Dystrophin and the Dystrophin Associated Glycoprotein Complex in the Zebrafish, *Danio rerio*

Stephen Peter Chambers

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Biological Sciences, The University of Auckland, August 2004.

Abstract

Mutations affecting the dystrophin gene and the dystrophin associated glycoprotein complex (DGC) genes in humans cause chronic muscle wasting diseases collectively termed muscular dystrophies (MDs), often with associated cardiomyopathy. The dystrophin gene is associated with the allelic, X-linked diseases Duchenne and Becker muscular dystrophy (D/BMD), of which DMD is fatal in early adulthood. The DGC is associated with the autosomal recessive limb girdle muscular dystrophies (LGMDs) which result from mutations of the human α-, β-, δ- and γ-sarcoglycan genes. Dystrophin and the DGC genes encode a sarcolemmal complex of proteins that forms the structural linkage between the muscle cytoskeleton and extracellular matrix, which is crucial for the maintenance of muscle integrity during normal activity. Dystrophin and the DGC genes are evolutionarily conserved, underpinning the important role of this complex.

In terms of DMD modeling, the mouse has served as a suitable vertebrate species but the pathophysiology of the disease in the mouse does not entirely mimic human DMD. LGMD modeling is served by a the BIO 14.6 hamster which is a traditional model of cardiomyopathy, and mouse sarcoglycan gene null strains. The zebrafish is an established model of vertebrate development, and is receiving increasing attention in terms of human disease modeling. This thesis sought to examine the potential of the zebrafish as a further comparative model for these MDs.

In order to provide experimental support to realise this modeling potential, this thesis describes the identification of apparent orthologues of many critical members of the DGC in the zebrafish. An apparent zebrafish orthologue of the human dystrophin gene was identified that expresses a 400kDa protein that is localised to the myofibre membrane surface of adult zebrafish muscle. Epitope mapping confirmed that zebrafish dystrophin retains all the conserved functional domains reported in the human protein. Dystrophin expression in the zebrafish embryo was observed at six hours post fertilisation (hpf), although maternal transcript was present in the zygote. Dystrophin protein was localised at the myosepta of zebrafish embryos at 24 hours post fertilisation (hpf).

Apparent zebrafish orthologues of human β-dystroglycan and β-, δ-, ε-, γ- and ζ-sarcoglycan were also identified, and in the case of β-dystroglycan and β-, and γ-sarcoglycan were shown to

express proteins localised to the myofibre membrane surface in adult zebrafish. Like dystrophin, zebrafish β-dystroglycan and β-sarcoglycan genes express proteins which localised to the myosepta of 24 hpf zebrafish embryos. Co-localisation of dystrophin with β-dystroglycan and F-actin confirmed that the zebrafish carries an intact DGC.

The DGC in the zebrafish retina was also studied to provide a basis for understanding the diversity of DGC expression and function in respect of the pathology of the human diseases modelled. A DGC which incorporates zebrafish utrophin or dystrophin with dystroglycan, was localised to the glial cell layer (GCL) and outer plexiform layer (OPL) of the adult zebrafish retina. During zebrafish retinal development, dystrophin and dystroglycan were first localised in the GCL at 48 hpf and the OPL at 72 hpf.

In terms of targeting gene expression to achieve disease modeling outcomes, two strategies were used targeting the carboxyl-terminus of zebrafish dystrophin transcript. Peptide nucleic acid (PNA) masking of the splice donor site of zebrafish dystrophin exon 71 resulted in mis-splicing of the transcript which conserved the reading frame of the transcript. Gene silencing by short interfering RNAs targeted against exon 53 and exon 68 of zebrafish dystrophin caused temporary loss of dystrophin transcript, and delayed dystrophin and β-sarcoglycan localisation at the myosepta of affected embryos.

Together, these data suggest that dystrophin and the DGC in zebrafish may play a highly conserved functional role in muscle architecture that, when disrupted, could offer insight into human neuromuscular disease processes.

Key words: disease modeling, dystroglycan, dystrophin, dystrophin-associated glycoprotein complex, sarcoglycans, muscular dystrophy, peptide nucleic acids, RNA interference, zebrafish

Acknowledgements

My first and heart-felt thank you goes to the closest partners in this endeavour, Professor Don Love, who managed to keep me going till the end, and my partner Lee who lived through it with me and is still there. No less so the many supporters who have helped in uncounted ways, but the closest collaboration and support in this work came from Andrew Dodd, I thank you too. This work would have been impossible without the collaborations of Professor Glenn Morris and Dr. Louise Anderson who both gifted antibodies and to whom I am deeply thankful. My sincere thank you also to Dr. David Bissett for his gift of utrophin antibodies and for discussion of his pre-publication results. Thank you also to Professor Colin Green, Dr. Shiva Reddy and Dr. David Crosssman who all helped this study with gifts of reagents and generous advice.

For your friendly advice and endless support; Dr Shane Lavery, Dr Adrian Turner, Beryl Davey, Dr David Saul and Dr Terry Gruijters, thank you all. For your special support with confocal microscopy, Hillary Holloway and Jaqui Ross of the Biomedical Imaging Research Unit, thank you too. Giselle Maguire, Rupert Overall and Tamara Sirey were also part of this work, thanks guys. Of course there are all the other people who keep one's life and sanity together by their friendship, who are so many that to list them all would endanger the rainforests, so my apologies if I appear to have missed anyone. Special thanks to Steve F and Marty; "Stevomix" Corporation and saved by the Lion; also Bro Rich, Michelle and Flynn; Steve P. and family; Paul and Megan, Richie M., Gus and all their familys; the Mahia "boys", and all the rest of you.

The sequence data used for gene predictions were produced by the Danio rerio Sequencing Group at the Sanger Institute and can be obtained from ftp://ftp.sanger.ac.uk/pub/zebrafish.

Grants from The University of Auckland Research Committee, the Maurice and Phyllis Paykel Trust and the Lottery Grants Board of New Zealand financially supported this work.

Table of contents

List of figures and tables

List of abbreviations

The most frequently used abbreviations in this work include:

2MeAO	2'-O-Methyl-antisense oligonucleotides
AAV	adeno associated virus
Amp	ampicillin
AP	anterior-posterior (axis)
ATP	adenosine triphosphate
bp	base pairs
BMD	Becker muscular dystrophy
BSA	bovine serum albumin
Ca^{++}	calcium ion
CK	creatine kinase
CMD	congenital muscular dystrophy
DGC	dystrophin associated glycoprotein complex
DMD	Duchenne muscular dystrophy
DMSO	dimethyl sulphoxide
DNA	deoxyribose nucleic acid
dNTP	deoxynucleoside triphosphates
dpf	days post fertilisation
DRP1	dystrophin-related protein 1 (utrophin)
DRP2	dystrophin-related protein 2
DTT	dithiothreitol
E. coli	*Escherichia coli*
EST	expressed sequence tag
EF	embryonic fissure
g	gram
GCL	glial cell layer
GZ	germinal zone
hpf	hours post fertilisation
IgG	immunoglobulin G (also G_1, G_{2a}, etc)
ILM	inner limiting membrane
INL	inner nuclear layer
IPL	inner plexiform layer

kb	kilobases
L	litres
LGMD	limb girdle muscular dystrophy
M	molar
MD	muscular dystrophy
mg	milligram
ml	millilitre
mM	millimolar
NEWI	North East Wales Institute
ng	nanogram
nl	nanolitre
OFL	optic fibre layer
oligo	oligonucleotide
ONL	outer nuclear layer
OPL	outer plexiform layer
PAGE	ployacrylamide gel electrophoresis
PE	pigmented epithelium
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
pH	potential hydrogen
PNA	peptide nucleic acid
Q-PCR	quantitative PCR
RNA	ribose nucleic acid
rpm	revolutions per minute
RT-PCR	reverse transcription PCR
siRNA	short interfering RNA
SDS	sodium dodecyl sulphate
SGCA	α-sarcoglycan
SGCB	β-sarcoglycan
SGCD	δ-sarcoglycan
SGCE	ε-sarcoglycan
SGCG	γ-sarcoglycan
SGCZ	ζ-sarcoglycan

TAE	Tris-acetate-EDTA
TBE	Tris-borate-EDTA
Tween 20	polyoxyethylene sorbitan monolaureate
Tet	tetracycline
U	units
µg	microgram
µL	microlitre
µm	micrometre (micron)
µM	micromolar
v/v	volume by volume
w/v	weight by volume