

# Journal of Integrated

# OMICS

a methodological journal

## Editors-in-Chief

Carlos Lodeiro-Espiño

Florentino Fdez-Riverola

Jens Coorssen

Jose-Luís Capelo-Martínez

# JIOMICS

## Journal of Integrated OMICS

---

### Focus and Scope

Journal of Integrated OMICS, JIOMICS, provides a forum for the publication of original research papers, preliminary communications, technical notes and critical reviews in all branches of pure and applied "-omics", such as genomics, proteomics, lipidomics, metabolomics or metallomics. The manuscripts must address methodological development. Contributions are evaluated based on established guidelines, including the fundamental nature of the study, scientific novelty, and substantial improvement or advantage over existing technology or method. Original research papers on fundamental studies, and novel sensor and instrumentation development, are especially encouraged. It is expected that improvements will also be demonstrated within the context of (or with regard to) a specific biological question; ability to promote the analysis of molecular mechanisms is of particular interest. Novel or improved applications in areas such as clinical, medicinal and biological chemistry, environmental analysis, pharmacology and materials science and engineering are welcome.

### Editors-in-Chief

**Carlos Lodeiro-Espiño**, University of Vigo, Spain

**Florentino Fdez-Riverola**, University of Vigo, Spain

**Jens R. Coorsen**, University of Western Sydney, NSW, Australia

**Jose-Luís Capelo-Martínez**, University of Vigo, Spain

### Regional editors

#### ASIA

---

**Gary Xiao**

Director of Functional Genomics and Proteomics Laboratories at Osteoporosis Research Center, Creighton University Omaha, Nebraska, USA

**Yogeshwer Shukla**

Proteomics laboratory at Indian Institute of Toxicology Research (Council of Scientific and Industrial Research), Lucknow, India

#### AUSTRALIA AND NEW ZEALAND

---

**Jens R. Coorsen**

University of Western Sydney, NSW, Australia

#### Europe

---

**Gilberto Igrejas**

University of Trás-os-Montes and Alto Douro, Life Sciences and Environmental School, Institute for Biotechnology and Bioengineering, Centre of Genetics and Biotechnology  
Department of Genetics and Biotechnology, 5001-801 Vila Real, Portugal

UFZ, Helmholtz-Centre for Environmental Research, Department of Proteomics, Permoserstr. 15, 04318 Leipzig, Germany

**Jan Ottervald**

Research and Development | Innovative Medicines Neuroscience, CNSP iMed Science Södertälje, AstraZeneca, Sweden

**Martin von Bergen**

#### North America

---

**Randen Patterson**

Center for Computational Proteomics, The Pennsylvania State University,

USA

**Oscar Alzate**

Associate Professor of Cell and Developmental Biology, Adjunct Associate Professor in Neurology, Director: Systems Proteomics Center

School of Medicine, The University of North Carolina at Chapel Hill, USA

**Yue Ge**

US Environmental Protection Agency, Research Triangle Park, USA

**South America**

---

**Eduardo Alves de Almeida**

Depto. de Química e Ciências Ambientais, IBILCE - UNESP, Brazil

University of Campinas - Unicamp

**Carlos H. I. Ramos**

Chemistry Institute – UNICAMP, Brazil

**Associated editors**

---

**ASIA**

---

**Amita Pal**

Division of Plant Biology, Bose Institute, Kolkata, INDIA

**Cheolju Lee**

Korea Institute of Science and Technology, Seoul, Korea

**Chii-Shiarng Chen**

National Museum of Marine Biology and Aquarium, 2 Houwan Road, Checheng, Pingtung, 944, Taiwan

**Chantragan Srisomsap**

Chulabhorn Research Institute, Bangkok, Thailand

**Chen Han-Min**

Department of Life Science, Catholic Fu-Jen University, Taipei, Taiwan

**Debmalya Barh**

Institute of Integrative Omics and Applied Biotechnology (IIOAB), India

**Dwaipayan Bharadwaj**

Genomics & Molecular Medicine Unit, Institute of Genomics & Integrative Biology (CSIR), Mall Road, Delhi, India

**Eiji Kinoshita**

Department of Functional Molecular Science, Graduate School of Biomedical Sciences, Hiroshima University, Japan

**Fan Chen**

Institute of Genetics and Developmental Biology, Chinese Academy of Sciences (CAS), China

**Feng Ge**

Institute of Hydrobiology, Chinese Academy of Sciences, China

**Ganesh Chandra Sahoo**

BioMedical Informatics Center of Rajendra Memorial Research Institute of Medical Science (RMRIMS), Patna, India

**Guangchuang Yu**

Institute of Life & Health Engineering, Jinan University, Guangzhou, China

**Hai-Lei Zheng**

School of Life Sciences, Xiamen University, China

**Hee-bal Kim**

Department of Food and Animal Biotechnology of the Seoul National University, Korea

**Hitoshi Iwahashi**

Health Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Japan

**Hong-Lin Chan**

National Tsing-Hua University, Taiwan

**HuaXu**

Research Resources Center, University of Illinois, Chicago

**Hui-Fen Wu**

Department of Chemistry, National Sun Yat – Sen University, 70, Lien-Hai

Road, 80424, Kaohsiung, Taiwan

**Hye-Sook Kim**

Faculty of Pharmaceutical Sciences, Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University, Japan

**Ibrokhim Abdurakhmonov**

Institute of Genetics and Plant experimental Biology Academy of Sciences of Uzbekistan, Uzbekistan

**Jong Won Yun**

Dept. of Biotechnology, Kyungsan, Kyungbuk 712-714, Republic of Korea

**Young-Gyu Ko**

College of Life Sciences and Biotechnology, Korea University, Korea

**Kazuaki Kakehi**

School of Pharmacy, Kinki University, Kowakae 3-4-1, Higashi-Osaka, 577-8502, Japan

**Kazuki Sasaki**

Department of Molecular Pharmacology, National Cerebral and Cardiovascular Center, Japan

**Kohji Nagano**

Chugai Pharmaceutical Co. Ltd., Japan

**Li Jianke**

Institute of Apicultural Research, Chinese Academy of Agricultural Science, Beijing, China

**Ling Zheng**

College of Life Sciences, Wuhan University, China

**Luk John Moonching**

National University of Singapore, Singapore

**Manjunatha Kini**

Department of Biological Sciences, National University of Singapore, Singapore

**Masaya Miyazaki**

National Institute of Advanced Industrial Science and Technology, 807-1 Shuku, Tosu, Saga 841-0052, Japan

**Ming-Fa Hsieh**

Department of Biomedical Engineering, Chung Yuan Christian University, Taiwan

**Suresh Kumar**

Department of Applied Chemistry, S. V. National Institute of Technology, Gujarat, INDIA

**Songping Liang**

Hunan Normal University, Changsha City, China

**Moganty Rajeswari**

Department of Biochemistry, All India Institute of Medical Sciences, Ansari Nagar, New Delhi, India

**Nam Hoon Cho**

Dept. of Pathology, Yonsei University College of Medicine, Korea

**Poh-Kuan CHONG (Shirly)**

National University of Singapore, Singapore

**Qian Shi**

Institutes of Biomedical Sciences, Fudan University, Shanghai, China

**Roger Beuerman**

Singapore Eye Research Institute, Singapore

**Sanjay Gupta**

Advanced Centre for Treatment, Research and Education in Cancer (ACTREC), Tata Memorial Centre, Kharghar, Navi Mumbai, India

**Sanjeeva Srivastava**

Indian Institute of Technology (IIT) Bombay, India

**Setsuko Komatsu**

National Institute of Crop Science, Japan

**Sameh Magdeldin Mohamed**

Niiigata prefecture, Nishi-ku, Terao, Niigata, Japan

**Sen-Lin Tang**

Biodiversity Research Center, Academia Sinica, Taipei, Taiwan

**Shaojun Dai**

Alkali Soil Natural Environmental Science Center, Key Laboratory of Saline-alkali Vegetation Ecology Restoration in Oil Field, Ministry of Education, Northeast Forestry University, P.R. China

**Shipin Tian**

Institute of Botany, Chinese Academy of Sciences, China

**Tadashi Kondo**

National Cancer Center Research Institute, Japan

**Taesung Park**

National Research Laboratory of Bioinformatics and Biostatistics at the Department of Statistics Seoul National University, Korea

**Toshihide Nishimura**

Guest Professor, Department of Surgery I, Tokyo Medical University, Tokyo, Japan, CSO, Biosys Technologies Inc., Tokyo, Japan, Corporate Officer, Medical ProteoScope Co. Ltd., Tokyo, Japan

**Wei-dong Zhang**

Lab of Natural Products, School of Pharmacy, Second Military Medical University, Shanghai, China

**Wenxiong Lin**

School of Life Sciences, Fujian Agriculture and Forestry University, China

**William Chen Wei Ning**

School of Chemical and Biomolecular Engineering Nanyang Technological University, Singapore

**Xiao LiWang**

Division of Cardiovascular Diseases, Mayo Clinic, Rochester, MN

**Xiao Zhiqiang**

Key Laboratory of Cancer Proteomics of Chinese Ministry of Health, Xiangya Hospital, Central South University, 87 Xiangya Road, Changsha, Hunan 410008, P.R. China

**Xiaoping Wang**

Key Laboratory of Molecular Biology & Pathology, State Bureau of Chinese Medicine, China

**Xuanxian Peng**

School of Life Sciences, Sun Yat-sen University, Guangzhou, China

**YasminAhmad**

Peptide and Proteomics Division Defence Institute of Physiological and Allied Research (DIPAS), DRDO, Ministry of Defence, Timarpur, Delhi-54, India

**Yin Li**

Institute of Microbiology, Chinese Academy of Sciences, Beijing, China

**Yoon-Pin LIM**

Department of Biochemistry, National University of Singapore, Singapore

**Yong Song Gho**

Department of Life Science, POSTECH, Pohang, Korea

**Youngsoo Kim**

Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, Republic of Korea

**Yu Xue**

Department of Systems Biology, College of Life Science and Technology Huazhong University of Science and Technology Wuhan, China

**Yulan Wang**

State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, Wuhan Centre for Magnetic Resonance, Wuhan Institute of Physics and Mathematics, The Chinese Academy of Sciences, China

#### AUSTRALIA AND NEW ZEALAND

---

**Bruno Catimel**

Epithelial laboratory, Ludwig Institute for Cancer Research, Post Office Royal Melbourne Hospital, Australia

**JoëlleCoumans-Moens**

School of Science and Technology, School of Medicine, University of New England, Australia

**Marc Wilkins**

University of New South Wales, Sydney, Australia

**Michelle Hill**

University of Queensland, Australia

**Michelle Colgrave**

CSIRO Livestock Industries, St Lucia, Australia

**Nicolas Taylor**

ARC Centre of Excellence in Plant Energy Biology & Centre for Comparative Analysis of Biomolecular Networks (CABiN), University of Western Australia, Perth, Australia

**Stefan Clerens**

Protein Quality &Function, AgResearch Ltd Christchurch, New Zealand

**Peter Solomon**

Research School of Biology College of Medicine, Biology and Environment, The Australian National University, Australia

**Phoebe Chen**

Department of Computer Science and Computer Engineering, La Trobe University, Melbourne, Australia

**Richard Christopherson**

School of Molecular Bioscience, University of Sydney, Australia

**Sham Nair**

Department of Biological Sciences, Faculty of Science, Macquarie University, NSW, Australia

**Valerie Wasinger**

Bioanalytical Mass Spectrometry Facility, Mark Wainwright Analytical Centre, The University of NSW, Australia

**Wujun Ma**

Centre for Comparative Genomics, Murdoch University, Australia

#### EUROPE

---

**AhmetKoc, PhD**

Izmir Institute of Technology, Department of Molecular Biology & Genetics,

Urla, Izmir, Turkey

**Alexander Scherl**

Proteomics Core Facility, Faculty of Medicine, University of Geneva, Geneva, Switzerland

**Alfonsina D'Amato**  
Politecnico di Milano, Department of Chemistry, Materials and Chemical Engineering "GiulioNatta", Italy

**Andrei Turtoi**  
University of Liege, Metastasis Research Laboratory, GIGA-Cancer Bât. B23, Belgium

**VittoriaMatafora**  
Biological Mass Spectrometry Unit, San Raffaele Scientific Institute, Milan, Italy

**Alfred Vertegaal**  
Molecular Cell Biology, Leiden University Medical Center, PO Box 9600-Mailbox S1-P, 2300 RC Leiden, Netherlands

**Almudena Fernández Briera**  
Dpt. Biochemistry Genetics and Immunology, Faculty of Biology –University of Vigo, Spain

**Ana Maria Rodríguez-Piñero**  
Institute of Biomedicine, University of Gothenburg, Sweden

**Ana Varela Coelho**  
Instituto de Tecnologia Química e Biológica (ITQB) Universidade Nova de Lisboa (UNL), Portugal

**Anna Maria Timperio**  
Dipartimento Scienze Ambientali Università della Tuscia Viterbo, Italy

**André Nogueira Da Costa**  
Molecular Carcinogenesis Group, Section of Mechanisms of Carcinogenesis International Agency for Research on Cancer - World Health Organization (IARC-WHO), Lyon, France

**Andrea Scaloni**  
Proteomics and Mass Spectrometry Laboratory, ISPAAM, National Research Council, via Argine 1085, 80147 Napoli, Italy

**Andreas Tholey**  
Division for Systematic Proteome Research, Institute for Experimental Medicine, Christian-Albrechts-University, 24105 Kiel, Germany

**Angel Manteca**  
Departamento de Biología Funcional and IUBA, Facultad de Medicina, Universidad de Oviedo, 33006 Oviedo, Spain

**Angel P. Diz**  
Department of Biochemistry, Genetics and Immunology, Faculty of Biology, University of Vigo, Spain

**Angela Bachi**  
Mass Spectrometry Unit DIBIT, San Raffaele Scientific Institute, Via Olgettina 58, I-20132 Milano, Italy

**Angela Chambery**  
Department of Life Science, Second University of Naples, Via Vivaldi 43, I-81100 Caserta, Italy

**António Sebastião Rodrigues**  
Departamento de Genética, Faculdade de Ciências Médicas, Universidade Nova de Lisboa, Portugal

**Arzu Umar**  
Department of Medical Oncology, Laboratory of Breast Cancer Genomics and Proteomics, Erasmus Medical Center Rotterdam Josephine Nefkens Institute, Rotterdam, The Netherlands

**Baggerman Geert**  
ProMeta, Interfaculty Center for Proteomics and Metabolomics, Leuven, Belgium

**Bart Devreese**  
Laboratory for Protein Biochemistry and Biomolecular Engineering, Department for Biochemistry and Microbiology, Ghent University, Belgium

**Bernard Corfe**  
Department of Oncology, University of Sheffield, Royal Hallamshire Hospital, Sheffield S10 2JF, United Kingdom

**Bernd Thiede**  
Biotechnology Centre of Oslo, University of Oslo, Blindern, 0317 Oslo, Norway

**Björn Meyer**  
Institut für Instrumentelle Analytik und Bioanalytik Hochschule Mannheim, Germany

**Cândido Pinto Ricardo**  
Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Av. da República-EAN, 2780-157 Oeiras, Portugal

**Carla Pinheiro**  
Plant Sciences Division, Instituto de Tecnologia Química e Biológica (ITQB), Universidade Nova de Lisboa, Portugal

**Andreas Boehm**  
Steigerfurtweg 8a, D-97084 Würzburg, Germany

**Carlos Gutiérrez Merino**  
Dept. Biochemistry and Molecular Biology University of Extremadura, Badajoz, Spain

**Celia Calado**  
N/A

**Celso Reis**  
Institute of Molecular Pathology and Immunology of the University of Porto, IPATIMUP, Portugal

**Celso Vladimiro Cunha**  
Medical Microbiology Department, Institute of Hygiene and Tropical Medicine, New University of Lisbon, Portugal

**Charles Steward**  
The Wellcome Trust Sanger Institute, Hinxton, United Kingdom

**Chris Goldring**  
Department of Pharmacology and Therapeutics, MRC Centre for Drug Safety Science, University of Liverpool, Ashton Street, Liverpool L69 3GE, United Kingdom

**Christiane Fæste**  
Section for Chemistry and Toxicology Norwegian Veterinary Institute, Oslo, Norway

**Christer Wingren**  
Department of Immunotechnology, Lund University, BMC D13, SE-221 84 Lund, Sweden

**Christophe Masselon**  
Laboratoire de Biologie a Grande Echelle (iRTSV/BGE), CEA Grenoble, France

**Cosima Damiana Calvano**  
Universita' degli Studi di Bari, Dipartimento di Chimica, Bari, Italy

**David Cairns**  
Section of Oncology and Clinical Research, Leeds Institute of Molecular Medicine, Leeds, UK

**Daniela Cecconi**  
Dip. di Biotecnologie, Laboratori di Proteomica e Spettrometri di Massa, Università di Verona, Verona, Italy

**David Honys**  
Laboratory of Pollen Biology, Institute of Experimental Botany ASCR, v. i., Rozvojová 263, 165 02 Prague 6, Czech Republic

**David Sheehan**  
Dept. Biochemistry, University College Cork (UCC), Ireland

**Deborah Penque**  
Departamento de Genética, Instituto Nacional de Saúde Dr Ricardo Jorge (INSA, I.P.), Lisboa, Portugal

**Ed Dudley**  
Institute of Mass Spectrometry, College of Medicine Swansea University, Singleton Park, Swansea, Wales, UK

**Elena Gonzalez**  
Complutense University of Madrid, Dept. Biochemistry and Molecular Biology IV, Veterinary Faculty

Madrid, Spain

**Elke Hammer**

Interfaculty Institute for Genetics and Functional Genomics, Ernst-Moritz-Arndt Universität, Friedrich-Ludwig-Jahn-Str. 15a, D-17487 Greifswald, Germany

**Edoardo Saccenti**

University of Amsterdam, Netherlands Metabolomics Centre, The Netherlands

**Eva Rodríguez Suárez**

Proteomics Core Facility - CIC bioGUNE, Parque tecnologico de Bizkaia, SPAIN

**Fernando J. Corrales**

Division of Hepatology and Gene Therapy, Proteomics Unit, Center for Applied Medical Research (CIMA), 31008 Pamplona, Spain

**Francisco J Blanco**

Platform of Proteomics, Proteo-Red-ISCIH INIBIC-Hospital Universitario A Coruña-Spain

**Francisco Javier Fernández Acero**

Laboratory of Microbiology, Marine and Environmental Sciences Faculty, University of Cádiz, Pol. RíoSan Pedro s/n, Puerto Real, Cádiz, Spain

**Francisco Torrens**

InstitutUniversitari de CiènciaMolecular, Universitat de València, Spain

**François Fenaille**

CEA, IBItecS, Service de Pharmacologie et DImmunoanalyse (SPI) France

**Georgios Theodoridis**

Department of Chemistry, Aristotle University, Greece

**GermanBou**

Servicio de Microbiología-INIBIC, ComplejoHospitalario Universitario la Coruña, As Xubias s/n, 15006 LaCoruña, Spain

**Gianfranco Mamone**

Proteomic and Biomolecular Mass Spectrometry Centre, Institute of Food Science CNR, Via Roma 52 A/C, I-83100 Avellino, Italy

**Gianfranco Romanazzi**

Department of Environmental and Crop Sciences, Marche Polytechnic University, Via BrecciaBianche60131 Ancona, Italy

**Gianluigi Mauriello**

Department of Food Science, University of Naples Federico II Naples, Italy

**Giuseppe Palmisano**

Department of Biochemistry and Molecular Biology University of Southern Denmark, Odense M, Denmark

**Hugo Miguel Baptista Carreira dos Santos**

REQUIMTE-FCT Universidade NOVA de Lisboa, Portugal

**Ignacio Casal**

FunctionalProteomicsLaboratory, Centro de Investigaciones Biológicas (CSIC), Ramiro de Maeztu 9, 28040 Madrid, Spain

**IñakiÁlvarez**

Institut de Biotecnologia i BiomedicinaVicentVillarPalasí, Universitat Autònoma de Barcelona, 08193, Bellaterra, Barcelona

**Isabel Liste**

Area de Biología Celular y delDesarrollo, Instituto de Salud Carlos III, Madrid, Spain

**IsabelleFournier**

University Lille Nord de France, Fundamental & Applied Biological Mass Spectrometry - EA 4550, Villeneuve d'Ascq, France

**Jacek Z. Kubiak**

CNRS UMR 6061, University of Rennes 1, Institute of Genetics and Development of Rennes, Rennes, France

**Jane Thomas-Oates**

Centre of Excellence in Mass Spectrometry and Department of Chemistry, University of York, Heslington, York YO10 5DD, UK

**Jatin Burniston**

Muscle Physiology and Proteomics Laboratory, Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, Tom Reilly Building, Liverpool, United Kingdom

**Jerry Thomas**

Technology Facility, Department of Biology, University of York, UK

**Jesús Jorrín Novo**

Agricultural and Plant Biochemistry, Proteomics Research Group, Department of Biochemistry and Molecular Biology, Córdoba, Spain

**Jesús Mateos Martín**

Osteoarticular and AgingResearch Lab, ProteomicsUnit INIBIC-ComplejoHospitalarioUniversitario de A Coruña, A Coruña, Spain

**Joan Cerdà**

Laboratory IRTA, Institute of Marine Sciences (CSIC), Passeigmarítim 37-49, 08003 Barcelona, Spain

**João Rodrigues**

Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, Portugal

**Johan Palmfeldt**

Research Unit for Molecular Medicine, Aarhus University Hospital, Skejby, Aarhus, Denmark

**Jose Andrés Fernández González**

Universidad del Pais Vasco, Facultad de Ciencia y Tecnología, Spain

**Jose Cremata Alvarez**

Department of Carbohydrate Chemistry, Center for Genetic Engineering and Biotechnology, Havana 10600, Cuba

**Jose Manuel Palma**

Departamento de Bioquímica, Biología Celular y Molecular de Plantas Estacion Experimental del Zaidin, CSIC, Granada, Spain

**José Manuel Bautista**

Departamento de Bioquímica y Biología Molecular IV, Universidad Complutense de Madrid, Spain

**José Moreira**

Danish Center for Translational Breast Cancer Research, Denmark

**Juraj Gregan**

Max F. Perutz Laboratories, University of Vienna, Dr. Bohr-Gasse 1, 1030 Vienna, Austria

**Karin Stensjö**

Department of Photochemistry and Molecular Science, Ångström laboratory, Uppsala University, Sweden

**Kathleen Marchal**

CMPG/Bioinformatics, Dep Microbial and Molecular Systems, Leuven, Germany

**Kay Ohlendieck**

Department of Biology, National University of Ireland, Maynooth, Co. Kildare, Ireland

**Keiryn Bennett**

CeMM - Center for Molecular Medicine of the Austrian Academy of Sciences Vienna, Austria

**Kjell Sergeant**

Centre de Recherche Public-Gabriel Lippmann, Department 'Environment and Agro-biotechnologies' (EVA), 41, rue du Brill, 4422 Belvaux, Luxembourg

**Lennart Martens**

Department of Medical Protein Research, VIB and Department of Biochemistry, Ghent University, B-9000 Ghent, Belgium

**Luis P. Fonseca**

Instituto Superior Técnico, Centro de Engenharia Biológica e Química, Institute for Biotechnology and Bioengineering, Av. Rovisco Pais, 1049-001 Lisboa, Portugal

**Manuel AvilésSanchez**

Department of Cell Biology and Histology, School of Medicine, University of Murcia, Espinardo, 30100 Murcia, Spain

**Marcello Donini**

ENEA -Casaccia Research Center, UTBIORAD-FARM, Biotechnology Laboratory, Italy  
**María de la Fuente**  
 Legume group, Genetic Resources, Mision Biologica de Galicia-CSIC, Pontevedra, Spain  
**Maria M. Malagón**  
 Department of Cell Biology, Physiology and Immunology, IMIBIC, Universidad de Córdoba, Spain  
**Maria Gabriela Rivas**  
 REQUIMTE/CQFB, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal  
**María Mayán**  
 INIBIC, LaCoruña, Spain  
**María Páez de la Cadena**  
 Department of Biochemistry, Genetics and Immunology, University of Vigo, Spain  
**Marie-Pierre Bousquet**  
 Institut de Pharmacologie et de Biologie Structurale, UPS/CNRS UMR5089, 205 route de Narbonne, Toulouse, France  
**Mario Diniz**  
 Dept. Química-REQUIMTE, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal  
**Martin Hajduch**  
 Department of Reproduction and Developmental Biology, Institute of Plant Genetics and Biotechnology, Slovak Academy of Sciences, Nitra, Slovakia  
**Martin Kussmann**  
 Faculty of Science, Aarhus University, Aarhus, Denmark  
**Martina Marchetti-Deschmann**  
 Institute of Chemical Technologies and Analytics, Vienna University of Technology, Vienna, Austria  
**Meri Hovsepyan**  
 Institute of Molecular Biology of Armenian National Academy of Sciences Yerevan, Armenia  
**Miguel Reboiro Jato**  
 Escuela Superior de Ingeniería Informática, Ourense, Spain  
**Monica Carrera**  
 Institute of Molecular Systems Biology, Zurich, Germany  
**Okay Saydam**  
 Molecular Oncology Laboratory, Division of Neuro-Oncology, Department of Pediatrics Medical University of Vienna, Austria  
**Ola Söderberg**  
 Department of Immunology, Genetics and Pathology, Uppsala University, Sweden  
**Patrice Francois**  
 Genomic Research Laboratory, Service of Infectious Diseases, Department of Internal Medicine, Geneva  
**Patrícia Alexandra Curado Quintas Dinis Poeta**  
 University of Trás-os-Montes and Alto Douro (UTAD), School of Agrary and Veterinary Sciences, Veterinary, Science Department, Portugal  
**Pantelis Bagos**  
 Department of Computer Science and Biomedical Informatics, University of Central Greece, Greece  
**Pedro Baptista**  
 Centre for Research in Human Molecular Genetics, Department of LifeSciences, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal  
**Pedro Santos**  
 CBMA-Centre of Molecular and Environmental Biology, Department of Biology, University of Minho, 4715-057 Braga, Portugal  
**Pedro S. Lazo**  
 Departamento de Bioquímica y Biología Molecular, Instituto Universitario de Oncología Del Principado de Asturias (IUOPA), Universidad de Oviedo, 33071 Oviedo, Spain  
**Philippe Castagnone-Sereno**  
 Interactions Biotiques et Sante Vegetale., Sophia Antipolis cedex France  
**Pierscionek Barbara**  
 School of Biomedical Sciences, University of Ulster, Cromore Road, Coleraine, BT52 1SA, United Kingdom  
**Pieter de Lange**  
 DipartimentodiScienzedellaVita, SecondaUniversità degli Studi di Napoli, Via Vivaldi 43, 81100 Caserta, Italy  
**Qi Zhu**  
 Dept. Electrical Engineering, ESAT/SCD, Katholieke Universiteit Leuven, Heverlee, Belgium  
**Ralf Hoffmann**  
 Institute of Bioanalytical Chemistry, Center for Biotechnology and Biomedicine, Faculty of Chemistry and Mineralogy, Leipzig University, Germany  
**Ricardo Gutiérrez Gallego**  
 Bioanalysis Group, Neuropsychopharmacology Program IMIM-Hospital del Mar & Department of Experimental and Health Sciences, University Pompeu Fabra, Spain  
**Roque Bru Martinez**  
 Plant Proteomics and Functional Genomics Group, Department of Agrochemistry and Biochemistry, Faculty of Sciences, Alicante University, Spain  
**Rune Matthiesen**  
 Institute of Molecular Pathology and Immunology, University of Porto, Rua Dr. Roberto Frias s/n, 4200-465 Porto, Portugal  
**Ruddy Wattiez**  
 Department of Proteomics and Microbiology, University of Mons (UMONS), Belgium  
**Ruth Birner-Gruenberger**  
 Medical University Graz, Austria  
**Christian Lindermayr**  
 Institute of Biochemical Plant Pathology, Helmholtz Zentrum München, German Research Center for Environmental Health, D-85764 Neuherberg, Germany  
**Sabine Luthje**  
 University of Hamburg, Biocenter Klein Flottbek, Hamburg, Germany  
**Salvador Ventura**  
 Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, Spain  
**Shan He**  
 Centre for Systems Biology, School of Biosciences and School of Computer Science, The University of Birmingham, England  
**Silvia Mazzuca**  
 Plan Cell Physiology Laboratory, Department of Ecology, University of Calabria, Italy  
**Sophia Kossida**  
 Biomedical Research Foundation, Academy of Athens, Department of Biotechnology, Soranou-Efesiou 4, 11527 Athens, Greece  
**Spiros D. Garbis**  
 Biomedical Research Foundation of the Academy of Athens, Center for Basic Research - Division of Biotechnology, 4 Soranou-Efesiou Street, 11527 Athens, Greece  
**Stefania Orrù**  
 University of Naples Parthenope, Naples, Italy  
**Stefano Curcio**  
 Department of Engineering Modeling, Laboratory of Transport Phenomena and Biotechnology University of Calabria, Italy  
**Susana Cristóbal**  
 Department of Clinical and Experimental Medicine Faculty of Health Science

Linköping University, Sweden

**Tâmará García Barrera**

Departamento de Química y Ciencia de los Materiales, Facultad de Ciencias Experimentales, Universidad de Huelva, Spain

**Theodore Alexandrov**

University of Bremen, Center for Industrial Mathematics, Germany

**Thole Züchner**

Ultrasensitive Protein Detection Unit, Leipzig University, Center for Biotechnology and Biomedicine, Institute of Bioanalytical Chemistry, Germany

**Tiziana Bonaldi**

Department of Experimental Oncology, European Institute of Oncology, Via Adamello 16, 20139 Milan, Italy

**Tsangaris George**

Proteomics Research Unit, Center of Basic Research II Foundation of Biomedical Research of the Academy of Athens, Athens, Greece

**Üner Kolukisaoglu**

Center for Plant Molecular Biology, Eberhard-Karls University Tübingen, Tübingen, Germany

**Valeria Bertagnolo**

Department of Morphology and Embryology University of Ferrara, Italy

**Vera Muccilli**

Dipartimento di Scienze Chimiche, Università di Catania, V.le A. Doria 6, 95126 Catania, Italy

**Virginie Brun**

French Atomic Energy Commission and *French National Institute for Health and Medical Research*, France

**Vladislav Khrustalev**

Department of General Chemistry, Belarussian, State Medical University, Dzerzhinskogo, 83 220029, Communisticheskaya 7-24, Minsk, Belarus

---

**SOUTH AMERICA**

**Deborah Schechtman**

Department of Biochemistry, Chemistry Institute, University of São Paulo, Brazil

**Fabio Ribeiro Cerqueira**

Department of Informatics and NuBio (Research Group for Bioinformatics), University of Vicosá, Brazil

**Luis Pacheco**

Institute of Health Sciences, Federal University of Bahia Salvador, Brazil

**Mário Hiroyuki Hirata**

Laboratório de Biologia Molecular Aplicado ao Diagnóstico, Departamento de Análises Clínicas e Toxicológicas, Faculdade de Ciências Farmacêuticas, Universidade de São Paulo, Brazil

**Jorg Kobarg**

Centro Nacional de Pesquisa em Energia e Materiais, Laboratório Nacional de Biociências, Brazil

**Marcelo Bento Soares**

Cancer Biology and Epigenomics Program, Children's Memorial Research Center, Professor of Pediatrics, Northwestern University's Feinberg School of Medicine

**Rinaldo Wellerson Pereira**

Programa de Pós Graduação em Ciências Genômicas e Biotecnologia Universidade Católica de Brasília, Brazil

**Roberto Bobadilla**

BioSigma S.A., Santiago de Chile, Chile

**Rossana Arroyo**

Department of Infectomic and Molecular Biology, Center of Research and Advanced Studies of the National, Polytechnical Institute (CINVESTAV-IPN), Mexico City, Mexico

**Rubem Menna Barreto**

Laboratório de Biologia Celular, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

**Vasco Azevedo**

Biological Sciences Institute, Federal University of Minas Gerais, Brazil

---

**NORTH AMERICA**

**Adam Vigil**

University of California, Irvine, USA

**Akeel Baig**

Hoffmann-La Roche Limited, Pharma Research Toronto, Toronto, Ontario, Canada

**Alexander Statnikov**

Center for Health Informatics and Bioinformatics, New York University School of Medicine, New York

**Anthony Gramolini**

Department of Physiology, Faculty of Medicine, University of Toronto, Canada

**Amrita Cheema**

Georgetown Lombardi Comprehensive Cancer Center, USA

**Anas Abdel Rahman**

Department of Chemistry, Memorial University of Newfoundland Newfoundland and Labrador St. John's, Canada

**Christoph Borchert**

Biochemistry & Microbiology, University of Victoria, UVic Genome British Columbia Proteomics Centre, Canada

**David Gibson**

University of Colorado Denver, Anschutz Medical Campus, Division of Endocrinology, Metabolism and Diabetes, Aurora, USA

**Eustache Paramithiotis**

Caprion Proteomics Inc., Montreal, Canada

**Fouad Daayf**

Department of Plant Science, University of Manitoba, Winnipeg, Manitoba, Canada

**Haitao Lu**

Washington University School of Medicine, Saint Louis, USA

**Hsiao-Ching Liu**

232D Polk Hall, Department of Animal Science, North Carolina State University Raleigh, USA

**Hui Zhang**

Johns Hopkins University, MD, USA

**Ing-Feng Chang**

Institute of Plant Biology, National Taiwan University, Taipei, Taiwan

**Irwin Kurland**

Albert Einstein College of Medicine, Associate Professor, Dept of Medicine, USA

**Jagjit Yadav**

Microbial Pathogenesis and Toxicogenomics, Laboratory, Environmental Genetics and Molecular, Toxicology Division, Department of Environmental Health, University of Cincinnati College of Medicine, Cincinnati, Ohio, USA

**Jianbo Yao**

Division of Animal and Nutritional Sciences, USA

**Jiaxu Li**

Department of Biochemistry and Molecular Biology, Mississippi State University, USA

**Jiri Adamec**

Department of Biochemistry & Redox Biology Center, University of Nebraska, Lincoln Nebraska, USA

**Jiye Ai**

University of California, Los Angeles

**Joshua Heazlewood**

Lawrence Berkeley National Laboratory, Berkeley, CA, USA

**Laszlo Prokai**

Department of Molecular Biology & Immunology, University of North Texas Health Science Center, Fort Worth, USA

**Lei Li**

University of Virginia, USA

**Leonard Foster**

Centre for High-throughput Biology, University of British Columbia, Vancouver, BC, Canada

**Madhulika Gupta**

Children's Health Research Institute, University of Western Ontario London, ON, Canada

**Masaru Miyagi**

Case Center for Proteomics and Bioinformatics, Case Western Reserve University, Cleveland, USA

**Michael H.A. Roehrl**

Department of Pathology and Laboratory Medicine, Boston Medical Center Boston, USA

**Ming Zhan**

National Institute on Aging, Maryland, USA

**R. John Solaro**

University of Illinois College of Medicine, USA

**Rabih Jabbour**

Science Application International Corporation, Maryland, USA

**Ramesh Katam**

Plant Biotechnology Lab, Florida A and M University, FL, USA

**Robert L. Hettich**

Chemical Sciences Division, Oak Ridge National Laboratory, Oak Ridge, USA

**Robert Powers**

University of Nebraska-Lincoln, Department of Chemistry, USA

**Shen S. Hu**

UCLA School of Dentistry, Dental Research Institute, UCLA Jonsson Comprehensive Cancer Center, Los Angeles CA, USA

**Shiva M. Singh**

University of Western Ontario, Canada

**Terry D. Cyr**

Genomics Laboratories, Centre for Vaccine Evaluation, Biologics and Genetic Therapies Directorate, Health Products and Foods Branch, Health Canada, Ontario, Canada

**Thibault Mayor**

Department of Biochemistry and Molecular Biology, Centre for High-Throughput Biology (CHiBi), University of British Columbia, Canada

**Thomas Conrads**

USA

**Thomas Kislinger**

Department of Medical Biophysics, University of Toronto, Canada

**Wayne Zhou**

Marine Biology Laboratory, Woods Hole, MA, USA

**Wei Jia**

US Environmental Protection Agency, Research Triangle Park, North Carolina, USA

**Wei-Jun Qian**

Pacific Northwest National Laboratory, USA

**Xiangjia Min**

Center for Applied Chemical Biology, Department of Biological Sciences Youngstown State University, USA

**Wan Jin Jahng**

Department of Biological Sciences, Michigan Technological University, USA

**William A LaFramboise**

Department of Pathology, University of Pittsburgh School of Medicine Shadyside Hospital, Pittsburgh, PA 15232 USA

**Xuequn Chen**

Department of Molecular & Integrative Physiology, The University of Michigan, Ann Arbor, USA

**Ying Xu**

Department of Biochemistry and Molecular Biology, Institute of Bioinformatics, University of Georgia, Life Sciences Building Athens, GA, USA



## CONTENTS OF VOLUME 1 | ISSUE 2 | DECEMBER 2011

### PRELIMINARY COMMUNICATIONS

---

- Preliminary serological secretome analysis of *Corynebacterium Pseudotuberculosis*. 193  
 Nubia Seyffert, Luis G.C. Pacheco, Wanderson M. Silva, Thiago L.P. Castro, Agenor V. Santos, Anderson Santos, John A. McCulloch, Maira R. Rodrigues, Simone G. Santos, Luiz M. Farias, Maria A.R. Carvalho, Adriano M.C. Pimenta, Artur Silva, Roberto Meyer, Anderson Miyoshi, Vasco Azevedo.

### ORIGINAL ARTICLES

---

- The effect of Zn<sup>2+</sup> on prostatic cell cytotoxicity caused by *Trichomonas vaginalis*. 198  
 Laura Isabel Vazquez-Carrillo, Laura Itze Quintas-Granados, Rossana Arroyo, Guillermo Mendoza Hernández, Arturo González-Robles, Bertha Isabel Carvajal-Gamez, M. Elizabeth Álvarez-Sánchez.
- Label-free protein quantification on tandem mass spectra in an ion trapping device. 211  
 HuiSong Pak, Carla Pasquarello, Alexander Scherl.
- Development of an in-house protocol for the OFFGEL fractionation of plant proteins. 216  
 Delphine Vincent, Peter S. Solomon.
- Quantitative Proteomic Analysis of Retina in Oxygen-Induced Retinopathy Mice using iTRAQ with 2D NanoLC-nanoESI-MS/MS. 226  
 Lei Zhou, Liu Xinling, Siew Kwan Koh, Li Xiaorong, Roger W Beuerman.
- A method for estimation of immunogenic determinants mutability: case studies of HIV1 gp120 and diphtheria toxin. 236  
 Khrustalev Vladislav Victorovich, Barkovsky Eugene Victorovich, Vasilevskaya Alla Evgenyevna, Skripko Svetlana Michailovna, Kolodkina Valentina Leonidovna, Ignatyev Georgiy Michailovich, Semizon Pavel Anatolyevich.
- Application of high content biology demonstrates differential responses of keratin acetylation sites to short chain fatty acids and to mitosis. 253  
 AQ Khan, L Shaw, PJM Drake, GJ Griffiths, SR Brown, BM Corfe.
- Bacterial two-hybrid analysis of the *Shewanella oneidensis* MR-1 multicomponent electron transfer pathway. 260  
 Jimmy Borloo, Lina De Smet, Jozef J. Van Beeumen, Bart Devreese.
- Extraction and purification of the lectin found in the tubers of *Eranthis hyemalis* (winter aconite). 268  
 Oliver George, Claudia Solscheid, Emilia Bertolo, Joanna Fell, David Lisgarten.
- Identification of spore specific allergens from *Penicillium chrysogenum*. 272  
 Verena Goodman, Andrea Müller, Helke Gröger-Arndt, Dietmar Schlosser, Nico Jehmlich, Regina

Treudler, Irina Lehmann, Jan-Christoph Simon, Martin von Bergen.

Differential expression of Outer membrane proteins in early stages of meropenem-resistance in *Acinetobacter baumannii*. 280

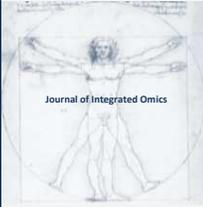
Jitendraa Vashistt, Vishvanath Tiwari, Arti Kapil, Rajeswari R. Moganty.

Proteomic approach for molecular physiological mechanism on consecutive monoculture problems of *Rehmannia glutinosa*. 287

Wenxiong Lin, Changxun Fang, Linkun Wu, Gailing Li, Zhongyi Zhang.

Hemoglobin subunit beta (HBB) is a potential biomarker for predicting response to Ge2tinib in NSCLC patients. 297

Poh Kuan Chong, Huiyin Lee, Marie Chiew Shia Loh, Wan-Teck Lim, Siew Pang Chan, Ross Andrew Soo, Yoon Pin Lim.



PRELIMINARY COMMUNICATION | DOI: 10.5584/jiomics.v1i2.54

## Preliminary serological secretome analysis of *Corynebacterium pseudotuberculosis*

Núbia Seyffert<sup>1,2</sup>, Luis G.C. Pacheco<sup>1,3,4</sup>, Wanderson M. Silva<sup>1</sup>, Thiago L.P. Castro<sup>1</sup>, Agenor V. Santos<sup>3,5</sup>, Anderson Santos<sup>1</sup>, John A. McCulloch<sup>1,5</sup>, Maira R. Rodrigues<sup>1</sup>, Simone G. Santos<sup>2</sup>, Luiz M. Farias<sup>2</sup>, Maria A.R. Carvalho<sup>2</sup>, Adriano M.C. Pimenta<sup>3</sup>, Artur Silva<sup>5</sup>, Roberto Meyer<sup>4</sup>, Anderson Miyoshi<sup>1</sup>, Vasco Azevedo<sup>1\*</sup>.

<sup>1</sup>Departamento de Biologia Geral, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais. Belo Horizonte 31.270-901, Brazil;

<sup>2</sup>Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais. Belo Horizonte 31.270-901, Brazil;

<sup>3</sup>Departamento de Bioquímica, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais. Belo Horizonte 31.270-901, Brazil;

<sup>4</sup>Instituto de Ciências da Saúde, Universidade Federal da Bahia, Salvador 40.110-902, Brazil; <sup>5</sup>Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém 66.075-900, Brazil.

Received: 02 November 2010 Accepted: 19 June 2011 Available Online: 20 June 2011

### ABSTRACT

Caseous lymphadenitis (CLA) is a chronic disease affecting small ruminants that is caused by *Corynebacterium pseudotuberculosis* and is responsible for significant economic losses. Various *C. pseudotuberculosis* secreted proteins are known to react with sera from infected goats. Mapping of the secretome would help us understand the pathogenesis of CLA. We identified six immunoreactive secreted proteins of *C. pseudotuberculosis* by 2D-Western blotting, using sera from goats with CLA, and characterized them by mass spectrometry. This preliminary information will give support to future studies aimed at the development of efficient vaccines and diagnostic kits.

**Keywords:** *C. pseudotuberculosis*, Caseous lymphadenitis, Secretome.

### 1. Introduction

Caseous lymphadenitis (CLA) is a chronic disease affecting small ruminants; it is caused by infection with *Corynebacterium pseudotuberculosis* and is responsible for significant worldwide economic losses due to decreases in both the productivity and the reproductive performance of infected animals [1]. The lack of efficient immunoprophylaxis against CLA results in ineffective management of this disease in animals, facilitating its dissemination [2]. Efficient vaccines against CLA and diagnostic kits for this disease are still not available, in part due to a lack of sufficient information concerning newly-characterized *C. pseudotuberculosis* virulence determinants [3].

Only a few genes and their products have been identified as factors that contribute to the virulence of *C. pseudotuberculosis*, including phospholipase D (PLD) [4,5], the *fagABC*

operon involved in iron acquisition by the cell [6] and the protease CP40 [7]. Chaplin et al. [8] developed a DNA vaccine encoding PLD to immunize sheep, but they achieved only partial protection against challenge with *C. pseudotuberculosis*. Similar results were obtained when sheep were immunized with a formalin-inactivated subunit vaccine [9]. CP40 protease has been reported as a possible candidate for the development of vaccines, based on Western blot analysis with serum samples from sheep experimentally infected with *C. pseudotuberculosis* [10].

To date, the search for immunogenic proteins has been carried out in a non-exhaustive manner, using various extraction and separation techniques [11,12]. A comprehensive analysis of the entire set of proteins expressed by *C. pseudotuberculosis* strains is needed in order to identify the best

\*Corresponding author: Dr. Vasco Azevedo. Departamento de Biologia Geral, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais. Belo Horizonte 31.270-901, Brazil. E-mail Address: vasco@icb.ufmg.br.

candidate proteins for immunoprophylactic or diagnostic applications. Bacterial secreted proteins have a various biological functions, ranging from toxicity to more subtle alterations of the host cell for the benefit of the invader; they are an important part of the pathogenic process [13].

There has not been much conclusive research concerning pathogenesis or even the immune response against *C. pseudotuberculosis* infection, compared to research on other important veterinary pathogens, such as *Mycobacterium tuberculosis* [14]. Mapping of the *C. pseudotuberculosis* secretome, followed by characterization of expressed proteins and assessment of their immunogenic potential would be ideal for shedding light on the pathogenesis of specific strains and the host immune response that they provoke, paving the way for the development of more efficient vaccines and diagnostic kits. We examined antigenic proteins of the secretome of *C. pseudotuberculosis* strain 1002 cultivated in chemically defined medium (CDM) using serological proteome analysis (SERPA) [15].

## 2. Material and Methods

### 2.1 Bacterial strain and growth conditions

*Corynebacterium pseudotuberculosis* strain 1002, originally isolated from an infected goat in Brazil [16], was routinely maintained in Brain Heart Infusion broth (BHI) and characterized by biochemical and molecular methods, as previously described [17]. For SERPA, bacteria were grown at 37°C under agitation (100 rpm), for 24 h in 1 L of chemically defined medium (CDM), until reaching the exponential growth phase (OD<sub>600nm</sub> = 1.3). The CDM contained 0.067M of phosphate buffer, 0.05% (v/v) Tween 80 (Sigma), 4% (v/v) 100X minimal Essential Medium (MEM) Vitamin Solution (Invitrogen), 1% (v/v) of 50X MEM Amino Acids Solution (Invitrogen), 1% (v/v) 100X MEM Non-Essential Amino Acids Solution (Invitrogen) and 1.2% (w/v) filter-sterilized glucose, as previously described [18].

### 2.2 Extraction of secreted proteins

*Corynebacterium pseudotuberculosis* exoproteins were obtained according to a previously described three-phase partitioning (TPP) protocol [19]. Briefly, bacterial cells were separated from the supernatant by centrifugation at 4,000 rpm for 20 min at 4°C. The supernatant was filtered through a 0.22 µm membrane (filter) and 30% (w/v) ammonium sulphate was added. The pH was adjusted to 4.0 and *n*-butanol was added at a ratio of 1:1, and the sample was vortexed. After 1h of incubation at room temperature, the precipitate at the interface was collected and re-suspended in 1 mL of 20 mM Tris-HCl buffer pH 7.4 with 10 µL of protease inhibitor (GE).

### 2.3 2D-PAGE-Western blot

Two-dimensional electrophoretic separation was carried

out, as previously described [20], with minor modifications. Secreted proteins (150 µg) were dissolved in 2-DE sample buffer (8 M urea, 2 M thiourea, 4% CHAPS, 1% (v/v) carrier ampholyte pH 3.0-5.6, 80 mM dithiothreitol (DTT), 40 mM Tris-base and bromophenol blue. The mixture was used for overnight rehydration of 11 cm immobilized pH gradient (IPG) strips (Immobiline DryStrip™ Gels, pH 3.0-5.6 NL [GE Healthcare]). Isoelectric focusing (IEF) was carried out at room temperature for 24.5 h (maximum voltage of 3,500 V and maximum current of 50 µA). After equilibration for 15 min in a 50 mM Tris HCl (pH 8.8) buffer solution containing 6 M urea, 2% (w/v) SDS, 30% (v/v) glycerol and 0.001% (v/v) bromophenol blue and 10 mg/mL DTT, the strips were equilibrated for 15 min in the same solution, except that the DTT was replaced by 25 mg/mL iodoacetamide. The proteins were resolved in 2D electrophoresis in 12% polyacrylamide gels under denaturing conditions, using a Protean IIXi system (Biorad). Protein spots were visualized by staining with Coomassie blue G-250 (GE Healthcare). For each protein sample, three 2D gels were stained to visualize proteins and six 2D gels were electroblotted onto polyvinylidene difluoride membranes (Owl system) for 1 h, with an electric current of 0.4 A. The membranes were blocked overnight at 4° C in 5% non-fat milk in phosphate buffered saline pH 7.5 with 0.05% Tween 20 (PBS-T). The membranes were then incubated at room temperature for 1 h in PBS-T with sera (at a proportion of 1:100 v/v PBS-T:serum) obtained from animals either infected or uninfected with *C. pseudotuberculosis*. The membranes were then washed with PBS-T three times for 5 min and incubated for 1h with an anti-goat IgG peroxidase antibody produced in rabbits (Sigma), diluted 1:1000 in PBS-T solution. Antibody-tagged protein spots were detected with DAB peroxidase substrate solution.

### 2.4 Identification of immunoreactive proteins

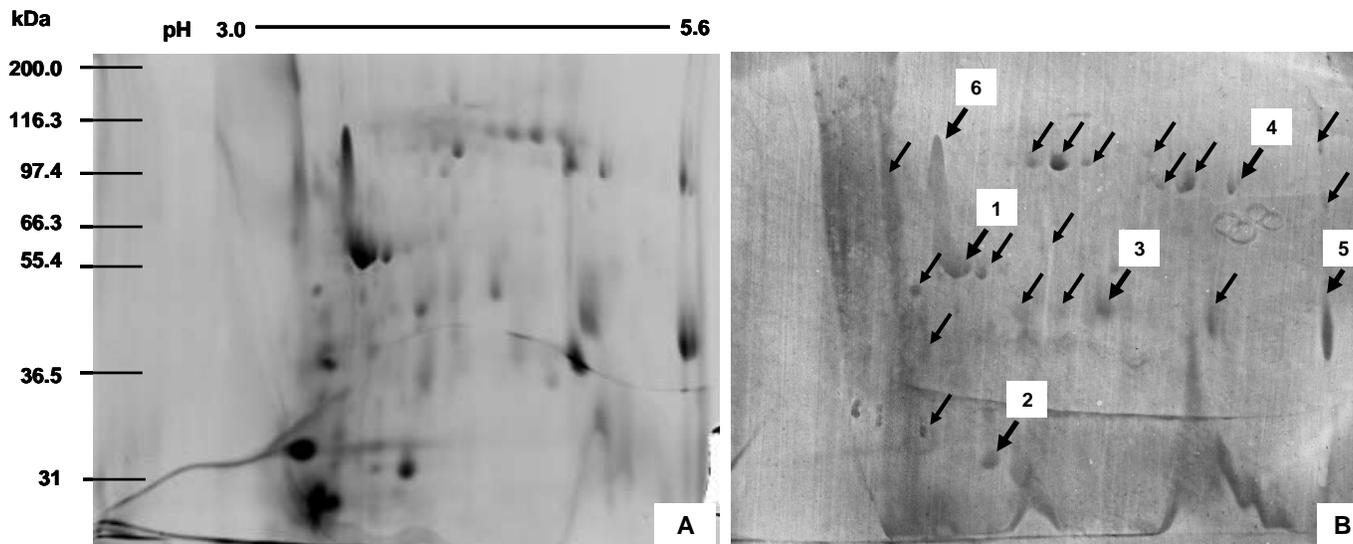
Membranes were digitally scanned and immunoreactive proteins matched to 2D gel images of the samples were identified using the Melanie software (GeneBio). All spots reactive in 2D-Western blots were selected from an analogous 2D stained gel and manually excised. The excised gel fragments were incubated overnight with 25mM bicarbonate/50% acetonitrile (ACN) solution until completely destained. After drying, gel fragments were placed in 50 mM ammonium bicarbonate solution with 20ng/µL sequencing-grade modified trypsin (Promega Biosciences, CA, USA). Digestion was run at 37° C overnight. The peptides were extracted using 5% formic acid/50% acetonitrile solution, concentrated in a SpeedVac (Savant, USA) to a volume of about 10 µL, desalted using ZipTip® C18 plates (C18 resin, P10; Millipore Corporation, Bedford, MA, USA) and eluted with 0.1% trifluoroacetic acid solution containing 50% ACN. The sample extract was mixed at a 1:1 ratio with matrix (10 mg/mL recrystallized α-

cyano-4-hydroxycinnamic acid) to a final volume of 1 µL and then spotted onto an MTP AnchorChip™ 600/384 (Bruker Daltonics) for matrix-assisted laser desorption/ionization time-of-flight tandem mass spectrometry (MALDI-TOF-MS/MS) (LIFT technology, Autoflex III™; BrukerDaltonics, Billerica, USA) analysis. Ionization was performed in MS/MS (PSD-LIFT technology) by irradiation of a nitrogen laser (337 nm) operating at 50 Hz. Data were acquired at a maximum accelerating potential of 25 kV in the positive and reflector modes. Trypsin and keratin contamination peaks were excluded from the mass spectra and MS/MS results were used to search the *C. pseudotuberculosis* 1002 (Gen Bank: CP001809.1) protein database using MASCOT software (<http://www.matrixscience.com/>).

**3. Results and Discussion**

Bacterial growth within a host resulting in infection is a consequence of colonization, adherence, invasion, evasion of the immune response and toxigenesis caused by the bacterial cell. This feat can be accomplished by a bacterial strain through temporal expression of a panoply of virulence genes (the virulon), in response to appropriate environmental stimuli. Characterization of when, which and what amounts of virulence factors are expressed in response to certain stimuli is necessary for understanding the pathogenesis of bacterial species [13]. The dynamics of the immune response to infection can only be fully understood if we characterize the bacterial proteins responsible for eliciting the immune response. The advent of genomics has made this approach feasible, since information concerning an immunogenic

protein can be traced back to the genome, and thence to the regulon that is involved. Several species of Actinobacteria have been the subject of proteomic analysis, including *Mycobacterium avium*, *Mycobacterium tuberculosis*, *Rhodococcus equi* and *Corynebacterium diphtheriae*, yielding insight into the relationships between the host and the bacterial parasite [21,22,23,24,25]. The secretome of *C. pseudotuberculosis* was first studied by Braithwaite et al. [12], who extracted proteins from a culture supernatant with ammonium sulphate; they found seven proteins with molecular weights between 14 and 64 kDa, five of which reacted with sera obtained from goats infected with *C. pseudotuberculosis*. A later study described the reaction of a pool of sera obtained from goats suffering from CLA against 11 *C. pseudotuberculosis* secreted proteins with molecular weights ranging from 24 to 125 kDa [19]; these proteins induced an increase in the serum concentration of IFN-γ in goats infected with this bacterium [26]. We made a follow up of that study. The excreted-secreted antigens of *C. pseudotuberculosis* were obtained by culturing the 1002 strain in CDM [18], with subsequent extraction of secreted proteins by TPP [19], and a 2D-PAGE-Western blot. Twenty-three immunoreactive spots were detected using sera obtained from animals with CLA; due to time and budget constraints, only six of these proteins were identified by MALDI-TOF-MS/MS (Figure 1 and Table 1). Immunoproteomic methods, such as SERPA [15], have been used to identify biomarkers and target antigens for developing diagnostic kits based on antibody/antigen detection, as well as to develop vaccines and treatments for various infectious diseases. Due to the unfeasibility of targeting many proteins simultaneously, our objective was to



**Figure 1.** Serological proteome analysis of secreted proteins of *Corynebacterium pseudotuberculosis* using serum from infected goats. A) 2D-PAGE with 150µg of sample for analysis of *C. pseudotuberculosis* secreted proteins. Spots were detected in the gels stained with Coomassie G-250. B) 2D-PAGE-Western blot analysis of *C. pseudotuberculosis* secreted proteins with 150µg sample. The black arrows indicate 23 immunoreactive spots detected by anti-goat IgG peroxidase antibody produced in rabbits. Numbers correspond to the proteins identified in Table 1.

**Table 1.** Antigenic proteins of *Corynebacterium pseudotuberculosis* identified by MALDI-TOF MS/MS. <sup>a</sup> Accession numbers in Entrez Protein (NCBI Genome CP001809.1). <sup>b</sup> Theoretical molecular weights (Mr). <sup>c</sup> Theoretical isoelectric points (pI)

Spot/Protein description	<sup>a</sup> Protein ID/NCBI	<sup>b</sup> Mr(kDa)/ <sup>c</sup> pI	MASCOT score (%)	Coverage (%)	Peptide Sequence	Ion score
<b>1</b> Resuscitation-promoting factor RpfB	gij302330380	40.31/5.06	298	13	K.AGYTVGDKDIVYPGLTEK.I	70
					K.TVFTGIAAATVKDVLAEK.G	114
					K.VQASQGWGAWPACTSK.L	122
<b>2</b> Putative secreted protein	gij302330462	24.39/5.34	174	27	K.AKDFADTLPEPLR.N	34
					K.DFADTLPEPLR.N	26
					K.LGPNHQAMNVHWFNLSTLQGGSTR.L	86
					R.VIALIEGTIATEASPCTFLPTAALFEVK.L	28
<b>3</b> NlpC/ P60 protein	gij302331099	36.63/5.62	115	6	R.GAVIDPLTNAVSAENPQNAIDR.A	115
<b>4</b> Putative efflux system protein	gij302331553	59.55/5.60	232	14	R.VLVEGTVEPIR.T	24
					R.DQLISAALDAAR.T	38
					K.TKPLYVPEIELTGNR.D	54
					K.NREPIKLPSEAVYGENNAK.K	26
					R.TVTVGNTTDIAEITGGELKPGDK.V	90
<b>5</b> Surface layer protein A (Spl A)	gij302329946	38.67/5.90	530	32	R.VVEAWAHSPSMNR.N	39
					K.ASSPDRPTVYLLNGDGGEGR.A	26
					R.GHATPEQMWGPMGSDYNR.Y	72
					R.YNDAVVMAEDLR.G	80
					R.GTEVYVSNASGVAGGHDLANPR.F	132
					R.LQSLNIPADFNLR.N	60
R.NTGTHSWSYWQDDL.R.A	120					
<b>6</b> Metalloendopeptidase-like protein	gij302205783	24.83/7.23	64	7	K.IVVHTPAMGTLTSPYGMK.VV	27

identify polypeptide chains that display immunoreactivity, thus narrowing the number of targets for further experiments. Six of the spots that displayed immunoreactivity by SERPA (Figure 1) and were identified by MALDI-TOF-MS/MS (Table 1) were consistent with five proteins identified by data-independent MS acquisition (LC-MSE) and six proteins predicted *in silico* by SurfG *plus* studies performed by our team to characterize the total *C. pseudotuberculosis* exoproteome, independent of the 2D-PAGE Western blot analysis [27]. Four immunoreactive secreted proteins that we identified, namely, resuscitation-promoting factor B (RpfB), Nlp/P60 protein, putative efflux system protein and surface layer protein A (SplA), have previously been reported from other bacterial species. Rpf homologues are widespread throughout the Actinobacteria [28] and have the ability to stimulate the growth of dormant mycobacteria, especially *Mycobacterium tuberculosis* [29]. NlpC/P60 belongs to the peptidase family and plays a role in turnover of the bacterial cell wall [30]. Previously characterized NlpC/P60 proteins include *Listeria monocytogenes* secreted autolysin P60 [31], *Bacillus subtilis* autolysins [32] and *Escherichia coli* membrane-associated lipoprotein [33]. The bacterial efflux system is composed of proteins that act as a continuous channel for the extrusion of substrates within the cell envelope into the external environment. This mechanism may be involved in various

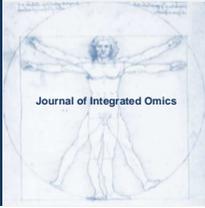
transport functions, including efflux of toxins, metabolites and drugs [34,35]. Surface layer protein A (Spl A) has been described from several bacterial species, in which it has various roles, including nutrient uptake, colonization, antiphagocytosis and exclusion of noxious substances [36]. Putative secreted protein (spot 2) was found to be similar to protein sequences of other Actinobacteria, based on protein BLAST/NCBI (National Center for Biotechnology Information), but it has not been described. Another secreted protein, metalloendopeptidase-like protein belongs to the family of metalloproteases [37], which are common in pathogenic bacteria, including *Pseudomonas aeruginosa* (degradation of host connective tissues) [38], *Clostridium spp.* (neurotoxin activity) [39], *Bacillus anthracis* (lethal toxin) [40] and *Listeria monocytogenes* (enzyme maturation) [41]. Secreted proteins of similar molecular weight have been found in other studies [12,19], through extraction of proteins by other methods and resolution by unidimensional electrophoresis [12,14], but they were not identified as being the CP40 protease [7]. Our preliminary results indicate that SERPA coupled with mass spectrometry analysis is a useful strategy for the identification of these six antigens. Studies are underway to develop a protocol for the detection of the other spots that remained unidentified. These findings may help identify proteins that can induce protective immunity or elicit immune responses with diagnostic value for CLA.

## Acknowledgements

This work was supported by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brasil), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil) and FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais, Brasil).

## References

- M.C.Fontaine, G.J.Baird, *Small Rumin. Res.* 76 (2008) 42–48. DOI:10.1016/j.smallrumres.2007.12.025
- N. Seyffert, A.S. Guimarães, L.G.C. Pacheco, R.W. Portela, B.L. Bastos, F.A. Dorella, M.B. Heinemann, A.P. Lage, A.M.G. Gouveia, R. Meyer, A. Miyoshi, V. Azevedo, *Res. Vet. Sci.* 88 (2010) 50-55. DOI:10.1016/j.rvsc.2009.07.002
- F.A.Dorella, L.G. Pacheco, N. Seyffert, R.W. Portela, R. Meyer, A. Miyoshi, V. Azevedo, *Expert. Rev. Vaccines* 8 (2009) 205-213. DOI:10.1586/14760584.8.2.205
- B.A. Lipsky, A.C. Goldberger, L.S. Tompkins, J.J. Plorde, *Clin. Infect. Dis.* 4 (1982) 1220–1235. DOI:10.1093/clinids/4.6.1220
- A.L. Hodgson, K. Carter, M. Tachedjian, J. Krywult, L.A. Corner, M. McColl, A. Cameron, *Vaccine* 17 (1999) 802–808. DOI:10.1016/S0264-410X(98)00264-3
- S.J. Billington, P.A. Esmay, J.G. Songer, B.H. Jost, *FEMS Microbiol. Lett.* 208 (2002) 41-45. DOI: 10.1111/j.1574-6968.2002.tb11058.x
- M.J. Wilson, M.R. Brandon, J. Walker, *Infect. Immun.* 63 (1995) 206-211.
- P. J. Chaplin, R. De Rose, J. S. Boyle, P. McWaters, J. Kelly, J. M. Tennent, A. M. Lew, J. P. Scheerlinck, *Infect. Immun.* 67 (1999) 6434-6438.
- K. Stanford, K.A. Brogden, L.A. McClelland, G.C. Kozub, F. Audibert, *Can. J. Vet. Res.* 62(1998) 38–43.
- J. Walker, H.J. Jackson, D.G. Eggleton, E.N.T. Meeusen, M.J. Wilson, M.R. Brandon, *Infect. Immun.* 62 (1994) 2562–2567.
- C.A. Muckle, P.I. Menzies, Y. Li, Y.T. Hwang, M. van Wesenbeeck, *Vet. Microbiol.* 30 (1992) 47-58.
- C.E. Braithwaite, E.E. Smith, J.G. Songer, A.H. Reine, *Vet. Microbiol.* 38 (1993) 59-70.
- K. Wooldridge, *Bacterial secreted proteins secretory mechanisms and role in pathogenics*, Norfolk, UK, 2009.
- I. Smith, *Clin. Microbiol, Rev.* 16 (2003) 463–496. DOI:10.1128/CMR.16.3.463-496.2003.
- N. Falisse-Poirrier, V. Ruelle, B. Elmoualij, D. Zorzi, O. Pierard, E. Heinen, E. De Pauw, W. Zorzi, *J. Microbiol. Methods.* 67 (2006) 593-596. DOI:10.1016/j.mimet.2006.05.002.
- R. Meyer, R. Carminati, R. Bahia, V. Vale, S. Viegas, T. Martinez, I. Nascimento, R. Schaer, J. Silva, M. Ribeiro, L. Régis, B. Paule, S. Freire, *J. Med. Biol. Sci.* 1(2002), 42-48.
- L.G.C. Pacheco, R.R. Pena, T.L.P. Castro, F.A. Dorella, R.C. Bahia, R. Carminati, M.N.L. Frota, S.C. Oliveira, R. Meyer, F.S.F. Alves, A. Miyoshi, V. Azevedo, *J. Med. Microbiol.* 56 (2007) 480-486. DOI: 10.1099/jmm.0.46997-0.
- L.F. Moura-Costa, B.J.A. Paule, V. Azevedo, S. M. Freire, I. Nascimento, R. Schaer, L.F. Regis, V.L.C. Vale, D.P. Matos, R.C. Bahia, R. Carminati, R. Meyer, *Rev. Bras. Saúde e Produção Animal*, 3(2002) 1-9.
- B.J.A. Paule, R. Meyer, L.F. Moura-Costa, R.C. Bahia, R. Carminati, L.F. Regis, V.L.C. Vale, S.M. Freire, I. Nascimento, R. Schaer, V. Azevedo, *Protein. Expr. Purif.* 34 (2004) 311-316. DOI: 10.1016/j.pep.2003.12.003.
- R. J. Simpson. Cold Spring Harbor: CSHL Press, 2003
- V. Hughes, J.P. Bannantine, S. Denham, S. Smith, A. Garcia-Sanchez, J. Sales, M.L. Paustian, K. MClean, K. Stevenson, *Clin. Vaccine Immunol.* 15(2008) 1824-1833. DOI: 10.1128/CVI.00099-08.
- P.R. Jungblut, U.E. Schaible, H.J. Mollenkopf, U. Zimny-Arndt, B. Raupach, J. Mattow, P. Halada, S. Lamer, K. Hagens, S.H. Kaufmann, *Mol. Microbiol.* 33(1999) 1103-17. DOI: 10.1046/j.1365-2958.1999.01549.x.
- S. Sinha, K. Kosalai, S. Arora, A. Namane, P. Sharma, A.N. Gaikwad, P. Brodin, S.T. Cole, *Microbiol.* 151 (2005) 2411-2419. DOI:10.1099/mic.0.27799-0.
- C. Barbey, A. Budin-Verneuil, S. Cauchard, A. Hartke, C. Laugier, V. Pichereau, S. Petry, *Vet. Microbiol.* 135 (2009) 334-45. DOI:10.1016/j.vetmic.2008.09.086.
- N. Hansmeier, T. Chao, J. Kalinowski, A. Pühler, A. Tauch, *Proteomics* 6 (2006) 2465–2476. DOI: 10.1002/pmic.200500360.
- R. Meyer, L. Regis, V. Vale, B. Paule, R. Carminati, R. Bahia, L. Moura Costa, R. Schaer, I. Nascimento, S. Freire, *Vet. Immunol. and Immunopathol.* 107 (2005) 249-54. DOI:10.1016/j.vetimm.2005.05.002.
- L.G. Pacheco, S.E. Slade, N. Seyffert, A.R. Santos, T.L. Castro, W.M. Silva, A.V. Santos, S.G. Santos, L.M. Farias, M.A. Carvalho, A.M. Pimenta, R. Meyer, A. Silva, J.H. Scrivens, S.C. Oliveira, A. Miyoshi, C.G. Dowson, V. Azevedo, *BMC Microbiol.* 2011 17;11(1):12. DOI:10.1186/1471-2180-11-12.
- A. Ravagnani, C. L. Finan, M. Young, *BMC Gen.* 6 (2005) 39. DOI:10.1186/1471-2164-6-39.
- E.C. Hett, M.C. Chao, L.L. Deng, E.J. Rubin, *PLoS Pathog.* 4 (2008):e1000001. DOI: 10.1371/ journal.ppat.1000001.
- V. Anantharaman, L. Aravind, *Genome Biol.* 4 (2003) R11. DOI:10.1186/gb-2003-4-2-r11.
- M. Kuhn, W. Goebel, *Infect. Immun.* 57 (1989) 55-61.
- T.J. Smith, S.A. Blackman, S.J. Foster, *Microbiol.* 146 (2000) 249–262.
- J.M. Aramini, P. Rossi, Y.J. Huang, L. Zhao, M. Jiang, M. Maglaqui, R. Xiao, J. Locke, R. Nair, B. Rost, T.B. Acton, M. Inouye, G.T. Montelione, *Biochem.* 47 (2008) 9715-9717. DOI: 10.1021/bi8010779.
- M.H. Jr. Saier, I. T. Paulsen, *Semin. Cell Dev. Biol.* 12 (2001) 205-213. DOI:10.1006/scdb.2000.0246.
- T.T. Tseng, K.S. Gratwick, J. Kollman, D. Park, D.H. Nies, A. Goffeau, M.H. Jr. Saier, *J. Mol. Microbiol. Biotechnol.* 1(1999) 107-125.
- M. Sara, U.B. Sleytr, *J. Bacteriol.* 169 (1987) 4092-4098.
- S. Miyoshi, S. Shinoda, *Microbes Infect.* 2, (2000) 91–98. DOI:10.1016/S1286-4579(00)00280-X.
- J.C. Olson, D.E. Ohman, *J Bacteriol.* 174 (1992) 4140–4147.
- F. Tonello, S. Morante, O. Rossetto, G. Schiavo, C. Montecucco, *Adv Exp Med Biol.* 389 (1996) 251–260.
- P. J. Hanna, *Appl Microbiol.* 87 (1999) 285–287.
- J. Raveneau, C. Geoffroy, J. L. Beretti, J. L. Gaillard, J. E. Alouf, P. Berche, *Infect Immun.* 60 (1992) 916–921.



ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v1i2.56

## The effect of $Zn^{2+}$ on prostatic cell cytotoxicity caused by *Trichomonas vaginalis*

Laura Isabel Vazquez-Carrillo<sup>1</sup>, Laura Itzel Quintas-Granados<sup>1</sup>, Rossana Arroyo<sup>2</sup>, Guillermo Mendoza Hernández<sup>3</sup>, Arturo González-Robles<sup>2</sup>, Bertha Isabel Carvajal-Gamez<sup>1</sup>, M. Elizabeth Álvarez-Sánchez<sup>1</sup>

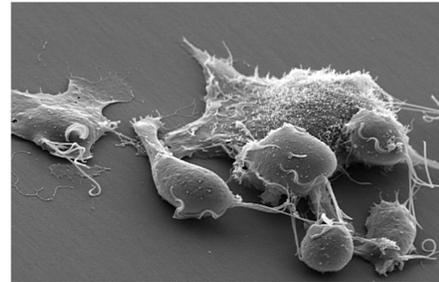
<sup>1</sup>Posgrado en Ciencias Genómicas, Universidad Autónoma de la Ciudad de México, plantel del Valle, San Lorenzo #290 Col. Del Valle, CP 03100 México DF, México. Telephone: +52 36912000 x 15306. Fax number: +52 55755805; <sup>2</sup>Departamento de Infectómica y Patogénesis Molecular, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional; <sup>3</sup>Departamento de Bioquímica, Facultad de Medicina, Universidad Nacional Autónoma de México, México, D.F., México.

Received: 1 December 2010 Accepted: 3 February 2011 Available Online: 15 February 2011

### ABSTRACT

Our investigation focused on the study of the proteome, morphology, and cytotoxicity of *T. vaginalis* during interactions with prostatic DU-145 cells. The results suggest that approximately 37 different proteins are expressed in the presence of  $Zn^{2+}$ , which also down-regulates the protein and transcriptional levels of TvCP65. The result is a negative effect on trichomonal cytotoxicity. The differentially expressed proteins were identified by mass spectrometry analysis.

**Keywords:** *Trichomonas vaginalis*; DU-145 prostatic cells; morphology;  $Zn^{2+}$  proteome; cytotoxicity; TvCP65.



### 1. Introduction

*Trichomonas vaginalis* is the causative agent of trichomoniasis, a common worldwide infection associated with important public health problems [1] and implicated as a cofactor in the transmission of the human immunodeficiency virus [2]. This infection causes infertility [3], low birth weight infants and preterm delivery [4], and a predisposition to cervical neoplasia [5]. Infections in women cause vaginitis, cervicitis, urethritis, a malodorous seropurulent vaginal discharge and infertility [3, 6]. The cytopathogenicity of *T. vaginalis* begins with cytoadherence to target cells, which is mediated by ligand-receptor type interactions [7-8]. After adherence to vaginal epithelial cells (VECs), *T. vaginalis* suffers a dramatic change from ellipsoid to amoeboid [9]. Five adhesins (AP65, AP51, AP33, AP23, and AP120) [10-12] and two cysteine proteinases (CPs) (TvCP30 and TvCP62) [1, 13-15] participate in trichomonal adherence. Following adherence, a cytotoxic effect on the VECs cells, due to two CPs

(TvCP65, 65 kDa, [16-17] and TvCP39, 39 kDa, [18-19]), is observed before the target cells are phagocytized by the parasite [20]. In addition, other processes are involved in host colonization and target cell damage, such as hemolysis [21], immune evasion [22-23], signal transduction [9], and apoptosis induction [24].

Studies about female trichomoniasis indicate that iron ( $Fe^{2+}$ ), which is a component of the vaginal microenvironment, is also an essential element up-regulating numerous trichomonad genes. The  $Fe^{2+}$  concentration in the female microenvironment changes during the menstrual cycle, which modulates the expression of crucial metabolic enzymes and several pathogenic properties, and also leads to the multiplication and cytoadherence of *T. vaginalis* [25]. Not only is adhesin synthesis positively regulated by  $Fe^{2+}$  [26], but some proteinases involved in C3b complement component degradation [22] and immune evasion through the regulation of

\*Corresponding author: Elizabeth Álvarez-Sánchez. Email Address: elizabethalvarezsanchez@yahoo.com.mx.

P270 phosphorylation [27] are also regulated by this cation. Furthermore,  $\text{Fe}^{2+}$  down-regulates proteinases involved in trichomonal cytotoxicity, TvCP65 and TvCP39, which also participate in the interaction between the parasite and host, e.g., HeLa cells and VECs [1, 18]. In particular, TvCP65 degrades collagen IV and fibronectin in a broad pH range [16]. In contrast, TvCP39 degrades collagens (I, III, IV, and V), fibronectin and hemoglobin, and human IgA and IgG [18]. Interestingly, the proteolytic activities of TvCP39 [19] and TvCP65 [28] are negatively regulated by  $\text{Fe}^{2+}$ , the transcript and protein expression levels of TvCP65 agree with this decreased proteolytic activity [28].

Although trichomoniasis affects men, most are asymptomatic [29]. However, mild cases of urethritis, chronic prostatitis, and epididymitis have been associated with this infection [30]. *T. vaginalis* has several consequences in men's health, such as reduced sperm cell motility due to cell surface interactions between the parasite and the head or tail of the sperm as well as due to the flagella of parasites [31], which leads to male infertility [32]. During male infections, *T. vaginalis* accesses the microenvironment of the prostate, which is the gland that surrounds the posterior urethra of a man and, along with the seminal vesicles, produces prostatic and seminal fluids. In this environment,  $\text{Zn}^{2+}$  is an important component of prostatic fluids, is present at high concentration (4.5 to 7 mM), and has been found to be critical in spermatogenesis [33]. Prostatic secretions are primary components of the antimicrobial defenses of the male genitourinary tract [32].  $\text{Zn}^{2+}$ , in particular, is principal to prostate defense due to its ability to prevent pathogen establishment in the male reproductive tract [33] and has, indeed, been found to have an antimicrobial spectrum towards bacteria, viruses, chlamydiae, and fungi [34]. This antimicrobial effect also affects *T. vaginalis*, which is sensitive to relatively low concentrations of  $\text{Zn}^{2+}$  chloride and sulfate, giving a minimal trichomonocidal concentrations (MTC) of 1.6 mM for both [33]. Nevertheless, a  $\text{Zn}^{2+}$  concentration of about 0.8 mM, obtained in patients with chronic bacterial prostatitis, is not trichomonocidal [35]. In these cases, *T. vaginalis* may persist for longer periods of time in the male genitourinary tract, and possibly progress to the prostate. *T. vaginalis* has been observed to infect the prostatic epithelium, cause chronic prostatitis, and elicit an inflammatory immune response [36]. Thus, *T. vaginalis* has been observed in prostate tissue near inflamed areas and epithelial hyperplasia, which suggests that *T. vaginalis* might be involved in prostate carcinogenesis [37].

Furthermore,  $\text{Zn}^{2+}$  plays an important role in the metabolism of parasites. It interferes in the function of the hydrogenosome, an organelle involved in the metabolism of pyruvate. The hydrogenosome is the main site of the initial  $\text{Zn}^{2+}$  effect in the pathogenic protozoan *Tritrichomonas foetus* [38].

The purpose of the present study was to investigate the morphology of *T. vaginalis* and the proteinases involved in its interaction with DU-145 prostatic cells and to examine the  $\text{Zn}^{2+}$ -dependent changes on the cytotoxicity and protein pro-

file of *T. vaginalis*. The  $\text{Zn}^{2+}$  effect on the protein and transcript levels of TvCP65 and the identification of differentially expressed proteins were of particular interest.

## 2. Material and methods

### 2.1 *T. vaginalis* and cells culture.

Trophozoites of *T. vaginalis* isolate CNCD 147 were axenically cultivated for 24 h in trypticase-yeast extract-maltose (TYM) medium pH 6.2 with 10% heat-inactivated horse serum (Gibco) and supplemented with or without 0.25 mM, 1.0 mM, and 1.6 mM  $\text{ZnCl}_2$  (Sigma). Samples were taken at several time points up to 24 h for parasite counting using a Neubauer-counting chamber, and viability was measured by the trypan blue exclusion method [28]. Cytotoxicity assays and proteomic profiles were performed with parasites cultivated under the same conditions.

Immortalized HeLa ATCC cells were grown in Dulbecco's Modified Eagle's Medium (DMEM; Life Technologies) supplemented with 10% horse serum and 10% penicillin-streptomycin (Gibco) at 37 °C in a 5%  $\text{CO}_2$  atmosphere for 24 h until confluent cell monolayers were obtained. DU-145 ATCC cells were grown in low glucose DMEM (pH 7.2) supplemented with 10% fetal bovine serum and 10% penicillin-streptomycin (Gibco) at 37 °C in a 5%  $\text{CO}_2$  atmosphere for 24 h until confluent cell monolayers were obtained.

### 2.2 2D gel electrophoresis.

For proteomic maps, we used a previously reported protocol [39] with modifications. *T. vaginalis* ( $1.2 \times 10^8$ ) grown in the presence or absence of 0.25 mM, 1.0 mM, and 1.6 mM  $\text{Zn}^{2+}$  were collected by centrifugation at 900 g for 5 min at 4 °C and washed three times with PBS pH 7.0. For the first dimension, parasites were lysed in a final volume of 200  $\mu\text{l}$  rehydration solution (7 M urea, 4% CHAPS, 70 mM DTT, 2% IPG buffer pH 4–7, trace bromophenol blue; Bio-Rad). The supernatant was centrifuged at 13000 g for 10 min at 4 °C to remove insoluble material, and samples of 120  $\mu\text{l}$  (corresponding to  $7.2 \times 10^7$  parasites) were applied to an IPG strip (7 cm, pH 4–7 linear; Bio-Rad) for passive rehydration for 12 h. All isoelectric focusing took place on a Protean IEF system (Bio-Rad) as follows: step 1—gradient from 1 to 225 V over 35 min; step 2—gradient from 250 to 2479 V over 185 min; step 3—gradient from 2466 to 15,434 V over 210 min. Before the second dimension, proteins were reduced (10 mg/ml DTT) and alkylated (25 mg/ml iodoacetamide) step-wise, 15 min for each step, in equilibration buffer (6 M urea, 2% SDS, 300 mM Tris-Cl pH 8.8, 20% glycerol, and 0.002% bromophenol blue) at room temperature. Equilibrated IPG strips were separated on 12% SDS-PAGE gels (7 cm  $\times$  8 cm  $\times$  1.0 mm) using the MiniProtean II Cell vertical system (Bio-Rad) and standard Tris/glycine/SDS buffer. Gels were run at 35 mA/gel at room temperature until the tracking dye left the gel and stained with Coomassie Brilliant Blue G-250 following procedures described by the manufacturer. Finally, gels were documented using Gel Doc EQ (Bio-Rad). Image analysis was performed using the pDQuest software (Bio-Rad). Three

independent protein preparations, each obtained from an independent parasite culture, were performed for comparisons of 2-DE maps. Differentially expressed proteins were determined by a tridimensional analysis using Melanie 7 Software and identified by tandem mass spectrometry analysis (MS/MS).

### 2.3 LC-ESI-MS/MS.

The MS/MS analysis of each fraction obtained from the offline separation steps was carried out on a 3200 Q TRAP hybrid tandem mass spectrometer (Applied Biosystems/MDS Sciex, Concord, ON, Canada) equipped with a nanoelectrospray ion source (NanoSpray II). The instrument was coupled online to a nanoAcquity Ultra Performance LC system (Waters Corporations, Milford, MA). Samples were desalted by injection onto a Symmetry C18 UPLC trapping capillary column (180  $\mu\text{m} \times 20$  mm, Waters Corporations) and washed with 0.1% formic acid in 100% MilliQ water at a flow rate of 15  $\mu\text{l}/\text{min}$ . After 3 min, the trap column was switched in line with a capillary analytical column. Peptides were separated on an ethylene-bridged hybrid, C18 UPL column (75  $\mu\text{m} \times 100$  mm, Waters Corporations) using a linear gradient of 2-70% acetonitrile, 0.1% formic acid over a 60 min period at a flow rate of 0.25  $\mu\text{l}/\text{min}$ . Spectra were acquired in automated mode using Information Dependent Acquisition (IDA). Precursor ions were selected in Q1 using the enhanced MS mode with a scan range of  $m/z$  400-1500 and 4000  $\text{amu}/\text{s}$ . Selected ions were subjected to an enhanced resolution scan at a low speed of 250  $\text{amu}/\text{s}$  over a narrow (30  $\text{amu}$ ) mass range and, then, to an enhanced product ion scan (MS/MS). Precursor ions were fragmented by collision activated dissociation (CAD) in a Q2 collision cell using rolling collision energy. The generated fragment ions were captured and mass analyzed in a Q3 linear ion trap.

### 2.4 Database search.

For protein identifications, the molecular mass of each tryptic fragment was used as query at the National Center for Biotechnology nonredundant database using the MASCOT search algorithm (Matrix Science, London, UK, available at <http://www.matrixscience.com>). Monoisotopic mass values, mass tolerance (peptide  $\pm 1.2$  Da and fragment  $\pm 0.6$  Da), possible methionine residue oxidation, and deamidation and carbamidomethylation at cysteine residues were considered as variable modifications. A maximum of one missed tryptic cleavage per protein was allowed, and no taxonomic restrictions were considered in the database search. For positive identifications, MASCOT individual ion scores  $>56$  indicated identity or extensive homology ( $p < 0.05$ ).

### 2.5 Cell-binding assay.

The cell-binding assay to detect proteinases with affinity to the host cell surface was performed as previously described [15]. Briefly, a clarified detergent extract from  $2 \times 10^7$  parasites, which was grown in the presence or absence of 0.25 mM, 1.0 mM, or 1.6 mM  $\text{Zn}^{2+}$ , was incubated for 18 h at 4  $^{\circ}\text{C}$  with  $1 \times 10^6$  fixed HeLa or DU-145 cells. Then trichomonad

proteinases bound to the surface of fixed cells were eluted in Laemmli buffer [40] for 20 min at 37  $^{\circ}\text{C}$ . Released proteinases were loaded onto a 10% SDS-PAGE gel copolymerized with 2% gelatin and run at 35 mA/gel. Gels were washed with 10% Triton X-100 for 10 min with gentle agitation. Finally, proteinase activation was performed in 100 mM sodium acetate buffer pH 4.5 with 0.1%  $\beta$ -mercaptoethanol for 18 h at 4  $^{\circ}\text{C}$ , 25  $^{\circ}\text{C}$ , 37  $^{\circ}\text{C}$ , and 43.5  $^{\circ}\text{C}$ . The gels were further stained with Coomassie Brilliant Blue for a visualization in which clear bands against a dark background indicate proteolytic activity. Densitometry analyses of activity bands were performed in triplicate using the software Quantity One ver. 4.6.3 (Bio-Rad).

### 2.6 Cytotoxicity assay.

The cytotoxicity assay was carried out using confluent HeLa and DU-145 cells monolayers in 48-well microtiter plates with  $3.5 \times 10^4$  HeLa or  $5.5 \times 10^4$  DU-145 cells/well, respectively. Briefly, parasites ( $2.75 \times 10^5$ ) grown with or without 0.25 mM, 1.0 mM, and 1.6 mM  $\text{Zn}^{2+}$  were resuspended in TYM-DMEM medium without serum, added to confluent cell monolayers of HeLa or DU-145 cells at a ratio of 5:1 (parasites:host cell), and incubated for several time points, up to 24 h, at 37  $^{\circ}\text{C}$  under a 5%  $\text{CO}_2$  atmosphere. Monolayer destruction was assessed using a colorimetric method and quantitated spectrophotometrically at 570 nm [14, 16, 28]. Each sample was performed in triplicate, and experiments were performed at least twice with similar results.

### 2.7 Western blot assay.

Total protein extract from parasites ( $2 \times 10^7$ ), which were grown in the presence or absence of 1.6 mM  $\text{Zn}^{2+}$ , were obtained by TCA-precipitation as previously described [16, 28, 41]. Solubilized proteins were resuspended in Laemmli buffer [40], boiled, and loaded onto a 10% polyacrylamide gel with an equivalent of  $4 \times 10^5$  parasites/lane. Protein extracts were blotted onto nitrocellulose membranes and blocked with 5% skim milk in PBS (pH 7.0) for 18 h at 4  $^{\circ}\text{C}$ . Membranes were incubated for 18 h at 4  $^{\circ}\text{C}$  with anti-TvCP65 primary antibody as reported [16], and an anti- $\alpha$ -tubulin monoclonal antibody (1:1000 dilution, Invitrogen) was used as a loading control. Then the blotted membrane was washed five times with a PBS pH 7.0-0.1% Tween 20 buffer. The primary antibody was detected with a secondary goat anti-mouse-IgG (H+L) horseradish peroxidase conjugate (1:3000 dilution, Invitrogen). The membrane was washed with PBS pH 7.0-0.1% Tween 20 and visualized using an enhanced chemiluminescence ECL Plus Western Blotting Detection System (GE Healthcare) according to the manufacturer's instructions.

### 2.8 RNA extraction and cDNA synthesis.

A sample containing  $2 \times 10^7$  parasites grown with or without 0.25 mM, 1.0 mM, and 1.6 mM of  $\text{Zn}^{2+}$  was collected by centrifugation at 900  $g$  for 5 min at 4  $^{\circ}\text{C}$  (Allegra<sup>TM</sup> X-22 Centrifuge, Beckman Coulter). The pellet was suspended in 1 ml of TRIzol<sup>®</sup> reagent (Invitrogen, Life Technologies, Carlsbad,

CA), and the total RNA was extracted as recommended by the manufacturer. RNA concentration was determined spectrophotometrically using a NanoDrop 2000 (Thermo Scientific). All 260/280 ratios were between 1.8 and 2.1. Finally, 1 µg of total RNA was reverse-transcribed using the Superscript Reverse Transcriptase Kit (Invitrogen) and the oligo-dT (dT<sub>18</sub>) (10 pmol/µl) primer.

2.9 Analysis of gene expression by semi-quantitative RT-PCR.

PCR was performed in 50 µl reactions containing 50 ng of cDNA, 10 pmol of each primer pair, and 0.25 U of Taq DNA polymerase (Invitrogen). PCR was carried out in a GeneAmp PCR System 9700 thermal cycler (Applied Biosystems). We used the following primer pairs to amplify: 370 bp of the *tvcp65* gene (accession number AY463696), forward: 5'-ACGCGATTACATCTGGAGAACTC-3', and reverse: 5'-ATAAGAGGAGCGTGATGGCACAT-3'; and 112 bp of the *β-tubulin* gene as reported [42]. The amplified products were analyzed on 2% agarose gels and visualized by ethidium bromide staining. Gene expression densitometry analyses were performed using the Quantity One Software (Bio-Rad). Data from densitometry quantification of the housekeeping gene (*β-tubulin*) were used to normalize the results.

2.10 Real-time qRT-PCR analysis of specific *T. vaginalis tvcp65* and *β-tubulin* genes.

Quantification of *tvcp65* expression was performed by real-time qRT-PCR. Oligonucleotide primers for real-time qRT-PCR were designed using Primer3 version 3.0 (www.primer3.sourceforge.net) and commercially synthesized (Instituto de Biotecnología, UNAM). To specific amplify 100 bp of the *tvcp65* gene, the forward: 5'-AATGTTGTTGAAGGCGATGAAA-3', and reverse: 5'-CTACAGCAGCTGGGCCATTT-3' were used, and previously reported primers were used to amplify 112 bp of the *β-tubulin* gene [42]. Each reaction was carried out in a total volume of 25 µl with 1 µl of cDNA (50 ng), 12.5 µl Maxima™ SYBR Green/ROX qPCR Master Mix (2×) kit (Fermentas), 1 µl of each primer (10 pmol/µl), and 9.5 µl of molecular biology grade water. qRT-PCR was performed using a 7500 Fast Real-Time PCR machine (Applied Biosystems). PCR conditions for *tvcp65* were as follows: 15 min at 95°C followed by up to 40 cycles of 95°C for 30 s, 57°C for 30 s, and 72°C for 30 s, and, finally, 95°C for 15 s, 60°C for 30 s, and 95°C for 15 s. The “housekeeping” *β-tubulin* gene was selected for a reference gene study because it did not vary in the presence of any of the tested Zn<sup>2+</sup> concentrations. Standardization of *tvcp65* messenger RNA (mRNA) expression was performed by dividing the value of each gene at different Zn<sup>2+</sup> concentrations by the value of the housekeeping gene found for all sample.

2.11 Indirect immunofluorescence assays.

Parasites were fixed using 4% paraformaldehyde for 1 h at 37°C and washed with PBS pH 7.0. Half of the fixed parasites were permeabilized using 1 M HCl for 2 h at room temperature. Parasites were then blocked with 0.2 M glycine for 1 h at

37°C followed by 0.2% fetal bovine serum for 15 min. Then trichomonads were incubated with polyclonal mouse anti-TvCP65 antibody (1:100 dilution) [16] or preimmune sera (PI) for 18 h at 4°C. Next, parasites were incubated with fluorescein isothiocyanate-conjugated anti-mouse immunoglobulins (1:90 dilution, Jackson ImmunoResearch) for 40 min at room temperature. Finally, Vectashield-DAPI mounting solution (Vector Lab) was added, and laser confocal microscopy was performed (Leica, DMLS).

2.12 Scanning electron microscopy.

Parasites cultivated in TYM media for 24 h at 37°C were incubated for different periods of time with DU-145 cell monolayers. Glutaraldehyde-fixed samples were dehydrated with increasing concentrations of ethanol, critical-point dried using a Samdri 780 apparatus (Tousimis, Rockville Maryland, USA), coated with gold using a JEOL JFC-1100 ion-sputtering device, and examined using a XL-30 ESEM scanning electron microscope (FEI Company, Eindhoven, The Netherlands).

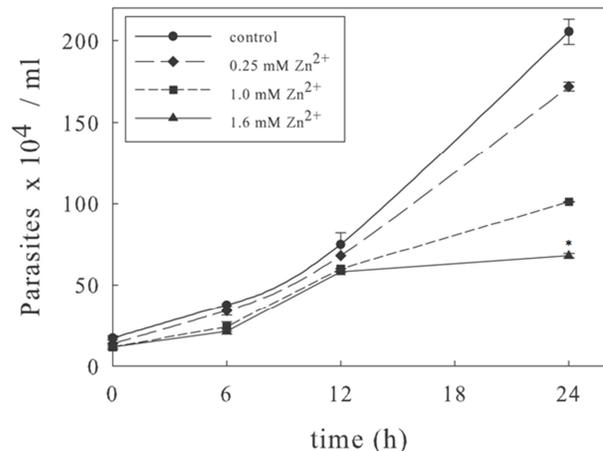
2.13 Statistical analysis.

All data are the means ± standard error of triplicate samples. A statistically significant change between means was determined by ANOVA (*p*) using Sigma Plot 11.0 Software.

3. Results

3.1 Growth kinetics of *T. vaginalis*.

To investigate the effects of Zn<sup>2+</sup> on *T. vaginalis* development, we first performed growth kinetics in the presence or absence of 0.25 mM, 1.0 mM, and 1.6 mM Zn<sup>2+</sup> (Fig. 1). After



**Figure 1.** Effect of Zn<sup>2+</sup> on the growth of *T. vaginalis*. Parasites were counted in a Neubauer-counting chamber at different time points (6, 12, and 24 h) after incubation at 37 °C in TYM medium supplemented with (0.25 mM, 1.0 mM, and 1.6 mM) or without Zn<sup>2+</sup> (control). In addition, parasite viability, as measured by trypan blue exclusion, was about 98% under all conditions. Each point indicates the mean of three experiments, and error bars indicate standard deviation. There were no statistically significant differences for data corresponding to 0.25 mM (*p*=0.970) and 1.0 mM Zn<sup>2+</sup> (*p*=0.960) compared to the control. In contrast a statistically significant difference were found for 1.6 mM Zn<sup>2+</sup> data (*p*=0.01) (\*) compared to the control.

24 h, the parasites cultivated without  $Zn^{2+}$  (control) had four duplications, while the parasites grown with 0.25 mM, 1.0 mM, or 1.6 mM  $Zn^{2+}$  had three and a half, three, and one duplication, respectively. Furthermore, viability for all samples was ~98%. Experiments were performed in triplicate with the same results. These findings indicate that  $Zn^{2+}$  had no effect on *T. vaginalis* viability, and trichomonad growth was only diminished in 1.6 mM  $Zn^{2+}$ .

### 3.2 *T. vaginalis* $Zn^{2+}$ proteomic map.

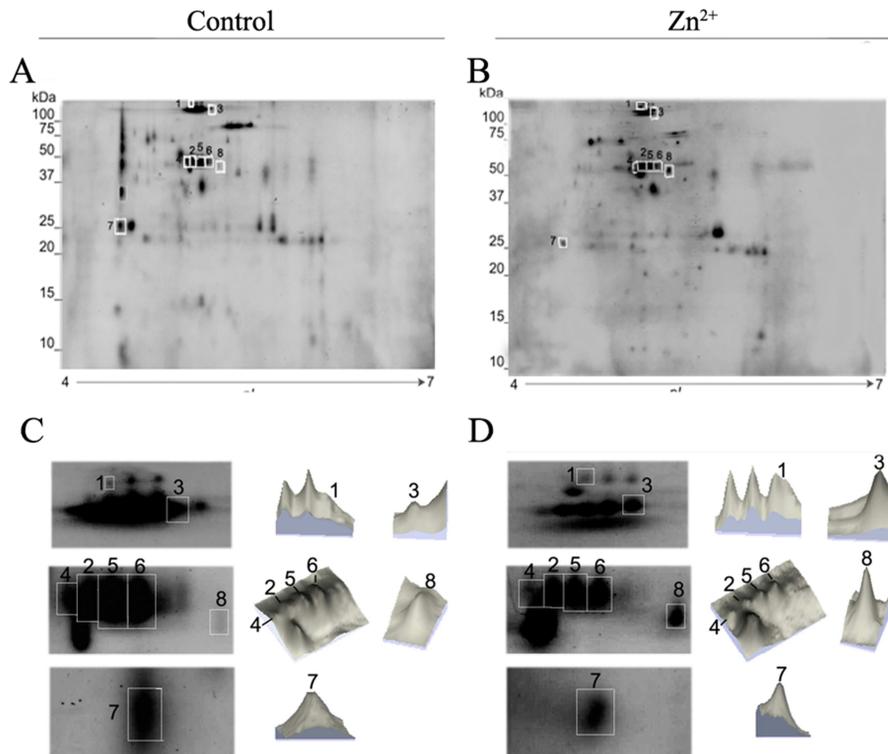
Images of three representative 2-DE gels obtained from three independent experiments using parasites grown in the presence and absence of 0.25 mM, 1.0 mM, and 1.6 mM  $Zn^{2+}$  were analyzed by pDQuest software (data not shown), and it was found that most of the protein profile changes were obtained with 1.6 mM  $Zn^{2+}$ . Therefore, we selected this specific concentration for further analyses. The 2-DE protein spot profiles were highly reproducible in terms of both the total number of protein spots and their relative positions and intensities. Coomassie Brilliant Blue G-250 stained gels showed approximately 150 spots in the presence of 1.6 mM  $Zn^{2+}$  (Fig. 2B) compared to 177 spots in its absence (Fig. 2A). Proteins detected in pI 4–7 had molecular masses between 20 and 250 kDa. The gels showed remarkable changes in protein profiles for parasites grown in the presence of  $Zn^{2+}$ . Indeed, we found at least 27 proteins that were differentially expressed in the

presence of  $Zn^{2+}$ . The major differences in expression protein profiles were obtained from 50 to 100 kDa.

Interestingly, we found some spots over-expressed in the presence of  $Zn^{2+}$  (Figs. 2C and 2D, spots 1, 3, 8) that correspond to fimbrin (gi|123493533), adenosylhomocysteinase (gi|123499896), and aminopeptidase P-like metalloproteinase (gi|123445672) (Table 1). Otherwise, the expression levels of several spots were diminished in the presence of  $Zn^{2+}$  (Fig. 2C and 2D). Spots 2, 4, 5, 6, and 7 correspond to fimbrin (gi|123397260) (spots 2 and 4, probably due to degradation), two adenosylhomocysteinases (gi|123499896 and gi|123488577), and asparaginyl endopeptidase-like cysteine peptidase (gi|123408789) (Table 1). These findings suggest that  $Zn^{2+}$  had both positive and negative effects on the expression levels of ~27 proteins and no effect on 111 proteins. Interestingly, the asparaginyl endopeptidase-like cysteine peptidase (gi|123408789) and the adenosylhomocysteinase (gi|123499896) had already been identified in the active degradome of *T. vaginalis* [39].

### 3.3 Proteinases involved in the interaction of *T. vaginalis* and prostatic cells.

We performed ligand-protease assays using parasites grown in the presence or absence of 0.25 mM, 1.0 mM, and 1.6 mM  $Zn^{2+}$  to detect possible *T. vaginalis* proteinases interactions with DU-145 cells. We observed differences between



**Figure 2.**  $Zn^{2+}$  proteomic map of *T. vaginalis*. The total proteins of *T. vaginalis* grown in the A) absence or B) presence of 1.6 mM  $Zn^{2+}$  were separated in the first dimension by isoelectric focusing over a pH range from 4.0 to 7.0 followed by 12% SDS-polyacrylamide gel electrophoresis. Protein spots were Coomassie Brilliant Blue G-250 stained. Enlarged images of the gels shown in A or B and three-dimensional analyses of 8 differentially expressed proteins in parasites cultivated in the absence C) or presence D) of  $Zn^{2+}$ . The left panel shows a representative image of a Coomassie Brilliant Blue G-250 gel, whereas the right panel shows a landscape representation of spots. Numbers 1, 3, and 8 indicate spots that are over-expressed in the presence of  $Zn^{2+}$ , while numbers 2, 4, 5, 6, and 7 indicate spots that are down-regulated in the presence of  $Zn^{2+}$ .

the proteinases that interact with DU-145 cells in comparison with those that interact with HeLa cells. [16]. Fig. 3A shows the activities of all proteinases from protein extracts of *T. vaginalis* grown in the absence (Fig. 3A, lane 1) or presence (lane 2) of 1.6 mM Zn<sup>2+</sup>. These extracts interacted with fixed DU-145 cells, and the zymograms show the activities of at least five trichomonad proteinases (TvCP70, TvCP65, TvCP39, TvCP25, and TvCP20) bound to the surfaces of fixed DU-145 cells in the absence (Fig. 3B, lane 1) or presence (lane 2) of 1.6 mM Zn<sup>2+</sup>. Interestingly, the proteolytic activity profile of proteinases that interact with prostatic cells was similar at all temperatures analyzed (data not shown), which suggests that these proteinases may be active up to 43.5 °C. Nevertheless, the best resolution of proteinase activity bands was obtained with activation at 4 °C (Fig. 3B). According to

the densitometry analyses of the TvCP70, TvCP65, TvCP39, TvCP25, and TvCP20 activity bands bound to the DU-145 cells (Fig. 3C), reductions of ~12%, 16%, 13%, 25%, and 11%, respectively, were observed when the parasites were grown in the presence of 1.6 mM Zn<sup>2+</sup> in comparison to the activity bands of untreated control parasites (taken as 100%). Interestingly, the activities of CPs involved in cytotoxicity (TvCP65 and TvCP39) were reduced in the presence of Zn<sup>2+</sup>, however this reduction was minimal. This result may be related to the reduction in trichomonal cytotoxicity observed in the presence of this cation. We also performed Western blot and indirect immunofluorescence assays and real-time PCR experiments to corroborate if the expression levels of these CPs were affected by Zn<sup>2+</sup>.

**Table 1.** Zn<sup>2+</sup> differentially expressed proteins identified by mass spectrometry from clinical isolate *T.vaginalis* CNCD 147 were identified by searching the National Center for Biotechnology non-redundant database using the MASCOT search algorithm.

Protein	Accession number <sup>a</sup>	Spot number <sup>b</sup>	MASCOT score <sup>c</sup>	Ion score <sup>d</sup>	Peptide sequence <sup>e</sup>	% Coverage <sup>f</sup>	Predicted MW Da /pI						
Fimbrin	gi 123493533	1	120	49	K.LSPEQILLR.W	5	67,839 /5.17						
				32	R.LLKPGETLADLLK.L								
				39	K.TVNLTNHPELFR.L								
Fimbrin	gi 123397260	2	125	41	K.LSPEQILLR.W	7	50,761 /6.43						
				43	R.LLKPGETLADLLK.L								
				41	K.TVNLTNHPELFR.L								
				37	R.ALTWPDGK.G								
				37	R.ASDVMIGGK.T								
				58	K.HLDEEVAR.L								
				48	R.HSLIDGINR.A								
				57	K.EMPGLMVLRE								
				22	K.EMPGLMVLRE								
				52	K.SPEGAPFEYR.I								
				49	K.TALVMGYGDVGK.G			30	54,064 /5.61				
				23	R.IADINLHVLGR.K								
				55	R.IADINLHVLGR.K								
39	R.LHLGSLDVHLTK.L												
56	K.LLFPAINVNDAVTK.S												
26	K.GFEFENAGAVDPQK.G												
25	K.GDNYEYTCVLAVLK.Q												
Adenosylhomocysteinase	gi 123499896	4	402	43	K.GPQQIVDDGGDATLLIQK.G	22	54,064 /5.61						
				31	K.VYTLPK.H								
				44	K.HLDEEVAR.L								
				32	K.SPEGAPFEYR.I								
				69	K.TALVMGYGDVGK.G								
				66	R.IADINLHVLGR.K								
				31	R.LHLGSLDVHLTK.L								
				65	K.LLFPAINVNDAVTK.S								
				27	K.EGTPEKPAGIPVFAWK.G								
				39	K.GPQQIVDDGGDATLLIQK.G								
				Adenosylhomocysteinase	gi 123499896			5	576	36	R.ALTWPDGK.G	33	54,064 /5.61
										53	K.HLDEEVAR.L		
										32	R.HSLIDGINR.A		
40	K.FDNIYGCR.H												
60	K.SPEGAPFEYR.I												
31	K.TALVMGYGDVGK.G												
59	R.IADINLHVLGR.K												
66	R.IADINLHVLGR.K												
24	R.LHLGSLDVHLTK.L												
54	R.LHLGSLDVHLTK.L												
62	K.LLFPAINVNDAVTK.S												
28	K.QADYINVPVEGPK.S												
45	K.EGTPEKPAGIPVFAWK.G												
39	K.AIVGNIGHFDNEIDTEGLK.N												

				45	R.VIITEVDPICALQAAMEGYQVR.R		
				35	K.VYTLPK.H		
				26	K.KVYTLPK.H		
				45	R.ASDVMIGGK.T		
				39	R.ALTWPDGK.G		
				43	K.HLDEEVAR.L		
				24	R.HSLIDGINR.A		
				47	K.FDNIYGCR.H		
				30	K.EMPGLMVLRE		
				59	K.SPEGAPFEYR.I		
				53	K.TALVMGYGDVGK.G		
	6	850		69	R.IADINLHVLGR.K	48	54,064
				35	K.TALVMGYGDVGK.G		/5.61
				34	R.LHLGSLDVHLTK.L		
				73	K.LLFPAINVNDVTK.S		
				22	K.QADYINVPVEGPK.S		
				53	K.GFEFENAGAVDPQK.G		
				30	K.GETLPEYWENTYR.A		
				32	K.EGTPEKPAIPVFAWK.G		
				42	K.GPQQIVDDGGDATLLIQK.G		
				37	K.AIVGNIGHFDNEIDTEGLK.N		
				67	R.VIITEVDPICALQAAMEGYQVR.R		
				44	K.HLDEEVAR.L		
				37	R.HSLIDGINR.A		
				22	K.EMPGLMVLRE		
Adenosylhomocysteinease	gi 1234488577	4	295	34	K.SPAGAPFEYR.I	19	53,873
				60	R.IADINLHVLGR.K		/5.53
				29	R.LHLGSLDVHLTK.L		
				37	K.LLFPAINVNDVTK.S		
				36	R.VIITEVDPICALQAAMEGYQVR.R		
				42	K.VVAGVPK.L		
Clan CD, family C13, asparaginyl endopeptidase-like cysteine peptidase	gi 1234408789	7	198	57	K.VTATNFYK.V		
				38	R.SLDHLNVYPGR.A	14	43,916
				36	R.SLDHLNVYPGR.		/5.95
				36	K.QSHVMEYGDTSLK.T		
				28	K.IILMCYDDIVNDAENPFK.G		
				61	K.LADEVGGIRI		
Clan MG, family M24, aminopeptidase P-like metallo-peptidase	gi 123445672	8	222	38	K.EVTGVDNVK.F		
				37	R.NVYSILLEK.Q		
				23	K.TPYEIEQIK.K	13	50,416
				31	K.DVYGALDQIK.M		/5.34
				38	K.AAELTSEAIHVMK.N		
				39	K.AAELTSEAIHVMK.N		

<sup>a</sup> NCBI nr database.

<sup>b</sup> Spot number in agreement with Figs. 2A and 2B.

<sup>c,d</sup> Ion score is  $-10 \cdot \log(P)$ , where P is the probability that the observed match is a random event. Individual ion scores > 56 indicate identity or extensive homology ( $p < 0.05$ ). Protein scores are derived from ion scores as a non-probabilistic basis for ranking protein hits.

<sup>e</sup> M = methionine modified by oxidation; N = deamidated; C = cysteine modified by carbamidomethyl.

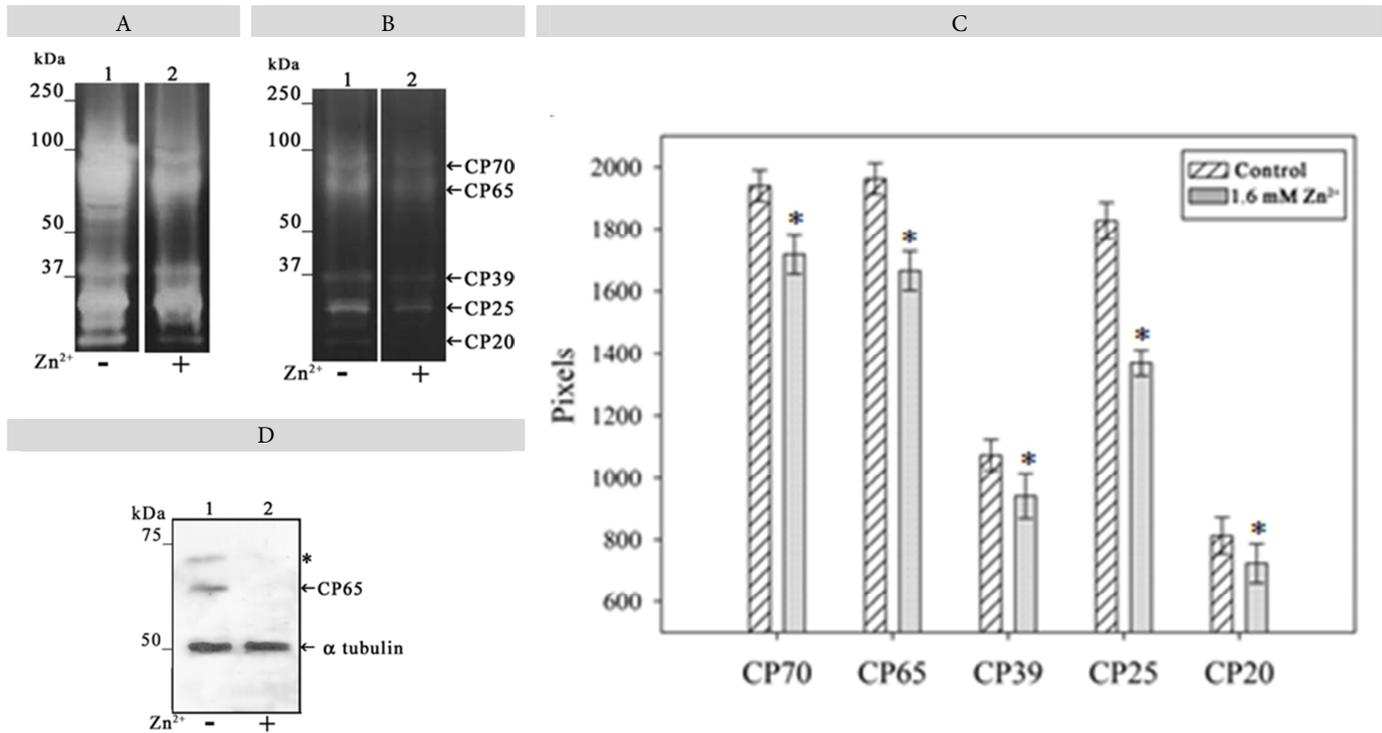
<sup>f</sup> %Coverage indicates the protein sequences or parts of these sequences that were identified by MS/MS.

Type of search:	MS/MS Ion search
Enzyme:	trypsin
Mass values:	Monoisotopic
Protein Mass:	Unrestricted
Peptide Mass Tolerance:	$\pm 1.2$ Da
Fragment Mass Tolerance:	$\pm 0.6$ Da
Max Missed Cleavages:	1
Instrument type:	ESI-4SECTOR

### 3.4 *T. vaginalis* interaction with DU-145 prostatic cells.

Using scanning electron microscopy, we evaluated the interaction between *T. vaginalis* and DU-145 prostatic cells at several incubation time points (Fig. 4A). The morphology of the parasites in contact with the DU-145 cell monolayer was similar to that of the parasites in contact with HeLa cells. However, their morphology was different in comparison to

those in contact with VEC cells [43]. We observed rounded parasite morphology but never a complete amoeboid transformation during interactions with DU-145 cells (Fig. 4A). After 5 min of incubation, *T. vaginalis* showed adherence to DU-145 cells (Fig. 4A, b). Parasites displayed an oval form with a rough surface, four anterior flagella, an undulating membrane, and posterior axostyle. At the 15 and 30 min time

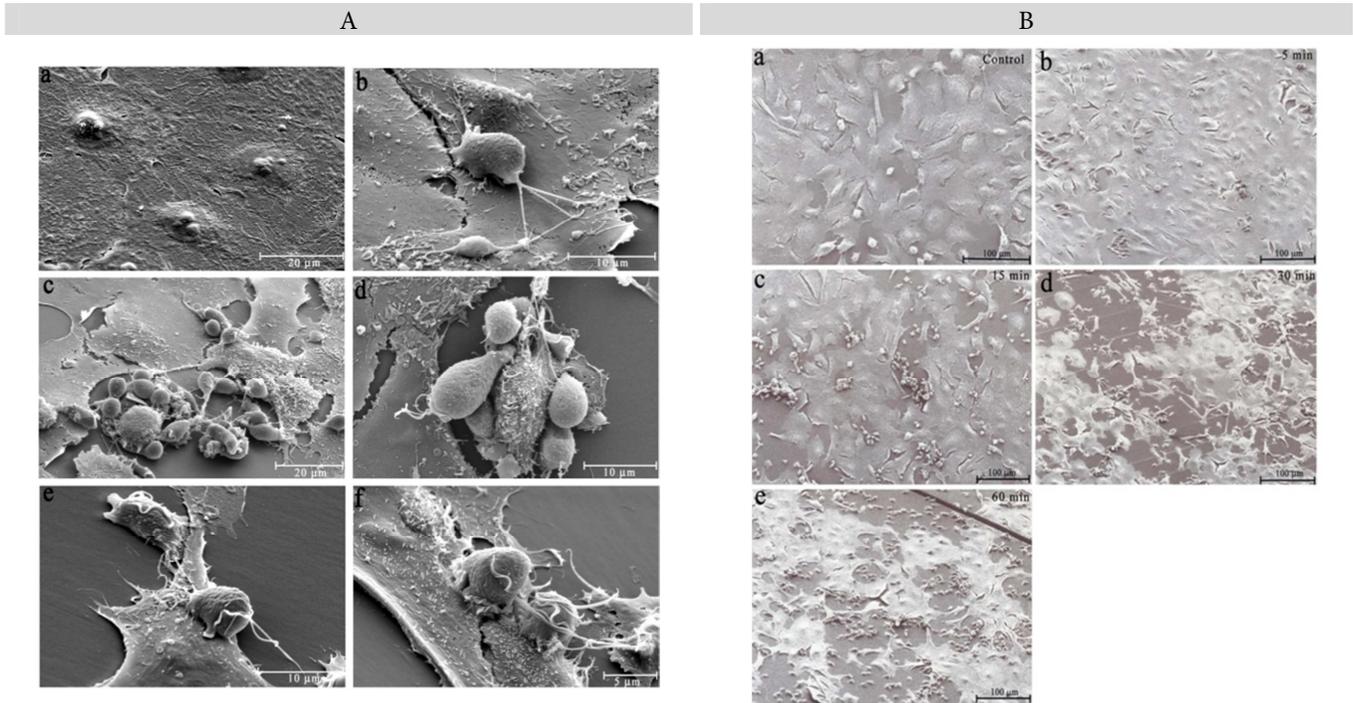


**Figure 3.** Proteinases involved in the interaction of *T. vaginalis* with prostatic cells. A) Zymograms correspond to trichomonad lysates from parasites cultivated in the absence (lane 1) or presence of 1.6 mM Zn<sup>2+</sup> (lane 2). B) Lysates from parasites cultivated in the presence or absence of Zn<sup>2+</sup> were used for ligand-proteinase assays and activated at 4 °C as described in the Materials and Methods section. Proteinase patterns correspond to the trichomonad proteinases with affinity for DU-145 cell surface from parasites cultivated in the absence (lane 1) or presence of 1.6 mM Zn<sup>2+</sup> (lane 2). Arrows show clear bands of the proteolytic activities of 70, 65, 39, 25 and 20 kDa (TvCP70, TvCP65, TvCP39, TvCP25, and TvCP20), respectively. C) Densitometry analyses of each activity band from three different gels activated at 4 °C. Bars indicate the mean intensity of activity bands, and error bars represent the standard deviations *p*=0.01 (\*). D) The Zn<sup>2+</sup> effect on the amount of TvCP65 and α-tubulin. Total protein extracts from *T. vaginalis* grown in the absence (lane 1) or presence of Zn<sup>2+</sup> (lane 2) were blotted onto nitrocellulose membranes for a Western blot assay with anti-TvCP65 and anti-α-tubulin (control) antibodies. Arrow indicates the immunodetected protein for each antibody employed. Asterisk indicates a higher sized band also recognized by the anti-TvCP65 antibody that might correspond to immunodetection of the TvCP65 precursor band.

points, we observed parasites in close proximities to the same target cells (Fig. 4A, c-d), which was probably due to a chemotaxis effect. After 60 min, we further observed empty areas that became progressively larger after 90 min, which corresponded to monolayer disruptions (Fig. 4A, e). After 90 min, we observed *T. vaginalis* pseudopods forming at the contact site (Fig. 4A, f). Interestingly, amoeboid trophozoites were not observed in contact with DU-145, however amoeboid trophozoites have been observed in contact with VECs [43]. In addition, the integrity and appearance of DU-145 prostatic cells without trichomonads did not change (Fig. 4A, a). We then kinetically analyzed the cytotoxicity of *T. vaginalis* towards DU-145 cells at several different time points (Fig. 4B). With scanning electron microscopy, we observed *T. vaginalis* attached to DU-145 cells after 5 min of incubation (Fig. 4B, b). After 15 min, the monolayer of DU-145 cells began to suffer damage, which might have been due to contact-dependent mechanisms (Fig. 4B, c). After 30 and 60 min, we observed irreversible cell damage due to the trichomonads (Fig. 4B, d, e). The no trichomonad control showed no changes in the appearance of DU-145 cells (Fig. 4B, a).

### 3.5 Cytotoxicity.

First, we determined the Zn<sup>2+</sup> effect on trichomonad cytotoxicity towards HeLa or DU-145 cells. After 1 h, the cytotoxicity exhibited by parasites, which were cultivated with 1.6 mM Zn<sup>2+</sup>, towards HeLa cells was 98% of that exhibited by parasites cultivated in absence of Zn<sup>2+</sup> (Fig. 5A). However, trichomonad cytotoxicity towards DU-145 cells was only 20% of what was seen with HeLa cells, and the presence or absence of Zn<sup>2+</sup> had no effect (Fig. 5A). These findings suggested that Zn<sup>2+</sup> had no effect on *T. vaginalis* cytotoxicity towards HeLa cells; however, trichomonad cytotoxicity towards DU-145 cells at this same time point was 80% less in comparison to HeLa cells (Fig. 5A). We then determined the cytotoxicity of *T. vaginalis* towards prostatic DU-145 cells at different time points and found 100% cytotoxicity after 8 h (Fig. 5B), which suggested trichomonad cytotoxicity was different for HeLa compared to DU-145 cells. In addition, we analyzed the effects of 0.25 mM, 1.0 mM, and 1.6 mM Zn<sup>2+</sup> on the cytotoxicity of DU-145 cells and found that major cytotoxic differences were obtained with 1.6 mM Zn<sup>2+</sup> (data not shown). The kinetic data show that cytotoxicity was 20% at 1



**Figure 4.** Interaction of *T. vaginalis* trophozoites with DU-145 prostatic cell monolayers. Morphological appearance of *T. vaginalis* in contact with DU-145 prostatic cells at different time points, and kinetic studies of the cytotoxicity of *T. vaginalis* towards DU-145 cells. A) Trophozoites were incubated with DU-145 cells for b) 5 min, c) 15 min, d) 30 min, e) 60 min, and f) 90 min. a) DU-145 prostatic cells before interacting with *T. vaginalis*. The majority of parasites attached to DU-145 cells retain a pear-like shape with four flagella, an undulating membrane, and axostyle. Some showed a few pseudopods (panel d-e). After 5 min of interaction, the lytic activity of the parasites is evident as seen by the disruption of the cell monolayer. B) Kinetics of *T. vaginalis* cytotoxicity towards DU-145 cell monolayers. Parasites were added to confluent DU-145 cell monolayers and observed under low magnification by scanning electron microscopy at different time points. Cells show a similar lytic activity at b) 5 min, c) 15 min, d) 30 min, and e) 60 min, as previously described. a) Control DU-145 cells without parasites.

h for parasites cultivated with or without 1.6 mM of Zn<sup>2+</sup> (Fig. 5B). After 8 h, the cytotoxicity of parasites cultivated with 1.6 mM of Zn<sup>2+</sup> was 72% (Fig. 5B), while in the absence of Zn<sup>2+</sup> the cytotoxicity was 100%. These findings suggested that Zn<sup>2+</sup> had some negative effect on trichomonal cytotoxicity towards DU-145 cells.

### 3.6 The effect of Zn<sup>2+</sup> on the amount of TvCP65.

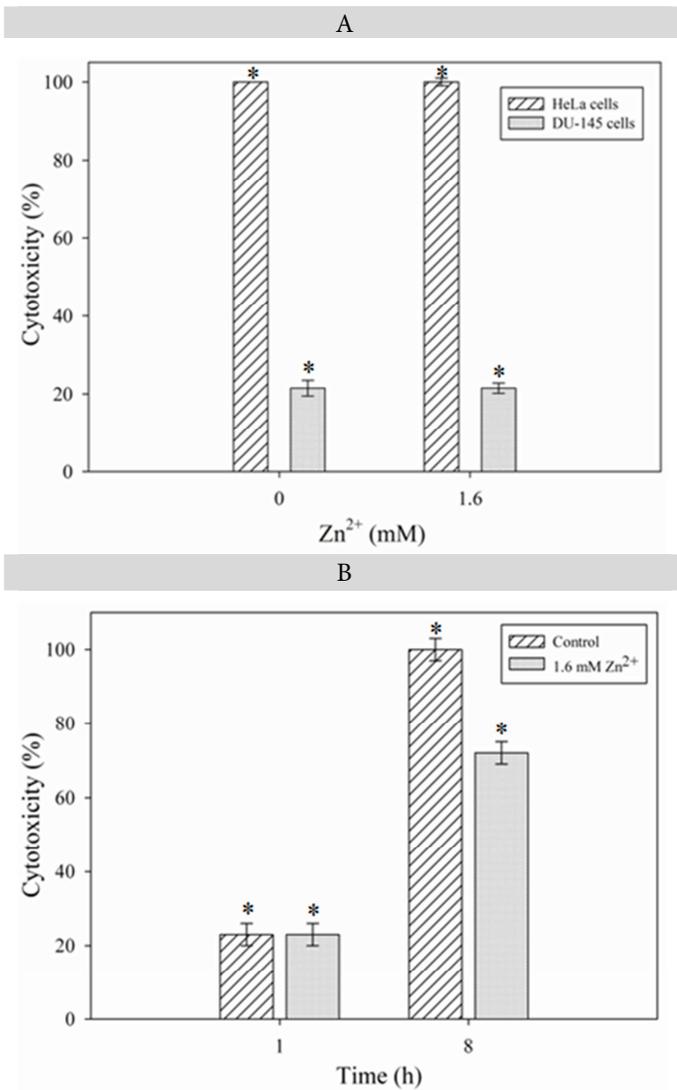
By Western blot analysis, an anti-TvCP65 antibody [16] immunodetected two bands that correspond to a precursor (Fig. 3D, lane 1 asterisk) and active TvCP65 enzyme (Fig. 3D, lane 1 arrow) in total protein extracts from parasites grown without Zn<sup>2+</sup>. In contrast, parasites grown in the presence of Zn<sup>2+</sup> lacked these bands (Fig. 3D, lane 2), while an anti- $\alpha$ -tubulin antibody detected a single band with the same intensity in the absence (Fig. 3D, lane 1) or the presence of Zn<sup>2+</sup> (lane 2). As expected, the amount of  $\alpha$ -tubulin was not changed in the presence of Zn<sup>2+</sup>, which indicated the Zn<sup>2+</sup> effect was specific for the TvCP65 protein level. These observations suggest that the amount of TvCP65 was reduced in the presence of Zn<sup>2+</sup>, which could be related to the reduction in the cytotoxicity level.

To confirm the decrease in TvCP65 protein level in *T. vaginalis*, indirect immunofluorescence assays were performed using permeabilized and non-permeabilized fixed parasites. Fig. 6 shows a reduction of fluorescence intensity

from surface (panels k and l) and cytoplasmic (panels h and i) TvCP65 in Zn<sup>2+</sup>-treated trichomonads compared to the control trichomonads, which show fluorescence on the surface (panels e and f) and in the cytoplasm (panels b and c). No fluorescence was observed from trichomonads treated identically but incubated with PI serum, a negative control (panels n and o). These data strongly suggested that the presence of Zn<sup>2+</sup> affects the amount of TvCP65, which resulted in a reduction in the TvCP65 proteolytic activity.

### 3.7 The effect of Zn<sup>2+</sup> on tvcp65 mRNA.

We then performed semi-quantitative and quantitative RT-PCR (Fig. 7) using cDNA from parasites grown in the presence or absence of Zn<sup>2+</sup> to determine the effect of Zn<sup>2+</sup> on the expression of the TvCP65 transcript. We obtained a 370-bp amplicon using cDNA from parasites grown in the absence (Fig. 7A, lane 1) or presence of 1.6 mM Zn<sup>2+</sup> (lane 2). Based on densitometry analyses (data not shown), an 80% reduction, compared to the no Zn<sup>2+</sup> control, was observed for *tvcp65* transcript. The 112-bp  $\beta$ -*tubulin* RT-PCR product was used as an internal control. As expected, the quantity of this transcript was not affected by the Zn<sup>2+</sup> concentration (Fig. 7B, lanes 1-2). No detectable PCR product from DNaseI-treated RNA from parasites cultivated in the presence or absence of Zn<sup>2+</sup> was obtained in the negative controls (RT(-) experiments)(Figs. 7A and 7B, lanes 3-4), which showed that ampli-



**Figure 5.** Zn<sup>2+</sup> effect on *T. vaginalis* cytotoxicity. Kinetic cytotoxicity of parasites cultivated in the presence or absence of 1.6 mM Zn<sup>2+</sup>. A) Cytotoxicity of *T. vaginalis* cultivated in the presence and absence of Zn<sup>2+</sup> towards HeLa and DU-145 cells after 1 h,  $p=0.056$  (\*). B) Data from trichomonal cytotoxicity towards DU-145 cells after 8 h without Zn<sup>2+</sup> were normalized to 100% for comparison. Cytotoxicity of *T. vaginalis* cultivated with 1.6 mM Zn<sup>2+</sup> was reduced 30%. Bars indicate the percent cytotoxicity of three experiments with triplicate samples, and error bars represent standard deviations,  $p=0.050$  (\*).

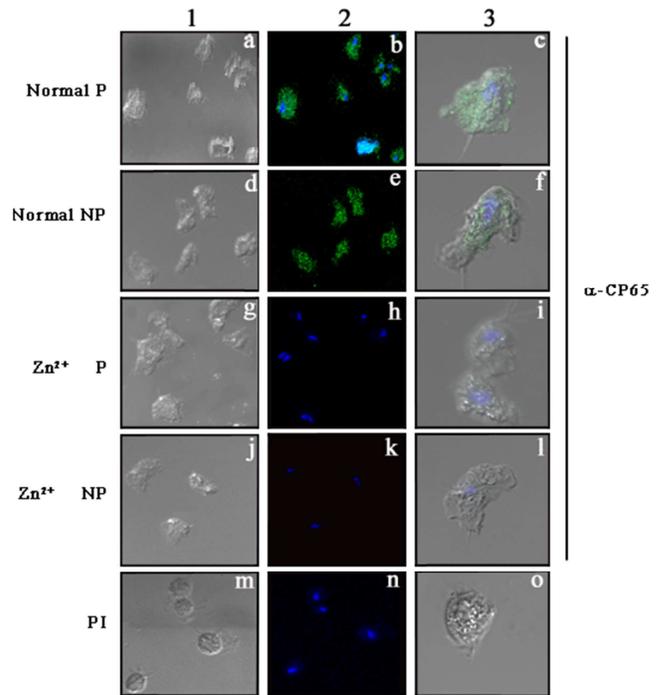
cons from *tvcp65* and  $\beta$ -*tubulin* genes corresponded to the quantity of mRNA of each gene. Consistent with these findings, the qRT-PCR results indicated that the *tvcp65* transcript was reduced 80% in the presence of Zn<sup>2+</sup> (Fig. 7C).

#### 4. Discussion

Zn<sup>2+</sup> is a catalytic component of over 300 enzymes [44]. This metal is an essential nutrient and plays a crucial role in several biological functions in the male microenvironment. Zn<sup>2+</sup> is a part of several proteins as both a structural and catalytic component, such as DNA-binding proteins, which contain Zn<sup>2+</sup> fingers, and the CCCH-type Zn<sup>2+</sup> finger protein, which participates in the adenylation and export of metazoan nuclear mRNAs [45]. Moreover, Zn<sup>2+</sup> has many functions in

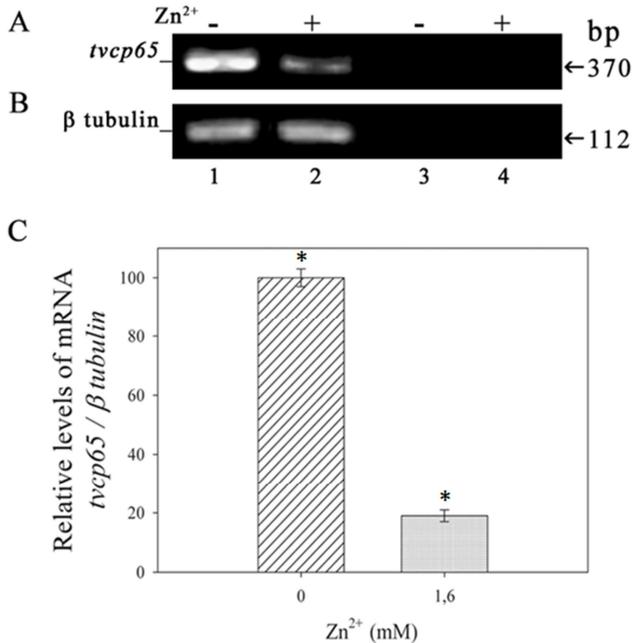
proteins, which include metalloregulation and Zn<sup>2+</sup> signaling [46]. Human prostatic fluid from healthy men and men with chronic bacterial prostatitis or trichomoniasis has different Zn<sup>2+</sup> concentrations [34]. Indeed, Zn<sup>2+</sup> levels found in the prostate could be important in resolving trichomoniasis in men [29, 34].

Our results with human DU-145 prostatic cells showed morphological and molecular changes of *T. vaginalis* that might contribute to understanding how this parasite causes damage to men. We analyzed different Zn<sup>2+</sup> concentrations and found that the major differences in cytotoxicity, duplication, and protein profile were obtained at 1.6 mM Zn<sup>2+</sup>, which is the minimum trichomonicidal concentration (MTC) [34] for *T. vaginalis*. In fact, survival of parasites in the presence of Zn<sup>2+</sup> was proportional to the Zn<sup>2+</sup> concentration. Nevertheless, parasite viability was not affected by Zn<sup>2+</sup>. As pointed out by Krieger and Rein, normal human prostatic fluid Zn<sup>2+</sup> concentrations kill *T. vaginalis in vitro* and may limit or solve trichomoniasis in most infected men [34]. Studies indicate that *T. vaginalis* survival in canine prostatic fluid is not significantly different from survival in Zn<sup>2+</sup> with a similar cation concentration [34]. This is not the case for *Tritrichomonas foetus*, which is unaffected by bovine seminal fluid. *T. foetus*



**Figure 6.** Expression and localization of TvCP65 in the presence of Zn<sup>2+</sup>. Immunofluorescence analyses of fixed, non-permeabilized (NP; d-f, j-l) and permeabilized (P; a-c, g-i) parasites grown in the presence (g-l) or absence (a-f, m-o) of 1.6 mM Zn<sup>2+</sup>, incubated with the anti-TvCP65 antibody (a-l) or preimmune sera (PI; m-o) followed by a secondary anti-mouse conjugated to a fluorescein isothiocyanate (Jackson) antibody (1:90 dilution), and mounted with Vectashield-DAPI. Indirect immunofluorescence (b, e, h, k), and the corresponding phase contrast microscopy at 60 $\times$  magnification (a, d, g, j) and 1.58 zoom (c, f, i, l). As a negative control, permeabilized parasites were incubated with PI (m-o) at a 1:20 dilution.

has an MTC of 200 mM for  $Zn^{2+}$ , which is far higher than the  $Zn^{2+}$  concentration found in the reproductive tract of a bull [29].



**Figure 7.**  $Zn^{2+}$  effect on the expression level of *T. vaginalis tvcp65* mRNA. Semi-quantitative RT-PCR analyses to detect the A) 370-bp *tvcp65* or B) 112-bp  $\beta$ -tubulin transcript using cDNA from parasites cultivated in the absence (lanes 1) or presence 1.6 mM (lanes 2) of  $Zn^{2+}$  and RT(-) a negative control from parasites cultivated in the absence (lanes 3) or presence of 1.6 mM  $Zn^{2+}$  (lanes 4). Arrows indicate the 370-bp and 112-bp RT-PCR products for the *tvcp65* and  $\beta$ -tubulin transcripts, respectively. These experiments were performed at least three times with similar results. C) qRT-PCR to quantify levels of *tvcp65* mRNA in trichomonads using cDNA from parasites cultivated in the absence or presence of 1.6 mM  $Zn^{2+}$ . Data were normalized to the amount of  $\beta$ -tubulin (quantity of *tvcp65* mRNA/quantity of  $\beta$ -tubulin mRNA). Bars show means of relative amounts of RT-PCR products from three separate experiments, and errors bars indicate standard deviation,  $p=0.01$  (\*).

Evidence indicates that *T. vaginalis* is not the only human parasite affected by  $Zn^{2+}$  levels. *Entamoeba histolytica* replication and adhesion is decreased in the presence of  $Zn^{2+}$ , which results in an inhibition of amoebic pathogenicity *in vivo* [47]. The activities of amoebic CPs are specifically inhibited by this cation. Interestingly, this  $Zn^{2+}$  inhibitory effect is reversible [48]. In men, prostate  $Zn^{2+}$  levels might inhibit the activities of *T. vaginalis* CPs, which may affect its cytotoxicity. This could be one of the reasons why male trichomoniasis is less severe.

When *T. vaginalis* attaches to VECs, it undergoes a rapid transformation to an amoeboid morphology [9]. Nevertheless, the parasites were not transformed when they came into contact with DU-145 prostatic cells. Furthermore, we observed extensive membrane interdigitations between parasite membranes and prostatic cells. This behavior has also been observed in parasites in contact with VECs [9]. There is a possibility that trichomonads that undergo morphological

transformations to amoeboid forms might have enhanced virulence capabilities [9]. Thus, the *T. vaginalis* amoeboid transformation did not occur with DU-145 prostatic cells, which resulted in lower cytotoxicity levels in those cells. This observation agrees with the diminished cytotoxicity levels of parasites grown in a similar prostate microenvironment due to the presence of  $Zn^{2+}$ . In addition, certain components of the vaginal microenvironment, such as  $Fe^{2+}$ , also influence the morphology of *T. vaginalis*.  $Fe^{2+}$  depletion induces morphological changes in this parasite from ellipsoid or amoeboid trophozoites to rounded ones, which is followed by flagella internalization and axostyle invagination [49], a pseudocyst morphology.

In addition,  $Zn^{2+}$  also affects the protein profile of *T. vaginalis*. Some proteins were up- or down-regulated in the presence of  $Zn^{2+}$ , while others showed no changes. A similar behavior was observed in the proteome of *T. vaginalis* in the presence or absence of  $Fe^{2+}$  [49], which is crucial in the female microenvironment and produces changes in its proteome. Forty-five proteins were found to be differentially expressed [49]. Because *T. vaginalis* is able to infect men, it can possibly adapt to the prostate microenvironment, survive, and establish an infection. We have demonstrated that *T. vaginalis* differentially expresses 27 proteins in the presence of  $Zn^{2+}$ , which suggests this parasite has the capability to adapt to different environments. These differences in protein expression profiles correlated with changes in some of its virulence properties, such as cytotoxicity.

We found that fimbrin (gi|123493533) and the aminopeptidase P-like metalloproteinase from Clan MG, family M24 (gi|123445672) were over-expressed in the presence of  $Zn^{2+}$ . In contrast, another fimbrin (gi|123397260), the asparaginyl endopeptidase-like cysteine peptidase (gi|123408789), and a different isoform of adenosylhomocysteinase (gi|123499896) were down-expressed with  $Zn^{2+}$ .

Fimbrins belong to a class of actin-bundling proteins that are components of the actin cytoskeleton and involved in many biological phenomena [50]. As an example, mutant *Saccharomyces cerevisiae* lacking the fimbrin gene display temperature sensitivity defects in growth, morphology, endocytosis, and sporulation [51-52]. However, over-expression of this protein is lethal [53]. In the male microenvironment, *T. vaginalis* faces a high  $Zn^{2+}$  concentration that induces the over-expression of fimbrin. This may be related to the morphological appearance of the parasite in the male microenvironment. In contrast, another fimbrin is down-regulated in the presence of this cation. Further studies are in progress to determinate the role of these two fimbrins in the biology of *T. vaginalis*.

Only two metalloproteinases (142 and 220 kDa) from *T. vaginalis* grown under normal culture conditions have been described [54]. This is the first report of a 50 kDa metalloproteinase from *T. vaginalis* expressed in the presence of  $Zn^{2+}$ . Moreover, *T. vaginalis* proteases are involved in many biological functions, including virulence, as virulence factors [16-17], participating in hemolysis [21], complement resistance

[22], cytotoxicity [18-19], apoptosis induction [24], nutrient acquisition [21, 55], and immune evasion [22-23]. It is possible that in the presence of  $Zn^{2+}$  *T. vaginalis* expresses this metalloproteinase for survival and colonization of the male microenvironment. We are in the process of purifying and characterizing this enzyme and determining its cellular localization and function.

Our findings indicate that it takes the parasite more time to cause damage to the prostatic DU-145 cell monolayers as compared to HeLa cell monolayers. This may be due to different cellular receptors. *T. vaginalis* presents a specific behavior according to its host. For example, although trichomonad adherence to HeLa and VEC cells occurs through the same adhesins [12], only VECs are capable of signaling for amoeboid transformation [9]. In addition, the density of particular receptors might also be involved in the responses from the parasite to different cell types.

Trichomonal cytotoxicity towards cervical cells includes the activity of different molecules, including proteinases. Its expression is induced under iron-depleted conditions [16, 18-19, 28], which increase trichomonal cytotoxicity levels towards HeLa cell monolayer due to an increase in TvCP65 proteolytic activity [28].  $Zn^{2+}$  particularly negatively affects the cytotoxicity of *T. vaginalis* towards DU-145 cells. Our results suggest that five *T. vaginalis* CPs (TvCP70, TvCP65, TvCP39, TvCP25, and TvCP20) were involved in contacting prostatic cells, and their proteolytic activities were negatively affected by  $Zn^{2+}$ . Interestingly, the CPs involved in cytotoxicity (TvCP65 and TvCP39) [16, 18-19, 28] also interacted with DU-145 cells, and their activities were negatively affected by  $Zn^{2+}$ . These findings might explain the reduction in cytotoxicity levels towards prostatic cells in the presence of  $Zn^{2+}$ , i.e., because of its inhibitory effect towards the CPs involved in cellular damage. Furthermore,  $Zn^{2+}$  inhibits a great number of proteins and functions, including CPs, signaling by phosphorylation, mitochondrial respiration, and neurotransmission. All of these are examples of the biological importance of  $Zn^{2+}$  inhibition [56].

One of the CPs involved in the *T. vaginalis* interaction with prostatic cells is TvCP65. It has previously been reported that TvCP65, which is located on the plasma membrane and in cytoplasm, has an important role in trichomonal cellular damage towards HeLa cell monolayers [16]. However, in the presence of  $Zn^{2+}$ , this molecule was not observed by immunofluorescence assays. The activity, protein level, and transcript of TvCP65 were negatively affected by  $Zn^{2+}$ , which led to a reduction in TvCP65-dependent cytotoxicity. TvCP65 proteolytic activity depends on the pH and the substrate. TvCP65 degrades collagen IV and human fibronectin at pH 3.6-7.5 and 3.6-6.0, respectively [16]. During *T. vaginalis* infections, the prostate pH changes. At the beginning of a male infection, the pH is about 7.0-7.5. When the infection is already established, the microenvironment pH changes to 8.0-8.5 [57]. Thus, TvCP65 might be implicated in prostatic cell damage at the beginning of the infection due to its activity at that pH range [16]. pH variations during male infections

might affect the activity of proteinases such as TvCP65 involved in the interaction of trichomonads and prostatic cells. TvCP65 degrades many substrates and interacts with DU-145 cells, this proteinase might bind to a specific receptor on the prostatic cell surface and be involved in nutrient acquisition mechanisms from different sources found in the male microenvironment.

According to the qRT-PCR results, the *tvcp65* mRNA level diminished in the presence of 1.6 mM  $Zn^{2+}$ . Although the proteolytic activity and amount of TvCP65 and its transcript diminished by 80% in the presence of  $Zn^{2+}$ , its cytotoxicity only decreased by 30%. It is important to mention that trichomonal cytotoxicity is a multifactorial process that includes the activity of porins for pore formation [58], contact-dependent disruption of the host cell membrane cytoskeleton [59], a cell-detaching factor [60], and the TvCP39 proteolytic activity [18-19], which might aid in trichomonal cytotoxicity in the presence of  $Zn^{2+}$ . We are in the process of determining the effect of  $Zn^{2+}$  on the amount of protein and transcript of TvCP39 to solve this question. Nevertheless, the possible mechanism of  $Zn^{2+}$  regulation is still unknown. However, it is known that several motifs found in transcriptional regulatory proteins are stabilized by  $Zn^{2+}$ , including the  $Zn^{2+}$  finger,  $Zn^{2+}$  clusters, RING finger, and LIM domains, and proteins that contain these domains are very common [44]. Evidence indicates that  $Zn^{2+}$  sensing in eukaryotes involves occupancy of the  $Zn^{2+}$  finger domains and their interactions with DNA [44]. The metal response element (MRE)-binding transcription factor-1 senses cellular  $Zn^{2+}$  ion concentrations in multicellular eukaryotes and activates the expression of proteins involved in the mechanisms of  $Zn^{2+}$  homeostasis. These types of sensors have been reported in *S. cerevisiae* and termed the  $Zn^{2+}$  responsive activator protein 1 (ZAP1), which has seven  $Zn^{2+}$  finger motifs [46]. Target genes regulated by ZAP1 include the  $Zn^{2+}$  transporter genes (*ZRT1* and *ZRT2*) and the *ZAP1* gene itself [46]. Authors have proposed that *ZAP1* encodes a transcriptional activator that binds to the promoters of these genes and activates their transcription when intracellular  $Zn^{2+}$  levels are low [46]. A similar  $Zn^{2+}$  sensing mechanism might be present in *T. vaginalis* that could result in different profile protein expressions in the presence of  $Zn^{2+}$ . However, there is no experimental evidence that *T. vaginalis* has a  $Zn^{2+}$  transcriptional activator or repressor. Nevertheless, a post-transcriptional regulatory mechanism mediated by an IRE/IRP-like system for TvCP4, a proteinase from *T. vaginalis*, has recently been described [61]. This post-transcriptional mechanism of  $Fe^{2+}$  regulation is based on the interactions of cytoplasmic  $Fe^{2+}$  regulatory proteins (IRPs) with  $Fe^{2+}$ -responsive elements (IREs) located in the untranslated regions (UTRs) of mRNAs of  $Fe^{2+}$ -regulated proteins [62]. A similar mechanism might regulate *tvcp65* mRNA in the presence of  $Zn^{2+}$ . Further studies are needed to solve this question.

## 5. Conclusions

Our results suggest that  $Zn^{2+}$  negatively affects growth, cy-

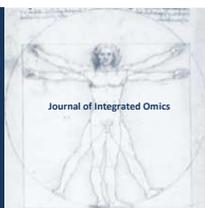
toxicity, and the expression of *T. vaginalis* proteinases related to the interaction with prostatic cells, which include those involved in trichomonal cytotoxicity, such as TvCP65. This is the first report of a 50 kDa metalloproteinase from *T. vaginalis* expressed in the presence of Zn<sup>2+</sup>. Further studies are necessary to elucidate the mechanism of *T. vaginalis* Zn<sup>2+</sup> regulation.

### Acknowledgements

The authors thank the Posgrado en Ciencias Genómicas at Universidad Autónoma de la Ciudad de Mexico (UACM) and CINVESTAV-I.P.N. We thank UACM for a scholarship to LIVC. LIQG was supported by a scholarship from CONACyT, Mexico. BICG was supported by a scholarship from ICyTDF, Mexico We thank M.Sc Eduardo Carrillo for technical assistance and Javier Mora for proofreading this manuscript. This work was supported by ICyTDF (PIFUTP08-150) and CONACyT (83808) to EAS.

### References

- Arroyo, R. and Alderete, J.F., *Infect. Immun.*, (1989)57: p. 2991-2997.
- Guenther, P.C., Secor, W.E., and Dezzutti, C.S., *Infect. Immun.*, (2005)73: p. 4155-4160.
- El-Shazly, A.M., et al., *J Egypt Soc Parasitol.*, (2001)31: p. 545-553.
- Cotch, M.F., et al., *Sex. Transm. Dis.*, (1997)24: p. 353-360.
- Viikki, M., et al., *Acta Oncol.*, (2000)9: p. 71-75.
- Schwebke, J.R. and Burgess, D., *Clin. Microbiol. Rev.*, (2004)17: p. 794-803.
- Alderete, J.F. and Garza, G.E., *Infect. Immun.*, (1985)50: p. 701-708.
- Alderete, J.F., et al., *Infect. Immun.*, (1988)56: p. 2558-2562.
- Arroyo, R., et al., *Mol. Microbiol.*, (1993)7: p. 299-309.
- Moreno-Brito, V., et al., *Cellular Microbiology*, (2005)7: p. 245-258.
- Garcia, A.F., et al., *Molecular Microbiology*, (2003)47: p. 1207-1224.
- Arroyo, R., Engbring, J., and Alderete, J.F., *Mol. Microbiol.*, (1992)6: p. 853-862.
- Hernández, H., et al., *Parasite Immunol.*, (2004)26: p. 119-125.
- Mendoza-Lopez, M.R., et al., *Infect. Immun.*, (2000)68: p. 4907-4912.
- Arroyo, R. and Alderete, J.F., *Arch. Med. Res.*, (1995)26: p. 279-285.
- Alvarez-Sanchez, M.E., et al., *Microbial Pathogenesis*, (2000)28: p. 193-202.
- Solano-González, E., et al., *International J. of Biochem. and Cell Biol.*, (2006)38: p. 2114-2127.
- Hernández-Gutiérrez, R., et al., *Exp. Parasitol.*, (2004)107: p. 125-135.
- Hernandez-Gutierrez, R., Ortega-López, J., and Arroyo, R., *J. Euk. Microbiol.*, (2003)50: p. 696-698.
- Rendon-Maldonado, J.G., et al., *Exp. Parasitol.*, (1998)89: p. 241-251.
- Dailey, D.C., Chang, T.H., and Alderete, J.F., *Parasitol.*, (1990)101: p. 171-175.
- Alderete, J.F., Provenzano, D., and Lehker, M.W., *Microb. Pathog.*, (1995)19: p. 93-103.
- Provenzano, D. and Alderete, J.F., *Infect. Immun.*, (1995)63: p. 3388-3395.
- Sommer, U., et al., *J. Biol. Chem.*, (2005)280: p. 23853-23860.
- Gorrell, T.E., *J. Bacteriol.*, (1985)161: p. 1228-1230.
- Lehker, M.W., Arroyo, R., and Alderete, J.F., *J Exp Med*, (1991)174: p. 311-318.
- Alderete, J.F., *Infect. Immun.*, (1999)67: p. 4298-4302.
- Alvarez-Sánchez, M.E., et al., *Microbes and Infection*, (2007)9: p. 1597-1605.
- Langley, J.G., Goldsmid, J.M., and Davies, N., *Genitourin Med.*, (1987)63: p. 264-267.
- Krieger, J.N., *Invest Urol.*, (1981)18: p. 411-417.
- Benchimol, M., et al., *Parasitol. Res.*, (2008)102: p. 597-604.
- Krieger, J.N., *Sex. Transm. Dis.*, (1995)22: p. 83-96.
- Krieger, J.N. and Rein, M.F., *J. Infect. Dis.*, (1982)146: p. 341-345.
- Krieger, J.N. and Rein, M.F., *Infect. Immun.*, (1982)37: p. 77-81.
- Fair, W.R., Couch, J., and Wehner, N., *Urology*, (1976)7: p. 169-177.
- Gardner, W.A.J., Culberson, D.E., and Bennet, B.D., *Arch. Pathol. Lab. Med.*, (1986)110: p. 430-432.
- Sutcliffe, S., et al., *International J. Cancer*, (2009)124: p. 2082-2087.
- Benchimol, M., et al., *Antimicrob. Agents Chemother.*, (1993)37: p. 2722-2726.
- Ramon-Luing, L.A., et al., *Proteomics*, (2010)10: p. 435-444.
- Laemmli, U.K., *Nature*, (1970)227: p. 680-685.
- Carvajal-Gamez, B., et al., *Infection, Genetics and Evolution*, (2010)10: p. 284-291.
- Leon-Sicairos, C.R., Leon-Felix, J., and Arroyo, R., *Microbiol.*, (2004)150: p. 1131-1138.
- Arroyo, R., et al., *Molecular Microbiology*, (1993)7: p. 299-309.
- Eide, D. and Guerinot, M.L., *ASM News*, (1997)63: p. 199-205.
- Hurt, J.A., et al., *J. of Cell Biol.*, (2009)185: p. 265-277.
- Zhao, H., et al., *J. Biol. Chem.*, (1998)273: p. 28713-28720.
- Vega Robledo, G.B., Carrero, J.C., and Ortiz-Ortiz, L., *Parasitol. Res.*, (1999)85: p. 487-492.
- Franco, E., Araujo Soares, M., and Meza, I., *Arch. Med. Res.*, (1999)30: p. 82-88.
- De Jesus, J.B., et al., *Proteomics*, (2007)7: p. 1961-1972.
- Delanote, V., Vandekerckhove, J., and Gettemans, J., *Acta Pharmacologica Sinica*, (2005)26: p. 769-779.
- Kubler, E. and Riezman, H., *Embo J*, (1993)12: p. 2855-2862.
- Adams, A.E.M., Botstein, D., and Drubin, D.G., *Nature*, (1991)354: p.
- Sandrock, T.M., et al., *Genetics*, (1999)151: p. 1287-1297.
- Bozner, P. and Demes, P., *Parasitology*, (1990)102: p.
- Lehker, M.W., et al., *J. Exp. Med.*, (1990)171: p.
- Wolfgang, M., Yuan, L., *Chem. Rev.*, (2009)109: p. 4682-4707.
- Fair, W.R. and Cordonnier, J.J., *J. Urology*, (1978)120: p. 695-698.
- Fiori, P.L., et al., *FEMS Microbiol. Lett.*, (1993)109: p. 13-18.
- Fiori, P.L., et al., *Infect. Immun.*, (1997)65: p. 5142-5148.
- Garber, G.E., Lemchuk-Favel, L.T., and Bowie, W.R., *J. Clin. Microbiol.*, (1989)27: p. 1548-1555.
- Solano-Gonzalez, E., et al., *FEBS Letters*, (2007)581: p. 2919-2928.
- Torres-Romero, J.C., Arroyo, R., *Infection, Genetics and Evolution*, (2009)9: p. 1065-1074.



ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v1i2.45

## Label-free protein quantification on tandem mass spectra in an ion trapping device

HuiSong Pak<sup>1</sup>, Carla Pasquarello<sup>1</sup>, Alexander Scherl<sup>1\*</sup>

<sup>1</sup>Proteomics Core Facility, Faculty of Medicine, University of Geneva, Geneva, Switzerland

Received: 27 October 2010 Accepted: 23 February 2011 Available Online: 14 March 2011

### ABSTRACT

Label free quantification using liquid chromatography and electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS) is widely used in quantitative proteomics. However, data-dependent bottom-up proteomics suffers from low reproducibility due to semi-random selection of precursor ions for tandem mass spectrometry. In addition, this acquisition mode is biased towards abundant peptides. To overcome these problems, alternative precursor-ion selection methods were developed, such as data-independent acquisition and *pseudo*-multiple selected reaction monitoring (p-mSRM). With these methods, several tandem mass spectra are acquired over the chromatographic elution time of precursor ions. In this report, we investigated if the acquired tandem mass spectra can be used for label-free quantification. For this, extracted fragment ion currents were correlated to relative protein concentration. A linear relationship between ion current and proteins concentration was observed over five orders of magnitude. Thus, we conclude that relative label-free peptide and protein quantification can be performed in an ion trap using the data-independent acquisition mode.

**Keywords:** *pseudo*-multiple selected reaction monitoring; quantitative analysis, ion trap, peptides.

### Abbreviations

**B-LG:**  $\beta$ -Lactoglobulin; **CID:** Collision induced dissociation; **DDA:** Data-dependent acquisition; **DIA:** Data-independent acquisition; **LC-ESI-MS/MS:** Liquid chromatography electrospray tandem mass spectrometry; **mSRM:** Multiple selected reaction monitoring; **PACIFIC:** Precursor Acquisition Independent From Ion Count; **p-mSRM:** *Pseudo*-multiple selected reaction monitoring; **SD:** Standard deviation; **SRM:** Selected reaction monitoring; **Trp II:** Trypsin inhibitor type II.

### 1. Introduction

Over the past years, direct interfacing of liquid chromatographs with tandem mass spectrometers (LC-MS/MS) has become a very popular tool for qualitative and quantitative analysis of complex peptide mixtures, such as enzymatic digestion product of complex protein samples. In particular, data-dependent acquisition (DDA), or on-the-fly precursor ion selection for isolation and subsequent activation and tandem mass analysis is widely used [1]. MS survey scans are acquired over the full mass range of peptide precursor ions and over the entire chromatographic elution time. The peptide precursor ion signals can then be used for

label-free relative quantification [2,3]. With this quantification method, ion currents from identical peptides from different samples are directly compared to each other.

Although dynamic exclusion [4] prevents redundant acquisition of the most abundant peptides, DDA is still biased towards abundant species. In addition, the selection contains a random component limiting reproducibility of identified peptides [5,6]. Thus, a major effort is focused on developing alternative precursor-ion selection methods. With the so-called "Precursor Acquisition Independent from Ion Count" (PACIFIC) or "Data-Independent

\*Corresponding author: Alexander Scherl, Proteomics Core Facility Centre Medical Universitaire Rue Michel-Servet 1 1211 Genève 4; tel: +41 22 379 5494 fax: +41 22 379 5926; E-mail Address: Alexander.Scherl@unige.ch

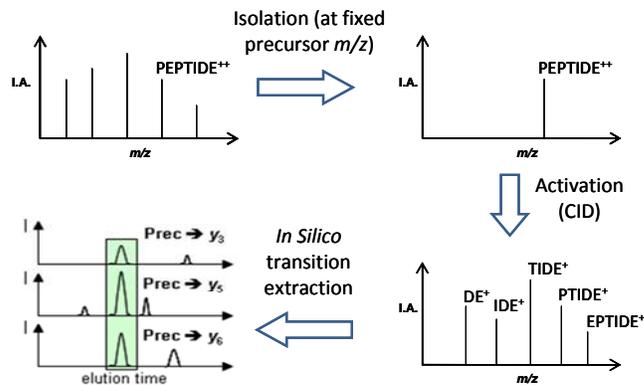
Acquisition” (DIA) methods, continuous, non overlapping mass-to-charge ratio ( $m/z$ ) windows are selected for isolation and subsequent activation [6,7]. Alternatively, no isolation is performed and all precursor ions are co-fragmented. The precursor-fragment ion lineage is then reconstituted post-acquisition using the chromatographic elution profile of all species [8,9]. A major advantage of these data-independent methods is that several tandem mass spectra are acquired over the entire chromatographic elution profile of all ions. Thus, fragment ion currents can be extracted for each precursor ion and used for quantification. Similarly, precursor-fragment ion transitions can be reconstructed post-acquisition, a term referred to as “*pseudo*-multiple selected reaction monitoring” or p-mSRM [10,11]. This mode of operation is described in Figure 1.

The term “*pseudo*” refers to the computer calculated post-acquisition reconstruction of the transition. Indeed, the full fragment ion spectrum is recorded here, by opposition to classical mSRM acquired in triple quadrupole instruments. *Pseudo*-mSRM has been showed to provide excellent quantitative measurement over a large dynamic range for small molecules [12] and peptides [13,14] in simple matrices and for non-complex mixtures. Here, we show that label-free MS2-based quantification is possible over at least five orders of magnitudes in very complex matrices, *e.g.* digested human plasma.

## 2. Materials and methods

Liquid plasma,  $\beta$ -Lactoglobulin (B-LG), trypsin inhibitor type II (Trp II), iodoacetamide (IAA) and acetonitril (AcN) were purchased from Sigma (St.louis, MO, USA). Urea, ammonium bicarbonate (AB), dithioerythritol (DTE) and water for chromatography and dilution were from Merck (Darmstadt, Germany). Porcine trypsin and formic acid (FA) was respectively from Promega (Madison, WI, USA) and Biosolve (Valkenswaard, the Netherlands). Stationary phases for columns were from Michrom (Auburn, CA). Analytical column (OD = 375  $\mu$ m, ID = 75  $\mu$ m, L=150 mm) and pre-column (OD = 375  $\mu$ m, ID = 100  $\mu$ m, L=20 mm) was made from fused silica tubing from BGB Analytik AG (Boeckten, Switzerland)

Human plasma and standard proteins were digested as previously reported [15]. In short, 500 $\mu$ g of B-LG and Trp II were dissolved in 200 $\mu$ l of 6M Urea and 50mM AB. 10 $\mu$ l of DTE 38mM was added and the solution was incubated at 37 $^{\circ}$ C for 60 min (reduction). Then 20 $\mu$ l of IAA 108mM was added for alkylation during 60 min in the dark. Liquid digestion was performed overnight, by adding 25 $\mu$ l of trypsin (0.2 $\mu$ g/ $\mu$ l). The digested solution was desalted with a C18 micro-spin column (Harvard Apparatus, Holliston, MA, USA) and dried. In order to have aliquots of 10pmol/ $\mu$ l, dried solutions were dissolved in AcN/FA/H<sub>2</sub>O 5/0.1/94.9%. These digests were spiked at various concentrations in trypsin digested, non-depleted human plasma. The concentrations were adjusted so that a total amount of 1, 10, 100 attomoles, 1, 10,



**Figure 1.** Principle of Selected Reaction Monitoring (SRM) and *pseudo*-multiple Selected Reaction Monitoring (p-mSRM). From a complex mixture of peptides, precursor ions are isolated and activated, typically by CID. All product ions are collected in the ion trap and scanned out according to their  $m/z$  ratio. This operation is repeated during the entire chromatographic elution of the peptides. Once all the data is collected, the ion currents corresponding to all desired transitions are reconstructed *in silico*.

100 femtomoles and 1 picomole standard proteins were injected in our LC-MS/MS system. The injected plasma amount was kept constant of 0.126  $\mu$ g per injection (corresponding approximately to one picomole of albumin).

The LC-MS/MS system consisted of a NanoAcquity chromatograph (Waters, Milford, MA) interfaced with an LTQ-Orbitrap velos mass spectrometer (Thermo Scientific, San Jose, CA). Peptides were trapped on a home-made, 20 mm long precolumn of 100  $\mu$ m inner diameter and separated on a 150 mm analytical column of 75  $\mu$ m inner diameter. The analytical separation was run for 65 min using a gradient of H<sub>2</sub>O/FA 99.9%/0.1% (solvent A) and CH<sub>3</sub>CN/FA 99.9%/0.1% (solvent B). The gradient was run as follows: 0–1 min 95% A and 5% B, then to 65% A and 35% B at 55 min, and 20% A and 80% B at 65 min at a flow rate of 220 nL/min.

All samples were analyzed from the most diluted to the most concentrated, and new chromatographic columns were used for each technical replicate. The mass spectrometer was operated in the following conditions: for DDA, full MS spectra were acquired in the Orbitrap detector from  $m/z = 400 - 2000$ . The target ion population was 500,000 ions. Tandem mass spectra were acquired in a data-dependent manner in the linear ion trap on the five most abundant precursors (if present). Dynamic exclusion was set to one minute. Precursor isolation window was set to 2.0  $m/z$  units. Normalized collision energy was set to 35%. For p-mSRM acquisitions, full MS spectra were acquired in the Orbitrap detector from  $m/z = 400 - 2000$ . The target ion population was 500,000 ions. Tandem mass spectra were acquired on the three  $\beta$ -lactoglobulin peptide precursor ions and on the three trypsin inhibitor peptide precursor ions listed in table 1, in a data independent manner. In other words, the precursor-ion  $m/z$  was isolated and fragmented over the full chromatographic analysis, no matter if precursor ions were present or not. The scan sequence was Full MS1 (orbitrap acquisition), CID of

$m/z=858.406$ , CID of  $m/z=545.929$ , CID of  $m/z=623.294$ , CID of  $m/z=588.316$ , CID of  $m/z=600.858$ , CID of  $m/z=928.475$ . All CID spectra were acquired in the linear ion trap. The target ion population was set to 10,000 ions. The precursor isolation window was set to 2.0  $m/z$  units. Normalized collision energy was set to 35%. All analysis were run in technical replicates (n=2). The first three DDA analyses, spiked with respectively 10 and 100 amol of standard proteins were used for database search.

Protein identification peak lists were generated from raw data using the embedded software from the instrument vendor (extract\_MSN.exe). The monoisotopic masses of the selected precursor ions were corrected using an in-house written Perl script [16]. The corrected mgf files were searched against the SwissProt/Uniprot database (release 15.10 of 03-Nov-2009) using Phenyx (GeneBio, Geneva, Switzerland). Homo sapiens taxonomy was specified for database searching (34785 sequences) and the two standard protein sequences were added. The parent ion tolerance was set to 10 ppm. Variable amino acid modification was oxidized methionine. Fixed amino acid modification was carbamidomethylation of cysteines. Trypsin was selected as the enzyme, with one potential missed cleavage, and the normal cleavage mode was used. Only one search round was used with selection of "turbo" scoring. The peptide p value was 1 E-2 for LTQ-OT data. False-positive ratios were estimated using a reverse decoy database [17]. All datasets were searched once in the forward and once in the reverse database. Separate searches were used to keep the database size constant. Protein and peptide score were then set up to maintain the false positive peptide ratio below 1%. This resulted in a slight overestimation of the false-positive ratio [17]. For all analyses, only proteins matching two different peptide sequences

were kept.

Quantitative data (extracted ion chromatograms and chromatographic peak integration) were extracted with Xcalibur 2.6 (Thermo Scientific). Base-to-base peak integration was performed manually using the "add peak" function of Xcalibur. The peak area value calculated by the software was used.

### 3. Results and discussion

In a first experiment, digested human plasma and the two digested standard proteins  $\beta$ -lactoglobulin (B-LG) and trypsin inhibitor type II (Trp II) were separately analyzed by DDA ESI-LC-MS/MS. This resulted with the identification of an average of 926 (SD=12) unique peptides from human plasma, corresponding to 413 (SD=18) unique proteins identified with at least two peptides. These numbers demonstrate the quality of the plasma digestion and the reproducibility of the data-dependent analysis. Then, for each standard protein the three peptides giving raise to the three most abundant precursor ion signals were selected for further analysis (Table 1). These 3 peptides correspond to sequence coverage of 15% for B-LG and 19% of Trp II.

In the next experiment, the digested proteins were mixed with digested plasma so that the injected amount varied from one attomole to one picomole. The quantity of plasma was kept constant, at 0.126  $\mu$ g per injection (corresponding to approximately 1 pmol of injected albumin). The peptide mixture was then analyzed using DDA and *pseudo*-mSRM. For each concentration, the acquired MS1 spectra were inspected for the presence of the three peptide precursor of each standard protein. The isotopic cluster corresponding to peptide LFSNPTQLEEQCHI was visible at injected amounts of 1 femtomole and above. The two other peptides were dis-

**Table 1.** Observed peptides for MS1 and MS2-based quantification and their detection limit

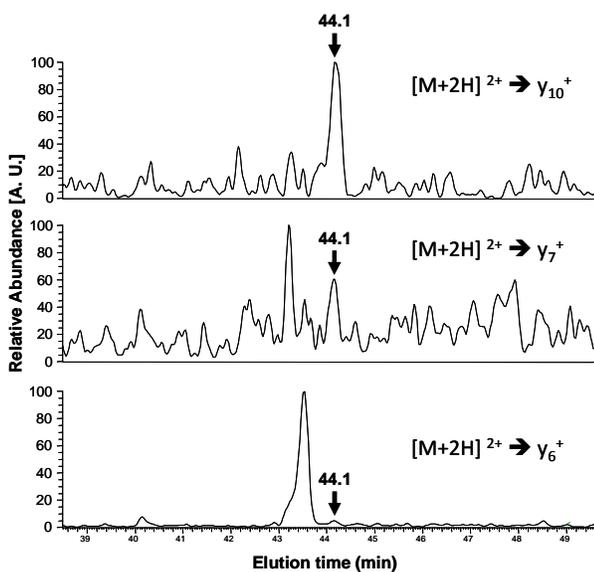
Protein	Peptide	MS1 detection limit	p-SRM transition	MS2 detection limit	retention time (min)
$\beta$ -Lactoglobulin	LFSNPTQLEEQCHI	1 fmol	$[M+2H]^{++} \rightarrow y_6^+$	100 amol	43,04
			$[M+2H]^{++} \rightarrow y_7^+$	100 amol	43,04
			$[M+2H]^{++} \rightarrow y_{10}^+$	10 amol	43,04
$\beta$ -Lactoglobulin	TPEVDDEALEKFDK	10 fmol	$[M+3H]^{+++} \rightarrow y_{10}^{++}$	100 amol	33,97
			$[M+3H]^{+++} \rightarrow y_{11}^{++}$	100 amol	33,97
			$[M+3H]^{+++} \rightarrow y_{12}^{++}$	10 amol	33,97
$\beta$ -Lactoglobulin	TPEVDDEALEK	10 fmol	$[M+2H]^{++} \rightarrow y_7^+$	100 amol	26,56
			$[M+2H]^{++} \rightarrow y_8^+$	100 amol	26,56
			$[M+2H]^{++} \rightarrow y_{10}^{++}$	100 amol	26,56
Trypsin Inhibitor	NELDKGIGTIISPPYR	1 fmol	$[M+3H]^{+++} \rightarrow y_5^+$	100 amol	39,98
			$[M+3H]^{+++} \rightarrow y_6^+$	100 amol	39,98
			$[M+3H]^{+++} \rightarrow y_{14}^{++}$	1 fmol	39,98
Trypsin Inhibitor	NKPLVVQFQK	10 fmol	$[M+2H]^{++} \rightarrow y_5^+$	100 amol	28,27
			$[M+2H]^{++} \rightarrow y_6^+$	100 amol	28,27
			$[M+2H]^{++} \rightarrow y_8^+$	10 amol	28,27
Trypsin Inhibitor	AAPTGNERCPLTWQSR	100 fmol	$[M+2H]^{++} \rightarrow y_8^+$	1 fmol	28,00
			$[M+2H]^{++} \rightarrow y_9^+$	1 fmol	28,00
			$[M+2H]^{++} \rightarrow y_{10}^{++}$	1 fmol	28,00

played clear isotopic clusters at 10 femtomoles and above. Similarly, one trypsin inhibitor peptide was visible at 1 femtomole of injected protein, the other two at 10 and 100 femtomoles. Thus, we concluded that relative quantification using an MS1 based label-free approach can be performed from amounts of 1 femtomoles and above. If the upper limit is arbitrarily set to 1 pmol, relative quantification could be done over a concentration range of three orders of magnitude (1 femtomole to 1 picomole).

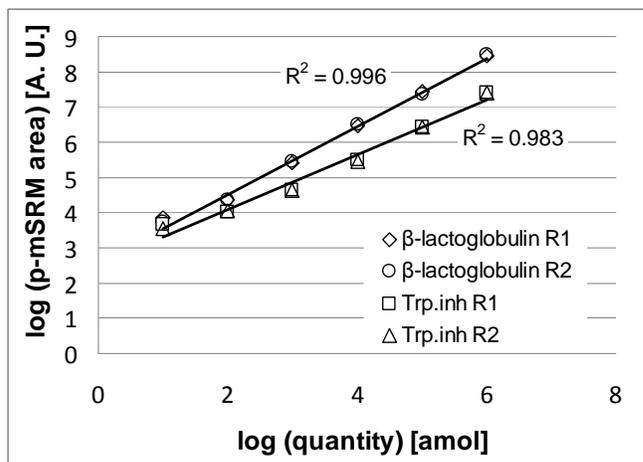
Individual fragment ion currents were extracted for the same standard protein peptides from the *pseudo*-mSRM data. For three out of the six peptides, fragment ions were clearly visible at injected amounts of 10 attomoles and above (Table 1). The extracted fragment ion chromatograms for  $\beta$ -lactoglobulin peptide LFSNPTQLEEQCHI are shown in Figure 2.

Next, we investigated if the extracted fragment ion chromatograms could be used for label-free MS2-based relative quantification. However, this strategy is complicated by the fact that multiple fragment-ion chromatograms can be reconstructed from a single peptide precursor ion. Thus, we applied an empirical method similar to the one showed by Silva and co-workers. Indeed, this group showed that the average MS signal response of the three most abundant peptides is constant for each protein [9]. Based on this, the sum of the three most abundant fragment ions was calculated for all three peptides per standard protein. The protein abundance values calculated with this method showed excellent linearity with protein concentration (Figure 3).

Coefficients of correlation are above 0.98 for both measured proteins and all replicates over the full range of concentration. This implies that relative quantification using an MS2-based label-free approach can be performed at lowest protein concentrations. With an arbitrary upper limit of one



**Figure 2.** Extracted *p*-mSRM transitions ion currents from 10 amol of  $\beta$ -lactoglobulin peptide LFSNPTQLEEQCHI spiked into digested human plasma. The peptide elutes at 44.1 minutes.



**Figure 3.** Combined *pseudo*-mSRM area as function of protein concentration for  $\beta$ -lactoglobulin and trypsin inhibitor spiked into human plasma. R1 = replicate 1, R2 = replicate 2. The correlation coefficient  $R^2$  is calculated from the  $\log_{10}$  values of *p*-mSRM area and concentration.

picomole, relative quantification could be done over five orders of magnitude (10 attomoles to 1 picomole). It represents an increase in dynamic range of two orders of magnitude compared to DDA acquisition, traditionally used for label-free quantification. Indeed, if the DDA mode is used on a very complex sample, the dynamic range of the analysis is close to the intra-spectrum dynamic range of the analyzer (in our case the Orbitrap analyzer). Makarov and co-workers have shown that the intra-spectrum dynamic range of an orbitrap was around 5,000 for full MS1 survey scans [18]. By contrast, the high dynamic range using the *pseudo*-mSRM mode can be explained by the combined dynamic range of the mass analyzer (in this case the linear ion-trap) and the ion injection time to fill the trap. The dynamic range of an ion trap is around  $10^2 - 10^3$ , with a varying injection time between 0.1-100 ms. Consequently, the resulting dynamic range with *pseudo*-mSRM is the product of this two values. This indicates that a dynamic range of five orders of magnitude should be possible, which corresponds to the value found in our experiments.

#### 4. Conclusions

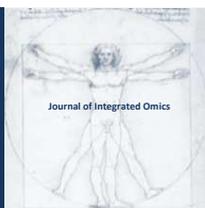
Our data shows that relative label-free quantification is possible in ion trapping devices, using the data-independent or *p*-mSRM mode of operation. Quantification can be performed directly, in complex samples over a large dynamic range. Moreover, combination of large-scale data-independent protein identification and label-free quantification is also possible. Data-independent acquisition strategies such as the recently published PACIFIC provide good results in terms of number of identifications and dynamic range. The data format itself corresponds to nothing else than a large-scale *pseudo*-mSRM experiments. The obtained data can therefore be directly used for MS2-based relative quantification.

**Acknowledgments**

A. S. thanks the Swiss National Science Foundation for support (grant 315230\_130830).

**References**

1. D.C. Stahl, K.M. Swiderek, M.T. Davis, T.D. Lee, J. Amer. Soc. Mass Spectrom. 7 (1996) 532-540.
2. W. Wang, H. Zhou, H. Lin, S. Roy, T.A. Shaler, L.R. Hill, S. Norton, P. Kumar, M. Anderle, C.H. Becker, Anal Chem 75 (2003) 4818-4826.
3. D. Radulovic, S. Jelveh, S. Ryu, T.G. Hamilton, E. Foss, Y. Mao, A. Emili, Mol Cell Proteomics 3 (2004) 984-997.
4. C.L. Gatlin, J.K. Eng, S.T. Cross, J.C. Detter, J.R. Yates, 3rd, Anal Chem 72 (2000) 757-763.
5. H. Liu, R.G. Sadygov, J.R. Yates, 3rd, Anal Chem 76 (2004) 4193-4201.
6. A. Panchoaud, A. Scherl, S.A. Shaffer, P.D. von Haller, H.D. Kulasekara, S.I. Miller, D.R. Goodlett, Anal Chem 81 (2009) 6481-6488.
7. M. Bern, G. Finney, M.R. Hoopmann, G. Merrihew, M.J. Toth, M.J. MacCoss, Anal Chem 82 833-841.
8. S. Purvine, J.T. Eppel, E.C. Yi, D.R. Goodlett, Proteomics 3 (2003) 847-850.
9. J.C. Silva, M.V. Gorenstein, G.Z. Li, J.P. Vissers, S.J. Geromanos, Mol Cell Proteomics 5 (2006) 144-156.
10. A. Scherl, S.A. Shaffer, G.K. Taylor, H.D. Kulasekara, S.I. Miller, D.R. Goodlett, Anal Chem 80 (2008) 1182-1191.
11. V. Lange, P. Picotti, B. Domon, R. Aebersold, Mol Syst Biol 4 (2008) 222.
12. S. Yang, J. Cha, K. Carlson, Rapid Commun Mass Spectrom 18 (2004) 2131-2145.
13. J.H. Baek, H. Kim, B. Shin, M.H. Yu, J Proteome Res 8 (2009) 3625-3632.
14. P. Shipkova, D.M. Drexler, R. Langish, J. Smalley, M.E. Salyan, M. Sanders, Rapid Commun Mass Spectrom 22 (2008) 1359-1366.
15. L. Dayon, C. Pasquarello, C. Hoogland, J.C. Sanchez, A. Scherl, J Proteomics 73 (2010) 769-777.
16. A. Scherl, Y.S. Tsai, S.A. Shaffer, D.R. Goodlett, Proteomics 8 (2008) 2791-2797.
17. J.E. Elias, S.P. Gygi, Nat Methods 4 (2007) 207-214.
18. A. Makarov, E. Denisov, O. Lange, S. Horning, J Am Soc Mass Spectrom 17 (2006) 977-982.



ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v1i2.59

## Development of an in-house protocol for the OFFGEL fractionation of plant proteins

Delphine Vincent<sup>\*1</sup>, Peter S. Solomon<sup>1</sup>.

<sup>1</sup>Research School of Biology, The Australian National University, Australia.

Received: 13 January 2011 Accepted: 02 March 2011 Available Online: 04 March 2011

### ABSTRACT

Protein sample prefractionation using off-gel technology coupled to mass spectrometry-based proteomics has proven very efficient in identifying novel proteins of low abundance in various biological systems with the exception of plant tissues. In this study, we have prefractionated protein samples from wheat leaves and apoplastic fluid using the Agilent 3100 OFFGEL fractionator along 3-10 pH range into 24 fractions. Two methods were compared: 1/ the Agilent method based on the use of a proprietary starter kit which comprises all the necessary consumables, and 2/ an in-house method in which focusing solutions were prepared using electrophoresis grade chemicals, and immobilised pH gradient (IPG) strips and electrode pads were acquired elsewhere. Method comparisons and quality of focusing were assessed using pH distributions, as well as one- and two-dimensional patterns. Both methods produced comparable well-resolved electrophoretic patterns. This study will allow other laboratories using OFFGEL device to consider in-house solutions.

**Keywords:** *Triticum aestivum*, leaf; apoplast; IPG strip; SDS-PAGE; Agilent 3100 OFFGEL fractionator.

### 1. Introduction

Many researchers adopt proteomics to tackle a biological question because it allows for the exploration of the cell acting components. As diverse as the proteomic field has become, two-dimensional electrophoresis (2-DE) which separates the proteins according to first their isoelectric point (pI) using isoelectric focusing (IEF) and second their molecular weight (MW) [1] remains a technique widely used. Reproducibility of 2-DE was dramatically improved upon the introduction of immobilised pH gradient (IPG) strips during the first dimension [2]. This technique can resolve thousands entities on a single gel, which reflects only a very small proportion of the proteome. Most of the technological improvements aim at increasing protein coverage, accuracy and sensitivity, thus providing the researchers with a variety of strategies with enhanced resolution. One of these strategies is sample prefractionation. The two main approaches are the chromatographic and electrophoretic prefractionations. Both exploit the physico-chemical proper-

ties of proteins, such as their charge, hydrophobicity, mass, *etc.*, to group them into discrete fractions prior to other analytical resolving techniques, thus optimising both resolution and sensitivity. Electrophoretic prefractionation based on electrokinetic migration such as IEF is gaining popularity as it is fully compatible with 2-DE, unlike chromatographic fractions, and offers the advantage of concentrating proteins of low abundance according to their pI. Many devices exist (reviewed in [3, 4]); one of them is off-gel electrophoresis (OGE).

OGE uses IEF and was developed by Ros *et al.* [5] who adapted a multicompartiment chamber on top of an IPG strip providing the desired pH gradient; solubilised proteins migrate through the strip until they reach their pI at a given compartment and then return into solution. By adjusting the pH range along IPG strips, OGE can be tailored to the protein sample requirement and achieve the high resolution of 0.1 pH unit [6]. An additional benefit of OGE is the desalting

**\*Corresponding author:** Delphine Vincent. Present address: Plant Sciences Division, Research School of Biology, Robertson Building 46, The Australian National University, Canberra, 0200 ACT, Australia. Telephone: +61 2 61254026. Fax: +61 2 61254331. Email Address: delphine.vincent@anu.edu.au.

of protein sample [7]. OGE has also been successfully used to separate antibody charge isoforms prior to capillary IEF [8]. Initially devised for fractionation of proteins prior to gel-based techniques, OGE was soon employed to fractionate peptides prior to reverse phase (RP) liquid chromatography (LC) coupled to tandem mass spectrometry (MS/MS) [9]. In such shotgun proteomics approaches, proteins can be enzymatically digested after [10] or before [11] OGE prefractionation; alternatively a sample can be fractionated twice, first using proteins and second using tryptic peptides [12]. Protein prefractionation using first OGE, second RP-LC, then followed by trypsin digestion and LC-MS/MS analysis was found to be a reproducible strategy to study complex samples [13]. The OGE device allowing unbiased fractionation of both proteins and peptides is now commercialised under the trade name of OFFGEL fractionator (Agilent Technologies) and was first reported by Hörth *et al.* [11]. A highly efficient quantitative strategy combines peptide labelling with isobaric tag for relative and absolute quantitation (iTRAQ) and OGE prefractionation prior to LC-MS/MS analysis [14-16]; it also enhances the identification of low abundant proteins [17]. Very recently, OGE was employed to explore phosphoproteomes [18, 19].

More publications, thus far limited to industrial or medical research, have made full use of this additional resolving step, mostly in a shotgun approach capacity, either fractionating proteins [20, 21], peptides [22-24] or both [25, 26]. However, to date, OGE has yet to be used on plant tissues. Our lab is interested in elucidating wheat responses to fungal pathogens; proteomics is one of the strategies adopted through the use of both gel-based and gel-free approaches. Here we report for the first time the OGE prefractionation of plant proteins. Soluble proteins were recovered from wheat whole leaves and separated using Agilent 3100 OFFGEL fractionator either using the purchased Agilent's kit or in-house solutions. The quality of protein prefractionation was assessed by 1-DE and 2-DE which helped comparing both the manufacturer's and our in-house protocols. Our in-house protocol was applied to apoplastic fluids from wheat leaves. Whether using whole leaf or apoplast samples, two-dimensional patterns were greatly enriched in spots following OGE prefractionation.

## 2. Materials and Methods

### 1. Wheat culture and sampling

Wheat (*Triticum aestivum* L. cv. Grandin) was grown from seeds. Prior to sowing, seeds were sterilised by incubation in 0.1% v/v sodium hydroxide/5% v/v ethanol for 10 min followed by three washes in sterile water. Ten sterilised seeds were sown per 2 L pot full of vermiculite grade 3 (Australian Perlite, Sydney NSW Australia) with 10 pellets of slow release Osmocote Exact fertiliser (Scotts Australia, Baulkham hills NSW Australia). Pots were placed in a sunlit greenhouse kept at 20°C and fully watered daily. All seeds

had germinated after 4 days and the first leaf emerged from the cotyledon after 9 days. First true leaf leaves were collected after 12 days. Leaf samples for whole protein extraction were pooled into a tube and instantly frozen by immersion in liquid nitrogen and stored at -80°C until use. Leaf samples for apoplast recovery were used fresh, immediately after sampling.

### 2. Apoplast recovery

Apoplastic fluid was recovered according to Solomon and Oliver [27]. Briefly, leaves were cut into 3cm-long pieces and placed into a 35 mL syringe barrel along with 15 mL ddH<sub>2</sub>O. By adapting the tip into a rubber stopper and operating the plunger, ddH<sub>2</sub>O was forced into the leaf pieces through the stomata until the leaves turned dark green. Leaf pieces were retrieved using tweezers and the residual water on their surface was blotted with paper towels. The dry leaf pieces were transferred into a 10 mL syringe barrel without adjusting the plunger. A 1.5 mL Eppendorf tube was adapted at the end of the syringe tip and both the syringe and the Eppendorf were inserted into a 50 mL Falcon tube. The apoplastic fluids were recovered by centrifugation using a swing rotor for 5 min at 4000 xg at room temperature. Ten leaves produced approximately 0.5 mL of fluid. Apoplastic samples were stored at -80°C until use.

### 3. Protein extraction from whole leaves

Frozen leaf samples were finely ground in liquid nitrogen using a chilled mortar and pestle. Soluble proteins were extracted using a trichloroacetic acid (TCA)/ 2-mercaptoethanol (2-ME)/acetone method devised by Damerval *et al.* [28]. Briefly, the frozen powder was transferred into chilled 50 mL Nalgene Oak Ridge tubes (Thermo Fisher Scientific, Scoresby VIC Australia) and filled up with ice-cold 10% w/v TCA/0.007% v/v 2-ME/acetone solution. Following homogenisation, tubes were left to incubate at -20°C overnight. Tubes were centrifuged for 30 min at 12,000 xg and -10°C and the supernatant was discarded. Pellets were rinsed three times in ice-cold 0.007% v/v 2-ME/acetone solution and recovered by centrifugation (30 min, 12,000 xg, -10°C). Pellets were dried at room temperature overnight and solubilised in 1 mL resuspension (R) solution (7M urea, 2M thiourea, 4% w/v 3-[3-(cholamidopropyl)dimethylammonio]-1-propane-sulfonate, 1% w/v dithiothreitol, 1% v/v 2-ME, 10mM tris-(2-carboxyethyl)-phosphine-HCl, 0.5% ampholites 3-10). Protein extracts were transferred into 2 mL Eppendorf tubes and stored at -20°C until use.

### 4. Protein content assay

The protein content of both leaf protein extracts and apoplastic fluid samples was estimated using 2-D Quant Kit (GE Healthcare, Ryldamere NSW Australia) following the

manufacturer's instructions.

5. OFFGEL fractionation

Proteins were fractionated into 24 fractions by liquid isoelectric focusing (IEF) using the Agilent 3100 OFFGEL fractionator (Agilent Technologies, Forest Hill VIC Australia). Two methods were tested for OFFGEL fractionation of whole leaf samples. 1/ In the Agilent's method, fractionation was performed using the Starter kit (Agilent Technologies Part No. 5188-6444) which includes all necessary consumables. 2/ In our in-house method, fractionation was performed by, whenever possible, replacing all the consumables provided by Agilent's starter kit with their lab equivalents listed in Table 1. The 1.25X OFFGEL stock solution (50 mL) for high resolution separation along 3-10 pH range was prepared following the manufacturer's instructions, aliquoted into 2 mL Eppendorf tubes and stored at -20°C until use.

Parts were assembled, Immobilised pH gradient (IPG) strips were rehydrated, and protein samples were diluted into the 1.25X OFFGEL stock solution as instructed in the Agilent Quick Start Guide. The same wheat leaf sample was separated on two different IPG strips placed in the same tray following the Agilent or the in-house method. A total of 0.5 mg of leaf proteins were loaded onto each IPG strip. The apoplast sample was fractionated during a separated experiment following our in-house method only, by loading 1 mg of apoplastic proteins.

Both OFFGEL fractionation runs for leaf and apoplast samples respectively were performed using the preset program OG24PR00 (64kVhrs, 8000V, 50µA, 200mW). At the end of the runs, fractions were transferred into individual 1.5 mL Eppendorf tubes. Collected fractions were stored at -20°C until further use.

6. pH measurement of the OFFGEL fractions

Fraction pH values were estimated by diluting 2µL into 20µL ddH<sub>2</sub>O and wetting 0-14 pH indicator strips (Merck, Kilsyth VIC Australia). Values are displayed in Table 2 and illustrated in Figure 1. The expected pH values *per* fraction were calculated according to the IPG strip supplier data (24 cm 3-10 pH) and the OFFGEL well dimensions (24 x 1 cm long wells). The 7 pH units along 3-10 range were divided by 24 to give a value of approximately 0.3 pH units for the range covered by each fraction.

7. One-dimensional electrophoresis (1-DE)

Two micrograms of unfractionated protein samples or 10 µL of each fraction were diluted into up to 20 µL of Laemmli Sample Buffer (Bio-Rad, Gladesville, NSW, Australia). Broad Range SDS-PAGE Molecular Weight (MW) Standards (Bio-Rad) were used for MW reference. Samples were boiled for 5 min and loaded onto 4% homecast stacking polyacrylamide gels on top of 12% homecast resolving polyacrylamide gels

Table 1. Consumables used for each OFFGEL fractionation method on wheat leaf protein samples.

consumable	supplier	part number	Agilent method		in-house method			
			grade purity	amount	supplier	part number	quality	amount
urea	Agilent	5188-6435	Electrophoresis grade	2 x 25.2 g	AMRESCO	0568-1KG	Ultra pure grade >99.9%	1 kg
thiourea	Agilent	5188-6436	ReagentPlus grade >99.9%	2 x 9.1 g	Sigma	T7875-100G	ReagentPlus grade >99%	100 g
DTT	Agilent	5188-6439	Electrophoresis grade	2 x 0.6 g	Sigma	D9779-5G	Molecular biology grade >99%	5 g
50% glycerol	Agilent	5188-6440	GC grade >99%	2 x 10 mL	Merck	4.10015.0500	GR grade >99.5%	500 mL
3-10 OFFGEL buffer	Agilent	5188-6437	n.a.	1 mL	GE Healthcare	17-6000-87	n.r.	1 mL
mineral oil	Agilent	5188-6443	n.a.	100 mL	Sigma	M5904-500ML	Molecular biology grade	500 mL
ddH <sub>2</sub> O	Millipore	n.r.	MilliQ grade	n.r.	Millipore	n.r.	MilliQ grade	n.r.
24 cm 3-10 IPG strip	Agilent	5188-6442	n.r.	8 units	GE Healthcare	12-6002-44	n.r.	12 units
electrode pad	Agilent	5188-6434	n.r.	90 units	Bio-Rad	165-4071	n.r.	500 units
24 well frame	Agilent	G3100-44503	n.r.	6 units	Agilent	G3100-44503	n.r.	6 units
24 well frame lid	Agilent	G3100-47100	n.r.	6 units	Agilent	G3100-47100	n.r.	6 units
tray	Agilent	G3100-60007	n.r.	2 units	Agilent	G3100-60007	n.r.	2 units
electrode	Agilent	G3100-60002	n.r.	2 sets	Agilent	G3100-60002	n.r.	2 sets

n.a. not available, n.r. not relevant

**Table 2.** pH values of OFFGEL fractions. Difference with expected value is indicated in bracket.

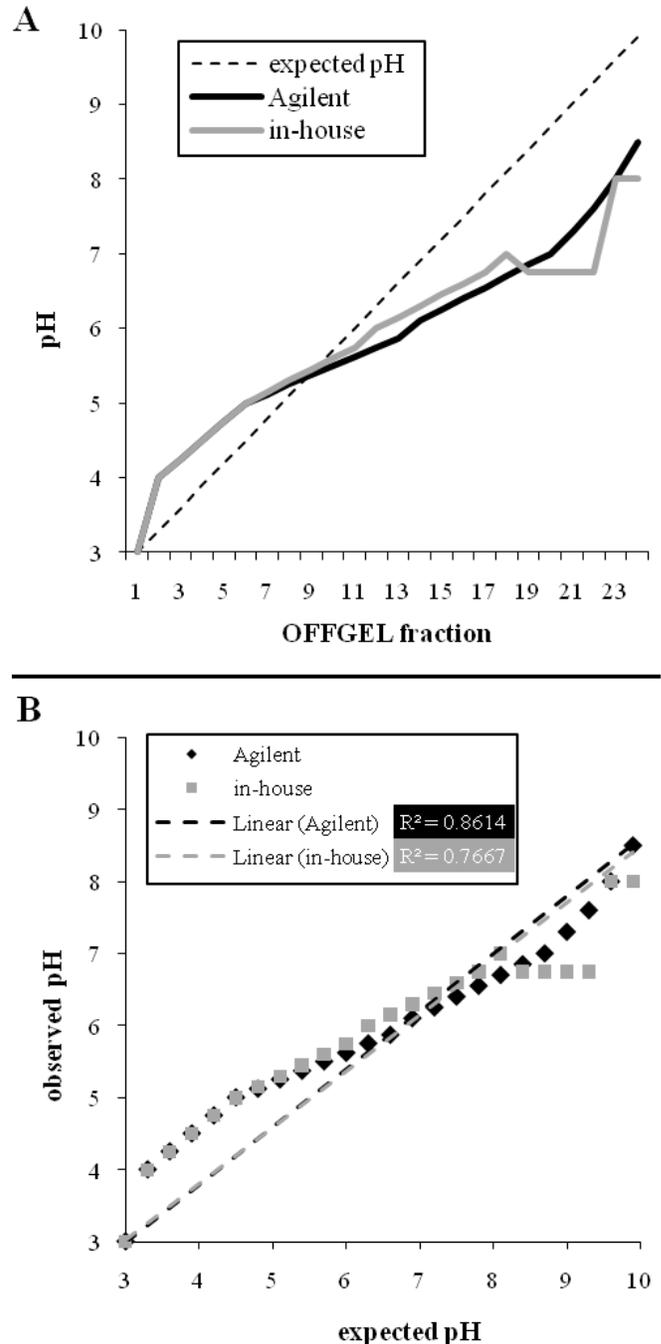
Fraction	Expected pH	Observed pH	
		lamina/Agilent method	lamina/in-house method
1	3.00	3.00 (+0.00)	3.00 (+0.00)
2	3.30	4.00 (+0.70)	4.00 (+0.70)
3	3.60	4.25 (+0.65)	4.25 (+0.65)
4	3.90	4.50 (+0.60)	4.50 (+0.60)
5	4.20	4.75 (+0.55)	4.75 (+0.55)
6	4.50	5.00 (+0.50)	5.00 (+0.50)
7	4.80	5.12 (+0.32)	5.15 (+0.35)
8	5.10	5.25 (+0.15)	5.30 (+0.20)
9	5.40	5.37 (-0.03)	5.45 (+0.05)
10	5.70	5.50 (-0.20)	5.60 (-0.10)
11	6.00	5.62 (-0.38)	5.75 (-0.25)
12	6.30	5.75 (-0.55)	6.00 (-0.30)
13	6.60	5.87 (-0.73)	6.15 (-0.45)
14	6.90	6.10 (-0.80)	6.30 (-0.60)
15	7.20	6.25 (-0.95)	6.45 (-0.75)
16	7.50	6.40 (-1.10)	6.60 (-0.90)
17	7.80	6.55 (-1.25)	6.75 (-1.05)
18	8.10	6.70 (-1.40)	7.00 (-1.10)
19	8.40	6.85 (-1.55)	6.75 (-1.65)
20	8.70	7.00 (-1.70)	6.75 (-1.95)
21	9.00	7.30 (-1.70)	6.75 (-2.25)
22	9.30	7.60 (-1.70)	6.75 (-2.55)
23	9.60	8.00 (-1.60)	8.00 (-1.60)
24	9.90	8.50 (-1.40)	8.00 (-1.90)

(10 wells each, 0.75 x 10 x 8 cm). Runs were performed using a Mini-PROTEAN Tetra Cell (Bio-Rad) system at room temperature for 10 min at 40V followed by 60 min at 180V.

#### 8. Two-dimensional electrophoresis (2-DE)

A total of 100 µg proteins from unfractionated samples were diluted into R solution to reach a volume of 470 µL and then loaded onto 24 cm 4-7 or 7-10 IPG Strips (Bio-Rad) through in-gel rehydration. Fractions were pooled so that their pH values was included into the pH ranges offered by Narrow Range 24 cm ReadyStrip IPG strips (Bio-Rad) as indicated in Table 3. Pooled fractions were diluted into R solution to reach a volume of 470 µL and were loaded onto IPG strips whose pH ranges overlapped theirs through in-gel rehydration. IPG-IEF was performed using IPGphor II (GE Healthcare) and the following program: 0V for 1h after which mineral oil was added, 50V for 12h, 200V for 30min, 500V for 30min, 1000V for 1h, 8000V for 90,000Vhrs. IPG strips equilibration was performed according to Görg *et al.* [29] by incubation into 1% DTT for 15 min following by incubation in 2.5% iodoacetamide for 15 min. Equilibrated IPG strips were transferred onto 12% homecast polyacrylamide/ bisacrylamide (37.5:1) gels (0.1 x 24 x 20 cm) and

sealed in 1% (w/v) agarose (Invitrogen, Mulgrave VIC Australia) in Laemmli [30] running buffer with 0.002 % (w/v) bromophenol blue (BPB). Second dimension was performed using the Ettan DALT six Electrophoresis Unit (GE Healthcare) at 10°C at 40V for 30min followed by 450V for 4h30 (0.4mA, 100W).



**Figure 1.** pH values of the OFFGEL fractions. A) Distributions of observed pH values. The dotted line represents the expected pH distribution. B) Scatterplots of the observed values of the fractions relative to the expected values. Dotted lines represent linear trendlines with intercept set to 3. Corresponding correlation rates (R<sup>2</sup>) are indicated.

**Table 3.** Fraction pooling and IPG-IEF conditions

Sample	OFFGEL method	Fraction pooling	Fraction volume	pH range	Figure number
wheat leaf	Agilent	1 to 5	0.64 mL each	3-6	3A
wheat leaf	Agilent	6 to 11	0.54 mL each	4-7	3B
wheat leaf	Agilent	2 to 13	0.27 mL each	4-7	3C
wheat leaf	Agilent	12 to 17	0.54 mL each	5-8	3D
wheat leaf	Agilent	17 to 24	0.51 mL each	7-10	3E
wheat leaf	in-house	1 to 5	0.64 mL each	3-6	3F
wheat leaf	in-house	6 to 11	0.54 mL each	4-7	3G
wheat leaf	in-house	2 to 13	0.27 mL each	4-7	3H
wheat leaf	in-house	12 to 17	0.54 mL each	5-8	3I
wheat leaf	in-house	17 to 24	0.51 mL each	7-10	3J
wheat leaf	none	none	-	4-7	3K
wheat leaf	none	none	-	7-10	3L
wheat apoplast	in-house	1 to 10	0.32 mL each	3-6	4B
wheat apoplast	in-house	1 to 15	0.20 mL each	4-7	4C
wheat apoplast	in-house	8 to 17	0.32 mL each	5-8	4D
wheat apoplast	in-house	17 to 24	0.45 mL each	7-10	4E
wheat apoplast	none	none	-	4-7	4F
wheat apoplast	none	none	-	7-10	4G

### 9. Gel staining

Both one- and two-dimensional gels were stained following the silver nitrate staining method developed by Rabilloud and Charmont [31]. Stained gels were scanned using Molecular Imaging PhorosFX Plus system (Bio-Rad) at 100  $\mu$ m resolution with the densitometry mode. Images were converted into *TIF*-format (16-bit greyscale, 254 dots per inch).

### 3. Results and Discussion

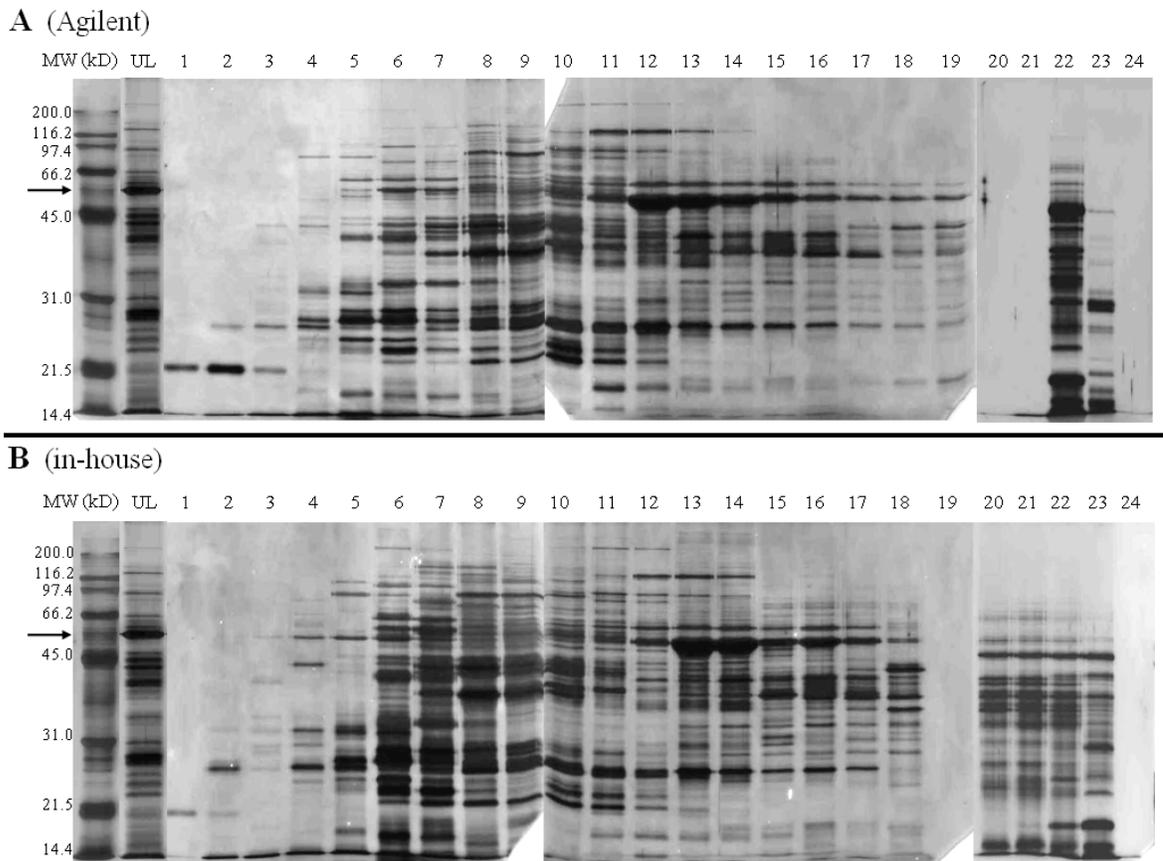
A number of recent proteomics projects incorporating OGE attests to the ever-increasing interest in the use of this novel technology. Despite its growing acceptance, OGE has yet to be reported for plant systems. The first objective of this study was to trial OGE prefractionation of a complex plant protein sample by comparing the manufacturer's method with an in-house protocol, with the final aim of minimising consumable expenses. To this end, using the Agilent 3100 OFFGEL fractionator high resolution mode (24 fractions *per* 24 cm 3-10 IPG strip), two protocols were tested: 1/ the Agilent protocol based on the use of the Agilent's proprietary starter kit, and 2/ our in-house protocol for which all required solutions were prepared using electrophoretic grade consumables routinely used in our laboratory, IPG strips and electrode pads were purchased separately from another supplier. The difference in methods is described in Table 1. The second objective of this study was to protein-enrich two-dimensional patterns from two different plant tissues, whole leaf and leaf apoplastic fluids, using an initial protein pre-fractionation step, compared to two-dimensional patterns obtained from unfractionated samples. Both the Agilent

protocol and our in-house method are assessed and compared using three approaches: pH distribution, one-dimensional profiles and two-dimensional patterns.

#### 1. The Agilent and in-house methods produce comparable electrophoretic patterns.

Leaf samples were prefractionated into 24 fractions along a 3-10 pH range using OGE. Table 2 (graphically represented in Figure 1A) lists the pH values of each OGE fraction and indicates how different they are from the expected values, calculated by dividing the seven pH units of 3-10 gradient by the number of compartments (24). Considering leaf samples, the pH distribution obtained following the Agilent method skews less from the expected distribution than that obtained with our in-house method, differing at the most by 1.7 pH unit (fractions 20 to 22). A consistent trend appears across methods (Agilent and in-house) with the acidic fractions (up to pH 5.60) always displaying a pH value superior to the expected one, while the pH of basic fractions is constantly below the expected value. Basic fractions present the greatest pH difference with the expected distribution, as high as 2.55 pH unit (in-house method, fraction 22). The pH of the fractions distributes linearly as illustrated by the trendlines fitting the scatterplots of the expected pH values *versus* the observed ones (Figure 1B). The Agilent method ( $R^2=0.86$ ) produces slightly more linearly distributed fraction pH from leaf sample than our in-house method ( $R^2=0.77$ ). A linear distribution was also reported in *P. falciparum* system; pH values were mostly above the expected distribution, especially within very acidic fractions [20].

To further assess the methods' efficiency, each fraction was



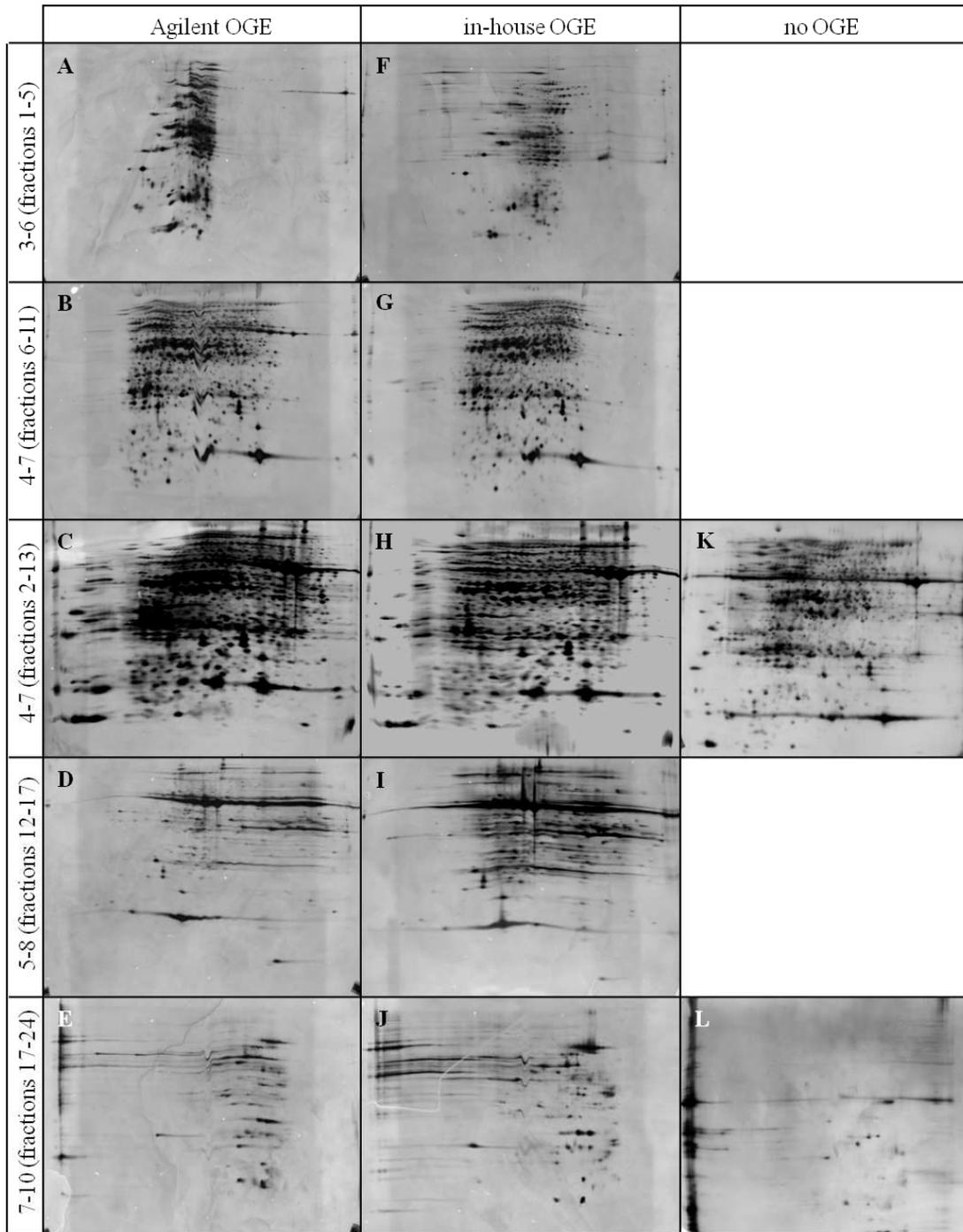
**Figure 2.** One-dimensional profiles of the OFFGEL fractions relative to those of unfractionated samples. A) Leaf sample fractionated according to Agilent's method. B) Leaf sample fractionated according to our in-house method. MW, molecular weight standards; UL, unfractionated leaf sample; 1 to 24, OFFGEL fractions 1 to 24. Arrows indicate the large subunit of RuBisCO.

separated using 1-DE, along with the initial unfractionated sample used as a reference (Figure 2). Comparing one-dimensional profiles of fractions obtained following the Agilent method to those obtained using our in-house protocol, it was evident that both methods produce excellent electrophoretic patterns, displaying well-resolved bands even within the most complex fractions. Proteins bands that are faint in unfractionated leaf samples (lanes UL) become much more prominent in some of the fractions, especially for proteins of very high or very low MW. In both methods (Figure 2A and B), very acidic fractions (1 to 5) are depleted in proteins, whereas acidic to neutral fractions (6 to 16) contain most of the proteins. The greatest difference in method efficiency arises within basic fractions (17 to 24). Using Agilent's protocol, fractions 17 to 19 resolve only few protein bands, and fractions 20, 21 and 24 contain no visible proteins; most of the basic proteins are resolved within fraction 22 and, to a lesser extent, fraction 23 (Figure 2A). Our in-house method proves superior in this alkaline range, with all the fractions displaying many protein bands, except fractions 19 and 24 which contain no visible proteins (Figure 2B).

Protein depletion within fractions of extreme pH was also observed along a 3-10 pH gradient [20]. The most abundant

proteins, such as RuBisCO (large subunit indicated by an arrow in Figure 2), are resolved across several OGE compartments. This was also reported for highly prominent animal proteins such as myosin and actin [19, 21]. The limit of resolving power of OGE when high protein loads are used has been evidenced [32]. Based on the literature, we used settings recommended by the manufacturer to perform the OGE and did not attempt to optimise them. Perhaps plant samples necessitate longer focusing periods and/or higher voltages. Therefore OGE in itself is not resolving enough; however, combined to other analytical techniques such as 2-DE, it greatly improves protein resolution as illustrated below.

The high resolution offered by 2-DE was our third approach to assess which of the two methods employed during OGE was the most efficient in fractionating proteins from wheat leaf samples. In our study, 2-DE was performed using IPG-IEF in the first dimension and SDS-PAGE in the second dimension. For the first dimension, four pH gradients were employed, 3-6, 4-7, 5-8, and 7-10, along 24 cm IPG strips. Following pre-fractionation and pH measurement, fractions were pooled according to the pH ranges covered by the IPG strips (Table 3). Regardless of the pH range, both Agilent (Figure 3A-E left panel) and our in-house (Figure 3F-I),



**Figure 3.** Two-dimensional patterns from leaf samples following either Agilent (A-E) or our in-house (F-I) method along 3-6 (A,F), 4-7 (B-C,G-H,K), 5-8 (D,I) and 7-10 (E,J,L) IPG strips with (A-J) or without (K-L) OGE prefractionation.

middle panel) methods produced excellent two-dimensional patterns bearing many well-resolved spots and a low background noise. A prefractionation using the Agilent method followed by IPG-IEF along a 3-6 pH gradient generated two-dimensional gels displaying incomplete protein focusing as attested by horizontal streaks within medium to high MW range (Figure 3A). All the spots resulting from the in-house method were well-resolved along this acidic gradient (Figure 3F). When fractions 6 to 11 corresponding to 4-7 pH range

were pooled and further resolved by 2-DE, the edges of the two-dimensional patterns became spot-depleted (Figure 3B,G), possibly due to inaccurate pH readings. By pooling more fractions (2 to 13), we were able to fully exploit the whole gel area, even if the background noise (dark areas around the spots) increased (Figure 3C,H). Patterns along 5-8 pH range (Figure 3D,I) present horizontal streaks possibly because it corresponds to the focusing range of the most abundant protein of our sample, RuBisCO. Consistent with

1-DE observations, the in-house method seems to improve protein focusing within neutral to basic ranges as can be seen along pH gradients 5 to 8 (Figure 3I) and 7 to 10 (Figure 3J) relative to the patterns produced using the Agilent method (Figures 3E and 3J, respectively).

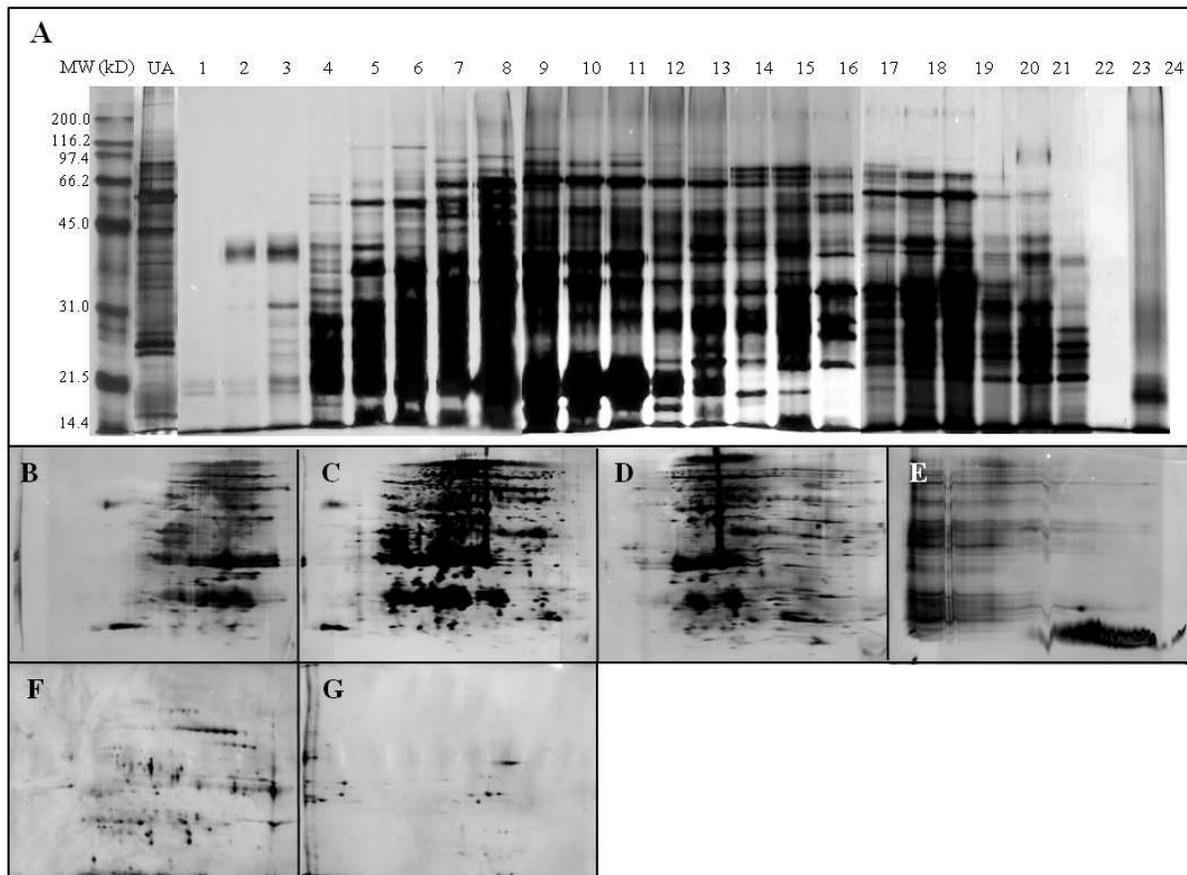
The primary aim of this paper was to provide an alternative to the Agilent kit. To this end, we did not attempt to optimise OGE solutions or program settings. However, we can anticipate improved protein separation through the optimisation of both the composition of the OGE solution and the OGE run. This warrants further investigation. The results presented here are significant as we estimate that our in-house protocol is ten-fold cheaper than Agilent proprietary method. Being perfectly adjusted to the OFFGEL fractionator, light plasticware equipment (electrodes, tray, well frames, and lids) must be purchased from Agilent Technologies. Although disposable, plasticware can be re-used several times provided pieces are thoroughly cleaned (isopropanol for the electrode, 10% SDS for the rest). Acquiring plasticware consumables through the manufacturer and re-using them will satisfy most laboratory's needs. Focusing solutions are very quick, easy and cheap to prepare; a lab familiar with proteomic procedures will already possess all the necessary chemicals, which must be of electrophoresis grade. Mineral oil or alternatively paraffin oil can be obtained through several suppliers. Although we have used electrode pads supplied by Bio-Rad (too wide we had to re-cut them), pads can be self-made using filter paper. We have not tested IPG strips supplied by manufacturers other than GE Healthcare; however because IPG strip format is standard, we anticipate that any manufacturer will provide compatible material. Thus far, the Agilent kit is amenable to OGE fractionation along 4 to 7 and 3 to 10 linear pH ranges only. For the purpose of this study, we have only tested OGE pre-fractionation along a linear 3-10 pH range which suited our subsequent 2-DE steps. Whether home-made or manufactured, linear or not, several pH ranges are available. They can be used during OGE to improve the focusing of proteins with extreme or overlapping pI. Only through the use of an in-house protocol can the end-user benefit from such an array of pH ranges.

## *2. An OGE pre-fractionation steps improves 2-DE focusing and sensitivity and allows recovery of proteins of low abundance in dilute samples*

Initial unfractionated leaf samples were also subject to 2-DE along 4-7 and 7-10 pH ranges and used as a reference (Figure 3K-L). Compared to two-dimensional gels obtained from pre-fractionated samples using OGE, the enrichment in spots and gain in resolution are considerable, particularly within basic ranges (Figure 3). OGE technology relies on soluble- or liquid-phase IEF. To the best of our knowledge, OGE was performed prior to 2-DE only once on human plasma [12]. Introducing a pre-fractionation step through a liquid-phase IEF device prior to 2-DE has proven a very

successful strategy to increase protein coverage and sensitivity in few biological systems such as rodent livers [33, 34], human parasite [35], human cancer [34, 36-38], none dealing with plants. Two-dimensional electrophoresis coupled to mass spectrometry (MS) still constitutes the almost unique platform utilized in plant proteome analysis [39], in particular abiotic stress response [40] which concerns most of plant research. With the gain in MS sensitivity, there is an urgent need in maximising spot resolution on two-dimensional gels in order to achieve the ideal focusing level of one protein *per* spot, as well as to augment the proportion of proteins of low abundance or highly hydrophobic. To date electrophoretic pre-fractionation appears to be one of the best strategies but has yet to be embraced by the plant proteomic community.

Because OGE proved successful on leaf samples, we applied the technology to apoplast fluid recovered from wheat leaves. Apoplast fluid has a low protein content (ten times less than leaf samples), and, being a subcellular fraction of leaves, are less complex than whole leaf samples. It is also rich in non protein components, such as sugars and organic acids [41]. Apoplastic fluid warrants further studies as it is a dynamic compartment involved in key processes such as defense, signalling, and solute reallocation [42]. Taking into account dilution issues, we have opted for a minimal manipulation of apoplastic fluid through a simple lyophilisation in order to minimise protein loss. Lyophilised apoplast samples were then resuspended in the focusing solution. Although unfractionated apoplast samples produce clean complex one-dimensional profile (lane UA in Figure 4A), the fractions obtained following OGE separation are smeared especially when rich in proteins (fractions 4 to 22). We hypothesise that lyophilised apoplastic fluids comprise too many interfering compounds, and might need further cleaning steps, such as dialysis. Nonetheless, a pre-fractionation step greatly helps enriching in proteins of low MW and abundance. This enrichment phenomenon was also reported on dilute fungal secretome samples [43]. Two-dimensional gels from pre-fractionated apoplast samples were produced by pooling more fractions than merely dictated by measured pH values along pH ranges 3-6, 4-7, 5-8, and 7-10 (Table 3 and Figure 4B-E). When compared to unfractionated samples (Figure 4F-G), the enrichment in proteins is obvious despite the high background noise, especially within neutral to basic ranges. Protein content is low in apoplastic fluids, and direct 2-DE without OGE pre-fractionation is only able to resolve the most prominent apoplastic proteins, which are mainly acidic (Figure 4F) since only few spots are visible along basic gradient (Figure 4G). Patterns spanning acidic to slightly basic gradients displayed many proteins (Figure 4B-D); however, very basic apoplastic proteins remained unfocused (Figure 4E). Because protein enrichment driven by OGE is performed indiscriminately, regardless of abundance, prominent proteins reach amounts that fall outside the detection range of silver nitrate staining and become saturated (very dark areas on Figure 4C in particular). This issue could be resolved by depleting the samples of the most abundant



**Figure 3.** Two-dimensional patterns from leaf samples following either Agilent (A-E) or our in-house (F-I) method along 3-6 (A,F), 4-7 (B-C,G-H,K), 5-8 (D,I) and 7-10 (E,J,L) IPG strips with (A-J) or without (K-L) OGE prefractionation.

proteins. Depletion steps combined to OGE were successfully applied to samples characterised by an enormous concentration range [13, 22, 44].

#### 4. Conclusion

For the first time, proteins from plant tissues were pre-fractionated using the 3100 OFFGEL fractionator. Two methods were compared by pre-fractionating proteins from wheat leaves and apoplastic fluids along a 3-10 pH range into 24 fractions. The Agilent method relied on the complete use of the proprietary starter kit which comprised all necessary chemicals and consumables. Our in-house method only differed from the manufacturer's method by the preparation of the focusing solution using our laboratory electrophoresis grade chemicals and the acquisition of IPG strips and electrode pads from other suppliers. The methods were compared and assessed using pH distribution, one- and two-dimensional patterns. Although OGE fractions obtained with the Agilent method better fitted the expected pH distribution than fractions produced using our in-house protocol, it did not affect protein focusing. Indeed both methods yielded excellent electrophoretic profiles of similar quality; spot resolution and number visibly slightly increased with our in-house method. When applied to very dilute samples

such as apoplastic fluids, OGE allowed to recover proteins of low abundance. OGE pre-fractionation should help us discovering novel proteins involved in wheat defense response. We are currently investigating peptide pre-fractionation on those tissues. Such strategy can be indiscriminately applied to any proteomic project, plant scientists should definitely consider it.

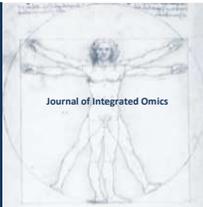
#### Acknowledgements

This research was supported by the Australian Research Council (DP0986139).

#### References

1. P.H. O'Farrell, *J. Biol. Chem.* 250 (1975) 4007-21.
2. B. Bjellqvist, K. Ek, P.G. Righetti, E. Gianazza, A. Gorg, R. Westermeier, W. Postel, *J. Biochem. Biophys. Methods* 6 (1982) 317-39.
3. P.G. Righetti, A. Castagna, B. Herbert, G. Candiano, *Bioscience Reports* 25 (2005) 3-17.
4. P.G. Righetti, A. Castagna, B. Herbert, F. Reymond, J.S. Rossier, *Proteomics* 3 (2003) 1397-1407.
5. A. Ros, M. Faupel, H. Mees, J. Van Oostrum, R. Ferrigno, F. Reymond, P. Michel, J.S. Rossier, H.H. Girault, *Proteomics* 2 (2002) 151-156.

6. P. Michel, F. Reymond, I.L. Arnaud, J. Josserand, H.H. Girault, J.S. Rossier, *Electrophoresis* 24 (2003) 3-11.
7. I.L. Arnaud, J. Josserand, H. Jensen, N. Lion, C. Roussel, H.H. Girault, *Electrophoresis* 26 (2005) 1650-1658.
8. C.D. Meert, L.J. Brady, A. Guo, A. Balland, *Anal. Chem.* 82 (2010) 3510-3518.
9. M. Heller, M. Ye, P.E. Michel, P. Morier, D. Stalder, M.A. Jünger, R. Aebersold, F. Reymond, J.S. Rossier, *J. Prot. Res.* 4 (2005) 2273-2282.
10. P.E. Michel, D. Crettaz, P. Morier, M. Heller, D. Gallot, J.D. Tissot, F. Reymond, J.S. Rossier, *Electrophoresis* 27 (2006) 1169-1181.
11. P. Hörth, C.A. Miller, T. Preckel, C. Wenz, *Mol. Cell. Proteomics* 5 (2006) 1968-1974.
12. M. Heller, P.E. Michel, P. Morier, D. Crettaz, C. Wenz, J.D. Tissot, F. Reymond, J.S. Rossier, *Electrophoresis* 26 (2005) 1174-1188.
13. D. Stalder, A. Haerberli, M. Heller, *Proteomics* 8 (2008) 414-424.
14. J. Chenau, S. Michelland, J. Sidibe, M. Seve, *Proteome Sci.* 6 (2008)
15. E. Ernoult, A. Bourreau, E. Gamelin, C. Guette, *J. Biomed. Biotechnol.* 2010 (2010) 927917.
16. C.M. Warren, D.L. Geenen, D.L. Helseth Jr, H. Xu, R.J. Solaro, *J. Proteomics* 73 (2010) 1551-1561.
17. E. Ernoult, E. Gamelin, C. Guette, *Proteome Sci.* 6 (2008)
18. K. Nagano, T. Shinkawa, N. Yabuki, H. Mutoh, N. Inomata, Y. Watanabe, M. Ashihara, S. Nagahashi, N. Ishii, Y. Aoki, M. Haramura, *J. Proteomics* (2010)
19. S.B. Scruggs, R. Reisdorph, M.L. Armstrong, C.M. Warren, N. Reisdorph, R.J. Solaro, P.M. Buttrick, *Mol. Cell. Proteomics* 9 (2010) 1804-1818.
20. R.D.M. O'Cuailain, J.E. Hyde, P.F.G. Sims, *Malaria J.* 9 (2010) 286.
21. M.A. Sentandreu, P.D. Fraser, J. Halket, R. Patel, P.M. Bramley, *J. Prot. Res.* 9 (2010) 3374-3383.
22. A.R. Vaezzadeh, A.C. Briscoe, H. Steen, R.S. Lee, *J. Prot. Res.* 9 (2010) 6082-6089.
23. L.N. Waller, K. Shores, D.R. Knapp, *J. Prot. Res.* 7 (2008) 4577-4584.
24. Y. Zhang, Y. Li, F. Qiu, Z. Qiu, *Electrophoresis* 31 (2010) 3797-3807.
25. S. Fraterman, U. Zeiger, T.S. Khurana, N.A. Rubinstein, M. Wilm, *Proteomics* 7 (2007) 3404-3416.
26. J. Mulvenna, B. Hamilton, S.H. Nagaraj, D. Smyth, A. Loukas, J.J. Gorman, *Mol. Cell. Proteomics* 8 (2009) 109-121.
27. P.S. Solomon, R.P. Oliver, *Planta* 213 (2001) 241-9.
28. C. Damerval, D. de Vienne, M. Zivy, H. Thiellement, *Electrophoresis* 7 (1986) 52-4.
29. Görg, W. Postel, S. Gunther, C. Friedrich, *Electrophoresis* 9 (1988) 57-9.
30. U.K. Laemmli, *Nature* 227 (1970) 680-5.
31. T. Rabilloud, S. Charmont, in: T. Rabilloud (Eds), *In Proteome research: Two-dimensional gel electrophoresis and identification methods*, Springer, Berlin, New York, 2000, pp.107-26.
32. N.C. Hubner, S. Ren, M. Mann, *Proteomics* 8 (2008) 4862-4872.
33. G. Gazzana, J. Boriak, *J. Prot. Res.* 6 (2007) 3143-51.
34. K. Lee, K. Pi, *Anal. Bioanal. Chem.* 396 (2010) 535-539.
35. R.K. Brobey, L. Soong, *Proteomics* 7 (2007) 116-20.
36. B. Seshi, *Proteomics* 7 (2007) 1984-99.
37. X. Zuo, D.W. Speicher, *Proteomics* 2 (2002) 58-68.
38. K.B. Lee, K.B. Pi, *J. Sep. Science* 32 (2009) 1237-1242.
39. J.V. Jorrín, A.M. Maldonado, M.A. Castillejo, *Proteomics* 7 (2007) 2947-2962.
40. D. Vincent, M. Zivy, in: J. Samaj and J.J. Thelen (Eds), *Plant proteomics*, Springer, Heidelberg, 2007, pp.386-94.
41. A.F. Lopez-Millan, F. Morales, A. Abadia, J. Abadia, *Plant Physiol.* 124 (2000) 873-84.
42. R.P. Haslam, A.L. Downie, M. Raveton, K. Gallardo, D. Job, K.E. Pallett, P. John, M.A.J. Parry, J.O.D. Coleman, *Ann. Appl. Biol.* 143 (2003) 81-91.
43. D. Vincent, M.H. Balesdent, J. Gibon, S. Claverol, D. Lapaillerie, A.M. Lomenech, F. Blaise, T. Rouxel, F. Martin, M. Bonneu, J. Amselem, V. Dominguez, B.J. Howlett, P. Wincker, J. Joets, M.H. Lebrun, C. Plomion, *Electrophoresis* 30 (2009) 4118-36.
44. J.A. Burgess, P. Lescuyer, A. Hainard, P.R. Burkhard, N. Turck, P. Michel, J.S. Rossier, F. Reymond, D.F. Hochstrasser, J.C. Sanchez, *J. Prot. Res.* 5 (2006) 1674-1681.



ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v1i2.36

## Quantitative Proteomic Analysis of Retina in Oxygen-Induced Retinopathy Mice using iTRAQ with 2D NanoLC-nanoESI-MS/MS

Lei Zhou<sup>1,2,a</sup>, Liu Xinling<sup>3,a</sup>, Siew Kwan Koh<sup>1</sup>, Li Xiaorong<sup>3\*</sup>, Roger W Beuerman<sup>\*1,2</sup>.

<sup>1</sup>Singapore Eye Research Institute, Singapore; <sup>2</sup>Department of Ophthalmology, National University of Singapore, Singapore; <sup>3</sup>Tianjin Medical University Eye Center, China. <sup>a</sup>These authors contributed equally.

Received: 01 September 2011 Accepted: 22 March 2011 Available Online: 22 March 2011

### ABSTRACT

Analysis of the retina proteome during hypoxia-induced retinal neovascularization (RN) process may be helpful to elucidate pathogenesis of related diseases, such as diabetic retinopathy (DR) and retinopathy of prematurity (ROP). Retinal neovascularization was induced in 7-day old, C57BL/6 mice by exposure to 80% oxygen for 5 days followed by 5 days in room air. Retinas from mice at postnatal day 17 from control and oxygen-induced retinopathy (OIR) groups were used for proteomic analysis. We have employed a quantitative proteomic approach, iTRAQ (isobaric Tagging for Relative and Absolute protein Quantification) coupled with 2D nanoLC-nanoESI-MS/MS to quantitatively compare the relative changes in the retina proteome from control and OIR mice. In total, 264 protein groups were identified at a 95% confidence level. Among them, OIR induced significant changes in 28 proteins (14 up-regulated and 14 down-regulated). Obvious changes include the up-regulation of a few plasma proteins (i.e. serum albumin, hemoglobin), indicating the breakdown of the blood-retina barrier. Vimentin, ribosomal proteins, some proteases, neural cell adhesion molecule (NCAM) 180, retinoschisin were found to be up-regulated and several crystallins such as isoform 1 of  $\alpha$  crystallin A chain, isoform 2 of  $\alpha$  crystallin A chain,  $\alpha$  crystallin B chain,  $\gamma$  crystallin D and  $\beta$ -A3/A1 crystallin were down-regulated. The iTRAQ result of  $\alpha$  crystallin B chain was also verified by Western blot analysis. The proteomic results from this study provide new avenues for understanding the pathogenesis of OIR induced retinal neovascularization and related retina diseases.

**Keywords:** Retina neovascularization; Oxygen-induced retinopathy; Retina proteome; iTRAQ; 2D nano LC-nanoESI-MS/MS.

### Abbreviations

**iTRAQ:** isobaric tagging for Relative and Absolute protein Quantification; **OIR:** oxygen-induced retinopathy; **ROP:** retinopathy of prematurity; **VEGF:** vascular endothelial growth factor; **SCX:** strong cation exchange column; **PBS:** phosphate buffered saline; **MS/MS:** tandem mass spectrometry; **GO:** gene ontology; **FITC:** fluorescein isothiocyanate; **ESI:** electroSpray Ionization; **2D nano LC:** two-dimensional nano-liquid chromatography; **RN:** retinal neovascularization.

### 1. Introduction

Retina neovascularization is the hallmark of proliferative retinopathies, such as diabetic retinopathy (DR), retinopathy of prematurity (ROP), central and branch retina vein occlusion (CRVO and BRVO), all of which constitute primary causes of blindness. Inappropriate proliferation of vessels derived from preexisting retinal vessels results from hypoxic

conditions including non-perfusion of the retina or a decrease in oxygen tension, which may lead to more serious complications, such as vitreous hemorrhage, tractional retinal detachment, and pre-retinal membrane formation [1, 2].

Because of the difficulty of obtaining human retina samples, an animal model of retina neovascularization is of criti-

\*Corresponding author: Prof. Roger W Beuerman, Singapore Eye Research Institute, 11 Third Hospital Avenue, Singapore 168751. E-mail Address: rwbeuer@mac.com. Prof. Li Xiaorong, Tianjin Medical University Eye Center, 64 Tongan Road, Heping district, Tianjin, China 300070. Fax: +86-22-23346434. E-mail Address: xiaorli@163.com.

cal importance. The mouse model of oxygen-induced retinopathy (OIR) is widely used to induce retinal neovascularization and serves as a model of the childhood disease of retinopathy of prematurity [3,4]. Neonatal C57BL/6 mice are exposed to a constant high (75%) oxygen level from postnatal day 7 to 12. By postnatal day 17, after five days of recovery in room air, revascularization of a central avascular area is seen with associated marked retinal neovascularization at the border between the central avascular and peripheral vascular retina [5]. Similar to proliferative diabetic retinopathy (PDR) and ROP, OIR is characterized by hypoxia-induced retinal neovascularization.

Retinal neovascularization is suggested to develop as result of an imbalance between angiogenic stimulators and inhibitors, which involves numerous growth factors, inhibiting factors, extracellular matrix proteins, vascular endothelial cells and pericytes [6, 7]. Vascular endothelial growth factor (VEGF), insulin-like growth factor (IGF), fibroblast growth factor (FGF), placental growth factor (PIGF), integrins, transforming growth factor (TGF), tissue factor (TF), and VE-cadherin all promote angiogenesis. On the other hand, pigment epithelium derived factor (PEDF), arrestin, endostatin, TIMPs, angiostatin may function as anti-angiogenic factors [6, 8-12]. Studies using this well-known model of angiogenesis with analysis or the retinal proteome may lead to the identification of novel drug targets for the inhibition of retina neovascularization. Blocking VEGF functions by the use of intra-vitreous injections of VEGF neutralizing antibodies, VEGF receptor antibodies, or VEGF receptor chimeric proteins have achieved clinical success [13]. Other anti-angiogenic approaches include anti-integrin, anti-IGF-1 and anti-proteinase antibodies [14] have also been considered. However, complete prevention of retinal neovascularization does not always occur. There still are many unknown features about the protein changes and the interactions of these factors in the process of RN.

In recent years, proteomics has been used to analyze the protein during the development of the mouse retina, mouse photoreceptor sensory cilium complex and a mouse model of retinal photoreceptor degeneration [15-17]. Additional studies have begun to examine the proteomic analysis of various retina diseases, such as age-related degeneration, diabetic retinopathy and retinal degeneration using rat or human samples [18-21]. However, there is no report on the differential proteomic analysis of mouse OIR.

In this paper, we describe the quantitative proteomic analysis of the mouse retina after varying oxygen tension to induce retinal neovascularization to retinas from normal mice using 2D nanoLC-nanoESI-MS/MS combined with iTRAQ (isobaric Tagging for Relative and Absolute protein Quantification). We have uncovered novel proteins, which may have mechanistic significance for understanding retinal neovascularization.

## 2. Materials and methods

### *Rodent Model of ROP*

All experimental procedures and use of animals followed the protocol approved by the Tianjin Medical University Animal Research Committee and were in accordance with the ARVO (The Association for Research in Vision and Ophthalmology) Statement for the Use of Animals in Ophthalmic and Vision Research.

Forty female C57BL/6 mice were provided by the Experimental Animal Center of the Chinese Science Academy. Mothers and pups received standard laboratory diet and water ad libitum. Light was cycled on a 12-hour-on, 12-hour-off, and the room was maintained at approximately 21°C. Oxygen-induced retinopathy was developed in mice pups according to a protocol previously established for the mouse model of ROP [5]. In brief, at postnatal day 7, mice pups and their nursing mothers were exposed to hyperoxic conditions (80% oxygen) for 5 days in an infant incubator. On postnatal day 12, the pups were returned to room-air (normoxic conditions) for 5 days until postnatal day 17. Control groups were age-matched animals maintained in room air for the duration of the experiment.

Neo-vascularization onset and severity were measured in retinas using fluorescein-labeled dextran ( $2 \times 10^6$  average MW; Sigma, USA) staining. At the conclusion of the experiments, animals were anesthetized and the right atrium used as the route to perfuse the right ventricle with 1 ml phosphate-buffered saline containing 50 mg/mL fluorescein-labeled dextran. After fixing for 1 hour in 10% phosphate-buffered formalin, the entire retina was carefully dissected from the posterior eye cup, radially cut from the edge of the retina to the equator in all four quadrants. Histology was carried out to determine the presence of neovascularization, briefly, 20 hematoxylin and eosin-stained serial sections (ten on each side of the optic nerve) per eye, at an interval of 20  $\mu$ m, were examined in masked fashion for the presence of neovascular buds projecting into the vitreous from the retina. In total 34 eyes from both control group (17 eyes) and OIR group (17 eyes) were assessed using a retinal scoring system [22]. In that scoring system, blood vessel growth, blood vessel tufts, extra retinal neovascularization, central vasoconstriction, retinal hemorrhage, and tortuosity of the blood vessels were scored (0 to 4) according to certain criteria [22]. The neovascular score was defined as the mean number of neovascular tufts per section. The mean scores and standard deviations (SDs) for the control and OIR groups were calculated and statistical analysis (student t-test) was performed to determine the group differences.

### *Extraction of retinal protein*

At postnatal day 17, 40 mice including 20 OIR treated mice and 20 age-matched control mice were sacrificed by cervical dislocation. Retinas were separated from retinal pigment epithelium (RPE), after which 1.5 times the volume of

lysis buffer was added (20% Sucrose, Tris acetate 20mM, MgCl<sub>2</sub> 2mM, Glucose 10mM, 2% CHAPS, pH7.2). Samples were homogenized at 4°C for 30 min and centrifuged for 10 min at 6,000 g and the supernatant was kept. Total protein concentration was measured on a Model 550 Microplate Reader using a Micro BCA™ Protein Assay Kit (Pierce, USA).

#### *iTRAQ labeling of retina proteins*

For quantitative proteomic analysis, 20 OIR retinal samples were pooled to form OIR group and 20 control retinal samples were pooled to form the control group.

Briefly, 50µg total protein from both control and OIR groups were digested in parallel with trypsin and then labeled with one of the four-iTRAQ reagents following the manufacturer's instructions (Applied Biosystems, Foster City, CA), in the following manner: 114 and 115 reagent for age-matched control (duplicated), 116 and 117 for OIR (duplicated). After labeling with iTRAQ, these four samples were combined (a total of 200µg protein) and subjected to 2D nanoLC-nanoESI-MS/MS (Dionex, Sunnyvale, CA and Applied Biosystems, Foster City, CA) analysis.

#### *Proteomic analysis by 2D nanoLC-nano-ESI-MS/MS*

Two-dimensional nanoLC-nanoESI-MS/MS was used for the analysis. The first dimension was a SCX column (300 mm i.d. x 10 cm porosity 10S SCX, Dionex) with 10 steps of salt plug (20µl injection) elutions (10mM, 20 mM, 30 mM, 40 mM, 50 mM, 75 mM, 100mM, 250mM, 500mM and 1000mM ammonium acetate) all at a flow rate of 30µl/min and using a loading solvent of 0.1% formic acid/ACN (95:5, v:v). The second dimension was RP separation. The 10 cm x 75 µm i.d. microcapillary LC column was self-packed using PicoFrit (New Objectives, Woburn, MA). This column had an integrated spray tip which was directly coupled with the nano-spray interface (Protana, Odense, Denmark) into Applied Biosystems' QSTAR mass spectrometer. The packing material was Luna C18, 3µm, 100 Å (Phenomenex Torrance, CA). Samples were loaded onto a trapping cartridge (C18, 0.3 x 5 mm, from Dionex) from Famos autosampler (Dionex) at 30 µL/min. After a 5 min wash with acetonitrile/water (2/98, v/v with 0.1% formic acid), the system was switched (Switchos, Dionex) into line with the C18 analytical capillary column. Using an Ultimate solvent delivery system (Dionex), a linear gradient of acetonitrile (0.1% formic acid) from 20% to 95% over 85 min at flow rate of 300 nL/min was used to analyze. Parameters for the nanospray and other instrumentation were set as follows: ionspray voltage (IS) = 2200 V, curtain gas (CUR) = 20, declustering potential (DP) = 60 V, focusing potential (FP) = 265 V, collision gas setting (CAD) = 5 for nitrogen gas, DP2=15. All data was acquired using information-dependent acquisition (IDA) mode with Analyst QS software (Applied Biosystems). TOF-MS survey scan parameters were set as follows: 1 sec TOF MS survey

scan in the mass range of 300~1200 Da followed by two product ion scans of 3 sec each in the mass range of 100~1500 Da. The "enhance all" function was used in the IDA experiments. Switching criteria were set to ions greater than  $m/z = 350$  and smaller than  $m/z = 1200$  with charge state of 2 to 4 and an abundance threshold of > 20 counts/s. Former target ions were excluded for 60s. IDA collision energy (CE) parameter script was used for automatically controlling the CE.

#### *Database searching*

The data output from the MS/MS was processed and searched against the International Protein Index (IPI, Version 3.01) protein database, under mouse taxonomy, using ProteinPilot™ software (Applied Biosystems, Foster City, CA). All protein identification was based on the criteria of Unused ProtScore > 1.3 (95% confidence). The Unused ProtScore is a measurement of all the peptide evidence for a protein that is not better explained by a higher-ranking protein. It is the true indicator of protein confidence. Protein quantitation was also carried out using ProteinPilot software (Applied Biosystems). Protein quantitation was achieved by averaging iTRAQ ratios of all peptides identified; normalization, using a Gaussian distribution with median of 1 when all peptides were considered between control and experimental groups, was performed after iTRAQ ratios were calculated.

#### *Western Blot*

Protein concentration was determined using BCA method with bovine serum albumin as standard. Protein samples (20 µg of total protein for each sample) were subjected to sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE) using 15% SDS – polyacrylamide gels, followed by electrophoretic blotting (overnight at 0.02A) onto nitrocellulose membranes. The blots were blocked with 5% skim milk in Tris-buffered saline – 0.2% Tween 20 for 1 hr. The blots were probed with primary polyclonal antibodies against  $\alpha$  crystallin B chain (using dilution of 1:500), for 3 mouse control retina samples and 3 mouse OIR retina samples for 3 hrs. It was followed by incubation with the corresponding horseradish peroxidase-conjugated secondary antibodies (using dilution of 1:10000). Actin protein expression was detected on the same membranes as a loading control. Specific protein bands were detected with Immobilon Western Chemiluminescent HRP Substrate (Millipore) using the enhanced chemiluminescence western blot analysis system (Bio-Rad, USA). The membrane was exposed to x-ray films and developed (Fujifilm).

### **3. Results**

#### *Retina Histopathology*

Representative retinal fluorescein-labeled dextran stained

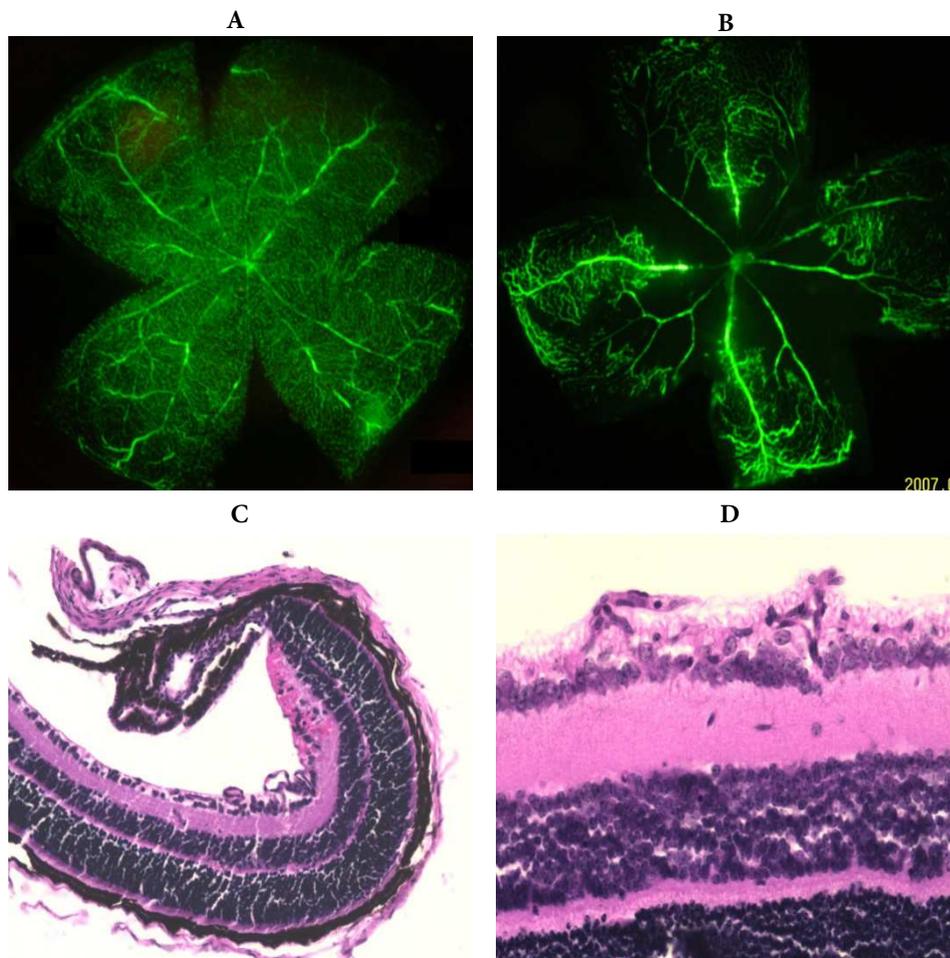
flatmounts are illustrated in Figure 1A. At postnatal day 17, the OIR group exhibited the characteristic central hypoperfusion area with neovascular tufts extending into the vitreous which were not seen when compared with age-matched controls (Fig. 1B). In paraffin cross-sections of OIR treated mouse retinas, large clusters of blood vessels were adherent to the internal limiting membrane (ILM), and many blood vessels were observed in the inner retina. In some cases, retinal hemorrhage in the peripheral retina was also noticeable (Fig. 1C and 1D). The neovascular scores for control group and OIR group at postnatal day 17 were  $0.23 \pm 0.18$  and  $28.24 \pm 2.51$  ( $p < 0.001$ ).

#### Mouse retina proteome

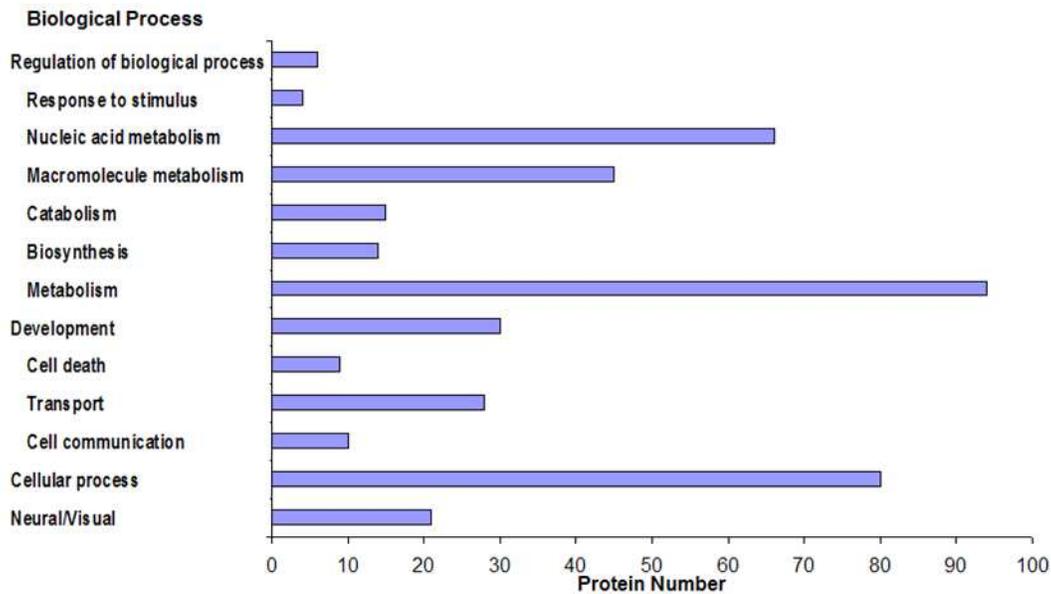
In total 2912 distinct peptides (peptide confidence level: 95%) were identified which matched 1041 proteins before grouping. After grouping, 264 protein groups in mouse reti-

na were identified from the iTRAQ experiments (Table 1 – Supplementary File). The threshold we used for identification of proteins was 95% confidence (or unused ProtScore > 1.3) in ProteinPilot software. Sequence coverage (%) and number of unique peptides for each protein group are also listed in the Table 1. Of the 261 proteins with quantitative information, 218 proteins have more than one peptide contributing to the quantitative results. In many cases, same peptide fragment could be acquired repeatedly by mass spectrometer and contributed to the average protein ratio.

Gene ontology (GO) annotations were analyzed using online tools, CateGORizer [23]. Based on the GO classification, among these 264 proteins, there are 209 proteins which have at least one identifier under “biological process” (Figure 2). Genes representing other functional groups of identified proteins belonged to cell metabolism (35.6%,  $n=94$ ), cellular process (30.0%,  $n=80$ ), development (11.4%,  $n=30$ ) and neuronal/visual (8.0%,  $n=21$ ) respectively. Some proteins had



**Figure 1.** Fluorescein-dextran perfused retinas of mice exposed to room air or hyperoxia. (A) the retina from P17 room air-raised mouse showing the normal pattern of retina vessels. (B) retina from P17 mouse exposed to 80% oxygen from P7 to P12. (C) and (D) 10  $\mu$ m paraffin cross-sections of C57BL/6 mice retinas stained with PAS and hematoxylin. Arrows indicate neovascular tufts extending into the vitreous. Retinal hemorrhage in the peripheral retina. V vitreous; ILM internal limiting membrane; GCL, ganglion cell layer; IPL inner plexiform layer.



**Figure 2.** Gene Ontology (GO) annotation under “biological process”. X-axis represents the number of proteins under each subcategory.

more than one GO identifier.

#### *Protein profile changes in OIR mice retina*

Relative quantification of the identified peptides was based on the individual ratios from signature ion ( $m/z = 114, 115, 116, \text{ or } 117$ ) peak areas of the iTRAQ reagent tags of the identified peptides from OIR samples compared with those of the age-matched control group. The iTRAQ ratios of all peptides for each protein were grouped, averaged and were expressed as ratios at the protein level. In this study, control and OIR groups were duplicated, i.e. iTRAQ reagents 114 and 115 were used to label two duplicated control samples and iTRAQ reagents 116 and 117 were used to label two duplicated OIR samples. If we used a 20% increase or decrease for iTRAQ ratios as the cutoff, the levels of 14 proteins were found to be increased in OIR group and the levels of 14 proteins decreased in OIR group. The list of proteins found to be increased or decreased by 20% between the OIR and control groups is seen in Table 2. Figure 3 shows two representative MS/MS spectra of peptide fragments originating from vimentin (IPI00227299.5) and  $\alpha$  crystallin B (IPI00138274.1), and relative quantitation in the mass range for the reporter ions ( $m/z = 114, 115, 116, \text{ or } 117$ ).

#### *Verification by Western Blot Analysis*

One of the proteins,  $\alpha$  crystallin B chain, showed significant down-regulated in OIR group as indicated by iTRAQ was subsequently verified using Western blot analysis. Figure 4 shows Western blot results of 3 control and 3 OIR retina samples. Compared with control retina samples, the levels of  $\alpha$  crystallin B chain were significantly decreased in OIR treated retina samples.

#### **4. Discussion**

The aim of this study was to obtain a mouse retinal proteome that responded to changes in oxygen tension causing pathological neo-vascularization. Previous studies [24] and our study shows that neovascularization was most prominent at P17. Thus, we used P17 tissues from both normal and OIR mice for the quantitative proteomic analysis. Pooled sample were used in this study for several reasons, including reducing biological variation, saving analysis time and cost, better suitability for small tissue size and sample volume. However, the information for individual biological variation is missing.

We have identified 264 protein groups (Unused ProtScore  $>1.3$ , 95% confidence) in P17 normal and OIR mice retina using 2D nanoLC-nanoESI-MS/MS. Gene ontology (GO) classification showed that the three largest functional categories are related to metabolism, cellular process and development. Twenty-one proteins are associated with neuronal/visual (8%). Among these 264 identified proteins, the levels of 14 proteins were found to be at least 20% higher in the OIR group when compared to the control group, whereas, the levels of 14 other proteins were at least 20% lower in the OIR group compared to the control group.

#### *Proteins increased in OIR*

As expected, some plasma proteins including hemoglobin and serum albumin are significantly increased in the OIR retina which may be due to breakdown of the blood-retinal barrier or an increase in leaky retinal capillaries. Retinal hemorrhage was observed in the peripheral retina of OIR mice (Fig.1c), which can directly increase the content of blood constituent proteins. Other proteomic studies showed

**Table 2.** Differential expression of retina proteins in oxygen-induced retinopathy (OIR) compared to age-matched control group.

N	Accession #	Name	Ratio 1	Ratio 2	Avg. Ratio	Molecular function	Biological Process	Cell component
1	IPI:IPI00469114.4	Hemoglobin subunit alpha	2.215	2.126	2.170	erythrocyte development / in utero embryonic development/ oxygen transport	heme binding/ iron ion binding/ oxygen binding/ oxygen transporter activity	hemoglobin complex
2	IPI:IPI00465823.3	similar to 60S ribosomal protein L18	1.982	1.996	1.989	Structural constituent of ribosome	translation	Intracellular/ ribosome
3	IPI:IPI00122971.1	Isoform N-CAM 180 of Neural cell adhesion molecule 1, 180 kDa isoform precursor	1.407	1.432	1.420	Protein binding/ heparin binding	Cell adhesion/cell surface receptor linked signal transduction/ positive regulation of calcium-mediated signaling/ homotypic cell-cell adhesion	anchored to membrane/ integral to membrane/ external side of plasma membrane/ axon
4	IPI:IPI00775948.1	60S ribosomal protein L7	1.509	1.285	1.392	RNA binding/ structural constituent of ribosome/ transcription regulator activity	translation	Ribosome/ cytosolic large ribosomal subunit
5	IPI:IPI00131695.3	Serum albumin precursor	1.413	1.332	1.372	DNA binding /drug binding / protein binding / pyridoxal phosphate binding /toxin binding	cellular response to starvation / hemolysis by symbiont of host red blood cells / maintenance of mitochondrion localization/ negative regulation of apoptosis	cytoplasm/protein complex/
6	IPI:IPI00227299.5	Vimentin	1.378	1.361	1.370	protein binding / structural molecule activity	intermediate filament-based process	cytoplasm / type III intermediate filament
7	IPI:IPI00785308.1	Prss3 protein (Fragment)	1.272	1.360	1.315	serine-type endopeptidase activity	proteolysis	
8	IPI:IPI00676858.1	similar to 40S ribosomal protein SA	1.276	1.342	1.309			
9	IPI:IPI00404551.1	B6-derived CD11 +ve dendritic cells cDNA, RIKEN full-length enriched library, clone:F730002E02 product: cathepsin D, full insert sequence	1.348	1.239	1.293	cathepsin D activity	proteolysis	lysosome / mitochondrion
10	IPI:IPI00130252.1	Retinoschisin precursor	1.222	1.363	1.290	Protein binding	Cell adhesion	Extracellular region
11	IPI:IPI00169463.1	Tubulin beta-2C chain	1.245	1.329	1.286	GTP binding/ GTPase activity/ structural molecule activity	Microtubule-based process/ protein polymerization	tubulin complex/ protien complex/ microtubule
12	IPI:IPI00459374.3	similar to Carbonyl reductase [NADPH] 1	1.230	1.210	1.220			
13	IPI:IPI00471441.1	Ptms protein	1.215	1.204	1.210	translation initiation factor activity	translational initiation	
14	IPI:IPI00230394.4	Lamin-B1	1.209	1.201	1.205	structural molecule activity		lamin filament
15	IPI:IPI00756474.1	similar to Splicing factor, arginine/serine-rich 3	0.781	0.800	0.790			
16	IPI:IPI00626755.4	9 kDa protein	0.757	0.799	0.778	nucleus/ DNA binding		chromatin/ cytoplasm
17	IPI:IPI00123313.1	Ubiquitin-activating enzyme E1 X	0.761	0.790	0.776	ATP binding/ ligase activity/ small protein activating enzyme activity	protein modification process/ ubiquitin-dependent protein catabolic process	
18	IPI:IPI00403058.1	Isoform 1 of Neutral alpha-glucosidase AB precursor	0.779	0.7288	0.754	glucan 1,3-alpha-glucosidase activity/ protein binding	N-glycan processing	Golgi apparatus/ alpha-glucosidase II complex/ melanosome
19	IPI:IPI00415403.1	Isoform 2 of Syntaxin-binding protein 1	0.749	0.744	0.746	protein binding	Protein transport/synaptic vesicle maturation/vesicle docking during exocytosis	mitochondrion
20	IPI:IPI00776358.1	acetyl-Coenzyme A acetyltransferase 3	0.741	0.737	0.739	transferase activity	metabolic process	

21	IPI:IPI00113536.2	Anp32b protein	0.765	0.706	0.735	protein binding		nucleus
22	IPI:IPI00606097.1	transmembrane protease, serine 13	0.578	0.493	0.534	scavenger receptor activity/serine-type endopeptidase activity	proteolysis	integral to membrane
23	IPI:IPI00113845.1	Proteasome subunit beta type 1 precursor	0.340	0.697	0.486	threonine-type endopeptidase activity	ubiquitin-dependent protein catabolic process	cytosol/ nucleus/ proteasome core complex
24	IPI:IPI00134845.1	Beta-A3/A1 crystallin protein	0.294	0.325	0.309			
25	IPI:IPI00420923.2	Gamma crystallin D	0.286	0.262	0.274		eye development	
26	IPI:IPI00138274.1	Alpha crystallin B chain	0.231	0.2451	0.238	structural constituent of eye lens	camera-type eye development /muscle development /transmembrane receptor protein tyrosine kinase signaling pathway	insoluble / fractionplasma membrane /soluble fraction /Z disc
27	IPI:IPI00108737.1	Isoform 1 of Alpha crystallin A chain	0.203	0.217	0.210			
28	IPI:IPI00109729.1	Isoform 2 of Alpha crystallin A chain	0.136	0.105	0.120		camera-type eye development/M phase specific microtubule process	Cellular component: cytoplasm /

that the vitreous body in proliferative diabetic retinopathy had higher levels of albumin, hemopexin, and IgG, IgA, complement C3, C4 and other blood constituent proteins due to the breakdown of the blood-retina barrier [25-27]. In our study, the surrounding tissues were carefully rinsed with cold saline solution to avoid the possibility of blood contamination.

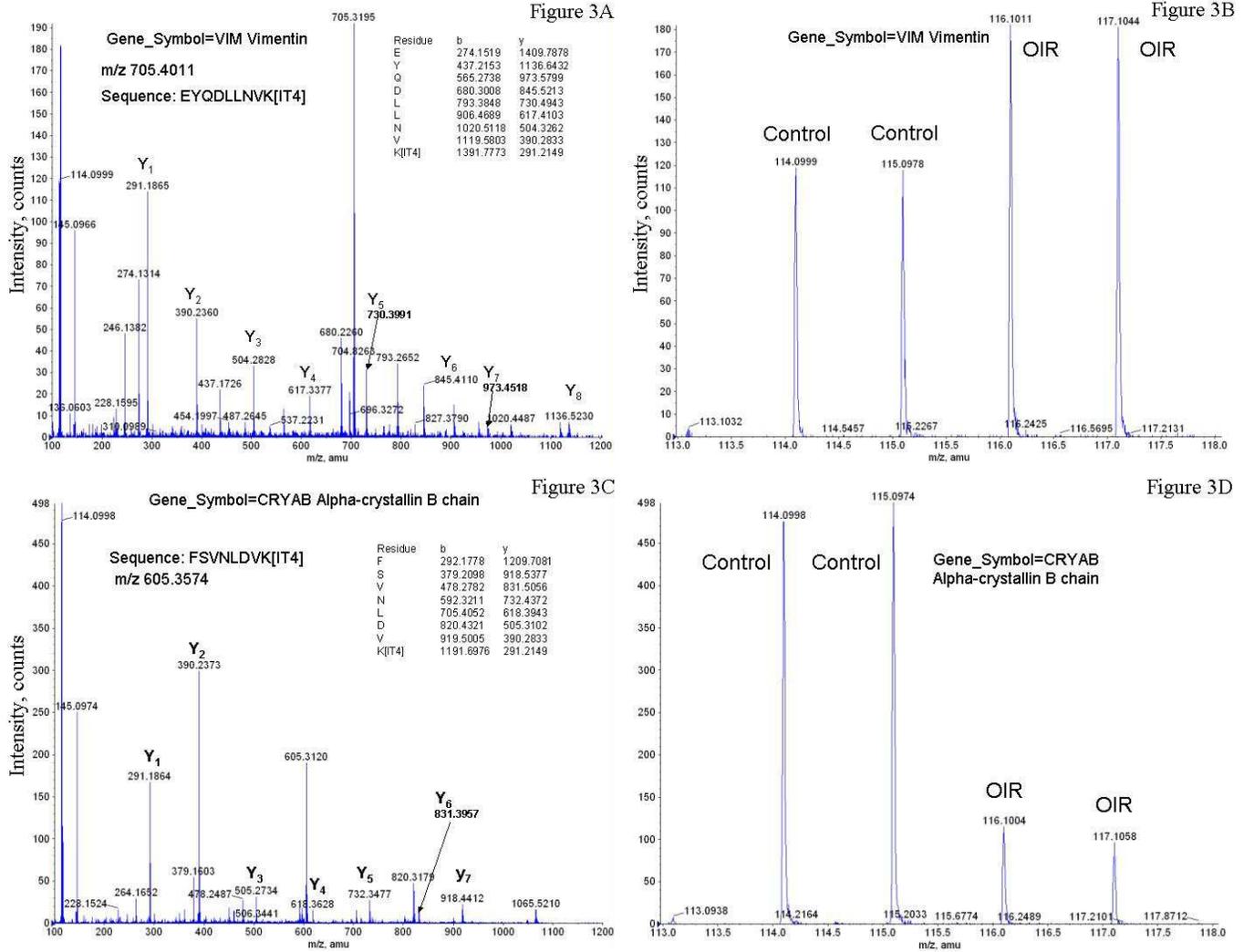
Retina diseases or injury including choroidal neovascularization, diabetic retinopathy, retinal detachment (RD), and retinal trauma can trigger reactive gliosis. In such situations, the increased production of intermediate filament proteins, vimentin, and glial fibrillary acidic protein (GFAP) are often observed [28]. In this study, the increased expression of vimentin was observed (Table 2 and Figure 3).

Our results showed the levels of three ribosomal proteins (60S ribosomal protein L18, 60S ribosomal protein L7 and 40S ribosomal protein SA) increased in the OIR treated mouse retina. Ribosomal proteins are major constituents of ribosomes that catalyze protein synthesis in the cytoplasm [29]. The eukaryotic ribosome is composed of a large (60S) and a small (40S) subunit, which consists of three RNAs and 46 proteins (60S) and one RNA and 33 proteins (40S) [30]. Under normal conditions, ribosomal proteins are synthesized precisely with rRNA to produce equimolar amounts of RNA and protein. Under altered conditions, including events surrounding cellular growth and proliferation, the expression levels of ribosomal proteins are altered. The large ribosomal subunit proteins L7 and the small ribosomal subunit proteins 40S exhibit decreased expression during neuronal differentiation of human embryonic carcinoma cells [31]. Independent alterations in ribosomal protein synthesis suggest that individual ribosomal proteins may have functions beyond the simple structural makeup of the ribosome or protein synthesis. For example, L7 can function as a co-regulator of nuclear receptors and has also been implicated in apoptotic pathways in cataract formation [32]. The 40S ribosomal protein and laminin receptor precursor protein are encoded by the same gene (*37LRP/p40*) [33], suggesting

it has dual functions as both a cell surface receptor and a ribosomal protein. It is overexpressed on cancer cells [34].

Proteases have been increasingly recognized as important factors in the pathophysiology of proliferative retinopathy and tumors. These are involved in the proteolytic degradation of the extracellular matrix, as well as the release and activation of growth and angiogenic factors. Prss3 protein is a kind of serine proteinase (mesotrypsin), belong to S1 peptidase family, encoded by the PRSS3 gene. Some studies suggest that it evokes a transient and pronounced Ca<sup>2+</sup> mobilization in both primary rat astrocytes and retinal ganglion RGC-5 cells, suggesting a physiological role in the brain [35]. Cathepsin D (CatD) acts as a lysosomal cysteine and aspartic proteinase, which are found in most mammalian cells. CatD is one of the most important lysosomal enzymes in RPE cells and can be changed by several posttranslational modifications to produce the biologically active form. It has been proposed that the high presence of inactive forms of CatD in RPE cells can accelerate RPE debris accumulation, RPE atrophy, proliferation and the accumulation of basal lamina and linear deposits associated with retina degeneration. It may play an important role in the development and progression of malignant tumors and overexpression of cathepsin D in aggressive cancers is associated with a poor prognosis [36]. Some studies indicate it has other effects which are independent of its proteolytic activity in cancer cell proliferation and tumor angiogenesis [37].

Neural cell adhesion molecule (NCAM) is a transmembrane protein which mediates cell adhesion and migration. NCAM 180 is one of three isoforms. The level of NCAM 180 isoform was reported to be elevated in astrocytes of glaucomatous optic nerve head [38]. Retinoschisin is a photoreceptor-secreted protein which mediates interactions/adhesion between photoreceptor, bipolar, and Müller cells and contributes to the maintenance of the integrity of the cytoarchitectural of the retina [39]. The mutation of retinoschisin causes a retina disease called X-linked juvenile retinoschisis, which is characterized by morphological and electrophysio-



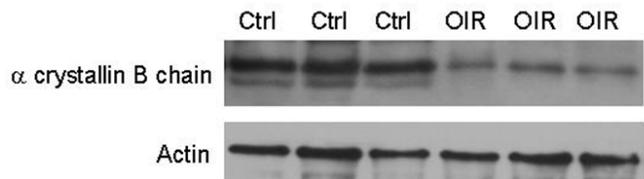
**Figure 3.** (A) MS/MS spectrum of one doubly charged peptide ion (EYQDLLNVK at m/z = 705.40) originated from vimentin and (B) relative quantification for vimentin between OIR samples and control samples. (C) MS/MS spectrum of one doubly charged peptide ion (FSVNLDVK at m/z = 605.36) originated from alpha-crystallin B chain. (D) relative quantification for alpha-crystallin B chain between OIR samples and control samples.

logical defects of the retina.

*Proteins decreased in OIR*

In this study, the crystallins were found to show the most striking changes in the retina proteome occurring between OIR treated animals and controls. We found the levels of five crystallins were down-regulated in OIR, including isoform 1 of  $\alpha$  crystallin A chain, isoform 2 of  $\alpha$  crystallin A chain,  $\alpha$  crystallin B chain,  $\gamma$  crystallin D and  $\beta$ -A3/A1 crystallin. One of them,  $\alpha$  crystallin B chain was also verified by Western blot analysis. Crystallins were originally considered to be static, abundant proteins providing transparency to the lens, and in vertebrates, three major classes of crystallins,  $\alpha$ ,  $\beta$ , and  $\gamma$  accumulate in the lens in a spatially and temporally regulated manner. Now however it is generally accepted that they also have non-lens roles and are retained in multiple tissues of the same organism [40]. Jinghua Xi [41] used

microarray analysis followed by quantitative RT-PCR and found that mouse retinal cells express transcripts for 20 different members of the crystallin gene family.  $\alpha$ A,  $\alpha$ B,  $\beta$ -, and  $\gamma$ -crystallins are detected in the outer and inner nuclear layers,  $\alpha$ B and  $\beta$ -crystallins are detected in the photoreceptor inner segments by immunoblot analysis and



**Figure 4.** Western blot results of alpha-crystallin B chain in 3 mouse control retina samples and 3 mouse OIR retina samples. The bottom panel is the expression of actin used as a loading control.

immunofluorescence [42]. Several reports have identified the expression of one or more of these crystallins in the retina and showed altered crystallin transcript levels resulting from chronic elevation of intraocular pressure, diabetes, light injury, mechanical injury, and age-related macular degeneration [18, 43-46]. It has been established that  $\alpha$ -crystallins are small heat shock proteins that act as molecular chaperones, and are distinguished from other chaperone families by their ability to prevent the non-specific aggregation of denatured proteins and their lack of ATP consumption. However, less is known about the cellular functions of  $\beta/\gamma$ -crystallins. Some studies suggest that  $\alpha$ A and  $\alpha$ B-crystallins may be involved in fundamental processes such as genomic stability. Farjo used gene microarrays and identified many genes of the crystallin family which were significantly down-regulated in the retinas of mice with diabetic retinopathy [47] which supports an implication of our findings.

## 5. Conclusion

In summary, we have used 2D nanoLC-nanoESI-MS/MS combined with iTRAQ to simultaneously determine relative changes in the proteome of retina tissues obtained from OIR mice (at postnatal day 17) compared to age-matched controls. In total, 264 protein groups were identified with high confidence (ProtScore>1.3, >95% confidence). Our data suggests a series of functional changes in proteins associated with retinal angiogenesis. The results from this study may provide a basis for new insights in RN research and closer examination of these factors may ultimately generate novel therapeutic strategies of retinal diseases associated with neovascularization.

## 6. Supplementary material

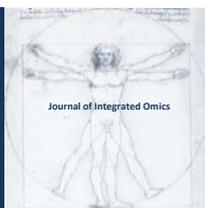
Supplementary material regarding this manuscript is online available in the web page of JIOMICS.

<http://www.jiomics.com/index.php/jio/rt/suppFiles/36/>

## References

1. Ray F, Gariano, Thomas W, Gardner. Retinal angiogenesis in development and disease. *Nature*. 2005; 438(15): 960-966.
2. Sjolie AK, Stephenson J, Aldington S, Kohner E, Janka H, et al. Retinopathy and vision loss in insulin dependent diabetes in Europe: the EURODIAB IDDM Complications Study. *Ophthalmology*. 1997; 104:252-260.
3. Kramerov A, Saghizadeh M, Pan H, Kabosova A, Montemar M. Expression of protein kinase CK2 in astroglial cells of normal and neovascularized retina. *Am J Pathol*. 2006;168(5):1722-36.
4. Kimura T, Takagi H, Suzuma K, Kita M, Watanabe D, et al. Comparisons between the beneficial effects of different sulphonylurea treatments on ischemia-induced retinal neovascularization. *Free Radic Biol Med*. 2007; 43(3):454-61.
5. Smith LE, Wesolowski E, McLellan A, Kostyk SK, D'Amato R, et al. Oxygen-induced retinopathy in the mouse. *Invest Ophthalmol Vis Sci*, 1994, 35(1):101-111.
6. Michael Dorrell, Hannele Uusitalo-Jarvinen, Edith Aguilar, and Martin Friedlander. *Ocular Neovascularization: Basic Mechanisms and Therapeutic Advances*. *Survey Of Ophthalmology*, 2007; 52: S3--S19
7. Sarah X. Zhang, Jian-xing Ma. Ocular neovascularization: Implication of endogenous angiogenic inhibitors and potential therapy. *Progress in Retinal and Eye Research* 2007; 26:1-37.
8. Smith LE, Shen W, Perruzzi C, Soker S, Kinose F, et al. Regulation of vascular endothelial growth factor-dependent retinal neovascularization by insulin-like growth factor-1 receptor. *Nat Med*. 1999; 5:1390-1395.
9. Sarlos S, Rizkalla B, Moravski CJ, Cao Z, Cooper ME, et al. Retinal angiogenesis is mediated by an interaction between the angiotensin type2 receptor, VEGF, and angiopoietin. *Am J Pathol*, 2003; 163(3):879-887.
10. Dawson DW, Volpert OV, Gillis P, Crawford SE, Xu H, et al. Pigment epithelium derived factor: a potent inhibitor of angiogenesis. *Science*. 1999; 285:245-248.
11. Kvanta A. Ocular angiogenesis: the role of growth factors. *Acta Ophthalmol Scand*. 2006; 84(3):282-8.
12. Yoshida S, Yoshida A, Ishibashi T. Induction of IL-8, MCP-1, and bFGF by TNF-alpha in retinal glial cells: implications for retinal neovascularization during after ischemic inflammation. *Graefes Arch Clin Exp Ophthalmol*, 2004; 242:409-413.
13. Aiello, L.P., Pierce, E.A., Foley, E.D., Takagi, H., Chen, H., Riddle, L., Ferrara, N., King, G.L., Smith, L.E., Suppression of retinal neovascularization in vivo by inhibition of vascular endothelial growth factor (VEGF) using soluble VEGF-receptor chimeric proteins. *Proc. Natl. Acad. Sci. USA* 1995 92, 10457-10461.
14. Das, A., McGuire, P.G., Retinal and choroidal angiogenesis: pathophysiology and strategies for inhibition. *Prog. Retin. Eye Res*. 2003; 22, 721-748.
15. Liu Q, Tan G, Levenkova N, Li T, Pugh EN, Rux J, Speicher DW, Pierce EA. The proteome of the mouse photoreceptor sensory cilium complex. *Mol Cell Proteomics*. 2007 May 9;
16. Haniu H, Komori N, Takemori N, Singh A, Ash JD, Matsumoto H. Proteomic trajectory mapping of biological transformation: Application to developmental mouse retina. *Proteomics*. 2006; 6(11):3251-61.
17. Cavusoglu N, Thierse D, Mohand-Said S, Chalmel F, Poch O, Van-Dorsselaer A, Sahel JA, Léveillard T. Differential proteomic analysis of the mouse retina: the induction of crystallin proteins by retinal degeneration in the rd1 mouse. *Mol Cell Proteomics*. 2003; 2(8):494-505.
18. Crabb JW, Miyagi M, Gu X, Shadrach K, West KA, et al. Drusen proteome analysis: an approach to the etiology of age related macular degeneration. *Proc Natl Acad Sci*, 2002; 99:14682-14687.
19. Haniu H, Komori N, Takemori N, Singh A, Ash JD, et al. Proteomic trajectory mapping of biological transformation: Application to developmental mouse retina. *Proteomics*. 2006; 6 (11):3251-61.
20. Godfrey J, Quin, Alice C. L. Len, Frank A. Billson, Mark C. Gillies. Proteome map of normal rat retina and comparison with the proteome of diabetic rat retina: New insight in the pathogenesis of diabetic retinopathy. 2007; 7:2636-2650.
21. Curtis L, Nordgaard, Kristin M, Berg, Rebecca J, Kapphahn, Cavan Reilly, Xiao Feng, et al. Proteomics of the Retinal Pig-

- ment Epithelium Reveals Altered Protein Expression at Progressive Stages of Age-Related Macular Degeneration. *Invest Ophthalmol Vis Sci*, 2006, 47;3: 815-822.
22. Higgins RD, Yu K, Sanders RJ, et al. Diltiazem reduces retinal neovascularization in a mouse model of oxygen induced retinopathy. *Curr Eye Res* 1999; 18: 20-7
  23. Zhi-Liang Hu, Jie Bao and James M. Reecy (2008) "CateGORizer: A Web-Based Program to Batch Analyze Gene Ontology Classification Categories". *Online Journal of Bioinformatics*. 9 (2):108-112.
  24. Sato T, Kusaka S, Hashida N, Saishin Y, Fujikado T, Tano Y. Comprehensive gene-expression profile in murine oxygen-induced retinopathy. *Br J Ophthalmol*. 2009 Jan;93(1):96-103.
  25. Bresgen M, Baum U, Esser P, Wiedemann P, Heimann K. Protein composition of the vitreous body in proliferative diabetic retinopathy. An analysis with 2D-electrophoresis. *Ophthalmologe*. 1994; 9l: 758—762.
  26. Nakanishi T, Koyama R, Ikeda T, Shimizu A. Catalogue of soluble proteins in the human vitreous humor: comparison between diabetic retinopathy and macular hole. *J Chromatogr B Analyt Technol Biomed Life Sci*, 2002; 776:89-100.
  27. García-Ramírez M, Canals F, Hernández C, Colomé N, Ferrer C, et al. Proteomic analysis of human vitreous fluid by fluorescence-based difference gel electrophoresis (DIGE): a new strategy for identifying potential candidates in the pathogenesis of proliferative diabetic retinopathy. *Diabetologia*. 2007; 50 (6):1294-1303.
  28. Nakazawa T, Takeda M, Lewis GP, Cho KS, Jiao J, Wilhelmsen U, Fisher SK, Pekny M, Chen DF, Miller JW. Attenuated glial reactions and photoreceptor degeneration after retinal detachment in mice deficient in glial fibrillary acidic protein and vimentin. *Invest Ophthalmol Vis Sci*. 2007 Jun;48(6):2760-8.
  29. Mager WH. Control of ribosomal protein gene expression. *Biochim.Biophys Acta*. 1988; 949:1-15.
  30. Wool IG, Chan Y-L, Gluck A. Structure and evolution of mammalian ribosomal proteins. *Biochem Cell Biol*. 1995; 73: 933-947.
  31. Bevort M, Leffers H. Down regulation of ribosomal protein mRNAs during neuronal differentiation of human NTERA2 cells. *Differentiation*. 2000; 66:81-92.
  32. Ben-Ishai R, Scharf R, Kaptan I. A human cellular sequence implicated in trk oncogene activation is DNA damage inducible. *Prod Natl Acad Sci USA*. 1990;87: 6039-6043.
  33. Jamieson KV, Wu J, Hubbard SR, Meruelo D. Crystal structure of the human laminin receptor precursor. *J Biol Chem*. 2008 Feb 8;283(6):3002-5. Epub 2007 Dec 6.
  34. Mafune K, Ravikumar TS, Wong JM, et al. Expression of a Mr 32,000 laminin-binding protein messenger RNA in human colon carcinoma correlates with disease progression. *Cancer Res*, 1990; 50:3888-3891.
  35. Yingfei Wang, Weibo Luo, Thomas Wartmann, Walter Halangk, Miklós Sahin-Tóth and Georg Reiser. Mesotrypsin, a brain trypsin, activates selectively proteinase-activated receptor -1, but not proteinase-activated receptor-2, in rat astrocytes. *Journal of Neurochemistry*, 2006; Volume 99 Issue 3 Page 759-769.
  36. Foekens JA, Look MP, Bolt-de Vries J, Meijer-van Gelder ME, van Putten WL, et al. Cathepsin-D in primary breast cancer: prognostic evaluation involving 2810 patients. *Br J Cancer*. 1999 Jan;79(2):300-7.
  37. Laurent-Matha V, Maruani-Herrmann S, Prébois C, Beaujouin M, Glondou M, Catalytically inactive human cathepsin D triggers fibroblast invasive growth. *J Cell Biol*. 2005 Jan 31;168 (3):489-99.
  38. Ricard CS, Kobayashi S, Pena JD, Salvador-Silva M, Agapova O, Hernandez MR. Selective expression of neural cell adhesion molecule (NCAM)-180 in optic nerve head astrocytes exposed to elevated hydrostatic pressure in vitro. *Brain Res Mol Brain Res*. 2000 Sep 30;81(1-2):62-79.
  39. Reid SN, Yamashita C, Farber DB. Retinoschisin, a photoreceptor-secreted protein, and its interaction with bipolar and muller cells. *J Neurosci*. 2003 Jul 9;23(14):6030-40.
  40. Sax C, Piatigorsky J. Expression of the  $\alpha$ -crystallin/small heat-shock protein/molecular chaperone genes in the lens and other tissues. In: Meister A, editor. *Advances in enzymology and related areas of molecular biology*. Vol 69. New York: Wiley Interscience; 1994:155-201.
  41. Jinghua Xi, Rafal Farjo, Shigeo Yoshida, et al. A comprehensive analysis of the expression of crystallins in mouse retina. *Molecular Vision* 2003; 9:410-9
  42. Ahmed F, Brown KM, Stephan DA, Morrison JC, Johnson EC, Tomarev SI. Microarray analysis of changes in mRNA levels in the rat retina after experimental elevation of intraocular pressure. *Invest Ophthalmol Vis Sci* 2004; 45:1247-58.
  43. Steele MR, Inman DM, Calkins DJ, Horner PJ, Vetter ML. Microarray analysis of retinal gene expression in the DBA/2J model of glaucoma. *Invest Ophthalmol Vis Sci* 2006; 47:977-85.
  44. Sakaguchi H, Miyagi M, Darrow RM, Crabb JS, Hollyfield JG, Organisciak DT, Crabb JW. Intense light exposure changes the crystallin content in retina. *Exp Eye Res* 2003; 76:131-3.
  45. Vazquez-Chona F, Song BK, Geisert EE Jr. Temporal changes in gene expression after injury in the rat retina. *Invest Ophthalmol Vis Sci* 2004; 45:2737-46.
  46. Jones SE, Jomary C, Grist J, Thomas MR, Neal MJ. Expression of alphaB-crystallin in a mouse model of inherited retinal degeneration. *Neuroreport* 1998; 9:4161-5.
  47. Farjo R, Kern TS, Andley UP, Swaroop A. Microarray analysis of mouse models of diabetic retinopathy reveals dramatic modulation of crystallin genes in the retina. *ARVO Annual Meeting*, 2003 May 4-9.



ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v1i2.64

## A method for estimation of immunogenic determinants mutability: case studies of HIV1 gp120 and diphtheria toxin

Khrustalev Vladislav Victorovich<sup>\*1</sup>, Barkovsky Eugene Victorovich<sup>1</sup>, Vasilevskaya Alla Evgenyevna<sup>2</sup>, Skripko Svetlana Michailovna<sup>2</sup>, Kolodkina Valentina Leonidovna<sup>3</sup>, Ignatyev Georgiy Michailovich<sup>3</sup>, Semizon Pavel Anatolyevich<sup>4</sup>.

<sup>1</sup>Department of General Chemistry, Belarussian State Medical University, Belarus, Minsk, Dzerzinskogo, 83; <sup>2</sup>Laboratory of HIV/AIDS Diagnostic, Republican Center of Hygiene, Epidemiology and Public Health, Belarus, Minsk, Kazinca, 50; <sup>3</sup>Immunoprophylaxis Laboratory, Republican Research & Practical Centre for Epidemiology and Microbiology, Belarus, Minsk, Filimonova, 23; <sup>4</sup>Laboratory of Biotechnology and Immunodiagnosics of Highly Dangerous Infections, Republican Research & Practical Centre for Epidemiology and Microbiology, Belarus, Minsk, Filimonova, 23.

Received: 14 April 2011 Accepted: 10 June 2011 Available Online: 13 June 2011

### ABSTRACT

There is a need for the method which helps to choose the less mutable immunogenic determinant for the design of recombinant or synthetic vaccines and ELISA test-systems. Our method based on the directional mutational pressure theory includes two steps: estimation of symmetric and asymmetric mutational pressure directions in a gene coding for a protein of interest; and selection of regions coding for its immunogenic determinants which are less prone to missense mutation occurrence and so, to immune escaping. Three original computer algorithms ("VVK Sliding Window", "VVK VarInvar" and "VVK Protective Buffer" available via [www.barkovsky.hotmail.ru](http://www.barkovsky.hotmail.ru)) have been created to perform all the necessary calculations and tests. "VVK Sliding Window" calculates nucleotide usage in fourfold and twofold degenerated sites, as well as usage of missense, nonsense and synonymous sites for each kind of nucleotide mutation along the length of a coding region, while "VVK Protective Buffer" calculates those indexes in a set of sequences. "VVK VarInvar" calculates percentage of variable sites in a set of aligned sequences, as well as nucleotide usage in invariable sites. Our method has been tested on HIV1 gp120 protein and on diphtheria toxin. The less mutable epitopes have been found for both proteins. Finally, it has been shown that antibodies recognizing the less mutable epitope of gp120 can be found in 80.22% of HIV1-infected persons.

**Keywords:** Mutational pressure; Sequence analysis; B-cell epitopes; HIV1 vaccine; gp120; Diphtheria toxin.

### Abbreviations

**G4f; C4f; A4f; T4f** – usage of guanine, cytosine, adenine and thymine, respectively, in fourfold degenerated sites; **G2f3p; C2f3p; A2f3p; T2f3p** – usage of guanine, cytosine, adenine and thymine, respectively, in twofold degenerated sites from third codon positions.

### 1. Introduction

Recombinant and synthetic vaccines are thought to be less dangerous than life attenuated vaccines [1]. In case of immunization with the most conserved immunogenic determinant (determinants) protective immunity will be formed against the most of the strains from the whole population of the pathogenic microorganism (against those strains among

which the immunogenic determinant is conserved). So, determination of the less variable and the less mutable immunogenic determinant (B-cell epitope or T-cell epitope) is one of the most important steps in recombinant or synthetic vaccine design [2].

It is possible to determine those parts of a coding region

\*Corresponding author: Vladislav Victorovich Khrustalev. Address: Belarus, Minsk, 220029, Communisticheskaya 7-24; Telephone: 80172845957; E-mail Address: [vkhrustalev@mail.ru](mailto:vkhrustalev@mail.ru)

which are less prone to missense mutation occurrence than other ones with the help of our method based on the mutational pressure theory [3].

If the number of variable sites between sequences of the protein of interest is low, the borders of conserved and variable regions cannot be determined, while levels of mutability can still be estimated [2]. The information on mutability of regions coding for immunogenic determinants is important even in case if the volume of data on amino acid substitutions in the protein of interest is sufficient. Immunization against the most conserved and the less mutable epitope (epitopes) of the pathogenic microorganism should be more efficient than immunization against the conserved and, at the same time, mutable epitope (epitopes) [2].

Envelope glycoprotein gp120 from Human immunodeficiency virus type 1 (HIV1) has been chosen to represent benefits of our method for the protein with sufficient amount of data on its variability. Diphtheria toxin has been chosen to represent benefits of our method for the protein with low number of available sequences.

3D B-cell epitopes have already been mapped on the surface of HIV1 gp120 in our previous work [2], while levels of their mutability have been determined by the less accurate method than that described in the present work. The less mutable and the most conserved predicted 3D B-cell epitope of HIV1 gp120 [2] was shown to be immunogenic in the present work. Antibodies to that predicted conformational B-cell epitope of gp120 can be found in 80.22% of HIV-infected persons using ELISA test-system with synthetic peptide (NQ21) corresponding to its consensus sequence. The fact that IgG from the serum of HIV1-positive person is able to cross-react with the NQ21 peptide has been confirmed by affinity purification.

Diphtheria toxin is the product of a gene (*tox*) which can be found in genomes of lysogenic corynephages infecting *Corynebacterium diphtheriae* and *Corynebacterium ulcerans* [4]. The toxin is lethal for susceptible persons in doses of 100 ng/kg [5]. In 1929, Ramon demonstrated that diphtheria toxin can be turned to its nontoxic, but antigenic equivalent (toxoid) by formaldehyde. The main disadvantage of diphtheria toxoid usage is the complicated procedure of its production [6].

Although diphtheria toxin is thought to be quite conserved protein [7], there are 54 variable amino acid sites in seven sequences studied. These data are not enough to separate conserved regions from variable ones for a sequence of 560 amino acids in length.

Our method requires the usage of three original computer algorithms available via our web page ([www.barkovsky.hotmail.ru](http://www.barkovsky.hotmail.ru)). These relatively simple algorithms included in MS Excel spreadsheets are described in our SciTopics Page [8].

The first step of the method is in the estimation of mutational pressure direction. The second step is in the comparison of nucleotide usage indexes in regions coding for immunogenic determinants with the aim to choose the less muta-

ble one.

Symmetric mutational GC-pressure exists when the rates of AT to GC nucleotide mutations occurrence are higher than the rates of GC to AT nucleotide mutations occurrence [3]. Symmetric mutational AT-pressure exists when the rates of GC to AT mutations occurrence are higher than the rates of AT to GC mutations occurrence [3]. Those types of mutations which occur more frequently are fixed by random genetic drift in synonymous sites more frequently than those types of mutations which occur rarely [3]. As long as mutational GC-pressure leads to the almost complete saturation of synonymous sites with G and C, while mutational AT-pressure leads to the almost complete saturation of synonymous sites with A and T, it finally leaves no substrate for synonymous nucleotide mutations [9].

Directional mutational GC-pressure increases levels of usage of those amino acid residues which are encoded by GC-rich codons [9, 10]. Acrophilic (prone to be located on a surface of a protein [11]) proline and glycine are among those four amino acid residues [10]. Such strongly hydrophobic amino acid residues as isoleucine, phenylalanine, tyrosine and methionine [12] are encoded by GC-poor codons. Indeed, according to our computer simulations with BepiPred 1.0 algorithm [13], mutational GC-pressure frequently leads to formation of new linear B-cell epitopes and elongation of previously existing ones [9], while mutational AT-pressure frequently leads to disappearance of epitopes or their parts from the surface of proteins [14]. Percent of highly immunogenic amino acid residues forming linear and discontinuous B-cell epitopes is usually higher for homologous proteins encoded by GC-rich genes [15].

There may be significant symmetric mutational bias in twofold degenerated and fourfold degenerated sites even in coding genomes with 3GC levels close to 50% [9]. Asymmetric mutational pressure exists due to the differences in rates of occurrence of different types of mutations in leading and lagging strands of DNA, as well as due to the differences in rates of occurrence of different types of mutations in transcribed and nontranscribed strands of DNA [16, 17]. Even in genomes of prokaryotic organisms and viruses direction of symmetric mutational pressure may be different for different genes [9].

“VVK Sliding Window” and “VVK VarInvar” algorithms were designed for estimation of both symmetric and asymmetric components of mutational pressure, as well as for separation of biases in rates of transitions from biases in rates of transversions [8].

“VVK Protective Buffer” algorithm was designed for fast calculation of those indexes characterizing the usage of nucleotides prone to frequent mutations [8] which are used during selection of the region coding for the less mutable epitope.

## 2. Material and Methods

### 2.1 Description of nucleotide sequences used

Thirty four sets of *env* gene sequences coding for HIV1 gp120 protein have been used as a material. The total number of sequences is equal to 689. Each set of sequences has been obtained from a single HIV1-infected person. Codes of infected persons and GenBank accession numbers of *env* gene sequences are listed below. H1 – H5: **EU743973** – **EU744175** [18]; S33 and S35: **EU604549** – **EU604642** [19]; DM1 – DM9: **EF575363** – **EF575486**; C61, C62, C93, C94, C96, C98, C109, ES2, ES4, ES7 – ES9: **DQ410040** – **DQ410649** [20]; 9F, 605F, 605M, 32M, 183M and 120F: **EU852934** – **EU853141** [21].

For a second example of the application of our method nine nucleotide sequences coding for *Corynebacterium diphtheria* toxin have been used. Three sequences were from integrated corynebacteriophages: one of them was from the reference genome of *Corynebacterium diphtheriae* (NP\_938615.1), two of them were from other strains of *Corynebacterium diphtheriae* (AJ576101.1; AY820132.1). Six sequences were from different corynebacteriophages: four were from *Corynebacteriophage beta* (K01722.1; EU069362.1; K01723.1; D78299), one was from *Corynebacteriophage omega* (V01536.1) and the last one was from unclassified *Corynebacteriophage* (X00703.1). Eight sequences of *tox* gene from integrated *Corynebacterium ulcerans* phages (AB304279.1; FJ858272.1; AB498872.1; AB304280.1; AB304278.1; AY141014.1; AY703827.1; AY141013.1) have also been used in the study.

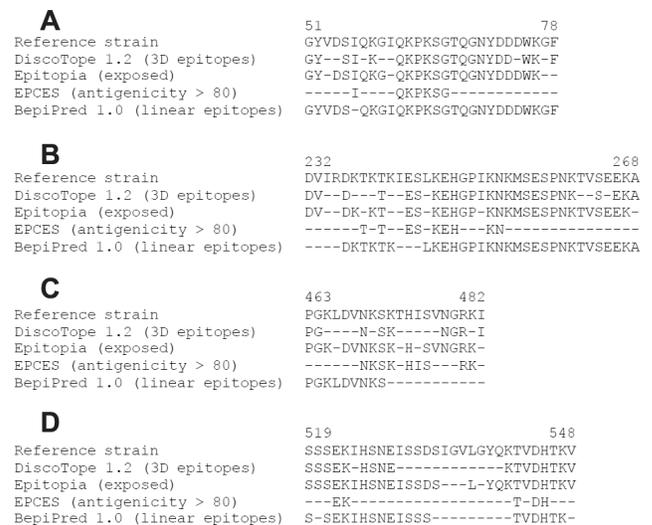
### 2.2 Algorithms for 3D and linear B-cell epitopes prediction

DiscoTope 1.2 algorithm [22] has been used to map 3D epitopes on the structure of diphtheria toxin deposited in PDB database (its accession number is “1SGK”). Then these results have been confirmed with the help of Epitopia [23] and Epces [24] algorithms. Sequences of the four most immunogenic regions of the diphtheria toxin are represented in Figure 1. For example, in Figure 1A amino acid sequence of the toxin from reference strain is written in the first line, amino acid residues included in 3D epitopes by DiscoTope 1.2 are written in the second line, residues which are exposed to solvent according to Epitopia prediction are written in the third line, residues with very high antigenicity score (according to Epces results) are written in the fourth line.

It has to be noted that there are always corresponding linear B-cell epitopes predicted by BepiPred 1.0 algorithm [13] for every 3D epitope from diphtheria toxin (see fifth lines in Figures 1A – 1D).

Seventeen sequences coding for *Corynebacterium diphtheria* and *Corynebacterium ulcerans* toxin have been aligned with sequences of its four most immunogenic regions (3D epitopes 1, 2, 3 and 4, respectively). Then four separate alignments of sequences (each of them is coding for one of the four epitopes) have been made to perform calculations in them.

Five most immunogenic regions of HIV1 gp120 protein have been predicted by us in the previous work [2] with the



**Figure 1.** Amino acid sequences of four diphtheria toxin regions with predicted B-cell epitopes. Amino acid residues included in 3D B-cell epitopes by DiscoTope 1.2 algorithm, exposed amino acid residues (according to Epitopia prediction), highly antigenic amino acid residues (according to Epces algorithm prediction) and linear B-cell epitopes (according to BepiPred 1.0 prediction) are designated.

help of DiscoTope 1.2 [22], Epitopia [23] and Epces [24] algorithms.

### 2.3 Original computer algorithms

“VVK Siding Window” algorithm (which is a new version of “VVK in length” algorithm [25]) calculates nucleotide usage in fourfold and twofold degenerated sites along the length of a coding region in sliding windows, as well as probabilities to be synonymous, missense and nonsense for every type of nucleotide mutation [8]. These kinds of analyses help to determine main directions of mutational pressure and to perform screening test to find the less mutable regions. The algorithm requires a single nucleotide sequence as an input.

“VVK VarInvar” algorithm (which is a new version of “VVK Consensus” algorithm [25]) calculates the percentage of variable sites in a set of aligned sequences [8]. This algorithm requires alignment with at least 109 variable sites to give a reproducible percentage (see section 5.3). Relative frequencies of different types of nucleotide mutations can be estimated dealing with the percentage of variable sites. Average levels of nucleotide usage in third codon positions should be compared with nucleotide usage in invariable sites from third codon positions to confirm directions of those mutations [8]. The average usage of the “stable” nucleotide should be increased in invariable sites from third codon positions. The usage of “mutable” nucleotide should be decreased in invariable sites from third codon positions.

“VVK Protective buffer” algorithm (which is a new version of “VVK in group” algorithm [25]) calculates usage of synonymous, nonsense and missense sites for each kind of

nucleotide mutation in the set of sequences (up to 50 sequences can be used as an input) [8].

#### 2.4 Peptide synthesis

Peptides of 21 amino acid residues in length have been synthesized for us by “Peptide 2.0 Inc.” company. The “Symphony” solid-phase peptide synthesizer (“Protein Technologies, Inc.”) has been used. The purity of both peptides was higher than 95% (according to the results of HPLC analyses). One of those peptides has been conjugated with biotin (via its N-terminal). Amino acid sequences of both peptides are the same as the consensus sequence for 3D epitope 1 from gp120 protein [2].

#### 2.5 Modification of ELISA test-system

Commercial “ELISA-Recombinant-HIV 1,2” test-system produced by “PharmLand LLC” company (<http://www.pharmland.by/en/products/elisa-recombinant-hiv-1>) has been modified by us. This test-system includes a 96-well plaque with adsorbed recombinant gp160 and gp140 proteins, solution of recombinant gp160 (from HIV1) and gp140 (from HIV2) proteins conjugated with biotin and standard components, such as solutions of streptavidin conjugated with horseradish peroxidase, TMB, hydrogen dioxide and sulfuric acid, as well as wash solution. Solution of recombinant gp160 and gp140 conjugated with biotin from this test-system was replaced with the solution of the peptide NQ21 conjugated with biotin (0.04 mg/ml in 0.1M PBS with pH 7.4). The protocol of the analysis was as follows: 0.05 ml of the peptide biotin-NQ21 solution and 0.1 ml of serum have been simultaneously added to each well (0.002 mg of the peptide per well) and incubated for 30 minutes (37°C). During this period of time asymmetric sandwiches (epitope 1 of gp160 – antibody – biotin-NQ21) have been formed in certain wells. The rest of the protocol was standard (according to manufacturer’s protocol): 7 washes; incubation with streptavidin conjugated with horseradish peroxidase for 30 minutes (37°C); 7 washes; incubation with TMB and hydrogen dioxide during 15 minutes (37°C); addition of sulfuric acid and measurement of optical density by the automatic 96-well plaque spectrophotometer (at the wavelength of 450 nm).

Serums from 91 persons with recently revealed HIV1-infection have been tested. Their diagnosis has been confirmed by two different ELISA test-systems and by immunoblotting in the Laboratory of HIV/AIDS Diagnostic of the Republican Center of Hygiene, Epidemiology and Public Health (Minsk, Belarus). All of those persons did not receive antiretroviral therapy. Serums from 234 HIV-negative persons have also been tested.

To separate positive and negative results in the modified ELISA test-system the following calculations were performed. Average level of optical density (OD) for all the wells with serums from HIV-negative persons has been cal-

culated for each of the four 96-well plaques, as well as the standard deviation. All the levels of OD higher than three standard deviations were considered to be positive (in a given plaque). In case of parallel analysis of a single serum (in two wells) an average OD for those two wells was calculated. For each of the four plaques an average level of OD for serums from HIV1-positive persons was significantly higher than that for serums from HIV-negative persons (according to the results of t-test).

#### 2.6 Affinity purification protocol

The column from AminoLink Plus Immobilization Trial Kit (“Thermo scientific Inc.”) was used to immobilize the peptide NQ21 via its NH<sub>2</sub> groups. There are two amine groups in the peptide NQ21: one is N-terminal and another one is from the side chain of lysine. Modified agarose contains aldehyde groups that react specifically with primary amines. After the spontaneous formation of semi-stable Schiff base bonds, reduction with sodium cyanoborohydride results in stable secondary amine bonds. Immobilization protocol (which can be found at <http://www.piercenet.com/instructions/2160491.pdf>) includes incubation with the peptide dissolved in pH 10 buffer during 4 hours; incubation with sodium cyanoborohydride dissolved in pH 7.4 PBS during 4 hours; incubation with quenching buffer (1M tris-HCl) during 30 minutes; incubation with sodium cyanoborohydride dissolved in pH 7.4 PBS during 30 minutes and washing the column with wash buffer (1M NaCl). About 2.3 mg of NQ21 have been finally immobilized.

1.5 ml of the serum from HIV1-infected person dissolved in 0.5 ml of PBS pH 7.4 has been added to the column with immobilized peptide NQ21 and incubated during 60 minutes. Then the column has been washed 11 times with PBS pH 7.4 (2 ml of PBS for each wash). Acetic acid (1M) has been used as the eluent. This procedure has been repeated for 1.5 ml of serum from HIV-negative person dissolved in 0.5 ml of PBS pH 7.4. Concentration of proteins in eluates has been estimated with the help of Hitachi 650-60 spectrofluorometer (excitation wavelength was equal to 296 nm, emission wavelength was equal to 345 nm).

#### 2.7 SDS-PAGE analysis

Eluates 1, 3 and 7 obtained during affinity purification of the serum from HIV1-infected person were analyzed in SDS-PAGE. Composition of the sample buffer was as follows: 1M Tris-HCl, pH 6.8; 4% SDS; 2% 2-mercaptoethanol; 20% glycerol; 0.02% bromophenol blue. Samples were heated for 2 minutes at 95°C in a water bath. Three dilutions of each eluate have been prepared (1:1; 1:2 and 1:4). PageRuler™ Unstained Protein Ladder molecular mass markers from “Fermentas Life Sciences” were used.

### 3. Results

#### 3.1 Nucleotide usage biases in fourfold and twofold degenerated sites from third codon positions along the length of a coding region

The usage of thymine in twofold degenerated sites from third codon positions (T2f3p) is always higher than the usage of cytosine in them (C2f3p) along the length of the region of HIV1 *env* gene coding for gp120 (see Figure 2A). The usage of adenine in twofold degenerated sites from third codon positions (A2f3p) is higher than the usage of guanine (G2f3p) in the most of “sliding windows” from the above-mentioned coding region.

The difference between adenine usage in fourfold degenerated sites (A4f) and cytosine usage in them (C4f) is very high (see Figure 2C), especially in comparison with the difference between thymine usage (T4f) and guanine usage (G4f) in those sites of the region of *env* gene coding for gp120 (see Figure 2D).

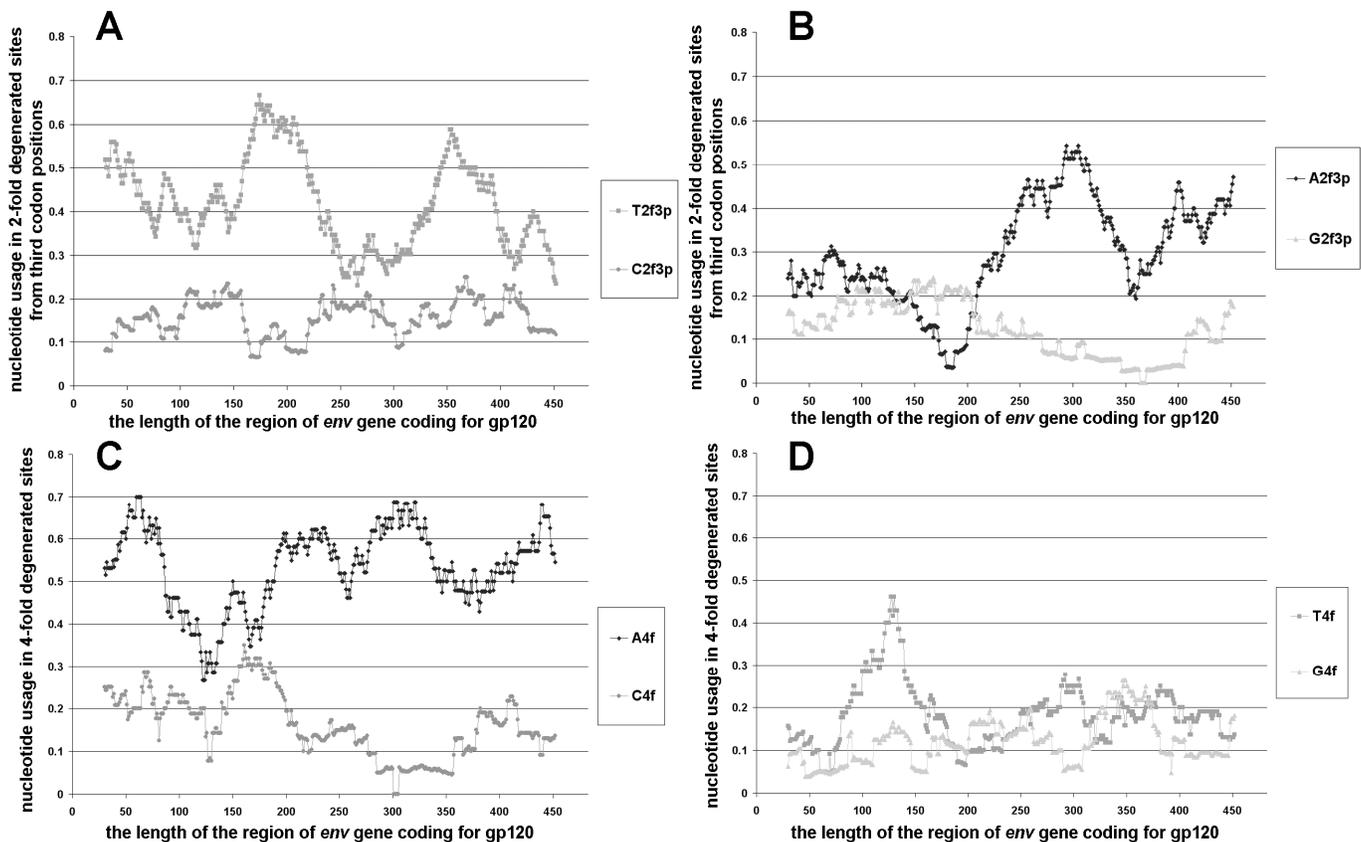
The level of C2f3p is significantly lower than the level of T2f3p along the most of the length of *tox* gene (see Figure 3A). The level of G2f3p is significantly lower than the level of A2f3p along the most of the length of *tox* gene (see Figure 3B).

The level of A4f is much higher than that of C4f in the most of “sliding windows” along the length of *tox* gene (see Figure 3C). The level of T4f is much higher than that of G4f (see Figure 3D). Moreover, T4f usage is significantly higher than A4f usage (average paired difference in t-test is equal to  $0.158 \pm 0.006$ ,  $P < 0.001$ ). This should be the evidence that bias

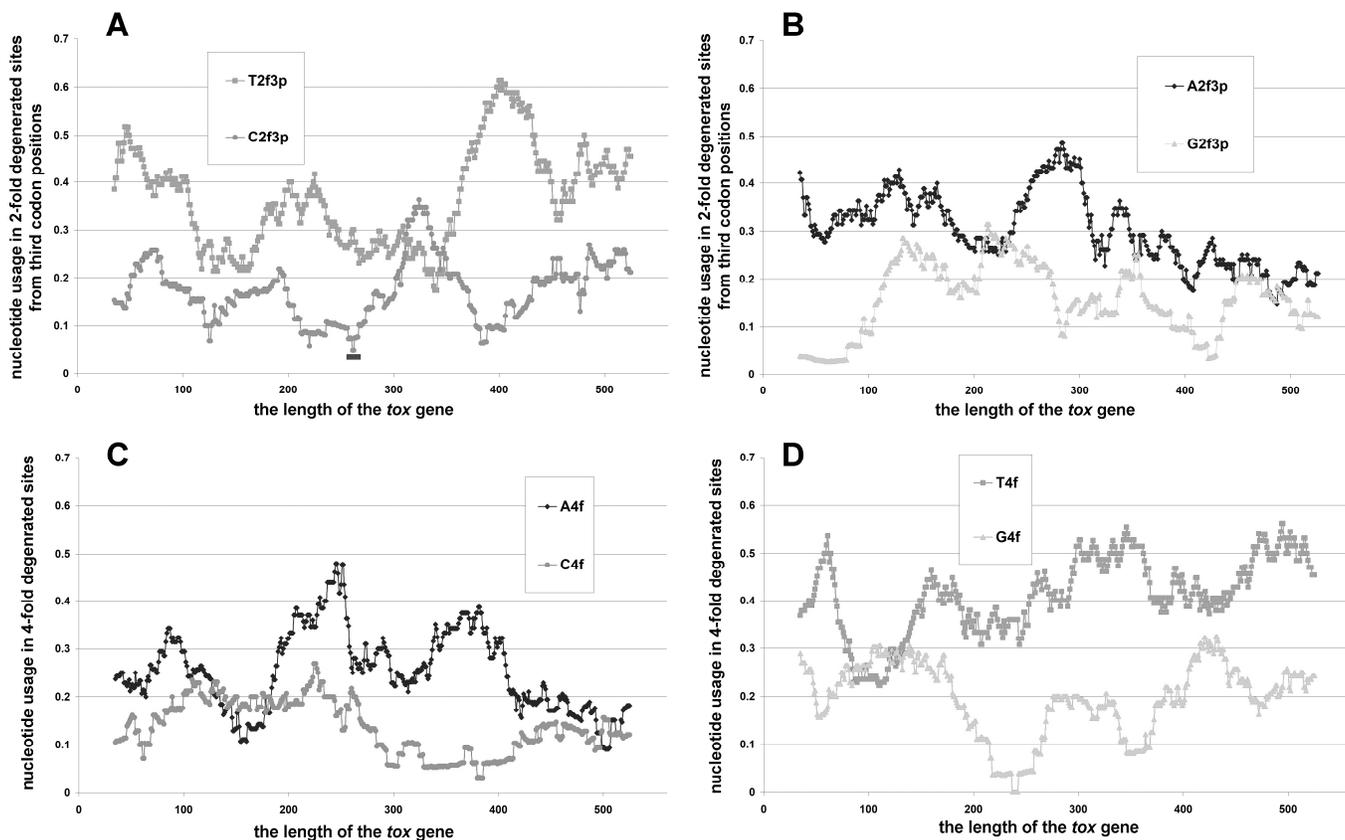
between G to T and T to G transversions have been stronger than the bias between C to A and A to C transversions in *tox* gene, unlike those in the region of HIV1 *env* gene coding for gp120.

Taking together, data represented in Figures 2 – 3 approve that mutational AT-pressure in both *env* and *tox* genes is caused by both GC to AT transitions and GC to AT transversions. However, the rates of C to A transversions in HIV1 *env* gene should be higher than the rates of G to T transversions. In the gene coding for diphtheria toxin the rates of C to A transversions should be lower than the rates of G to T transversions.

According to the results of codon-based Z-test for positive selection, nonsynonymous distance (calculated by Kumar method) [26] is significantly higher than synonymous distance ( $dN > dS$ ) between two sequences of *tox* gene from *Corynebacterium ulcerans* phages (for those with the following GenBank identifiers: **AY703827.1** and **FJ858272.1**). Indeed, there are three amino acid replacements between them, while the number of synonymous nucleotide mutations is equal to zero. Two of amino acid substitutions are specific to **AY703827.1** sequence. One of them (T262I) took place in Epitope 2. This substitution caused by a single C to T transition resulted in the replacement of relatively acrophilic and hydrophilic threonine [11] with strongly hydro-



**Figure 2.** Thymine and cytosine (A) and adenine and guanine (B) usage in twofold degenerated sites from third codon positions along the length of the region of *env* gene coding for gp120. Adenine and cytosine (C) and thymine and guanine (D) usage in fourfold degenerated sites along the length of the *tox* gene. The length of sliding window is equal to 70 codons.



**Figure 3.** Thymine and cytosine (A) and adenine and guanine (B) usage in twofold degenerated sites from third codon positions along the length of the *tox* gene. Adenine and cytosine (C) and thymine and guanine (D) usage in fourfold degenerated sites along the length of the *tox* gene. The length of sliding window is equal to 70 codons. The region with the minimum of C2f3p usage is designated.

phobic isoleucine [12]. This substitution might lead to the loss of affinity of previously synthesized antibodies to this epitope, since scores of thirteen amino acid residues surrounding T262I to be included in linear B-cell epitope according to BepiPred 1.0 prediction, became lower than those in chain B of the reference strain, while they are still above the threshold.

Interestingly, the region of 73 codons in length surrounding the codon number 262 has the lowest level of C2f3p (see Figure 3A). This fact can be interpreted as yet another reason to estimate probabilities of synonymous mutation occurrence in regions coding for B-cell epitopes. Moreover, this fact is one more evidence that low probability of synonymous mutation occurrence (due to the strong mutational pressure) may sometimes lead to the situation when dN becomes significantly higher than dS.

### 3.2 Distribution of synonymous, missense and nonsense sites for different types of mutations along the length of a gene

Nonsynonymous sites can be divided into two groups: missense sites and nonsense sites. Nucleotide mutation in missense site (for a given type of nucleotide mutation) leads to the amino acid replacement. Nucleotide mutation in nonsense site leads to the formation of stop-codon.

There are three stop-codons in the universal genetic code

(TAA, TAG and TGA). All of them are relatively GC-poor. That is why mutations of GC to AT direction have much higher probability to be nonsense than mutations of AT to GC direction [27]. Nonsense mutation leads to the translation of truncated protein. A protein may lose its C-terminal (if nonsense mutation occurred near the end of open reading frame), or N-terminal (in case if nonsense mutation occurred near the beginning of open reading frame, translation may start from the alternative initiation codon). Truncated protein may still perform its function. In this case nonsense mutations may lead to the immune escaping even more likely and more efficiently than missense mutations. In case if nonsense mutation leads to the complete loss of function for the given protein, one may not be afraid of immune escaping due to this type of mutation.

Distribution of synonymous, missense and nonsense sites along the length of the region of *env* gene coding for gp120 from the reference HIV1 strain is quite variable (see Figure 4). The lowest level of missense sites for C to T transitions is characteristic to the region coding for 3D epitope 1 of gp120 (see Figure 4A). The highest level of the “protective buffer” against G to A transitions (it includes both synonymous and nonsense sites) is also characteristic to that region (see Figure 4B). The lowest level of missense sites for G to A transitions is characteristic to the region coding for 3D epitope 3 of gp120, while highest levels are characteristic to regions

coding for 3D epitopes 5 and 2 (see Figure 4B). The region coding for 3D epitope 1 seems to be well protected from C to A transversions because of the low level of cytosine usage (see Figure 4C). Interestingly, the level of nonsense sites for G to T transversions is much higher than the level of nonsense sites for C to A transversions in both *env* (see Figures 4C and 4D) and *tox* genes (see Figures 5C and 5D), as well as in the most of bacterial coding regions [17].

Taking into account that the most frequent types of nucleotide mutations in HIV1 *env* gene are G to A transitions, as well as C to T transitions and C to A transversions, the first 3D epitope (from C1 region) was considered to be the less mutable 3D epitope of gp120 [2]. This statement has been made after the calculation of cytosine and guanine content in first, second and third codon positions of regions coding for gp120 3D epitopes. The level of “protective buffer” in that study was measured as the nucleotide usage in third codon positions. In the present work that level was calculated as accurate as it possible (for each type of GC to AT nucleotide mutation).

The region coding for 3D epitope 1 from the diphtheria toxin is well protected from missense C to T transitions (see Figure 5A), while levels of “protective buffer” against G to A (see Figure 5B), C to A (see Figure 5C) and G to T (see Figure 5D) mutations are too low for that region. There are no

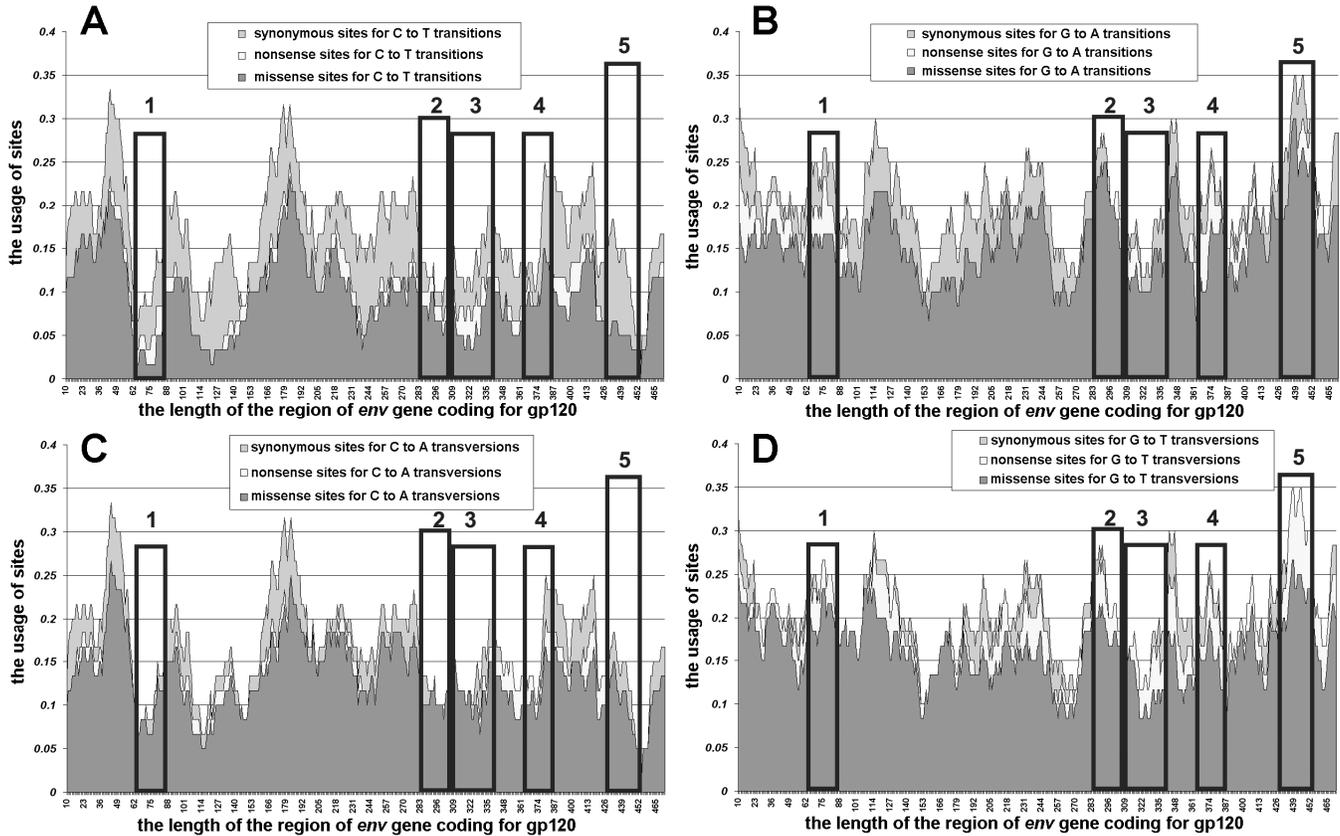
synonymous sites for G to T transversions in the region coding for 3D epitope 2 of the diphtheria toxin (see Figure D). Regions coding for 3D epitopes 3 and 4 seem to be less mutable than those coding for 3D epitopes 1 and 2 of the diphtheria toxin.

### 3.3 Percentage of different types of variable sites in sets of aligned sequences

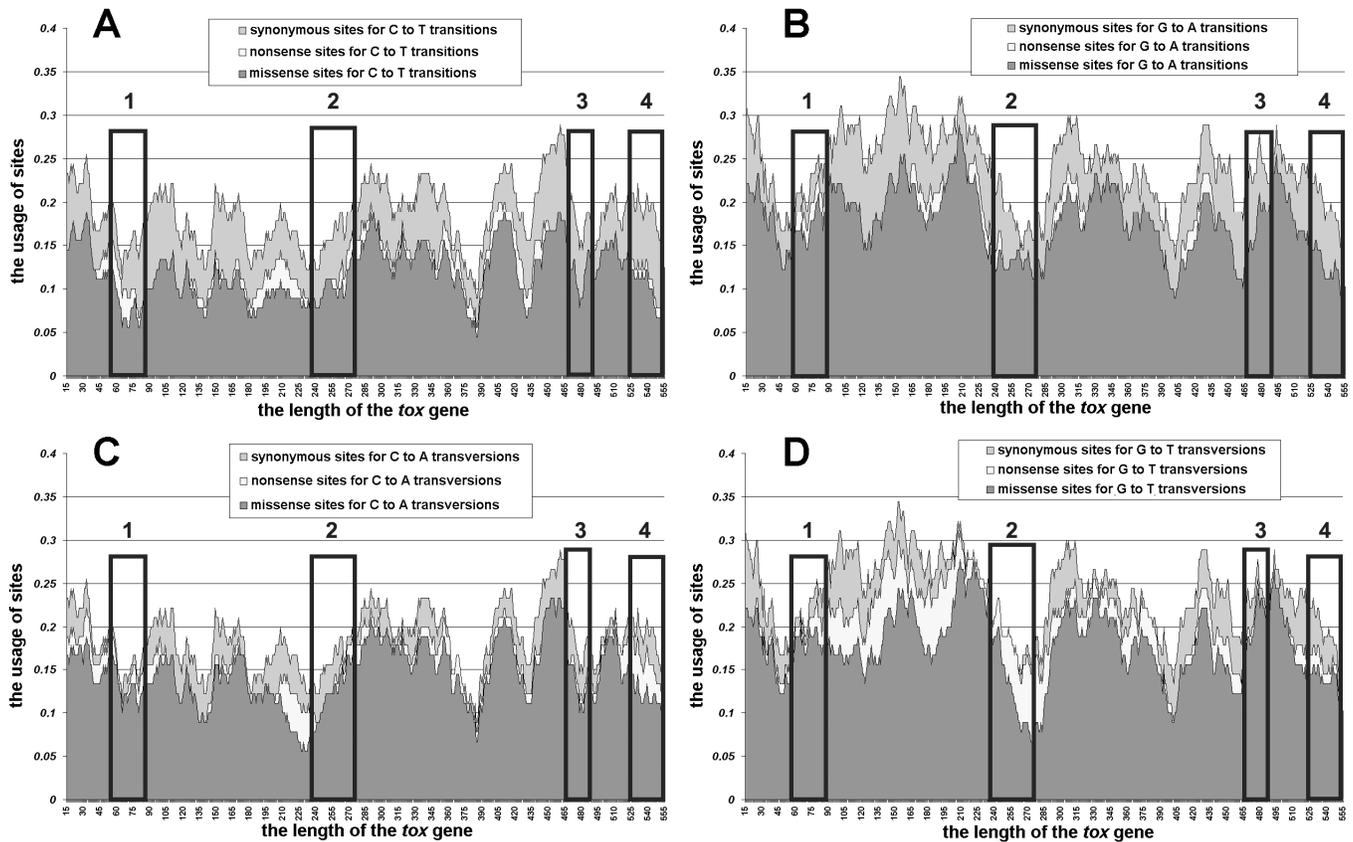
The information on the percentage of different types of variable sites in the alignment of homologous sequences helps to estimate relative frequencies of different types of mutations.

In Figure 6A average percentage of different variable sites for 34 sets of regions coding for gp120 is shown. The most frequent type of variable site among HIV1 *env* gene sequences contains adenine and guanine. The type of variable site containing thymine and cytosine is approximately two times less frequent among the sequences studied. The third place belongs to the type of variable site containing cytosine and adenine. The percent of sites containing cytosine and adenine is approximately two times higher than the percent of sites containing guanine and thymine.

In general, information given in Figure 6A is consistent with data from Figure 2. Indeed, C to A transversions should



**Figure 4.** Usage of synonymous, nonsense and missense sites for C to T transitions (A), G to A transitions (B), C to A transversions (C) and G to T transversions (D) along the length of the region of *env* gene coding for gp120. The length of sliding window is equal to 70 codons. Regions coding for 3D B-cell epitopes are designated by numbered boxes.



**Figure 5.** Usage of synonymous, nonsense and missense sites for C to T transitions (A), G to A transitions (B), C to A transversions (C) and G to T transversions (D) along the length of the *tox* gene. The length of sliding window is equal to 70 codons. Regions coding for 3D B-cell epitopes are designated by numbered boxes.

occur more frequently than G to T transversions. However, it is impossible to suggest that G to A transitions occur more frequently than C to T transitions during analysis of nucleotide usage biases in twofold degenerated sites (see Figures 2A and 2B).

The percentage given in Figure 6A can be reproduced from the most of the small groups of sequences. In all the 34 groups of sequences the percent of  $G \leftrightarrow A$  sites is higher than the percent of  $C \leftrightarrow T$  sites. In 33 from 34 groups of sequences the percent of  $C \leftrightarrow A$  sites is higher than the percent of  $G \leftrightarrow T$  sites. In the group of HIV1 *env* gene sequences from the patient designated as “ES2” (see Supplementary Material, Table 1) the percent of  $G \leftrightarrow T$  sites is higher than the percent of  $C \leftrightarrow A$  sites. This deviation is the consequence of the low number of variable sites (64) among sequences from that group. The number of transversions in that set is just 11.

The main requirement for appropriate work of “VVK VarInvar” algorithm is not the sufficient number of sequences but the sufficient number of variable sites between them.

There are 6 from 34 groups of sequences in which the percent of variable  $C \leftrightarrow A$  sites is higher than the percent of variable  $C \leftrightarrow T$  sites. This kind of deviation should not be the consequence of the low number of variable sites (see Supplementary Material, Table 1).

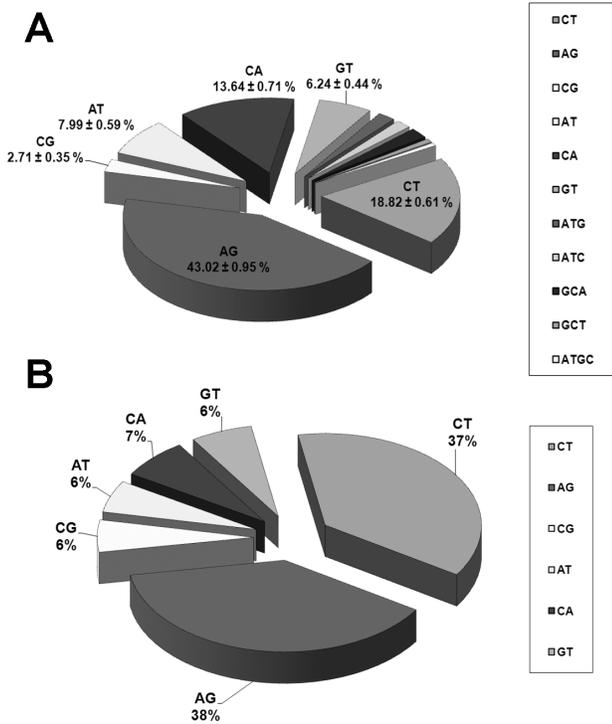
About 75% of variable sites (see Figure 6B) from the align-

ment of seventeen sequences coding for diphtheria toxin contain transitions (the total number of variable sites is equal to 141 in this alignment). The number of sites with  $A \leftrightarrow G$  transitions is approximately equal to the number of sites with  $T \leftrightarrow C$  transitions. Only 25% of variable sites contain transversions. Numbers of sites containing different types of transversions are close to each other. This information may be interpreted as an evidence that transitions occur approximately three times more frequently in *tox* gene than transversions. More attention should be paid to the mutability of regions coding for diphtheria toxin epitopes under the pressure of GC to AT transitions than to their mutability under the pressure of GC to AT transversions.

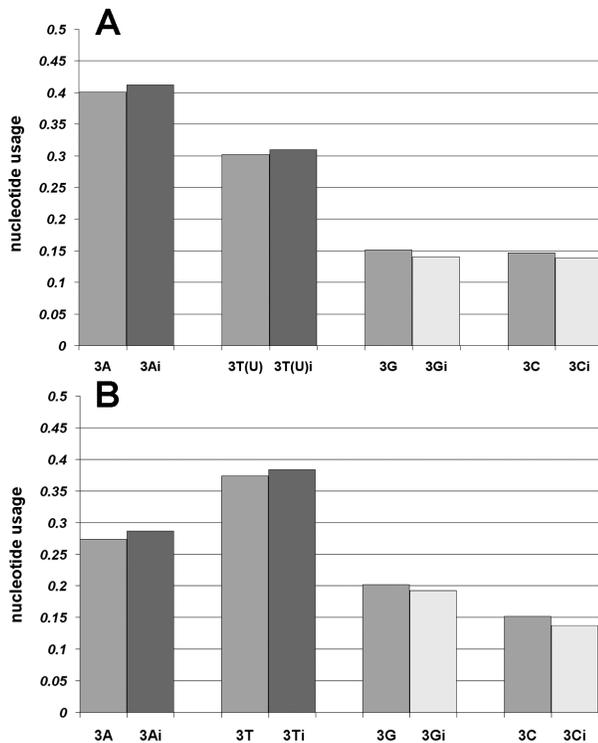
### 3.4 Nucleotide usage in invariable sites from sets of aligned sequences

According to the mutational pressure theory [3], adenine and thymine should be more stable nucleotides in both *env* and *tox* genes than guanine and cytosine. Mutations of guanine and cytosine should happen more frequently than mutations of adenine and thymine.

Indeed, the usage of guanine in invariable sites from third codon positions is significantly lower than its average usage in those positions (see Figure 7). The same situation is char-



**Figure 6.** Average percentage of different types of variable sites for 34 alignments of sequences coding for gp120 (A); percentage of different types of variable sites between seventeen sequences coding for diphtheria toxin (B).



**Figure 7.** Average levels of nucleotide usage in third codon positions and nucleotide usage in invariable sites from third codon positions of sequences coding for gp120 (A) and diphtheria toxin (B).

acteristic for cytosine usage (see Figure 7). Levels of adenine and thymine in invariable sites from third codon positions are significantly higher than their average levels of usage in third codon positions. These facts approve that adenine and thymine are much more stable than guanine and cytosine in both *env* and *tox* genes. Average levels of nucleotide usage for 34 groups of sequences are given in Figure 7A. Differences between average nucleotide usage in third codon positions and nucleotide usage in invariable sites are significant, although there are a few exceptions from common tendencies in certain groups (see Supplementary Material, Table 2).

Levels of thymine and adenine are high in third codon positions. However, the number of invariable sites containing adenine and thymine is higher than the number of invariable sites containing guanine and cytosine. This situation is possible only in case of mutational AT-pressure existence. The change of the symmetric mutational pressure direction will lead to the opposite situation (adenine and thymine will become quite instable and the number of invariable sites containing them will decrease). As it has been shown previously [9], nucleotide usage indexes will remain the same for some time after the change of mutational pressure direction. So, the calculation of nucleotide content in invariable sites from third codon position is the most sensitive indicator of the recent change in mutational pressure direction.

### 3.5 Probabilities to be missense, synonymous and nonsense for mutations of GC to AT direction, as well as amount of substrate for them, in regions coding for 3D B-cell epitopes of HIV1 gp120 and diphtheria toxin

Mutability levels of regions coding for five most immunogenic regions of HIV1 gp120, as well as for four most immunogenic regions of diphtheria toxin, have been compared in this section. There are three criterions to compare (for each type of GC to AT mutation): i) amount of the substrate for nonsynonymous (or missense) mutation; ii) amount of the substrate for synonymous mutation and iii) probability to be synonymous (or synonymous or nonsense). The region with the lowest amount of the substrate for nonsynonymous (or missense) mutation, with the highest amount of the substrate for synonymous (or synonymous or nonsense) mutation and with the highest probability of synonymous (or synonymous or nonsense) mutation has the lowest level of mutability under the pressure of certain mutations.

Amount of the substrate for synonymous (or synonymous or nonsense) mutation is a kind of “protective buffer” against amino acid replacements caused by certain type of nucleotide mutations.

In case if protein loses its function completely due to nonsense mutation, amount of the substrate for nonsense mutations should also form “protective buffer” together with amount of the substrate for synonymous mutations.

Average (for 689 sequences) usage of missense sites for C to T transitions in the region coding for 3D epitope 1 of HIV1 gp120 is very low (see Figure 8A). The highest proba-

bility to be synonymous or nonsense for G to A transition is also characteristic for that region (see Figure 8B). Results of the comparisons between five epitopes are combined in the Table 1. The first criterion for comparison is the probability to be synonymous or nonsense for the given type of mutation; the second criterion is the usage of missense sites for the given type of mutation; and the third criterion is the usage of sites for synonymous and nonsense sites for the given type of mutation.

Regions coding for epitopes 1 and 3 seem to be less mutable than regions coding for epitopes 2, 4 and 5. However, the region coding for epitope 1 is better protected from missense transitions than from transversions, while the region coding for epitope 3 is better protected from missense transversions than from transitions (see Table 1).

The amount of guanine which is prone to missense mutation is much higher in a region coding for diphtheria toxin epitope 1 than in other three regions (see Figure 9B). The amount of guanine in synonymous and nonsense sites for G to A transitions is much lower in the region coding for Epitope 1 than in three other regions. Obviously, the lowest probability to be synonymous for G to A transition is characteristic for the region coding for Epitope 1.

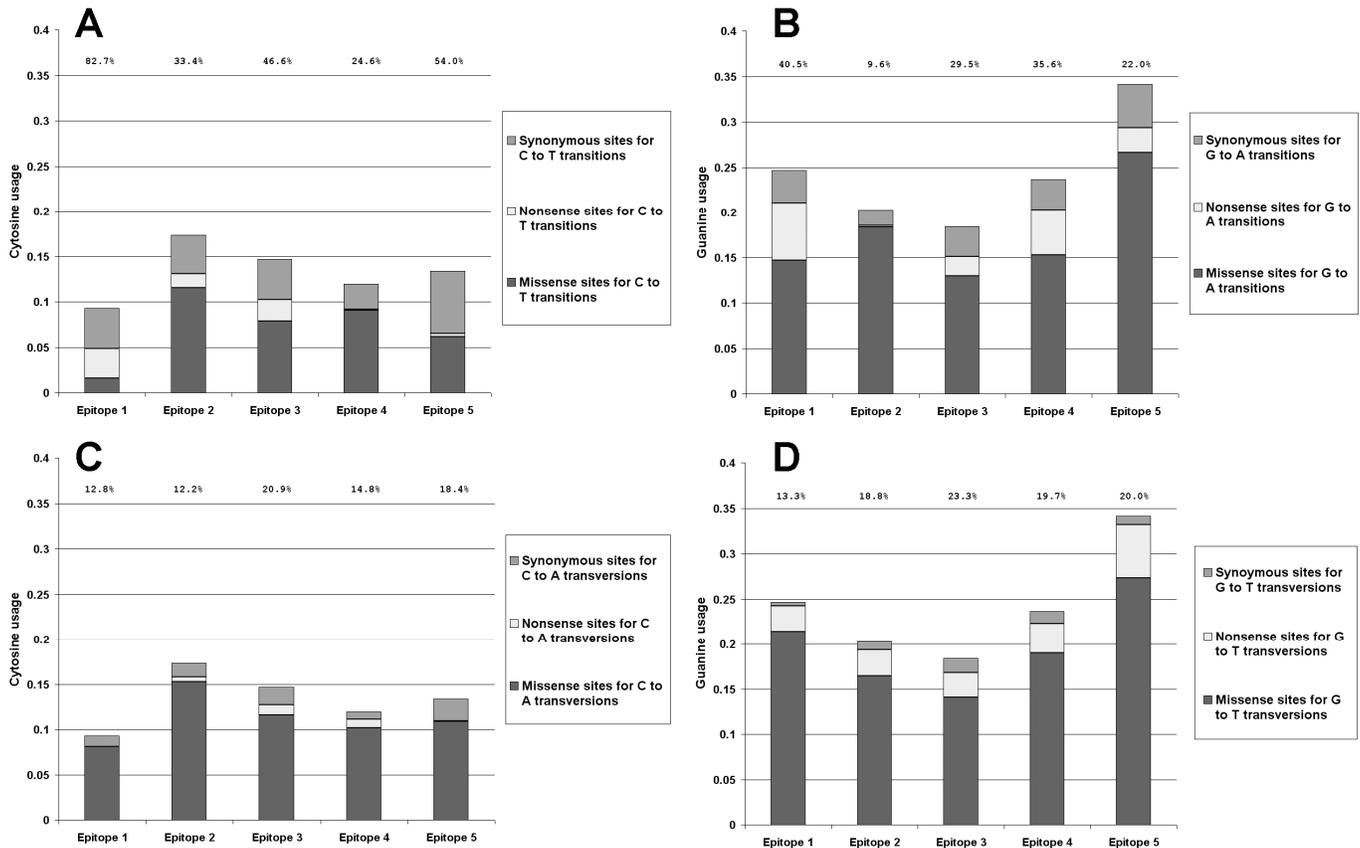
As to the mutability under the pressure of C to T muta-

tions, the region coding for epitope 1 is looking some better than three other regions (see Figure 9A and Table 2). The mutability of the region coding for epitope 2 is higher than that for the region coding for epitope 1: all three criteria show preference for the last one. However, regions coding for epitope 3 and epitope 4 have higher amount of “protective buffer” than the region coding for epitope 1.

The lowest amount of “protective buffer” against missense C to A transversions is characteristic to the region coding for epitope 1 (see Figure 9C). The highest probability to be synonymous or nonsense for C to A mutation is characteristic to the region coding for epitope 4, as well as the highest amount of the substrate for missense C to A mutations.

There is no substrate left for synonymous G to T transversions in the region coding for epitope 2 (see Figure 9D), and so its “protective buffer” is represented only by nonsense sites for G to T mutations. It is clear that the region coding for epitope 4 is the less mutable one under the pressure of G to T mutations (see Table 2), while regions coding for epitope 2 and epitope 3 are the most mutable ones.

The final conclusion is as follows: the region coding for epitope 4 is the less mutable one, since i) it is protected from missense G to T transversions better than other three ones, ii) it is protected from missense G to A transitions better



**Figure 8.** Average usage of synonymous, nonsense and missense sites for C to T transitions (A), G to A transitions (B), C to A transversions (C) and G to T transversions (D) in regions coding for five 3D B-cell epitopes of gp120. Probabilities to be synonymous or nonsense are written above the columns.

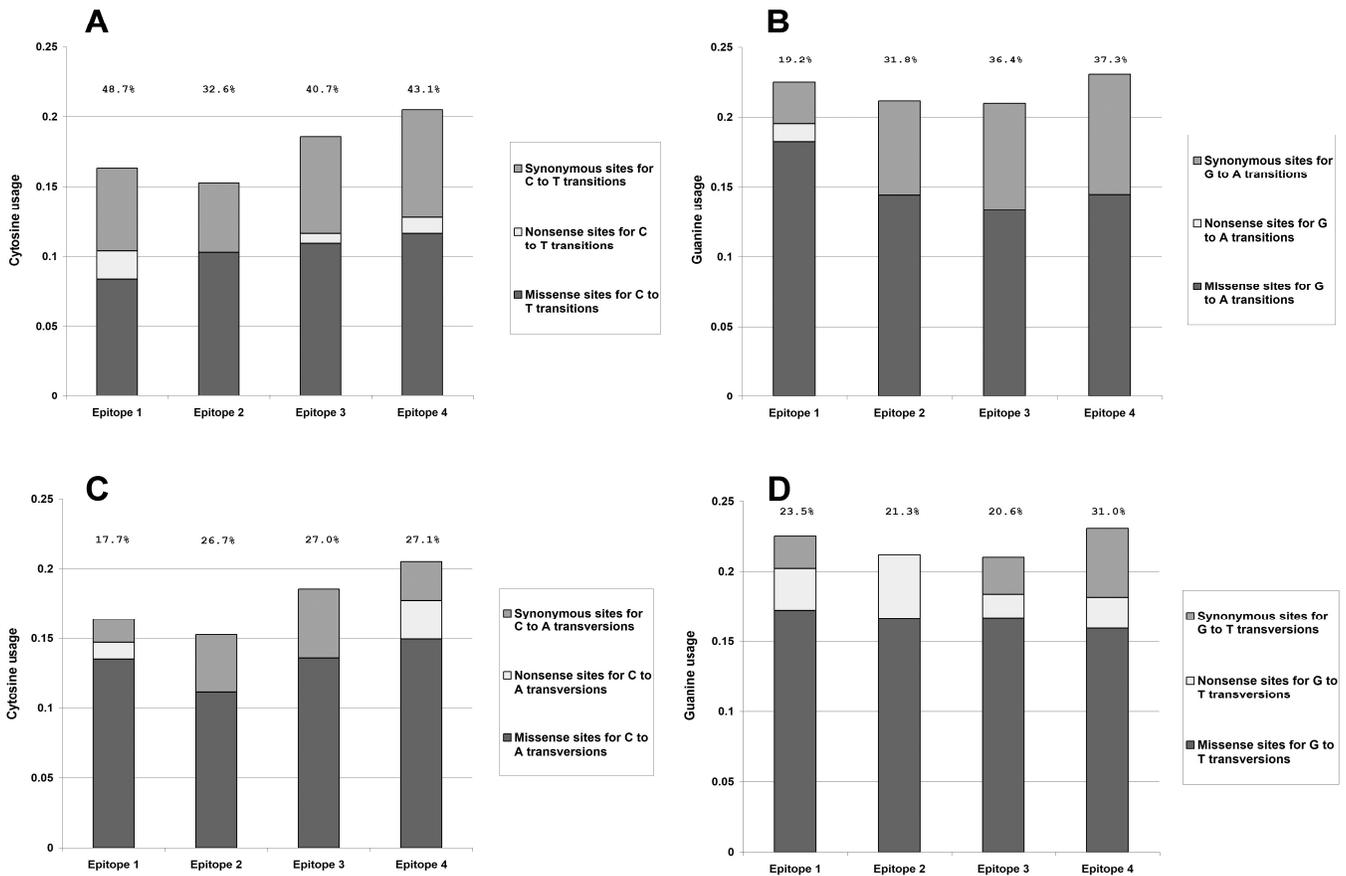
**Table 1.** Results of the application of our method to five regions of *env* gene coding for the most immunogenic 3D B-cell epitopes of gp120. For each type of mutation the region with the highest (according to the results of t-test) probability of synonymous or nonsense mutation is marked by the first “+”, the region with the lowest level of missense sites is marked by the second “+”, the region with the highest level of “protective buffer” (synonymous sites and nonsense sites) is marked by the third “+”.

Regions coding for 3D epitopes	G to A		C to T		G to T		C to A	
epitope 1	+	+	+	+	+			+
epitope 2								
epitope 3		+				+	+	+
epitope 4								
epitope 5						+		

than regions coding for epitope 1 and epitope 2, iii) it has the highest amount of “protective buffer” against nonsynonymous C to A transversions.

It has to be noted that there are no proline residues in consensus sequences of the epitope 1 of HIV1 gp120 and the

epitope 4 of diphtheria toxin, while there are a few of them in other epitopes. Mutations of proline residues (especially those caused by C to T transitions) usually have a drastic effect on length of linear B-cell epitopes [14]. It means that epitope 1 of HIV1 gp120 and epitope 4 of diphtheria toxin



**Figure 9.** Average usage of synonymous, nonsense and missense sites for C to T transitions (A), G to A transitions (B), C to A transversions (C) and G to T transversions (D) in regions coding for four 3D B-cell epitopes of diphtheria toxin. Probabilities to be synonymous or nonsense are written above the columns.

**Table 2.** Results of the application of our method to four regions of *tox* gene coding for the most immunogenic 3D B-cell epitopes of diphtheria toxin. For each type of mutation the region with the highest (according to the results of t-test) probability of synonymous or nonsense mutation is marked by the first “+”, the region with the lowest level of missense sites is marked by the second “+”, the region with the highest level of “protective buffer” (synonymous sites and nonsense sites) is marked by the third “+”.

Regions coding for 3D epitopes	G to A		C to T		G to T		C to A	
epitope 1			+	+				
epitope 2								+
epitope 3		+						+
epitope 4	+	+			+	+	+	+

are protected from those drastic effects simply because of the absence of proline.

### 3.6 Variability of four 3D B-cell epitopes from HIV1 gp120

689 amino acid sequences of HIV1 gp120 have been aligned with the help of PAM-matrix included in MEGA4 program [26]. 3D epitope 1 is quite conserved (see Table 3). There are nineteen amino acid residues (from twenty one) which can be found in a given site in more than 95% of sequences. Six of them are invariable. Either asparagine (81%) or aspartic acid (17%) can be found in the position 6; either glutamic acid (79%) or aspartic acid (18%) can be found in the position 9. These amino acid substitutions should have quite neutral consequences for the structure of the 3D epitope 1. Relatively radical amino acid substitution can be found only in the position 16. However, lysine instead of isoleucine has been found in this position only in one from 34 groups of sequences.

Analogous tables with percentage of amino acid substitutions have been created for four other 3D epitopes. Because of the high levels of variability those tables were included in Supplementary Material. There are fifteen amino acid residues (from thirty eight sites, including gaps) which can be found in a given site in more than 95% of 3D epitope 2 sequences (see Supplementary Material, Table 3). Actually, 3D epitope 2 is V3-loop, which is well-characterized immunogenic determinant of gp120 [2]. Relatively conserved part of V3-loop can be found in its N-terminal (CTRPNNNTR).

There are nine relatively conserved amino acid residues (from thirty six sites, including gaps) in 3D epitope 3. The most conserved motif from 3D epitope 3 (SSGGD) is situated in its C-terminal (see Supplementary Material, Table 4). This motif is actually a part of CD4 binding receptor of gp120 [28].

Interestingly, sequences corresponding to the epitope 4 cannot be aligned at all. There are no conserved amino acid residues in them. Gap can be found in each of the forty two

**Table 3.** Consensus sequence of the predicted 3D epitope 1 from gp120 of HIV1. Coordinates are given relatively to the length of gp160 and gp120.

94 / 66	95 / 67	96 / 68	97 / 69	98 / 70	99 / 71	100 / 72
<b>N99,71%</b>	<b>M97,54%</b>	<b>W100%</b>	<b>K96,96%</b>	<b>N100%</b>	N81,01%	<b>M99,71%</b>
K0,14%;	A2,17%;		E1,59%;		D16,52%;	V0,29%;
D0,14%;	V0,29%;		R1,16%;		G2,17%;	
			T0,29%;		H0,29%;	
101 / 73	102 / 74	103 / 75	104 / 76	105 / 77	106 / 78	107 / 79
<b>V99,57%</b>	E79,13%	<b>Q100%</b>	<b>M99,86%</b>	<b>H97,54%</b>	<b>E97,54%</b>	<b>D100%</b>
I0,29%;	D17,54%;		I0,14%;	Q2,46%;	G1,3%;	
A0,14%;	Q2,46%;				T1,16%;	
	N0,58%;					
	K0,14%;					
	G0,14%;					
108 / 80	109 / 81	110 / 82	111 / 83	112 / 84	113 / 85	114 / 86
<b>I96,52%</b>	I93,91%	<b>S98,12%</b>	<b>L100%</b>	<b>W100%</b>	<b>D98,84%</b>	<b>Q97,25%</b>
V3,48%;	K5,8%;	N1,74%;			E0,87%;	E2,75%;
	V0,29%;	G0,14%;			N0,14%;	
					G0,14%;	

sites of the alignment at least in one sequence (see Supplementary Material, Table 5).

There are thirteen conserved amino acid residues (from twenty nine sites, including gaps) in the alignment of 689 sequences corresponding to the 3D epitope 5 of gp120 (see Supplementary Material, Table 6). The most conserved (and relatively long) motif can be found in the C-terminal of that epitope (FRPGGGDMRDNR).

In general, the less mutable epitope of HIV1 gp120 (3D epitope 1) is the less variable one. However, relatively conserved motif from the 3D epitope 5 seems to be highly mutable. This situation can be caused by the strong negative selection preventing fixation of amino acid substitutions in the C-terminal of the 3D epitope 5. Nonfunctional gp120 proteins containing amino acid mutations in the abovementioned part of the 3D epitope 5 should occur frequently. It means that mutability of the region coding for the 3D epitope 5 may still take part in the “deception” of the immune system.

### 3.7 Modified ELISA test-system with the peptide NQ21 corresponding to the consensus sequence of 3D epitope 1 of gp120

To confirm that the less mutable and the less variable 3D epitope of gp120 predicted by DiscoTope 1.2 is really immunogenic the peptide (NQ21) corresponding to its consensus sequence has been synthesized. Commercial ELISA test-system has been modified to check the presence of antibodies cross-reacting with the peptide NQ21 in serums of HIV1-infected persons. Peptide NQ21 conjugated with biotin via its N-terminal has been added to that ELISA test-system instead of biotinylated recombinant gp160 and gp140 proteins. Levels of optical density (OD) for wells with serums from 234 HIV-negative persons were not higher than 3 standard deviations from their average level for each from four plaques. Levels of OD higher than 3 standard deviations (relatively to the average level of OD for serums from HIV1-negative persons) have been found only in serums from HIV1-positive persons. The presence of antibodies recognizing 3D epitope 1 of the recombinant gp160 protein (adsorbed in plaques from the commercial ELISA test-system) which are able to cross-react with the peptide NQ21 have been confirmed in serums of 80,22% (in 73 from 91) HIV1-infected persons.

This data approves that predicted epitope 1 of gp120 is a real B-cell epitope. Moreover, this epitope is conserved enough to be included in polyvalent vaccines against HIV1.

### 3.8 Affinity purification of antibodies cross-reacting with the peptide NQ21

2.3 mg of the peptide NQ21 has been immobilized on the column from AminoLink Plus Immobilization Trial Kit (“Thermo scientific Inc.”). Serum from HIV-negative person showed no cross-reactivity with immobilized peptide corresponding to the 3D epitope 1 of gp120 (see Figure 10A). There was a clear peak of fluorescence measured by Hitachi

650-60 spectrofluorometer in the third eluate for the serum from HIV1-positive person (see Figure 10B). According to the results of SDS-PAGE with 2-mercaptoethanol (see Figure 10C), there were four types of immunoglobulins in the third eluate. According to our results, IgG, IgM, IgA and IgE molecules can be synthesized against the epitope 1 of gp120.

## 4. Discussion

### 4.1 Mutability and variability

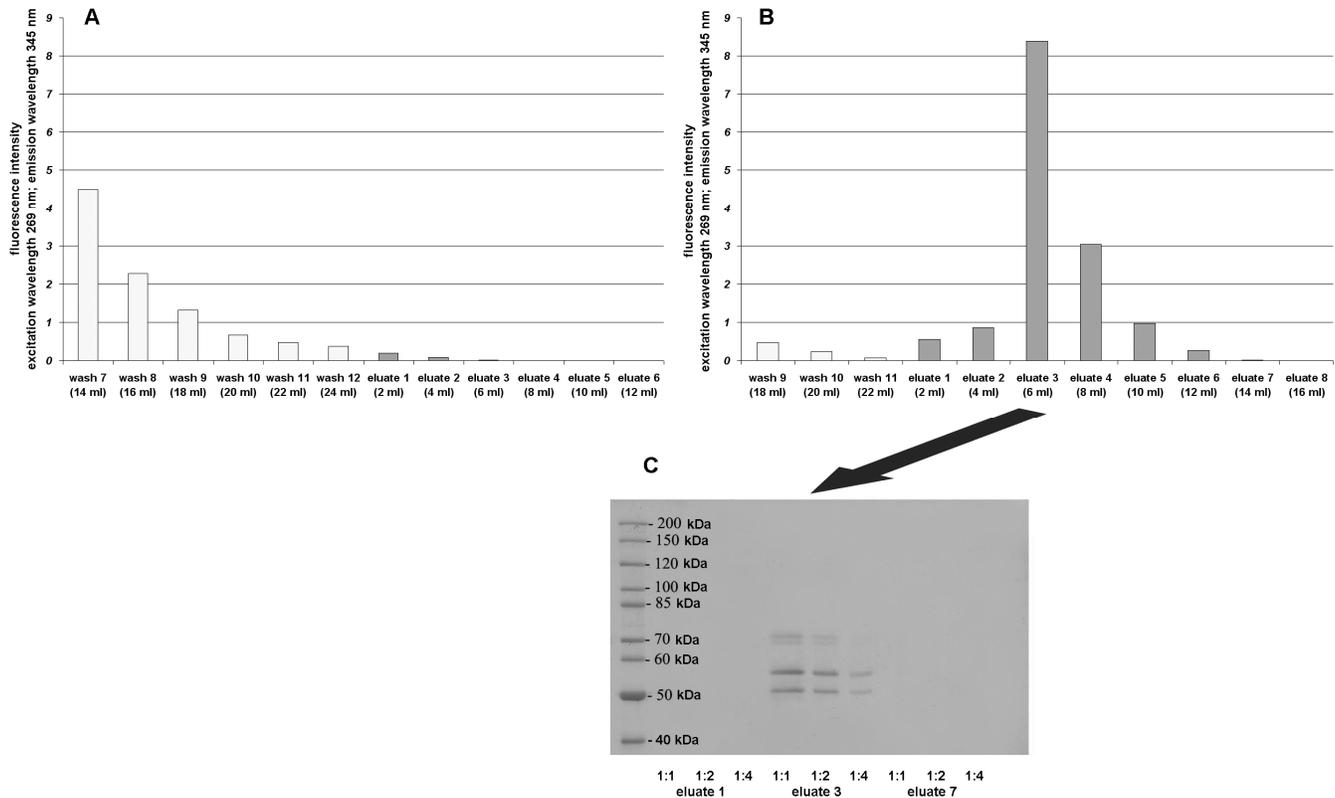
Our method has been created for estimation of mutability under the influence of mutational pressure. Some regions of a gene are usually more prone to missense mutations caused by mutational pressure than others. In our terminology these regions are “mutable” [2]. It means that missense nucleotide mutations will happen in them at a higher probability than in less mutable regions. However, mutable regions may theoretically be highly conserved [2]. In this case missense mutations will occur in them frequently, while all the amino acid replacements caused by them will be eliminated by natural selection.

In our previous study on HIV1 gp120 3D epitopes it has been found that its less mutable 3D epitope is the most conserved one [2]. Method which is able to predict level of mutability was used in that study [2], while in the present study improved and much more accurate method was introduced. It is impossible to check whether predicted 3D epitopes of diphtheria toxin are conserved or variable simply because of the small number of sequences and variable sites between them. That is why data obtained with the help of our method should be especially beneficial for studies on diphtheria toxin vaccine design.

### 4.2 Experimental data on gp120 immunogenicity

The statement that conserved regions of HIV1 gp120 are poorly immunogenic, while highly immunogenic regions of this protein are extremely variable can be found in many sources [29]. Several regions of gp120 that were shown to be recognized by monoclonal antibodies can be found in the HIV molecular immunology database ([http://www.hiv.lanl.gov/content/immunology/tables/ab\\_summary.htm](http://www.hiv.lanl.gov/content/immunology/tables/ab_summary.htm)). Some of those regions include N-terminal, C-terminal or central part of the epitope 1 predicted by us, but never the whole immunogenic determinant. Some of those antibodies (recognizing the epitope which includes N-terminal of the NQ21) were even shown to be neutralizing [30]. However, antibodies to the C1 region of gp120 are thought to bind monomeric and not oligomeric protein [29]. From this point of view, epitopes from C1 region of gp120 were considered to be bad targets for protective immunity development [29].

Attention should be paid to the fact that the main part of 3D epitope 1 is presented by long alpha-helix (see Figure 11). This alpha-helix is amphiphilic: one half of it is hydrophobic, while another half is hydrophilic. Synthesis of smaller pep-



**Figure 10.** Intensity of fluorescence in washes and eluates collected during affinity purification of the serum from HIV1-negative person (A) and HIV1-positive person (B); results of SDS-PAGE analysis of eluates 1, 3 and 7 collected during affinity purification of the serum from HIV1-positive person (C).

tides might lead to the partial destruction of that alpha-helix and to the loss of their cross-reactivity with many types of antibodies recognizing the corresponding part of native molecule.

Neutralizing antibodies to gp120 (preventing interactions between gp120 and CD4 molecules) were shown to stimulate infection of macrophages [28]. Since macrophages are thought to be a good reservoir for HIV-infection, the idea of the creation of vaccine stimulating synthesis of neutralizing antibodies may seem to be compromised.

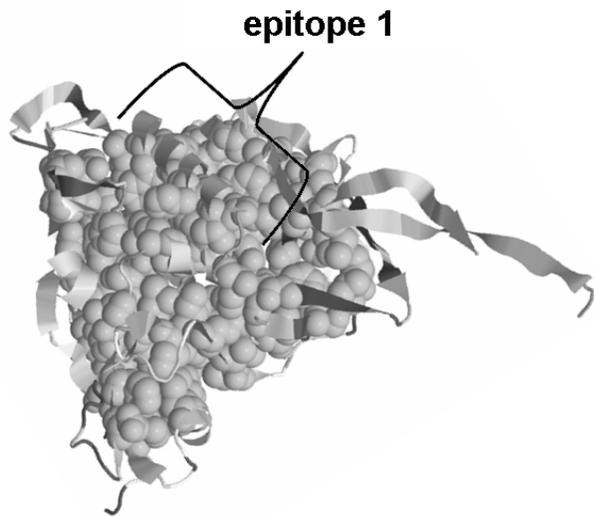
In the present study it has been shown that antibodies to the most immunogenic part of C1 region can be found in more than 80% of HIV1-infected persons. The fact that those antibodies were not found in serums of approximately 20% of HIV1-infected persons can be explained by at least three hypotheses. It is important to highlight that serums from persons with recently revealed HIV1-infection were tested in the present work. Probably, antibodies against gp120 3D epitope 1 in some of those persons have not been synthesized yet. In other words, those antibodies may not be synthesized in the yearly period of HIV1-infection in certain persons. On the other hand, antibodies against gp120 3D epitope 1 may not cross-react with the NQ21 peptide due to mutations in that region of viral protein. Even though 3D epitope 1 is the most conserved one, mutations disturbing cross-reactivity with NQ21 may still happen inside it. Once

again it has to be noted that mutations should happen in 3D epitope 1 much less frequently than in other epitopes. Finally, antibodies to certain epitopes of gp160 may somehow disturb or totally prevent binding of antibodies against 3D epitope 1 to adsorbed molecules in our modified ELISA test-system.

In case if antibodies cross-reacting with NQ21 are synthesized mostly against monomeric (and not oligomeric) viral glycoprotein, they should be able to bind gp120 molecules situated on the membrane of infected cells. In one of the experimental works C1 region of gp120 was shown to be the best target for antibodies with antibody-dependent cellular cytotoxic activity (ADCC) among other regions of this protein [31, 32]. Indeed, NK-cells recognize antibodies bound to viral antigens situated on the cellular membrane and kill infected cells. In our opinion, immunization against the peptide NQ21 should cause production of antibodies with antibody-dependent cellular cytotoxic activity. High titer of antibodies with ADCC is one of the predictors of slow progression of HIV-infection [33, 34].

#### 4.3 Experimental data on diphtheria toxin immunogenicity

Diphtheria toxin consists of two chains connected by a single disulfide bond [35]. Precursor of diphtheria toxin encoded by *tox* gene is cleaved by proteases into Chain A (N



**Figure 11.** Epitopia output for the X-ray structure (PDB ID: 3JWO) of gp120. Buried amino acid residues are shown as balls.

-terminal) and Chain B (C-terminal). There is also a short signal peptide encoded by the first 25 codons of *tox* open reading frame [36].

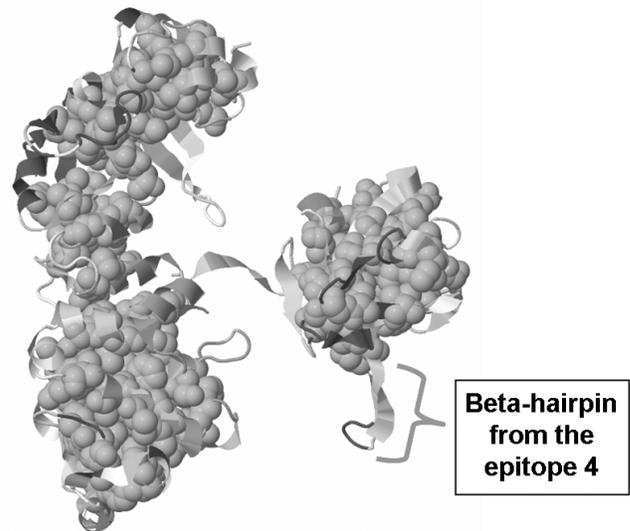
Chain A forms catalytic domain of the toxin. Catalytic domain blocks protein synthesis by causing an ADP-ribosylation of elongation factor 2, thus provoking cell death [35]. Epitope 1 is situated on the surface of chain A.

There are two distinct domains in chain B of the toxin (see Figure 12). One of them (N-terminal part of chain B) is predominantly alpha-helical one. This domain is called “translocation domain” or “transmembrane domain”. Indeed, long alpha-helices from this domain are involved in translocation of the catalytic domain from endosome to cytoplasm. Actually, acidic pH existing in endosome causes conformational changes in the structure of translocation domain leading to the penetration of endosome membrane by two long alpha-helices [35]. Epitope 2 is situated on the surface of translocation domain.

C-terminal part of B Chain is predominantly beta-structural (see Figure 12). This domain is able to bind its specific receptor on the surface of cells [35]. Both epitope 3 and epitope 4 are situated on a surface of receptor-binding domain. Epitope 4 really consists of two immunogenic regions situated close to each other in the primary structure, but relatively far from each other in the tertiary structure of the toxin (see Figure 12).

Many researchers tried to use recombinant proteins representing certain parts of diphtheria toxin in vaccine design studies [1, 6, 37, 38]. It has been shown [6, 38] that recombinant Chain B of the toxin is able to induce toxin-neutralizing antibodies in laboratory animals (rabbits and guinea pigs). Moreover, recombinant receptor-binding domain of Chain B is also able to induce formation of toxin-neutralizing antibodies [1].

#### 4.4 Solutions for antigen design for future recombinant diph-



**Figure 12.** Epitopia output for the X-ray structure (PDB ID: 1SGK) of diphtheria toxin. Buried amino acid residues are shown as balls.

#### *theria toxin vaccines*

According to our results, the target for neutralizing antibodies (receptor-binding domain) possesses the less mutable 3D epitope. It means that the usage of recombinant receptor-binding domain in new vaccine against the diphtheria toxin is well justified. Abovementioned recombinant receptor-binding domain may also be a good antigen for production of ELISA diagnostic tests to control the level of neutralizing antibodies.

N-terminal part of the epitope 4 is situated on a surface of the receptor-binding domain, while C-terminal part of it forms a protruding structure (see Figure 12). This structure is formed by two beta strands connected together (in a beta structure) and a loop between them (see Figure 12). This loop containing a 3/10 helix is recognized as linear B-cell epitope by BepiPred 1.0 [13], and as 3D epitope by Disco-Topo 1.2 [22] (see Figure 1). This loop is accessible to a solvent according to Epitopia [23] prediction (see Figure 12). In our opinion, antibodies against synthetic peptide with the amino acid sequence of this beta-hairpin (SIGVLGYQKTVDHTKVNSKLSLF) may be able to cross-react with the subsequent region of diphtheria toxin and *vice versa*.

There are a few amino acid substitutions between *C. diphtheria* and *C. ulcerans* toxins in epitope 1 and epitope 2 (see Figure 13). Epitope 3 is quite invariable at least in seventeen sequences studied. There are six amino acid substitutions between sequences of *C. diphtheria* and *C. ulcerans* toxins in epitope 4 [4]. All of them are concentrated in the center of this region, which is not recognized as highly antigenic by all the methods used (see Figure 1). However, there is a possibility that at least certain antibodies against this epitope of *C. diphtheria* toxin are not able to bind *C. ulcerans* toxin. So, antibodies to this region of the diphtheria toxin may be used

	51		78
<i>C. diphtheriae</i> phages	<b>GYVDS</b> IQKGIQKPKSGTQGNYYDDWKG <b>F</b>		
		T	
<i>C. ulcerans</i> phages	<b>GYVDS</b> IQKGIQKPKSGAQQNYYDDWKG <b>F</b>		
		T	
	232		268
<i>C. diphtheriae</i> phages	<b>DVIRDKTKTKIESLKEHGPIKNKMS</b> EPNKTV <b>SE</b> EKA		
<i>C. ulcerans</i> phages	<b>DAIRDKTKTKIESLKEHGPIKNKMS</b> EPNKTV <b>SE</b> EKA		I
	463		482
<i>C. diphtheriae</i> phages	<b>PGKLDV</b> NKSKTHISVNGR <b>KI</b>		
<i>C. ulcerans</i> phages	<b>PGKLDV</b> NKSKTHISVNGR <b>KI</b>		
	519		548
<i>C. diphtheriae</i> phages	<b>SSSEKI</b> HSNEISSDSI <b>G</b> V <b>LG</b> YQKTVDHT <b>KV</b>		
<i>C. ulcerans</i> phages	<b>SSSEKI</b> HSDE <b>T</b> PL <b>S</b> SI <b>D</b> V <b>LG</b> YQKTVDHT <b>KV</b>		

**Figure 13.** Alignment of amino acid sequences with predicted 3D-epitopes. Invariable amino acid residues are written in bold. Amino acid variations are written separately for toxins from *Corynebacterium diphtheriae* and *Corynebacterium ulcerans* phages.

in ELISA assays to discriminate between *C. diphtheriae* and *C. ulcerans* toxins. The beta-hairpin described above (at least its highly antigenic loop) seems to be conserved between *C. diphtheriae* and *C. ulcerans* toxins.

## 5. Concluding Remarks

A method for estimation of mutability levels for immunogenic determinants includes two steps: i) estimation of mutational pressure direction and ii) selection of the less mutable immunogenic determinant (determinants). The second step is based on the estimation of the amount of mutable nucleotides in missense, synonymous and nonsense sites of regions coding for immunogenic determinants for each type of the most commonly occurring nucleotide mutations in the given gene. The region coding for the less mutable immunogenic determinant should satisfy three criteria: i) a probability to be synonymous (or synonymous or nonsense) for the most common types of nucleotide mutations should be the highest one inside it; ii) it should have the lowest amount of mutable nucleotides in nonsynonymous (or in missense) sites for the most common types of nucleotide mutations; iii) it should have the highest amount of mutable nucleotides in synonymous (or in synonymous and nonsense) sites for the most common types of nucleotide mutations.

Our *in silico* method showed a good performance on HIV1 gp120 protein. Predictions have been successfully confirmed *in vitro*.

## 6. Supplementary material

Supplementary material regarding this manuscript is online available in the web page of JIOMICS.

<http://www.jiomics.com/index.php/jio/rt/suppFiles/64/0>

Table 1. Percentage of variable sites among sequences coding for HIV1 gp120 from each of the 34 monophyletic sets.

Table 2. Average levels of nucleotide usage in third codon

positions (3A; 3T(U); 3G and 3C) in comparison with nucleotide usage in invariable sites from third codon positions (3Ai; 3T(U)i; 3Gi and 3Ci) for each from 34 sets of monophyletic sequences coding for HIV1 gp120.

Table 3. Consensus sequence of the predicted 3D epitope 2 from gp120 of HIV1.

Table 4. Consensus sequence of the predicted 3D epitope 3 from gp120 of HIV1.

Table 5. Consensus sequence of the predicted 3D epitope 4 from gp120 of HIV1.

Table 6. Consensus sequence of the predicted 3D epitope 5 from gp120 of HIV1.

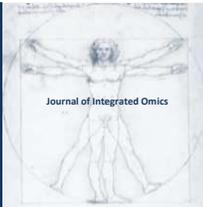
## Acknowledgements

Authors thank Dr. Oleg Vladimirovich Stoma, the General Director of the MedVax Company (Minsk, Belarus) and Professor Alexander Stanislavovich Vladiko, the Head of the Laboratory of Biotechnology and Immunodiagnosics of Highly Dangerous Infections from Republican Research and Practical Centre for Epidemiology and Microbiology (Minsk, Belarus) for support and useful consultations.

## References

1. K. Lobeck, P. Drevet, M. Léonetti, C. Fromen-Romano, F. Ducancel, E. Lajeunesse, C. Lemaire, A. Ménez, *Infect. Immun.* 66 (1998) 418–423.
2. V.V. Khrustalev, *Immunol. Invest.* 39 (2010) 551–569. DOI: 10.3109/08820131003706313
3. N. Sueoka, *Proc. Natl. Acad. Sci. USA* 85 (1988) 2653–2657.
4. A. Sing, M. Hogardt, S. Bierschenk, J. Heesemann, *J. Clin. Microbiol.* 41 (2003) 4848–4851.
5. A. M. Jr. Pappenheimer, *Ann. Rev. Biochem.* 46 (1977) 69–94.
6. A. A. Kaberniuk, O. S. Oliinyk, D. V. Kolybo, S. V. Komisarenko, *Ukr. Biokhim. Zh.* 81 (2009) 92–101.
7. H. Nakao, K. Mazurova, T. Glushkevich, T. Popovic, *Res. Microbiol.* 148 (1997) 45–54.
8. V. V. Khrustalev, E. V. Barkovsky, *SciTopics* (2011) Retrieved from [http://scitopics.com/Original\\_computer\\_algorithms\\_for\\_studies\\_on\\_directional\\_mutational\\_pressure.html](http://scitopics.com/Original_computer_algorithms_for_studies_on_directional_mutational_pressure.html)
9. V. V. Khrustalev, E. V. Barkovsky, *Genomics Proteomics Bioinformatics* 8 (2010) 22–32. DOI: 10.1016/S1672-0229(10)60003-4
10. V. V. Khrustalev, *Immunol. Invest.* 38 (2009) 613–623. DOI: 10.1080/08820130903062202
11. T. P. Hopp, K. R. Woods, *Mol. Immunol.* 20 (1983) 483–489.
12. J. Kyte, R. Doolittle, *J. Mol. Biol.* 157 (1982) 105–132.
13. J. E. P. Larsen, O. Lund, M. Nielsen, *Immunome Res.* 2 (2006) 2.
14. V. V. Khrustalev, *Molecular Immunology* 47 (2010) 1635–1639. doi:10.1016/j.molimm.2010.01.006
15. V. V. Khrustalev, E. V. Barkovsky, *Journal of Theoretical Biology* 282 (2011) 71–79. doi:10.1016/j.jtbi.2011.05.018
16. J. R. Lobry, N. Sueoka, *Genome Biol.* 3 (2002) 0058.
17. V. V. Khrustalev, E. V. Barkovsky, *Genomics* 96 (2010) 173–

180. doi:10.1016/j.ygeno.2010.06.002
18. E. M. Bunnik, L. Pisas, A. C. van Nuenen, H. Schuitemaker, J. Virol. 82 (2008) 7932-7941.
  19. S. Franssen, G. Bridger, J. M. Whitcomb, J. Toma, E. Stawiski, N. Parkin, C. J. Petropoulos, W. Huang, Antimicrob. Agents Chemother. 52 (2008) 2608-2615.
  20. J. R. Bailey, K. G. Lassen, H. C. Yang, T. C. Quinn, S. C. Ray, J. N. Blankson, R. F. Siliciano, J. Virol. 80 (2006) 4758-4770.
  21. M. Sagar, O. Laeyendecker, S. Lee, J. Gamiel, M. J. Wawer, R. H. Gray, D. Serwadda, N. K. Sewankambo, J. C. Shepherd, J. Toma, W. Huang, T. C. Quinn, J. Infect. Dis 199 (2009) 580-589.
  22. P. H. Andersen, M. Nielsen, O. Lund, Protein Science 15 (2006) 2558-2567.
  23. N. D. Rubinstein, I. Mayrose, E. Martz, T. Pupko, BMC Bioinformatics 10 (2009) 287.
  24. S. Liang, D. Zheng, C. Zhang, M. Zacharias, BMC Bioinformatics 10 (2009) 302.
  25. V. V. Khrustalev, E. V. Barkovsky. International Journal of Bioinformatics Research and Applications 7 (2011) 82-100.
  26. K. Tamura, J. Dudley, M. Nei, S. Kumar. Mol. Biol. Evol. 24 (2007) 1596-1599.
  27. E. V. Barkovsky, V. V. Khrustalev, Mol. Gen. Microbiol. Virol. 24 (2009) 17-23.
  28. R. L. Dunfee, E. R. Thomas, D. Gabuzda, Retrovirology 6 (2009) 69.
  29. J. P. Moore, Q. J. Sayntientau, R. Wyatt, J. Sodroski, J. Virol., 68 (1994) 469-484.
  30. G. R. Nakamura, R. Byrn, K. Rosenthal, J. P. Porter, M. R. Hobbs, L. Riddle, D. J. Eastman, D. Dowbenko, T. Gregory, B. M. Fendly, AIDS Res Hum Retroviruses 8 (1992) 1875-1885.
  31. A. Chung, I. Stratov, E. Rollman, S. J. Kent, 4th IAS Conference on HIV Pathogenesis, Treatment and Prevention, Sydney, Australia, (2007) abstract [TUPEA015].
  32. A. Chung, E. Rollman, L. J. Center, S. J. Kent, I. Stratov, J. Immunol. 182 (2009) 1202-1210.
  33. D. N. Forthal, G. Landucci, T. B. Phan, J. Becerra, J. Virol. 79 (2005) 2042-2049.
  34. V. R. Gómez-Román, C. Cao, Y. Bai, H. Santamaría, G. Acero, K. Manoutcharian, D. B. Weiner, K. E. Ugen, G. Gevorkian, J. Acquired Immune Deficiency Syndromes. 31 (2002) 147-153.
  35. C. E. Bell, D. Eisenberg, Biochemistry 36 (1997) 481-488.
  36. V. Kolodkina, L. Titov, T. Sharapa, O. Drozhzhina, Mol. Gen. Mikrobiol. Virusol. 1 (2007) 22-29.
  37. V. Y. Perera, M. J. Corbel, Epidemiol. Infect. 105 (1990) 457-468.
  38. D. V. Nascimento, E. M. B. Lemes, J. L. S. Queiroz, J. G. J. Silva, H. J. Nascimento, E. D. Silva, R. J. Hirata, A. A. S. O. Dias, C. S. Santos, G. M. B. Pereira, A. L. Mattos-Guaraldi, G. R. G. Armoa, Brazilian Journal of Medical and Biological Research 43 (2010) 460-466.



ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v1i2.57

# Application of high content biology demonstrates differential responses of keratin acetylation sites to short chain fatty acids and to mitosis

AQ Khan<sup>1,2</sup>, L Shaw<sup>1</sup>, PJM Drake<sup>1</sup>, GJ Griffiths<sup>3</sup>, SR Brown<sup>2</sup>, BM Corfe<sup>1\*</sup>.

<sup>1</sup>Department of Oncology, University of Sheffield, Medical School, Beech Hill Road, SHEFFIELD S10 2JF, UK; <sup>2</sup> Department of Gastroenterology, Northern General Hospital, SHEFFIELD, S5 7AU, UK; <sup>3</sup> Imagen Biotech Ltd, 48 Grafton Street, Manchester, M13 9XX, UK.

Received: 05 December 2010 Accepted: 02 July 2011 Available Online: 14 July 2011

### ABSTRACT

The intermediate filament cytoskeleton in epithelial tissues is formed of keratin heterodimers. Keratins are highly post-translationally modified proteins, with tyrosine phosphorylation, serine phosphorylation, and glycosylation amongst reported modifications. We and others have recently reported multiple acetylation sites on keratin 8 and we have previously shown that these sites are responsive to butyrate. In this study, we report the application of cellomic approaches to demonstrate differential responses of three lysine acetylations (lys 10, lys 471 and lys 482) to different short-chain fatty acids. The data imply no fixed hierarchy of acetylation on keratin 8, and furthermore imply different ranges of histone deacetylase (HDAC) inhibitory specificities for short chain fatty acids (SCFA). Furthermore we have used the functionality of the High Content Analysis (HCA) platform to show that the acetylation sites are differentially modified in cells undergoing mitosis. Taken together the data imply distinct roles for keratin acetylations in function.

**Keywords:** Keratin; Acetylation; Colorectal cancer; Short chain fatty acids; High content biology.

### 1. Introduction

Short-chain fatty acids (SCFAs) are the principal by-products of fibre fermentation in the gastrointestinal epithelium. The majority of SCFAs comprise of acetate, propionate, and butyrate with contributions of valerate and branched chain fatty acids (BCFA). All SCFAs are weak acids existing predominantly in the anionic, dissociated form in the colonic lumen(1). Butyrate has been shown to induce apoptosis (2), inhibit proliferation (3, 4), and promote a more differentiated phenotype (3, 4). This is considered to be through its ability to regulate gene expression by inhibition of the histone deacetylases (HDAC) (5, 6), which also regulate many non-histone proteins including transcription factors such as Sp1 (7), nuclear structural proteins (8), and p53 (9). We have recently shown keratin 8 (K8), an intermediate filament (IF) protein of the colonocyte cytoskeleton, is acetylated and its acetylation is butyrate-responsive (10).

Intermediate filaments (IFs) are one component of the eukaryotic cytoskeleton. They are formed from heterodimers of type 1 and type 2 keratins. Among various keratins pre-

sent in colonocytes, keratin 8 and 18 (K8, K18) is the principal pair (11). Other keratins include K7, K19, and K20. Dimerisation of keratins is mediated through pairing along their coiled-coil domains and is regulated, at least in part, through post-translational modification of the globular domains outside these regions. Keratins provide mechanical strength to epithelia in general and enable colonocytes to resist various chemical and mechanical stresses. They are also involved in various regulatory functions of cell including cell cycle, cell differentiation and apoptosis (12). Importance of K8 in intestinal epithelium was demonstrated through generation of K8 knockout mice. K8 null mice reaching to adulthood showed high rates of gastrointestinal epithelial inflammation and colonic hyperplasia (13) and also showed impaired electrolyte and fluid transport (14). In humans a subset of patients with inflammatory bowel disease (ulcerative colitis) has been shown to carry missense mutations in the Keratin 8/18 genes. Reconstructions of these mutations *in vitro* (K8: G62C, I63V, K464N; K18: S230T)

\*Corresponding author: Dr. Bernard Corfe, Department of Oncology, The University of Sheffield, The Medical School, Beech Hill Road, SHEFFIELD, S10 2JF, UK; Email Address: b.m.corfe@shef.ac.uk.

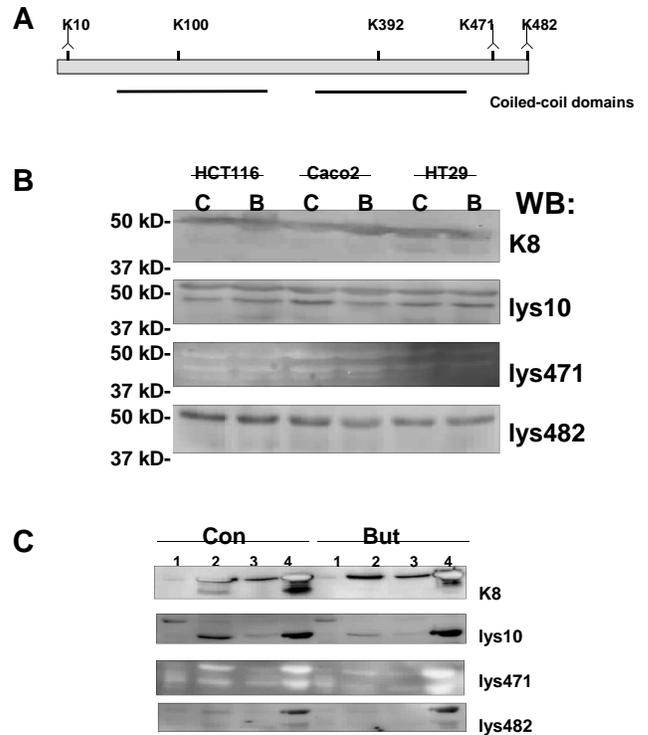
were associated with reduced filament assembly (15). An important characteristic of keratins is their relative stability of expression even after transformation to pathological state including transformation of normal cells into malignant cells. This property has enabled keratins to be applied as tumour markers (16). Changes in K8 expression during colorectal adenomagenesis and carcinogenesis have been studied by Polley *et al.* (17) by comparing protein expressions in normal and tumour (adenoma, carcinoma) tissue. They found the up regulation of various isoforms of K8 in morphologically normal mucosa from polyp and cancer patients compared with normal mucosa of patients with no pathology. Also, four K8 isoforms were over expressed in polyp mucosa relative to normal mucosa and seven K8 isoforms were over expressed in cancer mucosa relative to healthy mucosa. These findings indicate that expression or modification of K8 alters as the adenoma-carcinoma sequence progresses. Another study showing differential K8 expression in colorectal carcinoma has shown reduced expression of K8 in colorectal cancer is significantly associated with shorter patient's survival (18).

The structural and regulatory functions of keratins are in turn regulated, at least in part, by various post translational modifications (PTMs). Indeed K8 is reported as a highly post-translationally modified protein (Fig 1). The majority of PTMs occur in the globular N- and C-terminal domains, although several occur in the coiled-coil domain (Fig 1). The PTMs of keratins include: phosphorylation, glycosylation and transglutamination which occur on either C or N-termini. These modifications have many suggested implications on the filament assembly and solubility (19). K8 phosphorylation has been shown to regulate filament formation, cell signalling, and protein-protein binding (20) in response to cellular stress, mitosis, and apoptosis (21, 22). We have reported K8 and K18 as highly acetylated (10) and associated alterations in acetylation with depolymerisation (23). Further acetylations have been reported for K8 (24), and K8 now features 11 empirically determined acetylation sites. We have shown that butyrate causes an K8 acetylation using a high-content biology approach (23).

High-content biology (HCB, also known as High-Content Analysis (HCA), sometimes termed cellomics) is a high-throughput microscopy approach where multiple endpoints may be analysed in both a qualitative and quantitative way integrating the functionalities of both microscopy and flow cytometry (25). HCB is increasingly used in the pharmacology sector for analysis of multiple endpoints in screening cellular responses to compound libraries (26).

Although cellular responses to butyrate have been widely studied, the response to other SCFAs which also occur in the colon lumen at pharmacologically relevant levels is less explored. The HCB approach to analysis of keratin acetylation previously reported (23) was here applied to the analysis of response to other SCFAs.

## 2. Material and Methods



**Figure 1.** Different isoforms of K8 are acetylated at different positions. Panel-A shows the five positions identified to be acetylated. Two lysine residues were identified to be within the coiled-coil domain (lys100, lys392) and three positions (lys10, lys471, lys482) occur outwith the coiled-coil domains. Panel-B shows the immunoblotting for immunoprecipitation of fraction 4 of cytoskeletal isolation from three cell lines i.e., HCT-116, Caco2, HT-29 [either untreated (C), or treated with butyrate (B) 5mM for 17 hours]. After separation of proteins gels were immunoprobed for K8 and three antibodies against lysine residues at positions 10, 471, and 482. To identify which fraction contains maximum K8 and acetylated lysines HCT-116 cells were treated with 5mM for 17 hours and then cytoskeletal isolation was performed for both treated and untreated cells. All four fractions were immunoprobed with K8 and three antiacetylate antibodies (panel C).

### 2.1 Cell Culture

Cells (HCT-116, HT-29, and Caco2) were cultured in DMEM (Dulbecco/Vogt-modified Minimal Essential Medium) containing 1 g/L D-glutamine, 4 mM L-glutamine, 110mg/L sodium pyruvate and 25 mM HEPES (all supplied by Gibco). Added to the DMEM was penicillin/streptomycin (Gibco) and 10% FCS (Biosera, Sussex, UK). Cells were incubated at 37 °C, 5% CO<sub>2</sub> in humidified air.

### 2.2 Cytoskeletal isolation

Protocol for intermediate filament cytoskeletal isolation was as described by Herman *et al.* (27). For cytoskeletal isolation from cell lines, either treated with butyrate (5 mM, 17 hours) or otherwise, adherent cells were washed in warm PBS (Oxoid, UK) containing 2 mM MgCl<sub>2</sub> and 50mM pe-fabloc (Roche, Welwyn garden City, UK). Low Detergent Buffer (0.5X PBS; 50 mM MOPS pH 7; 10 μM MgCl<sub>2</sub>; 2μM

pefabloc; 1 mM EGTA; 0.15% v/v Triton x-100) was then added. After 90 seconds buffer was removed and cells were centrifuged and supernatant was removed which is membrane bound/soluble protein fraction (Fraction 1). Cold extraction was performed with High detergent Buffer (0.5X PBS; 50 mM MOPS pH 7; 10  $\mu$ M MgCl<sub>2</sub>; 2  $\mu$ M Pefabloc; 1% v/v Triton X-100, 1 M NaCl) which yielded the cytoskeleton (Fraction 2). Remaining sample was centrifuged to yield the high salt soluble fraction (Fraction 3) as supernatant and the remaining pellet was the cytoskeletal proper (Fraction 4). Samples were analysed by immunoblotting. Primary antibodies used were anti K8 (mouse monoclonal Abcam, 9023) and in-house antibodies raised against acetylated lysine residues of K8: lys10 (rabbit), lys482 (rabbit), and lys471 (chicken) (for validations of antibodies see supporting information in Drake *et al* (23)). After SDS-Page and Western transfer [standard method as described previously (28)] the cross-reaction was visualized using HRP-conjugated secondary antibodies (Dako rabbit for lys10 & lys482, Dako chicken for lys471, and Dako mouse for K8) and Western Lightening Chemiluminescence reagent plus (PerkinElmer, Boston, USA). Imaging was undertaken with Chemigenius Bio-Imaging system (Syngene).

### 2.3 Immunocytochemistry

HCT-116 cells were seeded in 96-well plate at  $3 \times 10^3$  per well and then grown for 48 hours. Cells were then treated with 0-20 mM of SCFAs (butyric acid, propionic acid, valeric acid, or valproic acid; Sigma). Internal triplicates were used in each experimental pass. After 24 hours of treatment cells were fixed at -20 °C with 100% methanol for 5 minutes before immunocytochemical (ICC) staining. 100  $\mu$ l of PBS was left in each well at the end of ICC staining. Plates were then sealed and sent for HCA.

During ICC primary antibodies used were same as in Western Blot (see section 2.2). Cells were dual immunostained for K8 and one of the three antibodies to normalize the acetylation intensity against the total K8 for each field while quantifying specific acetylation. All antibodies were diluted in digitonin according to Cellomics' HCA protocol. Secondary antibodies used were anti-mouse Alexa fluor 555-red for K8, anti-rabbit Alexa fluor 444-green for lys10 and lys482. Nuclei were stained with Hoechst 33342.

During HCA analysis, images were captured on an Arrayscan II (Cellomics) and analysed using the proprietary Arrayscan compartmental analysis algorithm. Regions of interest (ROIs), which corresponded to cells, were obtained based on the Hoechst nuclear staining profile. A mask was then created so that only cytoplasmic staining was measured, thereby excluding any non-specific nuclear fluorescent signal (23). The analysis was corrected for background fluorescence by using the program's spot thresholding algorithm. The data generated measured the total staining intensity of the identified spots for K8, acetylated lys10, acetylated lys482, and acetylated lys471. A single value for each well was ob-

tained by averaging the total staining intensities across the ROI population for the well.

In a different experiment, to compare acetylation in dividing and non-dividing cells, Caco2 cells were grown for 48 hours and then treated with various butyrate concentrations for 24 hours. Same primary and secondary antibodies were used as described above and same ICC protocol was followed. All control experiments were analyzed for increase in acetylated lysines in mitotic cells by using HCA.

### 2.4 Statistical analysis

GraphPad Prism5 was used to calculate p values by using paired T test. Graphs were generated through GraphPadPrism5.

## 3. Results

### 3.1 Keratin 8 is identified to have two isoforms which are differentially acetylated

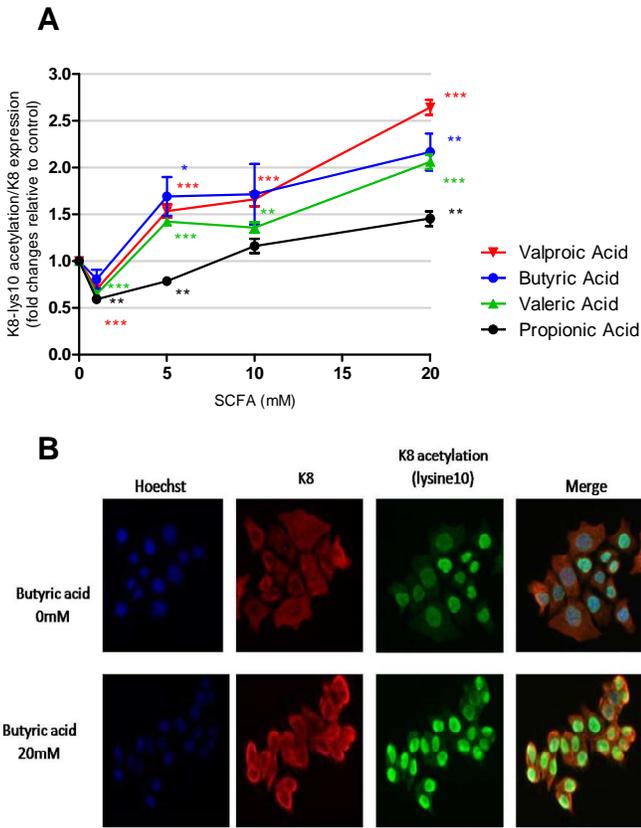
We have previously identified the five lysine residues which are acetylated in K8 (10) and further acetylation sites have been identified in a subsequent high-throughput study (24). Figure 1A shows the positions of these lysine residues relative to coiled-coil domains of K8. Three acetylation sites (lys10, lys471 & lys482) were found to be outside the coiled-coil domain at the N- and C-termini. Polyclonal antibodies have previously been described to these modified sites (23).

We first sought to establish whether K8 acetylation occurred in multiple colon cell lines. HCT-116, Caco2, and HT-29 were treated with butyrate (5 mM for 17 hours) and then insoluble fractions of treated and untreated cells were immunoprobed with antibodies against K8 and acetylated lysines (lys10, lys482, lys471) (Fig 1B). The panel indicates that lys10 and lys482 showed reactivity in a band corresponding to that of K8. Also, it is evident that K8 acetylation at lys482 was more marked in the higher molecular weight K8 band and lys10 acetylation was present in both forms of K8.

In order to determine in which fraction keratin 8 was most highly acetylated, fractions of HCT-116 cells were immunoprobed with K8, lys10, lys471, and lys482 (Fig 1C). K8 is most abundant in the insoluble fraction (fraction 4). The lower panels further demonstrate that most reactivity with anti-acetyl antibodies is seen in fraction 4 which demonstrates that K8 is acetylated when in the insoluble intermediate filament. Another interesting observation was that K8 appeared to exist as a doublet.

### 3.2 SCFAs increase K8 acetylation in vitro

Our previous work (10) suggested that butyrate can increase acetylation of K8. To demonstrate the effect of butyric acid and other SCFAs on K8 acetylation, HCT-116 colon cancer cells were treated with different SCFAs at five different concentrations and then the changes in K8 acetylation



**Figure 2.** Changes in K8-lys10 acetylation in response to SCFAs. HCT-116 cells were grown for 48 hours and then treated for 24 hours with increasing doses of four different SCFAs (butyric acid, propionic acid, valeric acid, valproic acid). After 24 hours cells were fixed with cold methanol. For immunocytochemistry cells were treated with antibody against K8 (Abcam 9023) and antiacetyl antibody for acetylated lysine residue at position 10 (lys10). Binding of K8 antibody was detected with a secondary antibody (Alexa flour 555-red) and binding of lys10 was detected with another secondary antibody (Alexa Flour 444-green). Cell nuclei were stained with Hoechst. High Content Analysis (HCA) was performed on a Cellomics Arrayscan. Nine fields were scored for each concentration of individual SCFA treatment. Total K8 was quantified for each field along with the intensity of Alexa-444 staining. Lys10 acetylation was normalized against total K8 expression for each field and then data was normalized to the control (no treatment). Figure2-A shows the fold changes in lys10 acetylation for each dose of SCFA compared to control. Data represents the mean  $\pm$  SEM for three experiments with three internal repeats in each experiment. P values were calculated for lys10 acetylation at each SCFA concentration. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$  relative to control (Paired t-test with GraphPad Prism5). Figure 2B shows the cells with or without treatment of butyric acid. Nuclei were localized with Hoechst (blue). Cytoplasmic staining of K8 (red) was colocalize with the anti-acetyl antibody (green) as shown by the yellowish merge in the last picture on the right.

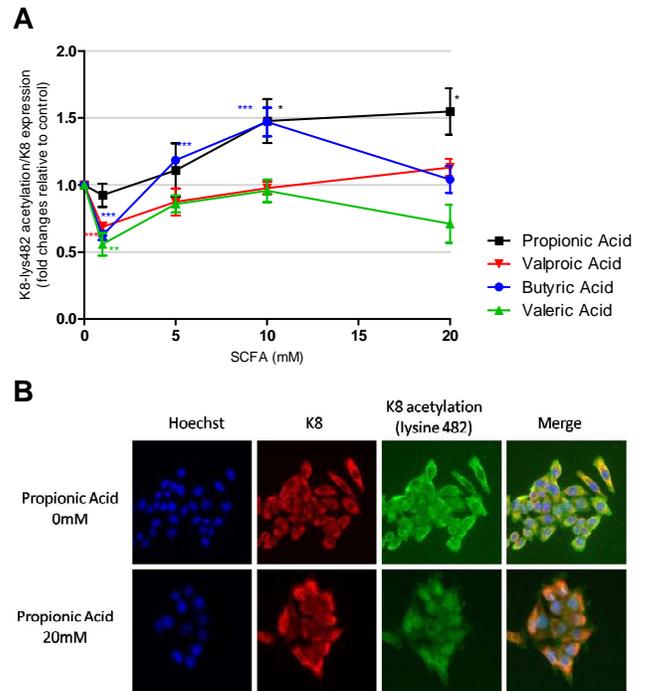
were analysed by HCA. HCA was used to analyze the results which has already been proved to be an effective and quick way of quantify total K8 and K8 acetylation (23).

HCT-116 cells treatment with 1 mM of all SCFAs showed a significant decrease ( $p < 0.01$  for propionic acid and  $p < 0.001$  for valeric and valproic acid) in acetylation at lys10, except for butyric acid (Figure 2A). Exposure of HCT-116 cells to 5

mM of each SCFA caused a significant increase ( $p < 0.05$ ) in lys10 acetylation with all SCFAs. There was not a marked difference of response between SCFAs treatment at 5 mM. At 10 mM of each SCFA a plateau in the concentration response was reached. Maximum response was observed at 20 mM treatment for all SCFAs and valproic acid caused the largest increase ( $p < 0.001$ ) in acetylation at 20 mM among all SCFAs (Figure 2A).

The response in lys482 acetylation with SCFAs was not as profound as for lys10 acetylation (Figure 3A). At 1 mM concentration all SCFAs caused a significant decrease ( $p < 0.05$ ) in K482 acetylation except propionic acid. Valproic acid and valeric acid did not increase acetylation at any concentration when compared with control (0 mM). Propionic acid and butyric acid treated cells at 10 mM showed a significant increase in lys482 acetylation ( $p < 0.01$  for propionic acid and  $< 0.001$  for butyric acid). At 20 mM exposure, cells treated with propionic acid showed a significant increase ( $p < 0.05$ ) in lys482 acetylation when compared to the control but response with butyric acid was not significant at 20 mM.

Similar to response with lys10 & 482, lys471 also showed a significant decrease in acetylation with valproic acid



**Figure 3.** Changes in K8-lys482 acetylation in response to SCFAs. Figure3-A shows the fold changes in lys482 acetylation for each dose of SCFA compared to control. Data represents the mean  $\pm$  SEM for three experiments with three internal repeats in each experiment. P values were calculated for lys10 acetylation at each SCFA concentration. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$  relative to control (Paired t-test with GraphPad Prism5). Figure 3B shows the cells with or without treatment of propionic acid. Nuclei were localized with Hoechst (blue). Cytoplasmic staining of K8 (red) was colocalize with the anti-acetyl antibody (green) as shown by the yellowish merge in the last picture on the right.

( $p < 0.05$ ) and propionic acid ( $p < 0.01$ ) at 1 mM treatment but butyric acid and valeric acid increased lys471 acetylation at 1 mM (not statistically significant) (Figure 4). Propionic acid and valproic acid caused a decrease in lys471 acetylation at all concentrations except at 20 mM where valproic acid increased acetylation compared to control, although this was not statistically significant. Butyric acid caused a significant increase ( $p < 0.01$ ) in lys471 acetylation at 20 mM treatment.

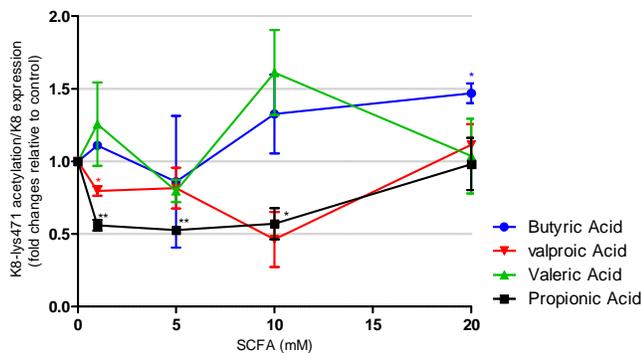
Figure 5 shows changes in acetylation at lys10, lys482, and lys470 in response to each SCFA. It can be observed that following SCFA treatment. Lys10 acetylation consistently preceded the lys482 acetylation in butyric-, valeric- and valproic-acid treated cells while lys482 acetylation only preceded the lys10 acetylation in propionic acid-treated cells. This may suggest that K8 acetylation is generally hierarchical at the three but is not inconsistent with different HDACs (with differing inhibition profiles by the SCFA) are involved in governing different acetylation sites.

### 3.3 Keratin 8 lys10 cell acetylation increases during mitosis

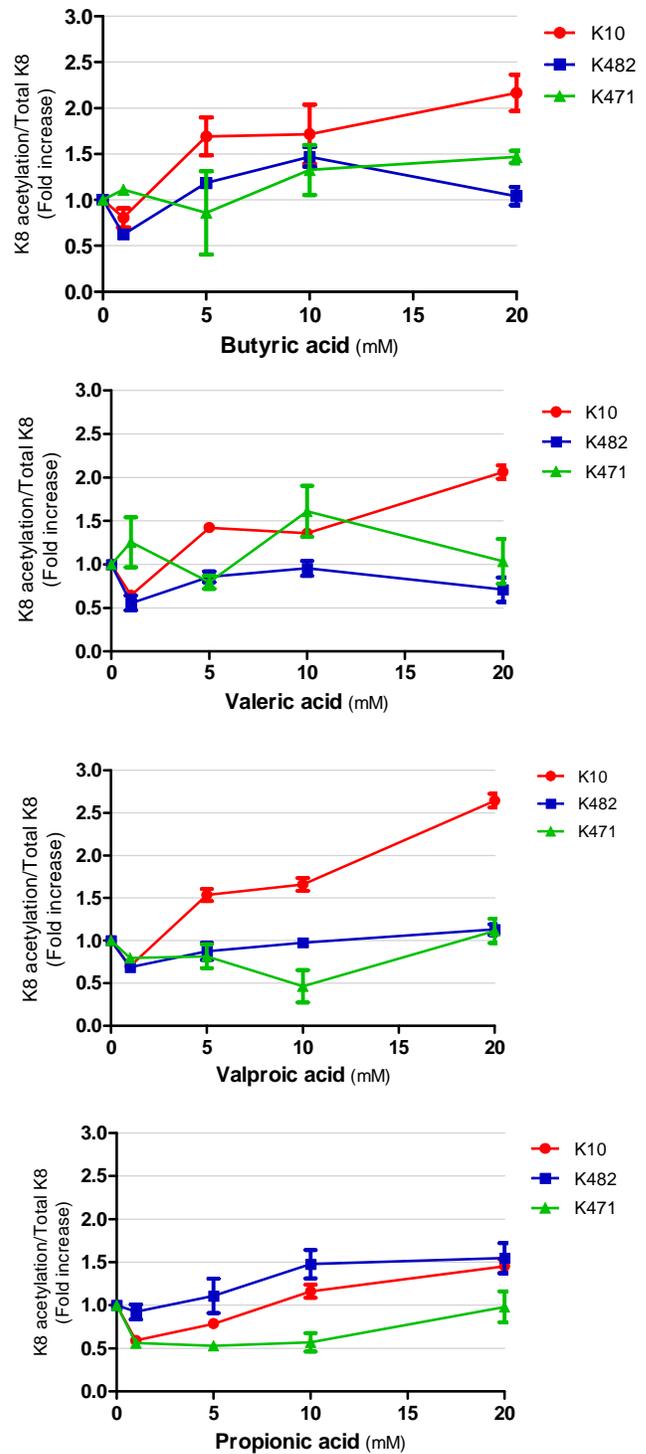
Mitosis is an important step during cell division and also various factors could affect this process. We investigated whether there is a differential change in K8 acetylation during mitosis and compared the acetylation in both dividing and non-dividing cells. Interestingly, microscopy revealed that there was an increase in K8 acetylation in mitotic cells (Fig 6A). To study this further K8 acetylation was analysed by HCA, by gating cells with the greatest intensity of nuclear staining (Fig 6B). Mitotic cells were found to have a fivefold increase in lys10 acetylation by comparison with non-mitotic cells, whereas there was a more modest increase at lys482 and lys471 (Fig 6C).

## 4. Discussion

K8 has been shown to have many PTMs which could influ-

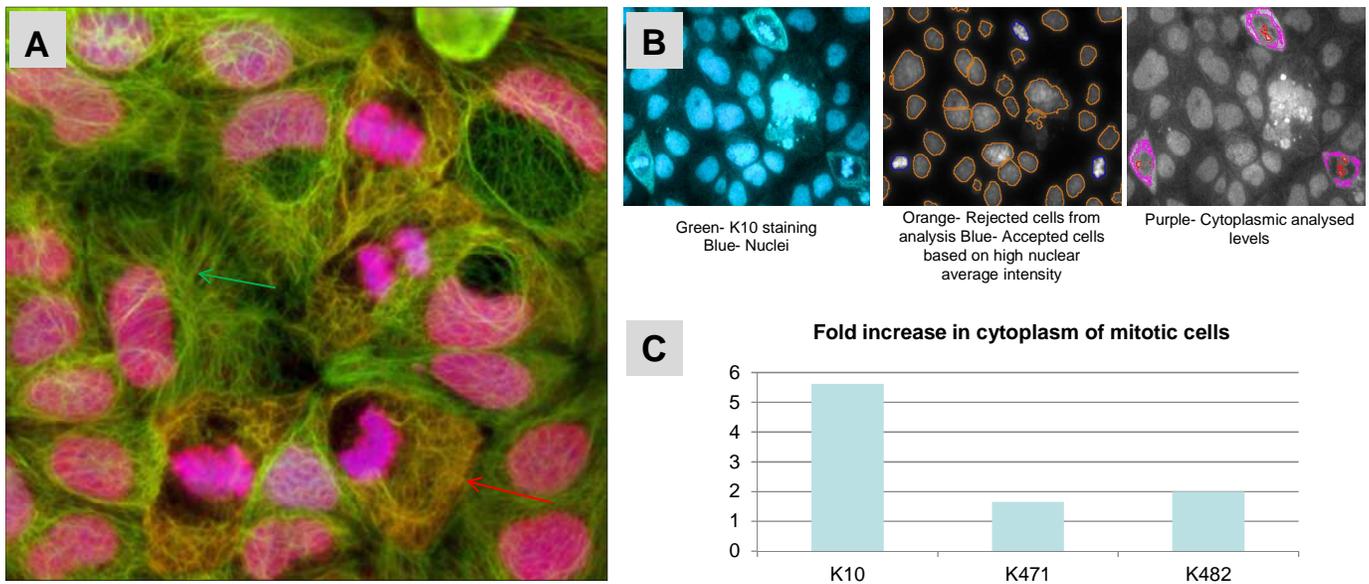


**Figure 4.** Changes in K8-lys471 acetylation in response to SCFAs. The fold changes in lys471 acetylation for each dose of SCFA compared to control. Data represents the mean  $\pm$  SEM for one experiment with three internal repeats. P values were calculated for lys471 acetylation at each SCFA concentration. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$  relative to control (Paired t-test with GraphPad Prism5).



**Figure 5.** Differential response of K8 acetylation to short chain fatty acids.

ence its structure and function. We previously showed that butyrate can increase K8 acetylation. In the present study we demonstrate the different sites of acetylation of K8 are modified differentially in response to treatment with different SCFAs, and to mitosis. We demonstrated that HCB is a useful tool to study K8 acetylation.



**Figure 6.** Keratin 8 is more acetylated in mitotic cells. Caco2 cells were grown for 48 hours and then treated with butyrate for 24 hours before fixing in ice cold methanol. Immunocytochemical staining for K8 and anti lysine antibodies was undertaken. Top picture (Fig 6A) shows that the filaments (K8) are greenish in non dividing cells (as shown with a green arrow) but the cell with a dividing nucleus (mitotic cell) shows the more yellowish (as indicated by a red arrow) fibrous network (K8) which indicated more acetylated K8 in mitotic cells. Figure 6B shows the scheme how acetylation was measured only in mitotic cells by rejecting the non-dividing cells. All control experiments were analysed for increase in acetylated lysines in mitotic cells (Figure 6C).

When the insoluble fractions of various colon cancer cell lines were immunoblotted we identified two isoforms of K8 which were differentially acetylated. This suggests that various isoforms are differentially modified. As lys10 acetylation occurs on the shorter form, and lys482 appears strongly acetylated in the short form and weakly acetylated in the longer form there is no reason to invoke loss of either end of the protein in production of these isoforms, which may potentially arise as a consequence of alterations in mass of the protein due to different burdens of post-translational modification.

Among various HDAC inhibitors, valproic acid has some attractive characteristics which make it suitable for a potential candidate to treat CRC. It has already been used for long time as an anticonvulsant and therefore its pharmacokinetic properties have been tested. It is available orally, well tolerated and has a longer *in vivo* half life compared with other HDAC inhibitors (29, 30). We showed, in this experiment, that all four studied SCFAs at 20 mM treatment (butyric acid, propionic acid, valeric acid and valproic acid) increase K8 acetylation at lys10. These observations again reinforce the theory that butyrate and other SCFAs can change the cell homeostasis through increase in K8 acetylation. Although the consequences of K8 acetylation are beyond the scope of this particular study but as mentioned in our previous report this acetylation could ultimately affect the cell apoptosis or could affect the polymerisation of cytoskeleton.

In conclusion, the present work provides evidence that acetylation is another PTM for K8 and that there are, at least, two isoforms of K8 which are differentially acetylated. SCFAs increase K8 acetylation and maximum lys10 acetyla-

tion is induced with valproic acid and for lys482 propionic acid produced maximal acetylation.

The results were achieved using a high content biology approach. Through application of this methodology, it was possible to measure the total intensity of K8 expression and also changes in K8 acetylation in order to allow us to find the relative increase in K8 acetylation. The technique was unique in allowing us to investigate both quantitatively and qualitatively the spatial alterations in keratin acetylation, for example in mitotic cells, representing a significant advantage over either microscopy or flow cytometry. This technique can successfully be applied to study acetylation response with various HDAC inhibitors. Future studies should address the relationship between K8 acetylation and structure and function.

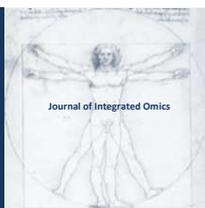
#### Acknowledgements

This was supported by a University of Sheffield Knowledge Transfer Award to BMC and BMI Thornbury Hospital.

#### References

1. Velazquez OC, Lederer HM, Rombeau JL. Butyrate and the colonocyte. Production, absorption, metabolism, and therapeutic implications. *Adv Exp Med Biol* 1997;427:123-34.
2. Chirakkal H, Leech SH, Brookes KE, Prais AL, Waby JS, Corfe BM. Upregulation of BAK by butyrate in the colon is associated with increased Sp3 binding. *Oncogene* 2006 Nov 16;25(54):7192-200.

3. Comalada M, Bailon E, de Haro O, Lara-Villoslada F, Xaus J, Zarzuelo A, et al. The effects of short-chain fatty acids on colon epithelial proliferation and survival depend on the cellular phenotype. *J Cancer Res Clin Oncol*2006 Aug;132(8):487-97.
4. Hodin RA, Meng S, Archer S, Tang R. Cellular growth state differentially regulates enterocyte gene expression in butyrate-treated HT-29 cells. *Cell Growth Differ*1996 May;7(5):647-53.
5. Gibson PR, Rosella O, Wilson AJ, Mariadason JM, Rickard K, Byron K, et al. Colonic epithelial cell activation and the paradoxical effects of butyrate. *Carcinogenesis*1999 Apr;20(4):539-44.
6. Daly K, Shirazi-Beechey SP. Microarray analysis of butyrate regulated genes in colonic epithelial cells. *DNA Cell Biol*2006 Jan;25(1):49-62.
7. Waby JS, Bingle CD, Corfe BM. Post-translational control of sp-family transcription factors. *Curr Genomics*2008 Aug;9(5):301-11.
8. Sgarra R, Rustighi A, Tessari MA, Di Bernardo J, Altamura S, Fusco A, et al. Nuclear phosphoproteins HMGA and their relationship with chromatin structure and cancer. *FEBS Lett*2004 Sep 10;574(1-3):1-8.
9. Luo J, Su F, Chen D, Shiloh A, Gu W. Deacetylation of p53 modulates its effect on cell growth and apoptosis. *Nature*2000 Nov 16;408(6810):377-81.
10. Leech SH, Evans CA, Shaw L, Wong CH, Connolly J, Griffiths JR, et al. Proteomic analyses of intermediate filaments reveals cytokeratin8 is highly acetylated--implications for colorectal epithelial homeostasis. *Proteomics*2008 Jan;8(2):279-88.
11. Moll R, Franke WW, Schiller DL, Geiger B, Krepler R. The catalog of human cytokeratins: patterns of expression in normal epithelia, tumors and cultured cells. *Cell*1982 Nov;31(1):11-24.
12. Magin TM, Vijayaraj P, Leube RE. Structural and regulatory functions of keratins. *Exp Cell Res*2007 Jun 10;313(10):2021-32.
13. Baribault H, Penner J, Iozzo RV, Wilson-Heiner M. Colorectal hyperplasia and inflammation in keratin 8-deficient FVB/N mice. *Genes Dev*1994 Dec 15;8(24):2964-73.
14. Toivola DM, Krishnan S, Binder HJ, Singh SK, Omary MB. Keratins modulate colonocyte electrolyte transport via protein mistargeting. *J Cell Biol*2004 Mar 15;164(6):911-21.
15. Owens DW, Wilson NJ, Hill AJ, Rugg EL, Porter RM, Hutcheson AM, et al. Human keratin 8 mutations that disturb filament assembly observed in inflammatory bowel disease patients. *J Cell Sci*2004 Apr 15;117(Pt 10):1989-99.
16. Chu PG, Weiss LM. Keratin expression in human tissues and neoplasms. *Histopathology*2002 May;40(5):403-39.
17. Polley AC, Mulholland F, Pin C, Williams EA, Bradburn DM, Mills SJ, et al. Proteomic analysis reveals field-wide changes in protein expression in the morphologically normal mucosa of patients with colorectal neoplasia. *Cancer Res*2006 Jul 1;66(13):6553-62.
18. Knosel T, Emde V, Schluns K, Schlag PM, Dietel M, Petersen I. Cytokeratin profiles identify diagnostic signatures in colorectal cancer using multiplex analysis of tissue microarrays. *Cell Oncol*2006;28(4):167-75.
19. Barak V, Goike H, Panaretakis KW, Einarsson R. Clinical utility of cytokeratins as tumor markers. *Clin Biochem*2004 Jul;37(7):529-40.
20. Ku NO, Zhou X, Toivola DM, Omary MB. The cytoskeleton of digestive epithelia in health and disease. *Am J Physiol*1999 Dec;277(6 Pt 1):G1108-37.
21. He T, Stepulak A, Holmstrom TH, Omary MB, Eriksson JE. The intermediate filament protein keratin 8 is a novel cytoplasmic substrate for c-Jun N-terminal kinase. *J Biol Chem*2002 Mar 29;277(13):10767-74.
22. Omary MB, Ku NO, Liao J, Price D. Keratin modifications and solubility properties in epithelial cells and in vitro. *Subcell Biochem*1998;31:105-40.
23. Drake PJ, Griffiths GJ, Shaw L, Benson RP, Corfe BM. Application of high-content analysis to the study of post-translational modifications of the cytoskeleton. *J Proteome Res*2009 Jan;8(1):28-34.
24. Choudhary C, Kumar C, Gnad F, Nielsen ML, Rehman M, Walther TC, et al. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science*2009 Aug 14;325(5942):834-40.
25. Zanella F, Lorens JB, Link W. High content screening: seeing is believing. *Trends Biotechnol*2010 May;28(5):237-45.
26. Zock JM. Applications of high content screening in life science research. *Comb Chem High Throughput Screen*2009 Nov;12(9):870-76.
27. Herrmann H, Kreplak L, Aebi U. Isolation, characterization, and in vitro assembly of intermediate filaments. *Methods Cell Biol*2004;78:3-24.
28. Griffiths GJ, Dubrez L, Morgan CP, Jones NA, Whitehouse J, Corfe BM, et al. Cell damage-induced conformational changes of the pro-apoptotic protein Bak in vivo precede the onset of apoptosis. *J Cell Biol*1999 Mar 8;144(5):903-14.
29. Blaheta RA, Cinatl J, Jr. Anti-tumor mechanisms of valproate: a novel role for an old drug. *Med Res Rev*2002 Sep;22(5):492-511.
30. Blaheta RA, Michaelis M, Driever PH, Cinatl J, Jr. Evolving anticancer drug valproic acid: insights into the mechanism and clinical studies. *Med Res Rev*2005 Jul;25(4):383-97.



ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v1i2.68

## Bacterial two-hybrid analysis of the *Shewanella oneidensis* MR-1 multi-component electron transfer pathway

Jimmy Borloo<sup>1</sup>, Lina De Smet<sup>1</sup>, Jozef J. Van Beeumen<sup>1</sup>, Bart Devreese<sup>1\*</sup>.

<sup>1</sup>Department of Biochemistry and Microbiology, Ghent University, Laboratory for Protein Biochemistry and Biomolecular Engineering (L-ProBE), K.L. Ledeganckstraat 35, 9000 Ghent, Belgium.

Received: 05 December 2010 Accepted: 02 July 2011 Available Online: 25 July 2011

### ABSTRACT

Understanding cellular systems requires profound analysis of the protein interaction networks. Protein interaction mapping is performed mainly by co-purification strategies or two-hybrid systems. Recently, we expanded the tools for analyzing protein-protein interactions in bacteria with a two-hybrid technique based on beta-galactosidase complementation, and demonstrated its potential to explore interactions of membrane systems and to study transient interactions between redox partners (Borloo et al. 2007a). We demonstrate here the functionality of this assay to reveal the interactions within the complex electron transfer chain of the dissimilatory metal reducing bacterium *Shewanella oneidensis* MR-1. Specifically, we identified the cytoplasmic membrane-bound CymA as a key component, after which the electron transport chain is found to bifurcate towards several periplasmic proteins. It again congregates at the crucial MtrA, which in turn forms an electron transfer complex with the outer membrane localized MtrB and the terminal ferric reductases MtrC and OmcA. These data are complemented by mutant screening and confirm previous kinetic analyses (Borloo et al., 2007b).

**Keywords:** *Shewanella oneidensis*; bacterial two-hybrid; protein interactions; beta-galactosidase.

### Abbreviations

NTA nitrilo triacetic acid; FR fumarate reductase; IPTG isopropyl- $\beta$ ,d-thiogalactopyranoside; ONPG o-nitrophenyl- $\beta$ ,d-galactose; NIF non-interacting fusion proteins; IF interacting fusion proteins.

### 1. Introduction

Several bacterial strains have been found capable of reducing heavy metals and to link this process to energy generation in order to grow and survive [1,2]. One of the most potent and striking examples of such strains is the Gram-negative facultative anaerobe *Shewanella oneidensis* MR-1 [3]. During anaerobic metal respiration, *S. oneidensis* MR-1 applies a sophisticated multi-component cascade, often referred to as the 'metal respiratory system' or the 'Mtr respiratory pathway'. This cascade transfers electrons from the cytoplasmic electron pool over the cytoplasmic membrane and periplasm, ultimately to the outer membrane localized terminal metal reductase(s). Many proteins, most of which are multi-haem cytochromes c, have been identified and characterized in the past as playing important roles in dissimilatory metal reduction [4,5]. Among these proteins is

the centrally positioned cytoplasmic membrane-attached tetrahaem cytochrome c CymA [5-7] that collects excess electrons from carbon sources and shuttles them to MtrA, a decahaem periplasmic c-type cytochrome [8]. MtrA facilitates electron transfer to the outer membrane decahaem cytochromes c MtrC and OmcA [9], which are dependent on the non-haem integral outer-membrane protein MtrB for proper localization in the outer membrane [10]. MtrC and OmcA in turn reduce the terminal electron acceptors, which range from organic flavins to iron or manganese oxide, onto chelated iron complexes.

Despite the fact that the Mtr respiratory pathway has already been quite well-documented, some prominent questions remain. Curiously, the *Shewanella oneidensis* MR-1 genome encodes over 40 c-type cytochromes. It includes

\*Corresponding author: Dr. Bart Devreese, Laboratory for Protein Biochemistry and Biomolecular Engineering, Ghent University, K.L. Ledeganckstraat 35, B-9000 Ghent, Belgium. Phone : +32 9 264 52 73. Fax : +32 9 264 53 38. E-mail Address : bart.devreese@ugent.be

a number of paralogues, in particular and most interestingly to this study, MtrC (MtrF), MtrA (MtrD) and MtrB (MtrE). It has been shown that the MtrC-OmcA pair is largely responsible for chelated and insoluble iron reduction [11-13], but that MtrF is modular with MtrC and even more active than OmcA during flavin reduction in the absence of MtrC [14]. Additionally, Coursolle and Grainick [14] proved that, upon deletion of MtrA, alternative electron flow routes are recruited and that primarily MtrD, and to a lesser extent DmsE (the periplasmic component of the DMSO respiratory system), at least in part take over its function. Furthermore, involvement of the periplasmic iron-induced flavocytochrome IfcA [15], the small tetrahaem cytochrome c (CctA) [16,17], fumarate reductase [18] and the MtrA paralogue MtrD have been shown to be implicated in the Mtr respiratory pathway, indicating that also periplasmic electron shuttling is modular. Even though the key players in anaerobic dissimilatory metal reduction have already been identified, the exact organization of the electron transport chain, that is, which protein interacts with which protein and in what order and stoichiometry, has not yet been elucidated.

In this work we provide crucial insights in the organization of the chelated iron respiration cascade by: i) assaying several well-selected *S. oneidensis* MR-1 insertion mutants (i.e. which have been shown to be part of the electron transport chain) for their Fe(III)-NTA reduction capacity, thereby verifying and evaluating their importance or redundancy in the involved electron transport chain, and comparing them to other mutant studies, and ii) applying a recently developed  $\beta$ -galactosidase based bacterial two-hybrid system [19] in order to assess the interaction behavior of all previously identified components of the electron transport chain. Finally, we propose a comprehensive model that represents the *S. oneidensis* MR-1 Mtr respiratory pathway used during chelated metal reduction.

## 2. Material and Methods

### 2.1 Bacterial Strains, Plasmids, Media and Growth Conditions

*S. oneidensis* MR-1 was originally isolated from Oneida Lake sediments [5] and was obtained from the LMG culture collection (LMG 19005; Ghent, Belgium). The strain MR-1R, used in this study, is a spontaneous rifampicin-resistant mutant of strain MR-1 that was isolated in-house. The *S. oneidensis* MR-1 *ccmE*, *cymA*, *STC*, *ifcA*, *FR*, *mtrA*, *mtrB*, *mtrD*, *mtrE*, *mtrF*, *omcA* and *omcB* disruption mutants, as well as the *omcA/omcB*, *mtrA/mtrD*, *mtrB/mtrE*, *mtrB/mtrF* and *omcB/mtrF* double mutants, were obtained during a previous study [20]. *Escherichia coli* strain XL-1 Blue (Stratagene, La Jolla, CA, USA) was used in all subcloning steps, whereas recombination deficient strain JM109 (The Coli Genetic Stock Center, New Haven, CT, USA) was applied in all subsequent cloning experiments.  $\beta$ -Galactosidase activity tests were performed using *E. coli* strain MC1061

(The Coli Genetic Stock Center), which lacks the entire *lacZ* locus.

The pCOLADuet-1 vector (Km<sup>R</sup>) was obtained from Novagen (Darmstadt, Germany). The pB2H $\Delta\alpha$  (Cm<sup>R</sup>) and pB2H $\Delta\omega$  (Cb<sup>R</sup>) plasmids were obtained from a previous study [19].

Media and growth conditions for MR-1R and all *S. oneidensis* mutant cultures were as described previously [11]. *E. coli* cultures were grown aerobically in Luria Bertani broth on a rotary shaker (200 rpm) at 37°C [21]. Growth media were supplemented with appropriate antibiotics when necessary, including chloramphenicol (Cm) at 25  $\mu$ g/mL, carbenicillin (Cb) at 100  $\mu$ g/mL, and kanamycin (Km) at 25  $\mu$ g/mL. When required, IPTG (Duchefa, Haarlem, The Netherlands) was added to a final concentration of 20 mM.

### 2.2 DNA Manipulations

To allow cytochrome c maturation in *E. coli*, the *ccm* genes were cloned from the pEC86 vector into the pCOLADuet-1 vector by XbaI-SalI restriction digest. A list of the synthetic oligonucleotides used in this study is presented in Table 1. Oligonucleotides included the restriction sites used to clone the DNA fragments into the vectors. Restriction digests, cloning and DNA electrophoresis were performed using standard techniques. DNA ligations were performed using T4 DNA Ligase (Promega, Madison, WI, USA). Isolation of plasmid DNA was accomplished using the QIAprep Plasmid Midi Kit 100 (Qiagen, Hilden, Germany). All DNA constructs were confirmed by DNA sequencing (GENOME Express, Meylan, France).

### 2.3 Fe(III)-NTA Reduction Assay

*S. oneidensis* MR-1R and all mutant strains were assayed for Fe(III)-NTA reduction activity according to previous reports [11], with the exception that 1.5 mM Fe(III)-NTA was used. Figure 1 presents the results of triplicate experiments.

### 2.4 $\beta$ -Galactosidase-based Bacterial Two-Hybrid Assay

All steps in the application of the two-hybrid system, with the exception of protein sample preparation and the enzymatic  $\beta$ -galactosidase assay, were carried out as previously described [19].

Protein sample preparation was carried out as follows. Cells were centrifuged for 10 minutes at 10,000 x g. The pellet was resuspended in 1 mL phosphate buffered saline (PBS, pH 7.4) with Complete Protease Inhibitor (Roche Diagnostics Corporation, Indianapolis, IN, USA) (1 tablet per 200 mL sample), followed by sonication for 30 seconds using a Branson Digital Sonifier Model 250-D at 10% of its maximum force. The sample was subsequently clarified by centrifugating for 30 seconds at 10,000 x g in a tabletop centrifuge (Eppendorf, Hamburg, Germany). The obtained super-

**Table 1.** *S. oneidensis* MR-1 proteins applied in this study.

PROTEIN	DESCRIPTION	OLIGONUCLEOTIDE NAME AND SEQUENCE <sup>a</sup>
CymA	Tetrahaem cytochrome c	A1 5'-GCATGCAGGAGGACAGCTATGAACTGGCGTGCCTATTTAAACCC-3'
		A2 5'-GGATCCTCCTTTGGATAGGGGTGAGCGAC-3'
IfcA	Tetrahaem flavocytochrome c	B1 5'-GCATGCAGGAGGACAGCTATGTTGAATACCAAATTATTACCGTTA-3'
		B2 5'-GGATCCTTAATAGAATTAGCTACTTGTTC-3'
STC	Tetrahaem cytochrome c	C1 5'-GCATGCAGGAGGACAGCTATGAGCAAAAACTATTAAGTGTGCTT-3'
		C2 5'-GGATCCCTTCTTCAGAACAGACGCAGAAGT-3'
FR	Tetrahaem flavocytochrome c	D1 5'-GCATGCAGGAGGACAGCTATGTTCCACAAGAAAGATTCAAAAAACA-3'
		D2 5'-GGATCCATTATCTTTAGCGAATTTAGCGGC-3'
MtrA	Decahaem cytochrome c	E1 5'-GCGGCCGAGGAGGACAGCTATGAAGAACTGCCTAAAAATGAAAAACCTACT-3'
		E2 5'-GGATCCGCGCTGTAATAGCTTGCCAGATGG-3'
MtrB	75 kDa, outer membrane	F1 5'-GCATGCAGGAGGACAGCTATGAAATTTAAACTCAATTTGATCACT-3'
		F2 5'-GGATCCGAGTTTGTAACCTATGCTCAGCAT-3'
MtrD	Decahaem cytochrome c	G1 5'-GCATGCAGGAGGACAGCTATGCTTACATTAATGTTATCGATTCTC-3'
	40 kDa, periplasm	G2 5'-GGATCCTCTCTGCAGCAACTGCCGGATGG-3'
MtrE	77 kDa, outer membrane	H1 5'-GCGGCCGAGGAGGACAGCTATGCA AATAGTGAATATATCGACTCT-3'
		H2 5'-GGATCCCATTTGGTAGCTTAAGGTCAAACC-3'
MtrF	Decahaem cytochrome c	I1 5'-GCATGCAGGAGGACAGCTATGAATAAGTTTGCAAGCTTACCACG-3'
		I2 5'-GGATCCGTTTATTGGATGGACTTTGAGTAC-3'
OmcA	Decahaem cytochrome c	J1 5'-GCGGCCGAGGAGGACAGCTATGATGAAACGGTTCAATTTCAATACC-3'
	83 kDa, outer membrane	J2 5'-AGATCTGTTACCGTGTGCTTCCATCAATTG-3'
OmcB	Decahaem cytochrome c,	K1 5'-GCATGCAGGAGGACAGCTATGATGAACGCACAAAAATCAAAAATCGCAC-3'
		K2 5'-GGATCCCATTTTCACTTTAGTGTGATCTGC-3'

<sup>a</sup> Underlined regions indicate the added restriction endonuclease sites to facilitate cloning: *Bam*HI for A2, B2, C2, D2, E2, F2, G2, I2 and K2; *Sph*I for A1, B1, C1, D1, F1, G1, I1 and K1; *Not*I for E1, H1 and J1; *Bgl*II for J2

nantant was used in the assay.

$\beta$ -Galactosidase activity was assayed quantitatively at room temperature by following ONPG hydrolysis and 2-nitrophenol formation at 415 nm in a Bio-Rad model 680 microplate reader (Bio-Rad, Hercules, CA), in a total volume of 200  $\mu$ L using  $\beta$ -Galactosidase Assay Buffer (Pierce, Rockford, IL, USA). Half of the volume was protein sample while the other half consisted of the  $\beta$ -galactosidase substrate ONPG. Spectrophotometric measurements were started immediately after mixing. All values were subsequently normalized toward the total protein content per sample. The experiment was repeated five times for every fusion protein couple.

During analyses, an internal NIF set was used consisting of the fusion protein couple AtpE- $\Delta\alpha$  and AtpB- $\Delta\omega$  [19]. The  $\beta$ -galactosidase activity value of this NIF set was  $3.49 \pm 2.62$  nmol/(min.mg) and is arbitrarily represented here as  $1.00 \pm 0.75$ . The activity levels of the analyzed fusion protein couples were subsequently expressed relative toward this NIF set (Table 2).

### 2.5 Statistical Analysis

In the application of the two-hybrid system, statistical analysis was used to attribute IF and NIF status to different sets of fusion protein couples. Since only comparisons between the mean values of small data sets were made, the student *t*-test was applied. Since we expected the IF sets to have at least an equal or larger mean value ( $\beta$ -galactosidase activity) than the NIF sets, on the one hand, combined with the fact that a larger mean value for the NIF sets compared to the IF sets can only be attributed to chance, on the other hand, it was appropriate to choose a one-tail P-value. For all two-hybrid experiments,  $n = 5$  and the alpha level ( $\alpha$ -level) was set at 0.05. All statistical analyses were carried out using GraphPad Prism Version 4.00 and on-line GraphPad statistical software (GraphPad Software, Inc., San Diego, CA, USA).

### 2.6 Chemical Cross-linking

*S. oneidensis* MR-1R was grown to mid-log phase in mini-

**Table 2.**  $\beta$ -Galactosidase based two-hybrid analysis results <sup>a</sup>

	CymA	IfcA	STC	FR	MtrA	MtrD	MtrB	MtrE	OmcA	OmcB	MtrF
CymA	+										
IfcA	+	-									
STC	-	-	-								
FR	++	++	+	++							
MtrA	++	++	++	++	++						
MtrD	+	++	-	++	++	++					
MtrB	-	-	-	-	++	+	-				
MtrE	-	-	-	-	+	-	-	-			
OmcA	-	-	-	-	++	++	++	++	-		
OmcB	-	-	-	-	+	++	++	++	++	-	
MtrF	-	-	-	-	-	++	-	+	-	-	-

<sup>a</sup> “+” or “++” was given when  $P < 0.05$  compared to the NIF set values. “+”, meaning interaction, when  $\text{NIF} \pm \text{SEM} < \text{value} \pm \text{SEM} < 10$  and “++”, meaning strong interaction, when  $10 < \text{value} \pm \text{SEM}$

mal medium [11] with 20 mM lactate and 50 mM ferric citrate as the electron donor and acceptor, respectively. Cells were centrifuged at 16,000 x g for 1 minute and then washed with an equal volume of phosphate buffered saline (PBS, pH 7.0) before suspending them in PBS to an optical density of 0.4 at 550 nm. Cells were cross-linked for 1 hour at room temperature with 1% formaldehyde. Following cross-linking, the total membrane fraction was prepared for both cross-linked and non-cross-linked cells, as described previously [19]. Equal amounts of the resulting samples were loaded on a 12% SDS-PAGE gel for analysis. After running the gel, it was stained specific for haem containing proteins as described by Thomas *et al.* [22]. The protein content of the resulting bands was subsequently analyzed by MALDI-TOF mass spectrometry.

### 2.7 SDS-PAGE, Western Blotting and Haem Staining Analyses

Protein content was verified by denaturing protein gel electrophoresis (SDS-PAGE) in 12% gels according to the method of Laemmli [24]. Western blotting using an antibody specific for  $\beta$ -galactosidase from *E. coli* (Invitrogen; Carlsbad, CA, USA) was performed by blotting the proteins to a Hybond ECL membrane for 4 hours at 40 V followed by an overnight non-fat dry milk blocking step at 4°C, extensive washing with PBS + 0.1% Tween20, and incubating the membrane with the anti- $\beta$ -galactosidase antibody for 1 hour at room temperature. The washing steps were repeated and the membrane was subsequently incubated with horse radish peroxidase (HRP) coupled to an antibody raised against the anti- $\beta$ -galactosidase antibody. HRP chemiluminescence was then used to detect the fusion proteins containing the  $\beta$ -galactosidase fragments. Staining specific for haem containing proteins was carried out as described by Thomas *et al.* [22]. For all analyses, equal amounts of total protein (approx.

20  $\mu\text{g}$ ) were consistently loaded on the gels.

### 2.8 Miscellaneous Procedures

Protein concentrations were determined by the Bradford assay [23] using the Bio-Rad Protein Assay Solution (Bio-Rad, Hercules, CA, USA). Image acquisition was carried out using software packages Corel® version 9 and Paintshop Pro version 5 (Corel, Berkshire, UK).

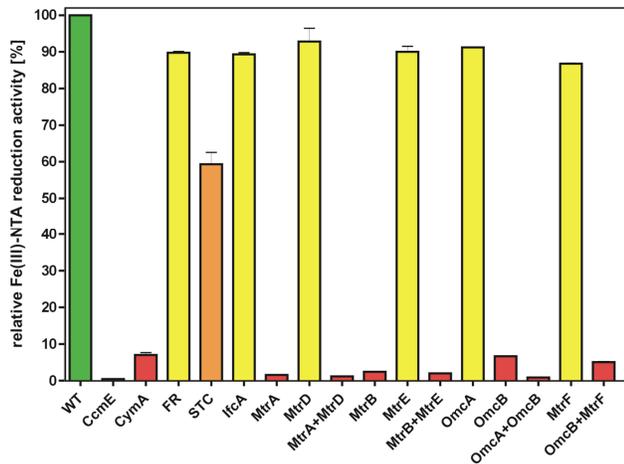
## 3. Results

### 3.1 Fe(III)-NTA Reduction Activity of the Insertion Mutants.

When comparing to *S. oneidensis* MR-1R, the *STC* insertion mutant has its Fe(III)-NTA reduction capacity being diminished by approximately 40% (Fig. 1). Disrupting the *omcA*, *FR*, *ifcA*, *mtrD* and *mtrF* genes leads to a slight drop in Fe(III)-NTA reduction activity, of approx. 10%. Whereas the *omcB* and *cymA* mutants still bear some residual reduction activity (10% of the MR-1R level at most), disrupting *mtrA* and *mtrB* almost completely abolishes Fe(III)-NTA reduction. Figure 1 also shows that when two (often modulatory) genes are simultaneously interrupted, particularly the *omcA/omcB*, *mtrA/mtrD*, *mtrB/mtrE*, *mtrB/mtrF* and *omcB/mtrF* double mutants, no significant levels of reduction activities are observed either. Importantly, a *ccmE* mutant (CcmE, SO\_0259), which is defective in cytochrome *c* maturation, also lacks Fe(III)-NTA reduction activity

### 3.2 $\beta$ -Galactosidase-based Bacterial Two-Hybrid Assay

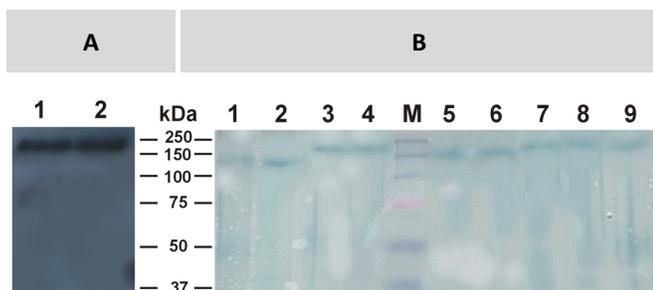
To assess their protein-protein interaction behaviour, *cymA*, *FR*, *ifcA*, *STC*, *mtrA*, *mtrB*, *mtrD*, *mtrE*, *mtrF*, *omcA* and *omcB* were all separately cloned both into the pB2H $\Delta\alpha$  and pB2H $\Delta\omega$  constructs as described previously [19]. A list of



**Figure 1.** *Fe(III)-NTA reduction assay of the mutants.* Bar representation of the Fe(III)-NTA reduction activities of the single and double insertion mutants applied in this study. Values are presented relative to the reduction capacity of *S. oneidensis* MR-1R, for which the level was set at 100%. Error bars indicate the standard error of the mean (SEM).

these proteins and their as of yet known properties is provided in Table 1.

Correct synthesis of the fusion proteins was controlled by Western blotting (immunoblot) for those containing MtrB and MtrE, and haem staining for those carrying haem-proteins. Figure 2 shows the expression results for pB2HΔα-containing fusion proteins. Haem staining reveals successful haem attachment for all haem-carrying fusion proteins after proper translocation to the periplasm (Fig. 2A). The immunodetection using anti-β-galactosidase antibodies verified that the MtrB- and MtrE-containing chimeras are properly



**Figure 2.** *Controls regarding the proper synthesis of the fusion proteins.* A. Western blot using an antibody specific for β-galactosidase, equal amounts of protein were loaded (~20 μg total protein). The molecular weight marker is indicated at the centre of the figure. Lane 1, pB2HΔαΩmtrB (193 kDa); Lane 2, pB2HΔαΩmtrE (195 kDa). B. Specific staining for haem proteins, equal amounts of protein were loaded (~20 μg total protein). The molecular weight marker is indicated at the centre of the haem-stained gel. Lane 1, pB2HΔαΩcymA (139 kDa); Lane 2, pB2HΔαΩSTC (132 kDa); Lane 3, pB2HΔαΩIfcA (183 kDa); Lane 4, pB2HΔαΩFR (181 kDa); Lane 5, pB2HΔαΩmtrA (161 kDa); Lane 6, pB2HΔαΩmtrD (158 kDa); Lane 7, pB2HΔαΩmtrF (192 kDa); Lane 8, pB2HΔαΩomcA (201 kDa); Lane 9, pB2HΔαΩomcB (195 kDa).

synthesized at low, *quasi* physiological levels when IPTG is added to the growth medium (Fig. 2B).

Our two-hybrid analysis, the results of which are presented schematically in Figure 3, and in a complete though condensed form in Table 2, reveals the interaction of the cytoplasmic membrane-anchored CymA with itself and with the periplasm-located IfcA, FR and MtrA. The latter two display a strong interaction with CymA, whereas the MtrA-paralogue MtrD surprisingly yielded no positive signal. The periplasmic pool, consisting of the multi-haem c-type cytochromes IfcA, STC, FR, MtrA and MtrD, furthermore reveal an extensive divergence of interactions, that is, besides IfcA interacting with FR, both of these proteins interact strongly with MtrA and its paralogue MtrD. In contrast, STC seems to be more limited in its interaction behaviour since it only interacts with FR and MtrA. Of the periplasmic proteins, FR and MtrA both exhibit positive interaction values with all other periplasm-directed cytochromes, and with themselves. However, given their interaction profiles with the outer membrane localized proteins MtrB, MtrE, OmcA and OmcB, apparently only MtrA, and to a lesser extent MtrD, are capable of making the link to the membranous protein pool. MtrB and its paralogue MtrE were additionally identified as to strongly interact with both OmcA and OmcB. Finally, although not found to form homo-oligomers, a convincingly positive interaction signal was also observed when the terminal ferric reductases OmcA and OmcB were co-expressed.

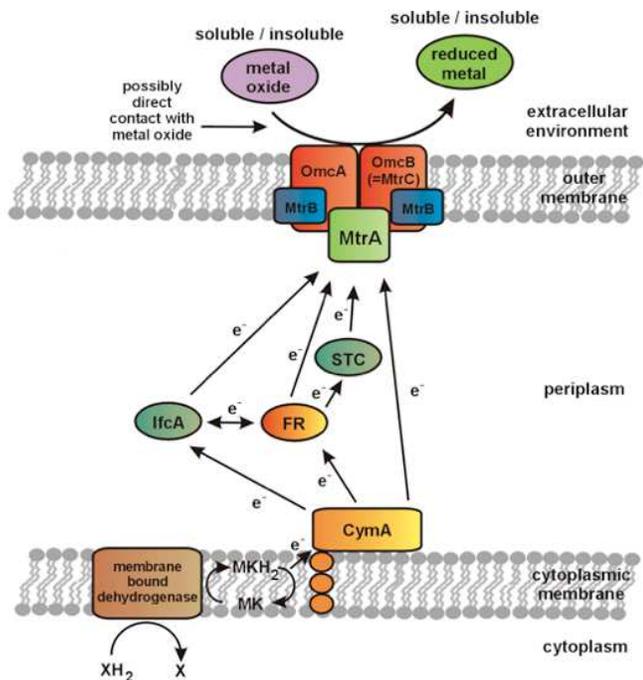
### 3.3 Chemical cross-linking.

The membrane fractions of both formaldehyde cross-linked and non-cross-linked *S. oneidensis* MR-1R cells were applied to SDS-PAGE and visualized by haem-staining. The gel showed two prominent bands of 80 and 73 kDa (Fig. 4) corresponding to OmcA and OmcB, respectively. Addition of 1% formaldehyde to undisrupted cells resulted in the appearance of a new haem band corresponding to a molecular weight of approximately 200 kDa, as shown in Figure 4. MALDI-TOF mass spectrometric analysis identified the cross-linked proteins as OmcA, OmcB and MtrB, which is in accordance with the observed molecular weight of the proteins on the haem-stained gel.

## 4. Discussion

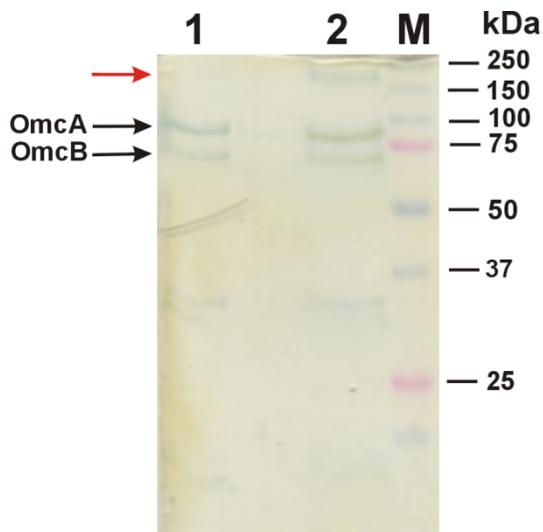
The fundamental importance of cytochromes *c* in anaerobic metal reduction has been recognized for quite some years [4, 25, 26] and is here confirmed through our experiments in that a *S. oneidensis* MR-1 mutant defective in cytochrome *c* maturation (*ccmE* mutant) does not hold any significant Fe(III)-NTA reduction activity (Fig. 1).

The central position of CymA (Fig. 3), being cytoplasmic membrane-anchored and faced to the periplasm, already suggests its pivotal role in the electron transfer process. This is again reflected by the lack of any substantial Fe(III)-NTA



**Figure 3.** Model for the metal reduction mechanism by *S. oneidensis* MR-1. The arrows indicate the possible routes of electron transfer starting from the cytoplasm, via the cytoplasmic membrane, periplasm and outer membrane, and ending at the reduction of the targeted metal species. See text for further details.

reduction capacity of the *cymA* mutant. The fact that it interacts with all periplasmic components of the electron transport chain, except with STC, is decisive in our conclu-



**Figure 4.** Heme-stain after the formaldehyde cross-linking experiment. Total membrane fractions originating from non-cross-linked *S. oneidensis* MR-1R cells (Lane 1), and formaldehyde cross-linked cells (Lane 2). Black arrows specify the positions of OmcA and OmcB, whereas the red arrow indicates the cross-linked high-molecular weight species observed only in Lane 2. The molecular weight marker is indicated on the right in units of kDa.

sion that CymA, apparently as a homo-oligomer, forms the crucial link between the menaquinone pool in the cytoplasmic membrane and the several periplasmic electron transfer proteins, which is in agreement with previous studies [14,27]. From CymA onwards, the electron transport chain further branches off towards three periplasmic electron transporters, i.e. IfcA, FR and MtrA (Fig. 3), of which the latter seems to take the leading role, given the effect of disrupting the *mtrA* gene on Fe(III)-NTA reduction (Fig. 1). The redundancy of IfcA and FR in the electron transfer cascade is reflected by i) the minor decrease in Fe(III)-NTA reduction activity that is observed when disrupting the respective genes (Fig. 1), and ii) the fact that both IfcA and FR interact with each other and, more importantly, with MtrA (and its paralogue MtrD). This indicates that after the bifurcation, the electron transport chain again congregates at some point, more particularly, at the key component MtrA, as was also proposed by Coursolle and Gralnick [14].

In *Shewanella frigidimarina* the iron-induced flavocytochrome  $c_3$  is coded by the *IfcA* gene, and although it displays fumarate reductase activity it is only induced when *S. frigidimarina* is grown on soluble Fe(III) [15]. Since IfcA might behave similarly in *S. oneidensis* MR-1, its redundancy in the electron transport chain seems plausible. Previous reports that Ifc<sub>3</sub> forms dimers in the psychrophilic *S. frigidimarina* [14] cannot be supported by our study of the *S. oneidensis* MR-1 IfcA.

Besides its previously reported [17] and in the present study confirmed interaction with CymA, the soluble fumarate reductase (FR) is one of two periplasmic proteins (the other being MtrA) that interacts with the small tetrahaem cytochrome *c* (STC), indicating yet another point where the electron transfer cascade splits up. This non-central position of STC, together with the limited effect in Fe(III)-NTA reduction observed from the *STC* *S. oneidensis* MR-1 mutant strain, makes us deduce that STC plays a secondary role in anaerobic metal respiration.

From the results given in Table 2 it is clear that MtrA, and to a lesser extent its paralogue MtrD, functions as the central electron courier in the periplasm, possibly as a homo-oligomer and as part of a greater outer membrane-associated complex, thus corroborating the observations made by Ross *et al.* [27,28]. This, together with the total loss of Fe(III)-NTA reduction activity in the *mtrA* mutant (but not in the *mtrD* mutant), is in contrast to earlier reports which stated that membrane fractions of a *mtrA* deletion mutant were still able to reduce Fe(III)-oxides [2].

Making a bridge from the periplasm to the outer membrane, we find MtrA again playing a decisive role in that it clearly interacts with the outer membrane situated  $\beta$ -barrel porin MtrB. These findings are in line with those of Ross *et al.* [28] and Hartshorne *et al.*[9], the latter of whom attribute a sheath-like role to MtrB, enabling it to transfer electrons from the periplasmic MtrA to the outer membrane-situated MtrC (= OmcB). Besides such crucial task and the fact that it serves as a component responsible for anchoring

or orienting proteins in/to the outer membrane [10], chances that MtrB and MtrE would serve as an electron transporter are inexistent, since they carry no haem or any other redox-active moieties. The presence or absence of MtrB thus indirectly affects the ability of *S. oneidensis* reducing certain electron acceptors, a statement that is corroborated by the results of our Fe(III)-NTA reduction assays with the *mtrB* mutant (Fig. 1). Surprisingly, the *mtrE* mutation merely caused a ~10% drop in Fe(III)-NTA reduction activity. Apart from MtrA and MtrB interacting, both proteins, and their paralogues MtrD and MtrE, displayed a positive interaction signal with both outer membrane localized decahaem cytochromes *c* OmcA and MtrC (= OmcB), and the MtrC paralogue MtrF. Also hinted at other studies [9,28], these results strengthen the theory that MtrA, MtrB, OmcA and MtrC might form a multi-component complex that would be responsible for metal reduction. It should be noted at this point that, using the techniques that we applied, no assumptions can be made concerning the stability of these interactions, since electron transfer only requires transient interactions (approx.  $10^{-9}$  s). Nevertheless, from our two-hybrid and cross-linking experiments we can deduce that the two outer-membrane decahaem cytochromes *c* OmcA and MtrC interact and very likely form a stable complex, as was previously also suggested by Shi *et al.* [29]. Our results support their suggestion that OmcA and MtrC interact in a 2:1 ratio.

Even though the results of our two-hybrid analyses suggest that MtrD, MtrE and MtrF play significant roles in the electron transport chain, the Fe(III)-NTA reduction assays refute this conclusion, since inactivating the respective genes brings the reduction activities down by only approximately 10% (Fig. 1). However, this seems logical, since their paralogues MtrA, MtrB and MtrC, respectively, are still present and take over their roles; a feature referred to as the 'modularity' by Coursolle and Gralnick [14]. The lack of reduction activity in the *mtrA*, *mtrB* and *omcB* mutants, in contrast, may find its origin in the low or non-expression of their respective paralogues.

## 5. Concluding Remarks

In conclusion, we have shown that our bacterial two-hybrid system is a powerful tool in protein-protein interaction studies. As the case of *S. oneidensis* MR-1 exemplifies, anaerobic metal reduction was found to require cytoplasmic membrane, periplasmic and outer membrane components, the most important of which being CymA, MtrA and OmcA, MtrC and MtrB, respectively. It has previously been shown that OmcA and OmcB both function as terminal ferric reductases, the latter (MtrC) being the principal enzyme (Fig. 1) [11]. We believe therefore that the electron transport chain ends at the OmcA/MtrC complex, accompanied by MtrA and MtrB, and that the ultimate step in metal reduction is the transfer of electrons to the targeted metal species, as is depicted in Figure 3.

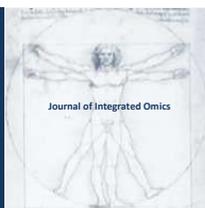
## Acknowledgements

This work was supported by a personal grant to J.B. from the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT-Vlaanderen). J.V.B. and B.D. are indebted to the Fund for Scientific Research (FWO-Vlaanderen) for granting research project G.0190.04, as well as to the Bijzonder Onderzoeksfonds of Ghent University for Concerted Research Action GOA 120154 and to BELSPO for funding IAP Grant PROFUSA. The authors are grateful to Bart Motte and Sarah De Keulenaer for technical help. We thank L. Thöny-Meyer for kindly providing us the pEC86 vector.

## References

1. D. J. Lonergan, H. L. Jenter, J. D. Coates, E. J. Phillips, T. M. Schmidt, D. R. Lovley, *J. Bacteriol.* 178 (1996) 2402-2408.
2. D. R. Lovley, E. J. Phillips, *App. Environ. Microbiol.* 54 (1988) 1472-1480.
3. K. Venkateswaran, D. P. Moser, M. E. Dollhopf, D. P. Lies, D. A. Saffarini, B. J. MacGregor, D. B. Ringelberg, D. C. White, M. Nishijima, H. Sano, J. Burghardt, E. Stackebrandt, K. H. Neilson, *Int. J. Sys. Bacteriol.* 49 (1999) 705-724.
4. A. S. Beliaev, D. A. Saffarini, J. L. McLaughlin, D. Hunnicutt, *Mol. Microbiol.* 39 (2001) 722-730.
5. C. R. Myers, J. M. Myers, *J. Bacteriol.* 179 (1997) 1143-1152.
6. J. M. Myers, C. R. Myers, *J. Bacteriol.* 182 (2000) 67-75.
7. C. Schwalb, S. K. Chapman, G. A. Reid, *Biochemistry* 42 (2003) 9491-9497. DOI: 10.1021/bi034456f
8. A. S. Beliaev, D. A. Saffarini, *J. Bacteriol.* 180 (1998) 6292-6297.
9. R. S. Hartshorne, C. L. Reardon, D. Ross, J. Nuester, T. A. Clarke, A. J. Gates, P. C. Mills, J. K. Fredrickson, J. M. Zachara, L. Shi, A. S. Beliaev, M. J. Marshall, M. Tien, S. Brantley, J. N. Butt, D. J. Richardson, *Proc. Nat. Acad. Sci. USA* 106 (2009) 22169-22174. DOI: 10.1073/pnas.0900086106
10. C. R. Myers, J. M. Myers, *Appl. Environ. Microbiol.* 68 (2002) 5585-5594.
11. J. Borloo, B. Vergauwen, L. De Smet, A. Brige, B. Motte, B. Devreese, J. Van Beeumen, *FEBS J.* 274 (2007b) 3728-3738. DOI: 10.1111/j.1742-4658.2007.05907.x
12. L. Shi, T. C. Squier, J. M. Zachara, J. K. Fredrickson, *Mol. Microbiol.* 65 (2007) 12-20. DOI: 10.1111/j.1365-2958.2007.05783.x
13. Y. Xiong, L. Shi, B. Chen, M. U. Mayer, B. H. Lower, Y. Londer, S. Bose, M. F. Hochella, J. K. Fredrickson, T. C. Squier, *J. Am. Chem. Soc.* 128 (2006) 13978-13979. DOI:10.1021/ja063526d
14. D. Coursolle, J. A. Gralnick, *Mol. Microbiol.* 4 (2010) 995-1008. DOI: 10.1111/j.1365-2958.2010.07266.x
15. P. S. Dobbin, J. N. Butt, A. K. Powell, G. A. Reid, D. J. Richardson, *Biochem. J.* 342 (1999) 439-448.
16. D. Leys, T. E. Meyer, A. S. Tsapin, K. H. Neilson, M. A. Cusanovich, J. J. Van Beeumen, *J. Biol. Chem.* 277 (2002) 35703-35711. DOI: 10.1074/jbc.M203866200
17. A. I. Tsapin, I. Vandenberghe, K. H. Neilson, J. H. Scott, T. E. Meyer, M. A. Cusanovich, E. Harada, T. Kaizu, H. Akutsu, D. Leys, J. J. Van Beeumen, *Appl. Environ. Microbiol.* 67 (2001)

- 3236-3244. DOI: 10.1128/AEM.67.7.3236-3244.2001
18. C. Schwalb, S. K. Chapman, G. A. Reid, *Biochem. Soc. Trans.* 30 (2002) 658-662. DOI: 10.1042/ bst0300658
  19. J. Borloo, L. De Smet, B. Vergauwen, J. J. Van Beeumen, B. Devreese, *J. Prot. Res.* 6 (2007a) 6, 2587-2595. DOI: 10.1021/ pr070037j
  20. A. Brige, B. Motte, J. Borloo, G. Buyschaert, B. Devreese, J. J. Van Beeumen, *Micr. Biotech.* 1 (2008) 40-52. DOI: 10.1111/ j.1751-7915.2007.00005.x
  21. L. De Smet, G. W. Pettigrew, J. J. Van Beeumen, *Eur. J. Biochem.* 268 (2001) 6559-6568.
  22. P. E. Thomas, D. Ryan, W. Levin, *Anal. Biochem.* 75 (1976) 168-176.
  23. M. M. Bradford, *Anal. Biochem.* 72 (1976) 248-254.
  24. U. K. Laemmli, *Nature* 227 (1970) 680-685.
  25. C.R. Myers, J.M. Myers, *FEMS Microbiol. Lett.* 108 (1993) 15-22.
  26. C.R. Myers CR, J.M. Myers, *J. Bacteriol.* 179 (1997) 1143-1152.
  27. D. E. Ross, J. M. Flynn, D. B. Baron, J. A. Gralnick, D. R. Bond, *PloS one* 6 (2011) e16649. DOI: 10.1371/ journal.pone.0016649.
  28. D. E. Ross, S. S. Ruebush, S. L. Brantley, R. S. Hartshorne, T. A. Clarke, D. J. Richardson, M. Tien, *Appl. Environ. Microbiol.* 73 (2007) 5797-5808. DOI:10.1128/AEM.00146-07
  29. L. Shi, B. Chen, Z. Wang, D. A. Elias, M. U. Mayer, Y. A. Gorby, S. Ni, B. H. Lower, D. W. Kennedy, D. S. Wunschel, H. M. Mottaz, M. J. Marshall, E. A. Hill, A. S. Beliaev, J. M. Zachara, J. K. Fredrickson, T. C. Squier, *J. Bacteriol.* 188 (2006) 4705-4714. DOI: 10.1128/JB.01966-05.



ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v1i2.72

## Extraction and purification of the lectin found in the tubers of *Eranthis hyemalis* (winter aconite)

Oliver George<sup>1</sup>, Claudia Solscheid<sup>1</sup>, Emilia Bertolo<sup>1</sup>, Joanna Fell<sup>1</sup>, David Lisgarten<sup>1</sup>.

<sup>1</sup>Ecology Research Group, Department of Geographical and Life Sciences, Canterbury Christ Church University, Canterbury, UK, CT1 1QU.

Received: 03 October 2011 Accepted: 17 October 2011 Available Online: 01 November 2011

### ABSTRACT

Lectin are proteins which play an important role in the defence mechanisms of plants against the attack of microorganisms and insects: this role has provoked particular interest in the fields of biotechnology and agriculture. This paper describes the extraction and purification of the lectin found in tubers of the winter aconite (*Eranthis hyemalis*), with the aim of improving and modernising the existing extraction protocol. The *Eranthis hyemalis* lectin (EHL) is a member of the type-2 Ribosome Inactivating Proteins (RIP) family, proteins which have the ability to inhibit *in vitro* protein synthesis. RIPs have been linked to plant defence by their antiviral, antifungal and insecticidal properties, and some have been found to be potent inhibitors of the Human Immunodeficiency Virus-1 (HIV-1) virus. EHL was purified using affinity column chromatography and ammonium sulphate precipitation; thiourea was used as antioxidant in order to prevent EHL denaturing during the extraction process. The presence of EHL in the extract was verified using a blood agglutination test with rabbit erythrocytes. Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) analysis was employed to determine the lectin size; EHL was found to be formed of two chains with molecular weights of approximately 31 kDa; the size of the whole protein was estimated as approximately 60 kDa. The concentration of the EHL in the post-column eluent, determined using the Bradford Assay, was 380.1  $\mu\text{g}\cdot\text{cm}^{-3}$ . This improved extraction protocol is the first step which will enable future research on the potential use of EHL in crop protection, by studying its insecticidal, fungicidal and bactericidal properties.

**Keywords:** Lectin, Ribosome Inactivating Proteins, Affinity chromatography, SDS-PAGE.

### 1. Introduction

Plant lectins are a class of proteins, non-immune in origin, which show a very specific interaction with carbohydrates. Different lectins bind with specific carbohydrate-containing compounds (e.g. polysaccharides, glycoproteins, and glycolipids), which can be free or bound in cell membranes. Another key characteristic of lectins is their ability to agglutinate erythrocytes, which provides an unambiguous indicator of their presence. Plant lectins have many functions including growth regulation, carbohydrate transport, and plant defence through interaction with microorganisms as well as insect and mammalian predators [1-3]. Additionally, lectins are involved in the detection of nitrogen-fixing bacteria on root surfaces, and the transport of hormones and

glycoproteins in plants [3, 4]. These properties make lectins particularly interesting in the context of crop protection, as biological control of pests and diseases. Lectin genes have been inserted into transgenic plants to initiate production of a desired lectin in large quantities, thereby inducing the plant resistance against insect, nematode or mammalian pests [2, 5, 6].

Historically, research on plant lectins has been concentrated in those found in the dry seeds of leguminous plant species. However, lectins have been shown to be present in a variety of plant tissues spanning a wide range of taxonomic groups [1-3]. Investigations into the structure and biochemical properties have identified and classified lectins from a

\*Corresponding author: Emilia Bertolo. Tel: +44 (0) 1227 78 2335; Fax: +44 (0) 1227 76 7531; Email Address: emilia.bertolo@canterbury.ac.uk

variety of plant species including *Canavalia gladiata* (swordbean) [7], *Viscum album coloratum* (mistletoe) [8], *Pisum sativum* (pea) [9], *Arum maculatum* L. [10] and *Eranthis hyemalis* (winter aconite) [11-14].

The lectin on which this study is focused is found in the tubers of winter aconite (*Eranthis hyemalis*), the first member of the Ranunculaceae family known to contain a lectin [11]. The *Eranthis hyemalis* lectin (EHL) makes up around 3% of the whole soluble protein content of the tubers; EHL is specific to N-acetyl-galactoseamine and blood group O erythrocytes [11]. EHL has been shown to have pesticidal activity against *Diabrotica undecimpunctata howardii*, a major insect pest of maize crops [12, 14].

Kumar, et al, [12] have identified EHL as a type-2 Ribosome Inactivating Protein (RIP), a class of proteins which are highly cytotoxic to eukaryotic organisms [15]. Type-2 RIP lectins provide plants with a defence mechanism against both vertebrate and invertebrate pests, and have also been found to have an inhibitory effect on plant viruses [12]. Other plant species containing RIP type-2 lectins include *Iris hollandica* (Dutch iris) [16] and *V. album coloratum* [8]. Another well known member of this group is the potent cytotoxic agent ricin, extracted from castor beans [12]. RIPs are formed by two polypeptide chains, a cytotoxic A chain and a carbohydrate-binding B chain. Binding of the B chain to a glycoconjugate on a cell surface mediates the entry of the cytotoxic A chain, which inactivates ribosomes. For EHL, initial studies suggest that the size of the A chain is between 28 and 30 kDa, the size of the B chain is 30 to 32 kDa, and the size of the whole lectin is between 55 and 62 kDa [11, 12].

Type-2 RIP lectins have been shown to have an inhibitory effect on viruses [12, 17], which has sparked considerable interest in the fields of medicine and pharmacology. Various plant lectins have been demonstrated to be useful in the diagnosis and treatment of human diseases such as Human Immunodeficiency Virus-1 (HIV-1); for example, the plant lectin jacalin can effectively block HIV-1 in vitro infection of lymphoid cells [17]. Lectins have also been used as diagnostic and therapeutic tools for cancer treatment [18]. Lectins from species such as mistletoe have a higher affinity to glycoconjugates found on the surface of tumour cells than those present on the surface of healthy cells; this anti-tumoral activity, demonstrated in-vitro, in-vivo and in human case studies, has been applied in cancer treatment in conjunction with conventional radiotherapy and chemotherapy [19, 20]. Lectins have also been studied as potential biomarkers for the early detection of pancreatic cancer, by detecting unique glycosylation patterns of proteins in the patient's serum [21].

Any potential applications of EHL in medicine, pharmacology or pest control rely heavily on the determination of the lectin's structure. Previous research groups were only partially successful, and so far only part of the amino acid sequence of the A chain has been determined [12]. Since purification is a crucial stage when aiming for crystals used

for structural analysis of any protein, improvements of the extraction protocol is a first vital step towards obtaining a suitable crystal for such structural analysis. Preliminary work carried out by our research group was successful in isolating EHL [22]. However, the yield obtained was very low, only 37  $\mu\text{g}\cdot\text{cm}^{-3}$ . Moreover, a brown deposit remained in the chromatography column after elution. This residue was attributed to the oxidation of phenols in the tissue by phenoloxidas-es, a problem often associated with the extraction of proteins and organelles from plant tissues. Phenols are oxidised to quinones which covalently bind to proteins, modifying structure and function; this reaction can be inhibited by the use of a phenoloxidase inhibitor [23]. Thiourea has been proposed as the most appropriate phenoloxidase inhibitor for the extraction of biologically active protein from phenol-rich plant tissues inhibitor [23].

Herein we present the revised extraction and purification of EHL using affinity column chromatography; thiourea was used as an antioxidant in order to prevent the lectin denaturing during the extraction process. The presence of EHL was determined by agglutination assays, using rabbit erythrocytes, since agglutination of red blood cells is a characteristic associated with lectins which provides an unambiguous indicator of lectin presence. Concentration of the extracted EHL was determined using a Bradford assay. SDS-PAGE (sodium dodecyl sulphate polyacrylamide gel electrophoresis) analysis was used to determine the size of the isolated lectin.

## 2. Materials and methods

### 2.1 Isolation and Purification of EHL

For the extraction and purification of the lectin from the winter aconite tubers the protocol described by Kumar et al. [12] was modified as follows. Tubers of the winter aconite (*Eranthis hyemalis*), purchased from MAS Seeds Ltd, were cleaned and peeled in an ice-bath before being finely diced. 40 g of the *E. hyemalis* tubers were homogenized in 500ml of 0.2 M sodium chloride solution (pH 5) containing 5 mM thiourea. The homogenized extract was allowed to settle for 30 minutes at 4°C, after which the supernatant was extracted and stored at 4°C. The remaining slurry was thoroughly resuspended in a further 500ml of 0.2 M sodium chloride solution with thiourea and allowed to settle for 30 minutes at 4°C; the supernatant was extracted again, added to the previous supernatant extraction, and stored at 4°C. The combined supernatant fractions were centrifuged at 14000 rpm at 4°C for 20 minutes. The fat layer was removed and discarded and the pH of the extractions was adjusted to 9 using 2M sodium hydroxide before being centrifuged again under the same conditions. The fat layer was again removed and the pH readjusted to pH 9. The extract was then filtered through 3MM Whatman chromatography paper.

Approximately 300ml of the raw extract was pumped through a CNBr-activated Sepharose affinity column (GE

Healthcare) coupled with fetuin from fetal calf serum (Sigma Aldrich). The column containing the bound lectin was washed with 150 mM phosphate-buffered saline solution (PBS), pH 7.2, to remove any unbound protein. The column bound EHL was then eluted with 30 ml of 20 mM diaminopropanol (DAP). The presence of lectin in the eluent was determined using a UV path monitor at  $\lambda$  280 nm (Unicam 5625 UV/VIS Spectrometer). The eluent was neutralised with 9.5ml of 1M Tris-HCl.

## 2.2. SDS-PAGE analysis

Two samples of the post-column eluent were run on 15% Tris-HCl gels in a Mini-PROTEAN 3 electrophoresis module for 40 minutes at 200V. Both samples were incubated at room temperature for 5 minutes. One sample was used to determine the size of the lectin. The other sample was incubated at 95°C before the electrophoresis run, in order to denature EHL and determine the size of the constituent polypeptide chains. The SDS-PAGE gels were stained with Coomassie Brilliant Blue for one hour and de-stained with molecular biology grade water for 24 hours.

## 2.3. Agglutination test and Bradford assay

The presence of lectin in the post-column eluent was confirmed by mixing 0.05 cm<sup>3</sup> of defibrinated rabbit blood with 0.05 cm<sup>3</sup> of column eluent on a wetted microscope slide. A control sample was prepared by mixing 0.05 cm<sup>3</sup> of 150 mM PBS with 0.05 cm<sup>3</sup> of defibrinated rabbit blood. Agglutination, assessed under both a high-power and low-power microscope, was also visible to the naked eye.

The concentration of EHL in the column eluent was determined using the pre-programmed Bradford Assay method on a Thermo Scientific BioMate 3 UV-Vis Spectrophotometer. Standard protein solutions were prepared using 0.1, 0.2, 0.3, 0.4 and 0.5 mg.cm<sup>-3</sup> bovine serum albumin (BSA); absorption readings of these solutions at 595 nm were used to prepare a calibration curve. The absorption reading of a sample of post-column eluent was interpolated in the calibration curve in order to determine the concentration of EHL.

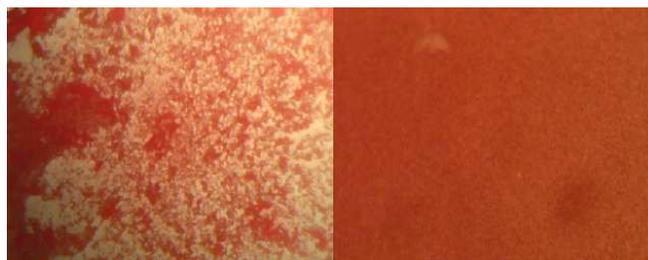
## 3. Results and discussion

Results suggest that EHL was successfully extracted and purified from the tubers of winter aconite. The sharp peak at  $\lambda$ =280 nm in the UV path monitor indicates the presence of protein in the post column eluent. Since the chromatography column is packed with fetuin, which shows a strong affinity for lectin proteins, any protein detected in the post-column eluent can be assumed to be lectin. The presence of lectin in the post-column eluent is also supported by the results of the agglutination test (see Figure 1). Agglutination of rabbit O-erythrocytes, visible even without the aid of a microscope, was observed in the assay containing the post-column

eluent. Agglutination of rabbit O-erythrocytes in the control assay was not observed.

Furthermore, the results from the SDS-PAGE analysis confirm that EHL was successfully extracted from the plant tubers, and corroborate that EHL is a dimeric lectin composed of two polypeptide chains of similar sizes. The size of the unreduced EHL was estimated as approximately 60 kDa. Incubation of the gel at 95°C before analysis produced the breakage of EHL into two chains of approximately 31 kDa in size (see Figure 2). These results are consistent with reported literature values [11, 12].

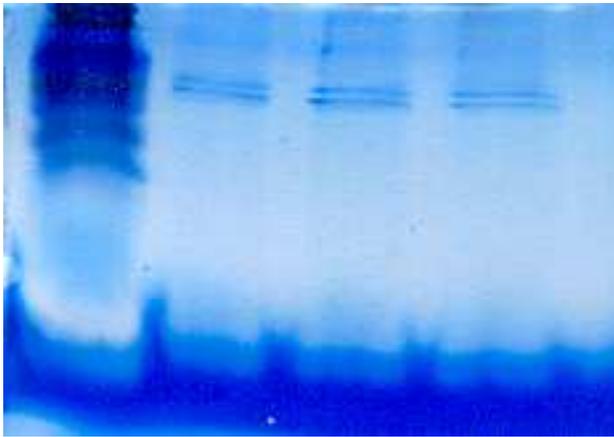
The concentration of EHL in the post-column eluent was measured using the Bradford assay. From the absorbance of the lectin eluent and those of the BSA standards, the concentration of EHL was determined as 380.1  $\mu$ g.cm<sup>-3</sup>. Preliminary work carried out by our research group [22], although successful in isolating EHL, produced a yield of only 37  $\mu$ g.cm<sup>-3</sup>. The low yield was attributed to protein oxidation during the extraction process, evidence of which was a



**Figure 1.** Blood smear (left) exhibits agglutination in the presence of the extracted EHL. Blood smear (right) exhibits even distribution of erythrocytes in the presence of the phosphate buffered saline (PBS) control; no agglutination is present (magnification x100).

brown deposit remaining in the chromatography column after elution. The use of thiourea as an antioxidant eliminated the presence of this brown residue in the affinity column. This suggests that the residue was, as suspected, the result of oxidation processes which resulted in the denaturing of the lectin protein. Beside the use of thiourea as antioxidant, additional measures taken in order to maintain the integrity of the EHL included ensuring that the sample was constantly maintained at approximately 4°C. These precautions could explain the significant improvement observed in the concentration of EHL in the post-column eluent.

Variations from the original method published by Kumar et al [12] include the use of PBS as buffer solution. Moreover, the original paper used two columns for the separation, one containing asialofetuin immobilised on agarose, followed by a Sepharose Q Fast flow anion exchange column. These have now been replaced with a single column step, using Fetuin bound to a Sepharose 4B column. Furthermore, tetraborate (used in the original paper) has been replaced by DAP, as the latter has been found to be more effective in eluting the lectin from the column, hereby reducing the volume of desorbant needed.



**Figure 2.** SDS-PAGE results for the gel incubated at 95°C before the electrophoresis run, in order to denature EHL. The gel shows the standard Kaleidoscopic marker (left hand side column) and EHL post-column eluent (three columns on the right). All EHL bands show two strands at approximately 31 kDa, corresponding to the reduced lectin.

#### 4. Conclusion

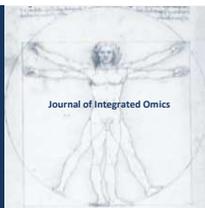
The lectin protein found in tubers of winter aconite (*Eranthis hyemalis*) has been successfully extracted and characterised. EHL was purified using affinity column chromatography and ammonium sulphate precipitation; thiourea was used as an antioxidant in order to prevent the lectin denaturing during the extraction process. The success of the extraction protocol was verified using a blood agglutination test with rabbit erythrocytes. Using SDS-PAGE analysis, the size of the unreduced EHL was established as approximately 60 kDa; furthermore, EHL was determined to be formed of two non-identical chains with molecular weights of approximately 31 kDa. The concentration of the EHL in the post-column eluent, determined using the Bradford Assay, was 380.1  $\mu\text{g}\cdot\text{cm}^{-3}$ . This is an order of magnitude improvement over our previous work, and enables us to undertake in-depth study of effect of the lectin on various plant pathogens and pests, and thus to determine its potential crop protection applications.

#### Acknowledgements

The authors thank Canterbury Christ Church University Research Fund for financial support. The authors would also like to thank Dr. David Ponsonby for assistance with SDS-PAGE queries.

#### References

1. A. Pusztai, Plant Lectins. Cambridge University Press, Cambridge, 1991.
2. E.J.M. Van Damme, W.J. Peumans, A. Pusztai, S. Bardocz, Handbook of Plant Lectins: properties and biomedical applications. John Wiley and Sons, New York, 1998.
3. N. Sharon, H. Lis, Lectins, second edition, Springer, The Netherlands, 2007.
4. I.E. Liener, N. Sharon, I.J. Goldstein (Eds.) (1986) The Lectins: Properties, Functions and Applications in Biology and Medicine, Academic Press Inc., London.
5. H. C. Sharma, K. K. Sharma, N. Seetharama, R. Ortiz, Electronic Journal of Biotechnology 3(2), Issue August 15 (2000) DOI: 10.2225/vol3-issue2-fulltext-3 (Accessed 24 September 2011).
6. Y.-Z. Pang, J.-H. Yao, G.-A. Shen, H.-X. Qi, F. Tan, Acta Botanica Sinica 46(7) (2004) 767-772.
7. P. Delatorre, B.A.M. Rocha, E.P. Souza, T.M. Oliveira, G.A. Bezerra, F.B.M.B. Moreno, B.T. Freitas, B.T. Santi-Gadelha, A.H. Sampaio, A.H. Jr Azevedo, B.S. Cavada, BMC Structural Biology 7(52) (2007) D.O.I. 10.1186/1472-6807-7-52 (Accessed 24 September 2011).
8. W. Ye, R.P.R. Nanga, C.B. Kang, J. Song, S.K. Song, H.S. Yoon, Journal of Biochemistry and Molecular Biology 39(5) (2006) 560-570.
9. S.A. Al-Sohaimy, E.E. Hafez, A.E. Abdelwahab, M.A. El-Saadani, Australian Journal of Basic and Applied Sciences 1 (3) (2007) 213-219.
10. E.J.M. Van Damme, K. Goossens, K. Smeets, F. Van Leuven, P. Verhaert, W.J. Peumans, Plant Physiology 107(4) (1995) 1147-1158.
11. Cammue, B.P., Peeters, B., Peumans, W.J., Biochem. Journal 227 (1985) 949-955.
12. M.A. Kumar, D.E. Timm, K.E. Neet, W.G. Owen, W.J. Peumans, A.G. Rao, J. Biol. Chem. 268(33) (1993) 25176-25183.
13. W.J. Peumans, E.J.M. Van Damme, Plant Physiology 109(2) (1995) 347-352.
14. D. Dent, Insect Pest Management, second edition. CABI Publishing, Oxon, 2000.
15. E. Gorelik, in: Lectins methods and protocols, J. M. Rodhes, J. D. Milton (Eds), Humana Press, New Jersey, 1998.
16. Q. Hao, E.J.M. Van Damme, B. Hause, A. Barre, Y. Chen, P. Rouge, W.J. Peumans, Plant Physiology 125 (2001) 866-876.
17. J. Fevero, P. Corbeau, M. Nicolas, M. Benkirane, G. Trave, J.F.P. Dixon, P. Aucouturier, S. Rasheed, J.W. Parker, P. Liautard, C. Devaux, J. Dornane, European Journal of Immunology 23(1) (2005) 179-185.
18. R. Mody, S. H. Joshi, W. Chaney, Journal of pharmacological and Toxicological Methods 33 (1) (1995) 1-10.
19. E.G. De Mejia, V.I. Prisecaru, Critical reviews in food science and nutrition 45(6) (2005) 425-445.
20. K. Hostanska, V. Vuong, S. Rocha, M.S. Soengas, C. Glanzmann, R. Saller, S. Bodis, M. Pruschy, British Journal of Cancer 88(11) (2003) 1785-1792.
21. C. Li, D. M. Simeone, D. E. Brenner, M. A. Anderson, K. A. Shedden, M. T. Ruffing, D. M. Lubman, J. Proteome Res. 8 (2) (2009) 483-492.
22. C. Solscheid, BSc dissertation, Department of Geographical and Life Sciences, Canterbury Christ Church University, 2006.
23. E. Van Driessche, S. Beeckmans, R. Dejaegere, and L. Kanarek, Analytical Biochemistry 141 (1984) 184-188.



ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v1i2.75

## Identification of spore specific allergens from *Penicillium chrysogenum*

Verena Goodman<sup>1</sup>, Andrea Müller<sup>1</sup>, Helke Gröger-Arndt<sup>1</sup>, Dietmar Schlosser<sup>2</sup>, Nico Jehmlich<sup>1</sup>, Regina Treudler<sup>3</sup>, Irina Lehmann<sup>4</sup>, Jan-Christoph Simon<sup>3</sup>, Martin von Bergen<sup>1,5\*</sup>.

<sup>1</sup>UFZ – Helmholtz Centre for Environmental Research Leipzig-Halle, Department of Proteomics, Permoserstr. 15 D-0418 Leipzig, Germany; <sup>2</sup>UFZ – Helmholtz Centre for Environmental Research Leipzig-Halle, Department of Environmental Microbiology, Permoserstr. 15 D-0418 Leipzig, Germany; <sup>3</sup>Department of Dermatology, Venerology and Allergology, Leipzig University Medical Center, Leipzig, Germany; <sup>4</sup>UFZ – Helmholtz Centre for Environmental Research Leipzig-Halle, Department of Environmental Immunology, Permoserstr. 15 D-0418 Leipzig, Germany; <sup>5</sup>UFZ – Helmholtz Centre for Environmental Research Leipzig-Halle, Department of Metabolomics Permoserstr. 15, D-04318 Leipzig, Germany.

Received: 21 October 2011 Accepted: 09 December 2011 Available Online: 28 December 2011.

### ABSTRACT

**Background:** The common indoor mould *Penicillium chrysogenum* is reported to be associated with respiratory diseases and allergic asthma. Beside the already known specific allergens stemming from mixtures of hyphae and spores the question was which allergens are specifically abundant in the airborne spores.

**Methods:** Serum samples of 50 patients with allergic symptoms were screened for a positive reaction with spore proteins using 1D immunoblotting. The sera of 10 patients with different immunopositive patterns in the 1D blots were selected for 2D westernblotting and IgE reactive proteins were identified by mass spectrometry.

**Results:** 29 proteins specifically interacting with IgE from sera were identified. 5 of them showed a frequency higher than 50% and were identified as Glyceraldehydes-3-phosphate dehydrogenase, Catalase A, delta-1-pyrroline-5-carboxylate dehydrogenase, glucose-6-phosphate isomerase and transaldolase. The vacuolar serine protease, known as the major allergen Pen ch 18, was identified in 3 cases.

**Conclusions:** Beside the confirmation of known allergens like Pen ch 18, new spore specific IgE binding potential allergens were identified. The differentiation into hyphae and spore specific allergens might be helpful for a more specified diagnosis and to differentiate the source of exposure leading to sensitization.

**Keywords:** Fungi allergens; *Penicillium chrysogenum*; SDS Page.

### 1. Introduction

Factors affecting indoor air quality are of increasing concern because, in industrialized societies, people spend up to 80% of their lifetime indoors [1]. There they are exposed to various factors including a large variety of chemical substances and allergens from different sources. Beside the house dust allergens, consisting mostly of house dust mite allergens, moulds are a dominant source of allergenic substances [1]. Among them there are species which occur predominantly in the outdoor environment but enter households by ventilation like species of *Cladosporium* and *Alternaria*. These typical outdoor fungi often grow on rotting plant material with a seasonal peak in summer [2].

In temperate areas construction and leakage problems or poor ventilation can result in increased dampness indoors promoting mould growth on organic material. The fungal profile in houses with dampness and mould problems is shifting to species of *Penicillium* and *Aspergillus*, such as *Penicillium chrysogenum*, *P.expansum* and *Aspergillus versicolor* with higher spore concentration in winter [3, 4].

In recent years the health risk of living in damp and mouldy environment was analysed in numerous studies. Mould exposure was identified as risk factor for various respiratory diseases, allergic symptoms and non specific symptoms like headache, eye and skin irritation [1, 5, 6]. Especial-

\*Corresponding author: Martin von Bergen. Helmholtz Centre for Environmental Research Leipzig-Halle, Department of Proteomics, Permoserstr. 15 D-0418 Leipzig, Germany. E-mail Address: Martin.7vonbergen@ufz.de.

ly, exposure with *Penicillium* spores is associated with asthma attacks and wheezing [7, 8].

The extraction of proteins was performed in earlier studies from complete fungal mycelium, consisting mostly out of hyphae and containing only a minor fraction of spores. The spores of *P. chrysogenum* are relative small with 2.6  $\mu\text{m}$  as the average aero-dynamic diameter [9] resulting in a higher tendency to become airborne. This is in particular true in comparison to fungal hyphae, which are first most often attached to the substrate and second are much larger and therefore more prone to end up in the house dust rather than becoming airborne. Since the most direct contact of patients occurs via the airborne spores rather than by mycelium we investigated the IgE reactive proteins out of extracts from spores from *P. chrysogenum*.

From *P. chrysogenum* there are so far seven allergens described in the Allergome database [10]. The list includes the following allergens: Pen ch 13 (Subtilisin like serin protease; Uniprot:Q9URR2), Pen ch 18 (Vacuolar serine protease, Uniprot:Q9P8G3), Pen ch20 (N-Acetyl Glucosamidase; Uniprot:Q02352), Pen ch 31 (Calreticulin; Uniprot:Q2TL59), Pen ch 33 (an uncharacterized 16 kDa protein; Uniprot: B0L0W9), Pen ch 35 (Transaldolase) and Pen ch MnSOD (Mangan Superoxide Dismutase; Uniprot: B6H9W9). The best characterized allergen, the serine protease from the subtilisin-type [11] is described as Pen ch 13 [12]. The induction of airway inflammation is associated with changes in junctional structure alterations resulting in increased permeability [13]. Another serine protease, the vacuolar serine protease (Pen ch 18) [14] is described as allergen from *P. chrysogenum* [15] and also from several other fungi [16-18]. Among the different serine proteases there occurs a relative high amount of cross-reactivity of IgE antibodies [19]. The Pen ch 20 allergen, the N-acetyl glucosamidase [20] is involved in the modification of sugars, relevant for building the chitin layer of the fungal cell wall. Calreticulin (Pen ch 31) is a chaperone and as such of relative high abundance. A rather meager characterization is available for Pen ch 33, which was identified once and for which only the open reading frame without even an supposed function is known. In a recent study transaldolases, listed as Pen ch 35 were described as a family of cross-reacting allergens in *C. cladosporioides* and *P. chrysogenum* [21].

Mold spores can be associated with Type I – IV allergies, but the type I is the most prevalent [22]. Allergic rhinitis, allergic asthma and conjunctivitis are all symptoms of Type I allergies, which are caused by the interaction between IgE antibodies and the allergens [23]. The classical way to find new allergens is to blot protein extracts, probe with sera from affected patients and detect the interaction between IgEs and potential allergens by an IgE specific antibody. This works especially well, when the proteins from a crude mixture are separated to near homogeneity into spots by 2D-gel electrophoresis [24-26].

By this approach sera from 50 patients with symptoms of atopic diseases were screened by 1D gel IgE-immunoblotting

and the patients with the strongest signals were selected for 2D-gel IgE immunoblotting. Among the identified IgE-binding proteins were well known allergens from *P. chrysogenum* and some so far not described putative allergens. The identified proteins were analysed in terms of their frequencies in order to obtain a semi-quantitative ranking of their importance and to allow a comparison between the spore specific allergen spectra from the one described for extracts obtained from hyphae or mixtures of hyphae and spores.

## 2. Material and Methods

### *Serum samples*

Serum samples of 50 patients were randomly selected from a population of patients with allergy diagnoses. The samples were taken from a blood repository, more specifically from 33 women and 17 men, which age ranged between 12 and 73 years, with a mean of 46.2 years. These patients were diagnosed for different allergic diseases at the Medical Department for Dermatology and Allergology at the University of Leipzig. (urticaria/angioedema n=16, adverse drug reactions n=14, contact allergy n=5, anaphylaxis n=3, atopic eczema n=3, mastocytosis n=2, food allergy n=2, others n=5).

### *Cultivation of P. chrysogenum and preparation of spore extract*

*P. chrysogenum* was cultivated for 21 days at 20°C in plastic flasks for cell culture containing dicloran-glycerol agar with chloramphenicol (DG 18, Oxoid, Wesel, Germany). The DG 18 agar reduces the growth mycelium and boosts the development of the spores [27]. The cultivated strain belongs to the Centre of Environmental Research (UFZ), Leipzig, Germany. The produced spores were collected from surfaces of the cultures and suspended in a solution of 3 mM chloramphenicol and 1 mM phenylmethylsulfonylfluoride (PMSF) in 50 mM Tris/HCl (pH 7.5). This suspension was centrifuged at 15,000 g at 4°C for 30 min. The pellet was re-suspended in a solution of 3 mM chloramphenicol in 20 mM Tris/HCl (pH 7.5). The spores in this suspension were disintegrated using a bead mill (Retsch GmbH, Haan, Germany). The disintegration was conducted at 4°C for 1.5 h, adding 1 to 1.5 mg glass beads (d=0.75 to 1 mm) per 1 ml of suspension. The suspension resulted from the disintegration process was twice centrifuged at 16,000 g for 20 min at 4°C and the supernatant was frozen at -20°C. The protein concentration in this supernatant was determined using the DC Bio-Rad Protein Assay, (Bio-Rad Deutschland, Munich) which derives from the method of Lowry [28].

### *1D SDS-PAGE, 2D electrophoresis and Immunoblotting*

A discontinuous SDS-Page with two different acrylamide concentrations was used for more precise bands. A 12% resolving gel (12%, 0.05% APS- 0.1% SDS, 0.075% TEMED,

0.375 Tris-HCl) and a 4% stacking gel which contains in contrast to the resolving gel 4% Acrylamide-Bisacrylamide and 0.125 M Tris-Cl (pH8.8) were produced.

A solution of 20% ice-cold TCA was used to precipitate 50 µg proteins after incubation for 4 h and the pellets were washed twice with 100% ice-cold acetone. The obtained pellet was dissolved in 10 µL sample buffer (Bromophenol Blue 0.1% (w/v), 5% 2-mercaptoethanol, 10% Glycerol, 2% SDS, 125 mM Tris-Cl), heated for 5 min at 60°C and applied to the gel.

For 2D-gel electrophoresis, 500 µg protein were purified by phenol extraction with subsequent solvent precipitation and washing steps as described earlier<sup>29</sup>. The resulting protein pellets were dried at ambient temperature and dissolved in 135 µL of DeStreak rehydration solution with 0.5% IPG (immobilized pH gradient) pH 3–10 non-linear (NL) buffer (v/v) (GE Healthcare, Uppsala, Sweden). The electrophoresis was performed as described earlier<sup>30</sup>. Gels were stained by Coomassie Brilliant Blue G-250 (CBB) staining (Roth, Kassel, Germany).

A second gel, which was handled parallel with the same procedure, was used for immuno-blotting. The proteins were transferred by electroblotting to a nitrocellulose membrane with pores of 0.2 µm, (Schleicher & Schuell, Dassel, Germany) using a Western Blot chamber as described previously [31]. In brief the chamber XCell II Blot Module (Invitrogen, Karlsruhe, Germany) and 10 mM CAPS buffer (10 mM CAPS- *N*-cyclohexyl-3-aminopropanesulfonic acid, 10% (v/v) Methanol (pH 11) were used. A current of 1.5 mA/cm<sup>2</sup> was applied for 90 min at ambient temperature. The blotting efficiency was checked by reversible PonceauS staining (0.2% PonceauS in 3% TCA).

For immunological detection, the membranes were washed with distillate water and blocked with 3% BSA in TBS (20 mM Tris, 500 mM NaCl, pH 7.5) overnight, and incubated with or without the serum samples that were diluted 1:40 in antibody buffer (TBS with 0.05% Tween 20 and 1% BSA) for 120 min at ambient temperature. Afterwards, the membranes were washed twice with TTBS (0.05% Tween 20 in TBS) for 10 minutes each and incubated with goat anti-human IgE secondary antibody conjugated with alkaline phosphatase (Sigma-Aldrich, Germany) in a dilution of 1:2000. After incubation for 2 h, the membranes were washed twice with TTBS and once with TBS. The chemiluminescence emitted was detected with the AP Conjugate Kit (Bio-Rad, Geisenhofen, Germany).

#### Protein identification by mass spectrometry

Protein spots that react positively with the serum samples were cut from polyacrylamide gels and digested overnight with trypsin (Sigma-Aldrich, Geisenhofen, Germany) as described previously [32]. The resulting peptides were eluted, concentrated by vacuum centrifugation and thereafter separated by RP nano-LC 1100 series (Agilent Technologies, Palo Alto, USA), with an analytical column Zorbax 300SB-

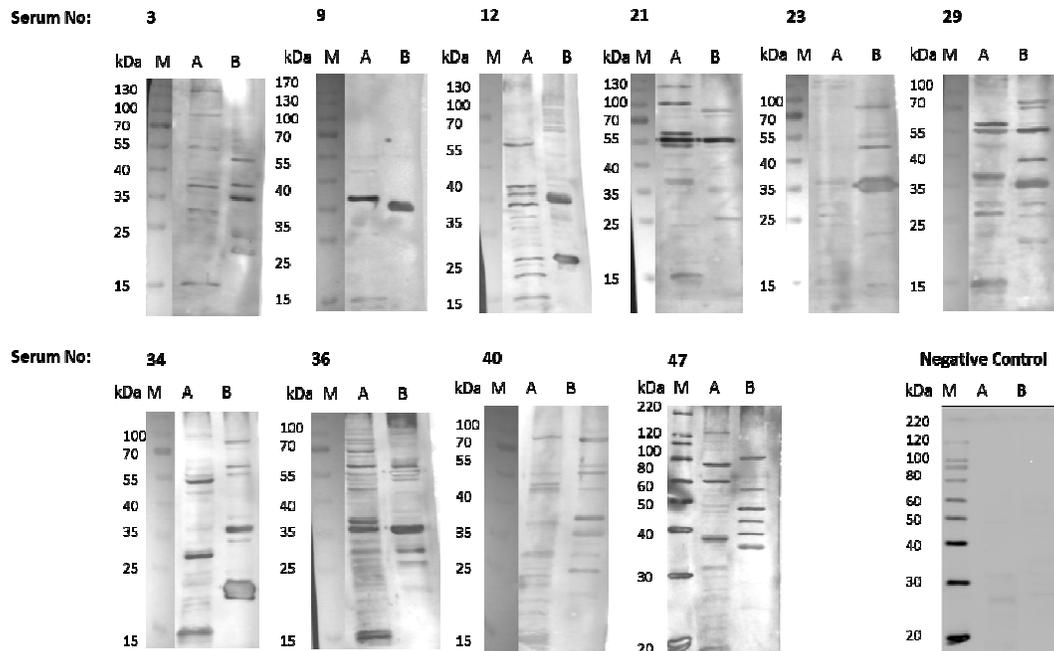
C18, 3.5 µm, 150 x 0.075 mm, linear gradient with 0-60% acetonitrile (ACN) in 35 min, using solvent A (97% water/3% ACN/0.1% formic acid) and solvent B (97% ACN/3% water/0.1% formic acid). The peptides were identified by on-line MS/MS (LC/MSD TRAP XCT mass spectrometer, Agilent Technologies), The peptide identifications were achieved by comparisons with the NCBI nr database, 20090310, all fungal entries, using Mascot in-house version 2.2.1 (Matrix Science, London, UK) with subsequent parameters: trypsin digestion, up to one missed cleavage, fixed modifications: carbamidomethyl (C) and with the following variable modifications: methionine oxidation and carbamidomethylation of cysteines, peptide tol.: ± 1.2 Da, MS/MS tol.: ± 0.6 Da and peptide charge: +1, +2 and +3. The amino acid sequences were compared for similarity with known proteins using the Allergome database with the algorithm Basic Local Alignment Search Tool protein (BLASTP) under the standard parameters BLOSUM62.

### 3. Results

All 50 sera were tested for a reaction with *P. chrysogenum* and additionally with *A. versicolor* in 1D blots. A set of 10 sera which were tested also in 2D gels, are presented in figure 1. For control we also incubated a blot without serum and solely with the secondary antibody. The result (lower right panel in Figure 1) indicate very weak bands for A at ~26 kDa and 32 kDa and for lane B at ~28 kDa and ~35 kDa. Since we used standardized conditions in terms of the amount of protein and serum and also for western blotting we are convinced that the background is very low and does not influence the signal intensity of the samples in a significant way. The sera showed a variety of IgE positive bands, for *A.versicolor* as well as for *P.chrysogenum*. Serum 12, 23, 29, 34, 36 and 47 are examples for a strong immune reactivity to both investigated spore extracts in terms of detected bands and staining intensity. It is remarkable that most sera exhibited different patterns in their reactivity. A dominant band for *P.chrysogenum* with an apparent molecular weight of 35 kDa was found in nearly all sera. A second band with a molecular weight approximately 55 kDa was also frequently detected.

From the 50 sera screened by 1D blots, 10 strongly reacting sera, which showed a variety of positive bands, were selected for detecting and identification of proteins in 2D blots.

After incubating 2D blots of extract from *P.chrysogenum* spores with patient sera, IgE-reacting spots were detected and in order to align the spots on the membrane with spots in the gel, the membranes were stained by Ponceau S and the resulting spot patterns on membranes and gels were aligned by the 2D-analysis software DECODON. Three representative examples are shown in Figure 2. The signals show detection of several groups of spots that resemble a pearl-chain appearance. There are spots at different molecular weight, also several at the apparent molecular weight of 35 kDa as



**Figure 1.** 1-D Immunoblot of selected sera, Immunological reaction against *A.versicolor* (A) and *P.chrysogenum* (B), M – unstained protein molecular weight marker, negative control without serum.

already seen in 1D blot analysis and at different pI. The horizontal pearl chain points to pI effective modifications or isoforms of one protein.

For identification of the immunoreactive proteins, a master gel was prepared and stained with Coomassie, shown in figure 3. All proteins spots, which reacted with one or more of the analyzed sera, were marked and the gels were aligned to the mastergel by DECODON. Coomassie stained spots corresponding to immune reactive spots were cut and proteins contained within were identified by mass spectrometry. The list of all detected allergens with decreasing frequency is given in table 1. All spots yielded unambiguous identifications and no spot revealed second hits, underlining the quality of the achieved resolution on the gel.

The immunoreactive spots were ranked according to their frequency within the tested 10 patients and the highest frequency (6/10) was found for Glyceraldehyde-3-phosphate dehydrogenase (GAPDH). A second group with a frequency of 5/10 included Catalase, 1-pyrroline-5-carboxylate dehydrogenase, Glucose-6-phosphate isomerase and Transaldolase. Pc21g01220 from the oxidoreductase family and fumarate hydratase were identified 4 times, a protein from the aldehyd-dehydrogenase family and hydralase 3 times in 10 sera. Pen ch 18, a vacuolar serine protease which is known as allergen from *P.chrysogenum*, was found also 3 times. We identified 18 more proteins which showed an immune reaction with human sera (Table 1).

#### 4. Discussion

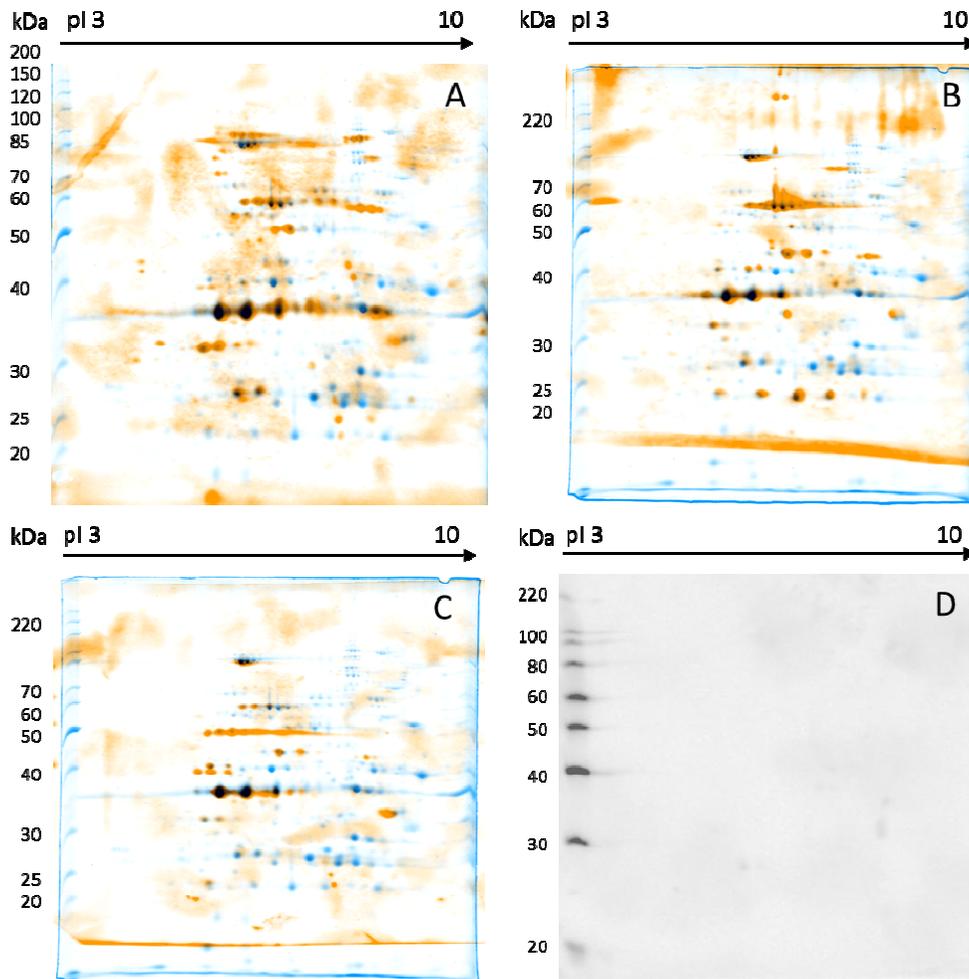
The advancement of this study compared to other quests to identify allergens of *P. chrysogenum* lies in the focus on

extracts from spores. This allows identifying spore specific IgE reactive proteins. Among them only one already known allergen, Pen ch 18, a vacuolar serine protease was detected in a minority of cases. Shen *et al.* [33] described Pen ch 18 as allergen with an association to asthma. Within the selected patients for this study, no asthmatic diseases were included. Nevertheless in three sera also Pen n 18 was identified. Other known allergens from *P. chrysogenum* like Pen ch 13, an alkaline serine protease and Pen ch 20, N-Acetyl glucosamidase were not found in this study underlining the difference between whole cell lysates and spore specific extraction [12, 20].

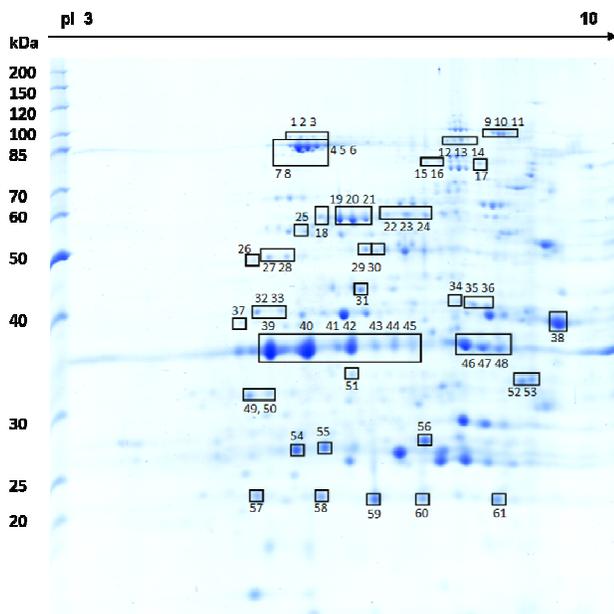
The major allergens are usually defined as those allergens which can be found in 50% of people who suffer from the corresponding allergy [34]. Since the patients analyzed in this study did not mainly suffer from fungal allergies this characterization cannot be used without problems. Additionally, the 1D blots were previously screened for strong bands and hence cannot be used as a reference point for overall frequency. Therefore, an arbitrary minimum limit of 50% should include most major allergens. However, the term major allergen is avoided in favor of high-frequency allergens. As high-frequency allergen the proteins GAPDH, catalase A, delta-1-pyrroline-5-carboxylate dehydrogenase, glucose-6-phosphate isomerase and transaldolase were identified.

GAPDH is an enzyme, which appears to be a main allergen in *P. chrysogenum* and *A. versicolor* [35]. Also the proteins stemming from *P. expansum* had shown a certain amount of cross-reactivity (unpublished data).

Interestingly, wheat glyceraldehyde-3-phosphate dehydrogenase has been identified as an allergen as well [36]. For the



**Figure 2.** Examples of layered 2 D gels Coomassie-stained gel (blue) and chemiluminescence detection (orange) A-C selected sera (A-serum 23, B-serum 29, C-serum 47, D-negative control without serum).



**Figure 3.** Positive reacting spots in Coomassie stained master gel. First Lane: Unstained Protein Marker. Remaining figure: Two dimensionally separated proteins of *P. chrysogenum*.

alignment, the nucleotide sequence of *Hordeum vulgare*, was taken since the exact sequence from the *Triticum aestivum* enzyme was not revealed in the respective paper [36]. A comparison between the enzyme of *Hordeum vulgare* and *P. chrysogenum* showed an identity of 71% and 84% positives. However, the length of both enzymes is similar. This points towards the epitope being in the 200 amino acid sequence which does not correspond to the *Alternaria* variant of the enzyme. This theory still has to be confirmed by epitope analysis.

The identity between the GAPDH of *P. chrysogenum* and humans is 70%, as determined by BlastP (84% positives and 0% gaps), thus only hardly allowing an allergic reaction according to the statistical estimation that proteins with more than 62% identity to human proteins are only rarely allergenic [37]. It will be assumed that a development of IgE reactivity to highly conserved allergenic molecules leads to cross-reactivity with divergent mould species [38]. Autoimmune responses caused by the production of IgE antibodies against fungal allergens have been described before in various papers [39-41]. The human variant of the analyzed enzymes is able to cause a strong reaction in skin prick tests. It has been sug-

**Table 1.** Name and molecular function of identified proteins.

Uniprot ass No.	Protein Molecular Function	Species	Protein scores	Mol W (Da)	Frequency	Spot
Pc21g14560	Glyceraldehyde-3-phosphate dehydrogenase	P.chrysogenum	1353, 1361, 1283, 1155, 1012, 1044, 1251	36149	6	39, 40, 41, 42, 43, 44, 45
Pc20g06360	Catalase	P.chrysogenum	1692, 1081, 624, 711	84305	5	4, 5, 6, 7, 8
Pc12g08900	1-pyrroline-5-carboxylate dehydrogenase activity	P.chrysogenum	1433	62610	5	18
Pc22g19730	Glucose-6-phosphate isomerase	P.chrysogenum	1357, 1403, 522	61256	5	19, 20, 21
Pc21g16950	Transaldolase	P.chrysogenum	1239, 1094, 1166	35589	5	46, 47, 48
Pc21g01220	oxidoreductase activity	P.chrysogenum	535,202	35392	4	54, 56
Pc22g23810	fumarate hydratase activity	P.chrysogenum	1007	58125	3	29
Pc06g01800	oxidoreductase activity	P.chrysogenum	763, 649	37416	3	32,33
Pc06g00180	aldehyddehydrogenase family	P.chrysogenum	840	54046	3	34
Pen ch 18	vacuolar serine protease	P.chrysogenum	612, 569	38489	3	49,50
Pc13g08730	hydrolase activity,	P.chrysogenum	804, 664	34666	3	52,53
Pc16g11860	Catalase	P.chrysogenum	839, 917, 808	79912	2	12, 13, 14
Pc22g11240	heat shock protein familie	P.chrysogenum	1285	69692	2	25
Pc22g05800	Saccharopine Reductase	P.chrysogenum	865	49177	2	30
Pc20g04720	Dihydrolipoyl dehydrogenase	P.chrysogenum	1171	54589	2	31
Pc21g17460	adenosine kinase activity	P.chrysogenum	953	43589	2	35,36
Pc13g07960	oxidoreductase activity	P.chrysogenum	799	36973	2	37
Pc20g11580	Superoxide dismutase	P.chrysogenum	428	24668	2	57
Pc22g25220	FMN binding oxidoreductase activity	P.chrysogenum	575, 342, 242	21307	2	59, 60, 61
Pc22g18630	methionine biosynthetic process	P.chrysogenum	1544, 1136, 1641	87540	1	1,2,3
Pc18g03470	metallopeptidase activity	P.chrysogenum	2005, 2183, 2105	98916	1	9, 10, 11
Pc22g15910	Proteolysis	P.chrysogenum	752, 696	67986	1	15,16
Pc22g16760	dipeptidyl-peptidase activity	P.chrysogenum	1078	78783	1	17
Pc18g00980	Not identified	P.chrysogenum	723, 585, 678	57162	1	22, 23,24
Pc13g12450	ATP binding heat shock protein 70 family	P.chrysogenum	305	74784	1	26
Pc18g00650	phosphotransferase activity	P.chrysogenum	550, 688	48009	1	27, 28
Pc13g09680	aspartic-type endopeptidase activity	P.chrysogenum	953	43589	1	38
Pc13g15120	oxidoreductase activity	P.chrysogenum	329	31448	1	51
Pc22g08970	Not identified	P.chrysogenum	324	21844	1	58

gested that auto antigens against manganese superoxide dismutase play a role in atopic dermatitis [42].

Catalase is an enzyme that is abundant in tetramers. It catalyzes the reaction of hydrogen peroxide to oxygen and water. Since hydrogen peroxide is a side product of challenge for cells exposed to oxygen. Therefore, almost all aerobic organisms have this enzyme. Its catalytic activity is attained through its heme groups. Each tetramer unit binds to one heme group [43].

Catalase A is another allergen found in all three species. It is the third most reactive allergen found in *A. versicolor* [35] and the fourth most abundant allergen in *P. chrysogenum*. The analyzed sequence contains the information for one of the subunits. In previous studies, it has been shown that humans react to catalase of *Metarhizium anisopliae*, a mold which infects insects. The positive reaction is attributed to cross-reactivity with different fungi [44]. This theory coin-

cides with the observation that catalase A is a major allergen in many indoor fungi. However, an alignment of catalase of *Metarhizium anisopliae* and catalase A of *P. chrysogenum* only shows 43% identity, 58% positives and 4% gaps. Catalase has been identified as an allergen in *Penicillium citrinum* [16].

50% of the chosen sera were from patients with urticaria, whereas the overall frequency was merely 20%. This might lead to the conclusion that patients with urticaria show a stronger reaction to indoor molds than the average patient with atopic disease. Nevertheless, 4 out of the 10 sera showed only a very slight reaction with the indoor mold proteins.

Urticaria is an allergic disease that can be caused either by contact, after inhalation or by digestion of allergens. After digestion the remaining peptides of the allergen are released into the blood stream. After skin contact, a localized reaction occurs which produces hives. The mast-cell activation in-

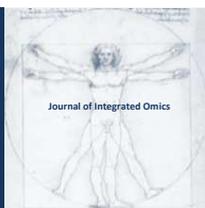
creases the vascular permeability and causes vasodilation of blood vessels [45]. A correlation between urticaria and dermatophysis has been found [46].

Beside the protein Pen ch 18 as a major allergen from *P. chrysogenum* related to asthma we identified other proteins with a high frequency. It can be assumed, that these patients were exposed to *P. chrysogenum* resulting in a sensitization. The Allergome database has introduced a classification system that illustrates the hierarchy in allergen characterisation ranging from first indications of allergenicity over IgE and several *in vivo* tests up to tests in epidemiological studies [10]. The identification of immune-reactive proteins from spores of *P. chrysogenum* will facilitate further clinical studies with the aim to provide an individual and source and allergen specific allergy profile that in turn will support prevention and treatment.

## References

1. WHO, Dampness and mould. In *WHO guidelines for indoor quality*, 2009.
2. Samson, R. A.; Hoekstra, E. S.; Frisvad, J. C.; Filtenborg, O., *Introduction to food- and airborne fungi*. 6 ed.; Centraalbureau voor Schimmelcultures: Utrecht, 2002.
3. Müller, A.; Lehmann, I.; Seiffart, A.; Diez, U.; Wetzig, H.; Borte, M.; Herbarth, O., Increased incidence of allergic sensitisation and respiratory diseases due to mould exposure: results of the Leipzig Allergy Risk children Study (LARS). *Int J Hyg Environ Health* 2002, 204, (5-6), 363-5.
4. Herbarth, O.; Schlink, U.; Müller, A.; Richter, M., Spatiotemporal distribution of airborne mould spores in apartments. *Mycol Res* 2003, 107, (Pt 11), 1361-71.
5. Frisk, M.; Magnuson, A.; Kiviloog, J.; Ivarsson, A. B.; Kamwendo, K., Increased occurrence of respiratory symptoms is associated with indoor climate risk indicators - a cross-sectional study in a Swedish population. *Respir Med* 2007, 101, (9), 2031-5.
6. Sahlberg, B.; Mi, Y. H.; Norback, D., Indoor environment in dwellings, asthma, allergies, and sick building syndrome in the Swedish population: a longitudinal cohort study from 1989 to 1997. *Int Arch Occup Environ Health* 2009, 82, (10), 1211-8.
7. Gent, J. F.; Ren, P.; Belanger, K.; Triche, E.; Bracken, M. B.; Holford, T. R.; Leaderer, B. P., Levels of household mold associated with respiratory symptoms in the first year of life in a cohort at risk for asthma. *Environ Health Perspect* 2002, 110, (12), A781-6.
8. Rosenbaum, P. F.; Crawford, J. A.; Anagnost, S. E.; Wang, C. J.; Hunt, A.; Anbar, R. D.; Hargrave, T. M.; Hall, E. G.; Liu, C. C.; Abraham, J. L., Indoor airborne fungi and wheeze in the first year of life among a cohort of infants at risk for asthma. *J Expo Sci Environ Epidemiol* 2010, 20, (6), 503-15.
9. Madelin, T. M.; Johnson, H. E., Fungal and actinomycete spore aerosols measured at different humidities with an aerodynamic particle sizer. *J Appl Bacteriol* 1992, 72, (5), 400-409.
10. Mari, A.; Scala, E., Allergome: a unifying platform. *Arb Paul Ehrlich Inst Bundesamt Sera Impfstoffe Frankf A M* 2006, (95), 29-39; discussion 39-40.
11. van den Berg, M. A.; Albang, R.; Albermann, K.; Badger, J. H.; Daran, J. M.; Driessen, A. J.; Garcia-Estrada, C.; Fedorova, N. D.; Harris, D. M.; Heijne, W. H.; Joardar, V.; Kiel, J. A.; Kovalchuk, A.; Martin, J. F.; Nierman, W. C.; Nijland, J. G.; Pronk, J. T.; Roubos, J. A.; van der Klei, I. J.; van Peij, N. N.; Veenhuis, M.; von Dohren, H.; Wagner, C.; Wortman, J.; Bovenberg, R. A., Genome sequencing and analysis of the filamentous fungus *Penicillium chrysogenum*. *Nat Biotechnol* 2008, 26, 1161-8.
12. Shen, H. D.; Lin, W. L.; Tam, M. F.; Wang, S. R.; Tzean, S. S.; Huang, M. H.; Han, S. H., Characterization of allergens from *Penicillium oxalicum* and *P. notatum* by immunoblotting and N-terminal amino acid sequence analysis. *Clin Exp Allergy* 1999, 29, (5), 642-51.
13. Chen, J. C.; Chuang, J. G.; Su, Y. Y.; Chiang, B. L.; Lin, Y. S.; Chow, L. P., The protease allergen Pen c 13 induce allergic airway inflammation and changes in epithelial barrier integrity and function in a murine model. *J Biol Chem* 2011.
14. Yu, C. J.; Chen, Y. M.; Su, S. N.; Forouhar, F.; Lee, S. H.; Chow, L. P., Molecular and immunological characterization and IgE epitope mapping of Pen n 18, a major allergen of *Penicillium notatum*. *Biochem J* 2002, 363, (Pt 3), 707-15.
15. Shen, H. D.; Chou, H.; Tam, M. F.; Chang, C. Y.; Lai, H. Y.; Wang, S. R., Molecular and immunological characterization of Pen ch 18, the vacuolar serine protease major allergen of *Penicillium chrysogenum*. *Allergy* 2003, 58, (10), 993-1002.
16. Simon-Nobbe, B.; Denk, U.; Poll, V.; Rid, R.; Breitenbach, M., The spectrum of fungal allergy. *Int Arch Allergy Immunol* 2008, 145, (1), 58-86.
17. Chou, H.; Tam, M. F.; Lee, L. H.; Chiang, C. H.; Tai, H. Y.; Panzani, R. C.; Shen, H. D., Vacuolar serine protease is a major allergen of *Cladosporium cladosporioides*. *Int Arch Allergy Immunol* 2008, 146, (4), 277-86.
18. Shen, H. D.; Tam, M. F.; Tang, R. B.; Chou, H., *Aspergillus* and *Penicillium* allergens: focus on proteases. *Curr Allergy Asthma Rep* 2007, 7, (5), 351-6.
19. Poll, V.; Denk, U.; Shen, H. D.; Panzani, R. C.; Dissertori, O.; Lackner, P.; Hemmer, W.; Mari, A.; Cramer, R.; Lottspeich, F.; Rid, R.; Richter, K.; Breitenbach, M.; Simon-Nobbe, B., The vacuolar serine protease, a cross-reactive allergen from *Cladosporium herbarum*. *Mol Immunol* 2009, 46, (7), 1360-73.
20. Shen, H. D.; Liaw, S. F.; Lin, W. L.; Ro, L. H.; Yang, H. L.; Han, S. H., Molecular cloning of cDNA coding for the 68 kDa allergen of *Penicillium notatum* using MoAbs. *Clin Exp Allergy* 1995, 25, (4), 350-6.
21. Chou, H.; Tam, M. F.; Chiang, C. H.; Chou, C. T.; Tai, H. Y.; Shen, H. D., Transaldolases are novel and immunoglobulin E cross-reacting fungal allergens. *Clin Exp Allergy* 2011, 41, (5), 739-49.
22. Gell, P.; Coombs, R., *Clinical Aspects of Immunology*. Oxford, England: 1963.
23. Roth, L.; Frank, H.; Kormann, K., *Giftpilze - Pilzgifte*. 1 ed.; ecomed: 1990.
24. Petersen, A.; Becker, W. M.; Schlaak, M., Epitope mapping of allergens and antigens of timothy pollen extract. *Appl Theor Electrophor* 1991, 2, (4-5), 129-33.
25. Sander, I.; Fleischer, C.; Meurer, U.; Bruning, T.; Raulf-Heimsoth, M., Allergen content of grass pollen preparations for skin prick testing and sublingual immunotherapy. *Allergy* 2009, 64, (10), 1486-92.
26. Gutierrez-Rodriguez, A.; Postigo, I.; Guisantes, J. A.; Sunen, E.; Martinez, J., Identification of allergens homologous to Alt a 1 from *Stemphylium botryosum* and *Ulocladium botrytis*. *Med Mycol* 2011.
27. Verhoeff, A. P.; van Wijnen, J. H.; Brunekreef, B.; Fischer, P.;

- van Reenen-Hoekstra, E. S.; Samson, R. A., Presence of viable mould propagules in indoor air in relation to house damp and outdoor air. *Allergy* 1992, 47, (2 Pt 1), 83-91.
28. Lowry, O. H.; Rosebrough, N. J.; Farr, A. L.; Randall, R. J., Protein measurement with the Folin phenol reagent. *The Journal of biological chemistry* 1951, 193, (1), 265-75.
  29. Benndorf, D.; Balcke, G. U.; Harms, H.; von Bergen, M., Functional metaproteome analysis of protein extracts from contaminated soil and groundwater. *Isme J* 2007, 1, (3), 224-34.
  30. Georgieva, D.; Risch, M.; Kardas, A.; Buck, F.; von Bergen, M.; Betzel, C., Comparative analysis of the venom proteomes of *Vipera ammodytes ammodytes* and *Vipera ammodytes meridionalis*. *J Proteome Res* 2008, 7, (3), 866-86.
  31. Morbt, N.; Mogel, I.; Kalkhof, S.; Feltens, R.; Roder-Stolinski, C.; Zheng, J.; Vogt, C.; Lehmann, I.; von Bergen, M., Proteome changes in human bronchoalveolar cells following styrene exposure indicate involvement of oxidative stress in the molecular-response mechanism. *Proteomics* 2009, 9, (21), 4920-33.
  32. Jehmlich, N.; Schmidt, F.; Taubert, M.; Seifert, J.; von Bergen, M.; Richnow, H. H.; Vogt, C., Comparison of methods for simultaneous identification of bacterial species and determination of metabolic activity by protein-based stable isotope probing (Protein-SIP) experiments. *Rapid Commun Mass Spectrom* 2009, 23, (12), 1871-8.
  33. Shen, H. D.; Lin, W. L.; Tam, M. F.; Wang, S. R.; Tzean, S. S.; Huang, M. H.; Han, S. H., Characterization of allergens from *Penicillium oxalicum* and *P. notatum* by immunoblotting and N-terminal amino acid sequence analysis. *Clin Exp Allergy* 1999, 29, (5), 642-51.
  34. Frieri, M.; Kettelhut, B., *Food hypersensitivity and adverse reactions : a practical guide for diagnosis and management*. M. Dekker: New York, 1999.
  35. Benndorf, D.; Müller, A.; Bock, K.; Manuwald, O.; Herbarth, O.; von Bergen, M., Identification of spore allergens from the indoor mould *Aspergillus versicolor*. *Allergy* 2008, 63, (4), 454-60.
  36. Sander, I.; Flagge, A.; Merget, R.; Halder, T. M.; Meyer, H. E.; Baur, X., Identification of wheat flour allergens by means of 2-dimensional immunoblotting. *J Allergy Clin Immunol* 2001, 107, (5), 907-13.
  37. Jenkins, J. A.; Breiteneder, H.; Mills, E. N., Evolutionary distance from human homologs reflects allergenicity of animal food proteins. *The Journal of allergy and clinical immunology* 2007, 120, (6), 1399-405.
  38. Mari, A.; Schneider, P.; Wally, V.; Breitenbach, M.; Simon-Nobbe, B., Sensitization to fungi: epidemiology, comparative skin tests, and IgE reactivity of fungal extracts. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology* 2003, 33, (10), 1429-38.
  39. Cramer, R.; Faith, A.; Hemmann, S.; Jaussi, R.; Ismail, C.; Menz, G.; Blaser, K., Humoral and cell-mediated autoimmunity in allergy to *Aspergillus fumigatus*. *J Exp Med* 1996, 184, (1), 265-70.
  40. Mayer, C.; Appenzeller, U.; Seelbach, H.; Achatz, G.; Oberkofler, H.; Breitenbach, M.; Blaser, K.; Cramer, R., Humoral and cell-mediated autoimmune reactions to human acidic ribosomal P2 protein in individuals sensitized to *Aspergillus fumigatus* P2 protein. *J Exp Med* 1999, 189, (9), 1507-12.
  41. Glaser, A. G.; Limacher, A.; Fluckiger, S.; Scheynius, A.; Scapozza, L.; Cramer, R., Analysis of the cross-reactivity and of the 1.5 A crystal structure of the *Malassezia sympodialis* Mala s 6 allergen, a member of the cyclophilin pan-allergen family. *Biochem J* 2006, 396, (1), 41-9.
  42. Schmid-Grendelmeier, P.; Fluckiger, S.; Disch, R.; Trautmann, A.; Wuthrich, B.; Blaser, K.; Scheynius, A.; Cramer, R., IgE-mediated and T cell-mediated autoimmunity against manganese superoxide dismutase in atopic dermatitis. *J Allergy Clin Immunol* 2005, 115, (5), 1068-75.
  43. Chelikani, P.; Ramana, T.; Radhakrishnan, T. M., CATALASE: A REPERTOIRE OF UNUSUAL FEATURES. *Indian Journal of Clinical Biochemistry* 2005, 20, (2), 131-135.
  44. Ward, M. D.; Donohue, M. J.; Chung, Y. J.; Copeland, L. B.; Shoemaker, J. A.; Vesper, S. J.; Selgrade, M. K., Human Serum IgE Reacts with a *Metarhizium anisopliae* Fungal Catalase. *Int Arch Allergy Immunol* 2009, 150, (4), 343-351.
  45. Janeway, C., *Immunobiology 5 : the immune system in health and disease*. 5th ed.; Garland Pub.: New York, 2001; p xviii, 732 P.
  46. Mendez, J.; Sanchez, A.; Martinez, J. C., Urticaria associated with dermatophytosis. *Allergol Immunopathol (Madr)* 2002, 30, (6), 344-5.



ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v1i2.67

## Differential expression of Outer membrane proteins in early stages of meropenem-resistance in *Acinetobacter baumannii*

Jitendraa Vashist<sup>1</sup>, Vishvanath Tiwari<sup>1</sup>, Arti Kapil<sup>2</sup> and Rajeswari R Moganty<sup>1\*</sup>

<sup>1</sup>Department of Biochemistry, <sup>2</sup>Department of Microbiology, All India Institute of Medical Sciences, Ansari Nagar, 110029, India.

Received: 18 April 2011 Accepted: 13 December 2011 Available Online: 28 December 2011

### ABSTRACT

*Acinetobacter baumannii* has emerged as one of the six most important drug-resistant microbes in the world. Resistance by *A. baumannii* to  $\beta$ -lactams and in particular to meropenem is a serious concern. In this connection, it is essential to understand the changes in the outer membrane proteome of *A. baumannii* in the initial stages of resistance. For this we have chosen one low resistant strain with minimal inhibitory concentration (MIC) of 32  $\mu\text{g/ml}$  and one intermediate strain with very low MIC of 0.8  $\mu\text{g/ml}$  of meropenem and compared their outer membrane profiles with that of sensitive strain, ATCC 19606 of *A. baumannii*. Decreased expression of porins, transporters and increased production of metabolic enzymes like Succinyl-CoA synthetase, enoyl-CoA hydratase is a common feature in both intermediate strain and low resistant strains. Interestingly, the differential protein expression levels showed a direct relationship with increasing meropenem resistance. It is clear that initial exposure to meropenem resistance drives *A. baumannii* to restrict the production of CarO and transporters, while the upregulation of genes of altered CarO, metabolic enzymes, peroxidases and antioxidant protein assist in the survival of the bacterium. Because of these unique features of adaptation combined with high metabolic changes in response to antibiotic pressure, *A. baumannii* poses challenges in therapeutic strategies.

**Keywords:** *Acinetobacter baumannii*; meropenem-resistance; DIGE; Outer membrane protein .

### 1. Introduction

Bacterial infections are often treated by  $\beta$ -lactam group of antibiotics which is considered as the most effective against a number of negative bacteria including *Acinetobacter baumannii*. *A. baumannii*, a very common hospital pathogen in Intensive Care Units

(ICUs) and wards has been identified as one of the six important and highly drug resistant hospital pathogens by the "Infectious Disease Society of America" (IDSA) [1, 2]. Therefore, worldwide emergence of antibiotic resistance in *A. baumannii* poses a serious threat to human health. Presently, the latest analogs of  $\beta$ -lactam, prescribed to treat patients affected by these gram negative bacteria are meropenem and imipenem. However, increasing number of meropenem resistant isolates strongly restricts the effective therapy options [3, 4].

It is known that *A. baumannii* develops resistance using a number of ways including expression of  $\beta$ -lactamases, alte-

rations in penicillin binding proteins (PBPs) [5, 6, 7], aminoglycoside-modifying enzymes in aminoglycosides resistant strains [8] etc. Bacterial membrane proteins are known to be key molecules in maintenance the permeability and efflux of antibiotic [9]. Differential expression of membrane proteins in susceptible and highly resistant strains of *A. baumannii* from different parts of the world clearly show a strong association with the emergence of the resistance phenotype [6, 10, 11, 12, 13, 14, 15]. Our recent report has clearly shown several altered outer membrane proteins in 'high  $\beta$ -lactam resistant strain' with minimal inhibitory concentration of 64  $\mu\text{g/ml}$  of meropenem [16]. However, the profiling of outer membrane proteins in low resistant strains is not yet understood and is not defined. Keeping in view of the rise of the meropenem resistance, we have undertaken the proteomic analysis of outer membranes of *Acinetobacter baumannii* in the initial stages of the meropenem resistance.

\*Corresponding author: Rajeswari R Moganty E. Adress: Department of Biochemistry, <sup>2</sup>Department of Microbiology, All India Institute of Medical Sciences, Ansari Nagar, 110029, India; Telephone : 91-11-26593416; Fax: 91-11-26588663; E-mail: rajeshwari3011@hotmail.com

## 2. Material and methods

### 2.1 Materials

Muller Hinton agar, MacConkey agar, and Luria broth were purchased from Himedia laboratories Ltd, India and Pronadisa Conda Laboratories Canada, respectively. N-lauroyl-sarcosine and ammonium bicarbonate were from Sigma chemical co. USA. Immobilized dry strips, phar-malytes, Cy2, Cy3, Cy5 dyes, dithiothritol and iodoacetamide were purchased from GE Health Care, USA. Acrylamide, bisacrylamide, ammonium persulphate, TEMED, sodium dodecyl sulphate, EDTA, coomassie brilliant blue, and bromophenol blue were purchased from Bio-Rad laboratories, USA. 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), acetonitrile and proteomic grade water were purchased from G. Biosciences, USA. Trypsin and dimethyl formamide were from Promega (USA) and Spectrochem (India), respectively. DNase was purchased from Promega, U.S.A. Acetic acid, glycerol, methanol were purchased from Qualigens, India and all other chemical were of analytical grade and purchased from Merck, India.

### 2.2 Bacterial strains

ATCC 19606 and forty four non-repetitive clinical strains of *A. baumannii* were collected from the Department of Microbiology, All India Institute of Medical Sciences, New Delhi (India). The clinical strains were confirmed as *A. baumannii* using standard biochemical [17].

### 2.3 Minimal Inhibitory Concentration

Agar dilution method was used to estimate the minimal inhibitory concentrations (MIC) of various  $\beta$  lactams for resistant strains of *A. baumannii* [18]. Concentrations of meropenem, in the range of 0.4  $\mu$ g/ml to 64.0  $\mu$ g/ml were used in estimating the MIC. The MICs of other antibiotics i.e. piperacillin, cefotaxime and ceftazidime were done up to 512  $\mu$ g/ml. The plate without any antibiotic was inoculated in parallel, to serve as a control.

### 2.4 Extraction of Outer Membrane proteins (Omps)

Bacteria were grown in Luria-Bertani broth (250 ml) at 37°C for 19.4 hours. Doubling time of *A. baumannii* ATCC19606 is 48min hence we grow bacteria for 25 generations. After 19.4 hours culture, bacteria were harvested by centrifugation. The pellet was suspended in 50mM Tris buffer, pH 7.5. Resuspended bacterial cells were sonicated using Misonix XL Ultrasonic processor under three pulses of 1min each at 50 Hz under cold conditions. Presence of nucleic acid (if any) can create vertical streaking in 2D PAGE and to avoid such interference from nucleic acid, lysed sample is conventionally treated with DNase [13]. Hence, the suspension containing the cell envelope was treated with

DNase (10  $\mu$ g/ml, for 15 min. at room temperature) and subjected to ultracentrifugation by Beckman Optima TL Ultracentrifuge for 30 min at 100000 g. After ultracentrifugation, total membrane fraction was obtained as pellet and this fraction was treated with 2% Sodium lauroyl sarcosinate for 30 min at room temperature which specifically solubilized inner membrane. The sample was further ultra- centrifuged for 30min. at 100000 g and outer membrane was obtained as pellet. The outer membrane fraction was stored at -700 C.

### 2.5 Differential In Gel Electrophoresis (DIGE)

Native, intermediate and low resistant strains of *A. baumannii* were grown three times for constant time period (19.4 hours) and in constant culture conditions (temperature 37°C, 250 ml culture) and outer membrane proteins were isolated as above given protocol. The membrane pellets were solubilized in modified rehydration buffer (7M urea, 2M thio-urea, 2% CHAPS) at room temperature and quantified by using the 2D Quant Kit (GE Healthcare) following the manufacturer's protocol. Three sets of experiments were conducted according to our previous published experimental design [16]. In each set, proteins from native and resistant strain were labeled with the fluorescent dyes Cy 3 and Cy 5 separately [16]. An internal control of Cy 2 labeled proteins was also added in each experiment which constitutes a mixture of 1/6th protein fraction of native and 1/6th protein fraction of resistant strain.

For each labeling reaction, 50  $\mu$ g of protein was incubated with 200 pmole of dye for 30 min in dark. To stop the reaction, 1  $\mu$ l of 10 mM lysine was added and incubated for 10 min. All labeling incubations were carried out on ice. The final volume of reaction mixture was adjusted to 250  $\mu$ l using rehydration buffer (7M urea, 2M thio-urea, 0.7 mg DTT, 2% CHAPS, 1.25  $\mu$ l IPG buffer) and rehydrated for 16 hours in dark with 4-7pH immobilized pH gradient (IPG) strip.

### 2.6 Two dimensional (2D) gel electrophoresis

Isoelectric focusing of rehydrated IPG strip was carried out on Ettan IPGphor 3 IEF system (GE healthcare) using the step protocol: 150 V for 1 h, 500 V for 1 hour, 1500 V to 3 hour, 4500 V for 4 hour to a total of 24,000 volt-hours. After completion of the first dimension, strips were equilibrated for 15 minutes in 5 ml of SDS-equilibration buffer (50 mM Tris-HCl (pH 8.8), 6 M urea, 30 % glycerol, 2 % SDS, 0.02 % bromophenol blue) and 0.05 % DTT in dark at room temperature [13]. After DTT treatment, the strips were treated with 1.25 % iodoacetamide solution prepared in SDS equilibration buffer for 15 min. in dark at room temperature. Second dimension was done on a 12 % polyacrylamide gel in SE 600 Ruby gel apparatus (GE healthcare). Gel was run at 15 mA for 30 min and then at 30 mA at 4°C till the bromophenol blue came out of the gel.

### 2.7 Image Acquisition and Analysis

After second dimension, DIGE gels were scanned for Cy2, Cy3 and Cy5 fluorescence labeled proteins using a Typhoon™ TRIO Variable Mode Imager (GE Healthcare, USA). Cy2 images were scanned at an excitation of 488 nm using 520BP40 emission filter; Cy3 images were scanned at an excitation of 532 nm using 580BP30 emission filter; Cy5 images were scanned at an excitation of 633 nm using 670BP30 emission filter. All gels were scanned with a photomultiplier tube (PMT) setting of 600 volt. Images were cropped using Image-Quant™ version 6.5 (GE Healthcare, USA) to remove areas extraneous to the gel image. The final expression levels in the gels were determined by the DeCyder software version 7.0 (GE Healthcare, USA). All nine images (3 different experiments) were uploaded to the workspace using image loader module. Three differential analysis (DIA) sets were created by DIA module. In each set, protein spots on DIGE image pairs (Cy 3 and Cy 5 labeled) were co-detected automatically and each gel image was intrinsically link to its in-gel standard (Cy2 labeled image). The number of spots for co-detection procedure was set to 1500. Further, all three DIA workspaces were then imported to biological variance analysis (BVA) workspace. The experimental setup and relationship between samples were assigned in the BVA workspace. Each individual Cy3 or Cy5 gel image was assigned an experimental condition, either native or resistant according to the labeling and all Cy2 images were classified as standards; gel-to-gel matching of the standard spot maps from each gel. The gel with the highest spot count was assigned as the master gel. Matching between gels was performed utilizing the in-gel standard from each image pair. Land-marking was done by manually identifying well defined spots along with their neighboring spots and matching these spots across the standard images.

The degree of difference in standardized abundance between two protein spot groups is expressed as average ratio (fold change). A fold change with a threshold value of minimum two fold increase or decrease was used. Student t-test was performed for every matched spot-set, comparing the average and standard deviation of protein abundance for a given spot. As the threshold value of protein expression was set to minimum of 2 fold for all experiments, therefore the proteins which had more value than the threshold and had significant p value ( $\leq 0.05$ ) of t-test were considered for identification [19].

### 2.8 In-gel digestion and Electrospray ionization mass spectrometry (ESI-MS)

After visualization with coomassie G-250, protein bands were excised from the 2D-SDS polyacrylamide gel and completely destained in 100 $\mu$ l destaining solution containing 1:1 100mM NH<sub>4</sub>HCO<sub>3</sub> and 100% acetonitrile till the bands appear colourless. Finally gels were dehydrated in 100% ACN. Reduction of proteins was carried out in 100 $\mu$ l of 10mM DTT in 50mM NH<sub>4</sub>HCO<sub>3</sub> for 45 minutes at 56°C. Alkylations of proteins were done in 100 $\mu$ l of 55mM Iodoacetamide (IAA)

prepared in 50mM NH<sub>4</sub>HCO<sub>3</sub>. Again washing of gels were carried out in 100 $\mu$ l of destaining solution containing 1:1 ratio of 50mM NH<sub>4</sub>HCO<sub>3</sub> and 100% ACN. Final dehydration of gels was done in 100% ACN for 15 minutes. Gels were centrifuged and supernatants were discarded and gels were completely dried in speed Vac for 20 min. Tryptic digest was started by the addition of 20  $\mu$ l from a 12.5 ng/ $\mu$ l trypsin solution in 25 mM NH<sub>4</sub>HCO<sub>3</sub> and kept on ice for 30 min for absorption. This was followed by further addition of 20  $\mu$ l of 25mM NH<sub>4</sub>HCO<sub>3</sub> solution in the reaction sample to overlay. The samples were kept at 37°C for 16 hours for digestion. Digested peptides were solubilized in 10  $\mu$ l 50% acetonitrile containing 0.1 % formic acid. 8  $\mu$ l of the peptide solution was used to load in a silica capillary (Proxeon Biosystem, USA) which was then fixed to a QSTAR-XL QTOF mass spectrometer. The progress of each run was monitored by recording the total ion current (TIC) for positive ions as a function of time for ions in the m/z range of 400-1600 for MS and 140-1600 for MS/MS. Mass spectra were acquired using information-dependent acquisition (IDA) method. Nanospray ionization method was used with an ionspray voltage of 900. The other parameters are as follows: interface temperature = 50°C, curtain gas flow = 1.13 L/min, declustering potential 1 = 60 V, declustering potential 2 = 15 V focusing potential = 280 V.

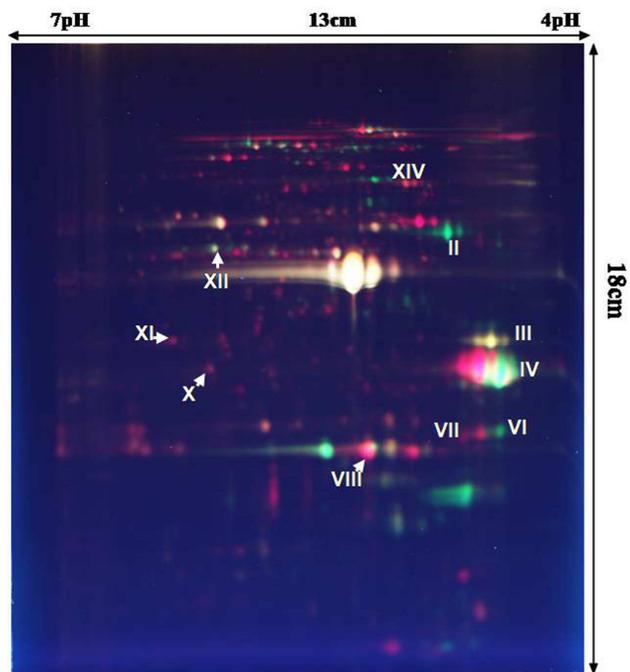
Database searching was done using Mascot (Version 1.6b4 Matrix Science, UK). Modifications considered were oxidation of methionine and carbamidomethylation of cysteine as variable and fixed manner. Search was further considered upto charged state ranging from +2 to +3 and limited to Eubacteria. The peptide mass tolerance range was  $\pm 1.0$  Da and fragment mass tolerance was  $\pm 0.3$  Da. All the Mowse score values reported are significant, p < 0.05. All spectra were searched online in NCBI nr database with 0-1 missed cleavage.

### 3. Results

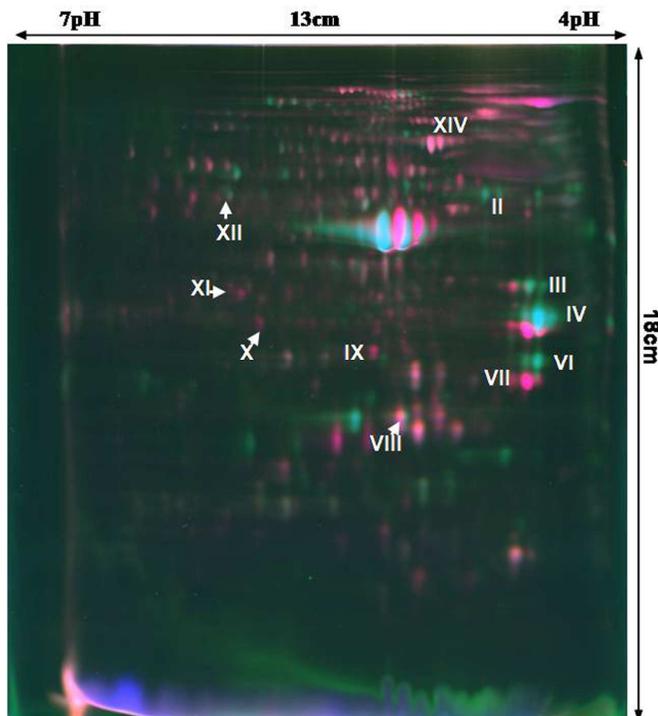
MICs of 44 clinical strains of *A. baumannii* were done for the routinely used  $\beta$ -lactams and also for the less common meropenem. Based on their MIC values, the clinical strains of *A. baumannii* were divided into two groups: intermediate (IR) strain with MIC 8.0  $\mu$ g/ml and low resistant (LR) with MIC  $\geq 16$  $\mu$ g/ml. The MIC of various  $\beta$ -lactams for the low resistant *A. baumannii* (LR122) obtained are: piperacillin and cefotaxime, 128 $\mu$ g/ml; ceftazidime, 64  $\mu$ g/ml and meropenem, 0.8  $\mu$ g/ml. However, MIC of same group of  $\beta$ -lactam for intermediate resistant strain (IR259) of *A. baumannii* showed much higher MICs as follows: piperacillin and cefotaxime,  $>512$  $\mu$ g/ml; ceftazidime, 128 $\mu$ g/ml and meropenem, 0.8 $\mu$ g/ml.

#### 3.1 Differential in gel electrophoresis (DIGE) and Mass spectrometry

Two representative strains RS 259 and RS 122 of *A. bau-*



**Figure 1a.** Differential In Gel Electrophoresis (DIGE) of outer membrane fraction of *Acinetobacter baumannii* ATCC19606 (50  $\mu$ g), intermediate strain RS 259 (50  $\mu$ g) and pooled fraction (50  $\mu$ g) using Cy3, Cy5 and Cy 2 dye, respectively. Isoelectric focusing (IEF) was done using 13 cm, 4-7 pH range IPG strip and second dimension was done on 16X18 cm gel- electrophoresis. Gels were analyzed by Typhoon Imager.



**Figure 1b.** Differential In Gel Electrophoresis (DIGE) of outer membrane fraction of *A. baumannii* ATCC19606 (50  $\mu$ g) and low resistant strain RS 122 (50  $\mu$ g) and pooled (50  $\mu$ g) using Cy3, Cy5 and Cy2 dye, respectively. The experimental conditions and labeling are same as in figure 1a.

*mannii* were chosen from intermediate and low resistant strain groups and are referred as IR and LR, respectively. 2D-DIGE profiles are given in figure 1a and figure 1b with respect to ATCC 19606.

In general, there is a high similarity of the outer membrane profiling amongst all the three strains ATCC, IR and LR (Figure 1a and b) which can be clearly seen from the superimposed images. The similarity profiles of *A. baumannii* native, intermediate and low resistant can be easily understandable as the bacteria do not go through a major shift in the protein profile in the early stages of resistance. However, a couple of new proteins in fact can be located for example like spot IX (Figure 1b). Biological variation analysis (BVA) of native and IR strain revealed 11 differentially expressed protein spots of which three were downregulated and five more were upregulated. Protein spots IV, XII and X are highlighted with \*. The proteins with non-significant expression in one strain (LR or IR) are reported as N.S. While in the case of LR strain, five proteins were down-regulated and five proteins were upregulated. Comparative differential expression of specific proteins of native ATCC is presented in Table 1.

### 3.2.1. Downregulated proteins

**CarO protein and their altered forms:** Two isoforms (protein spots IV and VI) were identified as CarO. In case of intermediate and low resistant strains, there is a clear decrease in the CarO proteins. The expression level of spot IV (29 kDa, 4.5 pI) showed a down regulation (-6.8 fold), while in case of comparison between the low resistant strain and ATCC the expression loss of the same protein had increased upto the 38 fold (Figure 1a and 1b). Similarly, down-regulation of spot VI was also noted in a sequential manner (-13 fold in IR and -57 fold in low resistant) (Table 1).

We also found the migratory differences in CarO isoforms of spot IV and VI (Figure 1a and 1b). Another major spot i.e. Spot VII was identified as the hypothetical protein of *A. baumannii* (gi|301512444) and the upregulation of spot VII is more than 5 folds in IR and LR strains (table 1, Figure 1a and 1b).

**34 kDa transporter:** Spot III showed a significant down-regulation only in low resistant strain (-6.5 fold) (Figure 1b) and later identified as 34 kDa Outer membrane protein of *A. baumannii*.

**Lipid transporter:** Protein spot II was down-regulated in both resistant strains in a similar manner. The down regulation of this protein was seen more in intermediate strain (-38 fold) as compared to low resistant strain (-22 fold) (Figure 1a and 1b). This protein was identified as outer membrane transporter protein of *A. baumannii*.

**OprD porin:** We also found spot XII as basic protein OprD homologous protein of *Pseudomonas spp.* OprD. An equal expression loss is seen in intermediate strain (-2.9 fold) and low resistant strain (-3.1 fold) (Figure 1a, 1b).

**Metabolic enzymes:** Two upregulated protein spots X and

**Table 1.** Identification of differentially expressed proteins of *Acinetobacter baumannii* intermediate strain (RS 259) and low resistant strain (RS 122) with reference to ATCC 19606. Differential expression is shown as fold change (minimum 2 fold, p-value  $\leq 0.05$ ).

SPOT No.	Mw. (IN KDA)	pI	FOLD CHANGE (IR)	FOLD CHANGE (LR)	IDENTIFICATION & (ACCESSION No.)	SCORE & % COVERAGE	FUNCTION
<b>DOWNREGULATED PROTEINS</b>							
II	50	4.9	(-) 38.9	(-) 22.9	O M transporter <i>A. baumannii</i> ATCC 19606 gi 260557248	116/12	Lipid Transport
III	34	4.7	N.S.	(-) 6.5	Putative 34 kDa Omp, <i>A. baumannii</i> ACICU gi 184159810	476/46	Non specific Transport
IV*	29	4.5	N.S.	(-) 38.6	Putative Omp, <i>A. baumannii</i> . gi 72535025	697/52	Non specific Transport
VI	26	4.5	(-) 13.6	(-) 57.9	Putative Omp, <i>A. baumannii</i> . gi 72535025	614/52	Non specific Transport
XII*	47	5.9	(-) 2.9	(-)3.1	OprD, <i>A. geno. spp.</i> 3 gi 193735465	203/11	Basic amino acid Transport
<b>UPREGULATED PROTEINS</b>							
VII	25	4.7	6.7	5.4	hypothetical protein AbauAB05_12702, <i>A. baumannii</i> AB058 gi 301512444	189/25	Transport
VIII	20	4.9	3.0	1.5	Peroxiredoxin, <i>A. baumannii</i> AB058 gi 301513596	241/40	Antioxidation
IX	24	5.1	-	2.56	Putative antioxidant protein, <i>A. baumannii</i> AYE gi 169794796	317/32	Antioxidation
X*	30	5.6	1.8	2.7	Succinyl-CoA synthetase $\alpha$ -subunit, <i>A. baumannii</i> SDF gi 169632628	442/29	Metabolism
XI	29	5.7	3.5	3.1	enoyl-CoA hydratase, phenylacetic acid degradation, <i>A. baumannii</i> AYE gi 169796408	153/13	Metabolism
XIV	57	4.9	3.8	N.S.	Chaperonin GroEL, <i>A. baumannii</i> SDF gi 169632653	362/18	Protein folding

\* Proteins were also found in high resistant strain in our earlier report [16]. Non-significant protein spots are shown by N.S.

XI were identified as enzymes integrated with energy producing reactions. Spot X had shown a sequential upregulation in both, intermediate (1.8 fold) and low resistant (2.7 fold) strains and identified as Succinyl-CoA synthetase protein [20]. Similar to Succinyl CoA synthetase, a threefold increment was seen in the expression level of enol-CoA hydratase enzyme (Protein spot XI).

**Chaperonin GroEL and antioxidant proteins:** Chaperonin GroEL protein was found in four fold elevated level in intermediate strain (spot XIV). Spot VIII protein had an upregulation of 3 fold and 1.5 fold in intermediate and low resistant strain, respectively. However, spot IX had a signifi-

cant differential expression in low resistant strain only, and intermediate had not shown the differential expression of this protein.

#### 4. Discussion

It is essential to identify the outer membrane proteins of *A. baumannii* which show their altered expression against different analogs of  $\beta$ -lactams. In the preceding years, a gradual increase in the resistance patterns of *A. baumannii* has been noticed in India as well as several other countries. The frequency of *A. baumannii* five year back was only 8-10%

while presently the presence of *A. baumannii* in our hospital has increased upto 30% (our unpublished data). MIC data clearly indicates a high resistance shown in all the 44 resistant strains for cefotaxime, piperacillin, and ceftazidime (MIC range up to 512 µg/ml). It is good to note that potency of meropenem is still high in these strains (MIC range 0.1 µg/ml to 8 µg/ml, n = 34) of *A. baumannii*. However, the resistance is slowly emerging out even to meropenem (MIC range 16-32µg/ml, n = 9). *A. baumannii* is rapidly emerging out as highly resistant organism and it will be very difficult to treat the infections as no high effective antibiotic is presently available.

#### 4.1 Down-regulation of porins as first defense in resistance

It is well evident that the intrinsic level of antibiotic resistance in gram-negative bacteria is directly regulated by porins [21]. Porins have the ability to diffuse even the large size molecules like antibiotics [22, 23]. Therefore, the bacterium, as a first step decreases the production of porins in order to control the entry of the antibiotic inside the bacterial system. This method of increasing resistance is a first line of defense and exerts a profound influence on the entry of the hydrophilic antibiotics like beta-lactams [24]. This feature has been described in majority of gram-negative bacteria like *E. coli* and *Pseudomonas*. In this category of porin, the first member is the CarO. The sequential downregulation in the expression clearly indicates the response of the antibiotic and it appears to be the initial step of the bacterium to resist against antibiotic load. CarO protein has got a special attention in the current decade and majority of the studies published in resistance mechanisms of *A. baumannii* stated its loss of expression or its alteration [10, 12, 13]. Based on the decreased expression more than 50 fold, it can be easily concluded that CarO isoform of 26 kDa is more susceptible for antibiotic load (Figure 1b). CarO isoforms is indeed responsible for nonspecific diffusion channel formation and plays important role in carbapenem resistance [12, 25].

Anomalous expression profile of 34 kDa protein is noticed in our study. Low resistant strain shows the significant down-regulation of 34 kDa protein while, neither IR strain nor high resistant strain showed a significant down-regulation. However, earlier reports suggested its potential role in β-lactam resistance [26]. A lipid transporter has been found to be downregulated in intermediate and low resistant strains. It is well known that β-lactams are hydrophilic in nature, however, the downregulation of this lipid transporter is due to resistance for antibiotics other than lactams.

OprD is a characteristic protein of porin family and was shown to be responsible for β-lactam resistance in gram negative bacteria [27, 11]. In the present study, the protein spot XII was found to be down regulated in both resistant strains (LR and IR). The spot XII was identified as OprD of *A. baumannii* (Figure 1a and 1b). However, a four fold upregulation was reported in high resistant *A. baumannii* [16]. Siroy et al. 2006 found no change in expression of OprD in multi-

drug resistant strain of *A. baumannii* [13]. We can say that there is no clear picture of direct involvement of OprD in resistance of *A. baumannii* and other gram negative bacteria because of conflicting reports [11,13,16,27]. Therefore, it may be speculated that *A. baumannii* during the initial stages of resistance, makes use several porins including OprD while in high resistance it uses other porins at the cost of OprD.

It is noticed that the down-regulation of porins is sequential and greater loss of expression has been noticed in low resistant strain as compared to intermediate strain. Therefore, it can be easily hypothesized that during the development of resistance, the bacterium uses downregulation or alterations in porins as a first step and these proteins gets downregulation in a sequential manner to combat the bactericidal effects of hydrophilic antibiotics.

Besides downregulated proteins, several Omps are upregulated as a means of adaptation of resistant bacteria in adverse conditions of antibiotic exposure. One such group of proteins is CarO isoforms.

**Altered CarO isoforms:** In LR strain CarO protein shows two isoforms with molecular weight 25 kDa and 26 kDa represented as the spot VII and spot VI, respectively. We would like to emphasize that this observation is not an artifact as the comparative DIGE study between the native and resistant bacterium was done three times and for each experiment set bacteria were grown. The anomalous migration of these proteins may be due to the change in the primary level of the protein which suggests modification in the primary role of this protein i.e. diffusion pathway may be modified. It appears that due to the meropenem load there is a modification or alteration of the internal residues of the protein spot VI which resulted in the shifting of these protein spots [13]. The presence of altered CarO isoforms is thus justifying the importance of porins in uplifting the resistance in pathogenic bacteria.

**Metabolic enzymes upregulation:** The previous comparative studies carried out between susceptible and resistant isolates of *A. baumannii* revealed high expression of metabolic enzymes and due to which a versatile metabolism is found in the resistant strains [28]. Concurrently, our data also found high expression of enzymes related to metabolism. In this regard, Succinyl-CoA synthetase and enol-CoA hydratase enzyme upregulation, clearly indicates that resistant bacteria have versatile metabolism as compared to susceptible bacterium.

Majority of the bacteria have multiple copies of the groEL gene which are active under different environmental conditions. It may be hypothesized that the antibiotic stress leads to the production of more production of this enzyme complex for proper folding of proteins in stress environment. It has been demonstrated that bactericidal antibiotics like beta-lactams induce the production of reactive oxygen species (ROS) for bacterial cell death [29]. Also, it was recently demonstrated that an up-regulation of several antioxidant proteins enhances the ability of bacteria to survive against ROS

and RNI damage [30]. The antioxidant proteins and peroxidases are evolved to combat the adverse effect of high antibiotic pressure on resistant bacteria [31]. Low resistant strain may use more of these proteins to defend itself against oxidative damage from human immune cells, such as macrophages. In fact, it supports the hypothesis that the presence of more than one differentially expressed protein influences the ability to infect and to spread in the population.

All these upregulated proteins of resistant bacteria support the fact of versatile metabolism of resistant *A. baumannii* which can survive in extreme conditions and have high adaptations as compared to the susceptible bacteria.

## 5. Conclusions

The transformation efficiency of *A. baumannii* is exceptionally high due to which it can acquire resistance in its early phases. Therefore, it is very important to monitor changes in the level of antibiotic susceptibility among clinical isolates. However, at present, it is unknown how the early phases of resistance in *A. baumannii* correspond to change in different protein expression profiles. The results described in this study permit to hypothesize that compensation for the meropenem-resistant phenotype may be responsible for the different protein expression in *A. baumannii*. Results appear to reveal that an interplay mechanism is present between antibiotic resistance, biological fitness and virulence. It can be concluded that porins, metabolic enzymes and antioxidative proteins have a potential role in the meropenem-resistance as most of the resistant strains harbor these specific proteins in *A. baumannii*.

## 6. Supplementary Material

Supplementary data and information is available at: <http://www.jiomics.com/index.php/jio/rt/suppFiles/67/0>

## Acknowledgements

We thank Department of Biotechnology India for financial assistance (BT/PR5149/MED/14/596/2004 and Clinical Proteomic facility at AIIMS). Dr. Jitendraa Vashist wishes to thank the Indian Council of Medical Research, India and Mr. Vishvanath wishes to thank the Council of scientific and industrial Research, India for providing research fellowships.

## References

- G.H. Talbot, J. Bradley, J.E. Edwards, D. Gilbert, M. Scheld, J.G. Bartlett, Clin. Infect. Dis. 42 (2006) 657-668.
- H.W. Boucher, G.H. Talbot, J.S. Bradley, J.E. Edwards, D. Gilbert, L.B. Rice, M. Scheld, B. Spellberg, J. Bartlett, Clin Infect Dis. 48 (2009) 1-12.
- L. Poirel, P. Nordmann, Clin Microbiol Infect. 12 (2006) 826-36.
- J. Vila, J. Pachon, Expert Opin. Pharmacother. 9 (2008) 587-99.
- A.Y. Peleg, H. Seifert, D.L. Paterson, Clin. Microbiol. Rev. 21 (2008) 538-582.
- F. Fernandez-Cuenca, L. Martinez-Martinez, M. C. Conejo, J.A. Ayala, E.J. Perea, A. Pascual, J. Antimicrob. Chemother. 51 (2003) 565-74.
- A. Kohlenberg, S. Brümmer, P.G. Higgins, D. Sohr, B.C. Piening, C. De Grahl, E. Halle, H. Rüdén, H. Seifert, J. Med. Microbiol. 58 (2009) 1499-1507.
- M. Van Looveren, H. Goossens, Clin. Microbiol. Infect. 10 (2004) 684-704.
- A. Kumar, H.P. Schweizer, Adv. Drug. Deliv. Rev. 57 (2005) 1486-1513.
- A.S. Limansky, M.A. Mussi, A. M. Viale, J. Clin. Microbiol. 40 (2002) 4776-4778.
- M. Dupont, J.M. Pages, D. Lafitte, A. Siroy, C. Bollet, J. Proteome Res. 4 (2005) 2386-90.
- A. Siroy, V. molle, C. lemaitre-guillier, D. vallenet, M. Pestel-Caron, A.J. Cozzone, T.Jouenne, E. Dé, Antimicrob Agents Chemother. 49 (2005) 4876-83.
- A. Siroy, P. Cosette, D. Seyer, C. Lemaitre-Guillier, D. Vallenet, A. Vandorselaer, S. Boyer-Mariotte, T. Jouenne T, E.Dé, J. Proteome Res. 5 (2006) 3385-3398.
- M. Fernandez-Reyes, M. Rodriguez-Falcon, C. Chiva, J. Pachon, D. Andreu, L. Rivas, Proteomics 9 (2009) 1632-45.
- S.H. Yun, C.W. Choi, S.O Kwon, G.W Park, K. Cho, K.H. Kwon, J.Y.Kim, J. S. Yoo, J. C. Lee, J.S.Choi, S. Kim, S. II Kim, J Proteome Res. 10 (2011) 459-469.
- J. Vashist, V. Tiwari, A. Kapil, M.R. Rajeswari, J Proteome Res. 9 (2010) 1121-1128.
- S. Constantiniu, A. Romaniuc, L.S. Iancu, R. Filimon, I. Tarași. The J of preventive medicine (2004) 12: 35-42.
- W.E. Schatten, R.F. Parker, Proc. Soc. Exp. Biol. Med. 82 (1953) 574-577.
- A. Alban, S.O. David, L. Bjorkesten, C. Andersson, E. Sloge, S. Lewis, I. Currie, Proteomics 3 (2003) 36-44.
- D. Vallenet, P. Nordmann, V. Barbe, L. Poirel, S. Mangenot, E. Bataille, C. Dossat, S. Gas, A. Kreimeyer, P. Lenoble, S. Oztas, J. Poulain, B. Segurens, C. Robert, C. Abergel, J.M. Claverie, D. Raoult, C. Médigue, J.Weissenbach, S. Cruveiller, Plos One 19, 3(2008) e1805.
- M. Viveiros, M. Dupont, L. Rodrigues, I. Couto, A. Davin-Regli, M. Martins, J.M. Page's, L. Amaral, PLoS ONE 2 (2007) e365.
- K. Jyothisri, V. Deepak, M.R. Rajeswari, FEBS Lett. 443 (1999) 57-60.
- J. Vashist, M.R. Rajeswari, J. Biomol. Struct. Dyn. 24 (2006) 243-253.
- H. Nikaido, Microbiol Mol Biol Rev. 67 (2003) 593-656.
- M.A. Mussi, A.S. Limansky, A.M. Viale, Antimicrob. Agents Chemother. 49 (2005) 1432-1440.
- M. del Mar Tomás, A. Beceiro, A. Pérez, D. Velasco, R. Moure, R. Villanueva, J. Martínez-Beltrán, G. Bou, Antimicrob Agents Chemother. 49 (2005) 5172-5175.
- M. M. Ochs, M. P. McCusker, M. Bains and R. E. W. Hancock. Antimicrob Agents Chemother 1999, 43(5):1085.
- N.C. Soares, M.Cabral, J. Parreira, C. Gayoso, M. Barba, G. Bou, Proteome Sci. 7 (2009) 37.
- M.A. Kohanski, D.J. Dwyer, B. Hayete, C.A. Lawrence, J.J. Collins, Cell 130 (2007) 797-810.
- N.C. Soares, M.P. Cabral, C. Gayoso, S. Mallo, P. Rodriguez-Velo, E. Fernandez-Moreira, G. Bou, J. Proteome Res. 9 (2010) 1951-1964.
- M. D. Adams, E. R. Chan, N. D. Molyneaux, and R.A. Bonomo, Antimicrob Agents Chemother. 54 (2010) 3569-3577.



ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v1i2.70

## Proteomic approach for molecular physiological mechanism on consecutive monoculture problems of *Rehmannia glutinosa*

Wenxiong Lin<sup>1,2\*</sup>, Changxun Fang<sup>2,#</sup>, Linkun Wu<sup>2,#</sup>, Gailing Li<sup>2,#</sup>, Zhongyi Zhang<sup>1</sup>

<sup>1</sup>Key Laboratory of Ministry of Education for Genetics, Breeding and Integrated Utilization of Crop (FAFU), Fuzhou 350002, Fujian, P. R. China; <sup>2</sup> School of Life Sciences, Fujian Agriculture and Forestry University, Fuzhou 350002, Fujian, P. R. China. # These authors contributed equally.

Received: 16 August 2011 Accepted: 13 December 2011 Available Online: 28 December 2011

### ABSTRACT

*Rehmannia glutinosa*, a famous Chinese medicinal plant, is not suitable for consecutive monoculture, because there are autotoxic metabolites excreted by its tuberous roots, which can greatly limit the plant growth and development. In this research, cultivar “Wen 85-5” *R. glutinosa* under three different cultivation modes, including the newly planted, the two-year and three-year consecutively monocultured, was used in the field test. The differential expression of leaf proteins, physiological changes and corresponding medicinal quality of tuberous roots at the early tuberous root enlargement stage were detected and compared in different years of consecutive monoculture. The results showed that consecutive monoculture resulted in decrement of chlorophyll content, photosynthetic capacity and root activity, but increases in free radicals and lipid peroxidation. Furthermore, the content of catalpol, the main medicinal ingredient in *R. glutinosa* tuberous roots was analyzed by FTIR and HPLC. The result showed that consecutive monoculture resulted in declined medicinal ingredients. Comparative proteomics analysis revealed 20 differentially expressed protein spots in response to increasing years of monoculture. Among them, ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) kinase, Rubisco, sedoheptulose-1,7-bisphosphatase related to Calvin cycle, and other proteins, i.e. proteasome, malonyl CoA-ACP transferase, antioxidases, pathogenesis-related protein and mRNA-binding protein were down-regulated with increasing years of monoculture. While energy metabolism related proteins (ATP synthase subunit  $\beta$ , ATPase, ATP-binding protein) and stress response related proteins (heat shock proteins) were up-regulated. Therefore it was concluded that consecutive monoculture of *R. glutinosa* remarkably affected the physiological reactions and induced the changes in the expression of leaf proteins, this in turn had a negative impact on the biomass and its quality of the medicinal plants.

**Keywords:** *R. glutinosa*; consecutive monoculture problem; proteomics; medicinal ingredient.

### Abbreviations

**AA-MA pathway** acetate-malonate pathway; **ACP** Acyl carrier protein; **Ci** intercellular CO<sub>2</sub> concentration; **CoA** coenzyme A; **2-DE** two-dimensional electrophoresis; **FTIR** Fourier Transform Infrared Spectroscopy; **Gs** stomata conductance; **HPLC** High-performance liquid chromatography; **MDA** malondialdehyde; **MS** Mass Spectrometry; **NP** the newly planted; **PMF** Peptide Mass Fingerprinting; **Pn** net photosynthetic rate; **RA** reduction amount of TTC; **RFW** fresh weight of roots; **Rubisco** ribulose-1,5-bisphosphate carboxylase/oxygenase; **SM** the two-year consecutively monocultured; **TM** the three-year consecutively monocultured; **TTC** Triphenyl-Tetrazolium Chloride; **Tr** transpiration rate.

### 1. Introduction

Consecutive monoculture problems refer to that when growing one crop consecutively in the same field, unfavorable results would be brought about even though under normal cultivation management, which include retardation in

plant growth, declines in crop yield, deterioration in growth conditions, and increases in pest damage, etc, [1]. The phenomenon is very common in the cultivation of Chinese medicinal plant, especially in *Rehmannia glutinosa* [1]. Accord-

\*Corresponding author: Wenxiong Lin, Agroecological Institute, Fujian Agriculture and Forestry University, Fuzhou, Fujian 350002, China. Tel: +86059183769440; Fax: +86059183769440; E-mail Address: lwx@fjau.edu.cn.

ing to statistics, about 70% medicinal plants suffer from various degrees of consecutive monoculture problems, which are adversely affecting plant yield and quality, limiting the sustainable utilization of Chinese medicinal resources, and further restricting the development of Chinese medicine industry. Previous researches suggested that there were three main reasons for plant consecutive monoculture problems, such as decline in soil fertility [2], serious pest damage caused by soil disease [3, 4], and autotoxicity of plant root exudates [5]. But many researchers believed that the latter is the resultant factors [6, 7, 8].

*R. glutinosa*, Scrophulariaceae perennial herb, is useful in many medicinal aspects, such as hemostat, diuretic, anti-inflammatory, anti-fungal, anti-radiation, and even in the treatment of liver disorders and diabetes. Thus it has been listed as one top-grade medicinal ingredient in China pharmacopoeias, and one of the highest consumed medicinal materials in China [9]. However, the consecutive monoculture problems in *R. glutinosa* are very prominent, which lead to retardation of plant growth and development, inhibition on enlargement of tuberous roots (the main medicinal part), thus decreases in yield and quality of medicinal plants. Furthermore, to overcome the obstacle, a necessary interval of 8-10 years from first cropping of *R. glutinosa* to its next cropping in the same field would really affect the corresponding production and regional economy development of *R. glutinosa* to a great degree [10].

There are many kinds of allelochemicals with autotoxicity in the *R. glutinosa* rhizospheric soil for first cropping [11]. These allelochemicals can be converted by soil microorganisms, and result in changed soil microflora and the toxicity of substance that remains in soil [12]. Thus the growth of next cropping is greatly suppressed by these autotoxic exudates from standing crop roots. Therefore, further research on self-poisoning-associated molecular mechanism of consecutively monocultured *R. glutinosa* can provide a theoretical basis for the application of ecological cultivation techniques in the regulation of *R. glutinosa* development.

## 2. Material and methods

### 2.1 Plant samples

*R. glutinosa* “Wen 85-5”, a cultivar produced largely in industry scale was used as crop material [13, 14]. In general, *R. glutinosa* was planted on April 15 and harvested on October 30. Three different kinds of soils from Jiaozuo city, Henan province, China (35°19'N, 113°51'E) were used for cultivation, including i) control soil with no *R. glutinosa* cultivation (No.1), ii) one-year cultivated soil (No.2), iii) two-year consecutively monocultured soil (No.3). The control plots were left fallow for the entire duration of the experiment (starting from April 15, 2006 to October 30, 2007). The one-year cultivated *R. glutinosa* plots were left fallow for the first year (2006) of the experiment, and then planted on April 15, 2007 and harvested on October 30, 2007. The two-year cultivated plots were planted on April 15, 2006, harvested on October 30, 2006, left fallow until April 14, 2007, replanted on April 15, 2007 and harvested on October 30, 2007. These three kinds of soils were collected at the same time (November 2, 2007). The physical and chemical properties of these three kinds of soils were shown in the Table 1. After dried at 70°C for 2 hours, the three different kinds of soils were pulverized and sieved (2mm mesh) respectively to remove rough granules and plant root residues, and then used separately for pot-cultivation of *R. glutinosa* in the greenhouse of Agroecological Institute, Fujian Agriculture and Forestry University on May 15, 2008. Each treatment had five replicates. The *R. glutinosa* plants grown in the corresponding three kinds of soils (No.1, No.2, No.3) were labeled respectively as the newly planted (NP), the two-year consecutively monocultured (SM) and the three-year consecutively monocultured (TM). Then we collected the plant samples from five replicates of each treatment at the early tuberous root enlargement stage and mixed together to make composite samples for further analysis.

### 2.2 Determination of root activity of *R. glutinosa*

Root activity could be reflected by the root dehydrogenase activity and measured by Triphenyl-Tetrazolium Chloride (TTC) method [15]. Optical density value in 485 nm ( $OD_{485}$ ) of various samples were measured and then converted to corresponding reduction amount (RA) of TTC through related formula to indicate root activity ( $mg/g \cdot h \text{ FW} = RA/$

**Table 1.** Physical and chemical properties of soils from the control, one-year and two-year cultivated plots.

Soils	OM (g/kg)	TN (g/kg)	TP (g/kg)	TK (g/kg)	AN (mg/kg)	AP (mg/kg)	AK (mg/kg)	pH
No.1	10.04c	2.46b	0.51c	1.14b	21.63b	34.27c	305.76b	13.37a
No.2	10.35b	1.49c	0.68b	1.18a	18.40c	37.37b	294.00c	13.27b
No.3	11.17a	2.81a	0.85a	1.09c	24.33a	54.20a	312.53a	13.28b

Note: No.1, control soil; No.2, one-year cultivated soil; No.3, two-year monocultured soil. OM, organic matter; TN, total nitrogen; TP, total phosphorus; TK, total potassium; AN, available nitrogen; AP, available phosphorus; AK, available potassium. Column values followed by different letters were significantly different ( $P \leq 0.05$ ,  $n = 3$ ) determined by ANOVA followed by Tucky's tests.

$RFW \cdot T$ , where,  $RFW$  = fresh weight of roots,  $T$  = incubation time.

### 2.3 Determination of main medicinal ingredients of *R. glutinosa* tuberous roots by Fourier Transform Infrared Spectroscopy (FTIR) and High-performance liquid chromatography (HPLC)

The samples were sieved separately by 100 mesh sifter (aperture 140  $\mu$ m) after ustulation and pulverization, then 1.5 mg of the sifted powders were placed in an agate mortar for rough grinding before adding 100 mg dried potassium bromide powders for further mixing and grinding. The mixture obtained above was put into the pressure machine for 1.5 min under 16 Mpa, then a piece of sample with 13 mm in diameter and 0.3 mm in thickness was got for Fourier Transform Infrared Spectroscopy analysis using AVATAR360-infrared spectrometer (Thermo Nicolet, USA).

The High-performance liquid chromatography setting parameters were as follows: chromatography column- 18 alkyl Silanes-bonding silica column, mobile phase-acetonitrile/0.1% phosphoric acid solution (volume ratio is 1/99), velocity- 1.0 mL/min, measuring wavelength- 210 nm, theoretical plates- 5000, column temperature- 25°C. Sample preparation was as follows: after drying under low pressure in 80°C for 24 hours, the sample was ground into coarse powder, into 0.4 g powders 25ml methanol was added before extracting by heating circumfluence for 1.5 h, then the liquid was filtrated, concentrated, and confected to 10 mL using mobile phase. The 10 ug/ml catalpol solution was used as standard control.

### 2.4 Determination of chlorophyll content of *R. glutinosa*

Three complete leaves per plant were collected for chlorophyll content determination by chlorophyll apparatus (SPAD-502, Konica Minolta, Japan) [16] at the early tuberous root enlargement stage of *R. glutinosa*. The determination was repeated three times, and statistic analysis of the results was carried out by DPS software (version 7.05) using one way analysis of variance (ANOVA) followed by Tucky's tests.

### 2.5 Determination of photosynthesis ability of *R. glutinosa*

In the meantime, net photosynthetic rate [ $P_n$ ,  $\mu$ mol/( $m^2 \cdot s$ )], transpiration rate [ $T_r$ , mmol/( $m^2 \cdot s$ )], stomata conductance [ $G_s$ , mmol/( $m^2 \cdot s$ )], and intercellular  $CO_2$  concentration [ $C_i$ ,  $\mu$ mol/mol] of the fourth leaf counted from center outward per plant were determined by portable photosynthesis apparatus (LI-6400, LI-COR, USA) and statistic analysis of the results was carried out by DPS software (version 7.05) using one way analysis of variance (ANOVA) followed by Tucky's tests.

### 2.6 Determination of malondialdehyde (MDA) content of *R. glutinosa* leaves

The detailed operation was conducted according to Esterbauer and Cheeseman [17]. Then  $OD_{532}$ ,  $OD_{600}$ , and  $OD_{450}$  value were respectively measured and applied to the formula [ $C$  ( $\mu$ mol/L) =  $6.45 \cdot (OD_{532} - OD_{600}) - 0.56 \cdot OD_{450}$ ], for further calculation of MDA content (nmol/g FW).

### 2.7 Protein Separation

After TCA/acetone extraction, the leaf proteins were separated through polyacrylamide gel with two-dimensional electrophoresis (2-DE) referring to O'Farrell's [18] and Wu *et al.*'s [19] methods. After electrophoresis, the gels were silver stained according to Vorum *et al.*'s [20] method, scanned with Imagescan and analyzed by ImageMaster software 5.0 (GE Healthcare, Uppsala, Sweden) to choose the differentially expressed protein spots for next in-gel digestion with Bergman *et al.*'s [21] method.

### 2.8 Mass Spectrometry (MS) analysis

The digested proteins were analyzed using an ABI 4700 Proteomics Analyzer (Applied Biosystems, Foster city, USA) mass spectrometer with TOF/TOF™ optics equipped with a 200-Hz frequency-tripled Nd: YAG laser operating at a wavelength of 355 nm. Mass spectra in the range of  $m/z$  700 to 3500 were acquired in the positive ion mode. The five most intense ions per spot were selected for subsequent MS/MS analysis. Collision-induced dissociation was performed using a collision energy of 1 keV in a collision cell with a gas pressure of  $6 \times 10^{-7}$  Torr. External mass calibration was applied using the ABI 4700 calibration mixture (Applied Biosystems, Framingham, MA, USA).

### 2.9 Database search

Based on the combined MS+MS/MS data, peptide and protein identifications were performed by searching against NCBI nr database (2009.09.15, 9694989 sequences) using GPS Explorer™ software 2.0 (Applied Biosystems, Foster city, USA) with MASCOT 2.2.03 search engine (Matrix Science Ltd, London, UK). The parameters were as follows: taxonomy was viridiplantae (Green Plants), trypsin digestion with one missed cleavage was selected, mass tolerance was 100 ppm for the precursors and 0.6 Da for the MS/MS ions, methyl methanethiosulfonate- labeled cysteines were set as fixed modifications, oxidized methionines were set as variable modifications. The self-degradation peak of trypsin and pollutants peak were removed manually before retrieval.

## 3. Results

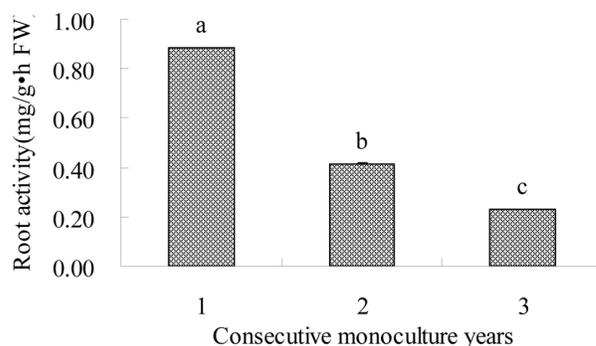
### 3.1 Root activity of *R. glutinosa*

The root activity of two-year (SM) and three-year (TM) consecutively monocultured *R. glutinosa* was significantly reduced by 53.1% and 73.9% respectively compared to the newly planted (NP) (Figure 1), suggesting that normal physiological metabolism of roots was greatly inhibited with poor performance of root activity and decline in root dehydrogenase activity under consecutive monoculture.

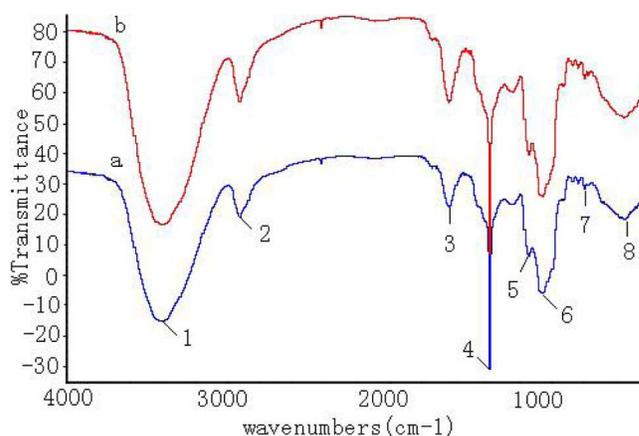
### 3.2 Analysis of main medicinal ingredients in *R. glutinosa* tuberous root

The results from FTIR analysis (Figure 2) showed that the main medicinal ingredients in terms of the species of functional groups were not significantly different between NP and SM. From FTIR detection, it also could be found that the main medicinal ingredients in *R. glutinosa* with different monoculture years shared the same chemical compositions, suggesting that consecutive monoculture had no significant effect on the main medicinal ingredients of *R. glutinosa* (Table 2).

Compared HPLC profiles of tested samples (relative concentration was 16 mg/ml) with that of catalpol standard liquid (10 µg/mL), we found that peak 3 in the retention time of 6.4 min was the characteristic peak of catalpol (Figure 3). Then the relative content of catalpol in tuberous roots of newly planted (NP) and two-year consecutively monocultured (SM) *R. glutinosa* was calculated according to that of standard catalpol liquid, with 0.88% and 0.63% (w/w) respectively, both in accordance with the content requirement mentioned in Pharmacopeia (no less than 0.2%, w/w). There were five peaks in total shared both by NP and SM, among them peak 1 was regarded as solvent peak emerging between 2 and 4 min, peak 5 showed no significant changes, the area of peak 2 and peak 4 of SM were both greater than NP, while peak 3 showed the opposite result,



**Figure 1.** The effect of consecutive monoculture on root activity of *R. glutinosa*. The numbers 1, 2 and 3 on the abscissa represent the plants from the newly planted (NP), the two-year (SM) and three-year (TM) consecutively monocultured, respectively. Different letters show significant differences at the 5% level according to its p-value by Tucky's test ( $P \leq 0.05$ ,  $n = 3$ ).

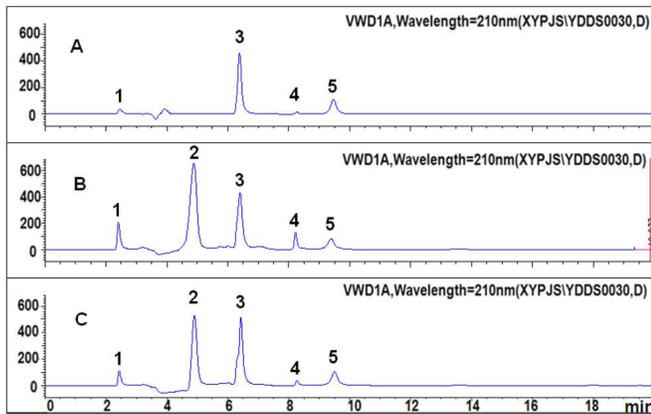


**Figure 2.** The FTIR analysis on the medicinal ingredients of newly planted (NP) and two-year monocultured (SM) *R. glutinosa* tuberous roots. a: newly planted *R. glutinosa*, b: two-year consecutively monocultured *R. glutinosa*.

**Table 2.** FTIR analysis of functional groups of medicinal ingredients in the tuberous roots of newly planted (NP) and two-year consecutively monocultured (SM) *R. glutinosa*

Peak Number <sup>a</sup>	Wavelength ( $\lambda\text{cm}^{-1}$ )		Functional group
	Newly planted (NP)	Consecutively monoculture (SM)	
1	3415.99	3407.65	—OH (association reaction)/ stretching vibration
2	2927.93	2927.93	methylenedi/ 2930、2850
3	1630.59	1643.11	C=C/ 1690~1600/ stretching vibration
4	1388.64	1388.64	methyl/ 1380
5	1142.53	1146.70	—S=O/ 1200~1040/ stretching vibration
6	1050.75	1059.10	C—O/ 1300~1000/ stretching vibration
7	796.29	796.29	C—Cl/ 800~600/ stretching vibration
8	554.35	554.35	C—Br/ 600~500/ stretching vibration

**Note:** <sup>a</sup> The number corresponds to the spectrum in Figure 2.



**Figure 3.** The HPLC analysis of the catalpol content of newly planted (NP) and two-year monocultured (SM) *R. glutinosa* tuberous roots. A: standard control, B: two-year consecutively monocultured *R. glutinosa*, C: newly planted *R. glutinosa*.

but no significant difference was shown in the sum of peak area 2, 3 and 4 between NP and SM. The result suggested that the effect of consecutive monoculture on medicinal quality of *R. glutinosa* was mainly reflected in the decline of content catalpol, a main medicinal ingredient in *R. glutinosa*.

### 3.3 Differential expression of *R. glutinosa* leaf proteins under different cultivation modes

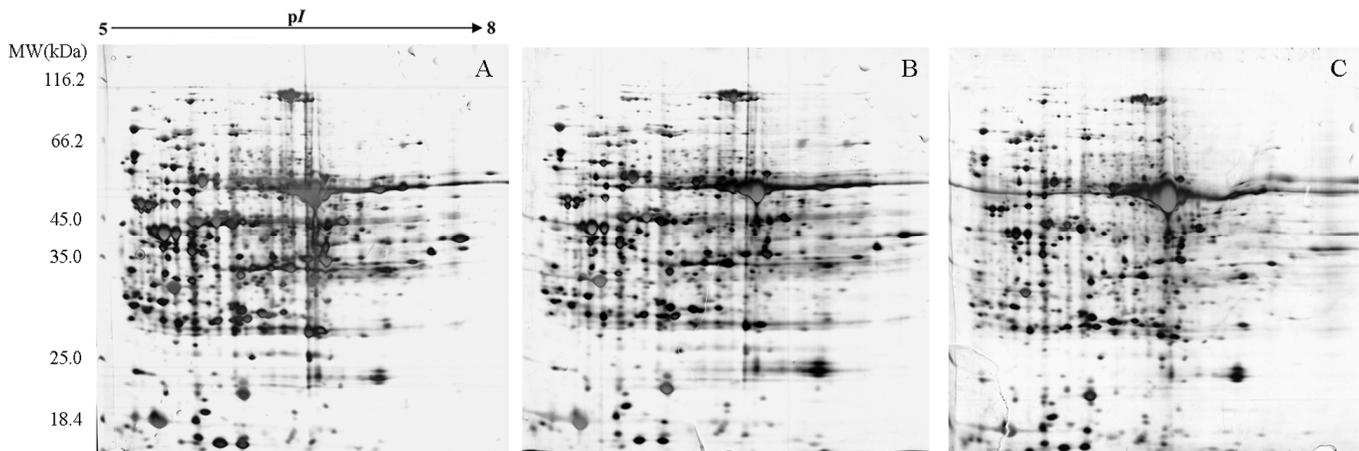
Using 2-D electrophoresis, differential expression profiles of proteins extracted from three different kinds of *R. glutinosa* leaves were established (Figure 4). Then imageMaster 5.0 software was applied to analyze the 2-D profiles to find out those proteins with differential expression (Figure 5). Finally, a total of 20 protein spots with *pI* in the range of 5- 8, MW about 14- 116 kDa, were differentially expressed, i.e., their intensities varied, at least

on one gel in comparison to the control (NP), by more than 1.5-fold. Twelve spots (spots 6, 7, 9, 10, 13, 16, 22, 24, 25, 27, 29, and 31 constituting 60% of the total differentially expressed proteins) were down-regulated with the increasing years of monoculture. Five spots (spots 4, 14, 21, 33 and 35, constituting 25% of the total differentially expressed proteins) were up-regulated with the increasing years of monoculture. Three spots (spots 12, 30 and 32, constituting 15% of the total differentially expressed proteins) were down-regulated only in the three-year monoculture, but none in the two-year monoculture.

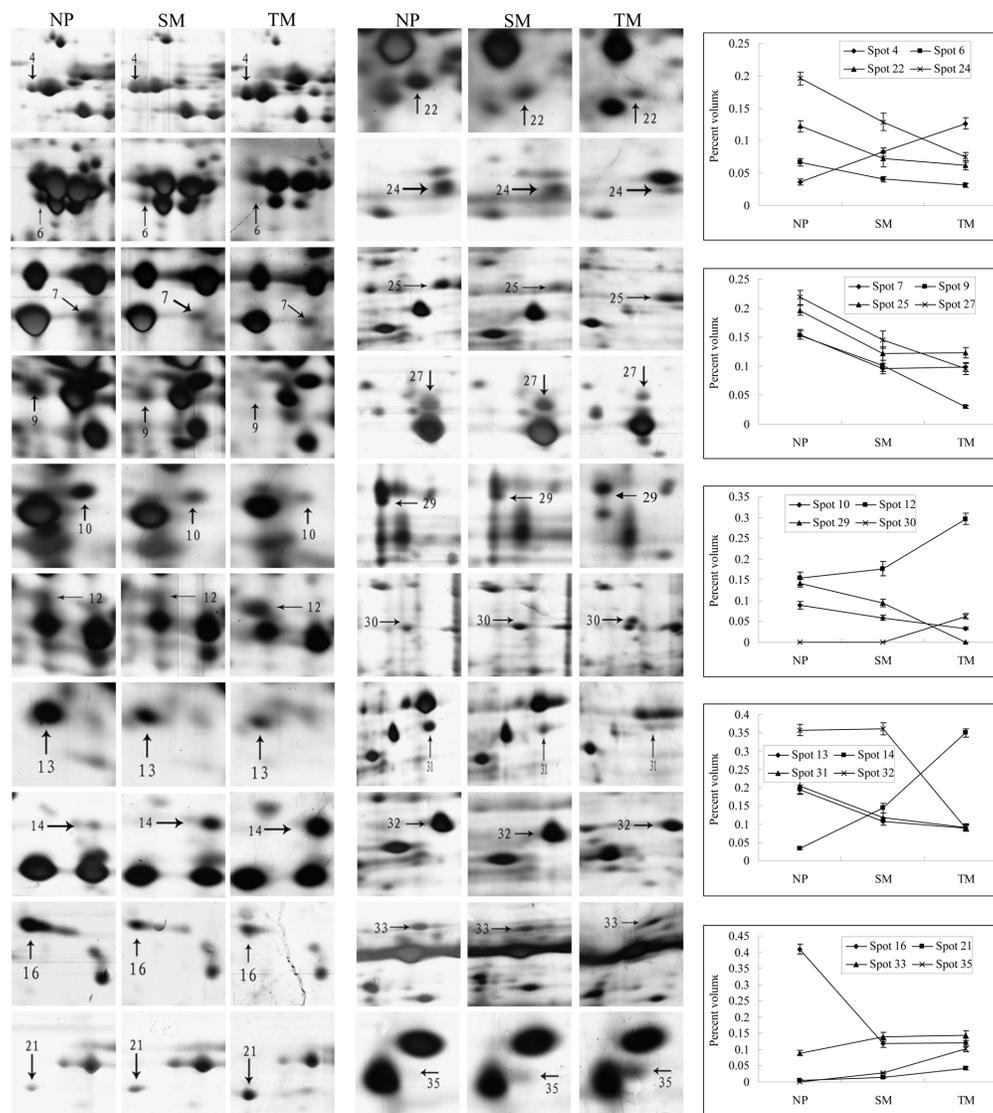
These 20 differentially expressed proteins were successfully identified by MS analysis (Table 3), and then divided into 5 classes according to their respective functions through bioinformatics searching (KEGG database, <http://www.genome.jp/kegg/>): (I) photosynthesis related proteins (spots 6, 7, 24 and 29, constituting 20%), (II) metabolism related proteins (spots 9, 10, 12, 13, 22, 25 and 32, constituting 35%), (III) energy conversion related proteins (spots 4, 21, 30 and 33, constituting 20%), (IV) resistance related proteins (spots 14, 16, 27 and 35, constituting 20%), (V) nucleotide metabolism related protein (spot 31, constituting 5%). The results showed that most proteins related to photosynthesis, mRNA binding (including the chloroplast stem-loop binding protein-41), protein degradation (including the proteasome) and pathogenesis were down-regulated with the increasing years of monoculture. However, the proteins involved in energy metabolism and stress/defense response (including the class-1 LMW heat shock protein and cytosolic class I small heat shock protein) were up-regulated with the extended monoculture.

### 3.4 Chlorophyll content of *R. glutinosa* leaves

The chlorophyll content of *R. glutinosa* leaves decreased with the increasing years of monoculture, showing that the



**Figure 4.** The silver stained 2-DE gel of proteins extracted from the leaves of newly planted (A), two-year (B) and three-year (C) consecutively monocultured *R. glutinosa*. In the first dimension, 160µg of protein was loaded on a 17 cm strip with a nonlinear gradient of pH 5-8. In the second dimension, a 5% stacking gel and a 10% separating gel were used. Unstained protein molecular weight markers ranging from 14.4 to 116.2 kDa (Promega, Madison, USA) were used in the second dimension for size standardization.



**Figure 5.** Close-up views of the differentially expressed proteins extracted from different *R. glutinosa* leaf samples in response to different-year consecutive monoculture (A) and relative abundance in terms of fold change of each spot in the gel (B). NP: proteins extracted from newly planted *R. glutinosa* leaves, SM: proteins extracted from two-year consecutively monocultured *R. glutinosa* leaves, TM: proteins extracted from three-year consecutively monocultured *R. glutinosa* leaves.

chlorophyll content of newly planted *R. glutinosa* leaves was significantly higher by 23.3% and 27.8% than those of two-year and three-year consecutively monocultured *R. glutinosa* leaves, respectively (Figure 6). The result suggested that consecutive monoculture caused the damage of photosynthesis system in *R. glutinosa* leaves, and had a great effect on the balance between chlorophyll synthesis and degradation, leading to a decline of photosynthesis capacity in plant (Figure 7).

### 3.5 Photosynthesis-associated indicators of *R. glutinosa*

As shown in Figure 7, photosynthetic rate (Pn), transpiration rate (Tr), stomata conductance (Gs) all significantly declined while intercellular CO<sub>2</sub> (Ci) increased with the increasing years of monoculture. It confirmed that photosyn-

thetic capacity of *R. glutinosa* was greatly affected by consecutive monoculture.

### 3.6 MDA content of *R. glutinosa*

As one of membrane lipid peroxidation indicators, variation of MDA content would reflect the corresponding changes in membrane lipid peroxidation and plant growth. It was shown that the MDA content of two-year (SM) and three-year (TM) consecutively monocultured *R. glutinosa* were higher than that of the newly planted (NP) by 43.1% and 73.1%, respectively (Figure 8). It suggested that membrane lipid peroxidation could be greatly intensified by consecutive monoculture. Lipid, a major component of cell membrane, is correlative to the normal intracellular metabolism to some extent, thus its peroxidation will directly lead to

**Table 3.** The differentially expressed proteins identified by MALDI-TOF/TOF-MS in *R. glutinosa* leaves under different cultivation conditions

Spot no.	Accession number <sup>a)</sup>	Protein name	Theoretical pI/ MW (kDa) <sup>b)</sup>	PMF/ Peptides <sup>c)</sup>	Score <sup>d)</sup>	Score C. I. % <sup>e)</sup>	Species <sup>f)</sup>
6	gi 13430334	Ribulose-bisphosphate carboxylase (RuBisCO) activase	6.7/37.0258	6/3	229	100	<i>Zantedeschia aethiopica</i>
7	gi 68565781	RuBisCO activase 2	6.78/48.0231	11/3	134	100	<i>Zantedeschia aethiopica</i>
24	gi 3116024	Ribulose-bisphosphate carboxylase large subunit	6.2/51.8881	12/0	70	97.326	<i>Pycreus nuerensis</i>
29	gi 5817374	Ribulose 1,5-bisphosphate carboxylase large subunit	6.57/48.6817	9/2	126	100	<i>Angelonia pubescens</i>
9	gi 115457386	Sedoheptulose 1, 7-diphosphate	5.64/42.2181	10/2	79	99.671	<i>Oryza sativa</i>
10	gi 12003283	Malonyl-CoA:ACP transacylase	5.91/39.6506	3/2	71	98.15	<i>Perilla frutescens</i>
12	gi 73808462	Putative S-adenosylmethionine decarboxylase proenzyme	5.04/37.7408	11/0	63	88.055	<i>Solanum lycopersicoides</i>
13	gi 12229923	Proteasome subunit alpha type 5 (20S)	4.7/25.9639	10/2	109	100	<i>Glycine max</i>
22	gi 46399269	Putative pyridoxine biosynthesis protein	5.93/33.068	18/2	181	100	<i>Nicotiana tabacum</i>
25	gi 82941449	Fructose-bisphosphate aldolase	6.47/38.1366	9/1	76	99.344	<i>Codonopsis lanceolata</i>
32	gi 1168408	Fructose-bisphosphate aldolase, cytoplasmic isozyme 1	6.38/38.4219	6/1	137	100	<i>Pisum sativum</i>
4	gi 15241847	ATP binding	5.03/71.3422	15/3	211	100	<i>Arabidopsis thaliana</i>
21	gi 18417676	ATP binding / ATPase/ nucleoside-triphosphatase/nucleotide binding / protein binding	5.93/108.8757	24/3	207	100	<i>Arabidopsis thaliana</i>
30	gi 115469766	UTP-glucose-1-phosphatettransferase	6.4/67.4665	6/3	70	97.617	<i>Oryza sativa</i>
33	gi 17224782	ATP synthase beta subunit	5.15/53.3608	18/1	148	100	<i>Stemona japonica</i>
14	gi 25044839	Class-1 LMW heat shock protein	6.77/17.5621	5/1	116	100	<i>Ananas comosus</i>
16	gi 38344034	OJ991214_12.15	4.67/18.1001	5/3	164	100	<i>Oryza sativa</i>
27	gi 510940	Pathogenesis related protein	7.19/16.4715	9/0	77	99.566	<i>Asparagus officinalis</i>
35	gi 37704433	Cytosolic class I small heat shock protein	5.8/14.0172	6/1	68	96.045	<i>Nicotiana tabacum</i>
31	gi 15229384	Chloroplast stem-loop binding protein-	8.54/43.9025	6/2	102	99.998	<i>Arabidopsis thaliana</i>

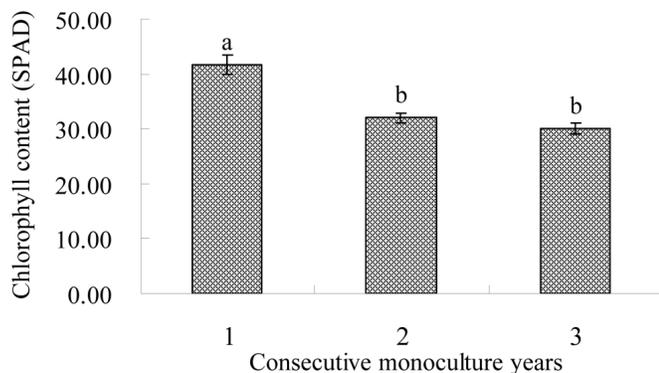
Note: <sup>a)</sup> GI number in NCBI. <sup>b)</sup> Theoretical pI and molecular weight. <sup>c)</sup> The number of peptides identified by MS or MS/MS. <sup>d)</sup> Protein scores were taken from the search results combined MS and MS/MS using GPS Explorer software (Version 2.0). In this program, a mascot score > 62 was considered significant ( $p < 0.05$ ). <sup>e)</sup> The Confidence Interval (C.I. %) for the protein score. <sup>f)</sup> The scientific name of green plants corresponding to the best-matching proteins.

a disturbance of intracellular balance and various metabolisms, and then have a further severe effect on plant's normal development. The results above confirmed that the normal development of consecutively monocultured *R. glutinosa* was suppressed compared with that of the newly planted [13], especially with the increasing years of monoculture.

#### 4. Discussion

Consecutive monoculture problems represent a major issue which greatly plagues the modern production of Chinese medicinal plant, thus the relevant research is very important for further elucidating its physiological mechanism at the molecular level. Our previous study found that vanillic

acid,  $\beta$ -sitosterol and some other second metabolites excreted by *R. glutinosa* roots were the main substances with autotoxicity and greatly limited the development of its next cropping [11]. In the present study, we found that the root activity of *R. glutinosa* decreased gradually with the increasing years of monoculture, which in turn intensified the damage degree on the whole plant growth. Furthermore, consecutive monoculture resulted in decrement of chlorophyll content, stomata conductance and photosynthesis rate, and increase in intercellular CO<sub>2</sub> concentration, which all together led to a decline of photosynthesis ability, whilst increment of MDA content and intensification of membrane lipid peroxidation in the leaves of consecutively monocultured *R. glutinosa* led to the deterioration of crops growth.

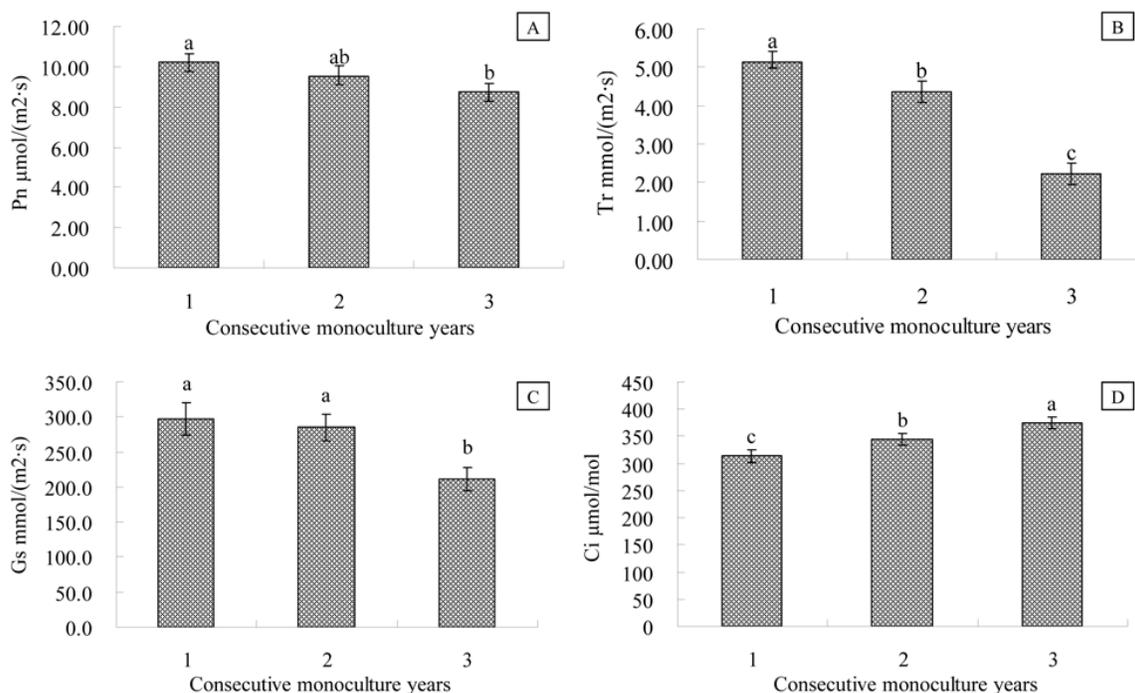


**Figure 6.** The effect of consecutive monoculture on chlorophyll content of *R. glutinosa*. The numbers 1, 2 and 3 on the abscissa represent the plants from the newly planted (NP), the two-year (SM) and three-year (TM) consecutively monocultured, respectively. Different letters show significant differences at the 5% level according to its p-value by Tucky's test ( $P \leq 0.05$ ,  $n = 3$ ).

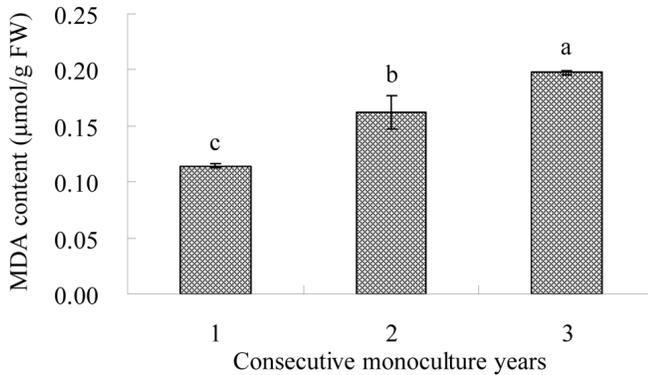
The results from comparative proteomic analysis showed that mRNA-binding protein was down-regulated in the consecutively monocultured *R. glutinosa*. Combined with some proteins, mRNA can prolong its life span and stability in the whole process from mRNA transcription to degradation though the half-life of mRNA in cell is comparatively short [22]. Downregulation of mRNA-binding showed that gene transcript process of the plant was inhibited under consecutive monoculture condition, which led to further influence on plant's physiology metabolism. Some proteins related to

Calvin cycle, such as ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) kinase, Rubisco and Sedoheptulose-1,7-bisphosphatase (SBPase), etc. were down-regulated with the extended monoculture. Down-regulation of these proteins in the consecutively monocultured *R. glutinosa* affected its Calvin cycle, leading to lack of enough substances and energy necessary for various life activities during plant normal development, further causing the consecutive monoculture problems. Previous researches reported that SBPase gene could promote crop resistance by intensifying the capacity of starch biosynthesis and sugar production after transferred into plants for overexpression [23, 24]. So a new idea has been proposed that the method can be used for overcoming the *R. glutinosa* consecutive monoculture problems. Moreover, two kinds of fructose 1, 6-bisphosphate aldolases (spots 25 and 32), both play a key role similar to SBPase in sugar synthesis, catalyze the reversible conversion of triosephosphate into fructose 1, 6-diphosphate in leaf chloroplast (starch synthesis) and cytoplasm sucrose biosynthesis [25, 26]. Therefore, inactivation of these enzymes would greatly affect the smooth process of Calvin cycle, starch synthesis in chloroplast and sucrose biosynthesis in cytoplasm, leading to retarded plant growth.

By Infrared Spectroscopy and HPLC analysis, it showed that consecutive monoculture had no significant effect on the medicinal ingredients of *R. glutinosa*, but mainly caused the quantity of main medicinal ingredient (catalpol) to decline. The increment of peaks 2, 4 and decrement of catalpol peak area (peak 3) in SM compared to the NP indicated that



**Figure 7.** The effect of consecutive monoculture on net photosynthetic rate (A), transpiration rate (B), stomata conductance (C), and inter-cellular  $\text{CO}_2$  concentration (D) of *R. glutinosa*. The numbers 1, 2 and 3 on the abscissa represent the plants from the newly planted (NP), the two-year (SM) and three-year consecutively monocultured (TM), respectively. Different letters show significant differences at the 5% level according to its p-value by Tucky's test ( $P \leq 0.05$ ,  $n = 3$ ).



**Figure 8.** The effect of consecutive monoculture on MDA content in *R. glutinosa* leaves. The numbers 1, 2 and 3 on the abscissa represent the plants from the newly planted (NP), the two-year (SM) and three-year (TM) consecutively monocultured, respectively. Different letters show significant differences at the 5% level according to its p-value by Tucky's test ( $P \leq 0.05$ ,  $n = 3$ ).

catalpol, a kind of terpenes produced by acetate-malonate pathway (AA-MA pathway) in the plant body, was changed into peak 2-related and 4-related substances in the consecutively monocultured *R. glutinosa*. In this study, malonyl CoA-ACP transferase, one key enzyme in AA-MA pathway, was also down-regulated in the consecutively monocultured *R. glutinosa*, leading the decrease in catalpol biosynthesis. The above results showed that main medicinal ingredient- catalpol in *R. glutinosa* was greatly affected by the consecutive monoculture. Catalpol is one of the main active components of *R. glutinosa* and is found in the roots, stems and leaves [27,28,29].

One down-regulated protein (spot 27), related to pathogenesis, was identified in this study. Previous researches found that a few proteins were produced when plants were infected by virus, like-virus, fungus or bacterium, which were not pathogen-specific but were dependent on the type of host reaction [30]. These proteins united together in the infected site to form a protective barrier against pathogens invasion, reduce plant susceptibility to disease. It was suggested that these proteins might have the anti-fungus or anti-bacterium activity. Thus the pathogenesis-associated protein with down-expression found in the consecutively monocultured *R. glutinosa* in this study indicated that plant defense mechanisms might be damaged because of consecutive monoculture.

Protein degradation is important for normal development of plants, and proteasome plays an important role in the degradation of unwanted proteins when damaged [31]. This kind of degradation mechanism is necessary for many kinds of cell processes, including cell cycle, regulation of gene expression, oxidative stress reaction and so on. The main role of subunit  $\alpha$  is to recognize a poly-ubiquitin chain which is connected to the target protein, and then initiates the corresponding degradation process. However, the protein (Proteasome subunit alpha type 5, spot 13) was down-regulated in consecutively monocultured *R. glutinosa*, which

resulted in weak ability of adaptation. Besides, we also found that pyridoxine biosynthesis protein isoform A (spot 22) and Malonyl CoA-ACP transacylase (spot 10) were down-regulated in consecutively monocultured *R. glutinosa*. They were linked to the biosynthesis of pyridoxine, a coenzyme of many enzymes, and fatty acid [32].

From above differentially expressed proteins, it showed that consecutive monoculture had a negative effect on *R. glutinosa* basic biosynthesis (such as the synthesis of sugar, pyridoxine, fatty acid and proteins), the degradation of proteins and so on, then affected the secondary metabolism and disease defence. Thus the growth of plant was trapped into a vicious cycle, resulting in the decline of plant defense ability against outside adverse factors such as drought, high temperature, diseases, and insect pests and so on.

However, three proteins (spots 4, 21 and 33) related to the energy metabolism, such as ATP-binding protein, ATPase and ATP synthase subunit  $\beta$ , were all up-regulated in the consecutively monocultured *R. glutinosa*. It indicated that consecutively monocultured *R. glutinosa* was indeed threatened by adversity, which needed large amounts of energy to survive. Besides, two proteins (spots 14 and 35, i.e. heat shock proteins, HSPs,) linked to stress/defense response were up-regulated in the SM and TM compared with those in the NP. They play a crucial role in protecting plant body and normal development of cell, even in protein folding, subunit assembling, intracellular transportation, protein degradation and so on [33, 34, 35]. Two kinds of HSPs were up-regulated in the consecutively monocultured *R. glutinosa*, suggesting that the plant initiated many kinds of HSPs-related physiology and energy pathways to pull through the hard times, such as ATP-synthesis system, selective degradation of ATP-dependent proteins, when it was in unfavorable conditions.

In summary, the self-poisoning substances excreted by *R. glutinosa* were accumulated in the soil through consecutive monoculture, directly or indirectly affected the substance and energy metabolism, the stress/defense response, and some other regulations at molecular level in the plant body, which are necessary or important for the normal development of next cropping, such as microbe interaction, leaf photosynthesis, root activity, adversity resistance, production of main medicinal ingredients, crop yields and so on. So investigation of consecutive monoculture problems at the molecular level will deeply uncover the relevant mechanism, which contributes to find a molecular genetic way to improve the crop yield and the production of main medicinal ingredients of *R. glutinosa*, further promotes the development of Chinese medicinal industry.

## 5. Concluding Remarks

In this study, the physiological reactions and changes in leaf protein expression of *R. glutinosa* in response to monoculture years, as well as the main medicinal ingredients (catalpol) were determined. The results showed that consec-

utive monoculture remarkably inhibited the photosynthesis rate, root activity and the contents of chlorophyll and catalpol. Further research on the comparative proteomics displayed that proteins involved in Calvin cycle (i.e. Rubisco), AA-MA pathway (i.e. malonyl CoA-ACP transferase), pathogenesis and mRNA binding (i.e. chloroplast stem-loop binding protein-41), etc. were down-regulated with the increasing years of monoculture. However, proteins related to energy metabolism (i.e. ATPase and ATP synthase) and stress/defense response (i.e. heat shock protein) were highly up-regulated with the extended monoculture, which might be a response to environmental inhibitory factors (such as toxic microbial metabolites, low pH of soil and nutrient deficiency, etc.) under consecutive monoculture. In further study, a genetic engineering method may be used to regulate the expression of above-mentioned key enzymes in plant metabolism, offering an efficient way to alleviate the consecutive monoculture problems existing in *R. glutinosa* monoculture system.

## Acknowledgements

This work was supported by grants from the National Science Foundation of China (Grant no. 30871494, 30772729) and the Natural Science Foundation of Fujian province (Grant no. 2008J0051).

## References

- Z.Y. Zhang, W.X. Lin, Chinese Journal of Eco-Agriculture. 17 (2009) 189-196. [in Chinese, English abstract.] DOI: 10.3724/SP.J.1011.2009.00189
- A. Freddy, Agron. J. 93 (2001) 556-563.
- C.H. Yang, D.E. Crowley, J.A. Menge, FEMS Microbiol. Ecol. 35 (2000) 129-136. DOI: 10.1111/j.1574-6941.2001.tb00796.x
- A. Mithufer, Trends Plant Sci. 7 (2002) 440-444. DOI:10.1016/S1360-1385(02)02336-1
- E.L. Rice, in: D.C. Washington (Ed.), Biochemical interactions among plants, National Academy of Sciences, USA, 1971, pp. 128-132.
- J.F. Du, W.J. Yin, Z.Y. Zhang, J. Hou, J. Huang, J. Li, Chinese Journal of Ecology. 28 (2009) 445-450. [in Chinese, English abstract.]
- Z.W. Wu, M.D. Wang, X.Y. Liu, H.G. Chen, X.C. Jia, Chinese Journal of Ecology. 28 (2009) 660-664. [in Chinese, English abstract.]
- W.J. Yin, J.F. Du, J. Li, Z.Y. Zhang, China Journal of Chinese Materia Medica. 34 (2009) 18-21. [in Chinese, English abstract.]
- X.S. Wen, S.L. Yang, J.H. Wei, J.H. Zheng, Chin Tradit Herbal Drugs. 33 (2002) 946-949. [in Chinese, English abstract.]
- X.E. Li, S.L. Chen, S.Q. Wei, J.H. Wei, J. Lan, China Journal of Chinese Materia Medica. 31 (2006) 344-346. [in Chinese, English abstract.]
- Z.F. Li, X.H. Qi, Q.S. Li, L.Q. Dai, M.Z. Lin, Z.Y. Zhang, W.X. Lin, Acta Ecologica Sinica. 30 (2010) 2576-2584. [in Chinese, English abstract.]
- G. Jilani, S. Mehmood, A.N. Chaudhry, I. Hassan, M. Akram, Ann. Microbiol. 58 (2008) 351-357. DOI: 10.1007/BF03175528
- L.K. Wu, H.B. Wang, Z.X. Zhang, R. Lin, Z.Y. Zhang, W.X. Lin, PLoS ONE. 6 (2011) e20611. DOI: 10.1371/journal.pone.0020611
- Y.H. Yang, X.J. Chen, J.Y. Chen, H.X. Xu, J. Li, Z.Y. Zhang, BMC Plant Biol. 11 (2011) 53. DOI: 10.1186/1471-2229-11-53
- H.S. Li, Principles and techniques of plant physiological experiment, Higher Education Press, Beijing, 2000.
- Q. Ling, W. Huang, P. Jarvis, Photosynth Res. 107 (2011) 209-214. DOI: 10.1007/s11120-010-9606-0
- H. Esterbauer, K.H. Cheeseman, Methods in Enzymology B. 186 (1990) 407-421. DOI: 10.1016/0076-6879(90)86134-H
- P.H. O'Farrell, J. Bio. Chem. 250 (1975) 4007-4021. DOI: 10.1007/978-1-60761-157-8\_2
- X.C. Wu, C.X. Fang, J.Y. Chen, Q.S. Wang, T. Chen, W.X. Lin, Z.L. Huang, J. Plant Biol. 54 (2011) 251-261. DOI: 10.1007/s12374-011-9162-y
- H. Vorum, H. Hager, B.M. Christensen, S. Nielsen, B. Honoré, Exp. Cell Res. 248 (1999) 473-481. DOI: 10.1006/excr.1999.4431
- A.C. Bergman, T. Benjamin, A. Alaiya, M. Waltham, K. Sakaguchi, B. Franzen, S. Linder, Electrophoresis. 21 (2000) 679-686. DOI: 10.1002/(SICI)1522-2683(20000201)21:3<679::AID-ELPS679>3.0.CO;2-A
- J. Nielsen, J. Christiansen, J. Lykke-Andersen, A.H. Johnsen, U.M. Wewer, F.C. Nielsen, Mol. Cell. Biol. 19 (1999) 1262-1270. PMID: PMC116055
- L. Feng, Y. Han, G. Liu, B. An, J. Yang, G. Yang, Y. Li, Y. Zhu, Funct. Plant Biol. 34 (2007) 822-834. DOI: 10.1071/FP07074
- L. Feng, K. Wang, Y. Li, Y. Tan, J. Kong, H. Li, Y. Li, Y. Zhu, Plant Cell Rep. 26 (2007) 1635-1646. DOI: 10.1007/s00299-006-0299-y
- M. Tamoi, Y. Hiramatsu, S. Nedachi, K. Otori, N. Tanabe, T. Maruta, S. Shigeoka, Photosynth. Res. 108 (2011) 15-23. DOI: 10.1007/s11120-011-9645-1
- J. Kossmann, U. Sonnewald, L. Willmitzer, Plant J. 6 (1994) 637-650. DOI: 10.1046/j.1365-313X.1994.6050637.x
- G.C. Liu, H.Q. Du, L. Liang, Chinese Traditional and Herbal Drugs. 23 (1992) 71-73. [in Chinese, English abstract.]
- C.H. Liu, G.S. Li, Y.X. Huang, Journal of Chinese Medicine Research. 14 (2001) 10-12. [in Chinese, English abstract.]
- T.X. Wang, J.Y. Li, Z.H. Hu, Chinese Traditional and Herbal Drugs. 35 (2004) 208-209. [in Chinese, English abstract.]
- F. Zhang, A. Simon, Virology. 312 (2003) 8-13. DOI: 10.1016/S0042-6822(03)00345-3
- H. Lodish, A. Berk, P. Matsudaira, C.A. Kaiser, M. Krieger, M.P. Scott, S.L. Zipursky, J. Darnell, Molecular Cell Biology, fifth ed., WH Freeman, New York, 2004.
- C. Oefner, H. Schulz, A. D'Arcy, G.E. Dale, Acta Crystallogr. D. Biol. Crystallogr. 62 (2006) 613-618. DOI: 10.1107/S0907444906009474
- S. Lindquist, Annu. Rev. Biochem. 55 (1986) 1151-1191. DOI: 10.1146/annurev.bi.55.070186.005443
- J.L. Key, C.Y. Lin, Y.M. Chen, Proc. Natl. Acad. Sci. USA. 78 (1981) 3526-3530. PMID: PMC319602
- W. Wang, B. Vinocur, O. Shoseyov, A. Altman, Trends Plant Sci. 9 (2004) 244-252. DOI: 10.1016/j.tplants.2004.03.006



ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v1i2.74

## Hemoglobin subunit beta (HBB) is a potential biomarker for predicting response to Gefitinib in NSCLC patients

Poh Kuan Chong<sup>1</sup>, Huiyin Lee<sup>2</sup>, Marie Chiew Shia Loh<sup>2</sup>, Wan-Teck Lim<sup>3</sup>, Siew Pang Chan<sup>4,5</sup>, Ross Andrew Soo<sup>2,6</sup>, Yoon Pin Lim<sup>\*1,7,8</sup>.

<sup>1</sup>Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore; <sup>2</sup>Cancer Science Institute of Singapore, National University of Singapore; <sup>3</sup>Department of Medical Oncology, National Cancer Centre of Singapore; <sup>4</sup>Division of Health Research, Faculty of Health Sciences, La Trobe University, Australia; <sup>5</sup>Stata Users Group (Singapore); <sup>6</sup>Department of Hematology/Oncology, National University Hospital, Singapore; <sup>7</sup>Department of Biological Sciences, Faculty of Science, National University of Singapore; <sup>8</sup>Bioinformatics Institute, Agency for Science, Technology and Research, Singapore.

Received: 20 October 2011 Accepted: 13 December 2011 Available Online: 30 December 2011

### ABSTRACT

EGFR mutation status has been reported to correlate well with the response of NSCLC patients to Gefitinib. However, EGFR mutation analysis is invasive in nature and recent studies supported the notion that EGFR mutation was unable to predict response to Gefitinib in some patients. We therefore conducted plasma proteomics to identify potential biomarkers that are less invasive and whose expressions correlate more significantly to response to Gefitinib. To identify protein candidates that correlate with response to Gefitinib, we profiled the relative expression levels of plasma proteins between responders and non-responders prior to Gefitinib treatment. Relative quantification of plasma proteins were analysed using Isobaric Tags for Relative and Absolute Quantification (iTRAQ) and liquid chromatography-electrospray ionization (ESI) tandem mass spectrometry. Proteins that were commonly upregulated or downregulated amongst responders but not the non-responders were selected for validation via immunoblotting. HBB protein was found to be significantly under-expressed in the plasma samples from 6 out of 7 gefitinib-responsive patients but over-expressed in a majority of the non-responders. Our finding showed that HBB is a potential biomarker for predicting response to Gefitinib that may be subject to a larger study to examine its role as a companion biomarker for Gefitinib therapy.

**Keywords:** NSCLC; Plasma; Proteomics; Gefitinib; iTRAQ; HBB

### 1. Introduction

The majority of lung cancers are non-small cell lung cancer (NSCLC) and most are diagnosed in late stage resulting in poor prognosis [1,2]. Tyrosine kinase inhibitors show great promises for cancer therapy [3]. Gefitinib (Iressa) is a highly selective tyrosine kinase inhibitor of the epidermal growth factor receptor (EGFR) that is used in the treatment of patients with advanced stage NSCLC. Initial studies have shown that gefitinib improved the quality of life in some NSCLC patients but the response rate was low [4, 5].

Subsequently, sensitizing somatic *EGFR* mutation, predominantly deletions in exon 19 and L858R point mutations in exon 21, was shown to be associated with increased sensitivity to gefitinib [6, 7]. Furthermore, in randomised studies,

the response rates of patients with sensitizing *EGFR* mutations treated with gefitinib was 62-74%, while the response rate was 1% for those without mutations [8-10]. This therefore argues for detection of *EGFR* mutations as a means to select patients for gefitinib. However, there are several challenges. First, access to tumour samples for analysis is limited. In many instances, diagnostic samples (e.g. fine needle aspirates) provide poor quality or insufficient amount after diagnostic pathology not to mention its invasive nature. Second, patients without *EGFR* mutation could still respond to gefitinib albeit at low lower rate [9]. This implies that factors other than *EGFR* mutation may predispose patients to gefitinib.

\*Corresponding author: Yoon Pin Lim. Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore. Email Address: bchlyp@nus.edu.sg

Conceivably, markers other than EGFR mutation exist that could predict response to gefitinib. Similarly, it is possible that surrogate marker(s) for EGFR mutation exists in biological fluids that are more easily accessible than tumor tissues. Blood sampling is minimally invasive and does not require repeated biopsies. The notion and feasibility of identifying markers for predicting drug response through blood-related proteomics has been demonstrated [11, 12]. Hence, we aim to profile the prior-to-treatment baseline expression level of proteins in the plasma of patients who responded to gefitinib versus those who did not.

We adopted the Isobaric Tags for Relative and Absolute Quantification (iTRAQ) approach to compare the plasma protein expression profiles of 7 responding and 6 non-responding individual patients. The objective of this pilot study is to identify candidate predictive biomarker that can be subject to larger-scale clinical trials to validate its utility in predicting response to gefitinib.

## 2. Material and Methods

### *Blood collection*

Patients with histologically or cytologically confirmed advanced stage NSCLC who were planned for gefitinib therapy were enrolled into the study. Blood collection was obtained at baseline (prior to Gefitinib treatment) and at the end of every treatment cycle. Basic patient demographics were collected. Tumour response was evaluated after every two cycles according to the RECIST criteria [13]. Disease control was defined as patients who had stable disease or better. The National University of Singapore Institutional Review Board approved the study and all patients gave written informed consent.

The protocol for blood collection and plasma preparation is based on findings from the Plasma Proteome Project (PPP) [14]. To ensure the consistency in plasma preparation, a quality control measure was implemented where all the blood samples must be processed into plasma in the laboratory within an hour after blood collection in the clinic. The integrity of each plasma sample is further verified by running 1D-SDS-PAGE and stained with SyproRuby fluorescent dye to ensure check for massive protein degradation. Samples that did not satisfy the specified time frame or integrity check would be stored away and will not be used for the study.

Samples from 13 patients were used amongst which 7 responded to gefitinib whereas the remaining 6 had disease progression following gefitinib treatment.

### *Plasma depletion and iTRAQ labeling*

One mL of plasma sample was first subject to delipidation by centrifugation at 130,000  $\times$  g at 4°C for 2 hr. The bottom layer of lipids-free plasma was collected and the total protein was estimated using BCA assay. Plasma sample from each

patient were then depleted using MARS Hu-7 affinity column (Agilent Technologies, USA) according to the manufacturer's protocol. Total protein after depletion was estimated using BCA assay. Protein samples were then reduced, alkylated, digested and labelled with iTRAQ reagents according to the recommended protocol (Applied Biosystems, Framingham, MA, USA).

### *LC-MS/MS analysis*

For each iTRAQ experiment, the labelled peptides were fractionated into 30 fractions using strong cation exchange using a PolySULFOETHYL™ A Column (PolyLC, Columbia, MD, USA) 5  $\mu$ m of 200mm length  $\times$  4.6mm ID, 200Å pore size. These fractions were cleaned-up using a C18 Discovery® DSC-18 SPE column (100mg capacity, Supelco, Sigma-Aldrich). The dried and cleaned fractions were then analyzed using Agilent 1100 nLC system (Agilent) coupled online to a quadruple time of flight mass spectrometer (QStar XL, Applied Biosystems), as described in reference [15]. Eluent from the reverse phase nLC was directly subjected to positive ion nanoflow electrospray analysis in an information dependant acquisition mode (IDA). A ToF MS survey scan was acquired ( $m/z$  370-1600, 0.5 sec), with the 3 most intense multiple charged ions (counts >70) sequentially subjected to MS/MS analysis. The time of summation of MS/MS events was set to be 2 seconds in the mass range of  $m/z$  100-1600.

Similarly to the previous study, protein identification and quantification were carried out using ProteinPilot™ software (version 2.0; Applied Biosystems, MDS-Sciex), searching against IPI human database (version 3.41) [15]. The search was performed using Paragon Algorithm™, which is discussed in detail elsewhere [16]. Only those proteins identified and quantified with at least 95% confidence were taken into account. All results were then exported into Excel for manual data interpretation.

### *Immunoblotting*

For validation of iTRAQ result, plasma samples used for iTRAQ analysis were subjected to immunoblotting for HBB and SAA using mouse monoclonal antibody (Santa Cruz), as described in previous studies [17, 18]. Triplicates blots were carried out for each sample to ensure that the generated data are reliable. For densitometry, images from x-ray film were first captured using Imager Scanner and its corresponding software LabScan version 5.0 (GE Healthcare). Image was then analyzed using ImageQuantTL software v2003.03 (GE Healthcare).

### *EGFR mutation analysis*

Formalin-fixed, paraffin-embedded tumour samples of the cases were obtained from the Pathology departments of the two participating centres i.e. National University Hospital and National Cancer Centre, Republic of Singapore. DNA

was extracted from 5 µm sections of each sample as described previously [19]. Mutations in exons 18–21 of *EGFR* were detected by direct sequencing as described previously [20].

*Statistical analysis*

To ascertain if there was a significant difference in the levels of HBB protein expressions (obtained from iTRAQ) between the responsive and non-responsive patients, the Mann-Whitney U test was applied. The Fisher exact tests were also performed to determine if there was significant association between *EGFR* mutation (0: Non wild-type, 1: Wild-type) and the clinical and demographic variables (e.g., gender, ethnicity, smoking status, histology, disease status and response towards treatment). The nonparametric tests were proposed in anticipation of the violation of normality assumption and cell sparsity owing to the small sample size. Using SPSS 16.0 (SPSS Inc, Illinois, USA) and Stata 10.0 (Stata Corp, Texas, USA), all statistical tests were conducted at 5% level of significance.

**3. Results and discussion**

*Proteome profiling of plasma samples from responders and non-responders before gefitinib treatment*

The characteristics of the 13 patients recruited for this study is shown in Table 1. Briefly, 7 responded to gefitinib

treatment and 6 did not. All the patients, of Asian origin, were diagnosed to suffer from stage IV diseases. They were mostly never smokers (84.6%) and had adenocarcinoma (69.2%). Six patients (46%) received gefitinib as first-line, 4 (31%) as second-line and 3 (23%) as third-line treatment. Mutation analysis for *EGFR* mutations was performed in ten patients in whom tumour tissues were available. *EGFR* mutations were detected in three of these patients.

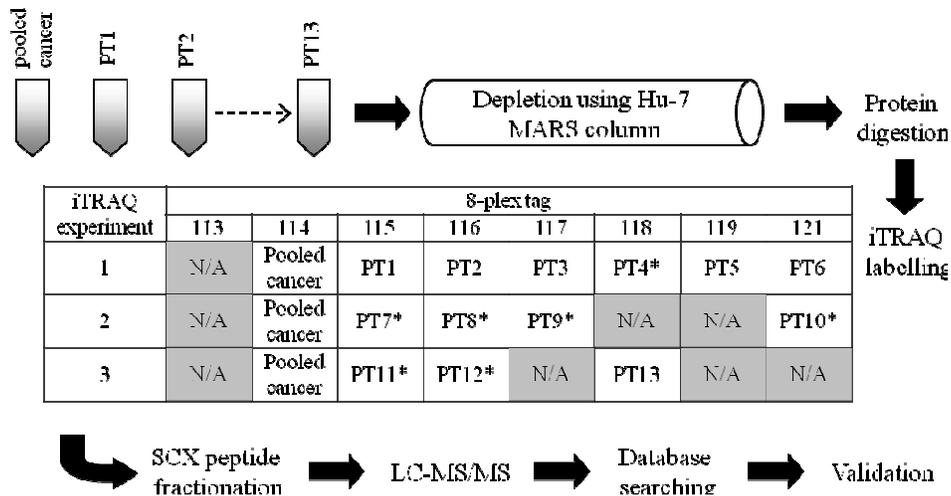
The plasma protein expression profile of individual patients before gefitinib treatment was analyzed in triplicates using quantitative iTRAQ approach. The experimental design showing how all 13 independent samples were analyzed using 8-plex iTRAQ is shown in Figure 1. Baseline plasma samples from all 13 patients were pooled and used as an internal control so that data could be normalized and compared across all the samples. Hence the relative quantification for each individual samples (7 responding and 6 non-responding individual patients) was relative to this pooled sample.

Only common proteins identified across the 3 iTRAQ datasets generated were selected for further consideration. A total of 115 common proteins were identified and quantified. The complete list of the proteins identified in these 3 iTRAQ datasets and their common proteins are tabulated in Supplementary Material 1. The 115 common proteins were then characterised based on their molecular function using Panther Classification (<http://www.pantherdb.org>) [21]. There were 177 protein assignments were obtained and sorted into 23 molecular classifications, since some of these proteins

**Table 1.** Sample Characteristics (including clinical data) of NSCLC patients

Patient ID	Age	Gender	Race	Smoking status	Histology	Disease Stage	Line of gefitinib treatment	Response	EGFR mutation status †
PT1	59	M	Malay	Never	Adenocarcinoma	IV	1 <sup>st</sup>	No	NA
PT2	55	F	Chinese	Never	Poorly differentiated	IV	1 <sup>st</sup>	No	NA
PT3	62	F	Chinese	Never	Adenocarcinoma	IV	1 <sup>st</sup>	No	exon 21, 858 point MT
PT4	69	M	Chinese	unknown	Large cell carcinoma	IV	2 <sup>nd</sup>	Yes	NA
PT5	75	M	Chinese	Former	Squamous cell	IV	3 <sup>rd</sup>	No	WT
PT6	51	F	Chinese	Never	Adenocarcinoma	IV	3 <sup>rd</sup>	No	exon 19 del 746 mt
PT7	64	F	Chinese	Never	Adenocarcinoma	IV	1 <sup>st</sup>	Yes	WT
PT8	55	M	Chinese	Never	Adenocarcinoma	IV	1 <sup>st</sup>	Yes	WT
PT9	61	F	Chinese	Never	Adenocarcinoma	IV	2 <sup>nd</sup>	Yes	Inconclusive
PT10	62	F	Chinese	Never	Adenocarcinoma	IV	2 <sup>nd</sup>	Yes	WT
PT11	57	F	Chinese	Never	Adenocarcinoma	IV	2 <sup>nd</sup>	Yes	WT
PT12	64	F	Chinese	Never	Poorly differentiated	IV	3 <sup>rd</sup>	Yes	exon 21, 858 point MT
PT13	60	F	Chinese	Never	Adenocarcinoma	IV	1 <sup>st</sup>	No	WT

† NA: No available tissue for *EGFR* analysis; WT: Wild type; Inconclusive : The analysed result was not conclusive.



**Figure 1.** The overview of experimental workflow employed in the study. A total of 3 8-plex iTRAQ datasets containing the baseline (before gefitinib treatment) plasma proteome profiles of the responders versus non-responders patients towards gefitinib treatment were generated. Patients who responded positively towards gefitinib treatment were denoted with \* in the figure.

have multiple classifications assigned (Supplementary Material 2). The majority of the common proteins were grouped under defence/immunity protein followed by protease and transport/carrier protein – typical of proteins found in the plasma. The details protein classification for each group is shown in Supplementary Material 2.

*Selection of candidate proteins that may segregate responders from non-responders*

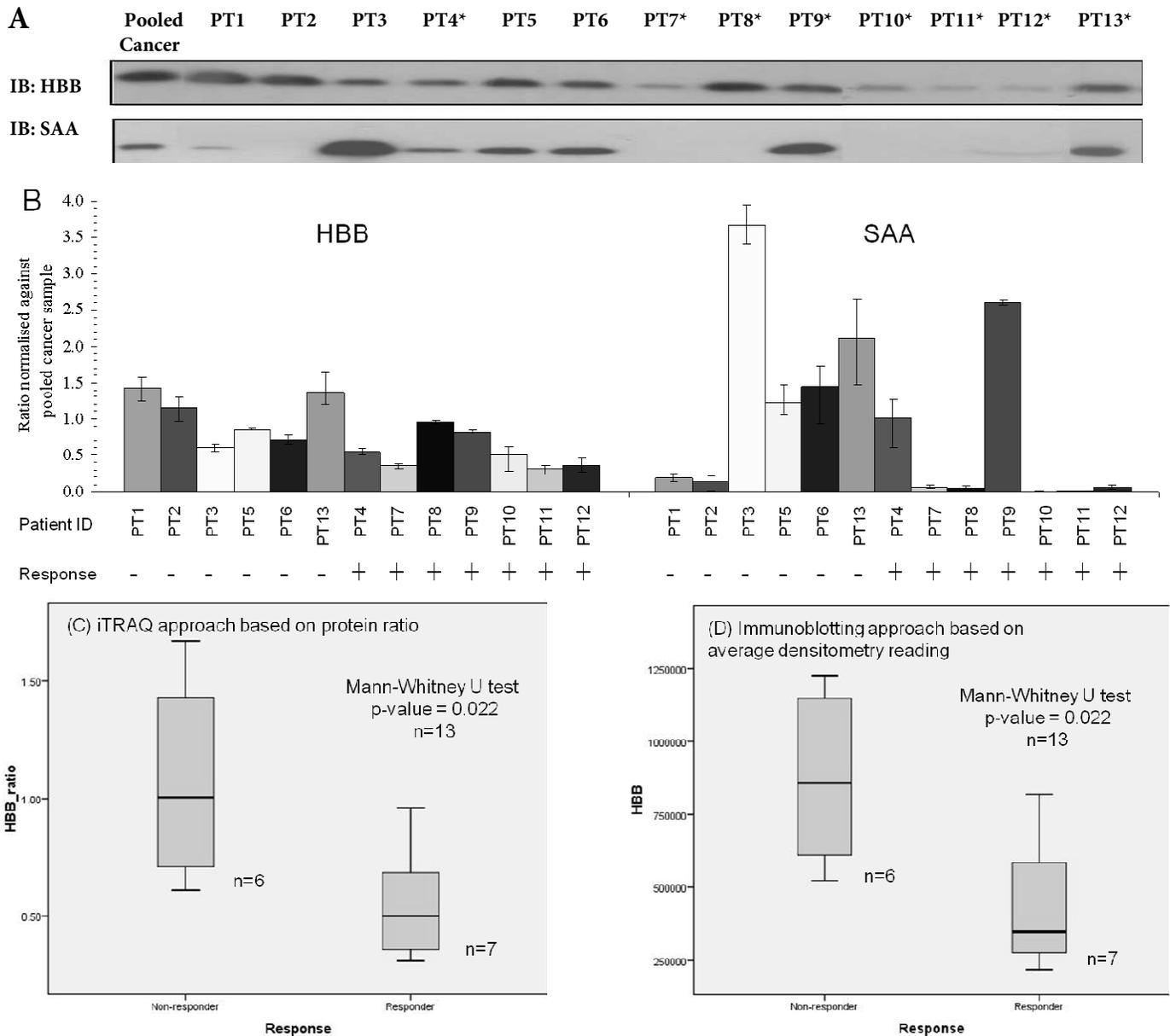
From the list of common proteins identified, proteins were considered up or down-regulated when their ratios were

>1.3 or <0.77, with their p-value <0.05. This cut-off value was determined from our preliminary analysis, which showed that the technical variation from duplicate sets of iTRAQ experiments was less than 30% (Supplementary Material 3). This technical variation has been consistently obtained in published and unpublished studies from our lab [22, 23]. The biological variations of protein expressions in the plasma of different subjects vary from one protein to another. Some are highly variable while others are not so. As such, we could only implement a cut-off based on technical variation, something that we can determine.

To help short-list proteins that might segregate responders

**Table 2.** Proteins that exhibited the same expression trend (over or under-expression) in at least 60% of the responder or non-responder group of patients. Proteins were considered differentially expressed (**bold**) when the protein ratio was found to be either >1.3 or <0.77, with p-value <0.05. Other details including p-value and error factor for these proteins can be tabulated in Supplementary Material 1.

Gene symbol	Protein name	Responsive patients						Non-responsive patients						
		PT4 : Pooled cancer	PT7 : Pooled cancer	PT8 : Pooled cancer	PT9 : Pooled cancer	PT10 : Pooled cancer	PT11 : Pooled cancer	PT12 : Pooled cancer	PT1 : Pooled cancer	PT2 : Pooled cancer	PT3 : Pooled cancer	PT5 : Pooled cancer	PT6 : Pooled cancer	PT13 : Pooled cancer
SAA2	Serum amyloid A2 isoform a	<b>0.44</b>	<b>0.56</b>	<b>0.27</b>	<b>1.81</b>	<b>0.20</b>	<b>0.75</b>	<b>0.63</b>	0.80	<b>0.44</b>	1.14	1.17	0.94	<b>0.77</b>
HBB	Hemoglobin subunit beta	<b>0.58</b>	<b>0.34</b>	0.97	<b>0.76</b>	<b>0.33</b>	<b>0.25</b>	<b>0.35</b>	<b>1.80</b>	<b>1.38</b>	<b>0.59</b>	<b>1.31</b>	0.91	<b>3.11</b>
LOC400682	Similar to hCG199685 8	<b>1.31</b>	<b>0.42</b>	<b>0.57</b>	0.78	<b>0.31</b>	<b>0.56</b>	0.89	0.87	1.15	1.21	<b>1.86</b>	<b>1.30</b>	1.25
FGB	Fibrinogen beta chain	<b>1.35</b>	<b>1.30</b>	<b>0.72</b>	<b>0.58</b>	<b>1.67</b>	<b>1.34</b>	1.07	0.96	1.28	<b>0.70</b>	1.12	0.84	<b>0.50</b>
F12	Coagulation factor XII	<b>1.46</b>	<b>1.33</b>	1.05	1.22	<b>1.39</b>	<b>1.32</b>	0.80	0.97	1.04	<b>0.77</b>	1.02	0.92	1.00
SER-PINA1	Isoform 1 of Alpha-1-antitrypsin	<b>1.52</b>	<b>0.31</b>	<b>0.47</b>	<b>0.69</b>	<b>0.72</b>	<b>1.63</b>	1.06	<b>0.72</b>	<b>1.34</b>	<b>1.72</b>	0.84	1.12	<b>0.24</b>
LUM	Lumican	<b>1.60</b>	<b>1.69</b>	0.80	0.96	<b>1.37</b>	<b>1.54</b>	1.27	0.82	<b>0.70</b>	<b>0.69</b>	<b>0.69</b>	0.91	<b>0.56</b>



**Figure 2.** (A) Representative immunoblots for both HBB and SAA protein validation. Patients who responded positively towards gefitinib treatment were denoted with \* in the figure. (B) Average densitometry readings of triplicate immunoblots for individual NSCLC patients in the study. Box plots showing the distribution of HBB protein expression obtained through (C) iTRAQ and (D) immunoblotting approaches among the responsive and non-responsive patients. Mann-Whitney U tests revealed that statistically significant difference (p-value<0.001) in the HBB expression level was observed for both plots.

from non responders, we first grouped the patients into responders or non responders. Next, we selected proteins that displayed the same expression trend (over- or under-expression) in at least 60% of the sample size within the responder or non responder group. This resulted in a total of 7 candidate proteins and they are presented in Table 2. Other proteins showed random expression trend across the responder and non-responders group. Hence they were not considered further in this study.

*Verification of HBB and SAA expression levels in the plasma*

*samples from responding and non-responding patients*

From Table 2, 2 proteins namely SAA and HBB were found to be under-expressed in 6 out of 7 patients (86%) who responded to gefitinib treatment. This is a higher frequency compared to the rest like LOC400682 and LUM, which were significantly under-expressed and over-expressed in 4 out of 7 cases, respectively. The striking differences in the expression levels of SAA and HBB between the responders and non-responder led us to investigate these two candidates more closely as they might represent bi-

omarkers that predict response to gefitinib. The rest of the candidates like LOC400682 and LUM could not be pursued further in part because of the lack of commercially available antibodies or because the antibodies were of poor quality that did not allow us to interpret the data accurately and with high confidence.

We first set out to examine the protein expression levels of SAA and HBB in the plasma using immunoblotting to determine whether the iTRAQ data could be verified. Figure 2A shows a representative blot each from immunoblotting of SAA and HBB. The densitometry readings for these 2 candidate proteins obtained through immunoblotting were shown in Figure 2B. Overall, the results obtained for both proteins using immunoblotting and iTRAQ approaches were congruent.

Next, statistical analysis was performed to examine whether the divergence in protein expressions of both SAA and HBB among the responsive and non-responsive patients was statistically significant. Using Mann-Whitney U test, the p-values for HBB protein expression level obtained through iTRAQ and immunoblotting were both 0.022 respectively. Figure 2C and 2D show the box plots of HBB protein expression obtained from iTRAQ and immunoblotting approach, respectively. The data indicates that the difference in HBB expression levels between responsive and non-responsive group of patients was significant. In contrast, no statistical significance was obtained for SAA protein (p-value > 0.05) (refer to Supplementary Material 4). Although the remaining 5 candidate proteins were not studied further in this study for reasons discussed above, the results of the statistical analyses of their expression levels between the responsive and non-responsive group of patients were nevertheless included in Supplementary Material 4 as a reference.

*EGFR mutation status and response to Gefitinib in the cohort studied*

Of the ten patients with EGFR mutation analysis performed, six were EGFR wild-type. Four out of 6 patients (67%) with EGFR WT responded to gefitinib. This was unexpectedly high given the response rate in patients EGFR WT to gefitinib can be as low as 1-3% [8, 24] and even higher range reported is about 38% [25]. Gefitinib has been shown to be more effective in never smokers amongst Asian NSCLC patients[26]. The fact that 5 out of 6 patients with no EGFR mutation in this study were never smokers could have contributed to the higher response rate to gefitinib. On the other hand, we observed only 1 out of 3 patients (33%) with EGFR mutation who responded to gefitinib. This is lower than the reported range of up to 75%. The discrepancy may result from the small sample size. It may also be due to the fact that 1 and 2 patients with EGFR mutation received gefitinib as a first line and third line treatment, respectively. So far, most of the studies on the correlation of EGFR mutation with response to gefitinib were conducted involving the use of gefitinib as 3-rd line treatment. These confounding factors

should be considered when designing future larger scale studies. It may also signal the need to discover population-specific biomarkers for predicting response of NSCLC patients to gefitinib therapy. Thus far, we are not aware that such a study has been conducted in Singapore.

To determine whether there exist a candidate surrogate marker for EGFR mutation, statistical correlation analysis between EGFR mutations and relative protein expression data was performed. The results of further assessments (Fisher exact test Mann-Whitney U test) concerning EGFR (0: Non wild-type including an inconclusive case, 1: Wild-type) are shown in Table 3. EGFR mutations status was not significantly associated with gender, age, smoking status, histology and response towards gefitinib treatment. No statistical test could be performed for disease status (all stage IV) and ethnicity (all Chinese). We caution that it is difficult to generalize the reported results of this pilot study owing to the small sample size. However, such a pilot study and the results, although preliminary in nature, are useful for planning a complete study involving more observations.

**4. Discussion**

There is sufficient evidence showing that not all NSCLC patients respond to gefitinib. As such, there is a need for

**Table 3.** Statistical analysis of EGFR mutation status.

Variable	MGFR Mutation Status		p-value
	Wild-type	Non wild-type	
Response towards gefitinib treatment			
No	2	2	0.999
Yes	4	2	
Histology			
Adenocarcinoma	5	3	0.667
Large cell carcinoma	1	0	
Poorly differentiated	0	1	
Disease stage			
IV	6	4	N.A.
Ethnicity			
Chinese	6	4	N.A.
Gender			
Female	4	4	0.467
Male	2	0	
Smoking status			
Ex-smoker	1	0	0.999
Non-smoker	5	4	
Age (years)	Median: 61 Range: 55-75	Median: 61.5 Range: 51-64	0.830

† Based on Mann-Whitney U test.

companion biomarkers to stratify patients to achieve cost effectiveness in cancer management. To this end, proteomics analysis of lung adenocarcinoma tissues from patients who showed different response to gefitinib was reported. Using a support vector machine algorithm, 9 proteins were selected that could distinguish responders from non responders. Differential expression of one of the nine proteins, heart-type fatty acid-binding protein was successfully validated [27]. While useful, the potential biomarkers identified via analysis of tissues means that invasive surgical methods have to be employed if these biomarkers were to be exploited. Biomarkers are best tested in body fluids like blood that are minimally invasive.

With respect to this, one study conducted serum protein profiling and obtained an algorithm based on 8 distinct mass peaks that could predict the outcomes following gefitinib treatment. In one of their datasets, the median survival in the predicted “good” and “poor” outcome groups was 207 and 92 days, respectively [28]. The identities of the 8 mass peaks are not known. A similar study also generated an algorithm of MALDI-mass peaks from serum profiling that could distinguish patients with different response to treatment with a combination of gefitinib and rofecoxib in platinum-pre-treated NSCLC patients [29]. While potentially useful, the use of such “biomarkers” represented by mass peaks with unknown identifies is risky since it is not clear whether they are specific to the sample preparation or analytical method used.

It is likely that an effective “biomarker” would consist of a panel of easy-to-access biomarkers with increased sensitivity and specificity than individual biomarkers. By combining various biomarkers identified from various studies, one may eventually be able to test such a combination of biomarkers for predicting drug response. To contribute to this cause, our study attempted to identify potential biomarkers that could distinguish between patients who respond versus those who don't respond to gefitinib. This was achieved by profiling the plasma proteins from 7 responders and 6 non-responders to gefitinib. There are limitations in this study. First, our study has a higher proportion of females (9 out of 13) than males (4 out of 13). However, the adjustment of gender imbalance is not feasible given the small sample size of this pilot study. A multivariate statistical technique which allows such demographic imbalances to be adjusted should be considered in future studies involving more observations. Second, the sample size of the study is small. Since this was meant to be a pilot study, we did not perform sample size calculation prior to data collection. The generated results, based on exploratory statistical techniques, would be useful for calculating the appropriate sample size for a complete study later.

Nonetheless, we have identified two proteins that were differentially expressed in gefitinib-responding and non-responding NSCLC patients were discovered. One of them is Hemoglobin subunit beta (HBB), one of the two polypeptide chains in adult haemoglobin. It plays an important role in oxygen transportation from lung to various peripheral tis-

ues. Various studies focused on the mutation of HBB since its defect can lead to numerous blood disorder diseases such as beta-thalassemia and sickle cell anemia [30]. On the other hand, HBB gene expression level was reported to be down-regulated in breast tumour tissue [31] and anaplastic thyroid cancer cell lines [32] compared with normal tissue/cells. In a glycoproteomic study on human lung adenocarcinoma tissue, HBB was also reported to be down-regulated compared to the normal tissue [33]. In this study, we showed that HBB protein expression was low in the majority of the plasma samples of NSCLC patients who were responsive to gefitinib treatment compared with the non-responders. This implies that HBB may be used to predict patient's response to gefitinib. It is not entirely clear how a lower HBB expression may influence drug response. It is conceivable that a lower HBB expression results in a hypoxic condition and that hypoxic tumours, which are metabolically stressed, may be more susceptible to gefitinib. While it is possible that HBB may be cancer cell-specific, we do not rule out the possibility that lower amount of HBB might be related to physiological processes such the oxygen carrying capacity of the erythrocytes.

## 5. Conclusion

In conclusion, our pilot study shows that HBB is a candidate discriminatory biomarker that is able to segregate between gefitinib-responding and non-responding NSCLC patients. The key finding from this study justifies the design of a future study with a larger sample size to validate the potential of HBB as a predictive biomarker for gefitinib therapy. Due to the expensive nature of Gefitinib, accrual of patients on this drug for future study will be a challenge. It is envisaged that such a large-scale validation study is best undertaken by the industry or a major cancer program/consortium.

## 6. Supplementary Material

Supplementary data and information is available at: <http://www.jiomics.com/index.php/jio/rt/suppFiles/74/0>

*Supplementary Material 1* - List of common proteins identified from 3 iTRAQ dataset generated from 13 NSCLC baseline plasma samples (before Iressa treatment); *Supplementary Material 2* - Molecular classification using Panther; *Supplementary Material 3* - Technical variation estimation in iTRAQ labeling; *Supplementary Material 4* - Statistical analysis of protein expression between responsive and non-responsive group of patients.

## Acknowledgements

This work was supported by funding from the Singapore Cancer Syndicate, Agency for Science, Technology and Research (A\*STAR). Dr Chong PK was supported by the Lee

Kuan Yew Endowment Funds.

## References

- O.S. Breathnach, B. Freidlin, B. Conley, M.R. Green, D.H. Johnson, D.R. Gandara, M. O'Connell, F.A. Shepherd, B.E. Johnson, *J Clin Oncol* 19 (2001) 1734-1742.
- N. van Zandwijk, G. Giaccone, *Curr Opin Oncol* 8 (1996) 120-125.
- L.K. Shawver, D. Slamon, A. Ullrich, *Cancer Cell* 1 (2002) 117-123.
- M. Fukuoka, S. Yano, G. Giaccone, T. Tamura, K. Nakagawa, J.Y. Douillard, Y. Nishiwaki, J. Vansteenkiste, S. Kudoh, D. Rischin, et al, *J Clin Oncol* 21 (2003) 2237-2246.
- M.G. Kris, R.B. Natale, R.S. Herbst, T.J. Lynch, Jr., D. Prager, C.P. Belani, J.H. Schiller, K. Kelly, H. Spirodonidis, A. Sandler, et al, *Jama* 290 (2003) 2149-2158.
- T.J. Lynch, D.W. Bell, R. Sordella, S. Gurubhagavata, R.A. Okimoto, B.W. Brannigan, P.L. Harris, S.M. Haserlat, J.G. Supko, F.G. Haluska, et al, *N Engl J Med* 350 (2004) 2129-2139.
- J.G. Paez, P.A. Janne, J.C. Lee, S. Tracy, H. Greulich, S. Gabriel, P. Herman, F.J. Kaye, N. Lindeman, T.J. Boggon, et al, *Science* 304 (2004) 1497-1500.
- T.S. Mok, Y.L. Wu, S. Thongprasert, C.H. Yang, D.T. Chu, N. Saijo, P. Sunpaweravong, B. Han, B. Margono, Y. Ichinose, et al, *N Engl J Med* 361 (2009) 947-957.
- T. Mitsudomi, S. Morita, Y. Yatabe, S. Negoro, I. Okamoto, J. Tsurutani, T. Seto, M. Satouchi, H. Tada, T. Hirashima, et al, *Lancet Oncol* 11 (2010) 121-128.
- M. Maemondo, A. Inoue, K. Kobayashi, S. Sugawara, S. Oizumi, H. Isobe, A. Gemma, M. Harada, H. Yoshizawa, I. Kinoshita, et al, *N Engl J Med* 362 (2010) 2380-2388.
- L. Pusztai, B.W. Gregory, K.A. Baggerly, B. Peng, J. Koomen, H.M. Kuerer, F.J. Esteva, W.F. Symmans, P. Wagner, G.N. Hortobagyi, et al, *Cancer* 100 (2004) 1814-1822.
- Z. Xiao, B.T. Luke, G. Izmirlian, A. Umar, P.M. Lynch, R.K. Phillips, S. Patterson, T.P. Conrads, T.D. Veenstra, P. Greenwald, et al, *Cancer Res* 64 (2004) 2904-2909.
- P. Therasse, S.G. Arbuck, E.A. Eisenhauer, J. Wanders, R.S. Kaplan, L. Rubinstein, J. Verweij, M. Van Glabbeke, A.T. van Oosterom, M.C. Christian, S.G. Gwyther, *J Natl Cancer Inst* 92 (2000) 205-216.
- G.S. Omenn, D.J. States, M. Adamski, T.W. Blackwell, R. Menon, H. Hermjakob, R. Apweiler, B.B. Haab, R.J. Simpson, J.S. Eddes, et al, *Proteomics* 5 (2005) 3226-3245.
- S. Lim, L.Y. Choong, C.P. Kuan, C. Yunhao, Y.P. Lim, *J Proteome Res* 8 (2009) 4062-4076.
- I.V. Shilov, S.L. Seymour, A.A. Patel, A. Loboda, W.H. Tang, S.P. Keating, C.L. Hunter, L.M. Nuwaysir, D.A. Schaeffer, *Mol Cell Proteomics* 6 (2007) 1638-1655.
- Y.P. Lim, L.S. Diong, R. Qi, B.J. Druker, R.J. Epstein, *Mol Cancer Ther* 2 (2003) 1369-1377.
- Y.P. Lim, B.C. Low, J. Lim, E.S. Wong, G.R. Guy, *J Biol Chem* 274 (1999) 19025-19034.
- R. Soong, B.J. Iacopetta, *Mod Pathol* 10 (1997) 252-258.
- T.M. Chin, D. Anuar, R. Soo, M. Salto-Tellez, W.Q. Li, B. Ahmad, S.C. Lee, B.C. Goh, K. Kawakami, A. Segal, et al, *Clin Chem* 53 (2007) 62-70.
- P.D. Thomas, M.J. Campbell, A. Kejariwal, H. Mi, B. Karlak, R. Daverman, K. Diemer, A. Muruganujan, A. Narechania, *Genome Res* 13 (2003) 2129-2141.
- Y. Chen, L.Y. Choong, Q. Lin, R. Philp, C.H. Wong, B.K. Ang, Y.L. Tan, M.C. Loh, C.L. Hew, N. Shah, et al, *Mol Cell Proteomics* 6 (2007) 2072-2087.
- P.K. Chong, H. Lee, J. Zhou, S.C. Liu, M.C. Loh, T.T. Wang, S.P. Chan, D.T. Smoot, H. Ashktorab, J.B. So, et al, *J Proteome Res* (in press) (2010).
- F.R. Hirsch, M. Varella-Garcia, P.A. Bunn, Jr., W.A. Franklin, R. Dziadziuszko, N. Thatcher, A. Chang, P. Parikh, J.R. Pereira, T. Ciuleanu, et al, *J Clin Oncol* 24 (2006) 5034-5042.
- G.Y. Ku, B.A. Haaland, G. de Lima Lopes, Jr., *Lung Cancer* 74 (2011) 469-473.
- S.T. Lim, E.H. Wong, K.L. Chuah, S.S. Leong, W.T. Lim, M.H. Tay, C.K. Toh, E.H. Tan, *Br J Cancer* 93 (2005) 23-28.
- T. Okano, T. Kondo, K. Fujii, T. Nishimura, T. Takanashi, Y. Ohe, K. Tsuta, Y. Matsuno, A. Gemma, H. Kato, et al, *Clin Cancer Res* 13 (2007) 799-805.
- F. Taguchi, B. Solomon, V. Gregorc, H. Roder, R. Gray, K. Kasahara, M. Nishio, J. Brahmer, A. Spreafico, V. Ludovini, et al, *J Natl Cancer Inst* 99 (2007) 838-846.
- K.J. O'Byrne, S. Danson, D. Dunlop, N. Botwood, F. Taguchi, D. Carbone, M. Ranson, *J Clin Oncol* 25 (2007) 3266-3273.
- G. Lettre, V.G. Sankaran, M.A. Bezerra, A.S. Araujo, M. Uda, S. Sanna, A. Cao, D. Schlessinger, F.F. Costa, J.N. Hirschhorn, S.H. Orkin, *Proc Natl Acad Sci U S A* 105 (2008) 11869-11874.
- I.B. Pau Ni, Z. Zakaria, R. Muhammad, N. Abdullah, N. Ibrahim, N. Aina Emran, N. Hisham Abdullah, S.N. Syed Hussain, *Pathol Res Pract* 206 (2010) 223-228.
- M. Onda, J. Akaishi, S. Asaka, J. Okamoto, S. Miyamoto, K. Mizutani, A. Yoshida, K. Ito, M. Emi, *Br J Cancer* 92 (2005) 2216-2224.
- J.H. Rho, M.H. Roehrl, J.Y. Wang, *Protein J* 28 (2009) 148-160.
- D.H. Liu, X.M. Wang, L.J. Zhang, S.W. Dai, L.Y. Liu, J.F. Liu, S.S. Wu, S.Y. Yang, S. Fu, X.Y. Xiao, D.C. He, *Biomed Environ Sci* 20 (2007) 33-40.
- S. Dai, X. Wang, L. Liu, J. Liu, S. Wu, L. Huang, X. Xiao, D. He, *Sci China C Life Sci* 50 (2007) 305-311.
- N. Khan, C.J. Cromer, M. Campa, E.F. Patz, Jr., *Cancer* 101 (2004) 379-384.
- W.M. Gao, R. Kuick, R.P. Orzechowski, D.E. Misek, J. Qiu, A.K. Greenberg, W.N. Rom, D.E. Brenner, G.S. Omenn, B.B. Haab, S.M. Hanash, *BMC Cancer* 5 (2005) 110.
- W.C. Cho, T.T. Yip, C. Yip, V. Yip, V. Thulasiraman, R.K. Ngan, T.T. Yip, W.H. Lau, J.S. Au, S.C. Law, et al, *Clin Cancer Res* 10 (2004) 43-52.