptide diluted to 2 µg ml^{-1} in PBS pH 7.4. All following steps were carried out at room temperature. After washing, the plates were blocked for one hour with 150 µl/well of assay buffer (3% milk in PBS pH 7.4, 0.05% Tween20, 0.5 mM biotin). The plates were washed and incubated for one hour with human sera diluted 1:200 in assay buffer (100 µl/well). Alkaline-phosphatase-conjugated goat anti-human IgG antibodies (Sigma product number A 3187) 1:1000 in assay buffer were used as secondary antibodies. After washing, 100 µl of conjugate dilution was added to each well and incubated for one hour, followed by a final wash. Wells were then incubated for 15 minutes with 100 µl of *p*-nitrophenyl phosphate (Sigma product number N 4645) at a concentration of 1 mg ml^{-1} in substrate buffer (13.2 mM Na_2CO_3, 35 mM $NaHCO_3$, 1mM $MgCl_2 \times 6H_2O$, pH 9.6). Absorbance values (A_{405nm}) were measured at 405 nm in a Tecan Sunrise microplate absorbance reader. All serum samples were tested in duplicates. The two values were averaged and blank-corrected. For distinguishing between positive and negative ELISA test results, 50 sera from healthy blood donors living in Switzerland were tested. The cut-off value was determined by the mean of the blood donor samples plus two standard deviations. Sensitivity and specificity of the single peptide candidates were calculated using test results from confirmed echinococcosis patients as true positives (TP). Blood donors and individuals infected by other helminths were taken as true negatives (TN). The working characteristics of the peptide antigens were explored by receiver-operating characteristics (ROC) plot analysis (Analyse-It, version 2.21).

To test the impact of pooling peptide candidates, LSP longD12, longD1 and longD8-9 were mixed in equal parts (mixW) and applied at the same conditions as described above. The cut-off was determined by the mean A_{405nm} of fifty blood donor samples plus 2 standard deviations. Sensitivity and cross-reactive behavior was tested with 30 echinococcosis sera and 15 sera from various helminth and amoeba infections that had previously shown cross-reactivity with the single peptides.

Results

Peptide selection

240 entries of *E. multilocularis* and 940 entries of *E. granulosus* proteins were retrieved from NCBI Entrez Protein Database including redundancies. Combining the outputs of PSORT II, TMHMM II, GPI-som, IUPred and Paircoil2, led to the selection of 11 protein candidates (table S1). According to the information available in GenBank, these proteins were isolated from metacestode or protoscolex tissue, with the possible exception of HSP70, for which no information on developmental stage was provided.

Six proteins had been reported previously to react with antibodies from patients' sera: EmII/3 [7], EM13 [6], P-29 [31], GRP [32], EG19 [33], AgB8/1 and AgB8/2 [1]. These earlier findings supported the inclusion of these antigens in our analysis. In total, 45 peptide sequences from 11 proteins were chosen to be chemically synthesized and spotted onto microarray. 31 peptides derived from *E. multilocularis* and 14 from *E. granulosus* (table S1).

Screening by peptide microarray

In silico selected peptides were screened for diagnostic potential on a microarray platform (Alta Bioscience, University of Birmingham, UK). From a total of 45 peptides contained in the microarray, 17 showed reactivity with pooled or single sera from echinococcosis patients (8/31 peptides from *E. multilocularis*, 9/14 peptides from *E. granulosus*). The peptides that reacted above the cut-off determined from blood donor sera (FI>1009) are listed in table 2. These 17 peptides were considered to have diagnostic potential and were therefore subjected to further testing of specificity and sensitivity. The predicted structure of the reactive peptide candidates was mostly alpha-helical coiled-coil (7 peptides) and alpha-helical (7 peptides). 3 peptides were predicted to be intrinsically unstructured.

Assessing peptide performance in ELISA

The 17 peptides reactive on microarray were carried forward for evaluating their application in ELISA-based serodiagnosis, the current routine technique. Test sera from AE and CE patients and sera from patients with infections other than echinococcosis were used. Sensitivity, cross-reactivity and impact of peptide length were evaluated (data not shown). With respect to sensitivity, long peptides (more than 30 residues) were superior to shorter ones. 3 out of 17 peptides were selected for extensive validation: longD12 (deriving from *E. multilocularis* EmII/3-10), longD1 (EmII/3-10) and longD8-9 (*E. multilocularis* antigen B8/1). All 3 peptides did not clearly differentiate between AE and CE infections as can be seen by overlapping standard deviations of A_{405nm} values obtained from AE and CE sera (figure 1). Peptides longD1 and longD12 derived from EmII/3-10, a well established antigen specific for the diagnosis of *E. multilocularis* infections [8]. In our hands, longD12 reacted with both, AE and CE sera (table 3: AE 10/44; CE 3/35). This contrasts with previous observations of a specific reactivity of EmII/3-10 with AE sera. The reactivity of peptide longD1 was similar to longD12 (AE: 11/44; CE: 2/35). Peptide longD8-9 (*E. multilocularis* antigen B8/1) reacted with more CE than AE sera (AE: 19/44; CE: 26/35) (table 3).

The cut-off values determined from 50 blood donors were similar for longD8-9 and longD1 with 0.165 and 0.200, respectively. The cut-off for longD12 was higher with an OD of 0.357. Compared to 90% sensitivity of a commercial *E. granulosus* hydatid fluid ELISA used in routine testing [34], the sensitivity of each single peptide alone was rather low, candidate longD8-9 being the most sensitive with 57%. LongD1 and longD12 both reached 16%. The specificities of longD8-9, longD1 and longD12 were 94%, 91% and 94% respectively, thus performing in the range required for a routine diagnostic assay. There were minor cross-reactivities observed with sera from helminth infections other than echinococcosis (table 3). Peptide longD8-9 was positive with 1/12 cysticercosis, 1/8 strongyloidiasis and 1/10 schistosomiasis

Table 2. Peptides with diagnostic potential identified on microarray.

	E.m.	*E.m.*	*E.m.*	*E.m.*	*E.m.*	*E.m.*	*E.m.*	*E.m.*	*E.g.*	*E.g.*	*E.g.*	*E.g.*	*E.g.*	*E.g.*	*E.g.*	*E.g.*	*E.g.*
Protein	**a)**	**a)**	**a)**	**b)**	**b)**	**b)**	**c)**	**c)**	**d)**	**d)**	**e)**	**e)**	**f)**	**f)**	**f)**	**g)**	**g)**
Peptide name	**A8**	**A9**	**A10**	**A4**	**A5**	**D12**	**B6**	**B9**	**C10**	**D2**	**D3**	**D5**	**D7**	**D8**	**D9**	**D10**	**D11**
Structure	**helical**	**helical**	**IUR**	**CC**	**CC**	**CC**	**CC**	**CC**	**helical**	**helical**	**IUR**	**IUR**	**helical**	**helical**	**helical**	**helical**	**CC**
Em-pool-1	1	1	0	0	0	1	1	0	0	1	0	1	1	1	0	0	0
Em-pool-2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Em-pool-3	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0
Em-pool-4*	1	1	1	0	0	1	1	1	0	0	1	0	0	1	0	0	1
Eg-pool-1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1
Em-3	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Em-11	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Em-12	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Em-23	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1
Em-24	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Em-2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Em-15	0	1	0	0	0	1	0	0	0	0	0	0	1	0	1	0	1
Em-18	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
Em-34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Em-35	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0

*Data incorporates merged results from repeated assays.
a) EM13 (accession Q07840); b) EmII/3 (accession AAA50580); c) Protoscolex specific coiled-coil protein (accession CAD44854); d) HSP70 (accession Q24789); e) EG19 (accession ABI24154); f) Antigen B8/1 (accession AAD38373); g) Antigen B8/2 (accession AAC47169).
1 = positive.
0 = neative.
Em = AE serum.
Eg = CE serum.
doi:10.1371/journal.pntd.0000771.t002

sera. Peptide longD1 was seropositive with 2/8 strongyloidiasis, 1/10 toxocariasis, 1/10 filariasis, 1/10 fascioliasis and 3/10 schistosomiasis sera. Peptide longD12 was positive with 1/8 strongyloidiasis, 1/10 toxocariasis, 2/10 schistosomiasis sera. Cross-reactions with *Taenia*, *Strongyloides* and filaria are also seen in conventional diagnostic assays. We further tested the specificity with sera from patients suffering from liver abscess due to *Entamoeba histolytica* infection. LongD12 was positive with 3/26, longD1 with 4/26 and longD8-9 with 3/26 amebiasis sera (table 3). To compare the diagnostic performance of each peptide, a receiver-operating characteristics (ROC) plot was generated (figure 2). Candidate longD8-9 emerged as clear favorite, although on its own not meeting the sensitivity required for antigens applied in routine diagnostics.

The use of 30–50mer peptides for serology impairs sensitivity compared to the full length antigen, because peptides likely harbor less epitopes. We calculated the cumulative sensitivity of the 3 single peptide ELISAs and measured the experimental combination of the 3 peptides in one antigen mix (mixW). The cut-off of mixW was 0.352, which is in the same range as the cut-off of one of its components, namely longD12. By cumulating the individual results of the 3 assays, the theoretical sensitivity indeed increased from 57% to 70%, but specificity decreased to 82% (table 4). Compared to cumulative positivity, pooling the 3 peptides in one well (mixW) led to a loss of positive test results. In return, the specificity increased due to a loss of cross-reactions. Positive test results obtained by single peptide ELISAs compared to pooled peptide ELISA are summarized in table 5.

Discussion

The use of synthetic peptides as substitutes of native antigen in immunoassays has been highly recommended for the diagnosis of infectious diseases by different authors [13,30,35]. Synthetic peptides are applied successfully in diagnosis of viral and bacterial infections [36–39]. Compared to a metazoan parasite, viruses and bacteria in general have smaller genomes and accomplish less complex post translational protein modifications, which might imply more rapid success in identification of immunodominant peptides. It even allows for analysis on whole proteome level [40]. Here, we investigated the use of synthetic peptides of 30–50 amino acids in length for the use in immunodiagnosis of human hydatid disease. We provided proof of principle for the selection of synthetic peptides by bioinformatic means and for screening on a microarray platform for diagnostic reactivity. 17 out of 45 peptides on the microarray proofed reactive with sera from echinococcosis patients.

The peptide performing best in our investigation was longD8-9, a 42mer peptide deriving from *E. multilocularis* antigen B8/1. It reached a sensitivity of 57% and a specificity of 94%. The reactivity of LSP longD8-9 with sera from AE as well as CE patients suggests its use for both, patient screening as well as for follow-up examinations of treated patients. For the latter purpose, it will be necessary to determine the proportion of patients seroconverting during treatment. The applicability of our peptide antigen for monitoring treatment follow-up samples is ongoing work in our laboratory.

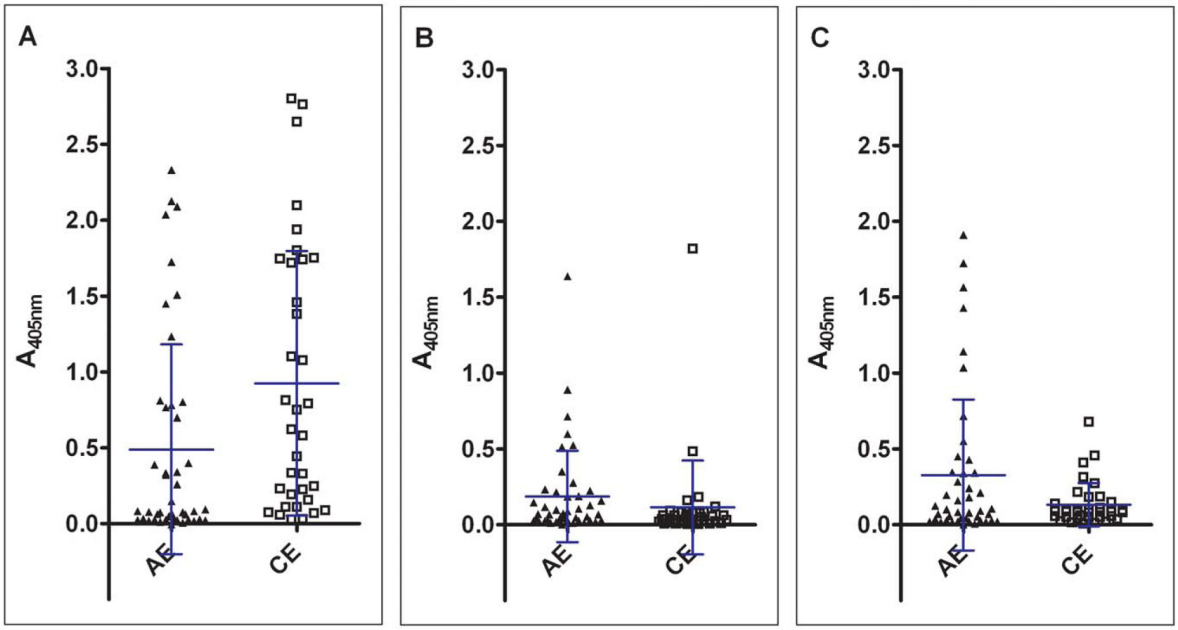

Figure 1. No clear discrimination of A_{405nm} values measured from alveolar echinococcosis (AE) and cystic echinococcosis (CE) sera. Following peptides served as antigens: 1A) longD8-9, cut-off 0.165. 1B) longD1, cut-off 0.200. 1C) longD12, cut-off 0.357. Bars represent mean and standard deviation. ▲ Alveolar echinococcosis (AE); □ Cystic echinococcosis (CE).
doi:10.1371/journal.pntd.0000771.g001

Table 3. Numbers of sera tested positive with peptides longD12, longD1 and longD8-9.

		Peptide	longD8-9	longD1	longD12	Cumulative result
		Cut-off A_{405nm}	0.165	0.200	0.357	
		Total tested	Positive	Positive	Positive	Positive
Healthy	Blood donors	50	4	4	4	11
Cestoda	AE	44	19	11	10	28
	CE	35	26	2	3	27
	Taeniasis	1	0	0	0	0
	Cysticercosis	12	1	0	0	1
	Diphyllobotriasis	1	0	0	0	0
Nematoda	Trichinellosis	2	0	0	0	0
	Trichuriasis	2	0	0	0	0
	Anisakiasis	1	0	0	0	0
	Hookworm	1	0	0	0	0
	Strongyloidiasis	8	1	2	1	2
	Toxocariasis	10	0	1	1	2
	Filariasis	10	0	1	0	1
	Ascariasis	8	0	0	0	0
Trematoda	Paragonimiasis	1	0	0	0	0
	Fascioliasis	10	0	1	0	1
	Schistosomiasis	10	1	3	2	5
Protozoa	Amoebiasis	26	3	4	3	8

doi:10.1371/journal.pntd.0000771.t003

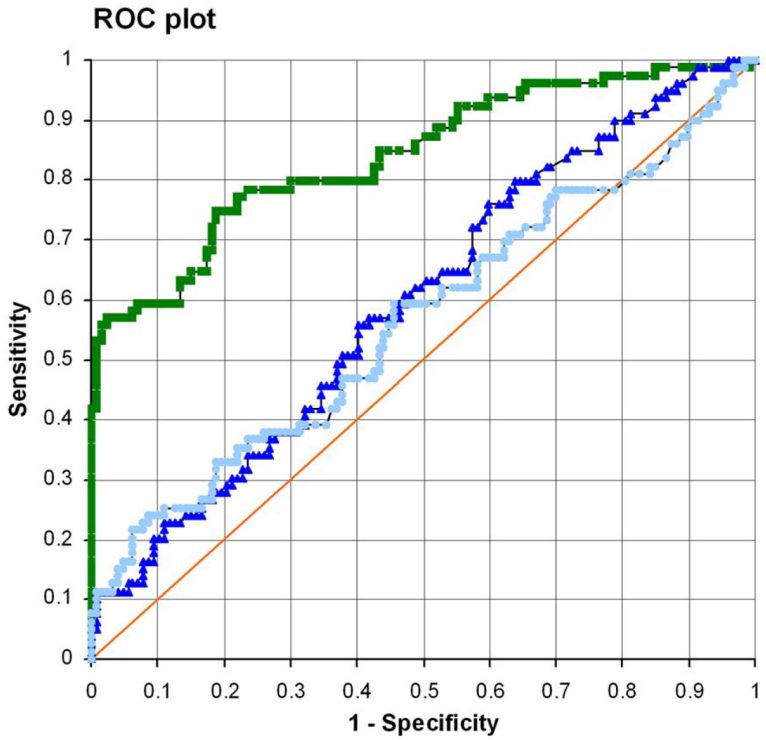

Figure 2. ROC plot illustrating sensitivity and specificity of LSP longD12, longD1, longD8-9. Diagonal line (orange): No discrimination; Square (green): longD8-9; Triangle (dark blue): longD1; Circle (light blue): longD12.
doi:10.1371/journal.pntd.0000771.g002

Our results confirm previous reports [38,41] that a test based on a single peptide needs to be complemented by additional peptides to reach a sensitivity comparable to the antigen extract. To compensate for limited sensitivity of individual peptides, we performed ELISAs with combinations of synthetic peptides applied as mixtures for coating. We compared these results to the "cumulative sensitivity", a theoretical value obtained by summing up the positive results from individual peptide ELISAs for each serum tested. The cumulative result amounted to a theoretical sensitivity of 70%. We did not detect this high sensitivity. In contrast, pooling of 3 peptides in one well led to a loss of positive signals otherwise obtained with individually tested peptides. 8/29 echinococcosis samples were lost through pooling. One reason was suggested by an increased cut-off for pooled peptides. In some sera similar ODs were observed in individual test and in peptides mixtures, but due to increased cut-off for pooled peptides, these sera turned negative. The loss of sensitivity obtained with pooled peptides in mixW could also be due to partial occupation of streptavidin binding sites with less reactive antigen. The same reasons may account for increased specificity. Combination of peptides in one well led to a loss of unwanted cross-reactivity with 5/11 previously false positive signals disappearing. Thus in our study, pooling peptides increased specificity, but lowered sensitivity. In future work peptide mixes might be optimized by combining peptides with equally good diagnostic sensitivity and specificity. In particular, the sensitivity of mixes is likely increased by continuous search for new peptides with improved diagnostic operating characteristics as substitutes for the least reactive candidates. The behaviour of peptide mixtures applied as antigens in ELISA needs further investigation.

Table 4. Sensitivity and specificity of peptides longD12, longD1 and longD8-9.

Peptide	total	AE+CE	TP	FN	TN	FP	Sensitivity	Specificity
longD8-9	79		45	34	120	7	57%	94%
longD1	79		13	66	115	12	16%	91%
longD12	79		13	66	127	8	16%	94%
cummulative	79		55	24	104	23	70%	82%

TP = true positive, FN = false negative, TN = true negative, FP = false positive.
doi:10.1371/journal.pntd.0000771.t004

To select peptides with broad recognition by human echinococcosis patients, our bioinformatic selection aimed at exported or parasite surface exposed antigens. Such proteins are likely to elicit an immune response. A major drawback was the scarce number of full-length *Echinococcus* spp. protein sequences deposited in public databases. Thus it might be worthwhile for further studies to adapt the bioinformatic selection criteria to the use of partial EST sequences. Our data points out the need for screening more peptide candidates. Further investigations on whole genome level are likely to provide additional candidates with sufficient sensitivity and specificity, which can eventually be combined to a synthetic antigen pool. Single or pooled LSPs are compatible with high throughput platform technologies such as luminex or biacore technology that could replace ELISA in the future.

Table 5. Comparison of positivity achieved by single peptides, cumulative results of single peptides and peptide pool mixW.

	longD8-9	longD1	longD12	cumulative result	mixW
Em-1	1	1	0	1	1
Em-5	1	0	0	1	0
Em-9	1	1	0	1	1
Em-14	1	1	0	1	1
Em-27	1	0	0	1	1
Em-31	0	0	1	1	0
Em-38	0	1	0	1	1
Em-39	1	0	1	1	1
Em-40	1	0	0	1	0
Em-41	1	1	0	1	1
Em-42	0	1	0	1	1
Em-44	1	0	0	1	1
Em-45	0	1	1	1	1
Em-47	0	0	1	1	1
Em-53	1	0	0	1	1
Eg-1	1	0	1	1	1
Eg-2	1	0	0	1	0
Eg-3	1	0	0	1	0
Eg-11	1	0	0	1	1
Eg-12	0	0	0	0	0
Eg-13	1	0	0	1	0
Eg-14	1	1	0	1	1
Eg-22	1	0	0	1	1
Eg-24	1	0	0	1	0
Eg-25	1	0	1	1	1
Eg-28	1	0	0	1	1
Eg-29	1	0	0	1	1
Eg-31	1	0	0	1	1
Eg-32	1	0	0	1	1
Eg-34	0	1	0	1	0
N-7	1	1	1	1	1
N-24	0	0	1	1	1
N-25	0	1	0	1	0
N-35	0	0	0	0	0
T-15	0	0	0	0	0
T-16	0	0	1	1	1
T-17	1	1	0	1	0
T-20	0	1	0	1	0
C-10	1	0	0	1	0
C-12	0	0	0	0	0
C-14	0	0	0	0	0
A-4	0	1	1	1	1
A-23	1	0	0	1	0
A-24	1	0	0	1	1
A-26	0	0	1	1	1

1 = positive.
0 = negative.
Em = *E. multilocularis* infection.
Eg = *E. granulosus* infection.
N = nematoda infection.
C = cestoda infection.
T = trematoda infection.
A = amoeba infection.
doi:10.1371/journal.pntd.0000771.t005

Additionally to surface or extracellular localization, one criterion during bioinformatic selection was prediction of secondary structures. Alpha-helical coiled-coils and intrinsically unstructured proteins are believed to adopt native conformation in aqueous solution and might therefore be well suited for immunoassays. It is widely accepted that the majority of epitopes are of the conformational type [42]. Our results showed that for diagnostic purposes the predicted alpha-helical and alpha-helical coiled-coil peptides were superior to peptides deriving from intrinsically unstructured regions. Among the serum-reactive peptides identified by microarray screening, 14 peptides were of predicted alpha-helical organization, while only 3 peptides derived from intrinsically unstructured regions. Furthermore, during our extensive testing of candidates, all IUR peptides were dismissed due to lack of sensitivity.

In addition to the secondary and tertiary structure of a peptide, a further determinant for recognition is the immobilization of peptides to the solid phase used in an immunodiagnostic test. Antibody capture by peptides crucially depends on the accessibility of key residues and therefore the introduction of a spacer molecule to spatially separate the peptide from its carrier protein is vitally important [43]. The peptides in our study were synthesized with an N-terminal AHX-spacer coupled to biotin. The biotin was used to immobilize the peptide antigen to a streptavidin-coated solid phase, i.e. microscope glass slides and 96-well ELISA plates. Coating the biotinylated peptides to non-streptavidin surfaces such as Immulon 2HB (Thermo Scientific) or Poly Sorp (NUNC) ELISA plates, led in those samples tested to discrepancies of duplicate values and mostly reduced A_{405nm} values, in a non-linear manner (data not shown). Immobilization of peptides by direct adsorbance to the polystyrene surface might lead to sterical blockage of key residues necessary for interaction with antibodies and thus decrease the A_{405nm} values. Similarly, HIV-1 peptides showed in ELISA increased test sensitivity and specificity when immobilized via biotin-streptavidin [44]. Surfaces functionalized with streptavidin guarantee a directional, highly dense and reproducible coating of biotinylated peptide antigens. It has been shown that the direction of peptide immobilization is of secondary importance as long as peptides deriving from internal protein sequences are investigated [43]. Interactions of internal sequences do not depend on free N- or C-terminus and therefore attachment of spacer and carrier molecules can be achieved successfully at both termini of the peptide.

Our approach represents an alternative to peptide selection by phage display [45,46]. The bioinformatic selection avoids construction and handling of phage display libraries and panning procedures. The most important limitation to successful panning consists in the lack of selective antibodies. Antibody purification from human sera requires pure and specific antigen, which generally is not yet available. Our bioinformatic approach to peptide selection reduces complex lab work and is compatible with screening on peptide microarray. In our hands, this platform proved highly suitable for investigation of antigen-antibody interaction.

Supporting Information

Figure S1 Alignment of *E. granulosus* and *E. multilocularis* antigenB8/1 sequences with location of peptide D8, D9 and longD8-9.
Found at: doi:10.1371/journal.pntd.0000771.s001 (0.42 MB TIF)

Figure S2 Alignment of *E. granulosus* and *E. multilocularis* antigenB8/2 sequences with location of D11 and longD11.
Found at: doi:10.1371/journal.pntd.0000771.s002 (0.44 MB TIF)

Table S1 Final set of peptides spotted onto microarray. Found at: doi:10.1371/journal.pntd.0000771.s003 (0.12 MB DOC)

Author Contributions

Conceived and designed the experiments: CL NM IF. Performed the experiments: CL EM NM. Analyzed the data: CL WQ EM NM IF. Contributed reagents/materials/analysis tools: BG IF. Wrote the paper: CL IF. Carried out the bioinformatic analysis: WQ.

References

1. Rott MB, Fernandez V, Farias S, Ceni J, Ferreira HB, et al. (2000) Comparative analysis of two different subunits of antigen B from *Echinococcus granulosus*: gene sequences, expression in *Escherichia coli* and serological evaluation. Acta Trop 75: 331–340.
2. Mamuti W, Yamasaki H, Sako Y, Nakao M, Xiao N, et al. (2004) Molecular cloning, expression, and serological evaluation of an, 8-kilodalton subunit of antigen B from *Echinococcus multilocularis*. J Clin Microbiol 42: 1082–1088.
3. Lorenzo C, Last JA, Gonzalez-Sapienza GG (2005) The immunogenicity of *Echinococcus granulosus* antigen 5 is determined by its post-translational modifications. Parasitology 131: 669–677.
4. Li J, Zhang WB, Wilson M, Ito A, McManus DP (2003) A Novel Recombinant Antigen for Immunodiagnosis of Human Cystic Echinococcosis. Jour Infect Dis 188: 1951–1960.
5. Hernández-González A, Muro A, Barrera I, Ramos G, Orduña A, et al. (2008) Usefulness of four different *Echinococcus granulosus* recombinant antigens for serodiagnosis of unilocular hydatid disease (UHD) and postsurgical follow-up of patients treated for UHD. Clin Vaccine Immunol 15: 147–153.
6. Frosch PM, Geier C, Kaup FJ, Müller A, Frosch M (1993) Molecular cloning of an echinococcal microtrichal antigen immunoreactive in *Echinococcus multilocularis* disease. Mol Biochem Parasitol 58: 301–310.
7. Vogel M, Gottstein B, Müller N, Seebeck T (1988) Production of a recombinant antigen of *Echinococcus multilocularis* with high immunodiagnostic sensitivity and specificity. Mol Biochem Parasitol 31: 117–125.
8. Müller N, Gottstein B, Vogel M, Flury K, Seebeck T (1989) Application of a recombinant *Echinococcus multilocularis* antigen in an enzyme-linked immunosorbent assay for immunodiagnosis of human alveolar echinococcosis. Mol Biochem Parasitol 36: 151–160.
9. Sako Y, Nakao M, Nakaya K, Yamasaki H, Gottstein B, et al. (2002) Alveolar echinococcosis: Characterization of diagnostic antigen Em18 and serological evaluation of recombinant Em18. J Clin Microbiol 40: 2760–2765.
10. Poretti D, Felleisen E, Grimm F, Pfister M, Teuscher F, et al. (1999) Differential immunodiagnosis between cystic hydatid disease and other cross-reactive pathologies. Am J Trop Med Hyg 60: 193–198.
11. Yamano K, Goto A, Nakamura-Uchiyama F, Nawa Y, Hada N, et al. (2009) Gal beta 1-6Gal, antigenic epitope which accounts for serological cross-reaction in diagnosis of *Echinococcus multilocularis* infection. Parasite Immunol 31: 481–487.
12. Müller N, Felleisen R (1995) Bacterial expression systems as tools for the production of immunodiagnostic parasite antigens. Parasitol Today 11: 476–480.
13. Noya O, Patarroyo ME, Guzman F, de Noya BA (2003) Immunodiagnosis of parasitic diseases with synthetic peptides. Curr Protein Pept Sci 4: 299–308.
14. Ghedin E, Wang S, Foster JM, Slatko BE (2004) First sequenced genome of a parasitic nematode. Trends Parasitol 20: 151–153.
15. Berriman M, Haas BJ, LoVerde PT, Wilson RA, Dillon GP, et al. (2009) The genome of the blood fluke *Schistosoma mansoni*. Nature 460: 352–358.
16. Zhou Y, Zheng HJ, Chen YY, Zhang L, Wang K, et al. (2009) The *Schistosoma japonicum* genome reveals features of host-parasite interplay. Nature 460: 345–U56.
17. Nakai K, Horton P (1999) PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends Biochem Sci 24: 34–35.
18. Krogh A, Larsson Br, von Heijne G, Sonnhammer ELL (2001) Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. J Mol Biol 305: 567–580.
19. Fankhauser N, Mäser P (2005) Identification of GPI anchor attachment signals by a Kohonen self-organizing map. Bioinformatics 21: 1846–1852.
20. Dosztanyi Z, Csizmok V, Tompa P, Simon I (2005) IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics 21: 3433–3434.
21. McDonnell AV, Jiang T, Keating AE, Berger B (2006) Paircoil2: improved prediction of coiled coils from sequence. Bioinformatics 22: 356–358.
22. Kohn WD, Hodges RS (1998) De novo design of alpha-helical coiled coils and bundles: models for the development of protein-design principles. Trends Biotechnol 16: 379–389.
23. McManus DP, Zhang WB, Li J, Bartley PB (2003) Echinococcosis. Lancet 362: 1295–1304.
24. Gottstein B (1992) Molecular and immunological diagnosis of echinococcosis. Clin Microbiol Rev 5: 248–261.
25. Eckert J, Deplazes P (2004) Biological, epidemiological and clinical aspects of echinococcosis, a zoonosis of increasing concern. Clin Microbiol Rev 17: 107–135.
26. Müller N, Frei E, Nuñez S, Gottstein B (2007) Improved serodiagnosis of alveolar echinococcosis of humans using an *in vitro* produced *Echinococcus multilocularis* antigen. Parasitology 134: 879–888.
27. Gottstein B (1985) Purification and characterization of a specific antigen from *Echinococcus multilocularis*. Parasite Immunol 7: 201–212.
28. Larsen J, Lund O, Nielsen M (2006) Improved method for predicting linear B-cell epitopes. Immunome Res 2: 2.
29. González-Sapienza G, Cachau RE (2003) Identification of critical residues of an immunodominant region of *Echinococcus granulosus* antigen B. J Biol Chem 278: 20179–20184.
30. González-Sapienza G, Lorenzo C, Nieto A (2000) Improved immunodiagnosis of cystic hydatid disease by using a synthetic peptide with higher diagnostic value than that of its parent protein, *Echinococcus granulosus* antigen B. J Clin Microbiol 38: 3979–3983.
31. González G, Spinelli P, Lorenzo C, Hellman U, Nieto A, et al. (2000) Molecular characterization of P-29, a metacestode-specific component of *Echinococcus granulosus* which is immunologically related to, but distinct from, antigen 5. Mol Biochem Parasitol 105: 177–185.
32. Mühlschlegel F, Frosch P, Castro A, Apfel H, Müller A, et al. (1995) Molecular cloning and characterization of an *Echinococcus multilocularis* and *Echinococcus granulosus* stress protein homologous to the mammalian 78 kDa glucose regulated protein. Mol Biochem Parasitol 74: 245–250.
33. Delunardo F, Ortona E, Margutti P, Perdicchio M, Vacirca D, et al. (2010) Identification of a novel 19kDa *Echinococcus granulosus* antigen. Acta Trop 113: 42–47.
34. Auer H, Stöckl C, Suhendra S, Schneider R (2009) Sensitivity and specificity of new commercial tests for the detection of specific *Echinococcus* antibodies. Wien Klin Wochensch 121: 37–41.
35. Gomara MJ, Haro I (2007) Synthetic peptides for the immunodiagnosis of human diseases. Curr Med Chem 14: 531–546.
36. Manocha M, Chitralekha KT, Thakar M, Shashikiran D, Paranjape RS, et al. (2003) Comparing modified and plain peptide linked enzyme immunosorbent assay (ELISA) for detection of human immunodeficiency virus type-1 (HIV-1) and type-2 (HIV-2) antibodies. Immunol Lett 85: 275–278.
37. Fachiroh J, Paramita DK, Hariwiyanto B, Harijadi A, Dahlia HL, et al. (2006) Single-assay combination of Epstein-Barr virus (EBV) EBNA1-and viral capsid antigen-p18-derived synthetic peptides for measuring anti-EBV immunoglobulin G (IgG) and IgA antibody levels in sera from nasopharyngeal carcinoma patients: Options for field screening. J Clin Microbiol 44: 1459–1467.
38. Shen GM, Behera D, Bhalla M, Nadas A, Laal S (2009) Peptide-Based Antibody Detection for Tuberculosis Diagnosis. Clin Vaccine Immunol 16: 49–54.
39. Morré SA, Munk C, Persson K, Krüger-Kjaer S, van Dijk R, et al. (2002) Comparison of three commercially available peptide-based immunoglobulin G (IgG) and IgA assays to microimmunofluorescence assay for detection of *Chlamydia trachomatis* antibodies. J Clin Microbiol 40: 584–587.
40. Zhu H, Hu SH, Jona G, Zhu XW, Kreiswirth N, et al. (2006) Severe acute respiratory syndrome diagnostics using a coronavirus protein microarray. Proc Natl Acad Sci U S A 103: 4011–4016.
41. Hernández M, Beltrán C, Garcia E, Fragoso G, Gevorkian G, et al. (2000) Cysticercosis: towards the design of a diagnostic kit based on synthetic peptides. Immunol Lett 71: 13–17.
42. Van Regenmortel MHV (2001) Antigenicity and immunogenicity of synthetic peptides. Biologicals 29: 209–213.
43. Andresen H, Grotzinger C, Zarse K, Kreuzer OJ, Ehrentreich-Forster E, et al. (2006) Functional peptide microarrays for specific and sensitive antibody diagnostics. Proteomics 6: 1376–1384.
44. Ivanov VS, Suvorova ZK, Tchikin LD, Kozhich AT, Ivanov VT (1992) Effective method for synthetic peptide immobilization that increases the sensitivity and specificity of ELISA procedures. J Immunol Methods 153: 229–233.
45. Ouyang L, Yi XY, Zeng XF, Zhou JC, Wang QL, et al. (2003) Partial protection induced by phage library-selected peptides mimicking epitopes of *Schistosoma japonicum*. Chin Med J 116: 138–141.
46. Hell RCR, Amim P, de Andrade HM, de Avila RAM, Felicori L, et al. (2009) Immunodiagnosis of human neurocysticercosis using a synthetic peptide selected by phage-display. Clin Immunol 131: 129–138.

Supporting information

Figure S1:

Alignment of *E. granulosus* and *E. multilocularis* antigen B8/1 sequences with location of peptide D8, D9 and longD8-9.

Figure S2:

Alignment of *E. granulosus* and *E. multilocularis* antigen B8/2 sequences with location of D11 and longD11.

Table S1: Final set of peptides spotted onto microarray.

Species	Protein	Accession number	Peptide name	Peptide sequence*	Length (amino acids)	Structure
E. multilocularis	EmII/3	AAA50580	A2	PIDRLITSKEQYDQTDEQWYERIIAYYKDH	30	helical
E. multilocularis	EmII/3	AAA50580	A3	EKSKINKRILALCTGNHELYMRRRKSDS	28	helical
E. multilocularis	EmII/3	AAA50580	A4	KEAERQRLKEERLQRMENEQKLRELRAQ	28	CC
E. multilocularis	EmII/3	AAA50580	D1	RAQMVEKESDLADMKNKASAYESKIAELE	29	CC
E. multilocularis	EmII/3	AAA50580	D12	KKEAEKAQAEAELRRMREKHDAKHK	25	CC
E. multilocularis	EmII/3	AAA50580	A5	NVRRTEESRVTAVSKNETLQTKLANLKME	29	CC
E. multilocularis	EmII/3	AAA50580	A6	STRDQSKMRDIDRRHEYNVREGNDKYKTLR	30	IUR
E. multilocularis	EM13	Q07840	A7	SIKEVKNFDSEFENAQKTWYKHYKNVNR	28	helical
E. multilocularis	EM13	Q07840	A8	HACKTVRSLQVQVQNAKNEPFGTPEQLRK	29	helical
E. multilocularis	EM13	Q07840	A9	EQLRKIEDKLRKGIMEEEKTRKAYEEALS	29	CC
E. multilocularis	EM13	Q07840	A10	HTTAYGSNSYDHGSEGATPSDYTS	24	IUR
E. multilocularis	EM13	Q07840	A11	DELSFNSGDLFEKLEDEDEQGWCKGRKD	28	CC
E. multilocularis	antigen 6	AAB61984	B1	DAFQKNTEKITTTDKLGTALEQVASQSEK	29	helical
E. multilocularis	antigen 6	AAB61984	B2	EAQKAKTKLEEVRLDLDSDKTKLKNAKTAE	30	CC
E. multilocularis	antigen 6	AAB61984	B3	KNAKTAEQKAKWEAEVRKDESDFDRVHQES	30	helical
E. multilocularis	protoscolex specific coiled-coil protein	CAD44854	B4	RQENQRLFEQFCQQIHNIQREKETVRLR	28	helical
E. multilocularis	protoscolex specific coiled-coil protein	CAD44854	B5	EIINLRGEVQQQKRRSGQRTQEHAETIQS	29	CC
E. multilocularis	protoscolex specific coiled-coil protein	CAD44854	B6	EHNAALQQKLDEANQSVTEVSVQMK	25	CC
E. multilocularis	protoscolex specific coiled-coil protein	CAD44854	B7	ESLRAEINHLKEDKATLEKKLQAIMDERDS	30	CC
E. multilocularis	protoscolex specific coiled-coil protein	CAD44854	B8	EDDWWYKHVHLDPSMTGFYEDDEIEIDDAS	30	CC
E. multilocularis	protoscolex specific coiled-coil protein	CAD44854	B9	SDLECNDSGAGGSTEEEFLRSFRTEVAE	28	helical
E. multilocularis	glucose regulated protein	Q24895	B10	DDRAVQKLRREVEKAKRTLSTEHSTMIEID	30	helical
E. multilocularis	glucose regulated protein	Q24895	B11	SAEDKGTGKKSNIVINKETNRLTPEEIER	29	IUR
E. multilocularis	glucose regulated protein	Q24895	B12	EIERMIQDAEKFSDQDKQVKERVEVRNDLE	30	IUR
E. multilocularis	glucose regulated protein	Q24895	C1	QVKDKEKMGGKLSDDEIKTIEDAADEAIK	29	IUR
E. multilocularis	glucose regulated protein	Q24895	C2	AIKWMENNPQAETSDYKKQKANLES	25	IUR

Table S1: continued

Species	Protein	Accession number	Peptide name	Peptide sequence*	Length (amino acids)	Structure
E. multilocularis	tropomyosin	CAC85552	C3	DFEKKEEEMNDWLSKVKNIQTEVDTVQES	29	CC
E. multilocularis	tropomyosin	CAC85552	C4	KLEETEKRATNAEAEVAAMTRRIRLLEED	29	CC
E. multilocularis	tropomyosin	CAC85552	C5	TKLDDASKAAEESERNRKTLETRSISDDER	30	CC
E. multilocularis	tropomyosin	CAC85552	C6	RKYDEAARRLAVTEVDLERAESRLETSESK	30	CC
E. multilocularis	tropomyosin	CAC85552	C8	ERLKTAEQRAAEAERQVSKLQNEVDRLEDE	30	CC
E. granulosus	HSP70	Q24789	C9	EVKSTAGDTHLGGEDFDSRLVNHFVEEFKR	30	helical
E. granulosus	HSP70	Q24789	C10	ELCSDLFRSTLDPVEKALRDAKLDKGAVHE	30	helical
E. granulosus	HSP70	Q24789	C11	SAVDKSTGKQNKITITRDKGRLSKEEIER	29	IUR
E. granulosus	HSP70	Q24789	C12	KSTVEDEKVKEKIGESDRRRIMEKCEETVK	30	IUR
E. granulosus	HSP70	Q24789	D2	GNQQAEKEEYEHRQKELESVCNPIIAK	27	helical
E. granulosus	EG19 antigen	ABI24154	D3	EAEAKCLRRPHQRVVKEGEVSKGDEVD	27	IUR
E. granulosus	EG19 antigen	ABI24154	D4	HEVSHEGKQESEDKDADKIAIEGVVRK	27	IUR
E. granulosus	EG19 antigen	ABI24154	D5	AHLGTGKSQHADEKALFYEEEAEDEGEDDE	30	IUR
E. granulosus	P-29	AAD53328	D6	KAAPQLSKMLTEASDVHQRMATARKNFNSE	30	IUR
E. granulosus	antigen B8/1	AAD38373	D7	QADDGLTSTSRSVMKMFGEVKYFFERDPLG	30	helical
E. granulosus	antigen B8/1	AAD38373	D8	RSVMKMFGEVKYFFERDPLGQKVVDLLKE	29	helical
E. granulosus	antigen B8/1	AAD38373	D9	DLLKELEEVFQLLRKKLRMALRSHLRG	27	CC
E. granulosus	antigen B8/2	AAC47169	D10	QAKDEPKAHMGQVVKKRWGELRDFFRNDPL	30	helical
E. granulosus	antigen B8/2	AAC47169	D11	QKLQLKIREVLKKYVKNLVEEKDDDSK	27	CC

* Additionally to the sequence listed, the peptides carried biotin and AHX-spacer at the N-terminus and were amidated at the C-terminus.

Chapter 2:

Mass spectrometry based design of synthetic antigens for the diagnosis of human echinococcosis

Claudia List [1,2], Ingrid Felger [1,2], Norbert Müller [3]

[1] Department of Medical Parasitology and Infection Biology, Swiss Tropical and Public Health Institute, Basel, Switzerland

[2] University of Basel, Basel, Switzerland

[3] Institute of Parasitology, University of Bern, Bern, Switzerland

Manuscript in preparation for publication.

Abstract

The use of chemically synthesized antigens for infectious disease diagnostics represents a promising alternative to current tests based on crude or partially purified antigen extracts. Benefits of synthetic peptides are unlimited availability, stability and reproducibility. We aimed at characterizing proteins in the 20-22 kDa protein fraction of *in vitro* cultivated *Echinococcus multilocularis* vesicle fluid by mass spectrometry. From thus identified proteins 30mer peptides were bioinformatically selected and screened for reactivity with sera from echinococcosis patients. We identified 34 proteins in this fraction and selected 41 peptides for synthesis. Of these 14 were reactive with echinococcosis sera, 12 were omitted due to reactivity with echinococcosis negative control sera. The resulting 2 validated peptides individually do not present sufficient sensitivity, but have potential for inclusion in multiplexed assays.

Introduction

Cystic and alveolar echinococcosis in humans are caused by the metacestode larvae of the closely related tapeworms *Echinococcus granulosus* and *E. multilocularis*. Successful establishment of the parasite leads to cyst development in liver, lungs or other organs. Cysts caused by *E. granulosus* are well-delineated, unilocular and fluid-filled. They consist of an inner germinal layer surrounded by a non-cellular laminated layer and host-produced fibrous adventitial tissue [1]. In contrast, *E. multilocularis* causes the growth of a dispersed, spongy metacestode tissue consisting of scattered, small cysts without limiting host-tissue barrier [1,2]. Early diagnosis is important, because radical removal of metacestode tissue by surgery is considered the most successful treatment. But diagnosis very often is delayed because cysts develop slowly and little specific symptoms, such as abdominal pain, nausea, jaundice and feeling of exhaustion. Up to 50% of AE and CE cases may remain asymptomatic and parasite lesions would incidentally be detected during examinations for other diseases [1].

Current routine diagnosis of human echinococcosis is based on imaging procedures (ultrasound, x-ray, computed tomography and magnetic resonance imaging) [1]. Immunodiagnostic techniques such as ELISA and immunoblotting are currently applied to confirm the presence of an *Echinococcus* cyst.

The antigens used in current immunodiagnostic screening tests mainly originate from native *E. granulosus* cyst fluid collected from naturally infected intermediate hosts, e.g. cattle or sheep. Native antigens are limited in availability and their standardization is difficult to achieve. To overcome the limitations of the current diagnostic tests the present study aimed at identifying proteins that are generally well recognized by sera of human echinococcosis patients and to select from these immunodominant regions. We focused on the diagnostically sensitive and specific 20-22 kDa banding pattern seen on immunoblots of *in vitro* cultivated *E. multilocularis* vesicle fluid described by

Müller and co-workers [3] and aimed to identify its protein content by tandem mass spectrometry (MS).

Our objective was to select protein domains for peptide synthesis that have the potential to structurally mimic native epitopes. We chose a bioinformatic approach for the selection of 30mer peptides whereby peptide length was limited by the micro-scale multi-peptide synthesis method (Alta Bioscience, University of Birmingham, UK). Within the proteins identified we searched for predicted alpha-helical coiled-coils (CC) and intrinsically unstructured regions (IUR). Synthetic peptides representing these two folding motifs are capable of adopting their native conformation *in vitro* [4]. Used as antigens in serological tests, peptides deriving from CC and/or IUR domains should therefore be able to present potential epitopes in their natural conformation.

We chose to synthesize peptides instead of recombinant expression of the proteins identified by MS, because this approach could potentially lead to a diagnostic product in a very fast and efficient way. Most importantly, with this approach purification problems in the production of diagnostics grade recombinant antigen could be avoided. Synthetic antigens are advantageous because they are highly pure, well-defined and can be produced in large amounts. Moreover, synthetic antigens can easily be standardized and consequently, immunodiagnostic tests using synthetic peptides may provide a basis for consistent diagnosis comparable among different laboratories. Synthetic peptides have been tested already in a wide range of diagnostic applications and proved valuable for diagnosis of viral diseases (reviewed e.g. in [5-7]).

New systems for parallel detection of individually labeled and multiplexed analytes are well suited for synthetic peptides, which by chemical modification can easily be coupled to microbeads or glass slides. The expected loss of sensitivity in peptide-based immunoassays might be compensated by multiplexing of analytes. Here we investigate first steps towards peptide-based serological assays by applying a proteomic strategy to identify immunologically relevant antigens. For *Echinococcus* as well as for many

other helminthic parasites very little genomic data is available. The presented strategy could pave the road for similar diagnostic tasks.

Materials and methods

Human sera

The echinococcosis sera derived from 33 CE and 43 AE adult patients living in Central Europe. All patients had active hepatic lesions of either CE1 or CE2 type (WHO-IWGE standardized classification). Echinococcosis infections were diagnosed as described by Müller et al. [3]. All samples were collected prior to therapeutic intervention, i.e. before surgery and/or chemotherapy.

To test whether the peptides selected were cross-reactive with sera from other parasitic infections, 74 sera were used. 2 patients were infected with *Trichinella spiralis*, 10 with *Toxocara canis*, 6 with *Ascaris* spec., 1 with *Anisakis* spec., 1 with hookworms, 7 with *Strongyloides stercoralis*, 4 with *Loa loa,* 1 with *Mansonella perstans,* 2 with *Onchocerca volvulus*, 9 with *Fasciola hepatica*, 1 with *Paragonismus westermani*, 7 with *Schistosoma mansoni*, 2 with S. *haematobium*, 1 with *S. mekongi*, 10 with *Taenia solium* and 10 with *Entamoeba histolytica.* Concomitant echinococcosis was ruled out by clinical and serological criteria.

A total of 89 serum samples from healthy blood donors living in Switzerland were used for cut-off calculation to discriminate between positive and negative test results (n=52), specificity testing (n=35) and as negative controls (n=2).

Ethical clearance for retrospective use of anonymized patient sera for test development and quality control was obtained from the ethical committee (Ethikkommission beider Basel).

Antigen, immunoblotting and NanoLC ESI-MS/MS

Vesicle fluid was obtained by aseptic aspiration from *E. multilocularis* metacestodes axenically cultivated as described by Hemphill et al. [8]. The

vesicle fluid (EmVF) was processed as described by Müller et al. [3] and stored at -80°C until use. EmVF was mixed with SDS sample buffer (10mM Tris/HCl pH 8.0, 9M urea, 10% SDS), incubated for 20 minutes at 65°C and electrophoretically separated in either a 5-20% linear gradient polyacrylamide gel or in a 15% polyacrylamide gel. One half of each gel was silver stained (Sigma SilverQuest LC6070). The separated proteins in the other half were electrotransferred onto nitrocellulose membrane as described by Poretti et al. [9]. Nitrocellulose membranes were incubated with serum from a patient infected with *E. multilocularis* to visualize the diagnostically sensitive and infection-specific 20-22 kDa banding pattern described by Müller et al. [3]. Using the immunoblot for orientation, three silver-stained bands corresponding to the region of the diagnostic banding pattern were excised from SDS-gels. One band (band A) derived from a 15% gel, two bands from a 5-20% gradient gel (band B and C). Gel slices were sent for NanoLC ESI-MS/MS analysis (Proteomics Core Facility, University of Geneva, Switzerland).

Immunoblots produced from EmVF separated in a 5-20% gradient gel as described above were cut into strips of 4mm width. 4 of these strips were incubated with rat anti-P29, negative rat control serum, mouse anti-14-3-3 and mouse negative control serum. All primary antibodies were diluted 1:400 in blocking solution (PBS, 2% skim milk, 0.1% Tween20 and NaN_3) and incubated over night at 4°C. Blots were washed 4 times for 5 minutes in PBS 0.1% Tween20. The secondary antibodies, IRDye800-conjugated goat anti-mouse IgG and IRDye800-conjugated goat anti-rat IgG (Li-Cor Biosciences, Lincoln, USA) were diluted 1:5’000 in Odyssey blocking buffer (Li-Cor Biosciences) for 1 hour at room temperature in the dark, followed by 2 washes with PBS-Tween and 2 washes with PBS prior to scanning in an Odyssey Infrared Imaging System (Li-Cor Biosciences).

Two more strips were incubated with human *E. multilocularis* positive and negative control serum and conventionally developed with anti-human IgG-horseradish peroxidase as described previously [9].

Database search

The spectra resulting from MS/MS analysis were searched against the Uniprot Protein Knowledgebase release 15.10 of November 3rd 2009 as well as against translations of the *E. multilocularis* and *E. granulosus* EST libraries available at the Wellcome Trust Sanger Institute (ftp://ftp.sanger.ac.uk/pub/pathogens/Echinococcus/) by using the MASCOT search engine (MATRIX Science, London, UK). Scaffold (version Scaffold_2_05_01, Proteome Software Inc., Portland, USA) was used to validate MS/MS based peptide and protein identifications. Peptide and protein identification probabilities were assigned by the PeptideProphet and the ProteinProphet algorithm [10,11]. Peptide and protein identifications were accepted if the probability of correct identification was >95% and at least two peptides per protein were detected.

Peptides

Protein and translated EST sequences identified by MS/MS analysis served as parent proteins for peptide design. 41 peptides between 24 and 30 amino acids in length were selected according to rules published previously [12]. Parent protein sequences were analyzed by different algorithms to identify regions of stable alpha-helical coiled-coils (Paircoil2 [13], StableCoil [14] and intrinsically unstructured regions (IUPred [15], RONN [16]). Protein sequences without CC predictions were further analyzed for potential B-cell epitopes (BepiPred [17]).

Peptides were produced by Fmoc solid phase synthesis (Alta Bioscience, University of Birmingham, UK). They contained an additional biotin used for immobilization onto streptavidin coated ELISA plates (NUNC immobilizer streptavidin) attached via a PEG spacer at the N-terminus. The C-terminus was amidated. The peptides were supplied as crude synthesis product. Peptide concentrations given in this work assume a consistent yield of 1mg of full-length peptide for each synthesis reaction.

ELISA

ELISAs were performed in 96-well plates (NUNC Immobilizer Streptavidin) as described elsewhere [12] with the following modifications: in initial screening ELISAs all 41 peptide candidates were coated at a concentration of 2ug/ml. Optimal coating concentration was determined for the most promising candidates C2 and B6 by peptide titration. For both peptides, the optimal coating concentration was 0.5ug/ml.

The absorbance was measured at 405nm (A_{405}) in a VERSAmax microplate reader. The raw data was transferred into Microsoft Excel and blank corrected. A cut-off value was calculated from results obtained from 52 healthy blood donors as average blood donor A_{405} value plus 2 standard deviations

The sensitivity (true positives / true positives + false negatives) was calculated for AE and CE sera together. All test blood donor samples (n=35) together with the samples from helminthic infections other than echinococcosis (n=64) served for calculation of specificity (true negatives / true negatives + false positives).

Results

NanoLC ESI-MS/MS

The proteins identified with a corresponding Uniprot entry varied between 24 and 49 kDa. All translated ESTs identified by MS/MS were subjected to a protein BLAST search (http://blast.ncbi.nlm.nih.gov/Blast.cgi). If the blastp search returned protein hits with more than 40% similarity to the EST query, the blastp hits were used to characterize the corresponding EST sequence, otherwise the translated EST was termed “unknown protein product”. Protein sequences identified and selected for peptide design are listed in table 1, together with the blastp hits of the corresponding EST queries.

Regarding the number of unique peptide spectra found by MS/MS analysis, the most abundant proteins in bands A, B and C were members of the 14-3-3 protein family. We designed therefore 15 peptides from 4 different conceptually translated 14-3-3-like EST sequences (table 2). The second most abundant protein was phosphoglycerate mutase, which was found in band A and B. It served as parent protein for 3 peptides (table 2). Antigen 6 (accession no. AAB61984) was the third most abundant protein, but was excluded from peptide design, because we had previously investigated Antigen 6 derived peptides [12]. In these analyses, antigen 6 peptides had shown only little reactivity with sera from echinococcosis patients. Following proteins served for the design of ≤ 3 peptides: thioredoxin peroxidase, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), TSP5, cytosolic malate dehydrogenase, lactate dehydrogenase, elongation factor 1 alpha, proliferating cell nuclear antigen (PCNA), purine nucleoside phosphorylase, uridine cytidine kinase I, proteasome subunit alpha 7 and three different unknown protein products (table 2).

For the selection of parent proteins, priority was given to *E. multilocularis* sequences, but the majority of the peptides designed could also be detected in *E. granulosus* EST sequences. Thus 17 peptides were 100% identical between *E. multilocularis* and *E. granulosus*. 9 differed in one, two or three amino acids. 10 peptides were *E. multilocularis*-specific, 2 peptides were *E. granulosus*-specific and 3 peptides could only partially be identified in *E. granulosus* libraries because of incomplete EST reads (table 2).

Immunoblot

Strips of EmVF immunoblots incubated with human echinococcosis positive and negative control sera are shown in figure 1. The arrow marks the main band visible in the 20-22 kDa diagnostic relevant region.

To confirm the presence of 14-3-3 and P29 identified by tandem mass spectrometry, we used hyperimmune sera from mouse and rabbit immunized with recombinant 14-3-3 and P29 respectively. Strips from the same EmVF

immunoblot as used for incubation with human control sera (see above) were incubated with the respective mouse and rabbit sera. Both, anti-recP29 and anti-rec14-3-3 hyperimmune sera showed a distinct banding pattern in the diagnostically relevant region (figure 2). The pattern of rat anti-recP29 was more alike the pattern produced by sera of human echinococcosis patients than the pattern of mouse anti-14-3-3. Depending on the batch of EmVF used for immunoblotting, human echinococcosis sera produce a more or less pronounced banding pattern of up to three bands with a rather fuzzy appearance. This pattern could be caused by the overlay of several different proteins reactive with echinococcosis sera, for example by the upper band of P29 and the lower band of 14-3-3. Since the diagnostic value of P29 has already been investigated [18,19] and because the performance of our P29 peptides investigated previously [12] was not satisfactory, we decided to leave P29 and to focus instead on 14-3-3 and the other proteins identified by MS/MS.

ELISA

The 41 peptide candidates were first assessed with 3 pools of AE and 3 pools of CE sera (5 sera in each pool). 14 out of 41 peptides yielded an A_{405}-value ≥ 0.150 after 20 minutes of incubation with substrate were selected for testing with individual sera from AE and CE infected patients (3 of each) and 1 blood donor serum. 8 of these 14 peptides were also reactive with single AE and CE sera and were further tested with an extended serum panel including additional echinococcosis sera, additional blood donor sera and additional sera from other helminthic and *Entamoeba histolytica* infections.

The most promising peptides, B6 deriving from phosphoglycerate mutase and C2 deriving from cytosolic malate dehydrogenase (cMDH), were tested with the entire serum panel. The results are listed in table 3. The cut-off A_{405} values were 0.196 for B6 and 0.114 for C2. Both peptides corresponded to an *E. multilocularis* sequence, but were equally reactive with sera from AE and CE patients and therefore were not able to discriminate between AE and CE infection. A blastp search in the collection of translated ESTs confirmed that

peptides C2 and B6 were completely identical in both *Echinococcus* species. Peptide C2 was detected in following EST libraries: Emme (*E. multilocularis* metacestode; ligation numbers 90724, 90725, 90726, 90727), EMmg (oligo-capped cDNA library from the metacestode (germinal layer) of *E. multilocularis*) and EGCWgr (oligo-capped cDNA library from the cyst wall of *E. granulosus*). Peptide B6 was found in Emme, EGPSPsl (spliced-leader cDNAs library from pepsin-treated *E. granulosus* protoscoleces) and EGPSPgr (oligo-capped cDNA library from the protoscolex stage of *E. granulosus*).

The sensitivities of both peptides were low, 15% for C2 and 18% for B6. However, their reactivity was complementary and thus the sensitivity increased to 27% if the results of the two peptide ELISAs were cumulated.

The specificities of the peptides were high, 90% for C2 and 91% for B6. Crossreactivity with neurocysticercosis was observed in both peptides: 2/10 samples were positive with C2 and 1/10 with B6. 2/34 sera from nematode infections, 2/20 sera from trematode infections and 2/10 sera from *Entamoeba histolytica* infections were tested positive in C2 ELISA. In B6, 4/34 nematode and 2/20 trematode samples were positive (table 3). Regarding the cumulated specificity of the two peptides, the value decreased to 82%.

Discussion

Considerable efforts have been made in order to improve the performance of antigens for the diagnosis of human echinococcosis by purifying fractions from crude extracts, e.g. the Em2 fraction from metacestode tissue or antigen 5 enriched cyst fluid. Recombinant antigens, such as EmII/3-10 and antigen B, were introduced into serological diagnosis. The application of recombinant antigens is advantageous because they can readily be produced in large amounts, while affinity purification of native antigens is time-consuming and often producing low yields [20]. A consequent further development would lead to replacement of both, native and recombinant antigens by chemically synthesized antigens, such as the 30 amino acids long peptides investigated

in this study. The added value of developing synthetic antigen consists in high reproducibility and purity, and also in the versatility of modified or labeled synthetic peptides, thus permitting the choice of alternative diagnostic detection systems, such as multiplex bead-based suspension assays. The recent developments of new platform technologies calls for defined analytes suitable for multiplexing.

Aiming at the discovery of peptides reactive with human echinococcosis sera, we followed a proteomic approach. We targeted the antigens seen at the 20-22 kDa region on immunoblots of *in vitro* cultivated *E. multilocularis* vesicle fluid. Reactivity of patients' sera with the 20-22 kDa banding pattern has been described as a marker for *Echinococcus* infection [3]. The protein content of this region cut out in 3 fragments was analyzed by tandem mass spectrometry and database mining.

One major challenge was the complexity of the protein mixture, 34 proteins were identified in the three bands subjected to mass spectrometry, 23 of which were found only by searching peptide spectra against our database of theoretical proteins conceptually translated from EST libraries, which are available at the Wellcome Trust Sanger Institute (ftp://ftp.sanger.ac.uk/pub/pathogens/Echinococcus/). Thus, in the case of an organism without (yet) sequenced genome, protein identification can be considerably improved by the use of EST libraries.

The number of trypsin fragments detected from a protein identified hints at its relative abundance in the excised gel fragment. But abundance does not correspond to antigenicity. Therefore a second round of immunological screening was required. The 41 peptide candidates were screened for reactivity with patients' sera in the well-proven ELISA system.

Among the proteins identified were the functionally diverse 14-3-3 family members. *Echinococcus* 14-3-3 homologue 1 and 2 were represented by the highest number of unique peptide spectra. Members of the 14-3-3 protein family were detected in all eukaryotes examined so far. The 14-3-3 family was

found to be involved in a wide range of functions including cell signaling, cell cycle regulation, metabolism control, apoptosis, gene transcription, intracellular trafficking and cytoskeletal structure [21]. In *E. multilocularis* the 14-3-3 protein was first detected by immunoscreening of an expression library of metacestode cDNA [22]. The protein was found to be 10-fold overexpressed in the metacestode stage when compared to expression levels in the adult tapeworm. In multiple sequence alignment the *Echinococcus* sequence clustered with the zeta isoforms, which are known to be related to tumor growth [23]. It has been suggested that the 14-3-3 protein might be involved in the infiltrative and potentially unlimited, tumor-like growth behavior of the *E. multilocularis* metacestode larva within the host tissue [22,23].

We also detected in the 20-22 kDa banding pattern proteins that were previously investigated for their antigenic properties in echinococcosis, such as antigen 6 (which shows 97% identity to *E. granulosus* antigen P29 [18]), cytosolic malate dehydrogenase [24,25], TSP5 [26], and thioredoxin peroxidase [27].

The detection of antigen 6 in *E. multilocularis* vesicle fluid was unexpected. Antigen P29 of *E. granulosus*, the putative antigen 6 orthologue, was identified in a study aimed deliberately at the discovery of antigens distinct from those contained in hydatid cyst fluid [18]. Immunoblotting with parasite extracts and immunofluorescence on parasite sections revealed the localization of P29 to protoscoleces and the germinal layer of the metacestode. It was absent from adult worms and from hydatid cyst fluid [18]. Thus, assuming *E. multilocularis* antigen 6 to be the orthologue of *E. granulosus* P29, our MS/MS-based identification approach gives evidence for the differential localization of a homologous protein within the two closely related tapeworms.

The cytosolic malate dehydrogenase (cMDH) of *Echinococcus granulosus* was as well found by immunoscreening of a protoscolex cDNA expression library [24]. All eukaryotes posses two forms of malate dehydrogenase, a cytosolic (cMDH) and a mitochondrial (mMDH) one. Malate dehydrogenases

are important enzymes of the citric acid cycle and gluconeogenesis. The *Echinococcus* cMDH was recombinantly expressed and assessed as diagnostic antigen in ELISA and immunoblot [25]. Due to its rather low sensitivity of 54%, the authors suggested to use the recombinant cMDH in combination with a panel of several defined antigens for the improved diagnosis of human echinococcosis [25]. Later on, recombinant cMDH was part of such an antigen panel tested in a multicenter study in South America. In this study, recombinant cMDH performed worse than expected and produced inconsistent results between the participating centers [28]. The authors therefore agreed the now commonly accepted fact that comparability of antigen performance needs standardization in respect to sera used and antigen preparation [28].

Three more proteins identified in the diagnostically relevant region were known from the glycolytic pathway and were reported antigens from other parasitic organisms: GAPDH in *Schistosoma mansoni* [29], lactate dehydrogenase in *Plasmodium* spec. [30] and phosphoglycerate mutase in *Candida albicans* [31]. Phosphoglycerate mutases are essential enzymes in glycolysis and are found in bacteria, fungi, plants and animals. Their benefit as diagnostic markers has not yet been investigated in detail. In contrast to few records on antigenicity, phosphoglycerate mutases and other enzymes of the glycolytic pathway appear as useful drug target in different parasites, such as flukes [32] and nematodes [33] but also *Trypanosoma* and *Leishmania* [34,35].

Further proteins identified were purine nucleoside phosphorylase which had been discovered as potential biomarker for trichloroethylene-induced autoimmune disease in humans [36] and proliferative nuclear cell antigen (PCNA) a protein considered as biomarker for human systemic lupus erythematosus [37]. The diagnostic banding pattern contained also few proteins that had not been investigated as potential antigens, namely elongation factor 1 alpha (BAF63674), uridine cytidine kinase I and proteasome subunit alpha 7. Some peptide spectra matched no Uniprot entry, but EST-sequences only. These sequences were of particular interest for

peptide design, because they represented proteins possibly unique to *Echinococcus* and thus were of great interest for the application in diagnostics.

The previous characterization of the 20-22 kDa diagnostic banding pattern with specific antibodies had shown that this region of the immunoblot did not contain any degradation forms of EmII/3, whereas the full-length EmII/3 appeared at the expected size of 65 kDa [3]. Our data supports this finding since EmII/3 was not detected in any of the bands analyzed.

In conclusion, it was not possible to tell from the list of proteins identified which was (were) the one(s) responsible for the specific reaction seen on immunoblot. As mentioned before, the estimated abundance of protein contained in the gel slices did not correspond with antigenicity. The 14-3-3-proteins identified by the highest number of unique peptides failed in ELISA. Only one 14-3-3-candidate, peptide A2, showed weak reactivity in first assessments, but had to be dismissed due to low sensitivity.

All other known *Echinococcus* proteins on the list which had previously been tested as antigens in immunoblot or ELISA showed inferior operating characteristics than the still unknown EmVF protein(s) in the 20-22 diagnostic banding pattern [18,24,27]. In addition, our screening ELISA was not able to identify peptides with comparable high sensitivity and specificity. A reason for this might be the nature of the epitopes sought after. In case the diagnostically relevant targets are mainly composed of carbohydrate epitopes, it will not be possible to reproduce these epitopes with simple synthetic peptides. On the other hand, the relevant epitopes might be of proteinaceous nature, but be part of a secondary structure that cannot be adopted by synthetic 30mer peptides. Another shortcoming of synthetic peptides might be the destruction of epitopes caused by the selection of peptide location. A peptide might start or end amidst the amino acids constituting an epitope and thereby destroy the site recognized by antibodies. The testing of several highly overlapping peptides might be chosen to counter this issue. Our results, in particular the

immunoblot with rat anti-recP29, strongly suggest considering overlapping peptides for a more detailed analysis of antigen 6 or P29, respectively.

Taken together, inference from the peptide antigens tested in the present study on the identity of the 20-22 kDa protein(s) cannot be made. Nonetheless two new peptides with potential for application in serodiagnosis of human echinococcosis have been discovered.

Besides the complexity of the protein content of the three bands tested by mass spectrometry, the second challenge of our approach consisted in the selection of antigenic domains within the identified proteins and translated ESTs. Our peptide design was guided by the predictions of alpha-helical coiled-coils (CC) and intrinsically unstructured regions (IUR). Proteins without predicted CC were further analyzed with BepiPred, an algorithm for the prediction of linear B-cell epitopes. Alpha-helical coiled-coils and intrinsically unstructured regions are believed to adopt native conformation if chemically synthesized and therefore synthetic peptides corresponding to these two motifs potentially could represent native epitopes.

The most reactive peptides discovered in our screening ELISA were peptide C2 from *Echinococcus* cMDH and peptide B6 originating from a member of the phosphoglycerate mutase family.

Regarding cMDH, we designed and tested two peptides: peptide C2 (27 amino acids in length) and C3 (28 amino acids). A blastp search within the translated EST libraries revealed that C2 was 100% identical in *E. multilocularis* and *E. granulosus* while peptide C3 differed in one amino acid. In ELISA, the sensitivity of peptide C2 was as low as 15% in, but specificity was high with 90%. C3 was discarded early in the screening procedure because it showed no reactivity with patients' sera. From the phosphoglycerate mutase sequence, we designed and tested three different 30mer peptides: B5, B6 and B7. Peptide B5 was non-reactive with patients' sera. B6 achieved a low sensitivity of 18%, but a high specificity of 91% and

B7 showed reactivity only with a pool made of CE sera, but not with single patients' sera.

Low sensitivity was the major shortcoming of the synthetic peptides identified and tested. This clearly precludes the use of a single peptide for diagnostic purposes, and combination of several peptides to create a complex antigen mixture is required. The limited length of the peptides could explain the observed low sensitivity because 30 amino acids potentially harbor less epitopes than 300 amino acids of an average full-length protein. This is supported by our previous characterization of another *E. multilocularis* peptide [12], a 42mer peptide originating from antigen B8/1. This peptide covers almost two thirds of the length of its parent protein (65 amino acids) and together with a closely related peptide p176 [38] represents to our knowledge the most sensitive synthetic peptides tested so far in the diagnosis of echinococcosis.

Our results demonstrate a new approach in the search for diagnostic peptides by MS combined with bioinformatics. While individual peptides do not provide sufficient sensitivity suitable for a diagnostic routine test, peptide-pools could have the potential to compensate this limitation.

Acknowledgment

We thank Prof. Bruno Gottstein from the Institute of Parasitology, University of Bern for kindly providing rat anti-P29 and mouse anti-14-3-3 pre- and hyperimmune sera. We are most grateful to Dr. Joachim Müller from the Institute of Parasitology, University of Bern, for technical guidance and Dr. Markus Spiliotis, Institute of Parasitology, University of Bern, for fruitful discussions on the topic. We are much obliged to Prof. Christoph Hatz and Dr. Stefanie Kramme from the medical and diagnostic services unit of the Swiss TPH in Basel for the provision of human sera from infections other than echinococcosis.

This project was funded by the Velux Foundation (http://www.veluxstiftung.ch). CL was supported by Rudolf Geigy Foundation (http://www.sti.ch/en/about-us/foundations.html), Fonds zur Förderung von Lehre und Forschung (http://www.fag-basel.ch/frameset.html) and Roche Research Foundation (http://www.research-foundation.org/rrf/).

References

1. Siles-Lucas M, Gottstein B: **Review: Molecular tools for the diagnosis of cystic and alveolar echinococcosis.** *Tropical Medicine & International Health* 2001, **6:**463-475.

2. Gottstein B: **Molecular and immunological diagnosis of echinococcosis.** *Clin Microbiol Rev* 1992, **5:**248-261.

3. Müller N, Frei E, Nuñez S, Gottstein B: **Improved serodiagnosis of alveolar echinococcosis of humans using an *in vitro* produced *Echinococcus multilocularis* antigen.** *Parasitology* 2007, **134:**879-888.

4. Corradin G, Villard V, Kajava AV: **Protein structure based strategies for antigen discovery and vaccine development against malaria and other pathogens.** *Endocrine Metabolic & Immune Disorders-Drug Targets* 2007, **7:**259-265.

5. Gómara MJ, Haro I: **Synthetic peptides for the immunodiagnosis of human diseases.** *Current Medicinal Chemistry* 2007, **14:**531-546.

6. Noya O, Patarroyo ME, Guzman F, de Noya BA: **Immunodiagnosis of parasitic diseases with synthetic peptides.** *Current Protein and Peptide Science* 2003, **4:**299-308.

7. Meloen RH, Langedijk JPM, Langeveld JPM: **Synthetic peptides for diagnostic use.** *Veterinary Quarterly* 1997, **19:**122-126.

8. Hemphill A, Gottstein B: **Immunology and Morphology Studies on the Proliferation of *in vitro* Cultivated *Echinococcus multilocularis* Metacestodes.** *Parasitology Research* 1995, **81:**605-614.

9. Poretti D, Felleisen E, Grimm F, Pfister M, Teuscher F, Zuercher C, Reichen J, Gottstein B: **Differential immunodiagnosis between cystic hydatid disease and other cross-reactive pathologies.** *Am J Trop Med Hyg* 1999, **60:**193-198.

10. Keller A, Nesvizhskii AI, Kolker E, Aebersold R: **Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search.** *Analytical Chemistry* 2002, **74:**5383-5392.

11. Nesvizhskii AI, Keller A, Kolker E, Aebersold R: **A statistical model for identifying proteins by tandem mass spectrometry.** *Analytical Chemistry* 2003, **75:**4646-4658.

12. List C, Qi W, Maag E, Gottstein B, Müller N, Felger I: **Serodiagnosis of *Echinococcus* spp. Infection: Explorative Selection of Diagnostic Antigens by Peptide Microarray.** *PLoS Negl Trop Dis* 2010, **4:**e771.

13. McDonnell AV, Jiang T, Keating AE, Berger B: **Paircoil2: improved prediction of coiled coils from sequence.** *Bioinformatics* 2006, **22:**356-358.

14. Tripet B, Hodges RS: **StableCoil: an algorithm designed to predict the location and relative stability of coiled-coils in native protein sequences.** In *Peptides: The Wave of the Future*; Edited by Lebl M, Houghten RA 365-366.

15. Dosztanyi Z, Csizmok V, Tompa P, Simon I: **IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content.** *Bioinformatics* 2005, **21:**3433-3434.

16. Yang ZR, Thomson R, McNeil P, Esnouf RM: **RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins.** *Bioinformatics* 2005, **21:**3369-3376.

17. Larsen J, Lund O, Nielsen M: **Improved method for predicting linear B-cell epitopes.** *Immunome Research* 2006, **2:**2.

18. González G, Spinelli P, Lorenzo C, Hellman U, Nieto A, Willis A, Salinas G: **Molecular characterization of P-29, a metacestode-specific component of *Echinococcus granulosus* which is immunologically related to, but distinct from, antigen 5.** *Molecular and Biochemical Parasitology* 2000, **105:**177-185.

19. Ben Nouir N, Gianinazzi C, Gorcii M, Müller N, Nouri A, Babba H, Gottstein B: **Isolation and molecular characterization of recombinant *Echinococcus granulosus* P29 protein (recP29) and its assessment for the post-surgical serological follow-up of human cystic echinococcosis in young patients.** *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2009, **103:**355-364.

20. Vogel M, Gottstein B, Müller N, Seebeck T: **Production of a recombinant antigen of *Echinococcus multilocularis* with high immunodiagnostic sensitivity and specificity.** *Molecular and Biochemical Parasitology* 1988, **31:**117-125.

21. Aitken A: **14-3-3 proteins: A historic overview.** *Seminars in Cancer Biology* 2006, **16:**162-172.

22. Siles-Lucas M, Felleisen RSJ, Hemphill A, Wilson W, Gottstein B: **Stage-specific expression of the 14-3-3 gene in *Echinococcus multilocularis*.** *Molecular and Biochemical Parasitology* 1998, **91:**281-293.

23. Siles-Lucas M, Nunes CP, Zaha A: **Comparative analysis of the 14-3-3 gene and its expression in *Echinococcus granulosus* and *Echinococcus multilocularis* metacestodes.** *Parasitology* 2001, **122:**281-287.

24. Rodrigues JJS, Ferreira HB, Zaha A: **Molecular cloning and characterization of an *Echinococcus granulosus* cDNA encoding malate dehydrogenase.** *Molecular and Biochemical Parasitology* 1993, **60:**157-159.

25. Ferreira HB, Zaha A: **Expression and analysis of the diagnostic value of an *Echinococcus granulosus* antigen gene clone.** *International Journal for Parasitology* 1994, **24:**863-870.

26. Dang Z, Watanabe J, Kajino K, Oku Y, Matsumoto J, Yagi K, Kouguchi H, Sugimoto C: **Molecular cloning and characterization of a T24-like protein in *Echinococcus multilocularis*.** *Molecular and Biochemical Parasitology* 2009, **168:**117-119.

27. Margutti P, Ortona E, Delunardo F, Tagliani A, Profumo E, Rigano R, Buttari B, Teggi A, Siracusano A: **Thioredoxin peroxidase from *Echinococcus granulosus*: a candidate to extend the antigenic panel for the immunodiagnosis of human cystic echinococcosis.** *Diagnostic Microbiology and Infectious Disease* 2008, **60:**279-285.

28. Lorenzo C, Ferreira HB, Monteiro KM, Rosenzvit M, Kamenetzky L, Garcia HH, Vasquez Y, Naquira C, Sanchez E, Lorca M et al.: **Comparative Analysis of the Diagnostic Performance of Six Major *Echinococcus granulosus* Antigens Assessed in a Double-Blind, Randomized Multicenter Study.** *J Clin Microbiol* 2005, **43:**2764-2770.

29. Charrier-Ferrara S, Caillol D, Goudot-Crozel V: **Complete Sequence of the *Schistosoma mansoni* Glyceraldehyde-3-Phosphate Dehydrogenase Gene Encoding A Major Surface Antigen.** *Molecular and Biochemical Parasitology* 1992, **56:**339-343.

30. Maltha J, Gillet P, Bottieau E, Cnops L, van Esbroeck M, Jacobs J: **Evaluation of a rapid diagnostic test (CareStart (TM) Malaria HRP-2/pLDH (Pf/pan) Combo Test) for the diagnosis of malaria in a reference setting.** *Malaria Journal* 2010, **9**.

31. Pardo M, Ward M, Pitarch A, Sanchez M, Nombela C, Blackstock W, Gil C: **Cross-species identification of novel *Candida albicans* immunogenic proteins by combination of two-dimensional polyacrylamide gel electrophoresis and mass spectrometry.** *Electrophoresis* 2000, **21:**2651-2659.

32. Martin RJ: **Modes of action of anthelmintic drugs.** *Veterinary Journal* 1997, **154:**11-34.

33. Raverdy S, Zhang Y, Foster J, Carlow CKS: **Molecular and biochemical characterization of nematode cofactor independent phosphoglycerate mutases.** *Molecular and Biochemical Parasitology* 2007, **156:**210-216.

34. Verlinde CLM, Hannaert V, Blonski C, Willson M, Perie JJ, Fothergill-Gilmore LA, Opperdoes FR, Gelb MH, Hol WGJ, Michels PAM: **Glycolysis as a target for the design of new anti-trypanosome drugs.** *Drug Resistance Updates* 2001, **4:**50-65.

35. Nowicki MW, Tulloch LB, Worralll L, Mcnae IW, Hannaert V, Michels PAM, Fothergill-Gilmore LA, Walkinshaw MD, Turner NJ: **Design, synthesis and trypanocidal activity of lead compounds based on inhibitors of parasite glycolysis.** *Bioorganic & Medicinal Chemistry* 2008, **16:**5050-5061.

36. Liu J, Xing X, Huang H, Jiang Y, He H, Xu X, Yuan J, Zhou L, Yang L, Zhuang Z: **Identification of antigenic proteins associated with trichloroethylene-induced autoimmune disease by serological proteome analysis.** *Toxicology and Applied Pharmacology* 2009, **240:**393-400.

37. Beyne-Rauzy O, Thébault S, Adoue D, Fortenfant F: **Anti-PCNA antibodies: prevalence and predictive value.** *Joint Bone Spine* 2005, **72:**432-435.

38. González-Sapienza G, Lorenzo C, Nieto A: **Improved immunodiagnosis of cystic hydatid disease by using a synthetic peptide with higher diagnostic value than that of its parent protein, *Echinococcus granulosus* antigen B.** *J Clin Microbiol* 2000, **38:**3979-3983.

Figures

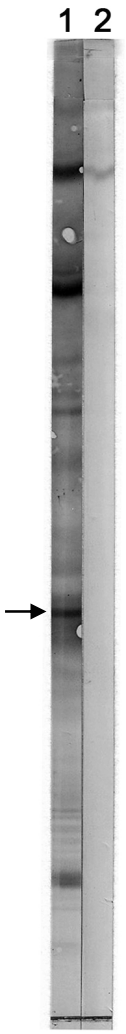

Figure 1:

Immunoblot of human echinococcosis positive (strip1) and negative (strip 2) control sera; the arrow indicates the main band visible in the diagnostically relevant region.

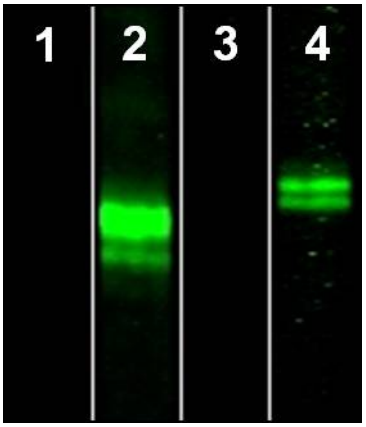

Figure 2:

Immunoblot incubated with 1) pre-immune rat serum, 2) serum from rat immunized with recombinant *E. granulosus* P29, 3) pre-immune mouse serum and 4) serum from mouse immunized with recombinant *E. granulosus* 14-3-3.

Tables

Table 1:
Selected sequences identified by MS/MS in bands A, B and C and corresponding blastp hits of EST queries.

Parent sequence for peptide design	Species[1)]	MS/MS[2)]	Blastp hit of EST query or name of Uniprot entry	Species	Accession no.	Identities
Emme_03b12.q1k	*E. m.*	A	14-3-3 family member	*C. elegans*	NP_509939	151/242 (62%)
EMmg_17b02.q1k	*E. m.*	A, B, C	Putative 14-3-3 protein	*E. granulosus*	AAX73175	245/247 (99%)
Emme_01f09.p1k	*E. m.*	A, B, C	14-3-3.2 zeta	*E. granulosus*	ABG02282	239/239 (100%)
Emme_05b10.p1k	*E. m.*	A, C	14-3-3 protein isoform 2	*A. caninum*	ACO59962	90/162 (55%)
Emme_36e09.q1k	*E. m.*	A, B	Phosphoglycerate mutase	*C. sinensis*	ABZ82035	180/233 (77%)
Emme_08g06.p1k	*E. m.*	B, C	Thioredoxin peroxidase	*E.multilocularis*	BAC11863	193/193 (100%)
Emme_01f11.p1k	*E. m.*	B, C	TSP5	*E. multilocularis*	ACJ02406	225/225 (100%)
EMmg_13f04.p1k	*E. m.*	C	Cytosolic malate dehydrogenase	*E. granulosus*	Q04820	221/228 (96%)
EMmg_10f08.q1k	*E. m.*	B, C	Lactate dehydrogenase	*C.sinensis*	AAV80238	158/266 (59%)
EMmg_17c12.q1k	*E. m.*	A	Putative PCNA	*S. japonicum*	CAX76520	209/258 (81%)
Emme_56h11.p1k	*E. m.*	A, C	Sj-Ts4 protein	*S. mansoni*	XP_002572587	51/137 (37%)
Emme_20a02.p1k	*E. m.*	A	Chloride intracellular channel 4	*S. japonicum*	CAX71022	78/210 (37%)
EMmg_15d06.q1k	*E. m.*	A, B	Expressed hypothetical protein	*T. adhaerens*	XP_002109451	74/242 (30%)
EGPSgr_9h05.q1k	*E. g.*	A	Purine nucleoside phosphorylase 5a	*D. rerio*	NP_998476	123/224 (54%)
EGPSgr_11e06.p1k	*E. g.*	A	Uridine cytidine kinase I	*S. mansoni*	XP_002577501	66/108 (61%)
EMmg_10f12.p1k	*E. m.*	A, B, C	Proteasome subunit alpha 7	*S. mansoni*	XP_002578798	160/208 (76%)
Accession no. A5LIC2	*E. m.*	A	Elongation factor 1 alpha	*E. multilocularis*	A5LIC2	448/448 (100%)
Accession no. Q27652	*E. m.*	A, C	GAPDH	*E. multilocularis*	Q27652	336/336 (100%)

[1)] Species origin of EST library or Uniprot entry.
[2)] Corresponding proteins were identified by MS/MS in one or several of the bands analyzed (bands A, B and C).

Table 2:
Sequences of peptides tested in ELISA.

Parent protein Name	Species	Peptide name	Sequence	Length	Identity to E.g. / E.m.	Prediction
14-3-3 protein	*E. m.*[1)]	MA2_A2	RRSSWRIISAIESKDEGNERVKNIR	25	100 %	IUR / B-cell epitope
14-3-3 protein	*E. m.*	MA2_A3	NERVKNIRACREQVEKELKQMANEILSLLE	30	100 %	CC
14-3-3 protein	*E. m.*	MA2_A4	NRYLAEIATASERDEVAKRSLDAYKVATD	29	100 %	B-cell epitope
14-3-3 protein	*E. m.*	MA2_A5	KDSTLIMQLLRDNLTLWRSDVDNDEAEKN	29	1 mismatch	IUR / B-cell epitope
14-3-3 protein	*E. m.*	MA2_A6	REENVYMAKLCEQCERYDEMVKAMKDVLES	30	E.m. specific	CC
14-3-3 protein	*E. m.*	MA2_A7	ERYDEMVKAMKDVLESGADLSVEERNLLS	29	E.m. specific	CC/IUR
14-3-3 protein	*E. m.*	MA2_A8	RVISSIEQKHDGDAKMQIAKKVREEIERE	29	E.m. specific	IUR / B-cell epitope
14-3-3 protein	*E. m.*	MA2_A9	TGDERKQASDNSLMAYKSATEVAEGDMQT	29	E.m. specific	IUR / B-cell epitope
14-3-3 protein	*E. m.*	MA2_A10	QLLRDNLTLWNSDAGDTDAAEPPKAD	26	E.m. specific	IUR / B-cell epitope
14-3-3 protein	*E. m.*	MA2_A11	KLAEQAERYEDMAVAMKTIAEMGNELNNEE	30	100 %	CC
14-3-3 protein	*E. m.*	MA2_A12	DIYLKKVEEELTKICNDVLALLSKNLITEK	30	100 %	CC
14-3-3 protein	*E. m.*	MA2_B1	EVQEGEQNDKSTEAAEEAYQKATSLAEAE	29	100 %	IUR / B-cell epitope
14-3-3 protein	*E. m.*	MA2_B2	KEVTRFEKELNNEERNLLSVAFKNVVGSRR	30	1 mismatch	IUR / B-cell epitope
14-3-3 protein	*E. m.*	MA2_B3	SRRNSYRVLSSRLARTQDPEKQALTKEYLD	30	1 mismatch	IUR / B-cell epitope
14-3-3 protein	*E. m.*	MA2_B4	RYKAENAKGEDHKQVVEASLKAYEEATE	28	100 %	B-cell epitope
Phosphoglycerate mutase	*E. m.*	MA2_B5	ENRFCGWHDADLSPKGEQEAQNAGKMIKER	30	100 %	IUR / B-cell epitope
Phosphoglycerate mutase	*E. m.*	MA2_B6	PVYKSWRLNERMYGGLQGLNKSETAAKHGE	30	100 %	IUR / B-cell epitope
Phosphoglycerate mutase	*E. m.*	MA2_B7	RAYDIPPPPLEVSDKRWPGLEEKYKLLDIE	30	1 mismatch	B-cell epitope
Thioredoxin peroxidase	*E. m.*	MA2_B8	NNVSRKEGGVQGMRIPMLADTNHKISRD	28	100 %	B-cell epitope
Thioredoxin peroxidase	*E. m.*	MA2_B9	DKHGEVCPANWHPGSKTFKPSAGDLKS	27	1 mismatch	IUR / B-cell epitope
GAPDH	*E. m.*	MA2_B10	KKVIISAPSADSSHVRVGVNHEKYDPSMK	29	n.a.***	IUR / B-cell epitope
GAPDH	*E. m.*	MA2_B11	TATQKVVDGPSKKAWRDGRTAAQN	24	100%	IUR / B-cell epitope
GAPDH	*E. m.*	MA2_B12	KLSKPATYDQIKAAVKRASESSALKRILE	29	3 mismatches	IUR / B-cell epitope
TSP5	*E. m.*	MA2_C1	DFVRLVGENLKEAIKGLESGSLSGSDPTLK	30	2 mismatches	CC
Cytosolic malate dehydrogenase	*E. m.*	MA2_C2	NVKIFKEQGEALDKYAKKTVKVLVVGN	27	100 %	B-cell epitope

Table 2 continued:
Sequences of peptides tested in ELISA.

Parent protein Name	Species	Peptide name	Sequence	Length	Identity to E.g. / E.m.	Prediction
Cytosolic malate dehydrogenase	*E. m.*	MA2_C3	DLSHAVVTKDGKQHPARELINDEKWVKE	28	1 mismatch	IUR / B-cell epitope
Lactate dehydrogenase	*E. m.*	MA2_C4	KIGQAGDPDDFASIHKAVVDSAYEIIRMK	29	100 %	CC
Elongation factor 1 alpha	*E. m.*	MA2_C5	DAVDYSEKRFQEISSEMKAYIKKVGYNPDT	30	100 %	B-cell epitope
Elongation factor 1 alpha	*E. m.*	MA2_C6	KNISVKDVRRGNVAGDSKNHPPREAAEFT	29	100 %	IUR / B-cell epitope
Elongation factor 1 alpha	*E. m.*	MA2_C7	KFAELKEKIDRRTGQVKETNPAKIKSGD	28	100 %	CC
PCNA	*E. m.*	MA2_C8	DSITLKAGDKADTITFLFESKNQEKVSEFE	30	E.m. specific	B-cell epitope
PCNA	*E. m.*	MA2_C9	DLGSGKIKLSPGGNADKPEESISIEMSES	29	E.m. specific	IUR / B-cell epitope
Unknown protein product 1	*E. m.*	MA2_C10	SESAASAIRDQAKALASTDAVTKLDDIACR	30	1 mismatch	IUR / B-cell epitope
Unknown protein product 2	*E. m.*	MA2_C11	RYMVEMDLIRLEVVPISYDNEPEDYK	26	100 %	CC
Unknown protein product 3	*E. m.*	MA2_C12	FTNEIRGTASRSNHPDDIVADFDTADRLKK	30	E.m. specific	IUR / B-cell epitope
Unknown protein product 3	*E. m.*	MA2_D1	DRMQDMVLEKFDAKKQPKLTSAELLGKLEQ	30	E.m. specific	IUR / B-cell epitope
Unknown protein product 3	*E. m.*	MA2_D2	QRRSLELARLDFDSARKKKESCTSEEKIR	29	E.m. specific	IUR / B-cell epitope
Purine nucleoside phosphorylase	*E. g.*[2)]	MA2_D3	NDDRFGTRFPDMTNLYTKELRELAKTVGKE	30	E.g. specific	B-cell epitope
Uridine cytidine kinase I	*E. g.*	MA2_D4	TRLSRKVIRDVVQRKRNLDAVLDNYFKFVK	30	E.g. specific	B-cell epitope
Proteasome subunit alpha 7	*E. m.*	MA2_D5	EVVQTGAKHMELGVMRRRSDSKPGEPIVE	29	n.a.[3)]	IUR / B-cell epitope
Proteasome subunit alpha 7	*E. m.*	MA2_D6	EEIEEYVKAIERERFEETEKKKQKKEASAS	30	n.a.[3)]	IUR / B-cell epitope

[1)] E. m. = *Echinococcus multilocularis*
[2)] E. g. = *Echinococcus granulosus*
[3)] n.a = not available either due to incomplete N- or C-terminus of the corresponding E.g. EST read

Table 3:
Assessment of peptides C2 and B6 in ELISA with sera from healthy blood donors, echinococcosis patients, patients infected with other helminths and amoebic liver abscess.

		C2	B6
Cut-off		**0.114**	**0.196**
Healthy	**n total**	n positive	n positive
Blood donors	**35**	4	2
Cestode infection			
Echinococcus multilocularis	**42**	8	6
Echinococcus granulosus	**32**	3	7
Taenia solium	**10**	2	1
Nematode infection			
Trichinella spiralis	**2**	0	0
Toxocara canis	**10**	1	2
Anisakis spec.	**1**	0	0
Hookworm	**1**	0	0
Strongyloides stercoralis	**7**	0	0
Loa loa	**4**	1	1
Mansonella perstans	**1**	0	0
Onchocerca volvulus	**2**	0	0
Ascaris spec.	**6**	0	1
Trematode infection			
Paragonymus westermani	**1**	0	0
Fasciola hepatica	**9**	0	1
Schistosoma mansoni	**7**	0	1
Schistosoma haematobium	**2**	1	0
Schistosoma mekongi	**1**	1	0
Amoeba infection			
Entamoeba histolytica	**10**	2	0

Chapter 3:

Diagnostic value of alpha-helical coiled-coil peptides for immunodiagnosis of human echinococcosis

Claudia List [1,3], Christoph Schmid [2,3], Ingrid Felger [1,3]

[1] Swiss Tropical and Public Health Institute, Molecular Diagnostics, Basel, Switzerland

[2] Swiss Tropical and Public Health Institute, Biostatistics and Computational Sciences, Basel, Switzerland

[3] University of Basel, Basel, Switzerland

Manuscript in preparation for publication.

Abstract

The main source of antigen used in routine serodiagnosis of human echinococcosis is hydatid fluid of *Echinococcus granulosus* cysts collected from naturally infected intermediate hosts, e.g. sheep and cattle. Native and recombinant antigens are valuable reagents because they contain a big variety of epitopes, but they generally harbor also cross-reactive epitopes which in turn compromise test specificity. To avoid any further dependency on non-standardized native parasite material, effective strategies were sought for *in silico* selection of immunodominant epitopes that can be mimicked by chemically synthesized peptides. We investigated the diagnostic potential of alpha-helical coiled coil (CC) peptides predicted from conceptually translated *Echinococcus* spp. EST reads available at the Wellcome Trust Sanger Institute (ftp://ftp.sanger.ac.uk/pub/pathogens/Echinococcus/). From the 2 E. *multilocularis* and the 5 *E. granulosus* EST libraries, a total of 158 non-redundant sequences were predicted to contain one or more CC domains of at least 30 amino acids in length. 53 different 30mer peptides were selected and screened for reactivity with patients' sera on a microarray platform. 18/53 peptides exhibited reactivity on microarray. 8/18 peptides were also reactive with sera from echinococcosis patients in ELISA. 7 of these peptides were discarded because of too little reactivity with echinococcosis sera or strong cross-reactivity with non-echinococcosis sera. Peptide MA2_H3 emerged as best performing candidate. It proved to be E. multilocularis specific and reached a sensitivity of 29% for the diagnosis of alveolar echinococcosis (AE) and a specificity of 96%. While such a low sensitivity seems inadequate for a diagnostic test based on this peptide alone, it is a valuable candidate for inclusion into a multi-peptide assay on the grounds of its high specificity.

Introduction

Future developments in serological diagnostics likely consists in highly pure and synthetic analytes and reagents in combination with new detection techniques compatible with multiplexing of diagnostic targets. The selection and production of synthetic antigens such as chemically synthesized peptides, which mimic immunodominant epitopes, would represent a logical step into the next generation of diagnostic tools.

Synthetic peptides are highly pure, well-defined and thus reproducible. Assays based on synthetic peptides can be standardized easily in any lab as the starting material for test production is of consistent quality and also relieves of dependency on parasite material for antigen production. In the past, considerable efforts have been made to develop such synthetic reagents for immunodiagnosis. Synthetic peptides have been applied in diagnosis of a wide range of infectious and autoimmune diseases (reviewed e.g. in [1,2]). Promising results were obtained especially in diagnostics of viral infections such as HIV [3], SARS-CoV (severe acute respiratory syndrome associated coronavirus) [4], *Rubella* virus [5] and others, but also in bacteria, e.g. *Chlamydia trachomatis* [6] and *Mycobacterium tuberculosis* [7] and the protozoan parasite *Toxoplasma gondii* [8]. Synthetic peptides for the diagnosis of helminthic infections have been reported for example from the nematode *Wuchereria bancrofti* [9], the trematodes *Fasciola hepatica* [10] and *Schistosoma mansoni* [11] and the cestodes *Taenia solium* [12] and *Echinococcus granulosus* [13].

Successful application of peptide antigens depends on their ability to mimic native epitopes. The majority of native epitopes is conformational [14]. In the present work, we have tested the diagnostic potential of peptides corresponding to alpha-helical coiled-coil domains. This folding motif is readily adopted by synthetic peptides *in vitro* [15], and hence likely mimics the native conformation of the synthesized sequence. Alpha-helical coiled-coils are among the most common conformations found in proteins and are either of

fibrous or globular type [16]. Current estimates indicate that approximately 5-10% of sequences from the various genome projects encode coiled-coil regions [17]. A survey of GenBank entries revealed that in pathogens and parasites many extended coiled-coil domains (>75 amino acids in length) were external and binding to host proteins [18]. Detailed immunological studies showed alpha-helical coiled-coil proteins to possess antigenic characteristics [19-21].

A heptad repeat underlies alpha-helical coiled-coils. This motif is denoted $(abcdefg)_n$, whereby positions a and d are usually occupied by nonpolar residues. This arrangement places the hydrophobic residues at one face of the alpha-helix promoting its association with a complementary surface to form a left-handed supercoil [18]. In absence of an interchain disulphide bridge, the minimal requirement in peptide length for formation of stable two-stranded alpha-helical coiled-coils was 3 heptad repeats [22], i.e. 21 amino acids. Such length can readily be achieved by chemical synthesis.

For this study all EST libraries currently available in the public domain were downloaded and conceptually translated. Because to date no assembly of any *Echinococcus* genome is available, the EST libraries constitute the only source for genome-wide data mining. Bioinformatic secondary structure prediction was applied to select for peptides with high predicted probabilities to form alpha-helical coiled-coils. The aim was to identify novel diagnostic reagents and to validate the genome-wide approach to antigen prediction for identification of diagnostic peptides. The underlying hypothesis was that not only the native conformation would be mimicked by these 30mer synthetic peptides, but also native epitopes that are reactive with patients' sera.

Materials and methods

Bioinformatics

All EST libraries available at the Wellcome Trust Sanger Institute (ftp://ftp.sanger.ac.uk/pub/pathogens/Echinococcus/) were downloaded. Two approaches were used for their conceptual translation: i) a Perl script, which translated all 6 frames and selected the longest open reading frame (ORF) as definitive translation and ii) the FrameDP algorithm, which was developed by Gouzy and co-workers to predict coding sequences in “noisy” transcripts with possible frameshifts and despite missing start/stop codons [23]. Output data of both algorithms were combined and used as theoretical protein database for both *Echinococcus* species. The quality of this database was assessed by a blastp query [24] for several arbitrarily selected *Echinococcus* spec. sequences: antigen B1 and B2 (Uniprot Q5EKQ4 and D1MH06), EmII/3 (Uniprot Q05768), P29 (Uniprot Q9U8G7), protoscolex specific coiled-coil protein (Uniprot Q8I7B6) and EM13 (Uniprot Q07840).

The highly redundant theoretical protein database was used as input for alpha-helical coiled-coil (CC) prediction by Paircoil2 [25]. To eliminate redundancy, each predicted CC domain of at least 30 amino acids in length was blasted against the full data set of predicted CC domains. Three categories were created: group 1 contained sequences for which both the Perl and FrameDP algorithm had produced identical translations and which matched to only one EST read. Group 2 contained sequences with identical Perl and FrameDP translations, but matches were found with more than one EST reads with few amino acids difference. Group 3 sequences were detected either in the Perl or the FrameDP translation only.

In total, 158 sequences with predicted CC-domains were selected from groups 1-3 and submitted to further analysis by StableCoil, an algorithm designed to predict not only the location, but also the relative stability of coiled-coils [26]. Peptide sequences for chemical synthesis were selected

from stretches of overlapping Paircoil2 and StableCoil prediction from group 1 and 2 proteins.

Peptide synthesis

Peptides were produced by Fmoc solid phase synthesis with C-terminal amidation (Alta Bioscience, University of Birmingham, UK). The batch synthesis procedure limited the length of the peptides to 30 amino acids. All peptides received an N-terminal biotin attached via PEG-spacer. Thus, both N-terminal and C-terminal peptide ends were uncharged, mimicking natural segments of internal protein sequences. Biotin was used to immobilize the peptides onto a streptavidin-coated surface, i.e. microscope glass slide for microarrays or 96-well plates for ELISA.

Peptides were supplied as crude synthesis product. In the absence of specification of peptide concentration, a consistent yield of 1mg of full-length peptide was assumed.

Human sera

Ethical clearance for retrospective use of anonymized patient sera for test development and quality control was obtained from the ethical committee (Ethikkommission beider Basel).

The echinococcosis sera used in this study consisted of samples from 33 CE and 43 AE adult patients living in Central Europe. They were the same serum samples as used in a previous study [27]. All patients had active hepatic lesions of either CE1 or CE2 type (WHO-IWGE standardized classification). Echinococcosis infections were diagnosed as described by Müller et al. [28]. All samples were collected prior to any therapeutic intervention, i.e. before surgery and/or chemotherapy.

In order to test whether peptides were also recognized by sera from patients infected with other helminths than *Echinococcus* spec., the following 74 sera were used: 2 trichinellosis, 10 toxocariasis, 6 ascariasis, 1 anisakiasis, 1 hookworm infection, 7 strongyloidiasis, 7 filariasis (*Loa loa, Mansonella perstans, Onchocerca volvulus*), 9 fascioliasis, 1 paragonimiasis, 10 schistosomiasis (*Schistosoma mansoni*, S. *haematobium*, S. *mekongi*), 10 neurocysticercosis (*Taenia solium*) and 10 amoebic liver abscess (*Entamoeba histolytica*). Concomitant echinococcosis in these 74 patients was ruled out by clinical and serological criteria.

Serum samples from healthy blood donors living in Switzerland were used for cut-off calculation to discriminate between positive and negative test results (n = 8 in microarray assay; n= 52 in ELISA), specificity testing (n = 35 in ELISA) and as negative controls in ELISA (n = 2).

Microarray

Microarray slides were spotted by Alta Bioscience, University of Birmingham, UK. Each slide contained in duplicate 53 different *Echinococcus* peptides plus a spotting control (TAMRA-labeled unrelated peptide), and human IgG as a control for the performance of the conjugate. Duplicates were arranged in two blocks. Signals were analyzed separately for each block. Therefore, per slide zero, one or two positive signals could be obtained for each peptide.

The microarray screening was performed as previously described [27]. Briefly, the microarray slides were blocked over night in assay buffer (1 x PBS pH 7.4, 0.05% Tween20, 3% milk powder), incubated for two hours with sera diluted 1:100 in assay buffer, washed extensively in PBS-T (1 x PBS pH 7.4, 0.05% Tween20) and incubated for one hour with a 1:100 dilution of Cy5-labeled goat anti-human IgG (H+L) conjugate (Jackson ImmunoResearch Laboratories product number 109-175-088) diluted in assay buffer. Slides were washed extensively in PBS-T first, then PBS and before drying with compressed air quickly rinsed with deionized water.

64 microarrays were incubated with following sera: 8 blood donor sera, 1 pool made of 5 AE sera, 20 single AE sera, 10 CE sera, 3 sera from other cestode infections (2 *Taenia solium*, 1 *T. saginata*), 13 sera from nematode infections (1 *Trichinella spiralis*, 1 *Trichuris trichura*, 2 *Ascaris* spec., 1 hookworm, 1 *Mansonella perstans*, 1 *Loa loa*, 1 *Onchocerca volvulus*, 1 *Strongyloides stercoralis*, 2 *Toxocara canis*), 6 from trematode infections (2 *Schistosoma mansoni*, 2 *S. haematobium*, 4 *Fasciola hepatica*) and 3 sera from patients with amoebic liver abscess.

The dried slides were scanned at 532 and 635 nm using a GenePix 4100A microarray scanner (Bucher Biotec AG, Basel, Switzerland). Cy5 and TAMRA images and local background corrected fluorescence intensities (FI) were acquired using the Axon GenePix Pro 6.0 software. Further analysis was done in Microsoft Excel. The cut-off fluorescence intensity (FI) was calculated for every peptide feature separately as average blood donor FI plus three times the standard deviation.

ELISA

ELISAs were done in 96-well plates (NUNC Immobilizer Streptavidin) as described elsewhere [27] with the following modifications: initially all peptide candidates were coated at a concentration of 4ug/ml. For the best performing peptide (MA2_H3) the optimal coating concentration was determined by peptide titration to be 0.025ug/ml.

The absorbance was measured at 405nm (A_{405}) in a VERSAmax microplate reader. The raw data was transferred into Microsoft Excel and blank corrected. The cut-off value was calculated as average blood donor A_{405} value plus 2 standard deviations from the results obtained from 52 healthy blood donors.

The sensitivity (true positives / true positives + false negatives) was calculated for AE and CE sera separately. Test blood donor samples (n=35) different from those used for cut-off calculation together with the samples from helminthic infections other than echinococcosis (n=64) served for calculation of specificity (true negatives / true negatives + false positives).

Results

Most translations of EST reads produced similar results by both algorithms with one obvious difference being the unequal length of protein sequences. FrameDP tended to shorten the protein sequences N-terminally to produce methionine at the starts. In addition, FrameDP corrected for frameshifts leading to extensions at the C-terminus when compared to the corresponding Perl-translated sequences. Considering the different capacities of these translation algorithms, we decided to combine the outputs. Thus, the 2 *E. multilocularis* and the 5 *E. granulosus* EST libraries containing a total of 23'684 EST reads produced 38'289 theoretical protein sequences by both translation algorithms together. To check the quality of these translations, our database was searched by blastp for the presence of 6 known *Echinococcus* antigens. All but one protein were detected.

All 38'289 theoretical protein sequences were subjected to alpha-helical coiled-coil prediction by Paircoil2. 626 protein sequences were predicted to contain one or more CC domains. Redundant sequences were removed leading to a final data set of 158 sequences (table 1). The corresponding translated EST reads were further analyzed by StableCoil. The Paircoil2 and StableCoil algorithms agreed in the prediction of CC domains in 127/158 sequences. From these regions of overlapping Paircoil2 and StableCoil predictions, 53 peptides were selected (table 2). The respective parent EST sequences were blasted against the translated protein database in order to identify reads with similar sequences. This made it possible to assess the species specificity of these peptide sequences. 18 peptides were specific for *E. multilocularis*, 25 were specific for *E. granulosus*, 5 were identical in both

species and 3 showed a difference of 2, 3 or 4 amino acids between the two species (table 2).

Due to the limited availability of patient's sera, the screening of the peptide candidates was performed by microarray. This technique uses only 2ul of sera for parallel screening of 53 different peptides. Fluorescence intensities from the microarray experiments were translated into a one digit value and the sum of positive signals is shown in table 3. Peptides MA2_F4 and MA2_F12 and MA2_H3, the three peptides most reactive with sera from echinococcosis patients (see peptide score 1 in table 3), were found to be also most cross-reactive with sera from infections other than echinococcosis (see peptide score 2 in table 3). 18/53 peptides yielded ≥1 positive signals in microarray. These 18 peptides were transferred into the 96-well ELISA format to further evaluate the potential diagnostic reagents in the format currently used in routine diagnostics.

In ELISA, the peptide candidates were first assessed with sera from 3 AE and 3 CE infections as well as 1 blood donor sample. 7 out of 18 peptides were reactive with one or more of these echinococcosis sera and hence were further tested with an extended serum panel of additional echinococcosis patients, additional blood donors and patients with other parasite infections.

Peptides MA2_F4 and MA2_F12 which were among the peptides with the highest reactivity in microarray assays with peptide score-1 of 9 and 5, respectively (table 3), repeatedly gave no signal in ELISA. Peptide MA2_F2 (peptide score-1 = 5) was excluded because of low sensitivity in ELISA. Peptides MA2_H3 (peptide score-1 = 5) and MA2_E8 (peptide score-1 = 4) and MA2_E10 showed promising results with echinococcosis sera. Peptides MA2_E8 and MA2_E10 disqualified due to high cross-reactivity with both, blood donor sera and sera from infections other than echinococcosis. Peptide MA2_H3 emerged as best candidate and was subsequently tested with a large serum panel.

A cut-off of A_{405} = 0.100 was determined for MA2_H3. This peptide proved to be reactive exclusively with sera from AE patients (table 4). 12/42 AE samples were tested positive resulting in a sensitivity of 29%. Limited cross-reactivity was observed with 1/10 neurocysticercosis sera and 1/9 fasciolosis sera and 2/35 blood donor sera (table 4) leading to a specificity of 96%.

Discussion

The literature holds records of synthetic peptides of good diagnostic value showing the same or even higher sensitivity than the respective native parent proteins [10,13], but reports also about reduced sensitivity [29,30]. Reduced sensitivity of peptides is likely due to a much shorter length compared to recombinant proteins or native antigen extracts and therefore potentially less epitopes. This hypothesis about antigen-length was supported by our previous work where a single 42mer *E. multilocularis* synthetic peptide had shown a rather high sensitivity of 57% in a comparable evaluation procedure [27]. This peptide derived from antigen B8/1 (accession number BAC77657). Native antigen B8/1 comprises 65 amino acids in length (without signal peptide) and thus represents a comparatively small protein. An advantageous coincidence, since this diagnostically well-proven major component of hydatid cyst fluid might be produced entirely by chemical synthesis.

Another explanation of decreased sensitivity of synthetic peptides was loss of conformational epitopes [30]. Using synthetic peptides in ELISA, adequate coating of such short antigens was recognized as key factor in test reproducibility. In the present work, we circumvented this problem by using a biotin-streptavidin system for peptide immobilization. Peptides were synthesized with an N-terminal biotin separated from the peptide by a PEG-spacer, thus providing full flexibility and preserving the entire peptide sequence for unrestricted antibody binding. Coating the peptide antigens to streptavidin ELISA plates prevented epitope masking or destruction caused by passive adsorption of the peptides to the polystyrene surface of conventional ELISA plates. Instead immobilization via biotin-streptavidin enabled a

directional and highly dense coating (every streptavidin has four biotin binding sites), resulting in reproducible test results.

The diagnostic value of peptide antigens also depends on their ability to mimic native epitopes. We aimed at the selection of 30mer peptides that are predicted to adopt a secondary structure that remains stable and reproducible under test conditions and in addition reflect the native conformation. The alpha-helical coiled-coil motif meets both of these requirements.

From a total of 23'684 redundant EST reads, 158 non-redundant sequences had one or more predicted alpha-helical coiled-coil domains. 53 peptides from stretches of predicted alpha-helical coiled-coils were selected for screening on a microarray platform for reactivity with patients' sera. Validation by ELISA, the assay format currently used in routine diagnostics, followed on the 18 peptides identified by microarray screening. 8/18 peptides were also reactive with echinococcosis sera in ELISA. Further validation with extended panels of sera revealed either low sensitivity with echinococcosis sera and/or high crossreactivity with sera from other helminthic infections, for example peptide MA2_E8 which derived from EST read EGPSPsl-5d04.p1k with similarity to the *Schistosoma mansoni* circulating cathodic antigen (CAA; accession number O02197). This peptide obtained promising results with echinococcosis sera, but had to be dismissed because of strong cross-reactivity not only with *Schistosoma mansoni*, *S. haematobium* and *S. mekongi*, but also with *Fasciola hepatica* and several different nematode infections. Monoclonal antibodies against CAA are used in urinary dip stick tests for the detection of schistosomiasis [31]. This test was shown to be of limited use for the detection of *S. haematobium* infection [32]. While in our tests the observed cross-reactivity made peptide MA2_E8 unsuitable for echinococcosis diagnostics, it nonetheless hints at the possibility to improve the dip stick assay for detection of *S. haematobium* infection.

The most promising peptide was MA2_H3 that corresponds to EST-read EGPSgr-14a07.p1k from an oligo-capped cDNA library of the protoscolex

stage of *E. granulosus* (Wellcome Trust Sanger Centre). In order to identify orthologs in both *Echinococcus* species, a blastp search was performed in all translated ESTs. No other match besides the original EST read EGPSgr-14a07.p1k was obtained. This suggested that the hypothetical protein identified likely is *E. granulosus* specific. But in contrast to this expectation, our immunologic screening indicated that peptide MA2_H3 was reactive only with AE sera (*E. multilocularis)*. 12/42 AE patients tested positive, while none of the 32 CE sera were positive in MA2_H3-ELISA. MA2_H3 had a sensitivity of 29% for AE infection. The panel of echinococcosis sera needs to be increased to elucidate this issue further and to confirm the observed *E. multilocularis* specific reactivity of MA2_H3.

Specificity for only one *Echinococcus* species was previously observed for antigen EmII/3, a protein transcribed and translated in both, *E. multilocularis* and *E. granulosus* [33,34]. EmII/3 was shown to localize to the undifferentiated germinal layer of the *E. multilocularis* metacestode tissue [35]. The orthologs have 98.6 % sequence similarity at amino acid level. But only about 6% of CE patients have antibodies against recombinant EmII/3-10, a fragment of EmII/3 [36]. Differing cyst biology of the well-delineated and encapsulated *E. granulosus* cyst as opposed to the infiltrative growing *E. multilocularis* metacestode, where parasite cells probably are in direct contact to host cells, was a suggested reason for the differential recognition of EmII/3-10 [33]. However, recombinant EmII/3-10 has been successfully used for differential diagnosis of *E. multilocularis* infection [37].

Specificity of MA2_H3 determined to be 96% based on extensive testing with control sera. Very few false positive results were obtained, i.e. in 1/10 patients infected with *Taenia solium*, 1/9 patients infected with *Fasciola hepatica* and 2/35 blood donors.

Despite the fact that no homologue of read EGPSgr-14a07.p1k was found in both *E. multilocularis* EST-libraries, a tblastn search at http://www.sanger.ac.uk/cgi-bin/blast/submitblast/Echinococcus against the *E.*

multilocularis contigs from shotgun sequencing returned a scaffold which contained our peptide sequence. In lack of any annotated *Echinococcus* genome data, we carried out gene prediction with GeneMark [38] using the prediction model for *Caenorhabditis elegans* (data not shown*)*. An ORF with two predicted exons was retrieved. The full-length predicted protein comprised 137 amino acids. GPI-som [39] predicted a C-terminal signal peptide, but no GPI-anchor site. No transmembrane domain was found (TMHMM [40]). PSORT II [41] predicted the protein to be most likely cytosolic. A blastp search at NCBI did neither find any conserved domains nor any significant similarity to any database entry (data not shown). The finding that this unknown protein seems to have no homology and is potentially even unique to *Echinococcus* spp., argues for further evaluation as diagnostic antigen. From such an antigen, fewer false positive results might be expected than from well conserved protein domains.

In conclusion, peptide MA2_H3, with a sensitivity of 29% and a specificity of 96%, has great potential to become a further component of a multiplexed diagnostic test for echinococcosis diagnosis. We envisage a test platform that can accommodate a panel of independent synthetic antigens, either targeting echinococcosis alone or in combination with other helminthic infections. New technologies, such as the Luminex microsphere based multiplexing system provide new options for parallel measurements. Synthetic peptides are ideal reagents for these perspectives, as their termini can be modified easily to facilitate coating of beads or microarrays. We anticipate an increase in sensitivity and specificity in multiplexed testing. In particular gains in sensitivity are likely as individual cut-offs for each peptide could be maintained. It is also conceivable that a greater number of peptides could be combined to allow diagnosis of several infectious agents in parallel in a single assay.

It is likely that using the here described methods for peptide identification could lead to a sensitive and specific serodiagnosis independent of native parasite extracts.

Acknowledgements

We are most grateful to Prof. Bruno Gottstein and Prof. Norbert Müller from the institute of parasitology in Bern for provision of human echinococcosis sera and to Prof. Christoph Hatz and Dr. Stefanie Kramme from the medical and diagnostic services unit of the Swiss TPH in Basel for provision of human sera from infections other than echinococcosis. Without their support, this study would not have been possible. We are also much obliged to Pascal Mäser for the Perl script.

Financial support was given by Velux Foundation (http://www.veluxstiftung.ch/). CL was supported by Rudolf Geigy Foundation (http://www.sti.ch/en/about-us/foundations.html), Fonds zur Förderung von Lehre und Forschung (http://www.fag-basel.ch/frameset.html) and Roche Research Foundation (http://www.research-foundation.org/rrf/).

References

1. Noya O, Patarroyo ME, Guzman F, de Noya BA: **Immunodiagnosis of parasitic diseases with synthetic peptides.** *Current Protein and Peptide Science* 2003, **4:**299-308.

2. Gómara MJ, Haro I: **Synthetic peptides for the immunodiagnosis of human diseases.** *Current Medicinal Chemistry* 2007, **14:**531-546.

3. Alcaro MC, Peroni E, Rovero P, Papini AM: **Synthetic peptides in the diagnosis of HIV infection.** *Current Protein & Peptide Science* 2003, **4:**285-290.

4. Chan PKS, To WK, Liu EYM, Ng TK, Tam JS, Sung JJY, Lacroix JM, Houde M: **Evaluation of a peptide-based enzyme immunoassay for anti-SARS coronavirus IgG antibody.** *Journal of Medical Virology* 2004, **74:**517-520.

5. Zrein M, Joncas JH, Pedneault L, Robillard L, Dwyer RJ, Lacroix M: **Comparison of A Whole-Virus Enzyme-Immunoassay (Eia) with A Peptide-Based Eia for Detecting Rubella-Virus Immunoglobulin-G Antibodies Following Rubella Vaccination.** *Journal of Clinical Microbiology* 1993, **31:**1521-1524.

6. Morré SA, Munk C, Persson K, Kruger-Kjaer S, van Dijk R, Meijer CJLM, van den Brule AJC: **Comparison of three commercially available peptide-based immunoglobulin G (IgG) and IgA assays to microimmunofluorescence assay for detection of *Chlamydia trachomatis* antibodies.** *Journal of Clinical Microbiology* 2002, **40:**584-587.

7. Scarpellini P, Tasca S, Galli L, Beretta A, Lazzarin A, Fortis C: **Selected pool of peptides from ESAT-6 and CFP-10 proteins for detection of *Mycobacterium tuberculosis* infection.** *Journal of Clinical Microbiology* 2004, **42:**3469-3474.

8. Kong JT, Grigg ME, Uyetake L, Parmley S, Boothroyd JC: **Serotyping of *Toxoplasma gondii* infections in humans using synthetic peptides.** *Journal of Infectious Diseases* 2003, **187:**1484-1495.

9. Madhumathi J, Pradiba D, Prince P, Jeyaprita P, Rao D, Kaliraj P: **Crucial epitopes of *Wuchereria bancrofti* abundant larval transcript recognized in natural infection.** *European Journal of Clinical Microbiology & Infectious Diseases* 2010,1-6.

10. Intapan PM, Tantrawatpan C, Maleewong W, Wongkham S, Wongkham C, Nakashima K: **Potent epitopes derived from *Fasciola gigantica* cathepsin L1 in peptide-based immunoassay for the serodiagnosis of human fascioliasis.** *Diagnostic Microbiology and Infectious Disease* 2005, **53:**125-129.

11. de Oliveira EJ, Kanamura HY, Takei K, Hirata RDC, Nguyen NY, Hirata MH: **Application of synthetic peptides in development of a serologic method for laboratory diagnosis of Schistosomiasis mansoni.** *Memorias do Instituto Oswaldo Cruz* 2006, **101:**355-357.

12. Fleury A, Beltran C, Ferrer E, Garate T, Harrison LJS, Parkhouse RME, Garcia E, Fragoso G, Costa-Cruz J, Biondi G et al.: **Application of synthetic peptides to the diagnosis of neurocysticercosis.** *Tropical Medicine & International Health* 2003, **8:**1124-1130.

13. González-Sapienza G, Lorenzo C, Nieto A: **Improved immunodiagnosis of cystic hydatid disease by using a synthetic peptide with higher diagnostic value than that of its parent protein, *Echinococcus granulosus* antigen B.** *J Clin Microbiol* 2000, **38:**3979-3983.

14. Huang J, Honda W: **CED: a conformational epitope database.** *BMC Immunology* 2006, **7:**7.

15. Su JY, Hodges RS, Kay CM: **Effect of chain-length on the formation and stability of synthetic alpha-helical coiled coils.** *Biochemistry* 1994, **33:**15501-15510.

16. Parry DAD, Fraser RDB, Squire JM: **Fifty years of coiled-coils and alpha-helical bundles: A close relationship between sequence and structure.** *Journal of Structural Biology* 2008, **163:**258-269.

17. Walshaw J, Woolfson DN: **SOCKET: A program for identifying and analysing coiled-coil motifs within protein structures.** *Journal of Molecular Biology* 2001, **307:**1427-1450.

18. Odgren PR, Harvie LW, Fey EG: **Phylogenetic occurrence of coiled coil proteins: implications for tissue structure in metazoa via a coiled coil tissue matrix.** *Proteins-Structure Function and Genetics* 1996, **24:**467-484.

19. Corradin G, Villard V, Kajava AV: **Protein structure based strategies for antigen discovery and vaccine development against malaria and other pathogens.** *Endocrine Metabolic & Immune Disorders-Drug Targets* 2007, **7:**259-265.

20. Jiz M, Wu HW, Meng R, Pond-Tor S, Reynolds M, Friedman JF, Olveda R, Acosta L, Kurtis JD: **Pilot-scale production and characterization of paramyosin, a vaccine candidate for schistosomiasis japonica.** *Infection and Immunity* 2008, **76:**3164-3169.

21. Tripet B, Kao DJ, Jeffers SA, Holmes KV, Hodges RS: **Template-based coiled-coil antigens elicit neutralizing antibodies to the SARS-coronavirus.** *Journal of Structural Biology* 2006, **155:**176-194.

22. Kwok SC, Hodges RS: **Effect of chain length on coiled-coil stability: Decreasing stability with increasing chain length.** *Biopolymers* 2004, **76:**378-390.

23. Gouzy J, Carrere S, Schiex T: **FrameDP: sensitive peptide detection on noisy matured sequences.** *Bioinformatics* 2009, **25:**670-671.

24. Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucl Acids Res* 1997, **25:**3389-3402.

25. McDonnell AV, Jiang T, Keating AE, Berger B: **Paircoil2: improved prediction of coiled coils from sequence.** *Bioinformatics* 2006, **22:**356-358.

26. Tripet B, Hodges RS: **StableCoil: an algorithm designed to predict the location and relative stability of coiled-coils in native protein sequences.** In *Peptides: The Wave of the Future*; Edited by Lebl M, Houghten RA 365-366.

27. List C, Qi W, Maag E, Gottstein B, Müller N, Felger I: **Serodiagnosis of *Echinococcus* spp. Infection: Explorative Selection of Diagnostic Antigens by Peptide Microarray.** *PLoS Negl Trop Dis* 2010, **4:**e771.

28. Müller N, Frei E, Nuñez S, Gottstein B: **Improved serodiagnosis of alveolar echinococcosis of humans using an *in vitro* produced *Echinococcus multilocularis* antigen.** *Parasitology* 2007, **134:**879-888.

29. Hernández M, Beltrán C, García E, Fragoso G, Gevorkian G, Fleury A, Parkhouse M, Harrison L, Sotelo J, Sciutto E: **Cysticercosis: towards the design of a diagnostic kit based on synthetic peptides.** *Immunology Letters* 2000, **71:**13-17.

30. Barbieri M, Fernandez V, Gonzalez G, Luaces VM, Nieto A: **Diagnostic evaluation of a synthetic peptide derived from a novel antigen B subunit as related to other available peptides and native antigens used for serology of cystic hydatidosis.** *Parasite Immunology* 1998, **20:**51-61.

31. van Dam GJ, Wichers JH, Ferreira TMF, Ghati D, van Amerongen A, Deelder AM: **Diagnosis of Schistosomiasis by Reagent Strip Test for Detection of Circulating Cathodic Antigen.** *J Clin Microbiol* 2004, **42:**5458-5461.

32. Stothard JR, Sousa-Figueiredo JC, Standley C, Van Dam GJ, Knopp S, Utzinger J, Ameri H, Khamis AN, Khamis IS, Deelder AM et al.: **An evaluation of urine-CCA strip test and fingerprick blood SEA-ELISA for detection of urinary schistosomiasis in schoolchildren in Zanzibar.** *Acta Tropica* 2009, **111:**64-70.

33. Felleisen R, Gottstein B: **Comparative analysis of full length antigen II/3 from *Echinococcus multilocularis* and *E. granulosus*.** *Parasitology* 1994, **109:**223-232.

34. Frosch PM, Mühlschlegel F, Sygulla L, Hartmann M, Frosch M: **Identification of a cDNA clone from the larval stage of *Echinococcus granulosus* with homologies to the *E. multilocularis* antigen EM10-expressing cDNA clone.** *Parasitology Research* 1994, **80:**703-705.

35. Felleisen R, Gottstein B: ***Echinococcus multilocularis*: Molecular and immunochemical characterization of diagnostic antigen II/3-10.** *Parasitology* 1993, **107:**335-342.

36. Gottstein B, Jacquier P, Bresson-Hadni S, Eckert J: **Improved primary immunodiagnosis of alveolar echinococcosis in humans by an enzyme-linked immunosorbent assay using the Em2^{plus} antigen.** *J Clin Microbiol* 1993, **31:**373-376.

37. Müller N, Gottstein B, Vogel M, Flury K, Seebeck T: **Application of a recombinant *Echinococcus multilocularis* antigen in an enzyme-linked immunosorbent assay for immunodiagnosis of human alveolar echinococcosis.** *Molecular and Biochemical Parasitology* 1989, **36:**151-160.

38. Besemer J, Borodovsky M: **GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses.** *Nucl Acids Res* 2005, **33:**W451-W454.

39. Fankhauser N, Maser P: **Identification of GPI anchor attachment signals by a Kohonen self-organizing map.** *Bioinformatics* 2005, **21:**1846-1852.

40. Krogh A, Larsson Br, von Heijne G, Sonnhammer ELL: **Predicting transmembrane protein topology with a hidden markov model: application to complete genomes.** *Journal of Molecular Biology* 2001, **305:**567-580.

41. Nakai K, Horton P: **PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization.** *Trends in Biochemical Sciences* 1999, **24:**34-35.

Tables

Table 1:
Processing of EST reads in numbers.

Library	Reads	Perl translations	Perl CC*	FrameDP translations	FrameDP CC*	CC without redundancy[1)]	Tested on microarray
EGCWgr	1850	1850	17	1188	22	10	1
EGPSPgr	2292	2292	14	1282	13	9	3
EGPSPsl	2891	2891	103	2267	89	41	15
EGPSgr	2348	2348	18	1271	17	9	3
EGPSsl	1602	1602	27	911	17	8	4
EMmg	2943	2943	47	1927	46	24	7
Emme	9758	9758	107	5759	89	57	20
Total	**23684**	**23684**	**333**	**14605**	**293**	**158**	**53**

[1)] Numbers of CC refer to sequences that contain one or more domains of alpha-helical coiled-coils predicted by Paircoil2.

Table 2:
Peptides spotted onto microarray and the species specificity of their sequence identified by blastp.

Peptide	Peptide sequence	Parent EST	EST origin	Species specificity of peptide
MA2_D7	RLTLMQAEESLGVLDARLNDARYQYERLRR	EGPSPsl-8a05.q1k	*E. granulosus*	n.a.[1)]
MA2_D8	ERSLMDLMNEESRLKSELSALNRRIEALVH	EGPSPsl-12f06.q1k	*E. granulosus*	*E. granulosus*
MA2_D9	EIQEKLNREIELRKEQMDIFTQDISELKGE	EGPSsl-3a11.p1k	*E. granulosus*	*E. granulosus*
MA2_D10	DRLSEVKEAYTQASGNIAERSKALAQLSEE	EGPSPsl-12d11.p1k	*E. granulosus*	*E. granulosus*
MA2_D11	TEDDIQTLKEALNVKIKHLVEIKKKLGHTD	EGPSPgr-15b04.q1k	*E. granulosus*	*E. granulosus*
MA2_D12	KRTIESLKKKLDYVSKGLADAQQGLRDLIS	EGPSPgr-3h11.p1k	*E. granulosus*	*E. granulosus*
MA2_E1	KGLESSLQKEKDAVKNLKQDLIRARHNKAT	EGPSPsl-12g01.q1k	*E. granulosus*	*E. granulosus*
MA2_E2	REVIETKASIASKMAELERLQNRLKAINQE	EGPSPsl-10a09.q1k	*E. granulosus*	*E. granulosus*
MA2_E3	KEFEAALSLNAEAQAEIRNLRREVTEFAAT	EGCWgr-7f07.p1k	*E. granulosus*	*E. granulosus*
MA2_E4	RDLRRQANRSRSRVTELSVAAQQANEELTT	EGPSPsl-4f02.q1k	*E. granulosus*	*E. granulosus*
MA2_E5	DDQGEQLDRINEGMDQINEDMKDAEKNLDD	EGPSPsl-8f10.q1k	*E. granulosus*	*E. granulosus*
MA2_E6	AKLNVKLKELRTQLTCLVDSENALKQKLVE	EGPSPsl-13h06.p1k	*E. granulosus*	*E. granulosus*
MA2_E7	TAWKSYNEELERILRKHEKDLLEMRRKVQE	EGPSsl-1a08.p1k	*E. granulosus*	*E. granulosus*
MA2_E8	KQLEQLKMQNRTLETSLEEHELTLEMIMSK	EGPSPsl-5d04.p1k	*E. granulosus*	*E. granulosus*
MA2_E9	NERMEKLEAIVKEMSLRVNELEVQIQRLGQ	EGPSgr-11a03.q1k	*E. granulosus*	*E. granulosus*
MA2_E10	QLKNENRKLYEENNLLKIKVDILLDMLAET	EGPSPsl-13g05.q1k	*E. granulosus*	*E. granulosus*
MA2_E11	TEKRQELENLESTNNKIEEATEKIRTKLEK	EGPSPsl-4h08.q1k	*E. granulosus*	*E. granulosus*
MA2_E12	DRLRELTMANASLTQERNDLQTKLDATLEE	EGPSsl-4d09.q1k	*E. granulosus*	*E. granulosus*
MA2_F1	QKQLKEVEFQETELKRLRDWRDSLKEQLES	EGPSPsl-5e11.p1k	*E. granulosus*	*E. granulosus*
MA2_F2	EEIDSLKTNLDDLKTTLNYLKSENKALVEK	EGPSPsl-13h06.q1k	*E. granulosus*	*E. granulosus*
MA2_F3	KDLADKVENAAINVDRMNFDNGRLQDKLAK	Emme-01d03.q1k	*E. multilocularis*	n.a.[1)]
MA2_F4	RKDLEQIKELKKHLVAEVKFHEEQIERHKE	Emme-22e06.p1k	*E. multilocularis*	Identical in both species
MA2_F5	QKARDRKKNSILLMEQSIRSLQSENQSLRQ	Emme-49a12.p1k	*E. multilocularis*	*E. multilocularis*
MA2_F6	NAIEHAKQVYAKSNVHLKQEISALKKKLIN	EMmg-11e11.q1k	*E. multilocularis*	*E. multilocularis*

Table 2 (continued):

Peptide	Peptide sequence	Parent EST	EST origin	Species specificity of peptide
MA2_F7	SLRHIRKQVTKLDEETALLSKHIDNLAGAE	Emme-16b09.p1k	*E. multilocularis*	*E. multilocularis*
MA2_F8	EWTKMSIERSRWEFERAELQARITVLQNEK	Emme-18g08.p1k	*E. multilocularis*	*E. multilocularis*
MA2_F9	KSKIKDHLSTIRTLQGDVKRRDTELNILRQ	Emme-61h07.q1k	*E. multilocularis*	*E. multilocularis*
MA2_F10	GAKAADLTKTVASLRDKHDDQESLNARLKA	EMmg-7b04.q1k	*E. multilocularis*	*E. multilocularis*
MA2_F11	KLQLSGKRQAIVILTQQLEDAKKEAYQFKL	Emme-22c01.p1k	*E. multilocularis*	*E. multilocularis*
MA2_F12	NLIRYMEAVYKSLTKEVEKLQEVESNIQKR	Emme-37a02.p1k	*E. multilocularis*	4 mismatches to *E.granulosus*
MA2_G1	ELTEDLKFLQETLKASVQDVSSLEEYTTEL	Emme-50a04.q1k	*E. multilocularis*	*E. multilocularis*
MA2_G2	DKVSRAESIYELVDRQIQRLDADMVEFKKA	Emme-09e09.p1k	*E. multilocularis*	*E. multilocularis*
MA2_G3	KAHKTKLEDELQVESATIGTLKDKMDCLNS	Emme-26g09.p1k	*E. multilocularis*	*E. multilocularis*
MA2_G4	SLANQLDRLRLEHADLTRRIICTRRRRQQK	Emme-38d05.p1k	*E. multilocularis*	*E. multilocularis*
MA2_G5	EHAKAKLKETQEELTEEAAVQVEKINKLSE	Emme-46a11.p1k	*E. multilocularis*	*E. multilocularis*
MA2_G6	NLSRLEADLRTRSTNYNNIKGALHALEKKQ	Emme-75a10.p1k	*E. multilocularis*	*E. multilocularis*
MA2_G7	ETRANAEKALFEEHTASFEQRISQLSQRLD	Emmg-19c04.q1k	*E. multilocularis*	*E. multilocularis*
MA2_G8	ELKAMDEEINNLTTNLQQLKDRIKSVESEL	Emme-26f02.p1k	*E. multilocularis*	*E. multilocularis*
MA2_G9	LKVQQENLQSQIEELNAEIKNAEDENYKLE	EGPSgr-11a10.q1k	*E. granulosus*	*E. granulosus*
MA2_G10	RAIQHHIKERLRLEAQLKEIEARAALVEEE	EGPSsl-3e03.q1k	*E. granulosus*	*E. granulosus*
MA2_G11	LLQRVAHLRLHNEELERQYAASEARVKSLT	EGPSPsl-13d11.q1k	*E. granulosus*	*E. granulosus*
MA2_G12	EMHSRILESELSNLRKRNAELMDDIALLKK	Emme-72h03.p1k	*E. multilocularis*	*E. multilocularis*
MA2_H1	KVKNEKLQGELNAAIQAFKDARKLYEIARG	Emmg-5g10.p1k	*E. multilocularis*	Identical in both species
MA2_H2	HEMSDQLVQKIDDMSTRIDDLEKNIGEILN	EGPSPgr-7c10.q1k	*E. granulosus*	*E. granulosus*
MA2_H3	HEFMMKEREKLNEMQLDLMEMLMDIQTMNE	EGPSgr-14a07.p1k	*E. granulosus*	*E. granulosus*
MA2_H4	TSLRRLVDDLAKEKAAKDVELKKHIDDMNE	Emme-44h11.q1k	*E. multilocularis*	Identical in both species
MA2_H5	RRKQIEDLQSEILRLESRFQDMERILVKAR	Emme-48e10.p1k	*E. multilocularis*	*E. multilocularis*
MA2_H6	EVLQTEITRLSKIEESLQRFISDIKALGVE	Emmg-4a08.q1k	*E. multilocularis*	*E. multilocularis*
MA2_H7	EQKAQLKDRLVTLESELSDCKATMATLKKE	Emme-08h08.p1k	*E. multilocularis*	Identical in both species

Table 2 (continued):

Peptide	Peptide sequence	Parent EST	EST origin	Species specificity of peptide
MA2_H8	DRLRADLAQLLEEKKLVEAELNTLQANRER	EMmg-16f03.p1k	*E. multilocularis*	Identical in both species
MA2_H9	SDMNTLSTEVSSLRNDIKRDLEDLRKLEGK	Emme-08d10.p1k	*E. multilocularis*	2 mismatches to *E.granulosus*
MA2_H10	TVKQIITLSAELNTSNKKLREHQDSIDESR	EMmg-12c05.q1k	*E. multilocularis*	3 mismatches to *E.granulosus*
MA2_H11	EELRSALSRLLAKVKEAEAEVEEHRSRRHR	EGPSPsl-6a09.q1k	*E. granulosus*	*E. granulosus*

1) n.a. = not available due to incomplete N- or C-terminus of EST read.

Table 3:
Peptides ranked according to their scores obtained in microarray experiments (peptide score 1= sum true positive results; peptide score 2 = sum of false positive results).

			Sera from AE and CE infections															Other infections									
Origin	Species	Peptide	Em-25	Em-15	Em-45	Em-7	Em-39	Em-36	Em-14	Em-22	Em-28	Em-26	Em-8	Em+	Eg-36	Eg-18	Eg-16	Peptide score-1	N-10	N-6	N-27	N-4	N-33	C-7	T-20	T-21	A-18	Peptide score-2
Emme_22e06.p1k	E.m.	MA2_F4	2	2	0	2	1	0	0	0	0	0	1	0	1	0	0	9	2	2	0	0	0	0	0	1	0	5
Emme_37a02.p1k	E.m.	MA2_F12	2	0	0	0	0	0	0	0	0	1	0	0	2	0	0	5	0	2	0	2	0	0	0	0	2	6
EGPSgr-14a07.p1k	E.g.	MA2_H3	2	0	0	0	0	0	2	0	1	0	0	0	0	0	0	5	1	0	0	0	0	2	0	0	0	3
EGPSPsl-13h06.q1k	E.g.	MA2_F2	0	2	1	0	0	0	0	2	0	0	0	0	0	0	0	5	0	0	0	0	0	0	1	0	0	1
EGPSPsl-5d04.p1k	E.g.	MA2_E8	0	0	2	0	0	0	0	0	0	0	0	0	0	0	2	4	0	0	0	0	0	0	0	0	0	0
Emme_48e10.p1k	E.m.	MA2_H5	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0
EGPSsl-3a11.p1k	E.g.	MA2_D9	0	2	0	0	0	0	0	0	0	0	0	0	0	1	0	3	2	0	0	0	0	0	0	0	0	2
EMmg-19c04.q1k	E.m.	MA2_G7	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	2	0	0	0	0	0	0	2
EGPSgr-11a03.q1k	E.g.	MA2_E9	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0
EMmg-12c05.q1k	E.m.	MA2_H10	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0
EMmg-7b04.q1k	E.m.	MA2_F10	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0
Emme_26g09.p1k	E.m.	MA2_G3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0
EGPSPgr-7c10.q1k	E.g.	MA2_H2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0
EGPSPsl-13g05.q1k	E.g.	MA2_E10	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1
Emme_38d05.p1k	E.m.	MA2_G4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Emme_16b09.p1k	E.m.	MA2_F7	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	0	1
EGCWgr-7f07.p1k	E.g.	MA2_E3	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
EGPSPsl-8a05.q1k	E.g.	MA2_D7	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0

Table 4:
Numbers of positive ELISA results obtained by peptide MA2_H3

		H3
Cut-off		**0.100**
Healthy	**total n**	positive n
Blood donors (cut-off)	**52**	3
Blood donors (test)	**35**	2
Cestode infection		
Echinococcus multilocularis	**42**	12
Echinococcus granulosus	**32**	0
Taenia solium	**10**	1
Nematode infection		
Trichinella spiralis	**2**	0
Toxocara canis	**10**	0
Anisakis spec.	**1**	0
Hookworm	**1**	0
Strongyloides stercoralis	**7**	0
Loa loa	**4**	0
Mansonella perstans	**1**	0
Onchocerca volvulus	**2**	0
Ascaris spec.	**6**	0
Trematode infection		
Paragonymus westermani	**1**	0
Fasciola hepatica	**9**	1
Schistosoma mansoni	**7**	0
Schistosoma haematobium	**2**	0
Schistosoma mekongi	**1**	0
Amoeba infection		
Entamoeba histolytica	**10**	0

Chapter 4:

Diagnostic value of synthetic peptide EmLD89 for follow-up of young cystic echinococcosis patients

Claudia List[1,2], Nadia Ben Nouir[3xa], Ingrid Felger[1,2], Hamouda Babba[4], Bruno Gottstein[3]

[1] Department of Medical Parasitology and Infection Biology, Swiss Tropical and Public Health Institute, Basel, Switzerland

[2] University of Basel, Basel, Switzerland

[3] Institute of Parasitology, University of Bern, Bern, Switzerland

[xa] Current address: Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany

[4] University of Monastir, Faculty of Pharmacy, Department of Clinical Biology B, Laboratory of Parasitology and Mycology, Monastir, Tunisia

Data in preparation for manuscript.

Abstract

The performance of a synthetic peptide as marker for serological follow-up of young cystic echinococcosis (CE) patients was explored by ELISA. From 40 young CE patients (3 to 14 years) from Tunisia, 226 consecutive serum samples were collected at the time point of diagnosis and in regular intervals after initiation of treatment (surgery and/or chemotherapy with albendazole). The patients were followed over 5 years. Based on clinical findings and supported by imaging methods, patients were retrospectively grouped into cured (CCE, n = 31) and non-cured (NCCE, n = 9) cases. All serum samples were comparatively tested with synthetic peptide EmLD89 ELISA versus EgHF-ELISA. In cured patients, 24/31 sera were reactive with peptide EmLD89 at diagnosis, compared to 30 positives with the EgHF-ELISA. All of these 24 patients showed a gradual decrease in antibody levels. 3 years after surgery and/or start of chemotherapy 12 (50%) CCE patients became sero-negative, compared to 5 seronegatives by EgHF-ELISA. All but one NCCE patients exhibited constantly high antibody titers against peptide EmLD89. This study provides proof-of-principle that synthetic peptides are suitable markers for the serological follow-up of CE patients.

Introduction

Cystic echinococcosis (CE) in humans is caused by the development of metacestode larvae of the taeniid tapeworm *Echinococcus granulosus* in the liver, lung, or any other organ. Up to 80% of patients suffer from infections of a single organ and present one solitary cyst [1]. Humans are accidental intermediate hosts and usually represent a dead end in the development of this parasite, since the life cycle of *E. granulosus* relies on infected intermediate hosts (herbivores such as e.g. voles) to be eaten by carnivorous final hosts, e.g. dogs. Intermediate hosts acquire infection by ingestion of eggs, released by adult tapeworms living in the small intestine of definitive hosts via faeces. *Echinococcus* eggs are highly resistant, and may remain infective for about one year in a suitable, moist environment at lower temperatures [2]. Sources of infection to humans may be the handling of infested definitive hosts (eggs have been shown to stick to the coat of dogs) or egg-contaminated soil as well as eating of contaminated vegetables, salads and fruits [2]. Cystic echinococcosis is a cosmopolitan zoonosis closely linked to animal husbandry, thus the highest prevalence rates among humans and animals occur where livestock production is extensive and where dogs have access to offal after uncontrolled slaughter [3].

The initial phase of infection is always asymptomatic and many, especially small cysts (<5cm) may remain without causing symptoms for years or even permanently (up to 60% of the cases) [1]. CE may become symptomatic if growing cysts start to exert pressure on adjacent tissue [1]. The diagnosis of *E. granulosus* infection is based on case history, clinical findings, lesions identified by imaging techniques such as ultrasound (US), computed tomography (CT) or magnetic resonance imaging (MRI) and serological means [4]. In fact, without confirmation by immunological tests, the diagnosis based on imaging methods remains doubtful for certain cysts, which can hardly be distinguished from other cyst-like “space occupying lesions” [5].

Over the years, a wide range of immunodiagnostic techniques have been applied and less adequate methods such as the complement fixation test have been replaced. Now commonly used in routine laboratory testing are the ELISA (enzyme-linked-immunosorbent assay), the IHA (indirect hem agglutination the IFAT (indirect fluorescence antibody test) and IB (immunoblotting).

The treatment of human CE consists of three options: surgery, PAIR (puncture, aspiration, injection, re-aspiration: percutaneous puncture of the cyst guided by ultrasound imaging, followed by aspiration of some cyst fluid and injection of a parasitocidal solution and finally reaspiration) and chemotherapy with benzimidazoles [6]. If the clinical condition of the patient and the location and type of cyst allow PAIR, this would be the treatment of choice. Removing the parasite by PAIR or surgery can lead to complete cure. However, the main complication is relapse due to incomplete removal or spillage of parasite material during surgery, leading to the formation of new cysts (secondary echinococcosis). Careful follow-up of echinococcosis patients after surgery is indicated for several years after treatment. Initial check-ups should be scheduled every 3-6 months and later on once per year [7]. Thus an appropriate strategy for reliable evaluation of treatment outcome is compulsory, leading to quick detection of recurring or newly growing cysts. The use of US and CT scans for follow-up of hydatid patients has become essential [8], but imaging techniques sometimes fail to show small changes or differences between viable, dying-out and even calcifying cysts. Immunodiagnostic tests are therefore highly appreciated tools to complement imaging procedures in the monitoring of treatment outcome, provided they yield the expected prognostic performance.

The major source of antigen applied in tests for both, primary as well as follow-up diagnosis, is *E. granulosus* hydatid cyst fluid [9-11]. Native antigen such as cyst fluid is extracted from various sources and is non-standardized in terms of quality and contents. To overcome limitations associated with native

antigen, different recombinant antigens for diagnosing primary and/or follow-up CE were used. Among these were recombinant B2t [11] and recP29 [12].

A consequent further development would consist in the design of synthetic antigens generated by chemical synthesis. Custom-made and commercial synthesis is widely available. Synthetic peptides can be routinely produced up to a length of 30-50 amino acid residues using a multi-channel peptide synthesizer or even up to 100-150 amino acids using a single-channel synthesizer [13]. Our previous work and that of others have demonstrated that synthetic peptides are well recognized by natively induced antibodies [14,15], probably by mimicking immunodominant epitopes. Previously we identified a 42mer synthetic peptide with a sensitivity of 74% for adult CE cases, which also strongly cross-reacted with alveolar echinococcosis patients infected with the closely related *E. multilocularis* [14]. Hence the peptide EmLD89 was considered as potential marker for serological follow-up of CE patients. We compared the peptide's diagnostic performance to results obtained by crude hydatid cyst fluid and to ELISA results obtained by using recombinant P29. The latter results were published previously [12]. To our knowledge, this is the first study to investigate the performance of a synthetic peptide for the use in immunodiagnostic follow-up.

Materials and methods

Human serum samples

The sensitivity and specificity of the EmLD89 batch used in this study was determined as described previously using an echinococcosis serum panel consisting of 33 CE and 43 AE adult Central European patients [14]. Echinococcosis in these patients had been diagnosed as described by Müller et al. [16]. All patients had active hepatic lesions of either CE1 or CE2 type (WHO-IWGE standardized classification). Blood samples were taken prior to any therapeutic intervention, i.e. before surgery and/or chemotherapy.

The cut-off value for differentiation of positive and negative test results was calculated from the A_{405}-values obtained from 50 healthy blood donors living in Switzerland (average plus 2 standard deviations).

For testing cross-reactivity, sera from 74 different patients were used. 10 patients were infected with *Taenia solium*, 2 with *Trichinella spiralis*, 10 with *Toxocara canis*, 1 with *Anisakis* spec., 1 with hookworms, 7 with *Strongyloides stercoralis*, 4 with *Loa loa,* 1 with *Mansonella perstans,* 2 with *Onchocerca volvulus*, 6 with *Ascaris* spec., 1 with *Paragonismus westermani*, 7 with *Schistosoma mansoni*, 2 with S. *haematobium*, 1 with *S. mekongi*, 9 with *Fasciola hepatica* and 10 with *Entamoeba histolytica.* Concomitant echinococcosis was ruled out by clinical and serological criteria.

The dynamics of the humoral anti-CE response was investigated with 226 serum samples from 40 young patients, aged between 3 and 14 years, diagnosed and treated for CE at the E. P. S. hospital of Monastir, Tunisia. These patients had been part of previous studies investigating the serological follow up of CCE versus NCCE patients [9,12,17]. Some of the patients presented with a relapse at begin of the study, i.e. they had previously been diagnosed with CE and treated accordingly, but treatment was not successful. Briefly, the study design was as follows: all cases were surgically, radiologically and/or histologically proven (table 1). The first serum sample (D0) was collected at the time of diagnosis of CE (or relapsing CE), i.e. before medical intervention in the course of the study which was either surgery and/or chemotherapy with albendazole. Both, regular blood samples were taken and radiological examinations were performed after 7 days (D7), 1 month (1M), 6 months (6M) and once at a later stage up to 5 years after surgery and/or chemotherapy (1Y - 5Y). The intervals and time points of sampling varied between the patients. The number of serum samples per patient ranged from 4 to 7. Hydatid cysts were classified into 5 categories according to WHO criteria [18]. At the endpoint of the follow-up period of 5 years, the patients were clustered into 2 groups according to the types of cyst present: group A consisted of cured or inactive hydatid disease cases. The

cysts had successfully been removed by surgery, had completely calcified (cyst type CE5) or remained inactive during the whole period of follow-up. Patients of group B presented with hydatid disease still ongoing (cyst type CE1, CE2 and CE3), either because surgery had not been feasible or had been incomplete.

Peptide

Peptide EmLD89 (biotin-PEG-KMLGEMKYFFERDPLGQKLVDLLKELEEVFQ MLRKKLRTALK) was custom made by Alta Bioscience, Birmingham, UK. The peptide was synthesized with C-terminal amide modification. “Crude” synthesis was chosen as purity level, because only full-length peptides were supposed to contain biotin and thus be able to bind to streptavidin coated ELISA plates. The shortfall of crude synthesis is an unknown amount of full-length peptide, thus making it necessary to determine optimal coating concentration for every newly synthesized batch of peptide. The peptide concentrations given here assume a consistent yield of 1mg of full-length peptide per synthesis reaction. According to the manufacturer’s recommendations, the lyophilized peptide EmLD89 was dissolved in 25% DMSO in dH_2O to produce stock solutions of a concentration of 4mg/ml.

ELISA

The optimal coating concentration for peptide EmLD89 of 0.25ug/ml was determined by peptide titration. ELISA was performed in 96 well plates (NUNC immobilizer streptavidin, Denmark). Coating was done in 1x PBS pH 7.4 (Sigma product number P4417). Plates were washed twice with PBS pH 7.4, 0.05% Tween20 and then coated over night at 4°C with 100ul/well of peptide dilution. All following steps were carried out at room temperature. The plates were washed twice and blocked for one hour with 150ul/well of assay buffer (3% milk in PBS pH 7.4, 0.05% Tween20, 0.5mM biotin). After three washes, the plates were incubated for one hour with human sera diluted 1:200 in assay buffer. Plates were washed again three times and incubated for one hour with 100µl/well of conjugate diluted 1:1000 in assay buffer (anti-human IgG

alkaline-phosphatase, Sigma product number A 3187) followed by three washes. Wells were then incubated for 10 minutes with 100ul of *p*-nitrophenyl phosphate (Sigma product number N 4645) at a concentration of 1 mg ml^{-1} in substrate buffer (13.2 mM Na_2CO_3, 35 mM $NaHCO_3$, 1mM $MgCl_2x6H_2O$, pH 9.6). Absorbance values (A_{405}) were measured at 405nm in a VERSAmax microplate reader.

All serum samples from a given follow-up patient were tested on the same plate in order to avoid inter-plate test variation. All samples were tested in duplicates. The two values were averaged and blank-corrected.

Data analysis

Raw data from the microplate reader was processed in Microsoft Excel. Average and standard deviation were calculated in Excel. Graphs were created using GraphPad Prism version 5.03 for Windows, GraphPad Software, San Diego California USA, www.graphpad.com.

EgHF-ELISA and recP29-ELISA data from the same patients were retrieved from previous studies documented earlier [9,12,17].

Results

Operating characteristics of peptide EmLD89

The operating characteristics of a newly synthesized batch of EmLD89 were highly similar to those of batch1 EmLD89 assessed in a previous study [14]. The cut-off value of the previous batch EmLD89 was A_{405} = 0.165, that of the current one was A_{405} = 0.166. The current EmLD89 showed a comparable and slightly higher sensitivity for the detection of CE infection in adult patients of 75% compared to 72% of the previous batch EmLD89.

Performance of peptide EmLD89 ELISA in patient follow-up

226 serum samples from 40 patients (31 CCE and 9 NCCE) collected at 9 different time points were tested with EmLD89. These samples had been tested previously in EgHF ELISA and Western blot and recP29 ELISA and Western blot [9,12,17]. A comparison of our EmLD89 results with these antigens is shown in table 2 for both, CCE and NCCE samples.

Baseline samples (D0) were available from 34 patients (27 CCE, 7 NCCE). 18 of these D0 serum samples tested positive in peptide EmLD89 ELISA resulting in a sensitivity of 53% for the detection of CE infection in children and adolescents. D0 sensitivities in these young patients for EgHF- and recP29-ELISA were 77% and 21%, respectively.

The development of antibody levels of CCE (31 patients) and NCCE (9 patients) in response to peptide EmLD89 is shown in figures 2 and 3, respectively. In general, the antibody levels increased after surgery/treatment and peaked at time point 1M followed by a decrease of antibody level visible in the CCE cases (figure 1), while the NCCE cases showed high antibody levels at all times, also at the end of the individual follow-ups (A_{405} 1.113 - 3.879) (figure 2). The peak of anti-*Echinococcus* antibodies 1 month after surgery/chemotherapy was also seen in EgHF- and recP29-ELISA.

24/31 CCE patients were positive for peptide EmLD89 at one or several time points tested. Among these 24 patients, 15 presented with liver cysts, 8 with lung cysts and 1 with multiple cysts in liver and peritonea. From 20 of these 24 patients, a D0 sample was available. 8/20 were EmLD89 negative at D0 (5 lung cases and 3 liver cases) but seroconverted during treatment. The dynamics of antibody recognition of the 24 EmLD89 positive CCE patients were as follows: 10/24 reached antibody titers below the cut-off within 3 years after surgery/chemotherapy (figure 3). One of these was only responding weakly to EmLD89. 10/24 CCE patients showed considerable decrease in anti-EmLD89 antibody levels. At the end of the follow-up period, 7 of these

were still clearly above cut-off while the remaining 3 patients reached A_{405} values only slightly above cut-off (A_{405} = 0.177, 0.192 and 0.172; cut-off = 0.166). 2/24 EmLD89 positive patients showed no major change in antibody levels, but the patients were monitored only for one and two years, respectively. Their corresponding EgHF A_{405}-values were also high at all time points. The remaining 2/24 CCE patients showed patterns not resembling any of the others: both were EmLD89 positive only at one time point during the entire follow up period, namely at 1M and 6M, respectively. 1 CCE patient was positive at D0 and 6M, all other samples from this patient were EmLD89 negative. The corresponding EgHF-ELISAs were highly positive at all time points.

7 CCE patients showed no reactivity at all with peptide EmLD89 throughout the entire period of follow-up (6 lung cases, 1 with multiple cysts in lung and peritonea). One of these 7 patients, patient no. 34, with a lung cyst, showed no reactivity with any antigen in any test applied at any time point tested. The remaining 6 patients were positive at one or more time points tested in EgHF-ELISA and 1 patient was additionally positive in recP29 ELISA.

Regarding the 9 NCCE patients, 8/9 showed high antibody levels at the first time point tested (DO, D7, 1M, respectively; A_{405} > 1.5). 1 NCCE patient remained negative until 2 years after surgery, but had been positive in EgHF ELISA from D7 onwards.

Discussion

Recurrence of cysts is frequently observed in human echinococcosis. A number of reports indicated that about 25% of all CE cases showed relapse following treatment by either surgery or chemotherapy [4,5,19,20]. Despite intensive research, no specific test for monitoring follow-up patients has been developed yet. Current routine diagnostic tests use the same antigens as for the detection of primary echinococcosis. These antigens generally derive from

crude or partially purified hydatid cyst extracts. Improved immunodiagnostic tools for treatment monitoring are still in need today.

The present work assessed the potential usefulness of a synthetic peptide for serological follow-up of CE patients. Peptide EmLD89 had previously been identified by screening on a microarray platform 45 different peptides originating mostly from known antigens and by combining two highly reactive, overlapping peptides (29mer and 27mer) in one long synthetic peptide (42mer) [14]. EmLD89 corresponding to antigen B8/1 of *E. multilocularis* (accession number BAC77657) reacted well with sera from both, adult AE (43% sensitivity) and CE cases (74% sensitivity) [14].

Differences in antibody levels have been attributed to cyst location and/or the number of cysts. Multiple cyst infestation was shown to lead to slower antibody decrease after successful treatment [17]. Cysts in the brain or eye and calcified cysts often induce low or no antibody response [1]. Hepatic cysts are thought to elicit higher antibody levels than pulmonary cysts [21,22]. We found that peptide EmLD89 did not differ in reactivity for sera from liver compared to lung cases. In contrast, 6 of the 7 CCE cases, all non-reactive with EmLD89, had pulmonary cysts. This confirmed previous reports that pulmonary cysts were weakly or not recognized by the host immune system [17].

The sera tested in this study were collected from young CE patients from 3 to 14 years of age [9,12,17]. CE may reach medical attention in all age groups with peaking numbers of cases in 6 to 15 years-old [1]. Diagnosis in children and adolescents is difficult, because their antibody response may be low [1]. Indeed, a lower response was observed also in our study participants. From our 40 patients 34 D0-samples were available, 18 of which were EmLD89-positive. This resulted in a sensitivity of 53% in young patients and compared to a sensitivity in adults of 74%.

From 31 CCE patients, 24 were positive with peptide EmLD89. All NCCE patients were positive for EmLD89. One NCCE patient tested positive only 2 years after treatment. In fact, this patient had a relapse two years after initial surgery. This patient can be considered non-responsive to peptide EmLD89 during the initial infection, whereas the delayed response was caused by the relapse giving rise to an anti-EmLD89 antibody development. The serological profile of this patient tested in standard EgHF-ELISA was considered normal. However, it showed negativity at baseline sampling, but followed by seroconversion after the first surgery. This patient therefore highlights the need for increasing the sensitivity of peptide-based testing. Increased sensitivity might be obtained by complementing EmLD89 with peptides that represent epitopes different from antigen B1 epitopes.

The major difference between the CCE and NCCE group was the gradual decline in antibody levels among the CCE patients measured by A_{405}-values. All but one NCCE case showed constantly high antibody levels even at the end of follow-up (5 years). The trend in antibody-level decrease measured in EmLD89-ELISA corresponded well to that seen in ELISA using native cyst hydatid fluid, but in EmLD89-ELISA more patients became sero-negative within 3 years after treatment compared to outcomes measured by native cyst fluid ELISA (12 patients versus 5 patients in our study population). This reflects a better resolution of antibody decline with peptide EmLD89, a single representative from the antigen B multi-gene-family, as compared to native cyst fluid containing a mixture of several antigen B subunits from different antigen B family members.

Recently, it was shown that the antigen B family comprises at least 10 genes and that all of them are conserved in parasites from different hosts and different geographical locations [23]. In addition, differential expression of antigen B family members in larval and adult developmental stages was found [23]. Antigen B3/1 was the predominant subunit expressed throughout the life cycle in all developmental stages investigated [23].

This recently published information on conservation and expression of antigen B highlights the suitability of this antigen for diagnostic purposes. In our previous study, we tested peptides from antigen B1 and antigen B2 [14]. Peptides corresponding to antigen B sequences were found to be superior in reactivity with patients' sera compared to peptides corresponding to any other sequence. Furthermore, we found that antigen B1 peptides showed higher sensitivities than antigen B2 peptides (data not shown). Regarding the diagnostic operating characteristics of different antigen B subunits, several authors reported opposing results. Our data support the findings of Lorenzo et al. and González-Sapienza et al., who reported recombinant and peptide antigen B1 superior to antigen B2 [24,25]. In contrast, the results of Rott et al., Virginio et al. and Hernández-González et al. made antigen B2 the better performing subunit [11,26,27]. Slight sequence variations and inclusion or exclusion of signal peptide sequences in the recombinant proteins were suggested reasons for the differing results [11]. B2t, the recombinant antigen B2 subunit, was proposed to perform better because of the missing signal sequence and hence representing the mature form of the protein [11]. However, EmLD89 does neither contain the signal peptide sequence and therefore revives the discussion about the reasons of the observed differences. We hypothesize that the discrepancy might be due to technical reasons such as different assay protocols leading for example to variable coating efficiencies of the different antigen B subunits. In addition, the serum panels used for evaluation were different between the studies, raising the question how the sensitivities might look like if the serum panels were exchanged between the opposing working groups.

In the present work, we have provided proof-of-principle for the value of a synthetic peptide as serological marker and strongly suggest the further investigation of more synthetic peptides from the whole spectrum of available antigen B sequences, especially of antigen B3/1. Consistent with reports from other diagnostic peptide studies [28,29], we expect the combination of several peptides to increase test sensitivity substantially. Peptide EmLD89 might represent the first step in the development of a highly sensitive and specific

synthetic antigen pool, resolving the limitations associated with native antigen extracts.

References

1. Pawlowski ZS, Eckert J, Vuitton DA, Ammann RW, Kern P, Craig PS, Dar KF, de Rosa F, Filice C, Gottstein B et al.: **Echinococcosis in humans: clinical aspects, diagnosis and treatment.** In *WHO/OIE manual on echinococcosis in humans and animals: a public health problem of global concern*. Edited by Edited by Eckert J, Gemmell MA, Meslin F-X, Pawlowski ZS. Paris: WHO/OIE; 2001:20-71.

2. Eckert J, Gottstein B, Heath D, Liu F-J: **Prevention of echinococcosis in humans and safety precautions.** In *WHO/OIE manual on echinococcosis in humans and animals: a public health problem of global concern*. Edited by Edited by Eckert J, Gemmell MA, Meslin F-X, Pawlowski ZS. Paris: WHO/OIE; 2001:238-247.

3. Romig T: **Epidemiology of echinococcosis.** *Langenbecks Archives of Surgery* 2003, **388:**209-217.

4. Buttenschoen K, Buttenschoen DC: ***Echinococcus granulosus* infection: the challenge of surgical treatment.** *Langenbecks Archives of Surgery* 2003, **388:**218-230.

5. Siracusano A, Teggi A, Ortona E: **Human cystic echinococcosis: old problems and new perspectives.** *Interdiscip Perspect Infect Dis* 2009, **2009:**474368.

6. Saimot AG: **Medical treatment of liver hydatidosis.** *World Journal of Surgery* 2001, **25:**15-20.

7. Brunetti E, Kern P, Vuitton DA: **Expert consensus for the diagnosis and treatment of cystic and alveolar echinococcosis in humans.** *Acta Tropica* 2010, **114:**1-16.

8. Wen H, New RRC, Craig PS: **Diagnosis and Treatment of Human Hydatidosis.** *British Journal of Clinical Pharmacology* 1993, **35:**565-574.

9. Ben Nouir N, Nuñez S, Gianinazzi C, Gorcii M, Miffler N, Nouri A, Babba H, Gottstein B: **Assessment of *Echinococcus granulosus* somatic protoscolex antigens for serological follow-up of young patients surgically treated for cystic echinococcosis.** *J Clin Microbiol* 2008, **46:**1631-1640.

10. Ortona E, Rigano R, Buttari B, Delunardo F, Ioppolo S, Margutti P, Profumo E, Teggi A, Vaccari S, Siracusano A: **An update on immunodiagnosis of cystic echinococcosis.** *Acta Tropica* 2003, **85:**165-171.

11. Hernandez-Gonzalez A, Muro A, Barrera I, Ramos G, Orduna A, Siles-Lucas M: **Usefulness of four different *Echinococcus granulosus* recombinant antigens for serodiagnosis of unilocular hydatid disease (UHD) and postsurgical follow-up of patients treated for UHD.** *Clin Diagn Lab Immunol* 2008, **15:**147-153.

12. Ben Nouir N, Gianinazzi C, Gorcii M, Müller N, Nouri A, Babba H, Gottstein B: **Isolation and molecular characterization of recombinant *Echinococcus granulosus* P29 protein (recP29) and its assessment for the post-surgical serological follow-up of human cystic echinococcosis in young patients.** *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2009, **103:**355-364.

13. Corradin G, Villard V, Kajava AV: **Protein structure based strategies for antigen discovery and vaccine development against malaria and other pathogens.** *Endocrine Metabolic & Immune Disorders-Drug Targets* 2007, **7:**259-265.

14. List C, Qi W, Maag E, Gottstein B, Müller N, Felger I: **Serodiagnosis of *Echinococcus* spp. Infection: Explorative Selection of Diagnostic Antigens by Peptide Microarray.** *PLoS Negl Trop Dis* 2010, **4:**e771.

15. Villard V, Agak GW, Frank G, Jafarshad A, Servis C, Nebie I, Sirima SB, Felger I, revalo-Herrera M, Herrera S et al.: **Rapid identification of malaria vaccine candidates based on alpha-helical coiled coil protein motif.** *Plos One* 2007, **2**.

16. Müller N, Frei E, Nuñez S, Gottstein B: **Improved serodiagnosis of alveolar echinococcosis of humans using an *in vitro* produced *Echinococcus multilocularis* antigen.** *Parasitology* 2007, **134:**879-888.

17. Ben Nouir N, Nuñez S, Frei E, Gorcii M, Müller N, Gianinazzi C, Mekki M, Nouri A, Babba H, Gottstein B: **Post-surgical follow-up (by ELISA and immunoblotting) of cured versus non-cured cystic echinococcosis in young patients.** *Parasitology* 2008, **135:**105-114.

18. Macpherson CNL, Vuitton DA, Gharbi HA, Caremani M, Frider B, Brunettii E, Perdomo R, Schantz PM, Felice C, Teggi A et al.: **International classification of ultrasound images in cystic echinococcosis for application in clinical and field epidemiological settings.** *Acta Tropica* 2003, **85:**253-261.

19. Horton J: **Albendazole for the treatment of echinococcosis.** *Fundamental & Clinical Pharmacology* 2003, **17:**205-212.

20. Franchi C, Di Vico B, Teggi A: **Long-term evaluation of patients with hydatidosis treated with benzimidazole carbamates.** *Clinical Infectious Diseases* 1999, **29:**304-309.

21. Babba H, Messedi A, Masmoudi S, Zribi M, Grillot R, Ambriosethomas P, Beyrouti I, Sahnoun Y: **Diagnosis of human hydatidosis: comparison between imagery and 6 serologic techniques.** *Am J Trop Med Hyg* 1994, **50:**64-68.

22. Verastegui M, Moro P, Guevara A, Rodriguez T, Miranda E, Gilman RH: **Enzyme-linked immunoelectrotransfer blot test for diagnosis of human hydatid disease.** *J Clin Microbiol* 1992, **30:**1557-1561.

23. Zhang WB, Li J, Jones MK, Zhang ZZ, Zhao L, Blair D, McManus DP: **The *Echinococcus granulosus* antigen B gene family comprises at least 10 unique genes in five subclasses which are differentially expressed.** *Plos Neglected Tropical Diseases* 2010, **4**.

24. Lorenzo C, Ferreira HB, Monteiro KM, Rosenzvit M, Kamenetzky L, Garcia HH, Vasquez Y, Naquira C, Sanchez E, Lorca M et al.: **Comparative Analysis of the Diagnostic Performance of Six Major *Echinococcus granulosus* Antigens Assessed in a Double-Blind, Randomized Multicenter Study.** *J Clin Microbiol* 2005, **43:**2764-2770.

25. González-Sapienza G, Lorenzo C, Nieto A: **Improved immunodiagnosis of cystic hydatid disease by using a synthetic peptide with higher diagnostic value than that of its parent protein, *Echinococcus granulosus* antigen B.** *J Clin Microbiol* 2000, **38:**3979-3983.

26. Rott MB, Fernandez V, Farias S, Ceni J, Ferreira HB, Haag KL, Zaha A: **Comparative analysis of two different subunits of antigen B from *Echinococcus granulosus*: gene sequences, expression in *Escherichia coli* and serological evaluation.** *Acta Tropica* 2000, **75:**331-340.

27. Virginio VG, Hernandez A, Rott MB, Monteiro KM, Zandonai AF, Nieto A, Zaha A, Ferreira HB: **A set of recombinant antigens from *Echinococcus granulosus* with potential for use in the immunodiagnosis of human cystic hydatid disease.** *Clinical and Experimental Immunology* 2003, **132:**309-315.

28. Hernández M, Beltrán C, García E, Fragoso G, Gevorkian G, Fleury A, Parkhouse M, Harrison L, Sotelo J, Sciutto E: **Cysticercosis: towards the design of a diagnostic kit based on synthetic peptides.** *Immunology Letters* 2000, **71:**13-17.

29. Shen GM, Behera D, Bhalla M, Nadas A, Laal S: **Peptide-Based Antibody Detection for Tuberculosis Diagnosis.** *Clinical and Vaccine Immunology* 2009, **16:**49-54.

Figures

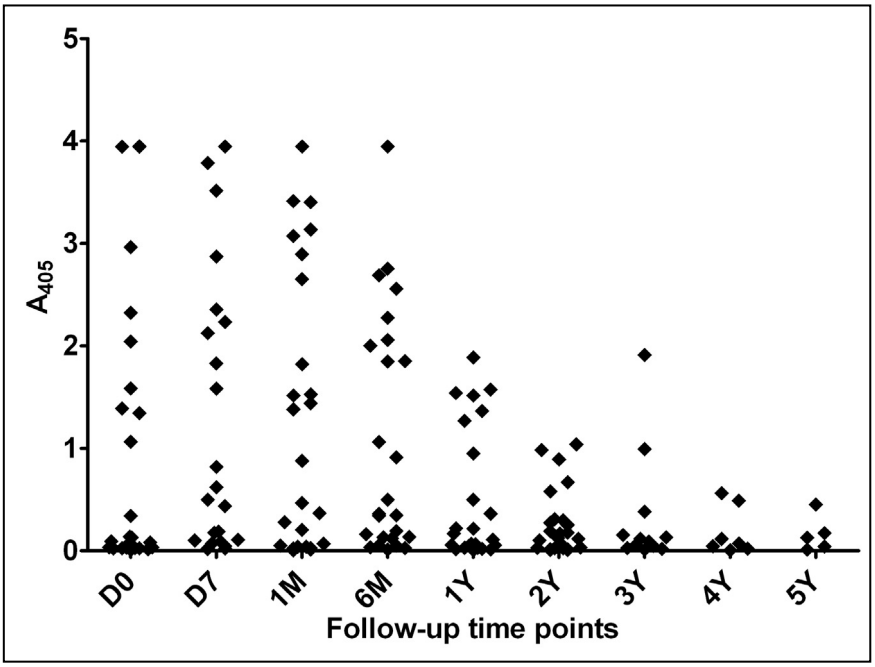

Figure 1:
Time course of antibody levels in CE patients (n=31) tested by EmLD89-ELISA.

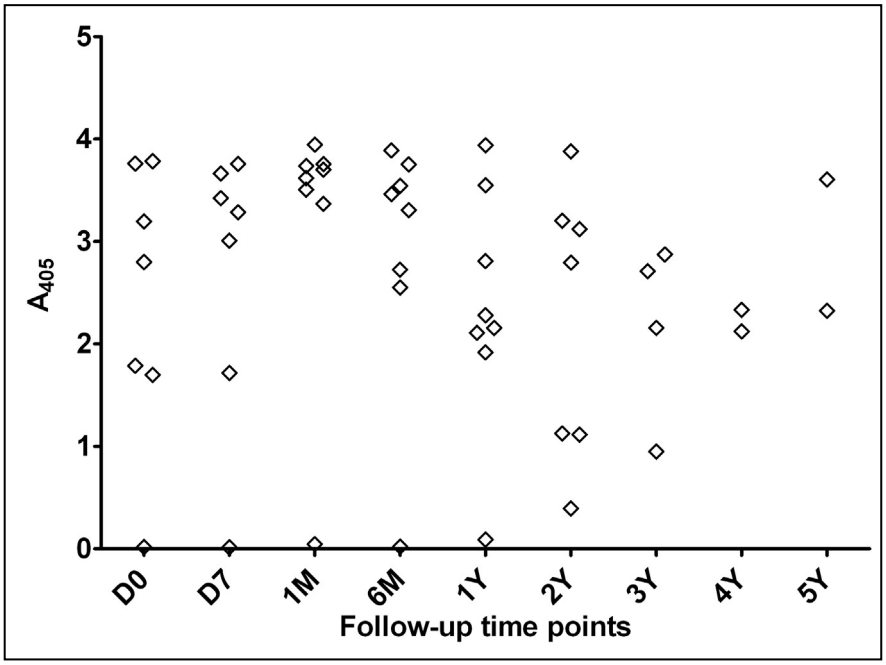

Figure 2:
Time course of antibody levels in NCCE (n=9) patients tested by EmLD89-ELISA.

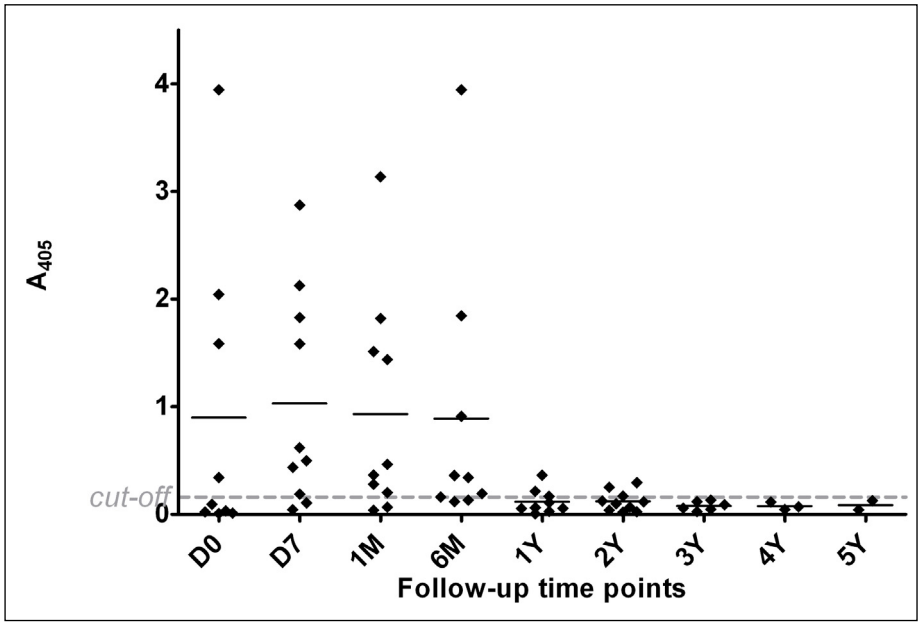

Figure 3:
Time course of antibody levels in CCE patients (n=12) reaching seronegativity within 3 years of follow-up. Bars represent mean A405 values.

Tables

Table 1:
Overview of the clinical features of the two groups of young cystic echinococcosis[a)] patients included in the present study. The sera used were identical to those employed in previous studies [9,12,17].

	Sex			Mean age (range) in years	Cyst location and type[c)]						Therapy and follow up[e)]				
	n[b)]	F	M		Lung		Liver		Multiple[d)]		Therapy period (months)	Surgery only	ABZ only	Surgery plus ABZ	Follow-up period (mean ± SD)
CCE	31	13	18	9 (3-14)	15	CE1(15)	13	CE1 (8) CE2 (2) CE3 (3)	2	CE1 (2) CE4 (1)	1 - 24	25	0	6	3±2 years
NCCE	9	4	5	9 (4-14)	0		6	CE1 (3) CE2 (1) CE3 (2)	3	CE1 (1) CE3 (2)	12 - 60	2	1	6	3±2 years

a) Cured cystic echinococcosis (CCE) and non-cured cystic echinococcosis (NCCE) patients.
b) Total number of patients.
c) Cyst type (CE1 to CE4; no CE5) corresponds to the sonographic classification carried out at initial diagnosis, according to WHO guidelines.
d) Multiple locations indicate that cysts were detected in the liver and peritoneum (2 cases), liver and lung (1 case), liver and spleen (2 cases).
e) For CCE patients, chemotherapy (ABZ; albendazole) was added to surgery when cysts were considered as fissured. In NCCE patients, ABZ was administered in cases exhibiting fissured cysts, multiple cysts and disseminated or inoperable cysts left after surgery. Cysts were considered to be fissured when the laminated layer and consequently germinal layer presented visible damages that no longer maintained the original healthy status of hydatid cysts. If respective cysts were punctured, the hydatid fluid appeared milky or colored, and high protein concentrations (i.e. ≥1mg protein/ml) were resolved.

Table 2:
Time course of positivity for antigens EmLD89, recP29 and EgHF by ELISA and/or Western Blot in 31 CCE and 9 NCCE patients.

	D0	D7	1M	6M	1Y	2Y	≥3Y
CCE	**n[a] =27**	**n[a] =23**	**n[a] =25**	**n[a] =25**	**n[a] =24**	**n[a] =25**	**n[a] =24**
EmLD89 ELISA positive (%)	12 (44)	15 (65)	17 (68)	16 (64)	12 (50)	12 (48)	7 (29)
EgHF ELISA positive (%)	20 (74)	18 (82) [b]	22 (92) [b]	22 (88)	19 (83) [b]	19 (76)	17 (71)
EgHF WB positive (%)	17 (63)	14 (61)	18 (72)	14 (56)	8 (33)	7 (28)	4 (17)
recP29 ELISA positive (%)	6 (22)	11 (48)	15 (60)	6 (24)	3 (13)	2 (8)	0 (0)
recP29 WB positive (%)	11 (41)	15 (65)	18 (72)	12 (48)	5 (21)	3 (13) [b]	1 (4)
NCCE	**n[a] =7**	**n[a] =7**	**n[a] =8**	**n[a] =8**	**n[a] =8**	**n[a] =7**	**n[a] =8**
EmLD89 ELISA positive (%)	6 (86)	6 (86)	7 (88)	7 (88)	7 (88)	7 (100)	8 (100)
EgHF ELISA positive (%)	6 (86)	6 (100) [b]	8 (100)	8 (100)	8 (100)	7 (100)	8 (100)
EgHF WB positive (%)	6 (86)	6 (86)	7 (88)	7 (88)	7 (88)	6 (86)	7 (88)
recP29 ELISA positive (%)	1 (14)	4 (57)	4 (50)	4 (50)	5 (63)	4 (57)	5 (63)
recP29 WB positive (%)	4 (57)	6 (86)	8 (100)	8 (100)	8 (100)	7 (100)	6 (75)

[a] Number of serum samples collected varied between patients; numbers (n) indicate the total amount of serum samples available from 31 CCE and 9 NCCE patients.
[b] One sample not available for testing

General discussion

This PhD thesis provides proof-of-principle for the applicability of synthetic peptides (20-50mer) as antigens in the diagnosis of human echinococcosis infection and it depicts three different approaches to tackle identification of the immunome of *Echinococcus multilocularis* and *E. granulosus* on the level of 25-45mer peptides.

Throughout the entire work, the peptide selection criteria remained the same. By means of bioinformatics, we selected structural motifs which can adopt native folding even as fragments of 25-45 amino acids in length and isolated from the context of the whole protein. Such motifs were on the one hand intrinsically unstructured (IUR) regions of proteins, potentially representing linear epitopes, and on the other hand alpha-helical coiled-coils, potentially displaying conformation-depending epitopes [1]. To further confine the regions for peptide design within the stretches of predicted CC, we used StableCoil [2] to find those CC sections with the highest relative stability. Sequences containing predicted IUR were further analyzed with BepiPred [3] for the prediction of linear B-cell epitopes. Peptides were finally selected from regions of highly stable CC or IURs with predicted B-cell epitopes.

Looking at the secondary structure predictions of the 6 peptide candidates identified and characterized in the course of this PhD thesis, 4 peptides were based on alpha-helices and only 2 peptides corresponded to predicted B-cell epitopes. The yield of B-cell epitope prediction might be considered poor, given the fact that about equal amounts of CC and IUR peptides have been tested. Our results support the general opinion that the current state of B-cell epitope prediction is far from ideal and that more research in this area is needed [4]. However, with more and more three dimensional structures of proteins being determined, attention is shifting from the prediction of linear B-cell epitopes to the prediction of discontinuous B-cell epitopes [5].

In contrast to given peptide selection criteria, we tested three fundamentally different approaches for the selection of parent proteins for peptide design.

The genomics approach (chapter 2) based on a genome-wide search for proteins potentially located to the host-parasite-interface, because such proteins are thought to potentially elicit an immune response. Bioinformatic algorithms were used to identify signal peptides (PSORT II [6]), transmembrane domains (TMHMM [7]) and GPI-anchor sites (GPI-som [8]). At the time of our bioinformatics search (done on October 16, 2007), NCBI Entrez Protein Database contained 240 entries of *E. multilocularis* and 940 entries of *E. granulosus* protein sequences, including redundancy. On the one hand, these 1'180 proteins were used as input sequences for prediction of localization and on the other hand for predictions of CC (Paircoil2 [9]) and IUR (IUPred [10]). The overlap of CC and/or IUR-containing sequences with predictions of a potentially extracellular or surface location was very small. Only 5 *E. granulosus* and 8 *E. multilocularis* sequences were identified (without redundancy).

Among the proteins predicted to contain stretches of CC and/or IUR only, a considerable number of known antigens were found, e.g. antigen II/3 [11], EM13 [12], P29 [13], EG19 [14]. We expanded our selection criteria and included all sequences containing CC and/or IUR, irrespectively of their potential localization. The set of parent proteins selected by our genomics approach therefore comprised protein sequences mainly from known antigens. Peptides selected from these parent proteins were screened for reactivity with patients' sera on microarray.

Screening on peptide microarrays is a desirable alternative to screening in ELISA, because considerably smaller amounts of patients' sera are needed. To screen 96 different peptides on microarray, 2µl of sera are sufficient, while the same screening in our ELISA system would necessitate 50µl of serum. We found that the availability of well-defined, parasitologically confirmed human serum samples for test development and validation was the most

limiting factor. Thus, a technique that preserves limited patient material offers a substantial advantage.

After screening on microarray reactive peptides were transferred into ELISA format, because ELISA has been routinely used in sero-diagnostics. Our method of peptide immobilization via biotin/streptavidin proved ideal as it was versatile to accommodate both formats. Peptides were attached to either solid phase by binding the N-terminal biotin of the peptides to streptavidin coated microscope glass slides or ELISA plates.

At the beginning of this project, we also addressed the question of optimal peptide length. Selected peptides found reactive with echinococcosis sera on microarray were synthesized again as longer versions of the same antigen domain (AltaBioscience, UK). Both, the “short” 27-30mers identified by microarray and their corresponding “long” 40-47mer counterparts were tested in ELISA. This parallel testing revealed that in general, the “longer” peptides performed better than the “shorter” ones.

This finding argued for a modified approach to peptide design. Longer peptides would be generally desirable, but this was not compatible with low cost peptide synthesis. In a genome wide screen synthesis costs represent a major investment and must be kept low. The Episcan peptide synthesis technology offered by AltaBioscience (University of Birmingham, UK) enables simultaneous synthesis of up to 192 different peptides in 1 to 2 mg quantities, but limits peptide length to 30 amino acids. However, the advantages of screening a batch of several different peptides at the same time simultaneously on microarray with as little as 2ul of patient serum outweighed the limitation of peptide length.

The peptide candidate performing best in our genomics approach was peptide longD8-9. It was developed on the basis of two overlapping peptides both of which were found to be reactive on microarray. The extended version peptide reached a sensitivity of 74% for the detection of CE infection and 43% for the

detection of AE infection in adult patients. 74% sensitivity reached by a single peptide is a considerable high value. The sensitivities of the second- and third-best candidates were much lower. Peptide longD1 detected 6% of CE and 25% of AE cases, peptide longD12 9% of CE and 23% of AE cases. But all three peptides reached a specificity ranging from 91 to 94% (chapter 2).

Due to the partial complementarity of their reactivity patterns, we decided to further characterize combinations of peptides that were applied as synthetic antigen pools. We found that by pooling several peptides, the sensitivity was indeed increased, but on the cost of specificity. Single peptides thus exhibited decreased sensitivity compared to native antigen but always obtained a high specificity of at least 90%, whereas a peptide pool reached higher sensitivity but decreased specificity compared to its individually tested components.

The possibility of increasing test sensitivity by combining different peptides targets the application of synthetic peptides rather to a screening test which requires maximal sensitivity than to a confirmatory test with utmost specificity.

Our second strategy applied to screen for antigenic proteins was a proteomics approach, which used tandem mass-spectrometry to identify proteins from *in vitro* cultivated *E. multilocularis* vesicle fluid (EmVF, [15]) (chapter 3). We focused on a 20-22kDa fraction that has been found to be diagnostically informative in immunoblots with patient's serum.

A major issue of the proteomics approach was the protein diversity in the 20-22kDa fraction. The standard search of peptide spectra against the Uniprot database identified 11 different proteins. 23 more proteins were identified with the help of conceptually translated EST libraries (Wellcome Trust Sanger Institute). To further down-select parent sequences for peptide design, we concentrated on those proteins highly abundant in the sample and hence represented by a large number of tryptic fragments. We designed a number of peptides corresponding to different EST sequences with similarity to 14-3-3. The evaluation of these peptides showed very little reactivity with sera from

patients. Only one 14-3-3-like peptide out of 15 was tested in more detail but finally had to be dismissed because of too little sensitivity with echinococcosis sera. None of the 14-3-3 peptides was considered for diagnostic use.

The low serological reactivity of *Echinococcus* 14-3-3 peptides might reflect the importance of this protein in many basic cellular events crucial for multiplication and survival of eukaryotic cells (reviewed e.g. by Siles-Lucas [16]). Antibodies directed against 14-3-3 might heavily interfere with parasite growth and therefore 14-3-3 proteins in parasites evolved to be little antigenic. In fact, the *Echinococcus* 14-3-3 protein produced promising results as vaccine candidate in a mouse model [17].

The two peptides performing best in our proteomics approach to antigen discovery were peptide C2 (from cytosolic malate dehydrogenase) and peptide B6 (from phosphoglycerate mutase). Peptide C2 reached a sensitivity of 15% and a specificity of 90%, while the sensitivity of peptide B6 was 18%, and the specificity 91%. These two peptides showed some degree of complementarity in their recognition of echinococcosis sera. Therefore C2 and B6 were considered candidates for inclusion in a synthetic antigen pool.

Our third strategy for the selection of parent proteins made use of conceptually translated EST sequences of *E. granulosus* and *E. multilocularis* and was therefore considered a transcriptomics approach (chapter 4). At the Wellcome Trust Sanger Institute, the sequences of 2 *E. multilocularis* and 5 *E. granulosus* EST libraries are available for download. The translation of the EST reads represented a challenge, because these libraries often constitute a collection of “noisy” data. Start/stop codons are often missing, frameshifts occur frequently. We decided to combine the outputs of two translation algorithms. On the one hand, we used a conservative Perl script translating each of the six possible frames followed by selection of the longest one as definitive open reading frame and on the other hand, a more advanced algorithm able to deal with incomplete sequences and frameshifts (FrameDP [18]).

This decision to use both outputs was supported by the proteomics approach. Certain peptide spectra measured by MS/MS could be matched either to translations made by FrameDP or by the Perl script. For example, the FrameDP algorithm tended to shorten the N-terminal region of translations to make methionine the first amino acid. In our MS/MS analysis, we found a peptide spectrum that matched to the protein region located upstream to a methionine that was set as start codon by FrameDP. The identification of this spectrum would have been lost using the FrameDP translation only, but it could be "rescued" by the Perl translation. On the other hand, FrameDP corrected properly for frameshifts leading to translations with a longer open reading frame than in the corresponding Perl translation which was truncated C-terminally due to the stop codon produced by the frameshift. We also found an example of a peptide spectrum that matched to a C-terminal region and could only be identified by the FrameDP translation.

As mentioned above, EST libraries often consist of incomplete sequences. Predictions of subcellular localization and GPI-anchor sites as carried out in the genomics approach were not possible, since these algorithms demand complete N- and C-termini. Instead, the main focus was on the prediction of stable alpha-helical coiled-coils. The CC motif was preferred over the IUR motif because it is generally accepted that the majority of epitopes are conformation-dependent epitopes [1].

The transcriptomics approach identified one promising peptide, namely peptide MA2_H3. This candidate was the first of all candidates tested that reacted exclusively with sera from AE patients. There was no cross-reactivity seen with sera from CE patients. Peptide MA2_H3 reached a sensitivity of 29% for the detection of AE infection and a specificity of 96%.

In summary, the three different selection approaches identified 6 peptides that were validated with the entire serum panel (AE, CE, other helminthic and *Entamoeba* infections and blood donor samples). The decreasing amount of peptides identified by the three different approaches might reflect the

decreasing amount of proven antigens contained in the sets of parent proteins selected for peptide design. The most promising peptide candidate longD8-9 emerged from a set of parent proteins based mainly on known antigens.

Despite CC and IURs being wide-spread in genomes of eukaryotes, not every protein sequence does contain either one or both of these domains. For the genomics and the proteomics approach, these structural selection criteria appeared to be too restrictive. For sequences with low probability predictions of CC and IUR domains, a different approach might be better suited, for example Pepscan epitope mapping (reviewed e.g. by Meloen et al. [19]). This method aims at the identification of linear B-cell epitopes by investigating highly overlapping peptides covering the entire protein sequence.

In chapter 5, the application of EmLD89, the most promising peptide identified in this PhD work, in a serological follow-up of young cystic echinococcosis patients is described. (For the purpose of the follow-up study, the peptide formerly termed longD8-9 was renamed to EmLD89.) This represents a novelty, because to date no synthetic peptide has been tested as marker for serological follow-up of CE patients.

The results obtained in EmLD89-ELISA showed a decrease of antibody titers comparable to the one seen in EgHF-ELISA. Titers remained high in non-cured CE patients with both antigens. Successful treatment can thus be monitored by decreasing antibody titers. Within 3 years after treatment, more cured CE patients reached seronegativity when tested in peptide ELISA compared to native EgHF-ELISA. This finding strongly recommends the further investigation of more synthetic peptides for application in serological follow-up.

Conclusions drawn from this study were i) microarray technology was well-suited for the screening of multiple peptides minimizing the use of patient serum, ii) longer peptides (40-50 amino acids) performed better than shorter ones (20-30 amino acids), iii) one single peptide was not sufficiently sensitive,

but sensitivity could be increased by combining several peptides and iv) peptide antigens also proved valuable tools in serological follow-up of CE patients but their application for this purpose needs further research.

References

1. Corradin G, Villard V, Kajava AV: **Protein structure based strategies for antigen discovery and vaccine development against malaria and other pathogens.** *Endocrine Metabolic & Immune Disorders-Drug Targets* 2007, **7:**259-265.

2. Tripet B, Hodges RS: **StableCoil: an algorithm designed to predict the location and relative stability of coiled-coils in native protein sequences.** In *Peptides: The Wave of the Future*; Edited by Lebl M, Houghten RA 365-366.

3. Larsen J, Lund O, Nielsen M: **Improved method for predicting linear B-cell epitopes.** *Immunome Research* 2006, **2:**2.

4. Greenbaum JA, Andersen PH, Blythe M, Bui HH, Cachau RE, Crowe J, Davies M, Kolaskar AS, Lund O, Morrison S et al.: **Towards a consensus on datasets and evaluation metrics for developing B-cell epitope prediction tools.** *Journal of Molecular Recognition* 2007, **20:**75-82.

5. Yang XD, Yu XL: **An introduction to epitope prediction methods and software.** *Reviews in Medical Virology* 2009, **19:**77-96.

6. Nakai K, Horton P: **PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization.** *Trends in Biochemical Sciences* 1999, **24:**34-35.

7. Krogh A, Larsson Br, von Heijne G, Sonnhammer ELL: **Predicting transmembrane protein topology with a hidden markov model: application to complete genomes.** *Journal of Molecular Biology* 2001, **305:**567-580.

8. Fankhauser N, Maser P: **Identification of GPI anchor attachment signals by a Kohonen self-organizing map.** *Bioinformatics* 2005, **21:**1846-1852.

9. McDonnell AV, Jiang T, Keating AE, Berger B: **Paircoil2: improved prediction of coiled coils from sequence.** *Bioinformatics* 2006, **22:**356-358.

10. Dosztanyi Z, Csizmok V, Tompa P, Simon I: **IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content.** *Bioinformatics* 2005, **21:**3433-3434.

11. Vogel M, Gottstein B, Müller N, Seebeck T: **Production of a recombinant antigen of *Echinococcus multilocularis* with high immunodiagnostic sensitivity and specificity.** *Molecular and Biochemical Parasitology* 1988, **31:**117-125.

12. Frosch PM, Geier C, Kaup FJ, Muller A, Frosch M: **Molecular cloning of an echinococcal microtrichal antigen immunoreactive in *Echinococcus multilocularis* disease.** *Molecular and Biochemical Parasitology* 1993, **58:**301-310.

13. González G, Spinelli P, Lorenzo C, Hellman U, Nieto A, Willis A, Salinas G: **Molecular characterization of P-29, a metacestode-specific component of *Echinococcus granulosus* which is immunologically related to, but distinct from, antigen 5.** *Molecular and Biochemical Parasitology* 2000, **105:**177-185.

14. Delunardo F, Ortona E, Margutti P, Perdicchio M, Vacirca D, Teggi A, Sorice M, Siracusano A: **Identification of a novel 19kDa *Echinococcus granulosus* antigen.** *Acta Tropica* 2010, **113:**42-47.

15. Müller N, Frei E, Nuñez S, Gottstein B: **Improved serodiagnosis of alveolar echinococcosis of humans using an *in vitro* produced *Echinococcus multilocularis* antigen.** *Parasitology* 2007, **134:**879-888.

16. Siles-Lucas M, Gottstein B: **The 14-3-3 protein: a key molecule in parasites as in other organisms.** *Trends in Parasitology* 2003, **19:**575-581.

17. Siles-Lucas M, Merli M, Mackenstedt U, Gottstein B: **The *Echinococcus multilocularis* 14-3-3 protein protects mice against primary but not secondary alveolar echinococcosis.** *Vaccine* 2003, **21:**431-439.

18. Gouzy J, Carrere S, Schiex T: **FrameDP: sensitive peptide detection on noisy matured sequences.** *Bioinformatics* 2009, **25:**670-671.

19. Meloen RH, Langedijk JPM, Langeveld JPM: **Synthetic peptides for diagnostic use.** *Veterinary Quarterly* 1997, **19:**122-126.

CURRICULUM VITAE

PERSONAL DETAILS

Name	Claudia List
Birthday	31.01.1978
Nationality	Swiss
Private address	Rosengässchen 3 CH-4310 Rheinfelden
Phone	079 574 57 31
Email	claudia.list@unibas.ch

EDUCATION AND WORK EXPERIENCE

Jan 2007 – Oct 2010	PhD at the Swiss TPH, Basel, supervised by Dr. Ingrid Felger. Applied methods: Bioinformatics, peptide microarray, ELISA, Western Blot.
2009, Nov 23-27	SiB Bioinformatics Workshop: Computational methods for proteomics data analysis.
2008, Sept 8-12	SiB Bioinformatics Workshop: Perl programming for biomedical researchers.
2007, Aug 27-31	7th BioSapiens European School in Bioinformatics.
2007, Apr 26	Swiss TPH course “diagnosis of helminthic infections in medical parasitology”.
Jul. 2005 – Dec. 2006	Scientific associate of the University of Basel for research of fungicide resistance at Syngenta Crop Protection AG.
Jun. 2004 – Jul. 2005	Internship at the bacterial screening department of Novozymes A/S in Bagsværd, Denmark. Applied methods: Cultivation of anaerobic bacteria, DNA extraction, construction and screening of genomic DNA libraries, PCR, thin layer chromatography, SDS-PAGE, enzyme activity assays.
Jan. 2003 – Apr. 2004	Diploma thesis at the University of Basel, supervised by Prof. Volker Schmid. Applied methods: RNA and DNA extraction, cDNA synthesis, homology PCR, RACE, cloning, sequencing, BLAST search, expression library screening, *in situ* hybridisation, immunohisto-chemistry, light and UV microscopy, maintenance of marine hydroid cultures.
1998 – 2002	Study of biology at the University of Basel.
1994 – 1997	Grammar school at Liestal.

PUBLICATION

Claudia List, Weihong Qi, Eva Maag Bruno Gottstein, Norbert Müller and Ingrid Felger; Serodiagnosis of *Echinococcus* spp. Infection: Explorative Selection of Diagnostic Antigens by Peptide Microarray, PLoS Negl Trop Dis. 2010 August; 4(8): e771.

POSTER

Claudia List, Elisabeth Frei, Weihong Qi, Norbert Müller, Ingrid Felger; *Selection of synthetic peptides for serologic diagnosis of helminth infections*. Joint meeting of the French Society for Parasitology, the German Society for Parasitology and the Swiss Society for Tropical Medicine and Parasitology, 2007, Strasbourg, France.

Claudia List, Weihong Qi, Norbert Müller, Ingrid Felger; *Selection of long synthetic peptides for serologic diagnosis of helminth infections*; Union of the Swiss Society for Experimental Biology (USGEB), 2007, Basel, Switzerland.

ORAL PRESENTATION

Claudia List, Ingrid Felger; *Selection and application of long synthetic peptides for serodiagnosis of helminth infections*; Meeting of the Swiss Society for Tropical Medicine and Parasitology, 2008; Jongny, Switzerland.

OTHER KNOWLEDGE

Languages	Fluent in German and English, moderate in French, basic knowledge in Danish.
Computer skills	Good knowledge of Microsoft Office, Adobe Photoshop; basic knowledge of Perl.

REFERENCES

Ingrid Felger	PhD thesis supervisor Swiss TPH, Basel; Switzerland Phone: +41 61 284 81 17 Email: ingrid.felger@unibas.ch
Norbert Müller	PhD thesis co-supervisor Institute of Parasitology, University of Bern, Switzerland Phone: +41 31 631 24 74 Email: norbert.mueller@ipa.unibe.ch