

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

The cDNA Sequence and Polymorphism of
Horse Transferrin

A thesis presented in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy in Veterinary Clinical Science
at Massey University, Palmerston North,
New Zealand.

Margaret Ann Carpenter

1992

ABSTRACT

Transferrin, the serum protein responsible for iron transport, is highly polymorphic in many species, including the horse. In this study, the cDNA sequence of horse transferrin was determined and used to identify sequence polymorphisms which distinguish some of the many variants of horse transferrin.

A horse liver cDNA library was constructed, and was screened using the human transferrin cDNA as a probe. The clones isolated gave 1800 bp of sequence. The remainder of the cDNA was obtained using PCR. The 2305 bp horse transferrin cDNA sequence included part of the 5' untranslated region and extended to the poly(A) tail. It had 80% sequence identity with the human transferrin cDNA, and encoded a protein of 706 residues, including a signal sequence of 19 amino acids. The amino acid sequence was compared to those of related proteins, i.e. the serum transferrins and lactoferrins of several species and human melanotransferrin. The horse transferrin sequence had the duplicated structure and conserved iron binding and cysteine residues which are characteristic of the transferrin family. This is consistent with the structural and functional similarities within the family. Horse transferrin has 73% amino acid sequence identity to human transferrin, 62% identity with human lactoferrin, and 53% identity with chicken transferrin.

Sequence polymorphisms distinguishing the variants of horse transferrin which are found in thoroughbreds (D, F₁, F₂, H₂, O, R and *) were identified. First, comparison of the horse transferrin cDNA and protein sequences with two reported amino acid substitutions distinguishing the D and R variants indicated that exons 12 and 15 were likely to be polymorphic. Therefore, these two regions were analysed by Southern blotting and by sequencing PCR products. The D and R variants differed by 10 nucleotide substitutions which encoded 6 amino acid substitutions. The F₁, F₂, H₂ and * variants were identical to D, and the O variant was very similar to R, in the regions studied. The data indicated that the horse transferrin variants comprise two distinct groups. Secondly, the positions of differences between the D and F₁ alleles were established by producing single stranded conformation polymorphisms. Sequencing then revealed 3 nucleotide substitutions that encode 2 amino acid substitutions. All 8 of the amino acid substitutions which were observed occurred at positions which are variable between members of the transferrin family. Location of the polymorphic residues on the 3-dimensional structure of human lactoferrin revealed that all were clustered at one end of the C-lobe.

ACKNOWLEDGEMENTS

I wish to thank Dr. Tom Broad (DSIR) for making this project possible and for providing unfailing support throughout, and also Dr. Eric Terzaghi (Department of Microbiology and Genetics, Massey University) and Dr. Brian Goulden (Department of Veterinary Clinical Sciences, Massey University) for their advice and encouragement. I gratefully acknowledge the Biotechnology Division of DSIR, DSIR Grasslands and the New Zealand Pastoral Agriculture Research Institute, Ltd. for providing facilities and funding.

Thanks to Dr. Ross MacGillivray, University of British Columbia, for providing the human transferrin cDNA; Dr. Ian Anderson, Bruce Farndale and other staff of the Equine Blood Typing and Research Centre, Massey University, for blood typing; Dr. John Tweedie, Department of Chemistry and Biochemistry, Massey University, for guidance with RNA preparation and providing pGEM plasmids; and Dr. Ted Baker, Department of Chemistry and Biochemistry, Massey University, for providing a photograph of human lactoferrin.

I am grateful to members of the Gene Mapping Unit of the N.Z. Pastoral Agriculture Research Institute (formerly of DSIR Grasslands and Biotechnology Division, DSIR), particularly Pauline Lewis, Sin Hen Phua and Andrea Ede for their advice and company in the lab. Thanks also to other staff of DSIR, such as John Forrest and Bruce Davidson for providing horse genomic DNA; Nick Ellison for providing the Clustal V package, λ gt10 and *E. coli* C600 and C600 *hflA*; Peter Spring for photographic services; and Paul Ealing for synthesis of PCR primers.

Thanks to friends, flatmates and family who listened, sympathised, supported and tried to understand, especially Tim and Grant.

TABLE OF CONTENTS

Abstract.....	ii
Acknowledgements.....	iii
Table of Contents.....	iv
List of Figures.....	ix
List of Tables.....	xi
Abbreviations.....	xii
CHAPTER 1: INTRODUCTION.....	1
1.1 FUNCTIONS OF TRANSFERRIN	1
1.1.1 Iron transport	1
1.1.2 Growth factor	3
1.1.3 Bacteriostasis	4
1.2 IRON BINDING	4
1.3 OCCURRENCE, SYNTHESIS AND DEGRADATION	5
1.4 STRUCTURE OF TRANSFERRIN	6
1.5 HETEROGENEITY OF TRANSFERRIN	10
1.5.1 Variation in the polypeptide chain	10
1.5.2 Glycosylation	11
1.5.3 Iron saturation	12
1.6 DISORDERS OF TRANSFERRIN FUNCTION	12
1.7 RELATED PROTEINS	14
1.7.1 Lactoferrin	14
1.7.2 Ovotransferrin	17
1.7.3 Melanotransferrin	17
1.7.4 Rat hemiferrin	18
1.7.5 B lym transforming gene proteins	18
1.7.6 Invertebrate transferrins	19
1.8 TRANSFERRIN GENES	19
1.8.1 Human transferrin	19
1.8.2 Chicken transferrin/ovotransferrin	20

1.8.3 Mouse lactoferrin	21
1.8.4 Chromosomal location of transferrin family genes	22
1.8.5 Transferrin gene regulation	22
1.9 EVOLUTION OF THE TRANSFERRIN FAMILY	24
1.10 HORSE TRANSFERRIN	28
1.11 AIMS OF THIS STUDY	32
CHAPTER 2: METHODS AND MATERIALS	33
2.1 CONSTRUCTION OF A HORSE LIVER cDNA LIBRARY	33
2.1.1 Tissue collection and storage	33
2.1.2 Isolation of RNA from horse liver	33
2.1.3 Preparation of poly(A)+ RNA	34
2.1.4 Northern blotting	35
2.1.5 cDNA synthesis	36
2.1.6 Methylation of cDNA	37
2.1.7 Ligation of <i>Eco</i> RI linkers to cDNA	38
2.1.8 Removal of linker fragments from cDNA	38
2.1.9 Preparation of λ gt10 vector	39
2.1.10 Ligation of cDNA into λ gt10	40
2.1.11 Packaging the cDNA library	40
2.1.12 Amplification of the cDNA library	41
2.1.13 Analysis of clones chosen at random from the library.....	41
2.2 ISOLATION AND ANALYSIS OF HORSE TRANSFERRIN cDNA	42
2.2.1 Probes used to screen the library	42
2.2.2 Screening the cDNA library	42
2.2.3 Analysis of positive clones	44
2.2.4 Subcloning horse transferrin cDNA fragments	45
2.2.5 Generation of subclones by exonuclease deletion	45
2.2.6 Amplification of the 3' end of horse transferrin cDNA	46
2.2.7 Cloning of PCR products	47
2.2.8 Sequencing	48
2.2.9 Sequence analysis	48

2.3 ANALYSIS OF KNOWN POLYMORPHIC REGIONS OF HORSE	
TRANSFERRIN cDNA	49
2.3.1 Exon 12: Southern blotting	49
2.3.2 Exon 12: cloning and sequencing	49
2.3.3 Exon 15: amplification and sequencing	50
2.4 IDENTIFICATION AND ANALYSIS OF OTHER POLYMORPHIC	
REGIONS	51
2.4.1 Samples for SSCP analysis	51
2.4.2 SSCP analysis	52
2.5 GENERAL METHODS	53
2.5.1 Quantitation and purity of DNA and RNA	53
2.5.2 Phenol/chloroform extractions	53
2.5.3 Spun column chromatography	54
2.5.4 DNA precipitation	54
2.5.5 Restriction enzyme digests	54
2.5.6 Isolation of genomic DNA from blood	55
2.5.7 Agarose gel electrophoresis	55
2.5.8 Isolation of DNA fragments from agarose gels	56
2.5.9 Southern transfer	56
2.5.10 Southern hybridization	57
2.5.11 Labelling of DNA probes	57
2.5.12 End-labelling λ HindIII size standards	58
2.5.13 Removal of probe from membranes	58
2.5.14 Autoradiography	58
2.5.15 Transformation of <i>E. coli</i> XL1 with plasmid DNA	59
2.5.16 Small scale preparation of plasmid DNA	59
2.5.17 Large scale preparation of plasmid DNA	60
2.5.18 Storage of bacteria	62
2.6 REAGENTS	62
2.6.1 Reagents used for RNA preparation and electrophoresis	62
2.6.2 Buffers used for the oligo(dT) cellulose column	63
2.6.3 Solutions used with bacteria and phages	63
2.6.4 Solutions used for hybridizations	63

2.6.5 Oligonucleotide primers	64
2.7 BACTERIAL STRAINS AND VECTORS	64
CHAPTER 3: RESULTS	65
3.1 CONSTRUCTION OF A HORSE LIVER cDNA LIBRARY	65
3.1.1 Isolation of RNA	65
3.1.2 Northern blotting	65
3.1.3 cDNA synthesis	68
3.1.4 Methylation of cDNA	68
3.1.5 Removal of linker fragments from cDNA	71
3.1.6 Ligation of cDNA into λ gt10	71
3.1.7 Titre of cDNA library	71
3.1.8 Assessment of library quality	74
3.2 HORSE TRANSFERRIN cDNA- λ GT10 CLONES	74
3.2.1 300 bp clones	74
3.2.2 1500 bp clones	77
3.2.3 Subclones for sequencing the 1500 bp clone	77
3.2.4 Screening for the 3' end of horse transferrin cDNA	80
3.3 HORSE TRANSFERRIN cDNA 3' REGION	83
3.3.1 Amplification of the 3' region	83
3.3.2 Sequencing of the 3' region	83
3.4 HORSE TRANSFERRIN cDNA AND PROTEIN SEQUENCES	86
3.4.1 Features of the cDNA sequence	86
3.4.2 Features of the amino acid sequence	86
3.4.3 Duplicated structure of horse transferrin	90
3.4.4 Comparison with other transferrins	90
3.5 POLYMORPHISM OF HORSE TRANSFERRIN	90
3.5.1 Exon 12	90
3.5.2 Exon 15	96
3.5.3 Detection of single stranded conformation polymorphisms	100
3.5.4 Sequence polymorphism in the 528 bp <i>Sau96AI</i> fragment	100
3.5.5 Sequence polymorphism in the 303 bp <i>EcoRI</i> fragment	103
3.5.6 Summary of polymorphisms analysed	103

CHAPTER 4: DISCUSSION AND CONCLUSIONS	108
4.1 EXPERIMENTAL CONSIDERATIONS	108
4.1.1 Quality of the cDNA library	108
4.1.2 Size of horse transferrin cDNA	109
4.1.3 Polymerase chain reaction	109
4.1.4 Direct sequencing of PCR products	110
4.1.5 Single stranded conformation polymorphisms	110
4.1.6 Blood typing	111
4.2 HORSE TRANSFERRIN cDNA AND PROTEIN SEQUENCES	111
4.2.1 Horse transferrin	111
4.2.2 Comparison of horse transferrin to other transferrin family sequences	112
4.2.3 Glycosylation	122
4.2.4 Iron and anion binding residues	123
4.2.5 Disulphide bonds	123
4.2.6 Structure of the region joining the two lobes	124
4.3 POLYMORPHISMS	127
4.3.1 Positions and types of amino acid substitutions	127
4.3.2 Polymorphism and electrophoretic mobility	129
4.3.3 Number of substitutions	130
4.3.4 Further polymorphism between D- and R-type alleles	132
4.3.5 Two groups of horse transferrins	133
4.3.6 Positions of polymorphisms in the primary structure	138
4.3.7 Positions of polymorphisms in the 3-dimensional structure	140
4.3.8 Interpretation of the polymorphism data	143
4.4 CONCLUSIONS	144
 BIBLIOGRAPHY	 145

LIST OF FIGURES

Figure 1. The structure of rabbit transferrin, represented by its α -carbon chain.....	8
Figure 2. Schematic representation of the iron binding site of the N-lobe of human lactoferrin.....	9
Figure 3. A scheme for the origin of transferrin genes, using the human transferrin gene as a model.....	25
Figure 4. Evolutionary scheme for genes encoding the transferrin family.....	27
Figure 5. Variants of horse transferrin as revealed by horizontal polyacrylamide gel electrophoresis of plasma samples.....	29
Figure 6. Restriction sites in the human transferrin cDNA which were used to create fragments for use as probes for screening the cDNA library.....	43
Figure 7. Analysis of preparations of total, poly(A) ⁺ and poly(A) ⁻ RNA.....	66
Figure 8. Elution of RNA from an oligo(dT) cellulose column.....	67
Figure 9. Products of the cDNA synthesis reaction.....	69
Figure 10. Trial methylation of λ DNA.....	70
Figure 11. Elution of cDNA and linker fragments from a Sepharose CL-4B column....	72
Figure 12. Ligation of cDNA to λ gt10 arms.....	73
Figure 13. <i>Eco</i> RI digests of ten clones chosen at random from the cDNA library.....	75
Figure 14. λ gt10-cDNA clones isolated from the cDNA library using the whole human transferrin cDNA as a probe.....	76
Figure 15. λ gt10-cDNA clones isolated from the cDNA library using the 5' 1500 bp of the human transferrin cDNA as a probe.....	78
Figure 16. Digests of the 1500 bp inserts from two λ gt10-cDNA clones	79
Figure 17. <i>Hinc</i> II digests of plasmids containing the 1500 bp fragment of the horse transferrin cDNA.....	81
Figure 18. Products of the exonuclease digestion of pUC18 carrying the 1500 bp fragment of horse transferrin cDNA.....	82
Figure 19. Amplification of the 3' region of horse transferrin cDNA by PCR.....	84
Figure 20. Restriction enzyme digests of the 800 bp PCR product.....	85
Figure 21. Horse transferrin cDNA clones and sequencing strategy.....	87
Figure 22. Horse transferrin cDNA sequence and its predicted translation product.....	88

Figure 23. Duplication in the amino acid sequence of horse transferrin.....	91
Figure 24. Polymorphic sites in exons 12 and 15.....	92
Figure 25. Southern blots of horse genomic DNA digested with <i>Sau3AI</i> and hybridized with an exon 12 fragment of the horse transferrin gene.....	93
Figure 26. Polymorphic positions in the sequence of exon 12 of the horse transferrin gene.....	95
Figure 27. Direct sequencing of PCR products from exon 15 of the horse transferrin gene.....	97
Figure 28. Polymorphic positions in the sequence of exon 15 of the horse transferrin gene.....	99
Figure 29. Fragments of horse transferrin cDNA screened for single stranded conformation polymorphisms.....	101
Figure 30. Single stranded conformation polymorphisms.....	102
Figure 31. Polymorphism in the sequence of the 528 bp <i>Sau96I</i> fragment.....	104
Figure 32. Polymorphisms in the sequence of the 303 bp <i>EcoRI</i> fragment.....	105
Figure 33. Positions of polymorphisms in the cDNA and amino acid sequences of horse transferrin.....	107
Figure 34. Multiple alignment of amino acid sequences of members of the transferrin family.....	113
Figure 35. Phylogenetic tree of members of the transferrin family.....	121
Figure 36. Disulphide bonding in the transferrins.....	125
Figure 37. Structure of the region joining the two lobes in serum transferrins and lactoferrins.....	126
Figure 38. Positions of horse transferrin sequence polymorphisms with respect to interspecies, interprotein and interlobe variation of transferrins.....	128
Figure 39. Positions of horse transferrin sequence polymorphisms on the 3- dimensional structure of human lactoferrin.....	141

LIST OF TABLES

Table 1. Features distinguishing the two groups of horse transferrin variants.....	31
Table 2. Polymorphisms in the cDNA and protein sequences of horse transferrin.....	106
Table 3. Amino acid sequence identity among members of the transferrin family.....	119
Table 4. Intra- and interspecies variation in exons 12 and 15 of mammalian serum transferrins.....	131
Table 5. Interspecies comparisons of mammalian serum transferrins and lactoferrins...	139

ABBREVIATIONS

ATP	adenosine triphosphate
bp	base pairs
BRL	Bethesda Research Laboratories
cAMP	cyclic adenosine monophosphate
cDNA	complementary DNA
cpm	counts per minute
cps	counts per second
dCTP	deoxycytosine triphosphate
DTT	dithiothreitol
EBTRC	Equine Blood Typing and Research Centre, Massey University
kb	kilobases
LF	lactoferrin
mRNA	messenger RNA
MTF	melanotransferrin
NTP	nucleotide triphosphates
PCR	polymerase chain reaction
PEG	polyethylene glycol
pfu	plaque forming units
RFLP	restriction fragment length polymorphism
rpm	revolutions per minute
rRNA	ribosomal RNA
SDS	sodium dodecyl sulphate
SSCP	single stranded conformation polymorphism
TEMED	tetramethylethylenediamine
TF	transferrin
U	units

Abbreviations for other reagents are defined in the Reagents section (Section 2.6)

ABBREVIATIONS**Amino Acids**

A	alanine
C	cysteine
D	aspartate
E	glutamate
F	phenylalanine
G	glycine
H	histidine
I	isoleucine
K	lysine
L	leucine
M	methionine
N	asparagine
P	proline
Q	glutamine
R	arginine
S	serine
T	threonine
V	valine
W	tryptophan
Y	tyrosine